Driving biomass breakdown through engineered cellulosomes

Sean P Gilmore, John K Henske, and Michelle A O’Malley*
Department of Chemical Engineering; University of California; Santa Barbara, CA USA

Extraction of sugar is the rate-limiting step in converting unpretreated biomass into value-added products through microbial fermentation. Both anaerobic fungi and anaerobic bacteria have evolved to produce large multi-cellulase complexes referred to as cellulosomes, which are powerful machines for biomass deconstruction. Characterization of bacterial cellulosomes has inspired synthetic “designer” cellulosomes, consisting of parts discovered from the native system that have proven useful for cellulose depolymerization. By contrast, the multi-cellulase complexes produced by anaerobic fungi are much more poorly understood, and to date their composition, architecture, and enzyme tethering mechanism remain unknown and heavily debated. Here, we compare current knowledge pertaining to the cellulosomes produced by both bacteria and fungi, including their application to synthetic enzyme-tethered systems for tunneled biocatalysis. We highlight gaps in knowledge and opportunities for discovery, especially pertaining to the potential of fungal cellulosome-inspired systems.

Introduction

Plant biomass is an abundant source of cellulose and hemicellulose, which are sugar-rich polymers that can be depolymerized and fermented into value-added chemicals.1 Many bioprocessing strategies employ metabolically engineered microbes like Saccharomyces cerevisiae or Escherichia coli to convert biomass hydrolysates into target products.2 However, sugar extraction from biomass relies on energy intensive chemical pretreatment to remove lignin and other recalcitrant biopolymers from substrates prior to hydrolysis.3,4 These steps are often performed in concert with expensive enzyme treatments,5 which limits the economic feasibility of this approach. Therefore, there is a critical need to develop enzyme systems that can act on unpretreated biomass, especially those that can be produced at high titers by fermentation capable microbes.

A wide variety of enzymes with complementary function are required to degrade plant biomass (Figure 1). While natural cellulosytic bacteria and aerobic fungi are a rich source of such enzymes, these microbes secrete a limited subset of enzyme types that cannot fully depolymerize crude plant material.6 To identify enzymes that degrade crude lignin-rich biomass one must look to the microbes that have evolved to degrade it. For example, large herbivores rely on a microbial consortia composed of anaerobic gut microbes (e.g. bacteria and fungi) to convert grasses and hay into sugar for the animal. Together, these anaerobic microbes secrete powerful enzymes capable of breaking down crude, unpretreated biomass.7

The high efficiency biomass breakdown associated with anaerobes stems from their ability to synthesize large multi-cellulase complexes called cellulosomes. These complexes link together all the diverse enzymes necessary for cellulose degradation through a “plug-and-socket” modular interaction via protein domains termed dockerin and cohesin. Logically, these tethered enzyme systems are suspected to increase degradation efficiency by concentrating active sites of the enzymes and targeting them toward the plant material, leading to substrate tunneling of the biomass toward free sugars. The well-studied bacterial cellulosome has demonstrated the power of these modular
enzyme complexes for biomass degradation. By comparison much less is known about fungal cellulosomes, yet early research suggests that they have functionalities equal to or greater than bacterial cellulosomes and can also be applied for bioprocessing applications. For example, anaerobic fungi produce a greater diversity of enzymes compared to anaerobic bacteria, including hemicellulases, such as xylanase and mannanase, other accessory enzymes responsible for lignin reorganization, such as polysaccharide deacetylases and esterases.9

**Bacterial Cellulosomes – From Native Parts to Synthetic Designer Cellulosomes**

Bacterial cellulosomes were first described in 1983 as "a discrete, cellulose-binding, multi-enzyme complex for the degradation of cellulotic substrates."10 They have since been found in many different bacterial species, primarily in the *Clostridium*, *Ruminococcus*, *Actinivi-brio*, and *Bacteroides* genera. Typically, these complexes in bacteria are built upon a large, non-catalytic protein called a scaffoldin.14 The size of bacterial scaffoldin proteins can vary widely, generally from 50kDa to 250kDa15; this size variation is related to the number of repeats of cohesin domains included in a particular scaffoldin. The cohesin domains associate strongly with dockerin domains on the individual cellulases,16 resulting in full complexes that range in size from 1.5 to 6MDa.14 and in bacteria the dockerin-cohesin interaction is highly species specific.17 Additionally, the scaffoldin very frequently contains one or more carbohydrate binding modules (CBM) to target the complex to its substrate.18 For further information on native bacterial cellulosomes there are several in depth reviews such as those by Bayer et al.14 and Doi et al.15

Following detailed studies on bacterial cellulose-degrading complexes, the concept of "designer cellulosomes" was first introduced by Bayer in 1994.19 Once cellulosomes were recognized to consist of modular parts, Bayer and colleagues proposed utilizing the native scaffoldin or cohesins with heterologous dockerin-fused enzymes to produce artificial cellulosomes, which would amplify cellulolytic capabilities for normally non-cellulolytic systems.19 Since then, many different reports have characterized "mini cellulosomes" inspired by bacterial cellulosomes.20-24 These studies have demonstrated that enzyme tethered complexes are much better than free enzymes at degrading low-accessibility, highly crystalline, insoluble substrates when produced in recombinant systems.20-24 However, very little improvement in activity is observed when complexes act upon well mixed, soluble substrates.20-24

Taken together, these observations suggest that the efficiency of cellulase complexes stems from CBM-facilitated enzyme targeting, as well as the relative organization of the enzymes within the complex. As shown in Figure 2, cellulose complexes are targeted to biomass substrate by the CBM. Once positioned, the cellulases act as a disassembly line to synergistically tunnel reactants and products toward sugars. In particular,
endoglucanases reduce the crystallinity of the substrate and free up free chain ends; these ends are then degraded by nearby processive exoglucanases, which release cellobiose as they move along the chain. Tethered β-glucosidases subsequently hydrolyze cellobiose to glucose. Such a model is supported by several reports, which noticed an increased rate of conversion of cellobiose to glucose when a β-glucosidase or β-xylosidase was included in synthetic mini cellulosomes. These results suggest that a mechanism similar to substrate channeling occurs, where the β-glucosidase acts on cellobiose as it is liberated from cellulose by a nearby exoglucanase. Indeed, other reports have demonstrated substrate channeling by fusing enzymes from a metabolic pathway to dockerins, and linking them together on a scaffoldin. This further demonstrates the broad applicability of the cellulosome system to any multi-enzyme biocatalytic process beyond those associated with cellulose degradation.

**Fungal Cellulosomes – Undercharacterized and Heavily Debated Complexes**

Although large multi-enzyme complexes have been documented in gut fungi since 1992, they are woefully under-studied compared to their bacterial counterparts. While they are believed to assemble through a modular cohesin-dockerin type interaction, the identity of the fungal cohesin domain, or a scaffoldin equivalent, remains elusive and is heavily debated. In fungi, dockerin domains are fused to catalytic enzymes, but these dockerins exist in tandem repeats at either the N or C-terminal of cellulases, compared to single copies often restricted to the C-terminal of cellulases in bacterial cellulosomes. The specificity of the dockerin-mediated interaction also appears to differ greatly from that found in anaerobic bacteria. Nagy et al. demonstrated through an ELISA that dockerin from one species can interact with cellulosomes from other species, suggesting that the dockerin-cohesin interaction is not species-specific as it is in anaerobic bacteria. Additionally, several reports estimate fungal cellulosomes to be greater than 1 MDa in size, although they have also been reported to be as small as 334kDa, and as large as 80MDa. This is similar to the bacterial system, where the size varies with the number of cohesins and particular type of enzyme associated.

Over 20 years ago, the first reported fungal cohesin was identified, yet there has not been convincing evidence since to substantiate this finding. At least 4 other reports have challenged this original finding, each proposing other proteins as fungal cohesins. By probing denatured fungal cellulosomes with an epitope-tagged recombinant dockerin, several studies have sought to find putative cohesin(s) through a Western Blotting approach. A short summary of the findings of these papers is detailed in Table 1. More recent reports have coupled this effort with Mass Spectrometry to identify the sequence of the interacting cellulosome-associated protein (represented by parentheses in the table). Interestingly, these proteins were all classified as catalytic proteins by sequence homology. In this regard, a catalytic scaffold would hold a distinct advantage over the bacterial scaffolding system because it would eliminate the need for the large, noncatalytic scaffold found in bacterial systems. However, as documented in Table 1, the protein identified varied with each study, therefore casting doubt on the results found in all of the studies. Furthermore, the method utilized must be called into question, since the cellulosome protein is denatured during SDS-PAGE before being transferred to the blot. Thus, such a technique is unlikely to fully replicate the native protein-protein interactions within fungal cellulosomes.

It was suggested by Nagy et al. that the fungal cellulosome interaction might be mediated by dockerin binding to post-translational modifications on the cohesin, which would not necessarily require a folded protein cohesin motif. They supported this claim with evidence that the cellulosomal proteins might be glycosylated, although they could not identify the exact nature of the glycans. However, this claim contradicts the findings of Raghothama, who identified several residues important for binding through an ELISA with mutant recombinant dockerins against native cellulosomes. These residues were aromatic amino acids, with flat edges of the aromatics presented as the likely interacting regions. Such regions are more indicative of protein-protein interaction than protein-glycan or other post-translational modifications.

Although much is still unknown regarding fungal cellulosome composition and structure, there are some preliminary findings from fungal cellulosomes that suggest that they may have distinct advantages over bacterial cellulosomes. The major degradation product of fungal cellulosomes is glucose, compared to cellobiose from bacterial cellulosomes. This is an attractive feature, since it removes the need to supplement costly β-glucosidases to cellulosomes. Two distinct classes of β-glucosidases have been identified in anaerobic fungi: freely diffusive (those without a dockerin domain) and cellulosome associated (with a dockerin domain). Finally, the enzymes identified to date from fungal cellulosomes comprise a long list with a diverse array of substrate specificities. A recent review by Hattori et al contains a complete list of glycoside hydrolase families and the species from which they were identified. There are close to 30 separate families represented across the various genera, which again reflects the large number of enzymes required to fully hydrolyze lignocellulose as demonstrated in Figure 1, indicating that fungal cellulosomes likely harbor complementary functions to their bacterial counterparts.

**Opportunities for New Discoveries and Synthetic Fungal Complexes**

While much has been learned about anaerobic fungi since they were first reported by Orpin in 1975, there is still a great deal of information that remains elusive, particularly regarding the cellulose-degrading complexes produced by the fungi and the sequence information encoding these enzymes. With the advent
of powerful techniques, such as Next Generation Sequencing (NGS) and Mass Spectrometry, many of the mysteries regarding the fungal cellulose should now begin to unfold. The most important information precluding our understanding of fungal cellulosomes is the identity of the cohesin and scaffoldin protein, including the conservation of these domains across fungal genera. Once known, it will undoubtedly become easier to determine the size, architecture, and potential diversity of anaerobic fungal cellulosomes. Finally, this knowledge can be applied to creating synthetic systems using the fungal cohesins and dockers to tether recombinant enzymes, which likely have desirable attributes distinct from those inspired by anaerobic bacteria as described above.

One exciting hypothesis to explain the wide range of size and compositional heterogeneity in fungal cellulosomes is that smaller cellulosomes associate into larger polycellulosomes, as has been demonstrated in some anaerobic bacteria. Given the sequence divergence of fungal cohesins-dockerin domains compared to those described above.

In conclusion, there is still much to learn about the cellulase complexes produced by anaerobic fungi. Compared to their bacterial relatives, fungal cellulosomes are capable of completely converting crude lignocellulosic biomass to its component sugars, due to the wide range of enzymes encoded within the complex. At the very least, they are an attractive resource for discovering new biomass degrading enzymes, novel modular protein-protein interaction domains, and potentially new enzyme superstructures from nature. Beyond this, their characterization could soon reveal a novel scaffolding system, which has applications in creating synthetic fungal enzyme complexes, as well as inspired complexes for any set of tandem biocatalytic processes.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Funding

The authors gratefully acknowledge funding support from the Office of Science (BER), US. Department of Energy (DE-SC0010352), the US. Department of Agriculture (Award 2011–67017–20459), and the Institute for Collaborative Biotechnologies through grant W911NF-09–0001 from the US. Army Research Office. The content of this information does not necessarily reflect the position or the policy of the Government, and no official endorsement should be inferred.

References

1. Sanderson K. A field in ferment. Nature 2006; 464:673-6; PMID:17151628; http://dx.doi.org/10.1038/444673a
2. Saxena RC, Adhikari DK, Goyal HB. Biomass-based energy fuel through biochemical routes: A review. Renewable Sustainable Energy Rev 2009; 13:167-78; http://dx.doi.org/10.1016/j.rser.2007.07.011
3. Linghong Z, Chunbao X, Champagne P. Overview of recent advances in thermo-chemical conversion of biomass. Energy Conversion Management 2010; 51:969-82; http://dx.doi.org/10.1016/j.enconman.2009.11.038
208 Volume 6 Issue 4

Acad Sci U S A 2012; 109:20431-6; PMID:23188794; http://dx.doi.org/10.1073/pnas.1211929109

Pages 7, Belach A, Belaich JP, Mostar E, Lamed R, Shoham Y, Bayer EA. Species-specificity of the cohesin-dock-

Page 10, Lamed R, Setter E, Kenig R, Bayer EA. The cellulosome system of Acetivibrzo-

Page 14, Bayer EA, Belaich JP, Shoham Y, Belaich A, Belaich JP, Morag E, Lamed R, Shoham Y, Belaich JP. Degradation of cellulose substrates by cellulase chimeras - Substrate targeting versus proximity of enzyme components. J Biol Chem 2001; 277:49621-30; PMID:12397974; http://dx.doi.org/10.1074/jbc.M100472200

Page 15, Doi RH. The rumen microbial ecosystem - some recent developments. Trends Microbiol 1997; 5:483-8; PMID:9447466; http://dx.doi.org/10.1016/S0966-842X(97)01359-1

Page 16, Stahl SW, Ciccek N, Sparling R, Berlin A, Levin DB. Biomass pretreatment: Fundamentals toward applica-

Page 17, Pages 7, Belach A, Belaich JP, Mostar E, Lamed R, Shoham Y, Bayer EA. Species-specificity of the cohesin-dock-

Page 20, Fierobe HP, Bayer EA, Tardif C, Czjzek M, Mechaly A, Lamed R, Shoham Y, Belaich JP. Characterization of a cellulolytic multienzyme complex - Substrate channeling in a self-assembled trifunctional enzyme complex. Proc Natl Acad Sci U S A 2002; 99:17452-7; PMID:12054094; http://dx.doi.org/10.1073/pnas.222074199

Page 21, You C, Zhang XZ, Zhang YHP. Mini-scaffoldin enhanced mini-cellulosome hydrolysis performance on low-accessibility cellulose (Avicel) more than on high-accessibility amorphous cellulose. Biochem Engineer J 2012; 63:57-65; PMID:22088649; http://dx.doi.org/10.1016/j.btej.2012.01.011

Page 22, Moriis S, Barak Y, Hadar Y, Wilson DB, Shoham Y, Lamed R, Bayer EA. Assembly of xylases into designer cellulosomes promotes efficient hydrolysis of the xylan component of a natural recalcitrant cellulose substrate. MBio 2011; 2: e00053-11; PMID:21624451; http://dx.doi.org/10.1128/mBio.00053-11

Page 23, Tsai S-L, DaSilva NA, Chen W. Functional display of 4-OMe-ethyl glucosidase in cellulosome of

Page 24, Sui L, Dalva NA, Chen W. Functional display of complex cellulosomes on the yeast surface via adaptive assembly. MBio 2012; 3: 1-3; http://dx.doi.org/10.1128/mBio.00047-12

Page 25, Gifen G, Anbar M, Morag E, Lamed R, Bayer EA. Enhanced cellulase degradation by targeted integration of a cohesin-fused β-glucosidase into the Clostridium thermocellum cellulosome. Proc Natl Acad Sci U S A 2010; 109:12098-303; PMID:22689561; http://dx.doi.org/10.1073/pnas.1007101109

Page 26, You C, Myung S, Zhang YHP. Facilitated substrate degradation in clostridium-like entities in bacteroides-cellulosolvens. Curr Genet 2003; 44:165-76; PMID:12834992; http://dx.doi.org/10.1007/s00294-002-0563-0

Page 27, Fanutti C, Ponyi T, Black GW, Hazlewood GP. A modular cellulase chimeric enzyme from the anaerobic fungus Panaeolus sp strain E2 on microcrystalline cellulose. Appl Environ Microbiol 1996; 62:20-5; PMID:8572896

Page 28, Ali BRS, Zhou LQ, Graves FM, Freedman RB, Black GW, Gilbert HP, Hazelwood GP. Cellulases and hemic-

Page 29, Nagy T, Tunnicliffe RB, Higgins LD, Waltera C, Gilbert HJ, Williamson MP. Characterization of a double docking from the cellulosome of the anaerobic fungus Piromyces eyasi. Proc Natl Acad Sci U S A 2002; 99:6210-2; PMID:11925190; http://dx.doi.org/10.1073/pnas.112100398

Page 30, Dijkerman R, Vervuren MFB, DenCamp H, van der-Wijst DR, van der Riet J, van der Drift C. Adsorption characteristics of cellulolytic enzymes from the anaerobic fungus Piromyces sp strain E2 on microcrystalline cellulose. Appl Environ Microbiol 1996; 62:20-5; PMID:8572896

Page 31, Ali BRS, Zhou LQ, Graves FM, Freedman RB, Black GW, Gilbert HP, Hazelwood GP. Cellulases and hemic-

Page 32, Zhong XZ, Zhang YHP. Mini-scaffoldin enhanced mini-cellulosome hydrolysis performance on low-accessibility cellulose (Avicel) more than on high-accessibility amorphous cellulose. Biochem Engineer J 2012; 63:57-65; PMID:22088649; http://dx.doi.org/10.1016/j.btej.2012.01.011

Page 33, Moriis S, Barak Y, Hadar Y, Wilson DB, Shoham Y, Lamed R, Bayer EA. Assembly of xylases into designer cellulosomes promotes efficient hydrolysis of the xylan component of a natural recalcitrant cellulose substrate. MBio 2011; 2: e00053-11; PMID:21624451; http://dx.doi.org/10.1128/mBio.00053-11

Page 34, Sui L, Dalva NA, Chen W. Functional display of complex cellulosomes on the yeast surface via adaptive assembly. MBio 2012; 3: 1-3; http://dx.doi.org/10.1128/mBio.00047-12

Page 35, Gifen G, Anbar M, Morag E, Lamed R, Bayer EA. Enhanced cellulase degradation by targeted integration of a cohesin-fused β-glucosidase into the Clostridium thermocellum cellulosome. Proc Natl Acad Sci U S A 2010; 109:12098-303; PMID:22689561; http://dx.doi.org/10.1073/pnas.1007101109

Page 36, Steenbakkers PJM, Harhangi HR, Bosscher MW, van den Drift C, Vogels GD, op den Camp HJ β-glucosidase in cellulose of the anaerobic fungus Piromyces sp strain E2 is a family 3 glycoside hydrolase. Biochem J 2003; 370:963-70; PMID:12485515; http://dx.doi.org/10.1042/BJ20021767

Page 37, Haitjema CH, Solomon KV, Henske JK, Theodorou HRJ, Campbell A, Bergman T, Gilbert HJ, Hazlewood GP. A modular 4-carboxymethyl ester hydrolyase from the anaerobic fungus Piromyces equi acts synergistically with xylanase and is part of a multiprotein cellu-

Page 38, Orpin CG. Studies on rumen flagellate neocallimastix-