In Situ Imaging of Single Carbohydrate-Binding Modules on Cellulose Microfibrils

Daryl J. Dagel,[†] Yu-San Liu,[‡,§] Lanlan Zhong,[†] Yonghua Luo,[‡] Michael E. Himmel,[‡,§] Qi Xu,[‡,§] Yining Zeng,[‡,§] Shi-You Ding,*[,‡,§] and Steve Smith*[,†]

Nanoscience and Nanoengineering, South Dakota School of Mines & Technology, 501 East St. Joseph Street, Rapid City, South Dakota 57701, United States, Biosciences Center, National Renewable Energy Laboratory, 1617 Cole Boulevard, Golden, Colorado 80401, United States, and Bioenergy Science Center, Oak Ridge National Laboratory, P.O. Box 2008, Oak Ridge, Tennessee 37831, United States

Received: October 12, 2010; Revised Manuscript Received: December 4, 2010

The low efficiency of enzymes used in the bioprocessing of biomass for biofuels is one of the primary bottlenecks that must be overcome to make lignocellulosic biofuels cost-competitive. One of the rate-limiting factors is the accessibility of the cellulase enzymes to insoluble cellulolytic substrates, facilitated by surface absorption of the carbohydrate-binding modules (CBMs), a component of most cellulase systems. Despite their importance, reports of direct observation of CBM function and activity using microscopic methods are still uncommon. Here, we examine the site-specific binding of individual CBMs to crystalline cellulose in an aqueous environment, using the single molecule fluorescence method known as Defocused Orientation and Position Imaging (DOPI). Systematic orientations were observed that are consistent with the CBMs binding to the two opposite hydrophobic faces of the cellulose microfibril, with a well-defined orientation relative to the fiber axis. The approach provides in situ physical evidence indicating the CBMs bind with a well-defined orientation on those planes, thus supporting a binding mechanism driven by chemical and structural recognition of the cellulose surface.

1. Introduction

Substantial progress in the conversion of biomass to biofuels at low cost is contingent on fundamental breakthroughs in our understanding of biomass structure and the mechanisms of cell wall degrading enzymes which convert lignocellulosic biomass to simple sugars suitable for fermentation. The chemical and structural properties that have evolved in biomass to prevent this disassembly of the plant cell wall (collectively referred to as biomass recalcitrance) are currently the subject of intensive study, motivated primarily by their potential as feedstocks for biofuels and other bio-based products.[1]

Lignocellulosic biomass consists chiefly of plant cell walls: an intricate mat of cellulose microfibrils embedded in a matrix of complex polysaccharides (hemicelluloses and pectin) and other polymers (mainly lignin).[2] While cellulose is the primary polysaccharide targeted for renewable feedstocks, its spatial-chemical structure and its interaction with degrading enzymes and microbes are not well understood at the nanometer scale. Toward this end, we investigate the interaction of carbohydrate-binding modules with cellulose microfibrils at the single molecule level, to elucidate details of the molecular mechanisms behind biomass conversion processes.

The primary plant-cell-wall-degrading microorganisms are bacteria and fungi. Although a surprisingly varied set of strategies have been found in different biomass degrading ecosystems in nature, the system commonly involves a community of microbes and their secreted enzyme systems, such as the cellulases, hemicellulases, and other related glycoside hydrolases; the polysaccharide lyases; and the carbohydrate esterases.[3] These enzymes are secreted as "free" enzymes, produced by most fungi or bacteria, or form multienzyme cellulosome complexes (produced by some anaerobic bacteria). The definitive enzymatic degradation of plant cell wall polysaccharides is generally accomplished by the synergistic action of enzymes with distinct glycosidic activities, which efficiently deconstruct the heterogeneous chemical structure which has evolved in the plant cell wall. These enzyme systems, secreted in an aqueous environment, must access the insoluble polymeric substrates in the plant cell wall to further catalyze hydrolysis. This key function is facilitated by the site-specific binding of the carbohydrate binding module (CBM), the component of the cellulase enzyme which directs the catalytic domain(s) to the cellulose surface.

CBMs are noncatalytic protein modules of glycoside hydrolases. They function as specific recognition modules that convey or harness the catalytic domain of the enzymes selectively to the polysaccharide substrate.[4] CBMs play a critical role to deliver the enzymatic components containing one or multiple catalytic modules (e.g., the entire complement of the cellulosome) collectively to the lignocellulosic substrate, thus fulfilling an important requirement for efficient degradation of insoluble cellulose and associated hemicelluloses. CBMs also potentially disrupt distinct regions of their target ligand,[5] align the polysaccharide chains, and function to enable the proximity of catalytic modules to allow their synergistic activities. The structure and ligand specificity of many CBMs have been determined experimentally, and several hundred others have been putatively identified and grouped into 59 families according to their amino acid sequence similarity (http://www.cazy.org/fam/acc_CBM.html). CBMs have also been classified into three types based on their function: Type A CBMs, referred to as surface-binding CBMs, bind specifically to insoluble crystalline cellulose. Types

* Corresponding authors. S.-Y. Ding. Phone: (303) 384-7758. Fax: (303) 384-7752. E-mail: shi.you.ding@nrel.gov. S. Smith. Phone: (605) 394-5268. Fax: (605)-394-2365. E-mail: steve.smith@sdsmt.edu.
[†] South Dakota School of Mines & Technology.
[‡] National Renewable Energy Laboratory.
[§] Oak Ridge National Laboratory.

10.1021/jp109798p

*Tr*CBM1-GFP
MTMITPSLHA*HHHHHH*GP**TQSHYGQCGGIGYSGPTVCASGTTCQVLNPYYSQCL**VPVEKMSKGEELFTG
VVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTFSYGVQCFSRYPDHMK
RHDFFKSAMPEGYVQERTISFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSH
NVYITADKQKNGIKANFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDH
MVLLEFVTAAGITHGMDELYK

*Ac*CBM2-GFP
MTMITPSLHACRGSS*HHHHHH*SSGLVPRGSHM**GVACRATYVVNSDWGSGFTATVTVTNTGSRATNGWTV
AWSFGGNQTVTNYWNTALTQSGASVTATNLSYNNVIQPGQSTTFGFNGSYSGTNAAPTLSCTAS**LVPVE
KMSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTFSYGV
QCFSRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNIL
GHKLEYNYNSHNVYITADKQKNGIKANFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSA
LSKDPNEKRDHMVLLEFVTAAGITHGMDELYK

*Ct*CBM3-GFP
MTMITPSLHACRGSS*HHHHHH*SSGLVPRGSHMG**NLKVEFYNSNPSDTTNSINPQFKVTNTGSSAIDLSK
LTLRYYYTVDGQKDQTFWCDHAAIIGSNGSYNGITSNVKGTFVKMSSSTNNADTYLEISFTGGTLEPGA
HVQIQGRFAKNDWSNYTQSNDYSFKSASQFVEWDQVTAYLNGVLVWGKEPG**LVPVEKMSKGEELFTGVV
PILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTFSYGVQCFSRYPDHMKRH
DFFKSAMPEGYVQERTISFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNV
YITADKQKNGIKANFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMV
LLEFVTAAGITHGMDELYK

Figure 1. Sequence and structure of recombinant CBM−GFPs. All three polypeptides contain an N-terminal 6xHis tag (shown in italic), a CBM (shown in bold), a short linker sequence (VPVEK), and a GFP (shown in underline). *Tr*CBM1 is a typical fungi family 1 CBM (36 amino acids) from *T. reesei* CBHI; *Ac*CBM2 is a bacterial family 2 CBM (101 amino acids) from the free cellulase enzyme of *A. cellulolyticus*; and *Ct*CBM3 is a bacterial family 3 CBM (156 amino acids) from the cellulosome scaffoldin of *C. thermocellum*. These three CBMs represent the three major natural cellulase paradigms.[3]

B and C bind to soluble polysaccharides, and they are chain-binding or end-binding CBMs, respectively. The surface-binding CBMs (type A) are particularly important, as they are commonly found in a wide distribution of different plant-cell-wall-degrading enzymes. It is thought that the planar binding sites of the type A CBMs are complementary to the planar (hydrophobic) (110) faces of the cellulose crystal.[6] In fact, this has been confirmed experimentally using TEM for the Iα allomorph of crystalline cellulose from *Valonia*.[4,7]

Fluorescence measurements that can achieve single-molecule sensitivity and nanometer localization (<10 nm) in sparse samples are now routine, due, in part, to advances in low-noise, single-photon detectors, postprocessing localization, and the development of total internal reflection fluorescence (TIRF) microscopy.[8,9] Recently developed techniques have removed the necessity of sparse coverage by utilizing photoactivatable fluorescent proteins (PA-FP) or by reducing the observation volume below the diffraction limit.[10] In addition to position, the 3D orientation of single molecules is important for several reasons, including its impact on photophysical parameters (e.g., fluorescence lifetime) and its role in deducing the orientation of the labeled biomolecule.[11] Defocused orientation and position imaging (DOPI) is a single-molecule method which exploits the anisotropy of dipole radiation outside of the focal plane.[12,13] In a homogeneous medium, fluorescing molecules can be considered as ideal electric dipole emitters, and when placed in a stratified medium with different refractive indices (e.g., air/sample/glass), the emitted radiation changes dramatically due to the self-interaction with the back-reflected electromagnetic field. This can be seen by slightly defocusing the microscope, which allows for unique identification of the 3D spatial orientation of the fluorescent dipole. DOPI has been applied in various areas, including imaging motor proteins,[14] photonic crystals,[15] and second harmonic generation.[16]

Previously, the binding specificity of CBMs to crystalline cellulose was demonstrated using single-molecule methods,[4,17,18] and processive motion of single fluorescently labeled family 2 CBMs along cellulose microfibrils was also observed.[19,20] Here, we utilize the DOPI method to observe the orientation of single fluorescently labeled CBMs bound to *Valonia* cellulose microfibrils. The approach provides in situ physical evidence of the CBM−cellulose molecular interactions—indicating their specificity for the hydrophobic crystalline facets of cellulose—and a well-defined cross-orientation of the CBM−green flourescent proteins (GFPs) relative to the fiber axis—suggesting a preferred orientation of the CBMs with respect to the cellulose crystal basis and contributing toward building a mechanistic model of the function of the CBMs and their associated cellulase enzymes.

We chose three families (e.g., 1, 2, and 3) of surface-binding CBMs for this study, representative of three very different types of cellulase systems. These CBMs are of interest due to their high binding affinity to crystalline cellulose and are commonly found in a wide variety of cellulase systems involved in breaking down plant cell walls. The family 1 CBM is found in fungi cellulases and are believed to play a critical role in accommodating the processive action of exoglucosidases, such as the cellobiohydrolases-I (CBH I), the key enzyme required for hydrolysis of crystalline cellulose.[21] The family 2 CBM has been found in bacterial free cellulase systems and could move directionally while bound to cellulose.[19] The family 3 CBM used in this study is derived from the scaffoldin protein that facilitates the assembly of multiple enzyme complexes to bind to the cellulolytic substrates (they may also be found in noncellulosomal cellulases), and the cellulosomes represent another important cellulolytic system found primarily in anaerobic bacteria.[22]

2. Materials and Methods

2.1. CBM−GFP Cloning, Expression, Purification, and Binding to Cellulose. Three CBMs were used in this study, and their sequences are illustrated in Figure 1. The detailed protocol for CBM cloning, expression, and purification was described in previous reports.[23,24] Briefly, TrCBM1 is a family 1 CBM from the Cellobiohydrolase (CBH) I gene of *Trichoderma reesei*. It was synthesized (GenScript Company, Piscataway, NJ) with genetic code optimization for expression in *Escherichia coli*. The gene also includes sequence encoding an N-terminal 6xHis tag. The synthesized DNA fragment was

amplified using a polymerase chain reaction (PCR) with primers, 5′-GATATAGCATGCCCATCATC-3′ and 5′-AGTTCTACCGGTACCAGGCACTGGC-3′, and cloned into pGFPuv (Clontech, Mountain View, CA) using restriction enzymes *SphI* and *AgeI* to generate expression plasmid pTrCBM1-GFP. *Ac*CBM2 is a family 2 CBM from *Acidothermus cellulolyticus*, and *Ct*CBM3 is a family 3 CBM from *Clostridium thermocellum*. Genomic DNAs of *A. cellulolyticus* and *C. thermocellum* (ATCC, Manassas, VA) were used as PCR templates to amplify the DNA fragments of *Ac*CBM2 and *Ct*CBM3, respectively. The primers used for PCR are 5′-GATATACATATGGGTGTGGCGTGCCGGGCGA-3′ and 5′-AGAGAGCTCGAGGCTGGCTGTGCAGCTGAGCGT-3′ for *Ac*CBM2 and 5′-GATATACATATGGGCAATTTGAAGGTTGAAT-3′ and 5′-AGAGAGCTCGAGACCGGGTTCTTTACCCCA-3′ for *Ct*CBM3. DNA fragments of *Ac*CBM2 and *Ct*CBM3 were cloned in a pET28b(+) vector (Novagen, Madison, WI, USA) using restriction enzymes *Nde*I and *Xho*I to generate the expression plasmids pET-*Ac*CBM2 and pET-*Ct*CBM3 that produce a fusion protein with a dual hexahistidine tag at its N- and C-termini. Primers 5′-ATGGG*CTGCAG*CCATCATCAT-3′ and 5′-TGGTGGGGTACCAGGCTGGCTGT-3′ for pET-*Ac*CBM2 and 5′-ATGGGCTGCAGCCATCATCAT-3′ and 5′-TGGTGG*GGTACC*AGACCGGGTTCT-3′ for pET-*Ct*CBM3 were used to amplify fragments including the CBM gene and an N-terminal His-tag from pET-*Ac*CBM2 and pET-*Ct*CBM3 plasmid. These PCR fragments were cloned into pGFPuv (Clontech, Mountain View, CA) using the restriction enzymes *Pst*I and *Kpn*I to generate expression plasmids p*Ac*CBM2-GFP and p*Ct*CBM3-GFP, respectively. p*Tr*CBM1-GFP, p*Ac*CBM2-GFP, and p*Ct*CBM3-GFP were then overexpressed in the *E. coli* strain of BL21 (DE3) (Stratagene, La Jolla, CA). Fusion proteins were purified using the QIAexpress Ni−NTA protein purification system (Qiagen, Valencia, CA). The CBM concentration was measured by NanoDrop 1000 (Thermo Scientific, Wilmington, DE, USA) and subsequently diluted with Tris buffer (50 mM Tris, pH 8.0) to optimize the resolution for single molecule detection. The purified protein was stored at 4 °C at a final concentration of 1 mg/mL in 20 mM Tris buffer (pH 8.0) containing 0.01% (w/v) NaN_3.

2.2. Assembling CBM−GFP and Cellulose Crystals. The cellulose crystals used for this study were isolated from the green algae, *Valonia ventricosa*.[23] *Valonia* crystalline cellulose microfibrils were exposed to dilute solutions of the CBM−GFP in aqueous buffer, with a concentration empirically determined to produce CBM−GFP binding frequency along the microfibril amenable to single molecule localization. About 1 μg of CBM−GFP protein in solution was incubated with 25 μg of *Valonia* cellulose crystals in 200 μL of Tris buffer (50 mM Tris, pH 8.0) by gentle mixing for 10 min, followed by centrifugation (10 000*g* for 5 min). These pellets were washed three times to remove unbound CBM−GFP in the solution. The resulting CBM−GFP-bound cellulose complexes (Figure 2) were suspended in 50 μL of Tris buffer for single molecule detection. An artists' conception of the CBM−GFP bound to the hydrophobic face of a cellulose microfibril is shown in Figure 2. Freshly prepared samples (2 μL volume) were placed between two glass coverslips. After the sample spreads out by the capillary effect, the coverslips were pressed together hard to minimize the optical depth of the sample and to obtain a thin and even distribution of cellulose substrate fibers, maintained in an aqueous environment.

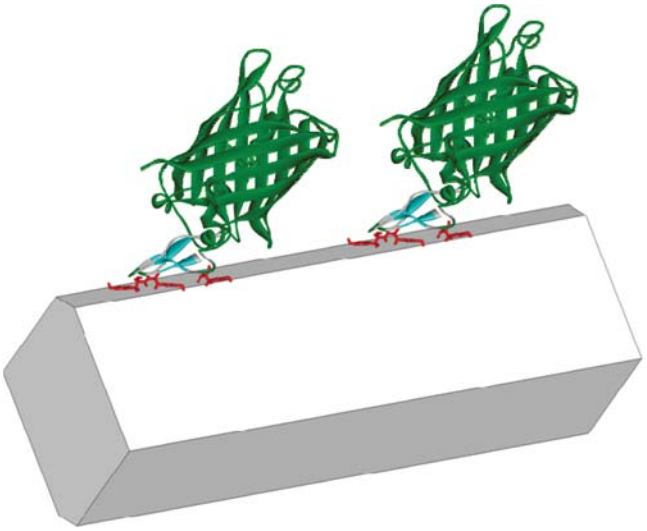

Figure 2. Schematic model of CBM−GFP bound to the crystalline cellulose microfibril. The recombinant CBM protein (red, turquoise, and gray) binds specifically to the (110) surfaces (narrow facets) of the cellulose Iα crystal. The GFPs (green) in turn are tagged to the CBM via a short linker polypeptide. *Valonia* cellulose crystals are ∼20 nm in diameter, while their planar faces are between 3 and 4 nm wide.[4] CBM−GFP dimensions are approximately to scale based on their published NMR structure of *Tr*CBM1 (PDB ID: 1CBH) and X-ray structure of GFP (PDB ID: 1Z1P).

3. Results and Discussion

The CBMs used in this study are type A surface-binding CBMs derived from three major cellulolytic systems. The *Tr*CBM1 is a family 1 CBM from the exoglucanse cellobiohydrolase-I (CBH I) of the fungus *Trichoderma reesei*. It contains 36 residues, and its NMR structure (PDB access # 1CBH) has been determined previously.[23] The *Ac*CBM2 is a family 2 CBM from the endoglucanase E1 of the bacterium *Acidothermus cellulolyticus*.[24] The *Ct*CBM3 is a family 3 CBM from the cellulosome scaffoldin CipA of the bacterium *Clostridium thermocellum*. Its X-ray structure (PDB access # 1NBC) has also been determined previously.[6] These three CBMs were genetically engineered to tag with a C-terminal green fluorescence protein (GFP) domain and a short linker peptide (Figure 1). The N-terminal peptide containing 6x histidines was used to purify these recombinant CBM-GFPs expressed in *E. coli* using a Ni−NTA protein purification system, and we previously demonstrated that the 6xHis tag did not affect the CBM binding to cellulose.[4] Details of the CBM−GFP cloning and sample preparation may be found in the Materials and Methods section.

The cellulose crystals used for this study were isolated from the green algae, *Valonia ventricosa*.[25] *Valonia* crystalline cellulose microfibrils were exposed to dilute solutions of the CBM−GFP in aqueous buffer, with a concentration empirically determined to produce CBM−GFP binding frequency along the microfibril amenable to single molecule localization. A schematic representation of the CBM−GFP bound to a crystalline cellulose microfibril is shown in Figure 2 (for further details of the sample preparation, see Materials and Methods).

GFP has been shown to have a well-defined transition dipole, registered to the protein,[26] with its fluorescence emission dipole oriented along the fixed axis of the incorporated chromophore (see Figure 2). It is this inherent polarization that makes the orientational imaging reported here possible. The CBM−GFP-bound cellulose samples were excited by circularly polarized 488 nm laser light under TIRF illumination. The subsequent

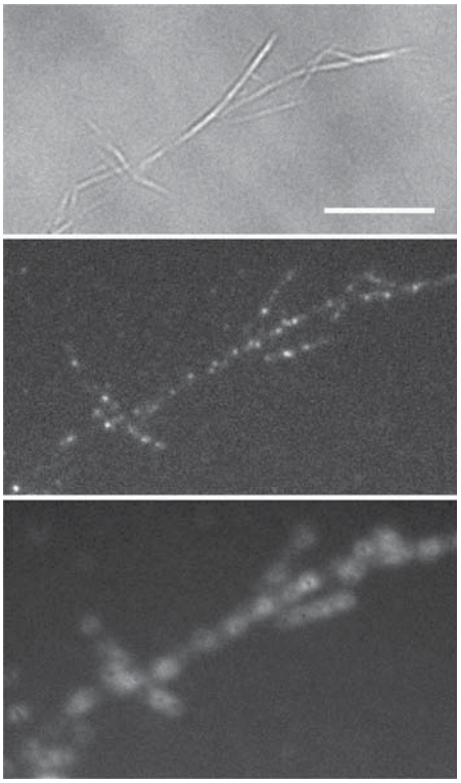

Figure 3. (Top) White light image of *Valonia* microfibrils; scale bar is 10 μm. (Center) TIRF image of GFP-labeled *Ct*CBM3 bound to microfibril. (Bottom) Defocused image, with structured patterns indicative of molecular orientation (see text).

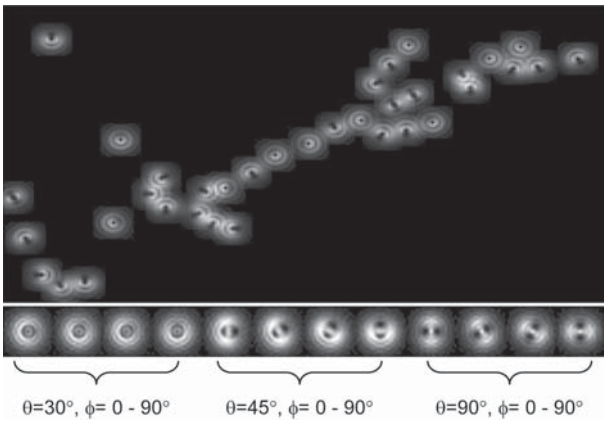

Figure 4. (Top) Computed image, based on extracted dipole orientations. (Lower) Example of computed DOPI images for varying polar and azimuthal angle.

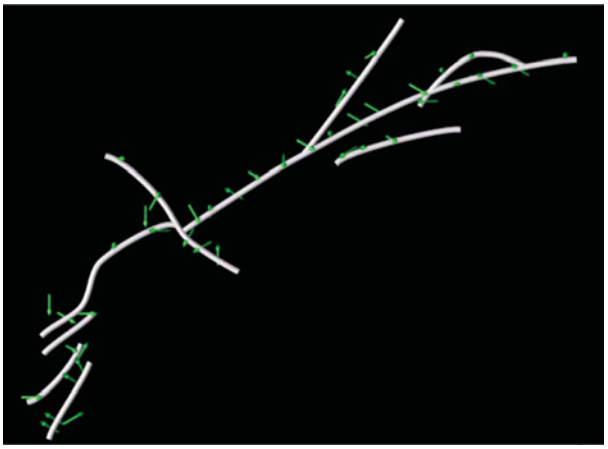

Figure 5. Three-dimensional model of *Valonia* microfibrils (gray tubes) showing orientation of GFP transition dipoles (green arrows).

fluorescence images (and white light bright field images) were obtained using an inverted frame fluorescence microscope (Olympus IX71) and an EMCCD camera (Andor iXon512). The system was equipped with a closed-loop piezoelectric stage (Mad City Laboratories LP200) which facilitated defocusing the fluorescence images with nanometer precision.

Figure 3 shows a white light image of a cluster of several *Valonia* cellulose microfibrils and in-focus and defocused fluorescence images of *Ct*CBM3-GFP bound to the same microfibrils. The defocused image was obtained by moving the piezo stage vertically by +0.6 μm. The white light and in-focus TIRF images show typical arrangements of microfibrils and the punctuated linear array of CBM−GFPs previously observed in these samples.[4] By defocusing the image slightly, the corresponding DOPI image shown in the lower part of Figure 3 was obtained. The structured patterns observed were stationary in time (as observed through time series of DOPI images, which showed little variation) and, as can be seen in the image and subsequent analysis, registered to the fibril axis. These structured patterns, and the analysis which follows, give the first experimental evidence of a preferred orientation of the CBMs relative to the cellulose crystal basis. A control experiment was done using GFP only (Clontech, Mountain View, CA), with the same protocol as was used for the CBM−GFP binding to cellulose, which resulted in no specific binding of GFP to cellulose, similar to previous reports.[19]

Single-molecule locations were extracted from the in-focus fluorescence images using standard techniques developed for particle tracking (http://www.physics.emory.edu/~weeks/idl/). At each centroid location, the corresponding defocused subimage was correlated with theoretical patterns generated from code provided by Enderlein (http://www.joerg-enderlein.de/?292). Figure 4 shows a summary of the results obtained through this analysis, based on a correlation analysis of the experimental data shown in Figure 3 and an array of theoretically predicted defocused images generated under conditions matched to the experiment (e.g., refractive indices and substrate thickness). The data obtained are represented in a 3D reconstruction of the dipole orientations (green arrows) superimposed on a model of the cellulose microfibril (gray tubes) shown in Figure 5. Inspecting the figure, it becomes clear that the dipole orientation of the GFP is not random and is registered to the microfibril.

To obtain a clearer understanding of the implications of the observed dipole orientations as they relate to the CBM−cellulose interactions, we performed a coordinate transformation wherein the dipole orientations inferred from the above analysis are referenced to the microfibrils to which they are associated through the CBM−cellulose interaction. To this end, the spherical polar (or elevation) angle θ and the azimuthal angle ϕ in the lab frame are related to the frame of the cellulose microfibril (θ_f, Ψ) by the following transformation: we first rotate the lab coordinate system about the z-axis (normal to the image plane) to align the x-axis along the long axis of the microfibril the subject CBM−GFP molecule is bound to. Next, we rotate ($90° - \theta_f$) about the x-axis to bring the normal of the plane parallel to the fiber axis and containing the transition dipole to the vertical or z-direction, where this angle is determined by requiring that the z-component of the dipole be zero in the new coordinate system. This angle then defines the plane in which the molecular transition dipole lies. In other words, we have rotated the microfibril about its axis of symmetry

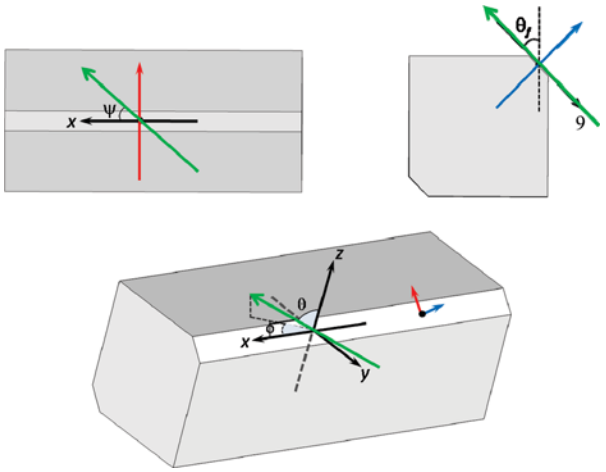

Figure 6. Relationship of lab frame angles (θ, ϕ) to the orientation angles (θ_f, ψ) referenced to the cellulose microfibril. Green arrow represents the GFP transition dipole, and red and blue arrows define principal axes registered to the microfibril and the hydrophobic (110) faces.

until the CBM−GFP transition dipole lies in the image plane. The angle Ψ is then the relative azimuthal angle $\Delta\phi$ between the fiber axis and the GFP transition dipole. The relationship between the two coordinate systems is illustrated in Figure 6, and the coordinate transformation can be written as

$$\tan \theta_f = \sin \Delta\phi \tan \theta$$

$$\cos \psi = \sin \theta \cos \Delta\phi$$

This coordinate representation was chosen to reveal the orientation of the plane parallel to the microfibril axis which contains the GFP transition dipole moment. While the structure of the CBM−GFP complex is yet to be determined, the observation that the GFP transition dipole is nonrandom and registered to the cellulose crystal basis tells us that the linkage between GFP and the CBMs is rigid or at least highly constrained, consistent with the short linker peptide sequence between CBM and GFP (only five amino acid residues, see Figure 1). Thus, these planes must, within some additive constant, identify the orientation of the binding face(s) of the cellulose microfibril. Similarly, the angle ψ must reveal details of the azimuthal orientation of the CBMs in that plane, relative to the cellulose crystal basis, again within an additive constant. As will be discussed in the following paragraphs, the data show that there are two equivalent planes to which the CBMs bind preferentially on the cellulose microfibril, consistent with previous studies, wherein type A CBMs bind specifically to the planar (110) faces of crystalline cellulose.[7,17] The azimuthal angle has not been observed before and gives the first experimental observation which indicates that the CBMs bind with a well-defined orientation on those planes, thus supporting a binding mechanism driven by chemical and structural recognition of the cellulose surface.

Several sets of defocused images were analyzed for each of the three CBM families, and the results are compiled in the histograms of Figure 7. The general trends for *Tr*CBM1, *Ac*CBM2, and *Ct*CBM3 on *Valonia* were similar: the polar face angle (θ_f) is distributed about 0° by approximately ±30°, and the azimuthal angle ψ peaks in the range ±60−80°. As discussed above, we interpret the polar angle as the orientation of the binding surfaces or facets of the microfibril, within an additive constant. In our analysis, GFP−CBMs bound to opposite sides of the microfibril are treated equivalently, as these are parallel planes, and transition dipoles rotated by 180° have equivalent radiation patterns. This can be seen clearly in the illustration of Figure 8, which shows the observed transition dipole orientation angles referenced to the microfibril. Were we to distinguish between these two equivalent cases, the face angle distributions would split into a bimodal distribution separated by 180°, similar to what is observed in the azimuthal angle distributions. The width of these distributions is to be expected from several factors, including structural disorder within the microfibril, bundling of multiple elementary microfibrils, causing apparent twisting of the microbril, and the random rotation of the individual fibers as they sit on the glass substrate. The width of these distributions may also be affected by possible motion of the CBM during the measurements. Similar distributions were reported for dried *Valonia* samples[7,17] and associated with the two distinct hydrophobic, "planar" (110) faces shown in Figure 8. However, these measurements are made in an aqueous environment while the protein is functional, thus removing ambiguity which may arise from the drying process associated with those measurements.

The apparent preference for a large relative azimuthal angle ψ in all the CBM−GFPs has not been observed before. As can be seen in the lower histograms of Figure 7, and in the graphical representations of Figures 5 and 8, all CBM−GFPs exhibit a preferred "cross" orientation near ∼±70°. The sign ambiguity is a consequence of the coordinate transformation, i.e., the sense of the rotation about the fiber axis, and the bimodal distribution observed further supports our interpretation that the CBMs bind to opposite faces of the microfibril. Subtle differences in the distribution of azimuthal angles shown in Figure 7 can be observed: for instance, the distributions for CBM2 and CBM3 appear somewhat narrower than for CBM1. This observation may indicate that the CBM1 is more "active"; namely, the CBM2 and CBM3 binding is tighter than that for CBM1. It has also been noted that samples of CBM1-bound cellulose are less stable. The unbound CBM1−GFP could be observed in solution after two days storage in 4 °C, whereas the CBM2 and CBM3 samples are stable for at least a week without detectable detachment. We believe that the fungal and bacterial CBM families may have different binding characteristics, and the subtle differences observed here may be a consequence of this. With a more detailed model of the structure of the CBM−GFP binding to cellulose, we may be able to discern these differences within the present data and future experiments. It is clear from the present analysis that the CBMs must recognize the cellulose surface in a way that is specific to the chemical structure of the CBM and the cellulose crystal and appears to be nonrandom. Such an observation is nontrivial and gives in situ experimental evidence for the biophysical mechanisms of how these proteins interact with cellulose. Such inputs are invaluable to theoretical and computational models of these complex biomolecules, as they provide realistic constraints that can be incorporated into simulations aimed at thermodynamic and kinetic descriptions of the CBMs[21] and the function of cellulases.

4. Conclusions

We applied the single molecule DOPI technique to extract 3D spatial information about GFP-tagged, surface-binding CBMs bound to *Valonia* crystalline cellulose microfibrils. The three families of surface binding CBMs studied represent components derived from fungal free cellulase, bacterial free

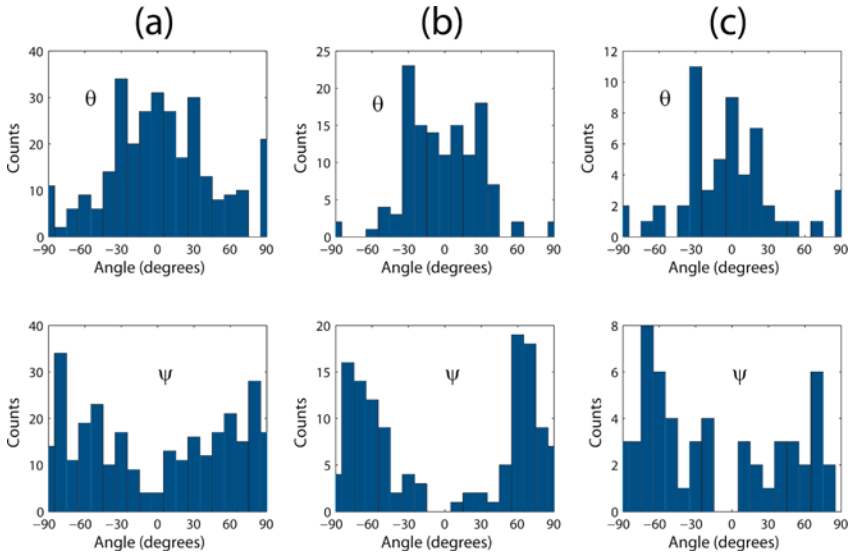

Figure 7. Histograms of the GFP−CBM polar face angle (θ_f) and relative azimuthal angle (ψ) for (a) *Tr*CBM1, (b) *Ac*CBM2, and (c) *Ct*CBM3 bound to *Valonia* microfibrils. These data represent the sum of several independent measurements from different samples.

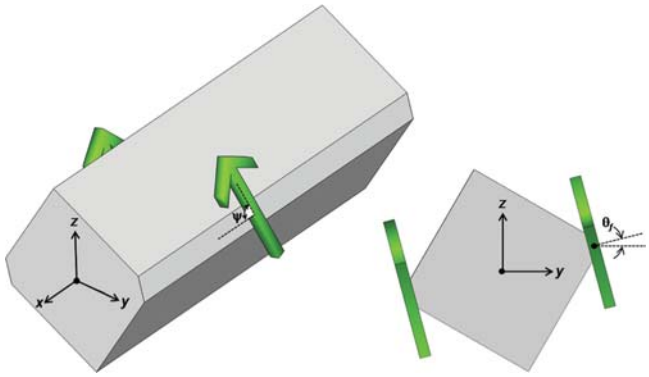

Figure 8. Dipole orientation referenced to the cellulose crystal. (Left) Azimuthal angle ψ describes the angle at which the dipole crosses the fiber axis (*x*-axis). (Right) Polar angle denotes the tilt of the dipole relative to the *z*-axis of the lab frame.

cellulase, and bacterial cellulosome, respectively. We observed systematic orientations consistent with the CBM−GFP binding to the two "hydrophobic" planar (110) faces of the cellulose microfibrils, with a well-defined crossed relative angle (±60−80°) between the GFP transition dipole and the cellulose microfibril long axis. The first observation confirms previous experiments that showed these proteins bind to the two hydrophobic faces of cellulose and gives the first in situ observation of this in an environment where the proteins are still functional (i.e., in an aqueous buffer), thus removing any ambiguity which may be introduced in dry sample preparation.[27] The second observation suggests that the binding of the CBMs is registered to the crystalline structure of the cellulose microfibril in three dimensions and supports a binding mechanism which is constrained within the plane of the hydrophobic surfaces, with surface recognition sensitive to chemical and structural properties of the CBMs and the cellulose substrate.[21] Further studies aimed at determining the structure of the CBM−GFP molecules used in this work and variations of different recombinant structures (i.e., using different linker peptides between CBM and GFP to mimic the natural variation of the linkers between CBM and catalytic modules) would allow a more detailed molecular-level description of the CBM−cellulose binding mechanisms. This work, combined with other single molecule methods, has the potential to observe CBM−cellulose interactions and their dynamics at the single molecule level.[28] Such observations could provide valuable insights toward building a molecular level understanding of the biophysical mechanisms which drive cellulase action and enzymatic biomass degradation processes.

Acknowledgment. This research was supported by the U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy Biomass Program through NREL subcontract ZCO-7-77379-01, by the National Science Foundation through award number DMR 0619890, and by the US DOE Office of Science, Office of Biological and Environmental Research, through the BioEnergy Science Center (BESC), a DOE Bioenergy Research Center. The authors acknowledge the 3D visualization and modeling efforts contributed by Thomas Smith.

References and Notes

(1) Himmel, M. E.; Ding, S. Y.; Johnson, D. K.; Adney, W. S.; Nimlos, M. R.; Brady, J. W.; Foust, T. D. *Science* **2007**, *315*, 804.
(2) Cosgrove, D. J. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 850.
(3) Wei, H.; Xu, Q.; Taylor, L. E.; Baker, J. O.; Tucker, M. P.; Ding, S. Y. *Curr. Opin. Biotechnol.* **2009**, *20*, 330.
(4) Ding, S. Y.; Xu, Q.; Ali, M. K.; Baker, J. O.; Bayer, E. A.; Barak, Y.; Lamed, R.; Sugiyama, J.; Rumbles, G.; Himmel, M. E. *Biotechniques* **2006**, *41*, 435.
(5) Boraston, A. B.; Bolam, D. N.; Gilbert, H. J.; Davies, G. J. *Biochem. J.* **2004**, *382*, 769.
(6) Tormo, J.; Lamed, R.; Chirino, A. J.; Morag, E.; Bayer, E. A.; Shoham, Y.; Steitz, T. A. *EMBO J.* **1996**, *15*, 5739.
(7) Lehtio, J.; Sugiyama, J.; Gustavsson, M.; Fransson, L.; Linder, M.; Teeri, T. T. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 484.
(8) Zander, C.; Enderlein, J.; Keller, R. A. *Single-molecule detection in solution: methods and applications*; Wiley-VCH: Berlin, 2002.
(9) Axelrod, D. In *Biophysical Tools for Biologists: In Vivo Techniques 89*; Correia, J. J., Detrich, H. W., III, Ed.; Academic Press: San Diego, 2008; p 169.
(10) Moerner, W. E. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 12596.
(11) Uji, I. H.; Deres, A.; Muls, B.; Melnikov, S.; Enderlein, J.; Hofkens, J. In *Fluorescence of Supermolecules, Polymers, and Nanosystems*; Berberan-Santos, M. N., Ed.; Springer: Berlin, 2008; p 257.
(12) Bohmer, M.; Enderlein, J. *J. Opt. Soc. Am. B* **2003**, *20*, 554.
(13) Patra, D.; Gregor, I.; Enderlein, J. *J. Phys. Chem. A* **2004**, *108*, 6836.
(14) Toprak, E.; Enderlein, J.; Syed, S.; McKinney, S. A.; Petschek, R. G.; Ha, T.; Goldman, Y. E.; Selvin, P. R. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 6495.
(15) Barth, M.; Schuster, R.; Gruber, A.; Cichos, F. *Phys. Rev. Lett.* **2006**, *96*, 4.
(16) Sandeau, N.; Le Xuan, L.; Chauvat, D.; Zhou, C.; Roch, J. F.; Brasselet, S. *Opt. Express* **2007**, *15*, 16051.

(17) Xu, Q.; Tucker, M. P.; Arenkiel, P.; Ai, X.; Rumbles, G.; Sugiyama, J.; Himmel, M. E.; Ding, S. Y. *Cellulose* **2009**, *16*, 19.

(18) Hong, J.; Ye, X.; Zhang, Y.-H. P. *Langmuir* **2007**, *23*, 12535.

(19) Liu, Y. S.; Zeng, Y. N.; Luo, Y. H.; Xu, Q.; Himmel, M. E.; Smith, S. J.; Ding, S. Y. *Cellulose* **2009**, *16*, 587.

(20) Liu, Y. S.; Luo, Y. H.; Baker, J. O.; Zeng, Y. N.; Himmel, M. E.; Smith, S. J.; Ding, S. Y. *Proc. SPIE* **2010**, *7571*, 757103.

(21) Bu, L. T.; Beckham, G. T.; Crowley, M. F.; Chang, C. H.; Matthews, J. F.; Bomble, Y. J.; Adney, W. S.; Himmel, M. E.; Nimlos, M. R. *J. Phys. Chem. B* **2009**, *113*, 10994.

(22) Ding, S. Y.; Xu, Q.; Crowley, M.; Zeng, Y.; Nimlos, M.; Lamed, R.; Bayer, E. A.; Himmel, M. E. *Curr. Opin. Biotechnol.* **2008**, *19*, 218.

(23) Kraulis, P. J.; Clore, G. M.; Nilges, M.; Jones, T. A.; Pettersson, G.; Knowles, J.; Gronenborn, A. M. *Biochemistry* **1989**, *28*, 7241.

(24) Ding, S. Y.; Adney, W. S.; Vinzant, T. B.; Decker, S. R.; Baker, J. O.; Thomas, S. R.; Himmel, M. E. In *Applications of Enzymes to Lignocellulosics*; ACS Symposium Series 855; Mansfield, S. D., Saddler, J. N., Eds.; American Chemical Society: Washington, DC, 2003; p 332.

(25) Imai, T.; Putaux, J. L.; Sugiyama, J. *Polymer* **2003**, *44*, 1871.

(26) Inoue, S.; Shimomura, O.; Goda, M.; Shribak, M.; Tran, P. T. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 4272.

(27) Igarashi, K.; Koivula, A.; Wada, M.; Kimura, S.; Penttila, M.; Samejima, M. *J. Biol. Chem.* **2009**, *284*, 36186.

(28) Sathitsuksanoh, N.; Zhu, Z. G.; Wi, S.; Zhang, Y.-H. P. *Biotechnol. Bioeng.* **2010**, DOI: 10.1002/bit.22964.

JP109798P