

1 Interrelationships between cellulase activity and 2 cellulose particle morphology

3 Johan P Olsen¹, Bryon S. Donohoe², Kim Borch³, Peter Westh¹ and Michael G. Resch^{4*}

4 ¹Research unit for Functional Biomaterials, Department of Science, Systems and Models,
5 Roskilde University, DK-4000 Roskilde, Denmark

6 ²Biosciences Center, National Renewable Energy Laboratory, Golden, CO

7 ³Novozymes A/S, Krogshøjvej 36, DK-2880 Bagsværd, Denmark

8 ⁴National Bioenergy Center, National Renewable Energy Laboratory, Golden, CO

9

10 *Corresponding author: Michael G. Resch, michael.resch@nrel.gov, (303) 384-7854

11 Acknowledgements

12 This work was supported by the Danish Council for Strategic Research, Program
13 Commission on Sustainable Energy and Environment (grant # 11-116772 to P.W.),
14 Augustinus Foundation (# 14-4639 to JO) and Oticon Foundation (# 14-3925 to JO).
15 BSD and MGR were supported by the DOE Office of Energy Efficiency & Renewable
16 Energy (EERE), Bioenergy Technologies Office (BETO). The authors gratefully
17 acknowledge associate professor Morten F. Pedersen and assistant professor Per M.
18 Jepsen for their assistance with statistical analyses and coulter counter measurements,
19 respectively.

20 **Keywords**

21 Imaging, Transmission Electron Microscopy, cellulose surface structure, cellulase,
22 cellobiohydrolase, endoglucanase

23 **Abstract**

24 It is well documented that the enzymatic hydrolysis of cellulose follows a reaction
25 pattern where an initial phase of relatively high activity is followed by a gradual slow-
26 down over the entire course of the reaction. This phenomenon is not readily explained
27 by conventional factors like substrate depletion, product inhibition or enzyme
28 instability. It has been suggested that the underlying reason for the loss of enzyme
29 activity is connected to the heterogeneous structure of cellulose, but so far attempts to
30 establish quantitative measures of such a correlation remain speculative. Here, we have
31 carried out an extensive microscopy study of Avicel particles during extended
32 hydrolysis with *Hypocrea jecorina* cellobiohydrolase 1 (CBH1) and endoglucanase 1 and
33 3 (EG1 and EG3) alone and in mixtures. We have used differential interference contrast
34 microscopy (DICM) and transmission electron microscopy (TEM) to observe and
35 quantify structural features at μm and nm resolution, respectively. We implemented a
36 semi-automatic image analysis protocol, which allowed us to analyze almost 3000
37 individual micrographs comprising a total of more than 300,000 particles. From this
38 analysis we estimated the temporal development of the accessible surface area
39 throughout the reaction. We found that the number of particles and their size as well as
40 the surface roughness contributed to surface area, and that within the investigated
41 degree of conversion ($< 30\%$) this measure correlated linearly with the rate of reaction.
42 Based on this observation we argue that cellulose structure, specifically surface area and
43 roughness, plays a major role in the ubiquitous rate loss observed for cellulases.

44 Introduction

45 Essentially all reported experiments have shown that the enzymatic hydrolysis of
46 cellulose exhibits a gradual loss of activity as the reaction progresses. This behavior is
47 only partially explained by product inhibition, substrate depletion and physical
48 instability of the enzymes (Zhang and Lynd 2004), and the underlying reasons for the
49 ubiquitous slowdown has been the subject of much debate (Bansal et al. 2009). In the
50 very early stage of enzymatic hydrolysis turnover slow-down might be explained by the
51 processive nature of cellulases (Jalak and Våljamäe 2010; Praestgaard et al. 2011), and
52 some reports have indicated that other enzyme effects like inactivation, unproductive
53 binding or enzyme crowding might influence activity during extended hydrolysis as well
54 (Eriksson et al. 2002; Yang et al. 2006). Nonetheless, most recent work has concluded,
55 that the continued rate loss is mainly substrate dependent (Arantes and Saddler 2011;
56 Bansal et al. 2012; Hong et al. 2007; Jeoh et al. 2007; Luterbacher et al. 2015) even
57 though relationships between substrate alterations and the reduced rate of hydrolysis
58 remain poorly understood. Undoubtedly part of the rate decrease can be attributed to
59 the heterogeneous nature of cellulose; in addition to being an insoluble polymer it is a
60 mixture of amorphous and crystalline regions forming fibrils of various sizes depending
61 on the cellulose source (Payne et al. 2015). This heterogeneity has been proposed to
62 contribute to the observed rate loss. Thus it has often been suggested that the
63 amorphous parts are degraded preferentially, leaving behind crystalline regions that are
64 more recalcitrant towards degradation (Zhang and Lynd 2004), but reports that the
65 crystallinity index of cellulose is unaffected by enzymatic degradation (Hall et al. 2010),
66 has directed attention towards more continuous effects like changes in accessible
67 surface area, pore size or available reactive sites on the substrate ~~has also been~~
68 ~~proposed~~ (Bansal et al. 2009; Grethlein 1985). Based on mechanistic models it has been
69 suggested that change in substrate surface area is an important factor for the decline in

70 hydrolysis rate (Levine et al. 2010). Hence, Bansal et al concluded that 90% of the rate
71 decline is caused by a decrease in substrate accessibility and hydrolysability – a
72 quantity the authors define as the fraction of enzyme binding sites that are available for
73 hydrolysis (Bansal et al. 2012).

74 Many studies have sought to clarify changes undergone by cellulosic substrates during
75 hydrolysis using high resolution imaging techniques like atomic force microscopy
76 (AFM), transmission or scanning electron microscopy (TEM, SEM) or fluorescence
77 microscopy. Whole pretreated biomass has a complex architecture across multiple
78 length scales, and while this architecture has been studied extensively for many
79 different substrates (Antal 1985; Ciesielski et al. 2014; Donohoe and Resch 2015), only
80 in rare cases have structural effects of enzymatic degradation of these substrates been
81 reported (Resch et al. 2014). In contrast numerous studies on enzyme induced changes
82 to the morphology of pure cellulose substrates like BMCC ribbons, *Valonia* fibrils and
83 Avicel have been published over the years (for an extensive review see Bubner et al.
84 2013). The general observation is that cellobiohydrolases (CBHs) – notably CBH1 –
85 target crystalline regions while endoglucanases (EGs) target amorphous regions, and
86 that the combined action of the two types of enzymes has a profoundly different effect
87 than each of them alone (Bubner et al. 2013; Payne et al. 2015). Furthermore it has been
88 found that CBHs degrades cellulose fibrils from the ends and cause narrowing and
89 sharpening of the fibril ~~and indeed entire bundles of cellulose fibrils~~ while EGs have no
90 apparent spatial preference and cause general surface disruption or fibrillation (Chanzy
91 and Henrissat 1985; Chanzy et al. 1983). Some studies have found that CBH1 acts to
92 clear away sub-fibrils created by EG (Sprey and Bochem 1992) while others have
93 observed that the action of CBH1 alone caused initial fibrillation, and only after
94 extensive hydrolysis left behind thinned, sharpened and recalcitrant fractions (Chanzy
95 et al. 1983; Imai et al. 1998; Jeoh et al. 2013). These studies have focused on structural
96 changes mostly from a qualitative approach, but some attempts have been made at

97 acquiring quantitative measures of changes in cellulose structure observed in AFM
98 (Bubner et al. 2012; Wang et al. 2013) and fluorescence microscopy (Luterbacher et al.
99 2015) to elucidate the progression over time. These studies have provided valuable
100 insights in the continuous morphological changes invoked by cellulases on cellulose
101 substrates. However, to make continued observation of a defined area possible, the
102 studied substrates were functionalized to a support material or solubilized and
103 recrystallized to obtain a smooth surface. These modifications might affect the
104 preferences for enzyme attack or alter structural integrity of the substrate and hence
105 influence the morphological effects of enzymatic degradation. In addition the general
106 approach is often incompatible with (or at least complicates interpretation of) bulk
107 biochemical measurements of substrate conversion that are required for direct
108 comparisons of substrate structure and enzymatic activity. For these reasons it is not
109 straightforward to draw conclusions from these studies regarding any effect substrate
110 alterations might have on cellulase slow-down.

111 Here, we have used differential interference contrast microscopy (DICM) and
112 transmission electron microscopy (TEM) to determine structural changes of Avicel
113 particles at the μm -nm scale during extended hydrolysis. Avicel was chosen primarily
114 due to the relatively narrow and well-defined size range of the particles making them
115 well suited for the image analysis presented here. The substrate was treated with two
116 different endoglucanases (EG1 and EG3) and a cellobiohydrolase (CBH1) from
117 *Trichoderma reesei* either alone or in mixtures. To get reliable, quantitative measures of
118 the observed structural changes we implemented automated procedures for image
119 analyses using the open source image processing software ImageJ. This allowed us to
120 analyze a very large number of micrographs, which was necessary due to the
121 heterogeneity of Avicel particles. In other words we did not attempt to identify distinct
122 structural features in individual samples but rather to obtain quantitative ensemble
123 average measures of structure changes analogous to what is obtained in biochemical

activity measurements. This approach, along with the microscopy study being blinded, ensured an unbiased and reliable quantification of the morphological effect of cellulase activity on unmodified Avicel and allowed us to correlate these changes to the loss of hydrolytic activity.

Materials and methods

Cellulose digestion

Unless otherwise noted, all chemicals were purchased from Sigma-Aldrich. All hydrolysis reactions were carried out in 30mM sodium acetate buffer, pH 5.0 with 2 mM CaCl_2 and 0.01% sodium azide. Avicel PH 101 was washed by centrifugation, once in milliQ water and twice in reaction buffer. Final dry matter concentration in the reaction mixture was 10 mg/mL (1% w/v). TrCel7A, TrCel7B and TrCel12A were heterologously expressed in *Aspergillus oryzae* and purified as described elsewhere (Westh et al. 2014). **To ensure a substantial conversion with the simple mixtures used here** the enzymes were dosed at 100 mg/g Avicel either alone or in mixtures of 5%, 25%, 50% or 80% EG to total enzyme (see Table 1). Control samples were made without added enzyme. All samples were supplemented with 10mg/g β -glucosidase (*Aspergillus niger*) and incubated at 50°C with end-over-end rotation at 10rpm. Samples were taken for activity measurements at 0, 8, 24, 48, 72, and 96 hours. A subset of each sample from 0, 8, 48 and 96 hours was stored at -20°C for imaging. Glucan conversion was determined by HPLC as previously described (Resch et al. 2014). All image processing was carried out using Fiji, a distribution of the open source image processing software ImageJ (Schindelin et al. 2012). The analyses described below were written into macros to allow automatic processing.

Table 1 Enzyme composition in studied samples. Total enzyme dosage was 100mg/g Avicel. All samples were added 10% β -glucosidase

	1	2	3	4	5	6	7	8	9	10	11	12
CBH1	100%	0%	0%	95%	75%	50%	20%	95%	75%	50%	20%	0%
EG1	0%	100%	0%	5%	25%	50%	80%	0%	0%	0%	0%	0%
EG3	0%	0%	100%	0%	0%	0%	0%	5%	25%	50%	80%	0%

Differential interference contrast microscopy:

Microscopy: Slurries of samples were transferred directly to a glass microscope slide and a cover slip was sealed around the edges to minimize evaporation. Images were taken with a SPOT RTKE CCD camera (Diagnostic Instruments, Sterling Heights, MI) on a Nikon C1 Plus microscope (Nikon, Tokyo, Japan) in bright field mode. All images were taken using a 60X 1.4 NA Plan Apo objective resulting in 1600 x1200 pixel images covering 195 x146 μ m (corresponding to a pixel size of 122nm).

Image processing: Representative examples of DCIM micrographs before and after processing in ImageJ are shown in Figure 2 and a step-by-step walkthrough of the processing is shown in supplementary Figure S1. Initially, the contrast was enhanced, and the 24-bit RGB image was converted to a binary mask by applying the “Yen” threshold method (Kapur et al. 1985). A shadow from uneven illumination of the samples distorted the image processing along the edges of many of the DICM micrographs. To exclude this from the analysis we removed the outermost 60 pixels along all four edges of the image. Since DICM enhances contrast many particles would appear as only a perimeter. To make particles solid we applied the “Fill Holes” command before the perimeter and area as well as the maximum and minimum Feret’s diameter (longest and shortest possible distance between two parallel tangents to the particle’s perimeter, here used as a measure of particle length and width, respectively) was determined for each particle using Fiji’s “Analyze Particles” function.

Transmission electron microscopy

Microscopy: 3 μ l sample was drop cast onto a 200 mesh carbon coated copper grid, negatively stained with 2% uranyl acetate and rinsed with water. The grids were imaged on a FEI Tecnai G2 20 Twin 200 kV LaB6 TEM (FEI, Hillsboro, OR) with a 4 megapixel Gatan UltraScan 1000 camera (Gatan, Pleasanton, CA). All images were acquired by the same operator to ensure consistent search criteria. To avoid bias the trial was blinded i.e. the operator was unaware which sample was on any particular grid. Grids were surveyed systematically at 1500x magnification and images were acquired at 3500x magnification, resulting in a frame size of 6.1 x 6.1 μ m. All particles large enough to be clearly identified at 1500x (more than $\sim$ 1 μ m long) and small enough to fit in the image frame at 3500x (less than $\sim$ 6 μ m) were imaged. As evident from the results section, most often a large number of much smaller particles were included in the field of view. Approximately 80-90 images were captured for each sample.

Image processing and analysis: Noise was minimized by the “Remove Outliers” option, which changes the value of any pixel to the median of all pixel values in a surrounding block of designated size (here, 10x10 pixel) if that pixel deviates from that median by more than a specified value (1x standard deviation). Subsequently the images were converted from grayscale to a binary mask by applying the “Triangle” threshold method (Rogers et al. 1977) and the particle perimeters and areas were measured using the “Analyze Particles” procedure. Note, however, that even though holes are subtracted from particle area they are not included in the perimeter. Consequently, in images including overlapping particles these were assessed as single large particles with holes, the perimeter being only the outer perimeter of the particle. To overcome this we implemented a macro that first measured the area and perimeter of the particles and subsequently inverted the mask so as to measure the perimeter of the holes. The resulting two data files were combined with a script in Matlab (R2013a v 8.1.0.604,

Natick, Ma, USA) to get the total perimeter and area. Thus the reported values are total perimeter and area for each image as opposed to each individual particle. A step-by-step review of the entire processing of a representative TEM micrograph can be found in the supplementary material (Figure S2).

Size determination by coulter counter

Samples with sufficient residual volume after microscopy (28 out of the total 36 samples) were analyzed on a Multisizer 4e coulter counter fitted with a 100 μ m aperture (Beckman Coulter, Brea, CA), using 5‰ NaCl (0.45 μ m vacuum filtered) as electrolyte. Samples were diluted 1:800 in the electrolyte solution and the particles were resuspended by repeated pipetting before 1ml was analyzed. Due to the limited amount of sample material we were unable to perform replicate measurements.

Statistics

As previously reported and detailed in the supplementary information, the size distribution of Avicel particles follows a log-normal distribution (Marshall and Sixsmith 1975) which makes the average size a poor measure of the typical particle in a given sample. For this reason we have mainly used the median values in the analysis presented below. However, in the case of DICM analysis, because each micrograph contained a large number of particles we could analyze the distribution of a given parameter within each micrograph, and use the 5 individual images of each sample as analytical replicates.

All analyses of correlation were carried out using the statistical software package JMP version 10 (SAS Institute Inc., Cary, NC). In the case of median length and circularity were data appeared to be linear (see Figure 3), we used Pearson's correlation analyses to determine if any such correlation could be considered significant. For particle count, spheroid surface area and surface roughness, the data was clearly non-linear and non-

monotone. To test whether the variation in each data set was random or significantly dependent on the conversion we applied Hoeffding's independence test. As detailed in the results section we divided the data into subsets but since these data sets were relatively small with considerable variation we did not find it reasonable to assume linearity (as implicitly done with Pearson's). For this reason we used the Spearman's rank analysis to test for correlation in these cases, although it should be noted that a similar outcome was found with Pearson's correlation analysis (not shown).

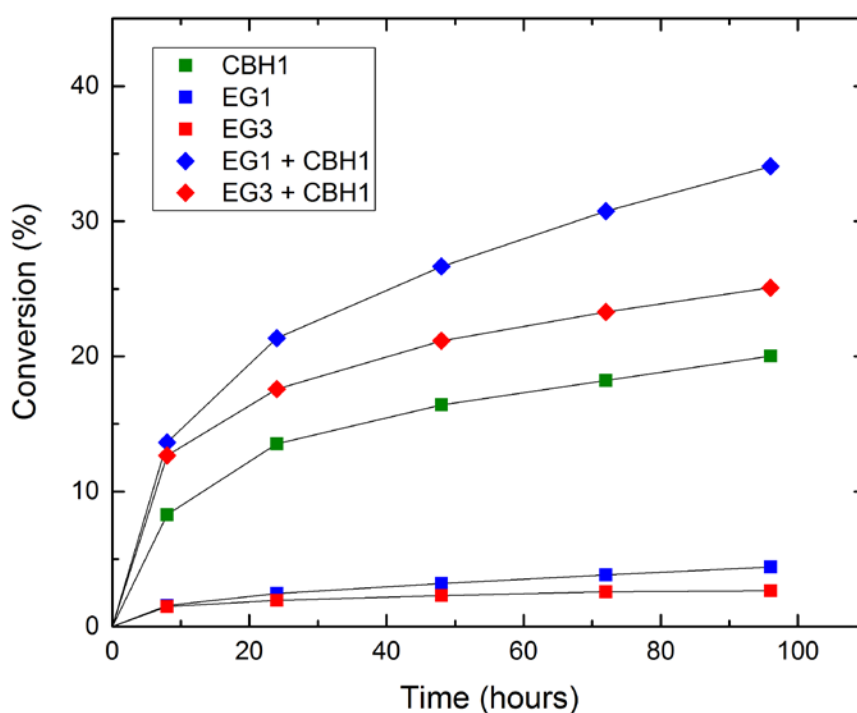


Figure 1 Activity measurements of monocomponent (squares) and 50/50 mixtures (diamonds) of CBH1 (green) with either EG1 (blue) or EG3 (red). Lines are shown to guide the eye.

Results

12 separate Avicel samples were digested by either mono- or bicomponent enzyme solutions (see Table 1), and aliquots were taken out at 5 time points for biochemical

analysis and 3 time points for imaging as described above. All samples had identical substrate and total enzyme loads at the beginning of the reaction. Conversion of substrate was calculated from the production of glucose and the progression over 96 hours is shown for selected samples in Figure 1. Five images were taken in the DICM for each sample at each time point. In the TEM 80-90 images were taken due to the much lower number of particles included in the image frame of this microscope at the magnification range used here. This resulted in 180 DICM and 2950 TEM micrographs (containing a total of approximately 10^5 and $2 \cdot 10^5$ detectable particles, respectively), which were subjected to the semi-automated image analysis described above.

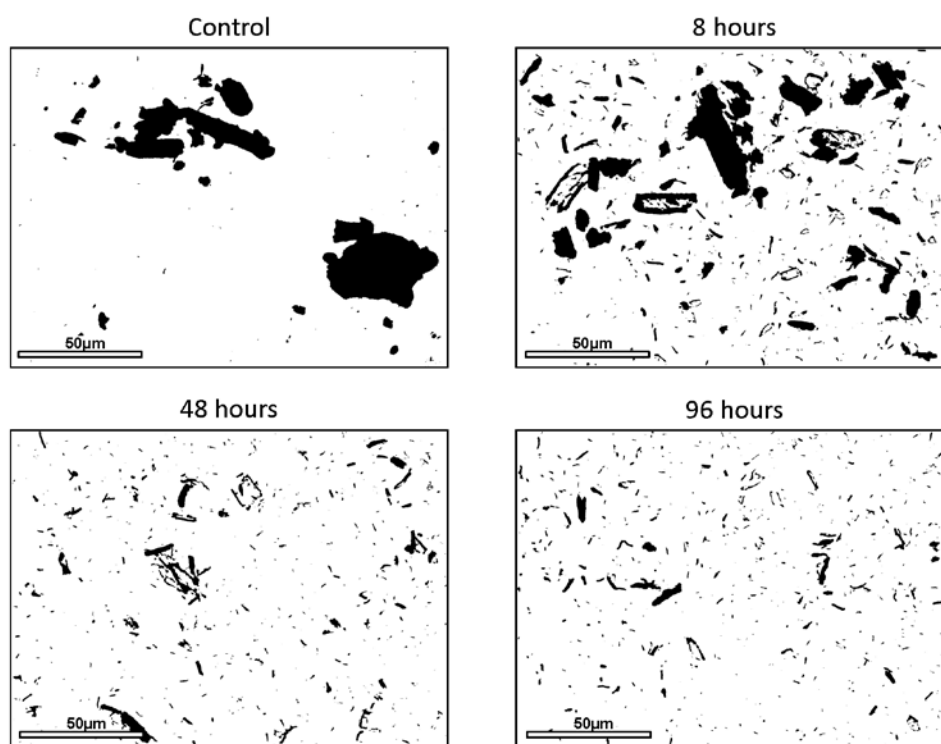


Figure 2 Detected particles in DICM micrographs of CBH1 digested Avicel after 8, 48 and 96 hours of hydrolysis, plus control incubated without enzyme for 8 hours (more details including raw images in Materials and methods and in the supplementary information). Scale bars = 50µm.

Differential interference contrast microscopy

The differential interference contrast microscopy (DICM) images covered a field of view of 195 x 146 μm with a pixel size of 122nm, and a typical micrograph contain hundreds of particles ranging from a single pixel to $\sim 50\mu\text{m}$ in length. Representative micrographs of CBH1 digested samples after 8, 48 and 96 hours of hydrolysis as well as an 8-hour control without added enzyme are shown in Figure 2. From visual inspection it is clear that the size distribution and number of particles change over time for this particular sample and similar trends were observed for all other samples (not shown). Generally we observed no obvious differences in the structural effects on this length scale among different enzymes and mixtures. ~~while different samples could often be distinguished from each other at a given time point, all samples followed similar trends when various structural parameters were considered in relation to the degree of conversion. This does not imply that the effect of the different enzymes is identical, but rather that the variation in particle sizes and shapes was large enough that any difference between samples was impossible to distinguish by eye. However we did find that control samples without added enzyme contained considerably fewer particles than enzyme treated samples and that the number of observed particles increased with conversion at least until 10% conversion after which no correlation was found (Figure 5, top). There was also a slight decrease in mean particle size that is apparent in the representative micrographs shown in Figure 2. However, the measured lengths fell into a log-normal distribution, and as such the mean does not adequately describe the typical particle (supplementary material).~~ When looking at the representative micrographs shown in Figure 2 it is clear that a relatively small number of large particles is present before the onset of the reaction, and that these break apart once enzyme is added. This was reflected in a slight decrease in the mean particle size (data not shown). However, as mentioned above the particle sizes were log-normal distributed and as such the mean

provides a very poor summary of the typical particle in a given sample. If medians were considered instead of means neither the particle perimeter (not shown) nor the length (Figure 3, top) changed with conversion. This might seem counter intuitive and at odds with what is readily observable from micrographs in Figure 2. However, the fact that the median particle length did not correlate with the degree of conversion was caused by the vast majority of particles being relatively small and similar in shape and size. In other words even though some very large particles obviously did break apart to form much smaller fragments, these were so few in numbers that they hardly affected the overall picture of the reaction mixture. Thus, if the 99th percentile lengths were plotted against conversion we observed a highly significant, negative correlation (Table 2). Conversely, for the 75th percentile, there was a very weak ($R = -0.17$) but non-significant correlation (Table 2), indicating that most of the size distribution did in fact not change during conversion, even though the very largest particles did. To ensure that this finding was not an artifact of the image analysis, we analyzed all samples with sufficient residual volume after microscopy on a coulter counter. While the results in terms of absolute values could not be compared between the two analyses, the observations that median particle size was unchanged throughout the experiment and that particle number increased initially but was unchanged during extended hydrolysis were confirmed (see supplementary material).

To determine whether the shape of the particles changed we computed their circularity. This size independent parameter, calculated as $4\pi \cdot \text{area} / \text{perimeter}^2$, is a number between 0–1 with 1 being a perfect circle. As seen in Figure 3 and Table 2 a somewhat similar albeit not identical picture was seen for this measure: We found a moderate, negative correlation for the median circularity, no significant correlation for the 25th percentile and a moderate, positive correlation for the 1st percentile. Since we see a strong, negative correlation between size and circularity (not shown) the results for the

1st percentile might reflect that the larger particles would often overlap, forming very non-circular aggregates as seen in the second and third panel of Figure 2. Conversely the shape of the typical particle, represented by the median circularity, became slightly less circular during conversion, which might be an indication of particle narrowing as previously reported for other cellulosic substrates (Boisset et al. 2000; Imai et al. 1998; Jeoh et al. 2013). For both length and circularity, the control samples were considerably different from enzyme treated samples irrespective of treatment time. Thus they were omitted from the correlation analysis and considered separately.

In summary the typical particle would not significantly shorten and only slightly narrow or in any other way change its general proportions on the μm length scale within the conversion range studied here ~~—an observation which will be addressed in more detail below.~~ Interestingly, this observation applied to all investigated enzymes and mixtures, which was in contrast to previous reports (reviewed in the introduction).

Table 2 Pearson correlation parameters for selected percentiles of particle length and circularity.

Note that the control samples without added enzyme were excluded from the regression analysis.

	Median		25th/75th percentile*		1st/99th percentile**	
	Correlation	Significance	Correlation	Significance	Correlation	Significance
Length	Very weak (R = 0.073)	None (p = 0.199)	Very weak (R = -0.17)	None (p = 0.344)	Moderate (R = -0.58)	Very high (p < 0.001)
Circularity	Moderate (R = -0.49)	High (p = 0.004)	Weak (R = -0.24)	None (p = 0.186)	Moderate (R = 0.46)	High (p = 0.007)

* 25th percentile for circularity, 75th for length. ** 1st percentile for circularity, 25th for length

~~One might expect that the typical particle size would decrease as more and more of the cellulose was solubilized, especially as this coincided with an increase in the number of particles. However, from a qualitative, visual survey of the micrographs it is apparent that a small number of relatively large particles (approximately 50 μm) were present at low degrees of conversion but disappeared as the reaction progressed (Figure 2). This~~

very small population being rapidly broken into many sub-particles might explain why we observed a constant median length but a slightly decreasing mean length.

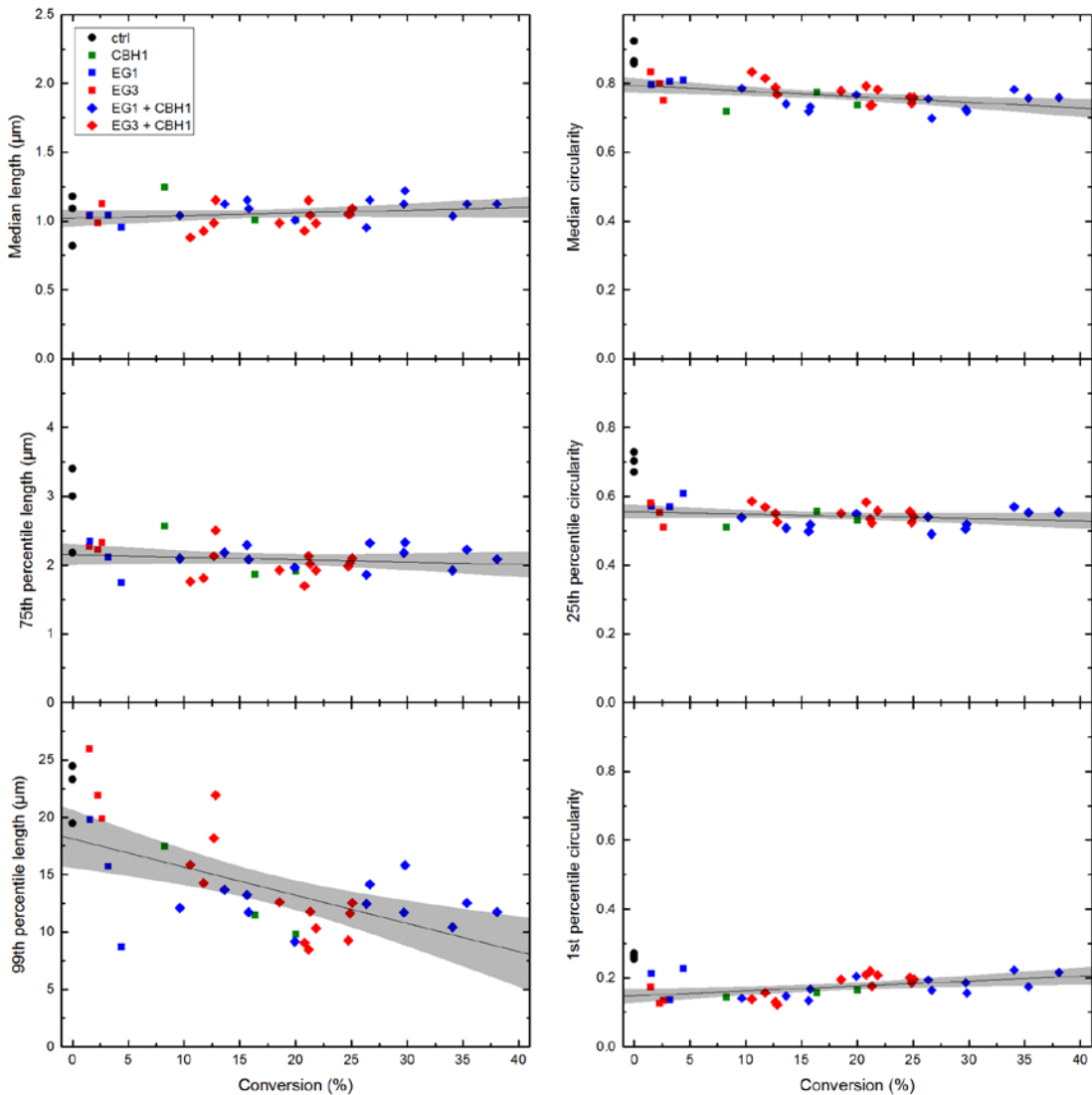


Figure 3 Particle length (left) and circularity (right) determined as the Feret's diameter measured with FIJI in DICM images plotted against conversion at 3 time points for 12 different digestions (control (black circles), CBH1 (green), EG1 (blue) and EG3 (red) alone (squares) or in 4 different mixing ratios (all represented by diamonds), see Table 1). Linear regression with 95% confidence bands is shown in dark grey. Top panels show median values, middle panels show 75th percentile length and 25th percentile circularity, bottom panels show 99th percentile length and 1st percentile circularity. Note that the control samples without added enzyme were excluded from the regression analysis. See text for further details.

330 **Transmission electron microscopy**

331 According to the DICM analysis the vast majority (>90%) of the observed particles had a
332 length of less than 5 μm . When studying the samples in the TEM we observed that the
333 particles from 1-5 μm exhibited very diverse morphologies. Many particles were
334 surrounded by tiny fibrils or subparticles, some less than 10 nm wide but with a length
335 of up to one μm and many of them overlapping to form relatively large aggregates
336 (Figure 4). Some particles appeared diffuse and still others appeared condensed with a
337 clearly defined perimeter. Importantly, no clear qualitative distinction between samples
338 was possible because of this heterogeneity; all the investigated samples (including
339 control samples without enzyme added) contained some examples of all particle
340 morphologies and no systematic changes with time or enzyme mixture could be
341 identified by visual inspection. This is not to say that the different samples looked
342 similar per se or that individual particles did not change during the hydrolysis. Rather,
343 the morphologies present in each individual sample were so diverse, that the overlap
344 between samples made unaided distinction impossible. As seen in Figure 4 the
345 individual subparticles could not be singled out in the automated analysis and for this
346 reason we were unable to quantify structural parameters on a particle level as we did
347 with DIC micrographs. Rather, for the TEM micrographs, we determined the total
348 perimeter and area of every identified particle in a single micrograph. This resulted in
349 80-90 replicates (corresponding to the number of images) for each sample.

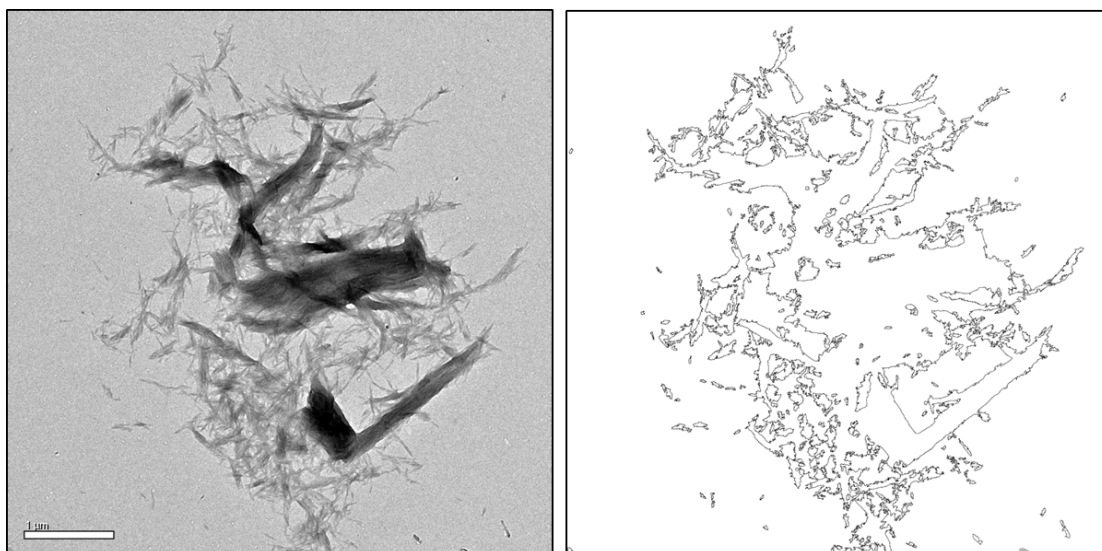


Figure 4 TEM micrograph of CBH1 digested Avicel after 8 hours of hydrolysis before image processing (left) and outlines of identified particles (right). Most sub-fibrils overlap making distinction between them impossible. Scale bar = 1μm.

Particle surface area and roughness

Micrographs present a 2D projection of particles settled on a slide or a sample grid, and based on this we cannot directly determine 3D structural information. However the hydrolysis took place in bulk suspension before the particles settled on the imaging support. Unless there was a connection between enzyme target sites and how the cellulose particles settled on the support, the particles will be randomly oriented and the sizes and shapes of their 2D projection will presumably suffice as an approximation of the 3D surface parameters. We argue that this is the case in the present study even though a thorough exploration is beyond the current scope. Using the measured length and width of the particles observed in DICM and assuming a prolate spheroid shape we approximated apparent surface areas as follows

$$A = 2\pi a^2 + \frac{2\pi ac}{e} \cdot \sin^{-1}(e)$$

366 where c is half the measured length, a is half the measured width and e is $\sqrt{c^2 - a^2} \cdot c^{-1}$
367 (for more details see supplementary material)(Zwillinger 1996). The apparent surface
368 area for each sample was calculated and as seen from Figure 5, this value increased
369 during the first $\sim 10\%$ conversion and then gradually decreased throughout the
370 investigated conversion range. In contrast to the median values reported above – which
371 gives information about the size and shape of the typical particle in a sample – the total
372 surface area per micrograph is a relative measure of the “concentration” of substrate.
373 With that in mind our results indicated that even though the μm scale size and shape of
374 a typical Avicel particle was similar throughout the reaction, the apparent total surface
375 area in the sample did change with conversion.

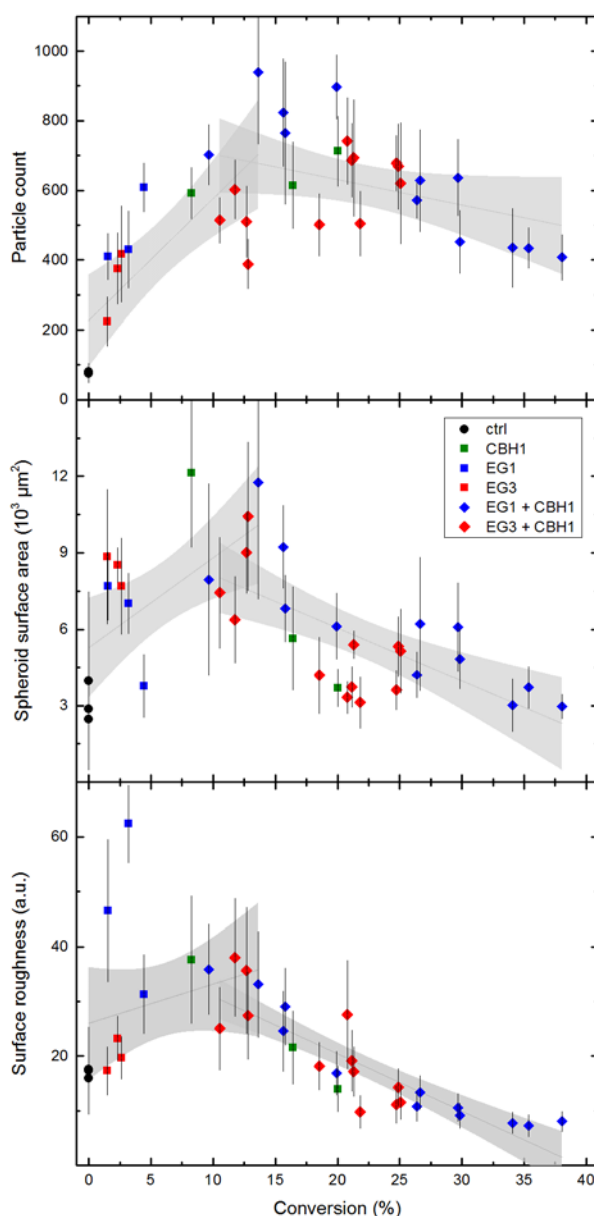


Figure 5 Average number of particles per DICM micrograph (top), average surface area calculated from DICM measurements (middle) and median surface roughness determined by TEM (bottom) all plotted against degree of conversion. Error bars represent standard deviation for 5 analytical replicates in the two uppermost panels and median absolute deviation in the bottom panel ($n \approx 80-90$). All three measures increase during the first 10-15% conversion, at which point the number of detected particles stay roughly constant while the surface area and – roughness start to decrease (linear regression with 95% confidence bands shown in gray. See Table 3 for non-linear correlation parameters). Symbols and colors have the same meaning as in Figure 3.

As described above, TEM revealed a high prevalence of composite structures comprising many small overlapping subparticles. We could not distinguish the individual subparticles in the automated analysis therefore we used the total perimeter and area in each micrograph to evaluate the relative surface roughness. We calculated the perimeter of a hypothetical circle with the same area as the total measured area for each micrograph (supplementary material). We then used the ratio of the observed perimeter to this theoretical perimeter as a relative measure of surface roughness on the nm scale. When plotted against overall substrate conversion (Figure 5) we observed a similar pattern to the approximated surface area calculated from DICM data; a slight initial increase followed by a marked decrease until it apparently leveled off at the highest conversions studied here (~30%). The progressions in particle count, surface area and surface roughness depicted in Figure 5 were all non-monotone, and neither of them was suited for any parameterized regression methods. Using a Hoeffding's independence test we were able to verify that the variations in the determined values were significantly dependent on the degree of conversion ($p < 0.001$ in all three cases). To analyze the correlation in detail we identified a maximum in all three data sets around 10-15% conversion and carried out separate correlation analyses for the data before and after this maximum. The results confirm our intuitive interpretation of the data, i.e. all three measures increased until 10-15% conversion, at which point the particle count stayed roughly constant while the surface area and – roughness started to decrease (see Figure 5 and Table 3).

Table 3 Spearman's correlation parameters for plots in Figure 5. Since the progressions in all three values were non-monotone the data set was divided into two parts which were analyzed separately.

	Conversion 0 - 15%		Conversion 10 - 40%	
	Correlation	Significance	Correlation	Significance
Particle count	Strong ($\rho = 0.78$)	Very high ($p < 0.001$)	Weak ($\rho = -0.30$)	None ($p = 0.14$)

Surface area	Moderate ($\rho = 0.59$)	Moderate ($p = 0.0155$)	Strong ($\rho = -0.68$)	Very high ($p < 0.001$)
Roughness	Moderate ($\rho = 0.56$)	Moderate ($p = 0.029$)	Very strong ($\rho = -0.91$)	Very high ($p < 0.001$)

408 Based on the results of the automated analysis we suggest that sub- μm fibrils are
409 predominantly present in the early stages of the reaction. This would indicate that the
410 actual surface area as experienced by an enzyme is decreasing in a manner, which is not
411 readily observed on the length scale used in the DICM analysis. The relative surface
412 roughness determined by TEM is a dimensionless number, equivalent to actual surface
413 area (nm resolution) per approximated surface area (μm resolution). Thus we used the
414 relative surface roughness determined by TEM to correct the surface area calculated
415 from DICM measurements by multiplying these two values. The resulting parameter was
416 used to **estimate the total** surface area within the studied samples. Plotting the rate of
417 the reaction for CBH1 alone and in mixtures against this measure gave a strong linear
418 correlation as seen in Figure 6 ($R = 0.81$, $p < 0.0001$). **Because the volume of sample in**
419 **the microscope image frame is unknown we were unable to relate the determined**
420 **surface area to the amount of cellulose, which makes validation by independent**
421 **measurement or literature reference impossible. However we can roughly estimate the**
422 **liquid volume in a single DIC micrograph to be on the order of 10^{-4} μl which, at a**
423 **cellulose concentration of $10 \mu\text{g}/\mu\text{l}$ and specific surface area of $1\text{-}25 \text{ m}^2/\text{g}$ Avicel**
424 **(Marshall and Sixsmith 1975), puts the expected area at $1\text{-}25 \text{ mm}^2$ per image.**
425 **Considering the approximations and measuring uncertainties we deem this to be in**
426 **relatively good agreement with our results (see Figure 6).**

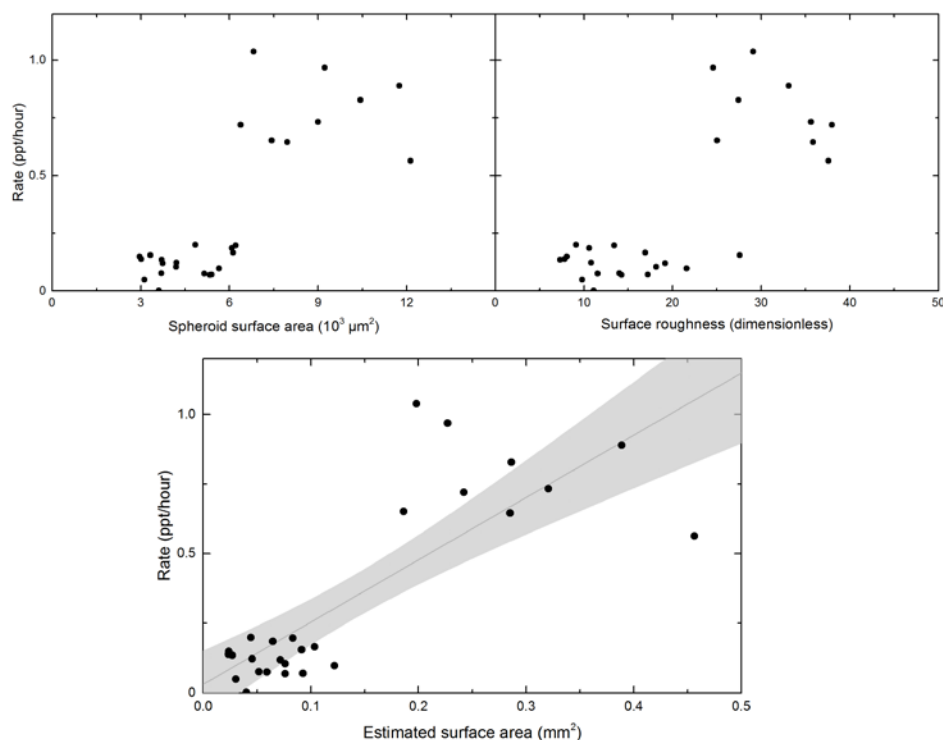


Figure 6 Reaction rate (percentage points per hour) plotted against calculated surface area determined by DICM (top, left) and surface roughness determined by TEM (top, right) as described in the text. The reaction rate in samples digested by CBH1 alone or in mixtures correlated linearly (dotted line) to the accessible surface area in the microscope field of view under the conditions studied here (bottom panel, 95% confidence bands shown in gray).

Discussion and Conclusions

We have used a combination of DICM and TEM to observe and measure changes in Avicel particle size and shape during enzymatic hydrolysis. We developed a method to perform a reliable and unbiased quantitation of size, shape and surface roughness of these particles on a μm-nm scale. Based on analyses of 180 DICM and 2950 TEM micrographs comprising a total of more than 300,000 particles we found that for the enzymes studied here, the changes in shape and surface structure of Avicel was not determined by the type of enzyme or composition of enzyme mixtures but rather they depended on the degree of conversion that had taken place at the time of observation.

442 We observed a marked increase in particle count with **initial** conversion but no change
443 in the median size of Avicel particles on a μm scale. **This somewhat counter intuitive**
444 **result (which was independently confirmed by coulter counter analysis – see**
445 **supplementary material) indicates that even though the very largest particles were**
446 **dramatically altered by enzymatic degradation, the general size distribution was by and**
447 **large unaltered during the experiment. Furthermore we observed only a slight change in**
448 **the shape of particles on this scale.** On a nm scale we observed a very large degree of
449 heterogeneity in the surface structure of Avicel, but by increasing the sampling size we
450 obtained reliable data on the surface roughness at this scale. **Note that the data**
451 **acquisition was done blind to avoid any bias in the selection of captured particles.** When
452 quantifying changes in relative surface area on both μm and nm scale we found a slight
453 initial increase followed by a continuous decrease throughout the studied range of
454 conversion. Based on these observations we suggest that most particles were in the
455 form of large superparticle aggregates prior to enzyme addition and that the initial
456 activity of the studied enzymes rapidly broke these aggregates apart to form a large
457 number of much smaller particles. In the following **(still early)** phase of reaction a large
458 fraction of the cellulose was in the form of sub- μm fibrils that were difficult to
459 distinguish by DICM. These fibrils appear to be relatively easy to degrade and their
460 presence decreased as the hydrolysis progressed, leaving compact (possibly
461 recalcitrant) particles behind in accordance with previous observations (Chanzy et al.
462 1983; Imai et al. 1998; Jeoh et al. 2013). By combining the DICM and TEM
463 measurements we obtained a relative measure of the **available** surface area over the
464 course of hydrolysis and, as was the case for the two values separately, this number
465 initially increased slightly but subsequently decreased as the hydrolysis progressed to
466 around 30% conversion where it leveled off **(not shown). It should be noted that for**
467 **experimental reasons our analysis does not go beyond this conversion and that the**
468 **results might not hold for higher degrees of conversion. Nonetheless, our observation is**

in agreement with the degree of conversion at which Bansal and colleagues determined a leveling off of both accessibility and hydrolysability (Bansal et al. 2012). Furthermore we found that the rate of reaction for CBH1 alone and in mixtures was closely correlated to the relative accessible surface area in the conversion range studied here (Figure 6). Based on these findings we conclude that surface alterations on both μm and nm length scales do indeed impact the rate loss that is unambiguously observed for cellulases in their hydrolysis of cellulose.

Electronic supporting material available

Supporting information includes details and examples on the automated image processing and analysis procedures for both DICM and TEM micrographs. Furthermore it includes results from coulter counter measurements of particle size distribution and number of particles in the reaction mixture. Finally, also found in the supporting information is details on the distribution of DICM data and on the geometry behind the conversion from 2D to 3D structural information.

References

- Antal MJ, Jr. (1985) Biomass pyrolysis: A review of the literature part 1 - carbohydrate pyrolysis. In: Böer KW, Duffie JA (eds) *Advances in solar energy*. Springer New York, pp 61-111. doi:10.1007/978-1-4684-8992-7_3
- Arantes V, Saddler JN (2011) Cellulose accessibility limits the effectiveness of minimum cellulase loading on the efficient hydrolysis of pretreated lignocellulosic substrates. *Biotechnology for Biofuels* 4(3)
- Bansal P, Hall M, Realff MJ, Lee JH, Bommarius AS (2009) Modeling cellulase kinetics on lignocellulosic substrates. *Biotechnology Advances* 27(6): 833-848
- Bansal P, Vowell BJ, Hall M, Realff MJ, Lee JH, Bommarius AS (2012) Elucidation of cellulose accessibility, hydrolysability and reactivity as the major limitations in the enzymatic hydrolysis of cellulose. *Bioresour Technol* 107: 243-250
- Boisset C, Fraschini C, Schülein M, Henrissat B, Chanzy H (2000) Imaging the enzymatic digestion of bacterial cellulose ribbons reveals the endo character of the cellobiohydrolase cel6a from *humicola insolens* and its mode of synergy with

498 cellobiohydrolase cel7a. *Applied and environmental microbiology* 66(4): 1444-
499 1452

500 Bubner P, Dohr J, Plank H, Mayrhofer C, Nidetzky B (2012) Cellulases dig deep - in situ
501 observation of the mesoscopic structural dynamics of enzymatic cellulose
502 degradation. *Journal of Biological Chemistry* 287(4): 2759-2765

503 Bubner P, Plank H, Nidetzky B (2013) Visualizing cellulase activity. *Biotechnology and*
504 *bioengineering* 110(6): 1529-1549

505 Chanzy H, Henrissat B (1985) Unidirectional degradation of valonia cellulose
506 microcrystals subjected to cellulase action. *FEBS Letters* 184(2): 285-288

507 Chanzy H, Henrissat B, Vuong R, Schülein M (1983) The action of 1,4- β -d-glucan
508 cellobiohydrolase on valonia cellulose microcrystals. *FEBS Letters* 153(1): 113-
509 118

510 Ciesielski PN, Wang W, Chen X, Vinzant TB, Tucker MP, Decker SR, Himmel ME, Johnson
511 DK, Donohoe BS (2014) Effect of mechanical disruption on the effectiveness of
512 three reactors used for dilute acid pretreatment of corn stover part 2:
513 Morphological and structural substrate analysis. *Biotechnol Biofuels* 7: 47

514 Donohoe BS, Resch MG (2015) Mechanisms employed by cellulase systems to gain
515 access through the complex architecture of lignocellulosic substrates. *Curr Opin*
516 *Chem Biol* 29: 100-107

517 Eriksson T, Karlsson J, Tjerneld F (2002) A model explaining declining rate in hydrolysis
518 of lignocellulose substrates with cellobiohydrolase i (cel7a) and endoglucanase i
519 (cel7b) of trichoderma reesei. *Applied Biochemistry and Biotechnology* 101: 41-
520 60

521 Grethlein HE (1985) The effect of pore-size distribution on the rate of enzymatic-
522 hydrolysis of cellulosic substrates. *Bio-Technol* 3(2): 155-160

523 Hall M, Bansal P, Lee JH, Realff MJ, Bommarius AS (2010) Cellulose crystallinity - a key
524 predictor of the enzymatic hydrolysis rate. *FEBS Journal* 277(6): 12-12

525 Hong J, Ye X, Zhang YHP (2007) Quantitative determination of cellulose accessibility to
526 cellulase based on adsorption of a nonhydrolytic fusion protein containing cbm
527 and gfp with its applications. *Langmuir : the ACS journal of surfaces and colloids*
528 23(25): 12535-12540

529 Imai T, Boisset C, Samejima M, Igarashi K, Sugiyama J (1998) Unidirectional processive
530 action of cellobiohydrolase cel7a on valonia cellulose microcrystals. *FEBS*
531 *Letters* 432(3): 113-116

532 Jalak J, Väljamäe P (2010) Mechanism of initial rapid rate retardation in
533 cellobiohydrolase catalyzed cellulose hydrolysis. *Biotechnology and*
534 *Bioengineering* 106(6): 871-883

535 Jeoh T, Ishizawa CI, Davis MF, Himmel ME, Adney WS, Johnson DK (2007) Cellulase
536 digestibility of pretreated biomass is limited by cellulose accessibility.
537 *Biotechnology and bioengineering* 98(1): 112-122

538 Jeoh T, Santa-Maria MC, O'Dell PJ (2013) Assessing cellulose microfibrillar structure
539 changes due to cellulase action. *Carbohydrate Polymers* 97(2): 581-586

540 Kapur JN, Sahoo PK, Wong AKC (1985) A new method for gray-level picture
541 thresholding using the entropy of the histogram. *Computer Vision, Graphics, and*
542 *Image Processing* 29(3): 273-285

543 Levine SE, Fox JM, Blanch HW, Clark DS (2010) A mechanistic model of the enzymatic
544 hydrolysis of cellulose. *Biotechnology and Bioengineering* 107(1): 37-51

545 Luterbacher JS, Moran-Mirabal JM, Burkholder EW, Walker LP (2015) Modeling
546 enzymatic hydrolysis of lignocellulosic substrates using confocal fluorescence
547 microscopy i: Filter paper cellulose. *Biotechnology and Bioengineering* 112(1):
548 21-31

549 Marshall K, Sixsmith D (1975) Some physical characteristics of microcrystalline
550 cellulose .1. Powders for pharmaceutical use. *Drug Dev Commun* 1(1): 51-71

551 Payne CM, Knott BC, Mayes HB, Hansson H, Himmel ME, Sandgren M, Stahlberg J,
552 Beckham GT (2015) Fungal cellulases. *Chem Rev* 115(3): 1308-1448

553 Praestgaard E, Elmerdahl J, Murphy L, Nymand S, McFarland KC, Borch K, Westh P
554 (2011) A kinetic model for the burst phase of processive cellulases. *FEBS Journal*
555 278(9): 1547-1560

556 Resch MG et al. (2014) Clean fractionation pretreatment reduces enzyme loadings for
557 biomass saccharification and reveals the mechanism of free and cellulosomal
558 enzyme synergy. *ACS Sustainable Chemistry & Engineering* 2(6): 1377-1387

559 Rogers EW, Zack GW, Latt SA (1977) Automatic measurement of sister chromatid
560 exchange frequency. *The Journal of Histochemistry and Cytochemistry* 25(7):
561 741-753

562 Schindelin J et al. (2012) Fiji: An open-source platform for biological-image analysis. *Nat*
563 *Methods* 9(7): 676-682

564 Sprey B, Bochem H-P (1992) Effect of endoglucanase and cellobiohydrolase from
565 *trichoderma reesei* on cellulose microfibril structure. *FEMS Microbiology Letters*
566 97(1-2): 113-117

567 Wang J, Quirk A, Lipkowski J, Dutcher JR, Clarke AJ (2013) Direct in situ observation of
568 synergism between cellulolytic enzymes during the biodegradation of crystalline
569 cellulose fibers. *Langmuir : the ACS journal of surfaces and colloids* 29(48):
570 14997-15005

571 Westh P, Kari J, Olsen JP, Borch K, Jensen K, Krogh KBRM (2014) Cellobiohydrolase
572 variants and polynucleotides encoding same. Patent no. WO 2014064115 A1

573 Yang B, Willies DM, Wyman CE (2006) Changes in the enzymatic hydrolysis rate of avicel
574 cellulose with conversion. *Biotechnology and Bioengineering* 94(6): 1122-1128

575 Zhang YHP, Lynd LR (2004) Toward an aggregated understanding of enzymatic
576 hydrolysis of cellulose: Noncomplexed cellulase systems. *Biotechnology and*
577 *Bioengineering* 88(7): 797-824

578 Zwillinger D (ed) (1996) Crc standard mathematical tables and formulae. 30th edn. CRC-
579 Press, Boca Raton, Florida
580

Supplementary material

Image processing – DICM:

Figure S1 shows a step-by-step review of the automated processing of a representative DICM micrograph as carried out in Fiji. The sample was Avicel digested with Cel7A for 8 hours. Contrast was increased by the “Enhance Contrast” function with saturated pixels set to 0.5% and “Normalize” option checked (A). Next the image was converted to binary by (B) applying the “Yen” threshold algorithm (Kapur et al. 1985), and the outermost 60 pixels along all four edges were cropped away (C). Finally the “Fill Holes” command was applied (D) and the particle perimeters and areas were determined by the “Analyze Particles” function (E).

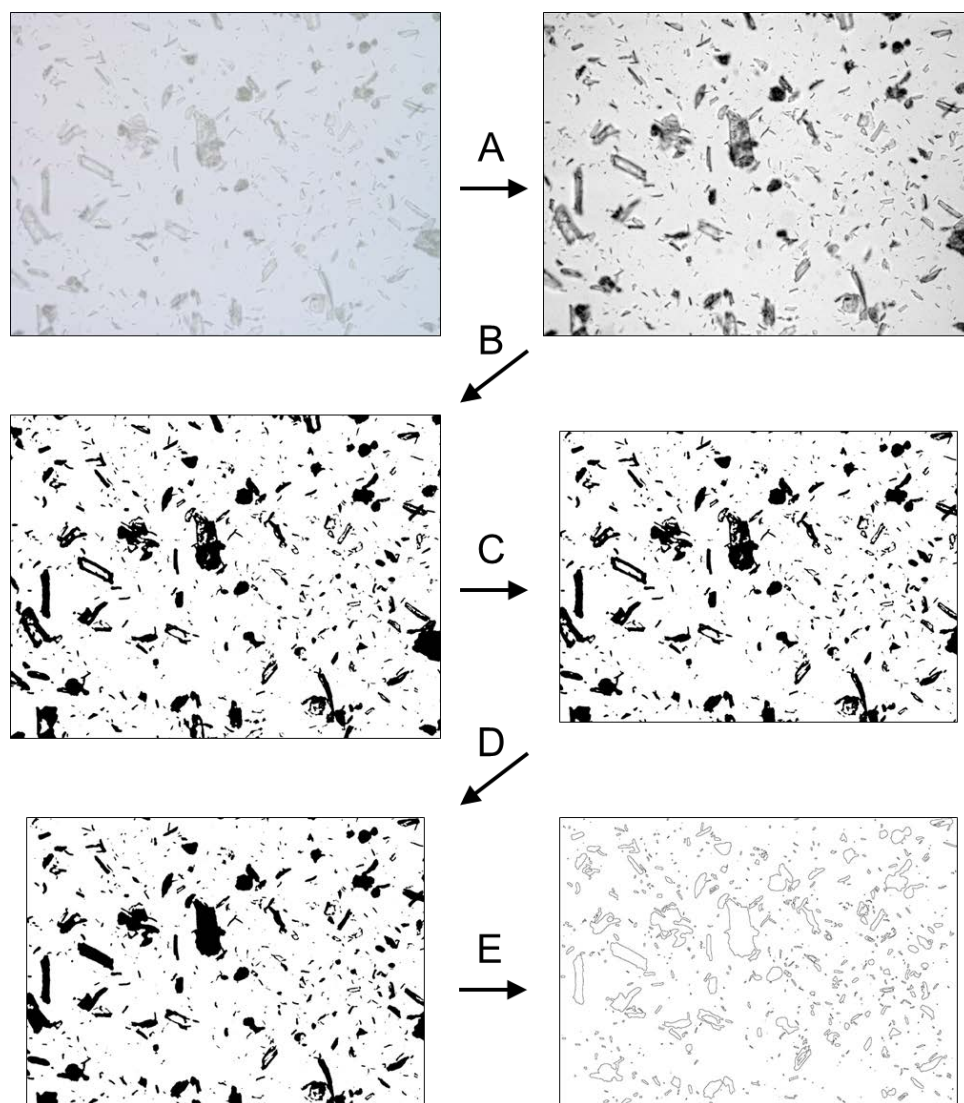


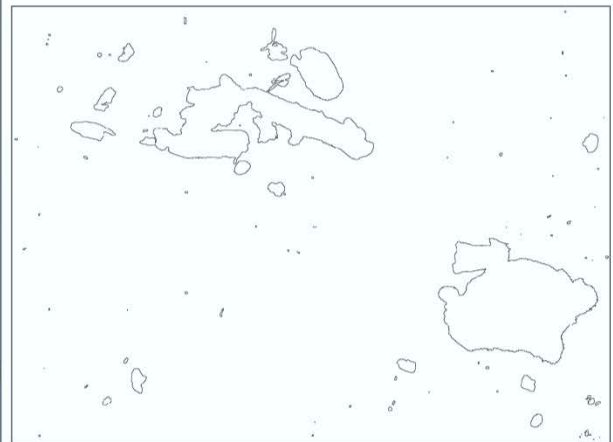
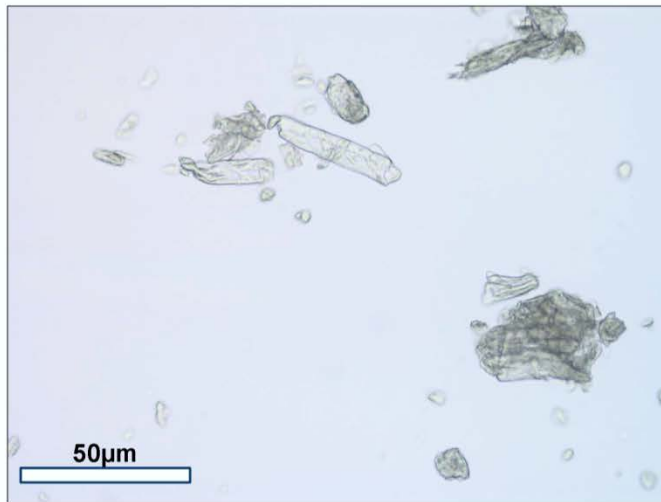
Figure S1 Step-by-step review of the processing of a DICM micrograph.

1

2 **Representative DIC micrographs:**

3 Figure S2 shows the same representative DIC micrographs as Figure 2 in the main text before and after image
4 processing

Control

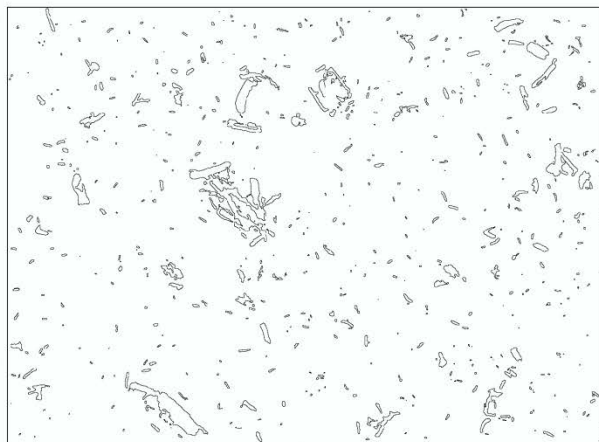
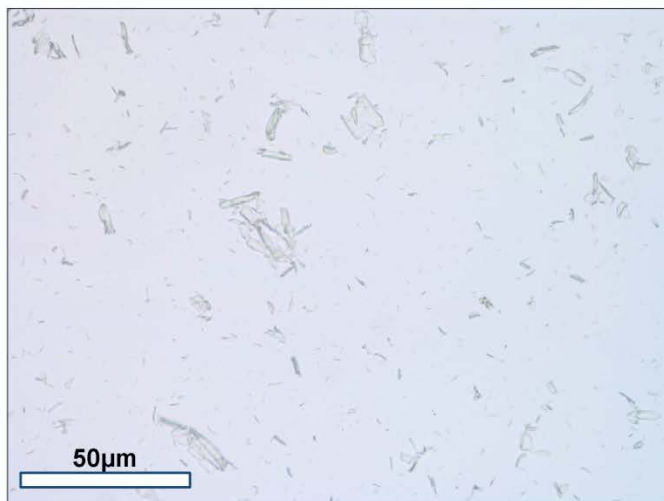


8 hours

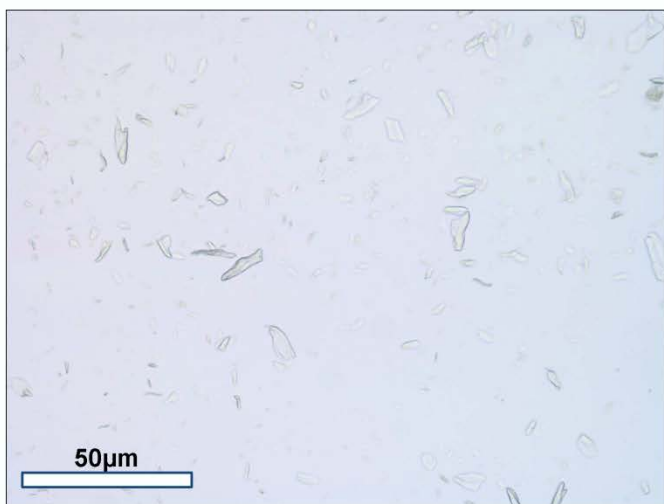


5

48 hours



96 hours



1

2 Figure S2 Representative micrographs shown in Figure 2 in the main text before and after processing.

3

Coulter counter results

Figure S3 shows results from coulter counter analysis of particle sizes analogously to Figure 3 in the main text. The observation from DICM that the median size did not change with conversion is even more pronounced for this measurement. Thus no correlation between diameter and conversion was observed for even for the 99th percentile (not shown), but for the 99.99th percentile a moderate correlation appeared (Figure S3, bottom). Correlation parameters are shown in Table S1.

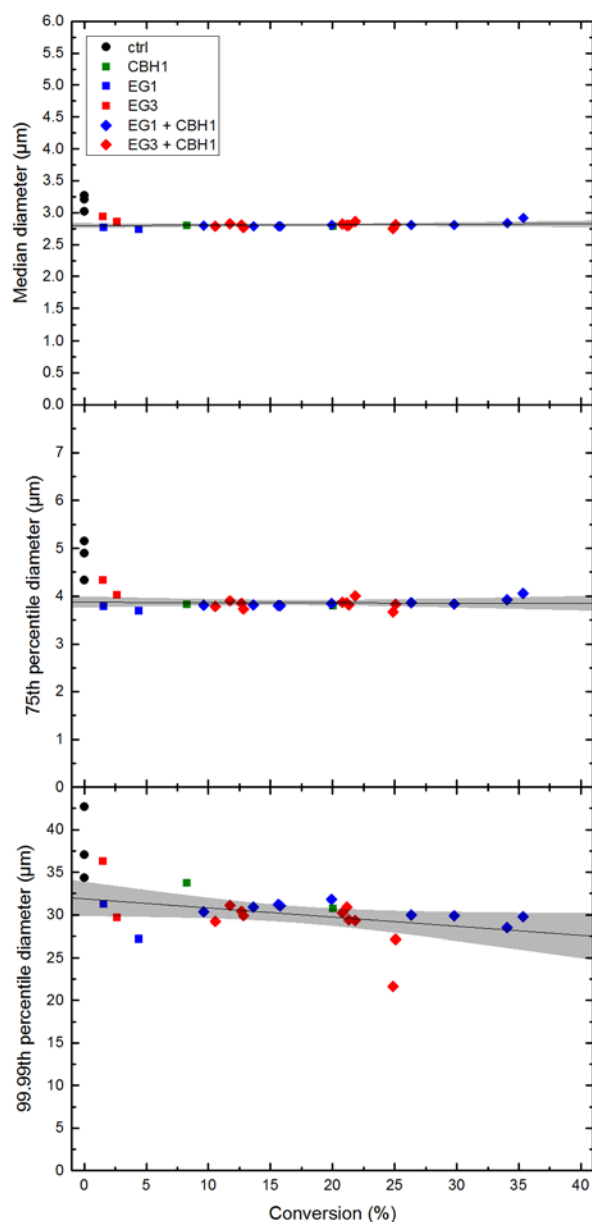


Figure S3 Median (top), 75th percentile (middle) and 99.99th percentile (bottom) particle diameter measured on coulter counter. Symbols and colors have the same meaning as in Figure 3 in the main text.

Table S1 Correlation analysis for particle diameter measured on coulter counter

	Median		75th percentile		99.99th percentile	
	Correlation	Significance	Correlation	Significance	Correlation	Significance
Length	Very weak (R = 0.14)	None (p = 0.51)	Very weak (R = -0.05)	None (p = 0.82)	Moderate (R = -0.40)	Moderate (p = 0.046)

Figure S4 shows the particle count determined by coulter counter plotted against conversion. Correlation parameters for the same data are shown in Table S2. In agreement with microscopy results reported in the main text, the number of particles increases during the early phase of the reaction but is unchanged throughout the rest of the experiment.

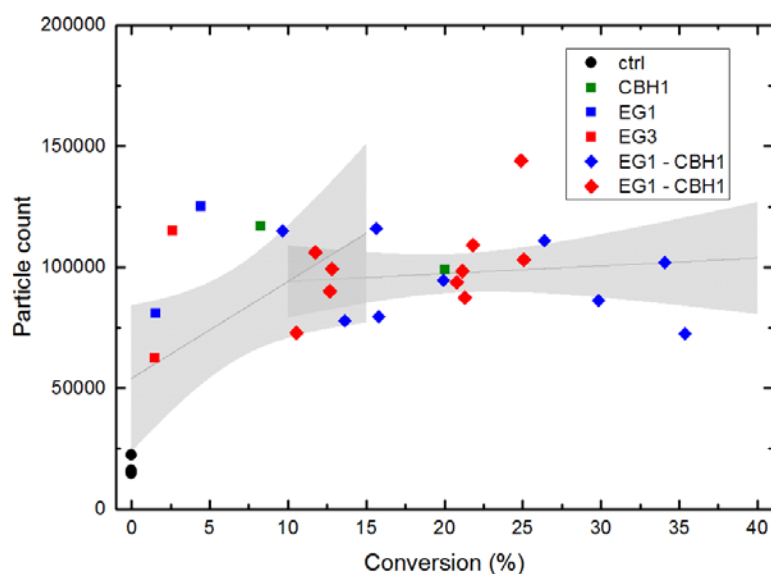


Figure S4 Particle count determined by coulter counter plotted against conversion. Symbols and colors have the same meaning as in Figure 3 in the main text.

Table S2 Correlation analysis for particle count determined by coulter counter

	Conversion 0 - 15%		Conversion 10 - 40%	
	Correlation	Significance	Correlation	Significance
Particle count	Moderate (p = 0.47)	Low (p = 0.088)	Very weak (p = 0.14)	None (p = 0.57)

Image processing – TEM:

Figure S5 shows a step-by-step review of the processing of a representative image from TEM. The sample was digested with CBH1 for 8 hr. Noise was minimized by the “Remove Outliers” option (A), which changes the value of any pixel to the median of all pixel values in a surrounding block of designated size (here, 10x10 pixel) if that pixel deviates from that median by more than a specified value (1x standard deviation). Subsequently the images were converted from grayscale to a binary mask by applying the “Triangle” threshold method (Rogers et al. 1977) (B) and the perimeter and area of the individual particles were analyzed by Fiji’s “Analyze Particles” procedure (C). The mask was also black/white inverted (D) allowing for a similar analysis of the holes (E). The resulting two data files were combined with a script in Matlab (R2013a v 8.1.0.604, Natick, Ma, USA), hence (F) is included for illustrative purposes.

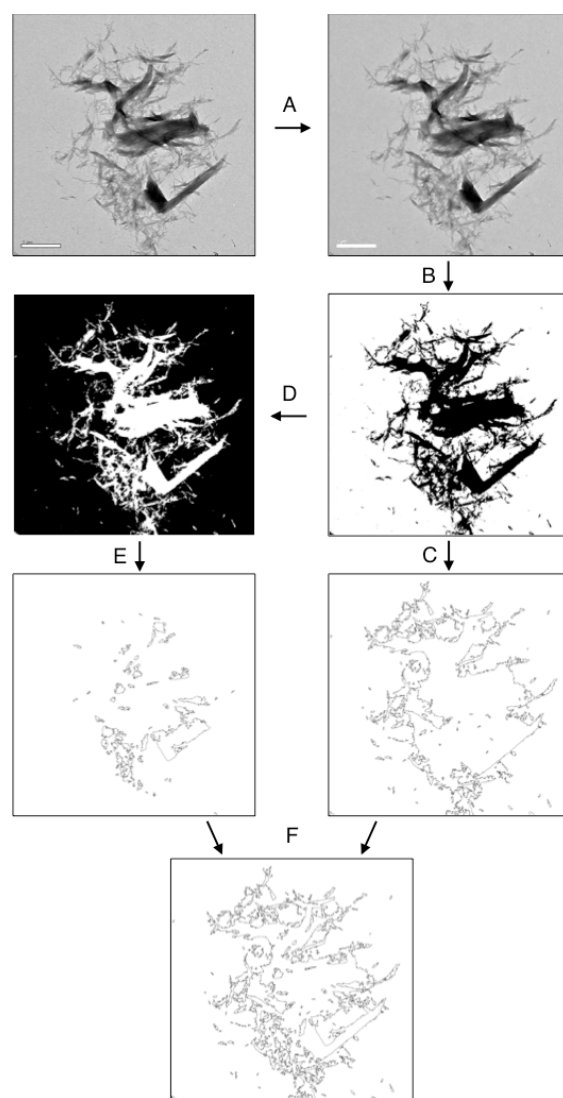


Figure S5 Step-by-step review of the processing of a TEM micrograph (See text for details).

Distribution of particle sizes

The particle lengths, widths, areas and perimeters measured in DICM did not follow a normal distribution. Plots and analyses shown in Figure S6 are for particle length after 8 hours digestion with CBH1, but similar patterns applied for all data. Data was clearly skewed with most observations falling into a relatively small quantile combined with a long “tail” of observations over a very large size distribution (histogram in Figure S6, left). When data was plotted against a theoretical normal distribution in a normal probability (Q-Q) plot, deviation from linearity clearly confirmed, that the data was not normal distributed. When the same plots were made for the log-transformed data (Figure S6, right), both histogram and Q-Q plot confirmed that data fell into a log-normal distribution.

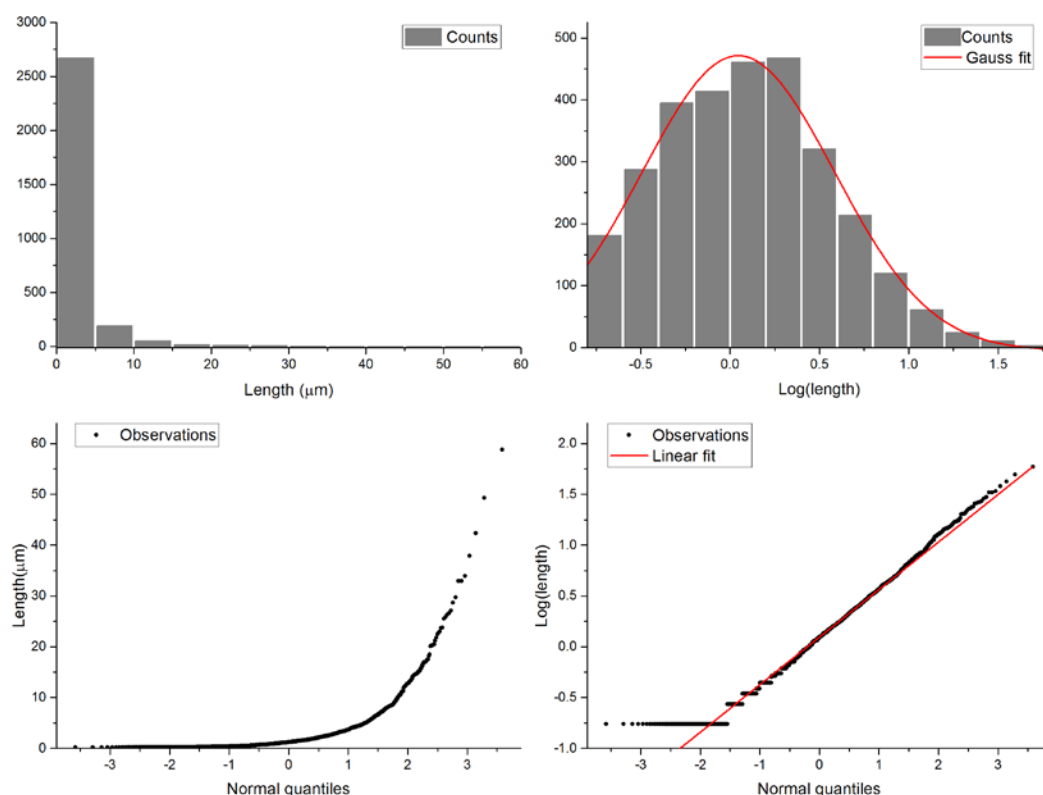


Figure S6 Left: Histogram and normal probability (Q-Q) plot of Avicel particle lengths after 8 hours of digestion by CBH1. From the histogram it is clear the data is skewed with most observations falling between 0 and 5 μm but some being as high as 60 μm . Right: Same as left after log-transformation of data. The histogram fits a normal distribution reasonably well and the log-normal distribution of the data was confirmed by the linear correlation with a theoretical normal distribution in the Q-Q plot. Results for the smallest particles were poorly resolved probably due to the image resolution.

Estimating surface area from 2D projection of a particle

In the analysis presented in the main text we assume that the surface of avicel particles can be approximated from DICM observations by treating each particle as a prolate spheroid, i.e. the 3-dimensional structure obtained by rotating an ellipse about its longest axis (see Figure S7).

The surface area of a spheroid can be calculated as follows (Zwillinger 1996):

$$A = 2\pi a^2 + \frac{2\pi ac^2}{\sqrt{c^2 - a^2}} \cdot \sin^{-1}\left(\frac{\sqrt{c^2 - a^2}}{c}\right) = 2\pi a^2 + \frac{2\pi ac}{e} \cdot \sin^{-1}(e)$$

Where $e = \sqrt{c^2 - a^2} \cdot c^{-1}$ while a and c are the minor and major radii, respectively.

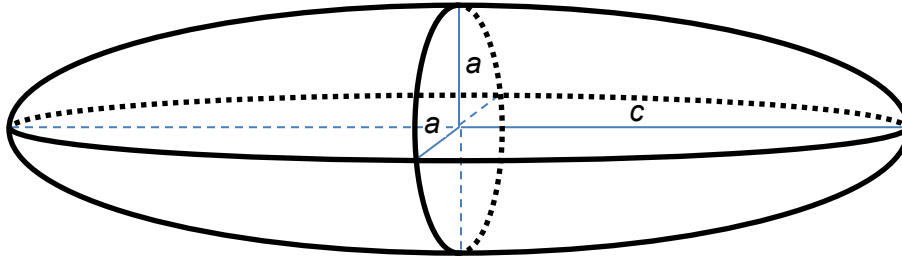


Figure S7 Particle shape was approximated as a prolate spheroid.

References

- Kapur JN, Sahoo PK, Wong AKC (1985) A new method for gray-level picture thresholding using the entropy of the histogram. *Computer Vision, Graphics, and Image Processing* 29(3): 273-285
- Zwillinger D (ed) (1996) *Crc standard mathematical tables and formulae*. 30th edn. CRC-Press, Boca Raton, Florida