Journal of proteome research

Article

pubs.acs.org/jpr

Development of a Multipoint Quantitation Method to Simultaneously Measure Enzymatic and Structural Components of the *Clostridium thermocellum* Cellulosome Protein Complex

Andrew B. Dykstra,[†,‡,#] Lois St. Brice,[§,∥] Miguel Rodriguez, Jr.,[†] Babu Raman,[†,⊥] Javier Izquierdo,[§,○] Kelsey D. Cook,[‡,●] Lee R. Lynd,[§] and Robert L. Hettich*[,†]

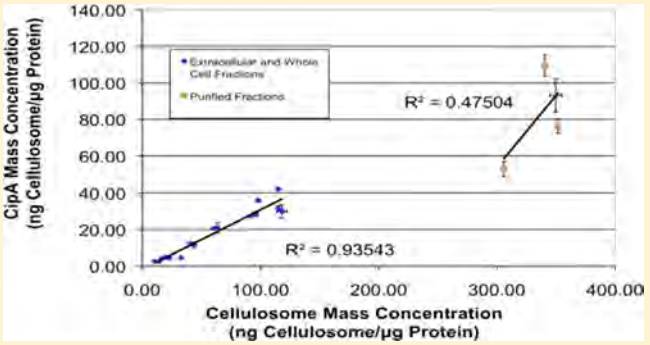
[†]Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-6131, United States

[‡]University of Tennessee, Knoxville, Tennessee 37996, United States

[§]Dartmouth College, Hanover, New Hampshire 03755, United States

Supporting Information

ABSTRACT: *Clostridium thermocellum* has emerged as a leading bioenergy-relevant microbe due to its ability to solubilize cellulose into carbohydrates, mediated by multicomponent membrane-attached complexes termed cellulosomes. To probe microbial cellulose utilization rates, it is desirable to be able to measure the concentrations of saccharolytic enzymes and estimate the total amount of cellulosome present on a mass basis. Current cellulase determination methodologies involve labor-intensive purification procedures and only allow for indirect determination of abundance. We have developed a method using multiple reaction monitoring (MRM-MS) to simultaneously quantitate both enzymatic and structural components of the cellulosome protein complex in samples ranging in complexity from purified cellulosomes to whole cell lysates, as an alternative to a previously developed enzyme-linked immunosorbent assay (ELISA) method of cellulosome quantitation. The precision of the cellulosome mass concentration in technical replicates is better than 5% relative standard deviation for all samples, indicating high precision for determination of the mass concentration of cellulosome components.



KEYWORDS: *protein quantitation, absolute quantitation, Clostridium thermocellum, cellulosome, enzyme-linked immunosorbent assay, multiple reaction monitoring mass spectrometry*

INTRODUCTION

The growing importance of protein quantitation for studies of biological systems has necessitated a concomitant need for sensitive and selective methods that afford high accuracy and precision measurements. However, quantitating specific proteins in a mixture is difficult to achieve with optical assays such as Bradford,[1] bicinchoninic acid (BCA),[2,3] and Lowry[4] assays because these assays quantitate all proteins present and lack any sort of selectivity toward specific proteins of interest. One of the most common methods of quantitating a single protein in a complex mixture is an enzyme-linked immunosorbent assay (ELISA).[5] This method has gained widespread use for targeted protein quantitation, particularly in clinical settings, where ELISA is often considered the gold standard for diagnoses of conditions ranging from pregnancy[6] to diseases such as human immunodeficiency virus (HIV)[7] and malaria.[8] However, antibodies specific to a protein of interest may not be commercially available at cost-effective prices or at all, and development of an antibody is a difficult, labor-intensive process. Furthermore, specificity of antibodies toward antigens varies, and binding of proteins with domains similar to the protein of interest cannot be completely prevented. In addition, ELISA is a single point measurement based on a single peptide sequence. This may be useful if only a single protein is targeted for quantitation, but use of a single protein to quantitate multicomponent protein complexes involves either making nontrivial assumptions or extensive calibration.

The use of mass spectrometry for protein quantitation has become increasingly prevalent. In particular, label-free (normalized spectral abundance factor (NSAF),[9] protein abundance index (PAI),[10] and exponentially modified protein abundance index (emPAIs)[11]) as well as label-based (isotope-coded affinity tag (ICAT),[12] stable isotope labeling by amino acids in cell culture (SILAC[13,14]), isobaric tag for relative and absolute quantitation (iTRAQ),[15] and tandem mass tag (TMT)[16]) methods of quantitation have gained widespread use in relative quantitation studies. However, mass spectra of

Received: August 1, 2013
Published: November 25, 2013

complex mixtures are not inherently quantitative,[17] so all of these methods provide only for the relative quantitation of proteins of interest under the specified sets of conditions.

In order to determine the actual quantity of a protein or number of protein molecules, internal standards must be employed. Quantitation with stable isotope-labeled (SIL) standards utilizes isotopically labeled peptides with very accurately known concentrations.[18] By comparing the ratios of chromatographic peak areas from unlabeled peptides to those of labeled peptides with precisely and accurately known amounts that have been spiked into the sample, the amount of the unlabeled peptide, and thus the unlabeled protein, can be derived. The operation of triple quadrupole (QqQ) mass spectrometers in selected reaction monitoring (SRM) mode increases the sensitivity and selectivity of these quantitation experiments, thereby allowing for the quantitation of multiple proteins in a single experiment with high accuracy and precision.[19]

Clostridium thermocellum has one of the fastest observed growth rates on cellulose,[20] due primarily to an extracellular multicomponent protein complex referred to as the cellulosome. A key component of the *C. thermocellum* cellulosome is the scaffoldin protein, also referred to as the cellulase-integrating protein A (CipA). The scaffoldin consists of a cellulose-specific carbohydrate binding module (CBM) that mediates cellulose binding, nine different Type I cohesin modules that bind with a high affinity to Type I dockerin domains of catalytic subunits, and a C-terminal Type II dockerin domain that binds with high affinity to Type II cohesin domains of anchoring proteins.[21] The three most prominent anchoring proteins, scaffoldin-dockerin binding component A (SbdA), open reading frame 2p (Orf2p), and outer layer protein B (OlpB), each have three repeated surface-layer homology (SLH) domains that anchor the proteins into the extracellular surface of the cell membrane. The dominant form of carbohydrate uptake in *C. thermocellum* has been observed as oligosaccharides with a mean length of four.[22] Cellulolytic enzymes associated with the cellulosome have a variety of functions, including exo- and endoacting enzymes that break β-glucosidic bonds, (CelA-CelW), as well as enzymes such as xylanases (Xyn), xyloglucan hydrolases (Xgh), and mannanases (Man) that hydrolyze carbohydrate sources found in hemicellulose. Of the 3236 entries in the *C. thermocellum* protein database, 84 proteins contain either cohesin and/or dockerin domains. This diversity of catalytic components and activities allows the cellulosome to solubilize carbohydrates from the full range of cellulose and hemicellulose constituents found in the heterogeneous cell wall, contributing to highly effective lignocellulose utilization and *C. thermocellum*'s role as a prime candidate for consolidated bioprocessing.[23]

To establish microbial cellulose utilization rates on a cell or cellulase-specific basis, it is necessary to independently quantitate the amount of cells, cellulose, and cellulase.[20,24,25] A previously developed method for mass quantitation of the *C. thermocellum* cellulosome complex employed ELISA,[26] in which a single structural protein, CipA, was quantitated, and the amount of CipA present in the sample was shown to be proportional to the amount of total cellulase under several conditions. In hopes of developing a less labor-intensive, more comprehensive, and more robust (e.g., in the presence of compounds arising from lignocellulose conversion) method, work was undertaken to evaluate MRM as an alternative approach to simultaneously quantitate multiple cellulosomal proteins in complex mixtures.

METHODS

Fermentations

C. thermocellum ATCC 27405 was grown anaerobically as previously described.[27] Samples were prepared using both Avicel PH105 (FMC Biopolymer, Philadelphia, PA) and D-(+)-cellobiose (Sigma Aldrich, St. Louis, MO) as substrates in separate fermentations. These samples were harvested during both the exponential and stationary growth phases as indicated. Samples were collected from 5 L fermentors in 40 mL aliquots that were centrifuged at 4500*g*, 4 °C for 1 h using a 5804 R centrifuge (Eppendorf, Hamburg, Germany) to separate the whole-cell (pellet) and extracellular (supernatant) fractions. Cellulosome from extracellular stationary growth phase samples was purified using affinity digestion as previously described[28] to obtain the purified cellulosome fraction.

Sample Preparation

All solutions were aqueous, except where noted. Extracellular fractions were concentrated to ~1 mL using 5 kDa spin filters (Sartorius Stedim Biotech GmbH, Goettingen, Germany) at 4500*g*, 4 °C. Whole cell fractions were lysed by resuspending in 5 mL of lysis buffer (4% w/v sodium dodecyl sulfate (SDS)[29] in 100 mM tris buffer, pH 8.0), heating at 95 °C for 5 min, and sonicating (Branson Ultrasonics, Danbury, CT) for 2 min at 20% amplitude (10 s on, 10 s off) in a water bath at ambient temperature. Samples were then heated again at 95 °C for 5 min. The concentrations of extracellular, whole cell, and purified cellulosome fractions were determined using a BCA assay. The remainder of each fraction was then snap-frozen with liquid nitrogen and stored at −80 °C.

After thawing on ice, 1 mL lysis buffer containing 10 mM dithiothreitol (DTT) was added to extracellular fractions and purified cellulosome samples.[29] This was followed by heating at 95 °C for 5 min, sonicating as described above, and heating again at 95 °C for 5 min. Whole cell fractions were adjusted to 10 mM DTT and heated at 95 °C for 10 min after thawing. Samples from all fractions were then adjusted to 20% trichloroacetic acid (TCA) and stored in a 4 °C refrigerator to precipitate proteins.[30] This was followed by centrifugation at 21000*g*, 4 °C for 15 min. Supernatant was carefully removed, and the resulting pellet was washed with 1 mL of cold acetone (−80 °C). This was followed by air-drying to produce a dry pellet of proteins.

Following drying, each pellet was brought up in 250 μL of 8 M urea in 100 mM tris buffer and was sonicated as described above in an ice water bath to resuspend the proteins. A BCA protein assay was performed at this point to determine the protein concentration prior to digestion. Following resuspension, 10 pmol each of 31 SIL peptides (Open Biosystems, Inc., Huntsville, AL) was added to each whole cell and extracellular fraction. Twenty picomoles of each of the 31 SIL peptides was added to purified cellulosome fractions. Samples were adjusted to 5 mM DTT, and proteins were denatured[31] under these reducing conditions at 37 °C for 1 h. Samples were then adjusted to 20 mM iodoacetamide and incubated in the dark for 15 min at room temperature to alkylate cysteine residues.[32]

Sequencing grade modified trypsin (Promega Co., Madison, WI) in 20 mM $CaCl_2$, 100 mM tris buffer was added to the samples at an enzyme/substrate ratio of 1:50 (20 μg of trypsin/1 mg of protein). Samples were incubated at room temperature

dx.doi.org/10.1021/pr400788e | J. Proteome Res. 2014, 13, 692−701

for 4 h. At the end of this predigestion, an additional 20 μg of trypsin was added to the sample. The sample was allowed to continue digesting overnight at room temperature (~20 h total digest time). Digestion was quenched by adjusting the sample to 200 mM NaCl, 0.1% formic acid. Undigested proteins and other cellular debris were removed by centrifugation at 21000*g*, 4 °C for 30 min with a 10 kDa spin filter (Sartorius Stedim Biotech GmbH, Goettingen, Germany). Concentration of filtrate was determined using a BCA assay, and the remainder was snap-frozen with liquid nitrogen and stored at −80 °C.

Chromatography

Nanospray emitters were prepared in house using a laser puller (model P-2000, Sutter Instrument Co., Novato, CA) with 100 μm i.d., 360 μm o.d. fused silica (Polymicro Technologies, Phoenix, AZ).[33] Emitters were bomb loaded with ~15 cm (2D global identification experiments) or ~10 cm (1D quantitation experiments) of C_{18} reversed-phase resin (Aqua 5 μm particle size, 125 Å pore size, Phenomenex, Torrance, CA) using a pressure cell (New Objective, Woburn, MA) with helium gas. In experiments utilizing two-dimensional chromatography, biphasic back columns were prepared from fused silica (150 μm i.d., 360 μm o.d., Polymicro Technologies, Phoenix, AZ) bomb loaded with ~5 cm of strong cation-exchange resin (Luna 5 μm particle size, 100 Å pore size, Phenomenex, Torrance, CA). ~5 cm of C_{18} reversed phase resin was then bomb loaded onto the back end of the strong cation-exchange resin as previously described.[34,35]

For two-dimensional separations utilized in global identification experiments, MudPITs (multidimensional protein identification technology)[35,36] were performed using an Ultimate 3000 HPLC (Dionex, Sunnyvale, CA) at a flow rate of ~300 nL/min, as previously described.[37−39] Solvent A consisted of 5:95% acetonitrile/water containing 0.1% formic acid, solvent B consisted of 70:30% acetonitrile/water containing 0.1% formic acid, and for MudPITs, solvent C consisted of 500 mM ammonium acetate in solvent A. For extracellular fractions, a 4-step, 7 h MudPIT was performed, but for more complex whole cell lysates, a 12-step, 23 h MudPIT was utilized. One-dimensional chromatography was utilized in QqQ experiments. Ten micrograms of digested peptides were bomb loaded directly onto the nanospray emitter. After loading of the sample, peptides were separated with a NanoLC-2D HPLC (Eksigent, Dublin, CA). The flow rate was 300 nL/min. The gradient used shifted from 98% solvent A to 80% solvent B over 5 min followed by the transition to 40% solvent A over the following 40 min.

Data Acquisition

A linear ion trap (LTQ XL, Thermo Scientific, Waltham, MA) was used for initial protein selection. Precursor scans were performed in the positive ion mode with full scans ranging from 400 to 1700 m/z followed by five data-dependent scans. For fragmentation, the isolation width was 3.0 Da, the normalized collision energy was 35.0%, the activation Q was 0.250, and the activation time was 30 ms.

A TSQ Quantum Discovery MAX QqQ mass spectrometer (Thermo Scientific, Waltham, MA) was used for quantitation experiments. The instrument was operated using a nanospray source (Proxeon Biosystems, Odense, Denmark) with a spray voltage of 1.75 kV. Data were acquired in the positive ion MRM mode, and three transitions were specified for each peptide. Peak widths for both Q1 and Q3 were set to 0.7 Da, and a scan width of 0.002 m/z was used. A scan time of 0.020 s, 1 microscan, and a chromatography filter value of 5 s were specified. Peaks were centroided. Skimmer offset was set to 15 V, and the transfer capillary temperature was optimized at 275 °C. Collision energies and tube lens voltages were optimized for each transition with breakdown curves performed on each isotopically labeled peptide. SIL peptides, tube lens voltages, collision energies, and fragments relevant to quantitation experiments are presented in Supplemental Table S1, Supporting Information.

For quantitation experiments, linear ion trap data used for preliminary protein and peptide selection were searched using DBDigger.[40] A precursor ion mass tolerance of 3 Da and a product ion mass tolerance of 0.5 Da were specified. Results were scored with MASPIC[41] and filtered with DTASelect.[42] A ΔCN value of 0.08 was used by DTASelect, and cross correlation (X_{corr}) values of 20, 25, and 40 were specified for +1, +2, and +3 ions, respectively. Data acquired on the QqQ for quantitation experiments were analyzed using Skyline[43] v. 0.7.

RESULTS AND DISCUSSION

Selection of Proteins for Quantitation

Proteomic profiling experiments were utilized to select a suitable set of diagnostic cellulosome proteins. Two extracellular stationary phase samples from Avicel- and cellobiose-grown cultures (Avi.Sta.E.1 and Cel.Sta.E.1 in Table 1) were analyzed with a 4-step, 7 h MudPIT, and whole cell lysates corresponding to the same substrates and growth phases (Avi.Sta.W.1 and Cel.Sta.W.1) were analyzed with a 12-step, 23 h MudPIT. Results were searched first with a database that consisted only of cellulosome proteins and then with the entire *C. thermocellum* protein database. Data were searched independently with the two databases to probe both the abundance of candidate proteins relative to only cellulosome proteins and relative to all detected proteins. Resulting NSAFs were used to assess the abundances of candidate proteins relative to only cellulosome proteins or to all detected proteins. The selected target proteins are listed in Table 2, along with NSAFs from all four profiling experiments. Although abundance was the primary factor for protein selection, protein function was also given consideration. For example, SbdA, OlpA, and Orf2p had relatively low NSAFs (<1% in extracellular fractions searched with the database containing only cellulosome proteins) but were selected for quantitation due to their critical role in cellulosome function as anchoring proteins that link CipA to the cell surface.[44]

The complexity of the samples used for quantitation is illustrated in Table 3, which lists the numbers of nonredundant peptide and protein identifications for the four samples based on database searches performed using the organism's entire protein database. On the basis of NSAF results in Table 2, the 14 proteins selected for quantitation constitute ~68% of the total cellulosomal protein abundance detected in extracellular samples, and ~76% of the cellulosomal proteins detected in whole cell lysate samples. Though results from these experiments were only rough estimates of relative protein abundance, they were consistent with previous semiquantitative global proteomic studies of *C. thermocellum*.[27,45] These 14 proteins correspond to 9−10% of total protein abundance in extracellular samples, and 1−2% in whole cell lysate samples. When the entire database was searched, a low amount of cellulosome was expected to be observed in stationary growth phase whole cell fractions because the cellulosome has been

Table 1. Samples Used for the Development of the Cellulosome Quantitation Method[a]

fermentation	culture age (h)	culture fraction	substrate	growth phase
Avi.P.1	36.25	purified extracellular cellulosome	Avicel	stationary
Avi.P.2	36.25	purified extracellular cellulosome	Avicel	stationary
Cel.P.1	32.25	purified extracellular cellulosome	Cellobiose	stationary
Cel.P.2	32.25	purified extracellular cellulosome	Cellobiose	stationary
Avi.Exp.E.1	15	extracellular	Avicel	exponential
Avi.Exp.E.2	15	extracellular	Avicel	exponential
Cel.Exp.E.1	15	extracellular	Cellobiose	exponential
Cel.Exp.E.2	15	extracellular	Cellobiose	exponential
Avi.Sta.E.1	36.25	extracellular	Avicel	stationary
Avi.Sta.E.2	36.25	extracellular	Avicel	stationary
Cel.Sta.E.1	32.25	extracellular	Cellobiose	stationary
Cel.Sta.E.2	32.25	extracellular	Cellobiose	stationary
Avi.Exp.W.1	15	whole cell	Avicel	exponential
Avi.Exp.W.2	15	whole cell	Avicel	exponential
Cel.Exp.W.1	15	whole cell	Cellobiose	exponential
Cel.Exp.W.2	15	whole cell	Cellobiose	exponential
Avi.Sta.W.1	36.25	whole cell	Avicel	stationary
Avi.Sta.W.2	36.25	whole cell	Avicel	stationary
Cel.Sta.W.1	32.25	whole cell	Cellobiose	stationary
Cel.Sta.W.2	32.25	whole cell	Cellobiose	stationary

[a]The fermentation column provides labels for each sample that are referred to throughout this manuscript. In this table, Avi = Avicel fermentation, Cel = cellobiose fermentation, Exp = exponential growth phase, Sta = stationary growth phase, P = purified extracellular cellulosome, E = extracellular fraction, W = whole cell fraction, and 1 or 2 refer to biological replicate number. Only a single time point representative of the exponential phase (mid-log, 15 h culture age) and stationary phase (late stationary phase, 36.25 h for Avicel fermentations, 32.35 h for cellobiose fermentations) were used in method development.

Table 3. Complexity of Samples Used for Quantification, as Illustrated by Total Nonredundant Peptide Identifications, Total Number of Assigned Spectra, and Total Number of Protein Identifications

sample	nonredundant peptides	assigned spectra	nonredundant proteins
Avi.Sta.W.1	15,484	58,652	1382
Cel.Sta.W.1	16,975	57,163	1540
Avi.Sta.E.1	6,336	14,883	593
Cel.Sta.E.1	5,693	17,259	597

shown to detach from the cell surface in the stationary growth phase.[46] Not only do these proteins represent the majority of the cellulosome in terms of NSAFs, but these proteins also represent a wide variety of cellulosomal functions, including the scaffoldin protein, anchoring proteins, and a range of cellulolytic enzymes of varying functions.

Choice of Labeled Peptides and MRM-MS Design

Given the significant cost of isotopically labeled peptides, their selection must be carefully planned once target proteins have been selected. In order to be considered for quantitation, potential peptides must have met all of the following conditions: the peptide sequence should be unique *both* to the protein of interest and unique within the protein sequence (i.e., the peptide occurs only once in the entire proteome), they should contain no missed cleavage sites, they should contain neither the protein N- or C-terminus due to the relatively high potential for protein degradation, and they should yield high signal quality across all substrates, growth phases, and cell fractions. Following protein selection, experiments were performed to establish an optimal set of peptides corresponding to the selected proteins that meet these criteria.

Eight of the samples outlined in Table 1 (Avi.Exp.E.1, Avi.Exp.W.1, Avi.Sta.E.1, Avi.Sta.W.1, Avi.Sta.W.1, Cel.Exp.W.1, Cel.Sta.E.1, Cel.Sta.W.1) were probed to ensure that peptides selected for quantitation were robust and detectable in all conditions used. Using peptides identified in profiling experiments described in the previous section as a guide to what peptides may be present in the samples, MRMs were performed

Table 2. NSAFs for Representative Samples Used to Select Proteins for Targeted Quantitation[a]

	Avi.Sta.W.1		Cel.Sta.W.1		Avi.Sta.E.1		Cel.Sta.E.1	
protein	cellulosome DB (%)	entire DB (%)	cellulosome DB (%)	entire DB (%)	cellulosome DB (%)	entire DB (%)	cellulosome DB (%)	entire DB
Cthe_3077 CipA	6.62	0.18	6.10	0.11	10.27	1.40	10.85	1.57
Cthe_0269 CelA	3.01	0.08	2.26	0.04	4.46	0.58	4.14	0.57
Cthe_2089 CelS	7.16	0.19	4.32	0.08	12.49	1.71	7.87	1.12
Cthe_0412 CelK	7.59	0.20	6.68	0.12	6.62	0.90	5.07	0.71
Cthe_0413 CbhA	2.58	0.07	2.08	0.04	2.40	0.33	1.66	0.23
Cthe_3078 OlpB	4.29	0.10	5.61	0.09	4.12	0.56	1.89	0.27
Cthe_0821	20.05	0.52	21.90	0.38	10.99	1.52	17.15	2.49
Cthe_0736	0.53	0.02	0.30	0.01	1.28	0.18	2.69	0.37
Cthe_1838 XynC	5.25	0.11	9.29	0.11	5.52	0.60	8.99	1.25
Cthe_0625 CelQ	2.65	0.05	1.89	0.02	4.26	0.61	3.99	0.58
Cthe_1307 SbdA	2.28	0.06	5.33	0.08	0.98	0.13	0.75	0.10
Cthe_0543 CelF	0.93	0.02	1.72	0.03	2.12	0.30	2.75	0.40
Cthe_3080 OlpA	9.94	0.25	5.92	0.11	0.50	0.05	0.13	0.01
Cthe_3079 Orf2p	2.93	0.08	2.56	0.04	1.12	0.15	0.28	0.04
total NSAF	75.82	1.91	75.96	1.25	67.12	9.00	68.20	9.70

[a]Data were searched with a database that consisted either only of cellulosome proteins or the organism's entire protein database.

dx.doi.org/10.1021/pr400788e | J. Proteome Res. 2014, 13, 692−701

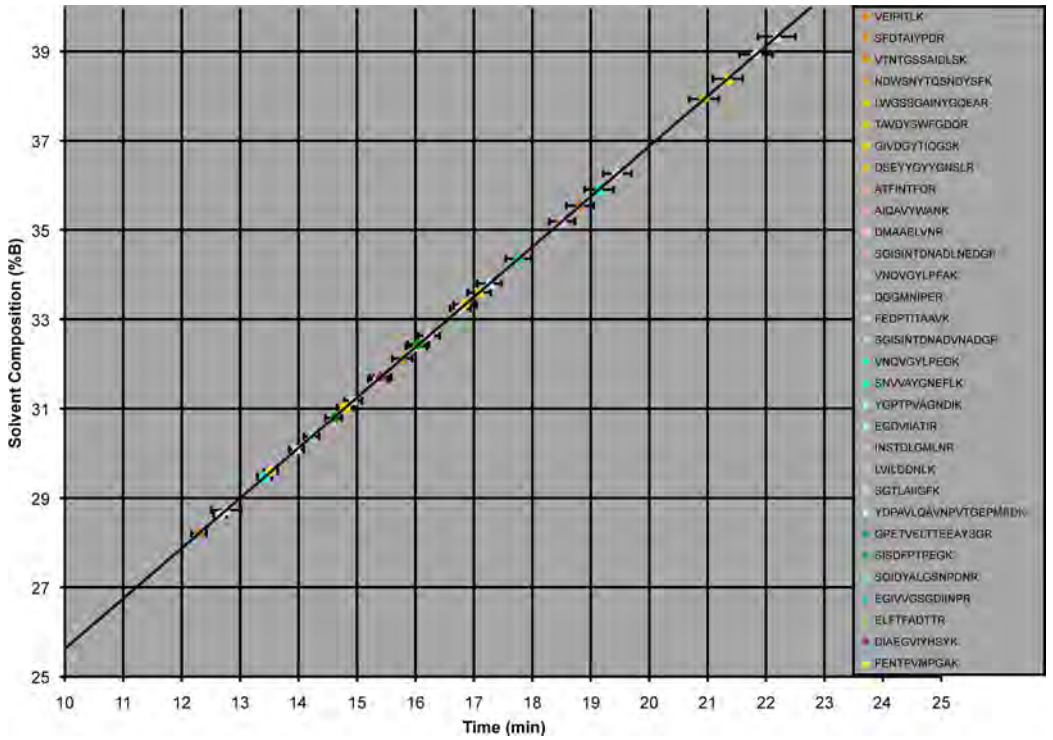


Figure 1. Retention times of all peptides averaged over all experiments. These results show that despite sample complexity ranging from hundreds of peptides in purified samples to tens of thousands of peptides in whole cell fractions retention times were remarkably reproducible.

on all eight samples to determine the best set of peptides corresponding to the particular proteins of interest. Doubly charged peptides were selected for all experiments, and up to nine different fragment ions were monitored for each precursor ion. All SRM chromatograms were manually inspected, and the three fragment ions for each precursor that exhibited the highest signal quality across different sample conditions were selected for potential quantitation.

A list of candidate peptides was compiled, and these were consolidated to a three-tiered list of proteins. Tier 1 consists of the four proteins (CipA, CelA, CelS, CelK) that yielded high NSAF values in Table 2 and were considered to be most biologically relevant to the cellulosome.[27,45] Four SIL peptides were obtained for each of these proteins due to their abundance and importance. Two SIL peptides were obtained for Tier 2 proteins, which included two anchoring proteins (OlpB and Cthe_0736), a xylanase (XynC), and two cellulases (CbhA and Cthe_0821). These proteins were of particular biological interest but were either not as abundant or did not provide as many peptides meeting all selection criteria. Only 1 SIL peptide was obtained for each protein in Tier 3, which included three low abundance anchoring proteins (SbdA, OlpA, and Orf2p) and two cellulases (CelQ and CelF) from glycoside hydrolase families that were already represented by either Tier 1 or 2 proteins. In total, 31 SIL peptides were purchased, corresponding to 14 cellulosome proteins in Table 2. Although the cellulosome quantitation assay could have been developed using only the proteins in Tier 1, peptides corresponding to proteins in Tiers 2 and 3 were included to expand the range of cellulosome representation and to probe stoichiometric ratios of proteins in the cellulosome, e.g., how many CelA cellulases are observed per CipA scaffoldin protein under different biological conditions.

Peptides were spiked into the sample after TCA precipitation but before denaturation, reduction, alkylation, and digestion. In these experiments, a bicinchoninic acid (BCA) protein assay was performed on an aliquot of the sample immediately after resolubilization. By spiking SIL peptides into the sample at this point, the mass of cellulosome protein (determined using MRMs) per mass of total protein in a particular fraction (purified cellulosome, extracellular fraction, whole-cell fraction, determined with a BCA assay following TCA precipitation) can be derived. Peptides spiked into the sample at this point experience all denaturing, reducing, alkylating, and digesting conditions that the actual samples experience, and sample preparation-induced sample losses should affect both the SIL peptides and the proteins. This should allow for accurate quantitation of proteins as they exist under biological conditions to the extent that these steps affect proteins and peptides similarly.

Reproducibility of peptide retention times was monitored across all samples outlined in Table 1, and averaged retention times and standard deviations for each peptide are presented in Figure 1. Standard error for peptide retention times was under 0.5 min for all 31 peptides and averaged 0.2 min. The average of precision across all peptides was 1.2% relative standard error (RSE). The average retention time was used to increase confidence in identification of a chromatographic peak assignment used for quantitation when peptide signal approached the limit of quantitation (LOQ at total ion current of 10^3 counts based on measurements of blanks).

Demonstration of the Multipoint Cellulosome Quantitation

Initial data analysis focused on the precision of ratios of unlabeled/labeled chromatographic peak areas (Supplemental Table S2, Supporting Information). Each of the three transitions monitored for each peptide provided an unlabeled/labeled ratio; Skyline reports the average of these three

dx.doi.org/10.1021/pr400788e | J. Proteome Res. 2014, 13, 692−701

Table 4. Cellulosome Quantitation Results for *Stationary Phase Purified Cellulosome Samples* Fermented on Avicel (Avi.P.1and Avi.P.2) and Cellobiose (Cel.P.1and Cel.P.2)

sample	ng of cellulosome protein/μg of total protein (technical replicates)	ng of cellulosome protein/μg of total protein technical replicate % RSE	ng of cellulosome protein/μg of total protein (biological replicates)	ng cellulosome protein/μg total protein biological replicate % RSE
Avi.P.1	306.04 ± 1.28	0.42	328.87 ± 13.19	4.91
Avi.P.2	351.70 ± 0.69	0.20		
Cel.P.1	349.83 ± 3.61	1.03	345.21 ± 3.07	0.95
Cel.P.2	340.58 ± 0.96	0.28		

Table 5. Cellulosome Quantitation Results for *Extracellular Fractions* Fermented on Avicel and Cellobiose

sample	ng of cellulosome protein/μg of total protein (technical replicates)	ng of cellulosome protein/μg of total protein technical replicate % RSE	ng of cellulosome protein/μg of total protein (biological replicates)	ng of cellulosome protein/μg of total protein biological replicate % RSE
Avi.Exp.E.1	63.53 ± 1.47	2.32	61.90 ± 1.13	1.87
Avi.Exp.E.2	60.26 ± 0.28	0.47		
Cel.Exp.E.1	92.16 ± 0.68	0.74	94.17 ± 1.21	1.51
Cel.Exp.E.2	96.18 ± 0.42	0.44		
Avi.Sta.E.1	118.24 ± 2.97	2.51	116.66 ± 1.55	0.96
Avi.Sta.E.2	115.08 ± 0.77	0.67		
Cel.Sta.E.1	97.98 ± 1.34	1.37	106.45 ± 4.94	5.63
Cel.Sta.E.2	114.91 ± 0.97	0.84		

Table 6. Cellulosome Quantitation Results for *Whole Cell Fractions* Fermented on Avicel and Cellobiose

sample	ng of cellulosome protein/μg of total protein (technical replicates)	ng of cellulosome protein/μg of total protein technical replicate % RSE	ng of cellulosome protein/μg of total protein (biological replicates)	ng of cellulosome protein/μg of total protein biological replicate % RSE
Avi.Exp.W.1	40.45 ± 0.28	0.70%	42.19 ± 1.05	2.91
Avi.Exp.W.2	43.93 ± 0.69	1.58%		
Cel.Exp.W.1	19.83 ± 0.10	0.49%	18.23 ± 0.95	6.21
Cel.Exp.W.2	16.63 ± 0.58	3.46%		
Avi.Sta.W.1	33.03 ± 0.69	2.07%	27.98 ± 2.93	12.75
Avi.Sta.W.2	22.94 ± 0.03	0.12%		
Cel.Sta.W.1	14.39 ± 0.19	1.29%	12.59 ± 1.04	10.11
Cel.Sta.W.2	10.79 ± 0.06	0.55%		

ratios for each peptide. Two technical replicates were measured for each sample. Precisions for unlabeled/labeled peptide ratios in a given sample type were determined as RSE for these two technical replicates. These values (Table S2, Supporting Information) ranged from 0.01% RSE for peptide LVILDDNLK (Cthe_0821) in sample Avi.Sta.W.1to 98% for peptide DIAEGVIYHSYK (OlpA) in sample Avi.Exp.E.1. Peak heights, background, and peak area ratios (unlabeled/labeled) were analyzed for the four peptides corresponding to a representative protein (CelA) in a single experiment on a whole cell sample (Avi.Exp.W.1). Peak heights for transitions were well above the background, typically at least 1−2 orders of magnitude above the background, and peak height values for transitions corresponding to the same peptide were all within an order of magnitude of each other. Signal-to-background (S/B) ratios ranged from 10 to 6000 with an average of ∼1000 for the 24 transitions (3 transitions per peptide, 4 peptides per protein, unlabeled and labeled versions of each peptide). When peak area ratios for the three transitions corresponding to each peptide were analyzed, % RSE values ranging from 6.1% (GIVDGYTIQGSK) to 21.8% (NDWSNYTQSNDYSFK) were observed. This suggests that the variance in measurements occurs at the level of the transitions. The peak heights and S/B ratios eliminate concentration effects and the possibility that the intensity of a single transition is swamping the intensities of other transitions as possible causes of the variation.

Though actual measurements were made at the peptide level, it is the protein level information that is of interest. Ratios of unlabeled/labeled peptide were determined from chromatographic peak areas for each transition for a given peptide, and these values were then averaged to determine a single unlabeled/labeled ratio for each peptide. These ratios were converted to picomoles of unlabeled peptide by multiplying the average ratio by the amount of labeled peptide spiked into the sample prior to digestion (10 pmol for whole cell and extracellular fractions, 20 pmol for purified cellulosome samples due to the increased amounts of cellulosome proteins expected in these samples). Protein mass was estimated from pmol of unlabeled peptide assuming a 1:1 peptide/protein mole ratio. For Tier 1 and 2 proteins, resulting microgram values derived from each peptide corresponding to a particular protein were averaged to deduce a microgram value for each protein. The total microgram amount of cellulosome protein for each sample was determined by summing the microgram values for each protein (microgram values for each protein and total microgram amount of cellulosome protein for each sample are shown in Supplemental Tables S3−S5). The microgram of total protein prior to digestion was calculated from averaged concentration values from a BCA assay performed at three different concentrations with three replicate measurements per concentration. Cellulosome mass concentrations are reported as nanograms of cellulosome proteins per microgram of total protein. The average cellulosome mass concentration is

dx.doi.org/10.1021/pr400788e | J. Proteome Res. 2014, 13, 692−701

reported in Tables 4−6 as well as precision for both technical replicates and biological replicates.

Despite the variability for individual peptides in technical replicates, precision for the cellulosome mass concentration was remarkably high. Precision for technical replicates ranged from 0.1% to 3.5% RSE, and the average precision for all technical replicates across all fractions and growth states was 1.1% RSE. These values indicated a high level of precision for the cellulosome mass concentration. Particularly surprising was the level of precision in complex whole cell fractions, which averaged 1.2% RSE across all eight whole cell lysate samples in Table 6. High precision of technical replicate measurements of cellulosome mass concentration suggests that even though measurements at the peptide level often lacked high precision (Supplemental Table S6, Supporting Information), variation at the cellulosome complex level is minimal and this method can be used for protein quantitation with high confidence.

As expected, precision for biological replicates was generally lower than technical replicates, ranging from 1.0% RSE to 12.8% RSE. This higher variation across biological replicates is most likely due to slightly different conditions during the course of the fermentations. The average precision for biological replicates across all cell fractions was 4.8% RSE. Higher % RSE values were observed in whole cell lysate samples compared to extracellular fractions (8.0% RSE averaged for all whole cell fraction biological replicates compared to 2.5% RSE averaged for all extracellular fraction biological replicates). However, variability for biological replicates was actually lower for whole cell lysate fractions compared to extracellular fractions (2.2 average biological replicate standard error for extracellular fractions versus 1.3 average biological replicate standard error for whole cell lysate fractions). The difference in precision between extracellular fractions and whole cell fractions results from cellulosome mass concentrations in extracellular fractions that were on average approximately 5 times larger than those from corresponding whole cell lysate samples. Variation in cellulosome mass concentration reflects variability both in the amount of substrate used for fermentation, fermentation conditions, recovery from TCA precipitation, trypsin digestion, and most importantly, natural biological variation, so overall precision of <5% RSE is reasonable for biological replicates.

Though the *entire* cellulosome complex was not quantitated in total in these experiments, for simplicity, "cellulosome mass concentration" in the remaining discussion will refer to the summed mass concentration of the 14 cellulosome components selected for quantitation. Cellulosome mass concentration results were generally consistent with expectations. In Table 5, higher cellulosome mass concentrations were observed in stationary phase extracellular fractions compared to exponential phase samples in Avicel fermentations. Conversely, in Table 6, (slightly) higher cellulosome mass concentrations were observed in the exponential phase whole cell lysates compared to stationary phase whole cell lysates in Avicel fermentations. These results were expected because as the bacterial growth phase in *C. thermocellum* proceeds from exponential phase to stationary phase, cellulosome complexes have been shown to detach from the cell surface[46] and so should be enriched in extracellular fractions and depleted in whole cell fractions. The overall enrichment in extracellular fractions compared to whole cell fractions is also expected because the cellulosome is an extracellular protein complex. Purified cellulosome fractions exhibited significantly higher cellulosome mass concentrations compared to both extracellular fractions and whole cell lysates, which was also expected considering the cellulosome had been affinity digested. Structural proteins such as Orf2p, OlpA, and SbdA that anchor into the cell surface were observed in extremely small amounts (if at all) in purified cellulosome samples. This was also expected because the affinity digestion used to purify the cellulosome only isolates those proteins that bind to cellulose, and the anchoring proteins do not bind the cellulose.

Examination of Table 4 shows that even though the *absolute* amount of cellulosome (μg) observed was significantly higher for the cellobiose fermentation compared to the Avicel fermentation, the *relative* amount of cellulosome (ng of cellulosome/μg of total protein) was relatively similar across all purified station phase samples regardless of cellulolytic substrate. It is interesting to note that while the average absolute amount of cellulosome detected for the cellobiose fermentations is 2.2 times larger compared to the average amount of cellulosome detected for the Avicel fermentations (557.5 $\pm$ 8.1 μg vs 249.0 $\pm$ 26.7 μg), the absolute amount of CipA in the cellobiose fermentations is 3.0 times larger compared to the absolute amount of CipA in the Avicel fermentations (163.4 $\pm$ 8.4 μg vs 49.5 $\pm$ 8.5 μg). This suggests a higher percentage of the total cellulosome results from CipA in cellobiose fermentations compared to Avicel fermentations. Analysis of Tables S3−S5 also reveals that the amount of CelS is greater than the amount of CelA across all fermentations and growth phases. The CelS/CelA ratio is relatively consistent across all fractions (1.8 $\pm$ 0.7 for purified fractions, 2.1 $\pm$ 0.7 for extracellular fractions, and 2.9 $\pm$ 0.6 for whole cell fractions). This suggests that regardless of which substrate is present, *C. thermocellum* produces more CelS than CelA. This is likely a result of the different cellulolytic functions of the two enzymes, as CelS is an exoglucanase while CelA is an endoglucanase. These observations combined with those involving absolute amounts of CipA as a function of substrate, have guided the formulation of a more extensive study to probe how the components of the cellulosome vary with varying substrate type and growth condition.

As mentioned above, the ELISA-based method for cellulosome quantitation previously developed[26] determined the amount of cellulosome present in a particular sample by analyzing the amount of CipA bound to an antibody. The assumption in this method is that CipA, as the scaffoldin protein that binds to cellulose, cellulolytic enzymes, and anchoring proteins, is representative of the cellulosome in general. This was tested by plotting the mass concentration of CipA (ng CipA/μg total protein) as determined by the method developed herein as a function of the summed mass concentrations of the 14 cellulosome components selected for quantitation (referred to as "summed cellulosome mass concentration") in Figure 2. For extracellular fractions and whole cell fractions, the observation of an R^2 value of 0.94 indicates strong correlation between CipA mass concentration and summed cellulosome mass concentration, which suggests the assumption that CipA is representative of the cellulosome as a whole generally holds true in whole cell and extracellular fractions. However, the correlation between CipA and *summed* cellulosome protein is much weaker for the purified cellulosome fractions, indicated by an R^2 value of 0.48. This reveals that for the much higher concentrations of the cellulosome in the purified samples, the use of CipA as the

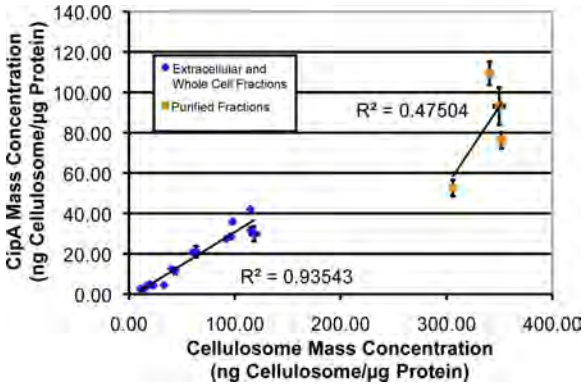


Figure 2. Comparison of CipA mass concentration as a function of total cellulosome mass concentration. These results indicate that there is a strong correlation between the CipA mass concentration and the summed cellulosome mass concentration for the extracellular and whole cell fractions, suggesting that CipA is a strongly representative protein for the cellulosome as a whole in these fractions. However, such a correlation between CipA mass concentration and summed cellulosome mass concentration is not observed in the purified fraction.

lone representative of this protein complex is nonlinear and may not accurately represent the cellulosome concentration.

Upon visual inspection of this graph, the three data points corresponding to purified samples with the highest cellulosome mass concentration appeared to have a strong negative correlation. When the purified cellulosome data point with the lowest cellulosome mass concentration is treated as an outlier, the R^2 value of the correlation is 0.86. However, the trend line (not shown) for these three data points has a negative slope, indicating that as summed cellulosome mass concentration increases, the CipA mass concentration decreases, in contrast to the expectation that CipA abundance is always directly correlated with total cellulome abundance. Regardless of whether the lowest purified data point was treated as an outlier, the correlation (or lack thereof) between CipA mass concentration and summed cellulosome concentration in purified samples indicated that using CipA to represent total cellulosome in purified cellulosome fractions can be problematic and should be approached with caution. While careful calibration and utilization of the ELISA-based method has been successfully used as the basis for many significant findings,[47−49] the multipoint measurement described herein allows for high throughput quantitation of both structural and enzymatic components of the *C. thermocellum* cellulosome with high precision and ease of use.

CONCLUSIONS

The MRM-MS method demonstrated in this report can be utilized to quantitate cellulosome mass concentration in samples ranging in complexity from purified cellulosome samples consisting of fewer than 100 proteins to whole cell fractions consisting of over 1500 proteins. However, the cellulosome mass concentration is determined by cellulosome peptides after digestion in mixtures that are perhaps an order of magnitude more complex than the protein samples, assuming each protein yields 10 peptides after digestion. The 14 proteins selected for this method correspond to the majority of cellulosome as determined by NSAFs. Precision for this method of quantitation is under 2% RSE for all technical replicates and on average below 5% RSE for biological replicates. As such, the method demonstrated herein greatly supplements the previously described[26] ELISA method for high throughput, multipoint quantitation of the most representative components of the cellulosome protein machine in complex samples with minimal purification.

ASSOCIATED CONTENT

Supporting Information

This material includes Table S1 − Peptides and parameters used in quantitation experiments; Table S2 − Ratios of chromatographic peak areas for unlabeled/labeled peptides for extracellular and whole cell fraction samples; Table S3 − Cellulosome quantitation results for stationary phase purified cellulosome samples; Table S4 − Cellulosome quantitation results for extracellular fractions; Table S5 − Cellulosome quantitation results for whole cell fractions; and Table S6 − Overall precision of peptide ratios determined from technical replicates across different fractions. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: hettichrl@ornl.gov. Phone: 865-574-4968. Fax: 865-576-8559.

Present Addresses

[#](A.D.) Univ. of California, San Francisco, CA.
[‖](L.S.) The Univ. of West Indies, St Augustine, Trinidad and Tobago.
[⊥](B.R.) Dow AgroSciences, Indianapolis, IN.
[○](J.I.) Research Triangle Institute, Raleigh-Durham, NC.
[●](K.C.) National Science Foundation, Arlington, VA.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Special thanks to Richard Giannone, Adriane Lochner, Paul Abraham, and Rachel Adams for technical advice. Thanks to Joel Kreps for his early involvement in this research. This research was sponsored by the U.S. DOE-BER, Bioenergy Research Program. Oak Ridge National Laboratory is managed by UT-Battelle, LLC, for the U.S. Department of Energy. Participation by K.D.C. while at the National Science Foundation was supported through the NSF Independent Research and Development program.

REFERENCES

(1) Bradford, M. M. Rapid and Sensitive Method for Quantitation of Microgram Quantities of Protein Utilizing Principle of Protein-Dye Binding. *Anal. Biochem.* **1976**, *72* (1−2), 248−254.

(2) Smith, P. K.; Krohn, R. I.; Hermanson, G. T.; Mallia, A. K.; Gartner, F. H.; Provenzano, M. D.; Fujimoto, E. K.; Goeke, N. M.; Olson, B. J.; Klenk, D. C. Measurement of Protein Using Bicinchoninic Acid. *Anal. Biochem.* **1985**, *150* (1), 76−85.

(3) Wiechelman, K. J.; Braun, R. D.; Fitzpatrick, J. D. Investigation of the Bicinchoninic Acid Protein Assay - Identification of the Groups Responsible for Color Formation. *Anal. Biochem.* **1988**, *175* (1), 231−237.

(4) Lowry, O. H.; Rosebrough, N. J.; Farr, A. L.; Randall, R. J. Protein measurement with the Folin phenol reagent. *J. Biol. Chem.* **1951**, *193* (1), 265−275.

(5) Engvall, E.; Perlmann, P. Enzyme-linked immunosorbent assay (ELISA). Quantitative assay of immunoglobulin G. *Immunochemistry* **1971**, *8* (9), 871−874.
(6) Bandi, Z. L.; Schoen, I.; DeLara, M. Enzyme-linked immunosorbent urine pregnancy tests. Clinical specificity studies. *Am. J. Clin. Pathol.* **1987**, *87* (2), 236−242.
(7) Maskill, W. J.; Crofts, N.; Waldman, E.; Healey, D. S.; Howard, T. S.; Silvester, C.; Gust, I. D. An evaluation of competitive and second generation ELISA screening tests for antibody to HIV. *J. Virol. Methods* **1988**, *22* (1), 61−73.
(8) Voller, A.; Huldt, G.; Thors, C.; Engvall, E. New serological test for malaria antibodies. *Br. Med. J.* **1975**, *1* (5959), 659−661.
(9) Paoletti, A. C.; Parmely, T. J.; Tomomori-Sato, C.; Sato, S.; Zhu, D. X.; Conaway, R. C.; Conaway, J. W.; Florens, L.; Washburn, M. P. Quantitative proteomic analysis of distinct mammalian Mediator complexes using normalized spectral abundance factors. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (50), 18928−18933.
(10) Rappsilber, J.; Ryder, U.; Lamond, A. I.; Mann, M. Large-scale proteomic analysis of the human spliceosome. *Genome Res.* **2002**, *12* (8), 1231−1245.
(11) Ishihama, Y.; Oda, Y.; Tabata, T.; Sato, T.; Nagasu, T.; Rappsilber, J.; Mann, M. Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol. Cell. Proteomics* **2005**, *4* (9), 1265−1272.
(12) Gygi, S. P.; Rist, B.; Gerber, S. A.; Turecek, F.; Gelb, M. H.; Aebersold, R. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* **1999**, *17* (10), 994−999.
(13) Ong, S. E.; Blagoev, B.; Kratchmarova, I.; Kristensen, D. B.; Steen, H.; Pandey, A.; Mann, M. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* **2002**, *1* (5), 376−386.
(14) Mann, M. Functional and quantitative proteomics using SILAC. *Nat. Rev. Mol. Cell Biol.* **2006**, *7* (12), 952−958.
(15) Ross, P. L.; Huang, Y. N.; Marchese, J. N.; Williamson, B.; Parker, K.; Hattan, S.; Khainovski, N.; Pillai, S.; Dey, S.; Daniels, S.; Purkayastha, S.; Juhasz, P.; Martin, S.; Bartlet-Jones, M.; He, F.; Jacobson, A.; Pappin, D. J. Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics* **2004**, *3* (12), 1154−1169.
(16) Thompson, A.; Schafer, J.; Kuhn, K.; Kienle, S.; Schwarz, J.; Schmidt, G.; Neumann, T.; Johnstone, R.; Mohammed, A. K.; Hamon, C. Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal. Chem.* **2003**, *75* (8), 1895−1904.
(17) Razumovskaya, J.; Olman, V.; Xu, D.; Uberbacher, E. C.; VerBerkmoes, N. C.; Hettich, R. L.; Xu, Y. A computational method for assessing peptide- identification reliability in tandem mass spectrometry analysis with SEQUEST. *Proteomics* **2004**, *4* (4), 961−969.
(18) Gerber, S. A.; Rush, J.; Stemman, O.; Kirschner, M. W.; Gygi, S. P. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100* (12), 6940−6945.
(19) Lange, V.; Picotti, P.; Domon, B.; Aebersold, R. Selected reaction monitoring for quantitative proteomics: a tutorial. *Mol. Syst. Biol.* **2008**, *4*, 222.
(20) Lynd, L. R.; Weimer, P. J.; van Zyl, W. H.; Pretorius, I. S. Microbial cellulose utilization: fundamentals and biotechnology. *Microbiol. Mol. Biol. Rev.* **2002**, *66* (3), 506−+.
(21) Kruus, K.; Lua, A. C.; Demain, A. L.; Wu, J. H. D. The anchorage function of cipA (Cell), a scaffolding protein of the *Clostridium thermocellum* Cellulosome. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92* (20), 9254−9258.
(22) Zhang, Y. H.; Lynd, L. R. Cellulose utilization by *Clostridium thermocellum*: bioenergetics and hydrolysis product assimilation. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (20), 7321−7325.
(23) Demain, A. L.; Newcomb, M.; Wu, J. H. D. Cellulase, clostridia, and ethanol. *Microbiol. Mol. Biol. Rev.* **2005**, *69* (1), 124−154.
(24) Holwerda, E. K.; Ellis, L. D.; Lynd, L. R. Development and evaluation of methods to infer biosynthesis and substrate consumption in cultures of cellulolytic microorganisms. *Biotechnol. Bioeng.* **2013**, *110*, 2380−2388.
(25) Holwerda, E. K.; Lynd, L. R. Testing alternative kinetic models for utilization of crystalline cellulose (Avicel) by batch cultures of Clostridium thermocellum. *Biotechnol. Bioeng.* **2013**, *110*, 2389−2394.
(26) Zhang, Y. H.; Lynd, L. R. Quantification of cell and cellulase mass concentrations during anaerobic cellulose fermentation: Development of an enzyme-linked immunosorbent assay-based method with application to *Clostridium thermocellum* batch cultures. *Anal. Chem.* **2003**, *75* (2), 219−227.
(27) Raman, B.; Pan, C.; Hurst, G. B.; Rodriguez, M., Jr.; McKeown, C. K.; Lankford, P. K.; Samatova, N. F.; Mielenz, J. R. Impact of pretreated Switchgrass and biomass carbohydrates on *Clostridium thermocellum* ATCC 27405 cellulosome composition: a quantitative proteomic analysis. *PLoS One* **2009**, *4* (4), e5271.
(28) Morgenstern, E. M.; Bayer, E. A.; Lamed, R. Affinity digestion for the near-total recovery of purified cellulosome from *Clostridium thermocellum*. *Enzyme Microb. Technol.* **1992**, *14* (4), 289−292.
(29) Wisniewski, J. R.; Zougman, A.; Nagaraj, N.; Mann, M. Universal sample preparation method for proteome analysis. *Nat. Methods* **2009**, *6* (5), 359−U60.
(30) Jiang, L.; He, L.; Fountoulakis, M. Comparison of protein precipitation methods for sample preparation prior to proteomic analysis. *J. Chromatogr., A* **2004**, *1023* (2), 317−320.
(31) Bennion, B. J.; Daggett, V. The molecular basis for the chemical denaturation of proteins by urea. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100* (9), 5142−5147.
(32) Sechi, S.; Chait, B. T. Modification of cysteine residues by alkylation. A tool in peptide mapping and protein identification. *Anal. Chem.* **1998**, *70* (24), 5150−5158.
(33) Delahunty, C.; Yates, J. R., Making Nanocolumns and Tips. In *Encyclopedia of Genetics, Genomics, Proteomics and Bioinformatics*, Dunn, M. J., Ed.; Wiley: Hoboken, NJ, 2005; Vol. *5*, pp 1882−1885.
(34) McDonald, W. H.; Ohi, R.; Miyamoto, D. T.; Mitchison, T. J.; Yates, J. R. Comparison of three directly coupled HPLC MS/MS strategies for identification of proteins from complex mixtures: single-dimension LC-MS/MS, 2-phase MudPIT, and 3-phase MudPIT. *Int. J. Mass Spectrom.* **2002**, *219* (1), 245−251.
(35) Washburn, M. P.; Wolters, D.; Yates, J. R. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* **2001**, *19* (3), 242−247.
(36) Wolters, D. A.; Washburn, M. P.; Yates, J. R. An automated multidimensional protein identification technology for shotgun proteomics. *Anal. Chem.* **2001**, *73* (23), 5683−5690.
(37) Baker, B. J.; Comolli, L. R.; Dick, G. J.; Hauser, L. J.; Hyatt, D.; Dill, B. D.; Land, M. L.; Verberkmoes, N. C.; Hettich, R. L.; Banfield, J. F. Enigmatic, ultrasmall, uncultivated Archaea. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (19), 8806−8811.
(38) Giannone, R. J.; McDonald, W. H.; Hurst, G. B.; Huang, Y.; Wu, J.; Liu, Y.; Wang, Y. Dual-tagging system for the affinity purification of mammalian protein complexes. *Biotechniques* **2007**, *43* (3), 296−302.
(39) VerBerkmoes, N. C.; Shah, M. B.; Lankford, P. K.; Pelletier, D. A.; Strader, M. B.; Tabb, D. L.; McDonald, W. H.; Barton, J. W.; Hurst, G. B.; Hauser, L.; Davison, B. H.; Beatty, J. T.; Harwood, C. S.; Tabita, F. R.; Hettich, R. L.; Larimer, F. W. Determination and comparison of the baseline proteomes of the versatile microbe Rhodopseudomonas palustris under its major metabolic states. *J. Proteome Res.* **2006**, *5* (2), 287−298.
(40) Tabb, D. L.; Narasimhan, C.; Strader, M. B.; Hettich, R. L. DBDigger: reorganized proteomic database identification that improves flexibility and speed. *Anal. Chem.* **2005**, *77* (8), 2464−2474.
(41) Narasimhan, C.; Tabb, D. L.; Verberkmoes, N. C.; Thompson, M. R.; Hettich, R. L.; Uberbacher, E. C. MASPIC: intensity-based tandem mass spectrometry scoring scheme that improves peptide identification at high confidence. *Anal. Chem.* **2005**, *77* (23), 7581−7593.

(42) Tabb, D. L.; McDonald, W. H.; Yates, J. R., 3rd. DTASelect and Contrast: tools for assembling and comparing protein identifications from shotgun proteomics. *J. Proteome Res.* **2002**, *1* (1), 21−26.
(43) MacLean, B.; Tomazela, D. M.; Shulman, N.; Chambers, M.; Finney, G. L.; Frewen, B.; Kern, R.; Tabb, D. L.; Liebler, D. C.; MacCoss, M. J. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* **2010**, *26* (7), 966−968.
(44) Fontes, C. M.; Gilbert, H. J. Cellulosomes: highly efficient nanomachines designed to deconstruct plant cell wall complex carbohydrates. *Annu. Rev. Biochem.* **2010**, *79*, 655−681.
(45) Gold, N. D.; Martin, V. J. Global view of the *Clostridium thermocellum* cellulosome revealed by quantitative proteomic analysis. *J. Bacteriol.* **2007**, *189* (19), 6787−6795.
(46) Bayer, E. A.; Kenig, R.; Lamed, R. Adherence of *Clostridium thermocellum* to cellulose. *J. Bacteriol.* **1983**, *156* (2), 818−27.
(47) Lu, Y.; Zhang, Y. H.; Lynd, L. R. Enzyme-microbe synergy during cellulose hydrolysis by *Clostridium thermocellum*. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (44), 16165−16169.
(48) Zhang, Y. H.; Lynd, L. R. Kinetics and relative importance of phosphorolytic and hydrolytic cleavage of cellodextrins and cellobiose in cell extracts of Clostridium thermocellum. *Appl. Environ. Microbiol.* **2004**, *70* (3), 1563−1569.
(49) Zhang, Y. H.; Lynd, L. R. Regulation of cellulase synthesis in batch and continuous cultures of Clostridium thermocellum. *J. Bacteriol.* **2005**, *187* (1), 99−106.