Electronic Supplementary Material (ESI) for ChemComm.

HaloTag mediated artificial cellulosome assembly on rolling circle amplification DNA template for efficient cellulose hydrolysis

Qing Sun and Wilfred Chen

Department of Chemical and Biomolecular Engineering, University of Delaware, Newark, DE 19716 United States.

Experimental procedures

Genetic manipulation and expression of all the proteins

Escherichia coli strain NEB 5-alpha (NEB #C2987I) was used as the host for genetic manipulations. *E. coli* strains BL21 (DE3) (Novagen) and BLR (DE3) (Novagen) were used for protein expressions.

Construction of expression vectors for sensor components

Escherichia coli strain NEB 5-alpha (*fhuA2 Δ(argF-lacZ)U169 phoA glnV44 Φ80Δ (lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17*) was used as the host for genetic manipulations.

ELP-HaloTag-His6 was constructed by PCR amplifying HaloTag with HaloTag-Forward, HaloTag-Reverse. PCR amplified HaloTag was digested with BamHI and XhoI and ligated into BamHI and XhoI digested pET24(a)-ELP[KV8F-40][1] to form pET24(a)-ELP-HaloTag.

CBM was then PCR amplified with CBM-Forward and CBM-Reverse, digested with Nde1 and ligated into Nde1 digested pET24(a)-ELP-HaloTag to form CBM-ELP-HaloTag. CelA, CelE and BglA were then individually PCR amplified with CelA-Forward, CelA-Reverse, CelE-Forward, CelE-Reverse, BglA-Forward, BglA-Reverse, digested with SacII and SpeI and inserted into SacII and SpeI digest pET24(a)-CBM-ELP-HaloTag to form CelA-ELP-HaloTag, CelE-ELP-HaloTag and BglA-ELP-HaloTag.

Protein expression

All the proteins were expressed in host *E. coli* BLR [*F- ompT hsdSB (r-B m-B) gal dcm(DE3) Δ (srl-recA)306::Tn10(TetR)*; Novagen, Madison,WI] in TB medium. All the four strains were inoculated in TB medium supplemented with 50 µg/mL kanamycin at 37 ^{0}C until OD reached 1. Then the culture was moved to 25 ^{0}C shaker for overnight leaky expression.

After protein expression, cells were harvested by centrifugation at 4200 g, resuspended in PBS and lysed by ultrasonic disruption using a sonicator. The cell debris was removed by centrifugation at 16.1 k rcf for 10 min at 4 ^{0}C. ELP purification of all the components were then conducted to purify the proteins from cell lysis.

Table S1. Primers used for recombinant gene construction

HaloTag-Forward	5’- CCG CAT GGA TCC GGC GGC AGC AGC CCG AGC ACC CCG CCG ACC CCG AGC CCG AGC ACC CCG CCG GGC GGC AGC GCA GAA ATC GGT ACT GGC TTT CCA T -3’
HaloTag-Reverse	5’- TTG GCC CAA GCT TGC CGG AAA TCT CGA GCG TCG -3’
CBM-Forward	5'- GGG AAT TCC ATA TGC CGC GGG ATC CGA CCA AGG GAG CAA CAC -3'
CBM-Reverse	5'- GGG AAT TCC ATA TGG CTG CCG CCC GGC GGG GTG CTC GGG CTC GGG GTC GGC GGG GTG CTC GGG CTG CTG CCG CCA CTA GTG TCG ACT ACT ACA CTG CCA CCG G -3'
CelE-Forward	5'- AAA TCC CCG CGG ATG CTT GTT GGG GCA GGA GAT TTG -3'
CelE-Reverse	5'- TTC TAG ACT AGT TGC AGC GAA GTC CAA TGC ATC C -3'
BglA-Forward	5'- AAA TCC CCG CGG TCA AAG ATA ACT TTC CCA AAA GAT TTC ATA TGG GGT -3'
BglA-Reverse	5'- TTC TAG ACT AGT AAA ACC GTT GTT TTT GAT TAC TTC TTT GTA CCA GTA G -3'
CelA-Forward	5'- AAA TCC CCG CGG GCA GGT GTG CCT TTT AAC ACA AAA

	TAC CC -3'
CelA-Reverse	5'- TTC TAG ACT AGT GTT TCC TGT TAT GTA CAA CAA AGT GAG CAG TCT C -3'

Protein purification and DNA conjugation

All fusion proteins were purified by taking advantage of the ELP inverse phase transition functionality.[2] After cell lysis, target proteins were precipitated by the addition of 1M Na_2SO_4, incubated at 37°C for 10 min, and centrifuged at 160,000g for 10 min at 37°C. After discarding the supernatant, the protein pellets were resolubilized in 4°C cold buffer. This thermal cycle was repeated one more time for better purity.

DNA linkers modified with a 5' amine group were ordered from Integrated DNA Technologies (Coralville, IA). To modify the linkers for HaloTag conjugation, the CH ligand (HaloTag Succinimidyl Ester (O4) Ligand) was mixed with the DNA linkers at a molar ratio of 30:1 and incubated at room temperature for 4 h. The mixture was then purified using a 3,000 Da ultrafiltration column (Vivaspin 500, Sartorius Stedim Biotech), to remove the excess CH ligand. The purified DNA linkers was mixed with the purified Halo fusions, using a 3x molar excess DNA linkers and incubated overnight at 4 °C. Unconjugated DNA was removed by utilizing the ELP tag by precipitating the protein-DNA linker conjugates as described above. All samples were analyzed by a 10% SDS-PAGE and stained with coomassie blue for analysis of labeling efficiency.

DNA linker sequence:
a AAAAATTCCTGACTGACTCTC
b AAAAATTCCCCTCCCTCCTTT
c AAAAATGTAGACGCTCGCTGT
d AAAAAGTTGGCTGGCTGGTGT
DNA linker blockers:
a* GAGAGTCAGTCAGGAATTTTT
b* AAAGGAGGGAGGGGAATTTTT
c* ACAGCGAGCGTCTACATTTTT
d* ACACCAGCCAGCCAACTTTTT
Template for four components:
GAGAGTCAGTCAGGAATTTTTAAAGGAGGGAGGGGAATTTTTACAGCGAGCGTCTA
CATTTTTACACCAGCCAGCCAACTTTTTTTTATGCATCGTCAGTTAG
Template for rolling circle amplification
CTAACTGACGATGCATAAAAAAAAGTTGGCTGGCTGGTGTAAAAATGTAGACGCTC
GCTGTAAAAATTCCCCTCCCTCCTTTAAAAATTCCTGACTGACTCTC

Functionality of the conjugated cellulosomal components

CelA activity was checked by mixing 80 µM of conjugated CelA with 1% CMC (Carboxymethyl Cellulose) from Sigma. Samples were collected after one h and immediately mixed with 0.5 mL of DNS reagents (10g/liter dinitrosalicylic acid, 10 g/L sodium hydroxide, 2 g/L phenol, 0.5 g/L sodium sulfite). After incubation at 95°C for 10 min, 1 mL of 40% Rochelle salts was added to fix the color before measuring the absorbance at 575 nm.

CelE activity was checked by mixing 200 μM of conjugated CelE with 0.5% PASC (5g/L Phosphoric Acid Swollen Cellulose), prepared from Avicel PH101 (Sigma).[3] Samples were collected after 1 hr and the amount of reducing sugars released was determined as described above.

The functionality of conjugated CBM was confirmed by the ability to bind tightly to cellulose. Conjugated CBM proteins were incubated with Avicel for 1 hr at room temperature. Avicel was then removed by centrifugation and washed once with the binding buffer. The bound proteins were eluted by boiling for 10 min in the elution buffer (1% SDS, 0.1M NaOH).

To assay the activity of conjugated BglA, 100 μL of 10 mM fluorescent substrate p-4-methylumbellifery- β-D-glucopyranoside was mixed with 5 μM of BglA and incubated at 37°C for 10 min. The activity was confirmed by detecting the fluorescence under the UV light.

Electromobility shift assay by native PAGE

Electromobility shift experiments were carried out in a 4.5% non-denaturing acrylamide gel. The protein only, protein plus DNA linkers and purified protein-DNA linkers were loaded and run for 45 min at a constant 90 V. The gel was then stained with ethidium bromide for 5 min before imaging.

Rolling circle amplification

A single stranded DNA template containing four binding sites (5'-GAGAGTCAGTCAGGAATTTTTAAAGGAGGGAGGGGAATTTTTACAGCGAGCGTCTA CATTTTTACACCAGCCAGCCAACTTTTTTTTATGCATCGTCAGTTAG-3') (1 □M) was ligated following a quick circular ligation protocol using the CircLigase ssDNA Ligase (epicentre). Then, 50 nM of the ligated oligonucleotide was treated with 0.5 mM dNTPs, 0.5 nM of linker DNA a and 0.4 U/μl of Phi29 DNA polymerase (NEB) in 50 μl of Phi29 buffer for 1 h, at 30°C before deactivating the enzymes at 65°C for 10 min. Single stranded DNA templates, circulated DNA templates, RCA products and purified RCA products were analyzed on 4.5% native PAGE gels.

Assembly of two- or four-component cellulosomes

For cellulosome assembly, equal concentrations of DNA-labeled cellulosomal components were mixed with the DNA template in a hybridization buffer. The mixture was heated to 37°C and incubated for 10 min. Next, 5X detergent buffer (25 mM DTT, 50% glycerol and 0.5% IGEPAL) was added into the mixture to prevent background interaction. After incubating at 4°C overnight, the mixture were taken out and mixed with 0.5% (5g/L) phosphoric acid-swollen cellulose (PASC). PASC was prepared from Avicel PH101 (Sigma) according to the method.[3] Samples were collected periodically to check reducing sugars or glucose concentration, which was determined with a glucose HK assay kit from Sigma.

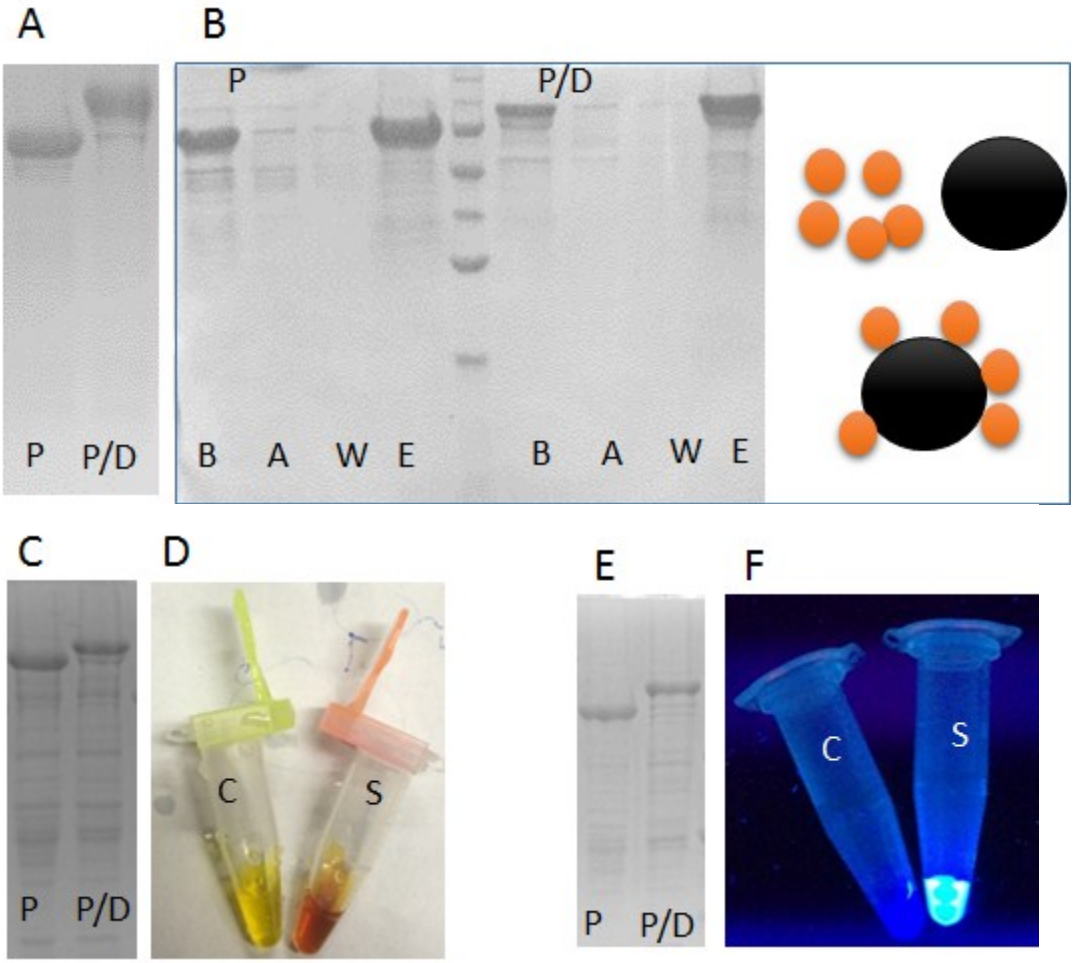

Figure S1. SDS-PAGE analysis (A, C, E) of DNA labeling of CBM-E-H, CelE-E- of CelA and activity check (B, D, F) of CBM, CelE and BglA after labeling and purification. In A, C and E, first lane are purified protein (P) and second lane are purified protein with DNA linker labeling (P/D). (B) The cellulose-binding function of CBM-E-H. Purified proteins (P) or DNA linker 2 modified proteins (P/D) were mixed with avicel and the bound proteins were removed by centrifugation. The amount of CBM-E-A in the solution (B) before or (A) after binding was analyzed. After wash (W), the bound proteins were eluted (E) by boiling.

Figure S2. Rolling circle amplification of DNA template. The discrete DNA template (T) was circulated (C) and rolling circle amplified to synthesize long repeating DNA sequence (R). Control was the rolling circle amplification conducted without circulated DNA template as shown (E). Nothing was detected in (E) demonstrating the synthesis of long repeating DNA template was specifically from circulated DNA template.

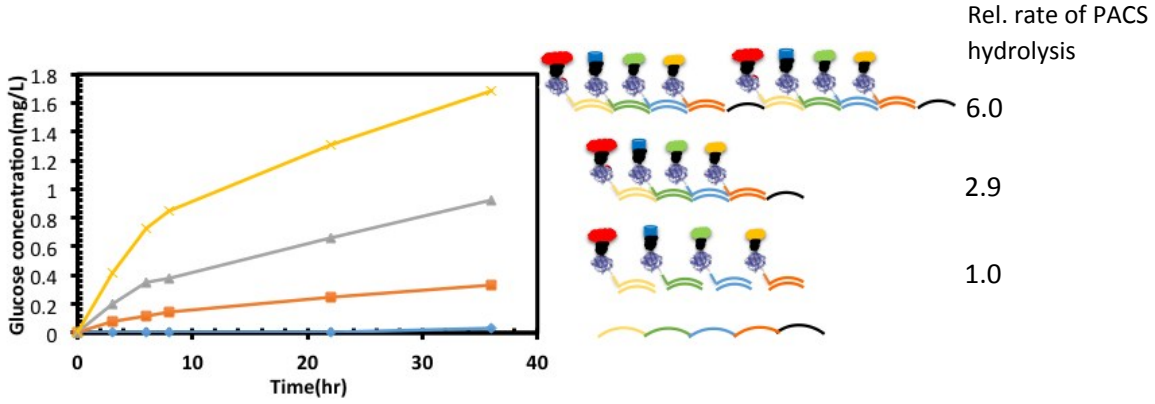

Figure S3. Enhancement in glucose production form PACS by free enzymes, the four-component cellulosome assembled on discrete and RCA templates. The same amount of enzymes and DNA scaffolds was used in both cases.

Reference:

1. F. Liu, S.-L. Tsai, B. Madan and W. Chen, *Biotechnology and Bioengineering*, 2012, **109**, 2829-2835.

2. D. W. Lim, K. Trabbic-carlson, J. A. Mackay and A. Chilkoti, *Biomacromolecules*, 2007, **8**, 1417–1424.

3. Y.-H. Zhang, J. B. Cui, R. L. Lee and R. K. Lana, *Biomacromolecules*, 2006, **7**, 644-648.