Supplementary Information

Creation of artificial cellulosomes on DNA scaffolds by zinc finger protein-guided assembly

Qing Sun,[a] Bhawna Madan,[b] Shen-Long Tsai,[c] Matthew P. DeLisa[d] and Wilfred Chen*[a]

[a] Department of Chemical and Biomolecular Engineering, University of Delaware, Newark, Delaware 19716, United States. Tel: +1 302 831 7650; E-mail: wilfred@udel.edu
[b] Amity Institute of Biotechnology, Uttar Pradesh, India. E-mail: bhawna.madan@gmail.com
[c] Chemical Engineering, National Taiwan University of Science and Technology, Taiwan. E-mail: stsai@mail.ntust.edu.tw
[d] School of Chemical and Biomolecular Engineering, Cornell University, 254 Olin Hall,Ithaca, New York 14853. E-mail: md255@cornell.edu

MATERIALS AND METHODS

Strains and plasmids

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. The ELP[KV_8F-40] polypeptide was constructed by overlapping oligonucleotides (Integrated DNA Technologies, Coralville, IA). The gene fragment was inserted into pET24a to generate pET24a-ELP[KV_8F]40 which contained several restriction sites for inserting gene of interest at the C- terminus. The DNA coding for Zif268 was ordered from Genescript (Piscataway, NJ) and then PCR amplified with the forward primer 5'- CCG GAATTC GGATCC GGC GGC AGC AGCCC and the reverse primer 5'- CCG CTC GAG CTG GTC TTC TTC AGA GAT AAG CTT -3'. The PCR product was digested and ligated into *Bam*H1 and *Xho*1

digested pET24a-ELP[KV_8F]40 to form pE-Z. The DNA coding for PE1A was ordered from Genescript (Piscataway, NJ), digested and ligated into *Sal*1 and *Xho*1 digested pET24a-ELP[KV_8F]40 to form pE-A.

CelA-E-Z was constructed by first PCR amplifying the CelA fragment (1) with the forward primer: 5'-AGTCC CCCGGG CCGCGG CATATGGCTAGC GCGGCC and the reverse primer: 5'- GGG AAT TCC ATA TGG GCG GCC GCC CGC GGG CTG CCG -3'. The amplified CelA fragment was digested and ligated into *Nde*1 digested pE-Z to form pCelA-E-Z. The orientation of CelA was confirmed by DNA sequencing. CBM-E-A was constructed similarly with primers 5'-AGTCC CCCGGG CCGCGG CATATGGCTAGC GCGGCC and 5'-GGG AAT TCC ATA TGG GCG GCC GCC CGC GGG CTG CCG -3' and restriction site *Nde*1 using pE-A as the backbone.

Expression and Purification of ELP-tagged target protein

All constructs were expressed in *E. coli* BLR [*F- ompT hsdSB (r-B m-B) gal dcm(DE3) Δ (srl-recA)306::Tn10(TetR)*; Novagen, Madison,WI]. Overnight cultures were inoculated into 20 mL Terrific broth (TB) medium supplemented with 100 μg/mL kanamycin and incubated at 37°C until OD600 reached 1. The cultures were moved to 25°C for overnight expression. Cells were harvested by centrifugation, resuspended in DNA binding buffer(100 mM Tris base, 90 mM KCl, 1 mM MgCl2, 100 μM ZnCl2, pH 7.5/5 mM DTT), and lysed by ultrasonic disruption using a sonicator.

Purification of the ELP fusion protein was achieved by two cycles of inverse phase transitions. NaCl was added to the cell lysates to a final concentration of 1 M and the mixture was incubated at 37 °C for 10 min before centrifuging for 30 min at 15,000 rpm at the same

temperature. The pellet was resuspended in ice-cold binding buffer and centrifuged for 30 min at 15,000 rpm at 4 °C to remove the insoluble cellular proteins. This precipitation and resolubilization process was repeated a second time and the purity of the protein was determined by 10% SDS-PAGE electrophoresis followed by Coomassie blue staining.

Electrophoretic Mobility Shift Assay

EMSA was performed with the Lightshift Chemiluminescent EMSA kit (Pierce Chemical Co., Rockford, IL) according to manufacturer's recommendations. All the DNA template was 5' biotinylated (Integrated DNA Technologies, Coralville, IA). Purified protein and DNA samples were incubated at 37°C for 1 h. The complexes were subjected to a 10% native acrylamide gel and run under 90 voltage for 30 min. Membrane transferring and chemiluminescence signaling were conducted with protocols suggested by manufacturer.

CBM functionality

CBM fusion proteins were incubated with Avicel for 1 h 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).

Dissociation constant calculation

The DNA concentration used was held constant as 1 nM while the protein concentration varied from 0 nM to 1000 nM. The binding affinities were calculated as follows. Let θ represent the moles of bound DNA per mole of total DNA, [P] represent unbound protein, [D] represent

unbound DNA, [PD] represent the bound protein and DNA complex, $[D]_{total}$ represent the total DNA:

$$K_D = \frac{[P][D]}{[PD]}$$

$$\theta = \frac{[PD]}{[D]_{total}} = \frac{[PD]}{[D]+[PD]} = \frac{\frac{[P][D]}{K_D}}{[D]+\frac{[P][D]}{K_D}} = \frac{[P]}{K_D+[P]}$$

The dissociation constant K_D is equal to the free protein concentration when the DNA bound fraction θ=50%. In this case, the DNA concentration is very low compared with the protein concentration, thus $[P] \approx [P]_{total}$. K_D is then measured as concentration of total protein concentration $[P]_{total}$ when θ=0.5 (Fig. S3).

Artificial cellulosome assembly and reducing sugars assay

Phosphoric acid-swollen cellulose (PASC) was prepared as previous described (1). Enzyme activity was assayed in the presence of a 0.1% (wt/vol) concentration of cellulose at 30°C in 20 nM Tris-HCl buffer (pH 6.0). Samples were collected periodically 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.

Supplemental Results

Figure S1 Purification of CelA-E-Z and CBM-E-A by two cycles of thermal precipitation and solubilization.

Figure S2 Determination of dissociation constants of CelA-E-Z and CBM-E-A by EMSA. All samples contained 1 nM DNA and 0, 10, 25, 50, 100, 300, 500, 1000 nM purified proteins. Controls contained 1 nM control DNA and 1000 nM protein.

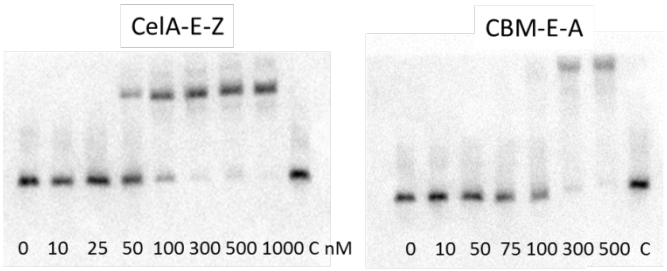
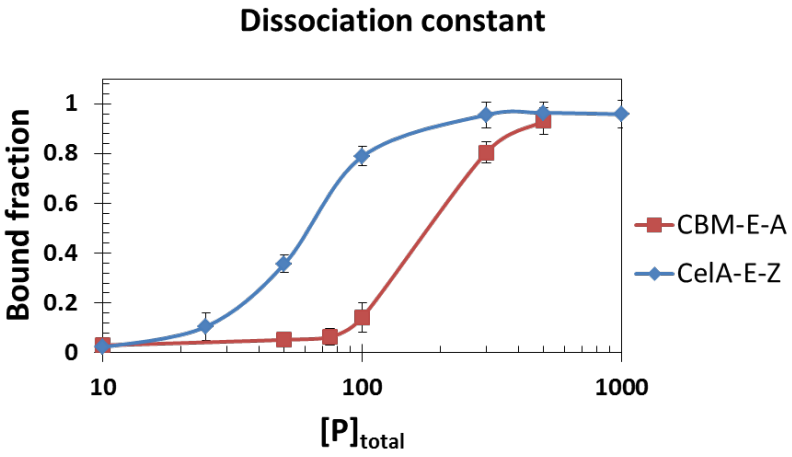


Figure S3 Dissociation constant calculations for CelA-E-Z and CBM-E-A



References

1. S.-L. Tsai, J. Oh, S. Singh, R. Chen, W. Chen, *Appl. Environ. Microbiol.* 2009, **75**, 6087.