Supplementary Material (ESI) for Chemical Communications
This journal is (c) The Royal Society of Chemistry 2013

Aligning an endoglucanase Cel5A from *Thermobifida fusca* on a DNA scaffold: potent design of an artificial cellulosome

Yutaro Mori[a], Shiori Ozasa[a], Shuhei Noda[b], Tsutomu Tanaka[b], Hirofumi Ichinose[c], Noriho Kamiya*[a,d]

[a] Department of Applied Chemistry, Graduate School of Engineering, Kyushu University, 744 Motooka, Fukuoka 819-0395, Japan. E-mail: nori_kamiya@mail.cstm.kyushu-u.ac.jp; Fax: +81 (0)92-802-2810; Tel: +81 (0)92-802-2806

[b] Department of Chemical Science and Engineering, Graduate School of Engineering, Kobe University, 1-1 Rokkodaicho, Nada, Kobe 657-8501, Japan

[c] Faculty of Agriculture, Kyushu University, 6-10-1 Hakozaki, Higashi-ku, Fukuoka, Japan

[d] Center for Future Chemistry, Kyushu University, Japan

Supporting Information

Materials	1
Experimental details	1
Recombinant cellulases prepared in this study	6
SDS-PAGE analysis of wild-type EG and EG-CK	7
RPLC chromatogram of degraded Z-QG-TEO-DNA	8
The calculated standard curve of dNTs	8
The calculated component ratio of Z-QG-TEO-DNA	9
SDS-PAGE analysis of Z-QG-TEO-$(EG)_n$	10
The calculated degree of conjugation of Z-QG-TEO-$(EG)_n$	10
The calculated number of labeled EGs on Z-QG-TEO-DNA	11
Degradation of cellulosic substrates with the artificial cellulosome	12
References	13

Materials

5-[3-(15-Z-Gln-Gly-amido-4,7,10,13-tetraoxapentadecanoylamido)-1-(*E*)-propenyl]-2'-deoxyuridine-5'-triphosphate (Z-QG-TEO-dUTP) was purchased from GeneACT (Fukuoka, Japan). 2'-deoxyadenosine (dA), 2'-deoxycytidine (dC), 2'-deoxyguanosine (dG), 2'-deoxythymidine (dT), nuclease P1 (*Penicillium citrinum*), phosphodiesterase I from *Crotalus adamanteus* venom, ethidium bromide, Avicel® PH-101 and 3,5-dinitrosalicylic acid (DNS) were obtained from Sigma-Aldrich (St. Louis, MO, USA). Sodium chloride, magnesium chloride hexahydrate, acetic acid, acetonitrile, Quick-CBB, agarose LE and carboxymethyl cellulose (CMC) were from Wako Pure Chemical (Osaka, Japan). N,N'-Dimethylformamide (DMF) and imidazole were from Kishida Chemical (Osaka, Japan). Alkaline phosphatase from calf intestine, In-Fusion HD Cloning kit, EX Taq®, PrimeSTAR HS and a pUC18 vector were purchased from Takara Bio. (Shiga, Japan). KOD FX® DNA polymerases were from Toyobo (Osaka, Japan). dNTPs were purchased from Roche Applied Science (Tokyo, Japan). The pET22b(+) plasmid vector was purchased from Novagen (Madison, WI, USA). HisTrap™ HP columns and PD-10 columns were from GE Healthcare UK Ltd (Buckinghamshire, UK). Sodium dodecyl sulfate (SDS) and tris(hydroxymethyl)aminomethane were from Nacalai Tesque (Kyoto, Japan). PCR primers were synthesized at GeneNet (Fukuoka, Japan). A QIAquick PCR purification kit was purchased from QIAGEN Inc. (Alameda, CA, USA). A fluorescein cadaverine and NuPAGE® Novex® 3-8% Tris-Acetate Gel was purchased from Invitrogen (New York, NY, USA). The KOD plus mutagenesis kit was purchased from Toyobo Co. Ltd. (Osaka, Japan). Microbial transglutaminase (MTG) was supplied by Ajinomoto Co., Ltd. (Tokyo, Japan). BCA protein assay kit was purchased from Pierce Biotechnology (Rockford, USA).

Experimental details

1. Expression and purification of recombinant EG(Tfu0901)

The vectors for protein expression using *Streptomyces lividans* as a host were constructed as follows. Polymerase chain reaction (PCR) was carried out using PrimeSTAR HS. The expression plasmids for EG (Tfu0901) and ΔCBM-EG(Tfu0901) contained an acyl-acceptor substrate peptide tag and a 6× histidine tag at the C-terminus (YPYDVDYA-KL-GGGS-MRHKGS-HHHHHH). EG(Tfu0901)-CK and ΔCBM-EG (Tfu0901)-CK were constructed as follows. ΔCBM-EG(Tfu0901) was amplified by PCR using pTONA4-*ps-tfu0901* as a template with the following primers: 5'-CCGGCCTTCGCTAGCGACGAAGGCTCC GAGCCGGG-3' (ΔCBM (Tfu0901)/Fw) and 5'-CGCTCAGTCGTCTCAGTGGTGGTGGTGGTGGTGGGAC TGGAGCTTGCTCCGCA -3' (Tfu0901 /Rv)[1]. The amplified fragment was introduced into the *Nhe*I and *Bgl*II sites of pUC702-pro-sig-term[2] with the In-Fusion HD Cloning kit. The resultant plasmid was called pUC702-*ps-ΔCBM(tfu0901)*. The *ps*-Tfu0901-Ktag-His_6 was amplified by PCR using pTONA4-*ps-tfu0901* as a template with the following primers: 5'- TCGTTTAAGGATGCAGCCATATGATGCTCCGCCACCGGCTC CG-3' (K-tag step1/Fw) and 5'-CGCTCAGTCGTCTCAGTGGTGGTGGTGGTGGTGGCGGGAGCCCTTGT

GGCGCATAAGCTTGGACTGGAGCTTGCTCCGCA-3' (K-tag step1 /Rv). The amplified fragment was introduced into the *Nde*I and *Hind*III sites of pTONA4 with the In-Fusion HD Cloning kit. The resultant plasmid was called pTONA4-Ktag-His6. EG (Tfu0901)-CK and ΔCBM-EG (Tfu0901)-CK were amplified by PCR using the pTONA4-*ps-tfu0901* or pUC702-*ps-ΔCBM(tfu0901)* as a template with the following primers: 5'-TCGTTTAAGGATGCAGCATGCTCCGCCACCGGCTCCGCCG-3' (If-plds/Fw) and 5'-GCCCTTGTGGC GCATGGAGCCGCCGCCAAGCTTGGCGTAGTCCGGGACGTCGTACGGGTAGGACTGGAGCTTGCTCC GCA-3' (Tfu0901-CK2 /Rv), respectively. The amplified fragments were introduced into the *Nde*I and *Hind*III sites of pTONA4-Ktag-His_6 with the In-Fusion HD Cloning kit. The resultant plasmids were called pTONA4-*eg(tfu0901)-CK* and pTONA4-*ΔCBM-eg(tfu0901)-CK*. Transformation, cultivation, and protein purification were conducted according to previous report[1] and the concentration was determined using BCA protein assay kit with a bovine serum albumin standard. Wild-type EG(Tfu0901), EG(Tfu0901)-CK and ΔCBM-EG(Tfu0901)-CK were successfully obtained.

MTG-reactive fluorescein substrates, fluorescein cadaverine or FITC-β-Ala-QG[2] (200 μM), and MTG (1.0 U/mL) were added to an aqueous solution of wild-type EG and EG-CK (10 μM in 25 mM Tris–HCl, pH 7.4) at 25 °C for 1 h. The reaction mixtures were mixed with the standard sample buffer for sodium dodecyl sulfate polyaclylamidegel electrophoresis (SDS-PAGE) analysis (12 vol.% 2-mercaptoethanol, 4 wt.% SDS, 20 vol.% glycerol in 100 mM Tris–HCl, pH 6.8) to terminate the MTG reaction. The reaction was followed by the increase on the fluorescence in the protein bands visualized by the fluorescent image of the SDS–PAGE gel. Before staining the gel with Coomassie Brilliant Blue (CBB) R-250, the fluorescent image of the gels was obtained using a Molecular Imager FX Pro (Bio-Rad Laboratories, Inc.). An excitation wavelength of 488 nm with a 530 (±15) nm band pass filter for the fluorescein derivatives was used (Fig. S1).

2. Expression and purification of *β*-glucosidase from *Themobifida fusca* (BGL(Tfu0937))

The vectors for protein expression using *Escherichia coli* BL21(DE3) as a host were constructed as follows. The plasmid pET22b(+)-*bgl(tfu0937)* was constructed as follows. The DNA fragment encoding BGL(Tfu0937) was amplified by PCR using *Thermobifida fusca* YX and cloned into the pET22b(+) plasmid vector with *Nde*I/*Hind*III double digestion to yield pET22b(+)-*bgl(tfu0937)*. The expression plasmids for BGL(Tfu0937) contained a 6× histidine tag at the C-terminus (KLAAALG-HHHHHH) (Table S1). The additional C-terminus peptide sequence other than the 6× histidine tag (KLAAALG) was derived from the DNA sequence between the *Hind*III site in the multiple cloning site and 6× histidine tag of the pET22b(+) vector.

Escherichia coli BL21(DE3) cells were transformed with the constructed plasmids and a positive clone was selected on LB agar plates containing 100 μg/mL ampicillin. For BGL(Tfu0937) expression from pET22b(+)-*bgl(tfu0937)*, LB medium containing 100 μg/mL ampicillin was inoculated with an overnight-cultured medium and grown until the OD_{600} reached ~0.6 at 37°C. Expression of the BGL(Tfu0937) was induced by the addition of isopropyl-*D*-thiogalactoside (IPTG) to a final concentration of 0.5 mM followed

by shaking for 16 h at 15°C. The cells were then collected by centrifugation at 5,000 *g* for 20 min at 4°C and washed three times in PBS (25 mM NaP, 137 mM NaCl, 2.68 mM KCl, pH 7.4) and subsequently resuspended in 10 mM NaP (pH7.4). The cell suspension was frozen with liquid nitrogen and stored at -80°C. The frozen cell suspension was then completely thawed on ice and sonicated for 10 min with a cooling period of 5 min after the first 5 min of sonication. The supernatant was separated from insoluble cell components by centrifugation at 5,800 g for 10 min at 4°C followed by further centrifugation at 20,400 g for 15 min at 4°C. The supernatant was filtered with a 0.45 μm membrane filter followed by filtration with a 0.2 μm membrane filter. From the filtered supernatant, BGL(Tfu0937) was purified by sequential chromatography as follows. Initially, BGL(Tfu0937) was purified using the 6× histidine tag on a Ni-NTA column and a linear gradient of 40–500 mM imidazole in 20 mM NaP, 0.5 M NaCl, pH 7.4 at 4°C. Fractions containing BGL(Tfu0937) were collected and desalted into 10 mM NaP (pH 7.4) on PD-10 Sepharose columns. The BGL(Tfu0937) solutions were applied to an anion-exchange column and purified using a linear gradient of 0–1 M NaCl in 10 mM NaP, pH 7.4 at 4°C. The fractions containing BGL(Tfu0937) were collected and pooled. All chromatography experiments were conducted on a BioLogic DuoFlow chromatography system (Bio-Rad Laboratories, Inc.) 280 nm. The concentration was determined by using BCA protein assay kit.

3. Preparation of Z-QG-TEO-DNA

We used the DNA fragment of pUC18 cDNA (310bp) as a template DNA. The pUC18 vector is a common used high copy number and general *E.coli* plasmid. This DNA fragment was specifically designed in a neutral area between pUC origin and LacZ gene. DNA containing a fragment of pUC18 cDNA (310 bp) was amplified by PCR with the following primers: 5'-GTAAAACGACGGCCAGTGCCAAG-3' (M13 /Fw) and 5'-AAGAGCGCCCAATACGCAAACC-3' (M13(310 bp) /Rv). The PCR reaction mixture contained 0.2 mM each of deoxynucleoside triphosphate, 1 mM of each primer, 1 ng/mL plasmid and 0.1 U/mL DNA polymerase. The PCR reaction protocol employed was 3 min of denaturation followed by 30 cycles of 98°C for 10 s, 55°C for 20 s and 72°C for 20 s. For the preparation of Z-QG-TEO-DNA, dTTP was substituted (0%, 20%, 40%, 60% and 80%) with Z-QG-TEO-dUTP in the PCR reaction mixtures. The Z-QG-TEO-DNAs were purified using a QIAquick PCR purification kit. The concentration of Z-QG-TEO-DNAs was calculated from the absorbance at 260 nm with UV-vis spectrometer Nano drop (Scrum. Inc.). Z-QG-TEO-DNAs were analyzed by 1.5% agarose gel electrophoresis in TAE buffer and stained with ethidium bromide.

The full sequence of the template DNA is as follows: 5'-GTAAAACGACGGCCAGTGCCAAGCTTGC ATGCCTGCAGGTCGACTCTAGAGGATCCCCGGGTACCGAGCTCGAATTCGTAATCATGGTCATAGC TGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGT GTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTT TCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGT TTGCGTATTGGGCGCTCTT-3'.

4. Determination of the incorporation ratio of Z-QG-TEO-deoxyuridine

Purified Z-QG-TEO-DNA from a PCR reaction mixture was enzymatically digested to deoxynucleosides. The Z-QG-TEO-DNA was incubated with nuclease P1 (2U) from *Penicillium citrinum* in sodium acetate (30 mM, pH 4.5) and zinc chloride (5 mM) for 3 h at 37°C. The solution was further incubated with phosphodiesterase I (0.002U) from *Crotalus adamanteus* venom and with alkaline phosphatase (20 U) from calf intestine in Tris-HCl (100 mM, pH 9.5) and $MgCl_2$ (10 mM) for 3 h at 37°C. After centrifugation, the supernatant was analyzed by using a 4.6 × 250 mm COSMOSIL $5C_{18}$-AR column (Nacalai Tesque) and a mobile phase of triethylammonium acetate (0.1 M, pH 7.0) and acetonitrile. The HPLC gradient conditions employed were as follows: 0% acetonitrile (10 min) followed by a linear gradient of 0 to 30 % acetonitrile (30 min) at a flow rate of 1 mL/min. The effluent was monitored by the absorbance at 300 nm for Z-QG-TEO-deoxyuridine and at 260 nm for the other deoxynucleosides (Figure S2). The relative concentration of each deoxynucleoside was determined using the peak area of each reference standard (Figures S3 and S4).

5. Cross-linking of K-tagged EG(Tfu0901) to the Z-QG-TEO-DNA

The Z-QG-TEO-DNA (5.22 μM Z-QG) was then mixed with 1 equiv. EG(Tfu0901)-CK or ΔCBM-EG(Tfu0901)-CK and 1.0 U/mL MTG in 50 mM phosphate buffer (pH 6.0). The final MTG reaction solution was incubated for 30 min at 40°C. After the incubation, *N*-ethylmaleimide was added to 1 mM to terminate the MTG reaction. The samples were analyzed by 1.5% agarose gels electrophoresis. To identify conjugations, the samples were mixed with the standard sample buffer for SDS-PAGE analysis and separated on a 3-8% gradient gel. After the electrophoresis, the gel was stained with CBB R-250 (Figures S5 and S6). The degree of conjugation was calculated by SDS-PAGE gel analysis using ImageJ software.

6. Enzymatic saccharification of cellulose

The activity was tested in a 1 mL final volume containing substrate (0.5 wt% Avicel or 0.05 wt% CMC), the reaction mixture of conjugated complexes called Z-QG-TEO-$(EG)_n$ (100 nM EG(Tfu0901)-CK or ΔCBM-EG(Tfu0901)-CK) and 100 nM BGL(Tfu0937) in 50 mM phosphate buffer (pH 7.0). The reaction was carried out at 50°C with stirring. The DNS assay was performed to measure the reducing ends of cellulose after the enzymatic reaction. At the end of the enzymatic reaction, 100 μL of DNS reagent containing 1.3 M DNS, 1.0 M potassium sodium tartrate, and 0.4 M NaOH was added to 100 μL of the reaction mixture containing substrate and incubated at 95°C for 5 min to label the reducing ends of the hydrolyzed cellulose. The DNS assay of cellulosic substrate without the conjugates was also conducted in the same way as a control experiment. The optical density was measured at 540 nm, and the activity was determined from a glucose standard curve (Fig. S7).

Table S1. Recombinant cellulases prepared in this study.

Name	Amino acid sequence of the C-terminal regions of the recombinant cellulases
EG(Tfu0901)-CK	----- YPYDVPDYA-KL-GGGS-MRHKGS-HHHHHH
ΔCBM-EG(Tfu0901)-CK*	
BGL(Tfu0937)	----- KLAAALGHHHHHH

*without CBM

Supplementary Material (ESI) for Chemical Communications
This journal is (c) The Royal Society of Chemistry 2013

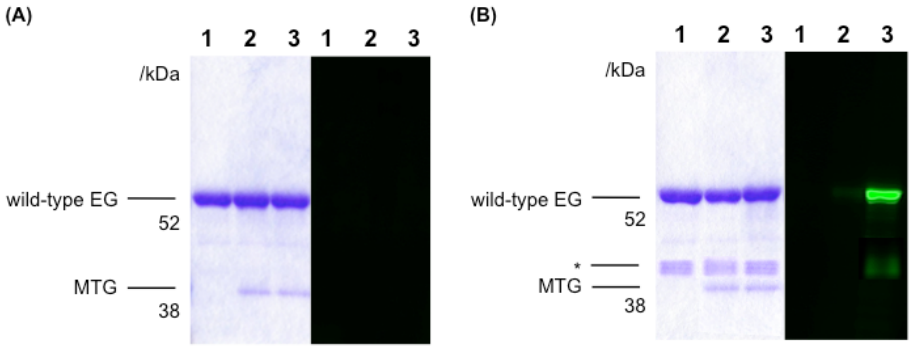

Figure S1. MTG-mediated labeling of wild-type EG (A) and EG-CK (B) with fluorescein substrate. SDS–PAGE analysis of EGs treated without MTG (lane 1), with MTG and fluorescein cadaverine (lane 2) or FITC-β-Ala-QG (lane 3). * shows the degraded EG(Tfu0901)-CK. Left: CBB stained image, right: the corresponding fluorescent image of the gel before CBB staining.

Supplementary Material (ESI) for Chemical Communications
This journal is (c) The Royal Society of Chemistry 2013

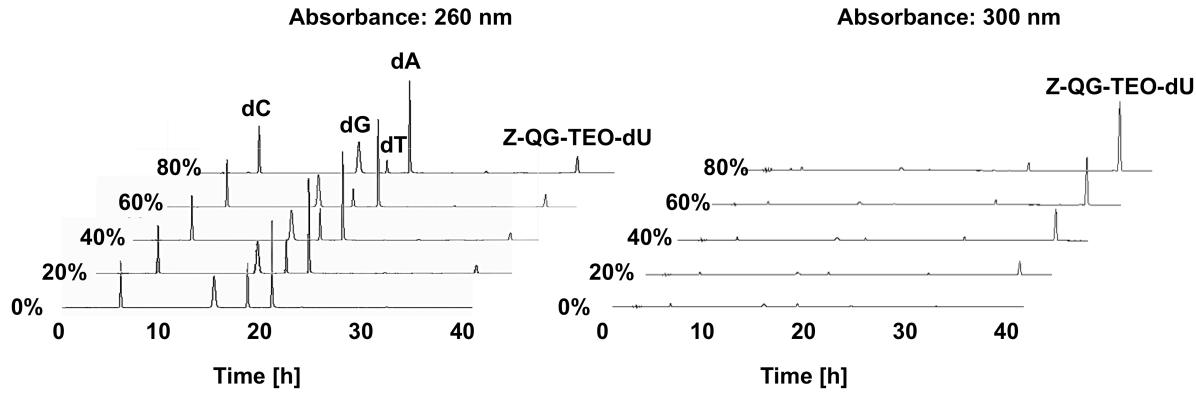

Figure S2. RPLC chromatograms of digested Z-QG-TEO-DNAs monitored by absorbance at 260 nm (left) and at 300 nm (right).

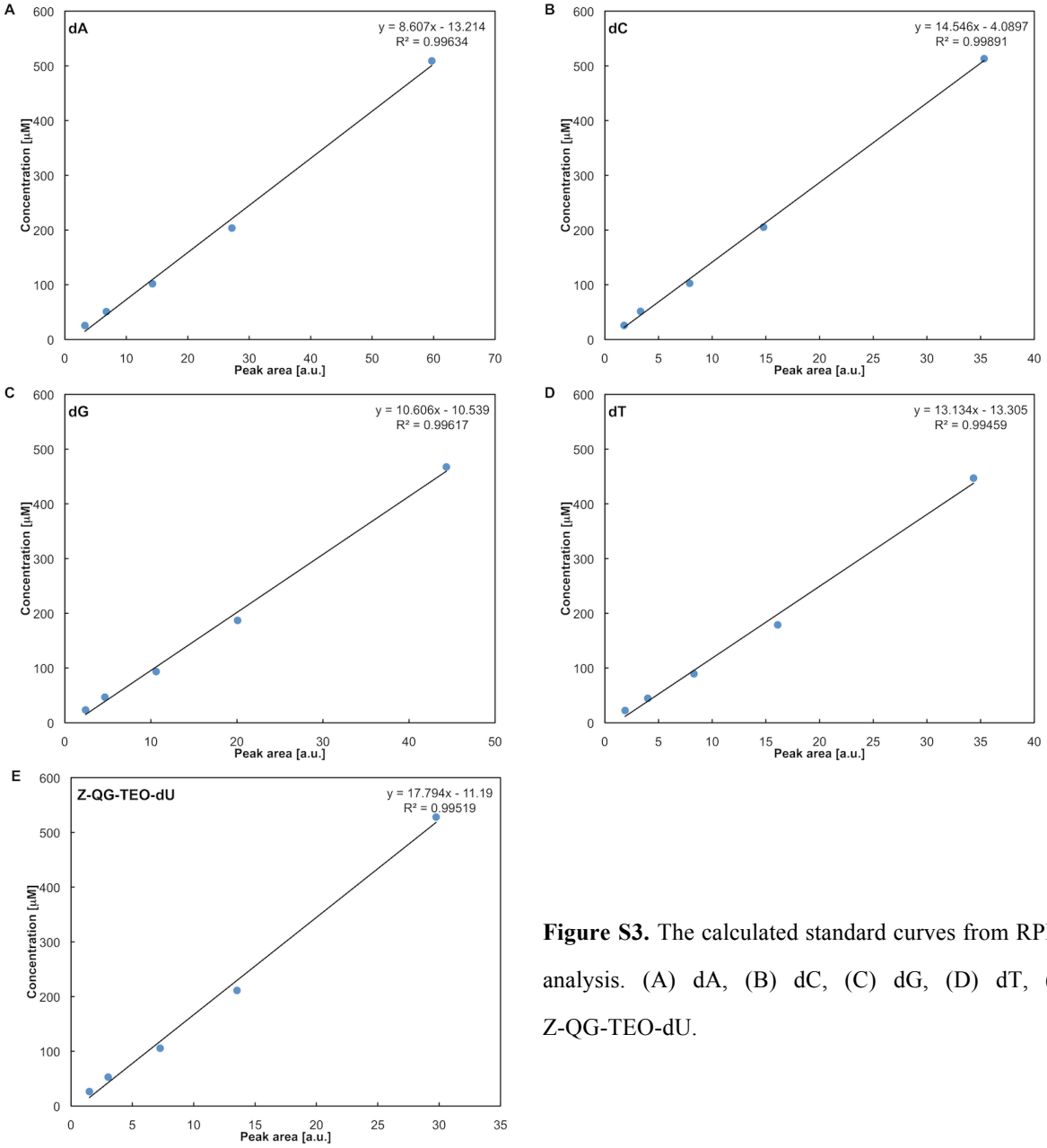

Figure S3. The calculated standard curves from RPLC analysis. (A) dA, (B) dC, (C) dG, (D) dT, (E) Z-QG-TEO-dU.

Supplementary Material (ESI) for Chemical Communications
This journal is (c) The Royal Society of Chemistry 2013

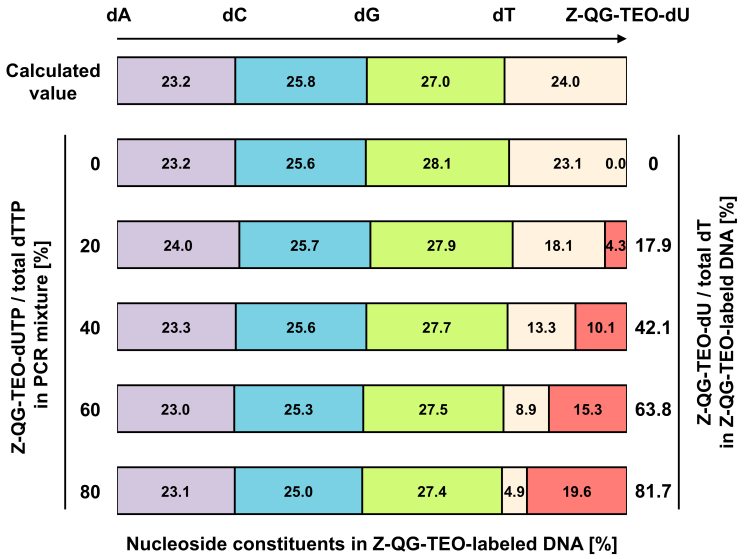

Figure S4. Fragmentation analysis of the incorporation ratio of Z-QG-TEO dUTP in Z-QG-TEO-DNAs. Percentages of nucleosides determined using reversed-phase HPLC analyses are presented from the left: dA, dC, dG, dT, and Z-QG-TEO-dU.

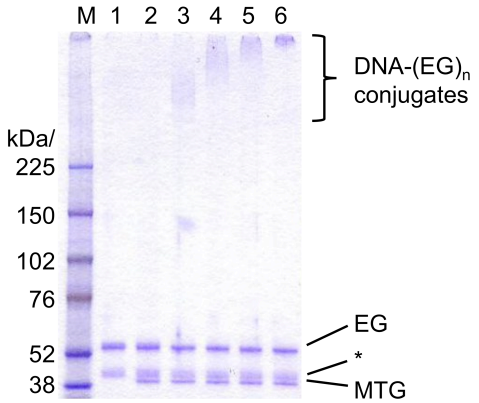

Table. The calculated degree of conjugation

Z-QG-TEO-dUTP /total dTTP [%]	the degree of conjugation [%]
20	22.2
40	22.1
60	21.0
80	21.5

Figure S5. (left) SDS-PAGE analysis of DNA-$(EG)_n$ conjugates without MTG (Lane 1) and with MTG (Lane 2-6). In the MTG reaction mixture, Z-QG-TEO(0%)-DNA (Lane 2), Z-QG-TEO(20%)-DNA (Lane 3), Z-QG-TEO(40%)-DNA (Lane 4), Z-QG-TEO(60%)-DNA (Lane 5) and Z-QG-TEO(80%)-DNA (Lane 6) were incubated with 1 equiv. EG(Tfu0901)-CK ([Z-QG-TEO]/[K-tagged EG] = 1). * shows the degraded EG(Tfu0901)-CK). (right) The degree of conjugation was calculated from SDS-PAGE analysis using ImageJ software.

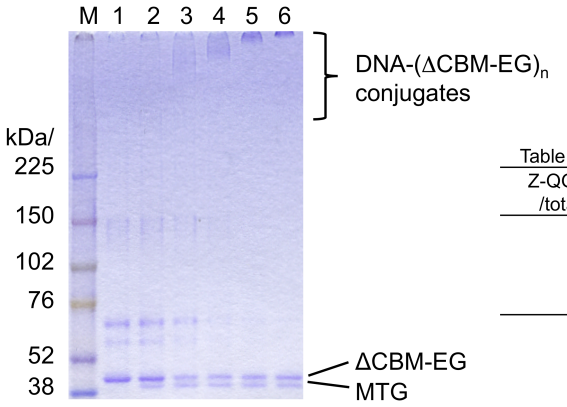

Table. The calculated degree of conjugation

Z-QG-TEO-dUTP /total dTTP [%]	the degree of conjugation [%]
20	61.4
40	64.6
60	56.3
80	51.6

Figure S6. (left) SDS-PAGE analysis of DNA-$(\Delta CBM\text{-}EG)_n$ conjugates without MTG (Lane 1) and with MTG (Lane 2-6). In the MTG reaction mixture, Z-QG-TEO(0%)-DNA (Lane 2), Z-QG-TEO(20%)-DNA (Lane 3), Z-QG-TEO(40%)-DNA (Lane 4), Z-QG-TEO(60%)-DNA (Lane 5) and Z-QG-TEO(80%)-DNA (Lane 6) were incubated with 1 equiv. ΔCBM-EG(Tfu0901)-CK ([Z-QG-TEO]/[K-tagged ΔCBM-EG] = 1). (right) The degree of conjugation was calculated from SDS-PAGE analysis using ImageJ software.

Table S2 The calculated number of labeled EGs* and the used equation.

	Z-QG-TEO-dUTP /total dTTP, X [%]	the incorporation ratio of Z-QG, A [%]	the degree of conjugation, B [%]	the number of labeled EGs, N [numbers / a DNA scaffold]
DNA-(EG)n	20	17.9	22.2	5.9
	40	42.1	22.1	13.8
	60	63.8	21.0	19.9
	80	81.7	21.5	26.1
DNA-(ΔCBM-EG)$_n$	20	17.9	61.4	16.4
	40	42.1	64.6	40.5
	60	63.8	56.3	53.4
	80	81.7	51.6	62.7

$$N(X\%) = L \times 2 \times (T/100) \times (A(X\%)/100) \times (B(X\%)/100)$$

N(X%) : the number of labeled EGs onto Z-QG-TEO(X%)-DNA, [numbers / a DNA scaffold]

L : the number of nucleotides of template DNA, 310 [nt]

T : the theoretical percentage of dTTP in double-strand DNA template, 24 [%]

A(X%) : the incorporation ratio of Z-QG-TEO dUTP in Z-QG-TEO(X%)-DNA, [%]

B(X%) : the degree of conjugation on Z-QG-TEO(X%)-DNA, [%]

X : the percentage of Z-QG-TEO-dUTP to dTTP in PCR mixture; 20, 40, 60, 80 [%]

*The natural cellulosome CipA has nine enzymes and one CBM[3].

Supplementary Material (ESI) for Chemical Communications
This journal is (c) The Royal Society of Chemistry 2013

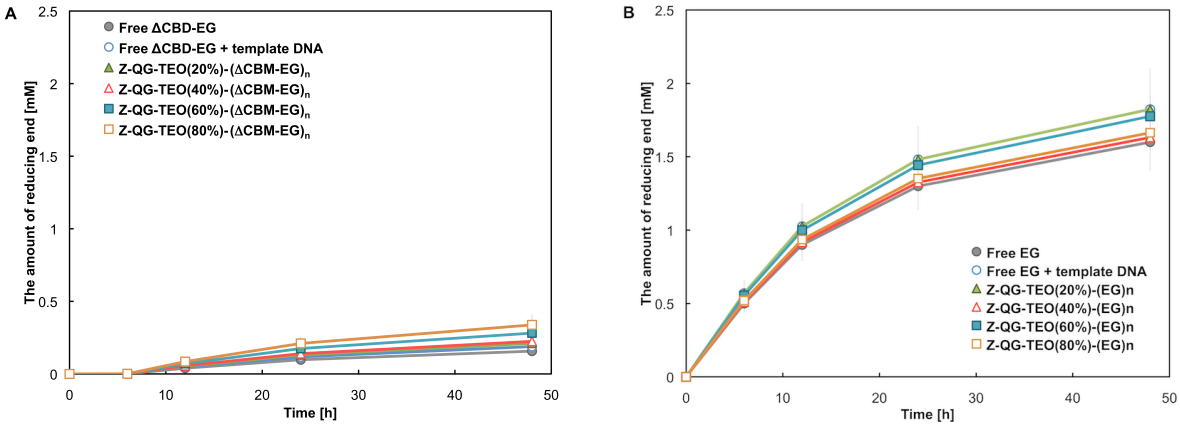

Figure S7. Degradation of cellulosic substrates by an artificial cellulosome with BGL(Tfu0937). The reaction mixture contained substrate (0.5 wt% Avicel (A) or 0.05 wt% CMC (B)), and the artificial cellulosome reaction mixture (100 nM ΔCBM-EG(Tfu0901)-CK or EG(Tfu0901)-CK) and 100 nM BGL(Tfu0937)).

References

1. a) S. Noda, E. Kitazono, T. Tanaka, C. Ogino, A. Kondo, *Microb. Cell Fact.*, 2012, **11**, 49-58. b) S. Noda, Ito Y, N. Shimizu, T. Tanaka, C. Ogino, A. Kondo, *Protein Expr. Purif.*, 2010, **73,** 198-202.
2. N. Kamiya, H. Abe, M. Goto, Y. Tsujic, H. Jikuya, *Org. Biomol. Chem.*, 2009, **7,** 3407–3412.
3. C. M. G. A. Fontes and H. J. Gilbert, *Annu Rev Biochem*, 2010, 79, 655-681.