Supplementary Information

Experimental Procedures

Biomass preparation: Napier grass was heat-pretreated with dilute sulfuric acid as described in a previous report,[S1] and the cellulose components were hydrolyzed with 40 FPU/gBM of Cellic CTec2 (Novozymes, Bagsvaerd, Denmark) in 50 mM sodium acetate buffer (pH 5.0) at 50°C for 60 h; consequently, more than 90 % celluloses were removed. The cellulase-treated biomass was washed with 6 M guanidine hydrochloric acid to remove the cellulases and was then washed with water to remove the denaturant.

Amorphous cellulose (PSC) was prepared from Avicel (PH-101, Sigma-Aldrich Inc., Tokyo, Japan) as described by Walseth et al.[S2]. Avicel (4 g) was dissolved in 100 ml of phosphoric acid. After stirring for 1 h at 4 °C, the solution was diluted with 1900 mL of cold water. After stirring for 1 h at 4°C, the extracted amorphous cellulose was collected by filtration with filter paper. The amorphous cellulose was washed 4 times with ultrapure water, 2 times with 1% $NaHCO_3$ (neutralization), and then 3 more times with ultrapure water. The cellulose paste was homogenized (2 min × 3) with a Multi-brander mill BLA-501 (Nihonseiki Kaisha Ltd., Japan). The resultant slurry was resuspended with 50 mM sodium acetate buffer (pH 5.0) for enzyme assay and stored at 4 °C.

Selection of biomass-binding peptides: A phage library with a 12-mer random peptide displayed on the surface (Ph.D.-12 Phage Display Peptide Library Kit, New England Biolabs, Beverly, MA) was used for the selection of peptide sequences with affinity for cellulase-degraded napier grass. The library scale of the peptides fused on the bacteriophages was 2×10^9. Approximately 10^{11} phages were mixed with 2 mg of the biomass in 100 μL of a 50 mM sodium acetate solution (pH 5.0) with 200 mM NaCl and 0.5% Tween 20 detergent, and the solution was incubated at room temperature with gentle agitation for 10 min. After exclusion of unbound phages from the

phage/biomass solution, phages with peptides bound to the biomass were separated from the biomass by 50 mM phosphate solution (pH 7.5) containing 2 M NaCl and 0.5% Tween to obtain the eluted peptide fractions. *E. coli* (K12 ER2738 strain) was infected with the eluted phages for amplification by 10^{11}, and the amplified phages were then mixed with the biomass suspension again. After four rounds of screening, DNA sequences encoding the peptides displayed on the phages eluted with 2 M NaCl and the residual phages on the biomass were analyzed for 27 phage clones.

Further, to obtain the phages that were not eluted with the 2 M NaCl solution, we infected *E. coli* with the residual phages on the biomass by mixing *E. coli* with the biomass treated with 2 M NaCl. The phages were amplified by 10^{11}, and the amplified phages were then mixed with the biomass suspension again. This screening step, in which *E. coli* was infected with the residual phages bound on the biomass even after 2 M NaCl treatment, was repeated 4 times, and DNA sequences encoding the peptides displayed on the phages which were not eluted with 2 M NaCl were analyzed for 50 phage clones.

Binding assay of identified peptides for biomass materials: A sample of each of the identified peptides with an FITC molecule and a glycine linker (GGGS) at the N terminus was purchased from Sigma-Aldrich Inc. The C-termini of all the peptides were amidated. Each FITC-labeled peptide solution (1 mL of a 50 mM sodium acetate solution (pH 5.0) with 200 mM NaCl, various concentrations) was mixed with 5 mg of each biomass (untreated napier grass, cellulase-degraded napier grass, Avicel, and PSC) for 10 min at room temperature. After centrifugation at 12,000×g for 10 min, the FITC fluorescence in the supernatants was measured on an F-2500 fluorescence spectrophotometer at an excitation wavelength of 495 nm (Hitachi Science & Technology, Tokyo, Japan) to estimate the unbound peptide concentration. BSA, which was labeled with a Fluorescein Labeling kit -NH_2 (DOJINDO MOLECULAR TECHNOLOGIES, INC., Japan), was also mixed with biomass materials, and the unbound BSA was estimated with the same method as used for peptides.

Enzyme activity assays: Synthesized peptides and BSA were mixed with 5 mg mL^{-1} heat-pretreated napier grass in a 50 mM sodium acetate solution (pH 5.0, 200 mM NaCl) at a concentration of 4, 40, or 400 μg mL^{-1}. The suspensions were allowed to stand for 12 h at 4°C, and then 2.5 μM recombinant endoglucanase D (CelD, prepared as described in a previous report[S3]) was added to degrade cellulose at 45°C. After an incubation interval of 1–96 h, 5 μL of the supernatant was mixed with 195 μL of tetrazolium blue chloride assay buffer (1 mg mL^{-1} tetrazolium blue chloride, 0.5 M sodium tartrate, 200 mM NaOH, pH 5.0) at 100°C for 3 min. After the reaction, solutions were rapidly cooled in an ice bath, and their absorbances at 655 nm were measured. The concentrations of produced reducing sugars were estimated from the absorbances by normalization with the absorbance of glucose reacted in tetrazolium blue chloride assay buffer.

References

S1. D. J. Schell, J. Farmer, M. Newman and J. D. Mcmillan, Appl. Biochem. Biotechnol., 2003, 105, 69.

S2. C. S. Walseth, *TAPPI*, 1952, **35**, 228.

S3. D.-M. Kim, H. Nakazawa, M. Umetsu, T.Matsuyama, N.Ishida, H. Takahashi, R. Asano and I. Kumagai, *Catal. Sci. Technol.*, 2012, **2**, 499.

Results

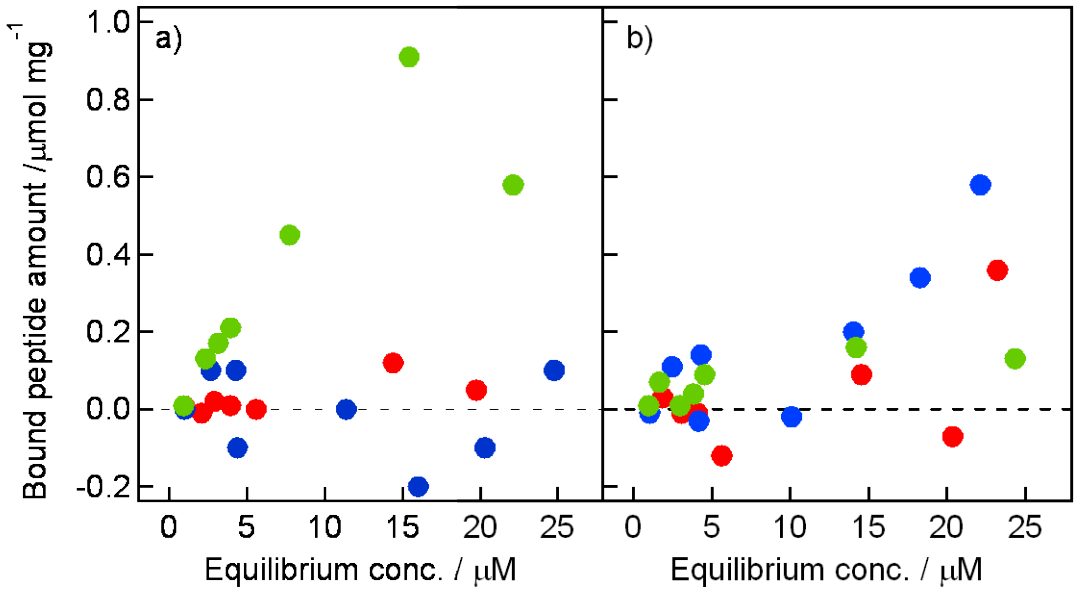

Fig. S1 Adsorption isotherms for BioMBP$_n$1 (red circles), BioMBP$_{nu}$1 (blue circles), and BioMBP$_{nu}$2 (green circles) against 5 mg mL^{-1} Avicel (a) and phosphoric acid swollen cellulose (b). The binding buffer was a 50 mM acetate solution (pH 5.0, 200 mM NaCl).

Table S1. Secondary structure prediction, hydropathy index, and isoelectric point (pI) for identified biomass-binding peptides

Peptide	Sequence	pI [b]	Hydropathy index
BioMBP$_n$1	S G H H N L H K T E H R	8.6	-26.3
secondary structure prediction[a]	C C C C C C C E E E E C		
BioMBP$_{nu}$1	S S L Q A H K P H H L R	11.0	-15.3
secondary structure prediction	C C C C C C C C C E E C		
BioMBP$_{nu}$2	K H V P R S P V E A L Y	8.6	-6.4
secondary structure prediction	C C C C C C C E E E E C		

[a] Secondary structures were predicted using GOR4 program at http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_gor4.html.
C, random coil; E; extended strand.
[b] Isoelectric point, pI, was calculated using the pI/Mass program at http://web.expasy.org/compute_pi/