Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

Modular Architecture of Protein Binding Units for Designing Properties of Cellulose Nanomaterials

*Jani-Markus Malho, Suvi Arola, Päivi Laaksonen, Géza R. Szilvay, Olli Ikkala, and Markus B. Linder**

anie_201505980_sm_miscellaneous_information.pdf

Supporting Information

Table of contents:

Nanofibrillated Cellulose

A dilute hydrogel (solid content 1.64%) of nanofibrillar cellulose was used as a starting material. The cellulose was mechanically disintegrated by ten passes through a M7115 Fluidizer (Microfluidics Corp., U.S.A.) essentially according to previous reports[1]. Nanofibrillated cellulose was then used for the films.

Gene Constructs

Three protein constructs, HFB-dCBM-12, -24, and -48, were designed where all contained an N-terminal hydrophobin HFBI[2] linked to two cellulose binding modules (dCBM)[3], but each had a different sized linker sequence between the two cellulose binding modules (CBM). The constructs were based on the previously published protein HFBI-DCBD described by Linder *et al*. [3-4] The linker sequences between the CBMs were designed so that HFB-dCBM-24 has a linker length of 24 amino acids (linker B) (identical to HFBI-DCBD[3]), HFB-dCBM-12 has a linker length of 12 amino acids (linker A) and is comprised of the first half of linker B, and HFB-dCBM-48 has a linker where the linker B was doubled to 48 amino acids (linker C). The linker between HFB and dCBMs was identical in all three constructs and contained two protease sites; one for Tobacco Etch Virus protease enzyme (TEV)[5] and one for trypsin. All linker sequences are summarized in Table S1.

Synthetic genes encoding HFB, dCBM-12, dCBM-24, and dCBM-48 flanking with compatible BsaI-sites were ordered from GenScript USA Inc. (NJ, USA) with *Trichoderma reesei* optimized codon usage in BsaI-free pUC57 plasmids. Golden Gate cloning [6] was used for assembling the *T. reesei* transformation cassettes from six pieces into a destination plasmid (pJJJ307, which is based on the pMK-RQ backbone) resulting in plasmids pHFB-TEV-dCBM-12, pHFB-TEV-dCBM-24 and pHFB-TEV-dCBM-48. The assembled parts were (in 5'-3' order) i) cbh1 promoter (pJJJ308), ii) HFBI coding sequence with signal sequence for secretion, iii) dCBM (12, 24 or 48) coding sequence, iv) cbh1 terminator (pJJJ311), v) hph hygromycin resistance marker gene under the gdpA promoter for strain selection (pJJJ312), and vi) a 3' flank sequence (pJJJ313) that together with the CBHI promoter serves to guide the gene recombination into the cbh1 locus in *T. reesei*. The ligation products were then transformed into *E. coli* XL1-Blue cells and plasmid containing transformants were selected on kanamycin plates and further selected by blue-white screening for insert-containing clones. Correct assembly of the transformation cassette was verified by restriction enzyme analysis and the correct protein coding sequence was verified by DNA sequencing. The transformation cassettes were then cut from the vectors using PmeI and purified by agarose gel electrophoresis followed by gel extraction using dialysis.

Table S1.

DNA and amino acid sequences of the linkers in the three fusion proteins HFB-dCBM-12, -24, and -48.

Linker	Sequence[a]
Linker 12	CCCGGCGCAAACCCGCCTGGCACCACCACCACCAGC PGANPPGTTTTS
Linker 24	CCCGGCGCAAACCCGCCTGGCACCACCACCACCAGCCGCCCA GCCACTACCACTGGAAGCTCTCCCGGACCT PGANPPGTTTTSRPATTTGSSPGP
Linker 48	CCCGGCGCAAACCCGCCTGGCACCACCACCACCAGCCGCCCA GCCACTACCACTGGAAGCTCTCCCGGACCTCCCGGCGCAAAC CCGCCTGGCACCACCACCACCAGCCGCCCAGCCACTACCACT GGAAGCTCTCCCGGACCT PGANPPGTTTTSRPATTTGSSPGPPGANPPGTTTTSRPATTTGSS PGP
Linker connecting HFB and dCBM	GGTGGAGGCTCTGGTGGAGGCTCAGGTGGAGGCAGTGAGAA CCTCTACTTCCAGGGCCCGGGCGCGAGCACCAGCACCGGCC GCGGCCCGGGCGGC GGGSGGGSGGGSENLYFQ/GGLQGTPGASTSTGR/GPGG

[a] DNA sequences are shown in black and amino acid sequences in blue.

Transformation of *Trichoderma reesei*

Biolistic PDS-1000/He Particle Delivery System (Bio-Rad, CA, USA) was used for the transformation of *T. reesei* M122 spores. The particle bombardment was performed under a helium pressure of 850 psi. Gold microcarrier particles (0.6 µm, Bio-Rad) were coated with the linear transformation cassette DNA according to the instruction manual of Bio-Rad laboratories. Gold particles without DNA were used as a control. After transformation the spores were plated on PD-plates with top-agar containing 1M $MgSO_4$, 1M $CaCl_2$ and 150 µg/ml hygromycin for selection and grown for 3 – 10 days in 28 °C. For selecting transformants several colonies were picked and streaked on PD-Triton plates containing hygromycin and grown for 5 – 8 days in 28 °C. The transformants were then re-streaked to new selective Triton plates and grown as above to ensure growth of only hygromycin resistance containing transformants. Insert containing transformants were identified by direct PCR (Phire® Plant Direct PCR kit, Finnzymes, F-130) using suitable oligonucleotide primers. Correct recombination localization into the cbh1 locus was verified by the absence of an amplicon using cbh1 specific primers.

Protein Production and Purification

The proteins were produced under the following conditions in 50 ml shake flask cultures: *T. reesei* minimal media, 4% lactose , 2% spent grain extract, 100 mM PIPPS (Piperazine-N,N′-*bis*(3-propanesulfonic Acid)) pH 5.5, 2.4 mM $MgSO_4$, 4.1 mM $CaCl_2$, 28°C, 7 days. To identify protein producing strains culture media and biomass extracts from growth time points were analysed by western blotting using rabbit anti-HFBI antibodies. The transformants selected for protein production were VTT-D-133335 (HFB-dCBM-12), VTT-

D-133336 (HFB-dCBM-24), and VTT-D-133337 (HFB-dCBM-48). The strains were then cultivated in 7 L bioreactors on media containing 50 vol-% spent grain extract, 60 g/L lactose, 1 g/L yeast extract, 4 g/L KH_2PO_4, 2.8 g/L $(NH_4)_2SO_4$, 0.6 g/L $MgSO_4 \cdot 7H_2O$, 0.8 g/L $CaCl_2 \cdot 2H_2O$, 2 ml/L trace solution. The pH was let to drop from 5 to about 3 during cultivation. At 24 h intervals 48 mg pepstatin A and 28 mg soy bean trypsin inhibitors (both from Sigma-Aldrich) were added to the cultures to minimize protein degradation. The culture supernatants were separated from the biomass by filtration through GF/A filters (Whatman). Protein expression levels were analysed by RP-UPLC and were 0.2 g/L, 0.4 g/L, and 3.0 g/L for HFB-dCBM-12, -24, and -48, respectively. The proteins were purified using aqueous two phase extraction and reverse-phase high-performance liquid chromatography (RP-HPLC) as described earlier [7] followed by lyophilisation.

Linker digestions

The fusion proteins were cleaved with sequencing grade modified trypsin (Promega) in 25mM Tris-HCl – 150mM NaCl buffer for 2 hours in room temperature. The trypsin digestion reaction was followed by RP-UPLC using a 2.1 x 100 mm, 1.7 µm, C4 Acquity BEH300 prST column and an Acquity I-Class system with a photodiode array detector (Waters, MA, USA). The proteins were eluted in a linear mobile phase gradient from 20 – 60% B using water (A) and acetonitrile (B), both containing 0.1% trifluoroacetic acid. Concentrations of the analysed proteins were determined using standard samples with protein concentrations determined by amino acid analysis (Amino Acid Analysis Lab, Uppsala University, Sweden).

Preparation of ^{3}H–Labeled dCBM.

All trypsin cleaved proteins were labelled with tritium for interaction studies by reductive methylation. 1.9 mg of lyophilized protein was dissolved in 1.9 ml of 0.2 M borate buffer, pH 8.96 and cooled on an ice bath. 13.2 ml of 0.37 % formaldehyde and 100 mCi of ^{3}H enriched $NaBH_4$ (10 Cimmol^{-1}, NET023H100MC, PerkinElmer) in 150 ml of 0.01 M NaOH were added and mixed for 30 minutes. The reaction was terminated by RP-HPLC. The specific activities were 0.52 Cimmol^{-1} for dCBD-12, 0.69 Cimmol^{-1} for dCBD-24 and 1.25 Cimmol^{-1} for dCBM-48.

The labelling did not affect CBM function. The reductive methylation occurs at amine residues which were only present at the free N-terminus. Different ratios of labelled and non-labelled protein gave identical results showing that protein function was not affected.

Determination of Binding Isotherm of dCBMs to NFC.

A 100 mM stock solution of dCBMs containing 10 % ^{3}H–labelled proteins in 100 mM sodium acetate buffer (pH 5.0) with 100 mM NaCl and 0.1% BSA was prepared. Solutions with different fusion protein concentrations were made from the stock solution by diluting with the same buffer. The BSA was used to minimize nonspecific binding of the dCBMs to test tube walls and filters. 100-200 ml of each protein solution was mixed with an equal volume of 1-2 g L^{-1} NFC in MilliQ–water and stirred in ambient temperature with 250 rpm for 1 h. After equilibration, the suspensions were filtered through disposable filters (Millipore, Millex®–GV filter unit, PVDF, hydrophilic, 0.22 mm, 13 mm, non–sterile) and the

amount of free protein was determined by liquid scintillation counter (PerkinElmer) from the filtrate.

Table S2.
The parameters for binding of different dCBMs to NFC

	Kd [uM]	Bmax [umolg^{-1}]
12-mer	1,50 +/- 0.19	8.63 +/- 0.24
24-mer	1,37 +/- 0.11	8,89 +/- 0.17
48-mer	1,32 +/- 0.20	3.62 +/- 0.15

Film Preparation and Humidity Control

The NFC concentration was kept at 2.0gl^{-1} in all films. Vacuum filtration was used to prepare films from dispersions mixtures containing NFC and proteins. The dispersions were filtrated using a Durapore membrane (GVWP, 0.22 µm, Millipore, U.S.A.) and an O-ring to define the film diameter. A gentle pressure was applied to the films using a 300 g load for 10 min to prevent wrinkling before oven drying overnight at +65 °C. Samples were then stabilized in a humidity controlled room at 50%RH prior to measurements.

Mechanical Tensile Testing

Tensile testing was performed on 5kN Tensile/compression module (Kammrath & Weiss GmbH, Germany) using 100N load cells. The elongation speed was 0.5 mm/min and the gauge length 10 mm. Samples were stabilized overnight in a humidity controlled room of 50%RH prior to the mechanical tensile testing. All of the tensile tests were performed at 50%RH. At least 4 specimens were measured from each sample. Specimen sizes were 2 cm × 2 mm × 7−13 µm, length, width, and thickness, respectively.

High resolution transmission electron microscopy

High resolution transmission electron microscopy imaging was performed using JEM-3200FSC field emission microscope (JEOL), which was operated at 300 kV in bright field mode with Omega-type Zero-loss energy filter. Ultrascan 4000 CCD camera (Gatan) was used to acquire the images of samples that were maintained at -187°C. 3 µl of sample dispersions were vitrified on c-flat grids under 100% humidity with FEI Vitrobot. The sample dispersions were blotted for 1.5 s with a filter paper before vitrification.

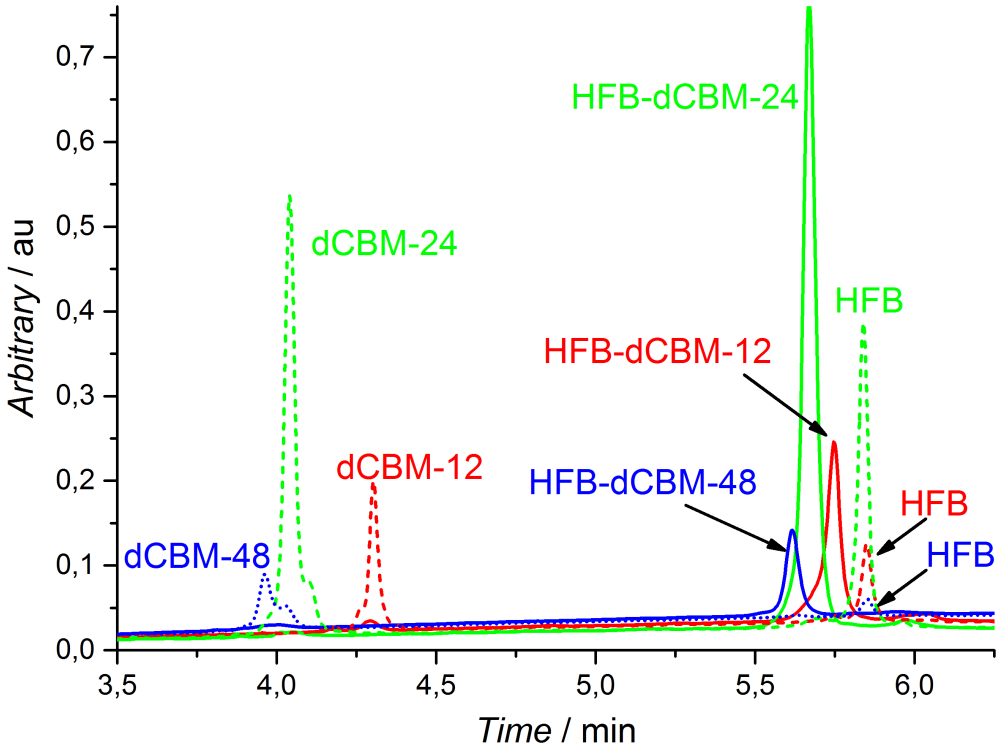

Figure S1. UPLC- chromatograms.

UPLC- chromatograms of the fusion proteins and the trypsin cleaved proteins. Bright red line represents HFB-dCBM-12 peak and the dark red line the trypsin cleaved dCBM-12 and the HFB peaks. Bright green line represents HFB-dCBM-24 and the dark green displays the cleaved dCBM-24 and HFBI. Dark blue line is the HFB-dCBM-48 and the lighter blue line the dCBM-48 and HFBI. The concentrations used are not the same across samples and thus the peaks heights vary, whereas the elution times are related to differences between the protein structures. Peaks that are located at 3.80-4.40 minutes are cleaved dCBM-proteins, the fusion protein peaks are located between 5.50 and 5.80 minutes, and the cleaved HFBI-proteins from different constructs around 5.82 minutes. The HFB-peaks from different fusion proteins elute almost at the same elution time, while the intact fusion proteins and dCBM protein -peaks show variation due to differences in their sequence and structure. The shoulders in the dCBM-peaks are expected to results from slight glycosylation occurring in the linker sequence between the two CBM-domains in the dCBM-structure.

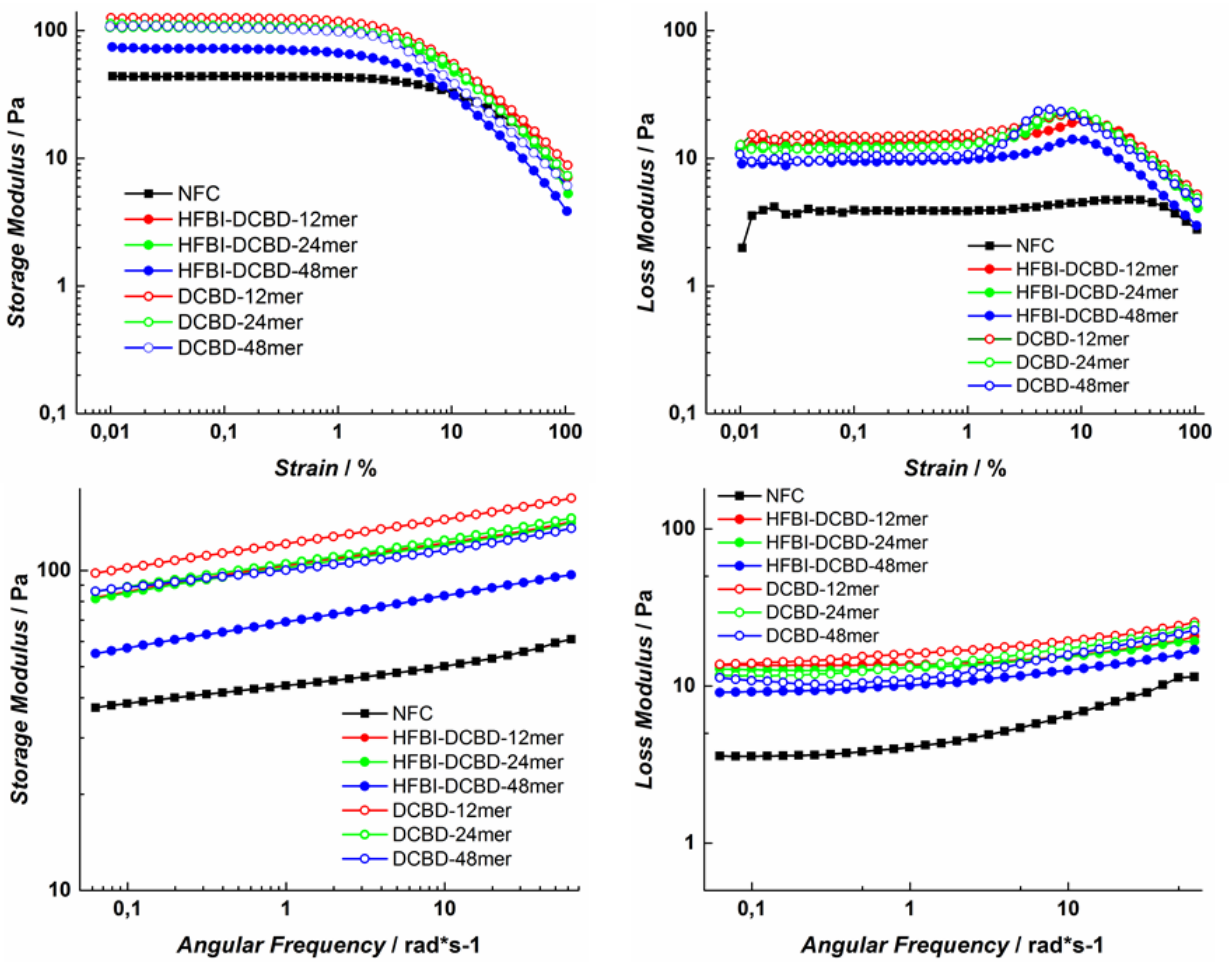

Figure S2. Rheology results

Rheological data with strain and angular frequency sweeps for 2 gl^{-1} NFC and 42% (w/w) protein.

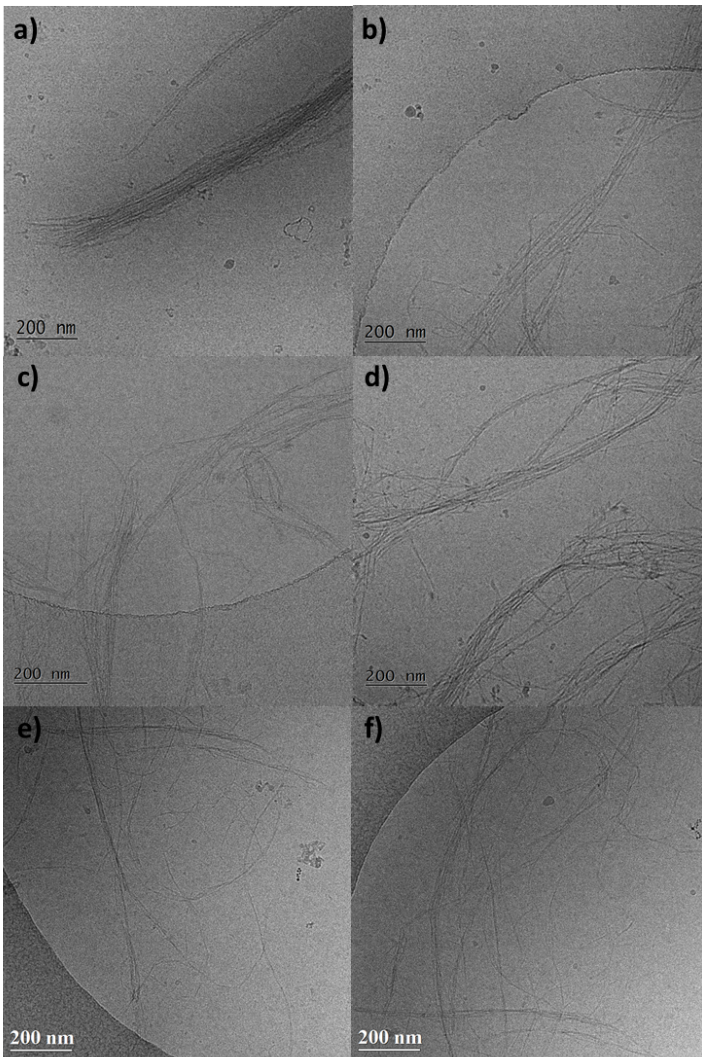

Figure S3. Cryo-TEM images

Cryo-TEM images of vitrified HFB-dCBM-48 and dCBM-48 dispersions with NFC. a-b) Represent HFB-dCBM-48-NFC dispersions, c-d) dCBM-48-NFC dispersions and e-f) unmodified NFC dispersions. Parts of individual nanofibers can be seen in all of the dispersions. Yet, in a-d) all of the nanofibers seem to be bound to other nanofibers, which is expected to result from biomolecular crosslinking of NFC nanofibers through HFB-dCBM-48 and dCBM-48 proteins. The amounts of HFB-dCBM-48 and dCBD-48 are 0.85gl^{-1} and the amount of NFC is 2.0gl^{-1}. The unmodified NFC nanofibers display also aggregations. However, individual nanofibers can be seen were as the protein-containing dispersion were more aggregated with no individually dispersed nanofibers. The results are qualitative as the NFC nanofibers tend to aggregate natively.

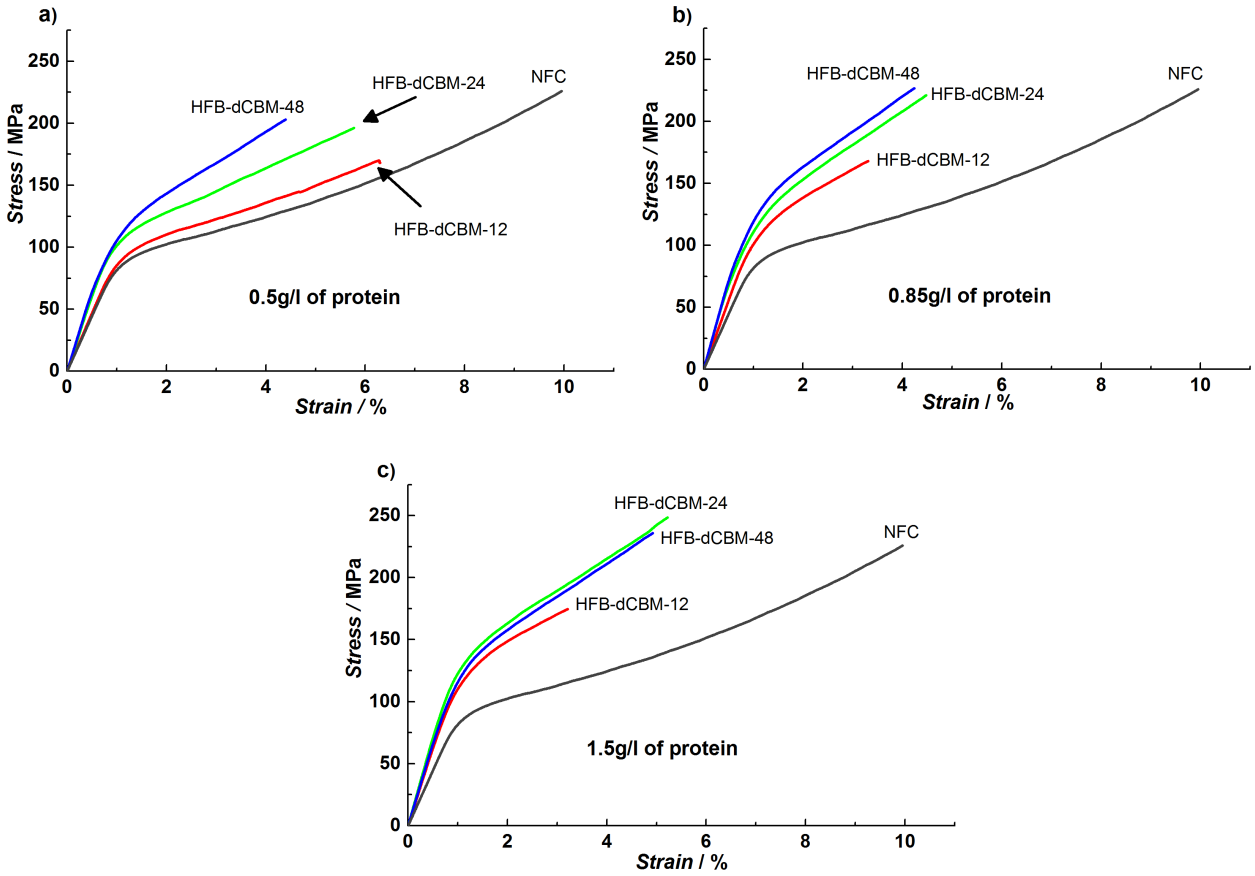

Figure S4. Stress-strain curves of fusion proteins

Representative stress-strain curves of fusion proteins HFB-dCBM-12, -24 and -48 in NFC matrix, where the red lines represent HFB-dCBM-12-NFC, green lines HFB-dCBM-24-NFC and blue lines HFB-dCBM-48-NFC. Unmodified NFC film is drawn with black line. a)-c) exhibit fusion proteins in three concentrations ($0.5gl^{-1}$, $0.85gl^{-1}$ and $1.5gl^{-1}$, respectively). The concentration of NFC is kept $2.0gl^{-1}$ in all of the films throughout the study.

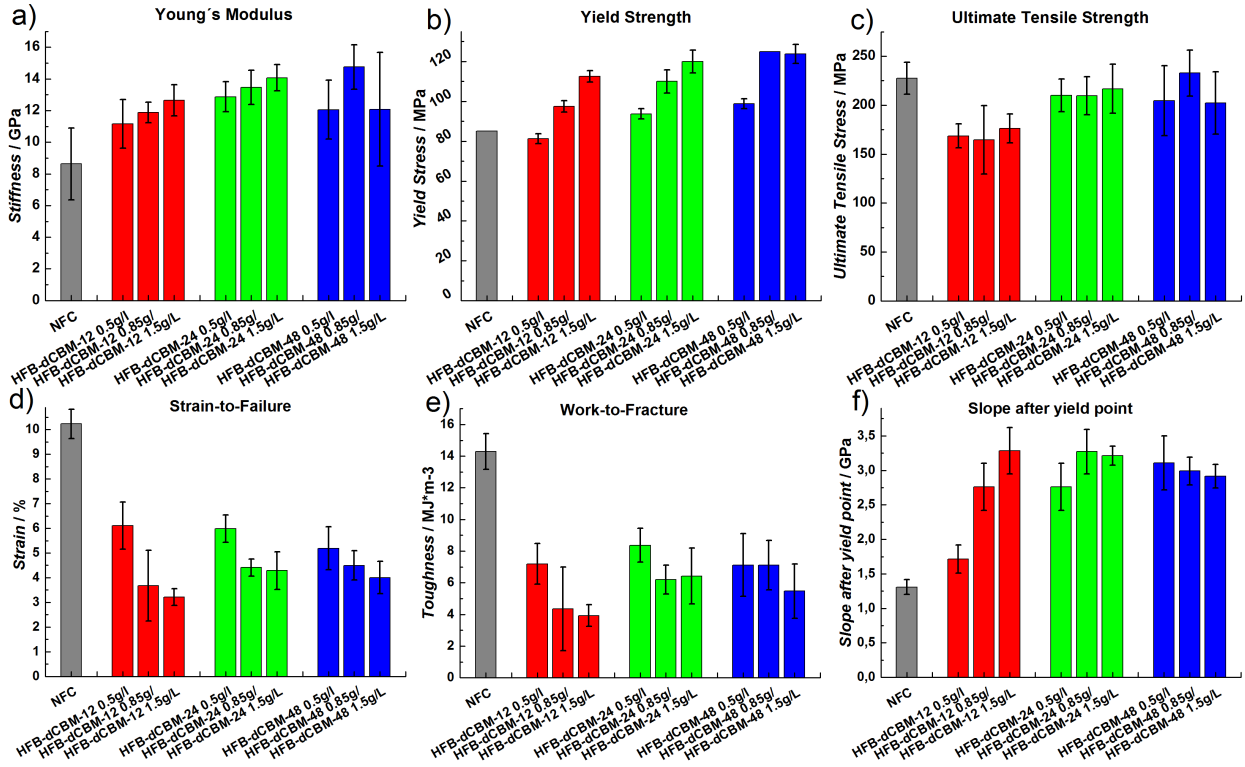

Figure S5. Summary of mechanical properties at different protein load.

Mechanical properties of three hybrids in three protein concentrations (0.5gl^{-1}, 0.85gl^{-1} and 1.5gl^{-1}) and an unmodified NFC film as a reference. a) Young´s moduli, b) yield strengths, c) ultimate tensile strengths, d) strain-to-failure e) toughness and f) slopes after yield point with standard deviations. Grey bars represent unmodified NFC films, red bars HFB-dCBD-12-NFC films, green bars HFB-dCBM-24-NFC films and blue bars HFB-dCBM-48-NFC films.

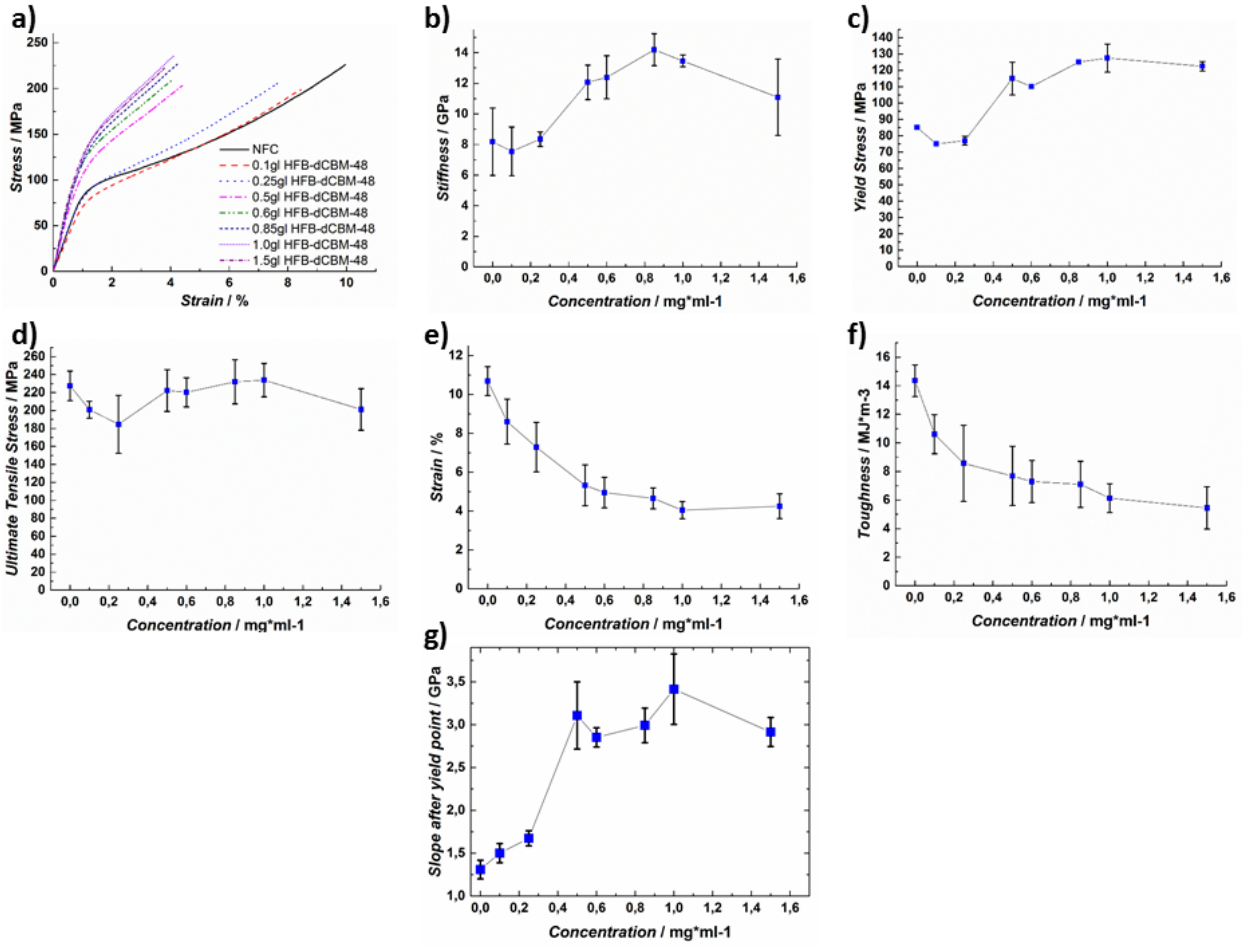

Figure S6 Mechanical properties of HFB-dCBM-48 as a function of concentration.
a) Representative stress-strain curves. b) Young´s moduli, e) yield strength, d) Ultimate tensile strength, e) strain-to-failure. f) work-to-fracture and g) slope after yield point. All data is presented with standard deviation.

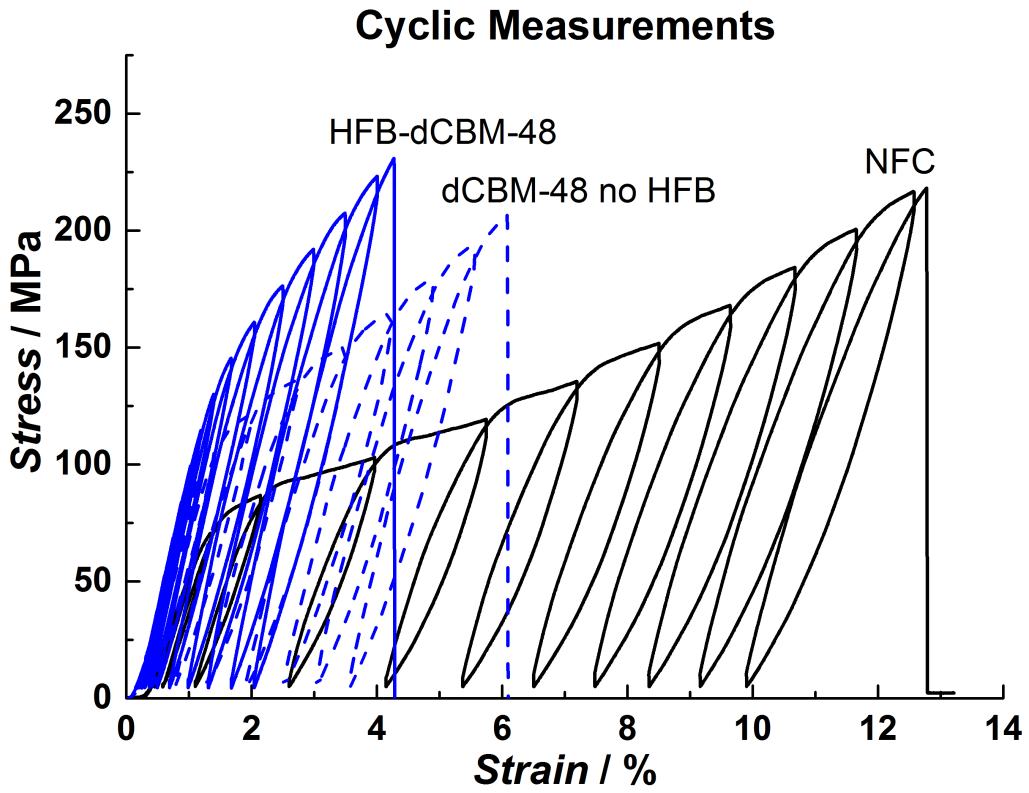

Figure S7. Representative stress-strain curves for HFB-dCBM-48 (solid blue line), dCBM-48 (dashed blue line) and unmodified NFC film (solid black line) from the cyclic tensile measurements. The results show that deformations after the yield point were not recovered after stress release and therefore the deformations were plastic but as the Young's modulus did not change, it is indicated that displaced interactions could reform after plastic deformation.

[1] M. Pääkkö, M. Ankerfors, H. Kosonen, A. Nykänen, S. Ahola, M. Österberg, J. Ruokolainen, J. Laine, P. T. Larsson, O. Ikkala, T. Lindström, *Biomacromolecules* **2007**, *8*, 1934-1941.

[2] J. Hakanpää, G. R. Szilvay, H. Kaljunen, M. Maksimainen, M. Linder, J. Rouvinen, *Protein Science* **2006**, *15*, 2129-2140.

[3] M. Linder, I. Salovuori, L. Ruohonen, T. T. Teeri, *Journal of Biological Chemistry* **1996**, *271*, 21268-21272.

[4] M. B. Linder, M. Qiao, F. Laumen, K. Selber, T. Hyytiä, T. Nakari-Setälä, M. E. Penttilä, *Biochemistry* **2004**, *43*, 11873-11882.

[5] J. Phan, A. Zdanov, A. G. Evdokimov, J. E. Tropea, H. K. Peters, R. B. Kapust, M. Li, A. Wlodawer, D. S. Waugh, *Journal of Biological Chemistry* **2002**, *277*, 50564-50572.

[6] A. Sarrion-Perdigones, E. E. Falconi, S. I. Zandalinas, P. Juárez, A. Fernández-del-Carmen, A. Granell, D. Orzaez, *PLoS ONE* **2011**, *6*, e21622.

[7] M. Linder, K. Selber, T. Nakari-Setälä, M. Qiao, M. R. Kula, M. Penttilä, *Biomacromolecules* **2001**, *2*, 511-517.