Voutilainen, S.; Paananen, Arja; Lille, Martina; Linder, Markus B.

Modular protein architectures for pH-dependent interactions and switchable assembly of nanocellulose

Published in:
International Journal of Biological Macromolecules

DOI:
10.1016/j.ijbiomac.2019.06.227

Published: 15/09/2019

Document Version
Publisher's PDF, also known as Version of record

International Journal of Biological Macromolecules 137 (2019) 270–276

Contents lists available at ScienceDirect

International Journal of Biological Macromolecules

journal homepage: http://www.elsevier.com/locate/ijbiomac

Modular protein architectures for pH-dependent interactions and switchable assembly of nanocellulose

Sanni Voutilainen [a,b], Arja Paananen [b], Martina Lille [b], Markus B. Linder [a,*]

[a] Department of Bioproducts and Biosystems, School of Chemical Engineering, Aalto University, Box 16100, 00076, Aalto, Espoo, Finland
[b] VTT Technical Research Centre of Finland Ltd., P.O. Box 1000, 02044 VTT Espoo, Finland

ARTICLE INFO

Article history:
Received 17 April 2019
Received in revised form 17 June 2019
Accepted 28 June 2019
Available online 29 June 2019

ABSTRACT

Protein engineering shows a wide range of possibilities for designing properties in novel materials. Following inspiration from natural systems we have studied how combinations or duplications of protein modules can be used to engineer their interactions and achieve functional properties. Here we used cellulose binding modules (CBM) coupled to spider silk N-terminal domains that dimerize in a pH-sensitive manner. We showed how the pH-sensitive switching into dimers affected cellulose binding affinity in relation to covalent coupling between CBMs. Finally, we showed how the pH-sensitive coupling could be used to assemble cellulose nanofibers in a dynamic pH-dependent way. The work shows how novel proteins can be designed by linking functional domains from widely different sources and thereby achieve new functions in the self-assembly of nanoscale materials.

1. Introduction

Biological materials are gaining attention among material scientists due to the wide range of possibilities for tuning their properties and integrating functionality [1–3]. The three main groups of biopolymers. i.e. proteins, polynucleotides, and carbohydrates, all have their own characteristics which are advantageous in different contexts. Carbohydrates, as for example cellulose and chitin are widely available and have excellent mechanical properties [4,5]. Polynucleotides (RNA and DNA) are superior in the storage and retrieval of information, and can be used for programming intricate structures in what is known as DNA-origami [6]. Proteins show a high versatility in binding interactions, molecular recognition, and switchable functions. One of the characteristics of proteins is how they often have evolved in natural systems by duplicating parts and recombining functionalities in a modular way [7]. Structural or specific binding elements are combined with molecular switches to create overall architectures that can perform complicated functions. Evolution has finetuned the interplay of the functional elements into astoundingly well-performing systems [8].

In this work we created proteins with novel functional architectures by combining the high affinity cellulose binding of proteins called cellulose binding modules (CBMs) and the reversibly pH-dependent dimerization function of the N-terminal parts of spidroin proteins. The new fusion-proteins were then used to make reversibly pH-switching assemblies with cellulose. The type of cellulose used in this study was a highly dispersed form, called cellulose nanofibrils (CNF).

CBMs are non-catalytic protein modules that are found as parts of enzymes acting on lignocellulosic substrates, such as glycoside hydrolases. The CBMs have a biological function to bind specifically to cellulose. Often the CBMs are connected to the catalytically active parts through extended linkers. The binding specificity of a binding module is determined by the binding site topology. The majority of CBMs that bind crystalline cellulose have flat hydrophobic binding faces containing aromatic residues [9–11]. CBMs can be produced by recombinant DNA technology as independent domains fused to other proteins and have found uses in applications such as immobilization of antibodies [12] or enzymes [13] to cellulose, and even as components of protein-based adhesives for cellulose [14].

Spider dragline silk that is produced by orb-weaving spiders is a unique fibrous material composed of proteins called spidroins. Spidroins have complex structures, consisting of three functionally and structurally different sections. The core of the spidroin contains highly repetitive sequences of poly-Ala stretches that are interrupted by Gly, Pro, Gln-rich regions. This part comprises the bulk of the spidroin and forms extensive β-sheet structures in the final assembled material [15]. At both ends of the repetitive core the spidroin has two conserved globular modules, one called the N-terminal module and the other the C-terminal module. The terminal modules are not structurally related and they have different tasks [16]. Both N- and C-terminal modules have a role when spidroin protein molecules are being assembled to form silk fibers. The C-terminal module forms covalent dimers and has been proposed to trigger the linkage of spidroins by unfolding into a

* Corresponding author.
E-mail address: markus.linder@aalto.fi (M.B. Linder).

https://doi.org/10.1016/j.ijbiomac.2019.06.227
0141-8130/

β-sheet amyloid form [17]. The N-terminal module forms a homodimer and firmly interconnects the spidroins in a pH-dependent way when the pH in the spider silk gland is lowered from pH 7 to pH 6–5.5. The 3D structure of the N-terminal module of major ampullate spidroin from the nursery web spider *Euprosthenops australis* has been solved [18]. The function of this NT module has been extensively studied through mutagenesis and the amino acids involved in the dimerization have been identified [19].

CNF is a form of cellulose that is typically obtained by disintegrating wood pulp by grinding and homogenizing, sometimes in the presence of enzymes [20]. This results in well dispersed fibrils with a crystalline packing of cellulose chains. They have a high aspect ratio with a diameter of around 5 nm and lengths of several micrometers. The high aspect ratio results in physical properties such as gelling by entanglement and shear thinning behavior. They can be assembled into for example fibers, films, or porous foams. These properties together with the possibility to functionalize NFC with other nanomaterials, chemically or through biotechnological means make NFC highly useful for functional materials [5]. The surface of CNF is relatively inert as cellulose is composed only of repeating glucose linked by β-1,4-glycosidic bonds. The surface shows only hydroxyl groups that are not very reactive. For functionalization, the hydroxyl groups are often oxidized, typically to carboxyl groups. These can then be further reacted and modified with polymers, proteins, or other chemical functionalities [21,22]. CNF gels and other forms of nanocellulose have found multiple uses as for example carriers of drug molecules or as three dimensional scaffolds for growing stem cells [23,24]. The use of CBMs for functionalizing cellulose has the advantage that no covalent modifications need to be made since the binding of CBM is spontaneous. The ability to produce CBMs as fusion proteins in which other functional proteins are linked has the advantage that no coupling chemistry has to be done to the protein either, and thus the CBM coupling is quick and efficient [12,25].

For our modular proteins we used a CBM from the *Clostridium thermocellum* cellulosome scaffoldin protein CipA and the spider silk N-terminal (NT) module from the *E. australis* major ampullate spidroin. This resulted in the fusion protein NT + CBM_{CipA} that we studied for functionalization of CNF. The idea was to utilize the pH-driven switching of the N-terminal module when the other end of the fusion protein is attached to cellulose through the CBM_{CipA} for pH dependent assembly of NFC. The binding properties of the CBM_{CipA} on nanocellulose were first determined for recombinantly produced single CBM_{CipA} and double-CBM_{CipA} proteins. This allowed an extensive understanding of the properties of spider silk N-terminal module containing CBM-fusion proteins. Finally, the effect of dimerization of the spider silk module was verified by rheological characterization of a CNF and NT + CBM_{CipA} mixture while slowly changing the pH.

2. Materials and methods

2.1. Materials

Restriction enzymes were purchased from New England Biolabs and ligase from Promega. The plasmid pET28a(+) was purchased from Novagen. Site directed mutagenesis was performed with a Phusion site-directed mutagenesis kit. DNA plasmid purification was done with Nucleospin plasmid purification kits. *E. coli* strains XL1-blue (Stratagene) and BL21 (DE3) (Novagen) were used as cloning and as expression hosts, respectively. EnPresso B (BioSilta) growth medium was used for protein production. Other chemicals were from Sigma Aldrich unless otherwise stated. Nanofibrillated cellulose (NFC) was prepared as described earlier [20].

2.2. Design of fusion proteins

The fusion proteins were designed to contain the N-terminal (NT) module of *E. australis* major ampullate spidroin 1 (MaSp1) [19] and the cellulose-binding module (CBM) from *C. thermocellum* cellulosome scaffoldin subunit CipA, CBM_{CipA} (Uniprot Q06851) [9] connected by a linker region PTPTPTTPTPTPTTPTPTPTS obtained from the *C. fimi* xylanase. A schematic illustration of the NT + CBM_{CipA} construct is shown in Fig. 1. The synthetized genes contained *Bsa*I restriction sites in their 5′ and 3′ ends. The multiple cloning site of pET28a(+) expression vector was modified to include two *Bsa*I restriction sites by inserting an adapter constructed from oligonucleotides 5′-CATGGG GAGACCGCGGATCCGAATTCGGGTCTC-3′ and 5′-TCGAGAGACCCGAATT CGGATCCGCGGTCTCCC-3′ into *Nco*I and *Xho*I restriction sites. The fusion proteins were constructed from pieces coding for NT, CBM_{CipA} and linker regions by utilizing the Golden gate method [26]. The genes were codon optimized for *E. coli* and synthetized by GenScript. The plasmids were transformed to *E. coli* XL1-blue –strain and plated on kanamycin (50μgmL^{-1}) containing LB-plates. The single CBM_{CipA} was cloned into the *Bsa*I sites alone. All the expressed proteins contained a 6-His-tag in the C-terminal end.

The same linker peptide was also used to connect two CBM_{CipA} proteins to form a double-CBM. An inactive variant, NT_{inact} + CBM_{CipA}, was made by site directed mutagenesis by replacing the Ala72 residue in NT by Arg to prevent the formation of a stable NT dimer [19]. In all, four different proteins were made, the single CBM_{CipA}, the double-CBM_{CipA}, the

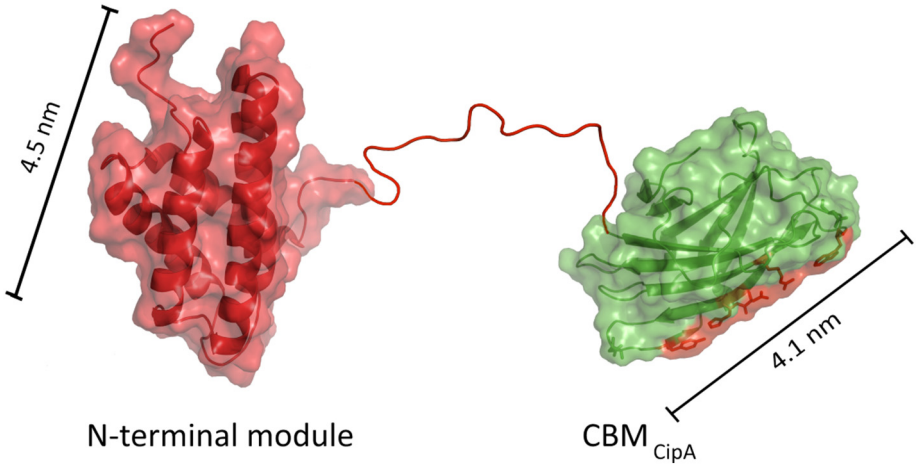

Fig. 1. Schematic representation of NT + CBM_{CipA} (PDB id. for N-terminal module and CBM_{CipA} are 2LPJ and 1NBC, respectively). The aromatic and charged residues of CBM_{CipA} forming the cellulose-binding face are marked in red. The bars are showing the approximated sizes of the molecules. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

NT + CBM_{CipA}, and the NT_{inact} + CBM_{CipA}. The full protein sequences are found in the supporting information.

2.3. Production and purification of the CBM and fusion proteins

The recombinant proteins were expressed in BL21(DE3) *E. coli* strain by cultivating the strains carrying the expression vectors in EnPresso B media (200 mL cultivations in 2 L Erlenmeyer flasks) according to the manufacturer's instructions. The protein production was induced by 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG, Roche Diagnostics). After 24 h induction the cells were collected by centrifugation at 3900*g* for 20 min. The cells were lysed by first freezing and thawing them once and then re-suspending them in 10 mL of 50 mM Tris-HCl pH 7.5 containing 1 gL^{-1} of lysozyme from chicken egg white, 2.5 mM $MgCl_2$ and DNAse 1 (Roche Diagnostics) in the presence of protease inhibitor cocktail (Complete, EDTA-free, Roche Diagnostics). The lysis reaction was carried out in room temperature for 2 to 3 h after which the lysis was completed by sonication with 0.5 s cycles for 3 min. The cell lysate was centrifuged for 30 min, 3900*g* and the supernatant was applied onto 5 mL His-Trap crude column (GE Healthcare) equilibrated with 20 mM sodium phosphate buffer pH 7.4 containing 0.5 M NaCl and 20 mM imidazole. After washing with the equilibration buffer, the bound proteins were eluted with a linear gradient of imidazole from 20 mM to 500 mM in the equilibration buffer. Fractions were analyzed on SDS–PAGE and those containing pure protein were pooled and the buffer was exchanged to 50 mM sodium phosphate buffer pH 7.0 using Econo-Pac columns (Bio-Rad). The production levels of all variants were excellent; the protein yield after purification varied from 400 to 500 mg from 1 L of culture. Purified proteins are shown on SDS-PAGE (18%) stained with Coomassie blue (Fig. 2).

2.4. Protein concentration determination

The protein concentrations of the purified proteins were determined by absorbance at 280 nm and calculated from the theoretical molar extinction coefficients calculated from the primary amino acid sequence [27]. CBM_{CipA} ε = 35140 M^{-1} cm^{-1}, NT + CBM_{CipA} and NT_{inact} + CBM_{CipA} ε = 40910 M^{-1} cm^{-1}, and double-CBM_{CipA} ε = 70820 M^{-1} cm^{-1}.

2.5. Size exclusion chromatography

Size exclusion chromatography (SEC) was performed with Äktapure system (Ge Healthcare) using Superdex 200 GL 10/300 column. The column was equilibrated with one of the following buffers: 50 mM citrate pH 3 or pH 3.5, 50 mM sodium acetate buffer pH 4, 4.5, 5 or 5.5, 50 mM sodium phosphate buffer pH 6, 6.5 or 7. All buffers contained 150 mM NaCl. Each sample was diluted in the equilibration buffer of the current run to protein concentration of about 0.5 to 0.7 gL^{-1} and 250 μL was injected into the column. The elution (0.5 $mLmin^{-1}$) was followed with a UV detector at 280 nm. SEC elution data were calibrated with the Gel Filtration Calibration Kit LMW (Ge Healthcare).

2.6. Protein labelling with ^3H

The single CBM_{CipA} and fusion proteins were labelled with ^3H by reductive methylation essentially as described earlier [28]. The buffer containing the proteins was changed to 200 mM borate (pH 8.5) using Econo-Pac gel-filtration columns. The concentration of amines (Lys and N-terminus) of each protein was calculated and 2.5 mg of each protein were mixed with formaldehyde in 5 times molar excess to the amine concentration. Tritium-enriched $NaBH_4$ (10.3 Ci $mmol^{-1}$; 100 mCi; Perkin-Elmer) was dissolved in 0.1 M NaOH and added to the protein sample. The reactions were incubated 30 min on ice and terminated by applying the sample onto an Econo-Pac gel filtration column. The protein was eluted with 50 mM sodium phosphate pH 7.

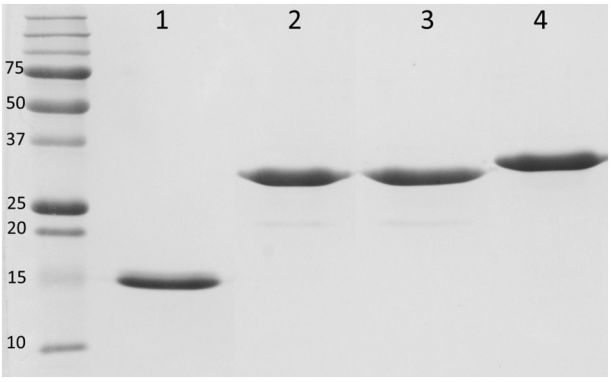

Fig. 2. SDS-PAGE analysis (18% gel) of the purified proteins. Lane 1: single CBM_{CipA}, lane 2: NT + CBM_{CipA}, lane 3: NT_{inact} + CBM_{CipA}, lane 4: double-CBM_{CipA}, leftmost lane: molecular weight marker.

2.7. Binding to NFC

Binding experiments were done in three different pH values using 50 mM sodium citrate pH 3.5, 50 mM sodium acetate pH 5.5, and 50 mM sodium phosphate pH 7 buffers. All buffers and protein dilutions contained 0.5% bovine serum albumin (BSA) to prevent non-specific binding. The initial protein concentrations ranged from 0.6 μM to 50 μM. Equal volumes (100 μL) of protein and NFC suspension (2 gL^{-1}) were dispensed in 24-well plates (Life Technologies) and mixed on orbital shaker in room temperature. After 60 min incubation, samples were filtered through Millex 0.22 μm GV13 filters (Millipore). The amount of ^3H in the filtrate was quantified by liquid scintillation counting using Ultima Gold XR (PerkinElmer) scintillation counting liquid and a Tri-Carb 2810TR (PerkinElmer) scintillation counter. The amount of bound enzyme was calculated by subtracting the measured free protein from the initial protein concentration. The reversibility of binding was tested by dilution experiments. A series of identical reaction mixtures were incubated for 60 min in order to obtain equilibrium (in pH 5.5). Then a 5-time excess of acetate buffer (0.5% BSA) without

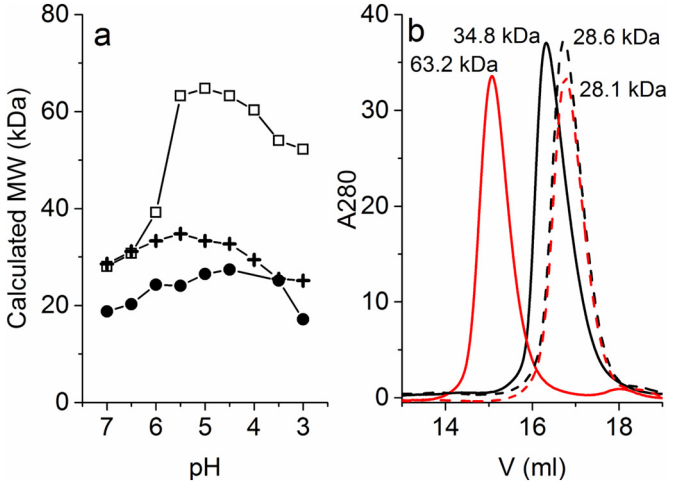

Fig. 3. Monomer/dimer analysis of NT + CBM fusion proteins. a) Molecular weights at different pH values calculated from size-exclusion chromatography; NT + CBM_{CipA} (□), NT_{inact} + CBM_{CipA} (+) and double-CBM_{CipA} (●) are shown. b) Size-exclusion chromatograms of NT + CBM_{CipA} at two different pH values; NT + CBM_{CipA} pH 5.5 (red), NT_{inact} + CBM_{CipA} pH 5.5 (black), NT + CBM_{CipA} pH 7 (dashed line, red), NT_{inact} + CBM_{CipA} pH 7 (dashed line, black), including the calculated molecular weights. The theoretical molecular weights calculated from the amino acid sequences are: NT + CBM_{CipA} = 34.6 kDa, NT_{inact} + CBM_{CipA} = 34.7 kDa, double-CBM_{CipA} = 37.9 kDa. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

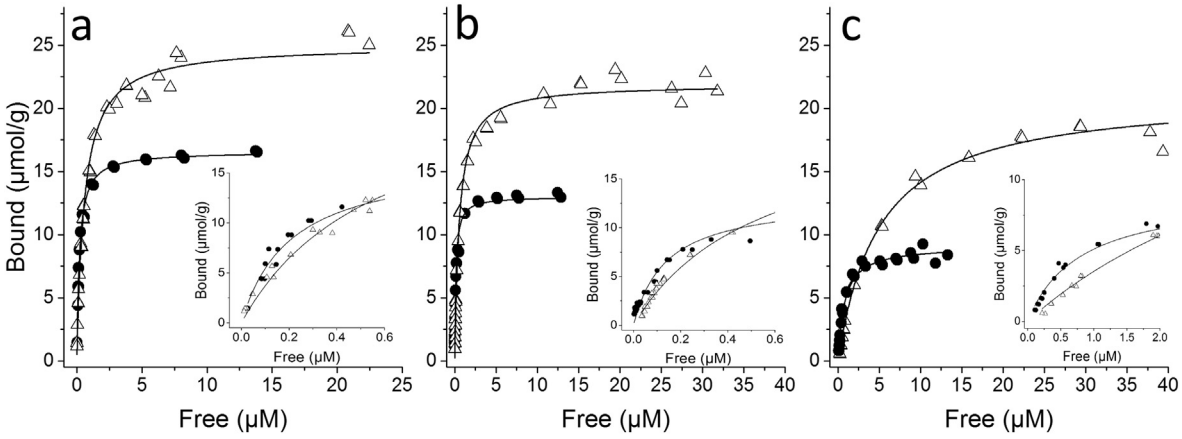

Fig. 4. Comparison of the binding isotherms of CBM_{CipA} (Δ) and double-CBM_{CipA} (●) in different pH values. a) pH 7, b) pH 5.5 and c) pH 3.5. The same isotherms with closer zoom to the low protein concentrations are shown in the inserts.

CBM_{CipA} or NT + CBM_{CipA} was added. After incubation for different times (15 s to 20 h), the mixtures were filtered, and the amounts of released protein were measured. Data were analyzed using the software Origin (Microcal).

2.8. Rheology

The pH dependent self-assembly of CNF and the protein variants NT + CBM_{CipA} and NT_{inact} + CBM_{CipA} were followed by rheological measurements. Samples were prepared by gently mixing 2 gL^{-1} CNF and 40 μM protein in 5 mM sodium phosphate buffer pH 7 in plastic tubes. The sample volume was 5 mL. Glucono-δ-lactone (GDL, Sigma Life Science) was added as a solid powder to the CNF-protein mixture (time point = 0, amount corresponds 0.4% final concentration) to change the sample pH gradually with time. Each sample was divided in two parts, one for pH measurement and the other for rheology. Rheological measurements were carried out at room temperature (22 °C) with a rheometer (AR-G2, TA instruments, UK) equipped with cross-hatched plate-plate geometry (diameter 40 mm). The viscoelastic properties were determined in small deformation oscillatory mode using 2.05 mL sample and a 1 mm gap. Time sweeps (frequency 0.1 Hz, strain 1%) were run for 2 h. Evaporation was prevented by using a cover. The pH profiles were measured simultaneously as the rheological measurements using a pH meter with a data logger (pH 3310, WTW). The rheological and pH measurement were synchronized. Results were plotted according to corresponding time points. As a reference, 2 gL^{-1} CNF in 5 mM sodium phosphate buffer pH 7 without protein was used.

3. Results and discussion

All fusion proteins, the single CBM_{CipA}, the double-CBM_{CipA}, the wild type spidroin NT-domain fused to the CBM (NT + CBM_{CipA}), and the reference with a mutated and inactive spidroin NT variant (NT_{inact} + CBM_{CipA}), were produced with the *E. coli* expression system and were readily purified (Fig. 2).

Initially we studied the pH dependent functionality of the two NT fusion proteins by SEC. At pH 7, the NT + CBM_{CipA} protein eluted with a mass corresponding to its monomeric form (Fig. 3a and b). Decreasing the pH to 6 led to dimerization and at pH 5.5 the protein was fully dimerized. This result supports previous reports of dimerization of the spidroin NT [18], and shows that the linkage to the CBM_{CipA}, does not interfere with the NT dimer formation. As expected, the inactive mutant Ala72Arg (NT_{inact} + CBM_{CipA}) did not show dimerization at the same pH-range. The Ala72 is located in the dimer interface and is in contact with the same residue of the other partner of the homodimer across the interface. When Ala72 was replaced by a bulky Arg residue, dimerization was prevented [19]. This mutant protein NT_{inact} + CBM_{CipA} was utilized as a negative control protein for the experiments.

We next determined the binding affinity and capacity of the isolated CBM_{CipA} and the double-CBM_{CipA} at pH 7 (Fig. 4A). The parameters of binding were obtained by fitting a one-site binding model (Eq. (1), where B is the amount of adsorbed protein, B_{max} is the maximum adsorbed protein, K_D is dissociation constant, and c is the concentration of free protein) to the data and are summarized in Table 1.

$$B = \frac{B_{max} * c}{K_D + c} \quad (1)$$

We noted that the linkage of two CBM_{CipA}-proteins to form the double-CBM_{CipA} affected the binding by increasing affinity K_D from 0.57 μM to 0.2 μM and lowering capacity (from 25 to 16.7 μmol g^{-1}). The measured value K_D for the isolated CBM_{CipA} fits very well with the previously reported K_D of 0.5 ± 0.2 μM [29] and 0.4 μM [30]. One-site Langmuir isotherms describe the binding data very well suggesting a single type of binding interaction. This becomes especially clear at high concentrations with the high-affinity double-CBM_{CipA}. In Fig. 4a, the insert shows the details of binding at low concentration, where the higher affinity of the double-CBM_{CipA} is clearly visible. Another way of comparing affinities is through the partitioning coefficient K_r, i.e., initial slopes of the curves (i.e. B_{max}/K_D). This value is the slope of the binding curve as the protein

Table 1
Affinity and capacity parameters obtained from binding isotherm fitting. Errors in B_{max} and K_D are standard errors non-linear curve-fit. B_{max} is the amount of adsorbed protein in saturation, K_D is dissociation constant. K_r is the partitioning coefficient which is defined as the ratio of B_{max} and K_D.

	pH 7			pH 5.5			pH 3.5		
	K_D (μM)	B_{max} (μmol g^{-1})	K_r (Lg^{-1})	K_D (μM)	B_{max} (μmol g^{-1})	K_r (Lg^{-1})	K_D (μM)	B_{max} (μmol g^{-1})	K_r (Lg^{-1})
CBM_{CipA}	0.57 ± 0.03	25.0 ± 0.4	**43.9**	0.54 ± 0.02	21.9 ± 0.2	**40.6**	5.0 ± 0.4	21.2 ± 0.4	**4.2**
Double-CBM_{CipA}	0.20 ± 0.01	16.7 ± 0.2	**83.5**	0.14 ± 0.01	13.0 ± 0.2	**92.9**	0.79 ± 0.05	9.1 ± 0.2	**11.5**
NT + CBM_{CipA}	2.7 ± 0.3	24.7 ± 0.9	**9.1**	1.9 ± 0.1	31.8 ± 0.5	**16.7**	1.3 ± 0.1	13.5 ± 0.2	**10.4**
NT_{inact} + CBM_{CipA}	3.4 ± 0.2	25.5 ± 0.5	**7.8**	3.3 ± 0.3	31.8 ± 1	**10.0**	4.7 ± 0.4	15.7 ± 0.6	**2.95**

Bold numbers signifies standard deviation.

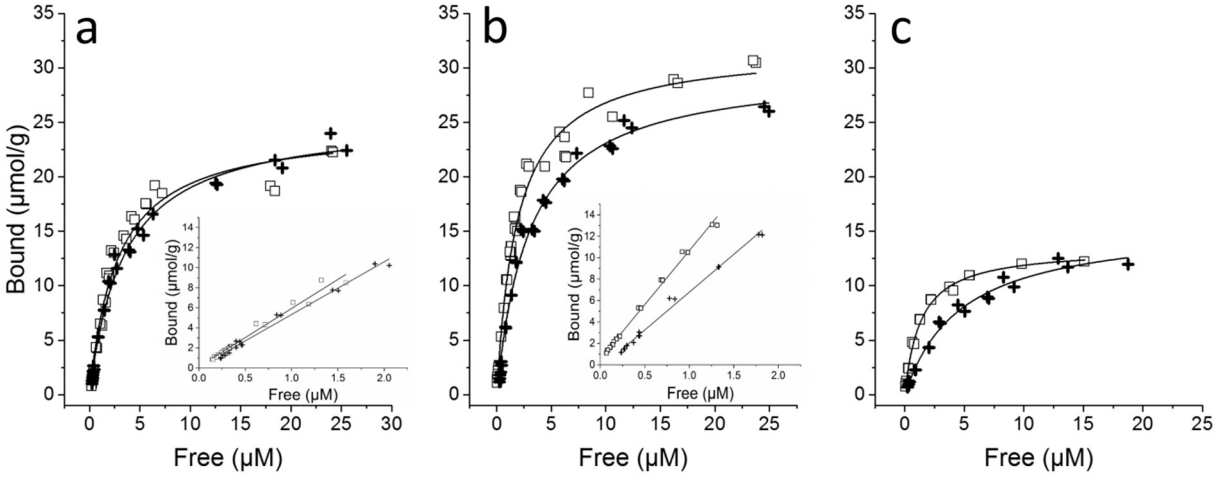

Fig. 5. Comparison of the binding isotherms of NT + CBM_{CipA} (□) and NT_{inact} + CBM_{CipA} (+) in different pH values. a) pH 7, b) pH 5.5 and c) pH 3.5. The same isotherms with closer zoom to the low protein concentrations are shown in the inserts.

concentration approaches zero [31]. The partitioning coefficient is analogous to the specificity constant commonly used in Michaelis-Menten analysis (k_{cat}/K_M, where k_{cat} is the turnover number and K_M is the Michaelis constant) [32]. Measured by K_r, the double-linkage resulted in an increased initial affinity from 43.9 to 83.5 Lg^{-1}. Double-linkage thus results in a clearly higher affinity and also a lower capacity due to its larger size. This increase is expected due to a colocalization effect [7,33]. The double-CBM_{CipA} binds first by one of its CBMs, and due to the colocalization of the second one, its local concentration increases, and it binds readily as also previously described [34].

The pH had a small effect on the CBM_{CipA} binding going from pH 7 to pH 5.5, but more notable when going as low as pH 3.5 (Fig. 4b and c, Table 1). The exact reason for the decrease of binding is not clear, but we note that the binding face of CBM_{CipA} does contain a His-residue, and that in engineered CBMs it has been shown that His residues lead to lower affinities at low pH due to their protonation [9,35]. The ratio of K_r values (i.e. the K_r for the double-CBM_{CipA} divided by the K_r for the single CBM_{CipA}) remained very similar at around 2 regardless of the pH (Fig. 4b and c, inserts, Table 1) That the ratio of K_r values does not change indicates that colocalization effects are not affected by pH [31].

Adding the N-terminal spidroin domain resulted in an overall decreased affinity (Fig. 5, Table 1), with the K_r dropping to 2.7 Lg^{-1} for the WT N-terminal domain and to 3.4 Lg^{-1} for the inactive N-terminal domain mutant. B_{max} values show that the drop was due to a reduction in affinity, not on a change in capacity. The reason for the decrease in affinity in the fusion construct may be due to a steric hindrance. The fact that the inactive mutant and the wild type have exactly the same binding at pH 7 where no dimerization of either occurs, shows that the effect is due to the presence of the N-terminal domain. This indicates that the steric hindrance caused by the presence of the N-terminal domain results in the decreased binding.

The two variants with wild-type NT and the inactivate mutant behave very differently when the pH decreased. At pH 3.5, the inactive mutant had a low K_r of 2.95 Lg^{-1} while the wild type had a K_r of 10.4 Lg^{-1}. This large increase of the NT + CBM_{CipA}, which followed a different trend comparing to all other CBM variants is expected due to the dimerization of the NT domain at low pH, resulting in a colocalization effect and thereby increased affinity. The non-dimerizing mutant, on the other hand, would not be expected to show colocalization at any pH.

For exploring the magnitude of the colocalization effects on binding we calculated the change in free energy of binding for the different variants using Eq. (2)Table 2.

$$\Delta G = RTlnK_D \quad (2)$$

We note that the cooperativity in case of the double construct (double-CBM_{CipA}) does not fully correspond to the sum of free energies of the single CBMs. This means that although there is a cooperativity in binding, it corresponds to less than the maximum possible. This observation is perfectly in line with studies on other cooperative systems, and as discussed in previously [33]. The likely reason for this commonly observed effect is that there is an energetic cost for orientation in binding and that the rearrangements of the geometries of the proteins is required. On the other hand, the observed cooperativity shows that the binding sites of CBMs on the cellulose surface are sufficiently close to each other for simultaneous binding of both domains, although the geometry of the binding sites on cellulose may require a precise positioning of the two CBMs in relation to each other. A precise positioning of CBMs implies higher structural order, and hence a higher entropy cost of binding. This would reduce the free energy available for the cooperative binding, which is in line with our observations. The free energy change during binding of double-CBM_{CipA} is slightly higher that for NT + CBM_{CipA} at pH 3 indicating a greater steric hindrance due to the presence of the NT. Looking at binding capacities (Table 1), we note that at pH 3, the binding capacity of double-CBM_{CipA} close to half of that of CBM_{CipA}. Because this is on a molar basis, it suggests that both CBMs of double-CBM_{CipA} are bound to the cellulose. The binding capacity of NT_{inact} + CBM_{CipA} is slightly higher than for NT + CBM_{CipA} suggesting that the dimerization interaction leads to a need for more space on the cellulose surface when the NT + CBM_{CipA} binds compared to the inactive mutant.

Another interesting property of CBM_{CipA} binding is if the protein shows only a very slow desorption from the cellulose surface one it has bound. This property was studied using experiments were a set of identical tubes where CBM_{CipA} and cellulose were mixed. Then buffer was added to some tubes and the bound amount in all tubes was measured. If the binding is reversible, the dilution would result in a lowering of bound amount according to Eq. (1) [36]. The results showed that

Table 2
Free energies of binding at different pH for all protein variants.

	pH 7	pH 5.5	pH 3.5
	ΔG	ΔG	ΔG
	$kJmol^{-1}$	$kJmol^{-1}$	$kJmol^{-1}$
CBM_{CipA}	−35.03	−35.16	−29.74
Double-CBM_{CipA}	−37.58	−38.45	−34.24
NT + CBM_{CipA}	−31.24	−32.10	−33.02
NT_{inact} + CBM_{CipA}	−30.68	−30.75	−29.89

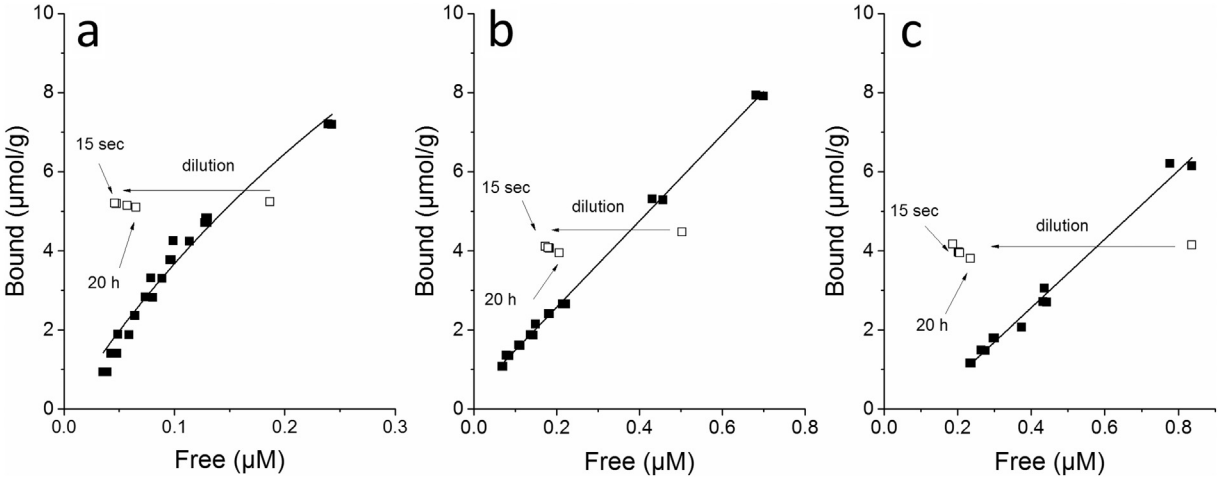

Fig. 6. Reversibility of CBM_{CipA} by dilution with buffer at pH 5.5, 22 °C. Reversible protein would return to a point on the isotherm, whereas CBM_{CipA} did not completely return to the isotherm during 20 h incubation. A) Single CBM_{CipA}, B) NT + CBM_{CipA}, C) NT_{inact} + CBM_{CipA}.

desorption happened only very slowly, around 10% in 20 h (Fig. 6). The NT domain linkages did not affect desorption rates.

Having demonstrated that the silk N-terminal domain functioned to dimerize the CBM-fusion proteins in a pH dependent manner, we continued to investigate how this pH sensitive switch could be used in the molecular design for novel materials functions. We studied how a gradual pH change would affect the viscoelastic properties of a network of NFC in the presence of the NT + CBM_{CipA}. As a control we used the NT_{inact} + CBM_{CipA} containing the inactive mutant. The pH change was accomplished by adding a lactone that slowly decomposes through hydrolysis when diluted into water. Through the carboxylic acid formed by the hydrolysis, the pH slowly decreased in the sample. This approach has the advantage that no additions of substances were needed during the experiment. The decrease in pH resulted in a dimerization of NT + CBM_{CipA}, but not NT_{inact} + CBM_{CipA} (Fig. 3), Accordingly, in the sample containing CNF and the functional NT + CBM_{CipA} an immediate increase in storage modulus was noted while the pH decreased, reaching an optimum at around pH 5.5 (Fig. 7). The control samples with either no added protein or with the inactive NT_{inact} + CBM_{CipA} mutant did not change the storage modulus appreciably. We therefore find that the dimerization of the N-terminal domain leads to crosslinking of the NFC fibrils. The pH-sensitive cross-linking of the NT in combination with the CBM binding to CNF could thus be used to modulate viscoelastic properties in a pH sensitive manner. The formation of a maximum peak at around pH 5.5 and the decrease of G′ at lower pH, where NT dimer formation is still strong, is probably due to the overall lower binding of the CBM_{CipA} at low pH (Fig. 5c). We could therefore demonstrate a silk-like assembly of NFC in an analogous way as in which silk proteins assemble due to a pH trigger [18]. We propose that such pH-dependent gel stiffening could be useful for applications such as switchable hydrogels for cell culture or release of drugs from NFC colloids [23,24]. On another level the work shows the great versatility for proteins as components in functional materials. Functionalities such as cellulose binding and pH-dependent binding can be combined in a straightforward way to achieve combined properties that would be very difficult to achieve by use of for example synthetic polymers or molecules. The simple dimerization of CBMs in the double-CBM_{CipA} shows that parameters such as affinity and binding capacity also can be engineered in a straightforward way and could find applications in areas such as immobilization, nanomaterials, or modification of cellulose degrading enzymes.

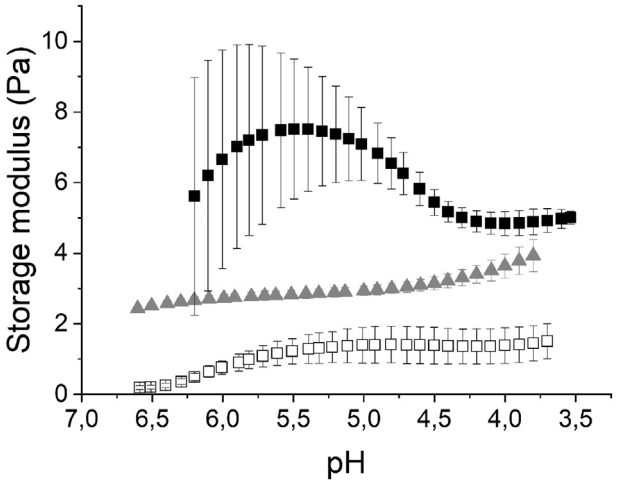

Fig. 7. The elasticity of NFC with addition of the hybrid spidroin dimerizing protein and cellulose binding modules (NT + CBM_{CipA}). The pH-dependent dimerization of the NT domain results in a macroscopic response in materials properties. Filled squares (■) represent the NT-CBM_{CipA}, open squares (□) the inactive variant NT_{inact} + CBM_{CipA}, and filled triangles (▲) the buffer control. The points are averages of three separate runs and error bars show the standard deviation.

4. Conclusions

The modularity of proteins such as spidroins and cellulolytic enzymes encourage us to explore modular architectures for building new functional biopolymers with novel combinations [37]. For material science this may be one of the most promising and versatile features of using proteins for molecular design. Functional properties from widely different sources can be combined, and resulting in proteins that can be produced as new multi-module polymers, allowing completely novel applications. Here we showed in detail how dimeric modules of CBMs interact with cellulose. The detection by tritium scintillation counting allowed highly precise binding data. Adding the pH sensitive spidroin NT-module resulted in a pH responsive mechanism for modulating protein interactions, which together with CNF gave a way to modulate viscoelastic properties of the combined protein-cellulose material. A detailed understanding of binding events and thermodynamics allowed us to understand the relation between cross-linking and cooperative binding.

Acknowledgements

The work was performed within the Academy of Finland Center of Excellence Programme (2014-2019) and Academy of Finland projects

307474, 317395 and 317019. We are grateful for the support by the FinnCERES Materials Bioeconomy Ecosystem.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijbiomac.2019.06.227.

References

[1] N.H.C.S. Silva, C. Vilela, I.M. Marrucho, C.S.R. Freire, C. Pascoal, A.J.D. Silvestre, Protein-based materials: from sources to innovative sustainable materials for biomedical applications, J. Mater. Chem. B 2 (2014) 3715–3740, https://doi.org/10.1039/c4tb00168k.

[2] M. Soikkeli, K. Kurppa, M. Kainlauri, S. Arpiainen, A. Paananen, D. Gunnarsson, J.J. Joensuu, P. Laaksonen, M. Prunnila, M.B. Linder, J. Ahopelto, Graphene biosensor programming with genetically engineered fusion protein monolayers, ACS Appl. Mater. Interfaces 8 (2016) 8257–8264, https://doi.org/10.1021/acsami.6b00123.

[3] P. Laaksonen, A. Walther, J.M. Malho, M. Kainlauri, O. Ikkala, M.B. Linder, Genetic engineering of biomimetic nanocomposites: diblock proteins, graphene, and nanofibrillated cellulose, Angew. Chem. Int. Ed. 50 (2011) 8688–8691, https://doi.org/10.1002/anie.201102973.

[4] M.A. Hubbe, O.J. Rojas, L.A. Lucia, M. Sain, T.A. Forest, Cellulosic nanocomposites: a review, BioResources. 3 (2008) 929–980.

[5] E. Kontturi, P. Laaksonen, M.B. Linder, A.H. Nonappa, A.H. Gröschel, O.J. Rojas, O. Ikkala, Advanced materials through assembly of nanocelluloses, Adv. Mater. (2018) 1703779, https://doi.org/10.1002/adma.201703779.

[6] S. Nummelin, J. Kommeri, M.A. Kostiainen, V. Linko, Evolution of structural DNA nanotechnology, Adv. Mater. 30 (2018), 1703721. https://doi.org/10.1002/adma.201703721.

[7] J. Kuriyan, D. Eisenberg, The origin of protein interactions and allostery in colocalization, Nature 450 (2007) 983–990, https://doi.org/10.1038/nature06524.

[8] F.H. Arnold, J.T. Meyerowitz, Evolving with purpose, Nature. 509 (2014) 166–167.

[9] J. Tormo, R. Lamed, A.J. Chirino, E. Morag, E.A. Bayer, Y. Shoham, T.A. Steitz, Crystal structure of a bacterial family-III cellulose-binding domain: a general mechanism for attachment to cellulose, EMBO J. 15 (1996) 5739.

[10] A.B. Boraston, D.N. Bolam, H.J. Gilbert, G.J. Davies, Carbohydrate-binding modules: fine-tuning polysaccharide recognition, Biochem. J. 382 (2004) 769–781, https://doi.org/10.1042/BJ20040892.

[11] M. Linder, T.T. Teeri, The roles and function of cellulose-binding domains, J. Biotechnol. 57 (1997) 15–28.

[12] M. Linder, T. Nevanen, L. Soderholm, O. Bengs, T.T. Teeri, Improved immobilization of fusion proteins via cellulose-binding domains, Biotechnol. Bioeng. 60 (1998)https://doi.org/10.1002/(SICI)1097-0290(19981205)60:5<642::AID-BIT15>3.0.CO;2-8.

[13] L.I. Crouch, A. Labourel, P.H. Walton, G.J. Davies, H.J. Gilbert, The contribution of non-catalytic carbohydrate binding modules to the activity of lytic polysaccharide monooxygenases, J. Biol. Chem. 291 (2016) 7439–7449, https://doi.org/10.1074/jbc.M115.702365.

[14] P. Mohammadi, G. Beaune, B.T. Stokke, J.V.I. Timonen, M.B. Linder, Self-coacervation of a silk-like protein and its use as an adhesive for cellulosic materials, ACS Macro Lett. 7 (2018) 1120–1125, https://doi.org/10.1021/acsmacrolett.8b00527.

[15] S. Keten, Z. Xu, B. Ihle, M.J. Buehler, Mechanical toughness of β -sheet crystals in silk, Nat. Mater. 9 (2010) 359–367, https://doi.org/10.1038/nmat2704.

[16] J.E. Garb, N.A. Ayoub, C.Y. Hayashi, Untangling spider silk evolution with spidroin terminal domains, BMC Evol. Biol. 10 (2010) 243.

[17] M. Andersson, G. Chen, M. Otikovs, M. Landreh, K. Nordling, N. Kronqvist, P. Westermark, H. Jörnvall, S. Knight, Y. Ridderstråle, L. Holm, Q. Meng, K. Jaudzems, M. Chesler, J. Johansson, A. Rising, Carbonic anhydrase generates CO_2 and H+ that drive spider silk formation via opposite effects on the terminal domains, PLoS Biol. 12 (2014), e1001921http://dx.plos.org/10.1371/journal.pbio.1001921.s002.

[18] G. Askarieh, M. Hedhammar, K. Nordling, A. Saenz, C. Casals, A. Rising, J. Johansson, S.D. Knight, Self-assembly of spider silk proteins is controlled by a pH-sensitive relay, Nature 465 (2010) 236–238, https://doi.org/10.1038/nature08962.

[19] K. Jaudzems, G. Askarieh, M. Landreh, K. Nordling, M. Hedhammar, H. Jörnvall, A. Rising, S.D. Knight, J. Johansson, pH - dependent dimerization of spider silk N-terminal domain requires relocation of a wedged tryptophan side chain, J. Mol. Biol. (2012) 477–487, https://doi.org/10.1016/j.jmb.2012.06.004.

[20] M. Pääkkö, M. Ankerfors, H. Kosonen, A. Nykänen, S. Ahola, M. Österberg, J. Ruokolainen, J. Laine, P.T. Larsson, O. Ikkala, T. Lindström, Enzymatic hydrolysis combined with mechanical shearing and high-pressure homogenization for nanoscale cellulose fibrils and strong gels, Biomacromolecules 8 (2007) 1934–1941http://pubs.acs.org/doi/abs/10.1021/bm061215p.

[21] T. Saito, S. Kimura, Y. Nishiyama, A. Isogai, Cellulose nanofibers prepared by TEMPO-mediated oxidation of native cellulose, Biomacromolecules 8 (2007) 2485–2491.

[22] S. Arola, T. Tammelin, A. Tullila, M.B. Linder, Immobilization — stabilization of proteins on nanofibrillated cellulose derivatives and their bioactive film formation, Biomacromolecules 13 (2012) 594–603.

[23] R. Kolakovic, L. Peltonen, A. Laukkanen, M. Hellman, P. Laaksonen, M.B. Linder, J. Hirvonen, T. Laaksonen, Evaluation of drug interactions with nanofibrillar cellulose, Eur. J. Pharm. Biopharm. 85 (2013) 1238–1244, https://doi.org/10.1016/j.ejpb.2013.05.015.

[24] Y.-R. Lou, L. Kanninen, T. Kuisma, J. Niklander, L.A. Noon, D. Burks, A. Urtti, M. Yliperttula, The use of nanofibrillar cellulose hydrogel as a flexible three-dimensional model to culture human pluripotent stem cells, Stem Cells Dev. 23 (2014) 380–392, https://doi.org/10.1089/scd.2013.0314.

[25] E. Ong, N.R. Gilkes, R.A.J. Warren, R.C. Miller, D.G. Kilburn, Enzyme immobilization using the cellulose-binding domain of a Cellulomonas Fimi exoglucanase, Bio/Technology 7 (1989) 604–607, https://doi.org/10.1038/nbt0689-604.

[26] C. Engler, R. Kandzia, S. Marillonnet, A one pot, one step, precision cloning method with high throughput capability, PLoS One 3 (2008), e3647.

[27] C.N. Pace, F. Vajdos, L. Fee, G. Grimsley, T. Gray, How to measure and predict the molar absorption coefficient of a protein, Protein Sci. 4 (1995) 2411–2423.

[28] B.F. Tack, J. Dean, D.E. Patrick, E. Lorem, A.N. Schechter, Tritium labeling of proteins to high specific radioactivity by reductive methylation, J. Biol. Chem. 255 (1980) 8842–8847.

[29] N. Georgelis, N.H. Yennawar, D.J. Cosgrove, Structural basis for entropy-driven cellulose binding by a type-a cellulose-binding module (CBM) and bacterial expansin, Proc. Natl. Acad. Sci. U. S. A. 109 (2012) 14830–14835, https://doi.org/10.1073/pnas.1213200109.

[30] E. Morag, A. Lapidot, D. Govorko, R. Lamed, M. Wilchek, E.A. Bayer, Y. Shoham, Expression, purification, and characterization of the cellulose- binding domain of the scaffoldin subunit from the cellulosome of Clostridium thermocellum, Appl. Environ. Microbiol. 61 (1995) 1980–1986.

[31] S. Arola, M.B. Linder, Binding properties of single and double cellulose binding modules reveal differences between cellulose substrate, Sci. Rep. 6 (2016) 35358, https://doi.org/10.1038/srep35358.

[32] J. Kuriyan, B. Konforti, D. Wemmer, The Molecules of Life, W.W. Norton, 2012.

[33] M. Mammen, S.-K. Choi, G.M. Whitesides, Polyvalent interactions in biological systems: implications for design and use of multivalent ligands and inhibitors, Angew. Chem. Int. Ed. 37 (1998) 2754–2794http://gmwgroup.harvard.edu/pubs/pdf/598.pdf.

[34] M. Linder, I. Salovuori, L. Ruohonen, T.T. Teeri, Characterization of a double cellulose-binding domain. Synergistic high affinity binding to crystalline cellulose, J. Biol. Chem. 271 (1996) 21268–21272.

[35] M. Linder, T. Nevanen, T.T. Teeri, Design of a pH-dependent cellulose-binding domain, FEBS Lett. 447 (1999) 13–16.

[36] G. Carrard, M. Linder, Widely different off rates of two closely related cellulose-binding domains from Trichoderma reesei, Eur. J. Biochem. 262 (1999) 637–643.

[37] P. Laaksonen, G.R. Szilvay, M.B. Linder, Genetic engineering in biomimetic composites, Trends Biotechnol. 30 (2012) 191–197http://linkinghub.elsevier.com/retrieve/pii/S0167779912000029.