Research review paper

Recombinant CBM-fusion technology — Applications overview

CrossMark

Carla Oliveira, Vera Carvalho, Lucília Domingues, Francisco M. Gama *

CEB — Centre of Biological Engineering, University of Minho, 4710-057 Braga, Portugal

ABSTRACT

Carbohydrate-binding modules (CBMs) are small components of several enzymes, which present an independent fold and function, and specific carbohydrate-binding activity. Their major function is to bind the enzyme to the substrate enhancing its catalytic activity, especially in the case of insoluble substrates. The immense diversity of CBMs, together with their unique properties, has long raised their attention for many biotechnological applications. Recombinant DNA technology has been used for cloning and characterizing new CBMs. In addition, it has been employed to improve the purity and availability of many CBMs, but mainly, to construct bi-functional CBM-fused proteins for specific applications. This review presents a comprehensive summary of the uses of CBMs recombinantly produced from heterologous organisms, or by the original host, along with the latest advances. Emphasis is given particularly to the applications of recombinant CBM-fusions in: (*a*) modification of fibers, (*b*) production, purification and immobilization of recombinant proteins, (*c*) functionalization of biomaterials and (*d*) development of microarrays and probes.

© 2015 Elsevier Inc. All rights reserved.

Contents

* Corresponding author at: CEB — Centre of Biological Engineering, Universidade do Minho, Campus de Gualtar, 4710-057 Braga, Portugal. Tel.: +351 253 604 400; fax: +351 253 604 429.

E-mail addresses: carlaoliveira@deb.uminho.pt (C. Oliveira), veracarvalho@deb.uminho.pt (V. Carvalho), luciliad@deb.uminho.pt (L. Domingues), fmgama@deb.uminho.pt (F.M. Gama).

1. Introduction

The molecular recognition of carbohydrates by proteins, namely glycoside hydrolases, is essential in several biological processes, including cell–cell recognition, cellular adhesion, and host-pathogen interactions. Therefore, understanding the structural basis of the ligand specificity of

carbohydrate-binding proteins is critical (Simpson et al., 2000). Many carbohydrate-active enzymes (CAZymes) are modular and constituted by a catalytic module and one or more non-catalytic carbohydrate-binding modules (CBMs) (Boraston et al., 2002, 2004; Hashimoto, 2006). Originally, the CBMs were classified as cellulose-binding domains (CBDs), because the first examples of these classes of protein domains bound to crystalline cellulose (Gilkes et al., 1988; Tomme et al., 1988). However, the more broad term CBM was proposed to depict the diverse ligand specificity of these sugar-binding modules derived from glycoside hydrolases. The Carbohydrate-Active enZYmes Database (CAZy) (Lombard et al., 2014) (www.cazy.org) describes the families of structurally related catalytic domains and CBMs (or functional domains) of enzymes that degrade, modify, or create glycosidic bonds such as glycoside hydrolases, glycosyltransferases, carbohydrate esterases and polysaccharide lyases.

1.1. CBM nomenclature

The nomenclature system for CBMs is based on the systematic nomenclature adopted for glycoside hydrolases (Henrissat et al., 1998). Briefly, a CBM is named by its family number, e.g. the family 10 CBM from *Cellvibrio japonicus* Xyn10A would be called CBM10, however to improve clarity, it could also be included the organism and even the enzyme from which it is derived. So, for this example, the CBM could be defined as *Cj*CBM10 or as *Cj* Xyn10ACBM10. If the glycoside hydrolase contains tandem CBMs of the same family, a number matching to the position of the CBM in the enzyme relative to the N-terminus is included.

1.2. CBM classification

In the CAZy database, a CBM is defined as a contiguous amino acid sequence within a carbohydrate-active enzyme with a discrete fold having carbohydrate-binding activity (Shoseyov et al., 2006). Numerous CBMs have been identified experimentally, and based on amino acid similarity many putative CBMs can be further identified. Currently, 48756 CBMs have been divided into 71 different families based on amino acid sequence, binding specificity, and structure (extensive data and classification can be found in the CAZy database www.cazy.org). The information on the CAZy database is constantly reviewed and reanalyzed in order to keep the information updated in such way that new families are frequently added (Lombard et al., 2014).

1.3. Fold families

Besides amino acid similarity classification, CBMs can also be divided into families according to their protein fold and tridimensional structure. In 2004, Boraston et al. classified the structures into "fold families" (Table 1 and Fig. 1).

The dominant fold among the CBMs is the β-sandwich fold in terms of the number of families and entries in databases. This fold is

Table 1
CBM fold families.
(Updated from Boraston et al., 2004; Guillen et al., 2010; Hashimoto, 2006).

Fold family	Fold	CBM families
1	β-Sandwich	2, 3, 4, 6, 9, 11, 15, 16, 17, 20, 21, 22, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 40, 41, 42, 44, 48, 47, 51, 70
2	β-Trefoil	13, 42
3	Cysteine knot	1
4	Unique	5, 12
5	OB fold	10
6	Hevein fold	18
7	Unique; contains hevein-like fold	14

characterized by two β-sheets, each consisting of three to six antiparallel β-strands (Fig. 1-A, B, C). All of the β-sandwich CBMs have at least one bound metal atom (with the exception of CBM2a from *Cellulomonas fimi* xylanase 10A).

The β-trefoil is the second most frequent fold. It comprises twelve strands of β-sheet, establishing six hairpin turns. A β-barrel structure is formed by six of the strands, associated with three hairpin turns. The remainder three hairpin turns form a triangular cap on one end of the β-barrel named 'hairpin triplet' (Fig. 1-E). The trefoil domain, defined by Boraston et al. (2004), is a contiguous amino acid sequence with four β-strands and two hairpin structures having a trefoil shape. Each trefoil domain provides one hairpin (two β-strands) to the β-barrel and one hairpin to the hairpin triplet.

The majority of the CBMs belonging to the fold family of OB (oligonucleotide/oligosaccharide-binding) have planar carbohydrate-binding sites containing aromatic residues (Fig. 1-H). Hevein domains are relatively small CBMs (~40 amino acids) and were originally identified as chitin-binding proteins in plants. The fold is mainly coil, however it has two small β-sheets and a tiny region of helix (Fig. 1-G).

1.4. Types of CBMs

The protein fold allowed CBMs to be grouped into fold families, however this is not predictive of function. The topology of CBM-ligand binding sites was also used by Boraston et al. (2004) to categorize CBMs into different "types" (Fig. 1), replicating the macromolecular structure of the target ligand. This organization was based on the structural and functional similarities of the CBMs, and three types of CBMs were described: type A ('surface-binding' CBMs), type B ('glycan-chain-binding' CBMs), and type C ('small-sugar-binding' CBMs).

Type A CBMs or surface-binding CBMs (Fig. 1-A, D, F, H) have a flat hydrophobic binding surface comprised of aromatic residues. The planar architecture of the binding sites is consistent with the flat surfaces of crystalline polysaccharides like cellulose and chitin (Boraston et al., 2004). The interaction of type A CBMs with crystalline cellulose is related with positive entropy, indicating that the thermodynamic forces that drive the binding of CBMs to crystalline ligands are relatively unique among CBMs. Another feature of type A CBMs is their little or no affinity for soluble polysaccharides (Bolam et al., 1998; Shoseyov et al., 2006).

Type B CBMs (Fig. 1-B) binding site architecture displays a cleft or groove arrangement and comprise several subsites able to accommodate the individual sugar units of the polysaccharide. The binding proficiency of the type B CBMs is assessed by the degree of polymerization of the carbohydrate ligand: several studies revealed high affinity toward hexasaccharides and negligible interaction with oligosaccharides with degree of polymerization of three or less. For this reason, type B CBMs are often called "chain binders". As in type A CBMs, the aromatic side chains play a critical role in ligand binding, and the orientation of the aromatic residues are crucial determinants of binding specificity (Boraston et al., 2004; Filonova et al., 2007a; Tomme et al., 1998).

Type C CBMs (Fig. 1-C, E, G, I), or lectin-like CBMs, have the lectin-like feature of binding optimally to mono-, di-, or tri-saccharides due to steric restriction in the binding site, lacking the extended binding site grooves of type B CBMs. (Abbott et al., 2008; Boraston et al., 2003a, 2004; Gregg et al., 2008). Due to their limited existence in plant cell wall of active glycoside hydrolases, in general, the identification and characterization of type C CBMs falls behind type A and B classification.
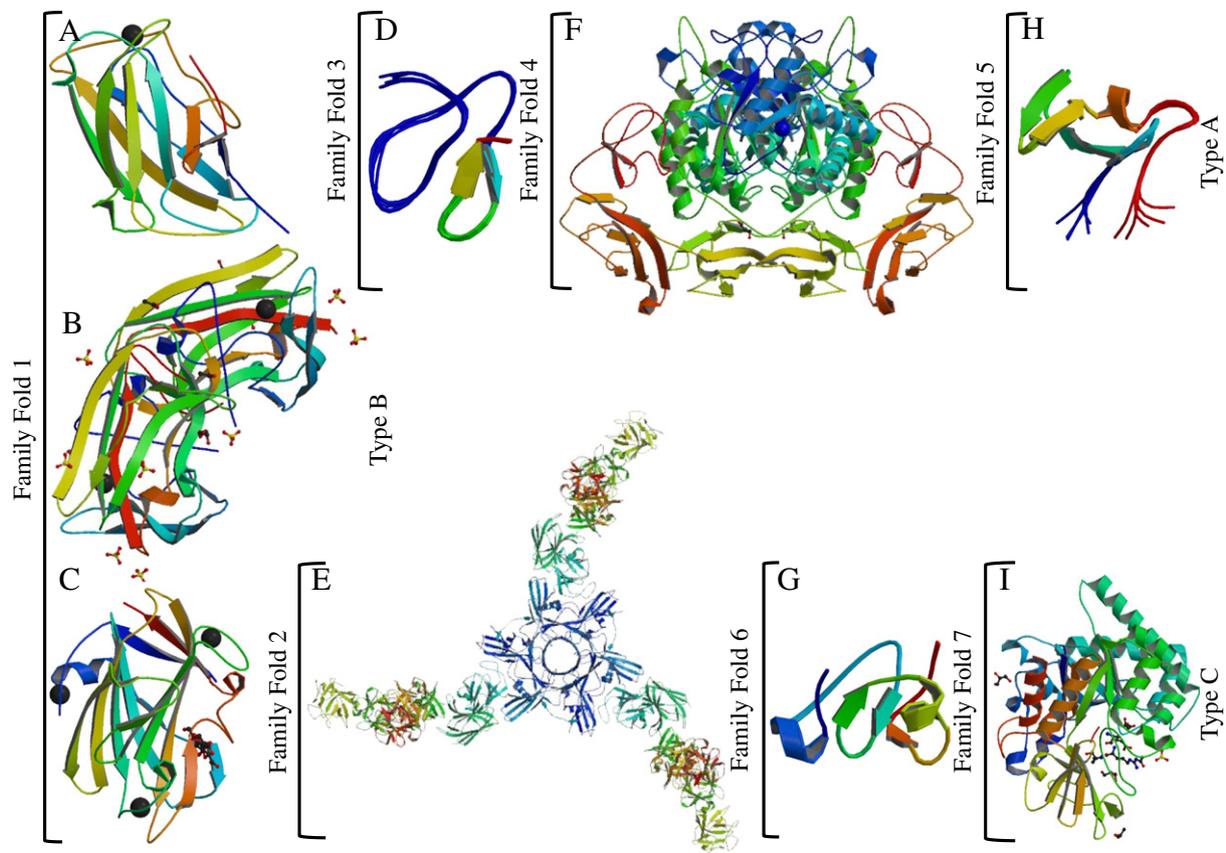
1.5. Role of CBMs in CAZymes

Several organisms produce CAZymes in order to perform various modifications to carbohydrates such as cleavage or formation of glycosidic bonds in polysaccharides and glycoconjugates. The CBMs present in most CAZymes have three general roles with respect to the function

Fig. 1. CBM fold families and functional types (CBMs represented as ribbon structures, assumed biological molecule). CBMs shown are as follows: (A) family 3 CBM from *Clostridium thermocellum* (CtCBM3), PDB code 4B9C (Yaniv et al., 2014); (B) family 4 from *C. thermocellum* (CtCBM4), PDB code 3P6B (Alahuhta et al., 2011); (C) family 9 from *Thermotoga maritima* (TmCBM9), PDB code 1I82 (Notenboom et al., 2001); (D) family 1 CBM from *Trichoderma reesei* (TrCBM1), PDB code 4BMF (Mattinen et al., 1998); (E) family 13 from *Clostridium botulinum* (CbCBM13), PDB code 3WIN (Amatsu et al., 2013); (F) family 5 from *Moritella marina* (MmCBM5), PDB code 4HMC (Malecki et al., 2013); (G) family 18 CBM from *Triticum kiharae* (TkCBM18) PDB code 2LB7 (Dubovskii et al., 2011); (H) family 10 CBM from *Cellvibrio japonicas* Ueda107 (CjCBM10), PDB code 1E8R (Raghothama et al., 2000); (I) family 14 CBM from *Homo sapiens* (HsCBM14) PDB code 1WAW (Rao et al., 2005).
Images from the RCSB PDB (www.rcsb.org) (Berman et al., 2000).

of their associated catalytic modules: proximity and avidity effect; substrate targeting; and microcrystallite disruptive function.

The CBMs promote the association of the enzyme with the substrate, securing a prolonged contact and effectively increasing the effective concentration on the polysaccharide surface thus enhancing its enzymatic activity (Bolam et al., 1998; Boraston et al., 2004; Guillen et al., 2010; Levy and Shoseyov, 2002; Shoseyov et al., 2006). This proximity effect has been demonstrated by several studies where it is noticed that CAZymes fail to act on their substrates when CBMs are genetically removed (Bolam et al., 1998; Boraston et al., 2003a; Shoseyov et al., 2006). This feature is mainly observed in enzymes that act on insoluble substrates and also in cellulosomes (Boraston et al., 2003a). CBMs can be present in single, tandem or even in multiple copies within the CAZymes architecture, and the avidity effect is related to this feature (Guillen et al., 2010). The same enzyme can be linked to several CBMs with similar or dissimilar binding specificity. Homogenous multimodularity could increase the avidity of the CAZyme for the matching substrate; on the other hand, heterogeneous multimodularity could permit the enzyme to bind different substrates. In 2009, Connaris et al. constructed a recombinant protein with tandem copies of the CBM40 from *Vibrio cholerae* sialidase. The sialidase from *V. cholerae* has two CBMs that flank the catalytic domain. The N-terminal CBM (CBM40) recognizes sialic acid with the highest reported affinity for a sialic acid-binding protein, K_d ~30 μM. Identical copies of *Vc*CBM40 were fused and manipulated in order to enhance its affinity through the avidity effect. Using four CBM40 a gain up to 3 times in affinity (due to the avidity effect) was observed when comparing with the single CBM40.

It has already been demonstrated that CBMs have specific substrate affinity, able to recognize different crystalline, amorphous, soluble and non-soluble polysaccharides (Linder and Teeri, 1997; Tomme et al., 1995; Wang et al., 2002). The data available nowadays suggests that the CBMs can be highly specific, targeting and distinguishing substrates with subtle structural differences. The substrate targeting effect has been shown in several studies and it has been used in biotechnological applications namely as tools for the elucidation of protein-carbohydrate interaction mechanisms and as molecular probes for polysaccharide localization in situ, as a means to identify different polysaccharides in plant cell-walls (McCartney et al., 2004).

Some CBMs have the ability to disrupt crystalline polysaccharides leading to an increase in substrate access. Din et al. (1991) first demonstrated this disruptive effect in 1991, but since then several CBMs were described as having this effect on polysaccharides (Gao et al., 2001; Pinto et al., 2004; Vaaje-Kolstad et al., 2005; Wang et al., 2008), however the mechanism of CBM-cellulose structure disruption, from a mechanistic stand point, is still unclear. Binding of CBMs to a crystalline substrate leads to polysaccharide chains disorganization and enhancement of substrate availability consequently increasing the CAZymes activity.

1.6. General applications of CBMs

CBMs can be obtained by two methods: enzymatic proteolysis of the CBM from the enzyme (Lemos et al., 2000; Pinto et al., 2004; Tilbeurgh et al., 1986) or using the recombinant DNA technology (recombinant CBMs) (Andrade et al., 2010a; Carvalho et al., 2008; Moreira et al., 2008). This last method allows overcoming the purity limitations of

the former one, namely avoiding the contaminant residual catalytic activity, but it is also a way to improve CBMs availability, and to obtain CBMs fused with other proteins (recombinant CBM-fusions) for specific applications, which will be the focus of this review. Recombinant CBM-fusions can either be produced in heterologous organisms or in the CBM original host. The heterologous host most used for producing recombinant CBMs is the bacterium *Escherichia coli*, followed by the yeast *Pichia pastoris*, due to their well-known genetics, fast growth, high-density cultivations and the availability of large numbers of compatible biotechnological tools. However, the choice of the host for the production of a given CBM is dependent upon the CBM properties and application for which it is intended. For example, production in yeast is considered when post-translational modifications (e.g. glycosylation) are required (Boraston et al., 2003b). Recombinant DNA technology also allows for structural and functional characterization of CBMs (Alahuhta et al., 2011; Dubovskii et al., 2011; Guo and Catchmark, 2013; Khan et al., 2013; Malecki et al., 2013; Moreira et al., 2010; Yaniv et al., 2014) and for other uses of CBMs, such as: expression in vivo for plant modulation studies (Safra-Dassa et al., 2006); protein engineering (consisting in the intrinsic addition, substitution or mutation of a CBM in order to improve the enzyme stability or hydrolytic activity) (Latorre-Garcia et al., 2005; Mahadevan et al., 2008); expression at the surface of cells, virus, and phages, for immobilization and targeting purposes (Fukuda et al., 2008; J.W. Kim et al., 2013; Tarahomjoo et al., 2008; Tolba et al., 2010); and synthetic construction of minicellulosomes (S. Kim et al., 2013). In this way, CBMs present a broad range of biotechnological applications (Bayer et al., 1994; Greenwood et al., 1992; Levy and Shoseyov, 2002; Ong et al., 1989; Shoseyov et al., 2006; Tomme et al., 1998; Volkov et al., 2004), some of which are already patented (Chang et al., 2010; Connaris and Taylor, 2011; Heerze et al., 2002; Schnorr and Christensen, 2005). Fig. 2 illustrates the main areas where CBMs have been used (Shoseyov et al., 2006).

Recombinant CBMs can find application in all the represented areas. In fact, most of the applications are reported for CBMs obtained from recombinant hosts as fusion proteins. These include, among others: the improvement of the cellulose fiber properties; production, purification and immobilization of recombinant proteins; functionalization of biomaterials; probes for protein-carbohydrate interaction and microarrays. There has been an increasing number of publications on this subject in the last few years; this review aims to summarize the applications of CBMs produced by recombinant methods, highlighting the most relevant and recent works.

2. Applications of recombinant CBMs

Recombinant CBMs are commonly produced from hosts fused with a partner (Fig. 3). Fusions aim to facilitate the production and purification of the CBM (e.g. in the case of small CBMs), or of the partner (e.g. peptides), the in situ detection of the CBM, and to obtain bi-functional proteins. CBM partners can be proteins, peptides, CBMs (different or the same), enzymes, or other molecules (e.g. antibodies). In most cases, the role of the CBMs is to bind the tagged protein to a target or support, taking advantage of specific carbohydrate-binding affinities.

2.1. Modification of fibers

The use of native CBMs for the modification of fibers used in the paper and textile industry has been demonstrated in several works (e.g. Cavaco-Paulo et al., 1999; Pala et al., 2001; Ramos et al., 2007). There are some reports on CBMs obtained from recombinant hosts concerning this application.

In the paper industry, enhancements in both pulp and paper properties using recombinant CBMs produced from *E. coli* have been reported. The papermaking process is essentially a very large drainage operation followed by press and dryer sections. While water removal in the press and dryer section is expensive, enhancements in the drainage section will have an important positive impact on the global performance of the process. Recombinant CBM3, originally from the bacterium *Clostridium thermocellum* CipA scaffolding protein, conjugated with polyethynel glycol (PEG), was able to improve the drainability of *Eucalyptus globulus*

Fig. 2. CBM applications and related areas.

PAPER INDUSTRY
Enhancement of the mechanical and physical properties of paper
(Cadena et al., 2010; Pala et al., 2001; Shi et al., 2014; Shoseyov et al., 2006)

TEXTILE INDUSTRY
Improvement of dye affinity Superior depilling
(Cavaco-Paulo et al., 1999; Degani et al., 2004; Ha et al., 2008; Ramos et al., 2007)

FOOD INDUSTRY
Improvement of nutritional value in animal feed Modulation of plant growth
(Ribeiro et al., 2008; Shoseyov et al., 2006; Tarahomjoo et al., 2008)

OTHER
Molecular probes, microarrays, immobilization of proteins and cells, enzyme engineering
(Haimovitz et al., 2008; Ofir et al., 2005; Shoseyov et al., 2006; Volkov et al., 2004)

ENVIRONMENT
Bioremediation
(Richins et al., 2000; Shoseyov et al., 2006; Xu et al., 2002)

MOLECULAR BIOLOGY
Recombinant protein production and purification
(Ong et al., 1989; Shoseyov et al., 2006; Tomme et al., 1998; Volkov et al., 2004)

BIOMEDICINE
Functionalization of biomaterials
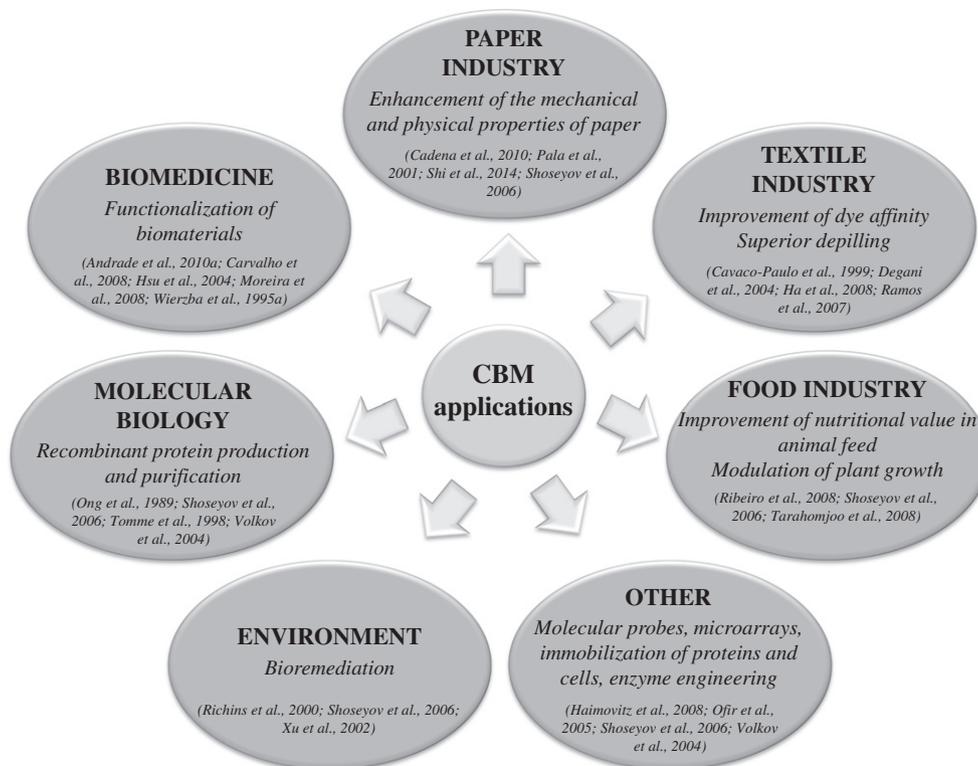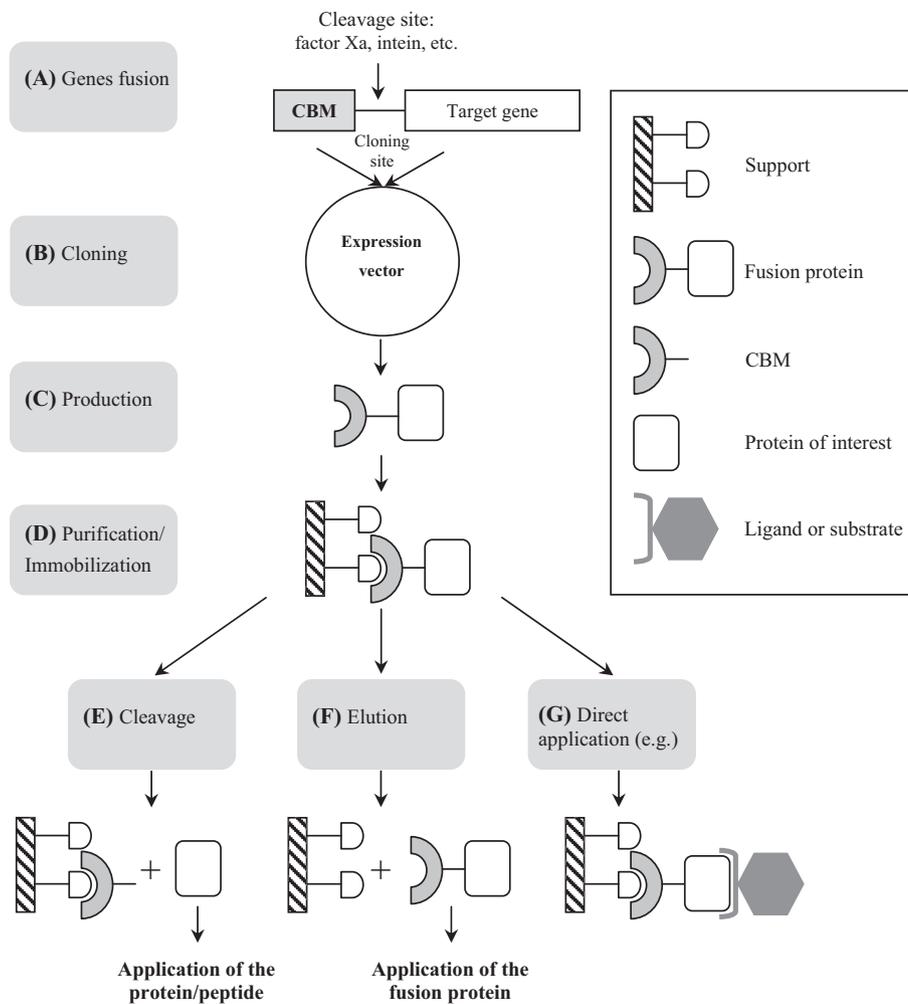(Andrade et al., 2010a; Carvalho et al., 2008; Hsu et al., 2004; Moreira et al., 2008; Wierzba et al., 1995a)

CBM applications

Fig. 3. Schematic representation of the recombinant CBM-fusion technology, from cloning to application. (A) Gene fusion between CBM and a gene of interest. CBM can either be cloned at the C-terminal or N-terminal of the target gene, the N-terminal fusion is represented. (B) Transformation of ligated plasmid vector into a prokaryotic or eukaryotic expression system. (C) Overexpression of the recombinant protein. (D) Purification/immobilization of the fusion protein on cellulose or other binding matrix. The fusion protein can also be purified by other means, as for instance, immobilized metal affinity chromatography (IMAC) using the attached his-tag (a tail of six histidines). (E) The immobilized protein is released from CBM and support by proteolytic cleavage of the engineered sequence located between the CBM and the protein (application of CBM-free protein or peptide). (F) The fusion protein is eluted from support for further application. (G) The immobilized fusion protein is directly used by the addition of a ligand or a substrate. (Figure adapted from Levy and Shoseyov, 2002).

and *Pinus sylvestris* pulps without affecting the physical properties of the paper sheets (Machado et al., 2009). This result was attributed to the presence of PEG, since the recombinant CBM alone was unable to modify pulp and paper properties, in spite of having high cellulose-binding affinity. It was suggested that the PEG mimetized the glycosidic fraction present in fungal CBMs, lacking in bacterial CBMs. Previously, highly glycosylated fungal CBMs showed to improve pulp drainability (Pala et al., 2001). Indeed, as PEG, the oligosaccharides attached to the fungal CBMs have high water affinity, and thus may lead to the hydration and stabilization of the fibers, thereby reducing the inter-fiber interactions and contributing to a better water drainage. Controversially, Cadena et al. (2010) reported that the un-glycosylated recombinant CBM3b, originally from *Paenibacillus barcinonensis endoglucanase* Cel9B, could alter cellulose fiber surface and influence paper properties. This recombinant CBM decreased the drainability of totally chlorine free kraft pulp from *E. globulus*, while it slightly increased paper strength properties (Cadena et al., 2010). In this work, the double amount of CBM was used (2 mg of CBM per gram of dried fibers).

Concerning recombinant CBM-fusions, Levy et al. (2002) constructed a bi-functional protein, containing two-fused cellulose-binding modules (CBM3), from *Clostridium cellulovorans*, to mimic the chemistry of cellulose cross-linking. The recombinant cellulose cross-linked protein

was applied onto Whatman cellulose filter paper by immersion. The fusion protein improved the treated paper's mechanical properties namely, tensile strength, brittleness, Young's modulus and energy to break. Recombinant CBM alone also improved the mechanical properties of paper, although to a lower extent. In addition, the fusion protein, in the range of concentrations of 0.025–2.5 mg/ml, conferred to the surface of the filter paper high hydrophobic nature transforming it into a more water-repellent paper, comparing to single CBM. Afterwards, the same authors constructed another bi-functional protein, a polysaccharide cross-bridging protein, containing a cellulose-binding domain from *C. cellulovorans* and a starch-binding domain from *Aspergillus niger* B1, fused *in frame* via a synthetic elastin gene (Levy et al., 2004). The fusion protein demonstrated cross-bridging ability in different model systems composed of insoluble or soluble starch and cellulose. The treatment of paper fibers with this recombinant protein, together with corn-starch, improved paper dry strength.

Very recently, Shi et al. (2014) engineered four double CBMs containing family 1 (from *Volvariella volvacea*) and/or family 3 (from *C. thermocellum*) CBMs, linked by family 1 native linker (NL) or the $(G_4S)_3$ linker (GS). The recombinant CBM3-GS-CBM3 was the most effective double CBM in enhancing paper mechanical properties in terms of folding endurance (27.4%) and tensile strength (15.5%), but

led to a slight increase in bursting strength (3.1%). The recombinant CBM1-NL-CBM1 achieved a significant simultaneous increase in tensile (12.6%) and burst (8.8%) strengths, and folding endurance (16.7%). The other two double CBMs, CBM3-GS-CBM1 and CBM3-NL-CBM1, had the lowest effective paper property improvement. The amount of 2.5 mg of double CBMs per gram of dried fibers was used.

For improving digital printing, Qi et al. (2008) constructed a recombinant triblock protein with dispersant and binding properties. The chimeric protein consisted in a carbon black binding peptide and a cellulose-binding peptide (both identified from phage display libraries through biopanning), jointed by a small, rigid, and hydrophilic interdomain linker (adapted from endoglucanase A of *Cellulomonas fimi*), whose function was to isolate the two peptides and allow the dual binding activity. The structured triblock protein was shown to disperse carbon black particles and attach it to paper surfaces.

CBMs have also been recombinantly fused with other proteins in *E. coli* for application in the manufacture of cotton fabrics. The scouring process to remove the cuticle layer of cotton fiber is one of the most important processes determining the fabric quality, as it affects the water absorption and dyeing efficiency of the cotton fabrics (Ha et al., 2008). Multidomain proteins able to monitorize the efficiency of cotton fabrics scouring were obtained by the fusion of a reporter moiety and cellulose-binding CBMs (Degani et al., 2004; Ha et al., 2008). Degani et al. (2004) used as reporter protein the enzyme β-glucuronidase (GUS) fused at the C-terminal of the CBM. An increase of CBM-GUS activity on a cotton fabric was correlated with the extent of the scouring process, since the amount of bound CBM-GUS increased proportionally when more cellulose fibers became available as the binding site for the multidomain protein (Degani et al., 2004). The fabric-bound GUS activity was visualized using the substrate 5-bromo-4-chloro-3-indolyl-β-D-glucuronic acid (X-Gluc), which yields an indigo-blue reaction product. The CBM-GUS test displayed higher sensitivity and repeatability than conventional water absorption techniques traditionally used in the textile industry, which require expensive equipment and special skills. Later, Ha et al. (2008) demonstrated that β-glycosidase (BglA) from *Thermus caldophilus*, fused at the N-terminal of *C. fimi* exoglucanase CBM, was the most appropriate reporter, showing a higher applicability and stability than GFP, DsRed2, or a tetrameric GUS produced from *E. coli*. Cotton fabrics with different scouring levels were treated with the recombinant BglA-CBM and incubated with X-Gal as the chromogenic substrate, resulting in a visible indigo color within 2 h, which intensity changed according to the conditions and extent of the scouring. Recently, the scouring efficiency of *Thermobifida fusca* cutinase was improved by the fusion of its C-terminus with the CBM from *T. fusca* cellulase Cel6A (CBMCel6A) or *C. fimi* cellulase CenA (CBMCenA) (Zhang et al., 2010). In addition, a strong synergistic effect between the recombinant fusion proteins and pectinase was observed. The cutinase-CBMCenA fusion protein was genetically modified in the CBM binding sites, by site-directed mutagenesis, to enhance its activity toward polyethylene terephthalate fiber, resulting in an improved binding and catalytic efficiency of 1.4–1.5 fold, when compared to that of the native enzyme (Zhang et al., 2013).

2.2. Recombinant protein production and purification

2.2.1. Affinity purification tools

The purification of recombinant proteins remains one of the most difficult and expensive tasks in biotechnology. The CBMs are attractive affinity tags for protein purification for several reasons (Wan et al., 2011): (i) the highly specific binding ability of the protein fused with a CBM tag; (ii) their low non-specific binding for other proteins; (iii) the efficient release of bound protein under non-denaturing conditions; (iv) its enhanced protein folding and secretion/solubility; and (v) its increased protein yield. In addition, they may allow for simultaneous purification and immobilization of the fused proteins leading to

other applications (see sections "Immobilization of recombinant proteins" and "Microarrays and probes"). Therefore, one of the most explored applications of CBMs has been their use as affinity purification tags of recombinant proteins. The majority of the CBM-based purification systems have been developed in the bacterium *E. coli* but also in yeast, namely *P. pastoris*, typically using a model protein to monitor production and purification steps (Table 2). However, *P. pastoris* glycosylation was reported to reduce the substrate affinity of the recombinant CBM in some cases (CBM2 from *C. fimi* (Boraston et al., 2001)). Other heterologous hosts have also been used, such as the bacteria *Streptomyces lividans* (Ong et al., 1995) and *Ralstonia eutropha* (Reed et al., 2006), and mammalian cells (Ong et al., 1995). In CBM-fusions, the CBM tag can either be placed at the N- or C-terminus of the target protein. Engineering gene fusions which introduce a specific cleavage site between the CBM and the target polypeptide are widely used to generate free CBM-tagged proteins. In this context, different cleavage methodologies have been adopted, such as protease action (e.g. Factor Xa (Kavoosi et al., 2004; Ong et al., 1995)), chemical cleavage with formic acid (Ramos et al., 2010), and self-cleavage of intein (Hong et al., 2008a; Liao et al., 2012; Wan et al., 2011).

The most used CBMs are those that bind to cellulose, CBM3 in particular (Table 2). Several works relating to production and application of biologically active CBM-protein fusions show the feasibility of employing such CBMs as affinity tags. These include the purification of enzymes, peptides, antibodies, etc. (e.g. Ofir et al., 2005; Ramos et al., 2010; Xu et al., 2002).

Reversible binding is an important requirement for protein purification. The interaction of CBMs with cellulose has long been characterized as "irreversible" (specifically, CBMs from families 2 and 3). CBMs with "irreversible" binding present limited usefulness as an affinity tag for protein purification, because desorption may require strong denaturing conditions (Rodriguez et al., 2004). However, recent works have described the interaction of a family 3 CBM with cellulose as a dynamic and reversible process (Machado et al., 2009), with desorption achieved at mild conditions (Liao et al., 2012). Furthermore, Lim et al. (2014) showed that a family 2a CBM could be readily reversible at 37 °C using methods previously established for a family 1 CBM (Linder et al., 1996). The "irreversible" nature of the CBM-cellulose interaction, that depends both on the CBM and cellulose properties, is not yet fully characterized.

The use of cellulose as purification media presents many advantages: inertness, low non-specific protein binding, commercial availability in a variety of inexpensive forms, and safety, as it has been approved for many pharmaceutical and human applications (Terpe, 2003). In addition, since CBMs adsorb spontaneously to cellulose, very little or no pretreatment of the samples is required prior to purification. CBMs bind to cellulose at a moderately wide pH range, from 3.5 to 9.5. The main driving forces for binding are hydrogen bond formation and van der Waals interaction (Tomme et al., 1998). The affinity is so strong that a CBM-immobilized fusion protein can only be released with buffers containing urea or guanidine hydrochloride, which require subsequent refolding of the target protein. Fused proteins with CBMs of families 2 and 3 can be eluted gently from cellulose with low-polarity solvents, such as ethylene glycol, which can be removed easily by dialysis. Alternatively, and as mentioned above, the target protein can be released from the CBM tag, and thus from the support, using a cleavage strategy. CBM-fusions have been mainly purified using Avicel – microcrystalline cellulose (e.g. Sugimoto et al., 2012) and RAC – regenerated amorphous cellulose (e.g. Hong et al., 2008b; Wan et al., 2011), but also CF11 – fibrous cellulose (e.g. Ramos et al., 2010; Sugimoto et al., 2012), according to CBMs preference. RAC has higher adsorption capacity than Avicel (Hong et al., 2007, 2008b). Other cellulose-based purification media have also been used. Bacterial microcrystalline cellulose has a high adsorption capacity (Hong et al., 2007). On the other hand, filter papers are easy to manipulate, but they only have one fourth of RAC's adsorption capacity (Hong et al., 2007). Kavoosi et al. (2007b)

Table 2
Properties and uses of CBMs as affinity tags for the purification of recombinant proteins.

CBMs family	Fold family	Molecular weight (kDa)	Isoelectric point	Target protein	Production host	References
1	3	3.8	7.0	RFP, red fluorescent protein	*Pichia pastoris*	Sugimoto et al. (2012)
2	1	8.8	7.7	Protein A from *Staphylococcus aureus*	*Escherichia coli*	Rodriguez et al. (2004)
				SFC, murine stem-cell factor	*P. pastoris*	Boraston et al. (2001)
		10.9	6.1	Interleukin-2	*E. coli*, *Streptomyces lividans*, mammalian COS cells	Ong et al. (1995)
3	1	17.5	6.8	k- and λ-carrageenases	*E. coli*	Kang et al. (2014)
				Protein A from *S. aureus*	*E. coli*	Shpigel et al. (2000)
				Heat shock protein hsp60	*E. coli*	Shpigel et al. (1998)
		17.1	4.8	MAG2, magainin-2	*E. coli*	Ramos et al. (2013)
				PPGK, polyphosphate glucokinase	*E. coli*	Liao et al. (2012)
				EGFP, enhanced green fluorescent protein	*P. pastoris*	Wan et al. (2011)
				LL37, cathelicidin derived human peptide	*E. coli*	Ramos et al. (2010)
				GFP, green fluorescent protein	*E. coli*	Hong et al. (2007; 2008a,b)
		17.6	4.7	Self-assembling peptides	*Ralstonia eutropha*	Reed et al. (2006)
9	1	21.3	4.8	GFP	*E. coli*	Kavoosi et al. (2004; 2007a,b,c; 2009)
				Self-assembling peptides	*R. eutropha*	Reed et al. (2006)
21[a]	1	11.5	5.1	EGFP and phytase	*E. coli*, *P. pastoris*	Lin et al. (2009)
30	1	9.8	4.2	Cis-epoxysuccinic acid hydrolase	*E. coli*	Wang et al. (2012)

[a] CBM21 has raw starch-binding affinity, while the other CBMs in the table have cellulose-binding affinity.

developed mechanically stable porous cellulose beads for cost-effective purification scale-up of recombinant proteins.

Several CBMs are already commercialized as recombinant protein production and purification systems. For example, cellulose-binding CBMs from *C. cellulovorans* and *C. fimi* are incorporated into some Novagen's *E. coli* pET expression vectors.

2.2.2. Production of peptides and enzymes

CBMs have shown to be effective fusion partners for the production and purification of recombinant antimicrobial peptides in *E. coli* (Guerreiro et al., 2008a; Klocke et al., 2005; Ramos et al., 2010, 2013). Antimicrobial peptides (AMPs) are part of the innate immune system, acting in a wide range of physiological defensive mechanisms developed to counteract bacteria, fungi, parasites and viruses. AMPs are generally defined as cationic, amphipathic peptides, with less than 50 amino acids, including multiple arginine and lysine residues. CBMs allow overcoming the difficulties of the recombinant production of this kind of molecules, due to their small size and potential toxicity for host, and may also be used for their immobilization. The CBM3 from *C. thermocellum* has been particularly employed in this respect. This CBM is an innocuous partner since it does not present antibacterial activity and does not bind to *E. coli* surface (Guerreiro et al., 2008a). Guerreiro et al. (2008a) reported the production of four AMPs fused to the N-terminal of CBM3, that presented cellulose-binding ability, and suggested their immobilization in cellulosic supports for the generation of novel bio-products possessing antimicrobial properties for biomedical application. However, the antimicrobial activity of the fusion proteins remained to be elucidated. Soon after, the cathelicidin derived human peptide LL37, which has a broad spectrum of antimicrobial and immunomodulatory activities, was also produced fused to either N- or C-terminal of CBM3 (Ramos et al., 2010). The fusion proteins were purified taking advantage of the CBM3 specific affinity for cellulose, and LL37 was cleaved from CBM3 with formic acid and further purified by reverse-phase HPLC. The recombinant LL37 obtained from the C-terminally fused protein showed antibacterial activity against *E. coli* K12, but not the one obtained from the N-terminally fused protein, presumably due to the presence of a methionine residue at its N-terminal (Ramos et al., 2010). Interestingly, the LL37 obtained by this methodology has been shown to preserve its immunophysiological properties in vitro and in vivo (Ramos et al., 2011). More recently, Ramos et al. (2013) produced Magainin-2 (MAG2), a polycationic antimicrobial peptide isolated from the skin of the African clawed frog *Xenopus laevis*, fused once again to the N-terminal of CBM3. Recombinant MAG2 was successfully cleaved and purified from the fusion partner, using a similar

protocol to the previously applied for LL37 recovery, and was functionally active against Gram-negative bacteria.

Another widely explored application of CBMs has been the production of enhanced enzymes, particularly cellulases, free (e.g. Hong et al., 2006; Kang et al., 2007; Martin et al., 2013; Reyes-Ortiz et al., 2013; Tang et al., 2013; Thongekkaew et al., 2013; von Ossowski et al., 2005; Yeh et al., 2005) or immobilized (see section "Immobilization of recombinant proteins"), by CBM fusion. It is obvious the large diversity of biological processes that can benefit from this CBM application. For example, a recombinant derivative of *C. thermocellum* xylanase, containing a CBM6, and a recombinant *C. thermocellum* cellulase fused to a family 11 β-glucan-binding domain, were used in the supplementation of animals feeding with success (Fontes et al., 2004; Guerreiro et al., 2008b; Ribeiro et al., 2008). The efficacy of the CBM6 could be noticed by the increased final body weight of birds fed on a wheat-based diet supplemented with the recombinant CBM-xylanase, when compared with birds only receiving the xylanase catalytic module (Fontes et al., 2004). As other example, among four recombinant fusions of different CBMs (families 3, 4–2, 6 and 9–2) with *C. thermocellum* cellodextrin phosphorylase (lacking an apparent CBM), the fusion with CBM9 resulted in a non-natural cellulose phosphorylase, which opened perspectives for developing a highly active enzyme for in vitro hydrogen production from cellulose by synthetic pathway biotransformation (Ye et al., 2011).

2.3. Immobilization of recombinant proteins

CBMs have been used as efficient immobilization tools of recombinantly fused proteins for different applications. One of these applications is bioremediation, aiming for the development of efficient pollutants removal systems. Some authors reported the degradation of toxic compounds using decontaminating enzymes fused with cellulose-binding CBMs, which enabled a single-step purification and immobilization of the fusion proteins into different cellulosic materials (Richins et al., 2000; Xu et al., 2002).

The immobilization of CBM-fusion proteins also allowed the enhancement of properties of some industrially important enzymes (Harris et al., 2010; Hwang et al., 2004; Myung et al., 2011; Rotticci-Mulder et al., 2001; Wang et al., 2012). For example, recently, Wang et al. (2012) fused five different CBMs to cis-epoxysuccinic acid hydrolase (CESH) and immobilized the chimeric enzymes on cellulose to improve their efficiency and stability. These chimeric enzymes were expressed in *E. coli* and purified using nickel or, for simultaneous immobilization, cellulose affinity. CESH may be used in the chemical industry

to produce tartaric acid, but its application is hampered due to instability of the purified form. The stability was slightly increased by the fusion with CBMs, and significantly by the immobilization on cellulose. Among the CBMs used, CBM30 (from *C. thermocellum*) was the best fusion partner for improving both the enzymatic efficiency and stability of CESH.

Enzyme recycling is an important issue in several biotechnological processes. A peculiar thermally reversible enzyme-binding system suitable for regenerating batch enzymatic processes was developed, in which a CBM from *C. cellulovorans* was fused with thermophilic enzymes from *Pyrococcus furiosus* (Nahalka and Gemeiner, 2006). The enzyme was active and free in the reaction mixture at 80–90 °C and deactivated and immobilized by affinity adsorption to cellulose at 30–40 °C. Craig et al. (2006) also developed a reversible high affinity interaction system but using the calcium-dependent cohesin-dockerin interaction from *C. thermocellum*. The CBM from *C. cellulovorans* fused to that cohesin was immobilized onto a cellulose matrix, allowing the binding of a complementary dockerin-tagged recombinant protein, an interaction which could be reversed with EDTA. Recently, a recombinant miniscaffoldin was constructed in *E. coli*, containing one family 3 CBM (from *C. thermocellum* CipA) and three types of cohesins for single-step purification and co-immobilization of a synthetic multi-enzyme complex also produced in *E. coli* (i.e. three complementary dockerin-tagged enzymes). This complex allowed the enhancement of the initial enzymatic reaction rate by facilitating substrate challenge (You and Zhang, 2013; You et al., 2012). The economics of affinity-tagging technologies depends in part on the cost and efficiency of the bioprocessing step used to remove the affinity tag. Kwan et al. (2002) used cellulose beads to immobilize a derivative of Factor X, recombinantly fused with a CBM in mammalian cells. The so immobilized Factor X was used for cost-effective application in tag removal from recombinant fusion proteins, thus facilitating protein purification (Kwan et al., 2002). The self-activating derivative enzyme retained approximately 80% of its initial activity after 30 days of continuous hydrolysis of a fusion protein substrate (Kwan et al., 2002). In another work, cellulose-containing magnetic nanoparticles were applied for immobilizing CBM-tagged enzymes, which can be employed for enzyme recycling/removal from enzymatic reactions by using a magnet (Myung et al., 2013; You et al., 2013).

Antibody-binding domains fused with CBMs have been used for studying antibody mediated cell or particle adhesion onto cellulose surfaces for mammalian cell culture applications (Craig et al., 2007; Lewis et al., 2006; Nordon et al., 2004; Pangu et al., 2007). For example, some authors described the construction and/or application of such fusions for direct immobilization of antibodies and cells onto regenerated cellulose made hollow fiber membranes (Craig et al., 2007; Nordon et al., 2004). Cell separation, based on the monoclonal antibody recognition, is achieved using these devices specifically for high-density cell culture, with the potential advantage to integrate cell selection and culture processes (Craig et al., 2007). Moreover, antibody-binding domains have been immobilized via CBMs fusion for biosensor applications. In an interesting work, several bispecific pentameric fusion proteins, consisting of five single-domain antibodies and five cellulose-binding modules linked via verotoxin B subunit (self-assembly facilitator of the 5 polypeptides), were engineered in *E. coli* and *P. pastoris* for simultaneous immobilization on cellulose and detection of the human pathogen *Staphylococcus aureus* (Hussack et al., 2009). Six bispecific expression cassettes were constructed by fusing three different CBMs (families 2 and 9) to the specific antibody in both orientations, resulting in six different pentamers. One of the proteins (containing N-terminal CBM9), when impregnated in cellulose filters, was able to recognize *S. aureus* cells in a flow-through detection assay. The ability of pentamerized CBMs to bind cellulose may form the basis of an immobilization platform for multivalent display of high avidity binding reagents on cellulosic filters for sensing of pathogens, biomarkers and environmental pollutants (Hussack et al., 2009).

Functionalization of biomaterials

Recombinant fusions of CBM-bioactive peptides produced and purified from *E. coli* have been employed to improve cell adhesion and proliferation on carbohydrate-based biomaterials for biomedical application. Among the bioactive cell adhesion motifs found in extracellular matrix proteins, the peptides Arg-Gly-Asp (RGD) (Andrade et al., 2010a; Carvalho et al., 2008; Hsu et al., 2004; Moreira et al., 2008; Wierzba et al., 1995a,1995b) and Ile-Lys-Val-Ala-Val (IKVAV) (Pertile et al., 2012) were used for such purpose. Andrade et al. (2010a) constructed different recombinant fusions of CBM3 from *C. thermocellum* cellulosome with RGD and GRGDY (RGD/CBM, RGD/CBM/RGD, GRGDY/CBM and GRGDY/CBM/GRGDY) in order to improve the affinity of mouse embryo fibroblasts 3T3 for bacterial cellulose (BC). In recent years, BC has emerged as a promising biomaterial for biomedical applications due to its properties such as high crystallinity, ultrafine fiber network, high tensile strength, and biocompatibility (Andrade et al., 2010b). Indeed, the recombinant bifunctional proteins improved the fibroblasts adhesion and spreading on BC, as compared with BC treated with the recombinant CBM3 alone (Andrade et al., 2010a). Furthermore, the proteins containing the RGD sequence showed a stronger effect than those containing the GRGDY sequence. More recently, the same CBM3 recombinantly fused with two different IKVAV sequences (IKVAV and (19)IKVAV), and ligated by the native glycanase 40 amino acid linker, was used to promote neuronal and mesenchymal stem cells (MSC) adhesion also on BC, to improve biocompatibility (Pertile et al., 2012). IKVAV means the respective 5 peptide amino acids, while (19)IKVAV corresponds to the extended amino-acid sequence (19 amino acids) based on the proteolytic laminin fragment PA-22 containing the sequence IKVAV. Both recombinant fusion proteins were stably adsorbed to the BC membranes (Pertile et al., 2012). The recombinant fusion protein (19)IKVAV-CBM3 was able to significantly improve the adhesion of both neuronal and mesenchymal cells, an effect that was dependent on the cell type. On the other hand, the recombinant fusion protein IKVAV-CBM3 only presented a marginal effect on MSC adhesion. Furthermore, the recombinant fusion protein (19)IKVAV-CBM3 allowed the release of nerve growth factor (a protein that belongs to the neurotrophins proteins family, which are known to induce the survival, development, and function of neurons) secreted by MSC to the culture medium, indicating that modified BC has the potential to be used in neuronal tissue engineering applications (Pertile et al., 2012). The previously constructed recombinant fusion protein RGD-CBM3 (Andrade et al., 2010a) was also tested in this last work, showing to improve the adhesion of part of the cell lineages studied, thus revealing a cell specific behavior (Pertile et al., 2012).

Recombinant fusions of CBMs with other affinities to the RGD tripeptide were also constructed to functionalize other biomaterials, also with appealing characteristics in the perspective of biomedical application, but with distinct success. A starch-binding module (SBM20 from *Bacillus* sp. TS-23 α-amylase) fused to RGD was effective in improving, by more than 30%, the adhesion of fibroblasts to dextrin-based hydrogel (Moreira et al., 2008). In fact, cell spreading on the hydrogel surface was only observed in the presence of the RGD-SBM, since the recombinant SBM alone did not have any effect on cells adhesion. Controversially, Carvalho et al. (2008) showed that a human chitin-binding module (ChBM) recombinantly fused to RGD affected negatively fibroblasts anchorage to reacetylated chitosan films, inhibiting cells adhesion and proliferation. The authors concluded that this unexpected effect was fundamentally due to the human ChBM, either fused or not with the RGD peptide, but the reason for this result remains unclear (Carvalho et al., 2008).

Concerning other potential biomedical applications, isolated works describe very interesting features of some recombinant CBMs, such as: a CBM from *C. cellulovorans* fused with a recombinant antigen protein (from *Aeromonas salmonicida*) as adjuvant for fish parenteral vaccination (Maurice et al., 2003), the recombinant CBM4 (a domain from the

Celk gene from *C. thermocellum*) as stabilizer of single-walled carbon nanotubes (SWNTs) in water (Xu et al., 2009), and a CBM from *C. fimi* fused to a murine stem cell factor as persistent activator of factor-dependent cells (Jervis et al., 2005). Furthermore, a bi-functional protein with self-assembly and graphene-binding properties (HFBI - hydrophobin I, from *Trichoderma reesei*) fused with two CBMs (from *T. reesei* Cel7A and Cel6A enzymes) was recombinantly produced in the original host for multiple uses: self-assembly of nanofibrillar cellulose (NFC) (Varjonen et al., 2011), construction of nanocomposites of graphene and NFC (Laaksonen et al., 2011), and preparation of drug nanoparticles and their immobilization in NFC for increased storage stability (Valo et al., 2011).

2.4. Microarrays and probes

The complexity of printing non-DNA microarrays, e.g., protein and peptide microarrays, is one of the major drawbacks of this approach. The use of affinity immobilization strategies may provide a general solution for fabrication of such microarrays. In this context, CBM-based microarray technology provides a technically simple but effective alternative to conventional technology. The intrinsic specificity of CBMs for individual carbohydrates and the facile modification with other molecules allow for efficient production of protein and peptide microarrays (Moreira and Gama, 2011). In CBM-based microarray technology, the probes (recombinant CBMs fused with proteins, peptides, or antibodies, produced from *E. coli*) are immobilized onto cellulosic supports via CBM adsorption. This technology offers fundamental advantages over current non-DNA microarray technology, such as retention of protein functionality after immobilization, ease of fabrication, extended stability of the printed microarray, integrated test for quality control and the capacity to print test proteins without purification steps (Ofir et al., 2005). CBM-based microarrays can be used in a variety of potential applications, technically impractical via conventional microarray technologies (Ofir et al., 2005). Nevertheless, only a couple of works describe the construction and application of such microarrays. Ofir et al. (2005) developed a CBM-based microarray system for serodiagnosis of human immunodeficiency virus (HIV) patients, using an affinity-based probe immobilization strategy, consisting in the combination of cellulose-coated glass slides with the family 3a CBM, from the cellulosome of *C. thermocellum*, recombinantly fused with HIV-related antibodies or peptides. This work took advantage of a large library of CBM3-fused single-chain antibodies constructed before by Azriel-Rosenfeld et al. (2004). Later on, Haimovitz et al. (2008) developed a CBM-based microarray system to analyze the cohesin-dockerin specificity. The specificity of cohesin-dockerin interactions is critically important for the assembly of cellulosomal enzymes into the multienzyme cellulolytic complex (cellulosome). Knowledge of the specificity characteristics of native and mutated members of the cohesin and dockerin families provides insight into the architecture of the parent cellulosome and allows selection of suitable cohesin-dockerin pairs for biotechnological and nanotechnological applications (Slutzki et al., 2012). In the above microarray, recombinant CBM-fused cohesins (CBM from the *C. thermocellum* scaffoldin CipA) were immobilized on cellulose-coated glass slides, to which recombinant xylanase-fused dockerin samples were applied (Haimovitz et al., 2008). Using this approach, extensive cross-species interaction among type-II cohesins and dockerins was shown for the first time. Furthermore, selective intraspecies binding of an archaeal dockerin to two complementary cohesins was demonstrated. In order to investigate the origins of the observed specificity, a similar microarray system was then described, consisting of CBM attached to cohesin mutants for screening of complementary binders (Slutzki et al., 2012). The advantages of this approach are that crude cell lysate can be used without additional purification, and the microarray can be used for screening both large libraries as initial scanning for "positive" plates, and for small libraries, wherein individual colonies are printed on the slide (Slutzki et al., 2012).

The innate binding specificity of different CBMs offers a versatile approach for mapping the chemistry and structure of surfaces that contain complex carbohydrates. Therefore, CBMs can be used as molecular probes for studying plant cell walls and carbohydrate-based substrates. Several works describe the use of recombinant CBMs for characterizing both native complex carbohydrates and engineered biomaterials. In these studies, based on fluorescent techniques, different types of CBM-probes are used: recombinant CBMs fused with fluorescent proteins (e.g. Ding et al., 2006; Kawakubo et al., 2010), recombinant CBMs labeled with a fluorochrome (e.g. fluorescein isothiocyanate-FITC (Filonova et al., 2007a; Filonova et al., 2007b)), or labeled by immunofluorescence (i.e. using antibodies, as for instance for his-tag detection) (e.g. Boraston et al., 2003a; Dagel et al., 2011; Jamal et al., 2004; Lehtio et al., 2003; McCartney et al., 2004; Siroky et al., 2012). Recently, a new technique was described (using scanning electron microscopy), using recombinant CBM44 from *C. thermocellum* and recombinant CBM2a from *C. fimi*, to track specific changes in the surface morphology of cotton fibers during amorphogenesis (i.e. non-hydrolytic "opening up" or disruption of a cellulosic substrate), one of the key steps in the enzymatic deconstruction of cellulosic biomass when used as a feedstock for fuels and chemicals production (Gourlay et al., 2012). More recently, the surface accessibilities of amorphous and crystalline celluloses in Avicel to cellulase were determined quantitatively using two recombinant fluorescent CBM-probes consisting of GFP fused with a CBM3 (that binds to both amorphous and crystalline celluloses) and mono-cherry fluorescent protein (CFP) fused with a CBM17 (that binds only to amorphous cellulose), thus revealing cellulose hydrolysis mechanism (Gao et al., 2014).

3. Perspectives

Recombinant expression systems, particularly *E. coli*, have boosted the application of several CBMs in many different areas. Most applications of CBMs are strictly dependent on the fusion with a partner or target protein. In recent years, besides the exploitation of CBMs as fusion tags for recombinant protein production and purification in different hosts, CBMs have been investigated as biomedical tools, namely through the functionalization of biomaterials. Recombinant CBMs have also been used as molecular probes for characterizing and monitoring alterations in native and engineered polysaccharides and for immobilizing and improving the properties of many enzymes.

There is no doubt that the trend is for the number of CBMs to continuously grow, and that recombinant DNA technology will expand their application. The design of recombinant fusions of CBMs with pathogens and toxins binders predicts the development of CBM-based biosensors. On the other hand, artificial CBMs with engineered biological activities (e.g. anti-viral, anti-bacterial, and anti-tumor) may settle the grounds for new synthetic drugs. Furthermore, recombinant production may help to elucidate the molecular evolution of CBMs, as well as several carbohydrate-protein interactions involved in many biological processes, some of which related with human diseases, in order to develop novel therapeutic strategies.

Cellulose is the most abundant natural polymer. It is traditionally used in different fields of activity, such as textile and pulp and paper. The availability of pure cellulosic materials such as nanocellulose (now produced at the industrial scale) and the emergence of bacterial cellulose as a sophisticated new material in different fields of application demonstrate the modernity of cellulose, with limitless applications in the biomedical, composites and electronics. On the other hand, nanotechnology is being able to construct complex functional structures made of biomolecules such as proteins or nucleic acids. The physiological role of protein-protein interactions is just being unraveled. CBMs thus show as promising tools for construction of complex puzzles in the surface of cellulosic biomaterials. The cellulosome structure will certainly inspire the creation of CBM based nanomachines.

Acknowledgements

References

Abbott DW, Eirin-Lopez JM, Boraston AB. Insight into ligand diversity and novel biological roles for family 32 carbohydrate-binding modules. Mol Biol Evol 2008;25:155–67.

Alahuhta M, Luo YH, Ding SY, Himmel ME, Lunin VV. Structure of CBM4 from Clostridium thermocellum cellulase K. Acta Crystallogr F 2011;67:527–30.

Amatsu S, Sugawara Y, Matsumura T, Kitadokoro K, Fujinaga Y. Crystal structure of Clostridium botulinum whole hemagglutinin reveals a huge triskelion-shaped molecular complex. J Biol Chem 2013;288:35617–25.

Andrade FK, Moreira SM, Domingues L, Gama FM. Improving the affinity of fibroblasts for bacterial cellulose using carbohydrate-binding modules fused to RGD. J Biomed Mater Res A 2010a;92:9–17.

Andrade FK, Pertile R, Dourado F, Gama FM. Bacterial cellulose: properties, production and applications. In: Lejeune A, Deprez T, editors. Cellulose: structure and properties, derivatives and industrial uses. Nova Science Publishers Inc; 2010b. p. 427–58.

Azriel-Rosenfeld R, Valensi M, Benhar I. A human synthetic combinatorial library of arrayable single-chain antibodies based on shuffling in vivo formed CDRs into general framework regions. J Mol Biol 2004;335:177–92.

Bayer EA, Morag E, Lamed R. The cellulosome — a treasure-trove for biotechnology. Trends Biotechnol 1994;12:379–86.

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The protein data bank. Nucleic Acids Res 2000;28:235–42.

Bolam DN, Ciruela A, McQueen-Mason S, Simpson P, Williamson MP, Rixon JE, et al. Pseudomonas cellulose-binding domains mediate their effects by increasing enzyme substrate proximity. Biochem J 1998;331:775–81.

Boraston AB, McLean BW, Guarna MM, Amandaron-Akow E, Kilburn DG. A family 2a carbohydrate-binding module suitable as an affinity tag for proteins produced in Pichia pastoris. Protein Expr Purif 2001;21:417–23.

Boraston AB, McLean BW, Chen G, Li AS, Warren RAJ, Kilburn DG. Co-operative binding of triplicate carbohydrate-binding modules from a thermophilic xylanase. Mol Microbiol 2002;43:187–94.

Boraston AB, Kwan E, Chiu P, Warren RA, Kilburn DG. Recognition and hydrolysis of non-crystalline cellulose. J Biol Chem 2003a;278:6120–7.

Boraston AB, Sandercock L, Warren RA, Kilburn DG. O-glycosylation of a recombinant carbohydrate-binding module mutant secreted by Pichia pastoris. J Mol Microbiol Biotechnol 2003b;5:29–36.

Boraston AB, Bolam DN, Gilbert HJ, Davies GJ. Carbohydrate-binding modules: fine-tuning polysaccharide recognition. Biochem J 2004;382:769–81.

Cadena EM, Chriac AI, Pastor FI, Diaz P, Vidal T, Torres AL. Use of cellulases and recombinant cellulose binding domains for refining TCF kraft pulp. Biotechnol Prog 2010;26:960–7.

Carvalho V, Domingues L, Gama M. The inhibitory effect of an RGD-human chitin-binding domain fusion protein on the adhesion of fibroblasts to reacetylated chitosan films. Mol Biotechnol 2008;40:269–79.

Cavaco-Paulo A, Morgado J, Andreaus J, Kilburn D. Interactions of cotton with CBD peptides. Enzyme Microb Tech 1999;25:639–43.

Chang MD, Lee Y, Huang R, Lin S, Chou W, Liu S. Anti-viral carbohydrate binding modules compositions and methods of use. WO Patent No. 2010/131114 A2.

Connaris H, Taylor GL. Carbohydrate binding molecules. US Patent No. 2011/0269670 A1.

Connaris H, Crocker PR, Taylor GL. Enhancing the receptor affinity of the sialic acid-binding domain of Vibrio cholerae sialidase through multivalency. J Biol Chem 2009;284:7339–51.

Craig SJ, Foong FC, Nordon R. Engineered proteins containing the cohesin and dockerin domains from Clostridium thermocellum provides a reversible, high affinity interaction for biotechnology applications. J Biotechnol 2006;121:165–73.

Craig SJ, Shu A, Xu Y, Foong FC, Nordon R. Chimeric protein for selective cell attachment onto cellulosic substrates. Protein Eng Des Sel 2007;20:235–41.

Dagel DJ, Liu YS, Zhong L, Luo Y, Himmel ME, Xu Q, et al. In situ imaging of single carbohydrate-binding modules on cellulose microfibrils. J Phys Chem B 2011;115:635–41.

Degani O, Gepstein S, Dosoretz CG. A new method for measuring scouring efficiency of natural fibers based on the cellulose-binding domain-beta-glucuronidase fused protein. J Biotechnol 2004;107:265–73.

Din N, Gilkes NR, Tekant B, Miller RC, Warren AJ, Kilburn DG. Non-hydrolytic disruption of cellulose fibers by the binding domain of a bacterial cellulase. Bio-Technol 1991;9:1096–9.

Ding SY, Xu Q, Ali MK, Baker JO, Bayer EA, Barak Y, et al. Versatile derivatives of carbohydrate-binding modules for imaging of complex carbohydrates approaching the molecular level of resolution. Biotechniques 2006;41:435–6. [8, 40 passim].

Dubovskii PV, Vassilevski AA, Slavokhotova AA, Odintsova TI, Grishin EV, Egorov TA, et al. Solution structure of a defense peptide from wheat with a 10-cysteine motif. Biochem Bioph Res Co 2011;411:14–8.

Filonova L, Gunnarsson LC, Daniel G, Ohlin M. Synthetic xylan-binding modules for mapping of pulp fibres and wood sections. BMC Plant Biol 2007a;7:54.

Filonova L, Kallas AM, Greffe L, Johansson G, Teeri TT, Daniel G. Analysis of the surfaces of wood tissues and pulp fibers using carbohydrate-binding modules specific for crystalline cellulose and mannan. Biomacromolecules 2007b;8:91–7.

Fontes CM, Ponte PI, Reis TC, Soares MC, Gama LT, Dias FM, et al. A family 6 carbohydrate-binding module potentiates the efficiency of a recombinant xylanase used to supplement cereal-based diets for poultry. Br Poult Sci 2004;45:648–56.

Fukuda T, Kato-Murai M, Kuroda K, Ueda M, Suye S. Improvement in enzymatic desizing of starched cotton cloth using yeast codisplaying glucoamylase and cellulose-binding domain. Appl Microbiol Biotechnol 2008;77:1225–32.

Gao PJ, Chen GJ, Wang TH, Zhang YS, Liu J. Non-hydrolytic disruption of crystalline structure of cellulose by cellulose binding domain and linker sequence of cellobiohydrolase I from Penicillium janthinellum. Acta Bioch Bioph Sin 2001;33:13–8.

Gao SH, You C, Renneckar S, Bao J, Zhang YHP. New insights into enzymatic hydrolysis of heterogeneous cellulose by using carbohydrate-binding module 3 containing GFP and carbohydrate-binding module 17 containing CFP. Biotechnol Biofuels 2014;7:24.

Gilkes NR, Warren RAJ, Miller RC, Kilburn DG. Precise excision of the cellulose binding domains from 2 Cellulomonas fimi cellulases by a homologous protease and the effect on catalysis. J Biol Chem 1988;263:10401–7.

Gourlay K, Arantes V, Saddler JN. Use of substructure-specific carbohydrate binding modules to track changes in cellulose accessibility and surface morphology during the amorphogenesis step of enzymatic hydrolysis. Biotechnol Biofuels 2012;5:51.

Greenwood JM, Ong E, Gilkes NR, Warren RAJ, Miller RC, Kilburn DG. Cellulose-binding domains — potential for purification of complex proteins. Protein Eng 1992;5:361–5.

Gregg KJ, Finn R, Abbott DW, Boraston AB. Divergent modes of glycan recognition by a new family of carbohydrate-binding modules. J Biol Chem 2008;283:12604–13.

Guerreiro CI, Fontes CM, Gama M, Domingues L. Escherichia coli expression and purification of four antimicrobial peptides fused to a family 3 carbohydrate-binding module (CBM) from Clostridium thermocellum. Protein Expr Purif 2008a;59:161–8.

Guerreiro CI, Ribeiro T, Ponte PI, Lordelo MM, Falcao L, Freire JP, et al. Role of a family 11 carbohydrate-binding module in the function of a recombinant cellulase used to supplement a barley-based diet for broiler chickens. Br Poult Sci 2008b;49:446–54.

Guillen D, Sanchez S, Rodriguez-Sanoja R. Carbohydrate-binding domains: multiplicity of biological roles. Appl Microbiol Biot 2010;85:1241–9.

Guo J, Catchmark JM. Binding specificity and thermodynamics of cellulose-binding modules from Trichoderma reesei Cel7A and Cel6A. Biomacromolecules 2013;14:1268–77.

Ha JS, Lee YM, Choi SL, Song JJ, Shin CS, Kim JH, et al. Thermostable beta-glycosidase-CBD fusion protein for biochemical analysis of cotton scouring efficiency. J Microbiol Biotechnol 2008;18:443–8.

Haimovitz R, Barak Y, Morag E, Voronov-Goldman M, Shoham Y, Lamed R, et al. Cohesin-dockerin microarray: diverse specificities between two complementary families of interacting protein modules. Proteomics 2008;8:968–79.

Harris JM, Epting KL, Kelly RM. N-terminal fusion of a hyperthermophilic chitin-binding domain to xylose isomerase from Thermotoga neapolitana enhances kinetics and thermostability of both free and immobilized enzymes. Biotechnol Prog 2010;26:993–1000.

Hashimoto H. Recent structural studies of carbohydrate-binding modules. Cell Mol Life Sci 2006;63:2954–67.

Heerze LD, Armstrong GD, Smith R. Anti-inflammatory, tolerogenic and immuno-inhibiting properties of carbohydrate binding-peptides. US Patent No. 2002/0054880 A1.

Henrissat B, Teeri TT, Warren RAJ. A scheme for designating enzymes that hydrolyse the polysaccharides in the cell walls of plants. Febs Lett 1998;425:352–4.

Hong SY, Lee JS, Cho KM, Math RK, Kim YH, Hong SJ, et al. Assembling a novel bifunctional cellulase-xylanase from Thermotoga maritima by end-to-end fusion. Biotechnol Lett 2006;28:1857–62.

Hong J, Ye X, Zhang YH. Quantitative determination of cellulose accessibility to cellulase based on adsorption of a nonhydrolytic fusion protein containing CBM and GFP with its applications. Langmuir 2007;23:12535–40.

Hong J, Wang Y, Ye X, Zhang YH. Simple protein purification through affinity adsorption on regenerated amorphous cellulose followed by intein self-cleavage. J Chromatogr A 2008a;1194:150–4.

Hong J, Ye X, Wang Y, Zhang YH. Bioseparation of recombinant cellulose-binding module-proteins by affinity adsorption on an ultra-high-capacity cellulosic adsorbent. Anal Chim Acta 2008b;621:193–9.

Hsu SH, Chu WP, Lin YS, Chiang YL, Chen DC, Tsai CL. The effect of an RGD-containing fusion protein CBD-RGD in promoting cellular adhesion. J Biotechnol 2004;111:143–54.

Hussack G, Luo Y, Veldhuis L, Hall JC, Tanha J, Mackenzie R. Multivalent anchoring and oriented display of single-domain antibodies on cellulose. Sensors 2009;9:5351–67.

Hwang S, Ahn J, Lee S, Lee TG, Haam S, Lee K, et al. Evaluation of cellulose-binding domain fused to a lipase for the lipase immobilization. Biotechnol Lett 2004;26:603–5.

Jamal S, Nurizzo D, Boraston AB, Davies GJ. X-ray crystal structure of a non-crystalline cellulose-specific carbohydrate-binding module: CBM28. J Mol Biol 2004;339:253–8.

Jervis EJ, Guarna MM, Doheny JG, Haynes CA, Kilburn DG. Dynamic localization and persistent stimulation of factor-dependent cells by a stem cell factor/cellulose binding domain fusion protein. Biotechnol Bioeng 2005;91:314–24.

Kang HJ, Uegaki K, Fukada H, Ishikawa K. Improvement of the enzymatic activity of the hyperthermophilic cellulase from Pyrococcus horikoshii. Extremophiles 2007;11:251–6.

Kang DH, Hyeon JE, You SK, Kim SW, Han SO. Efficient enzymatic degradation process for hydrolysis activity of the Carrageenan from red algae in marine biomass. J Biotechnol 2014;192:108–13.

Kavoosi M, Meijer J, Kwan E, Creagh AL, Kilburn DG, Haynes CA. Inexpensive one-step purification of polypeptides expressed in *Escherichia coli* as fusions with the family 9 carbohydrate-binding module of xylanase 10A from *T. maritima*. J Chromatogr B Analyt Technol Biomed Life Sci 2004;807:87–94.

Kavoosi M, Creagh AL, Kilburn DG, Haynes CA. Strategy for selecting and characterizing linker peptides for CBM9-tagged fusion proteins expressed in *Escherichia coli*. Biotechnol Bioeng 2007a;98:599–610.

Kavoosi M, Lam D, Bryan J, Kilburn DG, Haynes CA. Mechanically stable porous cellulose media for affinity purification of family 9 cellulose-binding module-tagged fusion proteins. J Chromatogr A 2007b;1175:187–96.

Kavoosi M, Sanaie N, Dismer F, Hubbuch J, Kilburn DG, Haynes CA. A novel two-zone protein uptake model for affinity chromatography and its application to the description of elution band profiles of proteins fused to a family 9 cellulose binding module affinity tag. J Chromatogr A 2007c;1160:137–49.

Kavoosi M, Creagh AL, Turner RF, Kilburn DG, Haynes CA. Direct measurement of the kinetics of CBM9 fusion-tag bioprocessing using luminescence resonance energy transfer. Biotechnol Prog 2009;25:874–81.

Kawakubo T, Karita S, Araki Y, Watanabe S, Oyadomari M, Takada R, et al. Analysis of exposed cellulose surfaces in pretreated wood biomass using carbohydrate-binding module (CBM)-cyan fluorescent protein (CFP). Biotechnol Bioeng 2010;105:499–508.

Khan MI, Sajjad M, Sadaf S, Zafar R, Niazi UH, Akhtar MW. The nature of the carbohydrate binding module determines the catalytic efficiency of xylanase Z of *Clostridium thermocellum*. J Biotechnol 2013;168:403–8.

Kim JW, Glasgow JN, Nakayama M, Ak F, Ugai H, Curiel DT. An adenovirus vector incorporating carbohydrate binding domains utilizes glycans for gene transfer. Plos One 2013;8:e55533.

Kim S, Baek SH, Lee K, Hahn JS. Cellulosic ethanol production using a yeast consortium displaying a minicellulosome and beta-glucosidase. Microb Cell Fact 2013;12.

Klocke M, Mundt K, Idler F, Jung S, Backhausen JE. Heterologous expression of enterocin A, a bacteriocin from *Enterococcus faecium*, fused to a cellulose-binding domain in *Escherichia coli* results in a functional protein with inhibitory activity against *Listeria*. Appl Microbiol Biotechnol 2005;67:532–8.

Kwan E, Guarna MM, Boraston AB, Gilkes NR, Haynes CA, Kilburn DG, et al. Self-activating factor X derivative fused to the C-terminus of a cellulose-binding module: production and properties. Biotechnol Bioeng 2002;79:724–32.

Laaksonen P, Walther A, Malho JM, Kainlauri M, Ikkala O, Linder MB. Genetic engineering of biomimetic nanocomposites: diblock proteins, graphene, and nanofibrillated cellulose. Angewandte Chemie 2011;50:8688–91.

Latorre-Garcia L, Adam AC, Manzanares P, Polaina J. Improving the amylolytic activity of *Saccharomyces cerevisiae* glucoamylase by the addition of a starch binding domain. J Biotechnol 2005;118:167–76.

Lehtio J, Sugiyama J, Gustavsson M, Fransson L, Linder M, Teeri TT. The binding specificity and affinity determinants of family 1 and family 3 cellulose binding modules. Proc Natl Acad Sci U S A 2003;100:484–9.

Lemos MA, Teixeira JA, Mota M, Gama FM. A simple method to separate cellulose-binding domains of fungal cellulases after digestion by a protease. Biotechnol Lett 2000;22:703–7.

Levy I, Shoseyov O. Cellulose-binding domains biotechnological applications. Biotechnol Adv 2002;20:191–213.

Levy I, Nussinovitch A, Shpigel E, Shoseyov O. Recombinant cellulose crosslinking protein: a novel paper-modification biomaterial. Cellulose 2002;9:91–8.

Levy I, Paldi T, Shoseyov O. Engineering a bifunctional starch-cellulose cross-bridge protein. Biomaterials 2004;25:1841–9.

Lewis W, Keshavarz-Moore E, Windust J, Bushell D, Parry N. Construction and evaluation of novel fusion proteins for targeted delivery of micro particles to cellulose surfaces. Biotechnol Bioeng 2006;94:625–32.

Liao H, Myung S, Zhang YH. One-step purification and immobilization of thermophilic polyphosphate glucokinase from *Thermobifida fusca* YX: glucose-6-phosphate generation without ATP. Appl Microbiol Biotechnol 2012;93:1109–17.

Lim S, Chundawat SP, Fox BG. Expression, purification and characterization of a functional carbohydrate-binding module from *Streptomyces* sp. SirexAA-E. Protein Expr Purif 2014;98:1–9.

Lin SC, Lin IP, Chou WI, Hsieh CA, Liu SH, Huang RY, et al. CBM21 starch-binding domain: a new purification tag for recombinant protein engineering. Protein Expr Purif 2009; 65:261–6.

Linder M, Teeri TT. The roles and function of cellulose-binding domains. J Biotechnol 1997;57:15–28.

Linder M, Salovuori I, Ruohonen L, Teeri TT. Characterization of a double cellulose-binding domain. Synergistic high affinity binding to crystalline cellulose. J Biol Chem 1996; 271:21268–72.

Lombard V, Ramulu HG, Drula E, Coutinho PM, Henrissat B. The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res 2014;42:D490–5.

Machado J, Araujo A, Pinto R, Gama FM. Studies on the interaction of the carbohydrate binding module 3 from the *Clostridium thermocellum* CipA scaffolding protein with cellulose and paper fibres. Cellulose 2009;16:817–24.

Mahadevan SA, Wi SG, Lee DS, Bae HJ. Site-directed mutagenesis and CBM engineering of Cel5A (*Thermotoga maritima*). FEMS Microbiol Lett 2008;287:205–11.

Malecki PH, Raczynska JE, Vorgias CE, Rypniewski W. Structure of a complete four-domain chitinase from *Moritella marina*, a marine psychrophilic bacterium. Acta Crystallogr D 2013;69:821–9.

Martin M, Wayllace NZ, Valdez HA, Gomez-Casati DF, Busi MV. Improving the glycosyltransferase activity of *Agrobacterium tumefaciens* glycogen synthase by fusion of N-terminal starch binding domains (SBDs). Biochimie 2013;95:1865–70.

Mattinen ML, Linder M, Drakenberg T, Annila A. Solution structure of the cellulose-binding domain of endoglucanase I from *Trichoderma reesei* and its interaction with cello-oligosaccharides. Eur J Biochem 1998;256:279–86.

Maurice S, Dekel M, Shoseyov O, Gertler A. Cellulose beads bound to cellulose binding domain-fused recombinant proteins; an adjuvant system for parenteral vaccination of fish. Vaccine 2003;21:3200–7.

McCartney L, Gilbert HJ, Bolam DN, Boraston AB, Knox JP. Glycoside hydrolase carbohydrate-binding modules as molecular probes for the analysis of plant cell wall polymers. Anal Biochem 2004;326:49–54.

Moreira SM, Gama FM. Carbohydrate binding modules: functions and applications. Handbook of carbohydrate polymers: development, properties and applications. Nova Science Pub Inc.; 2011. p. 201–35.

Moreira SM, Andrade FK, Domingues L, Gama M. Development of a strategy to functionalize a dextrin-based hydrogel for animal cell cultures using a starch-binding module fused to RGD sequence. BMC Biotechnol 2008;8:78.

Moreira SM, Castanheira P, Casal M, Faro C, Gama M. Expression of the functional carbohydrate-binding module (CBM) of human laforin. Protein Expr Purif 2010;74: 169–74.

Myung S, Zhang XZ, Zhang YHP. Ultra-stable phosphoglucose isomerase through immobilization of cellulose-binding module-tagged thermophilic enzyme on low-cost high-capacity cellulosic adsorbent. Biotechnol Prog 2011;27:969–75.

Myung SW, You C, Zhang YHP. Recyclable cellulose-containing magnetic nanoparticles: immobilization of cellulose-binding module-tagged proteins and a synthetic metabolon featuring substrate channeling. J Mater Chem B 2013;1:4419–27.

Nahalka J, Gemeiner P. Thermoswitched immobilization — a novel approach in reversible immobilization. J Biotechnol 2006;123:478–82.

Nordon RE, Shu A, Camacho F, Milthorpe BK. Hollow-fiber assay for ligand-mediated cell adhesion. Cytometry A 2004;57:39–44.

Notenboom V, Boraston AB, Kilburn DG, Rose DR. Crystal structures of the family 9 carbohydrate-binding module from *Thermotoga maritima* xylanase 10A in native and ligand-bound forms. Biochemistry-Us 2001;40:6248–56.

Ofir K, Berdichevsky Y, Benhar I, Azriel-Rosenfeld R, Lamed R, Barak Y, et al. Versatile protein microarray based on carbohydrate-binding modules. Proteomics 2005;5: 1806–14.

Ong E, Greenwood JM, Gilkes NR, Kilburn DG, Miller RC, Warren RA. The cellulose-binding domains of cellulases — tools for biotechnology. Trends Biotechnol 1989;7:239–43.

Ong E, Alimonti JB, Greenwood JM, Miller Jr RC, Warren RA, Kilburn DG. Purification of human interleukin-2 using the cellulose-binding domain of a prokaryotic cellulase. Bioseparation 1995;5:95–104.

Pala H, Lemos MA, Mota M, Gama FM. Enzymatic upgrade of old paperboard containers. Enzyme Microb Tech 2001;29:274–9.

Pangu G, Johnston E, Petkov J, Parry N, Leach M, Hammer DA. Targeted particulate adhesion to cellulose surfaces mediated by bifunctional fusion proteins. Langmuir 2007; 23:10682–93.

Pertile R, Moreira S, Andrade F, Domingues L, Gama M. Bacterial cellulose modified using recombinant proteins to improve neuronal and mesenchymal cell adhesion. Biotechnol Prog 2012;28:526–32.

Pinto R, Moreira S, Mota M, Gama M. Studies on the cellulose-binding domains adsorption to cellulose. Langmuir 2004;20:1409–13.

Qi M, O'Brien JP, Yang J. A recombinant triblock protein polymer with dispersant and binding properties for digital printing. Biopolymers 2008;90:28–36.

Raghothama S, Simpson PJ, Szabo L, Nagy T, Gilbert HJ, Williamson MP. Solution structure of the CBM10 cellulose binding module from *Pseudomonas xylanase* A. Biochemistry-Us 2000;39:978–84.

Ramos R, Pinto R, Mota M, Sampaio L, Gama FM. Textile depilling: superior finishing using cellulose-binding domains with residual enzymatic activity. Biocatal Biotransform 2007;25:35–42.

Ramos R, Domingues L, Gama M. *Escherichia coli* expression and purification of LL37 fused to a family III carbohydrate-binding module from *Clostridium thermocellum*. Protein Expr Purif 2010;71:1–7.

Ramos R, Silva JP, Rodrigues AC, Costa R, Guardao L, Schmitt F, et al. Wound healing activity of the human antimicrobial peptide LL37. Peptides 2011;32:1469–76.

Ramos R, Moreira S, Rodrigues A, Gama M, Domingues L. Recombinant expression and purification of the antimicrobial peptide magainin-2. Biotechnol Prog 2013;29: 17–22.

Rao FV, Houston DR, Boot RG, Aerts JMFG, Hodkinson M, Adams DJ, et al. Specificity and affinity of natural product cyclopentapeptide inhibitors against *A. fumigatus*, human, and bacterial chitinases. Chem Biol 2005;12:65–76.

Reed DC, Barnard GC, Anderson EB, Klein LT, Gerngross TU. Production and purification of self-assembling peptides in *Ralstonia eutropha*. Protein Expr Purif 2006;46:179–88.

Reyes-Ortiz V, Heins RA, Cheng G, Kim EY, Vernon BC, Elandt RB, et al. Addition of a carbohydrate-binding module enhances cellulase penetration into cellulose substrates. Biotechnol Biofuels 2013;6:93.

Ribeiro T, Ponte PI, Guerreiro CI, Santos HM, Falcao L, Freire JP, et al. A family 11 carbohydrate-binding module (CBM) improves the efficacy of a recombinant cellulase used to supplement barley-based diets for broilers at lower dosage rates. Br Poult Sci 2008;49:600–8.

Richins RD, Mulchandani A, Chen W. Expression, immobilization, and enzymatic characterization of cellulose-binding domain-organophosphorus hydrolase fusion enzymes. Biotechnol Bioeng 2000;69:591–6.

Rodriguez B, Kavoosi M, Koska J, Creagh AL, Kilburn DG, Haynes CA. Inexpensive and generic affinity purification of recombinant proteins using a family 2a CBM fusion tag. Biotechnol Prog 2004;20:1479–89.

Rotticci-Mulder JC, Gustavsson M, Holmquist M, Hult K, Martinelle M. Expression in *Pichia pastoris* of *Candida antarctica* lipase B and lipase B fused to a cellulose-binding domain. Protein Expr Purif 2001;21:386–92.

Safra-Dassa L, Shani Z, Danin A, Roiz L, Shoseyov O, Wolf S. Growth modulation of transgenic potato plants by heterologous expression of bacterial carbohydrate-binding module. Mol Breeding 2006;17:355–64.

Schnorr KM, Christensen LLH. Carbohydrate-binding modules of a new family. WO Patent No. 2005/042735 A1.

Shi X, Zheng F, Pan R, Wang J, Ding S. Engineering and comparative characteristics of double carbohydrate binding modules as a strenght additive for papermaking applications. Bioresources 2014;9:3117–31.

Shoseyov O, Shani Z, Levy I. Carbohydrate binding modules: biochemical properties and novel applications. Microbiol Mol Biol R 2006;70:283–95.

Shpigel E, Elias D, Cohen IR, Shoseyov O. Production and purification of a recombinant human hsp60 epitope using the cellulose-binding domain in Escherichia coli. Protein Expr Purif 1998;14:185–91.

Shpigel E, Goldlust A, Eshel A, Ber IK, Efroni G, Singer Y, et al. Expression, purification and applications of staphylococcal protein A fused to cellulose-binding domain. Biotechnol Appl Biochem 2000;31(Pt 3):197–203.

Simpson PJ, Xie HF, Bolam DN, Gilbert HJ, Williamson MP. The structural basis for the ligand specificity of family 2 carbohydrate-binding modules. J Biol Chem 2000;275: 41137–42.

Siroky J, Benians TAS, Russell SJ, Bechtold T, Knox JP, Blackburn RS. Analysis of crystallinity changes in cellulose II polymers using carbohydrate-binding modules. Carbohydr Polym 2012;89:213–21.

Slutzki M, Ruimy V, Morag E, Barak Y, Haimovitz R, Lamed R, et al. High-throughput screening of cohesin mutant libraries on cellulose microarrays. Methods Enzymol 2012;510:453–63.

Sugimoto N, Igarashi K, Samejima M. Cellulose affinity purification of fusion proteins tagged with fungal family 1 cellulose-binding domain. Protein Expr Purif 2012;82: 290–6.

Tang CD, Li JF, Wei XH, Min R, Gao SJ, Wang JQ, et al. Fusing a carbohydrate-binding module into the Aspergillus usamii beta-mannanase to improve its thermostability and cellulose-binding capacity by in silico design. Plos One 2013;8:e64766.

Tarahomjoo S, Katakura Y, Shioya S. New strategy for enhancement of microbial viability in simulated gastric conditions based on display of starch-binding domain on cell surface. J Biosci Bioeng 2008;105:503–7.

Terpe K. Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. Appl Microbiol Biotechnol 2003;60:523–33.

Thongekkaew J, Ikeda H, Masaki K, Iefuji H. Fusion of cellulose binding domain from Trichoderma reesei CBHI to Cryptococcus sp. S-2 cellulase enhances its binding affinity and its cellulolytic activity to insoluble cellulosic substrates. Enzyme Microb Technol 2013;52:241–6.

Tilbeurgh HV, Tomme P, Claeyssens M, Bhikhabhai R, Pettersson G. Limited proteolysis of the cellobiohydrolase I from Trichoderma reesei — separation of functional domains. Febs Lett 1986;204:223–7.

Tolba M, Minikh O, Brovko LY, Evoy S, Griffiths MW. Oriented immobilization of bacteriophages for biosensor applications. J Appl Environ Microbiol 2010;76:528–35.

Tomme P, Vantilbeurgh H, Pettersson G, Vandamme J, Vandekerckhove J, Knowles J, et al. Studies of the cellulolytic system of Trichoderma reesei Qm-9414 — analysis of domain function in 2 cellobiohydrolases by limited proteolysis. Eur J Biochem 1988;170: 575–81.

Tomme P, Driver DP, Amandoron EA, Miller RC, Antony R, Warren J, et al. Comparison of a fungal (Family-I) and bacterial (Family-II) cellulose-binding domain. J Bacteriol 1995; 177:4356–63.

Tomme P, Boraston A, McLean B, Kormos J, Creagh AL, Sturch K, et al. Characterization and affinity applications of cellulose-binding domains. J Chromatogr B Biomed Sci Appl 1998;715:283–96.

Vaaje-Kolstad G, Horn SJ, van Aalten DMF, Synstad B, Eijsink VGH. The non-catalytic chitin-binding protein CBP21 from Serratia marcescens is essential for chitin degradation. J Biol Chem 2005;280:28492–7.

Valo H, Kovalainen M, Laaksonen P, Hakkinen M, Auriola S, Peltonen L, et al. Immobilization of protein-coated drug nanoparticles in nanofibrillar cellulose matrices-enhanced stability and release. J Control Release 2011;156:390–7.

Varjonen S, Laaksonen P, Paananen A, Valo H, Hahl H, Laaksonen T, et al. Self-assembly of cellulose nanofibrils by genetically engineered fusion proteins. Soft Matter 2011;7: 2402–11.

Volkov IY, Lunina NA, Velikodvorskaya GA. Prospects for the practical application of substrate-binding modules of glycosyl hydrolases. Appl, Biochem Micro + 2004;40:427–32.

von Ossowski I, Eaton JT, Czjzek M, Perkins SJ, Frandsen TP, Schulein M, et al. Protein disorder: conformational distribution of the flexible linker in a chimeric double cellulase. Biophys J 2005;88:2823–32.

Wan W, Wang D, Gao X, Hong J. Expression of family 3 cellulose-binding module (CBM3) as an affinity tag for recombinant proteins in yeast. Appl Microbiol Biotechnol 2011; 91:789–98.

Wang JY, Stuckey JA, Wishart MJ, Dixon JE. A unique carbohydrate binding domain targets the Lafora disease phosphatase to glycogen. J Biol Chem 2002;277:2377–80.

Wang LS, Zhang YZ, Gao PJ. A novel function for the cellulose binding module of cellobiohydrolase I. Sci China Ser C 2008;51:620–9.

Wang S, Cui GZ, Song XF, Feng Y, Cui Q. Efficiency and stability enhancement of cis-epoxysuccinic acid hydrolase by fusion with a carbohydrate binding module and immobilization onto cellulose. Appl Biochem Biotechnol 2012;168:708–17.

Wierzba A, Reichl U, Turner RF, Warren RA, Kilburn DG. Adhesion of mammalian cells to a recombinant attachment factor, CBD/RGD, analyzed by image analysis. Biotechnol Bioeng 1995a;46:185–93.

Wierzba A, Reichl U, Turner RF, Warren RA, Kilburn DG. Production and properties of a bifunctional fusion protein that mediates attachment of vero cells to cellulosic matrices. Biotechnol Bioeng 1995b;47:147–54.

Xu Z, Bae W, Mulchandani A, Mehra RK, Chen W. Heavy metal removal by novel CBD-EC20 sorbents immobilized on cellulose. Biomacromolecules 2002;3:462–5.

Xu Q, Song Q, Ai X, McDonald TJ, Long H, Ding SY, et al. Engineered carbohydrate-binding module (CBM) protein-suspended single-walled carbon nanotubes in water. Chem Commun 2009;337–9.

Yaniv O, Fichman G, Borovok I, Shoham Y, Bayer EA, Lamed R, et al. Fine-structural variance of family 3 carbohydrate-binding modules as extracellular biomass-sensing components of Clostridium thermocellum anti-sigma(I) factors. Acta Crystallogr D 2014;70:522–34.

Ye X, Zhu Z, Zhang C, Zhang YH. Fusion of a family 9 cellulose-binding module improves catalytic potential of Clostridium thermocellum cellodextrin phosphorylase on insoluble cellulose. Appl Microbiol Biotechnol 2011;92:551–60.

Yeh M, Craig S, Lum MG, Foong FC. Effects of the PT region of EngD and HLD of CbpA on solubility, catalytic activity and purification characteristics of EngD-CBD(CbpA) fusions from Clostridium cellulovorans. J Biotechnol 2005;116:233–44.

You C, Zhang YHP. Self-assembly of synthetic metabolons through synthetic protein scaffolds: one-step purification, co-immobilization, and substrate channeling. Acs Synth Biol 2013;2:102–10.

You C, Myung S, Zhang YH. Facilitated substrate channeling in a self-assembled trifunctional enzyme complex. Angewandte Chemie 2012;51:8787–90.

You C, Chen H, Myung S, Sathitsuksanoh N, Ma H, Zhang XZ, et al. Enzymatic transformation of nonfood biomass to starch. Proc Natl Acad Sci U S A 2013;110:7182–7.

Zhang Y, Chen S, Xu M, Cavaco-Paulo A, Wu J, Chen J. Characterization of Thermobifida fusca cutinase-carbohydrate-binding module fusion proteins and their potential application in bioscouring. Appl Environ Microbiol 2010;76:6870–6.

Zhang Y, Wang L, Chen J, Wu J. Enhanced activity toward PET by site-directed mutagenesis of Thermobifida fusca cutinase-CBM fusion protein. Carbohydr Polym 2013;97: 124–9.