ELSEVIER

Journal of Chromatography B, 715 (1998) 283–296

JOURNAL OF CHROMATOGRAPHY B

Characterization and affinity applications of cellulose-binding domains[1]

Peter Tomme[a,b,c], Al Boraston[a,b,c], Brad McLean[a,b,c], Jeff Kormos[a,b,c], A. Louise Creagh[a,b,d], Ken Sturch[a,d], Neil R. Gilkes[a,c], Charles A. Haynes[a,b,d], R. Antony J. Warren[a,c], Douglas G. Kilburn[a,b,c,*]

[a]*Protein Engineering Networks of Centres of Excellence (PENCE), University of British Columbia, 6174 University Boulevard, Vancouver, B.C. V6T 1Z3, Canada*

[b]*The Biotechnology Laboratory, University of British Columbia, 6174 University Boulevard, Vancouver, B.C. V6T 1Z3, Canada*

[c]*Department of Microbiology and Immunology, University of British Columbia, 6174 University Boulevard, Vancouver, B.C. V6T 1Z3, Canada*

[d]*Department of Chemical Engineering, University of British Columbia, 6174 University Boulevard, Vancouver, B.C. V6T 1Z3, Canada*

Abstract

Cellulose-binding domains (CBDs) are discrete protein modules found in a large number of carbohydrolases and a few nonhydrolytic proteins. To date, almost 200 sequences can be classified in 13 different families with distinctly different properties. CBDs vary in size from 4 to 20 kDa and occur at different positions within the polypeptides; N-terminal, C-terminal and internal. They have a moderately high and specific affinity for insoluble or soluble cellulosics with dissociation constants in the low micromolar range. Some CBDs bind irreversibly to cellulose and can be used for applications involving immobilization, others bind reversibly and are more useful for separations and purifications. Dependent on the CBD used, desorption from the matrix can be promoted under various different conditions including denaturants (urea, high pH), water, or specific competitive ligands (e.g. cellobiose). Family I and IV CBDs bind reversibly to cellulose in contrast to family II and III CBDs which are in general, irreversibly bound. The binding of family II CBDs (CBD_{Cex}) to crystalline cellulose is characterized by a large favourable increase in entropy indicating that dehydration of the sorbent and the protein are the major driving forces for binding. In contrast, binding of family IV CBDs (CBD_{N1}) to amorphous or soluble cellulosics is driven by a favourable change in enthalpy which is partially offset by an unfavourable entropy change. Hydrogen bond formation and van der Waals interactions are the main driving forces for binding. CBDs with affinity for crystalline cellulose are useful tags for classical column affinity chromatography. The affinity of CBD_{N1} for soluble cellulosics makes it suitable for use in large-scale aqueous two-phase affinity partitioning systems. © 1998 Elsevier Science B.V. All rights reserved.

Keywords: Cellulose-binding domains; Cellulose affinity; Proteins

*Corresponding author. Address for correspondence: Biotechnology Laboratory and Dept. of Microbiology and Immunology, University of British Columbia, 237-6174 University Boulevard, Vancouver, B.C. V6T 1Z3, Canada. Tel.: (604) 822-4182; fax: (604) 822-2114; e-mail: kilburn@unixg.ubc.ca

[1]Presented at the 2nd Conference on Affinity Technology, Arlington, VA, USA, September 29–30, 1997.

0378-4347/98/$19.00 © 1998 Elsevier Science B.V. All rights reserved.
PII: S0378-4347(98)00053-X

1. Introduction

Purification and recovery of biologically active molecules are important aspects of biotechnology and are a major consideration in the design of fermentation processes. The continued growth and maturation of the pharmaceutical and biotechnology industries has created an increasing need for practical and cost-efficient large-scale processing techniques. The challenges of producing ever-increasing amounts of (recombinant) biomolecules of extremely high purity and bioactivity from a complex background of contaminants, have inspired the development of new or improved separation techniques. Desired characteristics for bioseparations include ease of use, reproducibility, high capacity and high selectivity for the target molecules.

Biospecific affinity purification holds an important place in the repertoire of protein and peptide purification techniques and is becoming increasingly popular and powerful due to rapid advances in molecular biology. In particular, gene fusion technology has expanded the range of useful and general biological interaction systems. Recombinant hybrids containing a polypeptide fusion partner, termed 'affinity tag', to facilitate the purification of the target polypeptides are now widely used [1,2]. Besides allowing the affinity purification of the target polypeptides, these affinity tags can also increase the in vivo proteolytic stability, modulate the solubility or control the cellular localization of the target polypeptides in the expression host [1]. In addition, these tags can be used for immobilization of biologically active molecules in continuous bioprocessing applications. To date, numerous affinity tags have been developed [3,4]. These vary from whole proteins and protein domains to poly- or single amino acid residues. Some of the more commonly used tags include: *Staphylococcus aureus* protein A or its synthetic two domain variant ZZ [5,6]; glutathione S-transferase from *Schistosoma japonicum* [7]; the maltose-binding protein MalE from *Escherichia coli* [8]; the S-peptide from ribonuclease A [9,10]; biotin [11,12]; Streptavidin [13]; and the popular polyamino acid affinity tags, in particular polyhistidines, which form the basis of immobilized metal ion affinity chromatography (IMAC) [14,15]. More recently, with the use of phage display, this repertoire has been extended considerably through the development and identification, from large combinatorial peptide libraries, of numerous peptide mimics with high, specific affinity for particular target molecules [16]. These technologies allow us to tailor the purification process to specific problems or to particular target molecules which are otherwise hard to purify to homogeneity by more traditional methods.

Although each affinity system offers particular advantages, many share the common drawback of the high cost of the affinity matrix which limits the use for large scale applications. This cost is associated with rather complex, expensive supports or the requirement for complicated chemical modifications to cross-link the solid supports or to covalently attach the ligands to them.

Cellulose has a number of advantages which makes it an ideal matrix for large scale affinity purposes; it is cheap, has excellent physical properties, is inert and has a low, nonspecific affinity for most proteins. Furthermore, cellulose is commercially available in many different forms: cotton, cloth, filters, membranes, powder, fibers, beads, hydrogels, sheets of defined porosity and as soluble polymers with various degrees of polymerization, substitution and viscosity. Finally, cellulosics are safe and have been approved for many pharmaceutical and human uses.

This paper describes the characterization and use of cellulose-binding domains (CBDs) as alternative and highly versatile tags for affinity applications based on their high and specific affinity for cellulose and the related polysaccharide chitin. Different CBDs, with different properties and specificities were characterized, and are described. Different systems and applications for the production, purification and immobilization of polypeptides, based on the different properties of these CBDs, are discussed.

2. Experimental

2.1. Chemicals and reagents

Microcrystalline cellulose (Avicel PH101) was from FMC International (Little Island, County Cork, Ireland). Cellulose CF1 and hydroxypropylmethyl cellulose (HPMC, viscosity ~50 cSt for a 2% (w/v)

solution) was from Sigma Chemical Company (St. Louis, MO, USA). Bacterial microcrystalline cellulose (BMCC) was prepared from cultures of *Acetobacter xylinum* (ATCC 23769) as described previously [17]. Acid-swollen cellulose (PASC) was obtained by phosphoric acid treatment of Avicel PH101, as reported previously [18]. Dextran T500 was from Pharmacia Biotech (Uppsala, Sweden). Guanidinium hydrochloride (Ultra pure) was from ICN Biomedicals (Aurora, OH, USA) and urea (AnalaR, min 99.5% pure) was obtained from BDH (Darmstadt, Germany). All buffers were purchased from Sigma Chemical Company or BDH and were of the highest purity available.

2.2. Equipment

A LCC-500 Plus fast system (Pharmacia Biotech, Uppsala, Sweden) equipped with an UV-M monitor, a FRAC-100 fraction collector and a Pharmacia LKB-Rec-102 recorder was used for low-pressure liquid protein chromatography. Conductivity of the column fractions and the buffer solutions was measured with a hand held model 604 conductivity meter (VWR Scientific, Mississauga, Canada). UV absorption measurements were performed in a Hitachi U-2000 spectrophotometer. Isothermal titration calorimetry experiments (ITC) for binding of CBD_{Cex} and CBD_{N1} to their respective sugar ligands were carried out in a Calorimetry Science (Provo, UT, USA) model 4200 isothermal titration calorimeter as described previously [19,20]. A Calorimetry Science model 4215 isothermal differential scanning calorimeter (DSC) was used for DSC measurements [21]. SDS–PAGE acrylamide electrophoresis was done on a Mini-PROTEAN II system (BioRad) according to the manufacture's instructions.

2.3. Cellulose—affinity purification of CBD_{CenA}-PhoA

A Pharmacia XK 50/30 column was packed with 20 g of cellulose (CF1, Sigma), previously washed with water to remove fines, giving a bed volume of approximately 100 ml. After equilibration (4°C) of the column with 50 m*M* Tris·HCl buffer, pH 7.5 (300 ml), *E. coli* culture supernatant or cell extract (90 ml) containing ~20 mg of fusion protein was passed through the column at a flow-rate of 1 ml·min^{-1}. The column was washed with 300 ml of 50 m*M* Tris·HCl buffer, pH 7.0 containing 1 *M* NaCl, 0.02% NaN_3 (high-salt buffer), followed by a linear gradient (100 ml) from high-salt buffer to 50 m*M* Tris. HCl buffer, pH 7.5, 0.02% NaN_3 (low-salt buffer), and finally 150 ml of low-salt buffer. Adsorbed protein was eluted with 200 ml of distilled water or with a 200-ml linear gradient from 0–6 *M* guanidinium hydrochloride (Gnd·HCl) in 50 m*M* Tris·HCl buffer, pH 7.5. Fractions of 15 ml were collected throughout and monitored for $A_{280\ nm}$ and conductivity. To test the performance of the cellulose column during multiple rounds of CBD_{CenA}-PhoA purification, columns were run as described above but protein was eluted with a 200-ml linear gradient from low-salt buffer to distilled water, followed by 200 ml of distilled water. The column was then re-equilibrated with 50 m*M* Tris·HCl buffer, pH 7.5 (300 ml) and a new cycle was started.

2.4. Purification of cellulose-binding domains

Overnight cultures of *E. coli* strain JM101, harboring pTugN1n [22] were diluted 500-fold in tryptone-yeast extract-phosphate medium (TYP) [23] supplemented with 100 μg kanamycin per ml, and grown at 30°C to an $A_{600\ nm}$ of 2.0. Transcription of the gene fragment encoding CBD_{N1} was induced by the addition of isopropyl-1-thio-β-D-galactopyranoside (IPTG) to a final concentration of 0.1 m*M* and growth was continued for a further 12 to 18 h at 30°C. To each liter of culture supernatant, clarified by centrifugation at 4°C for 10 min at 13 000 *g*, 50 g of Avicel washed in 50 m*M* potassium phosphate buffer, pH 7.0 was added. After incubation overnight at 4°C with stirring, Avicel was recovered by vacuum filtration on a Whatman GF/A or GF/C glass filter and washed with 100 ml of 1 *M* sodium chloride in 50 m*M* potassium phosphate buffer, pH 7.0 followed by 150 ml of 50 m*M* potassium phosphate buffer, pH 7.0. CBD_{N1} was eluted from the Avicel with distilled water (approximately 1 l). Recoveries were increased by recombining the initial filtrate and the buffer washes with the Avicel. The above process was repeated and the two fractions were combined, adjusted to pH 6.0 with potassium

phosphate buffer (20 m*M* final), and further purified by anion-exchange chromatography on MacroQ (BioRad, Mississauga, Canada). A Pharmacia XK26 column, packed with 50 ml of MacroQ resin, was equilibrated overnight at 4°C with 300 ml of potassium phosphate buffer (20 m*M*), pH 6.0. A volume of 250 ml of Avicel-purified CBD_{N1}, containing approximately 90 mg of protein, was passed through the column at a flow-rate of 1 ml·min^{-1}. The column was washed with 300 ml of equilibration buffer (1 ml·min^{-1}) and CBD_{N1} was eluted (1 ml·min^{-1}) with a linear gradient (600 ml) of 0–1 *M* sodium chloride in 20 m*M* potassium phosphate buffer, pH 6.0. Fractions (10 ml) were screened for protein by UV absorbance at 280 nm and analysed for purity by SDS–PAGE on 12% polyacrylamide gels. Pure CBD_{N1} fractions were pooled, desalted, exchanged into the appropriate buffer and concentrated by ultrafiltration on a 1-K filter (Filtron Technology).

Recombinant CBD_{Cex} containing a N-terminal hexa-histidine tail, was produced in *E. coli* JM101 essentially as described for CBD_{N1} except that cultures were grown at 37°C. The CBD was purified from the culture supernatant by immobilized metal affinity chromatography (IMAC). A small BioRad column (1.5×10 cm) was packed with 10 ml of Ni^{2+} agarose (Novagen, Madison, WI, USA) to give a final bed volume of ~5 ml. All subsequent operations were done at room temperature and at flow-rates of 2 ml·min^{-1} using a Pharmacia P1 peristaltic pump. The column was washed with 25 ml of distilled water, charged with 50 ml of charge buffer (50 m*M* $NiSO_4 \cdot 6H_20$) and equilibrated with 50 ml of binding buffer (20 m*M* Tris·HCl buffer, pH 7.9 containing 10 m*M* imidazole and 0.5 *M* NaCl). A volume of 100 ml of concentrated culture supernatant, diluted 1:10 in binding buffer, was loaded on the column and pure CBD was recovered by stepwise elution with 10 ml of 25, 30, 35, 40, 50, 75 and 150 m*M* imidazole in 20 m*M* Tris·HCl buffer, pH 7.9 containing 0.5 *M* NaCl. The column was then regenerated by stripping the Ni^{2+} off the column with 5 column-volumes of 100 m*M* EDTA in 20 m*M* Tris·HCl buffer, pH 7.9–0.5 *M* NaCl followed by 5 column-volumes of 6 *M* guanidinium hydrochloride in the same buffer. The column was then recycled as described above. Protein fractions were treated as described for CBD_{N1} except that 16% acrylamide gels were used to evaluate the purity of CBD_{Cex}.

CBD_{Clos} (CbpA) from *Clostridium cellulovorans* was a gift from CBDTechnologies and was purified using published strategies [24].

2.5. Cellulose-binding experiments and determination of equilibrium binding constants

Depletion isotherms were obtained by incubating 1 mg of cellulose (BMCC, Avicel or PASC) in a 1.5-ml centrifuge tube with various amounts of CBD (1–50 μ*M*) in 50 m*M* potassium phosphate buffer pH 7.0 (1 ml total volume). Enzyme controls without cellulose were included. All experiments were done in triplicate. Samples were incubated with end-over-end mixing at 4°C for 4 up to 24 h (unless stated otherwise). After incubation, the cellulose was pelleted by centrifugation at 4°C (2×10 min, 16 000 *g*) and the concentration of bound CBD was calculated from the difference in total and free CBD concentration, determined spectrophotometrically ($A_{280\ nm}$) in the supernatants before (enzyme controls) and after addition of cellulose, respectively.

Equilibrium binding constants (K_a) were obtained from the depletion isotherms (Plot of [B] versus [F]) after fitting (nonlinear regression) of the raw data to a Langmuir type adsorption model

$$[B] = [N_o]K_a[F]/1 + K_a[F]$$

where [B] is the concentration of bound CBD (mol per g cellulose), [F] is the concentration of free CBD (molar), $[N_o]$ is the total concentration of available binding sites on the cellulose surface (mol per g cellulose and K_a is the equilibrium association constant (M^{-1}).

To determine the reversibility of the binding reaction, all of the supernatant was carefully removed from the cellulose pellet after binding and quantitation of bound protein, and replaced by an equal volume (usually 1 ml) of 50 m*M* potassium phosphate buffer, pH 7.0. Samples were incubated for another 5 or 6 h at 4°C and treated as described above to determine the concentration of CBD released into the supernatant.

Conditions for desorption of CBDs bound to cellulose were evaluated as follows; 10 mg Avicel was incubated (4 h at 4°C) with 100 μg of pure CBD

in 50 m*M* potassium phosphate buffer, pH 7.0 (1 ml total volume) in 1.5-ml centrifuge tubes. The Avicel suspension was then packed into 1-ml syringes, plugged at the bottom with glass-wool, and the microcolumns were washed three times with 1-ml buffer (pH 7.0) using gravity flow. Desorption of bound CBD was attempted by washing the columns three times with 1 ml of distilled water; urea (1–8 *M*); Gnd·HCl (1–6 *M*); SDS (1–10% w/v); NaOH (0.1–2 N); triethylamine (1%); 50 m*M* Tris, pH 11.5; or cellobiose (0.1 *M*). After washing the cellulose three times with 1 ml of buffer, pH 7.0 to remove eluent, the Avicel was transferred to a 1.5-ml centrifuge tube, resuspended in 40 μl loading buffer, boiled for 5 min to desorp residual bound CBD, and centrifuged (5 min, 14 500 *g*) to pellet the cellulose. A 20-μl volume of the supernatant was assayed by SDS–PAGE on 12–16% polyacrylamide gels for the presence of CBD.

2.6. Construction of ternary phase diagrams

Ternary phase diagrams for mixtures of hydroxypropylmethyl cellulose and Dextran T500 in 50 m*M* phosphate buffer, pH 7.0 and 25°C were obtained according to the procedures described in [25].

3. Results

3.1. Families of cellulose-binding domains and their characteristics

Cellulose-binding domains are found in nature as discrete domains in cellulases and xylanases, proteins which are involved in the degradation of plant biomass [26,27], and in proteins without hydrolytic activity such, as the clostridial cellulose-integrating or binding proteins Cip and Cbp [24,27,28] (Fig. 1). In the polysaccharidases, CBDs are involved in targeting the enzymes to their substrates or particular substrate regions, thereby increasing the effective enzyme concentration at the surface of the insoluble cellulose. In Cbp or Cip, the CBD is part of a scaffolding subunit that organizes the catalytic subunits in a cohesive multienzyme complex called the 'cellulosome', which then adheres strongly to cellulose [29].

Currently, more than 180 putative CBD sequences

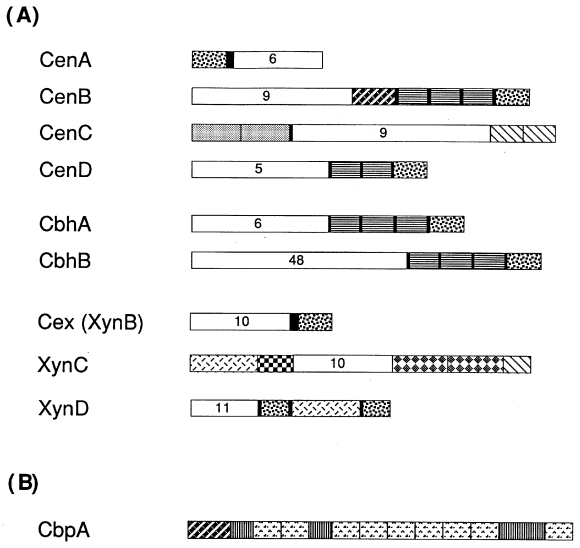

Fig. 1. Graphic representation of the domain structure of (A) *Cellulomonas fimi* cellulases and xylanases and (B) the *Clostridium cellulovorans* scaffolding protein CbpA. Each protein is represented as a linear map showing the arrangements of the various different structural and functional domains. CenA, CenB, CenC and CenD are β-1,4-glucanases, CbhA and CbhB are cellobiohydrolases, Cex (XynB) is a mixed exo-glucanase/xylanase and XynC and XynD are β-1,4-endoxylanases. CBDII, C:BDlII, CBDIV and CBDIX refer to the family II, III, IV or IX cellulose-binding domains. Fn3 and NodB indicate a fibronectin type III-like and a nodulation B-like domain, respectively. Catalytic domain; CBDII; CBDIII; CBDIV; CBDIX; Fn3 module; NodB; thermostabilizing domain; cohesion domain; hydrophilic domain (function unknown); other; linker.

can be classified in 13 different families [27,30] (Table 1). Some families contain only one (e.g. family V and VIII) or a few members, whereas others (e.g. family I, II and III) are large families of forty or more sequences. Members of the same family are expected to have very similar properties, although there may be subtle differences (see also Table 2). For instance, the family I CBD from *Trichoderma reesei* cellobiohydrolase II (CBD_{CbhII}) binds chitin with relatively strong affinity [31] whereas the homologous family I CBD from *T. reesei* cellobiohydrolase I (CBD_{CbhI}) has little or no affinity for this polysaccharide [31,32]. Members of different families seem to have different properties (see below) although overlap between some families cannot be ruled out.

Table 1
CBD families and their properties updated from [30]

Family	Size[a]	Characteristics
I	33–36	Exclusively CBDs from fungal enzymes. >40 members. One structure for CBD_{CbhI} (*T. reesei*). Some bind to chitin.
II	~100	Two sub-families (IIa and IIb). IIa contains two chitin-binding domains. ~40 members. One structure for CBD_{Cex} (*C.fimi*). Some IIa CBDs bind chitin, IIb CBDs show affinity for xylan.
III	130–170	Two sub-families (IIIa and IIIb). ~25 members. One structure for CBD_{Cip} (*Cl. thermocellum*). Some bind to chitin.
IV	125–170	5–6 members. One structure for CBD_{N1} (*C.fimi*). Does not bind crystalline cellulose.
V	63	CBD_{EGZ} from *Erwinia chrysanthemi* only. Structure solved.
VI	85–90	6 members. Low affinity for crystalline cellulose.
VII		Entry deleted.
VIII	152	CBD_{CelA} from *Dictyostelium discoidem* only.
IX	170–180	9 members, from thermostable xylanases. CBDs occur mainly as tandem repeats.
X	50–55	7 members almost exclusively from *Ps. fluorescens.*
XI	120–180	4 members, mainly from *Clostridium* and *Fibrobacter* enzymes.
XII	~50	>10 members. Mainly sequences derived from *Bacilli* endoglucanases and from various chitinases.
XllI	40–45	Contains triple repeated CBDs from xylanases and lectin like domains with different specificites.

[a] Approximate sizes (no. of amino acids) based on sequence similarities.

Table 2
Differences in properties of the homologous family III cellulose-binding domains isolated from the nonhydrolytic, clostridial scaffolding proteins

CBD	Organism	Expressed[a]	Elution[b]	Ref.
CBD_{CbpA}	*C. cellulovorans*	Insoluble	Denaturants[c]	[28]
$CBD_{CipA(B)}$	*C. thermocellum*	Soluble	Denaturants[d]	[39,40]
CBD_{CipC}	*C. cellulolyticum*	Soluble	Water	[38,55]

[a] Expressed as recombinant protein in *E. coli*.
[b] Elution conditions for desorption from Avicel.
[c] At least 1.5 *M* urea is required for desorption. Other denaturants and elution conditions are 6 *M* guanidinium hydrochloride or high pH (11.5).
[d] More than 3 *M* urea is required for desorption. Other denaturants and elution conditions are 10% SDS or 1% triethylamine (pH 11.5).

CBDs range in size from small, fungal domains of 33–36 amino acids (family I CBDs) to larger domains of up to 180 residues (family III). This wide range in domain sizes is advantageous when selecting tags to minimize interference with the biological activity of the target molecule or when selecting a fusion partner which can stabilize these target molecules. Small tags are usually most suited for this purpose but the larger CBDs can be used with equal success when small linker sequences are introduced between the two fusion partners [33,34] (Table 3). An added bonus is the proteolytic resistance and stability of the compactly folded cellulose-binding domains.

In nature, CBDs occur at various positions within the parental polypeptides (see Fig. 1), a useful feature when constructing CBD-hybrids. Depending on the molecule of interest, N- or C-terminal CBDs can be selected. This is particularly important for those biomolecules which can only tolerate foreign sequences at one end to retain activity. Thus, unlike most other protein tags such as protein A or glutathione S-transferase, CBDs can be shuffled around and offer the same flexibility as small peptide tags. Good candidates for a N- and C-terminal tag with similar properties are CBD_{CenA} from *Cellulomonas fimi* endoglucanase A and CBD_{Cex} from *C. fimi* exoglucanase/xylanase, respectively (Fig. 1, Table 3).

As shown by the molecular architecture of the *C. fimi* enzymes CenB and XynD, CBDs also occur internally (Fig. 1). This indicates that stable polyfunctional chimerics, with a functional internal CBD for immobilization to cellulose, can be created through gene fusion technology [35]. The close proximity of the different active centres (or enzymes) when encoded in one polypeptide, enhances the specificity and rate of the (coupled) reactions, while

Table 3
List of some of the recombinant CBD-hybrids produced for affinity applications

CBD	Fusion	Partner	Expression system
CenA	CBD_{CenA}-PhoA	Alkaline phosphatase	*E. coli*
	CBD_{CenA}-IL2	Human interleukin 2	Mammalian cells *Pichia pastoris*
	CBD_{CenA}-IL3	Human interleukin 3	Mammalian cells
	CBD_{CenA}-Benzonase™	Endonuclease	*E. coli*
	CBD_{CenA}-RGD	Adhesion peptide	*E. coli*
	XynA-CBD_{CenA}-CD_{CenA}	Xylanase	*E. coli*
	CBD_{CenA}-EGF	Epidermal Growth Factor	*E. coli*
	CBD_{CenA}-hCt	Human Calcitonin	*E. coli*
Cex	Abg-CBD_{Cex}	β-glucosidase	*E. coli*
	Cbg-CBD_{Cex}	Thermostable β-glucosidase	*E. coli*
	Glu oxidase-CBD_{Cex}	Glucose oxidase	chemically linked
	Prot A-CBD_{Cex}	Protein A	*E. coli*
	Avid-CBD_{Cex}	Streptavidin	*E. coli*
	Factor X-CBD_{Cex}	Factor X	*Pichia pastoris* Mammalian cells
	E. coli cells-CBD_{Cex}	Lpp-OmpA	*E. coli*
CenB	SF-CBD_{CenB}	Stem cell factor	*E. coli*
N1	CBD_{N1}-CD_{CenA}	CenA catalytic domain	*E. coli*
	CBD_{N1}-PhoA	Alkaline phosphatase	*E. coli*
CenD	SF-CBD_{CenD}	Stem cell factor	*Pichia pastoris*
Clos	CBD_{Clos}-Heparinase	Heparinase	*E. coli*
	CBD_{Clos}-GFP	Green fluorescent protein	Baculovirus

CBD_{CenA}, $_{CenB}$, $_{CenD}$, $_{N1}$, and $_{Cex}$ indicate CBDs from *Cellulomonas fimi* endoglucanase A, B, D, C and the exoglucanase/xylanase Cex, respectively. CBD_{Clos} refers to the CBD from the *Clostridium cellulovorans* cellulose-binding protein A (CbpA).

the affinity immobilization often enhances the thermal stability or recovery of the chimeric proteins. To add further to the already remarkable flexibility, circular permutations of CBDs; i.e. using a N-terminal CBD (e.g. CBD_{CenA} or CBD_{Clos}) as C-terminal or internal fusion partner, results in fully active hybrids containing CBDs with wild-type affinities [35].

Affinity tags can also modulate the solubility and control the cellular localization of the target molecules in the host organisms [1]. Expression of *C. fimi* CBDs or their hybrids, in *E. coli*, frequently results in the secretion of these proteins to the periplasm and consequently, a nonspecific leakage into the culture medium [32,36,37]. This process is not well understood but seems to be promoted, at least in part, by the CBDs. This property is useful for those proteins (fusion partners) which require the formation of disulfides for stability and thus benefit from being secreted. Furthermore, localization of the hybrids to the periplasm or culture media greatly simplifies the recovery and purification of the proteins. In some cases it is desirable to produce the fusion proteins intracellularly as insoluble aggregates or inclusion bodies. This strategy is useful for the production of small peptides or proteins which are prone to in vivo degradation when produced or secreted in soluble form. In *E. coli*, recombinant CBD_{Clos} from *Clostridium cellulovorans*, but not the homologous CBDs from *C. thermocellum* CipA (B) or *C. cellulolyticum* CipC, promotes the formation of these inclusion bodies [28,38–40] (Table 2).

Finally, the specificity of CBDs is not restricted to cellulosics, but extends to other glycans such as chitin (family I, II III and IV CBDs) and occasionally, xylans (family IIb) [28,30–32,36].

3.2. Interaction of CBDs with cellulose: binding and elution properties

From the existing 13 families, representatives of families I to V are characterized in most detail. In addition to a large body of biochemical and functional data, including binding specificities and affinities, a detailed three-dimensional structure has been solved for at least one member of each of these five families [41–45]. There are two classes of very different structural properties as exemplified by *C. fimi* CBD_{Cex} and CBD_{N1}:

(i) CBDs with high affinity for crystal cellulose (e.g. CBD_{Cex}) share a common feature; they are all β-proteins containing a ridge of linearly arranged, nearly regularly spaced, and solvent-exposed aromatics which are involved in the binding of the CBD to cellulose [46–48].

(ii) CBD_{N1} with affinity for soluble or amorphous cellulosics contains a small binding cleft rather than a linear array of binding residues. This carbohydrate-binding cleft is rich in polar residues and small hydrophobics but has relatively few aromatic residues [43]. Current results indicate that, with the exception of two tyrosines (Tyr19 and Tyr85), no other aromatic residues are involved in carbohydrate binding ([43]; J. Kormos and P. Tomme, unpubl.).

Although the structural features, as described above, will determine how the CBDs interact with cellulose, many of the details for interaction of CBDs with 'crystalline' cellulose remain poorly understood. Differences in the binding characteristics of (related) CBDs interacting with this insoluble matrix, are also hard to predict.

Table 4 summarizes the association constants K_a (M^{-1}) for binding of various CBDs on different cellulose allomorphs. CBDs with affinity for crystalline, bacterial cellulose (BMCC) bind equally well to Avicel (semicrystalline) and amorphous, phosphoric acid swollen cellulose (PASC). *T. reesei* CBD_{CbhI} has the lowest affinities for cellulose and CBD_{CipC} from *C. cellulolyticum* shows the highest affinities (Table 4). The binding interactions are fairly strong, with the average dissociation constants in the low micromolar range. In contrast, CBD_{N1} has markedly different specificities and binding affinities; it has a distinct preference for amorphous cellulose, with the affinity decreasing strongly with increasing crystallinity of the cellulose. No binding can be observed on BMCC (Table 4) [18,20]. Furthermore, CBD_{N1} binds small soluble cellooligosaccharides and larger soluble cellulosics with an affinity comparable to that for PASC [20,22].

Binding of CBD_{Cex} to BMCC is stable over a wide range of temperature and pH. As can be expected for the moderately exothermic binding (see thermodynamics), the binding affinity decreases slightly with increasing temperature [33] (Table 5).

Table 4
Equilibrium affinity constants K_a (M^{-1}) for binding of representatives from CBD families I to IV on different cellulosic matrices

Family/CBD	K_a (M^{-1})[a]		
	BMCC	Avicel	PASC
Family I			
T. reesei CBD_{CbhI} [b]	1.0×10^5	1.5×10^5	N.A
Family II			
C.fimi CBD_{Cex}	3.2×10^6	1.1×10^6	1.5×10^6
Family III			
C. cellulovorans CBD_{Clos} (CbpA)	1.7×10^6	1.6×10^6 [d]	N.A
C. thermocellum CBD_{CipC} [c]	2.9×10^7	7.7×10^6	6.7×10^7
Family IV			
C.fimi CBD_{N1}	No binding	2.3×10^3	4.1×10^5

[a] K_a values were obtained from depletion isotherms at 4°C (unless stated otherwise) after fitting of the raw data to a Langmuir-type adsorption model $[B]=[N_o]\,K_a[F]/1+K_a\,[F]$ where [B] is the concentration of bound CBD (mol per g cellulose), [F] is the concentration of free CBD (molar), $[N_o]$ is the total concentration of available binding sites on the cellulose surface (mol. g cellulose^{-1}) and K_a is the equilibrium association constant M^{-1}). [B] was calculated from the difference in total and free CBD concentration determined spectrophotometrically ($A_{280\ nm}$) in the supernatants before and after addition of cellulose, respectively (See Section 2 for details).
[b] Recalculated from the data in [31] using the Langmuir equation. Reported values are for binding at 22°C. Based on the reported temperature dependence [50] the K_a at 4°C is estimated to be $2.3\times10^5\ M^{-1}$.
[c] Affinities were determined for binding of miniCipC1 (CBD-HD_1-C_1). HD_1 and C_1 domains are believed not to interact with cellulose based on the similar K_a-values obtained for the isolated CBD and miniCipC1 on BMCC [54].
[d] The value was taken from [28]. The K_a-value was determined at 37°C. The published K_a for binding (20–22°C) of the homologous CBD_{CipB} from *C. thermocellum* to Avicel is $2.5\times10^6\ M^{-1}$ [40].
BMCC: Bacterial microcrystalline cellulose; PASC: phosphoric acid swollen cellulose; N.A: not assayed or not available.

However, the denaturation temperature of CBD_{Cex} increases 14°C from 64°C to about 78°C when bound to BMCC, indicating that the stability of the CBD is greatly enhanced upon immobilization (A.L. Creagh and C.A. Haynes, unpubl.). Although the magnitude of this stability increase is relatively large, it can be expected that the denaturation temperatures of most CBDs will be enhanced by adsorption to their cellulose ligands. Indeed, binding of CBD_{N1} to cellopentaose also moderately enhances its stability [21]. Binding of CBD_{CbhI} to cellulose exhibits a similar, albeit slightly more pronounced, dependency on pH and temperature as CBD_{Cex} [49,50].

Table 5
Effect of temperature and pH on the equilibrium constants for binding of CBD_{Cex} on BMCC

Temperature (°C)[a]	K_a (μM^{-1})	pH (22°C)	K_a (μM^{-1})
4°C	3.23	3.0	2.31
22°C	2.50	7.0	2.50
50°C	1.20	9.0	2.40

CBD_{Cex} (0.5–500 μM) was mixed with 1 mg BMCC in a final volume of 1 ml buffer. Buffers used were: 50 m*M* sodium citrate, pH 3.0; 50 m*M* sodium phosphate, pH 7.0; or 50 m*M* sodium carbonate, pH 9.0. Enzyme controls without BMCC were included. Samples were incubated (end-over-end) for 2 h at the given temperature. After centrifugation (2×5 min, 14 500 *g*) bound protein and K_a-values were determined as described under Section 2.
[a] Measured in 50 m*M* sodium phosphate buffer, pH 7.0.

Binding of CBD_{N1} to PASC or cellopentaose is somewhat more influenced by pH; the affinity changes little between pH 5.0 and 9.0 but drops 3–4 fold at higher pH. Binding cannot be measured below pH 5.0 because of the instability of the CBD [20]. The binding of CBD_{N1} to PASC and cellooligosaccharides decreases rapidly with increased temperature [J. Kormos and P. Tomme unpubl.] in accordance with the strong exothermic nature of these binding reactions [20].

Equally important as the binding affinity is the reversibility of the binding interaction because it determines the affinity application for that particular CBD. CBD_{Cex} (and CBD_{Clos}) shows irreversibility in binding to cellulose. A considerable debate as to the source and meaning of this 'irreversibility' is ongoing. Initially, the binding event seems to follow a

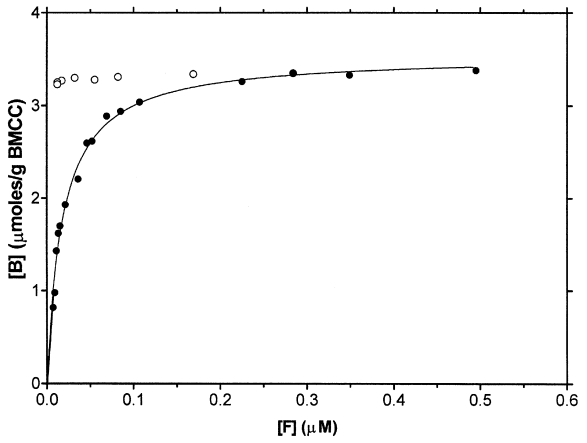

Fig. 2. Depletion isotherm for binding (4°C) of CBD_{Cex} to bacterial microcrystalline cellulose (BMCC) in 50 m*M* potassium phosphate, pH 7.0. Ascending (●) and descending (○) isotherms were obtained from the concentration of bound CBD [B] versus free, unbound CBD [F] measured as described under Section 2. The solid line represents the least-squares Langmuir model fit to the experimental data.

true equilibrium (reversibility) and indeed can be quantified assuming a Langmuir model (Fig. 2). However, despite the moderately high affinities ($K_a \sim 10^6$ to 10^7 *M*) (Table 4) no desorption can be observed when this equilibrium is disturbed, e.g. after dilution or by replacing the residual soluble CBD fraction for buffer and the descending and ascending binding isotherms do not superimpose (Fig. 2). In contrast, CBD_{CbhI} shows true reversibility in binding to BMCC [50,51]. Similarly, binding of CBD_{N1} to PASC or soluble cellulosics follows a true equilibrium and is completely reversible ([20]; J. Kormos and P. Tomme, unpubl.). It is not known if all members of the same family have a similar behaviour on cellulose or if these observations are restricted to the particular cases studied. Individual CBD_{Cex} (and to a certain degree CBD_{Clos}) molecules, have a tendency to interact with each other [42]. These protein–protein interactions are most obvious in the depletion isotherms which become biphasic at high protein concentrations and surface coverage. CBD_{CbhI} does not seem to show such a pronounced biphasic trend. It is thus tempting to speculate that the specific protein–protein recognition observed for CBD_{Cex} and CBD_{Clos} may, at least in part, be responsible for some of the apparent contradictory kinetics (reversibility) of binding. On the other hand, reversibility of binding may be directly related to the binding strength since CBD_{CbhI} (and CBD_{N1}) has a lower affinity for cellulose than CBD_{Cex} (or CBD_{Clos}) (Table 4). Nevertheless, caution should be exercised when interpreting these results obtained with the individual CBDs, since fusion partners can exert an effect on both the desorption characteristics and reversibility of the binding reaction [32,52].

Table 6 summarizes the most popular affinity tags and their characteristics. For bioseparations, an ideal affinity tag should be readily desorbed from the affinity matrix under mild conditions. However, the affinity interactions of many tags are so tight that extremes of pH or other protein denaturing conditions are needed to disrupt this interaction (Table 6). Conversely, these tight interactions are preferred for those applications which require a long-term immobilization of the fusions. For CBDs and CBD-hybrids, various elution conditions are required

Table 6
List of commonly used or commercially available affinity tags and their properties

Affinity tag	Size (kDa)	Matrix/ligand	Elution condition	Ref.
Cellulose-binding domain	4–20	cellulose	water or >4 *M* Guanidinium HCl	[18]
Maltose-binding domain	40	cross-linked amylose	10 m*M* maltose	[8]
Protein A	31	IgG	pH 2.8–3.0	[5,6]
Synthetic protein A (ZZ domain)	14	IgG	pH 2.8–3.0	[5,6]
Glutathion S-transferase	26	glutathione	5–10 m*M* reduced glutathione	[7]
Streptavidin	13	biotin	6 *M* urea pH 4.0	[13]
Biotin	0.24	monomeric avidin	20 m*M* biotin	[11,12]
Poly(His)	2–10 aa	IMAC	imidazole (<250 m*M*) or low pH	[14,15]
Flag peptide	8 aa	monoclonal antibody	pH 3.0 or 2–5 m*M* EDTA	[53]
S-peptide	15 aa	S-protein	2 *M* sodium thiocyanate	[9,10]

depending on the type of CBD and the matrix used. Some CBDs require denaturation to be eluted (Table 2) whereas others can be desorbed with water (Table 2) or with competitive ligands such as cellobiose [53]. Furthermore, binding and elution conditions can be modulated through mutagenesis of the binding residues.

3.3. Interaction of CBDs with cellulose: Thermodynamic analysis of binding

The thermodynamics for binding of CBD_{Cex} to insoluble, crystalline BMCC and of CBD_{N1} to a range of water-soluble cellulosics such as hydroxyethyl cellulose (HEC) were investigated by isothermal titration calorimetry (ITC). The results provide important insights into the driving forces for binding.

Analysis of the binding for CBD_{Cex} by ITC, against various putative binding models indicates the presence of two energetically different, independent classes of binding sites on the BMCC surface [19]. The high affinity site has an association constant K'_a of $6.3\times10^7\ M^{-1}$ and the low affinity site, a K'_a of $1.1\times10^6\ M^{-1}$.[2] The binding of CBD_{Cex} to either site is exothermic, but is driven by a large increase in (solvation) entropy ($T\ \Delta S^\circ = 30.6\ kJ\cdot mol^{-1}$). This, together with the large, negative differential heat capacity change ($\Delta Cp^\circ = -2.5$ to $-1.5\ kJ\cdot mol^{-1}\ K^{-1}$) associated with binding indicates that dehydration of both sorbent and protein make a significant contribution to the driving force for binding [19]. The small decrease in binding enthalpy ($\Delta H^\circ = -5.8\ kJ\cdot mol^{-1}$) suggests that some hydrogen bond formation between the protein residues and cellulose also takes place. Based on the similarity in structural features, CBDs that bind crystalline cellulose probably share many aspects of a common binding mechanism even if these CBDs belong to different families. If this holds true, then the driving forces for binding of these CBDs to the same substrate (i.e. BMCC) should be the same and the energetics of binding should follow a similar trend. Unfortunately, in the absence of a detailed thermodynamic binding analysis for these different CBDs, this cannot yet be proven with certainty. It is perhaps encouraging that the binding affinity of *T. reesi* cellobiohydrolase I (CBHI) to BMCC is enhanced by the presence of mono- and divalent cations, suggestive of a hydrophobic (dehydration) driving force [49].

The binding of CBD_{N1} to PASC, soluble cellooligosaccharides and hydroxy-ethyl cellulose is driven by a large exothermic standard enthalpy change (ΔH° ranges from $-60\ kJ\cdot mol^{-1}$ for glucan to $-32\ kJ\cdot mol^{-1}$ for PASC) with a slightly unfavourable (conformational) entropy contribution ($T\ \Delta S^\circ = -34$ to $-8.0\ kJ\cdot mol^{-1}$ for glucan and PASC, respectively). The small negative differential heat capacity change ($\Delta Cp^\circ = -430\ J\cdot mol^{-1}\ K^{-1}$ for HEC to $-210\ J\cdot mol^{-1}\ K^{-1}$ for cellooligosaccharides) and other thermodynamic data indicate that H-bonding and van der Waals interactions are the main driving forces for binding of the sugar to the polypeptide [20]. The K_a-values for interaction between CBD_{N1} and HEC or PASC are $2.2\times10^4\ M^{-1}$ and $1.5\times10^4\ M^{-1}$, respectively [20]. These values are somewhat lower than the one ($4.1\times10^5\ M^{-1}$) determined from the depletion isotherms (Table 4). CBD_{N1} binds small soluble oligosaccharides with 1:1 stoichiometry, whereas HEC and the larger cellulosics bind on average, 3 to 4 CBDs per polymer [20].

3.4. Affinity applications of CBD-hybrids

Two examples of applications in bioseparations, based on the different properties of CBDs, one using traditional column cellulose-affinity chromatography and the other a newly developed aqueous two-phase affinity partitioning system, are described below.

A hybrid between CBD_{CenA} and alkaline phosphatase (PhoA) was constructed with a small linker rich in proline and threonine (PT-linker) separating the two domains [34]. The CBD_{CenA}-PhoA hybrid was produced in *E. coli* and purified from the culture medium by affinity chromatography on cellulose CF1 (Sigma). After loading the column and washing

[2] These K'_a values (measured at 30°C) contain a term for the number of lattice units occupied by a single CBD molecule ($K'_a = a.K_a$). When this value ($a = 29$) is taken into consideration, the real K_a value for the high affinity site becomes $2.2\times10^6\ M^{-1}$ in good agreement with the values obtained (at 4°C) from other, independent methods (Table 4).

with high salt (1 *M* NaCl) to remove the contaminants, CBD_{CenA}-PhoA was recovered with an 84% yield by elution of guanidinium hydrochloride (Gnd·HCl). (Fig. 3) [34]. The hybrid was then refolded to its active dimers by dialysis against buffer. The hybrid could also be eluted with water but recovery yields (<60%) were somewhat lower than for elution with Gnd·HCl. Consequently, an additional washing step with Gnd·HCl increased recovery further by 29%. To demonstrate that the column could be reused, 10 consecutive purification cycles with water elution were run on the same column. From the second cycle onwards, consistent recovery yields of ~70% could be obtained throughout the complete recycling [34] (Fig. 4).

The selective binding of CBD_{N1} to a variety of water-soluble cellulosic polymers offers the possibility for use in a new, flexible aqueous two-phase affinity partitioning system.

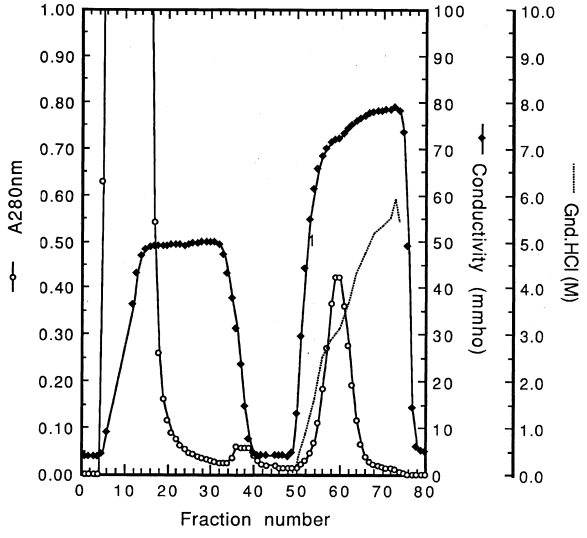

Fig. 3. Cellulose affinity purification of CBD_{CenA}-PhoA. A Pharmacia XK 50/30 column was packed with 20 g of cellulose (CF1, Sigma), previously washed with water to remove fines, giving a bed volume of approximately 100 ml. After equilibration (4°C) of the column with 50 m*M* Tris·HCl buffer, pH 7.5 (300 ml), *E. coli* culture supernatant (90 ml) containing ~20 mg of fusion protein was passed through the column at a flow-rate of 1 $ml \cdot min^{-1}$. The column was washed with 300 ml of 50 m*M* Tris·HCl buffer, pH 7.0 containing 1 *M* NaCl, 0.02% NaN_3 (high-salt buffer), followed by a linear gradient (100 ml) from high-salt buffer to 50 m*M* Tris·HCl buffer, pH 7.5, 0.02% NaN_3 (low-salt buffer), and finally 150 ml of low-salt buffer. Adsorbed protein was eluted with a 200-ml linear gradient from 0–6 *M* Gnd·HCl in 50 m*M* Tris·HCl buffer, pH 7.5. Fractions of 15 ml were collected throughout and monitored for $A_{280\ nm}$ and conductivity.

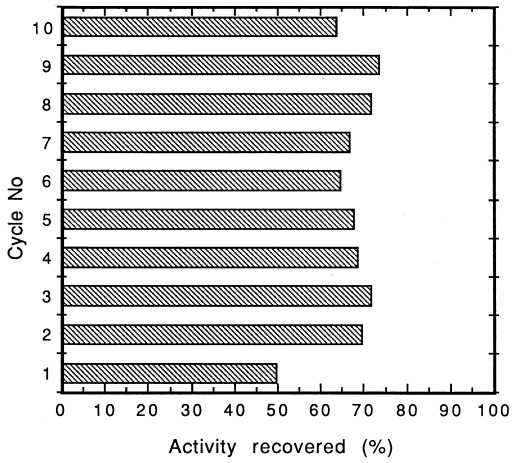

Fig. 4. Performance of the cellulose column during multiple rounds of CBD_{CenA}-PhoA purification. The experiment was done as described under Fig. 3 except that protein was eluted with a 200-ml linear gradient from low-salt buffer to distilled water, followed by 200 ml of distilled water. The column was then re-equilibrated with 50 m*M* Tris·HCl buffer, pH 7.5 (300 ml) and a new cycle was started.

Two-phase partitioning systems offer a number of unique advantages for large-scale purification of proteins and peptides, including high activity yields due to minimal inactivation of the proteins in the aqueous environment during purification, fast approach to equilibrium, easy and linear scale-up and continuous processing.

Two-phase systems are formed when two water soluble but incompatible polymers or a water soluble polymer and a strong electrolyte are mixed. Many two-phase systems contain polyethylene glycol (PEG) as one component and a glycan (e.g. dextran) as the second component. Systems containing two carbohydrates or polysaccharides are also frequently used. At equilibrium, each phase will be enriched in one separation-inducing component and will be nearly devoid of the other. When added to this two-phase system, proteins in a mixture (e.g. culture broth) will partition uniquely based on their relative affinities for the phase-forming components, as well as on size, surface chemistry and net charge.

Relatively low-partitioning coefficients and lack of sensitivity in the traditional partitioning systems, have motivated the development of 'affinity' parti-

tioning systems where an affinity ligand is covalently attached to one of the two polymers (usually PEG). However, these affinity partitioning systems are limited in their capacity and resolving power by low ligand densities which result from the presence of only one or two ligands per polymer chain. They are further limited by the expense of the chemistry needed to produce the polymer-ligand conjugates. The relatively high-binding affinity, combined with the potential for a single oligosaccharide chain to bind multiple CBD_{N1}-hybrids, suggests that both capacity and selectivity will be high in an affinity system based on water-soluble cellulosic polymers.

Several water-soluble polymers, including hydroxypropylmethyl cellulose (HPMC), methyl cellulose (MC), and hydroxyethylcellulose (HEC) were tested for their phase behaviour in different aqueous two phase systems. Fig. 5 shows the phase diagram obtained using the procedures of Haynes et al. [25] for mixtures of HPMC and Dextran T 500 in 50 m*M* phosphate buffer, pH 7.0 and 25°C. Phase separation occurs at low HPMC concentration and a stable partitioning system is formed at any total concentration above ca. 1% (w/w) HPMC and ca. 1.5% (w/w) Dextran T500, giving a large range of two-phase compositions and tie-line lengths useful for affinity partitioning. The favourable properties of a two-phase system formed with cellulosic polymers for which CBD_{N1} shows a high and specific affinity, suggests that affinity partitioning of CBD-hybrids is a feasible concept.

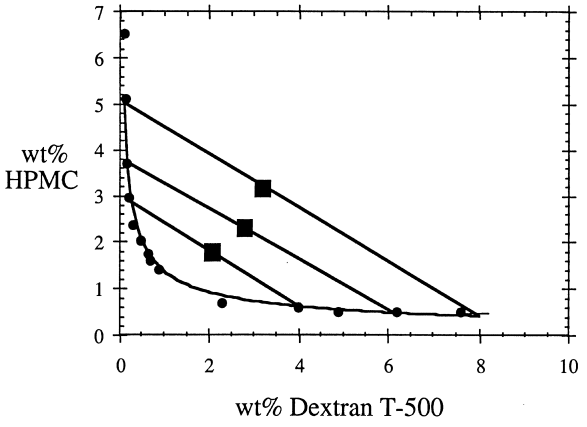

Fig. 5. Ternary phase diagram for mixtures of hydroxypropylmethyl cellulose and Dextran T500 in 50 m*M* phosphate buffer, pH 7.0 and 25°C. Phase diagrams were obtained according to the procedures described in [51].

4. Conclusions

Cellulose-binding domains have many desirable and flexible properties that make them ideal affinity tags: (1) they are relatively small and compact, yet come in a wide range of sizes, (2) they can be attached to the target protein without interference with the biological activity of the target molecule, (3) they can be attached to the target: at the N- or C-terminal or internal for multifunctional chimerics, (4) CBDs are stable and resistant to denaturation and proteolytic degradation in vivo and during the specific proteolysis step used to remove the tag from the target molecule, (5) in *E. coli*, some CBDs promote the export and leakage to the culture medium of recombinant hybrids whereas others induce the formation of insoluble aggregates, (6) they show a high, specific interaction with numerous cellulosic matrices in a background of low, nonspecific protein adsorption, (7) this affinity interaction is strong and stable under a variety of conditions which allow for an efficient and complete removal of contaminants or for stable, long term immobilization of the hybrids, and (8) CBDs can be selected for removal, under mild conditions, of hybrids from the affinity matrix.

The wide diversity in selectivities and affinities for different allomorphs or forms of cellulose and other polysaccharides (e.g. chitin) creates a large scope of applications for bioseparations and bioprocessing.

Finally, the large body of biochemical, functional and structural data available on selected CBDs and the vast number of as yet unexplored and uncharacterized families and CBDs suggests that CBDs can be manipulated or identified for more desirable or different properties.

Acknowledgements

This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada and the Network of Centres of Excellence Program and by a grant from CBD Technologies. We thank Oded Shoseyov, Philip Johnson and Lawrence Mclntosh for helpful discussions and insights. Special thanks to Stanley Hirsch, Tom Puskar and George Aaron for their support and for relentlessly promoting the technology.

References

[1] M. Uhlén, T. Moks, Methods Enzymol. 185 (1990) 129.
[2] E.R. La Vallie, J.M. McCoy, Curr. Opin. Biotechnol. 6 (1995) 501.
[3] C.F. Ford, I. Souminen, C.E. Glatz, Prot. Exp. Pur. 2 (1991) 95.
[4] B. Nilsson, G. Forsberg, M. Hartmanis, Curr. Opin. Struct. Biol. 2 (1992) 569.
[5] T. Moks, L. Abrahmsén, B. Nilsson, U. Hellman, J. Sjöquist, M. Uhlén, Eur. J. Biochem. 156 (1986) 637.
[6] B. Nillson, T. Moks, B. Jansson, L. Abrahmsén, A. Elmblad, E. Holmgren, C. Henrichson, T.A. Jones, M. Uhlén, Protein Eng. 1 (1987) 107.
[7] D.B. Smith, K.S. Johnson, Gene 67 (1988) 31.
[8] H. Bedouelle, P. Duplay, Eur. J. Biochem. 171 (1988) 541.
[9] F.M. Richards, P.J. Vithayathil, J. Biol. Chem. 234 (1959) 1459.
[10] J.-S. Kim, R.T. Raines, Protein Sci. 2 (1993) 348.
[11] J.E. Cronan Jr., J. Biol. Chem. 265 (1990) 10327.
[12] M. Wilchek, E.A. Bayer, Anal. Biochem. 71 (1983) 1.
[13] T. Sano, C.R. Cantor, Bio/Technology 9 (1991) 1378.
[14] J. Porath, J. Carlsson, I. Olsson, G. Belfrage, Nature 258 (1975) 598.
[15] E. Hochuli, W. Bannwarth, H. Döbeli, R. Gentz, D. Stüber, Bio/Technology 6 (1988) 1321.
[16] K. Nord, E. Gunneriusson, J. Ringdahl, S. Ståhl, M. Uhlén, P.-A. Nygren, Nature Biotech. 15 (1997) 772.
[17] N.R. Gilkes, M.L. Langsford, D.G. Kilburn, R.C. Miller, R.A.J. Warren, J. Biol. Chem. 259 (1984) 10455.
[18] J. Coutinho, N.R. Gilkes, R.A.J. Warren, D.G. Kilburn, R.C. Miller Jr., Mol. Microbiol. 6 (1992) 1243.
[19] A.L. Creagh, E. Ong, E. Jervis, D.G. Kilburn, C.A. Haynes, Proc. Natl. Acad. Sci. USA 93 (1996) 12229.
[20] P. Tomme, A.L. Creagh, D.G. Kilburn, C.A. Haynes, Biochem. 35 (1996) 13885.
[21] A.L. Creagh, J. Koska, P.E. Johnson, P. Tomme, M.D. Joshi, L.P. Mclntosh, D.G. Kilburn, C.A. Haynes, Biochem. (1997) in press.
[22] P.E. Johnson, P. Tomme, M.D. Joshi, L.P. Mclntosh, Biochem. 35 (1996) 13895.
[23] J. Sambrook, E.F. Fritsch, T. Maniatis, Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press, New York, 1989.
[24] O. Shoseyov, R.H. Doi, Proc. Natl. Acad. Sci. USA 89 (1990) 2192.
[25] C.A. Haynes, F.J. Benitez, H.W. Blanch, J.M. Pransnitz, AIChE J. 39 (1993) 1539.
[26] N.R. Gilkes, B. Henrissat, D.G. Kilburn, R.C. Miller, R.A.J. Warren, Microbiol. Rev. 55 (1991) 303.
[27] P. Tomme, R.A.J. Warren, N.R. Gilkes, Adv. Microb. Physiol. 37 (1995) 1.
[28] M.A. Goldstein, M. Takagi, S. Hashida, O. Shoseyov, R.H. Doi, I.H. Segel, J. Bacteriol. 175 (1993) 5762.
[29] P. Béguin, J.-P. Aubert, FEMS Microbiol. Rev. 13 (1994) 25.
[30] P. Tomme, R.A.J. Warren, R.C. Miller Jr., D.G. Kilburn, N.R. Gilkes, in: J.N. Saddler, M.H. Penner (Eds.), Enzymatic Degradation of Insoluble Carbohydrates, American Chemical Society Symposium Series, 1995, p. 142.
[31] M. Linder, I. Salovuori, L. Ruohonen and, T.T. Teerei, J. Biol. Chem. 271 (1996) 21272.
[32] P. Tomme, D.P. Driver, E.A. Amandoron, R.C. Miller Jr., R.A.J. Warren, D.G. Kilburn, J. Bacteriol. 177 (1995) 4356.
[33] E. Ong, N.R. Gilkes, R.A.J. Warren, R.C. Miller Jr., D.G. Kilburn, Bio/Technology 7 (1989) 604.
[34] J.M. Greenwood, N.R. Gilkes, R.C. Miller Jr., D.G. Kilburn, R.A.J. Warren, Biotech. Bioeng. 44 (1994) 1295.
[35] P. Tomme, N.R. Gilkes, R.C. Miller Jr., R.A.J. Warren, D.G. Kilburn, Protein Eng. 7 (1994) 117.
[36] E. Ong, N.R. Gilkes, R.A.J. Warren, R.C. Miller Jr., D.G. Kilburn, Biotech. Bioeng. 42 (1993) 401.
[37] D. Hazenwinkle, E. Jervis, O. Kops, C. Liu, C.A. Haynes, D.G. Kilburn, Biotech. Bioeng. 55 (1997) 854.
[38] S. Pagès, A. Bélaïch, C. Tardif, C. Reverbel-Leroy, C. Gaudin, J.-P. Bélaïch, J. Bacteriol. 178 (1996) 2279.
[39] K. Kruus, A.C. Lua, A.L. Demain, S.H.D. Wu, Proc. Natl. Acad. Sci. USA 92 (1995) 9254.
[40] E. Morag, A. Lapidot, D. Govorko, R. Lamed, M. Wilchek, E.A. Bayer, Y. Shoham, Appl. Environ. Microbiol. 61 (1995) 1980.
[41] P.J. Kraulis, G.M. Clore, M. Nilges, T.A. Jones, G. Petersson, J. Knowles, A.M. Gronenborn, Biochem. 28 (1989) 7241.
[42] G.-Y. Xu, E. Ong, N.R. Gilkes, D.G. Kilburn, D.R. Muhandiram, M. Harns-Brandts, J.P. Carver, L.E. Kay, T.S. Harvey, Biochemistry 34 (1995) 6993.
[43] P.E. Johnson, M.D. Joshi, P. Tomme, D.G. Kilburn, L.P. Mclntosh, Biochem. 35 (1996) 14381.
[44] J. Tormos, R. Lamed, A.J. Chirino, E. Morag, E.A. Bayer, Y. Shoham, T.A. Steitz, EMBO J. 15 (1996) 5739.
[45] E. Brun, F. Moriaud, P. Gans, M.J. Blackledge, F. Barras, D. Marion, Biochem. 36 (1997) 16074.
[46] T. Reinikainen, L. Ruohonen, T. Nevanen, L. Laaksonen, P. Kraulis, T.A. Jones, J.K.C. Knowles, T.T. Teeri, Proteins Struct. Funct. Genet. 14 (1992) 475.
[47] D.B. Poole, G.P. Hazlewood, N.S. Huskisson, R. Virden, H.J. Gilbert, FEMS Microbiol. Lett. 106 (1993) 77.
[48] N. Din, I.J. Forsythe, L.D. Burtnick, N.R. Gilkes, R.C. Miller Jr., R.A.J. Warren, D.G. Kilburn, Mol. Microbiol. 11 (1994) 747.
[49] T. Reinikainen, O. Teleman, T.T. Teeri, Proteins Struct. Funct. Genet. 22 (1995) 392.
[50] M. Linder, T.T. Teeri, Proc. Natl. Acad. Sci. USA 93 (1996) 12253.
[51] M.K. Bothwell, D.B. Wilson, D.C. Irwin, L.P. Walker, Enzyme Microb. Technol. 20 (1997) 411.
[52] T. Reinikainen, K. Takkinen, T.T. Teeri, Enzyme Microb. Technol. 20 (1997) 143.
[53] D. Wassenberg, H. Schurig, W. Liebl, R. Jaenicke, Protein Sci. 6 (1997) 1718.
[54] S. Pagès, L. Gal, A. Bélaïch, C. Gaudin, C. Tardif, J.-P. Bélaïch, J. Bacteriol. 179 (1997) 2810.
[55] T.P. Hopp, K.S. Frickett, V.L. Price, R.T. Libby, C.J. March, D.P. Cerretti, D.L. Urdal, P.J. Conlon, Bio/Technology 6 (1988) 1204.