Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Chromatography B, 807 (2004) 87–94
JOURNAL OF CHROMATOGRAPHY B
www.elsevier.com/locate/chromb

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*

Mojgan Kavoosi [a], Julia Meijer [a], Emily Kwan [b], A. Louise Creagh [a], Douglas G. Kilburn [b], Charles A. Haynes [a,*]

[a] *The Department of Chemical and Biological Engineering, The University of British Columbia, Vancouver, BC, Canada V6T 1Z3*
[b] *The Department of Microbiology and Immunology, The University of British Columbia, Vancouver, BC, Canada V6T 1Z3*

Available online 24 April 2004

Abstract

A novel inexpensive affinity purification technology is described based on recombinant expression in *Escherichia coli* of the polypeptide or protein target fused through its N-terminus to TmXyn10ACBM9-2 (CBM9), the C-terminal family 9 carbohydrate-binding module of xylanase 10A from *Thermotoga maritima*. Measured association constants (K_a) for adsorption of CBM9 to insoluble allomorphs of cellulose are between 2×10^5 and $8 \times 10^6\ M^{-1}$. CBM9 also binds a range of soluble sugars, including glucose. As a result, a 1 M glucose solution is effective in eluting CBM9 and CBM9-tagged fusion proteins from a very inexpensive commercially-available cellulose-based capture column. A processing site is encoded at the C-terminus of the tag to facilitate its rapid and quantitative removal by Factor X_a to recover the desired target protein sequence following affinity purification. Fusion of the CBM9 affinity tag to the N-terminus of green fluorescent protein (GFP) from the jellyfish, *Aquorin victoria*, is shown to yield >200 mg l^{-1} of expressed soluble fusion protein that can be affinity separated from clarified cell lysate to a purity of >95% at a yield of 86%.

Keywords: Purification; *Escherichia coli*; Polypeptides; Xylanase 10A; Carbohydrates

1. Introduction

The continued maturation of the pharmaceutical and biotechnology industries has created an increasing need for practical and economical large-scale processing techniques. Production methods such as fed-batch fermentations of recombinant microbes have advanced to a level where target biomolecules can be produced in g l^{-1} concentrations at relatively modest cost. As a result, downstream processing often accounts for more than 60% of the total operating cost, and as much as 70% of the capital cost of current biochemical production processes [1]. Purification of a target protein during manufacturing usually requires several chromatographic steps in series due to the relatively non-specific physico-chemical interactions that drive separations in these columns. Although product purities are often quite high, overall yields from multi-step chromatographic procedures are generally low due to the accumulated loss of product [2,3]. The challenge, therefore, is to reduce costs and increase overall yields by process simplification through elimination or combination of purification steps.

Toward this goal, a number of affinity separation systems have been developed in the past two decades to replace difficult multi-step chromatographic procedures with a highly selective binding step that serves to both purify and concentrate the product [4,5]. Polypeptide fusion tags that selectively bind a complementary ligand immobilized onto a suitable chromatographic matrix are now widely used at the laboratory scale to facilitate recombinant expression and purification of target proteins [6,7]. In addition to allowing rapid purification, affinity fusion tags have been shown in certain cases to increase in vivo proteolytic stability of the target protein, improve product solubility, and control product localization in or secretion from the expression host [8]. Skillful engineering of the fusion-tag/immobilized-ligand pair can, therefore, provide a robust and generic method for

* Corresponding author. Fax: +1-604-822-2114.
E-mail address: israels@chml.ubc.ca (C.A. Haynes).

1570-0232/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.jchromb.2004.03.031

efficient production and high-resolution affinity purification of recombinant protein targets.

Commercially available affinity tag systems include the calmodulin binding peptide [9,10], the glutathione *S*-transferase (GST) tag from *Schistosoma japonicum* [11,12], and various polyamino-acid affinity tags such as the polyhistidine tag [13–15] and its associated immobilized metal-ion affinity chromatography (IMAC) capture column. Each of these affinity tag systems has been used extensively at the laboratory scale, but only the GST and polyhistidine tag systems have found any appreciable use in manufacturing. More extensive use of affinity tags systems by industry has been thwarted in part by the complex chemical modifications required to cross-link the solid support or to graft the affinity-tag receptor to the resin surface, and by the relatively low tolerance of many of these affinity resins to repeated processing and sanitization (SIP) cycles. However, the dominant impediment to the use of current commercially available generic affinity tags in large-scale bioprocessing is cost. A new affinity tag that binds to an inexpensive, chemically and hydrodynamically robust chromatographic resin is therefore highly desirable.

Here, we present a generic and inexpensive affinity purification technology based on high-titer recombinant expression in *Escherichia coli* of fusion proteins containing the carbohydrate-binding module CBM9 attached to the N-terminus of the target protein or polypeptide. TmXyn10ACBM9-2 (henceforth referred to as CBM9), the C-terminal family 9 carbohydrate-binding module of xylanase 10A from *Thermotoga maritima* [16], binds specifically to the reducing ends of cellulose and soluble polysaccharides, a property that is currently unique to this CBM. Measured association constants (K_a) for adsorption of CBM9 to insoluble allomorphs of cellulose are between 2×10^5 and $8 \times 10^6\ M^{-1}$. CBM9 also binds a range of soluble sugars [17], including glucose, such that a 1-M glucose solution is effective in quantitatively eluting CBM9 and CBM9-tagged fusion proteins from a cellulose-based capture column. The presence of the CBM9 tag, therefore, allows for affinity capture and purification of a fusion protein on an inexpensive cellulose-based chromatography resin.

A unique processing site is encoded at the C-terminus of the tag to facilitate rapid and quantitative removal of the tag by Factor X_a to recover the desired target protein sequence following affinity purification [18]. Validation of the technology is provided by fusing the CBM9 affinity tag to the N-terminus of green fluorescent protein (GFP) from the jellyfish, *Aquorin victoria* [19,20]. The use of GFP as the target protein has the advantage that the natural fluorescence of GFP measured at 509 nm (excitation at 395 nm) offers a direct and convenient means of tracking the target fusion protein throughout the production and affinity purification process.

The generic CBM9 affinity-tag technology proposed here involves five distinct processing steps: (i) recombinant production (cytoplasmic) of the properly folded fusion protein in recombinant *E. coli* BL21 (DE3) cells; (ii) cell lysis and lysate resuspension; (iii) affinity purification (including elution) of the CBM9-tagged fusion protein on a suitable commercial cellulose-based chromatography resin; (iv) cleavage of the CBM9-linker-IEGR affinity tag sequence using immobilized recombinant Factor X_a; and finally (v) removal of CBM9 to obtain the purified target (GFP). Each of these essential processing steps is evaluated in terms of product yield, purity and concentration factor to provide a measure of the overall performance of the technology.

2. Materials and methods

2.1. Reagents

Kanamycin, glucose, and all other chemicals were purchased from Sigma Chemicals (St. Louis, MO, USA). All reagents were analytical grade unless stated otherwise. Restriction enzymes were purchased from New England Biolabs (Beverly, MA). T4-DNA ligase was obtained from Roche Molecular Biochemicals (Laval, Que.). Perloza MT100 chromatography resin having a nominal particle diameter distribution of 50–80 μm was purchased from Iontosorb Inc. (Czech Republic). *E. coli* BL21 (DE3) and Ni^{+2}-Sepharose resin were obtained from Novagen (Milwaukee, MI).

2.2. Cloning of CBM9–GFP fusion protein

All cloning procedures were performed using standard molecular biology techniques [21]. The *GFP* and *CBM9* coding regions were amplified from the vectors pGFPuv (Clontech, Palo Alto, CA) and pETCBM9, respectively. A *Bsp*HI restriction endonuclease site (underlined) was introduced at the 5′ end of *CBM9* gene fragment, using the oligonucleotide 5′-TTGCTAGCTTCATGACTAGCGGAATAATGGTAGC-3′ as primer. The sequence encoding for the S_3N_{10} linker (italic) and a Pvu I site (underlined) were introduced at the 3′ end of the CBM9 coding region using the oligonucleotide 5′-TCCCTCGAT*CGCGAGGTTGTTGTTATTGTTATTGTTGTTGTTGTTCGAGCTCGA*AAGCTTGATGAGCCTGAGGTTACC-3′ as primer. For the *GFP* gene fragment, the sequence encoding for the Factor X_a recognition site (IEGR) (italic) and a Pvu I restriction endonuclease site (underlined) were placed at the 5′ end, using the oligonucleotide 5′-CCG*ATCGAGGGTCGT*ATCATGAGTAAAGGAGA-3′ as primer. For the 3′ end, a Not I site (underlined) was introduced using the oligonucleotide 5′-TGCGGCCGCTTTGTAGAGCTCATCCATGCCATGTGTAATCCC-3′ as primer. Each PCR mixture (50 μl total volume) contained 50 ng of template, 30 pmol of each primer, 5% DMSO, 0.4 mM 2′-deoxy-nucleoside 5′-triphosphates, and 1 U of *PWO* DNA polymerase in buffer (Roche Molecular Biochemicals, Laval, Quebec). The following protocol for 25 successive PCR cycles was followed: denaturation at

94 °C for 30 s, annealing for 2 min by linearly increasing the temperature from 55 to 72 °C, and primer extension at 72 °C for 45 s. The resulting CBM9-S_3N_{10} and FX_a-GFP coding regions were digested with *Bsp*HI/*Pvu*I and *Pvu*I/*Not*I, respectively, and ligated (16 °C, 16 h) into the pET28b vector (Novagen) previously digested with *Nco*I and *Not*I to give the appropriate pET28-CBM9-S_3N_{10}-IEGR-GFP construct (hereafter referred to as pET28-CBM9-GFP). DNA sequencing was then completed to verify the construct (NAPS Unit, Biotechnology Laboratory, University of British Columbia).

2.3. Protein production

Overnight cultures of *E. coli* strain BL21/pET28-CBM9-GFP were diluted 100-fold in tryptone–yeast extract–phosphate medium (TYP) supplemented with 50 μg ml^{-1} of kanamycin and grown at 37 °C to a cell density ($OD_{600\,nm}$) of ∼1.0. Isopropyl-1-thio-β-D-galactoside (IPTG) was added to a final concentration of 0.3 mM. Incubation was then continued at 30 °C for a further 10–12 h. The cells were harvested by centrifugation (8,500 × *g*) at 4 °C for 20 min and then resuspended in high salt buffer (1 M NaCl, 50 mM potassium phosphate, pH 7.0) by gentle mixing. Cells were ruptured by two passages through a French pressure cell (21000 lb in^{-2}) and the cell debris removed by centrifugation for 30 min at 27,000 × *g* and 4 °C. CBM9–GFP fusion protein was purified by affinity chromatography as described below.

The stability of the fusion protein against proteolysis was assayed as follows. At 18 h post-induction, the culture was divided into two equal volumes and cells were harvested as described above. The cells in one container were resuspended in high salt buffer while the cells in the second container were resuspended in high salt buffer containing 1 mM phenylmethylsulfonyl fluoride (PMSF) (Sigma). Cells were disrupted (in the presence of fresh 1 mM PMSF or buffer) and the clarified cell extract (1.5 ml) incubated with 200 μl of Perloza MT100 (94.5 mg dry weight ml^{-1}). The mixture was mixed at room temperature for 3.5 h by rotating end-over-end. The resin was collected by centrifugation at 8,500 × *g* for 8 min, washed three times with 1 ml high salt buffer, 2× with low salt buffer (50 mM potassium phosphate pH 7.0) and 1× with TBS8 (15 mM NaCl, 10 mM Tris–HCl, pH 8.0). The protein was desorbed with 400 μl of 1 M glucose in TBS8 and analyzed by SDS–PAGE.

2.4. Affinity chromatography

A Pharmacia XK-16 column (8.5 cm × 1.6 cm i.d.) was packed by standard inclined pouring with Perloza MT100 resin to give a final bed volume of ca. 17 ml. All purification chromatograms were completed on a Pharmacia P-500 FPLC system (Amersham Biosciences) at 4 °C and with a flow rate of 0.5 ml min^{-1}. The Perloza MT100 affinity column was equilibrated with ∼10 column volumes (CV) of high salt buffer. Clarified cell extract (50 ml) was loaded onto the column and unbound protein was removed by washing the column with 10 CV of high salt buffer, 5 CV of low salt buffer and 4 CV of TBS8 buffer. CBM9–GFP was then eluted from the column with 5 CV of 1 M glucose in TBS8. Eluted protein fractions were analyzed for purity by 12% SDS–PAGE using 20% SDS sample buffer. Column regeneration was completed using 10 CV water followed by 10 CV of high salt buffer for equilibration.

2.5. Tag Cleavage by Factor X_a

In this experiment carried out at 21 °C and pH 7, a purified chimeric protein comprised of Factor X_a fused to CBM2a, the family 2a carbohydrate binding module (CBM2a) of xylanase 10A of the soil bacterium *Cellulomonas fimi*, was immobilized onto a Perloza MT500 column (henceforth called CBM2a-FXa_{im}) [22] and used to enzymatically remove the CBM9 affinity tag. Pure CBM9–GFP fractions were pooled and incubated with CBM2a-FXa_{im} at 21 °C, rotating end-over-end. After 108 h, CBM2a-FXa_{im} was removed by centrifugation (8,500 × *g*, 15 min) and washed extensively to collect all cleaved product. The cleaved products were buffer exchanged into low salt buffer and concentrated in a stirred ultrafiltration (UF) unit (Amicon, Beverly, MA) on a 1 K cutoff filter (Filtron, Northborough, MA). The concentrated protein solution was applied to a column (24 cm × 0.9 cm i.d.) packed with Perloza MT100 and washed with 15 CV of low salt buffer. Free GFP was collected in the flow through and concentrated by UF.

The processing time required for Factor X_a cleavage of the IEGR-terminal affinity tag was also assayed. Three milligrams of purified CBM9–GFP fusion protein was incubated with 3 μl CBM2a-FXa_{im} [22] at 21 °C (final [Factor X_a] to [fusion protein] ratio of 1:1000). The control experiments contained buffer in place of CBM2a-FXa_{im}. Samples were taken at the following time points (0, 0.5, 1, 2, 4, 6, 8, 10, 12, 14, 25.5, and 35 h post-incubation) and analyzed by SDS–PAGE.

2.6. Fluorescence calibration curves

Highly pure CBM9–GFP was obtained by sequentially purifying CBM9–GFP on the Perloza MT100 affinity column followed by immobilized-metal affinity chromatography (IMAC) on a Ni^{+2}-Sepharose IMAC resin (according to manufacturer's instructions). This highly pure protein was buffer exchanged into low salt buffer and concentrated as described above. Concentrations of the purified protein were determined by UV absorbance (280 nm) using a calculated molar extinction coefficient [23].

Varying concentrations of highly pure protein were mixed with either loading buffer (TBS8), elution buffer (1 M glucose in TBS8) or BL21 cell extract (A_{280} = 6 or 0.6) and the fluorescence measured (395 nm for excitation; 509 for emission) using a Cary Eclipse fluorescence

spectrophotometer (Varian, Palo Alto, CA). Linear calibration curves (0–0.35 μM CBM9–GFP) for measured fluorescence as a function of CBM9–GFP concentration in buffer and in BL21 cell extract were constructed from each data set.

2.7. Measurement of binding isotherms

Samples containing purified fusion protein at concentrations ranging from 1 to 30 μM were incubated with resin (1 mg (dry weight) of Perloza MT100, 5 mg CF31, 5 mg CF1 and 5 mg Avicel) in high salt buffer to a final volume of 1 ml. The samples were then incubated for 30 h at 4 °C (25 °C for Avicel samples) while mixing end-over-end. The cellulose was removed by centrifugation at 27,000 × *g* for 16 min at 4 °C. The supernatant was collected and the concentration of unbound protein was determined by UV absorbance (280 nm) using a Cary 100 Spectrophotometer (Varian).

An isotherm was generated by plotting the concentration of bound protein (μmol g^{-1} of resin) against the concentration of unbound protein (μM). The binding parameters were determined by fitting a non-linear regression of the Langmuir-type adsorption isotherm equation to the experimental data using GraphPad Prism 3.0 software.

Binding isotherms were also measured for CBM9–GFP in the presence of bacterial cell extract (A_{280} of 5.3). Samples were incubated for 16 h at 4 °C while mixing end-over-end. The cellulose was removed by centrifugation as described above and the fluorescence of the supernatant was measured. The concentration of unbound protein in the supernatant was determined from the calibration curve.

3. Results and discussion

Fig. 1 shows a block diagram for the CBM9–GFP fusion protein construct used in this work to validate the utility and performance of our CBM9 tag technology for inexpensive affinity purification of recombinant proteins and peptides in *E. coli*. In the pET28-CBM9-GFP vector, the coding sequence for the N-terminal CBM9 is followed by the gene fragment encoding an S_3N_{10} linker that serves to separate the CBM9 fusion tag from the target protein. The synthetic S_3N_{10} linker was used in this study because it has proven useful in our laboratory in the stable expression of a range of fusion proteins. The combined CBM9-S_3N_{10} fusion tag is separated from the N-terminal amino acid of GFP by the four amino acid IEGR processing site for the endoprotease recombinant human Factor X_a. The presence of the IEGR processing site allows Factor X_a catalyzed removal of the affinity tag following fusion protein purification to recover the pure target protein with its natural N-terminus.

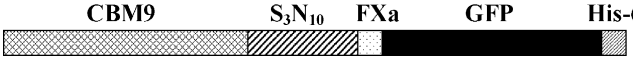

Fig. 1. Schematic representation of gene fragment coding for the CBM9-S_3N_{10}-IEGR-GFP fusion protein.

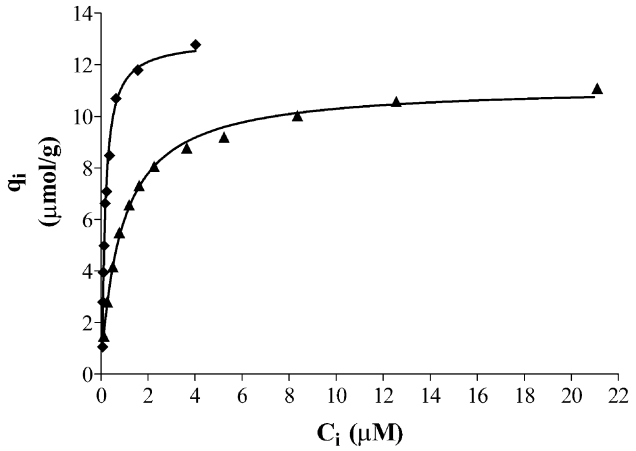

Fig. 2. Equilibrium adsorption isotherms for binding of CBM9 and CBM9–GFP to Perloza MT100 at 4 °C. CBM9–GFP binding at 4 °C to: (◆) Perloza MT100 in high salt buffer; and (▲) in high-salt buffer, where q_i is the bound protein concentration and C_i the equilibrium concentration of protein free in solution.

Table 1
Langmuir adsorption parameters (equilibrium association constant K_a and binding capacity N_o) for binding of CBM9 and CBM9–GFP to Perloza MT100 at 4 °C. Solvent contains pure protein in high-salt buffer

Protein	K_a (M^{-1})	N_o (μmol protein g^{-1} resin)
CBM9–GFP	7.3 (±0.92) × 10^6	13.0 (±0.35)
CBM9	1.1 (±0.08) × 10^6	11.2 (±0.14)

3.1. Binding isotherms and thermodynamics

Equilibrium adsorption isotherms at 4 °C for CBM9–GFP binding to the porous cellulose-based chromatography resin Perloza MT100 are shown in Fig. 2. As shown in Table 1, the Perloza MT100 stationary phase resin binds pure CBM9–GFP with a capacity of 13.0 (±0.35) μmol g^{-1} MT100. For the 53 kg mol^{-1} CBM9–GFP fusion protein, this equates to a saturation loading capacity of 690 mg protein bound per gram of resin, or ca. 150 mg ml^{-1} of column. In pure buffer at 4°C, CBM9–GFP binds Perloza MT100 with an affinity of 7.3 (±0.92) × 10^6 M^{-1}.

As shown in Table 2, CBM9–GFP also binds to a number of other commercially available cellulose-based resins.

Table 2
Binding affinity and capacity of CBM9–GFP on various cellulosic resins

Resin	K_a (M^{-1})	N_o (μmol protein g^{-1} resin)
Perloza MT100[a]	7.3 (±0.92) × 10^6	13.0 (±0.35)
CF1[a]	2.5 (±0.11) × 10^5	0.12 (±0.01)
CF31[a]	4.5 (±0.19) × 10^5	0.30 (±0.03)
Avicel[b]	4.1 (±0.68) × 10^5	0.52 (±0.02)

[a] Binding performed at 4 °C in high-salt buffer.
[b] Binding performed at 25 °C in high-salt buffer.

However, in each case, the resin capacity (and to a lesser extent the binding affinity) is significantly lower than observed for binding to Perloza MT100, indicating that the Perloza resin offers a relatively high concentration of entropically unhindered reducing ends for CBM9 binding.

The equilibrium adsorption isotherm at 4 °C for binding of pure CBM9 to Perloza MT100 is also shown (Fig. 2, Table 1). The binding properties of the isolated fusion tag (CBM9) are similar to those of the fusion protein (CBM9–GFP), indicating that the presence of the target protein does not significantly affect the performance of the CBM9 affinity tag.

The intrinsic fluorescence of GFP allowed us to also measure the binding isotherm for the CBM9–GFP fusion protein in the presence of the *E. coli* cell lysate from which it is purified. Although errors in CBM9–GFP fluorescence measurements are large when cell lysate components are present in the solution phase, the data suggest that neither the binding capacity of the resin or the affinity of the CBM9–GFP fusion protein for the resin is significantly altered by the presence of a large concentration of contaminant proteins (data not shown), indicating the specificity of the Perloza MT100 cellulose-based resin for CBM-tagged proteins.

3.2. Fusion protein expression and stability

Unoptimized batch fermentation yields of soluble CBM9–GFP in recombinant *E. coli* BL21 cultures were typically around 210 mg l^{-1} of culture, which represents a 40% increase in GFP yield over more standard expression systems [24,25]. The tendency for CBM fusion tags, including the more commonly used maltose binding protein, to increase soluble expression of otherwise low expressing proteins is well documented [26,27]. This ability to enhance titers of soluble protein is likely due, at least in part, to the relatively high solubility of CBMs, which allows them to serve as effective solubilizing agents for aggregation-prone polypeptides. In certain cases, fusion to a CBM can also promote the proper folding of the attached protein into its biologically active conformation. This chaperone-like quality distinguishes CBMs such as CBM9 and MBP from other affinity tags and greatly enhances their value as a fusion partner.

The performance of a fusion tag technology depends not only on the properties of the tag, but also on the stability of the amino-acid sequence that links the tag to the target protein. Spiking and incubation of purified CBM9 in an *E. coli* BL21 culture lysate resulted in no detectable degradation of the CBM as measured by SDS–PAGE gels. The stability of the S_3N_{10}IEGR linker against degradation by endogenous *E. coli* proteases present in the cytoplasm and cell lysate was, therefore, analyzed by SDS–PAGE following cell disruption and lysate clarification, either in the presence or absence of the protease inhibitor PMSF. As the CBM is not degraded significantly by endogenous proteases, proteolytic degradation within the linker will result in the appearance of a band

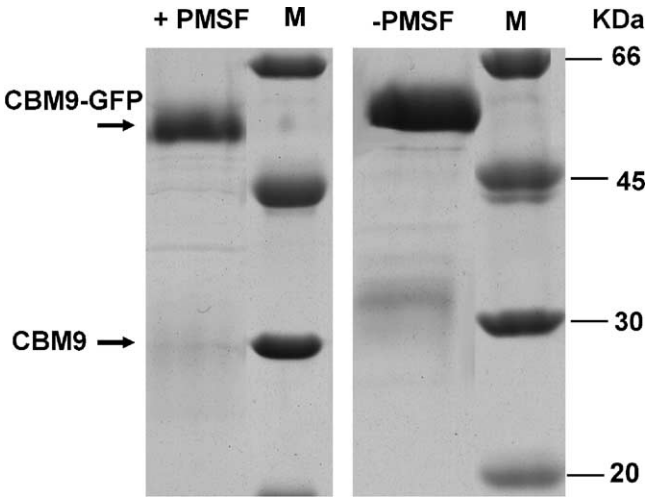

Fig. 3. Proteolytic stability of the S_3N_{10} linker in CBM9–GFP. 12% SDS–PAGE of CBM9–GFP purified with Perloza MT100 in a small batch system. PMSF treated cell extract containing CBM9–GFP was mixed end-over-end, washed with buffer and desorbed with 1 M glucose in TBS8.

on an SDS–PAGE gel corresponding to (or close to) the molecular mass of CBM9. As shown in Fig. 3, when PMSF is added to the washed cell suspension, a very small amount of proteolytic degradation of the S_3N_{10} linker occurs, either in vivo or during the cell processing and affinity purification steps. In the absence of a protease inhibitor, a slightly larger fraction of the recombinantly expressed CBM9–GFP fusion protein is lost due to degradation within the linker region. Under both processing conditions, however, the vast majority of expressed fusion protein remains intact through the induction, cell lysis and affinity chromatography steps.

3.3. Affinity purification on Perloza MT100 column

A typical chromatogram for affinity purification on a Perloza MT100 capture column of CBM9–GFP from an *E. coli* BL21 clarified cell lysate is shown in Fig. 4. No protease inhibitor (PMSF) was added to the cell suspension or lysate. The corresponding SDS–PAGE gel documentation of the purification process is shown in Fig. 5, and a summary of the fusion protein yield, purity, and concentration factor following elution from the Perloza MT100 column is provided in Table 3.

Table 3
Summary of purification of CBM9–GFP on Perloza MT100 at 4 °C

	Protein[a] (mg)	Yield (%)	Purity[b] (%)	Concentration factor
Cell extract	71.5	100		
Elution	61.5	86 (±3.6)	>95	45.7 (±9.9)
Free GFP (after tag removal)	60.2	84	>95	

[a] Protein concentration was quantified by fluorescence using a calibration curve.
[b] Purity determined by SDS–PAGE.

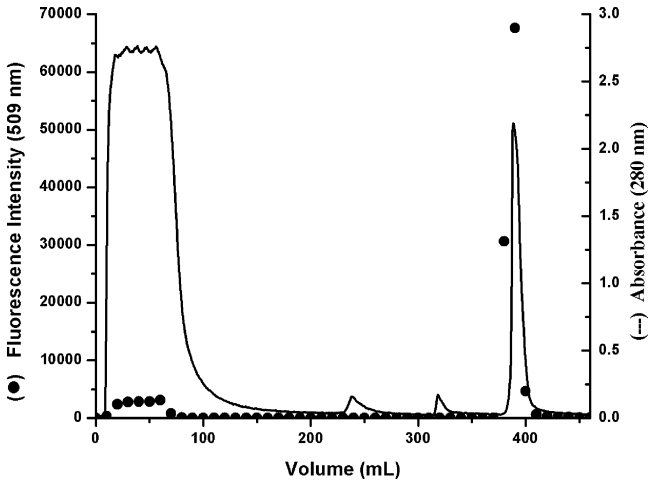

Fig. 4. Chromatogram of CBM9–GFP purification from an *E. coli* BL21 clarified cell lysate on Perloza MT100 at 4 °C. Fifty milliliter of clarified cell extract was loaded at 0.2 ml min^{-1} on a 17 ml column packed with Perloza MT100 resin, and then washed with 10 column volumes (CV) high salt buffer and 5 CV low salt buffer. Bound fusion protein was desorbed with 1 M glucose in TBS8. Ten milliliter fractions were collected and analyze by fluorescence (509 nm) and absorbance at 280 nm.

The intrinsic fluorescence of GFP allows us to monitor simultaneously the elution of contaminating proteins (UV absorbance at 280 nm) and the concentration (fluorescence intensity at 509 nm) of CBM9–GFP and its degradation products in each elution fraction. A small amount of CBM9–GFP or GFP within the clarified lysate load is lost in the column flowthrough. It is likely that most if not all of this fluorescent material represents the small amount of fusion protein that is degraded within the S_3N_{10} linker region, as shown in Fig. 3. Weakly bound contaminating proteins are sequentially removed in the column flowthrough and the two column wash steps. No loss of CBM9–GFP is detected in either wash step (Fig. 5, Lanes 3 and 4).

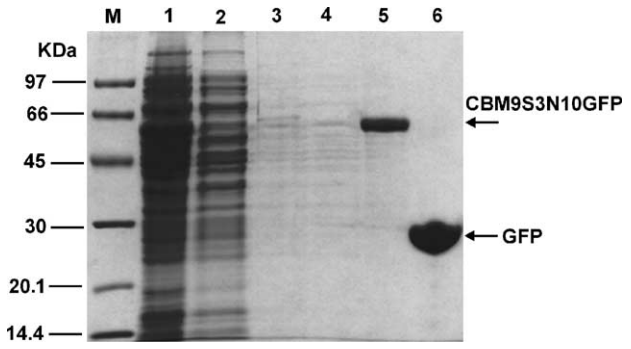

Fig. 5. SDS–PAGE documentation of the affinity purification of CBM9–GFP. 12% SDS–PAGE of CBM9–GFP purified on a 17 ml Perloza MT100 column. All samples dissolved in sample buffer containing 10% SDS. Lane M: molecular mass markers in kg mol^{-1}; Lane 1: clarified cell extract prior to column loading; Lane 2: column flow through; Lane 3: high salt wash; Lane 4: low salt wash; Lane 5: pure CBM9–GFP eluted in TBS8 containing 1 M glucose; Lane 6: purified GFP after affinity-tag removal by immobilized Factor X_a.

A 1 M glucose solution (in TBS8) is effective in quantitatively eluting all specifically bound fusion protein (Fig. 5, Lane 5). CBM9–GFP elutes from the column in a single sharp peak, as is evident from the overlapping A_{280} and fluorescence intensity peaks in the chromatogram. The purity of CBM9–GFP in the pooled fractions of the elution peak was >95% at an average yield of 86%. Both values are competitive with (in fact superior to) the published performance of other commercially available affinity tag systems, including the GST and poly-His fusion-tag technologies [28–33].

In these experiments, the Perloza MT100 column was loaded to less than half saturation capacity to guarantee capture of all CBM9–GFP from the clarified cell lysate. Despite operating the column at under-loaded conditions, a remarkably high concentration factor of ca. 46 was achieved, suggesting that the fusion protein loads, binds and elutes from the column in a reasonably tight band.

3.4. Column reusability

As the cost of any affinity chromatography technology is largely determined by the purchase price and reusability of the capture resin, we investigated the ability of the Perloza MT100 resin to provide acceptable and predictable purification performance with repeated column use. Six consecutive purifications were performed on a single Perloza MT100 column to identify any changes in column performance with increasing number of purification cycles. Very high product purity (>95%) was achieved in all six purification cycles. As shown in Table 4, product yield and concentration factor, however, were affected by repeated column use. An average yield of 86 (±3.6)% was observed for the affinity purification of CBM9–GFP from clarified cell lysate on a clean, freshly poured Perloza MT100 column. Slightly lower yields of ca. 79% were then consistently observed for each purification cycle thereafter. The product concentration factor followed the same trend, with a measured concentration factor of 46 (±9.9) for the first column cycle falling to a consistent value of ca. 28 for each subsequent cycle. The source of these modest changes is unclear. However, the repeatable good performance (>95% purity, 79% yield, concentration

Table 4
CBM9–GFP yield and purity for consecutive purification runs through the same column

Column cycle[a]	Yield[b] (%)	Concentration factor
1	86	45.7
2	78.4	28.8
3	81.2	29.2
4	79	28.8
5	78.8	26.8
6	78	27.5

[a] All runs gave a CBM9–GFP purity of >95% as determined by SDS–PAGE.

[b] Protein concentrations were quantified by fluorescence using a calibration curve.

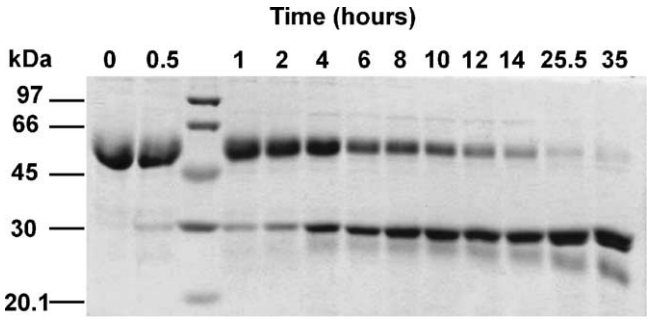

Fig. 6. Time course of CBM9–GFP cleavage by Factor X_a at 23 °C as shown on a 12% SDS–PAGE. A fusion protein to Factor X_a concentration ratio of 1000:1 was used.

factor of 28) of the column following the first column cycle suggests that our CBM9 fusion tag technology can provide a robust platform for affinity purification of recombinant proteins.

3.5. Removal of the CBM9-S_3N_{10}-IEGR affinity tag using an immobilized Factor X_a column

In certain cases, such as in the production of a human therapeutic protein, removal of the fusion tag following purification is required to recover the desired target protein with its natural N-terminus. We, therefore, have incorporated a Factor X_a processing site adjacent to the N-terminus of the target protein to facilitate tag removal by specific enzymatic cleavage. Fig. 6 is an SDS–PAGE gel showing the kinetics of tag cleavage when purified CBM9–GFP is processed at 21 °C and pH 8 with CBM2a-FXa_{im} at a fusion-protein to Factor X_a concentration ratio of 1000:1. To avoid stagnant settling of the CBM2a-FXa_{im} Perloza MT500 resin, the reaction mixture, which also contained 1 M glucose in the liquid phase, was mixed end-over-end in an orbital mixer. In the presence of 1 M glucose, CBM9 does not bind to Perloza MT500, while binding of CBM2a-FXa is irreversible at these conditions. Complete cleavage of the CBM9-S_3N_{10}-IEGR fusion tag was observed after 28 h.

The Factor X_a treated solution was then diafiltered on a 1 K cut-off filter to remove the 1-M glucose and loaded onto a second Perloza MT100 column to capture the cleaved CBM9 tag. Pure, N-terminally correct GFP was collected in the flow through with a yield of 98% and a purity of >95%. This resulted in an overall yield of the purified target protein (GFP) of 84% when a fresh Perloza MT100 column was used, or 77% when the same column was used for multiple purification cycles.

4. Conclusions

We have shown that proteins expressed in *E. coli* as fusions with the family 9 carbohydrate-binding module of xylanase 10A from *T. maritima* can be affinity purified on a cellulose-based Perloza MT100 column. The performance of our technology is competitive with all commercial fusion tag systems, and may offer advantages with respect to improving the expression of the target protein in a soluble form.

Acceptance and use of affinity tag systems in manufacturing of recombinant proteins have been slowed, at least in part, by the associated costs of the technology, particularly the cost of the resin. Perloza MT100 is a simple, highly porous regenerated cellulose/cellulose xanthate of uniform particle size and flow characteristics. The polymer bead structure is stabilized by hydrogen bonds only; there are no covalent cross-links within the resin. As a result, it is a durable and surprisingly inexpensive resin. When bought in bulk quantities, the cost of Perloza MT100 is ca. US$ 35 per liter of resin, which, for example, is close to 1/100 the cost of an equivalent volume of Ni^{2+}-NTA (IMAC) resin used to purify poly-His tagged proteins. The cost of Perloza MT100 resin also compares very favorably with the costs of those resins designed to capture fusion proteins tagged with GST or calmodulin binding protein. Direct capture on a packed column of Perloza MT100, which binds CBM9-tagged proteins with extraordinarily high capacity (in excess of 600 mg g^{-1} resin), therefore appears to offer a robust and inexpensive strategy for affinity purification of proteins expressed in soluble form as fusions with the CBM9 tag.

Acknowledgements

The authors would like to thank Professor Tony Warren for helpful discussions and use of his laboratory space and equipment. This work was supported by grants from the Protein Engineering Network of Centres of Excellence (PENCE) and NSERC.

References

[1] M.R. Ladisch, Bioseparations Engineering : Principles, Practice, and Economics, Wiley, New York, 2001.
[2] M.T. Hearn, D. Acosta, J. Mol. Recognit. 14 (2001) 323.
[3] H.A. Chase, Chem. Eng. Sci. 39 (1984) 1099.
[4] C.R. Lowe, A.R. Lowe, G. Gupta, J. Biochem. Biophys. Methods 49 (2001) 561.
[5] M. Wilchek, I. Chaiken, Methods Mol. Biol. 147 (2000) 1.
[6] J. Nilsson, S. Stahl, J. Lundeberg, M. Uhlen, P.A. Nygren, Protein Expr. Purif. 11 (1997) 1.
[7] K. Terpe, Appl. Microbiol. Biotechnol. 60 (2003) 523.
[8] P. Jonasson, S. Liljeqvist, P.A. Nygren, S. Stahl, Biotechnol. Appl. Biochem. 35 (2002) 91.
[9] P. Vaillancourt, T.G. Simcox, C.F. Zheng, Biotechniques 22 (1997) 451.
[10] C.F. Zheng, T. Simcox, L. Xu, P. Vaillancourt, Gene 186 (1997) 55.
[11] K.L. Guan, J.E. Dixon, Anal. Biochem. 192 (1991) 262.
[12] D.B. Smith, K.S. Johnson, Gene 67 (1988) 31.
[13] J. Crowe, H. Dobeli, R. Gentz, E. Hochuli, D. Stuber, K. Henco, Methods Mol. Biol. 31 (1994) 371.
[14] E. Hochuli, H. Dobeli, A. Schacher, J. Chromatogr. 411 (1987) 177.
[15] J. Porath, J. Carlsson, I. Olsson, G. Belfrage, Nature 258 (1975) 598.

[16] C. Winterhalter, P. Heinrich, A. Candussio, G. Wich, W. Liebl, Mol. Microbiol. 15 (1995) 431.
[17] A.B. Boraston, A.L. Creagh, M.M. Alam, J.M. Kormos, P. Tomme, C.A. Haynes, R.A. Warren, D.G. Kilburn, Biochemistry 40 (2001) 6240.
[18] K. Nagai, H.C. Thogersen, Nature 309 (1984) 810.
[19] A. Crameri, E.A. Whitehorn, E. Tate, W.P. Stemmer, Nat. Biotechnol. 14 (1996) 315.
[20] O. Shimomura, F.H. Johnson, Y. Saiga, J. Cell Comp. Physiol. 59 (1962) 223.
[21] J. Sambrook, E.F. Fritsch, T. Maniatis, Molecular Cloning : a Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989.
[22] E. Kwan, M.M. Guarna, A.B. Boraston, N.R. Gilkes, C.A. Haynes, D.G. Kilburn, R.A. Warren, Biotechnol. Bioeng. 79 (2002) 724.
[23] H. Mach, C.R. Middaugh, R.V. Lewis, Anal. Biochem. 200 (1992) 74.
[24] M. Chalfie, Y. Tu, G. Euskirchen, W.W. Ward, D.C. Prasher, Science 263 (1994) 802.
[25] M. Chalfie, S. Kain, Green Fluorescent Protein : Properties, Applications, and Protocols, Wiley-Liss, New York, 1998.
[26] J.D. Fox, R.B. Kapust, D.S. Waugh, Protein Sci. 10 (2001) 622.
[27] J.D. Fox, D.S. Waugh, Methods Mol. Biol. 205 (2003) 99.
[28] G. Chaga, D.E. Bochkariov, G.G. Jokhadze, J. Hopp, P. Nelson, J. Chromatogr. A. 864 (1999) 247.
[29] M.C. Baek, K.H. Choi, T.G. Oh, D.H. Kim, E.C. Choi, Protein Expr. Purif. 11 (1997) 257.
[30] M. Chang, J.L. Bolton, S.Y. Blond, Protein Expr. Purif. 17 (1999) 443.
[31] T.K. Littlejohn, O. Takikawa, D. Skylas, J.F. Jamie, M.J. Walker, R.J. Truscott, Protein Expr. Purif. 19 (2000) 22.
[32] K.B. Appa Rao, L.C. Garg, A.K. Panda, S.M. Totey, Protein Expr. Purif. 11 (1997) 201.
[33] A. Modesti, N. Taddei, M. Bucciantini, M. Stefani, B. Colombini, G. Raugei, G. Ramponi, Protein Expr. Purif. 6 (1995) 799.