

available at www.sciencedirect.comjournal homepage: www.elsevier.com/locate/aca

Bioseparation of recombinant cellulose-binding module-proteins by affinity adsorption on an ultra-high-capacity cellulosic adsorbent

Jiong Hong^{a,b}, Xinhao Ye^a, Yiran Wang^a, Y.-H. Percival Zhang^{a,c,d,*}

^a Biological Systems Engineering Department, Virginia Polytechnic Institute and State University, 210-A Seitz Hall, Blacksburg, VA 24061, USA

^b School of Life Science, University of Science and Technology of China, 443 Huangshan Road, Hefei, Anhui 230027, PR China

^c Institute for Critical Technology and Applied Science (ICTAS), Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA

^d DOE BioEnergy Science Center (BESC), Oak Ridge, TN 37831, USA

ARTICLE INFO

Article history:

Received 7 March 2008

Received in revised form

11 May 2008

Accepted 15 May 2008

Published on line 27 May 2008

Keywords:

Affinity adsorbent

Bioseparation

Cellulose-binding module

Protein purification

Regenerated amorphous cellulose

ABSTRACT

Low-cost protein purification methods are in high demand for mass production of low-selling price enzymes that play an important role in the upcoming bioeconomy. A simple protein purification method was developed based on affinity adsorption of a cellulose-binding module-tagged protein on regenerated amorphous cellulose (RAC) followed by modest desorption. The biodegradable cellulosic adsorbent RAC had a very high protein-binding capacity of up to 365 mg of protein per gram of RAC. The specifically-bound CBM-protein on the external surface of RAC was eluted efficiently by ethyl glycol or glycerol. This protein separation method can be scaled up easily because it is based on simple solid/liquid unit operations. Five recombinant proteins (CBM-protein), regardless of inter-cellular or periplasmic form, were purified successfully for demonstration purpose.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

It is challenging to develop simple, low-cost, and scalable methods for large-scale recombinant protein purification with a reasonable separation resolution [1–3]. Protein purification by affinity chromatography based on adsorption of a tag on the respective affinity resin is popular in labora-

tories and biotechnology companies [1,4,5]. But a number of factors, such as costly resin, batch operation, complicated scale-up, low adsorption capacity of affinity resin, slow association/dissociation rates of the target protein on resin, potential column fouling, and slow flow rate, economically prohibit affinity chromatography from separation of low-selling-price industrial enzymes. Low-cost scalable protein

Abbreviations: Avicel, commercial name of microcrystalline cellulose; BGL, β -glucosidase; CBM, cellulose binding module; CBP, cellobiose phosphorylase; CG, CBM-GFP; EG, ethylene glycol; GFP, green fluorescent protein; GNP, α -glucan phosphorylase; OCB, ompA-CBM-BG; PGM, phosphoglucomutase; RAC, regenerated amorphous cellulose.

* Corresponding author. Tel.: +1 540 231 7414; fax: +1 540 231 7414.

E-mail address: ypzhang@vt.edu (Y.-H.P. Zhang).

0003-2670/\$ – see front matter © 2008 Elsevier B.V. All rights reserved.

doi:10.1016/j.aca.2008.05.041

purification methods are of importance for implementation of biocommodity production catalyzed by enzymes or synthetic enzymatic pathways [6–10].

Adsorption, desorption, and solid/liquid separations (e.g., filtration, centrifugation) are among simple, low-cost, scalable unit operations. But traditional protein separation approaches based on these simple operations suffer from low product purity [2,3]. Recently, a precipitation-based protein separation method with a relatively high resolution has been developed by using an elastin (ELP) polypeptide tag. The elastin-containing fusion protein can precipitate or dissolve based on a temperature switch below or above the transition temperature [11]. But ELP tags, especially for longer elastin, inhibit recombinant protein expression [12]. Purification yields of low-expression ELP-tag protein are relatively low [12,13], especially for low-concentration recombinant protein (e.g., ~50% for 5 $\mu\text{g mL}^{-1}$, ~10% for 1 $\mu\text{g mL}^{-1}$) [14].

Cellulose-binding module (CBM) tags have been used for recombinant protein purification by using commercial cellulose matrix or powder (Avicel—microcrystalline cellulose or SigmaCell) [15–22]. The binding capacities of cellulosic materials are closely associated with the size of the adsorbed molecules [23]. Commercial microcrystalline cellulose has a low binding capacity, most of whose binding surface is internal [22,24]. Furthermore, the bound protein cannot be removed efficiently due to the adsorbed protein entrapment effect [25], resulting in relatively low protein recovery yields from the pores of cellulose matrix.

Regenerated amorphous cellulose (RAC) is made from microcrystalline cellulose through cellulose dissolution by concentrated phosphoric acid followed by cellulose regeneration by water precipitation [26]. It has a greater than 20-fold higher surface area of microcrystalline cellulose [23,24,27]. In addition, the entire binding surface of RAC is externally accessible to the target protein [24], implying that the binding capacity RAC could be nearly independent of protein size.

In this study, we developed a generic, low-cost, scalable protein purification method based on affinity adsorption on a low-cost and high adsorption capacity adsorbent RAC followed by modest desorption.

2. Materials and methods

2.1. Chemicals and strains

All chemicals were reagent grade or higher, purchased from Sigma (St. Louis, MO, USA) and Fisher Scientific (Pittsburgh, PA, USA), unless otherwise noted. Microcrystalline cellulose – Avicel PH105 (20 μm) – was purchased from FMC (Philadelphia, PA, USA). *Escherichia coli* DH5 α was used as a host cell for all DNA manipulation. *E. coli* BL21 Star (DE3) (Invitrogen, Carlsbad, CA, USA) containing a protein expression plasmid was used for producing the recombinant protein. Luria-Bertani (LB) medium was used for *E. coli* growth and protein expression with 100 $\mu\text{g mL}^{-1}$ ampicillin. *Paenibacillus polymyxa* ATCC 842 was grown on the ATCC potato medium. *Clostridium thermocellum* genomic DNA was gifted from Dr. Mielenz at the Oak Ridge National Laboratory (Oak Ridge, TN, USA). The oligonu-

cleotides were synthesized by Integrated DNA Technologies (Coraville, IA, USA) (Table 1).

2.2. Regenerated amorphous cellulose preparation [26]

Approximately 0.2 g of microcrystalline cellulose and 0.6 mL distilled water were added into a 50-mL centrifuge tube to form a cellulose-suspension slurry. Ten milliliters of ice-cold 86% H_3PO_4 (i.e., commercial 85% grade) was slowly added into the slurry with vigorous stirring so that the final phosphoric acid concentration was ca. 83.2%. The cellulose mixture turned transparent within several minutes. It stood for ca. an hour in an ice bath with occasional stirring. Approximately 40 mL of ice-cold water was added at a rate of approximately 10 mL per addition with vigorous stirring between additions, resulting in a white cloudy precipitate. The precipitated cellulose was centrifuged at $\sim 10,000 \times g$ and 4 °C for 20 min. The pellet was suspended in ice-cold water, followed by centrifugation to remove the phosphoric acid from the supernatant, four times. Approximately 0.5 mL of 2 M Na_2CO_3 and 40 mL of ice-cold distilled water were used to neutralize and suspend the cellulose pellet. After centrifugation, the pellet was washed by distilled waters twice or until pH 5–7. The carbohydrate concentration of RAC was calibrated by the phenol- H_2SO_4 method [28]. Other cellulose powders such as SigmaCell can be used to replace FMC PH105. The RAC slurry can be stored as a ~ 10 g RACL^{-1} suspension solution at 4 °C in the presence of 0.2% (w/v) sodium azide for at least 1 year.

2.3. Intracellular recombinant protein expression plasmid construction

The pCG plasmid encoding the CBM-GFP (CG) fusion protein was constructed based on the New England Biolabs plasmid pTWIN1 (Ipswich, MA, USA). The cellulose-binding module (*cbm*) gene fragment of *C. thermocellum* cellulosomal scaffoldin was amplified from the plasmid pNT02 [24] by PCR using the primers of CBM-F1 and CBM-R2; the *gfp* gene was amplified from the plasmid pNT02 by PCR with the primers of GFP-F2 and GFP-R1. Then *cbm* and *gfp* genes were connected by overlap PCR with the primers of CBM-F1 and GFP-R1. The *cbm-gfp* DNA fragment was digested by NdeI/PstI and then ligated with the NdeI/PstI-digested pTWIN1 to generate the plasmid pCG. The other intracellular target genes were amplified by the primers (Table 1), digested by XhoI/PstI, and then ligated with the XhoI/PstI-digested plasmid pCG. They included the *C. thermocellum* cellobiose phosphorylase (CBP) [29,30], phosphoglucomutase (PGM, CT1265), and putative α -glucan phosphorylase (GNP, CT0932).

2.4. Periplasmic recombinant protein expression plasmids

The pOCB plasmid encoding the ompA-CBM-BGL (OCB) fusion protein was constructed based on the plasmid pCG. The *ompA* signal sequence was amplified from the plasmid pFLAG-CTS (Sigma, St. Louis, MO, USA) with the primers of OmpA-F and OmpA-R; the *cbm* was amplified from plasmid pCG with the primers of CBM-F and CBM-R. Then the *ompA* and *cbm* DNA fragments were connected by overlap PCR with the primers of

Table 1 – The oligonucleotides for constructing recombinant protein expression plasmids

Name	Sequence
CBM-F1	5' GGTGGT <u>CATATGCCGGTATCAGGCAATTGAAGTTGAATTC</u> 3'
CBM-R2	5' CCCTCGAGGCCGCCAGGCCT GGGTTCTTTACCCATACAAGAACACCG 3'
GFP-F2	5' AGGCCTGGGGCCTCGAGGGCTCTTCCATGGTGAGCAAGGGCGAGAGCTGTTC 3'
GFP-R1	5' GGTGGTCTGCAGTTACTTGTACAGCTCGTCCATGCCGAG 3'
CBP-F2	5' ACTGCTCGAGATACCGCCGTCAGATGATCCGATGAAGTTCCGGTTTTTTTGTATGATGC 3'
CBP-R1	5' ACTGCTGCAGTTAGGTACCACTAGTATCGATCCATAATTACTTCAACTTTGTGAGTC 3'
ompA-F	5' ACTGACTGCATATGAAAAAGACAGCTATCGCGATTG 3'
ompA-R	5' ATTGCCTGATACCGG AGCTTGCGCAACGGTAGCGAAACCA 3'
BGL-F	5' ATGCCTCGAGATGACTATTTTCAATTTCCGCAGGACTTTATG 3'
BGL-R	5' ATCGCTGCAGTTAGCGTCTAGTCTCCAACCAAGTTATTAC 3'
GNP-F1	5' TGGTGGCTCGAGATGATCTTTTTGGAAAAATTAC 3'
GNP-R1	5' AAGAAGGGATCCTTACTGTACAATCCATCTGATAAGTCC 3'
PGM-F	5' GCATCGCTCGAG GGCTCTTCC ATGCGAAGTAGCGGCTTTAT 3'
PGM-R	5' ACGTGC <u>GGAT CCTCAGTCTT</u> TAAGAAGCGG TTCTATAAC 3'

Underline: the restriction enzyme site; italic: sequence for overlap PCR.

ompA-F and CBM-R. The *Paenibacillus polymyxa* β -glucosidase (BGL) gene was amplified by PCR from the genomic DNA with the primers BGL-F and BGL-R. The *ompA-cbm* and *bgl* fragments were digested by NdeI/XhoI and XhoI/PstI, respectively, and then ligated into the NdeI/PstI digested pCG in one step for the plasmid pOCB.

2.5. Protein expression and purification

The strain *E. coli* BL21 (DE3) containing the recombinant protein expression plasmid was used for protein expression and purification. After inoculation, 200 mL of LB medium in a 1-L Erlenmeyer flask was incubated at a rotary shaking rate of 160 rpm at 37 °C until the A_{600} reached between ~0.6 and 0.8. The protein expression was induced by adding IPTG (0.20–1.0 mM, final) and then the cultivation temperature was decreased to 18 °C for 9–12 h. The cells were harvested by centrifugation at 4 °C, washed by 20 mL of 50 mM Tris-HCl buffer (pH 8.0) once, and then re-suspended by ~20 mL of 50 mM Tris-HCl buffer (pH 8.0). The intracellular protein was released by sonication of Fisher Scientific Sonic Dismembrator Model 500 (3-s pulse, total 90 s, at the maximum strength) or the secreted periplasmic recombinant protein (CBM-BGL) was released by osmotic shock treatment, as described elsewhere [31]. After centrifugation, the supernatant of cell lysate or the supernatant containing the released periplasmic protein was mixed with a cellulosic adsorbent—RAC slurry. The RAC amount used was estimated at a ratio of 200 mg of target protein to 1 g of RAC (i.e., RAC was in excess to adsorb >90% of the target protein), where the amount of target protein was roughly estimated by the enzyme activity or SDS-PAGE. After adsorption at room temperature for ~10–15 min followed by centrifugation, the RAC pellet was suspended in 20 mL of 50 mM Tris-HCl buffer (pH 8.0) to remove impure proteins in the RAC matrix. After centrifugation, the RAC pellet containing the adsorbed CBM-target protein was suspended in 4 RAC pellet volumes of 100% ethyl glycol (EG), i.e., the final EG concentration was ~80% (v/v). After centrifugation, the purified CBM-target protein was obtained in the supernatant. The purified protein can be stored in EG solution at –20 °C. In order to obtain high concentration protein, a low volume of EG can be

used. In order to obtain high protein yields, the pellet can be washed several times by a small volume of EG. Ethylene glycol can be removed by dialysis, and then the dilute protein can be re-concentrated by Pall ultra-filtration centrifugal tubes. Similarly, glycerol can be used to replace EG for elution if EG inhibits enzyme activity.

2.6. Adsorption of CG on RAC and Avicel

The equilibrium adsorption of CBM-GFP on RAC or Avicel was conducted at room temperature. According to the Langmuir equilibrium, the maximum adsorption capacity (A_{max} , mg protein per g cellulose) of RAC or Avicel was calculated, as described elsewhere [23,24]. Dynamic adsorption of CG was measured at two concentrations of adsorbent (RAC and Avicel) at a ratio of $A_{max} \times Wt/CG$ of 2:1 or 1:1. RAC and Avicel slurry concentrations were 10 and 100 g L⁻¹ for sample transferring, respectively. After well mixing, the cellulose suspension solutions were withdrawn into 1.5-mL microcentrifuge tubes at different times, followed by immediate centrifugation. The GFP fluorescence in the supernatant was measured by a BioTek multi-detection microplate reader.

2.7. Low concentration protein capture

Two hundred microliters of the cell CG lysate (0.096 mg of CBM-GFP) was mixed with a 50 mM Tris-HCl buffer (pH 8.0), 2.0 mg of BSA mL⁻¹ in Tris buffer, or *E. coli* BL21 cell lysate at a, v/v, ratio of 1:2–200, respectively. Consequently, the concentrations of the dilute target protein solutions were 240 μ g mL⁻¹ (0.4 mL) to 2.4 μ g mL⁻¹ (40 mL). RAC was added at two ratios of 240 mg protein g⁻¹ RAC and 120 mg protein g⁻¹ RAC. After adsorption and washing, the adsorbed protein was eluted by using 0.2 mL of EG.

2.8. Enzyme assays and protein analysis

GFP fluorescence was detected (excitation at 485 nm, emission at 528 nm) by the Bio-Tek multi-detection microplate reader (Winooski, VT, USA). Cellobiose phosphorylase (CBP) activity was measured, as described elsewhere [29,32]. Putative α -glucan phosphorylase (GNP) activity was determined by

measuring an increase in inorganic phosphate in a reaction solution containing 10 mM soluble starch, 10 mM glucose-1-phosphate in a 50 mM HEPES buffer (pH 7.0) at 37 °C. One unit of CBP or GNP activity was defined as the amount of enzyme that produced 1 μ mol of phosphate per 15 min under the above conditions. β -Glucosidase activity was assayed at 37 °C using the substrate of *p*-nitro-phenyl- β -D-glucopyranoside (pNPG) [33,34]. PMG activity was measured in a 50 mM HEPES buffer (pH 7.5) with 5 mM glucose-1-phosphate, 0.5 mM Mn^{2+} , 5 mM Mg^{2+} , 1 mg mL⁻¹ BSA, and 0.1% Triton X-100 at 60 °C. The protein mass concentration was determined by the Bradford method based on a standard protein of bovine serum albumin. SDS polyacrylamide gel electrophoresis (SDS-PAGE) at 12% was performed in a Tris-glycine buffer [35].

3. Results

A simple protein separation method was designed based on high-affinity adsorption of the CBM-tagged protein on the surface of RAC followed by modest desorption. After sonication for intracellular protein or osmotic shock for periplasmic protein, followed by the first centrifugation, the CBM-target protein, along with other impure proteins, is mixed well with RAC slurry. After the second centrifugation, the supernatant containing most impure proteins is decanted. The impure proteins in the RAC matrix are washed away in a washing buffer once or several times, depending on purity requirement. The CBM-target protein can be easily eluted from RAC by using an elution buffer—ethyl glycol or glycerol. After the fourth centrifugation, the purified protein (CBM-target protein) can be obtained in the supernatant. This protein purification process involving only solid/liquid separation (centrifugation) can be easily scaled up.

Regenerated amorphous cellulose, which is prepared through cellulose dissolution and regeneration, has a binding capacity as high as 365 \pm 20 mg CG per gram (Fig. 1). The maximum protein binding capacity of RAC depends on the size of target protein, the size of CBM and the link between target protein and CBM [24,36]. Based on CG adsorption, RAC has a nearly 25 times of the binding capacity of commercial Avicel (14.8 \pm 0.4 mg g⁻¹) (Fig. 1). For another protein—theoredoxin-GFP-CBM, RAC binding capacity (369 \pm 19 mg TCG g⁻¹) is nearly 17.5 times higher than Avicel's (21.0 \pm 0.85 mg TCG g⁻¹) [24]. In addition, all binding surfaces of RAC are externally accessible to the CBM-tagged protein or cellulase [24,26], while most of the Avicel binding surface is internal [24].

We further studied the adsorption time required for reaching the equilibrium at different loadings of RAC and Avicel (Fig. 2). Fig. 2A shows that more CG is adsorbed on RAC than on Avicel when the $A_{max} \times Wt$ to CG ratio is 1:1 for both RAC and Avicel (Wt, weight of adsorbent, g), where the accessible surface of one gram of RAC equals that of 24.6 g of Avicel. Since adsorption of CBM-tagged protein obeys the Langmuir isotherm, a significant fraction of the target protein is not bound at $(A_{max} \times Wt)/CG = 1:1$ (Fig. 2A). In order to efficiently adsorb the protein, the adsorbent must be in excess. When the $A_{max} \times Wt$ to CG ratio is 2:1, the adsorption rate on RAC is much faster than that on Avicel (Fig. 2B). Ten-minute time is sufficient to adsorb nearly all of the target protein for RAC,

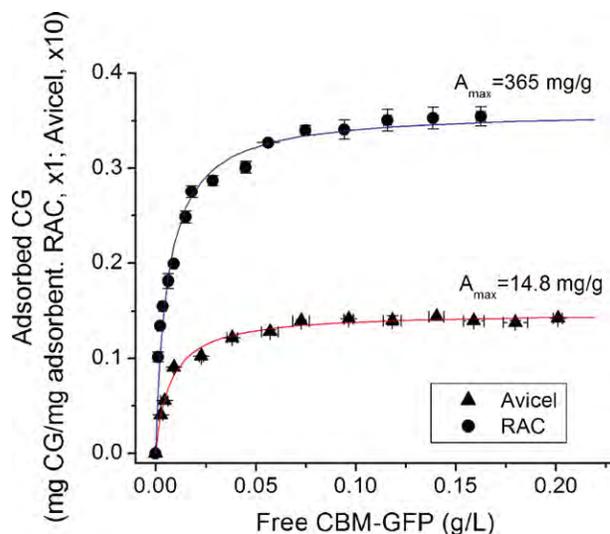


Fig. 1 – Various amounts of CBM-GFP protein adsorption on a fixed amount of RAC and Avicel at room temperature. The curves are fit and the A_{max} values are calculated based on the Langmuir equation.

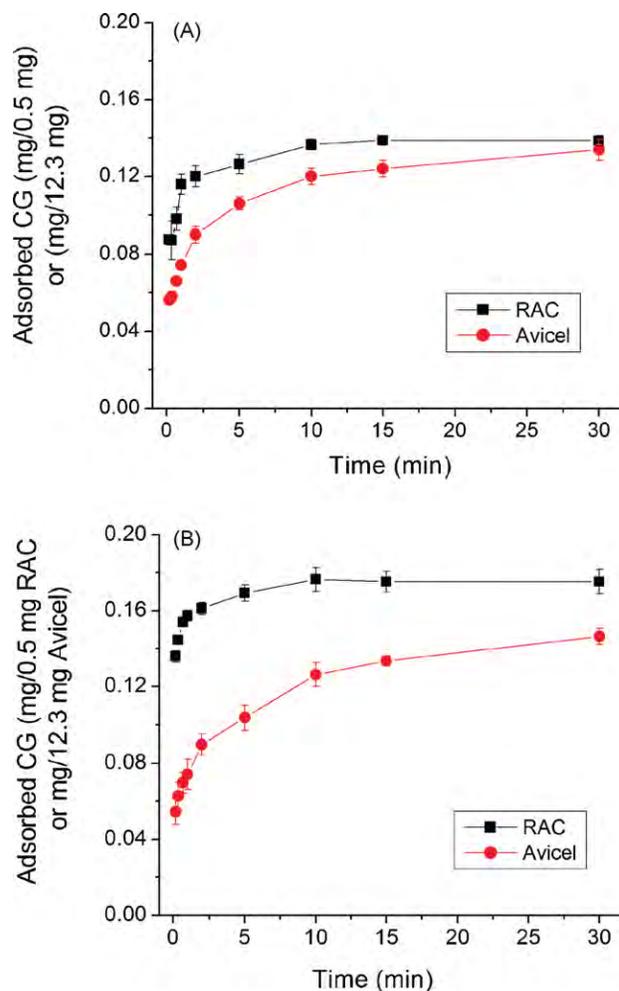


Fig. 2 – Dynamic adsorption of the CG protein on RAC and Avicel at the same $A_{max} \times Wt/CG$ ratios of 1:1 (A) and 1:2 (B) at room temperature for determining the time required for reaching the equilibrium.

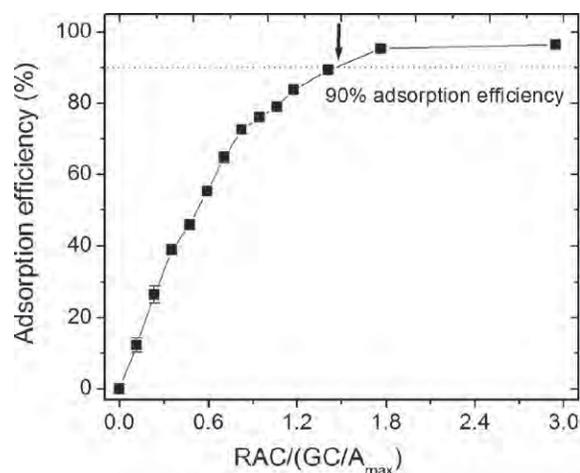


Fig. 3 – Adsorption of the CG protein in terms of $RAC/(CG/A_{max})$ for determining a minimal RAC loading for 90% adsorption of the CBM-target protein.

while a much longer time (~30 min) is needed for Avicel. The shorter adsorption time for RAC is consistent with the belief that protein mass transfer on the external surface of RAC is much faster than that occurring on the internal surface of Avicel. It is found that ~87% of Avicel's CBM-protein binding capacity is internal [24] and a much larger internal surface of Avicel is not accessibility to CBM-proteins but accessible to small molecules (e.g., nitrogen) [24]. In addition, higher protein adsorption yields are obtained on RAC than on Avicel at the same $(A_{max} \times Wt)/CG$ (Fig. 2B).

In order to determine an optimal amount of RAC added for the desired purification yield, various amounts of RAC were used to adsorb a definitive amount of target protein (Fig. 3). It is found that ~1.6-fold of CG/A_{max} RAC enables to adsorb 90% of the target protein, i.e., 1 g of RAC adsorbs ~228 mg of CG. The elution efficiency of the adsorbed CBM-GFP increased with increasing concentrations of EG or glycerol and the more volume used. For example, a 4-pellet volume of 70% EG (final concentration) can wash 90% of the target protein CG from RAC. Similarly, a 4-pellet volume of 80% glycerol can wash more than 80% of the CG. Higher protein yields can be achieved through multiple-step elution of a small volume EG.

Fig. 4 presents the images of CBM-GFP allocation during the purification processes. The cell lysate of *E. coli* BL21 (pCG) showed a strong fluorescence under UV radiation (tube 1), indicating the presence of GFP protein. After RAC addition and centrifugation, the RAC with adsorbed CG (tubes 2 and 3) exhibited a strong green color, while the color of the super-

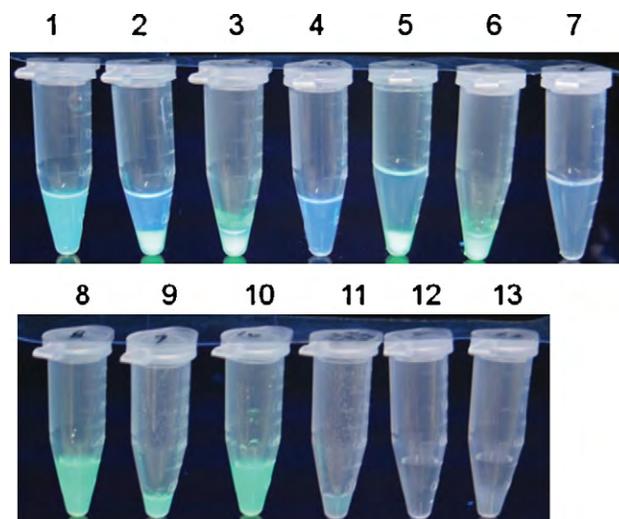


Fig. 4 – Images of the CBM-GFP purification processes under UV radiation. 1, the cell lysate of *E. coli* BL21 (pCG); 2, cell lysate + RAC after centrifugation; 3, the RAC pellet with the adsorbed target protein; 4, the supernatant after centrifugation; 5, the pellet suspended in ~0.5 mL of 50 mM Tris-HCl wash buffer (pH 8.0) followed by centrifugation; 6, the washed RAC pellet; 7, the used washing buffer; 8, RAC pellet with adsorbed protein suspended in 80% EG; 9, the RAC pellet after the 1st EG wash; 10, the purified CBM-target protein; 11, the RAC pellet washed by EG twice; 12, Tris-HCl buffer (negative control-1); 13, RAC-suspension solution (negative control-2).

natant (tube 4) was much weaker than that in tube 1. After washed by the buffer, the RAC pellets with the adsorbed CG were mixed with EG (tube 8). After centrifugation, most of desorbed CG was present in the supernatant (tube 10) and some fraction of the desorbed CG remained in the RAC matrix (tube 9). In order to release the CG efficiently, a second washing was recommended. In results, the fluorescence of RAC declined and more CG can be eluted by EG. The overall CG yield was 63%, and the purified fold was 2.19 (Table 2).

The purification process was analyzed by SDS-PAGE (Fig. 5). The results showed that RAC specifically adsorbed the CG (lane 3) and impure proteins remained in the supernatant (lane 4). Washing steps can remove more impure proteins from the RAC (lanes 5 and 6). The desorbed protein by using EG appeared to be a single protein band (lane 7).

The high-binding capacity RAC can bind the target protein specifically even at a low concentration, i.e., efficiently capture

Table 2 – The CBM-GFP protein purification process

Fraction	Vol. (mL)	Sp. fluorescence (mL ⁻¹)	Total fluorescence	Protein (mg mL ⁻¹)	Target protein (mg)	Sp. fluorescence	Yield (%)	Purification fold
Soluble cell lysate	20	6,520	130,400	1.32	26.37	4,945	100	1
Adsorbed protein ^a	3		122,140		11.49	10,630	94	2.15
Eluted CG protein	12	6,825	81,900	0.63	7.57	10,818	63	2.19

^a Adsorbed protein is calculated based on the difference between the initial protein used and the supernatant protein after the RAC adsorption.

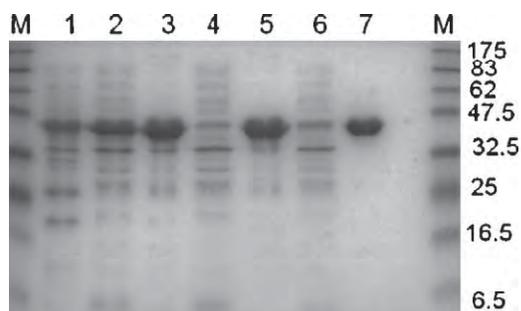


Fig. 5 – SDS-PAGE analysis of the CBM-GFP purification process. M, wide range pre-stain protein marker; (1) insoluble fraction of cell lysate; (2) soluble fraction of cell lysate; (3) RAC pellet with the adsorbed CG; (4) the supernatant after RAC adsorption; (5) the washed RAC pellet with the adsorbed CG; (6) the wash-out supernatant; (7) desorbed target protein by using EG. Each lane has 2.5 μg of total protein.

of protein from a dilute protein solution. Fig. 6 shows different CG yields for a large range of the target protein concentration from 2.4 to 480 $\mu\text{g mL}^{-1}$ in the presence of high concentration impure proteins ($\sim 2 \text{ mg mL}^{-1}$ BSA or *E. coli* cell lysate) or in the absence of added impure protein (i.e., dilution by Tris buffer). Higher protein yields can be obtained when the target protein concentrations are higher or more RAC is used. For most cases of recombinant protein expression where the target protein concentration could range around 50 $\mu\text{g mL}^{-1}$ to several hundred $\mu\text{g mL}^{-1}$, approximately 60–80% protein adsorption were expected to be obtained at the ratio of $\sim 200 \text{ mg}$ of target protein per gram of RAC (i.e., $\sim 55\%$ of binding surface of RAC occupied).

Four more enzymes (beta-glucosidase, BGL; cellobiose phosphorylase, CBP; glucan phosphorylase, GNP; and phosphoglucosmutase, PGM) were purified (Table 3). Among them, beta-glucosidase was expressed in a periplasmic form by the fusion with an outer membrane OmpA protein. Regardless of target protein location (intracellular or extracellular), target protein concentration, and relative ratio of target protein to total protein, this method can purify the desired proteins (Table 3). The higher protein yields were expected after optimization of amount of adsorbent used, volume of washing buffer, and so on.

4. Discussion

CBM is a powerful tag for protein separation because of (i) highly specific binding on cellulose, (ii) low non-specific bind-

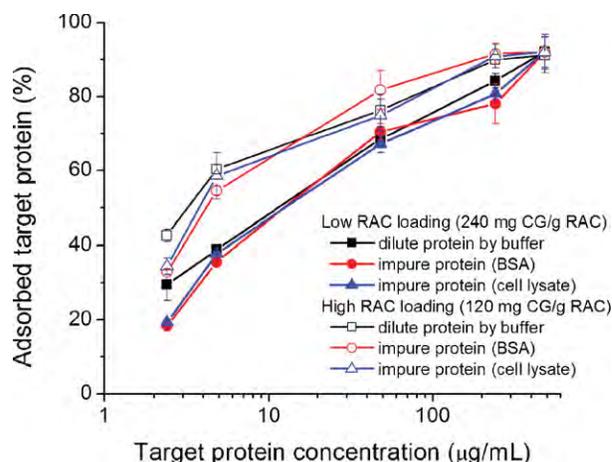


Fig. 6 – Adsorption efficiency of the different CG concentration solutions from 2.4 to 480 $\mu\text{g CG mL}^{-1}$ in the presence of impure proteins from the *E. coli* cell lysate ($\sim 2 \text{ mg}$ crude protein mL^{-1}), or 2 mg bovine serum albumin mL^{-1} , or in the absence of other added proteins at two different RAC loadings.

ing for other proteins, (iii) low-cost affinity matrix (cellulose) [15–18], (iv) enhanced protein folding [19], and (v) increased protein yields [20]. But the commercial cellulose matrix has a very low protein-binding capacity ($\sim 15 \text{ mg protein g}^{-1}$ of Avicel), and $\sim 87\%$ of its binding area is internal [24], resulting in slower adsorption rate and lower dissociation efficiency.

The use of the ultra-high-binding capacity adsorbent RAC for capturing CBM-tag proteins has several distinct advantages. (1) Relatively high protein purity. Based on the SDS-PAGE analysis results (Fig. 5), the purified CG is almost a single band after only one-step washing. (2) Reasonably high protein yield. In Tables 2 and 3, the yields obtained by one-step elution were between 29 and 63%. In favor of the $>90\%$ desorption (Table 2), the yield still can be improved by multi-step washing to recover the target proteins in the dead volume of RAC (each step using a small volume of elution buffer). (3) Shorter purification time due to faster protein adsorption and more efficient protein desorption for all external binding surface of RAC (Fig. 2). (4) Ultra-low-cost adsorbent RAC (e.g., $\sim 2 \text{ US cent g}^{-1}$ of RAC in the laboratory scale and as low as 0.02 US cent g^{-1} when large manufacturing is implemented [37]). Given 200 mg of bound target protein per gram of RAC, protein purification costs based on RAC could be as low as 10 or 0.1 US cents per gram of purified protein now or in the future, respectively. (5) Simple solid/liquid separation. The RAC and the eluted protein can be simply separated by centrifugation,

Table 3 – Purified protein quantification and activity assay

Protein	MW (kDa)	Purified protein ($\text{mg } 200 \text{ mL}^{-1}$ broth)	Specific activity	Yield (%)	Purification fold
CBM-GFP	45	7.6	$10,820 \text{ mg}^{-1}$	63	2.2
CBM-CBP	112	0.9	148 U mg^{-1}	36	6.3
CBM-PGM	84	2.4	415 IU mg^{-1}	55	2.0
CBM-GNP	117	1.2	353 U mg^{-1}	29	2.3
CBM-BGL	72	0.3	2.8 IU mg^{-1}	50	15.7

which can be scaled up easily with low capital investment. (6) Minimal waste treatment because only non-toxic chemicals are consumed and RAC is a biodegradable adsorbent. In addition, early capture of low concentration (secretory) protein enables rapid processing of large volumes of a very dilute target protein concentration.

CBM-containing proteins from cellulose matrix can be removed by a number of technologies, including low ion strength buffer or water washing [38,39], modest alkali pH [40,41], detergent plus an organic solvent [24,28], or ethylene glycol or glycerol [21]. But it is important to keep the desorbed protein active and to achieve relatively high yields. Ethylene glycol and glycerol are good choices. Most times, EG is preferable to glycerol because the former has lower viscosity (i.e., more easily handled) and is not consumed for most microorganisms. But glycerol could be used if EG is a strong inhibitor to the target enzyme, for example, β -glucosidase in this study. In addition, the eluted pure enzyme could be stored directly in EG or glycerol buffer at -20 or -80°C .

Product purity, yield, costs, and purification speed are four important consideration factors on the laboratory scale. More factors, such as process scalability, process complexity, capital investment, and waste treatment should be accounted for in industrial-scale product separation. Here a simple protein purification method has been developed to adsorb CBM-tag proteins by using the low-cost, ultra-high capacity adsorbent RAC followed by modest desorption. Five randomly picked CBM-tag proteins were purified by a simple solid/liquid unit operation–centrifugation.

5. Conclusion

A new CBM-tagged protein purification based on affinity adsorption on low-cost biodegradable cellulose was developed by using an ultra-high capacity cellulosic adsorbent RAC with a binding capacity of up to 365 mg protein per gram of RAC, enabling to efficiently capture (dilute) proteins. The adsorbed protein can be eluted efficiently by EG or glycerol. This scalable method can purify the CBM-tag protein by using a simple unit operation–centrifugation or filtration.

Acknowledgments

This work was supported by USDA-CSREES (2006-38909-03484), Air Force Young Investigator Award (FA9550-08-1-0145) and DuPont Young Faculty Award to YPZ.

REFERENCES

- [1] J. Arnau, C. Lauritzen, G.E. Petersen, J. Pedersen, *Protein Expr. Purif.* 48 (2006) 1.
- [2] T.M. Przybycien, N.S. Pujar, L.M. Steele, *Curr. Opin. Biotechnol.* 15 (2004) 469.
- [3] F. Hilbrig, R. Freitag, *J. Chromatogr. B* 790 (2003) 79.
- [4] D. Esposito, D.K. Chatterjee, *Curr. Opin. Biotechnol.* 17 (2006) 353.
- [5] J.L. Hartley, *Curr. Opin. Biotechnol.* 17 (2006) 359.
- [6] P. Lorenz, H. Zinke, *Trends Biotechnol.* 23 (2005) 570.
- [7] L.R. Lynd, C.E. Wyman, T.U. Germgross, *Biotechnol. Prog.* 15 (1999) 777.
- [8] Y.-H.P. Zhang, B.R. Evans, J.R. Mielenz, R.C. Hopkins, M.W.W. Adams, *PLoS One* 2 (2007) e456.
- [9] Y.-H.P. Zhang, M. Himmel, J.R. Mielenz, *Biotechnol. Adv.* 24 (2006) 452.
- [10] Y.-H.P. Zhang, *J. Ind. Microbiol. Biotechnol.* 35 (2008) 367.
- [11] D.E. Meyer, A. Chilkoti, *Nat. Biotechnol.* 17 (1999) 1112.
- [12] M.R. Banki, L. Feng, D.W. Wood, *Nat. Methods* 2 (2005) 659.
- [13] W.-Y. Wu, C. Mee, F. Califano, R. Banki, D.W. Wood, *Nat. Protocols* 1 (2006) 2257.
- [14] X. Ge, C.D.M. Filipe, *Biomacromolecules* 7 (2006) 2475.
- [15] C. Ramirez, J. Fung, R.C. Miller, R. Antony, J. Warren, D.G. Kilburn, *Nat. Biotechnol.* 11 (1993) 1570.
- [16] O. Shoseyov, Z. Shani, I. Levy, *Microbiol. Mol. Biol. Rev.* 70 (2006) 283.
- [17] M. Kavooosi, J. Meijer, E. Kwan, A.L. Creagh, D.G. Kilburn, C.A. Haynes, *J. Chromatogr. B* 807 (2004) 87.
- [18] P. Tomme, A.B. Boraston, B. McLean, J.M. Kormos, A.L. Creagh, K. Sturch, N.R. Gilkes, C.A. Haynes, R.A. Warren, D.G. Kilburn, *J. Chromatogr. B* 715 (1998) 283.
- [19] K. Murashima, A. Kosugi, R.H. Doi, *Proteins* 50 (2003) 620.
- [20] J. Ahn, E. Choi, H. Lee, S. Hwang, C. Kim, H. Jang, S. Haam, J. Jung, *Appl. Microbiol. Biotechnol.* 64 (2004) 833.
- [21] R.E. Novy, K.W. Yaeger, S.A. Miller, *in* *Novations* 7 (1997) 4.
- [22] J. Hong, Y. Wang, X. Ye, Y.-H.P. Zhang, *J. Chromatogr. A* 1194 (2008) 150.
- [23] Y.-H.P. Zhang, L.R. Lynd, *Biotechnol. Bioeng.* 88 (2004) 797.
- [24] J. Hong, X. Ye, Y.H.P. Zhang, *Langmuir* 23 (2007) 12535.
- [25] S.G. Desai, A.O. Converse, *Biotechnol. Bioeng.* 56 (1997) 650.
- [26] Y.-H.P. Zhang, J.-B. Cui, L.R. Lynd, L.R. Kuang, *Biomacromolecules* 7 (2006) 644.
- [27] Y.-H.P. Zhang, L.R. Lynd, *Biotechnol. Bioeng.* 94 (2006) 888.
- [28] Y.-H.P. Zhang, L.R. Lynd, *Biomacromolecules* 6 (2005) 1510.
- [29] Y.-H.P. Zhang, L.R. Lynd, *Appl. Environ. Microbiol.* 70 (2004) 1563.
- [30] Y.-K. Kim, M. Kitaoka, M. Krishnareddy, Y. Mori, K. Hayashi, *J. Biochem.* 132 (2002) 197.
- [31] N. Nossal, L. Heppel, *J. Biol. Chem.* 241 (1966) 3055.
- [32] J.K. Alexander, *J. Biol. Chem.* 243 (1968) 2899.
- [33] Y.-H.P. Zhang, J. Hong, X. Ye, *in*: J. Mielenz (Ed.), *Methods Mol Biol*, Humana Press, Totowa, 2008.
- [34] T.M. Wood, K.M. Bhat, *Methods Enzymol.* 160 (1988) 87.
- [35] Y.-H.P. Zhang, L.R. Lynd, *Anal. Chem.* 75 (2003) 219.
- [36] N.R. Gilkes, E. Jervis, B. Henrissat, B. Tekant, J.R.C. Miller, R.A.J. Warren, D.G. Kilburn, *J. Biol. Chem.* 267 (1992) 6743.
- [37] Y.-H.P. Zhang, S.-Y. Ding, J.R. Mielenz, R. Elander, M. Laser, M. Himmel, J.D. McMillan, L.R. Lynd, *Biotechnol. Bioeng.* 97 (2007) 214.
- [38] E. Morag, E.A. Bayer, R.L. Lamed, *Enzyme Microb. Technol.* 14 (1992) 289.
- [39] T. Reinikainen, O. Teleman, T.T. Teeri, *Proteins: Struct. Funct. Genet.* 22 (1995) 392.
- [40] D. Lee, A.H.C. Yu, J.N. Saddler, *Biotechnol. Bioeng.* 45 (1995) 328.
- [41] D.E. Otter, P.A. Munro, G. Scott, R. Geddes, *Biotechnol. Bioeng.* 34 (1989) 291.