MICROBIAL CELL FACTORIES

RESEARCH   **Open Access**

Identification of a heterologous cellulase and its N-terminus that can guide recombinant proteins out of *Escherichia coli*

Dongfang Gao[†], Shengjun Wang[†], Haoran Li, Huili Yu and Qingsheng Qi[*]

Abstract

Background: The Gram-negative bacterium *Escherichia coli* has been widely used as a cell factory for the production of proteins and specialty chemicals because it is the best characterized host with many available expression and regulation systems. However, recombinant proteins produced in *Escherichia coli* are generally intracellular and often found in the form of inclusion bodies. Extracellular production of proteins is advantageous compared with intracellular production because extracellular proteins can be purified more easily and can avoid protease attack, which results in higher product quality. In this study, we found a catalytic domain of a cellulase (Cel-CD) and its N-terminus can be employed as carriers for extracellular production of recombinant proteins.

Results: In this report, we identified the catalytic domain of a cellulase (Cel-CD) from *Bacillus* sp. that can be secreted into the medium from recombinant *E. coli* BL21 (DE3) in large quantities without its native signal peptide. By subcellular location analysis, we proved that the secretion was a two-step process and the N-terminal sequence of the full length Cel-CD played a crucial function in secretion. Both the Cel-CD and its N-terminal sequence can serve as carriers for efficient extracellular production of select target proteins.

Conclusions: Fusion of heterologous proteins with N20 from Cel-CD can carry the target proteins out of the cells with a concentration from 101 to 691 mg/L in flask cultivation. The extracellular recombinant proteins with a relative high purity. The results suggested that this system has a potential application in plant biomass conversion and industrial production of enzymes and therapeutic proteins.

Keywords: *Escherichia coli*, Cellulase, Extracellular protein production, N-terminal sequence

Introduction

The Gram-negative bacterium *Escherichia coli* has been widely used as a cell factory for the production of enzymes and therapeutic proteins, because it is the best characterized host with many available expression and regulation tools [1-3]. However, the common laboratory strains of *E. coli* are poor secretors of proteins under normal culture conditions, because this bacterium has a complex cell envelope with two layers [4-6]. Therefore, heterologous proteins produced in recombinant *E. coli* are generally intracellular and often in the form of inclusion bodies, from which the biologically active proteins can only be recovered by complicated and costly processes

[7]. Extracellular production of heterologous proteins in *E. coli* will not only provide a simple and convenient production and purification process, but also provide fast and direct screening capabilities for target therapeutic proteins or enzymes that are heterologously expressed in recombinant *E. coli* [8].

Significant effort to produce target proteins extracellularly in *E. coli* has been made, with the research efforts split into two categories. (1) Targeted accumulation of the heterologous protein in the periplasmic space through the inner membrane (IM) using a leader peptide, such as PelB, then the heterologous protein is released to the medium through the outer membrane (OM) using cell envelope mutants or lysis proteins [9-11]; (2) Fusion of the heterologous proteins to fusion partners that can be secreted from the cytosol out of the cells via known or unknown systems.

* Correspondence: qiqingsheng@sdu.edu.cn
[†]Equal contributors
State Key Laboratory of Microbial Technology, Shandong University, Jinan 250100, People's Republic of China

Gao *et al. Microbial Cell Factories* (2015) 14:49

Page 2 of 8

A number of signal peptides have been used for secretory production of recombinant proteins in both eukaryote and prokaryote [12,13]. The typical signal sequence most located in the amino terminal of proteins that functions as a targeting and recognition signal and contains a cleavage site which can be cleaved by a special signal peptidase after transportation [14]. However, the recombinant proteins fused with signal sequence are transported to the periplasmic space instead of culture medium due to the double membrane structure of *E. coli* [15]. To release the target proteins to the culture medium, cell envelope mutants or lysis proteins were employed. But it also suffers from the purity of the secreted proteins and cell sensitivity to the environment due to the leakage of the cell envelope.

Regarding the application of the second approach, several proteins, including heterologous proteins, were found to be secreted directly into the medium from recombinant *E. coli*. However, only a few of proteins were investigated as fusion partners for extracellular recombinant protein production, because the secretion efficiency and universality were not sufficient to scale to a larger expression volume [16-20]. In 2005, Majander *et al.*, modified the flagellar type III secretion apparatus of *E. coli*. Then, by fusing a heterologous protein between the 173-bp untranslated region upstream of the gene *fliC* (encoding flagellin) and a transcriptional terminator from *fliC*, they found that the target proteins can be secreted into the medium via the type III secretion apparatus at levels of 1 to 15 mg/L [21]. Subsequently, researchers identified a 13 kDa endogenous bacterial peptide in *E. coli*, YebF, which can be secreted into the medium in relatively large amounts by a two-step process via the periplasmic space [22]. Recently, an osmotically inducible protein Y (OsmY) from *E. coli* BL21 (DE3) was identified using a proteomic method. OsmY can also be secreted out of the cells with target proteins at concentrations ranging between 5 and 64 mg/L [20]. Furthermore, ESETEC®, WACKER's patented secretion system, a technology for producing proteins and antibody fragments on a patented strain of *E. coli* K12, was the only one commercial solution to our knowledge [23].

Above all, it is generally considered that no fusion secretion systems are suitable for all heterologous proteins, because current fusion partner systems have been shown to have limitations [24-26]. In this study, we report that the catalytic domain of a cellulase (Cel-CD) from *Bacillus* sp. Z-16 can be efficiently secreted from *E. coli* when it was heterologously overexpressed. The accumulation of Cel-CD in the culture medium reached 514 mg/L. As the cellulase plays an important role in cellulosic biomass transformation, the extracellular expression of Cel-CD in *E. coli* provides a platform for cellulose production. Both the Cel-CD and its N-terminal sequence have potential as fusion partners in the production of various recombinant proteins.
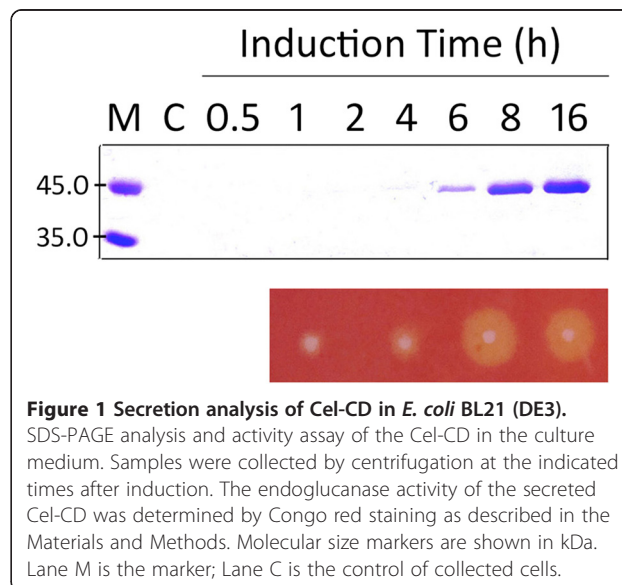
Results

Cel-CD is secreted from *E. coli* into the medium without typical signal peptides

A previously screened *Bacillus* sp. Z-16 was found to produce alkaline cellulase with endo-beta-glycosidase activity. This enzyme is composed of a 30-residue signal peptide at N-terminus, a catalytic domain (GH5, residues 30 to 404), followed by two carbohydrate-binding domains at C-terminal region (CBM_17_28, residues 405 to 821). The catalytic domain of this enzyme (Cel-CD) is 100% identical with a cellulase from *Bacillus subtilis* in the NCBI (DNA similarity is 99%, accession No.: M84963, 1,494 to 2,618). The full length Cel-CD consists of 375 amino acids with a molecular mass of 41.53 kDa.

We cloned the 1,125-bp coding region of the *cel-cd* gene from *Bacillus* sp. Z-16, without its upstream native signal sequence into the T7-driven expression plasmid pET28a, resulting in pET28a/cel. This recombinant plasmid was transformed into *E. coli* BL21 (DE3) cells for overexpression. During cultivation in the shake flask, we were surprised to find the Cel-CD in the culture medium not only in the cytosol. The Cel-CD progressively accumulated in the medium as the culture grew (Figure 1). After 24 h cultivation at 37°C, the final extracellular protein concentration reached 514 mg/L and the hydrolytic activity of Cel-CD towards carboxymethylcellulose (CMC) reached 558.4 U/L. Compare with the cells harboring pET28a, the overexpression of Cel-CD had no effect on the growth of *E. coli*.

Analysis of the N-terminal sequence of purified extracellular Cel-CD showed that it was MEGNTREDNF, which is the same as the predicted N-terminal sequence of the Cel-CD. The result proved that Cel-CD was

Figure 1 Secretion analysis of Cel-CD in *E. coli* BL21 (DE3). SDS-PAGE analysis and activity assay of the Cel-CD in the culture medium. Samples were collected by centrifugation at the indicated times after induction. The endoglucanase activity of the secreted Cel-CD was determined by Congo red staining as described in the Materials and Methods. Molecular size markers are shown in kDa. Lane M is the marker; Lane C is the control of collected cells.

Gao *et al. Microbial Cell Factories* (2015) 14:49

Page 3 of 8

translocated from *E. coli* BL21 (DE3) without removing any N-terminal sequence.

The secretion of Cel-CD is a two-step process via the periplasmic space

The lack of signal peptide in Cel-CD prompted us to determine its subcellular location. The *E. coli* BL21 (DE3) harboring Cel-CD was analysised using the cold osmotic shock method, as described in the Materials and Methods. We found that Cel-CD was in the cytoplasm, periplasmic space and medium, which proved that the secretion of Cel-CD was through the periplasmic space by western blot analysis with anti-Cel-CD antibodies (Figure 2A). In a subsequent experiment, we added chloramphenicol in the culture to inhibit nascent protein biosynthesis [27], and observed that the quantity of Cel-CD in the periplasmic space exhibited an initial increase and subsequently decreased in quantity, whereas it progressively accumulated in the medium (Figure 2B). These findings show that the secretion of Cel-CD is a two-step secretion process via the periplasmic space.
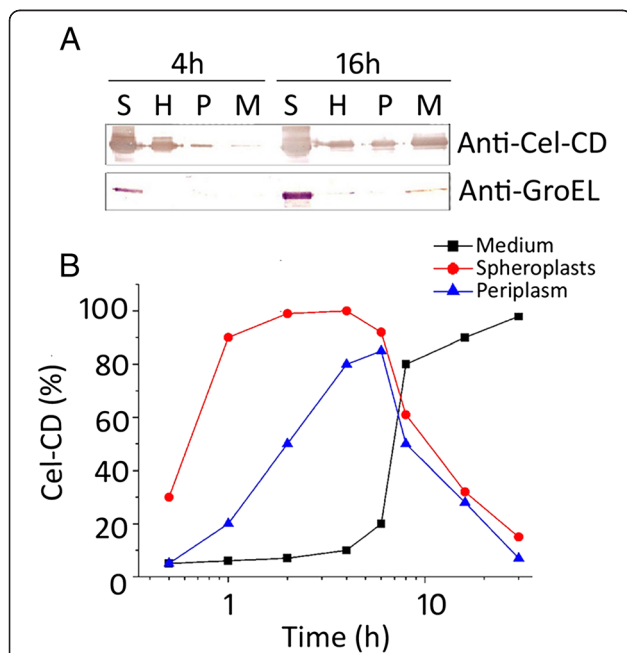

Figure 2 Subcellular localization and accumulation process analysis of Cel-CD. (A) Subcellular localization of Cel-CD. Cells were harvested 4 and 16 h after induction and washed with 1 mL of 100 mM MOPS (pH 7.0), then subjected to cold osmotic shock to separate the periplasm from spheroplasts. **(B)** The accumulation process of Cel-CD at different locations. Cells were harvested after 8 h induction and washed with 1 mL of 100 mM MOPS (pH 7.0), then suspended in 50 mL of fresh LB containing 80 μg/mL chloramphenicol. The quantification of Cel-CD was performed by Bio-Rad Quantity One Version 4.2.1. Lanes are marked as: S, spheroplasts; P, periplasm; H, hypertonic solution; M, medium.

Cell lysis determination

To exclude the possibility that extracellular Cel-CD accumulation was due to cell lysis or cell leakage through the outer membrane, we did western blot analysis against Cel-CD together with control proteins, the periplasmic maltose binding protein (MBP) [28] and the cytoplasmic protein GroEL [29]. The presence of control proteins in the medium would be an indication of cell lysis. As shown in Figure 3A, Cel-CD was detected in the culture medium at every time point, whereas MBP was not detected and GroEL was only detected in trace amounts after extended cultivation. This indicated that extracellular Cel-CD was not derived from cell lysis.

A recent report showed that the heterologously expressed cutinase from *Thermobifida fusca* was secreted from cells because of its hydrolytic activity toward phospholipids [30]. To exclude the possibility that Cel-CD secretion was also caused by its hydrolytic activity in the periplasmic space, the active site of Cel-CD was mutated according to the results of partial amino acid sequence alignment with cellulase. Two mutants, E160Q, and E160Q and E274Q, in which hydrolytic activity was completely lost, were still detected in the medium (Figure 3B). The morphology of *E. coli* BL21 (DE3) strains were analysised by Transmission electron microscopy (TEM). It showed that the cells expressing either Cel-CD or inactivated Cel-CD remained intact, supporting the aforementioned result (Figure 3C).

The N-terminus of Cel-CD plays a key role in secretion

As the secretion of Cel-CD in *E. coli* BL21 (DE3) was a two-step pocess without predicted signal peptide, we further detected if the N-terminus of mature Cel-CD functioned as a signal peptide. We constructed five protein mutants (Δ5Cel-CD, Δ10Cel-CD, Δ15Cel-CD, Δ20Cel-CD and Δ30Cel-CD) by gradual N-terminal truncation. The expression of the five mutants showed almost no difference, but the secretion level gradually decreased as the size of the truncation increased compared to the wild-type (Figure 4). This data demonstrate that the importance of the N-terminal sequence on Cel-CD secretion in *E. coli*. Moreover, the Δ20Cel-CD and Δ30Cel-CD mutants, which lack the N-terminal 20 or 30 residues, were completely inhibited the secretion. Taken together, these results indicated that the N-terminal 20 residues of Cel-CD may be the mini length for the export pocess.

Extracellular production of heterologous proteins in *E. coli*

The efficient and quantitative secretion of Cel-CD inspired us to test if it could be used as a fusion partner or signal peptide/protein to carry other proteins out of the cell in recombinant *E. coli*. We selected a series of proteins with different sizes and sources, including pectate lyase C (PelC, 24.3 kDa) from *B. subtilis*, human neuritin
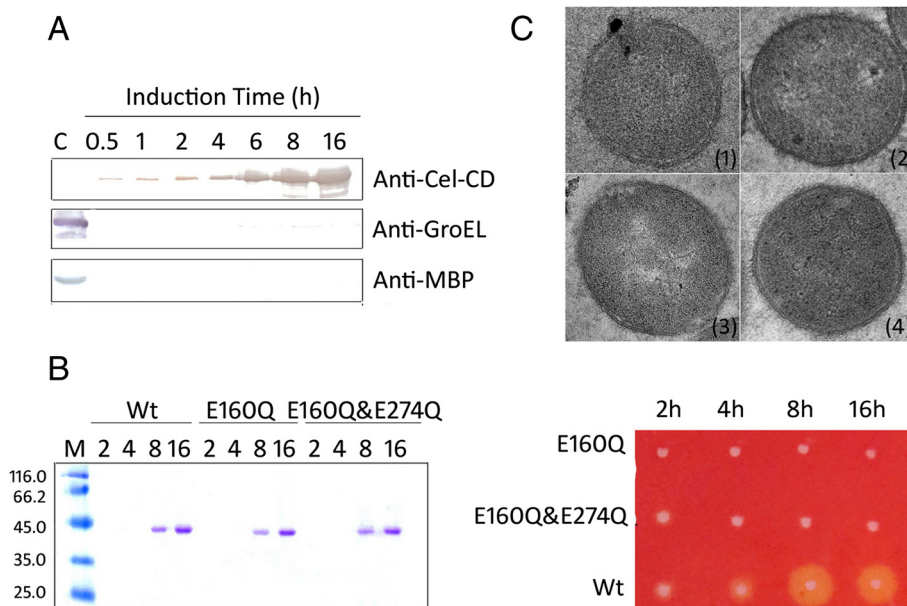
Gao *et al. Microbial Cell Factories* (2015) 14:49

Page 4 of 8

Figure 3 Cell lysis determination and inactive Cel-CD secretion analysis. (A) Immunoblotting detection of Cel-CD, GroEL and MBP in the culture medium. Samples were subjected to immunoblotting with an anti-Cel-CD antibody, an anti-GroEL antibody or an anti-MBP antibody. **(B)** SDS analysis of cellulase (wt) and its mutants (E160Q, E160Q&E274Q) secreted into the culture medium. The activities of the secreted mutants of Cel-CD (E160Q, E160Q&E274Q) were analyzed by Congo red staining method (4 h, 6 h, 8 h, 16 h). **(C)** TEM images of *E. coli* BL21 (DE3) that containing different plasmids. (1) pET28a/cel, (2) pET28a/E160Q, (3) pET28a/E160Q&E274Q, (4) pET28a. Protein samples were separated on 12% SDS-PAGE. Aliquots corresponding to 20 μL of the culture medium were loaded onto the gel. Molecular size markers are shown in kDa. Lane M is the marker; Lane C is the control of collected cells.
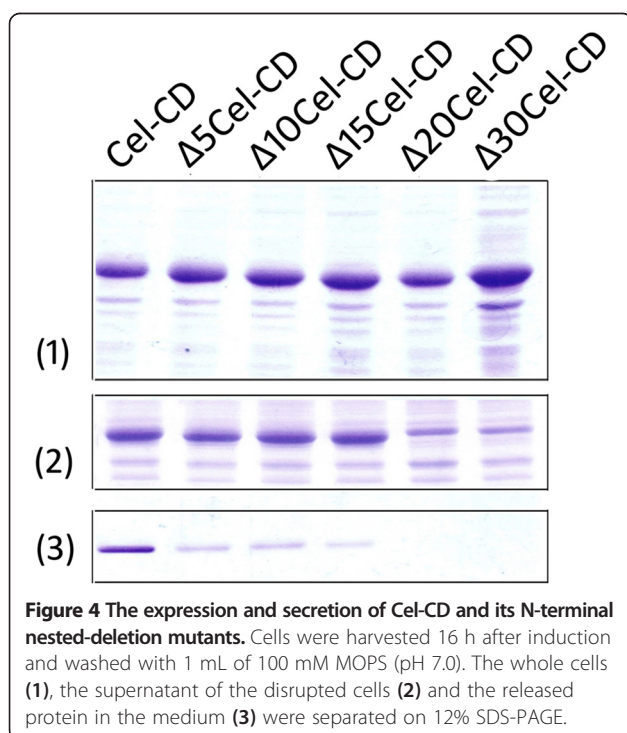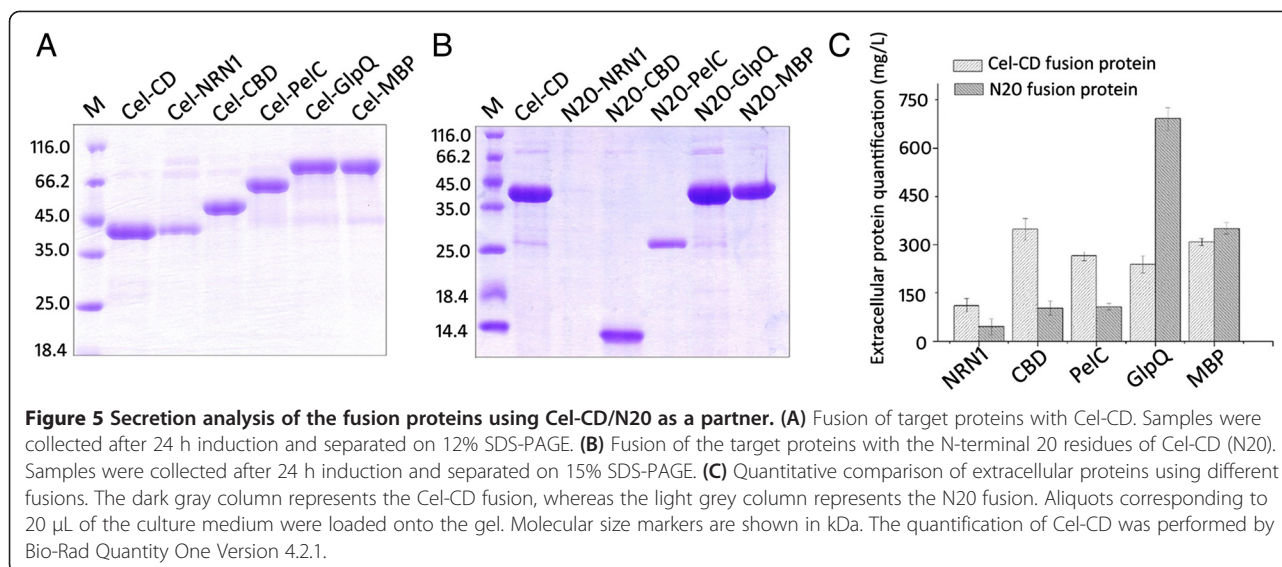

Figure 4 The expression and secretion of Cel-CD and its N-terminal nested-deletion mutants. Cells were harvested 16 h after induction and washed with 1 mL of 100 mM MOPS (pH 7.0). The whole cells **(1)**, the supernatant of the disrupted cells **(2)** and the released protein in the medium **(3)** were separated on 12% SDS-PAGE.

(NRN1, 15.3 kDa), the carbohydrate binding domain of cyclodextrin glycosyltransferase (CBD, 11.21 kDa) from *B. circulans*, maltose-binding protein (MBP, 43.4 kDa) and glycerophosphoryl diester phosphodiesterase (GlpQ, 40.8 kDa) from *E. coli.* These five proteins can only be detected in the cytoplasm when they were expressed without any signal sequences. When fused downstream of Cel-CD, all these proteins were detected in the medium after 24 h cultivation; however, the secretion levels varied (Figure 5A). The secretion of Cel-CBD was the highest, reaching 348 mg/L. The secretion of Cel-glpQ, Cel-MBP, Cel-pelC and Cel-NRN1 in the culture medium was also substantial with values of 266.8, 307.6, 264.6 and 211.3 mg/L, respectively.

Since the N-terminal region plays an important role in secretion, we tried the N-terminal 20 residues (N20) as a signal peptide to guide the target proteins out of *E. coli* (Figure 5B). The N-terminal 20 residues from Cel-CD were added to the target proteins by direct PCR ligation. Fusion of N20 with GlpQ and MBP led to a significant increase in secretion when compared with the Cel-CD fusion (Figure 5C). The extracellular protein was 691 and 350 mg/L, respectively, while the secretion of N20-CBD and N20-pelC reduced to 101 and 106.7 mg/L, respectively. However, N20-NRN1 was not secreted. Analysis of

Gao *et al. Microbial Cell Factories* (2015) 14:49

Page 5 of 8

Figure 5 Secretion analysis of the fusion proteins using Cel-CD/N20 as a partner. (A) Fusion of target proteins with Cel-CD. Samples were collected after 24 h induction and separated on 12% SDS-PAGE. **(B)** Fusion of the target proteins with the N-terminal 20 residues of Cel-CD (N20). Samples were collected after 24 h induction and separated on 15% SDS-PAGE. **(C)** Quantitative comparison of extracellular proteins using different fusions. The dark gray column represents the Cel-CD fusion, whereas the light grey column represents the N20 fusion. Aliquots corresponding to 20 μL of the culture medium were loaded onto the gel. Molecular size markers are shown in kDa. The quantification of Cel-CD was performed by Bio-Rad Quantity One Version 4.2.1.

the expression of N20-NRN1 showed that it was deposited as inclusion bodies (data not shown).

Discussion

In this study, we have shown that Cel-CD, the catalytic domain of a cellulase from *Bacillus* sp.Z-16, could be secreted into the culture medium without a typical signal sequence when expressed in *E. coli* BL21 (DE3). As a heterologous protein, the overexpression of Cel-CD was not found any interference to cell growth and metabolism. This may overcome the cell lysis caused by the overexpression of the fusion partner, thus extend the fermentation process to increase the extracellular production. Meanwhile, the secretion of Cel-CD and its recombinant proteins was a two-step pocess, the oxidizing environment and Dsb system of periplasm benefits the formation of disulfide bonds [31]. Further experiments showed that Cel-CD can also be expressed and secreted from *E. coli* K strains, such as *E. coli* BW25113, MG1655, though the level of secretion was much lower. Thus, the application of Cel-CD fusion partner is not strain dependent.

We have proven that the N-terminal sequence of full-length Cel-CD plays a key role in secretion. This indicated that N-terminal 20 amino acid residues of Cel-CD can serve as a carrier for the secretion of heterologous target proteins out of *E. coli*. Fused with Cel-CD or N20, five selected proteins found to secrete into the medium at high levels. The secretion efficiency depends on the properties of the target proteins. Cel-CD appears to be suitable for the secretion of small and low-solubility proteins because of its high solubility and large size in *E. coli*, whereas the N20 peptide is more suitable for large proteins. Notably, all secreted fusion proteins exhibited relative high solubility; fusion proteins with low solubility are not efficiently

secreted. We suggest that the solubility of the fusion protein is a prerequisite for secretion, while the N-terminal sequence and three dimensional structure of the protein is a crucial factor. Fusion of heterologous proteins with N20 from Cel-CD can carry the target proteins out of the cell. This is the first report to show that a short peptide can serve as a signal peptide and guide heterologous proteins across both the inner and outer membranes of *E. coli*.

Secretory production of proteins in *E. coli* has many possible applications. For example, a mutagenized protein library could be directly screened on a plate if the protein or its product is directly visible, which can greatly facilitate protein engineering research. Extracellular accumulation of Cel-CD with other cellulosic enzymes may also be useful for the creation of consolidated bioprocess for simultaneous biomass conversion product formation [32]. Finally, secretory production of proteins will simplify the protein production process and will decrease production costs of enzymes and antibodies.

Conclusions

Recombinant proteins produced in *Escherichia coli* are generally intracellular and often found in the form of inclusion bodies. In this study, a heterologous expressed catalytic domain of a cellulase (Cel-CD) in *E. coli* was found to be secreted out of the cells in large quantities with endo-beta-glycosidase activity. We demonstrated that the N-terminal sequence of Cel-CD played a crucial role in this secretion, which suggests that both Cel-CD and its N-terminus can be employed as carriers for extracellular production of recombinant proteins. Recombinant *E. coli* expressing the fusion protein of Cel-CD and other types of hydrolytic enzymes can be used for creation of consolidated bioprocess.

Gao *et al. Microbial Cell Factories* (2015) 14:49

Page 6 of 8

Materials and methods

Bacterial strains and plasmids construction

Strains and plasmids used in this study were shown in Table 1. *E. coli* strain DH5α (Invitrogen) was used for recombinant DNA manipulation. *E. coli* BL21 (DE3) (Novagen) was used as host for the expression of Cel-CD and other recombinant proteins. Cells were cultured in 5 mL Luria Broth (LB) medium for 12 h at 37°C. This seed culture was inoculated to 50 mL LB medium in a 300 mL flask. To induce the expression of the desired proteins, Isopropyl-β-D-thiogalactopyranoside (IPTG) was added to

a final concentration of 0.5 mM when OD_{600} reached 0.8. After induction, cells were culture until the indicated time. Antibiotics were added as follows: kanamycin (Kan) 50 μg/mL.

Cel-CD expression plasmids was created based on pET28a (+) (Novagen) containing the T7 promoter. The DNA sequence encoding the catalytic domain of cellulase (Cel-CD) used in this study was amplified from the genome of *Bacillus sp.* Z-16 (lab collection) by polymerase chain reaction using PrimeSTAR (TaKaRa) with the Cel-CD-*Nco*I-forward and Cel-CD-*Xho*I-Reverse as primers (Table 2). After digestion with *Nco*I and *Xho*I (Fermentas), the amplified DNA fragment was inserted into pET-28a using T4 DNA ligase (NEB), generating plasmid pCel. The DNA sequence encoding N20 amino acid residues of Cel-CD used in this study was amplified from the plasmids pCel by polymerase chain reaction using PrimeSTAR (TaKaRa) with the Cel-CD-*Nco*I-forward and Cel-20-*Bam*HI-Reverse as primers.

The DNA sequence encoding Cel-CD with 5 amino acid residues deletion at the N-terminus was amplified from the plasmids pCel by polymerase chain reaction using PrimeSTAR (TaKaRa) with the Cel-CD5-*Nco*I-Forward and Cel-CD-*Xho*I-Reverse as primers, generating plasmids

Table 1 The strains and plasmids used in this study

Description		Source
Strains		
E. coli DH5α	F⁻ *endA1 hsdR17*(r$_K$⁻ m$_K$+) *supE44 thi-l* λ⁻ *recA1 gyrA96 ΔlacU169* (Φ80dlacZM15)	Invitrogen
E. coli BL21 (DE3)	*hsdS, gal*(Δclts857 ind1, Sam7, nin5, lacUV5-T7, gene1	Novagen
Plasmids		
pET28a	5.4 kb, f1 ori, T7 promoter, Kan^R	Novagen
pCel	pET28a, *cel-cd* gene under T7 promoter	This study
pN20	pET28a, 60 bp at 5' region of *cel-cd* under T7 promoter	This study
pΔ5cel	pET28a, *cel-cd* gene that truncated 15 bp at 5' region	This study
pΔ10cel	pET28a, *cel-cd* gene that truncated 30 bp at 5' region	This study
pΔ15cel	pET28a, *cel-cd* gene that truncated 45 bp at 5' region	This study
pΔ20cel	pET28a, *cel-cd* gene that truncated 60 bp at 5' region	This study
pΔ30cel	pET28a, *cel-cd* gene that truncated 90 bp at 5' region	This study
pCelNRN1	pET28a, *cel-cd* gene fusion with *nrn1* from *Homo sapiens*	This study
pCelMBP	pET28a, *cel-cd* gene fusion with *male* from *Escherichia coli*	This study
pCelCBD	pET28a, *cel-cd* gene fusion with *cbd* from *Bacillus circulans*	This study
pCelGlpQ	pET28a, *cel-cd* gene fusion with *glpQ* from *Escherichia coli*	This study
pCelPelC	pET28a, *cel-cd* gene fusion with *pelc* from *Bacillus subtilis*	This study
p20NRN1	pET28a, *n20* gene fusion with *nrn1*	This study
p20MBP	pET28a, *n20* gene fusion with *male*	This study
p20CBD	pET28a, *n20* gene fusion with *cbd*	This study
p20GlpQ	pET28a, *n20* gene fusion with *glpq*	This study
p20PelC	pET28a, *n20* gene fusion with *pelc*	This study
E170Q	pET28a, the mutated *cel-cd* gene under T7 promoter	This study
E170Q&E275Q	pET28a, the mutated *cel-cd* gene under T7 promoter	This study

Table 2 The primers used in this study

Primers	Sequence
Cel-CD-*Nco*I-Forward	5'-TTTT*CCATGG*AAGGAAACACTCGTGAAGA-3'
Cel-CD-*Xho*I-Reverse	5'-TTTT*CTCGAG*AAGTACTTTCGTGTATTTTG-3'
Cel-CD-*Bam*HI-Reverse	5'-TTTT*GGATCC*AAGTACTTTCGTGTATTTTG-3'
Cel-20-*Bam*HI-Reverse	5'-TTTT*GGATCC*GCGTTTAACATTGTCATTAC-3'
Cel-CD5-*Nco*I-Forward	5'-TTTT*CCATGG*AAGACAATTTTAAACATTT-3'
Cel-CD10-*Nco*I-Forward	5'-TTTT*CCATGG*AACATTTATTAGGTAATGACAA-3'
Cel-CD15-*Nco*I-Forward	5'-TTTT*CCATGG*ACAATGTTAAACGCCCTTC-3'
Cel-CD20-*Nco*I-Forward	5'-TTTT*CCATGG*AACCTTCTGAGGCTGGCGCATT-3'
Cel-CD30-*Nco*I-Forward	5'-TTTT*CCATGG*AAGTCGATGGACAAATGACA-3'
NRN1-*Bam*HI-Forward	5'-TTTT*GGATCC*GCTGGTAAATGTGACGCCG-3'
NRN1-*Xho*I-Reverse	5'-TTTT*CTCGAG*TTAACCATTGCCCGAGCCG-3'
MBP-*Bam*HI-Forward	5'-TTTT*GGATCC*ATCGAAGAAGGTAAACTGG-3'
MBP-*Xho*I-Reverse	5'-TTTT*CTCGAG*TTACTTGGTGATACGAGTC-3'
GlpQ-*Bam*HI-Forward	5'-TTTT*GGATCC*GCGGACAGCAACGAAAAAA-3'
GlpQ-*Xho*I-Reverse	5'-TTTT*CTCGAG*TTACTCTTTATTAAGAAATT-3'
CBD-*Bam*HI-Forward	5'-TTTT*GGATCC*GACCAGGTCAGCGTCCGCTT-3'
CBD-*Xho*I- Reverse	5'-TTTT*CTCGAG*TTATGGCTGCCAATTCACGT-3'
pelC-*Bam*HI-Forward	5'-TTTT*GGATCC*GCAGATAAGGTTGTCCACGA-3'
pelC-*Xho*I- Reverse	5'-TTTT*CTCGAG*TTAGAACTGGGTGTTATTAT-3'
Glu160Gln-Forward	5'-TTATTTATGAGTTAGCGAAT*CAG*CCAAGTAGTA-3'
Glu160Gln-Reverse	5'- GATTCGCTAACTCATAAATAATGTGTGGATTGT-3'
Glu274Gln-Forward	5'-GAGTAGCGGTATTTGCGACAC*AA*TGGGGAACGA-3'
Glu274Gln-Reverse	5'-GTGTCGCAAATACCGCTACTCCGTTTTCTAACG-3'

Gao *et al. Microbial Cell Factories* (2015) 14:49

Page 7 of 8

p△5cel. The same method was used to construct p△10cel, p△15cel, p△20cel and p△30cel.

To fuse the target protein genes to the 3′ end of *cel-cd*, the DNA fragments were amplified by PCR, and were digested with restriction enzyme *Bam*HI and *Xho*I on the both ends. The resulting fragments were then ligated to the vector pET-28a/*cel* by T4 DNA ligase to obtain the objective plasmids. The same method was used to construct N20 fusion proteins.

Cel-CD activity assay

The hydrolytic activity of the secreted Cel-CD was determined by Dinitrosalicylic Acid (DNS) method. 0.1 mL culture medium was added to 2 mL Britton-Robinson buffer containing 1% CMC-Na in a test tube. Reaction mixture was incubated at 37°C for 30 min, then was added with 2 mL DNS reagent and boiled for 5 min before it was diluted to 12.5 mL. The enzyme activity was measured with a spectrophotometer at 540 nm.

Protein quantification

The total secretory proteins were quantified by Bradford method using Easy Protein Quantitative kit (Lot #G10322, TransGen Biotech). The proteins strained with Coomassie were detected by the UV spectrophotometer at 595 nm. The protein samples were loaded on 12% or 15% SDS-PAGE and stained with Coomassie dye. The stained gels were scanned by Scanner and the target protein bands were normalized and quantified by Bio-Rad Quantity One Version 4.0.

SDS-PAGE and western blotting

Cells were harvested at indicated time after induction and washed once with 1 mL of 100 mM MOPS (pH 7.0). The secreted protein samples were obtained directly by centrifugation. The cell pellets were then disrupted by ultrasonification and centrifuged to obtain the total soluble protein samples. The samples were then prepared by adding an equal volume of electrophoresis loading buffer. Proteins were separated on 12% SDS-PAGE and either stained with Coomassie dye or transferred to a PVDF membrane (100 V, 1 h) for immunoblotting, blocked with 5% (W/V) BSA in PBS/Tween, and probed using the appropriate antibody. Antibody to Cel-CD (Sigma, 1:2,000) was used for the detection of secreted Cel-CD in the culture medium and cells. Antibody to MalE (Sigma, 1:5,000) and GroEL (Sigma, 1:10,000) were used for the control of potential lysis or leakage from the periplasmic space (MalE) and cytoplasmic (GroEL).

Site-directed mutagenesis of Cel-CD

Site-directed mutagenesis mutants were constructed using Easy Mutagenesis System Kit (TransGene). The primers Glu160Gln-Forward/Reverse were included 5′end overlap region, mutation site and 3′ end extension region. In a certain system, the target DNA vectors containing the *cel-cd* gene were amplified by PCR and the templates were digested by DMT enzyme. After purified by gel extraction, the target vectors containing the E160Q mutant and E160Q&E274Q mutant was obtained. The Gln in the active sites of Cel-CD (Gln160 and Gln274) were all mutated to Glu. All the resulting constructs were confirmed by DNA sequencing and hydrolytic activity determination.

Transmission electron microscopy (TEM)

E. coli cells containing the target plasmids were collected by centrifuging at 4,000 g for 10 min. Pelleted cells were fixed by 2.5% glutaraldehyde and 1% osmic acid. After gradually dehydrated by ethyl alcohol, the cells were embedded in Epon-812 araldite. The samples were cut into 70 nm thick slice, stained using 2% uranyl acetate and lead citrate, and observed by the transmission electron microscope at 200 kV. The resulting microscopic images were photographed.

Analysis of the N-terminal sequence

The protein samples were purified and separated by 2D-SDS PAGE, and were transferred to polyvinylidene difluoride (PVDF) blotting membrane. The N-terminal sequence of the protein was analyzed by EDMAN degradation method [33] using the PROCISE491 protein sequencer by the Center for Protein Science, Peking University.

Congo red staining

To determine the hydrolytic activity of Cel-CD, an agarose plate-based assay was used. 1% (wt/vol) Carboxymethylcellulose (CMC-Na) was added in the agarose plate as a substrate (0.8%, wt/vol). Samples containing the Cel-CD were loaded into the hole of plate. The plate was incubated at 30°C for 3 h, then flooded with 0.1% (wt/vol) Congo red solution and stained for 15 min at room temperature. The dye was removed, and 5 mL of water was used to wash the plate. Finally, 5 mL 0.9% NaCl solution was applied for 15 min, and the plates were then dried and photographed.

Subcellular fractionation

Periplasmic protein was separated from cytoplasm by the osmotic shock method [34]. The cells were harvested by centrifugation (6,000 g for 5 min) and were suspended in 100 mM Tris–HCl containing 20% sucrose and 1 mM EDTA (pH 8.0), then pelleted by centrifugation (8,000 g for 5 min) followed by re-suspension in ice-cold water for 10 min. After the addition of $MgCl_2$ to a final concentration of 1 mM, the cell suspension was incubated on ice

Gao *et al. Microbial Cell Factories* (2015) 14:49

Page 8 of 8

for a further 10 min before being pelleted by centrifugation (8,000 g for 5 min).

References
1. Jong WS, Sauri A, Luirink J. Extracellular production of recombinant proteins using bacterial autotransporters. Curr Opin Biotechnol. 2010;21:646–52.
2. Ferrer-Miralles N, Domingo-Espin J, Corchero JL, Vazquez E, Villaverde A. Microbial factories for recombinant pharmaceuticals. Microb Cell Fact. 2009;8:17.
3. Langer ES. Capitalizing on novel expression systems. Genet Eng Biotechnol News. 2009;29:12–2.
4. Desvaux M, Hebraud M, Talon R, Henderson IR. Secretion and subcellular localizations of bacterial proteins: a semantic awareness issue. Trends Microbiol. 2009;17:139–45.
5. Wai SN, Lindmark B, Soderblom T, Takade A, Westermark M, Oscarsson J, et al. Vesicle-mediated export and assembly of pore-forming oligomers of the enterobacterial ClyA cytotoxin. Cell. 2003;115:25–35.
6. Pugsley AP, Francetic O. Protein secretion in *Escherichia coli* K-12: dead or alive? Cell Mol Life Sci. 1998;54:347–52.
7. Baneyx F, Mujacic M. Recombinant protein folding and misfolding in *Escherichia coli*. Nat Biotechnol. 2004;22:1399–408.
8. Choi JH, Lee SY. Secretory and extracellular production of recombinant proteins using *Escherichia coli*. Appl Microbiol Biotechnol. 2004;64:625–35.
9. Shin HD, Chen RR. Extracellular recombinant protein production from an *Escherichia coli* lpp deletion mutant. Biotechnol Bioeng. 2008;101:1288–96.
10. Robbens J, Raeymaekers A, Steidler L, Fiers W, Remaut E. Production of soluble and active recombinant murine interleukin-2 in *Escherichia coli*: high level expression, Kil-induced release, and purification. Protein Expr Purif. 1995;6:481–6.
11. Sevastsyanovich YR, Leyton DL, Wells TJ, Wardius CA, Tveen-Jensen K, Morris FC, et al. A generalised module for the selective extracellular accumulation of recombinant proteins. Microb Cell Fact. 2012;11:69.
12. von Heijne G. Signal sequences. The limits of variation. J Mol Biol. 1985;184:99–105.
13. Austen BM, Ridd DH. The signal peptide and its role in membrane penetration. Biochem Soc Symp. 1981;46:235–58.
14. von Heijne G. The signal peptide. J Membr Biol. 1990;115:195–201.
15. Mergulhao FJ, Summers DK, Monteiro GA. Recombinant protein secretion in *Escherichia coli*. Biotechnol Adv. 2005;23:177–202.
16. Gumpert J, Hoischen C. Use of cell wall-less bacteria (L-forms) for efficient expression and secretion of heterologous gene products. Curr Opin Biotechnol. 1998;9:506–9.
17. Jeong KJ, Lee SY. Excretion of human beta-endorphin into culture medium by using outer membrane protein F as a fusion partner in recombinant *Escherichia coli*. Appl Environ Microbiol. 2002;68:4979–85.
18. Teng C, Jia H, Yan Q, Zhou P, Jiang Z. High-level expression of extracellular secretion of a beta-xylosidase gene from Paecilomyces thermophila in *Escherichia coli*. Bioresour Technol. 2010;102:1822–30.
19. Nandakumar MP, Cheung A, Marten MR. Proteomic analysis of extracellular proteins from *Escherichia coli* W3110. J Proteome Res. 2006;5:1155–61.
20. Qian ZG, Xia XX, Choi JH, Lee SY. Proteome-based identification of fusion partner for high-level extracellular production of recombinant proteins in *Escherichia coli*. Biotechnol Bioeng. 2008;101:587–601.
21. Majander K, Anton L, Antikainen J, Lang H, Brummer M, Korhonen TK, et al. Extracellular secretion of polypeptides using a modified *Escherichia coli* flagellar secretion apparatus. Nat Biotechnol. 2005;23:475–81.
22. Zhang G, Brokx S, Weiner JH. Extracellular accumulation of recombinant proteins fused to the carrier protein YebF in *Escherichia coli*. Nat Biotechnol. 2006;24:100–4.
23. Muecke M, Leonhartsberger S. Increasing Production Process Efficiency. Genet Eng Biotechnol News. 2009;29:56–7.
24. Xia XX, Han MJ, Lee SY, Yoo JS. Comparison of the extracellular proteomes of *Escherichia coli* B and K-12 strains during high cell density cultivation. Proteomics. 2008;8:2089–103.
25. Zhou S, Yomano LP, Saleh AZ, Davis FC, Aldrich HC, Ingram LO. Enhancement of expression and apparent secretion of Erwinia chrysanthemi endoglucanase (encoded by celZ) in *Escherichia coli* B. Appl Environ Microbiol. 1999;65:2439–45.
26. Papi RM, Chaitidou SA, Trikka FA, Kyriakidis DA. Encapsulated *Escherichia coli* in alginate beads capable of secreting a heterologous pectin lyase. Microb Cell Fact. 2005;4:35.
27. Weber MJ, DeMoss JA. The inhibition by chloramphenicol of nascent protein formation in *E. coli*. Proc Natl Acad Sci U S A. 1966;55:1224–30.
28. Shuman HA, Panagiotidis CH. Tinkering with transporters: periplasmic binding protein-dependent maltose transport in *E. coli*. J Bioenerg Biomembr. 1993;25:613–20.
29. Horwich AL, Low KB, Fenton WA, Hirshfield IN, Furtak K. Folding in vivo of bacterial cytoplasmic proteins: role of GroEL. Cell. 1993;74:909–17.
30. Su L, Woodard RW, Chen J, Wu J. Extracellular location of Thermobifida fusca cutinase expressed in *Escherichia coli* BL21(DE3) without mediation of a signal peptide. Appl Environ Microbiol. 2013;79:4192–8.
31. de Marco A. Recent contributions in the field of the recombinant expression of disulfide bonded proteins in bacteria. Microb Cell Fact. 2012;11:129.
32. Bokinsky G, Peralta-Yahya PP, George A, Holmes BM, Steen EJ, Dietrich J, et al. Synthesis of three advanced biofuels from ionic liquid-pretreated switchgrass using engineered *Escherichia coli*. Proc Natl Acad Sci U S A. 2011;108:19949–54.
33. Edman P. A method for the determination of amino acid sequence in peptides. Arch Biochem. 1949;22:475.
34. Nossal NG, Heppel LA. The release of enzymes by osmotic shock from *Escherichia coli* in exponential phase. J Biol Chem. 1966;241:3055–62.