Protein Expression and Purification 104 (2014) 14–19

Contents lists available at ScienceDirect

Protein Expression and Purification

journal homepage: www.elsevier.com/locate/yprep

Expression and purification of soluble human cystatin C in *Escherichia coli* with maltose-binding protein as a soluble partner

Qing Zhang [a,1], Xiaozhi Zhao [a,1], Xiaoyu Xu [c], Bo Tang [c], Zhenglei Zha [a], Mingxin Zhang [a], Dongwei Yao [a], Xiaoxiang Chen [a], Xuhong Wu [a], Lin Cao [b,c,*], Hongqian Guo [a,*]

[a] Department of Urology, Drum Tower Hospital, Medical School of Nanjing University, 321 Zhongshan Road, Nanjing 210008, Jiangsu, PR China
[b] College of Food Science and Technology, Nanjing Agricultural University, Nanjing 210095, Jiangsu, PR China
[c] Vazyme Biotech Co., Ltd, Nanjing 210000, Jiangsu, PR China

ARTICLE INFO

Article history:
Received 16 July 2014
and in revised form 14 September 2014
Available online 26 September 2014

Keywords:
Cystatin C
Fusion expression
Soluble expression

ABSTRACT

Human cystatin C (CYSC) is a 13-kDa endogenous cysteine proteinase inhibitor and was investigated as a replacement for creatinine as a marker of renal function. However, expressing recombinant CYSC is difficult in *Escherichia coli* because of resulting low yield and insufficient purity and insolubility. Here, we cloned and fused *CYSC* to the C-terminus of three soluble partners – maltose-binding protein (MBP), glutathione S-transferase (GST) and translation initiation factor 2 domain I (IF2) – to screen for their ability to improve the solubility of recombinant CYSC when expressed in *E. coli*. MBP was best at enhancing the soluble expression of CYSC, with soluble fractions accounting for 92.8 ± 3.11% of all proteins. For scaled production, we purified the de-tagged CYSC by using a 3C protease-cleaved *MBP-T3-CYSC* fused protein with immobilized metal affinity chromatography and cation-affinity purification. The molecular weights of the de-tagged CYSC and human natural CYSC were similar, and the former could react specifically with CYSC polyclonal antibody. Moreover, the de-tagged CYSC displayed full biological activity against papain and cathepsin B, which was very similar to that of the human natural CYSC protein standard. We provide a method to produce large amounts of soluble recombinant human CYSC in *E. coli*.

Introduction

Cystatin C (CYSC),[2] a 13-kDa endogenous cysteine proteinase inhibitor, is a member of the family of proteins with an important role in intracellular catabolism of various peptides and proteins [1]. It was first investigated as a marker for glomerular filtration rate (GFR) in 1985 [2]. In the last decade, research on the use of CYSC to replace creatinine as a marker of renal function has been rapidly growing. Furthermore, CYSC is a predictor of heart failure, and increased levels of CYSC are independently associated with increased mortality in both chronic and acute heart failure [3]. Several other studies have shown CYSC as a strong risk factor of prognosis in venous thromboembolism [4] and a potential target for Alzheimer's treatment in various populations [5].

Therefore, recombinant production of CYSC is attractive for obtaining large amounts of protein for biological application. Recombinant production systems offer higher protein production, better growth, and consequently, improved productivity as compared with non-recombinant production systems [6]. *Escherichia coli* is commonly used as a host for rapid and economical production of recombinant proteins [7]. Several attempts have aimed to produce recombinant human CYSC in *E. coli*; however, the production was low, of mostly insoluble inclusion bodies [8].

Many attempts have been made to improve the soluble expression of recombinant proteins in *E. coli*. Fusion expression seems to be the most effective way to increase a protein's solubility. Compared to non-fusion proteins, fusion or chimeric proteins with "tags" linked to target proteins can improve the solubility of target proteins [9,10], protect against intracellular proteolysis [11], and sometimes function as specific expression reporters [12]. Furthermore, especially N-terminal fusion tags can provide reliable contexts for efficient translation initiation [13].

Some fusion partners worked effectively in certain situations but failed in others. In this work, we fused CYSC to three fusion tags, including maltose-binding protein (MBP), glutathione S-transferase (GST) and translation initiation factor 2 domain I

* Corresponding authors at: Department of Urology, Drum Tower Hospital, Medical School of Nanjing University, 321 Zhongshan Road, Nanjing 210008, Jiangsu, PR China (H. Guo). Fax: +86 25 83105987.
E-mail addresses: caolin@njau.edu.cn (L. Cao), dr.guohongqian@gmail.com (H. Guo).

[1] These authors contributed equally to this work.

[2] *Abbreviations used:* CYSC, cystatin C; MBP, maltose-binding protein; GST, glutathione S-transferase; IF2, initiation factor 2; IMAC, immobilized metal affinity chromatography; *E. coli*, *Escherichia coli*.

http://dx.doi.org/10.1016/j.pep.2014.09.010
1046-5928/

(IF2), to screen for their ability to enhance the soluble expression of CYSC in *E. coli* for large-scale production.

Materials and methods

Materials and reagents

E. coli BL21 (DE3) (Novagen, Billerica, MA, USA) was maintained in our laboratory. The expression vector pET-28a was from Novagen. Restriction enzymes NheI and BamHI were from New England BioLabs (Lpswich, MA, USA). The Ni-Sepharose FF resin was from GE (Piscataway, NJ, USA). The 3C protease with an N-terminus His tag was from Vazyme (Piscataway, NJ, USA). Human cathepsin B was from Calbiochem (La Jolla, CA, USA). Papain was from Sigma (Louis, MO, USA). Peptidyl proteinase substrates were from Bachem Feinchemikalien (Bubendorf, Switzerland). Other chemicals were of analytical or higher grade.

Construction of CYSC fusion vectors

Two 57-mer oligonucleotides (Linker sense and Linker antisense Table 1, from Genscript Biotechnology, Piscataway, NJ, USA) were annealed to produce a fragment containing an NheI sticky end at the 5′ end, a BamHI sticky end at the 3′ end and an SpeI restriction site adjacent to the NheI site to introduce a "T3" sequence that encodes the sequential TEV protease and 3C protease recognition sites between the fusion tags and target protein. The annealing product was then ligated to the pET-28a vector digested with NheI and BamHI. The resulting plasmid was named T3-28a (Fig. 1).

The three fusion partners were amplified by PCR (primers and restriction sites are in Table 2). The sources for templates for MBP and IF2 were *E. coli* BL21 [12,14] and for GST, *Schistosoma japonicum* [11] (Table 2). The PCR products were cloned into T3-28a between the NdeI and SpeI sites by use of the ClonExpress one-step cloning kit (Vazyme, Piscataway, NJ, USA) to obtain the pFusion (pMBP, pIF2 and pGST) vectors.

The complete coding sequence of mature CYSC (Ser27-Ala146) was synthesized by Genscript Biotechnology and cloned into the pUC57 plasmid. The CYSC open reading frame was PCR-amplified and cloned into pFusion vectors and the T3-28a plasmid between BamHI and Hind III by use of ClonExpress (Vazyme, Piscataway, NJ, USA). The resulting vectors were named *MBP-T3-CYSC*, *GST-T3-CYSC*, *IF2-T3-CYSC* and *T3-CYSC* (Fig. 1A).

Overexpression in small-scale and SDS–PAGE analysis

Expression plasmids were transformed into *E. coli* and grown overnight on LB plates containing 10 μg/ml kanamycin. Single-colony cultures were grown overnight at 37 °C, then inoculated into 3 ml LB medium with kanamycin at a ratio of 1:100 (v/v). To induce protein expression, IPTG was added to a final concentration of 0.5 mM when OD_{600} of cultures reached 0.8. Cultures were grown overnight at 16 °C to allow for protein expression, and cells were harvested by centrifugation and resuspended in phosphate buffered saline (PBS), followed by sonication and centrifugation. The samples of both insoluble (precipitation) and soluble fractions (supernatant) were placed on 12% SDS–PAGE gels, and protein bands were visualized by Coomassie Brilliant Blue staining. Gels were scanned by use of the G:BOX Gel imaging system (Syngene, Frederick, MD, USA).

Large-scale production of soluble MBP-T3-CYSC fusion protein

The *MBP-T3-CYSC* fusion protein was expressed at 16 °C as described above, except that the culture volume was shifted to 2 L. Cells from 2 L of shaken-flask cultures were harvested and resuspended in 100 ml buffer A (50 mM Tris–HCl, 500 mM NaCl, pH 8.0) containing 1 mM imidazole, 1 mM PMSF and disrupted in a high pressure homogenizer. After centrifugation (18,000*g* for 20 min, 4 °C), the supernatant was clarified by 0.45-μm filtration and loaded onto Ni-Sepharose FF resin. After a wash with buffer A containing 20 mM imidazole, the *MBP-T3-CYSC* fusion protein was eluted by use of buffer A plus 150 mM imidazole. The buffer of elution was then changed to buffer B (50 mM Tris–HCl, 150 mM NaCl, pH 7.5). The protein concentration was determined by use of the BCA Protein Quantification Kit (Vazyme, Piscataway, NJ, USA).

3C Protease cleavage and purification of de-tagged CYSC

3C protease and the *MBP-T3-CYSC* fusion protein (1 mg/ml) were mixed in buffer B at a ratio of 1:100 (w/w for enzyme/substrate), and incubated at 4 °C overnight. Mixtures were reloaded onto Ni Sepharose FF resin equilibrated with buffer B. The de-tagged CYSC was recovered in the flow-through (unbound) fraction and the remaining parts were recovered by washing the resin with buffer B supplied with 250 mM imidazole. The de-tagged CYSC protein was aliquoted and stored at −80 °C.

SDS–PAGE and Western blot analysis of de-tagged CYSC protein

Samples of de-tagged recombinant CYSC protein and natural CYSC protein standard were placed on two 12% SDS–PAGE gels. One protein band was visualized by Coomassie Brilliant Blue staining. The gels were scanned by use of the G:BOX Gel imaging system (Syngene). Proteins of the other SDS–PAGE gel were transferred to Immobilon-P polyvinylidene difluoride membranes (Millipore,

Table 1
Primers for molecular cloning.

Primer	Sequence	Introduced sites
Linker sense	ctagcactagtgaaaacttatattttcagggcctggaagttctgttccaggggcccg	SpeI
Linker antisense	gatccgggcccctggaacagaacttccaggccctgaaaatataagttttcactagtg	SpeI
CYSC-F	Ctgttccaggggcccggatcctccagtcccggcaagccgcc	BamHI
CYSC-R	Ctcgagtgcggccgcaagcttttaggcgtcctgacaggtgg	Hind III
MBP-F	Gtgccgcgcggcagccatatgaaaatcgaagaaggtaaactg	NdeI
MBP-R	aaaatataagttttcactagtagtctgcgcgtctttcaggg	SpeI
GST-F	gtgccgcgcggcagccatatgTCCCCTATACTAGGTTATTGG	NdeI
GST-R	aaaatataagttttcactagtACGCGGAACCAGATCCGATTTTG	SpeI
IF2-F	Gtgccgcgcggcagccatatgacagatgtaacgattaaaacg	NdeI
IF2-R	aaaatataagttttcactagtcactttgtctttttccgcagc	SpeI

Underline font represents restriction sites.