Hindawi Publishing Corporation
Journal of Biomedicine and Biotechnology
Volume 2012, Article ID 732017, 7 pages
doi:10.1155/2012/732017

Research Article

Codon Preference Optimization Increases Prokaryotic Cystatin C Expression

Qing Wang,[1] Cui Mei,[2] Honghua Zhen,[2] and Jess Zhu[2]

[1] *Department of Clinical Laboratory, The Affiliated Hospital of Qingdao University Medical College, Qingdao 266003, China*
[2] *Research Center, AI DE Diagnostic Co., LTD., Qingdao 266003, China*

Correspondence should be addressed to Qing Wang, wangqing0533@yahoo.com.cn

Received 12 June 2012; Revised 14 August 2012; Accepted 22 August 2012

Academic Editor: Jozef Anné

Gene expression is closely related to optimal vector-host system pairing in many prokaryotes. Redesign of the human *cystatin C* (*cysC*) gene using the preferred codons of the prokaryotic system may significantly increase *cysC* expression in *Escherichia coli* (*E. coli*). Specifically, *cysC* expression may be increased by removing unstable sequences and optimizing GC content. According to *E. coli* expression system codon preferences, the gene sequence was optimized while the amino acid sequence was maintained. The codon-optimized *cysC* (co-*cysC*) and wild-type *cysC* (wt-*cysC*) were expressed by cloning the genes into a pET-30a plasmid, thus transforming the recombinant plasmid into *E. coli* BL21. Before and after the optimization process, the prokaryotic expression vector and host bacteria were examined for protein expression and biological activation of CysC. The recombinant proteins in the lysate of the transformed bacteria were purified using Ni^{2+}-NTA resin. Recombinant protein expression increased from 10% to 46% based on total protein expression after codon optimization. Recombinant CysC purity was above 95%. The significant increase in *cysC* expression in *E. coli* expression produced by codon optimization techniques may be applicable to commercial production systems.

1. Introduction

The use of *E. coli* expression systems for foreign protein production by recombination has been well documented. Such expression systems possess superior characteristics, including fast growth rates, inexpensive fermentation media, and documented genetic code. The efficiency, and thus the cost of production of recombinant proteins in this microorganism, depends on highly variant protein expression levels [1, 2]. The protein expression level of foreign genes is impacted by the expression system, the specific nature of the foreign genes, and the regulation of protein expression. Each of these factors is also highly variable based on both the protein and host. Species-specific variations in codon usage are often cited as one of the major factors affecting protein expression, suggesting the effectiveness of codon optimization to suit a particular expression system, such as *E. coli* [3, 4]. Additionally, some rare codon varieties may produce only low levels of their cognate tRNAs, thus reducing the translational rate while simultaneously increasing the risk of translational errors. Because this may significantly reduce functional protein production, protein expression will be significantly influenced [5, 6]. Over the past decades, researchers have achieved success in redesigning many gene codons to improve their expression in certain systems, generating numerous commercially credible techniques [1–3, 7–10].

The protein Cystatin C (CysC) is an alkaline and non-glycosylated small protein, also known as the Y-trance protein or post-Y-globulin. CysC is composed of 122 amino acids, its isoelectric point is 9.3, and its molecular weight is 13.3×10^3 Daltons [11]. Although the CysC fusion protein can be expressed in *E. coli* [12], expression efficiency in eukaryotic genes using prokaryotic expression vectors is low. This suggests a preference for prokaryotic codons in the *E. coli* system. In many cases, specific bacterial and mammalian systems have shown a preference for the use of different codons with specific and unique characteristics [6]. Codon optimization is a genetic technique that has been previously used to achieve optimum expression of

a foreign gene based on the specific nature of the host system. These techniques are often based on previously identified preferred codons, relative to the system of interest. During the process of optimization, existent codons are replaced by a set of more suitable host codons [6, 13–15]. Because the gene of interest, *cysC*, is a mammalian gene that is poorly expressed in most bacterial systems, several alternatives exist to improve gene expression. Codon optimization or, alternatively, heterologous expression can improve expression by supplying the host with extra copies of tRNA molecules [16].

Typically, two strategies have been used for codon optimization. In the first strategy, commonly referred to as the "one amino acid-one codon" method, the most abundant host codon or a set of selected genes is assigned to all instances of a given amino acid in the target sequence [4, 8, 14, 17]. A variation of the first strategy is employed by the current study, termed the "codon randomization" method. This method applies translation tables based on the frequency distribution of the codons across an entire genome or a subset of highly expressed genes. These constructs are then used to assign a weighted value to each codon. Notably, random assignment of codons based on previously determined probability weights has proven successful in codon optimization for *E. coli* in previous studies [2, 7, 8, 18, 19].

According to the codon preferences previously observed in the *E. coli* expression system [6, 13–15, 20–22], the current study examines a method for codon optimization of the gene sequence for production of CysC in *E. coli*. This technique seeks not only to achieve optimization, but also to maintain the integrity of the amino acid sequence. Expression of the codon optimized (co-*cysC*) and wild type (wt-*cysC*) may be induced by cloning genes into a pET-30a plasmid, thus transforming *E. coli* into BL21 competent cells. Before codon optimization and after the procedure, the expression and biological activation of CysC in the prokaryotic expression vector and host bacteria were compared by purification of the total protein produced by the cell. The current study utilizes codon optimization strategies based on those developed for use in other protein systems in order to improve the expression of *cysC* in *E. coli*.

2. Materials and Methods

2.1. Materials. The pMD-18T vector was purchased from Takara (Shiga, Japan). *E. coli* BL21 competent cells, Platinum Pfx DNA Polymerase, Platinum HIFI Taq Polymerase, and pET30-a vectors were obtained from Invitrogen (California, USA). EcoRI and NotI were purchased from MBI (Fermentas, USA). Additionally, ligase, $Ni2^{+}$-NTA resin, and mouse anti-human CysC monoclonal antibodies were obtained from NEB (Beijing LTD., Beijing, China), QIAGEN (Dusseldorf, Germany), and Sigma (St. Louis, USA), respectively.

2.2. Construction of Recombinant wt-cysC Expression Vectors. The total cellular RNA was extracted using RNAiso reagent (Takara, Japan) according to the instructions provided by the manufacturer. The total RNA of human promyelocytic leukemia (HL-60) cells (Key laboratory molecular virology of Shandong Province, China) was analyzed using 1% agarose gel electrophoresis. Resultant bands and band intensities were observed and recorded. The A_{260} value was detected using an ultraviolet spectrophotometer (Lambda 45, USA). The CysC cDNA was obtained by reverse transcription polymerase chain reaction (RT-PCR) using the primers: 5′-GAATTCATGGCCGGGCCCCTGCGC-3′ (sense) (underlined portion represents the *EcoRI* enzyme cut site) and 5′-GCGGCCGCCTAGGCGTCCTGACAGGTGGA-3′ (antisense) (underlined portion represents the *NotI* enzyme cut site). The PCR products were purified, double-digested, extracted, and inserted into *EcoRI/NotI* sites of the pET-30a vector. The recombinant pET-30a-C plasmid was constructed and identified using double-digestion electrophoresis and DNA sequencing (Invitrogen Ltd., Shanghai, China).

2.3. Construction of Recombinant co-cysC Expression Vectors. The mRNA sequence for CysC provided by GenBank (Gene ID;1471) was analyzed, and the gene sequence was optimized in accordance with the codon preference characteristics previously determined for the *E. coli* expression system [2, 10, 23, 24]. The "codon randomization" method developed in this study employs translation tables. In this method, weighted values were assigned based on frequency distributions of each of the genomic codons. The strategy used for codon optimization was based on random assignment of a triplet for each amino acid according to an established preference table (http://www.kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=316407). Probability was based on the weight of each codon within the set encoding a given amino acid. Using this algorithm, one sequence was designed using the GeMS software package (KOSAN Biosciences Inc.) [18]. A complete comparison of the resultant gene sequence designs for co-*cysC* and wt-*cysC* are shown in Table 1.

According to analyses of these gene sequences, a single strand oligo was successfully designed and synthesized. *EcoRI* and *NotI* were added to the 5′ and 3′ ends of each sequence. The synthesized oligo was spliced to form an integrated gene using polymerase chain reaction (PCR). The synthesized sequence was inserted into a pMD-18T vector, and this vector was further transformed into *E. coli* BL21 competent cells. The validity of the gene sequence of the recombinant plasmid was confirmed by sequencing. A fragment of approximately 400 bp was obtained from the pMD-18T-*cysc* plasmid using *EcoRI* and *NotI*. The recombinant pET-30a-co plasmid was obtained by ligating this fragment with the pET-30a vector. The correct open reading frame (ORF) was verified, and the recombinant plasmid was transformed into the *E. coli* BL21 competent cells. The recombinant plasmids were extracted using alkaline lysis, and the gene sequence was confirmed using double digestion of *EcoRI* and *NotI* combined with gene sequencing.

2.4. Inducible Expression and Purification of Fusion Protein. The pET-30a-co and pET-30a-wt with correct sequences were transformed into the *E. coli* BL21 strain, and three

Table 1: Sequence comparison between wt- and co-*cysC* genes.

aa	S	S	P	G	K	P	P	R	L	V	G	G	P	M	D	A	S	V	E	E
co	TCT	TCT	CCG	GGT	AAA	CCG	CCG	CGT	CTG	GTT	GGT	GGT	CCG	ATG	GAC	GCT	TCT	GTT	GAA	GAA
wt	TCC	AGT	CCC	GGC	AAG	CCG	CCG	CGC	CTA	GTG	GGA	GGC	CCC	ATG	GAC	GCC	AGC	GTG	GAG	GAG
aa	E	G	V	R	R	A	L	D	F	A	V	G	E	Y	N	K	A	S	N	D
co	GAA	GGT	GTT	CGT	CGT	GCT	CTG	GAC	TTC	GCT	GTT	GGT	GAA	TAC	AAC	AAA	GCT	TCT	AAC	GAC
wt	GAG	GGT	GTG	CGG	CGT	GCA	CTG	GAC	TTT	GCC	GTC	GGC	GAG	TAC	AAC	AAA	GCC	AGC	AAC	GAC
aa	M	Y	H	S	R	A	L	Q	V	V	R	A	R	K	Q	I	V	A	G	V
co	ATG	TAC	CAC	TCT	CGT	GCT	CTG	CAG	GTT	GTT	CGT	GCT	CGT	AAA	CAG	ATC	GTT	GCT	GGT	GTT
wt	ATG	TAC	CAC	AGC	CGC	GCG	CTG	CAG	GTG	GTG	CGC	GCC	CGC	AAG	CAG	ATC	GTA	GCT	GGG	GTG
aa	N	Y	F	L	D	V	E	L	G	R	T	T	C	T	K	T	Q	P	N	L
co	AAC	TAC	TTC	CTG	GAC	GTT	GAA	CTG	GGT	CGT	ACC	ACC	TGC	ACC	AAA	ACC	CAG	CCG	AAC	CTG
wt	AAC	TAC	TTC	TTG	GAC	GTG	GAG	CTG	GGC	CGA	ACC	ACG	TGT	ACC	AAG	ACC	CAG	CCC	AAC	TTG
aa	D	N	C	P	F	H	D	Q	P	H	L	K	R	K	A	F	C	S	F	Q
co	GAC	AAC	TGC	CCG	TTC	CAC	GAC	CAG	CCG	CAC	CTG	AAA	CGT	AAA	GCT	TTC	TGC	TCT	TTC	CAG
wt	GAC	AAC	TGC	CCC	TTC	CAT	GAC	CAG	CCA	CAT	CTG	AAA	AGG	AAA	GCA	TTC	TGC	TCT	TTC	CAG
aa	I	Y	A	V	P	W	Q	G	T	M	T	L	S	K	S	T	C	Q	D	A
co	ATC	TAC	GCT	GTT	CCG	TGG	CAG	GGT	ACC	ATG	ACC	CTG	TCT	AAA	TCT	ACC	TGC	CAG	GAC	GCT
wt	ATC	TAC	GCT	GTG	CCT	TGG	CAG	GGC	ACA	ATG	ACC	TTG	TCG	AAA	TCC	ACC	TGT	CAG	GAC	GCC

Note: wt: wild-type; co: codon-optimized; aa: amino acid.

colonies were grown on ALB medium. Three single bacterial colonies were independently selected from separate colonies. A single bacterial colony of each clone was inoculated into 5 mL LB medium and cultured at 37°C for 10 h. The resultant bacteria were subsequently inoculated into 500 mL LB medium with 100 μg/mL ampicillin in the ratio of 1:100 and cultured at 37°C until the value of A_{600} reached 0.5. Protein expression in the bacteria was induced by 0.25 mmol/L isopropyl-β-D-thiogalactoside (IPTG), and the bacteria were harvested after 1, 2, 3, 4, and 5 h.

Cells in 50 mmol/L Tris-C1 (pH 8.0) were lysed using three 12 s rounds of ultrasonication. The precipitate was dissolved by inclusion in a body buffer (8 mol/L urea, l00 mmol/L NaCl, 1 mmol/L EDTA, 0.1 mmol/L PMSF, 50 mmol/L Tris-C1, pH 8.0), and the result was centrifuged at 4°C for 10 min at 6000 g. This resulted in an obvious supernatant for collection. The urea concentration was gradually decreased to 0.5 mmol/L using renaturation solution (10 mmol/L GSH, 2 mmol/L GSSG, 50 mmol/L Tris-C1, pH 8.0) every 1 h. The proteins were renatured at 4°C overnight. Remnant urea was removed by mixing with dialysate (50 mmol/L Tris-C1, pH 8.0) at 4°C. The mixture was centrifuged at 4°C for 10 min at 6000 g. Then, 1.5 mL supernatant (0.5 g/L) was loaded and purified using $Ni2^{+}$-NTA resin.

2.5. Detection of Recombinant Proteins. The purified recombinant proteins were diluted using 0.04 M phosphate buffer (PBS) and detected by sol particle immunoassays (SPIAs) (Leadmanbio, China) using a 7600 Chemistry Analyzer (Hitachi, Japan). The recombinant *Brucella* protein in the same vector and host cells was set as a negative control. This additionally served as a high-value and low-value quality control material.

2.6. Western Blot Analysis. The proteins produced by transformed bacteria were analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). The results were transferred onto a nitrocellulose film. The film was blocked using 5% skim milk in Tris-HCl buffer solution-Tween (TBST) (25 mmol/L Tris-HCI, 125 mmol/L NaCl, 0.1% Tween 20, pH 8.0) at 37°C for 2 h. The film was washed three times for 3 min and then incubated with mouse anti-human CysC monoclonal antibodies at 37°C for 1.5 h. The film was washed with TBST three times and incubated with horseradish-peroxidase- (HRP-)tagged goat anti-mouse IgG at 37°C for 1.5 h. Then, the film was washed and developed in tetramethylbenzidine (TMB) for 10 min.

3. Results

3.1. Optimization of the cysC Gene. The mRNA sequence of *cysC* was obtained from GenBank (Gene ID: 1471), and 63 of the 120 codons presented were optimized. Furthermore, the complex secondary structure and repeated sequences of the gene were screened. The rare codons in *cysC* were replaced by the preferred codons for *E. coli*, and the GC content was decreased from 60.88 % to 52.89%. Additionally, the AT-rich fragment was removed to avoid premature termination. Optimization was conducted to carefully avoid the inclusion of regions containing more than 5 A/T or G/C repeated sequences, regions that could potentially affect mRNA stabilization. The synthesized *cysC* gene was modified by synonymous codons while the integrity of the amino acid structure was maintained.

3.2. Identification of Prokaryotic Recombinant Plasmids. Recombinant pET-30a-co and pET-30a-wt plasmids were extracted by alkaline lysis and analyzed by double-digestion of *EcoRI* and *NotI*. Identification and sequencing were conducted for both plasmids that met the design requirements. Viable recombinant clones were successfully constructed.

3.3. Expression and Purification of CysC. The bacterial proteins of pET30a-co and pET30a-wt were analyzed, and the recombinant proteins were presented as an inclusion body. SDS-PAGE indicated that a 19-kilo Dalton (kDa) protein band was uniquely present in the transformed bacteria. Electrophoresis analysis indicated results in accordance with the theoretical value of 19.0 kDa.

At a time 5 h later, 0.25 mg/L IPTG induced the highest expression of fusion protein. The expression of pET30a-wt recombinant proteins accounted for 10% of the total protein present in the transformed bacteria, while the expression pET30a-co recombinant proteins in three colonies accounted for 47%, 49%, and 42% (average of 46%) of the total protein present in the transformed bacteria. Recombinant proteins found in the lysate of the inclusion body were purified using $Ni2^{+}$-NTA resin. The purified fusion proteins were presented at about 19000-Mr with a purity in excess of 90% (Figure 1).

3.4. SPIA Detection of the Diluted Recombinant Protein. Five hours after the bacteria was induced, the bacteria with PET30a-wt presented 36 mg/L CysC. The three colonies of pET30a-co generated yields up to 779 mg/L, 827 mg/L, and 770 mg/L (average of 792 mg/L). The standard errors of protein expression in the three colonies of pET30a-co were 30.64 mg/L, and the coefficient of variance (CV) was 3.97%. However, 0 mg/L CysC was present in the negative control. The co-cysC sequence produced significantly more protein than the wt-cysC sequence.

3.5. Identification of Recombinant Protein Antigenicity by Western Blot. The anti-human CysC monoclonal antibodies were shown to react with single recombinant proteins with an approximate molecular mass of about 19.0 kDa (Figure 2), suggesting that the recombinant proteins were human CysC.

4. Discussion

CysC, a member of cysteine proteinase inhibitor family, has an unclear biological function that has made it the target of numerous research studies. In order to improve and promote the study of this compound, more efficient and affordable synthetic production techniques for this protein

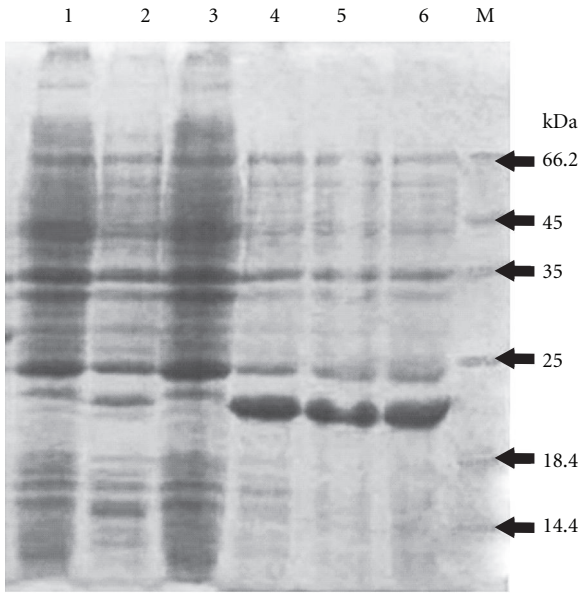

Figure 1: SDS-PAGE analysis of *cysC* expression in *E. coli*. (1) pET-30a-wt prior to induction; (2) 5 h after the induction of pET-30a-wt; (3) pET-30a-co prior to induction; (4–6) 5 h after the induction of pET-30a-co (three separate colonies); (M) Low molecular weight protein marker.

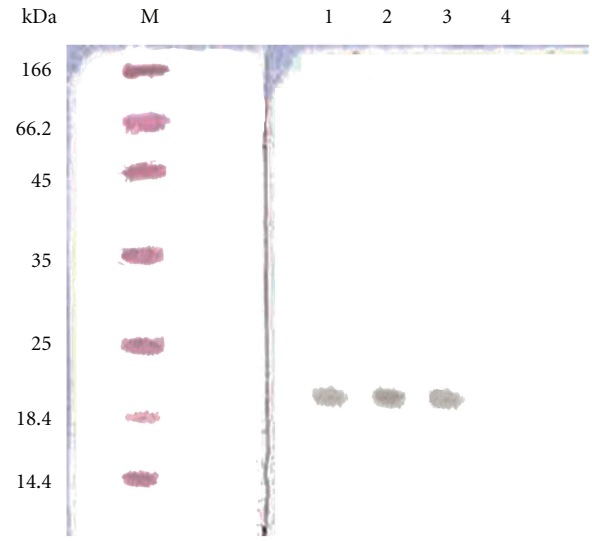

Figure 2: Western blot analysis of CysC protein expression. (M) Low molecular weight protein marker; (1–3) Proteins after induction of pET-30a-co (three separate colonies); (4) Proteins before induction of pET-30a-co.

in bacterial agents are required. Previous studies pertaining to the molecular structure and metabolism of CysC have demonstrated that it is present in virtually all nucleated cells. In addition, it demonstrates no tissue specificity, and stable production and cyclical levels are observed in many cell types. Most notably, the expression of CysC has not been shown to be related to pathology, age, gender, or metabolism in humans.

As one of the few proteins freely filtered by the glomeruli without significant reabsorption or secretion, CysC concentrations in human blood serum are an ideal index for glomerular filtration, a primary potential application for the protein [11, 23]. A study of 135 patients conducted by Grubb et al. found that the reciprocal of serum CysC was significantly related to the glomerular filtration rate, with a correlation coefficient reaching the notable level of 0.77 [24]. It has been reported that the accuracy and sensitivity of CysC for diagnosis of glomerular filtration rates were higher than those reported for either serum urea nitrogen or serum creatinine [11, 25–27]. Because extensive further research in development and testing of novel techniques for CysC utilization require large amounts of the protein, efficient bacterial synthesis is rapidly becoming a critical process in both research and clinical studies involving CysC.

The development of high-quality antibodies, such as egg yolk antibody (IgY), for the detection of CysC is a prominent goal in contemporary research. Preparation and purification of the protein antigen of CysC necessary for this process also increase the demand for bacterially produced CysC. Although CysC fusion protein was successfully expressed in *E. coli* [12], the low expression efficiency of the eukaryotic gene using prokaryotic expression vectors has remained problematic. Thus, codon optimization has recently become an increasingly important tool for the commercialization of bacterial gene expression. When researchers neglect to select appropriate expression vectors and host systems, ignoring optimal matching of the vector and host, production of certain proteins can be much less efficient. In many cases, this may result in resource utilization and costs that prove to be prohibitive for many researchers [28].

Each amino acid corresponds to a minimum of one codon and a maximum of six codons based on the known degeneracy of codons. This trend must be considered when matching eukaryotic proteins to bacterial hosts in order to achieve optimal expression. Notably, significant differences in codons also appear between different organism types and individual species. In protozoa, high expression of certain genes may be attributed to a preference for certain codons [29–31]. These preferences are generally related to the variations between prokaryotic and eukaryotic systems, which can cause some codons (often codons similar to the stop codon) to prematurely initiate breaks in the translational sequence. Additionally, certain codons may result in limited tRNA supplies, while abundant tRNA supplies are required to ensure translation efficiency [31]. The preferential selection of codons also correlates with the structure and function of the coded proteins.

Rare codons in mRNA are often associated with linkage areas, wherein translational rates exhibit notable decreases. These rare codons also result in unique protein domains and regular secondary unit structures. Similarly, the translational rate varies at different mRNA fragments, largely based on variations in secondary protein structures. Many organisms possess significantly different codon preferences, as indicated by the vast differences observed in real exons. In genetic

engineering, the target gene is generally a triplet sequence rather than the natural exon. Hence, the selected target gene must be analyzed in terms of its species-specific preference in order to obtain the highest expression of recombinant proteins. In bacterial cells, this process includes the removal of rare codons, utilization of preferential codons, minimization of secondary structure variation, and regulation of GC content. Cumulatively, these processes are generally referred to as "codon optimization."

Previous studies [12] have shown that eukaryotic wt-*cysC* was able to be successfully expressed in *E. coli*; however, common eukaryotic codons produced only low expression levels due to the absence of prokaryotic preferential codons. Through codon optimization, optimal expression levels can be achieved in these systems, resulting in much higher protein yields. In addition to the use of optimized codons, several studies [32–34] have also reported that the expression and ELISA titer of synthesized genes markedly increased upon the removal of rare codons and unstable sequences. Bagherpour et al. synthesized *fimH* of Uropathogenic *Escherichia coli* (*UPEC*) using mammalian codons [35]. Compared with the wild-type gene, the *fimH* gene of the mammalian codon has been shown to be compatible with eukaryotic expression systems. Therefore, the mammalian codon may be appropriated in a *fimH* construct as a DNA vaccine in COS7 cells. Anzor et al. reported that the expression of synthesized genes in *E. coli* was 3.4-fold greater than that of wt-PEDF in the native host when codon optimization was applied. Similarly, Menzella reported that the codon randomization method was a superior strategy for codon optimization [36]. Furthermore, Menzella demonstrated significant increases in chymosin protein expression, demonstrating the effectiveness of this strategy for reducing production costs in industrial enzyme processes that use microbial hosts.

The target gene, *cysC*, was redesigned over the course of the current study using the optimal techniques described by previous researchers [2, 10, 23, 24], including the removal of rare codons and unstable sequences. In addition, previously successful techniques in other similar expression systems were also applied, including utilization of optimized codons as well as reduction in AT- and GC-repeated sequences. The codons were modified, and the GC content of the target gene was decreased. In addition, the AT rich fragment was removed to avoid premature translational termination. This data suggested that the expression and titer of synthesized CysC were remarkably increased after optimization of induction concentration and time with IPTG.

5. Conclusions

The results of the current study confirm that production of CysC in prokaryotic systems, such as *E.* coli, may be improved through the application of codon optimization techniques. These findings demonstrate that expression of the recombinant protein (46%) was significantly higher in optimized systems than in the wild type (10%). The concentration of CysC dramatically increased from 36 mg/L to 792 mg/L. As the demand for CysC is expected to rise as an increasing number of researchers and clinicians begin to utilize the protein, increased expression may potentially represent significant gains for commercial enzyme producers. Though the original prokaryotic expression of *cysC* in *E. coli* was relatively low, selection of preferential codons for the prokaryotic systems enabled vast improvements in the design of the expression system. The expression of the fusion protein may also be increased by selecting an appropriate vector and host system, with special consideration applied to preferred codons. Researchers must consider that optimal gene and vector matching and host system matching are closely related to resultant protein expression levels. These considerations are important elements for optimal genetic engineering of these systems.

Conflict of Interests

The authors declare no conflict of interests.

Acknowledgment

This research was supported by the Shandong Science and Technology Development Project (no. 2011 YG21030).

References

[1] N. A. Burgess-Brown, S. Sharma, F. Sobott, C. Loenarz, U. Oppermann, and O. Gileadi, "Codon optimization can improve expression of human genes in *Escherichia coli*: a multi-gene study," *Protein Expression and Purification*, vol. 59, no. 1, pp. 94–102, 2008.

[2] M. Welch, S. Govindarajan, J. E. Ness et al., "Design parameters to control synthetic gene expression in Escherichia coli," *PLoS ONE*, vol. 4, no. 9, Article ID e7002, 2009.

[3] Z. Zhou, P. Schnake, L. Xiao, and A. A. Lal, "Enhanced expression of a recombinant malaria candidate vaccine in *Escherichia coli* by codon optimization," *Protein Expression and Purification*, vol. 34, no. 1, pp. 87–94, 2004.

[4] A. Villalobos, J. E. Ness, C. Gustafsson, J. Minshull, and S. Govindarajan, "Gene Designer: a synthetic biology tool for constructuring artificial DNA segments," *BMC Bioinformatics*, vol. 7, p. 285, 2006.

[5] T. Ikemura, "Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes: a proposal for a synonymous codon choice that is optimal for the *E. coli* translational system," *Journal of Molecular Biology*, vol. 151, no. 3, pp. 389–409, 1981.

[6] C. Gustafsson, S. Govindarajan, and J. Minshull, "Codon bias and heterologous protein expression," *Trends in Biotechnology*, vol. 22, no. 7, pp. 346–353, 2004.

[7] S. J. Kodumal, K. G. Patel, R. Reid, H. G. Menzella, M. Welch, and D. V. Santi, "Total synthesis of long DNA sequences: aynthesis of a contiguous 32-kb polyketide synthase gene cluster," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 44, pp. 15573–15578, 2004.

[8] X. Wang, X. Li, Z. Zhang, X. Shen, and F. Zhong, "Codon optimization enhances secretory expression of Pseudomonas

aeruginosa exotoxin A in *E. coli*," *Protein Expression and Purification*, vol. 72, no. 1, pp. 101–106, 2010.

[9] H. G. Menzella, J. R. Carney, and D. V. Santi, "Rational design and assembly of synthetic trimodular polyketide synthases," *Chemistry and Biology*, vol. 14, no. 2, pp. 143–151, 2007.

[10] H. G. Menzella, S. J. Reisinger, M. Welch et al., "Redesign, synthesis and functional expression of the 6-deoxyerythronolide B polyketide synthase gene cluster," *Journal of Industrial Microbiology and Biotechnology*, vol. 33, no. 1, pp. 22–28, 2006.

[11] E. Wasén, P. Suominen, R. Isoaho et al., "Serum cystatin C as a marker of kidney dysfunction in an elderly population [5]," *Clinical Chemistry*, vol. 48, no. 7, pp. 1138–1140, 2002.

[12] T. Chen, J. Feng, and J. Cao, "Prokaryotic expression, purification, identification of human cystatin C and preparation of its antiserum," *Acta Academiae Medicinae Militaris Tertiae*, vol. 29, no. 10, pp. 938–940, 2007.

[13] K. S. Sandhu, S. Pandey, S. Maiti, and B. Pillai, "GASCO: genetic algorithm simulation for codon optimization," *In Silico Biology*, vol. 8, no. 2, pp. 187–192, 2008.

[14] W. Gao, A. Rzewski, H. Sun, P. D. Robbins, and A. Gambotto, "UpGene: application of a web-based dna codon optimization algorithm," *Biotechnology Progress*, vol. 20, no. 2, pp. 443–448, 2004.

[15] Y. Cai, J. Sun, J. Wang, Y. Ding, and X. Liao, "One codon usage optimization software for protein heterologous expression," *Journal of Bioinformatics and Sequence Analysis*, vol. 2, no. 2, pp. 25–29, 2010.

[16] C. P. Carstens, "Use of tRNA-supplemented host strains for expression of heterologous genes in *E. coli*," *Methods in Molecular Biology*, vol. 205, pp. 225–233, 2003.

[17] F. Supek and K. Vlahoviček, "INCA: synonymous codon usage analysis and clustering by means of self-organizing map," *Bioinformatics*, vol. 20, no. 14, pp. 2329–2330, 2004.

[18] S. Jayaraj, R. Reid, and D. V. Santi, "GeMS: an advanced software package for designing synthetic genes," *Nucleic Acids Research*, vol. 33, no. 9, pp. 3011–3016, 2005.

[19] H. G. Menzella, R. Reid, J. R. Carney et al., "Combinatorial polyketide biosynthesis by de novo design and rearrangement of modular polyketide synthase genes," *Nature Biotechnology*, vol. 23, no. 9, pp. 1171–1176, 2005.

[20] N. M. Marlatt, D. E. Spratt, and G. S. Shaw, "Codon optimization for enhanced *Escherichia coli* expression of human S100A11 and S100A1 proteins," *Protein Expression and Purification*, vol. 73, no. 1, pp. 58–64, 2010.

[21] Z. Feng, L. Zhang, X. Han, and Y. Zhang, "Codon optimization of the calf prochymosin gene and its expression in Kluyveromyces lactis," *World Journal of Microbiology and Biotechnology*, vol. 26, no. 5, pp. 895–901, 2010.

[22] A. Fuglsang, "Codon optimizer: a freeware tool for codon optimization," *Protein Expression and Purification*, vol. 31, no. 2, pp. 247–249, 2003.

[23] M. Abrahamson, I. Olafsson, A. Palsdottir et al., "Structure and expression of the human cystatin C gene," *Biochemical Journal*, vol. 268, no. 2, pp. 287–294, 1990.

[24] A. Grubb, O. Simonsen, and G. Sturfelt, "Serum concentration of cystatin C, factor D and β2-microglobulin as a measure of glomerular filtration rate," *Acta Medica Scandinavica*, vol. 218, no. 5, pp. 499–503, 1985.

[25] M. Pergande and K. Jung, "Sandwich enzyme immunoassay of cystatin C in serum with commercially available antibodies," *Clinical Chemistry*, vol. 39, no. 9, pp. 1885–1890, 1993.

[26] K. Jung and M. Jung, "Cystatin C: a promising marker of glomerular filtration rate to replace creatinine [2]," *Nephron*, vol. 70, no. 3, pp. 370–371, 1995.

[27] D. J. Newman, H. Thakkar, R. G. Edwards et al., "Serum cystatin C: a replacement for creatinine as a biochemical marker of GFR," *Kidney International, Supplement*, supplement 47, pp. S-20–S-21, 1994.

[28] B. Watelet, M. Quibriac, D. Rolland et al., "Erratum to 'characterization and diagnostic potential of hepatitis B virus nucleocapsid expressed in *E. coli* and *P. pastoris*'," *Journal of Virological Methods*, vol. 102, no. 1-2, pp. 175–190, 2002.

[29] X. Chen, B. Gao, X. Y. Jia et al., "High-level expression of phenylalanine ammonia-lyase in Lactococcus Lactis via synthesized sequence based on bias codons," *Chinese Journal of Biotechnology*, vol. 22, no. 2, pp. 187–190, 2006.

[30] G. Sinclair and F. Y. M. Choy, "Synonymous codon usage bias and the expression of human glucocerebrosidase in the methylotrophic yeast, Pichia pastoris," *Protein Expression and Purification*, vol. 26, no. 1, pp. 96–105, 2002.

[31] P. Qian, X. Li, G. Tong, and H. Chen, "High-level Expression of the ORF6 Gene of Porcine Reproductive and Respiratory Syndrome Virus (PRRSV) in Pichia pastoris," *Virus Genes*, vol. 27, no. 2, pp. 189–196, 2003.

[32] C. Zhou, Y. Kuang, and L. Zeng, "Cod ons optimization and high-level expression of hepatitis B virus S gene in Pichia pastoris," *Chinese Journal of Biomedical Engineering*, vol. 13, no. 6, pp. 387–392, 2007.

[33] Z. Gao, Z. Li, Y. Zhang et al., "High-level expression of the Penicillium notatum glucose oxidase gene in Pichia pastoris using codon optimization," *Biotechnology Letters*, vol. 34, no. 3, pp. 507–514, 2012.

[34] A. G. Gvritishvili, K. W. Leung, and J. Tombran-Tink, "Codon preference optimization increases heterologous PEDF expression," *PLoS ONE*, vol. 5, no. 11, Article ID e15056, 2010.

[35] G. Bagherpour, A. A. Fooladi, J. F. Mehrabadi, M. R. Nourani, and B. Einollahi, "Evaluation of mammalian codon usage of fimH in DNA vaccine design," *Acta Microbiologica et Immunologica Hungarica*, vol. 58, no. 4, pp. 259–271, 2011.

[36] H. G. Menzella, "Comparison of two codon optimization strategies to enhance recombinant protein production in *Escherichia coli*," *Microbial Cell Factories*, vol. 10, p. 15, 2011.