J Ind Microbiol Biotechnol (2012) 39:1841–1849
DOI 10.1007/s10295-012-1185-y

Enhanced production of α-cyclodextrin glycosyltransferase in *Escherichia coli* by systematic codon usage optimization

Hua Liu · Jianghua Li · Guocheng Du · Jingwen Zhou · Jian Chen

Received: 28 June 2012 / Accepted: 31 July 2012 / Published online: 17 August 2012

Abstract Enhancing the production of α-cyclodextrin glycosyltransferase (α-CGTase) is a key aim in α-CGTase industries. Here, the mature α-*cgt* gene from *Paenibacillus macerans* JFB05-01 was redesigned with systematic codon optimization to preferentially match codon frequencies of *Escherichia coli* without altering the amino acid sequence. Following synthesis, codon-optimized α-*cgt* (coα-*cgt*) and wild-type α-*cgt* (wtα-*cgt*) genes were cloned into pET-20b(+) and expressed in *E. coli* BL21(DE3). The total protein yield of the synthetic gene was greater than wtα-*cgt* expression (1,710 mg L^{-1}) by 2,520 mg L^{-1}, with the extracellular enzyme activity being improved to 55.3 U mL^{-1} in flask fermentation. ΔG values at -3 to +50 of the *pelB* site of both genes were −19.10 kcal mol^{-1}. Functionally, coα-CGTase was equally as effective as wtα-CGTase in forming α-cyclodextrin (α-CD). These findings suggest that preferred codon usage is advantageous for translational efficiency to increase protein expression. Finally, batch fermentation was applied, and the extracellular coα-CGTase enzyme activity was 326 % that of wtα-CGTase. The results suggest that codon optimization is a reasonable strategy to improve the yield of α-CGTase for industrial application.

Keywords *Escherichia coli* · α-Cyclodextrin glycosyltransferase · Codon usage optimization · Flask fermentation · Batch fermentation

Electronic supplementary material The online version of this article (doi:10.1007/s10295-012-1185-y) contains supplementary material, which is available to authorized users.

H. Liu · J. Li · G. Du · J. Zhou · J. Chen
National Engineering Laboratory for Cereal Fermentation Technology (NELCF), Jiangnan University, 1800 Lihu Road, Wuxi 214122, Jiangsu, China

J. Zhou
State Key Laboratory of Bioreactor Engineering, East China University of Science and Technology, 130 Meilong Rd, Shanghai 200237, China

J. Zhou (✉) · J. Chen (✉)
School of Biotechnology and Key Laboratory of Industrial Biotechnology, Ministry of Education, Jiangnan University, 1800 Lihu Road, Wuxi 214122, Jiangsu, China
e-mail: zhoujw1982@gmail.com; zhoujw1982@jiangnan.edu.cn

J. Chen
e-mail: jchen@jiangnan.edu.cn

Introduction

Cyclodextrin glycosyltransferase (EC 2.4.1.19, CGTase) is a key extracellular enzyme in cyclodextrin (CD) production, capable of converting starch or starch derivatives into CD via a cyclization reaction [20]. CGTases are classified as α-, β-, and γ-CGTases according to the major CD products, differing in the number of glucose units (α-CD, β-CD, and γ-CD) [25]. Owing to its small internal cavity, high water solubility, and nonbioavailability [3, 18], α-CD has special applications in molecular recognition, nanomaterials, and food industries. Research to enhance the production of α-CGTase has recently attracted much interest.

Industrial applications of CGTase are not limited to production of CD. Many carbohydrate derivatives and other compounds, such as sugar alcohol, vitamins, glycosides, polyols, and flavonoids, can be used as acceptors in CGTase-catalyzed reactions, improving the properties and application of the products [27]; For instance, flavonoids such as hesperidin, rutin, and naringin can be transglycosylated into corresponding glycosyl flavonoids [8]. 2-*o*-α-D-Glucopyranosyl-L-ascorbic acid (AA-2G), which is

a widely used, stable, and nonreducible L-ascorbic acid derivative, can be synthesized enzymatically via transglycosylation with α-CGTase from L-ascorbic acid and β-cyclodextrin [11]. However, the low activity of the CGTase in this reaction is a barrier to the use of the enzyme for large-scale production of AA-2G [34, 35]. Further development of α-CGTase for production of AA-2G requires concentrated samples of the enzyme. Concentration of α-CGTase is a costly and time-consuming process. The cost and time requirements are problematic for any catalysis process that requires production of high-quality α-CD [33]. Therefore, enhancing the yields of α-CGTase is of interest.

Many strategies have been used to enhance the yield of CGTases. CGTase genes have been isolated from more than 30 bacteria, mainly *Bacillus* spp. [27]. The yield of CGTases from wild strains is relatively low. Production of a CGTase from *Bacillus* sp. in recombinant *E. coli* was performed first by Kato and Horikoshi [28]. Since then, many CGTases from various bacteria have been expressed in *E. coli*, including an α-CGTase from a *Paenibacillus macerans* strain isolated in our previous studies [18]. To enhance the heterologous production in *E. coli* of an extracellular enzyme, several molecular approaches and culture strategies have been used, such as manipulation of the signal peptide [14], enhancement of the ABC transporter system [24], and site-directed mutagenesis of the α-CGTase gene [21, 22]. In culture flask, changing the culture medium composition [18], concentration of the inducer [18], and medium supplements [7] (such as calcium [19] and glycine [17]) have also been studied. In fermentor culture, different fed-batch fermentation strategies were attempted to enhance the production of recombinant CGTase [5]. Previous investigations neglected the different codon usages between wild-type *Bacillus* strains and *E. coli*, which should be an important limiting factor for further enhancement of α-CGTase production.

Codon optimization is a method that adjusts codons in a target sequence to the high-frequency codons preferred in the expression host. Codon optimization is in fact based on synonymous mutations that change the nucleotide sequence of a gene without changing the amino acid sequence. Increasing evidence indicates that the choice of the synonymous codons can influence protein production by changing the messenger RNA (mRNA) secondary structure [1] and improving the correlation with the transfer RNA (tRNA) levels in the host [13]. However, few studies have directly examined whether a method in which a complete coding sequence is modified is capable of improving heterologous protein expression while retaining the main functions of the protein.

In this study, the systematic codon optimization strategy was employed to further improve the expression of α-CGTase in *E. coli*. The systematic codon optimization strategy designed here uses translation tables, based on the frequency distribution of the codons in an entire genome or a subset of highly expressed genes, to attach weights to each codon [16]. Compared with the "one amino acid–one codon" approach, the systematic codon optimization strategy also uses the second most preferred synonymous codons that may possibly be necessary when there is a need to pause translation, which avoids usage of rare codons of the host cell. The results show that codon optimization of the α-*cgt* gene significantly increased the yield of the protein compared with that of the wild-type gene in the *E. coli* BL21(DE3) strain. Codon-optimized α-CGTase is as functionally active as wild-type α-CGTase in enzyme activity and in forming α-CD. In addition, batch culture of *E. coli* with different plasmids in a 3-L fermenter was performed. By combining codon optimization of α-*cgt* with flask and batch culture, the extracellular α-CGTase activity was improved by 179 and 226 % compared with the original level, respectively.

Materials and methods

Codon optimization of α-CGTase

The codons of the α-*cgt* gene (2,061 bp, GenBank accession no. AAC04359.1) were optimized based on the codon preference of *Escherichia coli* using a codon algorithm with the GeMS software [15] and a codon table containing a fractional preference for each codon equal to that found in the genome of *E. coli*. The comparative result of the two sequences is presented in Online Resource 1. The DNA sequence was optimized for protein overexpression in *E. coli* without altering the encoded amino acid sequence. Modifications were made throughout the sequence. The codon-optimized α-*cgt* gene (coα-*cgt*, GenBank accession no. JX412224) was assembled successfully using overlapping polymerase chain reaction (PCR; Genscript, Nanjing, China). The free energies (ΔG values) of the 5′ translated *pelB*-coα-*cgt* and *pelB*-wtα-*cgt,* mRNA (representing positions −3 to +50 nt from the *pelB* start site) were calculated using the "mfold" RNA folding program (http://www.bioinfo.rpi.edu/applications/mfold) [36].

Bacterial strains and plasmids

The wild-type α-*cgt* (wtα-*cgt*) coding gene was cloned from *Paenibacillus macerans* JFB05-01 [18]. *Escherichia coli* JM109 (Invitrogen, Carlsbad, CA) was used for maintenance and manipulation of plasmids. *E. coli* BL21(DE3) (Novagen, Madison, WI), harboring the recombinant plasmids, was used for expressing recombinant α-CGTase. Plasmid pET-20b(+) (Novagen, Madison, WI) with the

pelB leader sequence was employed for the construction of pET-20b(+)-wtα-*cgt* and pET-20b(+)-coα-*cgt*.

Culture media

Luria–Bertani (LB) medium was used for seed culture, consisting of peptone 10 g, yeast extract 5 g, and NaCl 10 g/L. Modified Terrific Broth (TB) was employed in all shake flask and batch culture. The modified TB medium contains: glycerol 8 g, peptone 12 g, yeast extract 24 g, KH_2PO_4 2.32 g, K_2HPO_4 16.43 g/L.

Construction of plasmids for expression of coα-*cgt* and wtα-*cgt*

The *Paenibacillus macerans* strain JFB05-01 (CCTCC M2030062) and the synthesized α-*cgt* were used as PCR templates with the primers P1/P2, P3/P4 being used to amplify the wtα-*cgt* and coα-*cgt* genes, respectively (Table 1). The amplified coα-*cgt* and wtα-*cgt* genes were inserted downstream of the *pelB* signal peptide gene in the expression plasmid pET-20b(+), resulting in the plasmids pET-20b(+)-coα-*cgt* and pET-20b(+)-wtα-*cgt* (Fig. 1a, b), respectively. The coα-*cgt* and wtα-*cgt* plasmid constructs were used to transform chemically competent *E. coli* BL21(DE3) cells. Positive colonies were confirmed by colony PCR and DNA sequencing.

Flask fermentation for high-level expression

The transformants were grown in 20 mL LB medium containing 100 μg mL^{-1} ampicillin at 37 °C and 200 rpm for seed culture. Afterward, a 1 % (v/v) concentration of inoculum was inoculated into modified TB medium containing 100 μg mL^{-1} ampicillin for shaking flask culture. Isopropyl-β-D-thiogalactopyranoside (IPTG, 10 μM) was added into the growth medium when the OD_{600} reached 0.6, to induce the expression of the recombinant protein, and the mixture was incubated at 25 °C and 200 rpm. In addition, 1 % (w/v) glycine was added to the culture medium at the beginning of the culture. The total protein concentration and extracellular enzyme activity were assayed every 12 h during induction.

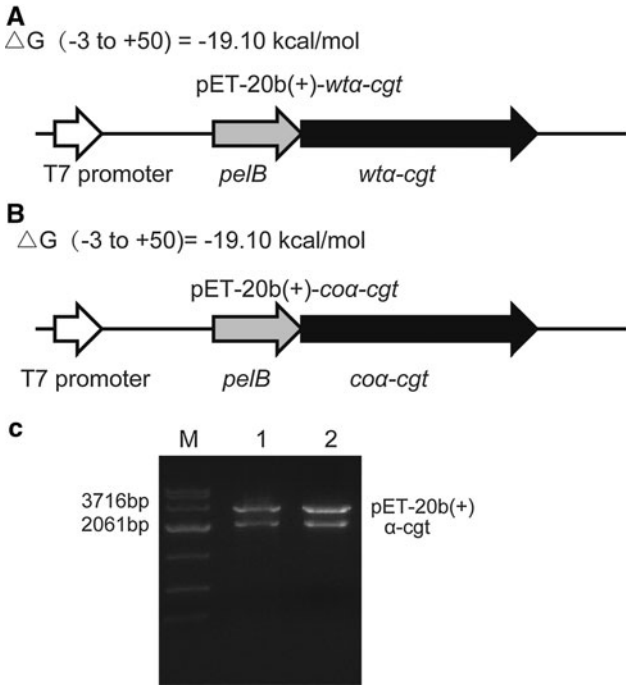

Fig. 1 pET-20b(+) constructs for wtα-CGTase and coα-CGTase protein expression in *E. coli*. The α-CGTase protein and expression vector constructs, designated as **a** and **b**, contained the wtα-*cgt* and the coα-*cgt* genes, respectively, cloned into the vector pET-20b(+) for expression of α-CGTase with the signal peptide *pelB*, and an mRNA folding energy of 19.10 kcal mol^{-1} at the *pelB* start site (−3 to +50). **c** A PCR gel showing the expected product insert size of 2,061 bp for both wtα-*cgt* and coα-*cgt*, and the size of the pET-20b(+) vector. M, DL10 000 DNA marker, 1, pET-20b(+)-wtα-*cgt*, 2, pET-20b(+)-coα-*cgt*

Batch culture

Batch culture was performed in a 3-L fermenter (BioFlo 115; New Brunswick Scientific Co.). Seed cultures (5 %) were inoculated into a 3-L fermentor containing 1.5 L modified TB medium. When the OD_{600} reached 8.0, 50 μmol L^{-1} IPTG was added, and the temperature was decreased to 25 °C. The cell density and the extracellular α-CGTase activity were measured every 3 h. During the entire process, the pH was kept at 7.0 by automatic addition of ammonia solution (25 %, v/v). To control excessive foaming, one or two drops of diluted antifoam were added

Table 1 Primers used in PCR reactions

Name	Primers (5′–3′)*	Enzyme
P1	GGCGAATTCGTCACCCGATACGAGCGTGGAC	*Eco*RI
P2	ATAAGAATGCGGCCGCATTTTGCCAGTCCACCGTCACC	*Not*I
P3	GGCGAATTCGTCACCGGACACCTCAGTGGAC	*Eco*RI
P4	ATAAGAATGCGGCCGCATTTTGCCAATCCACCGTCACCGTGCCAACGCC	*Not*I

* Underlined segments are restriction enzyme cutting sites

to the fermenter. To maintain the dissolved oxygen (DO) level around 30 % of air saturation, the agitation speed was varied from 400 to 900 rpm. The air flow rate was varied from 1.5 to 2.5 L min^{-1}. Ampicillin was added to all media at final concentration of 100 μg mL^{-1}.

Cell fractionation

Before cell fractionation, the OD_{600} of each culture was measured to guarantee that the same biomass was obtained. The sample was centrifuged at 8,000×*g* and 4 °C for 20 min, and the supernatant was used as the extracellular fraction. The resulting pellets containing the cells were resuspended in 500 μL phosphate-buffered saline (PBS) buffer and sonicated to release soluble intracellular proteins. Cell lysates were centrifuged at 8,000×*g* and 4 °C for 20 min to recover the soluble intracellular contents.

Assay of α-CGTase activity

The α-CGTase activity was determined by the methyl orange (MO) method, as described by Li et al. [19].

Miscellaneous methods

The concentration of the strain was determined by OD_{600} using a spectrophotometer (BioPhotometer Plus; Eppendorf, Hamburg, Germany). Samples were diluted approximately with 0.9 % (w/v) NaCl when the OD_{600} exceeded 0.8. To determine the dry cell weight (DCW), 10 mL culture broth was centrifuged at 8,000×*g* for l5 min. The pellet was washed and resuspended with 0.9 % (w/v) NaCl and centrifuged again, and dried to constant weight at 105 °C. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was performed as described previously [5]. Protein concentration was determined by the Bradford method. Production of cyclodextrins was analyzed by high-performance liquid chromatography (HPLC). The details were as described before [22]: 0.2 U mL^{-1} (α-CGTase cyclization activity) wt or coα-CGTase was incubated in 5 % (w/v) soluble starch in 50 mM phosphate buffer (pH 6.0) at 40 °C for 40 h. Samples were taken at regular time intervals and boiled for 10 min. Glucoamylase (Novozymes, Bagsvaerd, Denmark) was added to the boiled sample at final concentration of 1 U mL^{-1}, and then the mixture was incubated at 30 °C for 2 h to eliminate contaminating oligosaccharides while leaving the cyclodextrins intact. Then the samples were determined by the HPLC, using a LiChrosorb NH_2 column (Merck, Darmstadt, Germany) eluted with acetonitrile/water (65:45) at 1 mL/min. Products were detected using a Waters 410 refractive index detector (Waters Corp., Milford, MA, USA).

Results

Gene design, synthesis, and expression vector construction

Analysis of the native α-*cgt* gene revealed that some codons are rare codons for *E. coli* (Table 2); For example, 50 % of the codons encoding for arginine, and 21 out of the 37 codons encoding for leucine found in the native sequence are represented with frequencies below 8 and 11 %, respectively, in the *E. coli* genome. Thus, it is proposed that codon optimization of the α-*cgt* gene might result in an increase in protein expression. The systematic codon optimization strategy was used in this study, with the aid of a preference table (http://www.kazusa.or.jp/e/index.html) and a probability based on the weight of each codon within the set encoding a given amino acid. Using this algorithm, codon-optimized sequences were designed using the GeMS software package, and named coα-*cgt*. The systematic codon optimization strategy permits flexibility in codon selection to facilitate gene design by avoiding: (a) repetitive elements that may lead to gene deletions, (b) internal Shine–Dalgarno sequences, (c) secondary mRNA structures, and (d) restriction sites. The codon distributions of the co and wtα-*cgt* sequences are presented in Table 2. Modifications were made throughout the sequence. Codon-optimized α-*cgt* was assembled successfully using overlapping PCR strategy, and the wtα-*cgt* and coα-*cgt* genes cloned into the pET-20b(+) vector, under the control of the *pelB* peptide. PCR gels confirmed the expected product insert size of 2,061 bp for both wtα-*cgt* and coα-*cgt* genes, and the band of 3,716 bp representing the pET-20b(+) vector (Fig. 1c). After transformation of *E. coli* BL21(DE3) with the constructs, the positive bacterial colonies were identified for further analysis.

Expression of wtα-*cgt* and coα-*cgt* genes in flask cultures

The strains with different plasmids were cultured under the same condition described in the "Materials and methods" section. At 24 h after induction, there was improved protein secretion of the synthetic α-*cgt* gene compared with the native gene (Fig. 2a). Quantitative yields for proteins in the BL21(DE3) host cells are presented in Fig. 2b. At all comparable time points, the total yields of coα-*cgt* were greater than those for the wt gene. The highest levels were observed 48 h after IPTG induction, where the amount of coα-CGTase was 4,230 mg L^{-1} (515.9 mg g^{-1} DCW) and that of wtα-CGTase was 1,710 mg L^{-1} (204.8 mg g^{-1} DCW) (Fig. 2b), representing an approximate 252 % increase in the yield of the wtα-CGTase with codon optimization. No significant growth differences in the bacterial

Table 2 Codon preference in *E. Coli* and codon usage in the wild type and the optimized synthetic α-*cgt* gene

Amino acid	Codon	*f*[a]	wtα-*cgt*	coα-*cgt*	Amino acid	Codon	*f*[a]	wtα-*cgt*	coα-*cgt*
Ala	GCA	0.22	1	16	Leu	CUA	0.03	0	0
	GCC	0.25	22	13		CUC	0.10	8	0
	GCG	0.34	27	20		CUG	0.55	16	37
Arg	GCU	0.19	10	11		CUU	0.10	4	0
	AGA	0.04	0	0		UUA	0.15	1	0
	AGG	0.08	3	0		UUG	0.11	8	0
	CGA	0.05	2	0	Lys	AAA	0.76	13	27
	CGC	0.37	6	9		AAG	0.24	14	0
	CGG	0.08	4	0	Met	AUG	1.00	13	13
Asn	CGU	0.42	3	9	Phe	UUC	0.49	20	14
	AAC	0.61	33	28		UUU	0.51	13	19
Asp	AAU	0.39	21	26	Pro	CCA	0.20	1	1
	GAC	0.41	32	21		CCC	0.10	5	0
Cys	GAU	0.59	19	30		CCG	0.55	16	23
	UGC	0.57	0	0		CCU	0.16	2	0
Gln	UGU	0.43	0	0	Ser	AGC	0.27	15	13
	CAA	0.31	8	8		AGU	0.13	2	7
Glu	CAG	0.69	12	12		UCA	0.12	1	8
	GAA	0.70	11	19		UCC	0.17	15	7
Gly	GAG	0.30	8	0		UCG	0.13	11	6
	GGA	0.09	4	0		UCU	0.19	2	5
	GGC	0.40	43	41	Thr	ACA	0.30	5	0
	GGG	0.13	19	0		ACC	0.43	18	38
His	GGU	0.38	11	36		ACG	0.23	38	26
	CAC	0.48	3	6		ACU	0.21	3	0
Ile	CAU	0.52	8	5	Trp	UGG	1.00	13	13
	AUA	0.07	3	0	Tyr	UAC	0.47	17	16
	AUC	0.46	15	19		UAU	0.53	16	17
	AUU	0.47	19	18	Val	GUA	0.17	7	0
						GUC	0.20	13	12
						GUG	0.34	24	21
						GUU	0.29	6	17

[a] Relative frequency of each codon in *E. coli*

cultures were observed at OD_{600} after transfection of the plasmids.

coα-CGTase retains full biological activity compared with wtα-CGTase

Native α-CGTase can catalyze four reactions: cyclization reaction, coupling, disproportionation, and starch hydrolysis. Here the important reactions (cyclization reactions) were verified. During the induction phase, both the extracellular and intracellular α-CGTase activities were determined. After induction with IPTG for 48 h, the extracellular α-CGTase activity reached 55.3 U mL^{-1}, whereas the wtα-CGTase activity was 19.8 U mL^{-1} (Fig. 3a). However, wtα-CGTase and coα-CGTase were equally effective in the production of cyclodextrins in the reaction mixtures containing wt or coα-CGTase proteins (0.2 U mL^{-1} α-CGTase activity) and 5 % (w/v) soluble starch incubated at 40 °C (Fig. 3b). The production of α-CD was analyzed by HPLC every 10 h.

RNA folding energies

Some investigators have concluded that the upstream nucleotides of the transcript can significantly correlate with the stability of mRNA folding [31]. Using the "mfold" RNA folding program, the free energies of 5′ translated *pelB*-coα-*cgt* and *pelB*-wtα-*cgt*, representing positions −3 to +50 from the start site, were calculated. The calculations indicated that the 5′ regions of the above two genes were the same, with ΔG of −19.10 kcal/mol (Fig. 1a, b). The wt and coα-*cgt* genes were expressed under the control of the same peptide, and their mRNA folding energies were the same, thus being unlikely to contribute to the difference of the translation efficiency.

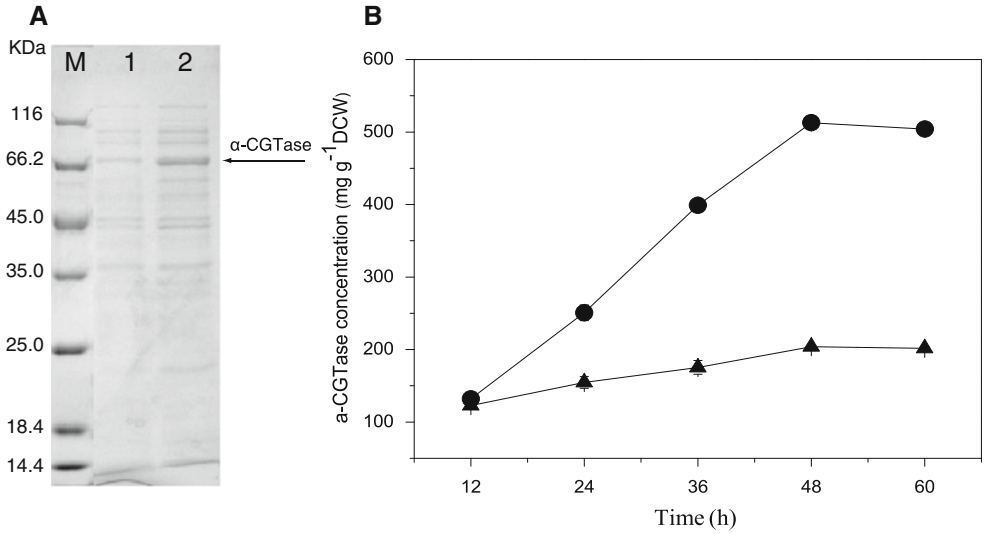

Fig. 2 Expression of α-CGTase in *E. coli* at 24 h postinduction. **a** Each *lane* was loaded with approximately 10 μL sample. The *arrow* shows the CGTase with molecular weight of approximately 72 kDa. M, Protein molecular weight standard; *1* Culture supernatant from *E. coli* BL21 (DE3) harboring pET-20b(+)-wtα-*cgt*; *2* Culture supernatant from *E. coli* BL21 (DE3) harboring pET-20b(+)-coα-*cgt*. **b** Time course of the total expression protein yields for wtα-CGTase (*triangles*) and coα-CGTase (*circles*) in *E. coli* BL21(DE3)

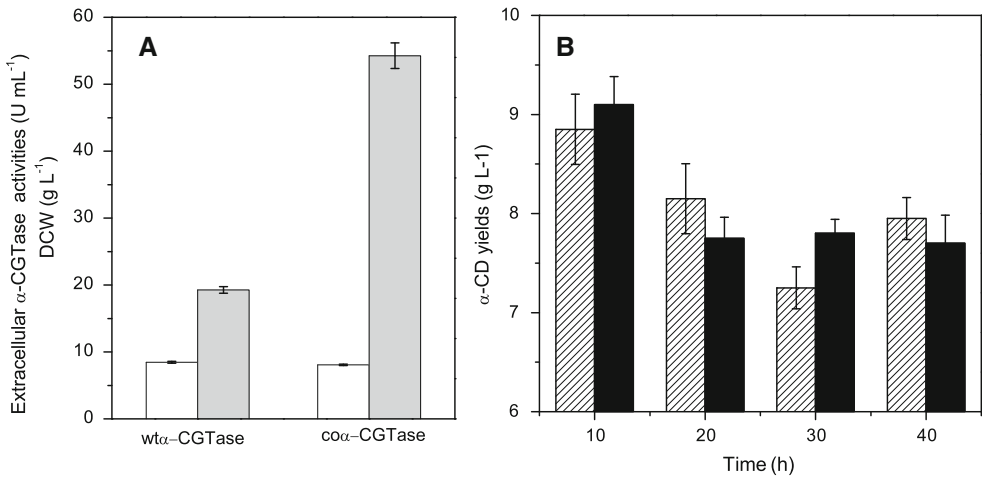

Fig. 3 Analysis of the enzyme activity and cyclodextrin-forming activity confirms that the wtα-CGTase and coα-CGTase proteins are identical. **a** Cell densities and extracellular α-CGTase activity of the two strains under the same culture conditions at 48 h after induction. The strains expressed either the wtα-*cgt* gene or the coα-*cgt* gene. *White bars* represent cell density, and *gray bars* represent extracellular α-CGTase activity. All experiments were evaluated under the same conditions and performed in triplicate. *Error bars* show the standard deviation. **b** α-CD formed during incubation of the wtα-CGTase and coα-CGTase. *Slash bars* represent wtα-CGTase, and *black bars* represent coα-CGTase

Secretion of α-CGTase in the batch fermentation process

Since, at the culture flask level, the α-CGTase activity increased significantly in *E. coli* with codon optimization for eventual industrial application, the large-scale expression of the enzyme by the two strains should also be compared. Therefore, the two strains were cultured with the same optimized batch protocol in 3-L fermentors. During the batch culture, the cell densities of the two strains were slightly different, and the OD_{600} reached about 20 (Fig. 4a, b). The highest extracellular α-CGTase activities appeared at 27 h after inoculation (Fig. 4a, b). Compared with the native enzyme gene expression (14.6 U mL^{-1}) (Fig. 4a) in *E. coli*, the secretion level of the codon-optimized gene (47.55 U mL^{-1}) (Fig. 4b) was increased

by 226 %. Again, the results indicated that the codon-optimized gene was improved in terms of expression level and was suited for industrial application.

Discussion

An alternative method to achieve high yield of α-CGTase has been urgently needed in the field of α-CD and AA-2G production. The specific aim of this work is to enhance the expression of α-CGTase in *E. coli* to obtain cost-effective production of α-CGTase for industrial applications.

In the process of heterologous expression, much work has been done to improve the secretion of heterologous proteins, including process optimization of host growth conditions and fermentation conditions, development of new promoters, new peptide of the vectors, and even the development of new host strains, organisms, and cell-free systems [9]. However, the topic of the codon bias of the gene has been skirted, although proteins that use more frequent codons have been shown to be highly expressed [23].

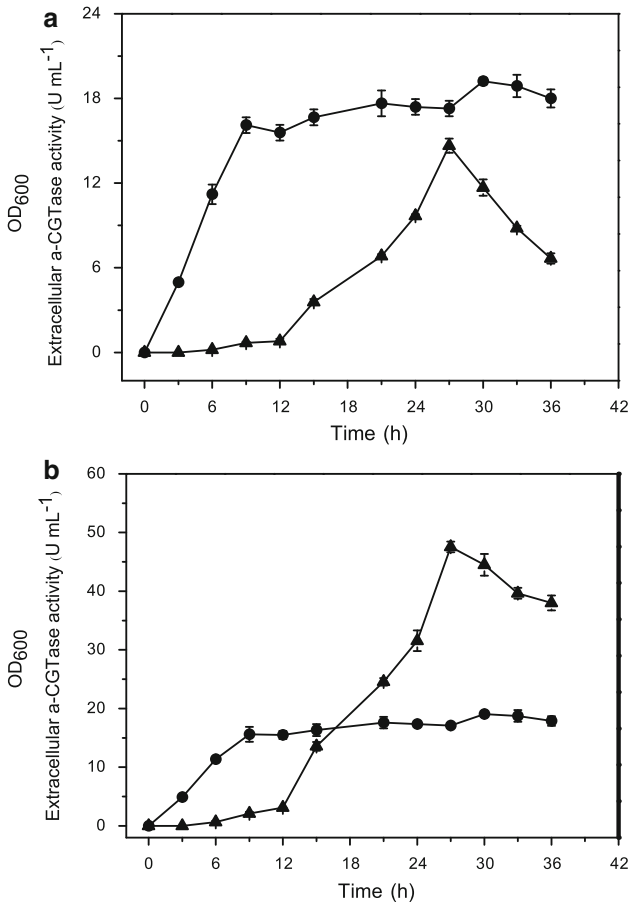

Fig. 4 Time course of extracellular α-CGTase activity and cell growth of the two strains in the same batch culture. Strains expressing either the wtα-*cgt* gene (**a**) or the coα-*cgt* gene (**b**) were cultured at 37 °C until OD_{600} reached 8.0. Gene expression was induced with 50 μM IPTG at 25 °C. *Black circles* represent cell density, and *black triangles* represent extracellular α-CGTase activity. All experiments were evaluated under the same conditions and performed in triplicate. *Error bars* show the standard deviation

Codon usage bias has evolved related to the regulation of expression levels. Firstly, high-usage-frequency codons, meaning high levels of the corresponding tRNAs in the host, will optimize translational efficiency [23]. Secondly, synonymous codon substitution of the gene, especially at the 5′-end, can impact the mRNA structure and stability, and thus affect translation initiation and elongation [30]. Various strategies have been used to minimize the bias for heterologous expression, as follows: (a) Supplying extra copies of rare tRNA genes to host cells [10]. However, this method can apply significant stress to host-cell metabolic processes. Besides, this method can resolve some, but not necessarily all, codon bias issues. (b) "Codon optimization" is the method that substitutes codons in a target sequence toward preferred high-frequency codons from the expression host, also called the "one amino acid–one codon" approach. This method has been successful for heterologous production of some proteins in certain host expression systems [32]. However, this may lead to tRNA traffic jams, or in some cases, the high levels of protein expressed may lead to the formation of insoluble products. (c) In the "codon harmonization" method a target gene sequence is recoded to "match" more closely the codon usage bias inherent in the native host when expressed in the heterologous host. This method was successfully applied to express several antigens in *E. coli* [6].

In this study, the systematic codon optimization method was introduced. The method is based on the "codon optimization" method. Therefore, the most preferred codons were used during codon optimization, which avoids usage of rare codons of the host cell. The second most preferred codons were used alternatively to prevent the probable depletion of tRNAs resulting from choosing consecutively the same codon for the frequently used amino acid. Furthermore, compared with the "one amino acid–one codon" approach, the systematic codon optimization strategy uses the unpreferred synonymous codons that may possibly be necessary when there is a need to pause translation.

In this study, the systematic strategy may have some advantages over the one amino acid–one codon method for two reasons: (a) In the synthesized α-*cgt* gene, there were replacements of several rare codons with those favored by *E. coli*, including for example, UUA(L) to CUG(L) and GUA(V) to GUU(V), avoiding the possibility that the ribosomes detach from the mRNA when encountering a rare codon. (b) The frequencies of other codons, such as AUU(I), GAU(D), GUU(V), and AGU(S), in the coα-*cgt*

gene were also balanced with the host preference, which may possibly be necessary when there is a need to pause translation, avoiding the "hungry codon syndrome" (which causes a ribosome progressing on the heterologous mRNA to stall at positions requiring a tRNA that is either depleted or largely uncharged with acyl-amino acid, resulting in an increased risk of translational error) [12].

Obviously, synthetic optimization of codon bias is a successful strategy to improve translation efficiency and enhance the production of α-CGTase. Regardless of the comparably higher yields of native α-CGTase reported in the literature, the synthetic codon-optimized gene resulted in protein levels that were 152 % greater than the amounts of the native gene. The increase in coα-CGTase over wtα-CGTase expression is most likely the result of the replacement of specific codons in the gene with those favored by the host, which ensures efficient translation and recycling of tRNA for the next round. Comparisons of codon-optimized genes with natural genes in identical systems at the protein expression level are frequent in recently published reports [2, 4, 10, 29]. The codon optimization methods vary widely, but all have replaced rare codons in the host with more frequently used codons. Expression of the well-adapted synthetic genes exceeds that of the native genes by 1-fold to 1,000-fold, representing levels suitable for industrial application. Synonymous codon substitutions have long been thought to be inconsequential, but evidence indicates that even a single synonymous codon substitution can have significant impact, not just on gene expression levels, but also on protein folding and enzyme activities [26]. The function of the two genes indicated that the well-adapted synthetic α-*cgt* gene retained the function of the native gene. The improvement of the extracellular enzyme activities was just the result of the higher expression level under the same culture conditions.

Some investigators have concluded that the first 40 nucleotides of the transcript can correlate significantly with the stability of the mRNA folding. However, the native and synthetic α-*cgt* gene were expressed under the control of the same peptide and their mRNA folding energies were the same, being unlikely to contribute to the difference in the translation efficiency.

In summary, an alternative method for obtaining high yield of α-CGTase has been urgently needed in the field of α-CD and AA-2G. The specific aim of this work is to improve expression of α-CGTase in *E. coli* to obtain cost-effective production of α-CGTase for industrial applications. Clearly, systematic optimization of codon bias is a successful strategy to improve translation efficiency and to enhance the production of α-CGTase. Regardless of the comparably higher yields of native α-CGTase published, heterologous α-CGTase expression was improved significantly by using codons that match those preferred by the host translational system, without altering protein function. Moreover, codon optimization has been a cost-effective way to increase the expression of α-CGTase in industry. With this established strain *E. coli* BL21(DE3) pET-20b(+)-coα-*cgt*, large-scale production of α-CGTase increased by 226 % (47.55 U mL^{-1}) in batch fermentation. Using fed-batch fermentation for high-cell-density cultivation, a much higher extracellular yield of enzyme would be achieved on an industrial scale.

Acknowledgments This work was supported by grants from the Major State Basic Research Development Program of China (973 Program, 2012CB720806), the National High Technology Research and Development Program of China (863 Program, SS2012A A023403, 2011AA100905), the Open Project Program of the Key Laboratory of Industrial Biotechnology, Ministry of Education, China (no. KLIB-KF201006), the National Natural Science Foundation of China (31000807), the Natural Science Foundation of Jiangsu Province (BK2011004), the Enterprise-university-research prospective program, Jiangsu Province (BY2009112), the Open Funding Project of the State Key Laboratory of Bioreactor Engineering, the Key Technologies R & D Program of Jiangsu Province, China (BE2011624), and the Fundamental Research Funds for the Central Universities (JUSRP211A29).

References

1. Angov E (2011) Codon usage: nature's roadmap to expression and folding of proteins. Biotechnol J 6:650–659
2. Angov E, Hillier CJ, Kincaid RL, Lyon JA (2008) Heterologous protein expression is enhanced by harmonizing the codon usage frequencies of the target gene with those of the expression host. PLoS ONE 3:e2189
3. Bar A, Krul CAM, Jonker D, de Vogel N (2004) Safety evaluation of an α-cyclodextrin glycosyltranferase preparation. Regul Toxicol Pharm 39:S47–S56
4. Burgess-Brown NA, Sharma S, Sobott F, Loenarz C, Oppermann U, Gileadi O (2008) Codon optimization can improve expression of human genes in *Escherichia coli*: a multi-gene study. Protein Expres Purif 59:94–102
5. Cheng J, Wu D, Chen S, Chen J, Wu J (2011) High-level extracellular production of α-cyclodextrin glycosyltransferase with recombinant *Escherichia coli* BL21 (DE3). J Agric Food Chem 59:3797–3802
6. Chowdhury DR, Angov E, Kariuki T, Kumar N (2009) A potent malaria transmission blocking vaccine based on codon harmonized full length Pfs48/45 expressed in *Escherichia coli*. PLoS ONE 4:e6352
7. Ding RR, Li ZF, Chen S, Wu D, Wu J, Chen J (2010) Enhanced secretion of recombinant α-cyclodextrin glucosyltransferase from *E. coli* by medium additives. Process Biochem 45:880–886
8. Go YH, Kim TK, Lee KW, Lee YH (2007) Functional characteristics of cyclodextrin glucanotransferase from alkalophilic *Bacillus sp* BL-31 highly specific for intermolecular transglycosylation of bioflavonoids. J Microbiol Biotechnol 17:1550–1553
9. Gustafsson C, Govindarajan S, Minshull J (2004) Codon bias and heterologous protein expression. Trends Biotechnol 22:346–353
10. Gvritishvili AG, Leung KW, Tombran-Tink J (2010) Codon preference optimization increases heterologous PEDF expression. PLoS ONE 5:e15056

11. Han RZ, Liu L, Li JH, Du GC, Chen J (2012) Functions, applications and production of 2-O-d-glucopyranosyl-l-ascorbic acid. Appl Microbiol Biotechnol 95:313–320
12. Huang YK, Chen YS, Mo DL, Cong PQ, He ZY (2012) Attenuated secretion of the thermostable xylanase xynB from *Pichia pastoris* using synthesized sequences optimized from the preferred codon usage in yeast. J Microbiol Biotechnol 22:316–325
13. Ikemura T (1985) Codon usage and tRNA content in unicellular and multicellular organisms. Mol Biol Evol 2:13–34
14. Ismail NF, Hamdan S, Mahadi NM, Murad AM, Rabu A, Bakar FD, Klappa P, Illias RM (2011) A mutant L-asparaginase II signal peptide improves the secretion of recombinant cyclodextrin glucanotransferase and the viability of *Escherichia coli*. Biotechnol Lett 33:999–1005
15. Jayaraj S, Reid R, Santi DV (2005) GeMS: an advanced software package for designing synthetic genes. Nucleic Acids Res 33:3011–3016
16. Kodumal SJ, Patel KG, Reid R, Menzella HG, Welch M, Santi DV (2004) Total synthesis of long DNA sequences: synthesis of a contiguous 32-kb polyketide synthase gene cluster. Proc Natl Acad Sci USA 101:15573–15578
17. Li ZF, Gu ZB, Wang M, Du GC, Wu J, Chen J (2010) Delayed supplementation of glycine enhances extracellular secretion of the recombinant α-cyclodextrin glycosyltransferase in *Escherichia coli*. Appl Microbiol Biotechnol 85:553–561
18. Li ZF, Li B, Gu ZB, Du GC, Wu J, Chen J (2010) Extracellular expression and biochemical characterization of α-cyclodextrin glycosyltransferase from *Paenibacillus macerans*. Carbohydr Res 345:886–892
19. Li ZF, Li B, Liu ZG, Wang M, Gu ZB, Du GC, Wu J, Chen J (2009) Calcium leads to further increase in glycine-enhanced extracellular secretion of recombinant α-cyclodextrin glycosyltransferase in *Escherichia coli*. J Agric Food Chem 57: 6231–6237
20. Li ZF, Wang M, Wang F, Gu ZB, Du GC, Wu J, Chen J (2007) γ-Cyclodextrin: a review on enzymatic production and applications. Appl Microbiol Biotechnol 77:245–255
21. Li ZF, Zhang JY, Sun Q, Wang M, Gu ZB, Du GC, Wu J, Chen J (2009) Mutations of lysine 47 in cyclodextrin glycosyltransferase from *Paenibacillus macerans* enhance β-cyclodextrin specificity. J Agric Food Chem 57:8386–8391
22. Li ZF, Zhang JY, Wang M, Gu ZB, Du GC, Li JK, Wu J, Chen J (2009) Mutations at subsite -3 in cyclodextrin glycosyltransferase from *Paenibacillus macerans* enhancing α-cyclodextrin specificity. Appl Microbiol Biotechnol 83:483–490
23. Lithwick G, Margalit H (2003) Hierarchy of sequence-dependent features associated with prokaryotic translation. Genome Res 13:2665–2673
24. Low KO, Muhammad Mahadi N, Abdul Rahim R, Rabu A, Abu Bakar FD, Murad AM, Md Illias R (2011) An effective extracellular protein secretion by an ABC transporter system in *Escherichia coli*: statistical modeling and optimization of cyclodextrin glucanotransferase secretory production. J Ind Microbiol Biot 38:1587–1597
25. Ong RM, Goh KM, Mahadi NM, Hassan OH, Rahman RNZRA, Illias RM (2008) Cloning, extracellular expression and characterization of a predominant β-CGTase from *Bacillus sp.* G1 in *E. coli*. J Ind Microbiol Biotechnol 35:1705–1714
26. Plotkin JB, Kudla G (2011) Synonymous but not the same: the causes and consequences of codon bias. Nat Rev Genet 12:32–42
27. Qi QS, Zimmermann W (2005) Cyclodextrin glucanotransferase: from gene to applications. Appl Microbiol Biotechnol 66:475–485
28. Kato T, Horikoshi K (1986) Cloning and expression of the *Bacillus subtilis* NO.313 γ-cyclodextrin forming CGTase gene in *Escherichia coli*. Agric Biol Chem 50:2161–2162
29. Tang LX, Jiang RX, Zheng K, Zhu XC (2011) Enhancing the recombinant protein expression of halohydrin dehalogenase HheA in *Escherichia coli* by applying a codon optimization strategy. Enzyme Microb Technol 49:395–401
30. Tuller T, Carmi A, Vestsigian K, Navon S, Dorfan Y, Zaborske J, Pan T, Dahan O, Furman I, Pilpel Y (2010) An evolutionarily conserved mechanism for controlling the efficiency of protein translation. Cell 141:344–354
31. Tuller T, Waldman YY, Kupiec M, Ruppin E (2010) Translation efficiency is determined by both codon bias and folding energy. Proc Natl Acad Sci USA 107:3645–3650
32. Yang Y, Malten M, Grote A, Jahn D, Deckwer WD (2007) Codon optimized *Thermobifida fusca* hydrolase secreted by *Bacillus megaterium*. Biotechnol Bioeng 96:780–794
33. Zehentgruber D, Lundemo P, Svensson D, Adlercreutz P (2011) Substrate complexation and aggregation influence the cyclodextrin glycosyltransferase (CGTase) catalyzed synthesis of alkyl glycosides. J Biotechnol 155:232–235
34. Zhang ZC, Li JH, Liu L, Sun J, Hua ZZ, Du GC, Chen J (2011) Enzymatic transformation of 2-O-α-d-glucopyranosyl-l-ascorbic acid (AA-2G) by immobilized α-cyclodextrin glucanotransferase from recombinant *Escherichia coli*. J Mol Catal B Enzym 68: 223–229
35. Zhang ZC, Li JH, Liu L, Sun J, Hua ZZ, Du GC, Chen J (2011) Enzymatic transformation of 2-O-α-D-glucopyranosyl-L-ascorbic acid by α-cyclodextrin glucanotransferase from recombinant *Escherichia coli*. Biotechnol Bioproc E 16:107–113
36. Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 31:3406–3415