Biochemical and Biophysical Research Communications 376 (2008) 647–652

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

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

Rare codon priority and its position specificity at the 5′ of the gene modulates heterologous protein expression in *Escherichia coli*

Chinnathambi Thangadurai, Pichaimuthu Suthakaran, Pankaj Barfal, Balaiah Anandaraj, Satya Narayan Pradhan, Harith Kamil Boneya, Subramanian Ramalingam, Vadivel Murugan *

Genetic Engineering Unit, Centre for Biotechnology, Sadar Patel Road, Anna University, Guindy, Chennai-600 025, India

ARTICLE INFO

Article history:
Received 24 August 2008
Available online 16 September 2008

Keywords:
Rare codon
Streptokinase
AGG codon
E.coli
Over-expression

ABSTRACT

Rare codons and their effects in heterologous protein expression in *Escherichia coli* were addressed by many investigators. Here, we propose that not all rare codons of a foreign gene have negative effect but selective codon among them and its specific position in the downstream of the start codon modulates the expression. In our study, streptokinase (47 kDa), encoded by *skc* gene of *Streptococcus equisimilis* was expressed in *E.coli*. The analysis of relative codon frequency of *skc* gene in *E.coli* reveals the presence of 30% of rare codons in it. Nevertheless, *E.coli* managed to yield over-expression of this target protein. To explore the codon bias in expression, we have introduced the selective AGG codon at different positions of *skc* gene such as +2,+3,+5,+8,+9 and +11. The results revealed that at +2 position "AGG" aided over-expression while shifting to +3 and +5 positions it rendered nil expression. In contrary, shifting of AGG codon to later positions like +9 and +11 the inhibitory effect was reversed and resulted in over-expression. The effect of 'AGG' rare codon was further studied in GFP expression. In conclusion, besides the choice of rare codons, their precise positions in the foreign gene dictate the level of protein expression.

© 2008 Elsevier Inc. All rights reserved.

The heterologous protein expression has brought radical changes in recombinant protein production. The large scale production of therapeutic, diagnostic and industrially important proteins or enzymes greatly depend on the protein expression strategies. The prime objective of heterologous protein expression is to obtain higher yield of target proteins [1]. But not all proteins are successfully expressed in *Escherichia coli*. Many factors such as vector and host, strength of promoters, inducers concentration, media composition etc are known to have influence in protein expression [2]. Despite the fact, it is difficult to draw a generalized conclusion which is universally applicable for the expression of any protein. Expression ability of each gene is unique in heterologous host. Even though all the above variables were made as constant, in some cases the target protein expression was not achieved. In those cases, the nucleotide sequences or codon contents of the target genes were identified to play a role in deciding the expression status [3–5]. The codons that lie immediately downstream of the initiation codons play a decisive role in target protein expression [6–11]. The presence of "AAA" codon at +2 position of the gene was reported to maximize the expression [12] whereas "NGG" codon at +2 position was found to have negative effect [13–15].

In addition to the preference for certain codons in early position of the gene, the codon usage that occurs in the gene in relation with the heterologous host was found to influence the expression level predominantly [16]. In this context the codon composition of the heterologous gene may be optimized for the expression host. For instance, the codon optimization of the archaeal gene was found to increase the protein yield in *E.coli* [6]. It was also noticed that the repeated CAT codons downstream of the NGG codon has synergistically enhanced the protein expression in *E.coli* [17].

In our study, we have over-expressed the Streptokinase (SK) in *E.coli.* It is 47 kDa protein encoded by *skc* gene of *Streptococcus equisimilis* which is a gram positive bacterium. Though *skc* gene is known to have many rare codons in its composition, over-expression was achieved instead of having negative effect. This gives an insight that although some codons (rare) are less frequently used in *E.coli*, all of them are not detrimental in target protein expression. To understand this, we have introduced the known inhibitory "AGG" codon in the early positions of *skc* gene which are otherwise not present in the *skc* gene. The substitution-driven effect of 'AGG' codon at early positions of Streptokinase (SK) and Green Fluorescent Protein (GFP) was experimentally analyzed. From our study it is clear that priority of AGG codon among the rare codons as well as its position at the proximity of the initiation codon plays a significant role in target protein expression.

* Corresponding author. Fax: +91 44 22350229.
E-mail address: vmurugan@annauniv.edu (V. Murugan).

0006-291X/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2008.09.024

Materials and methods

Plasmid, genes and reagents. The expression vector pRSETB and BL21 (DE3) were acquired from Invitrogen technologies (Cat. No. V351-20). The restriction enzymes NdeI (Cat. No. #R0111L) and EcoRI (Cat. No. #R0101L) were acquired from New England Biolabs. The PCR reaction was performed using the Proof Reading Polymerase (Vent DNA Polymerase—Cat. No. #M0254S) obtained from New England Biolabs. The Streptokinase gene (*Skc*) [GeneBank Accession No. K02986] was obtained as a gift from Dr. K.J. Mukherjee, Centre for Biotechnology, Jawaharlal Nehru University. The GFP gene (Cat. No. 107387) was procured from Bangalore Genie, Bangalore, India. The Gene bank accession number of GFP is EU048697. LB medium used for culturing of *E.coli* was obtained from Hi Media, Laboratories, Mumbai, India. Primers synthesis and DNA sequencing were done at Microsynth Corporation, Switzerland.

Cloning of SK and GFP genes and site-directed mutagenesis. The vector used for this study is pRSETB which is a T7 based expression vector. The NdeI and EcoRI restriction sites of this vector were exploited through out the cloning work of this study. All the constructs were made only in between these two sites.

PCR based site-directed mutagenesis was carried out for making the codon substitution. Towards this, primers were designed and synthesized with mismatch nucleotides at the necessary positions of the genes (Tables 1 and 2).

The gene, skc encoding streptokinase was cloned in the pRSETB vector in between NdeI and EcoRI restriction sites. The recombinant construct is named as pRSET-SK. The SKM1, SKM2 and SKM3 constructs were generated by substitution of AGG codon at +2, +3 and +5 positions of the SK gene, respectively. The AGG codon was then shifted from the previous positions to +8, +9 and +11 by inserting a common sequence, $(CAT)_6$ immediately next to the initiation codon. The resultant constructs were named as SKM1-His, SKM2-His and SKM3-His, respectively. Similarly, the SKM1-Tyr and SKM3-Tyr were constructed by inserting $(TAC)_6$ sequence immediately next to initiation codon of the SKM1 and SKM 3 mutants (Fig. 1). The GFP M2, M3and M4 construct were generated by inserting the AGG codon at the +2, +3 and +5 position of the GFP gene sequence (Fig. 1). The mutant constructs were confirmed after DNA sequencing from Microsynth Corporation, Switzerland.

Relative codon frequency analysis. The relative codon frequency analysis for the SK gene was done using graphical codon usage analyzer [18] using the option 'each triplet position vs. usage table' of the program. This tool is a component of ExPasy analysis tool which is accessible at http://www.gcua.schoedl.de/.

Expression analysis of the recombinant constructs. The constructs were transformed into BL21 (DE3) as per Mandel and Higa protocol [19] and the transformants were grown in LB medium till 0.6 OD at A_{600}. The cultures were induced with 1 mM IPTG at 37 °C for 3 h. The expression profile of the SK and GFP samples were analyzed by resolving in 10% SDS–PAGE as per Laemmli protocol [20]. Further, the gels were stained with Coomassie Brilliant Blue G-250.

Results

Analysis of relative codon frequency and SK expression

The relative codon frequency analysis has revealed that the Streptokinase gene (*skc*) of *S. equisimilis* is harboring 30% of rare codons for *E.coli* (Fig. 2) host. Among them, some of the codons were used less than 20% and 10% by the host translational machinery which is shown in blue and red colored bars respectively in the Fig. 2. Surprisingly, the gene was known to have consecutive rare codons even at early positions like +3, +4 and +5 (GCT, GGA, CCT) and several positions in the later part of the gene. Despite the presence of 30% rare codons, the gene favored over-expression of SK in *E.coli* BL21 (DE3) as shown in SDS–PAGE (Fig. 3A) which was further confirmed by Western blotting (Fig. 3B).

Mutation of skc gene by 'AGG' codon Substitution

The original codons of the *skc* gene at +2, +3 and +5 positions were replaced with the 'AGG' codon by site-directed mutagenesis. These constructs were designated as SKM1, SKM2 and SKM3 respectively. The effect of the 'AGG' codon on target protein expression is shown in Fig. 4A (*lane 3, 6 & 8*). The expression analysis revealed that the 'AGG' codon at +2 position aided over-expression, whereas at +3 and +5 positions it exhibited a negative effect in expression (nil expression).

Cloning and expression of SKM1-His, SKM2-His, SKM3-His

To understand the impact of early occurrence of 'AGG' codon on target protein expression we have shifted the 'AGG' codon to later positions of the gene. To achieve this, we have inserted $(CAT)_6$ in between the initiation and the +2 codons which in turn increased the physical distance of the 'AGG' codon from the initiation site. The insertion of $(CAT)_6$ in all three constructs namely SKM1, SKM2 and SKM3 has shifted the 'AGG' codon to +8, +9 and +11 positions (SKM1-His, SKM2-His and SKM3-His), respectively. The

Table 1
List of Primers

Sl. No	Name of Primer	Primer Sequence
1.	SK F1	5′ GGGATTC*CATATG*ATTGCTGGACCTGAG 3′
2.	SKM F1	5′ GGGATTC*CATATG*AGGGCTGGACCTGAG 3′
3.	SKM F2	5′ GGGATTC*CATATG*ATTAGGGGACCTGAG 3′
4.	SKM F3	5′ GGGATTC*CATATG*ATTGCTGGAAGGGAGTGGCTGCTA 3′
5.	SKM1 His-F	5′GGGATTC*CATATG*CATCATCATCATCATCATAGGGCTGGACCTGAG 3′
6.	SKM1 Tyr-F	5′GGGATTC*CATATG*TACTACTACTACTACTACAGGGCTGGACCTGAG 3′
7.	SKM2 His-F	5′GGGATTC*CATATG*CATCATCATCATCATCATATTAGGGGACCTGAG 3′
8.	SKM3 His-F	5′GGGATTC*CATATG*CATCATCATCATCATCATATTGCTGGAAGGGAGT GGCTGCTA 3′
9.	SKM3 Tyr-F	5′GGGATTC*CATATG*TACTACTACTACTACTACATTGCTGGAAGGGAGT GGCTGCTA 3′
10.	GFP F1	5′ GGGATTC*CATATG*GTGAGCAAGGGCGAG 3′
11.	GFPM F2	5′ GGGATTC*CATATG*AGGAGCAAGGGCGAG 3′
12.	GFPM F3	5′ GGGATTC*CATATG*GTGAGGAAGGGCGAGGAG 3′
13.	GFPM F4	5′ GGGATTC*CATATG*GTGAGCAAGAGGGAGGAGCTGTTC 3′
14.	SK R1*	5′ CCG*GAATTC*TTATTTGTCTTTAGG 3′
15.	GFP R1**	5′ CCG*GAATTC*TCACTCGTCCATGCCGAG 3′

The table lists the primers that were used for the sub-cloning and site directed mutagenesis of the SK and GFP in pRSETB vector. The restriction sites are italicized in each primer sequence. Only one reverse primer for SK indicated as * that was used in the combination with all the forward primers. Similarly, only one reverse primer for GFP was used in this study which is indicated as **. F represents forward primer and R represents reverse primer.

Download English Version:

https://daneshyari.com/en/article/1934414

Download Persian Version:

https://daneshyari.com/article/1934414

Daneshyari.com