Journal of Materials Science and Chemical Engineering, 2015, 3, 13-18
Published Online June 2015 in SciRes. http://www.scirp.org/journal/msce
http://dx.doi.org/10.4236/msce.2015.36003

Influence of Insertion of the Last Sense Codon on Expression Efficiency of Green Fluorescent Protein Gene in *Escherichia coli*

Xiaoliang Hao, Shota Inoue, Masahide Ishikawa

Department of Applied Chemistry, Graduate School of Engineering, Saitama Institute of Technology, 1690 Fusaiji, Fukaya, Saitama 369-0293, Japan
Email: a2002ldb@sit.ac.jp

Received April 2015

Abstract

We studied the relationship between insertion of the last sense codon (*i.e.*, the codon preceding the stop codon) and the efficiency of gene expression. We inserted 64 kinds of last sense codon at the 5' end of the stop codon of the green fluorescent protein (GFP) gene and introduced the modified GFP genes into *Escherichia coli* (*E. coli*). Measuring the fluorescence intensity of the GFP produced in *E. coli* showed that the last sense codon influenced GFP gene expression and when CCG was inserted as the last sense codon, fluorescence intensity of *E. coli* was increased to 2.09 fold. On the other hand, insertion of CUA caused decrease of fluorescence intensity to 0.33 fold. We hope that our findings, which may be applicable to gene engineering, will be useful for further studies of protein expression.

Keywords

Last Sense Codon, Gene Expression, Green Fluorescent Protein, Fluorescence Intensity

1. Introduction

Protein expression in *Escherichia coli* (*E. coli*) has been the most popular means of producing recombinant proteins. *E. coli* is a very useful host that offers efficient culturing efficiency, easy manipulation and low cost media. People have produced a wide variety of different types of proteins by *E. coli* [1]-[8]. Although there are many proteins are useful for the humans, these proteins can only rarely be obtained in sufficiently large quantities from their natural sources. Therefore, such proteins are often produced in *E. coli* via gene engineering.

The expression efficiency of genes in *E. coli* is very important in genetic engineering, and the codon composition of coding sequences plays an important role in the regulation of gene expression. To master the relationship between the gene sequence and protein expression is helpful for understanding principles of gene expression and effectively controlling the production of protein. The researches about codon, protein, and gene expression have attracted wide attention of people and made great achievements, for example some models had been constructed [9]-[11], some new methods had been detected and applied to research [12]-[14].

How to cite this paper: Hao, X.L., Inoue, S. and Ishikawa, M. (2015) Influence of Insertion of the Last Sense Codon on Expression Efficiency of Green Fluorescent Protein Gene in *Escherichia coli*. *Journal of Materials Science and Chemical Engineering*, **3**, 13-18. http://dx.doi.org/10.4236/msce.2015.36003

But there are still some issues have not been solved. For example, it is known that the sequences around the initiation codon influence gene expression in *E. coli* [15]. However, the influence of the sequences around the stop codon has not been clarified. So we used green fluorescent protein (GFP) gene for our study. GFP was discovered as a companion protein to aequorin in 1962, the famous chemiluminescent protein from Aequorea jellyfish [16]. GFP is composed of 238 amino acids and its molecular weight of roughly is 27 kDa [17]. The discovery of GFP led to a new revolution in molecular biology, medicine, and cell biology, it was particularly useful due to its stability and the fact that its chromophore was formed in an autocatalytic cyclization that does not require a cofactor, now has been a commonly used tool. The most common use of GFP is its remains as a reporter of gene transcription or a protein of interest in the determination of gene expression and protein localization [16], [18]-[23]. In just recent years, GFP has valued from obscurity to become one of the most widely studied and exploited proteins in biochemistry and cell biology [24].

In this study, we inserted 64 kinds of last sense codons (*i.e.*, the codon preceding the stop codon) at the 5' end of the stop codon of the GFP gene, tried to solve this problem by research. Because GFP is easy to be quantified and detected, so we compared the expression quantity of GFP, so we used the fluorescence of GFP to judge the influence of last sense codon on expression efficiency [25] [26]. This text aimed to research the correlativity between last sense codon and expression efficiency of protein in *E. coli* from the perspective of genetic engineering, probe how to gain more and more useful proteins to serve people.

2. Materials and Methods

2.1. Bacterial Strains and Plasmids

E. coli JM109 was used as the host strain, and the pKK223-3 plasmid was used as the expression vector for cloning. Primers for polymerase chain reaction (PCR) were purchased from Sigama-Aldrich Japan (Tokyo, Japan). Restriction enzymes (*Eco*RI, *Hin*dIII), a PCR amplification kit and T4 DNA ligase were purchased from TaKaRa Bio (Ohtsu, Japan).

2.2. Insertion of Last Sense Codon to GFP Gene

We inserted 64 kinds of last sense codons into GFP genes by PCR. The used primers were 5'Primer: CCCGAATTCTTTAACTTTAGGAAACACAATTCATGAGTAAAGGAGAAGAACTT and 3'Primer: CTCTAAGCTTTTANNNTTTGTATAGTTCATCCA.

The part with the underline was the stop codon, NNN was the 64 kinds of last sense codons. For example, when NNN was CGG, according to the rule of complementarity principle, it was understood that the last codon was "CCG". We constructed the plasmids by introducing modified GFP genes into pKK223-3 at a *Eco*RI site and a *Hin*dIII site. PCR was performed at 94˚C for 10 min, followed by 25 cycles of 94˚C for 45 s, 48˚C for 1 min and 68˚C for 45 s. After PCR, the plasmids were constructed by introduction of the modified GFP genes into expression vector. *E. coli* JM109 was transformed with the ligation mixture, the resulting clones were analyzed by sequencing.

2.3. Measurement of Fluorescence Intensity of GFP in *E. coli*

We compared the fluorescence intensity among the 64 GFP genes with different last sense codons. Cells were cultivated in LB medium supplemented with 0.1 mg/mL ampicillin and 40 μM isopropyl-*β*-D-thiogalactopyranoside (IPTG) at 37˚C for 18 h. We measured the absorbance at 600 nm of the *E. coli* culture diluted with LB medium, and the fluorescence intensity of the same culture at 508 nm excited at 396 nm using the Gemini fluorescent microplate reader (Nihon Molecular Devise corp., Tokyo, Japan). We then compared the expression efficiency of the modified GFP genes by measuring fluorescence intensity/absorbance at 600 nm, because the fluorescence intensity increased in proportion to the absorbance at this wavelength.

2.4. SDS-PAGE Analysis of Extracts of *E. coli* Having Modified GFP Genes

After the fluorescence intensities of acquired modified GFP genes were known, we wanted to compare the expression quantities of proteins, and analyzed the influences of the modified GFP genes on protein expression efficiency, therefore sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) experiment was

done. *E. coli* JM109 was transformed by the expression vector including GFP gene, and cultured in the LB medium of 20 mL under the presence of 40 μM IPTG at 37˚C for 18 h. DNase was added into the solution and the solution was incubated at 37˚C for 1 h to remove the remained DNA of the solution. The insoluble parts were separated from the solution by centrifugation (4˚C, 13,000 rpm, 10 min) and soluble proteins were analyzed by SDS-PAGE.

3. Results

As shown in **Figure 1**, the fluorescence intensity at 508 nm excited by 396 nm of *E. coli* culture was measured and compared the expression efficiency of the modified GFP gene as fluorescence intensity/absorbance at 600 nm, because the fluorescence intensity increased in proportion to the absorbance at this wavelength. And the relative fluorescence intensity (RFI) was normalized to that of wild type GFP. Error bars are based on the results of 3 independent experiments. The strongest relative fluorescence intensity (2.09 fold) was observed with a CCG last sense codon, while the lowest relative fluorescence intensity (0.33 fold) was observed with a CUA last sense codon. The relative fluorescence intensities of the other 62 codons were 0.7 - 1.5 fold.

As shown in **Table 1**, the relative fluorescence intensity differed between proteins expressed by genes with a modified last sense codon normalized to that of the wild type GFP gene. For example, for phenylalanine (coded

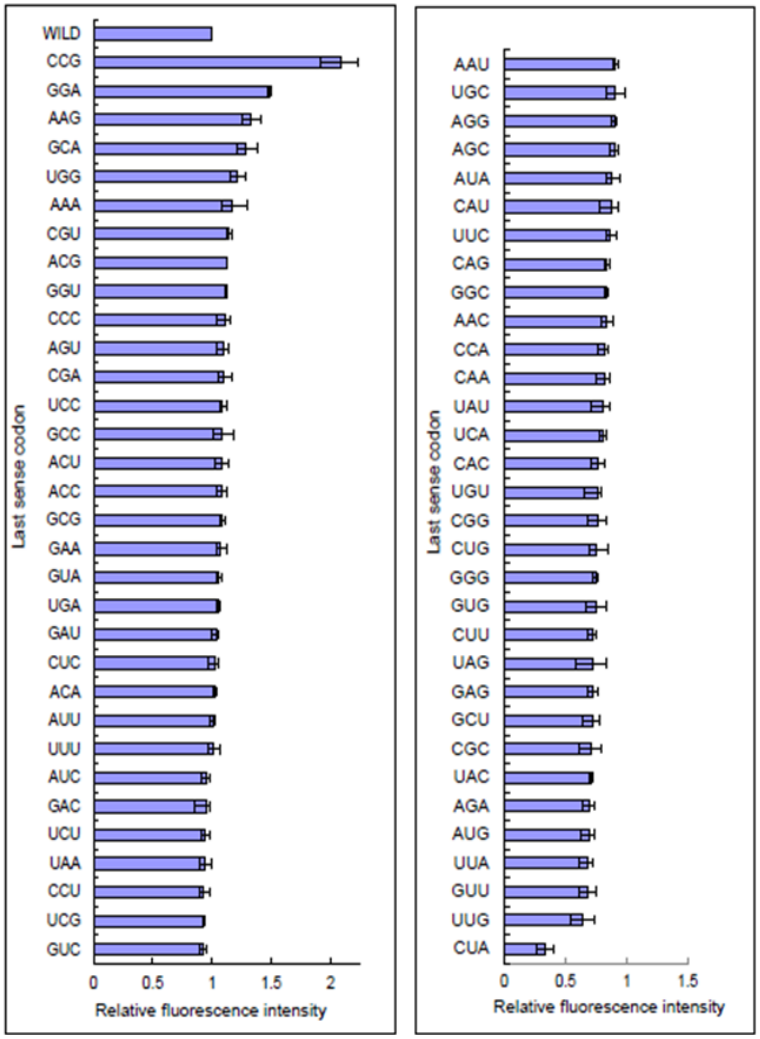

Figure 1. RFI of green fluorescent protein (GFP) expressed by genes with a insertion of last sense codon in *E. coli*. RFI was normalized with that of the wild type GFP gene.

Table 1. Relationship between the last sense codon and RFI.

Amino acid	Codon	RFI	Amino acid	Codon	RFI	Amino acid	Codon	RFI	Amino acid	Codon	RFI
Phe	UUU	±	Ser	UCU	±	Tyr	UAU	±	Cys	UGU	−
	UUC	±		UCC	±		UAC	−		UGC	±
Leu	UUA	−		UCA	±	STOP	UAA	±	STOP	UGA	±
	UUG	−		UCG	±		UAG	−	Trp	UGG	+
Leu	CUU	−	Pro	CCU	±	His	CAU	±	Arg	CGU	±
	CUC	±		CCC	±		CAC	−		CGC	−
	CUA	−−		CCA	±	Gln	CAA	±		CGA	±
	CUG	−		CCG	++		CAG	±		CGG	−
Ile	AUU	±	Thr	ACU	±	Asn	AAU	±	Ser	AGU	±
	AUC	±		ACC	±		AAC	±		AGC	±
Met	AUA	±		ACA	±	Lys	AAA	±	Arg	AGA	−
	AUG	−		ACG	±		AAG	+		AGG	±
Val	GUU	−	Ala	GCU	−	Asp	GAU	±	Gly	GGU	±
	GUC	±		GCC	±		GAC	±		GGC	±
	GUA	±		GCA	+	Glu	GAA	±		GGA	+
	GUG	−		GCG	±		GAG	−		GGG	−

Comparison of RFI: ++, >1.5; +, 1.2 - 1.5; ±, 0.8 - 1.2; −, 0.5 - 0.8; − −, <0.5.

by UUU and UUC), the relative fluorescence intensity was close, indicating differences in the quantity of GFP produced, but not variation in the amino acid sequences. Similar tendencies were seen for the other amino acids.

However, for most of the amino acids, the differences were clear; for example, the relative fluorescence intensities were different among the six last sense codons for leucine (UUA, UUG, UCU, UCC, UCA, and UCG).

Our experiments revealed that last sense codon influenced the expression efficiency of the GFP gene in *E. coli*; however, the relationship between the last sense codon and the gene expression efficiency of the GFP gene was not clear.

To confirm that the relative fluorescence intensity was related to the expression efficiency of the GFP gene, we performed SDS-PAGE of the extracts of each *E. coli* variant (**Figure 2**). A CCG last sense codon showed the strongest relative fluorescence intensity, and the band around 27 kDa derived from GFP was very thick. On the other hand, a CUA last sense codon showed the lowest relative fluorescence intensity, and the band around 27 kDa was very thin. Therefore, the bands around 27 kDa derived from GFP were related to the values of fluorescence intensity.

4. Discussion

Insertion of the last sense codon influenced the gene expression of GFP. Although the amino acid was the same, to our surprise, insertion the last sense codon changed the quantity of gene expression. For example, as both CTC and CTA are codons for leucine, the effect of the last sense codon in the GFP gene did not change the C-terminus amino acid of GFP. A similar phenomenon was seen in the other amino acids, but no clear relationship was determined.

Codon usage frequency has always been considered in the research and production of protein expression. When we compared our research results with the usage frequency of the codon commodities, we found that some rules were the same; for example, the CCG last sense codon, had high usage frequency and the lowest relative fluorescence intensity. Other last sense codons, such as AGA and GGG, showed similar trends between relative fluorescence intensity and usage frequency. However, great differences were also noted between the relative fluorescence intensity and usage frequency for other codons. For example, the GGA last sense codon had high relative fluorescence intensity, but low usage frequency. Furthermore, the relative fluorescence intensities of the AUA, and CGA last sense codon were almost same as that of the wild type, although they, like CUA,

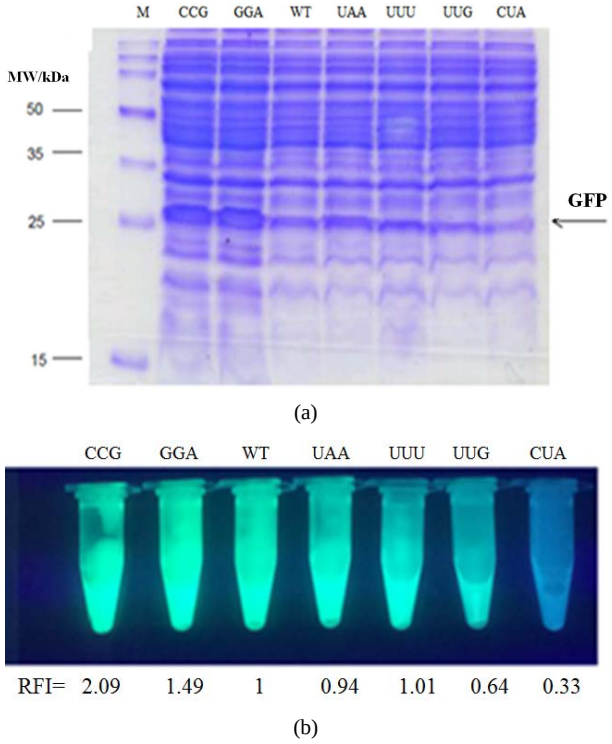

Figure 2. (a) Results of SDS-PAGE of the extracts of *E. coli* transformed by the expression vectors with the different last sense codon in the GFP gene. M: size marker; WT: wild type. Arrow shows the molecular weight of GFP (27 kDa); (b) Photograph of the extracts of *E. coli* having the same last sense codon as (a) on transilluminator.

are rare codons.

SDS-PAGE confirmed that relative fluorescence intensity was associated with the quantity of synthesized protein. Since CCG had the highest quantity of synthesized protein, it could be used as the last sense codon in gene engineering to produce useful proteins. Our findings will be useful for further studies of protein expression and may be applicable to gene engineering.

Acknowledgements

The authors are grateful for financial support by the "Nano Project" for Private Universities from Ministry of Education, Culture, Sports, Science, and Technology of Japan.

References

[1] Tolia, N.H. and Joshua-Tor, L. (2006) Strategies for Protein Coexpression in *Escherichia coli*. *Nature Methods*, **3**, 55-64. http://dx.doi.org/10.1038/nmeth0106-55

[2] Rosano, G.L. and Ceccarelli, E.A. (2014) Recombinant Protein Expression in *Escherichia coli*, Advances and Challenges. *Frontiers in Microbiology*, **5**, 1-17. http://dx.doi.org/10.3389/fmicb.2014.00172

[3] Hannig, G., and Makrides, S.C. (1998) Strategies for Optimizing Heterologous Protein Expression in *Escherichia coli*. *Trends in Biotech*, **16**, 54-60.

[4] Baneyx, F. (1999) Recombinant Protein Expression in *Escherichia coli*. *Current Opinion in Biotechnology*, **10**, 411-421.

[5] Panda, A.K. (2003) Bioprocessing of Therapeutic Proteins from the Inclusion Bodies of *Escherichia coli*. *Adv Biochem Eng Biotechnol*, **85**, 43-93.

[6] Palomares, L.A., Estrada-Mondaca, S. and Ramirez, O.T. (2004) Production of Recombinant Proteins, Challenges and Solutions. *Methods in Molecular Biology*, ch.2, 15-52.

[7] Mergulhao, F.J., Summers, D.K. and Monteiro, G.A. (2005) Recombinant Protein Secretion in *Escherichia coli*. *Biotechnology Advances*, **23**, 177-202. http://dx.doi.org/10.1016/j.biotechadv.2004.11.003

[8] Yin, J., Li, G., Ren, X. and Herrler, G. (2007) Select What You Need: A Comparative Evaluation of the Advantages and Limitations of Frequently Used Expression Systems for Foreign Genes. *Journal of Bacteriology*, **127**, 335-347. http://dx.doi.org/10.1016/j.jbiotec.2006.07.012

[9] Gilchrist, M.A. and Wagner, A. (2006) A Model of Protein Translation Including Codon Bias, Nonsense Errors, and Ribosome Recycling. *Journal of Theoretical Biology*, **239**, 417-434. http://dx.doi.org/10.1016/j.jtbi.2005.08.007

[10] Kosiol, C., Holmes, I. and Goldman, N. (2007) An Empirical Codon Model for Protein Sequence Evolution. *Molecular Biology and Evolution*, **24**, 1464-1479. http://dx.doi.org/10.1093/molbev/msm064

[11] Jukes, T.H. and Cantor, C.R. (1969) Evolution of Protein Molecules: Mammalian Protein Metabolism. 3. Academic Press, New York, 21-132. http://dx.doi.org/10.1016/B978-1-4832-3211-9.50009-7

[12] Kimura, M. (1980) A Simple Method for Estimating Evolutionary Rates of Base Substitutions through Comparative Studies of Nucleotide Sequences. *Journal of Molecular Evolution*, **16**, 111-120. http://dx.doi.org/10.1007/BF01731581

[13] Bielawski, J.P. and Yang, Z.H. (2004) A Maximum Likelihood Method for Detecting Functional Divergence at Individual Codon Sites, with Application to Gene Family Evolution. *Journal of Molecular Evolution*, **59**, 121-132. http://dx.doi.org/10.1007/s00239-004-2597-8

[14] Suzek, B.E., Ermolaeva, M.D., Schreiber, M. and Salzberg, S.L. (2001) A Probabilistic Method for Identifying Start Codons in Bacterial Genomes. *Bioinformatics*, **17**, 1123-1130. http://dx.doi.org/10.1093/bioinformatics/17.12.1123

[15] Sato, T., Terabe, M., Watanabe, H., Gojobori, T., Hori-Takemoto, C. and Miura, K. (2001) Codon and Base Biases after the Initiation Codon of the Open Reading Frames in the *Escherichia coli* Genome and Their Influence on the Translation Efficiency. *Journal of Biochemistry*, **129**, 851-860. http://dx.doi.org/10.1093/oxfordjournals.jbchem.a002929

[16] Shimomura, O., Johnson, F.H. and Saiga, Y. (1962) Extraction, Purification, and Properties of Aequorin, a Bioluminescent Protein from the Luminous Hydromedusan, Aequorea. *Journal of Cellular and Comparative Physiology*, **59**, 223-239. http://dx.doi.org/10.1002/jcp.1030590302

[17] Prasher, D.C., Eckenrode, V.K., Ward, W.W., Prendergast, F.G. and Cormier, M.J. (1992) Primary Structure of the *Aequorea victoria* Green Fluorescent Protein. *Gene*, **111**, 229-233. http://dx.doi.org/10.1016/0378-1119(92)90691-H

[18] Miyawaki, A. and Tsien, R.Y. (2000) Monitoring Protein Conformations and Interactions by Fluorescence Resonance Energy Transfer between Mutants of Green Fluorescent Protein. *Methods in Enzymology*, **327**, 472-500. http://dx.doi.org/10.1016/s0076-6879(00)27297-2

[19] Wu, Y.F., Zhou, Y.B., Song, J.P., Hu, X.J., Ding, Y. and Zhang, Z.H. (2008) Using Green and Red Fluorescent Proteins to Teach Protein Expression, Purification, and Crystallization. *Biochemistry and Molecular Biology Education*, **36**, 43-54. http://dx.doi.org/10.1002/bmb.117

[20] Poppenborg, L., Friehs, K. and Flaschel, E. (1997) The Green Fluorescent Protein is a Versatile Reporter for Bioprocess Monitoring. *Journal of Biotechnology*, **58**, 79-88. http://dx.doi.org/10.1016/S0168-1656(97)00134-X

[21] Johnson, F.H., Shimomura, O., Saiga, Y., Gershman, L.C., Reynolds, G.T. and Waters, J.R., (1962) Quantum Efficiency of Cypridina Luminescence, with a Note on That of Aequorea. *Journal of Cellular and Comparative Physiology*, **60**, 85-103. http://dx.doi.org/10.1002/jcp.1030600111

[22] Morin, J.G. and Hastings, J.W. (1971) Energy Transfer in a Bioluminescent System. *Journal of Cellular Physiology*, **77**, 313-318. http://dx.doi.org/10.1002/jcp.1040770305

[23] Morise, H., Shimomura, O., Johnson, F.H. and Winant, J. (1974) Intermolecular Energy Transfer in the Bioluminescent System of Aequorea. *Biochemistry*, **13**, 2656-2662. http://dx.doi.org/10.1021/bi00709a028

[24] Tsien, R.Y. (1998) The Great Fluorescent Protein. *Annual Review of Biochemistry*, **67**, 509-544. http://dx.doi.org/10.1146/annurev.biochem.67.1.509

[25] Chiu, W., Niwa, Y., Zeng, W., Hirano, T., Kobayashi, H. and Sheen, J. (1996) Engineered GFP as a Vital Reporter in Plants. *Current Biology*, **3**, 325-330. http://dx.doi.org/10.1016/S0960-9822(02)00483-9

[26] Yang, F., Moss, G. and Maheswaran, G. (1996) The Molecular Structure of Green Fluorescent Protein. *Nature Biotechnology*, **14**, 1246-1251. http://dx.doi.org/10.1038/nbt1096-1246