Influence of Some Point Mutations of Green Fluorescent Protein Gene on Expression Efficiency in *Escherichia coli*

Xiaoliang Hao

Abstract—In this paper, I studied on the influence of some point mutations of green fluorescent protein (GFP) gene on expression efficiency in *Escherichia coli* (*E. coli*). It was found that some mutant genes of GFP had influenced on the expression efficiency of GFP gene. In this paper, I acquired six GFP mutants (A402G, A675C, A462G, A88G, T357C, A443G), compared the differences among the six mutant sites, and discussed the relationship between the mutant gene and protein expression efficiency. It was found that A443G had the lowest relative fluorescence intensity (0.04 fold), and A88G had the lower relative fluorescence intensity (1.2 fold), against wildtype of GFP. I hope that the findings, which may be applicable to genetic engineering, will be helpful for further studies of protein expression.

Index Terms—Mutant gene, gene expression, green fluorescent protein, fluorescence intensity.

I. INTRODUCTION

Green fluorescent protein (GFP) was discovered as a companion protein to aequorin, the famous chemiluminescent protein from Aequorea jellyfish [1]. GFP is composed of 238 amino acids and hence a rather small protein with a molecular weight of roughly 27 kDa [2]. Excitation at 396 nm results in an emission maximum at 508 nm [3]. The discovery of GFP led to a new revolution in molecular biology, whose different mutants had been engineered over the last few years [4], [5]. Meantime, *Escherichia coli* (*E. coli*) is a convenient host for protein expression and one of the organisms of choice for the production of recombinant proteins in high quantities and low production costs, it has become the most popular expression platform [6]–[13].

As we all know, gene contains a specific sequence of nucleotides which gives the instructions for the specific sequence of amino acids, if the gene is mutated, may change the formation of amino acids, and then influence the structure and function of the protein [14]. 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. GFP and its some mutants have many useful applications, such as real-time detection, no disruption or toxicity to the host cells, no requirement for cofactors, and feasibility of fusion with the target proteins, have become one of the most used fluorescent probes in cell biology and molecular biology [15]-[17].

The research about GFP and its some mutants has always been concerned. People had found mutant was a good way, seven strategies were used to generate mutations in GFP [18], a library had been constructed in *E. coli* of mutant GFP genes [15]. A new method had been developed to determine the mutagenic efficacy of a suspected mutagen by employing GFP as a direct biosensor for mutation detection [19]. A new strategy of molecular evolution aimed at generating a red-emitting mutant of GFP had been used and had succeeded in producing the first GFP mutant that substantially matures to the red-emitting state with excitation and emission maxima at 555 and 585 nm, respectively [20].

Extensive efforts had been applied to create GFP variants. As a result, blue and cyan mutants with Tyr66 substituted with Phe, His, or Trp were generated. Also, it was demonstrated that certain point mutations can strongly influence a proportion of neutral / anionic chromophore [20]-[22]. Two plasmid vectors had been constructed for the expression of the synthetic SGFP-TYG gene in ascomycete species, indicating that SGFP-TYG can be used as a highly effective vital marker in ascomycetes [23].

Relative to the wildtype of GFP, Emerald GFP was originally reported to have 5 amino acid changes (S65T, S72A, N149K, M163T, and I167T) [24]. Point mutations have been inserted, which resulted in much faster chromophore formation and up to a 100-fold increase in the fluorescence of the produced protein [22], [25], [26].

To research the properties of the mutations could be helpful for understanding the mechanism of mutant sites and finding the phenomenon in my favor. Compared to the newly discovered mutants, the previously characterized mutants GFP-S65T and GFP-Y66H had been employed [27]. Many properties relevant to quantitative imaging in living cells of wildtype, AlphaGFP (F99S/M153T/V163A), S65T, EGFP (F64L / S65T) , EBFP (F64L /S65T / Y66H / Y145F) had been investigated. The five GFP variants had been used extensively or are potentially useful [28].

This study was based on the point mutation. During our former research, my collaborators and I 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 and studied on influence of last sense codon on expression in *E. coli* [29]. In the process, I had acquired twenty mutant recombinant GFP genes. Among them, one special variant GFP(ACC) was found, whose inserted last sense codon was ACC, had three

Manuscript received April 15, 2015; revised June 30, 2015. The author is grateful for financial support by the "Nano Project" for Private Universities from Ministry of Education, Culture, Sports, Science, and Technology of Japan.

Xiaoliang Hao is with the Department of Life Science and Green Chemistry, Saitama Institute of Technology, 1690, Fusaiji, Fukaya, Saitama 369-0293, Japan and School of Chemical Engineering, University of Science and Technology LiaoNing, 185 Qianshan Middle Road, High-tech zone, Anshan, LiaoNing, 114501, China (e-mail: haoxiaoliang1980@163.com).

mutant sites, 56, 357 and 443 of the GFP gene. *E. coli* having this vector indicated 0.1 fold of fluorescence intensity against wildtype of GFP. From this variant, T357C and A443G were acquired, both of which had one mutant site only, without inserted last sense codon, whose fluorescence intensities were 1.01 fold and 0.04 fold, respectively.

II. MATERIALS AND METHODS

A. Constructions of Plasmid and Cloning Vectors

During our former researches, I had acquired twenty mutant recombinant GFP genes. Among these recombinant GFP genes, the GFP genes with special features were selected, such as had the higher or lower fluorescence intensity, easy to get and only had one mutant site. pKK223-3 was used for the expression vector in this study. In order to research the relationship between mutant gene and the protein expression efficiency, polymerase chain reaction (PCR) was used for removing these last sense codons of the modified GFP genes. PCR was performed at 94 °C for 10 min, followed by 25 cycles of 94 $\,$ $\!\!\!\!\!^{\circ}$ for 45 s, 48 $\,$ $\!\!\!\!\!^{\circ}$ for 1 min and 68 $\,$ $\!\!\!\!^{\circ}$ 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. Primers (PCR) were purchased from Sigma-Aldrich (Tokyo, Japan). Restriction enzymes (EcoRI, HindIII, HpaI, PmaCI), a PCR amplification kit and T4 DNA ligase were purchased from TaKaRa (Otsu, Japan).

The mutant sites of T357C and A445G were so special that the constructions were different from the other mutations. The restriction enzymes (*Pma*CI and *Hin*dIII) were used to digest the wildtype of GFP and created T357C. Latter the restriction enzymes (*Eco*RI and *Hpa*I) were used to digest the wildtype of GFP and created A443G. The used primers were seen below:

1) T357C (*Pma*CI and *Hin*dIII) GFP *Pma*CI Primer (5'Primer) 5'-AAGACACGTGCTGAAGTCAAGTTTGAAGGTG ATACCCTCGTTAATAGAATCGAGTAAA-3'

GFP *Hin*dIII New Primer (3'Primer) 5'-CTCTAAGCTTTTATTTGTATAGTTCATCCA-3'

2) A443G (*Eco*RI and *Hpa*I) GFP *Eco*RI Primer (5'Primer) 5'-CCCGAATTCTTTAACTTTAGGAAACACAATTC ATGAGTAAAGGAGAAGAACTT-3'

GFP *Hpa*I Primer (3'Primer) 5'-TGAAGTTAACTTTGATTCCATTCTTTGTTTG TCTGCCATGATGTATACATTGCGTGAGTTATAGTTG TATTCC-3'

When the relationship between mutant gene and protein synthesis efficiency was researched, apart from T357C and A443G, there were some other common mutant GFP genes, in order to remove the last sense codons and get the mutant GFP genes (A402G, A675C, A462G, A88G) with only one mutant site respectively, the primers (*Eco*RI and *Hin*dIII) below were used:

GFP *Eco*RI Primer (5'Primer):

5'-CCCGAATTCTTTAACTTTAGGAAACACAATTCAT GAGTAAAGGAGAAGAACTT-3'

GFP *Hin*dIII New Primer (3'Primer): 5'-CTCTAAGCTTTTATTTGTATAGTTCATCCA-3'

B. Measurement of GFP Relative Fluorescence Intensity in E. coli

GFP-fusion fluorescence intensity was an excellent indicator of over-expression potential [30]. Because fluorescence was one of the most convenient ways to follow a protein expression and purification procedure [31], so the fluorescence intensity was used to analyze the expression efficiency of proteins. The cells were cultivated in LB medium supplemented with 0.1 mg / mL ampicillin and 40 μ M isopropyl- β -D-thiogalactopyranoside (IPTG) at 37 °C for 18 h [32]. The culture of *E. coli* was measured its absorbance at 600 nm, and the fluorescence intensity at 508 nm was excited at 396 nm of the same culture and measured 3 times by the Gemini fluorescence microplate reader (Nihon Molecular Devise). The expression efficiency of GFP gene was compared as the value of the fluorescence intensity / the absorbance at 600 nm.

C. SDS-PAGE Analysis of Extracts of E. coli Having Mutant Site of GFP Gene

After the fluorescence intensities of acquired mutations were known, the mutant genes were expressed in the proteins, compared the quantities of proteins, and analyzed the influences of the mutant genes on protein expression. In order to be easily compared, sodium dodecyl sulfate polyacrylamide gel electropheresis (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. 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.

EcoRI site

3- Mutant siteof A443G

CACGTG: Recognition site of PmaCI GTTAAC: Recognition site of HpaI

Fig. 1. The mutant sites of GFP(ACC)

III. RESULTS AND DISCUSSION

From Fig. 1, it was seen that GFP(ACC) had three mutant sites. By the study, due to the experimental conditions and operational constraints, I was failure to acquire the first mutant site GFP(ACC-M1), finally the second mutant site (T357C) and the third mutant site (A443G) were acquired. CACGTG was the recognition site of primer *Pma*CI, GTTAAC was the recognition site of primer *Hpa*I. The digested sites of *Eco*RI and *Hin*dIII were located in the 5' and 3' side of GFP gene.



Fig. 2. The construction of T357C.

From Fig. 2, it was seen that the enzymes *Pma*CI and *Hin*dIII were used to digest wildtype of GFP to acquire the expression vector which had the *Pma*CI and *Hin*dIII digested sites, the primer *Pma*CI which had one mutant site and primer *Hin*dIII were used to acquire the GFP gene of T357C by PCR. After ligation, T357C was acquired, without last sense codon, but with only one mutant site. Similarly, the enzymes *Eco*RI and *Hpa*I were used to digest wildtype to acquire the expression vector which had the *Eco*RI and *Hpa*I digested sites, the primer *Hpa*I which had one mutant site and primer *Eco*RI were used to acquire the GFP gene of A443G by PCR. After ligation, A443G was acquired (see Fig. 3).



Fig. 3. The construction of A443G.

Compared to T357C and A443G, the acquisition method of A402G, A675C, A462G, A88G was different, because A402G, A675C, A462G and A88G came from different last sense codons and T357C and A443G came from GFP(ACC).

From Fig. 4, it was seen that the enzymes *Eco*RI and *Hind*III were used to digest the pKK223-3 to acquire the expression vector which was different from the expression vector (wildtype of GFP) of T357C and A443G. The primers *Eco*RI and *Hind*III were used to acquire the common GFP gene fragment by PCR. After ligation, A402G, A675C, A462G, A88G were acquired respectively.

In this study, these mutant genes A402G, A675C, A462G, A88G, T357C, and A443G were acquired (see Table I).



Fig. 4. The constructions of A402G, A675C, A462G, A88G.

TABLE I: THE MUTATION GENES IN THE RESEARCH			
No	Mutation	Change of codon	Amino acids substitution
1	A402G	GGA-GGG	Nothing
2	A675C	ACA-ACC	Nothing
3	T357C	CTT-CTC	Nothing
4	A462G	GCA-GCG	Nothing
5	A88G	AGT-GGT	S30G
6	A443G	CAC-CGC	H148R

From Table I, it was seen that each mutant GFP gene had only one mutant site. For now, getting the variant with a single mutant site was more easier to compare the impact of mutant site on protein expression efficiency than getting the variant with two and more mutant sites. As shown in Table I, for A402G, the normal nucleotide on the gene site 402 was GGA, after been mutated, the nucleotide was changed from GGA to GGG. The amino acids on this position were both Glycine and were not substituted. The representation of other mutant GFP genes was similar with A402G, but for A88G and A443G, the amino acid was substituted from Serine to Glycine in A88G, the amino acid of A443G was substituted from Histidine to Arginine in A443G, respectively.



Fig. 5. Comparison of the relative fluorescence intensities of mutants of GFP gene in *E. coli*.

As the result, six kinds of GFP mutants with only one mutant site were acquired and the fact that some mutant sites had affected on the protein expression efficiency by measuring the relative fluorescence intensity was tested (see Fig. 5). The expression efficiency of GFP gene was compared as the value of the relative fluorescence intensity / the absorbance at 600 nm normalized with that of the wildtype of GFP gene.

From our former results of the relative fluorescence intensities, it was seen that the fluorescence intensity of GFP(ACC) was 0.1 fold; GFP(A56T, T357C, A443G) was made by removing the last sense codon of GFP(ACC) by PCR, whose fluorescence intensity was also 0.1 fold, from the results of relative fluorescence intensity in this study, it was seen that T357C was 1.01 fold, A443G was 0.04 fold against wildtype, next I put these GFP genes do the SDS-PAGE experiment. When compared the band around 27 kDa, size marker was used to be referenced, it was seen that the bands of wildtype and T357C were the thickest, the bands of GFP(ACC), GFP(A56T, T357C, A443G), and A443G were the thinnest. The results of SDS-PAGE were the same with the results of fluorescence intensity (see Fig. 5). SDS-PAGE confirmed that the relative fluorescence intensity was associated with the quantity of synthesized protein (see Fig. 6). Based on the above results, it was learned that there were some differences in the expression efficiency among six kinds of GFP mutations and we could compare and analyze the reasons behind the phenomenon.



Fig. 6. SDS-PAGE of soluble proteins extracted from *E. coli* in GFP gene M: size marker. MW of GFP: ~ 27kDa.

It had shown that the A443G had the lowest relative fluorescence intensity (0.04 fold), and A88G had the lower relative fluorescence intensity (0.16 fold), it meaned that when the nucleotide was changed from A to G on the gene site 443, the relative fluorescence intensity was the lowest. Similarly when the nucleotide was changed from A to G on the gene site 88, the relative fluorescence intensity would be decreased to 0.16 fold against wildtype of GFP and acquired the lower expression efficiency. On the contrary, for the A402G, even the amino acid was not substituted, when the nucleotide was changed from A to G on the gene site 402, the relative fluorescence intensity was increased to 1.2 fold. That meaned that some special mutations could control the increase and decrease of the protein expression efficiency. In the future research, it is expected to find more mutant sites, especially some other important gene sites, finally achieve that the protein could be expressed by artificial control.

REFERENCES

- O. Shimomura, F. H. Johnson, and Y. Saiga, "Extraction, purification, and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea," *J. Cell. Comp. Physiol*, vol. 59, pp. 223-239, June 1962.
- [2] D. C. Prasher, V. K. Eckenrode, W. W. Ward, F. G. Prendergast, and M. J. Cormier, "Primary structure of the Aequorea victoria green-fluorescent protein," *Gene*, vol. 111, pp. 229-233, February 1992.
- [3] J. C. March, G. Rao, and W. E. Bentley, "Biotechnological applications of green fluorescent protein," *Appl Microbiol Biotechnol*, vol. 62, pp. 303-315, March 2003.
- [4] A. Miyawaki and R Y. Tsien, "Monitoring protein conformations and interactions by fluorescence resonance energy transfer between mutants of green fluorescent protein," *Methods Enzymol*, vol. 327, pp. 472-500, February 2000.
- [5] Y. Wu, Y. Zhou, J. Song, X. Hu, Y. Ding, and Z. Zhang, "Using green and red fluorescent proteins to teach protein expression, purification, and crystallization," *Biochemistry and Molecular Biology Education*, vol. 36, pp. 43-54, January 2008.
- [6] N. H. Tolia and L. Joshua-Tor, "Strategies for protein coxpression in *Escherichia coli*," *Nature Methods*, vol. 3, pp. 55-64, January 2006.
- [7] G. L. Rosano and E. A. Ceccarelli, "Recombinant protein expression in *Escherichia coli*, advances and challenges," *Frontiers in Microbiology*, vol. 5, pp. 1-17, April 2014.
- [8] G. Hannig and S. C. Makrides, "Strategies for optimizing heterologous protein expression in *Escherichia coli*," *Trends in Biotech*, vol. 16, pp. 54-60, February 1998
- [9] F. Baneyx, "Recombinant protein expression in *Escherichia coli*," *Curr Opin Biotechnol*, vol. 10, pp. 411-421, October 1999.
- [10] A. K. Panda, "Bioprocessing of therapeutic proteins from the inclusion bodies of *Escherichia coli*," *Adv Biochem Eng Biotechnol*, vol. 85, pp. 43-93, June 2003.
- [11] L. A. Palomares, S. Estrada-Mondaca, and O. T. Ramirez, Production of Recombinant Proteins, Challenges and Solutions: Methods Mol Biol, 2004, ch. 2, pp. 15-52.
- [12] F. J. Mergulhao, D. K. Summers, and G. A. Monteiro, "Recombinant protein secretion in *Escherichia coli*," *Biotechnol Adv*, vol. 23, no. 3, pp. 177-202, May 2005.
- [13] J. Yin, G. Li, X. Ren, and G. Herrler, "Select what you need: A comparative evaluation of the advantages and limitations of frequently used expression systems for foreign genes," *J Bacteriol*, vol. 127, pp. 335-347, January 2007.
- [14] G. Kudla, A. W. Murray, D. Tollervey, and J. B. Plotkin, "Coding-sequence determinants of gene expression in *Escherichia coli*," *Science*, vol. 324, pp. 255-257, April 2009.
- [15] B. P. Cormack, R. Valdivia, and S. Falkow, "FACS-optimized mutants of the green fluorescent protein (GFP)," *Gene*, vol. 173, pp. 371-374, February 1996.
- [16] C. Nolte, M. Matyash, T. Pivneva, C. G. Schipke, C. Ohlemeyer, U. K. Hanisch, F. Kirchhoff, and H. Kettenmann, "GFAP promoter controlled EGFP-expressing transgenic mice: a tool to visualize astrocytes and astrogliosis in living brain tissue," *Glia*, vol. 33, pp. 72-86, January 2001.
- [17] H. J. Cha, M. Q. Pham, G. Rao, and W. E. Bentley, "Expression of green fluorescent protein in insect larvae and its application for heterologous protein production," *Biotechnology and Bioengineering*, vol. 56, pp. 239-247, November 1997.
- [18] D. A. Zacharias and R Y. Tsien, Molecular Biology and Mutation of Green Fluorescent Protein, Green Fluorescent Protein: Properties, Applications, and Protocols, 2nd edition, 2006, pp. 83-118.
- [19] Y. K. Tak, P. K. Naoghare, K. H. Lee, S. S. Park, and J. M. Song, "Green fluorescent protein (GFP) as a direct biosensor for mutation detection: Elimination of false-negative errors in target gene expression," *Analytical Biochemistry*, vol. 380, pp. 91-98, May 2008.
- [20] A. S. Mishin, F. V. Subach, I. V. Yampolsky, W. King, K. A. Lukyanov, and V. V. Verkhusha, "The first mutant of the aequorea victoria green fluorescent protein that forms a red chromophore," *Biochemistry*, vol. 47, pp. 4666-4673, February 2008.
- [21] R. Heim and R. Y. Tsien, "Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer," *Curr. Biol*, vol. 6, pp. 178-182, February 1996.

- [22] R. Heim, D. C. Prasher, and R. Y. Tsien, "Wavelength mutations and posttranslational autoxidation of green fluorescent protein," in *Proc. Natl. Acad. Sci. U.S.A*, December 1994, vol. 91, pp. 12501-12504.
- [23] R. Maor, M. Puyesky, B. A. Horwitz, and A. Sharon, "Use of green fluorescent protein (GFP) for studying development and fungal-plant interaction in Cochliobolus heterostrophus," *Mycol. Res*, vol. 102, no. 4, pp. 491-496, April 1998.
- [24] R. Y. Tsien, "The green fluorescent," Annu. Rev. Biochem, vol. 67, pp. 509-544, April 1998.
- [25] R. Heim, A. B. Cubitt, and R. Y. Tsien, "Improved green fluorescence," *Nature*, vol. 373, pp. 663-664, February 1995.
- [26] W. Chiu, Y. Niwa, W. Zeng, T. Hirano, H. Kobayashi, and J. Sheen, "Engineered GFP as a vital reporter in plants," *Current Biology*, vol. 6, pp. 325-330, March 1996.
- [27] A. G. von Arnim, X. W. Deng, and M. G. Stacey, "Cloning vectors for the expression of green fluorescent protein fusion proteins in transgenic plants," *Gene*, vol. 221, pp. 35-43, August 1998.
- [28] G. H. Patterson, S. M. Knobel, W. D. Sharif, S. R. Kain, and D. W. Piston, "Use of the green fluorescent protein and its mutants in quantitative fluorescence microscopy," *Biophys J*, vol. 73, no. 5, pp. 2782-2790, November 1997.
- [29] X. L. Hao, S. Inoue, and M. Ishikawa, "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*, vol. 3, pp. 13-18, June 2015.

- [30] J. Hammon, D. V. Palanivelu, J. Chen, C. Patel, and D. L. Minor, "A green fluorescent protein screen for identification of well-expressed membrane proteins from a cohort of extremophilic organisms," *Protein Sci.*, vol. 18, no. 1, pp. 121-133, January 2009.
- [31] G. S. Waldo, B. M. Standish, J. Berendzen, and T. C. Terwilliger, "Rapid protein-folding assay using green fluorescent protein," *Nat Biotechnol*, vol. 17, pp. 691-695, July 1999.
- [32] K. R. Siemering, R. Golbik, R. Sever, and J. Haseloff, "Mutations that suppress the thermosensitivity of green fluorescent protein," *Current Biology*, vol. 6, pp. 1653-1663, December 1996.



Xiaoliang Hao was born in China on June 7, 1980. He is a teacher of University of Science and Technology Liaoning, China. Since 2012, he is an exchange PhD student, and now has studied in Saitama Institute of Technology, Japan. The major field of his study includes genetic microorganisms, life sciences areas. In particular, some researches about gene sequence changes and expression of green fluorescent protein, etc.