# Using error-prone PCR to create green fluorescent gene variants

Juliana Monson

BIO 489

April 9, 2016

## ABSTRACT

Many versions of GFP have been engineered by site-directed or random mutagenesis of native proteins. However, there are still areas where these proteins can be altered in order to provide the research community with more effective biotechnology tools. This experiment used error-prone PCR to introduce random mutations into the GFP gene in an attempt to shift the wavelength of emission or brightness of GFP. The end goal of this was to produce a laboratory exercise suitable for the undergraduate. A variety of primer pairs were used to amplify the GFP gene from pmaxGFP (Lonza). The 5' primer ended at the Met1 of the GFP gene to allow potential mutation to occur everywhere except for Met1, and annealing temperatures ranging from 47°C to 64°C were tried. However, due to time constraints and issues with primers, PCR attempts were unsuccessful at amplifying the GFP gene.



## KEYWORDS

Green fluorescent protein, error-prone PCR, random mutagenesis, primers


## INTRODUCTION

Fluorescent proteins were first isolated from the jellyfish *Aequorea victoria* (Shimomura and others 1962). They were called green fluorescent protein due to the colour of luminescence that the protein emitted in vivo (Johnson and others 1962).

The gene for GFP was found to code for a polypeptide composed of 238 amino acids, that is spontaneously fluorescent (Prasher and others 1992). GFP quickly became a tool that was utilized throughout the research community. Further study revealed that the protein is an eleven-stranded β-sheet, wrapped into a form of secondary structure called a β-can. The section of the protein responsible for fluorescence is called the chromophore, and is located near the center of the protein (Ormo and others 1996). During GFP translation, the protein first folds into its native

conformation, then undergoes self-sufficient post-translational modifications (Prasher and others 1992); the chromophore forms when an internal tripeptide cyclizes and then becomes oxidized (Nagai and others 2002). The β-can fold is a distinguishing feature of the superfamily of GFP-like proteins, which vary in the conjugation of the chromophore (Campbell 2008), and the electrostatic interactions that occur between the chromophore and the environment of the protein that surrounds it (Henderson and Remington 2005).

Recombinant expression of GFP was first applied experimentally in 1994 to *Caenorhabditis elegans* neurons (Chalfie and others 1994) and to *Escherichia coli* (Inouye and Tsuji 1994). It is now common practice to introduce the GFP gene into a cell to track transcription, act as a biosensor, or a partner for fluorescence resonance energy transfer (FRET).

Although GFP and the other proteins of this type that have been isolated are widely used in research, there are problems associated with the native proteins. For example, since most fluorescent proteins were isolated from marine organisms, the ideal protein folding temperature for the native protein is much cooler than 37°C. This problem has been resolved by the introduction of certain mutations to the GFP gene (Tsien 1998; Nagai and others 2002). While green were originally the only colour of fluorescent proteins (Tsien 1998), there are now violet, red, orange, yellow, and cyan variants available of differing brightness. However, there is still no bright far-red protein (Shaner and others 2005). Photostability is not usually an issue when an experiment requires less than 10 photographs, but if more are required, prolonged exposure to the excitation source can result in photobleaching (Shaner and others 2005). The production of proteins that can resist photobleaching without sacrificing brightness is still being researched. For FPs to be effective, it is important that environmental situations do not affect their fluorescence. Some of the yellow FPs are very chloride sensitive, making them difficult to use in

certain cellular situations. The proteins Citrine (Nagai and others 2002) and Venus (Griesbeck and others 2001) have been produced, addressing the problem of chloride sensitivity. Many of the proteins, including mOrange, GFPs and yellow FPs are vulnerable to low pH, limiting the applications of these proteins (Shaner and others 2005). Reducing the sensitivity of FPs to environmental conditions is important for the investigation of different cellular compartments.

These issues (and others) have pushed the development of new variants of FPs. Although many problems associated with the early FPs have been resolved, there are still things that need to be improved, and new applications of FPs to be discovered. The engineering of new proteins is an ongoing field of research.

The production of new FPs often requires the mutation of available FPs. Sometimes, a specific amino acid is targeted, such as A206K to produce monomer FPs (Zacharias and others 2002). This is called site directed mutagenesis, and is often done using mutagenic primers to perform PCR, overlap extension PCR (Ho and others 1989) or the ligation method: two separate PCR reactions are conducted with primers containing different target mutations, as well as the same restriction enzyme cut site at the start of the primers. Products are then digested and ligated (Shaner and others 2004). If there is no specific residue being targeted, then random mutagenesis is employed by manipulating PCR conditions to encourage the incorporation of incorrect base pairings. Sometimes combinations of these methods are used.

Polymerase chain reactions are most often used to accurately amplify a segment of DNA. However, there are applications of error-prone PCR as well. One of the most commonly used polymerases for PCR is *Taq* polymerase; however, this enzyme has a relatively large error rate, ranging from $10^{-5}$ (Eckert and Kunkel 1991) to $10^{-3}$ (Caldwell and Joyce 1994) per nucleotide. The reason for this low fidelity is the lack of 3'-5' exonuclease, or proofreading, activity of *Taq*.

Generally, using PCR conditions that combat the incorporation of incorrect nucleotides controls this characteristic, but the *Taq* polymerase quality of low fidelity is exploited in mutagenic PCR. In order to increase the mutation frequency, the number of cycles or error rate of the polymerase must be increased (Eckert and Kunkel 1991).

The error rate of the polymerase can be adjusted by the conditions of PCR. Increasing the rate of DNA production results in an increased probability that a strand of DNA will be extended from a mispaired base. The addition of a nucleotide onto an incorrect base is kinetically unfavourable, but this block can be overcome by increasing the concentration of dNTPs, increasing the extension period of the PCR cycle (Eckert and Kunkel 1991), or increasing the amount of enzyme (Caldwell and Joyce 1994). If the concentration of a single dNTP is decreased, the error rate of that particular base is increased. This has the potential to decrease the fidelity of the polymerase eightfold (Eckert and Kunkel 1991). The addition of $MnCl_2$ can also be used to decrease polymerase specificity (Caldwell and Joyce 1994).

In this experiment, the concentration of dATP was be reduced and $MnCl_2$ added to the PCR to promote random mutagenesis within the GFP gene as described by Griesbeck and others (2001). It was expected that the mutations would result in a shift of the wavelength or emission brightness of GFP.

## METHODS

Error-prone PCR was performed on a green fluorescent protein gene in the vector pmaxGFP (Lonza, Figure 1) using the protocol established by Greisbeck and others (2001).

In this experiment, a variety of primers (Table 1) and annealing temperatures (Ta) were used. All of the 5' primers included an *Age*I site and ended at the Met1 of the GFP gene. This allowed for potential mutation to occur everywhere except for Met1 (Griesbeck and others

2001). The PCR (35 cycles) was run in 50µl batches, containing 250 U/50 µl of *Taq* polymerase, 5 µl of 10 x PCR buffer with $Mg^{2+}$ (NEB, #B9014S), 0.5 µM of each primer, and 200 µM of the dNTPs.

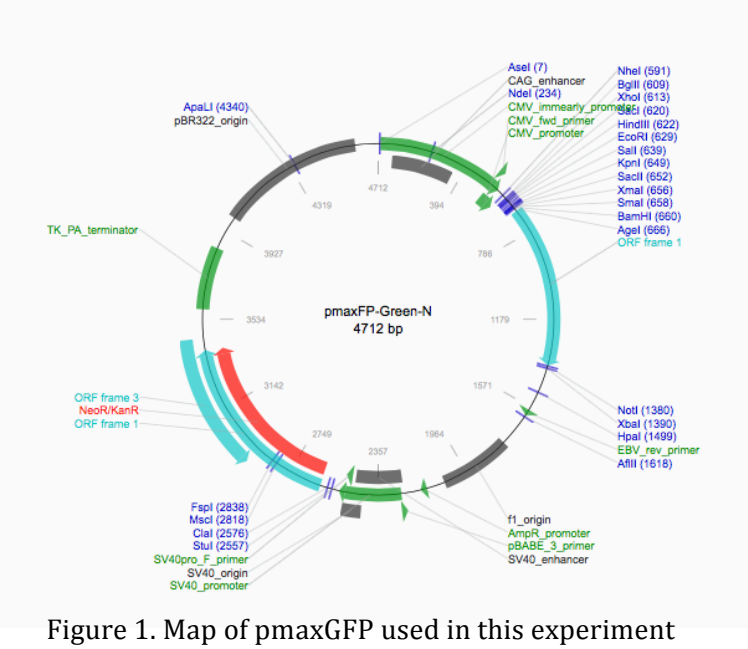


Figure 1. Map of pmaxGFP used in this experiment (Image from https://www.addgene.org/browse/sequence_vdb/3525/; accessed March 22, 2016).

Table 1. The sequences and features of the primers used in this experiment.

| Name of Primer | Primer Features | Sequence |
|---|---|---|
| Primer 1 | Forward primer, start codon, *Age*I cut site, Tm= 72°C | 5'- GATCCACCGGTCGCCACCATG |
| Primer 2 | Reverse primer, *Xba*I cut site, Tm= 66°C | 5'- ATCTAGAGTCGCGCCGGTTTA |
| Primer 3 | Reverse primer, *EcoR*I cut site, Tm= 57°C | 5'- ATTATGAATTCTAGAGTCGCGGCC |
| Primer 4 | Reverse primer, *Nde*I cut site, Tm= 47°C | 5'- TATGGCTCATATGATCTAG |
| Primer 5 | Forward primer, start codon, *Age*I cut site, Tm= 64°C | 5'- TCCACCGGTCGCCACCATG |

Polymerase chain reactions using primers 1 and 2 were performed with an annealing temperature (Ta) of 64°C and 2.5 µg and 0.25 µg of template DNA per reacion. All following reactions were performed using a 0.25 µg of template DNA per reaction. These reactions were treated with *Nde*I, an enzyme with a single restriction site on the pmaxGFP.

The reaction was then performed with primers 1 and 3 (Ta= 62°C). Primers 4 and 5 were used in PCR with three different annealing temperatures: 47°C, 50°C and 55°C. The final pair used was primers 3 and 5, at an annealing temperature of 55°C.

The PCR products were then purified by agarose gel electrophoresis and viewed using UV light.

## RESULTS

When primers 1 and 2 were run together (Ta= 64°C), there was one large band. The ladder (NEB, #N3233S) did not resolve, so it is not clear what size of DNA the band represents (Figure 2a). All of the following gels were run with a different ladder (NEB, #N0468S). The same primer pair was run again with a different ladder, and the following bands are visible: 4.3 kb, 3.3 kb, 2.5 kb, 1.6 kb, and a band much smaller than 0.5 kb (Figure 2b). The same reaction was also run with a smaller concentration of template DNA, and no bands can be seen in the PCR (Figure 2c). The primer pair was also tested using an annealing temperature of 55°C, but no band was seen (not shown).

Primers 1 and 3 were run in a PCR (Ta= 62°C), and no DNA was seen on the gel (Figure 2d).

Primers 4 and 5 were used in a PCR (Ta= 47°C) that resulted in a smear of DNA of varying sizes: less than 0.5 kb to greater than 3 kb. There were also two bands that were smaller

than 0.5 kb that appeared in the negative control (Figure 2e). In separate reactions (Ta= 50°C and 55°C), no bands were seen in the gels (not shown).

The final pair of primers tested was primers 3 and 5 (Ta= 55°C). Three differently sized bands can be seen in the PCR product: 2.7 kb, 1.6 kb and less than 0.5 kb. In the negative control, a similarly sized less than 0.5 kb band can be seen (Figure 2f).

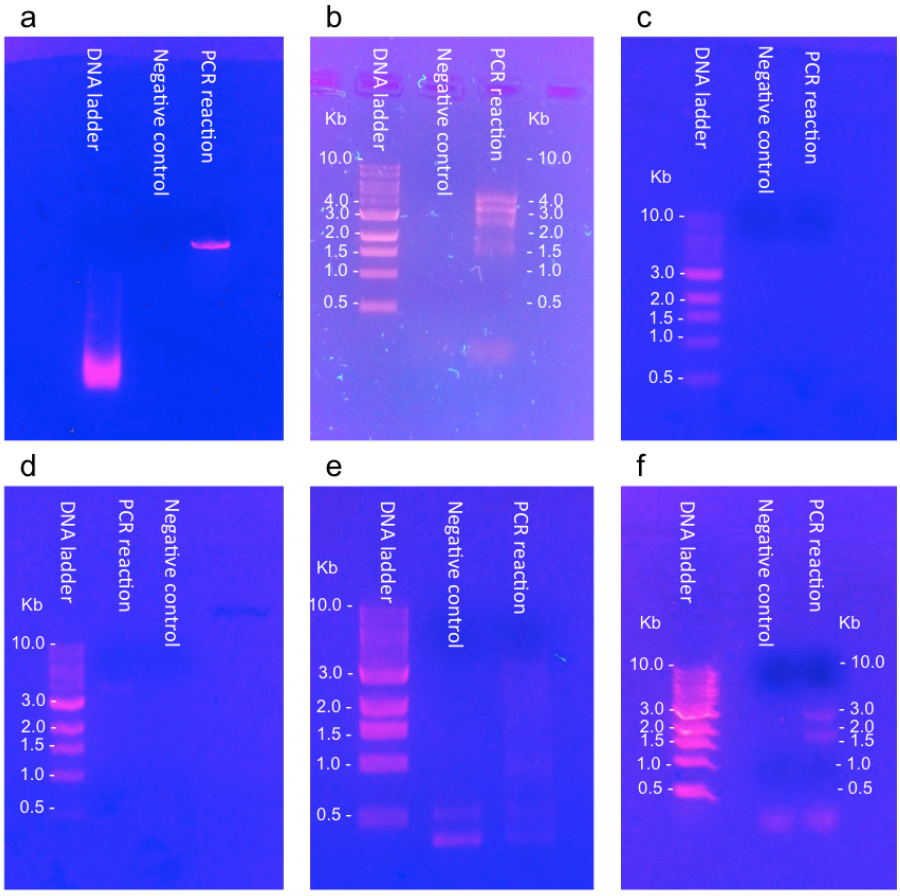


Figure 2. Agarose gels were run using PCR products run with different primers (shown in Table 1), and different annealing temperatures. a) PCR product from reaction using primers 1 and 2 (Ta= 64°C) using a DNA ladder (NEB, CAT: #N3233S) and 2.5µg of template DNA. All other gels were run with a different DNA ladder (NEB, #N0468S) b) Primers 1 and 2, Ta= 64°C with 2.5µg of template DNA. c) Primers 1 and 2, Ta= 64°C with 0.25µg of template DNA. d) Primers 1 and 3 (Ta= 62°C). e) Primers 4 and 5 (Ta= 47°C). f) Primers 3 and 5 (Ta=55°C).

## DISCUSSION

In this experiment, the targeted amplicon should have been 735-747 bp, depending on the primer pair used. However, in each PCR that was performed, the product fragments seen in the bands were either much larger or much smaller than the expected size.

Because the ladder did not resolve in the first gel that was run (Figure 2a), it is difficult to determine the band size. In the second PCR using primers 1 and 2, it was determined that the concentration of template DNA was high (Figure 2b). Three of the bands (4.3 kb, 3.3 kb and, 1.6 kb) correspond to the three topological forms of a plasmid: nicked/relaxed, linear, and supercoiled (Figure 2b). The smear of DNA between the 1.6 kb band and the larger fragments could be topoisomeric contaminants (Smith and others 2007). The small band is a primer-dimer. The initial amount of template DNA used in the PCR was 2.5 µg. When the concentration of template DNA is too high, it can poison in a couple of different ways: the concentration of Mg2+ is effectively lowered as it is used to stabilize the structure of the larger concentration of template DNA. The amount of Mg2+ is insufficient for *Taq* function, so the reaction will not proceed (Lorenz 2012). In addition, the initial molar ratio of primers to target sequence should be ~$10^8$:1 to ensure that the target DNA will bind to primers rather than each other. If the ratio is too low, the reaction will be compromised as was seen in Figure 2b (Cha and Thilly 1993). When the amount was reduced to 0.25 µg, the same reaction was run. There were no bands visible in the gel (Figure 2c). Likewise, there are no bands in the PCR product from the reaction using primers 1 and 3 (Figure 2d). This is likely due to the difference in melting temperatures in the primers. The difference in primer melting temperature should not be greater than 3°C so that the annealing step can work effectively. However, the difference in Tm between primers 1-2 and 1-3 is 6°C and 15°C, respectively.

The difference in Tm between primers 4-5 is 17°C. in order to combat this, annealing temperatures that are lower than average were used. The annealing temperature of 47°C resulted in a smear of DNA in the PCR product and the formation of primer-dimers (Figure 2e). With the lowered annealing temperature, primer 4 was able to bind to the template DNA, allowing DNA to be amplified. However, lowering the annealing temperature decreases the stringency of primer annealing, and so many nonspecific segments of the template being amplified (Malhotra and others 1997). When the annealing temperatures were raised, to 50°C and 55°C, there were no bands seen (not shown). Because primer 4 has a Tm of 47°C, it was probably unable to anneal to the template at these temperatures, and DNA could not be amplified.

When primers 3-5 were used, primer-dimers formed, as well as two other bands, 3.0 kb and 1.6 kb long (Figure 2f). The primers may have found another place of complimentarily on the plasmid, producing to wrong segments of DNA.

One possible element that could have contributed to the issues experienced with in this experiment is the GC-rich region in the plasmid directly before the start of the gene. The forward primers were chosen to include the Met1 of the GFP gene in order to prevent mutation from occurring at the start codon in the error-prone conditions. However, primer annealing is less specific in regions GC-rich segments of DNA (Sarkar and others 1990). Another result of this was that the starting primers used had higher melting temperature than anticipated. Finding an ideal annealing temperature to use for both primers was difficult.

Due to time constraints and difficulties with primers, the GFP gene was not successfully amplified. Future attempts should look at completely different primers rather than conserving the Met1, and designing primers with melting temperatures within 1-2°C of each other.

REFERENCES

Belousov VV, Fradkov AF, Lukyanov KA, Staroverov DB, Shakhbazov KS, Terskikh AV, Lukyanov S. 2006. Genetically encoded fluorescent indicator for intracellular hydrogen peroxide.

Caldwell RC, Joyce GF. 1995. Mutagenic PCR. PCR Methods and Applications. 3:5136- 5140.

Campbell RE. 2008. Fluorescent Proteins. Scholarpedia [Internet]. [Cited 2015 October 5]; 3(7):5410.

Cha RS, Thilly WG. 1993. Specificity, Efficiency, and Fidelity of PCR. PCR Methods and Applications. 3:S18-S29.

Chalfie M, Tu Y, Euskirchen G. 1994. Green fluorescent protein as a marker for gene expression. Science. 263(5148):802-805.

Eckert KA, Kunkel TA. 1991. DNA polymerase fidelity and the polymerase chain reaction. PCR Methods and Applications. 1:17-24.

Griebeck O, Baird GS, Campbell RE, Zacharias DA, Tsien RY. 2001. Reducing the Environmental Sensitivity of Yellow Fluorescent Protein. Journal of Biological Chemistry. 276(31):29188-29194.

Hanson GT, Aggeler R, Oglesbee D, Cannon M, Capaldi RA, Tsien RY Remington SJ. 2004. Investigating mitochondrial redox potential with redox-sensitive green fluorescent protein indicators. Journal of Biological Chemistry. 279:13044-13053.

Henderson JN, Remington SJ. 2005. Crystal structures and mutational analysis of amFP486, a cyan fluorescent protein from *Anemonia majano*. Proc. Natl. Acad. Sci. USA. 102:12712-12717.

Ho SN, Hunt HD, Horton RM, Pullen JK, Pease LR. 1989. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene. 77:51-59.

Inouye S, Tsuji FI. 1994. Aequorea green fluorescent protein. Expression of the gene and fluorescence characteristics of the recombinant protein. FEBS Lett. 341(2-3):77-280.

Jayaraman S, Haggie P, Wachter RM, Remington SJ, Verkman AS. 2000. Mecanism and cellular applications of a green fluorescent Protein-based halide sensor. Journal of Biological Chemistry. 275:6047-6050.

Johnson FH, Shomomura O, Saiga Y, Gershman LC, Reynolds GT, Waters JR. 1962. Quantum efficiency of Cybridina luminescences, with a note on that of Aequorea. J Cell Comp. Physiol. 60:85-103.

Lorenz TC. 2012. Polymerase chain reaction: basic protocol plus troubleshooting and optimization strategies. Journal of Visualized Experiments. 63:e3998.

Malhotra K, Foltz L, Mahoney W, Schueler P. 1997. Interaction and effect of annealing temperature on primers used in differential display RT-PCR. Nucleic Acids Research. 26(3):854-856.

Nagai T, Ibata K, Park ES, Kubota M, Mikoshiba K, Miyawaki A. 2002. A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. Nature Biotechnology. 20:87-90.

Nakai J, Ohkura M, Imoto K. 2001. A high signal-to-noise Ca(2+) probe composed of a single green fluorescent protein. Nature Biotechnology. 19:137-141.

Ormo M, Cubitt AB, Kallio K, Gross LA, Tsien RY, Remington SJ. 1996. Crystal structure of the *Aequorea victoria* green fluorescent protein. Science. 273:1392-1395.

Prasher DC, Eckenrode VK, Wared WW, Prendergast FG, Cormier MJ. 1992. Primary structure of the *Aequorea victoria* green-fluorescent protein. Gene. 111:229-233.

Sarkar G, Kapelner S, Sommer S. 1990. Formamide can dramatically improve the specificity of PCR. Nucleic Acids Research. 18(24):7465.

Shaner NC, Campbell RE, Steinbach PA, Giepmans BNG, Palmer AE, Tsien RY. 2004. Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. 22(12):1567-1572.

Shaner NC, Steinbach PA, Tsien RY. 2005. A guide to choosing fluorescent proteins. Nature Methods. 2(12):905-909.

Shimomura O, Johnson FH, Saiga Y. 1962. Extraction, Purification and Properties of Aequorin, a Bioluminescent Protein from Luminous Hydromedusan, *Aequorea.* J. Cell. Comp. Physiol. 59:223-239.

Smith C, DePrince RB, Dackor J, Weigl D, Griffith J, Persmark M. 2007. Separation of topological forms of plasmid DNA by anion-exchange HPLC: Shifts in elution order of linear DNA. Journal of Chromatography B. 854(1-2):121-127.

Tsien RY. 1998. The green fluorescent protein. Annu. Rev. Biochem. 67:509-544.

Zacharias DA, Violin JD, Newton AC, Tsien RY. 2002. Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells. Science. 296:913-916.