Site-Directed Mutagenesis of Position 204 Threonine to Isoleucine Failed to Generate Acid-Tolerant EGFP

Mimi Vignjevic

MacEwan University

BIOL 421: Techniques in Molecular and Cellular Biology

Dr. Kimberly Harcombe and Dr. Habib Rezanejad

7 December 2023

Abstract

Protonation of green fluorescent protein (GFP) in acidic conditions prevents the emission of fluorescent light and limits the ability to visualize, localize, and study acidic organelles. Therefore, it is critical to introduce mutations into enhanced GFP (EGFP) to generate acid-tolerant fluorescent proteins. This experiment aimed to replicate a threonine to isoleucine mutation at position 204 in EGFP and identify if acid-stable fluorescent proteins would be produced in the BL21(DE3) *Escherichia coli* system. Site-directed mutagenesis was utilized to generate T204I mutant EGFP. SDS-PAGE and fluorescence microscopy were employed to analyze induction success and fluorescence. Spectrofluorophotometry was used to determine the excitation and emission spectra of T204I mutant EGFP and whether acid-tolerant proteins were generated. Results illustrated that the mutation of threonine to isoleucine at position 204 produced fluorescent proteins at pH 7. However, at pH 6 and 5, proteins failed to fluoresce. The replication of the T204I mutation failed to generate acid-stable EGFP proteins in the BL21(DE3) *E. coli* system. There is significance in generating acid-stable cellular markers, as currently, no cellular markers thrive in acidic conditions. This limits the ability to study acidic organelles and acidic cellular processes. Creating acid-tolerant markers will permit a greater range of biological research.

Introduction

Green fluorescent protein (GFP) is a small, naturally occurring protein isolated from the jellyfish Aequorea victoria (Chalfie, 1995). GFP natively folds into an 11-stranded beta-barrel surrounding a central alpha helix (Kain, 1999). Within positions 65-67 of the helix lies a serine-tyrosine-glycine tripeptide sequence that cyclizes to form the fluorescent chromophore, the segment of protein that emits light (Kain, 1999; Remington, 2011). GFP's spectral abilities occur

when the chromophore absorbs ultraviolet (UV) radiation at its excitation maxima of 488 nm (Chalfie, 1995). A short, high-energy wavelength converts electrons within the chromophore into high-energy states that experience a transient excited lifetime with gradual energy losses as heat (Remington, 2011). Fluorescence occurs when excitation can no longer be sustained. Electrons emit residual energy as longer, lower energy wavelengths of green light, returning GFP to its stable ground state (Kain, 1999; Kennis et al., 2004; Tsien, 1998). Shielding the chromophore through the beta-barrel allows regulated absorption and emission of energy and protection from fluorescence quenching, contributing to GFP stability and functionality (Follenius-Wund et al., 2003; Jain et al., 2009). GFP's stable structure protects the internal protein environment, encouraging favourable events contributing to fluorescence production.

GFP's spectral abilities have granted researchers an essential experimental tool. Since the protein is not species-specific and does not require additional gene products from *A. victoria*, GFP expression in eukaryotes is possible (Chalfie et al., 1994). In vivo, GFP cDNA is stable, non-toxic, can be cloned, and can be detected non-invasively (Rosochacki & Matejczky, 2002). However, limitations arise following the expression of GFP proteins in vivo. Long periods of UV radiation alter the chromophore's structure and stability, fading fluorescent signals and causing photobleaching (Corish & Tyler-Smith, 1999). Additionally, substandard pH and temperature conditions in vivo modify the protein's hydrogen bond network, altering protein folding and structure (Ward, 2005). To prevent loss of fluorescence in vivo, mutations introduced near the chromophore can prevent structural changes and stabilize the hydrogen bond network (Falkow et al., 1996).

The enhanced green fluorescent protein (EGFP) is a behaviourally optimized GFP mutant. Mutagenesis of EGFP replaces large amino acids with smaller ones, generating more

efficient folding, a further stabilized structure, and more tightly packed proteins (Arpino et al., 2014; Tsien, 1998). EGFP contains two mutations: a phenylalanine to leucine mutation at position 64 and a serine to threonine mutation at position 65 (Falkow et al., 1996). Introducing leucine and threonine promotes tighter protein folding and a more favourable hydrogen bond network, improving GFP's overall stability and structure (Arpino et al., 2012). The enhanced protein structure results in a 35 times stronger fluorescent signal, improved solubility, and greater expression in mammalian systems (Falkow et al., 1996). EGFP development provided researchers with a more powerful, longer-lasting fluorescent signal, enabling the study of protein dynamics and localization, examination of gene expression at single and multiple cell levels, and visualization of tagged organelles (Cinelli et al., 2000; Jiang et al., 2007; Utrtna & O'Byrne, 2014). By creating more experimentally favourable GFP variants, better fluorescent signals can improve the quality and range of scientific research.

EGFP is limited in the imaging of acidic organelles. Deviation from EGFP's neutral pKa of 6 triggers protonation at position 204 tyrosine, disrupting the internal hydrogen bond network and preventing fluorescence in acidic conditions (Roberts et al., 2016; Shinoda et al., 2018). The lack of acid- tolerant cellular markers prevents understanding acidic organelles and cellular processes like lysosomes and endocytosis (Roberts et al., 2016). Consequently, there is a limited comprehension of acidic organelles' mechanisms, structures, and pathophysiology, clouding the greater understanding of cell biology. To develop an acid-tolerant marker, Kneen and colleagues (1998) introduced a threonine to isoleucine mutation at position 204 to prevent protein protonation. The authors found that isoleucine changed EGFP's pKa to 4.8, enabling fluorescence in acidic conditions by preventing protonation events. It is essential to recreate the threonine to isoleucine mutation identified by Kneen and researchers (1998), as the development

of an acid-tolerant marker can extend the current knowledge of EGFP and the field of molecular biology and provide a better understanding of acidic organelles and processes.

This experiment aimed to replicate Kneen and colleagues' (1998) previously acid-stable threonine to isoleucine mutation at position 204 to identify whether expression in the BL21 (DE3) *E. coli* system would produce the same acid-stable fluorescent proteins. If a threonine to isoleucine mutation at position 204 is introduced into EGFP, then acid-tolerant proteins will be generated since isoleucine will alter the protein pKa to 4.8 and prevent protonation events that inhibit fluorescence. To generate the T204I mutant, recombinant EGFP plasmids were subjected to site-directed mutagenesis. Nickel affinity chromatography was used to purify mutant EGFP proteins, and separation by molecular weight was carried out by SDS- PAGE. Fluorescence microscopy was used to analyze fluorescence, and spectrofluorophotometry was utilized to identify whether proteins were stable in acid.

Materials and Methods

To generate mutant T204I EGFP, recombinant pET15-EGFP plasmids were provided by Dr. Kimberly Harcombe and Dr. Habib Rezanejad. Recombinant plasmids were subjected to QuikChange site-directed mutagenesis using specific, designed forward and reverse primers following procedure outlined in the lab manual (figure 1) (Harcombe, 2023). T204I mutant EGFP was then transfected into the BL21(DE3) *E. coli* strain and induced using IPTG1 reagent. Since a hexapeptide tag was engineered into the recombinant EGFP plasmid, expressed mutant proteins were purified and collected by Ni-NTA affinity chromatography. BL21(DE3) *E. coli* expressing mutant EGFP protein were lysed, and protein separation was carried out using SDS-PAGE to confirm mutational success. Transfection, induction, Ni-NTA affinity chromatography, and SDS-PAGE was carried out following provided procedure (Harcombe, 2023).

Generation of the T204I mutation was confirmed by preparing an aliquot of QuikChange mutant EGFP sample for Sanger Sequencing. T7 terminator primers provided by Dr. Kimberly Harcombe and Dr. Habib Rezanejad were used to carry out sequencing (Harcombe, 2023). Samples were sent to the Molecular Biology Services Unit in the Department of Biological Sciences at the University of Alberta. The ABI 3730 Genetic Analyzer was used to separate cycle sequencing products. Sequence analysis was conducted using CLC Genomics Workbench 23 and FinchTV following lab manual procedure (FinchTV, 2023; Harcombe, 2023; QIAGEN, 2023;).

T204I mutant EGFP fluorescence was analyzed using fluorescence microscopy. Four different samples of BL21(DE3) *E. coli* were prepared following given procedure: uninduced wildtype EGFP culture, uninduced mutant T204I EGFP culture, induced wildtype EGFP culture, and induced T204I mutant EGFP (Harcombe, 2023). The Olympus BX51 TRF fluorescence microscope at MacEwan University was used to assess fluorescence, and cellSense Standard software was used for image capturing (Olympus, 2023).

Spectrofluorophotometry was used to assess the T204I mutational success. pH 7, 6, and 5 solutions were created to test mutant EGFP stability in acidic conditions. A pH 7 solution was prepared following procedure (Harcombe, 2023). A pH 6 solution was prepared by combining 3mL mQH2O, 0.263 mL NaH2PO4, and 0.036 mL Na2H2PO4, provided by Jennifer Bubenko of MacEwan University. 1mL of pH 6 solution was added to a mixture of 0.5 mL purified T204I mutant EGFP from QuikChange mutagenesis and 1.5 mL of buffer B. pH was tested, and NaH2PO4 was added in 0.500 mL increments until pH 6 was achieved. Preparation of pH 5 solution was carried out by combining 3mL mQH2O, 0.263 mL NaH2PO4, and 0.036 mL Na2H2PO4, with an additional 1 mL NaH2PO4 to achieve final pH 5. 1 mL of pH 5 solution was added to a mixture of 0.5 mL purified T204I mutant EGFP from QuikChange mutagenesis and

1.5 mL of buffer B. pH was tested, and NaH2PO4 was added in 0.500 mL increments until pH 5 was achieved. The Shimadzu RF-5301 PC Spectrophotometer at MacEwan University was used, and the RF-530XPC program was utilized to obtain maxima values (Shimadzu, 2023).

Results

DNA sequencing and sequence chromatography revealed the quality of mutant EGFP DNA (figure 2). Figure 2A illustrates only the threonine to isoleucine mutation at position 204, with no additional mutations in the sequenced EGFP DNA. Figure 2B illustrates tall, sharp, well-separated chromatography peaks. No evidence of secondary peaks, problematic sequences, artifacts, or ambiguities suggests good quality sequenced EGFP DNA.

SDS-PAGE provided a high-resolution separation of proteins in BL21 (DE3) *E. coli* samples (figure 3). EGFP proteins separated to 29 kDa (figure 3). No EGFP band was present in the *E. coli* cell pellet (figure 3, lane 1). Mutant EGFP proteins were found following IPTG induction (figure 3, lanes 3- 6). EGFP proteins were found in *E. coli* cell lysis and various Ni-NTA affinity chromatography washes (figure 3, lanes 7-9). Elution of EGFP proteins yielded the thickest EGFP band (figure 3, lane 10).

Fluorescence microscopy was utilized to identify differences in fluorescence between induced and uninduced *E. coli* and wild-type and mutant *E. coli* (Figure 4). Uninduced wild type and mutant EGFP produced faint fluorescent signals (figure 4A; 4C). Induced wildtype and mutant EGFP generated bright signals that were non-differentiable (figure 4B, 4D).

Spectrofluorophotometry compared the excitation and emission spectra between wild-type and mutant EGFP at pH 7, 6, and 5 (table 1). As depicted by table 1, expected excitation and emission maxima were obtained for wildtype EGFP. However, mutant EGFP illustrated

unexpected deviations, like two excitation maxima at pH 7. Table 1 indicated that mutant proteins were brighter at pH 7 and 6, while wild-type proteins were brighter at pH 5.

Discussion and Conclusion

To obtain T204I mutant EGFP proteins, BL21(DE3) *E. coli* was utilized. This strain was significant for EGFP expression because of its high transformation efficiency, rapid growth, and high quantity of proteins produced (Kim et al., 2017). Additionally, the BL21(DE3) strain contains the T7 promoter system. Upstream of the EGFP gene lies the T7 promoter, which silences gene activity until the activation of T7 RNA polymerase (Tabor, 1990; Du et al., 2021). The T7 RNA polymerase gene is regulated by the lac repressor, an inhibitor produced by the lac operon found within the *E. coli* chromosome (Harcombe, 2023). The lac repressor prevents the synthesis of the T7 RNA polymerase, regulating the expression of the gene downstream of the T7 promoter (Harcombe, 2023). Treating BL21(DE3) *E. coli* with IPTG results in EGFP expression, as IPTG inhibits the lac repressor, synthesizing the T7 RNA polymerase and binding to the T7 promoter (Harcombe, 2023). The BL21 (DE3) *E. coli* T7 promoter system was effective because no EGFP expression occurred without T7 RNA polymerase, and prevention of selection against plasmid maintenance occurred, generating a high yield of proteins when induced.

Mutagenic PCR was performed using designed mutagenic primers (figure 1). The mutagenic primers ensured the mutation of only threonine and prevented additional mutations from occurring. DNA sequencing confirmed the success of mutagenesis by analyzing quality and coverage. Figure 2A aligns the sequenced mutant EGFP DNA and mutagenesis PCR products with the desired EGFP sequence. The alignment illustrated the desired threonine to isoleucine at position 204, indicating the success of site- directed mutagenesis. No additional mutations were

present in the alignment, suggesting no unwanted fluorescence or protein structural changes will occur. Figure 2B sequence chromatogram indicates tall and well-separated peaks with no artifacts or ambiguities within the EGFP DNA. The observed chromatogram indicates good-quality DNA. However, there was 1X coverage because only the T7 terminating primer was used to sequence each nucleotide once. Reduced coverage limited the ability to conclude that the sequenced DNA was correct (Xu et al., 2018). If ambiguities and artifacts were present on the chromatogram, there was no secondary sequence to compare and identify mistakes. However, the lack of ambiguities, artifacts, and distinct, sharp peaks suggested that the EGFP sequence obtained was accurate and mutated correctly.

SDS-PAGE was used to separate EGFP proteins based on size to further identify mutagenesis success. It was expected that proteins would separate around 29 kDa because the mutation of threonine to isoleucine, along with the hexapeptide histidine tag, would increase the molecular weight from 26 kDa to 29 kDa (Kneen et al., 1998). As depicted in figure 3, the EGFP protein band was found at 29 kDa. This indicates no additional mutations that could alter molecular weight. If the band was found at a different location, it could be deduced that additional mutations were present to alter the molecular weight. Since the EGFP band separated to the expected weight, it was confirmed that mutagenesis was successful.

Figure 3 illustrates the success of EGFP protein production after induction. The uninduced cell pellet in lane 2 exhibited no significant 29 kDa band (figure 3). BL21 (DE3) *E. coli* were grown in the absence of IPTG, preventing the production of EGFP. The T7 RNA polymerase was inhibited by the lac repressor produced by the lac operon, preventing the synthesis of proteins. Adding IPTG to *E. coli* produced a 29 kDa band, as seen in lanes 3-6, indicating successful induction and protein expression (figure 3). IPTG prevented the production

of the lac repressor, encouraging T7 RNA polymerase production and binding to the T7 promoter upstream of the mutant EGFP gene (Du et al., 2021; Harcombe, 2023; Tabor, 1990). EGFP bands became thicker as time lapsed, indicating continuous induction and accumulation of proteins, with the most induction occurring after 120 minutes (figure 3).

Additionally, SDS-PAGE indicated successful purification of EGFP proteins during Ni-NTA affinity chromatography (figure 3). Mutant EGFP proteins expressed a hexapeptide histidine tag that interacted with the nickel column, encouraging EGFP protein binding and preventing non-specific interactions (Harcombe, 2023). Lane 10 illustrates that the disruption of the His-tag and nickel bond causes the elution of EGFP proteins off the column (figure 3). However, as illustrated in lanes 8 and 9, washing the column resulted in the loss of some EGFP proteins (figure 3). EGFP loss occurred due to contaminants outcompeting EGFP for nickel binding and breaking of weak histidine-nickel bonds (Pina et al., 2014). Although some EGFP protein was lost, the majority was collected during column elution, producing proteins ready for mutational success analysis.

Fluorescence microscopy was used to identify differences in fluorescence between BL21 (DE3) *E. coli* samples. Figures 4A and 4C illustrate that uninduced wild-type and mutant *E. coli* produced faint fluorescent signals. This is because of the leaky expression of the T7 promoter system in BL21 (DE3) *E. coli.* The T7 RNA polymerase has basal expression due to imperfect control of the lac operon (Spehr et al., 2000). Infrequently, the lac operon can naturally become repressed, producing the T7 RNA polymerase (Spehr et al., 2000). Consequently, there was minute EGFP protein production in the absence of IPTG, resulting in a faint fluorescent signal, as seen in figures 4A and 4C. Following IPTG induction, a strong, intense fluorescence signal was produced, as complete control of the lac operon occurred. The fluorescent signal produced

by induced wild type and mutant EGFP was non-distinguishable, both intense, bright green signals. There was no change in fluorescence signal observed at pH 7 because the threonine to isoleucine mutation did not change the structure of the chromophore (Kneen et al., 1998).

The acid tolerance of the mutant EGFP was tested by analyzing and comparing the excitation and emission maxima to wild-type EGFP in pH 7, 6, and 5 (table 1). Mutant EGFP proteins at pH 7 were excited by a 397 nm wavelength, which was shifted from the wild-type excitation maximum of 491 nm. This is because the isoleucine at position 204 altered the stability of the chromophore, requiring a higher energy wavelength to become excited (Kneen et al., 1998). Additionally, the production of two excitation maxima illustrated the presence of a neutral and ionized form of EGFP (table 1) (Walker et al., 2021). The 397 nm excitation maximum correlates to the neutral form of EGFP, while the 508 nm represents the ionized EGFP. Usually, ultraviolet radiation excites electrons, keeping EGFP neutral (Brejc et al., 1997; Drobizhev et al., 2015; Walker et al., 2021). However, the absorption of excess UV radiation triggered the loss of electrons within the chromophore, leading to EGFP ionization (Brejc et al., 1997; Drobizhev et al., 2015; Walker et al., 2021). The ionized form then required a higher wavelength of energy to become excited, as seen in the 508 nm wavelength in table 2 (Drobizhev et al., 2015). The emission maximum of mutant EGFP was 515 nm, different from the wild type 509 nm because the isoleucine destabilized the chromophore, requiring a higher energy wavelength for excitation (Kneen et al., 1998).

Further testing of mutant EGFP in acid depicted the failure of the proteins to fluoresce in acid. At pH 6 and 5, excitation shifts from the expected 397 nm to 503 nm and 509 nm, respectively, as expected (Kneen et al., 1998). This indicated the further destabilization of the chromophore, requiring more energetic wavelengths for excitation (Drobizhev et al., 2015). At

pH 6, fluorescence still occurred at the expected 515 nm range. However, at pH 5, fluorescence was lost as the excitation maximum was undetected (table 2). The significant loss of fluorescence in acid can be attributed to the complete destabilization of EGFP's hydrogen bond network. The original threonine 204 was required to form a hydrogen bond with histidine 148 (Brejc et al., 1997). This bond contributed to the stabilization of EGFP and the chromophore. Mutation of threonine removed the bond and destabilized the protein. Additionally, threonine 204 donated a hydrogen bond to tyrosine 66 of the chromophore, further stabilizing the protein (Agmon, 2005; Brejc et al., 1997). Introducing isoleucine at position 204 removed the hydrogen bond, destabilizing the chromophore. The mutation at position 204 also impacted the neighbouring serine 205 required to interact with tyrosine 66 (Agmon, 2005; Aslopovsky et al., 2023). Isoleucine altered the shape of EGFP, preventing serine 205 from creating hydrogen bonds with tyrosine 66 (Agmon, 2005). Finally, the tyrosine to isoleucine mutation blocked a natural proton pathway channel in EGFP (Agmon, 2005). The threonine 204 created a channel that allowed the movement of protons out of the internal protein environment (Agmon, 2005). Introducing isoleucine blocked the channel, causing an internal proton buildup that disrupts the chromophore's hydrogen bonding network (Agmon, 2005). The threonine to isoleucine mutation failed to generate acid tolerant EGFP proteins because of disruption of the protein’s hydrogen bond network.

In the present report, it is evident that the threonine to isoleucine mutation at position 204 fails to produce acid tolerant EGFP proteins. Position 204 threonine is required to stabilize and encourage hydrogen bonding within the chromophore and whole protein (Agmon, 2005; Brejc et al., 1997; Walker et al., 2021). Substituting threonine with isoleucine disrupts the hydrogen bonding network, resulting in protein destabilization and loss of function in acidic conditions

(Agmon, 2005; Brejc et al., 1997; Walker et al., 2021). Future directions include confirming the disrupted hydrogen bond network using X-ray crystallography and fluid chromatography to identify protein structure and folding (Goetz et al., 2014; Rahat et al., 2009;). Additionally, removing the hexapeptide histidine tag using protease can be pursued to determine if acid stability would differ. Identification of different mutations that successfully cause acid tolerance is essential. Currently, there is a lack of acid tolerant cellular markers that can be used to visualize acidic organelles and processes. As a result, there is a poor comprehension of the mechanisms and processes that occur in cellular conditions and organelles. Creating acid-tolerant cellular markers will enable researchers to study and understand acidic conditions deeply. The greater field of science will be improved as a better understanding of cells and physiology will be obtained, and researchers have access to better in vivo tools, bringing forth a greater possibility of research in different fields.

References

Agmon, N. (2005). Proton pathways in green fluorescent protein. *Biophysical Journal, 88*(4), 2452-2461. 10.1529/biophysj.104.055541

Arpino, J. A. J., Reddington, S. C., Halliwell, L. M., Rizkallah, P. J., & Jones, D. D. (2014). Random single amino acid deletion sampling unveils structural tolerance and the benefits of helical registry shift on GFP folding and structure. *Structure, 22*(6), 889-898. 10.1016/j.str.2014.03.014.

Aslopovsky, V. R., Scherbinin, A. V., Kleshcina, N. N., & Bochenkova, A. V. (2023). Impact of the protein environment on two-photon absorption cross-sections of the GFP chromophore anion resolved at XMCQDPT2 level of theory. *International Journal of Molecular Sciences, 24*(14), 1-14. https://doi.org/10.3390/ijms241411266

Brejc, K., Sixma, T. K., Kitts, P. A., Kain, S. R., Tsien, R. Y., Ormo, M., & Remington, S. J. (1997). Structural basis for the dual excitation and photoisomerization of the *Aequorea victoria* green fluorescent protein. *Proceedings of the National Academy of Sciences USA, 94*(6), 2306-2311. 10.1073/pnas.94.6.2306

Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W., & Prasher, D. C. (1994). Green fluorescent protein as a marker for gene expression. *Science, 263*(5148), 802-805. DOI: 10.1126/science.8303295

Chalfie, M. (1995). Green fluorescent protein. *Photochemistry and Photobiology, 64*(4), 651-656. https://doi.org/10.1111/j.1751-1097.1995.tb08712.x

Cinelli, R. A., Ferrari, A., Pellegrini, V., Tyagi, M., Giacca, M., & Beltram, F. (2000). The enhanced green fluorescent protein as a tool for the analysis of protein dynamics and

localization: local fluorescent study at the single-molecule level. *Photochemistry and Photobiology, 71*(6), 771-776. 10.1562/0031-8655(2000)071<0771:tegfpa>2.0.co;2

Corish, P., & Tyler-Smith, C. (1999). Attenuation of green fluorescent protein half-life in mammalian cells. *Protein Engineering, Design, and Selection, 12*(12), 1035-1040. https://doi.org/10.1093/protein/12.12.1035

Drobizhev, M., Callis, P. R., Nifosi, R., Wicks, G., Stoltzfus, C., Barnett, L., Hughes, T. E., Sullivan, P., & Rebane, A. (2015). Long- and short-range electrostatic fields in GFP mutants: implications for spectral tuning. *Scientific Reports, 5*(13223), 1-14. https://doi.org/10.1038/srep13223

Du, F., Liu, Y. S., Xu, Y. S., Li, Z. J., & Sun, X. M. (2021). Regulating the T7 RNA polymerase expression in *E. coli* BL21 (DE3) to provide more host options for recombinant protein production. *Microbial Cell Factories, 20*(189), 1-10. doi.org/10.1186/s12934-021-01680-6

Falkow, S., Cormack, B. P., & Valdivia, R. H. (1996). FACS-optimized mutants of the green fluorescent protein (GFP). *Gene*, *173*(1), 33-38. 10.1016/0378-1119(95)00685-0.

FinchTV. (2023). *FinchTV software* (Version 1.5.0) [Computer software]. https://digitalworldbiology.com/FinchTV

Follenius-Wund, A., Bourotte, M., Schmitt, M., Iyice, F., & Pigault, C. (2003). Fluorescent derivatives of the GFP chromophore give a new insight into the GFP fluorescence process. *Biophysical Journal, 85*(3), 1839-1850. 10.1016/S0006-3495(03)74612-8.

Goetz, G.H., Farrell, W., Shalaeva, M., Sciabola, S., Anderson, D., Yan, J., Philippe, L., & Shapiro, M. J. (2014). High throughput method for the indirect detection of

intramolecular hydrogen bonding. *Journal of Medicinal Chemistry, 57I*(7), 2920-2929. 10.1021/jm401859b

Harcombe, K. (2023). *Fall 2023 Lab Manual: Cloning, Mutagenesis, and Expression of Enhanced Green Fluorescent Protein (EGFP)*. MacEwan University.

Jain, A., Blum, C., & Subramaniam, V. (2009). Chapter 4 – fluorescence lifetime spectroscopy and imaging of visible fluorescent proteins. In P. Verdonck (Ed.), *Advances in Biomedical Engineering* (pp. 147-176). Elsevier Science.

Jiang, H., Zhang, J., Shi, B. Z., Xu, Y. H., Li, Z. H., & Gu, J. R. (2007). Application of EGFP-EGF fusions to explore mechanism of endocytosis of epidermal growth factor. *Acta Pharmacologica Sinica, 28*(1), 111-117.10.1111/j.1745-7254.2006.00481.x

Kain, S. R. (1999). Green fluorescent protein (GFP): applications in cell-based assays for drug discovery. *Drug Discovery Today, 4*(7), 304-312. 10.1016/S1359-6446(99)01330-6.

Kennis, J. T. M., Larsen, D. S., Van Stokkum, I. H. M., Vengris, M., Thor, J. J., & Van Grondelle, R. (2004). Uncovering the hidden ground state of green fluorescent protein. *Proceedings of the National Academy of Sciences, 101*(52), 17988-17993. 10.1073/pnas.0404262102.

Kneen, M., Farinas, J., Li, Y., & Verkman, A. S. (1998). Green fluorescent protein as a non-invasive intracellular pH indicator. *Biophysical Journal, 74*(3), 1591-1599. 10.1016/S0006-3495(98)77870-1.

Olympus. (2023). *Olympus cellSens Standard imaging software* (Version 4.2.1) [Computer software]. https://www.olympus-lifescience.com/en/software/cellsens/

Pina, A. S., Lowe, C. R., & Roque, A. C. A. (2014). Challenges and opportunities in the purification of recombinant tagged proteins. *Biotechnology Advances, 32*(2), 366-381. 10.1016/j.biotechadv.2013.12.001

QIAGEN (2023). *QIAGEN CLC Genomics Workbench computer software* (Version 23.0.1) [Computer software]. https://digitalinsights.qiagen.com/downloads/product-downloads/

Rahat, O., Alon, U., Levy, Y., & Schreiber, G. (2009). Understanding hydrogen-bond patters in proteins using network motifs. *Bioinformatics, 25*(22), 2921-2928. https://doi.org/10.1093/bioinformatics/btp541

Remington, S. J. (2011). Green fluorescent protein: a perspective. *Protein Science, 20*(9), 1509-1519. 10.1002/pro.684

Roberts, T. M., Rudolf, F., Meyer, A., Pellaux, R., Whitehead, E., Panke, S., & Held, M. (2016). Identification and characterisation of a pH-stable GFP. *Scientific Reports, 6*(28166), 1-9. 10.1038/srep28166

Rosochacki, S. J., & Matejczky, M. (2002). Green fluorescent protein as a molecular marker in microbiology. *Polish Journal of Microbiology, 51*(3), 205-216.

Shimadzu. (2023). *LabSolutions RF software for Spectrofluorophotometer* (Version 5) [Computer software]. https://www.shimadzu.com/an/products/molecular-spectroscopy/fluorescence/fluorescence-spectroscopy-software/ilabsolutions-rf/index.html

Shinoda, H., Ma, Y., Nakashima, R., Sakurai, K., & Nagai, T. (2018). Acid-tolerant monomeric GFP from *Olindias formosa*. *Cell Chemical Biology, 25*(3), 330-338. 10.1016/j.chembiol.2017.12.005.

Spehr, V., Frahm, D., & Meyer, T. F. (2000). Improvement of the T7 expression system by the use of T7 lysozyme. *Gene, 257*(2), 259-267. doi/10.1016/S0378-1119(00)00400-5

Tabor, S. (1990). Expression using the T7 RNA polymerase/promoter system. *Current Protocols in Molecular Biology, 16*(11), 1-16. 10.1002/0471142727.mb1602s11

Tsien, R. Y. (1998). The green fluorescent protein. *Annual Review of Biochemistry, 67*(1), 509-544. 10.1146/annurev.biochem.67.1.509.4

Walker, A., Silva, N., & Marcey, D. (2021). Green fluorescent protein. https://earth.callutheran.edu/Academic_Programs/Departments/BioDev/omm/jsmolnew/gfp_new/gfp.html#:~:text=III.,The%20GFP%20Chromopore,475nm%20in%20its%20ionized%20form.

Ward, W. W. (2005). Biochemical and physical properties of green fluorescent protein. In M. Chalfie, & S.R. Kain (Ed.), I*n Green Fluorescent Protein: Properties, Applications, and Protocols* (pp 39-65). John Wiley and Sons.

Xu, C., Wu, K., Zhang, J., Shen, H., & Deng, H. W. (2018). Low-, high-coverage and two-stage DNA sequencing in the design of the genetic association study. *Genetic Epidemiology, 41*(3), 187-197. 10.1002/gepi.22015.

Figures and Tables

Figure 1: Analysis of wildtype EGFP, T204I mutant EGFP, and designed primer sequences. CLC Genomics Workbench 23 was utilized to identify and mutate position 204 threonine, and design mutagenic primers for QuikChange (QIAGEN, 2023). The EGFP sequence was provided by Dr. Kimberly Harcombe and Dr. Habib Rezanejad. *1A*: protein sequence of wildtype EGFP. *1B:* desired T204I mutant EGFP protein sequence. *Boxes* indicate desired nucleotide. *Numbers* correspond to nucleotide position. *Asterix* illustrate stop codon. *1C*: designed forward primer nucleotide sequence. *1D:* designed reverse primer nucleotide sequence. *Capitalized Letters* illustrates mutant nucleotide. *5' and 3'* represent 5' and 3' ends of the primers. *Numbers* illustrate nucleotide position.

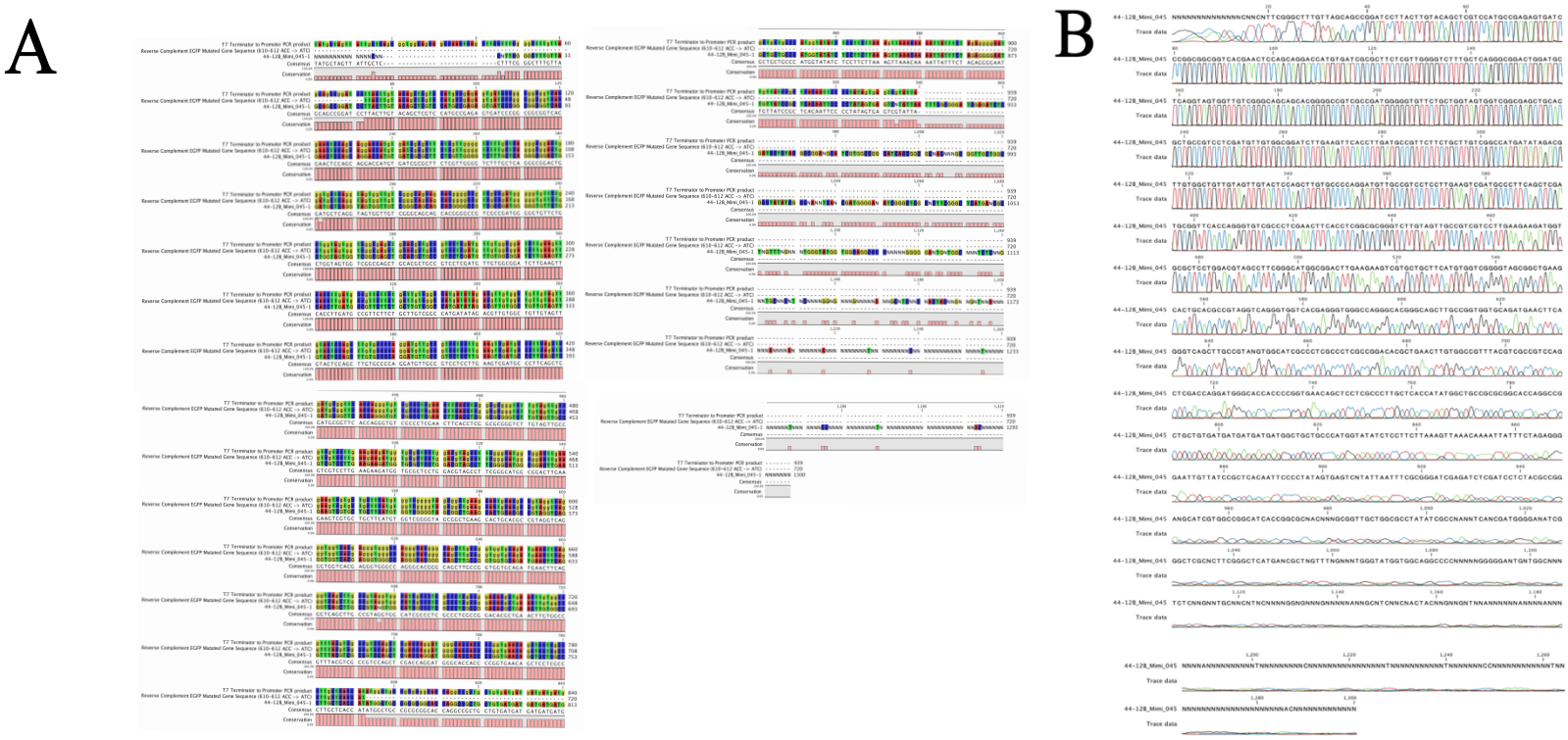

Figure 2: Sequence and quality analysis of mutant EGFP DNA. Sanger sequencing was utilized to determine mutant EGFP DNA sequence. Samples were sent to the Molecular Biology Services Unit in the Department of Biological Sciences at the University of Alberta. The ABI 3730 Genetic Analyzer was used to separate cycle sequencing products. CLC Genomics Workbench 23 was used to obtain the sequence alignment and chromatogram (QIAGEN, 2023). *1A:* alignment of PCR products using T7 terminator primers, reverse complement of mutant EGFP DNA, and Sanger sequenced 44-12B_Mimi_045-1 DNA. *Consensus* represents the most common nucleotide. *Conservation* indicates the percent of amino acid conservation. *1B:* sequence chromatogram of mutant EGFP. *44-12B_Mimi_045* represents DNA sequence. *Trace data* illustrates sequence chromatogram. *Numbers* represent nucleotide position.

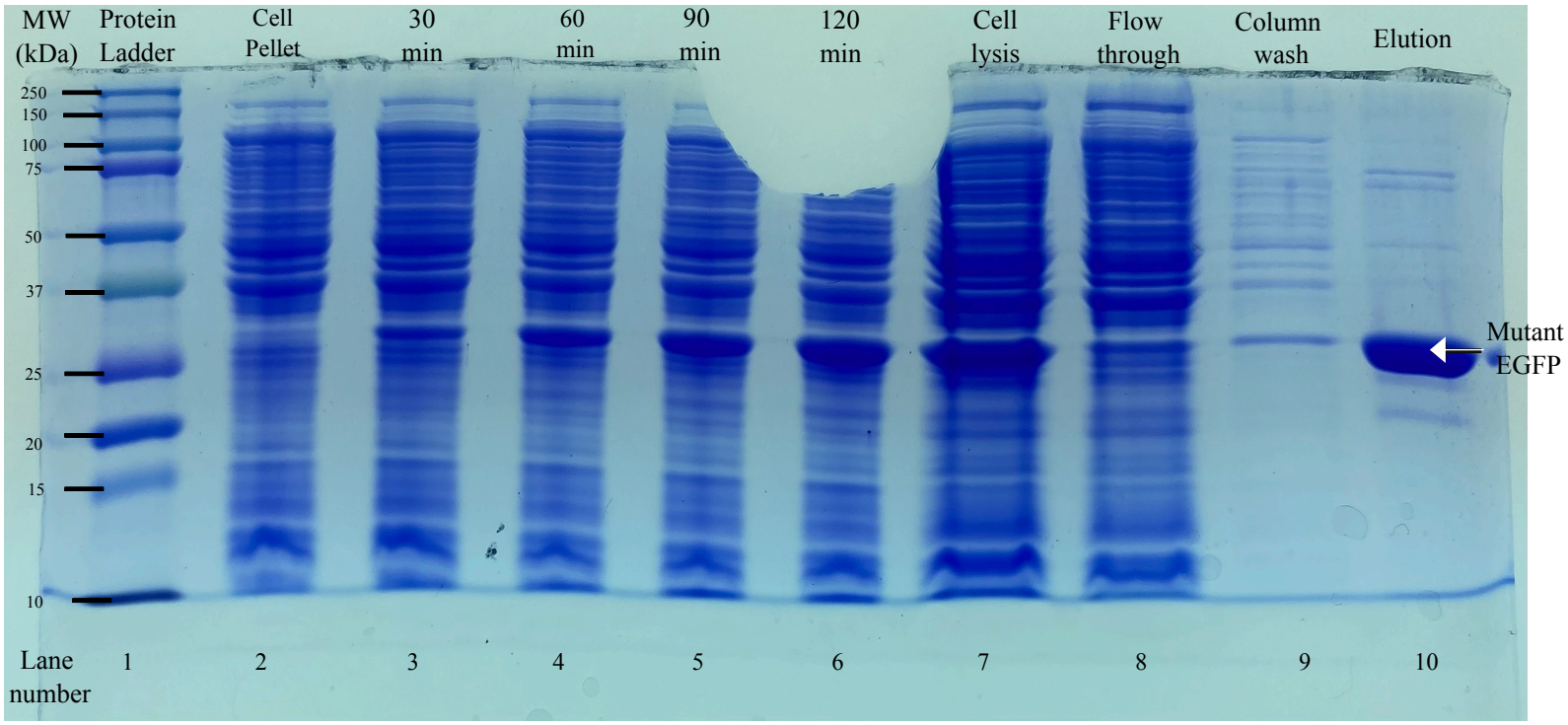

Figure 3: Analysis of induction and purification of mutant EGFP proteins from BL21 (DE3) *E. coli*. SDS-PAGE was utilized to stain and separate mutant EGFP proteins by size with Coomassie blue. *Dashes* correspond to molecular mass. *Arrow* indicates desired mutant EGFP. *Lane number* corresponds to lane numbers. Lanes are as follows: *ladder,* 10 uL Precision Plus Protein Kaleidoscope Standard; *cell pellet,* 10 uL *E. coli* pellet; *30–120-minute*, 25 uL of purification after 30-minute increments of IPTG induction; *cell lysis*, 25 uL *E. coli* lysis; *flow through,* 25 uL of flow through from nickel-affinity chromatography; *column wash,* 25 uL of column wash collected during nickel-affinity chromatography; *elution*, 25 uL of purified mutant EGFP following nickel-affinity chromatography.

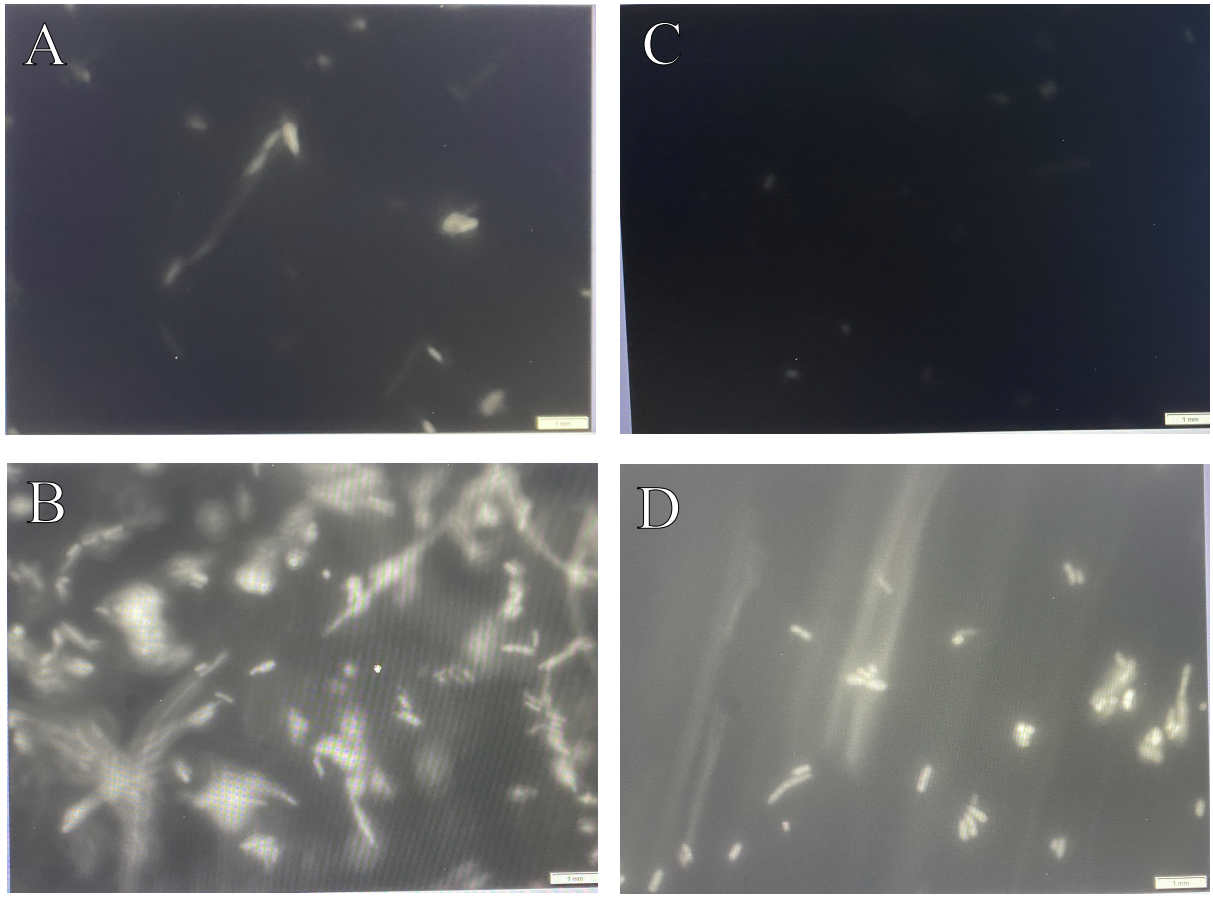

Figure 4: Examination of induced and uninduced BL21 (DE3) *E. coli* expressing either wildtype or acid-tolerant EGFP. An Olympus BX51 TRF fluorescence microscope was used to visualise fluorescence intensity. Letters are as follows: *A*, uninduced wild-type *E. coli*; *B*, induced wild-type *E. coli*; *C*, uninduced T204I mutant EGFP *E. coli*; *D*, induced T204I mutant EGFP *E. coli*.

Table 1: excitation maxima, emission maxima, and specific activity of wildtype and mutant DL21 (DE3) *E. coli* at pH 7, 6, and 5.

E. coli Type	pH	Excitation maxima (nm)	Emission maxima (nm)	Specific Activity
Wildtype	7	491	509	1.69
Wildtype	6	492	509	0.421
Wildtype	5	489	508	0.125
Mutant	7	397 508	515	3.08
Mutant	6	503	512	3.02
Mutant	5	509	0	0