

Gene Expression: A Timer for *In Vivo* Venus Maturation

Kirsten Frieda

Dr. Jie Xiao/Professor X. Sunney Xie Group
Harvard REU 2003

Abstract

Real time fluorescence imaging within living cells offers an intriguing glance into the basic mechanisms that provide a foundation for life. Such visualization techniques apply well to the proteins Venus (a variant of yellow fluorescent protein) and β -galactosidase (β -gal). Venus is self-fluorescing; however, it matures slowly to produce its fluorophore after translation. In contrast, β -gal generates fluorescence almost instantaneously by catalyzing the hydrolysis of a fluoregenic substrate. We propose to measure Venus's *in vivo* maturation time using a fusion protein of Venus and β -gal, in which the fluorescence from Venus will be measured against the fluorescence generated by β -gal. Toward this end, variants of a *venus-lacZ* fusion gene were constructed. Current tests are evaluating fluorescence signal increase upon induction of the fusion gene. Accordingly, adjustments in the regulation system to decrease basal level expression and increase overall induction are under way.

Introduction

In studying gene expression, fluorescence imaging is a useful technique for detecting proteins expressed in living cells. Two proteins that lend themselves well to this method of detection are Venus and β -galactosidase (β -gal). Venus is an improved variant of yellow fluorescent protein (YFP)¹ and, ultimately, of green fluorescent protein (GFP)² (Figure 1) that was originally isolated from the jellyfish *Aequorea victoria*. Like other GFPs, Venus is self-fluorescing. Its internal residues oxidize to produce its fluorophore with excitation/emission maxima of 515/528 nm¹.

β -gal is coded by the *lacZ* gene. Normally part of the *lac* operon, β -gal aids in the metabolism of lactose. Though it does not self-fluoresce, β -gal can be detected via its hydrolysis of the fluoregenic substrate DDAO-gal to free fluorescent DDAO with excitation/emission maxima of 645/660 nm (Figure 2).

While β -gal almost instantaneously frees DDAO to give a fluorescence signal upon translation, Venus takes some time to produce its internal fluorophore. Though Venus's maturation at 37°C is one and a half minutes as measured *in vitro*¹, much faster than previous YFP variants (about half an hour to two hours), no published data of its maturation time *in vivo* is available. We plan to use Venus in imaging gene expression and would like to

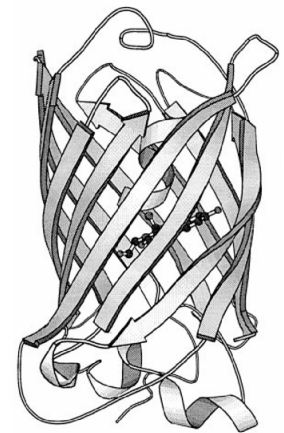


Figure 1. Structure of GFP with its internal fluorophore shown at the center²

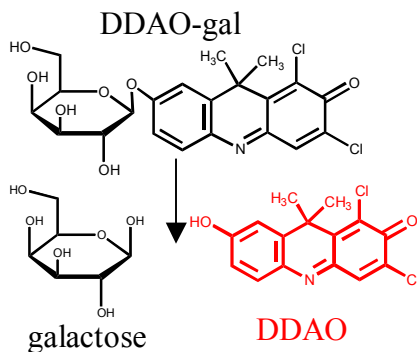


Figure 2. Hydrolysis of DDAO-gal to release fluorescent DDAO

understand the *in vivo* maturation of Venus, since the maturation kinetics may be different *in vivo*. Initially we hope to measure an ensemble average of Venus maturation times based on a simple kinetic model using a fluorometer. Ultimately, we would like to time maturation in single cells by fluorescence imaging using a microscope.

Toward this end, we have linked *venus* and *lacZ* to create a fusion gene. The two components of this gene will be transcribed and translated together to give a fusion protein of Venus and β -gal. β -gal's quick cleavage of DDAO-gal and the corresponding fluorescence signal can then mark the fusion protein's creation, thus serving as a starting point for timing Venus's maturation to generate its fluorophore. With these considerations in mind, I constructed this fusion gene and initially put it under the control of P_{tac} promoter and *lac* repressor. Under uninduced conditions (basal-level expression), the *lac* repressor protein binds to the *lacO* site in front of the gene, preventing the transcription of the gene by the RNA polymerase binding at the promoter. When an inducer is added, the binding affinity between the *lac* repressor and *lacO* site is greatly decreased, and thus the gene is massively transcribed and the protein produced.

An additional consideration in constructing this fusion protein involves its degradation. The microscope experiment aims to visualize gene expression via fluorescence bursts resulting from the stochastic event of the repressor temporarily dropping off the DNA at basal level (i.e., uninduced). This necessitates efficiently degrading previously-formed proteins so that their signal does not obscure detection of newly-formed ones. Naturally long-lived, β -gal, for example, has an *in vivo* half-life of ~20 hours or more. Use of Ubiquitin (Ub) and application of the N-end rule in the fusion gene construction can reduce protein half-life to minutes. Ubiquitin is fused at the protein's N terminus, and can then be cleaved off by ubiquitin-specific protease to expose a certain destabilizing amino acid at the protein's N-terminal. This specific amino acid helps signal the protein for degradation, as described by the N-end rule which relates a protein's stability to its terminal amino acid. The N-end amino acids Arginine and Leucine used in this work confer protein half-lives of ~2 and ~3 minutes respectively^{3,4}. To address degradation concerns, the fusion genes created here include *ubiquitin* and the codon for a specific amino acid before both *venus* and *lacZ*.

Methods and Materials

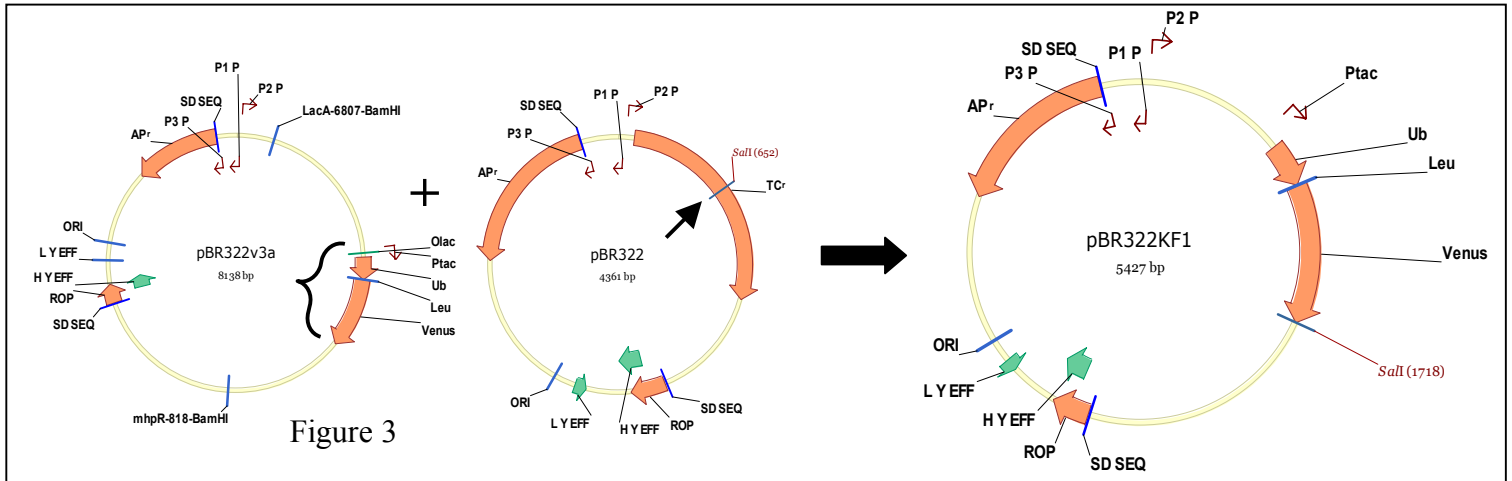
Gene Construction

The cloning work was carried out using standard protocol described in *Molecular Cloning: A Laboratory Manual*⁵. The typical procedure for gene insert preparation involved PCR amplification using Pfu DNA polymerase, gel purification and extraction of the desired amplification fragment, enzyme digestion of DNA fragments to produce overhangs complementary to the vector when appropriate, and phosphorylation of blunt ended inserts. Plasmid vectors were prepared for ligation with the insert by enzyme digestion to open the plasmid or reverse PCR followed by enzyme digestion, gel purification and extraction, and dephosphorylation. The gel extractions and the purifications following enzymatic reactions were accomplished with the QIAquick Gel Extraction and PCR Purification kits (Qiagen) respectively.

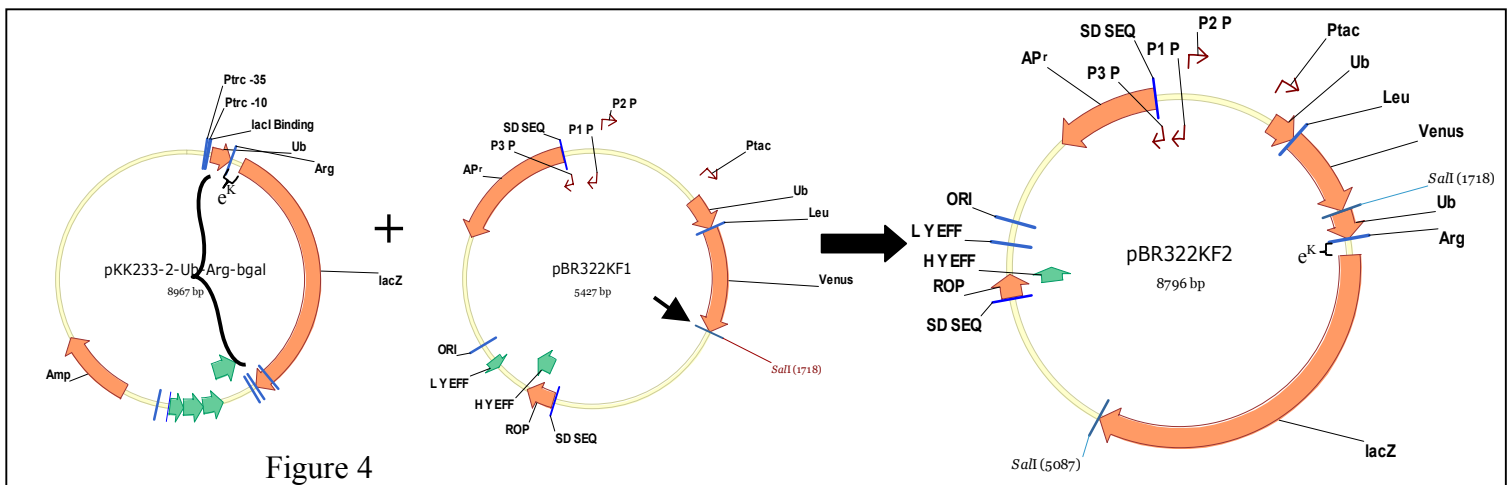
Ligation of the vector and insert for pBR322 derivatives was performed with New England Biolab's Quick Ligation protocol. Transformation of DH5 α competent cells with the ligation mixture followed standard procedure⁵ with heat shock for 45 seconds at 42°C. Ligation

and transformation for the pBAD202 plasmid proceeded according to the pBAD202 Directional TOPO Expression Kit (Invitrogen).

To begin, the *ub-venus* gene controlled by P_{tac} promoter and O_{lac} from pBR322v3a was first PCR amplified (primers: 5'-NNATTCTCGAGTCATAACGGTTCTGGCAAATATTC and 5'-NNATTGTCGACCTTGACAGCTCGTCCATGC). This fragment had a *XhoI* cutting site at its 5' end and a *Sall* site at its 3' end and was inserted into a pBR322 vector at the *Sall* cutting site. (pBR322 was used as a host for this cloning work because it is a common, relatively easy to work with, and low copy number plasmid.) The resulting intermediate-step plasmid is named pBR322KF1. (Figure 3)

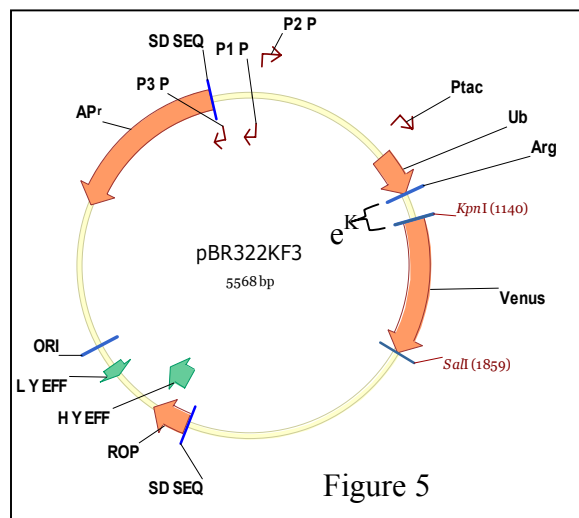


Immediately following *venus* on pBR322KF1, the *ub-e^K-lacZ* gene was inserted at the *Sall* site. This *lacZ* gene was copied from pKK233-2(Ub-Arg-β-gal) (primers: 5'-AATGGTCGACATGCAGATTTTCGTC AAGACTTTG and 5'-AATGGTCGACTTTTTGACACCAGACCAACTG). The resulting plasmid, pBR322KF2, contains the fusion gene *ub-venus-ub-e^K-lacZ* controlled under P_{tac} and O_{lac} . (Figure 4)

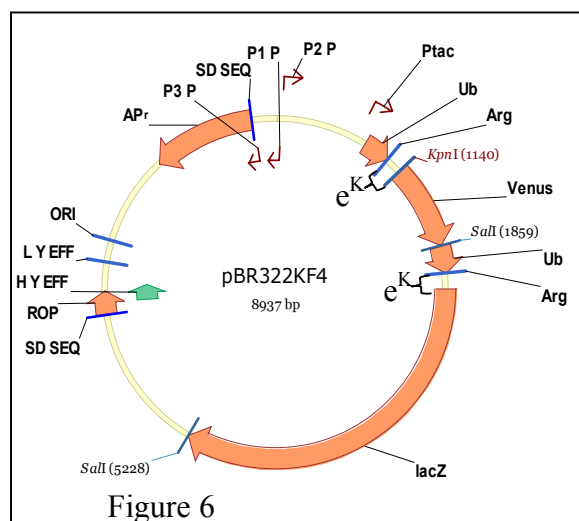


The e^K sequence between *ub* and *lacZ* on the fragment copied from pKK233-2(Ub-Arg- β -gal) is an extra, random sequence that was found to be important in yeast for degradation⁶. It may also, however, be important in *E. coli* to aid in degradation. By providing an extra, unfolded sequence that exhibits the destabilizing N-end rule amino acid at one end and connects to the protein at the other, it might help the protein's end insert into the narrow protease channel for degradation. The protease is thought to grab the protein and drag its unfolded form through the channel^{7,8}.

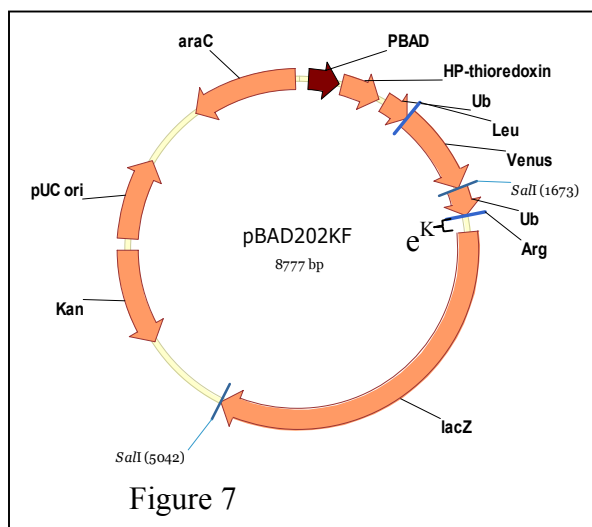
Since the e^K sequence may be beneficial, it was also added between *ub* and *venus*. (Later comparison of the degradation of the protein lacking the extra e^K sequence and that possessing it can further illustrate its actual importance.) Reverse PCR was used to copy the pBR322KF1 template minus *ub* (primers: 5'-TACCGAATTCTGTTTCCTGTGTGAAATTGTTATCCG and 5'-TAATGGTACCTTATCCAGCAAGGGCGAGGAG). *ub-e^K* from pKK233-2(Ub-Arg- β -gal) (primers: 5'-ATGCAGATTTTCGTCAAGACTTTGACCG and 5'-TAATGGTACCTCGGGAAACCTGTCGTGCC) was then inserted into the opened pBR322KF1 using one blunt end and an enzyme-digested *KpnI* site. The addition of the e^K sequence between *ub* and *venus* yielded pBR322KF3. It can be used for studies of Venus alone where the e^K sequence is desired and serves as the precursor to fusing the *lacZ* gene after *venus*. (Figure 5)



Insertion of *ub-e^K-lacZ* after *venus* proceeds by the same method yielding pBR322KF2. This gives the slightly modified fusion gene on pBR322KF4. (Figure 6)



The previously described gene constructions were under the control of the P_{lac} promoter and *lac* operator. A plasmid with the fusion gene under control of a different promoter, the P_{BAD} promoter, was also created to address issues described later. This work used the pBAD202 Directional TOPO Expression Kit (Invitrogen) to ligate the fusion gene from pBR322KF2 (primers: 5'-CACCATGCAGATTTTCGTCAAGACTTTGAC and 5'-GCACTATCGTGCTGACTTTTTGACAC CAGAC) into the provided vector. The new plasmid, pBAD202KF, thus contains a version of the fusion gene regulated by a different system. (Figure 7)



Bacterial Strains and Media

The cell strains used included DH5 α for cloning work with pBR322 variants, NovaBlue for the fluorescence studies of genes under P_{tac}/O_{lac} control (NovaBlue has *lacI^q* for increased repression of *lac* repressor), and TOP10 for pBAD cloning. LB plates and liquid medium with selective antibiotics were routinely used. Antibiotics were used in the following concentrations: 25 μ g/ml carbenicillin (for retention of pBR322 plasmid derivatives), 5 μ g/ml tetracycline (for retention of the F' plasmid with *lacI^q* gene in NovaBlue), and 34 μ g/ml kanamycin (for retention of pBAD202 plasmid derivatives). M9 medium was employed for growing cells used in fluorescence measurements.

Induction of the studied genes under *lac* operator control (P_{tac}/O_{lac}) was accomplished with 1mM IPTG. The fusion gene regulated by the P_{BAD} system, with repression from AraC binding, will be induced with \sim .2% L-arabinose in future investigations. These levels should saturate induction.

β -gal Assay

β -gal hydrolyzes ortho-nitrophenyl- β -galactoside (ONPG) to yellow ortho-nitrophenyl (ONP). The absorbance of this product over time was used to relate the amount of β -gal present in different cell samples since the rate of ONP production increases linearly with the amount of β -gal present to hydrolyze ONPG. The reactions were carried out in cuvettes with 1 ml Z Buffer and 0.2 ml ONPG. To this solution was added cell culture from induced and uninduced samples grown in identical conditions and made permeable to ONPG by toluenization. The amounts of

cell culture added were as follows: +20 μl induced culture for the induced low concentration test, +100 μl induced culture for the induced high concentration test, and the same amounts for the two uninduced culture tests. Absorbance readings were recorded by a UV/Vis spectrophotometer at 420 nm every 15 seconds for up to 70 minutes.

Spectrofluorometer and Microscope

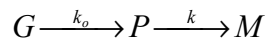
Fluorometer experiments were performed with a SPEX Fluorolog 3 spectrofluorometer. The cuvette containing the cell samples was placed in a thermal jacket cuvette holder with circulating waterbath set at 37°C. A microstirring bar was placed inside the cuvette to keep the sample thoroughly mixed during the measurement. The slit width for the excitation and emission was set at 1 nm and 5 nm respectively. Time trajectories of Venus were taken every 1 second with excitation at 480 nm (blue shifted from Venus's excitation maxima to avoid significant overlap with its emission) and emission measured at 530 nm. The integration time was 1 second.

The Zeiss Axiovert 135 microscope used for Venus visualization was set for bright field illumination with fluorescence excitation from an argon laser at 514 nm. Images were detected with a liquid nitrogen-cooled Princeton Instruments Oma V camera.

Results

With these plasmid constructs we have begun fluorescence tests to detect Venus and β -gal production. So far we have primarily focused on monitoring Venus fluorescence. At this point we want to observe the fluorescence difference between uninduced and induced cells. These efforts could lead to an ensemble measurement of Venus maturation time and will allow us to fine tune the protein's expression (discussed later) in order to prepare for the complete single cell microscope experiments.

Using a spectrofluorometer, which can monitor fluorescence increase of a cell culture, we would like to obtain a distinct curve upon protein induction. As mentioned this should allow us to fit Venus maturation to a simple model, providing an ensemble average measurement of maturation time. Such a model is as follows:



The gene, G , is transcribed and translated at rate k_o to give the protein, P , which becomes the mature protein, M , with rate constant k .

$$\frac{d[P]}{dt} = k_o - k[P]$$

$$\frac{d[M]}{dt} = k[P]$$

$$[P] = \frac{k_o}{k} - \frac{C}{k} e^{-kt}$$

$$[M] = k_o t + \frac{C}{k} e^{-kt} + a$$

where C and a are constants set for initial conditions. Fitting an increasing fluorescence curve (and specifically its beginning shape) with an equation describing mature protein, M , should yield rate constant k that defines an average maturation time.

Venus Fluorescence in Induced NovaBlue Cells with pBR322KF2

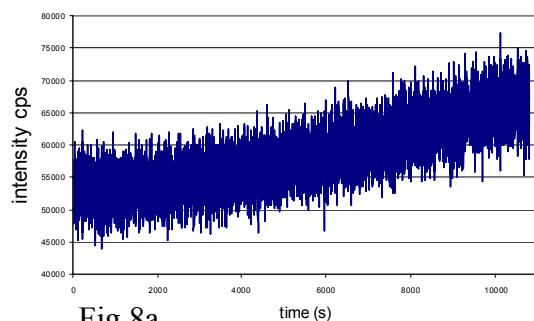


Fig 8a

Fluorescence Time Scan with OD Correction

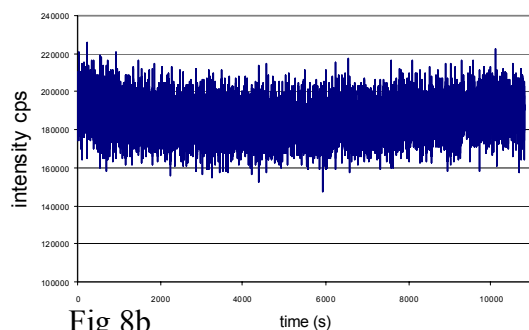


Fig 8b

Figure 8. Time course of Venus fluorescence over the three hours after induction (8a is the fluorometer reading; 8b is the attempted correction for cell growth based on linear OD increase.)

We have induced the *lacI^q* NovaBlue culture containing pBR322KF2 in order to observe fluorescence signal increase (Figure 8). Unfortunately, the signal increase was not dramatic. Figure 8a shows a minor rise in the fluorescence reading during the three hour induction course. This rise of a little less than 30% is presumably due in part to cell growth resulting in increasing cell numbers and thus an automatic potential for a larger population of fluorescent Venus proteins. A preliminary correction based on the optical density (OD_{600}) change of the culture attempts to account for cell growth. This OD correction with results shown in Figure 8b must overcorrect the reading, though, because it seems to indicate no real signal increase upon induction.

More convincingly, microscope visualization of Venus fluorescence within individual cells confirms that there is a noticeable intensity increase hours after induction (Figure 9). This rise of up to about 10-fold in the most intense areas of the cell, however, is not as large as expected. Proteins under P_{tac} promotion have yielded a 50-fold induced/uninduced ratio⁹ and generally provided a significant signal difference in previous published investigations. Further, very good induction in some systems yields a 1000-fold ratio between induced and repressed cell states.

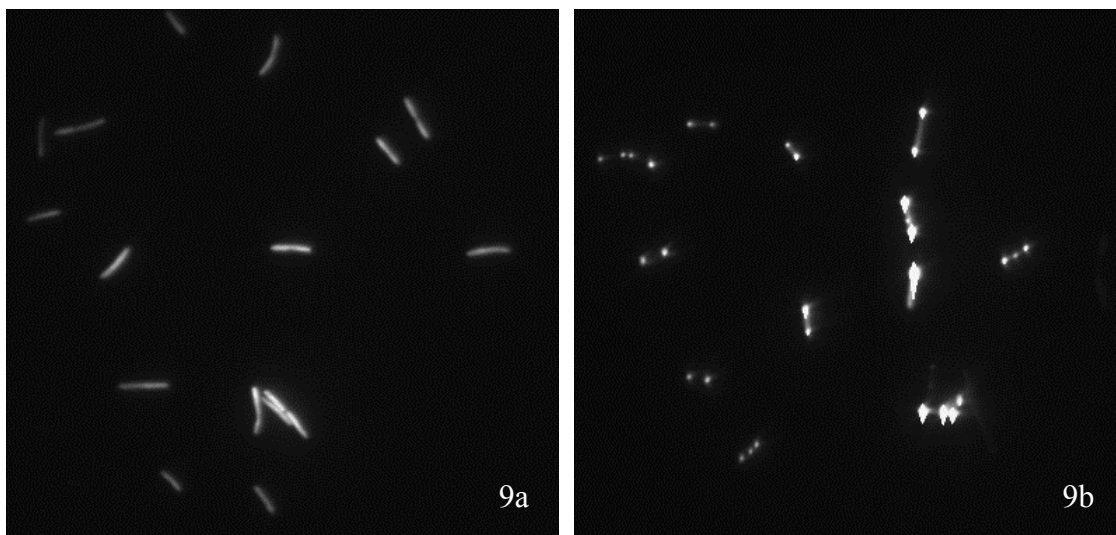


Figure 9. Venus fluorescence in 9a) uninduced and 9b) induced NovaBlue cells containing pBR322KF2 (These preliminary microscope experiment images are simply of the unaltered cells themselves and don't yet apply the protease for protein degradation.)

As an interesting note, we have observed localization of the fusion protein (Venus protein alone may also behave similarly) within the *E. coli* cells (seen in 9b). Polar location is particularly prominent in the induced cells, though more intense fluorescence spots that aren't always polar are often seen in induced and uninduced cells. This localization at the poles of the cell has been previously noted and attributed to environmental stress¹⁰. Further investigation of this occurrence would be intriguing.

The reasons for low induction are not clear at this point. Similar results have been attained for other gene constructs studied in our lab that involve *venus* and are under the control of the same P_{tac}/O_{lac} system. These constructs include the gene for simply *venus* without *lacZ* fused at its end as well as the *venus* gene placed as a single copy on the bacteria's chromosome instead of existing on multiple plasmids. Thus, it seems that the issue is not simply inherent to the fusion gene nor is it primarily an artifact of having too many plasmids carrying the gene within each cell.

Additional possibilities are that Venus is not folding efficiently and that P_{tac}/O_{lac} aren't controlling expression tightly enough. If Venus is not folding well, then much more Venus may be produced in the induced cells but may not mature to generate a detectable fluorophore. This possibility was investigated using the fusion protein of pBR322KF2 in order to detect β -gal's activity, which does not have the potential for inefficient folding and consequent inactivity that Venus does. A β -gal assay was performed to obtain a time curve of absorbance of the β -gal cleaved product of ONPG, ONP. Readings at short times were linear as expected, as the slope is proportional to the rate of ONP production and thus the amount of β -gal present. The slopes at longer times were not proportional to the amount of β -gal as they curved, reflecting ONPG depletion and absorbance signal saturation. Comparison of the samples' initial, linear slopes yielded ratios of induced/uninduced β -gal amounts of ~4.5-5 (Figure 10). Thus it seems Venus is being observed in quantities approximately equivalent to its translation since β -gal of the fusion protein has similar detectable levels of expression.

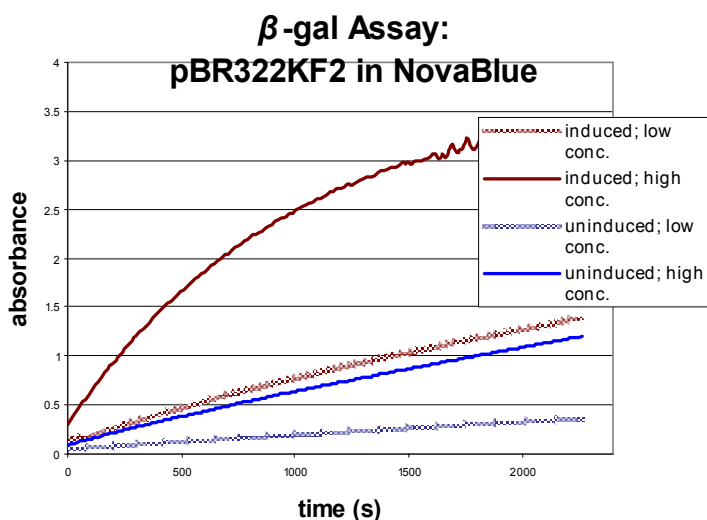


Figure 10. β -gal activity is measured by the rate of ONP production. It is proportional to the graph's slope, or absorbance over time. Specifically, ONP's rate of appearance, and thus a relative amount of β -gal, is calculated as $[Abs/time] * (vol/(\epsilon * l))$, which uses Beer's law to determine ONP concentration. A ratio of the slopes or, equivalently, of the units of β -gal, compares the quantity of β -gal present in the induced and uninduced cultures.

	slope (1/min)	units β -gal= $\mu\text{mol ONP}/\text{min}$	induced/uninduced ratio
induced; low conc.	0.000681	0.0111	4.54
induced; high conc.	0.003084	0.0535	
uninduced; low conc.	0.000150	0.0024	4.94
uninduced; high conc.	0.000625	0.0108	

If, as a second possibility, P_{tac}/O_{lac} aren't controlling expression tightly enough, then the P_{tac} promoter is probably too strong, resulting in an already high level of protein even at the uninduced, or repressed, level. This seems strange since P_{tac} has been successfully used in the past and given a low basal level in the absence of inducer¹¹. However, this may be consistent

with the high Venus fluorescence levels seen on the microscope; both induced and uninduced cells are easily seen fluorescing through the microscope eye piece alone (i.e., without the camera), though the induced cells are noticeably brighter and don't photobleach as visibly as the uninduced ones.

To test the fusion gene, as well as just Venus alone, under the control of a different regulation system, the arabinose-inducible P_{BAD} promoter was chosen. P_{BAD} is reportedly weaker than P_{tac} and thus better for limiting expression in the repressed state. Previously this system has shown controlled, variable gene regulation with expression that can be modulated by inducer (arabinose) concentration and lowered more with glucose. And, it exhibits the potential for more than a 1000-fold difference between repressed and induced protein production in minimal media⁹. The newly constructed pBAD202KF, which utilizes the P_{BAD} promoter, is now ready for fluorescence tests.

Discussion

As an initial step for single cell fluorescence imaging experiments, I have constructed variants of a fusion gene linking *venus* and *lacZ*. Generally this fusion will allow visualization of both expressed proteins via fluorescence detection and, specifically, will act as a timer for the measurement of *in vivo* Venus maturation. In order to prepare for the complete microscope imaging of both proteins, we are performing preliminary fluorescence tests in which we have monitored induction of the gene from its repressed state. Such work on the fluorescence increase resulting from induction could be used for calculating an average of Venus maturation time based on a rate model as described. This model will evolve as we determine the importance of other, more complex parameters, such as the relevance of a large basal Venus signal and natural Venus degradation.

Currently, however, induction doesn't produce the expected large fluorescence difference. This could be due to Venus's inefficient folding in small part, though not primarily as the β -gal assay results with ~5-fold difference agreed with the microscope Venus images showing up to ~10-fold induced/uninduced difference. Likely the regulation of the gene expression under P_{tac}/O_{lac} control is not tight enough. We are trying to decrease transcription of the fusion gene by using a weaker promoter P_{BAD} . As part of a tunable expression system this should also give us increased control of protein levels⁹. In addition, we will try to replace P_{tac} with the weaker P_{lac} while maintaining the *lac* operator. This should yield induction improvements similar to those expected for P_{BAD} .

These modified constructs may help in increasing the induction effect due to a better-controlled basal level expression. Addressing this issue is also important for the single cell microscope experiment because a high frequency of basal level expression that exceeds the fusion protein's degradation rate will obscure the resolvability of the individual expression events.

In proceeding to the microscope fluorescence imaging once the fusion gene is under a regulation system of desirable strength, some additional steps are necessary. The gene must be put on the chromosome or on a plasmid compatible with the ubiquitin-specific protease plasmid. *E. coli* cells then need to be transformed to contain both the fusion gene and protease plasmid. (This fusion gene can also be expressed in yeast, which, as a eukaryote, already has the protease gene naturally.)

We must also investigate Venus stability under high power excitation, as the protein photobleaches easily. Ideally this involves determining settings to maximize emission sensitivity in order to detect low numbers of Venus at a time while minimizing photobleaching and cell disruption. For equipment, we need to set up dual wavelength excitation on a microscope to image Venus and β -gal simultaneously.

These efforts represent the exciting opportunity to look at gene expression in real time. The constructed fusion gene leads toward a beautiful and deceptively simple means for timing Venus *in vivo* maturation and exploring the expression events within cells.

Acknowledgements

I thank Professor X. Sunney Xie for welcoming me into his lab this summer. I am also very grateful for the friendly aid and guidance of my wonderful mentor, Dr. Jie Xiao. Thank you as well to Dr. Ji Yu, Long Cai, and the rest of the Xie group for all of their kind help. Much appreciation to the Harvard REU program for funding and support.

References

1. Nagai, T., *et al.*, *A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications*. Nat Biotechnol, 2002. **20**(1): 87-90.
2. Tsien, R.Y., *The green fluorescent protein*. Annu Rev Biochem, 1998. **67**: 509-44.
3. Bachmair, A., D. Finley, and A. Varshavsky, *In vivo half-life of a protein is a function of its amino-terminal residue*. Science, 1986. **234**(4773): 179-186.
4. Varshavsky, A., *The N-end rule: functions, mysteries, uses*. Proc Natl Acad Sci USA, 1996. **93**(22): 12142-9.
5. Sambrook, J., D.W. Russell, *Molecular Cloning: A Laboratory Manual*. 3rd ed. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press, 2001.
6. Bachmair, A., A. Varshavsky, *The Degradation Signal in a Short-Lived Protein*. Cell, 1989. **56**:1019-1032.
7. Matouschek, A., *Protein unfolding—an important process in vivo?* Curr Opin Struct Biol, 2003. **13**: 98-109.
8. Hoskins, J.R., *et al.*, *Protein binding and unfolding by the chaperone ClpA and degradation by the protease ClpAP*. PNAS, 2000. **97**(16):8892-7.
9. Guzman, L.M., *et al.*, *Tight Regulation, Modulation, and High-Level Expression by Vectors Containing the Arabinose P_{BAD} Promoter*. J Bacteriol, 1995. **177**(14): 4121-30.
10. Santini, C.L., *et al.*, *Translocation of Jellyfish Green Fluorescent Protein via the Tat System of Escherichia coli and Change of Its Periplasmic Localization in Response to Osmotic Up-Shock*. J Biol Chem, 2001. **276**(11): 8159-64.
11. Cormack, B.P., R.H. Valdivia, S. Falkow, *FACS-optimized mutants of the green fluorescent protein (GFP)*. Gene, 1996. **173**:33-38.