The Dissertation Committee for Kathryn Lauren Moncivais Certifies that this is the approved version of the following dissertation:

Novel Tools for the Study of Protein-Protein Interactions in Pluripotent Cells

Committee:

Zhiwen Jonathan Zhang, Supervisor

George Georgiou

Pengyu Ren

Krishnendu Roy

Xiaojing John Zhang

Novel Tools for the Study of Protein-Protein Interactions in Pluripotent Cells

by

Kathryn Lauren Moncivais, B.S.Bio.E., M.S.E.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August 2011

Dedication

For Miciah, who might have been a doctor.

Acknowledgements

Without the support and/or inspiration of these people I would not have this degree: Dr. Jonathan Zhang, Dr. Peppas, Dr. Georgiou, Dr. Roy, Dr. Ren, Dr. John Zhang, mom, Chris, James, Jessica, Laura, Yinan, Eileen, my wonderful undergrads: Jay, Wonjae, Young, and Tiffany, Dr. Saterbak, Dr. Oden, Tracy Volz, Jie, Gabby, Aiko, Liang, Hassan, Jessica, Laura, Coach Eike, Mrs. Ziegler, Mrs. Mueller, Coach Fletcher, and Mrs. Brown.

Novel Tools for the Study of Protein-Protein Interactions in Pluripotent Cells

Kathryn Lauren Moncivais, Ph.D.

The University of Texas at Austin, 2011

Supervisor: Zhiwen Jonathan Zhang

Unnatural amino acids (UAAs) have been used in bacteria and yeast to pinpoint protein binding sites, identify binding partners, PEGylate proteins site-specifically (vs. randomly), and attach small molecule fluorophores to proteins. The process of UAA incorporation involves the manipulation of the genetic code, which is established by the proper function of aminoacyl tRNA synthetases (RSs) and their cognate transfer RNAs (tRNAs). It has been discovered that certain regions of RS proteins can either block or enable cross-species reactivity of RSs. In essence, a bacterial RS can function with a human tRNA by transferring the human CP1 region to the bacterial RS, and vice versa. This knowledge has been used to engineer a tRNA capable of recognizing a stop codon (tRNA*), rather than an amino acid codon, and a cognate RS capable of recognizing only tRNA* and no endogenous tRNAs. We have previously described the use of this methodology to engineer a UAA incorporation system capable of amber stop codon suppression in HEK293T cells. Since UAAs are so useful, and their use has now been enabled in mammalian systems, we applied UAA incorporation to pluripotent cells. Stem and pluripotent cells have been the focus of cutting edge research for years, but much of the work done on these cell lines is done in the ignorance of basic biological processes underlying differentiation, dedifferentiation, and tumorigenesis. In order to facilitate the study of these basic biological processes and enable more adept manipulation of differentiation, dedifferentiation, and tumorigenesis, the development and use of two separate UAA incorporation systems is described herein. The overarching goal of this project is to facilitate the study of protein-protein interactions in stem and pluripotent cells. Since we have also previously described the development of a mammalian two-hybrid system, the use of that system in pluripotent cells is also described.

Table of Contents

List of Tables

List of Figures

CHAPTER ONE

Introduction: Specific Aims and Overview

INTRODUCTION

In recent years "stem cell" has become somewhat of a buzz word in the scientific community, probably owing to the leaps and bounds by which the field has grown in that same time period of time. The New York Times ran a story in 1998 about the infrequent but invaluable use of cryopreserved cord blood for treatment of leukemia, and since that time the incidence of stem cell based therapies has only increased (1). Stem cells have been successfully used as a treatment for anemia and multiple types of leukemia (2), and laboratory research has demonstrated their use in the treatment of *in vivo* models of diabetes (3), cardiac muscle disorders, Parkinson's disease, and spinal cord injury (4, 5). Despite these encouraging results, stem cell transplantation techniques often result in the formation of benign tumors, a fact which is unsurprising in view of the proliferative nature of stem cells. Therefore, as with any new treatment modality, scientists must first characterize the behavior of these cells and why they develop benign tumors. Once that information is known, they can begin to capitalize on the advantages of stem cell based therapies.

To characterize the causes of stem cell behaviors, research must be devoted to studying the protein-protein interactions (PPIs) and activation pathways involved in stem cell differentiation, dedifferentiation, and tumor formation. Without this knowledge, stem cell based therapies are simply cancer time bombs waiting to go off, but with it, they stand to change the face of medicine forever. Scientists may choose to use siRNA, co-precipitation, two-hybrid systems, genetic mutants, cross-linking, RT-PCR, ELISA, or any of a number of commonly used techniques to study PPIs in stem cells.

Unfortunately, all of the most common methods for studying PPIs, like those just mentioned, suffer from a few obvious disadvantages. Many of these techniques rely upon *in vitro* reaction conditions which poorly mimic the native environment of the PPIs under study, and they frequently measure PPIs indirectly and can produce positive results due simply to non-specific interactions of other proteins with those of interest. For a PPI investigational technique to successfully identify PPIs involved in the complex processes of differentiation, dedifferentiation, and cancer development, it would need to overcome some of these disadvantages and be able to capture weak, transient, and spatiotemporally specific interactions.

In an attempt to fulfill this need for an advanced PPI investigational method, our lab has recently developed a novel technique for detection of protein-protein interactions *in vivo* in *Escherichia coli (E. coli)*. An unnatural amino acid (UAA), L-3,4-dihydroxyphenylalanine (L-DOPA), bearing an orthogonal functional group not found in any natural amino acid was genetically encoded in the gene for the *Staphylococcus aureus (S. aureus)* virulence factor Sortase A (SrtA). *E. coli* cells expressing SrtA with L-DOPA were lysed and subjected to cross-linking conditions which revealed that SrtA exists as a dimer *in vivo*(6). While other cross-linking methods do exist, they utilize functional groups which are found in more than one amino acid and therefore frequently occur more than once within any given protein. These types of cross-linkers work so well that they often cross-link random, non-interacting proteins, yielding a substantial amount of false positives. While scientists are certainly accustomed to verifying and re-verifying their results, in a field with as much forward momentum as that of stem cell research, it would be far more favorable to use more stringent screening tools in order to speed the rate at which dependable information is generated and thereby decrease the time differential between initial discovery of stem cell

cascades and the clinical application of the same. Our lab has previously demonstrated that L-DOPA does not randomly cross-link, and it will only cross-link when incorporated into the interacting surface of the protein(s) (6). Therefore L-DOPA generates information about both the existence and location of interaction. Since L-DOPA has these clear advantages over conventional cross-linking methodologies, it is a good candidate tool for elucidating stem cell PPIs. In the following chapters of this dissertation, the groundwork for using L-DOPA as a site-specific cross-linker in stem-like cells will be presented.

Various UAAs have previously been incorporated into proteins in vivo in yeast and bacterial systems, and a few have been used in mammalian cells(7).The incorporation efficiency in mammalian systems is dramatically lower than in yeast or bacterial systems, and at the same time the background is higher. Due to the extreme difficulty associated with stem cell transfection, the system for site-specific incorporation of L-DOPA must be improved to make it a viable tool in stem cells. Our group previously used a *Methanococcus jannaschii (M. jannaschii)* $tRNA^{Tyr}$/tryosyl-tRNA synthetase pair which suppresses an amber stop codon to incorporate UAAs. This *M. jannaschii* $tRNA^{Tyr}$ is recognized to some degree by unknown endogenous aminoacyl tRNA synthetases (RSs) in mammalian cells, and therefore produced a significant amount of background incorporation of random amino acids. While the system has been used in easily transfected cell lines like HEK293Ts, it was too crude to be used in stem cells. Substantial progress has been made in recent years in understanding how tRNAs are recognized by their cognate RSs, and many new tRNA sequences have been catalogued for a myriad of different species and amino acid specificities (8). Additionally, scientists have made great strides in understanding how RSs discriminate between their intended amino acid and all the rest and have successfully switched RS specificities for

both amino acids and tRNAs. Armed with that information and bioinformatics tools, this dissertation presents the optimization of a UAA incorporation system and its application in stem cells.

In addition to improving the system for incorporation of UAAs in mammalian cells, it was also necessary to improve upon existing transfection methods for stem cells. In the past, calcium phosphate precipitation and electroporation were the most effective techniques for stem cell transfection. However both methods result in high levels of cell death and mediocre transfection efficiencies. Newly available commercial products like Fugene 6 and Fugene HD have demonstrated exceptionally high transfection efficiencies in some hard-to-transfect cell lines, and newly developed electroporation techniques have achieved high transfection efficiencies as well. Additionally, polyethylenimine has been recently demonstrated as a cost-effective alternative to expensive commercial reagents like Fugene 6 and Fugene HD. Presented in this dissertation is the investigation and optimization of the best of these techniques in order to develop a viable transfection protocol for hard to transfect cells.

While UAAs can serve in a great many capacities beyond that of site-specific cross-linker, there are other PPI study methods which may also help in the delineation of differentiation/dedifferentiation/tumorigenesis pathways in stem cells. While they are plagued by a high incidence of false-positives, two-hybrid methods have worked well for some applications in the past. Since the our lab recently developed a novel two-hybrid system capable of detecting weak interactions in HEK293T cells, its application in stem-like cells has been investigated in this work (9). A two-hybrid system for use in stem or stem-like cells, though imperfect, could be an asset in the toolbox for PPI investigation and thus is worth examination.

SPECIFIC AIMS

The overall goal of this project is to develop and demonstrate the use of novel tools to study PPIs in proteins in stem cells which can be used in the future to elucidate signaling cascades involved in the processes of differentiation, dedifferentiation, and tumorigenesis. To accomplish that, the project has been broken down into three specific aims.

Aim 1: Use rational design to create a new tRNA, aminoacyl synthetase, and TAG-reporter protein system which exhibits lower background, increased orthogonality, and increased reporter gene expression in mammalian cells for subsequent use in stem cells.

We hypothesized that bioinformatics tools could be used to develop an acceptor stem which was unrecognized by endogenous RSs. Human tRNA sequences for all amino acids were compared with all known $tRNA^{Tyr}$ sequences from any other organism, and a new $tRNA^{Tyr}$ was developed. In parallel with that, an *E. coli* $tRNA^{Tyr}$ was tested for orthongality. A new reporter gene was created by inserting an amber stop codon into a GFP gene optimized for mammalian expression. In this way we were able to create a more orthogonal tRNA, use its cognate synthetase, and switch to a reporter protein more easily expressed by mammalian cells, which afforded an efficient UAA incorporation system appropriate for use in stem cells.

Aim 2: Overcome the stem cell transfection limitations of this system by (a) developing a transfection protocol capable of higher than previously published efficiencies, and (b) condensing the current triple plasmid system to a double or single plasmid system.

We hypothesized that current transfection methods were not effectively optimized for use in stem cells and that individual optimization of each technique or a combination of two or more techniques could dramatically increase transfection efficiencies in hard to transfect cell lines. (a) Fugene 6, Fugene HD, polyethylenimine, and electroporation were explored to determine the most efficient method for transfection of both single and double plasmids into P19 embryonal carcinoma cells. By optimizing these reagent based transfection methods outside the recommended limits of use it we were able to achieve higher transfection efficiencies than previously reported using the canonical methods. Further, by exploring extreme cell density, DNA concentration, and buffer conditions for electroporation we successfully exceeded the efficiency obtained using commercial products. We further hypothesized that decreasing the number of plasmids necessary for the system would allow for more efficient expression of the UAA incorporation system components. (b) After finding the maximum transfection efficiency for this cell line, the DNA constructs were altered such that the TAG-reporter protein and synthetase were expressed by a single bidirectional plasmid, and the tRNA was expressed by another plasmid.

Aim 3: Demonstrate application of the optimized UAA incorporation system and the two-hybrid system in P19s, and demonstrate that the UAA incorporation system does not cause differentiation.

We hypothesized that the tetracycline repressor based two-hybrid system could be used to detect PPIs in stem cells by introducing it to the system with the optimized transfection protocol developed in Aim 1. We used the two-hybrid system to detect a number of PPIs and found that strong interactions could be identified in this manner. We also hypothesized that the optimized UAA incorporation system would successfully

incorporate UAAs into proteins in P19s and would not cause differentiation. To test this theory, P19s were transfected with the components of a UAA incorporation system and immunostained for differentiation markers. It was found that the UAA incorporation system did not cause differentiation of this cell line and therefore could be useful in the study if PPIs in stem or stem-like cells. Cells from the same experiment were lysed and probed for UAA incorporation.

OVERVIEW

Chapter Two provides a broad overview of the background and significance of this work. *Chapter Three* discusses the lengthy process of developing a superior transfection method by optimizing a wide variety of methods and in one instance combining two. It reveals that a superior, consistent transfection method for P19s was developed and dramatically improved the outcome of all experiments in this work which involve transfection. *Chapter Four* details the development of a UAA incorporation system suitable for use in mammalian cells, while *Chapter Five* discusses the application of that system in P19s. *Chapter Six* gives a brief overview of the progress made on incorporating UAAs into proteins in P19s. *Chapter Seven* discusses development and use of a mammalian two-hybrid system, which was able to detect strong binding events. Finally, *Chapter Seven* discusses the broad conclusions which can be drawn from this work and recommendations for how it should proceed in the future.

REFERENCES

1. GRADY D. The Hope, and Hype, of Cord Blood. The New York Times. 1998 December 1, 1998.
2. Pfendler K, Kawase E. The potential of stem cells. OBSTET GYNECOL SURV. 2003;58(3):197-208.
3. Voltarelli JC, Couri CE, Stracieri AB, Oliveira MC, Moraes DA, Pieroni F, et al. Autologous nonmyeloablative hematopoietic stem cell transplantation in newly diagnosed type 1 diabetes mellitus. Jama. 2007;297(14):1568-76.
4. Johnson PJ, Tatara A, Shiu A, Sakiyama-Elbert SE. Controlled release of neurotrophin-3 and platelet-derived growth factor from fibrin scaffolds containing neural progenitor cells enhances survival and differentiation into neurons in a subacute model of SCI. Cell Transplant.19(1):89-101.
5. Liu S, Qu Y, Stewart TJ, Howard MJ, Chakrabortty S, Holekamp TF, et al. Embryonic stem cells differentiate into oligodendrocytes and myelinate in culture and after spinal cord transplantation. Proc Natl Acad Sci U S A. 2000;97(11):6126-31.
6. Umeda A, Thibodeaux G, Zhu J, Lee Y, Zhang Z. Site-specific Protein Cross-Linking with Genetically Incorporated 3,4-Dihydroxy-L-Phenylalanine. CHEMBIOCHEM. 2009;10(8):1302-4.
7. Monahan SL, Lester HA, Dougherty DA. Site-specific incorporation of unnatural amino acids into receptors expressed in Mammalian cells. Chem Biol. 2003;10(6):573-80.
8. Sprinzl M, Vassilenko KS. Compilation of tRNA sequences and sequences of tRNA genes. Nucleic Acids Res. 2005;33(Database issue):D139-40.
9. Thibodeaux GN, Cowmeadow R, Umeda A, Zhang Z. A tetracycline repressor-based mammalian two-hybrid system to detect protein-protein interactions in vivo. Anal Biochem. 2009;386(1):129-31.

CHAPTER TWO

Background and Significance

CURRENT TRENDS IN STEM CELL BASED THERAPIES

A brief survey of scientific literature reveals more stem cell based therapies and lines of research than could be described in this short work. In fact a simple search of clinicaltrials.gov for "stem cell" reveals 1,619 active clinical trials involving stem cell based therapies. Some of the most exciting stem cell based therapies currently in use or development are those focused on spinal and nerve damage treatment. Fetal stem cells have been used in the treatment of cerebellar ataxia, a disorder of caused by dysfunction or absence of Purkinje cells in the cerebellar cortex. Several studies have found that transplanted embryonic stem cells (ESCs) engraft and migrate to the adult cerebellar cortex but do resolve the functional problems of cerebellar ataxia (1, 2). In contrast, Triarhou and colleagues have demonstrated that grafting the cells directly into the desired location allows the ESCs to engraft and form synaptic contacts which result in reversal of some symptoms of cerebellar ataxia (3). While these results in a mouse model are very exciting, their application in a human setting is sobering. A 2009 study published in PLoS Medicine describes a stem cell transplantation procedure which was used to treat a child with ataxia telangiectasia. Four years after the procedure recurrent headaches were the first sign of the spinal and cerebral tumors which resulted from the stem cell transplantation (4). Unfortunately the authors did not comment on the possible benefits of the procedure, instead focusing on the nature of the tumors and their implications for the field of stem cell research. However, the authors of this study echo the sentiments in this dissertation that the field of basic stem cell biology needs further advancement before treatments of this kind can be used safely and effectively.

Despite the clear risks of ESC transplantation based therapies, there are still a number of ongoing studies investigating this strategy for the treatment of a number of disorders, and some have gone so far as to genetically modify these cells prior to transplantation. Xi and colleagues used purified embryonic stem cell derived cardiomyocytes (ESC-CMs) in conjunction with mouse embryonic fibroblasts (MEFs) to treat an *in vitro* model of myocardial infarction (5). They found that combining the ESC-CMs and MEFs allowed mechanical integration of the ESC-CMs and force transduction from the damaged tissue to the transplanted cells. Rizvanov et al transfected human umbilical cord blood cells (hUBCs) with a plasmid which simultaneously expressed vascular endothelial growth factor 165 (VEGF165) and human fibroblast growth factor 2 (HFGF2) (6). The transfected hUBCs were then injected into presymptomatic amyotrophic lateral sclerosis (ALS) transgenic mice, which lead to development of astrocyte like cells. Since astrocyte dysfunction is a known component of ALS, these results are considered to be promising.

No matter how many successful experiments are carried out *in vitro* or even in animal models of certain diseases, the problem of cancer development will still loom large in the field of stem cell therapies. Further compounding that problem are the ethical issues surrounding the use and obtainment of ESCs. In light of these issues, much work has been devoted to finding alternate sources of pluripotent cells. Some researchers have focused on more readily available sources of multipotent cells like mesenchymal stem cells (MSCs) while others have gone so far as to attempt to dedifferentiate adult cells into stem or progenitor cells and use them in place of ESCs. Researchers first seized upon the idea of using MSCs as differentiable and ultimately implantable tissue replacements due to their ready availability, ease of expansion, and lack of immune rejection risk. In 2004 Alhadlaq and colleagues reported the

development of a human-shaped articular condyle using a polyethylene glycol (PEG) based hydrogel and rat bone marrow derived mesenchymal stem cells (7). MSCs were expanded in the presence of chondrogenic or osteogenic factors, suspended in the unpolymerized hydrogel solution, and then the cell + hydrogel solution was photopolymerized into the shape of a human mandibular condyle. The cell-seeded hydrogels were implanted in severe combined immunodeficiency (SCID) mice and harvested four weeks later at which point they exhibited the intended osteogenic and chondrogenic regions. While this work certainly serves as a proof of concept, the implants would need to be monitored for a much longer period of time in order to draw any conclusions about longevity of the state of differentiation of these cells as well as their mechanical and biological properties. Macchiarini et al took a similar but more organic approach to engineering an implantable replacement airway for a young woman with end-stage bronchomalacia (8). A human donor trachea was decellularized, cleared of MHC antigens, and then seeded with the epithelial cells and MSC derived chondrocytes. The seeded trachea matrix was incubated for 96 hours in a novel bioreactor which rotated 90 degrees every 30 minutes to ensure even distribution of growth medium. The engineered trachea replacement was then implanted and demonstrated no issues with host rejection despite the recipient not taking immunosuppressive drugs. Additionally, the recipient of the engineered trachea experienced almost immediate restoration of a functional airway. In addition to these successes, MSCs have been differentiated into insulin-producing cells in the search for a viable beta-cell replacement modality in the treatment of diabetes (9-11). Researchers have also used cord blood derived MSCs to create insulin producing cells, which reinforces the prudence of cryo-preserving cord blood (12). It has also been suggested

that MSCs could be used in cancer therapy since they naturally migrate toward tumors. They could be used as vehicles for anticancer gene delivery (13).

It should be noted that in each of these cases in which MSCs have been used as a source of differentiable and later implantable cells, the differentiation methods employed are not based upon a global understanding of the underlying processes. Instead they are based upon the scientific body of work leading up to this point. In laymen's terms, these differentiation processes are akin to a child who writes a letter to Santa every year and always gets the present at the top of their list. They are ignorant of the underlying process of gift procurement and delivery (i.e. their parents), but they depend upon their experience that writing a letter to Santa always yields the same results. While this is not necessarily the worst position for the state of research, it certainly leaves a lot of room for improvement. However, as can be seen below, exploitation of the knowledge that A leads to B in ignorance of what happens in between, can produce some very exciting results.

On the dedifferentiation front, researchers have developed inducible pluripotent stem cells (iPSCs) using transcription factors known to play a role in pluripotency (14). Takahashi and Yamanaka used viral vectors to express Oct4, Sox2, Klf4, and c-Myc in adult fibroblasts, which successfully restored these cells to multipotency (15). A subsequent study using human somatic cells used Oct4, Sox2, Nanog, and Lin28 to induce pluripotency in a more clinically relevant manner (16).

While the monumental nature of these studies cannot be denied, the amount of time, effort, and money that it took to complete them is almost as monumental as the work itself. Owing to the lack of basic biological knowledge of the processes of differentiation and dedifferentiation, each of these studies began by screening much larger libraries of transcription factors and the effects of numerous combinations of

these transcription factors of interest. In short, researchers started by generating basic biological information and applying it to the process of dedifferentiation. While they did discover cocktails of transcription factors which effectively induce pluripotency, there are still no answers as to why they induce pluripotency and what other effects these transcription factors may be exerting. Regardless of the state of ignorance of the global processes involved in dedifferentiation, these studies served as the jumping off point for a number of very interesting experiments. In response to concerns of the viral methods of pluripotency originally used, scientists began investigating non-viral induction methods. The use of plasmid DNA, protein treatment, and RNA-based methods have each accomplished successful induction of pluripotency in adult cells (17-19). Unfortunately the highest reported efficiency with which all pluripotency induction methods act is at the most 0.005% (20). This dismal induction efficiency can be attributed to incomplete knowledge of the underlying mechanisms of dedifferentiation and differentiation, a fact which scientists seem to comment about profusely without actually acting to change the situation (14).

From the preceding information, one can draw a few conclusions. First of all, the field of stem cell research is flourishing. Second, stem cell research seems to be just on the cusp of developing incredibly useful, revolutionary, regenerative therapies for a number of diseases. Lastly though, and least excitingly, this field is lacking in fundamental, basic science knowledge of the molecular processes underlying dedifferentiation, differentiation, and tumorigenesis of pluripotent cells. While incredible tools and therapies may be developed in the absence of this knowledge, it cannot be argued that all of these therapies would be better, safer, and more accessible to the people they would most benefit if the blanks in our knowledge of the inner workings of these cells were filled.

PROTEIN-PROTEIN INTERACTIONS STUDIES IN STEM CELLS

It is a well established fact that most embryonic stem cells (ECs) require leukemia inhibitory factor (LIF) in order to maintain pluripotency (21), which made the LIF mechanism of action one of the first targets of protein-protein interactions (PPIs) studies in stem cells. This section will focus on the methods employed to study PPIs in stem cells rather than the interactions themselves, because the overall aim of this project is to create and achieve proof of concept of tools which will facilitate the study of protein-protein interactions. Ersnt and colleagues used a combination of stable cell line development, immunoprecipitation, Western Blotting, mobility shift assays, antisense RNA, and nucleotide exchange assays to determine a few useful aspects of the LIF mechanism (21). First, it became clear that phosphorylation of the Src-related kinase Hck was an almost immediate effect of LIF stimulation, which was further corroborated by the investigation of the effect of a constitutively active form of Hck. Cells expressing the constitutively active form of Hck had less dependence on LIF for the maintenance of pluripotency. Further, inhibition of the expression of members of the Janus kinase (Jak) family increased the dependence of ES cells on LIF, implicating Jak family members as part of the signaling cascade involved in pluripotency maintenance. In 1999 (3 years later) the same group used an almost identical tool set to delineate the cytoplasmic domain of the cytokine receptor chains involved in maintenance of pluripotency via the LIF pathway (22). They found that the carboxy-terminus (C-terminus) of gp130 was necessary for transduction of proliferation signals in LIF treated cells, and they further refined this information by determining the location within the C-terminus without which pluripotency could not be maintained. It should be noted that both of these studies were completed by the same group, that there was a three year gap between these publications, and that the information generated deals specifically, as far as we

know at this point, with pluripotency. Differentiation, dedifferentiation, and tumorigenesis likely involve far more complicated and transient protein-protein interaction networks than pluripotency, and thus can be expected to require more intense study if they are to be delineated.

Another popular PPI investigational technique is that of microarrays. Freemantle et al used the commercially available Human Unigene 1 Life microarray from Incyte Genomics, which is comprised of 9,128 independent clones, representative of 8524 unique genes, to study retinoic acid (RA) mediated differentiation in human embryonal carcinoma cells (23). They identified a total of 57 genes which were up-regulated and 37 genes which were down-regulated in response to RA treatment. It should be noted that, based upon the number of up-regulated or down-regulated genes and the total number represented in the microarray, at any given time less than 1% of the cDNAs studied were affected by RA treatment. This indicates that the number of proteins and transcriptional factors involved in RA mediated differentiation, and ostensibly other differentiation pathways, is very small. However, many proteins are known to have promiscuous interactions with multiple binding partners, so this does not mean that the total number of proteins and PPIs involved in these processes is the same size as the number of altered genes. Unsurprisingly, many of the upregulated genes in this study were already known participants in developmental pathways (i.e. transforming growth factor beta, Notch, Hedgehog, and Wnt). In contrast, many of the downregulated genes were involved in protein processing and turnover. While high throughput studies like this one generate a wealth of information, they do not give any reliable information as to which of the upregulated proteins is interacting with which other proteins. So it provides insight into which proteins are likely involved but does not reveal the molecular details of how they interact and achieve signal transduction.

P19 EMBRYONAL CARCINOMA CELLS

While developing a tool kit for the study of PPIs in stem or pluripotent cells, it is crucial that an appropriate cell line be employed. An appropriate cell line for this endeavor would have the following characteristics: 1) Capable of differentiating into multiple cell types, 2) Transfectable, 3) Well established protocols for differentiation assessment, 4) Highly proliferative, 5) Requires minimal effort to maintain pluripotency. Admittedly, that list is almost unreasonable, but there are cell lines that have all of those characteristics. We have chosen P19 mouse embryonal carcinoma cells (P19s) for the work undertaken in this project; their history and qualifications, as per the previous list, are discussed below.

Teratocarcinomas are germ cell tumors composed of embryonal carcinoma and teratoma cells and can form when early embryos are transplanted into ectopic sites (24, 25). McBurney and Rogers transplanted a 7.5 day old embryo into the testis and then cultured the cells which arose from that procedure, eventually developing the P19 cell line (26). This cell line was remarkable at the time as it grew easily and quickly in culture in the absence of feeder layers or differentiation maintenance supplements like LIF. With many stem cell lines, cell density cannot surpass a certain level without initiating differentiation, but conveniently P19s do not exhibit this same behavior (27, 28). This property alone makes maintaining this cell line much simpler and less time intensive than culturing a cell line that requires a feeder layer and/or differentiates in response to confluency. McBurney and colleagues demonstrated that P19s could be differentiated, but they had to be both dense and treated with certain drugs in order to efficiently direct them toward a certain lineage (29). In response to nonlethal concentrations of either dimethylsulfoxide (DMSO) or retinoic acid (RA), P19s can differentiate to neuronal cells or a variety of mesodermal cells including cardiac and skeletal muscle (30-32).

As previously mentioned, cells must be cultured at high density and in non-tissue culture treated vessels in order for DMSO or RA induced differentiation to occur with any efficiency (32). To differentiate P19s using RA, cells must be aggregated and exposed to 3x10-7M RA for as little as four hours to induce irreversible differentiation to neuronal cell types (30). six days after RA treatment up to 85% of cells express neuronal markers (33). While other cell types do occur in RA treated P19 cultures, neuronal cell types are by far the most abundant (30). DMSO induced differentiation begins with aggregation as well and requires only 0.5-1% (v/v) DMSO in the culture medium (32). Within 7 days of DMSO treatment, cardiac muscle cells which often beat rhythmically can be found in culture, accounting for roughly one quarter of the total cells in culture. These cells express a number of proteins which indicate their similarity to embryonic tissue rather than adult tissue (34), which makes them a good candidate for the study of PPIs involved in normal cardiac development.

In light of the many advantages P19 cells possess over their more totipotent peers, it is unsurprising that they have been used continuously since their discovery to study many aspects of development and differentiation (35-37). To investigate the role of bone morphogenic proteins (BMPs) in cardiac differentiation, P19s were transfected to over-express noggin, an inhibitor of BMPs (38). Noggin expression successfully prevented DMSO induced cardiac differentiation, but addition of BMP proteins to the media could restore cardiac differentiation. While DMSO may induce cardiac differentiation, it is not responsible for upregulation of BMP, and cardiac differentiation cannot proceed in the absence of BMP(38). In the same study as well as subsequent studies aimed at delineating the signals involved in DMSO induced differentiation of P19s, researchers were able to develop an incomplete picture of this process, shown in Figure 2.1 (39).

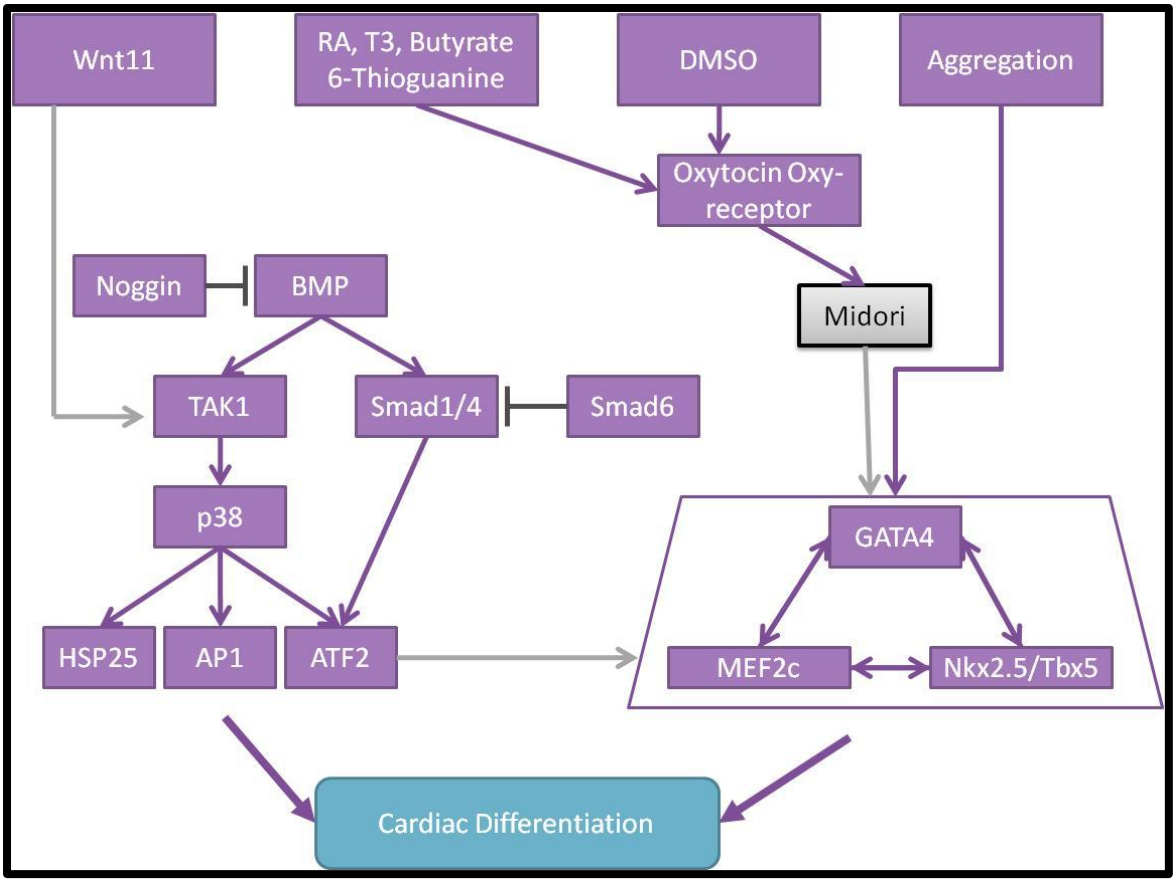


Figure 2.1: DMSO induced differentiation pathways in P19 embryonal carcinoma cells. Gray lines or boxes indicate speculated relationships or proteins that are unconfirmed at this time.

Similar experiments have been carried out with the end goal of defining the neural differentiation pathway activated by RA treatment in P19s. By overexpressing FoxA1, a known component of neural differentiation in P19s, and studying the expression of several other proteins in response to FoxA1 expression, Tan and colleagues were able to identify some of the downstream effects of RA induction (40). FoxA1 expression was followed by a dramatic and almost immediate decrease in expression of the pluripotency marker Nanog followed by a delayed increase in the expression of the neuronal protein Nestin (40). It was also found that siRNA knockdown

of FoxA1 prevented the upregulation of both Nestin and sonic hedgehog (Shh) in response to RA induction. These results indicate that FoxA1 expression is necessary for neuronal differentiation in response to RA treatment, and Nestin and Shh are downstream targets of the FoxA1 differentiation cascade.

In contrast to the action of FoxA1, Foxm1 expression is diminished during neuronal differentiation in response to RA (41). Knockdown of Foxm1 resulted in a corresponding decrease in expression of the well-known pluripotency marker Oct4, and overexpression of Foxm1 in P19s at day 4 of RA induced neuronal differentiation restored the expression of Oct4 as well as Nanog and Sox2, also well known pluripotency markers (41). The effect of Foxm1 knockdown was also assessed in teratoma formation, which yielded an interesting result: Foxm1 negative teratomas were limited in the lineages to which they could differentiate. Foxm1 knockdown caused spontaneous cardiac differentiation *in vivo*. This result begs for further investigation as it is one of few with the implication that the neuronal and cardiac differentiation pathways may involve some sort of molecular switch, most likely somewhere downstream of Foxm1.

The preceding information points to several basic truths about the investigations of PPIs in P19s. First, these cells have proven highly useful in the last 30 years, and therefore represent a solid choice for development of new tools for the study of PPIs in pluripotent cells. Second, the information gleaned from these cells in that time period is woefully incomplete. Third and finally, the usefulness of these cells and their potential impact on regenerative and transplantation medicine demands that scientists find ways to more quickly and efficiently exploit the unique advantages of this cell line and others like it.

STEM CELL TRANSFECTION METHODS

Stem cells have the potential to revolutionize the fields of medicine and biology, so it is unsurprising that scientists would be interested in transfecting them with exogenous DNA. In certain adherent cell lines, commercial reagents have been used to achieve close to 100% transfection efficiency, which is a monumental improvement over all techniques in use 20 years ago (unpublished data). Unfortunately, as is the case with most transfection methods, these reagents do not transfect stem cells as well as they do other cell lines. Because of this, most attempts at genetic manipulation of stem cells have utilized viral methods. Lentiviral transfection methods have achieved both high, stable gene expression and unmodified differentiability of infected cells (42, 43). While lentiviral and most viral methods yield superior results as compared to all other techniques, the difficulty and additional safety measures involved in producing the viruses prevents them from being the first choice for this project. However, they still represent a logical gene delivery vehicle, especially for investigations of the effect of single proteins or mutant proteins. The project presented within this dissertation involves the expression of two proteins and one exogenous tRNA. Because of that, the time it takes to use virus-based method in this endeavor may prove more costly than the results are worth. Thus, it will be investigated for efficacy and feasibility, but it will be quickly replaced with a simpler method if it appears to be more time consuming than its results can justify.

Electroporation, liposomal, and non-viral, non-liposomal methods have been the most successful transfection techniques aside from viral based methods. Fugene 6 and Exgen500, both non-liposomal, non-viral reagents, have reportedly achieved 11% and 16% transfection efficiency in stem cells, respectively. However Exgen500 did adversely affect both cell viability and rate of proliferation in comparison to Fugene 6 (44). The

proprietary method of Nucleofection has been used to achieve between 50 and 60% transfection efficiency in endothelial cells (45) and a whopping 80% transfection efficiency in P19 embryonal carcinoma cells (Amaxa Literature). Since Nucleofetion is a proprietary form of electroporation, studies related to non-proprietary electroporation could shed some light on how to achieve the same level of transfection. It has been reported that high salt, high voltage electroporation conditions can achieve between 6 and 10 fold better transfection efficiency than alternative methods (46).

LABELED PROTEINS AND THEIR USES

Green Fluorescent Protein (GFP) was discovered in 1962 by Shimomura, Johnson, and Saiga (47), and it is safe to say that this discovery revolutionized the *in vivo* and *ex vivo* study of proteins. In the near half century since its discovery, GFP has been transformed from an unknown greenish protein into a household name in biological science. In the following chapters, the use of an optimized form of GFP as a reporter protein will be discussed extensively, but one of the most frequent and successful applications of GFP has been as a fusion tag (48). Since the introduction of GFP as a fusion tag, researchers have attempted to increase the potency and decrease the size of fusion tags, because GFP was able to yield such a wealth of information about protein compartmentalization. In the following section, the uses and creation of protein tags will be discussed, as will their implications for the research presented in this dissertation.

The labeling of proteins in mammalian cells allows researchers to quickly and easily determine a number of things. First, tagging proteins with a fluorescent moiety allows for the microscopic evaluation of localization of a single protein (i.e. nuclear versus endoplasmic reticulum), and adding two different fluorescent moieties to two

proteins can reveal valuable information about colocalization of proteins in unperturbed cells or during specific cellular processes. This type of experiment is fast and can provide information for the refinement of subsequent studies. Second, tagging proteins with a cross-linker can allow determination of PPIs and conformational changes in living cells. A protein of interest can be tagged with a cross-linker, and at the chosen time or perhaps under the conditions of interest to the researcher, cross-linking can be catalyzed and subsequent protein purification, and possibly mass spectrometry as well, can reveal the identity of the cross-linking partner. Protein tagging can also facilitate PPI investigation through fluorescence resonance energy transfer (FRET) (49, 50). Third, tagging proteins can allow for immunological detection of proteins for which dependable antibodies do not exist. One example of this phenomenon is histidine (His) tagging of proteins. His-tagged proteins can be easily detected with commercially available antibodies directed against the His-tag rather than the protein itself. Fourth, proteins can be tagged to allow for easy purification. A His-tagged protein can efficiently bind nickel beads in a protein purification column, a technique which is incredibly frequently used by researchers in all biological fields. The fifth and final use of tagged proteins is to allow for chemical or post-translational modification. Tagged proteins can be selectively modified in a number of ways when certain functional groups are introduced, and site-specifically tagging said proteins allows the targeting of whatever modification is desired. Further, some tags can serve as pseudo post-translational-modifications (PTMs) and facilitate the study of the same (51).

UNNATURAL AMINO ACIDS AND THEIR APPLICATIONS

So-called unnatural amino acids (UAAs) are amino acid analogues which are unrepresented in the genetic code of most organisms. Because they are absent from the

genetic code, they present orthogonal functional groups to most systems, functional groups which can be specifically modified without interfering with other proteins necessary for function of the cell or organism under study (52). These functional groups have been used in a number of novel applications in recent years. Some UAAs are very similar to canonical amino acids but are heavier than their analogs. These UAAs can be extremely helpful in crystallography experiments. For example, methionine was replaced with selenomethionine in a protein, and that replacement facilitated crystallographic determination of the protein's structure(53).

A number of UAAs are small fluorescent molecules which are incredibly useful for protein tagging and visualization applications. Schultz and colleagues have used the small fluorescent UAA shown in Figure 2.2 to visualize nuclear localization of histones in Chinese hamster ovary (CHO) cells (52). This same type of UAA is known to exhibit high sensitivity to environmental changes which can be detected by excitation and emission spectra shifts (52). Mills et al successfully exploited this property, using a small fluorescent UAA to probe antibody-antigen binding (54). Summerer and colleagues employed dansylalanine in much the same strategy to detect folding/unfolding of proteins in yeast (55).

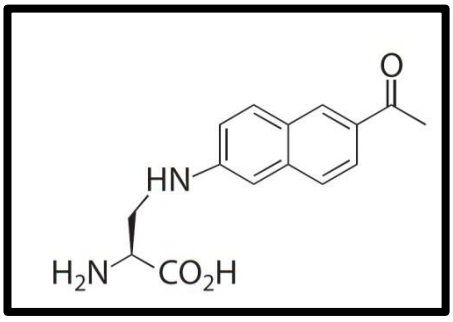


Figure 2.2: Small fluorescent UAA used in (52)

Another popular application of UAAs is in cross-linking of interacting proteins. Since many PPIs are transient and/or weak, covalent cross-linking is necessary to directly assess biologically relevant interactions, especially if those interactions are to be assessed *in vivo.* Two UAAs have been successfully used in this endeavor by multiple groups. L-3,4-dihydroxyphenylalanine (L-DOPA) has been incorporated into proteins in multiple organisms and cross-linked to their interacting partners. L-DOPA was genetically encoded in the gene for the *Staphylococcus aureus* (*S. aureus*) virulence factor Sortase A (SrtA). *E. coli* expressing SrtA with L-DOPA were lysed and subjected to cross-linking conditions which revealed that SrtA exists as a dimer *in vivo* (56). This same UAA has also been used to increase the affinity of a peptide-based antibody by cross-linking the peptide-antibody to its target on a nitrocellulose membrane (57). Benzoylphenylalanine (Bp-Ala) is a photoreactive cross-linker which covalently attaches to binding partners upon irradiation by certain wavelengths of light (52). Bp-Ala has been incorporated into the adapter protein Grb2, which is known to function in Ras signaling (58). When co-expressed in CHO cells with the EGF-receptor and subjected to photocrosslinking conditions, Bp-Ala containing Grb2 irreversibly linked to EGF receptors (58). It should be noted that L-DOPA mediated cross-linking is usually carried out with lysed cells, whereas Bp-Ala cross-linking can be done in living cells. This is important, because *in vivo* cross-linking combined with fluorescent probes can generate a wealth of information very quickly, whereas techniques which assess cell lysate rather

than live cells generate less information which may be less biologically relevant due to lysis conditions and their effect on PPIs.

The last and final application of UAAs to be discussed in this dissertation is as a non-fluorescent probe of protein structure and function in living cells. Since UAAs often differ from their canonical counterparts in the size and/or shape of their functional groups, they can be used to investigate the mechanical role of amino acids in wild type proteins. For example, o-methyl-L-tyrosine was used to investigate a proposed model of voltage dependent K^+ channel inactivation in neurons (59). Tyr19 was replaced with o-methyl-L-tyrosine in the K^+ channel Kv1.2, and its effect on activation/inactivation of Kv1.2 was assessed. It was found that introducing the bulky side chain of o-methyl-L-tyrosine slowed the inactivation by a factor ranging from five to seven in comparison to the wildtype Kv1.2 channel (59). This method could be applied to the mechanism of action of many ion channels as well as a number of other conformational change/protein structure investigations.

MECHANICS OF UNNATURAL AMINO ACID INCORPORATION

The process of UAA incorporation is much like that of the incorporation of natural amino acids. Each canonical amino acid within a protein is represented by a series of codons, each of which consists of a specific three nucleotide sequence. Each codon represents only one amino acid, though each amino acid is represented by more than one codon. In addition to the codons for amino acids, there are codons representing 'START' and 'STOP' locations for protein translation, effectively directing the ribosome where to begin and end translating a protein. Once a gene has been transcribed from DNA to messenger RNA (mRNA), it can be translated by the ribosome

complex. As shown in Figure 2.3, the ribosome is the location at which tRNA with previously covalently attached amino acid(s) base pairs with mRNA, bringing the correct amino acid into the ribosome active site and allowing it to be attached to the growing peptide chain. For this process of DNA→mRNA→protein to work correctly, the fidelity of the original genetic code must be maintained by RSs, which catalyze the attachment (aminoacylation) of amino acids to tRNAs. RSs must therefore distinguish both the correct amino acid and any of a set of isoacceptor tRNAs which bear the anticodon for their cognate amino acid. Therefore UAAs can be incorporated by hijacking the endogenous cellular protein synthesis machinery.

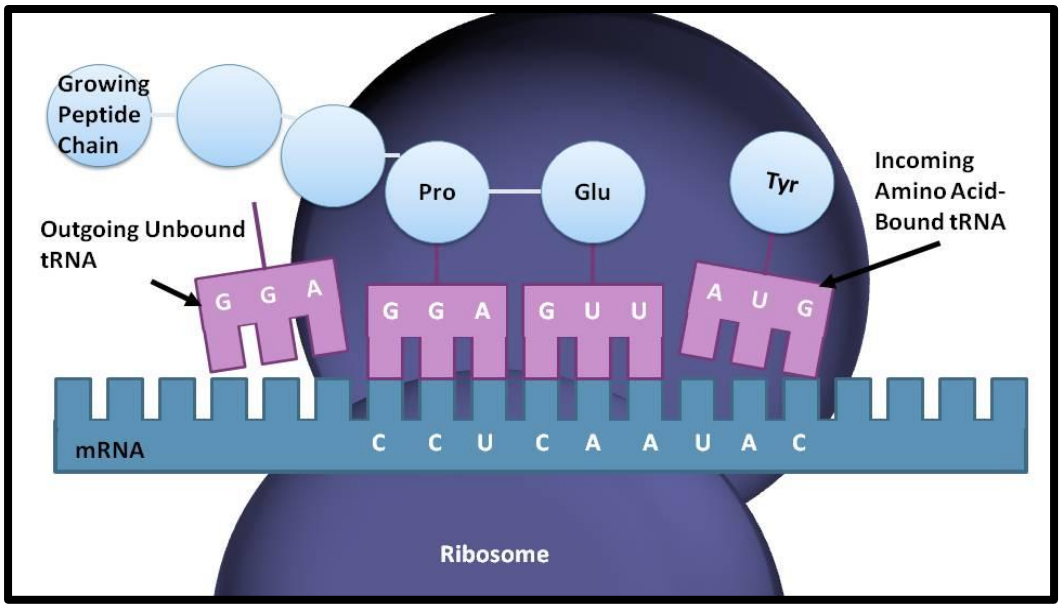


Figure 2.3: Protein synthesis basics

There are three basic methods by which this process of hijacking the endogenous cellular machinery can be accomplished.

1. *In vitro* aminoacylation

   TRNAs can be produced, purified, and aminoacylated with the UAA of

interest *in vitro*. They can then be introduced to cells in a variety of ways. Once in the cells, they will be incorporated into proteins by endogenous ribosomes just as any other aminoacylated tRNA would be.

2. Random Incorporation

   Cells can be starved of a natural amino acid then exposed to the UAA analog to that natural amino acid. They will then misincorporate the UAA indiscriminately in locations where the natural amino acid should have been incorporated. This technique alters all expressed proteins in a cell, not just specific intended proteins.

3. Site-specific Incorporation

   A mutant RS and tRNA are produced by the cell as an addition to its existing genetic code. This RS-tRNA pair can then incorporate the intended UAA at specific locations within proteins designated by the researcher.

Each of these techniques has been successfully employed in mammalian cells, and each has its own set of strengths and weaknesses which will be discussed in the following sections.

IN VITRO AMINOACYLATION OF UNNATURAL AMINO ACIDS IN LIVING CELLS

Monahan and colleagues developed a system for UAA incorporation in mammalian cells through *in vitro* aminoacylation of amber suppressor tRNA (60). Amber suppressor tRNA is any tRNA bearing the anticodon for an amber stop codon, whose DNA sequence is TAG. This stop codon has been used extensively in the field of UAA incorporation because it is the least frequently used stop codon in mammalian cells and

therefore will result in the least number of perturbations of endogenous protein expression. After aminoacylating amber suppressor tRNA *in vitro*, mRNA coding for the protein of interest, in this case a GFP-based reporter protein or a nicotinic acetylcholine receptor (nAchR), was co-electroporated with the aminoacylated tRNA into CHO cells. Successful incorporation of the UAA was evidenced by a shift in the acetylcholine dose response curve of cells expressing the UAA-receptor in comparison to those expressing the wildtype receptor (60) .

Six years after the Monahan manuscript was published, a new version of *in vitro* aminoacylation appeared in the literature. Instead of aminoacylating an amber suppressor tRNA, they used a tRNA with a frame-shift anticodon. In effect, a tRNA with a four base (versus the traditional three base) anticodon was used to translate proteins in *Xenopus* oocytes (61). The same *in vitro* aminoacylation approach was used, but instead of electroporating, the oocytes were simply injected with mRNA for the protein of interest and the aminoacylated frameshift tRNA. Since the chosen UAA in this investigation was a small-molecule fluorescent UAA, the success of incorporation was assessed via fluorescence microscopy (61).

While Monahan and colleagues achieved proof of concept of the idea of *in vitro* aminoacylation for UAA incorporation in living cells, Pantoja and colleagues demonstrated proof of concept of aminoacylation of frameshift tRNAs for UAA incorporation in living cells. Each investigation resulted in high incorporation efficiencies as well as quick results (60, 61). However, these experiments are short-lived and can only be used to study phenomenon occurring over small time frames. Since the cells studied are not producing the machinery with which to incorporate the UAA, they will only incorporate it as long as they have aminoacylated tRNA to use. Once that tRNA is exhausted, UAAs will no longer be incorporated. Additionally it should be noted that *in*

vitro aminoacylation for UAA incorporation in mammalian cells can be quite cumbersome without the appropriate transfection technology. Monahan and colleagues used a microelectroporator which required total electroporation volumes of less than 5 µL and a total amount of aminoacylated tRNA+DNA of 32.5 µg in that same volume (60). With a conventional electroporator which uses a volume between 100 and 400 µL, that same experiment would take between 0.65 and 2.6 milligrams of aminoacylated tRNA+DNA per condition. That is an enormous amount of DNA alone, but coupled with the necessary amount of aminoacylated tRNA, the materials costs alone for an experiment of that size would be prohibitive. As will be discussed in Chapter 3, large amounts of DNA are necessary for successful electroporation with conventional electroporators, but clearly, with the proper equipment, the raw amounts of DNA necessary for a successful experiment need not be so unreasonably gargantuan. So while *in vitro* aminoacylation can provide quick, effective, UAA incorporation, it is not the most versatile of techniques and is not an accessible technology for researchers without the necessary equipment.

RANDOM INCORPORATION OF UNNATURAL AMINO ACIDS IN LIVE CELLS

The Tirrell group at the California Institute of Technology has pioneered the use of random incorporation of UAAs into proteins in live cells (62-64). The general protocol involves depleting cells of the natural analog of the UAA to be used, treating them for a defined period of time with the UAA, and finally labeling the UAA with a fluorescent probe for visualization (63). Beatty et al successfully used this strategy to label proteins expressed during defined time points with different colors of fluorophores with reactive groups specific to the UAAs used (63). Though useful, this technique is imperfect. First, the labeling technique requires treatment with $CuSO_4$, which is toxic for most cells.

Second, it was found that cells with azidohomoalanine (Aha) incorporated and then reacted with the fluorescent probe bodipy-630-alkyne were indistinguishable from control cells unlabeled with Aha when using confocal microscopy, though fluorescence activated cell sorting (FACS) was able to differentiate between the two (63).

Using the same UAA, Aha, Beatty and colleagues were able to develop a more biologically inert labeling technique. Instead of copper catalyzed addition, azid-alkyne cycloaddition was used to label randomly incorporated UAAs in rat fibroblasts (62). Dimethylaminocoumarin (DMAC) was conjugated to a variety of cyclooctyne acids to create coumarin-cyclooctyne conjugate molecules capable of labeling Aha in live cells (62). Rat fibroblasts were depleted of methionine then labeled with Aha for four hours. Cells were then incubated with coumarin-cyclooctyne conjugates to afford fluorescently labeled Aha containing proteins. Confocal microscopy and FACs were used to confirm the presence of coumarin in these cells. Cell viability assays were also performed to assess the impact of the coumarin-cyclooctyne labeling method on cell survival, which was negligible (62).

Song and colleagues actually took this technology one step further by developing a fluorescent tetrazole-based probe capable of photo-crosslinking to an alkene bearing UAA (65). Homoallylglycine (HAG) was used as an analog for methionine in Hela cells previously deprived of methionine. Cells were lysed and tetrazole probes added to the lysate immediately before irradiating lysates with 305 nm light for ten minutes. Denaturing sodium-dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE) buffer was added to fluorescently labeled cell lysates, they were separated via SDS-PAGE, and fluorescence was assessed in the gel. It was found that 1 mM HAG resulted in efficient labeling of all proteins synthesized in the presence of HAG, and this fluorescence peaked twenty four hours post HAG addition (65). One very interesting

thing completed by this group and not others was a mass spectrometry analysis of the percentage of methionine sites actually occupied by HAG. A His-tagged protein was expressed under HAG labeling conditions, purified, digested, and analyzed via nanoLC-tandem mass spectrometry. It was found that each methionine site had a different HAG occupancy with no real pattern to which sites were more or less frequently labeled, and the frequency varied among sites between 42.3% and 83.7% (65). It was also found that proteins could be labeled in live cells by simply incubating them in tetrazole-probe containing buffer for ten minutes, then irradiating them with 302 nm light for five minutes. Using FACS, it was determined that HAG labeled cells were exhibited an eleven fold increase in mean fluorescence over non-HAG labeled cells.

The novelty and scientific elegance of the preceding studies cannot be argued, however this technique is not yet viable for the study of PPIs. Rather, it is an excellent tool for the study of temporally regulated global protein expression studied. Researchers can quite easily label proteins expressed during a given time period, and slight modifications of the experiments in the preceding paragraphs could generate a substantial amount of useful information. For instance, one can imagine that stem cells in the process of differentiation could be labeled during the first 2 days of differentiation, globally, with a UAA capable of blue fluorescent labeling. On day three, they could be labeled with a UAA capable of red fluorescent labeling. After labeling with the second UAA, cells could be conjugated with their fluorescent labels, lysed, and their protein expression profiles examined by simple SDS-PAGE. A series of blots could be completed by probing, stripping, and re-probing the same membrane with antibodies for proteins of interest in the differentiation process. By comparing the fluorescence and western blot images it would be possible to compare proteins highly expressed during the early phase of differentiation with those expressed during later phases.

Obviously the time points of two days with one label and the third day with another could be changed, even to a number of hours, as Beatty and colleagues saw efficient labeling within that period of time (63). While the technique of random UAA incorporation is not necessarily viable for the study of PPIs as it stands, it may yet find a place in the study of differentiation, dedifferentiation, and tumorigenesis.

SITE SPECIFIC INCORPORATION OF UNNATURAL AMINO ACIDS IN LIVE CELLS

The last and final type of UAA incorporation to be covered in this dissertation, and the one which this project has chosen to utilize, is site-specific incorporation of UAAs. This particular method of UAA incorporation is perhaps the most involved, but as will be discussed in the following pages, it may also be the most useful for the study of PPIs in mammalian cells. In order to site-specifically incorporate UAAs into proteins in living cells, a codon must be appropriated to stand for the UAA, an RS must be developed which specifically recognizes the desired UAA, and a tRNA must be developed which is recognized by the RS and not by any endogenous RSs in the cell line to be used. Basically, a site-specific UAA incorporation system must have the following characteristics in order to function as intended:

1) UAAs must be incorporated in response to a unique or unused codon.

2) The tRNA bearing the unique codon must not be recognized by endogenous synthetases.

3) The UAA specific RS must recognize only the intended tRNA and UAA, and it must discriminate against canonical amino acids and endogenous tRNAs.

While such a UAA incorporation system has been accomplished in bacteria and yeast (51, 56, 66-68), no global system satisfying all three requirements has yet been developed for use in multiple cell types. Thus individual researchers have developed

methods of evolving RSs and tRNAs with specificity for each other and various UAAs of interest (66). First, researchers start by selecting a tRNA and RS pair from a different organism than the one in which they want to incorporate the UAA. The tRNA from that organism is then mutated at locations known to be involved in tRNA specificity (discussed below) such that there is little to no reactivity of that tRNA with the endogenous RSs in the organism in which the tRNA will be used. The chosen RS is then mutated at residues selected based upon the crystal structure of a similar RS, and a library of mutant RSs is built. The RS library is then subjected first to a round of positive selection in *E. coli*, which only allows cells expressing an RS capable of aminoacylating the previous developed tRNA with the UAA of interest to survive. Cells which survive the positive selection are then subjected to negative selection in which cells containing RSs that aminoacylate the previously developed tRNA with a canonical amino acid survive. Colonies which die in the negative selection round are the ones that do not charge the tRNA with a canonical amino acid, and they are picked from a replica plate and expanded for further analysis (66). This technique and variations of it have been used to develop a number of tRNA-RS pairs capable of incorporating a wide variety of UAAs into proteins in a number of cell types (55, 56, 66-70), but those of the greatest importance for this work are the ones which have been applied to mammalian cells. These types of systems are unique in that they cannot be evolved in the cells in which they are intended to function. Mammalian cells are not capable of the high-throughput library screening employed in positive and negative selection necessary for tRNA-RS development. Thus, most researchers have chosen to evolve these pairs in yeast and then transfer them to the mammalian cell line of their choice.

One of the earliest examples of this methodology being applied to protein translation in mammalian cells came from Sakamoto et al in 2002 (71). A *Bacillus*

stearothermophilus amber suppressor tRNA ($tRNA^{B.stear}$) was used with a previously evolved *E. coli* tyrosine tRNA synthetase which recognizes $tRNA^{B.stear}$ and charges it with the UAA 3-iodo-L-tyrosine (71). A reporter protein was created by adding an amber stop codon to the sequence of cyan fluorescent protein (CFP), and expression of CFP was assessed visually. As another measure of UAA incorporation, a separate reporter protein based upon Ras was used in two forms: wildtype and amber stop codon mutated. Cells expressing both the wildtype and amber stop codon mutated versions of Ras were cultured in media containing 0.3 mM 3-iodo-L-tyrosine, and UAA incorporation was assessed via western blot and mass spectrometry (71). It is important to note that the highest expression level achieved in this study was about one fourth the level of expression seen in the wildtype protein, indicating that while UAA incorporation is possible in mammalian cells, it is inefficient. Examination of the western blots produced in this experiment may lead one to conclude that UAA incorporation is only slightly inefficient, but this would be an erroneous conclusion as the surface area of cells used for the amber stop codon mutated Ras gene was five times the surface area used for wildtype Ras protein purification (71). Despite the obvious inefficiency of UAA incorporation in this case, it did serve as a clear proof of concept that exogenous tRNAs and RSs could be functionally expressed in mammalian cells.

Another instance of site-specific UAA incorporation in mammalian cells was published in 2005. Zhang et al used a slightly different methodology to incorporate 5-hyrdoxytryptophan (5-Htrp) into proteins in mammalian cells (72). Instead of utilizing the amber stop codon, as many have done in the past, Zhang and colleagues designed a system that incorporated 5-Htrp into proteins in response to an opal stop codon with the sequence TGA (72). Cells were co-transfected with opal suppressor tRNA, an exogenous RS specific for 5-Htrp, and reporter protein with an opal suppressor stop

codon inserted into the middle of its DNA sequence. 5-Htrp incorporation was assessed via western blot and mass spectrometry, which confirmed the incorporation of 5-Htrp into the reporter protein at the opal stop codon (72). Since 5-Htrp exhibits substantial absorbance at 310 nm, the absorbance spectra of wildtype and 5-Htrp containing proteins were compared. This revealed an eleven fold increase in 310 nm absorbance of the 5-Htrp containing protein in comparison to the wildtype protein. Additionally, it was demonstrated the 5-Htrp can be redox cross-linked by applying a positive potential to a solution of 5-Htrp containing protein (72). Of note is the assertion by the authors of this study that "the exact mechanism of the protein crosslinking mediated by [5-Htrp] is not yet clear" (72). It seems that the field of stem cell research is not the only one willing to use incomplete information in order to further their scientific endeavors.

My colleagues and I have recently developed and demonstrated the use of a new method for development of a UAA incorporations system in mammalian cells (69). Both tRNAs and RSs have specific domains within their structure that have been implicated in species as well as UAA specificity. These exact identifiers will be discussed in greater detail below. Our system uses rational design to manipulate these identifiers and quickly create orthogonal, functional UAA incorporation systems for use in mammalian cells. A previously evolved *Methanocaldococcus jannaschii (M. jannaschii)* tRNA-RS pair served as the jumping off point for this system. It is known that the wildtype *M. jannaschii* tRNA shares some species specific identifiers with human tRNA and thus would not be a good candidate for use in mammalian cells without alteration of these identifiers (69). To create tRNA orthogonality, the species specific identifiers in *M. jannaschii* tyrosyl tRNA ($tRNA^{MJtyr}$) were mutated to match those of *E. coli* tRNA ($tRNA^{Ectyr}$). Once that mutation was introduced, the $tRNA^{MJtyr}$ no longer functioned with its native RS. To restore RS function, the tRNA recognition site of *M. jannaschii* tyrosine-RS (RS^{mjtyr}) was

replaced with that of the *E. coli* tyrosine-RS (69). The success of this cut-paste approach was demonstrated by successful expression of two separate TAG mutated genes. In the initial development phase, when a quick readout of tRNA-RS function was desired, a TAG mutated eGFP gene was used. This allowed visual assessment of orthogonality and function. Once this was established, another TAG mutated protein was overexpressed in HEK293T cells and assessed via western blot. While no quantification of suppression efficiency was done, the authors concluded that suppression efficiency was approximately 20% of wildtype expression of un-TAG-mutated proteins. This efficiency will be important in the final stages of the work presented in Chapter Six. Of note is the fact that a UAA was not incorporated in this work, but rational design of a mammalian UAA incorporation system was demonstrated by amber stop codon suppression (69). From this study and the studies mentioned in the preceding paragraphs it is clear that site-specific UAA incorporation can be achieved in mammalian cells, though it is unclear how versatile this technology may prove to be.

DETERMINANTS OF tRNA SPECIES SPECIFICITY

All tRNAs have the same basic cloverleaf structure shown in Figure 2.4, which indicates that RS discrimination of desired versus undesired tRNA cannot be based upon secondary structure alone. Though the anticodon is involved in recognition, its effect is frequently less than that of the acceptor stem nucleotides, specifically nucleotides 1 and 72 (73, 74). Wakasugi et al have demonstrated that switching the human wild type (WT) $tRNA^{Tyr}$ from C1:G72 to G1:C72 completely abolishes aminoacylation by human RS^{Tyr} (75). A subsequent study demonstrated that mutation of *M. jannaschii* $tRNA^{Tyr}$ from C1:G72 to G1:C72 prevents aminoacylation by RS^{mjTyr}(14). It has also been shown that the eukaryotic RS^{Tyr} from *Pneumocystis carinii* is dependent upon the 1:72 nucleotide

identity (76). So at least in the case of $tRNA^{Tyr}$, the 1:72 base pair is able to confer or block aminoacylation in a species specific manner. In addition to the 1:72 base pair, *E. coli* RSs for Met, Cys, Trp, Val, Ile, Glu, Gln, and Arg are known to bear identifiers within their acceptor stems further than just the 1:72 pair (77). It has been posited that these interactions are based on a destabilization (or lack) of the 1:72 base pair which confers flexibility on the acceptor stem (77, 78). Thus, previous research has clearly demonstrated that it is possible to confer or knock down tRNA species specificity by mutating the nucleotides within the acceptor stem.

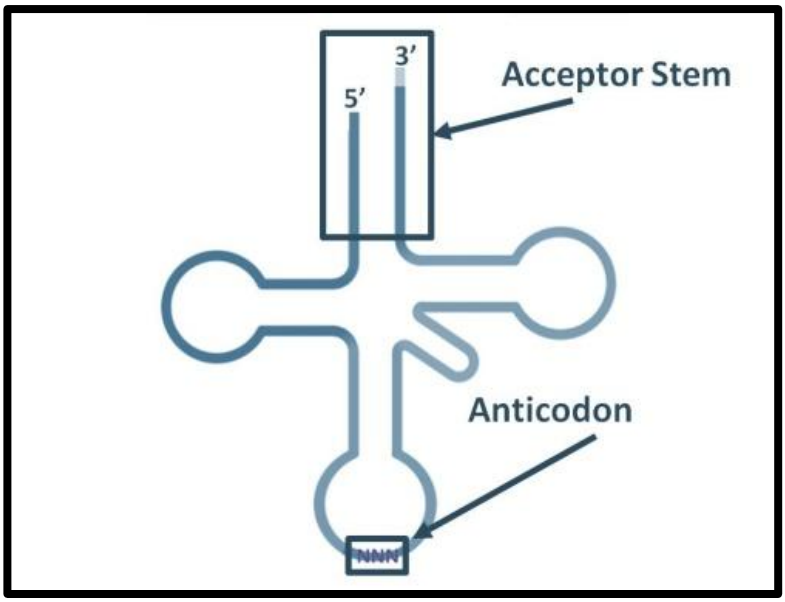


Figure 2.4: tRNA Secondary Structure

Though the effect of the anticodon sequence is less dramatic than that of the acceptor stem nucleotides, they do serve as identifiers for some tRNA/RS pairs. Since the anticodons used in UAA incorporation are necessarily different than those used for canonical amino acids, mutant UAA tRNAs with mutations only in the anticodon usually display some loss of activity as compared to their wildtype predecessors. Kobayashi et

al. solved the crystal structure for the *M. jannaschii* $tRNA^{Tyr}$/ RS^{mjTyr} pair, revealing that RS^{mjTyr} does use the anticodon in tRNA identification (79). As shown in Figure 2.5, RS^{mjTyr} interacts with the anticodon, but only at the 286th residue, and the only interaction is between RS-D286 and tRNA-G34. When creating an amber suppressor tRNA from *M. jannaschii* $tRNA^{Tyr}$, G34 is mutated to C34, switching the 34th nucleotide from a purine to a pyrimidine base. This switch effectively increases the distance over which interaction must occur between nucleotide 34 and D286, which explains the decrease in activity between wildtype and amber suppressor *M. jannaschii* $tRNA^{Tyr}$. Fortunately, mutations in the cognate RS can accommodate this anticodon mutation, but those will be discussed in the following section. Based upon the preceding information, an amber suppressor tRNA can be designed for use in a mammalian system by designing a tRNA sequence with a non-mammalian 1:72 nucleotide pair and acceptor stem sequence.

Figure 2.5: Crystal structure of M. jannaschii tRNA and RS^{Tyr} demonstrating interaction between nucleotide G34 and residue D286 (PDB 1j1u)

DETERMINANTS OF AMINOACYL-TRNA SYNTHETASE SPECIFICITY

In 1999 Schimmel and Steer identified an RS whose unique properties make it a perfect candidate for mutation and use with UAA incorporation (80). Each species carries genes for the RSs of most, if not all, of the 20 canonical amino acids, the protein products of which are organized into two classes (81). Class I RSs are defined by a Rossman nucleotide binding fold and highly conserved HIGH and KMSKS motifs, while the class II RSs are characterized by three highly conserved sequence motifs within a 7 stranded beta structure + 3 alpha helices (82). The RS identified by Schimmel and Steer is the previously mentioned class I tyrosyl RS (RS^{mjTyr}) from the archaebacterium *M. jannaschii*. This RS*mjTyr* includes the characteristic Rossman fold and HIGH and KMSKS motifs, but it lacks a C-terminal portion of the protein implicated in *E. coli* (83) and *Bacillus stearothermophilus* RS^{Tyr}s (84-86) as interacting with the anticodon (77) region of tRNA. Since mutation of anticodon nucleotide U36 to G36 only yielded a 6 fold decrease in aminoacylation of G36-tRNA versus wildtype, Schimmel and Steer suggested that RS^{mjTyr} does not directly interact with the anticodon of its cognate tRNA. As previously mentioned, Kobayashi and colleagues have demonstrated that RS^{mjTyr} does interact with the anticodon, but in a different manner than other class I RSs (79, 80). As previously mentioned, mutation of the anticodon of *M. jannaschii* $tRNA^{Tyr}$ from G34 to C34 increases the distance over which the interaction between anticodon and RS^{mjTyr} must occur. To offset the decrease in aminoacylation activity created by this mutation, a series of mutant RS^{mjTyr} enzymes were created by mutating D286 to residues with a larger functional group than aspartic acid. In this way, the distance between C34 and residue 286 could be decreased. Ultimately it was found that mutation of D286 to arginine increased aminoacylation of the amber suppressor tRNA by 8 fold, achieving 22% of the activity of the wildtype *M. jannaschii* $tRNA^{Tyr}/RS^{mjTyr}$ pair (79). The preceding

work has established that it is possible to dramatically increase amber suppressor tRNA recognition by mutating D286 of RSmjTyr to R286.

As previously mentioned, nucleotides 1 and 72 can confer or knock-down aminoacylation in a species specific manner. This indicates that the acceptor stem nucleotides interact with their cognate RS. One group investigating the role of nucleotides 1 and 72 in RS recognition successfully located the region within the RS responsible for that interaction. Wakasugi et al. used chimeric proteins to demonstrate the region within class I RS^{Tyr}s responsible for discrimination of nucleotides 1 and 72 (75). An *E. coli* RS^{Tyr} was engineered to carry an N-terminal segment of the human enzyme, called connective protein 1 (CP1) (75).The aminoacylation activity of this chimeric enzyme (Figure 2.6, *E. coli* RS^{TyrEHE}) was compared to that of the wild type (E. coli RS^{Tyr}) enzyme in aminoacylating either C1:G72 tRNA or G1:C72 tRNA, revealing that the chimeric enzyme could aminoacylate the C1:G72 tRNA but not the G1:C72 tRNA, shown in Figure 2.6 (75). By sequence comparison, the chimeric EHE enzyme is ~90% homologous with *E. coli* but was able to aminoacylate tRNA with human 1:72 nucleotides and not the tRNA with *E. coli* nucleotides in the same position. This result suggests that the CP1 region of RS^{Tyr}s is capable of conferring species specific recognition of nucleotides 1 and 72 on chimeric proteins. Since the RS^{mjTyr} investigated by Schimmel and Steer can readily accommodate anticodon mutations, and the CP1 regions of RS^{Tyr}s can be switched out to confer species specificity according to 1:72 nucleotide identities, we engineered *M. jannaschii* tRNA/ RS^{mjTyr} system in which the mutated tRNA acceptor stem was unrecognized by mammalian cells, and the RS was given a CP1 region which recognized the mutated tRNA. The development of this system is described in full detail in Chapter Four, and its application to UAA incorporation in stem cells is described in Chapter Five.

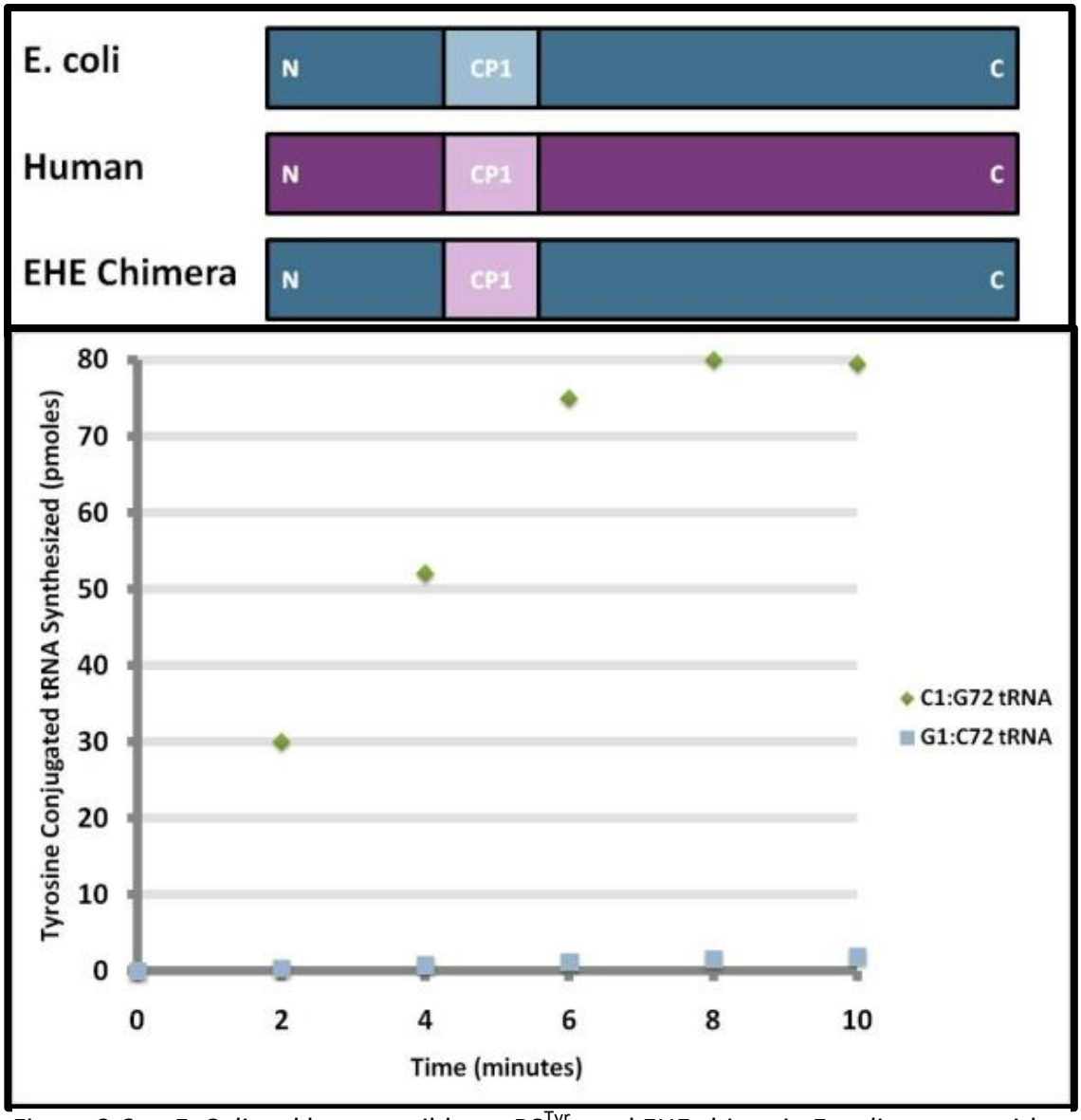


Figure 2.6: *E. Coli* and human wildtype RS^{Tyr}s and EHE chimeric *E. coli* enzyme with human CP1 region; Aminoacylation of C1:G72 versus G1:C72 tRNAs by EHE chimeric RS, adapted from Wakasugi et al (75).

A Two-Hybrid System for Use in Stem Cells

Since the aim of this entire project is to facilitate the study of PPIs in stem cells, an attempt at the use of a two-hybrid system was a logical addition to the work as a whole. A two-hybrid system is an elegant method for assessing PPIs *in vitro* and has been successfully applied to high throughput studies of PPIs (87). The basic design of a two-hybrid system consists of a reporter gene under the control of a modular transcriptional activation domain. Proteins are fused to two separate parts of the transcriptional activators, and PPIs are detected by expression of the reporter protein (87). Two-hybrid systems were initially used in yeast, owing to their adaptability to high throughput screening of two-hybrid libraries (88). However, in terms of clinical relevance, a two-hybrid system which directly screens PPIs in mammalian cells would be more advantageous, though at this time it would be incapable of easy high-throughput application. Two-hybrid systems have previously been converted for use in mammalian cells, but they were little more than simple transfers of the yeast system into mammalian cells (89) and frequently suffered from high background expression of reporter gene. My colleagues developed a mammalian two-hybrid system which improved upon previous versions, yielding lower levels of background and detecting weak interactions (90).

The system developed by my colleagues, a tetracycline-repressor based mammalian two-hybrid system (tr-M2H), is based upon the highly active tetracycline repressor (TetR), a dimeric transcriptional regulator which binds specifically to the tet operator (TetO) to inhibit expression of a downstream gene. It was developed by modification of the commercially available Tet-Off® Advanced Inducible Gene Expression System (Clontech, Mountain View, CA). The Tet-Off system is composed of the full length TetR C-terminally fused to three transcriptional activation domains from

the herpes simplex virus VP16 (90). TetR-VP16 fusion protein dimer allows the TetR dimer to bind the Tet-Responsive-Element (Tre) while the VP16 domain recruits transcriptional machinery. In the trM2H system, the TetR c-terminal dimerization domain (alpha8-alpha10) has been replaced with bait and prey molecules, shown in Figure 2.7. When bait and prey proteins interact, the reporter protein, GFP in this case, is expressed. This system was successfully used to assess interactions of peptide pairs with binding constants ranging from 0.99 nM to 55 µM (90). HEK293T cells were transfected with the components of the tr-M2H system and imaged using a fluorescence microscope 48 hours after transfection. Cells were harvested, lysed, and reporter protein expression was assessed via western blot. Both the fluorescence images and western blots confirmed that peptide pairs known to interact with each other resulted in expression of the reporter protein, while non-interacting pairs did not cause expression of the reporter protein (90). These results lead us to conclude that the system may be appropriate for use in stem cells as it exhibits very low background and high sensitivity. As part of the body of work represented in this dissertation, we investigated the use of the tr-M2H in stem cells. This work will be discussed in detail in Chapter Seven.

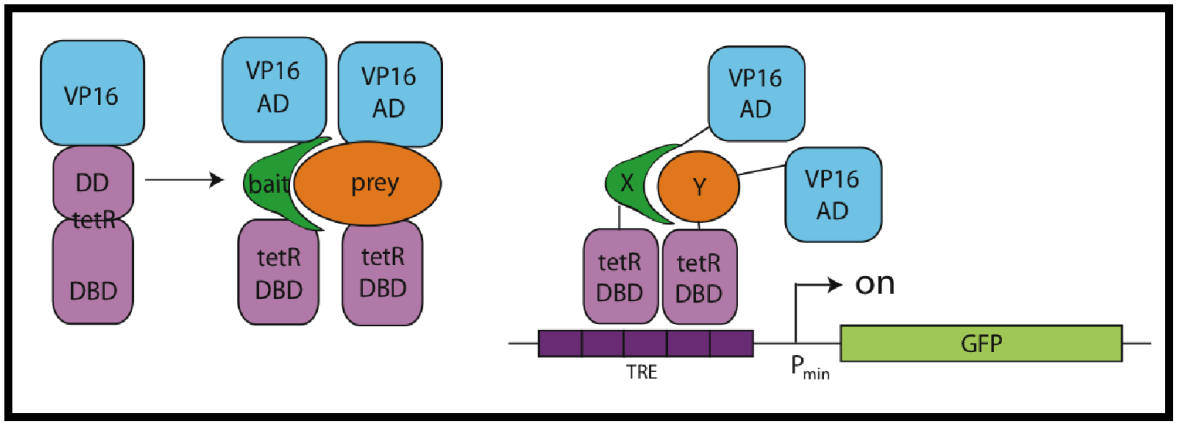


Figure 2.7: Design of tr-M2H

By applying the body of knowledge previously generated by our group and others in the field of UAA incorporation systems as well as our own investigations of stem cell transfection techniques, this dissertation discusses attempts to transfer UAA incorporation systems into the stem cell environment and thereby facilitate the study of basic biological processes in these cells. The tr-M2H system's application to the stem cell environment will be discussed briefly. The ultimate goal of this research is to speed the study and enhance the quality of knowledge of PPIs in stem cells by utilizing novel technologies.

REFERENCES

1. Carletti B, Grimaldi P, Magrassi L, Rossi F. Specification of cerebellar progenitors after heterotopic-heterochronic transplantation to the embryonic CNS in vivo and in vitro. J Neurosci. 2002;22(16):7132-46.
2. Sotelo C, Alvarado-Mallart RM. The reconstruction of cerebellar circuits. Trends Neurosci. 1991;14(8):350-5.
3. Triarhou LC, Zhang W, Lee WH. Amelioration of the behavioral phenotype in genetically ataxic mice through bilateral intracerebellar grafting of fetal Purkinje cells. Cell Transplant. 1996;5(2):269-77.
4. Amariglio N, Hirshberg A, Scheithauer BW, Cohen Y, Loewenthal R, Trakhtenbrot L, et al. Donor-derived brain tumor following neural stem cell transplantation in an ataxia telangiectasia patient. PLoS Med. 2009;6(2):e1000029.
5. Xi J, Khalil M, Spitkovsky D, Hannes T, Pfannkuche K, Bloch W, et al. Fibroblasts support functional integration of purified embryonic stem cell-derived cardiomyocytes into avital myocardial tissue. Stem Cells Dev.20(5):821-30.
6. Rizvanov AA, Guseva DS, Salafutdinov, II, Kudryashova NV, Bashirov FV, Kiyasov AP, et al. Genetically modified human umbilical cord blood cells expressing vascular endothelial growth factor and fibroblast growth factor 2 differentiate into glial cells after transplantation into amyotrophic lateral sclerosis transgenic mice. Exp Biol Med (Maywood).236(1):91-8.
7. Alhadlaq A, Elisseeff JH, Hong L, Williams CG, Caplan AI, Sharma B, et al. Adult stem cell driven genesis of human-shaped articular condyle. Ann Biomed Eng. 2004;32(7):911-23.
8. Macchiarini P, Jungebluth P, Go T, Asnaghi MA, Rees LE, Cogan TA, et al. Clinical transplantation of a tissue-engineered airway. Lancet. 2008;372(9655):2023-30.
9. Xie QP, Huang H, Xu B, Dong X, Gao SL, Zhang B, et al. Human bone marrow mesenchymal stem cells differentiate into insulin-producing cells upon microenvironmental manipulation in vitro. Differentiation. 2009;77(5):483-91.
10. Tayaramma T, Ma B, Rohde M, Mayer H. Chromatin-remodeling factors allow differentiation of bone marrow cells into insulin-producing cells. Stem Cells. 2006;24(12):2858-67.
11. Sun Y, Chen L, Hou XG, Hou WK, Dong JJ, Sun L, et al. Differentiation of bone marrow-derived mesenchymal stem cells from diabetic patients into insulin-producing cells in vitro. Chin Med J (Engl). 2007;120(9):771-6.

12. Mabed M. The potential utility of bone marrow or umbilical cord blood transplantation for the treatment of type I diabetes mellitus. Biol Blood Marrow Transplant. 2011;17(4):455-64.

13. Dai LJ, Moniri MR, Zeng ZR, Zhou JX, Rayat J, Warnock GL. Potential implications of mesenchymal stem cells in cancer therapy. Cancer Lett. 2011;305(1):8-20.

14. Ao A, Hao J, Hong CC. Regenerative chemical biology: current challenges and future potential. Chem Biol. 2011;18(4):413-24.

15. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell. 2006;126(4):663-76.

16. Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, Tian S, et al. Induced pluripotent stem cell lines derived from human somatic cells. Science. 2007;318(5858):1917-20.

17. Cho HJ, Lee CS, Kwon YW, Paek JS, Lee SH, Hur J, et al. Induction of pluripotent stem cells from adult somatic cells by protein-based reprogramming without genetic manipulation. Blood. 2010;116(3):386-95.

18. Okita K, Nakagawa M, Hyenjong H, Ichisaka T, Yamanaka S. Generation of mouse induced pluripotent stem cells without viral vectors. Science. 2008;322(5903):949-53.

19. Warren L, Manos PD, Ahfeldt T, Loh YH, Li H, Lau F, et al. Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. Cell Stem Cell. 2010;7(5):618-30.

20. Hasegawa K, Zhang P, Wei Z, Pomeroy JE, Lu W, Pera MF. Comparison of reprogramming efficiency between transduction of reprogramming factors, cell-cell fusion, and cytoplast fusion. Stem Cells.28(8):1338-48.

21. Ernst M, Oates A, Dunn AR. Gp130-mediated signal transduction in embryonic stem cells involves activation of Jak and Ras/mitogen-activated protein kinase pathways. J Biol Chem. 1996;271(47):30136-43.

22. Ernst M, Novak U, Nicholson SE, Layton JE, Dunn AR. The carboxyl-terminal domains of gp130-related cytokine receptors are necessary for suppressing embryonic stem cell differentiation. Involvement of STAT3. J Biol Chem. 1999;274(14):9729-37.

23. Freemantle SJ, Kerley JS, Olsen SL, Gross RH, Spinella MJ. Developmentally-related candidate retinoic acid target genes regulated early during neuronal differentiation of human embryonal carcinoma. Oncogene. 2002;21(18):2880-9.

24. Damjanov I, Andrews PW. The terminology of teratocarcinomas and teratomas. Nat Biotechnol. 2007;25(11):1212; discussion.

25. Stevens LC. The development of transplantable teratocarcinomas from intratesticular grafts of pre- and postimplantation mouse embryos. Dev Biol. 1970;21(3):364-82.

26. McBurney MW, Rogers BJ. Isolation of male embryonal carcinoma cells and their chromosome replication patterns. Dev Biol. 1982;89(2):503-8.

27. McBurney MW. Clonal lines of teratocarcinoma cells in vitro: differentiation and cytogenetic characteristics. J Cell Physiol. 1976;89(3):441-55.

28. Nicolas JF, Dubois P, Jakob H, Gaillard J, Jacob F. [Mouse teratocarcinoma: differentiation in cultures of a multipotential primitive cell line (author's transl)]. Ann Microbiol (Paris). 1975;126(1):3-22.

29. McBurney MW. P19 embryonal carcinoma cells. Int J Dev Biol. 1993;37(1):135-40.

30. Jones-Villeneuve EM, McBurney MW, Rogers KA, Kalnins VI. Retinoic acid induces embryonal carcinoma cells to differentiate into neurons and glial cells. J Cell Biol. 1982;94(2):253-62.

31. Mullen RJ, Buck CR, Smith AM. NeuN, a neuronal specific nuclear protein in vertebrates. Development. 1992;116(1):201-11.

32. McBurney MW, Jones-Villeneuve EM, Edwards MK, Anderson PJ. Control of muscle and neuronal differentiation in a cultured embryonal carcinoma cell line. Nature. 1982;299(5879):165-7.

33. McBurney MW, Reuhl KR, Ally AI, Nasipuri S, Bell JC, Craig J. Differentiation and maturation of embryonal carcinoma-derived neurons in cell culture. J Neurosci. 1988;8(3):1063-73.

34. Rudnicki MA, Sawtell NM, Reuhl KR, Berg R, Craig JC, Jardine K, et al. Smooth muscle actin expression during P19 embryonal carcinoma differentiation in cell culture. J Cell Physiol. 1990;142(1):89-98.

35. Edwards MK, Harris JF, McBurney MW. Induced muscle differentiation in an embryonal carcinoma cell line. Mol Cell Biol. 1983;3(12):2280-6.

36. Jasmin, Spray DC, Campos de Carvalho AC, Mendez-Otero R. Chemical induction of cardiac differentiation in p19 embryonal carcinoma stem cells. Stem Cells Dev.19(3):403-12.

37. Brown K, Legros S, Artus J, Doss MX, Khanin R, Hadjantonakis AK, et al. A comparative analysis of extra-embryonic endoderm cell lines. PLOS ONE.5(8):e12016.

38. Monzen K, Shiojima I, Hiroi Y, Kudoh S, Oka T, Takimoto E, et al. Bone morphogenetic proteins induce cardiomyocyte differentiation through the mitogen-activated protein kinase kinase kinase TAK1 and cardiac transcription factors Csx/Nkx-2.5 and GATA-4. Mol Cell Biol. 1999;19(10):7096-105.

39. van der Heyden M, Defize L. Twenty one years of P19 cells: what an embryonal carcinoma cell line taught us about cardiomyocyte differentiation. CARDIOVASC RES. 2003;58(2):292-302.

40. Tan Y, Xie Z, Ding M, Wang Z, Yu Q, Meng L, et al. Increased levels of FoxA1 transcription factor in pluripotent P19 embryonal carcinoma cells stimulate neural differentiation. Stem Cells Dev. 2010;19(9):1365-74.

41. Xie Z, Tan G, Ding M, Dong D, Chen T, Meng X, et al. Foxm1 transcription factor is required for maintenance of pluripotency of P19 embryonal carcinoma cells. Nucleic Acids Res. 2010;38(22):8027-38.

42. Ma Y, Ramezani A, Lewis R, Hawley RG, Thomson JA. High-level sustained transgene expression in human embryonic stem cells using lentiviral vectors. Stem Cells. 2003;21(1):111-7.

43. Gropp M, Itsykson P, Singer O, Ben-Hur T, Reinhartz E, Galun E, et al. Stable genetic modification of human embryonic stem cells by lentiviral vectors. Mol Ther. 2003;7(2):281-7.

44. Tinsley RB, Faijerson J, Eriksson PS. Efficient non-viral transfection of adult neural stem/progenitor cells, without affecting viability, proliferation or differentiation. J Gene Med. 2006;8(1):72-81.

45. Zernecke A, Erl W, Fraemohs L, Lietz M, Weber C. Suppression of endothelial adhesion molecule up-regulation with cyclopentenone prostaglandins is dissociated from IkappaB-alpha kinase inhibition and cell death induction. Faseb J. 2003;17(9):1099-101.

46. Yan CN, Li F, Patterson C, Runge MS. High-voltage and high-salt buffer facilitates electroporation of human aortic smooth-muscle cells. Biotechniques. 1998;24(4):590-2.

47. Shimomura O, Johnson FH, Saiga Y. Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea. J Cell Comp Physiol. 1962;59:223-39.

48. Tsien RY. The green fluorescent protein. Annu Rev Biochem. 1998;67:509-44.

49. Miyawaki A, Llopis J, Heim R, McCaffery JM, Adams JA, Ikura M, et al. Fluorescent indicators for Ca2+ based on green fluorescent proteins and calmodulin. Nature. 1997;388(6645):882-7.

50. Giepmans BN, Adams SR, Ellisman MH, Tsien RY. The fluorescent toolbox for assessing protein location and function. Science. 2006;312(5771):217-24.

51. Liu CC, Schultz PG. Recombinant expression of selectively sulfated proteins in Escherichia coli. Nat Biotechnol. 2006;24(11):1436-40.

52. Liu CC, Schultz PG. Adding new chemistries to the genetic code. Annu Rev Biochem. 2010;79:413-44.

53. Yang W, Hendrickson WA, Crouch RJ, Satow Y. Structure of ribonuclease H phased at 2 A resolution by MAD analysis of the selenomethionyl protein. Science. 1990;249(4975):1398-405.

54. Mills JH, Lee HS, Liu CC, Wang J, Schultz PG. A genetically encoded direct sensor of antibody-antigen interactions. CHEMBIOCHEM. 2009;10(13):2162-4.

55. Summerer D, Chen S, Wu N, Deiters A, Chin JW, Schultz PG. A genetically encoded fluorescent amino acid. Proc Natl Acad Sci U S A. 2006;103(26):9785-9.

56. Umeda A, Thibodeaux G, Zhu J, Lee Y, Zhang Z. Site-specific Protein Cross-Linking with Genetically Incorporated 3,4-Dihydroxy-L-Phenylalanine. CHEMBIOCHEM. 2009;10(8):1302-4.

57. Umeda A, Thibodeaux GN, Moncivais K, Jiang F, Zhang ZJ. A Versatile Approach to Transform Low-Affinity Peptides into Protein Probes with Co-Translationally Expressed Chemical Cross-Linker. Anal Biochem.

58. Hino N, Okazaki Y, Kobayashi T, Hayashi A, Sakamoto K, Yokoyama S. Protein photo-cross-linking in mammalian cells by site-specific incorporation of a photoreactive amino acid. Nat Methods. 2005;2(3):201-6.

59. Wang W, Takimoto JK, Louie GV, Baiga TJ, Noel JP, Lee KF, et al. Genetically encoding unnatural amino acids for cellular and neuronal studies. Nat Neurosci. 2007;10(8):1063-72.

60. Monahan SL, Lester HA, Dougherty DA. Site-specific incorporation of unnatural amino acids into receptors expressed in Mammalian cells. Chem Biol. 2003;10(6):573-80.

61. Pantoja R, Rodriguez EA, Dibas MI, Dougherty DA, Lester HA. Single-molecule imaging of a fluorescent unnatural amino acid incorporated into nicotinic receptors. Biophys J. 2009;96(1):226-37.

62. Beatty KE, Fisk JD, Smart BP, Lu YY, Szychowski J, Hangauer MJ, et al. Live-cell imaging of cellular proteins by a strain-promoted azide-alkyne cycloaddition. CHEMBIOCHEM. 2010;11(15):2092-5.

63. Beatty KE, Tirrell DA. Two-color labeling of temporally defined protein populations in mammalian cells. Bioorg Med Chem Lett. 2008;18(22):5995-9.

64. Dieterich DC, Hodas JJ, Gouzer G, Shadrin IY, Ngo JT, Triller A, et al. In situ visualization and dynamics of newly synthesized proteins in rat hippocampal neurons. Nat Neurosci.13(7):897-905.

65. Song W, Wang Y, Yu Z, Vera CI, Qu J, Lin Q. A metabolic alkene reporter for spatiotemporally controlled imaging of newly synthesized proteins in Mammalian cells. ACS Chem Biol.5(9):875-85.

66. Wang L, Brock A, Herberich B, Schultz PG. Expanding the genetic code of Escherichia coli. Science. 2001;292(5516):498-500.

67. Wang L, Schultz PG. Expanding the genetic code. Chem Commun (Camb). 2002(1):1-11.

68. Xie J, Schultz PG. An expanding genetic code. Methods. 2005;36(3):227-38.

69. Thibodeaux G, Liang X, Moncivais K, Umeda A, Singer O, Alfonta L, et al. Transforming a Pair of Orthogonal tRNA-aminoacyl-tRNA Synthetase from Archaea to Function in Mammalian Cells. PLOS ONE. 2010;5(6):-.

70. Ye S, Kohrer C, Huber T, Kazmi M, Sachdev P, Yan EC, et al. Site-specific incorporation of keto amino acids into functional G protein-coupled receptors using unnatural amino acid mutagenesis. J Biol Chem. 2008;283(3):1525-33.

71. Sakamoto K, Hayashi A, Sakamoto A, Kiga D, Nakayama H, Soma A, et al. Site-specific incorporation of an unnatural amino acid into proteins in mammalian cells. NUCLEIC ACIDS RES. 2002;30(21):4692-9.

72. Zhang Z, Alfonta L, Tian F, Bursulaya B, Uryu S, King DS, et al. Selective incorporation of 5-hydroxytryptophan into proteins in mammalian cells. Proc Natl Acad Sci U S A. 2004;101(24):8882-7.

73. Schimmel P, Giege R, Moras D, Yokoyama S. An operational RNA code for amino acids and possible relationship to genetic code. Proc Natl Acad Sci U S A. 1993;90(19):8763-8.

74. Giege R, Puglisi JD, Florentz C. tRNA structure and aminoacylation efficiency. Prog Nucleic Acid Res Mol Biol. 1993;45:129-206.

75. Wakasugi K, Quinn CL, Tao N, Schimmel P. Genetic code in evolution: switching species-specific aminoacylation with a peptide transplant. Embo J. 1998;17(1):297-305.

76. Quinn CL, Tao N, Schimmel P. Species-specific microhelix aminoacylation by a eukaryotic pathogen tRNA synthetase dependent on a single base pair. Biochemistry. 1995;34(39):12489-95.

77. Giege R, Sissler M, Florentz C. Universal rules and idiosyncratic features in tRNA identity. Nucleic Acids Res. 1998;26(22):5017-35.

78. Alexander RW, Nordin BE, Schimmel P. Activation of microhelix charging by localized helix destabilization. Proc Natl Acad Sci U S A. 1998;95(21):12214-9.

79. Kobayashi T, Nureki O, Ishitani R, Yaremchuk A, Tukalo M, Cusack S, et al. Structural basis for orthogonal tRNA specificities of tyrosyl-tRNA synthetases for genetic code expansion. Nat Struct Biol. 2003;10(6):425-32.

80. Steer BA, Schimmel P. Major anticodon-binding region missing from an archaebacterial tRNA synthetase. J Biol Chem. 1999;274(50):35601-6.

81. Webster T, Tsai H, Kula M, Mackie GA, Schimmel P. Specific sequence homology and three-dimensional structure of an aminoacyl transfer RNA synthetase. Science. 1984;226(4680):1315-7.

82. Fabrega C, Farrow MA, Mukhopadhyay B, de Crecy-Lagard V, Ortiz AR, Schimmel P. An aminoacyl tRNA synthetase whose sequence fits into neither of the two known classes. Nature. 2001;411(6833):110-4.

83. Himeno H, Hasegawa T, Ueda T, Watanabe K, Shimizu M. Conversion of aminoacylation specificity from tRNA(Tyr) to tRNA(Ser) in vitro. Nucleic Acids Res. 1990;18(23):6815-9.

84. Bedouelle H. Recognition of tRNA(Tyr) by tyrosyl-tRNA synthetase. Biochimie. 1990;72(8):589-98.

85. Bedouelle H, Guez-Ivanier V, Nageotte R. Discrimination between transfer-RNAs by tyrosyl-tRNA synthetase. Biochimie. 1993;75(12):1099-108.

86. Bedouelle H, Winter G. A model of synthetase/transfer RNA interaction as deduced by protein engineering. Nature. 1986;320(6060):371-3.

87. Phizicky EM, Fields S. Protein-protein interactions: methods for detection and analysis. Microbiol Rev. 1995;59(1):94-123.

88. Dang CV, Barrett J, Villa-Garcia M, Resar LM, Kato GJ, Fearon ER. Intracellular leucine zipper interactions suggest c-Myc hetero-oligomerization. Mol Cell Biol. 1991;11(2):954-62.

89. Luo Y, Batalao A, Zhou H, Zhu L. Mammalian two-hybrid system: a complementary approach to the yeast two-hybrid system. Biotechniques. 1997;22(2):350-2.

90. Thibodeaux GN, Cowmeadow R, Umeda A, Zhang Z. A tetracycline repressor-based mammalian two-hybrid system to detect protein-protein interactions in vivo. Anal Biochem. 2009;386(1):129-31.

CHAPTER THREE

Transfection Optimization in P19 Embryonal Carcinoma Cells

INTRODUCTION

Recent advances in the field stem cell research and therapy make it an attractive scientific problem for many researchers. In Chapter Two a brief overview of stem cell based therapies was discussed. This included an implanted tracheal graft that had been cellularized with the patient's own cells (1) as well as tissue engineered articular condyles (2). There have also been cases of embryonic stem cells (ESCs) being transplanted from a donor to a disease model recipient, and some of these cases resulted in disappearance or improvement of some of the disease related symptoms (3, 4). However, in some cases of ESC transplantation, including the one case of transplantation of ESCs into a human being, cancer can develop from the transplanted cells (5). With the current successes in the field of stem cell science, one can only imagine the therapies that could result from a more controlled, successful manipulation of stem cell behavior. To that end, this dissertation will discuss the development and implementation of tools for the study of protein-protein interactions (PPIs) in stem cells.

In Chapter Two unnatural amino acid (UAA) incorporation was discussed as a means of studying PPIs in living cells, as was the tetracycline-repressor based mammalian two-hybrid system (tr-M2H). In order to successfully apply either of those systems to stem cells, a method must first be established for introduction of foreign DNA to these cells. Obviously, this is not the first line of research to attempt transfection of stem cells, therefore we started by investigating the methods used in the literature for stem cell transfection and designed our optimization experiments from there. Since

P19 embryonal carcinoma cells have been widely used, there are several previous examples of transfection in this cell line.

The earliest transfection experiments in P19s used either calcium phosphate precipitation or electroporation as their method of choice (6), but commercial reagents and proprietary electroporation based technologies have replaced them in recent years. Despite the advent of new 'better' technologies, viral transfection methods continue to be the most effective if not the most widely used. In 1992 Schmidt et al used a retroviral system to transfect P19s with a gene encoding c-src (7). Though they did not report measuring transfection efficiency directly, western blots of the desired protein products confirmed successful transfection and expression of detectable levels of protein (7). There is one fundamental difference between the goals and requirements of the Schmidt project and those of the project discussed in this dissertation: This project requires expression of two separate proteins and a tRNA. This means that viral transfection methods require much more preparation and time for this project than for those investigating expression or knockdown of a single gene. Thus, we chose to investigate non-viral methods before investigating viral methods for simplicity's sake.

In other work in our lab, HEK293T cells are routinely transfected with Fugene6 and FugeneHD, each of which achieve close to 100% transfection efficiency in this cell line when used at the appropriate confluency. Thus it was convenient when Tan and colleagues published their work in which they used Fugene6 to transfect two separate DNA constructs into P19s (8), effectively giving us scientific license to use the materials already present in the lab. They too did not comment on the transfection efficiency of Fugene6, but it is noteworthy that they also used an adenovirus based transfection method for more protein intensive components of their work (8). In contrast to the preceding studies which do not include quantitative measurements of transfection

efficiency, Lonza's Amaxa® Cell Line Nucleofector Kit provides readily accessible information regarding its success in transfecting P19s (9). This documentation includes images of transfected P19s expressing green fluorescent protein (GFP) along with the assertion that the transfection efficiency is higher than 80% (9). The images do not seem to represent 80% transfection efficiency, but fluorescence activated cell sorting (FACS) is more sensitive than fluorescence microscopy. So perhaps the image did not capture as much information as FACS. Regardless, even 50% transfection efficiency would be remarkable. Thus, Nucleofection seemed like a good method to test. Unfortunately, the Nucleofector is prohibitively expensive, and each transfection requires the use of a kit, further adding to the cost of this system. In light of this and the preceding information, we chose to test and optimize Fugene6, FugeneHD, and electroporation to start. We eventually tested a lentiviral transfection method as well.

MATERIALS AND METHODS

Cell Lines and Routine Cell Culture

P19 embryonal carcinoma cells were obtained from American Type Culture Collection (ATCC). Cells were routinely maintained in Alpha-Mem (HyClone, Logan, UT) with 10% (v/v) fetal bovine serum (Atlanta Biologicals, Lawrenceville, GA) and nonessential amino acids (Sigma-Aldrich, St. Louis, MO) in a humidified incubator at 37° Celsius with 5% CO_2 atmosphere. Cells were passaged every 1-3 days at a ratio between 1:4 and 1:20. For passaging, cells were rinsed in phosphate buffered saline (PBS) (Sigma-Aldrich, St. Louis, MO), incubated with 0.5% trypsin (Gibco, Invitrogen, Carlsbad, CA) for 3-5 minutes, and centrifuged to pellet. Cells were then resuspended in complete media (as described earlier in this paragraph) and seeded into clean flasks with fresh media.

DNA Constructs

The plasmid pEGFP-N1 was obtained from Clontech (Mountainview, CA).

Fugene6 and FugeneHD Transfection

Fugene6 and FugeneHD were obtained from Roche (Switzerland) and used according to directions. Previous experiments had shown that a ratio of 2 µL Fugene(6 or HD) to 1 µg DNA was the most effective, so that ratio was used for all experiments. Complexes were formed in Opti-mem (Gibco, Invitrogen, Carlsbad, CA). All transfection experiments were done in the absence of antibiotics, and they were also done while the cells were in complete media with 10% FBS. Media was changed 24 hours after transfection, and every 24 hours thereafter until experiments were completed.

Electroporation

Cell densities were previously optimized. Instrument parameters were then optimized using the previously determined cell density. The combinations of parameters tested are shown in Table 3.1. From there, several total amounts of DNA were tested to find the best DNA concentration. In preparation for electroporation, cells were harvested using the same procedure as used for passaging, washed once in PBS, spun down, and resuspended in PBS. Cells were counted using a hemacytometer. Cell-DNA mixtures were made in Eppendorf tubes and then transferred to electroporation cuvettes with a 2 mm gap (Fisher Scientific, Waltham, MA). Cell-DNA mixtures were incubated on ice for no less than 5 minutes and no more than 15 minutes before electroporation, and they were kept on ice after electroporation for at least 3 minutes before plating. Cell-DNA mixtures were electroporated using a Bio-Rad Gene Pulser with capacitance extender and pulse controller (Bio-Rad, Hercules, CA) using the settings shown in Table 3.1. After electroporation (and previously mentioned short ice

incubation), cells were resuspended in complete media and plated in tissue culture vessels of the appropriate size (Corning, Lowell, MA). Media was changed 24 hours post-transfection and every 24 hours thereafter until experiments were completed.

Sample	Voltage (volts)	Resistance (Ohms)	Capacitance (uF)	Cell Density (cells/uL)	DNA (ug)
A	230	infinite	1010	12500	50
B	270	infinite	1010	12500	50
C	230	0	1010	12500	50
D	270	0	1010	12500	50
E	230	infinite	1010	12500	100
F	270	infinite	1010	12500	100
G	230	0	1010	12500	100
H	270	0	1010	12500	100

Table 3.1: Initial Electroporation parameters

Hybrid FugeneHD + Electroporation Transfections

A modified version of Roche's FugeneHD protocol was used for the hybrid transfection procedure. Instead of 50 µL of diluent per µg of DNA, 25 µL of diluent was used per µg of DNA. To ensure clarity of the procedure, the detailed protocol for transfection of 15 µg of a single construct will be discussed. Assuming the concentration of DNA is 1 µg/µL, the following protocol would be used. First, PBS is added to an eppendorf tube such that the final volume of PBS + DNA + FugeneHD will be 25 uL/ug DNA. In this case that is 330 µL of PBS. Then 15 µg of DNA is added to the PBS, followed by 30 µL of FugeneHD. The tube is closed securely, vortexed, and incubated while cells are harvested. P19s at 80-90% confluency are harvested from a single T-75 flask, washed in PBS, spun down, and resuspended in 900 µL of PBS. 300 µL of the cell solution is then aliquotted into an eppendorf tube and spun down. Once the supernatant has been aspirated from the eppendorf tube of P19s, the FugeneHD complexes should have been incubating for about 15 minutes. Once the FugeneHD complexes have incubated

at room temperature for at least 15 minutes, the P19 cell pellet in the eppendorf tube is resuspended in Fugene complexes and transferred to an electroporation cuvette. Electroporation is then completed as described for protocol D in Table 1.

Calculation of Transfection Efficiency

Transfection efficiency was calculated by determining the percentage of green pixels in each fluorescent image. Images were converted to monochromatic images in Adobe Photoshop. The magic wand tool was set to a tolerance of 5 and used to select the non-fluorescent part of the image. Transfection efficiency was defined as (fluorescent pixels/total pixels)*100.

Fluorescence Imaging

GFP expression was assessed 24 hours post-transfection and every 24 hours thereafter (until experiments concluded) via fluorescence microscopy using a Nikon Eclipse TE2000-S microscope with a FITC HyQ filter (Chroma, Rockingham, VT).

RESULTS AND DISCUSSION

Figure 3.1 shows the transfection efficiencies calculated from P19s 24 hours post transfection, and Figure 3.2 shows the fluorescent images from which those efficiencies were calculated. Transfection efficiency steadily rose as the amount of DNA was increased, up to 4.5 ug, after which point increasing the amount of DNA used resulted in lower transfection efficiency. Overall, transfection efficiencies varied between 1.5% and 24.4% for Fugene6 transfection. Morphologically there is little difference between Fugene6 transfection conditions, but as can be seen in Figure 3.2, cells are small and rounded in all conditions. This indicates that Fugene6 has a negative impact on

morphology and general health of P19s, but it also achieves reasonably efficient transfection.

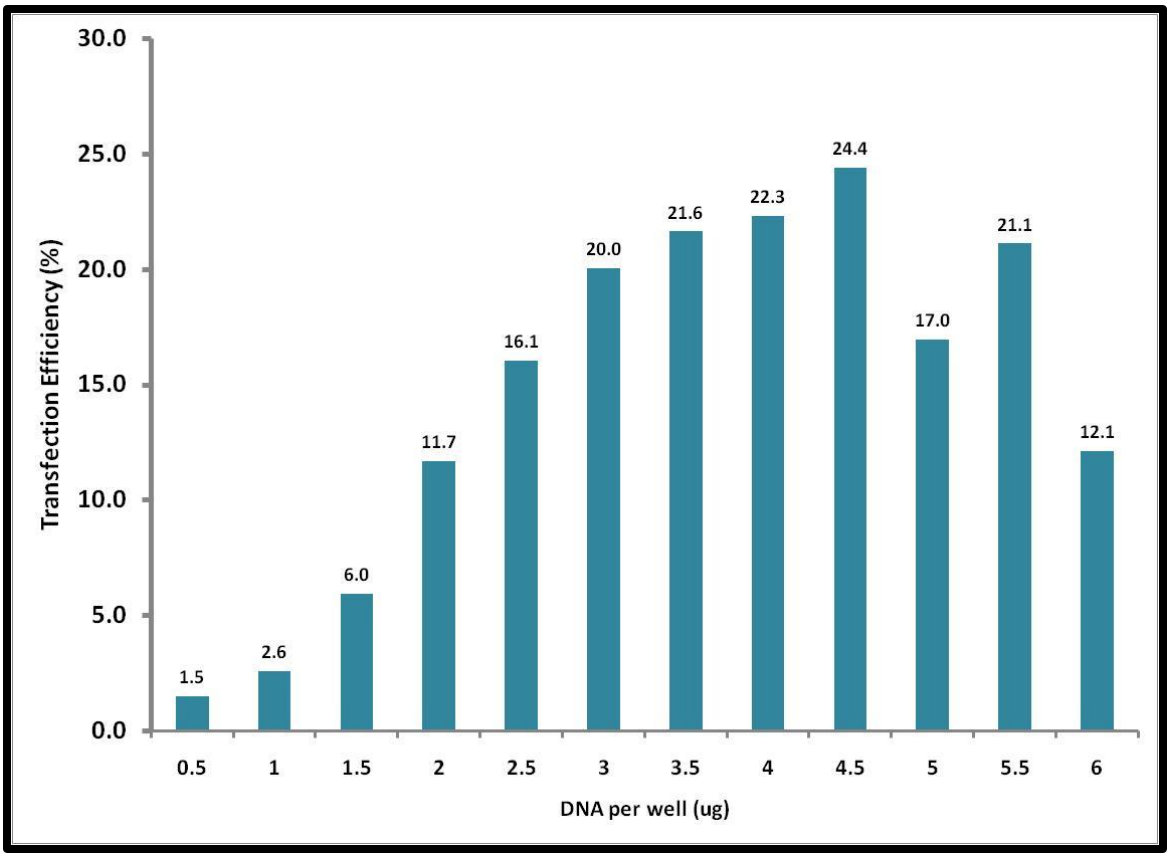


Figure 3.1: Efficiency of Fugene6 transfection of P19 embryonal carcinoma cells, assessed 24 hours post-transfection.

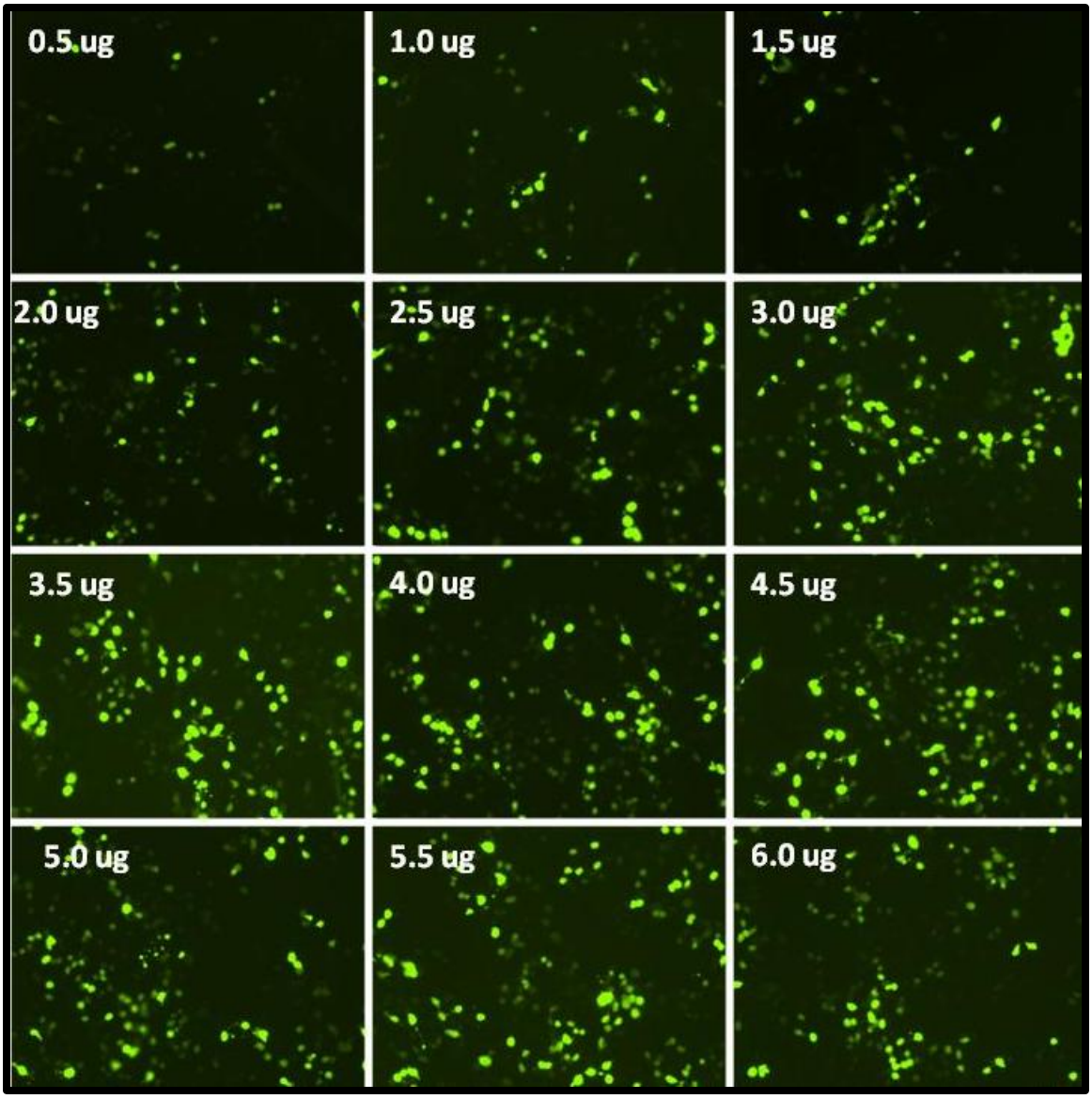


Figure 3.2: P19 embryonal carcinoma cells expressing GFP 24 hours after transfection by Fugene6 using 0.5 – 6.0 µg of pEGFP-N1 plasmid DNA.

P19s transfected with FugeneHD maintained much healthier morphology, as can be seen in Figure 3.3. Figure 2.3also indicates that 5 µg of DNA transfected using FugeneHD results in the highest transfection efficiency achieved by either Fugene6 or FugeneHD. Figure 3.4 further corroborates that data. Interestingly, Figure 3.4, which shows transfection efficiencies calculated from the fluorescent images shown in Figure 3.3 indicates that Fugene6 transfection efficiency for equivalent amounts of DNA outperforms FugeneHD until the amount of DNA surpasses 3 µg, at which point FugeneHD clearly yields higher transfection efficiencies. For FugeneHD, transfection efficiencies ranged from 1.3% to 35.7% across all variations of this technique.

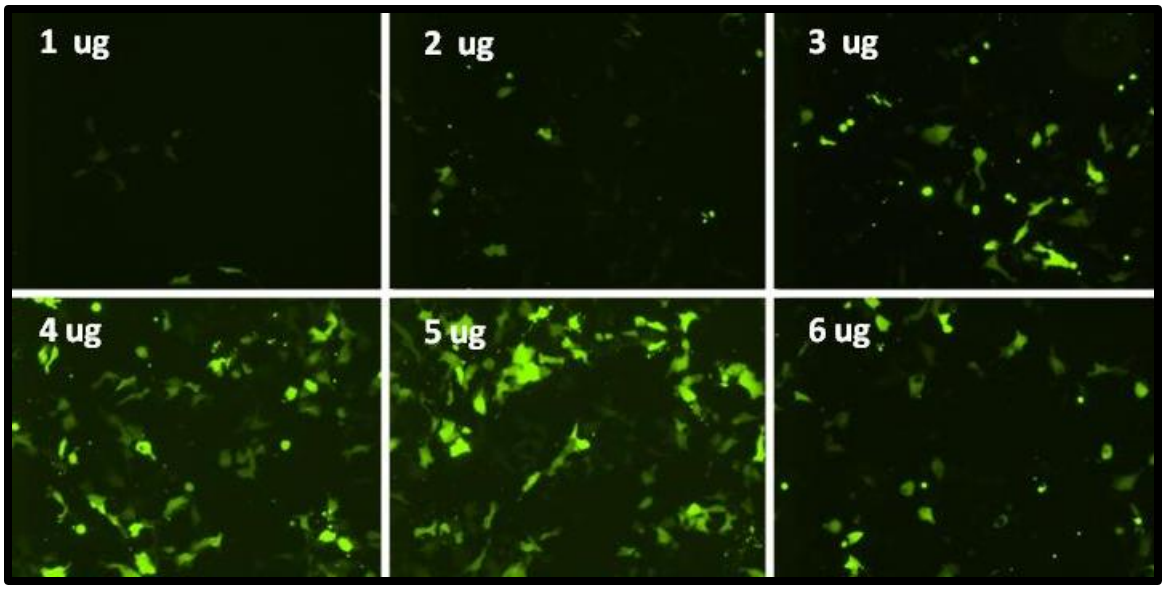


Figure 3.3: P19 embryonal carcinoma cells expressing GFP 24 hours after transfection by FugeneHD using 1 – 6 µg of pEGFP-N1 plasmid DNA.

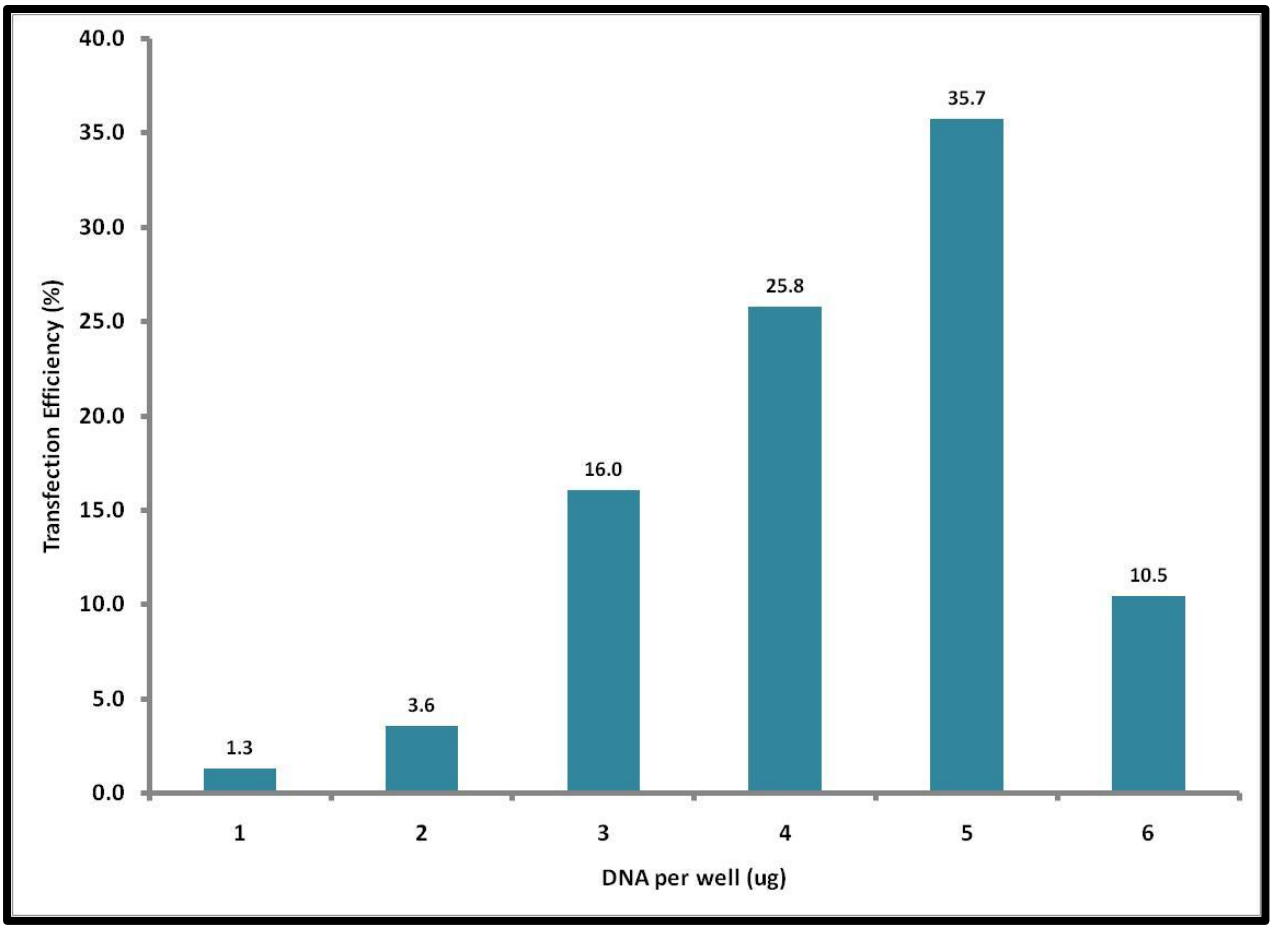


Figure 3.4: Efficiency of FugeneHD transfection of P19 embryonal carcinoma cells, assessed 24 hours post-transfection.

Fluorescent images of cells transfected using the electroporation parameters detailed in Table 1 are shown below in Figure 3.5, and the calculated transfection efficiencies for each of those conditions is shown in Figure 3.6. Overall transfection efficiencies for electroporation with the parameters detailed in Table 3.1 ranged from 6.8% - 74.1%. However, the brightfield images in Figure 3.5 demonstrate that the mortality rate for each of these conditions is dramatically different. The best conditions for cell survival were conditions C and D, but D clearly achieves higher transfection efficiency.

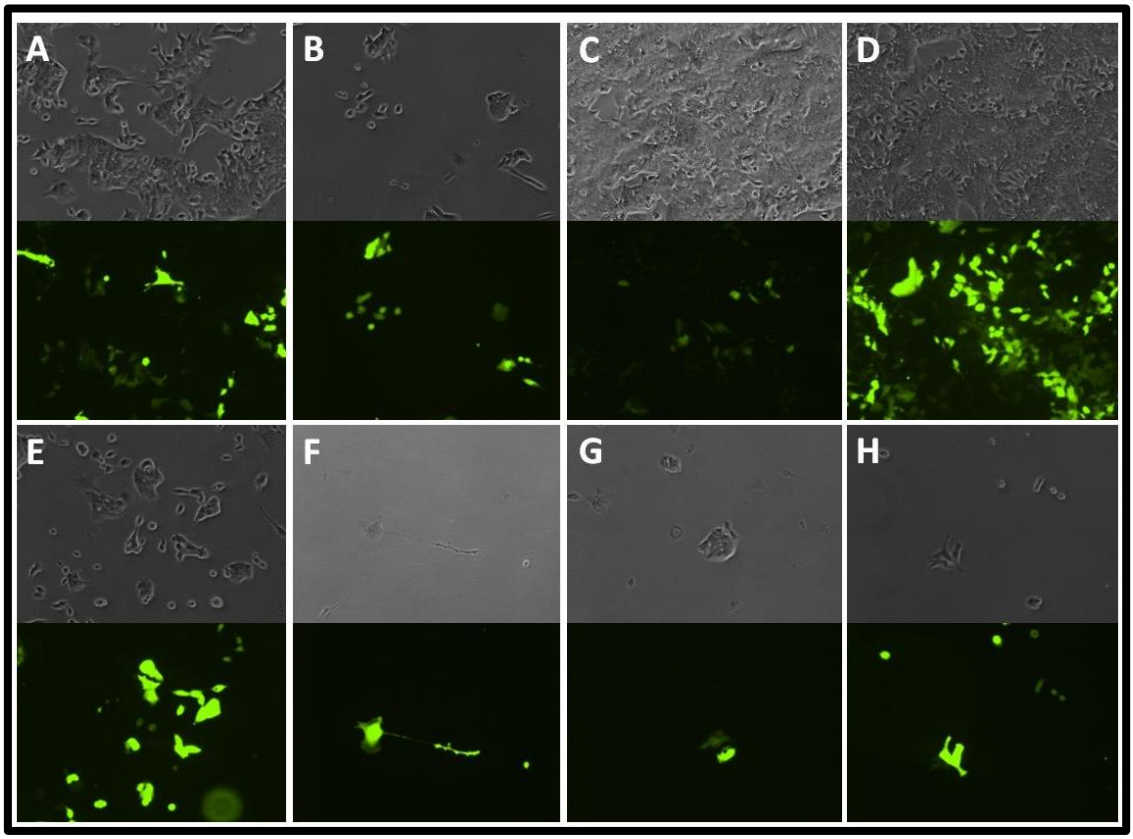


Figure 3.5: P19 embryonal carcinoma cells expressing GFP 24 hours after transfection by electroporation protocols detailed in Table 2.1. Brightfield images are on top of the fluorescent images of each condition. Each condition is clearly labeled with the name of the protocol used (A-D).

Parameter set H did yield the highest transfection efficiency of the initial electroporation experiments, and it corresponds to transfection efficiencies commonly achieved in HEK293Ts (unpublished data). Unfortunately, the mortality rate of this protocol precludes its use as a transfection technique for the purposes of this project, as it would not produce enough live cells to use for protein expression assays. On the other hand, parameter set D achieved 42.4% transfection efficiency and, by visual inspection, excellent cell viability. After completing the Fugene6, FugeneHD, and electroporation experiments A-H, electroporation with parameter set D represented higher transfection efficiency than both Fugene methods and the best combination of transfection

efficiency, cell viability, and impact on morphology of all the electroporation protocols. With that in mind we chose to further explore optimization of protocol D.

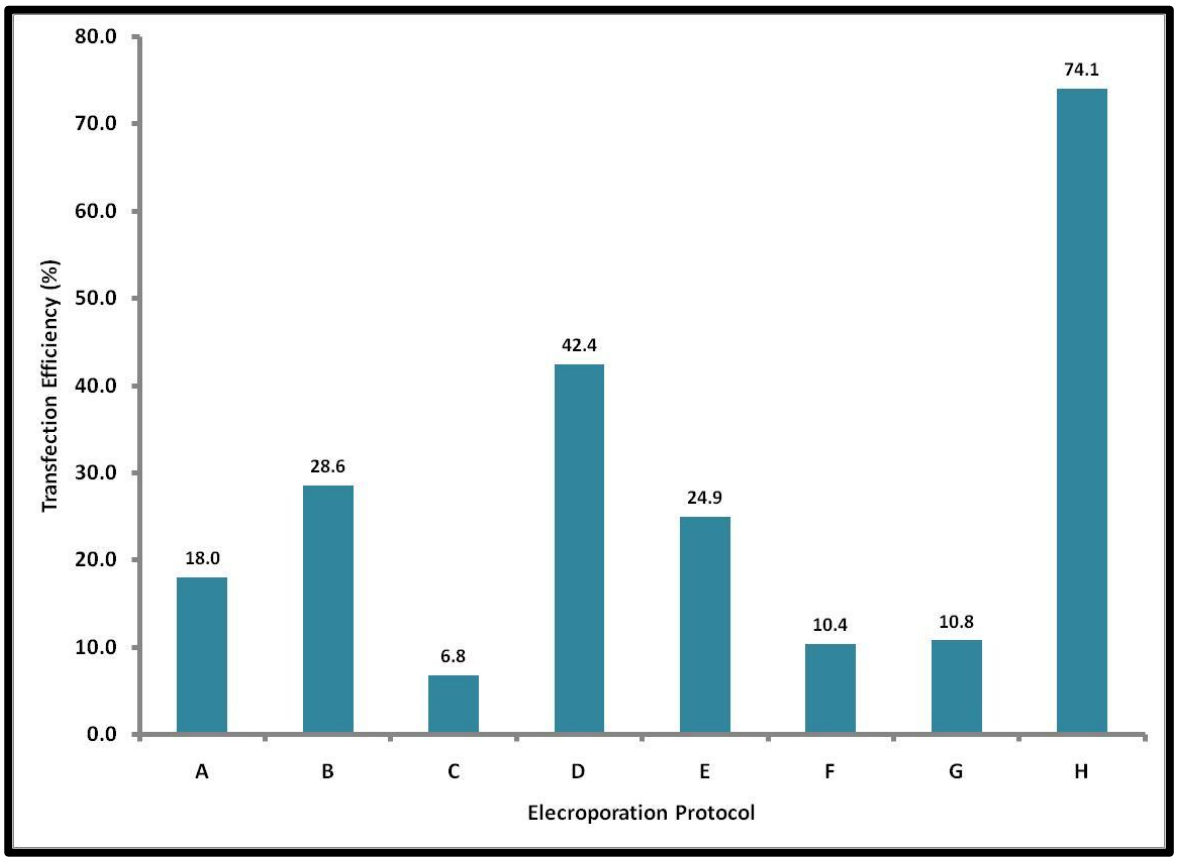


Figure 3.6: Efficiency of electroporation based transfection of P19 embryonal carcinoma cells using parameters detailed in Table 2.1, assessed 24 hours post-transfection.

Figure 3.7 shows P19s electroporated using parameter set D and various amounts of DNA. As can be seen from those images as well as the transfection efficiencies calculate from them, shown in Figure 3.8, increasing the amount of DNA to 100 µg dramatically increased the transfection efficiency without negatively impacting cell viability. The brightfield images for these experiments are not shown, because all conditions were at 100% confluency when the images were taken.

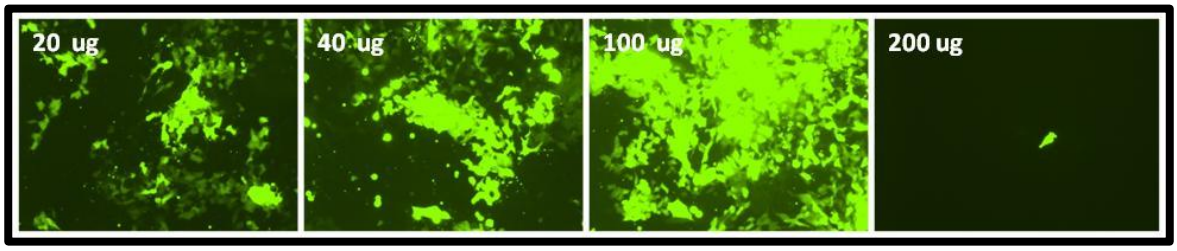


Figure 3.7: P19 embryonal carcinoma cells expressing GFP 24 hours after transfection by electroporation protocol D from Table 2.1 with various amounts of DNA (noted in each frame).

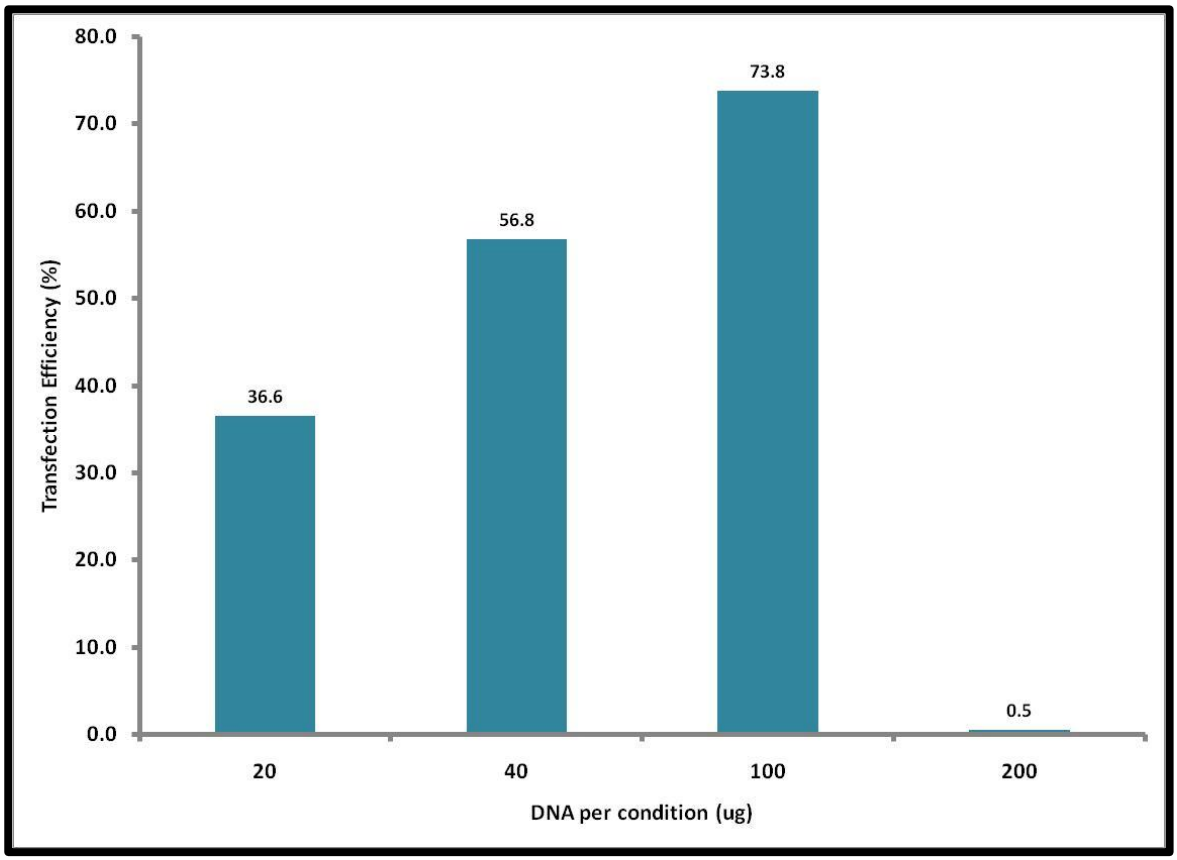


Figure 3.8: Efficiency of electroporation based transfection of P19 embryonal carcinoma cells using parameter set D from Table 3.1 and various amounts of DNA, assessed 24 hours post-transfection.

As shown in Figure 3.8, the 100 µg condition achieved a remarkable 73.8% transfection efficiency. At this point, we considered that we had found the best possible transfection technique for this cell line and began using it to test the UAA incorporation system in P19s. Fortunately we continued running eGFP positive controls for all of those experiments, because we learned that the instrument parameters, cell density, and amount of DNA were not the only factors in the success of this technique. Another critical factor is the amount of time between when cells are living happily in culture to when they are electroporated, and the amount of time between electroporation and re-plating is also critical. Several representative positive controls are shown in Figure 3.9. While analyzing what could be the underlying causes of the stark differences between transfection efficiencies obtained during protocol development (Figure 3.7) and those obtained while attempting to implement the UAA incorporations system, it became clear that the only real difference was the number of conditions in each experiment. Experiments done in developing the protocol involved, at the most, nine conditions whose sole difference was the amount of a single DNA construct used, while experiments done with the UAA incorporation system involved 8 conditions at a minimum, and each of those conditions used different amounts of DNA and different numbers of at least four different constructs. The protocol itself was fine, but the amount of time between passaging cells, electroporating, and replating them was exerting a strong negative influence on transfection efficiency. To test this theory, cells were harvested from two flasks and were not counted. DNA was pre-aliquotted into eppendorf tubes, ready to receive cells as soon as they were passaged, washed, and resuspended in PBS. Each T-75 flask was used for three different conditions. Thus the amount of time necessary for the whole electroporation process was dramatically

reduced. As can be seen in Figure 3.10, this change brought transfection efficiencies back into the range expected based upon our optimization experiments.

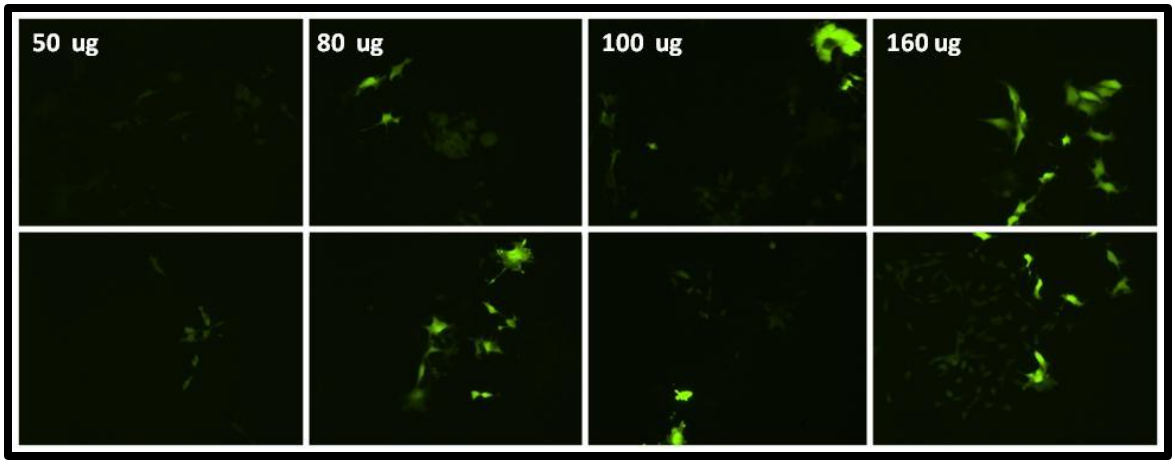


Figure 3.9: P19 embryonal carcinoma cells expressing GFP 24 hours after transfection by electroporation protocol D from Table 3.1 with various amounts of DNA (noted in each frame). Each image is from a separate experiment.

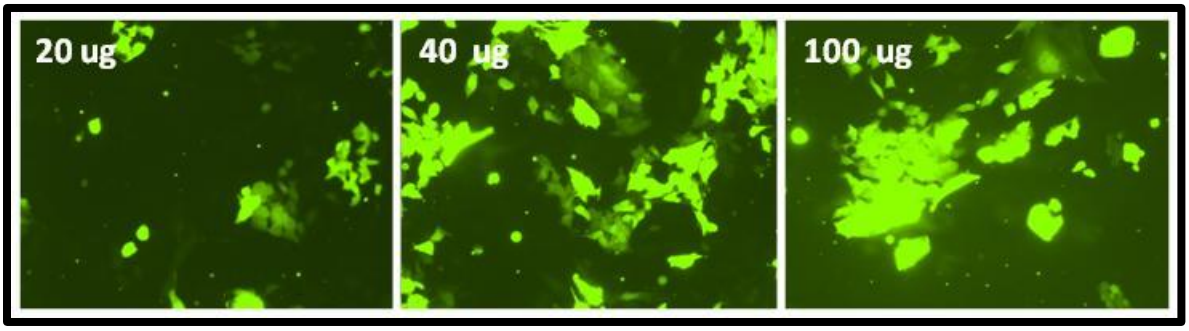


Figure 3.10: P19 embryonal carcinoma cells expressing GFP 24 hours after modified transfection by electroporation protocol D from Table 3.1 with various amounts of DNA (noted in each frame).

Though the alterations made to the electroporation protocol yielded transfection efficiencies high enough to use for UAA incorporation efficiency, they used milligrams of DNA for each experiment, which became problematic toward the end of this project. In an effort to finish the necessary experiments with a limited amount of DNA, we tried to recreate the phenomenon of Nucleofection. Presuming that the name was a combination of the words nucleus and transfection, we posited that Nucleofection had somehow achieved direct delivery to the nucleus via electroporation. As far as we know both Fugene methods and electroporation are simply methods of getting DNA past the lipid bilayer and not necessarily into the nucleus. However, if a liposomal based reagent could be delivered into the cytosol whole, then it may be able to cross the nuclear membrane and deliver DNA directly to the nucleus. Thus we chose to combine electroporation with FugeneHD. It is impossible at this stage to comment with any certainty on whether DNA actually was transferred directly to the nucleus or if the complexes were delivered to the cytosol whole. However, Figure 3.11 shows fluorescence microscopy images of cells transfected with electroporation + FugeneHD, and it is clear from these images that they achieve transfection efficiencies as high as the previous electroporation alone protocol. In contrast to electroporation alone, this protocol uses approximately $1/4^{th}$ the amount of DNA per condition and resulted in far more stable electroporation results. In all of the electroporation experiments, it became clear that the time constant τ for each electroporation event was indicative of the success of transfection as well as mortality rate. FugeneHD in combination with electroporation consistently resulted in time constants below 30, which was indicative of consistently successful transfection and low cell death. Prior to that, each experiment

was done with an n of at least three to ensure that one of the three repeats would result in a time constant less than 30.

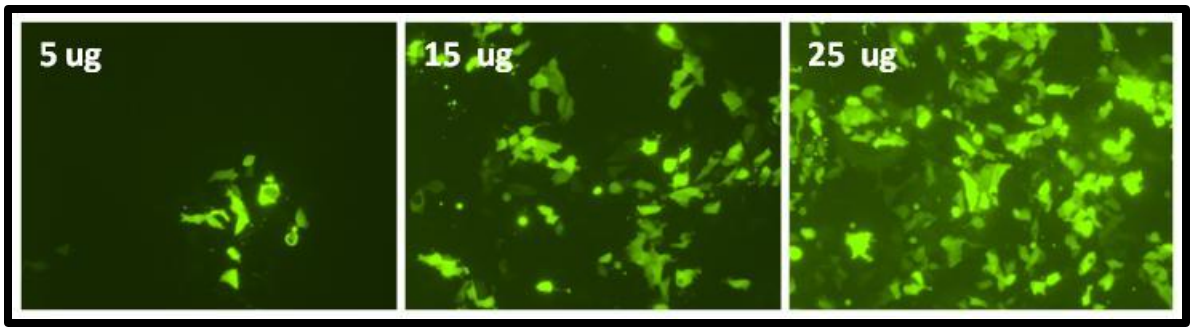


Figure 3.11: P19 embryonal carcinoma cells expressing GFP 24 hours after transfection via a hybrid method of FugeneHD + electroporation protocol D from Table 3.1.

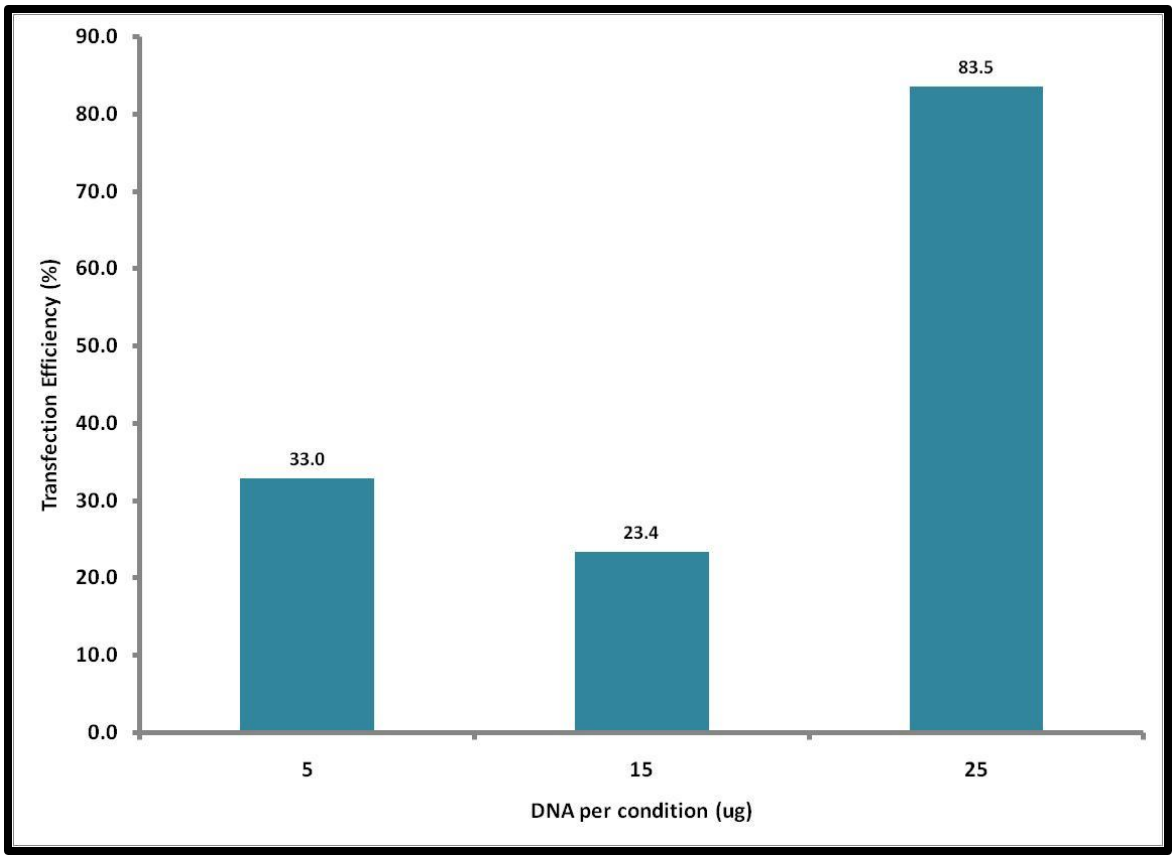


Figure 3.12: Efficiency of hybrid FugeneHD + electroporation method of transfection of P19 embryonal carcinoma cells.

As is evidenced by the transfection efficiencies in Figure 3.12, combining FugeneHD with electroporation yielded higher transfection efficiencies than electroporation alone. In comparing electroporation with FugeneHD or Fugene6 alone it is important to keep in mind the amount of cells used in each case. Each electroporation condition uses approximately 75 cm^2 of cells at 85% confluency while FugeneHD and Fugene6 use 9.6cm^2 of cells at 60% confluency. A simple calculation indicates that each electroporation condition transfects approximately 11 times the number of cells that a FugeneHD or Fugene6 condition transfects. Meaning that the best FugeneHD adherent cell transfection method, which used 5 µg to transfect a single well in a 6-well plate, would require 55 µg of DNA to transfect the number of cells in an electroporation condition but would only transfer DNA to 35% of those cells. In contrast to that, FugeneHD + electroporation used only 25 µg to transfer DNA to 83.5% of the same amount of cells. Thus, FugeneHD + electroporation is the most efficient means of DNA transfer for large quantities of cells. Since UAA incorporation systems usually achieve about 20% of the expression level of wildtype proteins (10-12), a large number of cells is necessary to obtain enough protein for western blot, UAA incorporation assays, and mass spectrometry. Therefore the hybrid FugeneHD + electroporation method is the most appropriate method to use for UAA incorporation systems in stem cells.

Each transfection method tested could be optimized to achieve transfection of at least 24% of the cells in a condition. While this level of transfection may be high enough for some applications, it is too low for something like site-specific UAA incorporation. Combining the efficacy of FugeneHD with that of electroporation yielded a hybrid transfection method capable of transfection efficiencies on par or in excess of other methods tested while using a fraction of the DNA. In experiments where the amount of DNA is not an issue, it may be more useful to use straight electroporation, like protocol

D from Table 3.1. This method of transfection is capable of high efficiency DNA transfer, but requires milligram levels of DNA and is not the most efficient in terms of transfection efficiency per μg of DNA. Since parts of the work presented in this dissertation were done in a research environment in which DNA was unlimited, this protocol was used for some experiments. However, when DNA was available in limited quantities, the hybrid FugeneHD + electroporation technique was used to achieve acceptable transfection efficiencies with limited amounts of DNA.

REFERENCES

1. Macchiarini P, Jungebluth P, Go T, Asnaghi MA, Rees LE, Cogan TA, et al. Clinical transplantation of a tissue-engineered airway. Lancet. 2008;372(9655):2023-30.
2. Alhadlaq A, Elisseeff JH, Hong L, Williams CG, Caplan AI, Sharma B, et al. Adult stem cell driven genesis of human-shaped articular condyle. Ann Biomed Eng. 2004;32(7):911-23.
3. Carletti B, Grimaldi P, Magrassi L, Rossi F. Specification of cerebellar progenitors after heterotopic-heterochronic transplantation to the embryonic CNS in vivo and in vitro. J Neurosci. 2002;22(16):7132-46.
4. Sotelo C, Alvarado-Mallart RM. The reconstruction of cerebellar circuits. Trends Neurosci. 1991;14(8):350-5.
5. Amariglio N, Hirshberg A, Scheithauer BW, Cohen Y, Loewenthal R, Trakhtenbrot L, et al. Donor-derived brain tumor following neural stem cell transplantation in an ataxia telangiectasia patient. PLoS Med. 2009;6(2):e1000029.
6. McBurney MW. P19 embryonal carcinoma cells. Int J Dev Biol. 1993;37(1):135-40.
7. Schmidt JW, Brugge JS, Nelson WJ. pp60src tyrosine kinase modulates P19 embryonal carcinoma cell fate by inhibiting neuronal but not epithelial differentiation. J Cell Biol. 1992;116(4):1019-33.
8. Tan Y, Xie Z, Ding M, Wang Z, Yu Q, Meng L, et al. Increased levels of FoxA1 transcription factor in pluripotent P19 embryonal carcinoma cells stimulate neural differentiation. Stem Cells Dev. 2010;19(9):1365-74.
9. AG LC. Amaxa (R) Cell Line Nucleofector Kit V.
10. Thibodeaux GN, Cowmeadow R, Umeda A, Zhang Z. A tetracycline repressor-based mammalian two-hybrid system to detect protein-protein interactions in vivo. Anal Biochem. 2009;386(1):129-31.
11. Umeda A, Thibodeaux G, Moncivais K, Jiang F, Zhang Z. A versatile approach to transform low-affinity peptides into protein probes with cotranslationally expressed chemical cross-linker. ANALYTICAL BIOCHEMISTRY. 2010;405(1):82-8.
12. Zhang Z, Alfonta L, Tian F, Bursulaya B, Uryu S, King DS, et al. Selective incorporation of 5-hydroxytryptophan into proteins in mammalian cells. Proc Natl Acad Sci U S A. 2004;101(24):8882-7.

CHAPTER FOUR

An Unnatural Amino Acid Incorporation System for Mammalian Cells

INTRODUCTION

Chapter Two briefly discussed tRNA and aminoacyl-tRNA-synthetase (RS) specificity. In the following pages, the background information pertinent to engineering a tRNA/RS pair will be discussed in detail, beginning with the basic concepts of the genetic code and protein translation. With that information in mind, the application of those concepts to the creation of an unnatural amino acid (UAA) incorporation system for use in mammalian cells and then stem cells will be discussed.

The Genetic Code and Protein Synthesis

All genetic information is represented at the molecular level by DNA, but that information must be translated before it presents itself as proteins, cells, organisms, physiological phenomena, or behavioral modifications (1). Therefore all biological processes are dependent upon the fidelity of the genetic code as it is represented at the level of DNA. Since the focus of this work is on protein-protein interactions (PPIs), the most important part of this process is the basic idea of DNA translation to protein. This process is a multiple-step process, each of which has several mechanisms for maintaining the faithful representation of DNA at the protein level.

Each of the twenty canonical amino acids is represented at the DNA level by a sequence of three nucleotide bases, called a codon. Each codon only represents one amino acid, but each amino acid can be represented by several codons; there are codons for both STOP and START translation signals. The first step in the process of converting DNA to protein is transcription. In this process, RNA polymerase ‘reads’ DNA

in the nucleus and creates a strand of messenger RNA (mRNA) representative of the DNA. In some eukaryotic cells, there can be post-transcriptional modifications like end-capping and intron splicing which occur before mRNA is transcribed to protein. In any case, when the final mRNA product is produced, it can then be transcribed by a ribosome in the cytosol.

As previously mentioned, each amino acid is coded for by a number of codons at the DNA level, so a ribosome must have some method of reading those codons and converting them from nucleic acids to the proper amino acid. The bridge between those two points in the process of protein translation is aminoacylated tRNA. Each tRNA has a region called the anticodon which bears the complementary sequence to its intended amino acid's codon. In this way, a tRNA bearing an amino acid can come into the ribosome, base pair with the correct codon, and then allow attachment of the correct amino acid to the growing peptide chain. The process of aminoacylated tRNA base pairing with mRNA as it is read by the ribosome and amino acids being added to proteins in the correct sequence is depicted in Figure 4.1.

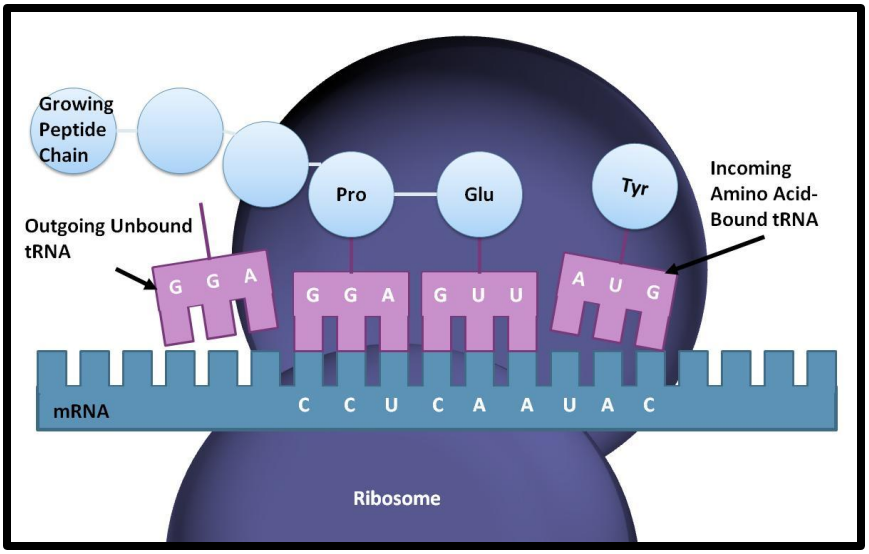


Figure 4.1: Translation of mRNA to protein within the ribosome.

Figure 4.1 includes the protein translation concepts most pertinent to UAA incorporation systems, but it is missing one very important part of the concept: charging of tRNA with the correct amino acid. Ribosomes do not actually discriminate between codon and amino acid; they do not have a proofreading mechanism based upon the mRNA codon and amino acid match. Instead they depend upon the tRNA anticodon base pairing with its cognate mRNA codon and being preattached to the appropriate amino acid. This means that maintenance of the genetic code is achieved prior to the ribosome's role in protein translation at the point of tRNA aminoacylation. Therefore aminoacyl-tRNA synthetases (RSs), which conjugate amino acids to tRNAs, are responsible for maintaining the fidelity of the genetic code. RSs catalyze amino acid attachment to tRNAs (aminoacylation) in a two step reaction, shown in Figure 4.2. If an amino acid is attached to a tRNA bearing an anticodon intended for a different amino acid, then the protein for which this tRNA is used will not be produced as it is coded at the DNA level. It is this very property of protein translation and maintenance of the genetic code which the UAA incorporation system presented in this dissertation exploits.

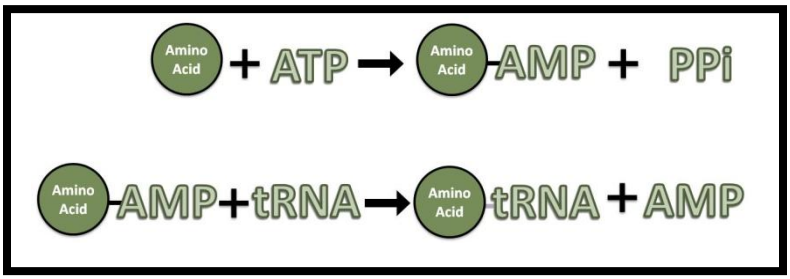


Figure 4.2: Two step aminoacylation of tRNA by RS

The UAA incorporation system developed and implemented in this work intervenes at the level of tRNA aminoacylation to reprogram the genetic code such that

it includes the desired UAA(s). In light of the facts in the preceding paragraph, it must be logical that RSs have mechanisms by which they distinguish correct tRNAs and amino acids from incorrect tRNAs and amino acids, ensuring that they aminoacylate only their intended tRNAs with the correct amino acid. As has been demonstrated previously by a number of groups, RSs recognize sequences and structures of tRNA molecules (2-4) as well as structures and functional groups of amino acids (5, 6). Therefore, RSs can be engineered to recognize mutant tRNAs and UAAs and thus can supplement the genetic code of cells or organisms to which they are added. The sequence and structural motifs by which RSs recognize correct tRNAs will be discussed in detail in the following section.

tRNA Identity Determinants

The genetic code is enforced by RS discrimination of correct tRNAs and amino acids, and correct tRNAs bear anticodons complementary to the codon by which their cognate amino acid is represented. Therefore the anticodon is an obvious candidate for a tRNA identity determinant. The first experimental evidence for the role of the anticodon in tRNA identity was presented in 1964 when chemical modification of the anticodon region of yeast tRNA abolished aminoacylation (7). This region has since been confirmed as a strong identity determinant for most tRNAs, but it is not the only one (8). Since amino acids are covalently attached to the 3' terminal CCA on tRNA molecules, the nucleotides in that region were also strong candidates for identity determinants as they must physically interact in some way with the RS in order for the proper amino acid to be attached. In the 1970s, several groups used a number of experimental techniques to demonstrate that nucleotides in the tRNA acceptor stem region (See Figure 4.3 for acceptor stem location) do play a role in tRNA recognition by RSs and therefore can be classified as identity determinants (9-11).

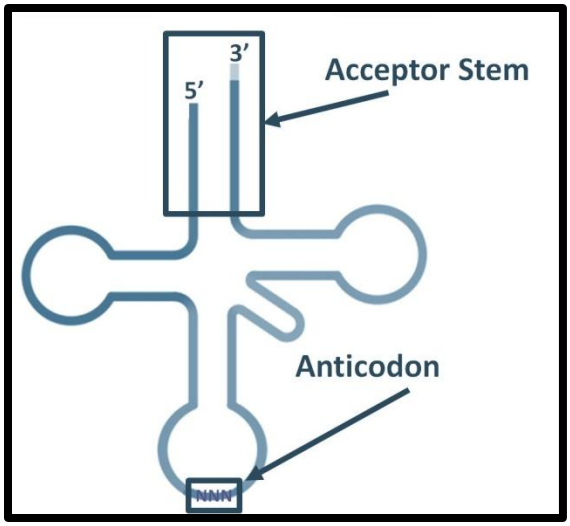


Figure 4.3: tRNA secondary structure

Despite the fact that anticodon mutations can exert some effect on tRNA identity, they are not the most potent determinants of tRNA identity (5, 12). Instead, the nucleotides in the acceptor stem are stronger identity determinants as compared to those of the anticodon (13). Incidentally, this fact is instrumental in making UAA incorporation systems possible, because site-specific UAA incorporation requires the use of anticodon mutated tRNAs. If the anticodon were a very strong determinant of tRNA identity, this would present a problem. Himeno and colleagues used a series of tRNA mutants containing changes to the anticodon, the acceptor stem, and the variable loop of *Escherichia coli (E. coli)* $tRNA^{ser}$ to determine which changes would allow charging of $tRNA^{ser}$ by a tyrosyl tRNA-synthetase (5). They found, as previously, that the anticodon did play a role in recognition, but for this particular tRNA-RS pair, other nucleotides were far stronger determinants of identity than the anticodon. For clarity's sake it should be noted that tRNAs are numbered from 5' to 3'. So if a tRNA has a nucleotide G3, that is the third nucleotide from the 5' end, and it is a guanosine. Himeno and colleagues reported a strong effect on identity stemming from the so called discriminator base, located at position 73, directly 5' of the terminal CCA (5). Anticodon mutation did not allow acylation of mutant tRNA by a tyrosyl RS, but anticodon mutation combined with

discriminator base mutation and the addition of a $tRNA^{ser}$ sequence in the variable loop of the mutant tRNA did allow charging of the mutant $tRNA^{ser}$ by a tyrosyl-RS (5).

Further corroborating the idea that acceptor stem nucleotides are strong determinants of tRNA identity, the Schimmel group has done a variety of experiments using RNA fragments to demonstrate that acceptor stem nucleotides alone can allow aminoacylation of RNA molecules (14-18). Perhaps the most important implication of this work is the idea that, without the effect of the anticodon, aminoacylation can be achieved by manipulation just the acceptor stem nucleotides. While there are effects from other nucleotides, like the previously mentioned instance in which the variable loop nucleotides played a role in tRNA identity, it is possible to create RNA or tRNA molecules which are easily usable substrates for RS aminoacylation by manipulation of only the acceptor stem nucleotides. This allows mutant amber suppressor tRNAs to be tailored to the specificity either of a canonical amino acid or UAA by choosing appropriate acceptor stem sequences.

Since the project detailed in this dissertation is specifically designed for use in mammalian cells, the tRNA and RS used should be non-reactive with mammalian tRNAs and RSs. Therefore, it is important to know the specific identifiers commonly found in mammalian/eukaryotic tRNAs versus prokaryotic tRNAs. As mentioned in Chapter Two, Wakasugi et al demonstrated that, much like the $tRNA^{ser}$ and $tRNA^{ala}$ mentioned in the preceding paragraphs, the acceptor stem nucleotides are powerful identity determinants in human tRNAs (19). *E. coli* $tRNA^{tyr}$ ($tRNA^{ectyr}$) has acceptor stem nucleotides G1:C72, while human $tRNA^{tyr}$ ($tRNA^{hstyr}$) has acceptor stem nucleotides C1:G72 (19). *Methanocaldococcus jannaschii (M. jannaschii)* has the same 1:72 base pair as $tRNA^{hstyr}$. In both human and *M. jannaschii* based experiments, mutations which switched the identity of the 1:72 base pair were able to abolish aminoacylation by

wildtype enzymes (19, 20). It is this experimental evidence on which the orthogonality of tRNA used in the UAA incorporation presented in this dissertation is based.

OVERALL EXPERIMENTAL DESIGN

The experiments detailed in this chapter pertain only to the initial design and creation of the system, while Chapter Five and Chapter Six cover its implementation in P19 embryonal carcinoma cells. The design, optimization and implementation of a UAA incorporation system for use in mammalian cells is a multifaceted project. As such, it was broken down into several parts, some of which were completed in parallel. Those parts are described briefly below, and the experimental details of each will follow.

1. Reporter Plasmid Optimization

   Since previous experiments using HEK293T cells and the TAG reporter protein received from the Schultz group indicated background expression of green fluorescent protein in the absence of either tRNA or RS constructs, it was necessary to create a new reporter plasmid.

2. tRNA and RS Creation and Optimization

   Since previously evolved RSs with specificities for a number of useful UAAs are based upon the *M. jannaschii* tRNA/RS pair, modifications were made to that pair to transfer it to mammalian cells. The tRNA was made orthogonal to mammalian systems and given a new promoter to increase tRNA levels, and the CP1 region of the *M. jannaschii* RS^{Tyr} was replaced with that of *E. coli*. The D286R mutation was also introduced to increase aminoacylation efficiency of the amber suppressor tRNA.

3. tRNA Acceptor Stem Optimization

   Since nucleotides beyond simply the discriminator base and the 1:72 pair

have been implicated in RS recognition of various other tRNAs from multiple species, a tRNA database was used to find an acceptor stem that would be completely orthogonal to the human system. This was done to eliminate the background amber suppression seen in previous versions of this system by making the amber suppressor tRNA completely orthogonal to the mammalian system.

4. Condensation of UAA incorporation System to a 2-Plasmid System

   Since transfection efficiencies in P19s are expected to be lower than those achieved in other cell lines, it could be helpful to put the necessary DNA constructs into two bidirectional plasmids instead of three one-directional plasmids. To that end, the genes encoding the reporter protein, the tRNA, and the RS were transferred to bidirectional plasmids.

Completion of these four goals will result in successful stop codon suppression, which is not the same as UAA incorporation. However, stop codon suppression in this case is the first step toward UAA incorporation, and indicates feasibility of UAA incorporation. Chapter Six will discuss the details involved in progressing from stop codon suppression to actual incorporation of UAAs.

MATERIALS AND METHODS

New Reporter Protein Plasmid Construction

The plasmid p-EGFPN1 was purchased from Clontech. The 40th codon was mutated from TAC to TAG using site-directed mutagenesis (SDM). Forward and reverse primers were designed using Stratagene's online primer design program, and sequences are below. All DNA sequences are listed in the normal convention of 5'-3'.

Forward Primer: GCGAGGGCGAGGGCGATTAGACCTACGGCAAGC

Reverse Primer: GCTTGCCGTAGGTCTAATCGCCCTCGCCCTCGC

The QuickChange II Site-Directed Mutagenesis Kit was purchased from Stratagene (Stratagene, Agilent Technologies, Inc., Santa Clara, CA), and SDM was completed according to the provided protocol. Plasmids were sequenced by the University of Texas at Austin ICMB DNA Sequencing Facility. The new plasmid was termed 40TAG-peGFPN1.

Routine Cell Culture

HEK293T cells were purchased from Invitrogen (Carlsbad, CA). Complete HEK293T media was composed of high glucose DMEM with 4500 mg/L of sodium bicarbonate and L-glutamine (Sigma-Aldrich, St. Louis, MO) supplemented with 10% (v/v) fetal bovine serum (FBS) (Atlanta Biologicals, Lawrenceville, GA) and 1X nonessential amino acids (Sigma-Aldrich, St. Louis, MO). Cells were maintained in a humidified incubator at 37° Celsius with 5% CO_2 atmosphere. Cells were passaged every 2-3 days at ratios between 1:10 and 1:20. Passaging was accomplished by washing cells with phosphate buffered saline (PBS) (Sigma-Aldrich, St. Louis, MO), treating with 0.5% trypsin (Sigma-Aldrich, St. Louis, MO) for 2-5 minutes, and harvesting via centrifugation. In between passaging, fresh media was added to cells every 24 hours.

Testing of New Reporter Protein Plasmid

After sequencing confirmed successful creation of 40TAG-peGFPN1, it was tested by introduction to mammalian cells and fluorescence assisted cell sorting (FACS). HEK293T cells were grown in 6-well plates to between 70% and 90% confluency for each experiment. On the day of transfection, DNA for both 40TAG-peGFPN1 and the previously used TAG reporter protein encoding plasmid, Lital40TAG, was introduced to

HEK293T cells with FugeneHD (Roche, Switzerland). FugeneHD was used according to the manufacturer's protocol using a 2:1 ratio of µL FugeneHD to µg DNA. Complexes were formed in Opti-Mem (Gibco, Invitrogen, Carlsbad, CA) and added dropwise to wells.

24 hours after transfection, cells were imaged using a Nikon Eclipse TE2000-S microscope with a FITC HyQ filter (Chroma, Rockingham, VT). Cells were then harvested for FACS. Briefly, the passaging protocol was followed up to the point of centrifugation. After centrifugation, cells were washed once in PBS and recentrifuged. PBS supernatant was aspirated, and cells were resuspended in 4% paraformaldehyde in PBS (USB Corp. Cleveland, OH). Cells were incubated in paraformaldehyde for 15 minutes, then re-centrifuged. Paraformaldehyde was aspirated, and cells were resuspeded in PBS then subjected to FACS. FACS was accomplished using a BD FacsCalibur, and data analysis was done using Cyflogic. For all FACS experiments, dead cells were excluded using forward and side scatter. GFP gating was done by designating the area of the GFP histogram occupied by untransfected cells as "GFP Negative," and all cells exhibiting higher fluorescence than untransfected cells were classified as "GFP Positive."

Creation of Orthogonal tRNA from *M. jannaschii* tRNA

All restriction enzymes were purchased from New England Biolabs (Ipswich, MA), and all oligonucleotides were purchased from Sigma-Aldrich (St. Louis, MO) unless otherwise indicated. The orthogonal tRNA for *M. jannaschii* was created by ordering oligonucleotides which included all the desired mutations, including the G1:C72 base pair and the amber anticodon. Those primers are shown below.

Forward: GAAGATCTCCGGCGGTAGTTCAGCCTGGTAGAACGGCGGACTCTAA

ATCCGCATGTCGCTGGTTCAAATCCGGCCCGCCCGAGACAAGTGCG

GTTTTTTT

Reverse: CCAATGCATTGGTTGCCCGCTCGAGTAGAAAAAAACCGCACTTGTC

TCCGGCGGGCCGGATTTGAACCAGCGACATGCGGATTTAGAGTCCG

CCGTTCTA

The above oligonucleotides were boiled and allowed to come to room temperature unaided. They were then extended using the Klenow fragment of *E. coli* DNA polymerase (New England Biolabs, Ipswich, MA). This DNA fragment was double-digested with BglII and PstI and cleaned up using a Qiagen PCR Cleanup Kit (Valencia, CA). The plasmid pTRE-Tight was purchased from Clontech (Mountainview, CA), double digested with BglII and PstI then treated with calf-intestinal-phosphatase (CIP) (New England Biolabs, Ipswich, MA). The resultant fragments were separated on a 1% agarose gel. The larger fragment was purified out of the gel using a Qiagen Gel Extraction Kit (Valencia, CA). The annealed and digested tRNA DNA was then ligated into the pTRE-Tight fragment using T4 DNA Ligase to afford the plasmid pTRE-Tight-MJtRNATyr.

Since pTRE-Tight-MJtRNATyr has no promoter for tRNA synthesis, the human H1 promoter was inserted 5’ of the tRNA. The human H1 promoter was constructed in a similar manner to the tRNA DNA fragment. The forward and reverse primers for the human H1 promoter are below.

Forward: CAACCCGCTCCAAGGAATCGCGGGCCCAGTGTCACTAGGCGGGAAC

ACCCAGCGCGCGTGCGCCCTGGCAGGAAGATGGCTGTGAGGGACAG

GGGAGTGGCGCCCTGCAA

Reverse: GAACTTATAAGATTCCCAAATCCAAAGACATTTCACGTTTATGGTGA

TTTCCCAGAACACATAGCGACATGCAAATATTGCAGGGCGCCACTC

CCCTGTCCCTCACAGCC

The above oligonucleotides were annealed as explained for the tRNA DNA above. The fragment was then amplified by polymerase chain reaction (PCR) using the primers below:

Forward: GGAATTCCAATTCGAACGCTGACGTCATCAACCCGCTCCAAGGAA

TC

Reverse: GAAGATCTGTGGTCTCATACAGAACTTATAAGATTCCCA

The resultant DNA fragment was digested with KpnI and BglII while pTRE-Tight-MJtRNATyr was simultaneously digested with the same enzymes. Digested pTRE-Tight-MJtRNATyr was treated with CIP, and then subjected to agarose gel electrophoresis, as mentioned above. The larger fragment was then gel purified, and the human H1 promoter fragment was purified as well, both as previously mentioned for construction of the original pTRE-Tight-MJtRNATyr. The human H1 promoter fragment was ligated into the purified backbone of pTRE-Tight-MJtRNATyr to afford the plasmid 312tRNA which includes the human H1 promoter upstream of the newly orthogonal amber suppressor tRNA based upon *M. jannaschii* $tRNA^{tyr}$.

The same process was used to create a plasmid encoding the wildtype *M. jannaschii* $tRNA^{tyr}$, which was used in multiple experiments as a pseudo positive control as well as an earmark of the level at which tRNA is produced in mammalian cells. This was done using oligonucleotides encoding the wildtype acceptor stem and the amber anticodon and afforded the plasmid wtMJtRNA. The correct sequence of this plasmid was confirmed by sequencing (UT ICMB DNA Sequencing Facility, TX).

Orthogonality Testing of tRNA Constructs

HEK293T cells were grown to 70-80% confluency in 6-well plates (Corning, Lowell, MA) and then cotransfected 40TAG-peGFPN1 and each of the various tRNA

constructs mentioned previously (i.e. each tested tRNA was transfected along with 40TAG-peGFPN1, not with other tRNAs). FugeneHD was used for transfection, and complexes were formed in Opti-Mem using a ratio of 2 µL FugeneHD to 1 µg DNA. Media was changed every 24 hours after transfection. Cells were imaged using a Nikon Eclipse TE2000-S microscope with a FITC HyQ filter (Chroma, Rockingham, VT) every 24 hours after transfection. At 48 hours post-transfection, cells were harvested for FACS. Briefly, the passaging protocol was followed up to the point of centrifugation. After centrifugation, cells were washed once in PBS and recentrifuged. PBS supernatant was aspirated, and cells were resuspended in 4% paraformaldehyde in PBS (USB Corp. Cleveland, OH). Cells were incubated in paraformaldehyde for 15 minutes, then re-centrifuged. Paraformaldehyde was aspirated, and cells were resuspeded in PBS then subjected to FACS. FACS was accomplished using a BD FacsCalibur, and data analysis was done using Cyflogic.

Creation of RS for Charging Orthogonal tRNA with Tyrosine

The wildtype *M. jannaschii* tyrosyl-tRNA synthetase (RS^{mjwt}) was created by introducing the D286R mutation mentioned previously and swapping the *E. coli* CP1 region for that of *M. jannaschii (21)* to afford the EME chimeric protein shown in .

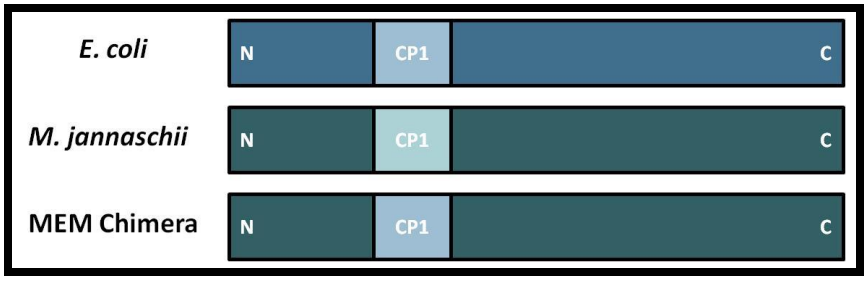


Figure 4.4: Schematic of chimeric RS construction depicting orthogonal RS with N and C termini from *M. jannaschii* and CP1 region from *E. coli*

Briefly, the gene for RS^{mjwt} was amplified via PCR, digested, and ligated into pEF6-V5-TOPO (Invitrogen, Carlsbad, CA) using the

manufacturer's protocol.

The N- and C-terminal portions of RS^{mjwt} were amplified via PCR using the following primers:

N-Forward: AAGGATCCACCATGGACGAATTTGAAATGAT

N-Reverse: ACATTCATATTGCCGAACCACTGGAATTCACTTCCAT

C-Forward: AGGGGATTTCGTTCACTGAGGTTATCTATCCAATAATGCA

C-Reverse: CCCGAATTCTAATCTCTTTCTAATTGGCT

The CP1 region of *E. coli* tyrosyl-tRNA synthetase was amplified from an in-house plasmid using the following primers:

Forward: ATGGAAGTGAATTCCAGTG GTTC GGCAATATGAATGT

Reverse: TGCATTATTGGATAGATAACTTCAGTGAACGAAATCCCCT

1 ug of each PCR product (N-terminus of RS^{mjwt} , C-terminus of RS^{mjwt}, and CP1 region of *E. coli* tyrosyl-tRNA synthetase) was combined, denatured for 15 minutes at 85° Celsius, and elongated with the Klenow fragment of *E. coli* DNA polymerase (New England Biolabs, Ipswich, MA). The fragment was then amplified using the N-Forward and C-Reverse primers mentioned on the preceding page. The resultant fragment was then digested, as was pEF6/V5, with HindIII and EcoRI. The digestion products were purified and cleaned up as mentioned in the preceding sections, and they were then ligated to afford the plasmid 309RS. This plasmid encodes the chimeric MEM protein shown in Figure 4.4. The correct sequence of this plasmid was confirmed by DNA sequencing (UT ICMB DNA Sequencing Facility, TX).

Amber Stop Codon Suppression Assay

HEK293T cells were grown to 70%-80% confluency as previously described. They were transfected with the elements of the UAA incorporation system described in the

preceding pages. The various combinations of constructs employed in those transfections are detailed in Table 4.1.

	40TAG-peGFP-N1	312tRNA	309RS	eGFP
40TAG	+	-	-	-
tRNA	+	+	-	-
Suppress	+	+	+	-
Positive	-	-	-	+

Table 4.1: DNA included in amber stop codon suppression assay conditions. The left hand column is the name of each condition, while the top row lists the DNA used. '+' means the DNA was included, '-' means the DNA was not included.

Transfections were carried out as previously described with FugeneHD. Complexes were made and added to cells according to the manufacturer's instructions using Opti-Mem as the diluents and a ratio of 2 µL FugeneHD to 1 µg DNA. Media was changed every 24 hours after transfection, and cells were imaged immediately prior to media changes using a Nikon Eclipse TE2000-S microscope with a FITC HyQ filter (Chroma, Rockingham, VT). After 72 hours, cells were harvested for FACS. As described previously, the passaging protocol was followed up to the point of centrifugation. After centrifugation, cells were washed once in PBS and recentrifuged. PBS supernatant was aspirated, and cells were resuspended in 4% paraformaldehyde in PBS (USB Corp. Cleveland, OH). Cells were incubated in paraformaldehyde for 15 minutes, then re-centrifuged. Paraformaldehyde was aspirated, and cells were resuspeded in PBS then subjected to FACS. FACS was accomplished using a BD FacsCalibur, and data analysis was done using Cyflogic.

Acceptor Stem Optimization and Creation of Super Orthogonal tRNA

To find an acceptor stem completely unlike any human tRNA, the tRNAdb 2009 was used (22). The acceptor stem sequences for every human tRNA were compiled, and then a search was done using the database to find an acceptor stem sequence for a tyrosyl tRNA with 0 sequence homology to any human tRNA. Only one tRNA was found which fit this description, belonging to the malaria causing parasite *Plasmodium falciparum (P. falciparum)*. To test the orthogonality of this acceptor stem, oligonucleotides encoding an amber suppressor form of the *P. falciparum* tRNA were ordered, shown below:

Forward: GGAAGATCTAAGTTAATGCCTGAGTGGTTAAAGGAATGGACTCTAAATCCA

TTGATAATATATCTACATCAGTTCAAATCTGATTTAACTTACCAAAGTTTCTCGAGCGG

Reverse: CCGCTCGAGAAACTTTGGTAAGTTAAATCAGATTTGAACTGATGTAGATATA

TTATCAATGGATTTAGAGTCCATTCCTTTAACCACTCAGGCATTAACTTAGATCTTCC

The forward and reverse oligonucleotides were combined in equimolar quantities, boiled, and allowed to come to room temperature unaided. They were then digested with BglII and PstI, as was the plasmid 312tRNA. The digested 312tRNA was separated via agarose gel electrophoresis and gel purified. The digested and purified *P. falciparum* tRNA fragment was then ligated into the 312tRNA backbone to afford the plasmid PfalctRNA. The correct sequence of this plasmid was confirmed by sequencing (UT ICMB DNA Sequencing Facility, TX).

Creation of 2-Plasmid UAA incorporation System

The plasmid pBi-CMV4 was purchased from Clontech (Mountainview, CA). The chimeric MEM synthetase was amplified by PCR out of the plasmid 309RS using the primers below.

Forward: CTTAAGCTTCCATGGACGAATTTGAAATG

Reverse: TCCGATATCGCTTATAATCTCTTTCTAATTGG

The resultant PCR product and the plasmid pBi-CMV4 were separately digested with EcoRV and HindIII. pBiCMV4 was also treated with CIP. Both double digest products (pBiCMV4 and PCR product) were separated via agarose gel electrophoresis as previously described. They were then gel purified using a Qiagen Gel Purification Kit (Valencia, CA). The PCR product was then ligated into pBiCMV4 using T4 DNA Ligase (Fermentas Inc., Thermo Scientific, Glen Burnie, MD) to afford the plasmid pBi-309.

The DNA encoding eGFP with an amber stop codon at the 40^{th} position was amplified by PCR out of 40TAG-peGFPN1 using the primers below.

Forward: TGGAGAATTCTGCAGTCGACGGT

Reverse: CCTCTAGAGTCGCGGCCGCTTTACTTGTACAGCTCGT

The resultant PCR product, along with pBi-309, was digested with EcoRI and XbaI. pBi-309 was then treated with CIP to prevent self ligation. Both digestions were then separated by agarose gel electrophoresis, and the appropriate bands were removed then cleaned up as described above. The 40TAG-GFP segment was ligated into the pBi-309 backbone using T4 DNA Ligase (Fermentas Inc., Thermo Scientific, Glen Burnie, MD) to afford the plasmid 40TAG-bi-309. The correct sequence of this plasmid was confirmed by sequencing (UT ICMB DNA Sequencing Facility, TX).

In order to remove both the human H1 promoter as well as the previously designed orthogonal tRNA, the plasmid 312tRNA was double digested with BglII and XhoI, as was pBiCMV4. The digested pBiCMV4 was treated with CIP, and then separated using agarose gel electrophoresis, as was the digested 312tRNA plasmid. The appropriate bands were cut out of each lane and purified as previously described. The 312tRNA insert was ligated into the pBiCMV4 backbone using T4 DNA Ligase (Fermentas

Inc., Thermo Scientific, Glen Burnie, MD) to afford the plasmid pBi-312. Since pBiCMV4 already has a gene in it for a variant of red fluorescent protein (RFP), this plasmid was created so that cells could monitored for GFP expression, indicative of stop codon suppression, and for RFP, indicative of transfection with the tRNA carrying plasmid. The correct sequence of this plasmid was confirmed by sequencing (UT ICMB DNA Sequencing Facility, TX).

Testing of 2-Plasmid UAA Incorporation System

The components of the two-plasmid incorporation system were tested in HEK293T cells. The cells were plated in 6-well plates and grown to 70%-80% confluency. They were transfected with the respective components of the 3-plasmid UAA incorporation system and the 2-plasmid UAA incorporation system using Fugene HD as previously described. Media was changed 24 hours after transfection and every 24

	40TAG-peGFP-N1	312tRNA	309RS	pBi-312	40TAG-Bi-309	eGFP
40TAG	+	-	-	-	-	-
tRNA	+	+	-	-	-	-
Suppress	+	+	+	-	-	-
40TAG-Bi	-	-	-	-	+	-
Bi-tRNA	+	-	-	+	-	-
Bi-Suppress	-	-	-	+	+	-
Hybrid	-	+	-	-	+	-
eGFP	-	-	-	-	-	+

Table 4.2: DNA included in comparison of 3-plasmid and 2-plasmid UAA incorporation systems. The left hand column is the name of each condition, while the top row lists the DNA used. '+' means the DNA was included, '-' means the DNA was not included.

hours after that. Cells were imaged using a Nikon Eclipse TE2000-S microscope with a FITC HyQ filter (Chroma, Rockingham, VT) every 24 hours as well. Depending upon the

experiment, cells were harvested for FACS at either 48 hours or 72 hours post-transfection. FACS harvesting was done as previously described on page 84. The DNA constructs used in each condition are shown in Table 4.2.

RESULTS AND DISCUSSION

Reporter Protein Plasmid Optimization

Since the plasmid encoding the reporter protein used by another group consistently yielded background expression of reporter protein, a new one was created based upon the peGFP-N1 plasmid from Clontech. This plasmid was optimized for expression in mammalian cells and had a strong constitutive promoter. The promoter on the previous plasmid, Lital40TAG, was unknown, and sequencing attempts consistently failed. Thus, SDM was used to create the new fluorescent reporter protein encoding plasmid, 40TAG-peGFPN1, which is the same as the original plasmid except for a mutation from a tyrosine codon to an amber stop codon at the 40th position. In order to confirm that the new reporter protein encoding plasmid did actually yield lower

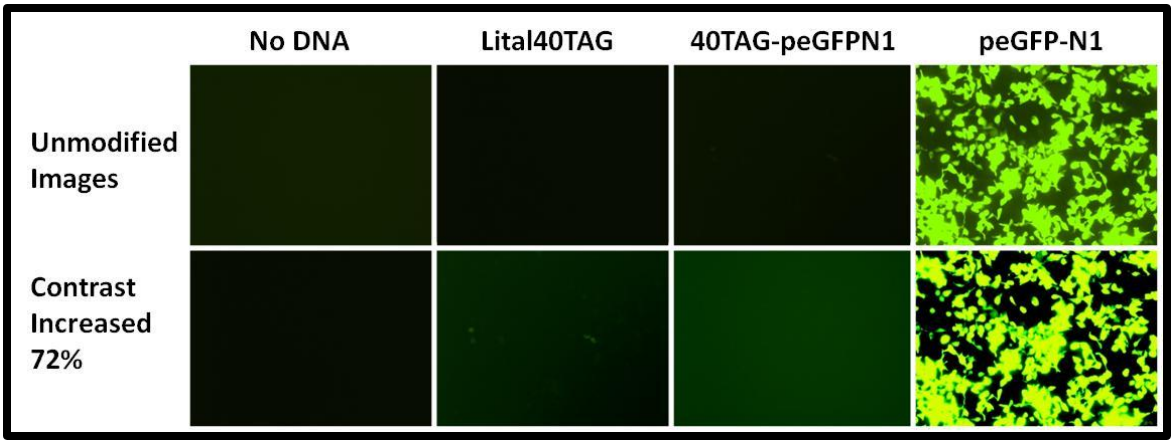


Figure 4.5: Fluorescent images demonstrating decreased background expression of 40TAG GFP from new plasmid 40TAG-peGFPN1 as compared to Lital40TAG. Unmodified images were taken 48 hours after transfection with gain set to 1 and 1 second exposure time. Contrast Increased images (bottom row) are the original images with contrast increased 72% using Adobe Photoshop.

background expression than the earlier version, HEK293T cells were transfected with each plasmid and expression of GFP was assessed via fluorescence microscopy as well as FACS. In order to facilitate visual appreciation of the background expression of reporter protein in the Lital40TAG condition, the original images were modified by increasing contrast by 72% using Adobe Photoshop. As evidenced by Figure 4.5, 40TAG-peGFPN1 produces less background expression of TAG-GFP protein than its predecessor Lital40TAG. Though this difference is difficult to see, the contrast enhanced version of the Lital40TAG image clearly shows four fluorescent cells. Neither the negative control, nor the 40TAG-peGFPN1 condition exhibits this same level of fluorescence.

To further confirm that 40TAG-peGFPN1 produces less background than Lital40TAG, cells were harvested and subjected to FACS using a BD FacsCalibur. As previously mentioned, dead cells were excluded using forward and side scatter, and the category of GFP positive was assigned to cells exhibiting higher fluorescence than those in the No DNA negative control condition. Figure 4.6 shows the FACS results, confirming the information presented in Figure 4.5, that 40TAG-peGFPN1 produces less background expression of viable GFP protein. Though the mechanism behind this background expression from the Lital40TAG plasmid, my colleagues and I have speculated that there may be more than one version of TAG-GFP gene in that plasmid, which may explain the consistent sequencing failures. If the second version of the TAG GFP gene had the amber stop codon in a place downstream of the UV core of GFP, it may be possible for some percentage of that protein to fold well enough to exhibit fluorescence. Regardless of the mechanism behind the background expression found in Lital40TAG transfected cells, the new plasmid 40TAG-peGFPN1 clearly produces less background and thus was used for all subsequent experiments.

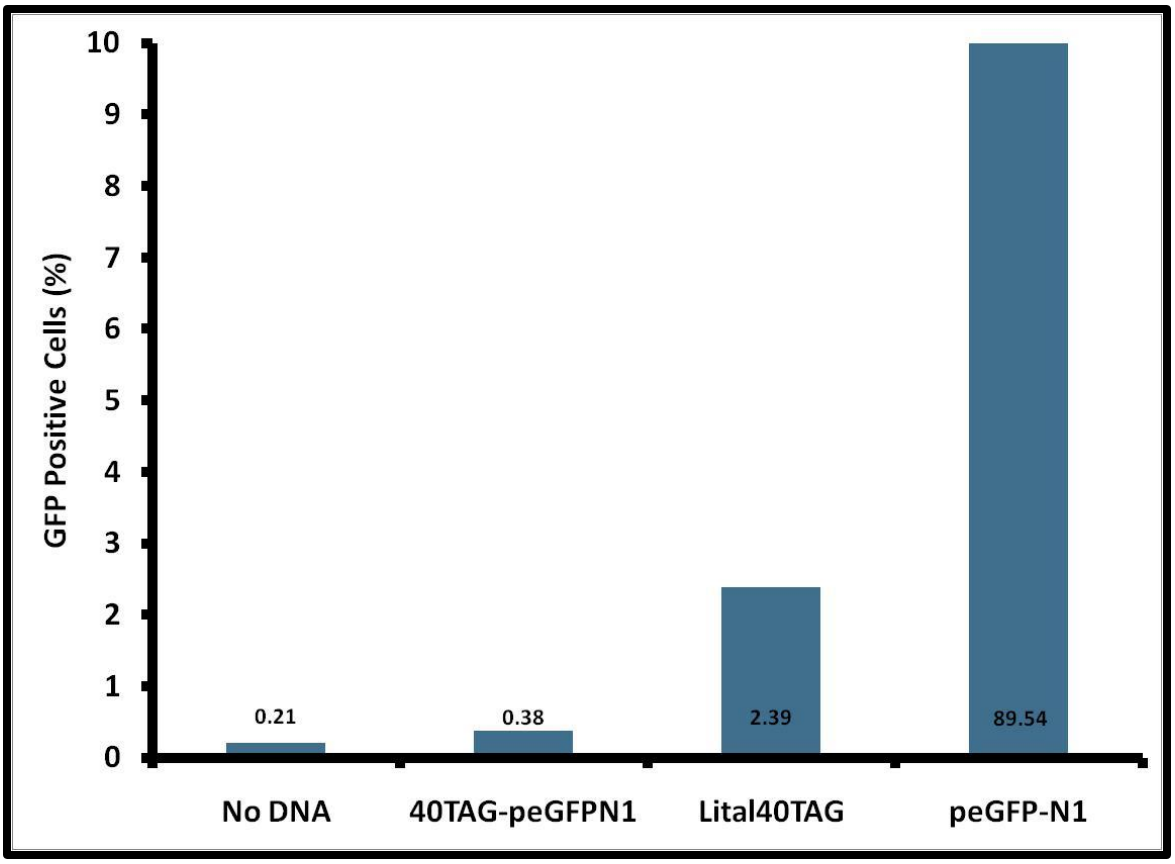


Figure 4.6: HEK293T cells exhibiting GFP fluorescence as determined by FACS. The new plasmid 40TAG-peGFPN1 exhibits 80% less background fluorescence than Lital40TAG and less than 50% more than the negative control.

Testing of Orthogonal tRNA

The plasmids encoding both the newly created orthogonal tRNA (312tRNA plasmid) as well as the non-orthogonal tRNA (wtMJtRNA plasmid) were transfected into HEK293T cells along with 40TAG-peGFPN1. If full length GFP were produced in either condition, it would indicate that the tRNA encoded within the transfected plasmid was not orthogonal. Therefore, if the tRNA encoded by the plasmid 312tRNA is as orthogonal as intended, it should not produce full length GFP. Since the tRNA encoded by the plasmid wtMJtRNA is not intended to be orthogonal, it was used as a measure of tRNA

transcription levels as well as the capacity of HEK293T cells to produce functional exogenous tRNA from plasmid DNA. Both **Figure 4.7** and Figure 4. 8 demonstrate successful expression of amber suppressor tRNA (wtMJtRNA) and orthogonality of the newly created 312tRNA.

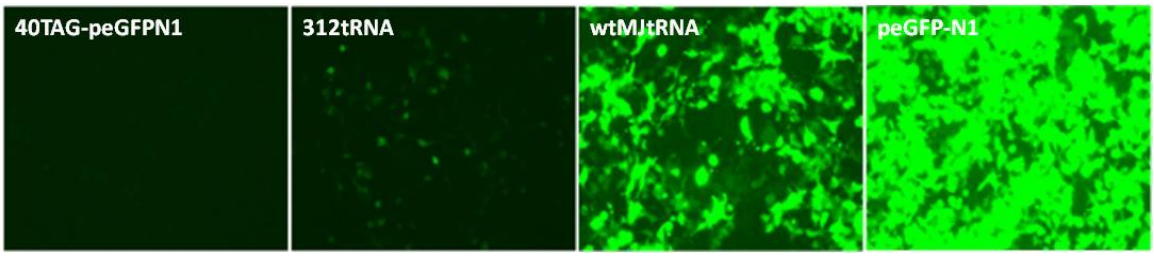


Figure 4.7: Fluorescent images demonstrating orthogonality of 312tRNA as well as tRNA expression (wtMJtRNA) in HEK293T cells. Images were taken 48 hours after transfection with gain set to 1 and 1 second exposure time

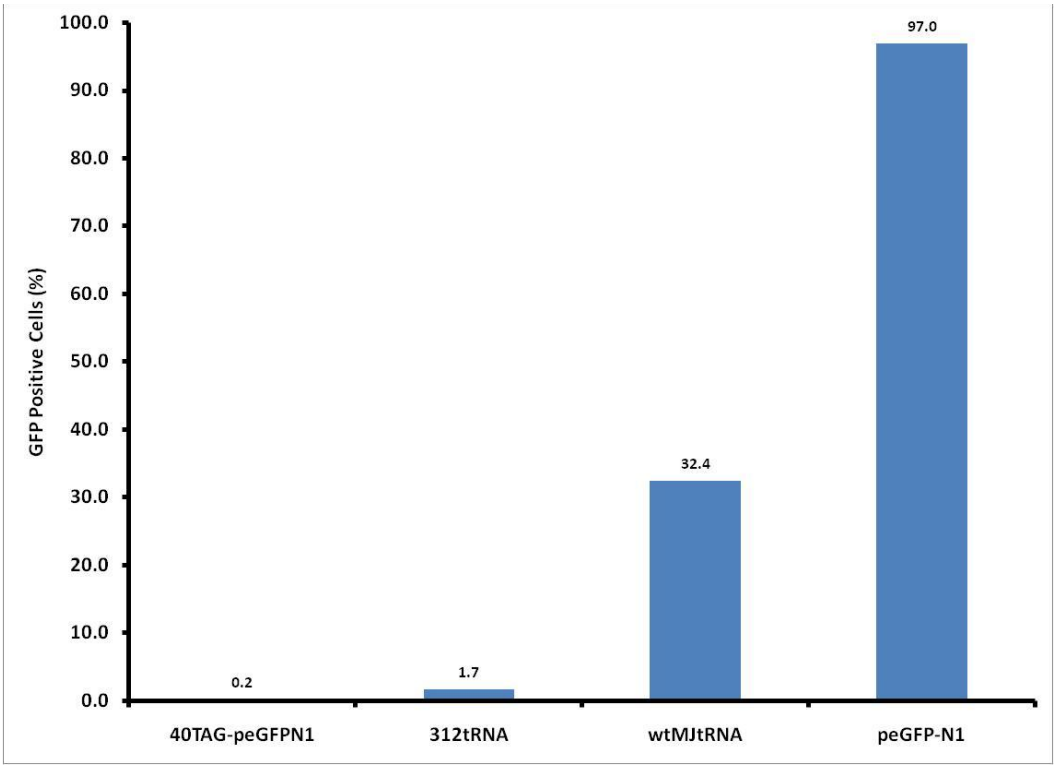


Figure 4. 8 : HEK293T cells exhibiting GFP fluorescence as a result of amber suppression as determined by FACS.

Testing of Complete UAA Incorporation System

In order to confirm the intended functionality of the complete UAA incorporation system, including the products of plasmids 40TAG-peGFPN1, 312tRNA, and 309RS, the components of the system were transfected into HEK293T cells as detailed in Table 4.1. Images taken 48 hours after transfection are shown in Figure 4.7 below.

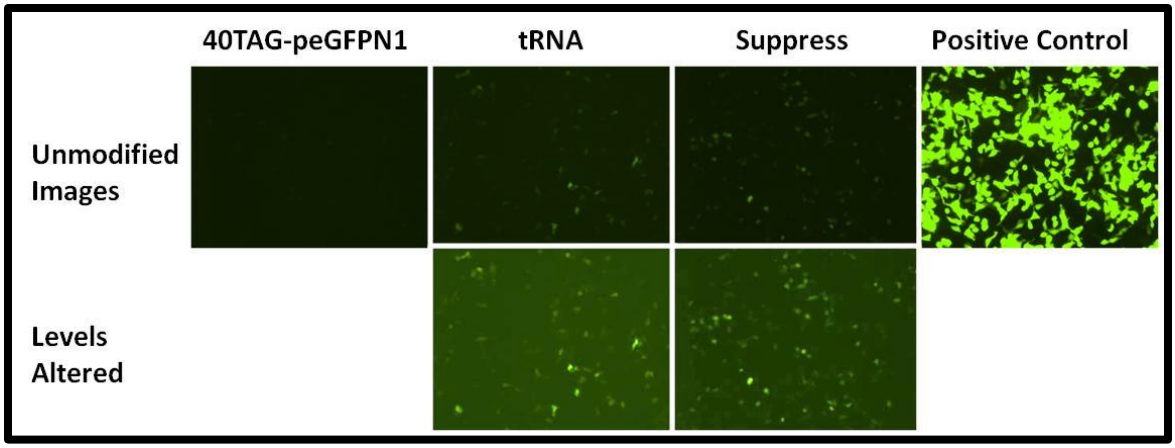


Figure 4.9: Fluorescent images demonstrating functionality of UAA Incorporation System in HEK293T cells transfected as described in Table 4.1. Unmodified images were taken 48 hours after transfection with gain set to 1 and 1 second exposure time. Levels Altered images (bottom row) are the original images with the levels changed from 0-255 to 0-110 using Adobe Photoshop. The Levels Altered version of the Suppress condition clearly demonstrates are higher percentage of fluorescent cells than the Levels Altered version of the tRNA control condition.

Since the differences between the tRNA control condition and the Suppress condition were subtle in the original pictures, the levels of each image were changed from 0-255 to 0-110 using Adobe Photoshop. This technique brings out the green cells more clearly than the original images. In the Levels Altered row, the Suppress Condition clearly has about twice the percentage of fluorescent cells as the tRNA condition, which

indicates a number of things. First, the engineered chimeric synthetase encoded by the 309RS plasmid appears to successfully charge the tRNA encoded by the 312tRNA plasmid. This is indicated by the increased percentage of fluorescent cells in the Suppress condition versus the tRNA condition. Second, the tRNA is not completely orthogonal. Though this was previously established by the tRNA orthogonality assays, it serves to further solidify the idea that creating an orthogonal tRNA for use in mammalian cells may be more difficult than previously anticipated. This was just one of the many reasons why it was prudent to use the tRNAdb 2009 and find an acceptor stem completely different than any human acceptor stem, as discussed in the following section.

Creation of Super Orthogonal tRNA

Since many groups have previously established that nucleotides in the acceptor stem other than the 1:72 base pair and discriminator nucleotide often play a role in identity determination (3, 4, 6, 9, 23, 24), we attempted to make a super orthogonal tRNA by finding an acceptor stem sequence with 0% homology with any human tRNA acceptor stem. Since the pilot test of this tRNA occurred in human cells, it was to be a proof of concept that screening for acceptor stem homology within tRNAs native to the host cell can improve orthogonality. In the event that this technique worked, it could then be applied to other systems. Thus, by comparing acceptor stem sequences of all known human tRNAs to acceptor stem sequences from all known tyrosyl-tRNAs from any organism, a tRNA from the parasite *Plasmodium falciparum* was found which had a completely different acceptor stem sequence than any human tRNA. Oligonucleotides were ordered for an amber suppressor tRNA based upon this sequence, and its DNA was cloned into the previously developed expression cassette (see page 82). To test the

orthogonality of this tRNA, it was cotransfected into HEK293T cells along with 40TAG-peGFPN1. Thus, orthogonality could be assessed by fluorescence as with previous orthogonality assays. For comparison, cells were also cotransfected with the previously created 312tRNA and 40TAG-peGFPN1. Though once again the fluorescent images do not make the distinction between conditions overt, there are subtle differences as evidenced by Figure 4.8.

Figure 4.10: Comparison of *P. falciparum* amber suppressor tRNA and tRNA produced by 312tRNA plasmid. Higher fluorescence in the *P. falciparum* condition indicates that it is less orthogonal than the tRNA encoded by the 312tRNA plasmid. Images were taken 48 hours after transfection, each with gain set to 1 and a 1 second exposure time.

Interestingly, the difference between the *P. falciparum* condition and the 312tRNA condition is difficult to discern. However it is clear that neither the *P. falciparum* condition nor the 312tRNA condition is entirely orthogonal. From this information alone, one cannot determine conclusively whether or not the *P. falciparum*

tRNA is a suitable candidate for further investigation or optimization, but the FACS data aid tremendously in this endeavor.

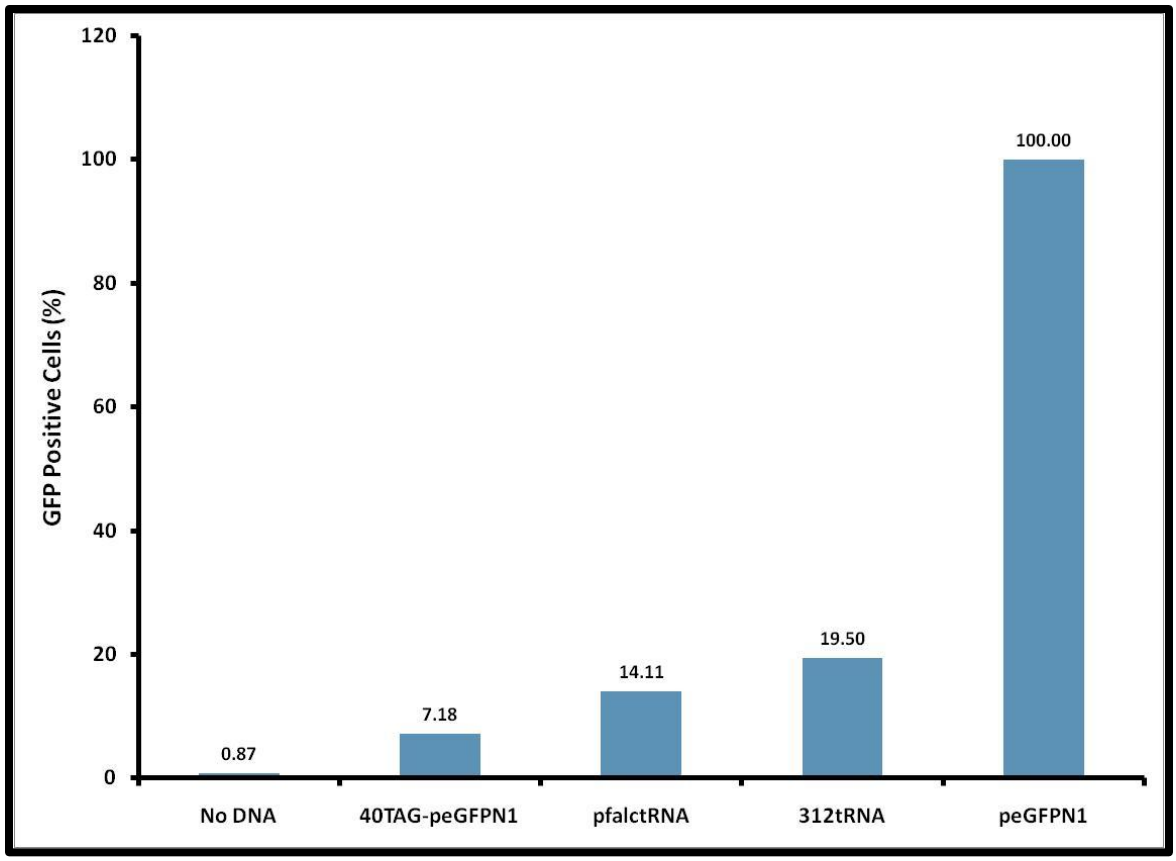


Figure 4.11: HEK293T cells exhibiting GFP fluorescence as a result of amber suppression, determined by FACS.

As shown in Figure 4.9, the pfalctRNA does actually produce less fluorescence than the 312tRNA, though both of them produce suboptimal amounts of amber suppression in the absence of synthetases engineered to charge them with either amino acids or UAAs. In these experiments, the voltage on the blue laser used to excite GFP fluorescence was turned up in order to highlight the differences between the two tRNAs tested and the 40TAG-peGFPN1 negative control. If compared only to each other, the pfalctRNA seems to exhibit about 25% less fluorescence than the 312tRNA, but that

comparison is not the most faithful metric for comparing the two tRNAs. Instead, they should each be compared to the 40TAG-peGFPN1 condition. When that comparison is done, the pfalctRNA produces about twice the level of fluorescence as the reporter plasmid alone, and the 312tRNA produces almost three times the fluorescence of the reporter plasmid alone. If this information alone was used to judge the usefulness of the pfalctRNA, then it would indicate that the pfalctRNA is a better candidate for further development than the 312tRNA.

Unfortunately, orthogonality is not the only consideration in this project. The *M. jannaschii* system was initially chosen because multiple synthetases have already been developed and in fact already exist in the lab for the incorporation of multiple UAAs. The process involved in developing new synthetases is anything but simple, but would be worthwhile if a completely orthogonal tRNA could be found. In that case, development of a new system using that tRNA as a jumping off point would make sense both scientifically as well as financially. However, since the *P. falciparum* tRNA did not represent even a 50% decrease in background amber suppression, it seemed an unwise decision to pursue it any further. Since it did decrease the background to a degree, we decided to test it with a number of in-house synthetases to see if perhaps it would fit with one of them. If so, then it would be a useful addition to the project, and if not, the project could move forward with the suboptimally orthogonal 312tRNA.

One important piece of information indicated by the *P. falciparum* tRNA is that orthogonality is more complex than we currently understand. Though the *P. falciparum* tRNA contains an acceptor stem which is completely foreign to the mammalian system, there is some endogenous synthetase, and perhaps multiple synthetases, which recognize this foreign acceptor stem and successfully charge it with an amino acid. Perhaps the most incredible part of this phenomenon is the fact that, taken at face

value, it may lead one to believe that endogenous RSs are simply sloppy and capable of making mistakes. Once again, that would be a hasty and incorrect conclusion from this information. First of all, we have no metric with which to gauge the amount of tRNA produced by the 312tRNA or the pfalctRNA plasmids. It could be that these tRNAs are produced in quantities much higher than endogenous tRNAs, making the cells statistically more prone to RS recognition errors than they would be in a normal situation. Even if these tRNAs are produced in the same quantities or less than endogenous tRNAs, there is still little evidence to support the idea that endogenous RSs could be making recognition mistakes. This is simply due to the fact that, if they were, certainly they would be making those mistakes in the absence of exogenous tRNAs, and those mistakes would be reflected in an inability to produce and then analyze the sequence of proteins expressed in mammalian cells. Since proteins are successfully expressed all the time in mammalian cells without frequently misincorporated amino acids, the reason for mischarging of the exogenous tRNAs in this project is very unlikely to be mistakes made by endogenous RSs. Thus, what we can conclude from this portion of the project is that tRNA identity elements are still poorly understood, making it difficult to create any completely orthogonal UAA incorporation system. Keeping that in mind, the 312tRNA, which exhibits about 33% more background than the pfalctRNA, is an acceptably orthogonal solution to a poorly understood problem.

Testing of 2-plasmid UAA Incorporation System

The UAA incorporation system initially developed on pages 84 and 85 of this Chapter is housed in three separate plasmids. In order to decrease the burden of optimization on stem cell transfection (described in Chapter Three), this system was condensed into two bidirectional plasmids. This two plasmid UAA incorporation system

was then tested in HEK293T cells to assess its functionality as compared to the original three plasmid UAA incorporation system. HEK293T cells were transfected with either the two-plasmid or the three-plasmid system and assessed via fluorescence microscopy and FACS. Figure 4.10 shows the images of cells transfected with the different UAA incorporation systems 72 hours post-transfection. From this figure, it is clear that the two-plasmid system performs poorly in comparison to the three-plasmid system, which is somewhat surprising. Unfortunately, this conclusion is further corroborated by the FACS data in Figure 4.10, which shows again that the two-plasmid system achieves a lower level of amber stop codon suppression than the three-plasmid system. On one hand, the two-plasmid system exhibits little to no background, but it also exhibits little to no suppression. Thus, in choosing which system to move forward with, the logical conclusion is to choose the three-plasmid system, though a more ideal situation would be to improve the performance of the two-plasmid system.

In order to improve the performance of the two-plasmid system, it is first necessary to identify the possible reasons for its surprisingly bad performance. Though time was insufficient to investigate these reasons, there are a few possible culprits that may be targets for future improvements of this system. The original 40TAG-peGFPN1 plasmid contains a cytomegalovirus (CMV) promoter, but the pBi-CMV4 plasmid which houses the two-plasmid system has a minimal CMV promoter. It is possible that the minimal CMV promoter simply does not produce enough protein in HEK293T cells to be useful for a UAA incorporation system. Furthermore, the 312tRNA plasmid has tet-responsive element (TRE) upstream of the human H1 promoter. While the human H1 promoter was transferred to the two-plasmid system, no tests were ever done to assess the influence of the TRE based promoter on tRNA transcription levels. Thus, it is difficult to say whether or not tRNA is being produced by the two-plasmid system. With any

tRNA tested thus far, there has been some level of recognition and charging of that tRNA by endogenous synthetases. Since the two-plasmid system does not appear to exhibit any background at all, it is likely that it is not producing enough tRNA, if it is producing any at all.

If the two-plasmid version of the UAA system had excluded the H1 promoter, depending only on the CMV promoter to drive tRNA transcription, it would seem likely that very little if any tRNA was being produced. Boden et al compared the ability of the CMV promoter, the human H1 promoter, and several others to produce RNA in HEK293T cells (25). They found that the CMV promoter was a poor choice for transcription of RNA, and in fact its ability to produce silencing RNA was almost identical to a negative control (25). Thus, if the two-plasmid system had used only a CMV promoter to drive tRNA transcription, that would have constituted a very serious design flaw and easily explained the poor performance of the system. However, as with many of the tRNA producing plasmids previously developed in our lab, this one left the original protein promoter intact, and inserted the tRNA promoter downstream of it. Thus, it is difficult to ascertain exactly why the two-plasmid system did not function as highly as the three-plasmid system. Without further investigation, a definitive conclusion is impossible to make. However, removing the CMV promoter from the two-plasmid system and testing a variety of promoters, including full-length CMV promoters for 40TAG-GFP and 309RS genes, is a logical next step.

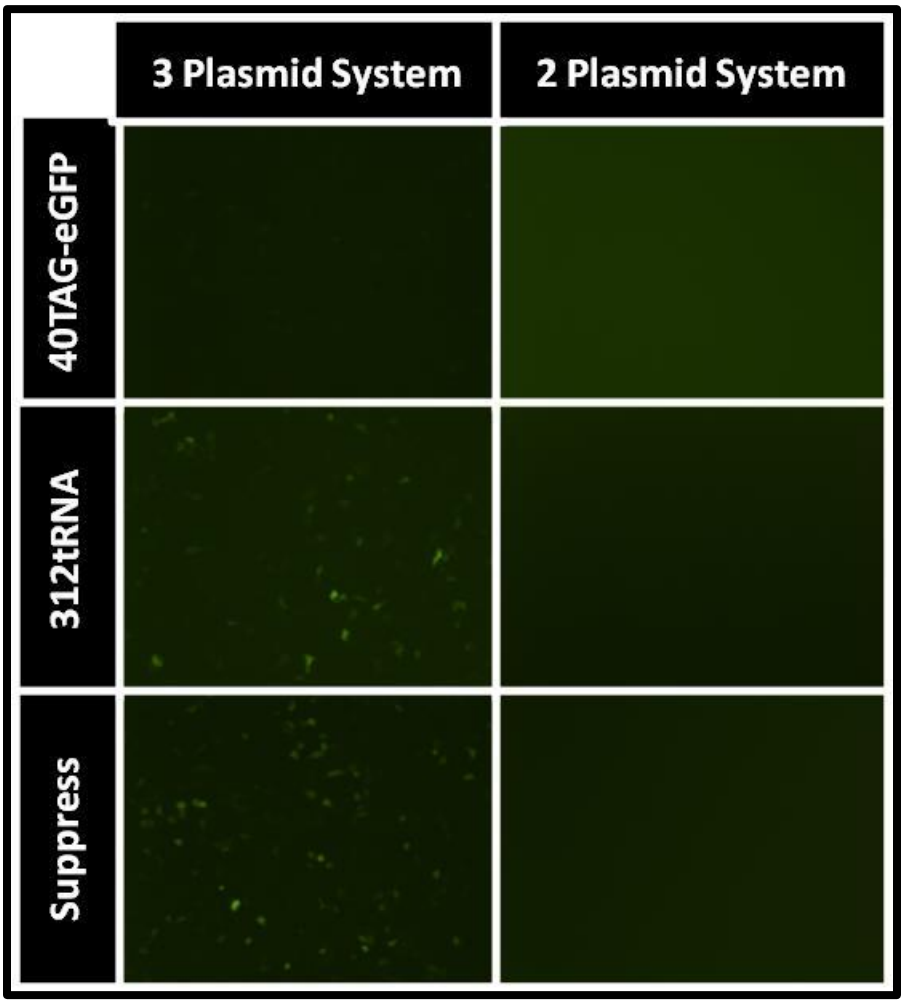


Figure 4.12: Comparison of two-plasmid versus three-plasmid UAA Incorporation System in HEK 293T cells indicating decreased performance of the two-plasmid system as compared to the three plasmid system. (A) GFP fluorescence at 72 hours post-transfection. Images captured with gain of 1 and 1 second exposure. (B) GFP positive cells as determined by FACS.

REFERENCES

1. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P, editors. Molecular Biology of the Cell. 4 ed. New York: Garland Science; 2002.

2. Giege R, Puglisi JD, Florentz C. tRNA structure and aminoacylation efficiency. Prog Nucleic Acid Res Mol Biol. 1993;45:129-206.

3. Bedouelle H, Winter G. A model of synthetase/transfer RNA interaction as deduced by protein engineering. Nature. 1986;320(6060):371-3.

4. Bedouelle H. Recognition of tRNA(Tyr) by tyrosyl-tRNA synthetase. Biochimie. 1990;72(8):589-98.

5. Himeno H, Hasegawa T, Ueda T, Watanabe K, Shimizu M. Conversion of aminoacylation specificity from tRNA(Tyr) to tRNA(Ser) in vitro. Nucleic Acids Res. 1990;18(23):6815-9.

6. Bedouelle H, Guez-Ivanier V, Nageotte R. Discrimination between transfer-RNAs by tyrosyl-tRNA synthetase. Biochimie. 1993;75(12):1099-108.

7. Kisselev LL. The role of the anticodon in recognition of tRNA by aminoacyl-tRNA synthetases. Prog Nucleic Acid Res Mol Biol. 1985;32:237-66.

8. Giege R, Sissler M, Florentz C. Universal rules and idiosyncratic features in tRNA identity. Nucleic Acids Res. 1998;26(22):5017-35.

9. Chambers RW. On the recognition of tRNA by its aminoacyl-tRNA ligase. Prog Nucleic Acid Res Mol Biol. 1971;11:489-525.

10. Ebel JP, Giege R, Bonnet J, Kern D, Befort N, Bollack C, et al. Factors determining the specificity of the tRNA aminoacylation reaction. Non-absolute specificity of tRNA-aminoacyl-tRNA synthetase recognition and particular importance of the maximal velocity. Biochimie. 1973;55(5):547-57.

11. Shimura Y, Ozeki H. Genetic study on transfer RNA. Adv Biophys. 1973;4:191-226.

12. Buechter DD, Schimmel P. Aminoacylation of RNA minihelices: implications for tRNA synthetase structural design and evolution. Crit Rev Biochem Mol Biol. 1993;28(4):309-22.

13. Alexander RW, Nordin BE, Schimmel P. Activation of microhelix charging by localized helix destabilization. Proc Natl Acad Sci U S A. 1998;95(21):12214-9.

14. Francklyn C, Schimmel P. Aminoacylation of RNA minihelices with alanine. Nature. 1989;337(6206):478-81.

15. Martinis SA, Schimmel P. Enzymatic aminoacylation of sequence-specific RNA minihelices and hybrid duplexes with methionine. Proc Natl Acad Sci U S A. 1992;89(1):65-9.

16. Shi JP, Francklyn C, Hill K, Schimmel P. A nucleotide that enhances the charging of RNA minihelix sequence variants with alanine. Biochemistry. 1990;29(15):3621-6.

17. Shi JP, Martinis SA, Schimmel P. RNA tetraloops as minimalist substrates for aminoacylation. Biochemistry. 1992;31(21):4931-6.

18. Shi JP, Schimmel P. Aminoacylation of alanine minihelices. "Discriminator" base modulates transition state of single turnover reaction. J Biol Chem. 1991;266(5):2705-8.

19. Wakasugi K, Quinn CL, Tao N, Schimmel P. Genetic code in evolution: switching species-specific aminoacylation with a peptide transplant. Embo J. 1998;17(1):297-305.

20. Steer BA, Schimmel P. Major anticodon-binding region missing from an archaebacterial tRNA synthetase. J Biol Chem. 1999;274(50):35601-6.

21. Thibodeaux G, Liang X, Moncivais K, Umeda A, Singer O, Alfonta L, et al. Transforming a Pair of Orthogonal tRNA-aminoacyl-tRNA Synthetase from Archaea to Function in Mammalian Cells. PLOS ONE. 2010;5(6):-.

22. Juhling F, Morl M, Hartmann RK, Sprinzl M, Stadler PF, Putz J. tRNAdb 2009: compilation of tRNA sequences and tRNA genes. Nucleic Acids Res. 2009;37(Database issue):D159-62.

23. Beuning P, Musier-Forsyth K. Transfer RNA recognition by aminoacyl-tRNA synthetases. BIOPOLYMERS. 1999;52(1):1-28.

24. Fabrega C, Farrow MA, Mukhopadhyay B, de Crecy-Lagard V, Ortiz AR, Schimmel P. An aminoacyl tRNA synthetase whose sequence fits into neither of the two known classes. Nature. 2001;411(6833):110-4.

25. Boden D, Pusch O, Lee F, Tucker L, Shank PR, Ramratnam B. Promoter choice affects the potency of HIV-1 specific RNA interference. Nucleic Acids Res. 2003;31(17):5033-8.

CHAPTER FIVE

Amber Stop Codon Suppression in P19 Embryonal Carcinoma Cells

INTRODUCTION

Chapter Four described the creation of an unnatural amino acid (UAA) incorporation system for use in mammalian cells by CP1 swap methodology, and Chapter Three described the development of a transfection protocol capable of efficiently transfecting P19 embryonal carcinoma cells (P19s). In this chapter, the use of the UAA incorporation system in P19 embryonal carcinoma cells using the previously developed transfection method will be described. First, some relevant background will be reviewed and expanded upon, and then our achievement of amber stop codon suppression in P19s will be described in detail.

Overall Experimental Design

Since suppression in HEK293T cells was successful, we chose to directly transfer the components of the UAA incorporation system to P19s. Both the original electroporation method and the hybrid FugeneHD+electroporation method (see Chapter Three) were used, because work on many aspects of this project was completed in parallel. Since the suppression in HEK293T cells was not remarkably successful, it was uncertain whether or not the UAA incorporation system developed in Chapter Four would function in P19s. Undifferentiated cells are known to respond differently to various promoters than other cell lines, and in fact some are unresponsive to cytomegalovirus (CMV) based promoters prior to differentiation (1, 2). Thus, we held it as a possibility that the UAA incorporation system might not work in P19s. Further, we concede the fact that CP1 swap methodology does work in HEK293T cells (3), but that

does not mean that protein translational machinery and protein folding pathways are identical in HEK293T cells and P19s. This means that while UAA incorporation systems for use in regular, easily transfectable, non-differentiable mammalian cells lines can be created via CP1 swap methodology and simple tRNA engineering, again, this does not necessarily transfer perfectly to P19s. To that end, we reasoned that another, more robust (though less accessible), UAA incorporation system should be tested for successful amber stop codon suppression in mammalian cells. So we created the components for using a UAA incorporation system based upon the *Escherichia coli (E. coli)* tyrosyl-tRNA/aminoacyl-tRNA synthetase pair. Since this pair has been evolved to incorporate a number of UAAs in eukaryotic cells, it is possible that the active site and tRNA recognition elements of the synthetase will function at a higher level than those of the chimeric enzyme created in Chapter Four. Therefore, both the *E. coli* and *Methanocaldococcus jannaschii (M. jannaschii)* based UAA incorporation systems were introduced to P19s using the previously developed transfection methods. Success of either or both UAA incorporation systems in suppressing an amber stop codon would allow progression to incorporation of useful UAAs into proteins in P19s.

MATERIALS AND METHODS

Routine Cell Culture

HEK293T cells were purchased from Invitrogen (Carlsbad, CA). Complete HEK293T media was composed of high glucose DMEM with 4500 mg/L of sodium bicarbonate and L-glutamine (Sigma-Aldrich, St. Louis, MO) supplemented with 10% (v/v) fetal bovine serum (FBS) (Atlanta Biologicals, Lawrenceville, GA) and 1X

nonessential amino acids (Sigma-Aldrich, St. Louis, MO). Cells were maintained in a humidified incubator at 37° Celsius with 5% CO_2 atmosphere. Cells were passaged every 2-3 days at ratios between 1:10 and 1:20. Passaging was accomplished by washing cells with phosphate buffered saline (PBS) (Sigma-Aldrich, St. Louis, MO), treating with 0.5% trypsin (Sigma-Aldrich, St. Louis, MO) for 2-5 minutes, and harvesting via centrifugation. In between passaging, fresh media was added to cells every 24 hours.

P19 embryonal carcinoma cells were obtained from American Type Culture Collection (ATCC). Cells were routinely maintained in Alpha-Mem (HyClone, Logan, UT) with 10% (v/v) fetal bovine serum (Atlanta Biologicals, Lawrenceville, GA) and nonessential amino acids (Sigma-Aldrich, St. Louis, MO) in a humidified incubator at 37° Celsius with 5% CO_2 atmosphere. Cells were passaged every 1-3 days at a ratio between 1:4 and 1:20. For passaging, cells were rinsed in phosphate buffered saline (PBS) (Sigma-Aldrich, St. Louis, MO), incubated with 0.5% trypsin (Gibco, Invitrogen, Carlsbad, CA) for 3-5 minutes, and centrifuged to pellet. Cells were then resuspended in complete media (as described earlier in this paragraph) and seeded into clean flasks with fresh media.

Creation of *E. Coli* Based UAA Incorporation System

Primers and stuff were used to create the plasmids ectRNA and ecRS, encoding wildtype *E. coli* tRNAtyr and wildtype *E. coli* aminoacyl tyrosyl tRNA synthetase respectively. Briefly. Correct sequences for each plasmid were confirmed by DNA sequencing (UT ICMB DNA Sequencing Facility, TX).

DNA Constructs

The plasmids 40TAG-peGFPN1, 312tRNA, and 309RS encoding the components of the *M. jannaschii* based UAA incorporation system were developed in Chapter Four and used for the investigations in this chapter.

Transfection and Amber Stop Codon Suppression Assays

One important note about electroporation techniques: DNA must be quite concentrated in order to use the amounts necessary in the volume of a cuvette. Thus, all DNA prep utilized either Maxi or Mega Prep kits from Qiagen (Valencia, CA), and DNA was suspended at 1 µg/uL at a minimum. All DNA was sterile filtered in a laminar flow hood prior to use in transfection experiments.

FugeneHD Transfection of HEK293T Cells

FugeneHD was purchased from Roche (Switzerland). HEK293T cells were grown to 70-80% confluency in 6-well plates (Corning, Lowell, MA) and then cotransfected 40TAG-peGFPN1 and each of the various tRNA constructs mentioned previously (i.e. each tested tRNA was transfected along with 40TAG-peGFPN1, not with other tRNAs). FugeneHD was used for transfection, and complexes were formed in Opti-Mem using a ratio of 2 µL FugeneHD to 1 µg DNA. Media was changed every 24 hours after transfection.

Electroporation Transfection

P19 embryonal carcinoma cells at 80% - 90% confluency were harvested using the same procedure as used for passaging, washed once with PBS (Sigma-Aldrich, St. Louis, MO), spun down again, and resuspended in 600 µL of PBS per harvested T-75 flask of cells. For each condition in an experiment, 200 µL of cells was aliquotted into an eppendorf tube. Then DNA appropriate for each condition was added to the cells. The total volume for each electroporation was then brought up to 450 µL by addition of PBS. Cell DNA mixtures were then transferred to electroporation cuvettes with a 2 mm gap (Fisher Scientific, Waltham, MA), and cuvettes were incubated on ice for no less than 5 minutes and no more than 15 minutes before electroporation. Cell-DNA mixtures were

electroporated using a Bio-Rad Gene Pulser with capacitance extender and pulse controller (Bio-Rad, Hercules, CA) with settings of 270 Volts, 100 Ω, and 960 μF. After electroporation, cells were incubated on ice for at least 3 minutes and then resuspended in complete media. Cells were then plated in tissue culture vessels of the appropriate size (Corning, Lowell, MA). Media was changed 24 hours post-transfection and every 24 hours thereafter until experiments were completed.

Hybrid FugeneHD + Electroporation Transfection

Fugene HD was purchased from Roche (Switzerland). A modified version of Roche's FugeneHD protocol was used for the hybrid transfection procedure. Instead of 50 μL of diluent per μg of DNA, 25 μL of diluent was used per μg of DNA. Assuming the concentration of DNA were 1 ug/μL, the following protocol would be used. First, PBS is added to an eppendorf tube such that the final volume of PBS + DNA + FugeneHD will be 25 uL/ug DNA. In this case that is 330 μL of PBS. Then 15 ug of DNA is added to the PBS, followed by 30 μL of FugeneHD. The tube is closed securely, vortexed, and incubated while cells are harvested. P19s at 80-90% confluency are harvested from a single T-75 flask, washed in PBS, spun down, and resuspended in 900 μL of PBS. 300 μL of the cell solution is then aliquotted into an eppendorf tube and spun down. Once the supernatant has been aspirated from the eppendorf tube of P19s, the FugeneHD complexes should have been incubating for about 15 minutes. Once the FugeneHD complexes have incubated at room temperature for at least 15 minutes, the P19 cell pellet in the eppendorf tube is resuspended in Fugene complexes and transferred to an electroporation cuvette.

Amber Stop Codon Suppression Assays

HEK293T cells used for initial testing of the *E. coli* tRNA/RS pair were transfected with FugeneHD as described above. P19s were transfected with the elements of the two UAA incorporation systems using either the electroporation protocol, or the hybrid transfection protocol. Though DNA amounts had been previously optimized for each protocol, various amounts of DNA were tested for each UAA incorporation system to ensure the best triple transfection possible. In each case, media was changed 24 hours after transfection and every 24 hours thereafter. Cells were imaged every 24 hours after transfection using a Nikon Eclipse TE2000-S microscope with a FITC HyQ filter (Chroma, Rockingham, VT). Cells were harvested at either 72 hours or 96 hours post-transfection for fluorescence activated cell sorting (FACS). Briefly, the passaging protocol was followed up to the point of centrifugation. After centrifugation, cells were washed once in PBS and recentrifuged. PBS supernatant was aspirated, and cells were resuspended in 4% paraformaldehyde in PBS (USB Corp. Cleveland, OH). Cells were incubated in paraformaldehyde for 15 minutes, then re-centrifuged. Paraformaldehyde was aspirated, and cells were resuspended in PBS then subjected to FACS. FACS was accomplished using a BD FacsCalibur, and data analysis was done using Cyflogic. For all FACS experiments, dead cells were excluded using forward and side scatter. GFP gating was done by designating the area of the GFP histogram occupied by untransfected cells as "GFP Negative," and all cells exhibiting higher fluorescence than untransfected cells were classified as "GFP Positive."

Immunostaining for Differentiation Assays

To ascertain the effect of the exogenous components of the UAA incorporation system(s), P19s were transfected with the *E. coli* tRNA/RS as per the conditions shown in

Figure 5.1. One condition was also transfected with the non-orthogonal *M.* jannaschii based amber suppressor tRNA. 48 hours after transfection, all conditions were harvested as detailed in the preceding section, and then they were sterile sorted with a BD Biosciences FacsAria to isolate and keep the GFP positive cells. As with previous FACS experiments, dead cells were excluded using forward and side scatter. GFP gating was done by designating the area of the GFP histogram occupied by untransfected cells as "GFP Negative," and all cells exhibiting higher fluorescence than untransfected cells were classified as "GFP Positive." GFP positive cells were collected in sterile, complete media with 20% (v/v) FBS. While sorting was still going on, sorted cells were kept on ice. The entire sorting process for these samples took approximately 4 hours. After sorting, cells were spun down and resuspended in complete media + Penstrep (Gibco, Invitrogen, Carlsbad, CA). Cells were then seeded in sterile, chambered coverglass (Fisher Scientific, Waltham, MA). Media was changed every 24 hours after cells were seeded. 48 hours after sorting, media was aspirated from cells. Each chamber was washed, very gently, with PBS. After aspirating the PBS, 4% paraformaldehyde in PBS was added to each chamber, and cells were incubated in paraformaldehyde for 10 minutes. Paraformaldehyde was then aspirated, and cells were washed twice with PBS. Cells were then permeabilized by incubating in 0.25% Triton-X in PBS for 15 minutes at room temperature. Cells were then washed twice with PBS, followed by blocking with 1% BSA (NEB, Ipswich, MA) in PBS for 1 hour at room temperature.

The following procedure was followed for immunostaining. PBS + BSA was aspirated, and cells were then incubated with primary antibodies for neural cell adhesion molecule (NCAM), α-actinin, and Oct4. Mouse monoclonal anti- α-actinin was purchased from Sigma-Aldrich (St. Louis, MO), rat monoclonal anti-NCAM was purchased from GeneTex (Irvine, CA), and rabbit polyclonal anti-Oct4 was purchased

from Millipore (Billerica, MA). Primary antibodies were diluted in PBS according to the manufacturer's instructions, and cells were incubated with the primary antibody solution overnight at 4°C. Samples were then washed three times with PBS, followed by incubation in the secondary antibody solution for 2 hours at room temperature. The secondary antibody solution was prepared by dissolving secondary antibodies in PBS according to the manufacturer's instructions. All of the following secondary antibodies were purchased from Invitrogen (Carlsbad, CA): AlexaFluor® 594 chicken anti-rat, AlexaFluor® 532 goat anti-rabbit, and AlexaFluor® 405 goat anti-mouse. Secondary antibody solutions were protected from light as much as possible. After incubation, the secondary antibody solution was aspirated, and the samples were washed three times with PBS.

Confocal Microscopy Imaging

After immunostaining, samples were imaged with a Leica SP2 AOBS confocal microscope (Leica Microsystems Inc., Buffalo Grove, IL) in the UT ICMB Core Facility. AlexaFluor® 405 fluorescence was captured with the 405 nm laser at 24% power, and the corresponding photomultiplier (PMT) set to 493 volts. AlexaFluor® 532 fluorescence was captured with the 543 nm laser at 25% power, and the corresponding photomultiplier (PMT) set to 597 volts. AlexaFluor® 594 fluorescence was captured with the 594 nm laser at 24% power, and the corresponding photomultiplier (PMT) set to 570 volts. All images were averaged over eight frame captures to reduce noise, and all images were captured with a 40x oil objective during one microscope session. Fluorescence images were captured sequentially to minimize fluorophore cross-talk.

RESULTS AND DISCUSSION

Amber Stop Codon Suppression Using the *E. coli* System

To confirm the functionality of the *E. coli* tRNA and RS, they were first used in HEK293T cells. Figure 5.1 demonstrates the excellent functionality of the *E. coli* tRNA/RS pair.

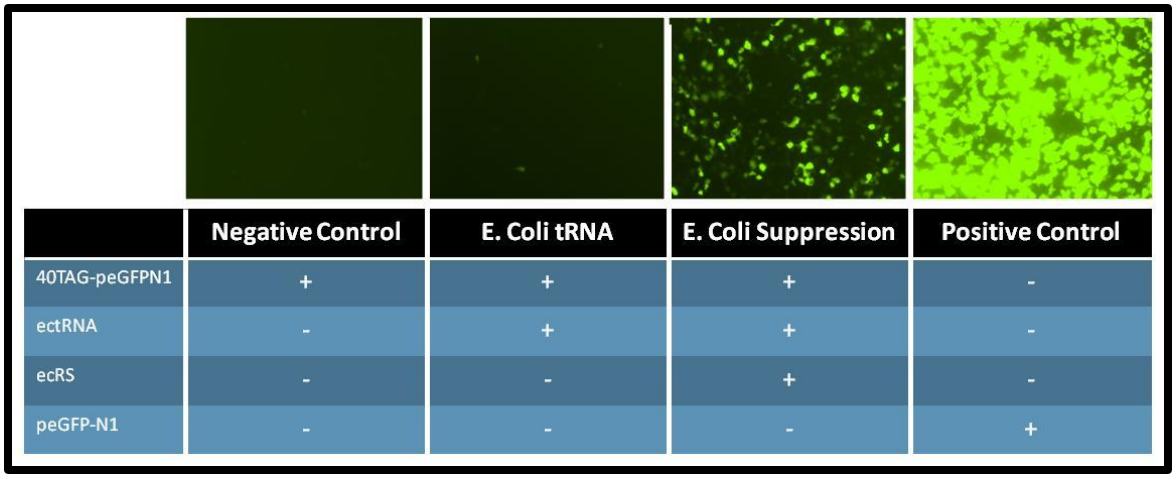


Figure 5.1: Fluorescent images demonstrating functionality of *E. coli* based UAA Incorporation System in HEK293T cells transfected with the DNA constructs below each image, as described on page 109. All images were captured 48 hours after transfection with gain set to 1 and exposure time set to 1 second.

Unsurprisingly, the *E. coli* tRNA condition exhibits some fluorescence. Since FACS was not done on these cells, it is difficult to quantify the level of background created by the *E. coli* tRNA. Even without the quantitative data, this particular experiment demonstrates that exogenous, bacterial tRNA is recognized by some endogenous RSs in HEK293T cells, once again pointing to the hitherto underestimated complexity of tRNA identity determinants. Despite the imperfect orthogonality of the *E. coli* tRNA, when in the presence of *E. coli* RS^{tyr}, this tRNA is clearly charged and able to suppress amber stop codons, as evidenced by the suppression image in Figure 5.1. In fact, the suppression

condition appears to obtain between 30% and 50% of the eGFP expression achieved by transfection with peGFP-N1.

Without directly comparing this data to the previously developed *M. jannaschii* tRNA/RS pair, it could have been difficult to gauge whether or not to pursue one or both UAA incorporation systems. Thus, another experiment was done so that the cells could be assessed via FACS. In order to compare the *E. coli* and *M. jannaschii* systems directly, the respective experiments were normalized to the fluorescence level measured from cells transfected only with 40TAG-peGFPN1.

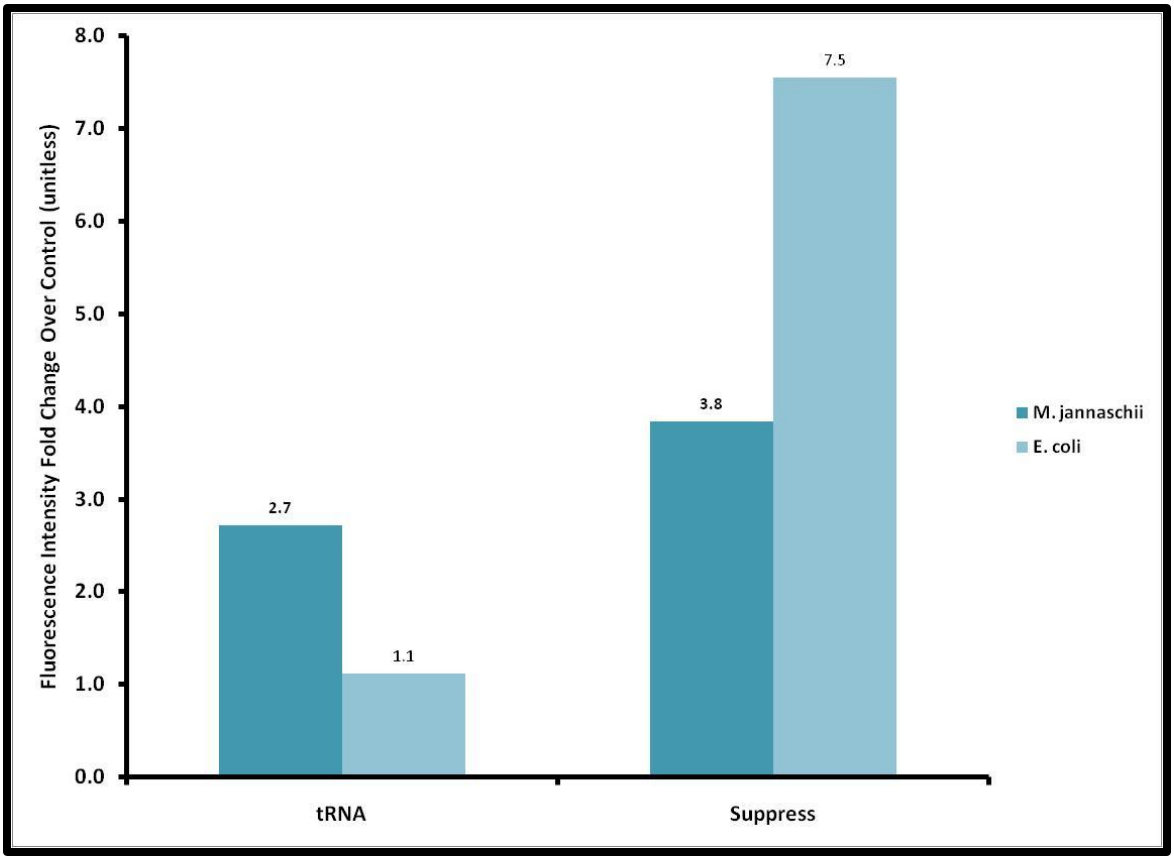


Figure 5.2: Fluorescence intensity fold change as determined by analysis of FACS data for HEK293T cells transfected with the *E. coli* or *M. jannaschii* based UAA incorporation system components. Cells for each sample were harvested 48 hours post transfection, and fold change is defined as (sample fluorescence)/(40TAG-peGFPN1 fluorescence).

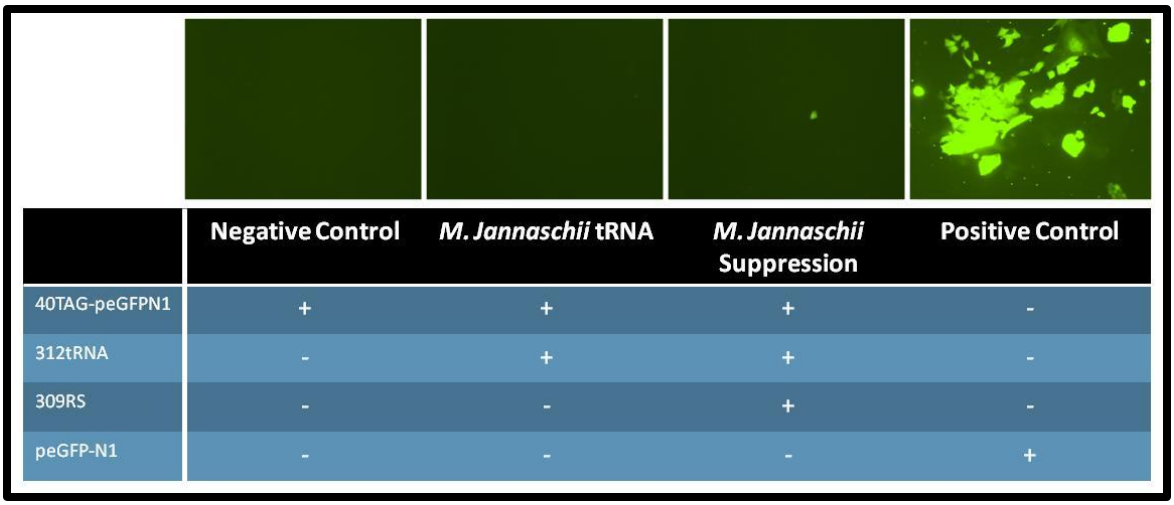
Figure 5.2 demonstrates the clear differences between the *E. coli* tRNA/RS pair and the previously developed *M. jannaschii* based tRNA/RS pair. While the $tRNA^{mjtyr}$ exhibits almost 3 times the fluorescence of the 40TAG-peGFPN1 control, the $tRNA^{ectyr}$ exhibits almost exactly the same fluorescence as that control. In contrast, the *E. coli* pair exhibits almost eight times the fluorescence of the control, whereas the *M. jannaschii* based pair achieves just less than four times the control. Even more stark is the difference between the respective tRNA background and suppression efficiencies of the two systems. The *E. coli* suppression condition achieves almost eight times the fluorescence caused by tRNA background, whereas the *M. jannaschii* suppression condition achieves less than twice the fluorescence caused by its tRNA background. Clearly then, the *E. coli* pair is a better choice, except for the lack of readily available synthetases. While the amino acid sequences for *E. coli* based synthetases are known, none of those synthetases are preexisting in our lab. Thus, based upon the information presented in Figure 5.1and Figure 5.2, it seemed the most expedient route was to pursue the use of both systems. Thus, in the event that the *M. jannaschii* pair was successful in P19s, we could move immediately to UAA incorporation in those cells. If not, we would be working on making the UAA specific synthetases from the *E. coli* synthetase while testing the two systems in P19s.

There is one more key point that must be made about the preceding data. In HEK293T cells, when coexpressing an amber suppressor tRNA based upon a wildtype tRNA and a wildtype tRNA synthetase, those two elements are unable to achieve the expression efficiency of endogenous tRNAs. I.e. endogenous tRNAs are able to produce far more wildtype GFP from the peGFP-N1 plasmid than the single exogenous amber suppressor tRNA and cognate synthetase can produce from a nearly identical plasmid with a single point mutation. This is an important limitation of a UAA incorporation

system as it highlights the difficulty of introducing foreign DNA or proteins to a cell or organism and coaxing to behave as endogenous materials, and it also points to the fact that alterations to tRNAs and synthetases logically decrease the efficiency of their interactions. Regardless, both the *M. jannaschii* based UAA incorporation system and the *E. coli* system functioned reasonably well in HEK293T cells, so they were both tested in P19s.

Amber Stop Codon Suppression in P19s with the *M. jannaschii* Based System

P19s were electroporated with the components of the *M. jannaschii* based UAA incorporation system, 312tRNA, 309RS, and 40TAG-peGFPN1. As with all other transfection based experiments, GFP expression was assessed via fluorescence microscopy at 24 and 48 hours post transfection, and cells were harvested and subjected to FACS at 48 hours post transfection. As shown in Figure 5.3, the *M. jannaschii* based system performs very poorly in P19s.



	Negative Control	*M. Jannaschii* tRNA	*M. Jannaschii* Suppression	Positive Control
40TAG-peGFPN1	+	+	+	-
312tRNA	-	+	+	-
309RS	-	-	+	-
peGFP-N1	-	-	-	+

Figure 5.3: Fluorescent images demonstrating functionality of *M. jannaschii* based UAA Incorporation System in P19s electroporated with 100 µg each of the DNA constructs below each image. All images were captured 48 hours after transfection with gain set to 1 and exposure time set to 1 second. In each image, cells are approximately 50% confluent.

Only one fluorescent cell is evident in the suppression condition shown in Figure 5.3. This result could have been indicative of low transfection efficiencies, but the positive control effectively demonstrates that transfection efficiency was not a problem. In fact, the cells in each condition are approximately 50% confluent, which means that, based upon the images in Figure 5.3, the transfection achieved higher than 70% efficiency. If the *M. jannaschii* tRNA and synthetase pair function in P19s as they do in HEK293T cells, then one could reasonably expect fluorescence exhibited by suppression conditions to proportionately decrease with overall transfection efficiency. In Chapter Four, data was presented indicating that HEK293T cells transfected with the components of the *M. jannaschii* tRNA/RS pair achieve approximately 27% of the fluorescence exhibited by positive controls. Thus, in the case of P19s, we might expect 27% of 70% (fluorescent cells in positive control) of the total cells to express enough GFP to be detected by fluorescent microscopy. That means that if the system works in P19s as it does in HEK293T cells, we would see about 19% of the total cells in a given image as fluorescent. If there were only five cells in the suppression condition image, that could indicate functionality of the *M. jannaschii* pair in P19s. Unfortunately there are hundreds of cells in that image, and yet only one has expressed a detectable amount of GFP.

We posited that perhaps P19s did not respond to the promoters in the system, those for protein as well as for tRNA. While this may have been the case for the tRNA plasmid, it probably was not the case for the RS or 40TAG-peGFPN1 plasmids, as they have either identical or almost identical promoters to the peGFP-N1 plasmid which clearly produced functional GFP protein in this cell line. Thus, we tested the tRNA transcription by creating a non-orthogonal tRNA, specifically a wildtype *M. jannaschii* $tRNA^{tyr}$ with an amber anticodon. By transfecting cells with this nonorthogonal amber

suppressor tRNA and 40TAG-peGFPN1, GFP fluorescence could be used as an indicator of successful tRNA synthesis in P19s.

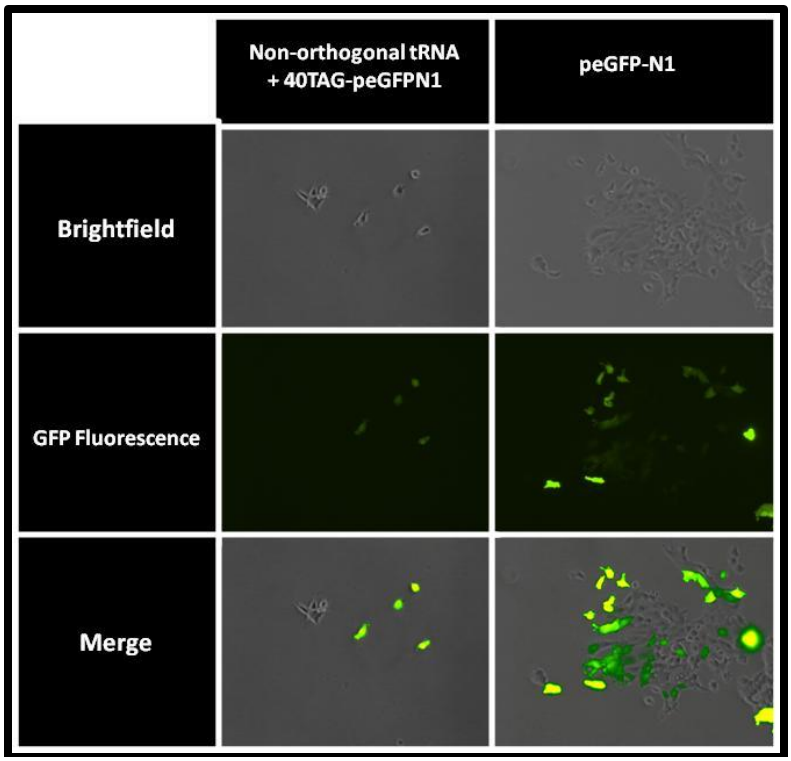


Figure 5.4: Fluorescent images demonstrating non-orthogonal tRNA synthesis by P19 cells transfected with 50 µg each of non-orthogonal tRNA and 40TAG-peGFPN1, or in the case of the peGFP-N1 column, 100 µg of peGFP-N1. Fluorescence images were captured 48 hours after transfection with gain set to 1 and exposure time set to 1 second. The contrast in the merged images FITC channel has been increased by 88% to facilitate visualization.

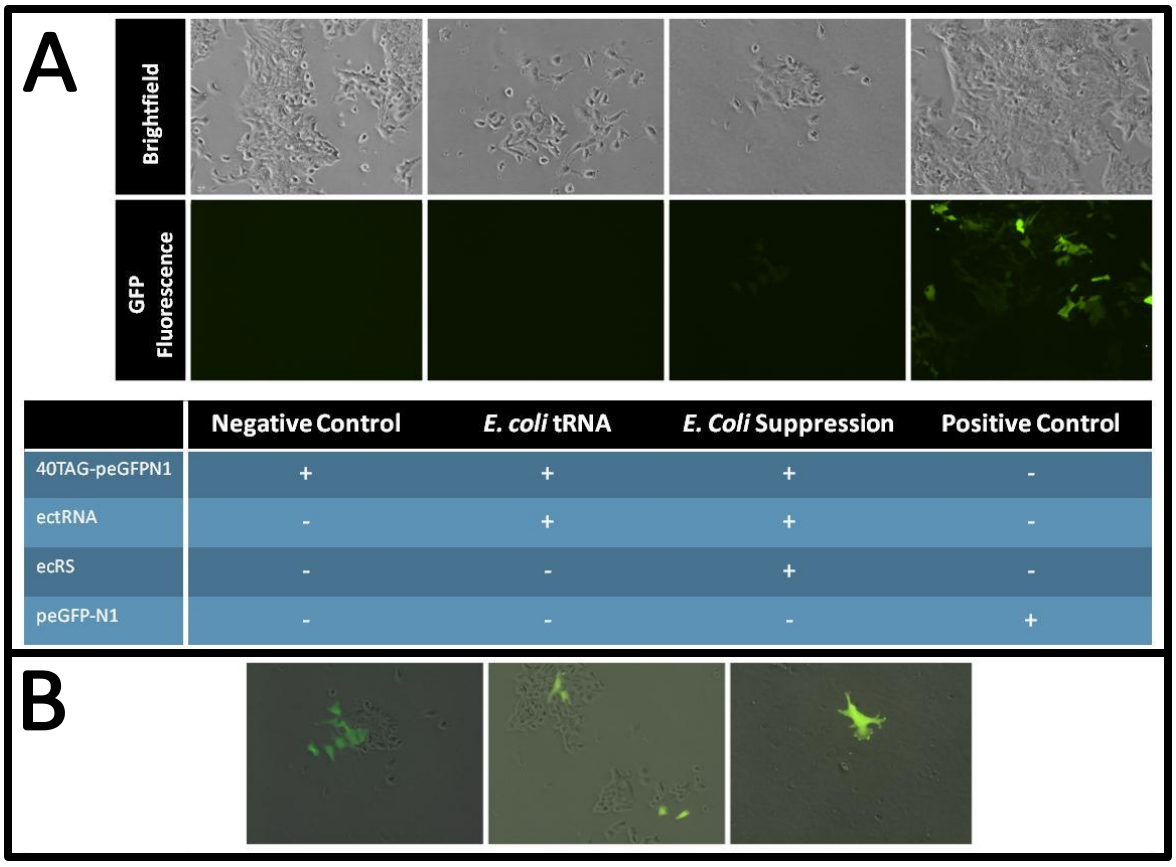
As evidenced by the Non-orthogonal tRNA column in Figure 5.4, P19s do synthesize tRNA when transfected with a plasmid bearing the human H1 promoter upstream of a tRNA gene. In comparison to the peGFP-N1 positive control, the non-orthogonal tRNA + 40TAG-peGFPN1 condition appears to achieve approximately the same level of fluorescence as the positive control. However, cell viability in the non-orthogonal tRNA condition does appear to be compromised. This may indicate a

hitherto unexpected complication of expressing amber suppressor tRNA in pluripotent cell lines, namely, that overexpression of amber suppressor tRNA compromises cell viability. Though we have not seen this same phenomenon in HEK293T cells, it is conceivable that amber suppression releases fusion proteins or transcription factors which can only effectively activate apoptotic pathways in pluripotent cell lines.

Successful amber suppression via tRNA recognition by endogenous P19 synthetases indicates that there are fundamental differences between the protein translational machinery components in P19s and HEK293Ts. Since exogenous tRNA can clearly be produced and charged with amino acids, the problem cannot be simply a complete lack of tRNA transcription. Though it may be that P19s do not complete post-transcriptional modifications to tRNA fragments in the same manner as HEK293Ts, rendering tRNAs produced within these cells compatible only with P19 synthetases or only moderately compatible with exogenous synthetases. To fully understand the mechanism behind the stark difference in amber suppression with a non-orthogonal tRNA and endogenous synthetase versus amber suppression using the previously developed *M. jannaschii* based orthogonal tRNA/RS pair would require a great deal of further study. Since it became clear that P19s could produce functional exogenous tRNA, we chose to move forward and test the *E. coli* tRNA/RS pair since they performed better in HEKs (better meaning achieved higher fluorescence) than the *M. jannaschii* pair.

Amber Stop Codon Suppression in P19s with the *E. coli* Based System

P19s were electroporated with the *E. coli* amber suppressor tRNA and wildtype synthetase (ectRNA and ecRS) to determine their functionality in P19s. GFP expression was assessed using fluorescence microscopy and FACS.



	Negative Control	*E. coli* tRNA	*E. Coli* Suppression	Positive Control
40TAG-peGFPN1	+	+	+	-
ectRNA	-	+	+	-
ecRS	-	-	+	-
peGFP-N1	-	-	-	+

Figure 5.5: A) Fluorescent images demonstrating amber stop codon suppression by P19s transfected with the *E. coli* based plasmids listed. B) Multiple merged brightfield+fluorescence images taken from the sample shown in the *E. coli* suppression condition of part A. All images were captured 48 hours after transfection with gain set to 1 and exposure time set to 1 second.

As shown in Figure 5.5, above, the *E. coli* tRNA/RS pair does successfully suppress amber stop codons in P19s, with a much greater degree of success than the *M. jannaschii* pair. Without doing any further experiments, these images demonstrate squarely that the amber suppressor tRNA and RS from *E. coli* can be functionally expressed in P19s, and this tRNA/RS pair can work together to suppress amber stop codons. As evidenced by the brightfield images in Figure 5.5 A, there is a negative effect on cell viability which correlates with exogenous amber suppressor tRNA transfection.

Though some studies have indicated increased cell mortality rates when mammalian cells are transfected with high levels of DNA, the positive control in this case rules out the role of 'DNA poisoning' in the increased cell mortality rates. The suppression condition and the positive control condition each used the same total amount of DNA, while the 40TAG-peGFPN1, *E. coli* tRNA, and *E. coli* suppression conditions each utilized the same amount of DNA per construct. Since this phenomenon of decreased cell viability in tRNA containing transfections is consistent across experiments with *M. jannaschii* orthogonal and non-orthogonal tRNA and *E. coli* tRNA, there likely is some mechanism for which tRNA synthesis is serving as the impetus. Regardless, the number of surviving cells as well as the efficiency of suppression appears high enough to warrant continued pursuit of the *E. coli* based system.

To that end, P19 cells transfected with the components of the *E. coli* based UAA incorporation system, exactly as shown in Figure 5.5, were harvested 48 hours post-transfection and subjected to FACS. Figure 5.6 below shows the results of the first of many identical experiments. There are several important pieces of information in this figure, the first of which is that the *E. coli* suppression condition exhibits 27 times the fluorescence intensity of the 40TAG-peGFPN1 condition and slightly less than 4 times the fluorescence intensity of the tRNA control condition. The first number is of little consequence, but quadrupling the fluorescence signal from the tRNA control condition is a substantial accomplishment, especially in light of the fact that this increase over control is fully half of that seen in HEK293T cells. Since people have successfully used systems like these in HEK293T cells, this result is encouraging, and demanded further pursuit of the *E. coli* based system. Additionally, the suppression condition achieved approximately 10% of the fluorescence level of the positive control. While this number may seem low, we consider it an accomplishment. For our system to produce full length

GFP requires the functional expression of both tRNA and synthetase prior to expression of GFP, while the positive control can express GFP using the preexisting endogenous protein translational machinery. Thus, we chose to further investigate the effects of these exogenous materials on the differentiation state of P19s.

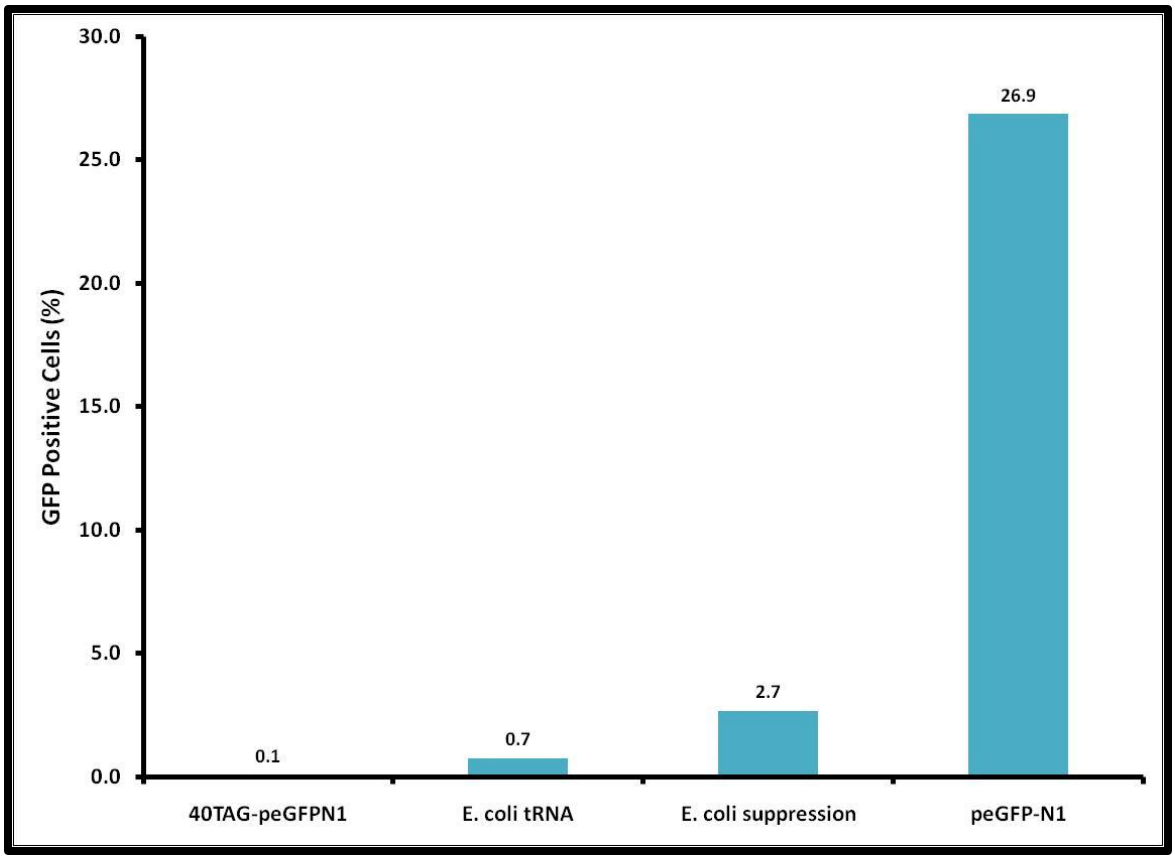


Figure 5.6: Fluorescence intensity of P19s expressing components of the *E.* coli based UAA incorporation system, as determined by analysis of FACS.

Differentiation Assays

To assess the effect of the *E. coli* based UAA incorporation system on the differentiation state of P19s, cells were electroporated with the components of the system, and then they were sterile sorted by FACS. Cells expressing GFP were grown for

another day on chambered coverslips. They were then fixed, permeabilized, and immunostained for differentiation and pluripotency markers. For conditions with very small or nonexistent populations of GFP expressing cells, cells not expressing GFP were used for staining. For instance, the cells transfected with ectRNA and 40TAG-peGFPN1 were not expected to fluoresce upon successful expression of the protein/tRNA encoded by their plasmids. Thus, testing the non-fluorescent cells in those conditions was the only real and logical option.

As shown in Figure 5.7, cells in all conditions stain positively for Oct4, which is a marker of pluripotency in many stem cell lines, including P19s (4). Cells were not counterstained with any nuclear or membrane permeant stains, because all available stains would have interfered with either the secondary antibody fluorescence or the GFP fluorescence. Since we wanted to investigate the relationship between cells successfully expressing the exogenous tRNA and/or RS, the GFP signal needed to be differentiable from all others. Thus, it is difficult to discern cell boundaries in these images. Furthermore, brightfield images were unable to effectively capture cell boundaries either. Thus, only GFP and the AlexaFluor® dyes are visible in these images. Nevertheless, they provide a wealth of information. From the GFP column in Figure 5.7, one can ascertain that it was impossible to find cells exhibiting GFP fluorescence in either the *E. coli* tRNA condition or the No DNA condition, but all three other conditions did contain GFP expressing cells. This further corroborates the previous fluorescence images and FACS data indicating that the non-orthogonal amber suppressor tRNA can be synthesized and used by endogenous synthetases to suppress the TAG codon in 40TAG-peGFPN1, and the *E. coli* derived tRNA/RS pair can achieve the same.

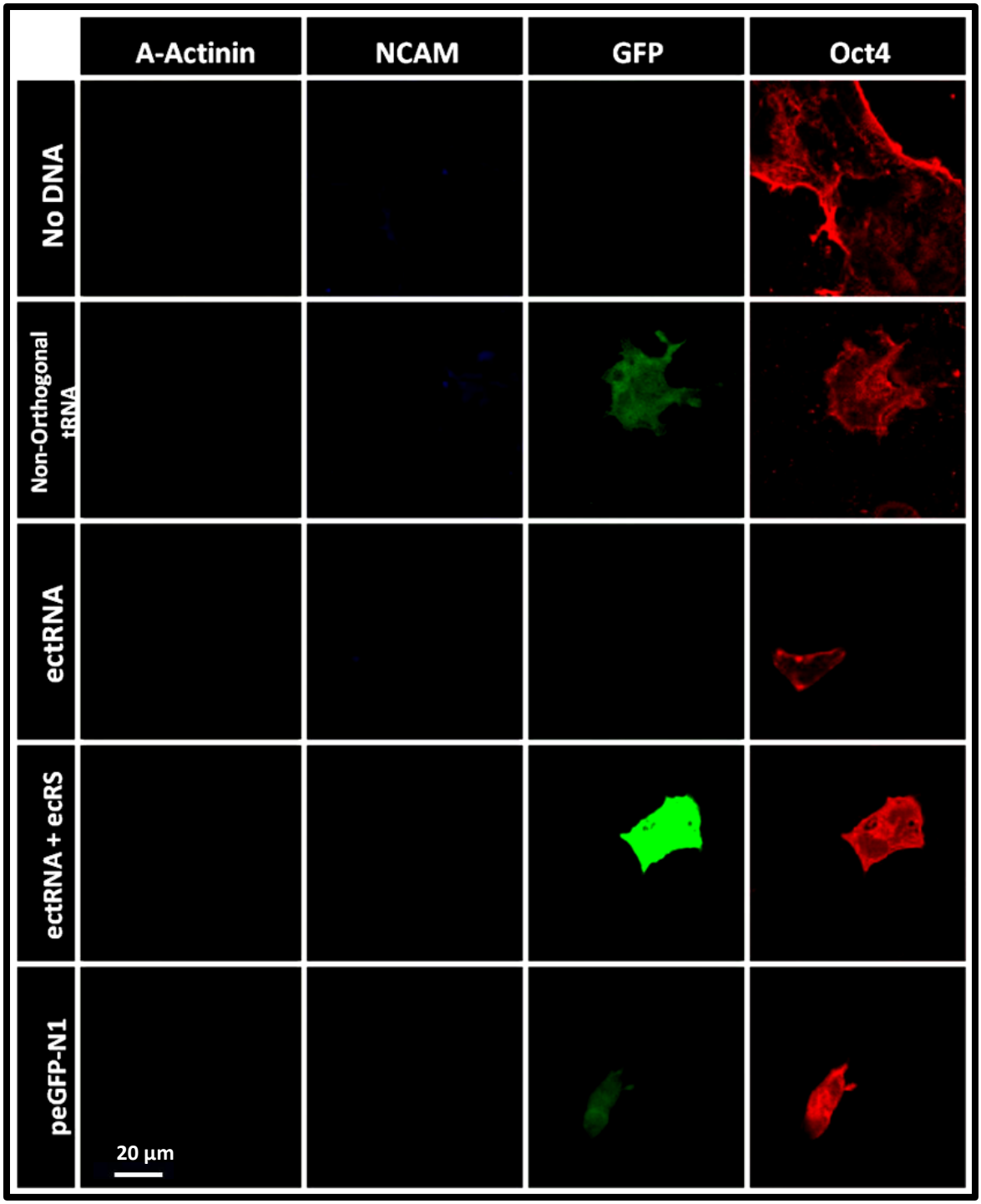


Figure 5.7: Pseudocolored confocal images of immunostained P19s transfected with exogenous tRNA(s) and synthetases, or appropriate controls, as noted on the left hand column. All images were captured with the same microscope settings from the same experiment.

Though there is some very light fluorescence in the range of the NCAM antibody dye in the non-orthogonal tRNA condition, that is most likely not indicative of differentiation as the same light staining appears in the negative control (No DNA). There was no discernible staining due to α-actinin, indicating that these cells are most likely not differentiation down any skeletal muscle pathways. Overall, there appears to be no difference in staining in the control conditions as compared to the exogenous tRNA/RS conditions. This indicates that P19s can produce and use exogenous tRNAs and RSs to suppress stop codons, and that process most likely is not inducing differentiation.

Some may argue that 96 hours post-transfection is too soon to assess differentiation. However, in terms of P19 differentiation, it is well established in the literature that P19s exhibit discernible and often irreversible differentiation as early as 24 hours after induction of differentiation (4, 5). Thus, assessing differentiation 96 hours post-transfection, as far as the relevant literature is concerned, should be a fairly reliable method. However, since these cells have to actually express several components before they can have an effect on differentiation, if they have an effect, it could be that differentiation markers would be slightly delayed. While further study could confirm this, our controls appeared no different than the tRNA/RS conditions, so we felt it unnecessary to further investigate their ability to differentiate P19s at this time. It is possible that certain UAAs could exert some effect on differentiation state, at which point it will be logical to pursue another differentiation study.

MORPHOLOGICAL AND VIABILITY CHANGES RESULTANT FROM AMBER SUPPRESSION

In both the *E. coli* and *M. jannaschii* amber suppression conditions shown in Figures 5.3 - 5.7, there are dramatically fewer surviving cells than in the negative and positive control conditions. While no cell proliferation assays were performed, the

images clearly demonstrate that amber suppression has deleterious effects on P19s. While these effects did not prevent completion of the experiments detailed in this dissertation, they do give some pause as to the future of projects along these lines. Amber stop codons in the 40TAG-peGFPN1 gene are suppressed by these systems, but so too are amber stop codons within the P19 genome, resulting in protein polymers which exceed the length of normal proteins. This situation logically results in proteins which do not fold properly or in fact have completely different functions than intended, resulting in decreased cell viability. Without further study, it is impossible to determine exactly how much amber suppression P19s can accomplish without compromising cell viability to the degree seen thus far in this project. Until such work is done, the usefulness and feasibility of UAA incorporation in response to amber stop codon suppression in P19s will remain unknown.

CONCLUSION

In the preceding pages, data has been presented indicating each of the following:

1. *E. coli* based amber suppressor tRNA and wildtype RS can achieve amber stop codon suppression in P19s.
2. *E. coli* based amber suppressor tRNA and wildtype RS are substantially more efficient in P19s than the *M. jannaschii* derived pair.
3. Expression of exogenous tRNAs and/or synthetases does not induce discernible differentiation by 4 days post-transfection.
4. Amber suppression appreciably compromises P19 cell viability.

From this information we can draw a few more conclusions. First of all, the *E. coli* system is certainly worth pursuing, as it works so well and produces enough UAA

incorporation system expressing cells that a population of them can be isolated for further testing by FACS. Implementation of the *E. coli* based UAA incorporation system in mammalian cells could swiftly enable a variety of UAA based studies in the P19 or other pluripotent cell environments, especially if that work involves collaborations with other investigators with experience using the *E. coli* synthetases. Second, exogenous synthetases and tRNAs do not appear to be inducing differentiation, meaning that UAA incorporation systems could be employed in the study of normal differentiation processes. Finally, since it would take time to develop the UAA specific synthetases for *E. coli* tRNAs, and there were fluorescent cells in the *M. jannaschii* system, we chose to pursue both the *E. coli* system and the *M. jannaschii* system concurrently. The reason for this stems from the experience of several lab members with UAA incorporation in other organisms. It has been their experience that UAA specific synthetases often perform better than expected based upon those evolved simply for amber anticodon permissivity. Thus, we chose to pursue actual incorporation of UAAs (as opposed to the amber suppression studies presented in the preceding pages) using the *M. jannaschii* system to see if a UAA specific synthetase could charge the amber suppressor tRNA more efficiently than the tyrosyl-RS. Thus, UAA incorporation using the *M. jannaschii* based synthetases, and creation of *E. coli* based UAA specific synthetases will be described in Chapter Six.

REFERENCES

1. LaFemina R, Hayward GS. Constitutive and retinoic acid-inducible expression of cytomegalovirus immediate-early genes in human teratocarcinoma cells. J Virol. 1986;58(2):434-40.
2. Gonczol E, Andrews PW, Plotkin SA. Cytomegalovirus replicates in differentiated but not in undifferentiated human embryonal carcinoma cells. Science. 1984;224(4645):159-61.
3. Thibodeaux G, Liang X, Moncivais K, Umeda A, Singer O, Alfonta L, et al. Transforming a Pair of Orthogonal tRNA-aminoacyl-tRNA Synthetase from Archaea to Function in Mammalian Cells. PLOS ONE. 2010;5(6):-.
4. Xie Z, Tan G, Ding M, Dong D, Chen T, Meng X, et al. Foxm1 transcription factor is required for maintenance of pluripotency of P19 embryonal carcinoma cells. Nucleic Acids Res. 2010;38(22):8027-38.
5. Tan Y, Xie Z, Ding M, Wang Z, Yu Q, Meng L, et al. Increased levels of FoxA1 transcription factor in pluripotent P19 embryonal carcinoma cells stimulate neural differentiation. Stem Cells Dev. 2010;19(9):1365-74.

CHAPTER SIX

Incorporation of Unnatural Amino Acids into Proteins in P19 Embryonal Carcinoma Cells

INTRODUCTION

Unnatural amino acids (UAAs) have been used in a variety of applications including, but not limited to, protein-tagging, site-specific post-translational modification mimicking, and protein structure/function assays (1-5). The preceding chapters have detailed the development and testing of UAA incorporation systems for use in all mammalian cells, but in the scope of this dissertation they have been applied to P19 embryonal carcinoma cells (P19s) with the intent that they will eventually be deployed in other pluripotent or stem cell lines. Chapter Five described the use of a *Methanocaldococcus jannaschii (M. jannaschii)* based amber suppressor tRNA and aminoacyl tRNA-synthetase (RS) pair and an *Eschericia coli* (*E. coli)* based amber suppressor tRNA/RS pair for amber stop codon suppression in both HEK293T cells and P19 embryonal carcinoma cells. Based upon the success of the *E. coli* tRNA/RS pair, we chose to continue investigation along those lines by developing a UAA specific RS from the *E. coli* RS, after which we could use that RS to incorporate UAA(s) into proteins in P19s. Since the *M. jannaschii* tRNA/RS pair functioned, though at a low level, we chose to move forward and test the previously developed *M. jannaschii* derived UAA specific synthetases in P19s, because previous experience has shown that some UAA specific synthetases are actually more efficient at charging amber suppressor tRNAs than their canonical amino acid charging counterparts (personal communications).

Overall Experimental Design

To accomplish the goal of incorporating UAAs into proteins in P19s, the following experimental goals were set:

1. Create and test an *E. coli* derived RS specific for the incorporation of the UAA m-acetyl-L-phenylalanine (pketo) in P19s.
2. Test the previously developed *M. jannaschii* tRNA/RS pair specific for pketo incorporation.

The first goal will involve site-directed mutagenesis of the RS developed in Chapter Five. The second will require simple application of the previously developed *M. jannaschii* tRNA with a cognate RS specific for incorporation of pketo. Hydrazide staining will be used to confirm the incorporation of pketo into proteins in P19s.

MATERIALS AND METHODS

Routine Cell Culture

P19 embryonal carcinoma cells were obtained from American Type Culture Collection (ATCC). Cells were routinely maintained in Alpha-Mem (HyClone, Logan, UT) with 10% (v/v) fetal bovine serum (Atlanta Biologicals, Lawrenceville, GA) and nonessential amino acids (Sigma-Aldrich, St. Louis, MO) in a humidified incubator at 37° Celsius with 5% CO_2 atmosphere. Cells were passaged every 1-3 days at a ratio between 1:4 and 1:20. For passaging, cells were rinsed in phosphate buffered saline (PBS) (Sigma-Aldrich, St. Louis, MO), incubated with 0.5% trypsin (Gibco, Invitrogen, Carlsbad, CA) for 3-5 minutes, and centrifuged to pellet. Cells were then resuspended in complete media and seeded into clean flasks with fresh media.

DNA Constructs

The plasmids encoding 40TAG-eGFP (40TAG-peGFPN1 plasmid) and *M. jannaschii* based amber suppressor tRNA (312tRNA plasmid) and chimeric tyrosyl RS (309RS plasmid) were described in the preceding chapters. The pketo specific chimeric RS (plasmid pketoMJRS) was a gift from a colleague.

Hybrid FugeneHD + Electroporation Transfection

Fugene HD was purchased from Roche (Switzerland). A modified version of Roche's FugeneHD protocol was used for the hybrid transfection procedure. Instead of 50 µL of diluent per µg of DNA, 25 µL of diluent was used per µg of DNA. Assuming the concentration of DNA were 1 ug/µL, the following protocol would be used. First, PBS is added to an eppendorf tube such that the final volume of PBS + DNA + FugeneHD will be 25 uL/ug DNA. In this case that is 330 µL of PBS. Then 15 ug of DNA is added to the PBS, followed by 30 µL of FugeneHD. The tube is closed securely, vortexed, and incubated while cells are harvested. P19s at 80-90% confluency are harvested from a single T-75 flask, washed in PBS, spun down, and resuspended in 900 µL of PBS. 300 µL of the cell solution is then aliquotted into an eppendorf tube and spun down. Once the supernatant has been aspirated from the eppendorf tube of P19s, the FugeneHD complexes should have been incubating for about 15 minutes. Once the FugeneHD complexes have incubated at room temperature for at least 15 minutes, the P19 cell pellet in the eppendorf tube is resuspended in Fugene complexes and transferred to an electroporation cuvette.

Unnatural Amino Acid Incorporation Assays

Transfection and Fluorescence Microscopy Incorporation Assessment

P19s were transfected using the above hybrid FugeneHD + electroporation method with 40TAG-peGFPN1, 312tRNA, and pketoMJRS. Media was changed every 24 hours after transfection, and fluorescent images were captured daily immediately prior to media changes. Fluorescent images were captured using a Nikon Eclipse TE2000-S microscope with a FITC HyQ filter (Chroma, Rockingham, VT). 48 hours after transfection, cells were harvested and split into two samples per experimental condition, one for fluorescence activated cell sorting (FACS), and one for western blot/hydrazide modification. For FACS sorting, the passaging protocol was followed up to the point of centrifugation. After centrifugation, cells were washed once in PBS and recentrifuged. PBS supernatant was aspirated, and cells were resuspended in 4% paraformalehyde in PBS (USB Corp. Cleveland, OH). Cells were incubated in paraformaldehyde for 15 minutes, then re-centrifuged. Paraformaldehyde was aspirated, and cells were resuspeded in PBS then subjected to FACS. FACS was accomplished using a BD FacsCalibur, and data analysis was done using Cyflogic. For all FACS experiments, dead cells were excluded using forward and side scatter. GFP gating was done by designating the area of the GFP histogram occupied by untransfected cells as “GFP Negative,” and all cells exhibiting higher fluorescence than untransfected cells were classified as “GFP Positive.”

Western Blot and Hydrazide Modification

For western blot/hydrazide staining, cells were pelleted after harvest and resuspended in M-PER mammalian protein extraction reagent (Thermo-Scientific, Waltham, MA) with 1X Complete mini protease inhibitor (Roche, Switzerland). Both M-

Per and the complete mini protease inhibitor were used according to the manufacturer's instructions. Lysis mixtures were incubated with gentle shaking at room temperature for 10 minutes followed by centrifugation at 14,000xg for 15 minutes to pellet cell debris. At this point, all samples not intended for hydrazide modification were refrigerated at 4°C until gel electrophoresis could be completed.

For hydrazide modification intended samples, supernatant was removed from centrifuged samples and dialyzed into PBS overnight. After dialysis, samples were hydrazide modified as follows. AlexaFluor® 633 hydrazide was purchased from Invitrogen (Carlsbad, CA) and dissolved in 200mM KCl. Dialyzed samples were removed from dialysis membranes and placed in microcentrifuge tubes. AlexaFluor® 633 hydrazide solution was then added to microcentrifuge tubes, and the samples were incubated at 4°C for 16 hours. After incubation with hydrazide, samples were dialyzed into ammonium bicarbonate buffer (pH 7.2) for 24 hours to remove unreacted hydrazide. Samples were then lyophilized overnight and resuspended in PBS. After resuspension, samples were subjected to SDS-polyacrylamide gel electrophoresis (SDS-PAGE). Total protein concentrations were the same in each lane, and non-hydrazide reacted samples were subjected to SDS-PAGE simultaneously with hydrazide reacted samples.

Protein was blotted onto nitrocellulose using an iBlot (Invitrogen, Carlsbad, CA). Membranes were blocked in 1% dry milk in TBST followed by probing with Clontech Living Colors anti-eGFP primary antibody (Clontech, Mountainview, CA). The secondary antibody, goat-anti-mouse horseradish peroxidase (HRP) conjugate was purchased from BioRad (Hercules, CA). GFP was detected using Amersham ECL Plus detection reagents (GE Healthcare, Piscataway, NJ).

Creation of pketo Specific E. coli RS

Site-directed mutagenesis was used to introduce the necessary mutations to the previously developed ecRS plasmid.

RESULTS AND DISCUSSION

Pketo Incorporation with the *M. jannaschii* tRNA/RS Pair

Figure 6.1 below shows the FACS results of the first two pketo incorporation.

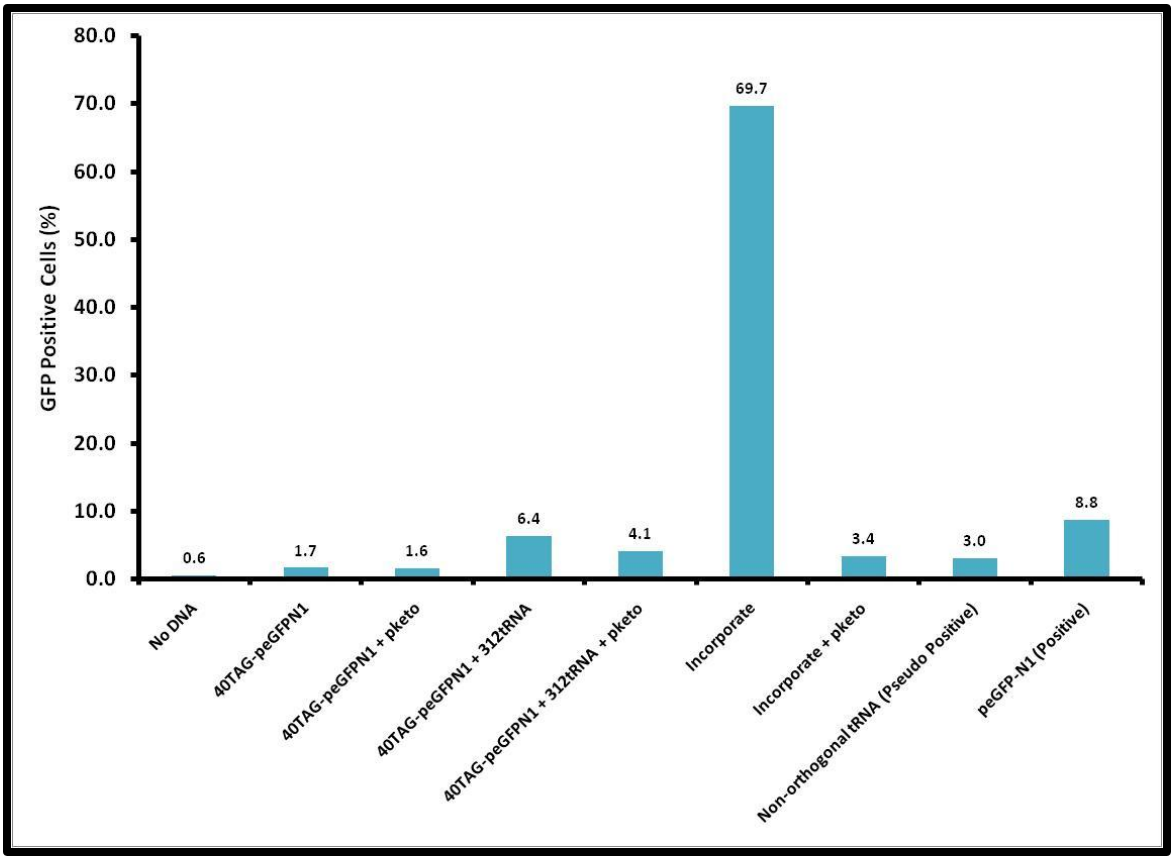


Figure 6.1: Fluorescence intensity of P19s expressing components of the *M. jannaschii* based pketo incorporation system, as determined by analysis of FACS. Only columns denoted with '+ pketo' were grown in the presence of pketo.

experiments. From this data (Figure 6.1), it appears that addition of pketo to conditions with exogenous tRNA dramatically decreases the fluorescence of these cells. Furthermore, as evidenced by the almost unbelievably high percentage of fluorescent cells in the Incorporate condition (without pketo), the pketo RS seems to be highly capable of charging the amber suppressor tRNA with canonical amino acids. In fact, in all previous experiments, not even positive controls (except for those used during single plasmid optimization experiments detailed in Chapter Three) have achieved GFP expression this high. Thus, armed with data indicating that the pketo specific RS was not specific for pketo, and that addition of pketo to cells somehow decreased the detectable levels of GFP, we came to the only logical conclusion: the samples had somehow been mislabeled. That must be followed by the assertion that two separate experiments were FACS sorted at the same time, and all samples are labeled meticulously. Thus, the likelihood of this type of mistake was almost 0. Nonetheless, it was more likely, scientifically speaking, than the pketo RS functioning almost completely incorrectly, and the pketo UAA inexplicably decreasing overall fluorescence of treated samples. Therefore, the samples were run again, and labeling of tubes was triple checked prior to transporting them to the FACS room.

The second measurement of carefully labeled samples revealed that the first round of FACS samples were labeled correctly. Again, the conditions with pketo in the media exhibited substantially lower fluorescence than their corresponding non-pketo conditions, and the incorporation condition without pketo neatly outpaced the fluorescence of the positive control. Of note is the contradictory nature of this data in comparison to the fluorescence images. In the images, the most fluorescent condition is the positive control, followed by the non-orthogonal tRNA condition and then the incorporation with and without pketo. Since FACS is more sensitive than fluorescence

microscopy, it is possible that it is legitimately detecting GFP expression levels which simply cannot be appreciated in the images. After discussing this with several colleagues it was determined that multiple experiments had been completed using this same

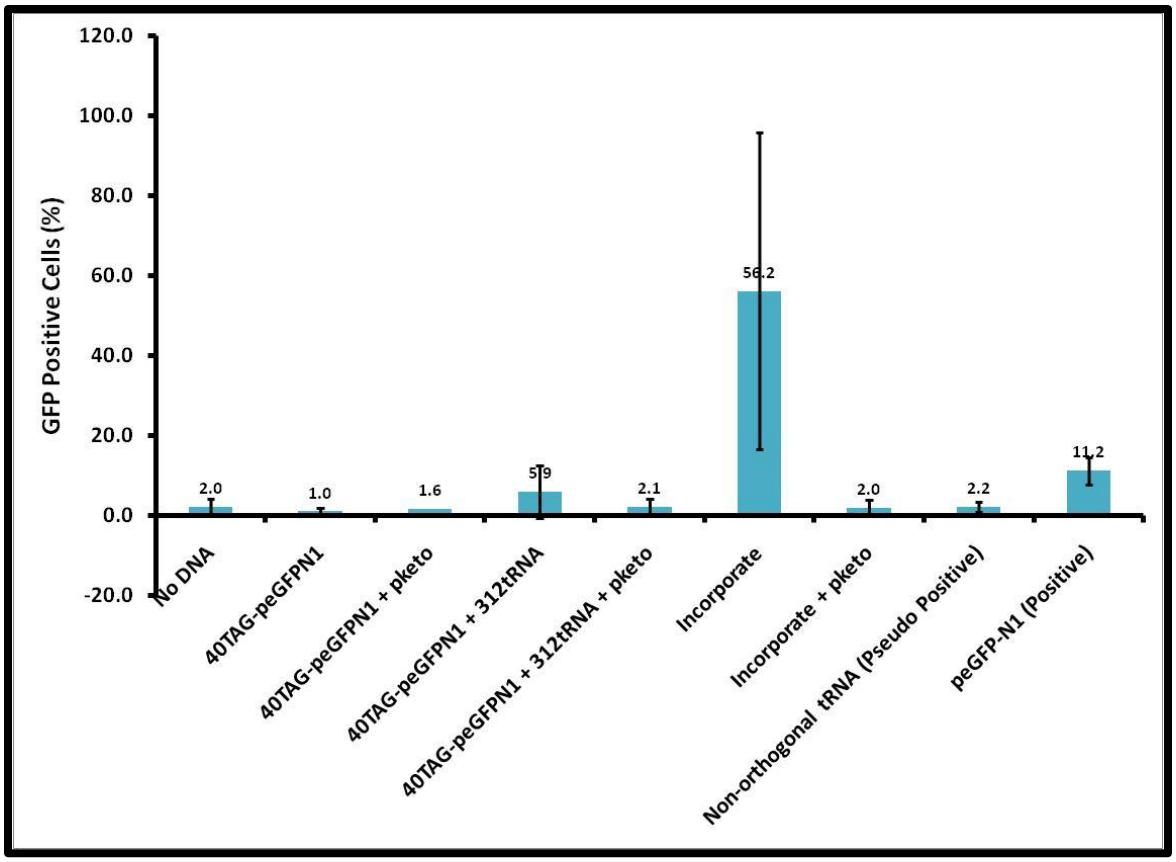


Figure 6.2: Fluorescence intensity of P19s expressing components of the *M. jannaschii* based pketo incorporation system, as determined by FACS analysis. Fluorescence intensities were averaged over two experiments, each measured twice. Only columns denoted with '+ pketo' were grown in the presence of pketo.

synthetase and tRNA pair and the same pketo, but across different cell lines with different transfection techniques, and all had yielded the same unexpected results. Fluorescence microscopy detected very little difference between incorporations with and without pketo, but FACS consistenly detected dramatically higher fluorescence in the incorporate condition without pketo than in any other condition, including positive controls for all samples. With that in mind, it became evident that the *M. jannaschii* based pketo specific RS was somehow increasing the fluorescence of cells in which it was expressed, and pketo was legitimately decreasing fluorescence in all cells transcribing exogenous amber suppressor tRNA.

While there are certainly a variety of explanations for this phenomenon, the most likely is that GFP is being expressed at low levels in cells transfected with 40TAG-peGFPN1, 312tRNA, and pketoMJRS. Either GFP fluorescence or GFP folding is then disrupted in cells which are grown in the presence of pketo. The design of this UAA incorporation system, at this stage, directs incorporation of pketo at the 40th amino acid in eGFP. This amino acid is not part of the UV core and is surface accessible. Thus, incorporation at the designated location would not be expected to decrease fluorescence. However, it has been shown that mutations in or near the UV core of GFP can disrupt folding and effectively knock down fluorescence despite expression of full length protein. Thus, if pketo were somehow being incorporated in locations other than the designated 40th amino acid, then it is logical that this mutation could effectively prevent proper folding of GFP and subsequent fluorescence measurements.

The question then becomes: could pketo be incorporated at locations other than the designated 40th amino acid? As the Tirrell group has aptly demonstrated, cells deprived of a certain amino acid will incorporate unnatural analogs of that amino acid into all synthesized proteins, provided that they have no means of synthesizing the

deprived amino acid (6). In these deprivation type experiments, the concentration of UAA in the media is 1 mM. In contrast, the experiments represented by Figure 6.1 and Figure 6.2 used 5 mM pketo, but we did not deprive the cells of any natural amino acids. However, according to the manufacturer's product information, L-tyrosine is at 0.231 mM in the media, and L-phenylalanine is at 0.194 mM in the media. Thus, a concentration of 5 mM pketo in the media is more than 20X the concentration of L-tyrosine and 25X the concentration of L-phenylalanine in the media. Since pketo is most similar to phenylalanine, it could be that an endogenous phenylalanyl synthetase is mistakenly charging phenylalanyl tRNAs with pketo. Since the concentration of pketo is so many fold higher than that of phenylalanine, it could be 'flooding the system' and increasing the likelihood of this mischarging. Since the Tirell group has demonstrated repeatedly that endogenous synthetases can mistakenly charge endogenous tRNAs with UAAs if they are present while a canonical amino acid is absent, there is some evidence to suggest that mischarging of endogenous tRNAs with pketo could be happening in these experiments. If that were the case, then GFP fluorescence could be compromised by indiscriminate/random incorporation of pketo. This could explain the apparent decrease in fluorescence between non-pketo and pketo containing conditions. However, this still does not account for the high fluorescence exhibited by incorporation conditions in the absence of pketo.

In other UAA incorporation experiments, researchers have spent months, some years, in multiple rounds of positive and negative selections in which a whole enzyme was screened with libraries in only certain portions of the enzyme (7, 8). That is in contrast to what has been done in this research where pieces of enzymes were stitched together to create orthogonality as well as UAA specificity. Since the pketo RS used in the preceding experiments is this chimeric stitched together enzyme, incorporating

pieces of various other enzymes, it has not undergone any positive or negative selections. That being the case, this enzyme is unproven. Based upon the data shown in Figure 6.1 and Figure 6.2, this enzyme may have lost specificity for pketo or it may have been made permissive such that it accepts multiple amino acids or amino acid analogues as substrates. If this were the case, it would explain the increased fluorescence in the incorporation without pketo conditions. Coupled with the possibility that the concentration of pketo is high enough to persuade endogenous synthetases to charge endogenous tRNAs with it, these two possibilities together could explain the curious results shown thus far. However, it must be stated that there is no conclusive evidence by which one could draw a reasonably justified conclusion as to the mechanisms underlying the results. It can be determined that the amber suppressor tRNA and pketo RS do not function as anticipated and certainly do not appear to be site-specifically incorporating pketo into the amber stop codon in the 40th position of GFP.

In light of these confusing results, we chose to pursue the one available avenue to clarify the mechanism behind these results. With two final experiments, we attempted to incorporate pketo site-specifically into 40TAG-eGFP in P19s, and we chose to probe for site-specific incorporation (or lack thereof) by hydrazide modification of the keto functional group. As with the previous experiments, P19s were transfected with the components of the *M. jannaschii* based pketo incorporation system, and conditions were split after electroporation and treated with either no pketo or 10 mM pketo. A small sample of each was subjected to FACS, results shown in Figure 6.3, and an entire T-75 flasks's worth of cells for each condition were then used for hydrazide modification. After hydrazide modification, successful site-specific incorporation of pketo would show the presence of AlexaFluor® conjugated hydrazide at the molecular weight of GFP after samples were subjected to SDS-PAGE, and the same band would

appear in a western blot probing for GFP. If pketo was incorporated into no proteins, then there should be no hydrazide presence in the gel, and finally if pketo were incorporated randomly due to the overwhelming concentration of pketo in the cells,

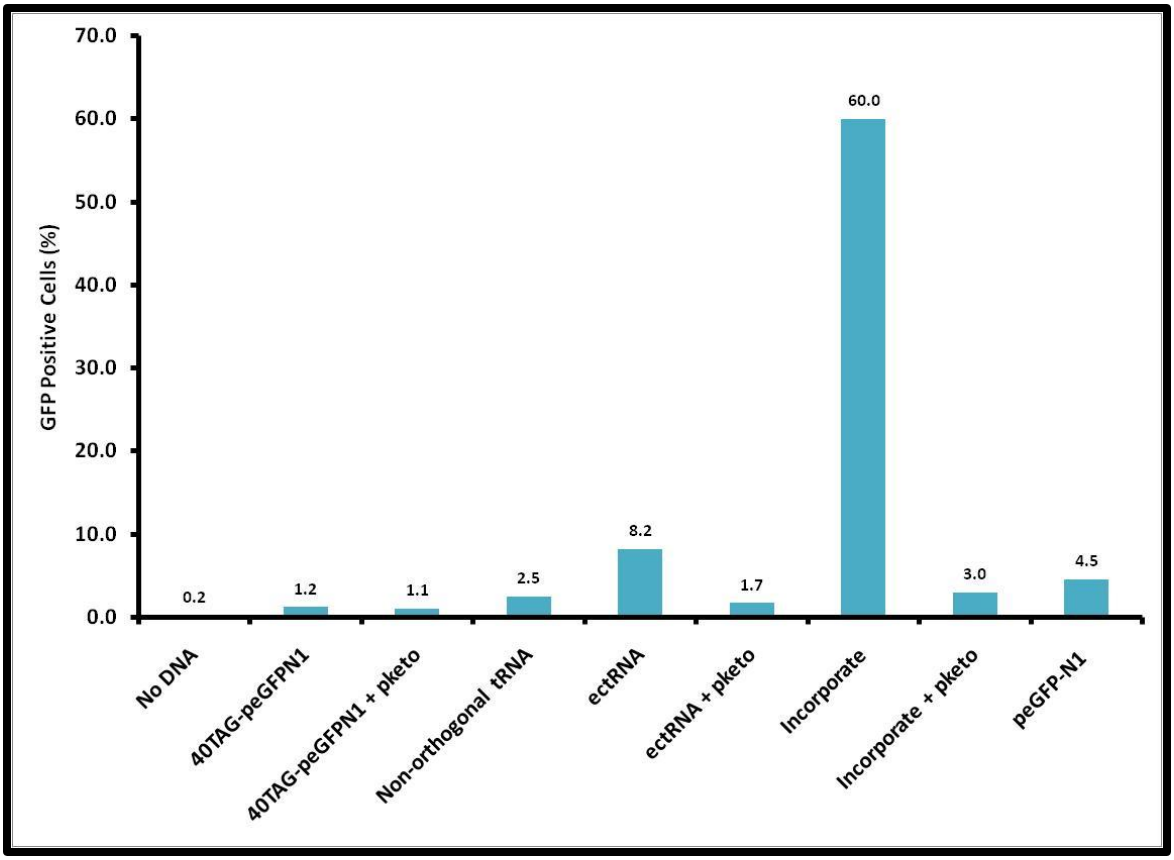


Figure 6.3: Fluorescence intensity of P19s expressing components of the *M. jannaschii* based pketo incorporation system, as determined by FACS analysis. Fluorescence intensities were averaged over two experiments, each measured twice. Only columns denoted with '+ pketo' were grown in the presence of pketo.

there should be hydrazide staining in a smear across multiple molecular weights in the cell lysate.

Based upon the percentage of fluorescent cells shown in Figure 6.3 we can ascertain, once again, that addition of pketo to the medium of P19s transfected with

exogenous amber suppressor tRNAs reduces measureable fluorescence. Once again, the incorporation condition without pketo in the media exhibited the highest percentage of GFP positive cells. Of note is the fact that the ectRNA + pketo condition exhibits half the fluorescence of the incorporate + pketo condition, indicating that co-expression of pketoMJRS with 312tRNA increases amber stop codon suppression. This gave us hope that hydrazide modification would demonstrate that pketo was in fact site-specifically incorporated into the reporter protein encoded by 40TAG-peGFPN1.

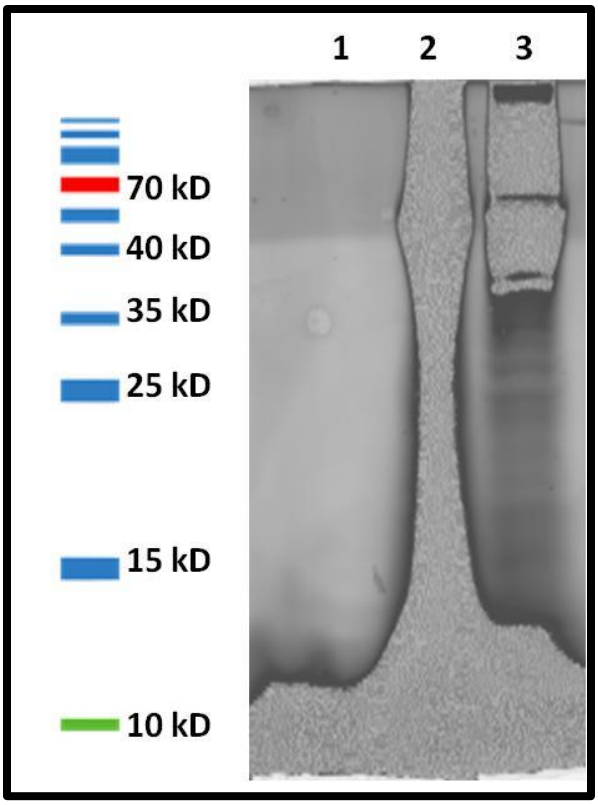


Figure 6.4: Fluorescence scan of SDS-PAGE gel used to separate hydrazide labeled proteins (1) Cell lysate from non-orthogonal tRNA condition (2) hydrazide modified cell lysate from incorporate + pketo condition (3) hydrazide modified cell lysate from peGFP-N1 + pketo condition. Scans were acquired with a Typhoon Trio scanner with 532 nM excitation, a 610nm bandpass filter, and power set to 300 V.

The grey, pixilated portions of lanes 1 and 2 in Figure 6.4 are indicative of extremely high fluorescence and are an artifact of the imaging technique. The high fluorescence in these two lanes demonstrates that cell lysate from conditions with and without exogenous tRNAs or RSs are hydrazide modified when pketo is included in the media. Since the signal in this image was so high, as the protein loading was very high, it was difficult to discern whether or not there was a distinguishable band of hydrazide labeling in the range of GFP molecular weight. Since we are interested in finding the clear existence, or lack thereof, of a band with both GFP and hydrazide modification, we ran another gel from a separate but identical experiment, using $1/10^{th}$ the amount of protein. That gel was imaged using the same technique as previously. Figure 6.5A is the unaltered image, demonstrating AlexaFluor® coupled hydrazide fluorescence. Lanes 2 and 3 are from the incorporation + pketo and the peGFP-N1 + pketo conditions, and they were both treated with hydrazide. At the bottom of this gel, there is a dark band corresponding to un-reacted hydrazide which did not dialyze out, but electrophoresis effectively separated it from the rest of the cell lysate. Just below the 15 kD standard band, there is dark, heavy band indicative of a large amount of overexpressed protein creating a bubble. If this band had corresponded to the molecular weight of GFP, it might have indicated successful site-specific incorporation of pketo in response to amber stop codon suppression. Unfortunately, it does not correspond to the molecular weight of GFP, rendering this particular experiment of very little value. Since hydrazide can react with various types of carboxyl groups, it can react with multiple proteins in the cell lysate. If this experiment had used purified GFP instead of lysate, it might have given more information and conclusively indicated presence or absence of pketo, even if the yield was low. Since the western blot in Figure 6.5D shows no GFP, it is unsurprising that the hydrazide staining was not localized in a band corresponding to GFP.

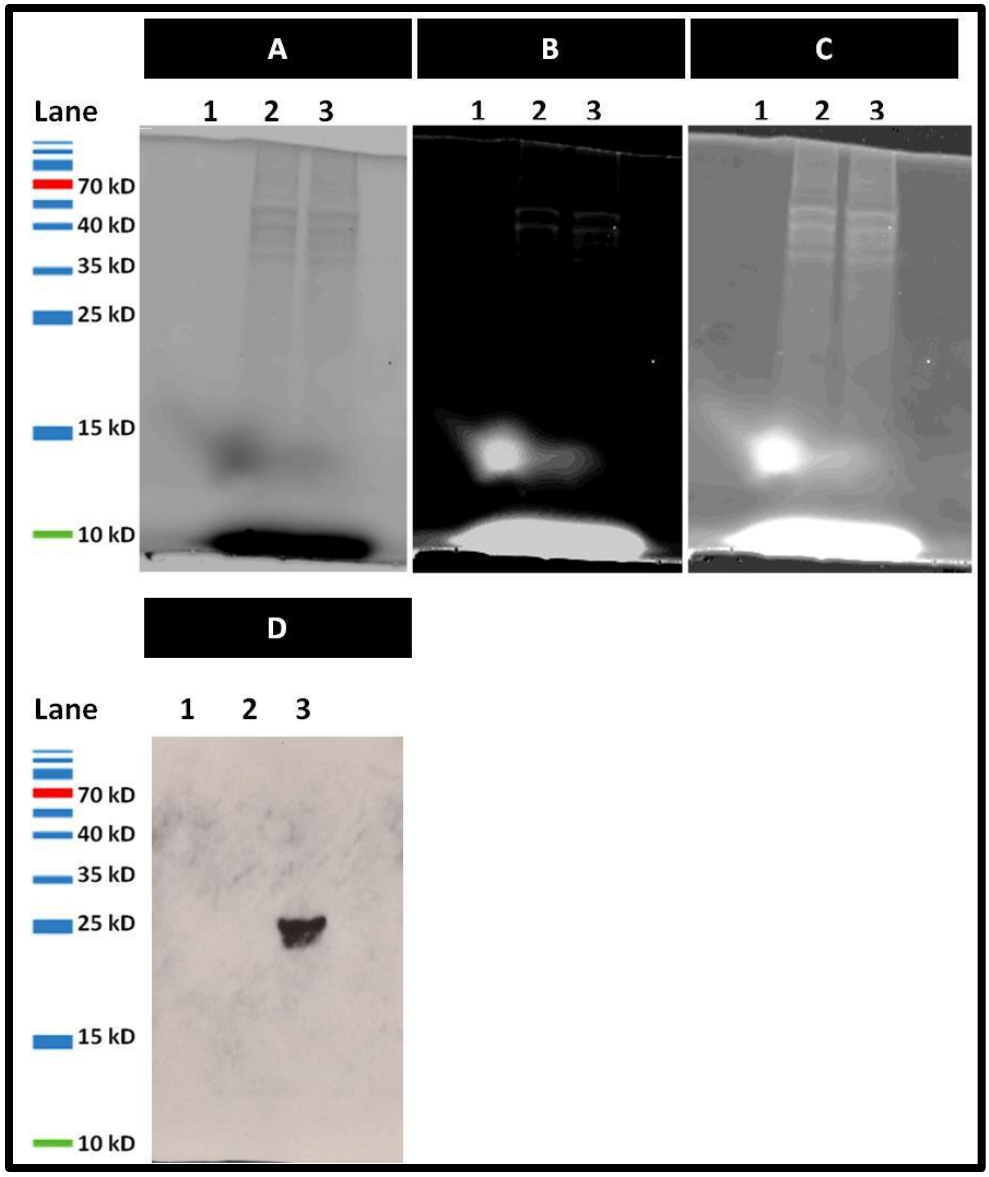


Figure 6.5: **[A-C]** Fluorescence scan of SDS-PAGE gel used to separate hydrazide labeled proteins (1) Cell lysate from non-orthogonal tRNA condition (2) hydrazide modified cell lysate from incorporate + pketo condition (3) hydrazide modified cell lysate from peGFP-N1 + pketo condition. Scans were acquired with a Typhoon Trio scanner with 532 nM excitation, a 610nm bandpass filter, and power set to 400 V. (A) unaltered image (B) Image inverted, contrast increased (C) image inverted, exposure increase 8X using Adobe Photoshop. **[D]** Western blot detecting GFP expression, from the same gels shown in A-C.

CONCLUSION

In the preceding pages, data has been presented indicating that the function of the *M. jannaschii* based pketo specific amber suppressor tRNA/chimeric RS pair is compromised as compared to the function of the previously evolved pketo specific RS. Testing of this system in stem cells revealed unexpected fluorescence intensity increases caused by expression of the *M. jannaschii* chimeric pketo specific RS as well as decreases in fluorescence intensity upon addition of pketo to the media of amber suppressor tRNA synthesizing cells. Figure 6.4 and Figure 6.5 do not provide conclusive evidence of site-specific pketo incorporation. Further, the inability to detect GFP expression from non-orthogonal tRNA synthesizing cells or pketo incorporating cells indicates that larger amounts of cells and possibly protein purification will be necessary to detect GFP expression and/or pketo incorporation. Thus, the system needs further refinement before it can be successfully applied to the incorporation of UAAs in P19s or stem cells. The most direct approach is probably directed evolution using library screening in yeast to create UAA specific RSs from whole enzymes as opposed to the stitched together nature of those used in this investigation.

REFERENCES

1. Tanaka Y, Bond M, Kohler J. Photocrosslinkers illuminate interactions in living cells. MOL BIOSYST. 2008;4(6):473-80.
2. Umeda A, Thibodeaux G, Moncivais K, Jiang F, Zhang Z. A versatile approach to transform low-affinity peptides into protein probes with cotranslationally expressed chemical cross-linker. ANALYTICAL BIOCHEMISTRY. 2010;405(1):82-8.
3. Wang L, Schultz PG. Expanding the genetic code. Chem Commun (Camb). 2002(1):1-11.
4. Ye S, Kohrer C, Huber T, Kazmi M, Sachdev P, Yan EC, et al. Site-specific incorporation of keto amino acids into functional G protein-coupled receptors using unnatural amino acid mutagenesis. J Biol Chem. 2008;283(3):1525-33.
5. Wang W, Takimoto JK, Louie GV, Baiga TJ, Noel JP, Lee KF, et al. Genetically encoding unnatural amino acids for cellular and neuronal studies. Nat Neurosci. 2007;10(8):1063-72.
6. Song W, Wang Y, Yu Z, Vera CI, Qu J, Lin Q. A metabolic alkene reporter for spatiotemporally controlled imaging of newly synthesized proteins in Mammalian cells. ACS Chem Biol.5(9):875-85.
7. Zhang Z, Alfonta L, Tian F, Bursulaya B, Uryu S, King DS, et al. Selective incorporation of 5-hydroxytryptophan into proteins in mammalian cells. Proc Natl Acad Sci U S A. 2004;101(24):8882-7.
8. Wang L, Brock A, Herberich B, Schultz PG. Expanding the genetic code of Escherichia coli. Science. 2001;292(5516):498-500.
9. Beatty KE, Tirrell DA. Two-color labeling of temporally defined protein populations in mammalian cells. Bioorg Med Chem Lett. 2008;18(22):5995-9.

CHAPTER SEVEN

A Mammalian Two-Hybrid System for Use in P19 Embryonal Carcinoma Cells

INTRODUCTION

We have previously described the development of a tetracycline repressor-based mammalian two-hybrid (trM2H) system (1), and in an attempt to further the field of protein-protein interactions (PPIs) studies in stem and pluripotent cell lines, we chose to optimize it and apply it to the study of PPIs in P19 embryonal carcinoma cells (P19s). This system was briefly discussed in Chapter Two, but a few more details of the system will be outlined below, after which optimization of the system for use in sensitive cell lines will be described. Finally, the use of this system in P19s will be discussed.

The previously developed trM2H is based upon the highly active tetracycline repressor (TetR), a dimeric transcriptional regulator which binds specifically to the tet operator (TetO) to inhibit expression of a downstream gene. It was developed by modification of the commercially available Tet-Off® Advanced Inducible Gene Expression System (Clontech, Mountain View, CA). The Tet-Off system is composed of the full length TetR C-terminally fused to three transcriptional activation domains from the herpes simplex virus VP16 (1). TetR-VP16 fusion protein dimerization allows the TetR dimer to bind the Tet-Responsive-Element (Tre) while the VP16 domain recruits transcriptional machinery. My colleagues replaced the TetR c-terminal dimerization domain (alpha8-alpha10) with bait and prey molecules, shown in Figure 7.1A. When bait and prey proteins interact, the reporter protein, GFP in this case, is expressed (Figure 7.1C). This system was successfully used to assess interactions of peptide pairs with

weak binding constants (1). Due to the sensitivity of this system, it was a good candidate for testing in P19s. Thus it was tested in P19s using the transfection methods described in Chapter Three.

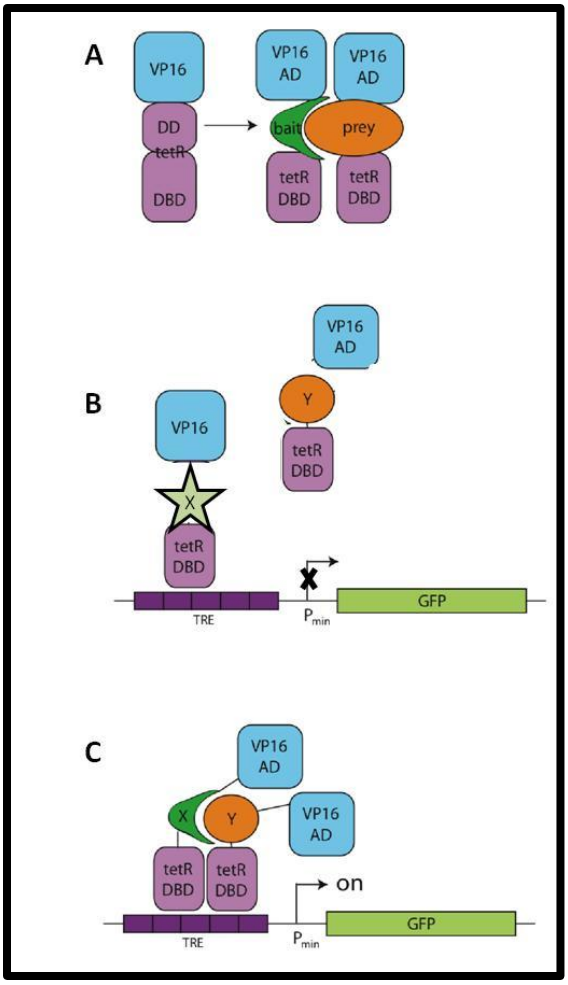


Figure 7.1 Basics of the trM2H system (A) replacement of dimerization domain with bait and prey molecules (B) bait/prey that do not interact will not cause GFP expression (C) interacting bait and prey turn on GFP expression

OVERALL EXPERIMENTAL DESIGN

P19 embryonal carcinoma cells were transfected with the components of the previously developed trM2H system, and its function was assessed using fluorescence microscopy and fluorescence activated cell sorting (FACS).

MATERIALS AND METHODS

Routine Cell Culture

P19 embryonal carcinoma cells were obtained from American Type Culture Collection (ATCC). Cells were routinely maintained in Alpha-Mem (HyClone, Logan, UT) with 10% (v/v) fetal bovine serum (Atlanta Biologicals, Lawrenceville, GA) and nonessential amino acids (Sigma-Aldrich, St. Louis, MO) in a humidified incubator at 37° Celsius with 5% CO_2 atmosphere. Cells were passaged every 1-3 days at a ratio between 1:4 and 1:20. For passaging, cells were rinsed in phosphate buffered saline (PBS) (Sigma-Aldrich, St. Louis, MO), incubated with 0.5% trypsin (Gibco, Invitrogen, Carlsbad, CA) for 3-5 minutes, and centrifuged to pellet. Cells were then resuspended in complete media (as described earlier in this paragraph) and seeded into clean flasks with fresh media.

DNA Constructs

The plasmids pTet-Off and pTRE-Tight-AcGFP were purchased from Clontech (Mountainview, CA).

Electroporation

P19 embryonal carcinoma cells at 80% - 90% confluency were harvested using the same procedure as used for passaging, washed once with PBS (Sigma-Aldrich, St. Louis, MO), spun down again, and resuspended in 600 µL of PBS per harvested T-75 flask of cells. For each condition in an experiment, 200 µL of cells was aliquotted into an eppendorf tube. Then DNA appropriate for each condition was added to the cells. The total volume for each electroporation was then brought up to 450 µL by addition of PBS. Cell DNA mixtures were then transferred to electroporation cuvettes with a 2 mm gap (Fisher Scientific, Waltham, MA), and cuvettes were incubated on ice for no less than 5 minutes and no more than 15 minutes before electroporation. Cell-DNA mixtures were

electroporated using a Bio-Rad Gene Pulser with capacitance extender and pulse controller (Bio-Rad, Hercules, CA) with settings of 270 Volts, 100 Ω, and 960 μF. After electroporation, cells were incubated on ice for at least 3 minutes and then resuspended in complete media. Cells were then plated in tissue culture vessels of the appropriate size (Corning, Lowell, MA). Media was changed 24 hours post-transfection and every 24 hours thereafter until experiments were completed.

FugeneHD Transfection

FugeneHD was purchased from Roche (Switzerland). P19s were grown to 50-60% confluency in 6-well plates (Corning, Lowell, MA) and then transfected with the components of the trM2H system. FugeneHD was used for transfection, and complexes were formed in Opti-Mem using a ratio of 2 μL FugeneHD to 1 μg DNA. Media was changed every 24 hours after transfection.

Hybrid FugeneHD + Electroporation Transfection

Fugene HD was purchased from Roche (Switzerland). A modified version of Roche's FugeneHD protocol was used for the hybrid transfection procedure. Instead of 50 μL of diluent per μg of DNA, 25 μL of diluent was used per μg of DNA. Assuming the concentration of DNA were 1 ug/μL, the following protocol would be used. First, PBS is added to an eppendorf tube such that the final volume of PBS + DNA + FugeneHD will be 25 uL/ug DNA. In this case that is 330 μL of PBS. Then 15 ug of DNA is added to the PBS, followed by 30 μL of FugeneHD. The tube is closed securely, vortexed, and incubated while cells are harvested. P19s at 80-90% confluency are harvested from a single T-75 flask, washed in PBS, spun down, and resuspended in 900 μL of PBS. 300 μL of the cell solution is then aliquotted into an eppendorf tube and spun down. Once the supernatant has been aspirated from the eppendorf tube of P19s, the FugeneHD

complexes should have been incubating for about 15 minutes. Once the FugeneHD complexes have incubated at room temperature for at least 15 minutes, the P19 cell pellet in the eppendorf tube is resuspended in Fugene complexes and transferred to an electroporation cuvette.

RESULTS AND DISCUSSION

trM2H Testing With Adherent FugeneHD Transfection

Before the stem cell transfection optimization had reached a definite conclusion, the trM2H was tested in P19s using adherent FugeneHD transfection. pTRE-Tight AcGFP was used as a negative control, pTet-Off was served as the positive control, and the homo-dimerizing construct gcn4 and its non-dimerizing mutant gcnxx were used to test the functionality of the system in P19s. The gcn4 construct proved to be the easiest to detect in previous experiments with HEK293T cells, so we posited that testing with gcn4 would give a quick assessment of the viability of this technique in P19s.

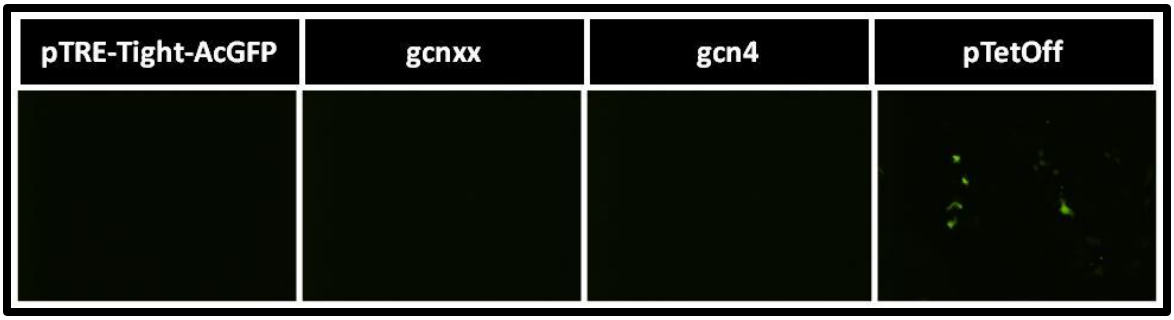


Figure 7.2: Fluorescence microscopy images of P19s expressing dimerizing and non-dimerizing trM2H proteins. pTRE-Tight-AcGFP is a negative control, while pTetOff is a positive control. Gcnxx is a non-dimerizing protein, which should not cause GFP expression, and gcn4 is a homodimerizing protein which should cause GFP expression. All images were taken 48 hours post-transfection with exposure of 1 second and gain set to 1.

Figure 7.2 above demonstrates the inefficacy of the trM2H system in initial experiments. Since the pTetOff positive control failed to produce a significant number of

fluorescent cells, we posited that transfection efficiency was limiting the use of this system in P19s. Previous experiments in HEK293T cells yielded pTetOff positive controls with GFP expression levels similar to those of plain peGFP-N1 positive controls, around 80-95%. This is in stark contrast to the very low GFP expression level shown in the pTetOff frame of Figure 7.2. Thus, we attempted to optimize transfection efficiency with FugeneHD by simply using 2-4 times the amount of DNA complexes to test the system again, but this did not increase transfection efficiency or the rate of appearance of fluorescent cells in positive controls or conditions with homodimerizing proteins. Therefore electroporation was explored as a way to increase transfection efficiency and enable the use of the trM2H in P19s.

trM2H Testing With Electroporation

P19s were electroporated with the components of the trM2H system according to the protocol on page 149. Figure 7.3 below demonstrates the results of that

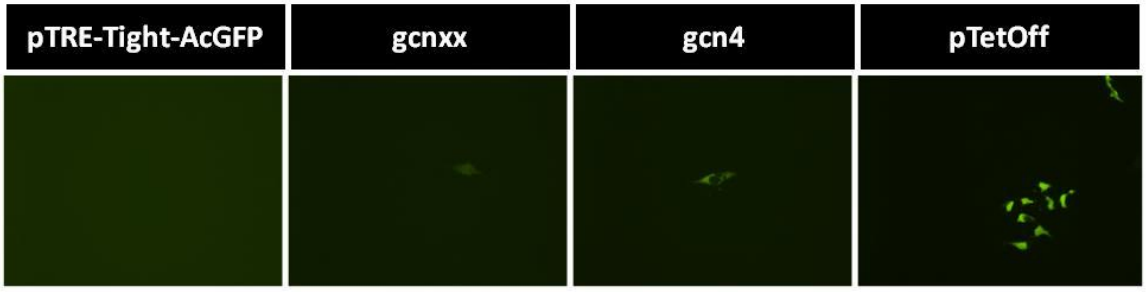


Figure 7.3: Fluorescence microscopy images of P19s expressing dimerizing and non-dimerizing trM2H proteins after electroporation. pTRE-Tight-AcGFP is a negative control, while pTetOff is a positive control. Gcnxx is a non-dimerizing protein, which should not cause GFP expression, and gcn4 is a homodimerizing protein which should cause GFP expression. There were no fluorescent cells in the pTRE-Tight-AcGFP well, one in the gcnxx well, and eleven in the gcn4 well. The pTetOff well had too many fluorescent cells to count by hand. All images were taken 48 hours post-transfection with exposure of 1 second and gain set to 1.

experiment. As stated in the figure caption, the gcnxx condition had a single fluorescent cell in the well, while the gcn4 condition had 11 fluorescent cells. If subjected to a more quantitative fluorescence assessment, this experiment might have yielded a signal-to-noise ratio analogous to that achieved in HEK293T cells, in which case this system would be an excellent tool for the investigation of PPIs in pluripotent cell lines. Unfortunately fluorescence microscopy was the only fluorescence assessment method available to us at the time. Since the method of counting fluorescent cells by hand was time consuming and certainly not cost-effective in terms of man-hours spent versus information generated, the trM2H was not further tested until FACS sorting was available, and the hybrid FugeneHD + electroporation method was developed.

trM2H Testing With FugeneHD + Electroporation

Figure 7.4 below demonstrates the infeasibility of using the trM2H system in its current state to study PPIs in P19s or other pluripotent cell lines. The positive control yielded 1% fluorescent cells, which is barely more than twice the signal from the negative control pTRE-Tight-AcGFP. Similarly disappointing were the percentages of fluorescent cells for gcnxx, fosjun, and fosjunmu conditions as there was little difference between all three, and in fact a nondimerizing fosjunmu condition exhibited more fluorescent cells than the dimerizing fosjun condition. These results indicate that the data in Figure 7.4 is little more than noise. Thus, we concluded that the trM2H may be driven by tightly regulated promoters which cause very low levels of protein expression in P19s, making it nearly impossible to use this system as a means of investigating PPIs in P19s or other pluripotent cell lines.

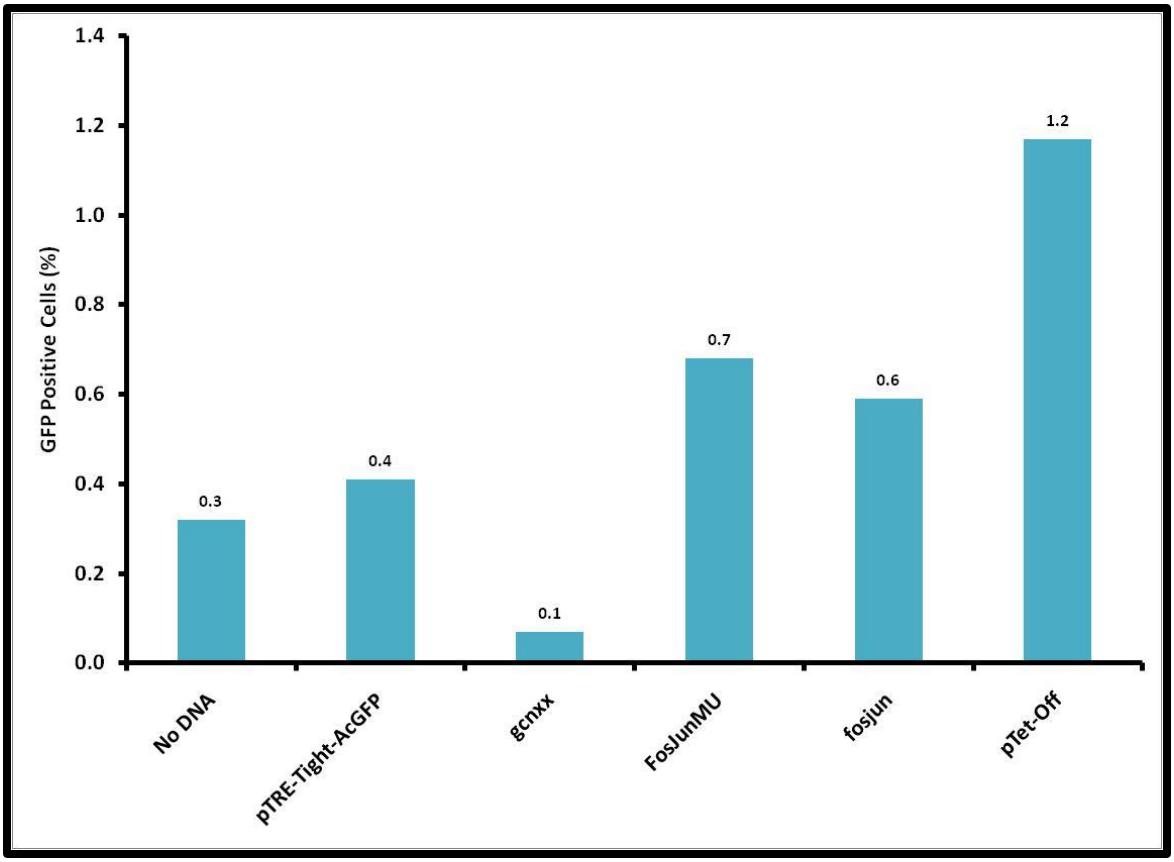


Figure 7.4: Percentage of P19s expressing GFP as a result of dimerization of trM2H proteins, as determined by FACS.

CONCLUSION

Though use of the trM2H has successfully detected low affinity PPIs in HEK293T cells (1), it does not function in P19s as well. This may be caused by poor protein expression in P19s in response to the trM2H's tightly regulated promoters, indicating that the emphasis of the trM2H system on decreasing noise may in fact be preventing its function altogether in P19s. While the trM2H does not appear at this time to be a viable technology for use in P19s or other pluripotent cell lines, that is not to say that other two-hybrid systems could not produce high quality data in this cell line. It is possible that other systems with lower signal to noise ratios than the trM2H could

produce better results in P19s, especially if they utilize promoters other than the Tet operon.

REFERENCES

1. Thibodeaux GN, Cowmeadow R, Umeda A, Zhang Z. A tetracycline repressor-based mammalian two-hybrid system to detect protein-protein interactions in vivo. Anal Biochem. 2009;386(1):129-31.

CHAPTER EIGHT

Facilitating Protein-Protein Interactions Studies in Stem Cells with UAA Incorporation and a Mammalian Two-Hybrid System: Conclusions and Future Directions

INTRODUCTION

In Chapter Two the many uses and possibilities of stem cells for research and therapeutic applications were discussed. It was also mentioned that many promising lines of stem cell research are based upon incomplete knowledge of the biological processes underlying differentiation, dedifferentiation and tumorigenesis. Protein-protein interactions (PPIs) studies could yield a wealth of useful information about those processes and facilitate better, more efficient manipulation of stem cells, thus paving the way for swift development of stem-cell based therapies. In order to facilitate the process of studying PPIs in stem cells, the aim of this dissertation has been to develop tools appropriate and efficient for studying PPIs in P19 embryonal carcinoma cells (P19s), a pluripotent cell line.

P19 EMBRYONAL CARCINOMA TRANSFECTION METHODS

Most, if not all PPI investigations require introduction of exogenous DNA and/or proteins to cells. Since pluripotent and stem cell lines are difficult to transfect, the first part of this project involved development of protocols for efficient DNA transfer to P19s. In Chapter Three, multiple methods were investigated in order to find the most successful transfection method for use in P19s. Commercially available reagents Fugene6 and FugeneHD were capable of transfecting P19s, but the efficiency was low. Electroporation could be used to achieve high transfection efficiencies, on the order of

80%, but electroporation was very sensitive to several experimental parameters that are difficult to control including the length of time between harvest, electroporation, and replating. Additionally, highly efficient electroporation require milligram amounts of DNA for each experiment, which made this technique prohibitive for anyone, including us, without unlimited amounts of DNA. So electroporation was used in many of the experiments in this work, but we chose to investigate the efficacy of a hybrid transfection method involving FugeneHD complexes and electroporation. This method was able to increase the consistency of transfection success while at the same time dramatically decreasing the amount of DNA necessary for each condition. So two methods for efficient transfection of P19s were developed – electroporation with our protocol is appropriate for highly efficient transfection when large quantities of DNA are available, and the hybrid method is appropriate for acceptably efficient transfection when DNA is in limited supply.

CREATION OF AN UNNATURAL AMINO ACID INCORPORATION SYSTEM FOR MAMMALIAN CELLS

Chapter Four described the creation of an unnatural amino acid (UAA) incorporation system for use in mammalian cells. An orthogonal amber suppressor tRNA was created, as was an aminoacyl-tRNA synthetase (RS) capable of charging the amber suppressor tRNA. A reporter protein plasmid with a stop codon was also created in order to decrease background signal from a previous version of the same reporter plasmid. It is important to note that the RS was created by 'cutting and pasting' modular protein domains from both *Methanocaldococcus jannaschii (M. jannaschii)* and *Escherichia coli (E. coli)* based enzymes, and this system functioned quite well in HEK293T cells. The development of this system was recently published (1), and its success lead us to believe that it could be successfully implemented in P19s.

Amber Stop Codon Suppression in P19 Embryonal Carcinoma Cells

Since the system developed in Chapter Four worked well in HEK293T cells we chose to apply it to P19s in its current state and described that application in Chapter Five of this dissertation. We also tested an *E. coli* based tRNA/RS pair and a non-orthogonal tRNA to confirm that these pluripotent cells were in fact able to synthesize and use exogenous amber suppressor tRNAs. Both the non-orthogonal tRNA and the *E. coli* tRNA/RS pair were able to suppress amber stop codons well in P19s, but the *M. jannaschii* pair performed poorly when used in P19s. We hypothesize that this is due to the chimeric nature of the *M. jannaschii* based enzyme, but have not been able to confirm that as yet. Since some of our colleagues have experienced that UAA specific RSs are more efficient than their canonical amino acid counterparts, we moved forward with both the *E. coli* pair and the *M. jannaschii* pair.

Unnatural Amino Acid Incorporation in P19 Embryonal Carcinoma Cells

Since we already had an *M. jannaschii* derived RS specific for the UAA m-acetyl-L-phenylalanine (pketo), we used it in an attempt to incorporate pketo into the protein produced by the plasmid 40TAG-peGFPN1. Fluorescence activated cell sorting (FACS) data contradicted fluorescence microscopy images so consistently that we were forced to recognize that this pketo tRNA/RS pair derived from *M. jannaschii* was exerting some unknown influence on fluorescence. Though we have yet to determine exactly the mechanism behind the fluorescence intensities measured by FACS, we have posited that the chimeric pketo RS may have a compromised ability to recognize the intended tRNA, while at the same time pketo may be used as a substrate by endogenous RSs. In the given time frame we were unable to further investigate the GFP range fluorescence phenomena exhibited by these samples, but we were able to probe for pketo

incorporation using hydrazide modification. Fluorescently imaged protein gels indicated that pketo may be incorporated at random into all proteins, in a genome-wide fashion, and western blots failed to detect GFP expression in all samples except the positive controls. This leads us to believe that the system is not yet efficient enough to detect reporter protein expression in cell lysate, but purification of reporter proteins from larger cell samples could fix this problem. Also indicated by these results is the fact that P19s are probably good candidates for spatiotemporal protein labeling as routinely performed by the Tirrell group (2).

At the same time we were developing an *E. coli* based RS specific for the UAA o-methyl-L-tyrosine. Unfortunately, the final mutations in that RS were unable to be completed in time to be included in this dissertation. The sum of our results in amber stop codon suppression and UAA incorporation in P19s is that, while definitely possible at this point in time, these endeavors would benefit greatly from directed evolution of the current RSs or new RSs. This could ensure that the UAA specific RSs are recognizing the intended amber suppressor tRNAs and increase the efficiency with which they charge those tRNAs with UAAs. This could alleviate the necessity of using such high concentrations of UAAs in the growth medium so that genome-wide labeling of proteins would not proceed as efficiently as it seems to have done in the experiments described herein.

A Two-Hybrid System in P19 Embryonal Carcinoma Cells

Though the tetracycline-repressor based mammalian two-hybrid (trM2H) system did work quite well in HEK293Ts (3), it was not successful in detecting protein-protein interactions in P19s. This may be due to the well-documented fact that undifferentiated cell lines frequently respond differently than terminally differentiated cell lines to

various promoters. The trM2H uses very tightly regulated promoters, which may not function well in P19s. Despite its lack of success in P19s, the trM2H could be a useful tool if modified for use in stem cells, as could other two hybrid systems. This could be done by optimizing the promoters used on the bait and prey molecules as well as finding a better dimerization dependent promoter for use in P19s.

OVERALL CONCLUSION

The overall goal of the work presented in this dissertation was to facilitate the study and discovery of PPIs in stem and/or pluripotent cells by developing novel methods and tools for use in these cell types. A successful transfection method was created, and an *M. jannaschii* based UAA incorporation system was created and its use demonstrated in HEK293T cells. This same system did not work perfectly in P19s, but the number of cells successfully expressing components of this system and suppressing amber stop codons is certainly high enough in the stem cell and P19 community to be considered quite successful. While we would like to have achieved a higher efficiency of amber suppression with this system, we believe that the door is now quite open, and future work can proceed in a more directed manner now that this work has clarified the areas in need of improvement.

We successfully used the *E. coli* based UAA incorporation system to suppress amber stop codons in P19s with great success, and it was incredibly regrettable that we were unable to incorporate a UAA using this same system. We believe the next step in UAA incorporation in P19s involves the *E. coli* pair, and that this pair will yield the fastest, most exciting applications of UAA incorporation in P19s at this time. Therefore the next logical line of work in this vein would be to finish creation of the o-methyl-L-

tyrosine specific RS and/or to collaborate with other researchers who have already developed these enzymes in order to quickly exploit their utility.

Since the overall goal of this research was to facilitate the study of PPIs in stem cells, we believe that goal has certainly been achieved. While this work is just beginning, it has the potential to greatly impact the field of stem cell science, and it is our belief that the work presented in this dissertation amounts to being just on the cusp of using UAAs in a variety of stem cell based applications.

REFERENCES

1. Thibodeaux G, Liang X, Moncivais K, Umeda A, Singer O, Alfonta L, et al. Transforming a Pair of Orthogonal tRNA-aminoacyl-tRNA Synthetase from Archaea to Function in Mammalian Cells. PLOS ONE. 2010;5(6):-.
2. Song W, Wang Y, Yu Z, Vera CI, Qu J, Lin Q. A metabolic alkene reporter for spatiotemporally controlled imaging of newly synthesized proteins in Mammalian cells. ACS Chem Biol.5(9):875-85.
3. Thibodeaux GN, Cowmeadow R, Umeda A, Zhang Z. A tetracycline repressor-based mammalian two-hybrid system to detect protein-protein interactions in vivo. Anal Biochem. 2009;386(1):129-31.

References

1. Grady D. The Hope, and Hype, of Cord Blood. The New York Times. 1998 December 1, 1998.
2. Liu CC, Schultz PG. Adding new chemistries to the genetic code. Annu Rev Biochem. 2010;79:413-44.
3. Pfendler K, Kawase E. The potential of stem cells. OBSTET GYNECOL SURV. 2003;58(3):197-208.
4. Voltarelli JC, Couri CE, Stracieri AB, Oliveira MC, Moraes DA, Pieroni F, et al. Autologous nonmyeloablative hematopoietic stem cell transplantation in newly diagnosed type 1 diabetes mellitus. Jama. 2007;297(14):1568-76.
5. Johnson PJ, Tatara A, Shiu A, Sakiyama-Elbert SE. Controlled release of neurotrophin-3 and platelet-derived growth factor from fibrin scaffolds containing neural progenitor cells enhances survival and differentiation into neurons in a subacute model of SCI. Cell Transplant.19(1):89-101.
6. Liu S, Qu Y, Stewart TJ, Howard MJ, Chakrabortty S, Holekamp TF, et al. Embryonic stem cells differentiate into oligodendrocytes and myelinate in culture and after spinal cord transplantation. Proc Natl Acad Sci U S A. 2000;97(11):6126-31.
7. Umeda A, Thibodeaux G, Zhu J, Lee Y, Zhang Z. Site-specific Protein Cross-Linking with Genetically Incorporated 3,4-Dihydroxy-L-Phenylalanine. CHEMBIOCHEM. 2009;10(8):1302-4.
8. Monahan SL, Lester HA, Dougherty DA. Site-specific incorporation of unnatural amino acids into receptors expressed in Mammalian cells. Chem Biol. 2003;10(6):573-80.
9. Sprinzl M, Vassilenko KS. Compilation of tRNA sequences and sequences of tRNA genes. Nucleic Acids Res. 2005;33(Database issue):D139-40.
10. Thibodeaux GN, Cowmeadow R, Umeda A, Zhang Z. A tetracycline repressor-based mammalian two-hybrid system to detect protein-protein interactions in vivo. Anal Biochem. 2009;386(1):129-31.
11. Carletti B, Grimaldi P, Magrassi L, Rossi F. Specification of cerebellar progenitors after heterotopic-heterochronic transplantation to the embryonic CNS in vivo and in vitro. J Neurosci. 2002;22(16):7132-46.
12. Sotelo C, Alvarado-Mallart RM. The reconstruction of cerebellar circuits. Trends Neurosci. 1991;14(8):350-5.

13. Triarhou LC, Zhang W, Lee WH. Amelioration of the behavioral phenotype in genetically ataxic mice through bilateral intracerebellar grafting of fetal Purkinje cells. Cell Transplant. 1996;5(2):269-77.

14. Amariglio N, Hirshberg A, Scheithauer BW, Cohen Y, Loewenthal R, Trakhtenbrot L, et al. Donor-derived brain tumor following neural stem cell transplantation in an ataxia telangiectasia patient. PLoS Med. 2009;6(2):e1000029.

15. Xi J, Khalil M, Spitkovsky D, Hannes T, Pfannkuche K, Bloch W, et al. Fibroblasts support functional integration of purified embryonic stem cell-derived cardiomyocytes into avital myocardial tissue. Stem Cells Dev.20(5):821-30.

16. Rizvanov AA, Guseva DS, Salafutdinov, II, Kudryashova NV, Bashirov FV, Kiyasov AP, et al. Genetically modified human umbilical cord blood cells expressing vascular endothelial growth factor and fibroblast growth factor 2 differentiate into glial cells after transplantation into amyotrophic lateral sclerosis transgenic mice. Exp Biol Med (Maywood).236(1):91-8.

17. Alhadlaq A, Elisseeff JH, Hong L, Williams CG, Caplan AI, Sharma B, et al. Adult stem cell driven genesis of human-shaped articular condyle. Ann Biomed Eng. 2004;32(7):911-23.

18. Macchiarini P, Jungebluth P, Go T, Asnaghi MA, Rees LE, Cogan TA, et al. Clinical transplantation of a tissue-engineered airway. Lancet. 2008;372(9655):2023-30.

19. Xie QP, Huang H, Xu B, Dong X, Gao SL, Zhang B, et al. Human bone marrow mesenchymal stem cells differentiate into insulin-producing cells upon microenvironmental manipulation in vitro. Differentiation. 2009;77(5):483-91.

20. Tayaramma T, Ma B, Rohde M, Mayer H. Chromatin-remodeling factors allow differentiation of bone marrow cells into insulin-producing cells. Stem Cells. 2006;24(12):2858-67.

21. Sun Y, Chen L, Hou XG, Hou WK, Dong JJ, Sun L, et al. Differentiation of bone marrow-derived mesenchymal stem cells from diabetic patients into insulin-producing cells in vitro. Chin Med J (Engl). 2007;120(9):771-6.

22. Mabed M. The potential utility of bone marrow or umbilical cord blood transplantation for the treatment of type I diabetes mellitus. Biol Blood Marrow Transplant. 2011;17(4):455-64.

23. Dai LJ, Moniri MR, Zeng ZR, Zhou JX, Rayat J, Warnock GL. Potential implications of mesenchymal stem cells in cancer therapy. Cancer Lett. 2011;305(1):8-20.

24. Ao A, Hao J, Hong CC. Regenerative chemical biology: current challenges and future potential. Chem Biol. 2011;18(4):413-24.

25. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell. 2006;126(4):663-76.

26. Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, Tian S, et al. Induced pluripotent stem cell lines derived from human somatic cells. Science. 2007;318(5858):1917-20.

27. Cho HJ, Lee CS, Kwon YW, Paek JS, Lee SH, Hur J, et al. Induction of pluripotent stem cells from adult somatic cells by protein-based reprogramming without genetic manipulation. Blood. 2010;116(3):386-95.

28. Okita K, Nakagawa M, Hyenjong H, Ichisaka T, Yamanaka S. Generation of mouse induced pluripotent stem cells without viral vectors. Science. 2008;322(5903):949-53.

29. Warren L, Manos PD, Ahfeldt T, Loh YH, Li H, Lau F, et al. Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. Cell Stem Cell. 2010;7(5):618-30.

30. Hasegawa K, Zhang P, Wei Z, Pomeroy JE, Lu W, Pera MF. Comparison of reprogramming efficiency between transduction of reprogramming factors, cell-cell fusion, and cytoplast fusion. Stem Cells.28(8):1338-48.

31. Ernst M, Oates A, Dunn AR. Gp130-mediated signal transduction in embryonic stem cells involves activation of Jak and Ras/mitogen-activated protein kinase pathways. J Biol Chem. 1996;271(47):30136-43.

32. Ernst M, Novak U, Nicholson SE, Layton JE, Dunn AR. The carboxyl-terminal domains of gp130-related cytokine receptors are necessary for suppressing embryonic stem cell differentiation. Involvement of STAT3. J Biol Chem. 1999;274(14):9729-37.

33. Freemantle SJ, Kerley JS, Olsen SL, Gross RH, Spinella MJ. Developmentally-related candidate retinoic acid target genes regulated early during neuronal differentiation of human embryonal carcinoma. Oncogene. 2002;21(18):2880-9.

34. Damjanov I, Andrews PW. The terminology of teratocarcinomas and teratomas. Nat Biotechnol. 2007;25(11):1212; discussion.

35. Stevens LC. The development of transplantable teratocarcinomas from intratesticular grafts of pre- and postimplantation mouse embryos. Dev Biol. 1970;21(3):364-82.

36. McBurney MW, Rogers BJ. Isolation of male embryonal carcinoma cells and their chromosome replication patterns. Dev Biol. 1982;89(2):503-8.

37. McBurney MW. Clonal lines of teratocarcinoma cells in vitro: differentiation and cytogenetic characteristics. J Cell Physiol. 1976;89(3):441-55.

38. Nicolas JF, Dubois P, Jakob H, Gaillard J, Jacob F. [Mouse teratocarcinoma: differentiation in cultures of a multipotential primitive cell line (author's transl)]. Ann Microbiol (Paris). 1975;126(1):3-22.

39. McBurney MW. P19 embryonal carcinoma cells. Int J Dev Biol. 1993;37(1):135-40.

40. Jones-Villeneuve EM, McBurney MW, Rogers KA, Kalnins VI. Retinoic acid induces embryonal carcinoma cells to differentiate into neurons and glial cells. J Cell Biol. 1982;94(2):253-62.

41. Mullen RJ, Buck CR, Smith AM. NeuN, a neuronal specific nuclear protein in vertebrates. Development. 1992;116(1):201-11.

42. McBurney MW, Jones-Villeneuve EM, Edwards MK, Anderson PJ. Control of muscle and neuronal differentiation in a cultured embryonal carcinoma cell line. Nature. 1982;299(5879):165-7.

43. McBurney MW, Reuhl KR, Ally AI, Nasipuri S, Bell JC, Craig J. Differentiation and maturation of embryonal carcinoma-derived neurons in cell culture. J Neurosci. 1988;8(3):1063-73.

44. Rudnicki MA, Sawtell NM, Reuhl KR, Berg R, Craig JC, Jardine K, et al. Smooth muscle actin expression during P19 embryonal carcinoma differentiation in cell culture. J Cell Physiol. 1990;142(1):89-98.

45. Edwards MK, Harris JF, McBurney MW. Induced muscle differentiation in an embryonal carcinoma cell line. Mol Cell Biol. 1983;3(12):2280-6.

46. Jasmin, Spray DC, Campos de Carvalho AC, Mendez-Otero R. Chemical induction of cardiac differentiation in p19 embryonal carcinoma stem cells. Stem Cells Dev.19(3):403-12.

47. Brown K, Legros S, Artus J, Doss MX, Khanin R, Hadjantonakis AK, et al. A comparative analysis of extra-embryonic endoderm cell lines. PLOS ONE.5(8):e12016.

48. Monzen K, Shiojima I, Hiroi Y, Kudoh S, Oka T, Takimoto E, et al. Bone morphogenetic proteins induce cardiomyocyte differentiation through the mitogen-activated protein kinase kinase kinase TAK1 and cardiac transcription factors Csx/Nkx-2.5 and GATA-4. Mol Cell Biol. 1999;19(10):7096-105.

49. van der Heyden M, Defize L. Twenty one years of P19 cells: what an embryonal carcinoma cell line taught us about cardiomyocyte differentiation. CARDIOVASC RES. 2003;58(2):292-302.

50. Tan Y, Xie Z, Ding M, Wang Z, Yu Q, Meng L, et al. Increased levels of FoxA1 transcription factor in pluripotent P19 embryonal carcinoma cells stimulate neural differentiation. Stem Cells Dev. 2010;19(9):1365-74.

51. Xie Z, Tan G, Ding M, Dong D, Chen T, Meng X, et al. Foxm1 transcription factor is required for maintenance of pluripotency of P19 embryonal carcinoma cells. Nucleic Acids Res. 2010;38(22):8027-38.

52. Ma Y, Ramezani A, Lewis R, Hawley RG, Thomson JA. High-level sustained transgene expression in human embryonic stem cells using lentiviral vectors. Stem Cells. 2003;21(1):111-7.

53. Gropp M, Itsykson P, Singer O, Ben-Hur T, Reinhartz E, Galun E, et al. Stable genetic modification of human embryonic stem cells by lentiviral vectors. Mol Ther. 2003;7(2):281-7.

54. Tinsley RB, Faijerson J, Eriksson PS. Efficient non-viral transfection of adult neural stem/progenitor cells, without affecting viability, proliferation or differentiation. J Gene Med. 2006;8(1):72-81.

55. Zernecke A, Erl W, Fraemohs L, Lietz M, Weber C. Suppression of endothelial adhesion molecule up-regulation with cyclopentenone prostaglandins is dissociated from IkappaB-alpha kinase inhibition and cell death induction. Faseb J. 2003;17(9):1099-101.

56. Yan CN, Li F, Patterson C, Runge MS. High-voltage and high-salt buffer facilitates electroporation of human aortic smooth-muscle cells. Biotechniques. 1998;24(4):590-2.

57. Shimomura O, Johnson FH, Saiga Y. Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea. J Cell Comp Physiol. 1962;59:223-39.

58. Tsien RY. The green fluorescent protein. Annu Rev Biochem. 1998;67:509-44.

59. Miyawaki A, Llopis J, Heim R, McCaffery JM, Adams JA, Ikura M, et al. Fluorescent indicators for Ca2+ based on green fluorescent proteins and calmodulin. Nature. 1997;388(6645):882-7.

60. Giepmans BN, Adams SR, Ellisman MH, Tsien RY. The fluorescent toolbox for assessing protein location and function. Science. 2006;312(5771):217-24.

61. Liu CC, Schultz PG. Recombinant expression of selectively sulfated proteins in Escherichia coli. Nat Biotechnol. 2006;24(11):1436-40.

62. Yang W, Hendrickson WA, Crouch RJ, Satow Y. Structure of ribonuclease H phased at 2 A resolution by MAD analysis of the selenomethionyl protein. Science. 1990;249(4975):1398-405.

63. Mills JH, Lee HS, Liu CC, Wang J, Schultz PG. A genetically encoded direct sensor of antibody-antigen interactions. CHEMBIOCHEM. 2009;10(13):2162-4.

64. Summerer D, Chen S, Wu N, Deiters A, Chin JW, Schultz PG. A genetically encoded fluorescent amino acid. Proc Natl Acad Sci U S A. 2006;103(26):9785-9.

65. Umeda A, Thibodeaux GN, Moncivais K, Jiang F, Zhang ZJ. A Versatile Approach to Transform Low-Affinity Peptides into Protein Probes with Co-Translationally Expressed Chemical Cross-Linker. Anal Biochem.

66. Hino N, Okazaki Y, Kobayashi T, Hayashi A, Sakamoto K, Yokoyama S. Protein photo-cross-linking in mammalian cells by site-specific incorporation of a photoreactive amino acid. Nat Methods. 2005;2(3):201-6.

67. Wang W, Takimoto JK, Louie GV, Baiga TJ, Noel JP, Lee KF, et al. Genetically encoding unnatural amino acids for cellular and neuronal studies. Nat Neurosci. 2007;10(8):1063-72.

68. Pantoja R, Rodriguez EA, Dibas MI, Dougherty DA, Lester HA. Single-molecule imaging of a fluorescent unnatural amino acid incorporated into nicotinic receptors. Biophys J. 2009;96(1):226-37.

69. Beatty KE, Fisk JD, Smart BP, Lu YY, Szychowski J, Hangauer MJ, et al. Live-cell imaging of cellular proteins by a strain-promoted azide-alkyne cycloaddition. CHEMBIOCHEM. 2010;11(15):2092-5.

70. Beatty KE, Tirrell DA. Two-color labeling of temporally defined protein populations in mammalian cells. Bioorg Med Chem Lett. 2008;18(22):5995-9.

71. Dieterich DC, Hodas JJ, Gouzer G, Shadrin IY, Ngo JT, Triller A, et al. In situ visualization and dynamics of newly synthesized proteins in rat hippocampal neurons. Nat Neurosci.13(7):897-905.

72. Song W, Wang Y, Yu Z, Vera CI, Qu J, Lin Q. A metabolic alkene reporter for spatiotemporally controlled imaging of newly synthesized proteins in Mammalian cells. ACS Chem Biol.5(9):875-85.

73. Wang L, Brock A, Herberich B, Schultz PG. Expanding the genetic code of Escherichia coli. Science. 2001;292(5516):498-500.

74. Wang L, Schultz PG. Expanding the genetic code. Chem Commun (Camb). 2002(1):1-11.

75. Xie J, Schultz PG. An expanding genetic code. Methods. 2005;36(3):227-38.

76. Thibodeaux G, Liang X, Moncivais K, Umeda A, Singer O, Alfonta L, et al. Transforming a Pair of Orthogonal tRNA-aminoacyl-tRNA Synthetase from Archaea to Function in Mammalian Cells. PLOS ONE. 2010;5(6):-.

77. Ye S, Kohrer C, Huber T, Kazmi M, Sachdev P, Yan EC, et al. Site-specific incorporation of keto amino acids into functional G protein-coupled receptors using unnatural amino acid mutagenesis. J Biol Chem. 2008;283(3):1525-33.

78. Sakamoto K, Hayashi A, Sakamoto A, Kiga D, Nakayama H, Soma A, et al. Site-specific incorporation of an unnatural amino acid into proteins in mammalian cells. NUCLEIC ACIDS RES. 2002;30(21):4692-9.

79. Zhang Z, Alfonta L, Tian F, Bursulaya B, Uryu S, King DS, et al. Selective incorporation of 5-hydroxytryptophan into proteins in mammalian cells. Proc Natl Acad Sci U S A. 2004;101(24):8882-7.

80. Schimmel P, Giege R, Moras D, Yokoyama S. An operational RNA code for amino acids and possible relationship to genetic code. Proc Natl Acad Sci U S A. 1993;90(19):8763-8.

81. Giege R, Puglisi JD, Florentz C. tRNA structure and aminoacylation efficiency. Prog Nucleic Acid Res Mol Biol. 1993;45:129-206.

82. Wakasugi K, Quinn CL, Tao N, Schimmel P. Genetic code in evolution: switching species-specific aminoacylation with a peptide transplant. Embo J. 1998;17(1):297-305.

83. Quinn CL, Tao N, Schimmel P. Species-specific microhelix aminoacylation by a eukaryotic pathogen tRNA synthetase dependent on a single base pair. Biochemistry. 1995;34(39):12489-95.

84. Giege R, Sissler M, Florentz C. Universal rules and idiosyncratic features in tRNA identity. Nucleic Acids Res. 1998;26(22):5017-35.

85. Alexander RW, Nordin BE, Schimmel P. Activation of microhelix charging by localized helix destabilization. Proc Natl Acad Sci U S A. 1998;95(21):12214-9.

86. Kobayashi T, Nureki O, Ishitani R, Yaremchuk A, Tukalo M, Cusack S, et al. Structural basis for orthogonal tRNA specificities of tyrosyl-tRNA synthetases for genetic code expansion. Nat Struct Biol. 2003;10(6):425-32.

87. Steer BA, Schimmel P. Major anticodon-binding region missing from an archaebacterial tRNA synthetase. J Biol Chem. 1999;274(50):35601-6.

88. Webster T, Tsai H, Kula M, Mackie GA, Schimmel P. Specific sequence homology and three-dimensional structure of an aminoacyl transfer RNA synthetase. Science. 1984;226(4680):1315-7.

89. Fabrega C, Farrow MA, Mukhopadhyay B, de Crecy-Lagard V, Ortiz AR, Schimmel P. An aminoacyl tRNA synthetase whose sequence fits into neither of the two known classes. Nature. 2001;411(6833):110-4.

90. Himeno H, Hasegawa T, Ueda T, Watanabe K, Shimizu M. Conversion of aminoacylation specificity from tRNA(Tyr) to tRNA(Ser) in vitro. Nucleic Acids Res. 1990;18(23):6815-9.

91. Bedouelle H. Recognition of tRNA(Tyr) by tyrosyl-tRNA synthetase. Biochimie. 1990;72(8):589-98.

92. Bedouelle H, Guez-Ivanier V, Nageotte R. Discrimination between transfer-RNAs by tyrosyl-tRNA synthetase. Biochimie. 1993;75(12):1099-108.

93. Bedouelle H, Winter G. A model of synthetase/transfer RNA interaction as deduced by protein engineering. Nature. 1986;320(6060):371-3.

94. Phizicky EM, Fields S. Protein-protein interactions: methods for detection and analysis. Microbiol Rev. 1995;59(1):94-123.

95. Dang CV, Barrett J, Villa-Garcia M, Resar LM, Kato GJ, Fearon ER. Intracellular leucine zipper interactions suggest c-Myc hetero-oligomerization. Mol Cell Biol. 1991;11(2):954-62.

96. Luo Y, Batalao A, Zhou H, Zhu L. Mammalian two-hybrid system: a complementary approach to the yeast two-hybrid system. Biotechniques. 1997;22(2):350-2.

97. Schmidt JW, Brugge JS, Nelson WJ. pp60src tyrosine kinase modulates P19 embryonal carcinoma cell fate by inhibiting neuronal but not epithelial differentiation. J Cell Biol. 1992;116(4):1019-33.

98. AG LC. Amaxa (R) Cell Line Nucleofector Kit V.

99. Umeda A, Thibodeaux G, Moncivais K, Jiang F, Zhang Z. A versatile approach to transform low-affinity peptides into protein probes with cotranslationally expressed chemical cross-linker. ANALYTICAL BIOCHEMISTRY. 2010;405(1):82-8.

100. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P, editors. Molecular Biology of the Cell. 4 ed. New York: Garland Science; 2002.

101. Kisselev LL. The role of the anticodon in recognition of tRNA by aminoacyl-tRNA synthetases. Prog Nucleic Acid Res Mol Biol. 1985;32:237-66.

102. Chambers RW. On the recognition of tRNA by its aminoacyl-tRNA ligase. Prog Nucleic Acid Res Mol Biol. 1971;11:489-525.

103. Ebel JP, Giege R, Bonnet J, Kern D, Befort N, Bollack C, et al. Factors determining the specificity of the tRNA aminoacylation reaction. Non-absolute specificity of tRNA-aminoacyl-tRNA synthetase recognition and particular importance of the maximal velocity. Biochimie. 1973;55(5):547-57.

104. Shimura Y, Ozeki H. Genetic study on transfer RNA. Adv Biophys. 1973;4:191-226.

105. Buechter DD, Schimmel P. Aminoacylation of RNA minihelices: implications for tRNA synthetase structural design and evolution. Crit Rev Biochem Mol Biol. 1993;28(4):309-22.

106. Francklyn C, Schimmel P. Aminoacylation of RNA minihelices with alanine. Nature. 1989;337(6206):478-81.

107. Martinis SA, Schimmel P. Enzymatic aminoacylation of sequence-specific RNA minihelices and hybrid duplexes with methionine. Proc Natl Acad Sci U S A. 1992;89(1):65-9.

108. Shi JP, Francklyn C, Hill K, Schimmel P. A nucleotide that enhances the charging of RNA minihelix sequence variants with alanine. Biochemistry. 1990;29(15):3621-6.

109. Shi JP, Martinis SA, Schimmel P. RNA tetraloops as minimalist substrates for aminoacylation. Biochemistry. 1992;31(21):4931-6.

110. Shi JP, Schimmel P. Aminoacylation of alanine minihelices. "Discriminator" base modulates transition state of single turnover reaction. J Biol Chem. 1991;266(5):2705-8.

111. Juhling F, Morl M, Hartmann RK, Sprinzl M, Stadler PF, Putz J. tRNAdb 2009: compilation of tRNA sequences and tRNA genes. Nucleic Acids Res. 2009;37(Database issue):D159-62.

112. Beuning P, Musier-Forsyth K. Transfer RNA recognition by aminoacyl-tRNA synthetases. BIOPOLYMERS. 1999;52(1):1-28.

113. Boden D, Pusch O, Lee F, Tucker L, Shank PR, Ramratnam B. Promoter choice affects the potency of HIV-1 specific RNA interference. Nucleic Acids Res. 2003;31(17):5033-8.

114. LaFemina R, Hayward GS. Constitutive and retinoic acid-inducible expression of cytomegalovirus immediate-early genes in human teratocarcinoma cells. J Virol. 1986;58(2):434-40.

115. Gonczol E, Andrews PW, Plotkin SA. Cytomegalovirus replicates in differentiated but not in undifferentiated human embryonal carcinoma cells. Science. 1984;224(4645):159-61.

116. Tanaka Y, Bond M, Kohler J. Photocrosslinkers illuminate interactions in living cells. MOL BIOSYST. 2008;4(6):473-80.

Vita

Katy was born and raised in Bryan/College Station, Texas, home of the Fighting Texas Aggies. She attended Bryan High School before beginning her degree in biomedical engineering at Rice University. While at Rice she interned at the Texas Heart Institute under the supervision of Dr. Doreen Rosenstrauch, working on a cellular engineering project. After graduating from Rice, Katy enrolled at UT Austin where she rotated through several labs before finding a permanent home in the Zhang Lab under the supervision of Dr. Zhiwen Jonathan Zhang. She is interested in cellular and molecular engineering and their application to gene therapy and tissue engineering.

Permanent email: katy.l.m@gmail.com

This dissertation was typed by Kathryn Lauren Moncivais.