7-20-2004

Development Of A DNA Probe Assay For Rare Transfer RNAs And Its Use In Measuring The Expression Of The ArgX Gene In Escherichia Coli

Katie M. Cavnar

THE FLORIDA STATE UNIVERSITY

COLLEGE OF ARTS AND SCIENCES

DEVELOPMENT OF A DNA PROBE ASSAY FOR RARE TRANSFER RNAS AND ITS USE IN MEASURING THE EXPRESSION OF THE *ARGX* GENE IN *ESCHERICHIA COLI*

By

KATIE CAVNAR

A Thesis submitted to the
Department of Biological Science
in partial fulfillment of the
requirements for the degree of
Master of Science

Degree Awarded:
Fall Semester, 2004

The members of the Committee approve the Thesis of Katie Cavnar defended on July 20, 2004.

Robert H. Reeves
Professor Directing Thesis

Thomas Keller
Committee Member

Lloyd Epstein
Committee Member

Approved:

Timothy Moerland, Chair, Department of Biological Science

The Office of Graduate Studies has verified and approved the above named committee members.

For my husband Peter, for believing in me and giving me Grace

ACKNOWLEDGEMENTS

I would foremost like to thank Dr. Robert Reeves for taking me as a graduate student, allowing for my unusual circumstances those first years and for helping me always with patience and kindness. Thanks to his wife Jane for her friendship and generosity and for always keeping the candy basket full. Thank you Lauren Porter for all your help through the years and the help that will come in the years ahead. I would have made it through graduate school without you but it wouldn't have been half as fun. Thanks to Megan Lowenberg for your continual optimism and encouragement, especially near the end. I would like to acknowledge and thank my committee members, Drs. Keller and Epstein for their time and their advice concerning this project

I would like to thank my boss Carolyn Lewis at Tallahassee Memorial Hospital for allowing me the chance to attend to graduate school by forever working around my schedule when I was full time and giving me hours when I was part time. I would like to thank the Ouimet lab, Dr. Charles Ouimet for hiring me and working around my school schedule and to the graduate students Karen and David Dietz, Orenda Johnson and Mark Darcy for their help with questions, lab supplies and most important their time. A special thanks to Claire Griffin for her help with the plates.

I thank my family and friends for their encouragement and support; I couldn't have done it without them.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

In *Escherichia coli*, the *argX* gene is the first gene in a cluster with three other tRNA genes. The tRNA products of the three downstream genes, *hisR*, *leuT* and *proM* are abundant in *E. coli*, and all four of the tRNAs are processed from the same precursor RNA transcript. The *argX* promotor appears to be a strong promotor and is preceded by an Up element. By all accounts, the gene product of *argX*, tRNAarg3 should be a major tRNA. Like the three downstream genes in this cluster, *argX* must be transcribed at a high level because of its location at the beginning of the gene cluster. However, tRNAarg3 is considered a rare tRNA in *E. coli*. Its low level is correlated with the low codon usage of its corresponding codon, CGG. It appears that the nucleotide sequence just upstream of *argX* can base pair with the 5'- end of the tRNA. This alternative secondary structure in the precursor RNA transcript may prevent the maturation of the tRNAarg3 by disrupting the base pairing of the aminoacyl-tRNA stem. This structure is required by ribonuclease P for maturation of the 5'-ends of all tRNAs and also prevents the 3'-exonucleases from digesting into the tRNA from the CCA (3'-) end. Different constructs of the *argX* gene have been cloned into high copy number plasmids. These constructs differ in the nucleotides just upstream of the *argX* gene. Deletion of nucleotides results in higher expression of tRNAarg3 and the loss of an alternative secondary structure involving the 5'–end of the tRNA. The data suggests alternative secondary structure formation during transcription may in part be responsible for tRNAarg3 abundance in *E. coli.* A tRNA assay has been developed to measure the levels of rare tRNAs from extracts. This assay is both sensitive and specific and demonstrates its ability to detect low levels of tRNAs.

INTRODUCTION

Transfer RNA (tRNA) is the adapter molecule in protein synthesis. It interacts with a codon, a three-nucleotide sequence in the mRNA, which base pairs with the tRNA's complementary anticodon. A specific amino acid is covalently bound to the 3'-end of the tRNA with the help of an aminoacyl-tRNA synthetase, one of 20 enzymes that "activate" each of the amino acids (Lodish et al., 1998). The activated amino acid forms an amino acid adenylate (AA-AMP) and is bound to the synthetase until it collides into a tRNA molecule specific for that particular amino acid. Once the impact occurs, the aminoacyl-tRNA synthetase transfers the amino acid to the terminal adenosine of the tRNA and the tRNA is now charged, becoming an aminoacylated-tRNA (Watson et al., 1987). On the ribosome the aminoacylated-tRNA then binds to the codon in the mRNA via the anticodon and transfers its attached amino acid to the growing polypeptide chain. This process is called translation because the sequence of nucleotides on the mRNA is translated into a sequence of amino acids that are bound together to create a protein (Snyder and Champness, 1997).

The nucleotide sequences vary distinctly for different tRNA genes, but all have in common four or five inverted repeats that define the regularly spaced stem loop elements of tRNA molecules known as the clover leaf model devised by Holley et al. (1965). There are several well-conserved structures of the tRNA cloverleaf model (Figure 1). The 5' end always has a 5' terminal monophosphate group, usually guanylic acid. The 5' end base pairs with the 3' end to form the aminoacyl stem. The D loop has eight to twelve unpaired bases and contains the modified base dihydro-U. The anticodon loop contains seven unpaired bases, which includes the anticodon, three adjacent bases that bind to successive bases that make up a codon in mRNA. The anticodon is flanked by a purine on the 3'end and a uridine on the 5' end. There is always a variable loop, although highly variable in size from one tRNA to the next. The TΨ (pseudouridine) C loop contains seven unpaired bases and almost always contains the 5' to 3' sequence TΨC. The 3' terminus of all tRNAs end in CCA plus a fourth variable nucleotide, which extends beyond the aminoacyl stem. (Watson et al., 1987) The function of the CCA trinucleotide

sequence is to position the attached amino acid in the peptidyl transferase center of the ribosome, allowing for transpeptidation to occur during translation (O'Connor et al., 1993). The aminoacyl-tRNA synthetase correctly aminoacylates its analogous tRNA by recognizing the anticodon and CCA stem in addition to occasionally recognizing nucleotides in the variable stem loop (Saks and Sampson, 1995).

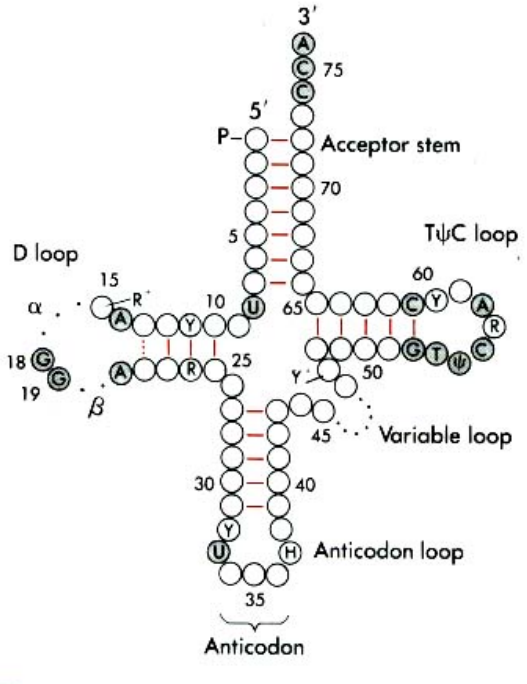


Figure 1. Secondary structure of tRNAs. Conserved base pairs are shaded in this cloverleaf model of tRNA (Watson et al., 1987). .
R = purine, Y = pyrimidine, H = modified pyrimidine

The tertiary structure of tRNA consists of double helical stems that conform to an A type helix structure expected for RNA:RNA duplexes (Watson et al., 1987). Additional hydrogen bonds bend the cloverleaf structure into a stable tertiary structure (Figure 2). Conserved bases that exist in all tRNAs are involved in tertiary hydrogen bonding activities showing that all tRNAs basically have the same conformation. The

structure resembles an upside down "L" (Watson et al., 1987; Mans et al., 1991) with the amino acid acceptor CCA group at one end of the "L" opposite the anticodon. The D and TΨC loops form the corners. Similarity between tRNA tertiary structures is so great that only the angle between the two helical arms of the "L" vary slightly creating a hinge formation that may be flexible (Watson et al., 1987). This logically explains the fact that all tRNAs attach onto ribosomes. However, there must be a slight variation between the tertiary structures if particular selection enzymes (aminoacyl-tRNA synthetases, RNases, exonucleases, etc) can differentiate between tRNAs (Watson et al., 1987).

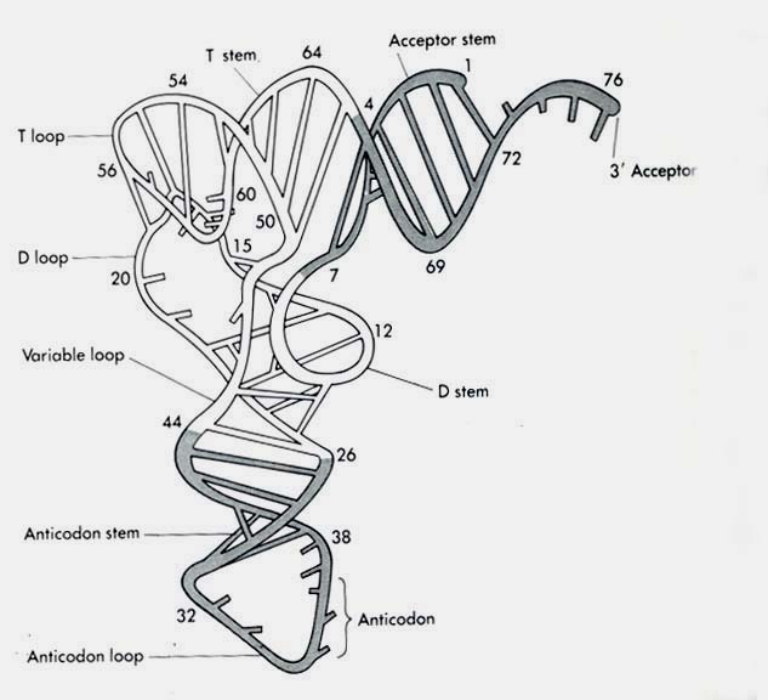


Figure 2. Tertiary structure of tRNAs. The anticodon and acceptor arm are shaded in this tertiary structure of tRNA (Watson et al., 1987).

Escherichia coli tRNA genes and promotors

The genome of *Escherichia coli* contains 86 tRNA genes. Each gene is 75 to 95 base pairs long. These genes lack introns and they carry a CCA sequence at the 3'

terminus. Of the 86 genes, most occur in clusters. Only 20 exist as a single genes and the remaining genes are organized as polygenic clusters containing identical or unrelated tRNAs, rRNAs or even protein coding genes (Fournier and Ozeki, 1985). In *E. coli*, redundancy of tRNA genes differentiates them from protein coding genes because of this common occurrence of multiple gene copies and gene copy numbers perform an important role in assuring an equalized population of acceptor tRNA species (Fournier and Ozeki, 1985).

Gene copy number and promotor efficiency are important determinants of tRNA levels (Bauer et al., 1993). In general the promoters for tRNA genes are consensus, strong promoters. They are typical σ^{70} types, with –35 and –10 elements found in the upstream region of the transcriptional units (Travers, 1984; Jinks-Robertson, 1987; Inokuchi and Yamao, 1994), meaning that the –35 and –10 sequences are recognized by the σ^{70} subunit of *E. coli* RNA polymerase (Ross et al., 1998). The space between –35 and –10 ranges from 15 to 18 nucleotides (Inokuchi and Yamao, 1994). In addition many tRNA promoters have an Up element, a sequence of about 20 A-T rich base pairs upstream of the -35 site (Ross et al., 1993; Keener and Nomura, 1996). These elements enhance transcription about 30-fold (Ross et al., 1993). The canonical sequences of Up elements contain two conserved regions: 1) from 5' -57 to –47 the sequence is AAAa/ta/tTa/tTTTT and 2) from –44 to –41 the sequence is AAAA (Ross 1998). The terminators are usually Rho-independent. Rho is a 50 kilodaltan (kDa) protein that can be involved in termination of prokaryotic transcription, although the mechanism of how it exactly works is unknown (Lodish et al., 1998).

Different promotor strengths and/or differences in the survival of various processed intermediate precursors can account for different levels of mature tRNA produced (Watson et al., 1987). All mature *E. coli* tRNAs are derived by the processing of larger precursor molecules containing extensions of accompanying nucleotides at both the 5' and 3' ends of the tRNA transcript (Fournier and Ozeki, 1985; Li and Deutscher, 1994). These precursors can be unstable and short-lived and can also contain anywhere from one to seven different tRNAs in one transcript. The maturation of tRNAs from their newly transcribed precursors relies on a minimum of eight to ten enzymes involved in converting the primary transcript into the mature tRNA (Fournier and Ozeki, 1985). A failure to

process tRNAs correctly can result in the termination of protein synthesis and cell growth (Watson et al., 1987; Ow and Kushner, 2002).

tRNA maturation

There are two types of enzymatic activity involved in processing the precursor tRNA transcript. Exoribonucleases, such as RNases II and T, are responsible for the removal of excess nucleotides at the 3'-end. Endoribonucleases, such as RNases P and E, are responsible for the maturation of both the 5' and 3' ends, respectively. The 3'-exoribonucleases catalyze the hydrolysis of the tRNA precursors in the 3' to 5' direction by removing one nucleotide at a time until they reach the CCA sequence at the 3' terminus (Watson et al., 1987; Deutscher, 1994) and are responsible for the final 3'-end processing of tRNA precursors. The secondary structure of the aminoacyl-tRNA stem is presumed to be the major factor that halts the progression of these 3' exoribonucleases, keeping them from destroying the tRNA (Deutscher, 1984; Deutscher, 1994). Endoribonucleases make one cut at a critical place on the tRNA precursor transcript. There are two types of tRNA precursors. The type-I precursor already has the CCA sequence present at the 3' end and needs and exo- and/or endoribonuclease for processing. Type-II precursors do not have a CCA sequence and therefore require the enzyme tRNA nucleotidyltransferase to insert the three nucleotides. tRNA nucleotidyltransferase uses CTP and ATP to synthesize the CCA terminus even though there is no nucleotide template available (Deutscher, 1990). Some prokaryotic genes encode for the CCA sequence, some do not and some have multiple genes of both types. All *E. coli* tRNA genes do encode for the CCA sequence.(Deutscher, 1990)

There are several 3'-exoribonucleases found in *E. coli* and other prokaryotes and eukaryotes: RNases II, D, BN, T, PH and PNPase (polynucleotide phosphorylase) which all remove nucleotides from the 3' end one way or another (Li and Deutscher, 1994). RNase II and PNPase shorten longer 3' sequences to two to four extra 3' residues. RNase T and PH trim the last few nucleotides of the precursor residues making them responsible for helping mature the 3' end of the final tRNA product. (Li and Deutscher, 1994; Li and Deutscher, 1996) RNase D and BN have less major roles. Cells lacking in RNase II, D,

BN, T and PH are inviable, however, the removal of any one of these enzymes alone will still lead to tRNA processing and maturation at fluctuating degrees of effectiveness (Li and Deutscher, 1994) and mutant strains of *E. coli* deficient in RNase D, II and BN still produced mature tRNA 91-100% of the time (Li and Deutscher, 1996).

Endonucleolytic cleavage is also required for tRNA maturation. tRNA precursors contain long, highly structured 3' trailer sequences that are difficult to remove by 3'-exoribonucleases alone. It is also necessary to separate tRNAs from each other in polycistronic operons by endonucleolytic cleavage. Endonucleolytic cleavage occurs rapidly, during or immediately following transcription (Deutscher, 1984). Because RNase P is inhibited by the occurrence of a long 3' precursor sequence, 3' endonucleolytic cleavage is required for 5' maturation (Li and Deutscher, 2002). RNase E is a 3'endoribonuclease essential for cell viability (Li and Deutscher, 2002; Ow and Kushner, 2002). Processing of the *argX* cluster as well as other polycistronic operons and monocistronic transcripts begins with RNase E; furthermore it is the rate-limiting step in maturation of tRNA (Ow and Kushner, 2002). In an experiment done by Ow et al. (2002), it was shown that inactivation of RNase E, using mutated RNase E genes, leads to an increase in the amount of precursor tRNAs for polycistronic operons and monocistronic transcripts. Before RNase P and the 3'to 5' exoribonucleases can complete tRNA maturation, cleavage by RNase E, usually within a few nucleotides of the CCA sequence, is required; therefore RNase P cleavage at the 5' site of the tRNA precursor is dependent on prior RNase E processing (Ow 2002). An experiment done by Li and Deutscher (2002) using two $tRNA^{TYR}$ separated by a spacer region spanning 209 nucleotides demonstrated that when RNase E and P were absent, a product containing both tRNAs accumulated.

RNase P must recognize common features of tRNA structure in order to mature the 5' end (Abelson, 1979; Deutscher, 1984; Deutscher, 1994). It was first detected in the *E. coli* $tRNA^{TYR}$ 5'-end precursor due to its enzymatic activity of removing the 5' nucleotide sequence prior to the tRNA gene (Deutscher, 1984; Deutscher, 1994). It makes a single cut, producing a 5' phosphate end of the tRNA (Watson et al., 1987). RNase P is a ribonucleoprotein made up of a small protein and RNA of 375-377 nucleotides (M1 RNA), which are both necessary for enzyme activity. The M1 RNA has a 5' nucleotide sequence in a loop region that is complementary to the 5'-GTΨCPi-3' loop found in all *E.*

coli tRNAs and a second sequence was found to interact with a common sequence in the dihydrouridine loop (Deutscher, 1984). These sequences suggest that the structures of the tRNA domain itself rather than the precursor regions are what are crucial for RNase P recognition (Deutscher, 1984). Furthermore, secondary and tertiary structure of tRNAs is of the up most importance since mutations that alter base pairing in the aminoacyl stem and effect overall tRNA conformation can have substantial consequences on RNase P activity (Deutscher, 1984). Overall, it is critical for the appropriate aminoacyl-tRNA synthetase to be able to recognize its specific, mature tRNA in order for accurate translation of mRNA from nucleic acids into proteins (Perret et al., 1990).

Gene copy number

There is normally a good correspondence between gene copy number and tRNA content. For major tRNAs that are found abundantly in *E. coli*, the corresponding genes may be triplicated and quadruplicated. There are no duplicated genes for minor tRNAs. For instance tRNAala1B, a major tRNA, is located in three *rrn* clusters, which probably have the most efficient promoters in *E. coli*. Not only do these promoters have consensus sequences, there are Up elements preceding them, which enhance their efficiencies. Another example of this is found in the leucine family of tRNAs in *E. coli*, there are five different tRNAs with different anticodons. Only one of these is a major tRNA, tRNAleu1, and it is transcribed from four separate genes, leuV, leuP, leuQ and leuT. The other four tRNAs, considered to be in minor amounts are all transcribed from single genes. There are differences in the levels of these other four leucine tRNAs. TransferRNAleu3 is considered a rare tRNA, and the others, tRNAleu2, tRNAleu4 and tRNAleu5 are found in moderate amounts. These differences are assumed to be due to differences in the transcriptional efficiencies or strength of their respective promoters. Therefore, the abundance of individual tRNAs in the cell depends on both gene dosage and the strength of the promoters from which they are transcribed.

Codon usage

The concentration of tRNA species can be rate-limiting for protein synthesis and it is important in determining the variation of translation rates at individual codons

(Anderson, 1969; Kurland, 1991). It has been demonstrated that tRNA levels in the cell correlate with codon usage (Ikemura, 1981a; Gouy and Gautier, 1982; Brinkmann et al., 1989; Sorensen et al., 2003) (Table 1). tRNAs for minor codons exist in a very limited amount, even in exponentially growing cells (Ikemura, 1981a). Codons occur in triplets, which contain three ribonucleotides specific for one amino acid. The codon is recognized by its cognate tRNA as a result of the corresponding anticodon in the tRNA sequence. Codons recognized by abundant tRNAs are being used more often than those recognized by rare tRNA, especially in highly expressed genes, like the ribosomal protein genes (Ikemura, 1981b; Gouy and Gautier, 1982; Bulmer, 1988). This is known as codon bias. This biased codon usage may exercise selective pressure on tRNA levels and could have

Table 1. *Escherichia coli* codon usage table (www.kazusa.or.jp/codon/).

fields: [triplet] [frequency: per thousand] ([number])

UUU 21.8(92794)	UCU 9.2(39368)	UAU 16.4(69657)	UGU 5.1(21880)
UUC 16.6(70740)	UCC 8.9(37939)	UAC 12.3(52559)	UGC 6.4(27050)
UUA 13.6(57701)	UCA 7.7(32822)	UAA 2.0(8397)	UGA 1.0(4067)
UUG 13.2(56041)	UCG 8.7(36910)	UAG 0.3(1081)	UGG 14.4(61189)
CUU 11.2(47576)	CCU 7.1(30220)	CAU 12.6(53787)	CGU 20.9(89058)
CUC 10.6(45276)	CCC 5.3(22586)	CAC 9.7(41487)	CGC 21.4(90911)
CUA 3.9(16619)	CCA 8.5(35974)	CAA 14.6(62363)	CGA 3.6(15316)
CUG 51.5(219391)	CCG 22.6(96363)	CAG 29.0(123370)	CGG 5.6(23884)
AUU 29.7(126636)	ACU 9.5(40314)	AAU 18.3(77780)	AGU 9.0(38223)
AUC 24.9(106060)	ACC 22.9(97419)	AAC 21.8(92629)	AGC 15.6(66538)
AUA 5.1(21651)	ACA 7.7(32962)	AAA 34.4(146383)	AGA 2.6(11131)
AUG 27.3(116328)	ACG 14.0(59631)	AAG 11.3(47980)	AGG 1.6(6686)
GUU 19.0(80902)	GCU 16.1(68716)	GAU 32.2(137042)	GGU 25.3(107559)
GUC 14.9(63252)	GCC 25.1(106678)	GAC 19.5(82941)	GGC 28.9(123244)
GUA 11.2(47715)	GCA 20.5(87304)	GAA 39.8(169392)	GGA 8.4(35950)
GUG 25.7(109400)	GCG 32.6(138586)	GAG 18.5(78669)	GGG 11.1(47128)

Coding GC 51.54% 1st letter GC 58.69% 2nd letter GC 40.72% 3rd letter GC 55.20%

resulted from the accumulation of a large number of mutations occurring in both proteins and tRNA genes (Ikemura, 1981b; Bulmer, 1988). Codon bias and tRNA isoacceptor concentrations have coevolved and the correlation is more pronounced in genes with higher expression levels (Elf et al., 2003). Another reason for codon bias is due to non-random choices between pyrimidine ending codons. For example, if the first two bases are both A or U, C is the more likely choice for the third position than U (Gouy and Gautier, 1982). If the first two bases are either C or G, than U is usually chosen over C (Gouy and Gautier, 1982). These choices involve the energy levels of codon-anticodon interactions (Gouy and Gautier, 1982). Codons that base pair to more abundant tRNA species form nearly perfect Watson-Crick base pairs (Fuglsang, 2003). These codons are "optimal" codons (Ikemura, 1985; Fuglsang, 2003). Several observations about codon usage have been made. There is a bias toward one or two codons for almost all synonymous codon families and certain codons are more frequently used by many different genes, regardless of abundance in protein (Makrides, 1996). Also, highly expressed genes demonstrate a greater degree of codon bias than poorly expressed ones (Makrides, 1996; Kanaya et al., 2001). Finally, to reiterate, the frequency of use of synonymous codons usually correlates to the abundance of their cognate tRNAs.

Rare codon gene expression can result in translational errors that can lead to ribosomal stalling at a position requiring the addition of an amino acid coupled to a minor tRNA (Wu et al., 2004). This can cause codon misreading, which can manifest as amino acid substitutions or translation errors, particularly in rare arginine, leucine, isoleucine and proline codons, all leading to unfavorable proteins (Wu et al., 2004). A decrease in RNA stability correlated to the amount of rare codons in yeast has been demonstrated to be a major contribution of lower expression levels of certain genes (Wu et al., 2004). Then why doesn't every gene use codons that are specific for abundant tRNAs? There may be restrictions because of peptide function. Certain sequences of amino acids may be required for the function and structure of the protein being translated (Ikemura, 1981a).

There is a considerable effect of codon position in highly expressed genes (Bulmer, 1988), which can have different effects of level of protein production (Zdanovsky and Zdanovskaia, 2000). Most *E. coli* proteins preferentially use minor codons in the first 25 codons of the transcript (Chen and Inouye, 1990), however the

presence of rare codons too close to the translational start point can lead to strong negative effects on genes expression (Zahn and Landy, 1996). For example, slow translation through low usage codons early in the mRNA transcripts may allow ribosomes to detach before they finish reading (Goldman et al., 1995). This is also dependent on the number of repeat rare codons found early in the transcript. Chen and Inoye (1990) found that a rare codon, AGA (tRNA *argU*) placed in five repeats at the beginning of the transcript caused an inhibition in the production of beta galactosidase, however when the repeats were moved further down the transcript, there was an increase in the level of protein.

Heterologous Proteins

The recent interest in rare tRNAs of *E. coli* is the result of problems that have been encountered with the expression of heterologous proteins. Expression of heterologous genes in *E. coli* is one of the most frequently used techniques in the laboratory and industry. A number of heterologous proteins have been made in *E. coli*. Some of these have been human proteins, such as human growth hormone, that have had great therapeutic value. These proteins are expressed from recombinant plasmids that contain the heterologous gene behind a compatible promoter. An efficient promotor for gene expression should be strong, strictly regulated, easily transferable to other *E. coli* strains and its induction should be simple and cost effective (Makrides, 1996). Because codon usage differs between humans (or other organisms) and *E. coli*, the expression of heterologous proteins that demand a high use of rare codons quickly exhausts the supply of the corresponding rare tRNA and the protein cannot be expressed (Imamura H, 1999; Sorensen et al., 2003). Factors that effect production of proteins from heterologous genes in *E. coli* include biased codon usage leading to truncated proteins, gene product toxicity, mRNA secondary structure and stability and proteolytic degradation. (Kane, 1995; Makrides, 1996; Sorensen et al., 2003; Wu et al., 2004). Since rare *E. coli* codons are often preferred in other organism genes being used to create heterologous proteins, the limited existence of cognate charged tRNAs can cause translational stalling, frame shifts or amino acid misincorporation (Imamura et al., 1999; Sorensen et al., 2003). This gives rise to translational termination prior to complete protein synthesis leading to poor expression of full-length heterologous proteins. In a number of cases this problem is

solved just by raising the level of the rare tRNA, i.e., inserting a copy of the rare cognate tRNA onto a plasmid gene behind a strong promoter (Chen and Inouye, 1994; Sorensen et al., 2003). This simple solution to the "codon usage problem" has allowed the expression of a number of recombinant proteins in *E. coli,* and has become an invaluable tool for basic research and industrial applications. An experiment by Acosta-Rivero et al. (Acosta-Rivero et al., 2002) showed that an increase in the amount of the rare tRNA that binds codons AGA and AGG (tRNAs *argU* and *argW* respectively), caused an increase in the amount of human interferon IFNαC and in the hepatitis C viral core protein. In another experiment by Tan et al. (Tan et al., 2003), the supplementation of the rare tRNA *argU* (coding for AGA) increased the yield of the hepatitis B core antigen protein (HbcAg) in *E. coli.* Data from an experiment by Zdanovsky and Zdanovskaia (2000) demonstrated that amplification of certain tRNA genes for rare codons in *E. coli* improved the expression of clostridia neurotoxins. However, not all of these rare tRNA genes caused an improvement. There may have been several other factors afore mentioned that could have been inhibiting the expression of these neurotoxins as well. Another solution is to use site directed mutagenesis by constructing codons in the target sequence of the heterologous protein that reflect the tRNA pool in the host (Sorensen et al., 2003). In an experiment by Ma (Ma et al., 1993), the valine codon at position 75 was mutated from GUG to AGG (a rare arginine codon) in a heterologous protein dihydrofolate reductase (DHFR) gene using site-directed mutagenesis by PCR. They then created a synthetic tRNA with the anticodon CCU on a plasmid by altering the anticodon of the $tRNA^{Ala/UCG}$ gene from *E. coli* to $tRNA^{Ala/CCU}$. They expressed both the mutated and wild type DHFR proteins in *E. coli*. There was no discernable difference between the amounts of protein expressed in the mutated DHFR gene compared to the amount made using the wild type gene. Misincorporation of one codon for another rare codon can also occur in the expression of heterologous proteins. An experiment by Calderon et al. (Calderone et al., 1996) showed the misincorporation of a lysine codon for a rare arginine codon (AGA) in a heterologous protein being expressed in *E. coli*. This error was corrected by replacing the minor AGA codons in the sequence with a more abundant arginine codon (CGC), which in turn increased the production of the protein. They also helped to resolve the misincorporation error by co-expressing the *argU* tRNA gene on a plasmid.

argX and the argX operon

There are several codons in E. coli that are used very rarely, 5 in 1000 codons or less. In each case the single tRNA that reads these codons is considered to be a rare tRNA and in fact cannot be easily detected in RNA extracts. Three of these rarely used codons code for arginine. The three corresponding tRNA genes are single genes - *argX*, *argU* and *argW* -, and their gene products - tRNAarg3, tRNAarg4 and tRNAarg5 - are correspondingly found at very low or undetectable levels in extracts of *E. coli.* The genes for *argU* and *argW* found as single genes are assumed to have very inefficient promoters. The *argX* gene is found in a gene cluster with three other tRNA genes that have tRNAs in moderate and even major amounts, and therefore the reasons for the low levels of tRNAarg3 are more obscure.

In the cell, not all tRNAs that are transcribed from the same cluster exist at the same levels. Consequently, the sequence of the tRNA gene arrangement in the cluster does not reflect the levels of the gene products. This suggests that intracellular tRNA levels can be regulated, not only by promoter efficiencies for the tRNA genes, but also by differential processing of tRNA precursors. In a study on the use of anticodons in *E. coli*, the anticodon of *argX,* CCG, was used at the low rate 0.36% (Inokuchi and Yamao, 1994). Therefore the levels of intracellular tRNAs can be directly related to the processes that are affected by the predetermined overall codon usage of mRNAs (Inokuchi and Yamao, 1994). The abundance of the tRNAarg3 in *E. coli* is difficult to explain in terms of the location of its gene *argX*. It is the first gene in a cluster of four tRNA genes known as the *argX* gene cluster or operon, and is followed by *hisR*, *leuT* and *proM* respectively (Figure 3). The *argX* operon was isolated by direct cloning and hybridization. It was located on the chromosome with the help of prophage mapping data at 85.5 minutes (Fournier and Ozeki, 1985). The *argX* transcript is encoded first in the operon. The entire operon is about 480 nucleotides long (Hsu et al., 1984). The cluster is transcribed from a strong promotor that is preceded by an Up element. The –10 and –35 promotor elements are both in good agreement with consensus sequences for prokaryotic promotors. They are separated by 17 base pairs, which is optimal for most bacterial gene promotors (Hsu et al., 1984). The tRNAs from the latter three genes in this cluster are found in medium to high abundance in *E. coli,* corresponding to their respective codon usage, whereas the *argX*

gene product, tRNAarg3 is not (Hsu et al., 1984). It is one of the rarest tRNAs in *E. coli*, which is associated with the low codon usage of its matching codon, CGG. Hsu et al. (1984) described evidence showing all four genes are co-transcribed indicating that there should be a more balanced abundance of all four tRNAs and stated the need for a more in depth study on the expression and turnover rates of these tRNAs. Analysis of the sequence of the transcript directly preceding the 5'-end of the *argX* gene reveals some interesting anomalies that may account for the low abundance of tRNAarg3.

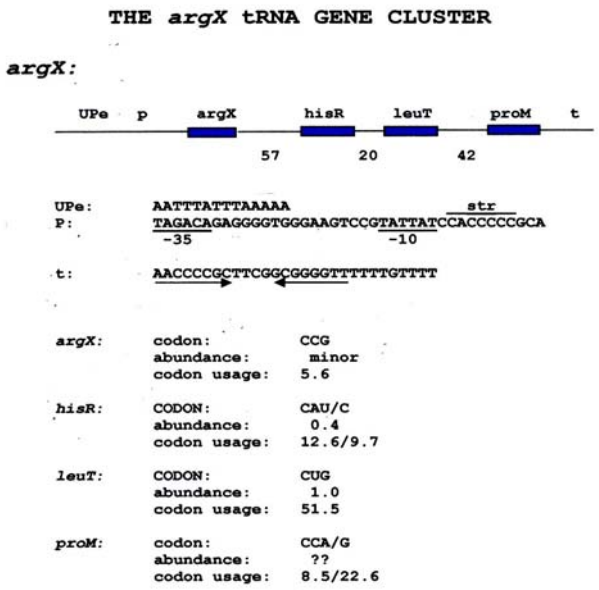


Figure 3: The *argX* operon is shown here along with the sequences of the Up element (Upe) the Promotor (P), and the terminator (t).

tRNAs normally form an aminoacyl stem, in which the first seven bases at the 5'end of the tRNA base pair to seven bases near the 3' end. However, the eight nucleotides just before the start of the 5' end of the *argX* precursor RNA can base pair with the beginning of the tRNA, creating an alternative structure (stem-loop) to the aminoacyl stem (Figure 4). The aminoacyl-tRNA stem is one of the major recognition factors in tRNA maturation. The endoribonuclease, RNase P, makes a precise cut at the 5'-end of the aminoacyl-tRNA stem in the precursor tRNA. If there is an alternative structure formed before maturation, RNase P will not recognize its normal cut site. In addition if the aminoacyl stem does not form correctly, 3'-exonucleases that generate the 3'- or CCA end of the tRNA will likely proceed past the CCA sequence and degrade the tRNA before it matures. *argX* is separated from the polycistronic operon by RNase E cleavage upstream, but near the 5' end of the *hisR* gene (Li and Deutscher, 2002). RNase E usually makes its cut site much closer to the 3' end of the tRNA being matured. This elongated cleavage site requires removal of a lengthy trailer sequence by the 3' exoribonucleases (Li and Deutscher, 2002) (Figure 5). The delay in removal could contribute to the proposed alternate base pairing that occurs at the 5' end of *argX*.

If there is an alternative structure formed before maturation of the tRNA *argX* so that the aminoacyl-tRNA stem is not formed, there could be a loss of recognition by the RNase P and no cutting at the 5' end. Due to the consequences of this proposed alternative structure, the 3'-exoribonucleases may continue to degrade the 3' end of the tRNA past the CCA sequence and into the region normally protected by secondary structure of the aminoacyl stem. The resulting partially degraded tRNA cannot mature and likely would be completely digested. The production of different constructs that manipulate the sequence just prior to the 5' end could show that it is in fact the abnormal base pairing that is inhibiting the complete maturation of the *argX* tRNA.

Several *argX*-containing plasmids, using plasmid pUC18, have been constructed and tested. It is anticipated that modifying the sequence preceding the 5'-end of the *argX* gene will affect the quantity of tRNAarg3. All *argX* plasmid constructs contain the *hisR* gene directly following the *argX* gene as it does in the normal gene cluster. The presence of the *hisR* gene allows for the rapid digestion of the 3' end of *argX* before ribonuclease P can cut the aminoacyl stem; as occurs naturally in the cell. There are no

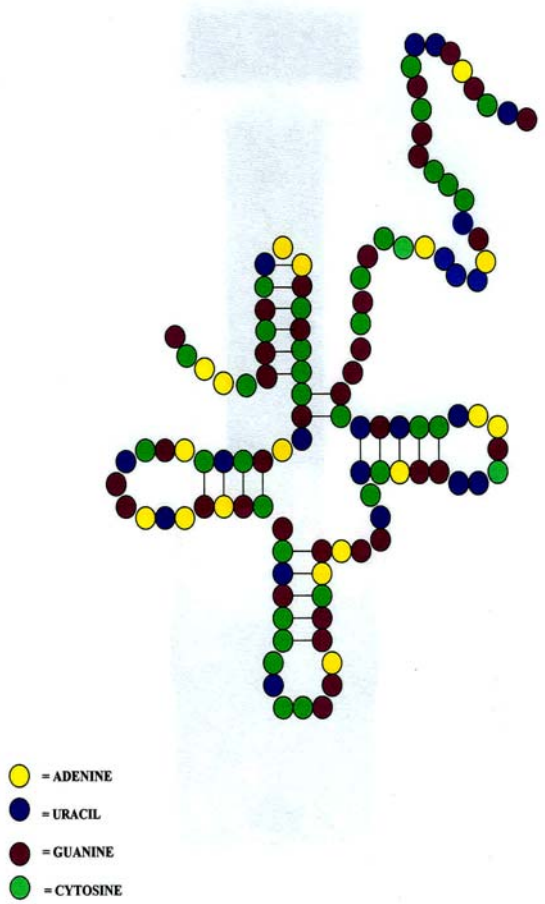


Figure 4. The proposed alternative structure of argX tRNA is shown above.

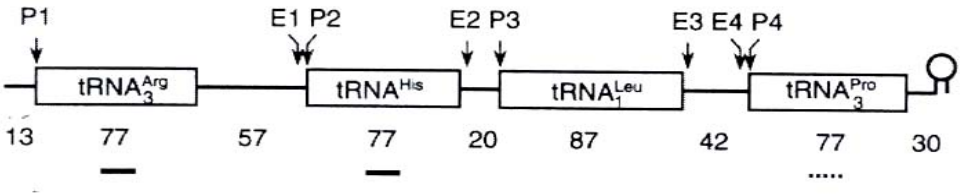


Fig. 5. This diagram shows the four cut sites for RNase E (E1-E4) and RNAse P (P1-P4). The RNAse E cut site for $tRNA_3^{ARG}$ (E1) is located close to the RNAse P (P1-P4) cut site (P2) for $tRNA^{HIS}$ (Li and Deutscher 2002).

other copies of the *hisR* gene in the *E. coli* genome; therefore it also serves as an internal control, showing the constructs are being transcribed.

tRNA detection and quantification

Early work on separation and quantification of tRNAs was done by two-dimensional (2-D) gels. Different concentrations of acrylamide with a ratio of usually 1:2 has been used to fractionate RNAs that range from 50-500 nucleotides (Ikemura, 1989). The use of different concentrations of urea in the first and second dimensional gels was mainly used for small RNA molecules such as tRNA (Ikemura, 1989). This combination of acrylamide and urea concentration gave good resolution for an extensive variety of RNA molecules (Ikemura, 1989). Ionic and conformational aspects were also involved in the separation of tRNAs by this method (Stein and Varricchio, 1974). Rare tRNAs were not easily detected by this method if they were detected at all. This method is time consuming and laborious. Other methods used to separate and quantify tRNA or tRNA genes included UV absorbance, ribosome binding, Northern and Southern blotting and non-radioactive hybridization assays using nylon transfer membranes and DNA probes.

We have developed a sandwich hybridization assay that can be used for the detection of rare tRNA products. The assay incorporates the use microtiter wells and capture and reporter probes to detect low levels of specific tRNAs. Both rapid and sensitive, it is a convenient method for detection and quantification of individual tRNA species.

MATERIALS AND METHODS

Bacterial strain and plasmid preparation

E.coli strain JM109 (RR539) was the host used for cloning the tRNA genes. The genotypes of these strains along with plasmid containing strains are listed in Table 2. *E. coli* strain K12 (RR34) was used to prepare genomic DNA that was used as template for the synthesis of the tRNA genes.

The plasmid pUC18 (Figure 6) was used as a vector for the inserted tRNA genes. It is a 2686 base pair (bp) plasmid of double stranded (ds) DNA that is commonly used for cloning and sequencing of inserted DNA. It is a pUC derivative containing the intergenic (IG) region of M13 that codes for production of ssDNA. It carries a β-lactamase gene, which makes the host strain resistant to ampicillin. The pUC plasmids are all high copy number cloning vectors, with more than 100 copies per genome. All the pUC plasmids exist in high copy numbers because their pMB1 ori is a modified colE1 ori designed for high copy replication.

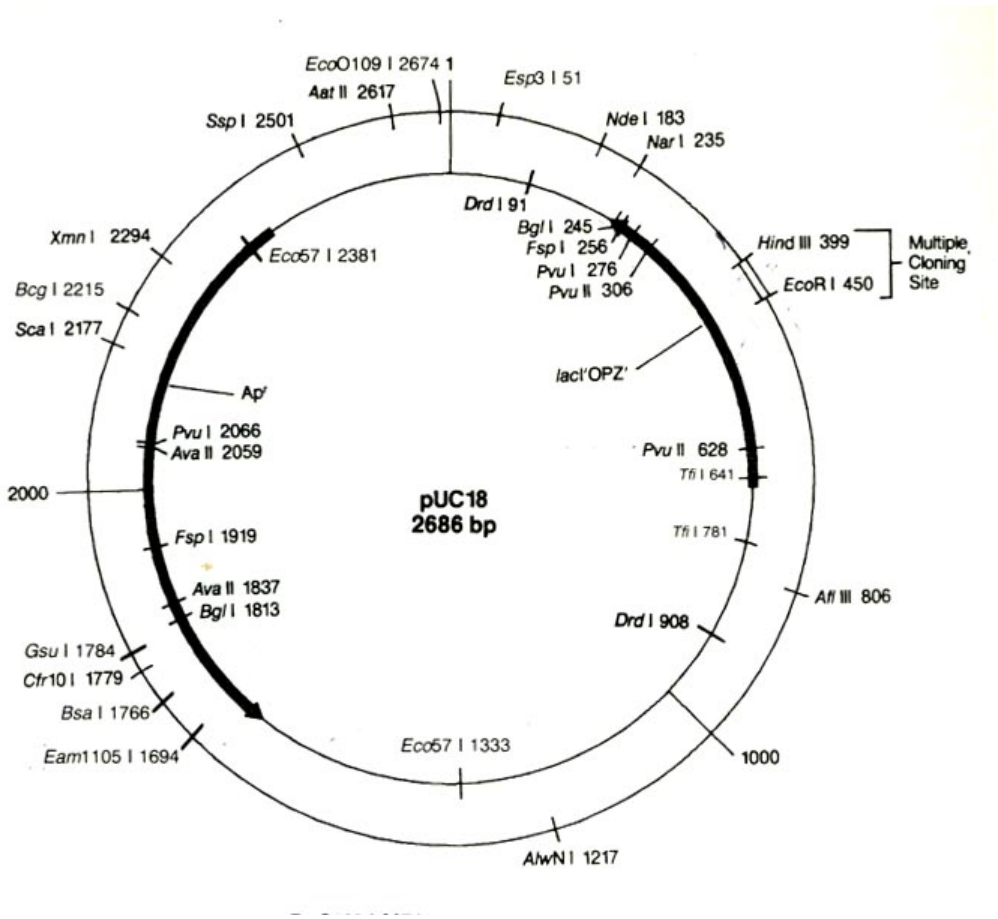
Bacterial strains were grown in Luria-Bertani (LB) broth (LB: 1% tryptone, 0.5% yeast extract, 1% NaCl, pH 7.0) at 37°C with vigorous shaking. The LB medium used to grow strains containing pUC18 plasmids was augmented with 60 μg/ml ampicillin. Plasmid DNA was prepared according to the lysis by alkali protocol described by Sambrook and Russell (Sambrook and Russell, 2001). The DNA was dissolved in 3 ml of TE solution (10 mM Tris-HCl, 1 mM EDTA, pH 8.0) and stored at -20°C. 400 μl samples were further purified through QIAquick PCR Purification Kits (QIAGEN Inc., Chatsworth, CA), eluted in 100 μl of water, and stored at -20°C. The concentration of the plasmid DNA was determined by gel electrophoresis analysis using a Gibco-BRL mass ladder as a standard.

Table 2. Bacterial Strains and Plasmids

Strain designation	Description
RR34	*E. coli*: K12 C600: piR T5r thr- leu- thi- suIII
RR 539	*E. coli*: JM109: *recA1 supE44 endA1 hsdR17 gyrA96 relA1 thi Δ*(*lac-proAB*) *F'*[*traD36proAB*$^{+}$*lac1*^{q}*lacZΔM15*]
RR750	JM109/pUC18
RR765	JM109/pRR765 (pUC18 + argXlong-hisR @ BamHI/PstI)
RR766	JM109/pRR766 (pUC18 + argXshort-hisR @ BamHI/PstI)
RR776	JM109/pRR776 (pUC18 + argX-C to G(+6)-hisR @ BamHI/PstI

Table 3. Primers used in PCR reactions and for DNA sequencing (bold regions indicate restriction endonuclease recognition sequences).

Name	Sequence 5' to 3'	Restriction Endonuclease Recognition
322 (M13/pUC) (Forward)	GTTTTCCCAGTCACGAC	
325 (M13/pUC) (Reverse)	AACAGCTATGACCATG	
1791 (argX-long Forward)	CCC**GGATCC**TATCCACCCCCGCAACG	BamHI
2947 (argX Reverse) *used to make controls only	CCC**CTGCAG**CACCGCAGCTCAAGCGC	PstI
2948 (argX-short Forward) *also used to make controls	CCC**GGATCC**GCGCCCGTAGCTCAGCT	BamHI
2994 (hisR Forward) *used to make controls only	CCC**GGATCC**GTAACAAGATTTGTAGT	BamHI
2995 (hisR Reverse) *used to make controls only	CCC**GGTACC**GTCACAACTTCTAATAA	RpnI
3706 (argX Forward) (+6, C to G)	CCC**GGATCC**TATCCACCCCCGCAACG GGGCTAAGCGCCCGTAGCT	BamHI
4001 (hisR Reverse)	CCC**CTGCAG**GTCACAACTTCTAATAA	PstI



pUC18 multiple cloning site and primer binding regions: 364–500

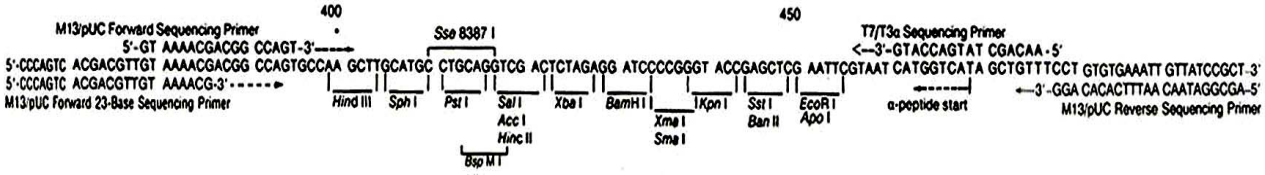


Figure 6. Diagram of the pUC18 plasmid and its multiple cloning site sequence (taken from BRL reference guide)

Colony PCR (Polymerase Chain Reaction): DNA preparation for PCR

Genomic DNA samples were all obtained from *E. coli* (RR34) and were prepared as follows: *E. coli* strain K12 (RR34) was grown on LB-Agar plates overnight at 37°C, and an isolated colony was inoculated into a 1.5 ml Eppendorf tube containing 50 µl of a 10 mM EDTA solution. The microfuge tube was boiled in a water bath (100°C) for 5 minutes. Following this denaturation step, the tube was vortexed briefly and centrifuged for 5 seconds at 14,000 RPM in a microfuge. This DNA preparation was then ready to serve as a template for PCR amplification of tRNA genes. If immediate usage was not possible, the lysate was stored at –20°C for subsequent use.

Polymerase Chain Reaction (PCR)

The DNA fragments containing the tRNA genes required for cloning were obtained through PCR reactions (Mullis, 1990; Saiki, 1990) using RR34 genomic DNA as the template. The master mix required for the PCR reaction contained all of the necessary ingredients for synthesis with the exception of Taq DNA polymerase and the template DNA. It was prepared using: 100 µl 10X Gibco-BRL Commercial Taq Buffer, 70 µl 50 mM $MgCl_2$, 10 µl 5 mM dNTP mixture (dATP, dTTP, dCTP and dGTP – final concentration 50 µM), 50 µl of each forward and reverse primer and 705 µl of H_2O. 0.5 µl of Taq DNA polymerase and 1 µl of prepared DNA extract (described above) were added to 99 µl of this master mix. The PCR reaction tubes were placed in an MJ Research PTC-100 thermal cycler programmed according to the following protocol: initial DNA denaturation for 5 minutes at 95°C, 30 cycles consisting of 1.) 40 seconds denaturation at 95°C, 2.) annealing of primers to DNA template for 25 seconds at 50°C and 3.) primer extension for 3 minutes at 72°C. A final 10-minute long extension at 72°C is followed by indefinite cooling at 4°C.

The forward primer used in each reaction depended upon how each tRNA *argX* construct was being manipulated and amplified (Table 3 and Figure 8) and the reverse primer served to ensure the *hisR* tRNA gene followed behind the *argX* gene as it does naturally. The primers used for amplification of the tRNA genes are shown in Table 3. Note that in each of these primers, restriction endonuclease sites (bolded sequence in

Table 3) near the 5'-end have been included for cloning each of the resultant DNA fragments. The primer sequences designated "forward" are from sequences from GenBank just upstream (transcriptional direction) of the respective tRNA gene from *E. coli* K12, and like wise the primers designated "reverse" are complementary to a sequence just downstream of the gene. Primers 322 and 325 in Table 3 are used for analyzing DNA fragments inserted into pUC18. They represent sequences in pUC18 just upstream and downstream of the cloning site. They are used to make PCR DNA fragments of the cloning site to verify insertion of tRNA genes. Verification is done by DNA sequencing of the entire cloning site.

DNA purification

QIAquick-spin DNA purification kits were employed to purify double stranded DNA from excess protein, primers and nucleotides. The protocol (from QIAGEN, Inc.) is as follows: 5 volumes of PB buffer were added to the solution containing DNA to be purified. This solution was then loaded into a QIAquick-spin column, which was placed in a 2 ml microfuge collection tube. The column and collection tube (containing the sample to be purified) was then centrifuged at 10,000 RPM in a Savan Speedfuge for 60 seconds. After spinning, the flow-through fraction in the collection tube was discarded. The column was then washed with 750 µl of PE buffer and spun again for 60 seconds at 10,000 RPM and the flow-through was discarded. This step was repeated twice more to ensure removal of any residual PE buffer. The QIAquick-spin column was then placed in a new 1.5 ml microcentrifuge tube. The purified DNA was then eluted from the column by adding 50 µl of sterile distilled H_2O and centrifuging for 60 seconds at 10,000 RPM. Purified samples that were not immediately needed were frozen at -20°C.

Agarose gel electrophoresis

Agarose gel electrophoresis was carried out in either a mini or wide-mini apparatus (BioRad) using 2X TBE buffer (prepared from a 10X stock solution containing 450 mM Tris-boric acid and 10 mM EDTA. The DNA bands migrated under an electric field of 100 volts, until the bromophenol blue tracking dye migrated within 1.5 cm of the

bottom edge of the gel. A total of 2 µl of DNA sample, 2 µl of tracking dye and 8 µl of 2X TBE was loaded into each well.

A 2% gel was used when analyzing DNA fragments (100-800 bp) obtained from PCR reactions. A standard 100 base pair ladder was used to determine fragment DNA base pair sizes. In cases where the concentration of DNA was being analyzed, a mass ladder was also loaded on the gel.

After electrophoresis, the gel was stained with ethidium bromide (0.5 µg/L) with gentle agitation for 30 minutes and then rinsed with distilled water. The gel was viewed with an ultraviolet transilluminator on a GelDoc system (BIO-RAD) and a digital image was captured.

Preparation of recombinant DNA

This section describes the method used to clone *argX*-short into pUC18. The cloning of other genes followed a similar protocol. The plasmids and PCR amplified tRNA genes containing restriction endonuclease recognition sites were cut with the appropriate enzymes. This cutting was done by adding equal concentrations by weight (µg/µl) of the plasmid and tRNA genes to one microfuge tube. This provided a 30:1 ratio of tRNA genes to plasmid. 35 µl (50 ng) of the *argX* -short gene fragment was added to 2 µl of pUC18 (50ng), 5 µl of RE_{ACT} 2 (Invitrogen), 6 µl sterile glass distilled H_2O and 2 µl of PstI restriction endonuclease (Invitrogen) and allowed to incubate for 2 hours at 37°C. The reaction mixture was purified on a QIAgen spin column. However, DNA was eluted in only 43 µl of water, instead of the usual 50 µl. To the purified mixture was added, 5 µl of RE_{ACT} 3 (Invitrogen) and 2 µl of BamHI restriction endonuclease (Invitrogen). This reaction mixture was incubated at 37°C for 2 hours and purified on a QIAgen spin column with a final elution of 40 µl of water. Lastly, 10 µl of T4 DNA Ligase buffer (Invitrogen) and 2 µl of T4 DNA Ligase (Invitrogen) were added to the doubly cut plasmid-insert mixture and allowed to incubate at room temperature for 2.5 hours. The ligated DNA (in 52 µl) was frozen at -20°C.

Preparation of competent JM109 cells

The methods used in the preparation of competent cells were adapted from Sambrook and Russell (2001). The protocol for the preparation of 2.0 ml of competent cells is as follows: the night before competent cells were needed, 2.0 ml of LB was stick inoculated with RR539 from frozen culture in a 15 ml tube and grown overnight at 37°C with shaking. After the cells were grown to full density, 0.5 µl of culture was added to 30 ml of LB in an Erlenmeyer flask. The culture was then grown to early log phase, 2 hours with shaking at 37°C. The cells were removed from the water bath, chilled in an ice slurry and then centrifuged in a Sorvall at 5,000 RPM for 10 minutes in an SS34 rotor, maintained at 0-4°C. The supernatant was decanted. Next, 10 ml of ice cold sterile 0.1 M $CaCl_2$ was added and the cell pellet was suspended and allowed to stand on ice for 10 minutes. The cells were again centrifuged to a pellet, the supernatant discarded. Then the cells were re-suspended in 2.0 ml of ice cold, sterile 0.1M $CaCl_2$ and placed on ice for 30 minutes. The cells were now competent for transformation and 0.2 ml of the cells were aliquoted into ten 1.5 ml Eppendorf tubes and maintained at 0°C.

Transformation of competent cells and transformant screening by PCR

Recombinant plasmids were inserted into competent cells by "heat pulse" transformation. The ligation reaction mixture (20 µl) was added to 200 µl of competent cells maintained on ice. Both positive and negative controls were included; cells with uncut plasmid, cells only, or DNA (ligation mixture) only. Each tube was placed in a 42°C heat block and heat shocked for exactly 90 seconds and then removed and chilled on ice for 2 minutes. 800 µl of LB was added to each tube, which were then placed in the 37°C block and allowed to express the antibiotic resistance gene for one hour. After an hour, 400-500 µl of each cell or control mixture was plated onto LB-ampicillin agar plates and grown overnight at 37°C.

From each successful transformation plate, up to 50 isolated, single colonies were chosen and grown on a patch plate overnight at 37°C. Several colonies from each plate were then prepared for screening using the colony PCR protocol (Sambrook and Russell, 2001). Primers 322 and 325, the pUC primers, were used to amplify the cloning site.

The PCR products were purified by QIAquick PCR purification kits and then analyzed by agarose gel electrophoresis using a 100 bp ladder to measure correct length of the desired insert and a mass ladder to quantitate amount of product present. Those transformants that gave PCR fragments of the correct length were grown to full density on LB-ampicillin media and then placed into frozen culture for further use. The PCR fragment from these transformants was sequenced using the pUC primer 322. Sequence analysis by the Sanger dideoxy method was carried out using an automated DNA sequencer (Applied Biosystems, Model 373A) at the Florida State University DNA sequencing facility.

Preparation of tRNA extracts

Strains from frozen culture were grown up overnight at 37°C in 2 ml of minimal media broth (E media containing thiamine, glycerol and ampicillin). The next day the strains were diluted into 30 ml flasks, 1 ml of cells per 30 ml E media. Two flasks were inoculated with each strain, and to one was added 60 µl of 1.0 M isopropylthio-β-D-galactoside (IPTG). IPTG induces the cloned tRNA genes which are inserted down stream from a *lac* promotor and operator. The cells were harvested when they reached an A600 of 0.6-1.0, which was measured on a Baush and Lomb Spectronic 20.

Washed cells were recovered from 30 ml of minimal media broth. Cells were lysed by brief treatment with lysozyme in TE buffer (prepared from a 10X stock containing 0.1 M Tris base and .01 M EDTA, pH 8.0) plus 1 M glucose. The mixture was incubated for 5 minutes at 37°C. 5 µl of 10% sodium dodecyl sulfate (SDS), 50 µl of 100 mM PIPES, pH 6.5 and 5 µl of 1 M $MgCl_2$ were added and the mixture was placed in a boiling water bath for 40 seconds. The mixture was centrifuged for 5 minutes and the supernatant was transferred to a clean microfuge tube. The RNA was precipitated by adding 20 µl of 5 M NaCl and 600 µl of cold isopropanol. The RNA was pelleted by centrifugation and the supernatant removed. The RNA precipitate was washed with 500 µl of cold 70% ethanol, centrifuged and the ethanol decanted. The RNA was dissolved by adding 500 µl of sterile 10 mM PIPES, pH 6.5 and 5 µl of 100 mM EDTA. The tRNA solution was heated in an apparatus (Speed-Vac) to evaporate any residual alcohol and

was stored at -20°C. The amount of RNA from each extract was measured by absorbance at 260 nm using a Pharmacia GeneQuant.

tRNA assay

There were several capture and reporter probes made by Integrated DNA Technologies, Inc. (Table 4). The capture probe was added to PE buffer (1.0 M NaH_2PO_4 and 200 mM EDTA) at a dilution of 1:200. 200 µl of this dilution was added to each well. The wells were incubated for 15 minutes at 37°C. The wells were then washed 3 times with 1 X Tris Buffered Saline (TBS) (prepared from a 10X stock containing 100 mM Tris, pH 8.8 and 1.5 M NaCl). This was followed by a blocking step using Blotto (milk protein) in TTBS (TBS with Tween 20 added). The wells were washed 2 times with glass distilled water. 200 µl of a prehybridization solution (0.2% SDS, 10X Denhardt's solution and 6X SSPE) was then added to each well and the wells were incubated for 30 minutes at 42°C. The wells were washed 2 times with glass distilled water. The reporter probe was added to hybridization buffer (0.2% SDS, 6X SSPE and 40% formamide) at a dilution of 1:200. 200 µl of this dilution along with up to 50 µl of corresponding tRNA was added to each well. The wells were incubated at 60°C for one hour. The wells were then washed 5 times with wash A (0.2% SDS and 6X SSPE), 3 times with wash B (0.2% SDS and 6X SSPE, different volume) and 3 times with 1X Buffer 1 (prepared from a 10X stock containing 0.1 M Tris, pH 8.8 and 0.15 M NaCl, pH 7.5). The wells were then incubated at 42°C for two 5 minute intervals in Buffer 2 (3% bovine serum albumin into 10 ml of 1X Buffer 1). Buffer two was discarded without washing and 200 µl of a streptavidin-alkaline phosphatase solution (SA-AP) (1:1000 dilution of SA-AP (Invitrogen) to Buffer 1) to each well and incubated for 10 minutes at room temperature. The wells were washed 5 times with Buffer 1 and 3 times with Buffer 3 (0.1 M Tris base, 0.1 M NaCl and 0.05 M $MgCl_2$, pH 9.5). 200 µl of TruBlue® solution was added to each well and incubated at room temperature for 10 minutes. The color change was read at an absorbance of 650 nm on a DU® 640 Spectrophotometer (Beckman).

argX and hisR tDNAs (tDNA controls)

Using asymmetric PCR, single-stranded DNA was made analogous to its single-stranded tRNA counterpart (McCabe, 1990). *E. coli* K12 (RR34) DNA was used as template, and the primers (Table 3) were used at a ratio of 1:10, reverse to forward, respectively. The single stranded DNA fragments were analyzed on an agarose gel to verify correct fragment size. The single stranded DNA fragments were used in the tRNA assay (see Figure 10 below) in increasing amounts of 0 µl, 10 µl, 25 µl, 50 µl and 100 µl. Standard curves, tDNA amount versus A650 reading (see Figures 11-14 below), were plotted and used to determine the amount of *argX* or *hisR* gene product (tRNAarg3 or tRNAhis) in cell extracts.

Table 4. Capture and Reporter probe sequences

Name	**Sequence 5' to 3'**	**Modification**
argX Capture Probe	TTTTTTTTTTGGCGCGCCCGACAGGATTCG AACCTGAGACCTCTGCC	5' Amino Modifier C12
hisR Capture Probe	TTTTTTTTTTGGGGTGGCTAATGGGATTCG GAACCCACGACAACTGA	5' Amino Modifier C12
argX Reporter Probe	TCCGGAGGGCAGCGCTCTATCCAGCTGAGC TACGGGCGCTTTTTTTTT	3' Biotin
hisR Reporter Probe	ATCACAATCCAGGGCTCTACCAACTGAGCT ATAGCCACCTTTTTTTTT	3' Biotin

RESULTS

Development of the tRNA Assay

The detection of rare tRNAs in cell extracts of *E. coli* has not been easily accomplished in the past. Two dimensional gel electrophoresis was used as an early method for detection of individual tRNAs and could even be used to determine relative amounts of the major tRNAs. However, the method was insufficient at detecting rare tRNAs. Other methods that are able to detect rare tRNAs include the use of column chromatography with specific tRNAs labeled with a radioactive amino acid. These methods are slow and laborious, and still may not be able to detect a rare tRNA when it is but one member of a family of tRNAs. In the arginine family of tRNAs in *E. coli*, there are seven genes making four tRNAs. By far the major tRNA is tRNAarg2, which is transcribed from four tandemly repeated genes. The other three arginine tRNAs, each represented by single genes, are minor or rare tRNAs (e.g., tRNAarg3 is a product of the *argX* gene) and can barely be detected by any method. In this study an assay was developed that was both sensitive and specific, did not employ the use of radioactivity and was not as time consuming as previous methods. The result was a sandwich hybridization assay that utilized microtiter wells, capture probes, and reporter probes that produces a detectable (and measurable) color change (Figure 7). The assay is both sensitive and specific and has the great advantage of measuring a single, minor member of a tRNA family (e.g., tRNAarg3 in the presence of tRNAarg2).

The first attempt at developing a non-radioactive assay involved the attachment of tRNA to nylon membranes and then used a biotinylated probe to detect the tRNA. This method was time consuming and gave inconsistent results. The attachment of tRNA to the nylon membranes seemed to be a limiting step. Since each specific tRNA is only between 70 and 90 bases in length, it was difficult to assure its attachment to the membrane. It appeared to be specific for the tRNA, but not sensitive and was unable to yield quantitative, reproducible results.

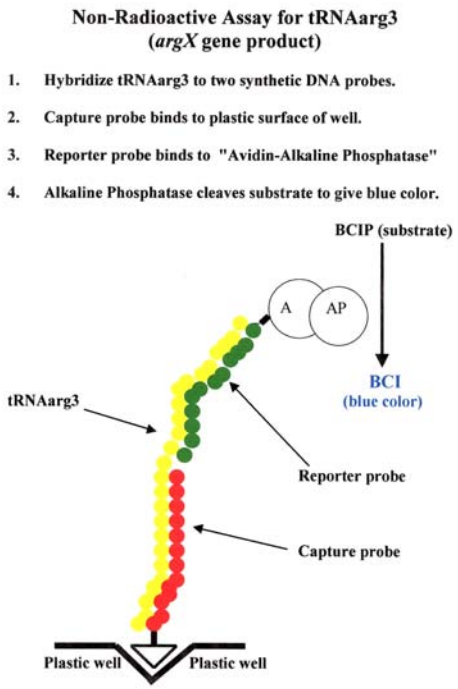


Figure 7. The general concept of the sandwich hybridization assay is shown above.

Our second attempt at developing a specific tRNA assay used microtiter plate technology with Versalinx 96-well plates coated with salicylhydroxamic acid (SHA). The capture DNA probe, which hybridizes to the 3'-end of the tRNA target, was modified at its 5'-end with phenylboronic acid (PBA). PBA forms a stable complex with SHA, and the capture probe is specifically bound to the plastic surface at its 5'-end (see Figure 7). In this assay a hybridization "sandwich" is made of the target tRNA between the capture probe and the biotinylated reporter probe, which hybridizes to the 5'-half of the tRNA. This assay seemed to work very well and was both specific and sensitive. Unfortunately, while we were developing this assay, the Versalinx microtiter plates were discontinued by the manufacturer, and an alternative capture method had to be employed.

The DNA-Bind$_{TM}$ (Corning, New York) 1 x 8 stripwell plates are designed for nucleic acid hybridization assays by enabling covalent binding of aminated nucleic acids to the surface of the plate. Reactive N-oxysuccinimide esters (NOS groups), which react with nucleophiles like primary amines, are covalently linked to the polystyrene surface of the wells. Oligonucleotides modified at the 3'- or 5'-end with primary amines can be

directly attached to the NOS surface. The primary amine attached to the 5'-end of the capture probe reacts with NOS groups causing the ester to undergo nucleophilic substitution, displacing the N-oxysuccinimide group and allowing for very specific coupling (Corning, 2003) (Figures 8 and 9). The reporter probe is tagged with biotin at its 3'-end, which specifically binds a streptavidin-alkaline phosphatase conjugate. The bound alkaline phosphatase portion of the conjugate catalyzes the removal of phosphate from the substrate, BluePhos ®, which develops a soluble blue reaction product in the well. The BluePhos ® substrate system consists of 5-bromo-4chloro-3-indolyl phosphate (BCIP) which is hydrolyzed by alkaline phosphatase. The indoyl product is oxidized to form a soluble blue chromophore by a tetrazolium compound, which is reduced to a blue-colored formazan (KPL, 2003). This mixture of chromophores has an absorbance that is easily detected with a spectrophotometer at 620-650 nm. Originally, the assay used a combination of BCIP and nitro blue tetrazolium chloride (NBT) which was partially insoluble and made measurement of absorbance more difficult. After extensive testing, the BluePhos ® substrate was introduced into the assay with more sensitive and consistent results. The solubility of the BluePhos ® substrate allowed for a homogeneous color change with no precipitation of the substrate.

Once the protocol for the assay was determined, it was tested on single stranded DNA controls (tDNAs). The controls were made by asymmetric PCR to simulate either arginine or histidine tRNAs, products of either *argX* or *hisR*. The controls were measured in units (µl) of tDNA added per well and the color reaction quantified on the spectrophotometer (Figures 10-12). In the *argX* assays, Figures 10A, 11 and 13 the specific *argX* capture and reporter probes (Table 4) were used to sandwich the target tDNA or tRNA. Color development with the tDNAarg3 substrate is both linear and reproducible. Figure 13 represents the results of nine separate assays, performed on three different days in triplicate, and clearly indicates the reproducibility of the assay. The assay was also specific for arginine tDNA since neither histidine tDNA nor partially purified *E. coli* valine tRNA (at a greater than 10-fold excess) produce any color change. The same also is true for the *hisR* assay as shown in Figures 10B, 12 and 14. Again the

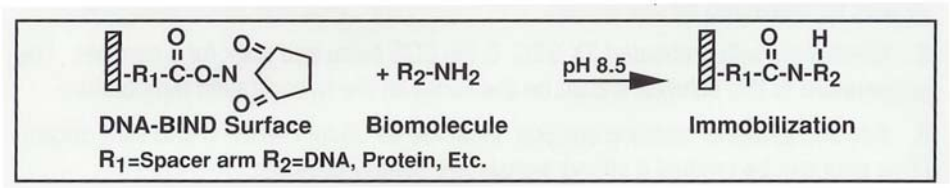


Figure 8. Primary amines react with NOS groups causing the ester to undergo nucleophilic substitution, displacing the N-oxysuccinimide group and allowing for very specific coupling (Corning 2003). This creates a stable binding to the polystyrene surface of the bottom of the well. This system is used is the DNA-BIND$_{TM}$ stripwell plates.

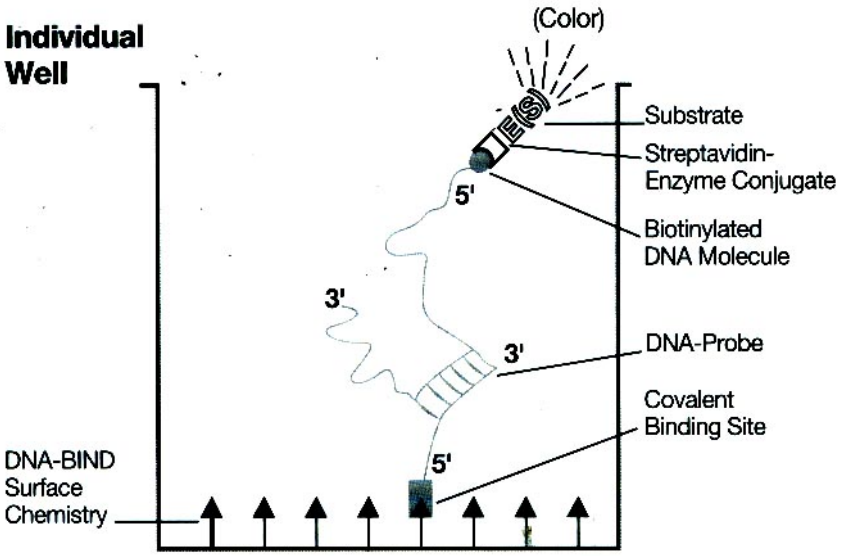
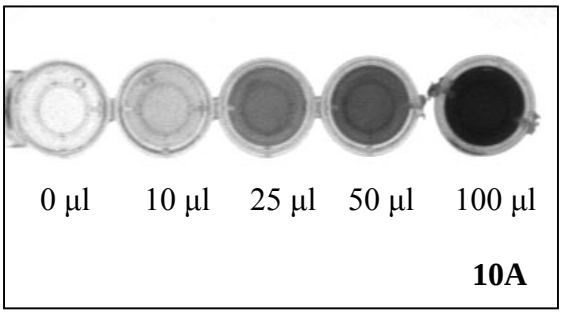


Figure 9. This figure is another depiction of the tRNA assay using DNA-BIND$_{TM}$ stripwell plates showing a capture probe binding to a biotinylated target (Corning 2003).

assay was linear over the substrate tDNAhis concentration that was used; it was reproducible and specific for the histidine tDNA. The slight increase in valine tRNA in Figure 12 for the *hisR* assay is most likely due to a background of tRNAhis found in the commercial preparation of the valine tRNA sample. The sample was enriched but not purified for valine and the assay is sensitive enough to detect the low background of tRNAhis.

These results with control tDNAs gave us confidence in the ability to accurately measure the amount both tRNAarg3 and tRNAhis in cell extracts of *E. coli.* As related above these two tRNA products are made from a single transcript, but appear to be present in the cell in very different amounts. The histidine tRNA is a considered a to be a relatively common tRNA and the arginine3 tRNA is considered a minor or rare tRNA.



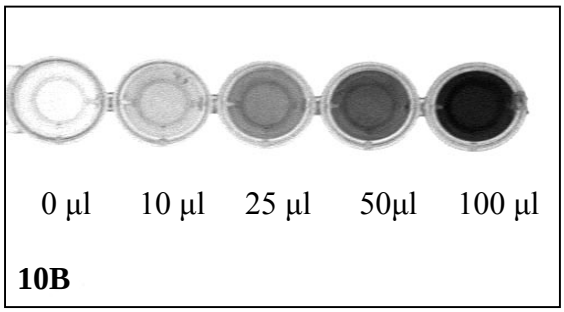


Figure 10 A and B. A. The *argX* control microtiter wells depict the increasing color change corresponding to the amount of single stranded tDNA added to each well. B. The *hisR* control microtiter wells depict the increasing color change corresponding to the amount of single stranded tDNA added to each well.

argX and hisR constructs (plasmids pRR765, 766 and 776)

Three different tDNA constructs were amplified by PCR using DNA from *E. coli* K-12 (RR34). The sequence for the *argX* gene cluster from GenBank was used to design primers that would hybridize to the 5'-end of the transcript (forward primers) as

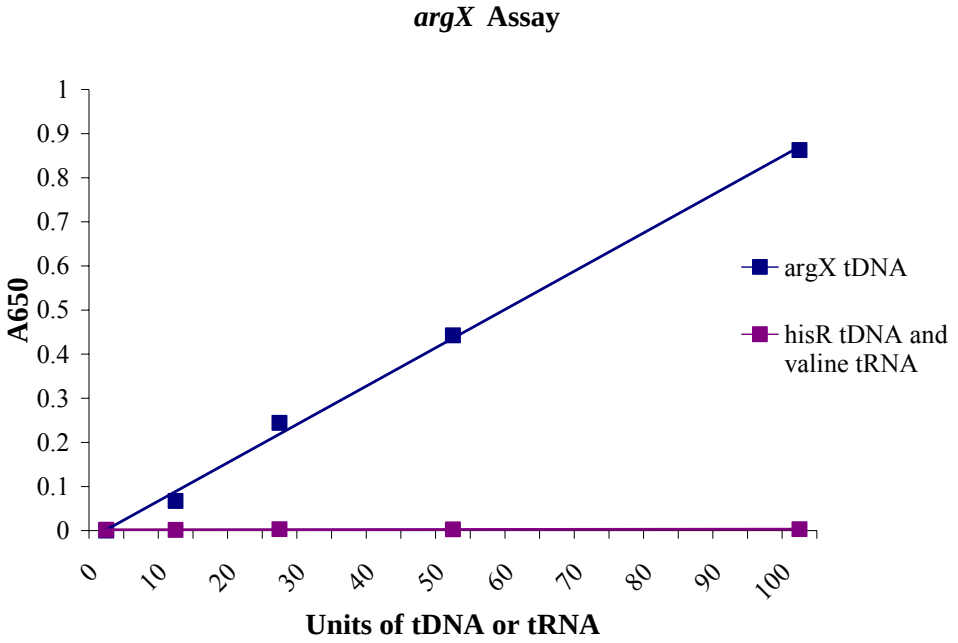


Figure 11. Standard curve created for *argX* tDNA controls using the *argX* assay. The specificity of the assay was tested using *hisR* tDNA and a commercial preparation of valine tRNA.

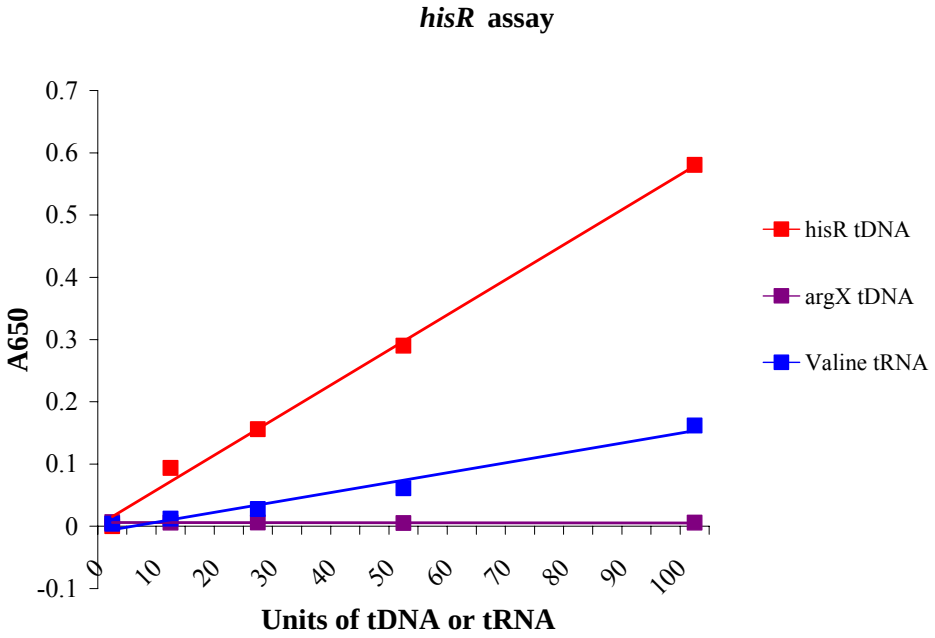


Figure 12. Standard curve created for *hisR* tDNA controls using the *argX* assay. The specificity of the assay was tested using *argX* tDNA and a commercial preparation of valine tRNA.

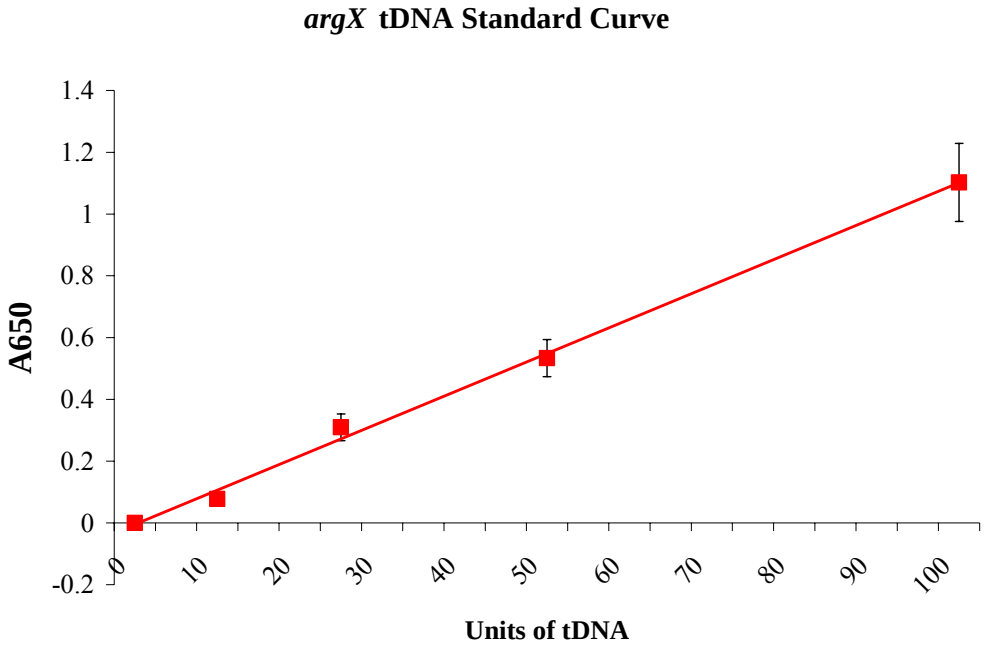


Figure 13. Standard curve created for all *argX* tDNA controls. This represents the mean of nine assays performed on three different days.

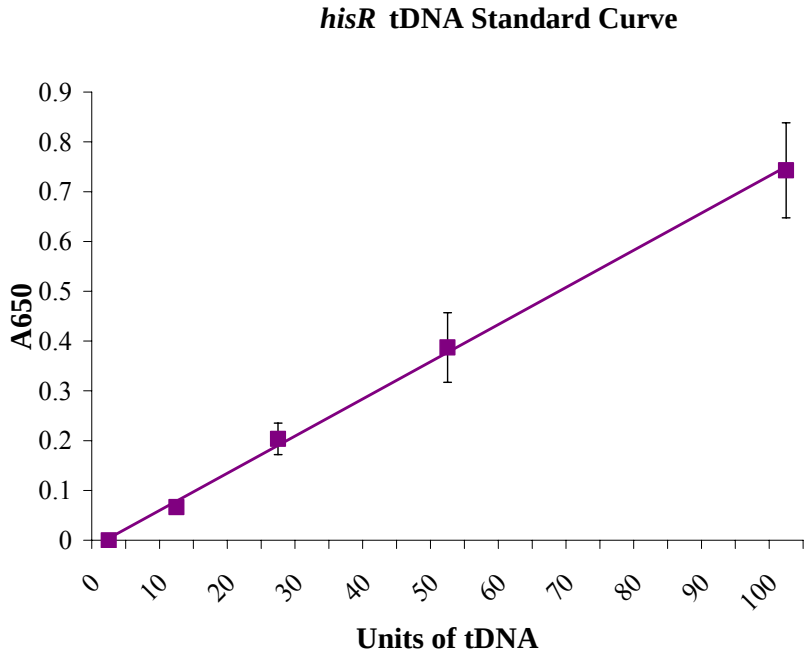


Figure 14. Standard curve created for all *argX* tDNA controls. This represents the mean of nine assays performed on three different days.

mentioned previously (see Table 4). One of these PCR fragments was the same as the wild type sequence and the other two were designed to change the sequence of nucleotides just in front of the argX gene. For all three constructs the 3'-end primer (reverse primer) was designed to hybridize to the sequence just downstream of the *hisR* gene, and, therefore, the sequence of the *hisR* gene and the sequence of base-pairs between the two tRNA genes in the operon are the same in all constructs (Table 5). The *argX*-short construct (insert in plasmid pRR766) has the thirteen wild type nucleotides directly preceding the 5'-end of *argX* completely removed. The *argX*-long construct (in pRR765) is the wild type sequence. A third construct (in pRR776) theoretically will cause a disruption in the proposed alternative base pairing structure (Figure 15C) by creating a base pair substitution of G/C for C/G at position minus six relative to the *argX* gene. Recombinant plasmids were constructed *in vitro* from each of the three PCR fragments inserted into the pUC18 cloning site, and each plasmid was cloned into *E. coli* JM 109 cells. Transformants were selected for screening, and those that had inserts in the pUC18 cloning site by PCR fragment size analysis were isolated and assigned strain and plasmid numbers (pRR765, 766 and 776) (Figure 15). The integrity of each construct was confirmed by DNA sequencing. A PCR fragment containing the entire pUC18 cloning site was amplified, purified and sent to the FSU DNA Sequencing Facility. The results in all three cases were as predicted by the cloning strategy. (see Tables 5 and 6).

Detection of tRNA arg3 using the tRNA assay

Three separate batches of tRNA were extracted from the constructed strains RR765, RR766 and RR776. These E. coli JM 109 strains are isogenic and differ only by the insert in the recombinant pUC plasmids (pRR765, pRR766 and pRR776, respectively) as described above. In addition, tRNA extracts were made from strain RR750 which is E.coli JM109 containing unaltered pUC18 (no tRNA genes inserted). The *argX* tRNA assay (capture and reporter probes for tRNAarg3) was performed three times on each batch of extracts, and controls were run in parallel (in triplicate) for each of the batches. (The *hisR* assay and controls were also run in parallel with each of the argX assays.) A standard curve of the controls for each run was determined based on the absorbance readings (Figures 13 and 14).

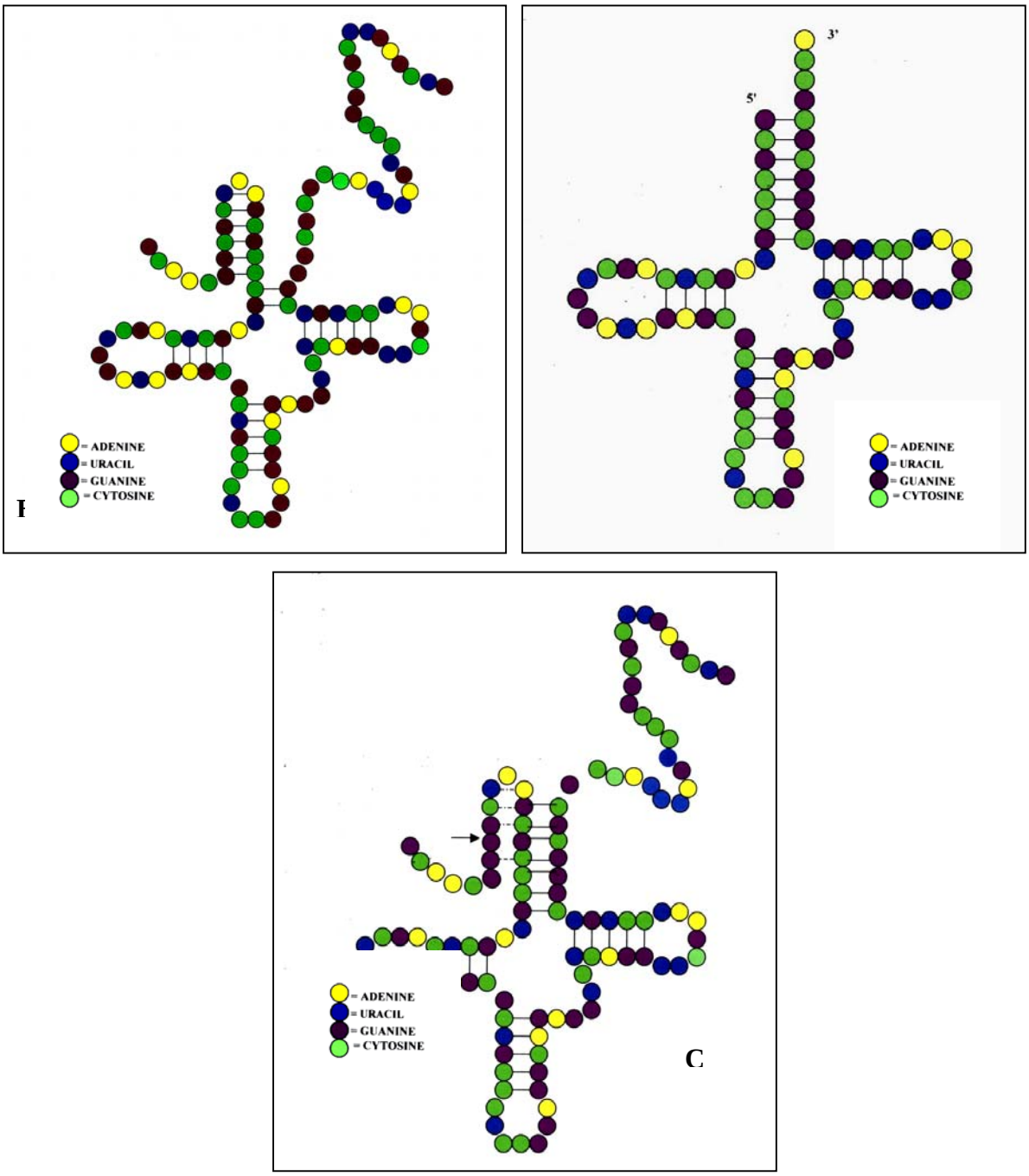


Figure 15 A, B and C. Theoretical drawings of the constructed tRNAs.
15A. pRR765 with no modification done to the preceding 5' sequence. 15B. pRR766 with the sequence prior to the 5'-end removed. 15C. pRR776 with the base pair change (arrow) creating a disruption in the hypothetical alternative base pairing.
Dashed lines indicate less likely base pairing situations.

Table 5. The *argX* operon sequence. The *argX* gene is in blue, the *his R* gene is in pink and the sequence that has been manipulated at the 5'-end of *argX* is in red (Blattner et al., 1997).

CGGGAAAGCG CATAAACTGG AGGAATAAGC AGCAAAACGC ACAAACCGTA
ACCAAACGCG CAATTTATTT AAAAAGGGAC TAGACAGAGG GGTGGGAAGT
CCGTATTATC CACCCCCGCA ACGGCGCTAA GCGCCCGTAG CTCAGCTGGA
TAGAGCGCTG CCCTCCGGAG GCAGAGGTCT CAGGTTCGAA TCCTGTCGGG
CGCGCCATTT AGTCCCGGCG CTTGAGCTGC GGTGGTAGTA ATACCGCGTA
ACAAGATTTG TAGTGGTGGC TATAGCTCAG TTGGTAGAGC CCTGGATTGT
GATTCCAGTT GTCGTGGGTT CGAATCCCAT TAGCCACCCC ATTATTAGAA
GTTGTGACAA TGCGAAGGTG GCGGAATTGG TAGACGCGCT AGCTTCAGGT
GTTAGTGTCC TTACGGACGT GGGGGTTCAA GTCCCCCCCC TCGCACCACG
ACTTTAAAGA ATTGAACTAA AAATTCAAAA AGCAGTATTT CGGCGAGTAG
CGCAGCTTGG TAGCGCAACT GGTTTGGGAC CAGTGGGTCG GAGGTTCGAA
TCCTCTCTCG CCGACCAATT TTGAACCCCG CTTCGGCGGG GTTTTTTGTT
TTCTGTGCAT TTCGTCACCC TCCCTTCGCA ATAAACGCCC GTAATA

Table 6. The constructed sequences prior to the 5'-end of the *argX* gene.

Construct	5'-end Sequence	Explanation
765	GGATCCTATC CACCCCCGCA ACGGCGCTAA GCGCCCGTAG	The sequence (in red) prior to the 5'-end of the *argX* gene(in blue) has no changes made and therefore is the wild type construct.
766	TACGAATTCG AGCTCGGTAC CCGGGGATCC GCGCCCGTAG	The first 13 base pairs of the argX transcript are removed resulting in a "shortened form" of *argX*. The sequence in black is the sequence upstream of the *BamH*I site in pUC18.
776	GGATCCTATC CACCCCCGCA ACGGCGCTAA GCGCCCGTAG	A base pair change at -6 (in black) prior to the 5'-end of *argX* causes a disruption in the proposed base pairing.

For the studies described here, the sequence just in front of the 5'-end of the *argX* gene was manipulated in an attempt to increase production of the gene product, tRNAarg3. The sequence was modified from the wild type gene to decrease the amount of alternative base pairing that could be inhibiting the formation of the aminoacyl stem, which could therefore prevent maturation (see Figure 15). The nucleotide sequences preceding the 5'-end of *argX* in each construct are shown in Table 6. Each batch of tRNA extracts consisted of growing each strain (pRR750, 765, 766 and 776) in duplicate, one culture was induced with IPTG and one was not induced. The tRNA inserts all follow a *lac* promoter and operator on pUC18, and transcription of the tRNAs is induced by the addition of IPTG. With the exception of pRR750, there is an obvious induction of tRNAarg3 (Figure 16). This demonstrates that the operator site is functioning and induction is normal. With each batch of extracts the amount of tRNA added to each assay well was determined by an A260 reading of total RNA on a GeneQuant (Pharmacia), and all extracts, induced and uninduced, in each batch were normalized to each other based on these readings. pRR750 contains the plasmid with no tRNA genes, and therefore, as expected there is not a discernable difference between pRR750 induced and uninduced (Figures 16 and 17). This assay also indicates the very small amount of genomic tRNAarg3 that is being produced.

pRR765 contained the wild type sequence of the gene (see Tables 5 and 6). The uninduced levels of pRR765 are comparable to those of pRR750 (Figure 17). There is a slight induction that occurs. These results for pRR765, demonstrate that those levels of tRNAarg3 being made are only slightly higher than genomic levels when the strain is induced.

The pRR766 construct ("*argX*-short")was made so that the first 13 bases of the wild type transcript were removed, creating a shorter version of the transcript. This was done in order to eliminate the possibility of alternate base pairing that may be occurring in the wild type transcript (Figure 15A). The amount of tRNA being made in the uninduced construct pRR766 is elevated compared to pRR765, and there is a 5.4-fold increase in the amount of tRNAarg3 made in induced extracts of RR766 compared to pRR765 (Figure 17). The lowest value of the error bar for induced pRR766 still shows a 4-fold increase when compared to the highest value of the error bar for induced pRR765

(Figure 17). These results demonstrate that the pRR766 construct ("argX-short") results in considerably more tRNA gene product than the wild type construct pRR765, and supports the idea that the nucleotides preceding the gene have an effect on the amount of product.

The construction of pRR776 involved a single base-pair change 6 base pairs in front of the *argX* gene (Figure 15C). It was expected that this would disrupt the alternative base pairing in the resulting transcript, and would result in tRNA yields similar to those of pRR766. The uninduced amount of tRNA produced is similar to the amount produced by uninduced pRR766. The amount of tRNA produced from the induced construct shows a 1.5-fold increase over induced pRR765 (Figure 17). This increase is smaller than expected for this construct, which may mean that changing only one nucleotide does not entirely prevent the alternative base pairing structure.

Detection of tRNAhis using the tRNA assay

All *argX* plasmid constructs contain the *hisR* gene directly following the *argX* gene as it does in the normal *argX* operon. The presence of the *hisR* gene and the normal intergene region between the two genes was included in all three constructs to allow for the normal maturation of the 3'-end of tRNAarg3. The endonuclease and exonuclease activities that mature the 3'-end of the tRNAarg3 precursor could proceed before ribonuclease P cuts the 5'-end of the precursor and rapid maturation at the 3'-end may be important in controlling the overall production of the tRNAarg3. If 3'-exonuclease activity proceeds beyond the CCA 3'-end as it could in the structure in Figure 15A, little mature tRNAarg3 would be produced. There are no other copies of the *hisR* gene in the *E. coli* genome; therefore, it also serves as an internal control, showing the constructs are being transcribed and induced. Assaying for the production of tRNAhis also gives us an indication of the relative production of tRNAarg3 since they are both matured from the same transcript.

It is important to note that the tRNA extracts had to be diluted 200-fold prior to addition to the wells in all the *hisR* assays. This was done to ensure that the absorbance readings would be within the values of the standard curve, and indicates that the

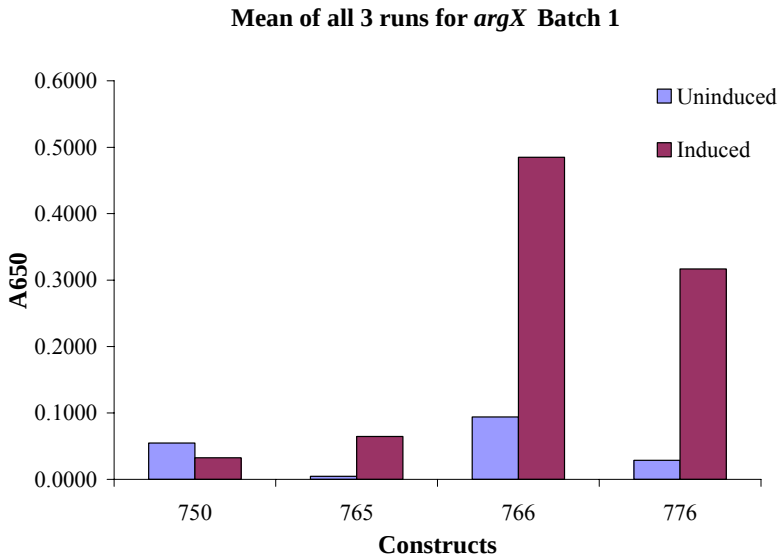


Figure 16. Batch 1 is a representation of results, the other two batches not shown produced similar results. This graph shows the amounts of *argX* gene product tRNAarg3 made by strains, uninduced or induced by IPTG, containing plasmids with either no tRNA genes added (750), the wild type *argX* gene (765), base pair deletions at the 5'-end to inhibit alternative structure formation (766) or a base pair substitution at the 5'-end to inhibit alternative structure formation (776).

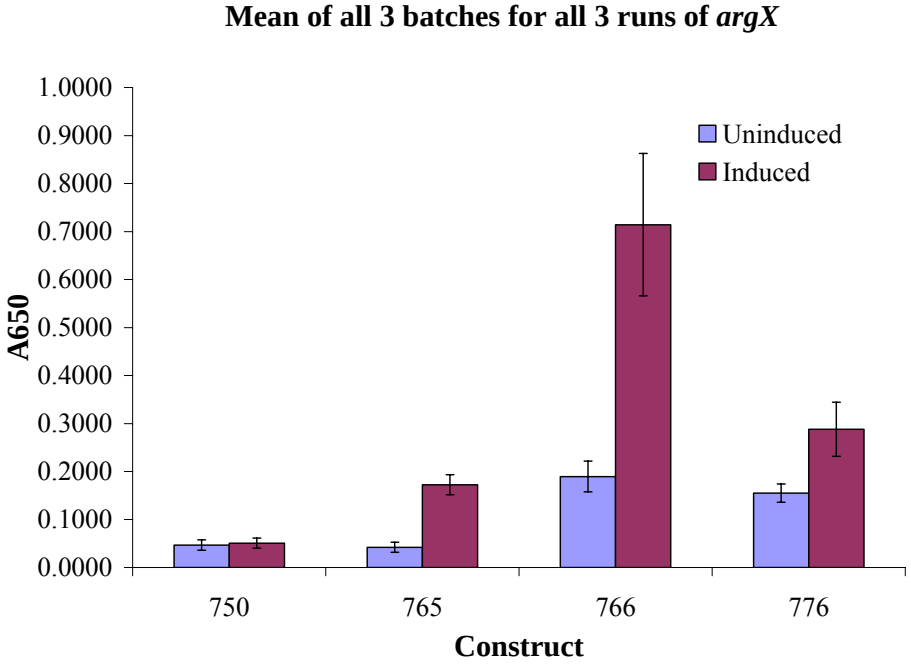


Figure 17. The means of all three batches for all three runs using error bars to illustrate the standard of error for each construct, induced and uninduced. This graph shows the amounts of *argX* gene product tRNAarg3 made by strains, uninduced or induced by IPTG, containing plasmids with either no tRNA genes added (750), the wild type *argX* gene (765), base pair deletions at the 5'-end to inhibit alternative structure formation (766) or a base pair substitution at the 5'-end to inhibit alternative structure formation. (776).

background levels of tRNAhis are approximately 200-fold higher than tRNAarg3, a very surprising result. When comparing the assay results in the Figures below, all *hisR* tRNA results need to be increased by a factor of 200 relative to the argX assays (except for results in Figure 21). The *hisR* gene is in its wild type form in all constructs, except of course for pRR750, which contains no tRNA genes inserts.

The results with strain RR750 shows the genomic amount of tRNAhis that is made (Figures18 and 19). There is not a distinguishable difference between induced and uninduced, as expected since the plasmid, pRR750, has no tRNA genes. There is a significant difference in induced versus uninduced amount of tRNAhis in strains RR765, 766 and 776 (Figures18 and 19). This shows that induction of the tRNA transcript occurred as expected. As predicted these three strains have similar amounts of tRNAhis in both uninduced and induced extracts. It is also expected that the uninduced constructs would have tRNAhis levels similar to those of the control construct, RR 750 (Figures 18 and 19). However, there seems to be an increase in the production of tRNAhis in these extracts even without induction, an unexpected result that will be discussed below. Uninduced levels of tRNAarg3 in these three extracts are also slightly elevated when compared to the level in extracts of RR750.

argX vs. hisR results and genomic tRNA results

The amount of *argX* tRNA made in comparison to *hisR* tRNA is very small (Figure 20). This result is also shown in an experiment where undiluted pRR750 was assayed for both *argX* and *hisR* (Figure 21). This experiment measured the amount of genomic tRNA being translated for each gene. As expected there is not much of a difference between uninduced and induced extracts for either tRNA. The results for RR750 extracts shows a 20-fold difference between *argX* and *hisR* genomic tRNAs, and since the hisR assay absorbance readings had to be diluted five-fold compared to the argX assay, there is really a 100-fold difference. In all other assays comparing the levels of the two tRNAs there is a 200-fold difference in their amounts. All extracts had to be diluted 200-fold in the *hisR* assays to be comparable to the results in the *argX* assays (Figures 16-20).

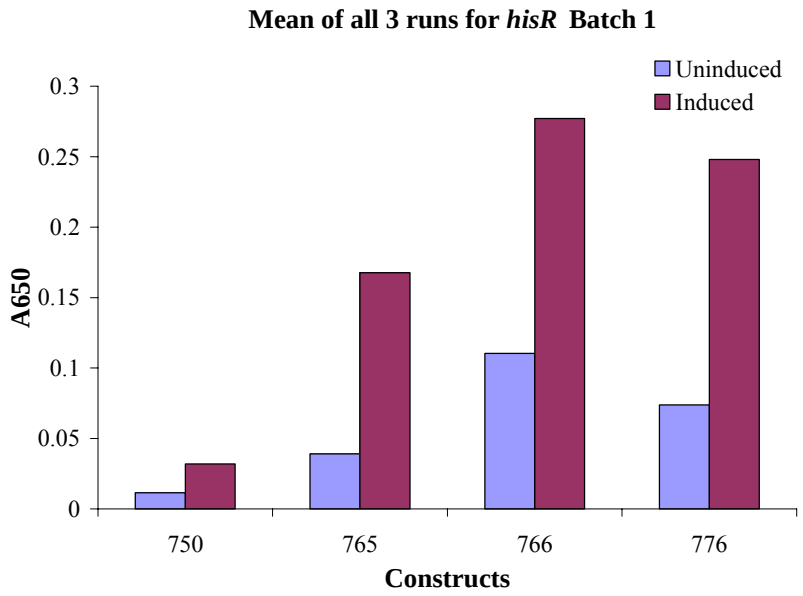


Figure 18. Batch 1 is a representation of results, the other two batches not shown produced similar results. This graph shows the amounts of *hisR* tRNA made by strains uninduced or induced by IPTG.
*All results should be increased by a factor of 200 compared to tRNAarg3

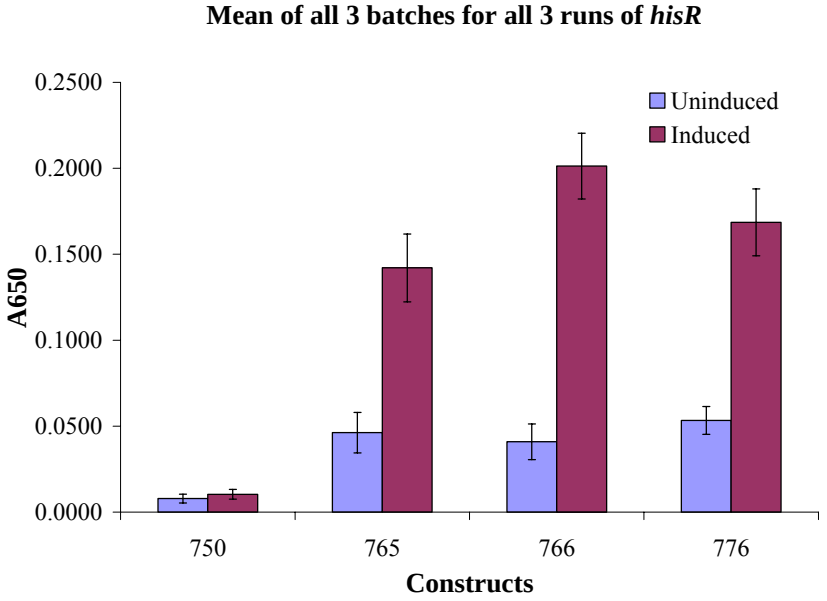


Figure 19. The means of all three batches for all three runs using error bars to illustrate the standard of error for each construct, induced and uninduced. This graph shows the amounts of *hisR* tRNA made by strains uninduced or induced by IPTG.
*All results should be increased by a factor of 200 compared to tRNAarg3

This result confirms earlier data on the abundance of tRNAhis, a fairly abundant tRNA in *E. coli* versus tRNAarg3, a very rare tRNA in *E. coli*. Inokuchi and Yamao (1994) measured the abundance of most of the tRNAs of *E.coli.* The abundance of *hisR* tRNA was given a relative value of 0.4 indicating that it is present in moderate amounts *in vivo* (The most abundant tRNAs had values of just over 1.0.). The amount of tRNAarg3 was too low to measure and was considered a minor (or rare) tRNA.

The tRNA assay that was developed in this study is clearly able to detect specific tRNAs and rare tRNAs, i.e. it is reproducible, specific and very sensitive. The genomic *argX* and *hisR* tRNA from Figure 21 was subtracted from the plasmid strains to illustrate how much tRNA was made from each of the plasmids (Figures 22 and 23). The difference is very subtle. The ability of the assay to even detect genomic tRNA, especially from the *argX* gene demonstrates its substantial sensitivity. The work with the different tDNAs used in the two assays illustrates the high specificity each assay has for its target tRNA species. For the first time a rare tRNA can be assayed in extracts where there are large amounts of other tRNA species accepting the same amino acid.

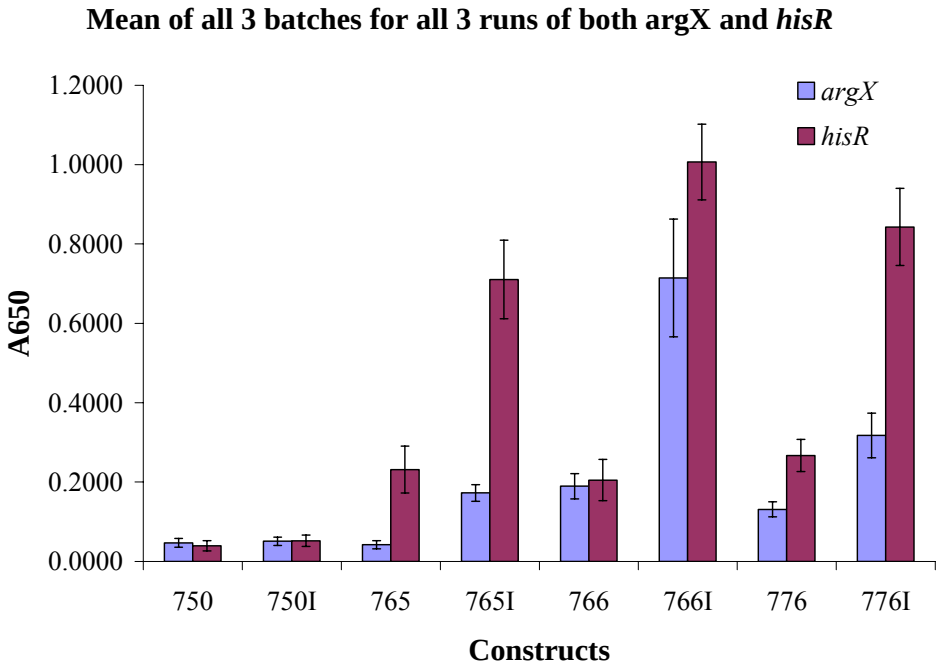


Figure 20. The means of all three batches for all three runs using error bars to illustrate the standard of error for each construct, induced and uninduced for both *argX* and *hisR*. This graph shows the amounts of *argX* and *hisR* tRNA made by strains uninduced or induced by IPTG.
*All results for *hisR* should be increased by a factor of 200.

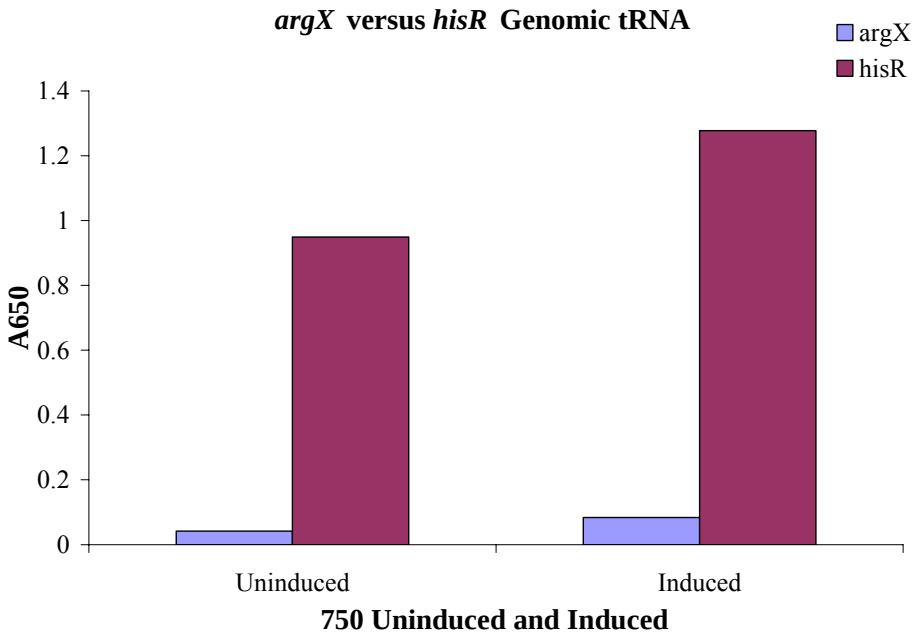


Figure 21. The amount of genomic tRNA made in *E. coli*. This graph shows these amounts using undiluted tRNA from pRR750, which does not contain any plasmid tRNA genes, only the pUC 18 plasmid itself.

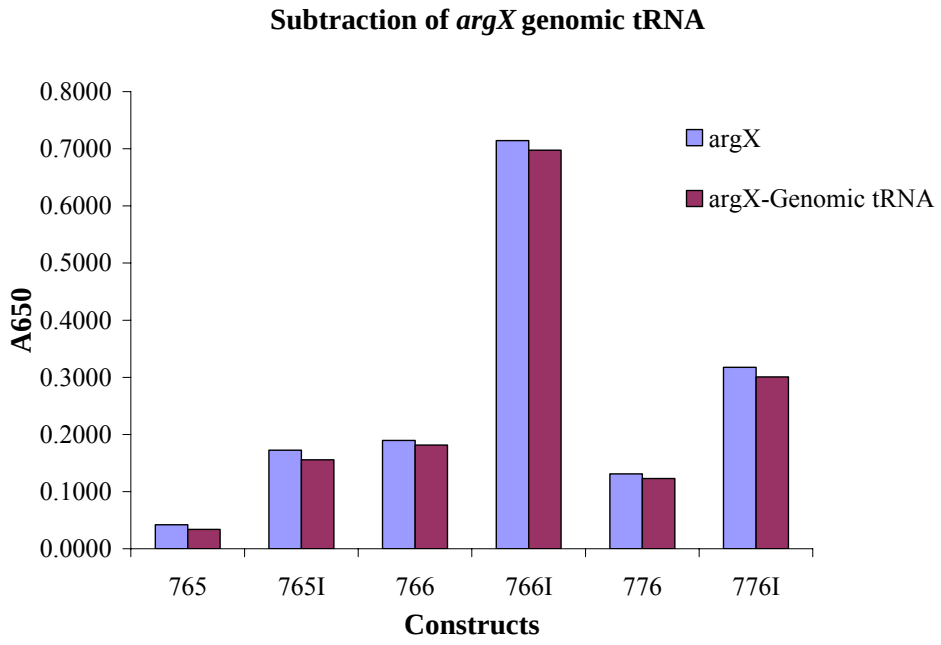


Figure 22. The subtraction of genomic *argX* tRNA from plasmid *argX* tRNA is very subtle for all constructs.

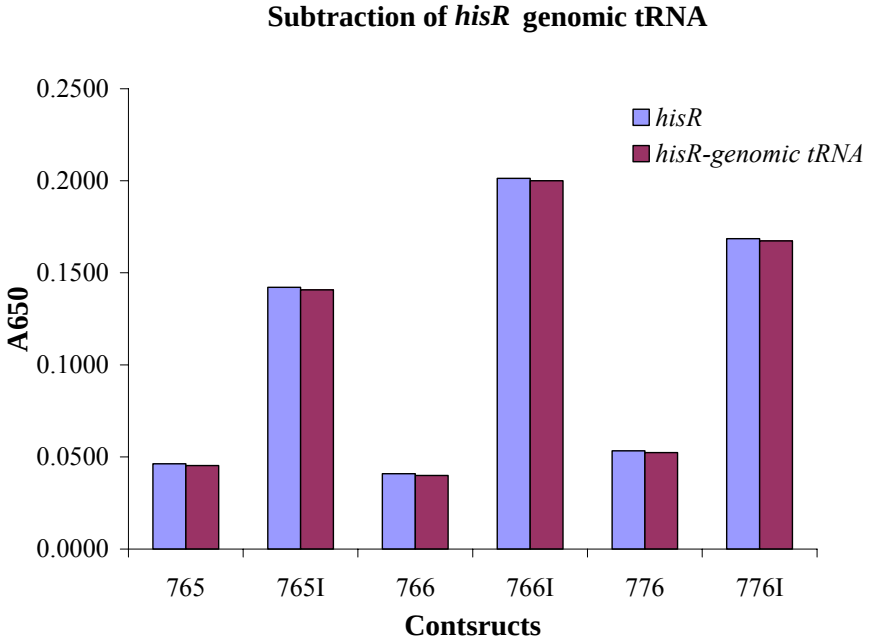


Figure 23. The subtraction of genomic *hisR* tRNA from plasmid *hisR* tRNA is very subtle for all constructs.

DISCUSSION

The tRNA assay developed in this study will be a useful tool in quantifying rare tRNAs. tRNA made in such low quantities may now be measurable using this assay. This could be useful for many applications including the expression of heterologous proteins in *E. coli* and other bacteria. The use of prokaryotic organisms to produce eukaryotic proteins can have its limits, many of which have been previously mentioned including biased codon usage. While the problem can sometimes be solved by adding a plasmid with a copy of the rare tRNA corresponding to the low usage codon, sometimes that is only a partial solution. It may be beneficial to be able to assay the rare tRNA once the plasmid has been transformed into its intended host. This assay offers a simple, non-radioactive solution that can be performed in less than a day once the tRNA extracts are made. Although some tRNAs only differ by several nucleotides, the use of capture and reporter probes that cover the entire length of the target tRNA make the assay very specific. The reporter probe can be modified to display a color change, to fluoresce or to be chemiluminescent depending on the chosen substrate method. Single stranded DNA controls (tDNAs) are uncomplicated to make and are specific to the probes as well. The necessary blocking and washing steps help make the assay sensitive by eliminating the possibility of non-specific binding and reducing background.

The results presented here demonstrate that manipulation of the sequence preceding the 5'-end of the *argX* gene can make a difference in the amount of tRNAarg3 produced. The increase in amount of *argX* tRNA from pRR766 shows that maturation does occur when the nucleotides prior to the 5'-end are removed allowing for the natural formation of the aminoacyl stem. The correct aminoacyl structure then allows for ribonuclease P to make its cut at the correct site which leads to maturation. The wild type construct, pRR765, had results similar to that of the control construct pRR750, which did not contain any tRNA genes on it plasmid. This supported the earlier report that genomic *argX* tRNA is indeed made in very small amounts and tRNAarg3 is a rare tRNA. The amount of tRNA made from pRR776 seemed to be intermediate between pRR765 and pRR766. It was expected that there would be almost if not as much tRNA produced from

this construct as from pRR766. Changing just one nucleotide to disrupt the predicted stem-loop structure involving the 5'-end of the tRNA may not have been enough to completely prevent formation of this alternative structure. It may have only partially destabilized this stem-loop structure (Figures 4 and 15) so that intermediate levels of the tRNAarg3 are matured correctly. However, it does demonstrate that even a minor change to sequence just in front of *argX* can increase the amount of tRNA made.

The results of the *hisR* assays supported these experiments in several ways. The increase in the amount of *hisR* tRNA produced when cultures were induced verified that the constructs on the plasmid were being transcribed and induction was successful. It was a valuable internal control. The *hisR* gene directly follows the *argX* gene *in vivo*, and in all the tRNA constructs used in this study, the normal or wild type sequence (including the *hisR* gene) was downstream of the *argX* gene. We felt that this was necessary for the proper maturation of the 3'-end of tRNAarg3. Also, there are no other copies of the *hisR* gene in the *E. coli* genome therefore eliminating the possibility of more than one copy of the gene creating a larger genomic background.

In results for both *argX* and *hisR* tRNA, the uninduced constructs for pRR766 and 776 all showed an increase in tRNA levels, when it was expected that the levels would be similar to that of pRR750 uninduced. An explanation for these elevated levels of both tRNAs is that the 100-200 copies of *lacO* (the *lac* operator) on pUC18 titrates the limited amount of lactose repressor made from the single copy of *lacI* on the genome (actually on an F' in JM109), and therefore, some of the *lac* operators on pUC18 are not bound by repressor. In other words because pUC18 is a high copy number plasmid, the host *lac* repressor levels may not be sufficient to efficiently repress the *lac* operator on the plasmids (Li, 1999).

When tRNA levels are compared between *argX* and *hisR*, there is a major difference in the levels of *hisR* tRNA over *argX* tRNA. If the graph in Figure 20 was made to scale after factoring in the 200-fold dilution, the levels of *argX* would barely even register. This difference in abundance between tRNAhis and tRNAarg3 was expected although the magnitude was much larger than anticipated. However, if our theory was correct that the alternative structure at the 5-end of the transcript was responsible for the low abundance of tRNAarg3, the induced levels of tRNAarg3 in

RR766 should have reached the levels of the induced tRNAhis. This did not happen. Instead of a 200-fold increase in tRNAarg3, only a 4 to5-fold increase was observed. These results indicate that something else besides the altered 5'-end may effect maturation of tRNAarg3. Recently it has been shown that RNAse E, an endonuclease, is important for the maturation of the 3'-ends of tRNA genes in *E. coli* (Li and Deutscher, 2002). The cleavage site by RNase E for *argX* is much further downstream from the 3'-end of the gene than is seen in most tRNA genes. The 3'-end of tRNA transcripts are cut by RNase E usually within 4-8 nucleotides of the CCA end, but the RNase E cut site for *argX* tRNA is 40 nucleotides downstream, and only 8 nucleotides upstream from the 5'end of *hisR* (Li and Deutscher, 2002) (Figure 5). This extended sequence between tRNAarg3 and the RNAse E cleavage site could contribute to a reduction in the maturation of the tRNA, but at this time there is no evidence to support this idea.

McNulty et al. (2003) coexpressed the *argX* gene to supplement rare arginyl $tRNA_{CGG}$ levels in the host. They were expressing a heterologous protein, the coding region for a protease domain from Herpes Simplex Virus 2 (HSV-2), which contained 11 rare arginine codons (CGG) for *argX*. Transforming a plasmid containing the *argX* gene into the *E.* coli host caused an increase in expression levels of authentic protein by up to 7-fold. The coexpression eliminated the misincorporation of glutamine, which uses the codon CAG, for arginine. However, while the plasmid increased relative amounts of correctly translated protein from 22% in the wild type host to 60% there was a lag of 1.5 hours before full-length HSV-2 protease was produced. During this lag, 78% of mistranslated protein was expressed. This could be the time frame it would take to process and accumulate enough mature *argX* tRNA from the transcript. However, using an altered form of the *argX* gene, like that in the pRR766 construct, this lag time may be reduced and may result in more authentic heterologous protein that is not mistranslated.

There are several experiments that could be performed for future studies. To test the tRNA assay's sensitivity and specificity, it would be beneficial to examine more tRNAs that are rare in *E. coli* such as *argU* and *argW* gene products. The use of low copy plasmids may not yield as much tRNA product but it may reduce the problem of a "leaky" *lac* operator, causing there to be only genomic tRNA made in uninduced constructs. Finally, the use of RNase E mutations in combination with *argX* constructs

modified on the 3-side might give some insight into whether RNase E plays a role in the abundance of *argX* tRNA. The modification of the 5'-end of the *argX* gene is not the whole story as to why *argX* tRNA is so rare nor is it the complete answer to the maturation process of the *argX* transcript. However, these results are a first step towards the complete answer as to why this particular tRNA is so rare, and developing a tRNA assay for rare tRNAs that is reliable, sensitive and specific, will allow us for the first time to measure the abundance of the minor tRNAs.

REFERENCES

Abelson J. 1979. RNA processing and the intervening sequence problem. Annu Rev Biochem 48:1035-1069.

Acosta-Rivero N, Sanchez JC, Morales J. 2002. Improvement of human interferon HUIFNalpha2 and HCV core protein expression levels in Escherichia coli but not of HUIFNalpha8 by using the tRNA(AGA/AGG). Biochem Biophys Res Commun 296:1303-1309.

Anderson WF. 1969. The effect of tRNA concentration on the rate of protein synthesis. Proc Natl Acad Sci U S A 62:566-573.

Bauer BF, Elford RM, Holmes WM. 1993. Mutagenesis and functional analysis of the Escherichia coli tRNA(1Leu) promoter. Mol Microbiol 7:265-273.

Blattner FR, Plunkett G, 3rd, Bloch CA, Perna NT, Burland V, Riley M, Collado-Vides J, Glasner JD, Rode CK, Mayhew GF, Gregor J, Davis NW, Kirkpatrick HA, Goeden MA, Rose DJ, Mau B, Shao Y. 1997. The complete genome sequence of Escherichia coli K-12. Science 277:1453-1474.

Brinkmann U, Mattes RE, Buckel P. 1989. High-level expression of recombinant genes in Escherichia coli is dependent on the availability of the dnaY gene product. Gene 85:109-114.

Bulmer M. 1988. Codon usage and intragenic position. J Theor Biol 133:67-71.

Calderone TL, Stevens RD, Oas TG. 1996. High-level misincorporation of lysine for arginine at AGA codons in a fusion protein expressed in Escherichia coli. J Mol Biol 262:407-412.

Chen GF, Inouye M. 1990. Suppression of the negative effect of minor arginine codons on gene expression; preferential usage of minor codons within the first 25 codons of the Escherichia coli genes. Nucleic Acids Res 18:1465-1473.

Chen GT, Inouye M. 1994. Role of the AGA/AGG codons, the rarest codons in global gene expression in Escherichia coli. Genes Dev 8:2641-2652.

Corning. 2003. DNA-BINDTM Application Guide. In:Cambridge: Corning Costar Corporation.

Deutscher M. 1994. Chapter 5: tRNA Processing Nucleases. In: Soll DaR, U., Editor. tRNA: Structure, Biosynthesis, and Function. Washington, D.C.: ASM Press.

Deutscher MP. 1984. Processing of tRNA in prokaryotes and eukaryotes. CRC Crit Rev Biochem 17:45-71.

Deutscher MP. 1990. Ribonucleases, tRNA nucleotidyltransferase, and the 3' processing of tRNA. Prog Nucleic Acid Res Mol Biol 39:209-240.

Elf J, Nilsson D, Tenson T, Ehrenberg M. 2003. Selective charging of tRNA isoacceptors explains patterns of codon usage. Science 300:1718-1722.

Fournier M, Ozeki H. 1985. Structure and organization of the transfer ribonucleic acid genes of Escherichia coli K-12. Microbiological Reviews 49:379 - 397.

Fuglsang A. 2003. The effective number of codons for individual amino acids: some codons are more optimal than others. Gene 320:185-190.

Goldman E, Rosenberg AH, Zubay G, Studier FW. 1995. Consecutive low-usage leucine codons block translation only when near the 5' end of a message in Escherichia coli. J Mol Biol 245:467-473.

Gouy M, Gautier C. 1982. Codon usage in bacteria: correlation with gene expressivity. Nucleic Acids Res 10:7055-7074.

Hsu LM, Klee HJ, Zagorski J, Fournier MJ. 1984. Structure of an Escherichia coli tRNA operon containing linked genes for arginine, histidine, leucine, and proline tRNAs. J Bacteriol 158:934-942.

Ikemura T. 1981a. Correlation between the abundance of Escherichia coli transfer RNAs and the occurrence of the respective codons in its protein genes. J Mol Biol 146:1-21.

Ikemura T. 1981b. Correlation between the abundance of Escherichia coli transfer RNAs and the occurrence of the respective codons in its protein genes: a proposal for a synonymous codon choice that is optimal for the E. coli translational system. J Mol Biol 151:389-409.

Ikemura T. 1985. Codon usage and tRNA content in unicellular and multicellular organisms. Mol Biol Evol 2:13-34.

Ikemura T. 1989. Purification of RNA molecules by gel techniques. Methods Enzymol 180:14-25.

Imamura H JB, Wakagi T, Matsuzawa H. 1999. High level expression of Thermococcus litoralis 4-alpha-glucanotransferase in a soluble form in Escherichia coli with a novel expression system involving minor arginine tRNAs and GroELS. FEBS Letters 457:393-396.

Imamura H, Jeon B, Wakagi T, Matsuzawa H. 1999. High level expression of Thermococcus litoralis 4-alpha-glucanotransferase in a soluble form in Escherichia coli with a novel expression system involving minor arginine tRNAs and GroELS. FEBS Lett 457:393-396.

Inokuchi H, Yamao F. 1994. Chapter 3: Structure and Expression of Prokaryotic tRNA Genes. In: Soll DaR, U., Editor. tRNA: Structure, Biosynthesis, and Function. Washington, D.C.: ASM Press.

Jinks-Robertson S. 1987. Chapter 85: Ribosomes and tRNA. In: Neidhart Fea, Editor. *Escherchia coli* and *Salmonella typhimurium* Cellular and Molecular Biology. Washington, D.C.: American Society of Microbiology.

Kanaya S, Kinouchi M, Abe T, Kudo Y, Yamada Y, Nishi T, Mori H, Ikemura T. 2001. Analysis of codon usage diversity of bacterial genes with a self-organizing map (SOM): characterization of horizontally transferred genes with emphasis on the E. coli O157 genome. Gene 276:89-99.

Kane JF. 1995. Effects of rare codon clusters on high-level expression of heterologous proteins in Escherichia coli. Curr Opin Biotechnol 6:494-500.

Keener J, Nomura M. 1996. Chapter 90: Regulation of Ribosome Syntheis. In: Neidhardt Fea, Editor. *Escherichia coli* and *Salmonella* Cellular and Molecular Biology. Washington, D.C.: ASM Press.

KPL KPL, Inc. 2003. BluePhos®Microwell Phosphatase Substrate System.

Kurland CG. 1991. Codon bias and gene expression. FEBS Lett 285:165-169.

Li H. 1999. Lecture 18: Extrachromosomal elements, plasmiids and selectable markers. In:Tallahassee: Florida State University.

Li Z, Deutscher MP. 1994. The role of individual exoribonucleases in processing at the 3' end of Escherichia coli tRNA precursors. J Biol Chem 269:6064-6071.

Li Z, Deutscher MP. 1996. Maturation pathways for E. coli tRNA precursors: a random multienzyme process in vivo. Cell 86:503-512.

Li Z, Deutscher MP. 2002. RNase E plays an essential role in the maturation of Escherichia coli tRNA precursors. Rna 8:97-109.

Lodish H, Baltimore D, Berk A, Zipursky S, Matsudaira P, Darnell J. 1998. Molecular Cell Biology. New York: Scientific American Books, Inc.

Ma C, Kudlicki W, Odom OW, Kramer G, Hardesty B. 1993. In vitro protein engineering using synthetic tRNA(Ala) with different anticodons. Biochemistry 32:7939-7945.

Makrides SC. 1996. Strategies for achieving high-level expression of genes in Escherichia coli. Microbiol Rev 60:512-538.

Mans RM, Pleij CW, Bosch L. 1991. tRNA-like structures. Structure, function and evolutionary significance. Eur J Biochem 201:303-324.

McCabe PC. 1990. Production of Single-Stranded DNA by Asymmetric PCR. In: White TJ, Editor. PCR Protocols: A Guide to Methods and Applications. San Diego, CA: Academic Press, Inc.

McNulty DE, Claffee BA, Huddleston MJ, Porter, ML, Cavnar KM, Kane JF. 2003. Mistranslational errors associated with the rare arginine codon CGG in Escherichia coli. Protein Expr Purif 27:365-374.

Mullis KB. 1990. The unusual origin of the polymerase chain reaction. Sci Am 262:56-61, 64-55.

O'Connor M, Willis NM, Bossi L, Gesteland RF, Atkins JF. 1993. Functional tRNAs with altered 3' ends. Embo J 12:2559-2566.

Ow MC, Kushner SR. 2002. Initiation of tRNA maturation by RNase E is essential for cell viability in E. coli. Genes Dev 16:1102-1115.

Perret V, Garcia A, Grosjean H, Ebel JP, Florentz C, Giege R. 1990. Relaxation of a transfer RNA specificity by removal of modified nucleotides. Nature 344:787-789.

Ross W, Aiyar SE, Salomon J, Gourse RL. 1998. Escherichia coli promoters with UP elements of different strengths: modular structure of bacterial promoters. J Bacteriol 180:5375-5383.

Ross W, Gosink KK, Salomon J, Igarashi K, Zou C, Ishihama A, Severinov K, Gourse RL. 1993. A third recognition element in bacterial promoters: DNA binding by the alpha subunit of RNA polymerase. Science 262:1407-1413.

Saiki R. 1990. Amplification of Genomic DNA. In: White TJ, Editor. PCR Protocols: A Guide to Methods and Applications. San Diego, CA: Academic Press, Inc.

Saks ME, Sampson JR. 1995. Evolution of tRNA recognition systems and tRNA gene sequences. J Mol Evol 40:509-518.

Sambrook J, Russell DW. 2001. Molecular Cloning: A Laboratory Manual. Plainview, NY: Cold Spring Harbor Laboratory Press.

Snyder L, Champness W. 1997. Molecular Genetics of Bacteria. Washington, D.C.: AMS Press.

Sorensen HP, Sperling-Petersen HU, Mortensen KK. 2003. Production of recombinant thermostable proteins expressed in Escherichia coli: completion of protein synthesis is the bottleneck. J Chromatogr B Analyt Technol Biomed Life Sci 786:207-214.

Stein M, Varricchio F. 1974. Separation of t-RNAs by two-dimensional polyacrylamide gel electrophoresis. Anal Biochem 61:112-119.

Tan WS, Dyson MR, Murray K. 2003. Hepatitis B virus core antigen: enhancement of its production in Escherichia coli, and interaction of the core particles with the viral surface antigen. Biol Chem 384:363-371.

Travers AA. 1984. Conserved features of coordinately regulated E. coli promoters. Nucleic Acids Res 12:2605-2618.

Watson JD, Hopkins NH, Roberts JW, Steitz JA, Weiner AM. 1987. Chapter 14: Involvement of RNA in Protein Synthesis. In: Molecular Biology of the Gene. Menlo Park, CA: The Benjamin/Cummings Publishing Company, Inc.

Wu X, Jornvall H, Berndt KD, Oppermann U. 2004. Codon optimization reveals critical factors for high level expression of two rare codon genes in Escherichia coli: RNA stability and secondary structure but not tRNA abundance. Biochem Biophys Res Commun 313:89-96.

www.kazusa.or.jp/codon/. Codon Usage Tables.

Zahn K, Landy A. 1996. Modulation of lambda integrase synthesis by rare arginine tRNA. Mol Microbiol 21:69-76.

Zdanovsky AG, Zdanovskaia MV. 2000. Simple and efficient method for heterologous expression of clostridial proteins. Appl Environ Microbiol 66:3166-3173.

BIOGRAPHICAL SKETCH

Education

1999-Present Florida State University Tallahassee, Florida
Degree Sought: Masters, Major: Molecular Biology, December 2004

Tallahassee Memorial Hospital Tallahassee, Florida
Licensure: Medical Technologist (ASCP) and State of Florida

1991-1995 Florida State University Tallahassee, Florida
Degree: Bachelor of Science, Major: Biological Science

Professional Experience

2000-2003 FSU Department of Biology Tallahassee, Florida
Graduate Assistant

- Teach Prokaryotic Laboratory for both biology majors and for the health sciences
- Teach DIS students laboratory techniques in molecular biology

1997-Present Tallahassee Memorial Hospital Tallahassee, Florida
Medical Technologist, Department of Microbiology

- Responsible for culture preparation, immunology, serology, and parasitology testing
- Responsible for blood and body fluid testing
- Responsible for interpretation of results and reporting results to physicians and pathologists

2003-Present FSU College of Medicine
Senior Biological Scientist

- Responsible for maintaining and managing neuroscience lab
- Perform neuroscience experiments on mouse and rat brain using immunohistochemical and molecular biological techniques
- Involved in grant writing and the publication of manuscripts for journals

Professional Memberships

American Society for Microbiology (ASM)
American Society of Clinical Pathologists (ASCP)

Presentations

19-23 May 2001:
Eggers, K, Porter, L and Reeves, RH.
"The expression of the *argX* tRNA gene in *Escherichia coli"*; Poster presentation
101st General Meeting of the American Society for Microbiology
Orlando, Florida

Publications

McNulty DE, Claffee BA, Huddleston MJ, Porter, ML, Cavnar KM, Kane JF. 2003. Mistranslational errors associated with the rare arginine codon CGG in Escherichia coli. Protein Expr Purif 27:365-374.