REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
11-08-2010	Final Report	22-Jul-2004 - 21-Jul-2010

4. TITLE AND SUBTITLE

Site Specific Incorporation of Amino Acid Analogues into Proteins In Vivo

5a. CONTRACT NUMBER
W911NF-04-1-0353

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
611102

6. AUTHORS
Uttam L. RajBhandary

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
Massachusetts Institute of Technology
Office of Sponsored Programs
Bldg. E19-750
Cambridge, MA 02139 -4307

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
45633-LS.4

12. DISTRIBUTION AVAILIBILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT
The goal of this project is to develop methods for the site-specific insertion in vivo of one or more amino acid analogues into proteins in eubacteria and in eukaryotes. The amino acids to be used include those that are photoactivatable, those that are fluorescent, those that carry reactive side chains such as keto groups, heavy atoms such as iodine, those that act as spectroscopic probes, and those that mimic phosphoserine, phosphothreonine, or phosphotyrosine. Besides providing a method for the production of new biomateirals with novel chemical and

15. SUBJECT TERMS
Suppressor tRNAs, Protein Synthesis, Unnatural Amino Acids, Mutant Aminoacyl-tRNA Synthetases

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Uttam RajBhandary
UU	UU	UU	UU		19b. TELEPHONE NUMBER 617-253-4702

Report Title

Site Specific Incorporation of Amino Acid Analogues into Proteins In Vivo

ABSTRACT

The goal of this project is to develop methods for the site-specific insertion in vivo of one or more amino acid analogues into proteins in eubacteria and in eukaryotes. The amino acids to be used include those that are photoactivatable, those that are fluorescent, those that carry reactive side chains such as keto groups, heavy atoms such as iodine, those that act as spectroscopic probes, and those that mimic phosphoserine, phosphothreonine, or phosphotyrosine. Besides providing a method for the production of new biomateirals with novel chemical and biological properties, proteins carrying such amino acid analogues will have wide applications in biology. These include studies on the folding, structure, stability and function of proteins, protein-protein interactions, protein localization and protein dynamics in vivo and signal transduction.

List of papers submitted or published that acknowledge ARO support during this reporting period. List the papers, including journal references, in the following categories:

(a) Papers published in peer-reviewed journals (N/A for none)

Number of Papers published in peer-reviewed journals: 0.00

(b) Papers published in non-peer-reviewed journals or in conference proceedings (N/A for none)

Number of Papers published in non peer-reviewed journals: 0.00

(c) Presentations

Number of Presentations: 0.00

Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Number of Non Peer-Reviewed Conference Proceeding publications (other than abstracts): 0

Peer-Reviewed Conference Proceeding publications (other than abstracts):

Number of Peer-Reviewed Conference Proceeding publications (other than abstracts): 0

(d) Manuscripts

NA

Number of Manuscripts: 0.00

Patents Submitted

Patents Awarded

Graduate Students

NAME	PERCENT_SUPPORTED
Christian Boehm	0.00
Mary Jane Tsang Mui Ching	0.00
Mayur Goyal	0.00
FTE Equivalent:	**0.00**
Total Number:	**3**

Names of Post Doctorates

NAME	PERCENT_SUPPORTED
FTE Equivalent:	
Total Number:	

Names of Faculty Supported

NAME	PERCENT_SUPPORTED	National Academy Member
Uttam RajBhandary	1.00	No
FTE Equivalent:	**1.00**	
Total Number:	**1**	

Names of Under Graduate students supported

NAME	PERCENT_SUPPORTED
Christian Boehm	0.00
Mayur Goyal	0.00
Mary Jane Tsang Mui Ching	
FTE Equivalent:	**0.00**
Total Number:	**3**

Student Metrics

This section only applies to graduating undergraduates supported by this agreement in this reporting period

The number of undergraduates funded by this agreement who graduated during this period:	1.00
The number of undergraduates funded by this agreement who graduated during this period with a degree in science, mathematics, engineering, or technology fields:.....	1.00
The number of undergraduates funded by your agreement who graduated during this period and will continue to pursue a graduate or Ph.D. degree in science, mathematics, engineering, or technology fields:.....	1.00
Number of graduating undergraduates who achieved a 3.5 GPA to 4.0 (4.0 max scale):	1.00
Number of graduating undergraduates funded by a DoD funded Center of Excellence grant for Education, Research and Engineering:	0.00
The number of undergraduates funded by your agreement who graduated during this period and intend to work for the Department of Defense	0.00
The number of undergraduates funded by your agreement who graduated during this period and will receive scholarships or fellowships for further studies in science, mathematics, engineering or technology fields:	1.00

Names of Personnel receiving masters degrees

NAME
Total Number:

Names of personnel receiving PHDs

NAME
Total Number:

Names of other research staff

NAME	PERCENT_SUPPORTED	
Caroline Koehrer	0.50	No
FTE Equivalent:	**0.50**	
Total Number:	**1**	

Sub Contractors (DD882)

Inventions (DD882)

List of Appendixes, Illustrations and Tables

Tables

Figures

Figure 3. Acid urea PAGE/Northern blot analysis of additional mutants derived from *hsup2am*, *hsup2oc* and *hsup2op* tRNAs. **(A)** amber suppressor series; **(B)** ochre suppressor series; **(C)** opal suppressor series. Suppressor tRNAs were visualized by RNA blot hybridization using a 5'-^{32}P-labeled oligonucleotide complementary to nucleotides 57-72 of $tRNA^{Gln}$. A 5-^{32}P-labeled oligonucleotide complementary to nucleotides 7-22 of the human $tRNA^{Ser}$ was used as internal standard (data not shown) for quantitation of RNA and aminoacylation levels by PhosphorImager analysis. QRS, *E. coli* GlnRS.

Figure 4. Firefly luciferase activity in cell extracts of HEK293T cells transfected with plasmids carrying the genes for *hsup2/C32A38am, hsup2/C32A38oc* and *hsup2/C32A38op* tRNA and *E. coli* GlnRS (QRS) or *E. coli* TrpRS (WRS) as indicated. Cells were also co-transfected with a plasmid encoding the reporter RLucFLuc fusion protein with the appropriate amber, ochre or opal mutation to measure suppression activity. Luciferase activities are given as relative luminescence units (RLU) per μg of total cell protein.

Figure 5. *E. coli* GlnRS and *E. coli* TrpRS-dependent amber, ochre and opal suppression in HEK293T cells. Immunoblot analysis of proteins isolated from cells co-transfected with plasmids carrying the genes encoding the luciferase reporter, *hsup2/C32A38am, hsup2/C32A38oc* or *hsup2/C32A38op* tRNA and, when present, *E. coli* GlnRS (EcQRS) or *E. coli* TrpRS (EcWRS). The RLucFLuc fusion protein was detected with an anti-RLuc antibody. *, protein cross-reacting unspecifically with anti-RLuc antibody.

Figure 6. Synthesis of 2-amino 5-keto hexanoic acid **(2)** from 6-diazo 2-amino 5-keto hexanoic acid **(1)** using hydroiodic acid.

Figure 7. NMR spectra of 6-diazo 2-amino 5-keto hexanoic acid **(A)** and 2-amino 5-keto hexanoic acid (crude product) after neutralization of hydroiodic acid with sodium hydroxide **(B)**.
Figure 8. Substrate binding in *E. coli* GlnRS. Schematic diagram showing the ternary complex of GlnRS with $tRNA^{Gln}$ and glutamine (re-printed from Rath *et al.*, 1998).

Figure 8. Substrate binding in *E. coli* GlnRS. Schematic diagram showing the ternary complex of GlnRS with $tRNA^{Gln}$ and glutamine (re-printed from Rath *et al.*, 1998).

Figure 9. Activities of *E. coli* GlnRS wild-type and mutants Y211F Q255L (YFQL) and Y211F Q255M (YFQM) towards various amino acids. *In vitro* aminoacylation was performed using 10 mM of amino acid as indicated; 1, Glutamine; 2, Leucine; 3, Isoleucine; 4, Methionine; 5, Methionine sulfoxide. aa, amino acid.

Figure 10. Random mutagenesis of *E. coli* GlnRS at specific sites. **(A)** Mutagenesis of D212, D219, S227, and C229 in a Y211F/Q255L background. **(B)** Mutagenesis of Y211, D212, D219, S227, C229, and Q255.

Figure 11. Positive and negative selections in *S. cerevisiae* for *E. coli* GlnRS mutants that use 2-amino 5-keto hexanoic acid. **(A)** Plasmids generated. **(B)** Overall scheme for selecting the desired mutant GlnRSs.

Figure 12. **(A)** Plasmids and *S. cerevisiae* strain for positive and negative selection of *E. coli* GlnRS mutants in yeast. The GlnRS gene and GAL4 (T44amber, R110amber) gene are expressed under the control of the constitutive ADH1 promoter. **(B)** Strategy for positive and negative selection of *E. coli* GlnRS mutants which utilize 2-amino 5-keto hexanoic acid instead of glutamine in yeast.

Figure 13. *In vivo* aminoacylation state of *Bst*-Yam suppressor tRNA in mammalian cells. Total RNA was isolated under acidic conditions from HEK293T cells transfected with plasmids carrying the genes for *Bst*-Yam and *E. coli* TyrRS (wild-type, AcpRS, or BzpRS). Cells were cultured in the absence and presence of Acp or Bzp as indicated.

tRNAs were analyzed by acid urea PAGE followed by Northern hybridization using a DNA oligonucleotide specific for *Bst*-Yam. An oligonucleotide specific for the human initiator $tRNA_i^{Met}$ was used as internal standard for quantitation of RNA and aminoacylation levels by PhosphorImager analysis. aa, amino acid; pre-Yam, *Bst*-Yam precursor.

Figure 14. Schematic representation of the secondary structure of G protein-coupled receptors (GPCRs). The 7-TM helices (I – VII) and the cationic amphipathic helix VIII are shown as cylinders. Positions in CCR5 (★) and rhodopsin (●) subjected to site-specific incorporation of unnatural amino acids are indicated.

Figure 15. Expression of functional CCR5 mutants containing Acp or Bzp at positions 28, 96, or 260. HEK293T cells were transfected with plasmids carrying the genes for CCR5-wt or CCR5 mutant with an amber mutation at position I28, F96, or F260. Plasmids encoding *Bst*-Yam and *E. coli* TyrRS (AcpRS or BzpRS) were co-transfected, and the corresponding unnatural amino acids (Acp or Bzp) were provided in the cell media as indicated. **(A)** Cell lysates were analyzed by immunoblot using an antibody against the C-terminal 1D4 epitope of CCR5 (top) and RANTES-induced calcium flux (bottom). Equal amounts of total protein were analyzed per lane. Calcium flux traces represent the levels of Fluo-3 fluorescence emitted over time by cells in suspension. The addition of RANTES (20 nM) is indicated by arrows. Relative CCR5 activities and EC_{50} values for RANTES-induced calcium flux are summarized underneath. **(B)** Cell surface expression of CCR5-wt and CCR5 mutants containing Acp. Fluorescence images of non-permeabilized cells labeled with a FITC-conjugated anti-CCR5 antibody against the extracellular N-terminal 2D7 epitope of CCR5 (N; top panel), and permeabilized cells labeled with an antibody against the cytoplasmic C-terminal 1D4 tag (C; bottom panel) of CCR5 followed by staining with an Alexa 488-conjugated secondary antibody. Blue, nuclei stained with DAPI. Green, labeled CCR5.

Figure 16. Site-specific incorporation of Acp into rhodopsin and *in vitro* labeling of rhodopsin mutants. **(A)** *In vitro* labeling of rhodopsin mutants (rho-Y29Acp, rho-Y102Acp, and rho-Y274Acp) with fluorescein hydrazide. Rhodopsin concentrations were confirmed by difference UV-vis spectra, and the loading volume was normalized to the same amount of rhodopsin based on the 500 nm absorbance. Samples were separated by 12% SDS-PAGE; the fluorescence image was taken with a Typhoon 9400 Image Scanner using an excitation/emission filter set optimized for fluorescein detection. **(B)** Stoichiometric ratios of fluorescein/rhodopsin (F/R) were

determined after normalizing the amount of rhodopsin based on the absorbance at 500 nm. **(C)** Difference spectra of fluorescein-labeled rohodopsin mutants were generated by subtracting the photobleached spectra from the dark spectra. wild-type (black); rho-Y29Acp (red); rho-Y102Acp (blue); rho-Y274Acp (green).

Figure 17. Effect of sodium butyrate. HEK293T cells were transfected with plasmid DNA (pTNT.FLuc.Y70am or wt; pSVB.Yam; pCDNA.TyrRS.wt). 3 hours post-transfection, fresh medium supplemented with serum, glutamine and sodium butyrate (1 mM final concentration) was added. Cells were harvested 48 h post-transfection and assayed for FLuc activity **(A)**. 10 ug of total protein were separated by 10% SDS-PAGE and analyzed by immunoblotting using a His4-antibody to detect FLuc protein; lane 1, mock; lane 2, FLuc.Yam; 3, FLuc.wt; molecular weight markers are shown on the right **(B)**.

Figure 18. Nonsense suppression in HEK293S *GnTI* cells. HEK293T and HEK293S *GnTI* cells were transfected with plasmid DNA (pTNT.FLuc.Y70am or wt; pSVB.Yam; pCDNA.TyrRS.wt). Cells were harvested 48 h post-transfection and assayed for FLuc activity **(A)**. 10 ug of total protein were separated by 10% SDS-PAGE and analyzed by immunoblotting using a His4-antibody to detect FLuc protein; lane 1, mock; lane 2; FLuc.am; lane 3, FLuc.wt; molecular weight markers are shown on the right **(B)**.

Figure 19. Plasmids for transient transfection into mammalian cells and generation of stable mammalian cell lines for site specific incorporation of unnatural amino acids. **(A)** Three-plasmid system for transient transfection of HEK293T cells: pSVBpUC carrying the tRNA gene (Yam), pCDNA3.1 carrying various aminoacyl-tRNA synthetase genes (*E. coli* aaRS), and pCMVTNT carrying the reporter gene (FLuc). **(B)** Plasmid for generation of stable cell lines. pBIG.H1-3xYam.aaRS carries three copies of the tRNA gene under the control of the H1 polymerase III promoter (H1-3xYam) and various aminoacyl-tRNA synthetases (TyrRS, AcpRS, AzpRS, BzpRS) under the control of a tetracycline inducible CMV promoter. Galactosidase is used for detection of positive clones during the selection of stable cell lines. The cell line used is the "HEK293S tet" cell line, generated in the Khorana laboratory for tetracycline-inducible high-level expression of opsin mutants (Reeves *et al.*, 2002).

Figure 20. Effect of copy number on tRNA expression. HEK293S tet cells were transfected in a 24-well plate with plasmids pBIG.H1-Yam.TyrRS or pBIG.H1-3xYam.TyrRS **(A)** and pTNT.FLuc.Y70am (0.5 μg each). 3 hours post-transfection, fresh medium supplemented with serum, glutamine, sodium butyrate and tetracycline (2.5 μg/ml final concentration) was added. Cells were harvested 48 h post-transfection and assayed for tRNA Yam expression using Northern blot analysis **(B)**.

Figure 21. Site-specific incorporation *in vivo* of benzoyl phenylalanine (Bzp) into MTF using a specialized suppressor tRNA/aminoacyl-tRNA synthetase (BzpRS) pair.

Figure 22. Two-plasmid system for site-specific incorporation *in vivo* of benzoyl phenylalanine (Bzp) into proteins in *E. coli*. The first plasmid, a pBAD derivative, carries the specialized amber suppressor tRNA/BzpRS pair; the second plasmid, a pACYC184 derivative, carries the protein gene of interest (MTF) with amber mutations at various positions.

Figure 23. Functional interactions between *E. coli* MTF and fMet-tRNA.

(A) Main points of contact between MTF and fMet-tRNA [taken from Schmitt *et al.*,1998].

(B) (C) Proximity of Gly41 to the A72A73 bases **(B)** and Lys206 to the 3'end of tRNA **(C)**. X-ray structure of *E.coli* MTF.fMet-tRNA complex drawn using RASMOL version 2.6 with coordinates obtained from the Protein Data Bank (PDB code 2FMT). MTF is colored gray, except Gly41 **(B)** and Lys206 **(C)** (red spacefill); the tRNA is in blue wireframe, except for positions A72A73 (**B**) and A76 **(C)** (green spacefill) [taken from Mayer *et al.*, 2002].

Figure 24. (A) Reaction of maleimide with a thiol group. **(B)** Some examples of maleimide derivatives that can be used for reaction with cysteine attached to tRNA.

Figure 25. Cloverleaf structures of amber and opal suppressor tRNAs derived from *M. jannaschii* $tRNA^{Cys}$ (MjaCysam/op). The mutated anticodon sequences, the U11 to C11 and the G37 to A37 mutations are indicated.

Figure 26. (A) Overexpression of MjaCysam/op tRNAs in *E. coli*. Lane 1, total *E. coli* tRNA; lane 2, MjaCysam; lane 3, MjaCysop. **(B)** Aminoacylation of MjaCysam/op tRNAs with *M. maripaludis* CysRS (160 pmoles of tRNA in each sample). **(C)** Suppression of amber and opal stop codons in the CAT gene by MjaCysam/op tRNAs in *E. coli*.

Figure 27. A) Overexpression of MjaCysam tRNA in *E. coli*. Lane 1, total *E. coli* tRNA; lane 2, MjaCysam. **(B)** Affinity purification of MjaCysam using a biotinylated DNA oligonucleotide coupled to streptavidin-agarose resin. Lane 1, starting material (total tRNA from *E. coli* overexpressing MjaCysam tRNA); lane 2, unbound tRNAs; lane 3, wash; lane 4, elution 1; lane 5, elution 2.

Native polyacrylamide gels stained with ethidium bromide are shown. Position of MjaCysam tRNA is indicated with an asterisk.

Figure 28. *In vitro* aminoacylation of MjaCysam tRNA. Suppressor tRNA was aminoacylated with cysteine using *M. maripaludis* CysRS (80 pmoles of tRNA in each sample; final concentration 1.6 μM) in the absence and presence of inorganic pyrophosphatase (PPase) as indicated.

Figure 29. MjaCysam tRNA is active in suppression in *in vitro* translation systems. **(A)** *E. coli* S30 extract. **(B)** Rabbit reticulocyte extract.

CAT, chloramphenicol acetyltransferase; β-lac, β-lactamase; FLuc, firefly luciferase; CysRS, *M. maripaludis* CysRS

Samples were analyzed by SDS-polyacrylamide gelelectrophoresis followed by autoradiography. 80 pmoles of tRNA (final concentration: 3.2 μM) in each sample as indicated. CysRS was added to a final concentration of 4 μM.

Statement of the problem studied

The focus of research supported by this grant has been the development of methods for the synthesis of proteins carrying one or more unnatural amino acids at specific sites *in vivo* and *in vitro*. Approaches used involved orthogonal aminoacyl-tRNA synthetase: suppressor tRNA pairs and import of suppressor tRNAs aminoacylated with unnatural amino acids into mammalian cells.

Summary of the most important results

1. We have identified *E. coli* GlnRS and *E. coli* $tRNA^{Gln}$ derived ochre and opal suppressors as orthogonal synthetase:suppressor tRNA pairs for use in yeast and in mammalian cells. Using these systems, we demonstrated concomitant suppression of all three termination codons in a mRNA in mammalian cells. We also showed that the $tRNA^{Gln}$ derived amber and opal suppressors can be aminoacylated with tryptophan by *E. coli* WRS. These results open up the possibility of using the same suppressor tRNAs for insertion of unnatural amino acid analogues related to either glutamine or tryptophan into proteins in mammalian cells.

2. The highly efficient incorporation of unnatural amino acids to two different GPCRs highlights the potential application of unnatural amino acid mutagenesis to this most ubiquitous of transmembrane receptors involved in diverse signaling pathways. It also highlights the need to isolate stale cell lines that can be used for this purpose. The availability of such cell lines will greatly expand the application of unnatural amino acid mutagenesis towards investigations of many important biological processes and make the technology available to a much broader community of investigators.

3. The identification of a new orthogonal synthetase:suppressor tRNA pair consisting of an archaeal CysRS and amber and opal suppressors derived from $tRNA^{Cys}$ opens up the possibility of a completely general method for the introduction of a variety of unnatural amino acids into proteins *in vitro* and possibly in mammalian cells (details follow).

REPORT

We are working on several approaches to the site-specific insertion of one or more unnatural amino acids into proteins in eubacteria and in eukaryotes. These include: (1) the identification of new orthogonal aminoacyl-tRNA synthetase: suppressor tRNA pairs; (2) import of suppressor tRNA aminoacylated with unnatural amino acids into mammalian cells and (3) synthesis of proteins carrying unnatural amino acids *in vitro*.

1. Identification of *E. coli* glutaminyl-tRNA synthetase (GlnRS) and ochre (UAA) and opal (UGA) suppressors derived from *E. coli* glutamine tRNA ($tRNA^{Gln}$) as orthogonal synthetase.tRNA pairs for use in yeast and in mammalian cells.

We showed previously that an amber (UAG) suppressor derived from *E. coli* $tRNA^{Gln}$ was an orthogonal suppressor tRNA in mammalian cells. We have now shown that ochre (UAA) and opal (UGA) suppressor tRNAs derived from *E. coli* $tRNA^{Gln}$ are also orthogonal in mammalian cells in that they are not aminoacylated by any of the mammalian aminoacyl-tRNA synthetases (Köhrer et al., 2004; Köhrer and RajBhandary, 2008). Through further mutagenesis of these suppressor tRNAs, we have generated amber, ochre and opal suppressor tRNAs, all of which are highly active in suppression in mammalian cells (Figure 1). The *E. coli* $tRNA^{Gln}$ derived amber suppressor tRNA has even higher activity in suppression than that of the homologous amber suppressor derived from human serine tRNA. Each suppressor tRNA is specific for its cognate codon (Table 1). Activity of the tRNAs in suppression in mammalian cells is dependent upon expression of *E. coli* GlnRS, thus none of the suppressor tRNAs are aminoacylated by mammalian aminoacyl-tRNA synthetases (Table 2 and Figures 2 and 3). Using these suppressor tRNAs, we have shown, for the first time, that it is possible to suppress concomitantly all three termination codons in a single mRNA (Table 3). This work has also provided us with suppressor tRNAs with a range of activities in suppression (36 fold increase for the amber suppressor, 156 fold for the ochre suppressor and 200 fold for the opal suppressor). The dependence of these suppressor tRNAs on *E. coli* GlnRS for activity in suppression also opens up the possibility of regulating the activity of the suppressor tRNAs in mammalian cells by regulating the expression of *E. coli* GlnRS. This can be achieved by using the tetracycline regulated system that we had described previously for regulated suppression of amber codons in mammalian cells.

We have found that the amber and opal suppressor tRNAs from *E. coli* $tRNA^{Gln}$ are also good substrates for *E. coli* tryptophanyl-tRNA synthetase (TrpRS). With these suppressor tRNAs, suppression of amber and opal codons in mammalian cells is dependent upon expression of either *E. coli* GlnRS or *E. coli* TrpRS. The amount of the firefly luciferase reporter protein obtained in mammalian cells by suppression of amber and opal codons in luciferase mRNA in cells expressing *E. coli* TrpRS is comparable to those expressing *E. coli* GlnRS (Figures 4 and 5). These results open up the possibility of using the same suppressor tRNAs for insertion of unnatural amino acid analogues related to either glutamine or tryptophan into proteins in mammalian cells. Schultz and coworkers showed previously that *Bacillus subtilis* TrpRS does not aminoacylate any of the tRNAs in mammalian cells and used a mutant TrpRS for introduction of 5'-hydroxy tryptophan, a fluorescent amino acid, site-specifically into proteins in mammalian cells. The same approach can now be used for the introduction of tryptophan

analogues into proteins in mammalian cells using the amber or the opal suppressor derived from *E. coli* $tRNA^{Gln}$.

2. Attempts at isolation of mutants of *E. coli* GlnRS, which aminoacylate amber suppressor tRNAs with the keto analogue of glutamine, 2-amino 5-keto hexanoic acid

Our goal is to use the mutant GlnRS in yeast or in mammalian cells for the site-specific insertion of the above keto analogue of glutamine into proteins. As a first step towards this goal, starting from 6-diazo 2-amino 5-keto hexanoic acid, we have synthesized the desired analogue in a single step reaction using hydroiodic acid (Figure 6). The NMR spectrum of the product (Figure 7) obtained after neutralization of hydroiodic acid with sodium hydroxide shows that the desired product has been formed.

The crystal structure of *E. coli* GlnRS complexed to its tRNA substrate and to glutamine (Figure 8) shows that specificity of the enzyme for glutamine derives mainly from hydrogen bonds between the carboxamide ($-CONH_2$) part of the side chain to Arg30 (not included in the Figure), a water molecule and Tyr211. The water 1050 and Tyr211 each form a hydrogen-bond with one of the two hydrogens of the $^{-}CONH_2$ group. In 2-amino 5-keto hexanoic acid, the $-CONH_2$ is replaced by $^{-}COCH_3$. Therefore, for mutagenesis of *E. coli* GlnRS, we focused on making the binding pocket in the region of the carboxamide group more hydrophobic by removing the groups involved in forming the hydrogen-bond with the $^{-}NH_2$ of the $-CONH_2$ group. We mutated separately Tyr211 to Phe, and Gln255, whose side chain is involved in positioning several of the water molecules including water 1050 to Leu and Met and also combined the mutations. The resulting mutant enzymes Y211F, Q255M, Q255L, Y211F Q255M and Y211F Q255L have been expressed in *E. coli* and purified. The double mutants Y211F Q255M and Y211F Q255L are much less active in *in vitro* aminoacylation of tRNA with glutamine than the wild type enzyme suggesting that the glutamine binding pocket of GlnRS has been perturbed enough to prevent its efficient use of glutamine as a substrate (Figure 9). These results fulfill one of the criteria necessary for use of these enzymes to incorporate analogues of glutamine into tRNA.

3. Attempts at selection of GlnRS mutants which utilize 2-amino 5-keto hexanoic acid instead of glutamine in the yeast *Saccharomyces cerevisiae*

In the next step, we introduced four random mutations in the amino acid binding pocket of *E. coli* GlnRS in a Y211F/Q255L background or introduced six random mutations at amino acid 211, 212, 219, 227, 229 and 255. DNA sequence analysis of 12 selected mutants (Figure 10) showed that the mutations were indeed random and that there was no bias in the nature of mutations at these sites

The library generated above was used to identify mutants of GlnRS which aminoacylate the $tRNA^{Gln}$ derived amber suppressor tRNA with 2-amino 5-keto hexanoic acid instead of glutamine. A combination of positive and negative selection schemes were used. The positive selection selects for those GlnRS mutants which aminoacylate the suppressor tRNA with either glutamine or 2-amino 5-keto hexanoic acid. The negative selection selects out those GlnRS mutants which aminoacylate the suppressor tRNA with glutamine. Figure 11A shows the two plasmids, one used for cloning of the amber suppressor tRNA gene and the library of GlnRS

mutants and the other one for the cloning of the URA3 gene, carrying an amber mutation used for positive and negative selections and an *E. coli* β-galactosidase (lacZ) gene used for final screening of transformants carrying the desired mutant GlnRS gene.

The strategy for isolation of the desired *E. coli* GlnRS mutant was as follows (Figure 11B). Transform a Δura leu$^-$ trp$^-$ yeast strain with the two plasmids. Select for growth in a medium lacking leu, trp and uracil but containing glutamine and 2-amino 5-keto hexanoic acid. This selects for GlnRS mutants which can use either glutamine or the glutamine analogue for aminoacylation of the suppressor tRNA and suppression of the amber mutation in the URA3 gene. Cells that grow in this medium are pelleted and then suspended in medium lacking leu and trp but containing glutamine, uracil, and 5-fluoroorotic acid (5-FOA). Cells carrying mutants of GlnRS, which still aminoacylate the amber suppressor tRNA with glutamine, will suppress the amber mutation in the URA3 gene. These cells will convert 5-FOA to the toxic compound 5-fluorouracil and die. In contrast, cells carrying mutants of GlnRS which can only aminoacylate the suppressor tRNA with 2-amino 5-keto hexanoic acid will survive. Colonies that grow in this medium can be further screened for their ability to suppress the amber mutation in the *E. coli* lacZ gene. We also used a different strategy of positive and negative selection described by Schultz and coworkers (Chin et al., 2003) for use in yeast (Figure 12).

Unfortunately, in spite of several attempts, we were unable to identify any mutants of GlnRS which utilized the amino acid 2-amino 5-keto hexanoic acid. It is possible that we need to go through several rounds of positive and negative selection to obtain the desired mutant. It is also possible that the altered amino acid binding pocket of the GlnRS mutants will accommodate larger analogues such as 2-amino 5-keto heptanoic acid or 2-amino 5-keto octanoic acid. This remains to be tested.

4. Site-specific incorporation of p-acetyl phenylalanine and p-benzoyl phenylalanine into G-protein coupled receptors (GPCRs)

Along with development of new systems, we are also working on optimizing existing systems for incorporating the unnatural amino acids p-acetyl phenylalanine (Acp) and p-benzoyl phenylalanine (Bzp) into proteins with the aim of applying it for the synthesis of membrane proteins such as G protein-coupled receptors (GPCRs). This work carried out in collaboration with Dr. Sakmar of Rockefeller University involves the chemokine receptor CCR5 (a co-receptor of human immunodeficiency virus-1; HIV-1) and rhodopsin (the photoreceptor of the eye).

Using a FLuc gene carrying an amber codon as the reporter, we have shown that Acp and Bzp can be incorporated site-specifically into FLuc at high efficiency (Ye, Köhrer et al., 2008; Köhrer and RajBhandary, 2009). This reaction requires the presence of the FLuc gene with a site-specific amber codon, the mutant *E. coli* tyrosyl-tRNA synthetases (AcpRS or BzpRS) and the unnatural amino acids Acp or Bzp (Table 4). As expected, aminoacylation of the suppressor tRNA is also dependent upon the presence of Acp or Bzp (Figure 13).

We have used this system to incorporate Acp or Bzp in mammalian cells to produce functional CCR5 or rhodopsin harboring reactive keto groups at one of three specific positions (Figure 14). Mutant CCR5 was obtained at levels of up to ~50% of wild type, based on

immunoblot analysis and ligand dependent calcium flux (Figure 15A) and shown to be expressed on the cell surface (Figure 15B). Rhodopsin containing Acp at three specific sites was also purified in high yield and reacted with fluorescein hydrazide *in vitro* to produce fluorescently labeled rhodopsin (Figure 16).

GPCRs are among the most important of transmembrane receptors on the cell surface with the human genome encoding ~865 such receptors. They are, therefore, the target of many drugs. Our work demonstrating the high efficiency of unnatural amino acid mutagenesis on two different GPCRs highlights the potential application of unnatural amino acid mutagenesis to this most ubiquitous of transmembrane receptors involved in diverse signaling pathways.

The incorporation of reactive keto groups at specific sites into GPCRs allows their reaction with different reagents to introduce a variety of spectroscopic and other probes. Furthermore, combination of cysteine labeling and unnatural amino acid mutagenesis can potentially be used to introduce two different fluorophores into a GPCR for FRET-based studies. Bzp also provides the possibility of photocrosslinking to identify precise sites of protein-protein interactions including GPCR binding to G proteins and arrestins and for understanding the molecular basis of ligand recognition by chemokine receptors.

5. Attempts at isolation of mammalian cell lines for incorporation of acetylphenylalanine, benzoylphenylalanine, or azidophenylalanine into proteins

With our success in incorporating the above unnatural amino acids into GPCRs, our efforts are now focused on isolation of stable mammalian cell lines that carry the *Bacillus stearothermophilus* tyrosine amber suppressor tRNA gene and the wild type or mutant tyrosyl-tRNA synthetase (TyrRS, AcpRS, AzpRS and BzpRS) genes. Such cell lines will have the capability of incorporating either Acp, Bzp or azido phenylalanine (Azp), a third chemically reactive amino acid, into any protein of interest and at any desired site. The only requirement would be a protein gene of interest carrying an amber codon at the desired site of incorporation of the unnatural amino acid. The availability of such cell lines will greatly expand the application of unnatural amino acid mutagenesis towards investigations of many important biological processes and make the technology available to a much broader community of investigators. Below, we describe the work so far towards this aim.

Optimization of protein production and increased frequency of amber suppression remain important goals. One of the reagents we have used for optimization of protein production in mammalian cells is Na butyrate. Figure 17 shows that HEK293T cells grown in the presence of Na butyrate, after transient transfection with a plasmid carrying the firefly luciferase gene, FLuc (wild type or an amber mutant), produce more FLuc activity (Figure 17A) and FLuc protein (Figure 17B) than cells grown in the absence of Na butyrate. Na butyrate will, therefore, continue to be used routinely during and after the isolation of stable cell lines.

Structure-function studies, which include techniques such as X-ray crystallography and NMR spectroscopy, require substantial amounts of highly purified and homogeneous material. However, integral membrane proteins often carry highly heterogeneous N-glycans that potentially interfere with such structure-function analyses. Therefore, the parental cells to be

used for isolation of cell lines will carry a mutation in the N-acetylglucosaminyl transferase I (*GnTI*) gene. This mutation blocks protein glycosylation at a specific step in the pathway and thereby yields proteins with homogeneous glycosyl modifications instead of complex and heterogeneous modifications. Figure 18 shows that efficiency of amber suppression in HEK293S *GnTI*$^{-}$ cells is essentially the same (31.48%) as in HEK293T cells (29.93%) and that yield of the suppressed FLuc protein is also approximately the same. Therefore, *GnTI*$^{-}$ cells will be used as parental cells for the isolation of stable cell lines.

For transient transfection work done till now, we have used three plasmids pSVBpUC, pcDNA 3.1 and pCMVTNT (Figure 19A). For the isolation of stable cell lines, the first two plasmids have been replaced by a single plasmid (Figure 19B) carrying the genes for the amber suppressor tRNA, the TyrRS gene and β-galactosidase. To prevent any possible toxic effects as a result of expression of the mutant TyrRS genes, our plan is to control expression of TyrRS using a tetracycline regulated system. Stable cell lines are isolated by co-transfecting a helper plasmid carrying the puromycin resistance gene and selecting for puromycin resistance. β-galactosidase will be used for the screening of positive clones.

For maximal production of the suppressor tRNA, we have inserted three copies of the tRNA gene under control of the RNA polymerase III HI promoter instead of a single copy (Figure 20A). This increase in tRNA copy number greatly increases the amounts of suppressor tRNA (Yam) and aminoacyl-suppressor tRNA (Tyr-Yam) produced in HEK293S cells (Figure 20B) and increases the amounts of FLuc activity in transiently transfected cells (Table 5). The production of aminoacylated suppressor tRNA (Figure 20B) and suppression of the amber codon in the FLuc gene (Table 5) show that the plasmid carrying the suppressor tRNA, TyrRS and β-galactosidase genes is fully functional in producing the tRNA and the TyrRS and can, therefore, be used to isolate the desired stable cell lines.

Finally, using puromycin resistance as a selectable marker, we have isolated pools of stable cell lines expressing the suppressor tRNA and wild type TyrRS or mutant forms such as AcpRS, AzpRS, BzpRS. Data from a preselected pool of puromycin resistant stable HEK293S cells show FLuc synthesis in the pool selected, suggesting that the pool of stable cell lines produces the suppressor tRNA and the aminoacyl-tRNA synthetase (Table 6). The next step is the isolation of individual clones, which produce increased amounts of FLuc.

6. Crosslinking of proteins carrying Bzp to RNA *in vivo* and *in vitro*

From assembly and packaging of RNA viruses to mRNA localization during development, the specific recognition of RNAs by proteins plays an important role in many biological processes. Examples of these biological processes include, among others, RNA processing, RNA splicing, RNA transport, ribosome assembly, translation, translation regulation, and assembly of silencing complexes. A new goal of ours, therefore, is to use proteins carrying photoactivatable groups such as Bzp to study the specificity and topology of RNA-protein interactions *in vitro* and *in vivo*. Upon irradiation at 350-360nm, benzophenones form crosslinks mostly by insertion into C-H bonds, and because they do not photodissociate, they give excellent crosslinking yields. Consequently, this work has the potential for very broad application in RNA biology. Therefore, as a proof of principle experiment on application of Bzp induced

photocrosslinking between a protein and an RNA, our aim is to introduce Bzp into specific sites, (Phe14, and amino acids 38-45) of *E. coli* methionyl-tRNA formyltransferase (MTF) (Figure 21), using the plasmids shown in Figure 22 and to study their crosslinking to the *E. coli* methionyl-tRNA *in vitro* and *in vivo*. MTF recognizes specifically the initiator methionyl-tRNA. The crystal structure of the MTF-initiator tRNA complex is known and Phe14 of MTF is in close proximity to the methionine moiety. We have shown before that amino acids 38-45, which are disordered in the structure of uncomplexed MTF, adopt a defined structure in presence of the initiator tRNA substrate and interact with nucleotides on the 3'-side of the tRNA acceptor stem, where most of the specificity determining contacts between MTF and the tRNA are (Figure 23). Crosslinking of MTF to the initiator tRNA *in vitro* and *in vivo* and identification of the crosslinking sites will provide strong evidence for the potentially powerful application of site-specific unnatural amino acid mutagenesis involving Bzp and make this approach an attractive one for studying RNA-protein interactions *in vitro* and *in vivo*.

Along these lines and in parallel, we are using a similar strategy to (i) study the topology of interaction between the *E. coli* protein synthesis initiation factor IF2 and the *E. coli* initiator tRNA and (ii) investigate whether IF2 forms a binary complex with the initiator fMet-tRNA *in vivo*.

7. Identification of a new orthogonal synthetase-suppressor tRNA pairs derived from methanogenic archaea for use in eubacteria and in mammalian cells

A general method for introduction of a variety of unnatural amino acids into proteins would be the use of an orthogonal suppressor tRNA that is aminoacylated *in vitro* with cysteine. The cysteine residue can, subsequently, be derivatized to introduce a variety of side chains (Figure 24). Towards this goal, we have cloned a cysteinyl-tRNA synthetase (CysRS) derived from *Methanococcus maripaludis* and amber and opal suppressor tRNAs derived from *Methanococcus jannaschii* cysteine tRNA in *E. coli* (Figure 25). The suppressor tRNAs are overproduced in *E. coli* (Figures 26A and 26B) and are quite active in suppression, but only in the presence of the *M. maripaludis* CysRS (Figure 26C). The requirement for expression of *M. maripaludis* CysRS shows that the amber and opal suppressor tRNAs are orthogonal in *E. coli* in that they are not aminoacylated by any of the *E. coli* aminoacyl-tRNA synthetases.

Both of the projects, import of aminoacylated suppressor tRNA and *in vitro* synthesis of proteins, require the availability of significant amounts of purified suppressor tRNAs. We have shown that the amber (Figure 27A) and opal (data not shown) suppressor tRNAs can be overproduced in *E. coli* and have purified significant amounts of the amber suppressor tRNA using a biotinylated oligonucleotide affinity column to capture the tRNA (Figure 27B).

The suppressor tRNA can be aminoacylated essentially quantitatively *in vitro* using purified *M. maripaludis* CysRS (Figure 28). The purified suppressor tRNA is active in suppression of amber codons in *E. coli* and in rabbit reticulocyte cell free systems using mRNAs carrying amber codons in chloramphenicol acetyl-transferase (CAT) (Figure 29A) or in firefly luciferase (FLuc) (Figure 29B).

Our next aim was the large scale aminoacylation of the suppressor tRNA with cysteine and derivatization of the Cys-tRNA with a fluorescent maleimide reagent for studies on import of the suppressor tRNA into mammalian cells or its use in *in vitro* protein synthesis systems. While aminoacylation of the suppressor tRNA with cysteine proceeded in good yield, we found, however, that modification of the cysteine side chain with a fluorescent maleimide reagent went poorly suggesting instability of the ester linkage of cysteine attached to the tRNA during the modification reaction.

Bibliography

1. Complete Set of Orthogonal 21st Aminoacyl-tRNA Synthetase–Amber, Ochre and Opal Suppressor tRNA Pairs: Concomitant Suppression of Three Different Termination Codons in an mRNA in Mammalian Cells. C. Köhrer, E. L. Sullivan and U. L. RajBhandary. *Nucleic Acids Res.* **32**, 6200-6211 (2004).
2. The Many Applications of Acid Urea Polyacrylamide Gel Electrophoresis to Studies of tRNAs and aminoacyl-tRNA synthetase. C. Köhrer and U.L. RajBhandary. *Methods* **44**, 129-138 (2008).
3. Progress Toward an Expanded Eukaryotic Genetic Code. J. W. Chin, A. Cropp, S. Chu, E. Meggers and P. G. Schultz. *Chem.& Biol.* **10**, 511-519 (2003).
4. Site-Specific Incorporation of Keto Amino Acids Into Functional G Protein-Coupled Receptors Using Unnatural Amino Acid Mutagenesis. S. Ye, C. Köhrer, T. Huber, M. Kazmi, P. Sachdev, E. C. Y. Yan, A. Bhagar, U. L. RajBhandary and T. P. Sakmar. *J. Biol. Chem.* **283**, 1525-1533 (2008).
5. Specialized Components of the Translational Machinery for Unnatural Amino Acid Mutagenesis: tRNA, Aminoacyl-tRNA Synthetases and Ribosomes. *In* Protein Engineering. C. Koehrer and U. L. RajBhandary (Eds.). Springer-Velag, Berlin, Heidelberg (2009) pp. 205-229.

Appendixes (tables 1-6; figures 1-29)

Table 1. Specificity of amber, ochre and opal suppression in HEK293T cells.

line	RLucFLuc	tRNA	QRS	FLuc activity (RLU/μg)
1	Q162am	---	---	1,487
2	Q162am	hsup2/C32A38am	QRS	28,510,124
3	Q162am	hsup2/C32A38oc	QRS	22,009
4	Q162am	hsup2/C32A38op	QRS	1,818
5	Q162oc	---	---	1,350
6	Q162oc	hsup2/C32A38am	QRS	1,396
7	Q162oc	hsup2/C32A38oc	QRS	3,755,288
8	Q162oc	hsup2/C32A38op	QRS	2,632
9	Q162op	---	---	9,871
10	Q162op	hsup2/C32A38am	QRS	16,589
11	Q162op	hsup2/C32A38oc	QRS	7,366
12	Q162op	hsup2/C32A38op	QRS	8,565,996

HEK293T cells were co-transfected with 0.5 μg of pRF plasmid carrying the luciferase reporter gene, 0.5 μg of pSVB plasmid carrying the tRNA gene and 5 ng (hsup2am, hsup2oc) – 10 ng (hsup2op) of pCMVTNT plasmid carrying the *E. coli* GlnRS (QRS) gene. Luciferase activities are given as relative luminescence units (RLU) per μg of total cell protein. Transfection of 0.5 μg of plasmid carrying the wild type RLucFLuc fusion gene yielded FLuc activities of 82.7×10^6 RLU/μg.

Table 2. Mutants of the orthogonal amber, ochre and opal suppressor tRNAs *hsup2* with enhanced suppressor activity in mammalian cells. The activity of all suppressor tRNAs is strictly dependent on co-expression of *E. coli* GlnRS (QRS).

line	RLucFLuc	tRNA	QRS	FLuc activity (RLU/μg)	fold increase
1	Q162am	---	---	1,487	---
2	Q162am	hsup2am	QRS	786,668	1x
3	Q162am	hsup2/C32A38am	QRS	28,510,124	36x
4	Q162am	hsup2/C32A38am	---	30,030	
5	Q162oc	---	---	1,350	
6	Q162oc	hsup2oc	QRS	24,065	1x
7	Q162oc	hsup2/C32A38oc	QRS	3,755,288	156x
8	Q162oc	hsup2/C32A38oc	---	2,108	
9	Q162op	---	---	9,871	
10	Q162op	hsup2op	QRS	42,735	1x
11	Q162op	hsup2/C32A38op	QRS	8,565,996	200x
12	Q162op	hsup2/C32A38op	---	9,063	
13	wt	---	---	82,683,171	

HEK293T cells were co-transfected with 0.5 μg of pRF plasmid carrying the luciferase reporter gene, 0.5 μg of pSVB plasmid carrying the tRNA gene and 5 ng (hsup2am, hsup2oc) – 10 ng (hsup2op) of pCMVTNT plasmid carrying the *E. coli* GlnRS (QRS) gene. Luciferase activities are given as relative luminescence units (RLU) per μg of total cell protein.

Table 3. Concomitant suppression of amber, ochre and opal codons in HEK293T cells.

line	RLucFLuc	tRNA amber	tRNA ochre	tRNA opal	QRS	FLuc activity (RLU/μg)
1	oc/am/op	---	---	---	---	26
2	oc/am/op	C32A38am	C32A38oc	C32A38op	---	256
3	oc/am/op	C32A38am	C32A38oc	C32A38op	QRS	49,987
4	wt	---	---	---	---	188,000,547

HEK293T cells were co-transfected with 0.5 μg of pRF plasmid carrying the luciferase reporter gene pRF.Y70ocY165amQ283op (oc/am/op), 0.5 μg of pSVB plasmid carrying the tRNA gene (each) and 10 ng of pCMVTNT plasmid carrying the *E. coli* GlnRS (QRS) gene. 3 hours post-transfection, cells were fed with fresh medium containing 10% serum, 10 mM sodium butyrate and 2 mM glutamine. Luciferase activities are given as relative luminescence units (RLU) per μg of total cell protein.

Table 4. Site-specific insertion of *p*-acetyl-phenylalanine (Acp) or *p*-benzoyl-phenylalanine (Bzp) into firefly luciferase (FLuc) in mammalian cells.

	FLuc	tRNA	TyrRS	aa	FLuc activity x10^6 (RLU/μg)	Relative FLuc activity (%)
1	Y70am	---	---	---	0.03	0.01
2	Y70am	Yam	---	---	0.88	0.35
3	Y70am	Yam	wt	---	86.60	34.53
4	wt	Yam	wt	---	250.82	100.00
5	Y70am	Yam	AcpRS	---	4.14	1.34
6	wt	Yam	AcpRS	---	308.92	100.00
7	Y70am	Yam	AcpRS	Acp	53.40	17.34
8	wt	Yam	AcpRS	Acp	307.88	100.00
9	Y70am	Yam	BzpRS	---	0.89	0.24
10	wt	Yam	BzpRS	---	361.53	100.00
11	Y70am	Yam	BzpRS	Bzp	15.34	7.75
12	wt	Yam	BzpRS	Bzp	198.75	100.00

HEK293T cells were transfected with plasmids carrying the genes for FLuc (wild-type or FLuc.Y70am), *Bst*-Yam, and *E. coli* TyrRS (wild-type, AcpRS, or BzpRS), and cultured in the absence and presence of Acp or Bzp at a final concentration of 0.3 mM. Cells were harvested 48 hours post-transfection and analyzed for FLuc activity. The average FLuc activities from two independent experiments are described as relative luminescence units (RLU)/μg of protein in a total cell extract and as percentage of wild-type activity (%). aa, amino acid.

Table 5. Effect of copy number on tRNA expression. HEK293S tet cells were transfected in a 24-well plate with plasmids pBIG.H1-Yam.TyrRS or pBIG.H1-3xYam.TyrRS and pTNT.FLuc.Y70am (0.5 μg each) **(see Figure 20)**. 3 hours post-transfection, fresh medium supplemented with serum, glutamine, sodium butyrate and tetracycline (2.5 μg/ml final concentration) was added. Cells were harvested 48 h post-transfection and assayed for FLuc and galactosidase activity. FLuc and galactosidase activities are given in relative luminescence units (RLU) per μg of protein in a total cell extract.

	pBIG.H1-Yam.TyrRS	pBIG.H1-3xYam.TyrRS
FLuc activity	8.7×10^6 RLU/μg	**15.3×10^6 RLU/μg**
Galactosidase activity	29.4×10^7 RLU/μg	27.8×10^7 RLU/μg

Table 6. Generation of stable cell lines expressing tRNA Yam and various aminoacyl-tRNA synthetases (e.g. TyrRS, AcpRS, AzpRS, BzpRS). Data from a pre-selected pool of stable HEK293S tet cells expressing tRNA Yam and TyrRS are shown. The day after transfection, growth medium was changed to fresh medium containing 0.5 μg/ml puromycin. Approximately 10 days later, surviving cells were transferred to fresh medium containing 0.75 or 1.0 μg/ml puromycin to increase selection pressure. A small aliquot of cells that were pre-selected at 0.5 μg/ml puromycin was transfected in a 24-well plate with pTNT.FLuc.Yam (0.5 μg) and induced with tetracycline. Cells were harvested 24 h post-transfection and assayed for FLuc and galactosidase activity. FLuc and galactosidase activities are given in relative luminescence units (RLU) per μg of protein in a total cell extract.

	pBIG.H1-3xYam.TyrRS
FLuc activity	**1.23×10^6 RLU/μg**
Galactosidase activity	8.87×10^7 RLU/μg

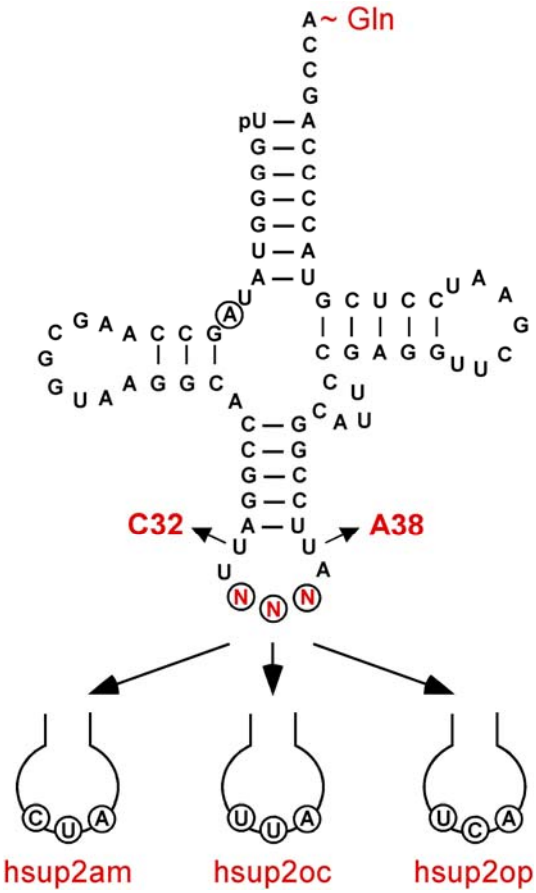


Figure 1. Cloverleaf structures of suppressor tRNAs derived from *E. coli* $tRNA^{Gln}$. The mutated anticodon sequences and the C9 to A9 mutation are circled.

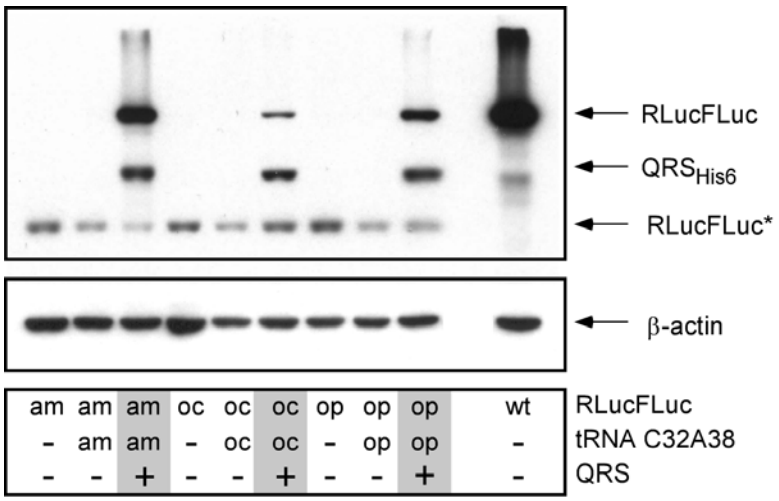


Figure 2. Amber, ochre and opal suppression in HEK293T cells. Immunoblot analysis of proteins isolated from cells co-transfected with plasmids carrying the genes encoding the luciferase reporter, *hsup2/C32A38am*, *hsup2/C32A38oc* or the *hsup2/C32A38op* tRNAs and, when present, *E. coli* GlnRS (QRS). The RLucFLuc fusion protein was detected with an anti-FLuc antibody and *E. coli* GlnRS was detected with an anti-His4-antibody. An antibody against β-actin was used as a loading control. RLucFLuc, full length fusion protein; RLucFLuc*, truncated RLucFLuc fusion protein.

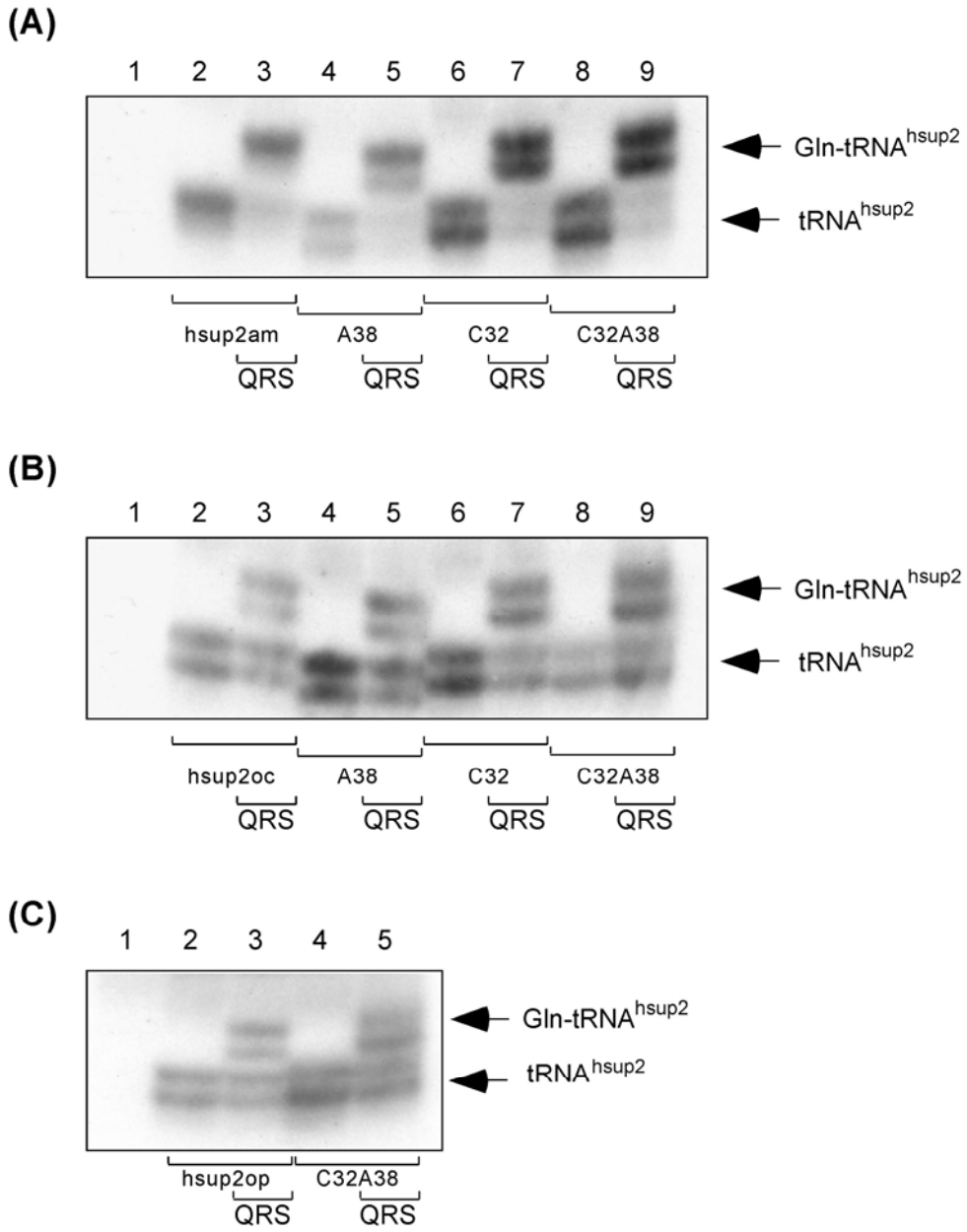


Figure 3. Acid urea PAGE/Northern blot analysis of additional mutants derived from *hsup2am*, *hsup2oc* and *hsup2op* tRNAs. **(A)** amber suppressor series; **(B)** ochre suppressor series; **(C)** opal suppressor series. Suppressor tRNAs were visualized by RNA blot hybridization using a 5'-^{32}P-labeled oligonucleotide complementary to nucleotides 57-72 of $tRNA^{Gln}$. A 5-^{32}P-labeled oligonucleotide complementary to nucleotides 7-22 of the human $tRNA^{Ser}$ was used as internal standard (data not shown) for quantitation of RNA and aminoacylation levels by PhosphorImager analysis. QRS, *E. coli* GlnRS.

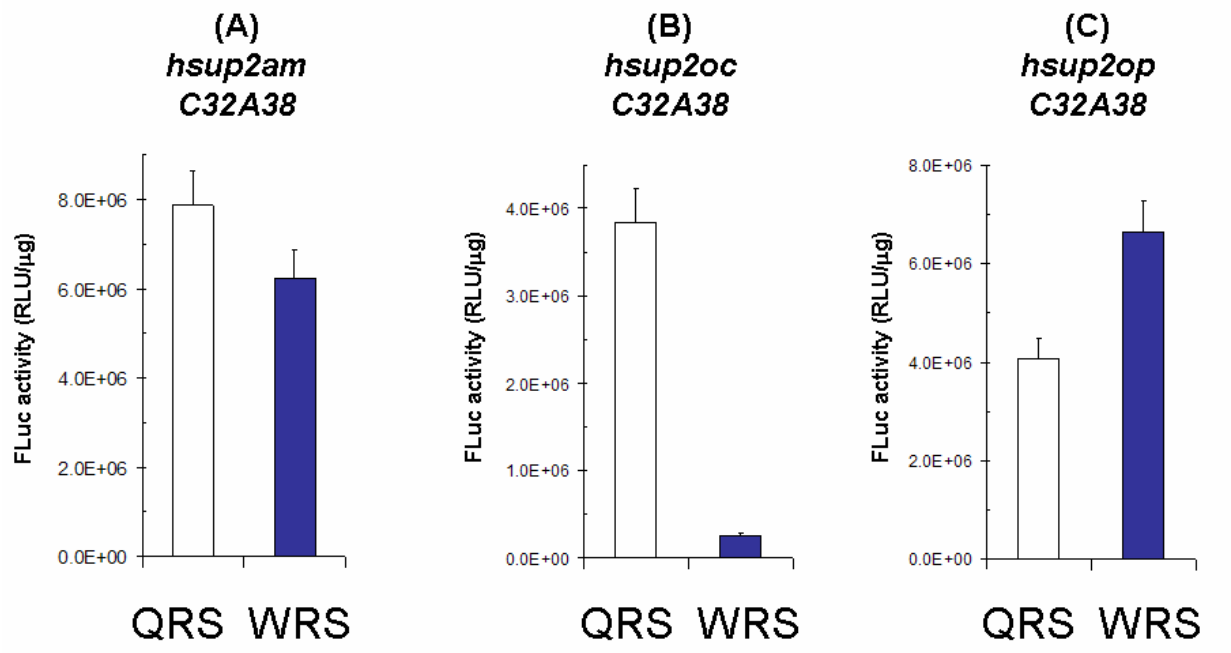


Figure 4. Firefly luciferase activity in cell extracts of HEK293T cells transfected with plasmids carrying the genes for *hsup2/C32A38am*, *hsup2/C32A38oc* and *hsup2/C32A38op* tRNA and *E. coli* GlnRS (QRS) or *E. coli* TrpRS (WRS) as indicated. Cells were also co-transfected with a plasmid encoding the reporter RLucFLuc fusion protein with the appropriate amber, ochre or opal mutation to measure suppression activity. Luciferase activities are given as relative luminescence units (RLU) per µg of total cell protein.

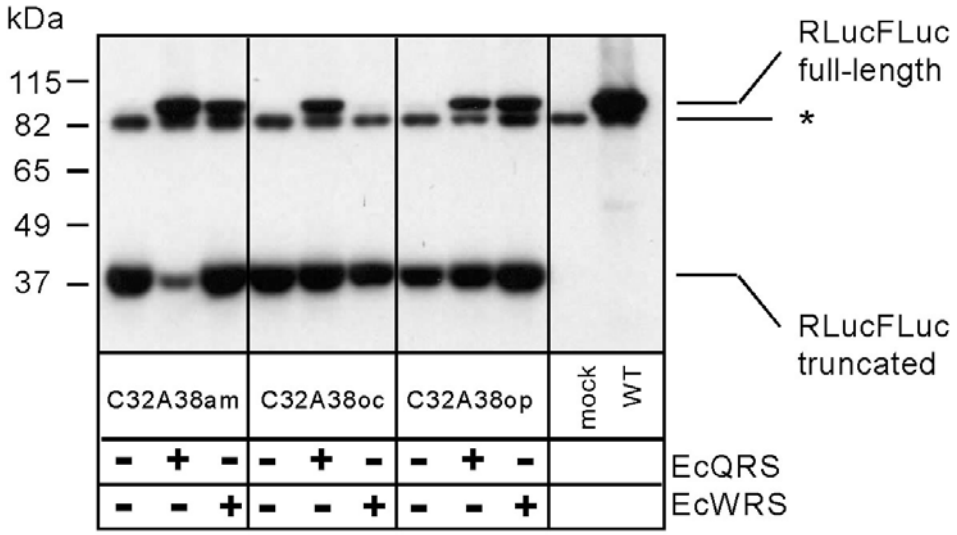
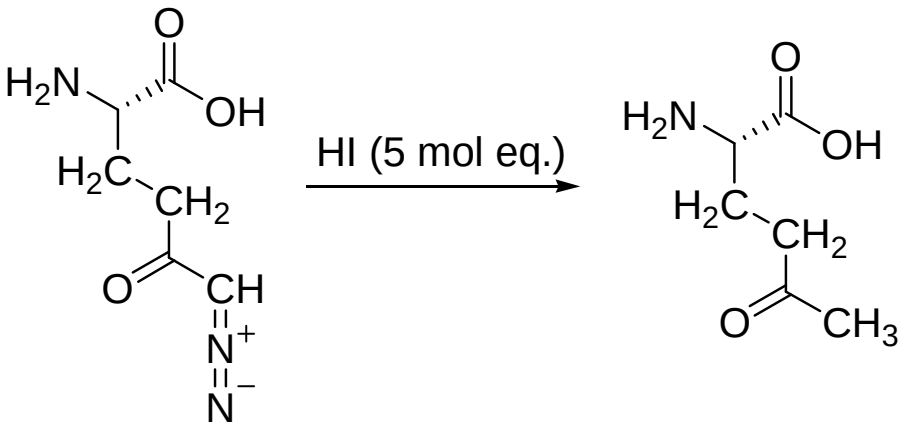


Figure 5. *E. coli* GlnRS and *E. coli* TrpRS-dependent amber, ochre and opal suppression in HEK293T cells. Immunoblot analysis of proteins isolated from cells co-transfected with plasmids carrying the genes encoding the luciferase reporter, *hsup2/C32A38am*, *hsup2/C32A38oc* or *hsup2/C32A38op* tRNA and, when present, *E. coli* GlnRS (EcQRS) or *E. coli* TrpRS (EcWRS). The RLucFLuc fusion protein was detected with an anti-RLuc antibody. *, protein cross-reacting unspecifically with anti-RLuc antibody.



(1) **(2)**

Figure 6. Synthesis of 2-amino 5-keto hexanoic acid **(2)** from 6-diazo 2-amino 5-keto hexanoic acid **(1)** using hydroiodic acid.

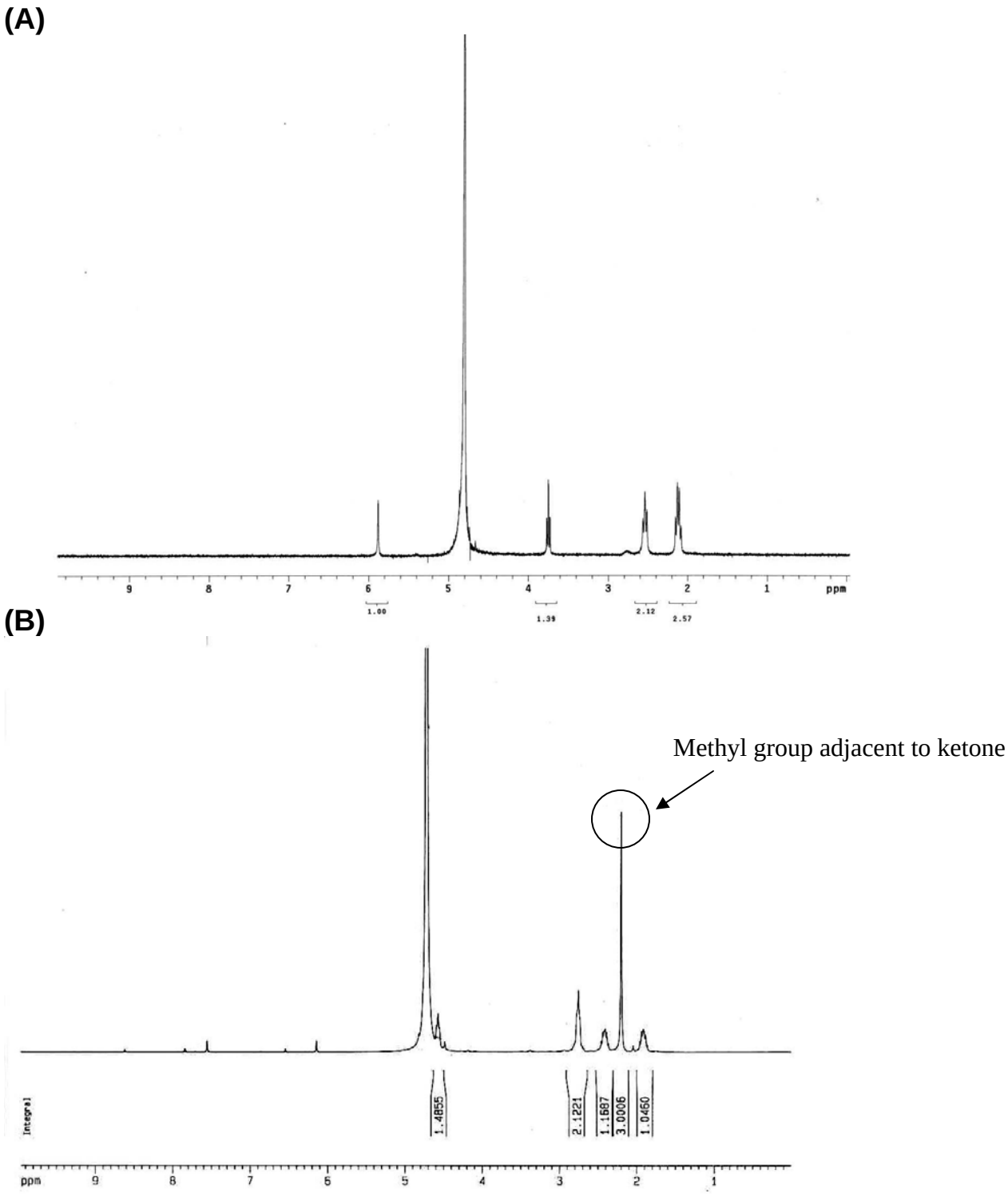


Figure 7. NMR spectra of 6-diazo 2-amino 5-keto hexanoic acid **(A)** and 2-amino 5-keto hexanoic acid (crude product) after neutralization of hydroiodic acid with sodium hydroxide **(B)**.

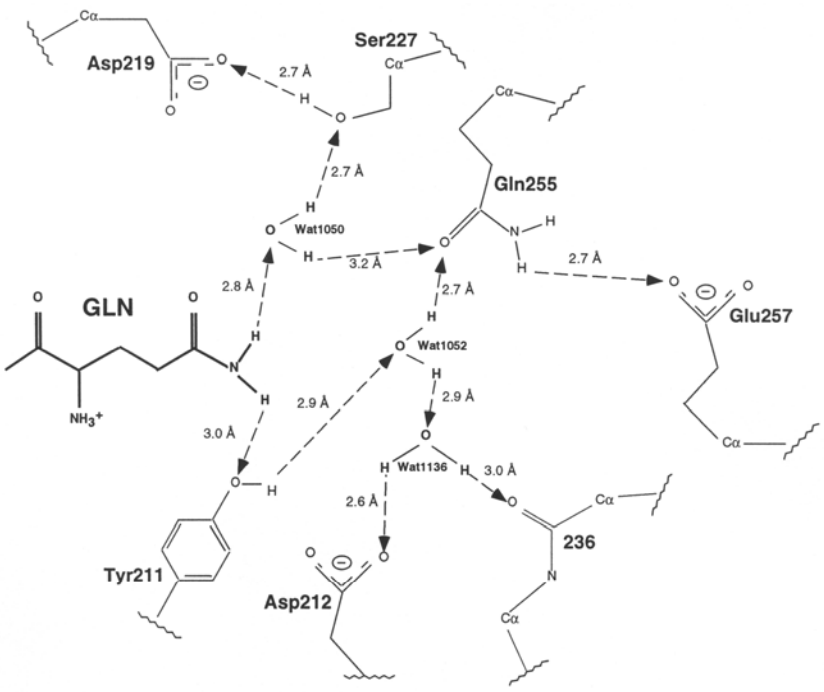


Figure 8. Substrate binding in *E. coli* GlnRS. Schematic diagram showing the ternary complex of GlnRS with $tRNA^{Gln}$ and glutamine (re-printed from Rath *et al.*, 1998).

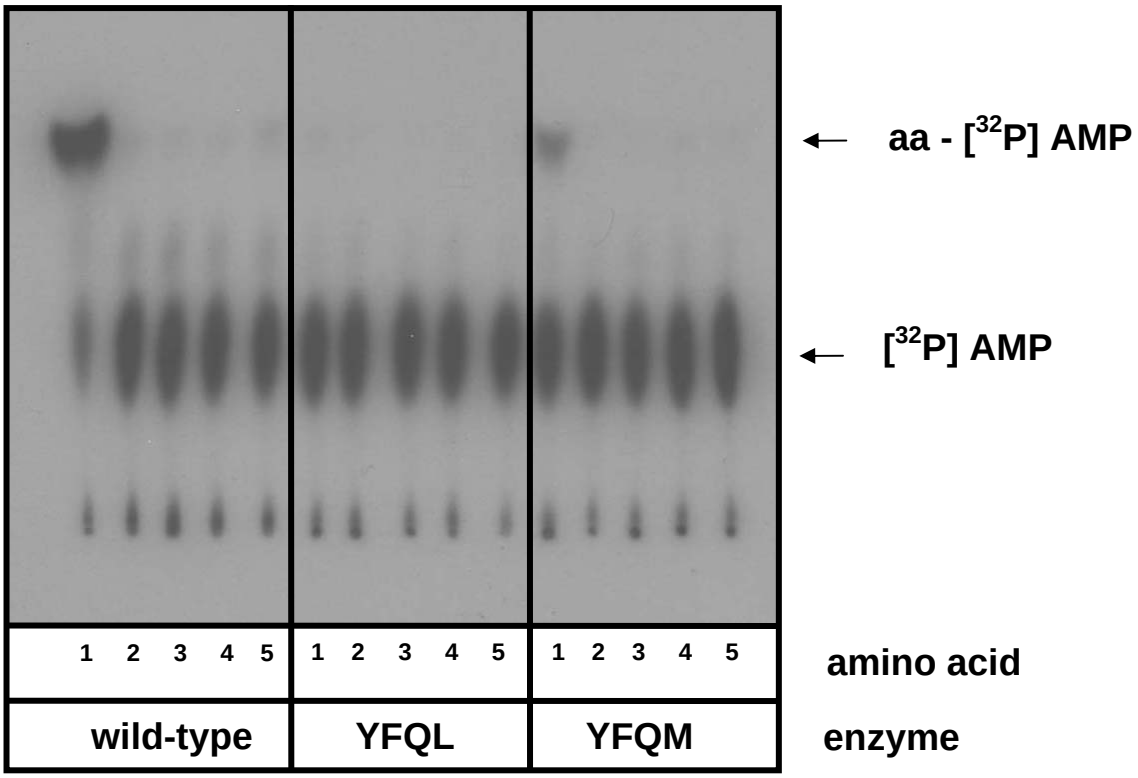


Figure 9. Activities of *E. coli* GlnRS wild-type and mutants Y211F Q255L (YFQL) and Y211F Q255M (YFQM) towards various amino acids. *In vitro* aminoacylation was performed using 10 mM of amino acid as indicated; 1, Glutamine; 2, Leucine; 3, Isoleucine; 4, Methionine; 5, Methionine sulfoxide. aa, amino acid.

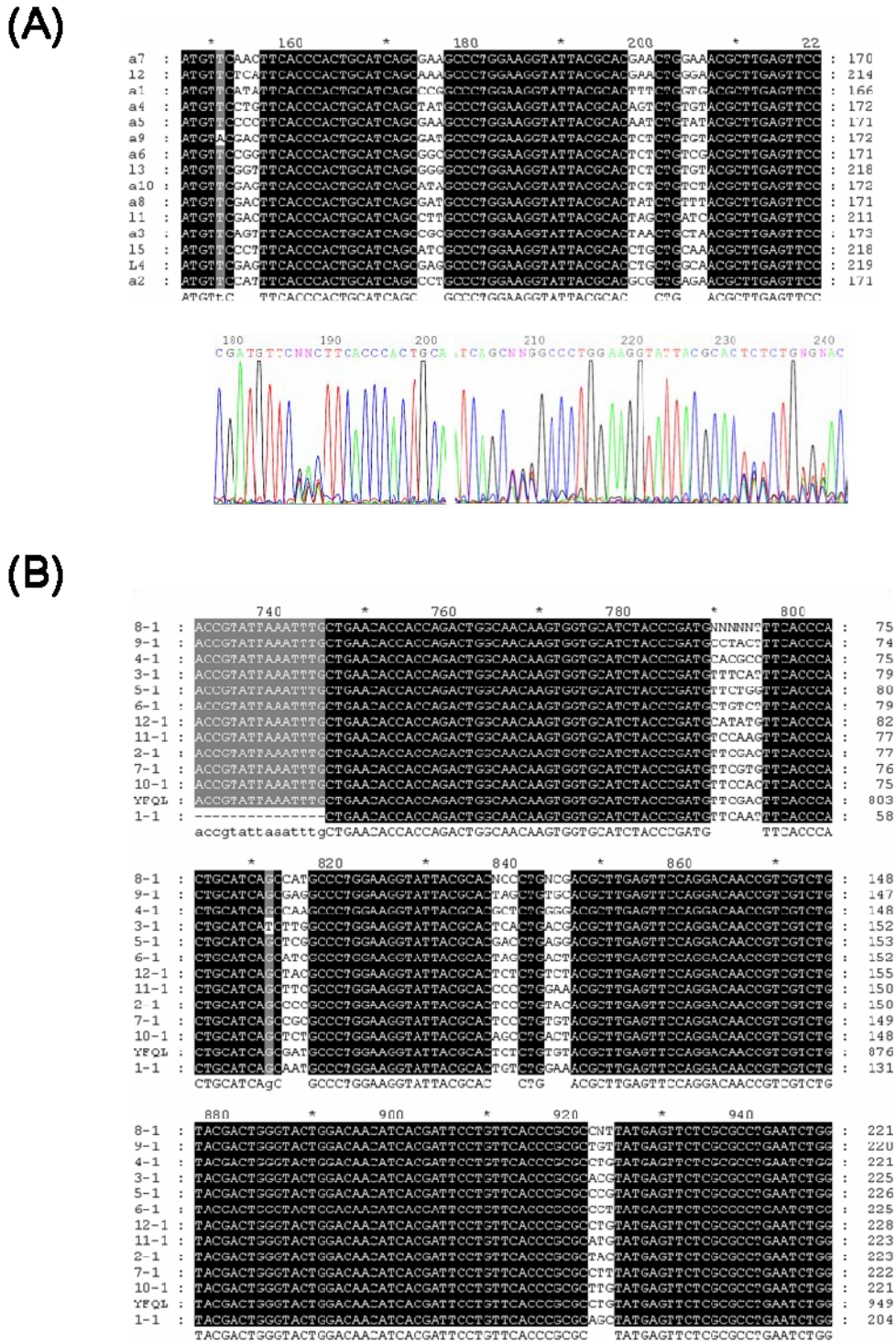


Figure 10. Random mutagenesis of *E. coli* GlnRS at specific sites. **(A)** Mutagenesis of D212, D219, S227, and C229 in a Y211F/Q255L background. **(B)** Mutagenesis of Y211, D212, D219, S227, C229, and Q255.

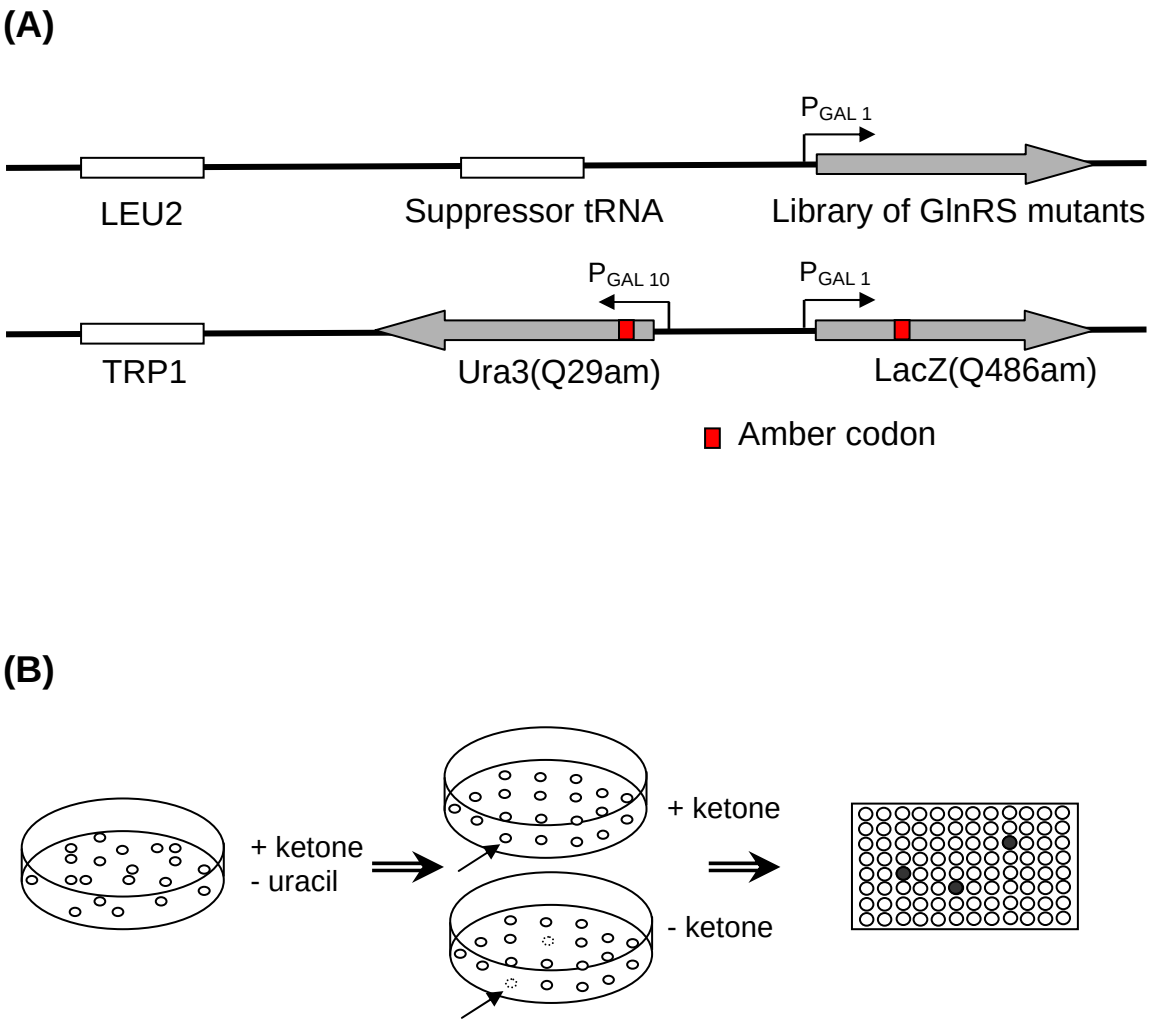
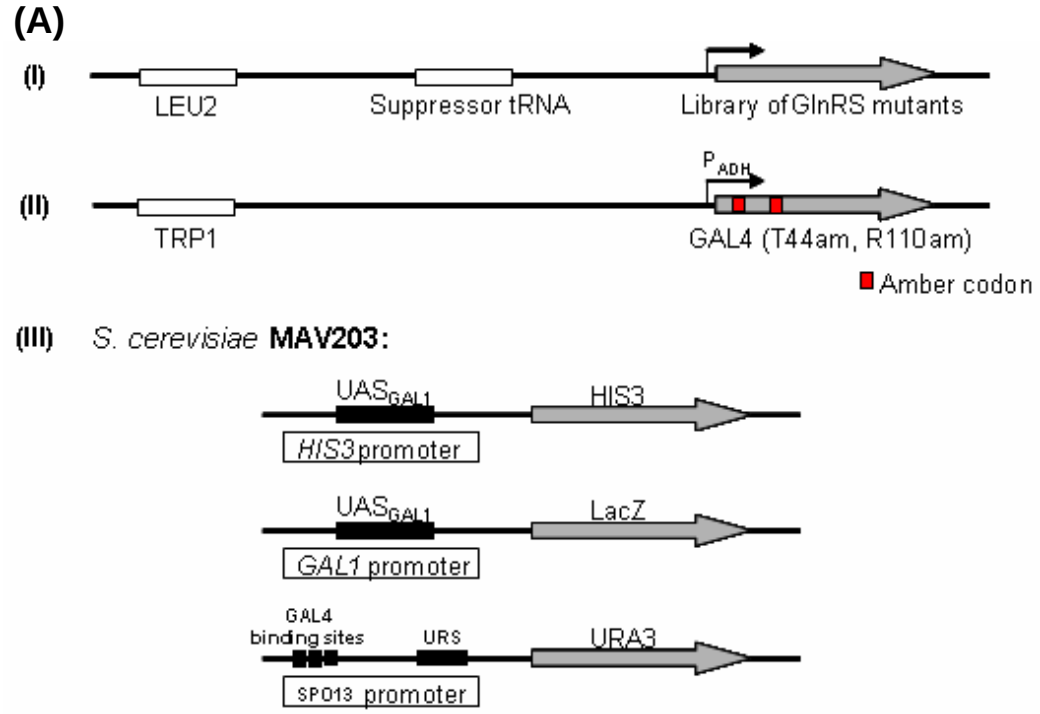


Figure 11. Positive and negative selections in *S. cerevisiae* for *E. coli* GlnRS mutants that use 2-amino 5-keto hexanoic acid. **(A)** Plasmids generated. **(B)** Overall scheme for selecting the desired mutant GlnRSs.



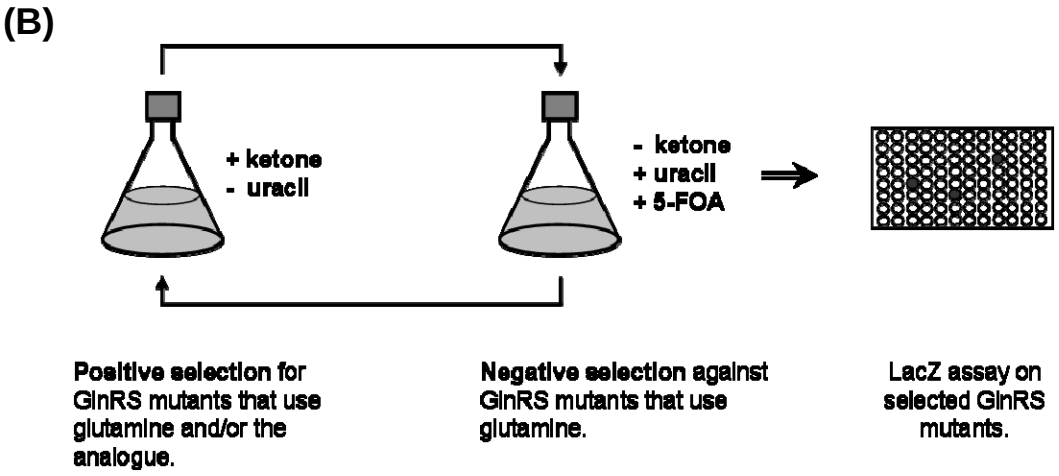


Figure 12. **(A)** Plasmids and *S. cerevisiae* strain for positive and negative selection of *E. coli* GlnRS mutants in yeast. The GlnRS gene and GAL4 (T44amber, R110amber) gene are expressed under the control of the constitutive ADH1 promoter. **(B)** Strategy for positive and negative selection of *E. coli* GlnRS mutants which utilize 2-amino 5-keto hexanoic acid instead of glutamine in yeast.

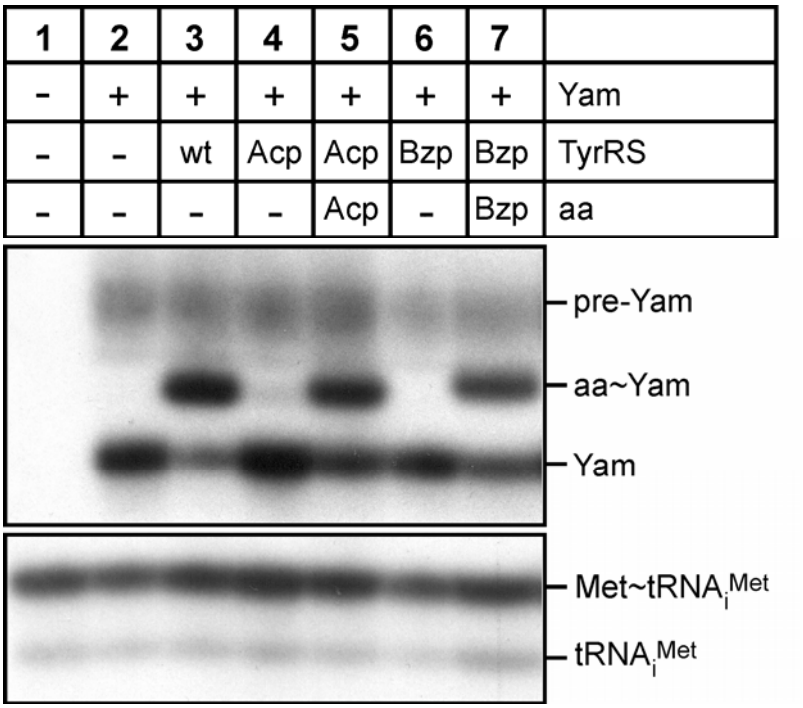


Figure 13. *In vivo* aminoacylation state of *Bst*-Yam suppressor tRNA in mammalian cells. Total RNA was isolated under acidic conditions from HEK293T cells transfected with plasmids carrying the genes for *Bst*-Yam and *E. coli* TyrRS (wild-type, AcpRS, or BzpRS). Cells were cultured in the absence and presence of Acp or Bzp as indicated.

tRNAs were analyzed by acid urea PAGE followed by Northern hybridization using a DNA oligonucleotide specific for *Bst*-Yam. An oligonucleotide specific for the human initiator tRNA-$_i^{Met}$ was used as internal standard for quantitation of RNA and aminoacylation levels by PhosphorImager analysis. aa, amino acid; pre-Yam, *Bst*-Yam precursor.

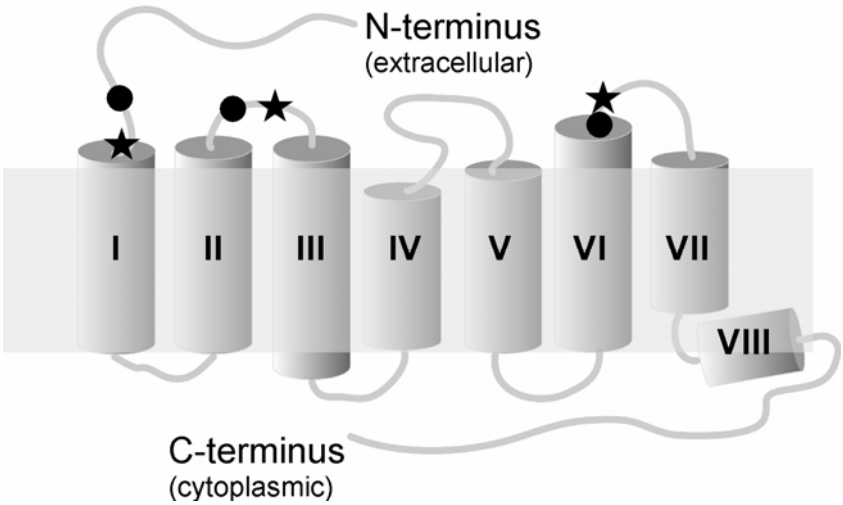
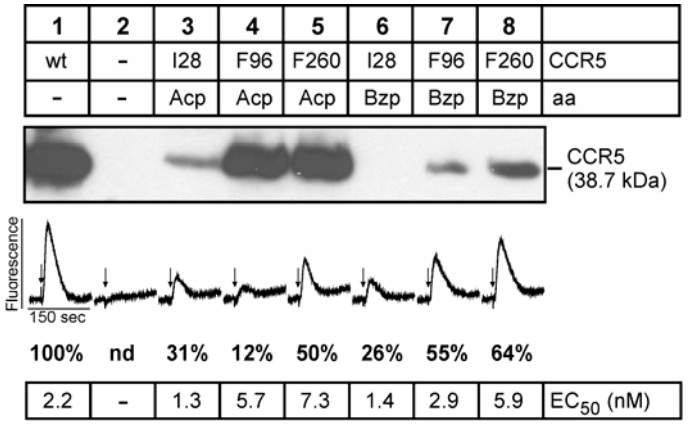


Figure 14. Schematic representation of the secondary structure of G protein-coupled receptors (GPCRs). The 7-TM helices (I – VII) and the cationic amphipathic helix VIII are shown as cylinders. Positions in CCR5 (★) and rhodopsin (●) subjected to site-specific incorporation of unnatural amino acids are indicated.

(A)



(B)

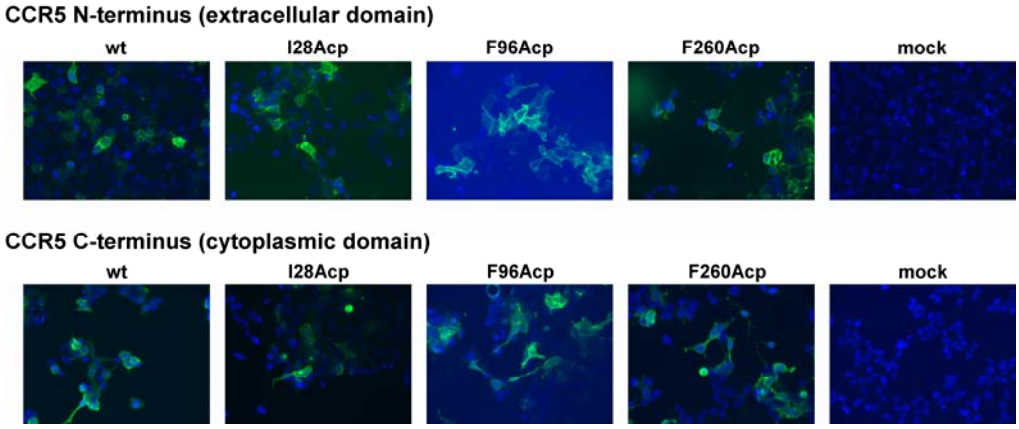
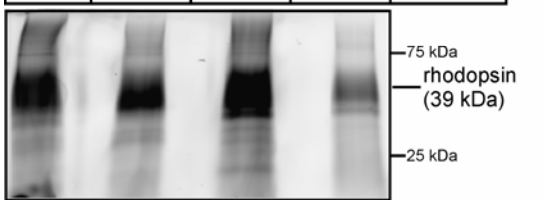


Figure 15. Expression of functional CCR5 mutants containing Acp or Bzp at positions 28, 96, or 260. HEK293T cells were transfected with plasmids carrying the genes for CCR5-wt or CCR5 mutant with an amber mutation at position I28, F96, or F260. Plasmids encoding *Bst*-Yam and *E. coli* TyrRS (AcpRS or BzpRS) were co-transfected, and the corresponding unnatural amino acids (Acp or Bzp) were provided in the cell media as indicated. **(A)** Cell lysates were analyzed by immunoblot using an antibody against the C-terminal 1D4 epitope of CCR5 (top) and RANTES-induced calcium flux (bottom). Equal amounts of total protein were analyzed per lane. Calcium flux traces represent the levels of Fluo-3 fluorescence emitted over time by cells in suspension. The addition of RANTES (20 nM) is indicated by arrows. Relative CCR5 activities and EC_{50} values for RANTES-induced calcium flux are summarized underneath. **(B)** Cell surface expression of CCR5-wt and CCR5 mutants containing Acp. Fluorescence images of non-permeabilized cells labeled with a FITC-conjugated anti-CCR5 antibody against the extracellular N-terminal 2D7 epitope of CCR5 (N; top panel), and permeabilized cells labeled with an antibody against the cytoplasmic C-terminal 1D4 tag (C; bottom panel) of CCR5 followed by staining with an Alexa 488-conjugated secondary antibody. Blue, nuclei stained with DAPI. Green, labeled CCR5.

(A)

1	2	3	4	
Y29	Y102	Y274	wt	rhodopsin
Acp	Acp	Acp	-	aa



(B)

1	2	3	4	
Y29Acp	Y102Acp	Y274Acp	wt	rhodopsin
1.41	0.63	1.59	0.28	F/R

(C)

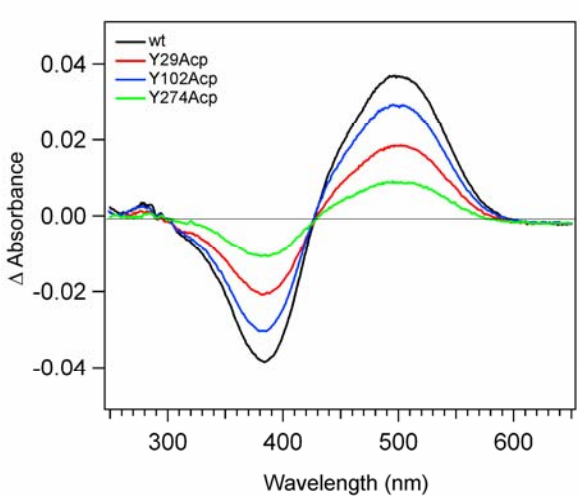
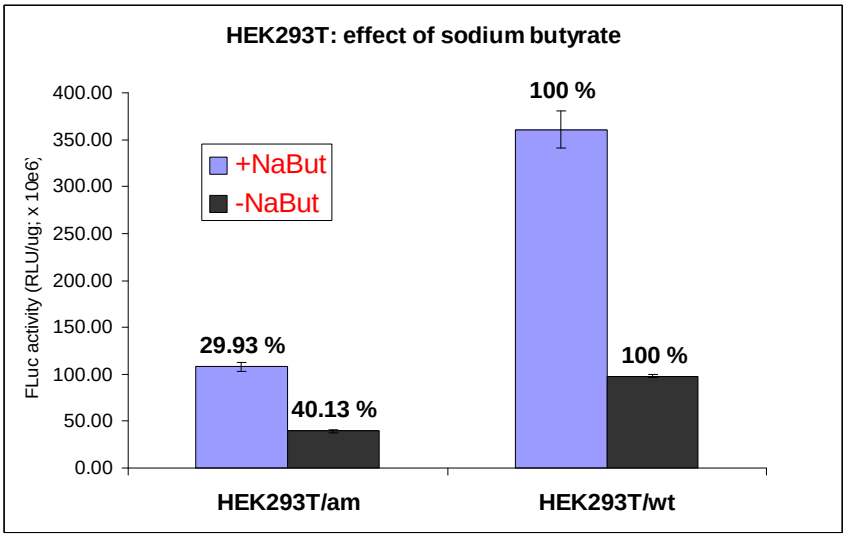


Figure 16. Site-specific incorporation of Acp into rhodopsin and *in vitro* labeling of rhodopsin mutants. **(A)** *In vitro* labeling of rhodopsin mutants (rho-Y29Acp, rho-Y102Acp, and rho-Y274Acp) with fluorescein hydrazide. Rhodopsin concentrations were confirmed by difference UV-vis spectra, and the loading volume was normalized to the same amount of rhodopsin based on the 500 nm absorbance. Samples were separated by 12% SDS-PAGE; the fluorescence image was taken with a Typhoon 9400 Image Scanner using an excitation/emission filter set optimized for fluorescein detection. **(B)** Stoichiometric ratios of fluorescein/rhodopsin (F/R) were determined after normalizing the amount of rhodopsin based on the absorbance at 500 nm. **(C)** Difference spectra of fluorescein-labeled rohodopsin mutants were generated by subtracting the photobleached spectra from the dark spectra. wild-type (black); rho-Y29Acp (red); rho-Y102Acp (blue); rho-Y274Acp (green).

(A)



(B)

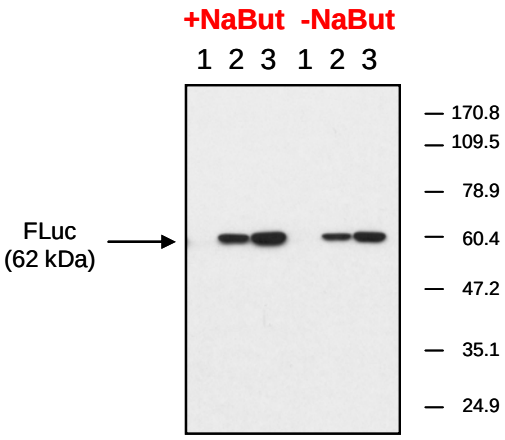
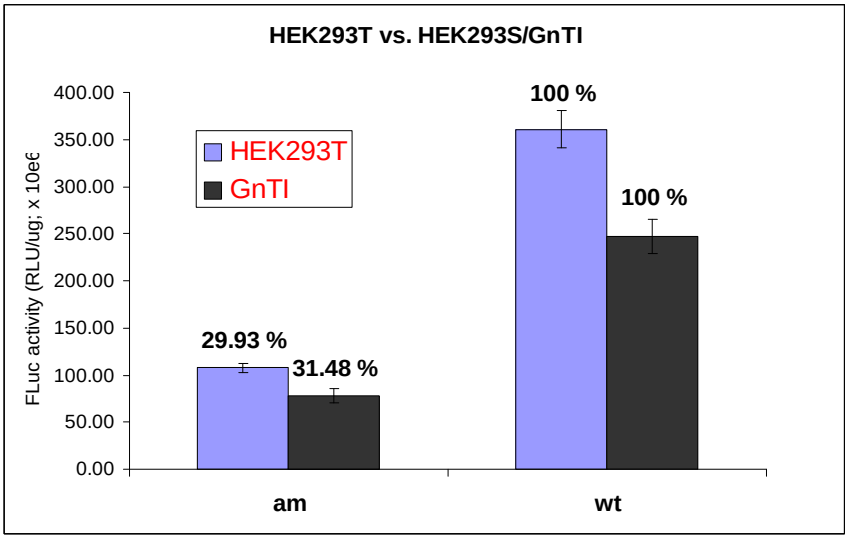


Figure 17. Effect of sodium butyrate. HEK293T cells were transfected with plasmid DNA (pTNT.FLuc.Y70am or wt; pSVB.Yam; pCDNA.TyrRS.wt). 3 hours post-transfection, fresh medium supplemented with serum, glutamine and sodium butyrate (1 mM final concentration) was added. Cells were harvested 48 h post-transfection and assayed for FLuc activity **(A)**. 10 ug of total protein were separated by 10% SDS-PAGE and analyzed by immunoblotting using a His4-antibody to detect FLuc protein; lane 1, mock; lane 2, FLuc.Yam; 3, FLuc.wt; molecular weight markers are shown on the right **(B)**.

(A)



(B)

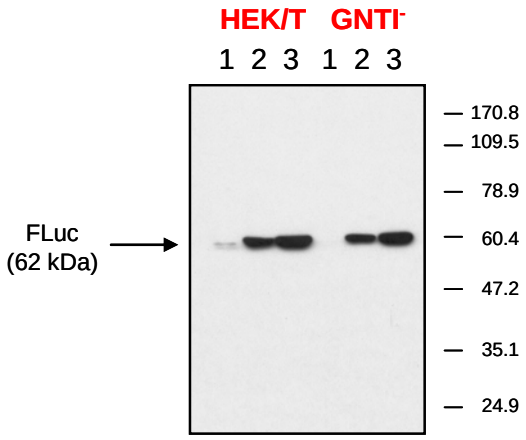


Figure 18. Nonsense suppression in HEK293S *GnTI* cells. HEK293T and HEK293S *GnTI* cells were transfected with plasmid DNA (pTNT.FLuc.Y70am or wt; pSVB.Yam; pCDNA.TyrRS.wt). Cells were harvested 48 h post-transfection and assayed for FLuc activity **(A)**. 10 ug of total protein were separated by 10% SDS-PAGE and analyzed by immunoblotting using a His4-antibody to detect FLuc protein; lane 1, mock; lane 2; FLuc.am; lane 3, FLuc.wt; molecular weight markers are shown on the right **(B)**.

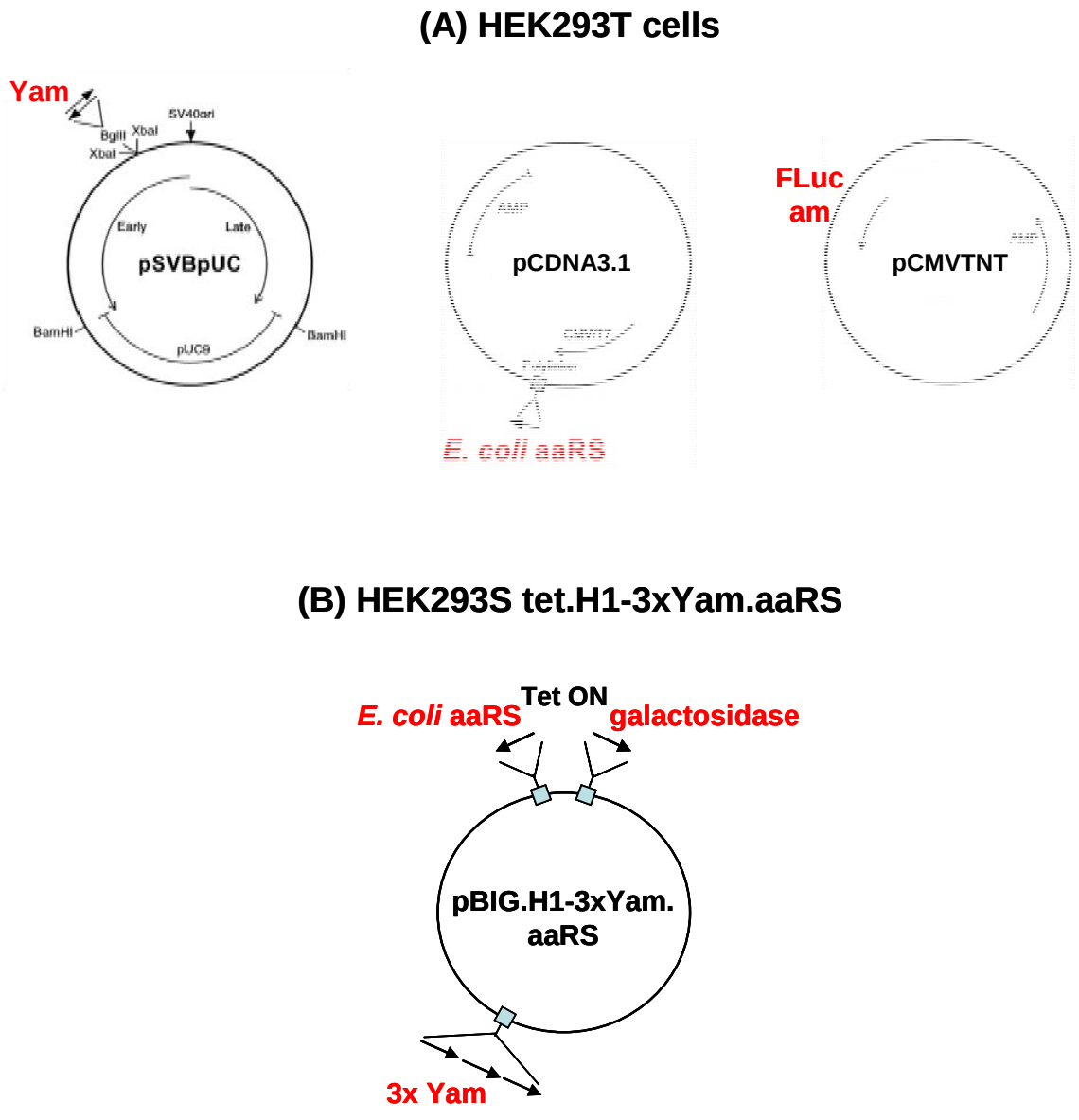


Figure 19. Plasmids for transient transfection into mammalian cells and generation of stable mammalian cell lines for site specific incorporation of unnatural amino acids. **(A)** Three-plasmid system for transient transfection of HEK293T cells: pSVBpUC carrying the tRNA gene (Yam), pCDNA3.1 carrying various aminoacyl-tRNA synthetase genes (*E. coli* aaRS), and pCMVTNT carrying the reporter gene (FLuc). **(B)** Plasmid for generation of stable cell lines. pBIG.H1-3xYam.aaRS carries three copies of the tRNA gene under the control of the H1 polymerase III promoter (H1-3xYam) and various aminoacyl-tRNA synthetases (TyrRS, AcpRS, AzpRS, BzpRS) under the control of a tetracycline inducible CMV promoter. Galactosidase is used for detection of positive clones during the selection of stable cell lines. The cell line used is the "HEK293S tet" cell line, generated in the Khorana laboratory for tetracycline-inducible high-level expression of opsin mutants (Reeves *et al.*, 2002).

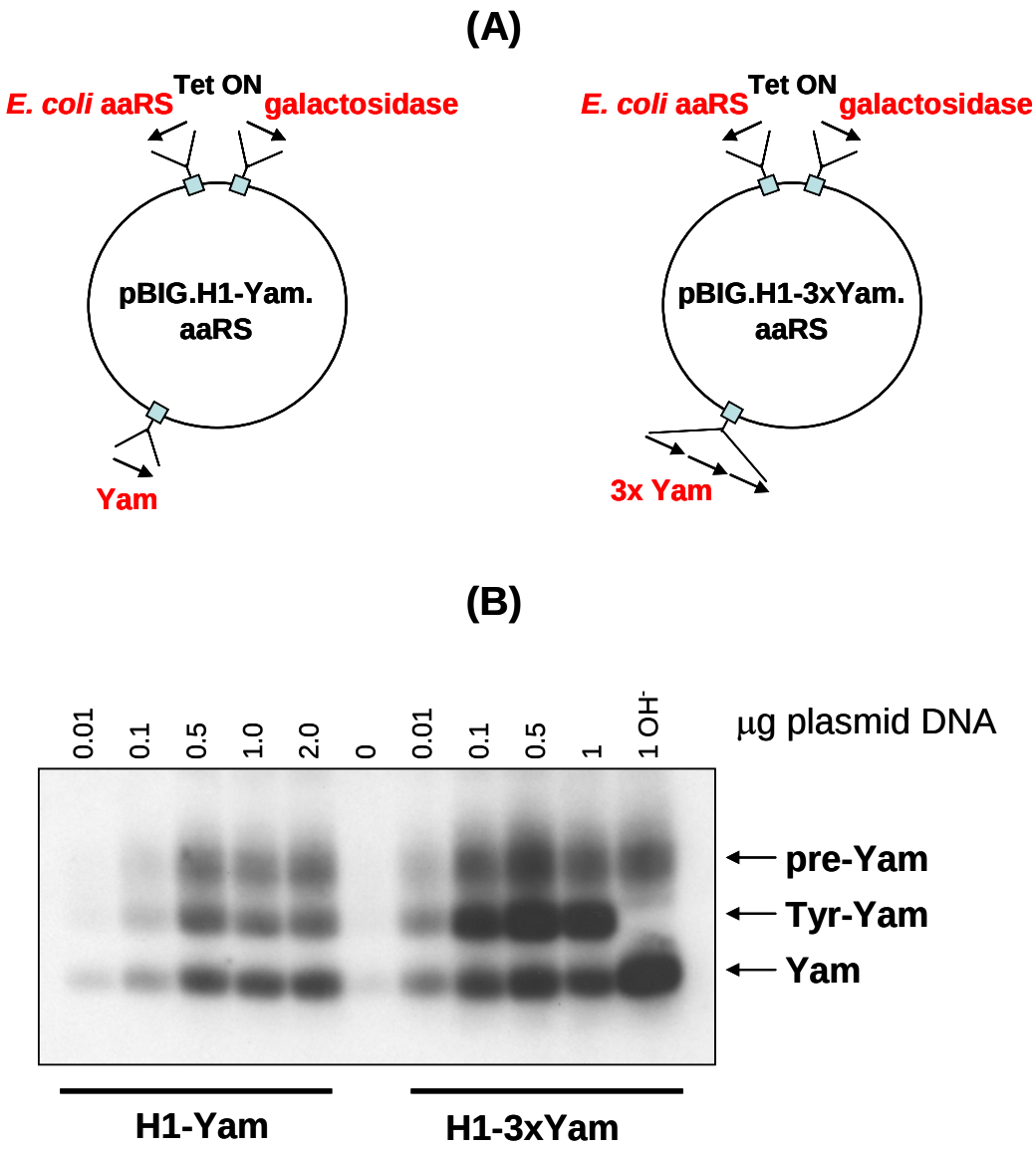


Figure 20. Effect of copy number on tRNA expression. HEK293S tet cells were transfected in a 24-well plate with plasmids pBIG.H1-Yam.TyrRS or pBIG.H1-3xYam.TyrRS **(A)** and pTNT.FLuc.Y70am (0.5 µg each). 3 hours post-transfection, fresh medium supplemented with serum, glutamine, sodium butyrate and tetracycline (2.5 µg/ml final concentration) was added. Cells were harvested 48 h post-transfection and assayed for tRNA Yam expression using Northern blot analysis **(B)**.

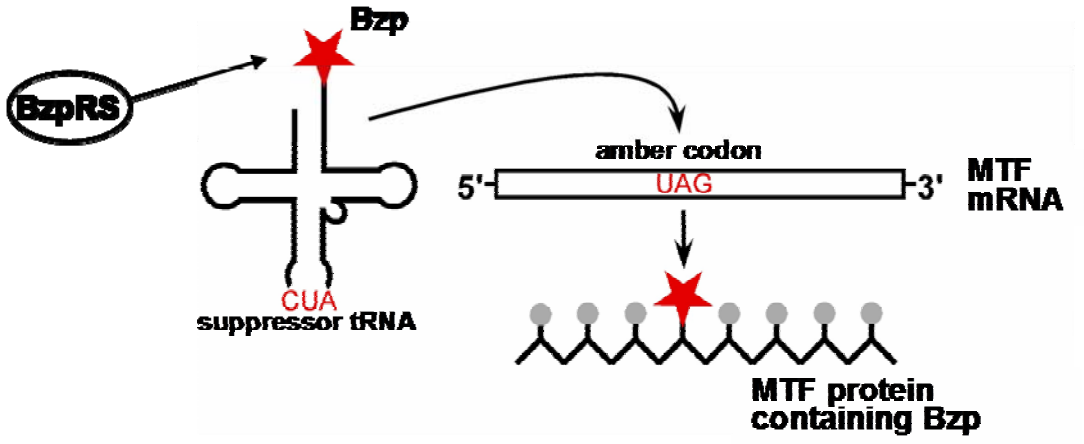


Figure 21. Site-specific incorporation *in vivo* of benzoyl phenylalanine (Bzp) into MTF using a specialized suppressor tRNA/aminoacyl-tRNA synthetase (BzpRS) pair.

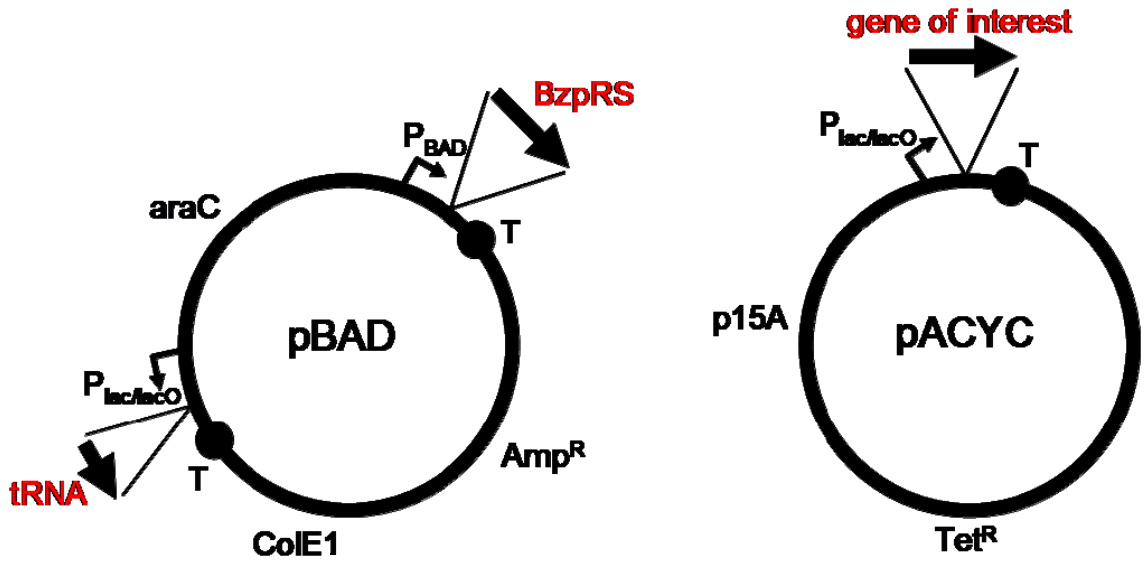


Figure 22. Two-plasmid system for site-specific incorporation *in vivo* of benzoyl phenylalanine (Bzp) into proteins in *E. coli*. The first plasmid, a pBAD derivative, carries the specialized amber suppressor tRNA/BzpRS pair; the second plasmid, a pACYC184 derivative, carries the protein gene of interest (MTF) with amber mutations at various positions.

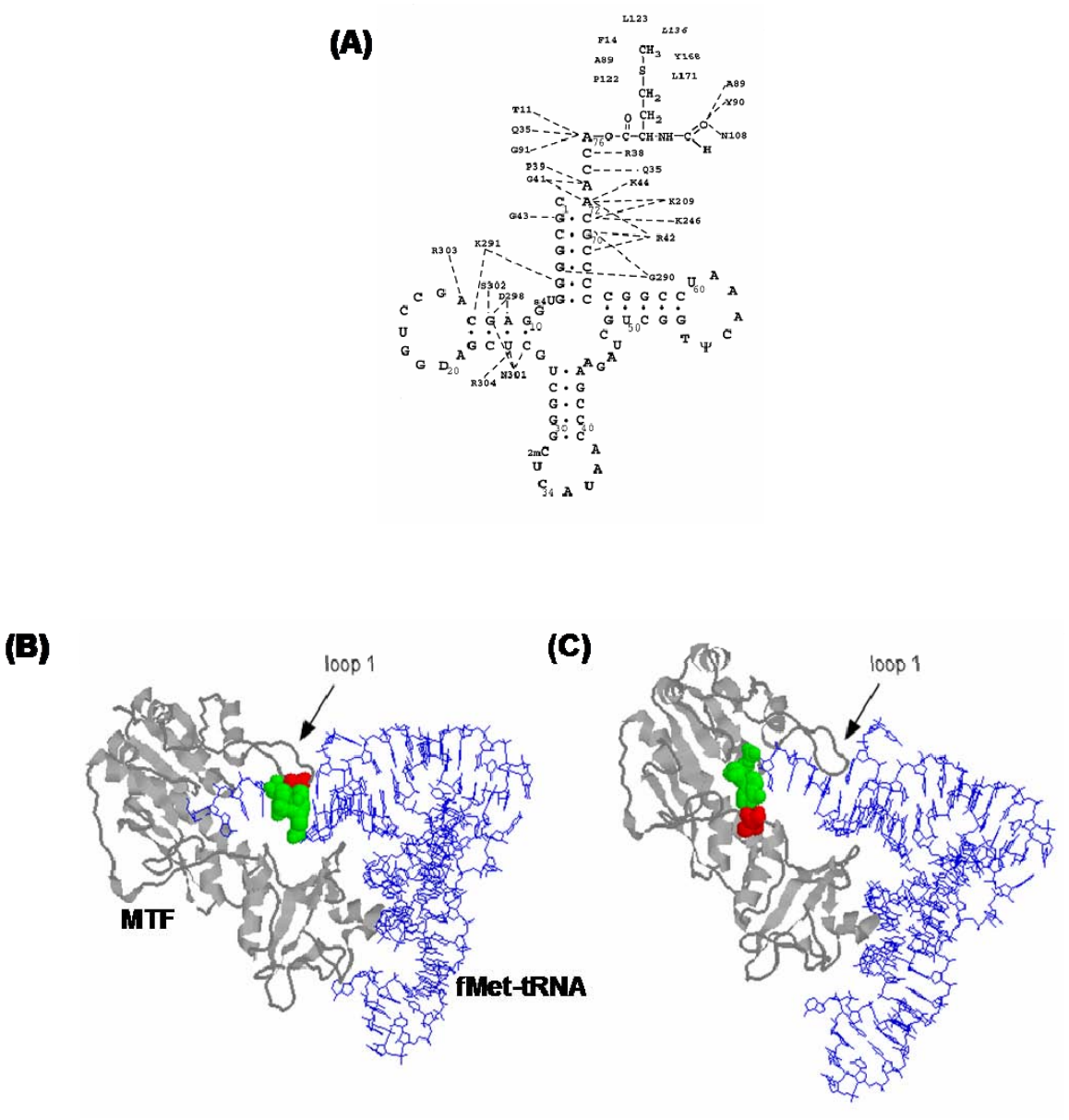
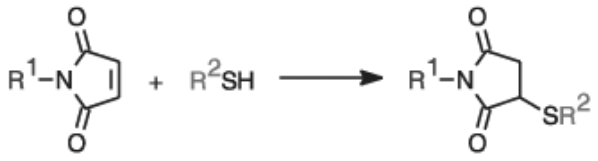


Figure 23. Functional interactions between *E. coli* MTF and fMet-tRNA.

(A) Main points of contact between MTF and fMet-tRNA [taken from Schmitt *et al.*,1998].

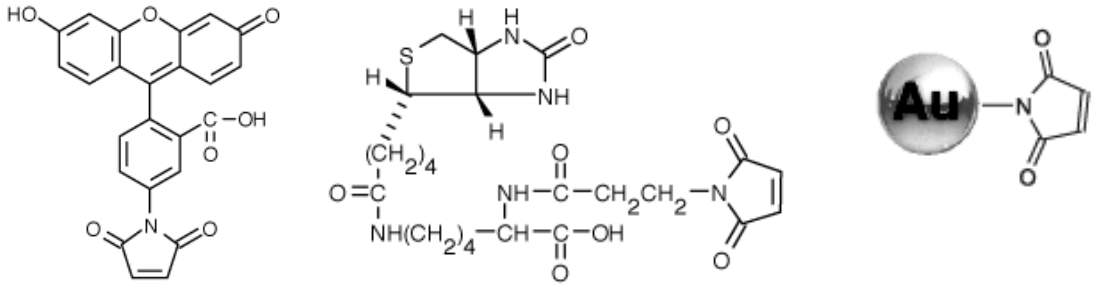
(B) (C) Proximity of Gly41 to the A72A73 bases **(B)** and Lys206 to the 3'end of tRNA **(C)**. X-ray structure of *E.coli* MTF.fMet-tRNA complex drawn using RASMOL version 2.6 with coordinates obtained from the Protein Data Bank (PDB code 2FMT). MTF is colored gray, except Gly41 **(B)** and Lys206 **(C)** (red spacefill); the tRNA is in blue wireframe, except for positions A72A73 (**B**) and A76 **(C)** (green spacefill) [taken from Mayer *et al.*, 2002].

(A)



Maleimide Thioether

(B)



Fluorescein Biotin Nanogold

Figure 24. (A) Reaction of maleimide with a thiol group. **(B)** Some examples of maleimide derivatives that can be used for reaction with cysteine attached to tRNA.

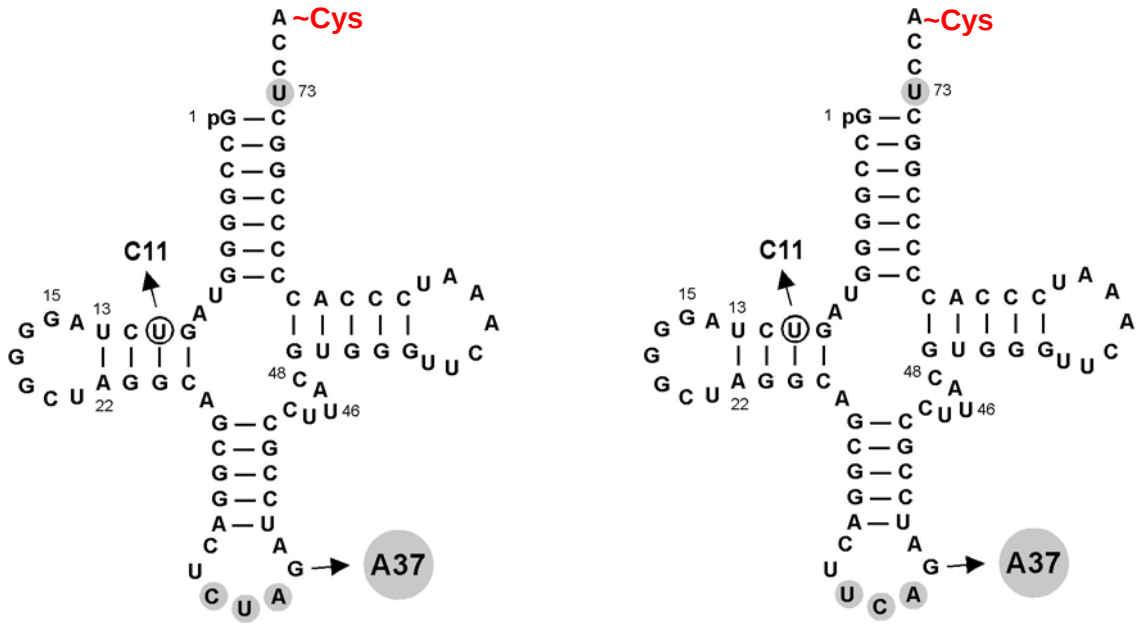
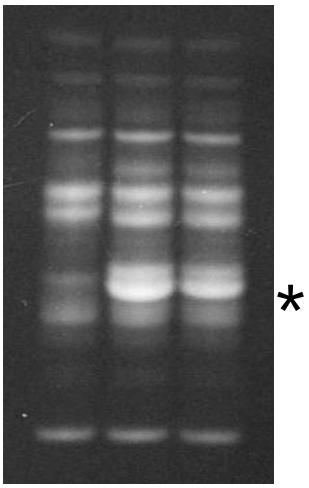
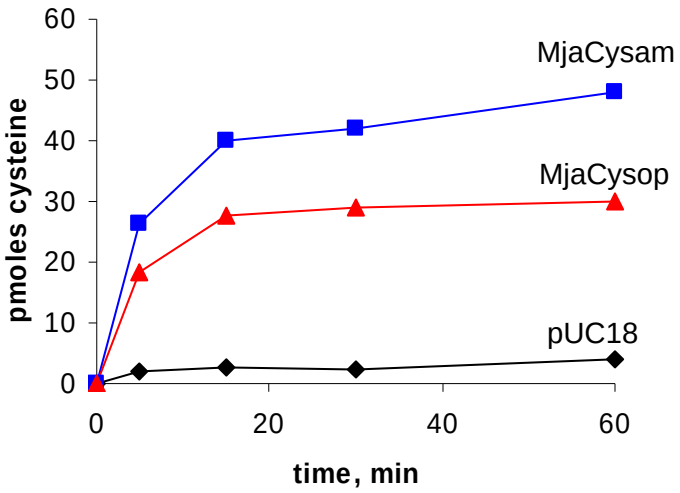


Figure 25. Cloverleaf structures of amber and opal suppressor tRNAs derived from *M. jannaschii* $tRNA^{Cys}$ (MjaCysam/op). The mutated anticodon sequences, the U11 to C11 and the G37 to A37 mutations are indicated.

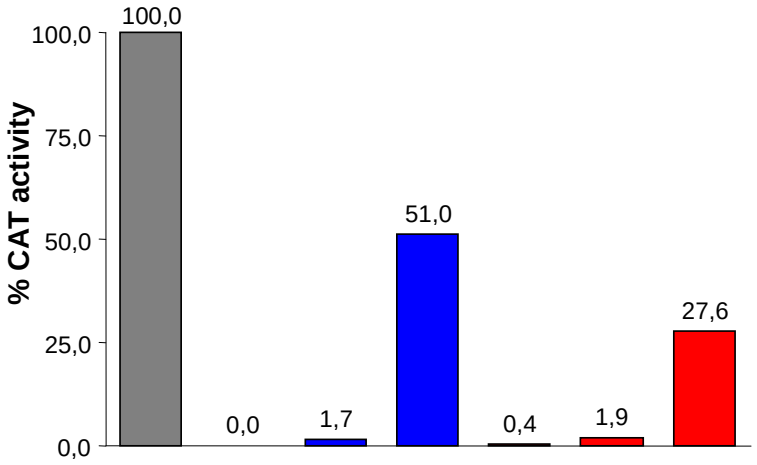
(A)



(B)



(C)



CAT	wt	am	am	am	op	op	op
MjaCys tRNA	-	-	am	am	-	op	op
MmaCysRS	-	-	-	+	-	-	+

Figure 26. (A) Overexpression of MjaCysam/op tRNAs in *E. coli*. Lane 1, total *E. coli* tRNA; lane 2, MjaCysam; lane 3, MjaCysop. **(B)** Aminoacylation of MjaCysam/op tRNAs with *M. maripaludis* CysRS (160 pmoles of tRNA in each sample). **(C)** Suppression of amber and opal stop codons in the CAT gene by MjaCysam/op tRNAs in *E. coli*.

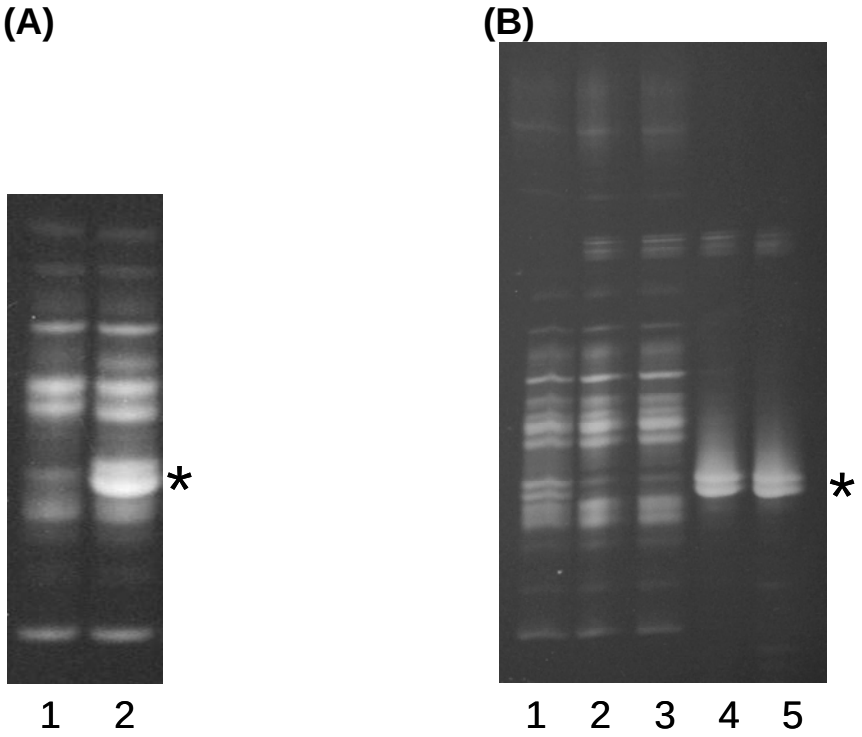


Figure 27. A) Overexpression of MjaCysam tRNA in *E. coli*. Lane 1, total *E. coli* tRNA; lane 2, MjaCysam. **(B)** Affinity purification of MjaCysam using a biotinylated DNA oligonucleotide coupled to streptavidin-agarose resin. Lane 1, starting material (total tRNA from *E. coli* overexpressing MjaCysam tRNA); lane 2, unbound tRNAs; lane 3, wash; lane 4, elution 1; lane 5, elution 2.

Native polyacrylamide gels stained with ethidium bromide are shown. Position of MjaCysam tRNA is indicated with an asterisk.

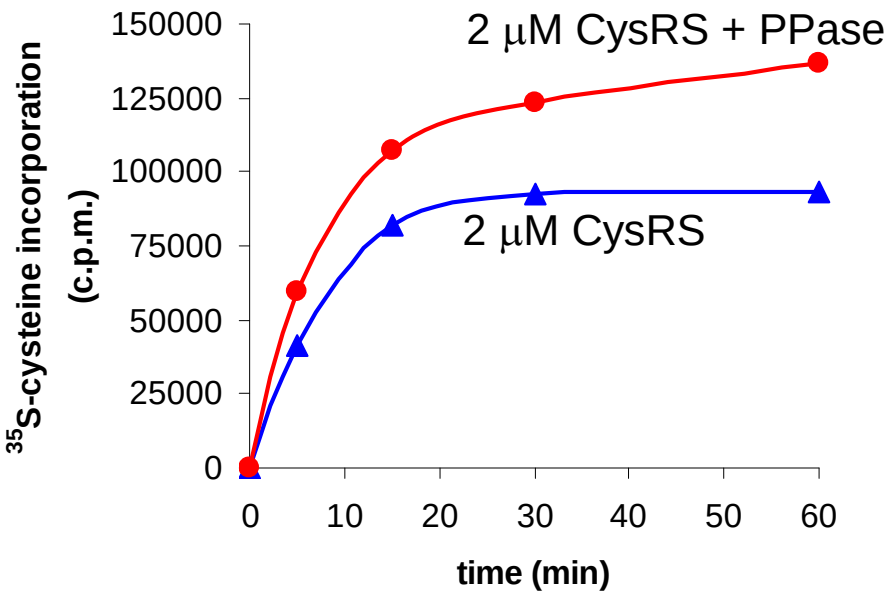
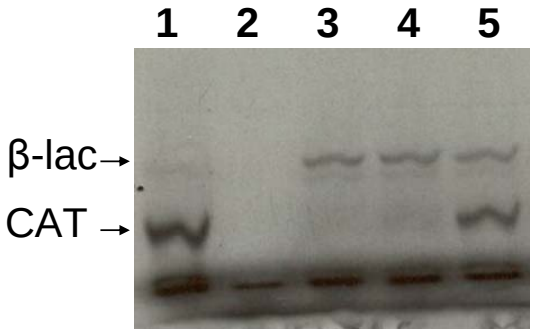


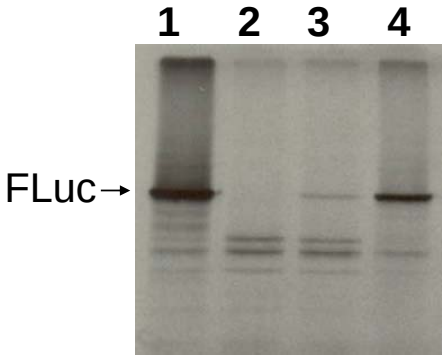
Figure 28. *In vitro* aminoacylation of MjaCysam tRNA. Suppressor tRNA was aminoacylated with cysteine using *M. maripaludis* CysRS (80 pmoles of tRNA in each sample; final concentration 1.6 μM) in the absence and presence of inorganic pyrophosphatase (PPase) as indicated.

(A)



CAT	wt	-	am	am	am
MjaCys				+	+
CysRS					+

(B)



FLuc	wt	am	am	am
MjaCys			+	+
CysRS				+

Figure 29. MjaCysam tRNA is active in suppression in *in vitro* translation systems. **(A)** *E. coli* S30 extract. **(B)** Rabbit reticulocyte extract.

CAT, chloramphenicol acetyltransferase; β-lac, β-lactamase; FLuc, firefly luciferase; CysRS, *M. maripaludis* CysRS

Samples were analyzed by SDS-polyacrylamide gelelectrophoresis followed by autoradiography. 80 pmoles of tRNA (final concentration: 3.2 μM) in each sample as indicated. CysRS was added to a final concentration of 4 μM.