Site Specific Non-native Amino Acid Replacement in Bacteriorhodopsin for Biomolecular Electronics

TABLE OF CONTENTS

Chapter 1: Protein Engineering is Limited by Twenty Naturally occurring Amino Acids

1.1 General Introduction to Protein Engineering and Organization of the Thesis

Protein engineering involves the creation of new or modified proteins, making it a powerful tool for generating novel biomaterials with precisely "tailor-made" properties. Today, protein engineering along with genetic engineering is impacting life on earth in areas ranging from agriculture, food, healthcare, environment and materials. Understanding biological structure and function at virtually any level requires an understanding of proteins, biopolymers of remarkable diversity composed of only twenty coded amino acids. Advances in structure determination have contributed greatly to this understanding. The field has dramatically enhanced our ability to probe the basis of structure and function in protein biochemistry and has enabled us to create proteins with entirely novel properties.

Table 1.1: Strategies for Protein Engineering

Methods for Creation/Isolation of proteins	**Methods for Modification of proteins**
Protein Isolation and Purification technologies	Chemical Modifications of Proteins
Recombinant DNA technology for heterologous expression	Crosslinking and immobilizations of proteins
Expression of chimeric proteins	Random mutagenesis and Selection of mutant proteins
	Site Directed mutagenesis
	Non-specific incorporation of unnatural amino acids
	Site specific incorporation of unnatural amino acids

One of the main goal of the present work is to provide a new array of tools to biochemists and cell biologists and thus giving access to large quantities of proteins with novel physical and chemical properties. These methods will not only allow us to engineer proteins with interesting new properties for a wide variety of applications such as biomolecular electronics, they will allow us to analyze the function of these mutant proteins in living cells.

The scientific work presented in this thesis focuses on site-specific incorporation of non-native amino acids and attempts to make it a more general technology for protein engineering. Recent advances in *in vitro* incorporation of non-native amino acids have expanded the scope of protein engineering and allow larger proteins in higher yields. Such modern methods of *in vitro* protein production allow us to study macromolecules with changes as small as a single atom or as significant as the addition of multiple biophysical probes. This exquisite ability to control protein composition is entering a new era with methods for the site-selective insertion of non-native amino acids into proteins in living cells, effectively expanding the genetic code.

It is important to understand that creation of such novel biomaterials would mainly involve intervening the naturally occurring protein biosynthesis process. In the last decade, the *in vitro* biosynthetic incorporation of unnatural amino acids has become especially robust for generating proteins of virtually any size, and novel insertion sequences like four-base codons and unnatural codons have proven useful (*1*). Methods for equipping cells with the essential biomolecules for genetic code expansion- "orthogonal tRNAs" and aminoacyl tRNA synthetases (aaRSs)- are reviewed in the subsequent sections.

The scientific work performed as a part of this thesis is presented in following chapters. In the current chapter, general concepts of protein synthesis are presented and issues involved in nature's restriction to 20 naturally occurring amino acids are discussed. This discussion thus, sets stage Chapter 2 where reasons, rationale and significance of expanding genetic code for site-specific incorporation of non-native amino acids in proteins are discussed.

Chapter 3 presents a literature review of the model protein used in this study, bacteriorhodopsin (bR), with special emphasis on the unsolved questions in its structure-function relationship. Chapter 4 establishes overall goal of this work in terms of specific aims, scope and study design.

Chapter 5-10 detail experimental work, results and discussion of the present study. Chapter 5 shows expression and purification of bR and its mutants used in the subsequent study. Chapter 6 describes how bacterioopsin (the apo-protein of bR) mRNA was generated and encapsulated into liposomes for *in vivo* delivery. Chapter 7 describes how misaminoacylated suppressor tRNA was generated for liposome encapsulations and delivered *in vivo*. Chapter 8 shows how *in vivo* deliveries of macromolecules have helped us elucidate the retinal biosynthetic pathway in *Halobacterium salinarium*.

Chapter 9 shows how the core goal of the present thesis was reached i.e. *in vivo* SNAAR into bR. Chapter 10 shows biophysical studies on SNAAR analogs of bR reveal a critical structural change during its function.

The entire approach towards this work is to intervene at the cellular protein synthesis machinery level. Hence, it is important to develop a complete understanding of the normal protein synthesis within a cell, especially at the translation level, to manipulate the non-native amino acids into the cellular protein synthesis machinery and eventually into the proteins. The following section, first of all, reviews the naturally occurring protein synthesis process.

1.2 Protein Synthesis

The central dogma defines the paradigm of molecular biology: genes are perpetuated as sequences of nucleic acid, but function by being expressed in the form of proteins. Three types of processes are responsible for the inheritance of genetic information and for it

conversion from one form to another. The central dogma of the cell involves three processes.

a.) DNA→DNA (replication),

b.) DNA→RNA (Transcription) and

c.) RNA→Protein (Translation).

Information is perpetuated by **replication**; a double stranded nucleic acid is duplicated to give identical copies.

Information is expressed by a two stage process. **Transcription** generates a single stranded RNA identical in sequence with one of the strands of the duplex DNA. Several different types of RNA are generated by transcription; the three principal classes involved in the synthesis of proteins are: **messenger RNA (mRNA), transfer RNA (tRNA)** and **ribosomal RNA (rRNA).** The process of transcription synthesizes a messenger RNA molecule containing the message encoded on the DNA. Translation converts the nucleotide sequence of RNA into the sequence of amino acids comprising of protein. An mRNA is translated into a protein sequence; tRNA and rRNA provide other components of the apparatus of protein synthesis. The entire length of an mRNA is not translated, but each mRNA contains atleast one coding region that is related to a protein sequence by the genetic code: each nucleotide triplet (codon) of the coding region represents one amino acid. Generally, most of the functional requirements of proteins have been provided by nature as the standard twenty amino acids. The following section of this chapter briefly summarizes the step involved in protein synthesis.

1.2.1. Transcription

Transcription is the process through which a DNA sequence is enzymatically copied (by an RNA polymerase) to produce a complementary RNA (*2*, *3*). In other words, it is the transfer of genetic information from DNA into RNA. In the case of protein-encoding DNA, transcription is the beginning of the process that ultimately leads to the translation of the genetic code (*via* the mRNA intermediate) into a functional peptide or

protein. The stretch of DNA that is transcribed into an RNA molecule is called *transcription unit*.

As in DNA replication, transcription proceeds in the 5' → 3' direction (i.e. the old polymer is read in the 3' → 5' direction and the new, complementary fragments are generated in the 5' → 3' direction). Transcription is divided into 3 stages: *initiation*, *elongation* and *termination*. DNA-dependent RNA polymerase, which catalyzes the polymerization of RNA from ribonucleotides in a DNA-dependent manner, is the enzyme responsible for the transcription of DNA into RNA. The RNA polymerase binds to a specific DNA sequence called the promoter, and unwinds the duplex DNA for about one turn of the helix to expose a short stretch of single stranded DNA so that complimentary base pairing can be made with the incoming ribonucleotides. The enzyme joins two of the ribonucleotide triphosphate monomers and then moves along the DNA strand, extending the growing RNA chain in the 5'→3' direction until it encounters a second special sequence called the terminator, which signals where the RNA synthesis should stop. After the transcription is completed, each RNA chain is released from the DNA template as a free, single stranded RNA molecule. In addition to RNA polymerase, other protein factors are required for efficient transcription, and they are responsible for determining the transcription efficiency of specific transcription units.

Transcription has some proofreading mechanisms, but they are fewer and less effective than the controls for DNA; therefore, transcription has a lower copying fidelity than DNA replication. Before the synthesis of a protein begins, the corresponding RNA molecule is produced by RNA transcription. One strand of the DNA double helix is used as a template by the RNA polymerase to synthesize a messenger RNA (mRNA). This mRNA migrates from the nucleus to the cytoplasm. During this step, mRNA goes through different types of maturation including one called splicing when the non-coding sequences are eliminated. The coding mRNA sequence can be described as a unit of three nucleotides called a codon.

1.2.2. Translation

Translation is the stepwise synthesis of a polypeptide with an amino acid sequence determined by the nucleotide sequence of the mRNA coding region. The genetic code relates each amino acid to a group of three consecutive nucleotides called as codons. Decoding of the mRNA takes place from the 5′→3′ direction, and the polypeptide is synthesized from the amino to the carboxyl terminus. Decoding of the message on the mRNA takes place *via* by means of **transfer RNA** molecules, each specific for one amino acid and for a particular **triplet** of nucleotides in mRNA called a **codon**. The family of tRNA molecules enables the codons in a mRNA molecule to be **translated** into the sequence of amino acids in the protein. Figure 1.1 A, B and C is a schematic representation of the translation process.

Initiation: The initiation of translation starts by formation of the initiation complex between the mRNA and the ribosomes. Three proteins, initiation factors IF-1, IF-2 and IF-3 are required for the initiation of the protein synthesis. All three factors bind to the 30S ribosomal subunit near the 3′ end of the 16S ribosomal RNA at adjacent sites that are located at the interface between the small and the big subunits of the ribosomes. Messenger RNA binds to the small ribosomal subunit immediately before the formation of the final initiation complex. This binding involves a molecular recognition mechanism between the 3′ end of the ribosomal RNA and a complementary region, usually consisting of 3 to 9 bases in the 5′ end of the mRNA initiation codon. This sequence is called the Shine-Dalgarno sequence. The final initiation complex takes place after the large ribosomal subunit joins the preinitiation complex. This complex consists of the ribosome bound to mRNA and occupied with the f-Met-tRNA on the ribosome.

Elongation: In the initiation complex, location of the charged initiator tRNA in the P site of the ribosome allows the transfer of the methionine residue to the amino group of another aminoacyl tRNA in the A site. During this stage, complexes, composed of an amino acid linked to tRNA, sequentially bind to the appropriate codon in mRNA by forming complementary base pairs with the tRNA anticodon. The ribosome moves from codon to codon along the mRNA. The first peptide bond is formed when the aminoacyl tRNA in the ribosomal A site is converted into the corresponding methionyl aminoacyl tRNA by transfer of the methionyl residues from the charged initiator tRNA in the P

site. Efficient entry of aminoacyl tRNA into the ribosomal A site requires the participation of an elongation factor, termed EF-Tu in prokaryotes and eEF-1 in eukaryotes alongwith GTP. The appropriate elongation factor forms a ternary complex with GTP and all aminoacyl tRNAs except initiator tRNA, thus ensuring that appropriately charged tRNAs are efficiently bound in the A site. This aminoacyl tRNA

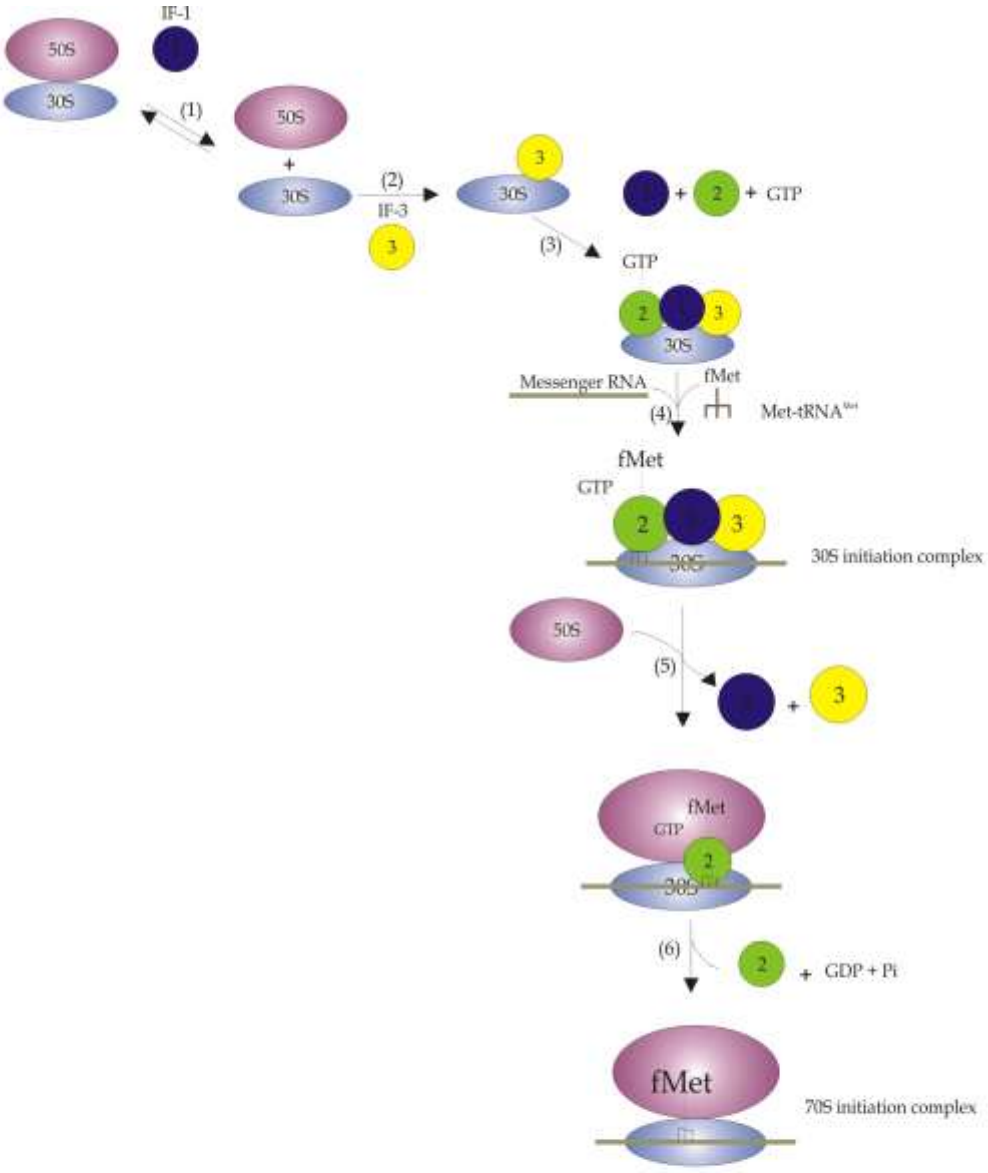


Figure 1.1. A Schematic diagram of translation in eukaryotes: Translation of mRNA into protein begins after assembly of initiator tRNA, mRNA and both ribosomal subunits. The complex initiation process that leads to 80S ribosome formation consists of several linked stages that are mediated by eukaryotic initiation factors. See text for details. The 40S ribosomal subunit is captured for initiation via complex arrays of protein-RNA and protein-protein interactions. In the cap-dependant mechanism, the pre-initiation complex binds to the mRNA at the 5' terminal cap structure with help of the eIF4F protein complex and then migrates along the mRNA until it encounters the initiation codon where the 80S ribosome is reconstituted. Upon release, the eIF are recycled

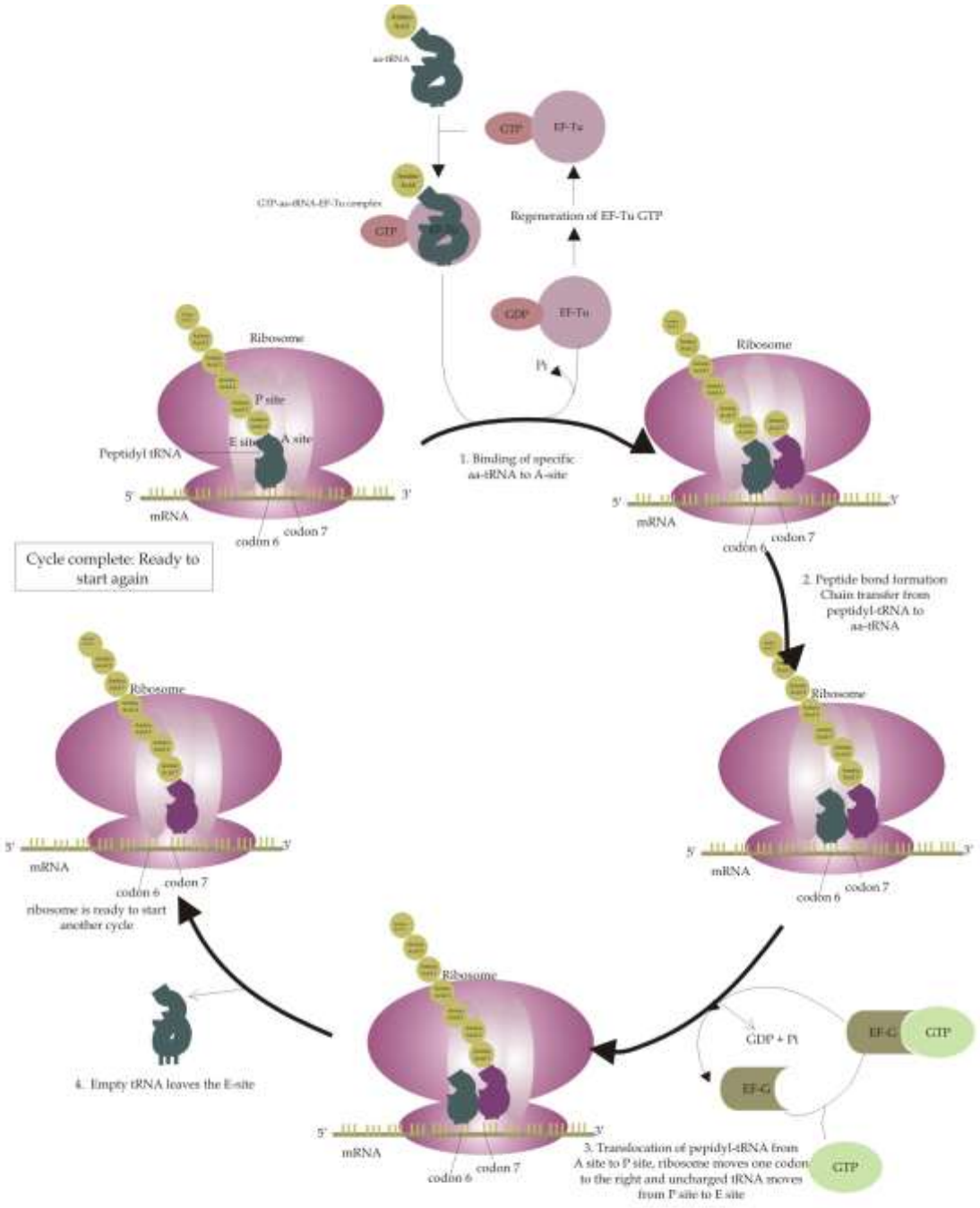


Figure 1.1B. Schematic diagram of translation in eukaryotes. The elongation process is depicted as a cycle. Following translocation (step 3) and empty tRNA release (step 4), the ribosome is ready to accept the next aminoacyl tRNA (aa-tRNA) and repeat the cycle. This cycle will continue until a termination codon is reached.

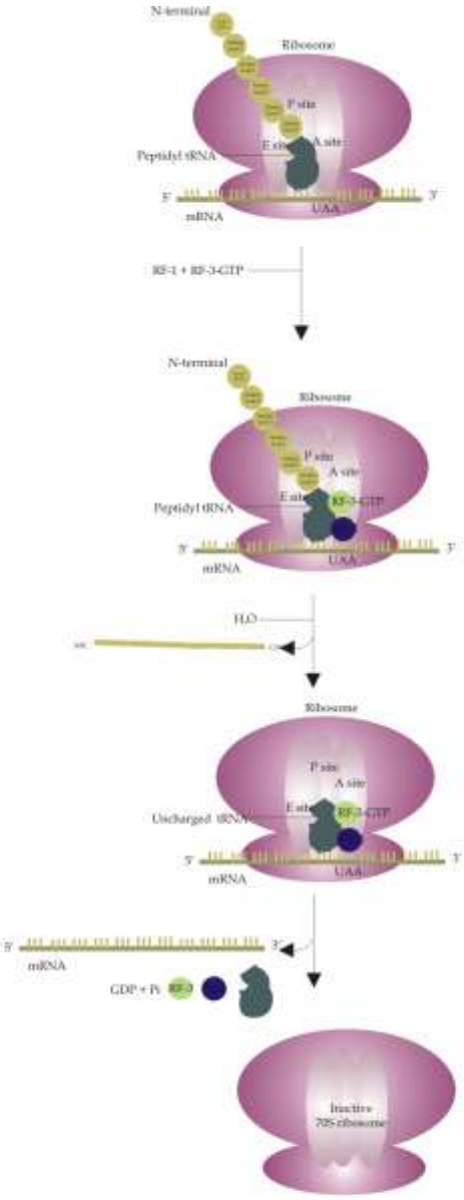


Figure 1.1 C Schematic diagram of translation in eukaryotes. The termination pathway in *E. coli* ribosomes: RF-1 recognizes the termination codons **UAA** and **UAG**, whereas RF-2 recognizes **UAA** and **UGA**.

binding reaction catalyzed by the EF-Tu is the rate-limiting step in the elongation cycle, peptide bond formation and translocation are much faster. The initial binding of the ternary complex to the ribosome is readily reversed, but the interaction is stabilized by the subsequent codon recognition, which induces the GTPase conformation of EF-Tu leading immediately to the hydrolysis of GTP component of the ternary complex to GDP. Hydrolysis of GTP moiety causes a further change in the conformation of EF-Tu from the GTP-binding form. This conformational change leads to the release of aminoacyl tRNA, allowing its CCA ends to align with the peptidyl transferase center of the ribosome and the instantaneous formation of the peptide bond. The elongation factor is later released from the ribosome as a complex with GDP.

Translocation: Translocation involves the movement of the ribosome along the mRNA in the 5′→3′ direction. Immediately after synthesis of the first peptide bond, the ribosomal a site contains dipeptidyl tRNA while the uncharged initiator tRNA remains in the P site. Thus, bothe these sites are occupied, and to allow the next aminoacyl tRNA to enter the A site, it is necessary to eject the uncharged tRNA and shift the dipeptidyl tRNA from the A into the P site. This translocation takes place as a concerted process involving movement of both mRNA and dipeptidyl RNA together into the P site, leaving the A site occupied by the next mRNA codon and free to accept the cognate aminoacyl tRNA. At the same time, the deacylated tRNA moves first into an E (exit) site with subsequent ejection when the next aminacyl tRNA enters the A site. Amino acids are added one by one, translated into polypeptidic sequences dictated by DNA and represented by mRNA.

Termination: The presence of one of the three termination codons, UAA, UAG, or UGA, in the A site results in the binding of a release factor instead of an aminoacyl tRNA to the ribosome. In prokaryotes, two release factors have been identified, one (RF1) recognizing UAA and UAG, the other (RF2), recognizing UGA. Ribosomal binding and release of RF1 and RF2 are stimulated by a third factor, RF3, which interacts with GTP and GDP. GTP hydrolysis appears to be required for the release factor from the ribosome. Thus, at the end of the ribosome cycle the coding sequence of messenger RNA has been translated to produce a particular polypeptide chain, and all components

involved become available for re-use in another round of cycle. Usually, several ribosomes, each at a different stage in completing the process of translation, are attached to one mRNA molecule, giving rise to polyribosomes.

1.2.3. Important Components of Protein Synthesis Machinery

1.2.3.1 Messenger RNA (mRNA)

The sequence information of a gene that is copied (transcribed) into the nucleotide sequence of RNA from the complementary strand of DNA is called the template strand. The primary transcript is a single strand of RNA, which is a faithful copy of the nontemplate strand of DNA with substitution of U residues in place of T residues found in DNA. Sometimes, the primary transcript is altered, as described below, before it functions as mRNA; in these cases, the original unmodified transcript is the precursor or pre-mRNA. Usually, mRNAs have nontranslated sequence at the 5′ and 3′ ends in addition to the coding domain. These noncoding sequences sometimes affect the efficiency of translation and the stability of mRNA. The structures of typical mRNAs are shown in Figure 1.2. The decoding process involves base pairing between three bases (designated a codon) in the mRNA and the three base anticodon of a transfer RNA (tRNA). In a separate reaction, each tRNA is first linked to a particular amino acid and thus the pairing of mRNA with tRNA determines the sequence of amino acids in the resulting protein.

Prokaryotic mRNA

In prokaryotes, pre-mRNA usually undergoes little or no modification so that pre-mRNA and mRNA are very similar, if not identical. Gene expression in prokaryotes usually involves the cotranscription of several adjacent genes, and translation of mRNA sequences into polypeptides may begin at the 5′ end of RNA while transcription is still in progress at the 3′ end.

Eukaryotic mRNA

In eukaryotes, the genetic information is stored mainly in the nucleus and to a minor degree in some organelles (mitochondria and chloroplasts). The description that follows

pertains only to nuclear genes. Eukaryotic genes are more complicatedthat prokaryotic genes because the coding region is ofter discontinuous: the coding sequence or exons are

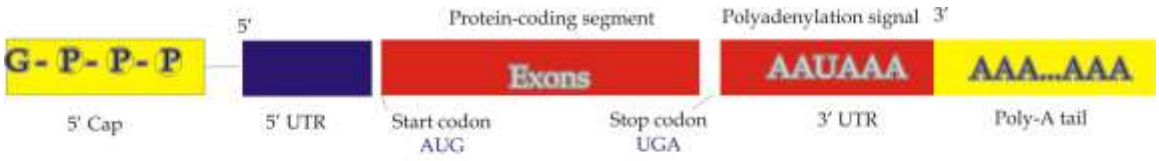


Figure 1.2: Simplified schematic illustration of mRNA structure

interrupted by intervening sequences (introns). Thus, genes and proteins are usually not collinear in eukaryotes. In the nucleus, a complicated set of splicing reactions removes all the introns from the pre-mRNA and fuses the exons into a continuous coding sequence. Other processing steps involve adding a "cap" to the 5′ end of the mrNA and a poly(A) "tail" to the 3′ end. After completion of these nuclear maturation steps, the mRNA is transported to the cytoplasm where it is translated. As is the case with prokaryotic mRNA, the coding region is flanked by 5′ and 3′ nontranslated sequences.

1.2.3.2 Transfer RNA (tRNA)

The tRNA molecules play a very crucial role in the synthesis of proteins. Structurally, it consists of a single strand of 77 ribonucleotides with a common 3′ terminal CCA sequence. Most of the bases are standard but some (e.g. pseudo, dihydroU, and T) are derived by modification after transcription of the transfer RNA genes. The secondary structure of tRNA is usually presented in two dimensions as a cloverleaf to highlight the regions of base pairing (See Figure 1.3 A and B). The CCA sequence carrying the amino acid is located distal to the anticodon.

The decoding process involves antiparallel base pairing between the three bases of mRNA codons and the complementary anticodons of tRNA during peptide bond formation. The first and middle bases of the codon form conventional base pairs with the third and middle base of the anticodon, respectively, but the third base of the codon pairs with the first base of the anticodon by a less stringent interaction, giving rise to degeneracy. This so-called "wobble" considerably reduces the number of tRNA species required to decode the 61 sense codons. Thus, the protein synthesis in the cytosol of eukaryotes contains only a few more than 40 different tRNAs, and in mitochondria, 22 to 24 species are sufficient.

The attachment of amino acid to tRNA involves the formation of an ester bond between the α–carboxyl group of the amino acid and the 3′ hydroxy group of the terminal

adenosine of the tRNA. It requires specific enzymes, the aminoacyl tRNA synthetases (AARSs). There are 20 different synthetases, each specific for one of the 20 amino acids,

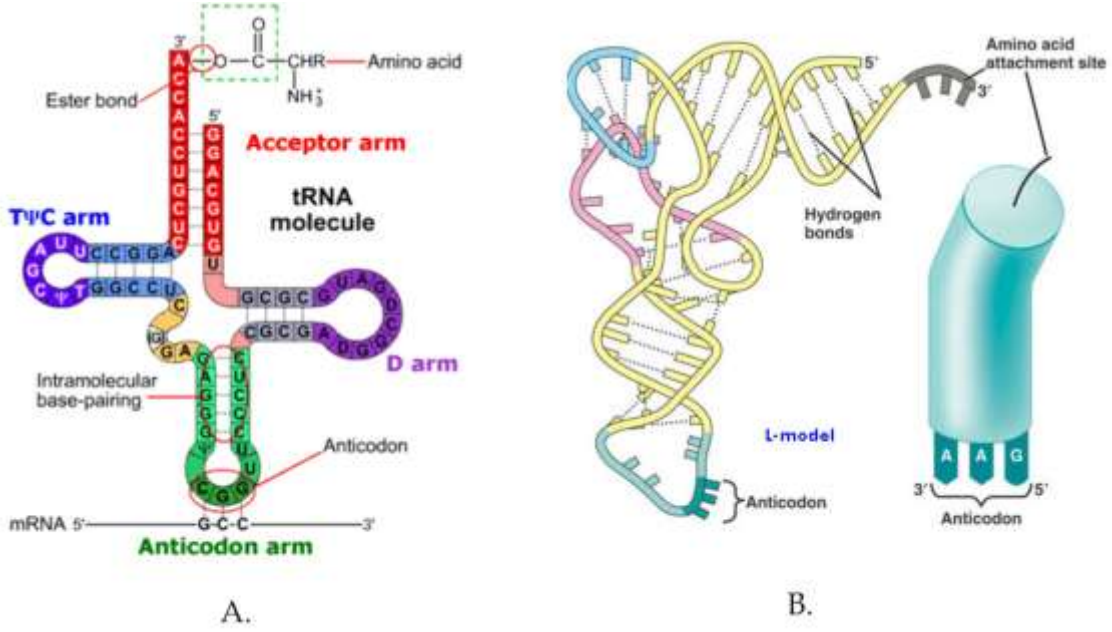


Figure 1.3: Simplified schematic illustration of tRNA structure. Panel A illustrates the two dimensional structure; Panel B illustrates the base pairing and hydrogen bonding of various domains within the tRNA molecule.

and each enzyme recognizes something unique in the structure of its cognate tRNA. The structural determinants that ensure accuracy of this charging reaction vary for different tRNAs. The anticodon may play a part but sometimes even a single base elsewhere is sufficient to determine the specificity of the tRNA-synthetase interaction. The accuracy of the synthetase reaction in attaching an amino acid to its cognate tRNA is critically important to the fidelity of the translation process. **Once the aminoacyl tRNA has been formed, the subsequent incorporation of the amino acid residue into a polypeptide does not depend upon the amino acid itself but only on the interaction between the anticodon of the aminoacyl tRNA and the codon of the mRNA.** Thus, an error in the synthetase reaction would lead to the incorporation of an inappropriate amino acid into the polypeptide. Synthesis of aminoacyl tRNA from amino acids requires activation of the amino acid carboxyl group with formation of an intermediate enzyme-bound aminoacyladenylate. The energy for the reaction (two high energy phosphate bonds) is provided by ATP and stored in the ester bond of the aminoacyl tRNA to be used subsequently for peptide bond synthesis.

A detailed review on tRNA and its interaction with its aminoacyl tRNA synthetase is presented in chapter 7.

tRNA nomenclature: The amino acid linked to a charged tRNA is indicated by a prefix and the specificity of the tRNA in the aminoacylation reaction is shown as a superscript on the right; for example, Phe-$tRNA^{Phe}$ indicates phenylalanine specific tRNA charged with phenylalanine. The anticodon maybe indicated as a right subscript, or alternatively, in the superscript after the amino acid, for example, $tRNA_{UGC}$ or $tRNA^{Ala}_{UGC}$

1.2.3.3 Ribosomes

Ribosomes are high molecular weight complexes of RNA and proteins, and the electron dense particles are easily visualized within cells by electron microscopy (Figure 1.4). Usually, several ribosomes become attached to one mRNA molecule, giving rise to polyribosomes (also called polysomes). Ribosomes from various sources (prokaryotes,

eukaryotic cytoplasm, mitochondria, chloroplasts and kinetoplasts) vary in size from 20 to 30 nm in diameter, but all are composed of a large and small subparticle or subunit

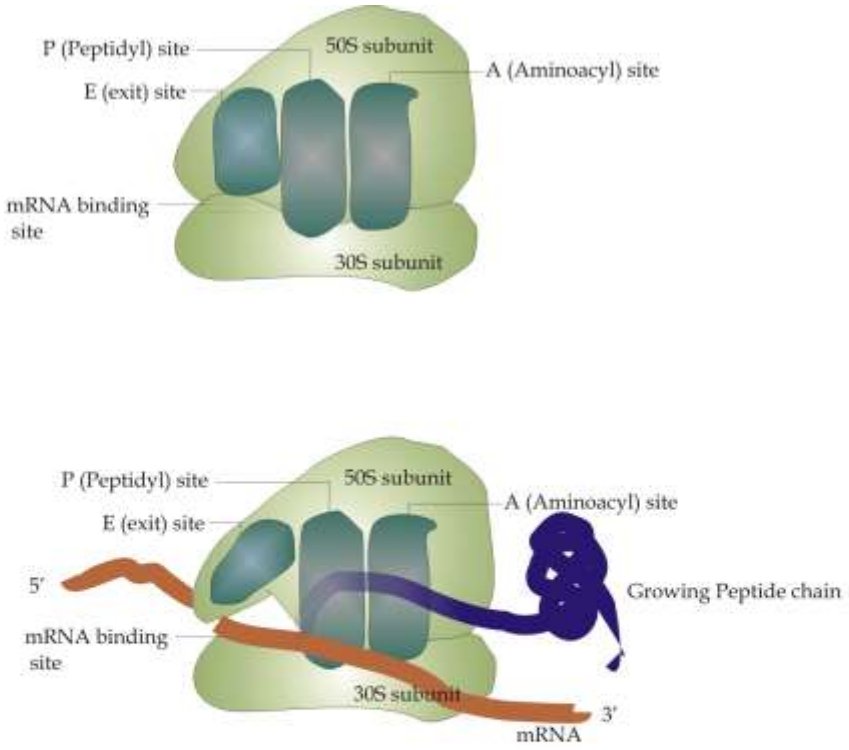


Figure 1.4: Simplified schematic illustration of structure of ribosomes. The model of ribosome structure shows the A (aminoacyl) and P (peptidyl) sites as cavities on the ribosome where charged tRNA (carrying an amino acid) molecules bind during polypeptide synthesis. The recently postulated E (exit) site is the site from which discharged tRNAs leave the ribosome. The mRNA-binding site binds a sequence near the 5′ end of the mRNA, placing the mRNA in the proper position for the translation of its first codon. The binding sites are all located at or near the interface between the large and small subunits.

and perform similar functions in protein synthesis. In general, cytoplasmic ribosomes of eukaryotes are larger than ribosomes of prokaryotes. The principle functional domains of the ribosome and associated components are given in figure 1.4. More detailed structure of the ribosome has allowed the placement of mRNA, aminoacyl tRNA, and the nascent polypeptide chain.

The small subunit comprises a single rRNA of 0.3-0.7 $x10^6$ Da and single copies of 20 to 30 unique proteins. It has a major function in binding the initiator tRNA and mRNA in the initiation of protein synthesis and decoding the genetic message.

The large subunit comprises a high molecular weight rRNA (0.6-1.7 x 10^6 Da) and often two smaller rRNAs (0.03-0.05 x 10^6 Da) and 30 to 50 different proteins are present, with one exception, as single copies. The large subunit binds the aminoacyl-tRNA at the A site, peptidyl tRNA at the P site and discharged tRNA at the E (exit) site. The large subunit contains the peptidyl transferase center, and unusually, this enzyme activity resides in the rRNA molecule itself rather than in the associated ribosomal proteins. This subunit is also involved in binding elongation factor G, which is required for translocation.

1.2.4. The Genetic Code

The genetic code is triplet, comma-less, and nonoverlapping. As a consequence, a nucleotide sequence has three possible reading frames. Because mRNA is normally translated into a unique polypeptide, an essential step in the translation process is the selection of the appropriate reading frame. This is achieved by starting translation at the initiation codon AUG or less frequently GUG, which ensures that the following codons are read in phase within the required reading frame.

Of the 64 theoretically possible triplets in the genetic code, 61 sense codons correspond to 20 genetically encoded amino acids found in all, or nearly all, proteins (Refer Table 1.2). When GUG is used as the initiation codon, it codes for methionine by interaction with the anticodon of the initiator Met-$tRNA^{met}$, whereare elsewhere it codes for valine.

All other codons specify only one amino acid but many amino acids are specified by two or more (upto six) codons; the code is unambiguous, but degenerate. The remaining three codons, termed nonsense codons, usually signify termination of synthesis and release of the finished polypeptide chain.

1.2.4.1 Deviations from the Standard Genetic code

One of the nonsense codons, UGA has an additional function in the synthesis of selenoproteins. The process involves the initial synthesis of selenocysteinyl-tRNA from a novel seryl-tRNAand selenium. This tRNA contains an anticodon that is able to decode UGA and insert selenocysteine residues into the growing polypeptide chain, but only at UGA codons in a particular context of neighbouring nucleotides. The insertion of selenocysteine residues into the polypeptide also requires a specific elongation factor T, which differs from the factor used for the incorporation of other aminoacyl tRNAs.

Mitochondria and chloroplasts, as well as certain organisms such as mycoplasma and ciliated protozoa, use a few non-standard genetic codes. For example, methionine is usually coded for by AUG (or occasionally GUG), but in human mitochondria this codon is replaced by AUA. Variation in the genetic code are thought to have arisen as a results of loss of some tRNA genes and mutantional pressure on DNA, giving rise to predominance of either AT- or GC- rich codons.

The non-universality of the genetic code was first observed in eukaryotic mitochondria and since then numerous examples have been described. For example, UGA can not only act both as terminator and a selenocysteine codon, depending on its context, but it can also serve as a tryptophan codon in mitochondria and *Mycoplasma capriocolum* and a cysteine codon in *Euplotes octocarinatus*. The other terminator codons, UAA and UAG, also have additional functions, for example in *Tetrahymena thermophila* where they are both able to encode glutamine. The extent and significance of these and other deviations from the genetic code have been reviewed by Osawa et al. who conclude that such deviations are both a result of and contribute to biodiversity. Variable translation of some codons has been observed during the production of recombinant proteins in *E.coli*

due to imbalances in the amino acid pools (*4*). The diversity of certain elements within the genetic code suggests it may be possible to accommodate additional amino acids. The most widespread example of this has been the amber suppression of the termination codon UAG, which is discussed in the following section. However, the incorporation of non-proteinogenic amino acids can most easily be achieved without any genetic manipulation simply by supplying cells with amino acid analogues, which mimic their natural counterparts during protein synthesis. However such incorporation would not be site-specific. With the exception of selenocysteine and formyl-methionine, none of these genetic code modifications result in the incorporation of amino acids outside the canonical twenty.

Table 1.2 shows the 64 codons and the corresponding coded amino acid

		2nd base			
		U	**C**	**A**	**G**
1st base	**U**	UUU (Phe/F) UUC (Phe/F) UUA (Leu/L) UUG (Leu/L)	UCU (Ser/S) UCC (Ser/S) UCA (Ser/S) UCG (Ser/S)	UAU (Tyr/Y) UAC (Tyr/Y) UAA Ochre (*Stop*) UAG Amber (*Stop*)	UGU (Cys/C) UGC (Cys/C) UGA Opal (*Stop*) UGG (Trp/W)
	C	CUU (Leu/L) CUC (Leu/L) CUA (Leu/L) CUG (Leu/L)	CCU (Pro/P) CCC (Pro/P) CCA (Pro/P) CCG (Pro/P)	CAU (His/H) CAC (His/H) CAA (Gln/Q) CAG (Gln/Q)	CGU (Arg/R) CGC (Arg/R) CGA (Arg/R) CGG (Arg/R)
	A	AUU (Ile/I) AUC (Ile/I) AUA (Ile/I) AUG (Met/M), *Start*	ACU (Thr/T) ACC (Thr/T) ACA (Thr/T) ACG (Thr/T)	AAU (Asn/N) AAC (Asn/N) AAA (Lys/K) AAG (Lys/K)	AGU (Ser/S) AGC (Ser/S) AGA (Arg/R) AGG (Arg/R)
	G	GUU (Val/V) GUC (Val/V) GUA (Val/V) GUG (Val/V)	GCU (Ala/A) GCC (Ala/A) GCA (Ala/A) GCG (Ala/A)	GAU (Asp/D) GAC (Asp/D) GAA (Glu/E) GAG (Glu/E)	GGU (Gly/G) GGC (Gly/G) GGA (Gly/G) GGG (Gly/G)

1.3 Restrictions within Natural Genetic Code for Amino Acid incorporation

1.3.1. Nature is restricted by twenty naturally occurring amino acids in proteins

Nature has restricted itself by using the twenty canonical amino acids in proteins. This restriction is contributed by the stringency of several components involved in the protein synthesis mechanism. Some of these factors include:

- The structure of transfer RNA and the aminoacylation reaction

- Aminoacyl-tRNA Synthetases and the Challenge of Accurate Amino Acid discrimination

The aminoacylation reaction and the concept of tRNA structure and Identity

Aminoacylation of tRNA occurs in two step reactions: first activation of the amino acids by ATP to form enzyme-bound aminoacyl-adenylates and second transfer of the activated amino acids to the 3'-terminal adenosine of tRNA. Amino acid attachment to tRNA occurs by ester bond formation with hydroxyl groups of the terminal ribose, either 2'-OH for the reactions catalysed by the class I synthetases or the 3'-OH grouping the case of class II enzymes. The activation step is tRNA independent, except for ArgRS, GlnRS and GluRS. Overall, aminoacylation reactions have to yield tRNAs correctly charged, otherwise wrong amino acids will be falsely incorporated into proteins (*5*). This implies correct recognition of both amino acids and tRNAs by the synthetases. However, synthetases can misactivate amino acids, are able to recognize non-cognate tRNAs and can catalyze mischarging reactions. The first answer to the dilemma was brought when it was realized that fidelity in tRNA aminoacylation, and consequently in protein synthesis, mainly relies to highes kinetic efficiency of the synthetases for their cognate substrates and is mostly governed by the k_{cat} of the tRNA charging reactions (*6*). This answer was refined when proofreading (or editing) mechanisms were discovered (*7*). Altogether, the phenomenological view of tRNA aminoacylation fidelity implies a strict correspondence between the charged amino acid and codon read by the carrier tRNA according to the rules of the genetic code. This correspondence is mediated by identity determinants within tRNAs and is defined by the tRNA identity rules, which constitute a "second genetic code".

Aminoacyl-tRNA Synthetases and the Challenge of Accurate Amino Acid discrimination

The most conserved feature of protein translation is the invariant set of twenty amino acids that is used to construct proteins and these form the basic building blocks of proteins. Remarkably, as a series, the encoded amino acids have many structural and

chemical redundancies (e.g. isoleucine vs. valine; threonine vs. serine; aspartate vs. glutamate) and yet they offer sufficient variety to generate the catalytic diversity of enzymes. This diversity is enhanced when almost near identical amino acids play dramatically different biological roles once they are incorporated into proteins.

Thus, despite the apparent redundancy of the encoded amino acids, the identity of a specific residue is often critical to a protein's cellular role, such that even a minor perturbation in protein translation would be intolerable. Error rates in protein translation are estimated to be less that 1 in 3000 (*8*) and perhaps as low as 1 in 38,000 (*9*). The mechanisms by which aminoacyl tRNA synthetases (AARS) guarantees accuracy are critical for the fidelity of protein translation.

In most cases, there is one AARS for each encoded amino acid. The twenty AARS can be subdivided into two different classes (class I and class II), (*10*, *11*) based on conserved sequence and structural motifs. Each enzyme efficiently attaches only its amino acid to the cognate tRNA(s) and a single catalytic or active site catalyzes the following reaction:

$$AA + ATP + AARS \rightarrow AA\text{-}AMP + PPi \qquad \text{.....(I)}$$

$$AA\text{-}AMP + tRNA^{AA} + AARS \rightarrow AA\text{-}tRNA^{AA} + AMP \qquad \text{.....(II)}$$

Precise substrate specificity is fundamental to the accuracy observed in translation because steps subsequent to aminoacylation are less specific for the amino acids.

1.3.2. Adaptor Hypothesis

Although protein translation is a stringent and tightly controlled phenomenon, nature has allowed certain relaxations within the system. In 1955 Francis Crick proposed his "Adaptor Hypothesis," which suggested that "some (so far unknown) structure carried the amino acids and put them in the order corresponding to the sequence in the nucleic acid strand" (*12*). As originally postulated in Crick's Adaptor hypothesis, the faithful synthesis of proteins from messenger RNA is dependent on the presence of perfectly

acylated tRNAs. The hypothesis also suggested that each aminoacyl-tRNA would be made by a unique enzyme. The hypothesis states:

"*The anticodon-codon recognition is found to be independent of the amino acid at the acceptor stem of the tRNA indicating that non-cognate amino acids could be incorporated into proteins*"

In short, the adaptor hypothesis predicts that aa-tRNA decoding on the ribosome depends solely on the tRNA anticodon and is independent of the amino acid structure. The ribosome is a large complex of RNA and protein that catalyzes the translation of RNA into protein. The ribosome is impressive in its ability to convert between nucleic acid and peptide, two structurally unrelated polymers. In addition, unlike most natural enzymes, the ribosome shows broad substrate specificity and uses amino acids with diverse chemical side chains. The adaptor hypothesis suggests that the amino acid specificity of the ribosome should be entirely determined by the tRNA anticodon and independent of amino acid structure.

1.3.3. Specificity of Protein Translation Machinery

The mechanism of protein synthesis is traditionally considered to have two phases with different specificities towards the 20 amino acid side chains. In the first phase, each amino acid is specifically recognized by its cognate aminoacyl-tRNA synthetase (aaRS) and sterified to the appropriate tRNA to form an aminoacyltRNA aa-tRNA). In the second phase, all of the different a-tRNAs are funneled into the translational machinery by binding elongation factor Tu-GTP (EF-Tu-GTP; EF-1 in eukaryotes) to form a ternary complex, which subsequently binds to the ribosome. The traditional view has been that the components of this second phase are not specific for the type of amino acid, and that tRNAs are generic adaptors that are entirely specified by the anticodon. As summarized by Woese (*13*), this view that tRNAs are adaptors which connect the amino acid with the anticodon was hypothesized on theoretical grounds by Crick in 1958 (*14*) and immediately accepted as the paradigm. This lack of specificity for the amino acid in the second phase of translation ensures that all amino acids are incorporated into protein with similar efficiencies and rates, despite their characteristic differences in size, charge

and hydrophobicity. In addition, the elongation factor and peptidyltransferase enzyme shows relatively broad specificity suggesting that a wide variety of amino acid side chains can be accommodated by the ribosome.

Considering all these loopholes in the protein synthesis machinery, this thesis is aimed at designing a novel strategy so that unnatural or non-canonical or non-native amino acids can be engineered into proteins by intervening the transcription-translation events.

Chapter 2: Expansion of the Genetic Code using Site-directed Non-Native Amino Acid Replacement (SNAAR)

2.1 General Introduction

Proteins are the central functional constituents in all living organisms ranging from viruses, bacteria, yeast, and plants to mammals. All of these biopolymers that are formed by natural biosynthetic pathways are composed of a genetically determined sequence of the 20 so-called natural amino acids. In the 1980s, *site-directed mutagenesis* had emerged as a powerful method to incorporate an amino acid in a site-specific position so that more subtle changes in the protein structure and function could be made for protein engineering. The method mainly involves replacing the existing codon with another codon coding for another naturally occurring amino acid (*15*, *16*). Such a gene when expressed in cellular systems would give a new protein containing the new amino acid at the position of interest. Such a technology is indeed extremely powerful in probing the functions of various amino acids, their interactions with other amino acids and their role in the structure of the protein molecule.

Proteins are elegant structures that catalyze various novel reactions because of temporal and spatial arrangement of active amino acid side chains. Extraordinary possibilities exist if these intelligent structures are fused with a vast knowledge of chemical reactivities accumulated over the last 100 years. Although chemical and biochemical methods for modifying amino acid moieties in proteins have been achieved, recent successes in incorporating unnatural amino acids open entirely new avenues for determining protein functions and for the creation of unnatural proteins with novel functionalities. For that purpose, we would have to expand genetic code beyond 20 naturally occurring amino acids.

Although site directed mutagenesis has become one of the most powerful tools for studying structure-function relationships of proteins, it is not without limitations. The most important limitation is restriction of the choice of amino acids to twenty naturally occurring amino acids.

Site-specific non-native amino acid replacement (SNAAR), on the other hand allows vast knowledge of chemistry to be incorporated into catalytic, structural and spatial properties of proteins in a highly defined (site-specific) manner. The physical and chemical properties of proteins are a reflection of the side chains of each of the component amino acids. For some purposes it would be very desirable to have amino acids with side chains of various selected physical chemical properties, such as a keto group, a crosslinker, or a NMR probe group, incorporated into the protein.

This chapter focuses on reviewing scientific literature so as to evolve strategies for SNAAR. The chapter focuses on:

1. Non-native amino acids found in nature
2. Chemical modifications of proteins to engineer non native amino acids non-specifically
3. Site specific Incorporation of Non-native Amino Acids into Proteins
    a. Gene modifications
    b. tRNA modifications
    c. Expression systems
4. Applications and examples of SNAAR analog proteins
5. Current status of site specific non-native amino acid replacement
6. Unique features and novelty elements of SNAAR strategy designed in this thesis

The methods for incorporation of non-native amino acids in proteins are schematically represented in Figure 2.1 and are reviewed in the subsequent sections.

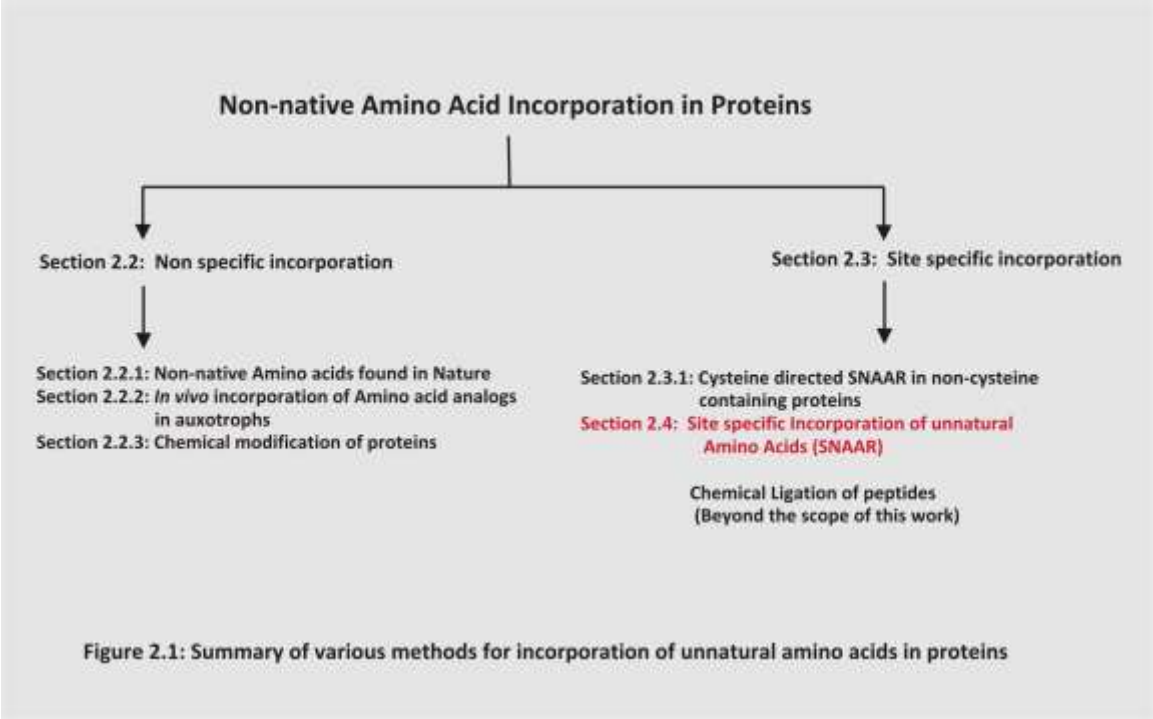


Figure 2.1: Summary of various methods for incorporation of unnatural amino acids in proteins

2.2.1 Non-native Amino acids found in Nature

Nature has found other means of achieving functional diversity by incorporating amino acids, which lay outside the canonical twenty. The non-universality of the genetic code was first observed in eukaryotic mitochondria and since then numerous examples have been described. For example, UGA can not only act both as terminator and a selenocysteine codon, depending on its context, but it can also serve as a tryptophan codon in mitochondria and *Mycoplasma capriocolum* and a cysteine codon in *Euplotes octocarinatus* (*17*). The other terminator codons, UAA and UAG, also have additional functions, for example in *Tetrahymena thermophila* where they are both able to encode glutamine (*18*, *19*). The extent and significance of these and other deviations from the genetic code have been reviewed by Osawa et al. (*20*) who conclude that such deviations are both a result of and contribute to biodiversity. Variable translation of some codons has been observed during the production of recombinant proteins in *E.coli* due to imbalances in the amino acid pools (*4*). The diversity of certain elements within the genetic code suggests it may be possible to accommodate additional amino acids. The most widespread example of this has been the amber suppression of the termination codon UAG, which is discussed in the following section. With the exception of selenocysteine and formyl-methionine, none of these genetic code modifications result in the incorporation of amino acids outside the canonical twenty.

2.2.2 *In vivo* incorporation of Amino Acid Analogs in auxotrophs

There are amino acid analogs, which have been incorporated into the proteins by simply growing the organisms in a medium containing the synthetic analogs. However, cells incorporate these unnatural amino acids in a random manner such that the analog is incorporated instead of its corresponding native amino acid. However, the proteins thus generated, do not contain the unnatural counterpart in a site-specific manner (e.g. azaleucine, (*21*) and plant anti-predatory compounds (e.g. canavanine, (*22*)). The incorporation of such analogues has found a number of applications, the more recent of which are outlined in Table 2.1.

Table 2.1. Examples of recent studies that have made use of the *in vivo* incorporation of amino acid analogs into proteins.

Amino Acid	Analogue	Application	Reference:
Arginine	Canavanine	Stress induction in *Saccharomyces cereviseae*	(*23*)
Cysteine	Selenocysteine	Introduction of doselinide bridges into thioredoxin	(*24*)
Isoleucine	Furanomycin	Incorporation studies	(*25*)
Leucine	Azaleucine	Phenotypic suppression of mutant thymidylate synthase	(*26*)
Methionine	Selenomethionine	Multiwavelength anomalous diffraction	(*27*)
Methionine	2-aminohexanoic acid	Modification of human EGF	(*28*)
Phenylalanine	3-Fluorophenylalanine	19F n.m.r. studies on cAMP receptor protein	(*29*)
Proline	Azetidine	Stress induction in *E.coli*	(*30*)
Tryptophan	5-Hydroxytryptophan	Spectral enhancement of proteins	(*31*)
Tyrosine	3-Fluorotyrosine	Electronic and dynamic studies of high potential iron protein	(*32*)

2.2.3 Chemical modifications of Proteins

One way of introducing non-native chemical moieties into proteins is to use chemical modification of proteins. The use of chemical modification of proteins has increased exponentially during the past two decades. Today the many different uses of chemical modification include determination of relative reactivities of side chain groups, the quantitation of individual amino acids, development of affinity reagents, mechanism-based reagents for pharmaceutical uses, cross-linking reagents, special techniques for bioprostheses, blocking reagents for peptide synthesis, and reagents for specific cleavages of peptide bonds. Table 2.2 shows some key proteins that have been modified using cysteine chemical modification and assayed.

Table 2.2: Examples of recent studies wherein cysteines in proteins are chemically modified

SR. NO.	PROTEIN	MODIFYING AGENT	ASSAY	REFERENCE
1.	*Escherichia coli* asparagine synthetase B.	p-(fluorosulfonyl)benzoyl adenosine (5'-FSBA)	Kinetic study	(*33*)
2.	H(+)-ATPase	Omeprazole	Inhibition study	(*34*)
3.	Ras proteins	nitric oxide (NO) modification	EPR	(*35*)
4.	Class II prolyl-tRNA synthetase (ProRS) from *Escherichia coli*	Sulfhydryl reagent	Kinetic study	(*36*)
5.	Glutamate dehydrogenase from *Clostridium symbiosum* is a homohexamer	Ellman's reagent	Fluoroscence	(*37*)
6.	human type I Ins(1,4,5)P3 5-phosphatase	N-ethylmaleimide (NEM), 5,5'-dithio-2-nitrobenzoic acid, iodoacetate	Inhibition study, Radiolabelling, CD, UV spectra	(*38*)
7.	Heparinase I from *Flavobacterium heparinum*	[3H]Iodoacetic acid	Radioactivity	(*39*)
8.	lac repressor	2-(bromoacetamido)-4-nitrophenol	Reactivity	(*40*)

The modification of **arginyl** residues in proteins continues to be based on the reaction of vicinal dicarbonyl compounds such as phenylglyoxal, 2,3-butanedione, and 1,2-cyclohexanedione to form an adduct. There has also been interest in the reaction of methylglyoxal with arginine residues in proteins (*41*, *42*). Modification of proteins with phenylglyoxal and p-hydroxy-phenylglyoxal has been used to identify arginyl residues involved in binding anionic allosteric modifiers (*43*) and substrates (*44*). There has been more extensive use of phenylglyoxal in the modification of membrane transport (*45-48*).

2.3 Site Specific Incorporations

2.3.1 Cysteine directed SNAAR in non-cysteine containing proteins

One very important avenue towards site-specificity can be through **cysteine mutagenesis**, which exploits the reactivity of this residue towards alkylating and oxidising agents. The approach is of limited usefulness when there are several

preexising cysteine residues. But otherwise, using site-directed mutagenesis, the residue of interest is first mutated to cysteine, and then the protein is purified and subjected to chemical modification. This approach has been successfully employed in biophysical characterisation of a number of proteins including steroid isomerase (*49*), aspartate aminotransferase (*50*) and bacteriorhodopsin (*51*). The latter example illustrates one of the most successful applications of cysteine mutagenesis, the site-specific introduction of spin-labels into membrane proteins and the subsequent determination of electronic paramagenetic resonance spectra.

The degree of specificity that can be attained with this technique is strictly dependent on the number and location of cysteine residues in the protein of interest and consequently it does not provide the flexibility of a method such as site-directed mutagenesis. An approach has been developed in recent years that combine the high specificity of site-directed mutagenesis with the flexibility of chemical modification to overcome these limitations. A large set of reagents, e.g. N-ethylmaleimide (NEM), iodoacetate, 5,5'-dithiobis-(2-nitrobenzoate) (DTNB), methyl methanethiosulfonate (MMTS), react with the site-specifically incorporated cysteine giving rise to a series of amino acid structural analogues at appreciable rates and in good overall yields (*52*). The selective incorporation of these analogues in place of important functional amino acids in a protein allows a more detailed examination of the role or location of that amino acid and serves as probes for structural and functional studies. Thus, with the advent of site-directed mutagenesis techniques, native cysteine residues can now be introduced at any desired position and subsequently modified giving rise to proteins containing unnatural cysteines (For examples, see Table 2.3).

Table 2.3: Example of Site-directed Cysteine Chemical Modification

SR. NO	PROTEIN	MODIFYING AGENT	ASSAY	REFERENCE
1.	Cytochrome b of *Rhodobacter sphaeroides* cytochrome bc1 complex	Nethylmaleimide (NEM)	Accessibility (through kinetic sudy)	(*53*)
2.	Sodium channels	Methanethiosulfonate ethylammonium.	Permeability	(*54*)
3.	Tat-TAR complex	psoralen derivative	Cross-linking	(*55*)
4.	Interleukin-8 (IL-8)	fluorescent N-methyl-N-(2-N-methyl, N-(7-nitrobenz-2-oxa-1, 3-diazol-4-yl)aminoethyl) acetamido (NBD)	Fluorescence	(*56*)
5.	Transmembrane proteins (H(+)-ATPase, K(+)-channel)	Fluorescein-5-maleimide (FM), benzophenone-4-maleimide (BM), N-ethylmaleimide (NEM),	Accessibility	(*57*)
6.	Escherichia coli aspartate aminotransferase	homolysine analogues (gamma-thiahomolysine and gamma-dithiohomolysine) carboxylate series [glutamate, (carboxymethyl)Cys, and (carboxyethyl) Cys]	Kinetic	(*58*)
7.	*Staphylococcal* nuclease	iodoacetic acid	CD Spectra	(*59*)
8.	lactose permease of *Escherichia coli*	[14C]-N-ethylmaleimide (NEM) beta,D-Galactopyranosyl 1-thio-beta,D-galactopyranoside (TDG) 2-(4'-maleimidylanilino)naphthalene-6-sulfonic acid (MIANS	Radioactivity Fluorescence spectra	(*60*)
9.	P-selectin	aziridine or nipsylcysteamine N-ethylmaleimide	Binding assay	(*61*)
10.	Tryptophan synthase	Nile Red	Fluorescence	(*62*)
11.	maltose binding protein (MBP)	4-[N-(2-(iodoacetoxy)ethyl)-N-methylamino]-7-nitrobenz-2-oxa-1,3-diazole (IANBD) and 6-acryloyl-2-(dimethylamino)-naphthalene (acrylodan)	Fluorescence intensity	(*63*)
12.	Rubisco activase.	2-bromoethylamine	Derivatisation	(*64*)
13.	voltage-dependent channel.(colicin Ia)	Biotin	Binding to streptavidin	(*65*)
14.	Glutamine synthetase	arginine "analog" (2-chloroacetamidine (CA) (producing a thioether) and 2,2'-dithiobis (acetamidine)(DTBA) lysine "analog"(3-	Catalytic activity measurements	(*66*)

		bromopropylamine (BPA))		
15.	Thermostable aspartate aminotransferase:	lysine sulfur analog residues(2-bromoethylamine, 3-bromopropylamine, and 2-mercaptoethylamine)	UV-Vis Absorbance, Kinetic	(*50*)
16.	gamma-aminobutyric acid type A (GABAA)receptor	charged, sulfhydryl reagents	Susceptibility to covalent modifications	(*67*)
17.	*Escherichia coli* thioredoxin	aliphatic thiosulfonates	CD spectra	(*68*)
18.	Bacterial sensory receptor	aqueous reagents	Reactivity Sulfhydryl bond formation	(*69*)
19.	ribulose bisphosphate carboxylase/oxygenase	Aminoethylcysteinyl	Enzyme activity	(*70*)
20.	Carboxypeptidase Y	various alkylating and thioalkylating reagents(e.g phenacyl bromide, benzyl methanethiolsulfonate)	Enzyme Activity	(*71*)

As a result, an extremely powerful tool for investigation of proteins has been developed. The vast number of examples cited above underlines its importance and the wide spectrum of systems in which such site-directed sulfhydryl chemistry has been applied with dependable results. However, this technology, besides being restricted to the amino acid cysteine, is also restricted to unique cases such as those proteins, **which do not have any native cysteines**. To synthesize proteins containing non-native amino acids, another alternative was to synthesize small peptides containing the unnatural counterpart. The main disadvantage is the restriction of the length or size of the peptide that can be synthesized chemically.

2.4 Site-directed Non-native Amino Acid Replacement (SNAAR)

For over a decade, *in vitro* biosynthetic methods have been employed to make proteins containing over 100 different unnatural amino acids varying widely in structure, including near analogs of natural amino acids, fluorophores, photoreactive moetities, and α–hydroxy acids (*72-76*).

SNAAR requires site-specific incorporation of non-native amino acids *via* misaminoacylated suppressor tRNA in response to well engineered amber codons in a gene of target protein. The method involves the following steps:

1. Incorporation of a unique codon in the target gene using site-directed mutagenesis: Three nonsense codons (opal (UGA); ochre (UAA); and amber (UAG)) can be used in this case. However, opal and ochre are commonly used stop codons in most living systems. Amber codon (UAG) is the stop codon in very rare cases. Hence amber mutations are commonly used in SNAAR.
2. Aminoacylation of suppressor tRNA with a non-native amino acid.
3. Introduction of the amber mutation containing gene (or mRNA) and misaminoacylated tRNA into a protein translation system (*in vitro or in vivo).*

Key Issues in SNAAR

Scientific literature reveals some of the important criteria that must be met for successful incorporation of an amino acid analogue during *in vivo* translation are:

i. *Uptake of amino acid analogue*: The analogue must be recognized and transported across the cytoplasmic membrane into the cell either by the machinery used for the uptake of its natural counterpart or by the general import machinery (*77*, *78*). In addition, it should not be degraded within the cell.

ii. *Stable formation of aminoacyl analogue tRNA complex*: The analogue must be a suitable substrate for aminoacylation by an aminoacyl-tRNA synthetase (for example p-fluorophenylalanine, (*79*)). Once formed, the aminoacyl-tRNA complex must avoid the editing pathways, which normally prevent misacylation of tRNA (*4*, *50*, *80*, *81*). It has been shown for a mutant of *E.coli* phenylalanine-tRNA synthetase (PheRS) that while both tyrosine and p-chlorophenylalanine are activated by the enzyme, only the naturally occurring non-cognate amino acid is edited (*82*).

iii. *Formation of a stable ternary complex with elongation factor Tu (EF-Tu)-GTP:* The final requirement is that the aminoacyl analogue-tRNA complex must be an efficient substrate for EF-Tu (*83*, *84*), which will then channel it to the ribosomal decoding site (*85*). Thus, the aminoacyl analogue-tRNA must avoid discrimination by EF-Tu, as has been observed for Glu-tRNA^{gln} (*83*), fMet-tRNA^{fMet} (*86*) and Sec-tRNA^{Sec} (*87*). Discrimination must also be avoided at the ribosomal site (*88*, *89*).

Given the extremely high fidelity of cellular protein synthesis, a reasonably large number of analogues have been identified which meet these criteria. This appears to result from *the comparative insensitivity of the aminoacyl-tRNA synthetase editing mechanism to amino acids not within the canonical 20 and the apparent promiscuity of EF-Tu towards the amino acid moiety of aminoacyl-tRNA*.

In summary, the critical elements for performing SNAAR are:

1. Gene sequence which generates the mRNA
2. Misaminoacylated tRNA
3. Expression system

Figure 2.2 gives a schematic representation of these critical elements and the strategies for each of them to achieve site-specific incorporation of non-native amino acids. The following section reviews the scientific literature for each of the strategy for a complete understanding of SNAAR related scientific achievements.

2.4.1 Target Gene Sequence leading to the mRNA sequence

Gene sequence of the target protein is of utmost importance and has to be engineered in such a way site-specific incorporation of non-native amino acid occurs on being translated. This gene sequence would be transcribed to yield the mRNA sequence having specific codons wherein the unnatural amino acids would be routed in the protein synthesis. In order to achieve this, there are three options for engineering the sites by chosing the codons:

1. Engineering suppressor codons
2. Engineering unique codons
3. Engineering Frameshift codons

2.4.1.1 Engineering suppressor codons

Suppressor codons are termination codons are never found within the coding sequence of the gene. There are three types of suppressor codons, amber (UAG), opal (UGA) and ochre (UAA). How do normal proteins terminate in cells with nonsense suppressors? A

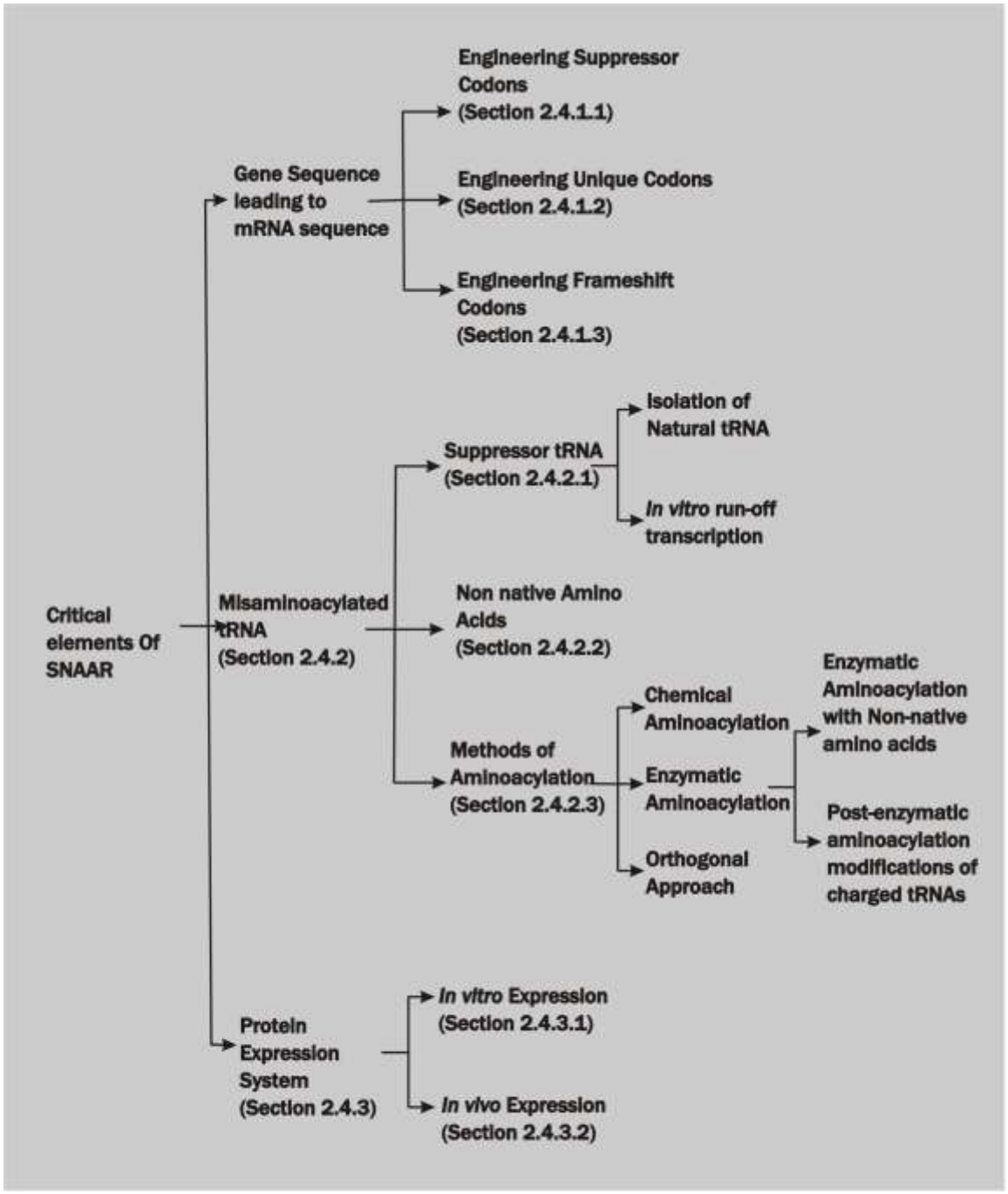


Figure 2.2: Schematic representation of key elements required for performing SNAAR. Various approaches for each work element has been described.

common explanation given for this problem is that there may be tandem nonsense codons at the end of most genes. However, now that we know the DNA sequence of multiple genomes, it is apparent that this explanation is inadequate -- although some genes end with two nonsense codons, many genes end with a single nonsense codon. An alternative explanation relies on the observation that ochre codons are the most common termination signal used at the end of genes. This idea is supported by the observation that although all known ochre suppressors are inefficient, cells containing ochre suppressors are quite sick (possibly because many translation of many genes fails to terminate properly). Amber codons, on the other hand, can suppress better and are rarely found to terminate the genes in most organisms. Hence, it is preferred codon of choice for suppression. Suppression of amber mutations can be achieved using suppressor tRNAs having the AUC recognition on its anticodon arm. This approach has been used extensively for generating SNAAR analog proteins (*73*, *74*, *90*, *91*).

Essentially, all of the *in vitro* and *in vivo* work with unnatural amino acid mutagenesis has, so far, involved the use of amber suppressor tRNA along with an amber stop codon defining the site of interest in the protein gene. The availability of another suppressor tRNA/nonsense codon pair would greatly add versatility of unnatural amino acid mutagenesis and allow site-specific insertion of two different unnatural amino acids in proteins. Hence, ochre mutation/ochre suppressor tRNA pair could be significant.

An ocher mutation results in the creation of a UAA (uracil-adenine-adenine) mRNA codon. Like amber codons, during translation the ocher codon is read as a stop command and the translation process stops. Ochre suppressors are mutant genes that result in the production of transfer ribonucleic acids (tRNA) that have anticodons altered so that they can read ocher codons (UAA sequence codons), and allow the continued synthesis of the protein. Ochre suppressors have shown not to be aminoacylated by any of the cytoplasmic AARSs and read through ochre codons efficiently (*92*).

2.4.1.2 Engineering unique codons

The use of the amber stop codon as an insertion signal limits us to a single unnatural amino acid per suppression reaction. To further expand the genetic code, it has been possible it has been possible to introduce unnatural amino acids in response to other nonsense suppressors, rare codons, or codons with unnatural bases. Missense suppression of the rare arginine codon AGG has proven to be highly efficient; however, insertion of arginine competes with insertion of unnatural amino acids and generates the protein of the same length (*93*). Benner and coworkers used an unnatural base to generate a novel (iso-dC) AG codon, which was decoded in an *invitro* translation reaction containing a chemically misacylated tRNA bearing a CU(iso-dG) anticodon. Although the suppression was efficient, the unnatural tRNA and mRNA has to be chemically synthesized (*94*, *95*). For such a process to be practical, it will be first necessary to develop means to produce the suppressor tRNA and mRNA enzymatically.

2.4.1.3 Engineering frameshift codons

The maintenance of the frame during translation is not absolute and the frameshift mutations can be externally suppressed. Early experiments in the 1960s and 1970s demonstrated that the mutant tRNA in *Salmonella* and yeast with extended anticodons can read non-triplet codons but their read-through efficiency was comparatively low. Recently, it was shown that the efficiency of UAGA quadruplet decoding for a $tRNA^{Leu}$ with engineered 3'AUCU5' anticodon could be elevated to the range of 20–40 % (*96*). An *in vitro* approach based on frameshift suppressions has been developed by Sisido and co-workers over the last decade (*97*). A combination of four base codons AGGN and the corresponding anticodons NCCU was used to examine the scope and possibilities for *in vitro* frameshift suppression for the incorporation of particular canonical amino acids. Indeed streptavidin was substituted with 2-anthranylalanine by using all combinations of quadruplets: AGGG, AGGA, AGGC, AGGU (*98*). The target gene sequence was optimized in a way that its mRNA contains extra downstream stop codons to terminate non-frameshifted products, that is, to abort protein synthesis in the absence of frameshift suppression (*98*). Sisido and co-workers (*88*) observed that incorporation of certain

noncanonical aromatic amino acids is dependent on their adaptability at the active center of the ribosomal A-site. The regularity that emerged from these experiments was that noncanonical amino acids with linearly expanded aromatic groups (e.g. 2-naphtylalanine, p-biphenylalanine, (p-phenyl-azo)phenylalanine) are more favorable than those with widely expanded or bent aromatic groups (e.g. 9-phenyl-anthrylalanine). The transfer of these components into a cell- free rabbit reticulocyte-based system resulted in essentially the same observations (*98*).

The use of quadruplets has aroused great interest because of the possibility of incorporating two or more different amino acids into a single protein. For example, streptavidin mRNA containing CGGG and AGGU quadruplets was translated in the presence of aminoacyl-tRNA$_{CCCG}$ and aminoacyl-tRNA$_{ACCU}$ charged with (7-nitrobenz-2-oxa-1,3- diazol-4-yl)lysine (CGGG signal in mRNA) and 2-naphthyl- alanine (AGGU signal in mRNA). Sisido and co-workers convincingly showed that the use of quadruplets *in vitro* protein biosynthesis provides a strategy for the introduction of noncanonical amino acids at several sites in a single protein(*98*). Although it was expected that the basic advantage of this strategy should be the circumvention of the competition with the release factors, the rules that govern the different suppression efficiencies for different quadruplets are unclear as well. For example, it was argued that some sites in mRNA cause more frameshifts than others (*99*). Therefore, this strategy has been limited both by the overall read- through efficiency of the suppressor tRNAs as well as by the occasional apparent misfolding of nascent proteins containing modified amino acids. The general aspects related to the limits of anticodon size that would help to identify efficient tRNA suppressors of two- , three-, four-, five-, or even six-base codons have recently been the subject of detailed studies (*99*, *100*). For example it was demonstrated that pentaplets (i.e. five-base codons) are indeed suitable for incorporation of special canonical amino acids into proteins (*100*). Extensive examinations performed by Schultz *et al.* have shown that the translation apparatus indeed permits decoding of codons consisting of three to five bases and that each codon type requires different tRNAs. However, six-base codons are incompatible with the translation machineryand pentaplets cannot be suppressed with the same efficiency as quadruplets. Such limits in

codon size are most probably governed by the ribosome itself where the tRNA is a sort of "molecular ruler" that measures codon size during translation (*99*).

2.4.2 Misaminoacylated tRNA

At least two factors are important in determining the accuracy of aminoacylation: a) 'identity elements' in tRNA denote nucleotides in certain positions crucial for protein interactions determining specificity, and b) the occurrence *in vivo* of competition between synthetases for a particular tRNA which may have ambiguous identity.

Misaminoacylated tRNA or wrongly charged tRNA forms an extremely critical element for SNAAR related work. As mentioned earlier, the strategy for adding the wrong amino acids into the proteins has to bypass the proofreading mechanisms of the cellular protein synthesis. Some of the key requirements are the following:

- Suppressor tRNA that is not aminoacylated by any of the endogenous AARSs in the cell
- An AARS that aminoacylates only the suppressor tRNA and no other tRNA in the cell (*101-104*)
- A mutant AARS that only aminoacylates the suppressor tRNA only with the unnatural amino acid and not the natural amino acid.

Using this approach, Schultz, Yokoyama and their coworkers have recently described the site-specific insertion of O-methyltyrosine, 2-naphthylalanine, p-azidophenylalanine, and iodotyrosine into proteins *in vivo*(*104-107*).

To generate a misaminoacylated tRNA, the following components are required, each is reviewed in the section below.

1. Suppressor tRNA
2. Non-native Amino acid
3. Aminoacyl tRNA synthetase
4. Method for misaminoacylation of tRNA

2.4.2.1 Suppressor tRNA

Suppressor tRNA is species of tRNA that are generated *in vitro* having anticodons corresponding to suppressor codons. Such suppressor tRNA is generated from WT tRNA e.g. cys-tRNA, lys-tRNA, phe-tRNA etc. by just mutating the anticodon to suppressor anticodon. Nonsense suppressors are produced by base substitution mutations in the DNA corresponding to the anticodon of a tRNA that cause the anticodon to pair with one of the terminarion (or "nonsense") codons, UAG (Amber), UAA (Ochre), or UGA (Opal).

Examples of nonsense suppressors produced by single base substitutions in *E. coli* are shown in the following table. Note that because of the wobble rules amber suppressing tRNAs can only read UAG codons, but ochre-suppressing tRNAs can read both UAA and UAG codons.

Table 2.4: Some Examples of Nonsense Suppressors in *E. coli*

Suppressor	Type	Anticodon change	tRNA	Gene	Efficiency [1]
supE	Amber	CUG --> CUA	$tRNA^{Gln}$	*glnV*	0.8-20%
supP	Amber	CAA --> CUA	$tRNA^{Leu}$	*leuX*	30-100%
supD	Amber	CGA --> CUA	$tRNA^{Ser}$	*serU*	6-54%
supU	Amber	CCA --> CUA	$tRNA^{Trp}$	*trpT*	
supF	Amber	GUA --> CUA	$tRNA^{Tyr}$	*tyrT*	11-100%
supZ	Amber	GUA --> CUA	$tRNA^{Tyr}$	*tyrU*	
supB	Ochre	UUG --> UUA	$tRNA^{Gln}$	*glnU*	
supL	Ochre	UUU --> UUA	$tRNA^{Lys}$	*lysT*	
supN	Ochre	UUU --> UUA	$tRNA^{Lys}$	*lysV*	
supC	Ochre	GUA --> UUA	$tRNA^{Tyr}$	*tyrT*	
supM	Ochre	GUA --> UUA	$tRNA^{Tyr}$	*tyrU*	
glyT	Opal	UCC --> UCA	$tRNA^{Gly}$	*glyT*	
trpT	Opal	CCA --> UCA	$tRNA^{Trp}$	*trpT*	0.1-30% [2]

[1] The suppression efficiency varies for different sites because the mRNA sequence following the nonsense codon influences how well the suppressor tRNA works.

[2] Note that this suppressor tRNA can still recognize the wild-type Trp codon as well as the UGA

codon.

Generation of such suppressor tRNA species would be possible by two routes:

Isolation of natural tRNA:

Suppressor tRNA is typically synthesized and cloned into *E.coli*. In this approach, the cells are induced to overexpress the tRNA followed by extraction of the total RNA. This total RNA has majority population of the suppressor tRNA. Suppressor tRNA, so obtained, can then be used to charge with non-native amino acids (*108*).

Run-off transcription of tRNA:

Suppressor tDNA cloned in *E.coli* can also be transcribed *in vitro* using the enzyme T4 RNA polymerase. This run-off transcribed tRNA can be used to misaminoacylate with non-native amino acid (*109*). The method of anticodon loop replacement has been used to make derivatives of yeast $tRNA^{Phe}$. By constructing tRNAs with a CUA anticodon, complementary to the amber (UAG) terminator, functional amber suppressor tRNAs were produced. The activity of these tRNAs was assayed in a mammalian cell-free protein synthesizing system. Aminoacylation kinetics using purified yeast phenylalanyl-tRNA synthetase has also been well studied. An efficient procedure for the replacement of the anticodon and the adjacent hypermodified nucleotide (residues 34-37) of yeast $tRNA^{Phe}$ with any desired oligoribonucleotide sequence has been developed (*110-116*).

2.4.2.2 Non-native Amino Acid

The physical and chemical properties of proteins are a reflection of the side chains of each of component amino acids. However, for some purposes it would be desirable to have amino acids with side chains of various selected physical and chemical properties, such a keto group, a crosslinker, or a NMR probe group, incorporated into a protein. This can be done using amino acid analogs or by modifiying post-aminoacylated tRNA.

Several options for various amino acids are available, a summary of which is presented below.

Tryptophan

Tryptophan is an attractive target amino acid for replacement since it is the source for almost all of the absorption and fluorescence of proteins. The diverse and rich indole chemistry offers numerous Trp analogues, which can be tested for translation activity. As the rarest amino acid with only about 1 % of all residues of globular proteins, Trp provides an almost site-specific intrinsic probe for studying protein structure, dynamics, and function. Some of the analogs that can be used as non-native amino acids are listed in the table below.

Table 2.5: Tryptophan analogs used for SNAAR

7-azatryptophan	(*117*)
2-azatryptophan	(*118*)
fluorotryptophan	(*119*)
5-hydroxyryptophan	(*120*)
aminotryptophan	(*121*, *122*)
4-fluorotryptophan	(*123*)
5-fluorotryptophan	(*124*)
6-fluorotryptophan	(*125*)
7-fluorotryptophan, 4-methyl-tryptophan, 7-azatryptopahan	(*117*)
4-hydroxytryptophan	(*122*)
5-hydroxytryptophan	(*126*)
5-aminotryptophan	(*127*)
l-b-(thieno[3,2-b]pyrrolyl)- alanine	(*128*)
l-b-(thieno[2,3-b]pyrrolyl)alanine	(*129*, *130*)
b- selenolo[3,2-b]pyrrolyl-l-alanine	(*131*)
b-selenolo[2,3- b]pyrrolyl-l-alanine	(*132*, *133*)
4-methyltryptophan	(*134*)

Tyrosine

The investigations into the modification of tyrosyl-tRNA synthetase (TyrRS) from *E. coli* yielded a mutant Phe130Ser- TyrRS capable of introducing the Tyr analogue azatyrosine (*135*) into cellular proteins. Yokoyama and co-workers (*136*) developed a suppressor tRNATyr and a mutant TyrRS that incorporate m-iodotyrosine (*137*) into proteins in mammalian cells. The research group of Schultz have evolved a couple of orthogonal *M. jannaschii* TyrRS/$tRNA^{Tyr}_{CUA}$ pairs with an impressive wide range of substrate specificity (*138*). Various amino acids listed in table 2.6 were incorporated into the green fluorescent proteins (*139*) glutathione-S-transferase (*140*) and dehydrofolate reductase (*141*) as a result of efficient suppression of the UGA termination codon.

Table 2.6: Tyrosine analogs used for SNAAR

m-iodotyrosine
O-methyltyrosine
O-acetyltyrosine
2-naphthylalanine
p-aminophenylalanine
p-bromophenylalanine
p-benzoylphenylalanine

Interestingly, a structural homology of TyrRS and TrpRS indicates that the two enzymes can be described as conformational isomers. It remains to be seen whether this feature can be used for a switch or relaxation in their substrate specificity and thus allow an additional expansion of the amino acid repertoire.

Phenylalanine

Fluorinated phenylalanine analogues were used in the initial incorporation experiments as well as in the last few years mainly for protein NMR spectroscopy. Yoshida (*142*)

found the stability of *Bacillus subtilis* α-amylase remains unchanged upon partial replacement of Phe with p-fluorophenylalanine but its activity was altered. Conversely, Richmond (*143*) found that *Bacillus cereus* exopenicillase changed its immunological properties upon substitution with p-fluorophenylalanine. On the other hand, for alkaline phosphatase from *E. coli*, replacement of 56% of the Phe groups with p-fluorophenylalanine had no detectable effects on stability and activity. More recently, o-, m-, and p- fluorophenylalanine (*138*, *144*) were systematically incorporated into recombinant proteins providing them with interesting optical (*145*) and thermal properties as well as with altered kinetics. Janecek and Rickenberg have shown that β-galactosidase expressed in the presence of the Phe surrogate 2-thienylalanine (*119*) is more labile than the normal enzyme. They also found that this 2-thienylalanine although a substrate for protein synthesis, inhibits permease action, that is, its active transport in the cytosol. This is the reason why incorporation of this analogue is quite poor in auxotrophic strains with intact permease function. More recently, the Phe analogue 3-thienylalanine was incorporated into a recombinant periodic protein that forms lamellar crystals comprising regularly folded β-sheets. Koide et al. (*146*) reported the incorporation of Phe analogues with pyridyl groups in recombinant human epidermal growth factor. These analogues, for example 2-azaphenylalanine, 3-azaphenylalanine , and 4-azaphenylalanine, are protonated in acidic solutions.

The discovery of a bacterial PheRS variant with relaxed substrate specificity by Kast and Hennecke (*147*) paved the way for further expansion of the translationally active Phe surrogates and analogues. Table 2.7 summarizes the phenylalanine analogs that have been used by far as non-native amino acids.

Table 2.7: Phenylalanine analogs used for SNAAR

p-aminophenylalanine
p-bromophenylalanine
p-chlorophenylalanine
p-cyanophenylalanine
p-ethynylphenylalanine

Further redesign of the active site of PheRS and its use for a coexpression in *E. coli* resulted in replacement of Phe residues by several phenylalanine analogs in dihydrofolate reductase (*148*, *149*). These designed PheRS variants allow efficient in vivo incorporation of aryl ketone functionality into proteins, which is normally rejected by the cellular translational machinery (*150*). Several Phe-based amino acid substrates such as p-bromophenylalanine and p-iodophenylalanine were incorporated into proteins as a response to the amber stop codon using engineered TyrRS from *M. jannaschii* (*138*). In this way, all these amino acids can have a "dual identity" serving as substrates for both systems.

Histidine

Histidine analogues 1,2,4-triazolyl-3-alanine and 2-methyl-histidine were incorporated into aspartate transcarbamylase and a mutant phospholipase A2. The products of these modifications differ dramatically from the native enzyme with regard to the activity–pH profiles (*151*). In fact, the replacement of His by its noncanonical counterpart 1,2,4-triazolyl-3-alanine yields an enzyme with high activity at acidic pH (*152*). A similar effect was achieved by the incorporation of 2-fluorohistidine (*153*) and 4-fluorohistidine; the pKa value reduced by approximately 5 pH units. Such a dramatic increase in the acidity of the imidazole ring dramatically alters the activity and properties of the fluorohistidine-containing proteins. This strong effect on pKa possibly induces a loss in enzymatic activity in many proteins in the physiological milieu making such His replacements with 2-fluorohistidine and 4-fluorohistidine responsible for the toxicity of these substances. These considerations were fully confirmed in the studies of ribonuclease S with semi-synthetically incorporated 4-fluorohistidine (*154*).

Recently, Ikeda *et al.* demonstrated the possibility of preparing a versatile set of histidine analogues as efficient substrates for protein synthesis (*155*). His residues are specially attractive for engineering enzyme activity and catalysis since they are often found in the active sites of different enzymes, functioning as catalysts in acid–base and nucleophilic processes. Isosteric replacements in the frame of some canonical amino acids such as

Ser/ Cys, Thr/Val represent an excellent tool to study the nature of the catalytic mechanism of residues directly involved in enzyme catalysis. However such replacements are not possible for His side chains. Therefore, engineering of histidyl-tRNA synthetase (HisRS) to extend its substrate specificity will greatly expand the scope and versatility of the enzyme catalysis.

Leucine, Valine, Isoleucine

Pioneering work by Anker and co-workers demonstrated that 5′,5′,5′-trifluoroleucine was incorporated by bacteria but not by eukaryotic systems. These early results were recently confirmed by Tirrell and Tang who in addition succeeded in incorporating hexafluoroleucine into peptides expressed *in vivo* (*156*). *Lactobacillus casei* dihydrofolate reductase expressed in *E. coli* was substituted with (2S,4S)-5-fluoroleucine to study conformational changes of the protein upon ligand binding with 19F NMR spectroscopy. The translational activity of threo-3-hydroxyleucine was reported by Hortin and Boime (*157*).

More recently, Apostol *et al*.(*158*) found that norvaline (*82*) can be incorporated in place of leucine in recombinant proteins. By using the T252Y mutant of leucyl-tRNAsynthetase with attenuated editing activity in E. coli under conditions of its constitutive overexpression, the noncanonical amino acids norvaline, norleucine, allylglycine, homoallylglycine, homopropargylglycine, and 2-butynylalanine were inserted efficiently at the leucine sites of recombinant proteins (*159*).

The valine analogues 2-amino-3-chlorobutyroneic acid, cyclobutaneglycine, penicillamine, and allo-isoleucine have been used to probe the process of membrane transversion of secretory proteins through the endoplasmatic reticulum (*160*). However, the extent of their incorporation to date is still unclear. For example, Porter *et al.* (*161*) reported that "cyclo-butaneglycine is activated by ValRS and forms cyclobutyl-glycyl–tRNAVal, and presumably is incorporated into proteins in lieu of valine". Therefore, additional experiments are necessary to resolve these ambiguities.

On the other hand, a recent experiment by Marliere and co-workers unambiguously demonstrated proteome-wide substitution of valine up to 24 % by a-aminobutyrate in an engineered *E. coli* strain with attenuated ValRS editing function (*162*). Isoleucine analogues, such as 4-thiaisoleucine, O-ethyl- threonine, O-methylthreonine, 4-fluoroisoleucine, and allo- isoleucine were used in pioneering experiments (*143*, *157*, *160*). Novel isoleucine analogues incorporated more recently into a few model target proteins include furanomycin and trifluoroisoleucine (*163*). The systems for the selection and tRNAs aminoacylation of leucine, valine, and isoleucine evolved in their synthetases additional domains for a "double sieve" discrimination of cognate from non-cognate, noncanonical and biogenic amino acids. Engineered valyl-tRNA synthetase or leucyl-tRNA synthetase with inactivated or attenuated editing function should have a great potential to fill up the second coding levels of leucine, valine, and isoleucine with a large number of analogues, surrogates, or similar amino acids.

Methionine

Methionine residues in proteins contribute to their structures with both hydrophobic interactions and hydrogen bonding; they are relatively rare (1.5 % of all side chains in proteins of known structures), and usually located in positions inaccessible to the bulk solvent with only 15 % of all Met residues exposed to the surface (*164*). Met (1) has always been an attractive replacement target in proteins and peptides because of its facile reversible conversion into a hydrophilic sulfoxide form upon oxidation. Using natural translation machinery in *E. coli* or even yeast and mammalian cells, it was possible to substitute Met residues with SeMet (*165*), Nle (*137*, *166*, *167*), ethionine (*143*, *168*) and even telluromethionine (*143*, *169*). Honek and co-workers demonstrated the possibility of incorporation of trifluoromethionine into phage lysozyme (*170*).

A target of the editing activity of the MetRS is the Met metabolic precursor homocysteine (*171*). In general, all Met-like amino acids smaller in size than Met itself are edited (with Nle on the border of this size-based exclusion) (*172*). More recently, Jakobowski showed that S-nitrosomethionine can be transferred to $tRNA^{Met}$ by MetRS and incorporated into proteins in *E. coli* and in the rabbit reticulocyte system (*173*).

Removal of the nitroso group from S-nitorosomethionine-tRNA^{Met} results in homocysteinyl-tRNA^{Met}, and subsequently, its *in vitro* incorporation into proteins in place of Met side chains. The increase in the intracellular MetRS concentration by coexpression experiments in the *E. coli* Met auxotroph CAG18491 and the model protein dihydrofolate reductase (*174*) allowed further extension of the number of Met-based amino acid analogues and surrogates : homoallylglycine, homo- propargyl-glycine, cis-crotylglycine, trans-crotylglycine, allyglycine, 2-butynylglycine, 2-aminoheptanoic acid, norvaline, and azidohomoalanine (*174*). Some of these amino acids have a "dual identity", such as norvaline, 2-butynylglycine, homoallyglycine , and allylglycine, since they can serve as substrates for both MetRS and LeuRS with attenuated editing activity (*159*).

Proline

Proline is a unique imino acid that has a cyclic side chain with a fixed angle, which acts as a classical breaker of both the α-helical and β-sheet structures in water-soluble proteins and peptides. In addition, the rate-limiting step in the slow folding reactions of many proteins involves cis/trans isomerization about peptidyl-prolyl-amide bonds (*175*). Thus, proline is able to play these important roles due to its unique ring size and geometry. In protein and peptide sequences it is often post-translationaly hydroxylated, but has no similar counter- parts among the remaining 19 canonical amino acids. Therefore, proline can be co-translationally substituted only with an expanded amino acid repertoire. Early reports assigned 3,4,-dehydroproline, 4-thiaproline, 4-selenoproline, *cis*-4-hydroxyproline, *cis*-4-fluoroproline, potential proline analogues in protein synthesis. Recent research confirmed that *cis*-4-fluoroproline, *trans*-4-fluoroproline, difluoroproline, and thiaproline can indeed be incorporated into recombinant proteins (*176*). In addition, the proline analogue 3,4-dehydroproline was incorporated in a repetitive periodic protein with lamellar morphology.

In peptide synthesis, pseudoprolines are well known from peptidomimetic studies as serine-, or threonine-derivatized oxazolidines, and cysteine-derived thiazolidines. Chemical aminoacylation of suppressor tRNAs revealed that various proline-based analogues and surrogates that influence the backbone structures of proteins can indeed

be incorporated into proteins. They include pipecolic acid, homopipecolic acid, cyclopropylglycine, azetidine-2-carboxylic acid, and hydroxyproline. Although *in vivo* incorporation of these analogues and surrogates and their derivatives has still not been achieved, there is no doubt that engineering or simple elevation of intracellular levels of prolyl-tRNA synthetase (ProRS) in the context of carefully designed experimental selective pressure, will result in their translational activity.

Arginine

Canavanine, a very potent toxin from the raw seeds of Canavalia ensiformis, is, so far, the only studied analogue of arginine as the substrate for protein synthesis. Pioneering studies from the 1950s and 1960s established that incorporation of canavanine into proteins resulted in altered conformation and disruption of its function. The production of functionally impaired, canavanyl-containing proteins affects cell development and has growth-inhibiting properties. Thus, anticancer and antiviral properties of this substance have been intensively studied (*160*).

Lysine

Lysine residues in proteins play an important role either as targets for chemical derivatization or post-translation modifications. Early studies reported a few analogues of lysine with the potential to be translated into proteins : S-2- aminoethylcysteine, 6-C-methyllysine, 5-hydroxylysine, trans- 4,5-dehydrolysine, 2,6-diamino-4-hexanoic acid, 4-oxalysine, 4-selenalysine. However, there are no reports about any recent research in this field. On the other hand, lysyl-tRNA synthetase (LysRS) could be an attractive target for engineering to relax of even exchange its substrate specificity in the desired direction. The specificity of the aminoacylation reaction (termed as discrimination or D-factor) with regards to the 20 canonical amino acids varies considerably among different aaRS *in vitro* (*171*). The highest D-factor (between 28 000 and > 500 000) is found for TyrRS, whereas the lowest values (between 130 and 1700) were observed for LysRS. Therefore one of the most important tasks for future experiments is to find out whether this observation of "intrinsically relaxed" LysRS substrate specificity can be generalized for further chemical diversification of the proteins based on lysine-derived analogues and surrogates. In addition, the existence of two unrelated forms of LysRS with the same

affinity for lysine but different preferences toward noncanonical amino acids such as thialysine (*177*) can be also exploited for expanding the lysine-like amino acid repertoire. Since Lys residues in proteins are mainly distributed at surfaces, it is reasonable to expect that LysRS-based orthogonal systems that utilize templates with optimized Lys-codons distributions, would most probably replace complicated suppressor-based approaches for chemical diversification of protein surfaces (*140*).

Cysteine

Cysteine continues to be an attractive target for site-specific modification in proteins. This is largely driven by the relative ease of specific modification of cysteine in proteins without concomitant modification of other nucleophilic sites such as lysine and histidine. As a result, a large number of reagents are available for the modification of cysteine. The insertion of cysteine residues into proteins (cysteine scanning) has been of particular value in providing sites for chemical modification. Differential modification with membrane-permeant and impermeant has been used to evaluate the solvent environment around reactive cysteinyl residues (*178*, *179*). Nitrosylation of cysteine residues is discussed below in the section on S-nitrosylation of proteins. Selective reduction of disulfide bonds in proteins has been used to study structure-function relations in proteins for some time (*180*). There are some recent studies on disulfide bond reduction in proteomics. There is significantly more work on the study of the insertion of cysteine residues to form disulfide bridges. The reduction of a disulfide bond in α-lactalbumin by ultraviolet light has been reported (*181*).

The successful modification of carboxyl groups in proteins has been challenging. Woodward's Reagent K has been used to modify a specific **glutamic acid** residue in human erythrocyte band 3 protein (*182*, *183*). N-(3-(dimethylamino)propyl)-N'-ethylcarbodiimide has been used to form an intra-molecular crosslink a protein between a lysine residue and a glutamic acid residue (*184*).

The above studies are cited as examples of the continuing use of site-specific chemical modification to the study of structure-function relationships in proteins. The

information gathered through the chemical modification of proteins over the last fifty years has been used to the manufacture of some novel biotherapeutics. The use of site-specific mutagenesis or protein engineering has been used more frequently of the study of structure-function relationships in proteins and a number of these studies have demonstrated complementarities with chemical modification studies (*185-187*).

2.4.2.4 Methods for Misaminoacylation of tRNAs

The main routes towards the expanded amino acid repertoire for ribosome-mediated protein synthesis are *via*:

1. Chemical Aminoacylation with non-native amino acids
2. Enzymatic Aminoacylation with non-native amino acids
3. Enymatic Aminoacylation with native amino acids and post-aminoacyl modifications of charged tRNA
4. Orthogonal Approach of Misaminoacylation of tRNA

Chemical Aminoacylation of tRNA with Non-native Amino Acids

Lipmann and co-workers demonstrated that misacylated tRNA could be chemically synthesized; they prepared Ala-$tRNA^{Cys}$ by chemical desulfurization (Raney nickel) of Cys-$tRNA^{Cys}$ (*5*). Indeed, misacylated tRNA could substitute Cys with Ala at all positions in haemoglobin, which was taken as a model protein. This finding supported the view that fidelity of protein synthesis was controlled largely at the level of amino acid activation and correct tRNA charging by the aaRS. Later, Hecht and co-workers (*188*) demonstrated the utility of the misacylated tRNAs for effecting amino acid substitutions at predetermined sites in polypeptides. In this way, it was shown that lysyl-tRNA acetylated at the N ε-amino group can be incorporated into the rabbit haemoglobin with nearly the same efficiency as the unmodified one (*188*). The general method for the preparation of misacylated (either suppressor or normal) tRNAs developed by Hecht and his group has been the subject of numerous reviews (*74*, *189*).

Chemical aminoacylation of tRNA was a very important approach and was used initially for unnatural amino acid incorporations into proteins. The structure of tRNA has been briefly described in Chapter 1, section 1.2.3.2, wherein the 3' terminal is known to contain the sequence CCA. This CCA sequence carrying the amino acid is located distal to the anticodon. For chemical aminoacylation of the tRNA, tRNA is initially

abbreviated to remove the terminal CA from the 3′ end of CCA either by the periodate

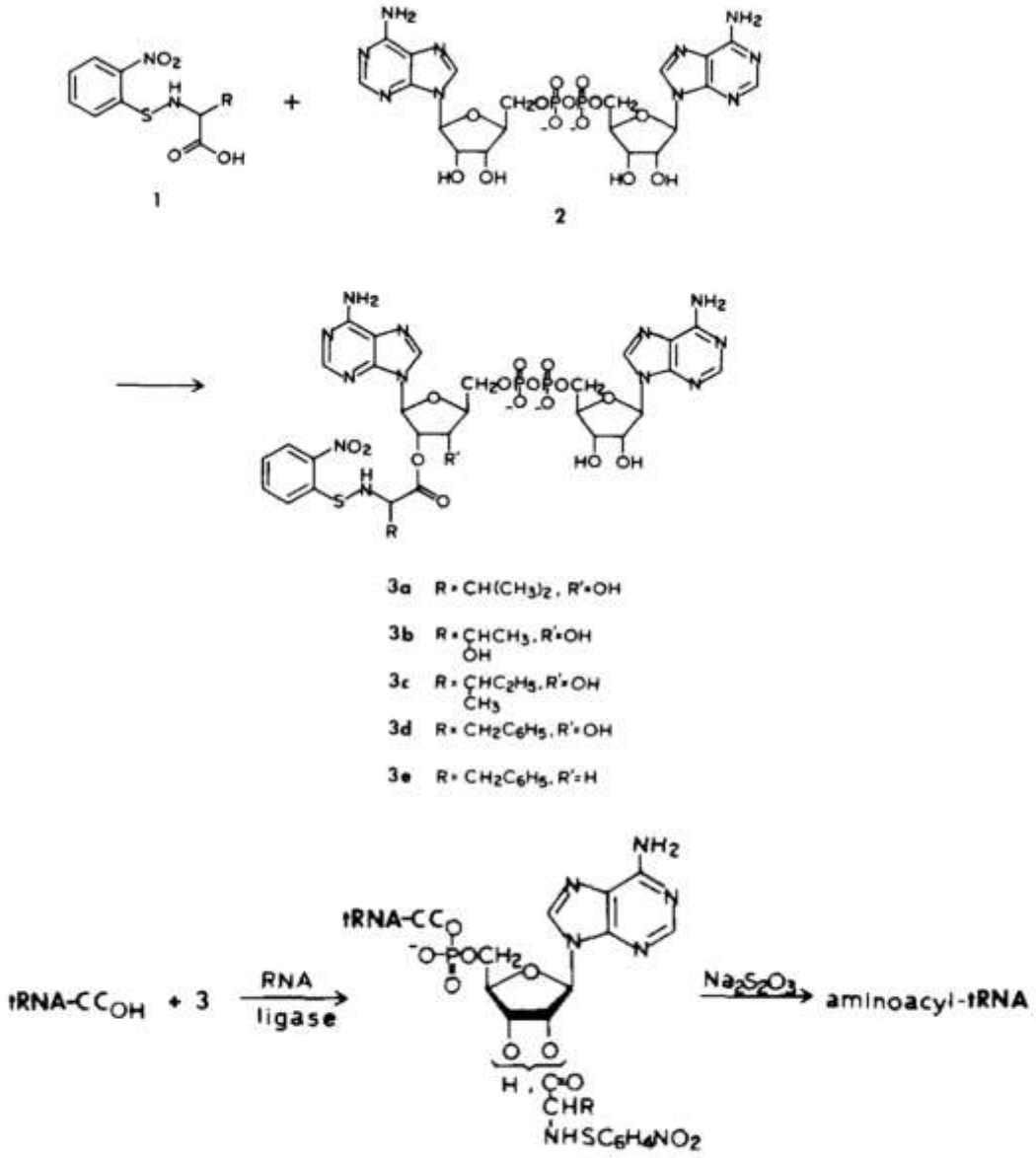


Figure 2.3: Chemical Aminoacylation of tRNA

method of by 3′-exonuclease. This tRNA obtained is used as a substrate for chemical aminoacylation (*90*) in which the non-natural amino acid of interest is first chemically acylated to the dinucleotide pdCpA and this species is then attached to the suppressor tRNA by T4 RNA ligase (*190*, *191*).

tRNA lacking the terminal adenosine moiety (tRNA-C-C_{OH}; abbreviated tRNA) is an acceptor substrate for RNA ligase and that preaminoacylated dinucleoside diphosphates 3a to 3e can be used as donors with quite reasonable efficiency (Figure 2.3). The aminoacylated dinucleoside diphosphates are each N-blocked with an o-nitrophenylsulfenyl group, which serves to diminish the rate of chemical hydrolysis of the aminoacyl moiety. NPS-phenylalanyl-tRNA's prepared by this procedure are deblocked with sodium thiosulfate and further purified by chromatographic separations (*192*). The resulting aminoacyl tRNA is then added to an *in vitro* translation reaction containing an mRNA where codon UAG has been inserted at the desired site of non-natural amino acid incorporation.

The *in vitro* experimental extension of missense or nonsense suppression phenomena for the expansion of the amino acid repertoire was accomplished by chemical misacylation of suppressor or cognate tRNAs. Their addition to a translation system programmed with a gene containing nonsense suppression sites results in the incorporation of the special noncanonical amino acid at the corresponding position in the protein.

Enzymatic Aminoacylation

Aminoacyl tRNA synthetases

The fidelity of protein biosynthesis rests not only on the proper interaction of the messenger RNA codon with the anticodon of the tRNA, but also on the correct attachment of amino acids to their corresponding (cognate) transfer RNA (tRNA) species. This process is catalyzed by the aminoacyl-tRNA synthetases which discriminate with remarkable selectivity amongst many structurally similar tRNAs. The basis for this highly specific recognition of tRNA by these enzymes (also referred to as 'tRNA identity') is currently being elucidated by genetic, biochemical and biophysical

techniques.

Aminoacyl tRNA synthetases form a direct connection between the trinucleotide codons of the genetic code and their corresponding amino acids. Each AARS catalyzes the biosynthesis of a specific, cognate set of AA-$tRNA^{AA}$ isoacceptors. In some cases, the cognate amino acid is structurally similar to one or more encoded amino acids and/or other available metabolites, leading to misactivation or misacylation of non-cognate amino acids. Proofreading mechanisms have now been identified in as many as nine different AARSs. The mechanisms by which each AARS guarantees accuracy are critical for the fidelity of protein translation. Precise substrate specificity is fundamental to the accuracy observed in translation because steps subsequent to aminoacylation are less specific for the amino acid. For some AARSs, the process of amino acid substrate selection is straightforward (e.g. tryptophanyl (*193*) or arginyl (*194*) tRNA synthetase), because the cognate amino acids are structurally dissimilar to other metabolites. In other cases such as IleRS, there is larger number of competing subtrates, such as valine. However, misaminoacylated $tRNA^{Ile}$ is never isolated. This is attributed primarily to the deacylation of activity of AARSs (*195, 196*).

Mutagenesis for Broader Enzyme Specificity

A lot of research has focused on modifying AARSs such that there is **broader enzyme specificity**. For e.g. mutations in the PheRS strains (Ala294→Ser) was found to relax the specificity of the enzyme and was able to charge p-fluorophenylalanine to the $tRNA^{Phe}$ (*147*). With the class I *Escherichia coli* isoleucine tRNA synthetase, which activates isoleucine and occasionally misactivates valine, as an example, a rationally chosen mutant enzyme was constructed that lacks entirely its normal strong ability to distinguish valine from isoleucine by the initial amino acid recognition sieve (*81*). Mutants of *E.coli* AspRS (Asp93→Lys) was able to acylate the *E.coli* $tRNA^{Asp}_{CUA}$. With the related Asp188→Lys mutation *in S.cereviseae* AspRS, it is able to acylate yeast $tRNA^{Asp}_{CUA}$ better in *E.coli*.

Enzymatic Aminoacylation with Non-native Amino acids

The chemical aminoacylation of tRNA or the chemical modification of aminoacylated

tRNA, is often used to allow amino acid analogs to enter the translational pathway. However, the enzymatic aminoacylation of tRNA by aminoacyl-tRNA synthetases (AARS) is much more convenient and practical, especially when considering large numbers of analogs. Effective methods for evolving AARS with relaxed or altered substrate specificities have been developed, and a growing number of interesting amino acid analogs can now be incorporated into peptides through the use of mutant AARS. Surprisingly, the substrate specificity of most wild-type AARS enzymes toward unnatural amino acids is poorly understood, and relatively few amino acid analogs have been reported to be clear AARS subtrates (*197*).

Editing and proofreading mechanisms of AARSs are extremely stringent and hence misaminoacylated tRNAs are generally deacylated intracellularly. Editing reactions are essential for the high fidelity of information transfer in processes such as replication, RNA splicing, and protein synthesis. Amino acid discrimination is achieved through sieves that may overlap with or coincide with the amino acid binding site. However, amino acid analogs which closely resemble the natural amino acids can be used to misaminoacylate tRNAs.

For example, the leucine analogue trifluoroleucine is about 200-fold less active than leucine in the ATP-pyrophosphate exchange activation assay (*198*). Nevertheless, it can easily be incorporated into relatively small coiled-coil proteins under normal expression conditions. On the other hand, the leucine analogue hexafluoroleucine does not support protein synthesis because it is activated approximately 4000-fold more slowly than leucine. The elevated intracellular concentrations of leucyl-tRNA synthetase proved to be a crucial factor for the incorporation of hexafluoroleucine as a response to the leucine-coding triplets in the mRNA sequence of coiled-coil protein HA1 (*156*).

Isotopic amino acids have been used by enzymatic aminoacylation of tRNAs. In a study by Sonar et al, two tyrosines in bacteriorhopdsin were substituted by isotopic tyrosines. Site-directed Isotope Labelling (SDIL) forms the first method of modifying proteins in a site-specific manner and resembling the native protein system (*199*). This can be

regarded as the least perturbed protein structure, since the modified amino acid would be an isotopic form of the same amino acid.

Recently, Hartman *et al* have developed an AARS assay based on mass spectrometry that can be used to rapidly identify unnatural monomers that can be enzymatically charged onto tRNA. By using this assay, they found 59 previously unknown AARS substrates. These include numerous side-chain analogs with useful functional properties. Remarkably, many β-amino acids, N-methyl amino acids, and α,α-disubstituted amino acids are also AARS substrates. The group has used this MALDI-based assay to screen 190 commercially available unnatural amino acids and have identified 90 unnatural backbone and side-chain analogs that can be enzymatically charged onto tRNA (*200*, *201*).

Post-Enzymatic Aminoacylation modification of charged tRNA

This approach is based incorporating chemically modified groups into proteins during their synthesis by modifying the side chain of an amino acid bound to its specific tRNA. In this method, the high selectivity of the aminoacyl tRNA synthetases, which normally only accepts the 20 naturally occurring amino acids, is overcome, by introducing the modification to the amino acid side chains after the tRNA has been misaminoacylated.

Table 2.9: Examples of Post Aminoacylation based modification of tRNA for SNAAR

Name of the protein	Name of the Unnatural amino acid	Type of study done	Reference
Hemoglobin	[14-C]Lys-tRNA	Total tRNA was charged with [14-C]Lys and acetylated at ε-amino nitrogen position by reaction with N-acetyl succinamide. This preparation was added to in vitro protein translation system of rabbit reticulocyte to find 82% incorporation of acetyl-[14-C]Lys.	(*202*)
Preprolactin	Lysin modified by the photoreactive group 4-(3-trifluoromethyldiazirino) benzoic acid	Modified lysine tRNA was used to introduce the photoactive group in the signal sequence of preprolactin. This was crosslinked to the	(*203*)

	interacting component by irradiation and thus the signal sequence-SRP interaction was studied.

Orthogonal approach for SNAAR

Orthogonal tRNA and orthogonal AARSs form an important component to achieve SNAAR and basically means that tRNA and AARS used for mischarging the tRNA should be from a different species from its expression system. Such tRNAs are not recognized by the cells endogenous synthetases and hence they do not get deacylated when charged with a wrong amino acid. Neither are they charged by the natural amino acid by the endogenous synthetases.

Thus, the addition of new amino acids to the genetic code of *Escherichia coli* requires an orthogonal suppressor tRNA that is uniquely acylated with a desired unnatural amino acid by an orthogonal aminoacyl-tRNA synthetase. Optimizing the anticodon recognition between orthogonal tRNA and synthetase significantly increases the incorporation efficiencies of various unnatural amino acids into proteins. An orthogonal tRNA/aminoacyl-tRNA synthetase pair can be generated to incorporate a desired unnatural amino acid into proteins in response to a blank codon. The suppressor tRNA must not only robustly suppress amber stop codons, but it must also be "orthogonal" to the aminoacyl tRNA synthetases in the translation system to prevent insertion of natural amino acids at the UAG codon. This problem has been solved principally by employing tRNAs derived from species of different kingdoms. In the case of *E.coli* S30 expression system, a yeast tRNA Phe modified with a CUA anticodon is competent for translation but is not misacylated by *E.coli* tRNA synthetases (*91*).

Hence, it is important to find a functionally "orthogonal" pair - a tRNA/synthetase pair that react with each other but not with endogenous *E. coli* pairs. A methodology was developed, with positive and negative selection, to evolve the specificity of the orthogonal synthetase to selectively accept unnatural amino acids. Following table

summarizes the pairs of orthogonal tRNA/synthetases that have been developed for SNAAR:

Table 2.8: Summarizes the pairs of orthogonal tRNA/AARSs that have been used for SNAAR

Sr. No.	Species of tRNA	Species of AARS	Species of Expression	Protein	Comments	Reference
1	*E.coli* $tRNA^{Gln}$	*E.coli* mutant GluRS	*In vitro*	*E.coli* surface protein LamB	1500-fold change in specificity found but enzyme could deacylate the mischarged amino acid	(*103*, *204*)
2	*S. cereviseae* $tRNA^{Gln}_{CUA}$	*S.cereviseae* GlnRS	*E.coli*	β-lactamase gene	Cells survived in 500μg/ml	(*205*)
3	*Methanococcus jannaschii* $tRNA^{Tyr}_{CUA}$	*Methanococcus jannaschii* TyrRS	*E.coli*	β-lactamase gene	Cells survived in 1200μg/ml	(*206*)
4	*E.coli* $tRNA^{f\text{-}met}_{CUA}$ mutant	*S.cereviseae* TyrRS	*E.coli*	Chloramphenicol acyltransferase (CAT)		
5	*Human* $tRNA^{f\text{-}met}$	*E.coli GlnRS*	*S.cereviseae*	-	-	(*102*)
6	*E.coli* $tRNA^{Gln}_{CUA}$	*E.coli* GlnRS	COS-1 and CV-1 cells	-	-	(*101*)

Thus by finding suitable orthogonal synthetase or mutants that load the orthogonal tRNA with only the desired unnatural amino acid, it is now possible to generate SNAAR analogs of protein *in vivo*. Similarly, Schultz and his colleagues made an engineered tRNA/synthetase orthogonal pair from the polar archean organism *Pyrococcus horikoshii* that recognizes the four-base codon AGGA. The tRNA has a four-base anticodon loop, and when a ribosome reading an mRNA within the *E. coli* cells encounter AGGA, it inserts the unnatural amino acid L-homoglutamine at that site (*207*, *208*).

By placing both of these systems within the same *E. coli* cell, Schultz and his colleagues have demonstrated, as a proof of principle, that it is technically possible to have mutually orthogonal systems operating at once in the same cell. This opens up the

possibility of doing multiple site substitution with additional unnatural amino acids in the future.

2.4.3 Protein Expression Systems

Expression of SNAAR analog containing proteins can be possible via two expression systems:

1. *In vitro* expression
2. *In vivo* expression

2.4.3.1 *In vitro* Expression

The classical experiments by Chapeville et al. (*5*) which clearly demonstrated that misacylated tRNA is capable of participating in protein synthesis, applied in vitro translation. In fact, cell-free systems with chemically aminoacylated tRNAs demonstrated remarkable ribosome plasticity in terms of the number, as well as the structural, steric, and chemical properties of the amino acids that can be translated into proteins.

To date, three different translation systems have been frequently employed. The most robust system is a supplemented S30 extract from *E.coli*, containing all the necessary components for transcription and translation (*209*). Typically, transcription from the T7 promoter is driven by the *in situ* addition of T7 RNA Polymerase to the extract. Alternative systems include rabbit reticulocyte expression system or wheat germ extract. Unfortunately, the method is technically demanding, and yields of mutant proteins rarely exceed 100μg/ml. Testing different suppressor tRNAs, Schultz and coworkers found that an amber suppressing derivative of E.coli $tRNA^{Asn}$ provided significantly higher yields in the S30 extract system for some amino acids (*210*). Various groups tried to change the compositions of the translation systems to increase the yields of the SNAAR analogs [Kim, 1996 #3003].

Chamberlin and co-workers (*211*) reported site-directed incorporation of 3-iodotyrosine into the 16-residue peptide product translated *in vitro* with a rabbit reticulocyte system. Similarly, by using a nonsense suppressor tRNA from yeast, Noren et al.(*73*)

incorporated several Phe analogues. This position-specific replacement strategy was first postulated by Kwok and Wong (*212*), who have shown that *E. coli* PheRS can aminoacylate yeast $tRNA^{Phe}$ with an efficiency of less than 1 % of that of *E. coli* $tRNA^{Phe}$. Thus, the amber suppressor tRNA derived from yeast-$tRNA^{Phe}$ was especially suitable for *in vitro* incorporation experiments with a coupled transcription/translation system based on *E. coli*. These tRNAs insert the desired amino acid efficiently into the β-lactamase ; positions were taken that correspond to those of the UAG codon on its mRNA. On the other hand, they were not acylated or deacylated by any of the *E. coli* aaRS present in the cell-free lysate. After the suppression reaction in the Zubay transcription/translation system, 2.8–7.5 μg/mL of the replaced β-lactamase was produced, which is about 15–20 % of the native protein yield (30–45 mg)(*73*). These successes were possible because four essential requirements for the application of suppression methodology were fulfilled: 1) the suppressible amber (UAG) mutation was generated in the gene of interest by conventional site-directed mutagenesis ; 2) the design of an efficient suppressor tRNA was achieved ; 3) the suppressor tRNA was chemically acylated, and 4) a suitable *in vitro* protein synthesis system was available.

Evidently, to make fairly exotic amino acids such as ε-(7-nitrobenz-2-oxa-1,3-diazol-4-yl)-l-lysine, or the biotin derivatives intracellularly bioavailable, experimental intervention for "rational re-design" of many components in the living cell is necessary. The range of the intervention is such that it would require simultaneous remodeling and non-invasive changes for the amino acid transport into cells. Furthermore, ways need to be found to avoid metabolic activation or modifications, as well as to pass translation-editing checkpoints of intact cellular physiology. It should always be kept in mind that protein synthesis by the *in vivo* gene expression is carried out in cells, which because of their cell walls and membranes form a closed system. The alternative, open system, is represented by *in vitro* protein synthesis. Thus, the *in vitro coupled transcription/translation*, which is devoid of the problems associated with the mechanisms of cellular physiological functions, could be optimized as a platform for engineering the genetic code with a significantly wider spectrum than in living organisms. This explains the significant interest in the development of high-yield cell-free translation systems.

Various improvements, including optimized batch reactions, addition of chaperones, replenishment of consumable factors, continuous flow or optimization of the relative concentrations of various lysate components, etc. are extensively described in current literature. Therefore, novel generations of even more sophisticated *in vitro* systems with improved performance in terms of protein yield and control of translation conditions would certainly serve as attractive platforms for the expansion of the amino acid repertoire.

2.4.3.2 *In vivo* Expression

Chemical aminoacylations are complex experiments, and *in vitro* expression systems still suffer from low protein yields. This certainly sped up efforts to transfer suppression-based methods to *in vivo* systems. Although the efforts to create an *E. coli*-based efficient GlnRS/tRNA2^{Gln} orthogonal pair (*205*) failed, these experiments yielded a methodological breakthrough in the development of an efficient system for positive/ negative selection of mutant aaRS enzymes and tRNAs. They also brought into focus other important aspects and issues that should be taken into consideration in host strain engineering. For example, amino acid uptake proved to be a critical issue since it is well known that transport of amino acids into cells is not purely passive. Indeed, most of the translationally active noncanonical amino acids are about the same size as canonical amino acids. Thus, it is not surprising that "exotic" amino acids such as (p-phenylazo)phenylalanine are hardly acceptable as substrates for cellular amino acid permease systems. Therefore, the systematic generation of the libraries of cell-permeable amino acids represents an important future research avenue (*213*).

Extra-chromosomally controlled heterologous expression systems in microorganisms are especially practical as platforms for expanding the amino acid repertoire because of their easy maintenance, propagation, and handling. These expression systems can be programmed either for position-specific or multiple-site incorporation of noncanonical amino acids by supplementation with additional translation components. These goals can be achieved in the context of living cells only if four basic requirements are fulfilled: 1) the availability of coding units programming the translation of noncanonical amino acids ; 2) tRNAs capable of decoding these units, and 3) an enzyme capable of charging

the tRNA. 4) Both aaRS and its corresponding tRNA should be free from cross-aminoacylation and compatible with host trans- lation machinery and physiology (*214*).

Furter (*215*) developed for the first time a general approach for the implementation of these criteria for position-specific *in vivo* incorporation of noncanonical amino acids. The Phe analogue p-fluorophenylalanine was incorporated at position 5 as a response to a UGA coding triplet in the mRNA sequence of recombinantly expressed dehydrofolate reductase. This was achieved by introducing yeast PheRS/amber suppressor $tRNA^{Phe}$ into the expression host *E. coli*. There is almost no cross-reactivity between the PheRS from yeast and *E. coli* since identity elements in their tRNAs evolved in a completely different manner. The *E. coli* strain used was a Phe-auxotrophic one and was resistant to p-fluorophenylalanine, that is, its PheRS prevents tRNAPhe from being charged with p-fluorophenylalanine. In contrast, for yeast PheRS, p-fluorophenylalanine is almost as good a substrate as Phe. Thus the system would not be sufficiently defined, since the UGA coding triplet at position 5 of dehydrofolate reductase would be read as both Phe and p-fluorophenylalanine. This problem was circumvented by a careful experimental design of the fermentation, resulting in replacement of 64–75 % at the amber position 5 in dehydrofolate reductase. This is indeed the first demonstration of the introduction of a new redundant aminoacylation pathway and a new design of the hybrid translation system. When such an expression host is exposed to the efficiently imposed experimental pressure, the non- canonical amino acid is preferentially incorporated over the canonical one.

A further advance was made by the use of the TyrRS/$tRNA^{Tyr}_{CUA}$ orthogonal pair from *M. jannaschii* (*206*) in the expression host *E. coli* for the position-specific incorporation of O-methyltyrosine into dihydrofolate reductase in yields of about 2 mg/L (*141*). This was demonstrated in the case of green fluorescent protein (GFP)[89] as well, where a variety of Tyr analogues were incorporated into the chromophore of this protein. The hybrid translation system with the TyrRS/ $tRNA^{Tyr}_{CUA}$ pair from *M. jannaschii* imported into *E. coli* is technically a much more advanced system than the one used by Furter (*216*) as it is much more stringent. In other words, the orthogonality is almost fully

achieved, since amber positions on the mRNA are translated with great fidelity into target proteins with minimal suppression-associated toxic effects. Finally, the use of substituted analogues of aromatic amino acids proved to be advantageous for this system. It remains to be seen whether these *in vivo* expression systems can be developed further for use in molecular biology and biochemistry. Limited success in experiments with GlnRS/$tRNA^{Gln}$ (*205*) poses a dilemma as to whether all synthetase/tRNA systems are generally suited for such design. In addition, the orthogonal synthetase/suppressor tRNA pair for each specific noncanonical amino acid has to be generated through a series of complicated mutagenesis and selection cycles. It is also remains to be clarified whether the *in vivo* transfer resolves the problems associated with context- dependent suppression phenomena (i.e. the problem of equal suppressability of all positions in a desired protein sequence). The *in vivo* approach also brings additional difficulties associated with the metabolic toxicity and bioavailability of the desired amino acids, that is, their transfer into the cytoplasm.

2.5 Site-Directed Isotope Labeling (SDIL) as a first step towards SNAAR

Site-Directed Isotope Labeling (SDIL) forms the first method of modifying proteins in a site-specific manner and resembling the native protein system. This can be regarded as the least perturbed protein structure, since the modified amino acid would be an isotopic form of the same amino acid. SDIL is non-invasive, non-perturbing strategy that allows one to gain better insight into protein structure and function. It is not difficult to come across several studies where site-directed mutagenesis of proteins gives rise to a perturbed protein structure and fails to provide a true picture of structure and function co-relation in proteins. Similarly, although chemical modification of proteins is a powerful technique, it does not have the ability to modify specific amino acids in the protein, thus giving a information about thus perturbed system rather than a targeted question. Cysteine mutagenesis followed by cysteine-chemical modification, forms a powerful technique to incorporate modified cyteines site-specifically into proteins and study them. However, this approach is possible only in proteins containing no native cysteines.

The essential elements required to accomplish site directed isotope labeling are a suppressor tRNA aminoacylated with an isotopic amino acid, a gene containing suppressor mutation at the desired site and a protein translation system. The isotopic aminoacylated tRNA delivers the isotopic amino acid into the nascent protein at the site of the suppressor codon (usually amber codon since it is not commonly used as stop codons in nature) and thus site-specificity is achieved (See Figure 2.4). Cell-free synthesis and exogenous addition of the aminoacylated suppressor tRNA are necessary to prevent aminoacylation of non-suppressor tRNAs with the isotopic amino acid.

SDIL has provided insight into the structure-function relation of various proteins. The most important protein that has been studied using SDIL is **bacteriorhodopsin**, an integral transmembrane found in the archeabacteria *Halobacterium salinarium*. Sonar *et al* reported some very insightful study by generating SDIL analogs of bacteriorhodopsin. These analogs were subjected to Fourier transform infrared difference spectroscopy to detect structural changes at the level of single residues in bacteriorhodopsin (bR). One paper describes the study on labeled tyrosines in bR. FTIR spectroscopy shows that out of 11 tyrosines, only Tyr185 is structurally active during the early photocylce and may be a part of the proton wire formed during the proton pumping activity, i.e. bR→M transition (*217*, *218*). In another work cited by the same group, SDIL and ATR-FTIR difference spectroscopy was used to investigate this conformational change. L-tyrosine containing 13C isotope at the carbonyl carbon was selectively incorporated at Tyr-57, Tyr-147 and Tyr 185 by SDIL. This involved the cell-free expression of bR in the presence of *E.coli* suppressor $tRNA^{Tyr}_{CUA}$ aminoacylated with L-[1-^{13}C]Tyr. ATR-FTIR difference spectroscopy reveals that of the 11 tyrosines, only one peptide carbonyl group of Tyr-185 undergoes a significant structural change during the bR→N transition (*1*, *219*). The main disadvantage of SDIL technology is that it is restricted to non-native amino acids which are isotopic in nature. However, both ***Site-directed isotope labeling (SDIL)*** is a new approach (*217*, *220*, *221*) which holds potential to overcome limitations of site-directed mutagenesis. SDIL is non-invasive and allows one to gain better insight into protein structure.

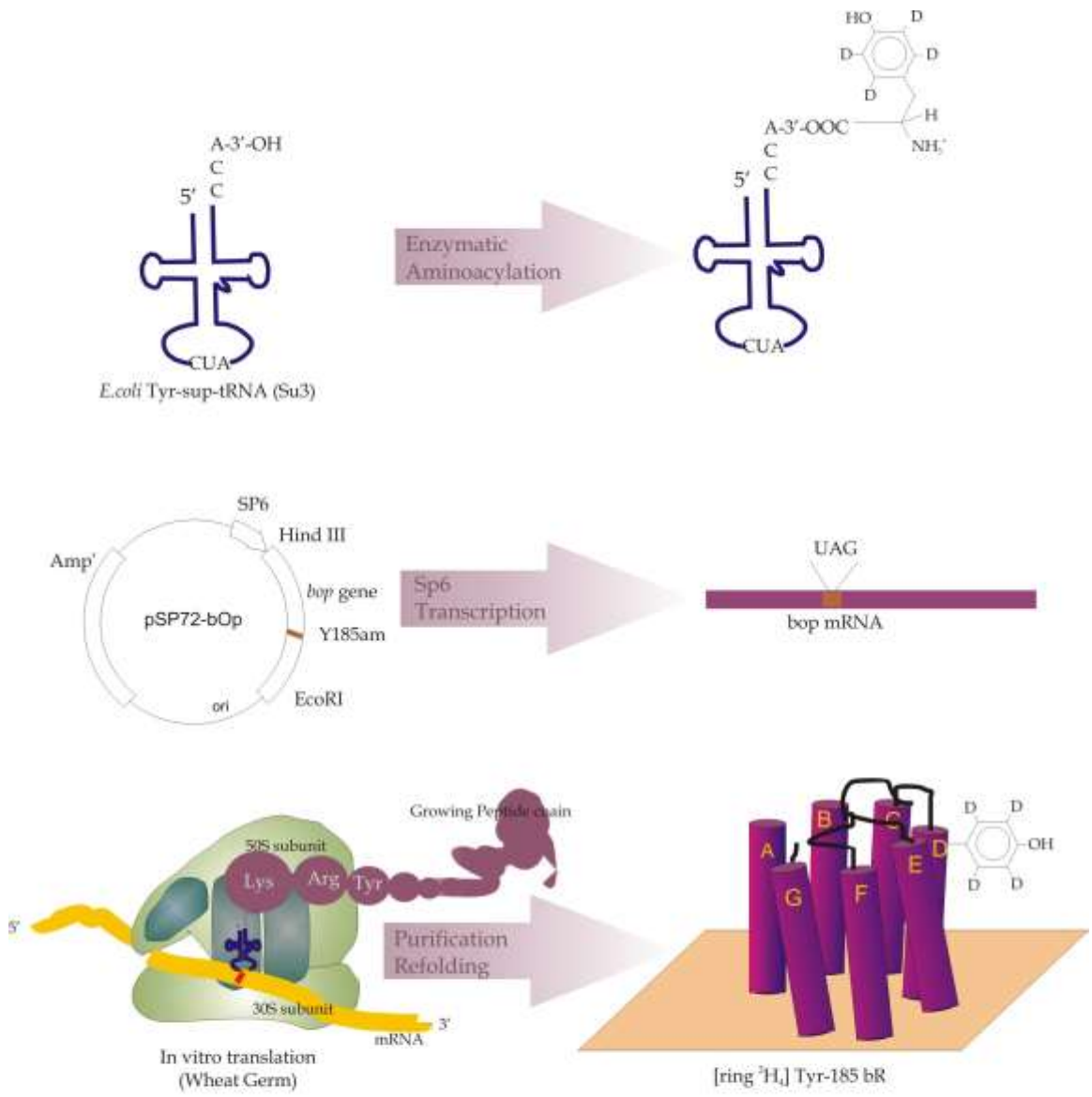


Figure 2.4: Schematic representation of the overall approach for Site-Directed Isotope Labeling (SDIL)

2.6 Applications and literature survey of SNAAR analogs

SNAAR analogs find a wide range of applications. Some of these include:

1. Protein Stability and Function: Unnatural amino acid incorporation can give enormous knowledge on protein stability and function. Since almost any amino acid side chain can be incorporated into proteins, a series of highly related amino acids can be used to examine the linear free energy relationships. Thorson *et al* took advantage of this ability to examine the strength of a tyrosine-glutamate hydrogen bond in staphylococcal nuclease (*222*, *223*).
2. Biophysical probes: A popular method of examining conformational changes in proteins has been the use of intrinsic tryptophan fluorescence. Unfortuantely tryptophan is a weak fluorophore, and such experiments are often complicated by presence of multiple tryptophan residues in the target proteins. Small fluorophores such as azatryptophan can be readily incorporated (*98*, *224*, *225*). Also spin labels have been incorporated into proteins (*226*). Photofunctional crosslinking groups like p-benzoylphenylalanine have been good substrates for translation. Such amino acids may be of great use to identify residues involved in protein-protein interactions.
3. Reactive Handles: Many important proteins contain modification that most likely cannot be introduced during translation due to their bulk. Most extracellular proteins in mammalian systems are heavily glycosylated by enzymes post-translationally. Such modifications are not performed in *E.coli*, making the recombinant production of these proteins difficult. One strategy to generate these chemical modifications would be to use unnatural amino acid mutagenesis to incorporate reactive handles into the protein. Several chemistries are available and sufficiently specific to allow such methodology. Schultz and coworkers demonstrated introduction of ketone modified tyrosine into T4 lysozyme at a surface exposed site, ALA82. The reaction of ketones with hydrazides and hydroxyl amines leads to stable linkages (hydrazones and oximes). It was possible to specifically label the ketones introduced in the T4 lysozyme with a fluorescein hydrazide (*227*). Some of the other examples of

SNAAR analogs are listed below in Table 2.10.

Table 2.10: Examples of various SNAAR analog proteins with their applications

Name of the protein	Name of the Unnatural amino acid	Type of study done	Reference
Lambda phage lysozyme	1,2,4-triazole-3-alanine	Stability of the prodein at lower pH	(*228*)
T4 Lysozyme	Nitric oxide spin labels	To study the protein tertiary fold, its equilibrium dynamics and time-dependent conformational changes	(*229*)
Aspartate aminotransferase	gamma-thiolysine	To check the increase in activity of enzyme	(*58*)
β-lactamase	Protein stability determinants		(*73*)
T4 Lysozyme	Novel backbone structures		(*230*)
T4 Lysozyme	Protein stability probes		(*231*)
Ras protein	Modified main chain, H-bonding and steric conformation		(*232*, *233*)
Staphylococcal nuclease	Glutamate analogs		(*234*)
T4 lysozyme	Biophysical probes like spin and fluorescent labels		(*226*)
T4 lysozyme	β-branched amino acids		(*235*)
Bacteriorhodopsin	Deuterated tyrosine		(*199*)
Dihydrofolate reductase			(*236*)

2.7 Limitations of the current SNAAR technology

The most critical "bottleneck factor" for making SNAAR a general, widely applicable technology is the lack of an easy and a general methodology for synthesizing a misaminoacylated suppressor tRNA. Most of the methods available to date are either very elaborate (e.g. chemical misaminoacylation) or not full-proof (e.g. post-aminoacylation modifications of amino acids such as lysine modify the the α- or the ε-amino group. However, it is important to note that the entire tRNA has large number of such amino groups, which are also accessible for modifications). The effects of these kind of modifications are yet unknown.

Expression of SNAAR proteins *in vitro* is possible by generating misaminoacylated tRNAs chemically and this is an elaborate procedure. In addition, *in vitro* expression

also faces problems such as uncertainty in obtaining the rightly folded protein structures after cell-free expression. As a result, SNAAR remains restricted to proteins.

- which are soluble
- which do not require any assistance in folding into native conformation or
- which can be renatured after complete denaturation
- which do not require any post-translational modification

In addition, cell-free systems are unstable and "fragile" to handle, i.e. a delicate balance has to be optimized for satisfactory protein production. Yields remain a big concern as far as *in vitro* SNAAR is concerned.

Scientific and technological potential of SNAAR can be greatly enhanced if *in vivo* methods are used for protein production instead of *in vitro* translations. For example, if misaminoacylated suppressor tRNA can be introduced directly into a cell for amber suppression of a desired, target protein, then the problem of protein folding or of yield can be successfully overcome. There has been few reports of in vivo SNAAR protein analog expressions, e.g. nicotinic acetylcholine receptor expression using microinjections in *Xenopus leavis* oocytes (*237*). The entire focus of expressing SNAAR analog proteins *in vivo* is centered upon the concept of orthogonality and developing suitable pairs of tRNA/AARSs. Due to this, the type of non-native amino acids incorporated are restricted to a few, e.g. phenylalanine, tyrosine or glutamine. It is important that a 'General Methodology for Mass Production' be developed such that a much more wider repertoire of chemical structures be incorporated into proteins with good yields.

Chapter 3: Bacteriorhodopsin

3.1 Bacteriorhodopsin (bR): A Model Protein

Bacteriorhodopsin (bR), the light driven proton pump from *Halobacterium salinarium* (formerly known as *H.halobium*), occupies an important place in current photobiology research and has enormous applications in advanced biomolecular electronics. Since the discovery in the early 1970s by Stoeckenius and co-workers of this retinal-containing protein, it has become one of the most intensively studied photoactive proteins and a focus for understanding the molecular mechanism of active ion transport and energy transduction in biological systems.

It is important to note the rationale behind choosing bR as a model protein for SNAAR applications. There are three major reasons why applications of SNAAR should be demonstrable in bacteriorhodopsin:

a) bR is an integral membrane protein with all-trans retinal as a prosthetic group. As would be clear from Chapter 5, its heterologous expression requires protein to be folded, reconstituted into membranes and provided with its prosthetic group for its functioning. More than 200 site directed mutants of bR have already been studied. Thus it poses all fundamental challenges that a protein engineer would be required to solve before SNAAR becomes a general approach.

b) Bacterriorhodopsin has a wide range of applications from photodesalination, molecular electronics, artificial retina and transient holography. As a result, bR is extensively studied but critical structure-function elements are still unclear. Being an integral membrane protein, crystal studies are extremely difficult and thus accurate structural information is difficult to elucidate. For the last 35 years, biophysicists, biochemists and protein engineers have concentrated substantial efforts in understanding the structure function relationship. Due to its photocycle with molecular events from femtosecond to millisecond time scales, obtaining functional information has become difficult and critical. Almost every possible biophysical technique ranging from ATR-FTIR, time resolved IR, visible, raman spectroscopies, ESR, NMR, photoacoustic spectroscopy have been applied

to bR and freeze frame time resolved picture of active site amino acid residual changes are evolved.

c) bR is a model protein for biophysical studies related to signal transduction in 7 helical G-protein coupled receptors and retinal binding receptor proteins. From literature study presented here, it would become obvious that a critical conformational change occurs in bR which drives the photocycle. But exact molecular nature of that change is not clear at all. SNAAR offers unique ability to a protein engineer to study such conformational change and extend it to other critical proteins.

Thus, bR and SNAAR have a lot to offer each other. This chapter reviews current scientific literature and attempts to distill out key problems that need to be addressed.

Structure of Bacteriorhodopsin

Primary and Secondary Structure

In 1975, a low resolution bacteriorhodopsin structure was obtained from electron diffraction (*238*) on purple membrane patches that revealed the existence of seven closely packed tubes arranged perpendicular to the membrane plane. The existence of a bundle of transversely packed alpha–helices was also directed by IR dichroism (*239*). Along with the primary sequence of bacteriorhodopsin (*240, 241*) and neutron diffraction (*242*), these data support the existence of seven distinct hydrophobic domains that fold into the membrane as a-helical segments. The transmembrane domain and loop regions have also been defined by immunologica1 (*243, 244*) and proteolytic studies (*245*). (See figure 3.1 for schematic structure of bR). Purple membrane has shown very high stability. While deletion of the carboxyl-terminus has shown to have no effect on purple membrane (PM) lattice dimensions, sheet size, or the electrogenic environment of the ground-state chromophore, truncation study of the entire carboxyl-terminus (*246*) suggest that the bR ground state exhibits two "domains" of stability: (1) a core chromophore binding pocket domain that is insensitive to carboxyl-terminal interactions and (2) the surrounding helical bundle whose contributions to protein stability and proton pumping are influenced by long-range interactions with the extramembranous carboxyl-terminus.

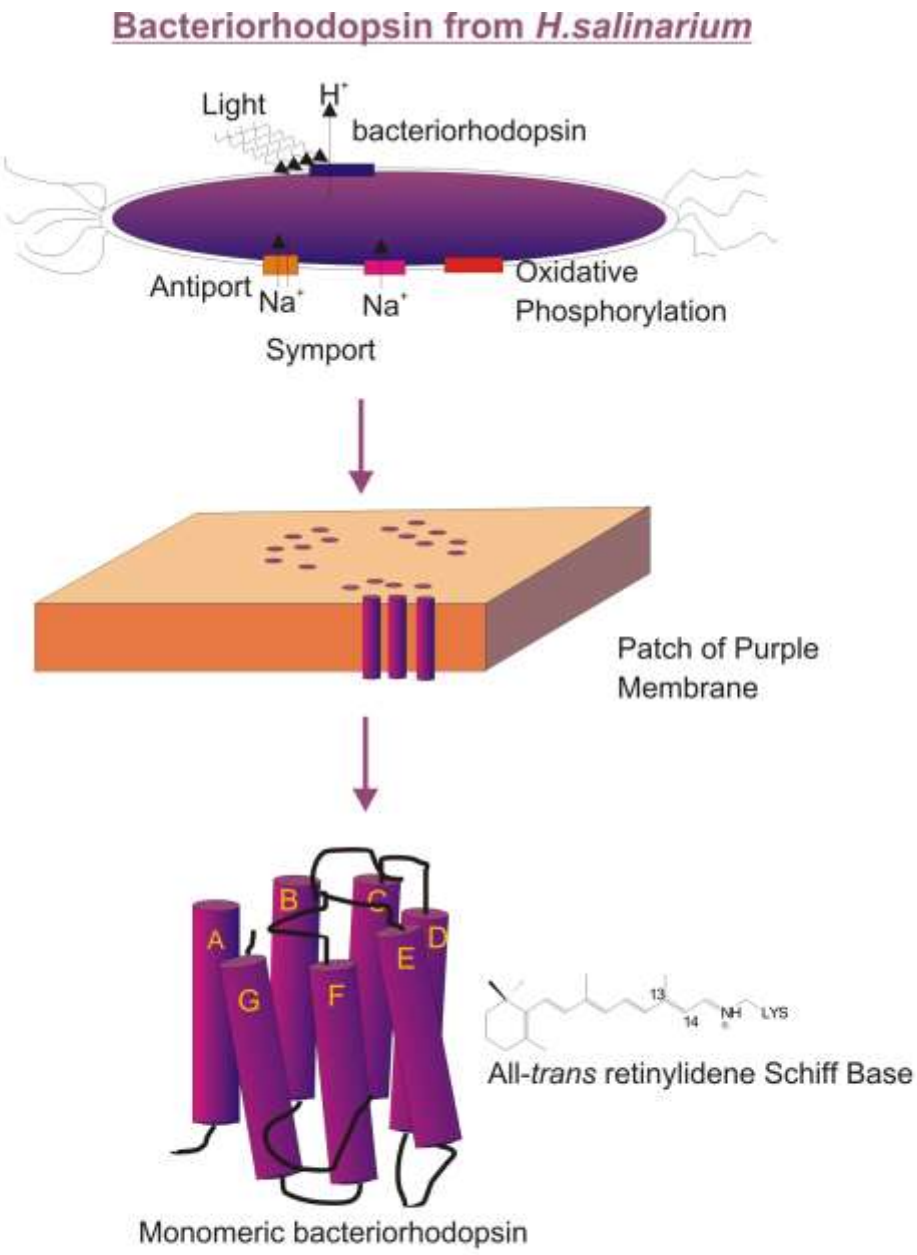


Figure 3.1: Schematic representation of the structure of bacteriorhodopsin. bR is transmembrane protein with seven alpha-helices having all-trans retinal as its prosthetic group.

Location and Orientation of the Retinal Chromophore

A variety of methods have helped establish the location and orientation of the retinylidene chromophore of bR. After initial disagreement regarding the site of attachment for retinylidene Schiff base linkage (*247*), it was conclusively established that it is Lys-216 on the G-helix (*248*) and that the attachment site does not change during the photocycle (*249*).

Because purple membrane can be oriented using a variety of methods including electric, magnetic fields, and isopotential spin drying (*250*), the orientation of the retinal can be probed. For example, visible linear dichroism measurements on oriented purple membrane showed that the polyene chain of retinal is tilted at approximately 20° relative to the membrane plane (*250*, *251*). By comparing the linear dichroism of bactcriorhodopsin containing a normal and analogue retinal, it was also found that the methyl groups on the polyene chain point toward the cytoplasmic surface of the membrane and that the Schiff base proton points down toward the extracellular surface (*252*). Neutron diffraction studies have localized the position of the Schiff base (*253*) and solid-state two-dimensional NMR has determined that the polyene chain is not straight but has in-plane curvature and possibly an out-of-plane twist (*254*). The orientation of the retinal and its direction in the membrane have also been established by elegant photochemical cross-linking experiments using a retinal analog with a diazirine group functionalized to the cyclohexane ring (*255*).

A Spectroscopic Model of the bR Active Site

On the basis of results obtained from site-directed mutagenesis (*256*, *257*), FTIR difference spectroscopy on bR mutants, and a variety of other biophysical data (*258-263*), a detailed three-dimensional structural model of the bR binding pocket referred to here as the "spectroscopic model" was proposed in 1988 (*258*, *264*). As shown in Figure 3.2A, a key prediction of the model is that the protonated Schiff base interacts with a complex counterion consisting of Asn-85, Asp-212, and Arg-82, thereby maintaining

charge neutrality in the active site. The positioning of Asp-85 also requires that Asp-96 be located close to the cytoplasmic surface of bR. Figure 3.2B shows the position ofTyr-

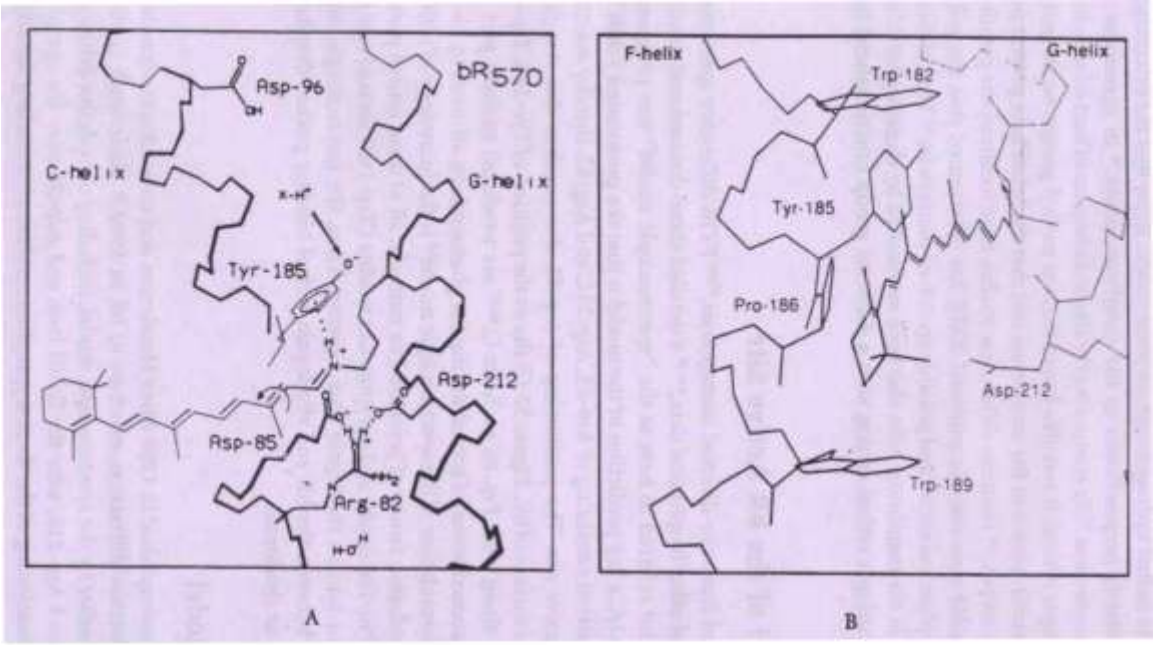


Figure 3.2: The structure of the light-adapted state of bR (bR570) proposed on the basis of FTIR-difference spectroscopy and site-directed mutagenesis. The Schiff base interacts with a complex counterion consisting of Asp-85, Asp-212 and Arg-82. Tyr-185 was proposed to also be partially ionized and undergo a net protonation on formation of the K-intermediate (see text). (B) Three dimensional structure of the retinal binding pocket containing residues Trp-182, Trp-189, and Pro-186 that form a part of a "box" that acts to constrain the possible conformations of retinal. Trp-86 on helix C (not shown) was also hypothesized to participate in the binding pocket. (Adapted from Reference 435).

185, Trp-182, Trp-189, and Pro-186 which, along with Trp-86 on helix-C (*261*, *265*) are predicted to form part of a "box" that acts to restrict the isomerization of retinal during the photocycle to *all-trans* to *13-cis*.

An additional factor in the formulation of the spectroscopic model is the homology of residues in the F-helix of bR and the rhodopsin family of proteins. For example, all of these retinal proteins have a Tyr-Pro pair (Tyr-185/Pro-186 in bR) and a tryptophan residue (Trp-182) located one turn higher in the F-helix, except for human rhodopsin blue pigment. Thus, the spectroscopic model predicts that similarities exist between the bR and rhodopsin retinal binding pocket despite the different retinal configuration in rhodopsin.

Electron Diffraction Model

A major step in this field was accomplished in 1990 when Henderson and co-workers succeeded in obtaining a high-resolution electron diffraction structure of bR in 1990 (*266*). Their work confirms many of the details predicted earlier by the spectroscopic model, including the relative position of the two counterions, Asp-85 and Asp-212, near the Schiff base and Asp-96 near the cytoplasm surface, as well as the relative positioning of the three tryptophan residues surrounding the retinal. While the exact position of the Arg-82 side chain was not specified, its position in the C helix allows it to interact with As-212 and Asp-85. Tyr-185 is also located close to the Schiff base and positioned to form a hydrogen bond with Asp-212, consistent with FTIR evidence that Asp-2L2 and Tyr-185 interact (*262*).

The Bacteriorhodopsin Photocycle

Central to the characterization of proton translocation by bacteriorhodopsin is its unique photocycle. Upon photon absorption, light-adapted bR (bR_{570}) rapidly converts into the first metastable intermediate, K_{630}. This is followed by thermal relaxation in the dark of the chromophore back to its initial state through various structural intermediates characterized by different absorption maxima. As shown in Figure 3.3, along with these spectral changes, a proton is released into the external medium during M formation and a second proton is most likely taken up from the inner medium during O formation.

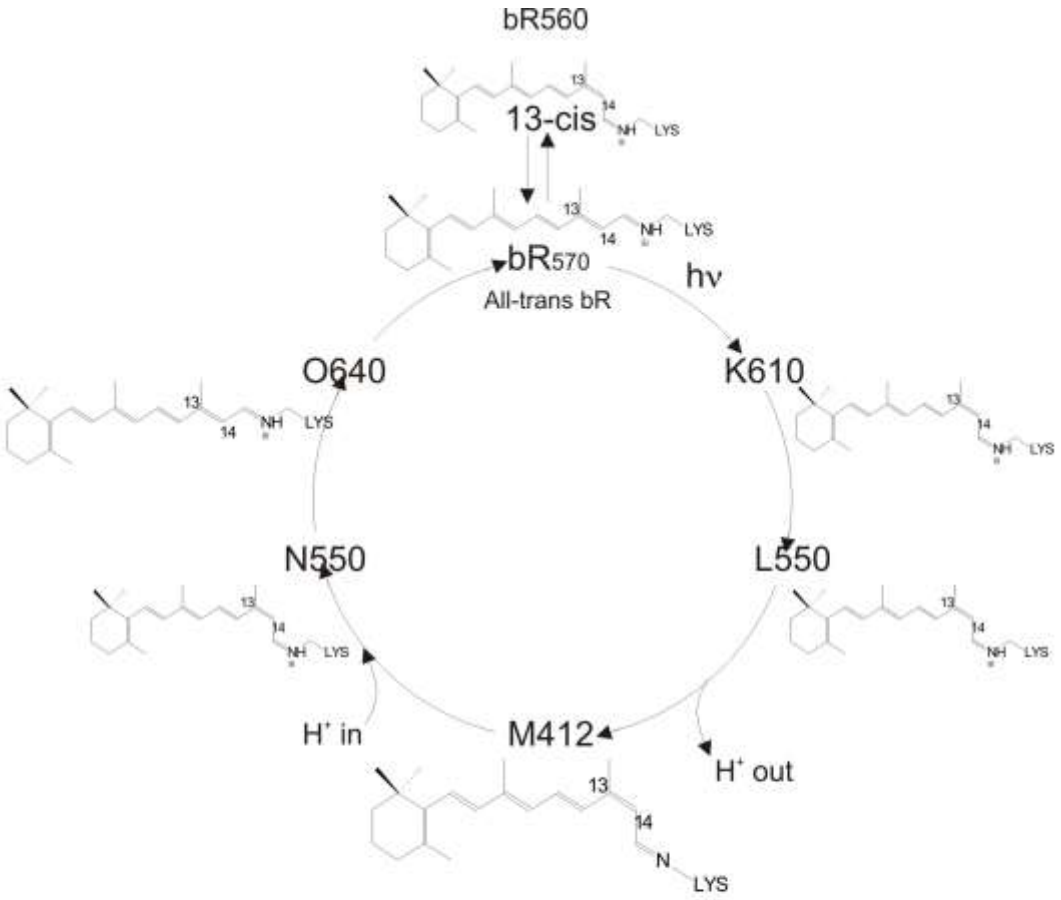


Figure 3.3: Photocycle of bacteriorhodopsin. Different intermediates in the photocycle of bacteriorhodopsin and dark-adapted bR along with the structure of the retinylidene chromophore. The subscripts indicate the approximate λmax of each intermediate. H^+ in and H^+ out denote protons taken up from the cytoplasmic medium and ejected into the extracellular medium, respectively.

Numerous visible absorption studies at low and room temperature have played a key role in the characterization of the main intermediates in the bR photocycle (*267-273*). Recently, femtosecond laser spectroscopy has provided unprecedented time resolution of the earliest events (*274*, *275*). While the existence of the intermediates shown in Figure 3.3 is well established, evidence has also been presented for the possibility of additional intermediates and branching (*271*, *276*).

In addition to visible absorption studies, a number of techniques, including resonance Raman spectroscopy solid-state NMR, FTIR difference spectroscopy, and FT-Raman spectroscopy, have contributed to our knowledge of the conformational changes of the retinylidene chromophore during the photocycle as described in the following sections.

Light and Dark Adaptation

Dark-adapted bR consists of two species (bR_{555} and bR_{570}). Illumination produces the functionally active light-adapted bR consisting only of bR_{570}. Extraction of the chromophore shows that the light-adapted state contains an all-trans-retinal configuration, whereas dark-adapted bR has a 13- cis: all-trans mixture (2:1) (*277*). Resonance Raman spectroscopy and solid-state NMR, along with the elegant use of isotope labels placed at specific positions on the retinal, have established unambiguously the structure of the chromophore in these two states (*278*). bR_{555} contains a 13-cis C=N *syn*-chromophore, while bR_{570} has an all-trans, C=N anti-chromophore (*279-282*). These two states also exhibit distinctly different photocycles (*271*).

The Primary Photochemical Event: bR—>K Transition

The primary photoreaction in the bacteriorhodopsin photocycle is the isomerization of the chromophore from an all-trans to 13-cis-conformation (*283-286*). However, the appearance of intensified bands in both the resonance Raman and FTIR difference spectrum (*287*) assigned to the hydrogen-out-of-plane wag mode of hydrogens on the polyene chain indicates that the polyene chain may not be planar due to twisting about C—C bonds. This might occur for example if the protein-chromophore interactions

impose constraints that prevent retinal from relaxing fully into a 13-cis-configuration. Indeed, several protein bands assigned to Tyr-185, Asp-115, Trp-86, Pro-186, and amide I and II bands appear in the bR→K FTIR difference spectrum, indicating that the protein responds to the initial photoisomerization of the chromophore (*288*).

The initial isomerization also weakens the interaction of the Schiff base with its counterion(s) during the bR→K transition. In particular, the frequency of the Schiff base C=N bond can be used as a probe of the environment and protonation state of the Schiff base during the bR photocycle. In the case of the K intermediate, this mode is assigned near 1609 cm^{-1} (*287*, *289*), a shift of approximately 30 cm^{-1} below that of the C=N stretch frequency in bR_{570}. In general, several studies show that both increased delocalization of electrons throughout the polyene chain and coupling with the NH in- plane bending mode can have a large effect on the C=N vibrational frequency (*289-291*).

The L Intermediate

The transition from the K to L intermediate involves a large blue shift in the chromophore from 630 to 550 nm. The resonance Raman spectrum demonstrates that the chromophore is still in a 13- cis-configuration with a protonated Schiff base (*292*, *293*). However, because of the absence of intensified HOOP modes, it appears to have relaxed from the twisted form present in the K intermediate (*294*). A rotation of the chromophore about the C14—C15 bond has been predicted on the basis of FTIR measurements and on theoretical grounds to explain the drop in the Schiff base P_{Ka} (*295*, *296*) However, resonance Raman studies on the L_{550} intermediate combined with isotope labeling indicate that its chromophore has a C14—C15 S-*trans*-structure (*293*).

The M Intermediate

Formation of the M intermediate is accompanied by a large shift in the λmax from 550 to 412 nm. Early work by Lewis and co-workers using resonance Raman spectroscopy demonstrated that this color shift is due to deprotonation of the Schiff base (*297*). As discussed below, there is strong evidence from FTIR difference spectroscopy (*264*, *298*) as well as a number of other biophysical studies that this proton is transferred to the

nearby counterion, Asp-85, thereby triggering a release of a proton to the external environment. The chromophore remains in a 13-cis- configuration (*297*, *299*) and has a 14-S-trans-C= N anti-configuration about the Schiff base (*279*, *300*). The possibility still exists, however, that more than one form of the M intermediate exists (*276*, *301*), although resonance Rarnan studies have not yet revealed multiple chromophore structures (*300*, *302*).

The N Intermediate

M decay to the N intermediate is accompanied by a shift of the λmax back to near 550 nm, similar to the λmax of the L intermediate. This intermediate was proposed to exist in early studies of the bR photocycle (*303*, *304*) (and references therein), and recently confirmed by several groups (*273*, *305-307*). Resonance Raman spectroscopy shows that it contains a 13-cis-chromophore similar to the M intermediates but with a protonated Schiff base (*293*). Thus, reprotonation of the bR Schiff base occurs during this step in the photocycle, with a decay time of several milliseconds. Several studies have shown that this proton originates from Asp-96, which deprotonates during this transition (*308-311*). A conformational change of the protein also appears to occur between the M and N intermediate (*310-313*).

The O Intermediate

The O intermediate has a red-shifted λmax near 640 nm and intensified HOOP modes (*306*, *314*) similar to the K intermediate. However, in this case, the chromophore exists in an *all-trans*-configuration. Thus, reisomerization of the chromophore occurs between the N and O intermediates. Recent FTIR studies on the bacteriorhodopsin mutant Y185F, which displays a slow O decay at high pH, indicate that this red-shift may be due to neutralization of the Asp-85 counterion (*315*). These observations are further supported by model independent kinetic analyses of FTIR difference spectra (*316*).

Evidence for an Equilibrium between bR_{570} and O

Fourier transform Raman (FT-Raman) spectroscopy has been introduced for studying bacteriorhodopsin (see Rath et al. (*317*) and references therein). This technique uses

exciting light in the near-IR region; thus, photoreactions of light-adapted and dark-adapted bR are avoided. The light-dark adaptation of bacteriorhodopsin and the mutant Tyr-185 →Phe (Y185F) was investigated at room temperature in solution. Interestingly, in comparison to wild-type bR, both the FT-Raman and resonance Raman spectrum of the light-adapted Y185F displayed new features characteristic of the vibrations of the O chromophore. The presence of the O intermediate along with the normal bR_{570} species in the light-adapted Y185F was also consistent with the results from visible absorption spectroscopy (*1*, *318*) and FTIR difference spectroscopy (*319*). Further evidence for the existence of an O-like species in Y185F comes from pump-probe Raman difference spectroscopy, where a red pump beam is found to produce a species very similar to the N intermediate in the photocycle (*317*). These results demonstrate that equilibrium exists between bR_{570} and the last intermediate in the bR photocycle, as demanded by simple thermodynamic considerations (*1*). In the case of Y185F, this equilibrium is shifted more toward the O intermediate than in normal bR. In support of this picture, recent studies show that high temperature can also produce appreciable levels of the O intermediate in wild-type (*320*). A similar phenomena may also occur in the mutant D85N (*321*).

Toward a Proton Pump Mechanism

In the sections below, key elements of a proton transport model were proposed on the basis of (1) the available FTIR data, (2) the electron diffraction-derived coordinates of bR," (3) information about the orientation of the retinal and its changes during the photocycle as discussed in the previous section, and (4) a variety of other data, including the results from extensive structure-function studies of bR using site-directed mutagenesis (*322*).

Key Molecular Events

The arrangement of several key residues investigated by FTIR difference spectroscopy is shown in Figure 3.4. This model retains most of the key elements of the spectroscopic model of bR (*264*) and adds several new features based on more recent studies (*288*).

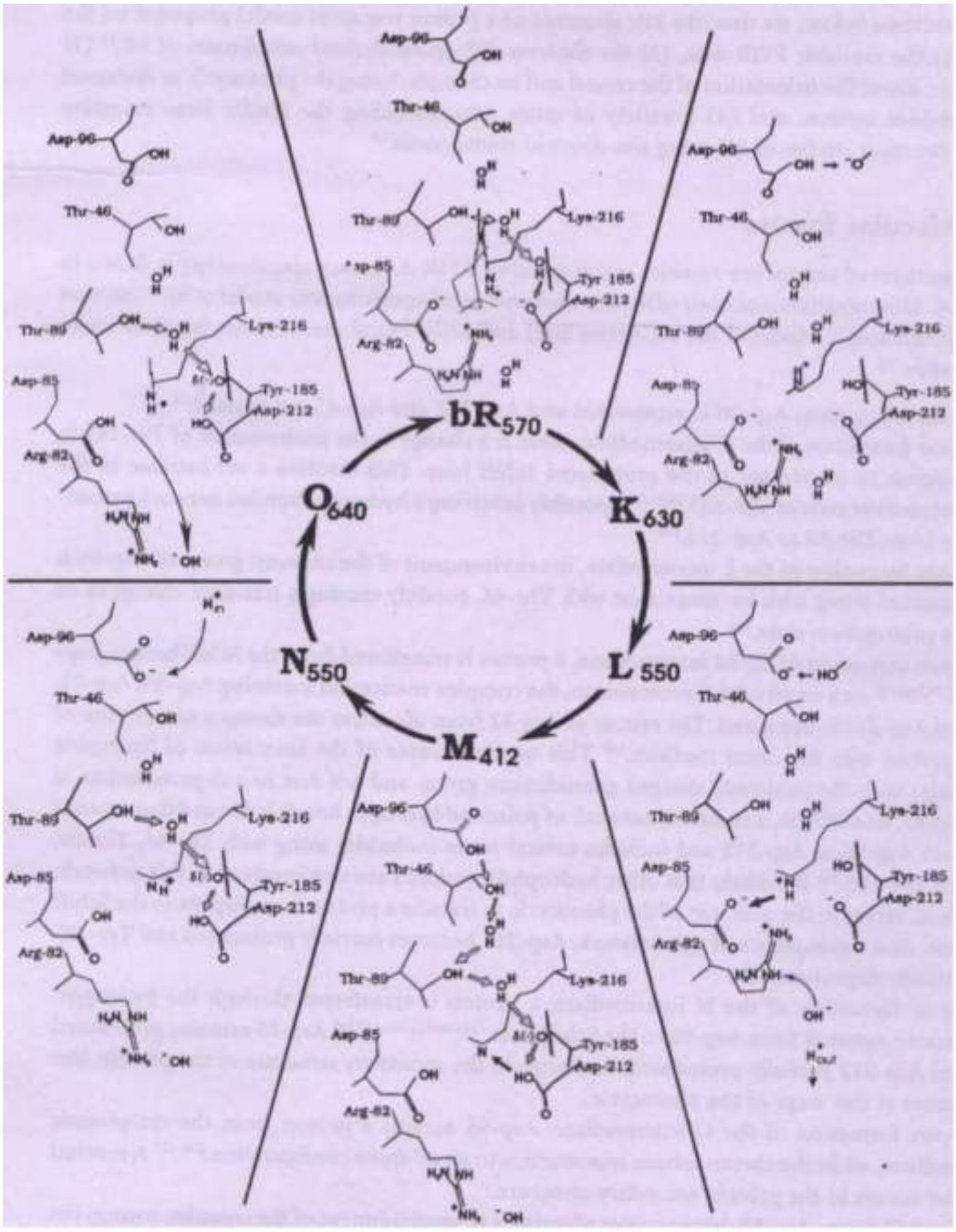


Figure 3.4: Model of hydrogen-bonded proton transport network and protonation changes that occur during the bR photocycle. The approximate position of the residues is based on the coordinate map of the electron diffraction-derived bR structure." (Adapted from Reference 435).

1. In the bR570 state, Asp-96 is protonated and Asp-212 and Asp-85 are ionized (*264*).
2. Upon formation of the K intermediate, there is a change in the environment of Tyr-185 in response to movement of the protonated Schiff base. This involves a net increase in the protonation state of Tyr-185 (*258*, *323*, *324*) possibly involving a hydrogen-bonded network extending from Thr-89 to Asp-212 (*325*).
3. Upon formation of the L intermediate, the environment of the carboxyl group of Asp-96 is disturbed along with its interaction with Thr-46, possibly causing a transient change in its net protonation state (*325*).
4. Upon formation of the M intermediate, a proton is transferred from the Schiff base to Asp85 (*258*, *325*, *326*). As a result of this protonation, the complex counterion involving Asp-85, Arg-82, and Asp-212 is disrupted. The release of Arg-82 from the active site causes a net ejection of a proton into the outer medium (*327*). This occurs because of the association of hydroxide ion(s) with the positively charged guanidinium group and not due to a deprotonation of Arg-82. In addition, a transient network of polarized hydrogen bonds is formed that extends from Asp-96 to Asp-212 and includes several water molecules along with Thr-46, Thr-89, and Tyr-185 (*325*). It is likely that other hydrophilic residues are also involved in this network, which serves in the next step of the photocycle to transfer a proton from Asp-96 to the Schiff base. As a consequence of this network, Asp-212 becomes partially protonated and Tyr-185 partially deprotonated.
5. Upon formation of the N intermediate, a proton is transferred through the hydrogen- bonded network from Asp-96 to the Schiff (*309*, *310*, *316*, *326*, *328*, *329*). Asp-85 remains protonated and Asp-212 partially protonated. A change in the secondary structure of the protein also occurs at this stage of the photocycle.
6. Upon formation of the O intermediate, Asp-96 accepts a proton from the cytoplasmic medium, while the chromophore reisomerizes to an *all-trans*-configuration (*314-316*). A reversal also occurs in the protein secondary structure.

7. Upon O decay, Asp-85 deprotonates allowing a re-establishment of the complex counterion environment around the Schiff base.

A Transient Hydrogen-Bonded Proton Wire in bR

On both theoretical and experimental grounds, it has been postulated that proton wires formed from a network of H-bonded residues and water molecules can serve as key functional elements in proteins (*330-333*). One characteristic of such a network of polarized hydrogen bonds is the delocalization of protons within the network (*331*). For example, even in the case of a hydrogen bond interaction between a high pK_a tyrosine residue and low pK_a aspartate residue, the tyrosine can donate a proton to aspartate and thus become partially deprotonated. For a similar reason, the high pK_a residues in the postulated hydrogen bonded network in bR, which includes Tyr-185, Thr-89, and Thr-46, can donate a proton to Asp-2l2, becoming partially deprotonated in the process.

In support of this model, studies have revealed partial protonation of several key residues in bR. For example, a band appearing at 1277 cm^{-1} in both the FTIR light-dark difference spectrum and photocycle difference spectra was assigned to the ionized form of Tyr-185 (*258*, *334*, *335*). However, the intensity of this band in the absolute IR absorption of light-adapted bR reveals that it reflects only a small fraction of ionized tyrosine (*336*). This low level, estimated at less than 10% at room temperature, may explain why ionized tyrosine has not yet been detected by recent UV resonance Raman (*337*, *338*) studies. Thus, while FTIR difference spectroscopy detects net changes in the protonation state of Tyr-185, it may bever become fully ionized.

A second example is Asp-212, which may be partially protonated in the M, N, and 0 intermediates (*315*). This is consistent with Asp-212 functioning as part of the relay system for proton transport between Asp-96 and the Schiff base as proposed previously (*264*). Partial protonation of Asp-2 12 might also explain why recent solid-state NMR studies (*335*, *339*) detected Asp-85 protonation during formation of the M intermediate but not a similar protonation ofAsp-212 (the N and 0 species were not examined in this study).

Additional evidence for the existence of a transient, H-bonded proton wire in bR comes from FTIR study of the mutant T89D. In this case, an increased protonation of Asp-212 during M formation is indicated by an increased intensity in a band assigned to Asp-212 at 1738 cm^{-1}. Along with this band is the appearance of a new negative band that corresponds to the deprotonation of Asp-89 (*325*) Such a transfer of a proton between Asp-89 and Asp-212 is expected if a proton pathway exists between Thr-89 and Asp-212 in wild-type bR. Recent hybrid Quantum Mechanics/Molecular Mechanics proton transfer calculations indicate that a transient [O] conformer with protonated Asp212 could be sampled during the long-distance proton transfer to the proton release group. These studies suggest that, in the starting proton transfer state O, the retinal is strongly twisted and at least three water molecules are present in the active site (*340*).

An attractive feature of an H-bonded proton wire is the initial movement of the Schiff base due to light-induced isomerization of retinal produces a charge separation that indirectly drives proton transport. Once the K intermediate is formed, its decay back to bR_{570} is prevented because of the high activation barrier for reisomerization of the chromophore to an all *trans*-configuration. Instead, the system moves to a lower energy state by transferring a proton from the Schiff base to Asp-85 during M formation (*341*). Vibrational spectroscopy shows the disruption of the hydrogen bonding of a water molecule, probably with Arg-82, appears. Loss of the interaction of the backbone carbonyl groups in helix G with Tyr57 and the Schiff base, and separation of Tyr57 from Arg82, may be causes of these spectral changes, leading to the stabilization of the protonated Asp85 in M. ^{13}C NMR spectra of [3-^{13}C]Ala-labeled wild-type bacteriorhodopsin (bR) and its mutants at Arg(82), Asp(85), Glu(194), and Glu(204) along the extracellular proton transfer chain show that Arg(82) may function as an information mediator (*342*). Charge flow occurs through two distinct domains, an extracellular pathway that allows the positive charge of the Schiff base that is no Longer shielded by its complex counterion to push a proton into the external medium and a second intracellular pathway that allows a negative charge, initially residing on Asp-

212, to flow toward Asp-96 which then acts to pull in a proton from the cytoplasmic medium.

X-ray Diffraction data from bR crystals where L intermediate was trapped indicate that the retinal chromophore, which is largely twisted in the K-intermediate, takes a more planar 13-*cis*, 15-*anti* configuration in the L intermediate. This configurational change, which is accompanied by re-orientation of the Schiff base N-H bond towards the intracellular side, is coupled with a large rotation of the side-chain of an amino acid residue (Leu93) making contact with the C13 methyl group of retinal. Following these motions, a water molecule, at first hydrogen-bonded to the Schiff base and Asp85, is dragged to a space that is originally occupied by Leu93. Diffraction data from a crystal containing the M intermediate showed that this water molecule moves further towards the intracellular side in the L-to-M transition. It is very likely that detachment of this water molecule from the protonated Schiff base causes a significant decrease in the pKa of the Schiff base, thereby facilitating the proton transfer to Asp-85. On the basis of these observations, it is argued that the vertical movement of a water molecule in the K-to-L transition is a key event determining the directionality of proton translocation in the protein (*343*).

Site-Directed Isotope Labeling of bR

FTIR difference spectroscopy has been limited by the absence of a general method for placing isotope labels at specific positions in a protein. A new approach based on site-directed isotope labeling (SDIL) was reported (*199*, *218*, *220*). SDIL-analogs of bR were produced by expressing bR in an in vitro system containing a tyrosine-suppressor tRNA aminoacylated with isotopically labeled tyrosine followed by refolding. SDIL-analogs containing ^{2}H and ^{13}C isotopes at both specific tyrosine an backbone peptide carbonyl groups were produced with properties almost identical to native bR. FTIR analysis of the bR $\rightarrow$K, L, M, and N steps in the photocycle of these essentially unperturbed SDIL-analogs has led to the identification of structurally active groups that may be involved in proton transport and in the coupling of chromophore isomerization to protein conformational changes.

In the future, it is likely that the SDIL approach will be combined with other biophysical methods besides FTIR to study bR. For example, rotational resonance NMR (*344*, *345*) can be used to determine distances (and changes in distances) between specific atoms such as in the Schiff base and Tyr-185. Although the FTIR measurements only require samples in the 10- to 25-μg range, it should be general, site-directed isotope labeling should also be applicable to a wide range of other proteins including those involved in enzyme catalysis, ion transport, and signal transduction.

Protein Conformational Changes and Chromophore Isomerization

A variety of neutron, X-ray and electron diffraction experiments have established that the transmembrane regions of bacteriorhodopsin undergo significant light-induced changes in conformation during the course of the photocycle ultimately resulting into vectorial proton pumping (*346*, *347*).

Molecular dynamics simulations reveal (*348*) that the M intermediate actually comprises a series of conformations involving a) a motion of retinal; b) protein conformational changes; and c) diffusion and reconfiguration of water in the space between the retinal Schiff base nitrogen and the Asp-96 side group. (a) turns the retinal Schiff base nitrogen from an early orientation toward Asp-85 to a late orientation toward Asp-96; (b) disconnects the hydrogen bond network between retinal and Asp-85 and tilts the helix F of bR, enlarging bR's cytoplasmic channel; (c) adds two water molecules to the three water molecules existing in the cytoplasmic channel at the bR_{568} stage and forms a proton conduction pathway. The conformational change (b) of the protein involves a 60 degrees bent of the cytoplasmic side of helix F and is induced through a break of a hydrogen bond between Tyr-185 and a water-side group complex in the counterion region.

Several earlier studies have focused on the nature of the secondary structural changes during the bR photocycle. Previously, such changes have been detected by circular dichroism (*349*) X-ray, and neutron diffraction experiments (*350*). Electron diffraction

studies of the mutant D96G revealed that the cytoplasmic portion of the F- and G-helices undergo a small tilt away from the membrane normal (*351*). Interestingly, this tilt and its reversal might facilitate movement of a proton from Asp-96 to the Schiff base and subsequent proton uptake by Asp-96 from the cytoplasmic medium. Electron diffraction studies also show that significant structural changes occur in wild- type bR upon formation of two different low-temperature substrates of the M intermediate (*352*).

Electron crystallographic analysis of light-driven structural changes in wild-type bacteriorhodopsin and a number of mutants (D38R, D96N, D96G, T46V, L93A and F219L, and the triple mutant D96G/F171C/F219L (*353*)) support a model for the photocycle of wild-type bacteriorhodopsin in which the structures of the initial state and the early intermediates (K, L and M) are well approximated by one protein conformation in which the Schiff base has extracellular accessibility, while the structures of the later intermediates (M, N and O) are well approximated by the other protein conformation in which the Schiff base has cytoplasmic accessibility (*354*). Widely varying views exist on the extent of conformational changes essential for proton translocation. In an EPR study wherein spin labels were attached to pairs of engineered cysteine residues at AB or CD interhelical loops of bR, distances between labels increase transiently by about 5 A during the photocycle. This opening occurs between proton release and uptake, and may be the conformational switch that changes the accessibility of the retinal Schiff base to the cytoplasmic surface after proton release to the extracellular side (*355*).

Unlike wild-type bR, the bR triple mutant D96G/F171C/F219L has been shown to undergo only minor structural rearrangements during its photocycle indicating that large conformational changes observed in the photocycle of the wild-type is not a prerequisite for vectorial proton transport (*356*).

In contrast, majority of studies indicate essential conformational changes during proton pumping. Involvement of F-helix in the proton translocation has been substantiated by various studies including work reported in this thesis. Several FTIR studies also show that a significant change occurs in the bR secondary structure and that it is correlated

with the M→N transition (*310*, *312*, *315*) (*316*) (*311*, *357*) (*358*). While these experiments were performed using FTIR transmission methods, which are restricted to thin, hydrated films, similar results were also obtained in the presence of a bulk aqueous medium by using attenuated total reflection (ATR) (*220*). This indicated that these changes arise from the tilting of the F helix about the Tyr-185/Pro-186 region of bR, which could serve as a hinge (*219*, *263*).

An electron crystallographic analysis of the N intermediate from the mutant F219L (*359*) gives a three-dimensional view of the large conformational change that occurs on the cytoplasmic side after deprotonation of the retinal Schiff base. Helix F, together with helix E, tilts away from the center of the molecule, causing a shift of approximately 3 A at the EF loop. The top of helix G moves slightly toward the ground state location of helix F. These movements open a water-accessible channel in the protein, enabling the transfer of a proton from an aspartate residue to the Schiff base. Time-resolved electron paramagnetic resonance (EPR) spectroscopy on site-directed spin-labeled bacteriorhodopsin studies (*360*) are consistent with a small movement of helix C and an outward tilt of helix F. These helix movements are accompanied by a rearrangement of the E-F loop and of the C-terminal turn of helix E.

The conformation of the structured EF interhelical loop of bacteriorhodopsin and its change in the M photointermediate were assessed by measuring the rate of reaction of 16 single engineered cysteine residues along the loop with water-soluble sulfhydryl reagents (*361*). The EF-loop should be affected by the well-known outward tilt of helix F in the M and N intermediates of the photocycle. A second mutation in each cysteine mutant, the D96N residue replacement, allowed full conversion to the M state by illumination. These studies indicate that the conformation of the EF-loop itself is changed. Thus, the motion of the loop in M is more complex than expected from simple tilt of helix F, and may include rotation that unwinds its twist. In another study involving time-resolved fluorescence depolarization and site-directed fluorescence labeling (*362*), the conformation of the AB-loop and EF-loop of bacteriorhodopsin and of the fourth cytoplasmic loop (helix VIII) of bovine rhodopsin were assessed. This study

suggests that the surface potential-based switch of the EF-loop is the missing link between the movement of helix F and the transient surface potential change detected during the photocycle of bacteriorhodopsin.

Low temperature FTIR studies of hydrated, glucose-embedded purple membrane at 170 K, 220 K, 230 K, and 240 K show that the F-helix tilt initiates during L→ M transition. Consistent with previous structural studies, an adjustment in the position or in the degree of ordering of helix G accompanies this motion (*363*). According to the current structural model of bacteriorhodopsin, Ile222 is located at the cytoplasmic end of helix G. Labeling of the single cysteine of the site-directed mutant Ile222 --> Cys with p-chloromercuribenzoic acid and determination of the position of the labeled mercury by x-ray diffraction in the unphotolyzed state, and in the M_N photointermediate (39), it is found that the position of the mercury at residue 222 is shifted by 2.1 +/- 0.8 A in the M_N intermediate. This agrees with earlier results that suggested a structural change in the G helix. The true conformational change associated with formation of the M intermediate was investigated by using a 3D crystal (space group P622) prepared by the membrane fusion method. Results indicate that, upon formation of the M intermediate, helix G moves towards the extra-cellular side by, on average, 0.5 angstroms. This movement is coupled with several reactions occurring at distal sites in the protein: (1) reorientation of the side-chain of Leu93 contacting the C13 methyl group of retinal, which is accompanied by detachment of a water molecule from the Schiff base; (2) a significant distortion in the F-G loop, triggering destruction of a hydrogen bonding interaction between a pair of glutamate groups (Glu194 and Glu204); (3) formation of a salt bridge between the carboxylate group of Glu204 and the guanidinium ion of Arg82, which is accompanied by a large distortion in the extra-cellular half of helix C; (4) noticeable movements of the AB loop and the cytoplasmic end of helix B. But, no appreciable change is induced in the peptide backbone of helices A, D, E and F.

FTIR measurement on mutants of bR also establish that the structural change that is normally correlated with the M—N transition in wild-type bR can also occur under conditions where the Schiff base does not deprotonate or reprotonate (*313*, *364*). For

example, the 1670- cm^{-1} band is detected in the photocycle of the mutant Y57D at 250 K, even though decay of the L intermediate is blocked, thereby preventing deprotonation/reprotonation of Schiff base and Asp-96 (*217*). One implication of these results is that the major structural change in bR is driven by steric interactions produced within the retinal-binding pocket as a direct consequence of chromophore isomerization**.** This conclusion was further substantiated by Fourier transform infrared/resonance Raman studies of the alkaline form of the mutant Asp-85→Asn which demonstrate that the M-->N conformational change can occur even in the photocycle of an unprotonated Schiff base form of bacteriorhodopsin (*365*).

Detection of Water Molecules in the Active Site of bR

Several lines of evidence indicate that one or more water molecules may be located in the active site of bR (*217*, *366*, *367*). This was confirmed directly through the assignment of bands in the region above 3500 cm^{-1} to the OH stretch mode of water (*368*) in the FTIR difference spectra of bR. One or more weakly hydrogen-bonded water molecules were detected that undergo a change in H-bonding as early as formation of the K intermediate (*369*). It was found that the mutation Asp85→Asn (D85N) and Y57D (*369*) cause the disappearance of some of these bands; thereby indicating that at least one water molecule is located in the active site. (A second may be located outside of this region (*369*). However, the exact position and number of water molecules still remains to be determined. An additional question is the role of these water(s), which may include participation in a proton wire (*325*).

3.2. Unsolved Questions in bR

While great progress has been made in the past few years in elucidating the mechanism of proton transport in bacteriorhodopsin, a complete understanding has not yet been achieved. For example, the path of proton transport through the protein has not yet been established, although as discussed in this chapter the outlines of a hydrogen-bonded proton transport network are emerging. Key questions also remain about the specific mechanism of chromophore reisomerization, protonation changes of Asp groups, and

the location and role of secondary structural changes that are detected at different stages of the photocycle.

Progress toward answering these questions is likely to depend on the further development of new biophysical techniques, including time-resolved FTIR (*310*, *312*, *370-372*) ultrafast time-resolved IR studies (*373*) time-resolved X-ray diffraction (*371*) and femtosecond visible absorption measurements (*374*). The ability of FTIR difference spectroscopy to probe changes in specific amino acid residues in bR during the photocycle should continue to provide details of proton translocation mechanism in bR, especially as methods improve to assign bands in the FTIR difference spectrum.

Chapter 4: Overall Goal of this Thesis: Expanding the Genetic Code using Site-specific Non-native Amino Acid Replacement (SNAAR)

4.1 Overall Goal of the Thesis

A study and review of scientific literature reveals necessity of site-specific non-native amino acid replacement for protein engineering, and thus expanding the genetic code beyond twenty naturally occurring amino acids. While the applications of SNAAR are enormous for genomics and proteomics, several limitations need to be overcome to make SNAAR a general technique. The study further reveals that SNAAR remains restricted to proteins a) which are soluble, b) which do not require any assistance in folding into native conformation or c) which can be renatured after complete denaturation and d) which do not require any post-translational modification.

One of the most important limitations that SNAAR faces is its protein expression system. Since SNAAR requires "intercepting" protein translation, till date, SNAAR analogs have been made *in vitro* in cell-free systems such as S30 extracts, wheat germ extracts, rabbit reticulocyte extracts etc. which are rich sources of protein translation machinery. *In vivo* expression has been carried out in *Xenopus leavis* oocyte by microinjections of misaminoacylated suppressor tRNA. As discussed earlier, SNAAR using cell-free protein biosynthesis system is faced with common problems such as low yields, tedious purification, protein folding and lack of post-translational modifications of the target protein. Similarly, *in vivo* SNAAR using *Xenopus* oocyte microinjections also suffers from low yields of target protein, requirement of an elaborate setup and specialized expertise. **Hence, this project is aimed at generalizing the methodology of SNAAR and applying it to other *in vivo* systems.**

4.2 Unique features and Novelty of our approach to SNAAR

The current work is aimed at simplifying SNAAR methodology and makes it a general method for incorporating non-native amino acids *in vivo*. This method has a number of

unique approaches for achieving this and the entire combination of simplified methods gives the novelty of this work.

To begin with, the approach of generating misaminoacylated suppressor tRNA needs to be simplified to a great extent since the chemical aminoacylation procedure is extrmeley tedious. In our approach for generation of misaminoacylated suppressor tRNA, suppressor tRNA is cloned into *E.coli* vectors and transcribed thereof, using *in vitro* transcription. The RNA isolated from the *in vitro* transcription reaction is primarily the suppressor tRNA of interest. Further, this tRNA is aminoacylated with cysteine and then subjected to post-aminoacylation modification. This is approach is simplified method for generating misaminoacylated tRNAs. By chosing cysteine as the modifying amino acid, we are introducing a large repertoire of chemical entities into proteins, e.g. nitrosylation reactions, reduction reactions etc.

The tRNA chosen in these examples is cysteine tRNA ($tRNA_{cys}$) and cysteine suppressor tRNA ($tRNA^{cys}_{CUA}$), since bacteriorhodopsin does not contain any cysteines and hence cysteine codons can be uniquely engineered into bacterio-opsin gene. However, the approach is not restricted and can be applied also to proteins containing cysteines, since cysteine suppressor tRNAs can be used for site specific delivery of non-native amino acid. However, the most significant advantage of using cysteinyl tRNA lies in the properties of cysteine modifications, most of which are based on the modification of the –SH side chain. This is advantageous since the tRNA skeletal structure does not contain any –SH groups and thus the modification of the –SH group remains restricted to only the cysteinyl amino acid. Post-aminoacylation reactions may not be possible when one uses suppressor tRNA of other amino acids such as glutamic acid, lysine, threonine since the tRNA skeletal backbone contains –COOH, $-NH_2$, -OH groups and these could also be subject to modifications leading to a completely new modified tRNA.

In addition, it is also now well documented that many *Archaea* bacteria possess only 16 of the 20 canonical AARSs. The absence of particular AARSs (such as AsnRS and GlnRS) is often compensated in indirect ways, that is, by recruitment of enzymes of

intermediary metabolism. Interestingly, *Halobacteria* are also devoid of CysRS, a very important criterion for the selection of *Halobacterial* cells as expression systems for the misaminoacylated cys-tRNA. This prevents the editing of the misaminoacylated tRNA within the *Halobacterial* cells. Furthermore, *Halobacteria* are much easy to cultivate and hence the expression of SNAAR analogs in *Halobacterium salinarium* as expression system could prove to be an interesting approach to generate SNAAR analog proteins.

In this study, we report lipofection as a method of *in vivo* introduction of misaminoacylated suppressor tRNA into large number of spheroplastic cells of *Halobacterium salinarium* so that SNAAR-analogs can be obtained in good yields. In order to perform SNAAR, we chose bacteriorhodopsin as the model protein. Bacteriorhodopsin (bR), an interesting photobiomaterial, is an extremely significant molecule due to its potential applications in biomolecular electronics (*375*).

This approach would help us in overcoming the limitations of the SNAAR technology in the following manner:

1. Higher yields: The main goal of this work is to obtain very good yields of SNAAR analogs. The thesis aims at making this a general technology for production of recombinant proteins containing non-native amino acids. Such novel proteins could be used and novel drugs. E.g. D-arginine containing insulin analogs which are resistant to gastric enzymes such as trypsin, carboxypeptidase etc. Such insulin SNAAR analog could be used to make an oral formulation making it convenient for administration.
2. Ease of Purification: Engineering affinity tags on the gene of interest or making the protein secretory can help us simplify the purification of SNAAR analogs to a large extent. Besides, since the expression is *in vivo*, one has to deal with fewer amount of impurities compared to the *in vitro* protein expression systems.
3. Protein Folding: The work reported here aims at producing SNAAR analogs *in vivo*. Since, SNAAR analogs are expressed in *Halobacterium salinarium* cells, the folding of the protein is taken care by the cellular machinery and only rightly folded functional SNAAR analogs are obtained.

4. Post-translational modifications of the target protein: If the SNAAR analogs are designed such that the expression system is native to the protein of interest, the SNAAR analogs that are generated would be post-translationally modified.

Thus, such a technology could become a powerful tool in generating novel protein analogs with enormous applications. This thesis aims at developing *in vivo* SNAAR as a general procedure that can be used routinely to generate novel proteins.

4.3 Study Design

The goal of the thesis is to achieve *in vivo* synthesis of SNAAR analogs so that the method can be simplified to such an extent that it can be performed in any laboratory. The most critical elements required for this work (see figure 4.1) include:

a. Target protein Gene construction: Gene construction and cloning of the target protein is the first requirement for obtaining SNAAR analogs. The target protein selected was bacteriorhodopsin. The *bop* gene and its amber mutants form one of the key critical components for completing this SNAAR work.
b. Suppressor tRNA: Gene construct for suppressor tRNA is the second critical component in this work. Synthesis of suppressor tRNA by *in vitro* transcription, its efficient aminoacylation and post-aminoacylation modification reactions become the important milestones to be achieved.
c. Co-injection of RNAs into cells: The goal of the thesis is to achieve SNAAR *in vivo*. In order to accomplish this, it is important that both misaminoacylated suppressor tRNA and amber codon containing bop mRNA be co-injected/ co-introduced into cells. The method of choice for co-introducing tRNA and mRNA was considered *via* lipofection.
d. SNAAR analog Protein Expression: The cells injected/ lipofected with the tRNA and mRNA would be channelized for protein expression. The expressed protein would have to be characterized for the presence of the non-native amino acid. The protein would be further characterized for it functions since the non-native amino acid should ideally not perturb the protein structure-function.

e. Biophysical study of SNAAR analog: The SNAAR analog would be studied for various parameters to obtain an in-depth understanding of structure-function relationship.

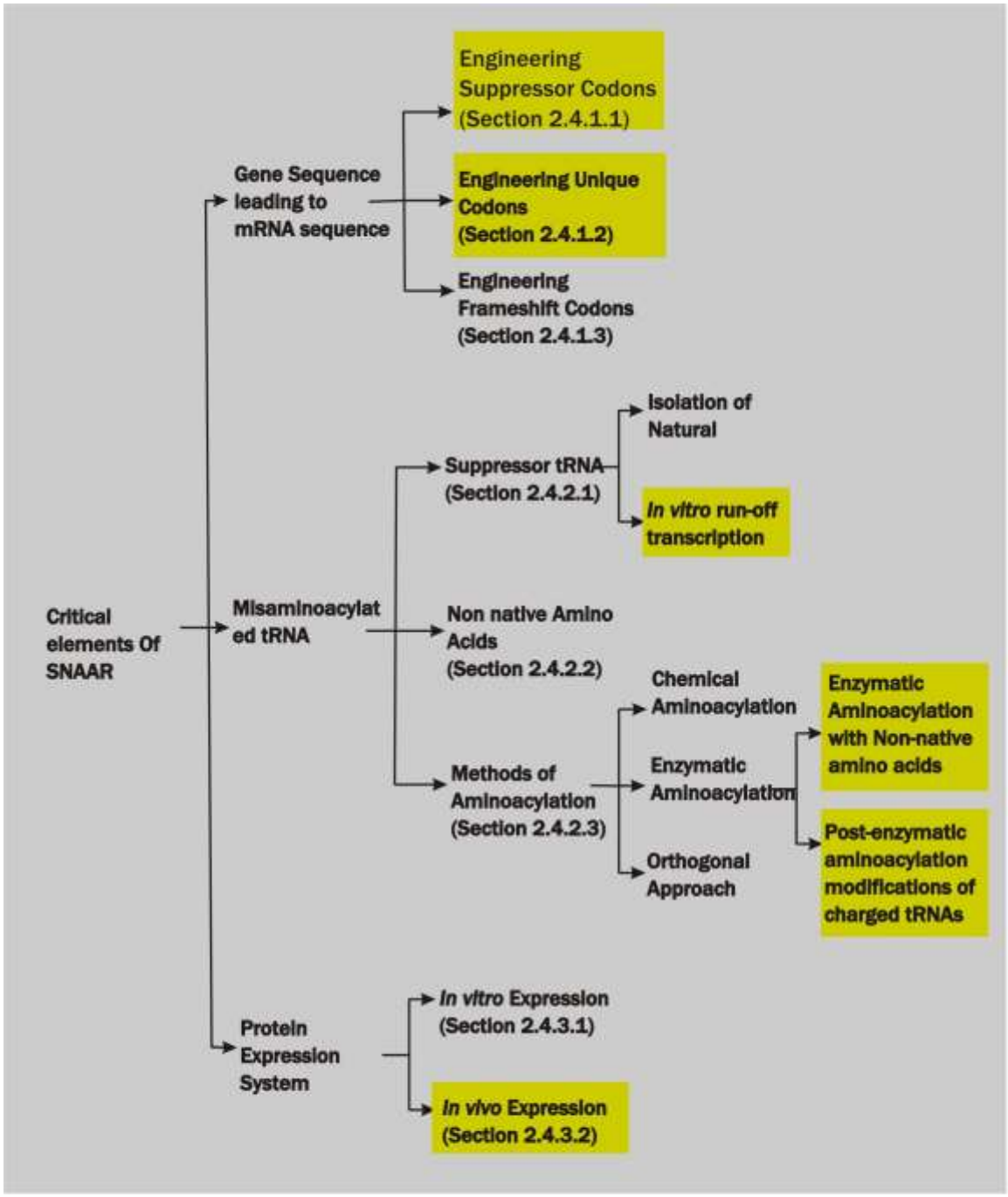


Figure 4.1: Schematic representation of key elements required for performing SNAAR. Yellow highlighted boxes indicate the approach taken in this work.

The target protein chosen, i.e. bacteriorhodopsin or bR has enormous applications in the biomolecular electronics and related sciences. The protein has one very important feature; it does not have any cysteines. Hence using cysteine mutagenesis followed by cysteine chemical modification of Cys-tRNA^{Cys}, would help us achieve the goal of site-directed amino acids replacement (SNAAR). In this case, we expect that modified cysteines (which are unnatural or non-native) would be incorporated in response to the engineered cysteine codons.

As discussed earlier, prokaryotic suppressor tRNAs are not recognized by the eukaryotic AARS synthetases and vice versa. Also important to note is that the bacteriorhodopsin is native to an archeabacteria called as *Halobacterium salinarium*. This archebacteria is very close to eukaryotes in the phylogenetic tree. Hence we decided to use *E.coli* suppressor tRNA in *Halobacterial* cells. This misaminoacylated *E.coli* suppressor tRNA would not be deacylated/recharged by any of the *Halobacterial* synthetases.

The overall scheme of the experiments is as shown in Figure 4.2. The suppressor tRNA would be misaminoacylated with the non-native amino acid and liposome encapsulated. The bop mRNA containing the cysteine mutation would also be liposome encapsulated. Together, both the liposomes containing the mRNA and tRNA would be co-lipofected into sheroplastic cells of Halobacterial cells. One very important point to be noted is that the *Halobacterial* strain chosen would have to be a mutant such that it does not have its indigenous capacity to produce bacteriorhodopsin. Such strain which lack the purple membrane (bR) are termed as Pum$^-$ strains.

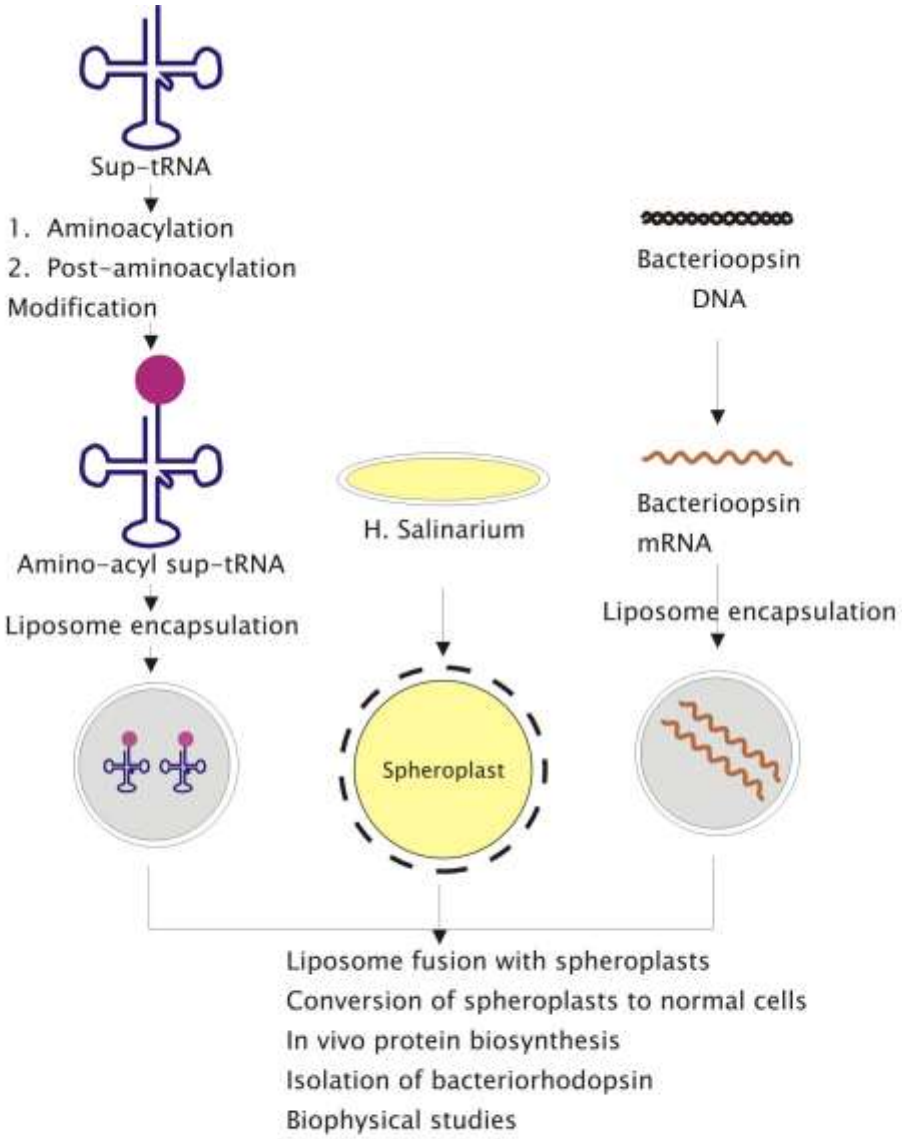


Figure 4.2: Schematic representation of the overall approach for achieving *in vivo* site-directed non-native amino acid replacement (SNAAR) in bacteriorhodopsin

4.4 Specific Aims of the Proposed Work

The detailed specific aims to achieve the above overall goal include:

1. Cloning of the *bop* gene into *E.coli* expression vector: The target protein gene, i.e. the *bop* gene is cloned into *E.coli* expression vectors. The protein is expressed, functionally refolded, and purified to obtain the functional protein.
2. Generation of amber and cysteine mutants of *bop* gene: Since bacteriorhodpsin protein does not contain any cysteines, cysteine codons can be engineered since they would be unique to the gene. Careful selection of amino acids to be mutated is made and codons corresponding to these amino acids are changed to cysteine codons. **Thus using amber mutagenesis as well as cysteine mutagenesis, two non-native amino acids could be simultaneously introduced into bR.**
3. Gene synthesis and cloning of *E.coli* cysteine-tRNA ($tRNA^{Cys}$) and cysteine tRNA suppressor ($tRNA^{Cys}_{UAG}$): *E.coli* cysteine tRNA and cysteine-suppressor tRNA are synthesized and cloned into an *E.coli* expression vector. This tRNA is generated by *in vitro* runoff transcription using RNA Polymerase enzyme. The tRNA so generated would be charged with cysteine using *E.coli* S100 extract. The cys-$tRNA^{cys}$ and the cys-$tRNA^{Cys}_{CUA}$ is subjected to cysteine chemical modifications with reactive groups such as fluorophores or chromophores.
4. Liposome encapsulations of RNAs: Both the RNAs, i.e. the misaminoacylated tRNAs and cysteine codon containing *bop* mRNA is encapsulated within liposomes. Liposomes are used to deliver the RNAs *in vivo* into the *Halobacterial* cells.
5. Lipofection of RNAs with *Halobacterial* spheroplasts: Liposome encapsulated tRNA and *bop* mRNA is lipofected with spheroplasts of Pum^- strains of *Halobacterium salinarium*.
6. Expression of the bOp SNAAR analog: The cells are then made to express the target protein by growing in nutrient rich media. The modified cysteines are incorporated site-specifically into bR in response to the engineered amber and cysteine codons.

7. Characterization of SNAAR analogs: The SNAAR analogs of bR, so obtained are then characterized for its function. Peptide mapping to prove the site-specific incorporation of modified cysteines is carried out.
8. Fluorescence studies: Further studies such as fluorescence measurements are carried out on the SNAAR analogs. FRET measurements are carried out at various stages of proton pumping of bR.

Chapter 5: Expression and purification of bR (wild type and mutants) and their functional refolding

5.1 Introduction

In vivo SNAAR requires the simultaneous introduction of amber codon carrying gene and misaminoacylated suppressor tRNA into cells so that a non-native amino acid is incorporated at the desired site when the cellular machinery translates the mRNA of the target protein of interest (See Figure 4.2). Thus, the first experimental goal was to introduce macromolecules such as nucleic acids (DNA and RNA) into cells. This is crucial as *in vivo* SNAAR requires the introduction of nucleic acids such as tRNA and mRNA in the ongoing protein synthesis of the target proteins.

Hence, the first work element of this thesis is the cloning of gene encoding the target protein, i.e. bacterio-opsin. The protein counter part, bacterio-opsin, is encoded as a gene in bacteria *Halobacterium salinarium*. This Halobacterial gene (*bop* gene) would ideally be cloned into *E.coli* vectors for the ease of cloning and expression. In order to efficiently express in *E.coli*, the gene needs to be optimized for codons in *E.coli*.

It is also important to establish that the cloned gene is completely identical to the wild type gene. Also it is extremely important to establish that the expressed protein is fully functional in terms of its activity. Hence it forms an important part of the thesis to express the protein, refold to its functional form and purify to a reasonable purity. As discussed in the previous chapter, bacteriorhodopsin is a light driven proton pump. The proton pumping activity of the refolded bacteriorhodopsin, thus, forms an important functional assay to check the activity of the protein.

In order to achieve SNAAR, one of the key components is the target protein gene containing the amber mutation at a specific site. In case of bR, since the cysteine codon is absent in the *bop* gene, a unique cysteine codon can be engineered site-specifically and used to generate double SNAAR analogs. In that case, an important part of the work

would be, to generate amber and cysteine mutants of *bop* gene at defined positions. Such engineered amber and cysteine codons would direct incorporation of the non-native cysteine in the protein.

From a detailed review of literature on structure-function relationship of bacteriorhodopsin (vide supra Chapter 3), it becomes clear that understanding of the exact nature of the light induced conformational change in bR is a key to understand not only bR but also several G-protein coupled receptors and retinal binding proteins. The role of C, F and G helices in conformational change has been indicated in some studies (See Chapter 3). With special focus on conformational changes within the helical structure of bR, key residues were identified for positioning amber and cysteine codons. In this thesis, we decided to elucidate this exact nature of conformational changes using SNAAR analogs (single and double) of bR and D96N mutant of bR. By using fluorescent analogs of cysteines, one could measure distances between two points within a protein. This could be done using **Fluorescence Resonance Energy Transfer or FRET**. Two amino acids on the end points of each helix (extracellular side and cytoplasmic side) were chosen as the site to incorporate the fluorescent cysteine analog (See Figure 5.1). Mutants of bR were generated for getting distance measurements on bR570 state. The D96N mutant of bR has a slow M decay and is known to have a prolonged M-state. Similar amber and cysteine mutants on end points of helices were generated on this D96N mutant to generate data on distance measurements in the M state. Table 5.1 elaborates the amino acids chosen on each of the seven transmembrane helices of bR.

Table 5.1 shows the amino acids chosen for amber and cysteine mutagenesis on the transmembrane C, F and G helices of bR

Helix	Cytoplasmic side	Extracellular side
C-helix	Valine-101	Tyrosine-79
G-helix	Leucine-224	Leucine-201
F-helix	Serine-169	Leucine-190

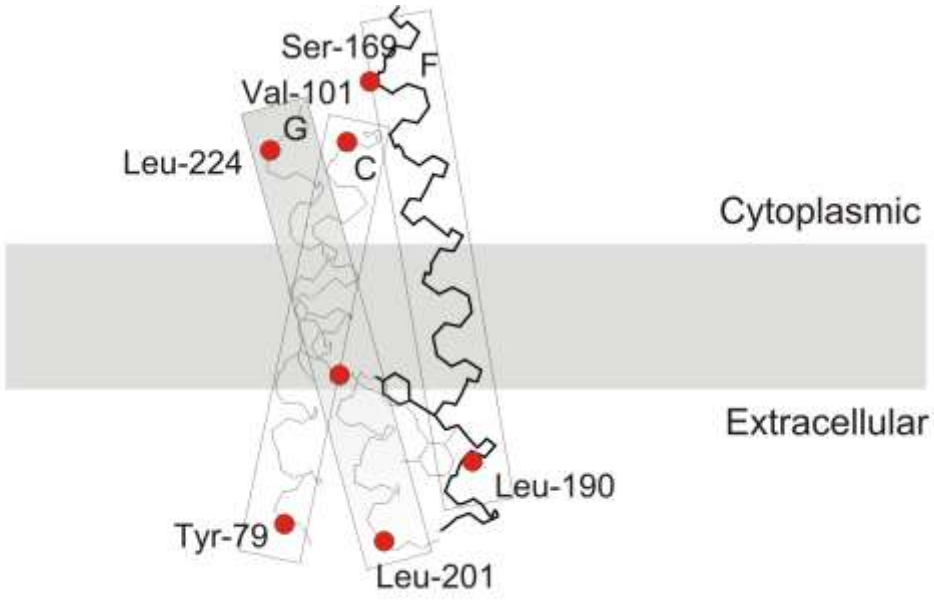


Figure 5.1: Schematic representation of the secondary structure of bR showing sites of modification. One amino acid was chosen at the cytoplasmic and extracellular end on C, F and G helix.

Various single and double mutants can be generated of the above amino acid residues. These mutants can be labeled with different fluorescent cysteine modifying agents (fluorescence donors and acceptor pairs) so that distance measurements can be carried out using FRET. Triple mutant also need to be generated as a part of this thesis work. Figure 5.2 a, b, c and d gives the location of the various mutations in the secondary structure of bacterio-opsin. Figure 5.2 a shows the single mutants generated for bacterio-opsin. Figure 5.2 b and c shows the double mutants generated as a part of this work. Figure 5.2 d shows the triple mutants that are generated. Table 5.2 given below lists the mutants of bR that have been generated as a part of this work.

Table 5.2: List of mutants of bR generated as a part of the thesis

Bacteriorhodopsin	**Type of Mutation**	**Mutation**	**Reference**
bR (M block)	Single mutant	D96N	Figure 5.2 a
bR (light adapted)	Single mutant	L201am L224am L190am S169am V101am Y79am	Figure 5.2 a
bR (M block)	Double mutants	D96N/L201am D96N/L224am D96N/L190am D96N/S169am D96N/V101am D96N/Y79am	Figure 5.2 b
bR (light adapted)	Double mutants	V101am/Y79C L224am/L201C S169am/L190C	Figure 5.2 c
bR (M block)	Triple mutants	D96N/V101am/Y79C D96N/L224am/L201C D96N/S169am/L190C	Figure 5.2 d

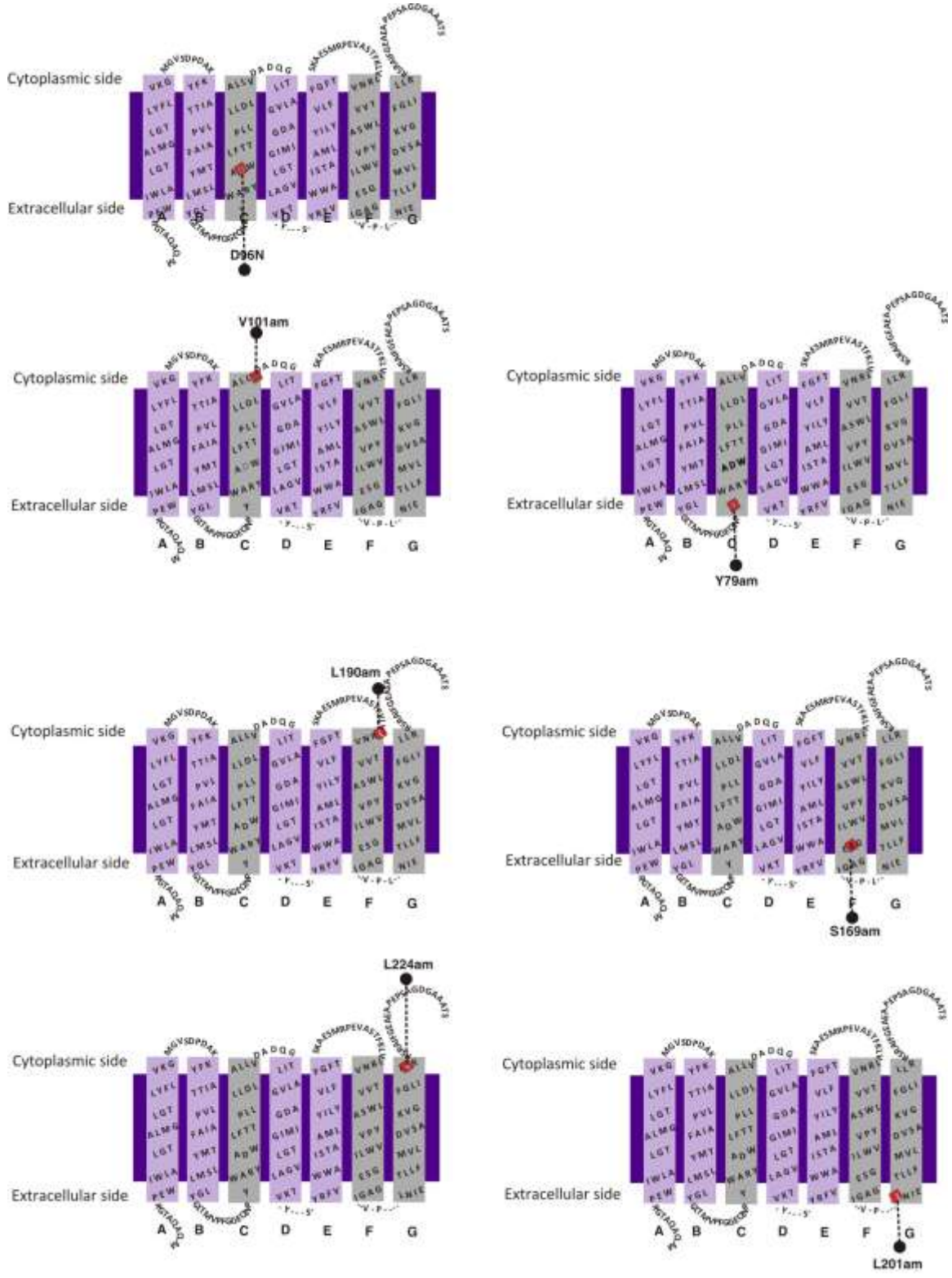


Figure 5.2 (a) Schematic Representation of Single Mutants of bR

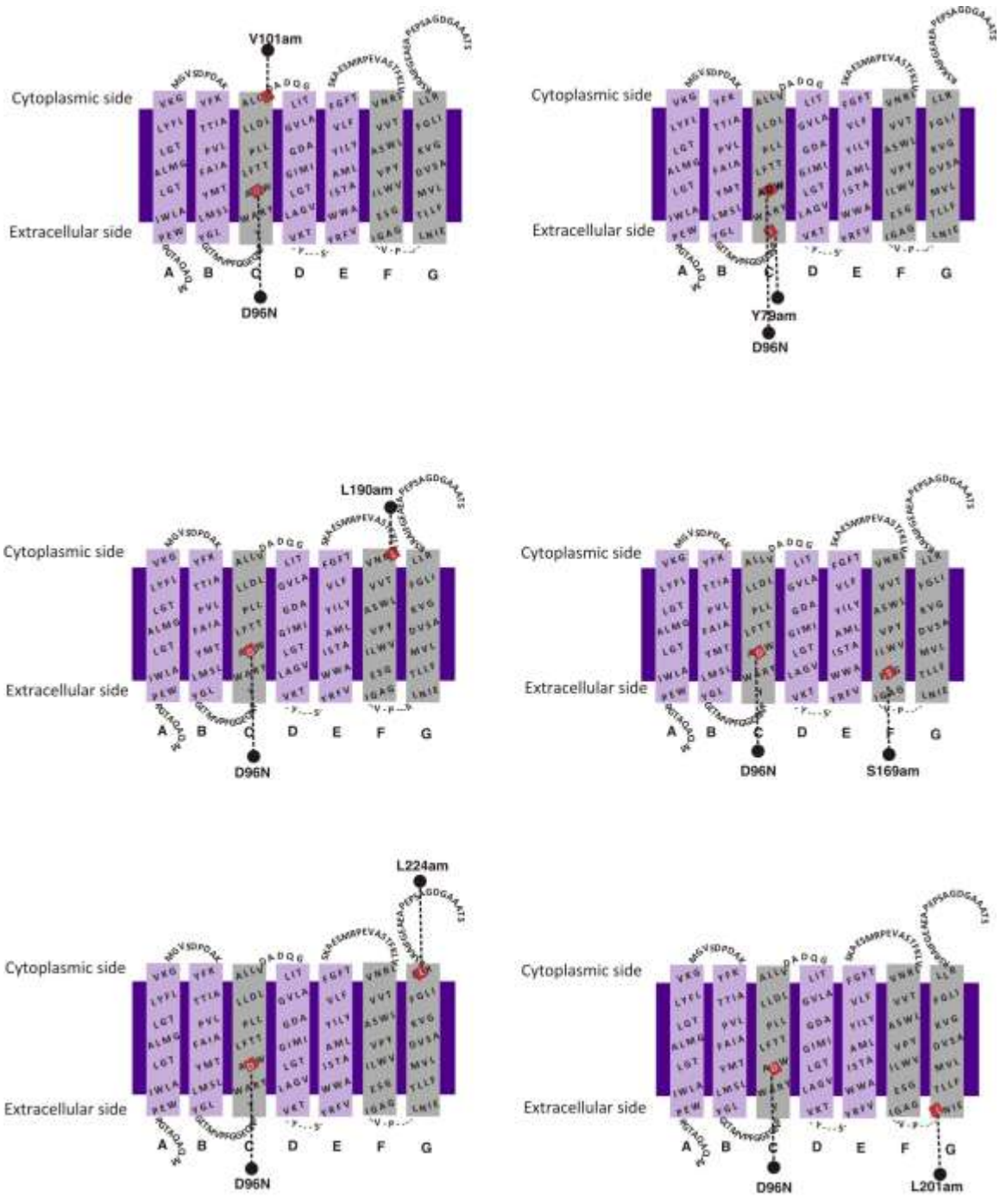


Figure 5.2 (b) Schematic Representation of Double Mutants of D96N bR (M-block mutants)

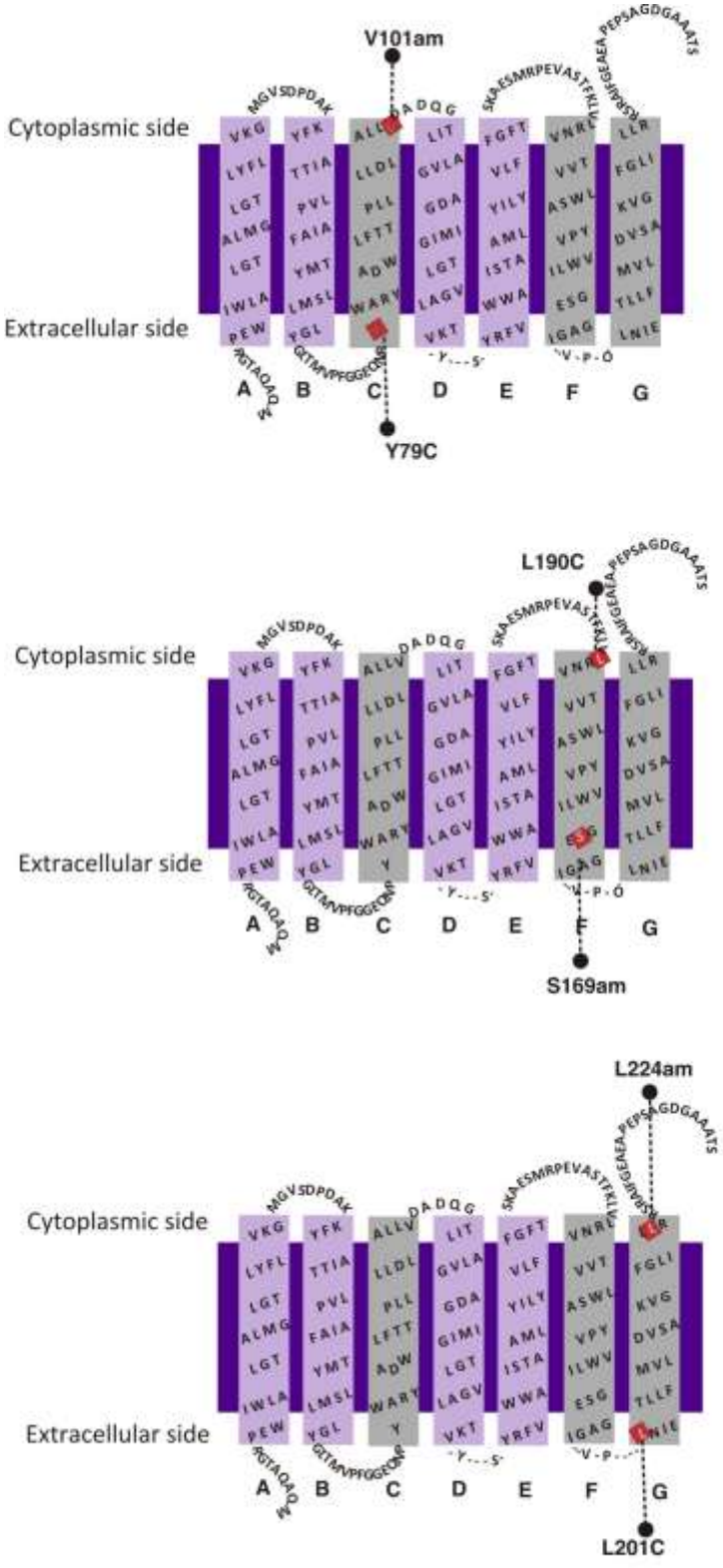


Figure 5.2 (c) Schematic Representation of Double Mutants of bR

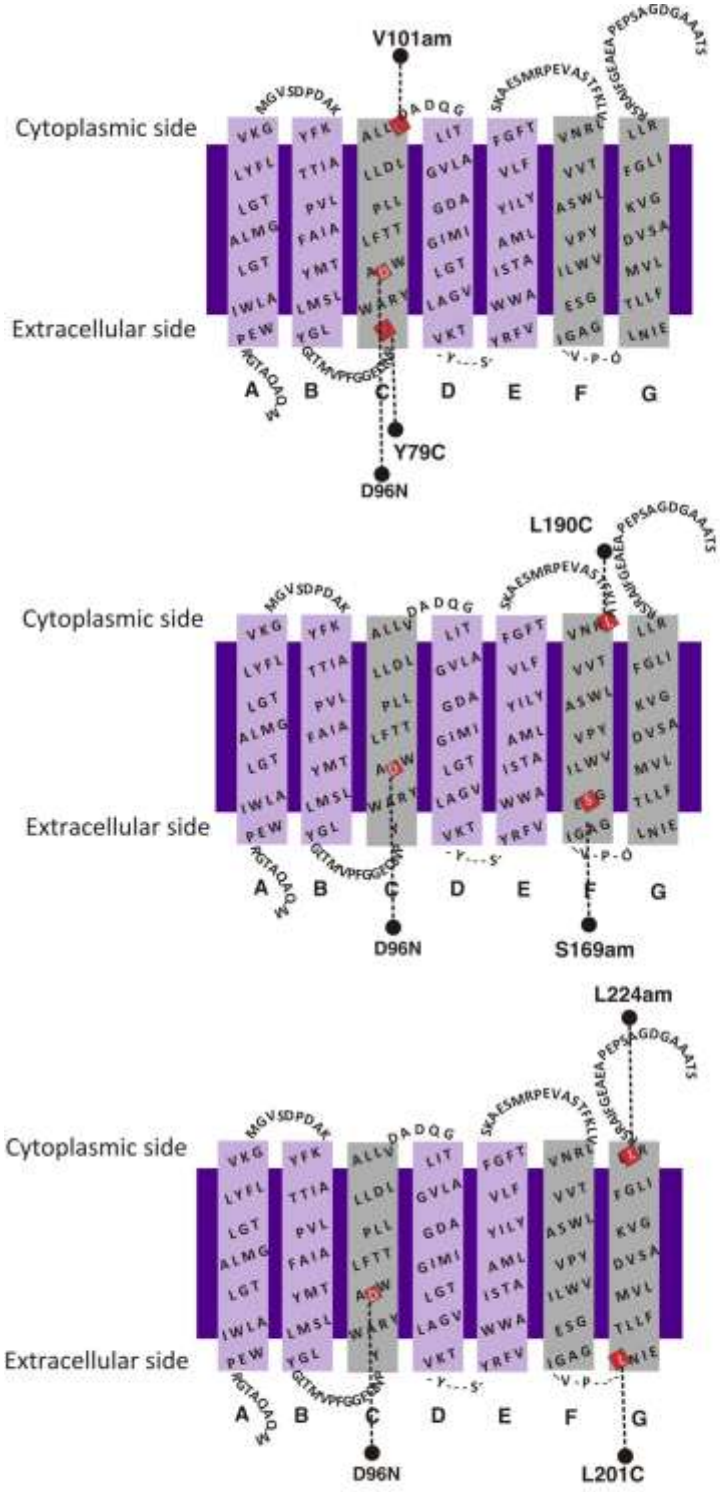


Figure 5.2 (d) Schematic Representation of Triple Mutants of D96NbR (M-block)

5.2 Methods and Materials

5.2.1 Materials

Synthetic oligonucleotides were custom synthesized from Genetix Associates (Mumbai, India). Restriction enzymes and other enzymes such as T7 RNA polymerase were purchased from Bangalore Genei (Bangalore, India). Other chemicals were obtained from Sigma. Bacto yeast extract and Bacto-tryptone were obtained from Difco. all-*trans*-Retinal was a kind gift from Dr. A. K. Singh (Chemistry Department, Indian Institute of Technology, Bombay, Mumbai, India).

5.2.2. Assembly of the bacterio-opsin (*bop*) gene

The bacterio-opsin gene was adapted as per Hackett *et al.* (*376*) wherein the gene was codon optimized for *E.coli* expression. Several restriction enzyme sites were also engineered into the gene in the process so as to facilitate further engineering and mutagenesis of the gene. The bacterio-opsin gene was synthetically constructed as represented in Figure 5.3. The entire gene was split into 14 oligonucleotides (Refer Table 5.3) and the gene was systematically constructed using the overlap PCR strategy as described in the following section. Figure 5.4 shows the gene sequence of bacterio-opsin aligned to its protein sequence.

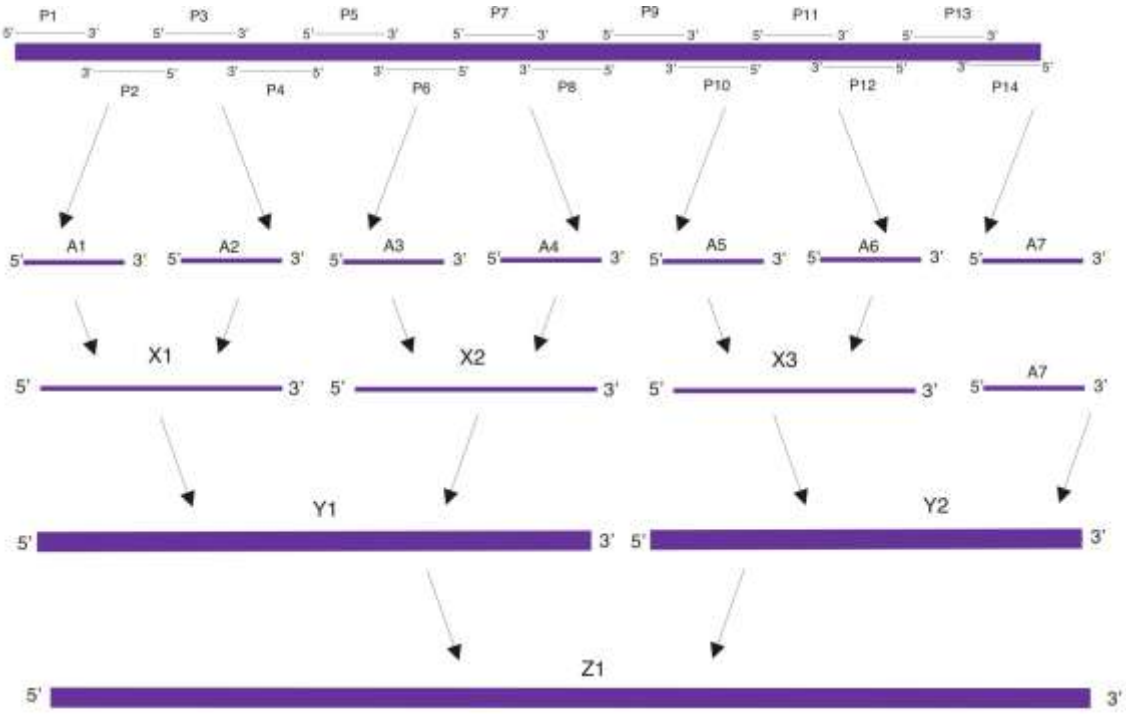
Table 5.3: List of Oligonucleotides for construction of the synthetic bacterio-opsin gene.

Sr. No.	Name	Sequence 5′→3′	No. of bases
1	P1	**AGCTTATGCA AGCTCAAATT ACTGGACGTC CGGAATGGAT CTGGCTAGCT CTGGGCACCG**	60
2	P2	**GGCCTAGGCT TTGTGGGTAT GGAAATTGGT CCTTCATGTC CCACGGGTCT GGGTAGTCTC GCCACGGGTC**	70
3	P3	**TCGGATCCGG ATGCGAAAAA ATTCTACGCT ATCACCACCC TGGTGCCGGC TATCGCATTC**	60
4	P4	**ATAGCGTAAG TGGTACATGG ACAGATACGA CGACCCAATG CCAGACTGGT ACCATGGCAA**	60
5	P5	**TGGTACCGTT CGGTGGTGAA CAGAACCCGA TCTACTGGGC CCGTTACGCT GACTGGCTGT**	60
6	P6	**ACCACGGGAC TAGTCGCAGT TGGTCGTCTC GGTCTAGATC GTCGTCGTCG CCCCACCACT TGTCGGTCAG**	60
7	P7	**CAGGGCACCA TCCTGGCTCT GGTTGGCGCC GACGGTATCA TGATCGGCAC CGGCCTGGTT**	60
8	P8	**TCGACGTCAT CTCTATCGGG TGGTTTGCTT TGCCATTCTC ATTTGGAACC AGTCGCGCGG TTGGTCCGGC**	70
9	P9	**TACTGCAGCT ATGCTGTACA TCCTGTACGT ACTGTTCTTC GGTTTCACCT CTAAAGCTGA**	60
10	P10	**ATTGCAATGC GTCATGAAAC TTCCAGCTGC GTTGAAGGCC TGCGTACGAA AGTCGAAATC**	60
11	P11	**CGTAACGTTA CCGTTGTTCT GTGGTCCGCT TACCCAGTTG TTTGGCTGAT CGGTTCTGAA**	60
12	P12	**TCTTTGCAGA TCTTGGTACT TGTCGTCCCA AAGTTATAAG TCGCCTTGTT ACGGCCGTGG AAGTCTTGGC**	70
13	P13	**AGACGTTTCT GCTAAAGTTG GTTTCGGTCT GATCCTGCTG CGTTCTCGAG CTATCTTCGG TGAAGCTGAA**	70
14	P14	**CGAGTAGTTC TCCATCGCCG GCGTGGCAGT GGGCGCCTGC CAAGGCCTCG AAGTCGAAGT**	60

Gene construction

To construct the synthetic bacterio-opsin gene from oligonucleotides, a PCR reaction was assembled with a set of oligonucleotides (i.e P1 and P2) with 10-15 complementary bases at 3' end. The 50 µl of PCR mix contained 1X PCR buffer, 2.5mM $MgCl_2$, 200mM dNTP each, 2.5units of Taq DNA polymerase and 25 pmoles of each oligonucleotide. The overlapping PCR reaction was carried out using the cycling conditions as follows; Initial denaturation at 96°C for 5minutes, followed by denaturation at 96°C of 20 sec, annealing at 60°C for 20 sec, extension at 72°C for 20 sec, for 35 cycles and final extension at 72°C for 5 minutes after the completion of 35 cycles. Seven PCR amplicons A1 to A7 of an average of 110-120 bp was obtained from seven pair of oligonucleotides using overlapping PCR method.

In the second set of PCRs, three amplicons X1, X2 & X3 were generated from the set of A1 & A2, A3 & A4, and A5 & A6 PCR amplicons respectively (Refer Figure 5.3). PCR reaction was set without forward and reverse oligonucleotides for first 10 cycles, to avoid nonspecific amplification and 25 pmoles of forward & reverse oligonucleotides were added to the PCR for the next 25 cycles. Cycling conditions were similar to the first PCR reaction. In the third set of PCR reaction, two amplicons Y1 and Y2 of 460 bp and 350 bp were obtained from overlapping X1and X2 amplicons and by overlapping of X3 and A7 amplicons.



Full length bacterio opsin gene (Final Amplicon)

Figure 5.3: Schematic representation of the assembly of synthetic bacterio-opsin gene

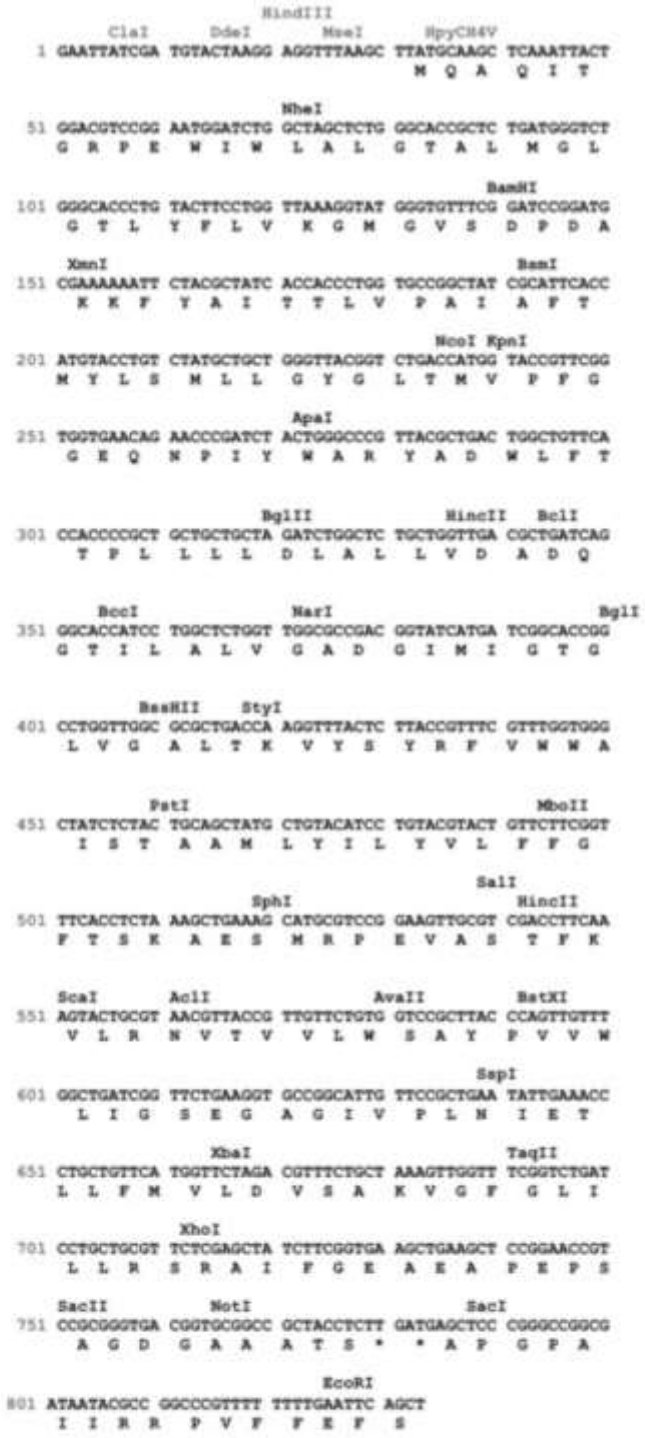


Figure 5.4: Sequence of the bacterio-opsin (*bop*) gene. The bacterio-opsin *(bop)* gene sequence with the amino acid sequence is shown in this figure. The restriction sites within the gene are also depicted.

The final 810 bp bacterio-opsin gene was obtained by overlapping 460 bp Y1 and 350 bp Y2 PCR amplicons. The PCR for final gene amplification was carried out without forward and reverse primers for 10 cycles and 25 pmoles of the forward P1 and reverse P14 was added for the next 25 cycles. The denaturation was at 96°C of 30 sec, annealing was at 60°C of 30 sec and extension was at 72°C of 45 sec. Final extension at 72°C for 5minutes was allowed for the complete synthesis of the gene.

Amplification of the Gene: To obtain a synthetic bacterio-opsin gene with the desired restriction ends, a PCR reaction was assembled with primers containing the restriction sites, forward primer containing *EcoR*I site and reverse primer containing *Hind*III site. Two forward primers were designed; BOPL+ (with the leader sequence) and BOPL- (without the leader sequence) for amplification of the bacterio-opsin gene.

The PCR mix contained 1X Taq buffer, 2.5mM $MgCl_2$, 200mM dNTP, 2.5 ng of F1 amplicon as template, 25 pmoles of forward and reverse primers and 2.5 units Taq DNA polymerase. The PCR conditions were 94°C for 5 min for initial denaturation, 10 cycles of 94°C for 20 sec denaturing, 60°C for 20 sec annealing and 72°C for 20 sec extension and 72°C for 5 min for final extension.

5.2.3. Construction of pGEM-SBOP-L+ and pGEM-SBOP-L-

The gene encoding for bacterio-opsin (*bop*) was synthesized as described in the previous section. The gene was optimized for expression in *E.coli*. The gene was designed so as to have the following elements in its sequence.

i) Signal sequence of the *bop* gene
ii) Ribosomal binding site
iii) T7 promoter

Genes encoding bacterio-opsin with and without leader sequences were synthesized using synthetic oligonucleotides and were assembled as shown in Figure 5.5. The 5'-sequence contains Shine-Dalgarno for *E.coli* expression (*377*). The leader sequence was

designed as described (*378*). The remaining *bop* sequence was adapted from Khorana and co-workers (*379*) with a minor modification of replacing the 3′-terminal *EcoR*I site

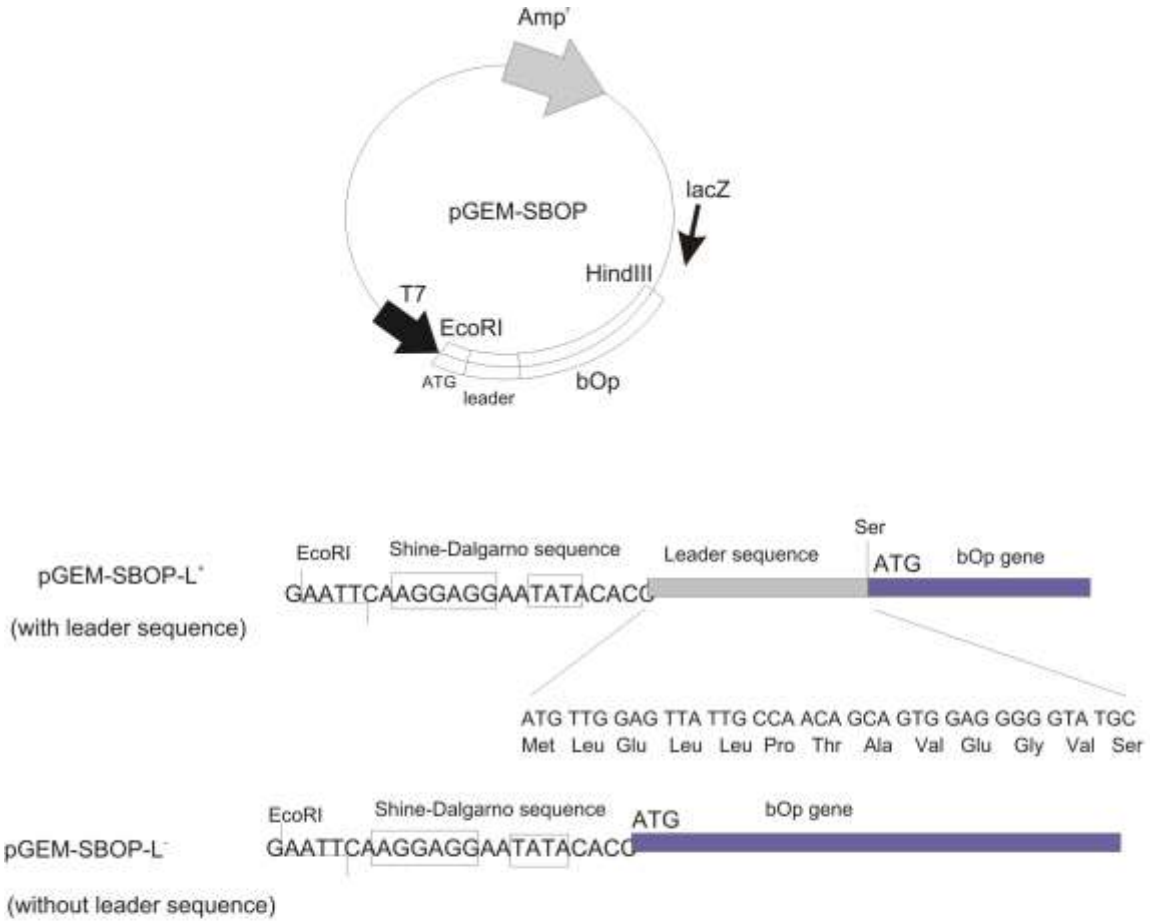


Figure 5.5: Construction of *bop* genes for expression of bR and for the preparation of bop mRNA. *bop* gene was amplified with two forward primers; one containing the EcoRI site alongwith a leader sequence and one with only the EcoRI site. The reverse primer was designed to contain a HindIII site. The amplicon obtained was subcloned cloned into pGEM3(z) in *EcoR*I / *Hind*III sites. The vector was linearised with NheI for *in vitro* transcription.

by *Hind*III. Both genes were cloned into pGEM3-(Z) vector (Promega, Madison, WI) at *Eco*RI-*Hind*III sites.

PCR amplicons obtained were digested with the restriction enzymes *EcoR*I and *Hind*III. The reaction assembled contained the 1 μg of PCR *bop* amplicon along with the RE buffer and 10 units each of *EcoR*I and *Hind*III. The reaction mixture was incubated at 37°C for 2 hours in a water batch. The vector pGEM-3z was also digested with *EcoR*I and *Hind*III. The *bop* genes (with and without leader sequence) were ligated with pGEM-3z in a ligation reaction with vector: insert ratio of 1:3 and 1:5. 3 μl of the ligation mixtures were independently transformed into competent cells of *E.coli* DH5α and screened for ampilcillin resistance. Transformants obtained were screened for the presence of plasmids. Plasmids were prepared from transformants and subjected to DNA sequencing. The DNA was sequenced on a Beckmann CEQ8000. The plasmid was transformed into *E.coli* BL-21 cells for expression.

5.2.2. Construction of amber and cysteine mutants of bR

The oligonucleotides for the construction of the amber and cysteine mutants (Refer Table 5.3) were custom synthesized from Sigma Aldrich, India. The *bop* gene has been codon optimized for *E.coli*, and has several restriction sites engineered into the gene (Refer Figure 5.4 for *bop* gene sequence). Mutagenesis of the *bop* gene was carried out using cassette mutagenesis. Oligonucleotide stretches between various restriction sites were synthetically prepared with the desired mutation and annealed. This annealed cassette is ligated with the RE digested pGEM-S-BOP-L+ (Refer to figure 5.6 for the cloning strategy). The primers designed for various amber and cysteine mutants of bR is listed in Table 5.4.

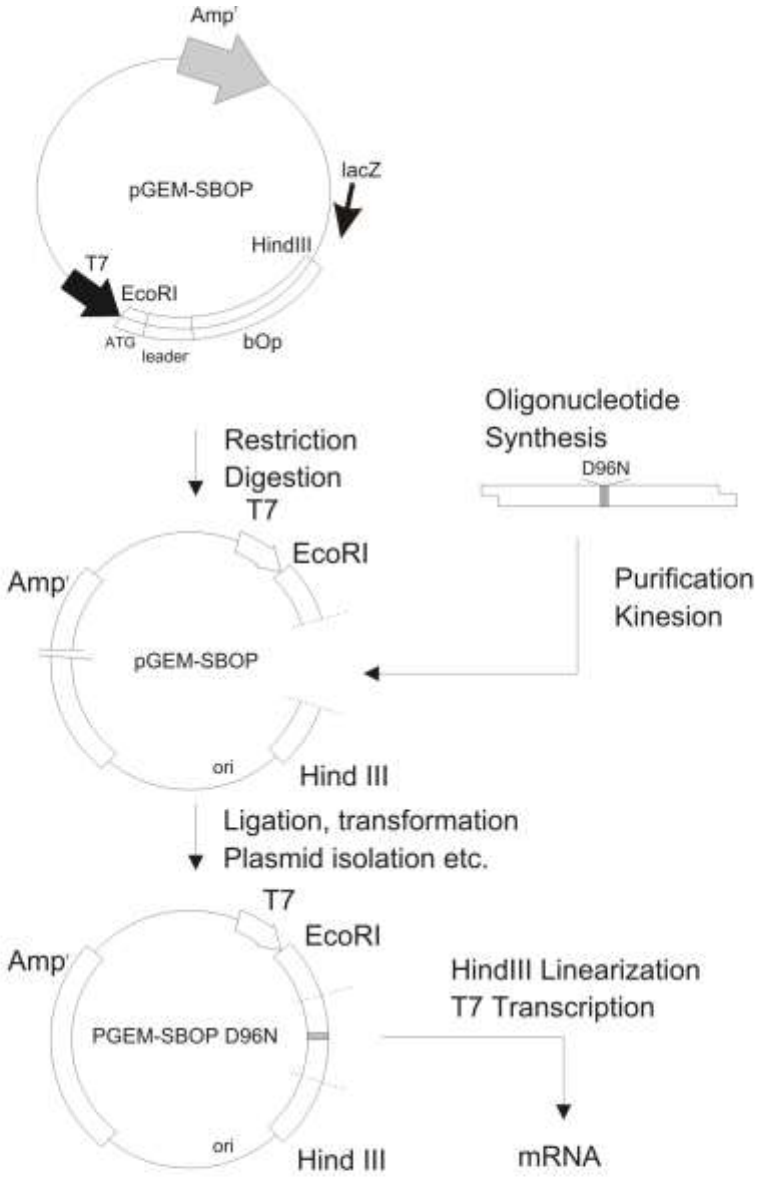


Figure 5.6: Cassette mutagenesis of the *bop* gene to generate mutants. For e.g. for generating amber mutants at position Leu-224, WT bop gene in pGEM-S-BOP-L+ was cut with restriction enzymes XbaI and XhoI. The insert containing the amber mutation at Leu-224 was prepared as synthetic oligonucleotides and the cassette was prepared by phosphorylation and annealing. The annealed cassette was ligated with the large fragment of pGEM-S-BOP-L+ to generate the *bop* gene with L224am mutation.

Table 5.4: List and sequences of primers custom synthesized for various amber and cysteine mutants of bR

Sr. No	Mutation	Primer sequence	Restriction Enzyme used for cloning
1	D96N	Sense strand 5′…CGTTACGCTGACTGGCTGTTCACCACCCCGCTGCTGCTGCTAAATCTGGCTCTGCTGGTT-3′ Antisense strand 5′…AACCAGCAGAGCCAGATTTAGCAGCAGCAGCGGGGTGGTGAACAGCCAGTCAGCGTAACGGGCC-3′	*Apa*I/*Hinc*II
2	L201am	Sense strand 5′…TTGGCTGATCGGTTCTGAAGGTGCCGGCATTGTTCCGTAGAAT-3′ Antisense strand 5′ATTCTACGGAACAATGCCGGCACCTTCAGAACCGATCAGCCAAACAA3′	*Bst*XI/*Ssp*I
3	L224am	Sense strand 5′CTAGACGTTTCTGCTAAAGTTGGTTTCGGTCTGATCCTGTAGCGTTC-3′ Antisense strand 5′TCGAGAACGCTACAGGATCAGACCGAAACCAACTTTAGCAGAAACGT-3′	*Xba*I/*Xho*I
4	L190am	Sense strand 5′…TTGGTAGATCGGTTCTGAAGGTGCCGGCATTGTTCCGCTGAAT-3′ Antisense strand 5′ATTCAGCGGAACAATGCCGGCACCTTCAGAACCGATCTACCAAACAA3′	*Bst*XI/*Ssp*I
5	S169am	Sense strand 5′-CGTCCGGAAGTTGCGTAGACCTTCAAAGT-3′ Antisense strand 5′-ACTTTGAAGGTCTACGCAACTTCCGGACGCATG-3′	*Sph*I/*Sca*I
6	V101am	Sense strand 5′-GATCTGGCTCTGCTGTAGGACGCT-3′ Antisense strand 5′-GATCAGCGTCCTACAGCAGAGCCA-3′	*Bgl*II/*Bcl*I
7	Y79am	Sense strand 5′-CGTTCGGTGGTGAACAGAACCCGATCTAGTGGGCC-3′ Antisense strand 5′-CACTAGATCGGGTTCTGTTCACCACCGAACGGTAC-3′	*Kpn*I/*Apa*I
8	Y79C	Sense strand 5′-CGTTCGGTGGTGAACAGAACCCGATCTGTTGGGCC-3′ Antisense strand 5′-CAACAGATCGGGTTCTGTTCACCACCGAACGGTAC-3′	*Kpn*I/*Apa*I
9	L201C	Sense strand	*Bst*XI/*Ssp*I

		5′…TTGGCTGATCGGTTCTGAAGGTGCCGGCATTGTTCCGTGTAA T-3′ Antisense strand **5′ATTACACGGAACAATGCCGGCACCTTCAGAACCGATCAGCCAA ACAA3′**	
10	L190C	Sense strand **5′…TTGGTGTATCGGTTCTGAAGGTGCCGGCATTGTTCCGCTGAA T-3′** Antisense strand **5′ATTCAGCGGAACAATGCCGGCACCTTCAGAACCGATACACCAA ACAA3′**	*Bst*XI/*Ssp*I

5.2.2.1 Cloning of amber and cysteine codon containing cassettes into pGEM-S-BOP-L+

The oligonucleotides pairs for the construction of the amber and cysteine mutants (Refer Table 5.4) were first annealed to form a double stranded cassette. The oligonucleotides were first heated separately to 95°C for about 30 mins. The two pairs of oligos were then mixed in equimolar ratio and gradually cooled to room temperature over 2 hours. This annealed cassette contains sticky ends of the desired restriction enzymes. The vector pGEM-S-BOP-L+ was then digested with restriction enzymes as shown in Table 5.4. The digested plasmid was run on a 0.7% agarose gel. The larger fragment of the plasmid was cut from the gel and further purified using QIAEX gel extraction kit. The purified vector fragment was quantitated and ligated with the respective cassette in a vector: insert ratio of 1:5.

The ligation mix was transformed into *E.coli* DH5α cells as per Maniatis et al (*380*). The transformants were selected on Luria-agar plates containing ampicillin. The transformants were picked from the plates and grown overnight at 37°C in Luria broth containing ampicillin. Plasmids were prepared from the growth for screening of the *bop* gene insert (*380*). Some of the growth was used to prepare glycerol stocks and stored at -80°C for further use. The plasmids prepared from each of the clones were digested with *Eco*RI/*Hind*III to throw out the *bop* cassette of 854 bp. The RE digests were run on a 1% agarose gel to confirm the size of the mutant cassettes in comparison with the control *bop* gene.

Table 5.5: Double mutants of bacterio-opsin gene

bR (M block)	Double mutants	D96N/L201am D96N/L224am D96N/L190am D96N/S169am D96N/V101am D96N/Y79am
bR (light adapted)	Double mutants	V101am/Y79C L224am/L201C S169am/L190C

For preparation of the D96N amber mutants, the plasmid of D96N was digested with the appropriate restriction enzyme. The cassette containing the desired amber mutation was generated in a manner similar to that described above. This cassette was ligated with the D96N plasmid and transformed into *E.coli* DH5α.

For generation of double mutants of amber and cysteine codons, plasmids of V101am, L224am and S169am were digested with the appropriate restriction enzymes. The cassette containing the cysteine mutations, e.g. Y79C, L201C and L190C, was ligated into the digested plasmids in the same order respectively. These ligated plasmids were transformed to obtain double mutants containing amber as well as cysteine mutations.

Table 5.6: Triple mutants of bacterio-opsin gene

bR (M block)	Triple mutants	D96N/V101am/Y79C D96N/L224am/L201C D96N/S169am/L190C

For generation of triple mutants, plasmids were prepared from the previously made double mutants, e.g. D96N/V101am, D96N/L224am and D96N/S169am, and digested with the appropriate enzyme listed in table 6. The cysteine mutation containing cassettes were generated by heating and annealing to give Y79C, L201C and L190C cassettes. These were ligated into the digested plasmids in the same order respectively to generate triple mutants of the *bop* gene.

5.2.2.3 Confirmation of *bop* mutant clones

The plasmids were subjected to DNA sequencing using a T7 promoter primer. The plasmids were purified and subjected to cycle sequencing using the mix provided by Beckmann Coulter. The samples were run on an 8-capillary Beckmann CEQ8000 machine. The data obtained was further analyzed and was aligned with the *bop* gene to screen for the desired mutation.

5.2.3. Expression and purification of bacterio-opsin protein in *E.coli* and its functional refolding into bacteriorhodopsin

All-trans retinal, RNAse A, DNAse I, phenylmethysulfonyl fluoride, Sephadex LH-60, and electrophoresis-grade SDS were obtained from Sigma. Soybean lipids (L-α-phosphatidylcholine, Type II-S) were obtained from Sigma and were purified by twice precipitating them from ether solution with acetone (*381*) using 50 μM dithiothreitol as an antioxidant. Octyl glucoside, DMPC and CHAPS were from Fluka. DEAE-Trisacryl was from Pall Lifesciences; and hydroxylapatite was from Bio-Rad. Chloroform and methanol were from SRL. Water was purified using a commercial deionizer. Other solvents are reagent grade.

Buffers- Buffer A contained 0.15M NaCl, 0.01 M NaPi, pH 7.0, 0.02% (w/v) NaN_3; Buffer B contained 0.1 NaPi, pH 6.0, 0.025% NaN3.

Organic Solvent Mixtures: Solvent A contained chloroform/methanol/water/TEA (100:100:35:1); Solvent B was solvent A without water (chloroform/methanol/TEA (100:100:1)); Solvent C was chloroform/methanol/acetic acid (100:100:1); Solvent D was chloroform/methanol/water/acetic acid (100:100:25:4); and Solvent E was chloroform/methanol/water (4:4:1). TEA acetate solutions of various concentrations were prepared by adding equimolar amounts of TEA and acetic acid to Solvent E.

The solvent compositions given above are volume ratios. Except where mentioned, all operations using these organic solvents were carried out in clean, dry glassware with Teflon fittings.

5.2.3.1 Expression of bacterio-opsin (bOp) protein in *E.coli*

Expression of the bacterio-opsin protein was carried out by first transforming the plasmids pGEM-S-BOP-L+ and pGEM-S-BOP-L- into *E.coli* BL-21(DE3) cells. For overexpressing the protein, the cells containing the recombinant plasmid were grown in a three-liter culture till mid-log phase (ie. OD 0.6). These cells were then induced with 1 mM IPTG. The cells were grown at 37°C for 5 hours post induction, after which they were harvested and then frozen. Thawed cells were resuspended in phosphate buffered saline containing $MgCl_2$ (1mM), $CaCl_2$ (0.1mM), DNAse (20 μg/ml), RNAse (20 μg/ml), and the protease inhibitor PMSF (200 μg/ml) and were lysed using a sonicator. Membranes were pelleted by centrifugation (60 mins, 40,000g, 4°C). Pellets of whole cells or membranes were stored at -80°C (*108*, *382*).

5.2.3.2 Isolation and regeneration of the bacterio-opsin protein

Purification and regeneration of bacterioopsin

- *Organic Solvent extraction of e-bO from E.coli membranes*

Lipids and highly lipophilic proteins, including denatured e-bO, were extracted from *E.coli* membranes or whole cells in Solvent A. Since solvent A contains saturating amount of water and therefore separates into two phases with an incremental addition of water, the wet pellet was mixed initially with water-free solvent (solvent B) until a single liquid phase was obtained. After centrifugation (20 mins, 8,000g, room temperature), the clear solution was decanted and the pellet was re-extracted three times with Solvent A, each time with 40 ml. The solvent extracts were combined (total volume 120 ml), and then a phase separation was effected by the addition of 220ml of Buffer B. After centrifugation (10 min, 8000g, 4°C), the liquid phases were both removed by decantation, and the solid interphase , which adhered to the walls, was washed with distilled water (2 x 100 ml). The wet pellet was stored at -20°C.

- *Hydoxylapaytite adsorption Chromatography in Organic solvent*

Dry hydroxyapatite (0.5g) was washed with solvent D (2 x 5 ml) and then with Solvent C (2 x 5 ml). Lipid depleted extract of *E.coli* membranes in solvent A (5 ml, 2-5 mg of protein/ml) was mixed with 2 volumes of the acidic solvent (Solvent C). A flocculent precipitate which formed was removed by centrifugation. The supernatant solution was mixed with half of the washed hydroxylapatite (0.25 g, dry weight). The hydroxylapatite suspension was poured into a column stoppered with glass wool, on top of a previously poured layer of the other half of the washed hydroxylapatite. After the loading solvent had flowed through the column, the column was washed with 5 ml of Solvent C and then with 5 ml of Solvent D. Finally, bOp was eluted with 10 ml of Solvent D to which water had been added to the solubility limit (15% by volume). Fractions of 0.5 ml were collected; those containing protein (as monitored by 280 nm absorbance) were pooled. Addition of 3 volumes of 0.2 M Na_2HPO_4 resulted in a phase separation, and bOp was recovered as a waxy solid at the interphase. The bOp was washed two times with water and then redissolved by adding 1 ml of solvent B (*382*).

- *Ion Exchange Chromatography on DEAE-Trisacryl in organic solvent*

DEAE Trisacryl was washed successively with 4 volumes each of water, methanol, and Solvent E. It was then packed into a 0.9 x 10 cm column. The column was equilibrated in Solvent E containing 50 mM TEA acetate at a flow rate of 15 ml/h. Lipid depleted extract of E.coli membranes in Solvent A was brought to 50 mM TEA acetate by the addition of TEA and acetic acid. The flocculent precipitate which formed was removed by centrifugation, and then the sample (5 ml, containing 2-5 mg/ml protein) was loaded on the column. The column was washed with 50 mM TEA acetate in Solvent E (~2 column volumes) until the 280 nm absorbance of the effluent reached a flat baseline. Elution was started using a linear gradient in TEA-acetate concentration (50-200 mM, using 2 column volumes each of initial and final concentrations in a glass gradient mixer). Fractions of 2 ml were colleced and analyzed by SDS gel electrophoresis. The fractions containing bOp were pooled. bOp was recovered by phase separation and redissolved in Solvent B (*382*).

- *Transfer of Proteins from organic solvent to SDS Solution*

To crude membrane extract or pure bOp in Solvent A (2-5 mg/ml, 1 ml) was added 100-250 µl of aqueous SDS solution (10% w/v). The exact amount of SDS was chosen to give a final SDS:protein ration of 5:1 (w/w). If a phase separation occurred, Solvent B was added until the solution was homogenous. The solvent was then evaporated, first in a vacuum centrifuge and subsequently in a lyophilizer, making sure that the vacuum reached 100 millitorrs for at least 1 hr. To the dry powder was added water to give a final SDS concentration of 0.5% (protein concentration, ~ 1 mg/ml). After vortexing gently to dissolve the SDS and protein, insoluble material was removed by centrifugation (*382*).

- *Renaturation and Chromophore Regeneration of bR*

bOp was renatured using procedures similar to those described previously (*383-385*). Typically, 0.2-0.4 ml of an aqueous SDS solution of bOp (0.5% SDS, protein concentration, 1 mg/ml) was added to 0.5 ml of Buffer B containing DMPC and CHAPS, each at a concentration of 2% (w/v) and the volume was brought to 1.0 ml by the addition of water. The actual protein concentration was determined precisely by measuring the UV absorbance at 280 nm ($\varepsilon 280 = 6.6 \times 10^4$ $cm^{-1}M^{-1}$) (*382*). Retinal (1.5 eq) in 10 µl ethanol was added, and the solution was kept overnight at room temperature.

5.2.4. UV-visible Absorbance and Light Dark Adaptation

Spectra were obtained on a Beckman DU-7 spectrophotmeter in 1-cm path length quartz cells. For dark adaptation, bR was kept in the dark at 25°C for 24 hours prior to obtaining the spectrum. Subsequently, light adaptation was accomplished by illumination (10 min, 20°C) with the 300-watt quartz halogen output of a slide projector using an orange filter. The spectrum was obtained within 2 min after removal of the sample from the illuminating light.

5.2.5. Proton pumping assay to check the functional refolding of bOp → bR

bR was renatured as described above except that the phosphate buffer concentration in the renaturation/regeneration mixture was decreased to 1 mM (pH 6.0). Also, for all bR

samples used in proton pumping measurements, the SDS concentration in the regeneration mixture was fixed at 0.2% following 24 h of incubation of the mixture, the concentration of the renatured bR was determined spectroscopically ($\varepsilon 280 = 5.2 \times 10^4$ $cm^{-1}M^{-1}$) (*382*, *386*). The mixture was then reconstituted into vesicles by using a modification of the detergent dilution procedure of Racker *et al* (*382*, *386*). This involved preparation of asolectin lipid vesicles (L-α-phosphatidylcholine, Type II-S, Sigma) by hydrating a film of the lipid (dried from a chloroform solution) with water and sonicating to clarity for 15-30 min under argon. Insoluble material was removed by centrifugation. To 30 ul of asolectin vesicles, 10 ul of 12.5% (w/v) octyl glucoside with 40 ul of WT bR was added and the solution was incubated on ice under argon for 5 mins. This 80 ul sample was then injected into 2 ml argon purged 2 M NaCl in a temperature-controlled reaction vessel equipped with a stirring bar and a micro-combination pH probe. After baseline stabilization (~ 10 mins), the sample was irradiated with a 300-watt projector lamp focused on the cell through a heat absorbing filter and a 435-nm long pass filter at saturating light intensity. Three light dependent proton uptakes were recorded with recovery intervals 3 mins. The system was calibrated with 2-5 ul injections of 1 mM HCl. Average values from 4 independently reconstituted vesicle preparations were calculated.

5.2 Results and Discussion

5.3.1. *bop* gene and its amber mutants were successfully cloned into *E.coli*

The bacterio-opsin gene was successfully synthesized synthetically using the overlap PCR strategy as described previously. Figure 5.7 shows the systematic assembly of the *bop* gene. Seven PCR amplicons A1 to A7 of an average of 110-120 bp was obtained from seven pair of oligonucleotides using overlapping PCR method. In the second set of PCRs, three amplicons X1, X2 & X3 were generated from the set of A1 & A2, A3 & A4, and A5 & A6 PCR amplicons respectively. In the third set of PCR reaction, two amplicons Y1 and Y2 of 460 bp and 350 bp were obtained from overlapping X1and X2 amplicons and by overlapping of X3 and A7 amplicons. The final 810 bp bacterio-opsin gene was obtained by overlapping 460 bp Y1 and 350 bp Y2 PCR amplicons. Finally, the

bop gene was amplified using synthetic primers designed for expression, i.e. contaiing the SD sequence and the leader peptide. Several clones of bacteriorhodopsin wild type

(A) **PCR I**

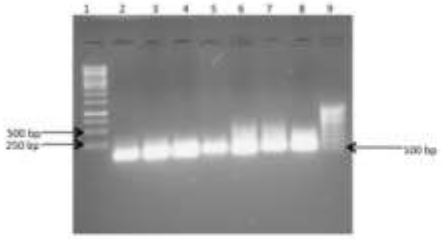


Figure 1: Agarose gel
Lane 1: 1Kb DNA Ladder
Lane 2: A1 amplicon
Lane 3: A2 amplicon
Lane 4: A3 amplicon
Lane 5: A4 amplicon
Lane 6: A5 amplicon
Lane 7: A6 amplicon
Lane 8: A7 amplicon
Lane 9: 100bp DNA Ladder

Table1: PCR amplicon size

Amplicon	Size
A1	120bp
A2	110bp
A3	110bp
A4	120bp
A5	110bp
A6	120bp
A7	120bp

(B) **PCR II**

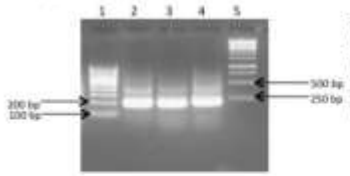


Figure 2: Agarose gel
Lane 1:100bp DNA Ladder
Lane 2: X1 amplicon
Lane 3: X2 amplicon
Lane 4: X3 amplicon
Lane 5::1Kb DNA Ladder

Table2: PCR amplicon size

Amplicon	Size
X1	230bp
X2	230bp
X3	230bp

(C) **PCR III**

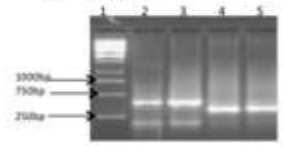


Figure 3: Agarose gel
Lane 1:1Kb DNA Ladder
Lane 2: Y1 amplicon
Lane 3: Y1 amplicon
Lane 4: Y2 amplicon
Lane 5: Y2 amplicon

Table3: PCR amplicon size

Amplicon	Size
Y1	460bp
Y2	350bp

(D) **PCR IV**

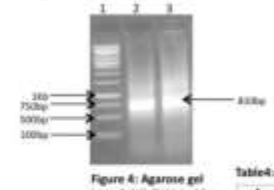


Figure 4: Agarose gel
Lane 1:1Kb DNA Ladder
Lane 2: F1 amplicon
Lane 3: F1 amplicon

Table4: PCR amplicon size

Amplicon	Size
F1	810bp

Figure 5.7: Assembly of the bacterio-opsin gene. (A) PCR 1: Generation of A1, A2, A3, A4, A5, A6 and A7 amplicons having sizes of the range 110-120 bp. (B) PCR 2: Generation of X1, X2 and X3 amplicons of size 230 bp. (C) PCR 3: Generation of Y1(460 bp) and Y2 amplicons (350 bp) and (D) Generation of final amplicon of *bop* of ~810 bp

and mutants were generated in expression vector pGEM-3z+. All the transformants were screened for the *bop* gene insert by PCR using the screening primers designed for the *bop* gene. Figure 5.8 shows the *bop* gene amplicon obtained after colony PCR of the transformants. Each of the transformants showed an amplicon of ~830 bp, corresponding to the positive control. Lane 1-24 shows the amplicon between 800-900 bp in all the clones of *bop* gene (WT and mutant).

Plasmids were purified from each of these clones and subjected to restriction digestion. Figure 5.9 shows agarose gel electrophoresis of restriction digestion of recombinant plasmids prepared from the transformants. Lanes 1-24 shows a *bop* gene cassette that has been thrown out of the recombinant plasmids of the clones of the *bop* gene (WT and mutant).

Each of the recombinant plasmids was subjected to DNA sequencing using the T7 forward primer. Figure 5.10 shows the DNA sequences that were obtained for each of the recombinant plasmids. The DNA sequences obtained are aligned with the wild type *bop* gene to observe the mutations that have been incorporated into each of the recombinant clones.

5.3.2. Wild type *bop* gene gave a functional expression of the protein whereas the amber mutants failed to express protein in *E.coli*

The vectors pGEM-SBOP-L+ and pGEM-SBOP-L- were independently transformed into *E.coli* BL-21 competant cells. The transformants so obtained were inoculated into 1000 ml of fresh Luria broth containing ampicillin and growth was continued at 37°C till an optical density of 0.6 was reached at 560 nm. The cells were induced using IPTG (1mM) and allowed to grow at 37°C for additional four hours. The cells were harvested and subjected to SDS-PAGE gel analysis for expression of the 27kDa bacterio-opsin (bOp) protein. Amber and cysteine mutants of *bop* were also expressed in a similar manner and expression was checked using SDS-PAGE. Bacterioopsin samples, with and without leader sequence, were isolated after *E. coli* expression as described earlier.

Figure 5.11 shows the expression of bacterio-opsin (WT and mutants) on SDS-PAGE from the various clones generated in section 5.2.2. Lane 1 shows the protein molecular weight marker. Lane 2 and Lane 3 show the expression of the bacterio-opsin protein (with and without the leader sequence respectively) with bands at 27kDa and 26kDa (Panel A). Lane 4 shows the presence of a band corresponding to bacterio-opsin, indicating that the D96N expresses the protein. Lanes 2-6 (Panel C) shows no band corresponding to the bacterio-opsin protein, indicating that these mutants do not express the protein due to the amber mutation. Figure 5.12 shows the chromatograms obtained for purification of bacteriorhodopsin on hydroxylapatite column as well as on DEAE-Trisacryl column chromatography. The bacteriorhodopsin obtained following the chromatography was of ~90% purity.

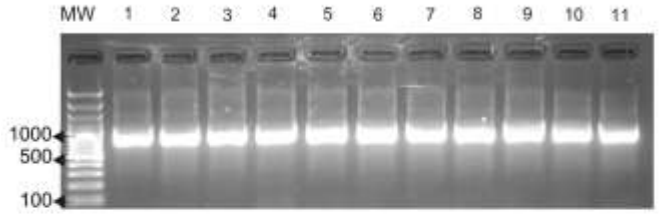
5.3.3. Wild type bOp protein can be refolded to fully functional bacteriorhodopsin

Bacterio-opsin protein was refolded using *all-trans* retinal to form bacteriorhodopsin, a purple colored protein solution. The folded bR was reconstituted into lipid vesicles and checked for its function of proton pumping. Table 5.7 shows the proton pumping activity of bR wild type (with and without leader sequence) and D96N mutant.

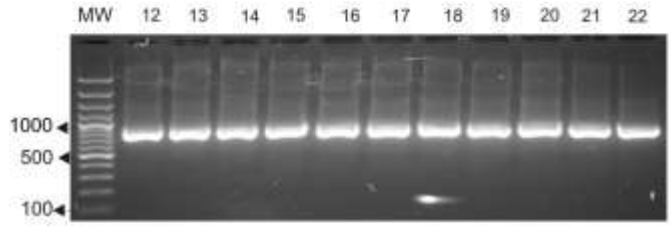
Table 5.7: Proton pumping activity of bR and its mutant at pH 6.0

Protein	Initial concentration (H+/bR/s)	Steady state (H+/bR)	% activity
bR from *Halobacterium salinarium*	3.54	36.3	100%
Wild type bR (with leader sequence)	3.51	35.8	98.6%
Wild type bR (without leader sequence)	3.48	35.6	98%
D96N mutant	0.104	1.074	3%

Lane No.	1	2	3	4	5	6	7	8	9	10	11
Clone	WT bR	D96N	L201am	L224am	L190am	S169am	V101am	Y79am	Y79C	L201C	L190C
PCR amplicon	+	+	+	+	+	+	+	+	+	+	+



Lane No.	12	13	14	15	16	17	18	19	20	21	22
Clone	D96N/ L201am	D96N/ L224am	D96N/ L190am	D96N/ S169am	D96N/ V101am	D96N/ Y79am	V101am/ Y79C	L224am/ L102C	S169am/ L190C	D96N/ V101am/ Y79C	D96N/ S169am/ L190C
PCR amplicon	+	+	+	+	+	+	+	+	+	+	+



Lane No.	23	24	25
Clone	D96N/ L224am L201C	WT bR (Positive)	Negative Control
PCR amplicon	+	+	-

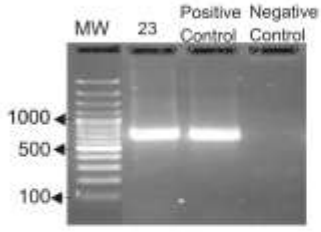
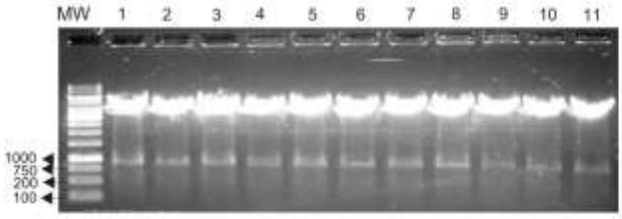
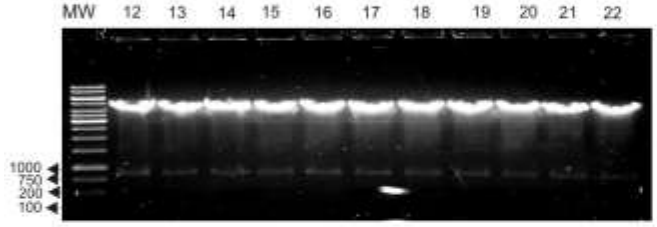


Figure 5.8: Colony PCR of mutant *bop* gene transformants. Transformants obtained for each *bop* mutant was grown in 50 ml. LB at 37C overnight. The growth obtained was centrifuged and cells were lysed by boiling. The clear lysates were used as templates for the *bop* gene PCR. PCR amplicon of ~830 bp was obtained in the wild type and as well as all the mutants. The negative control of lysed *E.coli* BL-21 cells showed absence of the amplicon at 830 bp.

Lane No.	1	2	3	4	5	6	7	8	9	10	11
Clone	WT bR	D96N	L201am	L224am	L190am	S169am	V101am	Y79am	Y79C	L201C	L190C
RE digested cassette	+	+	+	+	+	+	+	+	+	+	+



Lane No.	12	13	14	15	16	17	18	19	20	21	22
Clone	D96N/ L201am	D96N/ L224am	D96N/ L190am	D96N/ S169am	D96N/ V101am	D96N/ Y79am	V101am/ Y79C	L224am/ L102C	S169am/ L190C	D96N/ V101am/ Y79C	D96N/ S169am / L190C
RE digested cassette	+	+	+	+	+	+	+	+	+	+	+



Lane No.	23	24	25
Clone	D96N/ L201C L224am	WT bR (Positive)	Negative Control
RE digested cassette	+	+	-

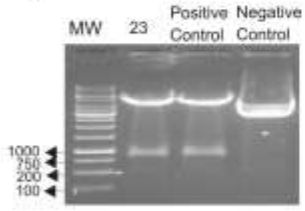
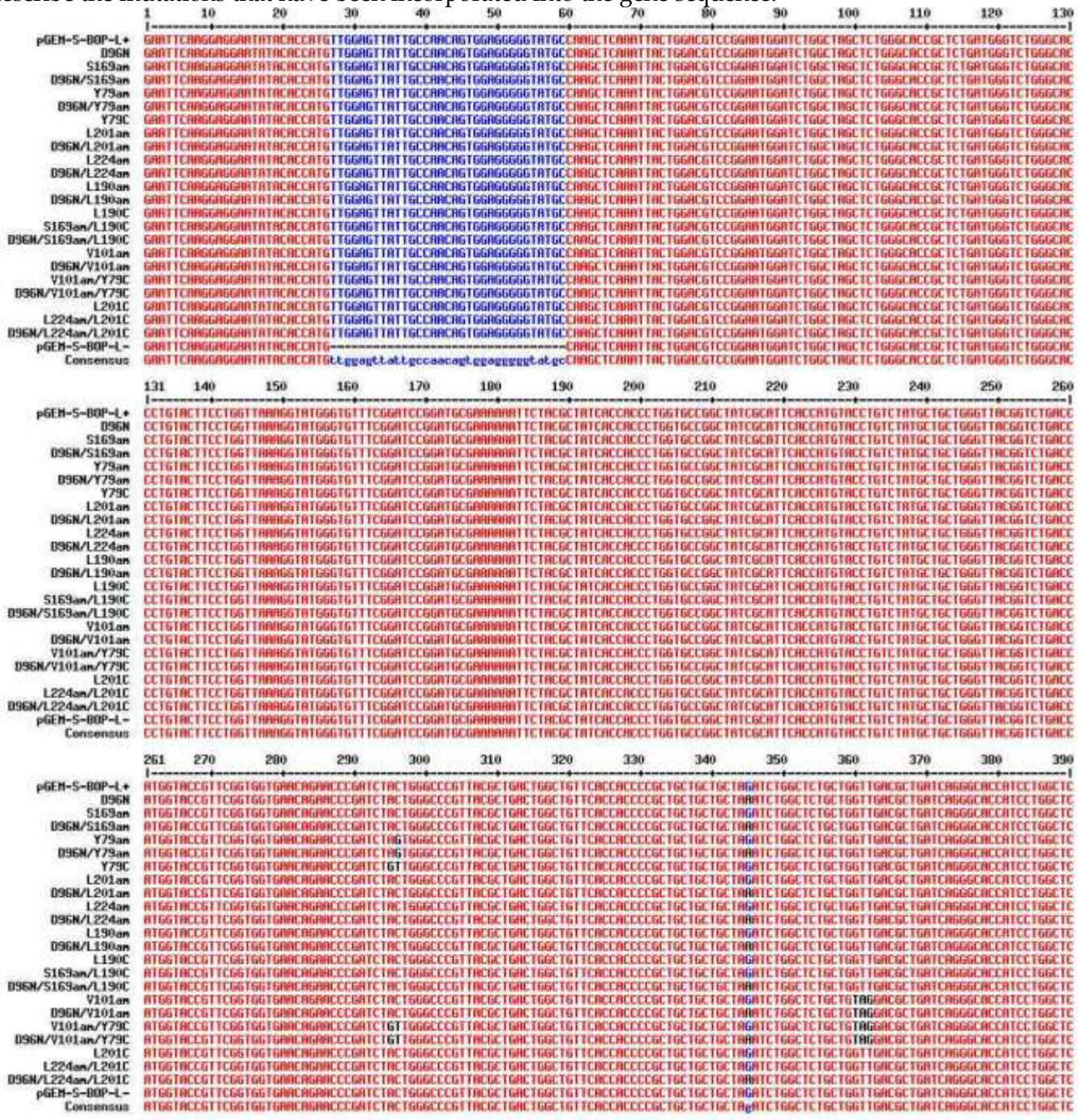
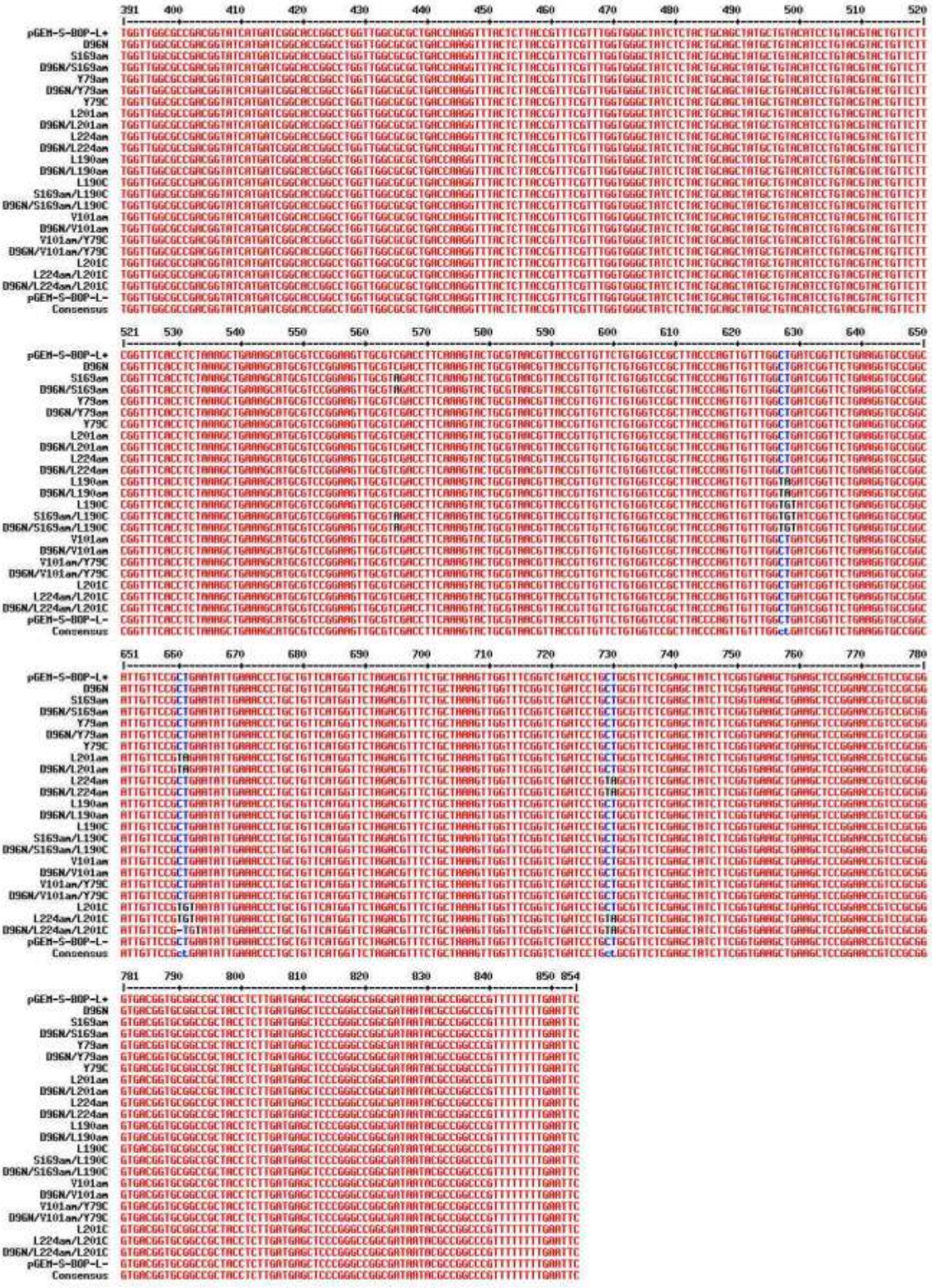


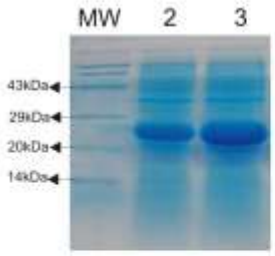
Figure 5.9: Restriction analysis of plasmids isolated from transformants. Transformants obtained for each *bop* mutant was grown in 50 ml. LB at 37C overnight. The growth obtained was centrifuged and plasmid was prepared using the mini-prep method (Maniatis et al). The plasmids were subjected to restriction digestion using the enzymes *Hind*III and *Eco*RI. Cassette of ~830 bp was thrown out in the wild type and as well as all the mutants. The negative control of pGEM-3z+ showed absence of the cassette at 830 bp.

Figure 5.10: DNA sequencing and alignment of bR WT and mutants. WT bR with and without the leader sequence was cloned and subjected to DNA sequencing. All the single, double and triple mutants were also subjected to DNA sequencing. The sequences obtained were then aligned using the software "Multalign" (www.bioinfo.genopole-toulouse.prd.fr/multalin/multalin.html). The blue letters describe the mutations that have been incorporated into the gene sequence.



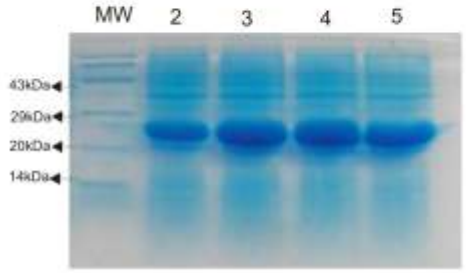

391 400 410 420 430 440 450 460 470 480 490 500 510 520
pGEM-S-BOP-L+
D96N
S169am
D96N/S169am
Y79am
D96N/Y79am
Y79C
L201am
D96N/L201am
L224am
D96N/L224am
L190am
D96N/L190am
L190C
S169am/L190C
D96N/S169am/L190C
V101am
D96N/V101am
V101am/Y79C
D96N/V101am/Y79C
L201C
L224am/L201C
D96N/L224am/L201C
pGEM-S-BOP-L-
Consensus
521 530 540 550 560 570 580 590 600 610 620 630 640 650
651 660 670 680 690 700 710 720 730 740 750 760 770 780
781 790 800 810 820 830 840 850 854

(A)



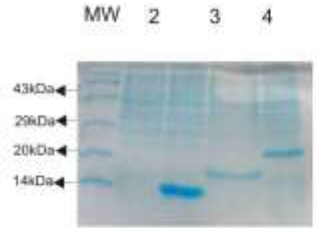
Sample	WT bOp with leader	WT bOp without leader
Expression at ~27kDa	+	+

(B)



Sample	D96N bOp	Y79C bOP	L201C bOp	L190C bOp
Expression at ~27kDa	+	+	+	+

(C)



Sample	S169am bOp	D96N/L201am bOp	V101am/Y79C bOp	L224am/L201C bOp	D96N/S169am/L190C bOp
Expression at ~27kDa	-	-	-	-	-

Figure 5.11: SDS-PAGE analysis for checking expression of bR. (A) WT bOp (with and without leader sequences) and mutant bOp clones in E.coli were expressed and lysates were loaded on a 12.5% SDS-PAGE and stained with Coomassie blue. Lane 2 and 3 shows the presence of bOp (with and without leader sequence respectively). (B) Lane 2-5 shows presence of band corresponding to bOp at 27kDa, comparable to the WT-bOp. (C) Lanes 2-5 shows absence of any band corresponding to bOp in the amber mutants.

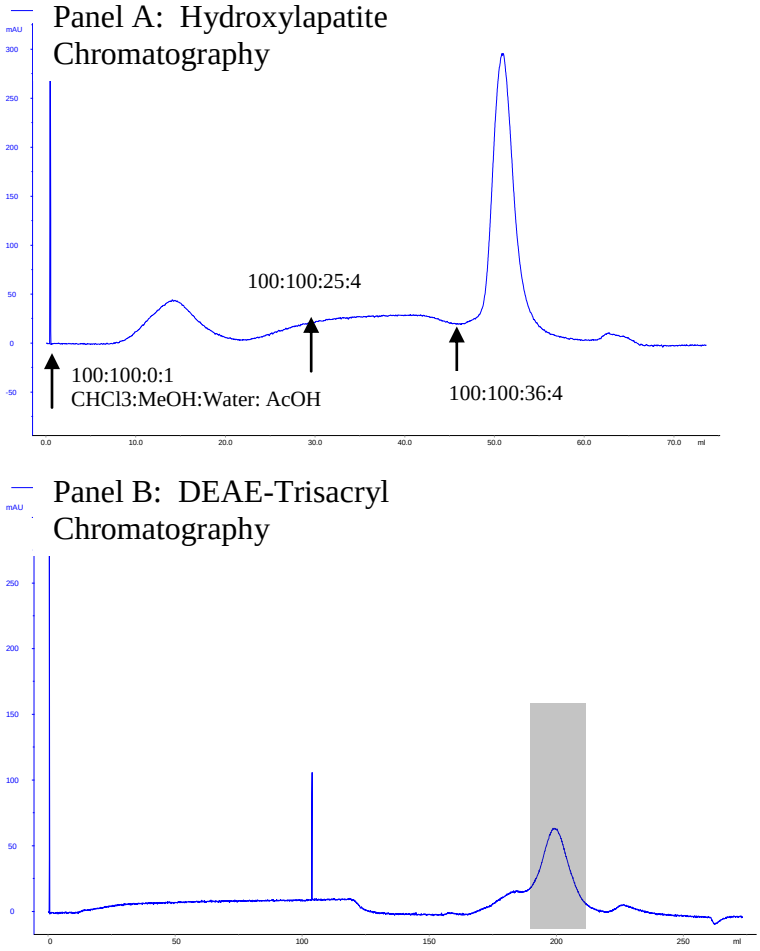


Figure: Purification of bacterioopsin: Panel A: Hydroxylapatite adsorption chromatography for purification of bacterioopsin: Column chromatography was carried out on AKTA FPLC system. See Materials and methods for detailed experimental conditions, solvent compositions were changed at the indicated points; Panel B: Ion exchange chromatography of E. coli expressed bacterioopsin on DEAE-Trisacryl: See Materials and methods for detailed experimental conditions.

5.3 Conclusions

Wild type bacterio-opsin gene was successfully cloned and expressed into *E.coli*. The protein was regenerated into bacteriorhodopsin using its chromophore retinal. This protein was found to be identical to the bacteriorhodopsin isolated from *Halobacterium salinarium* cells in terms of its activity. The protein was found to have 100% identical proton pumping efficiency when compared to the wild type bR from its natural source, *Halobacterium salinarium*. The cloned *bop* gene is thus ready to be used as an mRNA for further *in vivo* SNAAR work.

Chapter 6: *bop* mRNA: the first critical component of SNAAR

6.1. Introduction

One of the key components to *in vivo* SNAAR is mRNA of the target protein (in this case, bacterioopsin) containing the amber mutation. As described in previous chapter, wild type and amber mutation containing *bop* genes were successfully cloned into *E.coli*. This chapter details generation of the messenger RNA required for *in vivo* SNAAR work.

As, a next step towards SNAAR, this mRNA containing the amber codon at a specific site needs to be introduced into the cells along with the misaminoacylated suppressor tRNA. The theory that this amber suppression would allow the delivery of the non-native amino acid at the site of interest, and a novel protein containing a non-native amino acid would be synthesized *in vivo*, needs to be checked for in this work.

In this chapter, following work elements are described (see figure 6.1):

a) Generation of *bop* mRNA containing amber mutations using *in vitro* run off transcription
b) Encaspulation of the *bop* mRNA for lipofection
c) Lipofection of spheroplasts of *H. salinarium* with liposome encapsulated *bop* mRNA.
d) Expression and characterization of bR into *H. salinarium* using lipofected *bop* mRNA.

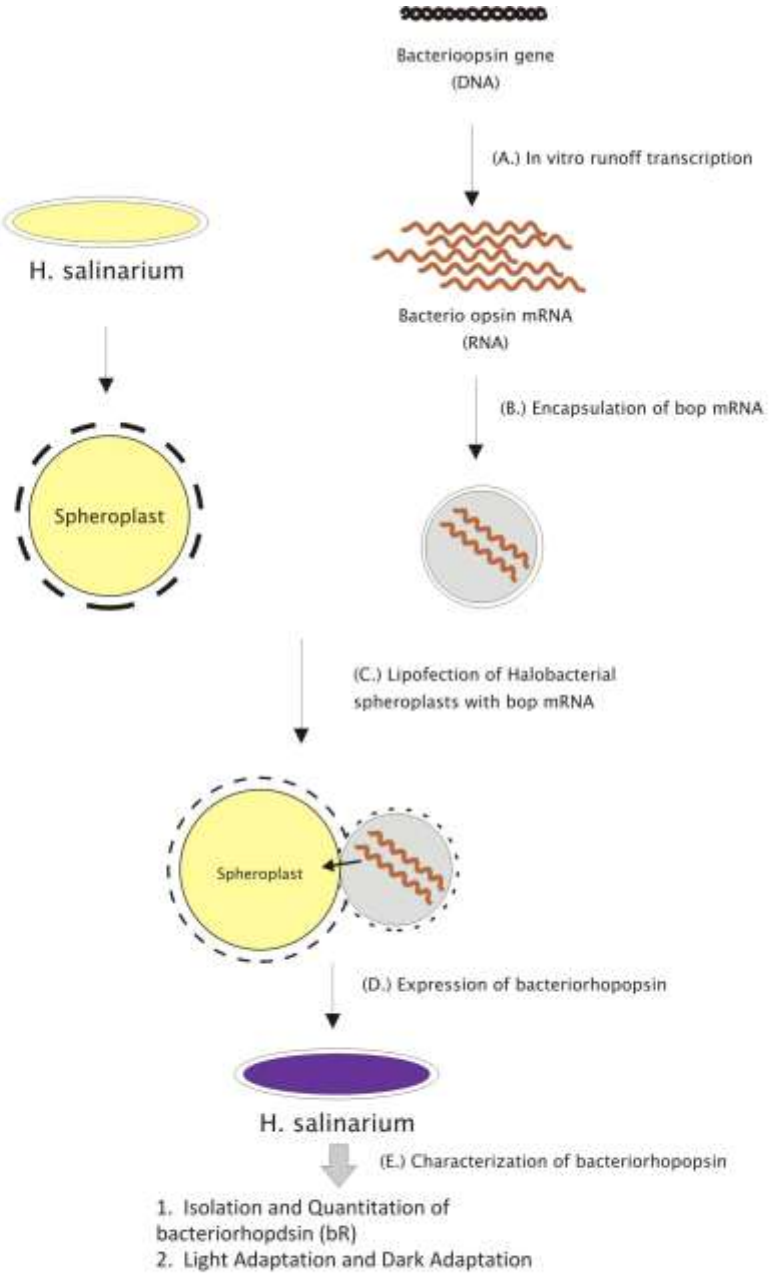


Figure 6.1: Schematic representation of the work elements in in vivo SNAAR using bacterio-opsin mRNA.

6.2. Methods and Materials

6.2.1. *In vitro* transcription of *bop* mRNA and it mutants

bop mRNA were obtained from *Hind*III linearized plasmids pGEM-SBOP-L+ and pGEM-SBOP-L- and T7 based transcription as described (*108*). For linearization of the vector, the recombinant plasmid was digested with *Hind*III and incubated at 37°C for 1.5 hours. This was followed by phenol: chloroform extraction and ethanol precipitation. The linearized plasmid was resuspended in 10 µl TE buffer. 10 µl of linearised DNA was mixed with T7 RNA transcription buffer, NTPs and T7 RNA polymerase enzyme. The reaction mixture was incubated at 37 degrees for one hour. The presence of RNA was checked on agarose gel and a UV-260 reading.

6.2.2. Liposome encapsulation of *bop* mRNA

Lipid:mRNA complex was prepared as described (*387*). Lipofectin complexed with *bop* mRNA (GIBCO BRL, USA) was used for lipofection [Lipofectin to mRNA ratio, 2.5:1; final concentration of lipids (100 µg/ml) and mRNA (40 µg/ml)] in 5 volumes of suspension of spheroplasting cells (see below) to 3 volumes of lipid:mRNA complex solution. The mixture was incubated at 37° C for 45 min.

6.2.3 Preparation of spheroplasts of *H. salinarium* and its lipofection with liposome encapsulated *bop* mRNA

Cell growth and spheroplast preparation were carried out essentially as described by Seehra and Khorana (*388*). Both Pum^+ and Pum^- strain was grown in medium [Per liter 250 g NaCl, 20 g of $MgSO_4.7H_2O$ and 3 g of trisodium citrate.$2H_2O$, 2 g KCl, 3 g of Bacto Yeast extract (Difco Laboratories, Detroit, Mich. USA) and 5 g of bacto tryptone (Difco)] at 37 °C (doubling time, 14-18 h). For anaerobic growth, cells were grown strictly in presence of nitrogen and 0.5% arginine as described (*389*). The cells were harvested in the midlog phase (turbidity ~0.7-1.0 units at 578 nm) by centrifugation (7000 x g) at 30°C. They were washed with the basal salt solution that contained per liter, 250 g of NaCl, 20 g of $MgSO_4.7H_2O$ and 2 g KCl. Washed cells from 250 ml culture grown were suspended in 10 ml of 4 M NaCl containing 25 mM KCl and 5 g/liter of L-alanine (pH

7.5). EDTA (2.5 ml of 0.5 M, pH 7.5) was then added and the suspension was incubated for 20 min at 37°C. At this time, spheroplast conversion was complete.

1.25 ml of Lipofectin (Life Technologies Inc., MD) complexed *bop* mRNA (2 mg) [Lipofectin to mRNA ratio, 2.5:1] was prepared based on procedures described earlier [Malone, 1989 #3025], and were mixed with 7.5 ml of suspension containing spheroplasts prepared from *H. salinarium* cells grown in 250 ml.

6.2.4 Expression and characterization of bR into *H. salinarium* (Pum$^-$ strain) using lipofected *bop* mRNA.

After liposome fusions and after appropriate incubations, spheroplast solution were washed with basal salts containing 0.1M $MgCl_2$, the cells were suspended in the growth medium containing peptone (volume same as during the original growth). The cells were grown for 12 hours and harvested by centrifugation (20,000 x g for 30 min) and washed with 4M NaCl. For anaerobic growth, cells were grown on 0.5% arginine and in presence of nitrogen as described (*389*). The pelleted cells were lysed by suspension in distilled water, and the membranous fraction was collected by centrifugation at 30,000 x g for 30 min, and washed with distilled water (3 x 20 ml). SDS-PAGE analysis was done at this point. After the final wash, the pellet was layered on sucrose density gradient (15-65% (w/v), 25 ml) and centrifuged at 150,000 x g for 16 h (*390*). The purple band of bR was isolated and washed with distilled water (4 x 25 ml) to remove the sucrose. Bacteriorhodopsin thus obtained was characterized using UV-visible spectroscopy and SDS-PAGE. For electrophoresis, equal amounts of cells were lysed with water, aliquots were separated by electrophoresis using 15% SDS-PAGE.

6.3. Results

6.3.1. *bop* mRNA was prepared from the cloned constructs

Bacterio-opsin mRNA prepared by *in vitro* transcription was checked in a 1.5% agaorse gel. Figure 6.2 shows the gel electrophoresis of samples prepared as *in vitro* runoff transcripts. Lane 1 shows the molecular weight markers, Lane 2 shows control reaction (without T7 RNA Polymerase), Lane 3 and 4 shows the sample from the *in vitro* transcription reaction of pGEM-S-BOP-L+ and pGEM-S-BOP-L- respectively.

All the bR mutants as listed in Table 5.2, were subjected similarly to *in vitro* runoff transcription for preparing the mRNA. This mRNA prepared contained the amber mutations required for amber suppression.

The mRNA so prepared was precipitated using ice-cold ethanol and 5M Na-acetate. The resultant RNA precipitate was dissolved in TE buffer and quantitated using absorbance at 260 nm. Typically, 500-700 μg RNA was obtained from a 100 ul *in vitro* transcription reaction.

6.3.2. Liposome encapsulated *bop* mRNA: A method for *in vivo* delivery.

Lipofectin mediated delivery is a known technique for *in vivo* delivery of nucleic acids in mammalian cells. On the phylogenetic tree, the archeabacteria *Halobacterium salinarium* closely resembles eukaryotes. Lipofectin encapsulated *bop* mRNA can hence be used to deliver this mRNA inside *Halobacterial* cells. An important point to note is that *Halobacterial* cells need to be converted to spheroplastic cells with the plasma membrane and without the cell wall so as to fuse with the liposomes.

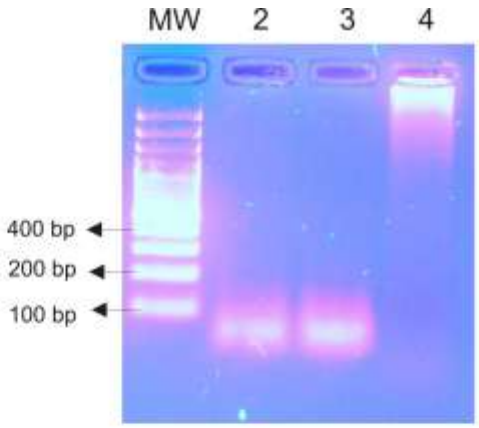


Figure 6.2: Agarose gel electrophoresis of *in vitro* runoff transcripts. Lane 2 and Lane 3 shows the runoff transcripts of obtained from pGEM-S-BOP-L- and pGEM-S-BOP-L+ respectively. Lane 4 shows the control reaction without the T7RNA polymerase enzyme.

6.3.3. Liposome mediated introduction of *bop* mRNA into spheroplasts of Pum- strains results in expression of bOp and subsequent formation of bR

Figure 6.3 shows that when lipofectin mediated introduction of *bop* mRNA was carried out in Pum$^-$ spheroplasts, bacteriorhodopsin is produced in correctly folded form as seen from SDS-PAGE (lane 4,5,6) and characteristic 558 nm absorption peak [see Figure 6.3 (inset)]. Sucrose density gradient ultracentrifugation shows that the bOp expression leads to its assembly as purple membrane lattice. In a control experiment, where only lipofectin lacking mRNA was used, no bOp production was observed (Figure 6.3, lane 3), thus indicating that the transfection protocol did not induce a latent/preexisting bop gene. Another control in which the *bop* mRNA was lipofected into Pum$^-$ cells and grown in the presence N_2 gas, showed that bR formation does not take place. This is due to the inhibition of the oxidation of β-carotene to retinal. SDS-PAGE analysis of membranes isolated from *bop* mRNA transfected spheroplasts shows presence of processed bR (presence of a small population of unprocessed bR cannot be ruled out from our analysis). This indicates that exogenously added *bop* mRNA is translated efficiently in spheroplasts of Pum$^-$ cells. Figure 6.3 inset shows absorption spectrum of bR formed after lipofection of *bop* mRNA into spheroplasts of Pum$^-$ cells. It can be seen that absorption spectrum exhibits characteristic absorption at 560 nm and is similar to native bR (*1*, *108*). The newly isolated bR showed characteristic light-dark adaptation and proton pumping activity. Our data shows that bOp translated from *bop* mRNA is integrated into membranes of halobacterial cells and binds to intrinsic retinal to form a functional bacteriorhodopsin.

6.3.4. Liposome mediated introduction of mutant *bop* mRNA into spheroplasts of Pum- strains results in no expression of bOp and subsequent formation of bR

Bacterio-opsin mRNA from mutants, when encapsulated in liposomes and fused with spheroplasts of Pum$^-$ strains of *Halobacterium salinarium*, no expression of purple membrane was seen. The mRNA failed to translate, resulting in no expression of

bacterioopsin and hence absence of bacteriorhodopsin. Table 6.1 summarizes the findings of lipofection experiments sone with various *bop* mRNAs.

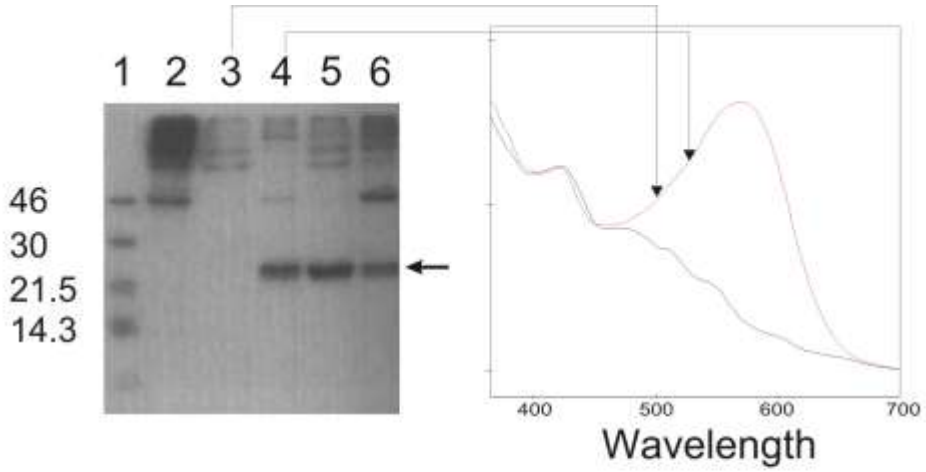


Figure 6.3: Coomassie-Blue stained SDS-PAGE analysis of membrane fractions isolated during bR purification: SDS-PAGE of various membrane fractions aliquoted before sucrose density gradient centrifugation step during isolation of bR from spheroplasts. The samples were loaded on 15% polyacrylamide gel followed by Coomassie blue staining. Lane 1: molecular weight markers; Lane 2: membranes isolated from spheroplasts of Pum$^+$; lane 3: membranes isolated from spheroplasts of Pum$^-$ cells treated with lipofectin lacking mRNA (see experimental details for text); Lane 4: membranes isolated from spheroplasts of Pum$^-$ cells treated with lipofectin associated bop mRNA; Lane 5: membranes isolated from spheroplasts of Pum$^-$ cells fused with liposomes containing bacterioopsin; Lane 6: membranes isolated from spheroplasts of Pum$^-$ cells fused with liposomes containing bR. Inset shows visible absorption spectra of the samples.

Table 6.1: Expression and Characterization of bR isolated after fusions of liposome encapsulated *bop* mRNA with Pum$^-$ spheroplastic cells of *Halobacterium salinarium*

Bacteriorhodopsin	**Type of Mutation**	**Mutation**	**Fusion of Liposome Encapsulated bop mRNA in Halobacterial spheroplasts**			
			Purple Membrane obtained (µg per 250 ml growth)	Abs. Max. (DA)	Abs. Max. (LA)	Proton Pumping Activity (%)
bR (wild type)	-	-	1200	558	570	100%
bR (with leader peptide) expressed in *E.coli*	-	-	1050	558	570	98.30%
bR (without leader peptide) expressed in *E.coli*	-	-	1138	558	570	98%
bR (M block)	Single mutant	D96N	1015	558	570	3%
bR (light adapted)	Single mutant	L201am	0	-	-	0%
		L224am	0	-	-	0%
		L190am	0	-	-	0%
		S169am	0	-	-	0%
		V101am	0	-	-	0%
		Y79am	0	-	-	0%
bR (M block)	Double mutants	D96N/L201am	0	-	-	0%
		D96N/L224am	0	-	-	0%
		D96N/L190am	0	-	-	0%
		D96N/S169am	0	-	-	0%
		D96N/V101am	0	-	-	0%
		D96N/Y79am	0	-	-	0%
bR (light adapted)	Double mutants	V101am/Y79C	0	-	-	0%
		L224am/L201C	0	-	-	0%

		S169am/L190C	0	-	-	0%
bR (M block)	Triple mutants	D96N/V101am/Y79C	0	-	-	0%
		D96N/L224am/L201C	0	-	-	0%
		D96N/S169am/L190C	0	-	-	0%

6.4. Discussion

One of the critical work elements in performing *in vivo* SNAAR is the delivery of amber suppressor mutation containing *bop* mRNA into cells. In this piece of work, bop mRNA has been successfully translated into bacterio-opsin and subsequently into bacteriorhodopsin. With this part of the work, this critical element has been addressed. The strategy involves the *in vitro* transcription mediated synthesis of mRNA from the amber mutation containing cloned constructs of the *bop* gene. The mRNA so prepared has to be successfully delivered into *Halobacterial* cells *via* hydrophobic lipids such as lipofectin. This mRNA has been translated to a functional protein, bacteriorhodopsin.

An important question to be addressed is whether the *bop* mRNA has been delivered within the cells. It is important to develop a good and reliable assay to monitor this *in vivo* delivery of mRNA. This issue would be discussed in greater details in Chapter 8.

Chapter 7: Misaminoacylated Suppressor tRNA: the second critical component for SNAAR

7.1 tRNA and Suppressor tRNA: An introduction

To expand the genetic code to allow for incorporation of additional, nonnatural, amino acids at predetermined sites in proteins that are synthesized *in vivo* requires a major intrusion into the tRNA world and some of the complexities that it contains.

As mentioned in Chapter 1, translation or protein synthesis machinery uses the genetic information in messenger RNA (mRNA) to synthesize proteins. Transfer RNA (tRNA) is charged with an amino acid in presence of ATP and the reaction is catalysed by aminoacyltRNA syntheytase. Aminoacylated tRNA's are then brought to the ribosome ((P-site), where they are paired with the corresponding trinucleotide codon in mRNA forming complex with initiation factors, ribosome, ATP and GTP. Peptide synthesis dose not start until the second aminoacyltRNA becomes bound to ribosome (A-site). The amino acid is then transferred on to the second amino acid on second tRNA in presence of peptidyl transferase, elongation factor, GTP and the ribosome moves on to the next codon. The cycle is then repeated to produce a full-length protein.

The maturation of tRNAs and mRNAs is monitored, as is the identity of amino acids attached to tRNAs. Accuracy is enhanced during the selection of aminoacyl-tRNAs on the ribosome and their base pairing with mRNA. So the tRNA world encompasses a large area of biology and chemistry (*391*). It includes all the reaction components of the translation apparatus that have tRNA-dependent interactions, aminoacyl-tRNA synthetases, ribosomal proteins, RNAs, mRNAs, initiation factors, elongation factors, translocation, peptidyl transfer, and peptide release. Therefore the fidelity of protein biosynthesis rests not only the proper interaction of the messenger RNA codon with the anticodon of the tRNA, but also on the correct attachment of amino acids to their corresponding (cognate) transfer RNA (tRNA) species. This process is catalyzed by the aminoacyl-tRNA synthetases, which discriminate with remarkable selectivity amongst

many structurally similar tRNAs. Correct recognition of transfer RNAs (tRNAs) by aminoacyl-tRNA synthetases is central to the maintenance of translational fidelity.

The hypothesis that synthetases recognize anticodon nucleotides was proposed in 1964 and considerable experimental support by mid 1970's (*392*). The genetic code is determined in aminoacylation reactions, whereby each amino acid is attached to the tRNA bearing the anticodon triplet of the codon that corresponds to that amino acid. These reactions are catalyzed by aminoacyl-tRNA synthetases. Typically there is one synthetase for each amino acid, although notable exception such as the formation of selenocysteinyl-tRNA (*393*, *394*) and glutaminyl- and asparaginyl-tRNA (*395*) exist. Aminoacyl-tRNA synthetases have a high specificity for their amino acids and their cognate tRNAs. In some case they also catalyze editing reactions to correct errors of aminoacylation. If misacylations occur that are not corrected, then toxicity result because an amino acid is incorporated into the wrong position in a growing polypeptide chain. Misacylations have been observed when mutations are introduced into either a synthetase or tRNA.

Major elements defining identity of all *E.coli* tRNAs have been deciphered and much is known about the identity elements of most yeast tRNAs and of a few tRNAs from other organisms (*392*, *396-399*). In short, identity of a tRNA is determined by a small number of nucleosides that often have been seen interacting with amino acids on the synthetases. In each tRNA, these nucleosides constitute the so-called 'identity set' that can be completed by structural elements of the nucleic acid. Negative elements that prevent a tRNA to be mischarged by noncognate synthetases can participate in identity. Some of the positive identity determinants can be considered as 'strong' since their mutation strongly reduces the aminoacylation capacity of the mutant tRNA, others are 'moderate' or 'weak'.

At first glance, several features emerge: (i) Identity elements are mainly located at the two distal extremities of the tRNA. (ii.) Except for glutamate and threonine identities, the discriminator base is a determinant, at least in *E.coli* tRNAs. (iii.) Specific structural

elements in tRNA often serve as identity determinants (i.e. the -1 residue in $tRNA^{His}$, the long extra-arm in $tRNA^{Ser}$, the G3-U70 wobble pair in $tRNA^{Ala}$. (iv.) For tRNAs specific for amino acids coded by more that four codons (leucine, serine, arginine), anticodon residues either do not participate in identity (leucine, serine) or, only for the middle C35 and semiconserved U/G36 positions (arginine).

In most systems, modified nucleosides do not participate in identity and thus are not recognized by the cognate synthetases. But they can play a major role in negative discrimination by preventing tRNAs to be recognized by noncognate synthetases, as was shown in isoleucine and aspartate systems (*400-402*). At present, the universal nature of the identity rules is rather well established and only faint differences distinguish identity sets for a given amino acid specificity along evolution.

For some AARS, the process of amino acid selection is straightforward (e.g. tryptophanyl (*193*) or arginyl tRNA synthetase (*194*), because the cognate amino acids are structurally dissimilar to other existing cognate amino acids. A single active site activates the cognate amino acid while concomitantly rejecting the others. In some instances substrate selection by a single site is challenged by competition from other substrates that are closely related in structure. For e.g. isoleucyl tRNA synthetase must select isoleucine while rejecting the closely related valine. The isopropyl side chain of valine differs from the isobutyl side chain of isoleucine. Pauling hypothesized that an enzyme active site designed to accommodate isoleucine would be unable to exclude valine from binding. He estimated that both amino acids would bind to the active site of IleRS with binding energies differeing only by 1-2 kcal/mole. Thus, IleRS should miscativate valine and transfer it to $tRNA^{Ile}$ with an efficiency that would lead to valine mistakenly replacing isoleucine in proteins at levels approaching 20% (or an error rate of 1 in 5). This prediction, however, is in direct contradiction with the high fidelity routinely observed in translated proteins. Consistently with Pauling's original prediction, IleRS misactivates Val-AMP (*403*). Remarkably, however, misaminoacylated $tRNA^{Ile}$ is never isolated. Later, Baldwin and Berg isolated a non-covalent complex of IleRS and Val-AMP, addition of $tRNA^{Ile}$ to this complex caused hydrolysis of Val-AMP

to valine and AMP. Using a special procedure, Val-tRNA^{Ile} was synthesized and challenged with IleRS. The mischarged tRNA^{Ile} was rapidly hydrolyzed by the enzyme. At the same time, Schreier and Schimmel (*195*) showed that this deacylation activity was shared by many tRNA synthetases. Thus, deacylation of mischarged tRNA appeared to be a major mechanism by which mistakes are cleared.

Experiments showing that misacylated tRNAs can be incorporated into protein supported the paradigm by suggesting that the translational machinery does not recognize the esterified amino acid of aa-tRNA. The classic Chapeville experiment using Raney Nickel to convert Cys-tRNACys to Ala-tRNACys found that alanine was incorporated at the cysteine codons in an in vitro translation assay (*5*). This result demonstrated that the translational machinery is unable to distinguish an incorrect from a correct amino acid. Many additional examples of incorporation of amino acids from misacylated tRNAs into protein have since been reported (*109*, *404*, *405*). Perhaps the most extensive experiments evaluating the incorporation of misacylated tRNAs relied on measuring the extent of suppression of nonsense codons by suppressor tRNAs.

For example, in the course of deducing the recognition rules of aaRSs, several amber-suppressor tRNA bodies were deliberately mutated such that they were aminoacylated by a different aaRS, and the resulting 'identity-swapped' tRNAs were shown to insert the new amino acid into protein (*392*, *399*). In addition, suppressor tRNAs esterified with O30 different unnatural amino acids have been successfully incorporated into protein (*104*). Together, these data suggest that the translational apparatus lacks specificity for different amino acids, once they are esterified onto tRNA. In a few isolated cases, however, the translation machinery seems to show specificity for the esterified amino acid. A prominent example occurs in the transamidation pathway, which is used as an alternative to GlnRS to produce Gln-tRNAGln in many bacteria and archaea (*406*, *407*). In this pathway, tRNAGln is first misacylated by GluRS to form Glu-tRNAGln and then reacted with a specific amidotransferase to produce Gln-tRNAGln. Because organisms using this pathway do not show misincorporation of glutamic acid at glutamine codons, it seems that the misacylated Glu-tRNAGln intermediate of this

pathway is not translated. However, *in vitro* experiments show that, although Gln-tRNAGln can bind EF-Tu, Glu-tRNAGln binds poorly (*83*). As a result, the amidotransferase can successfully compete with EF-Tu for the misacylated Glu-tRNAGln. Thus, in this case, the translational machinery seems to discriminate against certain misacylated tRNAs.

7.1.1. The use of "Amber Suppressors"

One approach is to charge a amber suppressor tRNA with a nonnatural amino acid and to create a premature stop codon at the desired position in the mRNA of interest. This approach has been used successfully with *in uitro* translation systems, (*73*, *199*, *226*, *232*, *233*, *237*, *408*). The suppressor tRNA can be classified into three groups on the basis of protein sequence information.

Class I suppressor tRNAs: tRNA($^{CUA}_{Ala2}$), tRNA($^{CUA}_{Gly1}$), tRNA($^{CUA}_{HisA}$), tRNA($^{CUA}_{Lys}$), tRNA($^{CUA}_{Tyr}$) and tRNA($^{CUA}_{ProH}$), insert the predicted amino acid.

Class II suppressor tRNAs: tRNA ($^{CUA}_{GluA}$), tRNA($^{CUA}_{Gly2}$) and tRNA($^{CUA}_{Ilet}$) were either partially or predominantly mischarged by the glutamine aminoacyl tRNA synthetase.

Class III suppressor tRNAs: tRNA ($^{CUA}_{Arg}$), tRNA($^{CUA}_{AspM}$), tRNA ($^{CUA}_{Ile2}$), tRNA ($^{CUA}_{Thr2}$), tRNA ($^{CUA}_{Met}$), and tRNA($^{CUA}_{Val}$) inserted predominantly lysine.

Many different non-natural amino acids have been introduced into a variety of proteins. In this process, the tRNA amber suppressor was charged by a combination of chemical and enzymatic methods that omitted an aminoacyl-tRNA synthetase. One problem is that naturally occurring tRNA synthetase can not activate the virtually unlimited number of nonnatural amino acids that could be introduced into a protein. The preparation of charged tRNAs *in vitro,* using these specialized methods, precludes doing protein synthesis *in vivo.* But, in spite of advances in the efficiency of *in vitro* translation systems, the yields are low while the labor and costs of materials are high, compared with *in vivo* systems. These considerations motivate the effort to develop

an *in vivo* aminoacylation system for incorporation of novel amino acids into proteins (*144*). In case of introduction aminoacyl tRNA into *in vitro* translation system it should be remembered that those aminoacylated tRNAs should not be edited by the aminoaclyl tRNA synthetases present in the system. Kwok and Wong have shown that E.coli phenylalanine-tRNA synthetases aminoacylates yeast tRNAPhe less than 1 percent as well as it acylates E.coli tRNAPhe (*212*). So it is recomended that prokaryotic tRNA should be used for eukaryotic translation systems and vice- versa.

Our search for suitable suppressor tRNAs is guided by following criteria:

1. While a suitable suppressor tRNA can always be chemically aminoacylated, the procedures are extremely chemistry intensive and are found to be "bottleneck factor" for wider applications.
2. Post-aminoacylation modifications of nucleophilic side chains of amino acids linked to the sup-tRNA is found to be an attractive and simple approach for easy SNAAR. A search for cysteine or lysine sup-tRNAs is therefore of prime priority.
3. While *in vivo* expressed suptRNAs would be ideal, their isolations from other tRNAs pose considerable challenges. This fact necessitates use of run off transcripts of suptRNAs. Evaluation of a variety of run off transcript becomes essential.

7.1.2. Our approach towards "tRNA mediated protein engineering"

We have chosen cysteine suppressor tRNAs as our focus because they would allow both chemical and enzymatic (post-aminoacyl modifications) pathways for SNAAR. We decided to evaluate E.coli tRNAs (both cognate and suppressor) in terms of aminoacylation efficiency and amber suppression in Halobacterial cells.

Proposed work elements are as follows:

- Gene construction and cloning of tRNA and suppressor tRNA of different organism.
- *In vitro* run off transcription of full length tRNA.

- Aminoacylation of the transcripts.
- Post-aminoacylation modification of amino acid residue on aminoacyl tRNA.
- Protein expression in presence of misaminoacylated sup-tRNA.

7.2 Methods and Materials

7.2.1. Gene constructs of cysteine tRNA (Cys-tRNA^{Cys}) and cysteine suppressor tRNA (Cys-tRNA$^{Cys}_{CUA}$)

Sequences of tRNA^{Cys} and tRNA$^{Cys}_{CUA}$ were adapted from Koumatsoulis et al (*409*). Transfer RNA gene constructs included a T7 RNA polymerase promoter, the tRNA gene ending with a unique downstream *Bst*NI site. Genes for tRNAs were constructed from six oligonucleotides (for each tRNA) and were cloned into *Eco*RI and *Hind*III sites of pSP65 (Promega Corpn. Wisconsin) to obtain plasmid constructs pSP65C and pSP65C-sup (See Figure 7.1).

7.2.2. *In vitro* transcription of Cys-tRNA^{Cys} and Cys-tRNA$^{Cys}_{CUA}$

Both tRNA constructs, pSP65-C, and pSP65-C-sup, were linearized and digested with *Bst*NI. *bop* gene construct was linearized with *Hind*III. Run-off transcripts of linearized DNA template were prepared using T7 RNA polymerase (Bangalore Genei, Bangalore, India) as described (*90*, *108*). The RNA generated by this process was quantitated using UV260 absorbance.

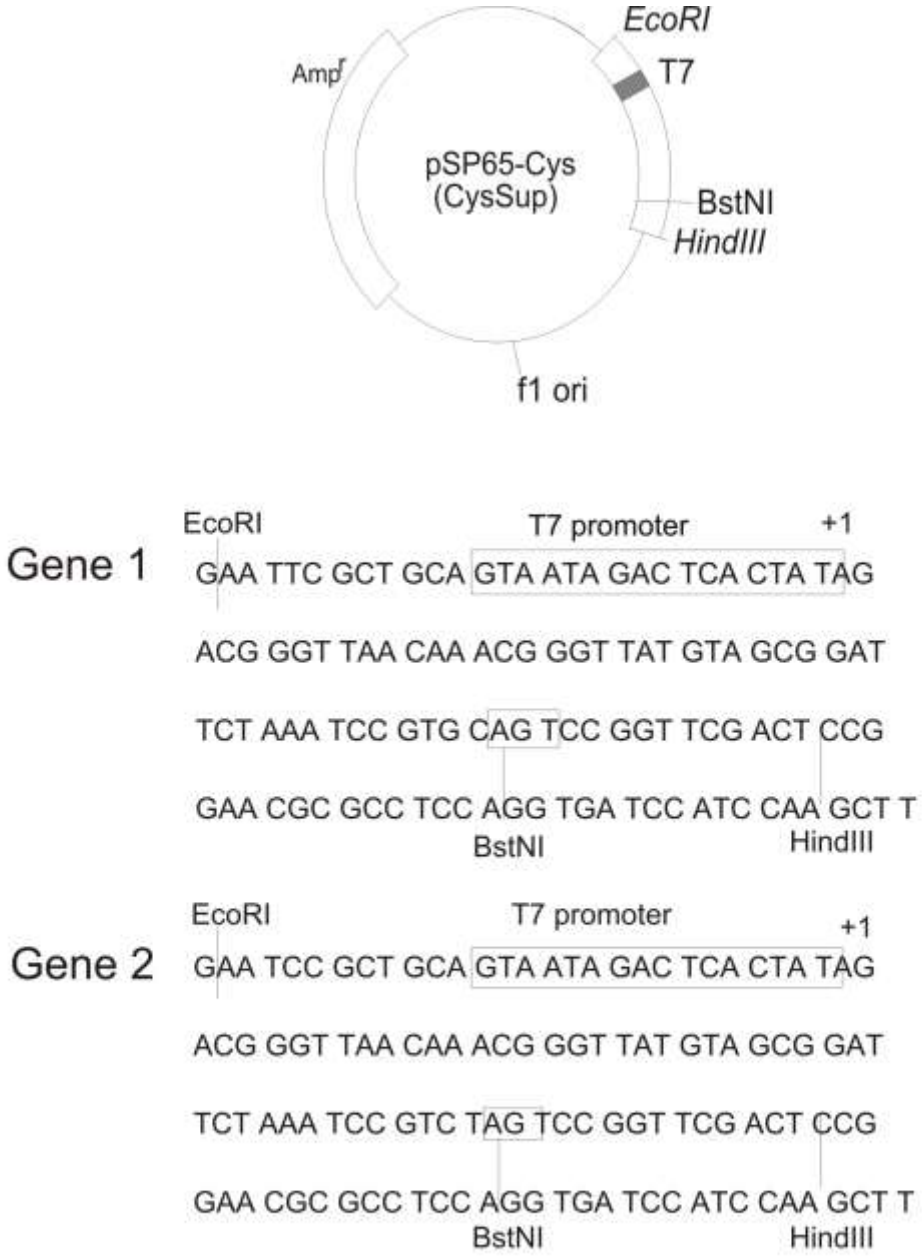


Figure 7.1: Construction of genes for tRNAcys and tRNAcysCUA: Synthetic genes for tRNAcys (Gene 1) and tRNAcysCUA (Gene 2) were synthesized and were cloned into EcoRI/HindIII sites. The vector was linearised with BstNI giving a CCA at the 3'-end.

7.2.3. Enzymatic Aminoacylations of Cys-tRNA^{Cys} and Cys-tRNA$^{Cys}_{CUA}$

Aminoacylation reactions were performed according to Sonar *et al* (*199*). For kinetic parameter determinations, tRNA concentration was varied from 0.2 μM to 1.5 μM with fixed *E.coli* S-100 concentration. K_m and V_{max} were calculated using ^{35}S-Cysteine from Michealis Menten plots (*409*).

7.2.4. Post-aminoacylation modifications of Cys-tRNA^{Cys} and Cys-tRNA$^{Cys}_{CUA}$

Typically 30 A_{260} units of extensively dialyzed (essentially free of DTT) aminoacylated tRNA (2 nmol Cys) were dissolved in 0.3 ml ice-cold 25 mM phosphate buffer (pH 6.0) and added to a 0.6 ml of prefrozen solution of N-(7-dimethylamino-4-methylcoumarin-3-yl)maleimide (DACM) (Molecular Probes, Eugene, OR) in DMSO. Ratio of tRNA to DACM is maintained at 1:10. Reaction was allowed to continue for 4 h at 4 °C and at the end of reaction, excess β-mercaptoethanol was added to scavange thiol reactive reagent. After additional 30 min, the reaction was stopped and "misaminoacyl-tRNAs" thus obtained were precipitated and resuspended in 5 mM NaOAc, pH 5.5. The products were characterized by fluorescence measurements (DACM: excitation at 383 nm, emission at 463 nm; ISA label: excitation at 355 nm and emission at 407 nm, BODIPY: excitation at 504 nm, emission at 530 nm). Labeled aminoacyl-tRNA linkage is stable for >48 hours at 25 °C and for weeks at -20 °C.

7.2.5. Determination of the extent of –SH modification on post-aminoacylated tRNA

Post-aminoacyl modified tRNAs were subjected to base hydrolysis by incubating in borate buffer (pH 10.0, 25 mM) for 1 h at 37°C. Total tRNA was alcohol precipitated and the supernatant was subjected to absorption measurement. Based on extinction coefficient of released labeled amino acid, the extent of modification was calculated.

7.2.6. Liposome encapsulation of misaminoacylated tRNA

Lipid:tRNA complex was prepared as described (*387*). Lipofectin complexed with *tRNA* (GIBCO BRL, USA) was used for lipofection [Lipofectin to mRNA ratio, 2.5:1; final concentration of lipids (100 μg/ml) and tRNA (40 μg/ml)] in 5 volumes of suspension of spheroplasting cells (see below) to 3 volumes of lipid:tRNA complex solution. The mixture was incubated at 37° C for 45 min.

7.2.7. Preparation of spheroplasts of *H. salinarium* and its lipofection with liposome encapsulated misaminoacylated tRNA.

Cell growth and spheroplast preparation were carried out essentially as described by Seehra and Khorana (*388*). Both Pum^+ and Pum^- strain was grown in medium [Per liter 250 g NaCl, 20 g of $MgSO_4.7H_2O$ and 3 g of trisodium citrate.$2H_2O$, 2 g KCl, 3 g of Bacto Yeast extract (Difco Laboratories, Detroit, Mich. USA) and 5 g of bacto tryptone (Difco)] at 37 °C (doubling time, 14-18 h). For anaerobic growth, cells were grown strictly in presence of nitrogen and 0.5% arginine. The cells were harvested in the midlog phase (turbidity ~0.7-1.0 units at 578 nm) by centrifugation (7000 x g) at 30°C. They were washed with the basal salt solution that contained per liter, 250 g of NaCl, 20 g of $MgSO_4.7H_2O$ and 2 g KCl. Washed cells from 250 ml culture grown were suspended in 10 ml of 4 M NaCl containing 25 mM KCl and 5 g/liter of L-alanine (pH 7.5). EDTA (2.5 ml of 0.5 M, pH 7.5) was then added and the suspension was incubated for 20 min at 37°C. At this time, spheroplast conversion was complete.

1.25 ml of lipofectin complexed with postaminoacylation SH-modified tRNA (e.g. DACM-Cys-$tRNA^{Cys}_{CUA}$) (2.5 mg tRNA) [Lipofectin to tRNA ratio, 2.5:1] was prepared based on procedures described earlier (*387*), and were mixed with 7.5 ml of suspension containing spheroplasts prepared from *H. salinarium* cells grown in 250 ml.

7.2.8. Determination of postaminoacylation modified $tRNA_{CUA}$ from cells

Cells lipofected with aminoacylated tRNA were lysed under acidic conditions using osmotic lysis. The total tRNA was isolated using the standard tRNA isolation protocol. The amount of postaminoacylation modified tRNA was quantitated by measuring the fluorescence of the label in this tRNA fraction. To check for the efficiency of recovery, a standard was run where unlipofected cells were broken and known amount of tRNA was added and reisolated.

7.3 Results

7.3.1. Cys-$tRNA^{Cys}$ and Cys-$tRNA^{Cys}_{CUA}$ were obtained by runoff transcriptions from their cloned genes

Both Cys-$tRNA^{cys}$ and cys-$tRNA^{cys}_{CUA}$ were prepared by *in vitro* transcription from their cloned genes. Figure 7.2 shows the gel electrophoresis of samples prepared as *in vitro* runoff transcripts. Lane 1 shows the molecular weight markers, Lane 2 shows control reaction (without T7 RNA Polymerase), Lane 3 and 4 shows the sample from the *in vitro* transcription reaction of pSP65-C and pSP65-C-sup respectively. The tRNA so prepared was quantitated using absorbance at 260 nm. Typically, 500ug RNA was obtained from a 100 ul *in vitro* transcription reaction.

7.3.2. A new method for non-native aminoacylation of $tRNA_{CUA}$

In this study, we have chosen a simpler route for generation of non-native aminoacyl linked amber suppressor tRNA (i.e. "misaminoacylated" $tRNA_{CUA}$) based on post-aminoacylation modification, instead of the currently used "chemical misaminoacylation" methods. Although, the method used here introduces non-native moieties as cysteine side-chain modifications, it avoids several synthetic organic chemistry based steps involved in chemical aminoacylation such as: (a) dinucleotide protections, b) synthesis of carboxyl activated and N-protected non-native amino acid,

and subsequent aminoacylation of the protected dinucleotide (c) T4 RNA ligase based ligation of truncated tRNA with aminoacyl dinucleotide, and finally (d) photo-

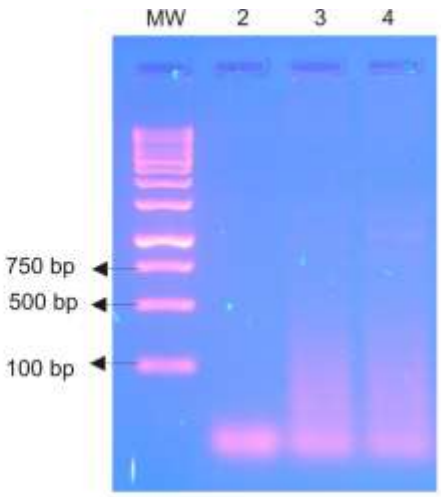


Figure 7.2: Agarose gel electrophoresis of *in vitro* runoff transcripts. Lane 2 shows the control reaction without the T7 RNA polymerase enzyme. Lane 3 and Lane 4 shows the runoff transcripts of obtained from pSP65-C and pSP65-C-sup respectively.

deprotection of the entire misaminoacylated suppressor tRNA. In our approach, $tRNA^{Cys}$ and $tRNA^{Cys}_{CUA}$ were prepared using *in vitro* run-off transcription (*410*). Aminoacylation kinetics was studies using ^{35}S-cysteine labels by varying the tRNA concentrations. S-100 extract was added along with excess ^{35}S cysteine to the reaction and the aminoacylated tRNA was precipitated and counted (see figure 7.3). Michealis-Menten curves of velocity (V) vs substrate concentration [S] was plotted to obtain the K_m and V_{max} values for both Cys-tRNA and Cys-$tRNA_{CUA}$ (See figure 7.4). Cysteine aminoacylation of both transcripts have shown comparable K_m (3.5 vs 7.15 μM, respectively) and V_{max} (0.54 vs 0.54 μmoles/l/min).

7.3.3. Post aminoacylation modification of –SH groups of Cys-$tRNA^{Cys}$ and Cys-$tRNA^{Cys}_{CUA}$ is a method of introducing non-native cysteines site specifically into protein

Bacterio-opsin protein, fortunately, does not contain any cysteines, and hence cysteine codons could be unique for protein engineering. The strategy used in this work can be an excellent method to introduce non-native cysteines in proteins which do not contain any cysteinyl groups. Also, the –SH group of cysteine can be exploited since it is reactive to a variety of reagents. These could be fluorogenic, chromophoric or even para-magnetic. Such modified cysteines, which are unnatural amino acids, can be incorporated site-specifically to give proteins with unique properties. Such labels can be used for a variety of studies or to alter the properties of these modified proteins.

FNon-native functionalities (DACM, ISA and BODIPY) are introduced as post-aminoacylation modifications of SH- groups of the cysteine present on aminoacyl tRNA. The extent of post-aminoacylation modifications is determined to be >95% using sulphydryl group determination and hydrolysis of aminoacyl-tRNA linkage (*411*). Table 7.1 gives the post-aminoacylation modification efficiencies of all the charged tRNAs.

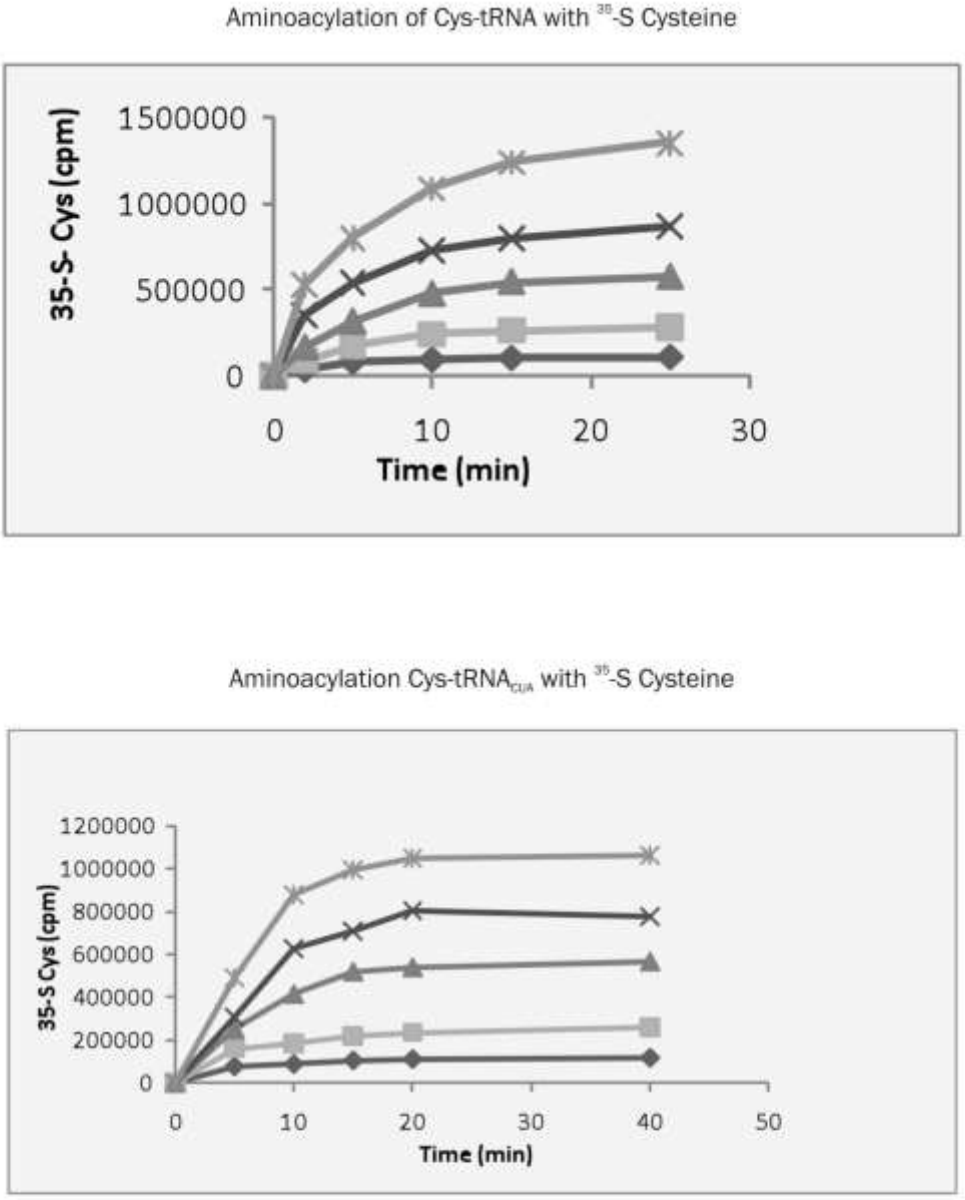


Figure 7.3: Aminoacylation of Cys-tRNA and Cys-tRNA$_{CUA}$. Aminoacylation was studied using 35-S Cysteine as substrate with varying amounts of tRNA and S-100 extracts. The reaction was arrested at various time-points by spotting the reaction on a filter paper, followed by washing the paper with TCA. The charged tRNA remains bound to the filter paper, whereas the other components gets washed away. This paper is then used for scintillation counting.

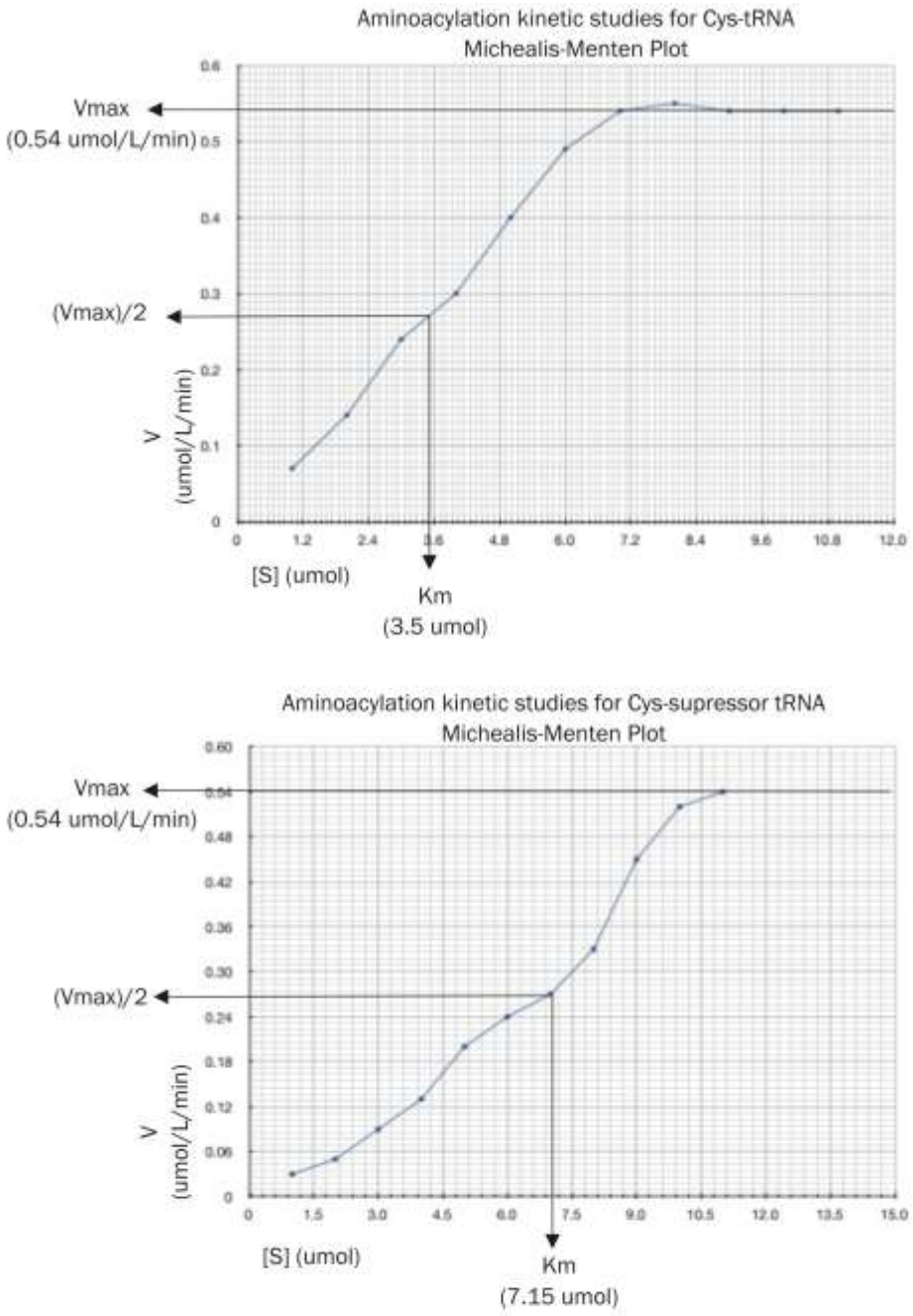


Figure 7.4: Michealis Menten Plot for determination of kinetic parameters for aminoacylation of Cys-tRNA and Cys-sup-tRNA. tRNA concentration was varied with fixed quantity of S-100 extract in presence of excess of ^{35}-S Cysteine. Km and Vmax values were calculated from the Michealis Menten plots.

Table 7.1: Labelling efficiency of post-aminoacyl modification of charged cys-tRNA^{cys} and cys-tRNA$^{cys}_{CUA}$

Mutation	Label	Labeling Efficiency	Absorption Maximum	Emission Maximum
Cys-tRNA$^{cys}_{CUA}$	DACM	98%	383	463
Cys-tRNA$^{cys}_{CUA}$	ISA	95%	355	407
Cys-tRNA^{cys}	BOPDIPY-Fl	96%	504	530

7.3.4. Liposome encapsulated misaminoacylated tRNA: A method for *in vivo* delivery.

Lipofectin mediated delivery of DACM modified cyteine tRNA into *Halobacterial* cells was checked by harvesting the spheroplastic cells, washing the cells and lysing the cells. The lysate was checked to fluorescence emission (DACM excitation: 383 nm; emission: 463 nm). DACM fluorescence was found in lysates indicating that the DACM modified tRNA had been delivered within the cells.

7.4 Discussion

The second critical work elements in performing *in vivo* SNAAR is the delivery of misaminoacylated suppressor tRNA into cells. With this part of the work, this critical element has been addressed. The strategy involves the *in vitro* transcription mediated synthesis of tRNA from the cloned constructs of the *suppressor tRNA* gene. The tRNA so prepared has to be successfully delivered into *Halobacterial* cells *via* hydrophobic lipids such as lipofectin.

Chapter 6 and Chapter 7 have thus described that the two most important components required for the *in vivo* SNAAR, i.e. mRNA and tRNA can be independently delivered into *Halobacterial* cells.

Chapter 8: *In-vivo* delivery of macromolecules such as *bop* mRNA, bacterioopsin protein and bacteriorhodopsin helped in elucidation of the Retinal Biosynthetic Pathway

8.1 Introduction

In vivo SNAAR requires the simultaneous introduction of amber codon carrying gene and misaminoacylated suppressor tRNA into cells so that a non-native amino acid is incorporated at the desired site when the cellular machinery translates the mRNA. Thus, the first experimental goal was to introduce macromolecules such as nucleic acids (DNA and RNA) into cells. This is crucial, as *in vivo* SNAAR requires the introduction of nucleic acids such as tRNA and mRNA. Several methods have been used to introduce nucleic acids *in vivo*. Some of these introduce nucleic acids *via* physical means such as electroporation whereas others use several chemicals such as calcium or lithium salts to make cells competant to accept the macromolecules. Introduction of nucleic acids into eukaryotic cells is done using the lipid vesicles. Phosphatidylcholine and phosphatidylethanolamine have been widely used to deliver mocromolecules *in vivo*. Cationic lipids such as lipofectamine (DOPE: DOTMA) have been found to extremely efficient in the delivery of macromolecules *in vivo*.

As a method of delivering nucleic acids to accomplish *in vivo* SNAAR, lipofectin was chosen as a carrier of macromolecules. Since we chose bacteriorhodopsin as a model protein to perform SNAAR, it was decided that the microorganism used to demonstrate *in vivo* SNAAR be the naturally occurring Halobacterial system. Thus, *Halobacterium salinarium* was the organism of choice to deliver the nucleic acids. *Halobacterium salinarium* is an archeabacterium and is supposed to be intermediate to bacteria and eukaryotes. Transformation in such archeabacteria can be achieved *via* fusion of liposomes made up of halobacterial membrane lipids with spheroplastic cells.

This organism contains three unique retinal proteins that control diverse functions. The study of such light activable retinal proteins (*412*) provides us with information on how

organisms adapt to their environments using light. The three proteins are bacteriorhodopsin involved in light induced energy transduction (*413*), sensory rhodopsin controlling light induced movement (*414*) and halorhodopsin for light induced chloride transport (*415*). This chapter describes an attempt to deliver nucleic acids *in vivo* into mutant halobacterial spheroplasts. This mutant (SD9 strain), also called Pum$^-$, was deficient in bacteriorhodopsin due to the disruption of the *bop* gene by an insertion element (*416*). When *bop* mRNA was transformed into such a mutant strain, it restored the characterisitics of the wild type S9 *Halobacterial* strain. A further analysis of this newly transformed strain showed some very interesting results in terms of retinal biosynthesis. As an offshoot from the main experiment and in order to study the role of the bacterio-opsin and bacteriorhodopsin protein in retinal biosynthesis, it was decided to deliver these protein macromolecules *via* liposomes. The results of these experiments are described in this chapter.

Current understanding of bR biosynthesis

Retinal acts as a chromophore and is essential for the light induced activity of these retinal proteins. In case of bacteriorhodopsin (bR)[1], the apoprotein bacterio-opsin (bOp) is attached to *all-trans* retinal *via* Lys-216 residue to form the chromoprotein bacteriorhodopsin (bR) (*249*). Intermediates in the pathway of retinal biosynthesis are well characterised and its final biosynthetic stages are known to involve the cyclization of lycopene to β-carotene followed by an oxidation of β-carotene to *all-trans* retinal (*417*) (See Figure 8.1). Biogenesis of purple membrane is inducible by limiting the oxygen supply (*418*), which turns on synthesis both of bacterio-opsin and retinal (*417*). In contrast, most of the lipid molecules necessary for purple membrane are synthesized long before the start of bacterio-opsin and retinal synthesis, i.e. the lipids are taken from the pool of the cell membrane.

[1] Abbreviations: bR: bacteriorhodopsin, bOp: bacterioopsin protein, *bop*: bacterioopsin mRNA, CHAPS: 3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate, DMPC: 1,2-Dimyristoyl-glycero-3-phosphocholine, *N*-(7-nitrobenz-2-oxa-1,3-diazol- 4-yl)-1,2-dihexadecanoyl-*sn*- glycero-3-phosphoethanolamine, triethylammonium salt (NBD-PE), N-Rh-PE: Rhodamine B 1,2-dihexadecanoyl-*sn*-glycero- 3-phosphoethanolamine, triethylammonium salt (rhodamine DHPE)Pum$^-$: lacking purple membrane, , SDS-PAGE: Sodium dodecyl sulfonate polyacrylamide gel electrophoresis.

Essentially all the retinal present in *Halobacteria* is associated with bacterio-opsin, so one observes nearly stoichiometric relation between bacterio-opsin and retinal content. This fact suggests the existence of a highly efficient regulation mechanism coordinating the bacterioopsin and retinal biosynthetic pathway (See Figure 8.1).

Early studies indicate that retinal biosynthesis in *Halobacteria* may be stringently regulated and factors involved in such a control remain poorly understood. There are four possible regulatory factors: transcriptional events leading to bacterioopsin mRNA, translation of *bop* mRNA to bacterioopsin protein, *all-trans* retinal and folded chromoprotein bacteriorhodopsin.

Earlier Sumper et al. have demonstrated that bacterio-opsin has some role to play in the tight regulation of retinal biosynthesis(*417*). When *Halobacterial* cells were grown in presence of protein inhibitor, puromycin, the retinal synthesis was found to be inhibited. Similarly, when retinal pathway was inhibited with nicotine, a five-fold excess of bacterio-opsin was found assembled in the form of brown membrane.

Other macromolecules such as proteins were also independently delivered *in vivo* in *Halobacteria* along with RNA and DNA, *via* lipofectin. The protein bacteriorhodopsin and the apo-protein bacterio-opsin were independently delivered *in vivo* in an attempt to probe the pathway involved in the regulation of it biosynthesis.

As a spinoff of the ongoing investigation for delivery of macromolecules into cells for SNAAR, we report a new approach to understand regulating factors involved in retinal biosynthesis in *halobacteria*. This approach involves *in vivo* introduction of mRNA or membrane proteins *via* liposome fusion into *halobacterial* spheroplasts. To elucidate the role of *bop* transcription, bacterio-opsin encoding mRNA is introduced in spheroplasts of Pum^- strain which initially contains neither bacterioopsin nor retinal. The translational product, bacterio-opsin was introduced into Pum^- strain using proteoliposome fusion. Results indicate that translation of *bop* mRNA results in expression of bacterioopsin,

which triggers retinal biosynthesis from lycopene. Importantly, delivery of liposome encapsulated apoprotein; bacterioopsin into spheroplasts of Pum- also triggers

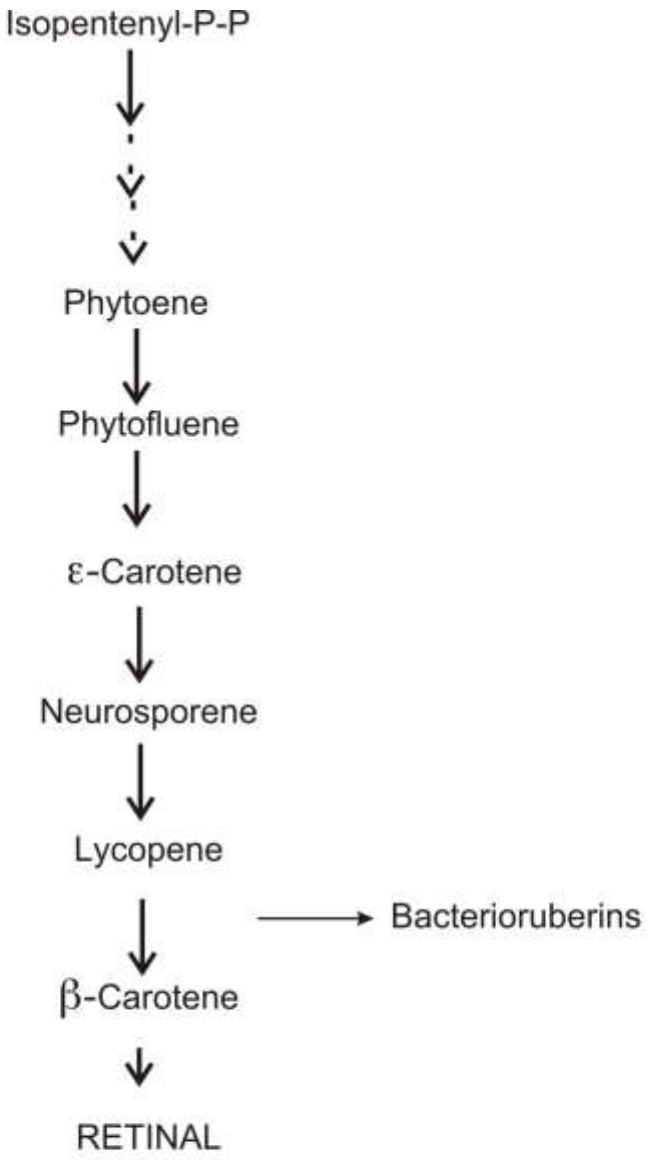


Figure 8.1: Retinal biosynthetic pathway: Retinal biosynthesis starts from isopentenyl pyrophosphate which converts into key intermediates such as lycopene, carotene followed by conversion to all-trans retinal. There exists a highly efficient regulation mechanism coordinating the bacterioopsin and retinal biosynthetic pathway.

biosynthesis of retinal and generation of bacteriorhodopsin. In order to determine the regulation leading to inhibition of retinal biosynthesis, both excess retinal and chromoprotein bR were exogenously introduced into Pum- spheroplasts. When chromoprotein, bacteriorhodopsin is delivered into spheroplasts of Pum-, lycopene to retinal conversion is inhibited whereas excess retinal does not inhibit lycopene utilization for retinal biosynthesis. Thus, these results show that while bacterioopsin triggers retinal biosynthesis, it is the formation of bacteriorhodopsin and not the excess of retinal that inhibits its biosynthesis.

8.2 Materials and Methods

8.2.1. Preparation of bOp and bR containing liposomes

bOp and bR containing liposomes were prepared as described (*419*). Both, bOp and bR were individually suspended in distilled water at 450 μM of protein content. Reconsitution of bOp and bR into liposomes is accomplished by sonication (*419*, *420*). A chloroform/methanol solution containing 2 mg of egg phosphatidylcholine and 18 mg of acetone ether-washed halobacterial lipids (*421*) is dried under N_2 gas, lyophilized and dispersed into 0.5 ml of 40 mM HEPES-KOH, pH 7.0 and 100 mM KCl. The lipid suspension is then diluted with an equal volume of distilled water containing 76.8 nmol of bR or bOp and sonicated (cycles of 15 sec sonication and 45 sec rest) at 4 °C for 20 min under N_2 gas. The liposomes proton pumping activity (in case of bR) was measured as described (*217*). Both, bOp (after regeneration by addition of *all-trans* retinal) and the bR are predominantly incorporated in an inside out direction, since light induced alkanization of bulk medium is observed (data not shown).

8.2.2. Fusion of proteoliposomes containing bOp and bR with spehroplasts of Halobacterial cells: a FRET assay

Proteoliposomes containing bOp or bR were fused with spheroplasts by combining bOp/bR proteoliposome solution (2.5 ml proteoliposomes prepared as described above) and spheroplast suspension (7.5 ml of supension containing spheroplasts prepared from *H. salinarium* cells grown in 250 ml). For the FRET fusion assay, 0.8 mol% of N-NBD-PE

and 0.8% of N-Rh-PE (when necessary) were incorporated into the bilayers of proteoliposomes (*422*). Fluorescence measurements were carried out at room temperature in a fusion buffer containing basal salt solution adjusted to pH 7.4. The final volume in the cuvette was 2 ml. Two minutes after adding 25 µl of the proteoliposome solution, 75 µl of the spheroplast solution was added. Fluorescence measurements (excitation at 465 nm, emission at 530 nm) were carried out as a function of time over 1 h. A cut off filter (<515 nm) was placed between the sample and the emission monochromator. To calibrate the fluorescence scale, the initial fluorescence of the liposomes was taken as the zero level, and the fluorescence at maximum probe dilution (determined by addition of 0.5% Triton X-100) was taken as 100% (*422*). FRET fusion assays were performed to determine efficiency of proteoliposome fusion for three systems: a.) N-NBD-PE + bR; b.) N-NBD-PE + N-Rh-PE; c.) N-NBD-PE + bOp + N-Rh-PE.

8.2.3. Fusion of liposome containing *bop* mRNA with *Halobacterial* spheroplasts

1.25 ml of Lipofectin (Life Technologies Inc., MD) complexed *bop* mRNA (2 mg) [Lipofectin to mRNA ratio, 2.5:1] was prepared based on procedures described earlier (*387*), and were mixed with 7.5 ml of suspension containing spheroplasts prepared from *H. salinarium* cells grown in 250 ml.

8.2.4. Protein expression and isolation

After liposome fusions and after appropriate incubations, spheroplast solution were washed with basal salts containing 0.1M $MgCl_2$, the cells were suspended in the growth medium containing peptone (volume same as during the original growth). The cells were grown for 12 hours and harvested by centrifugation (20,000 x g for 30 min) and washed with 4M NaCl. For anaerobic growth, cells were grown on 0.5% arginine and in presence of nitrogen as described (*389*). The pelleted cells were lysed by suspension in distilled water, and the membranous fraction was collected by centrifugation at 30,000 x g for 30 min, and washed with distilled water (3 x 20 ml). SDS-PAGE analysis was done at this point. After the final wash, the pellet was layered on sucrose density gradient

(15-65% (w/v), 25 ml) and centrifuged at 150,000 x g for 16 h (*390*). The purple band of bR was isolated and washed with distilled water (4 x 25 ml) to remove the sucrose. Bacteriorhodopsin thus obtained was characterized using UV-visible spectroscopy and SDS-PAGE. For electrophoresis, equal amounts of cells were lysed with water; aliquots were separated by electrophoresis using 15% SDS-PAGE.

8.2.5. Isolation of isoprenoids from spheroplasts

Isoprenoids are isolated from *H.salinarium* cells as described (*417*). Spheroplasts grown in 100 ml volume were lysed in 1 ml of water and treated with DNAse. Under vigorous stirring, the lysate was added into 9 ml acetone. After 20 min in the dark, 4 ml n-hexane and 1 ml water were added. The upper phase, containing the pigments, were evaporated under vacuum and the pigments were diluted in 100 µl toluene. The sample was applied to an alumina column (16 g, Activity grade II). Elution of β-carotene was carried out by adding 60 ml of 20% diethyl ether in n-hexane, elution of retinal was carried out by further addition of 140 ml of the same solvent mixture. Elution of lycopene was carried out with 40 ml of diethyl ether. Amount of each compound obtained was calculated spectrophotometrically (*423*). The values are expressed on a protein basis (*424*) as micrograms per gram of cellular proteins.

8.3 Results

8.3.1. Overall strategy

Figure 8.2 shows the overall 3-fold experimental strategy employed in this study. First, in order to see if the presence of bOp acts as a trigger for retinal biosynthesis, *bop* mRNA was translated into Pum$^-$ strain using lipofection as a method for *in vivo* introduction of *bop* mRNA. Cells have been lysed to assay the presence of bR and to determine the isoprenoid pigments such as lycopene, β-carotene and retinal. Lycopene is known to be a precursor to retinal (*425*) and its presence has been used to track retinal biosynthesis in halobacteria (*417*). Secondly, in order to see if exogenously added bOp triggers retinal biosynthesis, the apoprotein bOp was introduced into spheroplasts *via* liposome fusion.

Thirdly to distinguish between apoprotein and chromoprotein (retinal containing protein) mediated control, bR was introduced into spheroplasts *via* proteoliposome

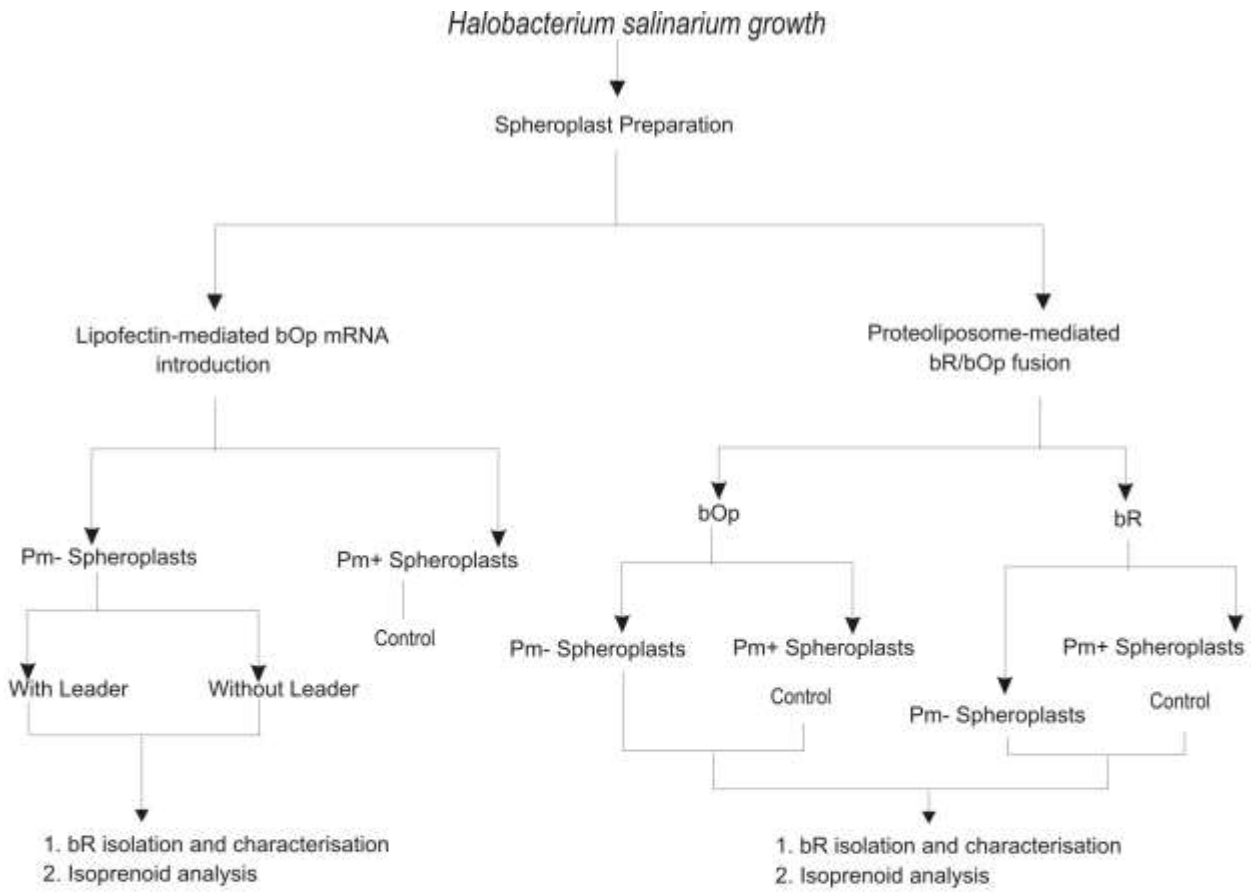


Figure 8.2: Overall experimental approach in the elucidation of the retinal biosynthetic pathway: Three-fold experimental strategy was employed in the study, wherein, *bop* mRNA, bacterio-opsin and bacteriorhodopsin were introduced separately into Pum- strains of *Halobacterial* spheroplasts via liposomes. Each of the cells was analyzed for presence of bacteriorhdopsin and the key isoprenoid pigments such as lycopene, carotene and retinal.

fusion. The state of bR (with or without retinal) and its effect on lycopene content were assessed by isolating bOp/bR and a further pigment analysis.

8.3.2. Spheroplasts of Pum$^-$ strains do not produce bR or bOp

Spheroplasts were prepared according to methods described earlier (*388*, *426*). The viability of spheroplasts was determined by titering them on spheroplast regeneration plates as described earlier (*427*, *428*) and it was found that spheroplasts were viable. Figure 8.3 shows SDS-PAGE analysis of partially purified membranes from spheroplasts. The samples subjected to SDS-PAGE were aliquoted before the sucrose density gradient step (*390*). Figure 8.3A shows clear absence of bOp in spheroplasts of Pum$^-$ origin (lane 2). In contrast, when Pum$^+$ spheroplasts were lysed and subjected to SDS-PAGE in a similar manner, presence of unprocessed and processed bacterioopsin (lanes 3 and 4) was seen as observed earlier (*388*). Absorption spectrum of sample subjected to SDS-PAGE also shows lack of characteristic absorption peak corresponding to bR at 560 nm (see figure 8.3 (inset)) suggesting that spheroplasts of Pum$^-$ strain do not produce bR or bOp.

8.3.3. Liposome mediated introduction of *bop* mRNA into spheroplasts of Pum$^-$ strains results in expression of bOp protein and subsequent formation of bR

Figure 8.3A shows that when lipofectin mediated introduction of *bop* mRNA was carried out in Pum$^-$ spheroplasts, bacteriorhodopsin is produced in correctly folded form as seen from SDS-PAGE (lane 4,5,6) and characteristic 560 nm absorption peak [see Figure 8.3 (inset)]. Sucrose density gradient ultracentrifugation shows that the bOp expression leads to its assembly as purple membrane lattice. In a control experiment, where only lipofectin lacking mRNA was used, no bOp production was observed (Figure 8. 3A, lane 3), thus indicating that the transfection protocol did not induce a latent/preexisting bop gene. Another control in which the *bop* mRNA was lipofected into Pum$^-$ cells and grown in the presence N_2 gas, showed that bR formation does not take place. This is due to the inhibition of the oxidation of β-carotene to retinal. SDS-PAGE analysis of membranes isolated from *bop* mRNA transfected spheroplasts shows presence of processed bR (presence of a small population of unprocessed bR cannot be ruled out from our

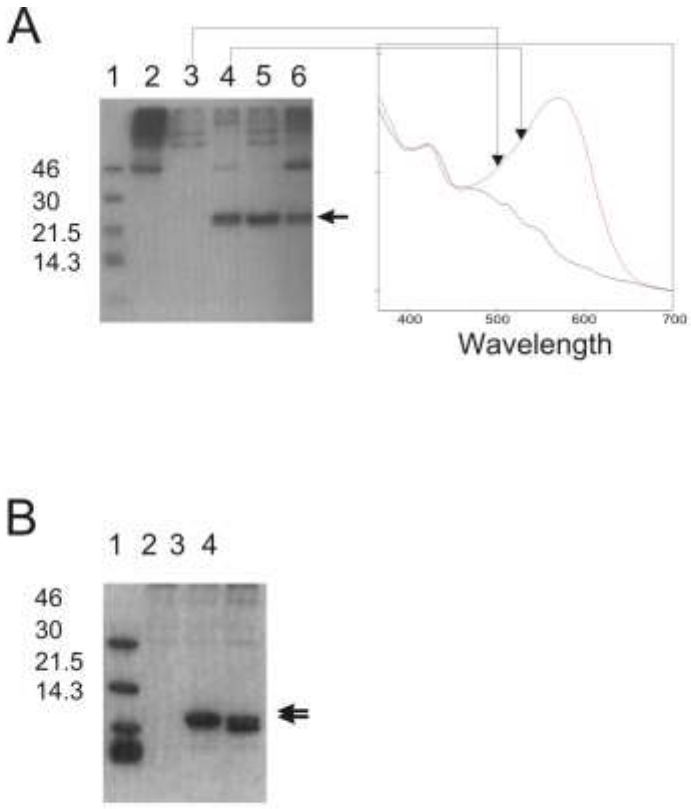


Figure 8.3: Coomassie-Blue stained SDS-PAGE analysis of membrane fractions isolated during bR purification: ***Panel A***. SDS-PAGE of various membrane fractions aliquoted before sucrose density gradient centrifugation step during isolation of bR from spheroplasts. The samples were loaded on 15% polyacrylamide gel followed by Coomassie blue staining. Lane 1: molecular weight markers; Lane 2: membranes isolated from spheroplasts of Pum$^-$; lane 3: membranes isolated from spheroplasts of Pum$^-$ cells treated with lipofectin lacking mRNA (see experimental details for text); Lane 4: membranes isolated from spheroplasts of Pum$^-$ cells treated with lipofectin associated bop mRNA; Lane 5: membranes isolated from spheroplasts of Pum$^-$ cells fused with liposomes containing bacterioopsin; Lane 6: membranes isolated from spheroplasts of Pum$^-$ cells fused with liposomes containing bR. Inset shows visible absorption spectra of the samples.

analysis). This indicates that exogenously added bop mRNA is translated efficiently in spheroplasts of Pum- cells. Figure 8.3A inset shows absorption spectrum of bR formed after lipofection of *bop* mRNA into spheroplasts of Pum- cells. It can be seen that absorption spectrum exhibits characteristic absorption at 560 nm and is similar to native bR (*199*, *217*). The newly isolated bR showed characteristic light-dark adaptation and proton pumping activity (data not shown). Our data shows that bOp translated from *bop* mRNA is integrated into membranes of halobacterial cells and binds to intrinsic retinal to form a functional bacteriorhodopsin.

It has been shown that spheroplasts are incapable of processing immature bR and only when they are returned to normal cells can they process bR to cleave off the leader sequence (*388*). Figure 8.3B shows presence of unprocessed bR in spheroplasts of Pum- cells (lane 3), which were lipofected with *bop* mRNA containing the leader sequence and where the spheroplasts were not allowed to revert to normal cells. Lane 4 shows presence of both unprocessed and processed bR in spheroplasts that were allowed to convert to normal cells. When mRNA lacking the leader sequence was used in lipofection, no detectable bOp expression was found (data not shown). This observation is in agreement with previously reported work (*378*, *429*), where it has been indicated that the leader sequence of *bop* mRNA is essential for binding to the ribosomal binding site in halobacteria. The hairpin structure formed by *bop* mRNA may also add to the intracellular stability of the exogenously introduced mRNA (*378*, *429*).

8.3.4. bOp and bR can be introduced *in vivo* into membranes of spheroplasts of *H. salinarium*

Fusion of proteoliposomes containing bOp and bR was done with spheroplasts of Pum$^-$ strain of *H. salinarium*. The isolation and purification of protein shows that exogenously introduced apoprotein bacterio-opsin, forms the chromoprotein bacteriorhodopsin and gets incorporated into the membranes of halobacteria (Figure 8.3, inset). After reconstitution, the isolated chromoprotein exhibited normal proton pumping activity indicating that the protein is functional.

Liposome fusion process was monitored using FRET. Fluorescent probe labeled lipids (NBD-PE, see materials and methods) were incorporated into liposomes and fused with spheroplasts. Fluorescence of the NBD-PE is quenched by bR due to its strong absorption at 560 nm. A FRET analysis was performed and results indicated that over a period of 1h, the fluorescence intensity increased as the label had moved large distances within the membrane. There was a decrease in the quenching of fluorescence as the distance between the label and the protein was increasing as a function of time (See Figure 8.4). This shows that liposome fusion had taken place with the spheroplast membranes of *H. salinarium*. Fusion efficiency was estimated as per (*430*) and was found to be in the range of 50-60 %.

8.3.5. Lycopene cyclization and subsequent retinal biosynthesis is induced by the presence of membrane bound bOp

Table 8.1 indicates analysis of key isoprenoids found in the retinal biosynthetic pathway. Analysis of Pum$^+$ cells shows characteristic absence of lycopene indicating its efficient conversion to β-carotene which in presence of oxygen, gets spontaneously converted into all-trans retinal (Table 8.1, column 1). In contrast, isoprenoid analysis of Pum$^-$ cells indicates accumulation of lycopene, and absence of β-carotene and retinal (Table 8.1, column 2). When Pum$^-$ cells are lipofected with bop mRNA, pigment analysis indicates that lycopene is consumed to form β-carotene and all-trans retinal in agreement with bR formation (Table 8.1, column 3). Introduction of the bOp protein via proteoliposomes in Pum$^-$ cells also shows an increase in concentration of β-carotene (Table 8.1, column 4).

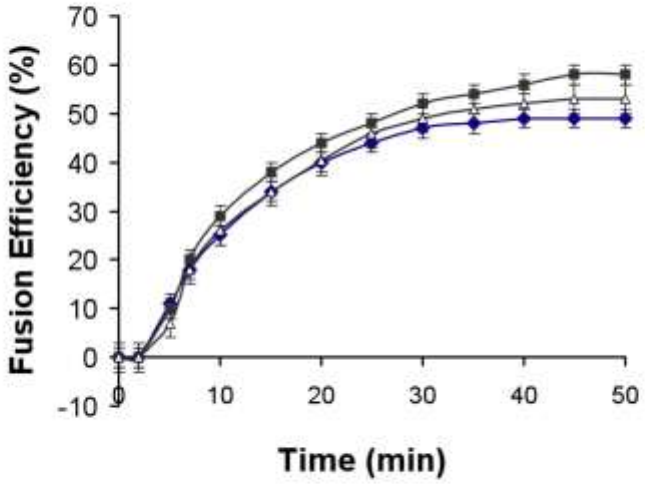


Figure 8.4: FRET analysis of fusion between spheroplasts of Pum- cells of *H. salinarium* and liposomes containing NBD-PE and bR. Values for the calculated from observed increase in NBD-emission intensity as described (430) (See also Materials and Methods). Time dependence fusion efficiency is shown for liposomes containing N-NBD-PE and Pum- spheroplasts (■); N-NBD-PE + N-Rh-PE + bOp and Pum- spheroplasts (▲) and, N-NBD-PE + bR and Pum- spheroplasts (◆).

8.3.6. Lycopene cyclization is inhibited by presence of bR

When spheroplasts of Pum$^-$ cells were fused with liposomes containing bacteriorhodopsin (bR), the protein was found to be incorporated into membranes of the regenerated cells. Pigment analysis of these cells showed similar isoprenoid composition as Pum$^-$ cells (Table 8.1, column 5), indicating that lycopene accumulation and hence inhibition of lycopene to β-carotene conversion by bR. The exogenously introduced bR was isolated and was found to be functionally unaltered when checked for proton pumping (data not shown).

8.3.7. Exogenously added retinal does not inhibit lycopene cyclization

It has been established that conversion of β-carotene to retinal is a spontaneous and an oxygen dependent process and is inhibited in anaerobic conditions (*431*). Under anaerobic conditions, β-carotene should accumulate and retinal should be absent. When spheroplasts of Pum$^-$ cells were lipofected with *bop* mRNA and grown under anaerobic conditions, although bOp induced lycopene to β-carotene conversion is seen (Table 8.1, column 6), β-carotene to retinal conversion is expected to remain inhibited. When excess retinal in introduced in Pum$^-$ cells lipofected with *bop* mRNA under anaerobic conditions, excess retinal does not seem to inhibit lycopene to β-carotene conversion (Table 8.1, column 7) as seen from increased β-carotene content in contrast to data shown in Table 1, column 8. It should be noted that if bR is introduced under similar anaerobic conditions in presence of excess retinal, no β-carotene is formed (Table 8.1, column 9), indicating that bR and not excess retinal is inhibiting lycopene to β-carotene conversion. In a control experiment where only lipofectin lacking mRNA was used for lipofection, no bR production was observed, thus ruling out possibility of lipofection protocol induced lycopene cyclization.

Table 1: Analysis of terminal isoprenoids in the retinal biosynthetic pathway

Isopentenyl-PP → →→ **Lycopene → β-carotene→ *all-trans* retinal**

	Pum^{+}-spheroplasts 1	**Pum^{-}-spheroplasts** 2	**Pum^{-}-spheroplasts lipofected with *bop* mRNA** 3	**Pum^{-}-spheroplasts fused with proteo-liposomes containing bacterio-opsin** 4	**Pum^{-}-spheroplasts fused with proteo-liposomes containing bacterio-rhodopsin** 5	**Pum^{-}-spheroplasts lipofected with *bop* mRNA, grown under anaerobic conditions** 6	**Pum^{-}-spheroplasts containing excess retinal, lipofected with *bop* mRNA, grown under anaerobic conditions** 7	**Pum^{-}-spheroplasts containing excess retinal, grown under anaerobic conditions** 8	**Pum^{-}-spheroplasts containing excess retinal, fused with proteo-liposomes containing bacterio-rhodopsin, grown under anaerobic conditions** 9
Lycopene	Trace	1700 (±30)	trace	trace	1600	1700 (±30)	1500 (±30)	1800 (±100)	1380 (±180)
β-carotene	560 (±30)	0	480 (±40)	450 (±34)	trace	1430 (±30)	1480 (±10)	trace	trace
Retinal	44 (±3) {54}[(b)]	0	40 (±8)	24 (±5) {27}[(c)]	25 (±5) {27}[(c)]	trace	1148 (±120) (*432*)	1250 (±180)	1480 (±210)
bR formation	Yes	No	Yes	Yes	Yes	No	Yes	No	Yes

(a) estimated according to Materials and Methods, expressed as μg of pigment per g of cellular proteins; (b) Value of retinal content expected from the isolated bacteriorhodopsin content; (c) Value of retinal expected from averaged 50% fusion efficiency of bOp or bR containing proteoliposome fusion.

8.4 Discussion

8.4.1. Elucidation of the retinal biosynthetic pathway

Purple membrane, a differentiated domain of the plasma membrane contains only one protein species, bacterioopsin, which is complexed with retinal in a 1:1 stoichiometry. Biogenesis of purple membrane is inducible by limiting oxygen supply (*418*) which turns on synthesis of both, bacterioopsin and retinal (*431*). In contrast, most of the lipid molecules necessary for purple membrane formation are synthesized long before the start of bacterioopsin and retinal synthesis, i.e. lipids are taken from the pool of the cell membrane (*431*).

Essentially most of the retinal present in halobacteria is associated with bacterioopsin, so one observes *nearly stoichiometric relation between bacterioopsin and retinal content*. This fact suggests the existence of a highly efficient regulation-mechanism co-ordinating the bacterioosopin and retinal synthetic pathways. In a classic piece of work by Sumper and Herrmann (*431*), it has been found that while bOp may be the trigger for retinal biosynthesis from its precursors, it is unclear which of the two: excess retinal or bacteriorhodopsin in the membrane, that generates feedback signal to stop retinal biosynthesis.

Following questions emerge on events that trigger retinal biosynthesis (Figure 8.5) out of existing data:

1. Which of the two steps, lycopene to β-carotene or β-carotene to retinal, is triggerred by bacterioopsin?
2. Which of the two events, bop transcription or bop translation, acts as a trigger of retinal biosynthesis?
3. Instead of possibility 2, if protein product bOp, is the trigger; which form of the protein, unfolded nascent polypeptide or folded polypeptide, acts as a trigger of retinal biosynthesis?

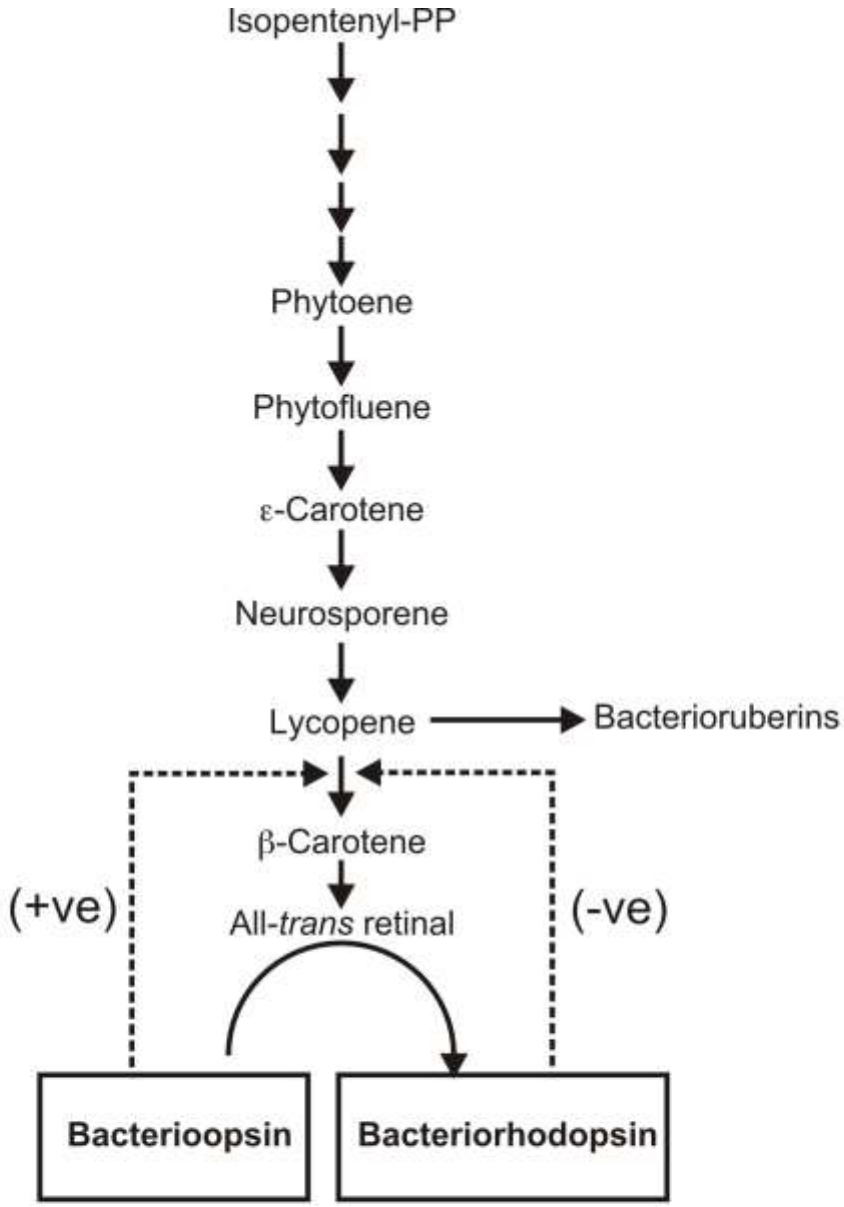


Figure 8.5: Schematic representation of retinal biosynthesis and its regulation by bacterioopsin and bacteriorhodopsin: Dotted lines indicate effect where bR or bOp acts in inhibitory (-ve) or stimulatory (+ve) manner on the biosynthetic pathway.

Similarly following questions emerge for trigger that inhibits or stops retinal biosynthesis:

1. Which of the two, excess retinal or bR, acts as a trigger for inhibition?
2. At what stage does the feedback signal operate, e.g. at lycopene to β-carotene conversion or β-carotene to retinal conversion?

It should also be noted that while bacterioopsin resides in brown membrane after translation and membrane insertion, it forms trimeric units only after retinal binding. The latter process is assumed to be under energy requirement (*431*). It is unclear if events associated with trimeric unit formation are involved in inhibition of retinal biosynthesis.

In the selected Pum$^-$ strain bOp, bR and retinal are absent. Both endogenously expressed as well as exogenously introduced bOp get converted into the retinal bound form, bacteriorhodopsin. Concomitant decrease in the lycopene content and formation of retinal suggests that the presence of the apoprotein bOp has triggered the conversion of lycopene to β-carotene and hence the formation of retinal. It should be noted that this trigger operates at the lycopene to β-carotene conversion and not the β-carotene to retinal. It has been established by nicotine binding and oxygen-dependence studies that β-carotene to retinal conversion is spontaneous (*417*, *431*).

The fact that *bop* mRNA as well as bOp have triggered lycopene to β-carotene conversion, it indicates that the trigger does not lie in the events associated with transcription or translation of the bop gene. The proteoliposome fusion is known to be a membrane associated with process (*419*) and the extreme hydrophobic nature of bOp rules out its direct introduction into the cytoplasm. Studies of Driessen and Konings (*419*) and our regeneration experiments on proteoliposomes suggest that bOp is in the correctly folded form. Thus, bOp-mediated retinal formation is not associated with the unfolded nascent polypeptide present in the cytoplasm. Instead, the trigger is clearly associated with the folded form of bOp and that too in the membrane integrated form. On the other hand, the fact that lycopene content remains unaltered after introduction of

chromoprotein bR suggests that the presence of bR triggers the inhibition of retinal biosynthesis.

In contrast, the trigger signaling inhibition of retinal biosynthesis does not lie with presence of excess retinal. When Pum- strain containing excess retinal was allowed to grow anaerobically in presence of bOp, it did not inhibit lycopene to β-carotene conversion. Thus we conclude that it is the presence of bR that inhibits and then regulates lycopene to β-carotene conversion and imparts stringency to retinal biosynthesis.

At present, neither the mechanism for how this triggering occurs or how retinal biosynthesis is initiated is clear. It is not clear, if the trigger lies with trimeric unit formation or not (check the data). It is also not clear what role, if any, do halorhodopsin and/or sensory rhodopsin play in retinal biosynthesis.

This is a first report of *in vivo*, liposome-mediated introduction of macromolecules for determination of factors regulating biosynthesis of cofactors and their binding to proteins. The approach is simple and novel and involves the *in vivo* introduction of mRNA and membrane apoprotein into spheroplasts lacking the protein as well as the cofactor. Such a study could provide information about the regulation of the pathway involved in the synthesis of its cofactor.

Chapter 9: Incorporation of Non-native Amino Acids into Bacteriorhodopsin, an integral membrane protein

9.1 Introduction

In this chapter, a new and potentially general technology that involves a large-scale, *in vivo* fusion of liposome-encapsulated misaminoacylated-suppressor transfer RNA along with the amber-codon containing target messenger RNA into a spheroplastic bacterial cell population has been described. As an application of the developed technology, SNAAR analogs of bacteriorhodopsin containing fluorescent amino acids are synthesized on a large scale. In this chapter, we demonstrate site-specific incorporation of fluorescent probes followed by fluorescence resonance energy transfer (*433*) (FRET) measurements (SNAAR coupled FRET). Using such SNAAR analogs, we demonstrate that detailed information regarding the nature of structural changes within a protein molecule during its function can be obtained.

We have chosen bacteriorhodopsin from *Halobacterium salinarium* (*434*) as a model protein for application of *in vivo* SNAAR. Bacteriorhodopsin, a 26 kDa protein, is found in the archeabacterium *Halobacterium salinarium*. The apoprotein bacterioopsin (bOp) is complexed with all-trans retinal to form the chromoprotein bacteriorhodopsin (bR). Retinal is linked to Lys-216 of the bOp via a Schiff base linkage. bR consists of seven closely packed trans-membrane alpha-helices. It acts as a light activable proton pump translocating two protons per photon of light absorbed (*435*). The proton motive force generated due to this is used for active fuel transport and ATP synthesis.

Bacteriorhodopsin (bR) undergoes a light-induced photocycle characterized by K, L, M, N and O intermediates, which ultimately results in proton transfer across the membrane (*435*). Upon photon absorption, light-adapted bR (bR_{570}) gets converted to its first metastable intermediate K. This K-state undergoes thermal relaxation to form L-state. A proton is released into the external medium during the L→M formation. M is converted into N intermediate. A second proton is taken up from the inner medium during O formation. O-decay results in the regeneration of bR, which then relaxes to

have the original conformation of the chromophore. Although this photcycle of bR has been studied by a number of biophysical techniques (*51*, *258*, *264*, *266*, *435-437*), the exact location and nature of gross protein conformational changes responsible for vectorial proton transfer remain unknown.

We have determined structural changes between bR and its key photocycle intermediate, M, using *in vivo* SNAAR coupled FRET. The experimental strategy used for this was as follows. To probe the helical movement during the functioning of bR, fluorescent donors were introduced at relatively immobile, membrane-embedded cytoplasmic and extracellular end-points of helices C, F and G (total 6) in a site-specific manner by introduction of amber codons (one at a time) into the *bop* gene (see Figure 9.1). Retinylidene Schiff base chromophore served as an intrinsic acceptor in these studies. Secondly, in order to probe intra-helical conformational changes, a fluorescent donor and an additional acceptor were simultaneously introduced at cytoplasmic and extracellular ends of each helix (see figure 9.1B). Incorporation of two non-native amino acid residues was achieved by introduction of amber and cysteine codons (since primary sequence of bR lacks cysteine) into the *bop* gene at codons corresponding to end-points of each helix (one helix at a time) followed by amber suppression combined with cysteine-directed chemical modification. The pairs of fluorescent donors and acceptors used in this study are (a) N-(7-dimethylamino-4-methylcoumarin-3-yl)maleimide (DACM) (Donor): retinal (Acceptor); (b) Iodosalicylic acid, (ISA) (Donor): retinal (Acceptor), and (c) DACM (Donor): N-(4,4-difluoro-5,7-dimethyl-4-bora-3a,4a-diaza-s-indacene-3-yl)methyl-N-(2-aminoethyl)maleimide(BODIPY) (Acceptor). Figure 9.2 briefly outlines the key steps and experiments performed to achieve *in vivo* SNAAR.

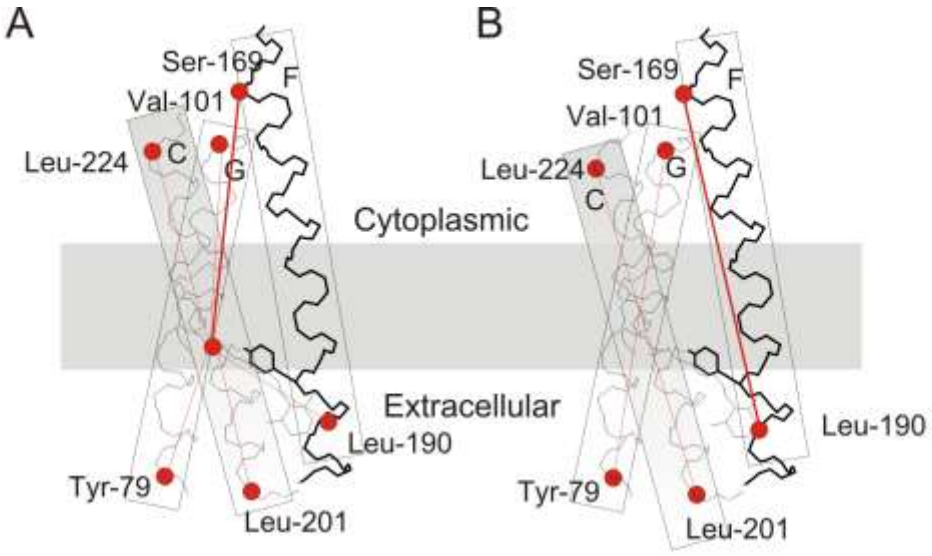


Figure 9.1: Schematic representation of the secondary structure of bR showing sites of modification for measurement of a) interhelical distances from end-point of each helix to retinyledene Schiff base and b) intrahelical distance; i.e. Between cytoplasmic and extracellular end-points

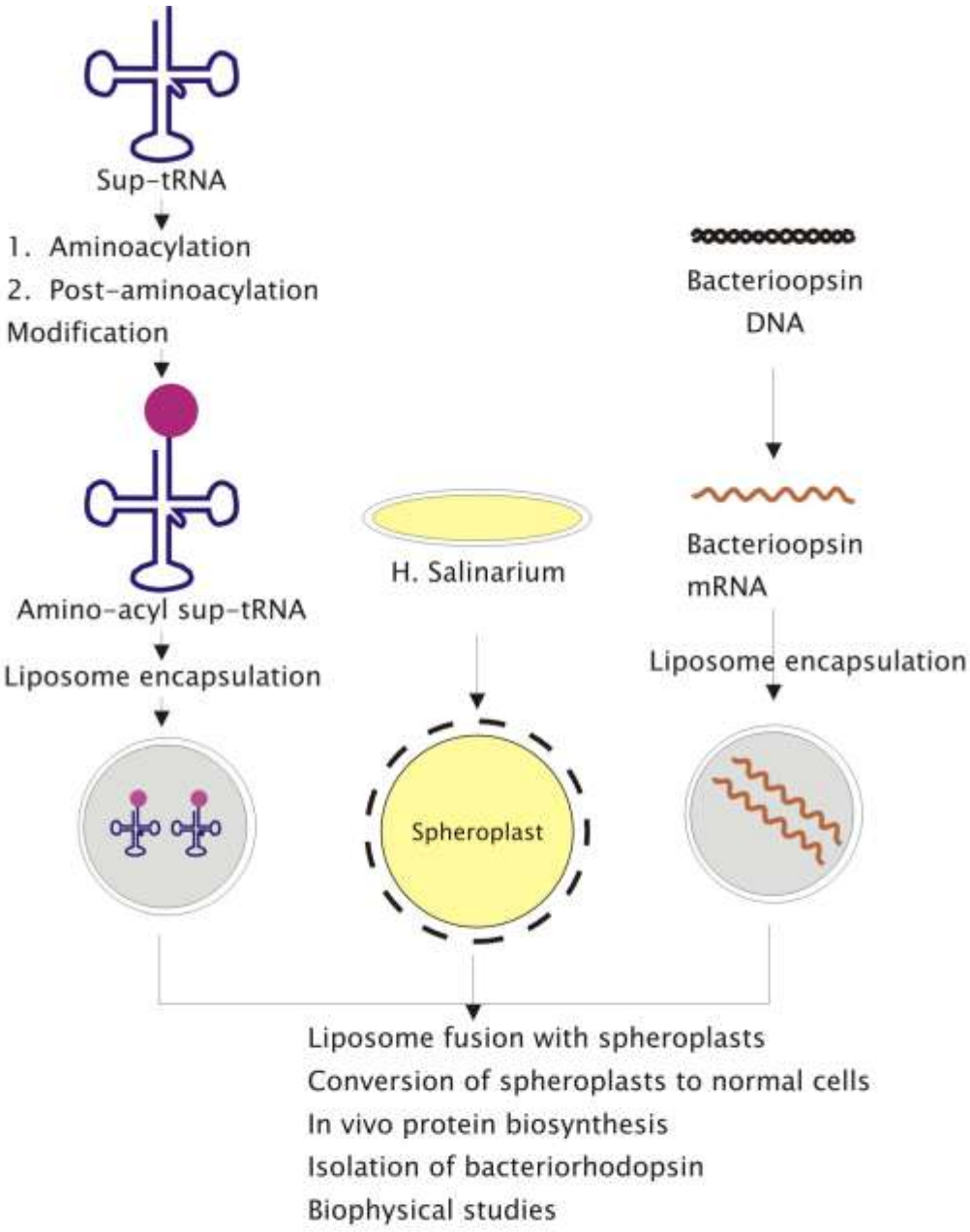


Figure 9.2: Schematic representation of the experimental strategy designed to achieve *in vivo* SNAAR. The critical elements involved are the misaminoacylated suppressor tRNA, suppressor codon containing mRNA and *Halobacterial* spheroplastic cells. The method of delivery of macromolecules into the cells is *via* lipofection.

The main experimental work elements performed were as follows:

1. Lipofections of non-native aminoacylated tRNA and amber codon containing *bop* mRNA into spheroplastic cells of *H.salinarium*
2. Expression and isolation of the SNAAR analog of bR:
3. Characterization of the SNAAR analog of bR:
   a) Proton pumping
   b) Fluorescence measurements
   c) V8 Protease digestion
4. Distance measurements using FRET as a spectroscopic ruler

9.2 Materials and Methods

9.2.1. *In vitro* transcription of tRNAs and mRNAs

As discussed in chapters 6 and 7, *in vitro* transcription of tRNA and mRNA was carried out. Both tRNA constructs, pSP65-C, and pSP65-C-sup, were linearized and digested with *Bst*NI. *bop* gene construct was linearized with *Hind*III. Run-off transcripts of linearized DNA template were prepared using T7 RNA polymerase (Bangalore Genei, Bangalore, India) as described (*90*, *108*).

9.2.2. Enzymatic aminoacylation reactions

As discussed in chapter 7, aminoacylation of tRNA was carried out. Aaminoacylation reactions were performed according to Sonar *et al* (*199*). For kinetic parameter determinations, tRNA concentration was varied from 0.2 μM to 1.5 μM with fixed *E.coli* S-100 concentration. K_m and V_{max} were calculated using ^{35}S-Cysteine from Michealis-Menten plots (*409*).

9.2.3. Postaminoacylation modification of Cys-tRNA^{Cys} and Cys-tRNA$^{Cys}_{CUA}$

As discussed in chapter 7, modification of aminoacylated tRNA was carried out. Typically 30 A_{260} units of extensively dialysed (essentially free of DTT) aminoacylated tRNA (2 nmol Cys) were dissolved in 0.3 ml ice-cold 25 mM phosphate buffer (pH 6.0)

and added to a 0.6 ml of prefrozen solution of N-(7-dimethylamino-4-methylcoumarin-3-yl)maleimide (DACM) (Molecular Probes, Eugene, OR) in DMSO. Ratio of tRNA to DACM is maintained at 1:10. Reaction was allowed to continue for 4 h at 4 °C and at the end of reaction, excess β-mercaptoethanol was added to scavange thiol reactive reagent. After additional 30 min, the reaction was stopped and "misaminoacyl-tRNAs" thus obtained were precipitated and resuspended in 5 mM NaOAc, pH 5.5. A similar procedure was adapted for ISA and BODIPY modification of $tRNA^{cys}_{CUA}$ and $tRNA^{cys}$ respectively. The products were characterized by fluorescence measurements (DACM: excitation at 383 nm, emission at 463 nm; ISA label: excitation at 355 nm and emission at 407 nm, BODIPY: excitation at 504 nm, emission at 530 nm). Labeled aminoacyl-tRNA linkage is stable for >48 hours at 25 °C and for weeks at -20 °C.

9.2.4. Lipofection of spheroplasts using bop mRNA and aminoacylated $tRNA_{CUA}$

Cell growth, preparation of spheroplasts of Pum^- strain of *H.salinarium* (lacking bR expression) and their lipofection were carried out as described (*388*, *438*) with following minor modifications. 1.25 ml of Lipofectin (Life Technologies Inc., MD) complexed *bop* mRNA (2 mg) [Lipofectin to mRNA ratio, 2.5:1] and 1.25 ml of lipofectin complexed with postaminoacylation SH-modified tRNA (e.g. DACM-Cys-$tRNA^{Cys}_{CUA}$) (2.5 mg tRNA) [Lipofectin to tRNA ratio, 2.5:1] was prepared based on procedures described earlier in chapter 6 and 7 (*387*), and were mixed with 7.5 ml of suspension containing spheroplasts prepared from *H. salinarium* cells grown in 250 ml.

9.2.5. Protein Isolation and characterization

SNAAR analogs of bR were isolated as described (*438*) and were characterized using UV-visible spectroscopy, proton pumping and SDS-PAGE (*108*, *199*). Further characterization of the protein to check for site-specific incorporation of modified cysteine was done by digestion of the protein using the V-8 protease (*439*). The SNAAR analog of bR was lyophilized and suspended in 2 ml 50 mM NaPi, 2mM EDTA (pH 7.8). SDS was added to give a final concentration of 0.25% (mass/vol.). Proteolytic digestion

was initiated by the addition of *S.aureus* V8 protease (Sigma) at a ratio of 0.03 mg protease/ mg bacteriorhodopsin. Proteolysis of the stirred suspension was carried out at 37°C for 18 h and terminated by boiling (for SDS-PAGE analysis). The gel obtained was cut into fine pieces of ~5 mm and the bands were eluted using the Hoeffer gel elctroeluter. The eluate was measured for fluorescence intensity.

9.2.6. Proton pumping assay of bR

Proton translocation activity was measured at room temperature as described previously (*440*). SNAAR analogs of bR were reconstituted into asolectin vesicles using a modification of the octyl glucoside dilution procedure of Racker et al (*386*). This involved preparation of asolectin lipid vesicles (L-α-phosphatidylcholine, Type II-S, Sigma) by hydrating a film of the lipid (dried from a chloroform solution) with water and sonicating to clarity (approximately 15-30 min) under nitrogen. Insoluble material was removed by a quick centrifugation in an eppendorf centrifuge. To 30 μl of the asolectin lipid suspension, 10 μl of 12.5% (w/v) octyl glucoside (Boehringer Manheim) plus 40 μl of the SNAAR analog (100 pmol) was added, and the solution was incubated on ice for 5 min. This sample was taken in a test tube and placed in front of the slit of a box containing a 40 W bulb. The light was made to pass through 5% $CuSO_4$ solution for about 15-20 mins. A pH electrode was dipped into the solution and changes in pH were monitored as a function of time. pH was measured in the on-state and off-state to the light and the initial rate (H^+/bR/s) and steady state (H^+/bR) was calculated. The system was calibrated 2-5 μl injections of 1 mM HCl/NaOH.

9.2.7. Cysteine chemical modification of SNAAR analog of bR

Purified SNAAR analog of bR containing engineered cysteines were further subjected to cysteine-directed chemical modification by incubating the purified SNAAR analog along with BODIPY-maleimide (1:4 molar ratio, in 50mM Tris buffer (pH 7.0), 150 mM KCl, 4 μM EDTA) for 1.5 h at room temperature. The efficiency of modification of cysteines in SNAAR analog was found to be almost 94% based on extinction coefficient of BODIPY.

9.3 Results

9.3.1. A new method for non-native aminoacylation of $tRNA_{CUA}$

In this study, we have chosen a simpler route for generation of non-native aminoacyl linked amber suppressor tRNA (i.e. "misaminoacylated" $tRNA_{CUA}$) based on post-aminoacylation modification, instead of the currently used "chemical misaminoacylation" methods. Although, the method used here introduces non-native moieties as cysteine side-chain modifications, it avoids several synthetic organic chemistry based steps involved in chemical aminoacylation such as: (a) dinucleotide protections, b) synthesis of carboxyl activated and N-protected non-native amino acid, and subsequent aminoacylation of the protected dinucleotide (c) T4 RNA ligase based ligation of truncated tRNA with aminoacyl dinucleotide, and finally (d) photo-deprotection of the entire misaminoacylated suppressor tRNA (*191*). In our approach, $tRNA^{Cys}$ and $tRNA^{Cys}_{CUA}$ were prepared using *in vitro* run-off transcription (*410*). Cysteine aminoacylation of both transcripts have shown comparable K_m (3.5 vs 7.15 μM, respectively) and V_{max} (0.54 vs 0.54 μmoles/l/min). Non-native functionalities (DACM, ISA and BODIPY) are introduced as post-aminoacylation modifications of SH-groups of the cysteine present on aminoacyl tRNA. The extent of post-aminoacylation modifications is determined to be >95% using sulphydryl group determination and hydrolysis of aminoacyl-tRNA linkage (*411*).

9.3.2. Liposome mediated introduction of bop mRNA and misaminoacylated $tRNA_{CUA}$ into spheroplasts of Pum- strains results in expression of bacteriorhodopsin containing non-native amino acid

Misaminoacylated suppressor tRNAs and the amber codon containing mRNA encoding bacterio-opsin (*bop)* were independently encapsulated into cationic lipids (lipofectin) and were simultaneously fused with spheroplasts of *H.salinarium* cells that inherently lacked bR expression (Pum^-). Figure 9.3(a) summarizes *in vivo* fate of aminoacylated

tRNA after fusion and during the expression of bacteriorhodopsin. Results show that aminoacylated native $tRNA^{Cys}$ introduced within spheroplasts *via* lipofection is

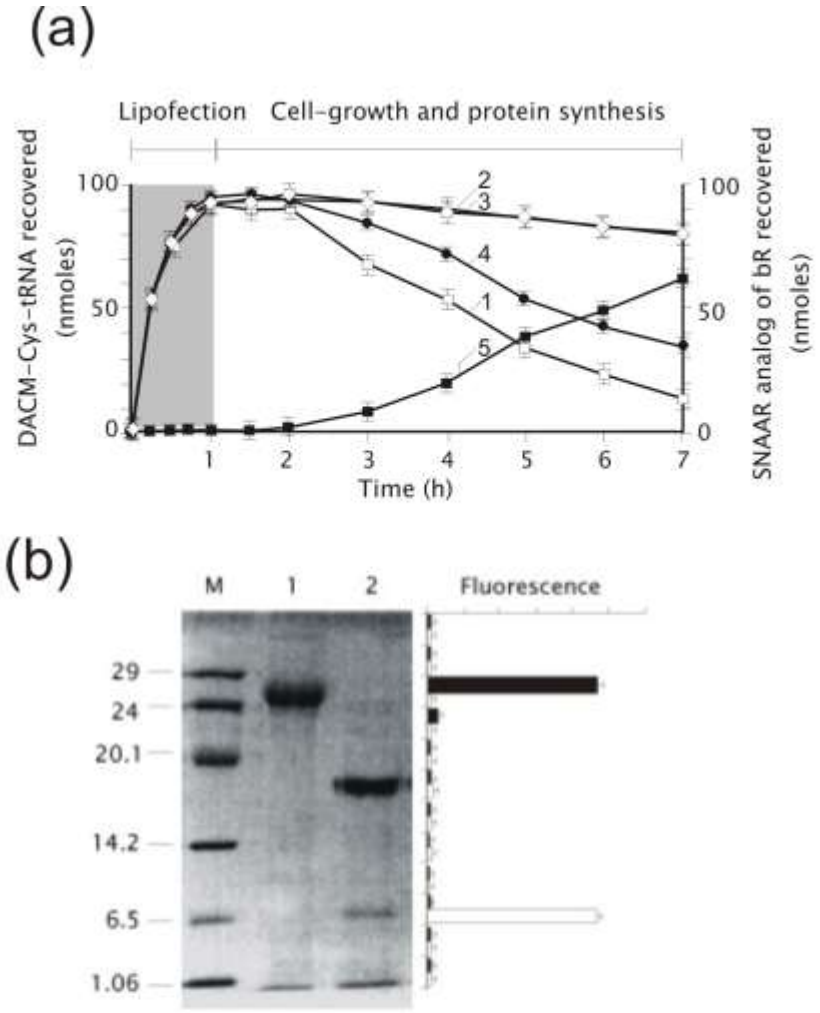


Figure 9.3: (a) Fate of misaminoacylated suppressor tRNA lipofected into spheroplasts of Pum-cells of H. salinarium during *in vivo* SNAAR of bR. Levels of misaminoacylated tRNA was estimated in the total tRNA isolated under acidic conditions 40 at various time intervals and quantitated using fluorescence (recovery efficiency was ~93-95%). (1) DACM-Cys-tRNA^{Cys} without *bop* mRNA, (2) DACM-Cys-tRNA$^{Cys}_{CUA}$ without *bop* mRNA, (3) DACM-Cys-tRNA$^{Cys}_{CUA}$ with WT *bop* mRNA, (4) DACM-Cys-tRNA$^{Cys}_{CUA}$ with S169am-*bop* mRNA and (5) SNAAR analog of bR (S169C-DACM) synthesized. **(b) Characterization of SNAAR analogs of bR.** SDS-PAGE analysis of S169am-ISA before (lane1) and after (lane 2) V-8 protease digestion.

consumed within 7 hours as spheroplasts revert and continue protein biosynthesis. On the other hand, levels of misaminoacylated tRNA$^{Cys}_{CUA}$ do not change indicating that it is not utilized for any other translation, unless amber codon containing *bop* mRNA is co-introduced. Its consumption in response to amber codon is found to be concomitant to production of SNAAR analog of bR showing an approximately 60% efficiency of amber suppression. Spheroplasting cells were reverted to normal cells and expressed SNAAR analogs of bR were purified using sucrose density gradient as described earlier (*390*, *438*). Typically 1.6 mg of SNAAR analogs of bR were obtained from 100 nmoles of misaminoacylated tRNAs used.

For obtaining double SNAAR analogs, spheroplastic cells were simultaneously fused with two populations of liposomes containing DACM-Cys-tRNA$^{Cys}_{CUA}$ and BODIPY-Cys-tRNA^{Cys} along with *bop* mRNA containing engineered amber and cysteine mutations*. Efficiency of incorporation of non-native amino acid (BODIPY-Cys) in response to cysteine codon was found to be poor probably due to the competition of cellular Cys-tRNA^{Cys} with BODIPY-Cys-tRNA^{Cys}. Therefore, to obtain complete modification of the engineered cysteine, each double SNAAR analog was further subjected to cysteine-directed chemical modification with BODIPY-maleimide

9.3.3. Non-native amino acids are introduced in a site-specific manner

Site-specificity of incorporation of fluorescent non-native amino acids was tracked by detecting their fluorescence in proteolytically degraded fragments of single SNAAR analogs. SNAAR analogs of bR were subjected to proteolytic digestion using V-8 protease and subsequent SDS-PAGE (Figure 9.3b). V-8 digestion cleaves bR as two fragments (1-166 (17kDa) and 167-232/4 (6.7 kDa)) (*439*). Presence of fluorescence in only one of the bands shows site-specificity of incorporation. For example, in case of

* Such misaminoacylated tRNA^{Cys} mediated cysteine chemical modification could prove useful in cases where uniquely engineered cysteines remain inaccessible to subsequent protein modification.

S169am-ISA, total fluorescence was detected in the smaller 6.7kDa band as expected. Purified SNAAR analogs were further characterized for their absorption and proton pumping properties and did not show significant perturbation as compared to their natural phenotype.

9.3.4. Characterization of the bR mutants and their SNAAR analogs in terms of absorbance and proton pumping

Wild type bR as well as its mutants were characterized for it's functioning in terms of absorption and proton pumping activity. The data has been tabulated in Tables 9.1-9.5. The data shows that the SNAAR analogs of bR are not significantly different from the wild type bR. Similarly, the SNAAR analogs of D96N are also not found to be significantly different in terms of absorption characteristics and proton pumping activity with respect to D96N. Also it can be seen that modifications such as DACM and ISA do not seem to alter the absorption and proton pumping characteristics. Thus, we could conclude that the protein structure does not seem to alter and hence the functioning of the protein is also unaltered. This is further confirmed from the distance calculations that have been tabulated in Tables 10.1 to 10.7 (for detailed discussion see Chapter 10).

Table 9.1: Absorption and proton pumping characterisitics of DACM-cysteine containing SNAAR analogues of bR.

Sr. No.	Mutant	Absorption data		Proton Pumping (%)
		DA	LA	
1	L201amcysDACM	561	571	98
2	L224amcysDACM	561	570	100
3	L190amcysDACM	562	572	98
4	S169amcysDACM	562	570	100
5	V101amcysDACM	561	571	110
6	Y79amcysDACM	561	570	101

Table 9.2: Absorption and proton pumping characterisitics of DACM-cysteine containing SNAAR analogues of D96N.

Sr. No.	Mutant	Absorption data		Proton Pumping (%)
		DA	LA	
1	L201amcysDACM	562	571	9
2	L224amcysDACM	563	572	10
3	L190amcysDACM	561	571	12
4	S169amcysDACM	561	570	11
5	V101amcysDACM	561	572	14
6	Y79amcysDACM	562	570	10

Table 9.3: Absorption and proton pumping characterisitics of ISA-cysteine containing SNAAR analogues of D96N.

Sr. No.	Mutant	Absorption data		Proton Pumping (%)
		DA	LA	
1	L201amcysISA	562	570	9
2	L224amcysISA	563	572	10
3	L190amcysISA	561	571	12
4	S169amcysISA	561	570	11
5	V101amcysISA	561	572	14
6	Y79amcysISA	562	570	10

Table 9.4: Absorption and proton pumping characterisitics of DACM-cysteine and BODIPY-cysteine containing SNAAR analogues of bR.

Sr. No.	Mutant	Absorption data		Proton Pumping (%)
		DA	LA	
1	L224amcysDACM/ L201cys-BODIPY	561	571	100
2	V101amcysDACM/ Y79cys-BODIPY	562	571	98
3	S169amcysDACM/ L190cys-BODIPY	562	570	99

Table 9.5: Absorption and proton pumping characterisitics of DACM-cysteine and BODIPY-cysteine containing SNAAR analogues of D96N.

Sr. No.	Mutant	Absorption data		Proton Pumping (%)
		DA	LA	
1	L224amcysDACM/ L201cys-BODIPY	562	570	10
2	V101amcysDACM/ Y79cys-BODIPY	561	572	10
3	S169amcysDACM/ L190cys-BODIPY	560	570	9

9.4 Discussion

Non-native functionalities were introduced into bacteriorhodopsin *in vivo* by using **Site-directed Non-native Amino Acid Replacement or SNAAR**. These unnatural moetites were introduced by modification of cysteines after aminoacylation of tRNAs. Such misaminoacylated tRNAs were able to successfully deliver the modified analogs of cysteine into the protein in a site-specific manner. The bacteriorhodopsin analog thus formed, contained the modified cysteine with labels such as DACM, ISA and BODIPY. These analogs of bacteriorhodopsin were completely characterized for its functioning for absorption and proton pumping. The modifications seemed to have no effect on the functioning of bacteriorhodopsin. The analogs of bacteriorhodopsin showed similar proton pumping and absorption characteristics as the wild type bR. Similarly, the analogs of the D96N mutant of bR showed similar patterns of absorption and proton pumping as the D96N mutant itself. This showed that modifications such as these do not cause any fundamental change in the structure and function of the proteins. Such SNAAR analogs would thus closely resemble the actual protein, and its behavior could be extrapolated to the native protein itself.

Flourescent probes such as DACM, ISA and BODIPY that were used to modify cysteines and delivered site-specifically into the protein bacteriorhodopsin, can be used for structure-function studies. FRET pairs such as DACM-retinal (from bR), ISA-retinal

(from bR), ISA-retinal (from D96N bR) and DACM-BODIPY (from bR) were used to map distances from helical endpoints. D96N analog allowed to trap bR in the M-state and thus distance measurements in the D96N state would closely resemble the M-state of bR. Chapter 10 discusses such an application of SNAAR analogs of bR.

Chapter 10: Fluorescence Resonance energy Transfer (FRET): a Spectroscopic Ruler for Biomolecules

10.1 Introduction

Fluorescence Resonance Energy Transfer (FRET) is the radiationless transmission of energy quantum from its site of absorption to the site of its utilization in the molecule, or system of molecules, by resonance interaction between chromophores, over distances considerably greater than interatomic, without conversion to thermal energy, and without donor and acceptor coming into direct collision. The donor is the dye that initially absorbs the energy, and the acceptor is the chromophore to which the energy is subsequently transferred. If a single fluorescence donor and a single absorbance acceptor exist within an assemblage (or a macromolecule), distance between donor and acceptor can be measured. Using FRET one could measure distances within 8 to 60Å.

For fluorescence resonance energy transfer to occur,

- the donor probe must have a high *quantum yield*
- the emission spectrum of the donor probe must overlap considerably with the absorption spectrum of the acceptor probe (i.e. *overlap integral*; *J*,)
- there is an appropriate alignment of the absorption and emission moments and their separation vector (i.e. *kappa square;* κ^2*, or orientation factor*)
- Donor and acceptor must be within 1 ± 0.5 *Ro* from each other (*Ro*, *Förster's distance*).

10.1.2. Some definitions

10.1.2.1. Donor quantum yield

The quantum yield (Q) of the donor is defined as the ratio of the number of photons emitted to the number absorbed a parameter which depends on the immediate environment of the probe.

10.1.2.2. FRET Efficiency

FRET Efficiency (E) can be obtained by measuring the fluorescence intensities of the donor with the acceptor (I_{DA}) and without the acceptor (I_D). FRET efficiency can also be measured using the lifetime of the donor in presence of acceptor (τda) and in the absence of the acceptor probe (τd).

$$E= (1-I_{DA}/I_D) = 1 - \tau_{DA}/\tau_D \qquad \text{------ (Equation 1)}$$

where I_{DA} and τ_{DA} are the intensity and lifetime respectively, of the donor in the presence of the acceptor, and I_D and τ_D, in the absence of the acceptor.

10.1.2.3. Fő́rster distance

The relationship between the transfer efficiency (E) and the distance between the two probes is given by the equation:

$$E = 1/ (1+R^6/Ro^6) \qquad \text{------ (Equation 2)}$$

where R is the distance between donor and acceptor and Ro is the Förster distance which is the distance between donor and acceptor probe at which 50% of the energy is transferred

Ro (Förster's distance) can be calculated using the formula:

$$Ro=[(8.79 \text{ x } 10\text{-}5) JQ_D n^{-4}\kappa^2]^{1/6} \text{ (Å)} \qquad \text{------ (Equation 3)}$$

where

J is the normalized spectral overlap of the donor emission (fD) and acceptor emission (e_A),

Q_D is the quantum efficiency for donor emission in the absence of acceptor,

n is the index of refraction (typically 1.3-1.4 for proteins) and

κ^2 is a geometric factor related to the relative angle of the two transition dipoles (2/3).

10.1.2.4. Overlap Integral, J

The overlap integral, J, represents the degree of overlap between the donor fluorescence spectrum and the acceptor absorbance spectrum and is given by

$$J = \int \varepsilon(\lambda) f_D(\lambda) \lambda 4 d\lambda \; / \; \int f_D(\lambda) d\lambda \text{ (in } M^{-1} cm^{-1} nm^4) \text{ ------ (Equation 4)}$$

where λ is the wavelength of light, $\varepsilon_{(\lambda)}$ is the molar extinction coefficient of the acceptor at that wavelength, and $f_D(\lambda)$ is the fluorescence spectrum of the donor normalized on the wavelength scale.

10.1.2.5. Kappa square κ^2, the orientation factor

The orientation factor is defined as the angle between the donor emission transition moment and the acceptor absorption transition moment. Kappa square varies between 0 and 4. In the Forster's equation, κ^2 (kappa square) assumes a numerical value of 2/3 provided that both the probes undergo unrestricted isotropic motion.

10.1.3. The strategy

Thus FRET becomes an important tool to study the organization of assemblies of biological macromolecules. The basic approach is to label one region of the assembly with a fluorescent energy donor, label a second region with an energy acceptor, measure the efficiency of nonradiative energy transfer between donor and acceptor, and then convert this efficiency to a distance between the two labeled regions (*433*).

With our technology of SNAAR, one could site-specifically introduce fluorescent donor and acceptor within a protein macromolecule. With this scheme, one could map the distance between functionally important loci on the assembly. A fluorescent energy donor is placed at the first site, and an energy acceptor at the second site. The efficiency of transfer is measured, enabling the distance between the two sites to be calculated. Such distances could be measured during the functioning of the molecule and thus one could precisely monitor the conformational changes occurring during its function. Also multiple labeling could give some information on conformational changes that takes place during functioning of the protein.

In this chapter, using FRET as a "spectroscopic ruler", we have determined following two types of distances for light-adapted bR (bR_{LA}) and its functional "M" intermediate: (a) individual distances from end-points of helix C, F, G and retinal, and (b) 'intra-helical' distances i.e. between chosen cytoplasmic and extracellular end-points of each helix (See Figure 9.1). Using the D96N mutant, the "M" intermediate can be trapped at low temperature (10°C), high salt concentration (150 mM KCl) and high pH (9.0) after illumination (*441*).

To find out the individual distances from the end-points of helices C, F and G from retinal, single donor-single acceptor FRET was carried out. In this case, a fluorescent donor such as DACM was introduced via cys-$tRNA^{cys}_{CUA}$ for site-specific incorporation of DACM-cys at the extracellular and cytoplasmic end-points of each of the above-mentioned helices. Retinal was the intrinsic acceptor in these cases. Fluorescence measurements were carried out on such SNAAR analogs. Similar measurements were carried out on SNAAR analogs containing another donor, ISA, and the retinylidene chromophore as the donor-acceptor pair.

To find out the intrahelical postional movement, single donor-double acceptor FRET was performed. In this case, DACM served as the single donor. However, in addition to the intrinsic acceptor retinal, another acceptor, BODIPY, was introduced at the other endpoint of the helix simultaneously, one helix at a time. (The second label could be introduced into bR molecule due to the characteristic absence of cysteine residues in the primary sequence of bR and hence its uniqueness giving rise to site-specificity). Such FRET measurements on proteins having multiple acceptors have been dealt with before (*442*). It has been found that for exact calculations, each acceptor should be considered independently.

10.2 Materials and Methods

10.2.1. Fluorescence measurements and distance determination

Fluorescent SNAAR analogs of bR carrying DACM or ISA labels at specific cysteine residues were expressed *in vivo* and characterized. Fluorescence emission measurements were carried out for all fluorescent labeled (Donor) SNAAR analogs of bR and D96N with or without retinylidene chromophore (acceptor). Fluorescence spectra were measured for all SNAAR analogs of bR in their light-adapted state. SNAAR analogs of bR without retinal were obtained by illuminating (>520 nm) SNAAR analogs in presence of 0.1 M hydroxylamine (final conc.) as described (*443*). Distance between fluorophore and the chromophore at the M intermediate stage was measured by first accumulating M intermediate under illumination (>520 nm) of D96N at pH 9.0, 10 ^{o}C (150 mM KCl) followed by immediate measurement of the fluorescence transfer efficiency and overlap integral, J (*433*, *442*).

10.2.2. Calculations for distance measurements

J and Ro values were calculated for each donor-acceptor pair as mentioned earlier. Efficiency of transfer (E) was calculated from the normalized fluorescence emission values. This value was substituted in another formula to obtain the distance between donor and acceptor labels.

10.3 Results

Fluorescence emission spectra were recorded for all the SNAAR analogs of bR and D96N with and without the acceptor (all-trans retinal). Two types of labels were introduced at the end-points of each helix; *viz*. DACM and ISA. The readings obtained were normalized to the absorbance and calculations were carried out to plot molar extinction spectra. Values of J and Ro were found out using the extinction coefficient values by substituting in the appropriate formulas. Figure 10.1 shows typical absorption spectra for the DACM-labeled bR, DACM-labeled D96N and ISA-labeled D96N (at M-state) respectively.

10.3.1. Distances calculated from FRET-based measurements are comparable to the distances obtained from X-ray crystallography

Distances of end-points of helices from retinal were calculated for bR , D96N and M-state using DACM and ISA labels. The data has been summarized in the following tables (See tables 10.1-10.7). Differences in the values obtained from X-ray crystallographic data could be due to the orientation of the donor fluorescent probe. However, *it is also important to note that conclusions of this study depend on change in distances as measured by FRET and not on absolute distances, which could be further dependent on exactness of factors such as* κ^2. In any case, we find that two FRET-based measurements obtained from two different fluorophores (DACM and ISA) are accurate within 1Å (see below). Tables 4 and 5 show calculations of distances and their comparisons with crystallographic data obtained from cryoelectron microscopy (*437*). The column entitled "DACM orientation" denotes location of DACM label w.r.t. alpha-carbon atom of the substituted amino acid.

Table 10.1: Calculations of distances from helical end-points to retinylidene chromophore using DACM-cys containing SNAAR analogs of bR

Sample	I_D	I_{DA}	E	R	R(cryst)	DACM orientation
L201amcysDACM	100	1.1	0.99	21.2	18.36	2.84
L224amcysDACM	100	0.4	1.00	18.1	16.74	1.36
L190amcysDACM	100	0.8	0.99	20.1	18.33	1.77
S169amcysDACM	100	2.5	0.98	24.4	14.56	-0.16
V101amcysDACM	100	1.9	0.98	23.2	19.35	3.85
Y79amcysDACM	100	0.7	0.99	19.6	17.11	2.49

$J = 2.13 \times 10^{15}$ (M^{-1} cm^{-1} nm^4) ; Ro = 44.96Å

Table 10.2: Calculations of distances from helical end-points to retinylidene chromophore using DACM-cys containing SNAAR analogs of D96N

Sample	I_D	I_{DA}	E	R	R(cryst)	DACM orientation
L201amcysDACM	100	1.4	0.99	22.1	18.36	3.74
L224amcysDACM	100	0.6	0.99	19.3	16.74	2.56
L190amcysDACM	100	1.4	0.99	22.1	18.33	3.77
S169amcysDACM	100	3.0	0.97	25.1	24.56	0.54
V101amcysDACM	100	2.0	0.98	23.4	19.35	4.05
Y79amcysDACM	100	0.7	0.99	19.6	17.11	2.49

$J = 2.13 \times 10^{15}$ (M^{-1} cm^{-1} nm^{4}) ; Ro = 44.86Å

In case of interhelical measurements, overlap integrals J, were determined from absorption and emission spectra (see Figure 10.1(a) shown typically for one SNAAR analogue, e.g. S169am-DACM). Fluorescence transfer efficiencies were calculated in the presence or absence of the acceptor, retinal (for absorption spectra, see for example curves 2, 4 Figure 10.1(a)). It is important to note that distances calculated for bR^{LA} and $D96N^{LA}$ do not show significant structural differences with respect to C, F and G helices and retinal position (See Tables 10.1 and 10.2). This shows that the D96N mutant is conformationally similar to the WT.

10.3.2. During the M-intermediate formation, F-helix moves 5Å away from retinal, while C and G helices do not show any appreciable conformational change

Table 10.3, 10.4 and 10.5 show distance determination between retinylidene chromophore and helical end-points using ISA (donor): retinal (acceptor) bR, D96N and D96N-M-state. These measurements confirm DACM based distances and show that both, wild-type and D96N, have unaltered structures.

Decay of the M-intermediate of bR is known to be reversibly blocked in case of bR mutant D96N at high pH and at high salt conditions. We have carried SNAAR coupled FRET measurements on accumulated M intermediate and have determined the nature of the conformational change during the bR→M transition. Since D96N and wild-type bR have similar initial structures, we argue that the conformational changes detected in the "D96N-M"are due to the M-state formation during the photocycle of bR. Thus, M intermediate is accumulated as described in materials and methods. ISA is a fluorophore that emits in the wavelength region where M intermediate absorbs, allowing us a suitable FRET donor:acceptor pair. Overlap integrals (curve 1, 2 and 3, Figure 10.1(c)), transfer efficiencies and distances were determined for bR^{LA}, $D96N^{LA}$ and the trapped "M" states of bR. The data (Table 10.1) shows that during the "M" intermediate, the cytoplasmic domain of F-helix moves 5 Å away from the retinylidene chromophore, while other reference points of C, F and G helices do not show appreciable distance changes.

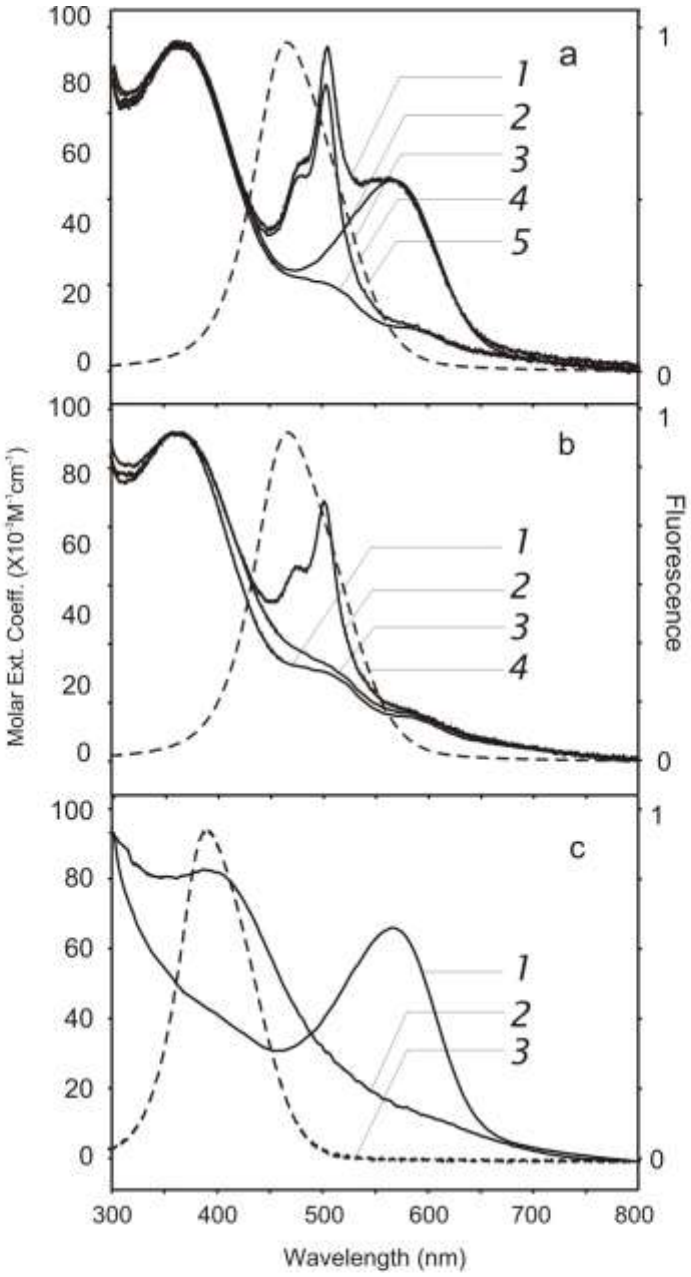


Figure 10.1: Characterization of SNAAR analogs of bR. (a) Normalized absorption spectra for (1)S169am-DACM/L190C-BODIPY, (2) S169am-DACM, (3) apo S169am-DACM/ L190C-BODIPY, (4) apo(S169am-DACM), and (5) emission spectra of S169am-DACM; (b)Normalized absorption spectra of (1) apo D96N:S169am-DACM, (2) D96N:S169am-DACM:L190am-BODIPY trapped as "M", (3) D96N:S169am-DACM trapped as "M", (4) emission spectrum of D96N:S169am-DACM; (c) Normalized absorption spectra of (1) D96N:S169am-ISA, (2) D96N:S169am-ISA trapped as "M" (see methods) and (3) emission spectra of S169am-ISA.

Table 10.3: Calculations of distances from helical end-points to retinylidene chromophore using ISA-cys containing SNAAR analogs of bR

Sample	I_D	I_{DA}	E	R	R(cryst)	ISA orientation
L201amcysISA	100	4.5	0.95	22.0	18.36	3.64
L224amcysISA	100	1.7	0.98	18.6	16.74	1.86
L190amcysISA	100	3.5	0.97	21.0	18.33	2.67
S169amcysISA	100	7.7	0.92	24.2	24.56	-0.36
V101amcysISA	100	7.4	0.93	24.0	19.35	4.65
Y79amcysISA	100	2.6	0.97	20.0	17.11	2.89

$J = 9.25 \times 10^{14}$ (M^{-1} cm^{-1} nm^{4}) ; Ro = 36.59Å

Table 10.4: Calculations of distances from helical end-points to retinylidene chromophore using ISA-cys containing SNAAR analogs of D96N

Sample	I_D	I_{DA}	E	R	R(cryst)	ISA orientation
L201amcysISA	100	7.4	0.93	22.6	18.36	4.24
L224amcysISA	100	4.3	0.96	19.8	16.74	3.06
L190amcysISA	100	6.7	0.93	22.4	18.33	4.07
S169amcysISA	100	28.7	0.71	24.9	24.56	0.34
V101amcysISA	100	9.6	0.90	24.0	19.35	4.65
Y79amcysISA	100	3.9	0.96	19.9	17.11	2.79

$J = 9.25 \times 10^{14}$ (M^{-1} cm^{-1} nm^{4}) ; Ro = 36.59Å

Table 10.5: Calculations of distances from helical end-points to retinylidene chromophore using ISA-cys containing SNAAR analogs of D96N-M state

Sample	I_D	I_{DA}	E	R	R(cryst)	ISA orientation
L201amcysISA	100	7.4	0.93	22.3	18.36	3.94
L224amcysISA	100	4.3	0.96	20.3	16.74	3.56
L190amcysISA	100	6.7	0.93	21.9	18.33	3.57
S169amcysISA	100	28.7	0.71	29.2	24.56	4.64
V101amcysISA	100	9.6	0.90	23.4	19.35	4.05
Y79amcysISA	100	3.9	0.96	19.9	17.11	2.79

$J = 5.95 \times 10^{15}$ (M^{-1} cm^{-1} nm^{4}) ; Ro = 33.99Å

As shown in Table 10.6 it is important to note that distances measured from two different donors show similar distance values with very high accuracy (within 1 A). These data indicate that protein structure may hold these fluorophores at fixed points (probably at similar hydrophobic loci) despite their structural differences.

Table 10.6: Comparison of distance values (R) obtained from two different donors; DACM and ISA

Mutant	From FRET		R (cryst)	Delta DACM	Delta ISA	DACM-ISA
	R (DACM)	R (ISA)				
WT L201am	21.2	22.0	18.36	2.84	3.64	-0.8
WT L224am	18.1	18.6	16.74	1.36	1.86	-0.5
WT L190am	20.1	21.0	18.33	1.77	2.67	-0.9
WT S169am	24.4	24.2	24.56	-0.16	-0.36	0.2
WT V101am	23.2	24.0	19.35	3.85	4.65	-0.8
WT Y79am	19.6	20.0	17.11	2.49	2.89	-0.4
D96N L201am	22.1	22.6	18.36	3.74	4.24	-0.5
D96N L224am	19.3	19.8	16.74	2.56	3.06	-0.5
D96N L190am	22.1	22.4	18.33	3.77	4.07	-0.3
D96N S169am	25.1	24.9	24.56	0.54	0.34	0.2
D96N V101am	23.4	24.0	19.35	4.05	4.65	-0.6
D96N Y79am	19.6	19.9	17.11	2.49	2.79	-0.3

Table 10.7 summarizes the distance measurements between helical end-points and retinal in the wild type-bR, D96N and D96N-M state. The data clearly shows a significant difference in the distance of S169 (located on the cytoplasmic side of the F-helix) from retinal. From this data, we conclude that the C and G helices may not undergo significant conformational change whereas the cytoplasmic side F-helix may be undergoing a conformational change which we have characterized to be ~5Å.

Table 10.7: Distance between retinylidene Schiff base and site-specifically introduced fluorophore (single donor-single acceptor) as measured by fluorescence resonance energy transfer measurements. (See Methods and Materials for additional details)[#]

Position of the fluorophore	Distance of the fluorophore from retinyledene chromphore (Å)			Difference in the distance of fluorophore and retinyledene chromophore (Å)	
	bR^{LA}	$D96N^{LA}$	In M-state	M state- bR^{LA}	M state- $D96N^{LA}$
L201am	22 (±1.2)	22.6(±1.1)	22.3(±1.3)	1.0 (±0.5)	-0.3 (±1.0)
L224am	18.6(±1.0)	19.8(±1.2)	20.3(±1.2)	1.7 (±1.1)	0.5 (±1.2)
L190am	21(±1.1)	22.4 (±1.2)	21.9 (±1.2)	0.9 (±1.21)	-0.5 (±1.2)
S169am	**24.2 (±1.1)**	**24.9 (±1.0)**	**29.2 (±1.3)**	**5 (±1.0)**	**4.3 (±1.1)**
V101am	24 (±1.2)	24 (±1.3)	23.4 (±1.3)	-0.6 (±1.2)	-0.6 (±1.3)
Y79am	20 (±1.2)	19.9 (±1.2)	19.9 (±1.2)	-0.1 (±1.2)	0 (±1.2)

10.3.3. F-helix of bacteriorhodopsin tilts by ~70° during its proton pumping function

In order to further refine the exact nature of distance changes, intra-helical distances were measured using FRET between single donor and two acceptors as described earlier. The distances between fluorophore (DACM) placed at one end of helix (on the cytoplasmic side), and the acceptor (BODIPY) at the other (on the extracellular side of the same helix), were calculated in presence of the retinylidene chromophore as the intrinsic acceptor. Such single donor-two acceptor calculations require independent measurement of J values of each donor: acceptor pair (*442*), i.e. (a) DACM (emission)/retinal (absorption) (curves 5 and 2, Figure 10.1(a)) in absence of BODIPY, and (b) DACM (emission)/BODIPY (absorption) (curves 5 and 3, Figure 10.1(a)), in absence of retinal. Fluorescence transfer efficiency in this case was determined by comparing the emission of the sample with both acceptors (for absorption spectrum, see curve 1, Figure 10.1(a)) and without any acceptor (for absorption spectrum, see curve 4, Figure 10.1(a)). Intra-helical distances were measured in case of all three states: light-adapted bR, light-adapted D96N and the trapped "M" and resulting data is summarized in Table 10.8.

[#] In each case, experiments were carried out in triplicate starting from preparation of the SNAAR analog.

Table 10.8: Intra-helical distances as determined by single donor-double acceptor FRET

Distance (Cytoplasmic end ↔ extracellular end)	**bR^{LA}** (Å)	**$D96N^{LA}$** (Å)	**D96N- "M-state"** (Å)
C Helix (V101am ↔ Y79Cys)	38.2 (±1.1)	37.9(±1.3)	38.2 (±1.3)
G Helix (L224am ↔ L201Cys)	34.1 (±1.2)	33.9 (±0.8)	34.4 (±1.0)
F Helix (S169am ↔ L190Cys)	**36.1 (±1.3)**	**35.6 (±1.0)**	**30.1 (±1.2)**

The data further confirms that the C and G helices do not any conformational change, however, the distance between end-points of helix F seems to change from 36 Å to 30.1 Å. From figure 10.2, it can be seen that the distances of the triangle $PF'F^M$ are known. Distance between PF^M and $F'F^M$ is based on the actual data obtained from the above FRET studies. The distance between F'P is assumed from the crystal structure of bR (assuming that P is the point of Tyr185-Pro186 residue). Based on simple triangular geometric calculation (Figure 10.2), it can be seen that the F helix should tilt by approximately 70° around some point (assumed to be Tyr-185/Pro-186 hinge, see below) within the helix, as the intra-helical end-points seem to come closer as a result of the "M" intermediate formation. Intra-helical distances for C and G helices are found to remain unaltered during the bR^{LA} to M transition. It should be noted that the retinylidene Schiff base is found to tilt by not more than 5° during the photocycle (*444*).

10.4 Discussion

Isomerization of retinal by light is the first step in the transduction of light energy of proteins in the rhodopsin family. This light-induced isomerization of retinal triggers changes in protein conformation so that a proton can be released into the extracellular medium followed by uptake of another proton from the cytoplasmic surface, net result being pumping out into the extracellular medium of a single proton. This release of proton into the extracellular medium has been shown to coincide with the M-intermediate formation (*445*). A combination of previous neutron diffraction studies (*350*), X-ray (*446*, *447*) and electron diffraction studies (*351*, *448*)have provided strong evidence that the transmembrane regions undergo significant light-induced changes in conformation during the course of the photocycle.

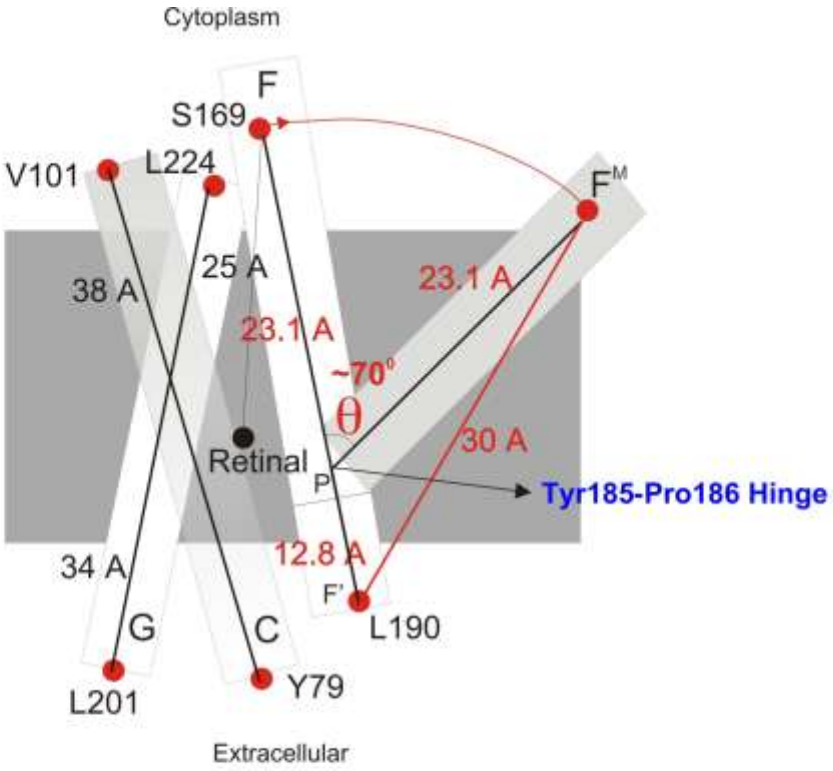


Figure 10.2: Schematic representation of outward tilt of helix-F in bR photocycle. Red filled circles indicate positions of the non-native amino acids introduced (one at a time). Black lines indicate distances measured using SNAAR-FRET in ground state of bR whereas red lines indicate distances measured in "M" form of bR.

Studies such as atomic force microscopy (*449*) and site-directed spin labeling by introducing cysteines into interhelical loops have revealed that movement of the protein moiety above the lipid bilayers which results in bending of the most prominent cytoplasmic E-F loop (*355*). This causes opening of the channel and the event occurs between proton release and uptake, and may be the conformational switch that changes the accessibilty of the retinal Schiff base to the cytoplasmic surface after proton release to the extracellular surface (*355*). Low temperature FTIR studies by Kluge *et al* have revealed that retinal isomerization induces conformational changes in the core structure of bR during early photocycle, which may involve an increase in the strength of intramolecular α-helical hydrogen bonds (*450*). Another study by Subramaniam et al was done in which electron crystallographic analysis of various conformations of bR was done by trapping single and triple mutants by freezing two-dimensional crystals in liquid ethane at varying times after illumination with a light flash (*437*). Electron diffraction patterns recorded from these crystals were used to construct projection difference Fourier maps to define light-driven changes in protein conformation. These maps showed that main structural change in the photocycle was largely localized to the central four helices lining the proton channel- i.e. F, G, B and C. They define the main structural changes as:

i) Ordering of the helix G at the cytoplasmic end and
ii) Outward tilt of F-helix with Pro-186, which is likely to serve as a "hinge" and thus causes a widening of the proton channel.

To summarise, when M→ N conformation change occurs, proton channel opens towards the cytoplasmic surface by the tilt of cytoplasmic end of the helix F and this is required for proton transfer from Asp-96 → retinal schiff. Also previously, SDIL-FTIR data has shown that very early in the photocycle of bR, Tyr-185/Pro-186 peptide bond gets perturbed due to retinal photoisomerization (*199*, *219*, *221*). SDIL-ATR-FTIR has shown that later in the photocycle, i.e. by M→N transition, conformational change within bR is complete (*219*). This work further confirms that the most prominent conformational change during the photocycle is in the F-helix whereas C and G helices do not show any appreciable difference. *It also defines the conformational change as a movement of the F-helix by 5 Å away from retinal by tilting at an angle of 45°*. Figure 10.2

shows a schematic representation of the outward tilt of the F-helix during the photocycle of light-adapted bR as determined by *SNAAR coupled FRET*. Further work based on flash-kinetic fluorescence measurements of such SNAAR analogs could allow *direct* observation of conformational changes in native bR instead of D96N.

In conclusion, this is the first report that demonstrates SNAAR on a large scale in an *in vivo* manner, and where more than one non-native functionalities are simultaneously introduced. The approach here is extendable to protoplasts of E. coli and mammalian cell-cultures. Simplicity of our approach of introduction of non-native moieties as post-aminoacylation of suppressor tRNA transcripts avoids tedious chemistry of chemical aminoacylation, and therein holds wider applicability. It should also be noted that this approach extends to all proteins including those bearing cysteines and allows *in vivo* protein folding. In these merits, we believe that this approach has ease and a simplicity that goes beyond traditional site-directed mutagenesis.

Applications of SNAAR analogs of bR in biomolecular electronics could be significant. It is expected that introduction of fluorophores/ chromophores could allow wavelength modulations required for the ON-OFF or 0-1 states required for binary electronics. Further work would bring these aspects of SNAAR into biomolecular electronics arena.

Bibliography

1. Sonar, S., Krebs, M. P., Khorana, H. G., and Rothschild, K. J. (1993) Static and time-resolved absorption spectroscopy of the bacteriorhodopsin mutant Tyr-185-->Phe: evidence for an equilibrium between bR570 and an O-like species, *32*, 2263-2271.
2. Krummel, B., and Chamberlin, M. J. (1989) RNA chain initiation by Escherichia coli RNA polymerase. Structural transitions of the enzyme in early ternary complexes, *28*, 7829-7842.
3. Siebenlist, U., Simpson, R. B., and Gilbert, W. (1980) E. coli RNA polymerase interacts homologously with two different promoters, *Cell 20*, 269-281.
4. Jakubowski, H., and Goldman, E. (1992) Editing of errors in selection of amino acids for protein synthesis, *Microbiol Rev 56*, 412-429.
5. CHAPEVILLE, F., LIPMANN, F., VON EHRENSTEIN, G., WEISBLUM, B., RAY WJ, J. r., and BENZER, S. (1962) On the role of soluble ribonucleic acid in coding for amino acids, *Proc Natl Acad Sci U S A 48*, 1086-1092.
6. Ebel, J. P., Giege, R., Bonnet, J., Kern, D., Befort, N., Bollack, C., Fasiolo, F., Gangloff, J., and Dirheimer, G. (1973) Factors determining the specificity of the tRNA aminoacylation reaction. Non-absolute specificity of tRNA-aminoacyl-tRNA synthetase recognition and particular importance of the maximal velocity, *Biochimie 55*, 547-557.
7. E.A., F. (1998) Catalysis of tRNA aminoacylation by class I and class II aminoacyl tRNA synthetases, in *Comprehensive Biological Catalysis* (M., i., Ed.), pp 573-607, New York: Academic Press.
8. Loftfield, R. B., and Vanderjagt, D. (1972) The frequency of errors in protein biosynthesis, *Biochem J 128*, 1353-1356.
9. Freist, W., Sternbach, H., and Cramer, F. (1988) Isoleucyl-tRNA synthetase from baker's yeast and from Escherichia coli MRE 600. Discrimination of 20 amino acids in aminoacylation of tRNA(Ile)-C-C-A, *Eur J Biochem 173*, 27-34.
10. Ludmerer, S. W., and Schimmel, P. (1987) Gene for yeast glutamine tRNA synthetase encodes a large amino-terminal extension and provides a strong confirmation of the signature sequence for a group of the aminoacyl-tRNA synthetases, *J Biol Chem 262*, 10801-10806.
11. Webster, T., Tsai, H., Kula, M., Mackie, G. A., and Schimmel, P. (1984) Specific sequence homology and three-dimensional structure of an aminoacyl transfer RNA synthetase, *226*, 1315-1317.
12. RICH, A., and CRICK, F. H. (1955) The structure of collagen, *176*, 915-916.
13. Woese, C. R. (1965) On the evolution of the genetic code, *Proc Natl Acad Sci U S A 54*, 1546-1552.
14. CRICK, F. H. (1958) On protein synthesis, *Symp Soc Exp Biol 12*, 138-163.
15. Trower, M. K. (1996) A protocol for site-directed mutagenesis employing a uracil-containing phagemid template, *Methods Mol Biol 58*, 469-476.
16. Zoller, M. J., and Smith, M. (1982) Oligonucleotide-directed mutagenesis using M13-derived vectors: an efficient and general procedure for the production of point mutations in any fragment of DNA, *Nucleic Acids Res 10*, 6487-6500.
17. Hatfield, D., and Diamond, A. (1993) UGA: a split personality in the universal genetic code, *Trends Genet 9*, 69-70.
18. Hanyu, N., Kuchino, Y., Nishimura, S., and Beier, H. (1986) Dramatic events in ciliate evolution: alteration of UAA and UAG termination codons to glutamine codons due to anticodon mutations in two Tetrahymena tRNAs, *EMBO J JID- 8208664 5*, 1307-1311.
19. Schull, C., and Beier, H. (1994) Three Tetrahymena tRNA(Gln) isoacceptors as tools for studying unorthodox codon recognition and codon context effects during protein synthesis in vitro, *Nucleic Acids Res 22*, 1974-1980.
20. Osawa, S., Jukes, T. H., Watanabe, K., and Muto, A. (1992) Recent evidence for evolution of the genetic code, *Microbiol Rev 56*, 229-264.
21. Stieglitz, B., and Calvo, J. M. (1971) Effect of 4-azaleucine upon leucine metabolism in Salmonella typhimurium, *J Bacteriol 108*, 95-104.
22. Pines, M., Rosenthal, G. A., and Applebaum, S. W. (1981) In vitro incorporation of L-canavanine into vitellogenin of the fat body of the migratory locust Locusta migratoria migratorioides, *Proc Natl Acad Sci U S A 78*, 5480-5483.

23. Heinemeyer, W., Kleinschmidt, J. A., Saidowsky, J., Escher, C., and Wolf, D. H. (1991) Proteinase yscE, the yeast proteasome/multicatalytic-multifunctional proteinase: mutants unravel its function in stress induced proteolysis and uncover its necessity for cell survival, *EMBO J 10*, 555-562.
24. Muller, S., Senn, H., Gsell, B., Vetter, W., Baron, C., and Bock, A. (1994) The formation of diselenide bridges in proteins by incorporation of selenocysteine residues: biosynthesis and characterization of (Se)2-thioredoxin, *33*, 3404-3412.
25. Kohno, T., Kohda, D., Haruki, M., Yokoyama, S., and Miyazawa, T. (1990) Nonprotein amino acid furanomycin, unlike isoleucine in chemical structure, is charged to isoleucine tRNA by isoleucyl-tRNA synthetase and incorporated into protein, *J Biol Chem 265*, 6931-6935.
26. Lemeignan, B., Sonigo, P., and Marliere, P. (1993) Phenotypic suppression by incorporation of an alien amino acid, *J Mol Biol JT - Journal of molecular biology 231*, 161-166.
27. Hendrickson, W. A., Horton, J. R., and LeMaster, D. M. (1990) Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure, *EMBO J 9*, 1665-1672.
28. Koide, H., Yokoyama, S., Kawai, G., Ha, J. M., Oka, T., Kawai, S., Miyake, T., Fuwa, T., and Miyazawa, T. (1988) Biosynthesis of a protein containing a nonprotein amino acid by Escherichia coli: L-2-aminohexanoic acid at position 21 in human epidermal growth factor, *Proc Natl Acad Sci U S A 85*, 6237-6241.
29. Hinds, M. G., King, R. W., and Feeney, J. (1992) 19F n.m.r. studies of conformational changes accompanying cyclic AMP binding to 3-fluorophenylalanine-containing cyclic AMP receptor protein from Escherichia coli, *Biochem J 287 (Pt 2)*, 627-632.
30. Kanemori, M., Mori, H., and Yura, T. (1994) Induction of heat shock proteins by abnormal proteins results from stabilization and not increased synthesis of sigma 32 in Escherichia coli, *J Bacteriol 176*, 5648-5653.
31. Ross, J. B., Senear, D. F., Waxman, E., Kombo, B. B., Rusinova, E., Huang, Y. T., Laws, W. R., and Hasselbacher, C. A. (1992) Spectral enhancement of proteins: biological incorporation and fluorescence characterization of 5-hydroxytryptophan in bacteriophage lambda cI repressor, *89*, 12023-12027.
32. Lui, S. M., Soriano, A., and Cowan, J. A. (1994) Electronic properties of the dissimilatory sulphite reductase from Desulfovibrio vulgaris (Hildenborough): comparative studies of optical spectra and relative reduction potentials for the [Fe4S4]-sirohaem prosthetic centres, *Biochem J 304 (Pt 2)*, 441-447.
33. Boehlein, S. K., Walworth, E. S., and Schuster, S. M. (1997) Identification of cysteine-523 in the aspartate binding site of Escherichia coli asparagine synthetase B, *36*, 10168-10177.
34. Lorentzon, P., Eklundh, B., Brandstrom, A., and Wallmark, B. (1985) The mechanism for inhibition of gastric (H+ + K+)-ATPase by omeprazole, *Biochim Biophys Acta 817*, 25-32.
35. Vanin, A. F., Serezhenkov, V. A., Mikoyan, V. D., and Genkin, M. V. (1998) The 2.03 signal as an indicator of dinitrosyl-iron complexes with thiol-containing ligands, *Nitric Oxide 2*, 224-234.
36. Stehlin, C., Heacock DH, n., Liu, H., and Musier-Forsyth, K. (1997) Chemical modification and site-directed mutagenesis of the single cysteine in motif 3 of class II Escherichia coli prolyl-tRNA synthetase, *36*, 2932-2938.
37. Hayden, B. M., and Engel, P. C. (2001) Construction, separation and properties of hybrid hexamers of glutamate dehydrogenase in which five of the six subunits are contributed by the catalytically inert D165S, *Eur J Biochem 268*, 1173-1180.
38. Communi, D., and Erneux, C. (1996) Identification of an active site cysteine residue in human type I Ins(1,4,5)P3 5-phosphatase by chemical modification and site-directed mutagenesis, *Biochem J 320 (Pt 1)*, 181-186.
39. Sasisekharan, R., Leckband, D., Godavarti, R., Venkataraman, G., Cooney, C. L., and Langer, R. (1995) Heparinase I from Flavobacterium heparinum: the role of the cysteine residue in catalysis as probed by chemical modification and site-directed mutagenesis, *34*, 14441-14448.
40. Chang, W. I., and Matthews, K. S. (1995) Role of Asp274 in lac repressor: diminished sugar binding and altered conformational effects in mutants, *34*, 9227-9234.
41. Takahashi, K. (1977) The reactions of phenylglyoxal and related reagents with amino acids, *J Biochem 81*, 395-402.
42. Takahashi, K. (1977) Further studies on the reactions of phenylglyoxal and related reagents with proteins, *J Biochem 81*, 403-414.

43. Carlson, C. A., and Preiss, J. (1982) Involvement of arginine residues in the allosteric activation of Escherichia coli ADP-glucose synthetase, *21*, 1929-1934.
44. Powers, S. G., and Riordan, J. F. (1975) Functional arginyl residues as ATP binding sites of glutamine synthetase and carbamyl phosphate synthetase, *Proc Natl Acad Sci U S A 72*, 2616-2620.
45. Beliveau, R., Bernier, M., Giroux, S., and Bates, D. (1988) Inhibition by phenylglyoxal of the sodium-coupled fluxes of glucose and phosphate in renal brush-border membranes, *Biochem Cell Biol 66*, 1005-1012.
46. Bjerrum, P. J. (1989) Chemical modification of the anion-transport system with phenylglyoxal, *Methods Enzymol 173*, 466-494.
47. Missiaen, L., Raeymaekers, L., Droogmans, G., Wuytack, F., and Casteels, R. (1989) Role of arginine residues in the stimulation of the smooth-muscle plasma-membrane Ca2+ pump by negatively charged phospholipids, *Biochem J 264*, 609-612.
48. Julien, T., Betakis, E., and Zaki, L. (1990) Chemical properties of the anion transport inhibitory binding site of arginine-specific reagents in human red blood cell membranes, *Biochim Biophys Acta 1026*, 43-50.
49. Holman, C. M., and Benisek, W. F. (1994) Extent of proton transfer in the transition states of the reaction catalyzed by the delta 5-3-ketosteroid isomerase of Comamonas (Pseudomonas) testosteroni: site-specific replacement of the active site base, aspartate 38, by the weaker base alanine-3, *33*, 2672-2681.
50. Kim, D. W., Yoshimura, T., Esaki, N., Satoh, E., and Soda, K. (1994) Studies of the active-site lysyl residue of thermostable aspartate aminotransferase: combination of site-directed mutagenesis and chemical modification, *J Biochem 115*, 93-97.
51. Steinhoff, H. J., Mollaaghababa, R., Altenbach, C., Hideg, K., Krebs, M., Khorana, H. G., and Hubbell, W. L. (1994) Time-resolved detection of structural changes during the photocycle of spin-labeled bacteriorhodopsin, *Science JT - Science (New York, N.Y.) 266*, 105-107.
52. Siksnys, V., and Pleckaityte, M. (1992) Role of the reactive cysteine residue in restriction endonuclease Cfr9I, *Biochim Biophys Acta 1160*, 199-205.
53. Tian, H., Yu, L., Mather, M. W., and Yu, C. A. (1997) The involvement of serine 175 and alanine 185 of cytochrome b of Rhodobacter sphaeroides cytochrome bc1 complex in interaction with iron-sulfur protein, *J Biol Chem 272*, 23722-23728.
54. Perez-Garcia, M. T., Chiamvimonvat, N., Ranjan, R., Balser, J. R., Tomaselli, G. F., and Marban, E. (1997) Mechanisms of sodium/calcium selectivity in sodium channels probed by cysteine mutagenesis and sulfhydryl modification, *Biophys J 72*, 989-996.
55. Wang, Z., Wang, X., and Rana, T. M. (1996) Protein orientation in the Tat-TAR complex determined by psoralen photocross-linking, *J Biol Chem 271*, 16995-16998.
56. Lusti-Narasimhan, M., Chollet, A., Power, C. A., Allet, B., Proudfoot, A. E., and Wells, T. N. (1996) A molecular switch of chemokine receptor selectivity. Chemical modification of the interleukin-8 Leu25 --> Cys mutant, *J Biol Chem 271*, 3148-3153.
57. Jones, P. C., Sivaprasadarao, A., Wray, D., and Findlay, J. B. (1996) A method for determining transmembrane protein structure, *Mol Membr Biol 13*, 53-60.
58. Gloss, L. M., and Kirsch, J. F. (1995) Examining the structural and chemical flexibility of the active site base, Lys-258, of Escherichia coli aspartate aminotransferase by replacement with unnatural amino acids, *34*, 12323-12332.
59. Wynn, R., Anderson, C. L., Richards, F. M., and Fox, R. O. (1995) Interactions in nonnative and truncated forms of staphylococcal nuclease as indicated by mutational free energy changes, *Protein Sci 4*, 1815-1823.
60. Sahin-Toth, M., and Kaback, H. R. (1993) Cysteine scanning mutagenesis of putative transmembrane helices IX and X in the lactose permease of Escherichia coli, *Protein Sci 2*, 1024-1033.
61. Hollenbaugh, D., Aruffo, A., and Senter, P. D. (1995) Effects of chemical modification on the binding activities of P-selectin mutants, *34*, 5678-5684.
62. Ruvinov, S. B., Yang, X. J., Parris, K. D., Banik, U., Ahmed, S. A., Miles, E. W., and Sackett, D. L. (1995) Ligand-mediated changes in the tryptophan synthase indole tunnel probed by nile red fluorescence with wild type, mutant, and chemically modified enzymes, *J Biol Chem 270*, 6357-6369.
63. Gilardi, G., Zhou, L. Q., Hibbert, L., and Cass, A. E. (1994) Engineering the maltose binding protein for reagentless fluorescence sensing, *Anal Chem 66*, 3840-3847.
64. Salvucci, M. E., and Klein, R. R. (1994) Site-directed mutagenesis of a reactive lysyl residue (Lys-247) of Rubisco activase, *Arch Biochem Biophys 314*, 178-185.

65. Qiu, X. Q., Jakes, K. S., Finkelstein, A., and Slatin, S. L. (1994) Site-specific biotinylation of colicin Ia. A probe for protein conformation in the membrane, *J Biol Chem 269*, 7483-7488.
66. Dhalla, A. M., Li, B., Alibhai, M. F., Yost, K. J., Hemmingsen, J. M., Atkins, W. M., Schineller, J., and Villafranca, J. J. (1994) Regeneration of catalytic activity of glutamine synthetase mutants by chemical activation: exploration of the role of arginines 339 and 359 in activity, *Protein Sci 3*, 476-481.
67. Xu, M., Covey, D. F., and Akabas, M. H. (1995) Interaction of picrotoxin with GABAA receptor channel-lining residues probed in cysteine mutants, *Biophys J 69*, 1858-1867.
68. Wynn, R., and Richards, F. M. (1993) Unnatural amino acid packing mutants of Escherichia coli thioredoxin produced by combined mutagenesis/chemical modification techniques, *Protein Sci 2*, 395-403.
69. Falke, J. J., Dernburg, A. F., Sternberg, D. A., Zalkin, N., Milligan, D. L., and Koshland DE, J. r. (1988) Structure of a bacterial sensory receptor. A site-directed sulfhydryl study, *J Biol Chem 263*, 14850-14858.
70. Smith, H. B., Larimer, F. W., and Hartman, F. C. (1988) Subtle alteration of the active site of ribulose bisphosphate carboxylase/oxygenase by concerted site-directed mutagenesis and chemical modification, *Biochem Biophys Res Commun 152*, 579-584.
71. Bech, L. M., and Breddam, K. (1988) Chemical modifications of a cysteinyl residue introduced in the binding site of carboxypeptidase Y by site-directed mutagenesis, *Carlsberg Res Commun 53*, 381-393.
72. Bain, J. D., Glabe, C. G., Dix, T. A., and Chhamberlin, R. (1989) Synthetic Site-Specific Incorporation of a Natural Amino Acid into a Polypeptide, *Journal of American Chemical Society 111*, 8013-8014.
73. Noren, C. J., Anthony-Cahill, S. J., Griffith, M. C., and Schultz, P. G. (1989) A general method for site-specific incorporation of unnatural amino acids into proteins, *Science JT - Science (New York, N.Y.) 244*, 182-188.
74. Mendel, D., Cornish, V. W., and Schultz, P. G. (1995) Site-directed mutagenesis with an expanded genetic code, *Annu Rev Biophys Biomol Struct 24*, 435-462.
75. Steward, L. E., and Chamberlin, A. R. (1998) Protein engineering with nonstandard amino acids, *Methods Mol Biol 77*, 325-354.
76. Dougherty, D. A. (2000) Unnatural amino acids as probes of protein structure and function, *Curr Opin Chem Biol 4*, 645-652.
77. Tam, R., and Saier MH, J. r. (1993) A bacterial periplasmic receptor homologue with catalytic activity: cyclohexadienyl dehydratase of Pseudomonas aeruginosa is homologous to receptors specific for polar amino acids, *Res Microbiol 144*, 165-169.
78. Tam, R., and Saier MH, J. r. (1993) Structural, functional, and evolutionary relationships among extracellular solute-binding receptors of bacteria, *Microbiol Rev 57*, 320-346.
79. Hennecke, H., and Bock, A. (1975) Altered alpha subunits in phenylalanyl-tRNA synthetases from p-fluorophenylalanine-resistant strains of Escherichis coli, *Eur J Biochem 55*, 431-437.
80. Gao, W., Goldman, E., and Jakubowski, H. (1994) Role of carboxy-terminal region in proofreading function of methionyl-tRNA synthetase in Escherichia coli, *33*, 11528-11535.
81. Schmidt, E., and Schimmel, P. (1994) Mutational isolation of a sieve for editing in a transfer RNA synthetase, *Science JT - Science (New York, N.Y.) 264*, 265-267.
82. Ibba, M., and Hennecke, H. (1994) Towards engineering proteins by site-directed incorporation in vivo of non-natural amino acids, *Biotechnology (N Y) 12*, 678-682.
83. Stanzel, M., Schon, A., and Sprinzl, M. (1994) Discrimination against misacylated tRNA by chloroplast elongation factor Tu, *Eur J Biochem 219*, 435-439.
84. Thompson, R. C. (1988) EFTu provides an internal kinetic standard for translational accuracy, *Trends Biochem Sci 13*, 91-93.
85. Barciszewski, J., Sprinzl, M., and Clark, B. F. (1994) Aminoacyl-tRNAs. Diversity before and unity after interaction with EF-Tu:GTP, *FEBS Lett 351*, 137-139.
86. Seong, B. L., and RajBhandary, U. L. (1987) Mutants of Escherichia coli formylmethionine tRNA: a single base change enables initiator tRNA to act as an elongator in vitro, *84*, 8859-8863.
87. Baron, C., and Bock, A. (1991) The length of the aminoacyl-acceptor stem of the selenocysteine-specific tRNA(Sec) of Escherichia coli is the determinant for binding to elongation factors SELB or Tu, *J Biol Chem 266*, 20375-20379.
88. Hohsaka, T., Sato, K., Sisido, M., Takai, K., and Yokoyama, S. (1993) Adaptability of nonnatural aromatic amino acids to the active center of the E. coli ribosomal A site, *FEBS Lett 335*, 47-50.

89. Roesser, J. R., Xu, C., Payne, R. C., Surratt, C. K., and Hecht, S. M. (1989) Preparation of misacylated aminoacyl-tRNA(Phe)'s useful as probes of the ribosomal acceptor site, *28*, 5185-5195.
90. Noren, C. J., Anthony-Cahill, S. J., Suich, D. J., Noren, K. A., Griffith, M. C., and Schultz, P. G. (1990) In vitro suppression of an amber mutation by a chemically aminoacylated transfer RNA prepared by runoff transcription, *Nucleic Acids Res JT - Nucleic acids research 18*, 83-88.
91. Ellman, J., Mendel, D., Anthony-Cahill, S., Noren, C. J., and Schultz, P. G. (1991) Biosynthetic method for introducing unnatural amino acids site-specifically into proteins, *Methods Enzymol 202*, 301-336.
92. Kohrer, C., Yoo, J. H., Bennett, M., Schaack, J., and RajBhandary, U. L. (2003) A possible approach to site-specific insertion of two different unnatural amino acids into proteins in mammalian cells via nonsense suppression, *Chem Biol 10*, 1095-1102.
93. Ma, C., Kudlicki, W., Odom, O. W., Kramer, G., and Hardesty, B. (1993) In vitro protein engineering using synthetic tRNA(Ala) with different anticodons, *32*, 7939-7945.
94. Piccirilli, J. A., Krauch, T., Moroney, S. E., and Benner, S. A. (1990) Enzymatic incorporation of a new base pair into DNA and RNA extends the genetic alphabet, *Nature JT - Nature 343*, 33-37.
95. Bain, J. D., Switzer, C., Chamberlin, A. R., and Benner, S. A. (1992) Ribosome-mediated incorporation of a non-standard amino acid into a peptide through expansion of the genetic code, *Nature JT - Nature 356*, 537-539.
96. Moore, B., Persson, B. C., Nelson, C. C., Gesteland, R. F., and Atkins, J. F. (2000) Quadruplet codons: implications for code expansion and the specification of translation step size, *J Mol Biol 298*, 195-209.
97. Hohsaka, T., Ashizuka, Y., Taira, H., Murakami, H., and Sisido, M. (2001) Incorporation of nonnatural amino acids into proteins by using various four-base codons in an Escherichia coli in vitro translation system, *40*, 11060-11064.
98. Hohsaka, T., Kajihara, D. Ashizaka, Y ., and et al. (1999) Efficient incorporation of nonnatural amino acids with large aromatic groups into streptavidin in invitro protein synthesizing systems, *J. Am. Chem. Soc. 121*, 34-40.
99. Magliery, T. J., Anderson, J. C., and Schultz, P. G. (2001) Expanding the genetic code: selection of efficient suppressors of four-base codons and identification of "shifty" four-base codons with a library approach in Escherichia coli, *J Mol Biol 307*, 755-769.
100. Hohsaka, T., Ashizuka, Y., Murakami, H., and Sisido, M. (2001) Five-base codons for incorporation of nonnatural amino acids into proteins, *Nucleic Acids Res 29*, 3646-3651.
101. Drabkin, H. J., Park, H. J., and RajBhandary, U. L. (1996) Amber suppression in mammalian cells dependent upon expression of an Escherichia coli aminoacyl-tRNA synthetase gene, *Mol Cell Biol 16*, 907-913.
102. Kowal, A. K., Kohrer, C., and RajBhandary, U. L. (2001) Twenty-first aminoacyl-tRNA synthetase-suppressor tRNA pairs for possible use in site-specific incorporation of amino acid analogues into proteins in eukaryotes and in eubacteria, *Proc Natl Acad Sci U S A 98*, 2268-2273.
103. Liu, D. R., Magliery, T. J., Pastrnak, M., and Schultz, P. G. (1997) Engineering a tRNA and aminoacyl-tRNA synthetase for the site-specific incorporation of unnatural amino acids into proteins in vivo, *Proc Natl Acad Sci U S A 94*, 10092-10097.
104. Wang, L., and Schultz, P. G. (2004) Expanding the genetic code, *Angew Chem Int Ed Engl 44*, 34-66.
105. Chin, J. W., Santoro, S. W., Martin, A. B., King, D. S., Wang, L., and Schultz, P. G. (2002) Addition of p-azido-L-phenylalanine to the genetic code of Escherichia coli, *J Am Chem Soc 124*, 9026-9027.
106. Sakamoto, K., Hayashi, A., Sakamoto, A., Kiga, D., Nakayama, H., Soma, A., Kobayashi, T., Kitabatake, M., Takio, K., Saito, K., Shirouzu, M., Hirao, I., and Yokoyama, S. (2002) Site-specific incorporation of an unnatural amino acid into proteins in mammalian cells, *Nucleic Acids Res 30*, 4692-4699.
107. Wang, L., Brock, A., and Schultz, P. G. (2002) Adding L-3-(2-Naphthyl)alanine to the genetic code of E. coli, *J Am Chem Soc 124*, 1836-1837.
108. Sonar, S., Patel, N., Fischer, W., and Rothschild, K. J. (1993) Cell-free synthesis, functional refolding, and spectroscopic characterization of bacteriorhodopsin, an integral membrane protein, *32*, 13777-13781.
109. Normanly, J., Kleina, L. G., Masson, J. M., Abelson, J., and Miller, J. H. (1990) Construction of Escherichia coli amber suppressor tRNA genes. III. Determination of tRNA specificity, *J Mol Biol JT - Journal of molecular biology 213*, 719-726.

110. Bruce, A. G., Atkins, J. F., Wills, N., Uhlenbeck, O., and Gesteland, R. F. (1982) Replacement of anticodon loop nucleotides to produce functional tRNAs: amber suppressors derived from yeast tRNAPhe, *79*, 7127-7131.
111. Bruce, A. G., and Uhlenbeck, O. C. (1982) Specific interaction of anticodon loop residues with yeast phenylalanyl-tRNA synthetase, *21*, 3921-3926.
112. Bruce, A. G., and Uhlenbeck, O. C. (1982) Enzymatic replacement of the anticodon of yeast phenylalanine transfer ribonucleic acid, *21*, 855-861.
113. Bare, L., Bruce, A. G., Gesteland, R., and Uhlenbeck, O. C. (1983) Uridine-33 in yeast tRNA not essential for amber suppression, *305*, 554-556.
114. Masson, J. M., Meuris, P., Grunstein, M., Abelson, J., and Miller, J. H. (1987) Expression of a set of synthetic suppressor tRNA(Phe) genes in Saccharomyces cerevisiae, *Proc Natl Acad Sci U S A 84*, 6815-6819.
115. Sampson, J. R., and Uhlenbeck, O. C. (1988) Biochemical and physical characterization of an unmodified yeast phenylalanine transfer RNA transcribed in vitro, *85*, 1033-1037.
116. Sampson, J. R., DiRenzo, A. B., Behlen, L. S., and Uhlenbeck, O. C. (1989) Nucleotides in yeast tRNAPhe required for the specific recognition by its cognate synthetase, *243*, 1363-1366.
117. Kipnis, D. M., Reiss, E., and Helmreich, E. (1961) Functional heterogeneity of the intracellular amino acid pool in mammalian cells, *Biochim Biophys Acta 51*, 519-524.
118. Kim, S. H., Suddath, F. L., Quigley, G. J., McPherson, A., Sussman, J. L., Wang, A. H., Seeman, N. C., and Rich, A. (1974) Three-dimensional tertiary structure of yeast phenylalanine transfer RNA, *185*, 435-440.
119. Sykes, B. D., Weingarten, H. I., and Schlesinger, M. J. (1974) Fluorotyrosine alkaline phosphatase from Escherichia coli: preparation, properties, and fluorine-19 nuclear magnetic resonance spectrum, *Proc Natl Acad Sci U S A 71*, 469-473.
120. Weber, A. L., and Miller, S. L. (1981) Reasons for the occurrence of the twenty coded protein amino acids, *J Mol Evol 17*, 273-284.
121. Bae, J. H., Rubini, M., Jung, G., Wiegand, G., Seifert, M. H., Azim, M. K., Kim, J. S., Zumbusch, A., Holak, T. A., Moroder, L., Huber, R., and Budisa, N. (2003) Expansion of the genetic code enables design of a novel "gold" class of green fluorescent proteins, *J Mol Biol 328*, 1071-1081.
122. Budisa, N., Rubini, M., Bae, J. H., Weyher, E., Wenger, W., Golbik, R., Huber, R., and Moroder, L. (2002) Global replacement of tryptophan with aminotryptophans generates non-invasive protein-based optical pH sensors, *Angew Chem Int Ed Engl 41*, 4066-4069.
123. Bronskill, P. M., and Wong, J. T. (1988) Suppression of fluorescence of tryptophan residues in proteins by replacement with 4-fluorotryptophan, *Biochem J 249*, 305-308.
124. Schuster, P. (2000) Taming combinatorial explosion, *Proc Natl Acad Sci U S A 97*, 7678-7680.
125. Prigogine, I. (1986) Life and physics. New perspectives, *Cell Biophys 9*, 217-224.
126. Barlati, S., and Ciferri, O. (1970) Incorporation of 5-methyl- and 5-hydroxy-tryptophan into the protein of Bacillus subtilis, *J Bacteriol 101*, 166-172.
127. Ardell, D. H., and Sella, G. (2001) On the evolution of redundancy in genetic codes, *J Mol Evol 53*, 269-281.
128. Forterre, P. (1997) Archaea: what can we learn from their sequences?, *Curr Opin Genet Dev 7*, 764-770.
129. Woese, C. (1998) The universal ancestor, *Proc Natl Acad Sci U S A 95*, 6854-6859.
130. Budisa, N., Alefelder, S., Bae, J. H., Golbik, R., Minks, C., Huber, R., and Moroder, L. (2001) Proteins with beta-(thienopyrrolyl)alanines as alternative chromophores and pharmaceutically active amino acids, *Protein Sci 10*, 1281-1292.
131. Crick, F. H. (1963) On the genetic code, *139*, 461-464.
132. Bae, J. H., Alefelder, S., Kaiser, J. T., Friedrich, R., Moroder, L., Huber, R., and Budisa, N. (2001) Incorporation of beta-selenolo[3,2-b]pyrrolyl-alanine into proteins for phase determination in protein X-ray crystallography, *J Mol Biol 309*, 925-936.
133. Boles, J. O., Henderson, J., Hatch, D., and Silks, L. A. (2002) Synthesis and incorporation of [6,7]-selenatryptophan into dihydrofolate reductase, *Biochem Biophys Res Commun 298*, 257-261.
134. Budisa, N., Pal, P. P., Alefelder, S., Birle, P., Krywcun, T., Rubini, M., Wenger, W., Bae, J. H., and Steiner, T. (2004) Probing the role of tryptophans in Aequorea victoria green fluorescent proteins with an expanded genetic code, *Biol Chem 385*, 191-202.

135. Igloi, G. L., von der Haar, F., and Cramer, F. (1977) Hydrolytic action of aminoacyl-tRNA synthetases from baker's yeast. "Chemical proofreading" of Thr-tRNA Val by valyl-tRNA synthetase studied with modified tRNA Val and amino acid analogues, *16*, 1696-1702.
136. Kiga, D., Sakamoto, K., Kodama, K., Kigawa, T., Matsuda, T., Yabuki, T., Shirouzu, M., Harada, Y., Nakayama, H., Takio, K., Hasegawa, Y., Endo, Y., Hirao, I., and Yokoyama, S. (2002) An engineered Escherichia coli tyrosyl-tRNA synthetase for site-specific incorporation of an unnatural amino acid into proteins in eukaryotic translation and its application in a wheat germ cell-free system, *Proc Natl Acad Sci U S A 99*, 9715-9720.
137. Anfinsen, C. B., and Corley, L. G. (1969) An active variant of staphylococcal nuclease containing norleucine in place of methionine, *J Biol Chem 244*, 5149-5152.
138. Santoro, S. W., Wang, L., Herberich, B., King, D. S., and Schultz, P. G. (2002) An efficient system for the evolution of aminoacyl-tRNA synthetase specificity, *Nat Biotechnol 20*, 1044-1048.
139. Wang, L., Xie, J., Deniz, A. A., and Schultz, P. G. (2003) Unnatural amino acid mutagenesis of green fluorescent protein, *J Org Chem 68*, 174-176.
140. Zhang, Z., Smith, B. A., Wang, L., Brock, A., Cho, C., and Schultz, P. G. (2003) A new strategy for the site-specific modification of proteins in vivo, *42*, 6735-6746.
141. Wang, L., Brock, A., Herberich, B., and Schultz, P. G. (2001) Expanding the genetic code of Escherichia coli, *292*, 498-500.
142. Yoshida, A. (1960) Studies on the mechanism of protein synthesis; incorporation of p-fluorophenylalanine into alpha-amylase of Bacillus subtilis, *Biochim Biophys Acta 41*, 98-103.
143. Richmond, M. H. (1962) The effect of amino acid analogues on growth and protein synthesis in microorganisms, *Bacteriol Rev 26*, 398-420.
144. Ibba, M., Kast, P., and Hennecke, H. (1994) Substrate specificity is determined by amino acid binding pocket size in Escherichia coli phenylalanyl-tRNA synthetase, *33*, 7107-7112.
145. Minks, C., Huber, R., Moroder, L., and Budisa, N. (2000) Noninvasive tracing of recombinant proteins with "fluorophenylalanine-fingers", *Anal Biochem 284*, 29-34.
146. Koide, H., Yokoyama, S., Katayama, Y., Muto, Y., Kigawa, T., Kohno, T., Takusari, H., Oishi, M., Takahashi, S., Tsukumo, K., and et al. (1994) Receptor-binding affinities of human epidermal growth factor variants having unnatural amino acid residues in position 23,*33*, 7470-7476.
147. Kast, P., and Hennecke, H. (1991) Amino acid substrate specificity of Escherichia coli phenylalanyl-tRNA synthetase altered by distinct mutations, *J Mol Biol 222*, 99-124.
148. Kirshenbaum, K., Carrico, I. S., and Tirrell, D. A. (2002) Biosynthesis of proteins incorporating a versatile set of phenylalanine analogues, *Chembiochem 3*, 235-237.
149. Sharma, N., Furter, R., Kast, P., and Tirrell, D. A. (2000) Efficient introduction of aryl bromide functionality into proteins in vivo, *FEBS Lett 467*, 37-40.
150. Datta, D., Wang, P., Carrico, I. S., Mayo, S. L., and Tirrell, D. A. (2002) A designed phenylalanyl-tRNA synthetase variant allows efficient in vivo incorporation of aryl ketone functionality into proteins, *J Am Chem Soc 124*, 5652-5653.
151. Beiboer, S. H., van den Berg, B., Dekker, N., Cox, R. C., and Verheij, H. M. (1996) Incorporation of an unnatural amino acid in the active site of porcine pancreatic phospholipase A2. Substitution of histidine by 1,2,4-triazole-3-alanine yields an enzyme with high activity at acidic pH, *Protein Eng 9*, 345-352.
152. Soumillion, P., and Fastrez, J. (1998) Incorporation of 1,2,4-triazole-3-alanine into a mutant of phage lambda lysozyme containing a single histidine, *Protein Eng 11*, 213-217.
153. Klein, D. C., Weller, J. L., Kirk, K. L., and Hartley, R. W. (1977) Incorporation of 2-fluoro-L-histidine into cellular protein, *Mol Pharmacol 13*, 1105-1110.
154. Dunn, B. M., DiBello, C., Kirk, K. L., Cohen, L. A., and Chaiken, I. M. (1974) Synthesis, purification, and properties of a semisynthetic ribonuclease S incorporating 4-fluoro-L-histidine at position 12, *J Biol Chem 249*, 6295-6301.
155. Ikeda, Y., Kawahara, S., Taki, M., Kuno, A., Hasegawa, T., and Taira, K. (2003) Synthesis of a novel histidine analogue and its efficient incorporation into a protein in vivo, *Protein Eng 16*, 699-706.
156. Tang, Y., and Tirrell, D. A. (2001) Biosynthesis of a highly stable coiled-coil protein containing hexafluoroleucine in an engineered bacterial host, *J Am Chem Soc 123*, 11089-11090.
157. Hortin, G., and Boime, I. (1983) Applications of amino acid analogs for studying co- and posttranslational modifications of proteins, *Methods Enzymol 96*, 777-784.

158. Apostol, I., Levine, J., Lippincott, J., Leach, J., Hess, E., Glascock, C. B., Weickert, M. J., and Blackmore, R. (1997) Incorporation of norvaline at leucine positions in recombinant human hemoglobin expressed in Escherichia coli, *J Biol Chem 272*, 28980-28988.
159. Tang, Y., and Tirrell, D. A. (2002) Attenuation of the editing activity of the Escherichia coli leucyl-tRNA synthetase allows incorporation of novel amino acids into proteins in vivo, *41*, 10635-10645.
160. Wilson, M. J., and Hatfield, D. L. (1984) Incorporation of modified amino acids into proteins in vivo, *Biochim Biophys Acta 781*, 205-215.
161. Porter, T. H., Smith, S. C., and Shive, W. (1977) Inhibition of valine utilization by cyclobutaneglycine, *Arch Biochem Biophys 179*, 266-271.
162. Doring, V., Mootz, H. D., Nangle, L. A., Hendrickson, T. L., de Crecy-Lagard, V., Schimmel, P., and Marliere, P. (2001) Enlarging the amino acid set of Escherichia coli by infiltration of the valine coding pathway, *292*, 501-504.
163. Wang, P., Tang, Y., and Tirrell, D. A. (2003) Incorporation of trifluoroisoleucine into proteins in vivo, *J Am Chem Soc 125*, 6900-6906.
164. Budisa, N., Huber, R., Golbik, R., Minks, C., Weyher, E., and Moroder, L. (1998) Atomic mutations in annexin V--thermodynamic studies of isomorphous protein variants, *Eur J Biochem 253*, 1-9.
165. Cowie, D. B., and Cohen, G. N. (1957) Biosynthesis by Escherichia coli of active altered proteins containing selenium instead of sulfur, *Biochim Biophys Acta 26*, 252-261.
166. Bogosian, G., Violand, B. N., Dorward-King, E. J., Workman, W. E., Jung, P. E., and Kane, J. F. (1989) Biosynthesis and incorporation into protein of norleucine by Escherichia coli, *J Biol Chem 264*, 531-539.
167. Gilles, A. M., Marliere, P., Rose, T., Sarfati, R., Longin, R., Meier, A., Fermandjian, S., Monnot, M., Cohen, G. N., and Barzu, O. (1988) Conservative replacement of methionine by norleucine in Escherichia coli adenylate kinase, *J Biol Chem 263*, 8204-8209.
168. Budisa, N., Steipe, B., Demange, P., Eckerskorn, C., Kellermann, J., and Huber, R. (1995) High-level biosynthetic substitution of methionine in proteins by its analogs 2-aminohexanoic acid, selenomethionine, telluromethionine and ethionine in Escherichia coli, *Eur J Biochem 230*, 788-796.
169. Boles, J. O., Lewinski, K., Kunkle, M., Odom, J. D., Dunlap, B., Lebioda, L., and Hatada, M. (1994) Bio-incorporation of telluromethionine into buried residues of dihydrofolate reductase, *Nat Struct Biol 1*, 283-284.
170. Duewel, H., Daub, E., Robinson, V., and Honek, J. F. (1997) Incorporation of trifluoromethionine into a phage lysozyme: implications and a new marker for use in protein 19F NMR, *Date of Input: 18/07/2007 36*, 3404-3416.
171. Jakubowski, H., and Goldman, E. (1992) Editing of errors in selection of amino acids for protein synthesis, *Microbiol Rev 56*, 412-429.
172. Fersht, A. R., and Dingwall, C. (1979) An editing mechanism for the methionyl-tRNA synthetase in the selection of amino acids in protein synthesis, *Date of Input: 18/07/2007 18*, 1250-1256.
173. Jakubowski, H. (2000) Translational incorporation of S-nitrosohomocysteine into protein, *J Biol Chem 275*, 21813-21816.
174. Kiick, K. L., van Hest, J. C., and Tirrell, D. A. (2000) Expanding the Scope of Protein Biosynthesis by Altering the Methionyl-tRNA Synthetase Activity of a Bacterial Expression Host, *Angew Chem Int Ed Engl 39*, 2148-2152.
175. Reimer, U., Scherer, G., Drewello, M., Kruber, S., Schutkowski, M., and Fischer, G. (1998) Side-chain effects on peptidyl-prolyl cis/trans isomerisation, *J Mol Biol 279*, 449-460.
176. Renner, C., Alefelder, S., Bae, J. H., Budisa, N., Huber, R., and Moroder, L. (2001) Fluoroprolines as Tools for Protein Design and Engineering , *Angew Chem Int Ed Engl 40*, 923-925.
177. Jester, B. C., Levengood, J. D., Roy, H., Ibba, M., and Devine, K. M. (2003) Nonorthologous replacement of lysyl-tRNA synthetase prevents addition of lysine analogues to the genetic code, *Proc Natl Acad Sci U S A 100*, 14351-14356.
178. Islam, M. S., Berggren, P. O., and Larsson, O. (1993) Sulfhydryl oxidation induces rapid and reversible closure of the ATP-regulated K+ channel in the pancreatic beta-cell, *FEBS Lett 319*, 128-132.
179. Lutsenko, S., Daoud, S., and Kaplan, J. H. (1997) Identification of two conformationally sensitive cysteine residues at the extracellular surface of the Na,K-ATPase alpha-subunit, *J Biol Chem 272*, 5249-5255.
180. Broillet, M. C. (2000) A single intracellular cysteine residue is responsible for the activation of the olfactory cyclic nucleotide-gated channel by NO, *J Biol Chem 275*, 15135-15141.

181. Schechter, Y., Patchornik, A., and Burstein, Y. (1973) Selective reduction of cystine 1-8 in alpha-lactalbumin, *12*, 3407-3413.
182. Jennings, M. L., and Anderson, M. P. (1987) Chemical modification and labeling of glutamate residues at the stilbenedisulfonate site of human red blood cell band 3 protein, *J Biol Chem 262*, 1691-1697.
183. Jennings, M. L., and Smith, J. S. (1992) Anion-proton cotransport through the human red blood cell band 3 protein. Role of glutamate 681, *J Biol Chem 267*, 13964-13971.
184. Matacic, S. S., and Loewy, A. G. (1979) Presence of the epsilon-(gamma-glutamic)lysine crosslink in cellular proteins, *Biochim Biophys Acta 576*, 263-268.
185. Ackers, G. K., and Smith, F. R. (1986) Resolving pathways of functional coupling within protein assemblies by site-specific structural perturbation, *Biophys J 49*, 155-165.
186. Dombroski, A. J., and Platt, T. (1988) Structure of rho factor: an RNA-binding domain and a separate region with strong similarity to proven ATP-binding domains, *Proc Natl Acad Sci U S A 85*, 2538-2542.
187. Miles, E. W., Kawasaki, H., Ahmed, S. A., Morita, H., and Nagata, S. (1989) The beta subunit of tryptophan synthase. Clarification of the roles of histidine 86, lysine 87, arginine 148, cysteine 170, and cysteine 230, *J Biol Chem 264*, 6280-6287.
188. Heckler, T. G., Chang, L. H., Zama, Y., Naka, T., Chorghade, M. S., and Hecht, S. M. (1984) T4 RNA ligase mediated preparation of novel "chemically misacylated" tRNAPheS, , 1468-1473.
189. Hohsaka, T., and Sisido, M. (2002) Incorporation of non-natural amino acids into proteins, *Curr Opin Chem Biol 6*, 809-815.
190. Robertson, S. A., Noren, C. J., Anthony-Cahill, S. J., Griffith, M. C., and Schultz, P. G. (1989) The use of 5'-phospho-2 deoxyribocytidylylriboadenosine as a facile route to chemical aminoacylation of tRNA, *Nucleic Acids Res JT - Nucleic acids research 17*, 9649-9660.
191. Robertson, A., Ellman, A., and Schultz, G. (1991) A general and efficient route fro chemical aminocylation of transfer RNAs, *Journal of American Chemical Society 113*, 2722-2729.
192. Hecht, S. M., Alford, B. L., Kuroda, Y., and Kitano, S. (1978) "Chemical aminoacylation" of tRNA's, *J Biol Chem 253*, 4517-4520.
193. Praetorius-Ibba, M., Stange-Thomann, N., Kitabatake, M., Ali, K., Soll, I., Carter CW, J. r., Ibba, M., and Soll, D. (2000) Ancient adaptation of the active site of tryptophanyl-tRNA synthetase for tryptophan binding, , 13136-13143.
194. Thiebe, R. (1983) Arginyl-tRNA synthetase from brewer's yeast. Purification, properties, and steady-state mechanism, *Eur J Biochem 130*, 517-524.
195. Schreier, A. A., and Schimmel, P. R. (1972) Transfer ribonucleic acid synthetase catalyzed deacylation of aminoacyl transfer ribonucleic acid in the absence of adenosine monophosphate and pyrophosphate, *11*, 1582-1589.
196. Yarus, M. (1972) Phenylalanyl-tRNA synthetase and isoleucyl-tRNA Phe : a possible verification mechanism for aminoacyl-tRNA, *69*, 1915-1919.
197. Budisa, N. (2004) Prolegomena to future experimental efforts on genetic code engineering by expanding its amino acid repertoire, *Angew Chem Int Ed Engl 43*, 6426-6463.
198. Tang, Y., Ghirlanda, G., Petka, W. A., Nakajima, T., DeGrado, W. F., and Tirrell, D. A. (2001) Fluorinated Coiled-Coil Proteins Prepared In Vivo Display Enhanced Thermal and Chemical Stability *Angew Chem Int Ed Engl 40*, 1494-1496.
199. Sonar, S., Lee, C. P., Coleman, M., Patel, N., Liu, X., Marti, T., Khorana, H. G., RajBhandary, U. L., and Rothschild, K. J. (1994) Site-directed isotope labelling and FTIR spectroscopy of bacteriorhodopsin, *Nat Struct Biol 1*, 512-517.
200. Hartman, M. C., Josephson, K., Lin, C. W., and Szostak, J. W. (2007) An expanded set of amino acid analogs for the ribosomal translation of unnatural peptides, *PLoS One 2*, e972.
201. Hartman, M. C., Josephson, K., and Szostak, J. W. (2006) Enzymatic aminoacylation of tRNA with unnatural amino acids, *Proc Natl Acad Sci U S A 103*, 4356-4361.
202. Johnson, A. E., Woodward, W. R., Herbert, E., and Menninger, J. R. (1976) Nepsilon-acetyllysine transfer ribonucleic acid: a biologically active analogue of aminoacyl transfer ribonucleic acids, *15*, 569-575.
203. Wiedmann, M., Kurzchalia, T. V., Bielka, H., and Rapoport, T. A. (1987) Direct probing of the interaction between the signal sequence of nascent preprolactin and the signal recognition particle by specific cross-linking, *J Cell Biol 104*, 201-208.

204. Liu, D. R., Magliery, T. J., and Schultz, P. G. (1997) Characterization of an 'orthogonal' suppressor tRNA derived from E. coli tRNA2(Gln), *Chem Biol 4*, 685-691.
205. Liu, D. R., and Schultz, P. G. (1999) Progress toward the evolution of an organism with an expanded genetic code, *Proc Natl Acad Sci U S A 96*, 4780-4785.
206. Wang, L., Magliery, T. J., Liu, D. R., and Schultz, P. G. (2000) A new functional suppressor tRNA/aminoacyl tRNA synthetase pair for the in vivo incorporation of unnatural amino acids into proteins., *J.Amer.Chem.Soc. 122*, 5010-5011.
207. Anderson, J. C., Wu, N., Santoro, S. W., Lakshman, V., King, D. S., and Schultz, P. G. (2004) An expanded genetic code with a functional quadruplet codon, *Proc Natl Acad Sci U S A 101*, 7566-7571.
208. Anderson, J. C., and Schultz, P. G. (2003) Adaptation of an orthogonal archaeal leucyl-tRNA and synthetase pair for four-base, amber, and opal suppression, *42*, 9598-9608.
209. Thorson, J. S., Cornish, V. W., Barrett, J. E., Cload, S. T., Yano, T., and Schultz, P. G. (1998) A biosynthetic approach for the incorporation of unnatural amino acids into proteins, *Methods Mol Biol 77*, 43-73.
210. Cload, S. T., Liu, D. R., Froland, W. A., and Schultz, P. G. (1996) Development of improved tRNAs for in vitro biosynthesis of proteins containing unnatural amino acids, *Chem Biol 3*, 1033-1038.
211. Bain, J. D., Diala, E. S., Glabe, C. G., Wacker, D. A., Lyttle, M. H., Dix, T. A., and Chamberlin, A. R. (1991) Site-specific incorporation of nonnatural residues during in vitro protein biosynthesis with semisynthetic aminoacyl-tRNAs, *30*, 5411-5421.
212. Kwok, Y., and Wong, J. T. (1980) Evolutionary relationship between Halobacterium cutirubrum and eukaryotes determined by use of aminoacyl-tRNA synthetases as phylogenetic probes, *Can J Biochem 58*, 213-218.
213. Koide, K., Finkelstein, J. M., Ball, Z., and Verdine, G. L. (2001) A synthetic library of cell-permeable molecules, *J Am Chem Soc 123*, 398-408.
214. Wang, L., and Schultz, P. G. (2002) Expanding the genetic code, *Chem Commun (Camb)*, 1-11.
215. Furter, R. (1998) Expansion of the genetic code: site-directed p-fluoro-phenylalanine incorporation in Escherichia coli, *Protein Sci 7*, 419-426.
216. Furter, R. (1998) Expansion of the genetic code: site-directed p-fluoro-phenylalanine incorporation in Escherichia coli, *Protein Sci 7*, 419-426.
217. Sonar, S., Marti, T., Rath, P., Fischer, W., Coleman, M., Nilsson, A., Khorana, H. G., and Rothschild, K. J. (1994) A redirected proton pathway in the bacteriorhodopsin mutant Tyr-57-->Asp. Evidence for proton translocation without Schiff base deprotonation, *J Biol Chem 269*, 28851-28858.
218. Liu, X. M., Sonar, S., Lee, C. P., Coleman, M., RajBhandary, U. L., and Rothschild, K. J. (1995) Site-directed isotope labeling and FTIR spectroscopy: assignment of tyrosine bands in the bR-->M difference spectrum of bacteriorhodopsin, *Biophys Chem JT - Biophysical chemistry 56*, 63-70.
219. Ludlam, C. F., Sonar, S., Lee, P., Ludlam, Nilson, Coleman, M., Herzfeld, J., Rajbhandary, U. L., and Rothschild, K. J. (1995) Involvement of Tyr-185 in The bacteriorhodopsin Eó>C Protein Switch, 1-21.
220. Ludlam, C. F., Sonar, S., Lee, C. P., Coleman, M., Herzfeld, J., RajBhandary, U. L., and Rothschild, K. J. (1995) Site-directed isotope labeling and ATR-FTIR difference spectroscopy of bacteriorhodopsin: the peptide carbonyl group of Tyr 185 is structurally active during the bR-->N transition, *34*, 2-6.
221. Sonar, S., Liu, M., Lee, P., Coleman, M., He, W., Pelletier, Herzfeld, J., RajBhandary, U. L., and Rothschild, K. J. (1995) Site-Directed Isotope Labelling and FT-IR Spectroscopy: The Tyr 185/Pro 186 Peptide Bond of Bacteriorhodopsin Is Perturbed during the primary Photoreaction, *Journal of American Chemical Society 117*, 11614-11615.
222. Thorson, J. C. E. M. E. C., and et al. (1995) Linear free energy analysis of hydrogen bonding in proteins, *J. Am. Chem. Soc. 117*, 1157-1158.
223. Thorson, J. C. E. S. P. G. (1995) Analysis of hydrogen bonding strength in proteins using unnatural amino acids, *J. Am. Chem. Soc. 117*, 9361-9362.
224. Hohsaka, T., Sato, K., Sisido, M., Takai, K., and Yokoyama, S. (1994) Site-specific incorporation of photofunctional nonnatural amino acids into a polypeptide through in vitro protein biosynthesis, *FEBS Lett 344*, 171-174.
225. Steward, L. E. C. C. S. G. M. A., and et al. (1997) In vitro site specific incorporation of fluorescent probes into beta-galactosidase, *J. Am. Chem. Soc. 119*, 6-11.

226. Cornish, V. W., Benson, D. R., Altenbach, C. A., Hideg, K., Hubbell, W. L., and Schultz, P. G. (1994) Site-specific incorporation of biophysical probes into proteins, *Proc Natl Acad Sci U S A 91*, 2910-2914.
227. Cornish, V. W. H. K. M. S. P. G. (1996) Site-specific protein modification using a ketone handle, *J. Am. Chem. Soc. 118*, 8150-8151.
228. Soumillion, P., and Fastrez, J. (1998) Incorporation of 1,2,4-triazole-3-alanine into a mutant of phage lambda lysozyme containing a single histidine, *Protein Eng 11*, 213-217.
229. Mchaourab, H. S., Lietzow, M. A., Hideg, K., and Hubbell, W. L. (1996) Motion of spin-labeled side chains in T4 lysozyme. Correlation with protein structure and dynamics, *35*, 7692-7704.
230. Ellman, J. A., Mendel, D., and Schultz, P. G. (1992) Site-specific incorporation of novel backbone structures into proteins, *Science JT - Science (New York, N.Y.) 255*, 197-200.
231. Mendel, D., Ellman, J. A., Chang, Z., Veenstra, D. L., Kollman, P. A., and Schultz, P. G. (1992) Probing protein stability with unnatural amino acids, *7 256*, 1798-1802.
232. Chung, H. H., Benson, D. R., Cornish, V. W., and Schultz, P. G. (1993) Probing the role of loop 2 in Ras function with unnatural amino acids, *Proc Natl Acad Sci U S A 90*, 10145-10149.
233. Chung, H. H., Benson, D. R., and Schultz, P. G. (1993) Probing the structure and mechanism of Ras protein with an expanded genetic code, *Science JT - Science (New York, N.Y.) 259*, 806-809.
234. Judice, J. K., Gamble, T. R., Murphy, E. C., AM, d. V., and Schultz, P. G. (1993) Probing the mechanism of staphylococcal nuclease with unnatural amino acids: kinetic and structural studies, *Science JT - Science (New York, N.Y.) 261*, 1578-1581.
235. Cornish, V. W., Kaplan, M. I., Veenstra, D. L., Kollman, P. A., and Schultz, P. G. (1994) Stabilizing and destabilizing effects of placing beta-branched amino acids in protein alpha-helices, *33*, 12022-12031.
236. Karginov, A. V., Lodder, M., and Hecht, S. M. (1999) Facile characterization of translation initiation via nonsense codon suppression, *Nucleic Acids Res 27*, 3283-3290.
237. Nowak, M. W., Kearney, P. C., Sampson, J. R., Saks, M. E., Labarca, C. G., Silverman, S. K., Zhong, W., Thorson, J., Abelson, J. N., Davidson, N., and et, a. l. (1995) Nicotinic receptor binding site probed with unnatural amino acid incorporation in intact cells, *Science JT - Science (New York, N.Y.) 268*, 439-442.
238. Unwin, P. N., and Henderson, R. (1975) Molecular structure determination by electron microscopy of unstained crystalline specimens, *J Mol Biol 94*, 425-440.
239. Rothschild, K. J., and Clark, N. A. (1979) Polarized infrared spectroscopy of oriented purple membrane, *Biophys J 25*, 473-487.
240. Ovchinnikov, A., Abdulaev, N. G., Feigina, Y., Kiselev, A. V. L. N. A., and Nazimov, I. V. (1978) Amino acid sequence of bacteriorhodopsin, *Bioorganic Chemistry 4*, 1573-.
241. Khorana, H. G., Gerber, G. E., Herlihy, W. C., Gray, C. P., Anderegg, R. J., Nihei, K., and Biemann, K. (1979) Amino acid sequence of bacteriorhodopsin, *Proc Natl Acad Sci U S A 76*, 5046-5050.
242. Engelman, D. M., Henderson, R., McLachlan, A. D., and Wallace, B. A. (1980) Path of the polypeptide in bacteriorhodopsin, *Proc Natl Acad Sci U S A 77*, 2023-2027.
243. Kimura, K., Mason, T. L., and Khorana, H. G. (1982) Immunological probes for bacteriorhodopsin. Identification of three distinct antigenic sites on the cytoplasmic surface, *J Biol Chem 257*, 2859-2867.
244. Ovchinnikov, A., Abdulaev, N. G., Vasilov, G., Vturina, I. Y., Kuryatov, B., and Kiselev, A. V. (1985) The antigenic structure and topography of bacteriorhodopsin in purple membranes as determined by interaction with monoclonal antibodies, *FEBS Lett JT - FEBS letters 179*, 343-.
245. Fimmel, S., Choli, T., Dencher, N. A., Buldt, G., and Wittmann-Liebold, B. (1989) Topography of surface-exposed amino acids in the membrane protein bacteriorhodopsin determined by proteolysis and micro-sequencing, *Biochim Biophys Acta 978*, 231-240.
246. Turner, G. J., Chittiboyina, S., Pohren, L., Hines, K. G., Correia, J. J., and Mitchell, D. C. (2009) The bacteriorhodopsin carboxyl-terminus contributes to proton recruitment and protein stability, *48*, 1112-1122.
247. Bridgen, J., and Walker, I. D. (1976) Photoreceptor protein from the purple membrane of Halobacterium halobium. Molecular weight and retinal binding site, *15*, 792-798.
248. Lemke, D., and Oesterhelt, D. (1981) Lysine 216 is a binding site of the retinyl moiety in bacteriorhodopsin, *FEBS Lett JT - FEBS letters 128*, 255-.
249. Rothschild, K. J., Argade, P. V., Earnest, T. N., Huang, K. S., London, E., Liao, M. J., Bayley, H., Khorana, H. G., and Herzfeld, J. (1982) The site of attachment of retinal in bacteriorhodopsin. A resonance Raman study, *J Biol Chem 257*, 8592-8595.

250. Rothschild, K. J., Sanches, R., Hsiao, T. L., and Clark, N. A. (1980) A spectroscopic study of rhodopsin alpha-helix orientation, *Biophys J 31*, 53-64.
251. Heyn, M. P., Cherry, R. J., and Muller, U. (1977) Transient and linear dichroism studies on bacteriorhodopsin: determination of the orientation of the 568 nm all-trans retinal chromophore, *J Mol Biol 117*, 607-620.
252. Lin, S. W., and Mathies, R. A. (1989) Orientation of the protonated retinal Schiff base group in bacteriorhodopsin from absorption linear dichroism, *Biophys J 56*, 653-660.
253. Heyn, M. P., Westerhausen, J., Wallat, I., and Seiff, F. (1988) High-sensitivity neutron diffraction of membranes: Location of the Schiff base end of the chromophore of bacteriorhodopsin, *Proc Natl Acad Sci U S A 85*, 2146-2150.
254. Ulrich, A. S., Watts, A., Wallat, I., and Heyn, M. P. (1994) Distorted structure of the retinal chromophore in bacteriorhodopsin resolved by 2H-NMR, *33*, 5370-5375.
255. Huang, K. S., Radhakrishnan, R., Bayley, H., and Khorana, H. G. (1982) Orientation of retinal in bacteriorhodopsin as studied by cross-linking using a photosensitive analog of retinal, *J Biol Chem JT - The Journal of biological chemistry 257*, 13616-13623.
256. Mogi, Stern, L. J., Marti, T., Chao, B. H., and Khorana, H. G. (1988) Structure-function studies on bacteriorhodopsin. VII. Aspartic acid substitutions affect proton translocation by bacteriorhodopsin, *Proceeding of National Acamedy of Sciences of the United States of America 85*, 4148-.
257. Stern, L. J., and Khorana, H. G. (1989) Structure-function studies on bacteriorhodopsin. X. Individual substitutions of arginine residues by glutamine affect chromophore formation, photocycle, and proton translocation, *J Biol Chem JT - The Journal of biological chemistry 264*, 14202-14208.
258. Braiman, M. S., Mogi, T., Stern, L. J., Hackett, N. R., Chao, B. H., Khorana, H. G., and Rothschild, K. J. (1988) Vibrational spectroscopy of bacteriorhodopsin mutants: I. Tyrosine-185 protonates and deprotonates during the photocycle, *Proteins 3*, 219-229.
259. Ahl, P. L., Stern, L. J., Mogi, T., Khorana, H. G., and Rothschild, K. J. (1989) Substitution of amino acids in helix F of bacteriorhodopsin: effects on the photochemical cycle, *28*, 10028-10034.
260. Mogi, T., Stern, L. J., Chao, B. H., and Khorana, H. G. (1989) Structure-function studies on bacteriorhodopsin. VIII. Substitutions of the membrane-embedded prolines 50, 91, and 186: the effects are determined by the substituting amino acids, *J Biol Chem JT - The Journal of biological chemistry 264*, 14192-14196.
261. Rothschild, K. J., Gray, D., Mogi, T., Marti, T., Braiman, M. S., Stern, L. J., and Khorana, H. G. (1989) Vibrational spectroscopy of bacteriorhodopsin mutants: chromophore isomerization perturbs tryptophan-86, *Date of Input: 10/04/2007 28*, 7052-7059.
262. Rothschild, K. J., Braiman, M. S., He, Y. W., Marti, T., and Khorana, H. G. (1990) Vibrational spectroscopy of bacteriorhodopsin mutants. Evidence for the interaction of aspartic acid 212 with tyrosine 185 and possible role in the proton pump mechanism, *J Biol Chem 265*, 16985-16991.
263. Rothschild, K. J., He, Y. W., Mogi, T., Marti, T., Stern, L. J., and Khorana, H. G. (1990) Vibrational spectroscopy of bacteriorhodopsin mutants: evidence for the interaction of proline-186 with the retinylidene chromophore, *Date of Input: 18/07/2007 29*, 5954-5960.
264. Braiman, M. S., Mogi, T., Marti, T., Stern, L. J., Khorana, H. G., and Rothschild, K. J. (1988) Vibrational spectroscopy of bacteriorhodopsin mutants: light-driven proton transport involves protonation changes of aspartic acid residues 85, 96, and 212, *Date of Input: 18/07/2007 27*, 8516-8520.
265. Rothschild, K. J., Braiman, M. S., Mogi, T., Stern, L. J., and Khorana, H. G. (1989) Conserved amino acids in F-helix of bacteriorhodopsin form part of a retinal binding pocket, *FEBS Lett JT - FEBS letters 250*, 448-452.
266. Henderson, R., Baldwin, J. M., Ceska, T. A., Zemlin, F., Beckmann, E., and Downing, K. H. (1990) Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy, *J Mol Biol JT - Journal of molecular biology 213*, 899-929.
267. Lozier, R. H., Xie, A., Hofrichter, J., and Clore, G. M. (1992) Reversible steps in the bacteriorhodopsin photocycle, *Proc Natl Acad Sci U S A 89*, 3610-3614.
268. Varo, G., and Lanyi, J. K. (1991) Thermodynamics and energy coupling in the bacteriorhodopsin photocycle, *Date of Input: 18/07/2007 30*, 5016-5022.
269. Varo, G., and Lanyi, J. K. (1991) Distortions in the photocycle of bacteriorhodopsin at moderate dehydration, *Biophys J 59*, 313-322.

270. Muller, H., Butt, J., Bamberg, E., Fedler, Hess, B., Siebert, F., and Engelhard, M. (1991) The reaction cycle of bacteriorhodopsin: An analysis using visible absorption,photocurrents and infrared techniques, *Eur Biophys J JT - European biophysics journal : EBJ 19*, 241-251.
271. Hofrichter, J., Henry, E. R., and Lozier, R. H. (1989) Photocycles of bacteriorhodopsin in light- and dark-adapted purple membrane studied by time-resolved absorption spectroscopy, *Biophys J 56*, 693-706.
272. Chernavskii, D. S., Chizhov, I. V., Lozier, R. H., Murina, T. M., Prokhorov, A. M., and Zubov, B. V. (1989) Kinetic model of bacteriorhodopsin photocycle: pathway from M state to bR, *Photochem Photobiol 49*, 649-653.
273. Dancshazy, Govindjee, R., Nelson, and Ebrey, G. (1986) A new intermediate in the photocycle of bacteriorhodopsin, *FEBS Lett 209*, 44-.
274. Pollard, T., Brito Cruz, C. H., Shank, V., and Mathies, R. A. (1989) Direct observation of the excited-state cis-trans photoisomerization of the bacteriorhodopsin: Multilevel line shape theory femtosecond dynamic hole-burning and its applications, *J.Chem.Phy 90*, 199-208.
275. Petrich, W., Breton, Martin, L., and A, A. A. (1987) Femtosecond absorption spectroscopy of light-adapted and dark-adapted bacteriorhodopsin, *Chem Phys Lett 137*, 369-.
276. Varo, G., and Lanyi, J. K. (1990) Pathways of the rise and decay of the M photointermediate(s) of bacteriorhodopsin, *Date of Input: 10/04/2007 29*, 2241-2250.
277. Scherrer, Mathew, M. K., Sperling, and Soeckenius. (1989) Retinal Isomer Ratio in Dark- Adapted Purple Membrane and Bacteriorhodopsin Monomers, *Date of Input: 10/04/2007 28*, 829-834.
278. Lugtenburg, J., Mathies, R. A., Griffin, R. G., and Herzfeld, J. (1988) Structure and function of rhodopsins from solid state NMR and resonance Raman spectroscopy of isotopic retinal derivatives, *Trends Biochem Sci 13*, 388-393.
279. Farrar, M. R., Lakshmi, K. V., Smith, S. O., Brown, R. S., Raap, J., Lugtenburg, J., Griffin, R. G., and Herzfeld, J. (1993) Solid state NMR study of [epsilon-13C]Lys-bacteriorhodopsin: Schiff base photoisomerization, *Biophys J 65*, 310-315.
280. Smith.S.O., Braiman, M. S., Myers, B., Pardoen, J. A. C., J.M.L., Winkel, Lugtenburg, J., and Mathies, R. A. (1987) Vibrational analysis of the all-trans-retinal chromophores in light-adapted bacteriorhodopsin, *Journal of American Chemical Society 1987*, 3108-.
281. Harbison, G. S., Smith, S. O., Pardoen, J. A., Courtin, J. M., Lugtenburg, J., Herzfeld, J., Mathies, R. A., and Griffin, R. G. (1985) Solid-state 13C NMR detection of a perturbed 6-s-trans chromophore in bacteriorhodopsin, *Date of Input: 18/07/2007 24*, 6955-6962.
282. Harbison, G. S., Smith, S. O., Pardoen, J. A., Winkel, C., Lugtenburg, J., Herzfeld, J., Mathies, R., and Griffin, R. G. (1984) Dark-adapted bacteriorhodopsin contains 13-cis, 15-syn and all-trans, 15-anti retinal Schiff bases, *Proc Natl Acad Sci U S A 81*, 1706-1709.
283. Braiman, M., and Mathies, R. (1982) Resonance Raman spectra of bacteriorhodopsin's primary photoproduct: evidence for a distorted 13-cis retinal chromophore, *Proc Natl Acad Sci U S A 79*, 403-407.
284. Rothschild, K. J., and Marrero, H. (1982) Infrared evidence that the Schiff base of bacteriorhodopsin is protonated: bR570 and K intermediates, *Date of Input: 23/04/2007 79*, 4045-4049.
285. Atkinson, G. H., Blanchard, D., Lemaire, H., Brack, T. L., and Hayashi, H. (1989) Picosecond time-resolved fluorescence spectroscopy of K-590 in the bacteriorhodopsin photocycle, *Biophys J 55*, 263-274.
286. Atkinson, G. H., Brack, T. L., Blanchard, D., and Rumbles. (1989) Picosecond time-resolved resonance Raman spectroscopy of the initial trans to cis isomerization in the bacteriorhodopsin photocycle, *Chem Phys 131*, 1-.
287. Rothschild, K. J., Roepe, P., Lugtenburg, J., and Pardoen, J. A. (1984) Fourier transform infrared evidence for Schiff base alteration in the first step of the bacteriorhodopsin photocycle, *Date of Input: 10/04/2007 23*, 6103-6109.
288. Rothschild, K. J. (1992) FTIR difference spectroscopy of bacteriorhodopsin: toward a molecular model, *J Bioenerg Biomembr 24*, 147-167.
289. Kakitani, H., Kakitani, T., Rodman, H., and Honig, B. (1983) Correlation of Vibrational Frequencies with Absorption Maxima in Polyenes, Rhodopsin, Bacteriorhodopsin and Retinal Analogues, *The Journal of Physical Chemistry 87*, 3620-3628.
290. Baasov, T., Friedman, N., and Sheves, M. (1987) Factors affecting the C = N stretching in protonated retinal Schiff base: a model study for bacteriorhodopsin and visual pigments, *Date of Input: 18/07/2007 26*, 3210-3217.

291. Gilson, H. S., Honig, B. H., Croteau, A., Zarrilli, G., and Nakanishi, K. (1988) Analysis of the factors that influence the C=N stretching frequency of polyene Schiff bases. Implications for bacteriorhodopsin and rhodopsin, *Biophys J 53*, 261-269.
292. Argade, P. V., and Rothschild, K. J. (1983) Quantative Analysis of Resonance Raman Spectra of Purple Membrane from Halobacterium halobium: L550 Intermediate, *Date of Input: 10/04/2007 22*, 3460-3466.
293. Fodor, S. P., Pollard, W. T., Gebhard, R., EM, v. d. B., Lugtenburg, J., and Mathies, R. A. (1988) Bacteriorhodopsin's L550 intermediate contains a C14-C15 s-trans-retinal chromophore, *Date of Input: 23/04/2007 85*, 2156-2160.
294. Sasaki, J., Maeda, A., Kato, C., and Hamaguchi, H. (1993) Time-resolved infrared spectral analysis of the KL-to-L conversion in the photocycle of bacteriorhodopsin, *Date of Input: 18/07/2007 32*, 867-871.
295. Gerwert, K., and Siebert, F. (1986) Evidence for light-induced 13-cis, 14-s-cis isomerization in bacteriorhodopsin obtained by FTIR difference spectroscopy using isotopically labelled retinals, *EMBO J JID- 8208664 5*, 805-811.
296. Tavan, P., and Schulten. (1986) Evidence for a 13,14-cis cycle in bacteriorhodopsin, *Biophys Journal 50*, 81-.
297. Lewis, A., Spoonhower, J., Bogomolni, R. A., Lozier, R. H., and Stoeckenius, W. (1974) Tunable laser resonance raman spectroscopy of bacteriorhodopsin, *Proc Natl Acad Sci U S A 71*, 4462-4466.
298. Fahmy, Siebert, F., Grossjean, M. F., and avan. (1989) Photoisomerization in bacteriorhodopsin studied by FTIR, linear dichroism, and photoselection experiments combined with quantum chemical theoretical analysis, *Journal of Molecular Structure 214*, 257-288.
299. Braiman, M., and Mathies, R. (1980) Resonance Raman evidence for an all-trans to 13-cis isomerization in the proton-pumping cycle of bacteriorhodopsin, *Date of Input: 18/07/2007 19*, 5421-5428.
300. Ames, J. B., Fodor, S. P., Gebhard, R., Raap, J., van den Berg, E., Lugtenburg, J., and Mathies, R. A. (1989) Bacteriorhodopsin's M412 Intermediate contains a 13-cis, 14-s-trans, 15-anti- Retianl Schiff Base Chromophore, *Date of Input: 10/04/2007 28*, 3681-3687.
301. Nakagawa, Ogura, Maeda, A., and Kitagawa. (1989) Transient Resonance Raman Spectra of Neutral and Alkaline Bacteriorhodopsin Photointermedites Observed with a Double-Beam Flow Apparatus: Presence of very Fast Decaying M412, *Date of Input: 10/04/2007 28*, 1347-1352.
302. Deng, H., Pande, C., Callender, R. H., and Ebrey, T. G. (1985) A detailed resonance Raman study of the M412 intermediate in the bacteriorhodopsin photocycle, *Photochem Photobiol 41*, 467-470.
303. Lozier, R. H., Bogomolni, R. A., and Stoeckenius, W. (1975) Bacteriorhodopsin: a light-driven proton pump in Halobacterium Halobium, *Biophys J 15*, 955-962.
304. Balashov, S. P., Imasheva, E. S., Litvin, F., and Hozier, H. (1990) The N Intermediate of bacteriorhodopsin at loe temperatures: Stabilization and photoconversion, *FEBS letters 271*, 93-96.
305. Kouyama, T., Nasuda-Kouyama, A., Ikegami, A., Mathew, M. K., and Stoeckenius, W. (1988) Bacteriorhodopsin photoreaction: identification of a long-lived intermediate N (P,R350) at high pH and its M-like photoproduct, *Date of Input: 18/07/2007 27*, 5855-5863.
306. Ames, J. B., and Mathies, R. A. (1990) The Role of Back- Reactions and Proton Uptake during the N -> O Transition in Bacteriorhodopsin's Photocycle: A Kinetic Resonance Raman Study, *Date of Input: 10/04/2007 29*, 7181-7190.
307. Lakshmi, K. V., Farrar, M. R., Raap, J., Lugtenburg, J., Griffin, R. G., and Herzfeld, J. (1994) Solid state 13C and 15N NMR investigations of the N intermediate of bacteriorhodopsin, *Date of Input: 18/07/2007 33*, 8853-8857.
308. Drachev, A., Kaulen, A. D., Khorana, H. G., Mogi, T., Otto, H., P., k. V., and Heyn, M. P. H. M. (1989) Participation of the Asp-96 carboxyl in hydrogen ion transfer along the inward proton-conducting pathway of bacteriorhodopsin, *Date of Input: 18/07/2007 54*, 1467-.
309. Otto, H., Marti, T., Holz, M., Mogi, T., Lindau, M., Khorana, H. G., and Heyn, M. P. (1989) Aspartic acid-96 is the internal proton donor in the reprotonation of the Schiff base of bacteriorhodopsin, *Proc Natl Acad Sci U S A 86*, 9228-9232.
310. Bousche, O., Braiman, M., He, Y. W., Marti, T., Khorana, H. G., and Rothschild, K. J. (1991) Vibrational spectroscopy of bacteriorhodopsin mutants. Evidence that ASP-96 deprotonates during the M----N transition, *J Biol Chem 266*, 11063-11067.
311. Pfefferle, J. M., Maeda, A., Sasaki, J., and Yoshizawa, T. (1991) Fourier transform infrared study of the N intermediate of bacteriorhodopsin, *Date of Input: 18/07/2007 30*, 6548-6556.

312. Braiman, M. S., Bousche, O., and Rothschild, K. J. (1991) Protein dynamics in the bacteriorhodopsin photocycle: submillisecond Fourier transform infrared spectra of the L, M, and N photointermediates, *Date of Input: 23/04/2007 88*, 2388-2392.
313. Cao, Y., Varo, G., Klinger, A. L., Czajkowsky, D. M., Braiman, M. S., Needleman, R., and Lanyi, J. K. (1993) Proton transfer from Asp-96 to the bacteriorhodopsin Schiff base is caused by a decrease of the pKa of Asp-96 which follows a protein backbone conformational change, *Date of Input: 18/07/2007 32*, 1981-1990.
314. Smith, O., Pardoen, J. A., Mulder, P. J., Lutenburg, B. J., and Mathies, R. (1983) Chromophore Structure in Bacteriorhodopsin's O640 Photointermediate, *Date of Input: 10/04/2007 26*, 6141-6147.
315. Bousche, O., Sonar, S., Krebs, M. P., Khorana, H. G., and Rothschild, K. J. (1992) Time-resolved Fourier transform infrared spectroscopy of the bacteriorhodopsin mutant Tyr-185-->Phe: Asp-96 reprotonates during O formation; Asp-85 and Asp-212 deprotonate during O decay, *Photochem Photobiol JT - Photochemistry and photobiology 56*, 1085-1095.
316. Hebling, B., Souvignier, G., and Gerwert, K. (1993) A model-independent approach to assigning bacteriorhodopsin's intramolecular reactions to photocycle intermediates, *Biophys J 65*, 1929-1941.
317. Rath, P., Krebs, M. P., He, Y., Khorana, H. G., and Rothschild, K. J. (1993) Fourier transform Raman spectroscopy of the bacteriorhodopsin mutant Tyr-185-->Phe: formation of a stable O-like species during light adaptation and detection of its transient N-like photoproduct, *Date of Input: 18/07/2007 32*, 2272-2281.
318. Dunach, M., Marti, T., Khorana, H. G., and Rothschild, K. J. (1990) Uv-visible spectroscopy of bacteriorhodopsin mutants: substitution of Arg-82, Asp-85, Tyr-185, and Asp-212 results in abnormal light-dark adaptation, *Date of Input: 23/04/2007 87*, 9873-9877.
319. He, Y., Krebs, M. P., Fischer, W. B., Khorana, H. G., and Rothschild, K. J. (1993) FTIR difference spectroscopy of the bacteriorhodopsin mutant Tyr-185-->Phe: detection of a stable O-like species and characterization of its photocycle at low temperature, *Date of Input: 18/07/2007 32*, 2282-2290.
320. Fukuda, K., and Kouyama, T. (1992) Photoreaction of bacteriorhodopsin at high pH: origins of the slow decay component of M, *Date of Input: 18/07/2007 31*, 11740-11747.
321. Turner, G. J., Miercke, L. J., Thorgeirsson, T. E., Kliger, D. S., Betlach, M. C., and Stroud, R. M. (1993) Bacteriorhodopsin D85N: three spectroscopic species in equilibrium, *Date of Input: 18/07/2007 32*, 1332-1337.
322. Krebs, M. P., Mollaaghababa, R., and Khorana, H. G. (1993) Gene replacement in Halobacterium halobium and expression of bacteriorhodopsin mutants, *Proc Natl Acad Sci U S A 90*, 1987-1991.
323. Dollinger, Eisenstein, and Lin, L. (1986) Fourier transform Infrared Difference Spectroscopy of Bacteriorhodopsin and Its Photoproducts Regenerated with Deuterated Tyrosine, *Date of Input: 10/04/2007 25*, 6524-6533.
324. Rothschild, K. J., Roepe, P., Ahl, P. L., Earnest, T. N., Bogomolni, R. A., Das Gupta, S. K., Mulliken, C. M., and Herzfeld, J. (1986) Evidence for a tyrosine protonation change during the primary phototransition of bacteriorhodopsin at low temperature, *Proc Natl Acad Sci U S A 83*, 347-351.
325. Rothschild, K. J., He, Y. W., Sonar, S., Marti, T., and Khorana, H. G. (1992) Vibrational spectroscopy of bacteriorhodopsin mutants. Evidence that Thr-46 and Thr-89 form part of a transient network of hydrogen bonds, *J Biol Chem 267*, 1615-1622.
326. Butt, J., Fendler, K., Bamberg, E., Tittor, and Oesterhelt, D. (1989) Aspartic acids 96 and 85 play a central role in the function of bacteriorhodopsin as a proton pump, *The EMBO Journal 8*, 1657-1663.
327. Otto, H., Marti, T., Holz, M., Mogi, T., Stern, L. J., Engel, F., Khorana, H. G., and Heyn, M. P. (1990) Substitution of amino acids Asp-85, Asp-212, and Arg-82 in bacteriorhodopsin affects the proton release phase of the pump and the pK of the Schiff base, *Proc Natl Acad Sci U S A 87*, 1018-1022.
328. Holz, M., Drachev, L. A., Mogi, T., Otto, H., Kaulen, A. D., Heyn, M. P., Skulachev, V. P., and Khorana, H. G. (1989) Replacement of aspartic acid-96 by asparagine in bacteriorhodopsin slows both the decay of the M intermediate and the associated proton movement, *Proc Natl Acad Sci U S A 86*, 2167-2171.
329. Tittor, Soell, Oesterhelt, D., Butt, J., and Bamberg, E. (1989) A defective proton pump,point-mutated bacteriorhodopsin Asp96 -> Asn is fully reactivated by azide, *The EMBO Journal 8*, 3477-3482.
330. Nagle, F., and Mille. (1981) Molecular models of proton pumps, *Journal of Chemical Phys 74*, 1367-.
331. Merz, H., and Zundel. (1983) Proton-transfer equilibriums in phenol-carboxylate hydrogen bonds. Implications for the mechanism of light-induced proton activation in bacteriorhodopsin, *Phys Lett 95*, 529-.

332. Brzezinski, Zundel, and Kraemer. (1987) Proton polarizability caused by collective proton motion in intramolecular chains formed by two and three hydrogen bonds. Implications for charge conduction in bacteriorhodopsin, *Journal of Phys.Chem 91*, 3077-.
333. Olejnik, J., Brzezinski, and Zundel. (1992) A proton pathway with large proton polarizability and the proton pumping mechanism in bacteriorhodopsin - Fourier transform difference spectra of a photoproducts of bacteriorhodopsin and of it's pentademethyl analogue, *Journal of Molecular Structure 271*, 157-173.
334. Rothschild, K. J., Roepe, P., Ahl, P. L., Earnest, T. N., Bogomolni, R. A., SK, D. G., Mulliken, C. M., and Herzfeld, J. (1986) Evidence for a tyrosine protonation change during the primary phototransition of bacteriorhodopsin at low temperature, *Date of Input: 23/04/2007 83*, 347-351.
335. Metz, G., Siebert, F., and Engelhard, M. (1992) Asp85 is the only internal aspartic acid that gets protonated in the M intermediate and the purple-to-blue transition of bacteriorhodopsin. A solid-state 13C CP-MAS NMR investigation, *FEBS Lett 303*, 237-241.
336. He, Y. W., Krebs, P., erzfeld, Khorana, H. G., and Rothschild, K. J. (1991) Tyrosine protonation changes in bacteriorhodopsin, A fourier transform infrared study of BR 548 and its primary photo product, *Proc SPIE-Int.Soc. Opt.Eng 1575*, 109-.
337. Netto, M. M., Fodor, S. P., and Mathies, R. A. (1990) Ultraviolet resonance Raman spectroscopy of bacteriorhodopsin, *Photochem Photobiol 52*, 605-607.
338. Ames, J. B., Ros, M., Raap, J., Lugtenburg, J., and Mathies, R. A. (1992) Time-resolved ultraviolet resonance Raman studies of protein structure: application to bacteriorhodopsin, *Date of Input: 18/07/2007 31*, 5328-5334.
339. Metz, G., Siebert, F., and Engelhard, M. (1992) High-resolution solid state 13C NMR of bacteriorhodopsin: characterization of [4-13C]Asp resonances, *Date of Input: 18/07/2007 31*, 455-462.
340. Phatak, P., Frahmcke, J. S., Wanko, M., Hoffmann, M., Strodel, P., Smith, J. C., Suhai, S., Bondar, A. N., and Elstner, M. (2009) Long-distance proton transfer with a break in the bacteriorhodopsin active site, *J Am Chem Soc 131*, 7064-7078.
341. Morgan, J. E., Vakkasoglu, A. S., Lugtenburg, J., Gennis, R. B., and Maeda, A. (2008) Structural changes due to the deprotonation of the proton release group in the M-photointermediate of bacteriorhodopsin as revealed by time-resolved FTIR spectroscopy, *Date of Input: 18/07/2007 47*, 11598-11605.
342. Tanio, M., Tuzi, S., Yamaguchi, S., Kawaminami, R., Naito, A., Needleman, R., Lanyi, J. K., and Saito, H. (1999) Conformational changes of bacteriorhodopsin along the proton-conduction chain as studied with (13)C NMR of [3-(13)C]Ala-labeled protein: arg(82) may function as an information mediator, *Biophys J 77*, 1577-1584.
343. Kouyama, T., Nishikawa, T., Tokuhisa, T., and Okumura, H. (2004) Crystal structure of the L intermediate of bacteriorhodopsin: evidence for vertical translocation of a water molecule during the proton pumping cycle, *J Mol Biol 335*, 531-546.
344. Creuzet, F., McDermott, A., Gebhard, R., K, v. d. H., Spijker-Assink, M. B., Herzfeld, J., Lugtenburg, J., Levitt, M. H., and Griffin, R. G. (1991) Determination of membrane protein structure by rotational resonance NMR: bacteriorhodopsin, *Science JT - Science (New York, N.Y.) 251*, 783-786.
345. Peersen, B., Yoshimura, Hojo, Aimoto, S., and Smith, O. (1992) Rotational resonance NMR measurements of intranuclear distances in an alpha-helical peptide, *Journal of American Chemical Society 114*, 4332-.
346. Subramaniam, S., Hirai, T., and Henderson, R. (2002) From structure to mechanism: electron crystallographic studies of bacteriorhodopsin, *Philos Transact A Math Phys Eng Sci 360*, 859-874.
347. Luecke, H., and Lanyi, J. K. (2003) Structural clues to the mechanism of ion pumping in bacteriorhodopsin, *Adv Protein Chem 63*, 111-130.
348. Xu, D., Sheves, M., and Schulten, K. (1995) Molecular dynamics study of the M412 intermediate of bacteriorhodopsin, *Biophys J 69*, 2745-2760.
349. Draheim, E., and Cassim, J. Y. (1985) Large scale global structure changes of the purple membrane during the photocycle, *Biophysical Journal 47*, 497-.
350. Dencher, N. A., Dresselhaus, D., Zaccai, G., and Buldt, G. (1989) Structural changes in bacteriorhodopsin during proton translocation revealed by neutron diffraction, *Proc Natl Acad Sci U S A 86*, 7876-7879.
351. Subramaniam, S., Gerstein, M., Oesterhelt, D., and Henderson, R. (1993) Electron diffraction analysis of structural changes in the photocycle of bacteriorhodopsin, *EMBO J 12*, 1-8.

352. Han, B. G., Vonck, J., and Glaeser, R. M. (1994) The bacteriorhodopsin photocycle: direct structural study of two substrates of the M-intermediate, *Biophys J 67*, 1179-1186.
353. Subramaniam, S., Lindahl, M., Bullough, P., Faruqi, A. R., Tittor, J., Oesterhelt, D., Brown, L., Lanyi, J., and Henderson, R. (1999) Protein conformational changes in the bacteriorhodopsin photocycle, *J Mol Biol 287*, 145-161.
354. Subramaniam, S., and Henderson, R. (2000) Crystallographic analysis of protein conformational changes in the bacteriorhodopsin photocycle, *Biochim Biophys Acta 1460*, 157-165.
355. Thorgeirsson, T. E., Xiao, W., Brown, L. S., Needleman, R., Lanyi, J. K., and Shin, Y. K. (1997) Transient channel-opening in bacteriorhodopsin: an EPR study, *J Mol Biol 273*, 951-957.
356. Tittor, J., Paula, S., Subramaniam, S., Heberle, J., Henderson, R., and Oesterhelt, D. (2002) Proton translocation by bacteriorhodopsin in the absence of substantial conformational changes, *J Mol Biol 319*, 555-565.
357. Ormos, P., Chu, K., and Mourant, J. (1992) Infrared study of the L, M, and N intermediates of bacteriorhodopsin using the photoreaction of M, *Date of Input: 18/07/2007 31*, 6933-6937.
358. Rothschild, K. J., Marti, T., Sonar, S., He, Y. W., Rath, P., Fischer, W., and Khorana, H. G. (1993) Asp96 deprotonation and transmembrane alpha-helical structural changes in bacteriorhodopsin, *J Biol Chem 268*, 27046-27052.
359. Vonck, J. (2000) Structure of the bacteriorhodopsin mutant F219L N intermediate revealed by electron crystallography, *EMBO J 19*, 2152-2160.
360. Rink, T., Pfeiffer, M., Oesterhelt, D., Gerwert, K., and Steinhoff, H. J. (2000) Unraveling photoexcited conformational changes of bacteriorhodopsin by time resolved electron paramagnetic resonance spectroscopy, *Biophys J 78*, 1519-1530.
361. Brown, L. S., Needleman, R., and Lanyi, J. K. (2002) Conformational change of the E-F interhelical loop in the M photointermediate of bacteriorhodopsin, *J Mol Biol 317*, 471-478.
362. Alexiev, U., Rimke, I., and Pohlmann, T. (2003) Elucidation of the nature of the conformational changes of the EF-interhelical loop in bacteriorhodopsin and of the helix VIII on the cytoplasmic surface of bovine rhodopsin: a time-resolved fluorescence depolarization study, *J Mol Biol 328*, 705-719.
363. Hendrickson, F. M., Burkard, F., and Glaeser, R. M. (1998) Structural characterization of the L-to-M transition of the bacteriorhodopsin photocycle, *Biophys J 75*, 1446-1454.
364. Sasaki, J., Shichida, Y., Lanyi, J. K., and Maeda, A. (1992) Protein changes associated with reprotonation of the Schiff base in the photocycle of Asp96-->Asn bacteriorhodopsin. The MN intermediate with unprotonated Schiff base but N-like protein structure, *J Biol Chem 267*, 20782-20786.
365. Nilsson, A., Rath, P., Olejnik, J., Coleman, M., and Rothschild, K. J. (1995) Protein conformational changes during the bacteriorhodopsin photocycle. A Fourier transform infrared/resonance Raman study of the alkaline form of the mutant Asp-85-->Asn, *J Biol Chem 270*, 29746-29751.
366. Rath, P., Marti, T., Sonar, S., Khorana, H. G., and Rothschild, K. J. (1993) Hydrogen bonding interactions with the Schiff base of bacteriorhodopsin. Resonance Raman spectroscopy of the mutants D85N and D85A, *J Biol Chem 268*, 17742-17749.
367. Deng, H., Huang, L., Callender, R., and Ebrey, T. (1994) Evidence for a bound water molecule next to the retinal Schiff base in bacteriorhodopsin and rhodopsin: a resonance Raman study of the Schiff base hydrogen/deuterium exchange, *Biophys J 66*, 1129-1136.
368. Maeda, A., Sasaki, J., Shichida, Y., and Yoshizawa, T. (1992) Water structural changes in the bacteriorhodopsin photocycle: analysis by Fourier transform infrared spectroscopy, *Date of Input: 18/07/2007 31*, 462-467.
369. Fischer, W. B., Sonar, S., Marti, T., Khorana, H. G., and Rothschild, K. J. (1994) Detection of a water molecule in the active-site of bacteriorhodopsin: hydrogen bonding changes during the primary photoreaction, *Date of Input: 18/07/2007 33*, 12757-12762.
370. Braiman, M. S., Ahl, P. L., and Rothschild, K. J. (1987) Millisecond Fourier-transform infrared difference spectra of bacteriorhodopsin's M412 photoproduct, *Date of Input: 23/04/2007 84*, 5221-5225.
371. Gerwert, K., Souvignier, G., and Hess, B. (1990) Simultaneous monitoring of light-induced changes in protein side-group protonation, chromophore isomerization, and backbone motion of bacteriorhodopsin by time-resolved Fourier-transform infrared spectroscopy, *Date of Input: 23/04/2007 87*, 9774-9778.

372. Weidlich, O., and Siebert, F. F. (1993) Time-resolved step- scan FT-IR investigation of the transition from KL to L in the bacteriorhodopsin photocycle: Identification of chromophore twists by assigning hydrogen out-of-plane (HOOP) bending vibrations, *Appl Spectrosc 47*, 1394-.
373. Diller, R., Iannone, M., Bogomolni, R., and Hochstrasser, R. M. (1991) Ultrafast infrared spectroscopy of bacteriorhodopsin, *Biophys J 60*, 286-289.
374. Mathies, R. A., Lin, S. W., Ames, J. B., and Pollard, W. T. (1991) From femtoseconds to biology: mechanism of bacteriorhodopsin's light-driven proton pump, *Annu Rev Biophys Biophys Chem 20*, 491-518.
375. Sonar, S., and Deshpande, A. (1997) Bacteriorhodopsin: A photoactive biomaterial for molecular electronics, *ISRAPS 7*, 11-18.
376. Hackett, N. R., Stern, L. J., Chao, B. H., Kronis, K. A., and Khorana, H. G. (1987) Structure-function studies on bacteriorhodopsin. V. Effects of amino acid substitutions in the putative helix F, *J Biol Chem JT - The Journal of biological chemistry 262*, 9277-9284.
377. Coleman, M., Sonar, S., Patel, N., and Rothschild, K. J. (1994) Transitioning Challenge to Commercial Application, in *Technologies and Applications conference for Emerging Technologies*, pp 326-341, Suny Institute of Technology.
378. Xu, Z. J., Moffett, D. B., Peters, T. R., Smith, L. D., Perry, B. P., Whitmer, J., Stokke, S. A., and Teintze, M. (1995) The role of the leader sequence coding region in expression and assembly of bacteriorhodopsin, *J Biol Chem 270*, 24858-24863.
379. Dunn, R. J., Hackett, N. R., McCoy, J. M., Chao, B. H., Kimura, K., and Khorana, H. G. (1987) Structure-function studies on bacteriorhodopsin. I. Expression of the bacterio-opsin gene in Escherichia coli, *J Biol Chem JT - The Journal of biological chemistry 262*, 9246-9254.
380. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1982) *Molecular Cloning: A Laboratory Manual*, Vol. 1, Cold Spring Harbor, N.Y.
381. Kagawa, Y., and Racker, E. (1971) Partial resolution of the enzymes catalyzing oxidative phosphorylation. XXV. Reconstitution of vesicles catalysing 32Pi- ATP exchange, *J. Biol. Chem. 246*, 10.
382. Braiman, M. S., Stern, L. J., Chao, B. H., and Khorana, H. G. (1987) Structure-function studies on bacteriorhodopsin. IV. Purification and renaturation of bacterio-opsin polypeptide expressed in Escherichia coli, *J Biol Chem JT - The Journal of biological chemistry 262*, 9271-9276.
383. Huang, K. S., Bayley, H., Liao, M. J., London, E., and Khorana, H. G. (1981) Refolding of an integral membrane protein. Denaturation, renaturation, and reconstitution of intact bacteriorhodopsin and two proteolytic fragments, *J Biol Chem JT - The Journal of biological chemistry 256*, 3802-3809.
384. London, E., and Khorana, H. G. (1982) Denaturation and renaturation of bacteriorhodopsin in detergents and lipid-detergent mixtures, *J Biol Chem 257*, 7003-7011.
385. Liao, M. J., London, E., and Khorana, H. G. (1983) Regeneration of the native bacteriorhodopsin structure from two chymotryptic fragments, *J Biol Chem JT - The Journal of biological chemistry 258*, 9949-9955.
386. Racker, E., and Stoeckenius, W. (1974) Reconstitution of purple membrane vesicles catalyzing light-driven proton uptake and adenosine triphosphate formation, *J Biol Chem JT - The Journal of biological chemistry 249*, 662-663.
387. Malone, R. W., Felgner, P. L., and Verma, I. M. (1989) Cationic liposome-mediated RNA transfection, *Proc Natl Acad Sci U S A 86*, 6077-6081.
388. Seehra, J. S., and Khorana, H. G. (1984) Bacteriorhodopsin precursor. Characterization and its integration into the purple membrane, *J Biol Chem JT - The Journal of biological chemistry 259*, 4187-4193.
389. Hartmann, R., Sickinger, H. D., and Oesterhelt, D. (1980) Anaerobic growth of halobacteria, *Proc Natl Acad Sci U S A 77*, 3821-3825.
390. Oesterhelt, D., and Stoeckenius, W. (1974) Isolation of the cell membrane of Halobacterium halobium and its fractionation into red and purple membrane, *Methods Enzymol 31*, 667-678.
391. Soll, D., RajBhandary, U.L. (1995) tRNA: structure, biosynthesis and function.
392. Saks, M. E., Sampson, J. R., and Abelson, J. N. (1994) The transfer RNA identity problem: a search for rules, *Science JT - Science (New York, N.Y.) 263*, 191-197.
393. Stadtman, T. C. (1996) Selenocysteine, *Annu Rev Biochem 65*, 83-100.
394. Wilting, R., Schorling, S., Persson, B. C., and Bock, A. (1997) Selenoprotein synthesis in archaea: identification of an mRNA element of Methanococcus jannaschii probably directing selenocysteine insertion, *J Mol Biol 266*, 637-641.

395. Curnow, A. W., Hong, K. W., Yuan, R., and Soll, D. (1997) tRNA-dependent amino acid transformations, *Nucleic Acids Symp Ser*, 2-4.
396. Normanly, J., and Abelson, J. (1989) tRNA identity, *Annu Rev Biochem 58*, 1029-1049.
397. Schulman, L. H. (1991) Recognition of tRNAs by tRNA-synthetases, *Prog Nucleic Acid Res Mol Biol 41*, 23-87.
398. McClain, W. H. (1993) Rules that govern tRNA identity in protein synthesis, *J Mol Biol JT - Journal of molecular biology 234*, 257-280.
399. Giege, R., Sissler, M., and Florentz, C. (1998) Universal rules and idiosyncratic features in tRNA identity, *Nucleic Acids Res 26*, 5017-5035.
400. Muramatsu, T., Nishikawa, K., Nemoto, F., Kuchino, Y., Nishimura, S., Miyazawa, T., and Yokoyama, S. (1988) Codon and amino-acid specificities of a transfer RNA are both converted by a single post-transcriptional modification, *Date of Input: 18/07/2007 336*, 179-181.
401. Perret, V., Garcia, A., Grosjean, H., Ebel, J. P., Florentz, C., and Giege, R. (1990) Relaxation of a transfer RNA specificity by removal of modified nucleotides, *Date of Input: 18/07/2007 344*, 787-789.
402. Triques, K., Coste, J., Perret, J. L., Segarra, C., Mpoudi, E., Reynes, J., Delaporte, E., Butcher, A., Dreyer, K., Herman, S., Spadoro, J., and Peeters, M. (1999) Efficiencies of four versions of the AMPLICOR HIV-1 MONITOR test for quantification of different subtypes of human immunodeficiency virus type 1, *J Clin Microbiol JT - Journal of clinical microbiology 37*, 110-116.
403. Baldwin, A. N., and Berg, P. (1966) Transfer ribonucleic acid-induced hydrolysis of valyladenylate bound to isoleucyl ribonucleic acid synthetase, *J Biol Chem 241*, 839-845.
404. Prather, N. E., Murgola, E. J., and Mims, B. H. (1984) Nucleotide substitution in the amino acid acceptor stem of lysine transfer RNA causes missense suppression, *J Mol Biol 172*, 177-184.
405. Tsai, F., and Curran, J. F. (1998) tRNA(2Gln) mutants that translate the CGA arginine codon as glutamine in Escherichia coli, *RNA 4*, 1514-1522.
406. Ibba, M., Becker, H. D., Stathopoulos, C., Tumbula, D. L., and Soll, D. (2000) The adaptor hypothesis revisited, *Trends Biochem Sci 25*, 311-316.
407. Ibba, M., and Soll, D. (2000) Aminoacyl-tRNA synthesis, *Annu Rev Biochem 69*, 617-650.
408. Liu, H., Peterson, R., Kessler, J., and Musier-Forsyth, K. (1995) Molecular recognition of tRNA(Pro) by Escherichia coli proline tRNA synthetase in vitro, *Nucleic Acids Res 23*, 165-169.
409. Komatsoulis, G. A., and Abelson, J. (1993) Recognition of tRNA(Cys) by Escherichia coli cysteinyl-tRNA synthetase, *Date of Input: 10/04/2007 32*, 7435-7444.
410. Daniel, V., Sarid, S., Beckmann, J. S., and Littauer, U. Z. (1970) In vitro transcription of a transfer RNA gene, *Proc Natl Acad Sci U S A 66*, 1260-1266.
411. Illangasekare, M., Sanchez, G., Nickles, T., and Yarus, M. (1995) Aminoacyl-RNA synthesis catalyzed by an RNA, *Science JT - Science (New York, N.Y.) 267*, 643-647.
412. Oesterhelt, D. (1998) The structure and mechanism of the family of retinal proteins from halophilic archaea, *Curr Opin Struct Biol 8*, 489-500.
413. Lanyi, J. K. (1997) Mechanism of ion transport across membranes. Bacteriorhodopsin as a prototype for proton pumps, *J Biol Chem JT - The Journal of biological chemistry 272*, 31209-31212.
414. Spudich, J. L. (1998) Variations on a molecular switch: transport and sensory signalling by archaeal rhodopsins, *Mol Microbiol 28*, 1051-1058.
415. Lanyi, J. K., Duschl, A., Varo, G., and Zimanyi, L. (1990) Anion binding to the chloride pump, halorhodopsin, and its implications for the transport mechanism, *FEBS Lett 265*, 1-6.
416. DasSarma, S., RajBhandary, U. L., and Khorana, H. G. (1983) High-frequency spontaneous mutation in the bacterio-opsin gene in Halobacterium halobium is mediated by transposable elements, *Proc Natl Acad Sci U S A 80*, 2201-2205.
417. Sumper, M., and Herrmann, G. (1976) Biosynthesis of purple membrane: control of retinal synthesis by bacterio-opsin, *FEBS Lett 71*, 333-336.
418. Oesterhelt, D., and Stoeckenius, W. (1973) Functions of a new photoreceptor membrane, *Proc Natl Acad Sci U S A 70*, 2853-2857.
419. Arnold, J., Dreiessen, M., and Konings, W. N. (1993) *Methods Enzymol. 88*.
420. van Dijck, P. W. M., van Dam, K. *Methods Enzymol. 88*, 17-.
421. Popot, J. L., Gerchman, S. E., and Engelman, D. M. (1987) Refolding of bacteriorhodopsin in lipid bilayers. A thermodynamically controlled two-stage process, *J Mol Biol JT - Journal of molecular biology 198*, 655-676.
422. Struck, D. K., Hoekstra, D., and Pagano, R. E. (1981) Use of resonance energy transfer to monitor membrane fusion, *Date of Input: 18/07/2007 20*, 4093-4099.

423. Kushwaha, S. C., Kates, M., and Weber, H. J. (1980) Exclusive formation of all-trans-phytoene by a colorless mutant of Halobacterium halobium, *Can J Microbiol 26*, 1011-1014.
424. LOWRY, O. H., ROSEBROUGH, N. J., FARR, A. L., and RANDALL, R. J. (1951) Protein measurement with the Folin phenol reagent, *J Biol Chem 193*, 265-275.
425. Danon, A., Brith-Lindner, M., and Caplan, S. R. (1977) Biogenesis of the purple membrane of Halobacterium halobium, *Biophys Struct Mech 3*, 1-17.
426. Cline, S. W., and Doolittle, W. F. (1987) Efficient transfection of the archaebacterium Halobacterium halobium, *J Bacteriol 169*, 1341-1344.
427. Mescher, M. F., and Strominger, J. L. (1976) Structural (shape-maintaining) role of the cell surface glycoprotein of Halobacterium salinarium, *Proc Natl Acad Sci U S A 73*, 2687-2691.
428. Jarrell, K. F. S. G. D. (1984) *Current Microbiology 10*, 147-152.
429. Betlach, M., Pfeifer, F., Friedman, J., and Boyer, H. W. (1983) Bacterio-opsin mutants of Halobacterium halobium, *Proc Natl Acad Sci U S A 80*, 1416-1420.
430. Arvinte, T., Hildenbrand, K., Wahl, P., and Nicolau, C. (1986) Lysozyme-induced fusion of liposomes with erythrocyte ghosts at acidic pH, *Date of Input: 23/04/2007 83*, 962-966.
431. Sumper, M., and Herrmann, G. (1976) Biogenesis of purple membrane: regulation of bacterio-opsin synthesis, *FEBS Lett 69*, 149-152.
432. Brown, A. E., Dolan, M. J., Michael, N. L., Zhou, S., Perfetto, S. P., Hawkes, C., Robb, M., Lane, J., Mayers, D., McNeil, J. G., Malone, J. D., Garner, R., Birx, D. L., and CN - RGP160 Phase IIA Vaccine Investigators. (2001) Clinical prognosis of patients with early-stage human immunodeficiency virus (HIV) disease: contribution of HIV-1 RNA and T lymphocyte subset quantitation, *Mil Med JT - Military medicine 166*, 571-576.
433. Fairclough, R. H., and Cantor, C. R. (1978) The use of singlet-singlet energy transfer to study macromolecular assemblies, *Methods Enzymol 48*, 347-379.
434. Stoeckenius, W., and Bogomolni, R. A. (1982) Bacteriorhodopsin and related pigments of halobacteria, *Annu Rev Biochem 51*, 587-616.
435. Rothschild, K. J., Sonar, S. (1995) *CRC Handbook of Organic Photochemistry and Photobiology*, CRC Press Inc., London.
436. Grigorieff, N., Ceska, T. A., Downing, K. H., Baldwin, J. M., and Henderson, R. (1996) Electron-crystallographic refinement of the structure of bacteriorhodopsin, *J Mol Biol JT - Journal of molecular biology 259*, 393-421.
437. Subramaniam, S., Lindahl, M., Bullough, P., Faruqi, A. R., Tittor, J., Oesterhelt, D., Brown, L., Lanyi, J., and Henderson, R. (1999) Protein conformational changes in the bacteriorhodopsin photocycle, *J Mol Biol JT - Journal of molecular biology 287*, 145-161.
438. Deshpande, A., and Sonar, S. (1999) Bacterioopsin-triggered retinal biosynthesis is inhibited by bacteriorhodopsin formation in Halobacterium salinarium, *J Biol Chem JT - The Journal of biological chemistry 274*, 23535-23540.
439. Sigrist, H., Wenger, R. H., Kislig, E., and Wuthrich, M. (1988) Refolding of bacteriorhodopsin. Protease V8 fragmentation and chromophore reconstitution from proteolytic V8 fragments, *Eur J Biochem 177*, 125-133.
440. Marti, T., Otto, H., Mogi, T., Rosselet, S. J., Heyn, M. P., and Khorana, H. G. (1991) Bacteriorhodopsin mutants containing single substitutions of serine or threonine residues are all active in proton translocation, *J Biol Chem JT - The Journal of biological chemistry 266*, 6919-6927.
441. Miller, and Oesterhelt, D. (1990) Kinetic optimization of bacteriorhodopsin by aspartic acid 96 as an internal proton donor, *Biochimica and Biophysica Acta 1020*, 57-64.
442. Gennis, R. B., and Cantor, C. R. (1972) Use of nonspecific dye labeling for singlet energy-transfer measurements in complex systems. A simple model, *Date of Input: 10/04/2007 11*, 2509-2517.
443. Cladera, J., Torres, J., and Padros, E. (1996) Analysis of conformational changes in bacteriorhodopsin upon retinal removal, *Biophys J 70*, 2882-2887.
444. Moltke, S., Wallat, I., Sakai, N., Nakanishi, K., Brown, M. F., and Heyn, M. P. (1999) The angles between the C(1)-, C(5)-, and C(9)-methyl bonds of the retinylidene chromophore and the membrane normal increase in the M intermediate of bacteriorhodopsin: direct determination with solid-state (2)H NMR, *Date of Input: 10/04/2007 38*, 11762-11772.
445. Grzesiek, S., and Dencher, N. A. (1986) Dependency of delta pH-relaxation across vesicular membranes on the buffering power of bulk solutions and lipids, *Biophys J 50*, 265-276.

446. Koch, M. H., Dencher, N. A., Oesterhelt, D., Plohn, H. J., Rapp, G., and Buldt, G. (1991) Time-resolved X-ray diffraction study of structural changes associated with the photocycle of bacteriorhodopsin, *EMBO J 10*, 521-526.
447. Kamikubo, H., Oka, T., Imamoto, Y., Tokunaga, F., Lanyi, J. K., and Kataoka, M. (1997) The last phase of the reprotonation switch in bacteriorhodopsin: the transition between the M-type and the N-type protein conformation depends on hydration, *Date of Input: 10/04/2007 36*, 12282-12287.
448. Subramaniam, S., Faruqi, A. R., Oesterhelt, D., and Henderson, R. (1997) Electron diffraction studies of light-induced conformational changes in the Leu-93 --> Ala bacteriorhodopsin mutant, *Proc Natl Acad Sci U S A 94*, 1767-1772.
449. Muller, D. J., Buldt, G., and Engel, A. (1995) Force-induced conformational change of bacteriorhodopsin, *J Mol Biol JT - Journal of molecular biology 249*, 239-243.
450. Kluge, T., Olejnik, J., Smilowitz, L., and Rothschild, K. J. (1998) Conformational changes in the core structure of bacteriorhodopsin, *Date of Input: 10/04/2007 37*, 10279-10285.