

# **Molecular mechanism of selenocysteine incorporation in bacterial translation**

Dissertation

for the award of the degree

“Doctor rerum naturalium” (Dr. rer. nat.)

Division of Mathematics and Natural Sciences

of the Georg-August-Universität Göttingen

Submitted by

Suresh Babu Kotini

from Rajampalem, India

Göttingen 2011

Members of the Thesis Committee:

Prof. Dr. Marina V. Rodnina, (1<sup>st</sup> reviewer)  
Department of Physical Biochemistry,  
Max Planck Institute for Biophysical Chemistry.

Prof. Dr. Holger Stark, (2<sup>nd</sup> reviewer)  
3D-Cryo Electron Microscopy,  
Max Planck Institute for Biophysical Chemistry.

Prof. Dr. Ralf Ficner,  
Department of Molecular Structural Biology,  
Göttingen Center for Molecular Biosciences.

Date of the oral examination: June 27<sup>th</sup>, 2011

**Affidavit**

I hereby declare that I prepared the PhD thesis “Molecular mechanism of selenocysteine incorporation in bacterial translation” on my own and with no other sources and aids than quoted.

A handwritten signature in blue ink. It starts with 'K.' followed by 'Suresh' in a cursive script, and ends with 'Babu' which is underlined.

Suresh Babu Kotini

Göttingen, April 26<sup>th</sup>, 2011

<b>1 Introduction .....</b>	<b>1</b>
<b>1.1 Co-translational insertion of selenocysteine (Sec) into bacterial proteins .....</b>	<b>1</b>
<b>1.2 Sec insertion into bacterial proteins .....</b>	<b>3</b>
1.2.1 <i>E. coli</i> tRNA <sup>Sec</sup> and their identity elements .....	6
<b>1.3 Sec insertion into archaeal proteins .....</b>	<b>8</b>
1.3.1 Archaeal tRNA <sup>Sec</sup> identity elements and recognition by <i>O</i> -phosphoseryl-tRNA kinase..	11
1.3.2 Sep-tRNA:Sec-tRNA (SepSecS).....	12
<b>1.4 Sec insertion into eukaryotic proteins .....</b>	<b>13</b>
1.4.1 Eukaryal tRNA <sup>Sec</sup> identity elements and recognition by <i>O</i> -phosphoseryl-tRNA kinase..	14
1.4.2 Interactions between eEFsec, SBP2, ribosomal protein L30 and SECIS- element .....	15
<b>1.5 The structure and identity elements of SECIS-RNA from organisms from three domains of life .....</b>	<b>16</b>
1.5.1 Bacterial SECIS-element .....	16
1.5.2 Archaeal and eukaryal SECIS-element .....	17
<b>1.6 Efficiency of Sec insertion into bacterial, archaeal and eukaryotic proteins .....</b>	<b>18</b>
1.6.1 Efficiency of Sec insertion into bacterial proteins.....	18
1.6.2 Efficiency of Sec insertion into archaeal proteins.....	19
1.6.3 Efficiency of Sec insertion into eukaryotic proteins.....	20
<b>1.7 Comparisons of EF-Tu and SelB.....</b>	<b>21</b>
1.7.1 Sequence and structural similarity.....	21
1.7.2 GDP, GTP and aa-tRNA binding properties .....	23
1.7.3 EF-Tu·EF-Ts complex.....	24
<b>1.8 Comparisons of SelB from prokaryotic, archaeal and eukaryotic organisms .....</b>	<b>25</b>
1.8.1 N-terminal domains (1-3).....	26

1.8.2 The SECIS-element binding, C-terminal domain (4).....	27
1.8.2.1 Prokaryotic, SelB domain (4).....	27
1.8.2.2 SECIS-RNA recognition by wing-helix motif .....	28
1.8.2.3 Mammals, SelB domain (4) .....	29
<b>1.9 Evolutionary relationship of eIF-2<math>\gamma</math> to bacterial SelB .....</b>	<b>31</b>
<b>1.10 Evolutionary conservation of Sec insertion machinery .....</b>	<b>32</b>
<b>1.11 Aim and scope of the work .....</b>	<b>35</b>
<b>2 Materials and Methods .....</b>	<b>36</b>
<b>2.1 Plasmid construction for <i>in vivo</i> experiments.....</b>	<b>36</b>
2.1.1 Cloning of <i>E. coli fdhF</i> DNA gene (Thr130-Asp179) into p94 plasmid .....	36
2.1.2 Stem-loop deletion mutagenesis .....	37
2.1.3 Site-directed mutagenesis of <i>fdhF</i> UGA <sub>140</sub> +SL to UUC <sub>140</sub> Quick-change PCR.....	37
2.1.4 Table: 7 Oligodeoxyribonucleotides used .....	38
2.1.5 Table: 8 Strain and plasmids of <i>in vivo</i> and <i>in vitro</i> experiments .....	39
<b>2.2 Expression and induction of fusion protein.....</b>	<b>40</b>
2.2.1 Induction of firefly- <i>fdhF</i> -renilla fusion protein expression at initial-log phase .....	40
2.2.2 Induction of firefly- <i>fdhF</i> -renilla fusion protein expression in the presence of selenium concentration dependence at initial-log phase, mid-log phase and stationary phase .....	40
2.2.3 Induction of firefly- <i>fdhF</i> -renilla fusion protein expression at various phases of cell growth (Initial-log, mid-log and stationary phase) .....	41
<b>2.3 Immunoblot with <math>\alpha</math>-firefly and <math>\alpha</math>-renilla antibodies .....</b>	<b>42</b>
<b>2.4 Dual reporter assay .....</b>	<b>43</b>
2.4.1 Luciferase activity measurements.....	43
2.4.2 Calculation of recoding efficiency .....	43

<b>2.5 Plasmid construction and <i>in vitro</i> experiments</b> .....	44
2.5.1 Cloning of <i>E. coli</i> full-length <i>fdhF</i> DNA gene into pUC19 vector .....	44
2.5.2 Cloning of <i>fdhF</i> DNA gene (Gly129-Ser180) into pTZ18R <i>in vitro</i> transcription vector ..	45
2.5.3 PCR amplification of <i>fdhF</i> DNA gene (Gly129-Ala156) for <i>in vitro</i> transcription .....	45
2.5.4 Cloning of <i>fdhF</i> DNA gene (Phe129-Ala156) into pBluescript II (SK+).....	45
2.5.5 Site-directed mutagenesis of <i>fdhF</i> UGA <sub>140</sub> +SL to UUC <sub>140</sub> +SL by Quick-change PCR .....	46
2.5.6 <i>In vitro</i> transcription and purification of model <i>fdhF</i> mRNAs ( <i>fdhF</i> SD) .....	47
2.5.7 <i>In vitro</i> transcription and purification of model <i>fdhF</i> mRNAs (strong SD).....	49
2.5.8 Preparation and purification of initiation complexes .....	51
2.5.9 Preparation and purification of aminoacyl-tRNAs .....	51
2.5.10 Expression and purification of wild type elongation factor SelB .....	52
2.5.11 Preparation of SelB·GTP·Sec-tRNA <sup>Sec</sup> .....	56
2.5.12 Preparation of EF-Tu·GTP·aa-tRNA .....	56
2.5.13 <i>In vitro</i> translation .....	57
2.5.14 HPLC separation of dipeptide (fMetGly) .....	57
2.5.15 HPLC separation of dipeptide (fMetPhe) .....	58
2.5.16 HPLC separation of tripeptide (fMetPheThr) and polypeptide up to 1 <sup>st</sup> Val .....	58
2.5.17 HPLC separation of polypeptide up to 2 <sup>nd</sup> Val .....	59
<b>3 Results</b> .....	60
<b>3.1 Efficiency of Sec incorporation in proteins <i>in vivo</i></b> .....	60
3.1.1 Efficiency of UGA recoding into selenocysteine by Sec-tRNA <sup>Sec</sup> .....	60
3.1.2 Selenium levels determine the efficiency of UGA recoding by Sec-tRNA <sup>Sec</sup> into selenocysteine .....	63
3.1.3 UGA recoding efficiency at various phases of cell growth.....	65

3.1.4 Determination of protein production levels at various phases of cell growth .....	67
3.1.5 Selenocysteine insertion efficiency does not depend on high selenium levels under normal physiological growth conditions of <i>E. coli</i> .....	68
<b>3.2 Design of the <i>in vitro</i> Sec incorporation system.....</b>	<b>70</b>
3.2.1 Design of model <i>fdhF</i> mRNA constructs .....	70
3.2.2 Optimization of translation of the mRNA sequence preceding the Sec codon .....	75
3.2.3 Incorporation of Sec into peptides <i>in vitro</i> .....	81
<b>4 Discussion.....</b>	<b>87</b>
<b>5 Summary .....</b>	<b>94</b>
<b>6 References.....</b>	<b>95</b>
<b>7 Abbreviations.....</b>	<b>107</b>
<b>8 Acknowledgements .....</b>	<b>108</b>
<b>9 Curriculum vitae.....</b>	<b>109</b>

## 1 Introduction

### 1.1 Co-translational insertion of selenocysteine (Sec) into bacterial proteins

Selenocysteine (Sec) has been recognized as the 21<sup>st</sup> genetically encoded amino acid that is incorporated into peptide during protein synthesis on the ribosome. Sec has a structure similar to that of cysteine (Cys), but with an atom of selenium taking the place of the sulfur, forming a selenol group. The selenol group in Sec is more nucleophilic than the thiol group in Cys, with pKa values of ~5.2 and 8.3, respectively (Huber and Criddle, 1967) (Table 1).

**Table 1:** Properties of the amino acids selenocysteine and cysteine.

	Selenocysteine	Cysteine
Codon	UGA*	UGU, UGC
Abbreviations	Sec, U	Cys, C
pKa	5.2	8.3
Structure	$  \begin{array}{c}  \text{H} \\    \\  {}^+\text{H}_3\text{N} - \text{C} - \text{COO}^- \\    \\  \text{CH}_2 \\    \\  \text{SeH}  \end{array}  $	$  \begin{array}{c}  \text{H} \\    \\  {}^+\text{H}_3\text{N} - \text{C} - \text{COO}^- \\    \\  \text{CH}_2 \\    \\  \text{SH}  \end{array}  $

\*In some organisms, e.g. *Mycoplasma*, and in mitochondria UGA codons designate tryptophan (Barrell et al., 1979; Christiansen et al., 1997).

This is probably the reason why Sec residue is primarily found at the active centre of enzymes that catalyze oxidation-reduction reactions (Chambers et al., 1986; Zinoni et al., 1986). Replacement of Sec with Cys resulted in ~1000-fold decreased activity of some enzymes (Axley et al., 1991). Proteins that contain one or more Sec residues are called selenoproteins (Kryukov and Gladyshev, 2004). Sec biosynthesis and insertion machineries are widely found in bacteria, archaea and eukaryotes. (Table 2) (Bock et al., 1991; Kryukov and Gladyshev, 2004; Low and Berry, 1996; Rother et al., 2001; Stadtman, 1996).



**Table 2:** Selenoproteins from three domains of life.

---

**Organisms and selenoproteins**

---

**Bacteria**

1. Formate dehydrogenases (*fdhF*, *fdnG*, and *fdoG*)
2. Hydrogenases
3. Selenoprotein A
4. Selenoprotein B
5. Proline reductase
6. Selenophosphate synthase (*selD*)
7. Peroxiredoxin

**Archaea**

1. Formate dehydrogenases (*fdhA*, *fdnG*, and *fdoG*)
2. Hydrogenases
3. Heterodisulfide reductase
4. Formylmethanofuran dehydrogenase
5. Selenophosphate synthase (*selD*)

**Eukaryotes****Group I (GPx group)** (protection against oxidative damage in the cell)

1. Cytosolic glutathione peroxidase
2. Gastrointestinal glutathione peroxidase
3. Plasma glutathione peroxidase
4. Phospholipid hydroperoxide glutathione peroxidase
5. Selenoprotein P

**Group II (TR group)** (catalyzing the reduction of thioredoxin)

12. Cytosolic thioredoxin reductase (TR1)
13. Thioredoxin reductase expressed intestis (TR2)
14. Mitochondrial thioredoxin reductase (TR3)

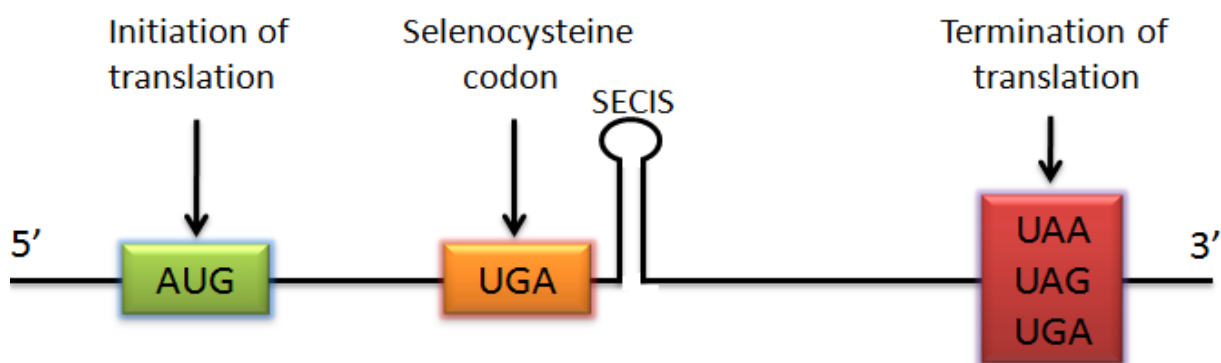
**Group III (other selenoproteins)**

16. Thyroid hormone deiodinase 1
  17. Thyroid hormone deiodinase 2
  18. Thyroid hormone deiodinase 3
  19. Selenoprotein R
  20. Selenoprotein N
  21. Selenophosphate synthase 2 (*selD*)
-

Unlike the twenty standard amino acids, Sec is not coded for by a sense triplet in the standard genetic code. Instead, it is encoded in a special way by a UGA codon, which normally serves as a signal for termination of protein synthesis. This was shown first for the *Escherichia coli* (*E. coli*) formate dehydrogenase-H (FDH-H) (Zinoni et al., 1986) and for the mammalian glutathione peroxidase gene (*GPx* gene) (Chambers et al., 1986).

## 1.2 Sec insertion into bacterial proteins

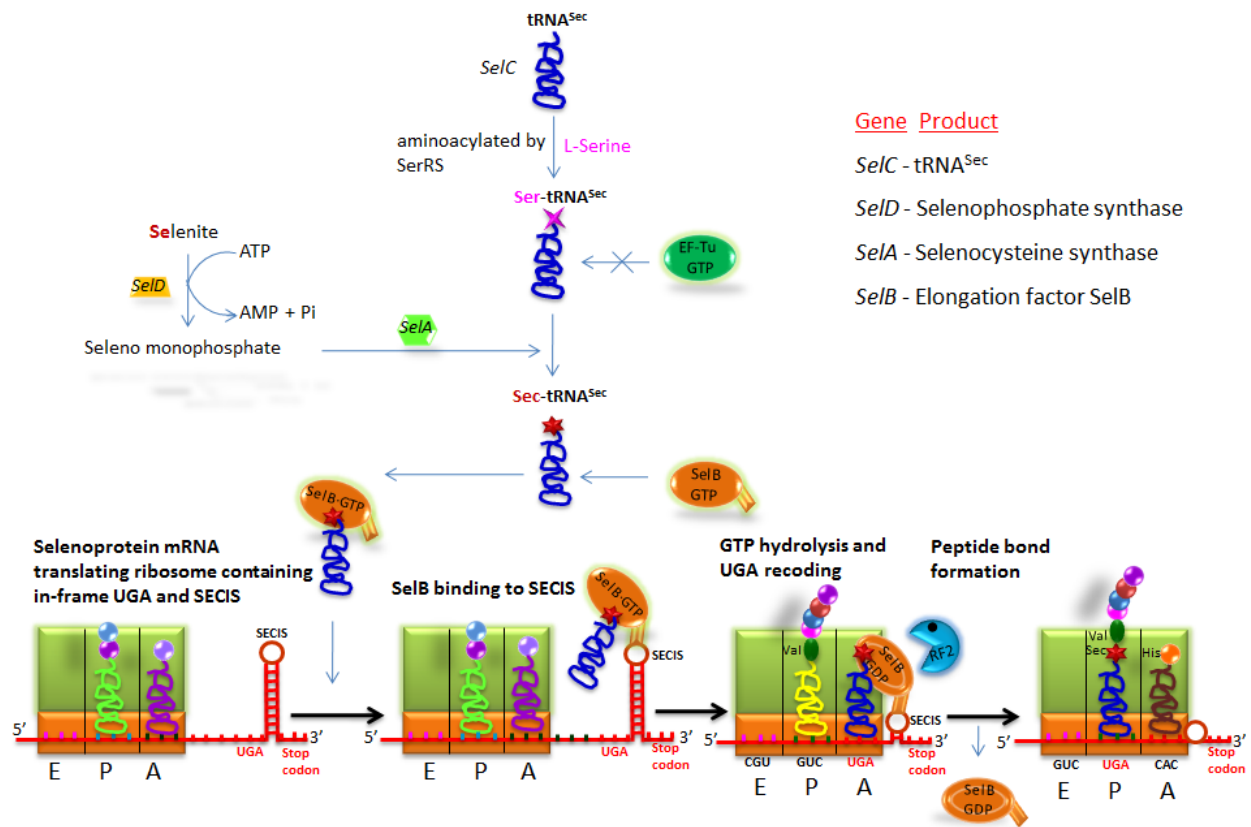
In bacteria, the biosynthesis and specific incorporation of Sec into proteins requires the function of two *cis* (a UGA codon, and the SECIS-element) and four *trans* elements (SelB, tRNA<sup>Sec</sup>, SelD & SelA) (Table 3). In bacteria, the SECIS-element is located immediately downstream the in-frame UGA codon (Zinoni et al., 1990) at which Sec is incorporated. The presence of the SECIS-element in the 3'-translated region also inhibits the termination of translation at the adjacent UGA codon (Baron and Bock, 1991) (Figure 1).



**Figure 1:** Bacterial selenoprotein mRNA contain in-frame UGA codon immediately followed by the SECIS-element.

The bacterial Sec incorporation system has been studied mainly using *E. coli* as the model organism. In all bacteria producing selenoproteins the Sec insertion is guided by a SECIS-element, however, the structure of the SECIS differs between gram-positive and gram-

negative bacteria and the SelB/tRNA<sup>Sec</sup> from *E. coli* is not compatible with the gram-positive *Eubacterium acidaminophilum* (Gursinsky et al., 2008). Unlike standard amino acids, no free Sec exists in the cell, but cells store selenium in the less reactive selenide form (H<sub>2</sub>Se). Selenophosphate synthase, SelD, encoded by the gene *selD*, plays an important role in the incorporation of selenide into the amino acid selenocysteine through a specific pathway (reviewed by Böck et al., 1991; Heider & Böck 1993; Böck & Sawers 1996). Selenophosphate synthase produces an activated form of selenium, selenomonophosphate, from ATP and selenide (Veres et al., 1992). In bacteria, Sec is synthesized on Sec-specific tRNA<sup>Sec</sup> through a multi-step process. tRNA<sup>Sec</sup> is first aminoacylated with serine by seryl-tRNA synthetase (SerRS) (Leinfelder et al., 1988) and Ser-tRNA<sup>Sec</sup> is directly converted to Sec-tRNA<sup>Sec</sup> by Sec synthase (SelA or SecS) (Forchhammer and Bock, 1991). In contrast, in eukarya and archaea, the hydroxyl group of seryl moiety is further phosphorylated by O-phosphoseryl-tRNA kinase (PSTK) (Carlson et al., 2004). The phosphate group is then converted to the selenol group by Sep-tRNA:Sec-tRNA synthase (SepSecS), to produce selenocystenyl-tRNA<sup>Sec</sup> (Sec-tRNA<sup>Sec</sup>) (Yuan et al., 2006; Xu et al., 2007).



**Figure 2:** Sec biosynthesis and insertion into bacterial proteins (Huttenhofer and Bock, 1998).

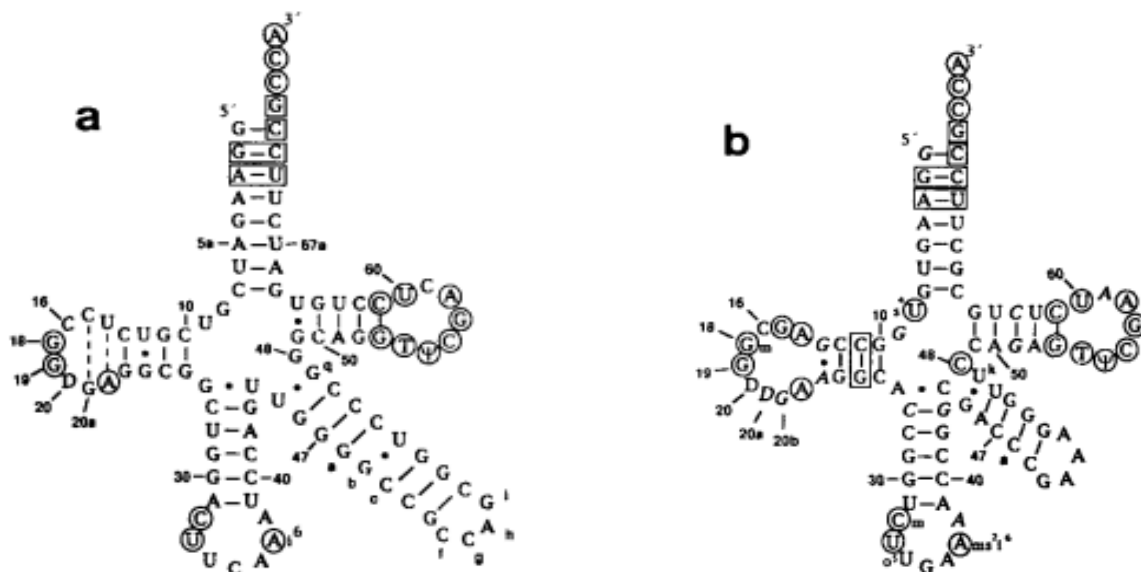
**Table 3:** Factors necessary for selenoprotein synthesis in *E. coli*.

<i>Cis</i> and <i>Trans</i> elements	Function
In-frame UGA codon	Specific site for Sec insertion
The SECIS-element	An mRNA secondary structure forming a hairpin loop that guides the Sec insertion
SelD	Selenophosphate synthase transforms selenide to monoselenophosphate
SelA	Selenocysteine synthase charges the selenium to make Sec from Ser on the tRNA <sup>Sec</sup>
SelC	Sec specific tRNA (tRNA <sup>Sec</sup> ), unique from other canonical tRNAs
SelB	Specialized elongation factor, specifically binds to Sec-tRNA <sup>Sec</sup>

### 1.2.1 *E. coli* tRNA<sup>Sec</sup> and its identity elements

In bacteria, the key player in the Sec incorporation machinery was originally identified as the product of gene *fdhC*, now designated as the *selC* gene, which codes for the Sec-specific tRNA<sup>Sec</sup> (Bock et al., 1991; Leinfelder et al., 1989). tRNA<sup>Sec</sup> has the anticodon ACU, complementary to the stop codon UGA (Leinfelder et al., 1988), and translates it in a selenoprotein mRNA to Sec, in response to a Sec-insertion sequence (SECIS) (Berry et al., 1993). The unusual sequence and structural properties of primary and secondary structure of tRNA<sup>Sec</sup>, which with 95 nucleotides is the largest tRNA in *E. coli*, differ from those of standard tRNAs in several respects, most notably in having an 8-base (bacteria) or 10-base (eukaryotes) pair acceptor stem, a long variable arm, and substitutions at several otherwise well-conserved base positions. A similarly long variable arm is also found in tRNA<sup>Ser</sup>, tRNA<sup>Leu</sup> and tRNA<sup>Tyr</sup>. There are several deviations from the consensus structure characteristic for canonical elongator tRNAs (e.g. tRNA<sup>Ser</sup>, Figure 3), namely a G instead of the canonical U at position 8, an A at position 14, a Y-R pair at the positions 10-25 and an R-Y base pair at positions 10-24. Enzymatic and chemical probing of the solution structure of tRNA<sup>Sec</sup> from *E. coli* in solution, compared with that of the canonical tRNA<sup>Ser</sup>, showed that these deviations, and the fact that the D stem of tRNA<sup>Sec</sup> is close to a six base pair helix minimizing the D loop to four nucleotides, restrict the types of tertiary interactions within the molecule (Baron et al., 1993; Schon et al., 1989) (Figure 3). Whereas the canonical G19-C56 interaction is still present, there are other interactions between C16 of the D loop and C59 of the T loop and

the canonical A21-(U8-A14) triple pair is substituted by a G8-(A21-U14) triple interaction.



**Figure 3:** Structures of (a) tRNA<sup>Sec</sup> and (b) tRNA<sup>Ser</sup> from *E. coli*.

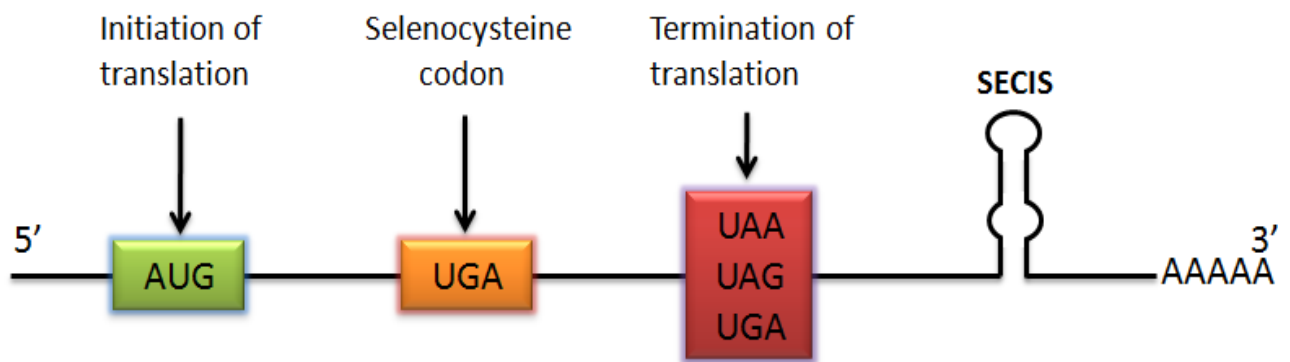
tRNA<sup>Sec</sup> is initially charged with L-serine by seryl-tRNA synthetase, but the resulting Ser-tRNA<sup>Sec</sup> is not used for translation because it is not recognized by the canonical translation elongation factor (EF-Tu in bacteria, eEF1A in eukaryotes). Rather, the tRNA-bound seryl residue is converted to a Sec residue by the pyridoxal phosphate-containing enzyme selenocysteine synthase (the *selA* gene product) using selenomonophosphate as the selenium donor substrate. The latter is synthesized from selenite and ATP by selenophosphate synthetase (the *selD* gene product). Finally, the resulting Sec-tRNA<sup>Sec</sup> binds to a specific translational elongation factor SelB (Forchhammer et al., 1989) which delivers the Sec-tRNA<sup>Sec</sup> in a targeted manner to the ribosomes translating mRNAs for selenoproteins (Figure 2). Sec is not the only amino acid to be biosynthesized on its tRNA, but cysteine can also be generated via Ser-tRNA<sup>Cys</sup> in certain methanogens (Sauerwald et al., 2005). In addition, Asn-tRNA<sup>Asn</sup> and Gln-tRNA<sup>Gln</sup> are generated from Asp-tRNA<sup>Asn</sup> (Feng et al., 2005b)

and Glu-tRNA<sup>Gln</sup> (Feng et al., 2005a) respectively, in many bacteria and essentially all archaea, and in some cases this is the sole route for asparagine biosynthesis (Feng et al., 2005b).

The proposed model of SelB action on the ribosome is largely deduced based on the wealth of knowledge on the mechanism of EF-Tu action on the ribosome. According to the proposed model, a quaternary complex of SelB·GTP·Sec-tRNA<sup>Sec</sup> and the SECIS-element of the mRNA is formed. Binding of Sec-tRNA<sup>Sec</sup> to SelB·GTP stabilizes the complex with the SECIS-element. During translation, the complex is translocated towards the in-frame UGA and the lower part (10-11 nucleotides) of the SECIS hairpin is expected to melt while the in-frame UGA codon is positioned in the A site, whereas the intact upper part with SelB bound to it through its domain 4 appears at the mRNA entrance of the ribosomal A site. The analogy to EF-Tu, the codon and anticodon interaction at the decoding center may induce the GTPase of SelB. After GTP hydrolysis, Sec-tRNA<sup>Sec</sup> is released from the factor, accommodates in the A site and Sec is incorporated into the nascent peptide, while efficiently competing the release factor 2 (RF2). Release of tRNA to the A site decreases the affinity of SelB·GDP for the mRNA SECIS-element and SelB-GDP dissociate from the ribosome and the SECIS-element to allow for the translation of the downstream mRNA sequence (Figure 2). After translation of the SECIS sequence, the RNA can refold and serve as a target for the formation of a new quaternary complex to assist the next oncoming ribosome in decoding UGA.

### **1.3 Sec insertion into archaeal proteins**

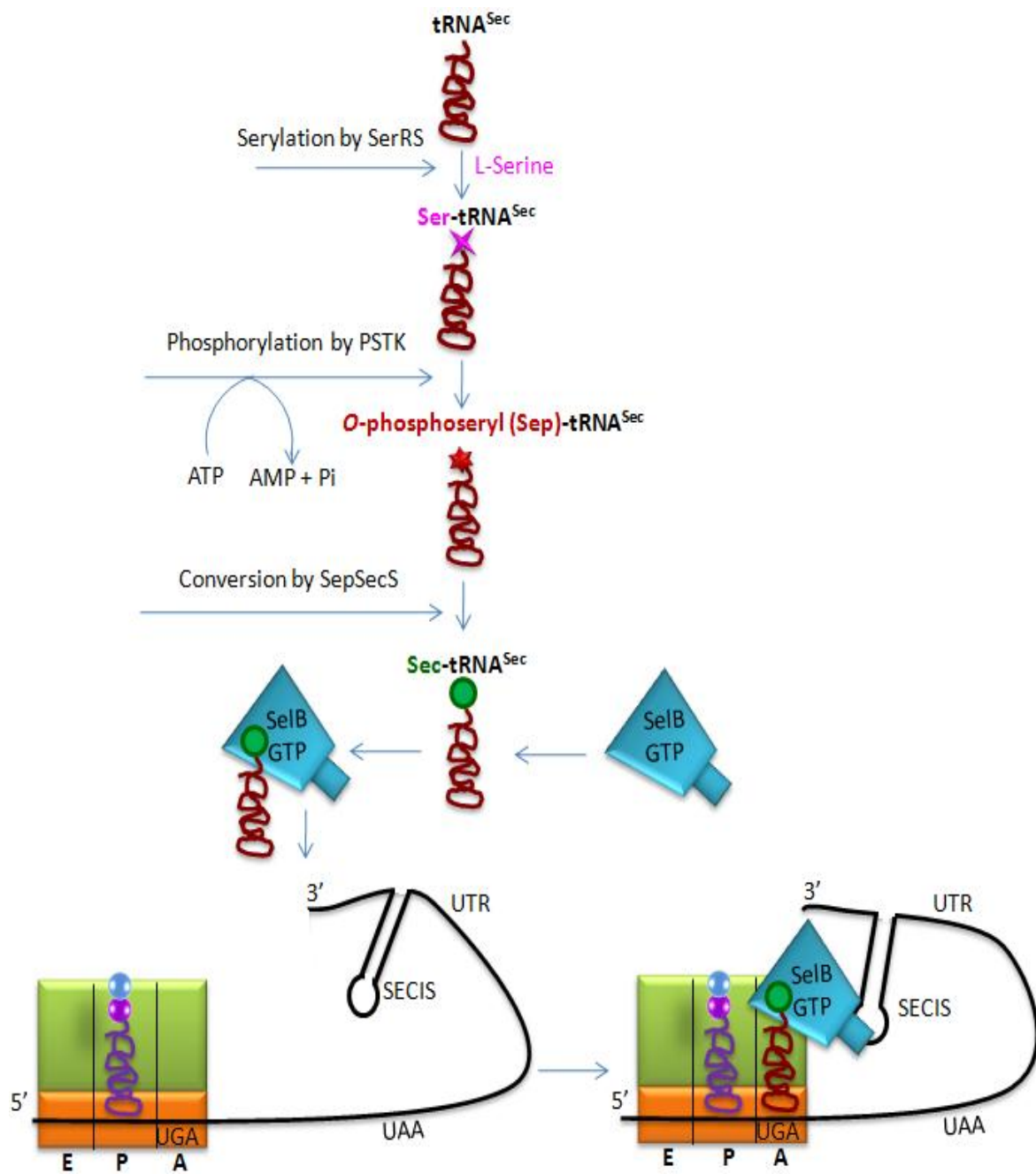
In archaeal selenoprotein mRNAs, the required signals are also stem-loop SECIS-elements, but they differ in sequence and structure from their prokaryotic counterparts and from each other (Rother et al., 2001) and they located at 3'-untranslated region (Figure 4).



**Figure 4:** Archaeal selenoprotein mRNA contains an in-frame UGA and SECIS-element in the 3'-untranslated region.

In all three domains of life  $\text{tRNA}^{\text{Sec}}$  is first charged with serine by seryl-tRNA synthetase producing  $\text{Ser-tRNA}^{\text{Sec}}$  as the initial precursor of  $\text{Sec-tRNA}^{\text{Sec}}$ . Bacteria convert this misacylated aminoacyl-tRNA species in a single step to  $\text{Sec-tRNA}^{\text{Sec}}$  by selenocysteine synthase (SelA). However, archaea and eukaryotes use two essential enzymes for this conversion; *O*-phosphoseryl- $\text{tRNA}^{\text{Sec}}$  kinase (PSTK) (Carlson et al., 2004) phosphorylates Ser bound to  $\text{tRNA}^{\text{Sec}}$  to generate *O*-phosphoseryl (Sep)- $\text{tRNA}^{\text{Sec}}$ , which is subsequently converted to Sec by Sep-tRNA:Sec-tRNA synthase (SepSecS) and make  $\text{Sec-tRNA}^{\text{Sec}}$  (Xu et al., 2007; Yuan et al., 2006) (Figure 5).

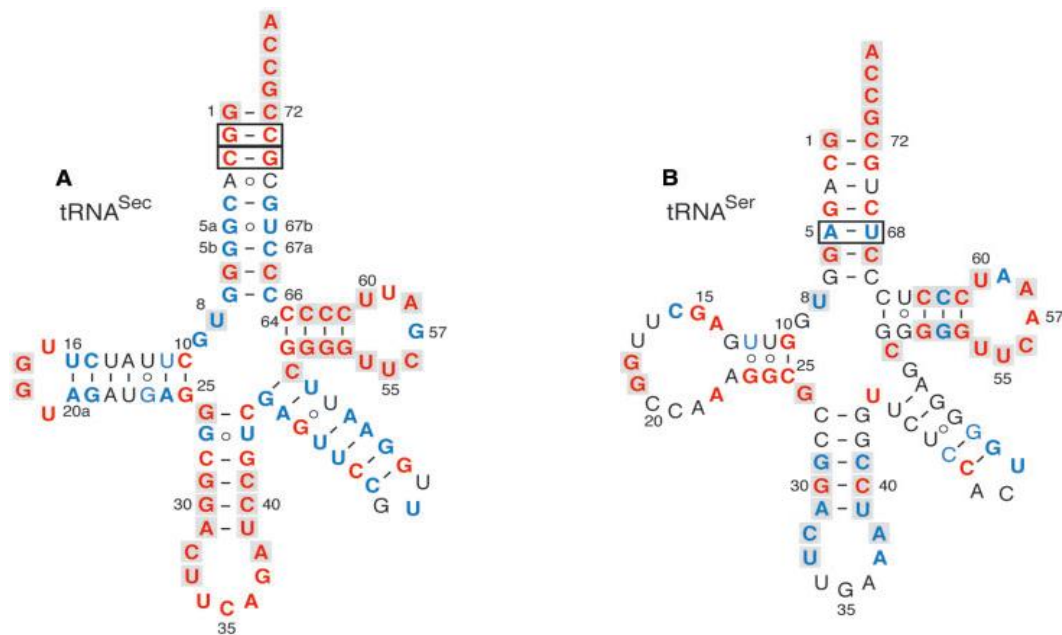




**Figure 5:** Sec biosynthesis and insertion into archaeal proteins.

### 1.3.1 Archaeal tRNA<sup>Sec</sup> identity elements and recognition by *O*-phosphoseryl-tRNA kinase

SerRS recognizes both tRNA<sup>Sec</sup> and tRNA<sup>Ser</sup>, but *O*-phosphoseryl-tRNA kinase (PSTK) must discriminate Ser-tRNA<sup>Sec</sup> and Ser-tRNA<sup>Ser</sup>. PSTK recognizes only tRNA<sup>Sec</sup> among all archaeal tRNA species; the tRNA identification by this enzyme requires a tRNA with a G1-C72 and a C2-G71 base pair, and also a pyrimidine in position 3 (Sherrer et al., 2008).



**Figure 6:** The cloverleaf structures of (A) *M. maripaludis* tRNA<sup>Sec</sup> and (B) tRNA<sup>Ser</sup>.

The crystal structure of the PSTK complex with tRNA<sup>Sec</sup> has revealed that not only involvement of acceptor stem base pair identity elements but also the specific interaction between the unique tRNA<sup>Sec</sup> D-arm and the PSTK CTD accounts for the strict tRNA<sup>Sec</sup> selectivity by PSTK (Chiba et al., 2010). The A5-U68 base pair in tRNA<sup>Ser</sup> has some anti determinant property for PSTK (Figure 6). Thus, PSTK has evolved to recognize the special features of tRNA<sup>Sec</sup>, which helps to ensure high fidelity encoding of Sec.

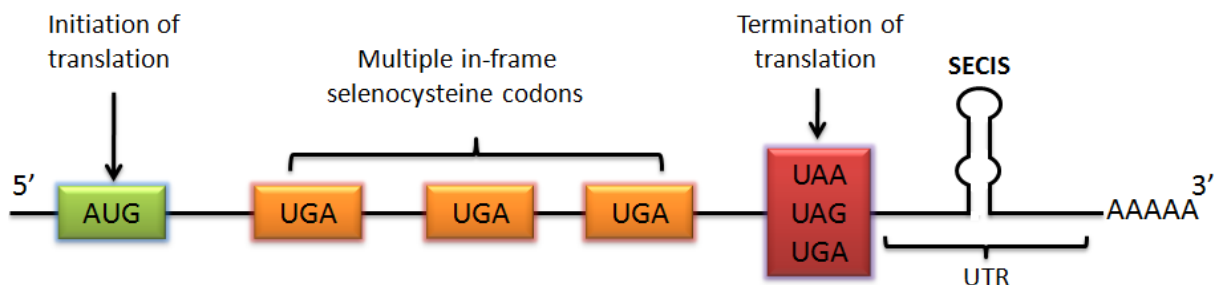
### 1.3.2 Sep-tRNA:Sec-tRNA (SepSecS)

In archaea and eukaryotes, Sec is formed in a tRNA dependent conversion of *O*-phosphoserine (Sep) by *O*-phosphoseryl-tRNA:selenocystenyl-tRNA synthase (SepSecS). Some methanogenic archaea harbor a gene that was thought to encode a SslA homolog (e.g. MJ0158), but its product is unable to synthesize Sec-tRNA<sup>Sec</sup> *in vitro* (Kaiser et al., 2005). The reactions catalyzed by PSTK and SepSecS are reminiscent of the indirect pathway of Cys-tRNA<sup>Cys</sup> synthesis in archaeal methanogens (Sauerwald et al., 2005), where Sep-tRNA<sup>Cys</sup> is converted to Cys-tRNA<sup>Cys</sup> by Sep-tRNA:Cys-tRNA synthase (SepCysS), a PLP-dependent enzyme carrying out a  $\beta$ -replacement on tRNA-bound Sep. The crystal structure of *Achaeglobus fulgidus* SepCysS has been reported (Fukunaga and Yokoyama, 2007). The crystal structure of archaeal and murine SepSecS apo-enzymes and phylogenetic analysis suggested that SepSecS forms its own branch in the family of fold-type I pyridoxal phosphate (PLP) enzymes. (Araiso et al., 2009). SepSecS catalyzes the conversion of the phosphoseryl moiety into the selenocysteinyl group by using selenophosphate as the selenium donor. An early observation that auto antibodies isolated from patients with type I autoimmune hepatitis targeted a ribonucleoprotein complex containing tRNA<sup>Sec</sup> led to the identification of the archaeal and the human SepSecS (Su, D et al 2009; Gelpi, C et al 1992). In the crystal structure, SepSecS binds only to the acceptor and variable arms of tRNA<sup>Sec</sup> and the most important binding element is an interaction between the discriminator base G73 of tRNA<sup>Sec</sup> and conserved Arg398 of the catalytic dimer. The discriminator base of tRNA<sup>Sec</sup> is universally conserved in archaea and eukaryotes. Replacing Arg398 with either alanine or glutamate renders the enzyme completely inactive, which suggest that the interaction between the discriminator base and the highly conserved Arg398 of the catalytic dimer is critical for

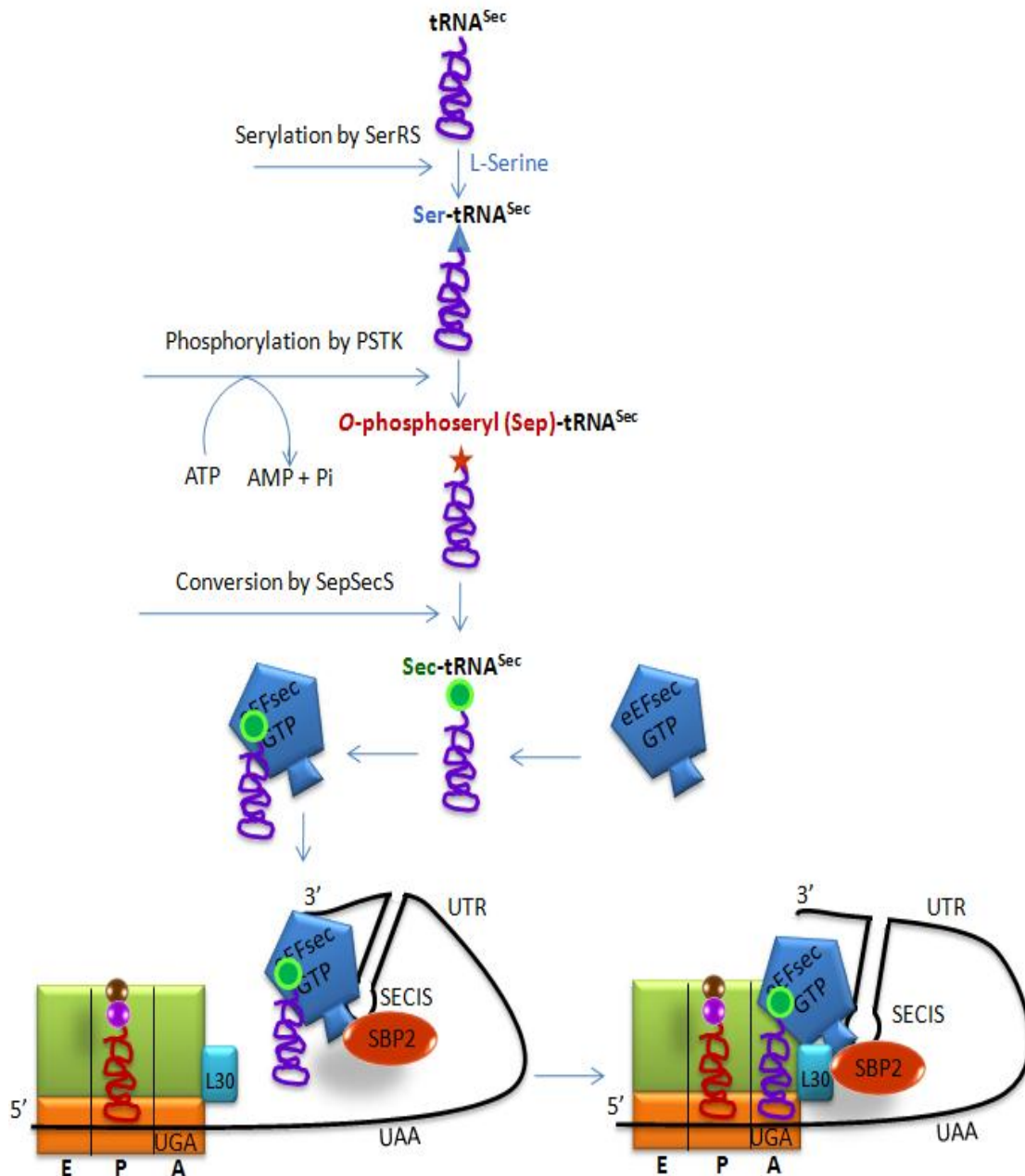
tRNA<sup>Sec</sup> recognition. Human tRNA<sup>Sec</sup> contains 90 nucleotides rather than the conventional 75 nucleotides of canonical tRNA molecules.

#### 1.4 Sec insertion into eukaryotic proteins

In eukaryal selenoprotein mRNAs, the required signals are also stem-loop SECIS-element, but they differ in sequence and structure from their prokaryotic counterparts and from each other (Rother et al., 2001). These SECIS-elements have been shown to recode several upstream in-frame UGA codons (Berry et al., 1993; Hill et al., 1993), a minimal spacing requirement is Met (Martin et al., 1996). It is known that it requires the existence of a SECIS-element in the 3'-untranslated region (UTR) of the mRNA with a distance from the in-frame UGA-element that naturally varies from 500 to 5300 nucleotides (Low and Berry, 1996).



**Figure 7:** Eukaryal selenoprotein mRNA contains multiple in-frame UGA codons and SECIS-element in the 3'-untranslated region.

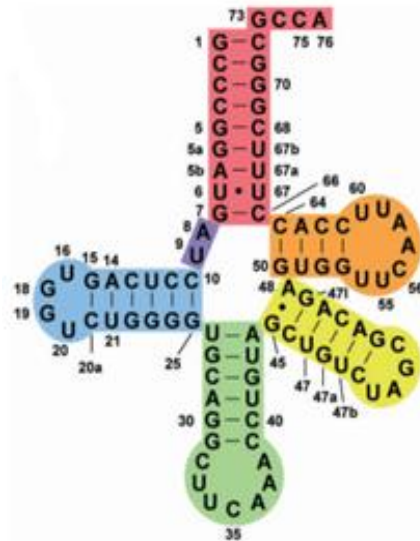


**Figure 8:** Sec biosynthesis and insertion into eukaryotic proteins.

#### 1.4.1 Eukaryal $\text{tRNA}^{\text{Sec}}$ identity elements and recognition by O-phosphoseryl-tRNA kinase

Eukaryal and archeal  $\text{tRNA}^{\text{Sec}}$  has a 9-bp acceptor stem and a 4-bp T stem, whereas bacterial  $\text{tRNA}^{\text{Sec}}$  has an 8-bp acceptor stem and a 5-bp T stem. The unique structure presumably allows  $\text{tRNA}^{\text{Sec}}$ , like  $\text{tRNA}^{\text{Ser}}$ , to function as a substrate for SerRS and in contrast as the exclusive target of PSTK, SepSecS (or SelaA), and EF-Sec. The recent crystal structure of

human tRNA<sup>Sec</sup> has revealed an unusual secondary structures of the acceptor, T and D stems, as well as unique tertiary interactions (Itoh et al., 2009). In contrast with the usual tRNAs, the long D stem of tRNA<sup>Sec</sup> does not interact with the extra arm (Itoh et al., 2009) (Figure 9).



**Figure 9:** Secondary structure of human tRNA<sup>Sec</sup>.

#### 1.4.2 Interactions between eEFsec, SBP2, ribosomal protein L30 and SECIS- element

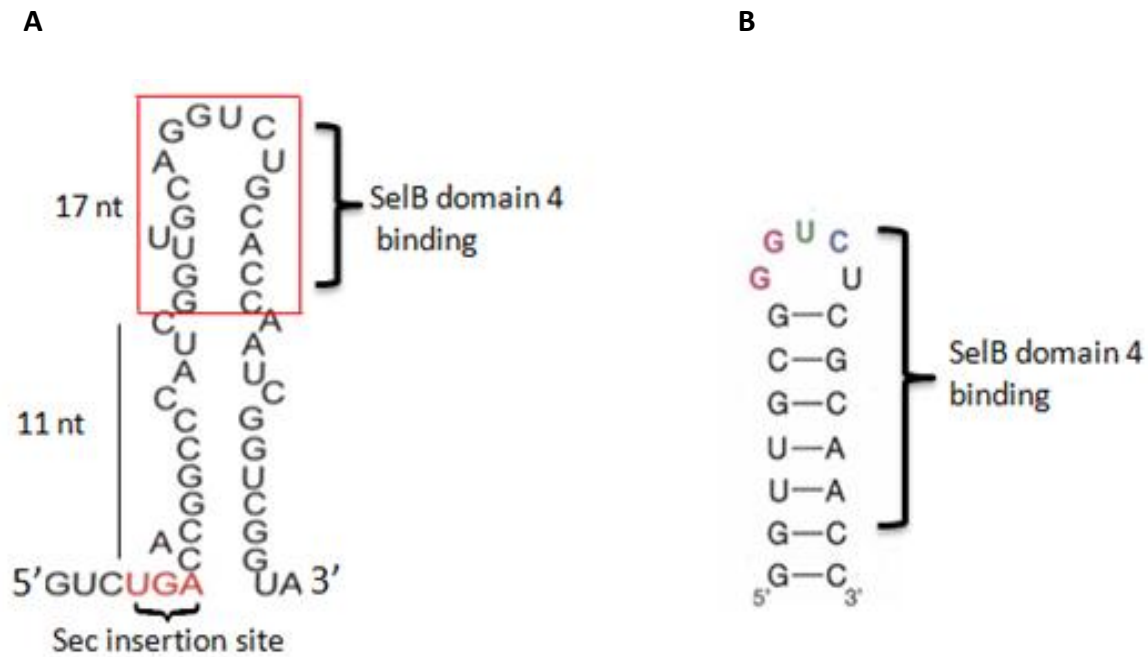
Eukaryotic SECIS elements were shown to have specific sequence and structural features, including a conserved G-A, A-G tandem purine pair at the base of a stem-loop, with the stem consisting of 10 to 12 base pairs (Martin et al., 1996) and this tandem purine pair is required for Sec insertion. Depletion of SBP2 resulted in abolition of Sec insertion, while addition of SBP2 restored Sec insertion, clear demonstrating that the SBP2 is required for Sec insertion. SBP2 did not show any sequence homology or activity of known elongation factors. Sec insertion directed from the 3'- UTR involves binding of SBP2 to the SECIS-element, followed by the recruitment of eEFsec-tRNA<sup>Sec</sup> by the SBP2-SECIS complex (Figure 8). Ribosomal protein L30 is an additional factor which may play a role in the Sec insertion process (Chavatte et al., 2005). Assembly of the complex at the SECIS-element, followed by delivery

of eEFsec-tRNA<sup>Sec</sup> to a UGA codon occupying the ribosomal A site, would allow translation of UGA codon at any site in the open reading frame (Figure 8).

## **1.5 The structure and identity elements of SECIS-RNA from organisms from three domains of life**

### **1.5.1 Bacterial SECIS-element**

The bacterial SECIS-element is an approximately 40 nucleotide long stem-loop structure that follows the UGA codon at the immediate 3' side. It is conserved by structure rather than by sequence in different organisms. Exceptions are organism like *Clostridium sticklandii* (Garcia and Stadtman, 1992) or *Eubacterium acidaminophilum* (Gursinsky et al., 2000) in which no such structures can be formed within the reading frames of mRNAs coding for selenopolypeptides. The solution structures of the SECIS-elements of the *E. coli fdhF* and *fdnG* mRNAs, were derived from chemical probing results combined with computer modeling. Despite differences in the sequence, identical bases within the upper part of the loop and stem of both hairpins interacted specifically with SelB indicative of identical tertiary structure. The lower region of the hairpin, which comprises the first 11 nucleotides following the UGA codon, is not required for SelB binding *in vitro* (Kromayer et al., 1996) and has been suggested to function by enhancing the stability of the loop structure and/or by preventing binding of RF2 (Huttenhofer et al., 1996), which recognizes UGA as a stop codon to terminate protein synthesis.



**Figure 10:** Secondary structure of SECIS. (a) *E. coli fdhF* SECIS-element (Huttenhofer et al., 1996). (b) *M. thermoacetica fdhA* SECIS-element (Selmer and Su, 2002).

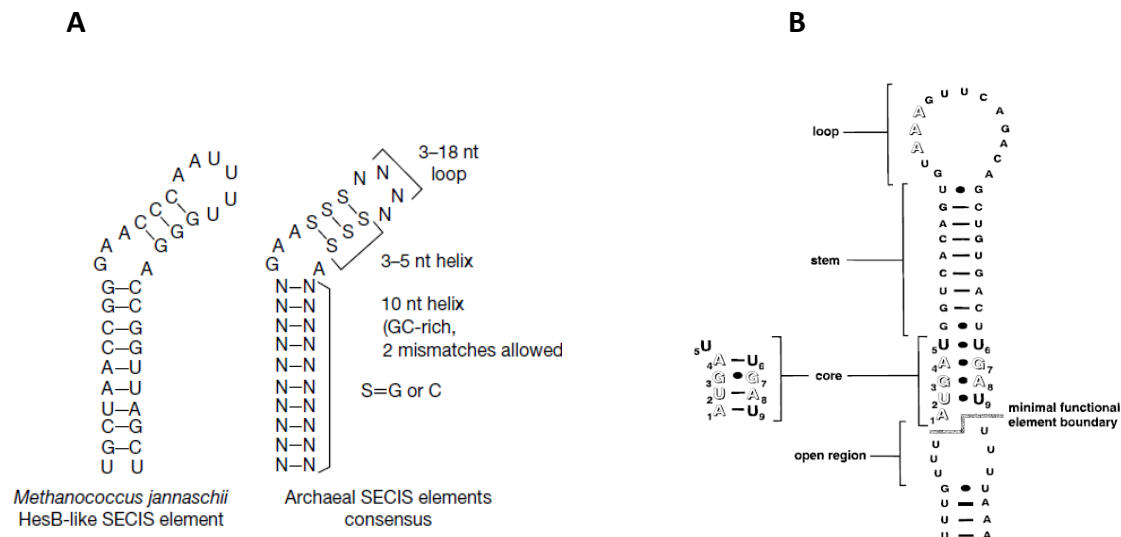
A displacement of the SECIS-element by more than one codon, or a displacement not preserving the reading frame results in a drastic reduction of selenocysteine insertion efficiency (Heider et al., 1992; Liu et al., 1998). The upper part of the stem-loop consisting of 17 nt is conserved within the mRNAs of formate dehydrogenases. This region is recognized and bound by SelB (Baron et al., 1993; Kromayer et al., 1996), with G23, U24 and U17/U18 bases playing a major role in SelB binding. G23 in the apical loop of the SECIS-element has been shown to bind SelB domain 4 and mutation at this region resulted in abolition of Sec insertion.

### 1.5.2 Archaeal and eukaryal SECIS-element

Eukaryotic SECIS-elements differ in sequence and structure from their prokaryotic counterparts and from each other (Rother et al., 2001). These SECIS-elements are found in



the 3' untranslated region (UTR) of the mRNA and can recode several internal UGA stop codons per gene (Hill et al., 1993).



**Figure 11:** (A) Archaeal SECIS-element structures. Archaeal SECIS element consensus sequence (right structure) and the SECIS-element in *M. jannaschii* HesB-like gene (left structure) are shown. (B). Eukaryotic SECIS-element of rat type-1 deiodinase; conserved structural features in the consensus structure are also indicated (Martin et al., 1998).

## 1.6 Efficiency of Sec insertion into bacterial, archaeal and eukaryotic proteins

### 1.6.1 Efficiency of Sec insertion into bacterial proteins

The components responsible for the recoding of a UGA stop codon as Sec have been studied intensively in bacteria employing *E. coli* as a model organism. According to the literature, the efficiency of UGA recoding into Sec by SelB-GTP-Sec-tRNA<sup>Sec</sup> in the presence of a SECIS-element in *E. coli* is very low, about 5% (Suppmann et al., 1999), when the cells are rapidly growing in rich media. Even when the Sec insertion machinery (SelB, tRNA<sup>Sec</sup> and SelA) is over expressed, the efficiency of Sec insertion remains low, about 10% (Suppmann et al.,

1999). When the cells are growing at the slowest growth rates in poor media, Sec insertion increased to 60% (Mansell et al., 2001). As growth rate decreases, the production of the overall number of RF2 molecules is reduced in the cell, although not apparently their free concentration (Adamski et al., 1994) and at this slow growth rates, the selection rate of Sec-tRNA<sup>Sec</sup> may exceed the RF2 selection rate. Genetic analysis indicated that the low efficiency of Sec incorporation is caused by termination at the UGA codon, rather than by the presence of a (stable) SECIS-element or the competition of the bulk of EF-Tu ternary complexes with the SelB·GTP·Sec-tRNA<sup>Sec</sup> (Mansell et al., 2001; Suppmann et al., 1999). Recoding of the UGA codon by Sec results in a translational pause of about 10 s (Suppmann et al., 1999), raising the possibility that recoding of UGA by SelB·GTP·Sec-tRNA<sup>Sec</sup> is an intrinsically inefficient process. However, the mechanism of discrimination between the specific insertion of Sec at a UGA codon by SelB·GTP·Sec-tRNA<sup>Sec</sup> and RF2-dependent termination is not clear.

### **1.6.2 Efficiency of Sec insertion into archaeal proteins**

Selenoproteins have been identified in *Methanococcus vannielii* and *Methanococcus voltae* and it has been shown that the insertion of Sec is directed by a UGA codon (Halboth and Klein, 1992). A search for open reading frames containing in-frame UGA codons in the genome of *Methanococcus jannaschii* (Bult et al., 1996) revealed the existence of seven putative selenoproteins (Wilting et al., 1997). Thus, as in bacteria and eukaryotes, the insertion of Sec is co-translational and directed by UGA. The efficiency of Sec insertion is thought to be similar to the one in eukaryotes, however, there is no clear literature information available on efficiency of Sec insertion into archaeal proteins.

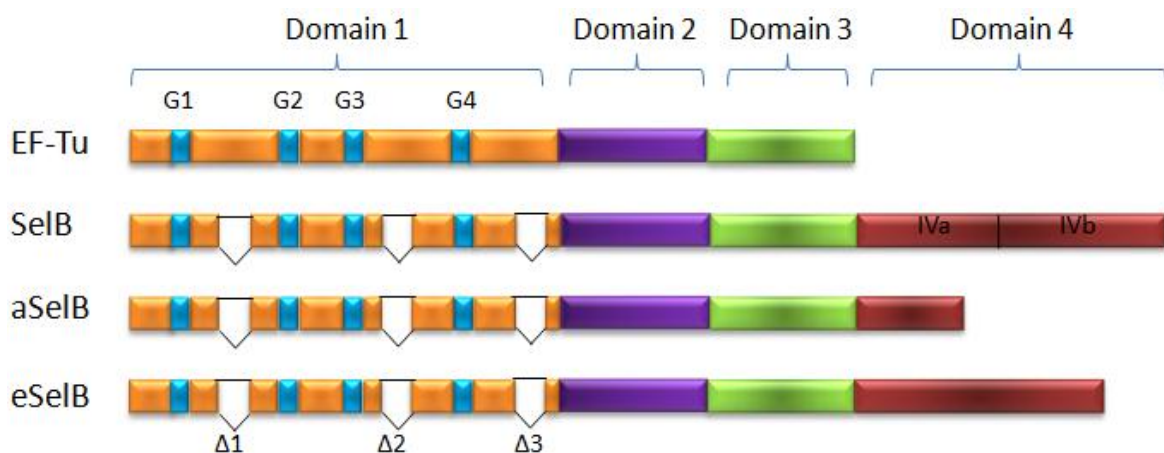
### 1.6.3 Efficiency of Sec insertion into eukaryotic proteins

Termination of translation in eukaryotes is catalyzed by release factors (RFs) eRF1 and eRF3. eRF1 confers recognition of all three termination codons (UAA, UAG and UGA) (Konecki et al., 1977), while eRF3 functions as eRF1 and ribosome-dependent GTPase (Frolova et al., 1996). Five components have thus far been identified that are necessary for the Sec insertion into ~25 eukaryotic selenoproteins. Two of these are *cis*-acting sequence, a SECIS-element and UGA codon. The three known *trans*-acting factors are eEFSec, Sec-tRNA<sup>Sec</sup> and the SECIS binding protein, SBP2. Other *trans*-acting factors have recently been implicated in the Sec insertion mechanism and its regulation, including the ribosomal protein L30 (Chavatte et al., 2005), nucleolin (Squires et al., 2007; Wu et al., 2000) and eIF4a3 (Budiman et al., 2009), each of which can interact directly with SECIS-elements. However, the exact roles of these latter proteins are not yet clear. If Sec insertion were in direct competition with termination, the prediction would be that the levels or activities of the RFs in cells would affect Sec incorporation efficiency, such that increase in termination efficiency would result in decreased synthesis of full-length selenoproteins. However, little is known about the levels of RFs in different tissues or cell lines or about circumstances that might affect RF levels or activity. Studies of the eukaryotic Sec incorporation mechanism suggest that Sec insertion is inefficient compared with termination. Nevertheless, selenoprotein P and several other selenoproteins are known to contain multiple selenocysteines. Studies on Sec incorporation efficiency in selenoprotein P have shown that the readthrough of the first UGA was inefficient with approximately 7% Sec incorporation efficiency. Readthrough of the second UGA codon occurred with much higher efficiency, approximately 67%, a nearly 10-fold increase (Fixsen and Howard, 2010).

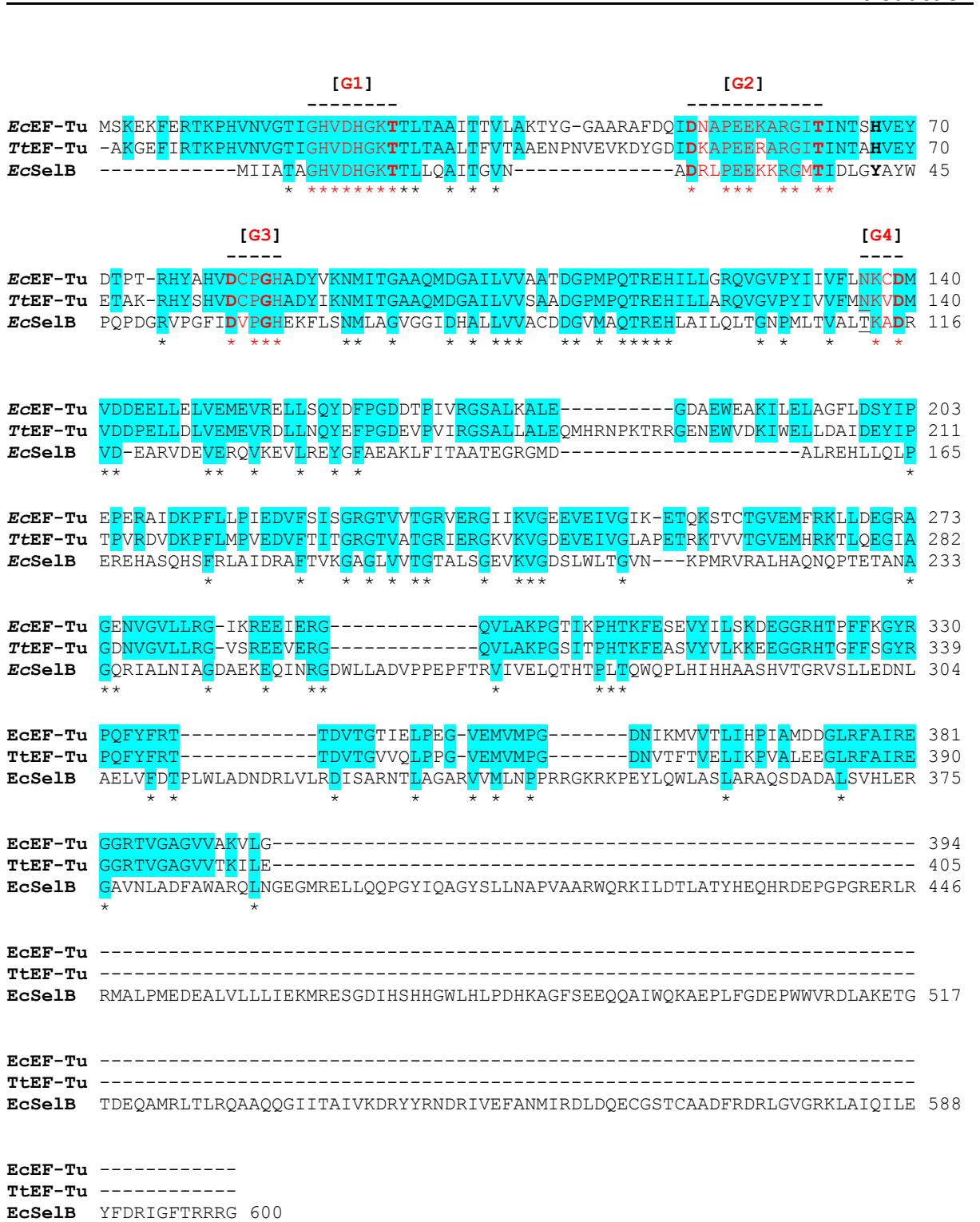
## 1.7 Comparisons of EF-Tu and SelB

### 1.7.1 Sequence and structural similarity

A structure of full-length SelB from *E. coli* is not available, but sequence comparisons and a homology model based on the available three-dimensional structure of full-length EF-Tu from *Thermus thermophilus* in the GTP bound conformation (Berchtold et al., 1993) suggest that SelB consists of 4 domains, where the N-terminal domains (1, 2, and 3) display sequence and structure similarity to EF-Tu and provide the binding site for GTP/GDP and Sec-tRNA<sup>Sec</sup> (Hilgenfeld et al., 1996) (Figure 12A, B). The C-terminal domain (4) did not show any sequence or structural similarity to known translational initiation or elongation factors.



**Figure 12:** Domain structure of the bacterial SelB and its archeal (aSelB) and eukaryal (eSelB) homologs, in comparison to EF-Tu. The G motifs involved in binding of the guanosine nucleotides are indicated by G1 to G4. Deletions within the SelB sequences relative to the EF-Tu primary structure are indicated.



**Figure 13:** ClustalW.2 sequence alignments of *E. coli* EF-Tu, *t. thermophilus* EF-Tu and *E. coli* SelB. Conserved amino acids are marked in blue.

### 1.7.2 GDP, GTP and aa-tRNA binding properties

SelB is a GTP-binding protein, which belongs to the family of translation factors (EF-Tu, EF-G, IF2, and RF3 and their eukaryotic homologs), and is evolutionary next to the translation initiation factor 2 (eIF-2γ) (Keeling et al., 1998).

**Table 4:** Comparison of guanosine nucleotide and aminoacyl-tRNA binding residues.

Motifs	<i>T. thermophilus</i> EF-Tu	<i>E. coli</i> SelB	Role
<b>Guanosine nucleotides</b>			
G1 (GXXXXGKT)	GHVDHGK <u>I</u> Thr- 25	GHVDHGK <u>I</u> Thr- 14	Phosphate binding loop 1 <sup>st</sup> coordination sphere of the Mg <sup>2+</sup> ion
G2	<u>D</u> KAPEERARG <u>I</u> Asp-51 Thr- 62	<u>D</u> RLPEEKRG <u>M</u> <u>I</u> Asp-26 Thr- 37	2 <sup>nd</sup> coordination sphere of the Mg <sup>2+</sup> ion 2 <sup>nd</sup> coordination sphere of the Mg <sup>2+</sup> ion
G3 (DXXG)	<u>D</u> CP <u>G</u> Asp-81 Gly-84	<u>D</u> VP <u>G</u> Asp-57 Gly-60	GTPase switch +GTPase switch
G4 (NKXD)	NKV <u>D</u> Asn-136 Asp-139	TKV <u>D</u> Thr-112 Asp-115	Guanosine base recognition
<b>tRNA of adenine 76- interactions (EF-Tu·Phe-tRNA<sup>Phe</sup>·GppNHp)</b>			
	Ile-231	Val- 185	
	Val-237	Val-191	
	Glu-271	Gln-224	
	Leu-289	Asn-240	
<b>Aminoacyl-moiety binding pocket (EF-Tu·Phe-tRNA<sup>Phe</sup>·GppNHp)</b>			
	His-67	Tyr-42	
	Glu-226	Asp-180	
	Asp-227	Arg-181	
	Phe-229	Phe-183	
	Thr-239	Thr-193	

<sup>+</sup>(Hilgenfeld et al., 1996)

The nucleotide-binding properties of SelB (Thanbichler et al., 2000) differ markedly from those of EF-Tu, but resemble those of EF-G and IF2 (Table 6) (Milon et al., 2006; Rodnina et al., 2000; Wilden et al., 2006; Gromadski K,B et al 2002; Mitkevich V, A et al., 2010).

### 1.7.3 EF-Tu·EF-Ts complex

The affinities of SelB for GTP and GDP are about 0.7  $\mu\text{M}$  and 13  $\mu\text{M}$ , respectively (Table 6) and the dissociation rate constant of GDP from SelB is high, 15  $\text{s}^{-1}$  (Thanbichler et al., 2000), which allows rapid and spontaneous nucleotide exchange. Thus, under *in vivo* conditions, nucleotide exchange is likely to occur rapidly and spontaneously without the help of nucleotide exchange factor (EF-Ts). Not surprisingly, SelB lacks the EF-Ts- interacting residues in domain 3 correlated well the finding that the affinity to GTP is significantly higher than to GDP which obviates the necessity for a guanosine nucleotide release factor. Due to the high affinity of EF-Tu for GDP, it is mostly the nucleotide exchange is carried out by EF-Ts. The crystal structure of *E. coli* EF-Tu in complex with EF-Ts revealed that the contact sites in EF-Tu are located in domain 1 and 3 (Kawashima et al., 1996). Only few of these residues in domain 1 are present in *E. coli* SelB, whereas the contacts in domain 3 (in *Tt* EF-Tu: Phe-335, Met-361, and Met-363) are absent (Table 5) (Hilgenfeld et al., 1996).

**Table 5:** Interaction with EF-Ts.

<i>E. coli</i> EF-Tu	<i>E. coli</i> SelB
<b>Domain 1</b>	<b>Domain 1</b> present
<b>Domain 3</b>	<b>Domain 3</b>
Phe-335	absent
Met-361	absent
Met-363	absent

## 1.8. Comparisons of SelB from prokaryotic, archaeal and eukaryotic organisms

```

EcSelB  -----MIIATAGVVDHGKTTLLQAITGVN----ADRLPEEKRGMTIDLGYAYWPQ-  47
MthSelB  -----MDYIVVGTAHVVDHGKTVLVKALTGVD----TRLKEEKERGISIELGFAPLTL-  50
MmaSelB  MHHHHHHSIEGRPHMDFKNINLGIFGHIDHGKTTLSKVLTEIASTSAHDKLPESQKRGITIDIGFSAFKL-  70
HsSelB  -----MAGRRVNVNVGVLGHIISGKTALARALSTTASTAAFDKQPQSRERGITLDLGFSCFSVP  59
          ** * *** *
          *

EcSelB  -----PDGRVPGFIDVPGHEKFLSNMLAGVGGIDHALLVACDDGVMAQTRE  94
MthSelB  -----PSGRQLGLVDVPGHERFIRQMLAGVGGMDLVMLVVAADGVMPOTRE  97
MmaSelB  -----ENYRIT-LVDAPGHADLIRAVVSAADIIDLALIVVDAKEGPKTQTGE  116
HsSelB  LPARLRSSLPEFQAAPAEPEPEPGEPLLQVTLVDCPGHASLIRTIIGGAQIIDLMMLVIDVTKGMQTQS-A  128
          * *** *
          *

EcSelB  HLAILQLTGNPMLTVALTKADRVDEARVDEVERQVKEVLREYG----FAEAKLFITAATEG-----  151
MthSelB  HLAIIDLLQIKKGIIIVITKIDLVLEADWLELVREEVRQAVKGTV----LEDAPLVEVSALTG-----  154
MmaSelB  HMLILDHFNIP-IIIVITKSDNAGTEEIKRTEMIMKSILQSTHN---LKNSSIIPIISAKTG-----  174
HsSelB  ECLVIGQIACQKLVVVLNKIDLLPEGKRQAIDKMTKKMQKTLENTKFRGAPIIPVAAKPGGPEAPETEA  198
          * *
          *

EcSelB  -RGMDAIR-EHLLQLPERE--HASQHSFRLAIDRAFTVKGAGLVVVTGTLALSSEVKVGDLSWLTGVNKPMPR  215
MthSelB  -EGIAELR-EQLDALAAVTPPRPAAGRVLPIDRVFSVTGFGTVVVTGTLWSGTIKVGDELEVQPEGLKTR  222
MmaSelB  -FGVDEIKNLIITTLNNAEIIIRNTESYFKMPLDHAFFPIKAGTVVTGTLINKGIVKVGDDELKVLPINMSTK  243
HsSelB  PQGIPELIELLTSQISIPT--RDPSPGFPLMSVDHCFISIKCGQTVMTGTILSGSISLGDVSEIPALKVVVK  266
          *
          * * * * *
          *

EcSelB  VRALHAQNQPTETANAGQRIALNIAGDAEKEQINRGDWLAD--VPPEPFTRVIVELQHTPLTQWQPL  282
MthSelB  ARNLQVHGRTVKEARAGQRAVANLAG-IETEAVHRGSSILTPGFLTPTYRLDASFLLNGARPLANRDRV  291
MmaSelB  VRSIQYFKESVMEAKAGRVGMAIQG-VDAKQIYRGXIITSKDTKLQTVDKIVAKIKISDIFKYNLTPKM  243
HsSelB  VKSMQMFHMPITSAMQEDRLGICVTQ-FDPKLLERG--LVCAPESLHTVHAALISVEKIPYFRGPLQTKA  265
          * * *
          *

EcSelB  HIIHHAASHVTGRVSLLEDN-----LAELVFDTPWLADNDRVLVLRDISARNTLAGARVVMNPPRRG  344
MthSelB  HFYLTSEALGRVVLDRDELNGGEEALIQLLMEKPVVASREDRFLRSYSFPMETIGGGI IIDPVPPKHR  361
MmaSelB  KVHLNVGMLIVPA-----VAVPFKKVTFGKTEENIILN-----EVISGNEXYXAFELEEK  293
HsSelB  KFHITVGHETVMG-----RLMFFSPAPDNFDQEPIIDLSFNFSQEYLFQEQYLSKDLTPA  321
          *

EcSelB  KRKPEYLQWLAS-LARAQSDADALSVHLERGAVNLAFAWARQLNGEGMRELLQ-----QPGYIQAG  344
MthSelB  RFQPEVLVSLQRRLEGSPEKILAQIIQEHREGLDWQEAATRASLSLEETRKLQSMAGQVTLRLRVEND  431
MmaSelB  VLAEVGDRVLITRLDLPPTTLRIXGHGLIEEFKPIKDLNIKKEVLREGKVKIDK-----  347
HsSelB  VTDNDEADKKAGQATEGHCPRQQWALVEFEKPVTCPRCLVIGSRLDADIHTNT-----  375

EcSelB  YSLLNAPVAARWQRKILDTLATYHEQHRDEPGPGRERLRMALPMEDEALVLLLLIEKMRESLDIHSHHGW  414
MthSelB  LYAISTERYQAWWQAVTRALEEFHSRYLPRPLGLAREELRSRYFSRLPARVYQALLEEWSREGLQLAANT  501
MmaSelB  -----GRTVIDGLAQS-----KVAAEKDIGEEISIEGKDIVGKIKGTFTGTC-----  389
HsSelB  -----CRLAFHGILLHGLEDRNYADSFLPRKLVYKLKHKHG-LVERAMDDYSVIG-----RS  426
          *
          *

EcSelB  LHLPDHKAGFSEEQQAIWQKAEPLFGDEPWVVRDLAKETG---TDEQAMRLTLRQAAQGGIITAIVKDRI  481
MthSelB  VALAGFTPSFSSETQKKLLKDLEDKYRVSRLQPPSFKEVAGSFNLDPSELEELLHYLVREGVLVKINDEFY  571
MmaSelB  LLTAEFSGNVENRDKVILNRLR-----RWG-----  414
HsSelB  LFKKETNIQLFVGLKVHLSTGELGIIDSFAFGQSGKFKIHIPGGLSPESKKILTPALKKR-----ARA  488

EcSelB  YRNDRIVEFANMIRDLQECGSTCAADFRDLGVGRKLAIQILEYFDRIGFTRRRG-----  537
MthSelB  WHRQALGEAREVIKNLASTG-PFGLAEARDALGSSRKYVLPLELYLDQVKFTRRVGDKRVVGN  631
MmaSelB  -----  530
HsSelB  GRGEATRQEESAERSEPSQHVVLSLTFKRYVFDTHKRMVQSP-----  530

```

**Figure 14:** ClustalW.2 sequence alignments of SelB from *E. coli*, *M.thermoacetica*, *M. maripalidus* and *H. sapiens*. Conserved amino acids are marked in green.



Sequence comparisons of SelB protein from prokaryotic (*Escherichia coli* and *Moorella thermoacetica*), archaeal (*Methanococcus maripaludis*) and eukaryotic (human) organisms shows that N-terminal domains are similar and domain 4 has completely unrelated sequence to each other.

### **1.8.1 N-terminal domains (1-3)**

In all three domains of life, SelB is a specialized translation elongation factor, which is responsible for the co-translational insertion of Sec into proteins by recoding the in-frame UGA stop codon in the presence of specific sequence at downstream hairpin structure (SECIS-element stem-loop structure). The detailed comparison of EF-Tu and SelB is discussed in the chapter (1.7). The identification of archeal SelB in *Methanococcus jannaschii* and *Methanococcus maripaludis* revealed an elongation factor with a shorter, 8 kDa C-terminal extension (Rother et al., 2000) (Figure 12). This, together with the finding of a gene harbouring two internal stop codons and only one SECIS-element in the 3' UTR, led to the proposal of a mechanism similar to the mammalian, adaptor mediated Sec incorporation (Rother et al., 2001). However, an SBP2 homologue has not been discovered in archaea so far. The crystal structure of full-length SelB from archaean *Methanococcus maripaludis* in the GDP-, GppNHp-, and apo-form shows that SelB has four distinct structural domains (1-4) that adopt a 'molecular chalice' arrangement. The arrangement of domains 1-3 in SelB is very similar in the presence of GTP and GDP (Leibundgut et al., 2005). The affinity of SelB·apo and SelB·GDP to Sec-tRNA<sup>Sec</sup> are similar and in the  $\mu\text{M}$  range but the affinity of SelB·GTP to Sec-tRNA<sup>Sec</sup> is in the pM range (available affinity measurements are summarized in Table 6). The

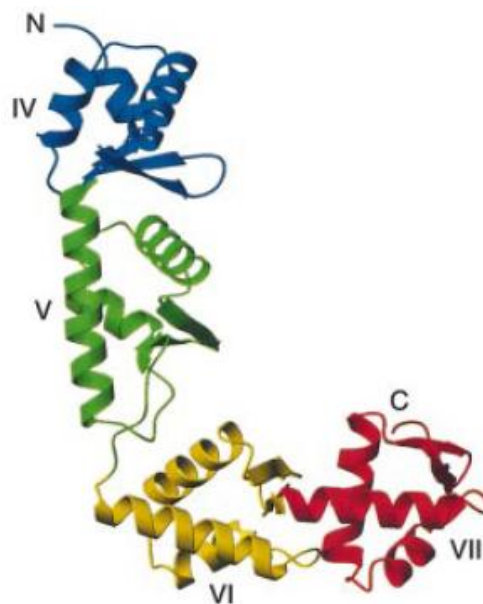
tight binding is achieved by formation of ion pairs which might involve in Sec moiety (Paleskava et al., 2010).

### 1.8.2 The SECIS-element binding, C-terminal domain (4)

An important property of the C-terminal domain (4) of SelB is that it binds to the SECIS-element (Kromayer et al., 1999; Kromayer et al., 1996); it has no analogies in other translational GTPases (Hilgenfeld et al., 1996; Kromayer et al., 1996).

#### 1.8.2.1 Prokaryotic SelB domain (4)

SelB binds to SECIS-elements very rapidly, close to the diffusion-controlled limit, and very tightly, with a  $K_d$  of the SelB·SECIS complex of about 1 nM (Thanbichler et al., 2000). The binding affinity does not depend on the nucleotide (GTP or GDP) bound to the G domain of SelB.

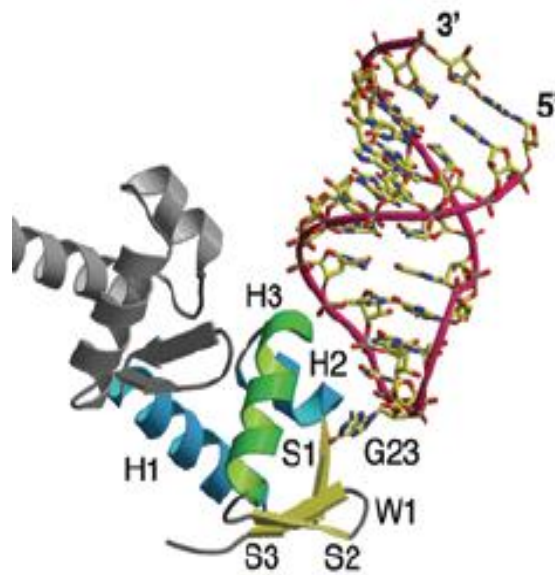


**Figure 15:** Overall structure of SelB domain 4.

Also the affinity of GTP/GDP to SelB does not depend on the presence of a SECIS-element, indicating that the respective binding sites do not influence each other (Table 6) (Thanbichler et al., 2000). But binding of Sec-tRNA<sup>Sec</sup> to SelB increased the affinity of the factor to SECIS-elements, indicating interplay between the two RNA binding sites (Table 6) (Thanbichler and Bock, 2001). The first crystal structure of the SECIS-binding fragment of SelB domain 4 from *Moorella thermoacetica* shows that the structure is L-shaped and each arm of L consists of two globular winged-helix (WH) domains (Selmer and Su, 2002) (Figure 14). The WH motif is a subfamily of the helix-turn-helix family (HTH) (Wintjens and Rooman, 1996) and is widely found in DNA- and RNA-binding proteins (Gajiwala and Burley, 2000). It has been suggested that conformational changes in SelB and the communication between the mRNA and the tRNA binding sites is essential to SelB function (Huttenhofer and Bock, 1998). The elongated L-shape of domain 4 is necessary to deliver the tRNA to the ribosomal A site and contact the mRNA SECIS hairpin at the solvent-exposed face of the 30S subunit (Selmer and Su, 2002).

#### 1.8.2.2 SECIS-RNA recognition by wing-helix motif

The crystal structure of SelB domain 4 from *M. thermoacetica* in complex with *fdhA* SECIS-element reveals backbone and loop recognition of a specific hairpin RNA by a WH motif (Figure 15). Out of four WH motifs only the C-terminal part interacts with the mRNA. Sequence-specific interactions are established with two conserved nucleotides (G23 and U24) at the epical loop of the SECIS hairpin (Yoshizawa et al., 2005) (Figure 16).



**Figure 16:** Crystal structure of SelB domain 4 complex with *fdhA* SECIS from *M. thermoacetica* (Yoshizawa et al., 2005).

#### 1.8.2.3 Mammals, SelB domain (4)

In mammals, SelB domain (4) is considerably shorter (Figure 12A) and binds to SBP2 (SECIS binding protein 2), a 94 kDa adaptor protein, which recognizes the SECIS-element and forms a quaternary complex together with SelB·GTP·Sec-tRNA<sup>Sec</sup> (Copeland et al., 2000; Tujebajeva et al., 2000).

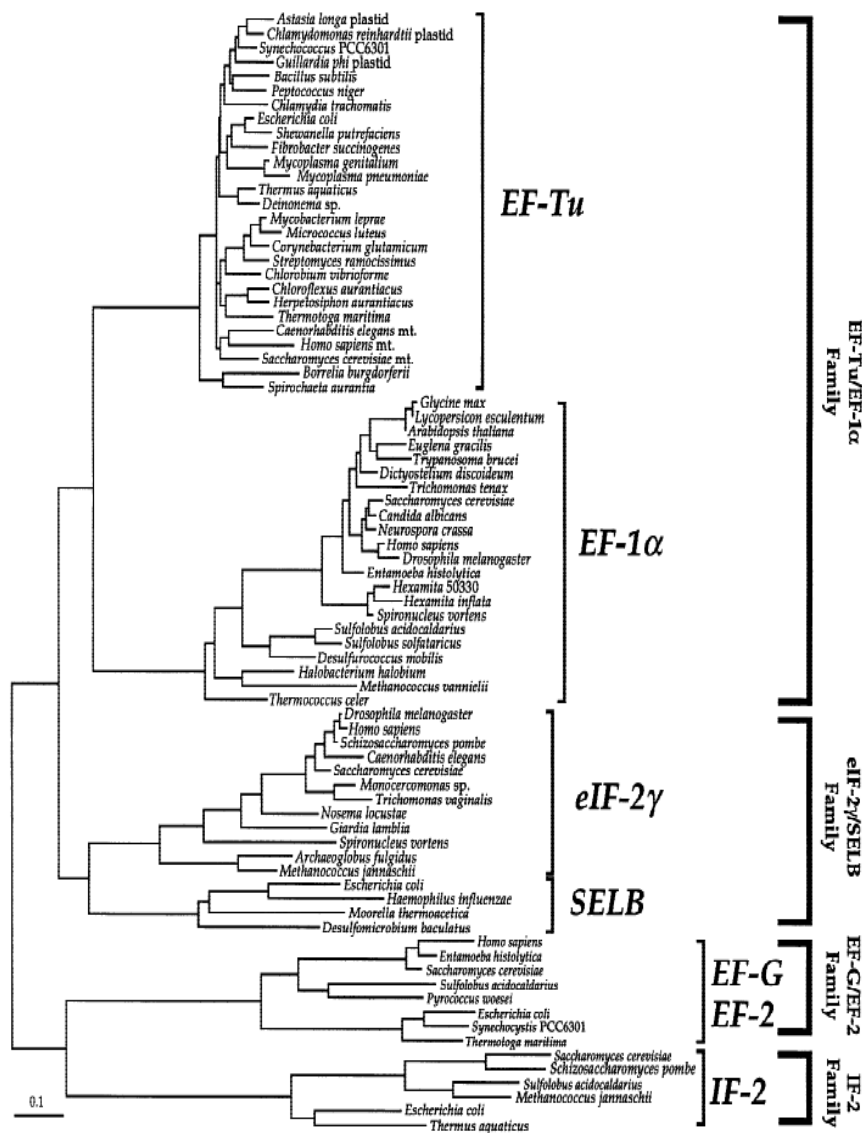
**Table 6:** Nucleotide binding properties of translational GTPases.

Source and factor	substrate	K <sub>d</sub> (nM)
<b><u>Escherichia coli:</u></b>		
SelB	GDP	13400 <sup>a</sup>
SelB	GTP	740 <sup>a</sup>
SelB·apo	Sec-tRNA <sup>Sec</sup>	420 ± 30 <sup>b</sup>
SelB·GDP	Sec-tRNA <sup>Sec</sup>	540 ± 60 <sup>b</sup>
SelB·GTP	Sec-tRNA <sup>Sec</sup>	(0.21 ± 0.06)·10 <sup>-3</sup> <sup>b</sup>
SelB·apo	<i>fdhF</i> SECIS-element (17 nt)	1.26 <sup>c</sup>
SelB·GDP	<i>fdhF</i> SECIS-element (17 nt)	1.19 <sup>c</sup>
SelB·GTP	<i>fdhF</i> SECIS-element (17 nt)	1.21 <sup>c</sup>
SelB· <i>fdhF</i> SECIS (17 nt)	GTP/GDP	N.D.*
SelB· <i>fdnG</i> SECIS (39 nt)	mantGDP	8750 <sup>d</sup>
SelB· <i>fdnG</i> SECIS (39 nt)	mantdGTP	530 <sup>d</sup>
EF-Tu	GDP	1 <sup>e</sup>
EF-Tu	GTP	60 <sup>e</sup>
EF-Tu·apo	Phe-tRNA <sup>Phe</sup>	N.D.*
EF-Tu·GDP	Phe-tRNA <sup>Phe</sup>	28,500 <sup>f</sup>
EF-Tu·GTP	Phe-tRNA <sup>Phe</sup>	0.85 <sup>f</sup>
IF2·apo	fMet-tRNA <sub>i</sub> <sup>fMet</sup>	300 <sup>g</sup>
IF2·GDP	fMet-tRNA <sub>i</sub> <sup>fMet</sup>	250 <sup>g</sup>
IF2·GTP	fMet-tRNA <sub>i</sub> <sup>fMet</sup>	180 <sup>g</sup>
<b><u>Eukaryotic:</u></b>		
eIF2-γ·apo	Met-tRNA <sub>i</sub> <sup>Met</sup>	115 <sup>h</sup>
eIF2-γ·GDP	Met-tRNA <sub>i</sub> <sup>Met</sup>	150 <sup>h</sup>
eIF2-γ·GTP	Met-tRNA <sub>i</sub> <sup>Met</sup>	9 <sup>h</sup>

\*N.D., not determined, <sup>a</sup>Thanbichler *et al.*, 2000, <sup>b</sup>Paleskava A *et al.*, 2010, <sup>c,d</sup>Thanbichler *et al.*, 2000, <sup>e</sup>Gromadski K,B *et al* 2002, <sup>f</sup>Dell V. A *et al.*, 1990, <sup>g</sup>Mitkevich V, A *et al.*, 2010, <sup>h</sup>Kapp, L. D *et al.*, 2004.

### 1.9 Evolutionary relationship of eIF-2 $\gamma$ and bacterial SelB

Translational GTPase are a family of proteins whose GTPase activity is stimulated by the large ribosomal subunit. The phylogenetic position of the eIF-2 $\gamma$  subfamily within the GTPase super family represented that the most closely related members of the EF-Tu, EF-1 $\alpha$ , and SelB (Figure 17) and in this tree eIF-2 $\gamma$  is most closely related to SelB. In all, there are four such groups, EF-2 and EF-G, EF-1 $\alpha$  and EF-Tu, the IF-2 like group, in addition to the eIF-2 $\gamma$  and SelB group (Keeling et al., 1998).



**Figure 17:** Evolutionary tree of translation factors (Keeling et al., 1998).

### 1.10 Evolutionary conservation of Sec insertion machinery

Although Sec-containing proteins have been identified in the three major domains of life (bacteria, archaea and eukaryotes), certain representatives of these organisms lack selenoprotein genes. Recent availability of complete sequences of many genomes allows assessment of selenoprotein genes encoded in these genomes. Interestingly, about half of the completely sequenced genomes in each of the three domains of life appear to lack selenoprotein genes. In these genomes, neither genes that are conserved among organisms that contain the Sec insertion system (e.g., SPS, tRNA<sup>Sec</sup>, SelB etc.) nor known selenoproteins can be found. Evolutionary implications of this observation are interesting and currently not fully understood. It was predicted that initial gain of Sec function may have been linked to the evolution of the UGN codons that inserted Cys in the primordial world. It was suggested that during development of the genetic code, the presence of a purine in the third position of the codon could indiscriminately code for Sec or Cys, while a pyrimidine in that position could conserve the Cys insertion function. It was also suggested that UGG was trapped by the newly evolved tryptophan and UGA evolved into a Sec codon. The hypothesis suggests that Sec evolved in later stages of evolution, after the 20 amino acids with their initially specified codons evolved (Gladyshev and Kryukov, 2001). It is now clear that Sec is strikingly different from the other 20 amino acids in its basic biosynthetic characteristics and its incorporation into protein. The differences can be summarized as follows: 1) the codon for Sec, UGA, serves a dual function in the genetic code which is the insertion of Sec into protein and the termination of protein synthesis (and the only other codon in the genetic code that serves a dual function is AUG which codes for the initiation of protein synthesis and the insertion of Met at internal positions of proteins); 2) a Sec insertion sequence (SECIS-element), which is the stem-loop structure located downstream of the Sec UGA codon, is the

only known mRNA structure that specifies insertion of an amino acid; 3) Sec is biosynthesized on its tRNA from serine which is attached to the tRNA by seryl-tRNA synthetase, an enzyme that most likely predated the use of Sec in protein as it must have evolved early for serine fidelity; 4) Sec tRNAs contain relatively few modified bases compared to other tRNAs; 5) Sec has its own specific elongation factor in bacteria, archaea and eukaryotes, while the other 20 amino acids share a common elongation factor (EF-Tu). These properties argue that Sec biosynthesis and its insertion into selenopolypeptides are unique compared to the other amino acids and that Sec was likely added to the already existing 20 amino acid genetic code. Irrespective of when the incorporation of Sec into protein originated, this amino acid serves as an example of numerous unique modifications that emerged for its specific use within the universal genetic code. It will be interesting to see if other examples of such novel variations in the genetic code will be found. Selenogenomes of various species have been characterized, i.e., those of important human pathogens (e.g. *Mycobacterium avium*). Their functional analysis is still rudimentary. However, they reveal that the use of the trace element changed dramatically during evolution. The machinery for selenoprotein biosynthesis was lost in yeast, terrestrial plants, and some bacteria. e.g., *Bacillus subtilis* has no tRNA<sup>Sec</sup> or SelB protein (Matsugi and Murao, 2004); hence, co-translational incorporation of Sec does not occur in this organism which may correlate to a high concentration of selenium in soil, the natural environment of *Bacillus subtilis*. It has been suggested that random incorporation of Sec into proteins occurs in this organism because cystenyl-tRNA synthetase cannot distinguish between cysteine and selenocysteine (Matsugi and Murao, 2004). The lack of SelB is not confined to soil bacteria. In fact, many human symbionts and pathogens do not contain SelB. On the other hand, *Pseudomonas putida*, also a soil bacterium, contains the *SelB* gene. In *Mycoplasma*, UGA



codons designate Trp (Christiansen et al., 1997). Another surprising feature of the distribution of SelB is its sporadic presence in several bacterial groups i.e, *Clostridium perferingens* contains the gene but *Clostridium tetani* does not; *Treponema denticola* has the gene but *Treponema pallidum* does not; *Mycobacterium avium* has it and other *Mycobacterium* species do not. Since the selenoprotein biosynthesis is a complex organism specific pathway, there are some barriers for the heterologous expression of genes encoding selenoprotein. On the positive side, it was found that there is a high degree of conservation of the selenoprotein biosynthesis throughout the enterobacteria (Heider et al., 1991).

**1.11 Aim and scope of the work**

In eubacteria, the co-translational mode of selenocysteine incorporation into proteins is a bypass of the classical way of protein synthesis since the recognition and delivery of the unique Sec-tRNA<sup>Sec</sup> to the in-frame UGA and stem-loop mRNA programmed ribosome is not by a canonical elongation factor Tu but rather by a specialized elongation factor SelB. It has been known until now, that the function of four gene products involve in selenocysteine synthesis and insertion. However, still there is lack of clear information on several mechanistic events of selenocysteine insertion into nascent peptide on the ribosome. The study of this unique mechanism is of basic interesting topic in the field of ribosomal translation and it has great importance in the evolutionary point of view as well. The aim of this work is to develop experimental assays to study selenocysteine incorporation into proteins *in vivo* and *in vitro*. Taking together all these results allowed us to study the efficiency of UGA recoding into selenocysteine by the bacterial translational machinery.

## 2 Materials and Methods

### 2.1 Plasmid construction for *in vivo* experiments

#### 2.1.1 Cloning of *E. coli fdhF* gene (Thr130-Asp179) into p94 plasmid

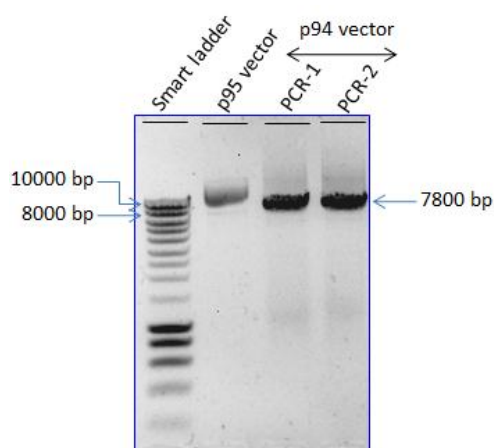
Blunt end cloning procedure was carried out according to the manufacturer's protocol (NEB).

p94 vector PCR condition

Forward primer: FP008

Reverse primer: FP007

Step 1:	Initial denaturation:	98°C, 2 minutes 30 seconds	1 cycle
Step 2:	Denaturation:	98°C, 15 seconds	
Step 3:	Primer annealing:	70°C, 15 seconds	
Step 4:	Primer extension:	72°C, 3 minutes	steps 2 to 4, 30 cycles
Step 5:	Final extension:	72°C, 10 minutes	1 cycle



**Figure 18:** Agarose gel (0.8%).

#### *E. coli fdhF* gene (Thr130-Asp179) UGA<sub>140</sub> + stem-loop

Insert PCR condition

Forward primer: SBK04

Reverse primer: SBK03

Step 1:	Initial denaturation:	98°C, 5 minutes	1 cycle
Step 2:	Denaturation:	98°C, 1 minute	

- Step 3:     Primer annealing:       70°C, 1 minute
- Step 4:     Primer extension:       72°C, 5 minutes       steps 2 to 4, 30 cycles
- Step 5:     Final extension:       72°C, 10 minutes       1 cycle

Insert PCR product was observed on agarose gel. Ligation reaction was performed according to the manufacturer's protocol (USB rapid ligation, USA). Transformation into *E. coli* NovaBlue competent cells were performed according to the manufacturer's protocol (Novagen).

### **2.1.2 Stem-loop deletion mutagenesis**

p95 vector containing *E. coli fdhF* gene (Thr130-Asp179) + stem-loop was used as a template.

Forward primer: SBK05

Reverse primer: SBK06

- Step 1:     Initial denaturation:   98°C, 2 minutes       1 cycle
- Step 2:     Denaturation:       98°C, 15 seconds
- Step 3:     Primer annealing:    70°C, 15 seconds
- Step 4:     Primer extension:    72°C, 3 minutes       steps 2 to 4, 30 cycles
- Step 5:     Final extension:    72°C, 10 minutes       1 cycle

PCR product was observed on agarose gel. Procedure as described above.

### **2.1.3 Site-directed mutagenesis of *fdhF* UGA<sub>140</sub> +SL to *fdhF* UUC<sub>140</sub> +SL by Quick-change PCR**

p94 vector PCR condition

Forward primer: fdhF.Frd-TGA-TTC

Reverse primer: FdhF.Rev-TGA-TTC

- Step 1:     Initial denaturation:   98°C, 2 minutes       1 cycle
- Step 2:     Denaturation:       98°C, 15 seconds

Step 3:	Primer annealing:	70°C, 15 seconds	
Step 4:	Primer extension:	72°C, 3 minutes	steps 2 to 4, 30 cycles
Step 5:	Final extension:	72°C, 10 minutes	1 cycle

p94 wild type plasmid *fdhF* UGA<sub>140</sub> +SL was amplified with *pfu* DNA polymerase (Promega). Amplified product was identified on gel and parental template DNA was digested with restriction enzyme Dpn1 (Fermentas) and 5 µl of amplified product was transformed into NovaBlue cells (Novagen) and plasmid DNA was purified according manufacturer's (Qiagen plasmid mini prep). Mutation was confirmed by DNA sequencing.

### 2.1.4 Table: 7 Oligodeoxyribonucleotides used.

Name	Sequence
<b>EfdhFHF1:</b>	5'- GTACAAGCTTATGAAAAAAGTCGTCACGG-3'
<b>EfdhFER2:</b>	5'- CGATGAATTCTTACGCCAGTGCCGCTTCG-3'
<b>fdhF129FSD:</b>	5'- ATGAATTCTTAAGTGGAGCGAGACCGATGGGTAACCAATAACGTTGAC-3'
<b>fdhF180Rev:</b>	5'- GCTGAAGCTTTTAGGAATCCGCCGGGTTGTACC-3'
<b>M13/pUC Rev:</b>	5'- CAGGAAACAGCTATGAC-3'
<b>fdhF.Rev156:</b>	5'- TCGCCGTTACCGACCGATTG-3'
<b>fdhF.Frd-TGA-TTC:</b>	5'- CGCTCGTGTCTTCCACGGCCCAT-3'
<b>fdhF.Rev-TGA-TTC:</b>	5'- ATGGGCCGTGGAAGACACGAGCG-3'
<b>SBK04:</b>	5'- ATGTTACCAATAACGTTGACTGCT-3'
<b>SBK03:</b>	5'- ATCCGCCGGGTTGTACCCG-3'
<b>SBK05:</b>	5'- AATGGCGCAATGAGGCAATG-3'
<b>SBK06:</b>	5'- TCAGACACGAGCGCAGCAGTC-3'
<b>No SL Frd.UUC:</b>	5'- TCGCTCGTGTCTTCAATGGCGCAATG-3'
<b>No SL Rev.UUC:</b>	5'- CATTGCGCCATTGAAGACACGAGCGCA-3'
<b>FP007</b>	Reverse primer: 5'- P-CAATTTGGACTTTCCGCCCTTCTTGCC-3'
<b>FP008</b>	Forward primer: 5'- P-ATGACTTCGAAAGTTTATGATCCAGAACAAAGGAAACG-3'

Underlined mutation position; P: phosphorylated.

**2.1.5 Table: 8** Strain and plasmids of *in vivo* and *in vitro* experiments.

Strain or plasmid	Genotype or phenotype	Reference or source
<b><i>E. coli</i> strains</b>		
Tuner (DE3)	$F^-$ <i>ompT hsdSB(rB<sup>-</sup> mB<sup>-</sup>) gal dcm lacY1 (DE3)</i>	Novagen
NovaBlue	<i>endA1 hsdR17 (r<sub>K12</sub><sup>-</sup> m<sub>K12</sub><sup>+</sup>) supE44 thi-1 recA1 gyrA96</i> <i>relA1 lac F'[proA<sup>+</sup>B<sup>+</sup> lacI<sup>q</sup>ΔM15::Tn10] (Tet<sup>R</sup>)</i>	Novagen
<b>Plasmids</b>		
pWL 194 SelB (wt)		(Forchhammer et al., 1990)
<b><u>In vitro:</u></b>		
pUC19 (full-length <i>fdhF</i> gene)		this study
pTZ18R (wt <i>fdhF</i> 129-180, <i>fdhF</i> SD followed by AUG, GGU)		this study
pBluescript II SK(+) ( <i>fdhF</i> 129-180)UGA <sub>140</sub> + Stem-Loop (strong SD, AUG UUC)		this study
pBluescript II SK(+) ( <i>fdhF</i> 129-180) UUC <sub>140</sub> + Stem-Loop (strong SD, AUG UUC)		this study
<b><u>In vivo:</u></b>		
p94 plasmid (firefly luciferase + renilla luciferase) parental plasmid for cloning		
p94 plasmid ( <i>fdhF</i> 130-179, UGA <sub>140</sub> +Stem-Loop)		this study
p94 plasmid ( <i>fdhF</i> 130-179, UGA <sub>140</sub> <sup>-</sup> Stem-Loop)		this study
p94 plasmid ( <i>fdhF</i> 130-179, UUC <sub>140</sub> <sup>+</sup> Stem-Loop)		this study
p94 plasmid ( <i>fdhF</i> 130-179, UUC <sub>140</sub> <sup>-</sup> Stem-Loop)		this study
p94 plasmid (T <sub>7</sub> followed by Firefly alone)*		this study
p94 plasmid (T <sub>7</sub> followed by Renilla alone)*		this study

\*Kind gift from Neva Caliskan

## 2.2 Expression and induction of fusion protein

### 2.2.1 Induction of firefly-*fdhF*-renilla fusion protein expression at initial-log phase

#### Preculture preparation

Glycerol stock of bacterial culture previously transformed with test constructs was streaked on an agar plate containing antibiotic kanamycin 30 µg/ml and plate was incubated overnight at 37°C. For preparation of preculture, a single colony was inoculated into LB media containing kanamycin 30µg/ml and grown over night at 37°C, 210 rpm (cell density reached an OD<sub>600</sub> of ~ 5.0-7.0 OD/ml).

#### Growth media: LB

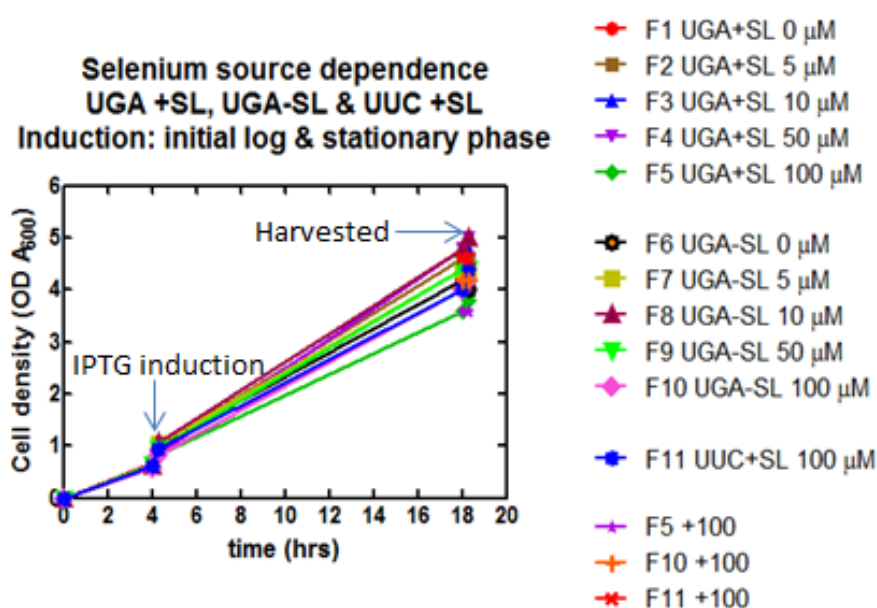
LB medium inoculated with preculture to the final concentration of 0.1%. Bacteria were grown aerobically at 37°C, 210 rpm to an OD<sub>600</sub> of ~ 0.5- 0.7, an aliquot was taken (which is equal 1 OD) for non-induction control and the expression of firefly-*fdhF*-renilla fusion protein was induced by addition of IPTG to the final concentration of 1 mM. Induction was continued for 30-40 minutes and then several aliquots of the cell culture were taken (which is equal 1 OD) and aliquots of cell pellets were stored at -20°C until use.

### 2.2.2 Induction of firefly-*fdhF*-renilla fusion protein expression in the presence of selenium concentration dependence at initial-log phase and stationary phase

**TPG growth media:** 1% Tryptone, 100 mM potassium phosphate pH 6.5, 0.5% glycerol, 1 mM MgSO<sub>4</sub>, 0.1 mM CaCl<sub>2</sub>, 0.4 µM H<sub>3</sub>BO<sub>3</sub>, 30 nM COCl<sub>2</sub>, 10 nM CuSO<sub>4</sub>, 10 nM ZnSO<sub>4</sub>, 80 nM MnCl<sub>2</sub>, 10 µM FeCl<sub>3</sub> and selenium source (0, 0.5, 1, 5, 10, 50 and 100 µM Na<sub>2</sub>SeO<sub>3</sub>).

Expression and induction until initial-log is same as above. Cells were grown further until reaches the OD<sub>600</sub> of ~ 5-6 (stationary phase) and induced the expression of firefly-*fdhF*-renilla fusion protein was induced by addition of IPTG to the final concentration of 1 mM and

continued the growth for another 30-40 minutes (Figure 19). Aliquots were taken (1 OD) and cell pellets were stored at -20°C until use.



**Figure 19:** Expression and induction of firefly-*fdhF*-renilla fusion protein at initial-log and stationary phase.

### 2.2.3 Induction of firefly-*fdhF*-renilla fusion protein expression at various phases of cell growth (Initial-log, mid-log and stationary phase)

**TPG growth media:** 1% Tryptone, 100 mM potassium phosphate pH 6.5, 0.5% glycerol, 1 mM MgSO<sub>4</sub>, 0.1 mM CaCl<sub>2</sub>, 0.4 μM H<sub>3</sub>BO<sub>3</sub>, 30 nM COCl<sub>2</sub>, 10 nM CuSO<sub>4</sub>, 10 nM ZnSO<sub>4</sub>, 80 nM MnCl<sub>2</sub>, 10 μM FeCl<sub>3</sub> and selenium source 50 μM Na<sub>2</sub>SeO<sub>3</sub>.

Rest of the expression is same as above. Aliquots were taken (1 OD) and cell pellets were stored at -20°C until use.



**2.3 Immunoblot with  $\alpha$ -firefly and  $\alpha$ -renilla antibodies**

0.5 OD<sub>600</sub> of cell extracts were boiled in SDS sample buffer and samples were separated on 12% SDS polyacrylamide gel (15V overnight). To detect induced firefly-fdhF-renilla fusion protein separately, two gels were ran in parallel one gel for detection of firefly-fdhF-renilla fusion protein using  $\alpha$ -firefly antibody and another is for detection of firefly-fdhF-renilla fusion protein using  $\alpha$ -renilla antibody. Proteins were transferred to a nitro cellulose membrane 0.2  $\mu$ m (Optitran BA-S 83, Whatmann) by semi-dry blotting (Peq-Lab) (0.8 mA/cm<sup>2</sup> 2 hours at room temperature). Transfer was checked by staining total protein with ponceau red (0.1% ponceau red and 1% acetic acid) about one minute and then destaining with 1% acetic acid 3 times. Membranes were blocked in 1 x PBS containing 20% skim milk powder and 0.1% Tween 20 overnight at 4°C. For detection of firefly luciferase, the blots were incubated with primary antibody Anti-Luciferase pAb, IgG, (Promega) (1:1000) raised in goat, 2 hours at room temperature and then washed 3 times with 1 x PBS containing 0.1% Tween 20 (1 x PBST). Then the blots were further incubated for 45 minutes at room temperature with secondary antibody Anti-Goat HRP (1:10,000). Then blots were washed 3 times with 1 x PBST. For detection of renilla luciferase, blots were incubated with primary antibody Anti-Renilla Luciferase IgG, (1:10,000) (Millipore) raised in mouse, 2 hours at room temperature and then washed 3 times with 1 x PBST. Then blots were incubated for 1 hour with secondary antibody Anti-Mouse IgG, (1:5,000) (Jackson ImmunoResearch) raised in goat. Then blots are washed 3 times with 1 x PBST. All antibodies were diluted in 1 x PBST containing 5% skim milk. The antigen-antibody complexes were detected using Super Signal West Pico Chemiluminescence Substrate (Pierce). Membranes were exposed to high performance chemiluminescence film (GE Healthcare) about 10 seconds and film was developed by standard procedure.

## **2.4 Dual reporter assay**

### **2.4.1 Luciferase activity measurements**

1 OD of cells were lysed in 70  $\mu$ l lyses buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA and 5 mg/ml Lysozyme) on ice for 15 minutes. Cells were frozen in liquid nitrogen followed by incubation at 37°C for 15 minutes. Cell lysates were centrifuged and supernatants were collected by centrifugation at 13,000 rpm for 10 minutes at 4°C. All the measurements were performed using luminometer (Sirius Berthold) and measurements delay time is 2 seconds, measure time is 5 seconds. First firefly luciferase and then renilla luciferase were measured. 5  $\mu$ l of supernatant were taken for measurements and mixed with substrate luciferin in the presence of ATP and activities were measured over time (0, 5, 10, 15, 20 minutes). 5  $\mu$ l of aliquots were taken for measurements and mixed with substrate renilla glow and Coelenterazine and activities were measured over time (0, 5, 10, 15, 20 minutes). All the measurements were performed at room temperature and each experiment was performed three individual measurements. Readthrough was calculated using following formula.

### **2.4.2 Calculation of recoding efficiency**

$$\text{Recoding efficiency \%} = \frac{\text{Rluc test} / \text{Fluc test}}{\text{Rluc control} / \text{Fluc control}} \times 100$$

## 2.5 Plasmid construction for *in vitro* experiments

### 2.5.1 Cloning of *E. coli* full-length *fdhF* gene into pUC19 vector

Sticky end cloning procedure was carried out according to the manufacturer's protocol (NEB).

Full-length *fdhF* PCR condition

Forward primer: EfdhFH1

Reverse primer: EfdhFER2

Step 1: Initial denaturation: 98°C, 5 minutes 30 seconds 1 cycle

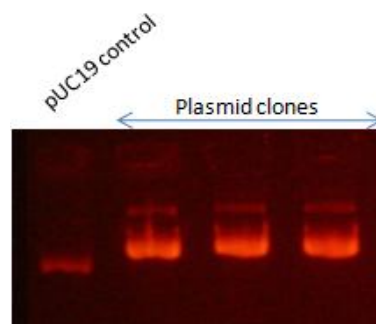
Step 2: Denaturation: 98°C, 1 minute

Step 3: Primer annealing: 70°C, 15 seconds

Step 4: Primer extension: 72°C, 5 minutes steps 2 to 4, 30 cycles

Step 5: Final extension: 72°C, 10 minutes 1 cycle

PCR product was observed on agarose gel (Figure 20) and Ligation reaction was performed according to the manufacturer's protocol (USB rapid ligation, USA). Plasmid DNA was purified according to the manufacturer's (Qiagen plasmid mini prep) (Figure 20). The *fdhF* insert was confirmed by DNA sequencing.



**Figure 20:** Plasmid mini prep of pUC19 vector contain full-length *fdhF* fragment.

**2.5.2 Cloning of *fdhF* DNA gene (Gly129-Ser180) into pTZ18R *in vitro* transcription vector**

Cloning procedure same as above

Insert PCR condition

Forward primer: fdhF129FSD

Reverse primer: fdhF180Rev

- |         |                       |                            |                         |
|---------|-----------------------|----------------------------|-------------------------|
| Step 1: | Initial denaturation: | 98°C, 2 minutes 30 seconds | 1 cycle                 |
| Step 2: | Denaturation:         | 98°C, 1 minute             |                         |
| Step 3: | Primer annealing:     | 70°C, 30 seconds           |                         |
| Step 4: | Primer extension:     | 72°C, 5 minutes            | steps 2 to 4, 30 cycles |
| Step 5: | Final extension:      | 72°C, 10 minutes           | 1 cycle                 |

PCR product was observed on agarose gel and Ligation reaction was performed according to the manufacturer's protocol (USB rapid ligation, USA). Plasmid DNA was purified according to the manufacturer's (Qiagen plasmid mini prep). The *fdhF* insert was confirmed by DNA sequencing.

**2.5.3 PCR amplification of *fdhF* fragment (Gly129-Ala156) for *in vitro* transcription template**

Forward primer: M13/pUC Rev

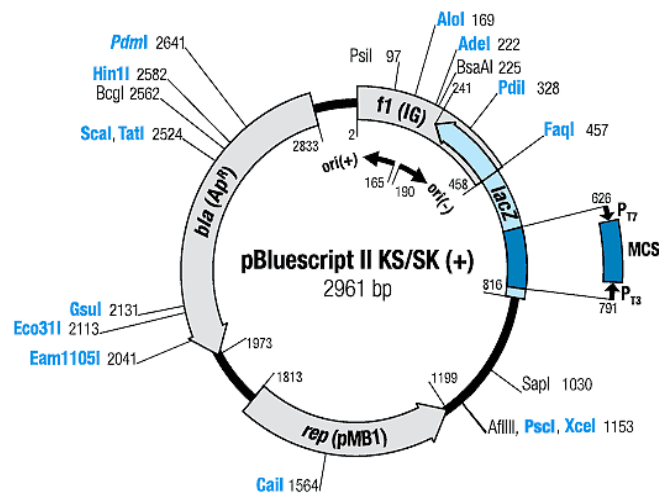
Reverse primer: FdhF.Rev156

The PCR was the same as above.

**2.5.4 Cloning of *fdhF* fragment (Phe129-Ala156) UGA<sub>140</sub> +SL into pBluescript II (SK+)**

The customized plasmid clone (Eurofins MWG Operon) pBluescript II (SK+) contain the *fdhF* fragment (Phe129-Ala156). The fragment contains strong Shine-Dalgarno (SD) sequence

followed by the ATG and TTC and natural coding sequence starts Thr 130 to Ala156. (Figure 21)



*fdhF* gene including restriction sites and strong SD sequence

5'-  
ACTCGAGGGCAAGGAGGTAATAATGTTACCAATAACGTTGACTGCTGCGCTCGTGTC<sup>TGA</sup>CACGG  
CCCATCGGTTGCAGGTCTGCACCAATCGGTCGGTAACGGCGCA<sup>GAATTC</sup>-3'

5'- XhoI 5'-CTCGAG-3'

3'- EcoRI 5'-GAATTC-3'

**Figure 21:** Scheme of pBluescript II (SK+) contains the *fdhF* fragment (Phe129-Ala156).

## 2.5.5 Quick-change site-directed mutagenesis of *fdhF* UGA<sub>140</sub> +SL into *fdhF* UUC<sub>140</sub> +SL by

### Quick-change PCR

Forward primer: fdhF.Frd-TGA-TTC

Reverse primer: fdhF.Rev-TGA-TTC

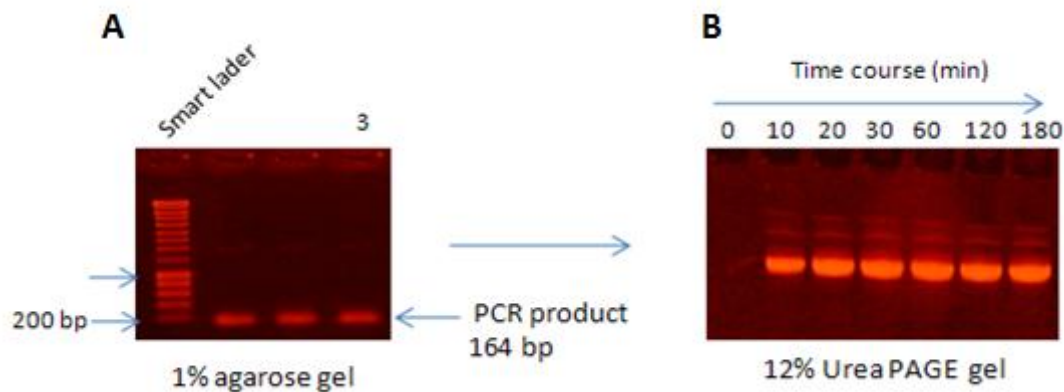
Step 1:	Initial denaturation:	98°C, 4 minutes	1 cycle
Step 2:	Denaturation:	98°C, 1 minute	
Step 3:	Primer annealing:	55°C, 1 minute	
Step 4:	Primer extension:	68°C, 6 minutes	steps 2 to 4, 18 cycles
Step 5:	Final extension:	68°C, 20 minutes	1 cycle

The amplified product was identified on an agarose gel and parental template DNA was digested with restriction enzyme DpnI (Fermentas). 5 µl of DNA was transformed into NovaBlue cells (Novagen) and plasmid DNA purified according manufacturer's (Qiagen plasmid mini prep). Mutation was confirmed by DNA sequencing.

### 2.5.6 *In vitro* transcription and purification of model *fdhF* mRNAs (*fdhF* SD)

#### *In vitro* transcription

*In vitro* transcription reaction was performed according to manufacturer's protocol (Fermentas). *fdhF* DNA was amplified by PCR as described previous. *In vitro* transcription was performed according to manufacturer's (Fermentas) and *in vitro* transcribed mRNA transcripts, i.e. *fdhF* mRNA UGA<sub>140</sub> and *fdhF* mRNA UUC<sub>140</sub> were identified on urea PAGE gel. (Figure 22A, B).



**Figure 22:** PCR amplification of template DNA (A) and *in vitro* transcription time course (B).

*In vitro* transcribed mRNA was subjected to phenol extraction followed by ethanol precipitation. Pellets were dissolved in loading buffer and loaded onto pre-equilibrated Nucleobond AX cartridge by gravitational flow. Washed was washed with washing buffer one column volume. Then mRNA was eluted with elution buffer three column volumes. This procedure was in house developed and its preparation procedure is different from the manufacturer's protocol. Gradient and elution profile was shown in figure 23. Purity of mRNA constructs was shown in figure 24.

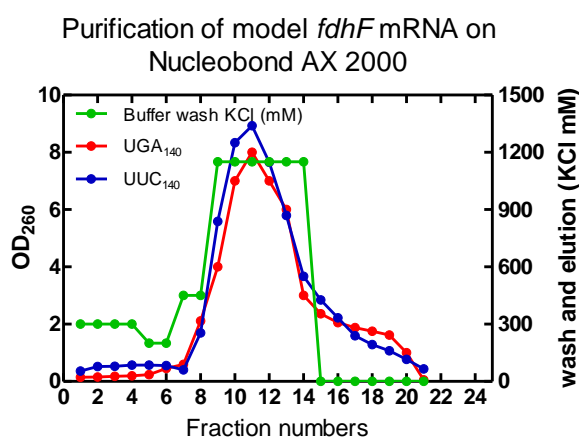
### Buffers

Equilibration buffer: 100 mM Tris/acetate, 300 mM KCl and 15% ethanol, pH 6.3

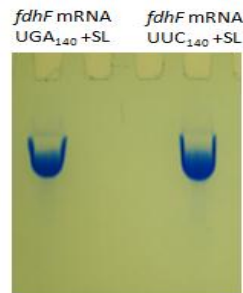
Loading buffer: 100 mM Tris /acetate, 200 mM KCl and 15% ethanol, pH 6.3

Wash buffer: 100 mM Tris /acetate, 450 mM KCl and 15% ethanol, pH 6.3

Elution buffer: 100 mM Tris /acetate, 1150 mM KCl and 15% ethanol, pH 6.3



**Figure 23:** Elution profile of model *fdhF* mRNAs (*fdhF* SD) UGA<sub>140</sub> and UUC<sub>140</sub>.

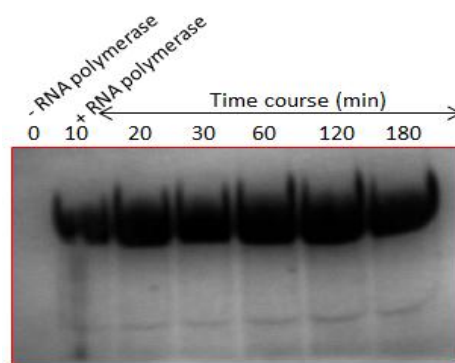


**Figure 24:** Purity of mRNA constructs on Urea PAGE gel.

### 2.5.7 *In vitro* transcription and purification of model *fdhF* mRNAs (strong SD)

#### *In vitro* transcription

*In vitro* transcription reaction was performed according to manufacturer's protocol (Fermentas). *In vitro* transcribed mRNA transcripts i.e. *fdhF* mRNA UGA<sub>140</sub> and *fdhF* mRNA UUC<sub>140</sub>, were identified on urea PAGE gel.



**Figure 25:** *In vitro* transcription.

#### Purification on MonoQ column

Buffer A: 20 mM Bis-Tris, 150 mM NaCl, 1 mM EDTA, pH 6.0

Buffer B: 20 mM Bis-Tris, 1000 mM NaCl, 1 mM EDTA, pH 6.0

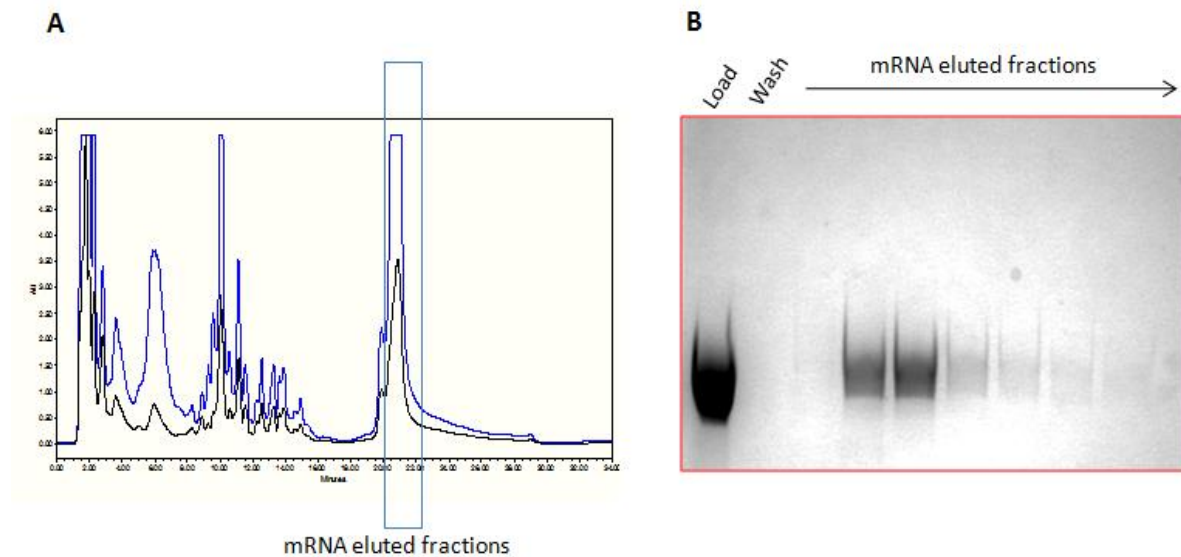
0 to 100% B 25 ml 25 minutes, Flow rate 1 ml/minute (Table 9).



**Table: 9** Gradient details.

Run time (min)	Gradient concentration (%)	
	A	B
0	100	0
5	100	0
25	0	100
27	0	100
28	100	0
35	100	0

Peak fractions were identified on HPLC chromatogram and urea PAGE gel (Figure 26A, B). Peak fractions were further subjected to ethanol precipitation at -20°C overnight. Recovered pellets were washed with 70% ethanol and then recovered pellet was dissolved in MilliQ water with 0.1 mM EDTA.

**Figure 26:** HPLC Elution profile (A) and purity of mRNA (UGA<sub>140</sub>) on Urea PAGE gel (B).

**2.5.8 Preparation and purification of initiation complexes**

Initiation complexes were prepared by incubating 70S ribosomes (1  $\mu$ M) with a 10-fold excess of *fdhF* mRNA; 1.5  $\mu$ M initiation factors 1, 2, and 3; 1.5  $\mu$ M f[ $^3$ H]Met-tRNA<sup>fMet</sup>; 1 mM GTP in TAKM<sub>7</sub> buffer (50 mM Tris-HCl, 70 mM NH<sub>4</sub>Cl, 30 mM KCl, 7 mM MgCl<sub>2</sub>, pH 7.5) at 37°C, 1 hour. Initiation complexes were purified by centrifugation through 400  $\mu$ l sucrose cushions (1.1 M sucrose in 1x TAKM<sub>7</sub>) at 260,000 X g for 2 hours (RC M120 GX ultracentrifuge, Sorvall). Pellets were dissolved in 1x TAKM<sub>7</sub> buffer to a final concentration of 5  $\mu$ M, stock frozen in liquid nitrogen, and stored at -80°C.

**2.5.9 Preparation and purification of aminoacyl-tRNAs**

300 A<sub>260</sub> units of total tRNA from *E. coli* MRE600 (Roche) was aminoacylated with the required 14C-labeled (25  $\mu$ M) and cold amino acids (0.3 mM); S100 (6%); ATP (3 mM) in buffer 1x TAKM<sub>7</sub> (pH 7.5) at 37°C, 30 minutes. Aminoacylation reaction was further subjected to phenol extraction and ethanol precipitation. Aminoacylated total tRNA was purified on MonoQ according to the standard protocol.

Buffer A: 50 mM NaAc, pH 4.5

Buffer B: 50 mM NaAc, 1.1 M NaCl, pH 4.5

Gradient: 0 to 100% B 45 ml, 45 minutes. Flow rate 1 ml/minute

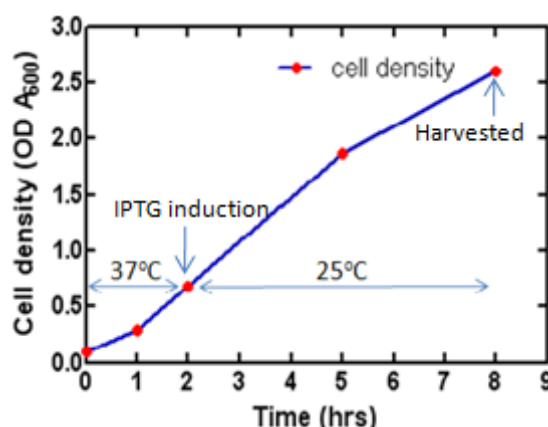
Peak fractions were identified on HPLC chromatogram and fractions were further subjected to ethanol precipitation at -20°C overnight. Recovered pellets were washed with 70% ethanol and the recovered pellet was dissolved in MilliQ water.

### 2.5.10 Expression and purification of wild type elongation factor SelB from *E. coli*

The expression and preparation procedure has been developed in such a way that SelB can be purified only in two chromatographic steps (Q-Sepharose followed by SP- Sepharose) in order to avoid precipitation during preparation.

#### Expression

Plasmid pWL 194 (harboring the *SelB* gene under control of a T7 promoter) was used in *E. coli* BL21 (DE3) cells for expression. A preculture (LB medium containing 100µg/ml ampicillin) was inoculated with a single colony and cells were grown over night at 37°C, 210 rpm. Before inoculation of the expression culture (LB medium containing 100µg/ml ampicillin), the preculture was centrifugated (4000 rpm, 15 min, 4°C) and the cells were resuspended in fresh LB medium. Cells were grown at 37°C, 210 rpm until ODA<sub>600</sub> reached 0.6 – 0.7, the culture was down to ~25°C, SelB expression was induced by the addition of IPTG to final 0.1 mM and cells growth was continued for another 6-8 hours and ODA<sub>600</sub> reached 2.0 – 2.5. (Figure 27). Cells were harvested at 4°C, frozen in liquid nitrogen and stored at -80°C until use.



**Figure 27:** Growth curve of *E. coli* BL21 (DE3), wild type SelB expression.

### Cell opening by Lysozyme treatment

Cells were resuspended in cold cell lysis buffer (50 mM Tris-HCl pH 8.0, 500 mM KCl, 10 mM MgCl<sub>2</sub>, 5 mM BME, 0.1 mg/ml PMSF and 7% glycerol) 1.8 ml buffer/g cell pellet. High concentrations of SelB tend to precipitate at low salt buffer, so it necessary to use 500 mM KCl in the cell lysis buffer. Cells were opened by the addition of lysozyme (lysozyme was dissolved in 10 mM Tris-HCl pH 8.0). 5 mg lysozyme/g cell pellet was optimal for cell opening. Cell opening was carried out on ice about 1 hour and then few crystals of DNase were added and lysis was continued for another 30 minutes. Cell debris was removed by centrifugation at 4°C, 18,500 rpm for 1 hour. The supernatant was collected (containing 500 mM KCl) and was diluted 7.5-fold with dilution buffer (50 mM Tris-HCl pH 8.0, 10 mM KCl, 10 mM MgCl<sub>2</sub>, 5 mM BME, 0.1 mg/ml PMSF and 7% glycerol) to reduce the KCl concentration to about 60 mM.

### Q-Sepharose

Supernatant was loaded onto a Q-Sepharose column (1.5 x 13 cm) which is pre-equilibrated with buffer A and the column was washed with buffer A until OD<sub>280</sub> reached base level (Figure 29A).

Buffer A: 50 mM Tris-HCl pH 8.0, 60 mM KCl, 10 mM MgCl<sub>2</sub>, 5 mM BME, 0.1 mg/ml PMSF, 7% glycerol

Buffer B: 50 mM Tris-HCl pH 8.0, 250 mM KCl, 10 mM MgCl<sub>2</sub>, 5 mM BME, 0.1 mg/ml PMSF, 7% glycerol.

0 to 100% B, flow rate 2 ml/minute

**Table: 10** Gradient details.

Run volume	Gradient concentration (%)	
	A	B
5 column volumes	0	100

SelB eluted at ~150-200 mM KCl and SelB containing fractions (Figure 28B) were pooled and buffer exchange to the buffer (50 mM Tris-HCl pH 8.0, 400 mM KCl, 10 mM MgCl<sub>2</sub>, 5 mM BME, 0.1 mg/ml PMSF, 7% glycerol) by using Vivaspin 30 (30 kDa cutoff) concentrators (Sartorius).

### SP-Sepharose

Buffer exchanged SelB fractions were loaded onto a SP-Sepharose (1.5 x 11 cm) column which was pre-equilibrated with buffer A. The column was washed buffer A until OD<sub>280</sub> reaches base level and the gradient was started (Figure 29B).

Buffer A: 50 mM K HEPES pH 6.8, 100 mM KCl, 10 mM MgCl<sub>2</sub>, 5 mM BME, 7% glycerol

Buffer B: 50 mM K HEPES pH 6.8, 400 mM KCl, 10 mM MgCl<sub>2</sub>, 5 mM BME, 7% glycerol

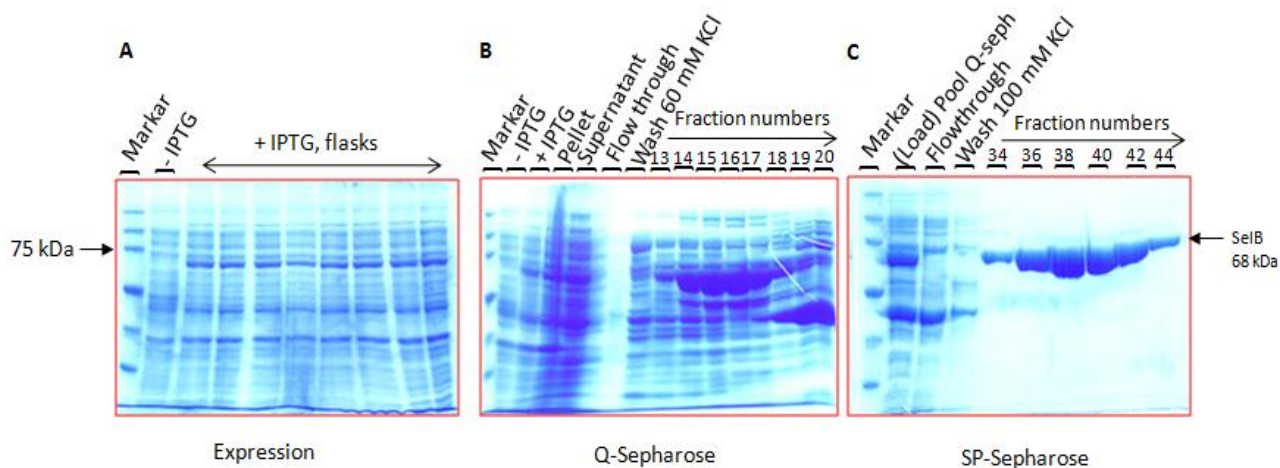
0 to 100% B, flow rate 2 ml/minute (Table 11).

**Table: 11** Gradient details.

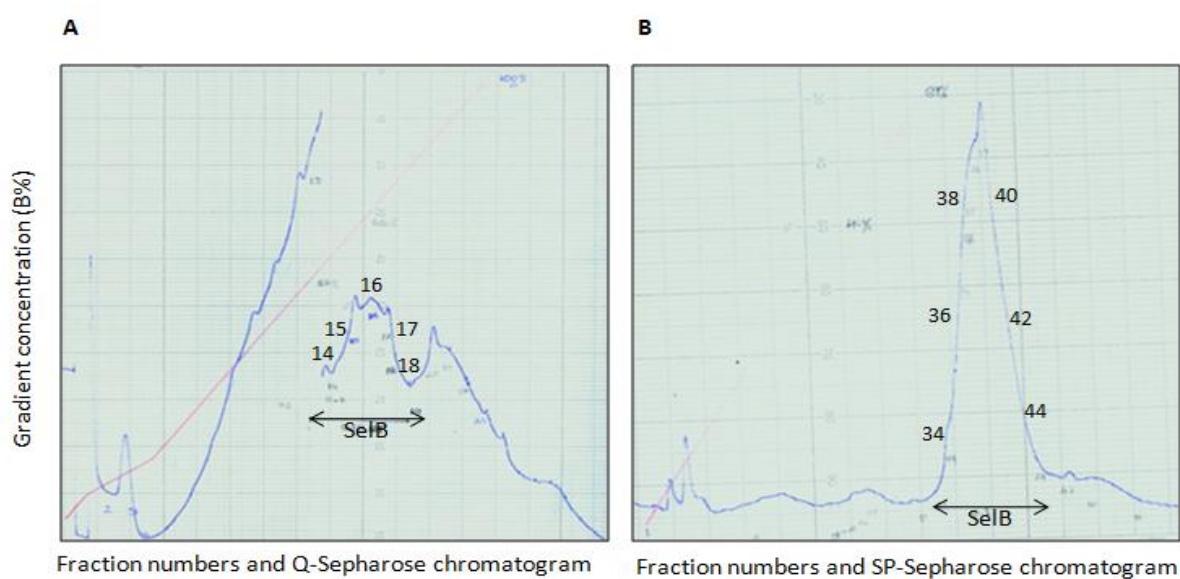
Run volume	Gradient concentration (%)	
	A	B
5 column volumes	0	100

SelB eluted at ~350-450 mM KCl and purity was identified on SDS PAGE (Figure 28C). SelB containing fractions were pooled and subjected to the final buffer exchange (50 mM K HEPES

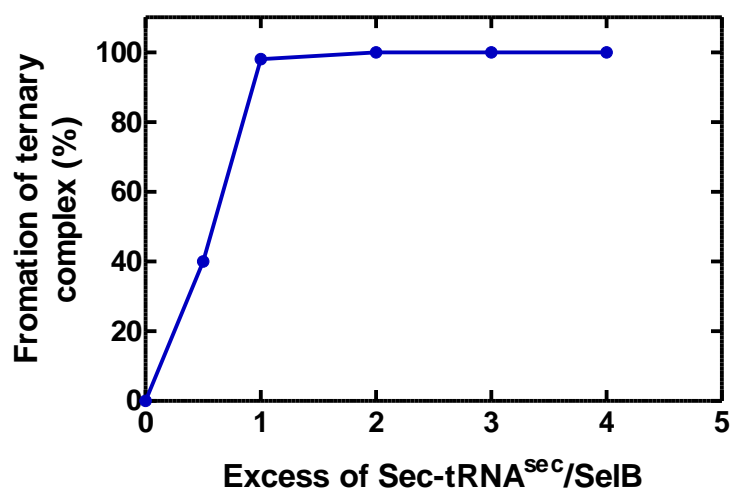
pH 7.5, 400 mM KCl, 10 mM MgCl<sub>2</sub>, 5 mM BME and 7% glycerol) for storage. Final glycerol concentration was 22% to avoid precipitation during storage. Aliquots were frozen in liquid nitrogen and stored at -80°C. The SelB concentration was measured spectroscopically at 280 nm ( $\epsilon_{\text{SelB}}$  81,080 M<sup>-1</sup> cm<sup>-1</sup>). SelB activity with respect to formation of ternary complex was measured (Figure 30).



**Figure 28:** Expression (A), purification of SelB; Q-Sepharose (B), SP-Sepahrose (C).



**Figure 29:** Elution profile of SelB on FPLC chromatogram. Q-Sepharose (A), SP-Sepahrose (B).



**Figure 30:** SelB activity with respect to formation of ternary complex.

#### 2.5.11 Preparation of SelB·GTP·Sec-tRNA<sup>Sec</sup>

SelB (5  $\mu$ M) was incubated with DTT (10 mM) and GTP (3 mM), in 1x TAKM<sub>3.5</sub>P<sub>8</sub> buffer (50 mM Tris-HCl, 70 mM NH<sub>4</sub>Cl, 30 mM KCl, 3.5 mM MgCl<sub>2</sub>, 8 mM putrescine, pH 7.5 at 37°C for 5 minutes. An equal amount of [<sup>3</sup>H]Sec-tRNA<sup>Sec</sup> (5  $\mu$ M) was added to the SelB binary complex and incubated at room temperature for 5 minutes.

#### 2.5.12 Preparation of EF-Tu·GTP·aminoacyl-tRNA

EF-Tu (200  $\mu$ M) was incubated with pyruvate kinase (0.1 mg/ml), phosphoenol pyruvate (3 mM) and GTP (1 mM), in 1x TAKM<sub>3.5</sub>P<sub>8</sub> buffer at 37°C for 15 minutes. Total aminoacyl-tRNA was added to the EF-Tu binary complex and incubated at 37°C for 1 minute.

### 2.5.13 *In vitro* translation

*In vitro* translation was initiated by the addition of EF-G (0.05  $\mu$ M), EF-Tu and SelB ternary complex premix in the presence and absence of RF2 to the initiated ribosomes. Translation reaction was carried out at 37°C for 1 hour.

### 2.5.14 HPLC separation of dipeptide (fMetGly)

Dipeptide was separated on Chromolith column and the gradient as follows (Table 12).

Buffer A 0.1% TFA, Buffer B 0.1% TFA, 65% acetonitril

**Table: 12** Gradient details.

Run time (min)	Gradient concentration (%)	
	A	B
2	100	0
6	97	3
10	95	5
16	92.5	7.5
20	87.5	12.5
26	70	30



### 2.5.15 HPLC separation of dipeptide (fMetPhe)

Column and buffers are same as above, flow rate 2 ml/min (Table 13).

**Table: 13** Gradient details.

Run time (min)	Gradient concentration (%)	
	A	B
0	100	0
1	100	0
0	90	10
3	85	15
4	80	20
5	70	30
6	60	40
7	50	50
8	25.4	74.6
9	19.2	80.8
10	0	100
14	100	0

### 2.5.16 HPLC separation of tripeptide (fMetPheThr) and polypeptide up to 1<sup>st</sup> Val

Tripeptide and polypeptide up to 1<sup>st</sup> Val was separated on Chromolith column and the gradient as follows (Table 14). Buffer A and B are same as above.

**Table: 14** Gradient details.

Run time (min)	Gradient concentration (%)	
	A	B
0	100	0
30	68	32
32	50	50
35	25	75
37	0	100
38	0	100
39	100	0
42	100	0

**2.5.17 HPLC separation polypeptide up to 2<sup>nd</sup> Val**

Buffer A and B same as above.

Polypeptide up to 2<sup>nd</sup> Val was separated on RP-8 column and the gradient as follows (Table 15).

**Table: 15** Gradient details.

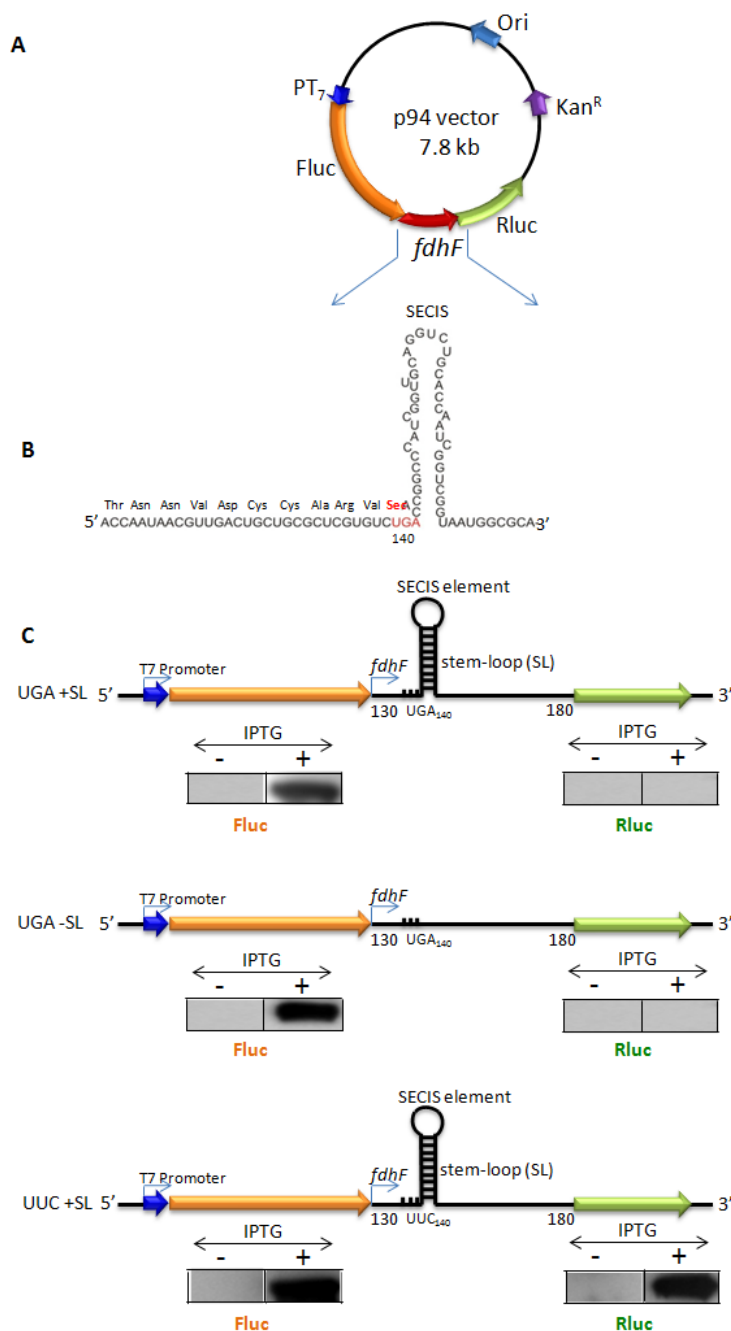
Run time (min)	Gradient concentration (%)	
	A	B
0	100	0
15	85	15
25	55	45
30	45	55
35	30	70
40	15	85
42	0	100
45	100	0

### 3 Results

#### 3.1 Efficiency of Sec incorporation in proteins *in vivo*

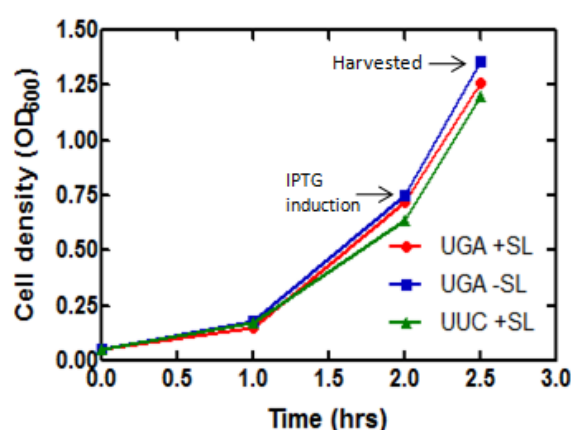
##### 3.1.1 Efficiency of UGA recoding into selenocysteine by Sec-tRNA<sup>Sec</sup>

In order to determine the efficiency of UGA recoding into selenocysteine *in vivo*, a plasmid-encoded reporter was constructed in the following way. The plasmid (Table 8) contained the gene for the Fluc controlled by an IPTG inducible T7 promoter, followed by the gene coding for Rluc. A fragment of wild-type *fdhF* (UGA<sub>140</sub>) followed by the SECIS (130 to 179) was inserted in between two reporter genes (Figure 31). Translation of the first gene was independent of Sec incorporation and switched on by the addition of IPTG. Translation of the second gene depended on the reading of the UGA<sub>140</sub> codon by SelB-Sec-tRNA<sup>Sec</sup>. To test the misreading efficiency in the UGA codon, which would result from a SelB-independent translation of the second reporter, stem-loop SECIS-element was deleted to abolish SelB binding. The maximum level of Rluc synthesis was determined by the expression from the control plasmid containing UUC codon (Phe) in the place of UGA<sub>140</sub> codon. We validated the experiments based on the quantification of translated products (Fluc and Rluc) by western blotting using the specific antibodies against firefly luciferase (Fluc) and renilla luciferase (Rluc). Firefly and renilla luciferase enzyme activity was measured. For the initial experiments western blots were used to validate the assay



**Figure 31:** Constructs for dual-reporter gene assay. (A) Vector p94 contains a fragment of the *fdhF* gene between a firefly (Fluc) and a renilla luciferase (Rluc) gene. Protein expression can be induced by IPTG addition. (B) Wild type fragment of the *fdhF* gene with a UGA codon and the SECIS element. (C) Test constructs: wild type fragment (UGA + SL), fragment without the SECIS element (UGA – SL) and fragment with replacement of the UGA codon (stop codon) with a UUC codon (Phe) (UUC + SL). Expression of the fluc and rLuc part of the fusion protein was shown by western blot analysis.

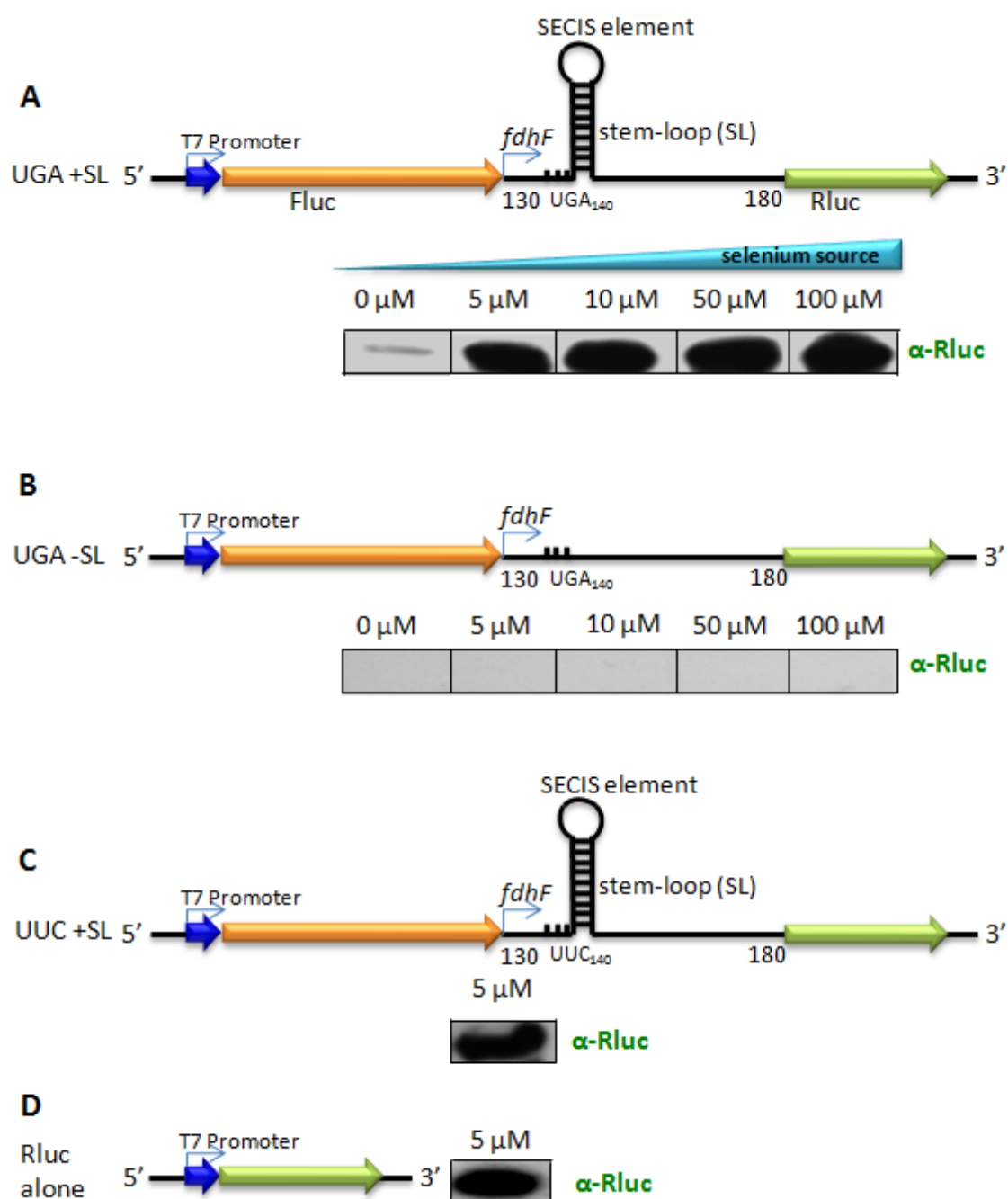
In the initial set of experiments, cells were grown at aerobic conditions in LB media at 37°C, induced by 1 mM IPTG at initial log phase ( $OD_{600} \sim 0.6-0.7$ ), and harvested 30 minutes after induction (Figure 32). Cells were opened by lysozyme treatment (Materials and Methods) and the expression of Fluc-fdhF-Rluc fusion protein was identified by western blotting with  $\alpha$ -Fluc and  $\alpha$ -Rluc antibodies (Materials and Methods). Western blotting analysis showed that there was no synthesis of second protein (Rluc) with constructs containing the wild type *fdhF* construct as well as the *fdhF* construct lacking the SECIS stem-loop (Figure 31) which implies that there was no UGA readthrough. In contrast, there was clear synthesis of Rluc when the UUC codon was replaced by UGA codon. While the low misreading level of the construct lacking the SECIS was expected, given the low decoding error frequency in bacteria (Drummond and Wilke, 2009), the inability of the Sec machinery to read the stop codon suggested that there might be a requirement of additional selenium source at rapid growth phase of cells for the internal synthesis of Sec-tRNA<sup>Sec</sup>.



**Figure 32:** Growth curves of *E. coli* transformed with test constructs. Times of protein expression induction (initial-log phase) and harvest are indicated. Cultures were grown in LB medium at 37°C.

### 3.1.2 Selenium levels determine the efficiency of UGA recoding by Sec-tRNA<sup>Sec</sup> into selenocysteine

In order to understand the requirement for selenium in rapid growing cells, we performed series of experiments where we studied selenium dependence of Sec incorporation. Cells were grown in TPG media (Materials and Methods) and induced by 1 mM IPTG for 30 minutes. Western blotting analysis with  $\alpha$ -Rluc showed that with the wild-type *fdhF* mRNA (UGA<sub>140</sub>) +SECIS there is a clear UGA readthrough upon increasing selenium levels (Figure 33 A). With the construct lacking the SECIS-element there is no detectable misreading of UGA (Figure 33B), whereas the UUC readthrough of *fdhF* mRNA (UUC<sub>140</sub>) +SECIS and the expression of Rluc alone did not depend of selenium levels (Figure 33C, D). These results strongly suggest that Sec incorporation into proteins in rapidly growing cells requires an additional selenium source.

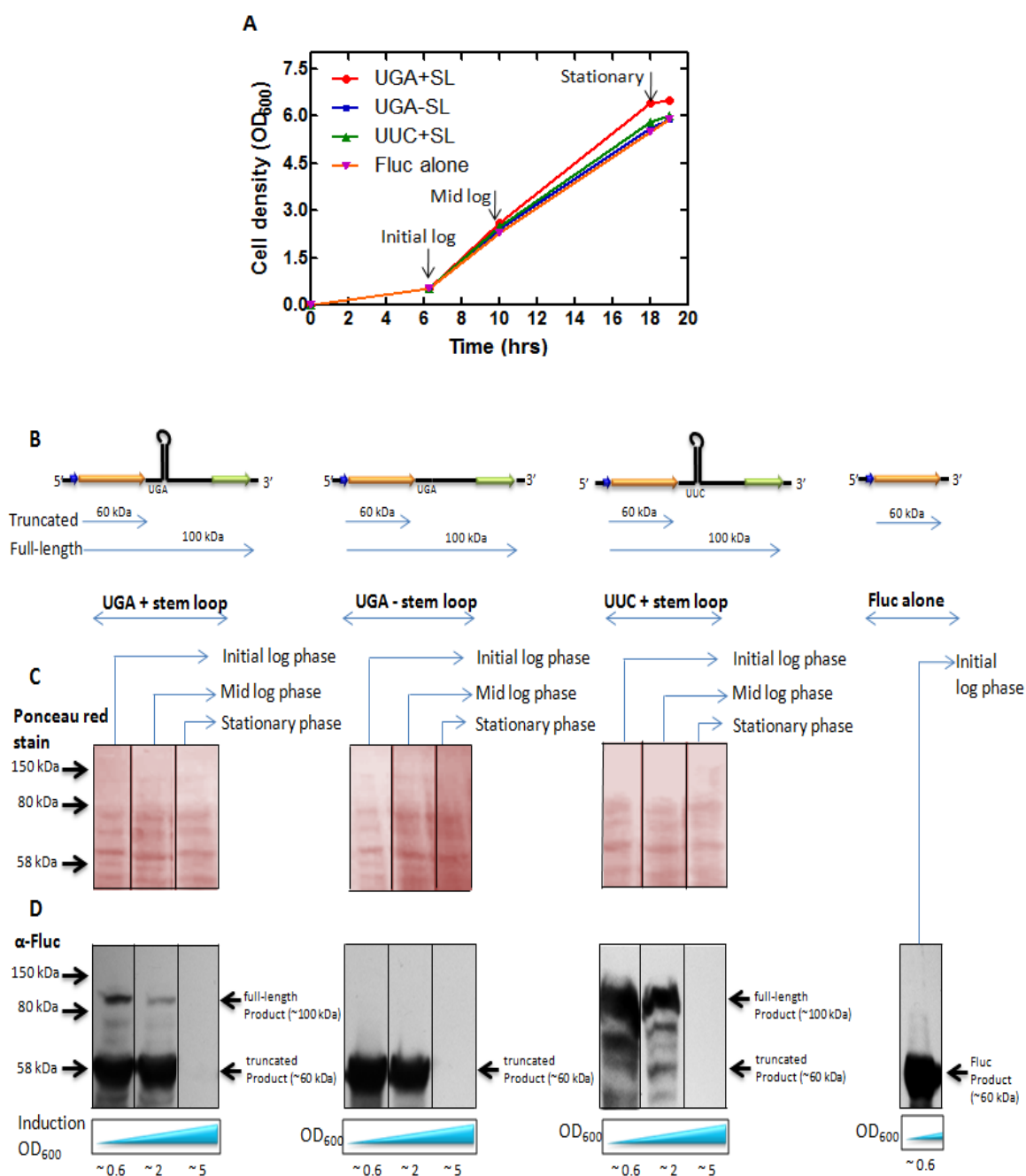


**Figure 33:** Selenium dependence of UGA recoding. For test constructs UGA + SL (A) and UGA – SL (B) UGA recoding was monitored at different concentrations of a selenium source (sodium selenite). Recoding was detected by western blotting using an anti-rluc antibody. Construct UUC + SL (C) and a construct with rluc alone (D) were used as controls at 5 μM selenium source. Cultures were grown in TPG medium and samples were taken at initial log phase as described.

### 3.1.3 UGA recoding efficiency at various phases of cell growth

Earlier reports suggested differences in the efficiency of Sec incorporation in cells at different phases of cell growth i.e. initial-log, mid-log and stationary phase. To test this notion, cells were grown in TPG media containing 50  $\mu$ M selenium source (Materials and Methods) and induced by 1 mM IPTG at 37°C at various phases of cell growth (Figure 34A). Synthesis of the full-length fusion and truncated proteins was analyzed by western blotting with  $\alpha$ -Rluc (antibody against Rluc) antibody (Figure 34B). Stem-loop deleted *fdhF* mRNA (UGA<sub>140</sub>) truncated protein was analyzed by western blotting with  $\alpha$ -firefly (anti-fluc antibody) antibody (Figure 34C). Firefly-fdhF-renilla fusion of full-length protein of *fdhF* mRNA (UUC<sub>140</sub>) +SECIS was analyzed by western blotting with  $\alpha$ -firefly (Figure 34D) and firefly alone as an internal control marker for the product size was analyzed by western blotting with  $\alpha$ -firefly antibody (Figure 34). The results showed that there is difference in the Sec-containing protein production levels at various phases of cell growth. Notably, we were not able to detect either full-length or truncated product at stationary phase. These results prompted us to test further the levels of protein production at each phase.

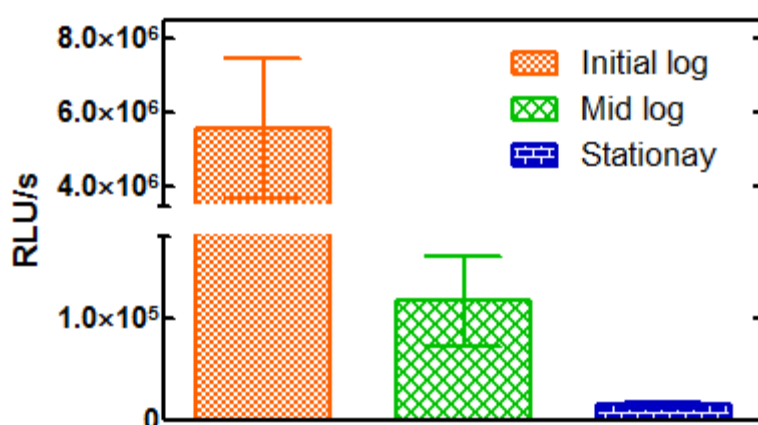




**Figure 34:** UGA recoding at various phases of cell growth. (A) Cells transformed with reporter constructs were grown in TPG medium supplemented with 50  $\mu$ M sodium selenite at 37°C. At indicated times samples were taken for protein expression induced by addition of IPTG, cells were harvested 30 min after induced. (B) Test constructs and expected product sizes. (C) Proteins in cell lysate were separated by SDS-PAGE and blotted. Ponceau red-stained membranes are shown. (D) Detection of expression products by western blotting with anti-fluc antibody.

### 3.1.4 Determination of protein production levels at various phases of cell growth

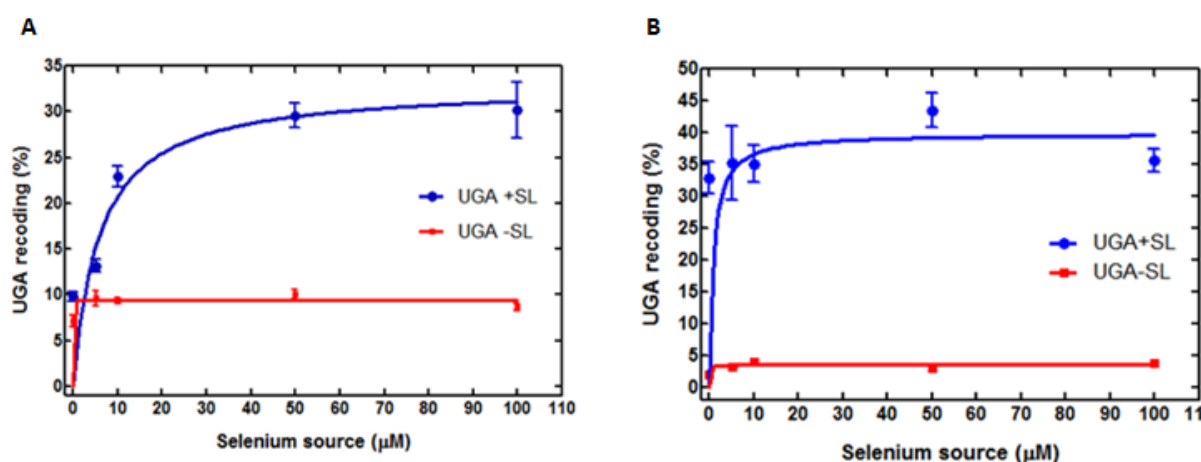
To compare the protein production levels in cells at different phases of cell growth, we have taken wild-type *fdhF* mRNA (UGA<sub>140</sub>) +SECIS as an example (Figure 34B). Cells were grown in TPG media, induced by 1 mM IPTG at various phases of cell growth and harvested after 30 minutes of induction (Figure 34A). The Luciferase activities (Relative Light Units- RLU/s) of the first protein in the reporter construct were measured (Materials and Methods) in samples collected from initial-log phase, mid-log phase and stationary phase. These results show that there is dramatic difference in protein production, suggesting a global shut-down of protein synthesis at the stationary phase (Figure 35).



**Figure 35:** Expression levels of fluc protein (UGA + SL) at different phases of cell growth. The activity of fluc as an indicator for protein expression was measured in a luciferase assay. Plotted are relative fluc units. Samples were taken in the same experiment as described in Figure 4A.

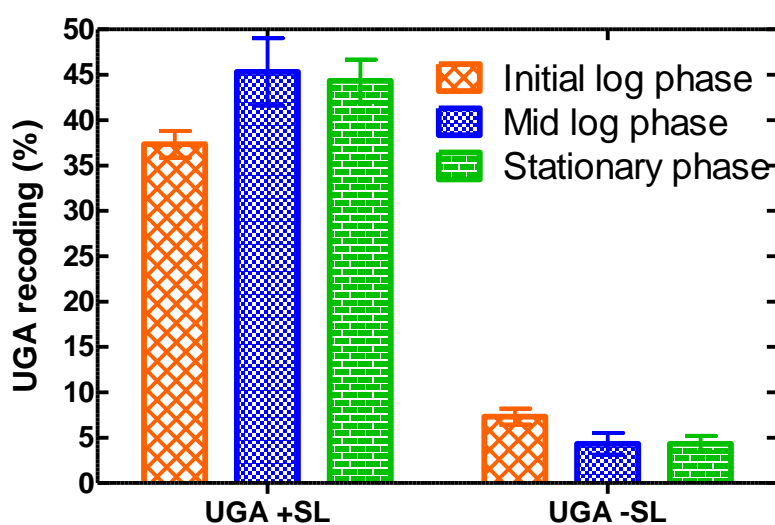
### 3.1.5 Selenocysteine insertion efficiency does not depend on high selenium levels under normal physiological growth conditions of *E. coli*

To test the selenocysteine insertion efficiency we conducted further experiments of selenium dependence selenocysteine insertion at initial log phase (rapid growth phase) and stationary phase (very poor growth phase). Induction and Fluc and Rluc measurements are same as above. These results of selenocysteine insertion at initial log phase clearly showed that there is increase in Sec insertion up 35% up on increasing selenium levels (with the apparent affinity of 6  $\mu\text{M}$ ). In comparison, misreading of UGA in the construct lacking the SECIS element was 5-10% (Figure 36A). The increase of UGA readthrough on wild type UGA +SL construct with selenium is consistent with western blot analysis (Figure 33A). Sec insertion does not depend on selenium levels at stationary phase and Sec insertion up to 40%, whereas the misreading is about 2-5% (Figure 36B).



**Figure 36:** Selenium-dependent Sec insertion at initial log phase (A) and stationary phase (B). Sec insertion is calculated from fluc and rluc activities in samples (Materials and Methods).

At any stage of cell growth, Sec is incorporated into proteins less efficiently (35-45%) than standard amino acids. At the same time, 5-10% UGA readthrough is due to misreading (Figure 37). It is believed that low efficiency of Sec insertion due to the competition with RF2 at UGA recoding site. The events and factors affecting the Sec insertion are discussed in the next chapter 3.2.3.



**Figure 37:** Efficiency of selenocysteine insertion at various phases of cell growth.

### 3.2 Design of the *in vitro* Sec incorporation system

#### 3.2.1 Design of model mRNA constructs.

*E. coli* produces three types of selenoproteins which contain selenocysteine in the enzyme active site (Figure 38A). In all three cases, a stop codon which is recoded as Sec is followed by a SECIS element.

**A**

##### *fdhF* mRNA

<----- SECIS stem-loop structure----->

Gly Thr Asn Asn Val Asp Cys Cys Ala Arg Val His Gly Pro Ser Val Ala Gly Leu His Gln Ser Val Gly Asn Gly Ala  
5'-GGU ACC AAU AAC GUU GAC UGC UGC GCU CGU GUC UGA CAC GGC CCA UCG GUU GCA GGU CUG CAC CAA UCG GUC GGT AAU GGC GCA-3'  
129 140 156

##### *fdnG* mRNA

Gly Met Leu Ala Val Asp Asn Glu Ala Arg Val His Gly Pro Thr Val Ala Ser Leu Ala Pro Thr Phe Gly Arg Gly Ala  
5'-GGG AUG CUG GCG GUA GAC AAC CAG GCG CGC GUC UGA CAC GGA CCA ACG GUA GCA AGU CUU GCU CCA ACA UUU GGU CGC GGU GCG-3'  
198 208 224

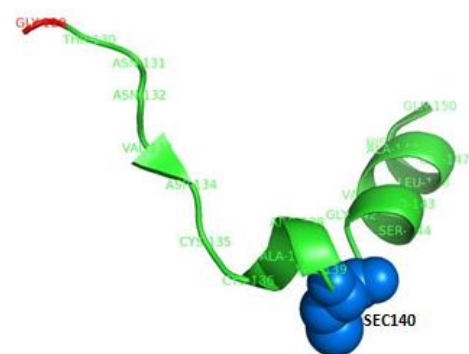
##### *fdoG* mRNA

Gly Met Leu Ala Val Asp Asn Glu Ala Arg Val His Gly Pro Thr Val Ala Ser Leu Ala Pro Thr Phe Gly Arg Gly Ala  
5'-GGU AUG CUC GCG GUC GAC AAC CAG GCG CGU GUC UGA CAC GGA CCA ACG GUA GCA AGU CUU GCU CCA ACA UUU GGU CGC GGU GCG-3'  
182 192 208

**B**

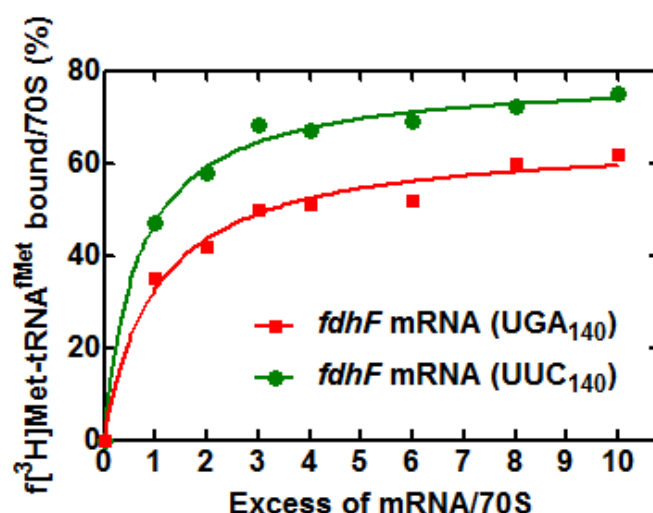
*fdhF* mRNA 5'-GGGAUUCUUAACUGGAGCGAGACCGAUGGGU ACC AAU AAC GUU GAC UGC UGC GCU CGU GUC UGA CAC GGC CCA UCG GUU GCA GGU CUG  
SD sequence spacer 129 140  
His Gln Ser Val Gly Asn Gly Ala  
CAC CAA UCG GUC GGU AAU GGC GCA-3'  
156

**C**



**Figure 38:** Choosing *E. coli* model *fdhF* mRNAs for *in vitro* translation. (A) Sequence of fragments of *E. coli* formate dehydrogenases *fdhF*, *fdnG* and *fdoG* containing Sec insertion site. (B) Design of *E. coli* model *fdhF* mRNA containing natural *fdhF* Shine-Dalgarno sequence followed by a Sec insertion fragment of natural coding sequence. (C) Stretch of fragment of formate dehydrogenase-H. SECIS identity elements are marked in blue.

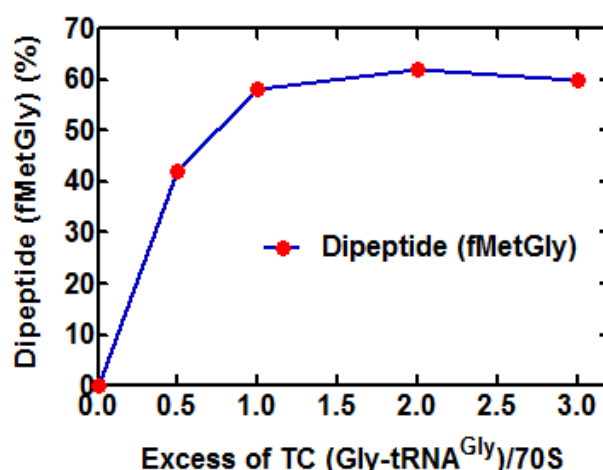
As the sequences of SECIS elements are very similar among three formate dehydrogenase, we have chosen a fragment of *fdhF* mRNA as a model mRNA for the *in vitro* study of Sec insertion. We have designed a model *fdhF* mRNA construct the following way. Natural *fdhF* Shine-Dalgarno sequence followed by a spacer and natural coding sequence positions 129 and 156 were included and with the Sec to be inserted at 140 position (Figure 38A, B, C). We have created control construct where the UGA<sub>140</sub> codon replaced with the UUC Phe codon. The fragment was chosen in such a way that formation of the ribosome initiation complex would not interfere with the SECIS-element. When the initiation codon is positioned in the P site of the ribosome, mRNA nucleotides up to +12-+14 are placed in the mRNA channel of the ribosome.



**Figure 39:** Testing translation initiation of model *fdhF* mRNA constructs.

*In vitro* transcribed and purified mRNAs were tested for their activity in the translation initiation in the presence of ribosomes, initiator tRNA and initiation factors (Materials and Methods). The binding of  $f[^3\text{H}]\text{Met-tRNA}^{\text{fMet}}$  to the ribosome was monitored by nitrocellulose filter binding assay at increasing mRNA concentrations (Materials and

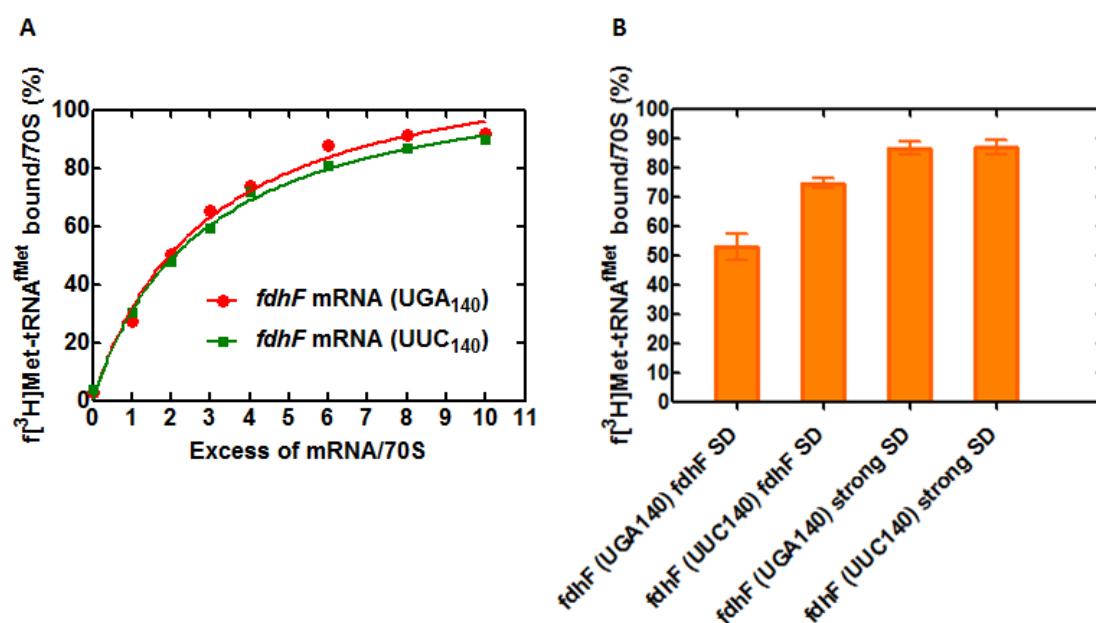
Methods). The ribosomes, which are well-characterized in routine assays, are close to 100% active in translation initiation in the presence of other model mRNAs (data not shown). The maximum binding of  $f[^3\text{H}]\text{Met-tRNA}^{\text{fMet}}$  to the ribosome in the presence of wild-type *fdhF* mRNA ( $\text{UGA}_{140}$ ) fragment or mutant *fdhF* mRNA ( $\text{UUC}_{140}$ ) was 60% and 70%, respectively (Figure 39).



**Figure 40:** Formation of the dipeptide (fMetGly). TC (Gly-tRNA<sup>Gly</sup>), ternary complex EF-Tu·GTP·Gly-tRNA<sup>Gly</sup>.

To further characterize the functional activity of the model mRNAs, we tested the efficiency of formation of the first dipeptide (fMetGly). The reaction was performed at increasing concentrations of ternary complex EF-Tu·GTP·Gly-tRNA<sup>Gly</sup> and fixed concentration of ribosome initiation complexes and incubated for 30 min at 37°C, which should ensure that the reaction proceeds to completion (Wohlgemuth et al., 2008). The analysis of dipeptides by HPLC showed that the maximum formation of fMetGly was 60% at about stoichiometric amounts of the ternary complex and initiation complex (Figure 40). Note that the amount of the formed initiation complex, rather than of total amount of added ribosomes, was taken into account in the experiments described above. Recalculating the values for the total

ribosomes gives only 36% activity (60% initiation x 60% dipeptide formation), which is relatively low concerning the high inherent activity of the ribosomes observed with other model mRNAs. This prompted us to modify the mRNA constructs such that the efficiency of initiation is improved.

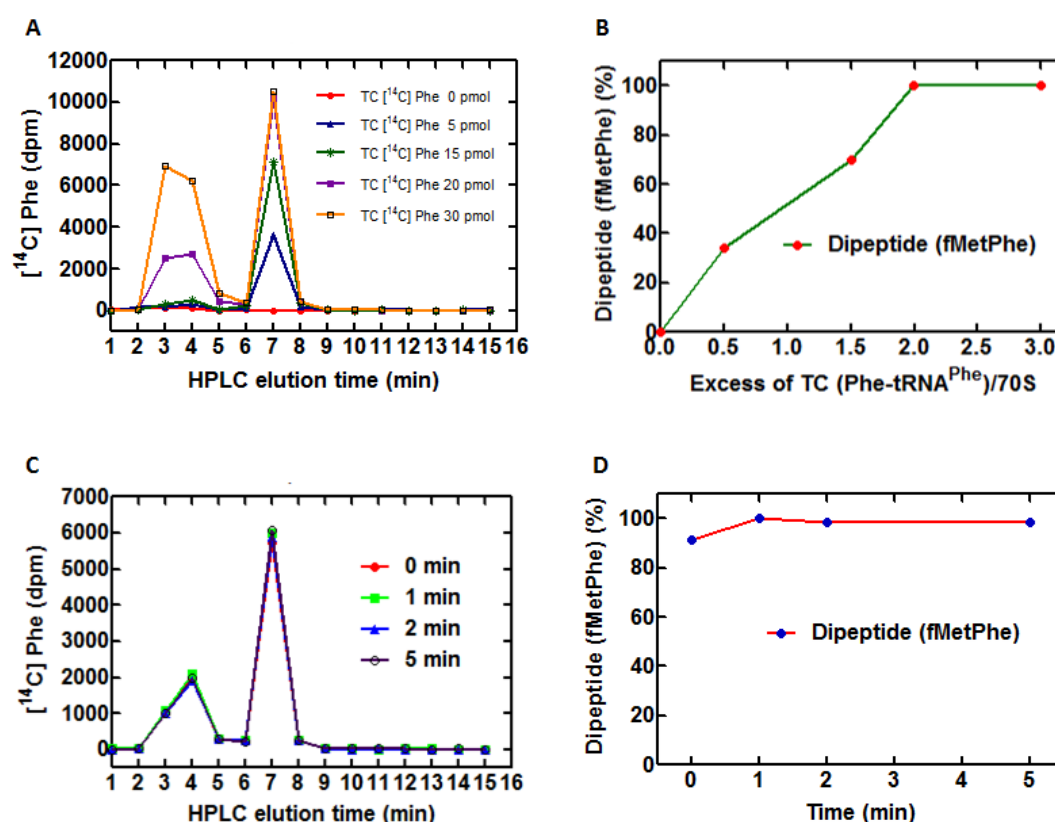


**Figure 41:** Testing translation initiation of model *fdhF* mRNA constructs with a strong Shine-Dalgarno sequence.

The most obvious way to improve initiation was to enhance the ribosome binding site of the model mRNA. We replaced the natural Shine-Dalgarno sequence of *fdhF* mRNA with the strong Shine-Dalgarno sequence followed by a spacer; we also replaced the Gly at the 2nd position (which turned out to be a rare codon) with Phe codon UUC for better reading of the 2nd codon and efficient dipeptide (fMetPhe) formation. The translation initiation was performed as previously described. Upon titration of wild-type *fdhF* mRNA (UGA<sub>140</sub>) and mutant *fdhF* mRNA (UUC<sub>140</sub>) the same initiation efficiency of up to 90% was achieved (Figure

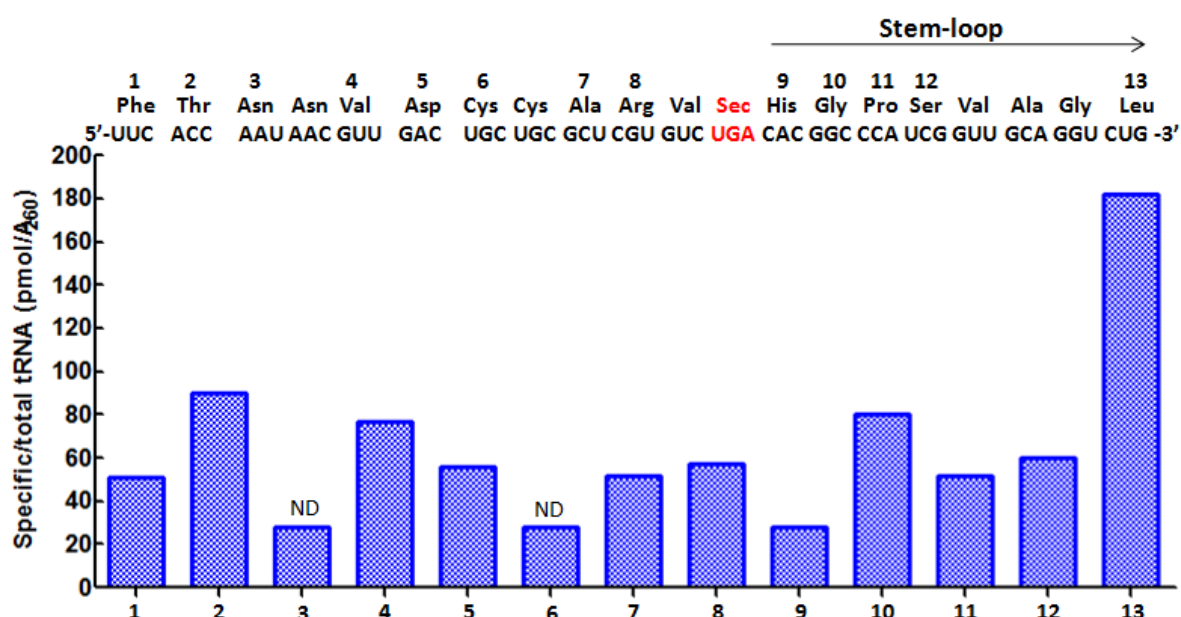


41A). The comparison of translation efficiencies with all model *fdhF* mRNA constructs is shown in Figure 41B.



**Figure 42:** Formation of dipeptide (fMetPhe).

To test the dipeptide (fMetPhe) formation, titrations with the ternary complex EF-Tu·GTP·Phe-tRNA<sup>Phe</sup> were performed with fixed concentration of initiated ribosomes and the dipeptide peak was analyzed by HPLC (Figure 42A). The maximum yield of dipeptide formation achieved with a 2-fold excess of ternary complexes over the initiation complex and was close to 100% (Figure 42B). The dipeptide formation was complete in less than one minute (Figure 42C, D), in agreement with the notion that the reaction is very fast (Gromadski and Rodnina, 2004).

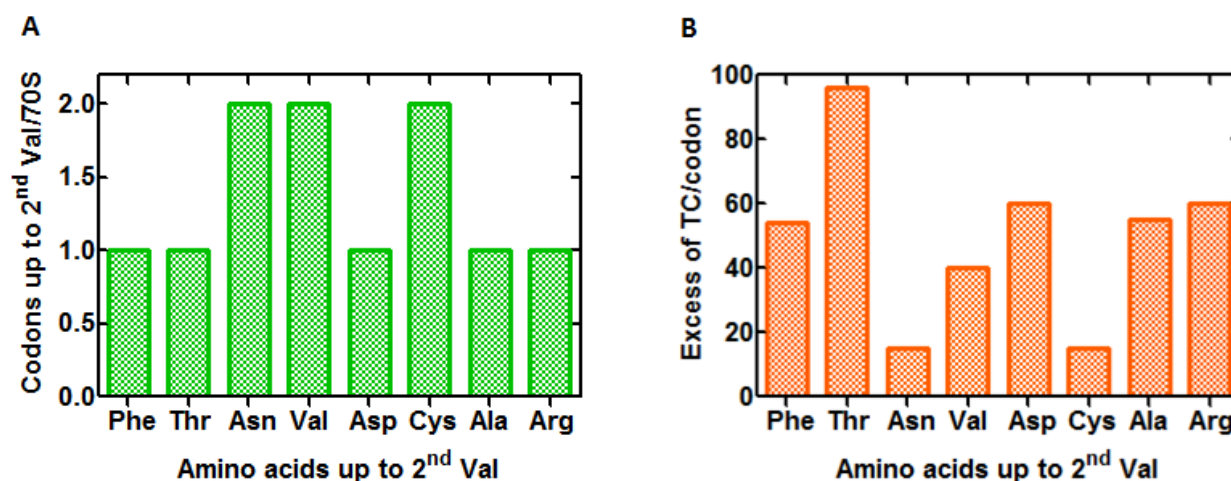


**Figure 43:** Estimation of individual tRNAs content in *E. coli* total tRNA (MRE 600). Upper panel, sequence of the *fdhF* mRNA (UGA<sub>140</sub>) construct that has to be translated. (ND) we assumed that the concentration of these tRNAs, which could not be measured directly, was similar to that of the lowest amount among tRNAs measured.

### 3.2.2 Optimization of translation of the mRNA sequence preceding the Sec codon

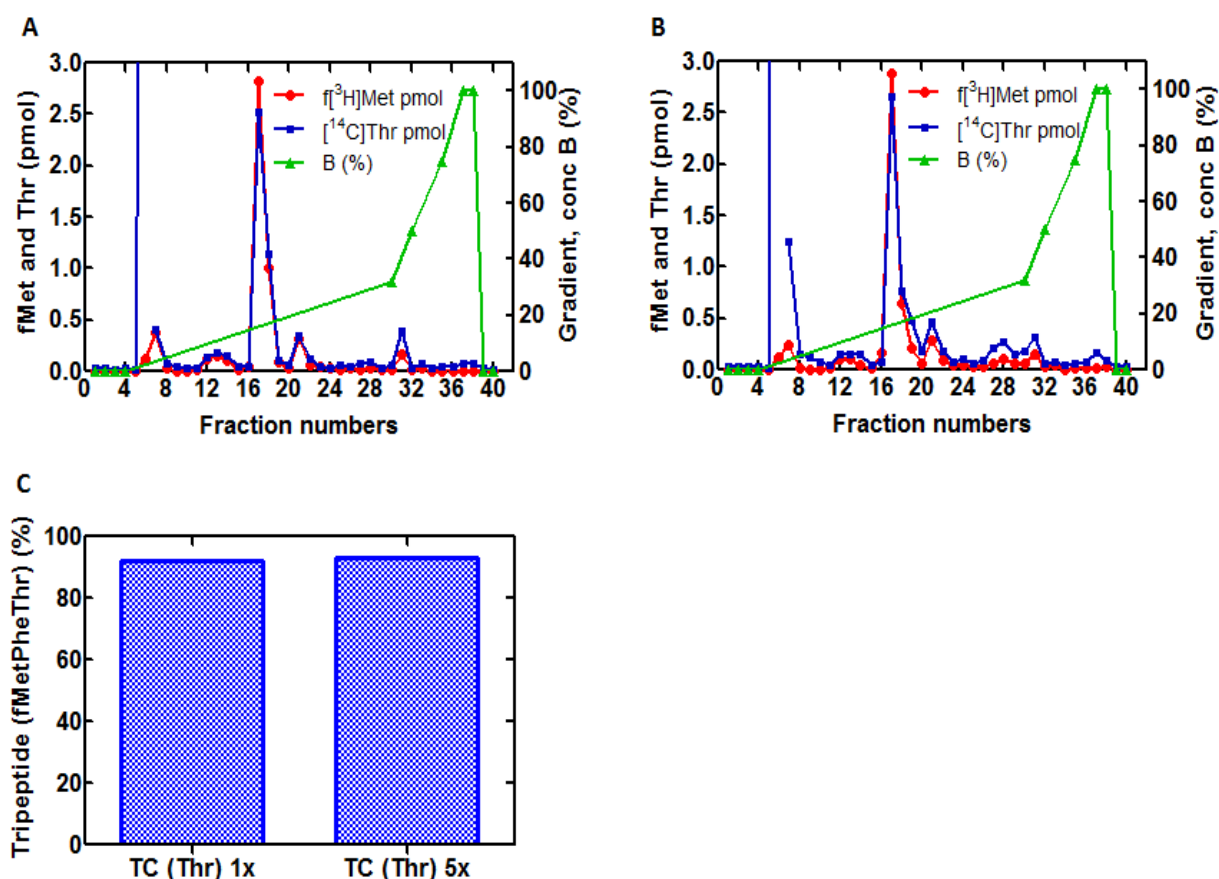
Before the recoding of the UGA codon by Sec, the ribosome has to translate a stretch of the *fdhF* mRNA with an amino acid sequence as indicated in Figure 43, upper panel. In order to prepare these ribosome nascent-chain complexes, we tested the amounts of the respective individual tRNAs in *E. coli* total tRNA (MRE600, Roche). Aminoacylation and TCA precipitation procedure is described in Materials and Methods using respective radioactively labeled amino acids. The content of each tRNA in the total tRNA pool is given in pmol/A<sub>260</sub> unit. For a 100% pure tRNA<sup>Phe</sup>, this value is 1750 pmol/A<sub>260</sub>, as calculated from its molecular weight and the ratio of single and double stranded regions. Thus, for a specific tRNA whose content in the total tRNA is 10%, 175 pmol/A<sub>260</sub> unit is expected. The content of different tRNAs varied between about 2% (tRNA<sup>His</sup>) and 10% (highly abundant tRNA<sup>Leu</sup>). For the two

tRNAs, aminoacylation of which could not be tested because the respective radioactive amino acids are not available, we assumed the same concentration as that of the least abundant tRNA.



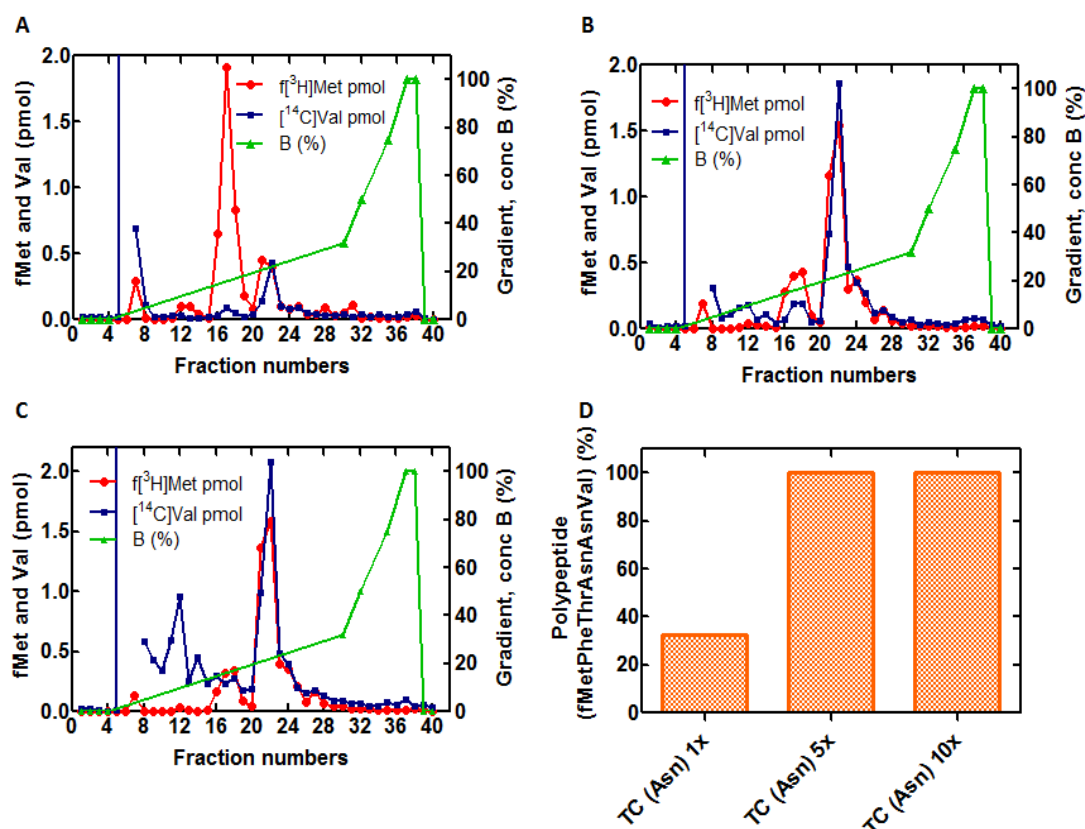
**Figure 44:** Codons to be translated up to the 2<sup>nd</sup> Val and the excess of the EF-Tu ternary complex over each codon.

The information on the abundance of individual tRNAs in the total tRNA pool allowed us to rationally design the optimization of the translation conditions. The idea of the experiment is to provide sufficient amounts of aminoacyl-tRNAs over each codon which has to be read up to the 2<sup>nd</sup> valine codon just before the UGA codon at position 140 (Figure 44A, B).



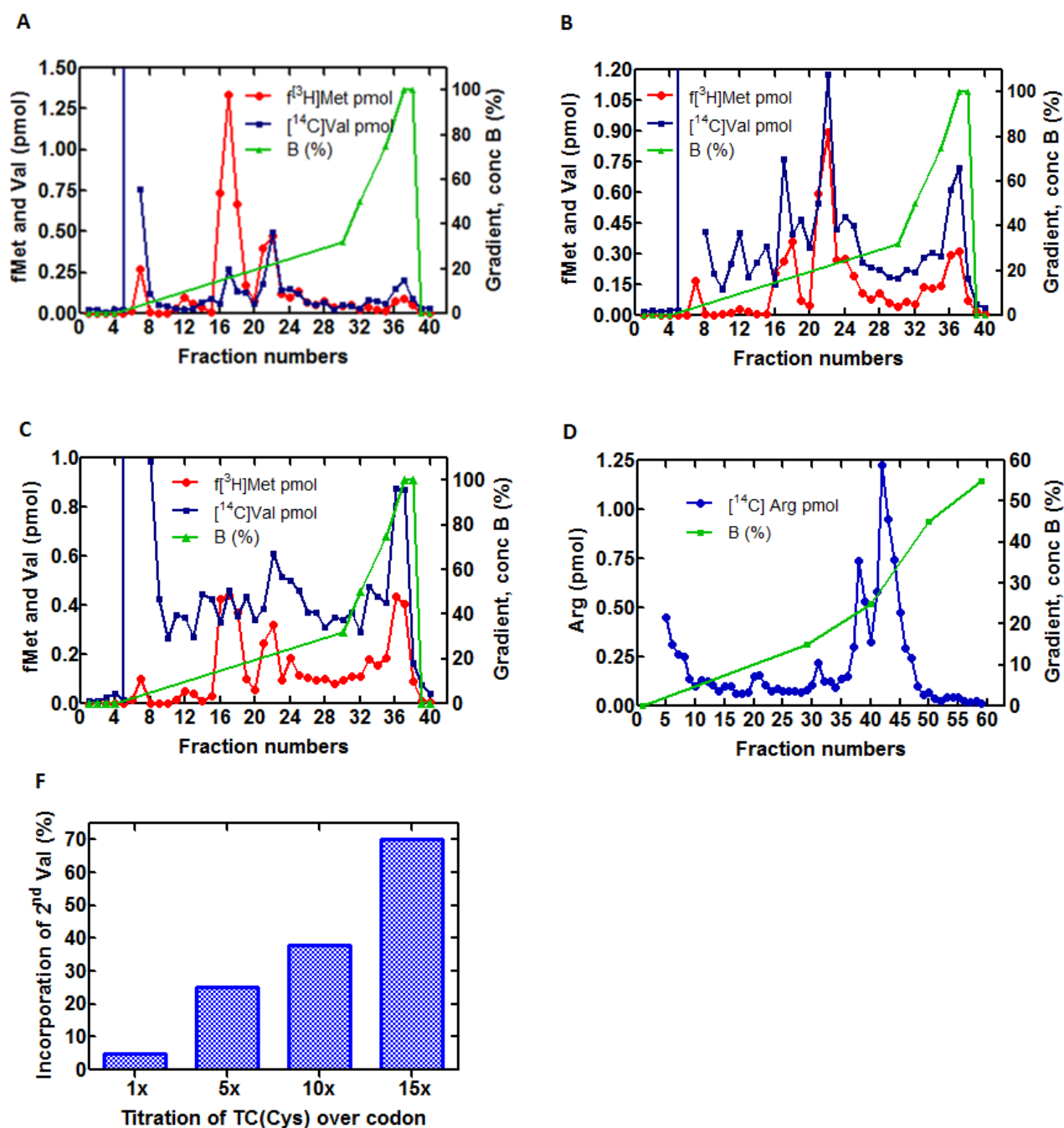
**Figure 45:** Optimization of translation. Formation of tripeptide (fMetPheThr).

In the following, we systematically tested the requirements for efficient translation of the mRNA sequence upstream the UGA codon. To check the incorporation of the 3<sup>rd</sup> amino acid, Thr, we have aminoacylated total tRNA with cold amino acid Phe and [<sup>14</sup>C]Thr (Materials and Methods). We added ternary complexes such that Thr ternary complex would be in a 1x or 5x excess over purified 70S·f[<sup>3</sup>H]Met-tRNA<sup>fMet</sup> initiation complex and the formation of the tripeptide fMetPheThr was monitored by HPLC (Figure 45A, B). The yield of the tripeptide was quite high, close to 90%, already at a stoichiometric amounts of the Thr ternary complex (Figure 45C).



**Figure 46:** Formation of polypeptide fMetPheThrAsnAsnVal.

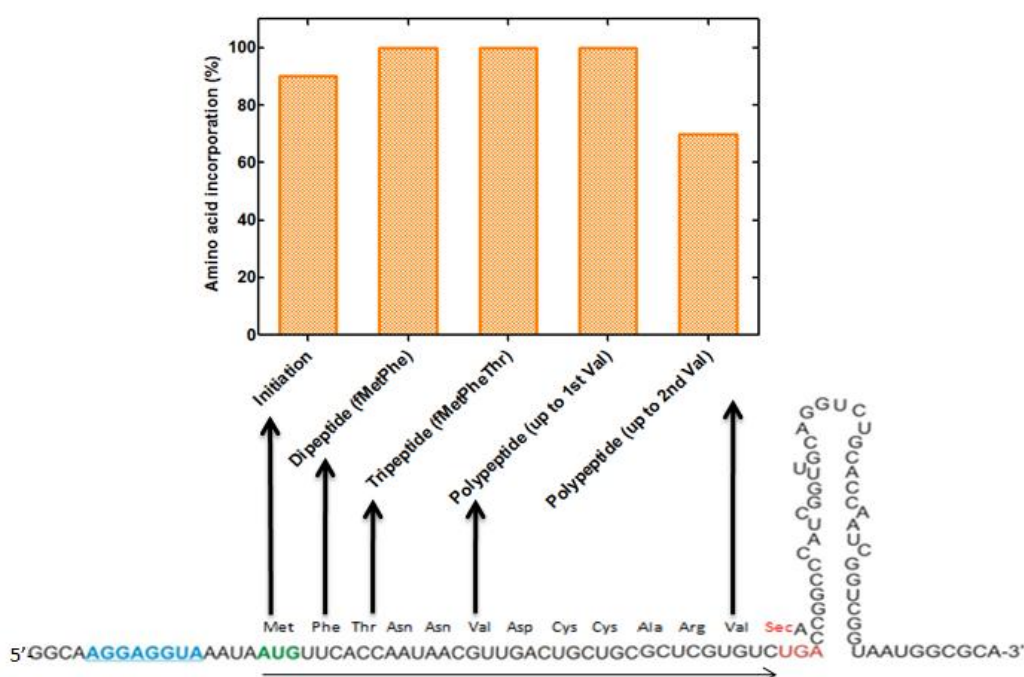
In the next step, we checked the translation of a peptide up to the 1<sup>st</sup> Val. We used aminoacylated total tRNA with cold Phe, Thr, Asn and [<sup>14</sup>C]Val. The amounts of ternary complexes were added such that the Asn ternary complex would be in a 1x (Figure 46A), 5x (Figure 46B) or 10x (Figure 46C) excess over the ribosomes and polypeptide was separated by HPLC. Analysis of peptides that contain both f[<sup>3</sup>H]Met and [<sup>14</sup>C]Val showed that a 5x excess of TC (Asn)/codon gives a 100% product up to 1<sup>st</sup> valine (Figure 46D).



**Figure 47:** Formation of polypeptide fMetPheThrAsnAsnValAspCysCysAlaArgVal.

In the final step, we tested the translation up to the 2<sup>nd</sup> Val, which immediately precedes the Sec codon. In this stretch, two Cys codons are found, for which the concentration of the individual tRNAs in the total tRNA pool is not known. We used aminoacylated total tRNA with non-radioactive Phe, Thr, Asn, Asp, Cys, Ala, Arg and  $[^{14}\text{C}]\text{Val}$ . and the ternary

complexes were added in concentrations that were estimated to be in a 1x (Figure 47A), 5x (Figure 47B), 10x (Figure 47C) over the Cys codons. and polypeptide was separated as previous gradient. For the experiment with a 15x excess of TC (Cys)/codon we prepared aminoacyl-tRNAs with cold Phe, Thr, Asn, Asp, Cys, Ala, Val and [ $^{14}\text{C}$ ]Arg and used a modified gradient for better separation of the peptide (Figure 47D). HPLC analysis shows that 15x excess of TC (Cys)/codon gives 70% readthrough up to 2<sup>st</sup> valine measured by incorporation of [ $^{14}\text{C}$ ]Arg (Figure 47F).



**Figure 48:** Summary of optimizations of *in vitro* translation codons up to 2<sup>nd</sup> Val.

The summary of the optimization of each translation step is shown in figure 48. The *in vitro* translation up to 2<sup>nd</sup> Valine yielded ribosome nascent chain complexes with up to ~70% homogeneity. This is a prerequisite for the following experiments where the Sec incorporation at the UGA codon is studied.

### 3.2.3 Incorporation of Sec into peptides *in vitro*

In order to understand the efficiency of Sec insertion at UGA codon against RF2, we measured Sec incorporation with increasing RF2 concentrations at conditions where all ternary complexes, including the SelB ternary complex and RF2 were added simultaneous to initiate translation. The concentrations of SelB and RF2 were close to the *in vivo* concentrations; note that SelB was not pre-bound to SECIS-element (Figure 49A). At these conditions, even in the absence of RF2 in the reaction only up to 30% of selenoprotein was synthesized. Upon increasing the concentration of RF2, Sec insertion was decreased even further (Figure 49B).

**Table 16:** Concentrations of translation machinery

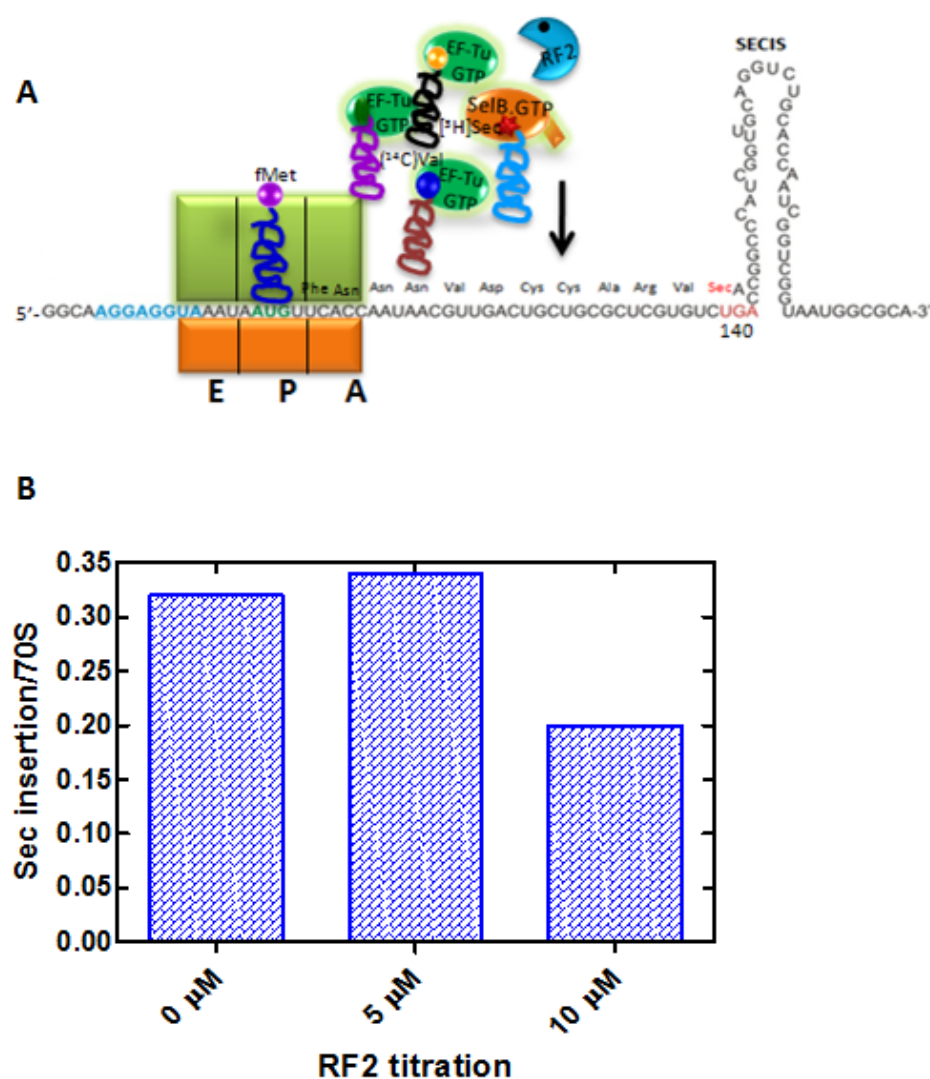
Components of translation machinery	<i>In vivo</i> concentration ( $\mu\text{M}$ ) <sup>a</sup>	<i>In vitro</i> concentration ( $\mu\text{M}$ ) <sup>c</sup>
Total ribosomes	10 <sup>b</sup>	0.14
Total EF-Tu	100	200
RF2	3 to 12	0 to 10
SelB	0.55	0.73
tRNA <sup>Sec</sup>	0.11	0.73

<sup>a</sup> (Adamski et al., 1994; Dong et al., 1996; Forchhammer et al., 1990; Tormay et al., 1996)

<sup>b</sup> Given that only three mRNA code for Sec in *E. coli*, the concentration of ribosomes programmed with these tRNAs is likely very low, <0.01  $\mu\text{M}$ .

<sup>c</sup> Concentration used in this work.



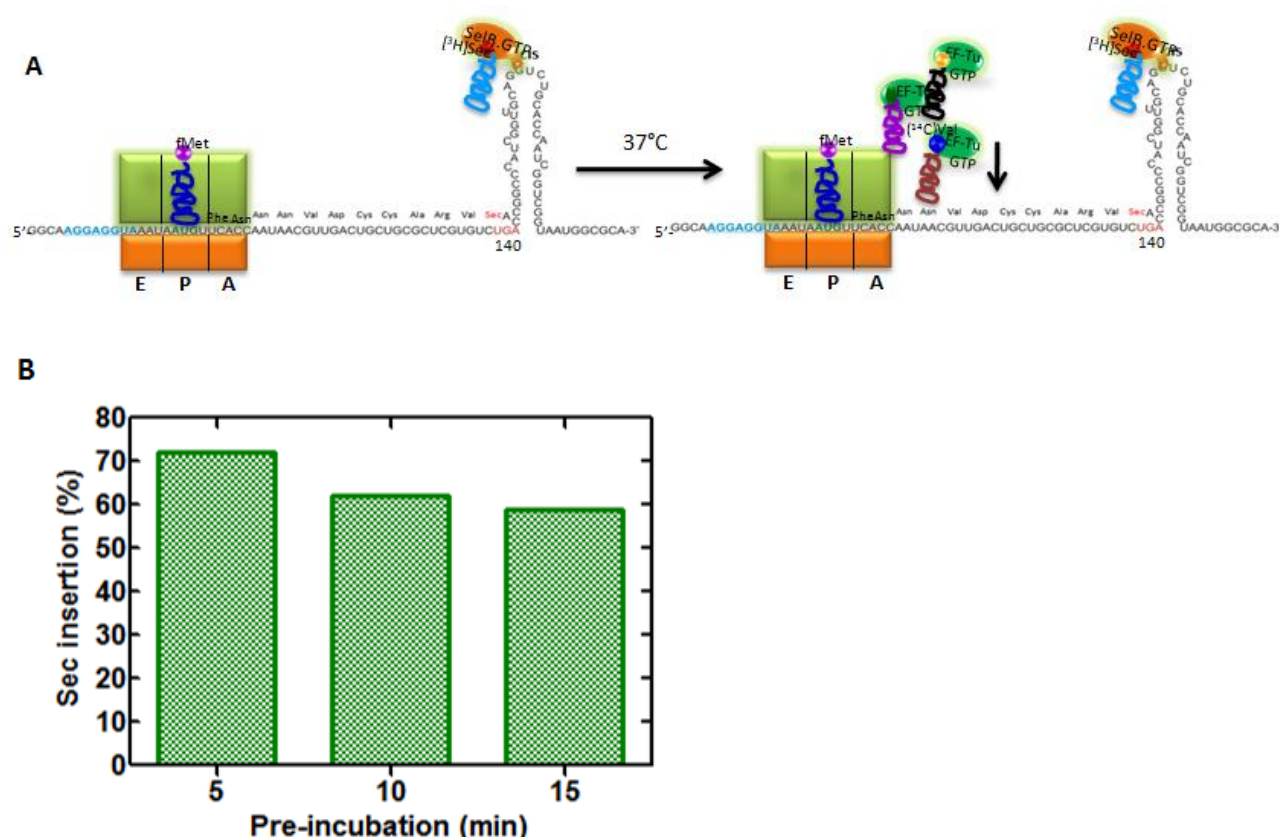


**Figure 49:** Efficiency of selenocysteine insertion in the presence and absence of RF2 without pre-incubation of SelB·GTP·Sec-tRNA<sup>Sec</sup> with SECIS-exposing initiated ribosomes.

These data suggest that the bulk of ternary complexes and RF2 are capable to compete with SelB·GTP·Sec-tRNA<sup>Sec</sup> for binding to the ribosome. Under the *in vivo* conditions SelB ternary complex may bind to SECIS in order to efficiently compete with RF2 against RF2 dependent termination at UGA codon.

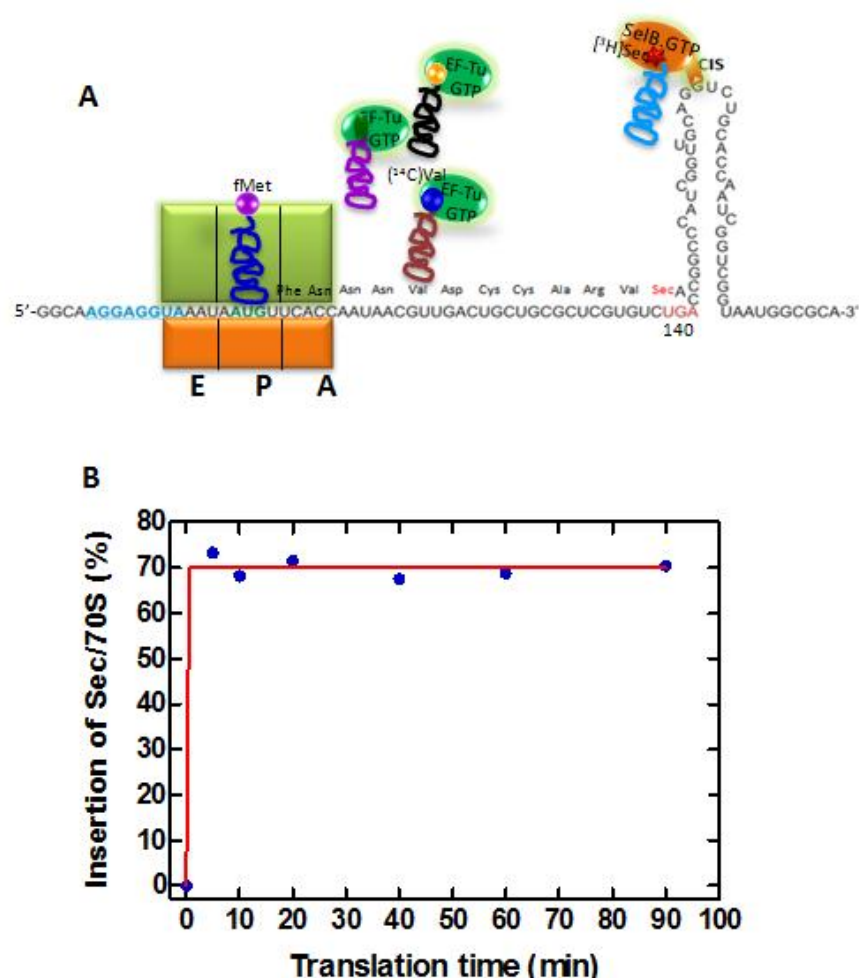
Incorporation of Sec into peptides entails two stages: the interaction of SelB·GTP·Sec-tRNA<sup>Sec</sup> with the SECIS element and delivery of the complex to the A site of the ribosome. In order to

optimize each of the stages, we tested how the pre-incubation of the SelB·GTP·Sec-tRNA<sup>Sec</sup> with the ribosome initiation complexes exposing SECIS affects Sec incorporation.



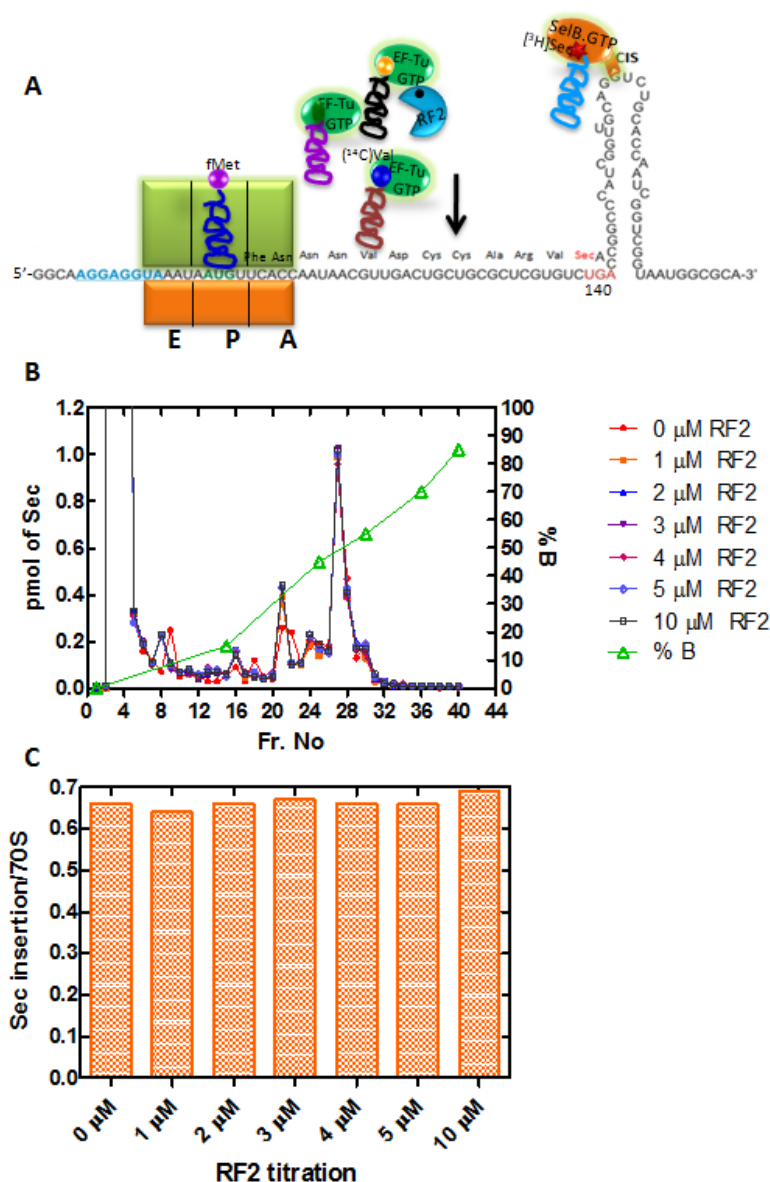
**Figure 50:** Optimization of pre-binding of SelB·GTP·Sec-tRNA<sup>Sec</sup> to SECIS of initiated ribosomes.

Purified ribosome initiation complexes programmed with *fdhF* mRNA (UGA<sub>140</sub>) and containing non-radioactive fMet-tRNA<sup>fMet</sup> in the P site were pre-incubated with a 5x excess of SelB·GTP·[<sup>3</sup>H]Sec-tRNA<sup>Sec</sup> at 37°C for different time intervals before addition of the EF-Tu ternary complex (Figure 50A). Translation was carried out for 90 min and polypeptide was analyzed by HPLC chromatogram. Pre-incubation of SelB ternary complex to the SECIS showed ~70% Sec insertion, i.e. practically all ribosomes that translated up to the Val codon preceding Sec were recruited to form selenopeptide.



**Figure 51:** Optimization of the translation time course.

We further followed the time course of translation at 37°C. Initiated ribosomes were pre-incubated with a 5x excess of SelB ternary complex for 5 min at 37°C. The Tu ternary complexes added and the time course of selenopeptide formation measured by HPLC (Figure 51A). The maximum possible amount of peptide was reached in less than 5 min (Figure 51B).



**Figure 52:** Efficiency of selenocysteine insertion in the presence and absence of RF2 when SelB-GTP-Sec-tRNA<sup>Sec</sup> pre-bound to SECIS of initiated ribosome.

The experiments shown so far were carried out in the absence of Release Factor 2 (RF2) which recognizes the UGA codon and may compete with the SelB-Sec-tRNA<sup>Sec</sup> for binding to the ribosome. In order to test whether the competition with RF2 affects Sec incorporation, we measure selenopeptide synthesis in the presence of increasing amounts of RF2. First we allowed the SelB ternary complex to pre-bind to SECIS (Figure 52A) and added Tu ternary

complexes together with RF2 (Figure 52B). The concentration of RF2 was close to its *in vivo* concentrations (3-10  $\mu$ M) (Table 16). High yield of selenopeptide was obtained regardless of RF2 concentration (Figure 52C). These results show when SelB is pre-bound to the SECIS element, RF2 does not compete with the Sec machinery for the stop codon.

## 4 Discussion

In bacteria, Sec incorporation into proteins works by bypassing the classical aa-tRNA delivery pathway of protein synthesis. Unlike the other standard amino acids, Sec is not coded directly in the genetic code. It is encoded in a special way by a UGA codon, which normally serves as a stop signal for termination of protein synthesis. UGA recoding into Sec takes place only when there is a special stem-loop, SECIS present immediately downstream of the UGA codon. Termination factor or release factor 2 (RF2) also recognizes this codon and competes with Sec insertion machinery at UGA recoding site. However, the mechanism of competition between RF2 and SelB-GTP-Sec-tRNA<sup>Sec</sup> at UGA recoding sites remains to be elucidated and the present work describes the first steps towards understanding the Sec-tRNA selection.

The mechanism of UGA decoding by Sec-tRNA<sup>Sec</sup> on the ribosome remains highly speculative, largely because detailed mechanistic studies were so far hampered by the lack of reliable *in vitro* systems. One unique property of SelB compared to other translation factors is its ability to interact with the SECIS-element even in the absence of the ribosome (Thanbichler et al., 2000). The secondary structure of SECIS is very stable and is likely to exist while the ribosome translates the upstream part of the mRNA. The ternary complex SelB-GTP-Sec-tRNA<sup>Sec</sup> forms very rapidly and is extremely stable (Paleskava et al., 2010). Two possible mechanisms were proposed, one of which focuses on the interaction of SelB-GTP-Sec-tRNA<sup>Sec</sup> to the SECIS element exposed to the ribosome (Huttenhofer and Bock, 1998) and second one deals with the potential mechanism of competition between the RF2 and Sec-tRNA<sup>Sec</sup> at the UGA recoding site (Mansell, J.B et al., 1999). The first model of SelB-GTP-Sec-tRNA<sup>Sec</sup> interaction with ribosome is purely based on the wealth of knowledge of mechanism of EF-Tu action on the ribosome.

According to the model proposed by Huttenhofer *et al.* (1998), SelB·GTP·Sec-tRNA<sup>Sec</sup> may interact first with the exposed SECIS stem-loop prior to the ribosome reaching the in-frame UGA codon. Upon translation, the lower part of stem-loop is expected to melt in order to position the UGA codon in the A site; concomitantly, SelB would be recruited to its interaction site on the ribosome. By analogy with EF-Tu·aa-tRNA, the UGA-Sec-tRNA<sup>Sec</sup> codon anti-codon interaction may activate the GTPase of SelB and deliver the Sec-tRNA<sup>Sec</sup> to the ribosome. GTP hydrolysis accelerates dissociation of Sec-tRNA<sup>Sec</sup> from SelB (Paleskava *et al.*, 2010). Upon delivery of Sec-tRNA<sup>Sec</sup> to the ribosome the affinity of SelB for its binding site would decrease (Thanbichler *et al.*, 2000) and then SelB dissociates from the ribosome; for the later steps, no clear experimental evidence is available.

Mansell *et al.* (1999) proposed a 'helical approach' model for the mechanism for how competition between Sec insertion and RF2 is mediated at the A site of the ribosome. According to this model, SelB·GTP·Sec-tRNA<sup>Sec</sup> is bound to the apical tip of the stem-loop SECIS and as the translating ribosome reaches towards to the UGA codon, the SelB complex must rotate about the axis of the helical stem as the secondary structure unwinds. The unwinding of the stem-loop structure is likely to cause a translational pause. Indeed, Supmann *et al.* (1999) has observed such translational pause upon UGA recoding by Sec-tRNA<sup>Sec</sup>. At the same time, SelB·GTP·Sec-tRNA<sup>Sec</sup> might be positioned in a way which allows it to block the entrance of RF2, thereby abolishing the competition with the termination factor. Again, direct experimental evidence for this model is lacking.

*E. coli* produces three types of formate dehydrogenases (Fdh) and these Fdhs contains intrinsic Sec residue at the active site. We have defined a model *fdhF* construct based on the available crystal structure of formate dehydrogenase-H (FdhH) (Boyington *et al.*, 1997) which contains Sec at position 140. The FdhH fragment has been defined the following way. First,

we looked for the *fdhF* native Shine-Dalgarno sequence and followed by the same spacer and canonical AUG initiation codon. Under *in vitro* translation conditions it is very difficult to prepare sufficiently large amounts of homogenous ribosome nascent chain complexes that reached the Sec codon at position 140. Therefore, we looked at the secondary structure of FDH-H (PDB 1FDI) to identify a shorter fragment which could be used to study UGA recoding while maintaining the protein secondary structure elements and keeping at least 10 natural amino acids upstream of the Sec 140 codon. The secondary structure of FdhH loop starts at Glycine 129 position followed by the natural amino acid sequence preceding the Sec 140 and up to a helix of Ala at position 156, which on the mRNA level also includes the SECIS-element. This stretch of fragment has no internal initiation codon. In the first experiments, we chose the natural Shine-Dalgarno sequence of the *fdhF* mRNA. However, as the efficiency of *in vitro* initiation was only up to 50%, and the yield of dipeptide (fMetGly) formation up to 60%, we redesigned the ribosome binding site and the first codon and replaced the *fdhF* natural Shine-Dalgarno sequence with the strong Shine-Dalgarno sequence followed by a spacer and the AUG initiation codon preceding a UUC codon Phe for a high yield of dipeptide (fMetPhe) formation. The efficiency of initiation and dipeptide formation with the improved construct was up to 90% and 100%, respectively. We have optimized the translation of codons up to the 2<sup>nd</sup> Val and were able to prepare up to 70% ribosome nascent-chain complexes (RNCs). The inherent high translation activity allowed us to further use the same *fdhF* mRNA UGA<sub>140</sub> including stem-loop SECIS for the *in vivo* assays, such that the same coding sequence of the *fdhF* mRNA UGA<sub>140</sub> including stem-loop SECIS could be used for all of *in vitro* and *in vivo* experiments. We note that here we describe the first example of the fully reconstituted, high-yield *in vitro* translation system designed to study the mechanism of selenoprotein synthesis.



One important question is to quantify the efficiency of Sec incorporation into proteins *in vivo* the dual reporter assay used for this study was carefully validated. Initial experiments on Sec insertion in LB media at the initial-log phase showed no detectable UGA<sub>140</sub> read through in the presence of the SECIS element, while a clear protein band was seen when UGA<sub>140</sub> replaced with UUC (Phe codon). This observation excluded the possibility of ribosomal drop-off at the stable stem-loop SECIS structure. However, it is unclear which factor(s)/step(s) limit the UGA<sub>140</sub> readthrough. It is widely accepted that RF2 is a factor that competes with Sec-tRNA<sup>Sec</sup> for reading the UGA codon, which is the canonical function of RF2 in termination. Earlier studies on levels of RF2 in *E. coli* show that the levels of RF2 can significantly vary depending on growth rate (Adamski et al., 1994). At high growth rates there are high levels of RF2 and at low growth rates the concentration of RF2 decreases, varying in the range between 3-12.5  $\mu$ M or 5900-24900 copies per cell. Earlier estimations suggested that in *E. coli* the efficiency of Sec insertion at UGA in the presence of SECIS-element is low, about 5%, at rapid growth phase (Suppmann et al., 1999). However, the efficiency of Sec insertion significantly varies dependent on the doubling time at initial-log phase (Mansell et al., 2001). When cells were at exponential growth phase (doubling time 20 min) in rich media UGA readthrough was <30%, at medium growth phase (doubling time 90 min) in minimal media Sec insertion raised to 40% and at slow growth phase (doubling time 125 min) in poor media Sec insertion increased to 60% (Mansell et al., 2001).

The present work indicates that in the presence of sufficient selenium source, the efficiency of Sec insertion *in vivo* was about 30%, essentially independent of the growth phase. It is possible that the low efficiency of Sec insertion in earlier reports at rapid growth phase of initial-log phase was due to the low selenium levels leading to shortage in the amounts of Sec-tRNA<sup>Sec</sup> required to recode the entire *fdhF* mRNA transcripts overproduced. In order to

understand the limiting factor for UGA readthrough, we conducted a series of experiments to check the requirement of an additional selenium source for the internal synthesis of Sec-tRNA<sup>Sec</sup> by SelD. Taken into consideration the *in vivo* concentration of tRNA<sup>Sec</sup>, 0.11  $\mu$ M (Dong et al., 1996; Tormay et al., 1996), and SelB, 0.55  $\mu$ M (Forchhammer et al., 1990), we performed experiments on selenium dependence of Sec insertion at initial-log phase (rapid growth phase). Upon increasing the selenium levels, there is increasing UGA readthrough resulting in Sec insertion levels of up to 30-35%. The 5-10% of readthrough at construct UGA<sub>140</sub> lacking SECIS-element can be attributed to miscoding, which was expected low decoding error frequency in bacteria (Drummond and Wilke, 2009). An alternative explanation, e.g. that the readthrough of the control construct was due to inefficient Sec incorporation in the absence of the SECIS-element, is less likely because the efficiency of the readthrough did not depend on selenium concentration. In contrast to the initial-log phase, Sec insertion at stationary phase (very poor growth phase) did not depend on selenium levels and was up to 35-45%. Notably, the overall protein synthesis was dramatically decreased at the stationary phase, such that the absolute amounts of Sec-containing proteins were strongly reduced compared to the initial-log phase. It is possible that trace amounts of selenium source in the medium (e.g. from the dead cells) are sufficient to support the readthrough at the stationary phase. However, the relative amount of the Sec-containing protein compared to the control construct were rather invariant at the optimum selenium conditions at any growth phase. This may suggest that this level of the readthrough reflects the inherent recoding efficiency, presumably determined by the competition between the SelB and EF-Tu ternary complexes and RF2.

To further understand the efficiency of Sec insertion at various phases of cell growth, we conducted a series of experiments at initial-log, mid-log and stationary phase, to test for the

potential correlation between Sec insertion and physiological status of the cell. If the Sec insertion machinery were in direct competition with RF2 dependent termination at the UGA<sub>140</sub> codon, the main prediction would be that the levels/activities of RF2 would affect the efficiency of Sec insertion, i.e. with decreasing RF2 concentrations at mid-log and stationary phases, Sec incorporation is expected to increase resulting in increased synthesis of full-length fusion protein product. However, this is not observed. This may suggest that RF2 does not compete with SelB·Sec-tRNA<sup>Sec</sup> for reading the stop codon in the context of the SECIS, at variance with the conclusion reached by Suppmann et al., 1999. Alternatively, the concentration of the Sec-insertion machinery may decrease with the growth rate as well. While it is not known whether the concentrations of SelB or tRNA<sup>Sec</sup> change with the growth phase, the concentrations of some tRNAs may vary up to 10-fold with the doubling time (Bremer and Dennis, 1987).

In order to directly address the question of competition between SelB, EF-Tu, and RF2, and to study the mechanism of Sec insertion into peptide, we could use a well defined *in vitro* translation system, which we used to test directly the proposed model of SelB complex interaction with the SECIS-element, i.e. testing whether pre-binding of SelB to the SECIS element affects the competition with RF2. In fact, there was a dramatic difference in Sec insertion efficiency depending on whether or not the SelB complex was allowed to pre-bind to the SECIS-element. When SelB was pre-bound to the SECIS-element, Sec insertion was up to 60-70% regardless of the excess of RF2. In contrast, when the SelB complex is added with all other components of the translational machinery, Sec insertion dropped to 30-35% even in the absence of RF2. Addition of RF2 up to moderate cellular concentration up to 5  $\mu$ M did not affect insertion but a RF2 concentration of 10  $\mu$ M showed an of Sec insertion readthrough dropped to 20%. These *in vitro* results are consistent with measured *in vivo*

readthrough up to 35%. The developed *in vitro* Sec insertion system allows us to further study the kinetics of UGA recoding into Sec and measuring the time of translation pause at UGA codon. The future goal of our experiments will be to further study the kinetic mechanism of UGA recoding into Sec by the bacterial machinery and the mechanism behind the inhibition of RF2 dependent termination.

## 5 Summary

Selenocysteine is the 21<sup>st</sup> amino acid which is incorporated into proteins by recoding a stop codon UGA followed by a selenocysteine insertion sequence (SECIS) of the mRNA. In bacteria, selenocysteine insertion requires specialized machinery which includes selenocysteine-specific tRNA<sup>Sec</sup>, translation factor SelB which delivers Sec-tRNA<sup>Sec</sup> to the ribosome, as well as proteins Sela, SelD, and seryl-tRNA synthetase which are required to produce Sec-tRNA<sup>Sec</sup>. The aim of this work is to develop experimental assays to study Sec incorporation into proteins *in vivo* and *in vitro*. As an *in vivo* assay, we designed the dual luciferase reporter assay and validated its performance using Western blots and luciferase reactions. The efficiency of Sec incorporation was 35% independent of growth conditions. Rapidly growing cells required additional selenium source for efficient Sec insertion. This level of UGA recoding could be reproduced in the fully reconstituted *in vitro* translation system upon synthesis of a fragment of a natural selenoprotein FdhH. The recruitment of SelB to the SECIS-element prior to translation prevented inhibition of Sec insertion by RF2, a termination factor which usually recognizes the UGA codon and competes with Sec-tRNA<sup>Sec</sup> for binding to the ribosome. These results shed light on the importance of the SECIS and on the mechanism by which a stop codon is redirected for efficient readthrough by a specific tRNA.

**6 References**

- Adamski, F.M., McCaughan, K.K., Jorgensen, F., Kurland, C.G., and Tate, W.P. (1994). The concentration of polypeptide chain release factors 1 and 2 at different growth rates of *Escherichia coli*. *J Mol Biol* 238, 302-308.
- Araiso, Y., Sherrer, R.L., Ishitani, R., Ho, J.M., Soll, D., and Nureki, O. (2009). Structure of a tRNA-dependent kinase essential for selenocysteine decoding. *Proc Natl Acad Sci U S A* 106, 16215-16220.
- Axley, M.J., Bock, A., and Stadtman, T.C. (1991). Catalytic properties of an *Escherichia coli* formate dehydrogenase mutant in which sulfur replaces selenium. *Proc Natl Acad Sci U S A* 88, 8450-8454.
- 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.
- Baron, C., Heider, J., and Bock, A. (1993). Interaction of translation factor SELB with the formate dehydrogenase H selenopolypeptide mRNA. *Proc Natl Acad Sci U S A* 90, 4181-4185.
- Barrell, B.G., Bankier, A.T., and Drouin, J. (1979). A different genetic code in human mitochondria. *Nature* 282, 189-194.
- Berchtold, H., Reshetnikova, L., Reiser, C.O., Schirmer, N.K., Sprinzl, M., and Hilgenfeld, R. (1993). Crystal structure of active elongation factor Tu reveals major domain rearrangements. *Nature* 365, 126-132.

- Berry, M.J., Banu, L., Harney, J.W., and Larsen, P.R. (1993). Functional characterization of the eukaryotic SECIS elements which direct selenocysteine insertion at UGA codons. *EMBO J* 12, 3315-3322.
- Bock, A., Forchhammer, K., Heider, J., and Baron, C. (1991). Selenoprotein synthesis: an expansion of the genetic code. *Trends Biochem Sci* 16, 463-467.
- Boyington, J.C., Gladyshev, V.N., Khangulov, S.V., Stadtman, T.C., and Sun, P.D. (1997). Crystal structure of formate dehydrogenase H: catalysis involving Mo, molybdopterin, selenocysteine, and an Fe<sub>4</sub>S<sub>4</sub> cluster. *Science* 275, 1305-1308.
- Bremer, H., and Dennis, P.P. (1987). Modulation of chemical composition and other parameters of the cell by growth rate. In *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology, F.C. Neidhardt, ed. (Washington, DC, American Society for Microbiology), pp. 1553-1569.
- Budiman, M.E., Bubenik, J.L., Miniard, A.C., Middleton, L.M., Gerber, C.A., Cash, A., and Driscoll, D.M. (2009). Eukaryotic initiation factor 4a3 is a selenium-regulated RNA-binding protein that selectively inhibits selenocysteine incorporation. *Mol Cell* 35, 479-489.
- Bult, C.J., White, O., Olsen, G.J., Zhou, L., Fleischmann, R.D., Sutton, G.G., Blake, J.A., FitzGerald, L.M., Clayton, R.A., Gocayne, J.D., *et al.* (1996). Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. *Science* 273, 1058-1073.
- Carlson, B.A., Xu, X.M., Kryukov, G.V., Rao, M., Berry, M.J., Gladyshev, V.N., and Hatfield, D.L. (2004). Identification and characterization of phosphoseryl-tRNA[Ser]<sup>Sec</sup> kinase. *Proc Natl Acad Sci U S A* 101, 12848-12853.

- Chambers, I., Frampton, J., Goldfarb, P., Affara, N., McBain, W., and Harrison, P.R. (1986). The structure of the mouse glutathione peroxidase gene: the selenocysteine in the active site is encoded by the 'termination' codon, TGA. *EMBO J* 5, 1221-1227.
- Chavatte, L., Brown, B.A., and Driscoll, D.M. (2005). Ribosomal protein L30 is a component of the UGA-selenocysteine recoding machinery in eukaryotes. *Nat Struct Mol Biol* 12, 408-416.
- Chiba, S., Itoh, Y., Sekine, S., and Yokoyama, S. (2010). Structural basis for the major role of O-phosphoseryl-tRNA kinase in the UGA-specific encoding of selenocysteine. *Mol Cell* 39, 410-420.
- Christiansen, G., Jensen, L.T., Boesen, T., Emmersen, J., Ladefoged, S.A., Schiotz, L.K., and Birkelund, S. (1997). Molecular biology of Mycoplasma. *Wien Klin Wochenschr* 109, 557-561.
- Copeland, P.R., Fletcher, J.E., Carlson, B.A., Hatfield, D.L., and Driscoll, D.M. (2000). A novel RNA binding protein, SBP2, is required for the translation of mammalian selenoprotein mRNAs. *EMBO J* 19, 306-314.
- Dong, H., Nilsson, L., and Kurland, C.G. (1996). Co-variation of tRNA abundance and codon usage in Escherichia coli at different growth rates. *J Mol Biol* 260, 649-663.
- Drummond, D.A., and Wilke, C.O. (2009). The evolutionary consequences of erroneous protein synthesis. *Nat Rev Genet* 10, 715-724.
- Feng, L., Sheppard, K., Tumbula-Hansen, D., and Soll, D. (2005a). Gln-tRNA<sup>Gln</sup> formation from Glu-tRNA<sup>Gln</sup> requires cooperation of an asparaginase and a Glu-tRNA<sup>Gln</sup> kinase. *J Biol Chem* 280, 8150-8155.



- Feng, L., Yuan, J., Toogood, H., Tumbula-Hansen, D., and Soll, D. (2005b). Aspartyl-tRNA synthetase requires a conserved proline in the anticodon-binding loop for tRNA(Asn) recognition in vivo. *J Biol Chem* 280, 20638-20641.
- Fixsen, S.M., and Howard, M.T. (2010). Processive selenocysteine incorporation during synthesis of eukaryotic selenoproteins. *J Mol Biol* 399, 385-396.
- Forchhammer, K., and Bock, A. (1991). Selenocysteine synthase from *Escherichia coli*. Analysis of the reaction sequence. *J Biol Chem* 266, 6324-6328.
- Forchhammer, K., Leinfelder, W., and Bock, A. (1989). Identification of a novel translation factor necessary for the incorporation of selenocysteine into protein. *Nature* 342, 453-456.
- Forchhammer, K., Rucknagel, K.P., and Bock, A. (1990). Purification and biochemical characterization of SELB, a translation factor involved in selenoprotein synthesis. *J Biol Chem* 265, 9346-9350.
- Frolova, L., Le Goff, X., Zhouravleva, G., Davydova, E., Philippe, M., and Kisselev, L. (1996). Eukaryotic polypeptide chain release factor eRF3 is an eRF1- and ribosome-dependent guanosine triphosphatase. *RNA* 2, 334-341.
- Fukunaga, R., and Yokoyama, S. (2007). Structural insights into the second step of RNA-dependent cysteine biosynthesis in archaea: crystal structure of Sep-tRNA:Cys-tRNA synthase from *Archaeoglobus fulgidus*. *J Mol Biol* 370, 128-141.
- Gajiwala, K.S., and Burley, S.K. (2000). Winged helix proteins. *Curr Opin Struct Biol* 10, 110-116.

- Garcia, G.E., and Stadtman, T.C. (1992). Clostridium sticklandii glycine reductase selenoprotein A gene: cloning, sequencing, and expression in Escherichia coli. *J Bacteriol* *174*, 7080-7089.
- Gladyshev, V.N., and Kryukov, G.V. (2001). Evolution of selenocysteine-containing proteins: significance of identification and functional characterization of selenoproteins. *Biofactors* *14*, 87-92.
- Gromadski, K.B., and Rodnina, M.V. (2004). Kinetic determinants of high-fidelity tRNA discrimination on the ribosome. *Mol Cell* *13*, 191-200.
- Gursinsky, T., Grobe, D., Schierhorn, A., Jager, J., Andreesen, J.R., and Sohling, B. (2008). Factors and selenocysteine insertion sequence requirements for the synthesis of selenoproteins from a gram-positive anaerobe in Escherichia coli. *Appl Environ Microbiol* *74*, 1385-1393.
- Gursinsky, T., Jager, J., Andreesen, J.R., and Sohling, B. (2000). A selDABC cluster for selenocysteine incorporation in Eubacterium acidaminophilum. *Arch Microbiol* *174*, 200-212.
- Halboth, S., and Klein, A. (1992). Methanococcus voltae harbors four gene clusters potentially encoding two [NiFe] and two [NiFeSe] hydrogenases, each of the cofactor F420-reducing or F420-non-reducing types. *Mol Gen Genet* *233*, 217-224.
- Heider, J., Baron, C., and Bock, A. (1992). Coding from a distance: dissection of the mRNA determinants required for the incorporation of selenocysteine into protein. *EMBO J* *11*, 3759-3766.
- Heider, J., Forchhammer, K., Sawers, G., and Bock, A. (1991). Interspecies compatibility of selenoprotein biosynthesis in Enterobacteriaceae. *Arch Microbiol* *155*, 221-228.

- Hilgenfeld, R., Bock, A., and Wilting, R. (1996). Structural model for the selenocysteine-specific elongation factor SelB. *Biochimie* 78, 971-978.
- Hill, K.E., Lloyd, R.S., and Burk, R.F. (1993). Conserved nucleotide sequences in the open reading frame and 3' untranslated region of selenoprotein P mRNA. *Proc Natl Acad Sci U S A* 90, 537-541.
- Huber, R.E., and Criddle, R.S. (1967). Comparison of the chemical properties of selenocysteine and selenocystine with their sulfur analogs. *Arch Biochem Biophys* 122, 164-173.
- Huttenhofer, A., and Bock, A. (1998). Selenocysteine inserting RNA elements modulate GTP hydrolysis of elongation factor SelB. *Biochemistry* 37, 885-890.
- Huttenhofer, A., Westhof, E., and Bock, A. (1996). Solution structure of mRNA hairpins promoting selenocysteine incorporation in *Escherichia coli* and their base-specific interaction with special elongation factor SELB. *RNA* 2, 354-366.
- Itoh, Y., Chiba, S., Sekine, S., and Yokoyama, S. (2009). Crystal structure of human selenocysteine tRNA. *Nucleic Acids Res* 37, 6259-6268.
- Kaiser, J.T., Gromadski, K., Rother, M., Engelhardt, H., Rodnina, M.V., and Wahl, M.C. (2005). Structural and functional investigation of a putative archaeal selenocysteine synthase. *Biochemistry* 44, 13315-13327.
- Kawashima, T., Berthet-Colominas, C., Wulff, M., Cusack, S., and Leberman, R. (1996). The structure of the *Escherichia coli* EF-Tu.EF-Ts complex at 2.5 Å resolution. *Nature* 379, 511-518.

- Keeling, P.J., Fast, N.M., and McFadden, G.I. (1998). Evolutionary relationship between translation initiation factor eIF-2gamma and selenocysteine-specific elongation factor SELB: change of function in translation factors. *J Mol Evol* 47, 649-655.
- Konecki, D.S., Aune, K.C., Tate, W., and Caskey, C.T. (1977). Characterization of reticulocyte release factor. *J Biol Chem* 252, 4514-4520.
- Kromayer, M., Neuhierl, B., Friebe, A., and Bock, A. (1999). Genetic probing of the interaction between the translation factor SelB and its mRNA binding element in *Escherichia coli*. *Mol Gen Genet* 262, 800-806.
- Kromayer, M., Wilting, R., Tormay, P., and Bock, A. (1996). Domain structure of the prokaryotic selenocysteine-specific elongation factor SelB. *J Mol Biol* 262, 413-420.
- Kryukov, G.V., and Gladyshev, V.N. (2004). The prokaryotic selenoproteome. *EMBO Rep* 5, 538-543.
- Leibundgut, M., Frick, C., Thanbichler, M., Bock, A., and Ban, N. (2005). Selenocysteine tRNA-specific elongation factor SelB is a structural chimaera of elongation and initiation factors. *EMBO J* 24, 11-22.
- Leinfelder, W., Stadtman, T.C., and Bock, A. (1989). Occurrence in vivo of selenocysteyl-tRNA(SERUCA) in *Escherichia coli*. Effect of sel mutations. *J Biol Chem* 264, 9720-9723.
- Leinfelder, W., Zehelein, E., Mandrand-Berthelot, M.A., and Bock, A. (1988). Gene for a novel tRNA species that accepts L-serine and cotranslationally inserts selenocysteine. *Nature* 331, 723-725.

- Liu, Z., Reches, M., Groisman, I., and Engelberg-Kulka, H. (1998). The nature of the minimal 'selenocysteine insertion sequence' (SECIS) in *Escherichia coli*. *Nucleic Acids Res* 26, 896-902.
- Low, S.C., and Berry, M.J. (1996). Knowing when not to stop: selenocysteine incorporation in eukaryotes. *Trends Biochem Sci* 21, 203-208.
- Mansell, J.B., Guevremont, D., Poole, E.S., and Tate, W.P. (2001). A dynamic competition between release factor 2 and the tRNA(Sec) decoding UGA at the recoding site of *Escherichia coli* formate dehydrogenase H. *EMBO J* 20, 7284-7293.
- Martin, G.W., 3rd, Harney, J.W., and Berry, M.J. (1996). Selenocysteine incorporation in eukaryotes: insights into mechanism and efficiency from sequence, structure, and spacing proximity studies of the type 1 deiodinase SECIS element. *RNA* 2, 171-182.
- Martin, G.W., 3rd, Harney, J.W., and Berry, M.J. (1998). Functionality of mutations at conserved nucleotides in eukaryotic SECIS elements is determined by the identity of a single nonconserved nucleotide. *RNA* 4, 65-73.
- Matsugi, J., and Murao, K. (2004). Genomic investigation of the system for selenocysteine incorporation in the bacterial domain. *Biochim Biophys Acta* 1676, 23-32.
- Milon, P., Tischenko, E., Tomsic, J., Caserta, E., Folkers, G., La Teana, A., Rodnina, M.V., Pon, C.L., Boelens, R., and Gualerzi, C.O. (2006). The nucleotide-binding site of bacterial translation initiation factor 2 (IF2) as a metabolic sensor. *Proc Natl Acad Sci U S A* 103, 13962-13967.
- Paleskava, A., Konevega, A.L., and Rodnina, M.V. (2010). Thermodynamic and kinetic framework of selenocysteyl-tRNA<sup>Sec</sup> recognition by elongation factor SelB. *J Biol Chem* 285, 3014-3020.

- Rodnina, M.V., Stark, H., Savelsbergh, A., Wieden, H.J., Mohr, D., Matassova, N.B., Peske, F., Daviter, T., Gualerzi, C.O., and Wintermeyer, W. (2000). GTPases mechanisms and functions of translation factors on the ribosome. *Biol Chem* 381, 377-387.
- Rother, M., Resch, A., Wilting, R., and Bock, A. (2001). Selenoprotein synthesis in archaea. *Biofactors* 14, 75-83.
- Rother, M., Wilting, R., Commans, S., and Bock, A. (2000). Identification and characterisation of the selenocysteine-specific translation factor SelB from the archaeon *Methanococcus jannaschii*. *J Mol Biol* 299, 351-358.
- Sauerwald, A., Zhu, W., Major, T.A., Roy, H., Palioura, S., Jahn, D., Whitman, W.B., Yates, J.R., 3rd, Ibba, M., and Soll, D. (2005). RNA-dependent cysteine biosynthesis in archaea. *Science* 307, 1969-1972.
- Schon, A., Bock, A., Ott, G., Sprinzl, M., and Soll, D. (1989). The selenocysteine-inserting opal suppressor serine tRNA from *E. coli* is highly unusual in structure and modification. *Nucleic Acids Res* 17, 7159-7165.
- Selmer, M., and Su, X.D. (2002). Crystal structure of an mRNA-binding fragment of *Moorella thermoacetica* elongation factor SelB. *EMBO J* 21, 4145-4153.
- Sherrer, R.L., Ho, J.M., and Soll, D. (2008). Divergence of selenocysteine tRNA recognition by archaeal and eukaryotic O-phosphoserine-tRNA<sup>Sec</sup> kinase. *Nucleic Acids Res* 36, 1871-1880.
- Squires, J.E., Stoytchev, I., Forry, E.P., and Berry, M.J. (2007). SBP2 binding affinity is a major determinant in differential selenoprotein mRNA translation and sensitivity to nonsense-mediated decay. *Mol Cell Biol* 27, 7848-7855.

- Stadtman, T.C. (1996). Selenocysteine. *Annu Rev Biochem* 65, 83-100.
- Suppmann, S., Persson, B.C., and Bock, A. (1999). Dynamics and efficiency in vivo of UGA-directed selenocysteine insertion at the ribosome. *EMBO J* 18, 2284-2293.
- Thanbichler, M., and Bock, A. (2001). Functional analysis of prokaryotic SELB proteins. *Biofactors* 14, 53-59.
- Thanbichler, M., Bock, A., and Goody, R.S. (2000). Kinetics of the interaction of translation factor SelB from *Escherichia coli* with guanosine nucleotides and selenocysteine insertion sequence RNA. *J Biol Chem* 275, 20458-20466.
- Tormay, P., Sawers, A., and Bock, A. (1996). Role of stoichiometry between mRNA, translation factor SelB and selenocysteyl-tRNA in selenoprotein synthesis. *Mol Microbiol* 21, 1253-1259.
- Tujebajeva, R.M., Copeland, P.R., Xu, X.M., Carlson, B.A., Harney, J.W., Driscoll, D.M., Hatfield, D.L., and Berry, M.J. (2000). Decoding apparatus for eukaryotic selenocysteine insertion. *EMBO Rep* 1, 158-163.
- Wilden, B., Savelsbergh, A., Rodnina, M.V., and Wintermeyer, W. (2006). Role and timing of GTP binding and hydrolysis during EF-G-dependent tRNA translocation on the ribosome. *Proc Natl Acad Sci U S A* 103, 13670-13675.
- 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.

- Wintjens, R., and Rooman, M. (1996). Structural classification of HTH DNA-binding domains and protein-DNA interaction modes. *J Mol Biol* 262, 294-313.
- Wohlgemuth, I., Brenner, S., Beringer, M., and Rodnina, M.V. (2008). Modulation of the rate of peptidyl transfer on the ribosome by the nature of substrates. *J Biol Chem* 283, 32229-32235.
- Wu, R., Shen, Q., and Newburger, P.E. (2000). Recognition and binding of the human selenocysteine insertion sequence by nucleolin. *J Cell Biochem* 77, 507-516.
- Xu, X.M., Carlson, B.A., Mix, H., Zhang, Y., Saira, K., Glass, R.S., Berry, M.J., Gladyshev, V.N., and Hatfield, D.L. (2007). Biosynthesis of selenocysteine on its tRNA in eukaryotes. *PLoS Biol* 5, e4.
- Yoshizawa, S., Rasubala, L., Ose, T., Kohda, D., Fourmy, D., and Maenaka, K. (2005). Structural basis for mRNA recognition by elongation factor SelB. *Nat Struct Mol Biol* 12, 198-203.
- Yuan, J., Palioura, S., Salazar, J.C., Su, D., O'Donoghue, P., Hohn, M.J., Cardoso, A.M., Whitman, W.B., and Soll, D. (2006). RNA-dependent conversion of phosphoserine forms selenocysteine in eukaryotes and archaea. *Proc Natl Acad Sci U S A* 103, 18923-18927.
- Zinoni, F., Birkmann, A., Stadtman, T.C., and Bock, A. (1986). Nucleotide sequence and expression of the selenocysteine-containing polypeptide of formate dehydrogenase (formate-hydrogen-lyase-linked) from *Escherichia coli*. *Proc Natl Acad Sci U S A* 83, 4650-4654.



Zinoni, F., Heider, J., and Bock, A. (1990). Features of the formate dehydrogenase mRNA necessary for decoding of the UGA codon as selenocysteine. *Proc Natl Acad Sci U S A* 87, 4660-4664.

**7 Abbreviations**

aa-tRNA	aminoacyl-tRNA
dNTP	deoxynucleoside-triphosphate
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EF-G	Elongation factor-G
EF-Tu	Elongation factor-Tu
GTP	Guanosine tri phosphate
LB	Luria Bertani
mRNA	messenger RNA
Na <sub>2</sub> SeO <sub>3</sub>	Sodium selenite
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffer saline
PCR	Polymerase chain reaction
PMSF	Phenylmethylsulfonyl fluoride
SDS	Sodium dodecyl sulphate
Sec	Selenocysteine
SECIS	Selenocysteine insertion sequence
tRNA	transfer RNA

### 8 Acknowledgements

First and foremost, I would like to thank my supervisor Prof. Dr. Marina V. Rodnina for giving me the opportunity to work in her lab and continuous support and guidance. I also thank to Prof. Dr. Wolfgang Wintermeyer for advices.

I am very grateful to my PhD thesis committee members Prof. Dr. Holger Stark and Prof. Dr. Ralf Ficner for their helpful suggestions and fruitful discussions.

Special thanks to Dr. Andrey L. Konevega and Dr. Frank Peske for their help.

I want to thank the GGNB program staff for the informative method courses and inspiring retreats.

I am very thankful to all of my lab members for their help.

## Curriculum vitae

Name	Suresh Babu Kotini
Date of birth	December 7 <sup>th</sup> , 1980
Place of birth	Rajampalem, India
2007-2011	Ph.D. thesis (GGNB PhD program) Prof. Dr. Marina V. Rodnina Department of Physical Biochemistry Max Planck Institute for Biophysical Chemistry Germany
2004-2007	Project assistant Centre for Cellular and Molecular Biology India
2002-2004	M.Sc. Microbiology Acharya Nagarjuna University India
2001-2002	Biochemist NVR Diagnostics & Research Centre India
1998-2001	B.Sc. Biochemistry Andhra University India
1996-1998	Intermediate education Board of Intermediate education India
1991-1996	High school School of secondary education India
1985-1991	Elementary school India