

**Structural and Biochemical Studies of the Human Selenocysteine
tRNA-Specific Elongation Factor eEFSec**

BY

MALGORZATA DOBOSZ-BARTOSZEK
M.S., Jagiellonian University, Krakow, Poland

THESIS

Submitted as partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Biochemistry and Molecular Genetics
in the Graduate College of the
University of Illinois at Chicago, 2016

Chicago, Illinois

Defense Committee:

Miljan Simonović, Chair and Advisor
Michael Caffrey
Karen Colley
Arnon Lavie
Alexander Mankin, Medicinal Chemistry and Pharmacognosy
Karl Volz, Microbiology and Immunology

This thesis is dedicated to my husband, Gregory. It is because of his presence by my side in both cheerful and challenging times, and continuous support and encouragement that this work could be accomplished.

ACKNOWLEDGMENTS

I would like to thank my thesis advisor, Prof. Miljan Simonović, for taking me under his mentorship during my doctoral studies, as well for his patience, guidance and advice. I am grateful for all his help and feedback, especially when working to improve my writing and oral presentation skills. Lastly, I am thankful to Prof. Simonović for his individual mentorship approach towards me, and finding the right balance between hands-on guidance and allowing me to be independent in my work.

Also, I would like to express my gratitude to all members of my thesis committee - Prof. Michael Caffrey, Prof. Karen Colley, Prof. Arnon Lavie, Prof. Alexander Mankin, and Prof. Karl Volz. I am grateful for their time, support, feedback on my research progress, and ideas for future experiments alongside the suggestions towards my personal development as an independent scientist.

I am thankful to all the collaborators and individuals who helped to advance my thesis work, especially Prof. Zbyszek Otwinowski, Dr. Srinivas Chakravarthy, Prof. Paul Copeland, and Prof. Dieter Söll.

I would also like to thank both the former and present members of the Simonović Lab for their research suggestions and for creating a pleasant working atmosphere.

Last but not the least, I would like to thank my family, both in Poland and the US, and all friends, particularly Dr. Agnieszka Mateja-Paduch, for their ongoing support and help, and for molding me into a person who I am now.

MDB

AUTHOR CONTRIBUTIONS

The human eEFSec-pET15b plasmid was a gift from the laboratory of Prof. Dieter Söll (Departments of Molecular Biophysics and Biochemistry, and Chemistry, Yale University, New Haven, CT). The experimental phase for SeMet-eEFSec:GDPNP crystals was calculated with help of Dr. Zbyszek Otwinowski (Department of Biochemistry, University of Texas Southwestern Medical Center, Dallas, TX). The SAXS data for eEFSec:GDPCP and eEFSec:GDP complexes were collected and the molecular envelopes were calculated by Dr. Srinivas Chakravarthy (Biophysics Collaborative Access Team/ Illinois Institute of Technology, Sector 18ID, Advanced Photon Source, Chicago, IL). Selenocysteine incorporation *in vitro* assays for wild type and mutant eEFSec constructs were completed by Mark H. Pinkerton and Dr. Paul R. Copeland (Department of Biochemistry and Molecular Biology, Rutgers – Robert Wood Johnson Medical School, Piscataway, NJ).

TABLE OF CONTENTS

<u>CHAPTER</u>	<u>PAGE</u>
I. BACKGROUND AND SIGNIFICANCE.....	1
I.1 The physiological role of selenium.....	1
I.2 The distinct features of selenocysteine.....	2
I.3 Selenocysteine tRNA: structure, function and role in health.....	5
I.4 Selenocysteine tRNA-specific elongation factor.....	8
I.5 Other factors involved in decoding of the selenocysteine UGA codon.....	9
I.5.1 SelenoCysteine Insertion Sequence (SECIS).....	10
I.5.2 SECIS-Binding Protein 2 (SBP2).....	11
I.6 The elongation phase of protein synthesis.....	12
II. MATERIALS AND METHODS	16
II.1 Cloning, expression and purification of recombinant human eEFSec.....	16
II.1.1 Wild-type eEFSec.....	16
II.1.2 The selenomethionine-labeled human eEFSec.....	17
II.1.3 His-Avi-TEV-eEFSec.....	18
II.2 Site-directed mutagenesis, expression, and purification of recombinant human eEFSec mutants	20
II.3 Crystallization, data collection and data processing.....	22
II.3.1 eEFSec:GDPNP.....	22
II.3.2 SeMet-eEFSec:GDPNP.....	22
II.3.3 eEFSec:GDPCP.....	23
II.3.4 eEFSec:GDP.....	23
II.4 Structure determination and refinement.....	24
II.5 Analysis of eEFSec:GDP and eEFSec:GDPCP by Small-angle X-ray scattering (SAXS).....	25
II.6 Analysis of nucleotide binding to eEFSec by Isothermal Titration Calorimetry (ITC)	26
II.7 <i>In vitro</i> selenocysteine incorporation assay.....	27
III. The crystal structure of the GTP-bound state of human eEFSec	28
III.1 Introduction.....	28
III.2 Experimental Results	30
III.2.1 Overall structure and domain organization of human eEFSec.....	30
III.2.2 A structural comparison with general translation elongation factors EF-Tu and EF1A.....	36
III.2.3 A structural comparison with SelB orthologs	39
III.2.4 Human eEFSec resembles the universal translation initiation factor IF2/eIF5B	44
III.3 Discussion.....	45

TABLE OF CONTENTS (continued)

<u>CHAPTER</u>	<u>PAGE</u>
IV. The EF-Tu-like domain of human eEFSec harbors the major functional sites: the GTPase site and Sec-binding pocket	48
IV.1 Introduction.....	48
IV.2 Experimental Results	49
IV.2.1 The GTPase site in human eEFSec	49
IV.2.2 The interaction of human eEFSec with guanine nucleotides and analogs	51
IV.2.3 The mutational analysis of the nucleotide-binding site in human eEFSec	54
IV.2.4 The characterization of the putative Sec-binding site in human eEFSec..	57
IV.3 Discussion.....	60
V. The GTP-to-GDP exchange in human eEFSec induces a non-canonical conformational change of D4, and not a canonical rearrangement of D1.....	63
V.1.. Introduction.....	63
V.2.. Experimental Results.....	64
V.2.1 The non-canonical conformational change in human eEFSec induced by the GTP-to-GDP exchange.....	64
V.2.2 Minor structural rearrangements in the EF-Tu-like domain of eEFSec upon the GTP-to-GDP exchange.....	73
V.3 Discussion.....	77
VI. The proposed mechanism of decoding of the Sec UGA codon.....	80
VI.1. Introduction.....	80
VI.2. Experimental Results.....	81
VI.2.1 A homology-based model of the eEFSec:tRNA ^{Sec} complex.....	81
VI.2.2 The conservation of the putative tRNA-recognition elements in eEFSec and SelB.....	85
VI.2.3 The model of human eEFSec:GDPCP complex with tRNA ^{Sec}	87
VI.3 Discussion.....	89
VII. CONCLUSIONS.....	93
VIII. FUTURE STUDIES.....	95
CITED LITERATURE.....	96
APPENDICES.....	109
APPENDIX A.....	109
APPENDIX B.....	111

VITA.....133

LIST OF TABLES

<u>TABLE</u>	<u>PAGE</u>
I. PRIMERS USED TO CLONE His-Avi-TEV INTO HUMAN eEFSec-pET15b	19
II. PRIMERS USED TO CREATE MUTANT eEFSec CONSTRUCTS	21
III. DATA COLLECTION AND REFINEMENT STATISTICS FOR eEFSec COMPLEXES WITH THE NON-HYDROLYZABLE GTP ANALOGS	32
IV. BINDING OF GUANINE NUCLEOTIDES AND ANALOGS TO HUMAN eEFSec.....	54
V. BINDING OF GUANINE NUCLEOTIDES TO GTPASE SITE MUTANTS OF HUMAN eEFSec	56
VI. DATA COLLECTION AND REFINEMENT STATISTICS FOR eEFSec:GDP	65
VII. STRUCTURAL ELEMENTS IN EF-Tu INVOLVED IN BINDING TO aa-tRNA ...	84
VIII. THE PUTATIVE STRUCTURAL ELEMENTS IN HUMAN eEFSec AND ITS ORTHOLOGS THAT MIGHT BE INVOLVED IN BINDING OF Sec-tRNA ^{Sec}	87

LIST OF FIGURES

<u>FIGURE</u>	<u>PAGE</u>
1. The synthetic pathway of Sec	4
2. The co-translational incorporation of an amino acid into a nascent protein chain.....	5
3. Secondary structure diagrams demonstrate differences between tRNA ^{Sec} and canonical tRNAs	7
4. Schematic representation of the secondary structure of SECIS elements.....	11
5. GTP and its analogs.....	30
6. The overall structure and domain organization of human eEFSec.....	34
7. The GTP analogs, GDPCP and GDPNP, induce the same conformation of human eEFSec.....	35
8. The GTP analogs, GDPCP and GDPNP, lock eEFSec in the GTP-bound state.....	37
9. The superimposition of human eEFSec:GDPCP onto the GDP-bound state of rabbit general translation elongation factor eEF1A2.....	38
10. The structural comparison of human eEFSec with SelB orthologs.....	41
11. The structure of the 'hinge' region in human eEFSec.....	42
12. Mutational analysis of the 'hinge' region in eEFSec.....	43
13. Structural conservation between human eEFSec and the universal translation initiation factor IF2/eIF5B.....	45
14. The structure of the GTPase site in human eEFSec.....	51

15.	The binding of guanine nucleotides and analogs to human eEFSec characterized by ITC.....	53
16.	The binding of guanine nucleotides to GTPase site mutants of human eEFSec characterized by ITC.....	55
17.	The <i>in vitro</i> Sec incorporation activity of the GTPase site mutants of human eEFSec.....	57
18.	The putative Sec-binding site in human eEFSec.....	59
19.	The GTP-to-GDP transition in human eEFSec induces a conformational change in D4, but not in D1.....	67
20.	Small-angle X-ray scattering (SAXS) analysis of the GTP- and GDP-bound states of human eEFSec.....	69
21.	The superimposition of SAXS molecular envelopes onto the corresponding crystal structures of the GTP- and GDP-bound states of eEFSec.....	70
22.	The comparison of the ‘canonical’ model of the GDP-bound state of human eEFSec with eEFSec:GDP SAXS envelope.....	71
23.	An enlarged loop in D3 might prevent the canonical conformational change in the EF-Tu-like domain of human eEFSec upon the GTP-to-GDP exchange	72
24.	Structure-based sequence alignment of eEFSec and SelB orthologs.....	73
25.	The structural rearrangements in the GTPase site in human eEFSec upon the GTP-to-GDP transition.....	75
26.	The structural rearrangements within the putative Sec-binding pocket of human eEFSec upon the GTP-to-GDP exchange.....	76

27.	A comparison between human eEFSec:GDPCP and the EF-Tu:GDPNP:Phe-tRNA ^{Phe} ternary complex.....	85
28.	A model of the human eEFSec:GDPCP:tRNA ^{Sec} ternary complex	89
29.	A model of decoding of the Sec UGA codon by human eEFSec.....	92

LIST OF ABBREVIATIONS

6xHis	Hexa-histidine tag
A	Adenine
aaRS	Aminoacyl-tRNA synthetase
aa-tRNA	Aminoacyl-tRNA
Ala	Alanine (one letter code: A)
ANL	Argonne National Laboratory
APCF	Advanced Protein Characterization Facility
APS	Advanced Photon Source
Asn	Asparagine (one letter code: N)
Asp	Aspartic acid (one letter code: D)
Arg	Arginine (one letter code: R)
ATP	Adenosine triphosphate
AU	Asymmetric unit
Avi	Biotinylation tag
BioCAT	Biophysics Collaborative Access Team
birA	Biotin ligase
β -ME	β -mercaptoethanol
bp	base pair
C	Cytosine
Cys	Cysteine (one letter code: C)
DIO	Iodothyronine deiodinase
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dNTP	Deoxynucleotide
E.coli	Escherichia coli
eEFSec	Eukaryotic elongation factor specific for selenocysteine
<i>EEFSEC</i>	Gene encoding eEFSec
eEF1	Eukaryotic translation elongation factor 1
eEF2	Eukaryotic translation elongation factor 2
EF-G	Elongation factor G
EF-Ts	Elongation factor thermo stable
EF-Tu	Elongation factor thermo unstable
EF1A	Elongation factor 1 subunit α
eIF5B	Eukaryotic translation initiation factor 5B
FOM	Figure of merit
G	Guanine
GDP	Guanosine-5'-diphosphate
GDPCP	Guanosine-5'-[(β , γ)-methylene]-triphosphate
GDPNP	Guanosine-5'-[(β , γ)-imido]-triphosphate
GEF	Guanine nucleotide exchange factor
Glu	Glutamic acid (one letter code: E)

LIST OF ABBREVIATIONS (continued)

Gln	Glutamine (one letter code: Q)
Gly	Glycine (one letter code: G)
GPx	Glutathione peroxidase
GTP	Guanosine-5'-triphosphate
GTPase	Enzyme that binds and hydrolyzes GTP
GTP _γ S	Guanosine-5'-(γ-thio)-triphosphate
HCl	Hydrochloric acid
H-bond	Hydrogen bond
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His	Histidine (one letter code: H)
IF2	Translation initiation factor 2
Ile	Isoleucine (one letter code: I)
IPTG	Isopropyl-β-D-1-thiogalactopyranoside
ITC	Isothermal titration calorimetry
KCl	Potassium chloride
K _d	Dissociation constant
KI	Potassium iodide
LB	Luria-Bertani medium
Leu	Leucine (one letter code: L)
LS-CAT	Life Sciences Collaborative Access Team
Lys	Lysine (one letter code: K)
mcm ⁵ U	Methylcarboxymethyl-5'-uridine
mcm ⁵ Um	Methylcarboxymethyl-5'-uridine 2'-O-methylribose
mRNA	Messenger RNA
Mg ²⁺	Magnesium (II) ion
MgCl ₂	Magnesium (II) chloride
Mn ²⁺	Manganese (II) ion
MnCl ₂	Manganese (II) chloride
NaI	Sodium iodide
NaCl	Sodium chloride
NH ₂	Amino group
Ni ²⁺	Nickel (II) ion
OD	Optical density
OH	Hydroxyl group
PAGE	Polyacrylamide gel electrophoresis
pKa	Acid dissociation constant
P(r)	Pair-distance distribution function
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
Phe	Phenylalanine (one letter code: F)
P loop	Phosphate-binding loop
PLP	Pyridoxal phosphate

LIST OF ABBREVIATIONS (continued)

PSTK	O-phosphoseryl-tRNA ^{Sec} kinase
PTC	Peptidyl-transferase center
RBD	RNA-binding domain
R _{free}	Residual factor for the subset of crystallographic data
R _g	Radius of gyration
r.m.s.d.	Root-mean-square deviation
RNA	Ribonucleic acid
R _{work}	Residual factor for the crystallographic data set
rRNA	Ribosomal RNA
SAD	Single-wavelength anomalous dispersion
SAXS	Small-angle X-ray scattering
SBC	Structural Biology Center
SBP2	SECIS-Binding Protein 2
SDS	Sodium dodecyl sulfate
Se	Selenium
Sec	Selenocysteine
SEC	Size-exclusion chromatography
SECIS	Selenocysteine insertion sequence
<i>SECISBP2</i>	Gene encoding SBP2
Sec-tRNA ^{Sec}	Selenocysteinyl-tRNA ^{Sec}
SelA	Selenocysteine synthase (bacteria)
SelB	Selenocysteine tRNA-specific elongation factor (bacteria/archaea)
SelH	Selenoprotein H
SelN	Selenoprotein N
SelP	Selenoprotein P
SeMet	Selenomethionine
<i>SEPN1</i>	Gene encoding SelN
SepSecS	O-phosphoseryl-tRNA ^{Sec} :selenocysteinyl-tRNA ^{Sec} synthase
Sep-tRNA ^{Sec}	Phosphoseryl-tRNA ^{Sec}
Ser	Serine (one letter code: S)
SerRS	Seryl-tRNA synthetase
Ser-tRNA ^{Sec}	Seryl-tRNA ^{Sec}
SID	Selenocysteine incorporation domain
SLiCE	Seamless ligation cloning extract
SPS2	Selenophosphate synthetase 2
SPR	Surface plasmon resonance
SRL	Sarcin-ricin loop
T	Thymine
TB	Terrific Broth
TCEP	Tris(2-carboxyethyl)phosphine
TEV	Tobacco Etch Virus protease
Thr	Threonine (one letter code: T)

LIST OF ABBREVIATIONS (continued)

Tris	Tris(hydroxymethyl)aminomethane
tRNA	Transfer RNA
tRNA ^{Phe}	Phenylalanine tRNA
tRNA ^{Sec}	Selenocysteine tRNA
tRNA ^{Ser}	Serine tRNA
Trp	Tryptophan (one letter code: W)
TrxR	Thioredoxin reductase
Tyr	Tyrosine (one letter code: Y)
U	Uracil
UTR	Untranslated region
WT	Wild type
w/v	Weight/volume
Val	Valine (one letter code: V)
ΔH	Enthalpy change
ΔS	Entropy change
ψ	Pseudouridine

SUMMARY

Selenium is the only essential micronutrient that is “genetically encoded” in all domains of life. It is found in proteins as the 21st amino acid selenocysteine. Mammals, and humans, have 25 ubiquitously expressed selenoproteins, many of which are essential. Selenoproteins and selenoenzymes are critical for the redox potential maintenance, protection of the genetic material and cell membrane from oxidative damage, regulation of the thyroid hormone metabolism, and control of gene expression and protein folding. The replacement of selenocysteine with either serine or cysteine renders selenoenzymes either completely inactive or significantly catalytically impaired. A tRNA^{Sec} knockout mutant mouse is embryonically lethal and mutations in enzymes facilitating selenoprotein synthesis cause systemic pathologies including severe early-onset neurodegeneration. This implies the accurate decoding of the selenocysteine codon and the correct placement of the selenocysteine residue within the nascent selenoprotein chain is a fundamental biological process. However, the complex biosynthesis of eukaryotic selenoproteins is still poorly understood.

Selenocysteine is the only proteinogenic amino acid encoded by an in-frame UGA codon that does not operate as the canonical translational stop codon. A specialized translation elongation factor, eEFSec in eukaryotes and SelB in prokaryotes, promotes incorporation of selenocysteine into selenoproteins by a still poorly understood mechanism. Here, I determined the crystal structures of the major physiological states of human eEFSec and characterized the functional sites of this important elongation factor. My results reveal that four domains of human eEFSec fold

into a chalice-like structure that resembles its archaeal ortholog. Surprisingly, unlike in EF-Tu and EF1A, the GTP-to-GDP exchange does not cause a major conformational rearrangement of domain 1 of eEFSec, but instead induces a movement of the appended domain 4. Furthermore, the binding studies show that eEFSec has similar binding affinities for GDP, GTP and GTP analogs, whereas the results of activity assays reveal that the mutations within GTPase and Sec-binding sites abolish eEFSec ability to promote selenocysteine incorporation *in vitro*. Based on my findings, I propose that eEFSec utilizes a non-canonical mechanism for release of the selenocysteinyl-tRNA during decoding of the selenocysteine codon.

THESIS ORGANIZATION

This PhD thesis is organized into eight chapters. Chapter I serves as an introduction of the thesis topic, demonstrates its physiological significance, and provides related background information. Chapter II describes experimental procedures and materials used to generate results presented and discussed in Chapters III-VI. Chapter VII summarizes the main conclusions of the work presented in this thesis, while Chapter VIII suggests the subsequent studies that would further expand the knowledge of the thesis subject.

In Chapter III, I present the crystal structures of human selenocysteine tRNA-specific elongation factor, eEFSec, in complex with the non-hydrolyzable GTP analogs, GDPNP and GDPCP. I describe the overall structure and domain organization of eEFSec, and perform the detailed comparison with structures of the general translation elongation factors EF-Tu/EF1A, SelB orthologs, and the universal translation initiation factor IF2/eIF5B.

Chapter IV focuses on the main functional sites of eEFSec: the GTPase site and the Sec-binding pocket. I describe the composition of each site as well as the results of mutational studies and *in vitro* activity assays. Also, I analyze interactions of eEFSec with guanine nucleotides and analogs.

In Chapter V, I report the crystal structure of the GDP-bound state of human eEFSec, and provide a detailed description of the conformational changes in eEFSec coupled to the GTP-to-GDP exchange. To confirm my findings, I analyze the GTP- and GDP-bound states of eEFSec in solution using Small-angle X-ray scattering (SAXS).

Lastly, I speculate about the structural elements in eEFSec that could play a role in preventing the canonical conformational change observed in EF-Tu/EF1A.

In Chapter VI, I propose the mechanism of decoding of the Sec UGA codon in humans. I present a model of the complex between human eEFSec:GDPCP and selenocysteine tRNA (tRNA^{Sec}), generated based on the known structure of the EF-Tu:GDPNP:Phe-tRNA^{Phe} complex. Furthermore, I analyze the structural elements involved in the tRNA binding that are conserved between human eEFSec, its orthologs and EF-Tu. I also suggest additional features present in the specialized but absent from general elongation factors, which might play a role in recognition of tRNA^{Sec}.

I. BACKGROUND AND SIGNIFICANCE

I.1 The physiological role of selenium

Selenium was discovered at the beginning of the 19th century by the Swedish chemist Berzelius, and named after Hellenic Moon Goddess Selene (Berzelius J 1818). The name was given due to chemical similarities with tellurium whose name was derived from a Latin word for Earth, tellus. Its significance as an essential micronutrient was established in mid-20th century (1, 2). The maintenance of selenium homeostasis is crucial for the development, survival and health of the organism. Excessive intake of selenium was shown to be toxic (3) and can even lead to cancer (4), whereas its deficiency has been linked to a variety of disorders and diseases such as male infertility (5), Keshan disease (6) and cretinism (7).

Selenium exerts its physiological role mainly in a form of the 21st genetically encoded amino acid - selenocysteine (Sec), which is as a constitutive component of a small group of proteins in all domains of life (8-11). For instance, the largest selenoproteome composed of 30 selenoproteins was identified in fish, while humans and rodents have 25 and 24 selenoproteins, respectively (12). However, in yeast and higher plants cysteine-containing proteins substitute for selenoproteins (13, 14).

Selenoproteins and selenoenzymes are involved in a variety of cellular processes such as the maintenance of selenium homeostasis (selenoprotein P, SelP) and cellular redox potential (thioredoxin reductases, TrxR), the removal of the reactive oxygen species (glutathione peroxidases, GPx), regulation of the thyroid hormone

activation/deactivation (iodothyronine deiodinases, DIO), gene expression (selenoprotein H, SelH) and protein folding (selenoprotein N, SelN) (15-17). However, the function of some identified Sec-containing proteins still remains unknown. Whereas selenoproteins typically contain one Sec residue important for their structure and fold, SelP harbors multiple Sec residues, importance of which is not clear. On the other hand, selenoenzymes harbor a single Sec residue, located in the active site and pivotal for their catalytic activity (18). The replacement, either accidental or deliberate, of Sec with either serine (Ser) or cysteine (Cys) was shown to cause misfolding of selenoproteins, and either complete inactivation or significant loss of catalytic activity in selenoenzymes (19-21). Therefore, the proper co-translational insertion of Sec into nascent protein chains is critical for their structure and function. A number of mutations have been identified in enzymes facilitating selenoprotein synthesis that lead to a variety of systemic pathologies (17, 22) including severe early-onset neurodegeneration (23, 24). However, the complex biosynthesis of eukaryotic selenoproteins is still a poorly understood process (25).

1.2 The distinct features of selenocysteine

Sec is distinct when compared to the 20 standard amino acids. Firstly, it is the only proteinogenic amino acid that contains an essential dietary microelement. Secondly, it is encoded by UGA codon (note: recent findings suggest that Sec is encoded by UAA, UAG and unused sense codons in microorganisms (26)), which otherwise signals the end of protein synthesis (27). Thirdly, Sec has a lower pKa value

for its side chain (5.3) than Cys (8.47), which makes it stronger nucleophile under physiological pH, and thus, highly reactive. This perhaps explains why there is no free pool of Sec in the cytosol (11). Moreover, Sec is the only proteinogenic amino acid that lacks the cognate aminoacyl-tRNA synthetase (aaRS). Therefore, a complex pathway has evolved to synthesize Sec on its cognate tRNA (tRNA^{Sec}) from a Ser precursor in all domains of life (28, 29).

The synthetic pathway of Sec comprises of several steps. In the first step, which is conserved in all organisms, seryl-tRNA synthetase (SerRS) charges tRNA^{Sec} with Ser yielding seryl-tRNA^{Sec} ($\text{Ser-tRNA}^{\text{Sec}}$). In the next step, $\text{Ser-tRNA}^{\text{Sec}}$ is converted into $\text{Sec-tRNA}^{\text{Sec}}$. The mechanism by which this process is achieved differs between bacteria and archaea/eukaryotes. In bacteria, the conversion of $\text{Ser-tRNA}^{\text{Sec}}$ to $\text{Sec-tRNA}^{\text{Sec}}$ is performed by one enzyme – homodecameric SelA (30). In contrast, in archaea and eukaryotes, the Ser-to-Sec conversion occurs in two steps (Figure 1). First, the seryl moiety of $\text{Ser-tRNA}^{\text{Sec}}$ is phosphorylated by *O*-phosphoseryl-tRNA^{Sec} kinase (PSTK) (31). The phosphoseryl-tRNA^{Sec} ($\text{Sep-tRNA}^{\text{Sec}}$) intermediate is then used as a substrate along with selenophosphate and a cofactor pyridoxal phosphate (PLP), in the terminal synthetic reaction catalyzed by *O*-phosphoseryl-tRNA^{Sec}:selenocysteinyl-tRNA^{Sec} synthase (SepSecS). SepSecS substitutes phosphoryl with selenol yielding $\text{Sec-tRNA}^{\text{Sec}}$ (28). Selenophosphate, a major selenium donor, is formed from selenide by selenophosphate synthetase 2 (SPS2) in the ATP-dependent manner (13).

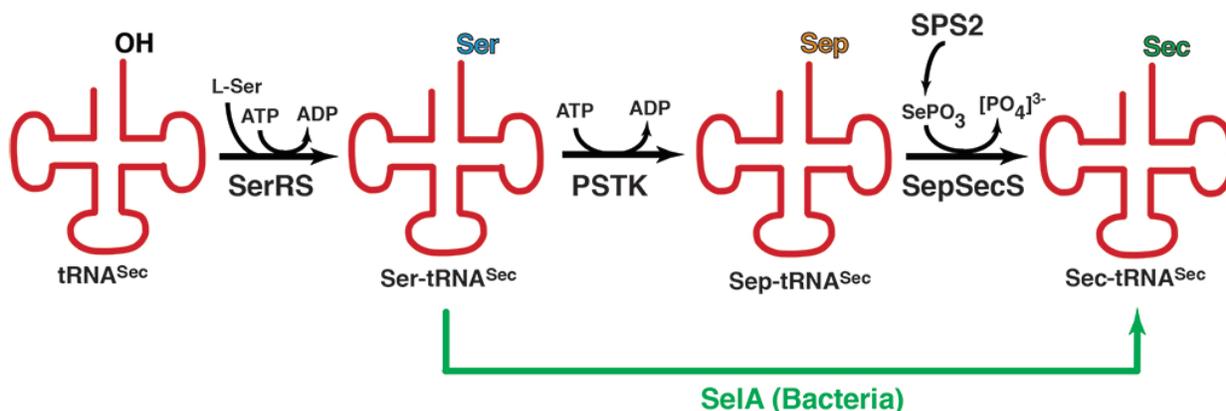


Figure 1. The synthetic pathway of Sec. Sec is synthesized on its own tRNA from a serine precursor. While the first step in the Sec synthetic pathway, which yields Ser-tRNA^{Sec}, is preserved in all domains of life, the following step, resulting in formation of Sec-tRNA^{Sec}, differs between bacteria and archaea/eukaryotes. In archaea and eukaryotes, the conversion of Ser moiety into Sec is achieved by the action of two enzymes, PSTK and SepSecS, whereas one enzyme (SclA) promotes Ser-to-Sec conversion in bacteria. A major selenium donor, selenophosphate (SePO₃), is formed from selenide by selenophosphate synthetase 2 (SPS2) in the ATP-dependent manner.

Lastly, in contrast to standard amino acids and pyrrolysine, which are all recognized and delivered to the ribosome by a general translation elongation factors EF-Tu or EF1A (Figure 2A), Sec-tRNA^{Sec} requires the specialized elongation factor, SelB in prokaryotes (32, 33) and eEFSec in eukaryotes (34, 35) (Figure 2B).

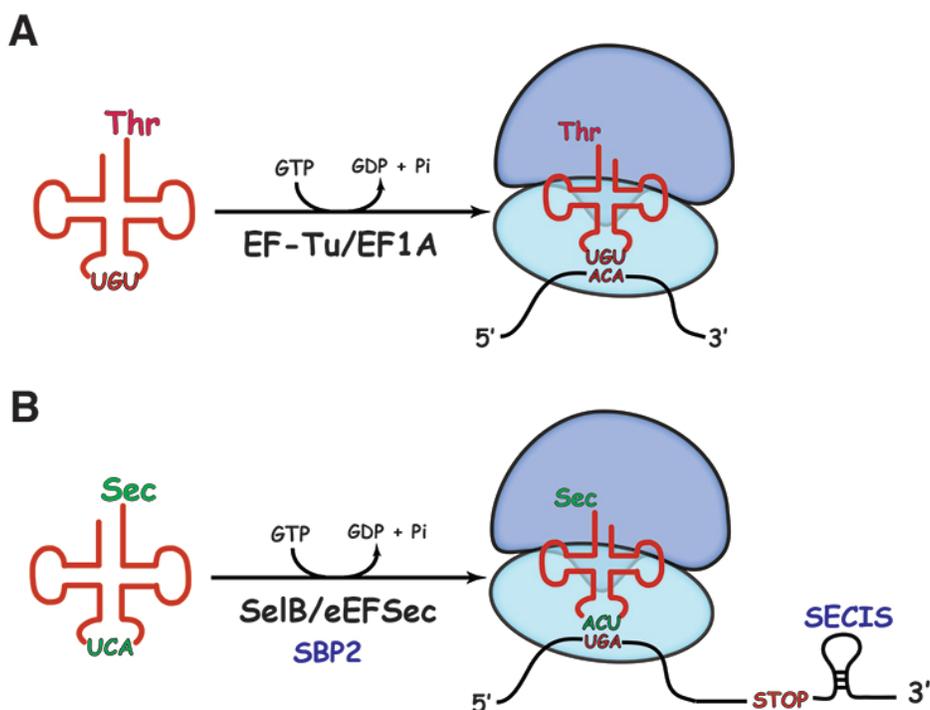


Figure 2. The co-translational incorporation of an amino acid into a nascent protein chain. (A) In a canonical pathway, aa-tRNA is recognized and delivered to the ribosome in the GTP-dependent manner by the general translation elongation factor, EF-Tu in bacteria and EF1A in archaea/eukaryotes. **(B)** Sec-tRNA^{Sec} requires SelB in bacteria/archaea or eEFSec in eukaryotes.

I.3 Selenocysteine tRNA: structure, function and role in health

Selenocysteine tRNA, denoted in the literature either as tRNA^{Sec} or tRNA^{[Ser]Sec}, has been extensively studied in various organisms. The most evident feature that distinguishes it from canonical tRNAs is its length. Bacterial tRNA^{Sec} is longer (95 nucleotides in *E.coli*, 100 nucleotides in *Moorella thermoacetica*) than a mammalian one (90 nucleotides). While a canonical tRNA adopts a 7/5 fold (*numbers represent the number of base pairs (bp) in the acceptor and T ψ C stems, respectively*), tRNA^{Sec} is characterized by 8/5 fold in bacteria and the 9/4 fold in archaea/eukaryotes (Figure 3).

Therefore, tRNA^{Sec} has 13 bp in the acceptor-T ψ C arm instead of the canonical 12 bp. The insertion of an additional bp causes the 5'- and 3' CCA-aa end of tRNA^{Sec} to be translated by 3.4 Å and rotated by ~33° when compared to the corresponding elements in other tRNAs. Additionally, tRNA^{Sec} has enlarged D- and variable arms. The D stem of tRNA^{Sec} is characterized by 6 bp instead of 3-4 bp found in canonical tRNAs. The extended variable arm, on the other hand, is found not only in tRNA^{Sec}, but also in tRNA^{Ser}, tRNA^{Tyr}, and tRNA^{Leu}. The structural divergence from canonical tRNAs might explain why tRNA^{Sec} cannot be recognized and delivered to the ribosome by general elongation factor, but requires the specialized elongation factor. Interestingly, human tRNA^{Sec} can interact with *E.coli* SelB *in vitro* and even complement for *E. coli* tRNA^{Sec} suggesting a certain level of structural conservation across species (36).

despite selenium deficiency (37). Both the amount and distribution of these isoforms varies in different tissues.

Several mouse models have illustrated the importance of tRNA^{Sec} for selenoprotein synthesis and the organism health. The most dramatic evidence was the embryonically lethal knockout mutant mouse (38). The conditional deletion of tRNA^{Sec} gene in the mammary epithelium was shown to reduce the selenoprotein level (39), whereas in liver it leads to the organ degeneration (40).

I.4 Selenocysteine tRNA-specific elongation factor

The co-translational insertion of Sec into a nascent selenoprotein is promoted by a specialized elongation factor, SelB in prokaryotes (32, 33) and eEFSec in eukaryotes (34, 35). SelB/eEFSec is a translational GTPase that binds Sec-tRNA^{Sec} with high affinity and stringent specificity, and delivers it to the site of translation in response to a particular in-frame UGA codon. The overall sequence identity between these protein factors is low (18-25%), and even lower (~13%) when only the C-terminal domain is used in calculation.

Studies on the archaeal (41) and bacterial orthologs (42) revealed that SelB is composed of four domains. The N-terminal part is composed of three domains (D1-3) resembling EF-Tu. In contrast, the C-terminal domain 4 (D4), which does not have its equivalent in EF-Tu and EF1A, is structurally divergent. The archaeal D4 adopts a β -barrel shape reminiscent of the oligonucleotide-binding (OB) fold and it is appended to the EF-Tu-like domain *via* a long linker region, while the bacterial D4 is formed by four

winged-helix motifs and is positioned in a completely different way relative to the EF-Tu-like domain. The appended D4 has been suggested to play a role in binding of Sec-tRNA^{Sec} across kingdoms (41, 43, 44). However, whereas the bacterial D4 directly interacts with the stem loop structure in the selenoprotein mRNA (45, 46), the archaeal and eukaryotic domains do not perform such function (see: Chapter I.5).

In contrast to EF-Tu and EF1A, both SelB and eEFSec have similar binding affinities for GTP and GDP (34, 35, 47). Also, both proteins lack sequences corresponding to the EF-Tu/EF1A elements involved in the interaction with the guanine nucleotide exchange factor (GEF). Therefore, it is thought that SelB and eEFSec do not require the GEF activity to cycle between the GDP- and GTP-bound states. Moreover, the low intrinsic GTPase activity of SelB is stimulated by the ribosome and selenoprotein mRNA (48). By contrast, eEFSec completely lacks an intrinsic GTPase activity in the absence of the ribosome. Interestingly, the enhancement of the GTPase activity was not observed even when all components of the Sec codon decoding machinery were present in the reaction mixture (44). These observations raised a question whether the specialized elongation factors promote Sec incorporation into a nascent selenoprotein by a mechanism distinct from that utilized by EF-Tu/EF1A.

I.5 Other factors involved in decoding of the selenocysteine UGA codon

Although a specialized elongation factor and Sec-tRNA^{Sec} play an important role in the process of Sec incorporation into selenoproteins, they are not sufficient for the read-through of an in-frame UGA codon. An essential RNA factor that serves to

differentiate the Sec UGA from the *opal* stop codon in all domains of life is a hairpin structure in the selenoprotein mRNA termed SelenoCysteine Insertion Sequence (SECIS) (49). In contrast to the prokaryotic system, the decoding of the Sec codon in eukaryotes is more complex and it requires SECIS-Binding Protein 2 (SBP2) and other RNA factors (50-52). Herein, I shall describe two most significant factors critical for Sec incorporation: the SECIS element and SBP2.

I.5.1 SelenoCysteine Insertion Sequence (SECIS)

The bacterial SECIS element is located within the coding region immediately downstream of the Sec UGA (53), while in archaea and eukaryotes, the poorly conserved SECIS is located in the 3'-untranslated region (UTR) of the selenoprotein mRNA (54). The sequence and structure of SECIS elements differ between the kingdoms (Figure 4). In eukaryotes, SECIS is composed of an apical loop, sometimes followed by additional mini-stem and bulge, and two helical stems that are separated by an internal loop (Figure 4C). Most importantly, eukaryotic SECIS contains three conserved sequence motifs essential for the decoding: the core formed by two non-Watson-Crick G-A base pairs, the AAR or CCR motif in the apical loop, and G or A nucleotide directly preceding the core of SECIS element.

The physiological significance of SECIS has been demonstrated by homozygous point mutation within the core of the SECIS element in the mRNA of SelN, which severely impairs the expression level *in vivo*, and consequently leads to SEPNI-related myopathy (55).

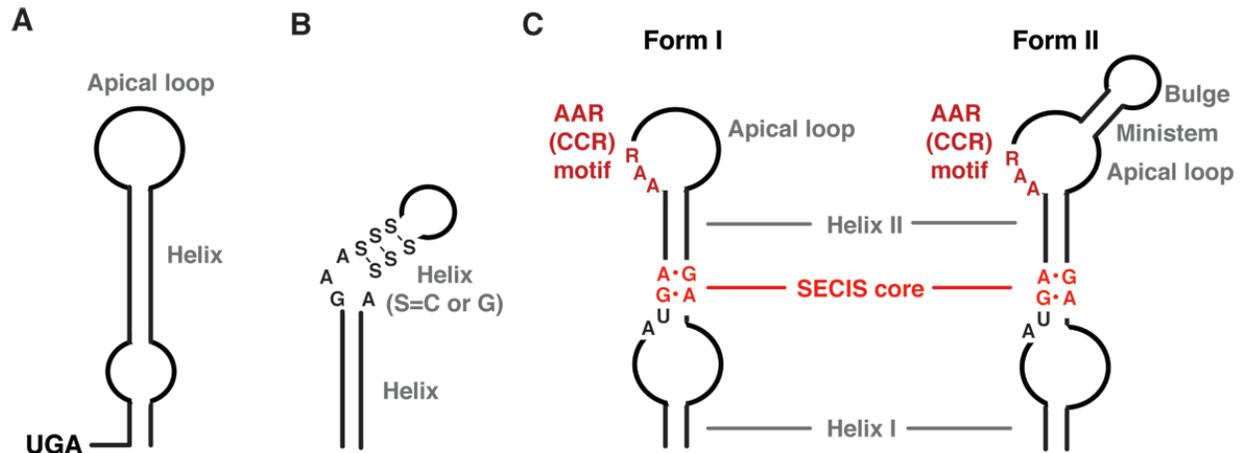


Figure 4. Schematic representation of the secondary structure of SECIS elements. Besides different location in selenoprotein mRNA, the SECIS elements from bacteria (A), archaea (B), and eukaryotes (C) differ in both sequence and structure. Moreover, the eukaryotic SECIS exists in two forms, depending on whether the apical loop contains an additional mini-stem followed by a bulge or not.

I.5.2 SECIS-Binding Protein 2 (SBP2)

Unlike the bacterial SelB, eEFSec cannot directly interact with SECIS. Instead, it requires an *in-trans* protein factor SBP2, which is composed of 854 amino acids. Currently, no structural information about SBP2 is available, most likely due to the fact that ~70% of SBP2 is disordered in the absence of its interaction partners (56). It is predicted that the protein is composed of three domains: the dispensable N-terminal domain (~400 aa), the central Selenocysteine Incorporation Domain (SID) that is critical for binding to eEFSec, and the C-terminal RNA-Binding Domain (RBD) that binds to SECIS and the ribosome through its L7Ae RNA-binding motif (57, 58). Both SID and RBD domains have been shown to be required for decoding of the Sec UGA codon. It is assumed that SBP2 serves to recruit eEFSec and Sec-tRNA^{Sec} near the ribosome,

which could be important for avoiding the premature termination of translation. However, it is not established whether eEFSec ternary complex interacts with already formed SBP2-SECIS assembly, or whether it first binds to SBP2 and then to SECIS. Further, the importance of SBP2 for human health is illustrated by the fact that mutations in *SECISBP2* lead to a variety of systemic diseases and disorders. For instance, the homozygous mutation of Arg531, the residue located within the RNA-binding domain of SBP2, to glutamine (R531Q), abolishes the binding of SBP2 to the SECIS element of iodothyronin deiodinase 1 (DIO 1). Consequently, the synthesis of DIO 1 is affected, causing abnormal thyroid hormone metabolism and delayed bone growth (59). The nonsense K438* mutation, carried on a paternal allele, yields a truncated SBP2 devoid of the RBD domain, which causes abnormal thyroid hormone metabolism, decreased activity of GPx3, and reduced serum levels of SeIP (60).

I.6 The elongation phase of protein synthesis

The mechanism of elongation phase of protein synthesis is highly conserved between bacteria and eukaryotes. In bacteria, elongation factors Tu, G and Ts promote the process. In eukaryotes, eEF1 substitutes for EF-Tu and EF-Ts, while eEF2 is analogous to EF-G. The current knowledge about the mechanism of elongation phase of protein synthesis comes mainly from the structural studies done on the bacterial model system (61-67).

The crystal structure of the *Thermus aquaticus* EF-Tu in complex with a non-hydrolyzable GTP analog, GDPNP, revealed that the protein adopts a globular shape

and is composed of three domains, D1-3 (62). The subsequent structural studies on the EF-Tu:GDPNP:Phe-tRNA^{Phe} (63) and EF-Tu:GDP (64) complexes revealed that binding of GTP to D1 induces a conformation in which the aminoacyl-binding pocket, located at the interface of D1 and D2, is properly formed. Therefore, only in the GTP-bound state can EF-Tu bind and deliver canonical aminoacyl-tRNAs (aa-tRNAs) to the ribosomal A site. The ternary complex has an elongated corkscrew shape where EF-Tu and the acceptor-T ψ C arm of aa-tRNA constitute a 'knob-like handle', while the anticodon arm composes a 'screw'. The 3' CCA-aa end of aa-tRNA is bound at the interface of D1 and D2, whereas the junction of all three EF-Tu domains forms a binding site for the 5'-end. D3 makes contacts with T-stem of aa-tRNA.

On the ribosome, the correct codon-anticodon base pairing and interaction with the sarcin-ricin loop (SRL) of the large ribosomal subunit stimulates the GTPase activity of EF-Tu and induces a major rearrangement of the protein factor structure. The $\sim 90^\circ$ rotation of D1 relative to D2 and D3 (61, 64) leads to a dramatic transformation of the globular structure of the GTP-bound state into a more elongated conformation of the GDP-bound state. Consequently, the amino acid-binding pocket and the tRNA-recognition surface are disrupted, and EF-Tu:GDP dissociates from aa-tRNA and the ribosome. The acceptor stem of the aa-tRNA accommodates in the A-site on the large ribosomal subunit and the aminoacyl group becomes properly positioned within the peptidyl-transferase center (PTC) for the reaction of peptide bond formation to occur (65, 66). After a peptide bond is formed, an extended peptidyl-tRNA is bound to the A site and the unacylated tRNA is bound to the P site. The EF-G translocase binds and

promotes translocation of the A-site peptidyl-tRNA and the P-site tRNA in a GTP-dependent manner. The shape of EF-G mimics the EF-Tu:aa-tRNA ternary complex; the arrangement of D1-3 resembles the GTP-bound state of EF-Tu, while the extended D4 is reminiscent of the anticodon arm of tRNA (68). Before the next round of translation elongation can occur, the GDP molecule bound to EF-Tu has to be exchanged for GTP. The nucleotide exchange is achieved by the elongation factor EF-Ts (69, 70).

Besides the knowledge of the protein and RNA factors involved in decoding of the Sec UGA codon, and a general scheme of the occurring events, the exact mechanism of the process is poorly understood. The current working model describing the co-translational incorporation of Sec is in part based on the analogy to the canonical mechanism for incorporation of standard amino acids, as well as on a number of biochemical studies done on the components of the Sec pathway. According to the model, in the first step the GTP-bound state of SelB/eEFSec selects Sec-tRNA^{Sec}. Then, the ternary SelB/eEFSec:GTP:Sec-tRNA^{Sec} complex is tethered near the translating ribosome. In bacteria, the delivery of Sec-tRNA^{Sec} to the ribosome occurs *via* a direct interaction between D4 of SelB and SECIS element in the selenoprotein mRNA. The corresponding step in the eukaryotic system requires SBP2 to act as a bridge between the eEFSec ternary complex and SECIS, but it is not well understood. Presumably, after the codon-anticodon interactions are established on the small ribosomal subunit, SelB/eEFSec hydrolyzes GTP and consequently releases Sec-tRNA^{Sec} and dissociates from the ribosome. Sec is then incorporated into a nascent selenoprotein chain.

To fully understand the mechanism by which Sec is incorporated into selenoproteins, further studies are needed in both the prokaryotic and eukaryotic systems. In this thesis I describe my contribution towards understanding of the mechanism of decoding of the Sec UGA codon in humans. Given that this process is more complex than the one in bacteria, I focused my structural and biochemical studies on human eEFSec.

II. MATERIALS AND METHODS

II.1 Cloning, expression, and purification of recombinant human eEFSec

II.1.1 Wild-type eEFSec

Human eEFSec gene was cloned into pET15b vector (Novagen) with an N-terminal 6xHis tag followed by a thrombin cleavage site. The plasmid was a gift from the laboratory of Prof. Dieter Söll (Departments of Molecular Biophysics and Biochemistry, and Chemistry, Yale University, New Haven, CT). The plasmid was transformed into *E. coli* Lemo21(DE3) competent cells (New England BioLabs, Inc.) following the manufacturer's protocol. The cells were grown at +37°C at 200 rpm in 8 L Luria-Bertani (LB) medium, supplemented with 0.1 mg/mL ampicillin, 0.025 mg/mL chloramphenicol, and 0.2 mM L-rhamnose. Once OD measured at 600 nm (OD_{600}) reached ~0.6 units, the overnight expression at +16°C was induced with 0.4 mM isopropyl β -D-1-thiogalactopyranoside (IPTG). The cells were harvested by centrifugation at 6,000 rpm at +4 °C and resuspended in the lysis buffer composed of 50 mM HEPES pH 7.5, 300 mM NaCl, 10% glycerol, 5 mM β -mercaptoethanol (β -ME). The buffer also contained a tablet of the protease inhibitor cocktail (Roche). After sonication, the lysate was cleared by centrifugation at 18,000 rpm at +4 °C for 40 min. The soluble fraction was loaded onto a HisTrap FF Crude column (GE Healthcare). The non-specifically bound protein was removed with the lysis buffer supplemented with 10 and 60 mM imidazole. The recombinant eEFSec was eluted with the lysis buffer supplemented with 300 mM imidazole. The eluate was diluted 3-fold in 20 mM HEPES pH 7.5, 50 mM NaCl, 0.5 mM

tris(2-carboxyethyl)phosphine (TCEP), and loaded onto the HiTrap SP HP ion-exchange column (GE Healthcare). eEFSec was eluted with a linear gradient of NaCl (0.1 – 1.0 M) in 20 mM HEPES pH 7.5, 0.5 mM TCEP. Fractions containing eEFSec were pooled and further purified on the HiLoad 16/600 Superdex 200 size-exclusion chromatography (SEC) column (GE Healthcare) in 20 mM HEPES pH 7.5, 150 mM NaCl, 0.5 mM TCEP, 5 mM MgCl₂ (or 5 mM MnCl₂). The average purification yield was ~0.7 mg of eEFSec per 1 L of culture. Pure eEFSec was concentrated to ~8 mg/mL and filtered through 0.22 μm filter. Protein aliquots were flash-frozen in liquid nitrogen and stored at -80°C.

II.1.2 The selenomethionine-labeled human eEFSec

The selenomethionine (SeMet)-labeled eEFSec (SeMet-eEFSec) was expressed by metabolic inhibition method. The Lemo21(DE3) cells harboring human eEFSec in pET15b (see section II.1.1) were grown at +37°C in 8L of M9 minimal medium supplemented with 0.4% glucose, 2 mM MgSO₄, 0.1 mM CaCl₂, 0.00005% thiamine-HCl, 0.04 g/L of each standard amino acid (except methionine), 0.1 mg/mL ampicillin, 0.025 mg/mL chloramphenicol, and 0.2 mM L-rhamnose. Once OD₆₀₀ reached ~0.6 units, the solution containing 0.12 g/L selenomethionine, 0.1 g/L threonine, 0.1 g/L phenylalanine, 0.1 g/L lysine, 0.05 g/L leucine, 0.05 g/L isoleucine, and 0.05 g/L valine, was added to the bacterial culture and the overnight expression at +16°C was induced with 0.4 mM IPTG. The cells were harvested and SeMet-eEFSec was purified using the same protocol as for the native eEFSec. The only difference was that higher concentration of reducing agents was used in buffers (e.g. 10 mM β-ME and 5 mM

TCEP instead of 5 mM and 0.5 mM, respectively). The average purification yield was ~0.6 mg of SeMet-eEFSec per 1 L of culture. Pure SeMet-eEFSec was concentrated to ~8 mg/mL and filtered through 0.22 μ M filter. Protein aliquots were flash-frozen in liquid nitrogen and stored at -80°C. Incorporation of 16 selenium atoms was confirmed by the mass spectrometry analysis using the Mass Spectrometry, Metabolomics and Proteomics Facility at UIC.

II.1.3 His-Avi-TEV-eEFSec

The part of His-Avi-TEV-eEFSec construct containing N-terminal 6xHis tag followed by a 15 amino acid long Avi tag (GLNDIFEAQKIEWHE) and TEV protease cleavage site, was made in polymerase chain reaction (PCR) using overlapping forward and reverse primers that were obtained from Integrated DNA Technologies (TABLE I). Following the digestion with NcoI and NdeI restriction enzymes (New England BioLabs, Inc.), His-Avi-TEV insert was cloned into pET15b vector containing human eEFSec gene. The replacement of thrombin cleavage site with Avi tag and TEV site was confirmed by DNA sequencing (DNA Services Facility at UIC).

TABLE I. PRIMERS USED TO CLONE His-Avi-TEV INTO HUMAN eEFSec-pET15b

Primer	Nucleotide sequence
Forward	5' ggagatataccatgggcagcagccatcatcatcatcacagcagcggcctgaacgacatcttc 3'
Reverse	5' aaacatatgaccctggaagtacaggtttctgccttcatgccattcaatcttctgagcttccaagatgctggtcaggcc 3'

The protein expression and purification was done using the same protocol as for the native eEFSec. The only difference was the composition of the buffer used during the size-exclusion chromatography step: 10 mM Tris HCl pH 8.0, 300 mM KCl, 0.5 mM TCEP, 5 mM MgCl₂. The average purification yield was ~0.6 mg of His-Avi-TEV-eEFSec per 1 L of culture. Pure His-Avi-TEV-eEFSec was concentrated to 3 mg/mL, and then biotinylated overnight at +4°C using the commercial biotinylation kit (Avidity). The biotinylated protein was purified on the HiLoad 16/600 Superdex 200 size-exclusion chromatography column in 10 mM Tris HCl pH 8.0, 300 mM KCl, 0.5 mM TCEP, 5 mM MgCl₂, to remove the excess of biotin and birA enzyme (biotin ligase). Pure biotinylated His-Avi-TEV-eEFSec was concentrated to ~2 mg/mL, aliquoted, flash-frozen in liquid nitrogen, and stored at -80°C. An attachment of a single biotin residue to His-Avi-TEV-eEFSec was confirmed by the Mass Spectrometry, Metabolomics and Proteomics Facility at UIC.

II.2 Site-directed mutagenesis, expression, and purification of recombinant human eEFSec mutants

The mutant eEFSec constructs were prepared using QuikChange Site-Directed Mutagenesis Kit (Agilent Technologies) (TABLE II) and primers obtained from Integrated DNA Technologies. The mutations were confirmed by DNA sequencing (done by DNA Services Facility at UIC). The mutant constructs were expressed in *E.coli* Lemo21(DE3) competent cells according to the same protocol as utilized for wild-type eEFSec. The Ni²⁺-affinity purification of the mutants was followed directly by the size-exclusion chromatography on S200 column in 50 mM HEPES pH 7.5, 250 mM NaCl, 0.5 mM TCEP, and 5 mM MgCl₂. The purification yield for eEFSec mutants varied between 0.4 mg and 0.8 mg of protein per 1 L of culture. In case of eEFSec H230A_R285A double mutant, the purification yield was significantly lower (0.14 mg per 1 L of culture), whereas the purification of eEFSec T48A yielded higher than expected amount of protein (1.1 mg) per 1 L of culture. The pure mutant eEFSec constructs were concentrated up to 2.7 – 3.4 mg/mL (40 – 50 μM), aliquoted, flash-frozen in liquid nitrogen and stored at -80°C.

TABLE II. PRIMERS USED TO CREATE MUTANT eEFSec CONSTRUCTS

Mutant eEFSec	Template	Primer	Nucleotide sequence^a
T48A	WT	Forward	cgcgagcgcggcatcgcgctcgatctgggcttctcg
		Reverse	cgagaagcccagatcgagcgcgatgccgctcgcg
D92A	WT	Forward	caggtcacgctggcgcctgccccgggcacgcc
		Reverse	ggcgtgccccggggcaggcgaccagcgtgacctg
H96A	WT	Forward	gtcgactgccccggggccgcctccctcatccgg
		Reverse	ccggatgagggaggcggccccggggcagtcgac
D229A	WT	Forward	ctcatgtctgtggcccactgtttctccatcaaaggc
		Reverse	gcctttgatggagaaacagtgggccacagacatgag
H230A	WT	Forward	ctcatgtctgtggacgcctgtttctccatcaaaggc
		Reverse	gcctttgatggagaaacaggcgtccacagacatgag
R285A	WT	Forward	gcatgcaaggagacgcgctgggcatctgcgtc
		Reverse	gacgcagatgccagcgcgtctccttgcatggc
R285N	WT	Forward	gcatgcaaggagacaacctgggcatctgcgtc
		Reverse	gacgcagatgccagggtgtctccttgcatggc
H230A_R285A	R285A	Forward	ctcatgtctgtggacgcctgtttctccatcaaaggc
		Reverse	gcctttgatggagaaacaggcgtccacagacatgag
K582A	WT	Forward	ctcagcctgactttcgcggttatgtcttcgac
		Reverse	gtcgaagacataacgcgcgaaagtcaggctgag
R583A_Y584A	WT	Forward	ctcagcctgactttcaaggctgctgtcttcgacaccac
		Reverse	gtgggtgtcgaagacagcagcctgaaagtcaggctgag
⁵⁸² KRYVF ⁵⁸⁶ →AAAAA	WT	Forward	ctcagcctgactttcgcggtgctgccgacaccacaagcgc
		Reverse	gcgcttggggtgtcggcggcagcagccgaaagtcaggctgag

^a The sequences of both forward and reverse primers are listed in the 5' to 3' direction.

II.3 Crystallization, data collection and data processing

II.3.1 eEFSec:GDPNP

Pure eEFSec was mixed with 1 mM GDPNP (Jena Bioscience). Equal volumes of the protein sample and reservoir solution were mixed and the crystals were grown at +12°C using the sitting drop vapor diffusion method. The best quality crystals of eEFSec:GDPNP:Mn²⁺ were obtained in 0.1 M HEPES pH 7.6, 0.15 M ammonium sulfate, 18% (w/v) PEG 3,350, and 0.02 M glycine. The crystals were dehydrated overnight at +12°C over the reservoir solution supplemented with 28% (w/v) PEG 3,350, then cryoprotected with the solution supplemented with 28% (w/v) PEG 3,350 and 25% sucrose, and flash-frozen in liquid nitrogen. The X-ray diffraction data were collected at the liquid nitrogen temperature ($\lambda = 0.97856 \text{ \AA}$) at the Life Sciences Collaborative Access Team (LS-CAT) beamline of the Advanced Photon Source, Argonne National Laboratory (APS-ANL, Darien, IL). The data were scaled and reduced in HKL-2000 (71). Because attempts to solve the crystal structure by molecular replacement were unsuccessful, the single-wavelength anomalous dispersion (SAD) phasing method based on selenomethionine was pursued.

II.3.2 SeMet-eEFSec:GDPNP

Pure SeMet-eEFSec was mixed with 1 mM GDPNP (Jena Bioscience), and the crystals of SeMet-eEFSec:GDPNP:Mn²⁺ were grown at +12°C in 0.1 M HEPES pH 7.6, 0.3 M ammonium sulfate, 16% (w/v) PEG 3,350, 0.02 M glycine, and 4% dextran sulfate, using the sitting drop vapor diffusion method. The crystals were cryoprotected in

the reservoir solution supplemented with 30% (w/v) PEG 3,350, and flash-frozen in liquid nitrogen. The X-ray diffraction data were collected at liquid nitrogen temperature ($\lambda = 0.97856 \text{ \AA}$) at the LS-CAT and the Structural Biology Center (SBC) beamlines of the APS-ANL, and scaled and reduced in HKL-3000 (71).

II.3.3 eEFSec:GDPCP

Pure eEFSec was mixed with 1 mM GDPCP (Jena Bioscience), and the crystals of eEFSec:GDPCP:Mg²⁺ were grown at +12°C in 0.1 M HEPES pH 7.6, 0.25 M ammonium sulfate, 16% (w/v) PEG 3,350, 0.02 M glycine, and 2% dextran sulfate, using the sitting drop vapor diffusion technique. The crystals were cryoprotected in the reservoir solution supplemented with 30% (w/v) PEG 3,350, and flash-frozen in liquid nitrogen. The X-ray diffraction data collection was performed at liquid nitrogen temperature ($\lambda = 0.97856 \text{ \AA}$) at the LS-CAT beamline of the APS-ANL. HKL-2000 was used for data scaling and reduction (71).

II.3.4 eEFSec:GDP

The crystals of eEFSec complexed with 1 mM GDP (Jena Bioscience) were grown at +12°C in 0.1 M Tris HCl pH 8.5, 0.2 M NaI, 0.2 M KI, and 18% (w/v) PEG 3,350, using the sitting drop vapor diffusion setup. The crystals were cryoprotected in the reservoir solution supplemented with 18% (w/v) ethylene glycol, flash-frozen in liquid nitrogen, and analyzed for X-ray diffraction at liquid nitrogen temperature

($\lambda = 0.97856 \text{ \AA}$) at the LS-CAT beamline of the APS-ANL. Scaling and reduction of the collected data was done in HKL-2000 (71).

II.4 Structure determination and refinement

The crystal structure of eEFSec:GDPNP was determined by single-wavelength anomalous dispersion (SAD) phasing based on SeMet. Positions of selenium atoms were determined and the initial estimate of the experimental phase was calculated to 3.4 \AA resolution in SHELX (72). Figure of merit (FOM) after SHELXD was 0.61 at 3.39 \AA . After density modification in DM, $R_{\text{cullis,ano}}$ and FOM were 0.75 and 0.84, respectively. Density modification was done in DM and an autobuild module of HKL3000 was used to trace the backbone in the experimental electron density map. Iterative model building was done in Coot (73) and structure refinement was done in Phenix (74). The crystal structures of eEFSec:GDPCP and eEFSec:GDP were determined by molecular replacement in Phaser (75) using the structure of SeMet-eEFSec:GDPNP as a search model. In case of the GDP-bound structure, the molecular replacement solution was identified only after the search model was divided into the EF-Tu-like domain and the C-terminal domain 4. The final models of eEFSec:GDPNP, eEFSec:GDPCP, and eEFSec:GDP refined to $R_{\text{work}}/R_{\text{free}}$ of 0.24/0.29 (TABLE III), 0.24/0.29 (TABLE III), and 0.30/0.34 (TABLE VI), respectively, and they were of excellent geometry. The Ramachandran plots showed that 84%, 91%, and 82% of residues of eEFSec:GDPNP, eEFSec:GDPCP, and eEFSec:GDP, respectively, were in preferred regions. Also, 16% (eEFSec:GDPNP), 9% (eEFSec:GDPCP), and 18% (eEFSec-GDP) of residues were in

allowed regions. All figures showing the crystal structures of eEFSec, its complexes with guanine nucleotides, the superimpositions with other translation elongation/initiation factors, as well as the model of eEFSec-tRNA^{Sec} complex, were generated in PyMol (Schrodinger, LLC. The PyMOL Molecular Graphics System, Version 1.8. (2015)). The protein domain motion analysis was performed using DynDom online tool (76). Sequences were aligned using MultAlin (77) and the figure was prepared using ESPript 3.0 (78).

II.5 Analysis of eEFSec:GDP and eEFSec:GDPCP by Small-angle X-ray scattering (SAXS)

The SAXS experiments were performed at the 18-ID Biophysics Collaborative Access Team (BioCAT) beamline at APS-ANL (79). The experiments were conducted at room temperature in 20 mM HEPES pH 7.5, 150 mM NaCl, 0.5 mM TCEP, 5 mM MgCl₂, and 20 μM nucleotide (GDPCP or GDP, Jena Bioscience). Wild-type eEFSec samples were in a similar buffer with 1 mM of the nucleotide and at a final concentration of ~5.4 mg/mL. An in-line setup in which a 24 ml Superdex 200 10/300 GL size-exclusion chromatography column (GE Healthcare) was directly connected to the SAXS cell was used. This configuration allowed the sample exposure to X-rays immediately after its elution from the column. The baseline scattering was established based on the measurements done before and after peak elution. IGOR Pro was used for data reduction (WaveMetrics, Inc., Lake Oswego, OR). The ATSAS package was employed for data processing and subsequent calculations (80, 81). Guinier analysis and

calculation of the radius of gyration (R_g) was done in PRIMUS (81, 82), while the calculation of the pair-distance distribution function ($P(r)$) was completed in GNOM (83). DAMMIF (84) was used to generate the low-resolution *ab initio* models, and DAMCLUST (85) - for model clustering and averaging. Superimposition of the SAXS *ab initio* models onto the X-ray crystal structures was done using SUPCOMB (85), whereas CRY SOL was used to compare theoretical SAXS curves derived from the crystal structure with the experimental data. Finally, images representing the SAXS envelopes and X-ray crystal structures were done in PyMOL (Schrodinger, LLC. The PyMOL Molecular Graphics System, Version 1.8. (2015)).

II.6 Analysis of nucleotide binding to eEFSec by Isothermal Titration Calorimetry (ITC)

The binding of guanine nucleotides (GDP, GTP, GDPNP, GDPCP, and GTP γ S; Jena Bioscience) to the wild type (WT) and mutant constructs of human eEFSec was monitored at +25°C on the MicroCal ITC200 instrument (GE Healthcare). The binding events were measured in 50 mM HEPES pH 7.5, 250 mM NaCl, 0.5 mM TCEP, and 5 mM MgCl₂, and the same buffer was used for the final purification step of all eEFSec constructs and to prepare the nucleotide stocks. eEFSec (40-50 μ M) was placed in the sample cell and then titrated with the nucleotide solution (400-500 μ M) while stirring. Heat changes resulting from productive binding were measured, integrated, and binding and thermodynamic parameters were calculated using MicroCal Origin software (GE Healthcare).

II.7 *In vitro* selenocysteine incorporation assay

The assay was done in the laboratory of Dr. Paul R. Copeland (Department of Biochemistry and Molecular Biology, Rutgers – Robert Wood Johnson Medical School, Piscataway, NJ, USA). The selenocysteine incorporation activity of recombinant WT and mutant eEFSec constructs was determined using an *in vitro* translation system with a luciferase reporter mRNA that harbors an in-frame UGA codon at position 258 followed by rat glutathione peroxidase 4 (GPx4) SECIS element in the 3' UTR. Each 12.5 μ L reaction contained 6.25 μ L wheat germ extract, 320 nM recombinant SBP2, 320 nM wild-type or mutant eEFSec, 20 μ M amino acid mix, 125 ng of luciferase mRNA, and 1.25 μ g total aminoacyl-tRNA pool from rat testes (a rich source of Sec-tRNA^{Sec}). The reactions were incubated at +25°C for 2h, and the luminescence intensity was measured using a 96-well plate luminometer. The luminescence intensity obtained for each eEFSec construct was normalized relative to the signal from a control sample, which did not contain eEFSec.

III. The crystal structure of the GTP-bound state of human eEFSec

Significant portions of Chapters 3, 4, 5, and 6 are derived from the following manuscript: Dobosz-Bartoszek M, Pinkerton MH, Otwinowski Z, Chakravarthy S, Söll D, Copeland PR, Simonović M (2016) Crystal structures of the human elongation factor eEFSec suggest a non-canonical mechanism of selenocysteine incorporation, ***under review, Nature Communications***.

III.1 Introduction

eEFSec is an eukaryotic elongation factor that specifically recognizes Sec-tRNA^{Sec} and no other elongator aminoacyl-tRNA (aa-tRNA). Its activity is required for rewiring of an in-frame UGA codon from the translational stop signal into a signal for selenocysteine incorporation (34, 35). The biological importance of eEFSec was demonstrated by the fact that selenoprotein synthesis was completely ablated in the *EEFSEC* knockout mutant of *Drosophila melanogaster* (86). The current knowledge about eEFSec is derived from studies completed on general elongation factors EF-Tu and EF1A (62, 64, 87-90), as well as from studies on the archaeal (41) and bacterial (32, 47, 91-94) SelB orthologs. However, because of poor sequence conservation (18-25%) and different requirements for decoding of the Sec UGA codon across kingdoms, it is likely that conclusions from those studies may not be applicable to eEFSec.

A handful of biochemical studies have been completed using the eukaryotic and mammalian orthologs as a model system (34, 35, 43, 44), but none were done on the human factor. eEFSec is a GTPase that lacks intrinsic activity in the absence of the

ribosome. Phylogenetic analysis suggested that eEFSec consists of four domains. The N-terminal domains 1-3 (D1-3) were predicted to play similar roles as the corresponding domains in EF-Tu and EF1A (e.g. GTPase activity and binding of aa-tRNA) (95), whereas poorly conserved C-terminal domain 4 (D4) was suggested to play several functions such as binding to Sec-tRNA^{Sec}, SECIS-dependent interaction with SPB2, and regulation of eEFSec GTPase activity (44). Given that none of those functions have been assigned to D4 of SelB, and that the overall sequence conservation between eEFSec and SelB is quite low, it is reasonable to suggest that the eukaryotic Sec tRNA-specific elongation factor is distinct. Therefore, the need for further extensive studies on eukaryotic eEFSec in general, and human ortholog in particular, is warranted.

To my knowledge, no attempts have been made to determine the structure of eEFSec. Such structural characterization would reveal similarities and differences between eEFSec, SelB and other elongation factors. Additionally, it would provide valuable insights into how eEFSec performs its functions. To address that, I determined the crystal structures of the intact human eEFSec in complex with the non-hydrolyzable GTP analogs, GDPNP and GDPCP (Figure 5). Here I present the first structural characterization of the eukaryotic specialized elongation factor and provide the detailed comparison with general translation elongation factors EF-Tu/EF1A, archaeal and bacterial SelB orthologs of eEFSec, as well the universal translation initiation factor IF2/eIF5B.

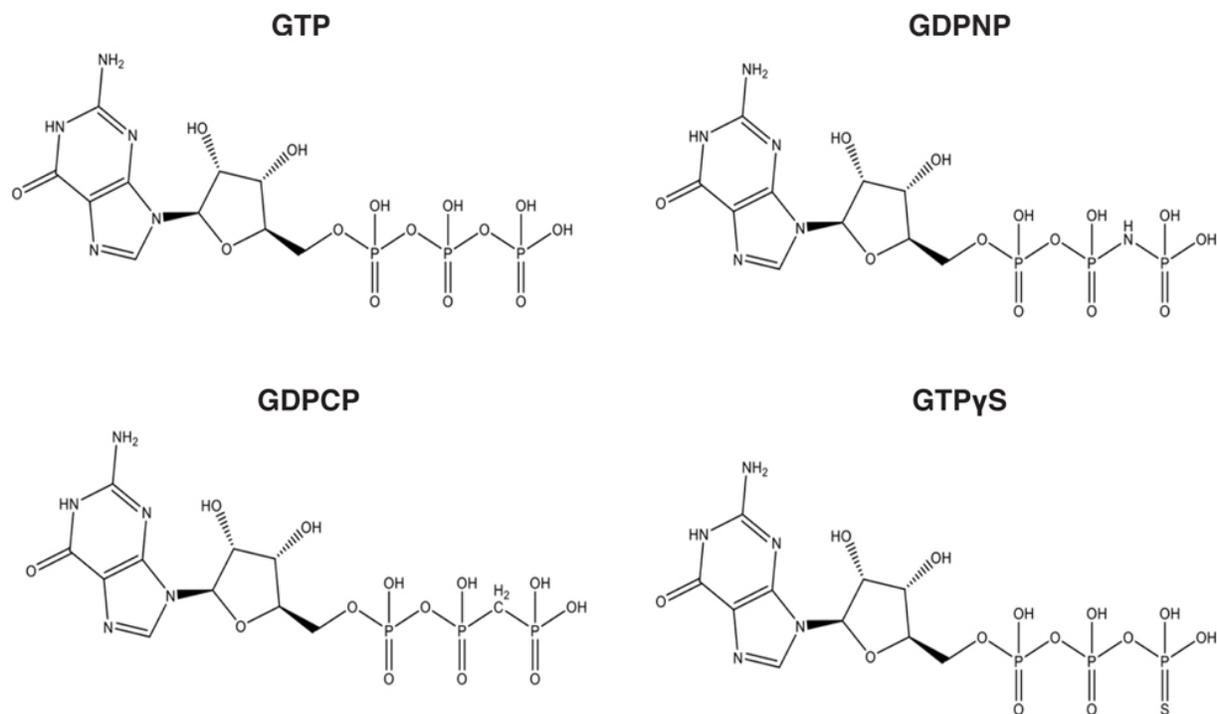


Figure 5. GTP and its analogs. The structures of GTP and commonly used non-hydrolyzable GTP analogs (GDPNP, GDPCP, and GTP γ S) are shown. GDPNP and GDPCP were used in crystallization of human eEFSec.

III.2 Experimental Results

III.2.1 Overall structure and domain organization of human eEFSec

Human eEFSec was co-crystallized with GDPNP (eEFSec:GDPNP) and GDPCP (eEFSec:GDPCP). The isomorphous crystals of eEFSec:GDPNP and eEFSec:GDPCP belonged to the orthorhombic space group (C222₁) and diffracted X-rays to 3.4 and 2.7 Å, respectively (TABLE III). The crystal structure of eEFSec:GDPNP was determined by single-wavelength anomalous dispersion (SAD) phasing based on selenomethionine (SeMet; see Methods), whereas the structure of eEFSec:GDPCP was solved by

molecular replacement using a complete monomer of eEFSec:GDPNP as a search model. With the exception of several flexible loops (residues 32-42, 70-80, 192-195, 383-403, 435-438, 524-526, and 544-569), the entire protein backbone was traced. The final models were refined to $R_{\text{work}}/R_{\text{free}}$ of 0.23/0.29 (eEFSec:GDPNP) and 0.24/0.29 (eEFSec:GDPCP) (TABLE III). The asymmetric unit (AU) contained two eEFSec molecules arranged in a head-to-head dimer. Given that monomers in each dimer are nearly indistinguishable and that eEFSec is a monomer in solution, the crystallographic dimer is most likely of no physiological significance. Strong positive peaks in the initial mFo-DFc electron density difference maps indicated that both monomers bound the appropriate GTP analog and a divalent metal ion, Mn^{2+} in case of eEFSec:GDPNP or Mg^{2+} in case of eEFSec:GDPCP. Mn^{2+} was used instead of Mg^{2+} in crystallization of eEFSec:GDPNP complex to ensure I obtained protein crystals and avoid crystallization of magnesium salts during an initial screening.

TABLE III. DATA COLLECTION AND REFINEMENT STATISTICS FOR eEFSec COMPLEXES WITH THE NON-HYDROLYZABLE GTP ANALOGS

	eEFSec:GDPNP ^a	eEFSec:GDPCP ^b
Data collection		
Space group	C 2 2 2 ₁	C 2 2 2 ₁
Cell dimensions a, b, c (Å)	92.3, 112.4, 327.7	94.3, 113.2, 329.7
Resolution (Å) ^c	46.00 – 3.38 (3.50 – 3.38)	49.00 – 2.72 (2.81 – 2.72)
<i>R</i> _{sym} or <i>R</i> _{merge}	0.09	0.23
<i>I</i> / σ ^c	19.7 (1.2)	14.5 (1.0)
Completeness (%) ^c	98.5 (88.9)	100 (100)
Redundancy ^c	9.5 (6.1)	19.4 (12.1)
Refinement		
Resolution (Å)	46.00 – 3.40	49.00 – 2.72
No. of reflections	21,650	46,503
<i>R</i> _{work} / <i>R</i> _{free}	0.24 / 0.29	0.24 / 0.29
No. of atoms		
Protein	6,604	7,259
Ligand/ion	66	66
Water		32
B-factors		
Protein	32.5	79.4
Ligand/ion	38.2	76.4
Water		77.0
R.m.s. deviations		
Bond lengths (Å)	0.010	0.014
Bond angles (°)	2.04	1.48

^a Two crystals were used in data collection.

^b Four crystals were used in data collection.

^c Values in parentheses are for the highest resolution shell.

Human eEFSec adopts a chalice-like structure composed of four domains: the N-terminal domains D1 (residues 1-215), D2 (residues 224-304) and D3 (residues 310-455), and the C-terminal domain D4 (residues 477-575). The linker region encompassing residues 469-476 and 576-582 connects D3 and D4 (Figure 5C). Domains 1-3 represent the cup of the chalice, the linker is the stem, and D4 is the base of the chalice (Figure 5A and 5B). The height of the chalice is ~ 100 Å, while its width varies along the molecule: it is the largest at the cup (60 Å), significantly smaller at the base (30 Å), and the smallest at the stem (20 Å). Domains D1-3 fold into an EF-Tu-like structure that harbors functional sites: the nucleotide-binding site and the putative Sec-binding pocket. The 6-stranded β -sheet of D1 is enclosed by 7 α -helices. An 8-residue long loop connects D1 with D2, which adopts a β -barrel structure composed of 8 antiparallel β -strands and a flanking, short α -helix. D3 harbors 7 antiparallel β -strands and continues into the linker region *via* a long helix $\alpha 8$. The linker is composed of two β -strands; the first strand, $\beta 23$, arises from D3 and continues into D4, which folds into a small β -barrel flanked by a pair of α -helices. The second strand of the linker region, $\beta 28$, runs antiparallel to $\beta 23$ and ends with an α -helical turn (residues 583-588) that sits below D3. The most C-terminal segment (residues 589-595) folds back below D3 and almost perpendicularly to the plane of the linker region while adopting a β -turn structure.

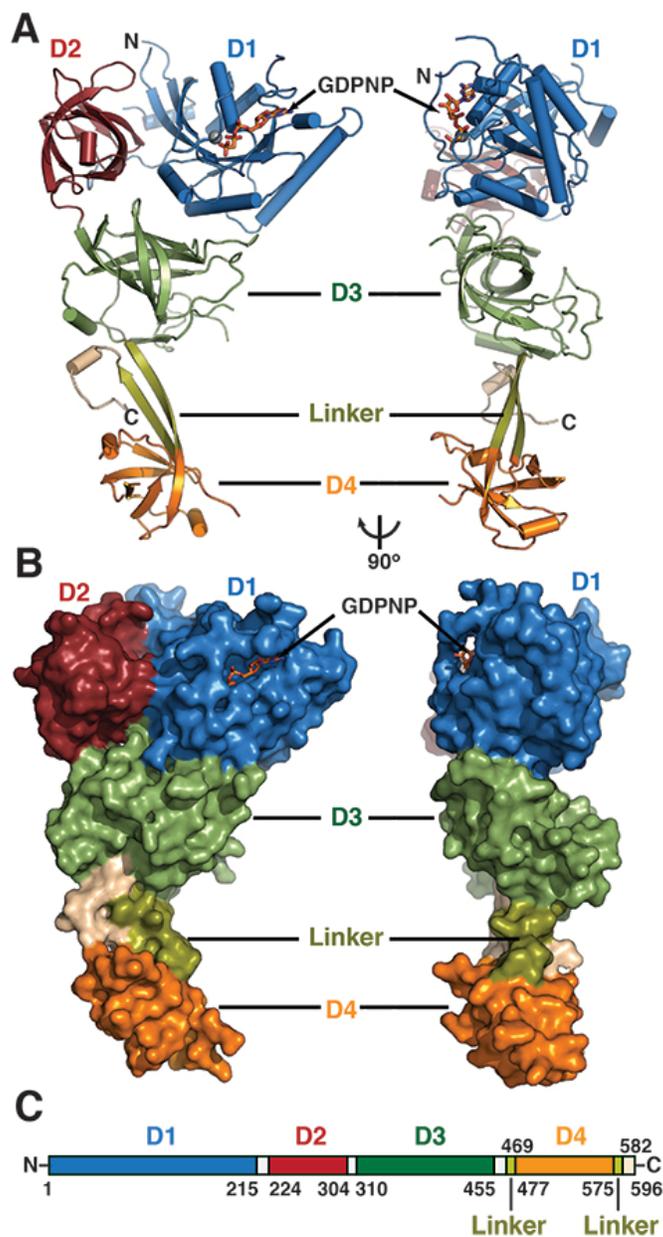


Figure 6. The overall structure and domain organization of human eEFSec. (A) Cartoon and surface (B) representation diagrams of the chalice-like structure of human eEFSec shown in two views rotated $\sim 90^\circ$ clockwise around vertical axis. Individual domains, linker and extreme C-terminus regions are labeled and colored according to the scheme in (C). GDPNP is shown as sticks and Mg^{2+} as a grey sphere. (C) Schematic diagram showing domain organization in human eEFSec. White bars denote regions connecting individual domains.

The superimposition of eEFSec:GDPNP and eEFSec:GDPCP revealed that the structures are identical (r.m.s.d. value of 0.7 Å for 488 residues), and suggested that both GTP analogs induce the same conformation of human eEFSec (Figure 6). This also argues that GDPCP and GDPNP are equally good mimics of GTP when bound to eEFSec.

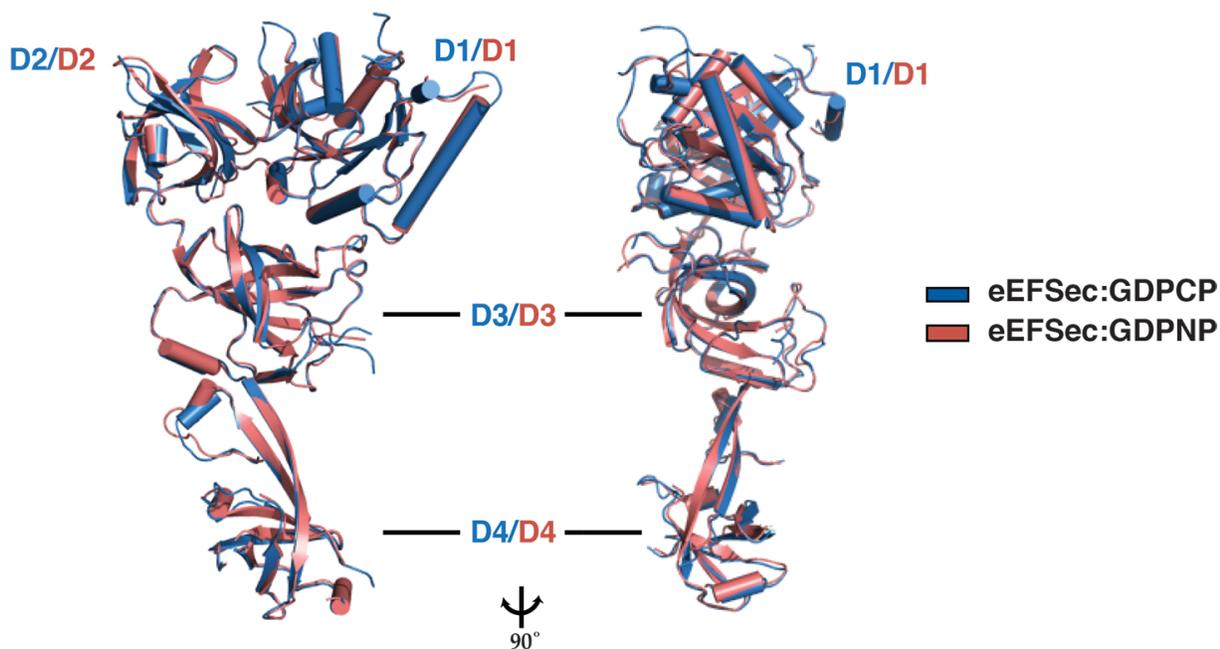


Figure 7. The GTP analogs, GDPCP and GDPNP, induce the same conformation of human eEFSec. The superimposition of eEFSec:GDPCP (blue) onto eEFSec:GDPNP (pink), seen from the two orientations related by $\sim 90^\circ$ clockwise rotation around the vertical axis, reveals the identical arrangement of domains D1-4 in both structures.

III.2.2 A structural comparison with general translation elongation factors EF-Tu and EF1A

I wished to establish whether eEFSec in complex with GTP analogs adopts a conformation that resembles the GTP-bound state of EF-Tu. Thus, I compared the structures of eEFSec:GDPNP and eEFSec:GDPCP with the structures of the major functional states of EF-Tu.

The superimposition of eEFSec:GDPNP and eEFSec:GDPCP onto the *Thermus aquaticus* EF-Tu:GDPNP (62) reveals a similar arrangement of D1-3 in all three structures (Figure 7A). Analogous comparison with the GTP-bound state of EF1A currently cannot be performed due to the lack of structural information on both archaeal and eukaryotic EF1A. By contrast, the domain arrangement is different when the same analysis is applied to *T. aquaticus* EF-Tu:GDP (64) (Figure 7B) and rabbit eEF1A2:GDP (90) (Figure 8). These results suggest that GDPNP and GDPCP trap eEFSec in a conformation that resembles the GTP-bound state of EF-Tu.

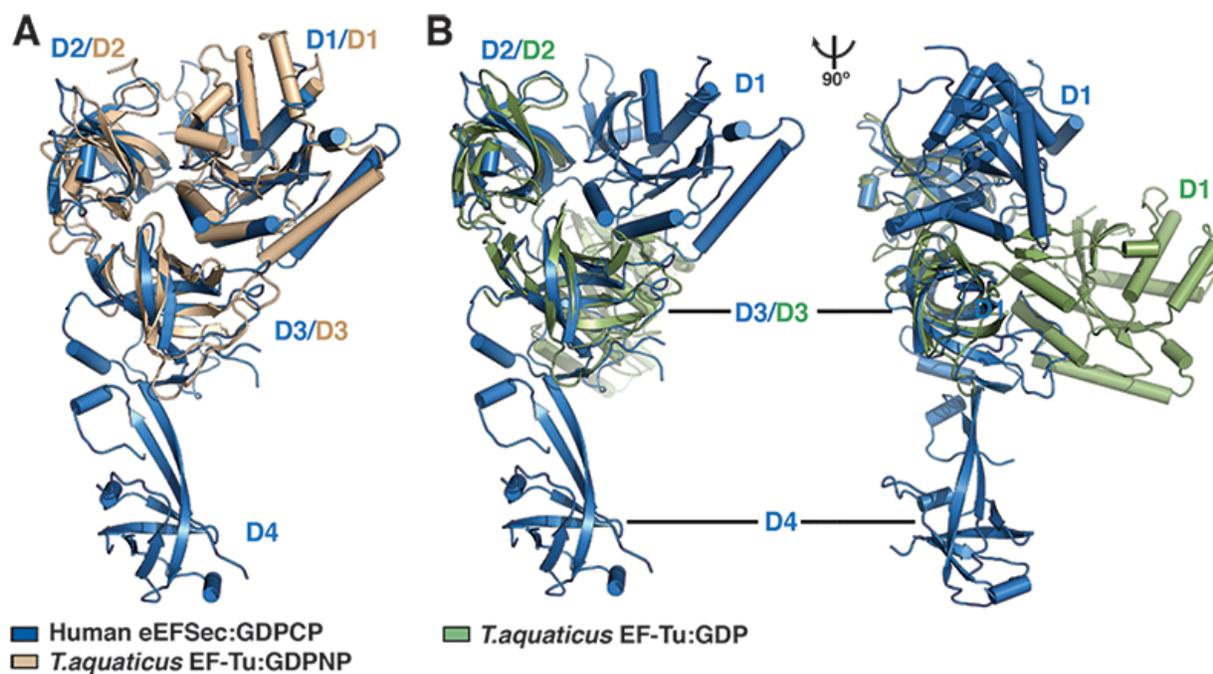


Figure 8. The GTP analogs, GDPCP and GDPNP, lock eEFSec in the GTP-bound state. The overlay of human eEFSec:GDPCP (blue) onto EF-Tu:GDPNP (**A**, beige) and EF-Tu:GDP (**B**, green), reveals that the arrangement of D1-3 in eEFSec:GDPCP is similar to that in the GTP-bound state of EF-Tu. Thus, GDPCP and GDPNP are optimal GTP mimics when bound to human eEFSec. The orientation of D1 in EF-Tu:GDP can only be seen when the view in (**A**) and (**B**, left) is rotated $\sim 90^\circ$ clockwise around the vertical axis.

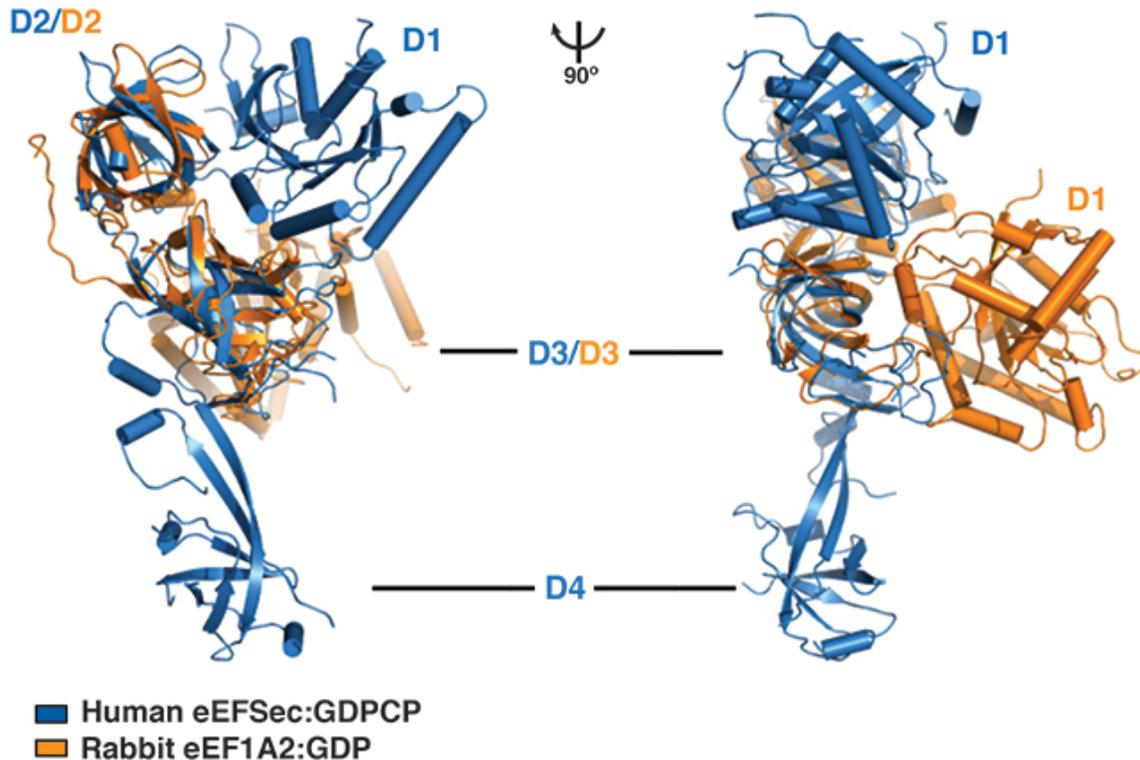


Figure 9. The superimposition of human eEFSec:GDPCP onto the GDP-bound state of rabbit general translation elongation factor eEF1A2. The overlay of eEFSec (blue) and eEF1A2 (orange) structures reveals different arrangement of D1-3. This further confirms that GDPCP and GDPNP trap eEFSec in the GTP-bound state. Two views are rotated $\sim 90^\circ$ clockwise around the vertical axis.

A more detailed analysis demonstrates that individual domains harbor differences that could be of functional significance. For instance, the structural overlay of D1, D2 and D3 of human eEFSec onto that of *T. aquaticus* EF-Tu yielded relatively small r.m.s.d. values spanning from 1.5-1.7 Å. However, D1 of EF-Tu harbors two well-ordered α -helical insertions that sit atop the GTPase site. By contrast, the corresponding regions in eEFSec (residues 32-42 and 186-202) are shorter and partially

disordered in my crystal structures. In addition, the dorsal side of eEFSec (i.e. the side of eEFSec opposite from the presumed tRNA-binding side) contains a partially disordered insertion (residues 57-87), while EF-Tu harbors a short and well-ordered loop. Likewise, loop β 10- β 11 in D2 is much shorter in eEFSec than in EF-Tu. Lastly, D3 of eEFSec contains insertions in several solvent-exposed loops. Perhaps the most obvious one is a well-ordered loop β 17- β 18 (residues 352-373), located at the dorsal face of eEFSec. Other enlarged loops in eEFSec are β 21- β 22 (residues 432-444) at the interface of D1 and D3, and loop β 18- β 19 (residues 378-410), but these are partially disordered in my structures.

III.2.3 A structural comparison with SelB orthologs

The crystal structures of the tRNA^{Sec}-specific elongation factors derived from different organisms provided a platform for analysis of the structural conservation across kingdoms. The overall structure of human eEFSec resembles its archaeal rather than the bacterial ortholog. The structural resemblance with the bacterial SelB is restricted to the EF-Tu-like domain, but it extends to D4 in case of the archaeal SelB (Figure 9). The global superimposition of the archaeal and bacterial SelB onto human eEFSec yielded r.m.s.d. values of 2.2 (356 residues) and 2.1 Å (323 residues), respectively. On the other hand, the overlay of D1 of human eEFSec onto those in the archaeal and bacterial SelB yielded r.m.s.d. values of 1.1 and 1.4 Å, respectively. The main differences are within switch 1 and around the GTPase site where several enlarged loops in eEFSec are partially disordered (i.e. residues 59-85 and loop β 6- α 6). An analogous comparison

using D2 also resulted in low r.m.s.d. values of 1.2 (archaeal SelB) and 1.3 Å (bacterial SelB). The differences are restricted to loop $\beta 7$ - $\beta 8$ (residues 233-238), which is oriented differently in eEFSec compared to SelB. The low r.m.s.d. values of 1.3 (archaeal SelB) and 1.6 Å (bacterial SelB) were also calculated after D3 was superimposed. The only differences are present in loops $\beta 17$ - $\beta 18$, $\beta 18$ - $\beta 19$, and $\beta 21$ - $\beta 22$, which are enlarged in eEFSec.

The most marked differences between the orthologs are present in the C-terminal D4. Although quite similar, the conservation of the archaeal and human D4 is not strict; the human enzyme harbors an additional α -helix and a longer C-terminal segment within the OB-fold (Figure 9A). In a striking contrast, the bacterial D4, which is the largest among the specialized elongation factors, consists of four winged-helix motifs and is rotated $\sim 90^\circ$ around the linker region when compared to the archaeal SelB and human eEFSec (Figure 9B).

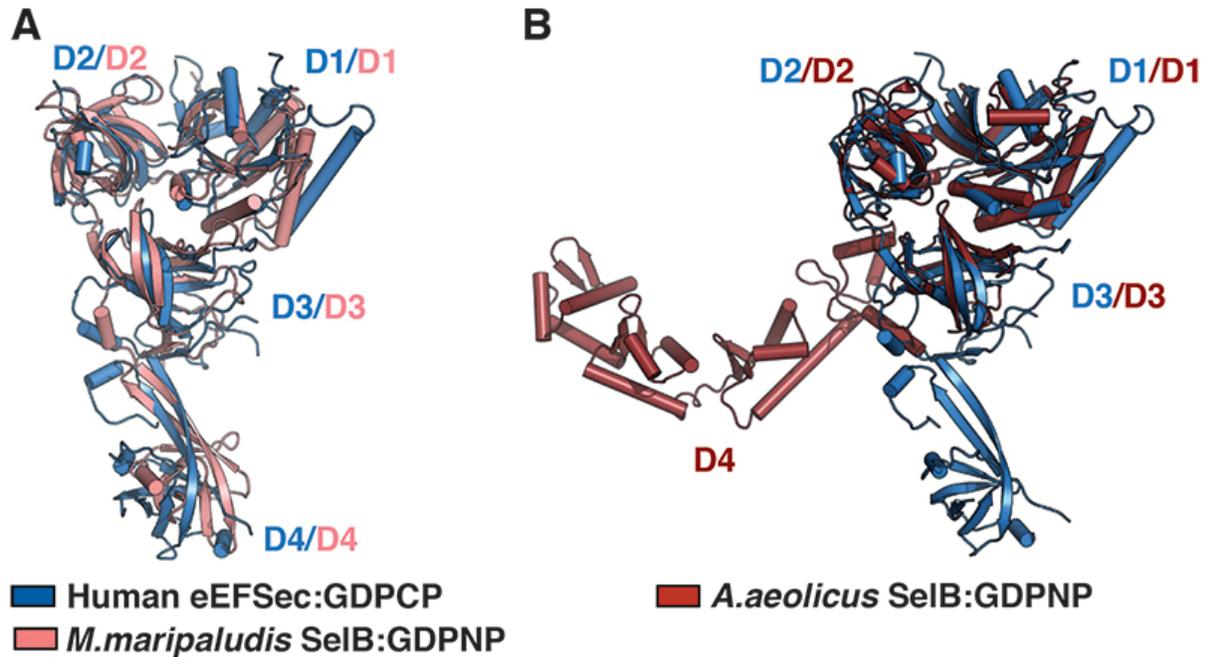


Figure 10. The structural comparison of human eEFSec with SelB orthologs. The global superimposition of human eEFSec:GDPCP (blue) onto *Methanococcus maripaludis* SelB:GDPNP (**A**, pink) and *Aquifex aeolicus* SelB:GDPNP (**B**, dark red), revealed that eEFSec resembles the archaeal rather than the bacterial ortholog. The most marked differences between eEFSec and its orthologs are present in the appended domain D4.

A closer inspection showed that the orientation of human D4 relative to EF-Tu-like domain is stabilized by interactions between loop $\beta 28-\alpha 11$ at the extreme C-terminus and residues in D3. In particular, a H-bond is formed between the conserved Glu372 (E372) from D3 and Lys582 (K582) from the C-terminal segment (Figure 10). The Glu-Lys pair is present in the archaeal (e.g. Glu325, Lys388), but not in the bacterial SelB. This led us to hypothesize that interactions within this ‘hinge’ region could be of importance for the interdomain interactions and perhaps domain orientation.

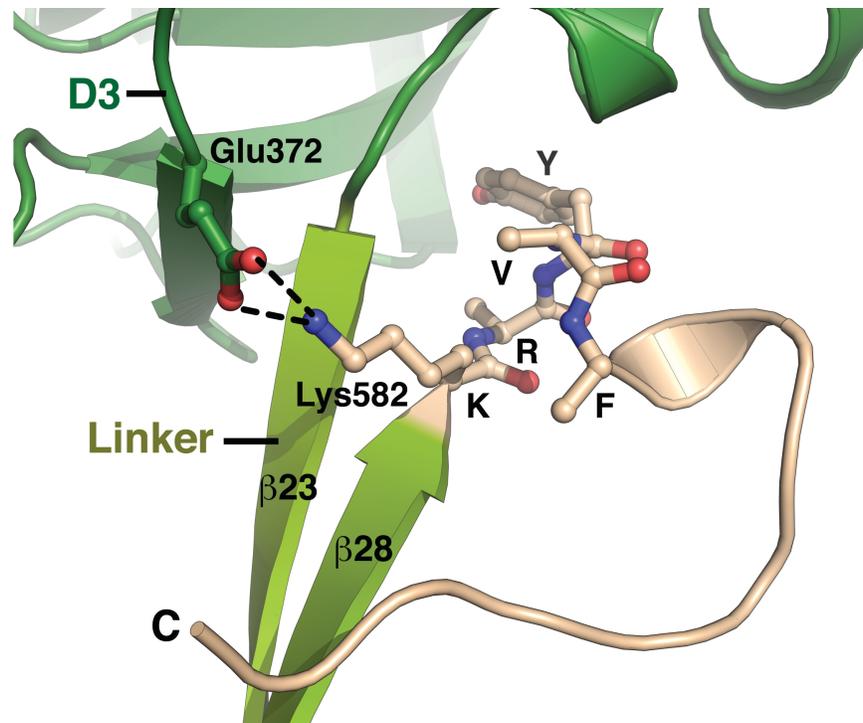


Figure 11. The structure of the ‘hinge’ region in human eEFSec. The ‘hinge’ is located at the interface between D3 (dark green), the linker region (light green), and the C-terminal segment (beige). The side chain of the conserved Glu372 (dark green sticks) forms H-bonds with the side chain of Lys582 (beige sticks) from the ‘KRVYF’ motif (beige sticks), which connects strand $\beta 28$ of the linker with the C-terminal segment. The interaction between Glu372 and Lys582 is conserved in the archaeal SelB, but absent from the bacterial ortholog, and it may be significant for stabilizing the interdomain interactions in eEFSec. The side chains of Arg583, Val585, and Phe586 in ‘KRYVF’ are disordered in my crystal structures.

I assessed the importance of the ‘hinge’ region by mutational and activity studies. I introduced single (K582A), double (R583A_Y584A) and penta ($^{582}\text{KRYVF}^{586} \rightarrow \text{AAAAA}$) alanine mutations within this region. The R583A_Y584A double mutation does not affect the expression level and the level of Sec incorporation (Figure 12C), while K582A is less stable than WT eEFSec. On the other hand, the penta alanine mutant elutes from the SEC column in two peaks (Figures 12A and 12B). The first peak

appears immediately after the void volume and probably contains higher molecular weight oligomers or aggregates. Although the second peak elutes as the WT monomer, it contains degraded mutant protein (Figure 11B).

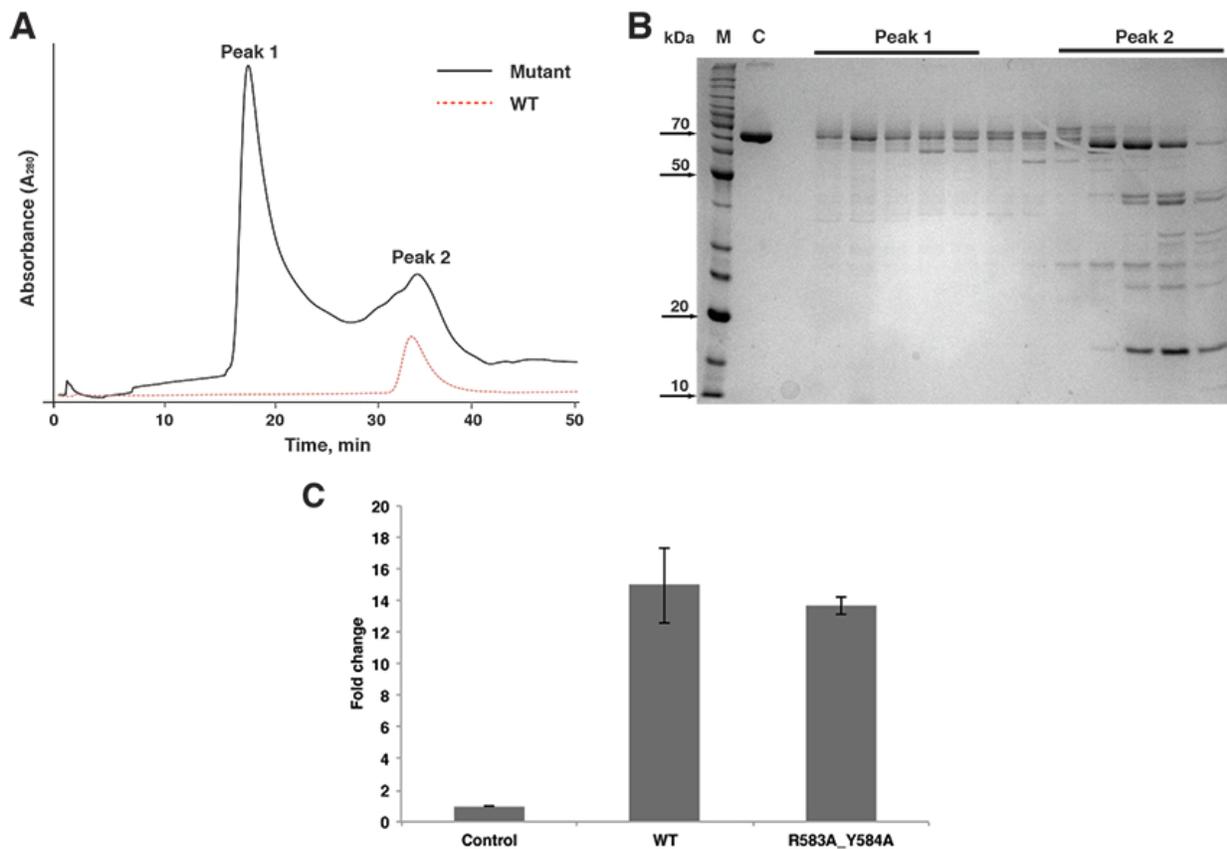


Figure 12. Mutational analysis of the ‘hinge’ region in eEFSec. (A) The SEC elution profile of the penta alanine mutant (black line) is different from that of the WT protein (red dashed line). **(B)** The SDS-PAGE analysis indicates that the majority of the mutant protein is found in Peak 1, which elutes from the SEC column immediately after the void volume. Also, Peak 2 contains predominantly degraded eEFSec. M: molecular weight markers, C: WT eEFSec. **(C)** The double mutation within the ‘hinge’ region (R583A_Y584A) does not impair the Sec incorporation activity of eEFSec *in vitro*.

III.2.4 Human eEFSec resembles the universal translation initiation factor IF2/eIF5B

Previous phylogenetic studies proposed that eEFSec and SelB are in a closer evolutionary relationship with the translation initiation protein factors rather than with the general elongation factors (96). The superimposition of individual domains of eEFSec:GDPCP onto the corresponding domains of the archaeal IF2/eIF5B:GDPNP yielded r.m.s.d. values of 2.7 Å (D1), 1.6 Å (D2) and 2.8 Å (D4). Because of significant differences, the linker region and D3 were omitted from this analysis. D3 of IF2/eIF5B folds into a mixed α/β structure that is linked to D4 by a long α -helical linker (97). By contrast, D3 of eEFSec has the β -barrel structure flanked by an α -helix, and is connected to D4 *via* two long antiparallel β -strands. However, the overall domain organization and the shape of human eEFSec resemble more IF2/eIF5B than EF-Tu/eEF1A (Figure 12).

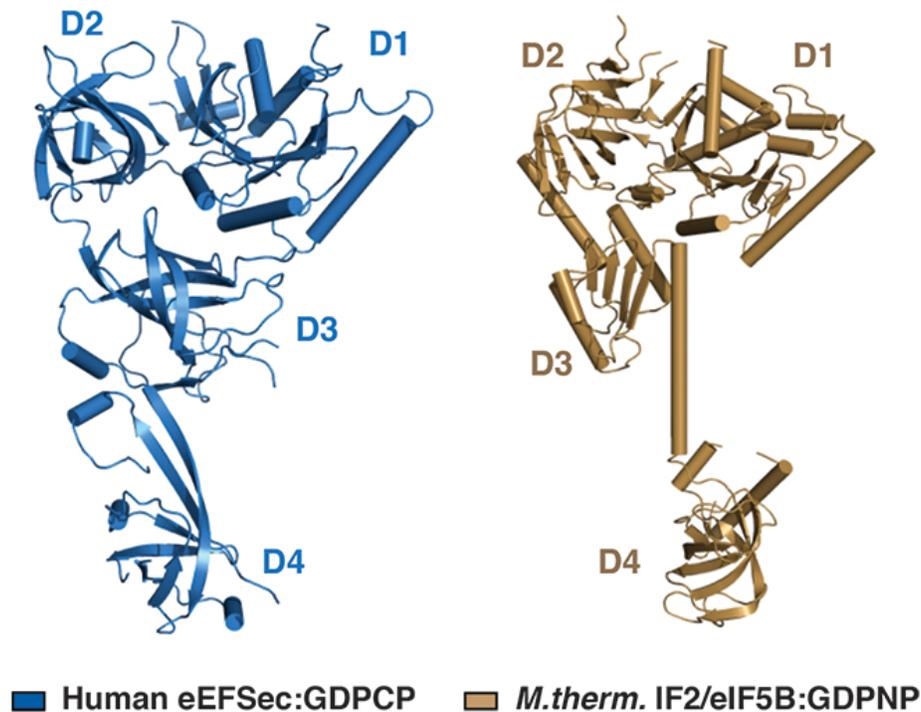


Figure 13. Structural conservation between human eEFSec and the universal translation initiation factor IF2/eIF5B. The overall domain arrangement and shape of human eEFSec:GDPCP (blue) and the universal translation initiation factor IF2/eIF5B from *Methanobacterium thermoautotrophicum* (gold) are quite similar.

III.3 Discussion

In spite of being essential for the development and survival of the healthy organism, synthesis of selenoproteins and incorporation of Sec in response to an in-frame UGA codon are poorly understood. As the first step towards deciphering how Sec is inserted into the nascent polypeptide, I determined the crystal structures of the intact human eEFSec in complex with the non-hydrolyzable GTP analogs, GDPNP and GDPCP. These structures practically represent the first structures of macromolecular

components responsible for promoting of the decoding of the Sec UGA codon in humans.

The structural studies revealed that four domains of eEFSec are arranged in a form of a chalice. The N-terminal domains D1-D3 fold into an EF-Tu-like structure, whereas the appended C-terminal domain D4, which is absent in general translation elongation factors, adopts a β -barrel structure. The overall architecture of human eEFSec resembles the archaeal rather than bacterial SelB. The structural conservation between the orthologs is most evident in the EF-Tu-like domain, whereas the most prominent differences are observed in D4. Both the sequence and domain size of D4 vary considerably across kingdoms, with the bacterial D4 being the largest. In addition, D4 folds into distinct structures among the orthologs, suggesting also a functional divergence. In particular, the bacterial D4 is composed of the winged-helix motifs often found in DNA- and RNA-binding proteins, which explains why bacterial SelB directly interacts with the SECIS element in the selenoprotein mRNA (42, 45). The corresponding domain in human eEFSec, which folds into a β -barrel flanked by two α -helices, cannot directly bind to SECIS element, but it interacts with an auxiliary protein factor SBP2 (98-100). However, the archaeal system does not require SBP2 and yet an overall architecture of the archaeal and eukaryotic D4 is similar. This suggests that common requirements other than SBP2 gave rise to a similar fold of D4 in the archaeal SelB and eEFSec. Additionally, the detailed structural analysis identified the 'hinge' region, located at the extreme C-terminus of eEFSec and between D4 and D3, which could be of structural importance. This particular segment may be significant for

stabilizing the orientation of the appended D4 relative to the EF-Tu-like domain in eEFSec and archaeal SelB. However, further examination of the potential functional role of this region is warranted.

Regardless of the differences and consistent with previous phylogenetic studies (96), my structural analysis suggests that human eEFSec and its orthologs are closely related to the translation initiation protein factors. It is not clear if this close evolutionary relationship is limited to the overall structure, or whether it is reflected at the functional level. Further mechanistic and structural studies on both systems are needed.

Moreover, my results show that eEFSec adopts the same structure in the presence of both GDPNP and GDPCP. Also, a structural comparison with EF-Tu and EF1A revealed that both GTP analogs trap eEFSec into the GTP-bound state, implying that GDPNP and GDPCP are equally good mimics of GTP when bound to human eEFSec. This finding is of a particular interest because of an ongoing debate as to which analog should be used to 'lock' the GTP-bound state of elongation factors. While some research groups used GDPNP as the GTP mimic to determine the crystal structures of the archaeal and bacterial SelB (41, 42), others raised a concern that GDPNP mimics GDP instead of GTP (47). However, I suggest that GDPNP and GDPCP could be used as GTP mimics in biochemical and structural analyses of eEFSec and SelB.

IV. The EF-Tu-like domain of human eEFSec harbors the major functional sites: the GTPase site and Sec-binding pocket

IV.1 Introduction

The current working model describing co-translational incorporation of Sec into a nascent selenoprotein chain is in part based on the mechanism utilized by EF-Tu and EF1A. The binding of GTP to D1 of EF-Tu has been shown to induce a conformation in which the amino acid-binding site, located at the interface of D1 and D2, is properly formed. As already mentioned, the ribosome-stimulated GTP hydrolysis by EF-Tu promotes a rotation of D1 away from D2 and D3, which leads to a dramatic transformation of the globular structure of the GTP-bound state into a more elongated conformation of the GDP-bound state (61, 64). This leads to disruption of the amino acid-binding pocket and the tRNA recognition surface, and then to the dissociation of EF-Tu from the ribosome. Given their importance for the function of elongation factors, I sought to position the main functional sites in eEFSec and to identify side chains critical for eEFSec function.

Following previous studies on EF-Tu and SelB, eEFSec is expected to contain two functional sites in the EF-Tu-like domain. The crystal structures of the GTP-bound state of human eEFSec provided an ideal platform for studying these sites. The binary complexes with GTP analogs revealed that the GTPase site is located within D1, just like in EF-Tu and SelB. Also, based on the structures of EF-Tu:GDPNP:Phe-tRNA^{Phe} (63) and SelB:GDPNP (41, 42), the putative Sec-binding pocket in eEFSec is predicted

to be at the interface of domains D1 and D2. This particular pocket is thought to be the main filter that discriminates the selenocysteinylyl moiety from other proteinogenic aminoacyl groups. Perhaps the most illustrative evidence supporting the significance of the Sec-binding site was provided by a study on the bacterial SelB which showed that the affinity for Sec-tRNA^{Sec} is $\sim 10^6$ -fold stronger than that for the unacylated tRNA^{Sec} (94). In this chapter, I describe the structure of both the GTPase and Sec-binding sites in human eEFSec, and present their biochemical characterization.

IV.2 Experimental Results

IV.2.1 The GTPase site in human eEFSec

The crystal structure of eEFSec:GDPCP and eEFSec:GDPNP unambiguously showed that the GTPase site in human eEFSec is the same when bound to GDPCP and GDPNP. Taken together with the results presented in sections III.2.1 and III.2.2, I concluded that both analogs trap eEFSec in the GTP-bound state. The GTPase site is composed of conserved elements found in other small GTPases: a phosphate-binding loop (P loop; residues 14-21; ¹⁴GxxxxGKT²¹), switch 1 (residues 32-47), switch 2 (residues 92-96; ⁹²DxxGH⁹⁶), the guanine-binding sequence (G-binding sequence; residues 146-149; ¹⁴⁶NKxD¹⁴⁹), and a divalent metal ion (Mg²⁺ in eEFSec:GDPCP, Mn²⁺ in eEFSec:GDPNP), which is an indispensable cofactor in all GTPases (Figure 13). Apart from a segment of switch 1, the GTPase site in my crystal structures is well ordered, and the electron densities for the GTP analog and the metal ion are well defined. The guanine ring of the GTP analog(s) is stacked between the aliphatic chain

of Lys147 and loop 190 (¹⁸⁸PGGP¹⁹¹). The Watson-Crick face interacts with the side chain of Asp149, while the Hoogsteen side forms H-bonds with the backbone amide of Ala186 and the side chain of Asn146. The ribose hydroxyls are within the H-bonding distance from the side chain carboxyl of Glu44. This interaction is, however, not present in the GDPNP complex, lessening its significance for complex stabilization. The backbone amides of the P loop (¹⁹GKTA²²) interact with non-bridging oxygens of α - and β -phosphates. The β - and γ -phosphates are held in place by interactions with Mg²⁺ ion and the backbone amide of Gly95 in switch 2. The side chains of Thr21 and Thr48 in switch 1, and two water molecules, W1 and W2, complete the octahedral coordination of Mg²⁺. W2 is oriented for interaction with Mg²⁺ by H-bonds with Asp92 from switch 2. The side chain of His96, whose corresponding residue in bacterial EF-Tu is pivotal for the GTPase activity (66, 101, 102), points away from the γ -phosphate.

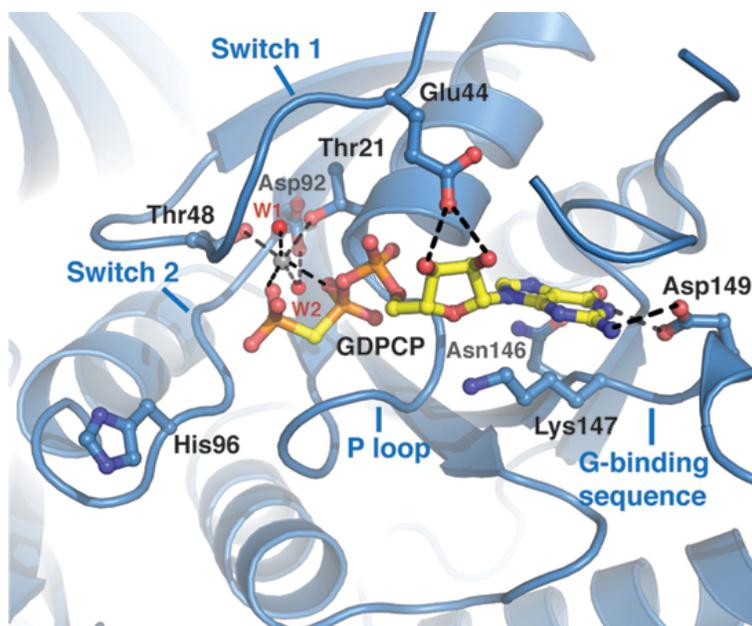


Figure 14. The structure of the GTPase site in human eEFSec. The GTPase site (blue cartoon) is located in D1 and is composed of the structural elements conserved in small GTPases: the P loop, switch 1, switch 2, the G-binding sequence, and a divalent metal ion. The binding of the GTP analog, GDPCP (gold sticks), causes partial ordering of switch 1, which orients Thr48 and Asp92 to interact with Mg^{2+} (grey sphere) and waters 1 and 2 (W1, W2; red spheres). The conserved residues in the GTPase site as well as the nucleotide are shown as sticks, and H-bonds are dashed lines.

IV.2.2 The interaction of human eEFSec with guanine nucleotides and analogs

I characterized the binding of GDP, GTP and the non-hydrolyzable GTP analogs (GDPNP, GDPCP, and $GTP\gamma S$) to human eEFSec by Isothermal Titration Calorimetry (ITC) (Figure 14). The measurements confirmed that eEFSec contains a single nucleotide-binding site and showed that the nucleotide binding is an enthalpically driven process. Perhaps more importantly, eEFSec binds GDP and GTP with similar binding affinities (K_d) of $0.19 \mu M$ and $1.21 \mu M$, respectively (TABLE IV). The binding affinities of eEFSec for the GTP analogs are similar to that for GTP, with K_d values ranging from

0.48 μM for $\text{GTP}\gamma\text{S}$, 1.20 μM in case of GDPNP , to 1.85 μM for GDPCP (TABLE IV). Interestingly, the enthalpy (ΔH) and entropy (ΔS) changes elicited by binding of GDPCP are similar to those of GTP , whereas those derived from the interactions with GDPNP and $\text{GTP}\gamma\text{S}$ are similar to GDP (TABLE IV). In case of the bacterial SelB , however, the thermodynamic parameters describing interactions with GDPNP and GDP were similar, whereas $\text{GTP}\gamma\text{S}$ resembled GTP (47). In other words, the ITC results suggested that GDPNP might trap SelB into a GDP -bound state, whereas $\text{GTP}\gamma\text{S}$ might be an adequate GTP mimic. A complex between GDPCP and SelB was not examined and thus I could not assess if the nature of such complex is the same across species. However, my crystal structures suggest that GDPNP and GDPCP are equally good mimics of GTP when bound to eEFSec in the absence of the ribosome.

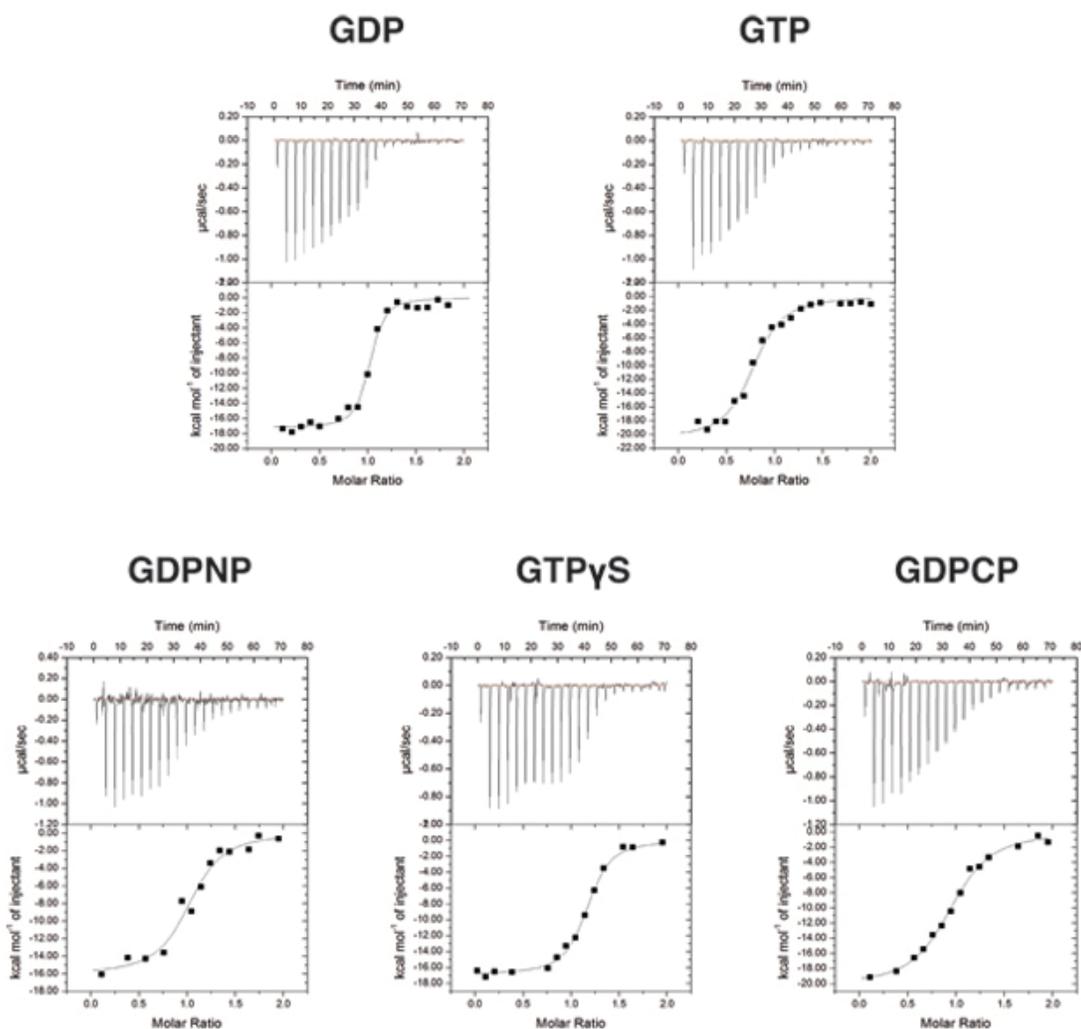


Figure 15. The binding of guanine nucleotides and analogs to human eEFSec characterized by ITC. The titration curves (upper panels) and binding isotherms (lower panels) derived from interactions of eEFSec with GDP, GTP and non-hydrolyzable GTP analogs are shown. The thermodynamic parameters calculated from the ITC data are presented in TABLE IV.

TABLE IV. BINDING OF GUANINE NUCLEOTIDES AND ANALOGS TO HUMAN eEFSec

Ligand	N	K _d (μ M)	Δ H (kcal/mol)	Δ S (cal/mol/deg)
GDP	0.98 \pm 0.01	0.19 \pm 0.05	-17.19 \pm 0.30	-26.90
GTP	0.77 \pm 0.02	1.21 \pm 0.26	-20.57 \pm 0.67	-41.90
GDPNP	1.01 \pm 0.03	1.20 \pm 0.47	-16.05 \pm 0.77	-26.70
GTPγS	1.14 \pm 0.01	0.48 \pm 0.06	-16.87 \pm 0.19	-27.70
GDPCP	0.95 \pm 0.02	1.85 \pm 0.24	-20.23 \pm 0.46	-41.60

IV.2.3 The mutational analysis of the nucleotide-binding site in human eEFSec

Based on biochemical studies on EF-Tu, and phylogenetic and structural analyses, I wished to determine if the side chains of Thr48, Asp92, and His96 within the GTPase site of eEFSec are important for nucleotide binding, GTP hydrolysis, Sec incorporation and selenoprotein synthesis. I designed three mutant constructs in which these residues were mutated into alanine – T48A, D92A, and H96A. All mutants expressed and purified like the wild-type (WT) eEFSec. ITC was used to characterize the interaction of eEFSec mutants with GDP and GTP (Figure 15).

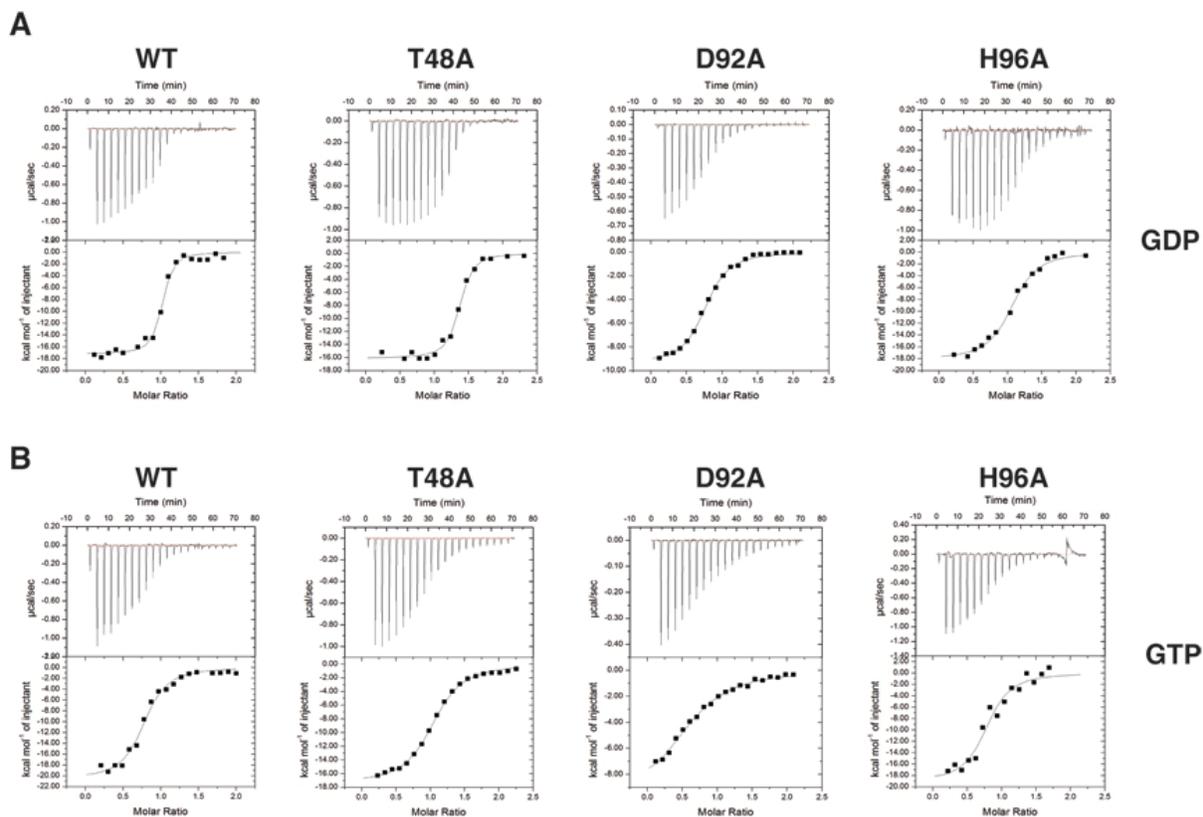


Figure 16. The binding of guanine nucleotides to GTPase site mutants of human eEFSec characterized by ITC. The titration curves (upper panels) and binding isotherms (lower panels) derived from interactions of eEFSec mutants with GDP (**A**) and GTP (**B**) are shown and compared to WT eEFSec. The thermodynamic parameters calculated from the ITC data are presented in TABLE V.

The binding affinity for GTP of T48A ($1.82 \mu\text{M}$) and H96A ($1.60 \mu\text{M}$) was similar to that of the WT protein ($1.21 \mu\text{M}$), whereas the D92A mutation caused a 7-fold decrease in the affinity ($8.47 \mu\text{M}$). Interestingly, D92A and H96A bound GDP with ~ 7 -fold weaker affinity ($1.50 \mu\text{M}$ and $1.30 \mu\text{M}$, respectively), while the affinity of T48A for GDP was similar to that of WT eEFSec ($0.19 \mu\text{M}$) (TABLE V). Although not affecting the

nucleotide binding, the GTPase site mutations severely impaired the read-through of the Sec UGA codon and selenoprotein synthesis *in vitro* (Figure 16).

TABLE V. BINDING OF GUANINE NUCLEOTIDES TO GTPASE SITE MUTANTS OF HUMAN eEFSec

Protein	N	K _d (μ M)	Δ H (kcal/mol)	Δ S (cal/mol/deg)
GDP binding				
WT	0.98 \pm 0.01	0.19 \pm 0.05	-17.19 \pm 0.30	-26.90
T48A	1.32 \pm 0.01	0.24 \pm 0.05	-16.11 \pm 0.25	-23.70
D92A	0.76 \pm 0.01	1.50 \pm 0.13	-9.41 \pm 0.13	-4.90
H96A	1.06 \pm 0.02	1.30 \pm 0.24	-18.14 \pm 0.45	-33.90
GTP binding				
WT	0.77 \pm 0.02	1.21 \pm 0.26	-20.57 \pm 0.67	-41.90
T48A	1.04 \pm 0.01	1.82 \pm 0.11	-17.41 \pm 0.18	-32.10
D92A	0.66 \pm 0.02	8.47 \pm 0.49	-9.88 \pm 0.49	-9.92
H96A	0.78 \pm 0.04	1.60 \pm 0.75	-19.06 \pm 1.37	-37.40

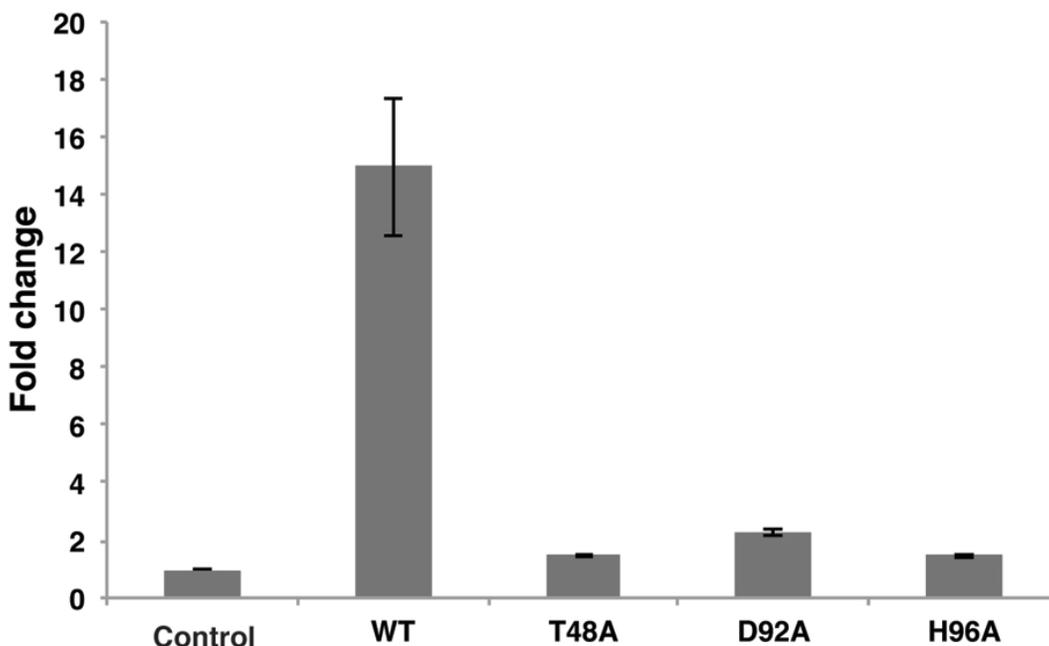


Figure 17. The *in vitro* Sec incorporation activity of the GTPase site mutants of human eEFSec. Mutations of conserved residues in GTPase site of human eEFSec (Thr48, Asp92, and His96) severely impaired read-through of the Sec UGA codon and selenoprotein synthesis *in vitro*. The luminescence intensities obtained for each eEFSec mutant and WT protein were normalized relative to the signal from a control sample, which did not contain eEFSec. Error bars represent standard deviation (s.d.) calculated from three replicates.

IV.2.4 The characterization of the putative Sec-binding site in human eEFSec

The predicted Sec-binding site is composed of residues from domains D1 (Phe53) and D2 (Asp229, His230, and Arg285) (Figure 17A). Asp229 and Arg285 are conserved across all kingdoms, while Phe53 and His230 are replaced with Tyr and Arg, respectively, in the bacterial SelB. The overall positive charge of the pocket, which is preserved in all organisms, is thought to complement the negatively charged selenocysteinyl moiety (41). By contrast, the amino acid-binding site in EF-Tu has the

overall negative charge, and it is comprised of Glu226, Asp227 and Asn285 from D2, and His67 from D1 (Figure 17B).

I tested the importance of the conserved residues and the overall charge of the pocket for eEFSec function by introducing single (D229A, H230A, R285A) and double alanine replacements (H230A_R285A). Additionally, the side chain of Arg285 was mutated to asparagine (R285N) to assess the effect of introducing the corresponding EF-Tu residue in that particular position. All mutants expressed and purified like WT eEFSec. However, a substitution of any amino acid within the Sec-binding pocket as well as the removal of a single positive charge was sufficient to completely abolish the ability of eEFSec to promote Sec incorporation and selenoprotein synthesis *in vitro* (Figure 17C). By contrast, the loss-of-function in the archaeal SelB was observed only upon removal of two positive charges from the pocket; the elimination of a single charge did not significantly affect SelB activity (41). Interestingly, the removal of both positive charges from the archaeal pocket did not shift the selectivity of SelB towards the near cognate Ser-tRNA^{Sec} (41).

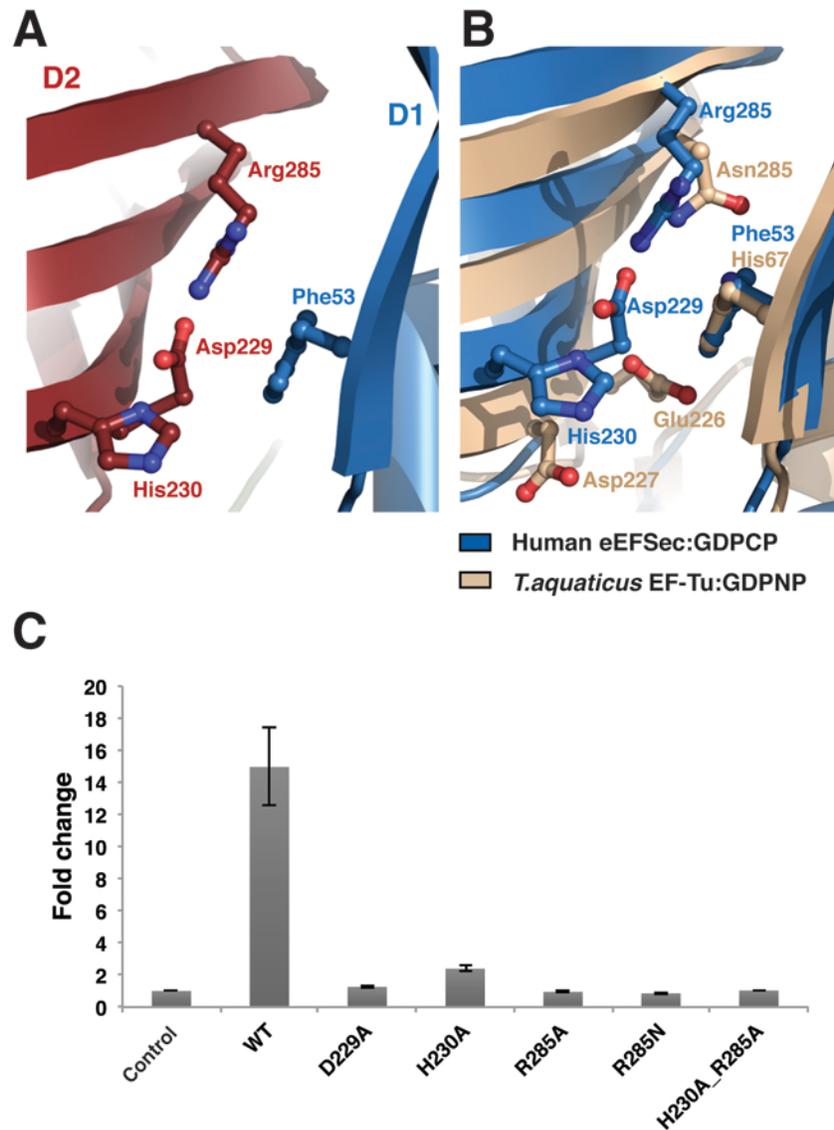


Figure 18. The putative Sec-binding site in human eEFSec. (A) The site is located at the interface of D1 (Phe53; blue) and D2 (Asp229, His230, Arg285; red). Asp229 and Arg285 are conserved across the kingdoms, whereas His230 and Phe53 are present in eukaryotic and archaeal orthologs, but are replaced by Arg and Tyr, respectively, in bacterial SelB. (B) The superimposition of the amino acid-binding sites reveals differences between eEFSec (blue) and EF-Tu (beige). The overall positive charge of the site in eEFSec is suggested to complement for negatively charged selenocysteinyl moiety of Sec-tRNA^{Sec}. (C) The replacement with alanine of any of the residues in the Sec-binding pocket (Asp229, His230, and Arg285) is sufficient to completely abolish the read-through of Sec UGA codon and selenoprotein synthesis *in vitro*. The luminescence intensities obtained for each mutant and WT eEFSec were normalized relative to the signal from a control sample, which did not contain eEFSec. Error bars represent standard deviation (s.d.) calculated from three replicates.

IV.3 Discussion

I characterized interactions between human eEFSec and guanine nucleotides and analogs using ITC. The results showed that eEFSec binds GDP and GTP with similar affinities, which is in good agreement with the previously reported findings for mouse eEFSec ($K_{d,GTP} = 0.3 \mu\text{M}$ vs $K_{d,GDP} = 0.6 \mu\text{M}$) (35), archaeal SelB ($K_{d,GTP} = 0.1 \mu\text{M}$ vs $K_{d,GDP} = 0.4 \mu\text{M}$) (33), and bacterial SelB ($K_{d,GTP} = 0.7 \mu\text{M}$ vs $K_{d,GDP} = 13 \mu\text{M}$) (103). In contrast, bacterial EF-Tu binds GDP with much higher affinity than GTP (1 and 60 nM, respectively) (104), and thus, it requires guanine nucleotide exchange factor (GEF) to cycle between the GDP- and GTP-bound states. Analogous to SelB, and in contrast to EF-Tu/EF1A, eEFSec lacks a consensus sequence required for GEF binding (34, 35, 105). Given that the GTP-to-GDP ratio in the cell is approximately 10:1, I suggest that eEFSec most likely does not require GEF to exchange GDP for GTP.

Further, though eEFSec has similar binding affinities for GTP analogs and GTP, the different ΔH and ΔS values were exerted upon binding. Similar differences were observed in the ITC study on the bacterial SelB (47). Therefore, the question was raised which analog is a faithful mimic of GTP when bound to the Sec tRNA-specific elongation factors? Based on the ΔH and ΔS values, one could argue that GDPCP would be the most appropriate GTP mimic for eEFSec, whereas GDPNP and GTP γ S would induce the GDP-bound conformation. However, my structural studies showed that both GDPCP and GDPNP lock eEFSec in the GTP-bound state. Therefore, I suggest that the ITC measurements alone may not be a reliable discriminator for the utility of the GTP analogs, at least in case of eEFSec and SelB, and that the observed discrepancies in

the thermodynamic parameters could be due to spontaneous hydrolysis of GTP γ S in solution, as well as different solvation around the β - γ -phosphate linkage in the analogs.

The detailed analysis of the GTPase site revealed that the T48A mutation did not affect the guanine nucleotide binding, whereas the mutations D92A and H96A had only modest impact. By contrast, the *in vitro* activity assays demonstrated that the GTPase site mutations completely abolished the ability of eEFSec to facilitate the Sec UGA codon read-through. Hence, my results suggested that the loss-of-function in the GTPase site mutants is not because of the impaired nucleotide binding, but most likely due to their inability to hydrolyze GTP. This may be particularly true for His96, the amino acid suggested to be essential for the ribosome-induced GTPase activity of EF-Tu. In my structures, the side chain of His96 is positioned relatively far from the nucleotide, which implies that this particular segment of the GTPase site may undergo an additional conformational change when bound to the translating ribosome. This may not be unprecedented, since in the structures of the ribosome-bound EF-Tu and EF-G the corresponding His86/His87 is repositioned towards the γ -phosphate of GTP through a H-bond with A2662 of the sarcin-ricin loop (SRL) of the 23S ribosomal RNA (66, 106). It is, therefore, reasonable to suggest that a similar interaction with the ribosome may stimulate the GTPase activity of eEFSec as well. Further structural and enzymatic studies utilizing a complete human decoding complex of the Sec UGA codon are needed to define the roles of Thr48, Asp92 and His96 in catalysis of GTP hydrolysis.

The mutational study showed that mutations in the Sec-binding pocket of eEFSec abolished the Sec incorporation activity *in vitro*. Because of a relatively large distance of

Asp229, His230 and Arg285 from the GTPase site (18, 16, 18.5 Å, respectively), I suggest that substitutions of those residues most likely did not affect the nucleotide binding or the GTPase activity of eEFSec. Rather, I speculate that the binding of Sec-tRNA^{Sec}, specifically recognition of the selenocysteinyl moiety, is affected by the mutations. Furthermore, my data showed that not only a single amino acid replacement, but also the removal of a single positive charge within the site completely disrupted the eEFSec function. A reasonable explanation is that elimination of any change disrupted a distinct environment required for recognition and binding of the selenocysteinyl moiety. To fully understand the mechanism by which eEFSec selects Sec-tRNA^{Sec} over its precursor Ser-tRNA^{Sec} and other elongator aa-tRNAs, and whether the eEFSec specificity for Sec can be changed towards another amino acid, further binding and activity assays involving Sec-tRNA^{Sec}, Ser-tRNA^{Sec}, and WT and mutant eEFSec constructs are warranted.

V. The GTP-to-GDP exchange in human eEFSec induces a non-canonical conformational change of D4, and not a canonical rearrangement of D1

V.1 Introduction

Studies on the bacterial protein synthesis established that a conformational change in EF-Tu coupled to GTP hydrolysis is critical for the elongation phase of translation. Only in the GTP-bound state EF-Tu can bind and deliver aa-tRNAs to the translating ribosome. After the codon-anticodon interactions are formed between the A-site aa-tRNA and the mRNA in the decoding center of the small ribosomal subunit, the GTPase activity and the conformational changes in EF-Tu are induced. In particular, minor structural adjustments in the aa-tRNA lead to a more intimate interaction between EF-Tu and the sarcin-ricin loop of the 23S rRNA. This causes the side chain of His84 to be closer to GTP. After GTP hydrolysis, D1 rotates $\sim 90^\circ$ away from D2 and D3, and the globular structure of the GTP-bound state is transformed into a more extended structure of the GDP-bound EF-Tu (64). This leads to disruption of the aminoacyl-binding pocket, which in turn causes EF-Tu:GDP to dissociate from the A-site aa-tRNA and the ribosome. The CCA-end of the aa-tRNA is then accommodated within the peptidyl-transferase center (PTC) where the aminoacyl group is poised for the reaction of peptide bond formation to occur. It was suggested that the analogous mechanism is employed by the archaeal (87) and mammalian EF1A (90).

The question, however, remained if the Sec tRNA-specific elongation factors, eEFSec and SelB, undergo the same conformational change upon the GTP-to-GDP

exchange and whether the same mechanism is utilized for the tRNA release. A structural study on the archaeal SelB revealed no major differences between the GTP- and GDP-bound states (41), whereas the bacterial SelB structure was determined only in the presence of a GTP analog (42). The observation that the archaeal SelB does not alter its structure when bound to GDP was scrutinized (42, 47) because ligand soaking was used to capture the functional states. In other words, it was suggested that the apo-SelB crystal form perhaps restricted the movement of D1 even in the presence of a nucleotide that would otherwise elicit such a movement. I obtained crystals of eEFSec:GDP that are non-isomorphous to crystals of eEFSec:GDPCP(GDPNP). Here, I present a detailed structural comparison of the GDP- and GTP-bound states of human eEFSec.

V.2 Experimental Results

V.2.1 The non-canonical conformational change in human eEFSec induced by the GTP-to-GDP exchange

Human eEFSec was co-crystallized with GDP (eEFSec:GDP). The crystals belonged to a monoclinic space group ($P2_1$) and diffracted X-rays to 3.3 Å resolution. The structure was determined by molecular replacement using individual domains of eEFSec:GDPNP as search models. The final model of eEFSec:GDP was refined to $R_{\text{work}}/R_{\text{free}}$ of 0.30/0.34 (TABLE VI). The asymmetric unit contained two eEFSec monomers in a head-to-tail arrangement that were almost indistinguishable (r.m.s.d. value of 0.9 Å).

TABLE VI. DATA COLLECTION AND REFINEMENT STATISTICS FOR eEFSec:GDP

eEFSec:GDP	
Data collection	
Space group	P 2 ₁
Cell dimensions	
<i>a, b, c</i> (Å)	58.7, 96.9, 125.4
	β = 90.25°
Resolution (Å) ^a	43.00 – 3.00
	(3.05 – 3.00)
R _{sym} or R _{merge}	0.13
I / σ ^a	7.6 (0.4)
Completeness (%) ^a	92.4 (51.8)
Redundancy ^a	3.7 (2.6)
Refinement	
Resolution (Å)	43.00 – 3.25
No. of reflections	22,047
R _{work} / R _{free}	0.30 / 0.34
No. of atoms	
Protein	5,983
Ligand/ion	56
Water	21
<i>B</i> -factors	
Protein	81.6
Ligand/ion	107.4
Water	68.2
R.m.s. deviations	
Bond lengths (Å)	0.009
Bond angles (°)	1.96

^a Values in parentheses are for the highest resolution shell.

Remarkably, a comparison of the GTP- and GDP-bound states of human eEFSec revealed that the structures of the main functional states are similar (Figure 18). The overlay of eEFSec:GDP onto eEFSec:GDPNP and eEFSec:GDPCP yielded r.m.s.d. values of 2.5 and 2.1 Å, respectively. Surprisingly, no significant conformational change within EF-Tu-like domain of eEFSec was observed upon GDP binding. Whereas GTP-to-GDP exchange causes $\sim 90^\circ$ rotation of D1 relative to D2 and D3 in EF-Tu and EF1A, the same event does not alter position of D1 in eEFSec. In other words, D1 adopts practically the same orientation relative to D2 and D3 in both functional states of eEFSec. Perhaps more importantly, when the view is oriented so that the presumed tRNA-binding (i.e. ventral) side is placed on the left hand side and perpendicular to the plane of paper, an unexpected observation is revealed that the C-terminal D4 of eEFSec swings towards the dorsal side by $\sim 26^\circ$ (Figure 18; right panel). The swing of D4 translates the entire domain by >15 Å.

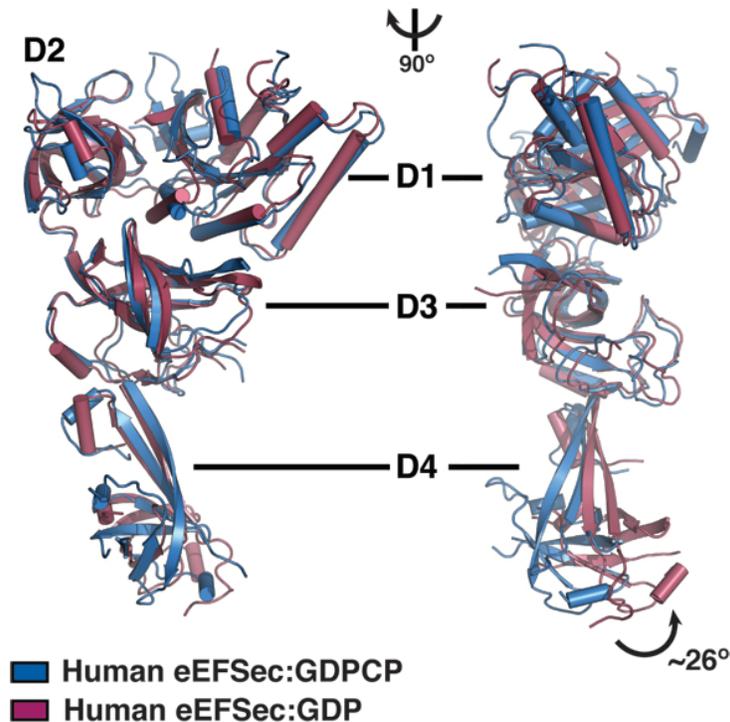


Figure 19. The GTP-to-GDP transition in human eEFSec induces a conformational change in D4, but not in D1. The global superimposition of eEFSec:GDPCP (blue) and eEFSec:GDP (light red) reveals a lack of the canonical conformational change in the EF-Tu-like domain. Instead, the C-terminal D4 swings $\sim 26^\circ$ towards the dorsal face of the molecule and away from the tRNA-binding site. Two views related by $\sim 90^\circ$ rotation around a vertical axis are shown. The view on the left is oriented so that the tRNA-binding site is in the paper plane. The view on the right is oriented so that the putative tRNA-binding site (or ventral face) is on the left and perpendicular to the paper plane.

It was suggested that D4 might ‘sense’ nucleotide binding to D1 and that it can regulate the GTPase activity of the mammalian eEFSec (44). Thus, the observed domain movement might be of functional importance. Similar structural rearrangements have been reported for the archaeal SelB (41) and the translation initiation factor IF2/eIF5B (97). Whereas the rearrangement in SelB was not discussed outside the realm of local

structure dynamics, the domain movement in IF2/eIF5B was considered functionally relevant.

However, the question could be raised whether the crystal packing in the eEFSec:GDP crystals hindered the movement of D1 relative to D1-3, or whether the D4 domain is inherently structurally flexible, which would lessen the significance of observed structural differences between the GTP- and GDP-bound states. To assess these possibilities, I analyzed eEFSec:GDP and eEFSec:GDPCP complexes in solution using size-exclusion chromatography (SEC) coupled to Small-angle X-ray scattering (SEC-SAXS). The results show that R_g of eEFSec does not alter upon the GTP-to-GDP exchange (R_g of $\sim 35\text{\AA}$) (Figure 19). Furthermore, eEFSec adopts a strikingly similar overall molecular shape irrespective of the bound nucleotide (Figure 20A and 20B, left panels). Remarkably, a side-to-side comparison of the experimental SAXS envelopes clearly shows that D4 of eEFSec adopts different orientations in the GDP- and GTP-bound states (Figure 20A and 20B, right panels). Importantly, the crystal structures agree with the corresponding SAXS envelopes, while the designed 'canonical' model of eEFSec:GDP could not be fitted into the eEFSec:GDP envelope (Figure 21). Taken together, my crystal structures and SAXS data suggest that the conformational change of D4 as well as the lack of canonical structural rearrangement of D1, observed upon the nucleotide exchange in the crystals, also occur in solution. Therefore, the conformational change in the GDP-bound state of eEFSec is not a crystallization artifact.

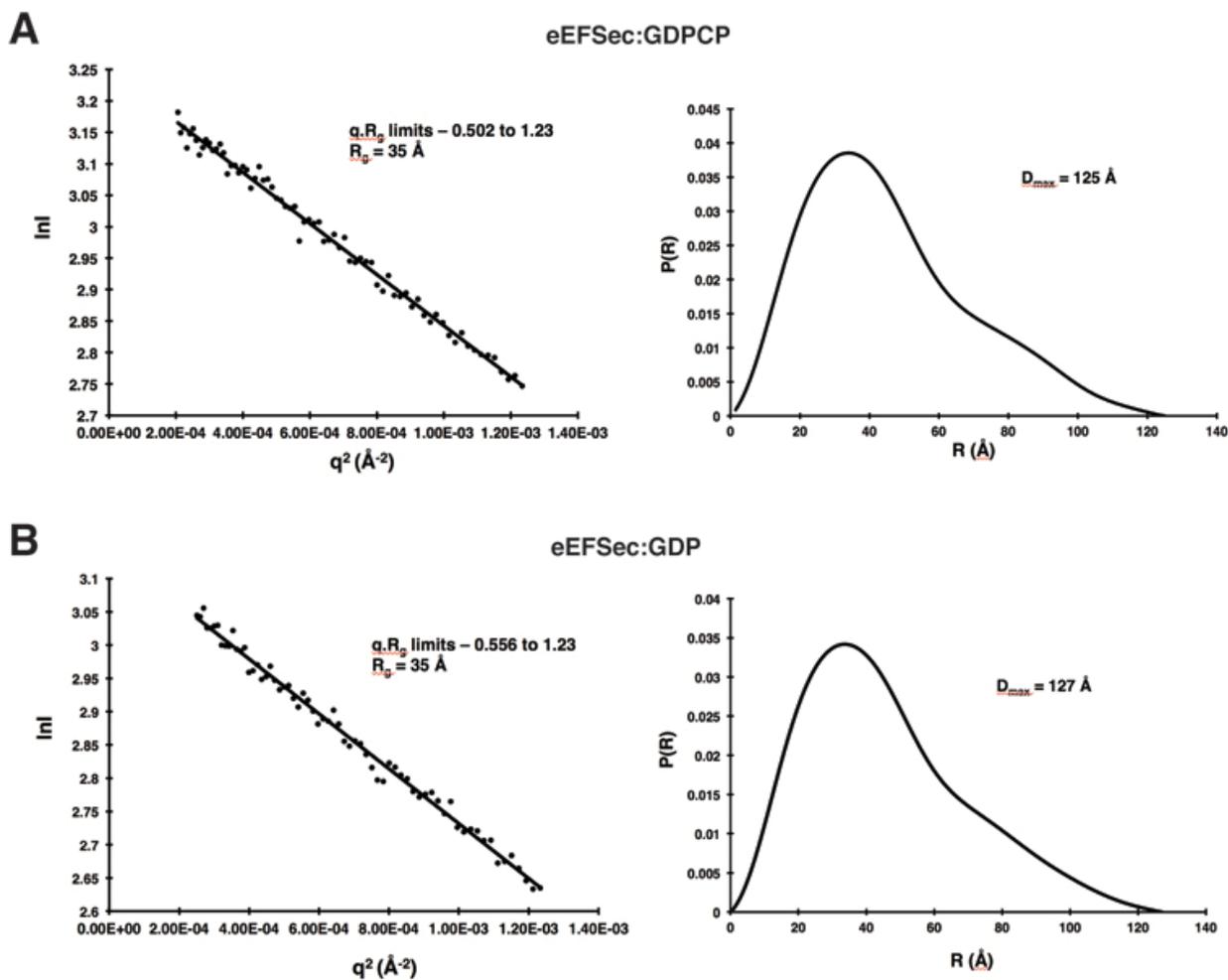


Figure 20. Small-angle X-Ray scattering (SAXS) analysis of the GTP- and GDP-bound states of human eEFSec. The Guinier (left panels) and Pair distance distribution function ($P(R)$, right panels) plots, generated for eEFSec:GDPCP (**A**) and eEFSec:GDP (**B**) complexes, reveal that R_g of eEFSec does not alter upon the nucleotide exchange.

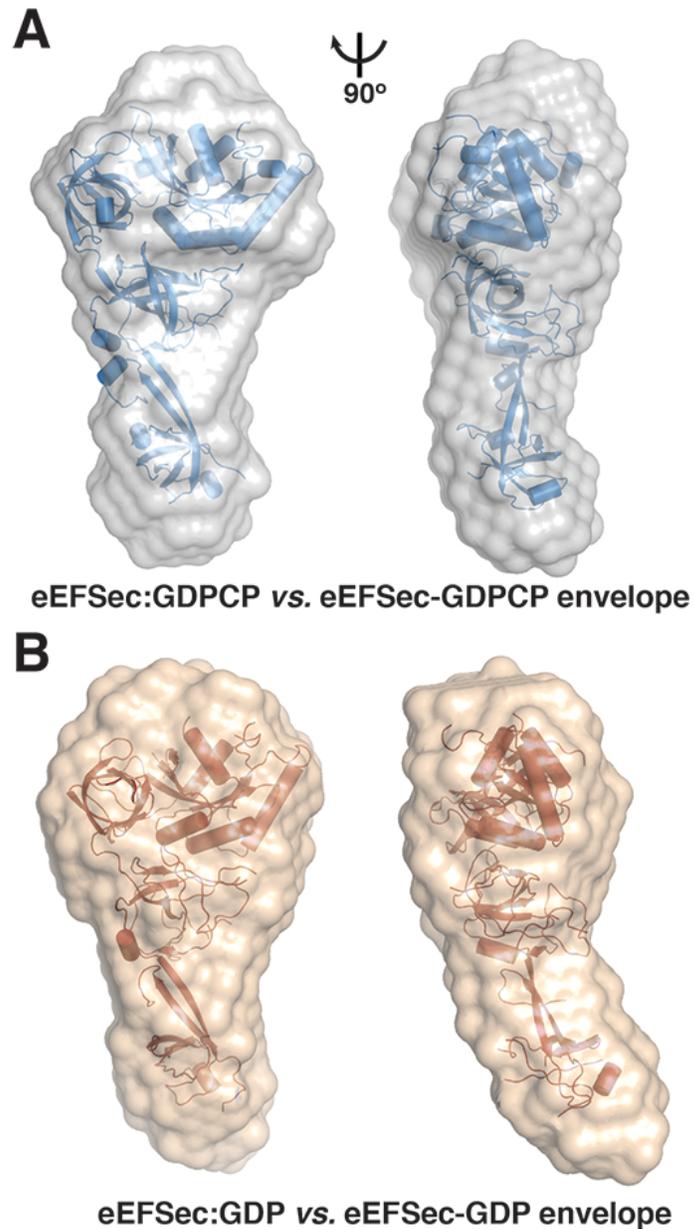
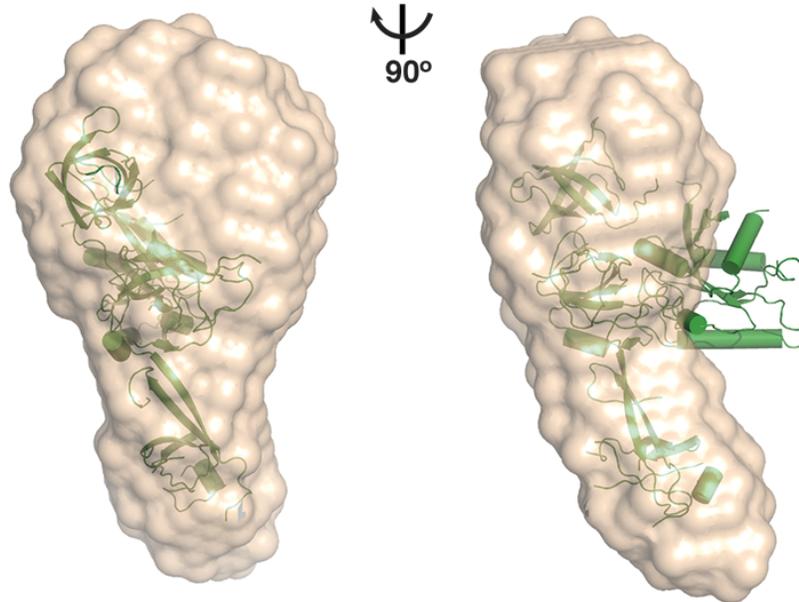


Figure 21. The superimposition of SAXS molecular envelopes onto the corresponding crystal structures of the GTP- and GDP-bound states of eEFSec. **(A)** The eEFSec:GDPCP crystal structure (blue) agrees well with its SAXS envelope (gray surface). **(B)** Also, the crystal structure of the GDP-bound eEFSec (red) superimposes well onto the envelope derived from the eEFSec:GDP complex (beige surface). Two views related by $\sim 90^\circ$ rotation around the vertical axis are shown. The EF-Tu-like domain is kept in a similar orientation in all panels.



'Canonical' eEFSec:GDP vs. eEFSec-GDP envelope

Figure 22. The comparison of the 'canonical' model of the GDP-bound state of human eEFSec with eEFSec:GDP SAXS envelope. The 'canonical' model of the GDP-bound eEFSec (green) does not fit into the SAXS envelope of eEFSec:GDP (beige surface). The canonical domain arrangement would place D1 outside of the molecular envelope. Two views related by $\sim 90^\circ$ rotation around the vertical axis are shown.

Lastly, the lack of canonical conformational change in eEFSec suggested that such movement might be prevented by the presence of unique structural elements that are relatively conserved among eEFSec orthologs, but absent in EF-Tu and EF1A. The structural analysis suggested that the loop $\beta 17$ - $\beta 18$ (residues 353-370), situated on the dorsal side of D3 and opposite from the putative tRNA-binding face of the molecule, as a good candidate for such element (Figure 21A). The loop occupies the same area where helix $\alpha 4$ of D1 would have been positioned if the canonical conformational change in eEFSec occurred (Figure 21B), thus leading to a steric clash. Furthermore, the sequence alignment suggested that the element corresponding to loop $\beta 17$ - $\beta 18$ is

also present in eEFSec from other eukaryotic organisms, archaeal SelB, but not in bacterial SelB (Figure 22). Therefore, I suggest that loop $\beta 17$ - $\beta 18$ might be an important structural element in eEFSec that prevents rotation of D1 during GTP-to-GDP exchange.

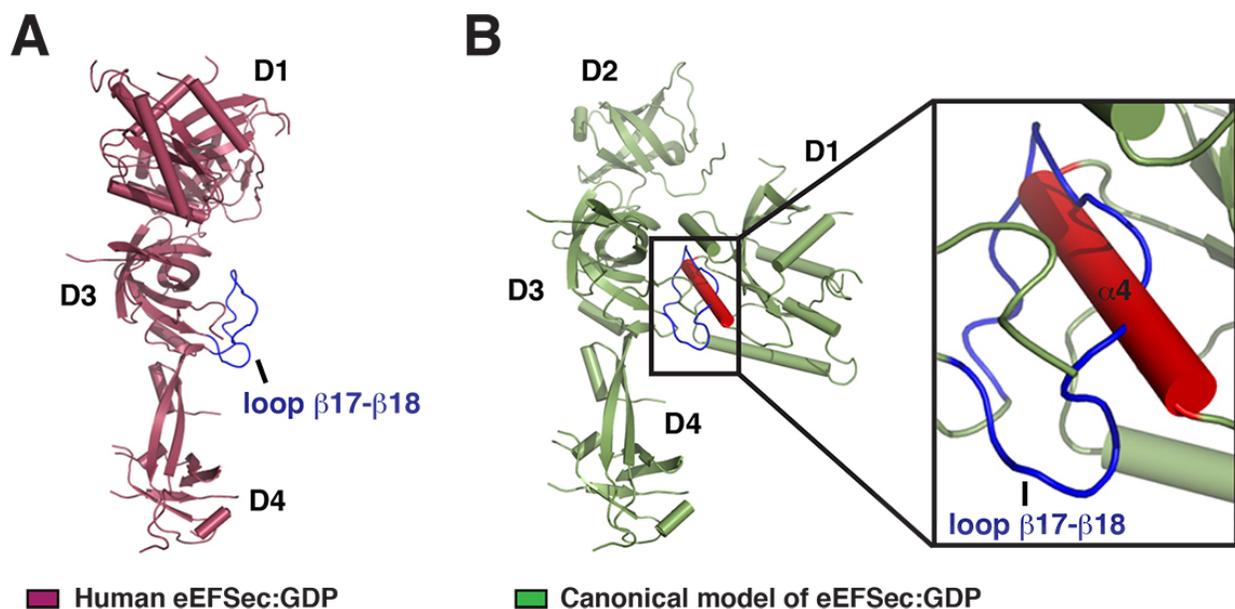


Figure 23. An enlarged loop in D3 might prevent the canonical conformational change in the EF-Tu-like domain of human eEFSec upon the GTP-to-GDP exchange. (A) The rotation of D1 relative to D2 and D3 induced by the GTP-to-GDP exchange, which was observed in EF-Tu and EF1A, might be prevented in eEFSec by the enlarged loop $\beta 17$ - $\beta 18$ (blue) located at the dorsal face of D3. **(B)** Modeling of the ‘canonical’ GDP-bound state of eEFSec revealed that the loop $\beta 17$ - $\beta 18$ (blue) would clash with helix $\alpha 4$ (red) from D1 if the canonical conformational change occurred in eEFSec upon GTP-to-GDP transition. The steric clash is highlighted with black box and shown in a close-up view.

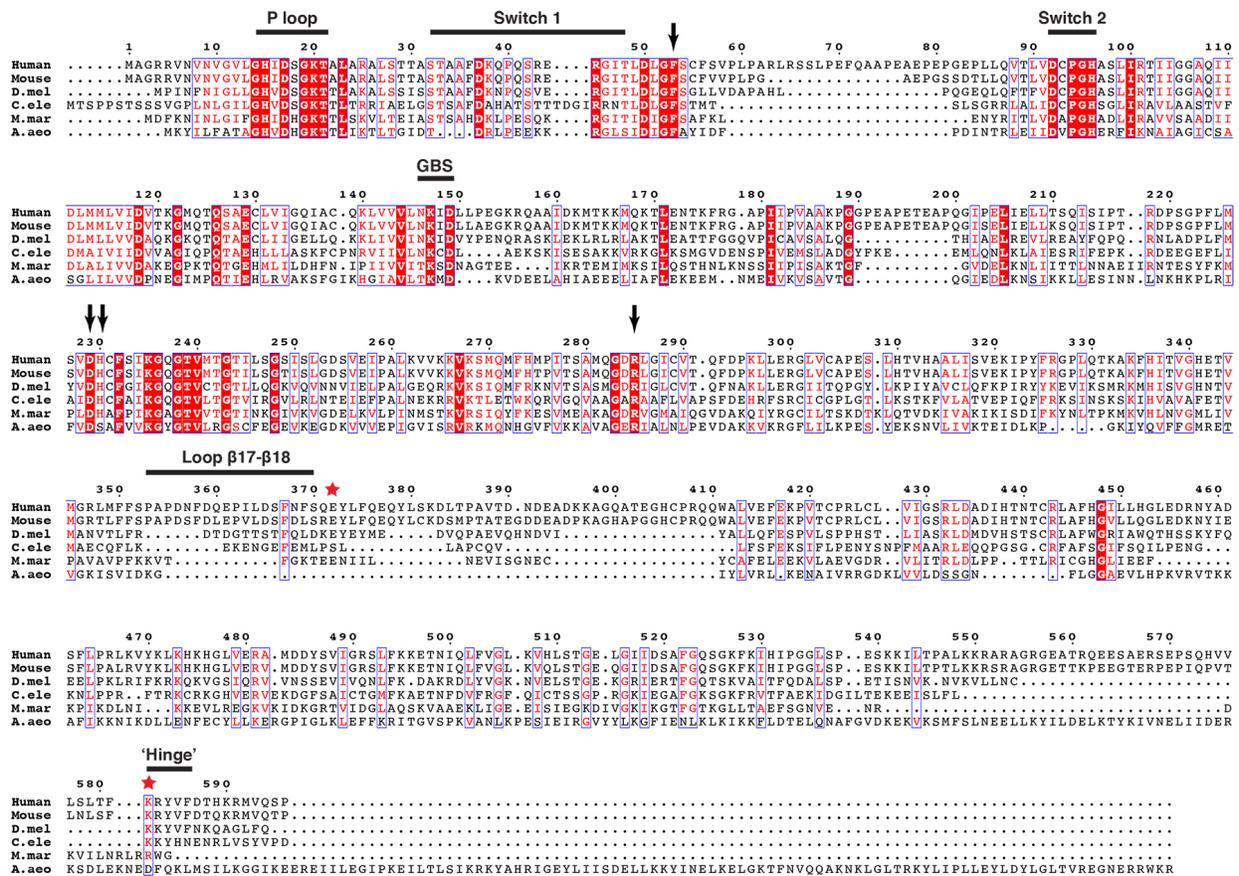


Figure 24. Structure-based sequence alignment of eEFSec and SelB orthologs. The components of the GTPase site (i.e. P loop, Switch 1, Switch 2, and guanine-binding sequence), loop β 17- β 18, and the 'KRYVF' sequence of the 'hinge' region are highlighted with black bars. The guanine-binding sequence is labeled as 'GBS'. Arrows point to residues of the Sec-binding pocket, and red stars highlight the ³⁷²Glu-Lys⁵⁸² salt bridge.

V.2.2 Minor structural rearrangements in the EF-Tu-like domain of eEFSec upon the GTP-to-GDP exchange

A series of minor structural rearrangements were observed in the EF-Tu-like domain of eEFSec after nucleotide exchange. Most evidently, D1 and D2 move in opposite directions in a ratchet-like movement. In particular, D1 moves towards the

ventral side, while D2 slightly rotates towards the dorsal side. These movements cause the GTPase site to assume a more open conformation and the Sec-binding pocket to constrict.

Also, a segment of switch 1 (residues 44-51), which harbors the catalytically important Thr48, is ordered in the GTP-bound state and it provides the roof for the γ -phosphate-binding pocket. By contrast, switch 1 and helix α 2 (residues 64-69) are disordered in the GDP-bound state, given the complete lack of electron density in this region. These elements could not be modeled (Figure 23B). Equally significant is the movement of switch 2 (residues 92-96), the adjacent β 3- α 3 loop, and helix α 3 (residues 101-109) upon GDP binding. Switch 2, which harbors Asp92, swings away from the GTPase site and towards the Sec-binding pocket. As a consequence, Asp92 is no longer optimally positioned to participate in Mg^{2+} coordination in the GDP-bound state. Interestingly, the Mg^{2+} ion is not present in the eEFSec:GDP structure, even though the protein sample used in crystallization contained 5 mM $MgCl_2$. This is consistent with the structure of the GDP-bound state of rabbit eEF1A isoform (90), which suggested that GDP binding is a Mg^{2+} -independent process. Concurrently, loop β 3- α 3 and the neighboring helix α 3 tilt $\sim 5\text{\AA}$ closer towards D2 and away from the GTPase site. This movement contributes to the constriction of the amino acid-recognition site.

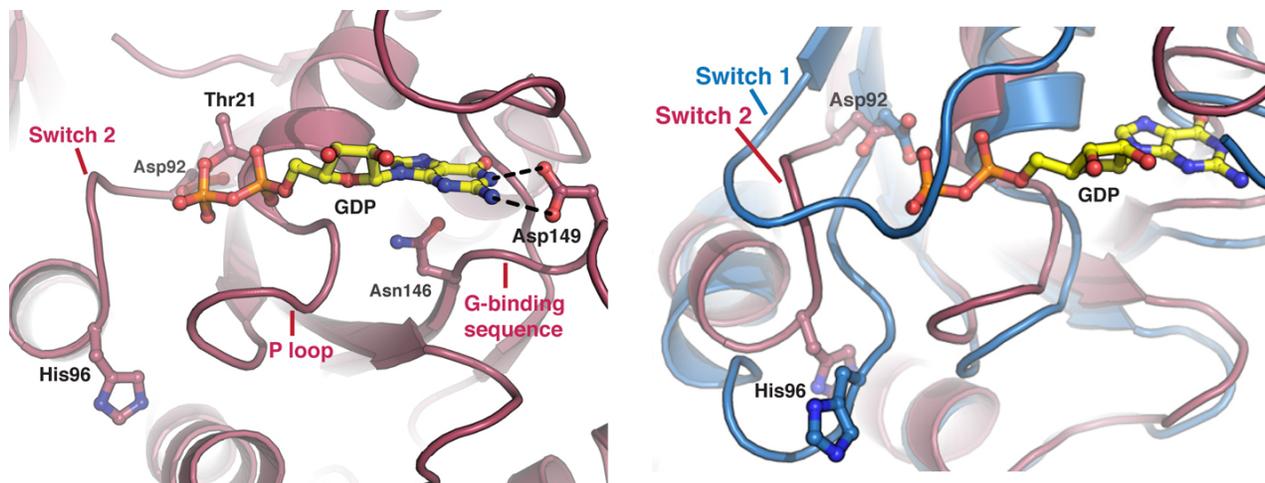


Figure 25. The structural rearrangements in the GTPase site of human eEFSec upon the GTP-to-GDP transition. (Left) The GTPase site in eEFSec:GDP. **(Right)** The GTP-to-GDP transition induces structural rearrangements mainly restricted to switch 1 and switch 2 regions. In the GDP-bound state (pink), switch 1 is completely disordered (see page 74 for details), while switch 2, containing Asp92 and His96, swings away from the GTPase site and towards the Sec-binding pocket. These elements are oriented differently in the GTP-bound state (blue).

The GDP binding also causes subtle structural changes in D2, with the elements constituting the Sec-binding pocket being particularly impacted. The side chain of His230, a segment of strand $\beta 7$, and the entire $\beta 7$ - $\beta 8$ turn are completely disordered in the GDP-bound state structure (Figure 24). In addition, turn $\beta 10$ - $\beta 11$ (residues 258-264), loop $\beta 14$ - $\beta 15$ (residues 304-309) and helix $\alpha 7$ (residues 290-300), which form the entrance of the Sec-binding site, move ~ 3.5 Å towards D1.

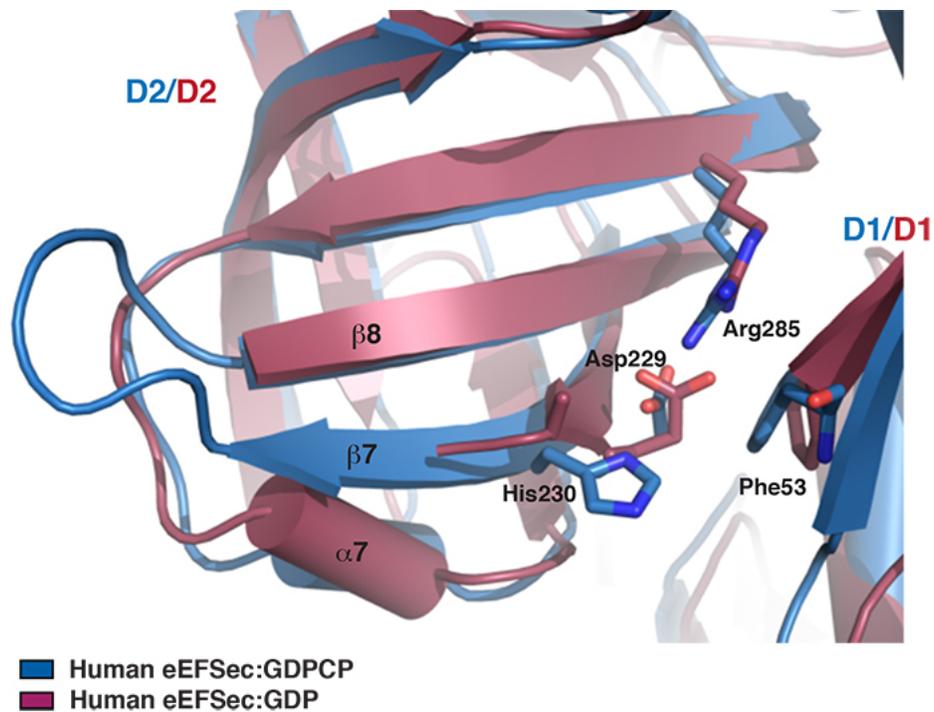


Figure 26. The structural rearrangements within the putative Sec-binding pocket of human eEFSec upon the GTP-to-GDP exchange. The GTP-to-GDP exchange causes subtle movements of the residues constituting the Sec-binding site, with His230 being particularly affected. The side chain of His230, the following segment of strand $\beta 7$ as well as the entire $\beta 7$ - $\beta 8$ turn, became completely disordered. Furthermore, GDP-binding to eEFSec caused the movement of several loops and helix $\alpha 7$ towards D1, which contributed to the constriction of the Sec-recognition site.

Lastly, minor conformational adjustments were also observed in D3. Loops $\beta 15$ - $\beta 16$ (residues 321-331) and $\beta 18$ - $\beta 19$ (residues 375-411) swing towards the dorsal side of eEFSec, contributing to the more open conformation of eEFSec, and possibly to the mechanism of release of Sec-tRNA^{Sec}.

V.3 Discussion

The structure of the GDP-bound state of human eEFSec enabled visualization of conformational rearrangements coupled to the nucleotide exchange, pivotal for elongation of selenoprotein synthesis. Although obtained in the absence of other components of the decoding complex (e.g. Sec-tRNA^{Sec}, 80S ribosome, SECIS and SBP2), I propose that structures of eEFSec bound to GTP analogs and GDP represent the true physiological states. This would be consistent with the fact that EF-Tu adopted the same structure in isolation (62) and in complex with Phe-tRNA^{Phe} (63) and the ribosome (65). A structural comparison of the GTP- and GDP-bound states of eEFSec revealed the unprecedented lack of conformational change within N-terminal EF-Tu-like domain, and the non-canonical movement of the appended C-terminal D4.

In contrast to EF-Tu and EF1A, D1 of eEFSec assumes the same orientation relative to D2 and D3 in both the GTP- and GDP-bound states. Only subtle rearrangements have been observed in D1-3 upon the nucleotide exchange event. These movements cause widening of the GTPase site and constriction of the Sec-binding pocket. In particular, a segment of switch 1 that is upstream of the lid of the selenocysteine-binding pocket is completely disordered, while switch 2 and helix α 3 move away from the GTPase site and towards D2, most likely contributing to the constriction of the amino acid-recognition site.

In addition, tightening of the Sec-binding pocket is achieved by the subtle rearrangements of the residues within the pocket (Asp229, His230, Arg285), as well as the movements of several structural elements in D2 (e.g. loop β 10- β 11, loop β 14- β 15

and helix $\alpha 7$) towards D1. Such changes perhaps cause a sufficient decrease in eEFSec binding affinity towards the selenocysteiny moiety, which in turn would lead to the release of Sec-tRNA^{Sec}. Furthermore, this could provide an explanation why the GTP-bound state of the mammalian eEFSec, and by extension bacterial SelB, exhibits a markedly stronger binding affinity towards Sec-tRNA^{Sec} (44, 94).

The signal that the nucleotide exchange occurred is transferred to D3 and then to D4. In D3, loops $\beta 15$ - $\beta 16$ and $\beta 18$ - $\beta 19$ move towards the dorsal face of eEFSec. Given that these loops may be involved in tRNA^{Sec} binding, these movements could contribute to the release of Sec-tRNA^{Sec}. Lastly, D4 swings in the same direction as solvent-exposed loops $\beta 15$ - $\beta 16$ and $\beta 18$ - $\beta 19$ of D3, albeit to the greater extent. It is important to emphasize that such movement of D4 was noted only after the GTP-bound state structure was overlaid onto that of the GDP-bound state. The swing of D4 was not present when comparing individual eEFSec monomers from the same asymmetric unit. By contrast, structural differences in the area of D4 were observed between archaeal SelB monomers within the asymmetric unit of the SelB:GDP binary complex structure. The authors suggested this to be caused by the inherent flexibility of D4. However, because GDP was soaked into the apo-SelB crystals and not co-crystallized with SelB, that the occupancy of GDP within the GTPase sites of SelB monomers was different. This would cause structural heterogeneity within the asymmetric unit, which would then be ascribed as the inherent structural feature of SelB.

Although in my structural studies I used the co-crystallization method rather than soaking crystals in the ligand containing solution, there was still a possibility that the

lack of the canonical conformational change of D1 and the observed swing of D4 might be due to the crystal packing. However, the SAXS results showed that the GTP- and GDP-bound states of human eEFSec are the same in solution and in the crystal. An agreement of the crystal structures with the corresponding molecular envelopes confirmed the lack of the canonical conformational change in the EF-Tu-like domain and the presence of the swing of D4 towards the dorsal face of the molecule. Moreover, a model of the 'canonical' GDP-bound state of eEFSec did not agree with the experimental SAXS envelope of eEFSec:GDP, further indicating that a rotation of D1 does not occur upon the GTP-to-GDP exchange in eEFSec.

I wondered whether the evolutionary pressure yielded structural elements unique to eEFSec (and SelB) that would prevent the rotation of D1 relative to D2 and D3. The structural analysis, combined with the structure-based sequence alignment and modeling studies suggested that loop β 17- β 18 in D3, located on the dorsal face of the molecule, could play a role in that process. It is not known whether deletion of the loop would allow the structural rearrangement of D1 upon the GTP-to-GDP exchange. To assess that, further studies are needed. Also, it is conceivable that additional elements hinder the canonical conformational change, but these remain to be determined.

The structures of the GDP- and GTP-bound states of eEFSec revealed the unprecedented conformational change promoted by nucleotide exchange and suggested that the release of Sec-tRNA^{Sec} during decoding of the Sec codon might occur *via* a mechanism that is distinct from the canonical mechanism utilized by EF-Tu and EF1A.

VI. The proposed mechanism of decoding of the Sec UGA codon

VI.1 Introduction

My results suggest that the nature of conformational changes in human eEFSec upon the GTP-to-GDP exchange is different from that in the general elongation factors, EF1A and EF-Tu (see: Chapter V). While the nucleotide exchange causes the major conformational change in the C-terminal D4 of eEFSec, only minor structural rearrangements occur within its N-terminal EF-Tu-like domain that lead to the opening of the nucleotide-binding site and the constriction of the Sec-binding pocket. Such findings suggest that eEFSec releases Sec-tRNA^{Sec} upon GTP hydrolysis by a distinct mechanism. Additionally, I have shown that the Sec-binding site constitutes a unique environment for selection and differentiation of the selenocysteinyl moiety from other aminoacyl groups (see Chapter IV). Taken together, I speculate that human eEFSec employs a non-canonical mechanism to facilitate the elongation phase of the selenoprotein mRNA translation.

However, the question remains as to how eEFSec recognizes only Sec-tRNA^{Sec}? So far, no structural studies have been reported for the complex between eEFSec (or SelB) and Sec-tRNA^{Sec}. A recent study on the bacterial ortholog presented a model of the ternary complex between SelB, the unacylated tRNA^{Sec}, and L-cysteine (a substitute for Sec), but the definitive conclusions about the recognition mechanism were not drawn (42). Here, I asked if any of the structural features in EF-Tu responsible for binding to the aminoacyl-tRNA are conserved in eEFSec. Also, I wondered if distinct structural

elements present in eEFSec and SelB, but absent in EF-Tu/EF1A, drive the productive interaction with tRNA^{Sec}. Here, I present a homology-based model of the ternary eEFSec:GDPCP:tRNA^{Sec} complex, and a structural comparison with the EF-Tu:GDPNP:Phe-tRNA^{Phe} ternary complex. Lastly, based on all my results presented in this thesis I propose the mechanism of decoding of the Sec UGA codon in humans.

VI.2 Experimental Results

VI.2.1 A homology-based model of the eEFSec:tRNA^{Sec} complex

I superimposed eEFSec:GDPCP onto EF-Tu:GDPNP:Phe-tRNA^{Phe} ternary complex (63), and analyzed the conserved structural elements between human eEFSec and *Thermus aquaticus* EF-Tu, which could be involved in binding to Sec-tRNA^{Sec} (Figure 25A). EF-Tu has been shown to interact with the 3' and 5' ends, and the T stem of the canonical aa-tRNA (63).

The 3' CCA-aminoacyl end of the aa-tRNA is bound in the groove located at the interface of domains D1 and D2. In the EF-Tu complex, the base of A76 is sandwiched between two loops, to which the corresponding regions in eEFSec are loops β 7- β 8 (residues 232-239) and β 11- β 12 (residues 272-277). Furthermore, A76 of the Phe-tRNA^{Phe} makes contacts with the side chains of Glu271, Leu289, Val237 and Ile231, which in eEFSec correspond to Gln271, Cys289, Val240 and Ile234, respectively (TABLE VII). Given that the GTP-to-GDP exchange in eEFSec causes disorder of the loop β 7- β 8 and the movement towards D1 of loop β 11- β 12, it is reasonable to suggest that those two structural elements are involved in binding of Sec-tRNA^{Sec}. While the

GTP binding to eEFSec perhaps stabilizes the binding site for A76, the GTP hydrolysis may contribute to the flexibility of the pocket, which in turn would decrease the affinity for the CCA-Sec of Sec-tRNA^{Sec}.

The 5'-end of Phe-tRNA^{Phe} binds to the junction between domains D1, D2 and D3 of EF-Tu, where it interacts with a loop ³⁰⁰RGQV³⁰³ from D2 and residues Lys90 and Asn91 from D1. The corresponding elements in eEFSec are the side chains of Arg101 and Thr102 from helix α 3 in D1, Arg300 from loop α 7- β 14, and a part of strand β 14 spanning residues Gly301, Lys302, and Val303 (TABLE VII). My analysis of the structural rearrangements coupled to the GTP-to-GDP exchange in eEFSec revealed that loop α 7- β 14 and strand β 14 move closer to D1, whereas helix α 3 tilts towards D2, thus narrowing the space between domains D1, D2, and D3. Such movements could presumably contribute to the release of the 5'-end of the Sec-tRNA^{Sec} upon GTP hydrolysis.

Furthermore, Arg330, His331, Gln341, Thr350, Lys376, and Gly391 in EF-Tu interact with the T-stem of Phe-tRNA^{Phe}. The analogous residues in eEFSec are Arg326, Gly327, His336, Val344, Lys418, and Leu433, respectively (TABLE VII). Arg326 and Gly327 are located in loop β 15- β 16, which moves away from the putative tRNA-binding (or ventral) side after the GTP-to-GDP exchange. However, the most interesting are the structural changes affecting His336 and Leu433. The side chain of His336, which resides in strand β 16, rotates $\sim 50^\circ$ towards D1 and away from the tRNA-binding site, whereas the side chain of Leu433, which is in loop β 21- β 22, slides ~ 5 Å and rotates $\sim 180^\circ$ in the same direction. Interestingly, loop β 21- β 22 (residues 431-445)

in eEFSec faces the tRNA-binding side, it is enlarged when compared to the corresponding loop in EF-Tu (residues 390-393), and is partially disordered in my structures (Figure 25B). Since the corresponding loop 390-393 in EF-Tu is involved in binding to the T-stem, I speculate that loop β 21- β 22 in eEFSec could be responsible for 'measuring' the length of the acceptor-T ψ C arm of tRNA^{Sec}. Given the results of my analysis, I propose that Arg326, Gly327, His336 and Leu433 in D3 of eEFSec could play a significant role in binding of the acceptor-T ψ C arm of Sec-tRNA^{Sec}.

TABLE VII. STRUCTURAL ELEMENTS IN EF-Tu INVOLVED IN BINDING TO aa-tRNA

Structural element of an aa-tRNA	EF-Tu	Human eEFSec
CCA-aa	Ile231	Ile234
	Val237	Val240
	Glu271	Gln271
	Leu289	Cys289
5'-end	Lys90	Arg101
	Asn91	Thr102
	Arg300	Arg300
	Gly301	Gly301
	Gln302	Leu302
	Val303	Val303
T-stem	Arg330	Arg326
	His331	Gly327
	Gln341	His336
	Thr350	Val344
	Lys376	Lys418
	Gly391	Leu433

Residues conserved in human eEFSec and *T. aquaticus* EF-Tu are shown in bold.

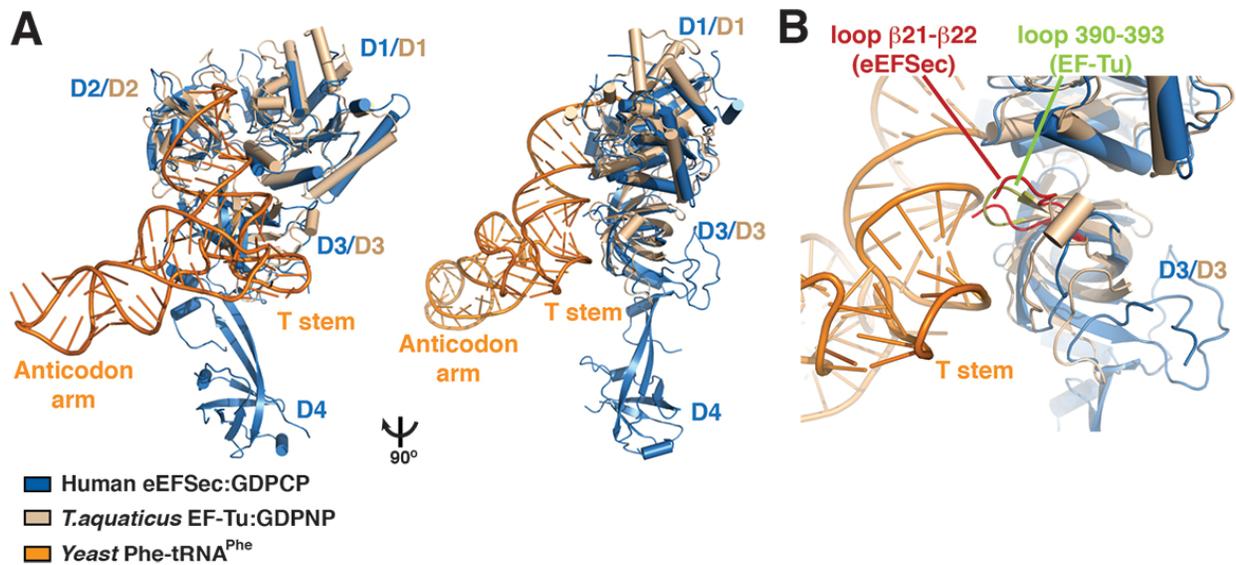


Figure 27. A comparison between human eEFSec:GDPCP and the EF-Tu:GDPNP:Phe-tRNA^{Phe} ternary complex. (A) The overlay of eEFSec (blue) onto the ternary EF-Tu complex (protein in beige and tRNA in orange) is shown in two views rotated $\sim 90^\circ$ clockwise around the vertical axis. This allowed identification of structural elements in eEFSec that could be involved in binding to Sec-tRNA^{Sec}. **(B)** A closer inspection suggests that D3 of eEFSec contains an enlarged loop β 21- β 22 (red), which corresponds to loop 390-393 (green) in EF-Tu. This particular loop might play a role in ‘measuring’ the length of the acceptor-T Ψ C arm of tRNA^{Sec}.

VI.2.2 The conservation of the putative tRNA-recognition elements in eEFSec and

SeIB

I wondered whether residues in human eEFSec, proposed to be important for Sec-tRNA^{Sec} recognition, are conserved in SeIB. The superimposition of eEFSec:GDPCP onto the archaeal (*Methanococcus maripaludis*) and bacterial (*Aquifex aeolicus*) SeIB revealed that side chains suggested to bind A76 (Val 240 and Gln271) and the 5'-end (Arg300) of Sec-tRNA^{Sec} are conserved across kingdoms. By contrast, Arg101 and Ile234 are conserved in eukaryotes and archaea, but replaced with Lys and

Val in bacteria (TABLE VIII). Further, residues in D3 that are proposed to bind to the T-stem of tRNA^{Sec} are conserved in eukaryotes and archaea, but not in bacteria (TABLE VIII). This is also reflected by the structural similarity of D3. For instance, the overlay of human and archaeal domains D3 yielded low r.m.s.d. value of 1.3 Å, while the same calculation with the bacterial domain resulted in a slightly higher value of 1.6 Å.

Furthermore, my structural analysis showed that the enlarged loop in D3, which corresponds to the partially disordered loop β 21- β 22 in eEFSec, is also present in the archaeal and bacterial SelB. The loop faces the putative tRNA-binding face and its length varies across species. It is the longest in human eEFSec (15 aa), slightly shorter in the archaeal (12 aa), and the shortest in the bacterial SelB (7 aa). Hence, I propose that loop β 21- β 22 in eEFSec and the analogous loop in SelB orthologs are involved in the interaction with the 13 bp long acceptor-T ψ C arm of tRNA^{Sec}.

TABLE VIII. THE PUTATIVE STRUCTURAL ELEMENTS IN HUMAN eEFSec AND ITS ORTHOLOGS THAT MIGHT BE INVOLVED IN BINDING OF Sec-tRNA^{Sec}

Structural element in Sec-tRNA ^{Sec}	eEFSec	Archaeal SeIB	Bacterial SeIB
CCA-Sec	Ile234	Ile196	Val190
	Val240	Val202	Val196
	Gln271	Gln233	Gln227
	Cys289	Ala251	Asn245
5'-end	Arg101	Arg74	Lys70
	Thr102	Ala75	Asn71
	Arg300	Arg262	Arg256
	Gly301	Cys264	Gly257
	Leu302	Ile265	Phe258
	Val303	Leu266	Leu259
T-stem	Arg326	Lys290	-
	Gly327	Tyr291	-
	His336	His300	Gln287
	Val344	Val308	Thr295
	Lys418	Glu345	Lys312
	Leu433	Leu361	-

Residues conserved across kingdoms are bold.

VI.2.3 The model of human eEFSec:GDPCP complex with tRNA^{Sec}

The putative structural elements of human eEFSec responsible for binding to Sec-tRNA^{Sec} are consistent with the recently proposed model of the bacterial ternary complex (42). However, that study did not discuss in detail a potential role of D4 in tRNA^{Sec} binding, which has already been suggested (43, 44). My structural studies of

the GTP- and GDP-bound states of human eEFSec showed that D4 responds to the nucleotide exchange by altering its orientation relative to the putative tRNA-binding side. Such findings led us to speculate that D4 of eEFSec plays an important role in binding and releasing of Sec-tRNA^{Sec} on the ribosome upon GTP hydrolysis. However, the question remains which part of tRNA^{Sec} this domain would bind to? To address this question, I modeled the ternary eEFSec:GDPCP:tRNA^{Sec} complex using the available structures of the GTP-bound state of human eEFSec, tRNA^{Sec} (107), and EF-Tu:GDPNP:Phe-tRNA^{Phe} (63) (Figure 26A).

Modeling analysis reveals that D4 of human eEFSec, specifically loop β 24- α 9 (⁴⁹¹RSLFKKETNI⁵⁰⁰), may be involved in binding of the long variable arm of tRNA^{Sec} (Figure 26B). While the identification of the corresponding loop in the bacterial SelB was not possible due to different structure of D4, the analogous region in the archaeal SelB is formed by a loop (⁴⁰⁷DGLAQ⁴¹¹) and an α -helix (⁴¹²SKVAAE⁴¹⁷). Additionally, the putative eEFSec-tRNA^{Sec} binding interface could include structural elements from D3 and the extreme C-terminus of eEFSec. The regions of D3, which could be in proximity with the variable arm of tRNA^{Sec}, are the loop β 22- α 8 (⁴⁵⁴LEDRN⁴⁵⁸) and helix α 8 (⁴⁵⁹YADSFLPR⁴⁶⁶) (Figure 26B). The corresponding elements in the archaeal SelB are the loop spanning ³⁷⁸EFKP³⁸¹ and α -helix encompassing ³⁸¹PIKDLN³⁸⁶, while in bacteria, the analogous region is loop ³⁴¹PKVRVTK³⁴⁶. Lastly, the C-terminal region of eEFSec, which could also interact with the variable arm of tRNA^{Sec}, includes a short α -helix (⁵⁸⁵VFDT⁵⁸⁸) and side chains of His589, Lys590 and Arg591 (Figure 26B). This

region is disordered in the bacterial SelB, while in archaea it is composed of Arg466, Trp467 and Gly468.

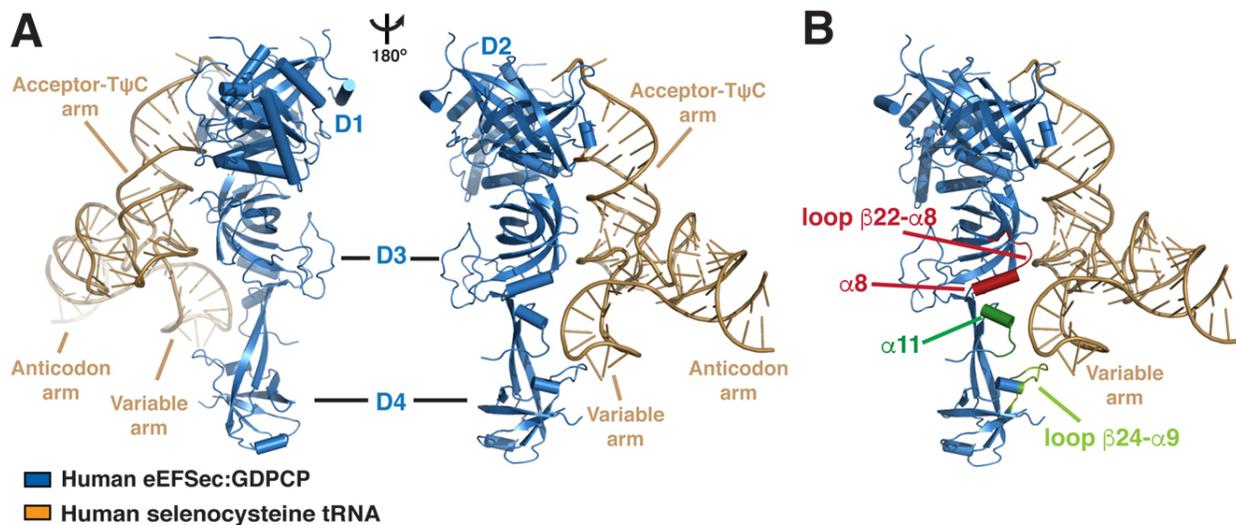


Figure 28. A model of the human eEFSec:GDPCP:tRNA^{Sec} ternary complex. (A) The model of the human ternary complex is shown in two views rotated $\sim 180^\circ$ counterclockwise around the vertical axis. In this model, D4 of eEFSec (blue) interacts with the variable arm of tRNA^{Sec} (gold). **(B)** The detailed analysis suggests that loop $\beta 22-\alpha 8$ and helix $\alpha 8$ (red) in D3, loop $\beta 24-\alpha 9$ (light green) in D4, and the extreme C-terminus (dark green) may be important motifs for binding to the variable arm of tRNA^{Sec}.

VI.3 Discussion

A comparison of the GTP-bound state of eEFSec with the structure of EF-Tu complexed with Phe-tRNA^{Phe} suggested that several structural motifs binding to aa-tRNA might be conserved in the Sec-specific elongation factors. In particular, EF-Tu

residues that interact with A76, the 5'-end, and the T-stem of Phe-tRNA^{Phe} are conserved in eEFSec and SelB. Moreover, human eEFSec shares a higher degree of conservation with the archaeal than with the bacterial ortholog. This is consistent with my observations that eEFSec and archaeal SelB are structurally homologous (see Chapter III). Thus, I propose that at least some aspects of the tRNA-recognition mechanism are preserved between the general and Sec-specific elongation factors.

My analysis of the theoretical model of the human eEFSec:GDPCP:tRNA^{Sec} ternary complex suggests that the appended D4 of eEFSec and archaeal SelB might play a significant role in recognizing Sec-tRNA^{Sec} through binding to its variable arm. Whether the same is true for the structurally divergent D4 of the bacterial SelB has yet to be determined. However, it is plausible that the bacterial D4 alters its conformation upon binding to tRNA^{Sec}, so that it could interact with the variable arm. Besides D4, my analysis suggests that D3 and the C-terminus of eEFSec and the archaeal SelB might interact with the variable arm as well. In case of the bacterial SelB, it has been proposed that D3 is necessary for tRNA^{Sec} binding (42), whereas the C-terminal part was shown to be important for binding to the SECIS element (45).

Taken together, I propose that human eEFSec employs a non-canonical mechanism for binding and release of Sec-tRNA^{Sec} during decoding of the Sec UGA codon (Figure 27). Such a mechanism perhaps arose due to specific and distinct requirements of the Sec UGA decoding machinery that are not present in the canonical system.

According to my model, the GTP-bound state of eEFSec recognizes and binds Sec-tRNA^{Sec}. The ternary eEFSec:GTP:Sec-tRNA^{Sec} complex is then tethered near the ribosome by the SBP2-SECIS complex. The mechanism governing interaction between these two complexes is not clear and further structural studies are needed. The mutational analysis of mouse eEFSec suggested that the region KKRAR, which in my structure is located at the base of D4 and corresponds to the residues 549-553, might be involved in binding of SBP2 (44). When ribosome reaches the Sec UGA codon, eEFSec delivers Sec-tRNA^{Sec} to the A-site. The codon-anticodon recognition as well as the interaction with the ribosomal RNA stimulates the GTPase activity of eEFSec in a mechanism that is perhaps analogous to EF-Tu. GTP hydrolysis induces a slight ratchet of D1 and D2, which results in opening of the GTPase site and constriction of the Sec-binding pocket (Figure 27). These movements might contribute to the decrease of the binding affinity of eEFSec towards the selenocysteiny moiety of Sec-tRNA^{Sec}. Concurrently, the lever-like movement of D4 towards the dorsal side of eEFSec together with the rearrangements within D3 leads to the release of Sec-tRNA^{Sec}, and dissociation of eEFSec:GDP from the ribosome (Figure 27). GTP hydrolysis may or may not be coupled to the dissociation of eEFSec from SBP2-SECIS complex.

I suggest that this particular mechanism may also be applicable to SelB with a significant distinction that the prokaryotic process does not involve SBP2. It could be that D4 of the prokaryotic SelB interacts more closely with SECIS and/or ribosomal proteins rather than eEFSec. However, the unexpected domain motions coupled to GTP hydrolysis observed in eEFSec might be conserved across species.

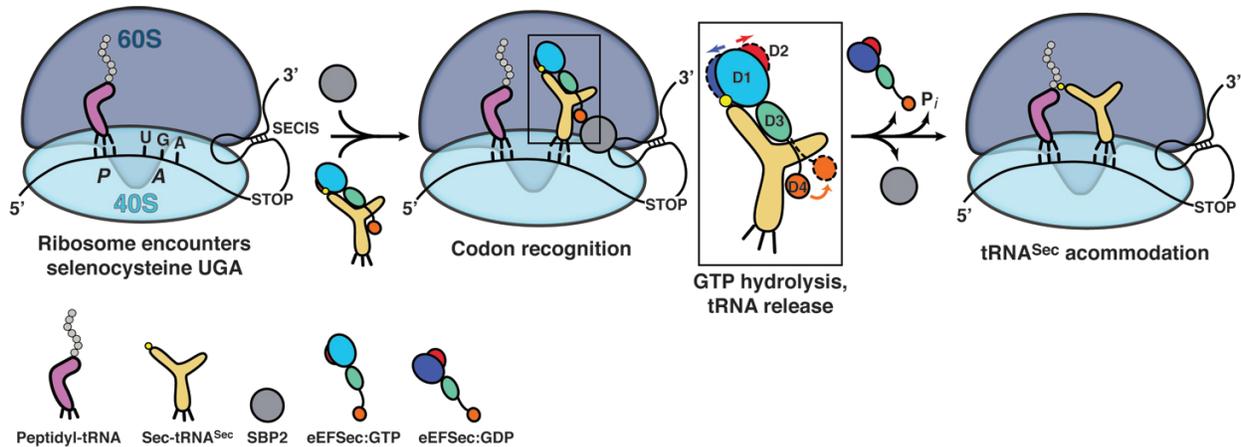


Figure 29. A model of decoding of the Sec UGA codon by human eEFSec. The 80S ribosome pauses when encountering the UGA codon. The ternary eEFSec:GTP:Sec-tRNA^{Sec} complex and SBP2 (grey sphere) bind to the ribosome through interactions with the SECIS element, which is located in the 3'UTR of the selenoprotein mRNA. After codon recognition, eEFSec hydrolyzes GTP and D1 (shades of blue) and D2 (red) move in a ratchet-like motion towards and away from the tRNA-binding (or ventral) side, respectively. Also, the domain D4 (orange and red) swings $\sim 26^\circ$ away from the ventral side. These movements are emphasized in the boxed inset; the GTP- and GDP-bound states of eEFSec are shown in solid and dashed lines, respectively. After eEFSec:GDP dissociation, Sec-tRNA^{Sec} accommodates and formation of the peptide bond occurs. Note: the variable arm of tRNA^{Sec} (pointing to the right), D4 of eEFSec, SBP2, and SECIS are most likely oriented perpendicularly to the plane of the paper and towards the reader. The proposed interactions between SBP2 and the ribosomal protein L30 are not shown for clarity.

VII. CONCLUSIONS

Here, the main findings and conclusions of this doctoral work are summarized:

- I determined the crystal structure of the intact recombinant human eEFSec by experimental phasing. Four domains fold into a chalice-like structure resembling the archaeal more than the bacterial SelB ortholog. The conservation is the most apparent in the N-terminal EF-Tu-like domain, whereas the appended C-terminal domain D4 is structurally divergent.
- As previously proposed for SelB, human eEFSec is more closely related to the translation initiation factor IF2/eIF5B than to the general elongation factors EF-Tu and EF1A.
- The recombinant eEFSec used in structural studies is a functional protein as it promotes Sec incorporation and selenoprotein synthesis *in vitro*.
- Unlike EF-Tu and EF1A, eEFSec binds GTP, non-hydrolyzable GTP analogs and GDP with similar affinities. Hence, eEFSec, just like SelB, does not require the GEF activity for the nucleotide exchange.
- Human eEFSec adopts the conformation of the GTP-bound state when bound to GDPNP and GDPCP. Thus, I conclude that GDPNP and GDPCP are equally good mimics of GTP when bound to human eEFSec.
- Whereas mutations in the GTPase site did not significantly affect the binding of GTP and GDP, they completely abolished the ability of eEFSec to promote the read-through of the Sec UGA codon, perhaps because of their inability to hydrolyze GTP.

- Mutations in the Sec-binding pocket of eEFSec abolished the Sec incorporation activity *in vitro*, most probably by disrupting recognition of the selenocysteinyl moiety in Sec-tRNA^{Sec}.
- The GTP-to-GDP exchange induces an unprecedented conformational change in the appended domain D4, and not the canonical rearrangement of D1. A subtle ratchet of D1 and D2 causes widening of the GTPase site and constriction of the Sec-binding pocket, which could contribute to the release of Sec-tRNA^{Sec}.
- Given the results of the solution-based studies, I conclude that structures of the GTP- and GDP-bound states are the same in solution and in the crystal.
- A comparison of the GTP-bound states of eEFSec and SelB with that of EF-Tu that is in complex with Phe-tRNA^{Phe} suggests that at least some aspects of the canonical tRNA-recognition mechanism are preserved in the Sec-specific elongation factors.
- I propose that eEFSec, and by analogy SelB, employs a non-canonical mechanism for the release of Sec-tRNA^{Sec} during decoding of the Sec UGA codon. I also suggest that D4, D3 and the extreme C-terminus are important structural elements involved in binding of the variable arm of Sec-tRNA^{Sec}.

VIII. FUTURE STUDIES

The work presented in this thesis provided the structural and biochemical characterization of the human Sec-specific elongation factor eEFSec, and answered a number of questions about the nature of conformational changes in eEFSec coupled to the GTP-to-GDP exchange. My results represent a foundation for future studies involving Sec-tRNA^{Sec}, SBP2, SECIS and the ribosome.

In order to fully understand the mechanism by which eEFSec selects Sec-tRNA^{Sec} among other aa-tRNAs, the crystal structure of the ternary complex between the GTP-bound state of eEFSec and an intact Sec-tRNA^{Sec} together with mutational analysis of both eEFSec and tRNA^{Sec} are warranted. Additionally, it would be interesting to establish whether mutations within the Sec-binding pocket lead to misincorporation of serine, or other near cognate amino acids, in place of selenocysteine. To accomplish this task, *in vitro* binding and activity assays involving Sec-tRNA^{Sec}, Ser-tRNA^{Sec}, and WT and mutant eEFSec constructs are needed.

Lastly, structural and biochemical studies utilizing a complete human and/or bacterial decoding particle will explain how eEFSec and SelB interact with SECIS, how rRNA stimulates the GTPase activity of eEFSec and SelB, how an unusual tRNA^{Sec} is accommodated in the decoding and peptidyl-transferase centers, and how SBP2 participates in the decoding process of Sec codon in eukaryotes.

CITED LITERATURE

1. Pinsent J. The need for selenite and molybdate in the formation of formic dehydrogenase by members of the coli-aerogenes group of bacteria. *Biochem J.* 1954;57(1):10-6. PubMed PMID: 13159942; PMCID: 1269698.
2. Schwarz K, Foltz CM. Selenium as an integral part of factor 3 against dietary necrotic liver degeneration. 1951. *Nutrition.* 1999;15(3):255. Epub 1999/07/17. PubMed PMID: 10408880.
3. Dudley HC, United States. Selenium as a potential industrial hazard. Washington,; U. S. Govt. print. off.; 1938. 1 p.l., 12 p. p.
4. Robertson DS. Selenium--a possible teratogen? *Lancet.* 1970;1(7645):518-9. Epub 1970/03/07. PubMed PMID: 4190197.
5. Gunn SA, Gould TC, Anderson WA. Incorporation of selenium into spermatogenic pathway in mice. *Proc Soc Exp Biol Med.* 1967;124(4):1260-3. Epub 1967/04/01. PubMed PMID: 6024846.
6. Observations on effect of sodium selenite in prevention of Keshan disease. *Chin Med J (Engl).* 1979;92(7):471-6. Epub 1979/07/01. PubMed PMID: 114371.
7. Vanderpas JB, Contempre B, Duale NL, Goossens W, Bebe N, Thorpe R, Ntambue K, Dumont J, Thilly CH, Diplock AT. Iodine and selenium deficiency associated with cretinism in northern Zaire. *Am J Clin Nutr.* 1990;52(6):1087-93. Epub 1990/12/01. PubMed PMID: 2239787.
8. Cone JE, Del Rio RM, Davis JN, Stadtman TC. Chemical characterization of the selenoprotein component of clostridial glycine reductase: identification of selenocysteine as the organoselenium moiety. *Proc Natl Acad Sci U S A.* 1976;73(8):2659-63. Epub 1976/08/01. PubMed PMID: 1066676; PMCID: 430707.
9. Kryukov GV, Castellano S, Novoselov SV, Lobanov AV, Zehtab O, Guigo R, Gladyshev VN. Characterization of mammalian selenoproteomes. *Science.* 2003;300(5624):1439-43. Epub 2003/05/31. doi: 10.1126/science.1083516 300/5624/1439 [pii]. PubMed PMID: 12775843.

10. Papp LV, Lu J, Holmgren A, Khanna KK. From selenium to selenoproteins: synthesis, identity, and their role in human health. *Antioxid Redox Signal*. 2007;9(7):775-806. Epub 2007/05/19. doi: 10.1089/ars.2007.1528. PubMed PMID: 17508906.
11. Lu J, Holmgren A. Selenoproteins. *J Biol Chem*. 2009;284(2):723-7. Epub 2008/09/02. doi: R800045200 [pii] 10.1074/jbc.R800045200. PubMed PMID: 18757362.
12. Castellano S, Lobanov AV, Chapple C, Novoselov SV, Albrecht M, Hua D, Lescure A, Lengauer T, Krol A, Gladyshev VN, Guigo R. Diversity and functional plasticity of eukaryotic selenoproteins: identification and characterization of the SelJ family. *Proc Natl Acad Sci U S A*. 2005;102(45):16188-93. Epub 2005/11/02. doi: 0505146102 [pii] 10.1073/pnas.0505146102. PubMed PMID: 16260744; PMCID: 1283428.
13. Hoffmann PR, Berry MJ. Selenoprotein synthesis: a unique translational mechanism used by a diverse family of proteins. *Thyroid*. 2005;15(8):769-75. Epub 2005/09/01. doi: 10.1089/thy.2005.15.769. PubMed PMID: 16131320.
14. Johansson L, Gafvelin G, Arner ES. Selenocysteine in proteins-properties and biotechnological use. *Biochim Biophys Acta*. 2005;1726(1):1-13. Epub 2005/06/22. doi: S0304-4165(05)00152-2 [pii] 10.1016/j.bbagen.2005.05.010. PubMed PMID: 15967579.
15. Rayman MP. Selenoproteins and human health: insights from epidemiological data. *Biochim Biophys Acta*. 2009;1790(11):1533-40. Epub 2009/03/31. doi: S0304-4165(09)00068-3 [pii]10.1016/j.bbagen.2009.03.014. PubMed PMID: 19327385.
16. Schmidt RL, Simonović M. Synthesis and decoding of selenocysteine and human health. *Croatian medical journal*. 2012;53(6):535-50. PubMed PMID: 23275319; PMCID: 3541580.
17. Labunskyy VM, Hatfield DL, Gladyshev VN. Selenoproteins: molecular pathways and physiological roles. *Physiol Rev*. 2014;94(3):739-77. doi: 10.1152/physrev.00039.2013. PubMed PMID: 24987004; PMCID: 4101630.
18. Chambers I, Frampton J, Goldfarb P, Affara N, McBain W, Harrison PR. The structure of the mouse glutathione peroxidase gene: the selenocysteine in the active site is encoded by the 'termination' codon, TGA. *EMBO J*. 1986;5(6):1221-7. Epub 1986/06/01. PubMed PMID: 3015592; PMCID: 1166931.

19. Axley MJ, Bock A, Stadtman TC. Catalytic properties of an *Escherichia coli* formate dehydrogenase mutant in which sulfur replaces selenium. *Proc Natl Acad Sci U S A*. 1991;88(19):8450-4. Epub 1991/10/01. PubMed PMID: 1924303; PMCID: 52526.

20. Rocher C, Lalanne JL, Chaudiere J. Purification and properties of a recombinant sulfur analog of murine selenium-glutathione peroxidase. *Eur J Biochem*. 1992;205(3):955-60. Epub 1992/05/01. PubMed PMID: 1577013.

21. Maiorino M, Aumann KD, Brigelius-Flohe R, Doria D, van den Heuvel J, McCarthy J, Roveri A, Ursini F, Flohe L. Probing the presumed catalytic triad of selenium-containing peroxidases by mutational analysis of phospholipid hydroperoxide glutathione peroxidase (PHGPx). *Biol Chem Hoppe Seyler*. 1995;376(11):651-60. Epub 1995/11/01. PubMed PMID: 8962674.

22. Dumitrescu AM, Refetoff S. Inherited defects of thyroid hormone metabolism. *Ann Endocrinol (Paris)*. 2011;72(2):95-8. doi: 10.1016/j.ando.2011.03.011. PubMed PMID: 21511232; PMCID: 3094475.

23. Agamy O, Ben Zeev B, Lev D, Marcus B, Fine D, Su D, Narkis G, Ofir R, Hoffmann C, Leshinsky-Silver E, Flusser H, Sivan S, Söll D, Lerman-Sagie T, Birk OS. Mutations disrupting selenocysteine formation cause progressive cerebello-cerebral atrophy. *Am J Hum Genet*. 2010;87(4):538-44. doi: 10.1016/j.ajhg.2010.09.007. PubMed PMID: 20920667; PMCID: 2948803.

24. Anttonen AK, Hilander T, Linnankivi T, Isohanni P, French RL, Liu Y, Simonović M, Söll D, Somer M, Muth-Pawlak D, Corthals GL, Laari A, Ylikallio E, Lahde M, Valanne L, Lonqvist T, Pihko H, Paetau A, Lehesjoki AE, Suomalainen A, Tynismaa H. Selenoprotein biosynthesis defect causes progressive encephalopathy with elevated lactate. *Neurology*. 2015;85(4):306-15. doi: 10.1212/WNL.0000000000001787. PubMed PMID: 26115735; PMCID: 4520820.

25. Shetty SP, Copeland PR. Selenocysteine incorporation: A trump card in the game of mRNA decay. *Biochimie*. 2015;114:97-101. doi: 10.1016/j.biochi.2015.01.007. PubMed PMID: 25622574; PMCID: 4458392.

26. Mukai T, Englert M, Tripp HJ, Miller C, Ivanova NN, Rubin EM, Kyripides NC, Söll D. Facile Recoding of Selenocysteine in Nature. *Angew Chem Int Ed Engl*. 2016;55(17):5337-41. doi: 10.1002/anie.201511657. PubMed PMID: 26991476; PMCID: 4833512.

27. Leinfelder W, Zehelein E, Mandrand-Berthelot MA, Böck A. Gene for a novel tRNA species that accepts L-serine and cotranslationally inserts selenocysteine. *Nature*. 1988;331(6158):723-5. Epub 1988/02/25. doi: 10.1038/331723a0. PubMed PMID: 2963963.

28. Lee BJ, Worland PJ, Davis JN, Stadtman TC, Hatfield DL. Identification of a selenocysteyl-tRNA(Ser) in mammalian cells that recognizes the nonsense codon, UGA. *J Biol Chem*. 1989;264(17):9724-7. Epub 1989/06/15. PubMed PMID: 2498338.

29. Leinfelder W, Stadtman TC, Böck A. Occurrence *in vivo* of selenocysteyl-tRNA(SERUCA) in *Escherichia coli*. Effect of sel mutations. *J Biol Chem*. 1989;264(17):9720-3. Epub 1989/06/15. PubMed PMID: 2524495.

30. Forchhammer K, Böck A. Selenocysteine synthase from *Escherichia coli*. Analysis of the reaction sequence. *J Biol Chem*. 1991;266(10):6324-8. Epub 1991/04/05. PubMed PMID: 2007585.

31. Carlson BA, Xu XM, Kryukov GV, Rao M, Berry MJ, Gladyshev VN, Hatfield DL. Identification and characterization of phosphoseryl-tRNA[Ser]Sec kinase. *Proc Natl Acad Sci U S A*. 2004;101(35):12848-53. Epub 2004/08/20. doi: 10.1073/pnas.04026361010402636101 [pii]. PubMed PMID: 15317934; PMCID: 516484.

32. Forchhammer K, Leinfelder W, Böck A. Identification of a novel translation factor necessary for the incorporation of selenocysteine into protein. *Nature*. 1989;342(6248):453-6. Epub 1989/11/23. doi: 10.1038/342453a0. PubMed PMID: 2531290.

33. Rother M, Wilting R, Commans S, Böck A. Identification and characterisation of the selenocysteine-specific translation factor SelB from the archaeon *Methanococcus jannaschii*. *J Mol Biol*. 2000;299(2):351-8. Epub 2000/06/22. doi: 10.1006/jmbi.2000.3756S0022-2836(00)93756-2 [pii]. PubMed PMID: 10860743.

34. Tujebajeva RM, Copeland PR, Xu XM, Carlson BA, Harney JW, Driscoll DM, Hatfield DL, Berry MJ. Decoding apparatus for eukaryotic selenocysteine insertion. *EMBO Rep*. 2000;1(2):158-63. Epub 2001/03/27. doi: 10.1038/sj.embor.embor604 embor604 [pii]. PubMed PMID: 11265756; PMCID: 1084265.

35. Fagegaltier D, Hubert N, Yamada K, Mizutani T, Carbon P, Krol A. Characterization of mSelB, a novel mammalian elongation factor for selenoprotein

translation. *EMBO J.* 2000;19(17):4796-805. Epub 2000/09/06. doi: 10.1093/emboj/19.17.4796. PubMed PMID: 10970870; PMCID: 302067.

36. Baron C, Sturchler C, Wu XQ, Gross HJ, Krol A, Böck A. Eukaryotic selenocysteine inserting tRNA species support selenoprotein synthesis in *Escherichia coli*. *Nucleic Acids Res.* 1994;22(12):2228-33. Epub 1994/06/25. PubMed PMID: 8036149; PMCID: 523678.

37. Hatfield DL, Carlson BA, Xu XM, Mix H, Gladyshev VN. Selenocysteine incorporation machinery and the role of selenoproteins in development and health. *Prog Nucleic Acid Res Mol Biol.* 2006;81:97-142. Epub 2006/08/08. doi: S0079-6603(06)81003-2 [pii]10.1016/S0079-6603(06)81003-2. PubMed PMID: 16891170.

38. Bösl MR, Takaku K, Oshima M, Nishimura S, Taketo MM. Early embryonic lethality caused by targeted disruption of the mouse selenocysteine tRNA gene (*Trsp*). *Proc Natl Acad Sci U S A.* 1997;94(11):5531-4. Epub 1997/05/27. PubMed PMID: 9159106; PMCID: 20812.

39. Holcik M, Sonenberg N. Translational control in stress and apoptosis. *Nat Rev Mol Cell Biol.* 2005;6(4):318-27. doi: 10.1038/nrm1618. PubMed PMID: 15803138.

40. Carlson BA, Novoselov SV, Kumaraswamy E, Lee BJ, Anver MR, Gladyshev VN, Hatfield DL. Specific excision of the selenocysteine tRNA[Ser]^{Sec} (*Trsp*) gene in mouse liver demonstrates an essential role of selenoproteins in liver function. *J Biol Chem.* 2004;279(9):8011-7. Epub 2003/12/09. doi: 10.1074/jbc.M310470200M310470200 [pii]. PubMed PMID: 14660662.

41. Leibundgut M, Frick C, Thanbichler M, Böck A, Ban N. Selenocysteine tRNA-specific elongation factor SelB is a structural chimaera of elongation and initiation factors. *EMBO J.* 2005;24(1):11-22. Epub 2004/12/24. doi: 7600505 [pii]10.1038/sj.emboj.7600505. PubMed PMID: 15616587; PMCID: 544917.

42. Itoh Y, Sekine S, Yokoyama S. Crystal structure of the full-length bacterial selenocysteine-specific elongation factor SelB. *Nucleic Acids Res.* 2015;43(18):9028-38. doi: 10.1093/nar/gkv833. PubMed PMID: 26304550; PMCID: 4605307.

43. Zavacki AM, Mansell JB, Chung M, Klimovitsky B, Harney JW, Berry MJ. Coupled tRNA(Sec)-dependent assembly of the selenocysteine decoding apparatus. *Mol Cell.* 2003;11(3):773-81. Epub 2003/04/02. doi: S1097276503000649 [pii]. PubMed PMID: 12667458.

44. Gonzalez-Flores JN, Gupta N, DeMong LW, Copeland PR. The selenocysteine-specific elongation factor contains a novel and multi-functional domain. *J Biol Chem.* 2012;287(46):38936-45. doi: 10.1074/jbc.M112.415463. PubMed PMID: 22992746; PMCID: 3493934.
45. Yoshizawa S, Rasubala L, Ose T, Kohda D, Fourmy D, Maenaka K. Structural basis for mRNA recognition by elongation factor SelB. *Nat Struct Mol Biol.* 2005;12(2):198-203. Epub 2005/01/25. doi: nsmb890 [pii] 10.1038/nsmb890. PubMed PMID: 15665870.
46. Selmer M, Su XD. Crystal structure of an mRNA-binding fragment of *Moorella thermoacetica* elongation factor SelB. *EMBO J.* 2002;21(15):4145-53. Epub 2002/07/30. PubMed PMID: 12145214; PMCID: 126154.
47. Paleskava A, Konevega AL, Rodnina MV. Thermodynamics of the GTP-GDP-operated conformational switch of selenocysteine-specific translation factor SelB. *J Biol Chem.* 2012;287(33):27906-12. doi: 10.1074/jbc.M112.366120. PubMed PMID: 22740700; PMCID: 3431659.
48. Huttenhofer A, Böck A. Selenocysteine inserting RNA elements modulate GTP hydrolysis of elongation factor SelB. *Biochemistry.* 1998;37(3):885-90. Epub 1998/02/10. doi: 10.1021/bi972298kbi972298k [pii]. PubMed PMID: 9454578.
49. Allmang C, Krol A. Selenoprotein synthesis: UGA does not end the story. *Biochimie.* 2006;88(11):1561-71. Epub 2006/06/02. doi: S0300-9084(06)00065-4 [pii] 10.1016/j.biochi.2006.04.015. PubMed PMID: 16737768.
50. Chavatte L, Brown BA, Driscoll DM. Ribosomal protein L30 is a component of the UGA-selenocysteine recoding machinery in eukaryotes. *Nat Struct Mol Biol.* 2005;12(5):408-16. Epub 2005/04/12. doi: nsmb922 [pii]10.1038/nsmb922. PubMed PMID: 15821744.
51. Xu XM, Mix H, Carlson BA, Grabowski PJ, Gladyshev VN, Berry MJ, Hatfield DL. Evidence for direct roles of two additional factors, SECp43 and soluble liver antigen, in the selenoprotein synthesis machinery. *J Biol Chem.* 2005;280(50):41568-75. Epub 2005/10/19. doi: M506696200 [pii]10.1074/jbc.M506696200. PubMed PMID: 16230358.
52. Small-Howard AL, Berry MJ. Unique features of selenocysteine incorporation function within the context of general eukaryotic translational processes. *Biochem Soc Trans.* 2005;33(Pt 6):1493-7. Epub 2005/10/26. doi: BST20051493 [pii]10.1042/BST20051493. PubMed PMID: 16246153.

53. Liu Z, Reches M, Groisman I, Engelberg-Kulka H. The nature of the minimal 'selenocysteine insertion sequence' (SECIS) in *Escherichia coli*. *Nucleic Acids Res.* 1998;26(4):896-902. Epub 1998/03/21. doi: gkb201 [pii]. PubMed PMID: 9461445; PMCID: 147357.
54. Berry MJ, Banu L, Harney JW, Larsen PR. Functional characterization of the eukaryotic SECIS elements which direct selenocysteine insertion at UGA codons. *EMBO J.* 1993;12(8):3315-22. Epub 1993/08/01. PubMed PMID: 8344267; PMCID: 413599.
55. Allamand V, Richard P, Lescure A, Ledeuil C, Desjardin D, Petit N, Gartioux C, Ferreira A, Krol A, Pellegrini N, Urtizberea JA, Guicheney P. A single homozygous point mutation in a 3'untranslated region motif of selenoprotein N mRNA causes SEPNI-related myopathy. *EMBO Rep.* 2006;7(4):450-4. Epub 2006/02/25. doi: 7400648 [pii]10.1038/sj.embor.7400648. PubMed PMID: 16498447; PMCID: 1456920.
56. Olieric V, Wolff P, Takeuchi A, Bec G, Birck C, Vitorino M, Kieffer B, Benjaminov A, Cavigliolo G, Theil E, Allmang C, Krol A, Dumas P. SECIS-binding protein 2, a key player in selenoprotein synthesis, is an intrinsically disordered protein. *Biochimie.* 2009;91(8):1003-9. Epub 2009/05/27. doi: S0300-9084(09)00146-1 [pii] 10.1016/j.biochi.2009.05.004. PubMed PMID: 19467292.
57. Copeland PR, Stepanik VA, Driscoll DM. Insight into mammalian selenocysteine insertion: domain structure and ribosome binding properties of Sec insertion sequence binding protein 2. *Mol Cell Biol.* 2001;21(5):1491-8. Epub 2001/03/10. doi: 10.1128/MCB.21.5.1491-1498.2001. PubMed PMID: 11238886; PMCID: 86695.
58. Kinzy SA, Caban K, Copeland PR. Characterization of the SECIS binding protein 2 complex required for the co-translational insertion of selenocysteine in mammals. *Nucleic Acids Res.* 2005;33(16):5172-80. Epub 2005/09/13. doi: 33/16/5172 [pii]10.1093/nar/gki826. PubMed PMID: 16155186; PMCID: 1214547.
59. Bubenik JL, Driscoll DM. Altered RNA binding activity underlies abnormal thyroid hormone metabolism linked to a mutation in selenocysteine insertion sequence-binding protein 2. *J Biol Chem.* 2007;282(48):34653-62. Epub 2007/09/29. doi: M707059200 [pii]10.1074/jbc.M707059200. PubMed PMID: 17901054.
60. Dumitrescu AM, Liao XH, Abdullah MS, Lado-Abeal J, Majed FA, Moeller LC, Boran G, Schomburg L, Weiss RE, Refetoff S. Mutations in SECISBP2 result in abnormal thyroid hormone metabolism. *Nat Genet.* 2005;37(11):1247-52. Epub 2005/10/18. doi: ng1654 [pii]10.1038/ng1654. PubMed PMID: 16228000.

61. Kjeldgaard M, Nyborg J. Refined structure of elongation factor EF-Tu from *Escherichia coli*. *J Mol Biol.* 1992;223(3):721-42. PubMed PMID: 1542116.
62. Kjeldgaard M, Nissen P, Thirup S, Nyborg J. The crystal structure of elongation factor EF-Tu from *Thermus aquaticus* in the GTP conformation. *Structure.* 1993;1(1):35-50. PubMed PMID: 8069622.
63. Nissen P, Kjeldgaard M, Thirup S, Polekhina G, Reshetnikova L, Clark BF, Nyborg J. Crystal structure of the ternary complex of Phe-tRNAPhe, EF-Tu, and a GTP analog. *Science.* 1995;270(5241):1464-72. Epub 1995/12/01. PubMed PMID: 7491491.
64. Polekhina G, Thirup S, Kjeldgaard M, Nissen P, Lippmann C, Nyborg J. Helix unwinding in the effector region of elongation factor EF-Tu-GDP. *Structure.* 1996;4(10):1141-51. PubMed PMID: 8939739.
65. Schmeing TM, Voorhees RM, Kelley AC, Gao YG, Murphy FVt, Weir JR, Ramakrishnan V. The crystal structure of the ribosome bound to EF-Tu and aminoacyl-tRNA. *Science.* 2009;326(5953):688-94. Epub 2009/10/17. doi: 1179700 [pii]10.1126/science.1179700. PubMed PMID: 19833920.
66. Voorhees RM, Schmeing TM, Kelley AC, Ramakrishnan V. The mechanism for activation of GTP hydrolysis on the ribosome. *Science.* 2010;330(6005):835-8. doi: 10.1126/science.1194460. PubMed PMID: 21051640; PMCID: 3763471.
67. Voorhees RM, Ramakrishnan V. Structural basis of the translational elongation cycle. *Annu Rev Biochem.* 2013;82:203-36. doi: 10.1146/annurev-biochem-113009-092313. PubMed PMID: 23746255.
68. Agrawal RK, Penczek P, Grassucci RA, Frank J. Visualization of elongation factor G on the *Escherichia coli* 70S ribosome: the mechanism of translocation. *Proc Natl Acad Sci U S A.* 1998;95(11):6134-8. PubMed PMID: 9600930; PMCID: 27598.
69. Wang Y, Jiang Y, Meyering-Voss M, Sprinzl M, Sigler PB. Crystal structure of the EF-Tu.EF-Ts complex from *Thermus thermophilus*. *Nat Struct Biol.* 1997;4(8):650-6. Epub 1997/08/01. PubMed PMID: 9253415.
70. Thirup SS, Van LB, Nielsen TK, Knudsen CR. Structural outline of the detailed mechanism for elongation factor Ts-mediated guanine nucleotide exchange on

elongation factor Tu. *J Struct Biol.* 2015;191(1):10-21. doi: 10.1016/j.jsb.2015.06.011. PubMed PMID: 26073967.

71. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. *Method Enzymol.* 1997;276:307-26. doi: Doi 10.1016/S0076-6879(97)76066-X. PubMed PMID: WOS:A1997BH42P00020.

72. Sheldrick GM. Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallogr D.* 2010;66:479-85. doi: 10.1107/S0907444909038360. PubMed PMID: WOS:000275941300017.

73. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D.* 2004;60:2126-32. doi: 10.1107/S0907444904019158. PubMed PMID: WOS:000225360500002.

74. Adams PD, Grosse-Kunstleve RW, Hung LW, Ioerger TR, McCoy AJ, Moriarty NW, Read RJ, Sacchettini JC, Sauter NK, Terwilliger TC. PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr D.* 2002;58:1948-54. doi: 10.1107/S0907444902016657. PubMed PMID: WOS:000178745000009.

75. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. *Journal of applied crystallography.* 2007;40:658-74. doi: 10.1107/S0021889807021206. PubMed PMID: WOS:000248077500003.

76. Girdlestone C, Hayward S. The DynDom3D Webserver for the Analysis of Domain Movements in Multimeric Proteins. *Journal of Computational Biology.* 2016;23(1):21-6. doi: 10.1089/cmb.2015.0143. PubMed PMID: WOS:000367401500003.

77. Corpet F. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* 1988;16(22):10881-90.

78. Robert X, Gouet P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research.* 2014;42(W1):W320-W4. doi: 10.1093/nar/gku316. PubMed PMID: WOS:000339715000053.

79. Fischetti R, Stepanov S, Rosenbaum G, Barrea R, Black E, Gore D, Heurich R, Kondrashkina E, Kropf AJ, Wang S, Zhang K, Irving TC, Bunker GB. The BioCAT undulator beamline 18ID: a facility for biological non-crystalline diffraction and X-ray absorption spectroscopy at the Advanced Photon Source. *J Synchrotron Radiat.* 2004;11(Pt 5):399-405. doi: 10.1107/S0909049504016760. PubMed PMID: 15310956.
80. Petoukhov MV, Franke D, Shkumatov AV, Tria G, Kikhney AG, Gajda M, Gorba C, Mertens HD, Konarev PV, Svergun DI. New developments in the program package for small-angle scattering data analysis. *Journal of applied crystallography.* 2012;45(Pt 2):342-50. doi: 10.1107/S0021889812007662. PubMed PMID: 25484842; PMCID: 4233345.
81. Petoukhov MV, Svergun DI. Analysis of X-ray and neutron scattering from biomacromolecular solutions. *Curr Opin Struct Biol.* 2007;17(5):562-71. doi: 10.1016/j.sbi.2007.06.009. PubMed PMID: 17714935.
82. Konarev PV, Volkov VV, Sokolova AV, Koch MHJ, Svergun DI. PRIMUS - a Windows-PC based system for small-angle scattering data analysis. *J Appl Cryst.* 2003;36:1277-82.
83. Svergun DI. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J Appl Cryst.* 1992;25:495-503.
84. Franke D, Svergun DI. DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J Appl Cryst.* 2009;42:342-6.
85. Kozin M, Svergun DI. Automated matching of high- and low-resolution structural models. *J Appl Cryst.* 2001;34:33-41.
86. Hirosawa-Takamori M, Chung HR, Jackle H. Conserved selenoprotein synthesis is not critical for oxidative stress defence and the lifespan of *Drosophila*. *EMBO Rep.* 2004;5(3):317-22. Epub 2004/02/24. doi: 10.1038/sj.embor.7400097 7400097 [pii]. PubMed PMID: 14978508; PMCID: 1299008.
87. Vitagliano L, Masullo M, Sica F, Zagari A, Bocchini V. The crystal structure of *Sulfolobus solfataricus* elongation factor 1alpha in complex with GDP reveals novel features in nucleotide binding and exchange. *EMBO J.* 2001;20(19):5305-11. doi: 10.1093/emboj/20.19.5305. PubMed PMID: 11574461; PMCID: 125647.

88. Andersen GR, Thirup S, Spremulli LL, Nyborg J. High resolution crystal structure of bovine mitochondrial EF-Tu in complex with GDP. *J Mol Biol.* 2000;297(2):421-36. doi: 10.1006/jmbi.2000.3564. PubMed PMID: 10715211.
89. Pedersen L, Andersen GR, Knudsen CR, Kinzy TG, Nyborg J. Crystallization of the yeast elongation factor complex eEF1A-eEF1B alpha. *Acta Crystallogr D Biol Crystallogr.* 2001;57(Pt 1):159-61. Epub 2001/01/03. doi: S0907444900015559 [pii]. PubMed PMID: 11134944.
90. Crepin T, Shalak VF, Yaremchuk AD, Vlasenko DO, McCarthy A, Negrutskii BS, Tukalo MA, El'skaya AV. Mammalian translation elongation factor eEF1A2: X-ray structure and new features of GDP/GTP exchange mechanism in higher eukaryotes. *Nucleic Acids Res.* 2014;42(20):12939-48. doi: 10.1093/nar/gku974. PubMed PMID: 25326326; PMCID: 4227793.
91. Förster C, Ott G, Forchhammer K, Sprinzl M. Interaction of a selenocysteine-incorporating tRNA with elongation factor Tu from *E.coli*. *Nucleic Acids Res.* 1990;18(3):487-91. Epub 1990/02/11. PubMed PMID: 2408012; PMCID: 333452.
92. Baron C, Böck A. 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.* 1991;266(30):20375-9. Epub 1991/10/25. PubMed PMID: 1939093.
93. Rudinger J, Hillenbrandt R, Sprinzl M, Giegé R. Antideterminants present in minihelix(Sec) hinder its recognition by prokaryotic elongation factor Tu. *EMBO J.* 1996;15(3):650-7. Epub 1996/02/01. PubMed PMID: 8599948; PMCID: 449983.
94. Paleskava A, Konevega AL, Rodnina MV. Thermodynamic and kinetic framework of selenocysteyl-tRNA^{Sec} recognition by elongation factor SelB. *J Biol Chem.* 2010;285(5):3014-20. Epub 2009/11/27. doi: M109.081380 [pii]10.1074/jbc.M109.081380. PubMed PMID: 19940162; PMCID: 2823455.
95. Noble CG, Song H. Structural studies of elongation and release factors. *Cell Mol Life Sci.* 2008;65(9):1335-46. doi: 10.1007/s00018-008-7495-6. PubMed PMID: 18213444.
96. Keeling PJ, Fast NM, McFadden GI. Evolutionary relationship between translation initiation factor eIF-2gamma and selenocysteine-specific elongation factor

SELB: change of function in translation factors. *J Mol Evol.* 1998;47(6):649-55. Epub 1998/12/16. PubMed PMID: 9847405.

97. Roll-Mecak A, Cao C, Dever TE, Burley SK. X-Ray structures of the universal translation initiation factor IF2/eIF5B: conformational changes on GDP and GTP binding. *Cell.* 2000;103(5):781-92. PubMed PMID: 11114334.

98. Berry MJ, Tujebajeva RM, Copeland PR, Xu XM, Carlson BA, Martin GW, 3rd, Low SC, Mansell JB, Grundner-Culemann E, Harney JW, Driscoll DM, Hatfield DL. Selenocysteine incorporation directed from the 3'UTR: characterization of eukaryotic EFsec and mechanistic implications. *Biofactors.* 2001;14(1-4):17-24. Epub 2001/09/25. PubMed PMID: 11568436.

99. Copeland PR, Driscoll DM. Purification, redox sensitivity, and RNA binding properties of SECIS-binding protein 2, a protein involved in selenoprotein biosynthesis. *J Biol Chem.* 1999;274(36):25447-54. Epub 1999/08/28. PubMed PMID: 10464275.

100. Copeland PR, Fletcher JE, Carlson BA, Hatfield DL, Driscoll DM. A novel RNA binding protein, SBP2, is required for the translation of mammalian selenoprotein mRNAs. *EMBO J.* 2000;19(2):306-14. Epub 2000/01/19. doi: 10.1093/emboj/19.2.306. PubMed PMID: 10637234; PMCID: 305564.

101. Daviter T, Wieden HJ, Rodnina MV. Essential role of histidine 84 in elongation factor Tu for the chemical step of GTP hydrolysis on the ribosome. *J Mol Biol.* 2003;332(3):689-99. PubMed PMID: 12963376.

102. Aleksandrov A, Field M. Mechanism of activation of elongation factor Tu by ribosome: catalytic histidine activates GTP by protonation. *RNA.* 2013;19(9):1218-25. doi: 10.1261/rna.040097.113. PubMed PMID: 23864225; PMCID: 3753929.

103. Thanbichler M, Böck A, Goody RS. Kinetics of the interaction of translation factor SelB from *Escherichia coli* with guanosine nucleotides and selenocysteine insertion sequence RNA. *J Biol Chem.* 2000;275(27):20458-66. Epub 2000/04/27. doi: 10.1074/jbc.M002496200M002496200 [pii]. PubMed PMID: 10781605.

104. Gromadski KB, Wieden HJ, Rodnina MV. Kinetic mechanism of elongation factor Ts-catalyzed nucleotide exchange in elongation factor Tu. *Biochemistry.* 2002;41(1):162-9. PubMed PMID: 11772013.

105. Hilgenfeld R, Böck A, Wilting R. Structural model for the selenocysteine-specific elongation factor SelB. *Biochimie*. 1996;78(11-12):971-8. Epub 1996/01/01. doi: S0300-9084(97)86719-3 [pii]. PubMed PMID: 9150874.

106. Tourigny DS, Fernandez IS, Kelley AC, Ramakrishnan V. Elongation factor G bound to the ribosome in an intermediate state of translocation. *Science*. 2013;340(6140):1235490. doi: 10.1126/science.1235490. PubMed PMID: 23812720; PMCID: 3836249.

107. Palioura S, Sherrer RL, Steitz TA, Söll D, Simonović M. The human SepSecS-tRNA^{Sec} complex reveals the mechanism of selenocysteine formation. *Science*. 2009;325(5938):321-5. Epub 2009/07/18. doi: 325/5938/321 [pii]10.1126/science.1173755. PubMed PMID: 19608919; PMCID: 2857584.

APPENDICES

APPENDIX A

Permissions of use for Springer. Dobosz-Bartoszek M, Simonović M. Structure and Mechanism of Selenocysteine Synthases. ***Selenium – Its Molecular Biology and Role in Human Health, 4th Edition.*** Editors: Hatfield DL, Schweizer U, Tsuji P, Gladyshev VN. Springer (2016).



01_03

PERMISSION LETTER

June 27, 2016

Springer reference***Selenium****Structure and Mechanism of Selenocysteine Synthases*

ISBN 978-3-319-41281-8

Your project**Requestor:** Malgorzata Dobosz-Bartoszek
malgosia.dobosz@gmail.com>**University:** University of Illinois at Chicago**Purpose:** Dissertation/Thesis

With reference to your request to reuse material in which Springer controls the copyright, our permission is granted free of charge under the following conditions:

Springer material

- represents original material which does not carry references to other sources (if material in question refers with a credit to another source, authorization from that source is required as well);
- requires full credit (Springer book/journal title, chapter/article title, volume, year of publication, page, name(s) of author(s), original copyright notice) to the publication in which the material was originally published by adding: "With permission of Springer";
- may not be altered in any manner. Abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of the author;
- Springer does not supply original artwork or content.

This permission

- is non-exclusive;
- is valid for one-time use only for the purpose of defending your thesis and with a maximum of 100 extra copies in paper. If the thesis is going to be published, permission needs to be reobtained.
- includes use in an electronic form, provided it is an author-created version of the thesis on his/her own website and his/her university's repository, including UMI (according to the definition on the Sherpa website: <http://www.sherpa.ac.uk/romeo/>);
- is subject to courtesy information to the co-author or corresponding author;
- is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without Springer's written permission;
- is only valid if no personal rights, trademarks, or competitive products are infringed.

This license is valid only when the conditions noted above are met.

Permission free of charge does not prejudice any rights we might have to charge for reproduction of our copyrighted material in the future.

Rights and Permissions
Springer
Tiergartenstr. 17
69121 Heidelberg
Germany

APPENDIX B

The content of this appendix will be published as Chapter 9, entitled “Structure and Mechanism of Selenocysteine Synthases”, in ***Selenium – Its Molecular Biology and Role in Human Health, 4th Edition***, (Editors: Hatfield DL, Schweizer U, Tsuji P, Gladyshev VN), Springer (2016).

Structure and Mechanism of Selenocysteine Synthases

Malgorzata Dobosz-Bartoszek and Miljan Simonović

Abstract. Selenocysteine synthases, SelA in bacteria and SepSecS in archaea and eukaryotes, catalyze the terminal reaction of selenocysteine synthesis during which either serine or phosphoserine is converted into selenocysteine while being attached to tRNA^{Sec}. The reaction mechanism is based on the PLP co-factor and requires selenophosphate as selenium donor. Although adopting different structures, SelA and SepSecS employ similar general principles for substrate recognition and chemical catalysis. Five SelA homodimers assemble into a ring structure that harbors ten tRNA-binding and catalytic sites. Each SelA subunit binds to the D-arm and T loop of tRNA^{Sec}. By contrast, two SepSecS homodimers arrange into a tetramer that binds to the acceptor and variable arms of tRNA^{Sec}. In both instances, at least two enzyme homodimers are needed to bind and act on one tRNA^{Sec} molecule. A 'non-catalytic' dimer employs its N-terminal domain to dock the tRNA to the enzyme, whereas residues of the C-terminal domain of the 'catalytic' dimer orient the aminoacyl group into the active site for catalysis to occur. Herein, the mechanism and structure of selenocysteine synthases are summarized.

1. Introduction

Selenium is the only essential dietary micronutrient that is found as a constitutive component in specific proteins and enzymes in all domains of life. Its precise location in

selenoproteins is imprinted in the DNA. The major physiological form through which selenium exerts its biological roles is selenocysteine, the 21st amino acid. Since its discovery in proteins (8, 108, 109), questions were raised about the importance and function of selenocysteine, and whether it could be substituted by isosteric amino acids, cysteine and serine. A typical selenoproteome is of limited size (9, 110, 111), but its members are important for the health, development, and survival of organism. Although composed of only 25 members (9), the human selenoproteome is vital for the maintenance of the cellular redox potential, protection of the membrane and DNA from oxidative damage, removal of reactive oxygen species, and regulation of the thyroid hormone homeostasis and metabolic rate (reviewed in (15, 16, 112)). Moreover, substitution of selenocysteine with cysteine or serine either completely diminishes or significantly reduces catalytic prowess of selenoenzymes (19-21, 113-118). Given that mutations affecting selenoprotein gene translation cause various pathologies including cancer (reviewed in (15, 16, 22, 112, 119-122)) and that the mouse tRNA^{Sec} knock-out mutant is embryonically lethal (38), it is reasonable to suggest that the efficient synthesis and insertion of selenocysteine are fundamental biological processes. It is thus not surprising that significant efforts were invested in determining identities of macromolecular components, the sequence of events, and the mechanism(s) of discrete steps governing selenocysteine synthesis across kingdoms.

Early studies indicated that selenium is an essential component of archaeal hydrogenase, bacterial glycine reductase (8), nicotinic acid hydroxylase (123), formate dehydrogenase (108) and xanthine dehydrogenase (124), and avian and mammalian

glutathione peroxidase (125-127). The assumption was that selenium was incorporated into proteins by an unknown post-translational mechanism. However, surprising observations that an in-frame UGA (*opal*) stop codon in the bacterial formate dehydrogenase and mouse glutathione peroxidase signals insertion of selenocysteine (18, 128), ignited an avalanche of extraordinary and unforeseen discoveries. Genetic and biochemical studies in bacteria have identified *selA*, *selB*, *selC* and *selD* as essential genes for selenoprotein synthesis. The monocistronic *selC* gene (129) was shown to encode tRNA^{Sec}, which is serylated by the cytosolic SerRS, 'reads' the UGA codon, and supports insertion of selenocysteine with the aid of specialized elongation factor SelB (27). tRNA^{Sec} was subsequently shown to adopt a novel fold, 8/5 in prokaryotes and 9/4 in eukaryotes, characterized by 13 base pairs in the acceptor-TΨC arm, and extended D- and variable arms (107, 130-133). This fold is in contrast to a canonical 7/5 fold and a 12-base pairs long acceptor-TΨC arm found in all other elongator tRNAs. Taken together, the results validated a proposition that selenocysteine is synthesized from serine (134) on its tRNA (135) in a reaction that requires SelD (135) and SelA (30, 129, 136-138). Concurrent studies on the bacterial and eukaryotic orthologs have firmly established that SelD is selenophosphate synthetase, the enzyme that converts selenide into selenophosphate in the presence of ATP (139, 140). On the other hand, phylogenetic, biochemical and enzymatic analyses suggested that SelA is a selenocysteine synthase that promotes conversion of Ser-tRNA^{Sec} into Sec-tRNA^{Sec} (30, 137). Hence, a novel two-step indirect aminoacylation pathway for selenocysteine synthesis was sketched out. In the first step, SerRS attaches L-serine onto tRNA^{Sec} at

the expense of one ATP molecule, and in the second, SelA promotes the serine-to-selenocysteine conversion while utilizing selenophosphate (30, 139).

Whereas the bacterial cycle of selenocysteine was delineated in the early 1990s, understanding of the analogous process in archaea and eukaryotes was significantly lagging. It took almost 20 years to ascertain that an unusual phosphoserine tRNA from mammalian brain, liver (141, 142) mammary glands (143), and avian liver (142), and the bovine liver *opal* suppressor serine tRNA (144, 145) represent, in fact, the eukaryotic tRNA^{Sec} (28). After purification of a specific kinase that phosphorylates the *opal* suppressor Ser-tRNA^{Sec} (146) it was proposed that the eukaryotic Sec-tRNA^{Sec} is synthesized from Ser-tRNA^{Sec} via Sep-tRNA^{Sec} intermediate (147-149). This proposal was thereafter reversed and it was suggested that the eukaryotic and bacterial processes follow the same reaction mechanism (150, 151). Around the same time, a report that autoantibodies from patients suffering from autoimmune hepatitis (152-154) precipitate a stable complex between a 48-kDa protein and tRNA^{Sec} went almost unnoticed because the precipitated protein factor was considered as a component that is somehow involved in co-translational incorporation of selenocysteine (155). The protein factor, known as soluble liver and pancreas antigen (SLA/LP) (152-154), was later sequenced, cloned and purified (156, 157). The bioinformatics analyses have revealed that SLA/LP is well conserved in archaeal and eukaryotic organisms that contained selenoproteins and that it is likely a PLP-dependent enzyme participating in a group transfer reaction (51, 158). Hatfield and co-workers then identified murine *O*-phosphoseryl-tRNA^{Sec} kinase (PSTK), which promotes the phosphoryl group transfer

from ATP onto Ser-tRNA^{Sec} and not onto Ser-tRNA^{Ser} (31). Not only that the discovery of PSTK lent the credence to the original proposal that synthesis of selenocysteine in higher organisms proceeds *via* the phosphoseryl intermediate, but it also explained why phosphoserine tRNA and the kinase activity were detected in the first place (141-143). In spite of this, the evidence that specific enzyme can form selenocysteine from the tRNA carrying phosphoserine was missing. That such enzymatic activity may exist was suggested by an unrelated study on the tRNA-dependent synthesis of cysteine in certain archaea. Söll and co-workers identified an archaeal enzyme, SepCysS, which converted Sep-tRNA^{Cys} into Cys-tRNA^{Cys} (159). The enzyme was classified as the Fold Type I PLP dependent enzyme, just like Sela. This prompted an intense search for an archaeal and eukaryotic enzyme that could utilize Sep-tRNA^{Sec} as a substrate. Two research groups independently identified murine and archaeal O-phosphoseryl-tRNA^{Sec}:selenocysteinyl-tRNA^{Sec} (SepSecS or SepS) that indeed supported selenocysteine and selenoprotein synthesis (160, 161). The enzyme promoted phosphoseryl-to-selenocysteinyl conversion in the presence of selenophosphate while retaining stringent specificity towards Sep-tRNA^{Sec}. Importantly, murine and archaeal SepSecS were shown to be homologs of human SLA/LP identified in the early 1990s, which, in turn, explained why SLA/LP formed a stable complex with Sep-tRNA^{Sec}. Taken together, all results suggested that the selenocysteine cycle in higher organisms is composed of three reactions. As in bacteria, SerRS first attaches serine onto tRNA^{Sec}. Subsequently, PSTK converts Ser-tRNA^{Sec} into Sep-tRNA^{Sec}, and in the terminal reaction, SepSecS substitutes selenol for phosphoryl yielding Sec-tRNA^{Sec}. Although

yielding the same reaction product, SelA and SepSecS use different tRNA-based substrates, implying that their catalytic mechanisms and structures are not completely conserved. In this chapter, our current understanding of selenocysteine synthases is presented.

2. The structure and architecture of the bacterial SelA

Following genetic and biochemical studies that identified SelA as the bacterial selenocysteine synthase, significant strides were made to determine the structure and architecture of both holo SelA and the SelA-tRNA^{Sec} binary complex. The CryoEM studies on *E. coli* and *Moorella thermoacetica* SelA established that 10 enzyme subunits, each containing one PLP molecule, arrange into a ~500-kDa fivefold symmetric structure (162, 163). However, the exact stoichiometry of the SelA-tRNA^{Sec} complex remained controversial (163, 164). Most recently, Yokoyama and co-workers provided a wealth of structural and biochemical information on this system by studying *Aquifex aeolicus* SelA and *Thermoanaerobacter tengcongensis* tRNA^{Sec} (131, 165). The 3.9-Å resolution crystal structure of holo SelA revealed that the enzyme is a homodecamer in which 10 subunits are arranged into a pentamer of dimers (Figure 1). Each SelA subunit is composed of three domains: an N-terminal domain (NTD; residues 1-66), a core domain (residues 90-338) and a C-terminal domain (CTD; residues 339-452), which fold into a structure typical of Fold Type I PLP-dependent enzymes (Figure 1a). A short linker (residues 67-89) connects NTD with the core domain. The enzyme belongs to a distinct taxon that is likely to form an independent group related to the

cystathionine γ -synthase family (165). Structurally, two SclA monomers form an intimate dimer at which interface two active sites are formed. Each active site contains the PLP cofactor covalently attached to a conserved Lys²⁸⁵ residue (see section 9.4). The mobile NTD protrudes from the plane of the ring structure and interacts with the core domain of the neighboring intimate dimer (Figure 1). On the other hand, CTD is oriented differently from that in other Fold Type I PLP enzymes and this enables interactions between neighboring dimers and formation of the pentameric SclA ring. The ring harbors large interdimer clefts that play an important function in tRNA binding and catalysis.

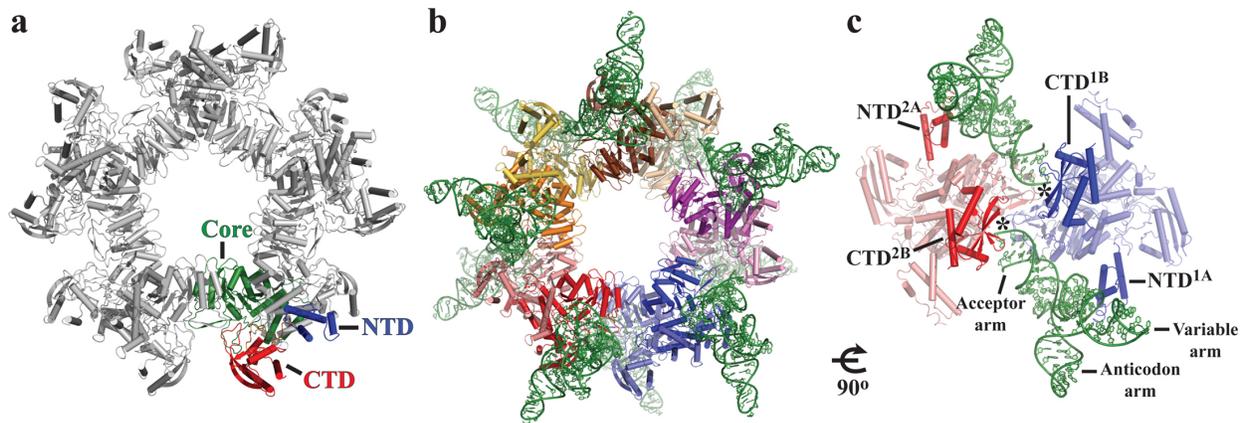


Figure 1. The bacterial SslA forms a ring structure that binds 10 tRNA^{Sec} molecules. (a) A ribbon diagram of holo SslA (PDBID: 3W1J), viewed above the ring plane, reveals that the enzyme adopts a pentameric ring-like structure. Domains in one SslA monomer are colored according to the following scheme: NTD is blue, the core domain is green, and CTD is red; other SslA subunits are grey. (b) The same orientation of the SslA-tRNA^{Sec} binary complex (PDBID: 3W1K) shows that 10 tRNAs bind to the SslA decamer. Intimate SslA dimers are in shades of similar colors (red and pink, orange and olive, brown and beige, purple and light purple, blue and light blue) and tRNA is green. (c) Two intimate dimers of SslA (pink and light blue) bind two tRNAs. NTD of monomer 1 (NTD^{1A}; blue) binds to the D-arm and T loop, and CTD of monomer 2 (CTD^{2B}; red) interacts with the tip of the acceptor arm. This is repeated on the other side of the dimer where NTD^{2A} (red) binds to the D-arm and T loop of the second tRNA, and the acceptor arm binds to CTD^{1B} (blue). The pattern continues around the ring structure. Asterisks designate locations of the catalytic grooves.

The crystal structure of the SslA-tRNA^{Sec} complex, solved at 7.5-Å resolution, provided first hints at how SslA recognizes tRNA^{Sec} (Figure 1b). The binary complex crystal contained 10 tRNA^{Sec} molecules bound to the SslA decamer thus suggesting 1:1 stoichiometry. tRNA^{Sec} interacts with the enzyme through its extended D- and acceptor-TΨC arms, whereas the variable and anticodon arms do not participate in binding (Figure 1c). Two intimate dimers bind one tRNA molecule in an arrangement in which one SslA dimer (e.g. subunits A, B or SslA^{AB}) binds the tRNA body while the other one

(e.g. subunits C, D or Sela^{CD}) positions and orients the aminoacylated CCA-end into the catalytic pocket (Figure 1c). This observation implies that only in a decameric arrangement can Sela form a productive complex with tRNA^{Sec}. Indeed, mutations that disrupted the dimer-dimer interface and the decamer structure inhibited the Sela activity, but retained ~90% of its tRNA-binding capacity (165). Intriguingly, the active site of the dimeric Sela mutant contains PLP and resembles that of other Fold Type I PLP enzymes including SepSecS, but in the absence of a stable pentameric ring the conformation of the active site is altered and catalysis does not occur (165).

Perhaps the most important structural element for tRNA recognition is NTD of Sela, which binds the D-arm and TΨC loop of tRNA^{Sec}. The deletion of NTD abolishes binding of Sela to both Ser-tRNA^{Sec} and unacylated tRNA^{Sec} (131). Moreover, this segment plays an important role in recognition and discrimination of tRNA^{Sec} from other tRNAs including a structural homolog tRNA^{Ser}. Particularly, the NTD interacts with the fifth and sixth base pairs G14:U21 and C15:G20a of the D-arm and with the triple base pair in the D loop. Because these base pairs are present in tRNA^{Sec} only, it is postulated that they may be specificity elements for all enzymes involved in selenocysteine synthesis. On the other hand, it was suggested that the length of the acceptor-TΨC arm of tRNA^{Sec} (8 nucleotides vs. 7 nucleotides found in canonical tRNAs) is not critical for recognition by Sela probably due to conformational mobility of the NTD. Further, an intimate dimer that binds to the D-arm and T loop does not interact with the CCA-end of tRNA^{Sec} and thus does not act on that particular tRNA (Figure 1c). Instead, this dimer positions tRNA so that the acceptor arm can interact with the C-terminal domain of the

neighboring intimate dimer. The tip of the acceptor arm binds to a large interdimer cleft where residues 423 and 424 of SelA might interact with the first base pair G1:C72 and G73 discriminator nucleotide. Mutational analyses have confirmed that recognition of the acceptor arm tip is important for selenocysteine synthesis (131). Furthermore, structural modeling suggested that the binding pocket for A76 is composed of residues from subunits A and C, Asn^{218A} and Phe^{224C}, which are in proximity only in the pentameric SelA. Hence, a model of tRNA recognition by SelA emerged in which one SelA dimer deploys NTD to dock tRNA^{Sec} onto the decamer, while the neighboring dimer uses CTD to interact and position the seryl-CCA into its catalytic groove. This pattern repeats on both sides of the SelA homodecamer yielding a large ribonucleoprotein assembly capable of simultaneously acting on 10 tRNA substrates.

3. The structure of the archaeal and eukaryotic SepSecS

The x-ray crystallography studies on the archaeal and murine holo SepSecS (166, 167), and the human SepSecS-tRNA^{Sec} binary complex (107) revealed the structure of selenocysteine synthase in higher organisms. SepSecS is a tetramer composed of a dimer of intimate dimers (Figure 2a). The intimate dimer interface encloses two active sites each carrying one PLP molecule covalently linked to a conserved Lys²⁸⁴ (see section 9.4). SepSecS is composed of three domains: an extended N-terminal domain (NTD; residues 1-130), a core domain (residues 131-315), and a C-terminal domain (CTD; residues 360-501) (Figure 2a). A linker composed of long α -helices (residues 316-359) joins the core domain and CTD. With the exception of the extreme C-terminus,

the overall structure of SepSecS is well conserved. The CTD of the archaeal SepSecS is shorter than the mammalian counterpart; it lacks a segment encompassing residues 473-493 which is identified as the antigenic region in patients suffering from chronic autoimmune hepatitis (reviewed in (119)).

SepSecS forms its own branch in the phylogenetic tree of Fold Type I PLP-dependent enzymes that may be related to sugar aminotransferase family (165, 167). This suggests that SepSecS is a primordial enzyme, probably present in the first common ancestor. A certain level of structural homology with the dimeric SepCysS and selenocysteine lyase is evident in the catalytic and CTD. On the other hand, both of these homologs lack the appended NTD, an element pivotal for oligomerization of SepSecS. In particular, N-terminal helices $\alpha 1$, $\alpha 2$, and $\alpha 4$ from each subunit interact with one another and establish a hydrophobic core of the tetramer. The importance of these interactions for tetramerization was shown by studies in which the deletion of helix $\alpha 1$ yielded a dimeric enzyme incapable of supporting selenocysteine synthesis (167). Thus, the appended NTD of SepSecS and the uniquely oriented CTD of SelA (131) are the major oligomerization elements in selenocysteine synthases.

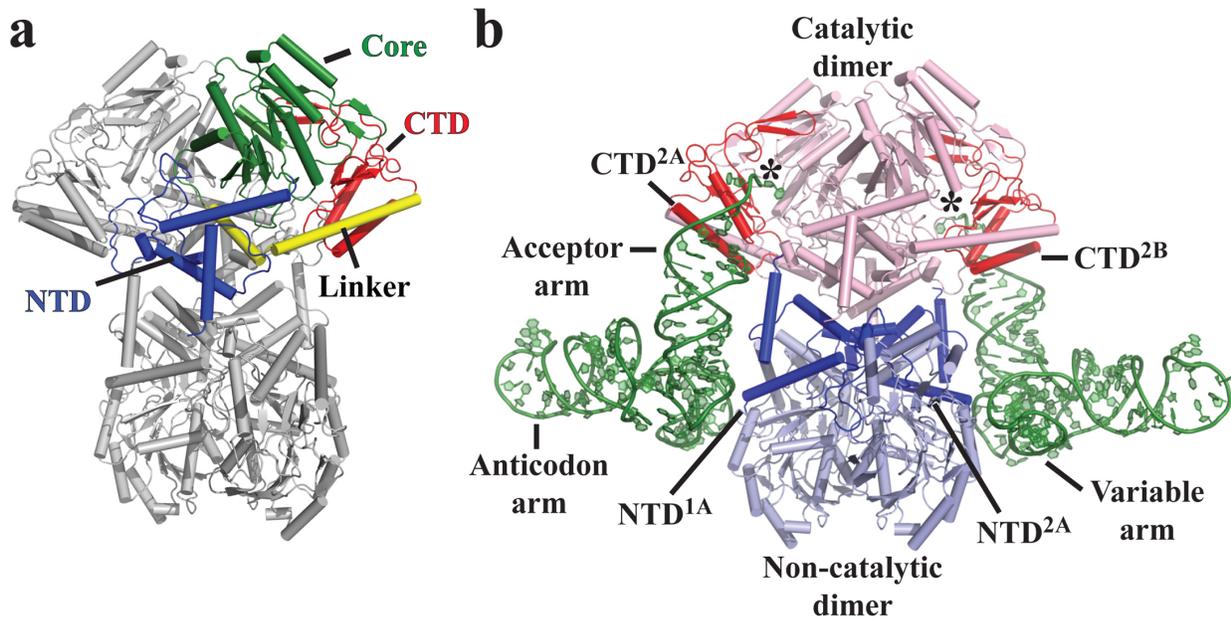


Figure 2. The SepSecS homotetramer binds up to two tRNA^{Sec} molecules. (a) Two SepSecS homodimers arrange into a tetramer. Domains of one subunit are colored: NTD is blue, the core domain is green, linker is yellow, and CTD is red. The rest of the tetramer is grey. **(b)** The binary complex structure (PDBID: 3HL2) shows that NTD (NTD^{1A}, NTD^{2A}) of the non-catalytic dimer (light blue) binds to the acceptor arm of tRNA, and that CTD (CTD^{2A}, CTD^{2B}) of the catalytic dimer (pink) binds to the tip of the acceptor arm. Asterisks mark the catalytic grooves (see Figure 9.3 for more detail).

The crystal structure of SepSecS-tRNA^{Sec} revealed the architecture of the binary complex, the mechanism of tRNA recognition, and provided hints about the complex stoichiometry and reaction mechanism (107). The structure showed that two tRNA^{Sec} molecules are bound to the SepSecS tetramer in a cross-dimer fashion (Figure 2b), implying that only the tetrameric SepSecS can bind and presumably act on Sep-tRNA^{Sec}. However, SepSecS binds to the opposite side of tRNA when compared to Sela. The N-terminal helix α 1 of the non-catalytic dimer binds to the long acceptor-T Ψ C

arm, which is the major recognition element, and a segment of its core domain interacts with the long variable arm (107). The tip of the acceptor arm binds to the residues in the C-terminal helices $\alpha 14$ and $\alpha 15$ of the catalytic dimer. In particular, three conserved arginine residues residing in helices $\alpha 14$ (Arg³⁹⁸) and $\alpha 15$ (Arg⁴⁵³, Arg⁴⁵⁶) interact with the G1:C72 base pair and the G73 discriminator nucleotide (107). Recognition of the G73 discriminator is important for selenocysteine synthesis (107), but the importance of the longer acceptor and variable arms for this process remains unclear. In addition, while there is no evidence that SepSecS undergoes a conformational change upon tRNA binding, the superimpositioning of the complexed (107) and unbound tRNA^{Sec} (133) revealed that the acceptor-, T Ψ C- and variable arms rotate around the vertical axis projecting through the anticodon arm by 24°, 17° and 33°, respectively, upon complex formation (168). This may be pivotal for orienting the phosphoseryl-CCA towards the catalytic groove.

Although SepSecS harbors four tRNA-binding sites, the binary complex crystal contained only two tRNAs bound to the tetramer. The same complex stoichiometry was observed when mimics of Sep-tRNA^{Sec} were used in binding assays (169). This suggests that SepSecS employs a half-sites activity, which is in contrast from the full-site activity exhibited by SelA. The results of fluorescence-quenching binding assays and small angle x-ray scattering analyses confirmed that SepSecS preferentially binds either one or two tRNA^{Sec} molecules at a time (170). However, it remains to be seen if the functional asymmetry of SepSecS is of physiological significance and if it is allosterically regulated.

4. Divergent active sites of selenocysteine synthases and a conserved catalytic mechanism

Both SelA and SepSecS harbor active sites formed at the interface of intimate dimers. Before atomic resolution structures were available, detailed enzymatic studies on SelA and SepSecS were pursued with the aim to delineate possible catalytic mechanism. Böck and co-workers established that the SelA-catalyzed substitution of selenol for hydroxyl proceeds through an anhydroalanyl intermediate covalently attached to PLP (30). In the presence of SelD, reduced selenium, ATP and sodium borohydride, the bacterial SelA converted Ser-tRNA^{Sec} into Ala-tRNA^{Sec} with a fraction of pyruvate, the breakdown product of anhydroalanine, being released into solution. The reaction was stoichiometric with 2 mol of enzyme acting on 1 mol of Ser-tRNA^{Sec}. Analogous studies on the archaeal (160), murine (161) and human (107) systems confirmed that the PLP cofactor is required for phosphoserine-to-selenocysteine conversion, but the stable anhydroalanyl intermediate remained elusive. This discrepancy was ascribed to the relative instability of reaction intermediates bound to SepSecS. More recent structural and mutational studies provided a wealth of information about the active sites of SelA and SepSecS, and possible catalytic mechanisms.

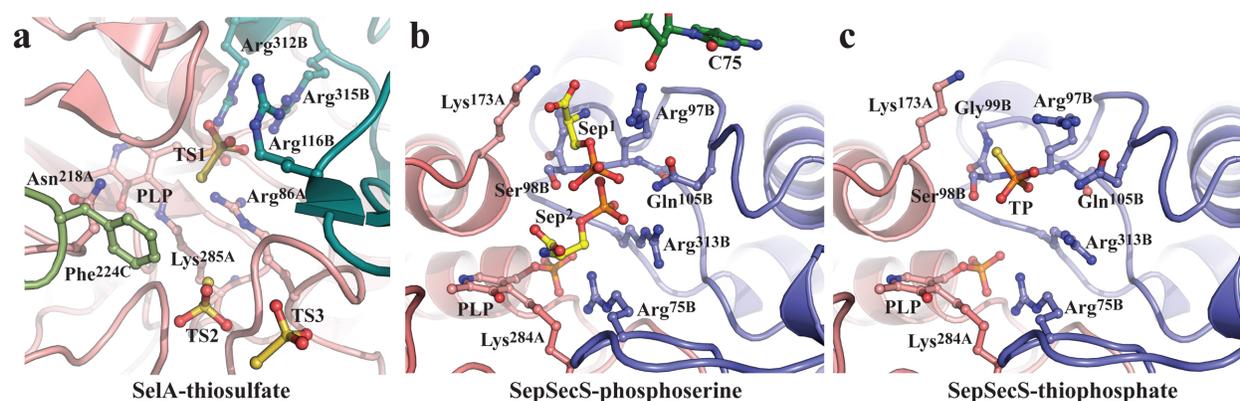


Figure 3. Catalytic sites in SelA and SepSecS, and binding pockets for phosphoserine and selenophosphate. (a) The active site in SelA is composed of residues from subunits A (teal) and B (pink) of the intimate dimer, and at least one residue from the neighboring dimer subunit C. Residues important for catalysis (Arg^{86A}, Arg^{116B}, Arg^{312B}, Arg^{315B}, Asn^{218A} and Phe^{224C}) are shown in sticks. Thiosulfate ion TS1 (gold sticks) designates the binding pocket for selenophosphate, and TS2 and TS3 mark the A76-binding pocket. (b) The active site in SepSecS is composed of residues only from subunits of the intimate dimer (pink and blue). Arg^{75B}, Ser^{98B}, Gln^{105B} and Arg^{313B}, which are important for catalysis, are shown in sticks. Free phosphoserine (Sep¹, Sep²; gold sticks) binds in two orientations, none of which are optimal for catalysis. (c) Thiophosphate (TP; gold sticks) binds to the P loop (Ser^{98B}-Gln^{105B}), just like TS1 in SelA.

The crystal structures of SelA in complex with thiosulfate (TS) (131), holo SepSecS (166, 167) and the human SepSecS-tRNA^{Sec} complex (107) revealed the three-dimensional structure of the active sites and suggested residues that might be pivotal for catalysis. In the SelA structure, TS1 binds to the putative selenophosphate-binding pocket, which is composed of Arg^{86A}, Arg^{312B} and Arg^{315B} (Figure 3a). The pocket sits atop PLP, which is lys coupled *via* Schiff base to Lys^{285A}. When a string of arginine side chains was mutated into alanines, the catalytic activity of SelA was markedly reduced. It was also shown that side chains of Arg^{119B} and Asp^{284A}, which

interact with Arg^{86A} and Arg^{312B}, might be important for catalysis (Figure 3a). Intriguingly, TS2 and TS3 bind near Asn^{218A} and Phe^{224J}, which are spatially proximal to the TS1-binding pocket. Because mutations of these two residues significantly diminished Sela activity, it was proposed that TS2 and TS3 might mimic the binding of A76. As a corollary, this proposal suggested that only as a pentameric ring could Sela bind Ser-tRNA^{Sec}, which was later shown to be correct (165).

Further, the crystal structures of SepSecS provided a comparable level of structural information. We shall briefly discuss only the human enzyme, but with a note that the same conclusions most likely apply to the archaeal and other eukaryotic orthologs. The SepSecS subunit of the catalytic dimer that interacts with the CCA-end (monomer A) provides the PLP cofactor linked to Lys^{284A}, Gln^{172A} that sits atop PLP, and Lys^{173A} (Figure 3b). The other subunit (monomer B) provides the P-loop (Ser^{98B} and Arg^{97B}), which binds thiophosphate, and a string of residues (Arg^{75B}, Gln^{105B}, and Arg^{313B}) important for catalysis (Figures 3b, c). The soaking experiments have shown that free phosphoserine binds to the active site of the catalytic dimer only, but in two orientations none of which are optimal for catalysis. This implies that the covalent attachment of phosphoserine to tRNA^{Sec} is essential for the proper placement of the phosphoseryl moiety into the active site and for orienting its amino group for attack onto the Schiff base of PLP. Thiophosphate, on the other hand, binds to the active site of the non-catalytic dimer and its binding pocket partly overlaps with that for free phosphoserine. This observation suggests that selenophosphate binds to SepSecS only after phosphate is eliminated from Sep-tRNA^{Sec}.

Taken together, a unified PLP-dependent mechanism for the terminal reaction of selenocysteine synthesis was proposed (Figure 4) (30, 107, 166, 167). The reaction begins with binding of Ser-tRNA^{Sec} and Sep-tRNA^{Sec} to SelA and SepSecS, respectively. After the seryl/phosphoseryl moiety is bound near PLP, its amino group attacks the Schiff base thus yielding the external aldimine. It is likely that the liberated Lys^{285/284} side chain abstracts the C α proton from serine/phosphoserine. The electron delocalization by the pyridine ring leads to a rapid β -elimination of water and phosphate from Ser-tRNA^{Sec} and Sep-tRNA^{Sec}, respectively. In both instances, the anhydroalanyl-tRNA^{Sec} intermediate is formed. Upon water/phosphate release, selenophosphate binds to the P-loop. The concomitant attack of water on selenophosphate and of the nucleophilic selenium onto the anhydroalanyl moiety yields an oxidized form of Sec-tRNA^{Sec} while releasing a second phosphate equivalent. Alternatively, selenophosphate attacks anhydroalanine and forms phosphoselenyl-tRNA^{Sec}, which subsequently breaks down to Sec-tRNA^{Sec} after water attack. Lastly, Lys^{284/285} re-establishes the internal aldimine and Sec-tRNA^{Sec} is released from the enzyme. It is suggested that, besides PLP, non-homologous arginine residues (Arg⁸⁶, Arg¹¹⁹, Arg³¹², and Arg³¹⁵ in SelA; Arg⁷⁵ and Arg³¹³ in SepSecS) play a critical role in catalysis. However, further structural and enzymatic studies are needed to both define precise role(s) for these residues and provide a complete picture about the remarkable mechanism of selenocysteine synthases.

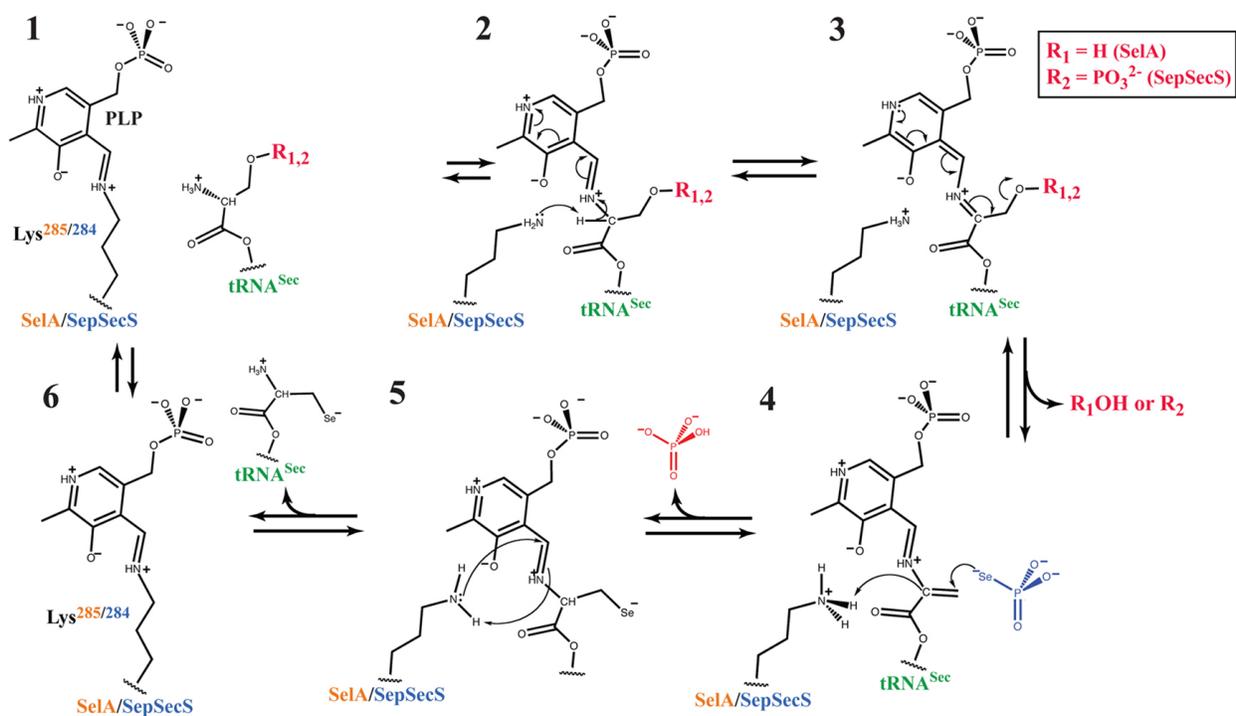


Figure 4. The proposed PLP-dependent mechanism of the terminal reaction of selenocysteine synthesis catalyzed by SelA and SepSecS. For details on the mechanism see section 9.4 of the main text and Figure 4 in references (107) and (131). The leaving groups in aminoacyl groups are highlighted in red: R₁ (SelA) is H and R₂ (SepSecS) is PO₃²⁻.

5. Future directions

Despite the remarkable progress, important questions about selenocysteine synthases and their biological roles remain unanswered. First, both SelA and SepSecS were crystallized with the unacylated tRNA^{Sec} and interactions of the seryl and phosphoseryl groups with the corresponding active sites were not visualized. Second, a stable anhydroalanyl intermediate bound to SepSecS has not been captured yet, leaving the exact mechanism in dark. Precise structural studies using substrate and intermediate analogs could provide more insight into the reaction mechanism. Third, given that SelA

and SepSecS are positioned downstream of selenophosphate synthetase, it is plausible that they play a role in regulating the overall selenium homeostasis. Fourth, given that SepSecS exhibits half-sites activity, it would be important to determine if this activity is allosterically regulated. Fifth, because SelaA and SepSecS bind tRNA^{Sec} from the opposite sides when compared to SerRS and PSTK, respectively, it is plausible that multi-enzyme 'selenosomes' facilitate selenocysteine synthesis. Lastly, given recent clinical reports (23, 24, 171, 172), studies on the role of SepSecS and selenoproteins in the development and maintenance of the healthy human brain are warranted.

Acknowledgements: This work was supported by a grant from the National Institute of General Medical Sciences of the National Institutes of Health R01 GM097042 (to M.S.)

References:

1. JB Jones et al 1979 *Arch Biochem Biophys* 195:255
2. JW Forstrom et al 1978 *Biochemistry* 17:2639
3. JE Cone et al 1976 *Proc Natl Acad Sci U S A* 73:2659
4. GV Kryukov et al 2003 *Science* 300:1439
5. AV Lobanov et al 2006 *Nucleic Acids Res* 34:496
6. Y Zhang et al 2006 *Genome Biol* 7:R94
7. FP Bellinger et al 2009 *Biochem J* 422:11
8. MP Rayman 2009 *Biochim Biophys Acta* 1790:1533
9. RL Schmidt, M Simonovic 2012 *Croat Med J* 53:535
10. C Rocher et al 1992 *Eur J Biochem* 205:955
11. M Maiorino et al 1995 *Biol Chem Hoppe Seyler* 376:651
12. VN Gladyshev et al 1996 *Proc Natl Acad Sci U S A* 93:6146
13. T Tamura, TC Stadtman 1996 *Proc Natl Acad Sci U S A* 93:1006
14. GG Kuiper et al 2003 *Endocrinology* 144:2505
15. HY Kim, VN Gladyshev 2004 *Mol Biol Cell* 15:1055
16. S Toppo et al 2008 *Antioxid Redox Signal* 10:1501
17. BC Lee et al 2009 *Biochim Biophys Acta* 1790:1471
18. MJ Axley et al 1991 *Proc Natl Acad Sci U S A* 88:8450

19. AM Dumitrescu, S Refetoff 2011 *Ann Endocrinol (Paris)* 72:95
20. S Palioura et al 2010 *Biol Chem* 391:771
21. A Lescure et al 2009 *Biochim Biophys Acta* 1790:1569
22. MP Rayman 2005 *Proc Nutr Soc* 64:527
23. GJ Beckett, JR Arthur 2005 *J Endocrinol* 184:455
24. MR Bosl et al 1997 *Proc Natl Acad Sci U S A* 94:5531
25. GL Dilworth 1982 *Arch Biochem Biophys* 219:30
26. R Wagner, JR Andreesen 1979 *Arch Microbiol* 121:255
27. SH Oh et al 1974 *Biochemistry* 13:1825
28. JT Rotruck et al 1973 *Science* 179:588
29. L Flohe et al 1973 *FEBS Lett* 32:132
30. I Chambers et al 1986 *EMBO J* 5:1221
31. F Zinoni et al 1986 *Proc Natl Acad Sci U S A* 83:4650
32. G Sawers et al 1991 *J Bacteriol* 173:4983
33. W Leinfelder et al 1988 *Nature* 331:723
34. Y Itoh et al 2013 *Nucleic Acids Res* 41:6729
35. Y Itoh et al 2013 *Science* 340:75
36. S Chiba et al 2010 *Mol Cell* 39:410
37. S Palioura et al 2009 *Science* 325:321
38. Y Itoh et al 2009 *Nucleic Acids Res* 37:6259
39. RA Sunde, JK Evenson 1987 *J Biol Chem* 262:933
40. W Leinfelder et al 1990 *Proc Natl Acad Sci U S A* 87:543
41. K Forchhammer, A Bock 1991 *J Biol Chem* 266:6324
42. K Forchhammer et al 1991 *Biochimie* 73:1481
43. K Forchhammer et al 1991 *J Biol Chem* 266:6318
44. P Tormay et al 1998 *Eur J Biochem* 254:655
45. A Ehrenreich et al 1992 *Eur J Biochem* 206:767
46. Z Veres et al 1992 *Proc Natl Acad Sci U S A* 89:2975
47. D Hatfield, FH Portugal 1970 *Proc Natl Acad Sci U S A* 67:1200
48. PH Maenpaa, MR Bernfield 1970 *Proc Natl Acad Sci U S A* 67:688
49. SJ Sharp, TS Stewart 1977 *Nucleic Acids Res* 4:2123
50. A Diamond et al 1981 *Cell* 25:497
51. D Hatfield et al 1982 *Proc Natl Acad Sci U S A* 79:6215
52. BJ Lee et al 1989 *J Biol Chem* 264:9724
53. T Mizutani, A Hashimoto 1984 *FEBS Lett* 169:319
54. T Mizutani, T Hitaka 1988 *FEBS Lett* 232:243
55. T Mizutani 1989 *FEBS Lett* 250:142
56. T Mizutani, Y Tachibana 1986 *FEBS Lett* 207:162
57. T Mizutani et al 1991 *FEBS Lett* 289:59
58. T Mizutani et al 1992 *Biochem J* 284 (Pt 3):827
59. P Berg et al 1981 *Verh Dtsch Ges Inn Med* 87:921
60. M Teufel et al 1983 *Eur J Pediatr* 140:30
61. M Manns et al 1987 *Lancet* 1:292
62. C Gelpi et al 1992 *Proc Natl Acad Sci U S A* 89:9739

63. M Costa et al 2000 *Clin Exp Immunol* 121:364
64. I Wies et al 2000 *The Lancet* 355:1510
65. T Kernebeck et al 2001 *Hepatology* 34:230
66. XM Xu et al 2005 *J Biol Chem* 280:41568
67. BA Carlson et al 2004 *Proc Natl Acad Sci U S A* 101:12848
68. A Sauerwald et al 2005 *Science* 307:1969
69. J Yuan et al 2006 *Proc Natl Acad Sci U S A* 103:18923
70. XM Xu et al 2007 *PLoS Biol* 5:e4
71. N Fischer et al 2007 *Biol Chem* 388:1061
72. H Engelhardt et al 1992 *Mol Microbiol* 6:3461
73. LR Manzine et al 2013 *FEBS Lett* 587:906
74. Y Itoh et al 2014 *J Mol Biol* 426:1723
75. OM Ganichkin et al 2008 *J Biol Chem* 283:5849
76. Y Araiso et al 2008 *Nucleic Acids Res* 36:1187
77. J Yuan et al 2010 *FEBS Lett* 584:342
78. L Rigger et al 2013 *Chemistry* 19:15872
79. RL French et al 2014 *J Biol Chem* 289:28783
80. AK Anttonen et al 2015 *Neurology* 85:306
81. O Agamy et al 2010 *Am J Hum Genet* 87:538
82. P Makrythanasis et al 2014 *Hum Mutat* 35:1203
83. B Ben-Zeev et al 2003 *J Med Genet* 40:e96

VITA

- NAME: Malgorzata Dobosz-Bartoszek
- EDUCATION: M.S., Chemistry, Jagiellonian University, Krakow, Poland, 2009
Ph.D., Biochemistry and Molecular Genetics, University of Illinois at Chicago, Chicago, Illinois, 2016
- RESEARCH EXPERIENCE: Graduate Research Assistant, Laboratory of Dr. Miljan Simonović, Department of Biochemistry and Molecular Genetics, University of Illinois at Chicago, Chicago, Illinois, May 2012 – August 2016
Research Technician, Laboratory of Dr. Robert Keenan, Department of Biochemistry and Molecular Biology, The University of Chicago, Chicago, Illinois, July 2009 – July 2011
Visiting Master Student, Laboratory of Dr. Robert Keenan, Department of Biochemistry and Molecular Biology, The University of Chicago, Chicago, Illinois, July 2008 – June 2009
Visiting Undergraduate Research Assistant, Laboratory of Guy Weinberg MD, Department of Anesthesiology, Jessie Brown VA Hospital, Chicago, Illinois, September 2006, July 2007 - September 2007
- PUBLICATIONS: **Dobosz-Bartoszek M**, Pinkteron M, Otwinowski Z, Chakravarthy S, Söll D, Copeland PR, Simonović M (2016) Crystal structures of human eEFSec suggest a non-canonical mechanism for selenocysteine incorporation (*under review, Nature Communications*).
Dobosz-Bartoszek M and Simonović M (2016) Structure and mechanism of selenocysteine synthases, in "Selenium", Hatfield DL, Schweizer U, Tsuji P, Gladyshev VN, Editors, *Springer Science+Business Media (in press)*.
Chavan T, Meyer J, Chisholm L, **Dobosz-Bartoszek M**, Gaponenko V (2014) A novel method for the production of fully modified K-Ras 4B, *Methods in Molecular Biology* **1120**, 19-32.

Mariappan M*, Mateja A*, **Dobosz M**, Bove E, Hegde R, Keenan R (2011) The mechanism of membrane-associated steps in tail-anchored protein insertion, *Nature* **477**, 61-66. **equal contribution*

Mateja A, Szlachcic A, Downing M, **Dobosz M**, Mariappan M, Hegde R, Keenan R (2009) The structural basis of tail-anchored membrane protein recognition by Get3, *Nature* **461**, 361-366.

SEMINARS:

Dobosz-Bartoszek M, Pinkteron M, Otwinowski Z, Söll D, Copeland PR, Simonović M, "Crystal structures of human eEFSec suggest a non-canonical mechanism for selenocysteine incorporation", RNA Club, The University of Chicago, Chicago, Illinois, April 2016

Dobosz-Bartoszek M, Otwinowski Z, Simonović M, "The crystal structure of human selenocysteine tRNA-specific elongation factor, eEFSec", Structural Biology Journal Club, University of Illinois at Chicago, Chicago, Illinois, October 2014

Dobosz-Bartoszek M, Otwinowski Z, Simonović M, "The crystal structure of human selenocysteine tRNA-specific elongation factor, eEFSec", The 25th tRNA Conference, Kyllini, Greece, September 2014

Dobosz M, Keenan R "The regulation of ATPase activity of *Saccharomyces cerevisiae* Get3", The 3rd Annual BMB Visiting Students Symposium, The University of Chicago, Chicago, Illinois, June 2009

POSTERS:

Jovanovic M, **Dobosz-Bartoszek M**, Simonović M "Characterization of functional mutants of human selenocysteine tRNA-specific elongation factor, eEFSec", The Honors College Research Symposium, University of Illinois at Chicago, Chicago, Illinois, November 2015

Dobosz-Bartoszek M, Otwinowski Z, Copeland PR, Simonović M "Structural and biochemical studies of the human selenocysteine tRNA-specific translational elongation factor eEFSec", The 8th Annual Departmental Retreat, Fontana, Wisconsin, October 2015

Dobosz-Bartoszek M, Otwinowski Z, Simonović M "The crystal structure of human selenocysteine tRNA-specific elongation factor, eEFSec", International Year of Crystallography Symposium, Chicago, Illinois, October 2014

Dobosz-Bartoszek M, Otwinowski Z, Simonović M “The crystal structure of human selenocysteine tRNA-specific elongation factor, eEFSec”, The 7th Annual Departmental Retreat, Fontana, Wisconsin, September 2014

Dobosz-Bartoszek M and Simonović M “Structural studies of human selenocysteine tRNA-specific translational elongation factor, eEFSec“, Quo Vadis Structural Biology? SBGrid/NE-CAT Computing School, Harvard Medical School, Boston, Massachusetts, June 2014

Dobosz-Bartoszek M and Simonović M “Structural studies of the human selenocysteine tRNA-specific elongation factor, eEFSec”, The 1st Annual MBRB Retreat, Chicago, Illinois, June 2014

Schmidt RL, **Dobosz-Bartoszek M**, Simonović M “Architecture of catalytic complexes essential for synthesis and co-translational insertion of selenocysteine in humans”, The 18th Annual Meeting of RNA Society, Davos, Switzerland, June 2013

Dobosz M and Simonović M “The role of specialized elongation factor, EFsec, in incorporation of selenocysteine into selenoproteins”, The 5th Annual Departmental Retreat, Fontana, Wisconsin, August 2012

Dobosz M, Mateja A, Downing M, Keenan R, “Get3 receptor mediated tail-anchored membrane protein targeting to the ER”, 2009 Molecular Biosciences Retreat, Galena, Illinois, November 2009

Dobosz M, Mateja A, Szlachcic A, Downing M, Keenan R, “The role of dimerization in tail-anchored membrane protein targeting to the ER by Get3”, 2008 Molecular Biosciences Retreat, Galena, Illinois, November 2008

TEACHING
EXPERIENCE:

Supervising undergraduate and junior graduate students working in the Simonovic laboratory, Department of Biochemistry and Molecular Genetics, University of Illinois at Chicago, Chicago, Illinois, 2015 – present

Chemistry teacher practice at High School level (within framework of Teaching Practicum), Department of Chemistry, Jagiellonian University, Krakow, Poland, November 2007 – March 2008

Chemistry teacher practice at Junior High School level (within framework of Teaching Practicum), Department of Chemistry, Jagiellonian University, Krakow, Poland, April 2007 – May 2007, September 2007

AWARDS:

First place poster prize, The 8th Annual Departmental Retreat, Fontana, Wisconsin, October 2015

Honorable mention, Women in Science and Engineering Graduate Student Award, University of Illinois at Chicago, Chicago, Illinois, April 2015

Poster award, International Year of Crystallography Symposium, Chicago, Illinois, October 2014

The RNA Institute Travel Award, SUNY, New York, to attend the 25th tRNA Conference, August 2014

The Graduate Student Council Travel Award, University of Illinois at Chicago, Chicago, Illinois, to attend the 25th tRNA Conference, August 2014