

Crystal structure analysis of selenocysteine biosynthesis components

Dissertation

for the award of degree

“Doctor of Philosophy” (Ph.D.)

Division of Mathematics and Natural Sciences

of the Georg-August-Universität Göttingen

submitted by

Oleg Ganichkin

from Makhachkala, Russia

Göttingen, 2010

Prof. Dr. Markus C. Wahl (Reviewer),
Structural Biochemistry Group, Freie Universität Berlin

Prof. Dr. Ralf Ficner (Reviewer),
Dept. of Molecular Structural Biology, University of Göttingen

Prof. Dr. Detlef Doenecke,
Dept. of Molecular Biology, Institute of Biochemistry and Molecular Cell Biology

Date of the oral examination:

AFFIDAVIT

Here I declare that my Ph.D. dissertation entitled “Crystal structure analysis of selenocysteine biosynthesis components” has been written independently and with no other sources and aids then quoted. I have made clear what was done by project collaborators.

Parts of this thesis were published in:

Ganichkin O.M., Xu X.M., Carlson B.A., Mix H., Hatfield D.L., Gladyshev V.N., Wahl M.C. 2008. Structure and catalytic mechanism of eukaryotic selenocysteine synthase. *J Biol Chem.* **283**:5849-65

Ganichkin O., Wahl MC. 2007. Conformational switches in winged-helix domains 1 and 2 of bacterial translation elongation factor SelB. *Acta Crystallogr D Biol Crystallogr.* **Pt 10**:1075-81

Oleg Ganichkin

Göttingen, 2010

TABLE OF CONTENTS

1 Summary	1
2 Introduction	3
2.1 Selenium.....	3
2.2 Selenocysteine biosynthesis and incorporation in prokaryotes.....	3
2.3 Selenocysteine biosynthesis and incorporation in eukaryotes.....	5
2.4 Selenocysteine specific elongation factors.....	8
2.5 Selenocysteine synthase.....	10
2.6 tRNA ^{Sec}	13
2.7 Aims of study.....	17
3 Materials and methods	18
3.1 Materials.....	18
3.1.1 Chemicals.....	18
3.1.2 Media and Antibiotics.....	18
3.1.3 Enzymes and Enzyme Inhibitors.....	19
3.1.4 Nucleotides and radionucleotides.....	19
3.1.5 Oligonucleotides.....	19
3.1.6 Plasmids.....	19
3.1.7 <i>E. coli</i> Strains.....	20
3.1.9 Commercial Kits.....	20
3.1.10 Crystallization screens.....	20
3.1.11 Working Materials.....	20
3.1.12 Technical Devices.....	21
3.2. Methods.....	21
3.2.1 General molecular biology methods.....	21
3.2.1.1 Concentration determination of nucleic acids.....	21
3.2.1.2 Phenol-chloroform-isoamylalcohol (PCI) extraction.....	22
3.2.1.3 Agarose gel electrophoresis of nucleic acids.....	22
3.2.1.4 DNA extraction from agarose gels.....	23
3.2.1.5 Polymerase chain reaction (PCR).....	23
3.2.1.6 Restriction digestion of DNA.....	24
3.2.1.7 DNA ligation.....	24

3.2.1.8 Site-directed mutagenesis.....	25
3.2.1.9 Transformation of <i>E. coli</i> cells by heat shock.....	25
3.2.1.10 Mini-preparation of plasmid DNA.....	26
3.2.1.11 DNA Sequencing.....	26
3.2.1.12 <i>E. coli</i> cells cultivation and storage.....	26
3.2.1.13 Preparation of radiolabelled transcripts.....	27
3.2.1.14 5'-End labeling of RNA or DNA-oligonucleotide.....	27
3.2.1.15 Autoradiography.....	27
3.2.1.16 Denaturing polyacrylamide gel electrophoresis (PAGE)	27
3.2.1.17 Native polyacrylamide gel electrophoresis of RNA	28
3.2.1.18 Silver staining of RNA	28
3.2.2 Standard methods for protein analysis.....	28
3.2.2.1. Protein denaturing SDS-polyacrylamide gel electrophoresis.....	28
3.2.2.2 Gel staining methods.....	29
3.2.2.3 Small scale solubility assay.....	29
3.2.2.4 Domain Mapping with Limited Proteolysis.....	31
3.2.3 Protein production.....	31
3.2.3.1 High scale expression of desired construct.....	31
3.2.3.2 Cell lysis.....	32
3.2.3.3 Purification and cleavage of His-tagged fusion proteins.....	32
3.2.3.4 Protein concentrating and concentration determination.....	32
3.2.4 RNA production.....	32
3.2.4.1 DNA fragment assembly.....	32
3.2.4.2 Preparation of DNA template for in vitro transcription.....	33
3.2.4.3 <i>In vitro</i> transcription from PCR product.....	33
3.2.5. Purification of selenocysteine biosynthesis components.....	34
3.2.5.1 Purification protocol for <i>mmuSecS</i> protein.....	34
3.2.5.2 Purification protocol for <i>mthSelB</i> ³⁷⁷⁻⁵¹¹ protein.....	35
3.2.5.3 Non-denaturing purification of <i>in vitro</i> transcribed RNA ^{Sec}	35
3.2.6 Crystallization experiments	36
3.2.6.1 Crystallization conditions screening.....	36
3.2.6.2 Data collection and structure determination.....	37
3.2.7 Bioinformatics methods.....	39
3.2.7.1 Alignment.....	39

3.2.7.2 Domain Prediction.....	39
3.2.7.3 Fold Prediction.....	39
3.2.8 Methods used to study RNA-protein interactions.....	40
3.2.8.1 Electrophoretic mobility shift assay (EMSA)	40
3.2.8.2 Filter binding assay.....	40
3.2.8.3 Analytical gel filtration.....	40
4 Results.....	41
4.1 Conformational switches in winged-helix domains 1 and 2 of bacterial translation elongation factor SelB.....	41
4.1.1 Putative complex between WH1/2 of SelB and helix 16 of 16S rRNA.....	41
4.1.2 Overall fold and global conformational changes.....	42
4.1.3 Local conformational switch in WH1.....	44
4.1.4 Emergence of a bona fide ligand-binding site.....	45
4.1.5 Flexibility of the connecting peptide between WH2 and WH3.....	46
4.2 Structure and catalytic mechanism of eukaryotic selenocysteine synthase	47
4.2.1 Limited proteolysis delineates a SecS core fully active in P-Ser-tRNA ^{Sec} to Sec-tRNA ^{Sec} conversion.....	47
4.2.2 SecS is a member of the fold type I family of PLP-dependent enzymes with distinct structural elements.....	50
4.2.3 Cross-strutting via the N terminus leads to homotetramers that exhibit surface properties suitable for binding P-Ser-tRNA ^{Sec}	53
4.2.4 The PLP cofactor is tightly anchored by non-canonical contacts to both protomers of a close dimer	55
4.2.5 Binding of phosphate triggers disorder-order transition in an active site loop.....	57
4.2.6 Phosphate binding is mediated by conserved and SecS-specific residues that are essential for selenocysteine synthesis.....	58
4.2.7 The P-loop could serve as a binding site for the P-Ser γ -phosphate and SeP.....	61
4.2.8 SecS discriminates against free <i>O</i> -phospho- <i>L</i> -serine.....	62
4.2.9 The inhibition profile of SecS resembles that of β -lyases.....	63
4.3 Crystal structure analysis of tRNA ^{Sec}	63
4.3.1 Co-crystallization of tRNA ^{Sec} with PSTK.....	63
4.3.2 Crystallization of tRNA ^{Sec}	66
4.3.3 Overall structure of tRNA ^{Sec}	67

4.3.4 The AD linker.....	68
4.3.5 The variable arm.....	69
4.3.6 The anticodon arm.....	70
4.3.7 The influence of 2'-O-methylation of U34 on anticodon loop conformation.....	71
4.3.8 Role of water molecules in tRNA ^{Sec} structure.....	72
4.3.9 Conformational flexibility.....	73
4.3.10 Metal binding sites.....	74
4.4 Purification protocol for crystallization-grade RNA.....	75
4.4.1 Template design for <i>in vitro</i> transcription.....	75
4.4.2 Template production.....	76
4.4.3 In vitro run-off transcription using T7 RNA polymerase.....	76
4.4.4 RNA purification.....	77
4.4.5 Biochemical activity of purified RNAs.....	80
4.4.6 Engineering of RNA constructs for crystallization.....	80
4.4.7 Crystallization of tRNA ^{Sec} constructs.....	82
5 Discussion.....	83
5.1 Conformational switches in winged-helix domains 1 and 2 of bacterial translation elongation factor SelB.....	83
5.2 Structure and catalytic mechanism of eukaryotic selenocysteine synthase	84
5.2.1 SecS orthologs constitute a unique subclass of fold type I PLP-dependent enzymes.....	84
5.2.3 tRNA selection strategy of SecS.....	85
5.2.4 Mechanisms for substrate binding and differentiation in the active site.....	86
5.2.5 Structure-based reaction mechanism	87
5.2.6 New insights into SecS action.....	89
5.3 Crystal structure analysis of tRNA ^{Sec}	90
5.4 Purification protocol for crystallization-grade RNA.....	92
6 References.....	94
7 Appendix.....	107
7.1 X-ray crystallography method.....	107
7.1.1 Crystallization and crystal growth.....	107
7.1.2 Data collection.....	107

7.1.3 Methods for resolving the phase problem.....	109
7.1.4 Model building and structure refinement.....	112
7.2 Supplementary Tables and Figures.....	113
7.3 List of abbreviations	117
8 Acknowledgements.....	119
9 Curriculum vitae.....	121

1 SUMMARY

Some organisms from all three domains of life eukaryotes, bacteria and archaea, incorporate selenocysteine (Sec) into fractions of their proteins. Sec synthesis takes place on its cognate tRNA^{Sec} which is initially aminoacylated with serine by seryl-tRNA^{Ser} synthetase (SerRS). For conversion of Ser-tRNA^{Sec} to Sec-tRNA^{Sec} in bacteria responsible selenocysteine synthase (SelA), whereas in eukaryotes/archaea *O*-phospho-L-seryl-tRNA^{Sec} kinase (PSTK) and selenocysteine synthase (SecS). In Sec incorporation during translation in response to a special UGA codon involved several factors: selenocysteine insertion sequence on selenoprotein mRNA (SECIS), elongation factor (SelB in bacteria or EFSec in eukaryotes) and eukaryotic SECIS-binding protein 2 (SBP2).

In the bacterial UGA-decoding process, elongation factor SelB, which specifically recognizes Sec-tRNA^{Sec} but non of the aminoacylated tRNAs, plays an important role. The N-terminal part harbors EF-TU homologous domains and the C-terminal extension consists of four winged-helix (WH) motifs arranged into two functional modules WH1/WH2 and WH3/WH4. The last two motifs specifically interact with the stem-loop structure located on selenoprotein mRNA right after the UGA Sec-codon, which functions as SECIS element. The role of WH1/WH2 was not well understood. Studies of the SelB-ribosome complex by cryo-EM demonstrate possible interaction of WH1/WH2 motifs with helix 16 of 16S rRNA. These two motifs from *Moorella thermoacetica* were co-crystallized together with helix 16 of 16S rRNA to uncover the mechanism underlying this interaction. In the crystal structure determined at 1.1 Å resolution only the WH1/WH2 motifs were found. However, compared to the previous structure of the two motifs in conjunction with the WH3 and WH4 modules, the first winged-helix motifs underwent a substantial conformational change, which was elicited by a change in the orientation of Trp396. Additionally, the C-terminal tail of the second functional module followed a different path compared to the previous structure. It is conceivable that these conformational switches constitute part of the molecular mechanism that underlies the communication between the N- and C-terminal functional parts of SelB.

In eukaryotes and archaea, SecS is the last enzyme in Sec biosynthesis. It converts *O*-phospho-L-seryl-tRNA^{Sec} into selenocysteyl-tRNA^{Sec} using selenophosphate as the selenium donor compound. A crystal structure of the 450-residue core of mouse SecS was determined at 1.65 Å. The fragment retains full SecS enzymatic activity. The protein is a member of the fold type I family of pyridoxal phosphate (PLP)-dependent enzymes. It is organized in three domains, which exhibit a number of features not found in other members of the family. Two

SecS monomers interact intimately and together build up two identical active sites around a PLP cofactor. The protomers reciprocally exchange a long SecS-specific insertion in the first domain, which remodels the active site compared to other fold type I enzymes. Two SecS dimers further associate to form a homotetramer via an N terminal region, which is unique to SecS orthologs. A detailed reaction mechanism for the enzyme based on activity profiles of mechanism-based inhibitors was formulated.

tRNA^{Sec} is a key molecule in Sec biosynthesis and incorporation that participates throughout all steps of whole pathways. Rational mutagenesis was applied to murine tRNA^{Sec} in order to obtain well diffracting crystals. Crystal structures of tRNA^{Sec} lacking nucleotides 72-76 of the acceptor stem were determined up to 2.0 Å resolution. The structures are globally identical to that of human tRNA^{Sec} published earlier. Our structure revealed details not seen before in the structure of human tRNA^{Sec} at 3.1 Å resolution. Structural alignment of the available tRNA^{Sec} molecules revealed that the variable arm and the part of the anticodon stem-loop are flexible regions, whereas the molecule core is rigid. Flexible parts of tRNA^{Sec} possibly play important roles in establishing contacts with interacting proteins. The anticodon loops of two tRNA^{Sec} molecules in the asymmetric unit adopted 5- and 3-nucleotide loop conformations. We suggested that fully modified tRNA^{Sec} favors a 5-nucleotide anticodon loop conformation based on modeling of 2'-*O*-hydroxymethylated ribose at U34. Water molecules in the first hydration shell were located in contact both with functional groups of the bases and with the backbone. Interestingly, guanosines are hydrated more than other tRNA^{Sec} bases which could reflect its more exclusive decoration chemical functionalities. Crystal soaking experiments of tRNA^{Sec} with Mn²⁺ provided evidence for five putative Mg²⁺ binding sites in the molecule. All metals were found in contact with N7 atoms of guanosines. However, the metal ions did not specifically stabilize the 3D structure of tRNA^{Sec}.

2 INTRODUCTION

2.1 Selenium

Selenium (Se) was discovered in 1817 by the Swedish chemist Berzelius. The similarity to Tellurium (meaning “earth”) expression allows him to name the new substance “Selenium” (from the Greek Σελήνη, “moon” (goddess)) (Weeks, 1932; Arnér, 2010). Selenium is poisonous to humans and animals in large amounts, but at the same time selenium is an essential trace element for mammals (Wilber, 1980; Schwarz and Foltz, 1957). Later it was shown that selenium is present in glutathione peroxidase enzyme as part of the amino acid selenocysteine (Sec) (Flohe *et al.*, 1973; Forstrom *et al.*, 1978). Today, Sec is known to be part of the active sites of a number of selenoproteins (25 in human), most of which carry out redox functions. Biochemical studies showed that enzymes containing Sec in the active site can be 10 to 100 times more active compared to their cysteine analogs (Lee *et al.*, 2000; Zhong and Holmgren, 2000). The higher activity of enzymes containing Sec compared to Cys can be explained by the higher reactivity of selenium due to its higher nucleophilicity and the lower pK_a of the selenolate compared to the thiolate group (Metanis *et al.*, 2006; Huber and Criddle 1967).

Selenocysteine differs from those of the other 20 proteinogenic amino acid by its biosynthesis and incorporation during the translational process in response to redefined UGA stop codon in mRNA (Böck and Stadtman, 1988; Allmang *et al.*, 2009). Organisms which contain selenoproteins are frequently encountered in all three domains of life: prokaryotes, archaea and eukaryotes. However plants and yeast are not able to incorporate Sec during translation at UGA Sec-codons. Nowadays total of 25 genes encoding selenoproteins in human and 24 in mouse were identified (Kryukov *et al.*, 2003).

2.2 Selenocysteine biosynthesis and incorporation in prokaryotes

Large contribution in Sec biosynthesis and incorporation mechanism understanding in bacteria were done by Bock and coworkers. Analysis of *E.coli* mutants which were unable to synthesize selenoproteins both formate dehydrogenase (FDH) N and FDH H led to the identification of four essential genes named *selA*, *selB*, *selC* and *selD* (Leinfelder *et al.*, 1988). The products of those genes were subsequently characterized and their roles in selenoprotein biosynthesis were established (Leinfelder *et al.*, 1988, 1990; Forchhammer *et al.*, 1989, 1991). Selenocysteine formation depends on the products of the *selA* and *selD* genes which

are, respectively, selenocysteine synthase and selenophosphate synthetase (SelD) (Leinfelder *et al.*, 1989). In addition, tRNA^{Sec} and SerRS are involved in the biosynthesis of Sec in bacteria (Gursinsky *et al.*, 2000). Sec synthesis takes place on its cognate tRNA^{Sec}, which is initially aminoacylated with serine by SerRS (Leinfelder *et al.*, 1988). Then the PLP-dependent enzyme, SelA, uses selenophosphate (SeP) to directly convert Ser-tRNA^{Sec} into Sec-tRNA^{Sec} (Forchhammer *et al.*, 1991). SeP is provided by SelD which synthesizes it from reduced selenium and ATP by transferring of the γ -phosphate residue of ATP to selenide concomitantly releasing AMP and inorganic phosphate (Leinfelder *et al.*, 1990; Ehrenreich *et al.*, 1992).

Selenocysteine incorporation in bacteria requires in addition to the Sec-tRNA^{Sec}, SECIS element and elongation factor SelB. The selenocysteyl-tRNA^{Sec} for Sec insertion at UGA codon in bacteria is recognized by a special elongation factor, SelB, which is escorting the aminoacylated Sec-tRNA^{Sec} to the ribosome (Forchhammer *et al.*, 1989). The N-terminal part of the SelB protein consists of domains, which are similar to EF-Tu. The C-terminal part, unlike EF-Tu, includes domains responsible for the second protein function, possessing a high binding specificity for a SECIS element - stem-loop structure located within the selenoprotein mRNAs coding region and directly downstream of a UGA Sec-codon (Heider *et al.*, 1992; Kromayer *et al.*, 1996). Such stem-loop structures function as Sec insertion elements that recode UGA, which normally is a translational termination signal. In other words, the SECIS element serves as a *cis*-acting element which recognizes a *trans*-acting factors (SelB in complex with Sec-tRNA^{Sec}) and directs them to the ribosome. Although SECISs have low sequence conservation they have a common secondary structure. The minimal requirement for Sec incorporation is a 17 nt-stem-loop structure located at a proper distance (11 nt) from the UGA Sec-codon (Liu *et al.*, 1998). Further studies revealed that the RNA backbone and two nucleotides of the SECIS loop are involved in the interaction with the C-terminal domains of SelB (Yoshizawa *et al.*, 2005; Ose *et al.*, 2007). The quaternary complex comprising SelB, SECIS, Sec-tRNA^{Sec} and GTP translocates towards the ribosome during translation. When the UGA reaches the A site, the lower helical part of the SECIS element is melted, SelB establishes contact with the ribosome and the charged Sec-tRNA^{Sec} after GTP hydrolysis is delivered in the A-site and released from the quaternary complex. After translation through the SECIS, the SECIS element refolds and can assist new cycles of decoding by the following ribosome (Böck, 2001). The process of selenocysteine biosynthesis and incorporation in bacteria is shown schematically (Figure 1).

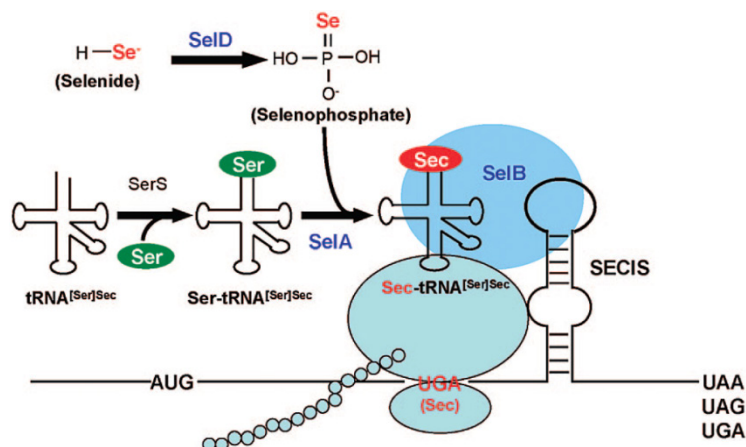


Figure 1. Sec incorporation in bacteria. Specific proteins involved in Sec biosynthesis and insertion into selenoprotein, are highlighted in blue (adopted from Zhang and Gladyshev, 2009)

2.3 Selenocysteine biosynthesis and incorporation in eukaryotes

Although globally Sec biosynthesis and decoding are common for both bacteria and eukaryotes they have clear differences (Böck *et al.*, 1991; Hatfield *et al.*, 2006). In comparison to bacteria, eukaryotes and archaea evolved a more complicated selenocysteine biosynthesis process which requires an additional step (Xu *et al.*, 2007a; Yuan *et al.*, 2006; Rother *et al.*, 2001). The identification of all players and the elucidation of the mechanism of selenocysteine synthesis took more than 10 years after the process in bacteria had been established. As in case of prokaryotes, Sec synthesis takes place on its cognate tRNA^{Sec} (Lee *et al.*, 1989). Again, in the first step tRNA^{Sec} is aminoacylated by SerRS with serine (Ohama *et al.*, 1994). Afterwards, Ser-tRNA^{Sec} is first phosphorylated by PSTK to produce P-Ser-tRNA^{Sec} (Carlson *et al.*, 2004). Although P-Ser-tRNA^{Sec} was known since the 1970s, its biological role remained unknown for a long time (Mäenpää *et al.*, 1970). One of the reasons was the suggestion by Mizutani and colleagues that Sec-tRNA^{Sec} in eukaryotes is directly synthesized from seryl-tRNA^{Sec} as in bacteria (Mizutani *et al.*, 1991). Using a comparative genomics approach, an ORF encoding PSTK was identified and the enzyme was characterized (Carlson *et al.*, 2004).

The selenium substrate for the Sec biosynthesis is supplied by selenophosphate synthetase (SPS). Two enzymes SPS1 and SPS2 were identified in eukaryotes. SPS1 was identified in human and originally thought to have a role in selenophosphate synthesis (Low *et al.*, 1995). Afterwards SPS2, which contains selenium in the active site, was identified in human, mouse and archaea (Guimarães *et al.*, 1996). Complementation studies of SelD⁻ *E. coli* cells with SPS1 or SPS2 showed that both proteins are able to restore selenoprotein

synthesis (Tamura *et al.*, 2004). But the results suggested that SPS2 was involved in the *de novo* synthesis of selenophosphate, whereas SPS1 may have a role in recycling selenocysteine by selenium salvage system (Tamura *et al.*, 2004). Later was demonstrated that only SPS2 synthesizes selenophosphate *in vitro* (Xu *et al.*, 2007a). Furthermore, knockout of the genes encoding SPS1 or SPS2 showed that indeed SPS2 is essential for biosynthesis of selenoproteins (Xu *et al.*, 2007b). The presence of SPS1 in organisms, which do not have selenoproteins, suggests a non-essential role in selenoprotein metabolism or an additional function in unrelated biochemical processes (Lobanov *et al.*, 2008; Xu *et al.*, 2007b). The last biochemical step of conversion P-Ser-tRNA^{Sec} to Sec-tRNA^{Sec} carried out by SecS using a selenophosphate as a selenium donor (Xu *et al.*, 2007a; Yuan *et al.*, 2006). More details about SecS protein are given in chapter 2.5.

Sec insertion in eukaryotes is also more complex than in bacteria and is orchestrated by a number of factors. The functions of bacterial SelB are distributed to two proteins: elongation factor EFSec which specifically recognizes Sec-tRNA^{Sec} and delivers it to the ribosome and SBP2 which binds SECIS element (Copeland *et al.*, 2000). Eukaryotic Sec insertion sequence elements in contrast to bacterial are located in the 3'-untranslated regions of selenoprotein mRNAs (Berry *et al.*, 1993). This positioning of the SECIS element allows not only to insert more than one Sec into a single polypeptide chain but also permits full flexibility in the sequence composition after the UGA codons. The secondary structures of SECIS elements in eukaryotes more complex than those of the bacterial counterparts and include a K-turn motif, which is an identity element for SBP2 (Cléry *et al.*, 2007). In addition in eukaryotes additional factors are involved in the UGA decoding process. Ribosomal protein L30 and initiation factor 4a3 compete with SBP2 for binding to SECIS elements (Chavatte *et al.*, 2005; Budiman *et al.*, 2009).

Currently, two different possible models of Sec incorporation in eukaryotes are discussed in the literature (Caban *et al.*, 2006; Chavatte *et al.*, 2005). The first Sec incorporation model (Figure 2A) is based on the notion that SBP2 could form a complex with the ribosome (Kinzy *et al.*, 2005). Such SBP2 pre-bound ribosomes stall upon arrival at a UGA codon, which would allow for a distant SECIS element to interact with SBP2. This interaction may trigger a conformational changes in the A-site of the ribosome, which could favor the binding of the eEFSec/GTP/Sec-tRNA^{Sec} complex and the subsequent insertion of Sec. After the incorporation event, L30 could displace SBP2 from the SECIS element and SBP2 could take its original position on the ribosome (Caban *et al.*, 2006).

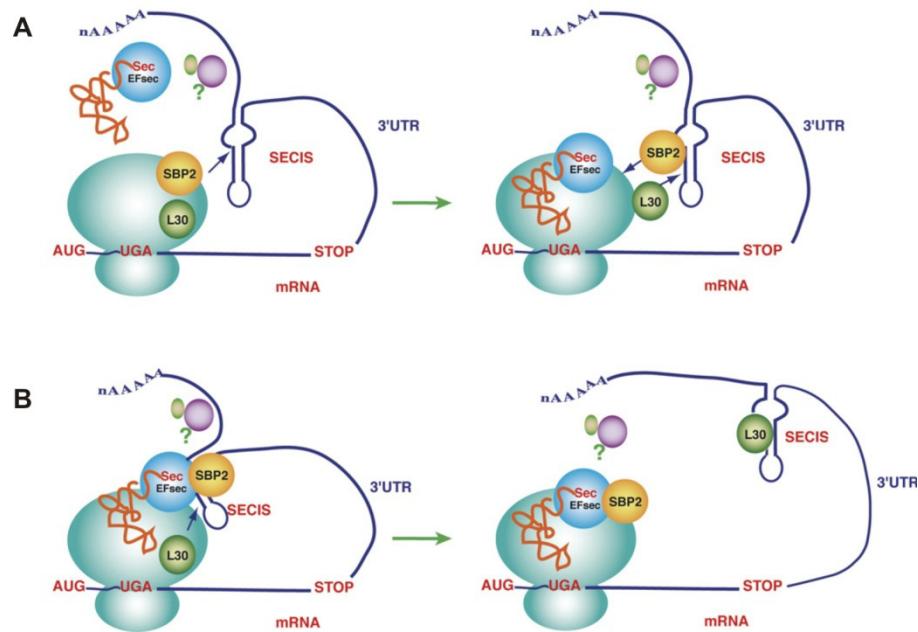


Figure 2. The selenocysteine incorporation models in eukaryotes. (A) SBP2 attached to the ribosome, interacts with the SECIS RNA and the EFSec/Sec-tRNA^{Sec} and delivers this complex to the A site of the ribosome (Kinzy *et al.*, 2005). L30 displaces the SECIS-bound SBP2. (B) The EFSec/Sec-tRNA^{Sec} complex is recruited at the SECIS RNA by SBP2. Ribosome-bound L30 displaces SBP2 (Chavatte *et al.*, 2005). In both models, L30 must leave the SECIS RNA to reset the system. Black arrows indicate factor reshuffling; as yet unidentified factors, possibly involved in the mechanism, are indicated with the question mark (adopted from Allmang and Krol, 2006)

The second model (Figure 2B) proposes that SBP2 in complex with a SECIS element recruits the eEFSec/GTP/Sec-tRNA^{Sec} complex. The association with the ribosome leads to exchange of SBP2 with L30 and as a result a kinked SECIS element is formed. This conformational change in the SECIS element is suggested to instigate the release of Sec-tRNA^{Sec} and GTP hydrolysis (Chavatte *et al.*, 2005).

Clearly, both models are still incomplete, as they not include the role of other factors involved in Sec incorporation process. For instance, SECp43 was found to associate with the EFsec/Sec-tRNA^{Sec} complex *in vitro* and enhance the interaction between EFSec and SBP2 *in vivo* (Small-Howard *et al.*, 2006). SECp43 is also important for the specific post-transcriptional modification of tRNA^{Sec} (Xu *et al.*, 2005). Furthermore, two independent groups found a stem-loop structure in a subset of selenoprotein mRNAs at a conserved distance (7 nucleotides) 3' of the UGA Sec-codon (Howard *et al.*, 2005; Pedersen *et al.*, 2006). These new RNA motifs were named selenocysteine codon redefinition elements (SRE) (Howard *et al.*, 2005). Biochemical studies revealed that such SRE motifs stimulate decoding of a UGA codon with Sec (Howard *et al.*, 2007)

In the following chapters I will discuss SelB, SecS and tRNA^{Sec} which were studied during Ph.D. projects in more detail.

2.4 Selenocysteine specific elongation factor

Historically selenocysteine incorporation machinery was first studied in prokaryotes. Bacterial SelB was identified and characterized by Böck's group (Forchhammer *et al.*, 1989). Is a multifunctional protein which recognizes a bacterial SECIS elements and Sec-tRNA^{Sec} (Baron *et al.*, 1993; Ringquist *et al.*, 1994). The N-terminal part (domains I–III) of the protein shows homology to elongation factor EF-Tu and binds GTP (Figure 3) (Forchhammer *et al.*, 1989; Kromayer *et al.*, 1996). In addition, these domains specifically recognize the specialized tRNA^{Sec} charged with Sec but neither the serylated nor the uncharged tRNA^{Sec} precursors (Kromayer *et al.*, 1996; Leibundgut *et al.*, 2005; Paleskava *et al.*, 2010).

For the second SelB function - binding to SECIS element is responsible the C-terminal extension which has no homologous part in EF-Tu and comprises of four winged-helix (WH) motifs, usually found in DNA binding proteins (Kromayer *et al.*, 1996; Gajiwala and Burley, 2000). The domain pairs WH1/WH2 and WH3/WH4 are linearly arranged and two pairs of domains adopt an overall "L" shape (Selmer and Su, 2002). Binding to SECIS element increases the concentration of Sec-tRNA^{Sec} on the selenoprotein mRNA in close proximity to the UGA Sec-codon.

Minimal fragment of SelB which binds to a SECIS element with high-affinity is represented by domains WH3/WH4 (Kromayer *et al.*, 1996; Yoshizawa *et al.*, 2005). The crystal structure analysis of the *Moorella thermoacetica* (*mth*) SelB WH3/WH4 in a complex with SECIS element uncovered the molecular basis of RNA recognition by the WH domains (Yoshizawa *et al.*, 2005). It was not clear whether the first two WH domains carried out a particular function apart from constituting a spacer between the terminal functional ends of SelB. A number of accumulated facts have suggested that the N-terminal EF-Tu-like domains and the SECIS-binding WH3/WH4 domains functionally interact with each other. For example, binding of Sec-tRNA^{Sec} to the SelB N-terminal domains increases its affinity to a SECIS element (Baron *et al.*, 1993; Thanbichler *et al.*, 2000). Moreover, interaction with a SECIS element initiates the SelB GTPase activity (Hüttenhofer and Böck, 1998). Little structural change was found in the SECIS-binding WH3/WH4 domains upon interaction with the mRNA, suggesting that they may not directly communicate SECIS binding to the N-terminal portion of SelB (Yoshizawa *et al.*, 2005). It has been proposed that the WH3/WH4

domains are flexibly hinged to the WH1/WH2 domains, which in turn may be flexibly connected with the EF-Tu homology portion. Therefore communication between the terminal portions could take place via these flexible connections (Selmer and Su, 2002).

The WH2-WH3 interdomain region forms a positively charged area that binds to the phosphate backbone of a neighboring SECIS. Therefore such recognition in the crystals may mimic a possible interaction with rRNA (helix 16 or 33) during selenocysteine incorporation (Ose *et al.*, 2007). However the structure of the quaternary SelB/GTP/Sec-tRNA^{Sec}/SECIS complex with the ribosome is not available and therefore molecular mechanism underlying functional cross-communication between them is not clear.

The SelB function in eukaryotes is carried out by EFSec, which also has two functional parts where the N-terminal part is homologous to eukaryotic elongation factor EF1-A (Fagegaltier *et al.*, 2000) (Figure 3). In contrast to SelB the C-terminal extension of EFSec is shorter and does not show any similarity with the SelB winged-helix motifs (Fagegaltier *et al.*, 2000). Moreover EFSec unable to bind the cognate SECIS element and therefore C-terminal part is dedicated for another function (Fagegaltier *et al.*, 2000; Tujebajeva *et al.*, 2000). The role of the C-terminal extension of EFSec was clarified by co-immunoprecipitation experiments. It was found that binding of EFSec with SBP2 is RNA dependent (Tujebajeva *et al.*, 2000). Further investigation revealed that tRNA^{Sec} is required for complex formation between EFSec and SBP2 and binding of tRNA^{Sec} to the elongation factor stabilizes its C-terminal domain (Zavacki *et al.*, 2003).

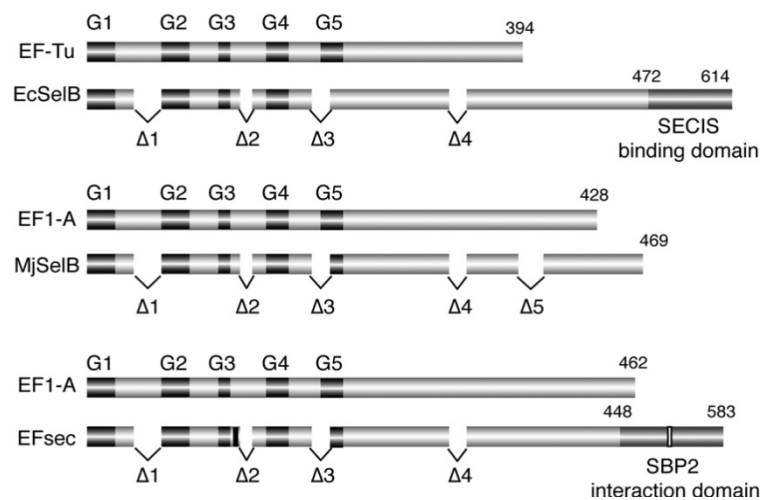


Figure 3. Schematic comparison of functional modules of Sec-tRNA^{Sec}-specific elongation factors from *E. coli* (EcSelB), *Methanococcus jannaschii* (MjSelB) and eukarya (EFsec) with EF-Tu or EF1-A. The C-terminal extensions of EcSelB and EFSec possess binding activity to the SECIS element or SBP2 respectively. The function of MjSelB C-terminal extension has not been assigned yet. The GTP binding domains are depicted (G1-G5); Δ1-Δ5 are the deletion regions relative to EFTu/ EF1-A (adopted from Kroll and Allmang, 2006).

Crystal structure of entire archaeal SelB revealed interesting features of the N-terminal part organization. Overall structure of SelB from *M. maripaludis* resembles the “chalice” observed for initiation factor IF2/eIF5B (Leibundgut *et al.*, 2005). Based on the functional studies of mutants, Leibundgut and colleagues show that two positively charged amino acids (arginine and histidine) in the aminoacyl-binding pocket of SelB are important for specific binding of Sec-tRNA^{Sec} and could interact with negatively charged selenium, thereby stabilizing the reactive selenol group (Leibundgut *et al.*, 2005). Another important amino acid is phenylalanine. It originates from domain I and protrudes into the aminoacyl-binding pocket of domain II. This hydrophobic residue is very flexible and may shield from the environment highly reactive selenium of Sec (Leibundgut *et al.*, 2005).

Other interesting findings come from tRNA modeling with SelB structure results (Leibundgut *et al.*, 2005). Authors suggest that conserved loop in the domain III of SelB, which is also present in bacteria and eukaryotes, may be responsible for specific interactions with tRNA^{Sec} and act as a ruler for measuring the 13 bp long acceptor/T arms (Leibundgut *et al.*, 2005). Interestingly, this loop is much shorter in EF-Tu and together with fact that eukaryotic tRNA^{Sec} can substitute the bacterial counterpart in *E.coli* for selenoprotein biosynthesis indicates a common recognition mechanism for Sec-tRNA^{Sec} between bacteria, archaea and eukaryotes (Nissen *et al.*, 1995, 1999, Baron *et al.*, 1994).

The C-terminal archaeal SelB extension is even shorter than eukaryotic EFSec. Since archaeal SelB does not show binding activity to a SECIS element and a protein with SBP2 function is not found, the role of the C-terminal SelB domain in archaea remains unclear.

2.5 Selenocysteine synthase

Selenocysteine synthase in bacteria was identified in the beginning of 90s (Forchhammer and Böck, 1991). The analysis of conversion Ser-tRNA^{Sec} to Sec-tRNA^{Sec} revealed two step reaction mechanism (Figure 4). At the first step takes place the formation of Schiff base between the serine α -amino group and the pyridoxal 5-phosphate carbonyl with the subsequent 2, 3-elimination of a water molecule and the generation of an aminoacrylyl-tRNA^{Sec} intermediate. Then the nucleophilic addition of selenide (from SeP) to the double bond of aminoacrylyl-moiety forms Sec-tRNA^{Sec} in the second step (Tormay *et al.*, 1998; Forchhammer and Böck 1991).

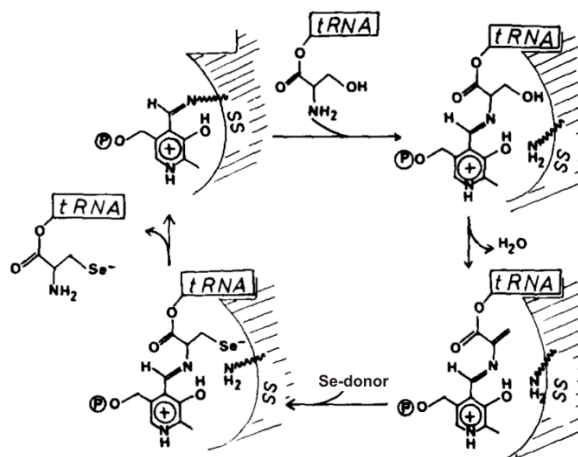


Figure 4. Scheme of Ser-tRNA^{Sec} to Sec-tRNA^{Sec} conversion reaction mechanism by bacterial selenocysteine synthase (modified from Forchhammer and Böck, 1991).

SelA architecture investigation using electron microscopy shows that it consists of 10 monomers yielding 500 kDa complex arranged in two rings with five-fold symmetry. Moreover studying of SelA in complex with tRNA^{Sec} revealed that one tRNA^{Sec} molecule binds per two subunits of enzyme (Engelhardt *et al.*, 1992).

Although understanding of selenocysteine biosynthesis in bacteria showed considerable progress in case of eukaryotes it was slower. For a long time it was unclear which protein displayed selenocysteine synthase function. It was shown that protein product of an archaeal ORF, designated as SecS, did not act on Ser-tRNA^{Sec} or P-Ser-tRNA^{Sec} (Kaiser *et al.*, 2005). The RNA-mediated interference technology provided the first direct evidence for an essential role of soluble liver antigen/liver and pancreas antigen (SLA/LP) in selenoprotein biosynthesis (Xu *et al.*, 2005). The protein was identified in precipitated complex with tRNA^{Sec} by autoantibodies from patients with a severe form of autoimmune chronic active hepatitis (Gelpi *et al.*, 1992; Costa *et al.*, 2000). Structural homology modeling predicted that SLA/LP is a PLP-dependent enzyme of the aspartate aminotransferase family (Kernebeck *et al.*, 2001). Finally direct involvement in conversion of P-Ser-tRNA^{Sec} to Sec-tRNA^{Sec} of SLA/LP was shown independently by the two groups of investigators (Yuan *et al.*, 2006; Xu *et al.*, 2007a). In order to identify candidate for selenocysteine synthase function both groups used different approaches. The first one revealed SecS in mammals using comparative genomics and *in vitro* activity assay (Xu *et al.*, 2007a). The second group proposed that selenocysteine synthase role belongs to SLA/LP based on analogy with cysteine formation pathway in some methanogens which is based on a tRNA-dependent conversion of *O*-phosphoserine to cysteine by PLP-dependent P-Ser-tRNA:Cys-tRNA synthase (SepCysS) (Yuan *et al.*, 2006). Its role was successfully confirmed by complementation studies in *E. coli*

(Yuan *et al.*, 2006; Sauerwald *et al.*, 2005). The identification of SecS both in eukaryotes and archaea shed light on the molecular mechanisms of catalytic activity. It was shown that SecS dephosphorylates P-Ser-tRNA^{Sec}, which indicated that aminoacrylyl-tRNA^{Sec} is a likely intermediate in the reaction (Xu *et al.*, 2007a). Furthermore was found that SeP produced by SPS2 serves as activated selenium donor for SecS (Yuan *et al.*, 2006; Xu *et al.*, 2007a) (Figure 5).

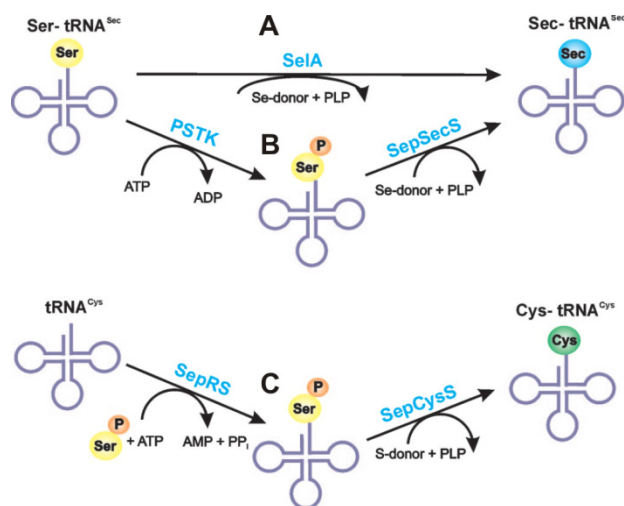


Figure 5. Schematic representation of tRNA-dependent amino acid transformations. Selenocysteine biosynthesis: (A) pathway in bacteria and (B) archaea/eukaryotes. (C) Cysteine biosynthesis pathway in methanogens (adopted from Yuan *et al.*, 2006).

Multiple sequence alignment and phylogenetic analysis of PLP-dependent proteins (Figure 6) revealed that Sela and archaeal Sela-like proteins belonging to the completely different families of PLP-containing enzymes compared to SecS (Kaiser *et al.*, 2005). Moreover this data suggest that their similar function is a result of convergent evolution. On other hand SecS is distantly homologous to SepCysS (Xu *et al.*, 2007a).

Although it was a breakthrough in understanding of selenocysteine biosynthesis the molecular mechanism of P-Ser-tRNA^{Sec} to Sec-tRNA^{Sec} by SecS conversion was not known. Later two independent groups unraveled molecular bases of the last Sec step maturation by analysis of the murine and archaeal SecS crystal structures (Araiso *et al.*, 2007, Ganichkin *et al.*, 2007). Recently analysis of the co-crystal structure of SecS and tRNA^{Sec} shed light on intimate interaction and recognition process of tRNA^{Sec} by SecS. In addition details of P-Ser-tRNA^{Sec} to Sec-tRNA^{Sec} conversion were refined (Palioura *et al.*, 2009).

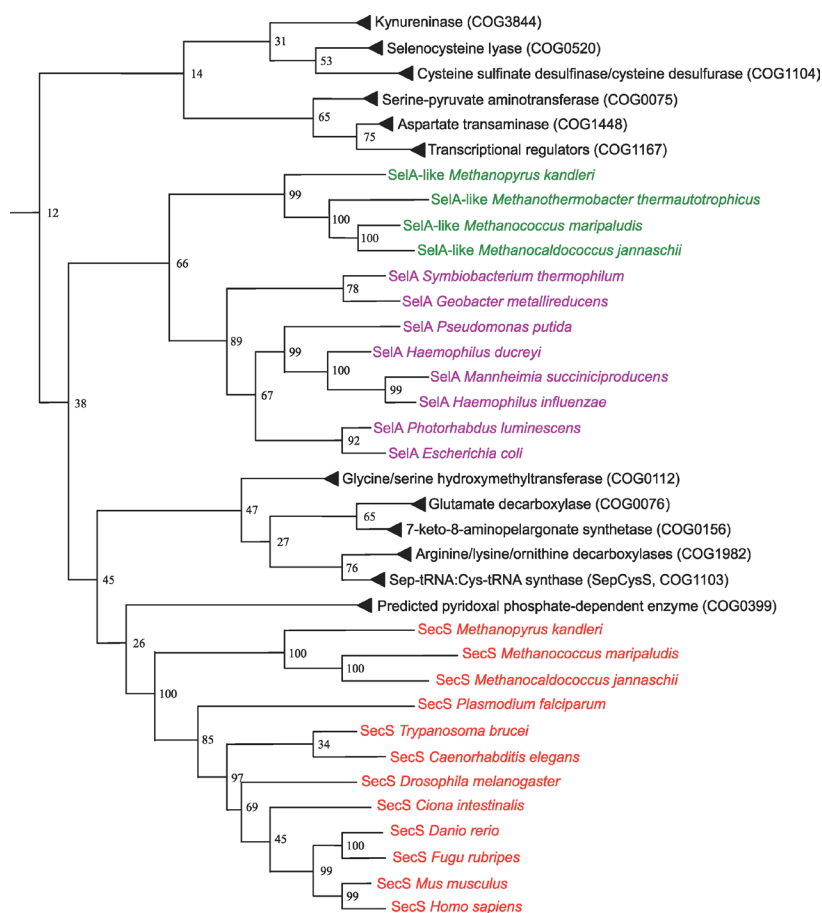


Figure 6. Phylogenetic tree of SecS, SelA, SelA-Like, and other PLP-dependent proteins (adopted from Xu *et al.*, 2007a).

2.6 tRNA^{Sec}

Suppressor tRNA^{Ser} was initially found as fraction of tRNA^{Ser} in mammals in the beginning of 70s (Hatfield and Portugal, 1970). Its involvement in selenocysteine incorporation in response to UGA codon was proved only in the late 80s (Leinfelder *et al.*, 1988; Lee *et al.*, 1989). Understanding that tRNA^{Sec} is only one factor which participates throughout whole process of Sec biosynthesis and its cotranslational incorporation makes it a key molecule in the biosynthetic process (Böck *et al.*, 1991). For interactions with other factors involved in Sec pathway tRNA^{Sec} evolved structural features distinct from canonical elongator tRNAs which discussed below.

Based on structural analysis with the use of enzymatic and chemical probes the bacterial tRNA^{Sec} model containing the 8-bp acceptor stem and the 5-bp T-stem (8/5 model) was proposed (Baron *et al.*, 1993). Coaxial staking of both elements give rise to a 13-bp stem, which is one base pair longer than canonical tRNAs with the 7/5 arrangement of the same elements. Such 13-bp stem of tRNA^{Sec} is identity element for SelB. The shortening of

tRNA^{Sec} acceptor stem to 7 base pairs cause loss binding activity between SelB and tRNA^{Sec} (Baron *et al.*, 1993).

Apart of the acceptor/T-stem organization tRNA^{Sec} has many other unique properties. For instance, the 6-bp D-stem (whereas in other tRNAs is 3 or 4) and long variable arm (Baron *et al.*, 1993) Presence of 13-bp stem formed by coaxial staking of acceptor and T-stem, and the unusually long variable arm makes tRNA^{Sec} the longest among elongator tRNA species known so far. In bacteria, tRNA^{Sec} is 95 nt long compared with canonical ones with average length of 76 nt (Baron *et al.*, 1993; Jühling *et al.*, 2009).

The tRNA^{Sec} solution structure of eukaryotes was debated for a long time and two models were proposed (Figure 8). Based on solution structure probing the tRNA^{Sec} model with the 9-bp acceptor and the 4-bp T-stem (9/4 model) was suggested (Hubert *et al.*, 1998; Sturchler *et al.*, 1993). The second model based on tRNA^{Sec} mutants studies and computer modeling proposed that length of the acceptor stem is a 7-bp and the T-stem is a 5-bp (7/5 model) like in canonical tRNAs (Ioudovitch *et al.*, 1998; Diamond *et al.*, 1981).

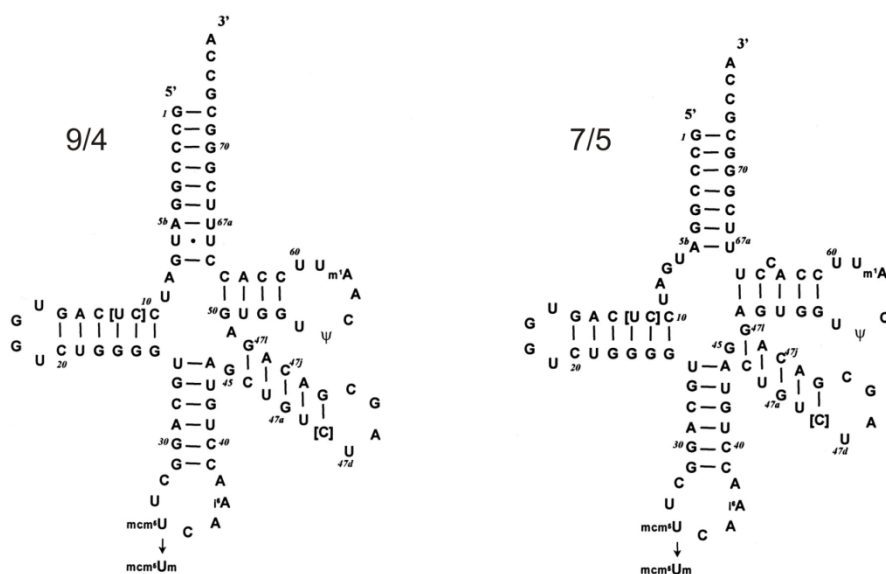


Figure 8. Mammalian tRNA^{Sec} shown in a 9/4 and a 7/5 cloverleaf model (adopted from Hatfield and Gladyshev, 2002).

The length of the 13-bp acceptor/T-stem helix the in eukaryotic tRNA^{Sec} is important for Ser to Sec conversion (Baron *et al.*, 1991; Sturchler-Pierrat *et al.*, 1995). The length of eukaryotic tRNA^{Sec} is 90 nt long which is still considerably longer than canonical tRNAs (Diamond *et al.*, 1981; Jühling *et al.*, 2009). As in case of bacteria eukaryotic tRNA^{Sec} possess the 6-bp D-stem and long variable loop (Wu and Gross, 1994). Interestingly, also the structures of archaeal and eukaryotic tRNA^{Sec} are very similar they have different identity elements for PSTK (Hubert *et al.*, 1998; Sturchler *et al.*, 1993). Kinase from HeLa S100 cell extracts

recognizes the D-stem length and secondary structure whereas archaeal PSTK recognizes the acceptor stem (Wu and Gross, 1994; Sherrer *et al.*, 2008).

Recently, the two eukaryotic tRNA^{Sec} structures become available. tRNA^{Sec} was crystallized alone and in complex with SecS at resolution 3.1 Å and 2.8 Å respectively (Itoh *et al.*, 2009; Palioura *et al.*, 2009). Both tRNA^{Sec} structures confirmed biochemical data in favor of 9/4 model (Figure 9). The D-arm included the 6 bp stem capped with 4 nt loop. The variable and D-arms do not form tertiary interactions leading to an open cavity formation in place of the tertiary core presented in canonical tRNAs (Figure 7) (Itoh *et al.*, 2009; Palioura *et al.*, 2009). Moreover the 13-bp acceptor-T-stem and the variable arm are structural features that serve as major recognition motifs for binding to SecS. Modeling studies suggest that SerRS cannot discriminate between both tRNA^{Ser} and tRNA^{Sec} as it recognizes their variable arms in the same way (Palioura *et al.*, 2009).

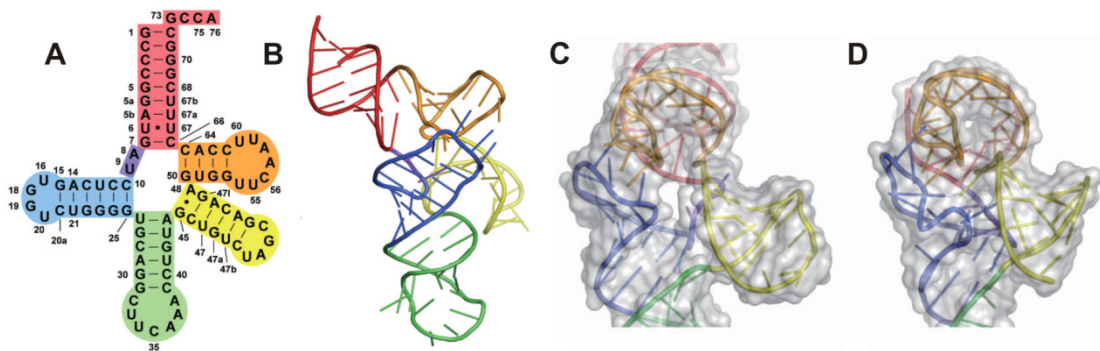


Figure 9. Structure of human tRNA^{Sec}. (A) cloverleaf model, (B) ribbon model [PDB ID: 3A3A (Itoh *et al.*, 2009)], (C) the D, T and extra arms of human tRNA^{Sec} formed open cavity (D) *T. thermophilus* tRNA^{Ser} (PDB ID: 1SER; Biou *et al.*, 1994) instead has tertiary interactions. Elements of tRNA^{Sec} on (C) and (D) shown with their surface models. The acceptor arm, AD linker, D-arm, anticodon arm, extra arm and T arm are colored red, purple, blue, green, yellow and orange, respectively (Figure modified from Itoh *et al.*, 2009).

tRNA^{Sec} has only a few modified nucleotides: pseudo-U55 (ψ) and N1-methyladenosine (m^1A58) in the T loop, 6-isopentenyl-A37 (i^6A37) and 5-methylcarboxymethyl-2'-O-methyluridine (mcm^5Um34) modification in the anticodon loop (Diamond *et al.*, 1993; Sturchler *et al.*, 1994). The eukaryotic tRNA^{Sec} has two isoforms with differences in U modifications at wobble position (Diamond *et al.*, 1993). The first contains 5-methylcarboxymethyluridine (mcm^5U) and the second one 2'-O-hydroxymethylated ribose (mcm^5Um). The modification mcm^5Um is unique and found only in tRNA^{Sec} so far (Hatfield *et al.*, 1991; Diamond *et al.*, 1993). The state of U34 modification depends on selenium diet. Cells or animals which were fed with adequate or enriched level of selenium contain higher level of mcm^5Um -containing Sec-tRNA^{Sec} compare to mcm^5U -containing Sec-tRNA^{Sec}. On the other hand lowering selenium level in diet followed redistribution of two tRNA^{Sec}

populations with prevailing of mcm^5U -containing Sec-tRNA^{Sec} compared to mcm^5Um -containing Sec-tRNA^{Sec} (Hatfield *et al.*, 1991). Modification mcm^5U to mcm^5Um is a highly specialized step which is depended on the correct primary, secondary and tertiary tRNA^{Sec} structure and probably it is the last step in maturation of tRNA^{Sec} (Kim *et al.*, 2000; Choi *et al.*, 1994). Recent studies of transgenic mutant tRNA^{Sec} mouse models demonstrate that mcm^5U -containing tRNA^{Sec} isoform responsible for synthesis of housekeeping selenoproteins whereas tRNA^{Sec} isoform with mcm^5Um involved in biosynthesis of stress-related selenoproteins (Carlson *et al.*, 2009). One of the candidates for the methylation of 2'-hydroxyribosyl moiety of tRNA^{Sec} could be the SECp43 found in complex with tRNA^{Sec} and SecS (Ding *et al.*, 1999; Xu *et al.*, 2005). RNA-mediated interference technology shows that knockdown of gene coded SECp43 leads to tRNA^{Sec} methylation level decrease (Xu *et al.*, 2005).

Apart of the unusual structural features which is described above quite interesting a transcription of tRNA^{Sec} gene. Normally transcription of elongator tRNA genes in eukaryotes initiates by RNA polymerase III from the sequence flanked at gene 5'-end sequence and terminates on cluster of thymidine nucleotides after the gene on 3'-end. After transcription 5'-leader sequence together with 3'-trailer sequence has to be removed yielding mature tRNA. As in case of elongator tRNAs, tRNA^{Sec} also has 3'-trailer sequence which is processed after transcription. However, tRNA^{Sec} does not have 5'-leader sequence and transcription starts directly from the first nucleotide of the gene. Therefore matured tRNA^{Sec} has 5'-triphosphate on its terminal residue (Lee *et al.*, 1987).

2.7 Aims of study

At the beginning of this Ph.D. project, little was known about the structural basis of selenocysteine biosynthesis and incorporation. In order to partially fill this gap I addressed the following specific aims during my Ph.D.:

1. What is the mechanism underlying interaction the winged-helix motifs 1 and 2 of SelB with helix 16 of 16S rRNA?
2. How does mammalian selenocysteine synthase convert P-Ser-tRNA^{Sec} to Sec-tRNA^{Sec}?
3. What is the structural organization of the unusual tRNA^{Sec}?

3 MATERIALS AND METHODS

3.1 Materials

3.1.1 Chemicals

Acrylamide	Roth, Karlsruhe
Rotiphorese Gel 30 (Acrylamide:Bis-Acrylamide 37.5:1)	
Rotiphorese Gel 40 (Acrylamide:Bis-Acrylamide 19:1)	
Rotiphorese Gel 30 (Acrylamide:Bis-Acrylamide 29:1)	
Agarose	Invitrogen, Karlsruhe
Ammoniumperoxodisulfate (APS)	Merck, Darmstadt
Bovine serum albumin (BSA)	Sigma, Taufkirchen
Bradford-protein staining solution	Bio-Rad, Muenchen
Bromphenolblue	Merck, Darmstadt
Coomassie Brilliant Blue R/G250	Roth, Karlsruhe
DNA-molecular weight marker (V, III)	Bioline, Luckenwalde
Dithiothreitol (DTT)	Roth, Karlsruhe
Ethylenediamine-N,N,N',N'-tetraacetic acid (EDTA)	Roth, Karlsruhe
Ethidiumbromide solution (10 mg/ml)	Boehringer, Mannheim
Gel filtration standard	BioRad, Muenchen
Glutathione, reduced	Sigma-Aldrich, Steinheim
Glycerol	Merck, Darmstadt
Imidazole	Merck, Darmstadt
Isopropyl β -D-1-thiogalactopyranoside (IPTG)	Sigma-Aldrich, Steinheim
Magnesium chloride	Merck Eurolab, Hannover
Precision protein standard marker	BioRad, Muenchen
Potassium chloride	Roth, Karlsruhe
Potassium phosphate	Sigma, Deisenhofen
Silver nitrate	Merck Eurolab, Hannover
Sodium chloride	Roth, Karlsruhe
Sodium dodecyl sulphate (SDS)	Serva, Heidelberg
Sypro-orange protein gel stain	Invitrogen, Karlsruhe
N,N,N',N'-Tetramethylenethylendiamin (TEMED)	Sigma, Taufkirchen
Tris-hydroxymethyl-aminomethan (Tris)	Roth, Karlsruhe
Triton X-100	Sigma, Taufkirchen
tRNA <i>E. coli</i>	Boehringer, Mannheim
Yeast extract	Roth, Karlsruhe
Xylencyanol FF	Fluka, Switzerland
4-2-hydroxyethyl-1-piperazineethanesulfonic acid (HEPES)	Calbiochem, USA

Any standard chemicals, organic substances and solvents (purification grade p.a.), which are not listed here, were ordered by one of the following companies: Merck (Darmstadt), Roth (Karlsruhe), Sigma (Taufkirchen), Serva (Heidelberg) or Fluka (Switzerland).

3.1.2 Media and Antibiotics

Antibiotics

Ampicillin	Sigma, Deisenhofen
------------	--------------------

Chloramphenicol Boehringer, Mannheim
 Kanamycin sulphate Sigma, Deisenhofen

Media

Luria-Bertani-broth (LB)-Agar Q-Biogene, USA
 LB-Medium Q-Biogene, USA
 Auto-inducing medium own production

3.1.3 Enzymes and Enzyme Inhibitors

Complete EDTA-freeTM tablets Roche, Mannheim
 Lysozyme Boehringer, Mannheim
 Pefa-Bloc Biomol, Hamburg
 Pfu DNA Polymerase Stratagene, Heidelberg
 Restriction endonucleases NEB, Frankfurt
 RNasin Promega, USA
 TEV-protease own production
 T4 DNA Ligase NEB, Frankfurt
 T4 Polynukleotid Kinase NEB, Frankfurt

Proteases for Limited Proteolysis and Mass Spectrometry

Chymotrypsin Roche, Mannheim
 Endoproteinase Asp-N Sigma-Aldrich, Steinheim
 Endoproteinase Glu-C Sigma-Aldrich, Steinheim
 Trypsin Sigma-Aldrich, Steinheim
 Elastase Sigma-Aldrich, Steinheim

3.1.4 Nucleotides and radionucleotides**Nucleotides**

Desoxynucleotide-5'-Triphosphate (dNTPs, 100 mM) GE healthcare, Freiburg
 Ribonucleotide-5'-Triphosphate (rNTPs, 100 mM) Jena Bioscience, Jena

Radionucleotide

[$\gamma^{32}\text{P}$]-ATP (6000 Ci/mmol, 10 $\mu\text{Ci/ml}$) GE healthcare, Freiburg

3.1.5 Oligonucleotides

Oligodeoxyribonucleotides were purchased from MWG Biotech (Ebersberg) and oligoribonucleotide obtained from Dharmacon (USA). The list of DNA oligos are in the Tables 1, 2 and 3. RNA-Oligonucleotide from *E.coli* helix16 of 16S rRNA was used in this study: 16h: 5'-GUAUGAAGAAGGCCUUCGGGUUGUAAAGUAC-3'

3.1.6 Plasmids

Vector	Promotor	Selection	Fusion tags	Protease site	Source
pET-M11	T7-lac	Kan	N-His, C-His	TEV	EMBL (Heidelberg)
pET-M13	T7-lac	Kan	C-His	none	EMBL (Heidelberg)
pET-M30	T7-lac	Kan	N-His, C-His, N-GST	TEV	EMBL (Heidelberg)

3.1.7 *E. coli* Strains

Strain	Genotype	Antibiotic resistance	Supplier
Rosetta2 (DE3)	F ⁻ <i>ompT hsdSB(r_B⁻m_B⁻)gal dcm</i> (DE3) pRARE2 (Cam ^R)	Cam	Novagen, Germany
XL-1 Blue	<i>endA1 gyrA96(nal^R) thi-1 recA1 relA1 lac glnV44 F' [::Tn10 proAB⁺ lacI^q Δ(lacZ)]M15 Amy Cm^R]</i> <i>hsdR17(r_K⁻ m_K⁺)</i>	Tet	Stratagene, US

3.1.9 Commercial Kits

Qiagen Plasmid Midi/Maxi Kit	Qiagen, Düsseldorf
QIAprep Spin Miniprep Kit	Qiagen, Düsseldorf
GFX Purification Kit	GE healthcare, Freiburg

3.1.10 Crystallization screens

Anions and Cations Suites	Qiagen, Hilden
Additive Screens I-III	Hampton Research, USA
Additive Screens	own production
Classics	Qiagen, Hilden
Crystal Screen I and II	Hampton Research, USA
Index I	Hampton Research, USA
Mb class I and II	Qiagen, Hilden
MPD suite	Qiagen, Hilden
JCSG screen	Qiagen, Hilden
Natrix Screen	Hampton Research, USA
(NH ₄) ₂ SO ₄ screen	Qiagen, Hilden
Nucleix Suite	Qiagen, Hilden
PACT screen	Qiagen, Hilden
PEG I and II	Qiagen, Hilden
pH clear I and II	Qiagen, Hilden
Pre-Crystallization test	Hampton Research, USA
Protein Complex screen	Qiagen, Hilden
SM I, II and III	Qiagen, Hilden
Buffers and reagents which were used for refinement (purification grade p.a.) were purchased at Merck, Sigma-Aldrich or Roth.	

3.1.11 Working Materials

CD-cuvette, 1 mm	Hellma
Dialysis membranes MWCO 3500-8000 Da	SpektraPor, USA
Electroporation cuvettes	Bio-Rad, Muenchen
Glutathione Sepharose 4B	GE healthcare, Freiburg
Ni-NTA agarose	Quiagen, Hilden
Pipettes (adjustable)	Eppendorf, Hamburg
Pipetting aid "pipettus-akku"	Hirschmann, Eberstadt
illustra TM MicroSpin TM G-50 Columns	GE Healthcare, Freiburg
Sterile filters 0.2 μm, 0.45 μm	Millipore, France
X-ray films Biomax MR	Kodak, USA

3.1.12 Technical Devices

Äkta explorer/prime/purifier and columns	GE healthcare, Freiburg
Autoklave, Varioklav 300-EPZ H+P	Labortechnik, Oberschleissheim
Biofuge (pico/fresco)	Heraeus, Hanau
Chirascan, CD-spectrometer	Applied Photophysics, UK
Crychem plates (sitting drop, 24-well)	Hampton Research, USA
Electroporation device, Gene Pulser	Bio-Rad, Muenchen
Eppendorf tubes (0.5, 1.5, 2 ml)	Eppendorf, Hamburg
Falkon tubes (5, 15, 50 ml)	Greiner, Kremsmünster
Image Plate, Mar345	Mar, Norderstedt
Geldocumentation device	Bio-Rad, Muenchen
Gelelectrophoresis device	Bio-Rad, Muenchen; Institutes workshop
Head over tail Rotor (HOT)	Cole-Parmer, USA
Heating block Dri-Block DB-3	Techne, UK
NanoDrop ND-1000	Thermo Fischer, USA
NanoDrop Robot	Cartesian Zinsser Analytic, USA
pH-meter	Mettler, Toledo
Phosphorimager Typhoon 8600	Molecular Dynamics
Rotating anode, Rigaku RU300	MSC, USA
Scintillation counter LS 1701/TRI-CARB 2100TR	Beckman/Packard, USA
Shaking incubator	New Brunswick, USA
SMART-system	GE healthcare, Freiburg
Sonifier (Branson 250D)	Heinemann Labortechnik
Sorvall RC 5B Centrifuge	Kendro, USA
Sorvall SA-300/600, SLC-.6000 rotor	Kendro, USA
Spectrophotometer Ultrospec 3000 pro	Biorad, Muenchen
VDX plates (hanging drop, 24-well)	Hampton Research, USA
Vivaspin, concentrator MWCO 5-100 kDa, 0.5 – 15 ml	Vivascience, Hannover
Vortex	Janke & Kunkel, Staufen
X-ray film developer X-Omat 2000	Kodak, USA
96-well crystallization plates	Greiner, Kremsmuenster
DNA Engine OPTICON™, real-time PCR	MJ Research, USA

3.2. Methods

3.2.1 General molecular biology methods

3.2.1.1 Concentration determination of nucleic acids

For nucleic acid concentration determination, 1 µl of a nucleic acid solution was measured at wavelength 260 nm in comparison to a reference on the NanoDrop ND-1000. The following equations were used to determine concentrations (Sambrook *et al.*, 1989; New England Biolabs catalogue 2006/07, USA):

$$1OD_{\lambda=260} = 50 \mu\text{g/ml double stranded DNA} = 0.15 \text{ mM (in nucleotides)}$$

$1OD_{\lambda=260} = 33 \mu\text{g/ml}$ single stranded DNA = 0.10 mM (in nucleotides)

$1OD_{\lambda=260} = 40 \mu\text{g/ml}$ single stranded RNA = 0.11 mM (in nucleotides)

3.2.1.2 Phenol-chloroform-isoamylalcohol (PCI) extraction

Phenol-chloroform-isoamylalcohol (PCI) extraction is used to purify and separate proteins and nucleic acids. The samples were mixed with an equal volume of PCI and vortexed for 1-2 min. If sample volume less than 100 μl , it was first adjusted to 100 μl with water or TE-buffer. Aqueous and organic phases were separated by centrifugation for 5 min in a tabletop microfuge (13000 rpm, RT). Upper phase was transferred into a new tube and nucleic acids were precipitated by adding 0.1 vol. of 3 M sodium acetate, pH 5.2 and 2.5 vol. of 100 % ethanol or 0.7 vol. of isopropanol. If the sample contained only traces of nucleic acids, a carrier was added (15 μg of glycogen or yeast tRNA). The samples were incubated for 2-16 hours at -20 °C in case of precipitation with ethanol or 20 min to 2 hours in case of isopropanol. The precipitate was subsequently collected by centrifugation (13000 rpm, 10 min, tabletop microfuge and at 4 or 20 °C depends of alcohol used for precipitation). Supernatant was decanted and the pellet washed twice with 1 ml of 80% (v/v) EtOH with centrifugation after each washing step ((13000 rpm, 5 min, tabletop microfuge, 4 °C). Pellet was dried and dissolved in MQ-grade water or TE-buffer.

3.2.1.3 Agarose gel electrophoresis of nucleic acids

To separate nucleic acids, agarose gel electrophoresis was performed according to standard protocols (Sambrook, *et al.*, 1989). Depending on the size of DNA fragments to be separated, gels were prepared with 0.8 - 2 % (w/v) agarose. Ethidium bromide (EtBr) (final concentration of 0.5 $\mu\text{g/ml}$) was added to the gel solution for visualization of DNA bands under UV light. As a molecular weight standards were used a 1-kb or 3-kb DNA ladder (New England Biolabs) or ladder (Biolane). The nucleic acid samples were 5:1 mixed with the 5x DNA loading dye. Electrophoretic separation was performed horizontally in 1x TBE running buffer at 5-8 V/cm at RT. Analysis and imaging was carried out under UV light with a wavelength of 312 or 365 nm. Short DNA fragments (100bp or shorter) were separated vertically on 12 % PAAG (AA:MBAA 29:1).

Gel solution

1x TBE

0.8-2 % (w/v) agarose

5xDNA loading dye

30 % glycerol (v/v)

0.25 % bromophenol blue (w/v)

0.25 % xylene cyanol FF (w/v)

3.2.1.4 DNA extraction from agarose gels

The samples (PCR products or enzyme-digested DNA samples) were electrophoresed through 0.8 -1 % agarose gel and stained with EtBr. To isolate DNA fragments, a gel was illuminated with UV light at 365 nm and bands of interest were excised with a sterile razor blade. Elution from the gel piece was done with the QIAquick Gel Extraction Kit (Qiagen) according to the manufacturer's protocol.

3.2.1.5 Polymerase chain reaction (PCR)

Plasmid or genomic DNA was used as a template for PCR. Templates for *in vitro* transcription and ORFs of proteins were amplified with Pfu-DNA polymerase. Forward and reverse primers (Table 1) for amplifying of DNA fragments for molecular cloning were designed to introduce compatible restriction enzyme sites with cloning vector. 3-6 additional bases were added at the 5' ends to allow efficient digestion (New England Biolabs catalogue 2006/07, USA). The annealing temperature was chosen 2-5 °C below of the calculated melting temperature by VNTI program. A typical PCR-protocol is shown below:

Standard 50 µl PCR reaction contained:

template DNA 5-50 ng
10 pmol primer (forward)
10 pmol primer (reverse)
1x PCR buffer
2 µl dNTP mix (2.5 mM each)
2,5 Units Pfu polymerase

Temperature protocol:

step 1 (initial denaturation) 2.5 min 94 °C
step 2 (denaturation) 30 sec 94 °C
step 3 (primer annealing) 30 sec 56 °C
step 4 (primer extension) 1-3 min 72 °C
step 5 (primer extension) 5 min 72 °C
steps 2-4 were repeated 25-35 times

The success of PCR reactions was checked by analytical agarose gel electrophoresis. To purify PCR products from the plasmid DNA template, primers and dNTPs, preparative agarose gel electrophoresis was performed in 1 x TBE buffer. The fragments of choice were isolated from the gel as described in Section 3.2.1.4.

Gene product	Sequence 5'→3'	Restriction enzyme	Vector
<i>mthSelB</i> ³⁷⁷⁻⁵¹¹	F: CATGCCATGGTTCCTCCCGGAAAAAATTC	NcoI	pET-M11
	R: CCGGAATTCAGAAACTGGGAGTAAAGCCG	EcoRI	
<i>mmuSecS</i>	F: TTCGATCGTCTCACATGAACCCGGAGAGCTTCG	BsmBI	pET-M13
	R: GCCAAACTCGAGTAGAGCAGGGCCCTGGCC	XhoI	
<i>mmuPSTK</i>	F: ATCGATCGTCTCACATGAAGACCGCGGCGGCTC	BsmBI	pET-M30
	R: CCGGAATTC TTAATGTTGCTTTGAAAAATACTTCTG	EcoRI	
<i>mmuPSTK</i> ¹⁴³⁻³⁵⁹	F: GTCGATCGTCTCACATGGTTCTAGACCTTTGTTTTGGTGTTAG	BsmBI	pET-M11

	R: CCGGAATTC	TTAATGTTGCTTTGAAAAATACTTCTGC	EcoRI	
<i>mmu</i> PSTK ¹⁴⁷⁻³⁵⁹	F: GTCGATCGTCTC	ACATGTTGTTTTGGTGTTAGATGACAACCTTTT	BsmBI	pET-M11
	R: CCGGAATTC	TTAATGTTGCTTTGAAAAATACTTCTGC	EcoRI	
<i>mmu</i> PSTK ¹⁵⁰⁻³⁵⁹	F: GTCGATCGTCTC	ACATGGTGTTAGATGACAACCTTTATTACCAAAG	BsmBI	pET-M11
	R: CCGGAATTC	TTAATGTTGCTTTGAAAAATACTTCTGC	EcoRI	
<i>mmu</i> PSTK ¹⁵⁸⁻³⁵⁹	F: GTCGATCGTCTC	ACATGCAAAGTATGAGATATGAAGCTACCAACT	BsmBI	pET-M11
	R: CCGGAATTC	TTAATGTTGCTTTGAAAAATACTTCTGC	EcoRI	
<i>mmu</i> PSTK ²²⁵⁻³⁵⁹	F: GTCGATCGTCTC	ACATGCTCATAATTCAGAGTTCAGCATGTTCT	BsmBI	pET-M11
	R: CCGGAATTC	TTAATGTTGCTTTGAAAAATACTTCTGC	EcoRI	
<i>mmu</i> PSTK ²³⁰⁻³⁵⁹	F: GTCGATCGTCTC	ACATGTCAGCATGTTCTCTGGAAGCC	BsmBI	pET-M11
	R: CCGGAATTC	TTAATGTTGCTTTGAAAAATACTTCTGC	EcoRI	
<i>mmu</i> PSTK ²³⁸⁻³⁵⁹	F: GTCGATCGTCTC	ACATGCTGGAGGTGACTGGTTTGTTC	BsmBI	pET-M11
	R: CCGGAATTC	TTAATGTTGCTTTGAAAAATACTTCTGC	EcoRI	
<i>mmu</i> PSTK ²⁵⁶⁻³⁵⁹	F: GTCGATCGTCTC	ACATGGAGGATAATACAGAACAAAAGGAAACTG	BsmBI	pET-M11
	R: CCGGAATTC	TTAATGTTGCTTTGAAAAATACTTCTGC	EcoRI	
<i>mmu</i> PSTK ²⁷⁰⁻³⁵⁹	F: GTCGATCGTCTC	ACATGTCTACTAACATCCTTCATAAAGCTGATGA	BsmBI	pET-M11
	R: CCGGAATTC	TTAATGTTGCTTTGAAAAATACTTCTGC	EcoRI	
<i>mmu</i> PSTK ²⁷²⁻³⁵⁹	F: GTCGATCGTCTC	ACATGAACATCCTTCATAAAGCTGATGAGAC	BsmBI	pET-M11
	R: CCGGAATTC	TTAATGTTGCTTTGAAAAATACTTCTGC	EcoRI	
<i>mmu</i> tRNA ^{Sec}	F: CCGGAATTC	TAATACGACTCACTATAGCC	EcoRI	pUC19
	R: GTAAGGATCC	TGGCGCCCGAAAGGTGG	BamHI	

Table 1. Oligonucleotide primers used for PCR amplification. The encoded gene product names, primers used, restriction sites inserted within the primers (also in red) and the vectors used for targeted cloning are indicated. For each primer pair, the forward primer is denominated F and the reverse R, respectively.

3.2.1.6 Restriction digestion of DNA

Restriction digestion of DNA was used to generate desired ends of the templates for *in vitro* transcription as well as vectors and inserts for subsequent subcloning. Preparative digestion reaction was designed to contain DNA and glycerol at final concentration less than 0.1 µg/µl and 5%, respectively. 2-3 U of restriction enzyme was used per 1 µg of DNA. Digestion was performed for 2-16 h at 37°C. If necessary, the enzyme was heat deactivated. The completeness of cleavage was ensured by analytical gel. If necessary, the fragments were gel purified.

3.2.1.7 DNA ligation

T4 DNA ligase was used to ligate DNA fragments containing 5'-phosphates and 3'-hydroxyl groups. For subcloning, a vector and an insert were cleaved with the same pair of restriction enzymes (the enzymes producing sticky ends were preferentially used). Buffers

and temperature conditions were chosen according to the manufactures instruction (New England Biolabs catalogue 2006/07, USA). To reduce self ligation of vectors, they were 5' dephosphorylated with alkaline phosphatase prior to ligation. After purification, vector and insert were mixed in 1:5 to 1:10 molar ratio (insert:plasmid), T4 DNA ligase buffer and T4 DNA ligase were added and the reaction was incubated overnight at 16°C or for 2h at RT. The circularised product was used for transformation of competent bacteria

3.2.1.8 Site-directed mutagenesis

Site-directed mutagenesis was performed according to the protocol provided with the QuikChange™ kit (Stratagene). The list of used oligonucleotides provided below (Table 2)

Gene product	Sequence 5'→3'	Vector
<i>mmuSecS</i> ^{R313K}	F: CAGCAAGATGTATCCAGGAAAGCCTCAGCCTCGCCGTC R: GACGGCGAGGCTGAGGCTTTCTCTGGATACATCTTGCTG	pET-M13
<i>mmuSecS</i> ^{R313E}	F: CAGCAAGATGTATCCAGGAGAAGCCTCAGCCTCGCCGTC R: GACGGCGAGGCTGAGGCTTCTCTGGATACATCTTGCTG	pET-M13
<i>mmuSecS</i> ^{R313S}	F: CAGCAAGATGTATCCAGGAAGCGCCTCAGCCTCGCCGTC R: GACGGCGAGGCTGAGGCGCTTCTCTGGATACATCTTGCTG	pET-M13
<i>mmuSecS</i> ^{Q105E}	F: CGGTGATATTTCTGCTGTGGAACAAAAAGCTGCAGGCTCTAG R: CTAGAGCCTGCAGCTTTTGGTTCCACAGCAGAAATATCACCG	pET-M13
<i>mmuSecS</i> ^{Q105S}	F: CGGTGATATTTCTGCTGTGTCCACAAAAAGCTGCAGGCTCTAG R: CTAGAGCCTGCAGCTTTTGGTGCACAGCAGAAATATCACCG	pET-M13
<i>mmuSecS</i> ^{E74Q}	F: GGGTAATTGTGGTGTGGGACAAAGGGAAGGGAGAGTGGCCTCTG R: CAGAGGCCACTCTCCCTCCCTTTGTCCCACACCACAATTACCC	pET-M13
<i>mmuSecS</i> ^{E74R}	F: GGGTAATTGTGGTGTGGGACGAAGGGAAGGGAGAGTGGCCTCTG R: CAGAGGCCACTCTCCCTCCCTTGTCCCACACCACAATTACCC	pET-M13
<i>mmutRNA</i> ^{Sec} (RNA 2)	F: GCTTCAAACCTGTAGCTGTCTTCGGACAGAGTGGTTCAATTCC R: GGAATTGAACCACTCTGTCCGAAGACAGCTACAGGTTTGAAGC	pUC19
<i>mmutRNA</i> ^{Sec} (RNA 3)	F: GCTTCAAACCTGTAGCTGTACAGGCGCGCAGACAGAGTGGTTCAATTCC R: GGAATTGAACCACTCTGTCTGCGCGCCTGACAGCTACAGGTTTGAAGC	pUC19

Table 2. Oligonucleotide primers used for mutagenesis. The gene product names, primers used, oligonucleotides inserted within the primers (in red) and vectors used are indicated. In each primer pair, the forward primer is denominated F and the reverse R, respectively.

3.2.1.9 Transformation of *E. coli* cells by heat shock

Chemicompetent (CaCl₂-method) cells of *E. coli* were prepared by standard methods (Sambrook *et.al.*, 1989). XL1blue strain was routinely used for *in vivo* plasmid amplification while the Rosetta (2) BL21(DE3) strain was utilized for protein over expression.

Transfection of chemicompetent bacteria was performed with the heat shock method. Prior to transformation, a 50 µl aliquot of competent cells was thawed on ice and mixed with 5 µl of ligation solution or 5-10 ng plasmid DNA in case of XL1blue strain and 100-200 ng of plasmid DNA in case of expression strain Rosetta(2) BL21(DE3). Mixture was incubated on ice for 30 min, heat shocked at 42 °C for 1 min and cooled on ice 1 min. 900 µl of LB medium were added and the mixture was incubated at 37 °C for 1h with gentle shaking. To distinguish bacteria that were successfully transformed the cells were transferred to agar plates containing appropriate antibiotic(s) and grown overnight.

3.2.1.10 Mini-preparation of plasmid DNA

For plasmid isolation with the Qiagen kits, single colonies the resulting after transformation, were used to inoculate the respective volumes of LB-medium, which are specified by the manufacturer. The cultures were grown overnight at 37 °C and alkaline lysis was performed according to the manufacturers protocol. Positive clones were identified by PCR reaction using the same conditions and set of primers as for insert production for cloning.

3.2.1.11 DNA Sequencing

For DNA sequencing 700 ng of plasmid DNA was mixed with 20 pmol of primer in total volume of 7 µl. Sequencing reaction was done by Seqlab company, Göttingen.

3.2.1.12 *E. coli* Cells cultivation and storage

E. coli cells were cultivated using standard methods (Sambrook *et al.*, 1989). Cells were either grown in liquid media like LB or auto-inducing medium in a shaking incubator at 37 or 16 °C depending on the experiment or they were grown overnight at 37 °C on LB-Agar plates. The auto-inducing medium was prepared as described by Studier (Studier, 2005). Depending on the selection marker of the plasmid, antibiotics in the following concentrations were added to the medium: Ampicillin 100 µg/ml, Chloramphenicol 34 µg/ml, Kanamycin 25 µg/ml. *E. coli* cells were storage at -80 °C for long term. Cells with an optical density 0.6 at wavelength 600 nm were mixed (v/v) 5:1 with 100% glycerol and flash frozen by liquid nitrogen.

3.2.1.13 Preparation of radiolabelled transcripts

The template for in vitro transcription was prepared as described (Carlson *et al.*, 2004). Uniformly ^{32}P -labeled tRNA^{Sec} was transcribed in 50 μl of transcription buffer (30 mM HEPES-NaOH, pH 8.0, 6.4 mM MgCl_2 , 2 mM spermidine, 40mM dithiothreitol, 400 μM each of GTP, CTP, and ATP, 200 μM UTP, 40 μCi of [α - ^{32}P]UTP (3000 Ci/mmol), 32 units of RNasin) containing 5 μg of linearized DNA template and 1.5 μl of T7 RNA polymerase. The transcription mixture was incubated for 2 h at 310 K, and synthetic tRNA^{Sec} was gel-purified.

3.2.1.14 5'-End Labeling of RNA or DNA-Oligonucleotide

10 pmol oligonucleotide were incubated for labeling in a volume of 10 μl with 1 μl (10U/ μl) T4 polynucleotidekinase (PNK), 1 μl 10 fold PNK-buffer in the presence of 3 μl [γ - ^{32}P]-ATP (6000 Ci/mmol; 10 $\mu\text{Ci/ml}$). The mixture was left at 37 °C for 1 h. The 5'-end labeled product was diluted to 50 μl and purified using a mini spin G25 gel filtration column according to the manufacturer's instruction.

3.2.1.15 Autoradiography

Gels containing radiolabelled molecules were exposed to Kodak BioMax X-ray film in a film cassette (typically with an intensifying screen) at -80°C . If the radioactivity of the sample was very high, the cassette was kept at RT for several minutes, rather than at -80°C . Film was developed in a Kodak X-Omat developing machine. Alternatively, gels were exposed to phosphorimager screens for 0.5-24h and scanned by Typhoon PhosphorImager.

3.2.1.16 Denaturing polyacrylamide gel electrophoresis (PAGE)

Denaturing urea PAGE is frequently used to analyze DNA and RNA fragments smaller than 2000 nucleotides. In this study, 10 % polyacrylamide gels (37.5: 1 acrylamide to bis-acrylamide ratio) containing 7-8M urea were routinely used to analyze snRNA and MINX splicing products. Low (3.5-7 %) and high percentage (14-18 %) polyacrylamide gels (19:1 acrylamide to bis-acrylamide ratio) were used to gel purify RNAs (150-400 nts long) and primers, respectively. Prior electrophoresis, samples were denatured by heating with RNA sample buffer at 95 °C for 3-5 min. Electrophoresis was performed vertically in 0.5 or 1xTBE running buffer.

7 M Urea-10 % Polyacrylamide Gel Solution (50ml):

RNA sample buffer:

16.7 ml Rotiphorese Gel 30 % 0.5 X TBE	0.5 X TBE
21 g urea	90 % formamide
5 ml 1X TBE buffer	0.05 % bromophenol blue
Per 50 ml gel solution: 300 µl 10% APS, 30 µl TEMED	0.05 % xylene cyanol

3.2.1.17 Native polyacrylamide gel electrophoresis of RNA

Nondenaturing PAGE is useful method to monitor RNA-protein as well as RNA-RNA complex formation. RNA-RNA interactions were analyzed on 10 % polyacrylamide gels (80:1) while RNA or DNA-protein complexes were separated on 6-8 % polyacrylamide gels. Gels were run at 25mA (7W) in the cold room. 0.5x TBE was used as a running buffer. After electrophoresis, the gels were dried under vacuum at 80 °C for 45-60 min.

Gel solution:

6-10 % (80:1) acrylamide solution 0.5x TBE
 Per 50 ml gel solution: 250 µl 25% (w/v) APS, 25 µl TEMED

Native RNA sample buffer:

50 % glycerol
 0.5x TBE
 0.01 % bromophenol blue

3.2.1.18 Silver staining of RNA

Following electrophoresis, the polyacrylamide gel containing RNA was fixed in 40 % (v/v) ethanol, 5 % (v/v) acetic acid solution for 30 min and then twice for 10 min each in a solution containing 10 % (v/v) ethanol, 5 % (v/v) acetic acid. Subsequently it was washed briefly with water. After incubating for 30 min in 1.2 mM AgNO₃, the gel was washed briefly twice with water, and then incubated with developing solution (280 mM Na₂CO₃, 0.018 % (v/v) formaldehyde) until the desired staining intensity was reached. The gel was incubated in formaldehyde until the desired staining intensity was reached. The gel was incubated in 5 % (v/v) acetic acid solution for 15 min and then dried on Whatmann paper under vacuum at 80°C for 30-60 min.

3.2.2 Standard methods for protein analysis

3.2.2.1. Protein denaturing SDS-polyacrylamide gel electrophoresis

The denaturing SDS-polyacrylamide gel electrophoresis (SDS-PAGE) was performed according to Laemmli *et al.* (Laemmli 1970). Depending on the molecular weight of protein-mixture that had to be separated, the crosslinking-degree was varied. In this study polyacrylamide gels of 12 % and 15 % (37.5:1 AA: MBAA, 1mm thickness) were used. As a standard for gel polymerization 0.25 % APS and 0.1 % TEMED in gel mixture were added.

Before application on the gel, the protein samples were mixed with protein loading buffer and incubated 5 min at 95 °C to ensure complete denaturation. After loading the samples on the gel, the proteins were focused in the stacking gel and subsequently separated in the resolving gel at 30-45 mA.

<u>Protein loading buffer</u> :	<u>Protein running buffer:</u>
75 mM Tris-HCl pH 6.8	25 mM Tris-HCl pH 8.8
1.25 mM EDTA	192 mM glycine
2.5 % (w/v) SDS	0.1 % (w/v) SDS
20 % (w/v) glycerol	
0.1 % (w/v) bromphenolblue	
50 mM DTT	
<u>Stacking gel buffer (4 fold):</u>	<u>Resolving gel buffer (4 fold):</u>
500 mM Tris-HCl pH 6.8	1.5 M Tris-HCl pH 8.8
0.4 % (w/v) SDS	0.4 % (w/v) SDS

Rotiphorese gel 30 (37.5:1 AA: MBAA) was used for the gel mixtures. To achieve the desired percentage of the gel, the respective amount of Rotiphorese gel 30 was calculated, diluted with water and mixed with 1/4th of the corresponding gel buffer. As a standard for polymerization of separation gel 0.12 % APS and 0.1 % TEMED was used. For stacking gel amounts of TEMED was the same and APS was doubled.

3.2.2.2 Gel staining methods

Proteins on SDS-polyacrylamide gels were visualized either by staining them with coomassie brilliant blue G250 and R250 (Sambrook *et al.*, 1989) or by silver-staining (Blum *et al.*, 1987). Whereas coomassie-staining reveals bands of up to 1 µg protein per band, silver-staining is far more sensitive and can detect up to 5 ng in a single protein band.

3.2.2.3 Small scale solubility assay

In order to find a soluble protein fragments cells of Rosetta(2) BL21(DE3) strain were transformed with appropriate expression vectors harboring constructs for expression. Then 3 ml of autoinduction medium were inoculated with a few colonies. After 48 h of expression with shaking at 16 °C 1.2 ml of cell culture was harvested. Pellet was snap-frozen in liquid nitrogen and thawed in cold water for 4 times. Partially lysed cells were resuspended in 500 µl of binding buffer with addition of 16 µl of lysozym (50 mg/ml). Mixture was incubated on ice 10' and then 10 µl of DNase (1 mg/ml) was added. Then tubes were transferred into sonication bath for 5 min to reduce sample viscosity. Cell debris was precipitated by

centrifugation 10 min at maximum speed. 400 μ l of supernatant were transferred into new tube with 50 μ l of Ni-NTA beads equilibrated with binding buffer. After 30 min of incubation at 4 °C on a head-over-tail rotor the supernatants were removed after a short centrifugation at 3000 rpm. Ni-NTA beads were washed with 1 ml of the washing buffer and resuspended in 50 μ l protein loading dye for electrophoresis. The samples were analyzed with SDS-PAGE and visualized with coomassie staining. Construct which had maximum expression and solubility level was used for further studies.

3.2.2.4 Domain mapping with limited proteolysis

Limited proteolysis and subsequent mass spectrometric analysis was applied to detect globular folded domains in the protein of interest. In order to find core structured fragments, which persist during proteolysis, incubation time of mixture protease with target protein was kept constant and protease concentration was varied. First of all several protease dilutions were performed: 1:10, 1:100 and 1:1000. Protease stock solution (1 mg/ml) was diluted with the buffer in which target protein was prepared. Then from each protease dilution 4 μ l was added to 16 μ l of 1 mg/ml protein and incubated for 30 min at 4 °C. The reaction was stopped by addition of 5 μ l of protein loading buffer with immediate heating at 95 °C for 5 min. Separation of the mixture was achieved by SDS-PAGE and bands were visualized with coomassie staining. In order to identify N- and C- end of stable fragments bands were excised and analyzed by mass spectrometry (analyzed by U. Plessmann and M. Raabe, Bioanalytical Mass Spectrometry Group, Max Planck Institute for Biophysical Chemistry, Germany)

3.2.3 Protein production

3.2.3.1 High scale expression of desired construct

E.coli strain Rosetta(2) BL21(DE3) and auto-inducing medium were used for production of all expression constructs. Usually, construct which had maximum expression and solubility level (according to small scale solubility assay) was used for expression in larger volume (usually from 2 to 6 l). Shaking flasks had at least 3 times volume of the culture. Medium was inoculated with 1/100 volume of overnight culture. In the beginning cell culture incubated with shaking at 37 °C and when the medium becomes turbid temperature was lowered to 16 °C. Cells were harvested when the maximum culture density was reached (10 to 12 OD _{$\lambda=600$} in ~48 h). Cells were harvested at 4000 rpm for 15 min, shock frozen in liquid nitrogen and stored at -80 °C.

3.2.3.2 Cell lysis

E. coli cells after over expression of target construct were resuspended in lysis buffer and disrupted by sonication. Proper cooling was accomplished with ice-bath and a number of short pulses (duration 1 sec) with pauses (duration 2 sec) to reestablish a low temperature. Subsequently, the lysate was centrifuged at 30000 rpm for 30 min to pellet the cell debris. The supernatant was supplied to affinity purification as described below.

3.2.3.3 Purification and cleavage of His-tagged fusion proteins

Ni-NTA agarose was equilibrated with A-buffer containing 2 mM β -mercaptoethanol and mixed with the lysate of the overexpressed cells. After $\frac{1}{2}$ -1 h of incubation on a head-over-tail rotor at 4 °C, the beads were pelleted at 2000 rpm for 3 min. The supernatant was removed and the beads were loaded on gravity flow column. Beads were washed consequently with A, A1, A2 buffer. Volume of each buffer was 10 times volume of beads. The elution step was done either with buffer A3 or consequently with several buffers A2 with 2 mM DTT containing gradual increase of Imidazole from 50 to 250 mM (in step of 50 mM). In the first case elution proceeded until no protein was detected by Bradford method. In the second case elution volume of each buffer was 5 times volume of beads. From each purification step starting at the cell lysis a sample was taken, applied to SDS-PAGE and visualized with coomassie staining.

Cleavage of the 6-His-fusion tag was achieved using a ratio of 1:50-1:100 TEV-protease: fusion-protein (mg/mg). In most cases the reaction was carried out during overnight dialysis against buffer A without Triton X-100. Completeness of the cleavage was controlled using SDS-PAGE with following coomassie staining. Protein sample was further purified after dialysis by passing through equilibrated gravity flow NI-NTA column. Bound to the beads impurities were eluted with buffer A2 containing 250 mM of Imidazole. Flow through, containing digested protein, and eluate, containing impurities and undigested protein, were analyzed on SDS-PAGE and visualized with coomassie staining.

A buffer:

20 mM HEPES NaOH pH 7.5

10 mM Imidazole

500 mM NaCl

0.02 % (v/v) Triton X-100

2 mM DTT

A1 buffer:

A2 buffer:

20 mM HEPES NaOH pH 7.5

50 mM imidazole

500 mM NaCl

2 mM β -mercaptoethanol

A3 buffer:

20 mM HEPES NaOH pH 7.5	20 mM HEPES NaOH pH 7.5
10 mM imidazole	250 mM imidazole
1 M mM NaCl	500 mM NaCl
2 mM β -mercaptoethanol	2 mM DTT

3.2.3.4 Protein concentrating and concentration determination

Protein solutions were concentrated using a membrane with a molecular weight cutoff, which is around 2 times less than the molecular weight of the respective protein. The concentration was determined with a Bradford assay: 20 μ l of protein sample was mixed with 1 ml of 1 time concentrated Bradford solution. After 5 min incubation, the absorption at wavelength 595 nm was measured by spectrophotometer. One optical unit was assumed as concentration 1 mg/ml of any protein sample. For precise estimation of protein concentration the sample was measured at wavelength 280 nm. Then based on extinction coefficient of target protein concentration was calculated according Lambert-Beer law.

To obtain the optimal concentration for crystallization, a pre-crystallization test was carried out. Alternatively the protein was crystallized in the Hampton classics screen and the number of drops, in which the protein precipitated, was counted. The optimal concentration was reached when around 1/3rd of all conditions showed precipitation 1 h after crystallization.

3.2.4 RNA production

3.2.4.1 DNA fragment assembly

The DNA fragment for the production of RNA by *in vitro* transcription was assembled by PCR. Oligodeoxynucleotides were created by Assembly PCR Oligo Maker program (<http://publish.yorku.ca/~pjohnson/AssemblyPCRoligomaker.html>). Parameters for assembly PCR were chosen according to Rydzanicz *et. al.*, 2005. Finally, PCR amplified DNA contained T7 RNA polymerase promoter adjacent at 5' end of the RNA gene and endonuclease restriction sites on both 5' and 3' ends. Oligonucleotides used for assembly tRNA^{Sec} gene are listed below (Table 3).

Oligonucleotides	Sequence 5'→3'
Assembly primers	ACTATAGCCCGGATGATCCTCA
	GGTTTGAAGCCTGCACCCCAGACCACTGAGGATCATCCGGGCTA
	GTGCAGGCTTCAAACCTGTAGCTGTTTAGCGACAGAGTGGTTCAATT

	GGCGCCCGAAAGGTGGAATTGAACCACTCTGTCGCTAAAC
Flanking primers	TAATACGACTCACTATAGCC 55
	TGGCGCCCGAAAGGTGG

Table 3. Oligonucleotides for constructing DNA fragment encoding tRNA^{Sec} gene. Assembly primers used for DNA fragment assembly and flanking primers for its amplification by PCR.

3.2.4.2 Preparation of DNA template for *in vitro* transcription

PCR product was used as a template for *in vitro* transcription. In order to generate a DNA template two steps PCR was performed using sense primer annealed to the T7 promoter and antisense primer with modification (two 2'-O-methyl rNTPs at the 5' end) which annealed to the 3' end of RNA gene (Table 4). The first PCR with 25 cycles was carried out in 25 μ l reaction volume containing 0.4 μ Mol of each sense and antisense oligodeoxynucleotides, 1x *Pfu* DNA polymerase buffer, 2.5 mMol of each dNTPs, 5 ng of plasmid DNA as a template and 1.25 U of *Pfu* DNA polymerase. Preparative second PCR was contained the same components and subjected for 35 cycles of amplification with the same reaction conditions as the first PCR. As a template was used 1 μ l of ten times diluted crude mixture from the first PCR. Amplified DNA fragment was phenol-chloroform-isoamylalcohol extracted and precipitated by isopropanol. The pellet was washed twice with 70 % EtOH, dried and dissolved in miliQ grade water.

Product	Sequence 5'→3'	PCR template
RNA1	F: CGACGTTGTAAAACGACGGCC R: TGGCGCCCGAAAGGTGG	pUC19-tRNA ^{Sec}
RNA2	F: CGACGTTGTAAAACGACGGCC R: TGGCGCCCGAAAGGTGG	pUC19-tRNA ^{Sec} (RNA 2)
RNA3	F: CGACGTTGTAAAACGACGGCC R: TGGCGCCCGAAAGGTGG	pUC19-tRNA ^{Sec} (RNA 3)
RNA4	F: CGACGTTGTAAAACGACGGCC R: GCCCGAAAGGTGGAATTGAAC	pUC19-tRNA ^{Sec}
RNA5	F: CGACGTTGTAAAACGACGGCC R: GCCCGAAAGGTGGAATTGAAC	pUC19-tRNA ^{Sec} (RNA 2)
RNA6	F: CGACGTTGTAAAACGACGGCC R: GCCCGAAAGGTGGAATTGAAC	pUC19-tRNA ^{Sec} (RNA 3)
RNA7	F: CGACGTTGTAAAACGACGGCC R: GCGCGCCCGAAAGGTGG	pUC19-tRNA ^{Sec}
RNA8	F: CGACGTTGTAAAACGACGGCC R: GCGCGCCCGAAAGGTGG	pUC19-tRNA ^{Sec} (RNA 2)
RNA9	F: CGACGTTGTAAAACGACGGCC R: GCGCGCCCGAAAGGTGG	pUC19-tRNA ^{Sec} (RNA 3)

Table 4. Generation of the DNA fragments encoding tRNA^{Sec} constructs for *in vitro* transcription. As a DNA templates for PCR were used tRNA^{Sec} constructs inserted into pUC19 vector. tRNA^{Sec} constructs named as RNA1 – RNA9 and shown in Figure 35.

3.2.4.3 *In vitro* transcription from PCR product

RNA was transcribed *in vitro* directly from PCR product using T7 RNA polymerase and established protocol (Pokrovskaya and Gurevich, 1994). A typical transcription reaction

mixture (1 ml) consist of 120 mM HEPES-Na (pH 7.5), 16 mM MgCl₂, 4 mM DTT, 3 mM each NTP (Jena Bioscience), 2 mM spermidine, 0.1 mg Ac-BSA, 1 U/ml inorganic pyrophosphatase (Fermentas), 0.4 U RNasin (Promega), 0.3 mg/ml T7 RNA polymerase (prepared in-house), 10-15 µg PCR product and deionised water (Millipore). The reaction was incubated at 37 °C for 3 h and then DNA template was subsequently digested with RQ1 RNase-free DNase (Promega) at 37 °C for 30 min. Transcribed RNA was analyzed by denaturing PAGE (15 % acrylamide; 8 M urea) and visualized by staining with 0.01 % methylene blue. The transcription mixture can be stored frozen at -20 °C until RNA purification.

3.2.5. Purification of selenocysteine biosynthesis components

3.2.5.1 Purification protocol for *mmuSecS* protein

A DNA fragment encoding the full-length *secS* gene from mouse (GenBank™ accession number NM_172490) was amplified by PCR and cloned into pETM-13 vector to allow the expression of a C-terminal His₆-tagged protein. The insert was verified by DNA sequencing. The resulting plasmid was termed pETM-13-*secS*.

Rosetta(2) BL21(DE3) cells were transformed with the pETM-13-*secS* expression construct. Overproduction of the target protein was carried out at 16 °C using auto-inducing medium. When the maximum culture density was reached cells from 6 l of culture were harvested by centrifugation at 6000 rcf for 15 min at 4 °C and stored in -80 °C. Bacterial pellet (~100 ml from 6 l of auto-inducing medium) was resuspended in buffer A (50 mM HEPES-NaOH, pH 7.5, 500 mM NaCl, 20 mM imidazole, 2 mM β-mercaptoethanol) in ratio 1 g of wet cells: 1,5 ml of buffer. 1 mg of Lysozym per 1 g of cell pellet was added and the mixture was incubated on ice 30 min with stirring. Then 300 µl of DNase I solution (1 mg/ml) and MgCl₂ till 2 mM final concentration were added. Resuspended cells were further incubated with stirring for 10 min. Cells were ruptured by sonication (Branson Sonifier 250) on ice and cleared lysate after centrifugation (30000 rcf, for 30 min at 4 °C) was incubated with Ni-NTA beads (1 ml of beads to 2 l culture volume) in order to capture target protein. Concentration of protein mixture after elution was estimated by Bradford assay. The eluate after concentrating on Milipore concentrator (cut off 100 kDa) was loaded on gel filtration Superdex 200 26/60 column equilibrated with buffer 20 mM HEPES-NaOH pH 7.5, 500 mM NaCl, 2 mM DTT.

3.2.5.2 Purification protocol for *mthSelB*³⁷⁷⁻⁵¹¹ protein

A DNA fragment encoding the WH1 and WH2 motifs of *mthSelB* (amino acids 377–511) was amplified from *M. thermoacetica* genomic DNA by PCR and cloned into plasmid pETM-11 to allow the expression of an N-terminally His₆-tagged protein. The insert was verified by DNA sequencing. Rosetta(2) BL21(DE3) cells were transformed with the *mthSelB*³⁷⁷⁻⁵¹¹ expression construct. Overproduction of the target protein was carried out at 289 K using auto-inducing medium (Studier, 2005). Cells were harvested when the maximum culture density was reached and resuspended in buffer A (10 mM HEPES–NaOH pH 7.5, 200 mM NaCl). Protein was purified from 2 l of medium. Cell lysis and purification on Ni-NTA was carried out in the same way as for SecS with minor differences in buffer composition:

A buffer:

10 mM HEPES NaOH pH 7.5

200 mM NaCl

2 mM DTT

A2 buffer (washing step):

10 mM HEPES NaOH pH 7.5

20 mM imidazole

1 M NaCl

2 mM β-mercaptoethanol

A3 buffer (one step elution):

10 mM HEPES NaOH pH 7.5

250 mM imidazole

200 mM NaCl

2 mM DTT

TEV protease was added to the eluate and the mixture was dialyzed against buffer A overnight. The sample was again passed over Ni-NTA Sepharose, collecting *mthSelB*³⁷⁷⁻⁵¹¹ without a tag in the flow through. Protein was concentrated, aliquot and flash-frozen in liquid nitrogen.

3.2.5.3 Non-denaturing purification of *in vitro* transcribed RNA^{Sec}

RNA transcripts were purified from crude transcription extracts on an anion-exchange Mono Q GL5/50 (GE healthcare) liquid chromatography column on ÄKTA purifier FPLC system (GE Healthcare). Sample was applied at flow rate 1 ml/min via loop or superloop (if sample volume was more than 5 ml). The MonoQ column was equilibrated with several column volumes of low-salt buffer A (0.4 M sodium acetate, 0.2 mM EDTA, 0.1 M Tris-HCl, pH 6.9). RNA transcripts were eluted with gradient (30 - 100 %) of high-salt buffer B (1.5 M sodium acetate, 0.2 mM EDTA, 0.1 M Tris-HCl, pH 7.3) (Kim *et al.*, 1995). The column was operated at 4 °C. Fractions were analyzed by denaturing PAGE (15 % acrylamide, 8 M urea). Fractions containing pure RNA products were pooled, concentrated by ultrafiltration using centrifugal device with a 10,000 MWCO (Millipore) and loaded on Superdex 75 GL 10/300

gel filtration column (GE Healthcare) equilibrated with 10 mM HEPES-NaOH, pH 7.5, 50 mM NaCl buffer. Purification was performed with following parameters: injection volume 0.5 ml, flow rate 0.5 ml/min and fraction size 0.5 ml. Fractions containing homogenous RNA were pooled, concentrated and snap frozen in liquid nitrogen and stored at -80°C .

3.2.6 Crystallization experiments

3.2.6.1 Crystallization conditions screening

Crystallization was performed on the in-house high-throughput facility, a nano drop robot (Cartesian Dispensing System MicroSys 4000XL, Genomic Solutions Ltd, UK). This apparatus was programmed to set up 100 nl-scale vapor diffusion sitting drop crystallization experiments in 96-well plates (containing 100 μl reservoir solutions). The process was controlled by a computer using AxSys software. An overview of the crystallization experiments performed with different protein complexes is provided in Table 5. Optimized reservoir formulations for crystallization listed in Table 6. The initial conditions that yielded crystals using the nano drop robot were subsequently scaled up to microliter range and refined by screening the effects of precipitant, additives and pH. Thus, droplets were set up by mixing 1 μl protein (or RNA) (8 - 20 mg/ml) with 1 μl reservoir and were equilibrated against 500 μl reservoir. The techniques used for the refinement of the crystallization conditions were: sitting drop vapor diffusion.

Macromolecule	Conditions screened	T $^{\circ}\text{C}$
<i>mthSelB</i> ³⁷⁷⁻⁵¹¹ :16h	576	4,20
<i>mmuSecS</i>	768	4,20
<i>mmuSecS</i> ^{elast}	576	20
<i>mmuPSTK</i> : <i>mmutRNA</i> ^{Sec}	768	4,20
<i>mmuPSTK</i> :AppNp: <i>mmutRNA</i> ^{Sec}	768	4,20
<i>mmuPSTK</i> ²⁵⁶⁻³⁵⁹ : <i>mmutRNA</i> ^{Sec}	768	4,20
<i>mmutRNA</i> ^{Sec}	768	20
<i>mmutRNA</i> ^{Sec} (constructs 1-9)	384	20

Table 5. High-throughput crystallization experiments with several proteins and protein:RNA complexes from *Mus Musculus* and *Moorella thermoacetica*.

Macromolecule	Refined crystallization conditions	T $^{\circ}\text{C}$	Crystal freezing
<i>mthSelB</i> ³⁷⁷⁻⁵¹¹ :16h	100 mM HEPES-NaOH pH 6.9-7.9, 3.9-4.5 M NaCl.		100 K N ₂ stream
<i>mmuSecS</i>	100 mM HEPES-NaOH pH 7.5, 1.6 M K/Na phosphate		Liquid N ₂

<i>mmuSecS</i> ^{elast}	11 % (v/v) ethylene glycol without other buffer component	100 K N ₂ stream
<i>mmuPSTK:AppNp:mmutRNA</i> ^{Sec}	100 mM Tris-HCl pH 8.2, 45 % (v/v) MPD	Liquid N ₂
<i>mmuPSTK</i> ²⁵⁶⁻³⁵⁹ : <i>mmutRNA</i> ^{Sec}	100 mM MES pH 5.8, 27% (v/v) PEG 400	Liquid N ₂
<i>mmutRNA</i> ^{Sec}	Have crystallized in many conditions	Liquid N ₂
<i>mmutRNA</i> ^{Sec} (RNA 4), Condition I ^a	100 mM MES pH 5.2, 10 mM MgAc, 2 M AmSO ₄	Liquid N ₂
Condition II ^b	100 mM Na citrate pH 5.2, 0.2 M K/Na tartrate, 1.8 M AmSO ₄	Liquid N ₂
Condition III ^c	0.2 M K ₂ PO ₄ , 40 % (v/v) MPD	Liquid N ₂

Table 6. Optimized reservoir formulations for crystallization of several proteins and protein:RNA complexes from *Mus Musculus* and *Moorella thermoacetica*

^a – Obtained crystals of P21 space group. After 20 % (v/v) Glycerol addition as a cryoprotector, crystals change space group to C2.

^b – Obtained crystals of P21 space group. Data from these crystals used for structure determination.

^c – Crystals of I4 space group.

3.2.6.2 Data collection and structure determination

- Diffraction data were collected from a cryocooled crystal at beamline PXII of the Swiss Light Source (Villigen, Switzerland) and processed using the HKL package (Otwinowski and Minor, 1997). The structure was solved by molecular replacement with MOLREP (Vagin and Teplyakov, 2000) using the coordinates for residues 377–511 of PDB entry 1lva (Selmer and Su, 2002) as a search model. Refinement proceeded by alternate cycles of model building and restrained positional/temperature-factor optimization with REFMAC5 (Murshudov *et al.*, 1997). All residues (377–511) could be unequivocally located in the electron density. Water molecules were automatically added with ARP/wARP (Perrakis *et al.*, 1999) and the water structure was checked and completed manually. Four chloride ions and two sodium ions (Figure 11D) were located based on the coordination spheres and positive difference densities after placement of water molecules at these positions. 14 amino-acid residues were refined with two alternative conformations. In the final refinement cycles, H atoms were automatically placed on the protein structure with REFMAC5 and temperature factors were refined anisotropically for all heavier atoms and ions. Refinement converged at acceptable crystallographic R factors and geometry (Suppl. Table 1)

- All data were recorded on a Bruker-Nonius FR591 rotating anode generator producing CuK_α X-radiation ($\lambda = 1.54179 \text{ \AA}$) at 45 kV and 100 mA equipped with Osmic mirrors and a MAR345 image plate (MAR Research). Data were processed with the HKL package (Otwinowski and Minor, 1997) (Suppl. Table 1).

The structure of *mmuSecS*^{elast} was solved via a single isomorphous replacement with anomalous scattering strategy. Crystals were soaked for 30 sec in cryo-protecting buffer supplemented with 0.5 M sodium iodide and immediately shock frozen in a 100 K nitrogen stream. Friedel pairs were kept separate during data reduction (Suppl. Table 2). The iodide-

soaked crystals proved isomorphous to the native crystals. 42 iodide positions were found by using SHELXD (Schneider and Sheldrick, 2002) and used for initial phase calculations (Supple. Table 2). Solvent flattening with SHELXE clearly indicated the correct hand of the heavy atom substructure (Supple. Table 2) and yielded a high quality experimental electron density map (Suppl. Figure 1).

424 of the 438 residues located in the final structure were positioned in the first round of automatic model building with ARP/wARP (Perrakis *et al.*, 1999). The structure was completed by manual model building and automatic refinement with Refmac5 (Murshudov *et al.*, 1997). Water molecules were automatically placed with ARP/wARP. One solute species was identified as a chloride ion based on the presence of 250 mM NaCl in the crystallization buffer and residual positive electron density after placement of a water molecule. Another solvent molecule was interpreted as an ethyleneglycol molecule originating from the crystallization or cryoprotecting buffer. TLS refinement (Winn *et al.*, 2001) was conducted to model differential global anisotropic displacements of the three domains of *mmuSecS*^{elast}. During all stages of refinement, a randomly selected set of 5% of the reflections was used for cross-validation (Suppl. Table 2). The iodide-soaked crystal structure was refined by the same strategy including the 42 iodide ions located by SHELXD. Additional, lower occupancy iodide positions were found in an anomalous difference Fourier map, obtained with phases calculated from the final refined native structure and the anomalous differences measured for the iodide data set.

For monitoring of phosphate binding, crystals were soaked for 1 min in cryo-protecting buffer supplemented with 0.5 M sodium phosphate, pH 7.5. Similarly, crystals could be derivatized with sulfate (not shown). Data were collected as described, and the structure of a phosphate-soaked crystal was solved by molecular replacement with MOLREP (Vagin and Teplyakov, 2000) using the structure coordinates of the native protein as a search model while omitting the solvent structure, the PLP cofactor, and alternative side chain conformations. Model building and refinement were conducted as described for the native protein (Suppl. Table 2). Coordinates and structure factors have been submitted to the Protein Data Bank.

- For diffraction data collection, crystals were transferred into cryo-protecting buffer (10 mM magnesium acetate, 2.0 M ammonium sulfate, 0.1 M MES, pH 5.2, 20 % glycerol) and shock-frozen in liquid nitrogen. For Mn²⁺-binding studies, crystal were soaked in a similar buffer, in which magnesium acetate was substituted with 100 mM manganese sulfate and 10 mM ammonium acetate for 2 hours and shock-frozen in liquid nitrogen.

Diffraction data for native and derivative crystals were collected at beamline 14.2 of BESSY (Berlin, Germany). Datasets were processed using the HKL package (Otwinowski and Minor, 1997) or XDS (Kabsch, 1988).

The structure was solved by molecular replacement with MOLREP (Vagin and Teplyakov, 2010) using sections of tRNA^{Sec} published earlier (Itoh *et al.*, 2009) as a search model. Molecular replacement searches using the full human tRNA^{Sec} model failed. The structure was completed by manual model building with COOT (Emsley *et al.*, 2010) and automatic refinement with Refmac5 (Murshudov *et al.*, 1997). Water molecules were automatically added with ARP/wARP (Perrakis *et al.*, 1999) and the water structure was checked and completed manually. During all stages of refinement, a randomly selected set of 5 % of the reflections was used for cross-validation (Suppl. Table 3).

3.2.7 Bioinformatics methods

Only the most frequently used online-servers and programs are listed and grouped according to application. Common databases for literature and structures are not listed.

3.2.7.1 Alignments

Clustal W - multiple sequence alignment (<http://align.genome.jp/>)

BLAST-Basic Local Alignment and Search Tool
(<http://www.ncbi.nlm.nih.gov/blast/Blast.cgi>)

TCoffee - structure-based multiple sequence alignment (Notredame *et al.*, 2000)

Vector-NTI program version 9-10 (Invitrogen)

3.2.7.2 Domain Prediction

SMART-Simple Modular Architecture Research Tool (<http://smart.embl-heidelberg.de/>)

PROSITE-Database (<http://www.expasy.ch/prosite/>)

3.2.7.3 Fold Prediction

Foldindex (<http://bip.weizmann.ac.il/fldbin/findex>)

IUPred-Prediction of Intrinsically Unstructured Proteins (<http://iupred.enzim.hu/>)

Secondary structure prediction server (<http://pbil.univ-lyon1.fr/>)

PHYRE version 2.0 – protein homology/analogy recognition server (Kelley and Sternberg, 2009)

3.2.8 Methods used to study RNA-protein interactions

3.2.8.1 Electrophoretic mobility shift assay (EMSA)

The EMSA was used to detect protein complexes with RNAs. Solutions of protein and ^{32}P -labeled RNA were combined. The resulting mixture was subjected to electrophoresis under native conditions through PAGE. Distribution of species containing nucleic acid was determined by autoradiography. Protein-nucleic acid complexes migrate more slowly than the corresponding free nucleic acid. To perform the assay, several components were combined in reaction volume of 10 μL : 20 fmol of the 5'-end ^{32}P -labeled RNA molecules, 1 μl of total yeast tRNA (10 mg/ml), 1 μL of 10 fold binding buffer and protein from 0.5 to 50 μM . Reaction mixture was incubated for 30 min at 4 $^{\circ}\text{C}$. Subsequently, 2 μl loading buffer for native were added to the reaction and 5 μl of the mixture were loaded on 6 or 8% PAGE (AA:MBAA was 29:1) containing 0.5 fold TBE. Electrophoresis was carried out at 4 $^{\circ}\text{C}$ in 0.5 fold TBE for 30' at 8 W. Thereafter, the gel was transferred to Whatman paper and dried at 80 $^{\circ}\text{C}$ for 1 h. The radioactively labeled RNA was visualized with autoradiography on X-ray films or by scanning exposed to the gel phosphoImager screen with Typhoon phosphoImager.

3.2.8.2 Filter binding assay

10,000 cpm (100 fmol) of ^{32}P -labeled unacylated tRNA^{Sec} were incubated 30 min on ice with 2.5, 5, 10, or 20 μM *mmuSecS* or *mmuSecS*^{elast} in 20 μl of buffer D (HEPES-NaOH, pH 7.5, 200 mM NaCl, 5 mM MgCl₂, 1 mM DTT, 0.5 mg/ml total *E. coli* tRNA). 10- μl aliquots of the reaction mixtures were loaded on a Protran BA 83 nitrocellulose membrane (Whatman) and washed with 50 ml of buffer C. For the detection of RNA-protein complexes, the membrane was exposed to a PhosphorImager screen overnight, which was then scanned using a Typhoon 8600 (GE Healthcare).

3.2.8.3 Analytical gel filtration

Analytical gel filtration runs were performed on a SMART-system at 4 $^{\circ}\text{C}$. Superdex 75 or 200 PC 3.2/30 columns were utilized, depending on the protein or RNA-protein complex size and equilibrated with 20 mM HEPES - NaOH pH 7.5, 150 mM NaCl, 5 mM MgCl₂, 2 mM DTT. Gel filtration standard was applied to the column in the mentioned above buffer. Normally the flow rate was set to 40 $\mu\text{l}/\text{min}$ and the fraction size was 40 μl . In the majority of the cases, 50 μl of 1 mg/ml protein or RNA-protein complex were subjected to gel filtration.

4 RESULTS

4.1 Conformational switches in winged-helix domains 1 and 2 of bacterial translation elongation factor SelB

4.1.1 Putative complex between WH1/2 of SelB and helix 16 of 16S rRNA

The bacterial translational elongation factor SelB is involved in incorporation of Sec into the nascent polypeptide chain in bacteria. With the N-terminal part it binds GTP and with high specificity the Sec-tRNA^{Sec}. The C-terminal extension harbours four WH motifs arranged in two functional pairs WH1/2 and WH3/4. The latter two motifs are responsible for SECIS binding, whereas the function of the first two motifs is not clear.

Possible role of the WH1/2 motifs in establishing contacts with the ribosome via interaction with the helix 16 of 16S rRNA (16h) was suggested based on observations of complex formation between SelB and the ribosome by cryo-EM (N. Fischer and H. Stark, 3D Electron Cryo-Microscopy Group, Max Planck Institute for Biophysical Chemistry, Germany, personal communication). In the crystal structure of a SECIS element with the entire C-terminal SelB domain it was found the WH1/WH2 and WH3/WH4 modules are positioned differently from known structures of that region in isolation indicating its high flexibility (Ose *et al.*, 2007). In addition the WH2-WH3 interdomain region forms a positively charged area that binds to the phosphate backbone of a neighboring SECIS. Interaction only with phosphate backbone facilitates sequence independent RNA binding. Therefore such recognition in the crystals may mimic a possible interaction with rRNA (helix 16 or 33) during selenocysteine incorporation (Ose *et al.*, 2007).

In order to shed light on the communication between SelB and the ribosome, I decided to study the putative interaction of WH1/2 with the 16h. Electrophoretic mobility shift assays (EMSA) showed that a complex between WH1/2 of *Moorella thermoacetica*, ranging from residue 377 to 511 (*mthSelB*³⁷⁷⁻⁵¹¹), and radioactively labeled 16h is formed in isolation. However, the affinity of *mthSelB*³⁷⁷⁻⁵¹¹ to the RNA substrate was low and unspecific (Figure 10).

Nonetheless, crystallization trials of a mixture *mthSelB*³⁷⁷⁻⁵¹¹ and 16h were setup. Crystals appeared in crystallization conditions containing high salt (for details see “Materials and Methods”) which is unfavorable for complex formation. Indeed, the solved structure determined at 1.1 Å resolution consists of only two WH domains [(Suppl. Table 1), Figure 11].

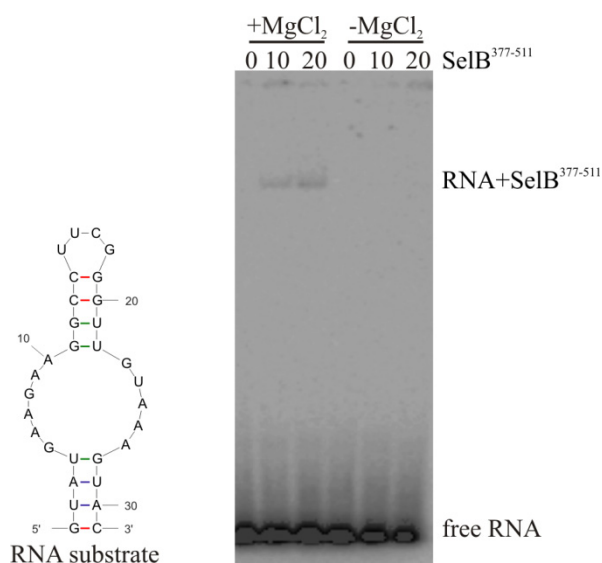


Figure 10. Analysis of complex formation by EMSA. The experiment was performed by titration of different SelB³⁷⁷⁻⁵¹¹ concentrations (0, 10 and 20 μM) to a constant concentration of [³²P] RNA (helix 16 of 16S rRNA from *E. coli*) in the presence or absence of magnesium chloride.

4.1.2 Overall fold and global conformational changes

Globally, the WH1 and WH2 domains resemble the structure of the domains in the framework of the entire C-terminal portion of SelB (residues 377-634; Selmer and Su, 2002; Figures 11A and 11B). Each domain is composed of three α -helices and a three-stranded antiparallel β -sheet (Figure 11C). The three α -helices form a short bundle that stands like a tripod on the β -sheet (Figure 11A). The β -sheet of the first domain rests on the helical bundle of the second domain, giving rise to alternating levels of α -helices and β -strands along the structure (Figures 11A and 11B).

Although the sequences of the proteins are identical, superimposition of the present structure and the corresponding portion of the WH1-4 structure (Selmer and Su, 2002) yielded a root-mean-square deviation (r.m.s.d.) of 1.75 \AA for the 135 common C $^{\alpha}$ atoms (Figures 11A and 11B), indicating a significant conformational difference. When the two structures are aligned with respect to the WH2 domains, the WH1 domains adopt different relative orientations. In part, the positional change in WH1 is a consequence of a bend of $\sim 7^{\circ}$ in the centre of helix $\alpha 1$ of WH2 in the WH1-4 structure (marked with an arrow in Figure 11B); the same element of the present structure is straight. More importantly, the helical bundles in the two WH1 domains adopt different positions with respect to the WH1 β -sheets (Figures 11 and 12). The entire three-helix bundle of WH1 moves like a rigid entity, since the α -helical subdomains of the two WH1 modules alone superimpose closely (r.m.s.d. of 0.81 \AA for 44 common C $^{\alpha}$ positions). In the present structure, the three-helix bundle of WH1 has separated

from the β -sheet at the side of helix α_3 , the apparent movement resembling the opening of a lid (Figure 11A).

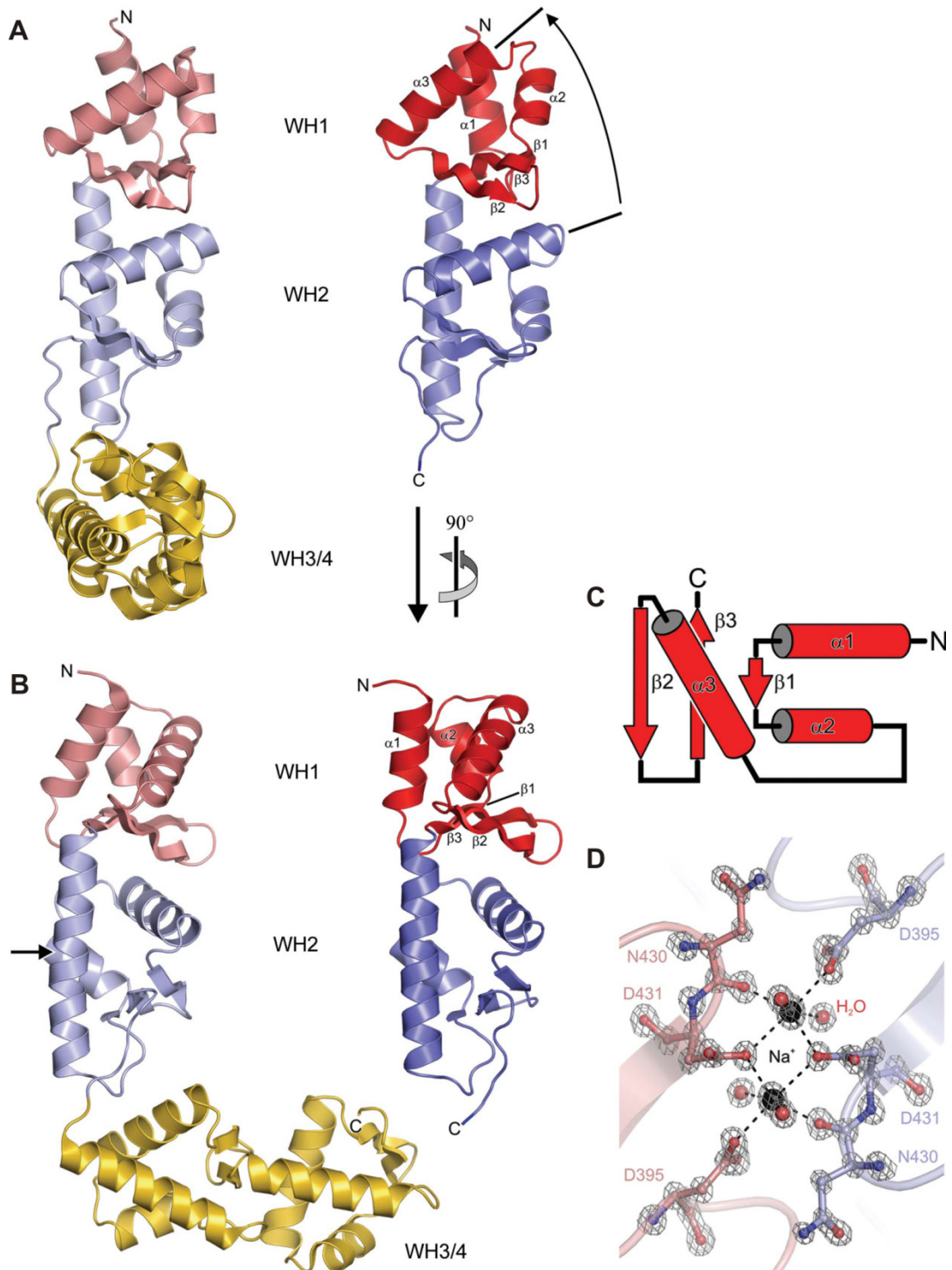


Figure 11. (A) and (B) Comparison of the previously determined WH1-4 structure of *mtbSelB* (left; PDB ID 1lva; Selmer and Su, 2002) with the present structure of domains WH1/2 (right). WH1, WH2 and WH3/4 of the WH1-4 structure are shown in light red, light blue and gold, respectively. WH1 and WH2 of the present structure are shown in red and blue, respectively. Secondary-structure elements of WH1 of the present structure are labelled. N, N-terminus; C, C-terminus. The view in (B) is orthogonal to that in (A), as indicated. The arrow in (A) emphasizes the change in the angle between the helical axes of the two neighbouring α_3 helices, representative of the conformational switch in WH1. An arrow in (B) indicates a bend of $\sim 7^\circ$ in helix α_1 of WH2

that is not seen in the present structure, in which the helix is straight. (C) Schematic representation of the topology of the WH domains. (D) $2F_o - F_c$ electron density (grey mesh) contoured at the 3σ level in the region of a crystal-packing motif involving two sodium ions (black spheres), water molecules (red spheres) and protein residues 395, 430 and 431 (ball-and-stick representation; C atoms are color-coded according to the respective protein; nitrogen, blue; oxygen, red).

4.1.3 Local conformational switch in WH1

The global conformational rearrangement can be traced to a local reorganization of a particular residue. At the interface region of helix α_3 and the β -sheet of WH1, the side chain of Trp396 adopts a different orientation compared with that in the previous structure (Figure 12A). The $C^\beta - C^\gamma$ dihedral angle of Trp396 has changed by almost 180° . As a consequence, the five-membered portion of the indole ring of Trp396 is proximal to helix α_3 in the previous structure, while the bulkier six membered part of the indole ring is rotated towards helix α_3 in the present structure (Figure 12A). Thus, in our structure the bulk of Trp396 is wedged between helix α_3 and the β -sheet. Despite the different orientation of the Trp396 side chain, it remains tightly packed in a hydrophobic pocket formed by Leu407, Thr410, Leu414 (all originating from α_3) and Leu432 (originating from β_3 ; Figure 12A). The latter four residues adjust slightly and together with the overall movement of the helices reshape the pocket for Trp396 (Figure 12B).

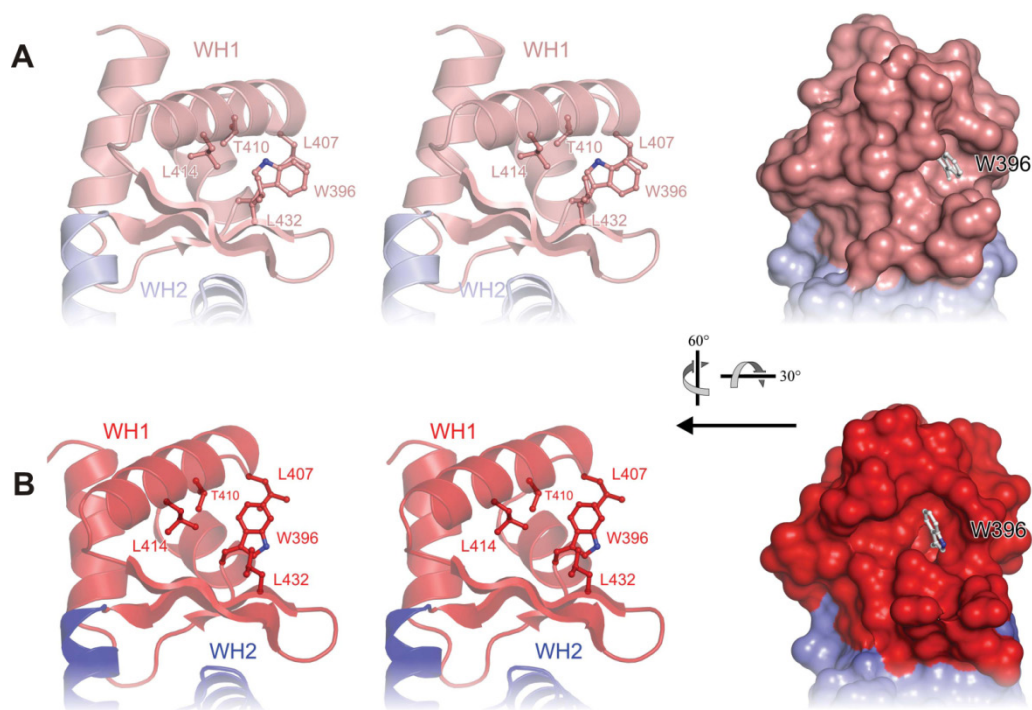


Figure 12. (A) Stereo ribbon plots comparing WH1 of the previous WH1-4 structure (top) and the present structure (bottom). The colouring of the domains is the same as in Figure 11. Trp396 and neighbouring residues are shown in ball-and-stick representation and are colour-coded as before. (B) Surface views of the same regions of the two crystal structures with Trp396 shown in ball-and-stick representation (carbon, grey). The Trp396 pocket adapts upon rearrangement of the Trp396 side chain. The view in (B) is the same as in Figure 11(A). In (B) the structures are rotated 30° about the horizontal axis and 60° about the vertical axis, as indicated.

4.1.4 Emergence of a bona fide ligand-binding site

Surface rendering of the molecules reveals that a short channel is opened up in the present structure directly neighbouring Trp396 and on the underside of helix $\alpha 3$ (Figure 12B). In the previous WH1-4 structure (Selmer and Su, 2002) this channel is completely occluded by the backbone of helix $\alpha 3$ and the side chains originating from this element (Figure 12B). The dimensions of the uncovered cleft appear to be suitable for the binding of a ligand, such as a short stretch of protein or nucleic acid. Indeed, the C-terminal tail of a neighbouring molecule in the crystal lattice is inserted snugly into this cleft (Figure 13A). The C-terminal Phe511 side chain of this neighbour is placed perpendicularly to the indole ring of Trp396 and may push it under helix $\alpha 3$ (Figure 13A).

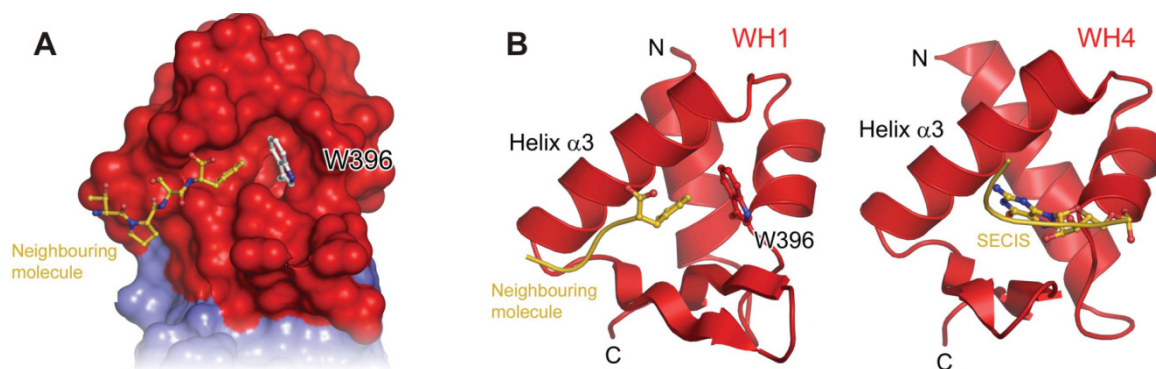


Figure 13. (A) Interaction of the WH1 domain of the present crystal structure (red surface; Trp396 in ball-and-stick representation; carbon, grey) with the C-terminal tail of a neighbouring molecule (ball-and-stick representation; carbon, gold). (B) Comparison of the interaction mode of WH1 in the present structure with the C-terminal tail of a neighbouring molecule (left) and the interaction mode of WH4 of *mthSelB* with a SECIS element (right; PDB ID: 2UWM; Ose *et al.*, 2007). WH domains are shown as red ribbon plots. Trp396 of WH1 is shown in ball-and-stick representation (carbon, red). Ligands are shown as a combination of ribbon and ball-and-stick representations (carbon and phosphorus, gold). Helices $\alpha 3$ and the N- and C-termini of the WH modules are labelled. Views are the same as in Figure 12B.

Interestingly, the location of the neighbouring peptide and in particular of Phe511 resembles the position of a guanine nucleotide from a SECIS element on the WH4 module in a previously determined SelB WH3/4-SECIS co-crystal structure (Figure 13B; Yoshizawa *et al.*, 2005). While results of the present work were in preparation for publication, two additional structures of C-terminal portions of SelB in complex with SECIS elements were reported (Soleret *et al.*, 2007; Oset *et al.*, 2007). In these structures, nucleotides bulged out from the stem or the apical loop of the SECIS elements again interact at the equivalent binding sites on the WH3 and WH4 elements of SelB, respectively.

Thus, although the conformational change observed in the WH1 domain correlates with a crystal lattice contact and presently it is not possible to attribute cause and effect, the contact site for a neighbouring molecule that emerges in the *mthSelB*³⁷⁷⁻⁵¹¹ crystal structure

resembles an authentic ligand-binding site in homologous domains. Together, these observations suggest that the WH domains of SelB harbour prominent ligand-binding sites underneath helix α_3 which may bind RNA or protein molecules.

4.1.5 Flexibility of the connecting peptide between WH2 and WH3

The two pairs of WH domains in the *mthSelB* WH1-4 structure (Selmer and Su, 2002) adopt an “L”-shape. This orientation of the WH-domain pairs is apparently stabilized by a salt bridge between Arg461 of WH2 and Glu552 of WH3 (Selmer and Su, 2002). Co-variation of the two interacting residues suggests that this salt bridge is conserved throughout bacterial SelB molecules (Selmer and Su, 2002). In *Escherichia coli* (*eco*) SelB, a mutation that destroys this salt bridge relaxed the stringency of the SECIS recognition (Kromayer *et al.*, 1999). Based on these findings and molecular modelling, which demanded a more elongated structure of SelB on the ribosome than that observed in the crystal, Selmer and Su suggested that SECIS binding leads to a functionally important conformational change in the factor (Selmer and Su, 2002).

In the present structure, the lack of WH3 occludes the possibility of a WH2-WH3 salt bridge. Under these conditions, the C-terminal tail of domain WH2 starting at residue 504, directly after the third strand of the WH2 β -sheet, follows a different path compared with the WH2-WH3 linker peptide of the WH1-4 structure (Figures 11A and 11B). Instead, the present WH2 terminus closely resembles the WH2-WH3 linker in a structure of *mthSelB* WH1-4 in complex with a SECIS element (Ose *et al.*, 2007). In this latter structure, the salt bridge connecting WH2 and WH3 is broken, yet the domains remain in a fixed relative orientation owing to crystal-packing interactions via a neighbouring SECIS element. The path of the WH2 C-terminal tails in the above structures appears to be additionally governed by a hydrogen bond from the Leu504 backbone NH to the side chain of His455 from the first helix of WH2 and by stacking of the side chains of His455 and Phe507. The different pH values under which the present crystals (pH 6.9–7.9), those of the *mthSelB* WH1-4 structure (pH 6.2) and those of the *mthSelB* WH1-4 in complex with a SECIS element (pH 5.5) were obtained may have influenced these ionic interactions.

Together, these observations corroborate the previous notion that the connection between the two pairs of WH domains, WH1/2 and WH3/4, is flexible and allows different relative orientations of the domain pairs after disruption of the Arg461-Glu552 salt bridge (Selmer and Su, 2002). In further agreement with this suggestion, the recent crystal structure of *ecoSelB*WH3/4 in complex with a SECIS element showed that SECIS binding sequestered

Arg524 (the equivalent of Glu552 in the reciprocal WH2-WH3 salt bridge in *ecoSelB*) in the formation of a SECIS-binding pocket (Soler *et al.*, 2007). By sequestration of Arg524, SECIS binding could lead to breakage of the WH2-WH3 salt bridge, giving rise to increased flexibility as required during the subsequent interaction of SelB with the ribosome.

4.2 Structure and catalytic mechanism of eukaryotic selenocysteine synthase

4.2.1 Limited proteolysis delineates a SecS core fully active in P^{Ser}-tRNA^{Sec} to Sec-tRNA^{Sec} conversion

The molecular identity of eukaryotic/archaeal SecS has only recently been elucidated. The RNA-mediated interference technology provided the first direct evidence for an essential role of soluble liver antigen/liver and pancreas antigen (SLA/LP) in selenoprotein biosynthesis (Xu *et al.*, 2005). The protein was identified in precipitated complex with tRNA^{Sec} by autoantibodies from patients with a severe form of autoimmune chronic active hepatitis (Gelpi *et al.*, 1992; Costa *et al.*, 2000). Unequivocal evidence that SLA/LP embodied the elusive eukaryotic/archaeal SecS was finally provided independently by two groups who directly demonstrated the conversion of P^{Ser}-tRNA^{Sec} to Sec-tRNA^{Sec} by the enzyme (Xu *et al.*, 2007a; Yuan *et al.*, 2006).

At the start of this project, it was not known how the enzyme recognizes Sec-tRNA^{Sec} and the SeP substrate and how SecS carries out its function. Therefore full-length SecS from mouse (*mmuSecS*) was expressed, purified, and crystallized, but diffraction was limited to ~6 Å resolution. To explore the possibility that flexible regions hindered generation of well ordered crystals, *mmuSecS* (~55 kDa) was digested with various proteases. Elastase gave rise to a stable fragment of about 49 kDa (Figure. 14A). Tryptic mass spectrometric fingerprinting (performed by M. Raabe and H. Urlaub, Bioanalytical Mass Spectrometry Group, Max Planck Institute for Biophysical Chemistry, Germany) showed that the elastase-resistant fragment encompassed residues 19-468 (not shown). Thus, the protease removed most of the C-terminal portion that carries the main SLA/LP antigenic epitope (Gelpi *et al.*, 1992; Wies *et al.*, 2000), also lacking in archaeal SecS (Figure 15). In the following work the elastase-resistant core of the enzyme is referred to as *mmuSecS*^{elast}.

Gel filtration analysis revealed that *mmuSecS* is tetrameric in solution (Figure 14B). Elastase treatment did not change the migration behaviour of the protein detectably, showing that the proteolytic treatment left the quaternary structure of the enzyme intact (Figure 14B).

Because *mmuSecS* binds significantly to unacylated tRNA^{Sec} (Xu *et al.*, 2007a), I examined the tRNA^{Sec} binding activity of *mmuSecS*^{elast} by nitrocellulose filter binding. The elastase-treated protein bound tRNA^{Sec} with an affinity and specificity comparable with those of the full-length enzyme (Figure 14C). We next tested the Sec synthesis activity of *mmuSecS*^{elast} using a paper chromatographic assay (Xu *et al.*, 2007a).

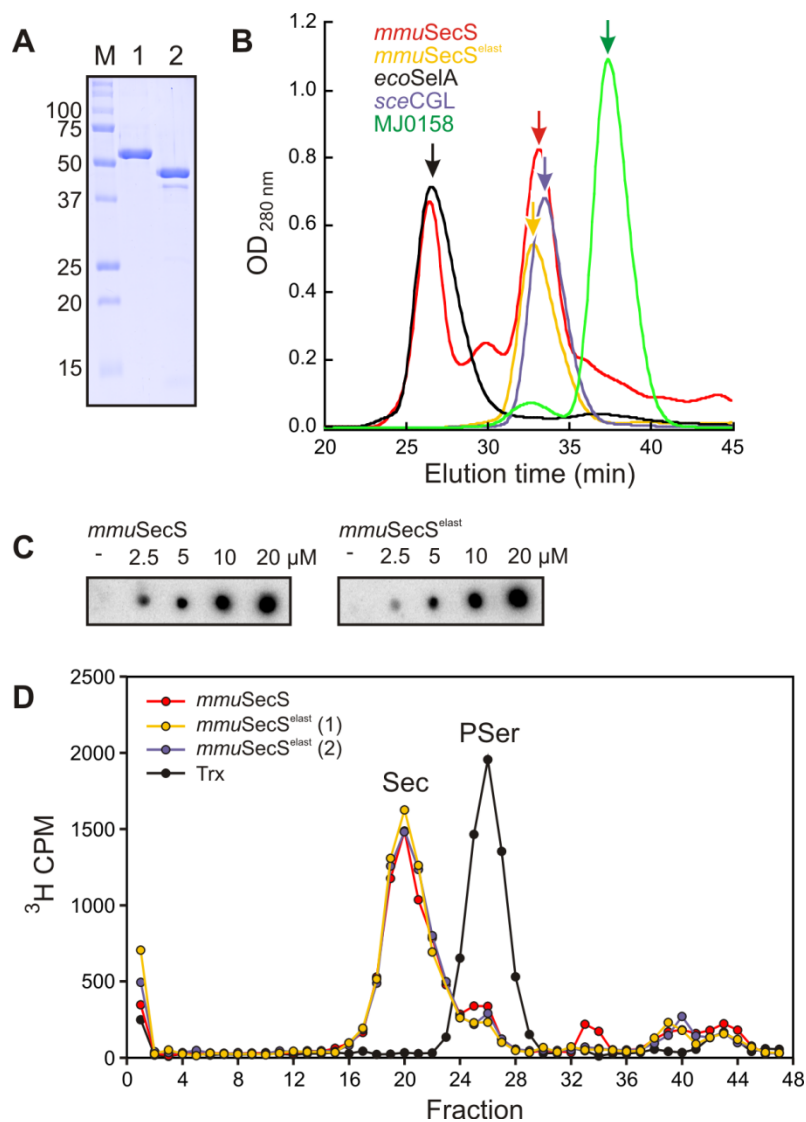


Figure 14. Characterization of *mmuSecS*^{elast}. (A), SDS-PAGE of *mmuSecS* (lane 1) and *mmuSecS*^{elast} (lane 2). M, molecular weight markers. (B), analytical gel filtration analysis of *mmuSecS* (red) and *mmuSecS*^{elast} (gold), *E. coli* SelA (*ecoSelA*) (black, ~506 kDa), *S. cerevisiae* cystathionine γ -lyase (*sceCGL*; blue, ~170 kDa), and MJ0158 (green, ~84 kDa) served as molecular mass markers. Peaks containing the respective proteins (arrows) were verified by monitoring the absorbance at 420 nm and by SDS-PAGE analysis (not shown). The high molecular weight peak (left) in the full-length *mmuSecS* run represents a contaminating nucleic acid fraction. (C), filter binding assay showing similar affinity of *mmuSecS* (top panel) and *mmuSecS*^{elast} (bottom panel) for deacylated tRNA^{Sec}. Experiments were conducted in the presence of unlabeled competitor tRNA and, thus, represent specific affinities. Concentrations of protein in the reactions are indicated. (D), conversion of [³H]P-Ser-tRNA^{Sec} to [³H]Sec-tRNA^{Sec} by *mmuSecS* (red) and *mmuSecS*^{elast} (gold and blue, which are duplicates, *mmuSecS*^{elast} (1) and *mmuSecS*^{elast} (2)). Trx (black) served as a negative control. Activity assay in section (D) performed by D. Hatfield and colleagues (department of biochemistry, NIH, USA).

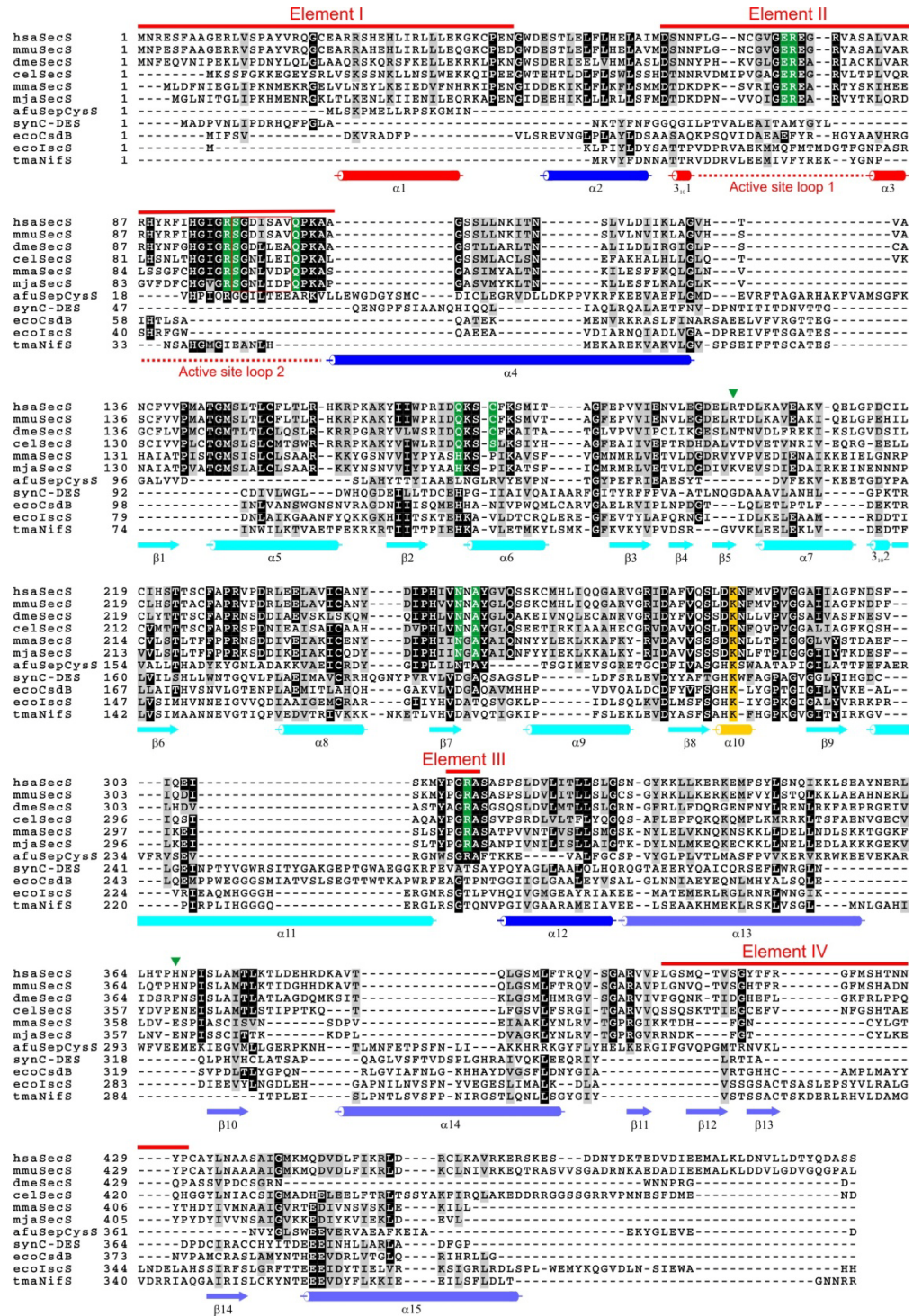


Figure 15. Structure-based multiple sequence alignment of SecS and enzymes of the fold type I family. Structure-based multiple sequence alignment generated by the 3DCoffee-option of Toffee (Notredame *et al.*, 2000) and shaded by BoxShadeSequences are numbered at the beginning of each line. The darker background represents higher conservation. The background of the PLP lysines is in gold; the background of residues whose side chains contact PLP and of residues contacting the P1 phosphate is in green. P-loop residues undergoing a disorder-order transition upon phosphate binding are boxed in red. Residues contacting the P2 phosphate are indicated by a green arrowhead. Elements I-IV discussed in the text are indicated by labeled red bars above the alignment (see also Figure 16). Active site loops 1 and 2 from element II are indicated by dashed lines below the alignment. Secondary structure elements as found in *mmuSecS*^{elast} are indicated below the alignment. Scaffolding elements are color-coded by domain; cyan, domain 1; blue, domain 2; steel blue, domain 3. Secondary structure elements belonging to elements I-IV are in red, and helix α_{10} (bearing the PLP cofactor) is in gold. hsa, *H. sapiens*; mmu, *M. musculus*; dme, *D. melanogaster*; cel, *C. elegans*; mma, *M. maripaludis*; mja, *M. jannaschii*; afu, *A. fulgidus*; syn, *Synechocystis*; eco, *E. coli*; tma, *T. maritima*.

After removal of elastase by gel filtration chromatography, *mmuSecS*^{elast} was incubated with buffer containing SeP and PSer-tRNA^{Sec} (obtained by phosphorylation of Ser-tRNA^{Sec} by PSer-tRNA^{Sec}) (Carlson *et al.*, 2004, Xu *et al.*, 2007a). As positive and negative controls, *mmuSecS* and Trx, respectively, were substituted for *mmuSecS*^{elast}. The efficiency at which *mmuSecS*^{elast} converted PSer-tRNA^{Sec} to Sec-tRNA^{Sec} was indistinguishable from that of the full-length enzyme (Figure 14D). The above results demonstrate that *mmuSecS*^{elast} structurally and functionally closely resembles the full-length enzyme and constitutes a suitable platform on which to explore the structure-activity relationships of SecS-dependent Sec biosynthesis.

4.2.2 SecS is a member of the fold type I family of PLP-dependent enzymes with distinct structural elements

mmuSecS^{elast} crystallized readily after the addition of 11 % (v/v) ethylene glycol at room temperature. The crystals diffracted to high resolution on a rotating anode x-ray generator and could be derivatized by quick-soaking in 0.5 M sodium iodide for structure solution by single isomorphous replacement with anomalous scattering (Suppl. Table 2, Suppl. Figure 1). Refinement converged at R/R_{free} factors of 16.8/19.8 % with the final model exhibiting good overall stereochemistry (Suppl. Table 2). The only amino acids lacking well defined electron density were residues 19-22 at the N terminus, residue 468 at the C terminus, and residues 98-104, constituting a flexible loop.

Structural homology searches (Holm and Sander, 1995) suggested that *mmuSecS*^{elast} exhibits significant similarity to the fold type I family of PLP-dependent enzymes (also referred to as the aspartate aminotransferase family) (Schneider *et al.*, 2000). Where appropriate, will be compared the structure of *mmuSecS*^{elast} to those of *Archaeoglobus fulgidus* PSer-cysteine synthase (*afuPSerCysS*) (Fukunaga *et al.*, 2007), members of the NifS family of Cys/Sec lyases (Fujii *et al.*, 2000; Kaiser *et al.*, 2000; Lima, 2002; Mihara *et al.*, 2002), and the cystine C-S lyase C-DES from *Synechocystis* (*synC-DES*) (Clausen *et al.*, 2000). A quantitative comparison with these proteins is given in Table 7. These enzymes act or can act on related substrates and may share some catalytic properties with SecS. In particular, PSerCysS from methanogenic Archaea affords a precedence for the tRNA-based amino acid synthesis via a PSer-tRNA intermediate. In these organisms the sole pathway for cysteine biosynthesis is via PSer-tRNA^{Cys}, obtained by direct aminoacylation of tRNA^{Cys} with PSer by PSer-tRNA^{Cys} synthetase, and subsequent PSer-tRNA^{Cys} to Cys-tRNA^{Cys} conversion by PSerCysS (Sauerwald *et al.*, 2005).

	afuPSerCys	tmaNifS-like protein	syncC-DES
PDB code	2E7J	1EG5	1ELU
Reference	Fukunaga and Yokoyama, 2007	Kaiser <i>et al.</i> , 2000	Clausen <i>et al.</i> , 2000
Sequence identity (%)	10.9	13.1	11.7
Matching C ^α atoms	257	283	266
r.m.s.d. (Å)	2.8	2.8	3.0

Table 7. Structural comparison of three fold type I PLP-dependent enzymes with *mmuSecS*^{elast}

mmuSecS^{elast} can be divided into three domains (Figure 16, A and B). Domain 1 (blue scaffold in Figure 16, A and B) is a composite of residues 23-130 and 313-330 and is purely α -helical (encompassing helices $\alpha 1$ - $\alpha 4$ and $\alpha 12$). The N-terminus (residues 23-44; element I in Figure 16, A and B) differs from that of other fold type I enzymes (Figure 16, C and D). In *mmuSecS*^{elast}, helix $\alpha 1$ is positioned at the protein surface, running approximately perpendicular to the scaffolding helices of the domain. Domain 1 also exhibits a long insertion between helices $\alpha 2$ and $\alpha 4$ (residues 62-108; element II in Figure 16, A and B), which encompasses two loops separated by helix $\alpha 3$. The corresponding element in other enzymes of the fold type I family is significantly shorter (Figure 16, C and D). Seven residues within the second loop of the insertion (residues 98-104) are disordered due to intrinsic flexibility (bordered by spheres in Figure 16A). Both the non-canonical N terminus and the unique insertion are conserved among SecS orthologs (Figure 15) and, therefore, are expected to confer unique functions on the enzyme.

Domain 2 of SecS is the largest module of the protein (residues 131-312; cyan scaffold in Figure 16, A and B). It comprises a $\alpha/\beta/\alpha$ sandwich fold encompassing a seven-stranded β -sheet ($\beta 1$ - $\beta 9$ - $\beta 8$ - $\beta 7$ - $\beta 6$ - $\beta 2$ - $\beta 3$) characteristic of the fold type I family. The β -sheet is parallel except for strand $\beta 9$ (Figure 16, A and B). A short helix ($\alpha 10$) between strands $\beta 8$ and $\beta 9$ carries a PLP cofactor in Schiff-base linkage to Lys-284 (Figure 16A and 17). Helices $\alpha 7$, $\alpha 8$ and $\alpha 9$ line the β -sheet at the convex outside, and helices $\alpha 5$, $\alpha 6$, and $\alpha 11$ lie at the concave inside. A short β -hairpin ($\beta 4$ and $\beta 5$) is inserted between strand $\beta 3$ and helix $\alpha 7$. Domain 2 is connected to the second part of domain 1 by a short loop (residues 311-314; element III in Figure 16, A and B).

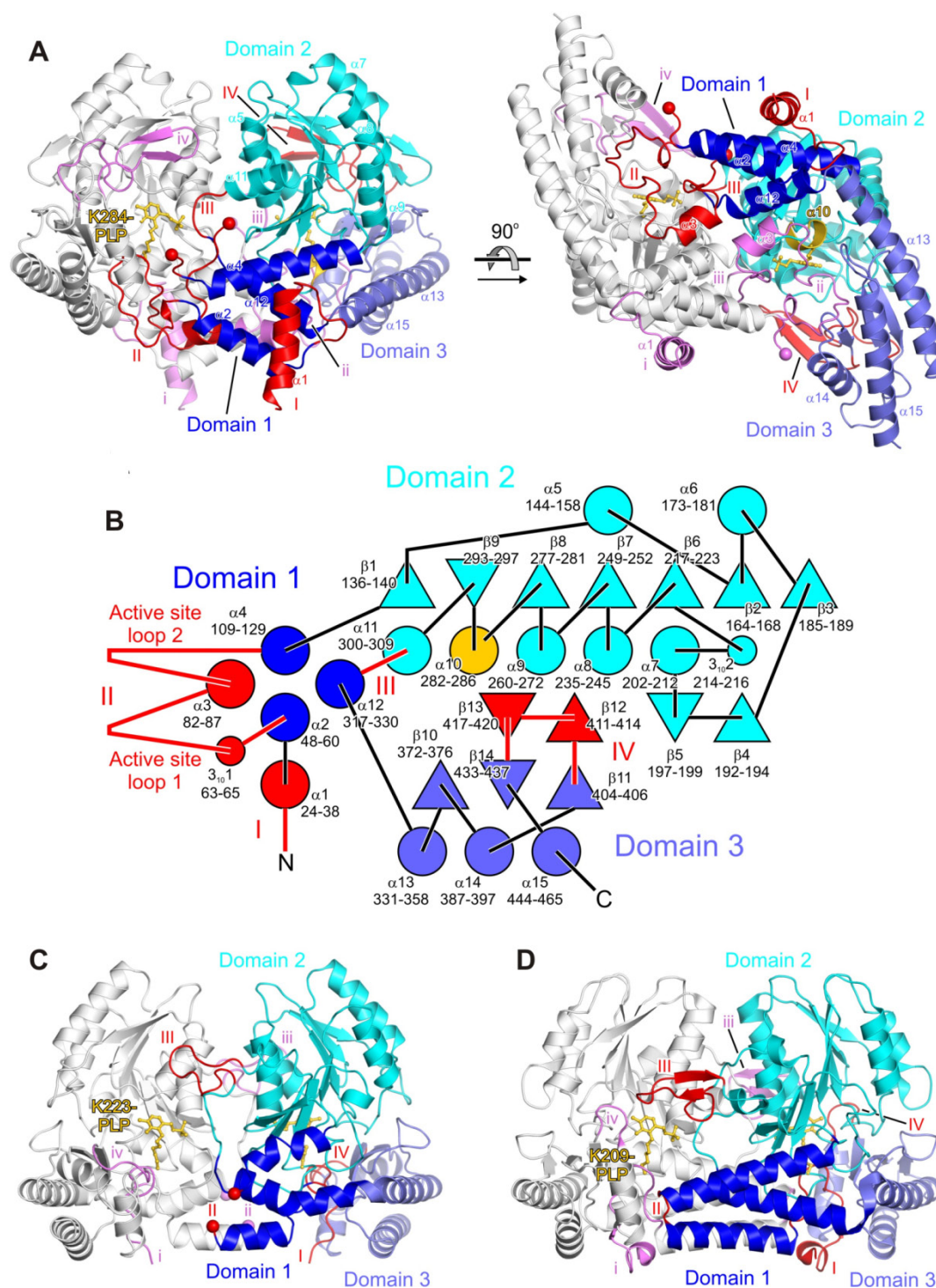


Figure 16. Domain organization and functional elements. (A), orthogonal ribbon plots (front and bottom) of *mmuSecS*^{elast}. One protomer is in light gray, and the other protomer is colour-coded according to domains (blue, scaffold of domain 1; cyan, scaffold of domain 2; steel blue, scaffold of domain 3). Selected elements (I-IV) as discussed in the text are shown in red and labelled with uppercase roman numerals for the reference molecule (Mol I) *viz.* in violet and labelled with lowercase roman numerals for the other protomer (Mol II). Helix α_{10} , bearing the PLP cofactor, is in gold. Selected α -helices are labelled as landmarks. (B), secondary structure topology (large circles, α -helices; small circles, 3_{10} -helices; triangles, β -strands). The secondary structure elements are labelled, and their residue extents are indicated. They are coloured as in (A). (C) and (D), ribbon diagrams of *afuPSer-CysS* (C), PDB ID: 2E7J (Fukunaga *et al.*, 2007) and *synC-DES* (D), PDB ID: 1ELU; (Clausen *et al.*, 2000) dimers in the front view after global superimposition of the proteins on *mmuSecS*^{elast}. Colouring is the same as for *mmuSecS*^{elast}.

Domain 3 of SecS (residues 331-467; steel blue scaffold in Figure 16, A and B) exhibits an α/β sandwich fold with three almost parallel helices ($\alpha 13$, $\alpha 14$, and $\alpha 15$) on the outside covering an antiparallel three-stranded β -sheet ($\beta 10$ - $\beta 14$ - $\beta 11$), which in turn rests on top of domain 2. A long loop (residues 408-430; element IV in Figure 16, A and B) with a β -hairpin ($\beta 12$ - $\beta 13$) at the tip is inserted between strands $\beta 11$ and $\beta 14$. The three-stranded β -sheet and the β -hairpin of the loop are at approximately right angles and encircle part of domain 2 (Figure 16A). Strands $\beta 11$ and $\beta 12$ thereby form one rim of an active site funnel leading from the surface to the PLP (Figure 16A). In the NifS-like enzymes, the analogue of the long domain 3 loop (element IV) is often disordered (Kaiser *et al.*, 2000) and bears a conserved Cys that can be charged in the active site with elemental sulphur. The resulting persulfide is thought to donate a S^0 building block for iron-sulphur cluster biosynthesis. Element IV of *mmuSecS*^{elast} does not contain a Cys.

4.2.3 Cross-strutting via the N terminus leads to homotetramers that exhibit surface properties suitable for binding P_{Ser}-tRNA^{Sec}

mmuSecS^{elast} crystals contained one protein molecule per asymmetric unit. Consistent with the gel filtration analysis, the orthorhombic crystal symmetry gave rise to tetramers in which the protomers are related by three orthogonal 2-fold axes (Figure 18A, left). Within a tetramer, two pairs of monomers (Mol I/II and Mol III/IV; Figure 18A, left) interact intimately, burying 7343 Å² of combined surface area upon association. Two of these close dimers further associate into tetramers via less extensive interactions between Mol I and Mol III, *viz.* Mol II/Mol IV (1891 Å² combined surface area buried in each contact), and between Mol I and Mol IV, *viz.* Mol II/III (276 Å² combined surface area buried in each contact; Figure 18A, left). The tetramers are held together by the formation of a short antiparallel coiled-coil between the $\alpha 2$ helices of Mol I and III (Mol II and IV), which is cross-strutted by helices $\alpha 1$ (Figure 18B). In contrast, *afuP_{Ser}CysS*, NifS relatives, or *synC*-DES lack the surface-exposed N terminus (elements I in Figure 16, C and D). Consistently, all these latter proteins exist as dimers. Figure 18C shows the electrostatic potential mapped to the surface of a *mmuSecS*^{elast} tetramer. Large patches of positive charge (blue) are visible, consistent with the overall basic pI of 8.3 calculated for the protein. In particular, the funnel leading to the active site is strongly positively charged. Therefore, the surface properties of *mmuSecS*^{elast} appear to be designed to contact the sugar-phosphate backbone of tRNA^{Sec} at multiple positions.

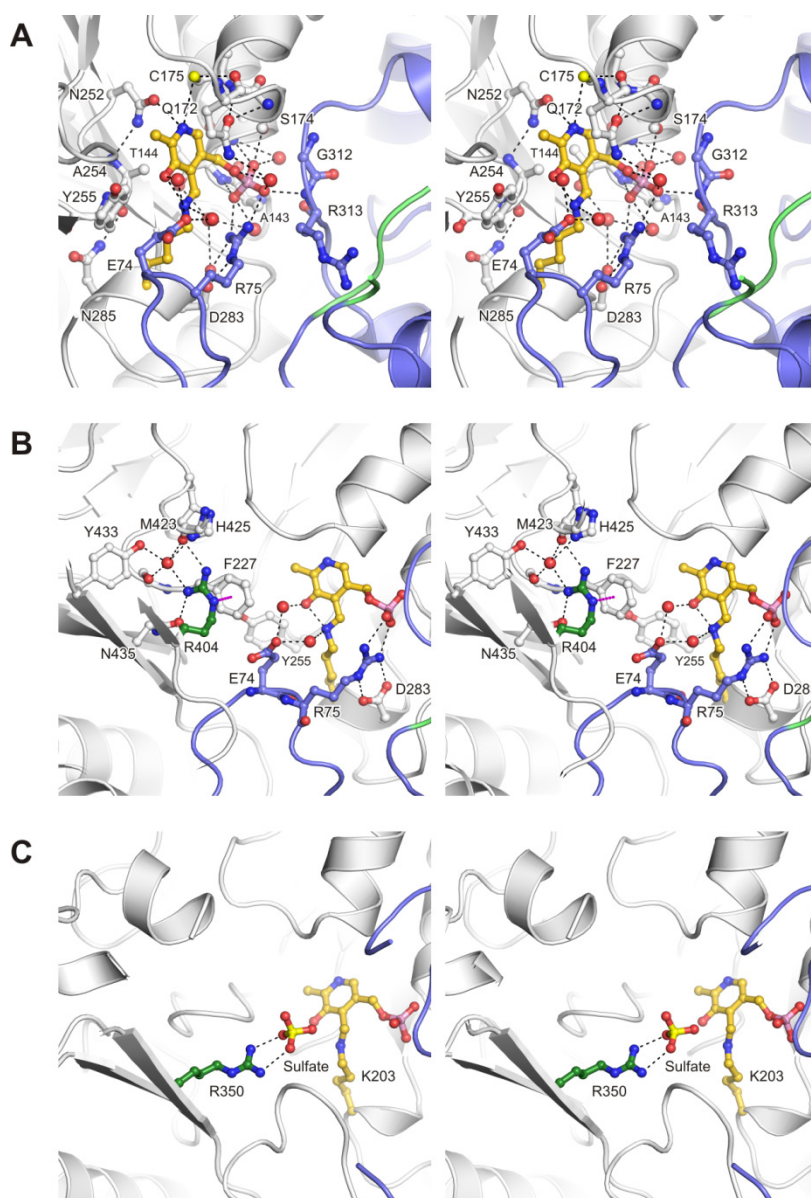


Figure 17. PLP environment. (A), stereo ribbon plot depicting the anchoring of PLP by two protomers of a close dimer in native *mmuSecS*^{elast}. The backbone of the molecule binding PLP in *cis* is shown in light gray, and the backbone of the molecule binding in *trans* is in steel blue. The base of the disordered loop (P-loop; see also Figure 19) is in green. The Lys-284-bound PLP and selected residues from both protomers are shown in ball-and-stick representation and are colour-coded by atom type (carbon, as the respective protein backbone; gold, carbon of Lys-284-PLP; blue, nitrogen; red, oxygen; yellow, sulphur; pink, phosphorus). Red spheres designate water molecules. Dashed lines indicate hydrogen bonds and salt bridges. The view is rotated 30° about the vertical axis compared with the front view in Figure 16A. (B), stereo ribbon plot of the remodelled neighbourhood of PLP. Selected residues are shown in ball-and-stick representations. Green, carbon of Arg404. Other colours are as in (A). Black dashed lines indicate hydrogen bonds and salt bridges. The magenta dashed line shows a cation- π stack of the side chains of Arg404 on Phe227. The view is rotated 90° about the vertical axis compared with the front view in Figure 16A (C), stereo ribbon plot of the PLP neighbourhood in a *T. maritime* NifS-like protein (PDB ID: 1EG5 (Kaiser *et al.*, 2000)) after global superimposition on *mmuSecS*^{elast}. The view and colour-coding are the same as in panel B. Arg350 of *T. maritime* NifS is the equivalent of Arg404 in *mmuSecS*^{elast}. In the absence of a substrate it binds to a sulphate ion mimicking a α -carboxylate.

Consistent with this view, was observed avid binding of anions to *mmuSecS*^{elast}. After soaking with NaI, was located 62 iodide ions per protomer bound to the surface of *mmuSecS*^{elast}

(Figure 18D). One of these positions was always occupied by a chloride ion in structures not treated with iodide (not shown).

4.2.4 The PLP cofactor is tightly anchored by non-canonical contacts to both protomers of a close dimer

A close dimer exhibits two identical active sites at the protomer interfaces pinpointed by a PLP cofactor (Figure 17A). Because the PLP was refined at full occupancy, leaving no residual difference density, all four subunits of a tetramer bear a cofactor. This situation is different from *afuPSerCysS*, where only one of two potential active sites in a dimer was equipped with PLP (Fukunaga *et al.*, 2007). In this work refer to the PLP attached to Lys284 of a reference molecule and the surrounding active site as “cis”; the PLP attached to the opposite protomer and its surrounding active site are referred to as “trans”.

In *mmuSecS*^{elast}, both monomers of a close dimer contribute side chains for PLP binding in an active site (Figure 17A). Apart from the covalent linkage to Lys284, PLP is additionally bound via multiple hydrogen bonds and electrostatic and van der Waals interactions in *cis*. These interactions exclusively involve residues from domain 2. The PLP phosphate group is positioned over the N terminus of helix $\alpha 5$, interacting favourably with the helix macro-dipole and engaging in hydrogen bonds to the backbone amides of Thr144 and Gly145. The pyridine nitrogen maintains hydrogen bonds to the side chains of Cys175 and Asn252. Asn252 is at variance with the vast majority of fold type I enzymes, in which an Asp at the equivalent position is the only strictly conserved residue apart from the PLP-bound Lys (Schneider *et al.*, 2000). An Asn is expected to support the electron sink character of the pyridine ring less than an Asp, possibly demanding a good leaving group such as a phosphate on the β -carbon. Interestingly, in *afuPSerCysS*, in which phosphate is also the leaving group, the pyridine nitrogen is again bound to an Asn (Figure 15) (Fukunaga *et al.*, 2007). The PLP pyridine ring of *mmuSecS*^{elast} is sandwiched between the side chains of Gln172 and Ala254 on the re and si faces, respectively. Archaeal SecS enzymes feature a His in place of Gln172 (Figure 15). A similar His in a NifS-like protein from *Thermotoga maritima* has been discussed as a tunable acid-base catalyst in the reaction mechanism (Kaiser *et al.*, 2000). Thus, the enzymatic mechanisms of archaeal and eukaryotic SecS may differ in detail.

Interactions with PLP in *trans* (Figure 17, A and B) involve residues from the SecS-specific insertion in domain 1 (element II) and from the short element III connecting domain 2 and the second part of domain 1 (Figure 16, A and B).

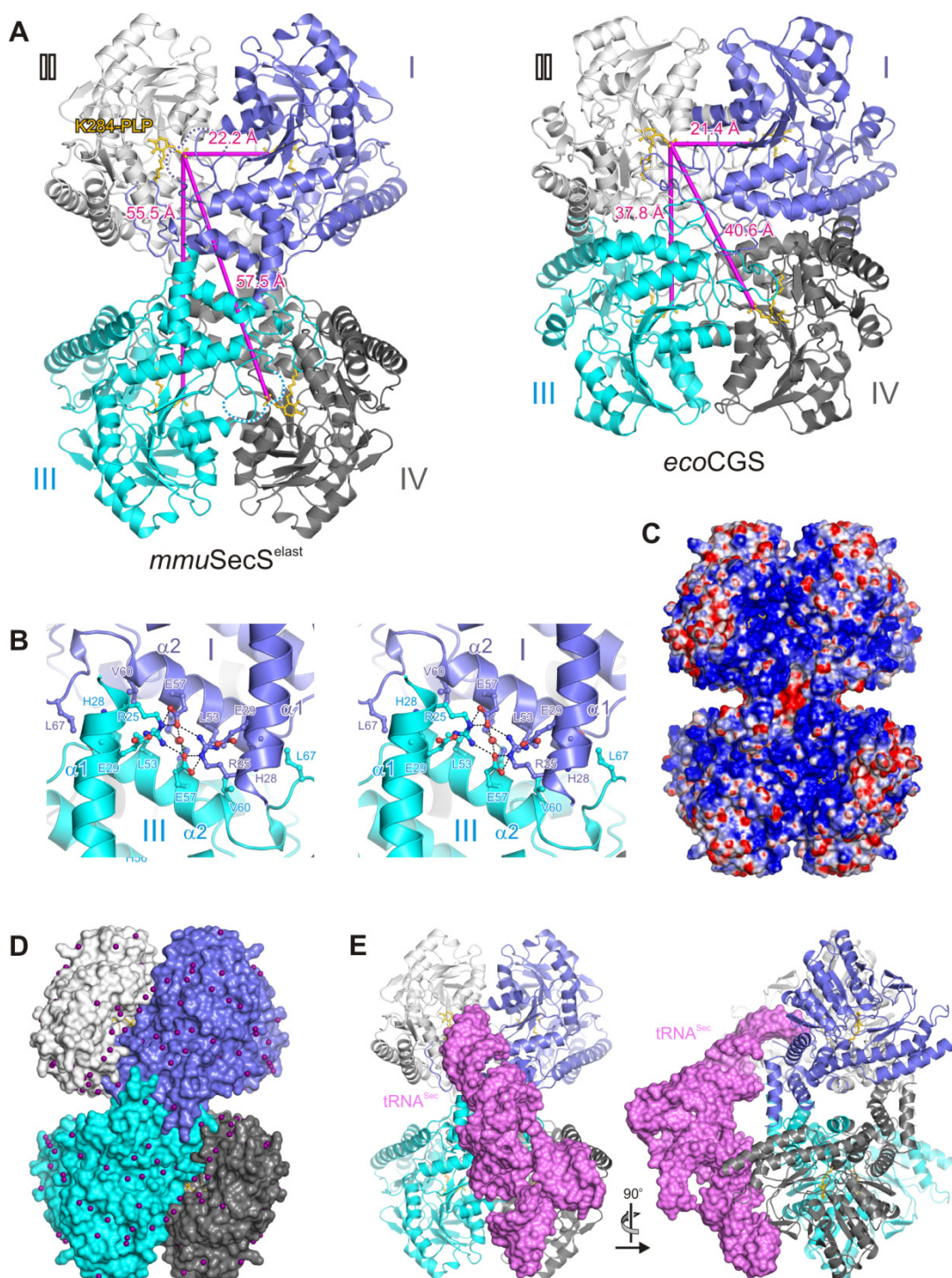


Figure 18. The *mmuSecS^{elast}* tetramer. A, left, overview of the *mmuSecS^{elast}* homotetramer, comprising a dimer of dimers (222 symmetry). Protomers (Mol I-IV) are shown in different colours. Distances between active sites (PLP phosphates) are indicated (magenta lines). Right, overview of the *E. coli* cystathionine γ -synthase (*ecoCGS*) homotetramer (PDB ID: 1CS1 (Clausen *et al.*, 1998)) drawn to scale and in the same orientation as the *mmuSecS^{elast}* tetramer. (B), stereo ribbon plot of the N-terminal tetramerization motif. Selected residues are shown as ball-and-stick representations and colour-coded by atom type. Dashed lines indicate hydrogen bonds and salt bridges. (C), electrostatic surface potential of the tetramer (calculated with PLP omitted). Blue, positive charge; red, negative charge. Large positively charged patches, in particular around the active site funnel, can be discerned. Lys284-bound PLP moieties are shown in gold at the bottom of the active sites funnels. (D), iodide ions (purple spheres) located in the structure of *mmuSecS^{elast}* soaked in 0.5 M sodium iodide. (E), hypothetical model of tRNA^{Sec} binding to *mmuSecS^{elast}*. No local conformational adjustments or energy minimizations were conducted in either the protein or the tRNA to optimize the fit.

Both elements primarily interact with the PLP phosphate. Arg75 (originating from the first loop of element II) is deposited on the phosphate side of the PLP pyridine ring where its side chain can engage in two charged interactions with anionic phosphate oxygens. Arg75 is appropriately positioned by a double salt bridge interaction with Asp283 from the *cis* protomer (Figure 17, A and B). The preceding residue of element II, Glu74, comes to lie on the opposite side of the pyridine ring and is connected via water molecules to the C3 hydroxyl group of PLP and the nitrogen of the Schiff base. It is kept in place by van der Waals contacts to Tyr255 of the *cis* protomer (Figure 17, A and B). In addition, the backbone nitrogen of Arg313 (from element III) hydrogen bonds to an anionic phosphate oxygen of the PLP.

In other fold type I PLP-dependent enzymes, element III is often significantly longer than in SecS (Figure 16, C and D) and provides additional residues for binding the PLP phosphate. In contrast, the region corresponding to element II in *afuPSerCysS* is much shorter than in SecS (21 *versus* 47 residues) and is completely disordered in the structure (Fukunaga *et al.*, 2007), failing to provide stable PLP anchoring (Figure 16C). In *synC-DES*, the equivalents of helices $\alpha 2$ and $\alpha 4$ are longer and place a very short element II, which is suspended between them, remote from the *trans* PLP and close to domain 3 of the other protomer (Figure 16D). Thus, in other fold type I enzymes, element II serves to reinforce dimerization but does not contribute directly to the active sites. Instead, in *afuPSerCysS*, NifS relatives and *synC-DES* a portion of the N terminus is positioned between domains 1 and 3 and harbors residues, which in some cases contact the *cis* PLP (Figure 16, C and D). In *mmuSecS*^{elast}, the first loop of the long element II replaces this N-terminal part in *trans* (Figure 16A).

In contrast to the situation within a close dimer, there is no cross-communication at the active sites between molecules belonging to the two different close dimers of a tetramer. This situation suggests that the close dimers are sufficient to provide the chemical microenvironment required for catalysis.

4.2.5 Binding of phosphate triggers disorder-order transition in an active site loop

Substrates of SecS contain a number of phosphates or phosphate-related groups, such as the phosphodiester backbone of tRNA^{Sec}, the γ -phosphate of the PSer moiety, and SeP. It was reasoned that phosphate could mimic binding of either of these groups at the active site of SecS and determined the crystal structure of *mmuSecS*^{elast} after soaking crystals for 30 s in 0.5 M phosphate buffer. Strikingly, was observed that a phosphate (P1 in Figure 19) is cradled in

the second loop of the domain 1 insertion (residues 98-104 of element II; green in Figure 19), which was previously disordered. Upon phosphate binding, this loop contracts and covers part of the *trans* active site (Figure 19, A and B). Arg313 originating from the loop that connects domain 2 to the second part of domain 1 (element III in Figure 16, A and B) forms the base of the P1 phosphate binding site. To bind the P1 phosphate, Arg313 and Gln105 (neighbouring the previously disordered element) are profoundly repositioned (Figure 19, C and D). In addition, the P1 phosphate interacts directly as well as via a water bridge with the side chain of Arg97 and with the side chain and backbone of Ser98, both of which were disordered in the absence of phosphate (Figure 19, C and D).

A poly-dentate anion appears to be required to engage in the observed interactions and elicit the structural transition. In agreement with this notion, a similar disorder-order transition was observed upon soaking with sulphate ions (not shown), whereas mono-dentate anions did not evoke any such change. For example, 250 mM NaCl were present in all crystallization and soaking experiments, and we did not observe any conformational changes in the structure soaked with an additional 0.5 M sodium iodide (Suppl. Table 2). Therefore, the loop between Gly96 and Pro106 of element II was referred to as the “phosphate loop” (P-loop).

Upon binding of the P1 phosphate, the P-loop closes off part of the active site of *mmuSecS*^{elast} (Figure 19, A and B). It is likely that active site closure accompanies catalysis and may serve, *e.g.* to locally exclude bulk water. Therefore, the 3'-end of tRNA^{Sec} most likely gains access to the PLP cofactor from the side opposite the P-loop (arrow in Figure 19B). It was observed a second phosphate binding site remote from the active site (P2 in Figure 19B). This phosphate location could indicate a site of contact to the tRNA^{Sec} phosphodiester backbone. However, in contrast to the P-loop and Arg313 (see below), residues contacting the P2 phosphate (Arg199, His368) are not conserved in SecS orthologs (Figure 15). It is possible that the binding of the P2 phosphate is not functional and restricted to mouse and a few other SecSs.

4.2.6 Phosphate binding is mediated by conserved and SecS-specific residues that are essential for selenocysteine synthesis

The P-loop is a highly conserved element of eukaryotic and archaeal SecS, which is lacking from related enzymes with a different function (Figure 15). In particular, Arg97, Ser98, and Gln105 from the P-loop as well as Arg313 from element III, which directly contact the P1 phosphate, are strictly conserved among SecS enzymes but not beyond.

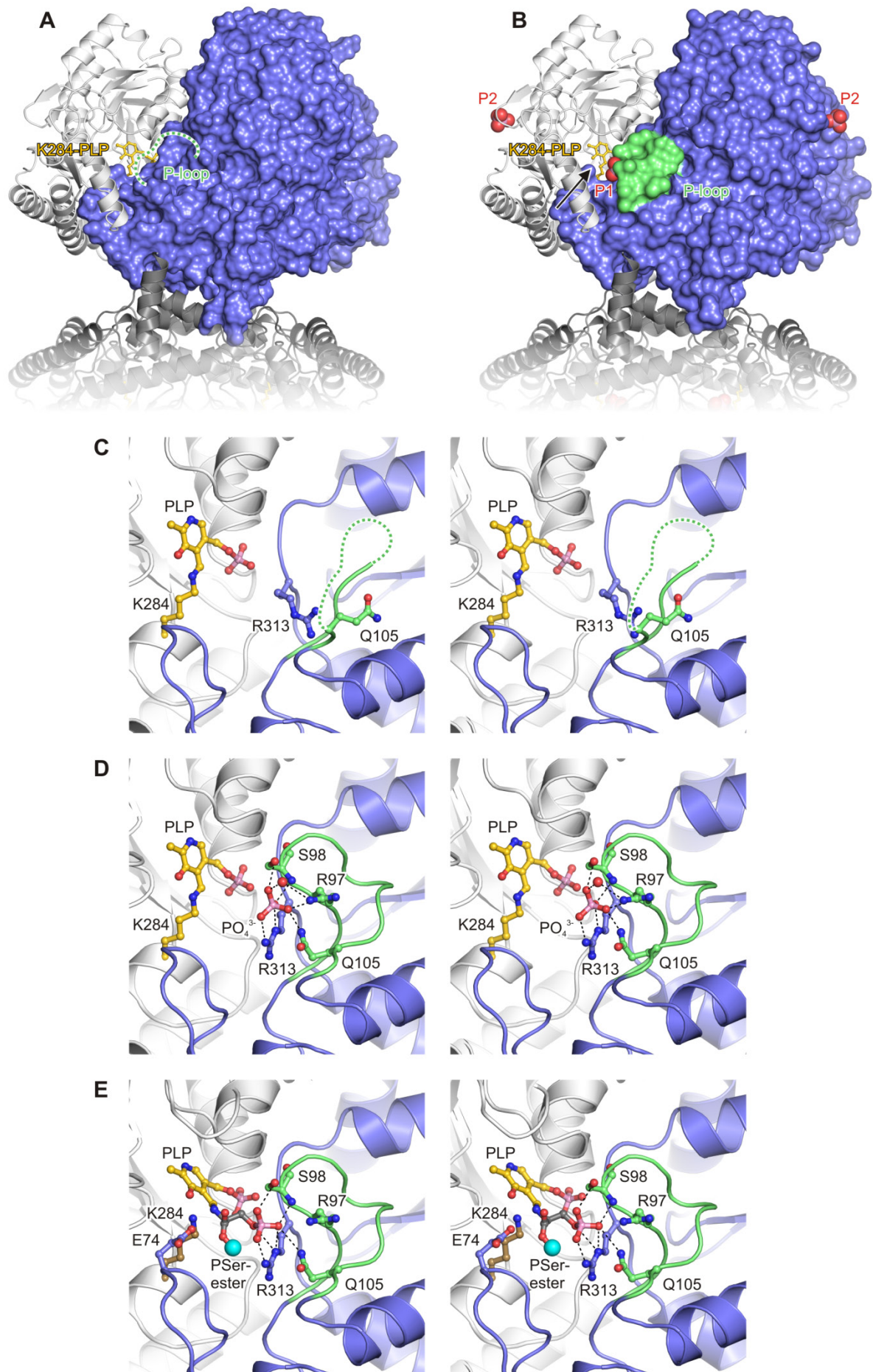


Figure 19. Phosphate-induced conformational changes. (A), native *mmuSecS*^{elast} close dimer with one molecule shown as a light gray ribbon and the other as a blue surface; the other protomers of the tetramer are shown as gray ribbons fading out. Residues 98-104 (dashed line) of the P-loop (green) are completely disordered. A Lys284-bound PLP is shown in gold ball-and-stick representation. The view is the same as in Figure 16A. (B), upon binding of the P1 phosphate, the previously disordered region of the P-loop becomes completely ordered and closes part of the active site like a lid (green surface). Bound phosphates are shown as space-filling models (red, oxygen; pink, phosphorus). The P2 phosphate may indicate a distal site of attachment of a portion of the tRNA^{Sec} backbone. The arrow indicates the likely access path of the tRNA^{Sec} 3'-end. (C), stereo ribbon plot of the active site formed around the Lys284-bound PLP of a reference molecule (light gray backbone) with participation by the second protomer of the close dimer (steel blue backbone). The partially disordered (dashed line) flexible P-loop of the second molecule is shown in green. The Lys284-bound PLP and selected residues are shown as ball-and-stick representations and are colour-coded by atom type as before. Note that the side chains of Gln105 and Arg313 are turned away from the Lys284-PLP moiety. The view is rotated 30° around the vertical axis compared with the views in panels (A) and (B). (D), stereo ribbon plot of the same active site region after binding of the P1 phosphate. Gln105 and Arg313 change their positions and orientations drastically to engage in hydrogen bonds *viz.* salt bridges to the phosphate oxygens. In addition, residues Arg97 and Ser98 of the P-loop engage in phosphate binding, leading to a complete ordering of the P-loop. Orientation and colour-coding are as in panel (C). Dashed lines indicate salt bridges and hydrogen bonds. (E), model for binding of P-Ser esterified at the α -carboxyl (cyan sphere). Without major adjustments of the protein matrix, the γ -phosphate can be cradled in the P-loop similar to the free phosphate in (D). Note that Glu74 closely approaches the α -carboxy ester.

These observations are consistent with the idea that the P-loop carries out an essential function that is specific for SecSs.

To directly probe the importance of residues contacting the P1 phosphate, we generated mutant *mmuSecS* proteins in which Gln105 was changed to Glu (*mmuSecS*^{Q105E}) or Arg313 was changed to Ser or Glu (*mmuSecS*^{R313S}; *mmuSecS*^{R313E}). All mutants were expressed as soluble proteins in *E. coli*, migrated as tetramers in gel filtration, and exhibited a PLP complement comparable with that of the wild type protein, as indicated by their absorption maxima at 334 nm (ketimine form) and 418 nm (aldimine form; not shown). We tested the mutants for their ability to convert P-Ser-tRNA^{Sec} to Sec-tRNA^{Sec}. Strikingly, each of the point mutants severely corrupted the activity of *mmuSecS* (Figure 20A). Although the *mmuSecS*^{Q105E} and *mmuSecS*^{R313S} mutants still exhibited activities of about 50 and 30 % of the wild type protein, respectively, the R313E exchange rendered *mmuSecS* virtually inactive (Figure 20A). The more severe effect with the negatively charged Glu compared with the neutral Ser in place of Arg313 directly supports the functional relevance of phosphate binding at this position. These results directly link the ability of *mmuSecS* to bind a phosphate or a related group via induced fit of the P-loop to its catalytic competence. Based on these observations, it was suggested that the P1 phosphate mimics binding of a substrate or of a functional portion of a substrate.

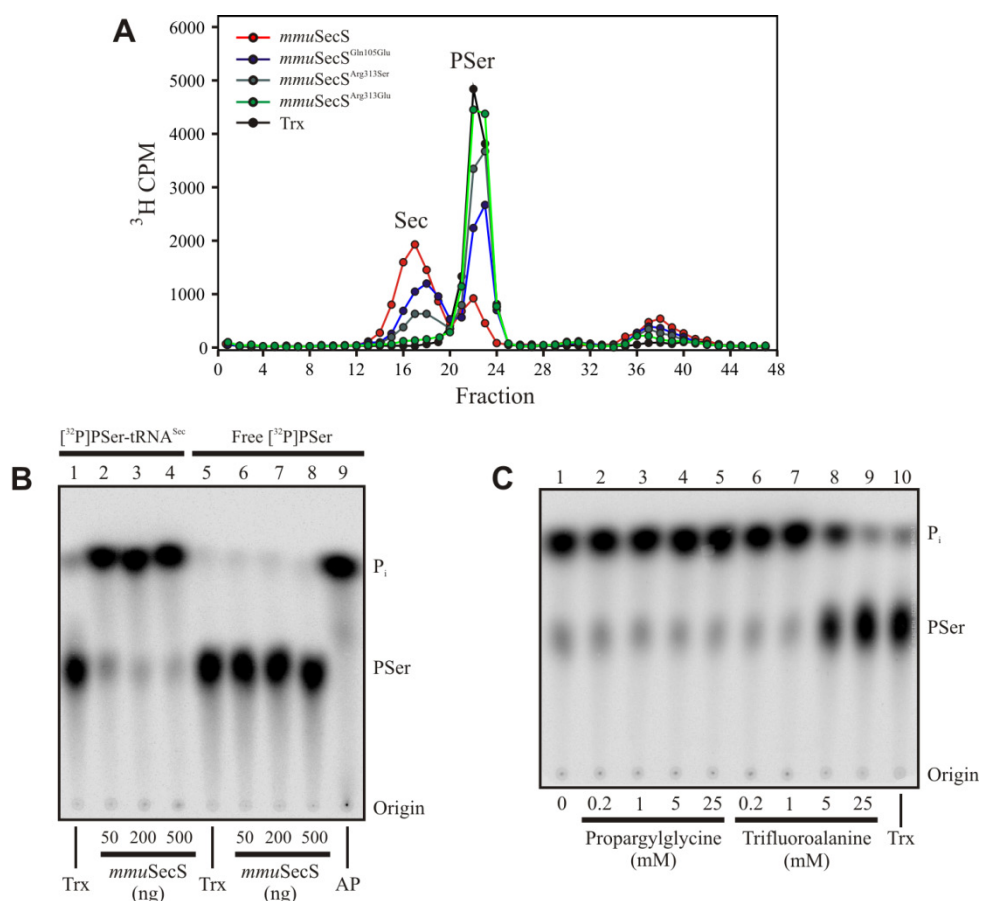


Figure 20. Functional analysis of *mmuSecS*. (A), mutational analysis of *mmuSecS*. Conversion of $[^3\text{H}]\text{PSer-tRNA}^{\text{Sec}}$ to $[^3\text{H}]\text{Sec-tRNA}^{\text{Sec}}$ by *mmuSecS* (red), *mmuSecS*^{Q105E} (blue), *mmuSecS*^{R313S} (dark green), and *mmuSecS*^{R313E} (light green). Trx (black) served as a negative control. (B), dephosphorylation activity of *mmuSecS*. Dephosphorylation of $[^{32}\text{P}]\text{PSer-tRNA}^{\text{Sec}}$ (lanes 1-4) and $[^{32}\text{P}]\text{PSer}$ (lanes 5-9) in the presence of varying amounts of *mmuSecS* is shown. Trx was used as a negative control (lanes 1 and 5), and alkaline phosphatase (AP) was used as a positive control for the generation of free $[^{32}\text{P}]\text{P}_i$ (lane 9). (C), inhibition studies of *mmuSecS*. $[^{32}\text{P}]\text{PSer-tRNA}^{\text{Sec}}$ dephosphorylation with *mmuSecS* (lane 1) in the presence of the mechanism-based inhibitors PG (lanes 2-5) and $\text{F}_3\text{-Ala}$ (lanes 6-9). Trx (lane 10) served as a negative control. Functional analysis of *mmuSecS* was performed by D. Hatfield and colleagues (department of biochemistry, NIH, USA).

4.2.7 The P-loop could serve as a binding site for the PSer γ -phosphate and SeP

We scrutinized the possibility that the γ -phosphate of $\text{PSer-tRNA}^{\text{Sec}}$ could be bound by the P-loop. To this end was modelled the structure of an external aldimine comprising PLP in a Schiff base linkage to a PSer esterified at the α -carboxylate. For modelling, was superimposed the structure of PSer-aminotransferase in complex with the substrate mimic α -methyl-L-glutamate (Hester *et al.*, 1999) onto the *mmuSecS*^{elast}-phosphate structure. The α -methyl-L-glutamate moiety with a α -carboxy ester of PSer was replaced, retained all side chain conformations as observed in the *mmuSecS*^{elast} structure in complex with phosphate, and allowed the PLP moiety to adopt a slightly more inclined orientation (Figure 19E). Even without adjustments of the protein matrix, the γ -phosphate of the PSer ester could be accommodated approximately at the P1 phosphate position.

These results suggest that the P-loop could serve to bind the γ -phosphate of P-Ser-tRNA^{Sec}. The similarity of phosphate and SeP suggests that the P1 phosphate could also mimic binding of the co-substrate SeP to the P-loop. Binding of the P-Ser moiety and SeP to the P-loop could occur sequentially and is not mutually exclusive (see “Discussion”). In contrast, it was not possible to fit a phosphate from the backbone of tRNA^{Sec} without clashes in the position of the P1 phosphate at the P-loop, consistent with suggestion above that the tRNA^{Sec} 3'-end approaches the active site distal to the P-loop (Figure 19B).

4.2.8 SecS discriminates against free *O*-phospho-*L*-serine

Free P-Ser is produced, for example, as an intermediate in the biosynthesis of serine by transamination from 3-phosphohydroxypyruvate. Therefore, SecS should be safeguarded against using free P-Ser as a substrate. This notion was tested by attempting to dephosphorylate free P-Ser with *mmu*SecS. Indeed, based on results of our collaborators (Dr. Hatfield and colleagues (department of biochemistry, NIH, USA).) the enzyme proved to be completely unreactive with respect to the free amino acid (Figure 20B).

The active site of *mmu*SecS^{elast} was inspected for possible filtering devices. Typically, PLP-dependent enzymes of the fold type I deploy a positively charged Arg in the neighborhood of the PLP to bind the negatively charged α -carboxylate of an amino acid substrate. In the absence of a substrate, a sulphate or phosphate group often binds at an equivalent position as the α -carboxylate (Kaiser *et al.*, 2000). The α -carboxylate binding Arg originates in *cis* from the β -sheet in domain 3. An equivalent Arg is also strictly conserved in SecS orthologs (Arg404; Figure 17B). However, in *mmu*SecS^{elast}, its side chain is turned away from the PLP. Arg404 interacts instead by a cation- π interaction with Phe227 (another residue strictly conserved only in the SecS orthologs; Figure 15). It is additionally fixed by hydrogen bonds to the side chain of Asn435 and the backbone carbonyl of Met423 and by water-mediated hydrogen bonds to the backbone carbonyl of Ala228 and to side chains of His425 and Tyr433 (Figure 17B). The side chain of Arg404 is thereby stably tugged away, since it does not change orientation in the presence of even 0.5 M phosphate (as seen in SecS phosphate-soaked structure), suggesting that Arg404 is not involved in binding of a α -carboxylate.

As pointed out above, *mmu*SecS^{elast} harbors Glu74 in *trans* next to the PLP moiety (Figure 17, A and B). Interestingly, Glu74 occupies the same spatial position as the equivalent of Arg404 in other fold type I enzymes (Figure 17, B and C). This observation and our model of an external aldimine of *mmu*SecS with a P-Ser ester (Figure 19E) demonstrate that Glu74

would strongly disfavour productive placement of a substrate with a free (negatively charged) α -carboxylate. In P $\text{Ser-tRNA}^{\text{Sec}}$, however, the α -carboxylate of P Ser is esterified to the 2 - or 3'-hydroxyl group of the 3-terminal adenosine and does not carry a negative charge. Therefore, we expect that Glu74 acts as a substrate filter by repelling compounds with a negatively charged α -carboxylate.

4.2.9 The inhibition profile of SecS resembles that of β -lyases

To gain additional insight into the reaction mechanism of SecS, we tested whether the dephosphorylation activity of *mmu*SecS is inhibited by the mechanism-based inhibitors PG and F₃-Ala. PG preferentially inhibits PLP-dependent enzymes such as cystathionine γ -lyase, which mediate lyase reactions at the γ -carbon (Lacourciere and Stadtman, 1998). In contrast, F₃-Ala preferentially inhibits enzymes such as cystathionine β -lyase, which mediate replacement of a substituent at the β -carbon (Clausen *et al.*, 1996). Strikingly, *mmu*SecS is unaffected by up to 25 mM PG (more than 12,500-fold excess over *mmu*SecS active sites; Figure 20C). In contrast, partial enzyme inhibition was detected in the presence of 5 mM F₃-Ala and above (Figure 20C). These inhibition profiles are in agreement with SecS catalyzing a β -replacement reaction with a high specificity. Furthermore, the data suggest that the lytic half-reaction of SecS follows the cystathionine β -lyase scheme, strongly supporting aminoacrylyl-tRNA^{Sec} as an intermediate (Xu *et al.*, 2007a)

Although our data do not allow us to derive exact inhibition constants, we note that compared with *E. coli* cystathionine β -lyase (Clausen *et al.*, 1996), the inhibition of *mmu*SecS by F₃-Ala is weak. For example, in a comparable setup, the halftime for inactivation of *E. coli* cystathionine β -lyase by 1 mM F₃-Ala was less than 2 min (Hubert *et al.*, 1998), whereas *mmu*SecS is not measurably affected under these conditions (Figure 20C). This observation is consistent with the suggested substrate discrimination by Glu74, since F₃-Ala exhibits a free α -carboxylate and is, therefore, expected to be discouraged from forming an external aldimine.

4.3 Crystal structure analysis of tRNA^{Sec}

4.3.1 Co-crystallization of tRNA^{Sec} with PSTK

tRNA^{Sec} is a key molecule in Sec biosynthesis and its incorporation during translation. All steps of Sec maturation take place on cognate tRNA^{Sec} which serves as recognition scaffold for PSTK and SecS. Then Sec-tRNA^{Sec} is recognized and delivered to the ribosome

only by special elongation factor EFSec and not by EF-Tu, like other tRNAs. It is interacting not only with single proteins like SerRS, PSTK, SecS but also was found in higher order complexes (Small-Howard *et. al.*, 2006). Thus, tRNA^{Sec} must have unique identity elements compared to other tRNAs. Therefore crystal structure analysis of tRNA^{Sec} definitely needed for understanding all its features.

In general crystallization of RNA molecules is difficult as they have high negative charge of phosphate backbone which produced repulsive effects. Therefore for crystallization of mouse tRNA^{Sec} was applied by two strategies: co-crystallization of tRNA^{Sec} in complex with one of the protein binding partners or alone.

For tRNA^{Sec} co-crystallization it is necessary to find the complex with the highest affinity. Synthetic full length tRNA^{Sec} from mouse was produced by *in vitro* transcription with T7 RNA polymerase and used for analytical gel filtration analysis of complex formation with several different binding factors: SerRS, PSTK and SecS. Based on experiments, PSTK showed the highest affinity to tRNA^{Sec}, whereas complexes with SecS or SerRS were hardly detectable (Figure 21).

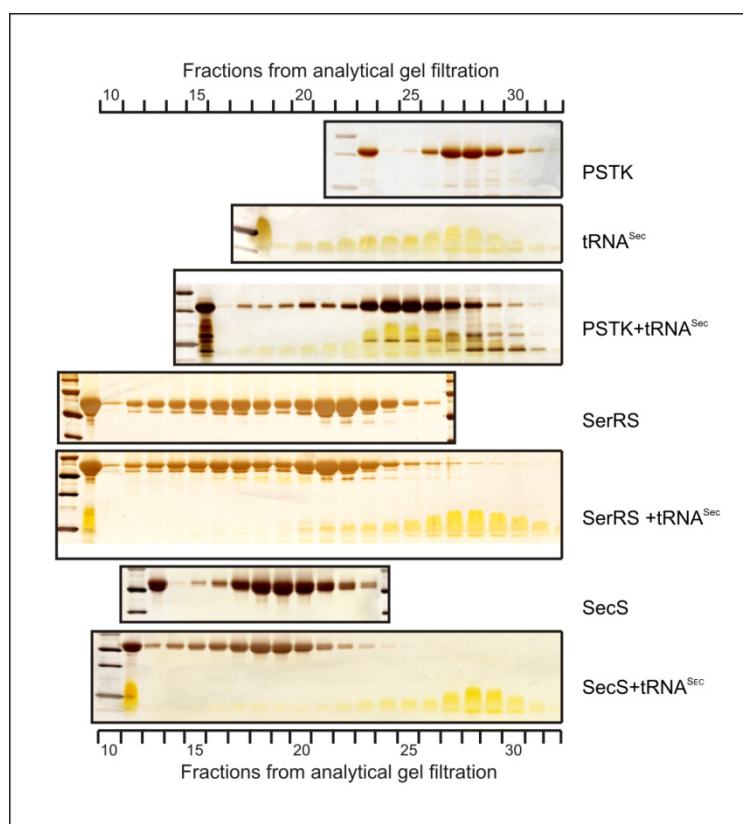


Figure 21. Interaction studies between tRNA^{Sec} and PSTK, SecS, SerRS. Fractions after analytical gel filtration were analyzed on PAGE SDS. Gels were stained first by coomassie and then by silver-containing solution.

Complex between tRNA^{Sec} and PSTK was assembled, purified and subjected to crystallization. Obtained crystals were of low diffraction quality and contained only the RNA compound.

To test the possibility that PSTK contains flexible regions that could interfere with the crystallization process, the kinase alone and in complex with tRNA^{Sec} were treated with various proteases. After elastase treatment several stable bands resistant to further digestion were observed (Figure 22).

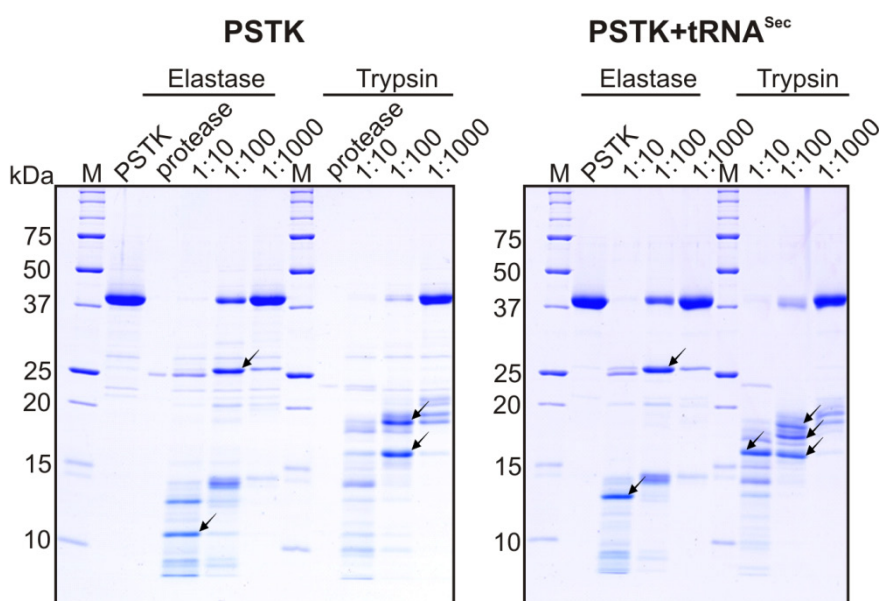


Figure 22. Limited proteolysis of PSTK. PSTK was treated by elastase and trypsin at different protease solution dilutions (1:10, 1:100 and 1:1000). After 30 min of incubation degradation products were analyzed on PAGE SDS. M – protein molecular weight markers. Amino acid sequence of protease-resistant protein bands (marked with arrow) were sent for tryptic mass spectrometric fingerprinting analysis.

Tryptic mass spectrometric fingerprinting (performed by M. Raabe and H. Urlaub, Bioanalytical Mass Spectrometry Group, Max Planck Institute for Biophysical Chemistry, Germany) showed that the elastase-resistant fragments located within the C-terminal part. Based on this result 10 new constructs of PSTK were designed (Figure 23A). Analytical protein solubility test revealed that only three of them have reasonable solubility level (Figure 23B).

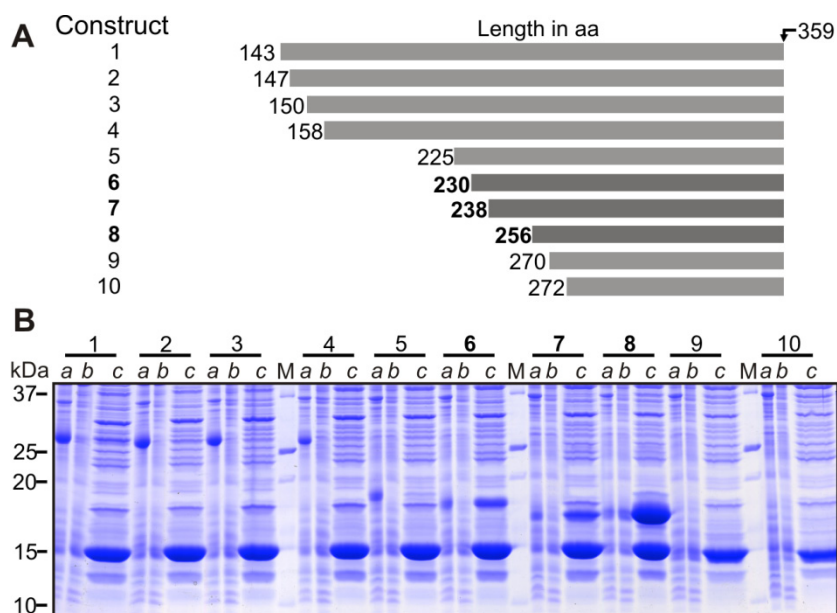


Figure 23. Solubility screening of PSTK constructs. (A) Design of PSTK truncated variants. Lines coloured grey schematically represent created PSTK constructs. (B) Expression profile and solubility level of PSTK constructs in *E.coli* cells analyzed on PAGE SDS. Three tracks on the gel *a*, *b* and *c* belongs to each analyzed construct and corresponds to crude cell lysate after expression, clarified cell lysate and lysate enriched on NI-NTA beads, accordingly. Constructs (# 6, 7 and 8, bold) with considerable solubility on (A) coloured dark grey and asterisk on (B).

The most soluble one encompassing 256-359 amino acid of PSTK (PSTK₂₅₆₋₃₅₉) was purified and tested for complex formation with tRNA^{Sec}. Truncated PSTK retained affinity to the tRNA^{Sec} (Figure 24A). Purified via gel filtration chromatography, the binary complex PSTK₂₅₆₋₃₅₉-tRNA^{Sec} was subjected to crystallization. The obtained crystals contained both compounds but exhibited low diffraction quality (Figure 24B).

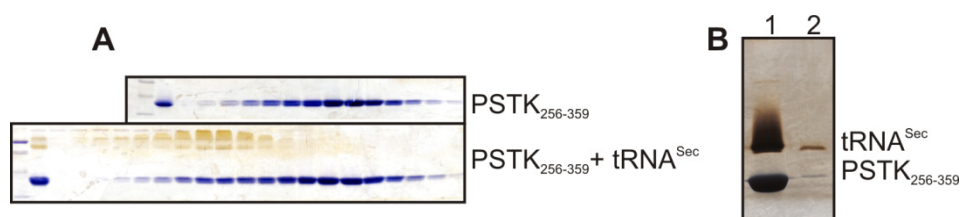


Figure 24. Complex formation between tRNA and PSTK₂₅₆₋₃₅₉. (A) Fractions after analytical gel filtration were analyzed on PAGE SDS. (B) Crystal contents of tRNA^{Sec}/PSTK₂₅₆₋₃₅₉ was analyzed on PAGE SDS. Track 1 - control (mixture of tRNA^{Sec} and PSTK₂₅₆₋₃₅₉), track 2 - dissolved crystal.

4.3.2 Crystallization of tRNA^{Sec}

In parallel crystallization of tRNA^{Sec} alone yielded many crystals derived from various crystallization conditions indicating high crystallization potency of that molecule. Unfortunately all of them showed low diffraction quality. In order to improve crystallization potential of tRNA^{Sec} alternative tRNA^{Sec} constructs were designed. Details of construct design will be described in chapter 4.4.6. In case of tRNA^{Sec} construct with deletion of nucleotides 73-76 of acceptor arm (Δ GCCA tRNA^{Sec}) were obtained reasonably diffracting crystals

belonging to P2₁, I4 and C2 space groups (below in the text the construct Δ^{GCCA} tRNA^{Sec} is referred to as tRNA^{Sec}). Initial crystals of P2₁ space group were improved after several rounds of micro seeding. In order to solve the tRNA^{Sec} structure crystals were soaked with various heavy metal salts (Table 8).

Substance	Concentration, mM	Incubation time
Ir(NH ₃) ₅ Cl ₂	10	5 min/overnight
Ir(NH ₃) ₆ Cl ₃	10	5 min/overnight
Os(NH ₃) ₄ Cl ₂	10	5 min/overnight
Os(NH ₃) ₆ Cl ₃	10	5 min/overnight
Ru(NH ₃) ₆ Cl ₃	10	5 min/overnight
CsCl	10/100	overnight
YbCl ₃	0.1/1	overnight
SmCl ₃	0.1/1	overnight
I3C	500	0.5-1 min

Table 8. Crystals of tRNA^{Sec} were derivatized at different heavy metal salt concentration and incubation time

Datasets collected from derivatized crystals did not contain a strong anomalous signal. At the same time, crystal structures of human tRNA^{Sec} alone and in complex with SecS were reported (Itoh *et al.*, 2009; Palioura *et al.*, 2009). Therefore using sections of human tRNA^{Sec} [PDB ID: 3A3A; (Itoh *et al.*, 2009)] as a search model, the structure of mouse tRNA^{Sec} was solved by molecular replacement. The final model was refined to 2.0 Å resolution (Suppl. Table 3). All 86 nucleotide bases of tRNA^{Sec} model derived from space groups P2₁ and C2 can be clearly located in the electron density map. Further in the text I refer only to the higher resolved structure derived from P2₁ space group.

4.3.3 Overall structure of tRNA^{Sec}

The tRNA^{Sec} has the canonical “L”-shape and its architecture is identical to other tRNA structures. The length of extra arm of tRNA^{Sec} resembles that of tRNA^{Leu}, tRNA^{Ser} and prokaryotic tRNA^{Tyr} (Steinberg *et al.*, 1993). Moreover the extra arm of tRNA^{Sec} is similar to tRNA^{Ser} allowing them to share SerRS for aminoacylation with serine. Our structure of murine tRNA^{Sec} at 2.0 Å revealed already known and new structural features which have not been seen in previous human tRNA^{Sec} structure at 3.1 Å resolution (Itoh *et al.*, 2009). The nucleotide sequences of mouse and human tRNA^{Sec} differ only in one position, 47c, (U in mouse and C in human). Therefore both structures can be directly compared. As in case with human tRNA^{Sec} structure (Itoh *et al.*, 2009), mouse tRNA^{Sec} consists of the 13 bp composite acceptor-TΨC helix (the ratio of acceptor and T-stem is 9/4). Whereas in canonical tRNAs the length of this element is 12 bp (7/5 secondary structure). The structure of complex tRNA^{Sec}

with SecS shows importance of 13 bp helix for recognition mechanism by SecS enzyme (Palioura *et al.*, 2009) The D-stem of tRNA^{Sec} consist of a six-bp stem and is capped with a four-nt loop (compared with a four-bp stem and a seven to eleven-nt loop in canonical tRNAs respectively). The length of the D-stem serves as identity element for PSTK (Wu and Gross, 1994). In the published structures the anticodon loop was either distorted probably due to crystal packing and most likely not physiologically relevant or was not included in the structure (Palioura *et al.*, 2009; Itoh *et al.*, 2009).

In the crystal belonging to space group P2₁ two tRNA^{Sec} molecules were found per asymmetric unit. Both molecules have different anticodon loop conformation and consist of three or five nucleotide loop. Dual conformational state of anticodon loop kept by crystal contacts with symmetry related molecule and discussed later in detail.

4.3.4 The AD linker

Two bases A8 and U9 connect the acceptor stem with the D-stem and were designated as AD linker (Itoh *et al.*, 2009). U9 participates in stacking interactions with the first base pair of the variable stem (G45:A48). The second base, A8, has a different orientation in the published structures of tRNA^{Sec} and in both tRNA^{Sec} chains A and B, derived from asymmetric unit of P2₁ crystal (Figure 25). The different position of A8 nucleotide in tRNA^{Sec} structures indicates its high flexibility. Moreover in our tRNA^{Sec} structures A8 interacts via water molecules with the first base pair of variable stem.

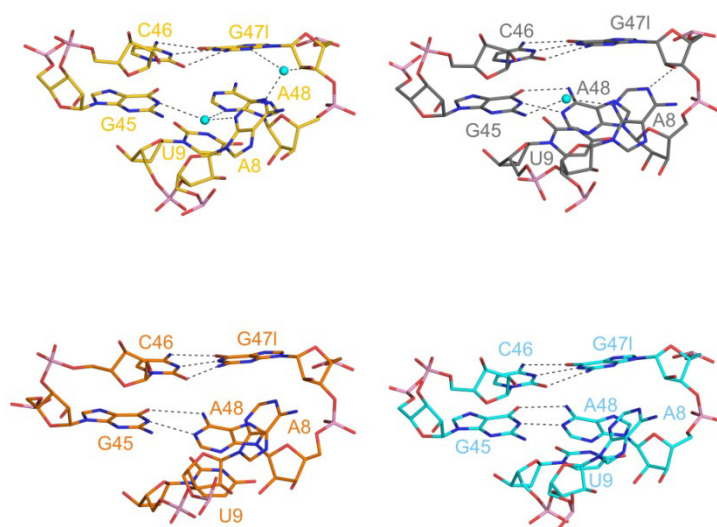


Figure 25. Tertiary interactions within the AD linker region. Drawn fragments derived from tRNA^{Sec} structures obtained in this work are colored gold and silver (chain A and chain B respectively), tRNA^{Sec} crystallized alone colored cyan (PDB ID 3A3A), tRNA^{Sec} derived from co-crystal structure with SecS colored orange (PDB ID 3HL2). Hydrogen bonds are shown as dashed lines. Water molecules represented as cyan spheres.

4.3.5 The variable arm

The variable arm adopts a stem-loop structure and consists of a 6 bp stem capped with a 4 nucleotide loop. The first base pair of the stem (G45:A48) makes a contact with the AD linker via water molecules and stacking interactions. The loop of the variable arm includes bases 47d-47g (UAGC) with U47c:G47h closing loop pair. The loop bases AGC stack on the 3'-side of the loop with their Watson-Crick functional groups turned outward. G47f and C47g formed crystal contacts based on Watson-Crick interactions with symmetry related C47g and G47f, respectively. A sharp turn of the uridine at position 47d is a characteristic structure designated as U-turn motif (Cabello-Villegas *et al.*, 2002). Structural alignment of the GAGA-tetraloop derived from 23S rRNA sarcin/ricin crystal structure (Correll *et al.*, 1999) with the loop of the variable arm (Figure 26) showed that the loop of variable arm mimics the 3-dimensional organization GNRA-tetraloops which possesses exceptional thermodynamic stabilities (Antao *et al.*, 1991). However, detailed inspection of important hydrogen bonds (H-bonds) for hyper stable tetraloop formation revealed that the distance between N3 of uridine 47d and 2OH group of cytosine 47g is too long for H-bond formation (3.9 Å in chain A or 4.3 Å in chain B). In addition, the loop-closing base pair in variable arm is GU which is less stable than GC in case of the GNRA-tetraloops. Despite of structural similarity with GNRA tetraloops, the loop of the variable arm most likely would not have hyper stability.

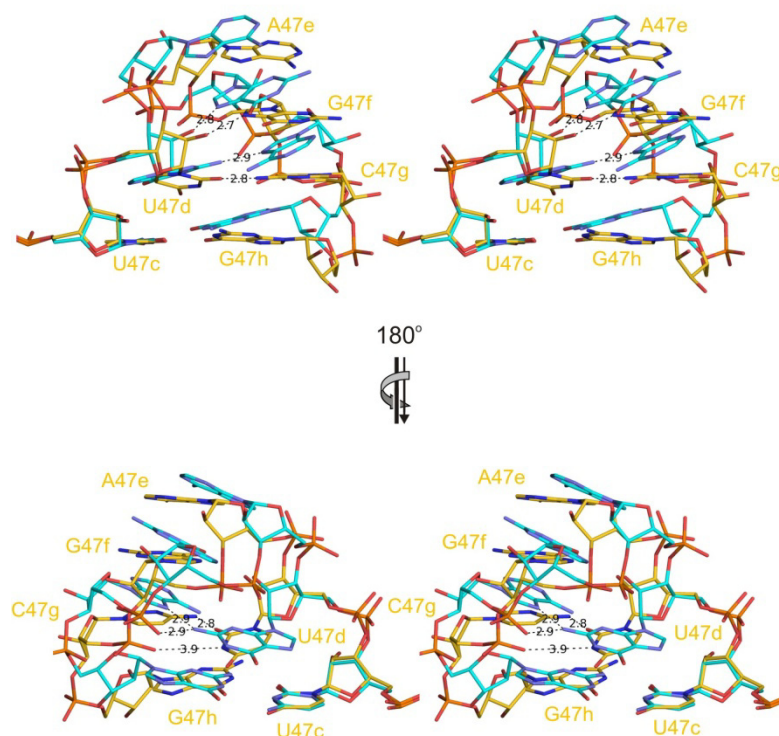


Figure 26. Structural alignment of the variable arm loop of tRNA^{Sec} and the GNRA-tetraloop. The loop of tRNA^{Sec} labeled in gold, GAGA-tetraloop from 23S rRNA sarcin/ricin crystal structure (PDB ID 483D) colored cyan. Hydrogen bonds are shown as dashed lines.

Comparison of our tRNA^{Sec} structure with the published ones showed a different orientation of the variable arm relative to the anticodon helix. Position of the first base pair of the stem (G45: A48) are identical in all tRNA^{Sec} structures. The remainder of the extra arm can adopt different conformations, from the most closed in our structure (turned toward the anticodon helix) to the most open (turned outward the anticodon arm) in tRNA^{Sec} from the cocrystal structure with SecS (Palioura *et al.*, 2009). The conformational difference of the variable loop in complex of the tRNA^{Sec} and SecS compared to our tRNA^{Sec} structure is most likely due to binding to the SecS which induced a conformational change within tRNA^{Sec}. The structure of tRNA^{Sec} from Itoh and coauthors occupy an intermediate position compared with mentioned above two structures.

4.3.6 The anticodon arm

The tRNA^{Sec} anticodon arm structure was not highlighted in detail so far. In the tRNA^{Sec} derived from cocrystal structure with SecS the anticodon loop could not be located in the electron density map (Palioura *et al.*, 2009). In another published tRNA^{Sec} structure, the anticodon loop is heavily involved in interaction with a neighbor tRNA^{Sec} molecule and adopts irregular structure (Itoh *et al.*, 2009). Although in our structure this loop also involved in crystal packing it adopts two relevant conformations (Figure 27, A and 3B).

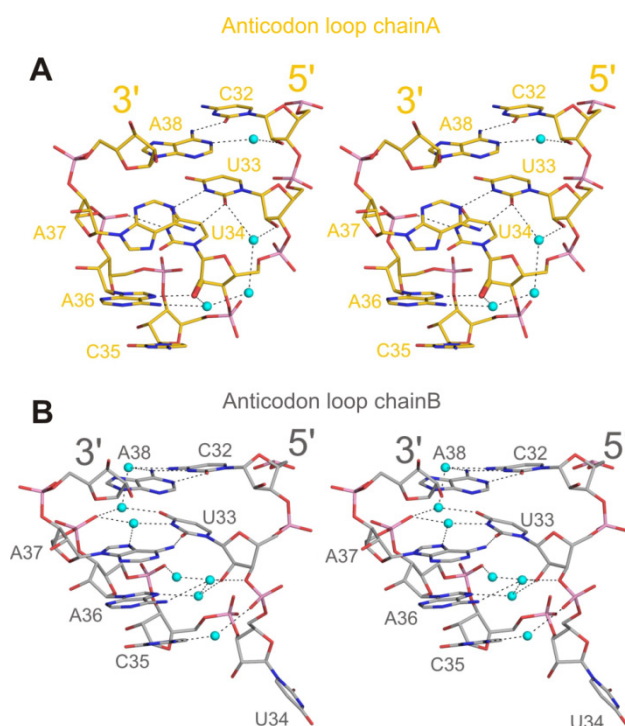


Figure 27. Anticodon loop conformation in chain A and B of tRNA^{Sec}. Stereo view of stick model of anticodon loop in chain A (A) and in chain B (B). 2'-oxygen of ribose of U34 in chain A which is undergoes 2'-O-hydroxymethylation is shown as bold stick. Water molecules represented as cyan spheres. Hydrogen bonds are drawn as dashed lines.

In chain A the anticodon stem consist of eight bp and a three-nt loop, whereas in chain B it comprise seven bp and five-nt loop. The well structured anticodon loop in chain A consists of three anticodon bases - UCA, where uridine 34 makes a sharp turn and interacts with the phosphate backbone of A 37 forming the U-turn motif. The loop bases CA partly stack on the 3'-side of the loop with their Watson-Crick functional groups turned outward. Both nucleotides participate in crystal contacts with the symmetry related chain A. The loop closes with U33:A37 reverse Watson-Crick base pair.

In the anticodon loop of chain B Uridine 34 flips out and form stacking interactions with A8 from the AD linker of the neighbor molecule in the crystal. As a consequence the base pairing of U33:A37 is altered. Namely, the A37 is turned outwards and involved in stacking interactions with A36 from the anticodon. Despite this, A37 still maintains connection with U33 via one direct H-bond and a second via water molecule at the Hoogsteen edge. Although U33 occupied a spatial position of U34 it doesn't form direct H-bond with phosphate backbone of A37 and this interaction is mediated via two water molecules. The overall structure of the anticodon loop of chain B with minor exception resembles a classical anticodon loop of tRNA^{Phe} (Klug *et al.*, 1974).

4.3.7 The influence of 2'-O-methylation of U34 on anticodon loop conformation

tRNA^{Sec} studied in frame of this work was produced by in vitro transcription without base modifications. In the literature, tRNA^{Sec} is distinguished in two isoforms based on modifications of U at the wobble position. The first contains a 5-methylcarboxymethyluridine at position 34 (mcm⁵U) and N6-isopentyladenosine (i⁶A) at position 37 and the second one has the same modifications and in addition a 2'-O-hydroxymethylated ribose at position 34 (mcm⁵Um) (Diamond *et al.*, 1993).

Based on our modeling studies, the presence of 2'-O-hydroxymethylated ribose at position 34 in the tRNA^{Sec} structure derived from chain A abolishes formation of two hydrogen bonds between O2 of U34 and N7 of A37 and via water molecule with N6 of A37 thereby destabilizing three- and favoring five-nt loop conformation. The biochemical data also revealed that the mcm⁵UmCA anticodon loop is stabilized tRNA^{Sec} in a relatively open conformation, whereas presence of the mcm⁵UCA anticodon shows a more compact tRNA^{Sec} structure (Diamond *et al.*, 1993). Modeling of i⁶A at position 37 in both three- and five-nt anticodon loops does not produce any conformational restrictions. Although it is not clear how this modification affect the loop conformation the presence of N6-isopentyladenosine at

position 37 is important for efficient translation of the Sec-codon by tRNA^{Sec} (Warner *et al.*, 2000).

NMR studies of A37N⁶-dimethylallyl modified and unmodified anticodon stem-loops of *Escherichia coli* tRNA^{Phe} revealed that absence of the modification leads to change from a five-loop to a tri-loop conformation of tRNA^{Phe} (Cabello-Villegas *et al.*, 2002). Moreover, molecular dynamics of tRNA^{Cys} shows the importance of the modified base ms²i⁶A at position 37 for disruption of base pair interaction which could be form between unmodified A37 and A38 with U32 and U33 (Alexander *et al.*, 2010). The five-bp loop conformation and the seven-bp stem are more favorable for A site requirements of the ribosome (Steitz, 2008; Vendeix *et al.*, 2008). Therefore from the data which stated above we suggest that conformation of anticodon loop in chain B reflects the conformation of the fully modified loop of tRNA^{Sec}.

4.3.8 Role of water molecules in tRNA^{Sec} structure

Water molecules are involved in many biological processes such as folding, catalysis and recognition of RNA or DNA molecules (Joachimiak *et al.*, 1994; Walter, 2007). Analysis of a number of nucleic acids X-ray structures allow to draw a conclusion that water is an integral part of RNA or DNA molecules (Westhof, 1988; Westhof and Beveridge, 1990).

It stands to mention that only in tRNA^{Sec} structure in complex with SecS a few water molecules bound to tRNA^{Sec} were found due to the limited resolution (Palioura *et al.*, 2009). In the present tRNA^{Sec} structure I was able to locate 195 water molecules for both chain A and B of tRNA^{Sec}. More than 25 % of hydrogen bonds formed by water molecules are involved with interaction of 2'-oxygens of the ribose and clearly participate in stabilization of tRNA^{Sec} structure. These interactions can be divided into three groups. The first fraction forms bridges between functional group of nucleotides and 2'-oxygens of the ribose. Example of such interaction is the 2'-oxygen of ribose to guanine N3 or cytosine O2 mediated by water molecules. Such an interaction was observed in the crystal structure of the RNA duplex [r(CCCCGGGG)]₂ earlier and well described (Egli *et al.*, 1996). The second group of water molecules participate in bridging two 2'-oxygens of the ribose from the same or neighboring molecule. In our structure we observed that 2 water molecules can bridge two neighbor 2'-oxygen of the ribose within the same duplex. The third group of water molecules was found to form bridges with 4'O of ribose or the phosphate backbone.

Around 25 % of the water molecules form a hydrogen bond with the phosphate backbone which is typical for nucleic acids (Kennard *et al.*, 1986). Half of all water molecules

participate in hydration of nucleotide bases and are bound directly to acceptor or donor groups. The binding modes are similar to those reported earlier (Egli *et al.*, 1996; Klostermann *et al.*, 1999). 47 % of bound water interacts with the guanine functional groups, while only 31 % of the residues of this construct are Gs. The fact that guanines are hydrated to a larger extent than other RNA bases, was observed previously in tRNA^{Phe} structure (Shi and Moore, 2000). Such high hydration level can be explained by the presence of more functional groups on its heterocyclic structure than in the other bases.

tRNA^{Sec} contains three GU base pairs: G5a:U67b in the acceptor stem, U12:G23 in the D-stem and G27:U43 in the anticodon stem (Figure 28). The hydration of these base pairs has a characteristic pattern which was observed earlier (Biswas *et al.*, 1997; Mueller *et al.*, 1999). This includes one water molecule in the minor groove which forms a H-bond with N2 of G and in addition two hydrogen bonds with 2'O and O2 of U.

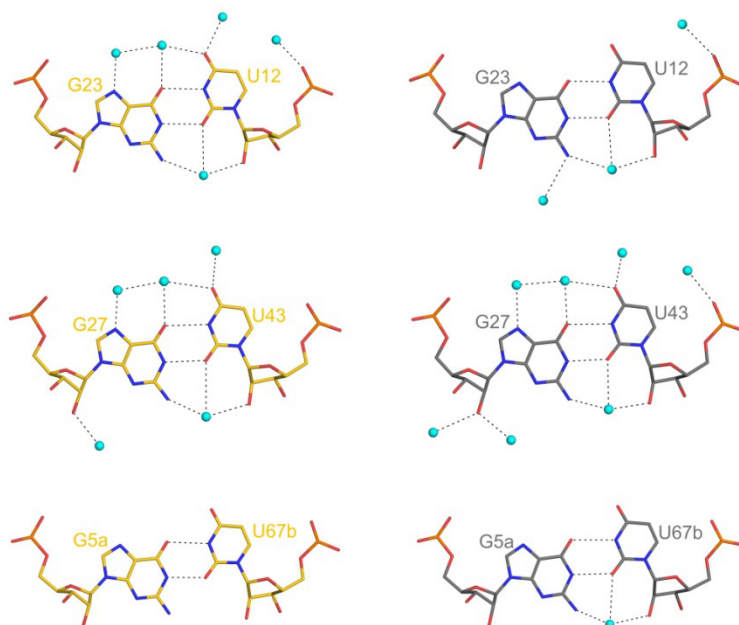


Figure 28. Hydration of GU base pairs. Hydration of GU base pairs in chain A (gold) and chain B (silver) of tRNA^{Sec}. H-bonds shown as dashed lines. Water molecules represented as cyan spheres.

From the side of major groove a water molecule forms hydrogen bonds with O4 of uridine, O6 of guanine and via neighbor molecule involved in interaction with N7. Although it is known that the GU pair forms a metal binding site, we were unable to locate any metal ions (Cate *et al.*, 1996).

4.3.9 Conformational flexibility

The crystalline polymorphism of RNA and particular of tRNA is well known (Dock *et al.*, 1987). Such polymorphism is present due to different conformational states of molecules

in a crystal. We suggest that conformational variations of tRNA^{Sec} found in our studies and compared with the published ones may represent different snapshots of conformational state of tRNA^{Sec} in solution captured by crystal contacts. Structural alignment of known tRNA^{Sec} structures clearly indicates regions of tRNA^{Sec} which undergo conformational changes (Figure 29). These are the variable loop and the anticodon loop. The remainder of tRNA^{Sec} structure seems quite rigid and not allows any deviations of backbone. The presence of flexible regions in some tRNAs was shown to be important for recognition by various enzymes involved in translation and modification processes as well (Alexandr *et al.*, 2010).

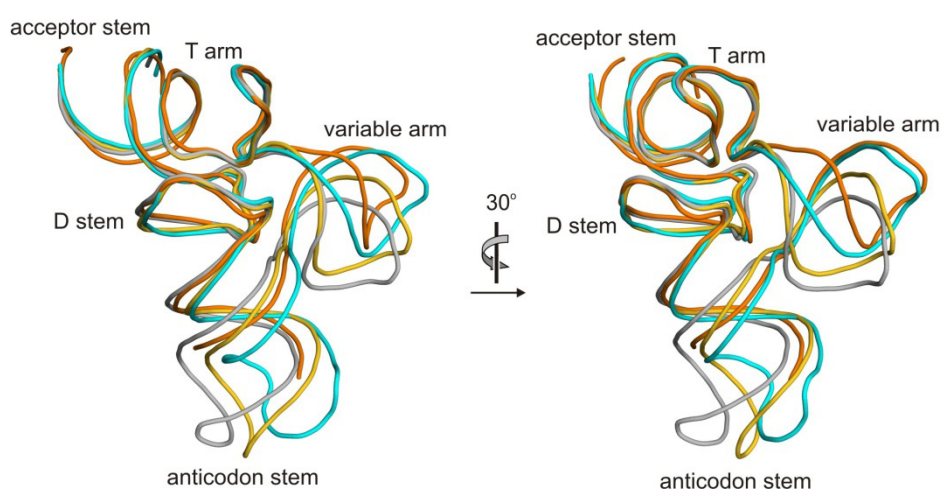


Figure 29. Structural alignment of tRNA^{Sec} crystal structures. tRNA^{Sec} crystallized alone colored cyan (PDB ID 3A3A; Itoh *et al.*, 2009), tRNA^{Sec} derived from cocrystal structure with SecS colored orange (PDB ID 3HL2; Palioura *et al.*, 2009), tRNA^{Sec} structures obtained in this study are colored gold and silver (Chain A and Chain B respectively).

The absence of tertiary interactions between the variable and the D-arms probably leads to increased flexibility of the variable arm. The presence of two H-bonds between the backbone of the variable arm and SecS (C46l:Arg271 and C46e:Ser260) in the co-crystal structure was sufficient to change the position of the variable arm compared to our structure as mentioned above. Such flexibility may be important for an induced fit binding of other factors to tRNA^{Sec}.

4.3.10 Metal binding sites

The negative charge of nucleic acids is responsible for the affinity to cations in general and particularly for divalent ions. Cations have a stabilization effect on DNA and RNA molecules due to neutralization of repulsive effect of the negatively charged phosphate groups (Record *et al.*, 1978; Manning, 1978). In order to assign divalent metal binding sites in DNA or RNA molecules Mn²⁺ is often used because of close properties to the Mg²⁺ in NMR or X-

ray crystallography studies (Gueron and Leroy, 1982; Ennifar *et al.*, 1999). It was observed that manganese could occupy in RNA structure position similar to magnesium and therefore used to locate probable magnesium ions in the structure. In the native tRNA^{Sec} structure, no Mg²⁺ ions could be identified. Therefore in order to locate putative magnesium binding sites in tRNA^{Sec} structure, we collected data set from crystals soaked in solution contained 100 mM manganese sulfate. Eight Mn²⁺ sites were found in the anomalous difference Fourier map. Three metal binding sites were located in chain A and 5 in chain B (Figure 29). Interestingly, all metal ions interact with N7 of guanine residues. Unfortunately, we cannot see the manganese hydration sphere and therefore it is not possible to establish contacts of the manganese ions via water molecules to other nucleotides in tRNA^{Sec}. Nonetheless at least one metal ion which is bound to G45 can reinforce interaction of U9 (AD-linker) with 45G:A48 (first base pair of the variable arm). Taken together, we suggest that divalent metal ions are probably not crucial for tRNA^{Sec} structure stabilization.

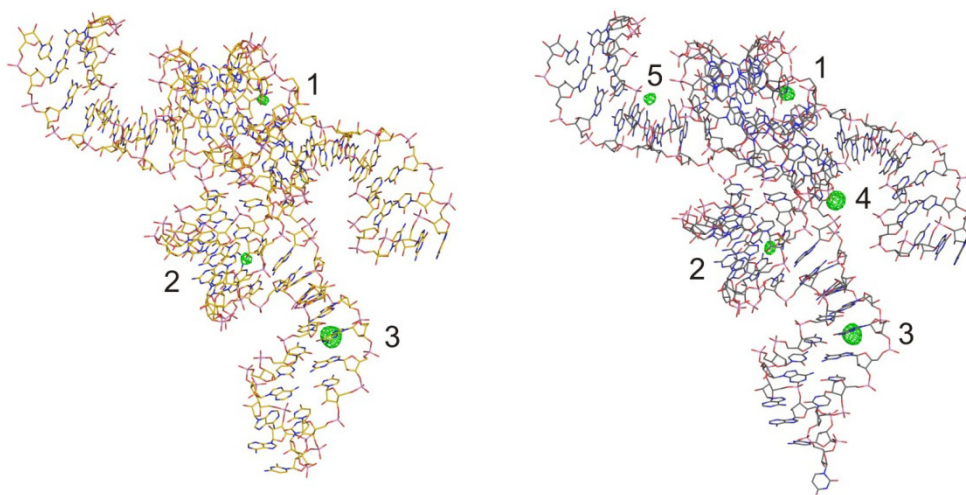


Figure 29. Location of manganese ions on tRNA^{Sec} structure. The Fourier difference map contoured at 5 σ is shown in green. tRNA^{Sec} derived from chain A and chain B colored gold and silver respectively. Three manganese sites are identical for both molecules. Sites 4 and 5 were found only in chain B.

4.4 Purification protocol for crystallization-grade RNA

4.4.1 Template design for *in vitro* transcription

Over the last two decades crystallization techniques together with implementation of the robotics systems significantly decrease amounts of material required for crystallization and allow to use nanoliter-range volume of the sample per crystallization condition. Apart of these, methods for production RNA sample suitable for crystallization are still actual.

For *in vitro* transcription with T7 RNA polymerase linearized plasmid DNA, double-stranded DNA generated by PCR or chemical synthesis are used as a templates. For preparation of long RNA molecules linearized plasmid DNAs are often a choice (Golden, 1997). PCR product as a template for synthesis of relatively short RNA fragments (~80 - 160 nt) has advantages over linearized plasmid DNA. Such as: possibility of insertion additional nucleotides via oligodeoxynucleotides used for PCR reaction or introduction two 2'-O-methyl modified nucleotides at 5'-end of the DNA template in order to suppress heterogeneity of RNA at 3'-end (Kao *et al.*, 1999, Sherlin *et al.*, 2001).

Desired DNA fragments for cloning and *in vitro* transcription included flanking sequences for restriction enzymes, T7 promoter and the gene of interest. Such fragment was assembled from oligodeoxynucleotides by PCR-based construction of long DNA molecules. The DNA oligos and annealing temperatures for PCR were calculated by “Assembly PCR Oligo Maker” program (Rydzanicz *et al.*, 2005).

4.4.2 Template production

In order to produce DNA template for *in vitro* transcription two step PCR reaction was implemented. The reverse primer used in the PCR contained two 2'-O-methyl modified nucleotides at 5'-end. For the first PCR step with 25 cycles a plasmid DNA with the desired cloned sequence was used as DNA template. Ten times diluted crude PCR mixture from the first reaction with amplified DNA was added in the reaction mixture for the second preparative PCR. Introduction of the two PCR steps allows to avoid plasmid linearization as total amount of plasmid is negligible compared to the amplified DNA fragment. Therefore involvement of plasmid DNA as template in transcription is neglected. Moreover use of amplified DNA fragment as template for preparative PCR with 35 cycles, significantly reduces unwanted non-specific PCR products. Before scaling up, several parameters of PCR (annealing temperature of forward and reverse primers, time for polymerization step and number of PCR cycles) have to be adjusted for obtaining maximal product yield. In order to decrease the number of point mutations which could accumulate in the amplified DNA fragment during PCR, it is important to use proof reading DNA polymerase. In our hands, best results were obtained with the *Pfu*-polymerase (Stratagene).

4.4.3 *In vitro* run-off transcription using T7 RNA polymerase

In this work, a transcription reaction elaborated by Pokrovskaya and co-authors for phage T7 RNA polymerase was used (Pokrovskaya *et al.*, 1994). We use PCR amplified

DNA template in the reaction mixture with a concentration of 10 ng/ μ l. To achieve maximum yield of transcription, the magnesium concentration in the reaction mixture should be optimized in analytical scale (50 μ l). Moreover, addition of the T7 polymerase after two hours of incubation increased the yield of the reaction. Based on this method, the general yield of purified RNA (70 – 90 nt) from 1 ml of transcription is about 1.2 to 1.5 mg.

4.4.4 RNA purification

The DNA template was removed from the transcription reaction by 30 min RQ1 RNase-free DNase treatment and then directly loaded on a strong anion-exchange MonoQ column connected to an FPLC system. Use of this column under moderate pressure allows to get very sharp peaks and outstanding separation within 30-50 min. Moreover as purification takes place under nondenaturing conditions, RNA species are separated on the basis of their secondary and tertiary structures (Guenther *et al.*, 1988). Thus different conformers can be separated within one run. RNA in the transcription mixture has the highest negative phosphate charge per molecule and therefore has the highest binding affinity to the column matrix compared to other components. All proteins which are present in the mixture do not bind to the column under our purification conditions. Therefore phenol/chloroform purification step is avoided.

Well defined buffer composition in which the sample is dissolved as well as homogenous state of the target molecule is important for successful crystallization and reproducibility. Therefore as a final purification step we employ size-exclusion chromatography. We used a 1 ml MonoQ 5/50 GL column connected to Äkta Purifier system (GE Healthcare). Sample was applied at a flow rate of 1 ml/min either via loop or superloop on the column equilibrated with low salt buffer (buffer A), which composition for MonoQ column adopted from (Kim R., *et al.*, 1995). During sample loading on column high absorbance peak always is monitored and corresponded to unincorporated ribonucleotides, T7 polymerase and other proteins from reaction mixture (Figure 30A; peak 1). Quality of RNA separation from proteins was analyzed on PAGE with SDS or 8M urea (Figure 30, B and C). It is useful to make a test purification in analytical scale for every new RNA molecule. For this pilot experiment, the gradient elution of RNA with high salt concentration (buffer B) from 0 to 100% have to be used in order to find a gradient point at which elution of target RNA is started. Once it is found for elution of preparative amounts of material can be applied salt concentration gradient around estimated value, which increases peak resolution. For our RNA molecules a salt concentration gradient from 30 to 100 % always works well.

When elution with salt concentration gradient is started small abortive synthesis products are eluted first (Figure 30A; peak 2). The next absorbance peak (Figure 30A; peak 3) usually corresponds to target RNA and after that aggregates or higher molecular weight nucleic acids are eluted (Figure 30A; peak 4). Usually the elution peak is narrow and the purified RNA after elution is obtained in a few milliliters which makes sample concentration by ultrafiltration before size-exclusion chromatography easier.

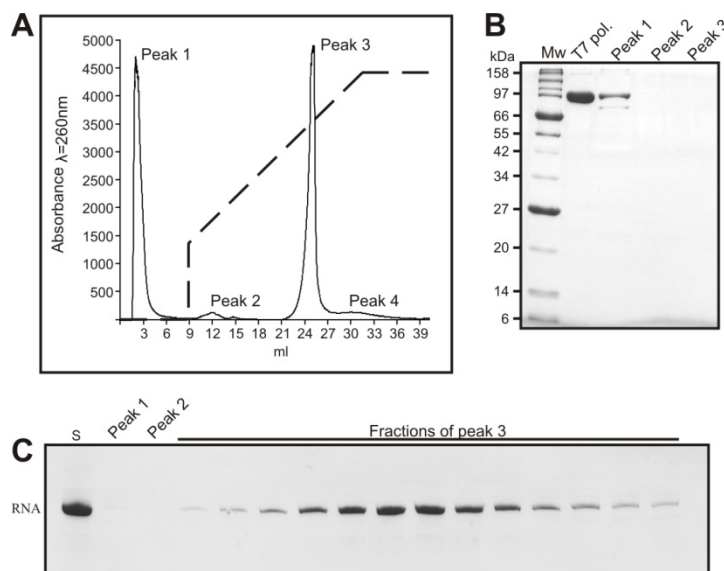


Figure 30. Evaluation of RNA purification from crude transcription mixture on anion-exchange chromatography. (A) Elution profile of tRNA^{Sec} from 1 ml transcription mixture obtained from monoQ column. Peak 1 corresponds to unincorporated rNTPs, T7 RNA polymerase and other proteins; Peak 2 corresponds to small abortive synthesis transcripts; Peak 3 corresponds to desired RNA sample. Concentration of solution B from 30 to 100 % was used for gradient elution and shown as a dashed line. (B) Denaturing SDS PAGE analysis of peak fractions TCA-precipitated pellets from peak 1, peak 2 and peak 3. T7 RNA polymerase and molecular weight markers (Mw) are loaded as references. Protein bands are stained by coomassie. (C) Denaturing PAGE analysis of peak fractions after monoQ column. S corresponds to crude transcription extract. RNA bands are stained by 0.01 % methylene blue.

A precipitation step by ethanol or isopropanol was avoided since it can cause aggregate formation. Gel filtration on column 10/300 GL is sufficient for amounts of RNA produced in 2 to 5 ml transcription reaction. For RNA obtained from larger scale of transcription (10 - 15 ml) reaction, HK16/60 gel filtration column was used. For mobile phase was used a buffer containing 10 mM HEPES-NaOH, pH 7.5, 50 mM NaCl. The quality of the purified RNA was monitored on native and denaturing PAGEs (Figure 31).

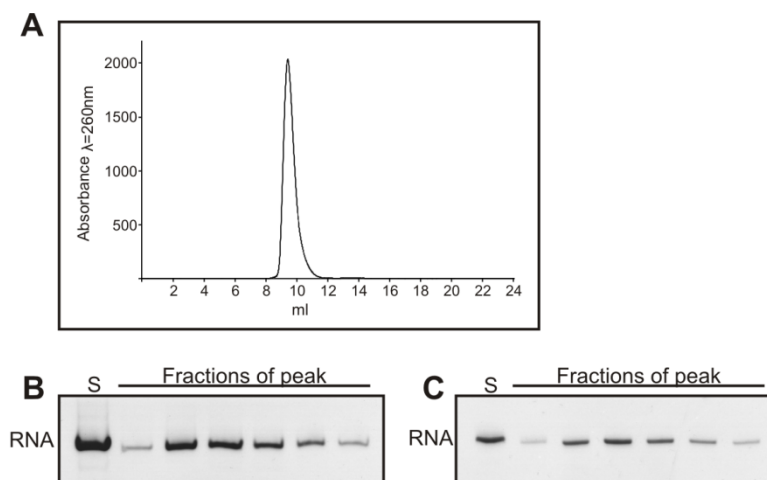


Figure 31. RNA purification by size-exclusion chromatography. (A) Elution profile of RNA sample obtained from Superdex 75 10/300 GL column. (B) Native and (C) denaturing PAGE analysis of individual 0.5 ml fractions collected after size-exclusion chromatography. S corresponds to concentrated RNA sample after anion-exchange chromatography.

In case of the Δ^{GCCA} tRNA^{Sec} construct, the sample also was analyzed by light scattering technique which is confirmed monomeric state of RNA (Figure 32).

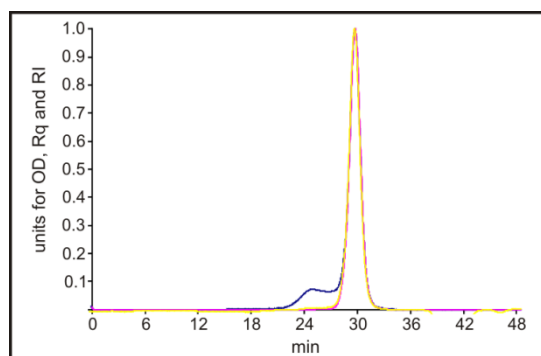


Figure 32. Analysis of the Δ^{GCCA} tRNA^{Sec} construct by light scattering. Optical density (OD) at 260 nm wavelength (pink curve), Rayleigh ratio (Rq) (blue curve) and Differential refractive index (RI) (yellow curve) of Δ^{GCCA} tRNA^{Sec} sample were measured during analytical gel filtration on Superdex 200 10/300 GL column. The measurement was done by R. Scherrers, Wyatt Technology Europe GmbH.

Fractions after gel filtration were concentrated by ultrafiltration (Millipore) to 8-9 mg/ml. A summary of our protocol is presented in Figure 33.

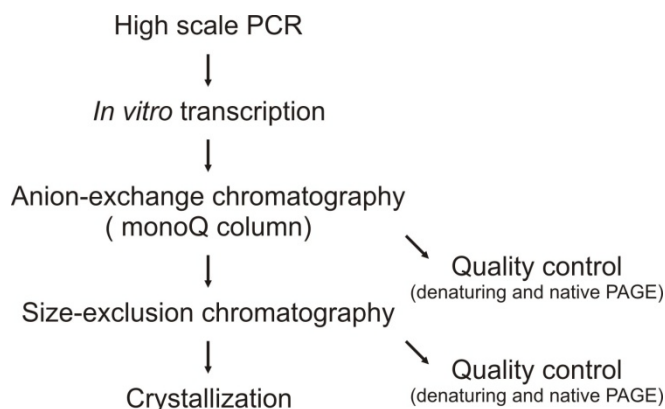


Figure 33. RNA production and purification scheme.

4.4.5 Biochemical activity of purified RNAs

In order to prove correct folding and biological activity of purified RNA samples, the RNAs were subjected to several biochemical experiments. Full length tRNA^{Sec} purified by our method showed binding activity to SecS protein based on a filter binding assay (Ganichkin *et al.*, 2008). It is known that tRNA^{Sec} and PSTK form a stable complex (Carlson *et al.*, 2004). Our results of analytical gel filtration experiments are consistent with published data (Sherrer *et al.*, 2008) (Figure 34). Purified SECIS RNA fragments derived from different selenoprotein mRNAs based on results of analytical gel filtration experiments showed complex formation with its binding partner the L30 protein (Figure 34) as reported earlier (Chavatte *et al.*, 2005).

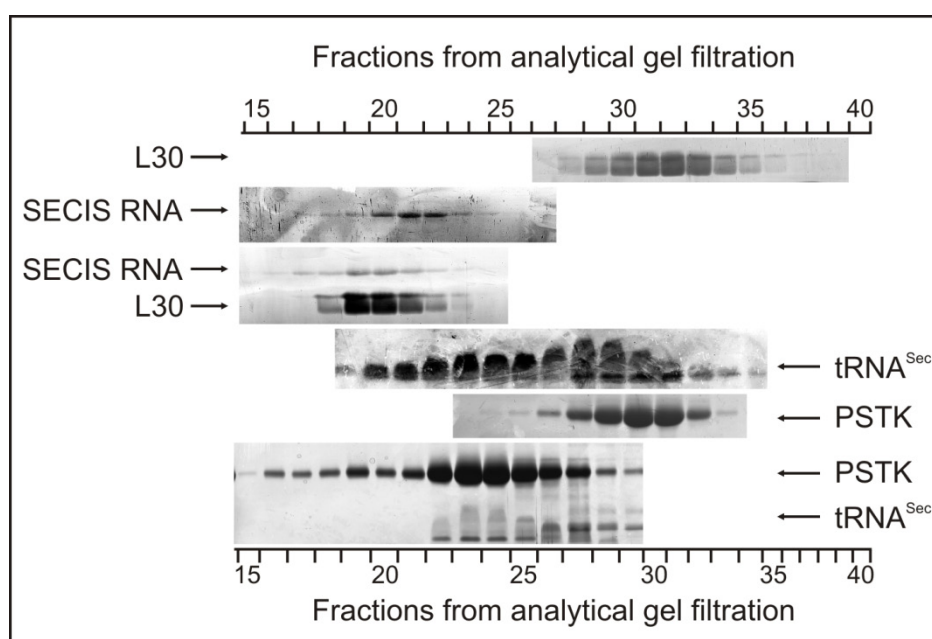


Figure 34. Functional analysis of purified RNAs. Denaturing SDS PAGE analysis of fractions derived from analytical size-exclusion chromatography of L30, SECIS RNA and its mixture. The same experiment was performed for tRNA^{Sec}, PSTK and its mixture. Shift of protein peak fractions in both experiments indicate complex formation between L30 and SECIS RNA or PSTK and tRNA^{Sec}. SDS PAGE stained with silver.

Finally, the crystal structure of ^{ΔGCCA}tRNA^{Sec} revealed that the purified construct is similar to other tRNA^{Sec} structures (Itoh *et al.*, 2009; Palioura *et al.*, 2009). Thus, results of complex formation between various RNAs and its binding partners clearly indicated that various different RNAs purified by our method are not only homogenous in length and conformational state but also biochemically active.

4.4.6 Engineering of RNA constructs for crystallization

Apart of screening of crystallization conditions it is important to find a crystallizable RNA construct. By number of successful examples the importance of rational designing of

RNA molecules for crystallization was demonstrated (Obayashi *et al.*, 2007; Ferre-D'Amare *et al.*, 1998; Golden *et al.*, 1997).

In this work, full length tRNA^{Sec} yielded crystals with low diffraction quality. Therefore I employed rational engineering for designing several alternative tRNA^{Sec} constructs (Figure 35), which were expected to form additional crystal contacts and thereby stabilize a crystal lattice. For modification of tRNA^{Sec} the variable loop sequence and the acceptor arm 3'GCCA-end were targeted. tRNA^{Sec} molecules are expected to form dimers in case of substitution of the variable loop with a "kissing loop" (exchange of UUAGC by CAGGCGCAG) or the 3'-end of acceptor arm GCCA with the GCGC sequence (Lodmell *et al.*, 2001). Substitution of the variable loop with the UUCG-tetraloop sequence (exchange of UUCGA by CAGGCGCAG) of acceptor arm will potentially increase stability of the molecule and deletion of 3'GCCA-end could facilitate stacking interactions in a crystal via acceptor stem (Ennifar *et al.*, 2000).

I generated the required templates using "QuikChangeTM" site-directed mutagenesis kit (Stratagene) and PCR. First of all we created two constructs (RNA 2 and RNA 3) containing sequence substitutions within the variable loop. Then using as a template for PCR plasmid DNA encoded tRNA^{Sec} with substituted variable loop or native tRNA^{Sec} sequence and different sets of primers, 9 different tRNA^{Sec} constructs were generated in total. As for transcription reaction we use the DNA template derived from PCR reaction and hence no additional cloning is required. Using our purification protocol from 2 ml of transcription reaction mixture is possible to produce at least 2.4 -3.0 mg of purified RNA. This amount of RNA concentrated up to 9 mg/ml is sufficient for high throughput crystallization screening and it is possible to set up as many as ~1000 crystallization conditions with use of robot dispensing 100 nl sample plus 100 nl precipitant solution.

Using PCR for generating DNA template for *in vitro* transcription and small scale transcription (2 ml) allows parallelizing production of DNA templates and then RNA constructs. Derivatives of native RNA molecule can be purified within the same conditions. Therefore no additional parameters for optimization for anion-exchange chromatography is required. The whole process for 9 constructs starting from template preparation to crystallization screening can be done within ten days. Once well diffracting crystals are found additional RNA for refinement of crystallization condition can be produced by scaling up the transcription reaction to 10 ml.

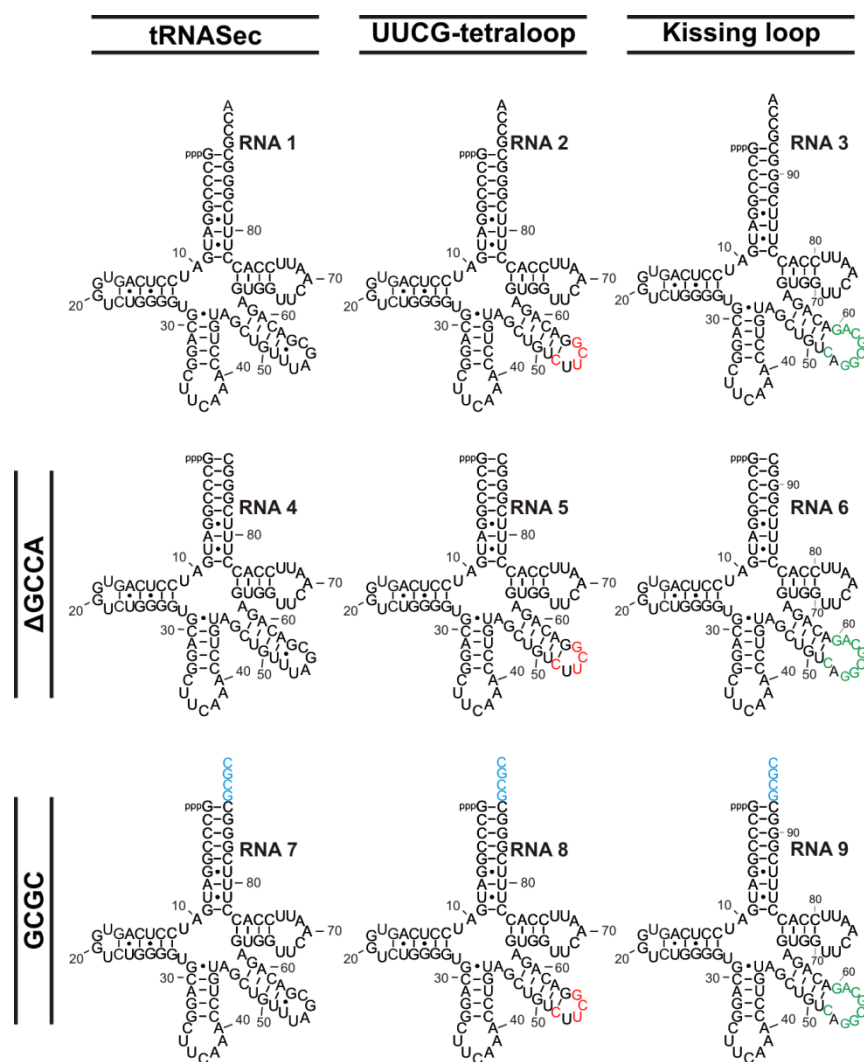


Figure 35. tRNA^{Sec} constructs designed for crystallization screening. RNA 1 represents full length tRNA^{Sec}, RNA 2 and 3 created by quick change mutagenesis technique and contain instead variable loop sequence UUCG (shown in red) and “kissing loop” (Lodmell *et al.*, 2001) (shown in green) accordingly. Using initial mutated constructs as templates for PCR and two different reverse primers were created DNA templates for *in vitro* transcription allowing to synthesize tRNA^{Sec} with deletion of 3'GCCA-end (RNA 4,5 and 6) or with substitution of GCCA to GCCG at 3'-end (RNA 7,8 and 9).

4.4.7 Crystallization of tRNA^{Sec} constructs

Six tRNA^{Sec} constructs (RNA 1, 2, 3, 4, 6, and 7) were purified by the method described above. Initial crystallization experiments of RNA^{Sec} constructs were set up by the Cartesian NanoDrop robot (Zinsser Analytik). Crystals appeared in many crystallization conditions for all constructs. They were harvested, shock frozen in liquid nitrogen and screened for diffraction using synchrotron radiation source [Swiss Light Source (Villigen, Switzerland)]. In case of RNA 4 construct (Δ GCCA-tRNA^{Sec}) diffracting crystals belonging to space groups P2₁, I4 and C2 were obtained. For a crystal of the monoclinic space group derived from initial crystallization screen data set with 2.7 Å resolution were collected. After several round of micro seeding size and diffraction quality of the initial crystals were

improved and they showed diffraction to 1.9 Å (Figure 36). Data collection statistics for tRNA^{Sec} are given in supplementary material (Suppl. Table 3).

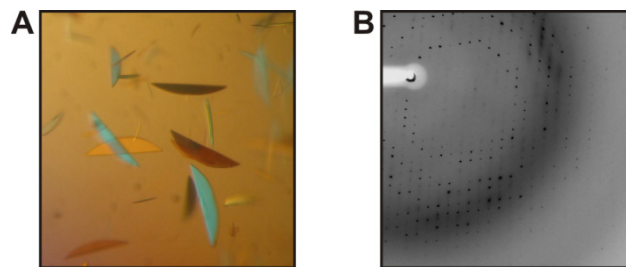


Figure 36. Monoclinic crystals resulting from purified $\Delta\text{GCCA}_t\text{tRNA}^{\text{Sec}}$ construct. (A) $\Delta\text{GCCA}_t\text{tRNA}^{\text{Sec}}$ crystals obtained using microseeding technique. (B) A 1° oscillation diffraction pattern from a crystal frozen to 100 K.

5 DISCUSSION

5.1 Conformational switches in winged-helix domains 1 and 2 of bacterial translation elongation factor SelB

The EF-Tu-like portion and the SECIS-binding WH3/4 part of bacterial SelB have to adopt defined relative orientations on the ribosome in order to allow concomitant interaction with Sec-tRNA^{Sec}, which approaches the ribosomal A-site where the UGA Sec-codon appears, and the SECIS element at the mRNA entrance channel. A defined but different orientation may also exist off the ribosome, explaining the cooperative binding of Sec-tRNA^{Sec} and the SECIS element to SelB. Most likely, the domain arrangement changes upon release of Sec-tRNA^{Sec} into the ribosomal A-site (Selmer and Su, 2002), which could lead to the somewhat reduced affinity of SelB for the SECIS element (Thanbichler *et al.*, 2000) and which could thereby facilitate the detachment of the factor from the mRNA. Otherwise, SelB would represent a roadblock for the progressing ribosome. No significant structural changes were seen in WH3/4 upon SECIS binding (Yoshizawa *et al.*, 2005). Instead, our results attest to the capacity of WH1/2 to undergo conformational changes. These conformational changes may be instrumental for functional communication between the N-terminal EF-Tu-like domains and the C-terminal SECIS binding domains of SelB. The different orientation of the C-terminal tail of WH2 (residues 504–511) in the present structure compared with the *mthSelB* WH1-4 structure directly supports the notion of a global repositioning of the two pairs of WH modules upon interaction of SelB with the ribosome. While different crystallization pH values may have contributed to the different orientations observed in the various crystal structures, it is conceivable that similar rearrangements are elicited by

changing molecular interactions in vivo. The relative movement of the domains would be enabled by conformational flexibility about the WH2-WH3 linker coupled to breakage of a conserved salt bridge connecting the two WH modules (Selmer and Su, 2002). The latter notion has been independently supported by recent crystal structures of *mt*hSelB WH1-4 in complex with SECIS elements (Soler *et al.*, 2007; Ose *et al.*, 2007). Interestingly, the WH1 domain was not visible in these latter crystal structures. The dynamic disorder of WH1 in these structures suggests that the link between the first two WH modules is also flexible. Therefore, rearrangement of the relative orientations of domains WH1 and WH2 may contribute to the conformational adjustments of SelB in different functional states. The conformational switch observed here within WH1 could also have long-range effects on the structure of SelB. When the previous WH1-4 structure is superimposed on WH1 helices α 1-3 or, alternatively, on WH2 of the present structure, the tip of the WH4 domain changes position by ~ 20 Å. It has been calculated that a conformational change in SelB WH1-4 that allows it to traverse an additional distance of 15-25 Å may be required on the ribosome in order to bridge between the ribosomal A-site (where the Sec-UGA codon appears) and the mRNA entrance channel (where the SECIS element is positioned; Selmer and Su, 2002). Possibly, part of this distance can be covered by a conformational response in WH1 as observed here. Domains I-III of SelB and portions of the 16S rRNA are candidate ligands that could interact at the putative ligand binding site disclosed by the conformational change in WH1 in the present structure. For example, it has been suggested that the WH1/2 region of SelB may interact with helix 16 or helix 33 of the 16S rRNA upon SelB binding to the ribosome (Ose *et al.*, 2007). In addition, a direct contact of domains I-III of SelB to WH1 has been discussed (Soler *et al.*, 2007). In summary, although the crystallization of complex *mt*hSelB377-511 and 16h was not successful the data presented here suggest that the WH1/2 element may serve as a malleable link rather than as a rigid spacer between the two business ends of bacterial SelB, thereby allowing them to adopt various functional orientations.

5.2 Structure and catalytic mechanism of eukaryotic selenocysteine synthase

5.2.1 SecS orthologs constitute a unique subclass of fold type I PLP-dependent enzymes

It was presented structural and functional analyses of mammalian SecS demonstrating how SecS orthologs are set aside from other PLP-dependent enzymes. Although *mmu*SecS^{elast} is composed of three domains whose structural scaffolds exhibit high similarity to the fold

type I family of PLP-dependent enzymes (Figure 16), distinguishing structural and functional characteristics are conferred by remodelled elements both within and outside of these conserved scaffolds (labelled I-IV in Figure 16). These novel elements clearly define SecS orthologs as a special subclass of the family.

Of paramount importance for the function of SecS are two motifs that are unique to and highly conserved in SecS orthologs. First, a special N terminus (element I) serves as a tetramerization device. It reinforces the interaction between two close dimers by cross-strutting (Figure 18, A and B). Neither *afu*PSerCysS nor NifS relatives nor *synC*-DES exhibit a comparable element, and all of these enzymes form dimers. As further detailed below, was suggested that tetramerization could be crucial for proper P-Ser-tRNA^{Sec} positioning. Second, a long insertion between helices $\alpha 2$ and $\alpha 4$ of domain 1 (element II) is involved in catalysis by SecS. Via a first loop, element II provides residues Glu74 and Arg75, which anchor the PLP cofactor in *trans*. It thereby positions the negatively charged Glu74 ideally to act as a substrate filter. Via a second loop it dispatches the P-loop to the *trans* active site, which mediates binding of SecS-specific substrates.

Evidently, the evolution of the unique N terminus and of the long domain 1 insertion went hand in hand. Placing the N terminus at the outside of the protein, where it can engage in tetramer formation, liberated a binding site between domains 1 and 3 as seen, *e.g.* in *afu*PSerCysS, NifS relatives, and *synC*-DES proteins (Figure 16). The first loop of the domain 1 insertion evolved to take advantage of the liberated binding site between domains 1 and 3 in *trans*. It thereby became ideally positioned to contribute residues to the *trans* active site (Figure 16A).

5.2.2 tRNA selection strategy of SecS

Results highlight a number of features that enable SecS to specifically recognize its substrates. The highly positively charged surface of *mmu*SecS^{elast} is apparently designed to interact at multiple sites with the phosphodiester backbone of tRNA^{Sec}. This expected mode of interaction is supported by observations of numerous anion (iodide) binding sites (Figure 18D), binding of a chloride ion in untreated crystals (not shown), and binding of the P2 phosphate ion distal to the active sites in phosphate-soaked crystals (Figure 19B). tRNA^{Sec} exhibits a number of unique structural characteristics compared with canonical tRNAs, such as an elongated (9 + 4 base pairs) helical stack between the acceptor stem and the T Ψ C stem and an unusually long variable arm (Commans and Bock, 1999; Hubert *et al.*, 1998). Multiple latching points on the surface of SecS would allow the enzyme to recognize these global

structural features and discriminate against other tRNAs, which are in vast excess in the cell. Thus, the tRNA selection strategy of SecS may resemble that of bacterial Sela, which very inefficiently converts Ser-tRNA^{Sec} mutants with a shortened, canonical acceptor stem (Baron and Bock, 1991).

The *mmu*SecS^{elast} active site environments are built up entirely by residues originating from the two protomers of a close dimer. Why then does *mmu*SecS form tetramers? One possibility is that the tetramer provides an effective binding platform for the large tRNA^{Sec} molecule. Although I have no direct evidence for the mode of P-Ser-tRNA^{Sec} binding to *mmu*SecS, portions of tRNA^{Sec} could extend beyond the borders of the molecule to whose active site its 3'-end is bound. To illustrate the relative sizes of the molecules and how the SecS tetramer could serve as a binding platform for tRNA^{Sec}, a hypothetical docking model (Figure 18E). A model of tRNA^{Sec} (as derived in Hubert *et al.*, 1998) was positioned on the *mmu*SecS^{elast} tetramer with the 3'-end approaching the active site of one monomer distal to the P-loop, leaving the P-loop available for accommodation of the P-Ser moiety. The body of tRNA^{Sec} was then adjusted by rotation about the 3'-terminal nucleotide to avoid clashes with the protein. Significantly, the unique mode of tetramerization provides the *mmu*SecS tetramer with a unique, elongated shape (shown by the characteristic distances between the active sites; Figure 18A). Other arrangements, such as the stubbier form of *E. Coli* cystathionine γ -synthase (Figure 18A), generate other relative dispositions of active sites (Clausen *et al.*, 1998). Thus, SecS may have evolved as a distinctly shaped tetramer to support efficient P-Ser-tRNA^{Sec} binding by one dimer to position its 3'-end appropriately in an active site of a neighbouring dimer. A similar principle may underlie the decameric organization of bacterial Sela. Further experiments are required to test these ideas.

5.2.3 Mechanisms for substrate binding and differentiation in the active site

Structural analyses have shown that the long, conserved domain 1 insertion (element II) of *mmu*SecS does not merely serve to support PLP anchoring in *trans*. Part of this insertion, which was referred to as the P-loop, can undergo disorder-order transitions coupled to binding of poly-dentate ions such as phosphate or sulphate. This observation was interpreted as direct evidence for the mode of substrate binding by *mmu*SecS. Molecular modelling suggests that the P1 phosphate could resemble the binding of the γ -phosphate of the P-Ser moiety of P-Ser-tRNA^{Sec} (Figure 19E). Evidently, SeP could also bind to the P-loop in a similar fashion as the P1 phosphate. In contrast, it was not possible to position the tRNA^{Sec} phosphodiester backbone in the same way at the P-loop. In any case, the P-loop in

combination with Arg313 (from element III) obviously binds and positions phosphate portions of substrates by an induced fit mechanism. This notion is perfectly borne out by performed mutational analyses, which showed that the phosphate-coordinating residues are crucial for *mmu*SecS activity (Figure 20A).

In addition, it was suggested that SecS has exploited the domain 1 insertion to install a filtering mechanism that allows it to exclude free P-Ser and other free amino acids from its active site. Glu74, positioned strategically next to the *trans* PLP (Figure 17, A and B), adopts a similar spatial position as the side chain of an Arg originating from the domain 3 β -sheet in related enzymes, which typically serves as an α -carboxylate recognition device (Figure 17C). In SecS, an equivalent Arg is present, but it is turned away via interactions with other conserved residues (Figure 17B), *e.g.* to engage in alternative interactions with the tRNA^{Sec} portion. It was suspected that Glu74 repels negatively charged carboxyl groups of free amino acids. Neutral ester moieties as in P-Ser-tRNA^{Sec} are presumably allowed to productively approach the PLP internal aldimine. Fact that SecS does not convert free P-Ser supports the role of Glu74 as a substrate discriminator. Similar to the P-loop, the elements constituting the putative substrate filter are highly conserved among SecS orthologs but not beyond. Thus, SecS has acquired specialized functional modules for substrate binding and differentiation.

γ -Phosphate binding at the P-loop provides a facile explanation for the observation that SecS binds P-Ser-tRNA^{Sec} preferentially over non-aminoacylated tRNA^{Sec} (Xu *et al.*, 2007a). However, presently it was possible only to speculate why Ser-tRNA^{Sec} is bound with least efficiency (Xu *et al.*, 2007a). One clue is provided by the observations that the P-loop tends to bind substrate mimics such as phosphate or sulphate and that P-Ser-tRNA^{Sec}, but not Ser-tRNA^{Sec}, can compete efficiently with these molecules. The unloaded 3'-end of tRNA^{Sec} may fit next to the contracted P-loop without having to compete for binding at that place. Another possibility is that γ -phosphate binding at the P-loop leads to proper accommodation of the P-Ser moiety and of the tRNA^{Sec} 3'-end in the active site, whereas Ser at the 3'-end may engage in alternative interactions, which could be mutually exclusive with proper fitting of tRNA^{Sec}. Clarification of these issues may require elucidation of a SecS-tRNA^{Sec} complex structure.

5.2.4 Structure-based reaction mechanism

NifS-like enzymes mobilize sulphur for iron-sulphur cluster biosynthesis via a protein-bound persulfide using a conserved cysteine that lies in a long loop of domain 3 (element IV in Figure 16, A and B) (Kaiser *et al.*, 2000). *synC*-DES employs a related strategy by

generating an external cysteine persulfide via cystine C-S cleavage, which remains non-covalently fixed at the active site (Clausen *et al.*, 2000). Furthermore, it has been discussed that PSerCysS could also employ a persulfide mechanism (Fukunaga *et al.*, 2007). Because NifS can support selenide delivery (Lacourciere and Stadtman, 1998), SecS could function in an analogous fashion by using a perselenide intermediate. However, although SecS exhibits a domain 3 loop analogous to the persulfide loop of NifS enzymes, no cysteine that could serve as attachment site for selenium is present in that loop. The only cysteine that is conserved in eukaryotes in the active site cavity is Cys226 (Figure 15). However, its sulphur atom is still more than 7 Å away from any atom of the PLP and remote from the modelled substrate. In addition, this residue is not conserved in archaeal SecS (Figure 15). Furthermore, in light of the observation that SeP delivered by selenophosphate synthetase 2 is the active selenium donor of SecS (Xu *et al.*, 2007a; Xu *et al.*, 2007b), internal or external perselenide production (via SeP) appears to be an off-pathway reaction. Under these circumstances it is unlikely that SecS functions via an intermediate perselenide moiety.

Instead, a mechanism that is consistent with all our findings evokes direct selenide delivery by SeP (Figure 37). As suggested by the inhibition results, the first part of the scheme is based on the *E. Coli* cystathionine β -lyase mechanism (Clausen *et al.*, 1996). In our working model the process is initiated by PSer-tRNA^{Sec} binding and positioning of the γ -phosphate at the P-loop (Figure 37, I). Lys284 is expected to be deprotonated after liberation upon external aldimine formation, constituting a strong base suitable for abstracting the α -hydrogen from the substrate (Figure 37, II). Phosphate release would generate aminoacrylyl-tRNA^{Sec} as an intermediate (Figure 37, III and IV). Next, the liberated phosphate is exchanged for SeP at the P-loop (Figure 37 IV), which would then be ideally situated to donate Se²⁻ to the β -carbon of the aminoacrylate moiety (Figure 37, IV and V). It was suggested that attack of the aminoacrylyl-tRNA^{Sec} intermediate by SeP involves concomitant attack by a water molecule on SeP, again giving rise to a phosphate leaving group (Figure 37, IV and V). A general base is proposed to activate this water molecule via a circular protonic shift involving Lys284 (Figure 37 IV). Subsequent reverse transaldimination (Figure 37, V and VI) liberates the product, Sec-tRNA^{Sec}, and regenerates the internal aldimine.

Details of the above model are still in the dark. For example, presently sequential binding and conversion of the substrates coupled to repeated contraction and relaxation of the P-loop was envisioned since phosphate-soaking experiments only revealed a single phosphate binding site. However, it cannot be rigorously excluded the possibility for a separate binding pocket for SeP at the present time. Furthermore, the identity of the water-activating general

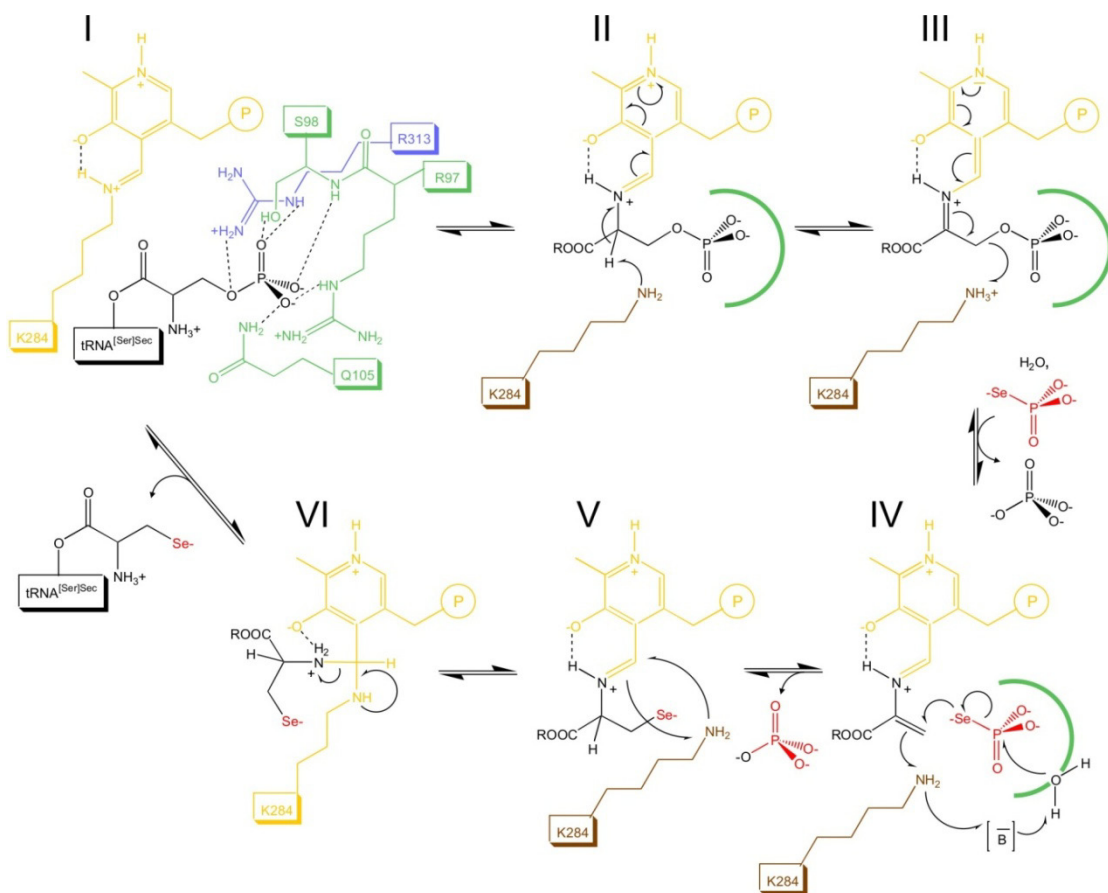


Figure 37. Proposed catalytic mechanism. Elements are colour-coded: gold, PLP and the PLP-bound Lys284; brown, liberated Lys284; green, residues of the P-loop; blue, Arg313; black, P-Ser-tRNA^{Sec}; red, SeP. R in II-VI represents tRNA^{Sec}. I, formation of the Michaelis complex, most likely involving binding of the γ -phosphate of P-Ser-tRNA^{Sec} to the P-loop/Arg313. In further steps, this binding site is symbolized by a green arch. II, abstraction of the α -proton from the internal aldimine by Lys284. III, phosphate elimination and generation of aminoacrylyl-tRNA^{Sec}; the evolving negative charge on the phosphate group is compensated by protonation via Lys284. IV, SeP replaces phosphate at the P-loop/Arg313 binding site. We suggest that attack of the aminoacrylyl tRNA^{Sec} intermediate by SeP involves concomitant attack by a water molecule on SeP. A general base (B) is proposed to activate this water molecule. Activation could be supported by a circular protonic shift involving Lys-284. V and VI, reverse transaldimination, regenerating the active site and release of the product, Sec-tRNA^{Sec}.

base is unknown at the time. This function could involve residues from the P-loop/Arg313 (for example, note the water bound to Arg97 and phosphate in (Figure 19)). In archaeal SecS, the proposed water activation could be provided by the histidine residue stacking on the PLP. It is also conceivable that Lys284 directly activates the water molecule without an intervening general base. Evoking a similar scenario, it has been shown that the PLP-binding lysine in *E. coli* cystathionine γ -synthase can reach to distal regions of the substrate (Clausen *et al.*, 1998).

5.2.5 New insights into SecS action

Recently a co-crystal structure of SecS and tRNA^{Sec} was reported (Palioura *et al.*, 2009). In that structure tRNA^{Sec} has a different binding mode compared to previously

suggested model (Figure 18E). One dimer of tetrameric SecS serves as a binding platform for two tRNA^{Sec}s (Figure 38), which point the 3'GCCA-ends into the active sites of the neighbouring dimer (Palioura *et al.*, 2009).

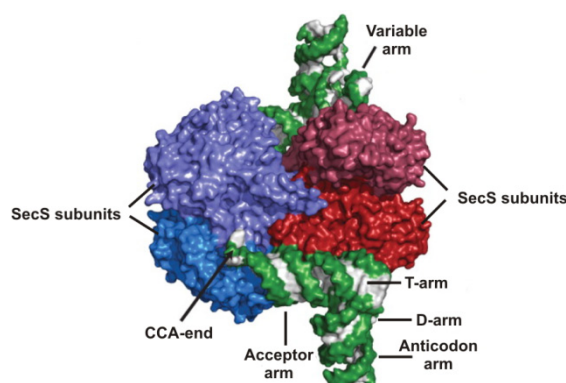


Figure 38. Co-crystal structure of human SecS with tRNA^{Sec}. Subunits of catalytic dimer are shown as dark and light blue surfaces and non-catalytic dimer as dark and light red respectively. tRNA^{Sec} coloured as grey and green surface (from Palioura *et al.*, 2009 with modifications).

As presented in Figure 38, human SecS binds only to the acceptor, TΨC, and variable stems of tRNA^{Sec}. Structure of binary complex of SecS and tRNA^{Sec} also confirmed our suggestions that enzyme has only one phosphate binding site in the reaction centre and therefore it can accommodate only one ligand at a time. Authors suggested another substrate-discriminating mechanism of P_{Ser}-tRNA^{Sec} against free P_{Ser} than was presented in our work. They show that tRNA^{Sec} binding to the SecS induced conformational change of the P-loop of the enzyme thereby allowing of the P_{Ser} binding. Moreover phosphate group of P_{Ser} important for its productive orientation within the active site of the enzyme. Therefore SecS can act only on P_{Ser} but not Ser. Notably, Palioura and co-authors suggested the same detailed reaction mechanism of P_{Ser} to Sec conversion as suggested in our earlier work.

5.3 Crystal structure analysis of tRNA^{Sec}

I determined the mouse tRNA^{Sec} structure at significantly higher resolution than reported previously (Itoh *et al.*, 2009; Palioura *et al.*, 2009). Therefore, in our structure of mouse tRNA^{Sec} was revealed already known and new structural features which have not been seen in previous tRNA^{Sec} structures. Structural analysis of tRNA^{Sec} allows to answer long-standing questions regarding the length of the stacked acceptor and the T-stem helices. Based on biochemical data two models were proposed: an unusual 9/4 nucleotide ratio between acceptor and T-stem and 7/5 ratio like in all elongator tRNAs (Sturchler *et al.*, 1993; Ioudovitch *et al.*, 1998). The crystal structure showed that tRNA^{Sec} has non-canonical 9/4 fold yielding a 13 bp composite acceptor-TΨC helix. The structure of complex tRNA^{Sec} with SecS

shows the importance of the 13 bp helix for the recognition mechanism by the SecS enzyme (Palioura *et al.*, 2009).

The D-arm of tRNA^{Sec} consist of a six-bp stem and is capped with four-nt loop (in contrast to a four-bp stem and a seven to eleven-nt loop in canonical tRNAs) serving as identity element for PSTK. The length of extra arm of tRNA^{Sec} resembles that of tRNA^{Leu}, tRNA^{Ser} and prokaryotic tRNA^{Tyr} (Steinberg *et al.*, 1993). Moreover the extra arm of tRNA^{Sec} is similar to that of tRNA^{Ser} allowing them to share SerRS for aminoacylation with serine. Compared to canonical tRNAs, tRNA^{Sec} has less tertiary interactions and forms an open cavity instead of an hydrophobic core (Figure 9) (Itoh *et al.*, 2009). The role of such cavity is not clear at this moment. Possibly, it lead to less rigidity of the entire molecule. Structural alignment of known tRNA^{Sec} molecules shows considerable flexibility of certain tRNA^{Sec} regions (Figure 29) and indirectly supports this idea. High flexibility of tRNA^{Sec} could be important for adopting an interaction surface of tRNA^{Sec} with potential binding partners. However to further support this hypothesis co-crystal structures of tRNA^{Sec} with different factors involved in tRNA^{Sec} modification, selenocysteine biosynthesis and its incorporation are required.

The variable arm consists of a six-bp stem capped with a four-nt loop. Uridine of the variable loop at position 47d makes a sharp turn which is a characteristic structural feature designated as U-turn motif (Cabello-Villegas *et al.*, 2002). The fold of the variable loop is reminiscent to those belonging to the s-Tltetraloop, classification established by Hsiao and colleagues and previously known as thermodynamically stable GNRA tetraloops (Hsiao *et al.*, 2006). However the distance between N3 of uridine 47d and 2'OH group of cytosine 47g is too long for H-bond formation therefore it is unlikely that the variable loop exhibit hyper stability as the GNRA tetraloops.

In the crystal structure obtained in this work, the anticodon loop adopted two different conformations in asymmetric unit. In chain A the anticodon stem consists of an eight-bp stem and a three-nt loop, whereas in chain B of a seven-bp stem and a five-nt loop. However it is most probable that only one conformation is physiologically relevant. Modeling of modified nucleotides within this loop shows that a 2'-O-hydroxymethylated ribose at position 34 (mcm⁵Um) favored a five-nt loop (Figure 27). Based on *in vivo* studies it was shown that modification from mcm⁵U to mcm⁵Um is a highly specialized step which is dependent on the correct primary, secondary, and tertiary structure of the tRNA^{Sec} and it is the last step in tRNA^{Sec} maturation (Kim *et al.*, 2000; Choi *et al.*, 1994). Moreover proportion of mcm⁵Um-containing Sec-tRNA^{Sec} compared to mcm⁵U-containing Sec-tRNA^{Sec} in the cell dependent

on Se diet suggesting that both isoforms could be responsible for synthesis of different fractions of selenoproteins (Hatfield *et al.*, 1991; Carlson *et al.*, 2009). However a link between 5 nucleotide loop conformation and such differentiation of selenoproteins biosynthesis is not clear and has to be elucidated.

5.4 Purification protocol for crystallization-grade RNA

A protocol for production and purification of RNA suitable for crystallization was elaborated. It was shown that with our approach it is possible to design new DNA templates for *in vitro* transcription without additional cloning.

One of the crucial steps in production of homogeneous RNA is the design of a DNA template for *in vitro* transcription. The use of a PCR product as a template for synthesis of RNA fragment (~80-160 nt) would be expected to have several advantages compared to the linearized plasmid DNA: 1) undesired flanking sequences that are often introduced as a part of the cloning procedure can be eliminated; 2) the possibility to incorporate a 2'-O-methyl modification at the 5'-end of the template can reduce 3'-end heterogeneity of the RNA (Kao *et al.*, 1999, Sherlin *et al.*, 2001). Manipulation of the sequence of the reverse primer for the PCR reaction provides flexibility to insert or delete certain nucleotides at 5'-end of the DNA template without cloning, thereby generating RNA constructs with different 3'-end.

It is known, that for effective transcription by T7 polymerase at least one G nucleotide at position +1 is required (Milligan *et al.*, 1987). Constructs that started with 4 or 5 consecutive guanosines could have one untemplated nucleotide at the 5'-end in fraction up to 30 % of their total transcriptional products (Pleiss *et al.*, 1998). The heterogeneity of the RNA product can be significantly reduced by decreasing the number of consecutive guanosines in the DNA template. A template starting with GCG showed no detectable 5'-end heterogeneity of the product (Pleiss *et al.*, 1998). To decrease heterogeneity on both ends of the RNA, I used one G or the GCG sequence starting at position +1 and incorporated two 2'-O-methyl modified nucleotides at 5'end of the DNA template.

It is well known that crystallization of RNA molecules is more complicated compared to proteins. One of the reasons is structural flexibility of RNA molecules which derived from a large number of rotatable bonds in the sugar-phosphate backbone. Such flexibility can lead to formation of alternative conformations which is unfavorable for crystallization. The Polyelectrolyte nature of RNA interferes with tight crystal packing of the molecules due to repulsive effects. Therefore apart from crystallization conditions screening it is important to find a crystallizable RNA construct. Rational engineering of RNA molecules for

crystallization proved to be a successful approach in a number of cases (Obayashi *et al.*, 2007; Ferre-D'Amare *et al.*, 1998; Golden *et al.*, 1997). In our case during crystallization of full length tRNA^{Sec} crystals with low diffraction quality were obtained. Employing rational engineering several tRNA^{Sec} constructs were designed (Figure 35), which were intended to form additional crystal contacts and thereby stabilize the crystal lattice. Indeed using this approach a construct which produced well diffracting crystals was found. This purification scheme is suitable for purification not only for tRNA^{Sec} constructs but also for other type of RNA molecules which indicates the wide applicability of this method.

6 REFERENCES

- Alexander R.W., Eargle J., Luthey-Schulten Z. 2010. Experimental and computational determination of tRNA dynamics. *FEBS Lett.* **584**:376-386.
- Allmang C., Krol A. 2006. Selenoprotein synthesis: UGA does not end the story. *Biochimie.* **88**:1561-71
- Allmang C, Wurth L, Krol A. 2009. The selenium to selenoprotein pathway in eukaryotes: more molecular partners than anticipated. *Biochim Biophys Acta.* **11**:1415-23.
- Antao V.P., Lai S.Y., Tinoco I. Jr. 1991. A thermodynamic study of unusually stable RNA and DNA hairpins. *Nucleic Acids Res.* **19**:5901-5.
- Araiso Y., Palioura S., Ishitani R., Sherrer R.L., O'Donoghue P., Yuan J., Oshikane H., Domae N., Defranco J., Söll D., Nureki O. 2008. Structural insights into RNA-dependent eukaryal and archaeal selenocysteine formation. *Nucleic Acids Res.* **36**:1187-99.
- Arnér E.S. 2010 Selenoproteins-What unique properties can arise with selenocysteine in place of cysteine? *Exp Cell Res.* **8**:1296-303.
- Baron C., Böck A. 1991. The length of the aminoacyl-acceptor stem of the selenocysteine-specific tRNA(Sec) of Escherichia coli is the determinant for binding to elongation factors SELB or Tu. *J Biol Chem.* **266**:20375-9.
- Baron C., Westhof E., Böck A., Giegé R. 1993. Solution structure of selenocysteine-inserting tRNA(Sec) from Escherichia coli. Comparison with canonical tRNA(Ser). *J Mol Biol.* **231**:274-92.
- Baron C., Heider J., Böck A. 1993. Interaction of translation factor SELB with the formate dehydrogenase H selenopolypeptide mRNA. *Proc Natl Acad Sci U S A.* **90**:4181-5.
- Baron C., Sturchler C., Wu X.Q., Gross H.J., Krol A., Böck A. 1994. Eukaryotic selenocysteine inserting tRNA species support selenoprotein synthesis in Escherichia coli. *Nucleic Acids Res.* **22**:2228-33.
- Berry M.J., Banu L., Harney J.W., Larsen P.R. 1993. Functional characterization of the eukaryotic SECIS elements which direct selenocysteine insertion at UGA codons. *EMBO J.* **12**:3315-22.
- Biou V., Yaremchuk A., Tukalo M., Cusack S. 1994. The 2.9 Å crystal structure of T. thermophilus seryl-tRNA synthetase complexed with tRNA(Ser). *Science.* **263**:1404-10.
- Biswas R., Wahl M.C., Ban C., Sundaralingam M. 1997. Crystal structure of an alternating octamer r(GUAUGUA)dC with adjacent G x U wobble pairs. *J Mol Biol.* **267**:1149-1156.
- Blum H., Beier H., Gross H.J. 1987. Improved silver staining of plant proteins, RNA and DNA in polyacrylamide gels. *Electrophoresis.* **8**: 93-99

- Böck A, Stadtman T.C. 1988. Selenocysteine, a highly specific component of certain enzymes, is incorporated by a UGA-directed co-translational mechanism. *Biofactors*. **1**:245-50.
- Böck A., Forchhammer K., Heider J., Baron C. 1991. Selenoprotein synthesis: an expansion of the genetic code. *Trends Biochem Sci*. **16**:463-7.
- Böck. A. Selenium metabolism in bacteria. in: Hatfield D.L. (Ed.). Selenium, Its Molecular Biology and Role in Human Health. 2001. *Kluwer Academic Publishers*, Norwood, MA. pp. 7–22
- Budiman M.E., Bubenik J.L., Miniard A.C., Middleton L.M., Gerber C.A., Cash A., Driscoll D.M. 2009. Eukaryotic initiation factor 4a3 is a selenium-regulated RNA-binding protein that selectively inhibits selenocysteine incorporation. *Mol Cell*. **35**:479-89.
- Caban K, Copeland PR. 2006. Size matters: a view of selenocysteine incorporation from the ribosome. *Cell Mol Life Sci*. **63**:73-81.
- Cabello-Villegas J., Winkler M.E., Nikonowicz E.P. 2002. Solution Conformations of Unmodified and A₃₇N⁶-dimethylallyl Modified Anticodon Stem-loops of *Escherichia coli* tRNA^{Phe}. *J. Mol. Biol*. **319**: 1015–1034
- Carlson B.A., Xu X.M., Kryukov G.V., Rao M., Berry M.J., Gladyshev V.N., Hatfield D.L. 2004. Identification and characterization of phosphoseryl-tRNA^{[Ser]^{Sec}} kinase. *Proc Natl Acad Sci U S A*. **101**:12848-53.
- Carlson B.A., Yoo M.H., Tsuji P.A., Gladyshev V.N., Hatfield D.L. 2009. Mouse models targeting selenocysteine tRNA expression for elucidating the role of selenoproteins in health and development. *Molecules*. **14**:3509-27.
- Cate J.H., Doudna J.A. 1996. Metal-binding sites in the major groove of a large ribozyme domain. *Structure*. **4**:1221-1229.
- Chavatte L., Brown B.A., Driscoll D.M. 2005. Ribosomal protein L30 is a component of the UGA-selenocysteine recoding machinery in eukaryotes. *Nat Struct Mol Biol*. **5**:408-16.
- Choi, I.S.; Diamond, A.M.; Crain, P.F.; Kolker, J.D.; McCloskey, J.A.; Hatfield, D.L. 1994. Reconstitution of the biosynthetic pathway of selenocysteine tRNAs in *Xenopus* oocytes. *Biochemistry*. **33**:601-605.
- Clausen T., Huber R., Laber B., Pohlenz H.D., Messerschmidt A. 1996. Crystal structure of the pyridoxal-5'-phosphate dependent cystathionine beta-lyase from *Escherichia coli* at 1.83 Å. *J Mol Biol*. **262**:202-24.
- Clausen T., Huber R., Prade L., Wahl M.C., Messerschmidt A. 1998. Crystal structure of *Escherichia coli* cystathionine gamma-synthase at 1.5 Å resolution. *EMBO J*. **17**:6827-38

- Clausen T., Kaiser J.T., Steegborn C., Huber R., Kessler D. 2000. Crystal structure of the cystine C-S lyase from *Synechocystis*: stabilization of cysteine persulfide for FeS cluster biosynthesis. *Proc Natl Acad Sci U S A*. **97**:3856-61.
- Cléry A., Bourguignon-Igel V., Allmang C., Krol A., Branlant C. 2007. An improved definition of the RNA-binding specificity of SECIS-binding protein 2, an essential component of the selenocysteine incorporation machinery. *Nucleic Acids Res*. **35**:1868-84.
- Commans S., Böck A. 1999. Selenocysteine inserting tRNAs: an overview. *FEMS Microbiol Rev*. **23**:335-51.
- Copeland P.R., Fletcher J.E., Carlson B.A., Hatfield D.L., Driscoll D.M. 2000. A novel RNA binding protein, SBP2, is required for the translation of mammalian selenoprotein mRNAs. *EMBO J*. **19**:306-14.
- Correll C.C., Wool I.G., Munishkin A. 1999. The two faces of the Escherichia coli 23 S rRNA sarcin/ricin domain: the structure at 1.11 Å resolution. *J Mol Biol*. **2**:275-87.
- Costa M., Rodríguez-Sánchez J.L., Czaja A.J., Gelpí C. 2000. Isolation and characterization of cDNA encoding the antigenic protein of the human tRNP(Ser)Sec complex recognized by autoantibodies from patients with type-1 autoimmune hepatitis. *Clin Exp Immunol*. **121**:364-74.
- Diamond A., Dudock B., Hatfield D. 1981. Structure and properties of a bovine liver UGA suppressor serine tRNA with a tryptophan anticodon. *Cell*. **25**:497-506.
- Diamond A.M., Choi I.S., Crain P.F., Hashizume T., Pomerantz S.C., Cruz R., Steer C.J., Hill K.E., Burk R.F., McCloskey J.A., Hatfield D.L. 1993. Dietary selenium affects methylation of the wobble nucleoside in the anticodon of selenocysteine tRNA([Ser]Sec). *J Biol Chem*. **268**:14215-14223.
- Ding F., Grabowski P.J. 1999. Identification of a protein component of a mammalian tRNA(Sec) complex implicated in the decoding of UGA as selenocysteine. *RNA*. **5**:1561-9.
- Dock A.C., Lorber B., Moras D., Pixa G., Thierry J.C., Giégé R. 1984. Crystallization of transfer ribonucleic acids. *Biochimie*. **66**:179-201
- Eargle J., Black A.A., Sethi A., Trabuco L.G., Luthey-Schulten Z. 2008. Dynamics of recognition between tRNA and elongation factor Tu. *J Mol Biol*. **377**:1382-1405
- Egli M., Portmann S., Usman N. 1996. RNA hydration: a detailed look. *Biochemistry*. **35**:8489-8494
- Ehrenreich A., Forchhammer K., Tormay P., Veprek B., Böck A. 1992. Selenoprotein synthesis in *E. coli*. Purification and characterisation of the enzyme catalysing selenium activation. *Eur J Biochem*. **206**:767-73.

- Emsley P., Lohkamp B., Scott W.G., Cowtan K 2010. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr. Pt 4*:486-501. Epub 2010 Mar 24.
- Engelhardt H., Forchhammer K., Müller S., Goldie K.N., Böck A. 1992. Structure of selenocysteine synthase from *Escherichia coli* and location of tRNA in the seryl-tRNA(sec)-enzyme complex. *Mol Microbiol.* **6**:3461-7.
- Ennifar E., Yusupov M., Walter P., Marquet R., Ehresmann B., Ehresmann C., Dumas P. 1999. The crystal structure of the dimerization initiation site of genomic HIV-1 RNA reveals an extended duplex with two adenine bulges. *Structure.* **11**:1439-49
- Fagegaltier D., Hubert N., Yamada K., Mizutani T., Carbon P., Krol A. 2000. Characterization of mSelB, a novel mammalian elongation factor for selenoprotein translation. *EMBO J.* **19**:4796-805.
- Ferre-D'Amare A.R., Zhou K., Doudna J.A. 1998. A general module for RNA crystallization. *J Mol Biol.* **3**:621-31
- Flohe L, Günzler W.A., Schock H.H. 1973. Glutathione peroxidase: a selenoenzyme. *FEBS Lett.* **1**:132-4.
- Forchhammer K., Böck A. 1991. Selenocysteine synthase from *Escherichia coli*. Analysis of the reaction sequence. *J Biol Chem.* **266**:6324-8.
- Forchhammer K., Leinfelder W., Boesmiller K., Veprek B., Böck A. 1991. Selenocysteine synthase from *Escherichia coli*. Nucleotide sequence of the gene (selA) and purification of the protein. *J Biol Chem.* **266**:6318-23.
- Forchhammer K., Leinfelder W., Böck A. 1989. Identification of a novel translation factor necessary for the incorporation of selenocysteine into protein. *Nature.* **342**:453-6.
- Forstrom J.W., Zakowski J.J., Tappel A.L. 1978. Identification of the catalytic site of rat liver glutathione peroxidase as selenocysteine. *Biochemistry.* **13**:2639-44.
- Fujii T., Maeda M., Mihara H., Kurihara T., Esaki N., Hata Y. 2000. Structure of a NifS homologue: X-ray structure analysis of CsdB, an *Escherichia coli* counterpart of mammalian selenocysteine lyase. *Biochemistry.* **39**:1263-73.
- Fukunaga R. and Yokoyama S. 2007. Structural insights into the second step of RNA-dependent cysteine biosynthesis in archaea: crystal structure of Sep-tRNA:Cys-tRNA synthase from *Archaeoglobus fulgidus*. *J Mol Biol.* **370**:128-41
- Gajiwala K.S., Burley S.K. 2000. Winged helix proteins. *Curr Opin Struct Biol.* **10**:110-6.
- Ganichkin O.M., Xu X.M., Carlson B.A., Mix H., Hatfield D.L., Gladyshev V.N., Wahl M.C. 2008. Structure and catalytic mechanism of eukaryotic selenocysteine synthase. *J Biol Chem.* **283**:5849-65.

- Gelpi C., Sontheimer E.J., Rodriguez-Sanchez J.L. 1992. Autoantibodies against a serine tRNA-protein complex implicated in cotranslational selenocysteine insertion. *Proc Natl Acad Sci U S A.* **89**:9739-43.
- Golden B.L., Podell E.R., Gooding A.R., Cech T.R. 1997. Crystals by design: a strategy for crystallization of a ribozyme derived from the Tetrahymena group I intron. *J Mol Biol.* **5**:711-23
- Guéron M, Leroy J.L. 1982. Significance and mechanism of divalent-ion binding to transfer RNA. *Biophys J.* **3**:231-6.
- Guimarães M.J., Peterson D., Vicari A., Cocks B.G., Copeland N.G., Gilbert D.J., Jenkins N.A., Ferrick D.A., Kastelein R.A., Bazan J.F., Zlotnik A. 1996. Identification of a novel seld homolog from eukaryotes, bacteria, and archaea: is there an autoregulatory mechanism in selenocysteine metabolism? *Proc Natl Acad Sci U S A.* **93**:15086-91.
- Guenther R.H., Gopal D.H., Agris P.F. 1988. Purification of transfer RNA species by single-step ion-exchange high-performance liquid chromatography. *J Chromatogr.* **444**:79-87.
- Gursinsky T., Jäger J., Andreesen J.R., Söhling B. 2000. A seldABC cluster for selenocysteine incorporation in Eubacterium acidaminophilum. *Arch Microbiol.* **174**:200-12.
- Hatfield D., Portugal F.H. 1970. Seryl-tRNA in mammalian tissues: chromatographic differences in brain and liver and a specific response to the codon, UGA. *Proc Natl Acad Sci U S A.* **67**:1200-6.
- Hatfield D.L., Carlson B.A., Xu X.M., Mix H., Gladyshev V.N. 2006. Selenocysteine incorporation machinery and the role of selenoproteins in development and health. *Prog Nucleic Acid Res Mol Biol.* **81**:97-142.
- Hatfield D., Lee B.J., Hampton L., Diamond A.M. 1991. Selenium induces changes in the selenocysteine tRNA^{[Ser]^{Sec}} population in mammalian cells. *Nucleic Acids Res.* **19**:939-43
- Hatfield D.L., Gladyshev V.N. 2002. How selenium has altered our understanding of the genetic code. *Mol Cell Biol.* **22**:3565-76.
- Heider J., Baron C., Böck A. 1992. Coding from a distance: dissection of the mRNA determinants required for the incorporation of selenocysteine into protein. *EMBO J.* **11**:3759-66.
- Hester G., Stark W., Moser M., Kallen J., Marković-Housley Z., Jansonius J.N. 1999. Crystal structure of phosphoserine aminotransferase from Escherichia coli at 2.3 Å resolution: comparison of the unligated enzyme and a complex with alpha-methyl-l-glutamate. *J Mol Biol.* **286**:829-50.
- Holm L., Sander C. 1995. Dali: a network tool for protein structure comparison. *Trends Biochem Sci.* **20**:478-80.

- Howard M.T., Moyle M.W., Aggarwal G., Carlson B.A., Anderson C.B. 2007. A recoding element that stimulates decoding of UGA codons by Sec tRNA[Ser]Sec. *RNA*. **13**:912-20.
- Howard M.T., Aggarwal G., Anderson C.B., Khatri S., Flanigan K.M., Atkins J.F. 2005 Recoding elements located adjacent to a subset of eukaryal selenocysteine-specifying UGA codons. *EMBO J*. **24**:1596-607.
- Huber R.E., Criddle R.S. 1967. Comparison of the chemical properties of selenocysteine and selenocystine with their sulfur analogs. *Arch Biochem Biophys*. **1**:164-73.
- Hubert N., Sturchler C., Westhof E., Carbon P., Krol A. 1998. The 9/4 secondary structure of eukaryotic selenocysteine tRNA: more pieces of evidence. *RNA*. **4**:1029-33.
- Hüttenhofer A., Böck A. 1998. Selenocysteine inserting RNA elements modulate GTP hydrolysis of elongation factor SelB. *Biochemistry*. **37**:885-90.
- Ioudovitch A., Steinberg S.V. 1998. Modeling the tertiary interactions in the eukaryotic selenocysteine tRNA. *RNA*. **4**:365-73.
- Itoh Y., Chiba S., Sekine S., Yokoyama S. 2009. Crystal structure of human selenocysteine tRNA. *Nucleic Acids Res*. **18**:6259-68.
- Joachimiak A., Haran T.E., Sigler P.B. 1994. Mutagenesis supports water mediated recognition in the trp repressor-operator system. *EMBO J*. **2**:367-72.
- Jühling F., Mörl M., Hartmann R.K., Sprinzl M., Stadler P.F., Pütz J. 2009. tRNADB 2009: compilation of tRNA sequences and tRNA genes. *Nucleic Acids Res*. (Database issue):D159-62.
- Kabsch W. 1988. Evaluation of single-crystal X-ray diffraction data from a position-sensitive detector. *J. Appl. Crystallogr.*, **21**, 916–924.
- Kaiser J.T., Gromadski K., Rother M., Engelhardt H., Rodnina M.V., Wahl M.C. 2005. Structural and functional investigation of a putative archaeal selenocysteine synthase. *Biochemistry*. **44**:13315-27.
- Kaiser J.T., Clausen T., Bourenkow G.P., Bartunik H.D., Steinbacher S., Huber R. 2000. Crystal structure of a NifS-like protein from *Thermotoga maritima*: implications for iron sulphur cluster assembly. *J Mol Biol*. **297**:451-64.
- Kennard O, Cruse W.B., Nachman J., Prange T., Shakked Z., Rabinovich D.J. 1986. Ordered water structure in an A-DNA octamer at 1.7 Å resolution. *Biomol Struct Dyn*. **3**:623-647.
- Kao C., Zheng M., Redisser S. 1999. A simple and efficient method to reduce nontemplated nucleotide addition at the 3 terminus of RNAs transcribed by T7 RNA polymerase. *RNA*. **9**:1268-72.

- Kelley L.A. and Sternberg M.J.E. 2009. Protein structure prediction on the web: a case study using the Phyre server. *Nature Protocols* **4**: 363-71
- Kernebeck T., Lohse A.W., Grötzinger J. 2001. A bioinformatical approach suggests the function of the autoimmune hepatitis target antigen soluble liver antigen/liver pancreas. *Hepatology*. **34**:230-3.
- Kim R, Holbrook E.L, Jancarik J., Kim S.H.1995. Synthesis and purification of milligram quantities of short RNA transcripts. *Biotechniques*. **6**:992-4.
- Kim L.K., Matsufuji T., Matsufuji S., Carlson B.A., Kim S.S., Hatfield D.L., Lee B.J. 2000. Methylation of the ribosyl moiety at position 34 of selenocysteine tRNA[Ser]^{Sec} is governed by both primary and tertiary structure. *RNA*. **6**:1306-15.
- Kinzy S.A., Caban K., Copeland P.R. 2005. Characterization of the SECIS binding protein 2 complex required for the co-translational insertion of selenocysteine in mammals. *Nucleic Acids Res.* **33**:5172-80.
- Klosterman P.S., Shah S.A., Steitz T.A. 1999. Crystal structures of two plasmid copy control related RNA duplexes: An 18 base pair duplex at 1.20 Å resolution and a 19 base pair duplex at 1.55 Å resolution. *Biochemistry*. **38**:14784-14792
- Klug A., Robertus J.D., Ladner J.E., Brown R.S., Finch J.T. 1974. Conservation of the molecular structure of yeast phenylalanine transfer RNA in two crystal forms. *Proc Natl Acad Sci U S A*. **71**:3711-5.
- Kromayer M., Wilting R., Tormay P., Böck A. 1996. Domain structure of the prokaryotic selenocysteine-specific elongation factor SelB. *J Mol Biol.* **262**:413-20.
- Kryukov G.V., Castellano S., Novoselov S.V., Lobanov A.V., Zehtab O., Guigó R., Gladyshev V.N. 2003. Characterization of mammalian selenoproteomes. *Science*. **300**:1439-43.
- Lacourciere G.M., Stadtman T.C. 1998. The NIFS protein can function as a selenide delivery protein in the biosynthesis of selenophosphate. *J Biol Chem.* **273**:30921-6
- Lee B.J., Worland P.J., Davis J.N., Stadtman T.C., Hatfield D.L. 1989. Identification of a selenocysteyl-tRNA(Ser) in mammalian cells that recognizes the nonsense codon, UGA. *J Biol Chem.* **264**:9724-7.
- Lee S.R., Bar-Noy S., Kwon J., Levine R.L., Stadtman T.C., Rhee S.G. 2000. Mammalian thioredoxin reductase: oxidation of the C-terminal cysteine/selenocysteine active site forms a thioselenide, and replacement of selenium with sulfur markedly reduces catalytic activity. *Proc Natl Acad Sci U S A*. **6**:2521-6.
- Lee B.J., de la Peña P., Tobian J.A., Zasloff M., Hatfield D. 1987. Unique pathway of expression of an opal suppressor phosphoserine tRNA. *Proc Natl Acad Sci U S A*. **84**:6384-8.

- Leibundgut M., Frick C., Thanbichler M., Böck A., Ban N. 2005. Selenocysteine tRNA-specific elongation factor SelB is a structural chimaera of elongation and initiation factors. *EMBO J.* **24**:11-22.
- Leinfelder W., Forchhammer K., Zinoni F., Sawers G., Mandrand-Berthelot M.A., Böck A. 1988. Escherichia coli genes whose products are involved in selenium metabolism. *J Bacteriol.* **170**:540-6.
- Leinfelder W., Stadtman T.C., Böck A. 1989. Occurrence in vivo of selenocysteyl-tRNA(SERUCA) in Escherichia coli. Effect of sel mutations. *J Biol Chem.* **264**:9720-3.
- Leinfelder W., Zehelein E., Mandrand-Berthelot M.A., Böck A. 1988. Gene for a novel tRNA species that accepts L-serine and cotranslationally inserts selenocysteine. *Nature.* **331**:723-5.
- Leinfelder W., Forchhammer K., Veprek B., Zehelein E., Böck A. 1990. In vitro synthesis of selenocysteinyl-tRNA(UCA) from seryl-tRNA(UCA): involvement and characterization of the selD gene product. *Proc Natl Acad Sci U S A.* **87**:543-7.
- Lima C.D. 2002. Analysis of the E. coli NifS CsdB protein at 2.0 Å reveals the structural basis for perselenide and persulfide intermediate formation. *J Mol Biol.* **315**:1199-208.
- Liu Z., Reches M., Groisman I., Engelberg-Kulka H. 1998. The nature of the minimal 'selenocysteine insertion sequence' (SECIS) in Escherichia coli. *Nucleic Acids Res.* **26**:896-902.
- Lobanov A.V., Hatfield D.L., Gladyshev V.N. 2008. Selenoproteinless animals: selenophosphate synthetase SPS1 functions in a pathway unrelated to selenocysteine biosynthesis. *Protein Sci.* **1**:176-82.
- Lodmell J.S., Ehresmann C., Ehresmann B., Marquet R. 2001. Structure and dimerization of HIV-1 kissing loop aptamers. *J Mol Biol.* **311**:475-90.
- Low S.C., Harney J.W., Berry M.J. 1995. Cloning and functional characterization of human selenophosphate synthetase, an essential component of selenoprotein synthesis. *J Biol Chem.* **270**:21659-64.
- Notredame C., Higgins D.G., Heringa J. 2000. T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol.* **302**:205-17.
- Mäenpää P.H., Bernfield M.R. 1970. A specific hepatic transfer RNA for phosphoserine. *Proc Natl Acad Sci U S A.* **67**:688-95.
- Manning G.S. 1978. The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotide. *Q. Rev. Biophys.* **11**: 179-246
- Metanis N, Keinan E, Dawson P.E. 2006. Synthetic seleno-glutaredoxin 3 analogues are highly reducing oxidoreductases with enhanced catalytic efficiency. *J Am Chem Soc.* **51**:16684-91.

- Mihara H., Fujii T., Kato S., Kurihara T., Hata Y., Esaki N. 2002. Structure of external aldimine of *Escherichia coli* CsdB, an IscS/NifS homolog: implications for its specificity toward selenocysteine. *J Biochem.* **131**:679-85.
- Milligan J.F., Groebe D.R., Witherell G.W., Uhlenbeck O.C. 1987. Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res.* **21**:8783-98.
- Mizutani T., Kurata H., Yamada K. 1991. Study of mammalian selenocysteyl-tRNA synthesis with [⁷⁵Se]HSe. *FEBS Lett.* **289**:59-63.
- Murshudov G.N., Vagin A.A., Dodson E.J. 1997. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr.* **53**:240-255.
- Nissen P., Kjeldgaard M., Thirup S, Polekhina G., Reshetnikova L., Clark B.F., Nyborg J. 1995. Crystal structure of the ternary complex of Phe-tRNAPhe, EF-Tu, and a GTP analog. *Science.* **270**:1464-72.
- Nissen P., Thirup S., Kjeldgaard M., Nyborg J. 1999. The crystal structure of Cys-tRNACys-EF-Tu-GDPNP reveals general and specific features in the ternary complex and in tRNA. *Structure.* **7**:143-56.
- Obayashi E., Oubridge C., Pomeranz Krummel D., Nagai K. 2007. Crystallization of RNA-protein complexes. *Methods Mol Biol.* **363**:259-76.
- Ohama T., Yang D.C., Hatfield D.L. 1994. Selenocysteine tRNA and serine tRNA are aminoacylated by the same synthetase, but may manifest different identities with respect to the long extra arm. *Arch Biochem Biophys.* **315**:293-301.
- Ose T., Soler N., Rasubala L., Kuroki K., Kohda D., Fourmy D., Yoshizawa S., Maenaka K. 2007. Structural basis for dynamic interdomain movement and RNA recognition of the selenocysteine-specific elongation factor SelB. *Structure.* **15**:577-86.
- Otwinowski, Z., and Minor W. 1997. Processing of X-ray diffraction data collected in oscillation mode. *Methods in Enzymology.* **276**: 307-326
- Palioura S., Sherrer R.L., Steitz T.A., Söll D., Simonovic M. 2009. The human SepSecS-tRNA^{Sec} complex reveals the mechanism of selenocysteine formation. *Science.* **325**:321-5.
- Paleskava A., Konevega A.L., Rodnina M.V. 2010. Thermodynamic and kinetic framework of selenocysteyl-tRNA^{Sec} recognition by elongation factor SelB. *J Biol Chem.* **285**:3014-20
- Pedersen J.S., Bejerano G., Siepel A., Rosenbloom K., Lindblad-Toh K., Lander E.S., Kent J., Miller W., Haussler D. 2006. Identification and classification of conserved RNA secondary structures in the human genome. *PLoS Comput Biol.* **4**:e33.
- Perrakis A., Morris R., Lamzin V.S. 1999. Automated protein model building combined with iterative structure refinement. *Nat Struct Biol.* **6**:458-463.

- Pleiss J.A., Derrick M.L., Uhlenbeck O.C. 1998. T7 RNA polymerase produces 5' end heterogeneity during *in vitro* transcription from certain templates. *RNA*. **10**:1313-7.
- Pokrovskaya I.D., Gurevich V.V. 1994. In vitro transcription: preparative RNA yields in analytical scale reactions. *Anal Biochem*. **220**:420-423.
- Record T.M., Anderson C.F., Lohman T.M. 1978. Thermodynamic analysis of ion effects on the binding and conformational equilibria of proteins and nucleic acids: the roles of ion association or release, screening, and ion effects on water activity. *Q. Rev. Biophys.* **11**: 103-178.
- Ringquist S., Schneider D., Gibson T., Baron C., Böck A., Gold L. 1994. Recognition of the mRNA selenocysteine insertion sequence by the specialized translational elongation factor SELB. *Genes Dev.* **8**:376-85
- Rother M, Resch A, Wilting R, Böck A. 2001. Selenoprotein synthesis in archaea. *Biofactors*. **14**:75-83.
- Rydzanicz R., Zhao X.S., Johnson P.E. 2005. Assembly PCR oligo maker: a tool for designing oligodeoxynucleotides for constructing long DNA molecules for RNA production. *Nucleic Acids Res.* 33(Web Server issue):W521-5.
- Sauerwald A., Zhu W., Major T.A., Roy H., Palioura S., Jahn D., Whitman W.B., Yates J.R. 3rd, Ibba M., Söll D. 2005. RNA-dependent cysteine biosynthesis in archaea. *Science*. **307**:1969-72.
- Sambrook, G., Fritsch, E.F., and Maniatis, T. 1989. Molecular cloning: A laboratory manual, second edition., New York.
- Sherlin L.D., Bullock T.L., Nissan T.A., Perona J.J., Lariviere F.J., Uhlenbeck O.C., Scaringe S.A. 2001. Chemical and enzymatic synthesis of tRNAs for high-throughput crystallization. *RNA*. **11**:1671-8.
- Schneider G., Köck H., Lindqvist Y. 2000. The manifold of vitamin B6 dependent enzymes. *Structure*. **8**:R1-6.
- Schneider T.R. and Sheldrick G.M. 2002. Substructure solution with SHELXD. *Acta Crystallogr D Biol Crystallogr.* Pt 10 Pt **2**:1772-9.
- Schwarz, K., and Foltz, C.M. 1957. Selenium as an integral part of factor 3 against dietary necrotic liver degeneration. *J. Am. Chem. Soc.* **79**: 3292 – 3293.
- Selmer M. and Su X.D. 2002. Crystal structure of an mRNA-binding fragment of Moorella thermoacetica elongation factor SelB. *EMBO J.* **21**:4145-53.
- Sherrer R.L., Ho J.M., Söll D. Divergence of selenocysteine tRNA recognition by archaeal and eukaryotic O-phosphoseryl-tRNA^{Sec} kinase. 2008. *Nucleic Acids Res.* **36**:1871-80.

- Sherrer R.L., O'Donoghue P., Söll D. 2008. Characterization and evolutionary history of an archaeal kinase involved in selenocysteinyl-tRNA formation. *Nucleic Acids Res.* **36**:1247-59.
- Shi H., Moore P.B. 2000. The crystal structure of yeast phenylalanine tRNA at 1.93 Å resolution: a classic structure revisited. *RNA.* **6**:1091-1105.
- Small-Howard A., Morozova N., Stoytcheva Z., Forry E.P., Mansell J.B., Harney J.W., Carlson B.A., Xu X.M., Hatfield D.L., Berry M.J. 2006. Supramolecular complexes mediate selenocysteine incorporation in vivo. *Mol Cell Biol.* **26**:2337-46.
- Studier F.W. 2005. Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif.* **41**:207-34.
- Sturchler C., Westhof E., Carbon P., Krol A. 1993. Unique secondary and tertiary structural features of the eucaryotic selenocysteine tRNA(Sec). *Nucleic Acids Res.* **21**:1073-9.
- Sturchler C., Lescure A., Keith G., Carbon P., Krol A. 1994. Base modification pattern at the wobble position of *Xenopus* selenocysteine tRNA(Sec). *Nucleic Acids Res.* **22**:1354-8.
- Sturchler-Pierrat C., Hubert N., Totsuka T., Mizutani T., Carbon P., Krol A. 1995. Selenocysteylation in eukaryotes necessitates the uniquely long aminoacyl acceptor stem of selenocysteine tRNA(Sec). *J Biol Chem.* **270**:18570-4.
- Steitz T.A. 2008. A structural understanding of the dynamic ribosome machine. *Nat Rev Mol Cell Biol.* **9**:242-53
- Tamura T., Yamamoto S., Takahata M., Sakaguchi H., Tanaka H., Stadtman T.C., Inagaki K. 2004. Selenophosphate synthetase genes from lung adenocarcinoma cells: Sps1 for recycling L-selenocysteine and Sps2 for selenite assimilation. *Proc Natl Acad Sci U S A.* **101**:16162-7.
- Thanbichler M., Bock A., Goody R.S. 2000. Kinetics of the interaction of translation factor SelB from *Escherichia coli* with guanosine nucleotides and selenocysteine insertion sequence RNA. *J Biol Chem.* **275**:20458-66.
- Tujebajeva R.M., Copeland P.R., Xu X.M., Carlson B.A., Harney J.W., Driscoll D.M., Hatfield D.L., Berry M.J. 2000. Decoding apparatus for eukaryotic selenocysteine insertion. *EMBO Rep.* **1**:158-63.
- Tormay P., Wilting R., Lottspeich F., Mehta P.K., Christen P., Böck A. 1998. Bacterial selenocysteine synthase--structural and functional properties. *Eur J Biochem.* **254**:655-61.
- Vagin A., Teplyakov A. 2010. Molecular replacement with MOLREP. *Acta Crystallogr D Biol Crystallogr.* **66**:22-25.
- Vendeix F.A., Dziargowska A., Gustilo E.M., Graham W.D., Sproat B., Malkiewicz A., Agris P.F. 2008. Anticodon domain modifications contribute order to tRNA for ribosome-mediated codon binding. *Biochemistry.* **47**:6117-29.

- Walter N.G. 2007. Ribozyme catalysis revisited: is water involved? *Mol Cell*. **6**:923-9
- Warner G.J., Berry M.J., Moustafa M.E., Carlson B.A., Hatfield D.L., Faust J.R. 2000. Inhibition of selenoprotein synthesis by selenocysteine tRNA^{[Ser]^{Sec}} lacking isopentenyladenosine. *J Biol Chem*. **275**:28110-28119.
- Weeks M. E. 1932. The discovery of the elements. VI. Tellurium and selenium. *J. Chem. Educ.* **3**: 474
- Westhof E. 1988. Water: an integral part of nucleic acid structure. *Annu. Rev. Biophys. Biophys. Chem.* **17**: 125–144.
- Westhof E. and Beveridge D.L. 1990. Hydration of nucleic acids. *Water Science Reviews* **5**, Cambridge University Press, Cambridge. 24–123.
- Wies I., Brunner S., Henninger J., Herkel J., Kanzler S., Meyer zum B?schenfelde K.H., Lohse A.W. 2000. Identification of target antigen for SLA/LP autoantibodies in autoimmune hepatitis. *Lancet*. **355**:1510-5.
- Winn M.D., Isupov M.N., Murshudov G.N. 2001. Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr D Biol Crystallogr*. Pt **1**:122-33
- Wilber C.G. 1980. Toxicology of selenium: a review. *Clin Toxicol*. **2**:171-230.
- Wu X.Q., Gross H.J. 1994. The length and the secondary structure of the D-stem of human selenocysteine tRNA are the major identity determinants for serine phosphorylation. *EMBO J*. **13**:241-8.
- Xu X.M., Mix H., Carlson B.A., Grabowski P.J., Gladyshev V.N., Berry M.J., Hatfield D.L. 2005. Evidence for direct roles of two additional factors, SECp43 and soluble liver antigen, in the selenoprotein synthesis machinery. *J Biol Chem*. **280**:41568-75.
- Xu X.M., Carlson B.A., Mix H., Zhang Y., Saira K., Glass R.S., Berry M.J., Gladyshev V.N., Hatfield D.L. 2007a. Biosynthesis of selenocysteine on its tRNA in eukaryotes. *PLoS Biol*. **1**:e4.
- Xu X.M., Carlson B.A., Irons R., Mix H., Zhong N., Gladyshev V.N., Hatfield D.L. 2007b. Selenophosphate synthetase 2 is essential for selenoprotein biosynthesis. *Biochem J*. **404**:115-20.
- Yoshizawa S., Rasubala L., Ose T., Kohda D., Fourmy D., Maenaka K. 2005. Structural basis for mRNA recognition by elongation factor SelB. *Nat Struct Mol Biol*. **12**:198-203.
- Yuan J., Palioura S., Salazar J.C., Su D., O'Donoghue P., Hohn M.J., Cardoso A.M., Whitman W.B., Söll D. 2006. RNA-dependent conversion of phosphoserine forms selenocysteine in eukaryotes and archaea. *Proc Natl Acad Sci U S A*. **103**:18923-7.

- Zavacki A.M., Mansell J.B., Chung M., Klimovitsky B., Harney J.W., Berry M.J. 2003. Coupled tRNA(Sec)-dependent assembly of the selenocysteine decoding apparatus. *Mol Cell*. **11**:773-81.
- Zhang Y., Gladyshev V.N. 2009. Comparative genomics of trace elements: emerging dynamic view of trace element utilization and function. *Chem Rev*. **109**:4828-61.
- Zhong L, Holmgren A. 2000. Essential role of selenium in the catalytic activities of mammalian thioredoxin reductase revealed by characterization of recombinant enzymes with selenocysteine mutations. *J Biol Chem*. **24**:18121-8.

7 APPENDIX

7.1 X-ray crystallography method

The ultimate goal of crystallography is to investigate functional properties of the molecule based on the knowledge of its three dimensional architecture. In order to obtain structural information, a long route has to be taken which can be divided into several parts: 1) crystallization, 2) data collection from a crystal, 3) solving phase problem 4) modeling of the structure and refinement (Rodes, 2006).

7.1.1 Crystallization and crystal growth

The aim of the first step is to obtain single, well-ordered crystal of the biological macromolecule (protein, RNA, DNA or their complexes) of interest. Crystallization is characterized as controlled precipitation of the target molecule from aqueous solution into a solid crystalline form. In practice, crystallization is done by mixing of concentrated protein solution (10-20 mg/ml) with an ca. equal volume of a certain precipitant solution (salt, organic solvent or organic polymer) and equilibration in a sealed container against a larger volume of precipitant solution (reservoir). The vapor pressures of both solutions slowly come into equilibrium due to redistribution of water from the protein mixture to the reservoir solution thereby increasing the protein concentration over time or by increasing volatile precipitant concentration in protein drop. During this equilibration process, the protein concentration reaches its solubility limit (supersaturation phase) and the protein molecules start to form ordered aggregates (nucleation phase) which turn into ordered crystals (crystal growth phase).

Unfortunately quite often biological macromolecules form disordered aggregates in form of an amorphous precipitate. The surface of every biological macromolecule has unique physico-chemical properties that influence crystallization. Therefore crystallization conditions cannot be predicted a priori and require tedious screening. By now many crystallization techniques have been developed and the most popular ones are hanging drop and sitting drop.

7.1.2 Data collection

Data collection describes the process of collecting all unique reflections on a detector produced by deviating X-rays emerging from the crystal positioned in an X-ray beam. There are three commonly used X-ray sources: sealed tubes, rotating anode tubes and synchrotrons.

The first two sources are routinely used for in-house data collection. The principles of X-ray generation in both variants are similar: electrons emitting from a hot filament (cathode) are accelerated by an electric field and encounter a metal anode. Upon collision, high-energy electrons from filament are displaced electrons from inner shell of atoms of metal anode. In result electrons from higher-energy orbital drops to lower-energy inner shell and the energy-differences emitted as X-rays. X-rays produced from the anode travel through a fine metal tube (collimator) forming a narrow X-ray beam. The rotating anode tube is ~ 10 times more powerful than sealed tubes due to better heat dissipation over a larger surface of the rotating metal disk (anode). Synchrotrons produce the brightest X-rays to date with the possibility to control the bandwidth, photon flux, beam dimensions, focus and collimation of the ray. Such control possible with help of optical devices including slits, attenuators, crystal monochromators and mirrors.

In order to collect a diffraction data set, the crystal has to be positioned in the X-ray beam between the X-ray source and the detector. The wavelength of the X-rays impinging on crystal is comparable to the distances between atoms in the molecules leading to X-ray diffraction on the crystal lattice. Diffraction be described as the reflection of X-rays from virtual parallel planes through the crystal lattice. If the path differences between two beams reflected from neighboring planes equals an integral multiple of the X-ray wavelength. Constructive interference and thus a reflection spot is observed. This process described by Bragg's law (Figure 39) with:

$$2d_{hkl} \sin\theta = n\lambda \quad (1)$$

where d_{hkl} is interplanar spacing, n is an integer, λ is X-ray wavelength and θ is an angle between parallel planes of atoms in a crystal and the impinges and emerges from them X-ray. In all other cases diffraction from parallel planes is out of phase and interference is destructive, thus no diffraction emerges.

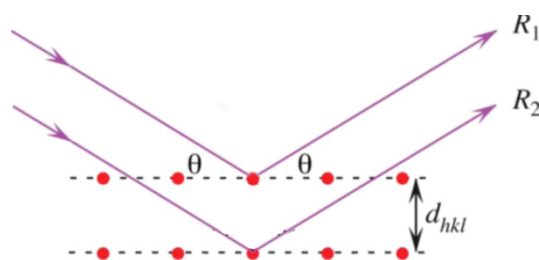


Figure 39. Conditions that produce strong diffracted rays. If the additional distance traveled by the more deeply penetrating ray R_2 is an integral multiple of λ , then rays R_1 and R_2 interfere constructively

Finally on a detector the diffraction pattern at given crystal orientation is recorded. In order to collect all possible reflections, the crystal is rotated within the X-ray beam and the

diffraction pattern is recorded in small angular slices (usually between 0.3° - 1°). The number of recorded images depends on the crystal symmetry and the unit cell size. A full series of collected diffraction images is called a data set. In the next step, known as scaling, a list of coordinates of all reflections and their intensities is generated. The position of reflections in the reciprocal space is defined by h , k , l coordinates, which describe a number of equivalent parallel planes in the set per unit cell in the x , y , z direction. The reflection intensity is proportional to the amplitude of the diffracted wave.

The relationship between atoms in the crystal and the diffraction pattern is described by the structure factor (F_{hkl}).

$$F_{hkl} = \sum_{j=1}^n f_j e^{2\pi i(hx_j + ky_j + lz_j)} \quad (2)$$

The structure factor is a sum of all individual atomic structure factors ($f_j e^{2\pi i(hx_j + ky_j + lz_j)}$), where f_j is the scattering factor of atom j , which determines the contribution of the atom; x_j , y_j and z_j are the coordinates of atom j in the unit cell, expressed as fractions of unit cell axis lengths, and establishes the phase of its contribution; h , k and l are the “frequencies” of the wave and also the indices of a specific reflection in the reciprocal lattice and i is $\sqrt{-1}$. The conversion of the data from the “reciprocal space” (h , k , l) to the value of the electron density at a “real space” lattice point (x , y , z) (designated as $\rho(x, y, z)$) is achieved by a Fourier transform operation:

$$\rho(x, y, z) = 1/V \sum_h \sum_k \sum_l |F_{hkl}| e^{[-2\pi i(hx + ky + lz) - \alpha(hkl)]} \quad (3)$$

where V is unit cell volume; each term in the series is a three-dimensional wave of amplitude $|F_{hkl}|$, phase α_{hkl} and frequencies h , k , l respectively.

Thus, for determination of an electron density map at any spatial point of the crystal information about the intensity, the phase and the frequency of scattered ray is required. Intensity and frequency can be measured directly but the phase information can not be obtained directly from the diffraction experiment and represents a problem on the way of structure determination.

7.1.3 Methods for resolving the phase problem

Several approaches were developed to obtain the missing phase information. The most important procedures for biological macromolecules are: MR (molecular replacement), SIR (single isomorphous replacement), MIR (multiple isomorphous replacement), SAD (single-

wavelength anomalous dispersion), MAD (multi-wavelength anomalous dispersion) and their combinations: SIRAS (single isomorphous replacement with anomalous scattering) and MIRAS (multiple isomorphous replacement with anomalous scattering). Apart of MR, all these methods are based on slight differences in diffraction patterns between collected datasets: one set collected from native crystal and another from derivatized crystal.

In order to see such perturbations in diffraction pattern, a small number of heavy atoms has to bind to identical sites of the protein in all units cells of a crystal. As another important condition for these methods, the unit cell parameters have to be nearly identical (isomorphous) between datasets. Depends on a method the datasets from derivatized crystal can be collected either at different wavelength or at the same, if it in the absorbance edge of the heavy metal in the crystal.

The differences between amplitudes of reflections from derivatized crystal (F_{PH}) and native crystal (F_P) yield approximately a contribution of heavy atoms (F_H). A few heavy atoms in the unit cell of the protein can be considered as a simpler structure. The coordinates of the heavy atoms are determined by a Fourier sum called the Patterson function. The Patterson vector map displays peaks at locations corresponding to vectors between atoms.

Once the positions of heavy atoms are found, using equation (2) both the amplitudes and the phases of structure factor F_H can be determined by the SIR method. Then by use of a Harker diagram, which represents the equation ($F_P = F_{PH} - F_H$) in the complex plane, possible to find the phase for F_P (Figure 40).

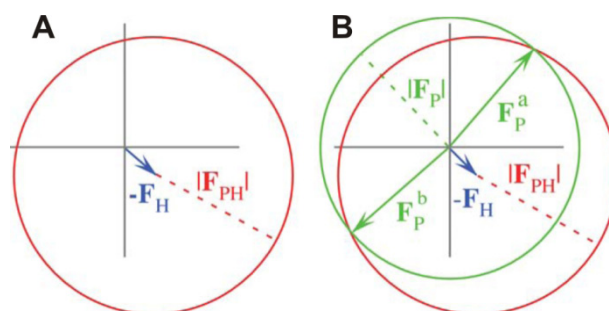


Figure 40. Vector solution of equation $F_P = F_{PH} - F_H$. (A) All points on the red circle equal the vector sum $|F_{PH}| - F_H$. (B) Vectors from the origin to intersections of the two circles are solutions to F_P . The structure factor of the heavy atom is blue, the structure factor of the heavy atom derivative of the protein is red and the structure factor of the protein is green. Circles depicting possible orientations of structure factors carry the same colors. (Adopted from Rhodes, 2006)

In order to do that the vector $-F_H$ has to be placed at the origin and circle with radius $|F_{PH}|$ drawn centered on the tip of the vector $-F_H$. Then the second circle with radius $|F_P|$ is drawn centered on the origin. The phase angles of the two vectors F_P^a and F_P^b that terminate at the

intersection points of both circles are the only possible phases for a given reflection. However, only one of these two phases is correct. Diagram construction by the same principle and their superimposition for all heavy atoms allow resolving the phase ambiguity.

Certain heavy atoms are able to absorb X-rays in a wavelength range available at synchrotron sources. As a result the intensity of the reflections h, k, l and symmetry-related $-h, -k, -l$, known as Friedel pairs, are not equal any more. The difference between Friedel pairs is called anomalous scattering or anomalous dispersion and is used by the SAD and MAD methods to obtain phase information with help of the Harker diagram. For the MAD method several data sets collected at different wavelength (heavy atom absorption-edge peak, inflection point of the absorption edge, and at a wavelengths distant from the absorption-edge peak). All of them have different values of the anomalous dispersion and each measurement of Friedel pairs at different wavelength provides the components for distinct sets of phasing equations. Superimposition of Harker diagram for data from all measured wavelength facilitate to solve the phase ambiguity. For the SIRAS method, the phase problem is solved by principles introduced for the SIR method in combination with the anomalous scattering strategy.

In molecular replacement phases from structure factors of a known protein (phasing model) are used as initial estimates of phases for a new structurally similar protein. In this method initial phases are obtained by positioning of the phasing model in the unit cell of the new protein. When the phasing model and target are superimposed then the phases from model can be directly used for calculation of the electron density from the intensities of the new protein. For superimposition, phasing model has to be rotated and translated in the unit cell of the new protein. The best orientation of the model can be found by superimposition of three dimensional Patterson maps of both molecules. Comparison of orientations is done by calculating a rotation function, which evaluates the correlation between the Patterson maps for both molecules in various orientations. Best orientation and position in the unit cell are found by comparison of overall agreement between amplitudes of the model $|F_{calc}|$ with the measured amplitudes $|F_{obs}|$ obtained from diffraction intensities of the new protein and expressed as R -factor:

$$R = \frac{\sum |F_{obs}| - |F_{calc}|}{\sum |F_{obs}|} \quad (4)$$

Usually for proteins, R -values of 0.3 to 0.4 for the best positioning of a phasing model often provided adequate initial estimates of phases.

The initial obtained electron density map by methods described above is hardly interpretable and its quality can be improved by several methods including solvent flattening or noncrystallographic symmetry (NCS) averaging and their variants/combinations. Those methods make advantage of the ability to distinguish protein and solvent regions in initial map or of internal symmetry.

7.1.4 Model building and structure refinement

The interpretation of the electron-density map by building and manipulating of the stick model of the known sequence within small sections of the electron-density map is intimately linked with the structure refinement process - cycles of map calculation and model building. Structure refinement improves the agreement of the calculated pattern from the model with the measured intensities. In order to reach better convergence of the intensities calculated from the model to the experimental ones, different factors have to be taken into account of the built molecule, like atoms oscillation and their occupancy.

7.2 Supplementary Tables and Figures

Data collection	
Wavelength (Å)	0.80
Temperature (K)	100
Space group	C2
Unit-cell parameters (Å, °)	a = 65.33, b = 46.01, c = 58.66, β = 120.43
Resolution (Å)	20.0–1.1 (1.2–1.1)
Unique reflections	61009 (13455)
Completeness (%)	100.0 (100.0)
Redundancy	7.0 (6.9)
I/σ(I)	10.0 (3.4)
Rsym(I) ^a (%)	8.7 (24.5)
Refinement	
Resolution (Å)	20.0–1.1 (1.128–1.1)
No. of reflections	57911 (3094)
Completeness (%)	100.0 (99.9)
Reflections in test set (%)	5.1
Rwork ^b (%)	12.0 (15.5)
Rfree ^b (%)	14.1 (18.2)
ESU ^c (Å)	0.015
Refined atoms	
Protein molecules/residues/atoms/protons	1/135/2387/1179
Water O atoms	358
Na ⁺ /Cl ⁻ ions	2/4
Mean B factors (Å ²)	
Wilson	6.5
Protein	8.4
Water	22.4
Na ⁺ /Cl ⁻ ions	12.7
Ramachandran plot ^d (%)	
Favoured	99.0
Allowed	1.0
Outliers	0
R.m.s.d. ^e from target geometry	
Bond lengths (Å)	0.014
Bond angles (°)	1.55
R.m.s.d. isotropic thermal factors (Å ²)	
Main-chain bonds	1.40
Main-chain angles	2.07
Side-chain bonds	2.96
Side-chain angles	4.15
PDB code	2v9v

Suppl. Table 1. Crystallographic data and refinement of SelB³⁷⁷⁻⁵¹¹. Values in parentheses are for the highest resolution shell.

^a $R_{\text{sym}}(I) = \frac{\sum_{\text{hkl}} \sum_i |I_i(\text{hkl}) - \langle I(\text{hkl}) \rangle|}{\sum_{\text{hkl}} \sum_i I_i(\text{hkl})}$; for n independent reflections and i observations of a given reflection; $\langle I(\text{hkl}) \rangle$ - average intensity of the i observations

^b $R = \frac{\sum_{\text{hkl}} ||F_{\text{obs}}| - |F_{\text{calc}}||}{\sum_{\text{hkl}} |F_{\text{obs}}|}$; R_{work} is for reflections that do not belong to the test set and R_{free} is for reflections that belong to the test set.

^c ESU, estimated overall coordinate error based on maximum likelihood.

^d Calculated with MolProbity (<http://molprobity.biochem.duke.edu/>; Davis *et al.*, 2004).

^e R.m.s.d. – root-mean-square deviation

Data collection	Native	Iodide	Phosphate
Wavelength (Å)	1.54	1.54	1.54
Temperature (K)	100	100	100
Space group	I222	I222	I222
Unit-cell parameters (a,b,c; Å)	59.2, 138.7, 141.7	59.2, 138.4, 141.5	59.2, 138.7, 141.7
Resolution (Å)	30.0-1.65 (1.74-1.65)	50.0-2.25 (2.38-2.25)	50.0-1.85 (1.95-1.85)
Unique reflections	68,502 (9,938)	52,394 (7,738)	49,649 (7,322)
Completeness (%)	97.3 (95.3)	98.3 (97.0)	99.3 (99.2)
Redundancy	4.7 (4.5)	4.7 (4.7)	3.5 (3.1)
I/σ(I)	20.5 (2.1)	13.4 (3.8)	16.3 (1.90)
R _{sym} (I) ^a (%)	4.5 (53.9)	9.3 (35.4)	4.6 (54.2)
Phasing			
Resolution (Å)	30.0-2.0		
Heavy atom sites	42		
Correlation coefficients ^b			
SHELXD CC _{all} /CC _{weak}	50.86/37.73		
SHELXE CC _{overall}	35.52		
CC _{free} left/right hand	62.69/26.92		
FOM ^c	0.582		
Refinement			
Resolution (Å)	20.0-1.65 (1.69-1.65)	30.0-2.25 (2.31-2.25)	20.0-1.85 (1.99-1.85)
No. of reflections	65,002 (4,662)	26,104 (1,839)	47,049 (3,157)
Completeness (%)	97.3 (95.0)	98.0 (96.3)	98.8 (93.0)
Reflections in test set (%)	5.1	5.0	5.1
R _{work} ^d (%)	16.8 (39.0)	18.3 (19.9)	16.6 (34.1)
R _{free} ^d (%)	19.8 (0.401)	23.7 (30.3)	20.2 (37.2)
Refined atoms			
Protein	1/135/2387/1179	-/-/-	1/441/-
molecules/residues/atoms/protons			
Water O atoms	723	264	509
PO ₄ ³⁻ /Cl ⁻ /I ⁻ ions	-/1/-	-/-/50	2/1/-
Mean B factors (Å ²)			
Wilson	20.0	32.2	23.9
Protein	15.1	26.9	21.0
Water	38.4	28.7	38.2
Ramachandran plot ^f (%)			
Favoured	97.9 (424/433)	97.7 (421/431)	97.7 (431/441)
Allowed	1.9 (8/433)	2.1 (9/433)	2.3 (10/441)
Outliers	0.2 (1/433)	0.2 (1/433)	0.0
R.m.s.d. ^g from target geometry			
Bond lengths (Å)	0.012	0.010	0.010
Bond angles (°)	1.34	1.22	1.19
R.m.s.d. B-factors (Å ²)			
Main-chain bonds	0.726	0.423	0.563
Main-chain angles	1.122	0.787	0.894
Side-chain bonds	2.005	1.402	1.604
Side-chain angles	3.152	2.318	2.540
PDB code	3BC8	3BCA	3BCB

Suppl. Table 2. Crystallographic data and refinement of SecS.

^a $R_{\text{sym}}(I) = \sum_{\text{hkl}} \sum_i |I_i(\text{hkl}) - \langle I(\text{hkl}) \rangle| / \sum_{\text{hkl}} \sum_i |I_i(\text{hkl})|$; for n independent reflections and i observations of a given reflection; $\langle I(\text{hkl}) \rangle$ - average intensity of the i observations

^b $CC = [\sum w E_o E_c \sum w - \sum w E_o \sum w E_c] / \{ [\sum w E_o^2 \sum w - (\sum w E_o)^2] [\sum w E_c^2 \sum w - (\sum w E_c)^2] \}^{1/2}$; w - weight (see http://shelx.uni-ac.gwdg.de/SHELX/shelx_de.pdf for full definitions).

^c FOM - figure of merit = $|F(\text{hkl})_{\text{best}}| / |F(\text{hkl})|$; $F(\text{hkl})_{\text{best}} = \sum_{\alpha} P(\alpha) F_{\text{hkl}}(\alpha) / \sum_{\alpha} P(\alpha)$

^d $R = \sum_{\text{hkl}} | |F_{\text{obs}}| - |F_{\text{calc}}| | / \sum_{\text{hkl}} |F_{\text{obs}}|$; R_{work} is for reflections that do not belong to the test set and R_{free} is for reflections that belong to the test set.

^e ESU, estimated overall coordinate error based on maximum likelihood.

^f Calculated with MolProbity (<http://molprobity.biochem.duke.edu/>; Davis *et al.*, 2004).

^g R.m.s.d. - root-mean-square deviation

Data Collection	Native	Manganese
Wavelength (Å)	0.91841	1.89218
Temperature (K)	100	100
Space Group	P2 ₁	P2 ₁
Unit Cell Parameters (a, b, c; Å)	48.9, 96.1, 71.8	48.4, 96.1, 71.8
Resolution (Å)	50.00-2.00 (2.03-2.00)	35-2.84 (2.91-2.84)
Reflections		
Unique	42189 (2072)	28639 (2067)
Completeness (%)	97.9 (95.8)	97.1 (95.6)
Redundancy	3.9 (3.7)	1.9 (1.8)
I/σ(I)	19.7 (1.2)	12.4 (2.6)
R _{sym} (I) ^a	4.7 (80.7)	5.3 (31.8)
Refinement		
Resolution (Å)	50.00-2.00 (2.05-2.00)	
Reflections		
Number	40040 (2700)	
Completeness (%)	97.4 (90.1)	
Test Set (%)	5.0	
R _{work} ^b	21.7 (37.6)	
R _{free} ^b	25.2 (42.3)	
ESU (Å) ^c	0.134	
Refined atoms		
RNA Molecules/Residues/Atoms	2 / 172 / 3888	
Water O atoms	195	
SO ₄ ²⁻ / Acetate/ Glycerol	2/2/2	
Mean B-Factors (Å ²)		
Wilson	46.8	
RNA	46.9	
Water	43.5	
R.m.s.d. ^d from Target Geometry		
Bond Lengths (Å)	0.010	
Bond Angles (°)	1.748	
R.m.s.d. B-Factors (Å ²)		
Main Chain Bonds	0.389	
Main Chain Angles	0.647	
Side Chain Bonds	1.597	
Side Chain Angles	2.330	
PDB ID	-	

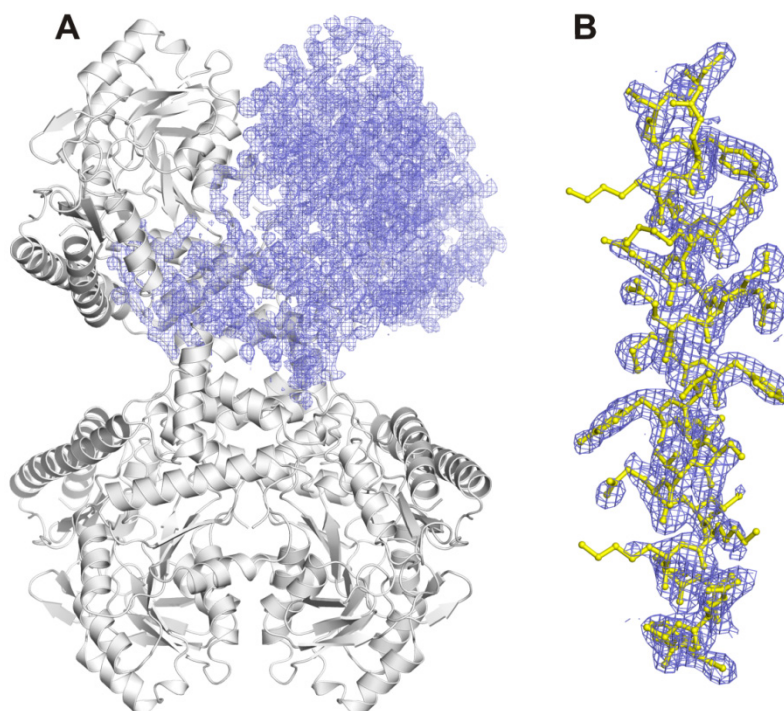
Suppl. Table 3. Crystallographic data and refinement of tRNA^{Sec}. Values in parentheses are for the highest resolution shell

^a $R_{sym}(I) = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$; for n independent reflections and i observations of a given reflection; $\langle I(hkl) \rangle$ – average intensity of the i observations

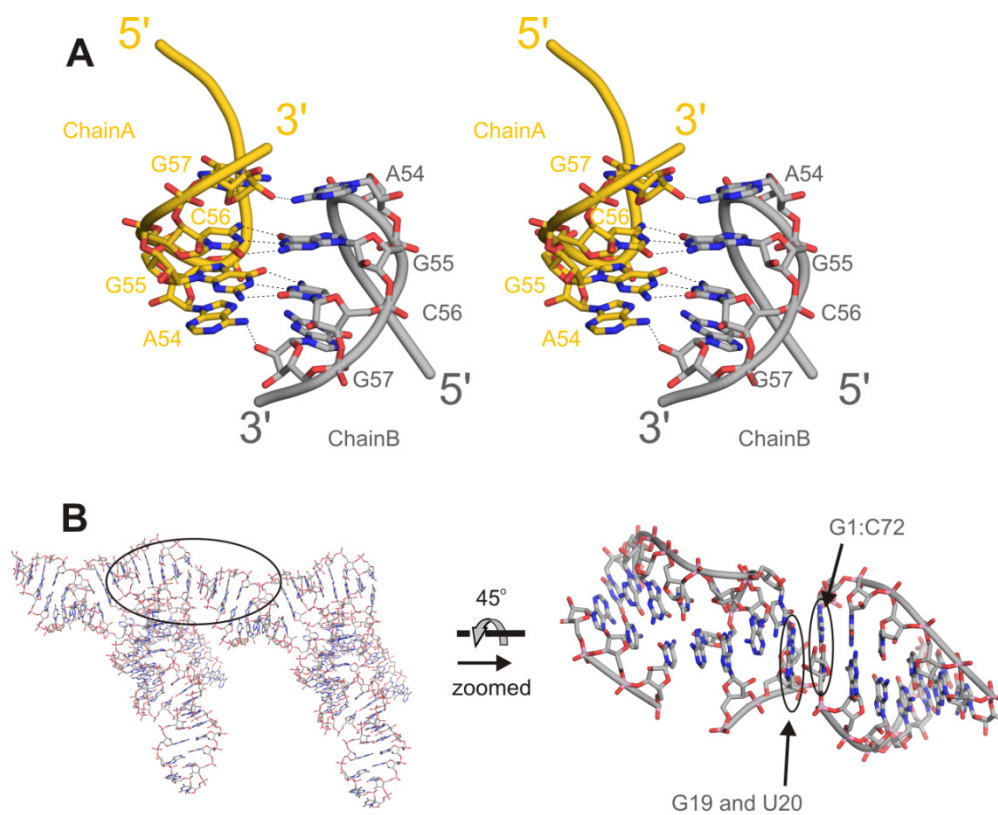
^b $R = \frac{\sum_{hkl} (|F_{obs}| - |F_{calc}|)}{\sum_{hkl} |F_{obs}|}$; R_{work} – $hkl \notin T$; R_{free} – $hkl \in T$; R_{all} – all reflections; T – test set

^c ESU – estimated overall coordinate error based on maximum likelihood

^d R.m.s.d. – root-mean-square deviation



Suppl. Figure 1. Electron density. Experimental SIRAS electron density (blue mesh) at 2.0 Å resolution after solvent flattening contoured at the 1σ level. (A). Density covering one molecule of a tetramer. The other three protomers are indicated as light gray ribbons. (B). SIRAS density of a α -helical element with the final model indicated as yellow ball-and-sticks.



Suppl. Figure 2. Crystal contacts of tRNA^{Sec} in $P2_1$ crystal form. Variable loop to loop interaction (A). Staking interaction of acceptor stem G1:C72 with D loop bases G19 and U20 (B). Chain A colored gold and chain B colored silver. Dashed lines represent H-bonds.

7.3 List of abbreviations

A	adenosine
aa	amino acid
AA	acrylamide
ATP	adenosine-5' - triphosphate
bp	base pair
BSA	bovine serum albumin
°C	degree celsius
C	cytidine
cDNA	complementary DNA
CTP	cytidine-5' - triphosphate
cpm	counts per minute
Cys	cysteine
Da	dalton
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide
DTT	1,4-dithiothreitol
EDTA	ethylenediamine-N, N, N', N'-tetraacetic acid
E.coli	Escherichia coli
EFSec	elongation factor
EMSA	electrophoretic mobility shift assay
fmol	femtomole
G	guanosine
GTP	guanosine -5'-triphosphate
HEPES	N-2-hydroxyethylpiperazine-N-2-ethane sulfonic acid
i6A	6-isopentenyl-A
kDa	kilodalton
l	liter
M	molar
MBAA	N,N'-methylenebisacrylamide
min	minute
ml	milliliter
mRNA	messenger RNA
Mw	molecular weight
m1A	N1-methyladenosine
mcm5U	5-methylcarboxymethyluridine
mcm5Um	5-methylcarboxymethyl-2'-O-methyluridine
nm	nanometer
nt	nucleotide
OH	hydroxyl
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PLP	pyridoxal-5'-phosphate
PSTK	O-phospho-L-seryl (PSer)-tRNA ^{Sec} kinase
RNA	ribonucleic acid
RT	room temperature
rpm	revolutions per minute
Se	selenium
Sec	selenocysteine amino acid
SECIS	selenocysteine insertion sequence

SecS	selenocysteine synthase
SBP2	Sec insertion sequence-binding protein 2
SDS	sodium dodecyl sulfate
SPS	selenophosphate synthase
SRE	selenocysteine codon Redefinition Element
T	thymidine
TBE	Tris-Borate-EDTA buffer
TEMED	N, N, N', N'-tetramethylethylenediamine
tRNA ^{Sec}	transfer RNA
U	uridine
UTP	uridine-5'-triphosphate
UV	ultraviolet
WH	winged-helix motifs
μ	micro
μl	microliter
ψ	pseudo uridine

8 ACKNOWLEDGEMENTS

At the end of my Ph. D. time I would like to thank peoples who have contributing to success of my work.

First of all I am very grateful to my supervisor Prof. Dr. Markus C. Wahl which guided me through difficulties and uncertainties of scientific life; for giving me a chance to work in his lab, for his fruitful discussions and advices of experimental results and continuous financial support.

I am thankful to Prof. Dr. Ralf Ficner and Prof. Dr. Detlef Doenecke for accepting to be my members of the thesis committee as well as for their useful suggestions during my annual progress reports.

Thanks a lot to Dr. Dolph L. Hatfield, Prof. Dr. Vadim N. Gladyshev and their group members for fruitful collaboration on SecS project.

I thank peoples from Max Planck Institute for biophysical chemistry (MPI) in Goettingen, particularly members of X-ray Crystallography Group and Cellular Biochemistry Department, Microscopy and Mass Spectrometry groups for creating nice working atmosphere, sharing ideas and lab protocols.

I would like to thank the members and former members of our crystallography group for their help and advice in every day life and the fun in “get together” on weekends.

I am thankful to Elke Penka, Jürgen Wawrzinek from MPI and Claudia Aligns from Free University of Berlin for their excellent technical assistance.

To the head of the Bioanalytical Mass Spectrometry Group, Dr. Henning Urlaub, and to Monika Raabe and Uwe Plessmann, I thank for the protein identifications.

Special thanks to Dr. Bernhard Loll for his enormous help with writing this manuscript.

Many thanks goes to:

...Dr. Steffen Burkhardt and coordination GGNB office for minimizing bureaucracy of everyday life

...Prof. Dr. Holger Stark and Dr. Niels Fischer for sharing ideas about SelB action on the ribosome.

...Dr. Alena Paleskava for introducing aminoacylation method.

...Juliane Moses for her help in solving everyday problems

...Russian friends in Cellular Biochemistry Department of MPI for their every day help, useful scientific discussions during “tee time”.

I am indebted to my family in Russia for their love and continuous encouragement, especially my dear parents Mikhail and Nadezda.

Lastly, my very special thanks to my beauty wife Ekaterina, essential part of my life, for her unconditional love and care.

9 CURRICULUM VITAE

Personal Information

Name Oleg Ganichkin

Born: March 4, 1982

Place of Birth: Makhachkala, Russian Federation

Education and research experience

- 07/2006 -present Ph.D., University of Goettingen, Germany.
(The research was carried out at the Max-Planck Institute for Biophysical Chemistry, Göttingen and at the Free University Berlin, Germany).
Project title: Crystal structure analysis of selenocysteine biosynthesis components.
Supervisor: Prof. Dr. Markus C. Wahl.
- 2004-2006 M.S. in Molecular Biology, Moscow State University, Russia
Project title: Study of properties of the novel two-subunit NAD⁺-dependent DNA ligase from bacteriophage T5.
Supervisor: Dr. Vladimir N. Ksenzenko
- 1999-2004 Diploma in Biology, Saratov State University, Russia
Project title: Characterization of the novel two-subunit NAD⁺-dependent DNA ligase from bacteriophage T5.
Supervisor: Dr. Vladimir N. Ksenzenko

Publications

Ganichkin O.M, Xu X.M., Carlson B.A., Mix H., Hatfield D.L., Gladyshev V.N., Wahl M.C. 2008. Structure and catalytic mechanism of eukaryotic selenocysteine synthase. *J Biol Chem.* **283**:5849-65

Ganichkin O, Wahl MC. 2007. Conformational switches in winged-helix domains 1 and 2 of bacterial translation elongation factor SelB. *Acta Crystallogr D Biol Crystallogr.* **Pt 10**:1075-81