Stretching the genetic code – incorporation of selenocysteine at specific UGA codons in recombinant proteins produced in *Escherichia coli*

Olle Rengby

Division of Biochemistry
Department of Medical Biochemistry and Biophysics
Karolinska Institutet

Stockholm 2008

Printed by Laserics Digital Print AB

ISBN 978-91-7409-260-8

To Sara

To My Family

The most exciting phrase to hear in science, the one that heralds new discoveries, is not “Eureka!” (I found it!) but “That's funny ...”

Isaac Asimov

(1920 - 1992)

Abstract

Selenocysteine (Sec) exists in all domains of life and represents the 21st naturally occurring amino acid. A Sec residue is co-translationally incorporated at a predefined opal (UGA) codon. UGA codons normally encode for translational stop via the protein release factor 2 (RF2). The incorporation mechanism of Sec into the selenoprotein involves complex machineries, dependent on several specific factors that differ between organisms. For both eukaryotes and prokaryotes, an mRNA secondary structure, called a Sec insertion sequence (SECIS) element, is required. Sec insertion systems for eukaryotes are different from that of bacteria. Due to the differences between species, recombinant expression of eukaryotic selenoproteins in *E. coli* is not a trivial task. However, our group has previously been able to overcome this species barrier and successfully expressed the mammalian selenoprotein thioredoxin reductase 1 (TrxR1) and other selenoproteins in *E. coli*. In this thesis, we have further developed this recombinant selenoprotein production system. We have also further characterized the recombinantly expressed rat TrxR1.

We have studied growth conditions affecting yield of the recombinant selenoprotein when expressing rat TrxR1, using various levels of the selenoprotein-encoding mRNA and growth in different types of medium. Guided by Principal Component Analysis (PCA), we discovered that the most efficient bacterial selenoprotein production conditions were obtained using high-transcription levels in the presence of the *selA, selB* and *selC* genes, with induction of production at late exponential phase. We also constructed an *E. coli* strain with the endogenous chromosomal promoter of the gene for relase factor 2 (RF2), *prfB*, replaced with the titrable P_{BAD} promoter. In a turbidostatic fermentor system the simultaneous impact of *prfB* knockdown on growth and on recombinant selenoprotein expression was studied, using rat TrxR1 as the model selenoprotein. This showed that lower levels of RF2 correlate directly to an increase of Sec incorporation specificity, while also affecting total selenoprotein yield concomitant with a slower growth rate.

Recombinant rat TrxR is expressed as a mixture of full-length and two-amino acid truncated subunits. Phenylarsine oxide (PAO) Sepharose can be used to enrich the Sec-containing protein. We investigated the mechanism of this purification by extensively purifying recombinant rat TrxR1, which gave an enzyme with about 53 U/mg in specific activity, which was higher than ever reported. Surprisingly, only about 65% of the subunits of this TrxR1 preparation contained Sec, which revealed a theoretical maximal specific activity of about 80 U/mg for TrxR with full Sec content. The high specific activity revealed that the inherent turnover capacity of rat TrxR1 must be revised, and that the efficiency of bacterial Sec incorporation may be lower than previously believed. We also discovered and characterized tetrameric forms of recombinant TrxR1, having about half the specific activity compared to the dimeric protein in relation to Sec content.

In conclusion, this thesis describes limiting factors for recombinant selenoprotein production in *E. coli* and shows how this production system can be optimized for higher yield and specificity. The results may prove to be of importance for the further development of *E. coli* as a useful source for synthetic selenoproteins. Results are also presented and discussed regarding the catalytic capacity of rat TrxR1 and novel multimeric states of the protein, which could represent unknown regulatory features of TrxR having potential physiological importance.

List of publications

This thesis is based on the following articles, which will be referred to in the thesis by their roman numbers.

I. **Olle Rengby**, Linda Johansson, Lars A. Carlson, Elena Serini, Alexios Vlamis-Gardikas, Per Kårsnäs and Elias S. J. Arnér. Assessement of production conditions for efficient use of *Escherichia coli* in high-yield heterologous recombinant selenoprotein synthesis. *Applied and Environmental Microbiology*, 2004 Sep:70(9):5159-67.

II. **Olle Rengby** and Elias S. J. Arnér. Titration and conditional knockdown of the *prfB* gene in *Escherichia coli*: Effects on growth and overproduction of a recombinant selenoprotein. *Applied and Environmental Microbiology*, 2007 Jan:73(2):432-41

III. **Olle Rengby**, Qing Cheng, Marie Vather, Hans Jörnvall and Elias S. J. Arnér. Highly active dimeric and low-activity tetrameric forms of selenium-containing rat thioredoxin reductase. *Submitted Manuscript* (2008).

Table of Contents

List of Abbreviations

AA	Amino Acid
Cys	Cysteine
DMPS	2,3-dimercaptopropane sulfonic acid
DTNB	5,5´-dithiobis (2-nitrobenzoic acid)
DTT	Dithiothreitol
FAD	Flavin adenine dinucleotide (oxidized form)
FDH	Formate dehydrogenase
GPx	Glutathione Peroxidase
GSH	Glutathione
GSSG	Glutathione disulfide
IPTG	Isopropyl-ß-D-thiogalactose
NADPH	Nocotinamide dinucleotide phosphate (reduced form)
nt	Nucleotide
ORF	Open reading frame
PAO	Phenylarsine oxide
RF1	Release factor 1
RF2	Release factor 2
Se	Selenium
Sec	Selenocysteine
SECIS	Selenocysteine insertion sequence
TrxR	Thioredoxin reductase
X	Unspecified amino acid
3´UTR	3´-untranslated region

1 Introduction

1.1 Selenium

Jöns Jacob Berzelius discovered selenium in the year 1817, when searching for the cause of an illness among workers at a sulfuric acid production facility. Copper pyrites, from the copper mine in Falun, were used as the source of sulfur and when Berzelius examined the lead chamber where the pyrites were processed, he found a red substance. A closer study showed that it was a new substance, and he named it Selenium after the Greek goddess of the moon, Selene.

Selenium was considered mainly a toxic agent for several decades. However, research performed in the latest half of the last century has revealed that selenium plays a very important role in life on earth. Selenium was identified as a essential trace element for bacteria in 1954 [1] and for mammals in 1957 [2].

Selenium is located between sulfur and tellurium in the periodic table and exists in six stable isotopes; 82 (9.19%), 80 (49.82%), 78 (23.52%), 77 (7.58%), 76 (9.02%) and 74 (0.87%). Some of the isotopes have special properties, ^{77}Se, for example, has a nuclear spin of ½ and is therefore suitable for *NMR* studies [3]. Another example is the artificially produced isotope ^{75}Se that is a gamma-emitter and therefore commonly used for diagnostic applications in medicine [4-6]. Selenium is at the 70th place in abundance among the naturally occurring elements in the earth's crust, with an average of 0.05 mg kg-1 [7], but it is unevenly distributed with values as high as 8 g per kg down to 0.005 mg Se per kg [8].Sulfur and selenium is compared in **Table 1**.

1.1.1 Selenium and health

The main part of our selenium intake comes from vegetables grown in soil rich of selenium, or from meat. Se contents in food from different countries are summarized elsewhere [9]. The selenium is absorbed by the sulfur pathways in plants and then taken up by humans or animals.

Table 1. Some physical properties of selenium and sulfur

Property	Selenium	Sulfur
Atomic number	34	16
Atomic weight	78.96	32.07
Atomic radius	1.17 Å	1.04 Å
Crystal structure	Hexagonal	Orthorhombic
Electron negativity	2.4	2.5
Melting point at 1 atm	220.5°C	112.8°C
Boiling point at 1 atm	685°C	445°C
Common compounds		
Inorganic	Selenite (SeO_3^{-2})	Sulfide (SO_3^{2-})
	Selenate (SeO_4^{2-})	Sulfate($SO4^{2-}$)
Organic	Sec (see **Table 4**)	Cys (see **Table 4**)
	Selenol (R-SeH)	Thiol (R-SH)
	Selenolate ($R\text{-}Se^-$)	Thiolate ($R\text{-}S^-$)
	Diselenide (R-Se-Se-R)	Disulfide (R-S-S-R)
	Selenenylsulfide (R-Se-S-R)	

The recommended daily intake of selenium is 55 μg. Diseases associated with selenium deficiency have great effects on human life in areas with soil containing low amount low selenium (especially the mountain areas of north east to south central China). Keshan disease is a type of cardiomyopathy characterized by abnormal ECG patterns, enlargement of the heart and congestive heart failure [10, 11]. The role of selenium in Keshan disease has been debated and it seems like the selenium deficiency causes changes in the viral genome of a normally non-pathogenic virus and changes it to a pathogenic one [12]. However, the symptoms are reversed by selenium supplementation.

Moderate selenium deficiency is connected with several conditions, such as increased cancer risk, decreased immune function and neurological conditions such as Parkinson's and Alzheimer's disease [13]. Selenium deficiency as well as overexposure to selenium, is toxic for mammals and the window is narrow between the hazardous nutritional levels and required levels. Selenium deficiency induces

symptoms and diseases mentioned above, whereas too high intake of selenium causes nail sloughing and hair loss in humans [9]. A recent case of selenium poisoning illustrates the danger with selenium toxicity. In 2006, a 75-year-old man died from cardiac arrest after eating 10 g of sodium selenite. He had read on the Internet that selenium had positive effects against prostate cancer [14].

Selenium intake has been shown to have impact on different types of cancer and this is reviewed elsewhere [15-17]. One of the first studies on the correlation between selenium and cancer was made in 1943 [18] where it was reported that selenium could cause cancer. This study has been criticized for poor experimental design, lack of controls and control of purity of the substance provided [13]. In a study published in 1996 by Clark et al the effects of selenium supplementation (200 μg selenium/day) significantly reduced the incidence of lung, colon, rectum or prostate cancer [19]. The highest preventive effect was in those persons with the lowest baseline selenium concentration.

Right now one of the largest investigations to prevent cancer is ongoing. The trial is called SELCT and is a randomized, double blinded and placebo controlled trial on selenium and vitamin E prevention on prostate cancer [20, 21]. The trial, which started in 2001, involves 32000 men and is planned to end in 2013.

1.2 Selenoproteins

Selenium is essential for human life and the main function of selenium is in the form of selenoproteins. Selenoproteins take part in some of the most fundamental functions in life, such as DNA synthesis and protection against oxidative damage. Normal selenium intake is therefore required to maintain function for these proteins.

Selenoproteins exist in all domains of life and for most of the higher organisms selenoproteins are essential. Some organisms have several selenoproteins such as the symbiotic deltaproteobacterium of gutless worm Olavius algarvensis where 57 selenoproteins [22] have been identified. However, in *C. elegance* only one selenoprotein has been identified, thioredoxin reductase [23]. Interestingly, some higher organisms seem to have lost their selenoproteins, such as plants and certain insects have also been reported not to have any selenoproteins, probably as an evolutionary responce to the change of the living environment [24].

The first selenoprotein was discovered in 1973; the mammalian gluthathione peroxidase [25]. The same year glycine reductase and formate dehydrogenase were found in prokaryotes [26, 27]. Selenoproteins have shown to be essential in mice where the knockout of Sec-encoding $tRNA^{Sec}$ results in early embryonic lethality [28].

1.2.1 Mammalian selenoproteins

In human, 25 selenoproteins gene were discovered by bioinformatic means in 2003 [29]. Several of the mammalian selenoproteins have yet unknown functions and still need to be thoroughly characterized. In **Table 2** some of the human selenoproteins are described shortly, for more information see references [30-32].

Table 2. Examples of selenoproteins from human

Selenoprotein	Comments	Ref.
Glutathione peroxidases (GPx)	GPx has the function to reduce peroxides and serves as a protection against oxidative damage in the cell	[25]
GPx1	Ubiquitous cytosolic form with limited role during normal cell development but with an active role in protection against selenium toxic and commonly used control of selenium status	[33, 34]
GPx2	Epithelium specific gastrointestinal GPx that is also expressed at elevated levels in several breast cancer lines	[35, 36]
GPx3	Is glycosylated and secreated to extracellular compartments. GPx3 has a wider range of substrates compared with GPx1 and GPx2	[37]
GPx4	Essential GPx localized in the cytosol, mitochondria and nuclei. Can reduce phosphoslipid- and cholestrol-hydroperoxides.	[38, 39]
Thioredoxin reductase (TrxR)	Catalyzing the reduction of Trx	
TrxR1	Predominantly cytosolic form of TrxR, important for cell development and protection against oxidative stress	[40]
TrxR2	Mitochondrial TrxR important for cell development and protection against oxidative stress	[41]
TGR	Expressed mainly in testis and contains an N-terminal monothiol glutaredoxin domain	[42, 43]

Table 2. Examples of selenoproteins from human (continued)

Iodothyronine deiodinases (IDO)	Three IDO´s has been identified with the size around 27-32 kDa. IDO function as an activator or inactivator of thyroid hormones.	[44]
Selenophsophate synthetase 2	50 kDa Se protein that catalyses phosphorylation of selenide with ATP to form selenophosphate (See Sec insertion section)	[45]
Selenoprotein P	A 43 kDa selenoprotein with up to 10 Sec residues. Acts as a selenium transporter protein but has also been shown to have antioxidant effects and a redox active site	[46, 47]
Sep 15	Has a C-X-U redox active motif and a classic Trx-fold. Function not fully determined	[48, 49]
SelH, SelM, SelO, SelP, SelT, SelV, SelW	C-XX-U redox motif	[29, 30]
Other selenoproteins		
SelN, SelS, SelI GPx6	Function unknown or reviewed elsewhere	[29, 30]

1.2.2 Selenoproteins of other organisms

As mentioned, selenoproteins exist in all domains of life, but not in all organisms. Based on a computer search, on 12 complete and 33 incomplete prokaryotic genomes made in 2004, about 20% of these contained genes that were concluded to encode for selenoproteins [50]. The number of genomes containing genes for selenoproteins are constantly increasing, though. In a study, made by Gladyshev and co-workers in 2008, based on data from the Global Ocean Sampling (GOS) expedition, more than 3600 selenoprotein genes were identified [51]. For the genes sequenced in that study, 90% were bacterial genes and 2.8% eukaryotic genes [52]. The amount of selenoprotein genes varies in prokaryotic organisms from one to ten [50], most of which are probably having a redox function. In **Table 3,** some characteristics of selenoproteins from some other organisms than mammals are described.

Table 3. Examples of selenoproteins in other organisms than mammals

Organism	Selenoprotein	Comments	Ref.
Haemophilus influenzae	Selenophosphate synthetase (*selD*)	Part of the selenoprotein synthesis	[53]
E. coli and other bacteria	Formate dehydrogenase H (*fdhF)*	Convert formate to carbondioxide. Essential in anaerobic growth	[54]
E. coli and other bacteria	Formate dehydrogenase N (*fdnG*)	Convert formate to carbondioxide.	[54]
E. coli and other bacteria	Formate dehydrogenase O (*fdoG*)	Convert formate to carbondioxide.	[55]
Eubacterium acidaminophilum	Glycine reductase (*grdA*)	Redox active. Two proteins; protein A (CXXU) and protein B UXXCXXC functioning by forming selenoether.	[56]
Clostridium sticklandii	Proline reductase (*prdB*)	The 26 kDa subunit of Proline Reductase contains a Sec residue and have redox function.	[57]
Methanococcus voltae	F420 non-reducing hydrogenase (*VhuU*)	Redox function, binds to Ni atom.	[58]

1.2.3 The specific selenoproteins in *E. coli*

In *E. coli*, three selenoproteins have been identified. All these are forms of formate dehydrogenase [55, 59, 60]. The three forms differ from each other in the way that they are expressed, with formate dehydrogenase encoded by the *fdoG* gene constitutively expressed in aerobic conditions whereas formate dehydrogenases encoded by the *fdhF* and *fdnG* genes are expressed in anaerobic conditions. *FdnG* is also induced by nitrate [60].

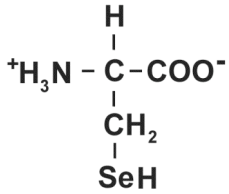
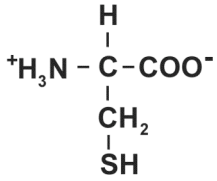
1.3 Selenocysteine

Selenium exists in selenoproteins in the form of selenocysteine, usually referred to as the 21st naturally occurring amino acid. Sec is a homolog to cysteine (Cys) with a selenium atom in place of the sulfur and the two amino acids differ due to the chemical properties of selenium and the biologically complexity of Sec insertion, as discussed herein.

1.3.1 Chemical Properties of Sec

Sec has a lower pKa than Cys (5.2 compared to 8.3) and that makes Sec, at physiological pH, more easily forming its anionic selenolate form than the corresponding thiolate formation of Cys, which instead is mainly present in its protonated thiol form. Sec is therefore, in combination with a high nucleophilicity, significantly more reactive than Cys. Sec was discovered as a unique amino acid in 1976 [61], but it was not until the sequencing of the first selenoprotein gene, in 1986, it was discovered that Sec was incorporated at the position of an in-frame UGA codon [59, 62]. Some features of Sec and Cys are listed in **Table 4**.

Table 4. Properties of the amino acids selenocysteine and cysteine

	Selenocysteine	**Cysteine**
Codon	UGA[a]	UGU, UGC
Abbreviations	Sec, U	Cys, C
pKa	5.2	8.3
Structure		

[a] For some organisms e.g. *Mycoplasma* and some mitochondria the UGA encodes tryptophan [63]

1.3.2 Function of Sec in proteins

A selenoprotein contains one or more Sec residues in the polypeptide chain. A Sec in selenoproteins provides some special properties due to the chemistry of Se and Sec. However, for most identified selenoproteins the specific functions of Sec are still unknown. Sec seems, however, to be mainly involved in redox reactions, while Cys has much wider spread in its functions, e.g. redox, structural Cys and metal binding [64]. Some of the known actions of Sec in protein are presented here.

Glutathione peroxidases (GPx) exist in human in seven different forms, five contain a Sec residue and two having a Cys in the place of the Sec, and are

expressed in different tissues. In human cells, GPx have a protective role against oxidative damage through the reduction of hydroperoxides. Sec in GPx forms an ionized selenolate that can react with H_2O_2, forming a selenenic acid, which is subsequently regenerated to a selenolate with glutathione [65]. A Sec to Cys mutation results in 1000-fold lower activity [66].

Iodothyronine Deiodinases 1 and 2 (DIO1 and DIO2) have the function to activate the thyroid hormone thyroxine (T_4) by converting it into triidothyronine (T_3) and T_3 into diidothyronine (T_2). DIO3 has the function of inactivating T_4 and T_3 [67]. Oxidized DIO1 forms a selenenyl-iodide intermediate with the thyroid hormone, whereas DIO2 and DIO3 do not [68]. Expressed Sec-to-Cys mutants of these enzymes showed a 100, 10 and 2-6 times decrease in turnover for DIO1, DIO2 and DIO3, respectively [69-71].

Formate dehydrogenase H has the function of converting formate to carbon dioxide. The role of Se of the Sec residue is to coordinate a molybdenum atom to molybdenum-pterins [72]. A Sec-to-Cys mutant of fdhH decreases the k_{cat} from 2800 sec^{-1} to 9 sec^{-1} [73].

Glycine reductase has the function of converting glycine to acetyl phosphate. The enzyme functions with two larger proteins of which the first protein forms a Schiff base with the glycine [74]. The Sec in glycine reductase attacks in an ionized selenolate form the Schiff base and the cleaved produced is transferred to the second protein that adds phosphate and forms acetyl phosphate. It is likely that the Sec is essential for activity of the enzyme [75].

1.4 Sec insertion – Stretching the genetic code

Sec insertion is in all organisms that express selenoproteins (gram-negative and gram-positive bacteria, mammals and archaea) incorporated at a UGA codon, directed by an mRNA secondary structure, called a Sec insertion sequence (SECIS) element. However, the structure and location of the SECIS vary in the different organisms and for different selenoproteins in the same organism. The fundamental difference in location of the SECIS elements between mammals and *E. coli* is schematically shown in **Figure 1**. The Sec insertion mechanisms also differ in other details and will be described in the following sections.

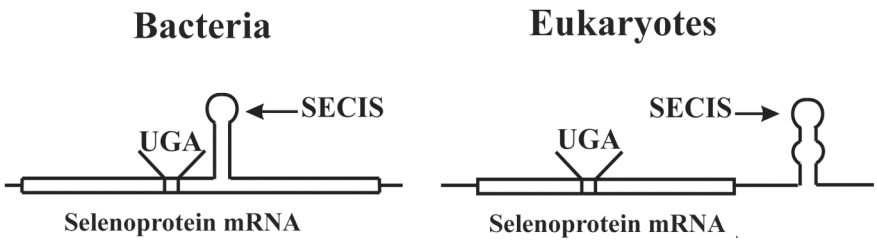


Figure 1. Schematic figure showing typical location of SECIS from bacteria and eukaryotes. The position of the bacterial SECIS is usually 11 nt downstream of the Sec-encoding UGA within the open reading frame while the eukaryotic can be hundreds of nt downstream in the 3´UTR.

1.4.1 Sec insertion and Selenoprotein synthesis in eukaryotes

The bacterial Sec incorporation system has been well known for several years thanks to a thorough work of Böck *et al* and also of other groups. The knowledge of the eukaryotic Sec incorporation system has made several advances in the last decade, but still there are factors in the eukaryotic system that need to be further studied.

Briefly, the mammalian Sec incorporation system consists mainly of the following components: The Sec specific transfer RNA, $tRNA^{(Ser)Sec}$, which is synthesized to its Sec containing form directly on the tRNA and is first charged with a serine [76, 77]. Sec is charged onto the tRNA using the action of Sec synthase, which also associates with a larger complex mediating Sec incorporation [78]. Selenophosphate synthetase (SPS) consists in two forms in mammalian cells, SPS1 and SPS2. SPS1 probably functions as a salvage pathway for Sec recycling. Knockdown of SPS2 has shown that SPS2 is the active selenium donor for Sec synthesis [76, 77]. A specific elongation factor, EFSec, delivers $tRNA^{(Ser)Sec}$ to the ribosomal A site for the insertion of Sec. eEFSec does, however, not bind directly to the SECIS; this connection is mediated by the SEICS binding protein 2 (SBP2). SBP2 also binds to the large subunit of the ribosome and a conserved part of the SECIS. Knockdown studies of SBP2 have shown to severely decrease the selenoprotein synthesis in cell lines [79]. Ribosomal binding protein L30 is an additional factor that binds the ribosome on its SBP2 binding site and takes the place of SBP2 after connecting with the ribosome [80]. The eukaryotic Sec insertion system is schematically shown in **Figure 2**.

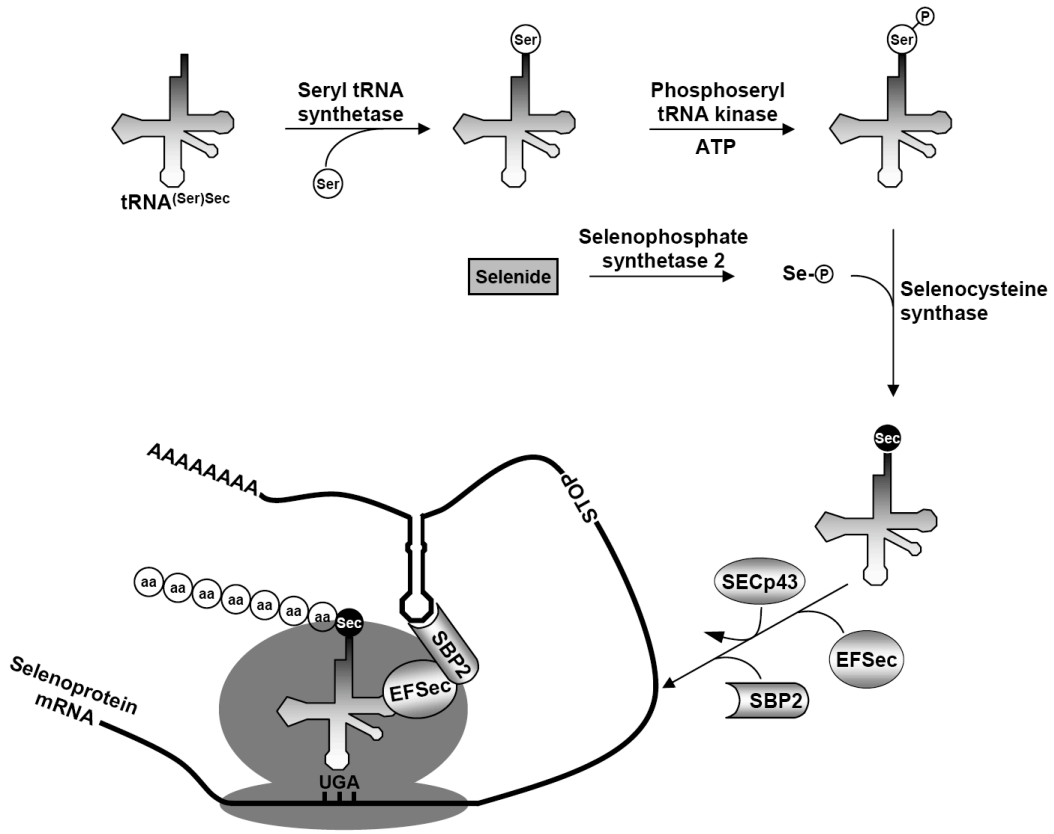


Figure 2. Schematic drawing of the eukaryotic Sec insertion mechanism. See text for details.

1.4.2 Sec insertion and selenoprotein synthesis in archaea

The archaeal Sec biosynthesis is reviewed elsewhere [81]. Archaea has, as the mammalian Sec incorporation system, the SECIS located in the untranslated region of the selenoprotein mRNA [82, 83]. Similarly to the eukaryotic eEFSec, the Sec-specific elongation factor (aSelB) does not bind the SECIS directly [81]. However, the aSelB N-terminal part has been confirmed to have an EF-Tu-like structure [84]. In mammalian systems, additional factors having the function of the C-terminal part of bacterial SelB have been discovered, called SBP2 [85], as also discussed above. A corresponding factor has not yet been found in Archaea.

1.4.3 Sec insertion and selenoprotein synthesis in bacteria

The bacterial Sec incorporation system has been studied mainly using *E. coli* as the model organism. In all bacteria producing selenoproteins the Sec insertion is guided by a SECIS element. However, the structure of the SECIS differs between gram-positive and gram-negative bacteria and the SelB/$tRNA^{Sec}$ from *E. coli* is not compatible with the gram-positive *Eubacterium acidaminophilum* [86].

This thesis has its main focus on selenoprotein production in *E. coli*, therefore the translational machinery of selenoproteins in *E. coli* will here be introduced in more detail.

Incorporation of Sec into the growing polypeptide chain is a complex procedure. The Sec incorporation has been thoroughly characterized for *E. coli* in the laboratory of August Böck [87]. Except for a selenium source, four gene products, SelA, SelB, SelC and SelD are requiered (these factors are summarized in **Table 5**).

Also the bacterial type SECIS element at the mRNA is required for a functional selenoprotein synthesis. This system is illustrated in **Figure 3** and briefly described in the following text.

Table 5. Factors necessary for selenoprotein synthesis in *E. coli*

Parameter	Comment	References
Selenium	Selenium can enter the bacteria via the sulfur pathway and/or reduced by the thioredoxin system	[88-90]
SECIS element	An mRNA secondary structure forming a hairpin loop that guides the Sec insertion	[83, 91]
SelA	Selenocysteine synthase charges the selenium to make Sec from Ser on the tRNASec	[92]
SelB	Sec-specific elongation factor	[93, 94]
SelC	Sec specific tRNA, $tRNA^{Sec}$, first charged with Ser	[95]
SelD	Selenophosphate synthetase transforms selenide to monoselenophosphate	[96-98]

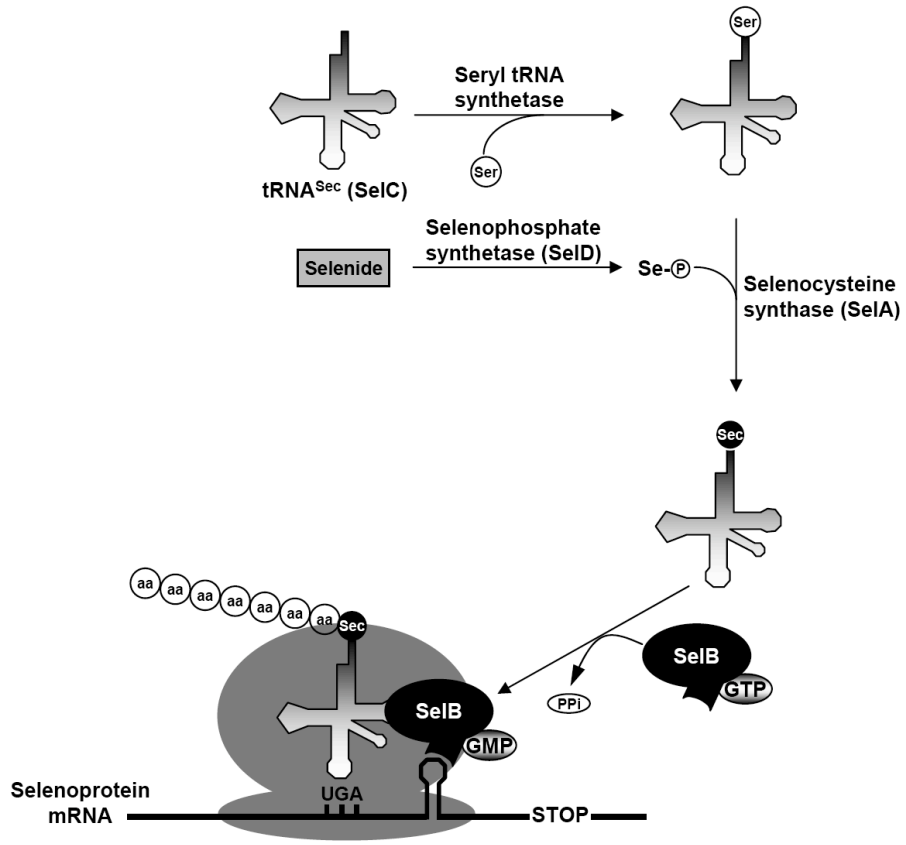


Figure 3. Schematic drawing of the *E. coli* Sec insertion mechanism. See text for details.

Selenium, after its uptake into the bacterial cell, is transformed into selenide by pathways not fully characterized. It is known, though, that the *E. coli* Trx-system is needed to convert selenite to selenide [89, 90]. Selenide is then transformed via the *SelD* gene product (selenophosphate synthetase) in an ATP dependent reaction, to selenophosphate [96, 98, 99], see **Reaction no.1**.

Reaction no.1: $ATP + HSe^- + H_2O \rightarrow H_3SePO_3 + P_i + AMP$

It has been shown that Cys17 of the 37 kDa *E. coli* SelD protein is essential for this reaction [97]. Selenium can also derive from the action of one of the *E. coli* Sec-lyases; IscS, CSD or CsdB, liberating selenium from Sec [100-102], see **Reaction no. 2**.

Reaction no.2: $L\text{-Sec} \rightarrow L\text{-alanine} + Se^0$

The selenophosphate is charged to the Sec-specific tRNA ($tRNA^{Sec}$), encoded by the *selC* gene. The structure of $tRNA^{Sec}$ has some specific characteristics as compared to those of standard tRNA. The *E. coli* $tRNA^{Sec}$ is shown in **Figure 4**. $tRNA^{Sec}$ has a slightly longer acceptor stem with one extra base pair and a long variable arm, and several base pairs that are conserved in other tRNA have here been changed [103-105].

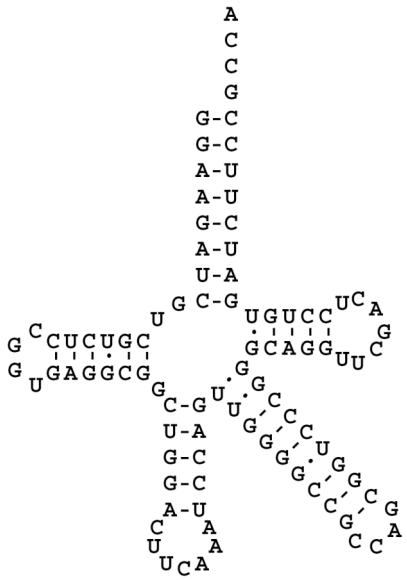


Figure 4. *E. coli* $tRNA^{Sec}$. The *E. coli* $tRNA^{Sec}$ has a slightly longer acceptor stem with one extra base pair and a long variable arm.

$tRNA^{Sec}$ is first aminoacylated (charged) with L-serine by seryl-tRNA synthetase (SerS) [95] and thereafter, via the action of the *SelA* gene product (selenocysteine synthase), converted into selenocysteyl-$tRNA^{Sec}$ [92]. Selenocysteine synthase is a pyridoxal phosphate dependent enzyme that esterifies the serine residue of seryl-$tRNA^{Sec}$, which then forms an intermediate state and the selenol takes the place on the amino acid, forming Sec-charged $tRNA^{Sec}$. Neither the $tRNA^{Sec}$ charged with Sec or the precursor $tRNA^{Sec}$ charged with seryl are, due to the structural differences, recognized by the *E. coli* elongation factor (EF-Tu) and are therefore not utilized in unspecific incorporation into growing polypeptides chains at the position of UGA codons (to which the $tRNA^{Sec}$ has an anticodon). However, the aminoacylated $tRNA^{Sec}$ binds, after the Se has been charged to form Sec, to a Sec-specific elongation factor (SelB) encoded by the *selB* gene. The SelB protein guides several important

steps in the process of Sec insertion into protein. The N-terminal part of SelB (SelB-N) shows similarities with EF-Tu and binds GTP and $tRNA^{Sec}$ charged with Sec. The C-terminal part of SelB (SelB-C) binds to the SECIS element (the bacterial SECIS is discussed in further detail below). The SECIS interaction domain of SelB is an extension that is a unique structural feature for SelB in bacteria; archaeal and eukaryotic Sec-specific elongation factors (eEFSec and aSelB) do not interact directly with their respective SECIS elements. The need of a SECIS element in the bacterial mRNA and its interaction with SelB also prevents unspecific incorporation of Sec at non-Sec-encoding UGA codons [91]. The SelB also binds to the ribosomal 30S subunit. When these connections have occurred, the $tRNA^{Sec}$ can deliver the Sec only at the predefined UGA in the A-site of the ribosome [106], via the action of SelB [107]. A structural overview of the proposed SelB/Ribosome/$tRNA^{Sec}$/SECIS complex is presented in **Figure 5**.

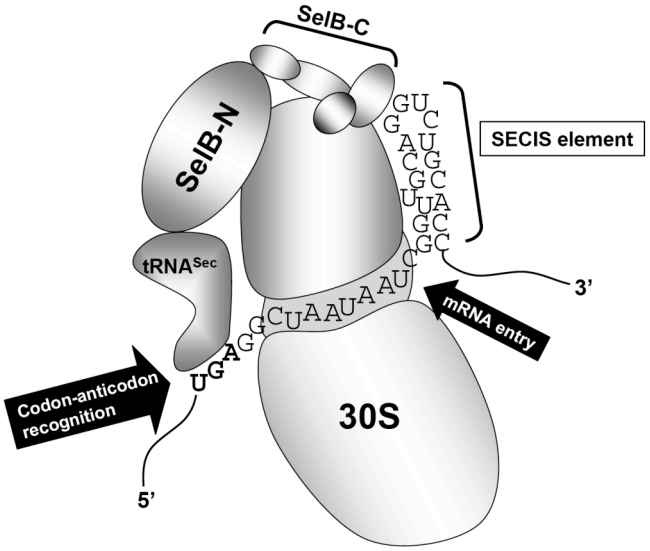


Figure 5. Sec insertion complex in *E. coli*. Insertion of Sec occurs when the mRNA UGA codon is in ribosomal A site. The SelB binds to both the ribosome, SECIS and $tRNA^{Sec}$ charged with a Sec. Interaction of SECIS with the ribosome occurs mainly with the phosphate backbone of the RNA chains of the SECIS. Adapted from reference [106].

The *sel*-genes are generally expressed constitutively [55]. The *selC* gene is monocistronic, whereas the other *sel* genes are expressed in operons. *selA* and *selB* are expressed as one transcriptional unit from the same operon, the *selAB* operon, from a weak promoter. The transcript from the *selAB* operon is additionally regulated by a

SECIS-like structure in the 5´-untranslated region of the operon that binds free SelB thus forming a SelB-mRNA complex that represses the translation of SelB [108]. This suggests that free SelB molecules regulate the expression of SelB so that only the amount of SelB that is needed in a bacterial cell is synthesized. *selD* is part of an operon containing two other genes. The stoichiometry of the different *sel* gene products (SelB and SelC) has been shown to be an important factor for the efficiency of the UGA decoding as Sec in the bacterial selenoprotein synthesis [109]. Overexpression of SelB gives lower UGA read-through; this effect is however reversed when $tRNA^{Sec}$ is also overexpressed. SelB is thus believed to form inactive complexes in the absence of its other binding partners.

1.4.4 The *E. coli* SECIS element

The *E. coli* SECIS element (see **Figure 6**) binds SelB and guides the Sec insertion to the correct predefined UGA codon. Studies have been performed, showing the minimal requirement for the natural SECIS element to be functional [110]. The natural distance of 11 nt from the UGA to the start of the stem-loop is of great importance; if this distance is changed by a single base pair to 10 or 12 nt, the Sec insertion rate is markedly decreased. The specific nucleotide bases of the SECIS binding to SelB were found to be mainly G23, U24 and U17/U18 [111] (Sec encoding UGA is no. 1, 2 and 3). A crystal structure of the SelB/SECIS complex from *Moorella thermoacetica* shows that the SelB protein interacts specifically with G23 and U24 of the SECIS, whereas the other bases involved in the connection to SelB bind with the phosphate backbone of the mRNA [106]. If G23 is changed to any other nucleotide, the SelB binding to SECIS will not take place [112, 113].

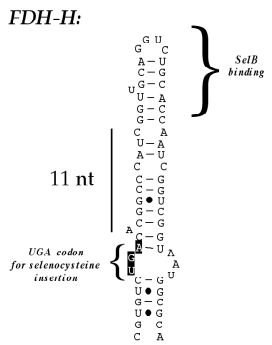


Figure 6. SECIS of *fdhF* from *E. coli*. The native SECIS from the *fdhF* is shown containing the Sec encoding. Adapted from reference [114].

1.4.5 Release factor 2 – opposing Sec insertion in *E. coli*

Release factor 2 (RF2) is an essential protein in *E. coli* [115] and is one of two proteins catalysing translational termination at the mRNA. This occurs at the ochre (UAG) or amber (UAA) codons via release factor 1 (RF1) or at the opal (UGA) or amber (UAA) codons via RF2 [116]. Thus, it is only RF2 that catalyses translational termination at UGA codons in *E. coli*. Translational stop in eukaryotes and archaea is performed by only one release factor that recognize all three stop codons [117-119]. The release factor proteins performing the actual translational termination at the stop codon (RF1 and RF2 in *E. coli* and eRF in eukaryotes) are called class 1 release factors. In addition to the class 1 release factors there is a second class of release factors, class 2 RFs, e.g. RF3 in prokaryotic and eRF3 in eukaryotic organisms, having the function of increasing the recycling of the class I release factors by assisting the removal of class 1 RFs from the ribosome [120, 121]. The eukaryotic and prokaryotic class 1 release factors differ in their amino acid sequences and also by the fact that the eukaryotic class 1 release factor can form complexes with the class 2 release factor in the absence of ribosome [122-124], while prokaryotic RF1 and RF2 lack this ability [121, 125].

A highly conserved sequence in the prokaryotic and eukaryotic class 1 release factors is a GGQ motif [126]. The CGQ motif is likely to be involved in the peptidyl-tRNA hydrolysis that eventually ends the translation, since mutation in that region disturbs the activity of the class 1 release factors [126].

E. coli RF1 and RF2 share similar structures and mechanisms for releasing the polypeptide chain [127, 128]. A second conserved sequence in these proteins is the domain that determines the identity of the release factor; Pro-Xxx-Thr (PXT) in RF1, and Ser-Pro-Phe (SPF) in RF2. Their general similarities in structure and activity have been studied by mutations of these particular conserved regions, resulting in switched function for RF1 and RF2 [129]. Also, mutants of RF2 have been produced resulting in an RF2 function to encode translational stop at all three normal stop codons [130].

1.4.6 RF 2 regulation

RF2 has some unique characteristics compared to RF1. RF2 expression is, in *E. coli,* regulated by an internal +1 frameshift at codon 26 [131, 132]. Interestingly,

this frameshift site involves a UGA codon, recognized by RF2 itself. The autoregulatory frameshift directs higher rate of RF2 truncation at codon 26 when levels of RF2 are high and when the levels of RF2 are critically low the ribosome undergoes a +1 frame shift on the mRNA, whereby an Asp residue is incorporated instead of the UGA truncation in the making of the RF2 protein (See **Figure 7**).

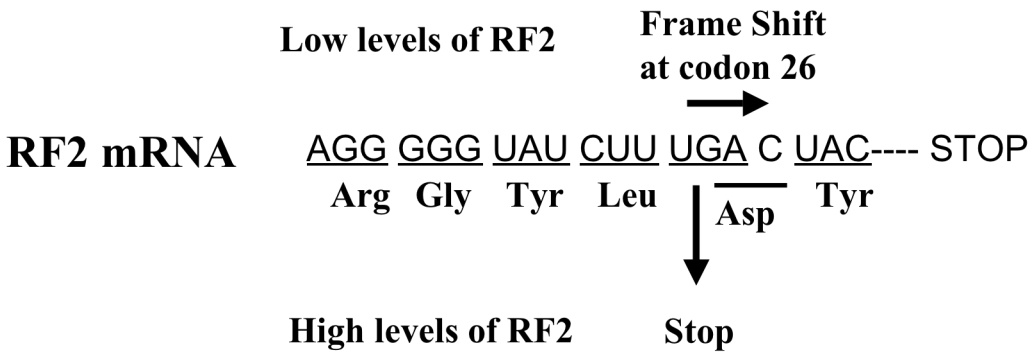


Figure 7. Frameshift sequence of RF2 mRNA. Frameshift occurs at the inframe codon 26 if limiting RF2 is present. If excess amount of RF2 is present truncation at codon 26 will occur.

Frameshifting during ribosomal translation of a normal mRNA is unlikely to occur, and the frequency is around once every 10^4 – 3 x 10^5 codon [133]. However, the efficiency of the programmed frameshift at codon 26 in RF2 varies and can be up to 50% in natural conditions [132, 134]. The occurence of frameshifning is not only regulated by the levels of RF2 present in the cell, but also by the codon context [135]. Codons 22 and 23 are Shine-Dalgarno (SD) like sequences (AGG-GGG). If the SD sequence is altered, the frameshifting occurs less frequent and codon 24-26 is referred to as a slippery site (UAU-CUU-UGA) that further enables the frame shift [131, 136-138].

The frame-shift event during RF2 translation occurs while the in-frame UGA codon enters the ribosomal A site, awaiting binding from RF2. If a non-functioning RF2 would occupy this site the frameshift increases to 100% [138]. Expressing the RF2 under a strong promoter gave a 90% frame-shifting [139]. This could be due to the fact that when the mRNA exists in higher concentration the

stoichiometry between RF2 and UGA becomes altered [140]. Overexpression of RF2 increases the levels 5-fold using a recombinant plasmid that carriesd the RF2 gene, [141].

The *prfB* gene is located in the same operon as the *lysS* gene encoding for lysyl-tRNA synthetase [142]. *lysS* is expressed constitutively to yield lysyl-tRNA synthetase, which charges lysyl tRNA with lysine [142]. An additional separately present gene on the *E. coli* chromosome, *lysU*, which is induced on demand under certain conditions, also encodes lysyl-tRNA synthetase [143, 144].

1.4.7 RF2 and Sec competition

It is of major interest how UGA codons can encode for both Sec and translational termination in the same organism and even in the same mRNA. A translational stop always occurs when a stop codon is in the A-site of the 30S ribosomal subunit, the same position as where the $tRNA^{Sec}$/SelB complex incorporates the Sec at a Sec-encoding UGA. The specific factors, such as the *sel*-gene products and SECIS, make Sec incorporation possible as described above. However, the activities of RF2 and Sec insertion, even at a predefined Sec-encoding UGA, will be in competition with each other [145]. Several components affect this competition and are crucial for the results of this thesis. Some major parameters of this RF2 vs SelB/$tRNA^{Sec}$ competition are given here:

- As discussed in the **Sec insertion and selenoprotein synthesis in bacteria** part (above), different SECIS elements provide different read-through activity, as a direct effect by the structure of the SECIS element and its affinity with SelB. A SECIS element that is less suitable for Sec insertion will allow RF2 to reach the UGA codon and thereby truncation will occur at higher rate [110].
- It is likely that the SelB-$tRNA^{Sec}$ forms a complex with the SECIS before the ribosome is present [109] and this probably makes it harder for the RF2 to compete with the binding to the UGA. Overexpression of SelB as the only overexpressed *sel*-gene will however give a lower UGA read-through because uncharged (or empty) SelB that binds to the SECIS fails to support the

incorporation of Sec and thereby RF2 has a possibility to bind and enable truncation.

- The kinetic parameters for binding the ribosome should also play a role. Lowering the affinity of either the SelB complex or RF2 to the ribosome should favor the other player in the competition between these two factors.
- Codon context of the UGA makes a difference for UGA read-through [139, 146]. UGA encodes translational termination with less efficiency if the following +4 position is a C. The natural sequences surrounding Sec-encoding UGA codons are usually of the character of bad translational terminator codons, e.g. the Sec-encoding UGA upstream sequence for bacterial formate dehydrogenases (*fdnG*, *fdhF* and *fdoG*) are Arg-Val encoded by CGT GTC and for rat TrxR GGA TGC (UGA) GGC UAA). It also seems like the nature of the penultimate amino acid influences UGA read-through [147-149].

All of these parameters affecting SelB and RF2 competition, as listed above, underlie the results of the work on optimized selenoprotein production, as presented in this thesis.

1.4.8 Non-Sec mediated UGA suppression in *E. coli*

Noren *et al* showed that the efficiency of Sec insertion is influenced by both the levels of selenium in the medium and of nucleotides directly downstream of the Sec-encoding UGA codon [150]. If no SECIS is present downstream of the UGA codon, either translational termination or Trp insertion may happen. If a SECIS is present, the rate of truncation or Trp incorporation is highly reduced and if a T or C follows the Sec-encoding UGA, Trp-mediated suppression is further reduced [150]. Of major importance for this thesis, however, is the fact that if both sufficient levels of selenium are present and a functional SECIS element is found in the mRNA, then no Trp-mediated suppression occurs, irrespective of codon context [150].

1.5 Rat thioredoxin reductase – the model selenoprotein produced and studied in this thesis

1.5.1 General properties of thioredoxin reductases

Thioredoxin reductase (TrxR) forms, together with thioredoxin (Trx) and NADPH, the thioredoxin system that serves an important role in the defense system against oxidative stress in most organisms of all domains of life and supports many redox-related cellular pathways [151]. The importance of the Trx-system has been studied in detail and is reviewed elsewhere [151-153].

TrxR proteins are divided into two groups; small-type TrxR and large-type TrxR, with their special characteristics presented in **Table 6**.

Table 6. Charecteristics of large- and small-type TrxR

	Large	**Small**
Organisms	Mammals and several other eukaryotic organisms	Archaea, bacteria, plant
Active sites	C-terminal -CU- or -CC- N-terminal CxxxxC	-CXXC-
Cofactors	FAD (enzyme-bound) NADPH	FAD (enzyme-bound) NADPH
Substrate specificity	Wide	Narrow
Subunit association	Dimer	Dimer
Size (as dimer)	≈110 kDa	≈70 kDa

The structure of large-type TrxR is more similar to glutathione reductase than that of small-type TrxR, while the Trx structure in all organisms consists of a central core of five B strands with four A helix surrounding them [154-156]. Both glutathione reductase and mammalian large-type TrxR have a -CVNVGC- redox active motif in the N-terminal domain. However, TrxR has also an additional C-terminal active site motif, either containing a dithiol (-CCX) [157] or a Sec residue as the second last amino acid in a –CUX motif that is essential for activity [154, 155, 158, 159].

More details of the features of rat TrxR1 are presented here, since this protein has been used as a model selenoprotein in this thesis. Sec, in rat TrxR1, is essential for its normal Trx-reducing activity and is located at the penultimate amino acid in the polypeptide chain [160]. The C-terminal end can be seen as an arm with high flexibility that probably is a prerequisite for the wide substrate specificity of the enzyme [155]. The C-terminal end is very reactive and the Sec-containing protein has been hard to crystallize for structural experiments (however, recently this was achieved by Qing Cheng and coworkers, showing a structure that essentially supports earlier models of the catalysis of the enzyme). A schematic drawing of the structure is show in **Figure 9**.

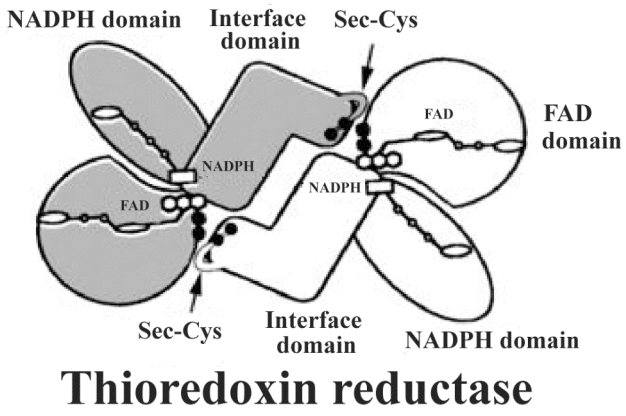


Figure 9. Schematic structure of dimer TrxR. The dimer consists of two subunits arraged in a head to tail conformation. The NADPH, FAD and interface domain are show in the figure aswell as the redox active flexible C-terminal Cys-Sec motif. Adapted from reference **[157]**.

TrxR is an NADPH dependent reductase and receives an electron that is transferred from NADPH to FAD of one subunit in the enzyme. Thereafter, the electrons are transferred to the redox active N-terminal site of the same subunit as that of binding the reduced FAD. This second redox active site is formed by two Cys residues located at positions 59 and 64 within a -CVNVGC- sequence. The electrons are then moved to the flexible C-terminus of the other subunit, to form a reduced selenolthiol from an oxidized selenenylsulfide found in a -GCUG- motif. The redox active selenolthiol can then reduce substrates of the enzyme, such as the active site disulfide of oxidized Trx.

1.5.2 Preparation of TrxR from rat liver

The native form of rat TrxR1 (produced recombinantly in this thesis) has been purified earlier from rat tissue. Obtaining the native selenoprotein TrxR has been an arduous task that was costly and time consuming. Mikaela Luthman and Arne Holmgren developed a preparation method in 1982 [161]. They started with 200 g rat liver and in crude protein lysates obtained from this amount of tissue they detected 305 units TrxR. After a purification process using ammonium sulfate precipitation, DEAE column, 2´,5´ADP Sepharose and (w-aminohexyl)agarose, 1.6 mg pure TrxR with a specific activity of 35 U/mg and a total of 55 U was collected. The features of the recombinant rat TrxR1 produced here will be discussed in comparison to the purified native rat TrxR towards the end of this thesis.

2 Present Investigation

2.1 Intentions

The overall aim of this study was to further characterize and develop the bacterial recombinant selenoprotein production system and to understand its limitations. To accomplish this we optimized growth conditions to achieve higher yield of recombinant selenoproteins and genetically created a bacterial strain that may function as an optimized host for selenoprotein production. Furthermore, we have wished to characterize the recombinant selenoprotein rat TrxR1 in more detail.

Specific aims, results and conclusions for each paper are, as an addition to what is discussed in the individual papers, presented below. The papers on which this thesis is based are found in their whole in the final section of this thesis.

2.2 Comments on the Results and Methods of the papers this thesis is based upon

2.3 Paper I.

Olle Rengby, Linda Johansson, Lars A. Carlson, Elena Serini, Alexios Vlamis-Gardikas, Per Kårsnäs and Elias S. J. Arnér. Assessement of production conditions for efficient use of *Escherichia coli* in high-yield heterologous recombinant selenoprotein synthesis. *Applied and Environmental Microbiology*, 2004 Sep:70(9):5159-67.

For **Paper I** we set out to increase the yield of the recombinant rat selenoprotein TrxR in *E. coli*, asking a number of general questions regarding this system:

- *Could titration of the selenoprotein mRNA levels result in increased selenoprotein yield?*

- *What would be a good way for producing a system capable of such titration?*
- *Can we lower the RF2 levels by a slower growth rate and thereby yield higher Sec insertion into the recombinant selenoprotein?*
- *What is the impact on the selenoprotein yield, when using co-expression of the sel-genes from the pSUABC plasmid under different growth conditions?*
- *Could diverging from standard protocols for recombinant protein production increase the yield of recombinant selenoproteins?*
- *Can multi-dimensional analysis be applied to determine optimal growth conditions for recombinant selenoprotein production?*

These questions will be answered in the conclusions of this thesis, but first some aspects of the methods and results of **Paper I** will be discussed.

2.3.1 Estimation of the E. coli Sec incorporation efficiency

In **Paper I** (and in **Paper II**), we used the specific activity of the produced recombinant rat TrxR as an indication of its Sec content. Since native rat TrxR had been purified from rat liver and shown to have a specific activity of 35 U/mg (see above), we assumed that this activity would correspond to a full Sec content, i.e. one Se atom per subunit. The specific activity as given in percent was thereby determined for different preparations of recombinant TrxR, using the enzymatic activity in U/mg from the standard DTNB assay [162] and setting 35 U/mg (the activity of native enzyme) as 100%. This correlation between specific activity and selenium content was later further studied in **Paper III** (see below).

2.3.2 Selenoprotein mRNA titration and the araBAD promoter

Since the stoichiometry of *sel*-genes is important for the UGA read-through as a Sec-encoding codon (see **Introduction**), we decided to analyze if changing the stoichiometry between the *sel*-genes and the selenoprotein mRNA could favor Sec incorporation. To do this, we needed a promoter capable of titrating the levels of selenoprotein mRNA expression. We chose the pBAD vector, based upon promoters from the *E. coli* araBAD operon. Since use of this system is crucial for major parts of this thesis, some important notes about the araBAD system will be discussed here.

The pBAD vector contains the P_{BAD} promoter from the araBAD operon as constructed in 1995 by Beckwith and co-workers [163]. Their intention was to design a promoter satisfying two conditions. First, the promoter should respond fast and efficient to a repressor. Second, it should be titrable within a wide range of inducer concentration. The P_{BAD} promoter is driven by arabinose that induces a conformational change in the AraC protein. AraC stimulates initiation of mRNA synthesis from the P_{BAD} promoter [164] and is a dimer containing an N-terminal dimerization domain that can bind arabinose and a C-terminal flexible DNA binding domain (**Figure 10.2**).

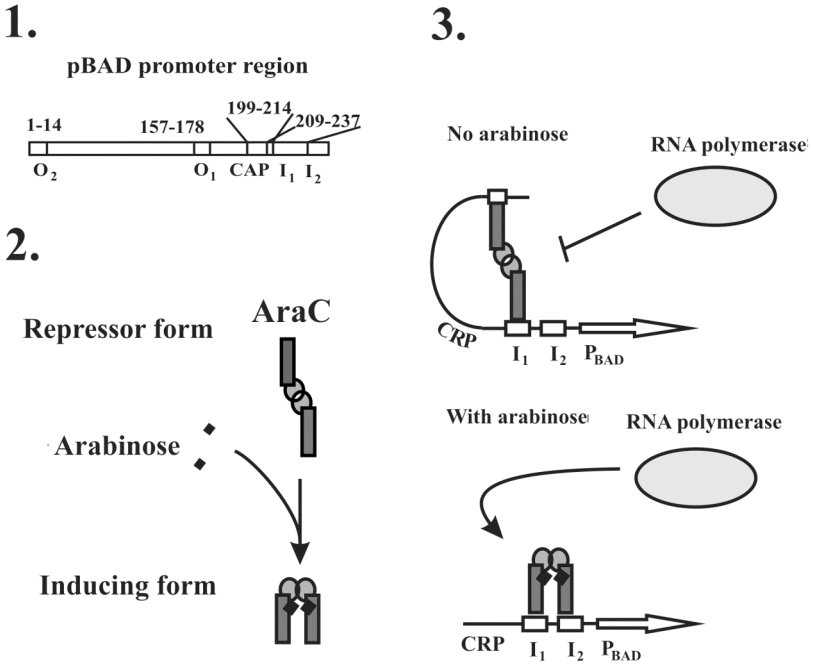


Figure 10. The pBAD promoter and regulation with AraC. 1) The araBAD promoter region from the pBAD vector, I_1 and I_2 are binding sites for araC inducing start of transcription while O_2 and I_1 are binding sites of araC for inhibition of transcription. cAMP activator protein (CRP) binding to CAP site stimulates araC to bind I_1 and I_2. **2**) AraC is a homodimer protein with arabinose binding capacity **3**) In the absence of arabinose the araC binds O_2 site and I_1 site and hiders thereby RNA polymerase. In the presents of arabinose the araC stimulates RNA polymerase for start of transcription.

In the absence of arabinose, AraC hinders the RNA polymerase to start transcription from the P_{BAD} promoter by binding the so called "O_2-site" (**Figure 10.1** and **10.3**). This results in very low transcriptional levels. When glucose is present, the transcription levels are depressed even further, down to virtually no transcription at all. Glucose lowers the levels of cyclic AMP (cAMP) that otherwise binds to the CAP site in the araBAD promoter region. The CAP site functions by stimulating AraC binding to the so called "I_1" and the "I_2"-sites and thereby increases promoter expression (**Figure 10.3**).

The P_{BAD} promoter can be regulated with a wide concentration range of arabinose as inducer. However, the expression from the promoter is dependent on uptake of arabinose from medium through the arabinose transporter protein AraE and AraE needs arabinose for induction of its own expression. Such a system is commonly referred to as an autocatalytical gene expression system [165]. The effect of arabinose titration on expression levels in such system, reflects a titrated regulation on the whole bacterial culture level, but not an actual titration of mRNA within each individual cell, which will be in a stocastic on-or-off regulation [166] (see **Figure 11**). For that reason we used two bacterial host strains, TOP10, recommended by the pBAD vector manufacturer, and BW27783, a bacterial host that was made to express AraE constitutively and thereby enable mRNA titration on the individual cell level [167].

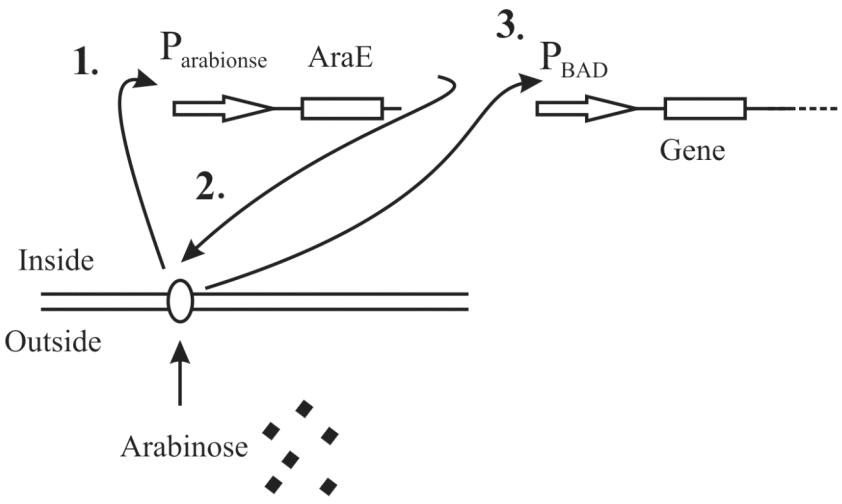


Figure 11. Flow-scheme of araE function and regulation. The arabinose transporter protein, AraE, suffers from an on-or-off regulation in normal conditions schematically described in this figure. **1**) The promoter of araE is arabinose dependent; arabinose cannot enter the cell without AraE. **2**) Once araE has been expressed arabinose can enter the cell and **3**) start transcription from the plasmid-born or chromosomal arabinose-dependent promoter such as P_{BAD} of the pBAD plasmid as well as **1**) inducing more AraE.

We placed the rat TrxR in the pBAD vector under the control of the P_{BAD} promoter. The gene for rat TrxR, carrying a variant of the bacterial SECIS element that enables TrxR expression in *E. coli* (as previously described for the pET vector, see section **Recombinant expression of selenoprotein in *E. coli*** below and [114]), was here expressed from a new plasmid called pBAD-TRS_{TER}. Titration experiments were performed using the two *E. coli* strains (TOP10 and BW27783) with the pBAD-TRS_{TER} plasmid, with or without co-transformation of the pSUABC plasmid that carries additional copies of the *sel* genes. Titration of mRNA levels was performed over a wide range of the inducer arabinose (0.0001%-1%). We then measured mammalian TrxR activity in the crude bacterial lysates using NADPH-dependent DTNB reduction, which was an efficient way of assessing Sec-containing TrxR levels since no endogenous bacterial enzyme reduces DTNB with NADPH at any appreciable level as compared to the overexpressed rat TrxR. For both strains, the mammalian TrxR activity reached a plateau, at an arabinose concentration of about 0.05% for the TOP10 culture and at 0.005% for the BW27783 culture. This is an interesting difference, since it is probably explained by the fact that the arabinose transporter protein araE is under arabinose control in the TOP10 but not in the BW27783 (where it is constitutively expressed). We also found that, in most conditions, the use of pSUABC was of great importance and when it was used, the yield was about doubled as compared to when it was not used.

2.3.3 Culture conditions

Previously published work showed that RF2-mediated translational termination at a predefined Sec-encoding UGA codon is in competition with $tRNA^{Sec}$ for decoding of the UGA as Sec [145]. We were therefore interested to determine the impact of RF2 on the selenoprotein yield, using our overproducing system.

It has also been shown that lowered growth rate in a culture decreases the levels of RF2 [168]. Could reduced growth rate increase the read-through of Sec-encoding UGA as Sec, compared to translational termination? We here wished to determine the effect of lower growth rate, and thereby lower RF2 levels, on the yield of Sec-containing recombinant rat TrxR. We performed studies on *E. coli* induced for selenoprotein production from the titrable rat TrxR1 overexpression plasmid (pBAD-

TRS_{TER}), where the growth rate was lowered by using minimal medium instead of previously used rich LB medium. The yield obtained thereby was however not favored and the yield of recombinant selenoprotein expressed under those conditions did not reach the levels that had been previously achieved [114]. Adamski showed that the number of RF2 molecules increased from 5900 to 24900 when the growth rate increased from 0.3 to 2.4 doublings per hour [168]. The percentage of RF2-containing ribosomes however increased only from 21% to 33% since the total number of ribosomes also increased [168]. In relation to our studies, it should likely be many components involved, resulting in lower Sec incorporation when selenoprotein production is induced in a culture grown in minimal medium, than just the number of RF2 molecules or the percentage of RF2-containing ribosomes. However, during production of selenoprotein in full medium, we found that the higher yield was obtained using high selenoprotein mRNA levels induced at late exponential phase/early stationary phase with a culture grown at room temperature. In this case, it should be reasonable to believe that lower RF2 levels in relation to SelB/$tRNA^{Sec}$ and selenoprotein mRNA, should be one of the factors explaining the higher yield of selenoprotein that was obtained.

2.3.4 The pET vector

The pET vector system uses the T7 phage RNA polymerase and a T7-specific promoter. The bacterial host must be a DE3 lysogenic strain to enable functional expression with pET vectors. A DE3 lysogenic strain contains the lac promoter for T7 RNA polymerase, which has been incorporated into the *E. coli* genome by phage recombination techniques. The T7 polymerase is functionally exclusively for T7 promoters and will not recognize any other promoter in the bacterial cell. The inducer, IPTG, functions by displacing the repressor of the lac operator and transcription can thereby initiate upon production of T7 polymerase. Both the promoter region of the pET vector and the T7 polymerase promoter region are induced by IPTG, see **Figure 12**.

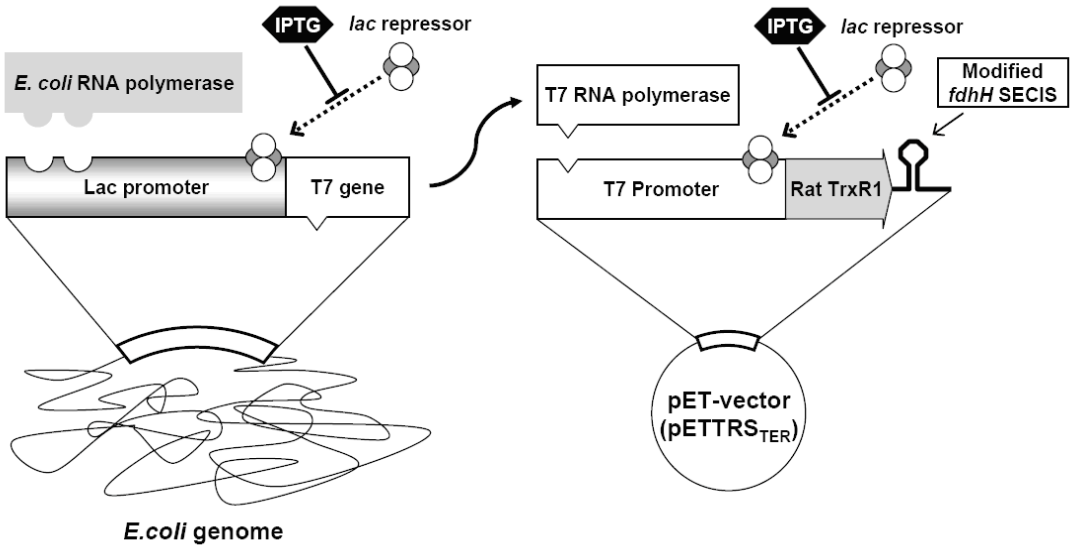


Figure 12. Overview of the pET vector system and pET-TRS$_{TER}$. The bacteria must carry a DE3 cassette encoding the IPTG induced T7 RNA polymerase. Once IPTG is present and T7 RNA polymerase has been produced it will function exclusively on the T7 promoter to produce copies of the gene of interest, in our case, rat TrxR.

2.3.5 Principal Component Analysis

How could the recombinant selenoprotein yield be improved in a system using overexpression of the selenoprotein mRNA levels, via a T7 promoter-driven transcription, together with overexpression of the *sel*-genes via the pSUABC plasmid? We decided to determine the impact of three different parameters on the selenoprotein yield, which we believed would have an impact on the selenoprotein production. We thus used principal component analysis (PCA) as a statistical tool for optimizing the following three parameters: temperature, induction at different growth phases and duration of induction. The optimization was preformed by randomly varying these production parameters, within a range that was considered to be reasonable. These parameters were investigated, using small-scale cultures that were induced for protein production at certain selected time points. Samples were then analyzed for protein amount and activity as well as cell density. The obtained data was then processed using PCA. The results of the PCA indicated co-variation between total amount of selenoprotein produced and selenoprotein produced per OD_{600}, suggesting that the total yield was not only highest at a certain condition because of more bacterial cells were present, but that each cell under those conditions also contained more selenoprotein.

The analysis showed that selenoprotein expression should, to attain optimal yield, start at a high OD_{600} (e.g. in the late exponential growth phase) and that expression should occur at lower temperature. The total duration of the incubation was, however, the parameter with the least impact on the yield. The growth parameters given from the PCA analysis were then tested in larger cultures induced with different time points of induction and different duration. The analysis finally resulted in a protocol where induction should be at late exponential phase ($OD_{600}\approx2.4$) with duration for about 24 hours at 24^{o}C, the so-called "2.4/24/24-protocol". We have subsequently found that this protocol gives good selenoprotein yield for a number of different recombinant selenoproteins, demonstrating the usefulness of the initial PCA analysis.

2.3.6 Unpublished control of the PCA

As a separate control to the published PCA prediction we also used a rule discovery system (RDS) to investigate if also such computer analysis would give the same prediction as the PCA. The data from **Table 3** in **Paper I** was analyzed and the resulting recommendation indeed supported the published predictions, but with some slight differences. Induction should, according to the RDS, occur at an $OD_{600}>1.75$ and incubation temperature should be lower than 28.5^{o}C. The RDS was not able to make a good prediction for the duration parameter, due to insufficient information input. However, this also correlated with the PCA since the duration had the lowest impact in that analysis. It should also be noted that our first interpretation of the PCA was to induce the culture at stationary phase, which gave very low yield, and this would probably not have been done if an RDS had also been performed at the same time.

2.3.7 Recombinant TrxR stability

Earlier unpublished data had suggested that recombinantly produced TrxR, induced in early exponential phase and at 37^{o}C, could loose activity if handled extensively in room temperature or upon repeated freezing and thawing. Having highly pure and larger amounts of the enzyme, we therefore wished to repeat the analysis of its stability. Two samples of recombinant rat TrxR were produced according to the "24/2.4/24-protocol" and purified over the 2´,5´ADP sepharose, whereupon they were

stored in the high-salt elution buffer (50 mM Tris, 2 mM EDTA, 500 mM NaCl) from the 2´5´ADP sepharose and treated in two different ways. One sample was stored at room temperature and activity measurement at certain time points were performed during several days of storage. The other sample was freezed-and-thawed several times with activity measurements in-between. The protein activity was only slightly affected by either of these treatments, showing that the selenoprotein produced was rather stable (**Figure 13**).

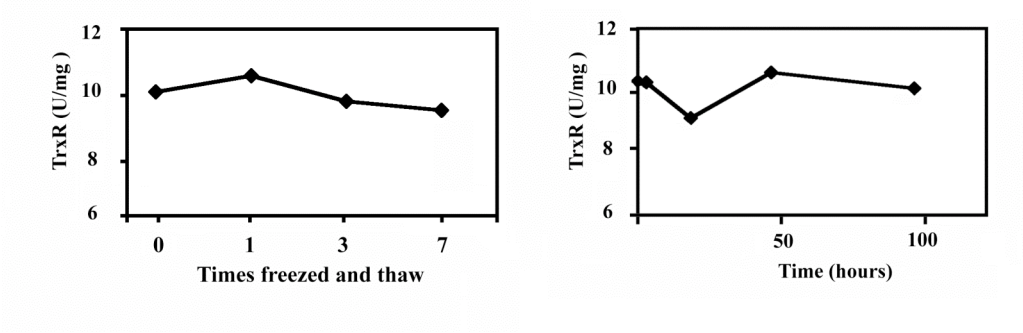


Figure 13. Recombinant TrxR stability. The recombinant rat TrxR1 was freezed-and-thawed or incubated at room temperature for a long period, with activity measurements using the DTNB assay as shown in the figure.

2.4 Paper II.

Olle Rengby and Elias S. J. Arnér. Titration and conditional knockdown of the *prfB* gene in *Escherichia coli*: Effects on growth and overproduction of a recombinant selenoprotein. *Applied and Environmental Microbiology*, 2007 Jan:73(2):432-41

In this study we wished to create an *E. coli* strain specialized for selenoprotein production, trying to answer a number of questions:

- *Could an E. coli strain specialized for selenoprotein production be made by knocking down the levels of RF2?*
- *Which existing technique would be suitable for replacing a native chromosomal promoter with a synthetic titrable promoter?*
- *Which would be the appropriate promoter to use for titration and knock down of the RF2 levels?*

- *How does overexpression of RF2 effect growth of the bacteria?*
- *How would knock down of the RF2 levels affect growth?*
- *What growth technique would allow us to knockdown RF2 simultaneously as recombinant selenoprotein is produced?*
- *How does knock-down of RF2 protein affect the recombinant yield and the specific activity of the selenoprotein?*

These questions will be answered in the conclusions of this thesis, but first some aspects of the methods and results of **Paper II** will be discussed.

2.4.1 Chromosomal exchange of the *prfB* promoter

RF2 is in competition with $tRNA^{Sec}$ for binding of the Sec-encoding UGA during selenoprotein mRNA translation in *E. coli* (see **Introduction** and [145]). Again we ask us, what effect could lowering the levels of RF2 have on the recombinant selenoprotein production yield and how could we deliberately achieve a specific RF2 knockdown? Due to the fact that RF2 is essential for *E. coli* growth [115] we could not simply knock out the *prfB* gene encoding for RF2 in order to achieve an RF2-free cell, which otherwise would likely be ideal for the purpose of selenoprotein production. At the time we raised these questions, there were no established methods published for knocking down specific genes in *E. coli* without the use of additional plasmids. We realized that we needed to replace the chromosomal promoter of *prfB* gene with a promoter capable of tight regulation at low levels of expression and we decided to attempt a recombination methodology previously successfully used for making null *E. coli* cells of non-essential genes.

The λ Red recombinase technique [169] was developed for knocking out specific genes using a PCR product containing one FRT (FLP recognition target) on each side of an antibiotic resistance gene (used for selection) and 36-50 nt extensions homologous to the regions adjacent to the gene of interest to be inactivated. The presence of exonucleases usually prevents the transformation of linear DNA in *E. coli*, but linear DNA transformation is possible upon expression of the gamma, λ, *exo* genes, thereby enabling use of λ Red recombination. Could the λ Red recombinase technique be used for not only knocking out, but also replacing genes (in our case the

promoter of a gene)? Indeed, we successfully used this technique with a PCR fragment containing not only an extension homologous to the sequence just upstream of the *prfB* gene and the FRT sites on one side of an antibiotic resistance gene, but also as the second homologous part, the first part of the *prfB* gene and an in-frame sequence of the P_{BAD} promoter. Once recombination occurred upon this approach, the native *prfB* gene became under P_{BAD} control and thereby RF2 levels were possible to be titrated using arabinose as an inducer and glucose as a repressor. The resulting host stain, after removal of antibiotic selection resistance gene and the helper plasmid for recombination, was called ORaa. We also made a DE3 lysogen of this strain in order to enable use of the pET vector systems for recombinant production (see above, under **Paper I**, for a discussion about the pET vector systems).

As also discussed for **Paper I**, arabinose-titrable regulation use of the P_{BAD} promoter requires expression of the arabinose transporter protein araE. We therefore chose the BW27783 strain, derived from the F- λ- *E. coli* BD792 [169, 170] strain that contains araE under a constitutively expressing promoter [167], as the parental strain for our experiment to construct the ORaa strain. This enabled titration of RF2 at the individual cell level (see discussion above, under **Paper I**).

2.4.2 The impact of lowered RF2 levels on selenoprotein production

How could we tightly regulate RF2 and estimate the effect of lowered RF2 levels on the recombinant selenoprotein yield, and the specific activity of Sec insertion as compared to RF2-mediated truncation? Importantly, we chose to perform the RF2 titration in a continuously turbidostatic fermentor system. In such system, the RF2 levels could be knocked down by constant addition of glucose and by continuously collecting samples for selenoprotein induction, without altering any other growth conditions of the culture. We were thus able to study the direct effect of reduced RF2 levels on the yield and specificity of recombinant selenoprotein production. Upon RF2 down-regulation the decoding of UGA as Sec clearly increased, as expected. However, growth of the bacteria and the total yield of TrxR were also lowered, resulting in a final selenoprotein yield that was very low (although with a high Sec insertion to UGA termination ratio). The decrease of RF2 termination was higher than the decrease of

yield in TrxR production, resulting in an expression of recombinant rat TrxR1 with a higher specific activity than previously reached, ending at about 40 U/mg, before the yield became to low to measure.

2.4.3 Continuously turbidostatic fermentor system

In a continuously turbidostatic fermentor, new medium is constantly added to the culture at the same rate as the culture is harvested, with the volume of the culture thereby held constant. A culture grown in a continuously turbidostatic fermentor differs from that of a batch culture in several ways. For batch-wise cultures, the medium composition changes constantly due to components in the medium being constantly consumed, such as antibiotics, nutritients and other supplements, while other components such as secondary metabolites or cell density are increasing. These changes during culturing makes *E. coli* of batch cultures enter new growth phases during cultivation and both the bacterial cell and the culture as a whole must adapt to the constantly changing parameters. Each of the growth phases can have different impacts on a recombinant protein being produced in such culture, as also demonstrated in ours and others work (see **Paper I**, **Paper II** and references [168, 171]). A culture in a continuously turbidostatic fermentor enters a steady state after some time of culturing, which can be sustained if sterile conditions are used, for a very long time. This is both useful for an industrial-scale production of recombinant protein products and for research purposes. In our case, we applied a continuously turbidostatic fermentor in the titration of the *prfB* gene because it allowed us to titrate and follow the levels of RF2 under tightly controlled conditions.

2.4.4 Knock-down of RF2 levels - effect on bacterial growth

As previously shown by others, the cellular level of RF2 protein is essential for bacterial growth [115]. We were therefore interested to study the effects on bacterial viability during the knock down of the RF2 protein in the ORaa strain. We grow the bacteria on plates containing high concentration of glucose, the P_{BAD} repressor, and showed that the bacteria displayed extremely slow growth under these conditions. However, the effect of glucose was rather bacteriostatic than bactericidal,

which was an interesting observation. It could possibly be explained by the fact that expression from the P_{BAD} promoter leaks even when grown in the presents of glucose, but that is less likely since glucose is a strong repressor and we used very high concentrations. Another explanation is that the RF2 present in the bacterial cells when placed onto the plate was sufficient for a continued slow growth, so that the glucose could be consumed giving the bacteria a chance to cultivate albeit with very slow growth. However, when the bacteria grown on glucose were transferred onto new plates containing arabinose, the recovery of growth was surprisingly fast, which strengthened the hypothesis that the effect was indeed bacteriostatic and that glucose had not been consumed in the first plate. The last explanation is thus that in the absence of RF2, *E. coli* has the possibility to enter some type of bacteriostatic "hibernating" state exemplified by a lack of growth, but having the capability to regain normal growth if RF2 levels are again increased.

2.4.5 Overexpression of RF2 - effect on growth

We also became interested in how overexpression of the RF2 protein could affect bacterial growth, and also how high overexpression was possible to reach, using ORaa grown in high arabinose concentrations. Overexpression of RF2 using high concentrations of arabinose gave elevated RF2 levels about 3-7 times higher compared to the host strain. This overexpression resulted in slightly lowered growth rates, but without other apparent phenotypes. It was surprising that the RF2 levels were not higher, considering the very high transcription achieved from the araBAD promoter in the presence of high arabinose. The overexpression of RF2 is probably limited by the internal frame shift located in the *prfB* gene, functioning as an auto-regulatory ability (see **Introduction** and [131]). Another factor could be that the P_{BAD} promoter from the pBAD vector was cloned into the chromosome without co-expression of the gene for the P_{BAD} activator protein AraC. AraC is, however, endogenously upregulated up to 10-fold upon the addition of arabinose [172].

2.5 Paper III

Olle Rengby, Qing Cheng, Marie Vather, Hans Jörnvall and Elias S. J. Arnér. Highly active dimeric and low-activity tetrameric forms of selenium-containing rat thioredoxin reductase. *Submitted Manuscript* (2008).

In this paper we wished to study the recombinant rat TrxR1 in more detail, specifically asking the following questions:

- *What is the molecular mechanism for purification of dimeric rat TrxR1 over a PAO Sepharose column?*
- *What is the explanation of the very high specific activity resulting from a PAO Sepharose purification of TrxR?*
- *In what form are recombinant rat TrxR expressed? Heterodimers or homodimers?*
- *What is the theoretical highest possible specific activity of recombinant rat thioredoxin reductase?*

The answers to these questions will be given in the conclusions of this thesis, but first some specific methods and results of **Paper III** will be discussed.

2.5.1 PAO Sepharose and the Sel-tag

The Phenylarsine oxide (PAO) Sepharose was originally developed as an affinity column for purification of vicinal dithiol-containing proteins [173]. The PAO Sepharose was subsequently found to have an even higher affinity for the C-terminal active site Cys-Sec of rat TrxR1. This property was a base for the development of the Sel-tag concept for recombinant proteins [174-177] and was used for purfication of Sec-containing TrxR in both **Paper I** and **Paper III**. A schematic drawing of how recombinant rat TrxR binds to PAO is shown in **Figure 14** and this purification is further discussed below.

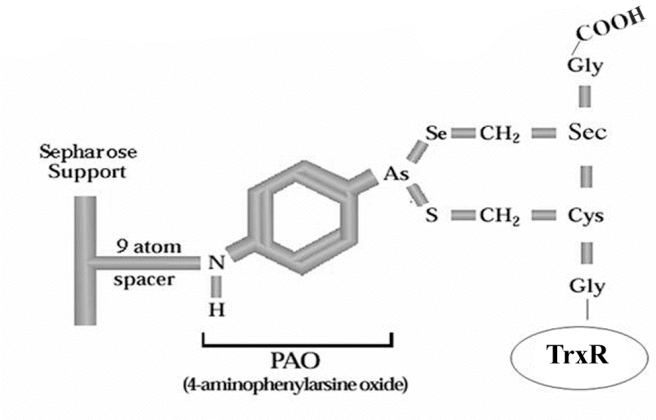


Figure 14. Binding of TrxR to PAO. PAO sepharose binding of rat TrxR. The selenylsulfide form a covalent binding to the arsenic atom of the PAO sepharose. The protein can be eluted by addition of DMPS.

2.5.2 Recombinant rat TrxR1; heterodimer or homodimer

PAO Sepharose is not able to bind directly to a UGA-truncated subunit of TrxR. The interface area of two subunits forming a dimer is approximately 3400 $Å^2$ [155] and should therefore not be easily separated. However, the appearance of 35% truncated subunits in a PAO Sepharose elution fraction of an extensively washed column, suggests that the UGA-truncated TrxR was co-purified and probably part of a heterodimer with a Sec-containing subunit (that binds the PAO Sepharose). This would lead to the conclusion that recombinant TrxR selenoprotein forms both homodimers and heterodimers when expressed, as further discussed in **Paper III**.

2.5.3 Specific activity of recombinant rat TrxR1

TrxR is expressed as a dimer and the electron flow for full enzyme activity is dependent on both subunits being present and acting jointly within one catalytic cycle (see **Introduction**). Recombinant expression of rat TrxR1 in *E. coli* clearly results in a mixture of UGA-truncated and full-length Sec-containing enzyme. The main focus of this thesis was, with significant success, to decrease the rate of UGA truncation (**Paper I** and **Paper II**). In those studies, the activity of the recombinant TrxR has been used as a direct indication of its selenium content. We have also used

PAO Sepharose for enrichment of the Sec-containing TrxR and the purified enzyme usually has had an activity similar to that of native enzyme purified from tissue. We have therefore assumed that this enzyme would have full Sec content. However, we have always asked ourselves the question of the exact molecular mechanisms explaining how the PAO-purification can obtain TrxR with full selenium content, considering that the dimer should likely form as heterodimers between UGA-truncated and Sec-containing subunits. How can those subunits be separated from each other during PAO Sepharose purification, so that the purified enzyme has full activity? To answer this question we purified a very large amount of recombinant rat TrxR and, as discussed in the paper, we outlined two different hypotheses for the molecular behavior of TrxR during purification over PAO. To our surprise we found that neither of our outlined hypotheses were correct. The explanation for the results of the purification was clearly that UGA-truncated TrxR subunits were indeed co-purified over the PAO Sepharose, probably being associated with a Sec-containing subunit, but that the existence of these subunits was masked by the remarkably high activity of the selenium containing enzyme. A batch of a PAO purified enzyme with a very high specific activity of 53 U/mg, surprisingly only had Sec in 65% of its subunits. In light of this finding, earlier conclusions concerning the efficiency of *E. coli* Sec incorporation into recombinant selenoprotein (**Paper I**, **Paper II** and [114]) should be revaluated since the Sec insertion has been less than previously believed (about half). Moreover, the maximal catalytic capacity of recombinant rat TrxR1 has to be revaluated for the fully Sec-containing enzyme, which should have a specific activity around 80 U/mg. This is significantly higher than that determined for any native TrxR purified from mammalian tissues. The possible explanations for lower activities in the native enzyme preparations are discussed in **Paper III**.

2.5.4 Tetrameric and multimeric forms of recombinant rat TrxR

As a first experimental plan to solve the question of how TrxR could be purified to such high specific activity, we performed crosslinking studies in combination with PAO Sepharose purification. The outcomes of those experiments were inconclusive, since neither of the outlined hypothesis turned out to be correct (see discussion above, and in **Paper III**). However, as an unexpected finding from these

experiments, a stable tetrameric form of TrxR1 was identified. Further analyses showed that this form of TrxR1 seemed to be made of subunits apparently associated in a unique and strong manner with each other, but not linked by disulfides (or selenenylsulfides). This tetrameric form of rat TrxR1 ($TrxR1_T$) had TrxR activity, but only approximately half that of the dimer form of the recombinant rat TrxR1 ($TrxR1_D$) when correlated to the selenium content. We also noted formations of multimers, but those were rather provoked by crosslinking or linked by disulfides (or selenenylsulfides) and were not as apparent or stable as the tetrameric TrxR1. The questions are why, where, how and when the tetrameric form of TrxR is formed. This has not yet been answered by us, but some possible functions are discussed in the **General Discussion** part (and in **Paper III**).

3 Conclusions of this thesis

The major conclusions of this thesis are presented below, answering the questions raised in connection with each of the papers forming the basis of the thesis:

3.1 Paper I – Conclusions

- *Could titration of the selenoprotein mRNA levels result in increased selenoprotein yield?*

  ⇨ Yes, to some extent. In the study we found that selenoprotein yield continuously surprisingly, increased, with higher mRNA levels (but Sec insertion specificity was still rather low).

- *What would be a good way for producing a system capable of such titration?*

  ⇨ The P_{BAD} promoter from the araBAD system was originally designed for a titration purpose and it suited well the purposes of our study. However, it is important to point out that araE needs to be constitutively expressed for a correct interpretation of the titration data, i.e. for P_{BAD}-driven transcription to be titrated on the individual cell level.

- *Can we lower the RF2 levels by a slower growth rate and thereby yield higher Sec insertion into the recombinant selenoprotein?*

  ⇨ Reducing the growth rate using minimal medium was not favorable for selenoprotein production, even if RF2 levels were lowered, due to a very low overall yield under such conditions. However, the increase in yield at late exponential phase may in part have been due to decreased RF2 levels.

- *What is the impact on the selenoprotein yield, when using co-expression of the sel-genes from the pSUABC plasmid under different growth conditions?*

⇨ Co-transformation with pSUABC results in increased selenoprotein yield in most growth conditions, with very few exceptions.

- *Could diverging from standard protocols for recombinant protein production increase the yield of recombinant selenoproteins?*

  ⇨ Yes, yield was quadrupled when induction was initiated at late exponential phase and continued at 24^{o}C for 24 hours.

- *Can multi-dimensional analysis be applied to determine optimal growth conditions?*

  ⇨ Yes, multidimensional analysis and data mining were useful tools for optimizing protein yield in a recombinant system.

3.2 Paper II - conclusions

- *Could an E. coli strain specialized for selenoprotein production be made by knocking down levels of RF2?*

  ⇨ Yes, knocking down RF2 levels clearly increased the specificity of Sec insertion into a recombinant selenoprotein. However, as shown in this study, the production protocols for expressing larger amounts of highly active selenoprotein must still be imporved, since lowering the RF2 levels also gave lower total yields.

- *Which existing technique would be suitable for replacing a native chromosomal promoter with a synthetic titrable promoter?*

  ⇨ The λ Red technique developed by Datsenko et al [169] has provend to be highly suitable for replacing genes in *E. coli* and tailoring specifically engineered variants of the chromosomal genes, including their promoter regions.

- *Which would be the appropriate promoter to use for titration and knock down of the RF2 levels?*

  ⇨ The P_{BAD} promoter has proven to be a useful tool because of its capacity of tightly shutting down the promoter as well as good titrational abilities, but for those purposes it must be used in a host strain constitutively overexpressing araE.

- *How does overexpression of RF2 affect growth of the bacteria?*

  ⇨ In a system where the *prfB* autoregulatory internal UGA codon is still intact, overexpression of RF2 up to 3- to 7-fold higher levels than normal were possible and at such conditions the growth rate decreased slightly.

- *How would knock down of the RF2 levels affect growth?*

  ⇨ Knock down of RF2 in the ORaa strain resulted in a bacteriostatic effect that severely reduced (or stopped) growth of the bacteria, but with maintained capabilities of regaining growth if RF2 levels were restored.

- *What growth technique would allow us to knockdown RF2 simultaneously as recombinant selenoprotein is produced?*

  ⇨ A continuously turbidostatic fermentor allowed us to induce selenoprotein production simultaneously with RF2 levels being knocked down, without altering other conditions for the culture.

- *How does knock-down levels of RF2 protein affect the recombinant yield and the specific activity of the selenoprotein?*

  ⇨ With lowering RF2 levels, the bacterial growth was much slower and led to expression of much less total amounts of recombinant selenoprotein. However, the specificity of Sec incorporation into recombinant TrxR increased significantly with decreased RF2 levels.

3.3 Paper III - conclusions

- *What is the molecular mechanism for purification of dimeric rat TrxR1 over a PAO-sepharose column?*

  ⇨ The PAO Sepharose bound only to Sec-containing subunits of TrxR1 heterodimers, with Sec-containing and UGA-truncated subunits not being separated from each other.

- *What is the explanation for the very high specific activity yielded from PAO-sepharose purification?*

  ⇨ It was revealed that the Sec-containing recombinant rat TrxR had significantly higher inherent specific activity than previously thought and that this high activity had previously been masked by co-purification of non-active UGA-truncated subunits.

- *In what form are recombinant rat TrxR expressed? Heterodimers or homodimers?*

  ⇨ The enzyme was expressed in the form of both heterodimers and homodimers, with the homodimers being either two Sec-containing subunits or two UGA-truncated subunits.

- *What is the theoretical highest possible specific activity of recombinant rat thioredoxin reductase?*

  ⇨ According to our selenium determinations and activity measurements of many different preparations, the specific activity of 100% Sec-containing dimeric recombinant rat TrxR1 should be around 80 U/mg.

4 Discussion

In this thesis we have further developed the recombinant selenoprotein production system in *E. coli*. In the following section, some other methods used for producing selenoprotein will be discussed in comparison to our system.

4.1 Different methods for selenoprotein production

High-yield production methods for selenoproteins have been and still are a major limiting factor in the field of selenoprotein research. The task of producing selenoproteins is not trivial, due to the complexity of Sec incorporation machineries and their species specificity. We have used *E. coli* as host organism for our studies, but several other techniques may be attempted for production of selenoproteins. Let us first, however, discuss advantages and disadvantages of the *E. coli* system.

4.1.1 Recombinant expression of selenoproteins in *E. coli*

It is truly astonishing what reserve capacity *E. coli* has to produce recombinant selenoproteins. In 1992, it was first shown that *E. coli* have a high capacity for production of recombinant selenoproteins, when its endogenous formate dehydrogenase H was overexpressed at high yields [178]. The expression of selenoproteins in *E. coli* necessitates the presence of a SECIS element positioned correctly in the mRNA, so that the *E. coli* SelB protein is compatible with that SECIS. This was taken in consideration when a Cys-to-Sec mutant of formate dehydrogenase from *Methanobacterium fromicicum* was expressed in *E. coli* where the gene was altered, not only from UGC (Cys) to UGA (Sec) but also so that it contained an *E. coli* SECIS downstream of the UGA [179]. Similar approaches were applied when rat TrxR1 was first expressed in *E. coli* [114], which is the technique that has been used in this thesis. For its expression, the rat TrxR1 gene was fused with a variant of the SECIS element from the *E. coli fdhH* gene. The Sec-encoding UGA in a gene for TrxR1 is located as the penultimate codon, which allowed the SECIS element to be positioned in the 3´UTR but still be recognized by SelB and without interfering with the coding sequence of the gene (See **Figure 15**). This strategy was also performed by others for a variant of human TrxR1 (KDRF) [180]. Moreover, the same strategy has

been successfully developed for production of Sel-tagged proteins, whereby the Sec-containing active site motif of TrxR is fused to other proteins in order to enable a range of biotechnological applications based upon the biochemical features of this motif [174-177, 181]. A similar methodology was also used in production of TGR from *Schistosoma mansoni*, which has proven to be promising key drug target for drug treatment of Bilharzia [182].

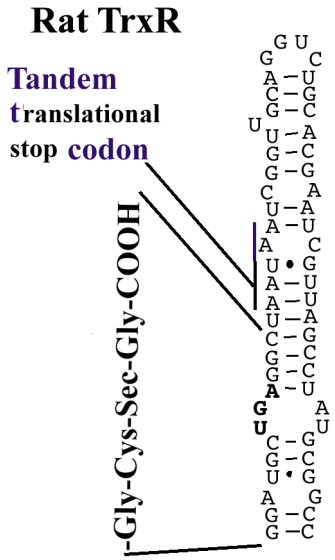


Figure 15. SECIS element used for expression of rat TrxR in *E. coli*. The altered SECIS element from the pET-TRS_{TER} plasmid. Included in the structure is the C-terminal end of TrxR and the SECIS element from *E. coli fdhF*.

Heterologous selenoproteins have also been expressed in *E. coli* with an internal Sec residue, with the *E. coli* SECIS placed 11 nt downstream of the UGA, within the open reading frame; the so-produced selenoproteins will then also necessarily have other amino acids than Sec dictated by the SECIS element as dictated from its invariable sequence corresponding to the loop region binding SelB (see the **Introduction** for details). The first heterologous protein expressed in *E. coli* with the Sec and SECIS placed within the open reading frame was phospholipid hydroperoxide glutathione peroxidase [183]. The effect of only having the SECIS introduced resulted in 28% activity loss, as compared to that of no mutation due to the SECIS element being engineered in the open reading frame. The full-length product produced by introduction of a UGA at the otherwise encoded Cys residue in the active site gave four times higher activity [183]. Glutathione S-transferase (GST) from *schistosoma*

japonica was expressed with its active site Tyr changed to Sec in *E. coli* using a similar approach [184]. In this case the introduction of a bacterial-type SECIS caused no activity loss, however, the active site Tyr-to-Sec mutation resulted in a loss of activity and no gain of any other measurable properties in terms of redox activity [184].

In conclusion, the *E. coli* system for recombinant selenoprotein production, as studied in this thesis and discussed above, has its major advantage in the high yields and relative ease for selenoprotein production. The major disadvantage of the E. coli system is the need for engineering of a SECIS element in the gene for the selenoprotein. In the case of TrxR or Sel-tagged proteins, this is no problem, whereas its requirement poses a major problem in production of selenoproteins containing an internal Sec residue.

4.1.2 Native Chemical Ligation

Native Chemical Ligation (NCL) is a technique that has been used for synthesis of many types of proteins [185]. The NCL methodology is based upon the native coupling with a C-terminal Cys, which undergoes trans-esterification with a second polypeptide N-terminal Cys. The formation of an acyl-group on the S is then migrated to the N atom to form a peptide bond, thereby joining two peptide fragments into one protein. Hilvert and colleagues have used this technique by replacing the N-terminal Cys of the second peptide with a Sec residue, thereby potentially producing selenoproteins [186]. The first experiment in that study showed a proof of concept, whereby they successfully ligated the pentapeptide LYRAG with L-selenocysteine with good yield. Furthermore, they successfully produced bovine pancreatic trypsin inhibitor (BPTI) with a Cys38-to-Sec mutation. The Sec mutant showed similar properties as the native enzyme [186].

The advantage of chemical ligation-based production of selenoproteins is the possibility to use non-natural peptide derivatives in the production, as well as the avoidance of adherence to biological limitations (such as the need for a SECIS in the *E. coli* production system). The disadvantage may be the chemical side-reactions that can be difficult to avoid, as well as the technical hurdles in making selenoproteins with a Sec residue that is located far from either end of the protein.

4.1.3 Auxotrophic expression systems

An additional technique that has been applied for the production of selenoproteins is the usage of auxotrophic strains where Sec can be incorporated by the sulfur pathway of $tRNA^{Cys}$-mediated insertion. Shen *et al* made a Ser9-to-Cys mutant of glutathione peroxidase and then expressed it in an auxotrophic strain in the presence of selenium in the medium, thereby turning Cys9 into Sec9. The Sec mutant activity was found to be comparable with that of naturally occurring GPx [187].

A similar approach was used to express mouse Grx with the active site residue Cys48 made into a Sec. The Sec-containing Grx active site was derived from mouse TGR and the Grx variant was introduced into the pET vector. The resulting enzyme was expressed in an auxotrophic strain in the presence of selenium and showed Grx and GPx activity [188]. Using similar methods, the two active site cysteine residues of *E. coli* Trx were replaced with Sec, whereby the expressed protein population contained 75-80% Sec-Sec, 5-10% Sec-Cys and 12-17% Cys-Cys [189].

The benefits of using auxotrophic strains for incorporation of Sec is that no SECIS or *sel*-genes are required. Drawbacks are that all Cys residues of an expressed protein have the ability to become converted to Sec and the technique is thereby not strictly Sec- or predefined residue-specific.

4.1.4 Production of selenoproteins using mammalian cells

Selenoproteins can also be produced by transfection of eukaryotic cells with selenoprotein-encoding genes compatible with the mammalian translation system, but this has so far only resulted in very low yield [190]. HEK293 cells transfected with human TrxR2 cDNA resulted in a 5.5-fold increase of total cellular TrxR activity [191] or similar effect with TrxR1. The benefits using eukaryotic cells to produce eukaryotic selenoproteins are several: lower risk for mis-folding of the protein, posttranslational modifications of mammalian selenoproteins can be obtained through natural routes, and no SECIS element has to be engineered inside the open reading frame while still allowing production of selenoproteins with internal Sec residues. However, retaining large quantities of selenoproteins using eukaryotic cells has so far not been possible and yields must be significantly increased before the mammalian expression system will be of any major or general use in selenoprotein research.

4.1.5 Intein-based selenoprotein production in combination with chemical ligation

Intein is the natural splicing of polypeptide chains after release from the ribosome, to form an active protein from a longer precursor. Intein-based techniques can be used to produce proteins with a C-terminal thiolester, that subsequently can be allowed to react with a Sec or Cys on the N-terminus of another peptide [192]. This technique was used to create a semisynthetic TrxR. A three-amino acid truncated form of TrxR was expressed with an intein. After a thiol-induced cleavage at the intein site, chemical ligation of the truncated TrxR with the synthetic tripeptide Gly-Sec-Cys was used to attack the N-terminus and thereby created a stabile TrxR protein [193].

4.2 General Concluding Discussion and Future Perspectives

The function of a selenoprotein can be implied, to some degree, by finding Cys homologues of the selenoprotein that have been studied in more detail, or through homology searches of preserved regions with known function, e.g. redox active sites. However, the detailed determination of function, kinetic parameters, structure, action of mechanism and other important biochemical parameters, requires a large quantity of the selenoprotein for it to be studied. Several of the different techniques used today for producing selenoproteins have limitations in yield or in the integrity of the protein (See the **Different methods for selenoprotein production**-section above and references therein). We have shown in this thesis that the capacity of *E. coli* for producing selenoproteins is very high (see **Paper I**, **Paper II** and [114]). In **Paper I** we showed that the already rather high level of recombinant selenoprotein production in *E. coli,* using previous protocols, could be increased further by altering the growth parameters in the culture used to express the selenoenzyme. The method used for determining optimal growth parameters for high selenoprotein yield (PCA) also demonstrated that multi-dimensional analysis can successfully be applied onto biological systems.

How could we improve the yield of recombinant selenoprotein production even further? There are several different approaches how this could potentially be achieved, of which some are listed here:

4.2.1 Optimization of growth parameters that have not yet been optimized.

- **Oxygen levels:** In fermentation process the oxygen levels are usually ranked as one of the most important parameter. Since the O_2 is a hydrophobic molecule, it can be difficult to get enough oxygen into the growth medium for optimal growth and protein production. The oxygen in a flask is rapidly consumed when a batch of aerobic growing bacteria used and parameters like formation of foam can have substantial effects on the oxygenation of the culture. It is possible that better controlled oxygen levels could increase yield in selenoprotein production – this has so far not at all been evaluated.
- **Medium usage:** The choice of carbon source could have great effect on a bacterial culture. In **Paper I** we showed that when glycerol was used as carbon source the effect of pSUABC was lower. This could be an effect of lower levels of the glucose-derived riboflavin precursor ribulose-5-phosphate [194]. Could some other medium additive be a limiting factor for selenoprotein production in *E. coli*? It is possible that much richer media than the full LB medium that we have used, would further increase selenoprotein yield.
- **pH:** Usually a stable pH in a fermentor process is to prefer; it could potentially be beneficial for selenoprotein yield if the pH was monitored and controlled throughout a cultivation and production process.
- **Stoichiometry between the individual *sel* -genes:** The use of co-expression of the *sel*-genes is very important for selenoprotein yield, however it has also been shown that too much *selB* lowers the Sec incorporation and that this effect can be reversed by overexpression of $tRNA^{Sec}$ [109] (see also discussion above). To regulate the stoichiometry between the *sel*-genes (increase all factors even more than with the pSUABC plasmid) might potentially increase the Sec incorporation as compared to the RF2-mediated truncation.

4.2.2 Further development of selenoprotein production system using the ORaa strain

In **Paper II,** we developed a strain that increased the Sec insertion as compared to RF2-mediated truncation by lowering the cellular RF2 levels. We believe that this strain can be a useful tool in future work expressing other selenoproteins in *E. coli,* since RF2 truncation is one of the major problems when producing selenoproteins. However, due to the difficulties in regulating the RF2 levels in a batch culture, it should be favorable to perform such experiments in a continuously turbidostatic fermentor similar to the one that we constructed in **Paper II.** A possible way is to use a fed-batch system where a larger culture is grown in glucose and when RF2 levels are knocked down, the culture is induced to produce the selenoprotein. Possible difficulties could be to find the right time to start the induction of the selenoprotein. Moreover, since the total yield was low when the RF2 levels were knocked down, these production experiments should preferentially be made in large culture volumes and with optimized conditions for both yield and Sec insertion specificity.

4.2.3 Further development using the PAO Sepharose

In **Paper III** we were interested in determining how the PAO Sepharose could function so efficiently in purification of recombinant rat TrxR1 expressed in *E. coli*. Our conclusions as described in **Paper III** and above, were that the TrxR we purified did not consist of only Sec-containing TrxR subunits but rather of a mixture of full-length and truncated subunits. As also discussed above, this means that the enzyme that was previously thought of as full-length still contained truncated subunits. One possible way to attain 100% full-length protein would be to bind the TrxR to the PAO Sepharose whereupon the column could potentially be treated with increasing concentrations of denaturing agents (such as urea) or high salt, to separate the two subunits from each other. The TrxR dimer would thereby release UGA-truncated subunits not bound to the PAO Sepharose, and the full-length subunits bound to the colmun could potentially form homodimers upon a subsequent elution. However, it is not clear wether this on-column refolding would be functional and it should also be

noted that the non-covalently bound FAD should not be lost from the full-length subunits due to this treatment.

4.2.4 Lowering the efficiency of RF2 by altering UGA codon context

As discussed in the **Introduction** under the RF2 part, the codon context of a UGA alters the efficiency of RF2-mediated translation termination. One possible approach would therefore be to alter the codons surrounding the Sec-encoding UGA to be less favorable for RF2 truncation. If this approach was combined with use of the ORaa strain, it may be possible to find good combinations of even higher yield and higher Sec insertion specificity during recombinant selenoprotein production.

4.3 Are the properties of recombinant rat TrxR found in Paper III of physiological relevance?

In **Paper III,** we showed that the recombinant rat TrxR had higher activity than we earlier had thought and that the selenium content of TrxR purified over the PAO sepharose was lower than previously believed (shown by measurement ration full-length and truncated enzyme using TOF-MS and by selenium content determination using ICP-MS). The question is, could the native rat TrxR1 have this high activity (around 80 U/mg)? Enzyme purified from rat liver was found to have 35 U/mg [161], however, the selenium content of that batch was not determined. To determine the activity of TrxR1 purified from rat tissue in correlation to its Se content should preferentially be done, as well as searching for possible inactive forms of the protein that could dilute the specific activity [195, 196].

In **Paper III**, a tetrameric form of TrxR ($TrxR_T$) was also discovered and characterized. A similar $TrxR_T$ have been detected in human cell lines [197], which might indicate that this tetrameric form might have a role in the cell. Possible functions could be an altered substrate specificity or, more likely, some activity-regulating function. It could also reflect some property not directly related to the enzymatic activity of TrxR1. More experiments must clearly be made to determine the potential physiological function of tetrameric TrxR, if any such exists.

4.4 Biotechnological use of Sec and selenoproteins

The Sel-tag is a multifunctional purification tag developed in our laboratory [176], as also discussed above. By introducing a small Sec-containing peptide at the C-terminal end of a recombinant protein expressed in *E. coli,* this can be used for protein purification and several other Sec-based applications. The peptide added as a Sel-tag corresponds to the four last amino acids of mammalian TrxR. The rat TrxR is thereby a naturally Sel-taged protein and purification of any protein labeled with the Sel-tag can be done over the PAO Sepharose, as described here in some detail for recombinant rat TrxR1 (see section above). The binding to the PAO Sepharose can only occur when the Sel-tag is reduced; the oxidized Sel-tag will not form a complex with the arsenic atom. The Sel-tag has several biotechnological applications; purification, Sec-specific fluorescent labeling, or labeling with either ^{75}Se as a gamma-emitter or ^{11}C as a positron emitter. Protocols for use of the Sel-tag and this type of labeling have been developed in our lab by Qing Cheng, Linda Johansson and coworkers [174-177] The purification using a Sel-tag has the advantage of beeing smaller and more specific than other common purification tags, including the His-tag, while it can also be used for both purification and labeling purposes.

4.5 Future perspectives – will recombinant selenoprotein production in E. coli be of importance?

The selenoprotein field is growing exponentially; increasing numbers of selenoproteins are discovered with obvious medical importance and new promising biotechnological applications are developed. As highlighted in this thesis, the need for efficient selenoprotein production systems is high in order to match the growing selenoprotein field. Production of recombinant proteins, selenoproteins and other factors in *E. coli* should be of great importance in the future. This thesis-work will hopefully and likely provide useful information and needed tools for the continued exploration of knowledge in the selenoprotein field.

5 Acknowledgments

The work presented in this thesis was performed at the Medical Noble Institute for Biochemistry, Department of Medical Biochemistry and Biophysics.
I would like to show my sincere gratitude to everyone that has helped me or in one way or another contributed to make this work possible.

There are several persons that I want thank in particular:

My supervisor **Associate Professor Elias Arnér**, thank you for your inspiring leadership and for never giving up. Your interest of science is truly contagious. Thanks.

Professor Arne Holmgren, thank you for sharing your knowledge, experience, laboratory and for all your nice stories around the coffee table.
Lena Ringdén for excellent secretary work and for your support.
Professor Peter Reichard for opening the field of the thioredoxin system and for being a source of inspiration

All the members of the EA-group. **Qing Cheng**, you are a truly good friend, I thank you for all the nice talks (science and other) and being been the nice person that you are (also thanks for all the help). **Stefanie Prast-Nielsen**, **Sofi Eriksson**, **Pascal Dammeyer** and **Marcus Cebula**, thanks your friendship. It has been so much fun. Also thanks for your support. Keep on doing good science!

The "Original" (or old) EA-group. **Dr. Linda Johansson**, **Dr. Anna-Klara Rundlöf** and **Dr. Karin Anestål**. Thanks for introducing me to the lab and all the fun. Those were the days ☺

My students: **Lisa Guldberg, Dr. Henrikas Nivinskas** and **Ewa-Maria Roswall,** for your good effort. An extra thanks to **Lisa**, it has been lovely to work with you.

Other member of the EA-group: **Dr. Tomas Nordman**, **Anne-Barbara Witte** and all the other persons, postdocs, students guest scientists (especially: **Dr. Oliver Rackham, Dr. Aleksandra Filipovska and Dr. Sharon Rozkovsky**), that have been a part of the group. I hope that you enjoyed it as much as I did.

Dr. Horst Lillig, for being a good roommate. **Horst** and **Dr. Carsten Berndt** for ice-war, metal, punk and science. Especially thanks for "das Sparnferkel", lecker. **Nicole** och **Diana.** Thanks for the cakes.

Dr. Maria Lönn, Christoph Hudeman, Tomas Gustafsson, Sergio Montano, Eva-Maria Hanschmann, Lars Bräutigam for good friendship. **Maria** and **Eva-Maria**, thanks for your help**, Dr. Jun Lu**, **Dr. Christina Carvalho. Dr. Yatao Du**. **Farnaz Zahedi Avval**, **Jack Andrzejewski, Dr. Javier Avila-Carino, Rolf Eliasson** Thank you all for the fun in the lab, for all the work that we have done. We are a good team!

Dr. Laura Papp, Dr. Aristi Potamitou Fernandez, Dr. Eng-Hui Chew, Dr. Isac Hashemy, Dr. Mari Enoksson, Dr. Johanna Ungerstedt, Dr. Catrine Johansson. Thank you all for being so helpful and for all the fun.
Dr. Alexios Vlamis-Gardikas thank you for your support and your friendship.

Professor Vera Bianca, Dr. Jian-Gou Fang, Dr. Liangwei Zhong, Monica Riveros, Anders Lundsjö, Sören och Svant. Thank for the help.
All MBB members. Especially: **Daniel, Emma, Yilams, Asi, Robert, Anderas, Essam, Susanne** and **Gunvor.**
Pedro and **Isabell** and everyone in the **Malaria-group**! Tanks
Dr. Sven Hamp and **Dr. Per Kårsnäs** and IBK.
Gustav Rengby. Thanks for all showing RDS and other systems.

To my **Friend**s:
Kemisterna från IBK bör få sin del av uppmärksamheten speciellt: **Magnus "Glappkäft" Janard, Robert "Robbelibobban bus" Steen, Sohrab "Sovab" Bellhav, Henrik "Frispel" Friberg och Daniel "Kung bland de korta" Hermansson.** Och alla andra! Tack för festerna och allt kul.

Grabbarna Grus från Bromsten. **Ulrik**, **Patrik** och **Roger**. Vänskap man kan lita på.
Tensta crew. **Eddie**, **Helen** och **Mats**.
Stefan och **Nettis.** Ni var världens bästa grannar! **Emil,** ett lite tack tack ska du ha. **Per Andersson, Anna, Pierre, Kita, Emma, Maria, Anna och ni alla andra**.
Tack till följande etablisimang/föreningar: **Flygande Käglorna**, **KÄK** och **Chrilles Bodega**.
Ulrik och **Mange**. Två helt suveräna polare.
Hela Familjen **Hedenskog**. Ni har stötta mig på alla möjliga sätt. Tack!
Hasse Magnét, du var en källa till glädje och insperation, du är saknad.

Familjen:

Mamma och **Pappa**. Tack för erat stöd och uppmuntran, detta skulle inte ha gått utan er. Min fina **syster** och **bror** och**, Karna** och **Gustav.** Jag älskar er alla.
Moster **Kerstin**. Du är helt underbar, tack för allt!
Tack till resten av hela familjen, kussarna **Malin** och **Larsa**. Tack **Malin** för de undebara London Weekends som du varit med att skapa. Faster **Kicki. Farmor** och **Farfar**
Lenart och **Maggi**, **Enar** och **Inga, Elsa, Marianne** och **Hjalmar**, tack för erat stöd.
Anders, det är helt suveränt att just du hittade till våran familj. Tjoho!
Lisa, **Christian** och **Malin** tack för allt skoj.
Sofia och **Fredrik** alla de trevliga stunder, erat stöd och alla roliga middagar har helt klart hjälp mig klara av detta.
Tack till min systerson **Leonard**, som hjälpte till att fatta några stora beslut.

Min älskade **Sara**. Ditt stöd och kärlek har varit oändligt och du är den finaste världen! Jag älskar dig.

6 References

[1] J. Pinsent, The need for selenite and molybdate in the formation of formic dehydrogenase by members of the coli-aerogenes group of bacteria, Biochem J 57 (1954) 10-16.

[2] K. Schwarz, and C. M. Foltz, Selenium as n itegral part of "Factor 3" against dietary necrotic liver degeneration, J Am Chem Soc 79 (1957) 3292-3293.

[3] H. Duddeck, Selenium-77 nuclear magnetic resonance spectroscopy, Progress in NMR Spectroscopy 27 (1995) 1-323.

[4] G. J. D'Angio, M. Loken, and M. Nesbit, Radionuclear (75Se) identification of tumor in children with neuroblastoma, Radiology 93 (1969) 615-617.

[5] R. Ferraris, G. Galatola, A. Barlotta, R. Pellerito, M. Fracchia, F. Cottino, and M. de la Pierre, Measurement of bile acid half-life using [75Se]HCAT in health and intestinal diseases. Comparison with [75Se]HCAT abdominal retention methods, Dig Dis Sci 37 (1992) 225-232.

[6] I. D. Whitton, and A. E. Houlder, 75Se-sodium selenite as a diagnostic aid in the assessment of peripheral pulmonary opacities, Thorax 34 (1979) 102-105.

[7] J. O. Nriagu, Global cycling of selenium, ed., CRC Press Inc., Boca Raton 1989.

[8] M. L. Berrow, and A. M. Ure, Geological materials and soil, ed., CRC Press Inc, Boca Raton 1989.

[9] G. F. Combs, Jr., Selenium in global food systems, Br J Nutr 85 (2001) 517-547.

[10] C. D. Thomson, Assessment of requirements for selenium and adequacy of selenium status: a review, Eur J Clin Nutr 58 (2004) 391-402.

[11] K. Ge, and G. Yang, The epidemiology of selenium deficiency in the etiological study of endemic diseases in China, Am J Clin Nutr 57 (1993) 259S-263S.

[12] M. A. Beck, O. A. Levander, and J. Handy, Selenium deficiency and viral infection, J Nutr 133 (2003) 1463S-1467S.

[13] M. L. Scott, The selenium dilemma, J Nutr 103 (1973) 803-810.

[14] K. A. See, P. S. Lavercombe, J. Dillon, and R. Ginsberg, Accidental death from acute selenium poisoning, Med J Aust 185 (2006) 388-389.

[15] R. Brigelius-Flohe, Selenium compounds and selenoproteins in cancer, Chem Biodivers 5 (2008) 389-395.

[16] M. P. Rayman, The importance of selenium to human health, Lancet 356 (2000) 233-241.

[17] M. P. Rayman, Selenium in cancer prevention: a review of the evidence and mechanism of action, Proc Nutr Soc 64 (2005) 527-542.

[18] A. Nelson, O. G. Fitzhugh, and H. O. Cavery, Liver Tumors following cirrhosis caused by selenium in rat, Cancer Res. 3 (1943) 230-236.

[19] L. C. Clark, G. F. Combs, Jr., B. W. Turnbull, E. H. Slate, D. K. Chalker, J. Chow, L. S. Davis, R. A. Glover, G. F. Graham, E. G. Gross, A. Krongrad, J. L. Lesher, Jr., H. K. Park, B. B. Sanders, Jr., C. L. Smith, and J. R. Taylor, Effects of selenium supplementation for cancer prevention in patients with carcinoma of the skin. A randomized controlled trial. Nutritional Prevention of Cancer Study Group, Jama 276 (1996) 1957-1963.

[20] E. A. Klein, I. M. Thompson, S. M. Lippman, P. J. Goodman, D. Albanes, P. R. Taylor, and C. Coltman, SELECT: the selenium and vitamin E cancer prevention trial, Urol Oncol 21 (2003) 59-65.

[21] S. M. Lippman, P. J. Goodman, E. A. Klein, H. L. Parnes, I. M. Thompson, Jr., A. R. Kristal, R. M. Santella, J. L. Probstfield, C. M. Moinpour, D. Albanes, P. R. Taylor, L. M. Minasian, A. Hoque, S. M. Thomas, J. J. Crowley, J. M. Gaziano, J. L. Stanford, E. D. Cook, N. E. Fleshner, M. M. Lieber, P. J. Walther, F. R. Khuri, D. D. Karp, G. G. Schwartz, L. G. Ford, and C. A. Coltman, Jr., Designing the Selenium and Vitamin E Cancer Prevention Trial (SELECT), J Natl Cancer Inst 97 (2005) 94-102.

[22] Y. Zhang, and V. N. Gladyshev, High content of proteins containing 21st and 22nd amino acids, selenocysteine and pyrrolysine, in a symbiotic deltaproteobacterium of gutless worm Olavius algarvensis, Nucleic Acids Res 35 (2007) 4952-4963.

[23] K. Taskov, C. Chapple, G. V. Kryukov, S. Castellano, A. V. Lobanov, K. V. Korotkov, R. Guigo, and V. N. Gladyshev, Nematode selenoproteome: the use of the selenocysteine insertion system to decode one codon in an animal genome? Nucleic Acids Res 33 (2005) 2227-2238.

[24] A. V. Lobanov, D. L. Hatfield, and V. N. Gladyshev, Selenoproteinless animals: selenophosphate synthetase SPS1 functions in a pathway unrelated to selenocysteine biosynthesis, Protein Sci 17 (2008) 176-182.

[25] L. Flohe, W. A. Gunzler, and H. H. Schock, Glutathione peroxidase: a selenoenzyme, FEBS Lett 32 (1973) 132-134.

[26] D. C. Turner, and T. C. Stadtman, Purification of protein components of the clostridial glycine reductase system and characterization of protein A as a selenoprotein, Arch Biochem Biophys 154 (1973) 366-381.

[27] J. R. Andreesen, and L. G. Ljungdahl, Formate dehydrogenase of *Clostridium thermoaceticum*: incorporation of selenium-75, and the effects of selenite, molybdate, and tungstate on the enzyme, J Bacteriol 116 (1973) 867-873.

[28] M. R. Bösl, K. Takaku, M. Oshima, S. Nishimura, and M. M. Taketo, Early embryonic lethality caused by targeted disruption of the mouse selenocysteine tRNA gene (Trsp), Proc Natl Acad Sci U S A 94 (1997) 5531-5534.

[29] G. V. Kryukov, S. Castellano, S. V. Novoselov, A. V. Lobanov, O. Zehtab, R. Guigo, and V. N. Gladyshev, Characterization of mammalian selenoproteomes, Science 300 (2003) 1439-1443.

[30] L. V. Papp, J. Lu, A. Holmgren, and K. K. Khanna, From selenium to selenoproteins: synthesis, identity, and their role in human health, Antioxid Redox Signal 9 (2007) 775-806.

[31] A. Kyriakopoulos, and D. Behne, Selenium-containing proteins in mammals and other forms of life, Rev Physiol Biochem Pharmacol 145 (2002) 1-46.

[32] L. Schomburg, U. Schweizer, and J. Kohrle, Selenium and selenoproteins in mammals: extraordinary, essential, enigmatic, Cell Mol Life Sci 61 (2004) 1988-1995.

[33] Y. S. Ho, J. L. Magnenat, R. T. Bronson, J. Cao, M. Gargano, M. Sugawara, and C. D. Funk, Mice deficient in cellular glutathione peroxidase develop normally and show no increased sensitivity to hyperoxia, J Biol Chem 272 (1997) 16644-16651.

[34] W. H. Cheng, Y. S. Ho, B. A. Valentine, D. A. Ross, G. F. Combs, Jr., and X. G. Lei, Cellular glutathione peroxidase is the mediator of body selenium to protect against paraquat lethality in transgenic mice, J Nutr 128 (1998) 1070-1076.

[35] F. F. Chu, R. S. Esworthy, Y. S. Ho, M. Bermeister, K. Swiderek, and R. W. Elliott, Expression and chromosomal mapping of mouse Gpx2 gene encoding the gastrointestinal form of glutathione peroxidase, GPX-GI, Biomed Environ Sci 10 (1997) 156-162.

[36] R. S. Esworthy, M. A. Baker, and F. F. Chu, Expression of selenium-dependent glutathione peroxidase in human breast tumor cell lines, Cancer Res 55 (1995) 957-962.

[37] R. Brigelius-Flohe, Tissue-specific functions of individual glutathione peroxidases, Free Radic Biol Med 27 (1999) 951-965.

[38] L. J. Yant, Q. Ran, L. Rao, H. Van Remmen, T. Shibatani, J. G. Belter, L. Motta, A. Richardson, and T. A. Prolla, The selenoprotein GPX4 is essential for mouse development and protects from radiation and oxidative damage insults, Free Radic Biol Med 34 (2003) 496-502.

[39] H. Imai, F. Hirao, T. Sakamoto, K. Sekine, Y. Mizukura, M. Saito, T. Kitamoto, M. Hayasaka, K. Hanaoka, and Y. Nakagawa, Early embryonic lethality caused by targeted disruption of the mouse PHGPx gene, Biochem Biophys Res Commun 305 (2003) 278-286.

[40] A. Miranda-Vizuete, J. Ljung, A. E. Damdimopoulos, J. A. Gustafsson, R. Oko, M. Pelto-Huikko, and G. Spyrou, Characterization of Sptrx, a novel member of the thioredoxin family specifically expressed in human spermatozoa, J Biol Chem 276 (2001) 31567-31574.

[41] S. R. Lee, J. R. Kim, K. S. Kwon, H. W. Yoon, R. L. Levine, A. Ginsburg, and S. G. Rhee, Molecular cloning and characterization of a mitochondrial selenocysteine-containing thioredoxin reductase from rat liver, J Biol Chem 274 (1999) 4722-4734.

[42] Q. A. Sun, Y. Wu, F. Zappacosta, K. T. Jeang, B. J. Lee, D. L. Hatfield, and V. N. Gladyshev, Redox regulation of cell signaling by selenocysteine in mammalian thioredoxin reductases, J Biol Chem 274 (1999) 24522-24530.

[43] Q. A. Sun, L. Kirnarsky, S. Sherman, and V. N. Gladyshev, Selenoprotein oxidoreductase with specificity for thioredoxin and glutathione systems, Proc Natl Acad Sci U S A 98 (2001) 3673-3678.

[44] D. Behne, A. Kyriakopoulos, H. Meinhold, and J. Kohrle, Identification of type I iodothyronine 5'-deiodinase as a selenoenzyme, Biochem Biophys Res Commun 173 (1990) 1143-1149.

[45] M. J. Guimaraes, D. Peterson, A. Vicari, B. G. Cocks, N. G. Copeland, D. J. Gilbert, N. A. Jenkins, D. A. Ferrick, R. A. Kastelein, J. F. Bazan, and A. Zlotnik, 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 (1996) 15086-15091.

[46] M. A. Motsenbocker, and A. L. Tappel, Selenium and selenoproteins in the rat kidney, Biochim Biophys Acta 709 (1982) 160-165.

[47] R. F. Burk, K. E. Hill, J. A. Awad, J. D. Morrow, T. Kato, K. A. Cockell, and P. R. Lyons, Pathogenesis of diquat-induced liver necrosis in selenium-deficient rats: assessment of the roles of lipid peroxidation and selenoprotein P, Hepatology 21 (1995) 561-569.

[48] A. D. Ferguson, V. M. Labunskyy, D. E. Fomenko, D. Arac, Y. Chelliah, C. A. Amezcua, J. Rizo, V. N. Gladyshev, and J. Deisenhofer, NMR structures of the selenoproteins Sep15 and SelM reveal redox activity of a new thioredoxin-like family, J Biol Chem 281 (2006) 3536-3543.

[49] D. Behne, A. Kyriakopoulos, M. Kalcklosch, C. Weiss-Nowak, H. Pfeifer, H. Gessner, and C. Hammel, Two new selenoproteins found in the prostatic glandular epithelium and in the spermatid nuclei, Biomed Environ Sci 10 (1997) 340-345.

[50] G. V. Kryukov, and V. N. Gladyshev, The prokaryotic selenoproteome, EMBO Rep 5 (2004) 538-543.

[51] Y. Zhang, and V. N. Gladyshev, Trends in selenium utilization in marine microbial world revealed through the analysis of the global ocean sampling (GOS) project, PLoS Genet 4 (2008) e1000095.

[52] S. Yooseph, G. Sutton, D. B. Rusch, A. L. Halpern, S. J. Williamson, K. Remington, J. A. Eisen, K. B. Heidelberg, G. Manning, W. Li, L. Jaroszewski, P. Cieplak, C. S. Miller, H. Li, S. T. Mashiyama, M. P. Joachimiak, C. van Belle, J. M. Chandonia, D. A. Soergel, Y. Zhai, K. Natarajan, S. Lee, B. J. Raphael, V. Bafna, R. Friedman, S. E. Brenner, A. Godzik, D. Eisenberg, J. E. Dixon, S. S. Taylor, R. L. Strausberg, M. Frazier, and J. C. Venter, The Sorcerer II Global Ocean Sampling expedition: expanding the universe of protein families, PLoS Biol 5 (2007) e16.

[53] G. M. Lacourciere, and T. C. Stadtman, Catalytic properties of selenophosphate synthetases: comparison of the selenocysteine-containing enzyme from Haemophilus influenzae with the corresponding cysteine-containing enzyme from *Escherichia coli*, Proc Natl Acad Sci U S A 96 (1999) 44-48.

[54] J. C. Cox, E. S. Edwards, and J. A. DeMoss, Resolution of distinct selenium-containing formate dehydrogenases from *Escherichia coli*, J Bacteriol 145 (1981) 1317-1324.

[55] G. Sawers, J. Heider, E. Zehelein, and A. Bock, Expression and operon structure of the sel genes of *Escherichia coli* and identification of a third selenium-containing formate dehydrogenase isoenzyme, J Bacteriol 173 (1991) 4983-4993.

[56] J. R. Andreesen, M. Wagner, D. Sonntag, M. Kohlstock, C. Harms, T. Gursinsky, J. Jager, T. Parther, U. Kabisch, A. Grantzdorffer, A. Pich, and B. Sohling, Various functions of selenols and thiols in anaerobic gram-positive, amino acids-utilizing bacteria, Biofactors 10 (1999) 263-270.

[57] U. C. Kabisch, A. Grantzdorffer, A. Schierhorn, K. P. Rucknagel, J. R. Andreesen, and A. Pich, Identification of D-proline reductase from Clostridium sticklandii as a selenoenzyme and indications for a catalytically active pyruvoyl group derived from a cysteine residue by cleavage of a proprotein, J Biol Chem 274 (1999) 8445-8454.

[58] M. Pfeiffer, R. Bingemann, and A. Klein, Fusion of two subunits does not impair the function of a [NiFeSe]-hydrogenase in the archaeon Methanococcus voltae, Eur J Biochem 256 (1998) 447-452.

[59] F. Zinoni, A. Birkmann, T. C. Stadtman, and A. Böck, Nucleotide sequence and expression of the selenocysteine-containing polypeptide of formate dehydrogenase (formate-hydrogen-lyase-linked) from *Escherichia coli*, Proc Natl Acad Sci U S A 83 (1986) 4650-4654.

[60] B. L. Berg, C. Baron, and V. Stewart, Nitrate-inducible formate dehydrogenase in *Escherichia coli* K-12. II. Evidence that a mRNA stem-loop structure is essential for decoding opal (UGA) as selenocysteine, J Biol Chem 266 (1991) 22386-22391.

[61] J. E. Cone, R. M. Del Rio, J. N. Davis, and T. C. Stadtman, Chemical characterization of the selenoprotein component of clostridial glycine reductase: identification of selenocysteine as the organoselenium moiety, Proc Natl Acad Sci U S A 73 (1976) 2659-2663.

[62] I. Chambers, J. Frampton, P. Goldfarb, N. Affara, W. McBain, and P. R. Harrison, The structure of the mouse glutathione peroxidase gene: the selenocysteine in the active site is encoded by the 'termination' codon, TGA, Embo J 5 (1986) 1221-1227.

[63] B. G. Barrell, A. T. Bankier, and J. Drouin, A different genetic code in human mitochondria, Nature 282 (1979) 189-194.

[64] D. E. Fomenko, S. M. Marino, and V. N. Gladyshev, Functional diversity of cysteine residues in proteins and unique features of catalytic redox-active cysteines in thiol oxidoreductases, Mol Cells 26 (2008) 228-235.

[65] F. Ursini, M. Maiorino, R. Brigelius-Flohe, K. D. Aumann, A. Roveri, D. Schomburg, and L. Flohe, Diversity of glutathione peroxidases, Methods Enzymol 252 (1995) 38-53.

[66] C. Rocher, J. L. Lalanne, and J. Chaudiere, Purification and properties of a recombinant sulfur analog of murine selenium-glutathione peroxidase, Eur J Biochem 205 (1992) 955-960.

[67] J. Köhrle, F. Jakob, B. Contempre, and J. E. Dumont, Selenium, the thyroid, and the endocrine system, Endocr Rev 26 (2005) 944-984.

[68] J. Köhrle, The deiodinase family: selenoenzymes regulating thyroid hormone availability and action, Cell Mol Life Sci 57 (2000) 1853-1863.

[69] M. J. Berry, J. D. Kieffer, J. W. Harney, and P. R. Larsen, Selenocysteine confers the biochemical properties characteristic of the type I iodothyronine deiodinase, J Biol Chem 266 (1991) 14155-14158.

[70] C. Buettner, J. W. Harney, and P. R. Larsen, The role of selenocysteine 133 in catalysis by the human type 2 iodothyronine deiodinase, Endocrinology 141 (2000) 4606-4612.

[71] G. G. Kuiper, W. Klootwijk, and T. J. Visser, Substitution of cysteine for selenocysteine in the catalytic center of type III iodothyronine deiodinase reduces catalytic efficiency and alters substrate preference, Endocrinology 144 (2003) 2505-2513.

[72] V. N. Gladyshev, S. V. Khangulov, M. J. Axley, and T. C. Stadtman, Coordination of selenium to molybdenum in formate dehydrogenase H from *Escherichia coli*, Proc Natl Acad Sci U S A 91 (1994) 7708-7711.

[73] M. J. Axley, A. Böck, and T. C. Stadtman, Catalytic properties of an *Escherichia coli* formate dehydrogenase mutant in which sulfur replaces selenium, Proc Natl Acad Sci U S A 88 (1991) 8450-8454.

[74] H. Tanaka, and T. C. Stadtman, Selenium-dependent clostridial glycine reductase. Purification and characterization of the two membrane-associated protein components, J Biol Chem 254 (1979) 447-452.

[75] G. E. Garcia, and T. C. Stadtman, *Clostridium sticklandii* glycine reductase selenoprotein A gene: cloning, sequencing, and expression in *Escherichia coli*, J Bacteriol 174 (1992) 7080-7089.

[76] X. M. Xu, B. A. Carlson, R. Irons, H. Mix, N. Zhong, V. N. Gladyshev, and D. L. Hatfield, Selenophosphate synthetase 2 is essential for selenoprotein biosynthesis, Biochem J 404 (2007) 115-120.

[77] X. M. Xu, B. A. Carlson, H. Mix, Y. Zhang, K. Saira, R. S. Glass, M. J. Berry, V. N. Gladyshev, and D. L. Hatfield, Biosynthesis of selenocysteine on its tRNA in eukaryotes, PLoS Biol 5 (2007) e4.

[78] A. Small-Howard, N. Morozova, Z. Stoytcheva, E. P. Forry, J. B. Mansell, J. W. Harney, B. A. Carlson, X. M. Xu, D. L. Hatfield, and M. J. Berry, Supramolecular complexes mediate selenocysteine incorporation in vivo, Mol Cell Biol 26 (2006) 2337-2346.

[79] L. V. Papp, J. Lu, F. Striebel, D. Kennedy, A. Holmgren, and K. K. Khanna, The redox state of SECIS binding protein 2 controls its localization and selenocysteine incorporation function, Mol Cell Biol 26 (2006) 4895-4910.

[80] L. Chavatte, B. A. Brown, and D. M. Driscoll, Ribosomal protein L30 is a component of the UGA-selenocysteine recoding machinery in eukaryotes, Nat Struct Mol Biol 12 (2005) 408-416.

[81] M. Rother, A. Resch, R. Wilting, and A. Böck, Selenoprotein synthesis in archaea, Biofactors 14 (2001) 75-83.

[82] M. Rother, A. Resch, W. L. Gardner, W. B. Whitman, and A. Böck, Heterologous expression of archaeal selenoprotein genes directed by the SECIS element located in the 3' non-translated region, Mol Microbiol 40 (2001) 900-908.

[83] M. J. Berry, L. Banu, Y. Y. Chen, S. J. Mandel, J. D. Kieffer, J. W. Harney, and P. R. Larsen, Recognition of UGA as a selenocysteine codon in type I deiodinase requires sequences in the 3' untranslated region, Nature 353 (1991) 273-276.

[84] M. Leibundgut, C. Frick, M. Thanbichler, A. Böck, and N. Ban, Selenocysteine tRNA-specific elongation factor SelB is a structural chimaera of elongation and initiation factors, Embo J 24 (2005) 11-22.

[85] M. J. Berry, R. M. Tujebajeva, P. R. Copeland, X. M. Xu, B. A. Carlson, G. W. Martin, 3rd, S. C. Low, J. B. Mansell, E. Grundner-Culemann, J. W. Harney, D. M. Driscoll, and D. L. Hatfield, Selenocysteine incorporation directed from the 3'UTR: characterization of eukaryotic EFsec and mechanistic implications, Biofactors 14 (2001) 17-24.

[86] T. Gursinsky, D. Gröbe, A. Schierhorn, J. Jäger, J. R. Andreesen, and B. Söhling, Factors and selenocysteine insertion sequence requirements for the synthesis of selenoproteins from a gram-positive anaerobe in Escherichia coli, Appl Environ Microbiol 74 (2008) 1385-1393.

[87] A. Böck, K. Forchhammer, J. Heider, W. Leinfelder, G. Sawers, B. Veprek, and F. Zinoni, Selenocysteine: the 21st amino acid, Mol Microbiol 5 (1991) 515-520.

[88] G. F. Kramer, and B. N. Ames, Isolation and characterization of a selenium metabolism mutant of Salmonella typhimurium, J Bacteriol 170 (1988) 736-743.

[89] M. Takahata, T. Tamura, K. Abe, H. Mihara, S. Kurokawa, Y. Yamamoto, R. Nakano, N. Esaki, and K. Inagaki, Selenite assimilation into formate dehydrogenase H depends on thioredoxin reductase in *Escherichia coli*, J Biochem 143 (2008) 467-473.

[90] S. Muller, J. Heider, and A. Böck, The path of unspecific incorporation of selenium in *Escherichia coli*, Arch Microbiol 168 (1997) 421-427.

[91] S. J. Klug, A. Huttenhofer, M. Kromayer, and M. Famulok, In vitro and in vivo characterization of novel mRNA motifs that bind special elongation factor SelB, Proc Natl Acad Sci U S A 94 (1997) 6676-6681.

[92] K. Forchhammer, W. Leinfelder, K. Boesmiller, B. Veprek, and A. Böck, Selenocysteine synthase from *Escherichia coli*. Nucleotide sequence of the gene (selA) and purification of the protein, J Biol Chem 266 (1991) 6318-6323.

[93] K. Forchhammer, W. Leinfelder, and A. Böck, Identification of a novel translation factor necessary for the incorporation of selenocysteine into protein, Nature 342 (1989) 453-456.

[94] K. Forchhammer, K. P. Rucknagel, and A. Böck, Purification and biochemical characterization of SELB, a translation factor involved in selenoprotein synthesis, J Biol Chem 265 (1990) 9346-9350.

[95] W. Leinfelder, E. Zehelein, M. A. Mandrand-Berthelot, and A. Böck, Gene for a novel tRNA species that accepts L-serine and cotranslationally inserts selenocysteine, Nature 331 (1988) 723-725.

[96] Z. Veres, L. Tsai, T. D. Scholz, M. Politino, R. S. Balaban, and T. C. Stadtman, Synthesis of 5-methylaminomethyl-2-selenouridine in tRNAs: 31P NMR studies show the labile selenium donor synthesized by the selD gene product contains selenium bonded to phosphorus, Proc Natl Acad Sci U S A 89 (1992) 2975-2979.

[97] I. Y. Kim, Z. Veres, and T. C. Stadtman, *Escherichia coli* mutant SELD enzymes. The cysteine 17 residue is essential for selenophosphate formation from ATP and selenide, J Biol Chem 267 (1992) 19650-19654.

[98] A. Ehrenreich, K. Forchhammer, P. Tormay, B. Veprek, and A. Böck, Selenoprotein synthesis in *E. coli*. Purification and characterisation of the enzyme catalysing selenium activation, Eur J Biochem 206 (1992) 767-773.

[99] R. S. Glass, W. P. Singh, W. Jung, Z. Veres, T. D. Scholz, and T. C. Stadtman, Monoselenophosphate: synthesis, characterization, and identity with the prokaryotic biological selenium donor, compound SePX, Biochemistry 32 (1993) 12555-12559.

[100] H. Mihara, T. Kurihara, T. Yoshimura, K. Soda, and N. Esaki, Cysteine sulfinate desulfinase, a NIFS-like protein of *Escherichia coli* with selenocysteine lyase and cysteine desulfurase activities. Gene cloning, purification, and characterization of a novel pyridoxal enzyme, J Biol Chem 272 (1997) 22417-22424.

[101] H. Mihara, M. Maeda, T. Fujii, T. Kurihara, Y. Hata, and N. Esaki, A nifS-like gene, csdB, encodes an *Escherichia coli* counterpart of mammalian selenocysteine lyase. Gene cloning, purification, characterization and preliminary x-ray crystallographic studies, J Biol Chem 274 (1999) 14768-14772.

[102] G. M. Lacourciere, H. Mihara, T. Kurihara, N. Esaki, and T. C. Stadtman, *Escherichia coli* NifS-like proteins provide selenium in the pathway for the biosynthesis of selenophosphate, J Biol Chem 275 (2000) 23769-23773.

[103] C. Baron, J. Heider, and A. Böck, Mutagenesis of selC, the gene for the selenocysteine-inserting tRNA-species in E. coli: effects on in vivo function, Nucleic Acids Res 18 (1990) 6761-6766.

[104] C. Baron, and A. Böck, 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 (1991) 20375-20379.

[105] A. Schön, A. Böck, G. Ott, M. Sprinzl, and D. Söll, The selenocysteine-inserting opal suppressor serine tRNA from E. coli is highly unusual in structure and modification, Nucleic Acids Res 17 (1989) 7159-7165.

[106] T. Ose, N. Soler, L. Rasubala, K. Kuroki, D. Kohda, D. Fourmy, S. Yoshizawa, and K. Maenaka, Structural basis for dynamic interdomain movement and RNA recognition of the selenocysteine-specific elongation factor SelB, Structure 15 (2007) 577-586.

[107] A. Huttenhofer, and A. Böck, Selenocysteine inserting RNA elements modulate GTP hydrolysis of elongation factor SelB, Biochemistry 37 (1998) 885-890.

[108] M. Thanbichler, and A. Böck, The function of SECIS RNA in translational control of gene expression in *Escherichia coli*, Embo J 21 (2002) 6925-6934.

[109] P. Tormay, A. Sawers, and A. Böck, Role of stoichiometry between mRNA, translation factor SelB and selenocysteyl-tRNA in selenoprotein synthesis, Mol Microbiol 21 (1996) 1253-1259.

[110] Z. Liu, M. Reches, I. Groisman, and H. Engelberg-Kulka, The nature of the minimal 'selenocysteine insertion sequence' (SECIS) in *Escherichia coli*, Nucleic Acids Res 26 (1998) 896-902.

[111] A. Hüttenhofer, E. Westhof, and A. Böck, Solution structure of mRNA hairpins promoting selenocysteine incorporation in *Escherichia coli* and their base-specific interaction with special elongation factor SELB, RNA 2 (1996) 354-366.

[112] C. Baron, J. Heider, and A. Böck, Interaction of translation factor SELB with the formate dehydrogenase H selenopolypeptide mRNA, Proc Natl Acad Sci U S A 90 (1993) 4181-4185.

[113] J. Heider, C. Baron, and A. Böck, Coding from a distance: dissection of the mRNA determinants required for the incorporation of selenocysteine into protein, Embo J 11 (1992) 3759-3766.

[114] E. S. J. Arnér, H. Sarioglu, F. Lottspeich, A. Holmgren, and A. Böck, High-level expression in *Escherichia coli* of selenocysteine-containing rat thioredoxin reductase utilizing gene fusions with engineered bacterial-type SECIS elements and co-expression with the *selA*, *selB* and *selC* genes, J Mol Biol 292 (1999) 1003-1016.

[115] K. Kawakami, T. Inada, and Y. Nakamura, Conditionally lethal and recessive UGA-suppressor mutations in the *prfB* gene encoding peptide chain release factor 2 of *Escherichia coli*, J Bacteriol 170 (1988) 5378-5381.

[116] E. Scolnick, R. Tompkins, T. Caskey, and M. Nirenberg, Release factors differing in specificity for terminator codons, Proc Natl Acad Sci U S A 61 (1968) 768-774.

[117] L. Frolova, X. Le Goff, H. H. Rasmussen, S. Cheperegin, G. Drugeon, M. Kress, I. Arman, A. L. Haenni, J. E. Celis, M. Philippe, and et al., A highly conserved eukaryotic protein family possessing properties of polypeptide chain release factor, Nature 372 (1994) 701-703.

[118] M. Dontsova, L. Frolova, J. Vassilieva, W. Piendl, L. Kisselev, and M. Garber, Translation termination factor aRF1 from the archaeon Methanococcus jannaschii is active with eukaryotic ribosomes, FEBS Lett 472 (2000) 213-216.

[119] W. P. Tate, A. L. Beaudet, and C. T. Caskey, Influence of guanine nucleotides and elongation factors on interaction of release factors with the ribosome, Proc Natl Acad Sci U S A 70 (1973) 2350-2355.

[120] A. V. Zavialov, L. Mora, R. H. Buckingham, and M. Ehrenberg, Release of peptide promoted by the GGQ motif of class 1 release factors regulates the GTPase activity of RF3, Mol Cell 10 (2002) 789-798.

[121] D. V. Freistroffer, M. Y. Pavlov, J. MacDougall, R. H. Buckingham, and M. Ehrenberg, Release factor RF3 in E.coli accelerates the dissociation of release factors RF1 and RF2 from the ribosome in a GTP-dependent manner, Embo J 16 (1997) 4126-4133.

[122] K. Ito, M. Uno, and Y. Nakamura, Single amino acid substitution in prokaryote polypeptide release factor 2 permits it to terminate translation at all three stop codons, Proc Natl Acad Sci U S A 95 (1998) 8165-8169.

[123] I. Stansfield, K. M. Jones, V. V. Kushnirov, A. R. Dagkesamanskaya, A. I. Poznyakovski, S. V. Paushkin, C. R. Nierras, B. S. Cox, M. D. Ter-Avanesyan, and M. F. Tuite, The products of the SUP45 (eRF1) and SUP35 genes interact to mediate translation termination in Saccharomyces cerevisiae, Embo J 14 (1995) 4365-4373.

[124] G. Zhouravleva, L. Frolova, X. Le Goff, R. Le Guellec, S. Inge-Vechtomov, L. Kisselev, and M. Philippe, Termination of translation in eukaryotes is governed by two interacting polypeptide chain release factors, eRF1 and eRF3, Embo J 14 (1995) 4065-4072.

[125] Y. Nakamura, K. Ito, and L. A. Isaksson, Emerging understanding of translation termination, Cell 87 (1996) 147-150.

[126] L. Y. Frolova, R. Y. Tsivkovskii, G. F. Sivolobova, N. Y. Oparina, O. I. Serpinsky, V. M. Blinov, S. I. Tatkov, and L. L. Kisselev, Mutations in the highly conserved GGQ motif of class 1 polypeptide release factors abolish ability of human eRF1 to trigger peptidyl-tRNA hydrolysis, Rna 5 (1999) 1014-1020.

[127] D. H. Shin, J. Brandsen, J. Jancarik, H. Yokota, R. Kim, and S. H. Kim, Structural analyses of peptide release factor 1 from Thermotoga maritima reveal domain flexibility required for its interaction with the ribosome, J Mol Biol 341 (2004) 227-239.

[128] B. Vestergaard, L. B. Van, G. R. Andersen, J. Nyborg, R. H. Buckingham, and M. Kjeldgaard, Bacterial polypeptide release factor RF2 is structurally distinct from eukaryotic eRF1, Mol Cell 8 (2001) 1375-1382.

[129] K. Ito, M. Uno, and Y. Nakamura, A tripeptide 'anticodon' deciphers stop codons in messenger RNA, Nature 403 (2000) 680-684.

[130] M. Uno, K. Ito, and Y. Nakamura, Polypeptide release at sense and noncognate stop codons by localized charge-exchange alterations in translational release factors, Proc Natl Acad Sci U S A 99 (2002) 1819-1824.

[131] W. J. Craigen, R. G. Cook, W. P. Tate, and C. T. Caskey, Bacterial peptide chain release factors: conserved primary structure and possible frameshift regulation of release factor 2, Proc Natl Acad Sci U S A 82 (1985) 3616-3620.

[132] W. J. Craigen, and C. T. Caskey, Expression of peptide chain release factor 2 requires high-efficiency frameshift, Nature 322 (1986) 273-275.

[133] J. Parker, Errors and alternatives in reading the universal genetic code, Microbiol Rev 53 (1989) 273-298.

[134] J. F. Atkins, R. B. Weiss, S. Thompson, and R. F. Gesteland, Towards a genetic dissection of the basis of triplet decoding, and its natural subversion: programmed reading frame shifts and hops, Annu Rev Genet 25 (1991) 201-228.

[135] V. Marquez, D. N. Wilson, W. P. Tate, F. Triana-Alonso, and K. H. Nierhaus, Maintaining the ribosomal reading frame: the influence of the E site during translational regulation of release factor 2, Cell 118 (2004) 45-55.

[136] R. B. Weiss, D. M. Dunn, A. E. Dahlberg, J. F. Atkins, and R. F. Gesteland, Reading frame switch caused by base-pair formation between the 3' end of 16S rRNA and the mRNA during elongation of protein synthesis in *Escherichia coli*, Embo J 7 (1988) 1503-1507.

[137] J. F. Curran, Analysis of effects of tRNA:message stability on frameshift frequency at the *Escherichia coli* RF2 programmed frameshift site, Nucleic Acids Res 21 (1993) 1837-1843.

[138] B. C. Donly, C. D. Edgar, F. M. Adamski, and W. P. Tate, Frameshift autoregulation in the gene for *Escherichia coli* release factor 2: partly functional mutants result in frameshift enhancement, Nucleic Acids Res 18 (1990) 6517-6522.

[139] E. S. Poole, C. M. Brown, and W. P. Tate, The identity of the base following the stop codon determines the efficiency of in vivo translational termination in *Escherichia coli*, EMBO J 14 (1995) 151-158.

[140] W. P. Tate, J. B. Mansell, S. A. Mannering, J. H. Irvine, L. L. Major, and D. N. Wilson, UGA: a dual signal for 'stop' and for recoding in protein synthesis, Biochemistry (Mosc) 64 (1999) 1342-1353.

[141] C. T. Caskey, W. C. Forrester, W. Tate, and C. D. Ward, Cloning of the *Escherichia coli* release factor 2 gene, J Bacteriol 158 (1984) 365-368.

[142] K. Kawakami, Y. H. Jonsson, G. R. Björk, H. Ikeda, and Y. Nakamura, Chromosomal location and structure of the operon encoding peptide-chain-release factor 2 of *Escherichia coli*, Proc Natl Acad Sci U S A 85 (1988) 5620-5624.

[143] R. A. Van Bogelen, V. Vaughn, and F. C. Neidhardt, Gene for heat-inducible lysyl-tRNA synthetase (*lysU*) maps near cadA in *Escherichia coli*, J Bacteriol 153 (1983) 1066-1068.

[144] I. N. Hirshfield, P. L. Bloch, R. A. Van Bogelen, and F. C. Neidhardt, Multiple forms of lysyl-transfer ribonucleic acid synthetase in *Escherichia coli*, J Bacteriol 146 (1981) 345-351.

[145] J. B. Mansell, D. Guevremont, E. S. Poole, and W. P. Tate, A dynamic competition between release factor 2 and the tRNA(Sec) decoding UGA at the recoding site of *Escherichia coli* formate dehydrogenase H, Embo J 20 (2001) 7284-7293.

[146] A. G. Cridge, L. L. Major, A. A. Mahagaonkar, E. S. Poole, L. A. Isaksson, and W. P. Tate, Comparison of characteristics and function of translation termination signals between and within prokaryotic and eukaryotic organisms, Nucleic Acids Res 34 (2006) 1959-1973.

[147] S. Mottagui-Tabar, and L. A. Isaksson, Only the last amino acids in the nascent peptide influence translation termination in *Escherichia coli* genes, FEBS Lett 414 (1997) 165-170.

[148] S. Mottagui-Tabar, A. Bjornsson, and L. A. Isaksson, The second to last amino acid in the nascent peptide as a codon context determinant, Embo J 13 (1994) 249-257.

[149] A. Bjornsson, S. Mottagui-Tabar, and L. A. Isaksson, Structure of the C-terminal end of the nascent peptide influences translation termination, Embo J 15 (1996) 1696-1704.

[150] K. E. Sandman, and C. J. Noren, The efficiency of *Escherichia coli* selenocysteine insertion is influenced by the immediate downstream nucleotide, Nucleic Acids Res 28 (2000) 755-761.

[151] E. S. J. Arnér, and A. Holmgren, Physiological functions of thioredoxin and thioredoxin reductase, Eur J Biochem 267 (2000) 6102-6109.

[152] C. H. Lillig, and A. Holmgren, Thioredoxin and related molecules--from biology to health and disease, Antioxid Redox Signal 9 (2007) 25-47.

[153] J. Nordberg, and E. S. Arnér, Reactive oxygen species, antioxidants, and the mammalian thioredoxin system, Free Radic Biol Med 31 (2001) 1287-1312.

[154] L. Zhong, E. S. Arnér, and A. Holmgren, Structure and mechanism of mammalian thioredoxin reductase: the active site is a redox-active selenolthiol/selenenylsulfide formed from the conserved cysteine-selenocysteine sequence, Proc Natl Acad Sci U S A 97 (2000) 5854-5859.

[155] T. Sandalova, L. Zhong, Y. Lindqvist, A. Holmgren, and G. Schneider, Three-dimensional structure of a mammalian thioredoxin reductase: implications for mechanism and evolution of a selenocysteine-dependent enzyme, Proc Natl Acad Sci U S A 98 (2001) 9533-9538.

[156] H. Eklund, F. K. Gleason, and A. Holmgren, Structural and functional relations among thioredoxins of different species, Proteins 11 (1991) 13-28.

[157] S. Gromer, L. Johansson, H. Bauer, L. D. Arscott, S. Rauch, D. P. Ballou, C. H. Williams, Jr., R. H. Schirmer, and E. S. Arnér, Active sites of thioredoxin

reductases: why selenoproteins? Proc Natl Acad Sci U S A 100 (2003) 12618-12623.

[158] S. Gromer, J. Wissing, D. Behne, K. Ashman, R. H. Schirmer, L. Flohe, and K. Becker, A hypothesis on the catalytic mechanism of the selenoenzyme thioredoxin reductase, Biochem J 332 (Pt 2) (1998) 591-592.

[159] L. Zhong, and A. Holmgren, Essential role of selenium in the catalytic activities of mammalian thioredoxin reductase revealed by characterization of recombinant enzymes with selenocysteine mutations, J Biol Chem 275 (2000) 18121-18128.

[160] L. Zhong, E. S. Arnér, J. Ljung, F. Åslund, and A. Holmgren, Rat and calf thioredoxin reductase are homologous to glutathione reductase with a carboxyl-terminal elongation containing a conserved catalytically active penultimate selenocysteine residue, J Biol Chem 273 (1998) 8581-8591.

[161] M. Luthman, and A. Holmgren, Rat liver thioredoxin and thioredoxin reductase: purification and characterization, Biochemistry 21 (1982) 6628-6633.

[162] E. S. J. Arnér, and A. Holmgren, Measurement of thioredoxin and thioredoxin reductase, ed., John Wiley and Sons, Inc., New York 2000.

[163] L. M. Guzman, D. Belin, M. J. Carson, and J. Beckwith, Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter, J Bacteriol 177 (1995) 4121-4130.

[164] C. M. Johnson, and R. F. Schleif, In vivo induction kinetics of the arabinose promoters in *Escherichia coli*, J Bacteriol 177 (1995) 3438-3442.

[165] A. Novick, and M. Weiner, Enzyme Induction As An All-Or-None Phenomenon, Proc Natl Acad Sci U S A 43 (1957) 553-566.

[166] D. A. Siegele, and J. C. Hu, Gene expression from plasmids containing the araBAD promoter at subsaturating inducer concentrations represents mixed populations, Proc Natl Acad Sci U S A 94 (1997) 8168-8172.

[167] A. Khlebnikov, K. A. Datsenko, T. Skaug, B. L. Wanner, and J. D. Keasling, Homogeneous expression of the P(BAD) promoter in *Escherichia coli* by constitutive expression of the low-affinity high-capacity AraE transporter, Microbiology 147 (2001) 3241-3247.

[168] F. M. Adamski, K. K. McCaughan, F. Jorgensen, C. G. Kurland, and W. P. Tate, The concentration of polypeptide chain release factors 1 and 2 at different growth rates of *Escherichia coli*, J Mol Biol 238 (1994) 302-308.

[169] K. A. Datsenko, and B. L. Wanner, One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products, Proc Natl Acad Sci U S A 97 (2000) 6640-6645.

[170] B. L. Wanner, Overlapping and separate controls on the phosphate regulon in *Escherichia coli* K12, J Mol Biol 166 (1983) 283-308.

[171] H. Makinoshima, A. Nishimura, and A. Ishihama, Fractionation of *Escherichia coli* cell populations at different stages during growth transition to stationary phase, Mol Microbiol 43 (2002) 269-279.

[172] M. J. Casadaban, Regulation of the regulatory gene for the arabinose pathway, araC, J Mol Biol 104 (1976) 557-566.

[173] E. Kalef, and C. Gitler, Purification of vicinal dithiol-containing proteins by arsenical-based affinity chromatography, Methods Enzymol 233 (1994) 395-403.

[174] Q. Cheng, L. Johansson, J. O. Thorell, A. Fredriksson, E. Samen, S. Stone-Elander, and E. S. Arnér, Selenolthiol and dithiol C-terminal tetrapeptide motifs for one-step purification and labeling of recombinant proteins produced in E. coli, Chembiochem 7 (2006) 1976-1981.

[175] Q. Cheng, S. Stone-Elander, and E. S. Arnér, Tagging recombinant proteins with a Sel-tag for purification, labeling with electrophilic compounds or radiolabeling with 11C, Nat Protoc 1 (2006) 604-613.

[176] L. Johansson, C. Chen, J. O. Thorell, A. Fredriksson, S. Stone-Elander, G. Gafvelin, and E. S. Arnér, Exploiting the 21st amino acid-purifying and labeling proteins by selenolate targeting, Nat Methods 1 (2004) 61-66.

[177] L. Johansson, G. Gafvelin, and E. S. Arnér, Selenocysteine in proteins-properties and biotechnological use, Biochim Biophys Acta 1726 (2005) 1-13.

[178] G. T. Chen, M. J. Axley, J. Hacia, and M. Inouye, Overproduction of a selenocysteine-containing polypeptide in *Escherichia coli*: the fdhF gene product, Mol Microbiol 6 (1992) 781-785.

[179] J. Heider, and A. Böck, Targeted insertion of selenocysteine into the alpha subunit of formate dehydrogenase from Methanobacterium formicicum, J Bacteriol 174 (1992) 659-663.

[180] R. Koishi, T. Nakamura, T. Takazawa, C. Yoshimura, and N. Serizawa, Production of functional human selenocysteine-containing KDRF/thioredoxin reductase in E. coli, J Biochem 127 (2000) 977-983.

[181] L. Johansson, L. Svensson, U. Bergstrom, G. Jacobsson-Ekman, E. S. Arnér, M. van Hage, A. Bucht, and G. Gafvelin, A mouse model for in vivo tracking of the major dust mite allergen Der p 2 after inhalation, Febs J 272 (2005) 3449-3460.

[182] A. N. Kuntz, E. Davioud-Charvet, A. A. Sayed, L. L. Califf, J. Dessolin, E. S. Arnér, and D. L. Williams, Thioredoxin glutathione reductase from Schistosoma mansoni: an essential parasite enzyme and a key drug target, PLoS Med 4 (2007) e206.

[183] S. Hazebrouck, L. Camoin, Z. Faltin, A. D. Strosberg, and Y. Eshdat, Substituting selenocysteine for catalytic cysteine 41 enhances enzymatic activity of plant phospholipid hydroperoxide glutathione peroxidase expressed in *Escherichia coli*, J Biol Chem 275 (2000) 28715-28721.

[184] Z. Jiang, E. S. Arnér, Y. Mu, L. Johansson, J. Shi, S. Zhao, S. Liu, R. Wang, T. Zhang, G. Yan, J. Liu, J. Shen, and G. Luo, Expression of selenocysteine-containing glutathione S-transferase in *Escherichia coli*, Biochem Biophys Res Commun 321 (2004) 94-101.

[185] P. E. Dawson, T. W. Muir, I. Clark-Lewis, and S. B. Kent, Synthesis of proteins by native chemical ligation, Science 266 (1994) 776-779.

[186] R. Quaderer, A. Sewing, and D. Hilvert, Selenocysteine-Mediated Native Chemical Ligation, Helvetica Chimica Acta 84 (2001) 1197-1206.

[187] H. J. Yu, J. Q. Liu, A. Böck, J. Li, G. M. Luo, and J. C. Shen, Engineering glutathione transferase to a novel glutathione peroxidase mimic with high catalytic efficiency. Incorporation of selenocysteine into a glutathione-binding scaffold using an auxotrophic expression system, J Biol Chem 280 (2005) 11930-11935.

[188] Y. Ge, Z. Qi, Y. Wang, X. Liu, J. Li, J. Xu, J. Liu, V. N. Gladyshev, and J. Shen, Engineered selenium-containing glutaredoxin displays strong glutathione peroxidase activity rivaling natural enzyme, Int J Biochem Cell Biol (2008).

[189] S. Müller, H. Senn, B. Gsell, W. Vetter, C. Baron, and A. Böck, The formation of diselenide bridges in proteins by incorporation of selenocysteine residues: biosynthesis and characterization of (Se)2-thioredoxin, Biochemistry 33 (1994) 3404-3412.

[190] A. Mehta, C. M. Rebsch, S. A. Kinzy, J. E. Fletcher, and P. R. Copeland, Efficiency of mammalian selenocysteine incorporation, J Biol Chem 279 (2004) 37852-37859.

[191] P. Y. Gasdaska, M. M. Berggren, M. J. Berry, and G. Powis, Cloning, sequencing and functional expression of a novel human thioredoxin reductase, FEBS Lett 442 (1999) 105-111.

[192] R. J. Hondal, and R. T. Raines, Semisynthesis of proteins containing selnocysteine, Methods Enzymol 347 (2002) 70-83.

[193] B. Eckenroth, K. Harris, A. A. Turanov, V. N. Gladyshev, R. T. Raines, and R. J. Hondal, Semisynthesis and characterization of mammalian thioredoxin reductase, Biochemistry 45 (2006) 5158-5170.

[194] A. Bacher, S. Eberhardt, M. Fischer, K. Kis, and G. Richter, Biosynthesis of vitamin b2 (riboflavin), Annu Rev Nutr 20 (2000) 153-167.

[195] S. Ma, R. M. Caprioli, K. E. Hill, and R. F. Burk, Loss of selenium from selenoproteins: conversion of selenocysteine to dehydroalanine in vitro, J Am Soc Mass Spectrom 14 (2003) 593-600.

[196] S. N. Gorlatov, and T. C. Stadtman, Human thioredoxin reductase from HeLa cells: selective alkylation of selenocysteine in the protein inhibits enzyme activity and reduction with NADPH influences affinity to heparin, Proc Natl Acad Sci U S A 95 (1998) 8520-8525.

[197] V. N. Gladyshev, K. T. Jeang, and T. C. Stadtman, Selenocysteine, identified as the penultimate C-terminal residue in human T-cell thioredoxin reductase, corresponds to TGA in the human placental gene, Proc Natl Acad Sci U S A 93 (1996) 6146-6151.