Thesis for the Master's degree in Molecular Biosciences

Main field of study in Immunology

Periplasmic targeting of "fluorobodies"

Lise Mirja Øieren

60 study points

Department of Molecular Biosciences

Faculty of Mathematics and Natural Sciences

UNIVERSITY OF OSLO 12/2009

Contents

Acknowledgement

Abbrevations

aa	Amino acid
Ab	Antibody
APC	Antigen presenting cell
E. coli	*Escherichia coli*
FACS	Fluorescence-activated cell sorting
GFP	Green fluorescent protein
Ig	Immunoglobulin
IPTG	Isopropyl-β-D-thio-galactopyranoside
mAb	Monoclonal Ab
ON	Overnight
PCC	Protein conducting channel
phOx	2-phenyloxazol-5-one
pMHC	peptide major histocompatibility complex
PPIase	Peptidylprolyl *cis/trans* isomerase
RE	Restriction endonuclease
RFU	Relative fluorescence units
RT	Room temperature
scFv	Single chain fragment variable
scTCR	Single chain TCR
Sec	Secretory
SFS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SRP	Signal recognition particle
Tat	Twin arginine transporter
TCR	T cell receptor
TF	Trigger factor
YFP	Yellow fluorescent protein

Abstract

Molecular evolution is a key strategy in the development of novel diagnostics and therapeutics. Libraries of mutant variants are generated and those with desired properties selected after screening. However, large scale screening is expensive and time-consuming and has become a bottleneck. The fluorescent properties of the green fluorescent protein (GFP) are used in quantitative and qualitative screening. An attractive approach is to take advantage of the fact that a correlation exists between the solubility of an upstream fusion partner to GFP and observed fluorescence when the fusion protein is expressed in *Escherichia coli* (*E. coli*).

The T-cell receptor (TCR) is a molecule with therapeutic and diagnostic potential. However, the single chain TCR (scTCR) shows toxic and aggregation-prone characteristics that may be alleviated by molecular evolution. When expressed in *E. coli*, scTCR and antibody (Ab) based molecules such as single chain fragment variable (scFv), need the oxidizing environment of the periplasm to form disulfide bonds and obtain a functional fold. The two scFvs, with specificity for two haptens phOx and NIP, are expressed in the periplasm at high and moderate levels, respectively. In this study, the scTCR as well as the NIP and phOx binders were genetically fused to GFP and three different mechanisms for translocation from the cytosol to the periplasm were compared. The results show that the two scFv fusions with pelB leader were transported to the periplasm, but cleaved in the linker region. The same result was obtained for anti-phOx-GFP with TorA leader. Neither over-expression of the chaperones GroEL-GroES nor FkpA, expression in host strains compensating for codon bias or lacking proteases, expression without induction of fusion proteins nor reducing the expression temperature to 16°C resulted in expression of full length fusion protein in the periplasm. The scTCR-GFP fusion was not transported to the periplasm with the use of neither pelB, DsbA nor TorA leader. Neither over-expression of the chaperones GroEL-GroES or FkpA, nor expression in a host strain compensating for codon bias resulted in periplasmic yield of fusion protein.

In conclusion, the expression profiles of the scFvs were not preserved when expressed as fusions to GFP. The scTCR-GFP fusions were not transported to the periplasm, regardless of leader sequence. Further optimization of expression system or construct design is needed to obtain a screening platform linking functional fold to observed fluorescence.

1 Introduction

1.1 Recombinant protein expression

1.1.1 General aspects

Recombinant protein expression is a necessity for production of proteins for research, diagnostics and therapeutics. The main goals of most recombinant protein expressions are to be fast, inexpensive, easy to purify and preferably produced in the milligram per liter range (Graslund, Nordlund et al. 2008). Many diverse strategies have been tested for tens of thousands of different proteins from a variety of classes, and there seems to be no expression strategy suitable for all proteins (Graslund, Nordlund et al. 2008). Therefore, the expression strategy for each individual protein has to be carefully chosen based on the properties and the downstream applications of the protein. In addition, the strategy might have to be optimized in order to obtain a product or maximize the functional yield where that is a goal.

When choosing an expression host for recombinant protein production, *Escherichia coli* (*E. coli*) usually stands out as the first choice because it is cheap and easy to grow, has an extensively studied genome, can be grown to high densities and has a very short generation time (Terpe 2006). It is therefore possible to test and optimize a variety of strategies in *E. coli* fast and inexpensively. *E. coli* has been successfully used for expression of recombinant proteins from a variety of classes and diverse species, ranging from prokaryotes to humans (Graslund, Nordlund et al. 2008). Generally, heterologous proteins are well suited for *E. coli* expression if they are expressed intracellularly and are soluble in their native context. However, expression of heterologous proteins in *E. coli* becomes more challenging when they are normally expressed extracellularly, contain disulfide bonds, are glycosylated or form hetero-multimers (Baneyx and Mujacic 2004). Expression of eukaryotic proteins with posttranslational modifications such as N-glycosylation, has proven difficult as *E. coli* has not been showen to perform this task naturally. However, recent identification of prokaryotic N-glycosylation systems and the transfer of these to *E. coli* are opening up the new field of glycoengineering with expression of recombinant glycoproteins (Langdon, Cuccui et al. 2009). Another important feature to consider when expressing recombinant proteins in *E. coli* is the reducing cytoplasm in contrast to the oxidizing periplasm that facilitates disulfide bond formation. Hence, expression of recombinant proteins that require disulfide bonds to obtain a

functional fold is either targeted to the periplasm by one of the secretory routes or expressed in a thioredoxin reductase (*trxB*) mutant *E. coli* strain that facilitates disulfide bond formation in the cytoplasm (Baneyx and Mujacic 2004; de Marco 2009). Furthermore, a large number of mutant *E. coli* strains are optimized for recombinant protein expression and are commercially available as shown in Table 1 (Terpe 2006).

Table 1: Commercially available *E. coli* strains used for heterologous protein expression and their key features. The table is adapted from (Terpe 2006).

E. coli strain	Derivation	Key features
		Some *E. coli* strains most frequently used for heterologous protein production and their key features
AD494	K-12	*trxB* mutant; facilitates cytoplasmic disulfide bond formation
BL21	B834	Deficient in *lon* and *ompT* proteases
BL21 trxB	BL21	*trxB* mutant; facilitates cytoplasmic disulfide bond formation; deficient in *lon* and *ompT* proteases
BL21 CodonPlus-RIL	BL21	Enhances the expression of eukaryotic proteins that contain codons rarely used in *E. coli*: AGG, AGA, AUA, CUA; deficient in *lon* and *ompT* proteases.
BL21 CodonPlus-RP	BL21	Enhances the expression of eukaryotic proteins that contain codons rarely used in *E. coli*: AGG, AGA, CCC; deficient in *lon* and *ompT* proteases.
BLR	BL21	*recA* mutant; stabilizes tandem repeats; deficient in *lon* and *ompT* proteases
B834	B strain	Met auxotroph; ^{35}S-met labeling
C41	BL21	Mutant designed for expression of membrane proteins
C43	BL21	Double mutant designed for expression of membrane proteins
HMS174	K-12	*recA* mutant; Rif resistance
JM 83	K-12	Usable for secretion of recombinant proteins into the periplasm
Origami	K-12	*trxB/gor* mutant; greatly facilitates cytoplasmic disulfide bond formation
Origami B	BL21	*trxB/gor* mutant; greatly facilitates cytoplasmic disulfide bond formation; deficient in *lon* and *ompT* proteases
Rosetta	BL21	Enhances the expression of eukaryotic proteins that contain codons rarely used in *E. coli*: AUA, AGG, AGA, CGG, CUA, CCC, and GGA; deficient in *lon* and *ompT* proteases
Rosetta-gami	BL21	Enhances the expression of eukaryotic proteins that contain codons rarely used in *E. coli*: AUA, AGG, AGA, CGG, CUA, CCC, and GGA; deficient in *lon* and *ompT* proteases; *trxB/gor* mutant; greatly facilitates cytoplasmic disulfide bond formation

Most strains are also available as DE3 and DE3 pLysS strains. Strains are commercially available from different manufacturers

1.1.2 High throughput screening

In development of new diagnostics and therapeutics, of the major bottlenecks in the identification of new or improved molecules is the single clone screening of pre-selected candidates. The need for efficient high quantity and quality screening forms the basis for high throughput screening as field and during the past 15 years, great improvements yielding increased performance have been achieved. One of the first issues that were addressed was to increase the screening capacity by using robotics and thereby automating large parts of the screening process. This makes the screening less laborious and much faster than manual screening. There has also been a simultaneous down-scaling in sample size to make the automation more resource effective. The use of microtiter plates reflects this development as many groups have gone from using 96 wells plates to the 384 and 1536 wells plate formats. There are examples of groups using 3456 wells microtiter plates (Brandish PE 2006), but the

small reaction volume (1-2 µl) and logistics makes them more challenging to use. Therefore, more recently there has been a turn in focus towards improving the quality of high throughput screening by optimizing more robust assays, reducing the false positives and false negatives fractions, improving other statistical parameters such as the Z'-factor (Zahng JH 1999) and increasing the sensitivity. According to Mayr and Bojanic (Mayr 2009) the three key factors influencing a successful high throughput screening is time, cost and quality. All three factors are connected and must be considered according to each other for each screening to achieve an optimal output balanced between quantity and quality.

Traditionally, high-throughput screening has been employed in classical drug discovery (e.g. combinatorial chemistry). However, the rise of functional proteomics has re-focused both academia and the industry through the last decade by embracing the next generation biotechnological solutions based on recombinant proteins. The challenge of tomorrow is then to accommodate next generation therapeutic and diagnostic scaffolds (Wu, Adams et al. 2008) to suitable screening platforms, preferably preserving the high throughput features.

A schaffold which has been extensively studied is the monoclonal antibody (mAb) and it has already resulted in approved therapeutics (Carter 2006). The mAb can be engineered to bind to almost anything, but there are still classes of ligands that are not effectively targeted by mAbs. One such class of ligands is the peptide/major histocompatibility complex (pMHC) which represents a key element in adaptive immunity. The natural receptor of the pMHC is the T-cell receptor (TCR) and it represents a molecule with great therapeutic potential (Molloy, Sewell et al. 2005). However, TCRs have proven difficult to express soluble due to the toxic and aggregation prone characteristics. Therefore, a screening platform is necessary for identification of TCR clones with improved characteristics.

1.1.3 Protein folding

Folding of recombinant proteins is a crucial step to obtain soluble and functional protein. In the cytoplasm of *E. coli*, a polypeptide can receive folding assistance from several molecular chaperones as seen in Figure 1. The general mechanism of molecular chaperones is to bind to nonpolar segments of polypeptides to shield the proteins from aggregating and stabilize intermediate folding conformations (Baneyx and Mujacic 2004). Trigger factor (TF) or DnaK-DnaJ are the first chaperones to bind the nascent polypeptide in need of folding

assistance as it emerges from the ribosome (Deuerling, Patzelt et al. 2003). After the release from TF or DnaK-DnaJ, the polypeptide can fold into its native conformation or halt at an intermediate state. In the latter case, the polypeptide can cycle between DnaK-DnaJ bound and unbound state, reaching new intermediates and finally the native conformation or receive aid from GroEL-GroES (Ewalt, Hendrick et al. 1997). The folding chaperones GroEL-GroES consist of a large chamber where the polypeptide is shielded from the environment and can reach intermediate conformations without aggregating (Hartl and Hayer-Hartl 2002). The cell is also equipped with holding chaperones (e. g. IbpA/B, Hsp31, Hsp33) that keep polypeptides unfolded and prevent aggregation until they can bind a folding chaperone. If aggregates are formed, they can be solubilized by disaggregation chaperones (e. g. ClpB) (Baneyx and Mujacic 2004). The chaperone level is often upregulated as a result of stress in the cell to prevent misfolding and promote refolding of proteins (Baneyx and Mujacic 2004). Overexpression of recombinant proteins in need of folding assistance induces stress on the cell and the folding machinery easily becomes overloaded, often resulting in inclusion body formation and proteolytic degradation. A well documented strategy to increase the folding capacity of the cell and overcome these problems is to over-express TF, DanK-DnaJ or GroEL-GroES, either alone or in combination (Nishihara, Kanemori et al. 2000; Baneyx and Palumbo 2003).

Also the periplasmic space contains chaperones which aid proteins reach their native fold, as seen in Figure 2. The main substrates of the folding chaperone Skp are outer membrane proteins which are assisted in the folding and membrane insertion processes (Chen and Henning 1996; Schafer, Beck et al. 1999; Walton and Sousa 2004). FkpA is a peptidylprolyl *cis/trans* isomerase (PPIase) which also acts as a periplasmic folding chaperone with broad substrate specificity (Missiakas, Betton et al. 1996; Arie, Sassoon et al. 2001). Previous studies have shown increased soluble periplasmic yields of secreted recombinant proteins when overexpressing Skp (Bothmann and Pluckthun 1998) or FkpA (Bothmann and Pluckthun 2000; Arie, Sassoon et al. 2001; Gunnarsen, Lunde et al. submitted). In addition to these two generic chaperones, the periplasm contains specialized chaperones (e.g. SurA, LolA, PapD, FimC), PPIases (e. g. SurA, PpiD, PpiA) and proteins involved in disulfide bond formation (e. g. the Dsb-family) (Baneyx and Mujacic 2004).
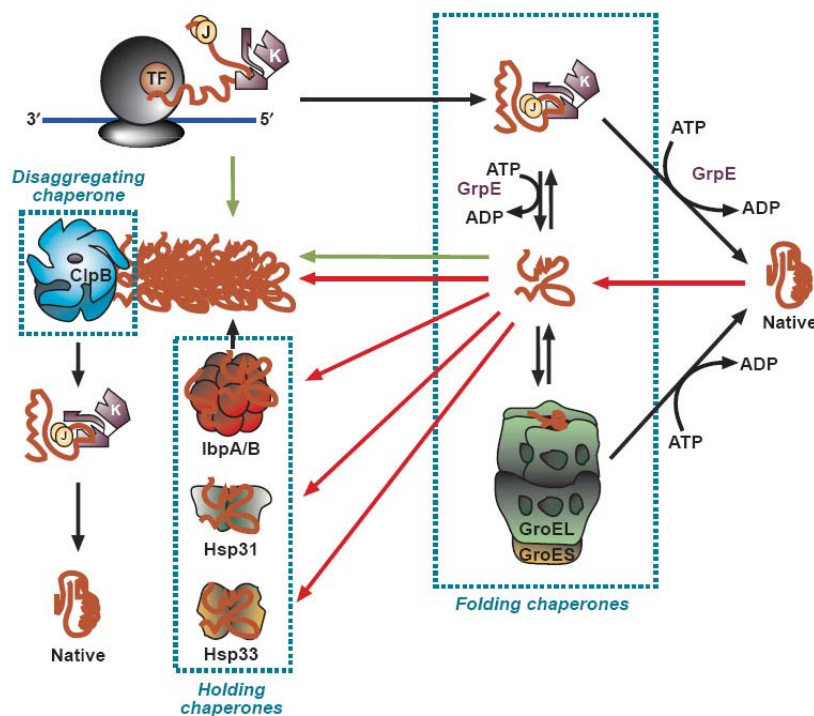
Figure 1: Schematic overview of cytoplasmic chaperones. Conventional folding routes for nascent polypeptides are indicated by black arrows. Unfolding and aggregation of thermolabile proteins induced by stress is indicated by red arrows. The green arrows indicate alternative routes for recombinant proteins that fail in the classical folding pathway. This figure is adapted from (Baneyx and Mujacic 2004).

1.1.4 Secretory pathways

Many proteins serve their function outside the cytoplasm and need to be inserted into a membrane, transported to the periplasm or secreted to the extracellular space. The main routes of transport to the periplasm are the general secretion pathway (Sec-pathway) (Driessen, Manting et al. 2001) and the twin arginine transporter pathway (Tat-pathway) (Berks, Palmer et al. 2005). In addition, there are more specialized pathways such as the YidC dependent pathway which is dedicated to insertion and folding of membrane proteins (Kol, Nouwen et al. 2008) and the type I-V protein secretion systems which mainly function in the translocation and secretion of proteins across the outer membrane in gram-negative bacteria such as *E. coli* (Henderson, Navarro-Garcia et al. 2004; Christie, Atmakuri et al. 2005; Backert and Meyer 2006; Cornelis 2006; Johnson, Abendroth et al. 2006). The main focus here will be on translocation of proteins to the periplasm, hence the YidC dependent pathway and the type I-V protein secretion systems will not be discussed further.

Proteins targeted for translocation to the periplasm contain an *N*-terminal signal sequence, also named leader peptide. Proteins containing such a signal sequence are preproteins and become a mature protein upon leader peptide cleavage by a signal peptidase after translocation (Dev and Ray 1990; Economou 1999). Signal sequences seem to have some general features and contain a positively charged *n*-region, a central hydrophobic *h*-region, and a *c*-region which is more polar than the *h*-region (von Heijne 1985), but still there is great variety in the known functional signal sequences. However, predictions of protein transport can be performed using bioinformatics tools, e.g. SignalP, TatP or SecretomeP (Center for biological sequence analysis, Technical university of Denmark), PSORTb (Brinkman Laboratory, Simon Fraser University) or PPSearch (European Bioinformatics Institute), based on computational algorithms such as artificial neural networks and hidden Markov models (Gardy, Laird et al. 2005; Emanuelsson, Brunak et al. 2007).
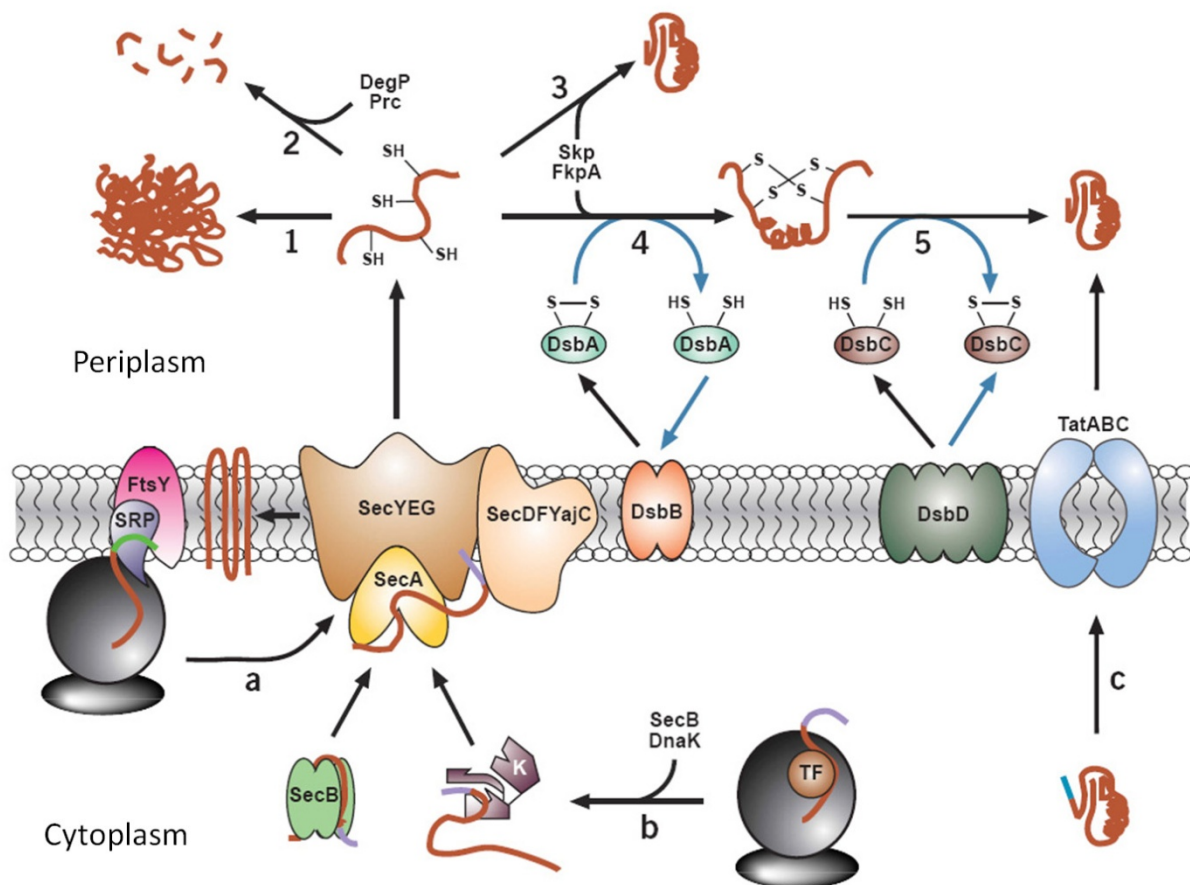

Figure 2: Overview of secretory pathways and periplasmic folding. (a) indicates the SRP-dependent Sec-pathway. (b) indicates the SecB/conventional Sec-pathway. (c) indicates the Tat-pathway. Folding pathways, degradation and aggregation are indicated in the periplasm. This figure is adapted from (Baneyx and Mujacic 2004).

The Sec-pathway

The Sec-pathway translocates unfolded polypeptides to the periplasm by two main targeting routes; the co-translational signal recognition particle (SRP)-dependent route or the post-translational SecB-dependent route (Driessen, Manting et al. 2001). Both routes utilizes the protein conducting channel (PCC) consisting of the heterotrimeric integral membrane protein complex SecYEG, and the peripheral associated ATPase SecA, a molecular motor which drives the translocation of the targeted proteins (de Keyzer, van der Does et al. 2003). SecA has been shown to bind both the signal sequence and mature regions of the preprotein (Papanikou, Karamanou et al. 2005). In the co-translational SRP-dependent pathway, the GTP-driven elongation of the polypeptide chain by the ribosome seems to be the main translocation force while translocation in the SecB-dependent pathway is powered by cycles of ATP hydrolysis and conformational changes of SecA (Natale, Bruser et al. 2008).

For the post-translational route, the emerging polypeptide is bound by TF as it emerges from the ribosome during synthesis (Buskiewicz, Deuerling et al. 2004). When the chain is released from the ribosome, it is bound by SecB which keeps the peptide unfolded and targets it to the Sec machinery (Driessen 2001). SecB then transfers the unfolded polypeptide chain to SecA (Driessen 2001). The polypeptide is threaded through the channel formed by SecYEG driven by energy from the proton motive force captured by SecDFYajC. In the translocation process the signal sequence is cleaved off by signal peptidases. Many different signal sequences have been used for periplasmic heterologous expression, e. g. OmpA, PelB and PhoA (Thie, Schirrmann et al. 2008).

The SRP-dependent pathway translocates co-translationally and is involved in both protein transport to the periplasm and insertion of proteins in the inner membrane. The SRP binds to the *N*-terminal signal sequence of the polypeptide as it emerges from the ribosome (Luirink, von Heijne et al. 2005). The ribosome/emerging-polypeptide/SRP complex then associates with the membrane-bound receptor FtsY. By GTP hydrolysis, the ribosome with the nascent peptide chain is transferred to SecYEG where the polypeptide chain is translocated to the periplasm (Driessen, Manting et al. 2001). Different leader sequences have been used for periplasmic heterologous expression, e. g. DsbA, TolB and TorT (Thie, Schirrmann et al. 2008).

Tat-pathway

The twin arginine translocator (Tat) differs from the SecB- and SRP- dependent pathways as it mainly translocates folded or partially folded proteins (DeLisa, Samuelson et al. 2002; Robinson and Bolhuis 2004). This translocation mechanism is thought to be important for the translocation of proteins that bind cofactors or fold fast and efficiently in the cytoplasm (Berks 1996; Berks, Palmer et al. 2005). The Tat-pathway was named because of the twin arginine motif found in signal peptides directing translocation through this transporter (Berks 1996). This twin arginine motif is highly conserved even though studies have revealed translocation despite substitution of one arginine with lysine (Hinsley, Stanley et al. 2001). Prior to translocation, the polypeptide is expected to fold and assemble with cofactors, if required. Polypeptide folding is aided by the folding chaperones previously described and in some cases it is also believed to be a quality control to prevent translocation of precursor protein, e. g. by the binding of a protein to the signal sequence to shield it from premature interactions with the translocation machinery (Pommier, Mejean et al. 1998; Jack, Buchanan et al. 2004). Although the translocation mechanism of the Tat-pathway is not fully understood, TatA, TatB and TatC have been identified as the essential translocation components (Bolhuis, Mathers et al. 2001). The primary signal sequence recognition occurs with TatC (Alami, Luke et al. 2003) and is followed by translocation of the Tat-substrate through a channel formed in the membrane with energy provided by the proton motive force (Natale, Bruser et al. 2008).

The Tat signal peptides are usually between 25 and 50 amino acids (aa) and are therefore on average longer than the signal peptides directing translocation through the Sec-pathway (Bruser 2007). A wide variety of recombinant proteins have been successfully targeted to periplasm via the Tat-transporter using the TorA signal sequence (DeLisa, Samuelson et al. 2002; Blaudeck, Kreutzenbeck et al. 2003; DeLisa, Tullman et al. 2003; Thammawong, Kasinrerk et al. 2006).

1.1.5 Quality control/proteolysis

Proteins unable to reach their native conformation are degraded by proteases to prevent accumulation of potentially deleterious peptides and to recycle aa residues. The proteases Lon, ClpYQ/HslUV, ClpAP, ClpXP and FtsH unfolds and degrades substrate polypeptide chains with the aid of peptidases in the cytoplasm of *E. coli* (Baneyx and Mujacic 2004). The

periplasm also contains proteases that impede aggregate formation. The primary periplasmic protease is considered to be DegP and is supplemented by the additional proteases DegS, DegQ, Protease III and OmpT (Baneyx and Mujacic 2004).

1.2 The T-cell receptor

The interaction between T cells and specialized antigen presenting cells (APCs) is one of the key interactions in both thymic selection and development of adaptive immunity. The specificity of this interaction is provided by the TCR on the T cell membrane and the pMHC on the surface of APCs. Production of soluble MHC tetramers has made it possible to track the course of an immune response in a specific manner. The counter approach, namely to specifically track specific antigen persentatino on an APC still remains elusive mainly due to technical challenges in soluble TCR expression. In addition, soluble TCRs have a great potential as therapeutics and diagnostics (Molloy, Sewell et al. 2005).

TCR is a glycosylated heterodimer comprised of an α- and a β-chain or a γ- and a δ-chain. Only a small subset of TCRs is composed of γ- and δ-chains. The main focus in this study will be on the αβ TCR. Each chain is comprised of two extracellular domains, namely a variable domain and a constant domain, in addition to a transmembrane segment. These domains contain inter-domain disulfide bonds that are important to obtain a functional fold. In addition, the two chains are covalently connected by a membrane proximal disulfide bond. A schematic view of the αβ TCR is shown in Figure 3.

The TCR interacts with the pMHC through the variable domains. Therefore, expression of soluble TCRs has been performed by genetically connecting the variable domains by a linker (Maynard, Adams et al. 2005; Gunnarsen, Lunde et al. submitted). This creates a single chain TCR (scTCR) molecule that is translated as one polypeptide that retains the pMHC specificity. The immunoglobulin (Ig) fold of the TCR variable domains shows a strong structural resemblance to the variable domains of antibodies (Abs) (Halaby, Poupon et al. 1999). Therefore, expression of the Ab based single chain fragment variable (scFv) molecules and scTCRs face many of the same challenges. Targeting expression of scFvs to the periplasm of *E. coli* facilitates disulfide bond formation and is a very successful strategy. However, functional expression of scTCRs is dependent of ectopic folding assistance to

obtain soluble periplasmic protein (Wulfing and Pluckthun 1994; Maynard, Adams et al. 2005; Gunnarsen, Lunde et al. submitted).
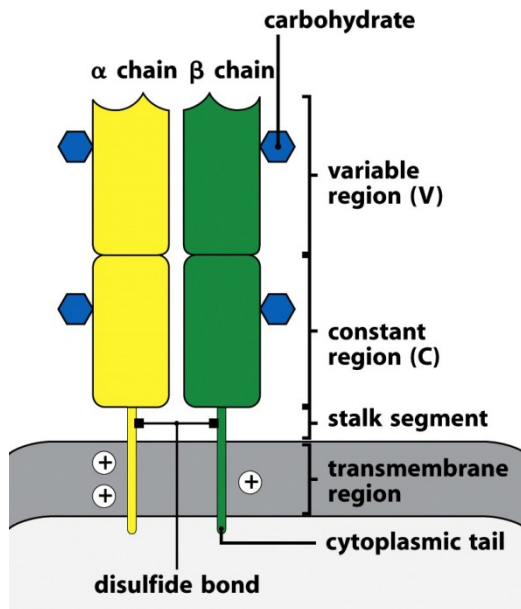
Figure 3: A schematic view of the αβ TCR structure. The TCR is a heterodimer composed of two glycosylated chains, the α- and β-chains. Each chain is composed of a variable domain, constant domain and a transmembreane segment. The two chains are covalently connected by a membrane proximal disulfide bond. The figure is adapted from Immunobiology 7th edition (Murphy, Travers et al. 2008).

The scTCR characteristic of the T-cell clone 4B2A1 from the murine MOPC315 myeloma model has been chosen to serve as a fusion partner to the green fluorescent protein (GFP). This TCR has been displayed on phage and expressed as a soluble entity, but as most TCRs it is toxic to the bacterial host cell and aggregation-prone. However, molecular evolution of the TCR is expected to improve functional expression.

1.3 GFP as a folding reporter

1.3.1 GFP

The green fluorescent protein (GFP) was first cloned from the jellyfish *Aequorea victoria* in 1992 (Prasher, Eckenrode et al. 1992). Since then, GFP has been subject to extensive mutagenesis and new clones have been identified with altered characteristics, in addition to the cloning and characterization of fluorescent proteins from several other species (reviewed in (Zimmer 2002; Shaner, Patterson et al. 2007)). The three-dimensional structure of GFP and

its derivatives is a beta-barrel consisting of 11 anti-parallel strands and a coaxial helix as shown in Figure 4 (Ormo, Cubitt et al. 1996). The fluorescence is due to the formation of an optically active chromophore with the key aas Ser65-Tyr66-Gly67 (*wild type*), which gives the major excitation and emission peaks at 395nm and 509nm, respectively. The isolation of a fluorescence-activated cell sorting (FACS) -optimized mutant, *GFPmut2*, with S65A, V68L and S72A substitutions gave a GFP with a red-shifted excitation maximum at 488nm, improved folding characteristics and increased solubility in *E. coli* (Cormack, Valdivia et al. 1996). Simultaneously, the "enhanced" GFP, *cycle 3*, was isolated with F99S, M153T and V163A mutations which was shown to reduce aggregation and increase chromophore activation when expressed in *E. coli* (Crameri, Whitehorn et al. 1996). Furthermore, an A206K mutation has been shown to reduce GFP's tendency to dimerize and produces a more monomeric GFP (Zacharias, Violin et al. 2002). Of particular interest are the folding reporter GFP (Waldo, Standish et al. 1999) and the superfolder GFP (Pedelacq, Cabantous et al. 2006). The folding reporter GFP was fused *C*-terminally to 20 proteins that had previously been expressed alone with different degrees of solubility. Cells expressing these fusion proteins showed a strong correlation between the fluorescence measured and the solubility of the upstream fusion partner. In contrast, the superfolder GFP has additional mutations giving greater resistance to chemical denaturants and improved folding kinetics. Therefore, this superfolder GFP folds well and fluoresce even when it is fused to poorly folded peptides.
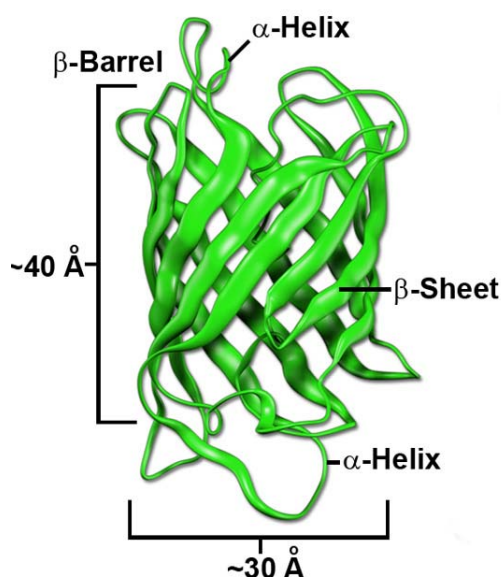

Figure 4: The structure and approximate dimensions of *Aequorea victoria* derivated fluorescent proteins. The figure is adapted from (Shaner, Patterson et al. 2007).

1.3.2 GFP fusion proteins

Many proteins are difficult to express in soluble form in *E. coli*. Different approaches to increase functional expression include over-expression of chaperones, altered expression conditions, targeting proteins for expression in the periplasm and expression in specialized host strains (Baneyx and Mujacic 2004; Graslund, Nordlund et al. 2008). Fusion of proteins to more soluble partners have also been shown to increase expression levels (Makrides 1996) and even altered the subcellular localization (Zhang, Brokx et al. 2006). However, these strategies do not alter the intrinsic sequence dependent folding characteristics of a protein and will not work for all. A putative strategy is therefore to fuse the protein of interest to a folding reporter protein which gives information on the folding status.

Protein structure is strongly correlated to the biological activity and misfolding of proteins can lead to loss of activity, aggregation, proteolysis and even cell death (Baneyx and Mujacic 2004). The most relevant approach to measure folding quality during protein expression is therefore to asses the biological activity. However, such assays often require a laborious experimental protocol. Therefore, fusing the protein of interest to GFP is a solution since the measured fluorescence have been shown to correlate with the folding ability of the protein (Waldo, Standish et al. 1999). This strategy has proven applicable as several studies have successfully used GFP as an optical folding reporter (Waldo, Standish et al. 1999; Nakayama and Ohara 2003; Wang and Chong 2003; Omoya, Kato et al. 2004).

Several attempts have been made to fuse GFP to scFv (Griep, van Twisk et al. 1999; Casey, Coley et al. 2000; Schwalbach, Sibler et al. 2000). The scFv requires intradomain disulfide bond formation to obtain a functional fold. Therefore, attempts were made to target the fusion protein for periplasmic expression. However, utilizing the Sec-pathway with a pelB leader yields were still low (Griep, van Twisk et al. 1999; Casey, Coley et al. 2000). Further approaches such as alternative leader sequences, over-expression of chaperones, expression in specialized host stains and altered expression conditions could improve periplasmic expression and functionality.

2 Aim of the project

As highlighted in the introduction, there is a need for a high throughput screening system to evaluate whether protein variants have obtained a functional fold during molecular evolution. When used as a protein-of-interest fusion partner, GFP may serve as an indicator of correct folding, since the intensity of fluorescence emitted from host cells producing fusion protein correlates with functional folding of the protein-of-interest.

The study was set up to investigate the feasibility for using GFP as fusion partner for scFv and scTCR variants, all of which contain Ig-domains that require the formation of an intradomain disulfide bridge to reach a functional fold. Such a bridge may be formed in the oxidizing environment of the *E.coli* periplasmic compartment. One scTCR and two different scFvs were chosen because they exhibit three distinct and different periplasmic expression profiles in *E. coli,* ranging from poor to very good.

The aim was separated into two sub-goals:

1. To evaluate the periplasmic expression profile of the three model proteins as fusions to GFP. A major issue was to investigate whether the GFP fusions retained the expression profiles earlier observed for "naked" scFvs and the scTCR. Three translocation mechanisms for periplasmic targeting were evaluated using both the Sec- and Tat-pathways for transport from cytoplasm to periplasm.

2. To develop the GFP-based Ig-fold tracker such that sufficient quantities of functional fusion protein were produced to allow downstream purification and use.

   To optimize periplasmic yields, the effect of several expression parameters were investigated, such as i) over-expression of the chaperones GroEL-GroES or FkpA, ii) expression in host strains that compensate for codon bias or lacking proteases, iii) expression without induction and iv) reducing the expression temperature to 16°C.

3 Materials and methods

3.1 General materials

3.1.1 Primers and other oligonucleotides

Synthetic oligonucleotides used for cloining in this study were purchased fromMWG-Biotech (Ebersberg, Germany) and are shown in Table 2.

Table 2: Over view of primers and other oligonucleotides used in this study.

Oligonucleotide	Sequence
TorAB7_EcoRI_new	5'-ATAGAATTCATTAAAGAGGAGAAATTAACCATGAACAATAACGATC TCTTTCAGAC-3'
TorAB7_1NcoI	5'-GGCCATGGCCGCAGTCGCACGTC-3'
TorAB7_2NcoI	5'- GGCCATGGCCGCTTGCGCCGCAGTCGCACGTC-3'
DsbA_1_Plus	5'-AATTCATTAAAGAGGAGAAATTAACCATGAAAAAGATTTGGCTGGCG CTGGCTGGTTTAGTTTTAGCGTTTAGCGC-3'
DsbA_1_Minus	5'-CATGGCGCTAAACGCTAAAACTAAACCAGCCAGCGCCAGCCAAATCT TTTCATGGTTAATTTCTCCTCTTTAATG-3'
DsbA_2_Plus	5'- AATTCATTAAAGAGGAGAAATTAACCATGAAAAAGATTTGGCTGGC GCTGGCTGGTTTAGTTTTAGCGTTTAGCGCATCGGC-3'
DsbA_2_Minus	5'-CATGGCCGATGCGCTAAACGCTAAAACTAAACCAGCCAGCGCCAGCC AAATCTTTTTCATGGTTAATTTCTCCTCTTTAATG-3'
DsbA_3_Plus	5'- AATTCATTAAAGAGGAGAAATTAACCATGAAAAAGATTTGGCTGG CG CTGGCTGGTTTAGTTTTAGCGTTTAGCGCATCGGCGGC-3'
DsbA_3_Minus	5'-CATGGCCGCCGATGCGCTAAACGCTAAAACTAAACCAGCCAGCGCC AGCCAAATCTTTTTCATGGTTAATTTCTCCTCTTTAATG-3'
pGB7NotI_fw	5'-ATATGCGGCCGCTGGCAGTAAAGGAGAAGAACTTTTCACTG-3'
pGB7BamHI_rev	5'- ATATGGATCCTTTGTATAGTTCATCCATGCCATG -3'
a1159g	5'- GGAAAACTACCTGTTCCGTGGCCAACACTTGTCAC-3'
g1607a_c1608a_c1609a	5'-AACCATTACCTGTCCACACAATCTAAACTTTCGAAAGATCCCAACGA AAAG-3'

Table 3: Overview of *E. coli* strains for protein expression.

E. Coli strain	Genotype	Antibiotics resistance
XL-1 Blue	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F' *proAB lacI^{q}ZΔM15* Tn*10* (Tet^r)]	Tet^r
XL-1 Blue pGro7	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F' *proAB lacI^{q}ZΔM15* Tn*10* (Tet^r)] pGro7	Cam^r, Tet^r
RosettaBlue^{TM}(DE3) pLysS	*endA1 hsdR17* (r$_{K12}^-$ m$_{K12}^+$) *supE44 thi-1 recA1 gyrA96 relA1 lac* (DE3) F'[*proA$^+$B$^+$ lacI^{q}ZΔM15*::Tn*10*] LysSRARE (Cam^R, Tet^R)	Cam^r, Tet^r
BL21 (DE3)	F$^-$ *ompT hsdS$_B$*(r$_B^-$m$_B^-$) *gal dcm* (DE3)	

3.1.2 Bacterial strains

Four different *E. coli* strains were used in this study and are shown in Table 3. XL-1 Blue was obtained in house and used for protein expression and cloning. XL-1 Blue pGro7 and BL21 (DE3) were obtained in house and used for protein expression. RosettaBlue™(DE3) pLysS (Darmstadt, Germany) was purchased from Novagen and used for protein expression.

3.1.3 Antibodies and other reagents

Ab and other reagents used in immunodetection were anti-GFP Ab (Abcam), SA-HRP (Amersham), anti-His-6-tag (C-term) Ab (Invitrogen), anti-mouse IgG-HRP Ab from sheep (Amersham), rabbit anti-GroEL Ab (Sigma), donkey anti-rabbit IgG Ab(Amersham), anti-Flag M2-HRP Ab (Sigma), anti-human lambda L-chain Ab (Dako) and donkey anti-rabbit IgG HRP Ab (Amersham).

All restriction endonucleases (RE) were purchased from New England Biolabs (NEB) (Ipswich, MA, USA). T4 DNA Ligases used in this study were purchased from NEB or Roche. The DNA modifying enzyme T4 Polynucleotide Kinase was purchased from NEB (Ipswich, MA, USA). Recombinant GFP (rGFP) was purchased from Clontech.

3.1.4 Description of parental plasmids

The pHOG21 derivatives (PMID: 9005945) harboring either a scFv anti-phOx (PMID: 1368267), or scFv H3 (*unpublished*) were kind gifts from Affitech AS (Oslo, Norway). In pHOG21 scFv H3, the scFv is found as an *N*-terminal fusion to yellow fluorescent protein (YFP) (Ole H. Brekke, *personal communication*). The pHOG21 derivative pSG1 harboring a scFv anti-NIP has been made *in-house* and described previously (S. Granum, M.Sc. 2000). The scTCR Vαβ 4B2A1 has been described previously (PMID: 17925331). The pFKPDN plasmid containing this scTCR has been described previously (Gunnarsen, K.S., *et al.*, In press). The pFKPEN is a pHOG-Dummy (PMID: 15914190) derivative containing a *fkpA* expression cassette as described for its phagemid counterpart pFKPDN (PIMD: 17925331). The pGALD7 plasmid (Løset, G.Å., *unpublished*) has been made *in-house* and contains a modified T7 terminator sequence. The pB7G plasmid harboring a GFPmut2 mutant (PMID: 8707053) and a *gain-of-function* mutant (B7) of the TorA signal sequence (PMID: 12021272)

was a kind gift from Dr. G. Georgiou (Department of Chemical Engineering, Institute for Cell and Molecular Biology, University of Texas, USA).

The scFvs anti-phOx and anti-NIP, as well as the scTCR Vαβ 4B2A1, were exchanged with the scFv H3 in the YFP-containing pHOG21 vector; thereby creating the plasmids pYFP-phOx, pYFP-NIP and pYFP-scTCR, respectively. The scFv anti-phOx and scTCR Vαβ 4B2A1 fused to YFP from these new plasmids were then inserted to the pFKPEN vector, thereby creating the pYFPFN-phOx and pYFPFN-scTCR, respectively. The construction of the pYFP and pYFPFN vectors described above was conducted by Dr. G.Å. Løset (*unpublished*).

3.2 Cloning methods

3.2.1 Purification of plasmid DNA

Purification of plasmid DNA was performed using Wizard® *Plus* SV Minipreps DNA Purification System (Promega) following the Centrifugation Protocol. Plasmid DNA was eluted from the spin column using destilled H_2O (dH_2O). DNA concentrations were estimated using NanoDrop ND-1000 Spectrophotometer or by Agarose gel Electrophoresis.

3.2.2 Precipitation of DNA

Precipitation of DNA was performed with seeDNA (Amersham Biosciences) or Pellet Paint Co-Precipitant (Novagen), using the manufacturer's protocol.

3.2.3 Restriction endonuclease digestion of DNA

RE digestion of DNA was performed using RE or High Fidelity (HF) RE from NEB. The REs were used with the appropriate reaction conditions described by NEB. The RE digestions were performed under optimal conditions for 1.5 hours or overnight (ON), depending on the subsequent application.

3.2.4 DNA modifying enzymes

To allow ligation of annealed oligos, 5'-phosphate was added by T4 Polynucleotide Kinase (NEB) with reaction conditions and procedure according to the manufacturer's protocol.

3.2.5 Agarose gel electrophoresis

Separation of DNA fragments was performed by agarose gel electrophoresis. The gels contained 0.8-3.0% agarose (Lonza) dissolved in 1xTAE buffer (40 mM Tris-acetate, 1mM EDTA) and 1μg/ml ethidium bromide. 1 kb DNA Ladder (NEB) or 100 bp DNA Ladder (NEB) was prepared and run in parallel to the samples. Proportionate amounts of 6xSB were added to the DNA Ladder and all samples before gel loading. Electrophoresis was carried out in 1xTAE buffer at 90V for 40-60 minutes. Visualization of DNA was performed by exposure to UV-light using BioDoc-It™ Transiluminator system (UVP).

3.2.6 DNA extraction from agarose gel

The DNA fragments were visualized on a Dual-Intensity Transilluminator (UVP) using the low setting of UV-light. The DNA fragments were cut out from the agarose gel with a clean scalpel. Extraction of the DNA fragments from the gel slice was performed using QIAquick Gel Extraction Kit (QIAGEN) following the manufacturer's protocol. DNA fragments were eluted from the spin column using dH$_2$O.

3.2.7 Ligation of DNA fragments

Ligation of DNA fragments was performed using T4 DNA Ligase (NEB or Roche) with supplemented buffer. Ligation took place under optimal reaction conditions according to the manufacturer's protocol, with an Insert:Vector ratio of 4:1 and a total reaction volume of 20μl. The ligation mix was incubated ON at room temperature (RT).

3.2.8 PCR amplification of DNA

Polymerase chain reaction (PCR) was used for amplification of DNA fragments using primers with tails containing RE sites. PCR was set up with Phusion DNA Polymerase (FINNZYMES) using optimal reaction conditions according to the protocol with a cycling program adjusted to the primers used. Controls for each primer were run in parallel. Validation of PCR product was done by agarose gel electrophoresis (Section 3.2.5). PCR products were purified using QIAquick PCR purification Kit (QIAGEN). Validation of purified PCR product was done by agarose gel electrophoresis (Section 3.2.5).

3.2.9 QuikChange Multi Site-Directed Mutagenesis

Multiple site-directed mutations were introduced using QuikChange® Multi Site-Directed Mutagenesis Kit (Stratagen) with optimal reaction conditions and procedure according to the protocol. Cycling program was adjusted according to the plasmid length. *Dpn* I from NEB and Electroporation competent XL-1 Blue cells were used instead of the equivalent materials supplied by the manufacturer.

3.2.10 Transformation of *E. Coli*-strains

Transformation of electroporation competent XL-1 Blue

XL-1 Blue cells were used both in cloning and protein production. Electroporation competent XL-1 Blue cells were thawed on ice. 39μl cells per transformation were transferred to a new, pre-chilled Eppendorf tube and kept on ice. 1 μl ligation mix (Section 3.2.7) or other DNA sample was added to the cells and mixed by stirring gently with the pipette tip. The sample was transferred to a pre-chilled electroporation cuvette (Molecular Bio Products). A Gene Pulser™ electroporator apparatus (BioRad) set on 1.3 kV, 200 Ω and 25 μFD was used to pulse the cells. 0.96 ml 2xYT was mixed with the sample in the cuvette and the total volume of 1 ml was transferred to a new Eppendorf tube and incubated at 37°C for 1 hour. Transformation mix was plated out on agar plates with 100 μg/ml ampicillin. The plates were incubated upside-down at 37°C ON.

Transformation of CaCl$_2$ competent RosettaBlue^{TM}(DE3)pLysS

CaCl$_2$ competent RosettaBlue^{TM}(DE3)pLysS (Novagen) cells were transformed according to the manufacturer's protocol.

Preparation of CaCl$_2$ competent XL-1 Blue pGro7 and BL21 (DE3)

XL-1 Blue pGro7 were made by transformation of electroporation competent XL-1 Blue (Section 3.2.10) with the pGro7 plasmid (Takara). CaCl$_2$ competent XL-1 Blue pGro7 and BL21 (DE3) were prepared by Ruby Sathiaruby Sivaganesh according to the ISA standard protocol.

Transformtaion of CaCl$_2$ competent XL-1 Blue pGro7 and BL21 (DE3)

CaCl$_2$ competent cells were thawed on ice. 100 µl cells were transferred to a pre-chilled Eppendorf tube and incubated on ice for 10 minutes. 5 µl of plasmid prep (undiluted or 10^2 diluted) was added to the cells and mixed by gently stirring the pipette tip. Incubation of the cells continued on ice for 30 minutes. The cells were heat pulsed at 42°C for 3 minutes followed by 2 minutes incubation on ice. 0.9 ml of pre-heated 2xYT medium was added to each tube of cells and the tubes were incubated at 37°C for 1 hour. 100 µl of sample was plated out on agar plates with appropriate antibiotics for selection. Plates were incubated upside-down at 37°C ON.

3.3 Construction of plasmids

All constructs created in this study were designed as shown by the schematic overview in Figure 5 A or B. The new plasmids were named with abbreviations describing the fusion proteins they express, following the *N-* to *C*-terminal order of the expressed entities, e.g. pT2scTCR-GFPFN encode the scTCR fused to GFP with a TorA2 leader sequence and the over-expression of FkpA. The expression vectors as shown in Figure 5 C.

3.3.1 Subcloning of pPscTCR-YFP

The scTCR in the Vαβ conformation in fusion to YFP was moved to an expression vector without co-expression of FkpA. The pYFPFN-scTCR Vαβ 4B2A1 and pYFP-phOx were digested with *Nco*I and *Not*I (Section 3.2.3). The DNA fragments from the digests were separated by agarose gel electrophoresis (Section 3.2.5). The scTCR cassette and the digested vector were extracted from the agarose gel (Section 3.2.6) and ligated together (Section 3.2.7). Ligation mix was transformed into Electroporation competent XL-1 Blue cells (Section 3.2.10). A random selection of clones were inoculated in 5 ml LB-medium supplied with 100mM glucose and 100µg/ml ampicillin and incubated at 37°C, ON with 250rpm in a shaking incubator (NBS InnOva 4000 Incubator Shaker or JEIO-TECH SI600R). Plasmids were isolated and clones were validated by sequencing performed by the Abi-lab (CEES, Department of Biology and Molecular Biosciences, University of Oslo).

(A) Expression vector elements without FkpA
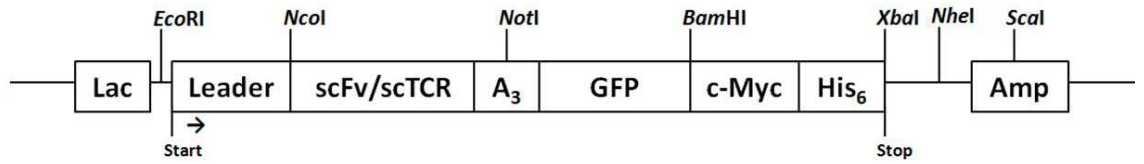
(B) Expression vector elements with FkpA
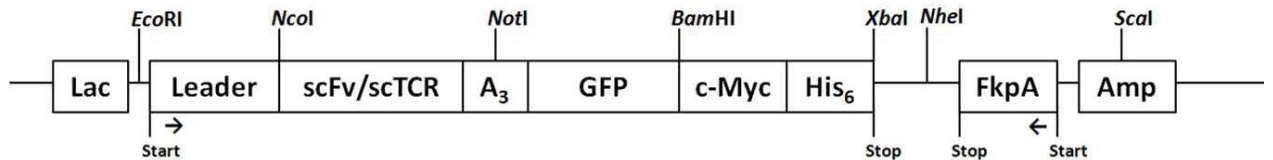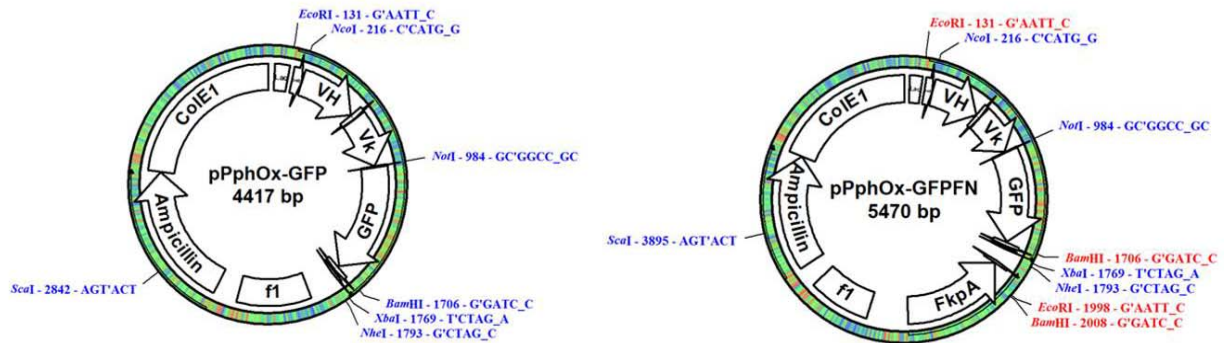
(C) Expression vectors

Figure 5: A schematic overview of the expression elements and vectors used. The figure shows the different exspression elements in (A) and (B): the Lac promoter (Lac); the leader sequence (Leader) being either the PelB-, TorA1-, TorA2-, DsbA1-, DsbA2- or DsbA3-leader; the scFv/scTCR cassette being scFv-anti phOx, scFv-anti NIP or scTCR; the linker (A$_3$) consisting of three alanine residues; the GFP (GFP); the c-Myc-tag (c-Myc); the His$_6$-tag (His$_6$); the periplasmatic chaperon FkpA (FkpA); the ampicillin resistance marker (Amp). The elements are not scaled according to actual size. (C) shows the overview of the two different expression vectors used, exemplified by pPphOx-GFP (without FkpA) and pPphOx-GFPFN (with FkpA).

3.3.2 Construction of scTCR-YFP fusion constructs with TorAB7 leaders

To create pT1-scTCR-YFP and pT2scTCR-YFP, the two leader versions TorA1 and TorA2 were created by PCR (Section 3.2.8) with the primer mix of TorAB7_EcoRI_new(frwd)/ TorAB7_1NcoI(rev) and TorAB7_EcoRI_new (frwd)/TorAB7_2NcoI(rev), respectively. The PCR was run with the following program: 98°C for 30 seconds, 30 repeats of denaturation (98°C for 10 seconds), annealing (54°C for 10 seconds) and elongation (72°C for 10 seconds) followed by 72°C for 5 minutes before cooling to 4°C. The primers contained tails with *Eco*RI or *Nco*I RE sites. The purified PCR products and pPscTCR-YFP were RE digested with

*Eco*RI and *Nco*I at 37°C ON (Section 3.2.3). Ligation, transformation and validation of clones were performed as described in Section 3.3.1.

3.3.3 Construction of scTCR-YFP fusion constructs with DsbA leader

To create pD1scTCR-YFP, pD2scTCR-YFP and pD3scTCR-YFP, the three leaders DsbA1, DsbA2 and DsbA3 were ordered as plus and minus strand oligonucleotides with overhangs as if digested with *Eco*RI and *Nco*I REs. Annealing of oligos was performed by mixing 2.5 µg of each DNA strand in a 1.5 ml eppendorf tube. The tubes were incubated at 95°C for 1 minute and cooled down to 4°C. The annealed oligos were treated with T4 Polynucleotide Kinase (NEB) with reaction conditions and procedure according to the manufacturer's protocol. DNA from the reaction mixes were precipitated with seeDNA (Section 3.2.2) and resuspended to a concentration of approximately 25ng/µl. Ligation of each leader into *Eco*RI/*Nco*I digested pPscTCR-YFP, transformation and validation of clones were performed as described in Section 3.3.1.

3.3.4 Construction of scTCR-GFP and scFv-GFP fusion constructs

Construction of pPphOx-GFPmut2

To make pPphOx-GFPmut2, GFPmut2 was amplified with PCR from pB7G using the primers pGB7NotI_fw and pGB7BamHI_rev which contains *Not*I and *Bam*HI RE sites in the tails, respectively. The PCR was run with the following program: 98°C for 30 seconds, 30 repeats of denaturation (98°C for 10 seconds), annealing (56°C for 15 seconds) and elongation (72°C for 10 seconds) followed by 72°C for 5 minutes before cooling to 4°C. The purified PCR products and pYFP-phOx were RE digested by *Not*I and *Bam*HI at 37°C ON as described in Section 3.2.3. Ligation, transformation and validation of clones were performed as described in Section 3.3.1. pPphOx-GFPmut2 was successfully validated by sequencing.

QuickChange of GFPmut2

GFPmut2 contains an *Nco*I RE site it was desirable to remove. In addition we wanted to introduce an A206K substitution in the GFP that has been shown to reduce the degree of dimerization of GFP (Zacharias, Violin et al. 2002; Shaner, Steinbach et al. 2005). These mutations were introduced by QuickChange Multiple Site-Directed Mutagenesis using the

primers a1159g and g1607a_c1608a_c1609a with pPphOx-GFPmut2 as template (Section 3.2.9). The cycling program used was 95°C for 1 minute, 30 repeats of denaturation (95°C for 1 minute), annealing (55°C for 1 minute) and elongation (65°C for 9 minutes and 12 seconds) followed by cooling to 4°C. Transformation and validation of clones were performed as described in Section 3.3.1. A validated clone with the desired mutations and the vector pGAL were RE digested with *Not*I and *Nhe*I (Section 3.2.3) in a two-step procedure with DNA precipitation (Section 3.2.2) in between in order to change optimal buffer for the REs. The modified GFPmut2 and the digested vector fragment were ligated, transformed and clones validated as described in Section 3.3.1. A clone with validated sequence from the *Not*I to the *Sal*I RE site was digested with these two enzymes in parallel to the vectors pYFP-phOx, pYFP-NIP, pPscTCR-YFP, pT1scTCR-YFP, pT2scTCR-YFP, pD1scTCR-YFP, pD2scTCR-YFP and pD3scTCR-YFP. The *Not*I/*Sal*I cassette was ligated into the vectors, transformed and validated as described in Section 3.3.1.

3.3.5 Assembly of pT2phOx-GFP

The pT2scTCR-GFP and pPphOx-GFP were digested with *Nco*I and *Not*I as described in Section 3.2.3. Ligation of the phOx cassette and the pT2scTCR-GFP digested vector, transformation, and validation of clones were performed as described in Section 3.3.1 in order to create pT2phOx-GFP.

3.3.6 Assembly of scTCR-GFP and scFv-GFP fusion constructs with FkpA

pPphOx-GFP, pPscTCR-GFP, pT2phOx-GFP, pT2scTCR-GFP and pYFPFN-phOx were digested with *Nhe*I-HF and *Sca*I-HF as described in Section 3.2.3. Ligation, transformation, and validation of clones were performed as described in Section 3.3.1 in order to create pPphOx-GFPFN, pPscTCR-GFPFN, pT2phOx-GFPFN and pT2scTCR-GFPFN.

3.4 Recombinant protein expression

3.4.1 Plasmids used for protein expression

Plasmids used in protein expression with the respective host strains are shown in Table 4.

3.4.2 Standard growth conditions for scTCR-GFP and scFv-GFP expression

Cells were inoculated from glycerol stock into 5 ml LB-medium in a 50 ml tube supplemented with 30μg/ml tetracycline, 100 μg/ml ampicillin and 0.1 M Glucose (LB$_{TAG}$-medium) for XL-1 Blue. LB$_{TAG}$-medium supplemented with 20μg/ml chloramphenicol (LB$_{TACG}$-medium) was used for Rosetta Blue (DE3) pLysS and XL-1 Blue pGro7. A LB$_{AG}$-medium containing 5 ml LB-medium supplemented with 100μg/ml ampicillin and 0.1 M glucose was used for BL21 (DE3). Cultures were incubated at 37°C ON with 250 rpm. The ON cultures were reinoculated into 5 ml of appropriate medium in 50 ml tubes with an OD$_{600nm}$~0.1. The cultures were incubated 37°C with 250 rpm until an OD$_{600nm}$ of 0.6-0.8 was reached. The cultures were centrifuged 4500g for 10 minutes and resuspended in 5 ml of respective medium without glucose and supplemented with 0.1 mM isopropyl-β-D-thio-galactopyranoside (IPTG) in 50 ml tubes. XL-1 Blue pGro7was in addition supplemented with 0.5 mg/ml L-arabinose. Cell cultures were incubated at 30°C ON with 250 rpm unless otherwise noted.

Table 4: Overview of plasmids used for protein expression.

Plasmid name	Leader	scTCR/scFv	Co-expression	Expression strains[a]
pPphOx-GFP	PelB	scFv anti-phOx		XB, XG, RB
pT1phOx-GFP	TorA1	scFv anti-phOx		XG
pT2phOx-GFP	TorA2	scFv anti-phOx		XB, XG, RB, BL
pPphOx-GFPFN	PelB	scFv anti-phOx	FkpA	XB, RB
pT2phOx-GFPFN	TorA2	scFv anti-phOx	FkpA	XB, RB
pPNIP-GFP	PelB	scFv anti-NIP		XB
pPscTCR-GFP	PelB	scTCR		XB, XG, RB
pT1scTCR-GFP	TorA1	scTCR		XB, XG
pT2scTCR-GFP	TorA2	scTCR		XB, XG, RB
pD1scTCR-GFP	DsbA1	scTCR		XB
pD2scTCR-GFP	DsbA2	scTCR		XB
pD3scTCR-GFP	DsbA3	scTCR		XB
pPscTCR-GFPFN	PelB	scTCR	FkpA	XB, RB
pT2scTCR-GFPFN	TorA2	scTCR	FkpA	XB, RB

[a] Expression hosts abbreviations: XB – XL-1 Blue, XG – XL-1 Blue pGro7, RB – Rosetta Blue (DE3) pLysS and BL – BL21 (DE3)

3.4.3 Cell fractionation

Cell materials were normalized according to OD_{600nm} and centrifuged at 4500 rpm for 15 minutes at 4°C. 1 ml supernatant was transferred to a 1.5 ml eppendorf tube, centrifuged at 13000 rpm for 10 minutes at 4°C and 750µl supernatant was transferred to a new 1.5 ml eppendorf tube and stored at 4°C (fraction I, medium fraction). The cell pellet was resuspended in 490 µl ice-cold sucrose solution (20% sucrose, 50 mM Tris/HCl pH 8.0 and 1 mM EDTA pH 8.0) together with 5µl lyzosyme (100 mg/ml in 50 mM Tris/HCl pH 8.0) and 5 µl RNaseA (10 mg/ml in Tris/HCl pH 8.0). The resuspended pellet was transferred to a new 1.5 ml eppendorf tube and incubated for 1 hour at 4°C with rotation. The cells were pelletted by centrifugation at 13000 rpm for 5 minutes at 4°C. The supernatant was transferred to new 1.5 ml eppendorf tube, centrifuged at 13000 rpm for 10 minutes at 4°C and 450µl supernatant was transferred to a new 1.5 ml eppendorf tube and stored at 4°C (fraction II, soluble periplasmic fraction). Residual liquid was removed from the cell pellet and the cells were resuspended in 475 µl 0.1 M Tris/HCl pH 8.0. 167 µl 4xSDS was thoroughly mixed with the cell solution and stored at 4°C (fraction III, cytoplasmic fraction). All fractions were finally stored at -20°C.

3.5 Analysis of expression level, subcellular localization and fluorescence

3.5.1 SDS-PAGE

Samples of expression culture or fractions were investigated by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE). Samples were mixed with proportionate amounts of 4xSDS. The samples and 10 µl Prestained Protein Marker Broad Range (7-175 kDa) (NEB) were heated at 95°C for 5 minutes prior to loading on a 4-12% Bis-Tris XT Precast gel (BioRad). The gel was run at 140 V for 1 hour and 45 minutes.

3.5.2 Coomassie staining

For visualization of the proteins in. The gel was washed 2 x 5 minutes in dH₂O, stained with Bio-Safe Coomassie Stain (BioRad) for 1 hour at RT with agitation, followed by washing 1 hour in dH₂O to visualize the protein bands. Pictures were taken with normal light using BioDoc-It™ Transiluminator system (UVP).

3.5.3 Western blot analysis

Western blot analysis was performed by semi-dry blotting. Samples separated on SDS-PAGE gel was blotted on to polyvinylidine fluoride membrane (Millipore) in a Trans-Blot SD Semi-Dry Transfer cell (Bio-Rad) in Tris/glycin buffer (25mM Tris, 192mM glycine, 20% methanol, pH 8.3). The apparatus was assembled and run at 25 V for 30 minutes. The membrane was blocked in 4% skimmed milk in PBST (PBS with 0,05% Tween 20) for 1.5 hours at RT or ON at 4°C, followed by washing 3 times for 5 minutes in PBST. Incubation with immunodetection reagents was performed in 4% skimmed milk in PBST for 1.5 hours at RT or ON at 4°C. The membrane was washed 3 times for 5 minutes in PBST after each incubation. Immunodetection of GFP was performed by incubating with anti-GFP Ab (Abcam) diluted 1:2000, followed by SA-HRP (Amersham) diluted 1:5000. Immunodetection of the His_6-tag was performed by incubating with anti-His-6-tag (C-term) Ab (Invitrogen) diluted 1:5000, followed by incubation with anti-mouse IgG-HRP Ab from sheep (Amersham) diluted 1:7000. Immunodetection of GroEL was performed by incubating with rabbit anti-GroEL Ab (Sigma) diluted 1:10000, followed by incubation with donkey anti-rabbit IgG Ab(Amersham) diluted 1:50000. Immunodetection of FkpA was performed incubating with anti-Flag M2-HRP Ab (Sigma) diluted 1:7000. Immunodetection of scFv anti-phOx was performed by incubation with rabbit anti-human lambda L-chain AB (Dako) diluted 1:4000, followed by incubation with donkey anti-rabbit IgG HRP Ab (Amersham) diluted 1:8000. When more than one immunodetection was performed on a membrane, it was incubated covered in Restore Western Blot Stripping Buffer (Thermo) for 30 minutes at 37°C to strip for detection reagents. Development of the western blots was performed incubation for 5 minutes in 5 ml Super Signal West Pico Lumino Enhancer Solution (Pierce) mixed with 5 ml Super Signal West Pico Stable Peroxide Solution (Pierce) and finally exposure of the membrane to Kodak BioMax MR film in a film cassette. The film was developed in an Optimax X-Ray Film Processor (PROTEC).

3.5.4 Qualitative analysis with fluorescence microscopy

Cell material was normalized according to OD_{600nm} and centrifuged 4500 g for 5 minutes at RT. Supernatants were discarded and the pellets gently resuspended in 1 ml PBS followed by centrifugation for 5 minutes at 4500 g at RT. The pellet was resuspended in a final volume of

500µl PBS. 2 µl cells were applied on a microscope slide and a cover slide was put on top. Pictures were taken using an Axioplan 2 Imaging fluorescence microscope (Zeiss) with filter 8 for fase contrast pictures (20 ms) and filter 2 (FITC-filter) for fluorescence pictures (100 ms).

3.5.5 Quantitative analysis of fluorescence with Victor fluorometer

Cell material was normalized and washed as described in Section 3.5.4. 100 µl cells or 100 µl recombinant GFP (Clonetech) (10 µg/ml) diluted in PBS was added to FluoroNunc 96-well plates, black (Nunc). Fluorescence detection in the plate was performed with a Victor 1420 Multilabel counter (Perkin Elmer) using the Fluorometry protocol at 485nm/535nm 1.0 sec. Triplicate of each sample was measured, however the variation within these measurements were less than 5%, and typically much smaller. Hence, average of the three samples is shown without standard deviations.

4 Results

4.1 Construct design

The Ig domain is dependent on disulfide bridge formation to obtain a functional conformation. Therefore, the three scaffolds chosen here, the scTCR and the scFvs with anti hapten specificity for NIP and phOx, denoted NIP and phOx, are targeted for periplasmic expression by a pelB leader. Due to differences in their solubility, the scTCR, NIP and phOx represents three different expression profiles in *E. coli*; only cytosolic; cytosolic and periplasmic; cytosolic, periplasmic and medium, respectively. The ultimate goal was to make a screening platform. These three scaffolds were chosen as "probes" and genetically fused to GFP to take advantage of the fluorescent properties of GFP in quantitative screening. To explore whether periplasmic expression of scTCR-GFP fusions could be obtained, three translocation mechanisms were compared using both the Sec- and Tat-pathways. Therefore, in addition to the pelB leader, the DsbA and TorA leader peptides were utilized for periplasmic targeting.
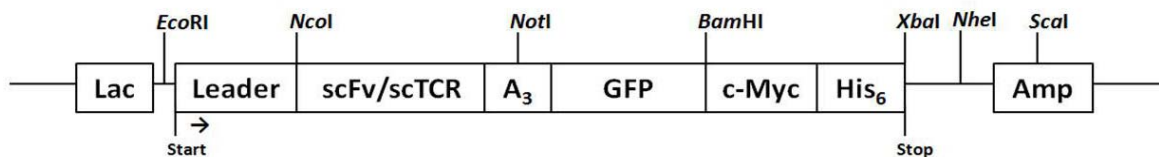
Figure 6: Schematic overview of key elements in the expression vector. The size of the elements in the figure is not scaled in proportion to sequence length.

4.1.1 Cloning strategy for the TorA and DsbA leader

The DsbA signal sequence targets polypeptides to the periplasm through the co-translational SRP-dependent secretory pathway and thereby prevents polypeptides from folding in the cytoplasm (Schierle, Berkmen et al. 2003). The TorA signal sequence targets proteins to the Tat-translocator for transport of folded or partially folded domains to the periplasm(Berks, Palmer et al. 2005). Premature folding of GFP in the cytoplasm should therefore not impede translocation by the TorA leader.

The parental pHOG vector used for protein expression in the current study contains RE-based cassettes, which are easily interchanged between the different constructs (Kipriyanov, Moldenhauer et al. 1997). The signal peptide is coded for in the 3'-end of the *Eco*RI/*Nco*I cassette (as shown in Figure 6).

Ineffective leader sequence removal after protein translocation may negatively affect the functionality of the mature polypeptide and the aa positions flanking the protease cleavage site is known to significantly influence the efficiency of the proteolysis (Shen, Lee et al. 1991). Therefore, the transition (the residues in the close vicinity of the interface between the signal sequence and the protein of interest) is crucial for functional transport of proteins. Hence, different transitions between the signal sequences chosen in the current study and the downstream scTCR cassette were tested *in silico* by SignalP 3.0 (Bendtsen, Nielsen et al. 2004) for the probability of 1) signal peptide recognition, 2) signal peptide cleavage efficiency and 3) cleavage site, to examine whether the new leader sequences could be designed compatible with the cassette system.

TorA leader design

A previous study isolated a gain-of-function mutant of the TorA leader (TorAB7) that shows a 6-fold increase in periplasmic expression of GFP compared to the *wild type* TorA leader (DeLisa, Samuelson et al. 2002). *In silico* analysis of the TorAB7 sequence fused directly to the scTCR sequence yielded low signal peptide and cleavage probability scores (Table 5) and was therefore not expected to be functional. In contrast, the TorAB7 sequence in context of its native partner, TorA, yielded a high probability score for signal sequence recognition and cleaving after position 39, which is the native cleavage site (Table 5).

Therefore, a possible strategy for designing a functional leader was to add residues from the TorA protein between the TorAB7 leader and the scTCR. Two sequences with different transitions were chosen for *in vivo* analysis and the leader sequences were named TorA1 and TorA2 (shown in Figure 7). To ensure that the resulting leader sequences would be functionally interchangeable with the pelB leader in other constructs in the expression system, the TorA1 and TorA2 leaders were analyzed *in silico* in fusion to scFv anti-phOx and anti-

NIP. This analysis gave similar probability scores and cleavage site as the scTCR (data not shown). The TorA1 and TorA2 leaders are showed in Figure 7.

The primer TorAB7_EcoRI_new (shown in Table 2) was designed to contain the stretch from the *Eco*RI site to the start of the leader peptide as a 5'-tag, in addition to the annealing to the 5' end of TorAB7 in pB7G. The primers TorAB7_1NcoI and TorAB7_2NcoI (shown in Table 2) were designed to anneal to the 3' end of the TorAB7 signal sequence and contain the respective transistion sequence and the *Nco*I RE site as a 5'-tags.

DsbA leader design

Based on a previous study (Steiner, Forrer et al. 2006), the DsbA leader sequence was chosen for co-translational translocation via the SRP-dependent pathway. The signal sequence was obtained from SwissProt (P0AEG4) and analyzed in fusion to scTCR *in silico* as previously described in the TorA design. Important considerations were peptide cleavage and whether the *Nco*I RE site could be preserved. Three sequences with different transistions were chosen for *in vivo* analysis. The chosen DsbA1, DsbA2 and DsbA3 leaders are shown in Figure 7. All three were analyzed *in silico* in fusion to scFv anti-phOx and anti-NIP and gave similar probability scores as the scTCR (data not shown). Three *Eco*RI/*Nco*I cassettes containing each of the DsbA leaders were ordered as plus and minus strand oligonucleotides (Table 2).

PelB	*N*-MKYLLPTAAAGLLLLAAQPAMA|-*C*
TorA1	*N*-MNNNDLFQTSRQRFLAQLGGLTVAGMLGPSLLTPRRATA|AMA-*C*
TorA2	*N*-MNNNDLFQTSRQRFLAQLGGLTVAGMLGPSLLTPRRATA|AQAAMA-*C*
DsbA1	*N*-MKKIWLALAGLVLAFSAMA|-*C*
DsbA2	*N*-MKKIWLALAGLVLAFSASMA|-*C*
DsbA3	*N*-MKKIWLALAGLVLAFSASAAMA|-*C*

Figure 7: Overview of the leader sequences."|"represents predicted cleavage site; aa in red are mutated or additional aa.

Table 5: *In silico* analysis of transitions between leader peptides and scTCR using the hidden Markov model (HMM) in SignalP3.0 (http://www.cbs.dtu.dk/services/SignalP/). This analysis was performed by Geir Åge Løset.

Input	Signal peptide probability score[a]	Cleavage site probability score[b]	Predicted cleavage site[c]
TorAB7-scTCR	0,056	0,030	37-38
TorA-TorA	0,963	0,950	39-40
TorAB7-TorA	0,941	0,929	39-40
TorAB7-AMA-scTCR	0,650	0,556	39-40
TorAB7-AAMA-scTCR	**0,829**	**0,772**	**39-40**
TorAB7-AAQMA-scTCR	0,874	0,863	39-40
TorAB7-AAQAMA-scTCR	0,890	0,854	39-40
TorAB7-AAQAAMA-scTCR	**0,935**	**0,910**	**39-40**
TorAB7-AAQAAMAT-scTCR	0,935	0,906	39-40
DsbA-scTCR	1,000	0,757	19-20
DsbA-A-scTCR	1,000	0,900	19-20
DsbA(S18M)-scTCR	**1,000**	**0,964**	**19-20**
DsbA(S18M)-A-scTCR	1,000	0,993	19-20
DsbA-AMA-scTCR	**1,000**	**0,973**	**22-23**
DsbA-MA-scTCR	**1,000**	**0,995**	**21-22**

[a] Probability of the sequence containing a signal peptide. [b] Probability of cleavage at the predicted cleavage site. [c] Predicted cleavage site between indicated residues.

Table 6: Over view of proteins and their theoretical molecular weights (MW).

Protein	Leader	scTCR/scFv	MW[a] preprotein	MW[a] without leader
PphOx-GFP	PelB	scFv anti-phOx	59kDa	57kDa
T1phOx-GFP	TorA1	scFv anti-phOx	61kDa	57kDa
T2phOx-GFP	TorA2	scFv anti-phOx	61kDa	57kDa
PNIP-GFP	PelB	scFv anti-NIP	59kDa	56kDa
PscTCR-GFP	PelB	scTCR	59kDa	56kDa
T1scTCR-GFP	TorA1	scTCR	61kDa	57kDa
T2scTCR-GFP	TorA2	scTCR	61kDa	57kDa
D1scTCR-GFP	DsbA1	scTCR	58kDa	56kDa
D2scTCR-GFP	DsbA2	scTCR	59kDa	56kDa
D3scTCR-GFP	DsbA3	scTCR	59kDa	56kDa
PphOx+c-myc+His$_6$	PelB	scFv anti-phOx	32kDa	30kDa
FkpA(+FLAG)			30kDa	27kDa
GroEL			~60kDa[b]	NA
rGFP			~27kDa[c]	NA
GFP+c-myc+His$_6$				29kDa
PphOx	PelB	scFv anti-phOx	29 kDa	27 kDa

[a] MW is calculated using http://www.expasy.ch/tools/pi_tool.html. [b] GroEL MW was adopted from Takara. [c] Recombinant GFP MW was adopted from Clonetech.

4.2 Comparison of YFP and GFP constructs

The YFP fluorophore encoded in the parental pHOG plasmid was exchanged with a gene encoding a mutant GFP fluorophore developed to yield increased fluorescence with standard FITC filters. This particular GFP mutant also exhibits increased functional expression in prokaryotic hosts (Cormack, Valdivia et al. 1996). To evaluate differences in performance, PphOx-GFP and PphOx-YFP was expressed in *E. coli* XL-1 Blue using identical conditions followed by whole cell fluorescent imaging. The results in Table 7 showed that the GFP construct yielded increased fluorescence as compared to the YFP construct when using the FITC-filter. The GFP vector was therefore chosen for further use in the current study.

4.3 Expression characteristics of scTCR and scFvs

The expression profile of the scTCR and the scFvs phOx and NIP, all with pelB leader, were validated by expression in XL-1 Blue followed by fractionation, SDS-PAGE and western bloting with anti-c-myc detection. Figure 8 shows that phOx was detected in the medium, periplasmic and cytosolic fractions, NIP was detected in the periplasmic and cytosolic fractions and the scTCR was only detected in the cytosolic fraction.
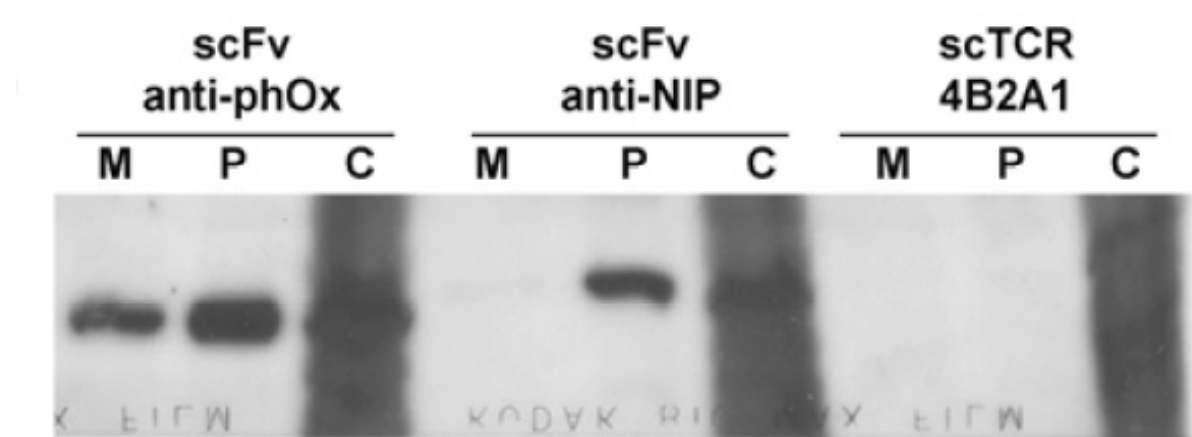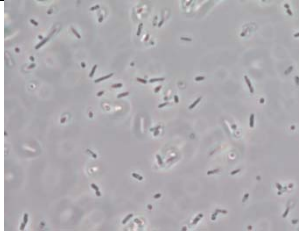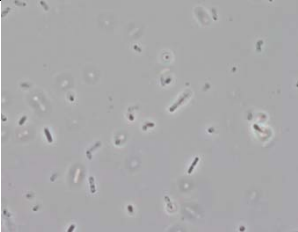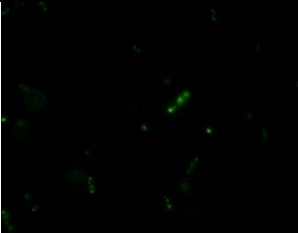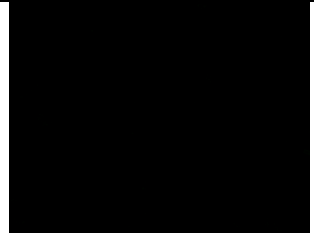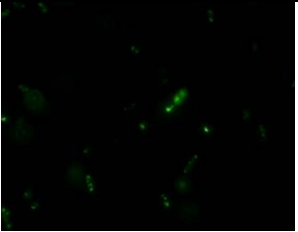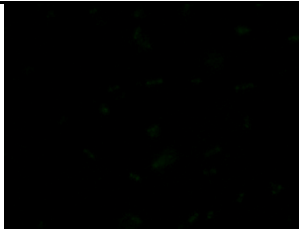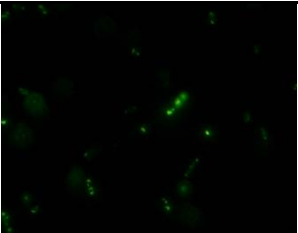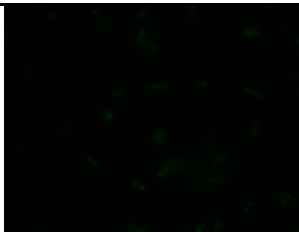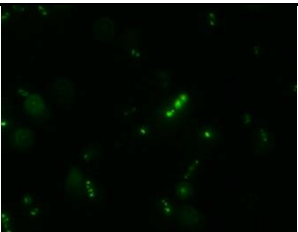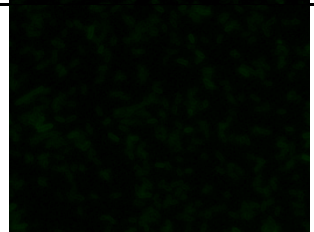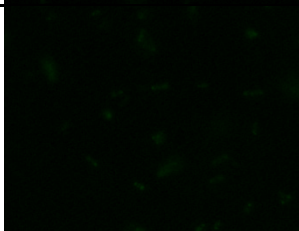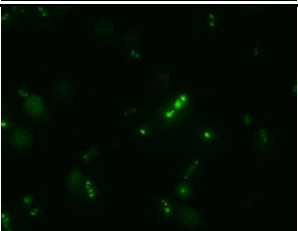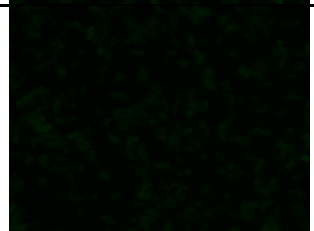

Figure 8: Expression of scTCR and the scFvs anti-phOx and anti-NIP followed by fractionation, SDS-PAGE and western bloting by anti-c-myc detection. M – Medium fraction; P – Periplasmic fraction; C – Cytosolic fraction. This experiment was performed by Geir Åge Løset.

Table 7: Fluorescence imaging of PphOx-GFP and PphOx-YFP expressed in XL-1 Blue.

Exposure time	PphOx-YFP	PphOx-GFP	XL1-Blue
20ms			
50ms			
75ms			
100ms			
150ms			
200ms			

4.4 Expression characteristics of scTCR-GFP and scFv-GFP fusions in XL-1 Blue

To analyze the expression profile of the three entities phOx, NIP and scTCR in fusion to GFP and to investigate the expression characteristics of the scTCR-GFP fusion with different leaders, the constructs were expressed in XL-1 Blue with standard expression conditions.

4.4.1 Effect of leader on total expression level

The total expression levels were analyzed by SDS-PAGE of normalized whole cell cultures followed by western blot (Figure 9). As seen in Figure 9, both the scFv-GFP fusions with pelB leader were detected at ~58 kDa, which is close to the theoretical predicted size. Anti-His$_6$ (Figure 9 A) and anti-GFP (Figure 9 B) were used as detection reagents. One and two weak bands, respectively, were shown which indicates the presence of cleavage products at ~30 kDa of phOx-GFP with pelB leader. This is about half the size of the full length fusion protein. Since the two entities that compose the fusion protein have approximately the same size, it indicates a cut in the linker region between the two modules.

For the scTCR-GFP fusions there were large differences in expression level between the various leader constructs. The pelB-construct is detected at expected size with an expression level similar to PNIP-GFP. Both the TorA-constructs are detected at approximately predicted size with expression levels far greater than that seen for pelB constructs. Some higher aggregates, possibly dimers, and bands at ~40kDa were also detected for the TorA-constructs. In contrast, of the three DsbA-constructs only the D1scTCR-GFP was detected with anti-His$_6$ (Figure 9 A) and none of the DsbA-constructs were detected with anti-GFP (Figure 9 B). The detected of D1scTCR-GFP with anti-His$_6$ might have been a result of overflow from the neighboring T2scTCR-GFP.

In summary, these results indicate that the signal peptide sequence had a profound effect on the total expression level of fusion proteins with TorA giving the highest level, pelB giving an intermediate level and DsbA not giving any detected level. The most striking observation was the very low expression levels seen with the DsbA leader construct.
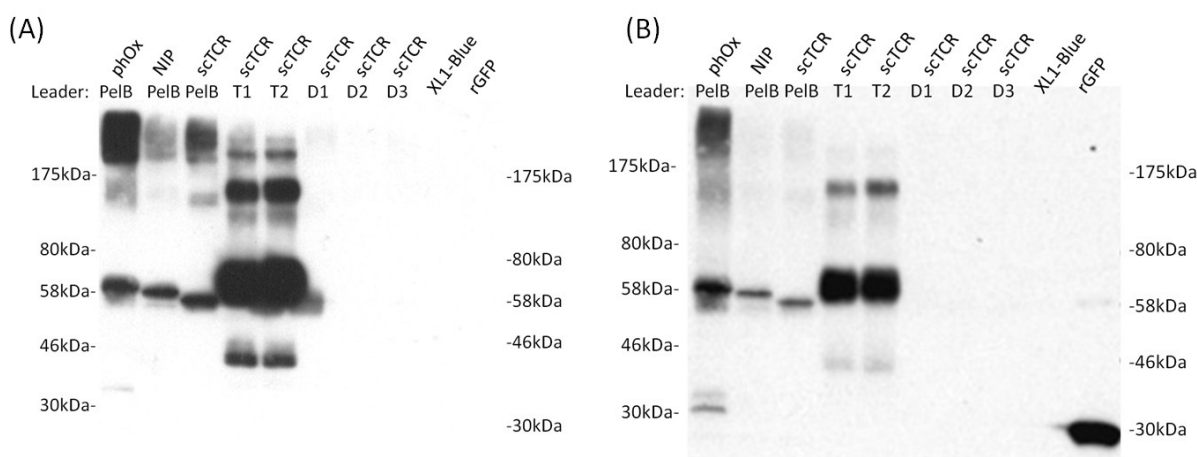
Figure 9: Normalized samples of cell cultures with expressed protein in XL1-Blue analyzed by SDS-PAGE and western blot. (A) Detection by anti-His$_6$ (Invitrogen) and anti-mouse IgG-HRP Ab from sheep (Amersham). (B) Detection by anti-GFP (Abcam) and SA-HRP (Amersham).

4.4.2 Effect of leader on subcellular localization

To investigate the effect of leader peptides on subcellular localization of scTCR-GFP fusions and the subcellurlar localization of the scFv-GFP fusions, normalized amounts of expression culture were fractionated and the medium (M), periplasmic (P) and cytoplasmic (C) were analysed by SDS-PAGE and western bloting (Figure 10).

Figure 10 shows that all three fusion proteins with pelB leader were mainly detected in the cytoplasmic fraction. Small amounts of scFv-GFP fusions transported to the periplasm and cleaved. Furthermore, neither the two TorA leaders, nor the three DsbA leaders increased the periplasmic or medium localization of the scTCR-GFP fusion proteins. Scrutiny of the blots revealed that PhOx-GFP in the cytoplasmic fraction is at the size of full length fusion proteins, with a smear of aggregated proteins at larger size and some degraded or cleaved proteins at smaller size. In the periplasmic and medium fractions, phOx-GFP detection with anti-GFP showed two bands at ~30kDa and ~34kDa (Figure 10 A and B), while detection with anti-His$_6$ only showed one band at ~34kDa (Figure 10 C and D). In the periplasmic fraction PNIP-GFP was only detected by anti-GFP in Figure 10 B with a band at ~30kDa. This indicated that phOx-GFP and NIP-GFP were transported to the periplasm where cleavage and possibly degradation took place.

Figure 10 A and C showed that the scTCR-GFP fusion protein were not in the medium or periplasmic fractions with anti-GFP or anti-His$_6$ detection, regardless of leader peptide. The

cytoplasmic fractions of scTCR-GFP in Figure 10 B and C, detected with anti-GFP and anti-His$_6$ Abs, respectively, showed a similar expression pattern as for the whole cell fractions in Figure 9: high expression level with TorA1 and TorA2 leaders, mainly of full length fusion proteins but also some higher aggregates and cleavage/degradation products; medium expression level with PelB leader of full length protein and with some higher aggregates and minor amounts of cleavage/degradation products; the DsbA-constructs are detected with a very low expression level in contrast to Figure 9 where no protein is detected. The low yield of detected DsbA constructs confirms functional open reading frames and can be explained by up concentration that occurred during fractionation.

Empty XL1-Blue cells show no detectable proteins with anti-His$_6$ or anti-GFP detection in Figure 10 A or B. However, in Figure 10 B a background detection at ~28kDa with anti-GFP is seen in all cytoplasmic fractions including the empty XL1-Blue cells and this was observed occasionally using the polyclonal anti-GFP ab. The scFv anti-phOx in pHOG with a PelB leader was expressed in parallel as a control since this protein is expressed in all three fractions with a detectable His$_6$-tag as seen in Figure 10 D.

4.4.3 Effect of leader on relative fluorescence in cells

To analyze if functionally folded protein is expressed, the relative fluorescence of cells was measured. There is a strong correlation between the folding state of the fusion partner and the GFP entity, as assessed by the fluorescence of the GFP (Zhang, Cantor et al. 2005). Thus, as the GFP is expressed as a *C*-terminal reporter in our fusion proteins, it is reasonable to assume that the *N*-terminal fusion partner has to fold correctly for the fusion proteins to exhibit fluorescence. Cells harboring the fusion proteins were therefore examined for their fluorescent properties.

After expression, normalized samples of cells were washed and relative fluorescence measured. Empty XL1-Blue cells were used to indicate autofluorescence in the bacteria and a standard of 10 µg/ml rGFP was used as a positive control. Triplicate of each sample was applied to the microtiter plate and measured in relative fluorescent units (RFU). The average of each sample is presented in Figure 11 without standard deviation as the variation within the triplicates was less than 5%.
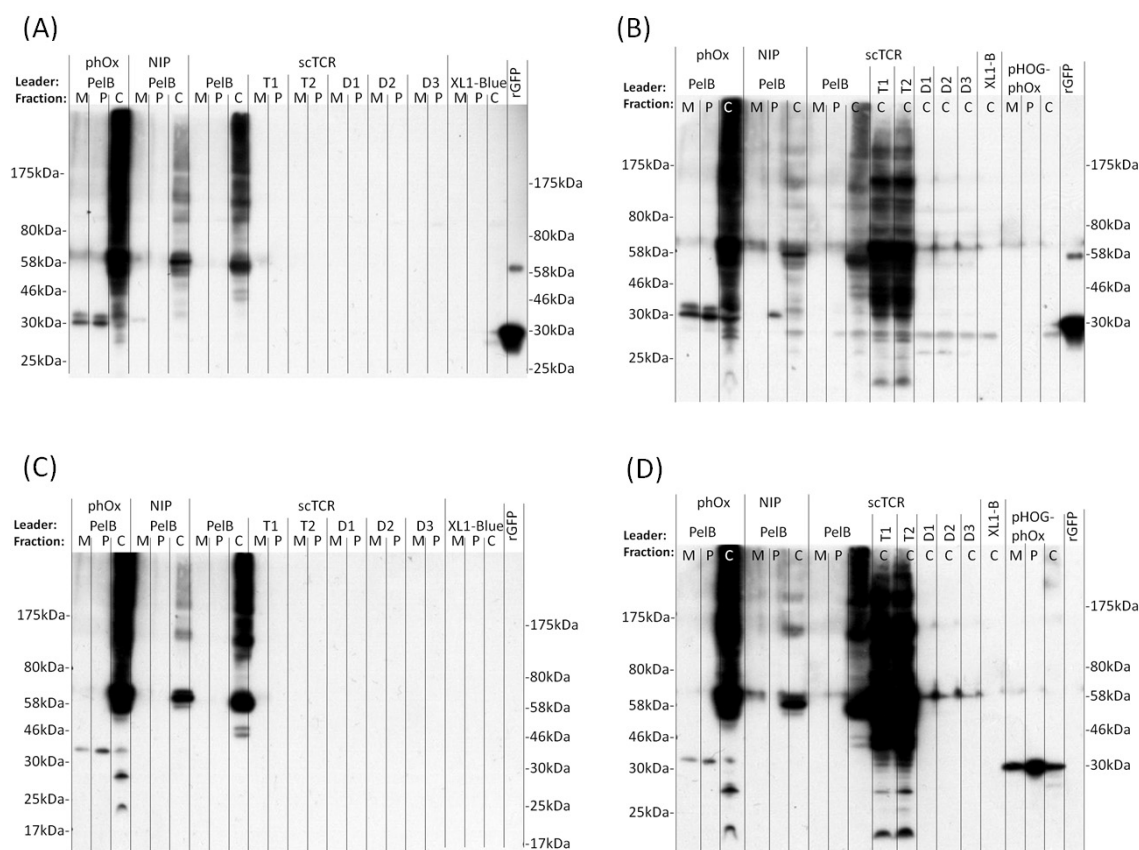
Figure 10: Expression analysis of subcellular localization in XL-1 Blue. Fraction M – medium, P – periplasm, C – cytoplasm. XL-1 Blue is cells without plasmid, pHOG-phOx is the parental scFv anti-phOx plasmid without GFP, rGFP is recombinant GFP. Detection with anti-GFP (A and B) or anti-His$_6$ (C and D). (A) Medium and soluble periplasmic fractions of scTCR-GFP with TorA and DsbA leaders and all fractions of phOx- and NIP-GFP. (B) Cytoplasmic fractions of scTCR-GFP with TorA and DsbA leaders and all fractions of phOx- and NIP-GFP. (C) Medium and soluble periplasmic fractions of scTCR-GFP with TorA and DsbA leaders and all fractions of phOx- and NIP-GFP. (D) Cytoplasmic fractions of scTCR-GFP with TorA and DsbA leaders and all fractions of phOx- and NIP-GFP.
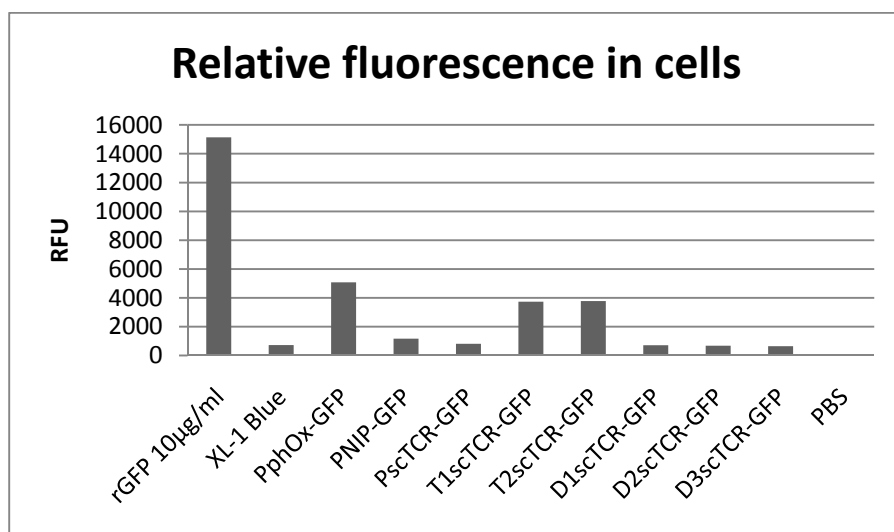
Figure 11: The relative fluorescence of normalized samples of cells, washed in PBS and analyzed in black 96-well microtiter plates using a Victor 1420 Multilabel counter (Perkin Elmer) using the Fluorometry protocol at 485nm/535nm 1.0 sec. The effect of leader peptide on scTCR-GFP fluorescence and characterization of the fluorescence of PphOx-GFP and PNIP-GFP.

As shown if Figure 11, scTCR-GFP with pelB or DsbA leaders and PNIP-GFP were measured to a RFU level similar to the autofluorescence measured in XL-1 Blue. Hence, these fusion proteins were not correctly folded. PphOx-GFP and scTCR-GFP with TorA leaders showed a 7 and 5 fold higher RFU, respectively.

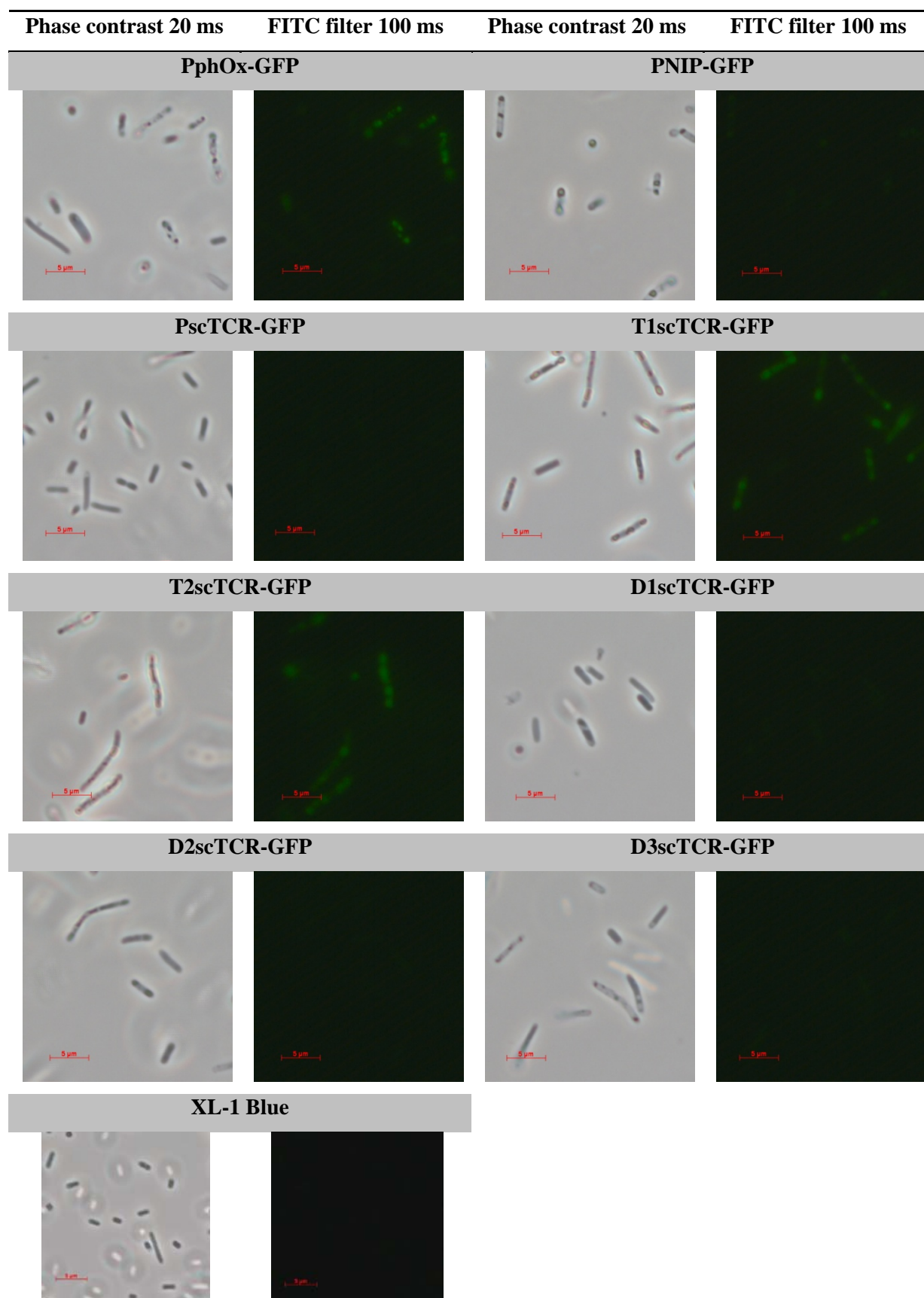## 4.4.4 Effect of leader on fluorescence microscopy of cells

To further investigate the fluorescent properties, cells were examined for the fluorescence distribution. Punctuated fluorescence has been shown to be an indication of periplasmic localization compared to evenly distributed fluorescence being a property of proteins expressed in the cytosol (Casey, Coley et al. 2000). The constructs were expressed in XL-1 Blue and cells washed prior to imaging.

The fluorescence images, presented in Table 8, visualize the differences in RFU measured and presented in Figure 11: PphOx-GFP, T1scTCR-GFP and T2scTCR-GFP show fluorescent properties while the other constructs, PNIP-GFP, PscTCR-GFP, D1scTCR-GFP, D2scTCR-GFP and D3scTCR-GFP, do not show fluorescence in comparison to the autofluorescence of empty XL-1 Blue cells. All the three fluorescent constructs display an uneven distribution of the fluorescence throughout the cell, with PphOx-GFP being slightly more punctuated than T1scTCR-GFP and T2scTCR-GFP.

In summary, the results so far show that the DsbA leader constructs were expressed at very low levels and these were therefore not investigated further. Both the phOx- and NIP-GFP fusions with pelB leader were detected as cleaved protein in the periplasm. However, of the two, the phOx fusion showed the most desirable expression characteristics and were examined further, whereas the NIP-GFP construct was excluded from the rest of this study.

Despite fluorescence of both phOx-GFP and scTCR-GFP with TorA leaders, the subcellular analysis showed that phOx-GFP is cleaved when translocated and the scTCR-GFP constructs were only localized in the cytosol. Hence, optimalization is needed to produce functional full length fusion proteins.

Table 8: Fluorescence imaging of constructs expressed in XL-1 Blue using an Axioplan 2 Imaging fluorescence microscope (Zeiss) with filter 8 for fase contrast pictures (20 ms) and filter 2 (FITC-filter) for fluorescence pictures (100 ms).

Phase contrast 20 ms	FITC filter 100 ms	Phase contrast 20 ms	FITC filter 100 ms
PphOx-GFP		PNIP-GFP	

| PscTCR-GFP | | T1scTCR-GFP | |

| T2scTCR-GFP | | D1scTCR-GFP | |

| D2scTCR-GFP | | D3scTCR-GFP | |

| XL-1 Blue | | | |

4.5 Effect of chaperones

4.5.1 Over-expression of GroEL-GroES

To examine the effect of over-expressing cytosolic folding assistance, expression was performed in XL-1 Blue pGro7 which harbors inducible over-expression of the cytosolic chaperones GroEl-GroES (Nishihara, Kanemori et al. 1998).

The TorA leader directs folded or partially folded domains to the Tat-pathway for translocation to the periplasm. Hence, TorA constructs are putative substrates of GroEL-GroES. Therefore, phOx-GFP and scTCR-GFP fusions with pelB, TorA1 and TorA2 leaders were expressed in XL-1 Blue pGro7. Proteins were expressed as previously described and performed in two parallels to test two different induction start points of GroEL and GroES expression: A – induction when ONC was reinoculated for expression and B – induction simultaneously as IPTG induction of fusion protein expression. Empty XL-1 Blue pGro7 cells were grown in triplicate as controls and validation of GroEL-GroES expression with induction A, B and an additional C without induction.

Validation of GroEL-GroES over-expression

To validate GroEL over-expression, analysis was performed of a coomassie stained SDS-PAGE gel and a western blot detected with anti-GroEL Ab.

The coomassie stained SDS-PAGE gel showed clear bands of GroEL expression (~60kDa) in the cytosolic fractions when expression was induced (A and B), but only endogenous chaperone levels without induction (C) (shown in Figure 12 A). The bands and expression levels were validated as GroEL by western blot analysis by detected with anti-GroEL Ab, hence confirming the chaperone over-expression (shown in Figure 12 B).

The low levels of GroEL detected in the medium and periplasmic fractions originates from cell lysis and overflow during the sample application to the gel. However, due to high sensitivity of the detection it represents minute amounts of protein.

Effect of GroEL-GroES

To validate effect of GroEL-GroES over-expression on periplasmic yield of scTCR- and phOx-GFP fusions with pelB, TorA1 and TorA2 leader were expressed in XL-1 Blue with over-expression of the chaperones. Normalized cell samples were fractionated and analyzed by SDS-PAGE and western bloting with anti-His$_6$ and anti-GFP detection.

Figure 13 shows that the scTCR-GFP constructs were not transported to the periplasm with pelB or TorA leaders when GroEL-GroES is over-expressed, as no bands are visible. GroEL-GroES had a negative effect on transport of phOx-GFP with pelB leaders as no periplasmic protein is detected. However, bands of cleaved periplasmic fusion protein at ~32kDa was detected for phOx-GFP with TorA1 and TorA2 leaders, as previously observed with pelB leader without chaperone over-expression. Hence, GroEL-GroES had no effect on periplasmic expression of scTCR-GFP fusions, independent of leader, or phOx-GFP fusions with TorA leaders. However, over-expression of GroEL-GroES seems to impede the transport of phOx-GFP fusions with pelB leader to the periplasm.

Collectively, the data so far showed a more favorable expression profile for the TorA2 leader constructs as compared to the TorA1 leader constructs. Thus, the TorA1 constructs were excluded from further analysis.
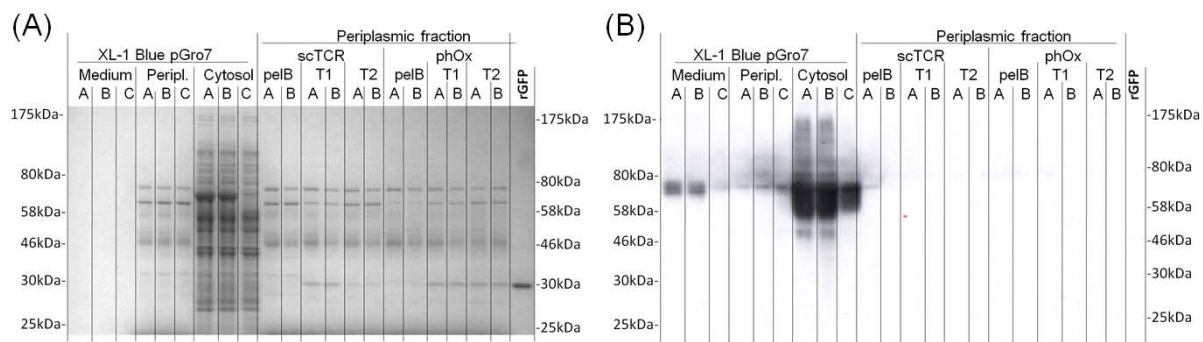
Figure 12: Coexpression of GroEL in XL-1 Blue pGro7 cells. The GroEL/ES expression is induced with L-arabinose at two different timepoints: A – when ONC is reinoculated to start the expression culture (OD600$_{nm}$~0.1), B – simultaneously as IPTG induction. C is without induction. (A) Fractions of XL-1 Blue pGro7 cells separated on SDS-PAGE and coomassie stained gel. (B) Fractions of XL-1 Blue pGro7 cells separated on SDS-PAGE and detection on western blot with anti-GroEL.
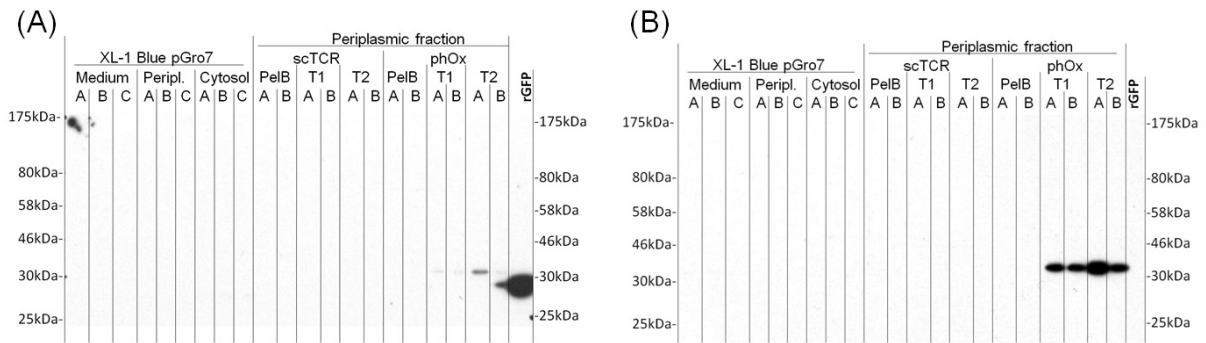
Figure 13: Expression in XL-1 Blue with coexpression of GroEL/ES, periplasmic fractions. The GroEL/ES expression is induced with L-arabinose at two different timepoint: A – when ONC is reinoculated to start the expression culture (OD600$_{nm}$~0.1), B – simultaneously as IPTG induction. C is without induction. (A) Fractions separated on SDS-PAGE and analyzed by western blot with anti-GFP. (B) Fractions separated on SDS-PAGE and analyzed by western blot using anti-His$_6$.

4.5.2 Over-expression of FkpA

Over-expression of the periplasmic chaperone FkpA rescues soluble periplasmic expression of the scTCR clone in this study (Gunnarsen, Lunde et al. submitted). To investigate the effect of FkpA over-expression, phOx- and scTCR-GFP fusions with pelB and TorA2 leader were expressed with and without FkpA over-expression.

Effect of FkpA on total expression level

To investigate the effect of FkpA over-expression on total expression level, normalized cell samples from protein expression in XL-1 Blue were analyzed using SDS-PAGE and Western blot as shown in Figure 14.

Expression of phOx-GFP with pelB leader yielded similar expression levels independent of FkpA over-expression. However, expression of phOx-GFP with TorA2 leader showed a reduced expression level when FkpA was over-expressed. Furthermore, expression of scTCR-GFP with either pelB or TorA leader showed litte effect on total expression level when FkpA was over-expressed. Hence, FkpA seems to have no effect on total expression level, with one exception where the effect was negative.
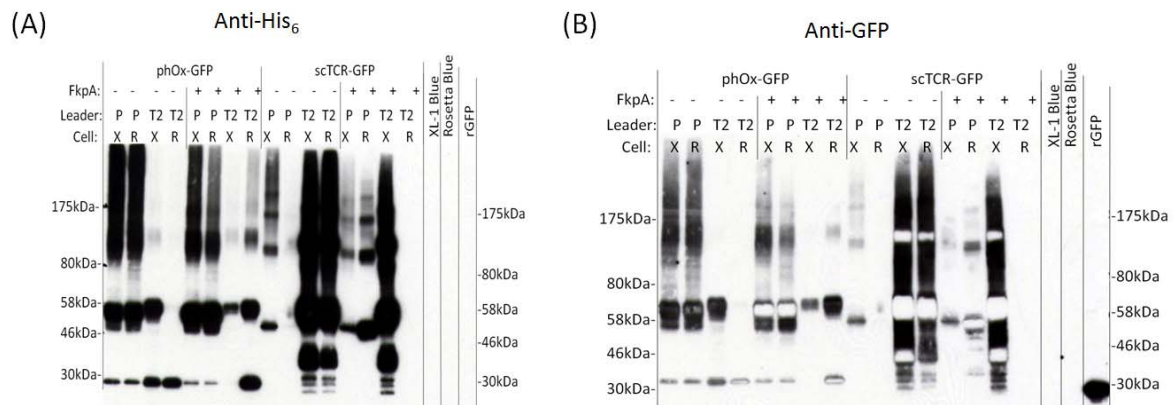
Figure 14: SDS-PAGE and Western blot of normalized cell samples from expression of phOx- and scTCR-GFP fusions with PelB (P) and TorA2 (T2) leader, with and without FkpA co-expression, in XL-1 Blue (X) and Rosetta Blue (DE3) pLysS (R). (A) Developed with anti-His$_6$. (B) Developed with anti-GFP.
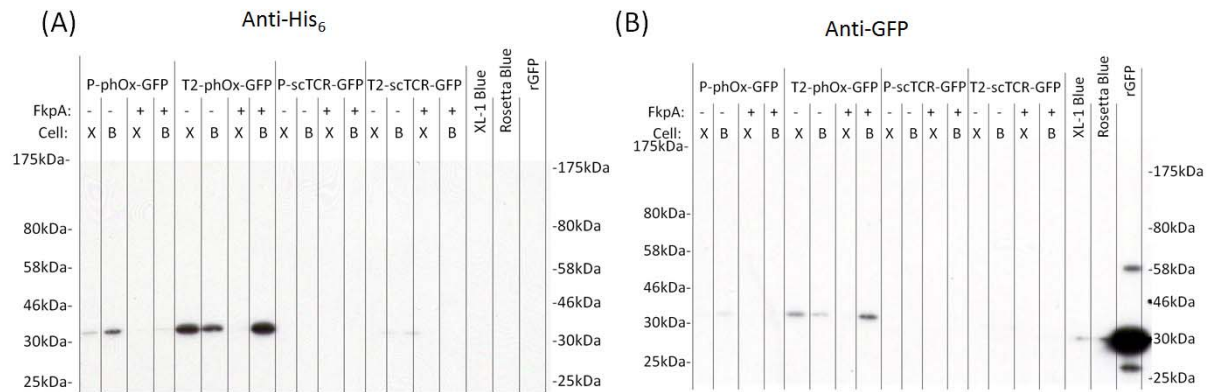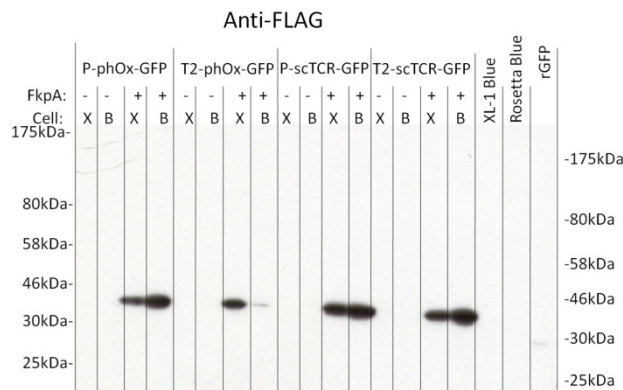

Figure 15: SDS-PAGE and western blot of normalized periplasmic fractions from expression of phOx- and scTCR-GFP fusions with pelB and TorA2 leader, with and without FkpA over-expression, in XL-1 Blue and Rosetta Blue (DE3) pLysS. (A) Detected with anti-GFP. (B) Detected with anti-His$_6$.

Figure 16: SDS-PAGE and western blot of normalized periplasmic fractions from expression of phOx- and scTCR-GFP fusions with pelB and TorA2 leader, with and without FkpA over-expression, in XL-1 Blue and Rosetta Blue (DE3) pLysS, detected with anti-Flag M2-HRP Ab.

Effect of FkpA on subcellular localization

To further investigate the effect of FkpA over-expression on subcellular localization, normalized periplasmic fractions from protein expression were analyzed using SDS-PAGE and western blot, as shown in Figure 15.

Expression of scTCR-GFP fusions did not yield periplasmic protein regardless of leader and over-expression of FkpA. Expression of phOx-GFP with pelB and TorA leaders resulted in detection of a cleaved product ~32 kDa in size in the periplasm, similar to what have been observed previously. However, the over-expression of FkpA reduced the periplasmic yield of both these construct. Periplasmic over-expression of FkpA was validated in Figure 16 for the constructs expressed in XL-1 Blue.

In summary, neither the scTCR-GFP nor the phOx-GFP construcs yielded periplasmic expression of full length fusion protein.

Effect of FkpA on relative fluorescence

To further investigate the effect of FkpA on the fluorescent properties of the constructs in XL-1 Blue, cells were washed and relative fluorescence measured.

Figure 17 shows the effect of FkpA on relative fluorescence in cells. For PscTCR-GFP and T2scTCR-GFP expression, the measurements showed similar levels of relative fluorescence with and without FkpA and therefore it was reasonable to believe that co-expression of FkpA had no effect these two construct. T2phOx-GFP expression in XL-1 Blue showed a more than four-fold reduction in fluorescence when FkpA is co-expressed. Expression of PscTCR-GFP gave a moderate increase in relative flourescende.

Effect of FkpA on fluorescence imaging

Fluorescence imaging of the constructs showed that FkpA did not notably increase the fluorescence or alter the properties in XL-1 Blue, as seen in Table 9.
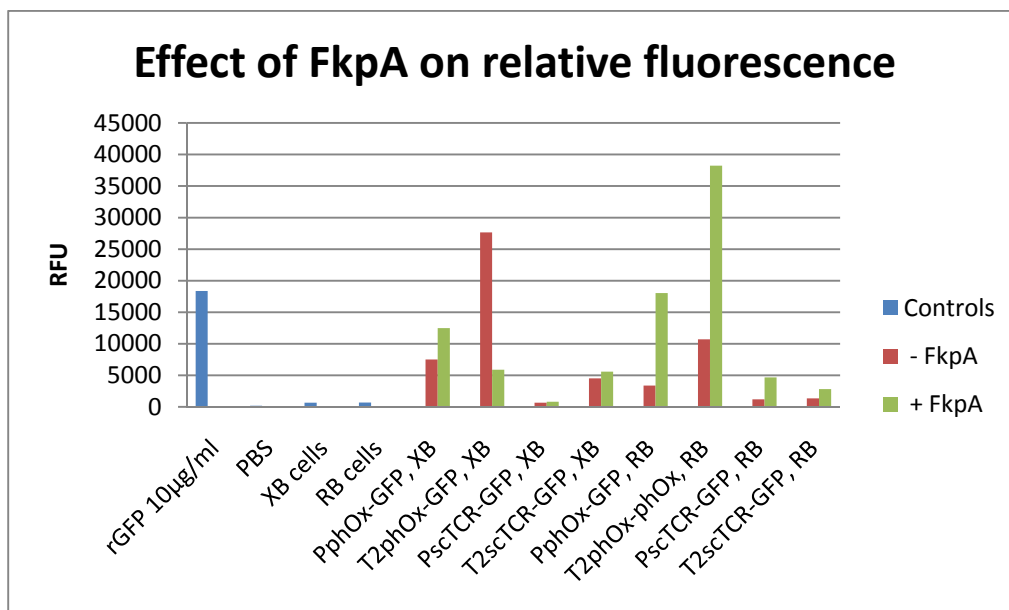
Figure 17: The relative fluorescence of normalized samples of cells, washed in PBS and analyzed in black 96-well microtiter plates using a Victor 1420 Multilabel counter (Perkin Elmer) using the Fluorometry protocol at 485nm/535nm 1.0 sec. The effect of co-expression of FkpA for each construct in each expression host, comparing with and without FkpA. XB – XL-1 Blue, RB – Rosetta Blue.
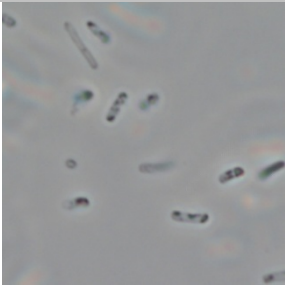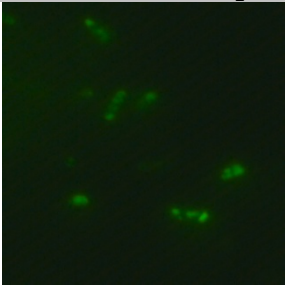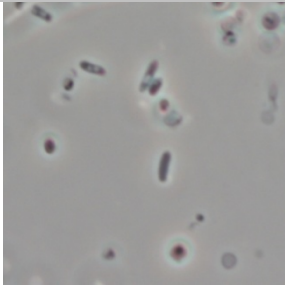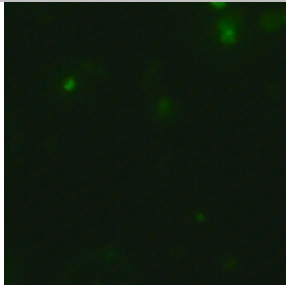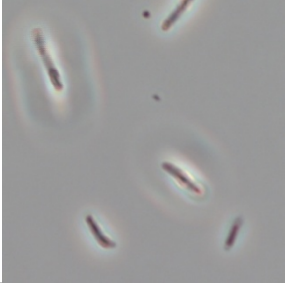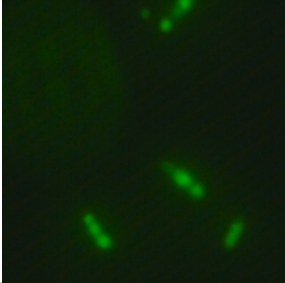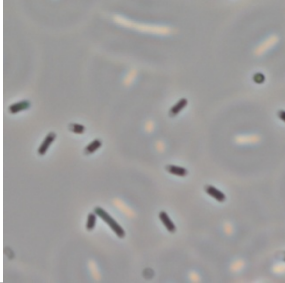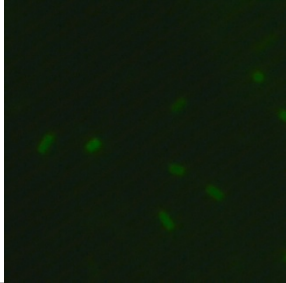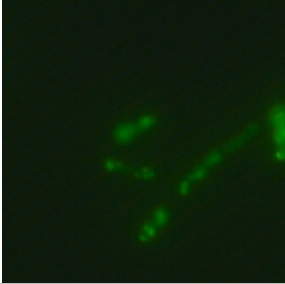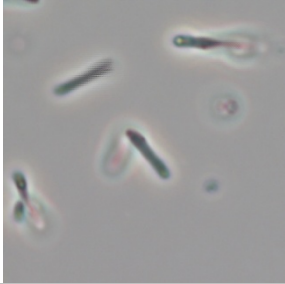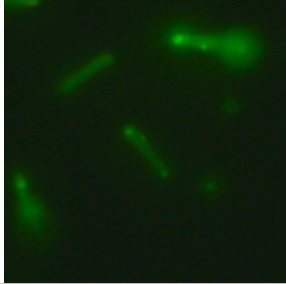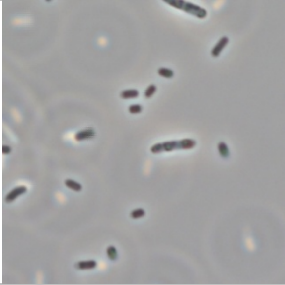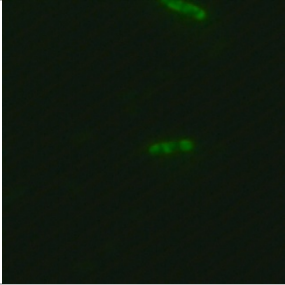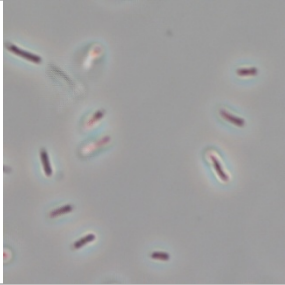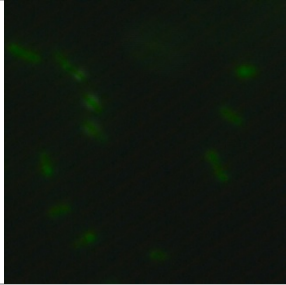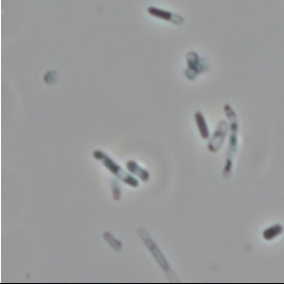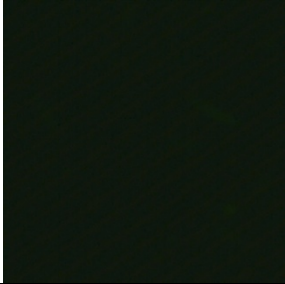
4.6 Effect of expression host

Expression of heterologous proteins in *E. coli* is challenging for many reasons and using host strains customized for protein expression has been one strategy.

4.6.1 Expression in Rosetta Blue (DE3) pLysS

Rosetta Blue (DE3) pLysS is a strain that contains extra tRNA genes for expression of proteins with codons rarely used in *E. coli*. To investigate the effect of extra tRNAs, phOx- and scTCR-GFP fusions with pelB and TorA2 leader were expressed with and without FkpA over-expression, were expressed in XL-1 Blue and Rosetta Blue (DE3) pLysS.

Table 9: Fluorescence imaging of constructs expressed in XL-1 Blue or Rosetta Blue, with and without FkpA.

XL-1 Blue		Rosetta Blue	
Fase contrast	**FITC filter**	**Fase contrast**	**FITC filter**
PphOx-GFP			
T2phOx-GFP			
PphOx-GFPFN			
T2phOx-GFPFN			
PscTCR-GFP			

T2scTCR-GFP

PscTCR-GFPFN

T2scTCR-GFPFN

Cells

Effect on total expression level

To examine the effect of extra tRNAs on total expression level, normalized cell samples from protein expression were analyzed using SDS-PAGE and Western blot as shown in Figure 14.

The phOx-GFP fusions with pelB leader showed similar total expression levels in XL-1 Blue and Rosetta Blue, independent of FkpA over-expression. Expression of phOx-GFP with TorA2 leader yielded a higher expression level in XL-1 Blue without FkpA, and contrary, a higher expression level in Rosetta Blue with FkpA.

For scTCR-GFP with pelB leader without FkpA, XL-1 Blue yielded a higher expression level than Rosetta Blue. However, scTCR-GFP with pelB leader and FkpA, showed a higher expression level in Rosetta Blue than in XL-1 Blue. Expression of scTCR-GFP with TorA2 leader without FkpA showed similar expression levels in both strains. In contrast, expression of scTCR-GFP with TorA2 leader with FkpA showed a higher expression level in XL-1 Blue as no potein was detected in Rosetta Blue.

In summary, expression in a host strain providing extra tRNAs for compensation of codon bias showed a variety of effects, depending on scaffold, leader and over-expression of FkpA.

Effect on subcellular localization

To further investigate the effect of extra tRNAs on subcellular localization, normalized periplasmic fractions from protein expression were analyzed using SDS-PAGE and western blot, as shown in Figure 15. Expression of scTCR-GFP fusions did not yield periplasmic protein regardless of leader, over-expression of FkpA and extra tRNAs.

The effect of extra tRNAs on the periplasmic expression of phOx-GFP with pelB leader did not significantly alter the periplasmic expression levels, with or without FkpA. Also the periplasmic expression of phOx-GFP with TorA2 leader without FkpA showed similar periplasmic expression levels with and without extra tRNAs. However, the periplasmic expression of phOx-GFP with TorA2 leader with FkpA showed a significant increase when expressed with extra tRNAs. Further investigation of FkpA expression in Figure 16 showed that the periplasmic FkpA level was very small, hence the two constructs were not directly comparable.

Effect on relative fluorescence in cells

To further investigate the effect of extra tRNAs on the fluorescent properties of the constructs, cells were washed and relative fluorescence measured. Figure 18 showed that phOx-GFP with pelB and TorA2 leader, without FkpA, showed reduced fluorescence when expressed in Rosetta Blue with extra tRNAs. In contrast, the same constructs with FkpA showed an increase in fluorescence. However, the T2phOx-GFPFN construct in Rosetta Blue

did not show over-expression of FkpA in the periplasm and can not be directly compared to the same construct expressed in XL-1 Blue.

 The scTCR-GFP fusion proteins did not show any great change in relative fluorescence, with the exception of scTCR-GFP with TorA2 leader which showed a reduced relative fluorescence when expressed with extra tRNAs.
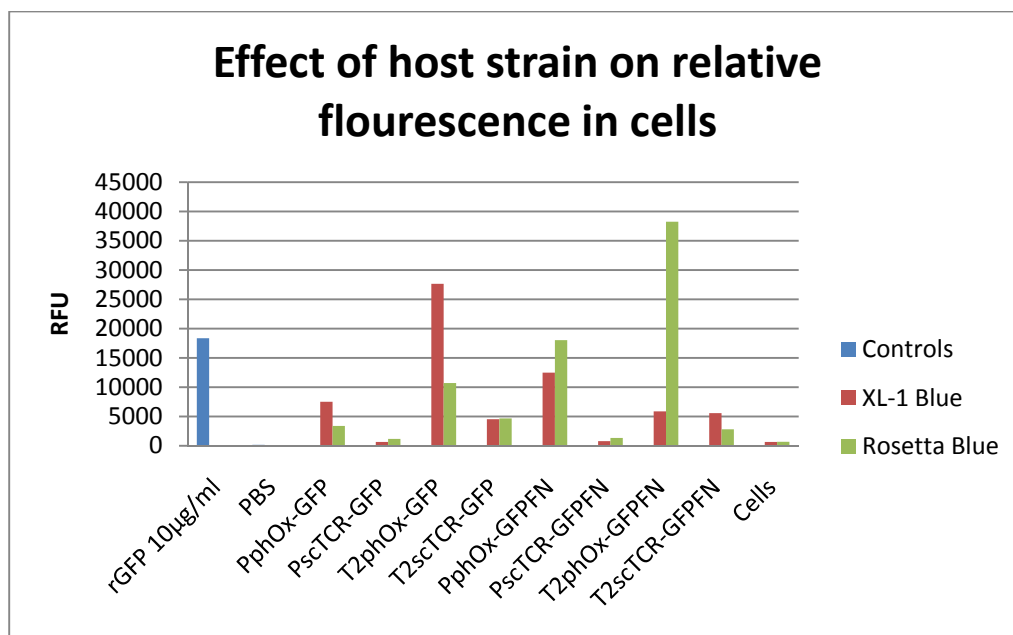
Figure 18: The relative fluorescence of normalized samples of cells, washed in PBS and analyzed in black 96-well microtiter plates using a Victor 1420 Multilabel counter (Perkin Elmer) using the Fluorometry protocol at 485nm/535nm 1.0 sec. The effect of host strain during expression with each construct expressed in XL-1 Blue put up agains expression in Rosetta Blue.

In Rosetta Blue, both phOx-GFP and scTCR-GFP, with pelB and TorA2, gave an increased relative fluorescence with the over-expression of FkpA.

Effect on fluorescence imaging

Fluorescence imaging of the constructs are presented in (Table 9). The images does not show any big difference in the fluorescent properties when extra tRNAs are expressed with the exception of T2phOx-GFP which show that more cells exhibit fluorescence when extra tRNAs are expressed. However, this is the previously mentioned construct which is not directly comparable due to lack of FkpA over-expression. In summary, expression of extra tRNAs did not yield periplasmic full length fusion proteins of neither phOx-GFP nor scTCR-GFP constructs.

4.6.2 Expression in BL21

To further investigate the possibility of obtaining full length periplasmic fusion protein, phOx-GFP with TorA2 leader was expressed in the host strain BL21 which is deficient in the *lon* and *ompT* proteases.

The expression of phOx-GFP with TorA2 leader in BL21 showed a decrease in total expression level and periplasmic expression level compared to XL-1 Blue, as shown in Figure 19 and Figure 20. Further, there was a decrease in relative fluorescence when expressed in BL21, as shown in Figure 21. Hence, no full length periplasmic fusion protein was obtained from expression in BL21.

4.7 Effect of induction and expression temperature

4.7.1 Induction versus no induction at 30°C

Induction of protein expression has been shown to give increased periplasmic yield of scTCR (Gunnarsen, Lunde et al. submitted). Therefore, the fusion proteins were expressed in XL-1 Blue with and without induction to investigate the effect of over-expression of fusion proteins versus expression at basal transcription, respectively.

Effect on total expression level

To investigate the effect of induction on total expression level, phOx-GFP with TorA2 and PelB leader, both with and without FkpA was expressed in XL-1 Blue with and without induction. Normalized cell samples from protein expression were analyzed using SDS-PAGE and Western blot.

As seen in Figure 19, induction did not significantly affect the total expression level of any of the four constructs when expressed at 30°C, the standard expression temperature.
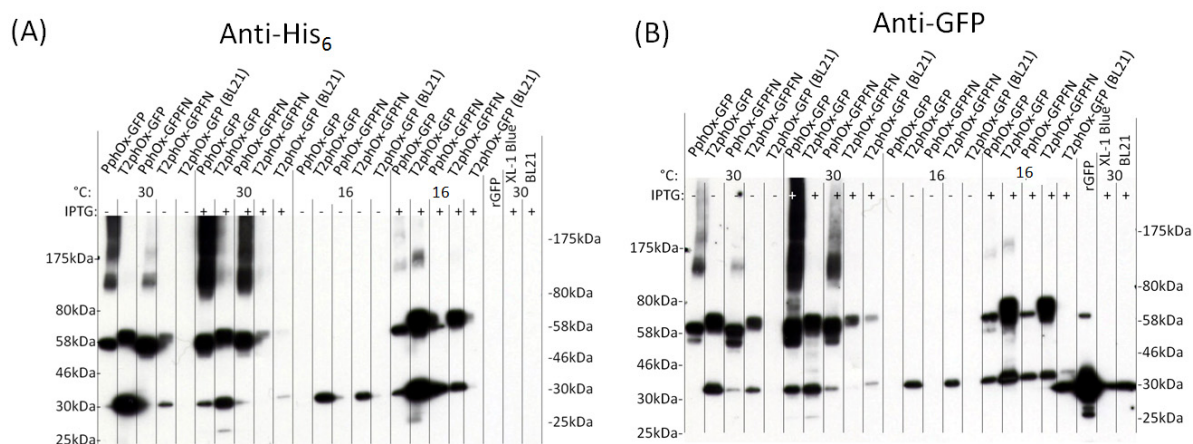
Figure 19: SDS-PAGE and Western blot of normalized cell samples from expression of of PphOx-GFP and T2phOx-GFP with and without FkpA in XL-1 Blue and T2phOx-GFP without FkpA in BL21 (A) Detection with anti-His₆ and anti-mouse IgG-HRP Ab from sheep. (B) Detection with anti-GFP and SA-HRP.

Effect on subcellular localization

To further examine the effect of induction of subcelluar localization, cells expressed at 30°C were fractionated and analyzed by SDS-PAGE and western blot, as shown in Figure 20.

No full length protein was obtained for any of the constructs. However, when phOx-GFP was expressed with pelB, cleaved fusion proteins were mainly detected with induction, independen of FkpA. The pelB constructs that was not induced show a reduction of periplasmic cleaved protein.

When phOx-GFP was expressed with TorA leader without FkpA, induction did not significantly affect the yield of cleaved protein in the periplasm. However, when expressed with FkpA, phOx-GFP with TorA leader showed a reduced yield of cleaved protein in the periplasm when not induced.

Effect on relative fluorescence in cells

The effect of induction was further investigated by measuring the relative fluorescence of cells with and without induction, as shown in Figure 21. Induction of constructs gave the highest relative fluorescence for phOx-GFP with pelB leader, both with and without FkpA. Expression of phOx-GFP with TorA leader, both with and without FkpA, obtained the highest relative fluorescence when not induced.
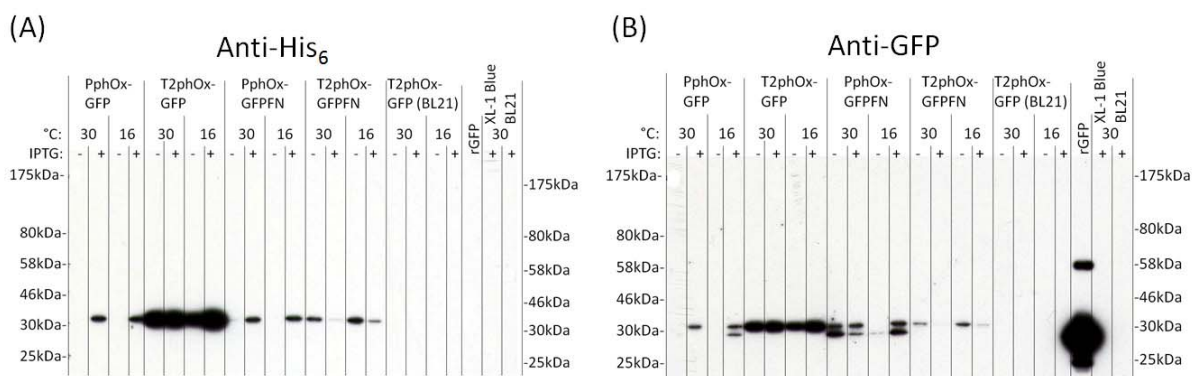
Figure 20: SDS-PAGE and Western blot of normalized periplasmic fractions from expression of PphOx-GFP and T2phOx-GFP with and without FkpA in XL-1 Blue and T2phOx-GFP without FkpA in BL21. (A) Detection with anti-His$_6$. (B) Detection with anti-GFP.
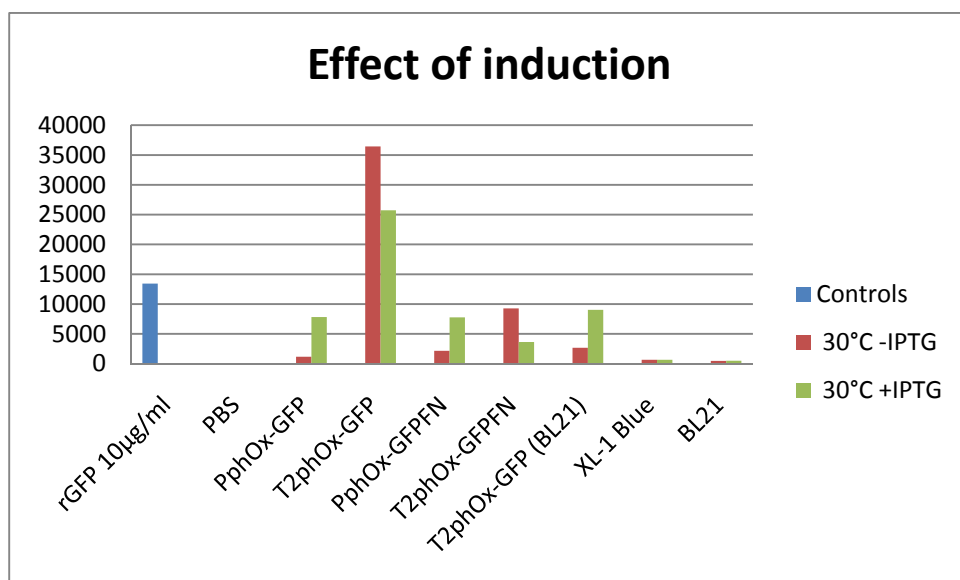
Figure 21: Effect of induction on relative fluorescence in cells.

Effect on fluorescence imaging

The effect of induction on the fluorescent properties of the cells was examined and the images are shown in Table 10. Induction seemed to be necessary for the phOx-GFP with pelB leader, with and without induction, to obtain fluorescent cells. None of the other constructs showed altered fluorescence.

In summary, induction at standard expression temperature 30°C did not significantly alter the total expression level. Furthermore, not inducing over-expression of fusion protein did not yield full length periplasmic fusion protein.

4.7.2 Expression at 30°C versus 16°C

A previous study has shown successful expression of periplasmic full length scFv-GFP fusion when performing the expression at 16 °C for 48 hours(Griep, van Twisk et al. 1999). Therefore, expression of the phOx-GFP fusions was tested by comparing 30°C ON to 16°C for 48 hours.

Effect on total expression level

Normalized samples of cell culture from protein expression were analyzed using SDS-PAGE and western blot as shown in Figure 19.

The effect of total expression level is shown in Figure 19. Full length fusion proteins of phOx-GFP with PelB and TorA leader, with and without FkpA over-expression, were detected when expressed at 30°C with and without IPTG induction, and when expressed at16°C with induction only. Cleaved fusion proteins were detected when expressed with a TorA leader without induction, at both 16°C and 30°C. When inducing at 16°C, cleaved fusion protein was detected regardless of leader and FkpA over-expression. When inducing expression at 30°C, phOx-GFP with pelB leader, with and without FkpA, was detected at cleaved fusion protein size. Expression of phOx-GFP with TorA2 leader with induction at 30°C was detected as a cleaved fusion protein without FkpA but not with FkpA over-expression.

Figure 19 B showed that application of the rGFP control to the SDS-PAGE gel gave overflow to the neighboring wells as a clear band is detected at ~30 kDa for T2phOx-GFP (BL21) at 16°C with induction, empty XL-1 Blue and empty BL21 cells. Due to limited number of wells in the gel, all the controls with XL-1 Blue and BL21 growth at 30°C and 16°C, with and without induction, with whole cell samples, periplasmic fractions and cytoplasmic fractions were analyzed by SDS-PAGE gels and western bloting as done for the expressions cultures. These western blots showed no protein expression with detection by anti-His$_6$ or anti-GFP.

Effect on subcellular localization

To further investigate the effect of incubation temperature, cells were fractionated and the periplasmic fractions analyzed using SDS-PAGE and Western blot, as shown in Figure 20.

For the constructs expressed, no full length fusion protein was detected in Figure 20. Without FkpA expression, phOx-GFP with TorA2 leader was detected as cleaved fusion protein at ~32 kDa at relatively high yields independent of expression temperature and induction. With FkpA over-expression, phOx-GFP with TorA2 leader was detected as cleaved fusion protein at ~32 kDa whithout induction at 30°C, and both with and without induction at 16°C. As previously observed, the expression level of phOx-GFP with TorA2 leader was reduced when FkpA is over-expressed. When phOx-GFP was expressed with pelB leader, cleaved fusion protein at ~32 kDa was detected when the cultures were induced, at 30°C and 16°C, and with and without FkpA. phOx-GFP expression with the pelB leader and without induction gave detection of cleaved fusion protein at ~32 kDa when FkpA was over-expressed at 30°C. Interestingly, when phOx-GFP with pelB leader was detected with anti-GFP in Figure 20 B, most of the detected cleaved protein was detected with doublet bands at ~30 kDa and ~32 kDa while the anti-His$_6$ detection only reveal one band.

Figure 22 A shows overexposure of the blot of periplasmic fractions detected with anti-His$_6$. The only band of full length fusion protein at ~60 kDa was detected for phOx-GFP with pelB leader expressed with FkpA over-expression at 30°C without induction. Under these conditions there was barely any detected cleaved fusion protein at 30 kDa.

The His$_6$-tag is located *C*-terminal in the fusion protein downstream of the GFP. The detection with anti-GFP and anti-His therefore only verify the presence of the *C*-terminal domain of the fusion protein. To detect the *N*-terminal domain of the scFv anti-phOx, the blot was incubated with rabbit anti-human lambda L-chain and donkey anti-rabbit IgG HRP and presented in Figure 22 B. Surprisingly, there was detection of bands at ~30 kDa when the phOx-GFP was expressed with pelB leader and not when expressed with TorA2 leader. When expressed with pelB leader, the scFv anti-phOx *N*-terminal entity was detected both with and without FkpA and when expressed at both 30°C and 16°C, with and without induction. This was surprising since the detection of the *C*-terminal entity with Abs in Figure 20 shows

highest expression levels with TorA2 leader. The expression of FkpA was verified by detection with anti-Flag M2-HRP in Figure 23.
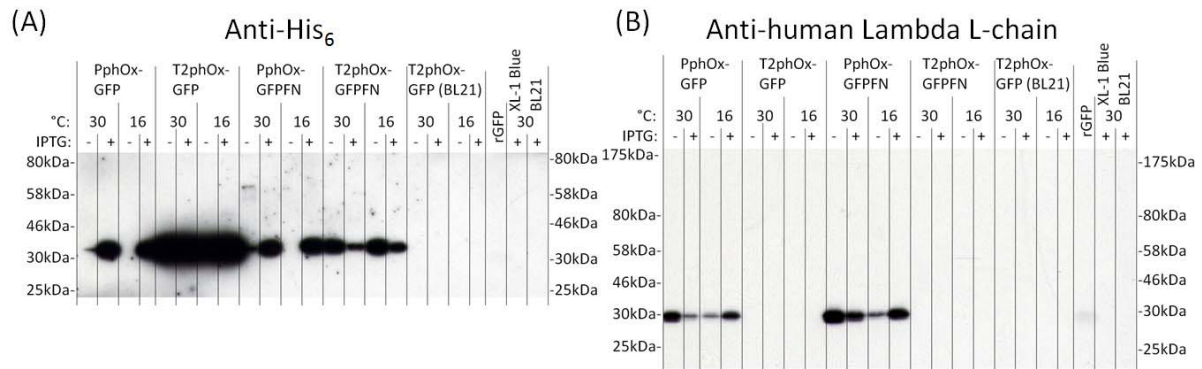

Figure 22: SDS-PAGE and Western blot of normalized periplasmic fractions from expression of PphOx-GFP and T2phOx-GFP with and without FkpA in XL-1 Blue and T2phOx-GFP without FkpA in BL21. (A) Detection with anti-His$_6$ and anti-mouse IgG-HRP Ab from sheep. (B) Detection with rabbit anti-human lambda L-chain of the scFv anti-phOx.
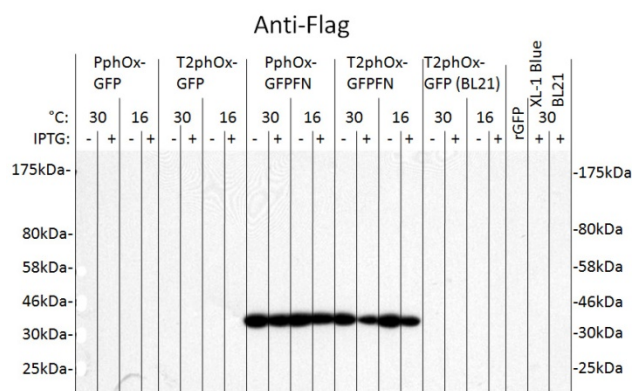
Figure 23: SDS-PAGE and Western blot of normalized periplasmic fractions from expression of PphOx-GFP and T2phOx-GFP with and without FkpA in XL-1 Blue and T2phOx-GFP without FkpA in BL21. Detection of FkpA with anti-Flag M2-HRP.

Effect on relative fluorescence in cells

The effect of temperature on the fluorescent properties was further investigated by measuring the relative fluorescence of cells expressed at 16°C and 30°C. Effect of temperature on fluorescence of cells expressed without induction is shown in Figure 24 and with induction is shown in Figure 25.

Figure 24 shows that without induction, reducing the temperature had little or no effect on the relative fluorescence, with the exception of phOx-GFP with TorA2 leader which shows that expression at 30°C results in four-fold increase in relative fluorescence.

Figure 25 shows that phOx-GFP with TorA2 leader without FkpA, and phOx-GFP with both leaders with FkpA showed an increase in relative fluorescence when the expression temperature was reduced to 16°C.

Effect on fluorescence imaging

To further examine the fluorescent properties of expression of phOx at different temperatures, cells were analyzed in a fluorescence microscopy. The images in Table 10and Table 11, visualizes the differences measured by relative fluorescence.
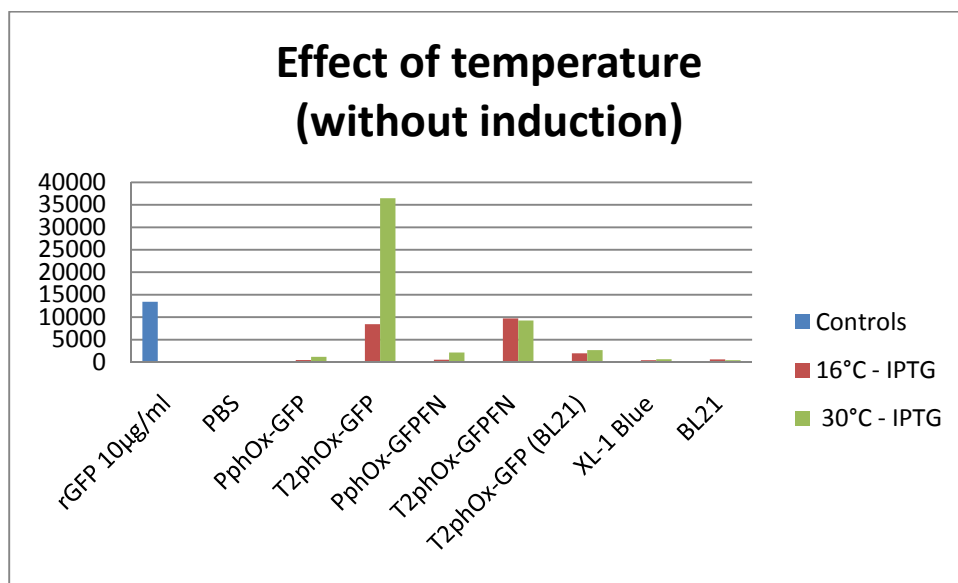
Figure 24: Effect of temperature on relative fluorescence in cells without induction.
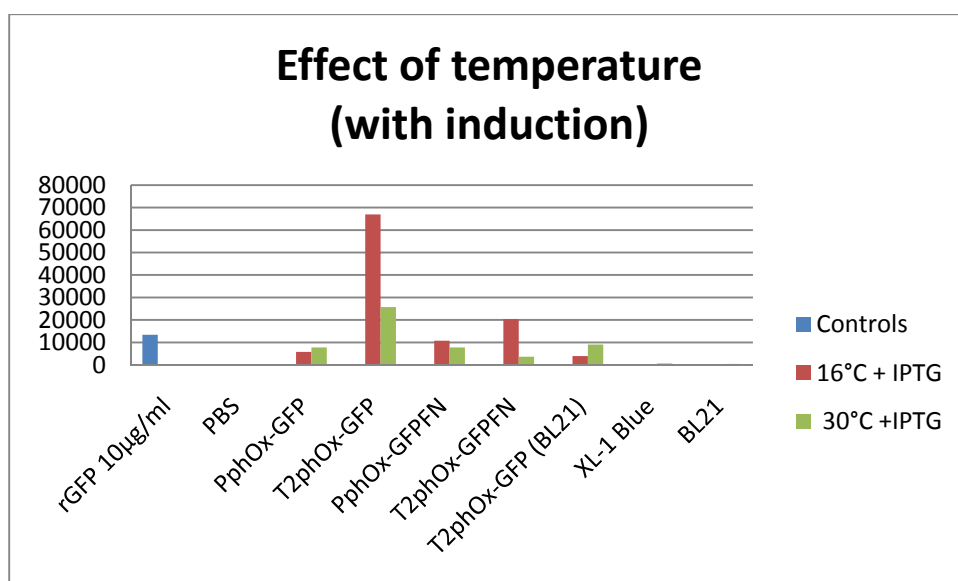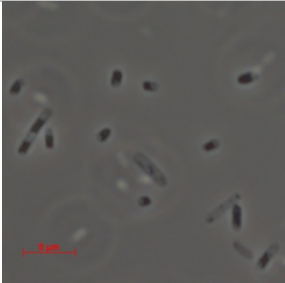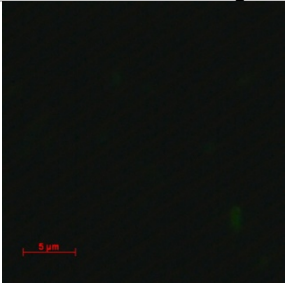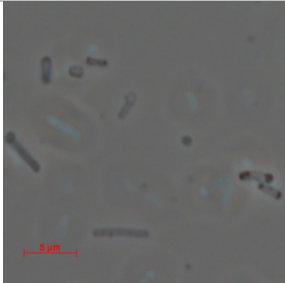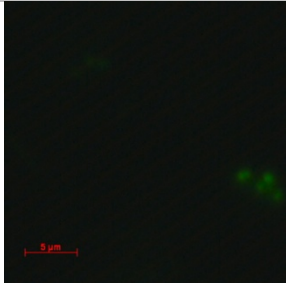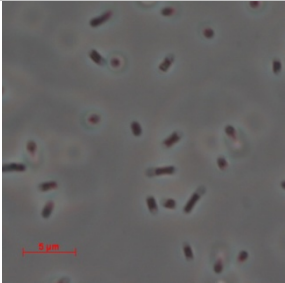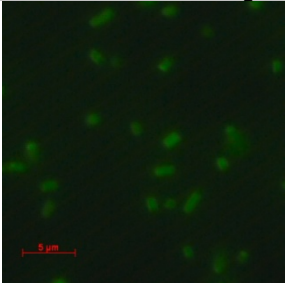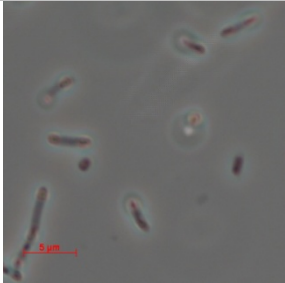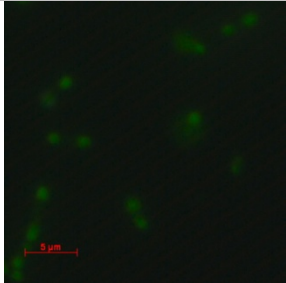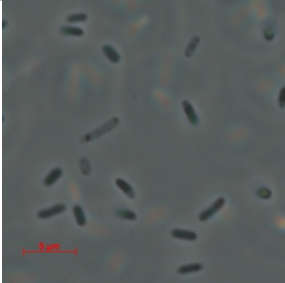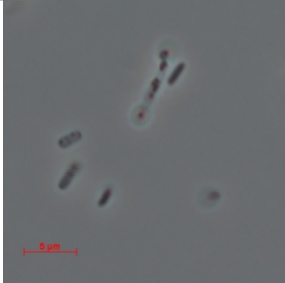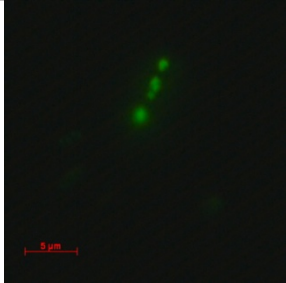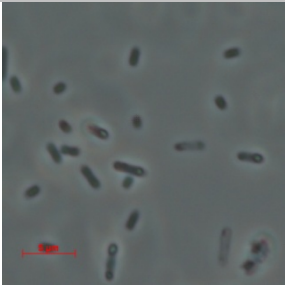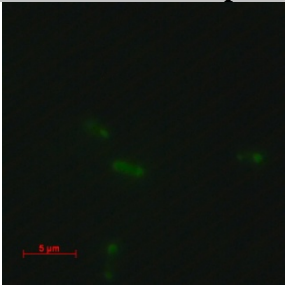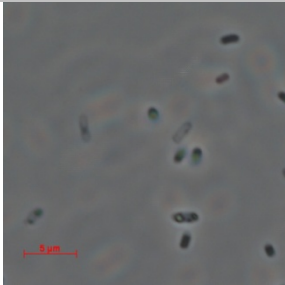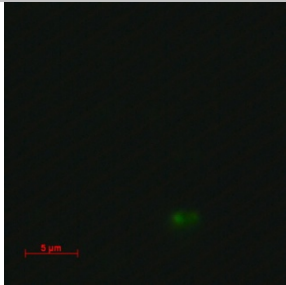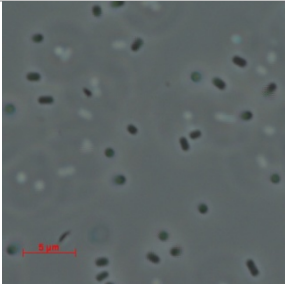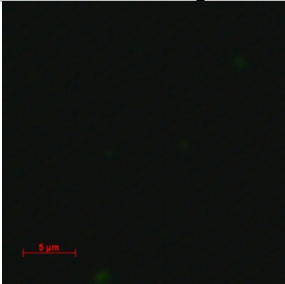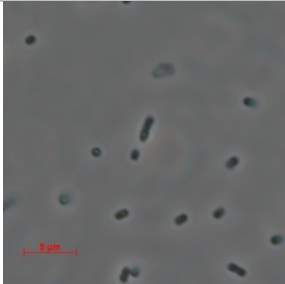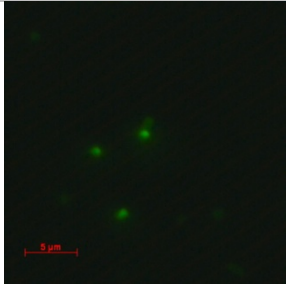
Figure 25: Effect of temperature on relative fluorescence in cells with induction.

Table 10: Fluorescence microscopy of cells expressed at 30°C.

Without IPTG		With IPTG	
Fase contrast	**FITC filter**	**Fase contrast**	**FITC filter**
PphOx-GFP			
T2phOx-GFP			
PphOx-GFPFN			
T2phOx-GFPFN			
T2phOx-GFP (BL21)			

XL1-Blue

BL21

Table 11: Fluorescence microscopy of cells expressed at 16°C.

Without IPTG		With IPTG	
Fase contrast	**FITC filter**	**Fase contrast**	**FITC filter**
PphOx-GFP			
T2phOx-GFP			
PphOx-GFPFN			
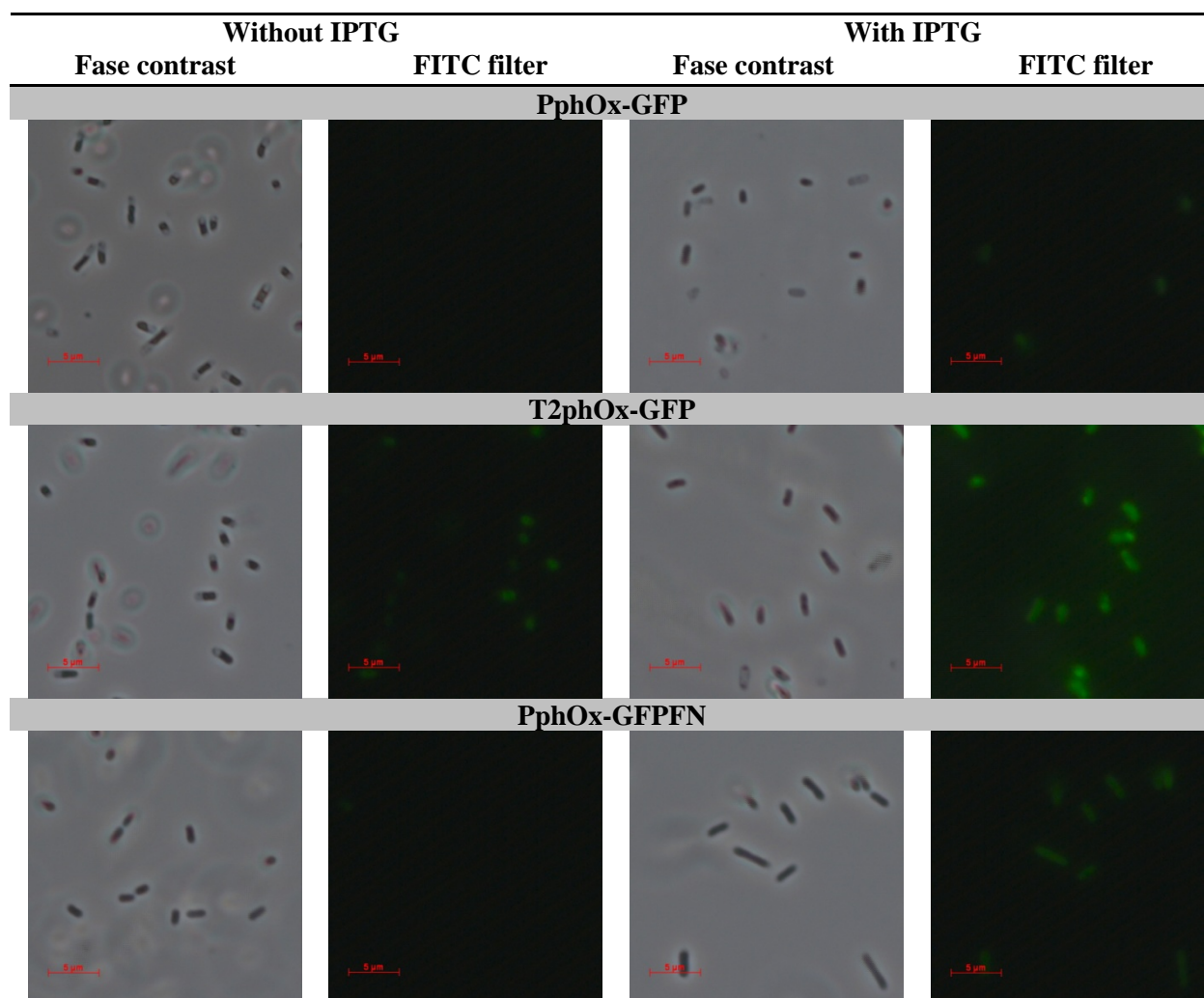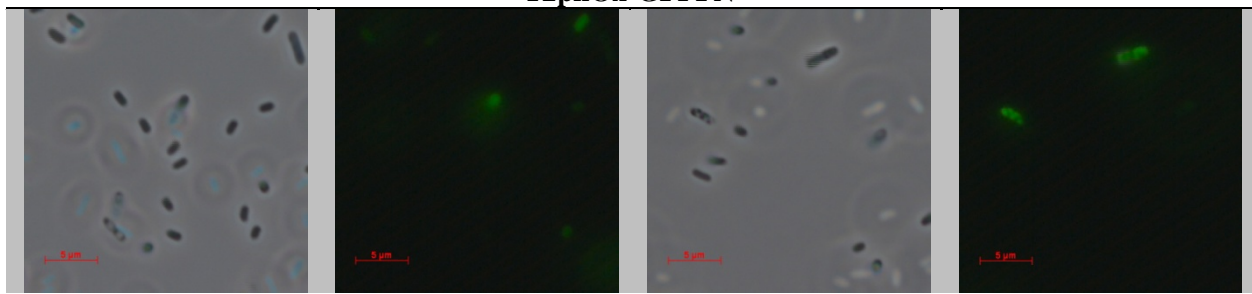
T2phOx-GFPFN

T2phOx-GFP (BL21)
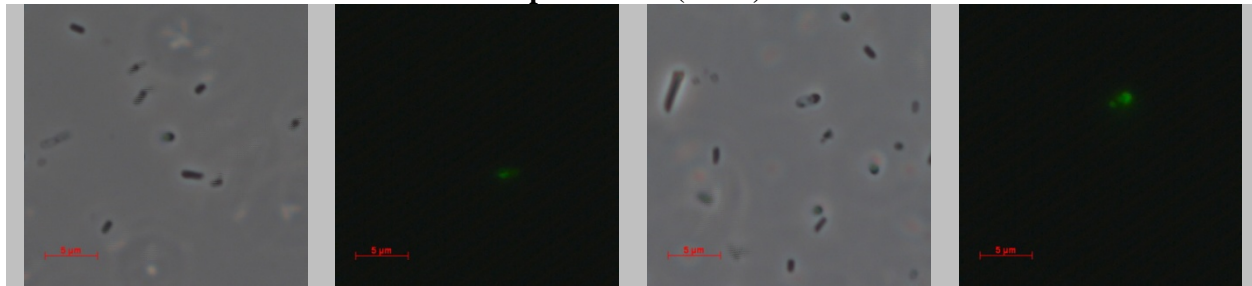
XL1-Blue
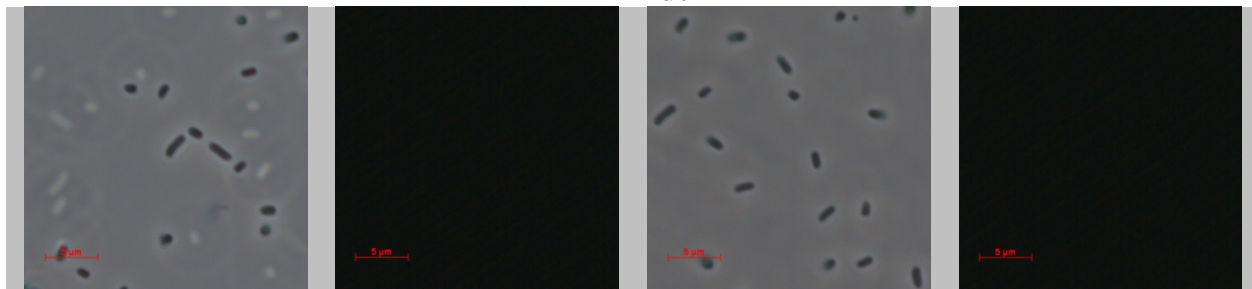
BL21
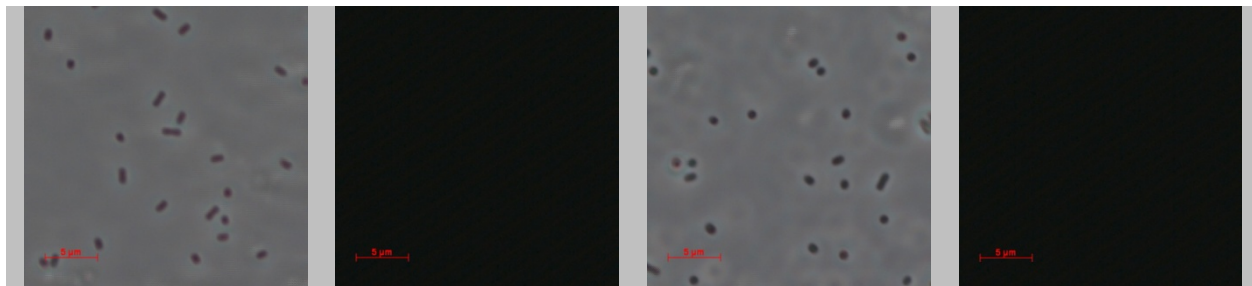
5 Discussion

Development of novel protein based diagnostics and therapeutics rely heavily on molecular evolution of polypeptides with desired properties. After generation of molecular libraries, selection and screening of candidate clones are performed. It is crucial to screen a large number of clones to identify those with desired properties. However, large scale screening is both expensive and time-consuming, even with access to semi-automated methods. Hence, screening becomes the bottleneck in selecting candidate clones.

There is a correlation between the solubility of an *N*-terminal fusion partner to GFP and fluorescence in the expressing host cell, namely *E. coli* (Waldo, Standish et al. 1999). Therefore, a feasible solution to generate a screening platform would be to genetically fuse GFP to the candidate polypeptide and take advantage of the fluorescent properties of GFP in screening. A class of molecules with great therapeutic potential is the soluble TCRs (Molloy, Sewell et al. 2005; Varela-Rohena, Molloy et al. 2008) and molecular evolution of soluble TCRs is expected to release this potential. Therefore, the TCR clone 4B2A1 from the murine MOPC315 myeloma model (Bogen, Malissen et al. 1986; Bogen, Ruffini et al. 2006) has been chosen as a fusion partner to GFP in this study. This TCR has been used in phage display (Loset, Lunde et al. 2007) and expressed in soluble form (Gunnarsen, Lunde et al. submitted). However, substantial manipulation of the expression conditions were required to allow functional expression in both cases, due to host cell toxicity and a tendency to aggregate, which is generally seen with TCRs.

Previous studies have shown expression of scFv-GFP fusions, but with very low functional yields. This appears to be the case whether expressed cytosolic (Schwalbach, Sibler et al. 2000; Oelschlaeger, Srikant-Iyer et al. 2002; Lu, Gong et al. 2005) or periplasmic, directed by a pelB leader sequence (Griep, van Twisk et al. 1999; Casey, Coley et al. 2000). Putative explanations for the low yields are that a scFv requires the oxidizing environment to fold (DeLisa, Tullman et al. 2003) or that premature folding of the GFP prior to translocation decreases the transport efficiency to the periplasm when using a leader sequence that requires post translational linear transport. Based on the structural resemblance between the scTCR and scFv formats as they both depend on the Ig fold topology (Halaby, Poupon et al. 1999), it

is reasonable to expect similar challenges for scTCR-GFP fusions. In addition, functional expression of the scTCR with a pelB leader has been found to depend on ectopic folding assistance e. g. over-expression of FkpA to obtain soluble periplasmic protein (Wulfing and Pluckthun 1994; Maynard, Adams et al. 2005; Gunnarsen, Lunde et al. submitted).

The DsbA signal sequence targets polypeptides to the periplasm through co-translational SRP-dependent secretory pathway and thereby prevents polypeptides from folding in the cytoplasm (Schierle, Berkmen et al. 2003). The TorA signal sequence targets proteins to the Tat-translocator for transport of folded or partially folded domains to the periplasm (Berks, Palmer et al. 2005). Premature folding of GFP in the cytoplasm should not impede translocation by the DsbA or TorA leader peptides. Thus, a putative solution for creating a stringent expression system of scTCR-GFP fusions would be to provide transport to the periplasm by the DsbA or TorA signal sequences. An expression system that allows periplasmic expression of protein-of-interest-GFP fusions and that also show correlation between the functional periplasmic yield and fluorescence has great potential as a screening platform. E. g. screening of the scTCR entity could be used to identify clones with improved solubility characteristics from a mutated candidate library merely by measuring cell fluorescence. Therefore, this study investigates the possibilities of scTCR-GFP expression by testing a panel of leaders for increased soluble periplasmic yield and fluorescence.

5.1 Effect of leader peptide

The expression profiles of the scTCR and scFv entities with pelB leader was investigated by expression in XL-1 Blue. We examined if the two scFvs and the scTCR kept their expression profiles when genetically fused to GFP and investigated if the scTCR-GFP expression could yield periplasmic fusion protein by utilizing three different leader peptides. All the three signal sequences have previously been successfully used for expression of recombinant protein in the periplasm (DeLisa, Samuelson et al. 2002; Thie, Schirrmann et al. 2008).

scTCR-GFP

Most striking was the difference in total and cytosolic expression level between the scTCR-GFP constructs. The DsbA leader constructs yielded extremely low protein levels compared to the pelB leader. In contrast, the TorA leader constructs gave high yields of cytosolic

protein. The only differences between these constructs are the leader sequences. Hence, the promoter should give similar transcription levels and the Shine-Dalgarno sequence should result in similar translation initiation. Therefore, it is reasonable to assume that the variations in total and cytosolic expression levels were caused by differences in the translocation pathways and how they affect folding and stability of the polypeptide.

A possible explanation for the extremely low total protein levels from the DsbA leader constructs could lie within the nature of the leader sequence: SRP-dependent co-translational transport through the Sec-pathway (Schierle, Berkmen et al. 2003). As the ribosome complex associates with the translocation units the polypeptide is directed to elongate through the transmembrane channel during translation, a failure to translocate could altogether interrupt the translation process and result in a degradation-prone polypeptide. Therefore, if the predicted transition between the leader and the fusion protein is not fully functional, the polypeptide would fail to translocate and thereby prevent translation of the polypeptide. Since the translation process seems to be coupled to the translocation process, a low cytoplasmic yield is therefore reasonable to expect for the DsbA constructs.

The higher yields of TorA leader constructs compared to the pelB leader construct might also be explained by the differences in the translocation pathways. Polypeptides with pelB leader are bound by SecB chaperones to maintain an unfolded conformation prior to translocation by the Sec-machinery (Driessen 2001). When over-expressing recombinant proteins in need of folding chaperones to obtain a correct fold, keeping the polypeptides unfolded might well be a limiting step. Specifically, overloading the components that stabilize the unfolded state might explain the lower yields of pelB fusions compared to TorA fusions. Vice versa, the higher yields of TorA leader constructs compared to the pelB could possibly be explained by the nature of the Tat-pathway. Natural substrates of the Tat-pathway are folded or partially folded proteins and many of these interact with folding chaperones prior to translocation (Robinson and Bolhuis 2004). Folding chaperones interacting with the Tat-pathway might aid the TorA fusion proteins to obtain a more stable conformation than the pelB leaders, and hence rescue them from degradation or allow a higher accumulation of proteins in the cytosol.

However, none of the leaders yielded scTCR-GFP fusions in the periplasm. Both the pelB and DsbA leaders have previously shown to give periplasmic yield of scFv proteins (Thie,

Schirrmann et al. 2008), and the TorA leader has shown translocation of a GFP (DeLisa, Samuelson et al. 2002) and other proteins (Blaudeck, Kreutzenbeck et al. 2003; DeLisa, Tullman et al. 2003). Despite the fact that these leaders have been shown to translocate similar domains separately, it is no guarantee for a functional translocation of the proteins as fusions.

Despite the lack of periplasmic expression, the TorA constructs showed a high relative fluorescence compared to the autofluorescence of empty XL-1 Blue cells. In addition, with the scTCR, the TorA versions were the only constructs to show fluorescence in the imaging analysis. This indicates that a very small fraction of GFP has reached a fold where it forms a functional fluorophore in the cytosol. This is implicated by the fact that PphOx-GFP showed a higher fluorescence than the TorA scTCR-GFP, even if the total amount of protein was much lower.

As for the fluorescent properties of the cells expressing phOx-GFP, the fluorescence could originate from both the cytosol and the periplasm, while the TorA2 scTCR constructs exhibited fluorescence despite only cytosolic protein localization. Whether the cleaved detected protein in the periplasm exhibited fluorescence was not investigated. Either way, the cleaved protein would have uncoupled the fluorescent property from the intended fusion partner making it inapplicable for screening purposes.

scFv-GFP

Initially, the phOx- and NIP-GFP fusions were expressed with pelB leader. We saw that recombinant protein was detected in the same fractions as for scFv. However, the fusion proteins were not detected as full length proteins, but as cleaved or degraded protein at approximately half the size of the fusions. Subsequent expression of phOx-GFP with TorA2 leader showed the same periplasmic detection of cleaved protein as with the pelB leader, only with a higher expression level.

The scFv and GFP entities are of approximately the same size. Since the detection was performed with anti-GFP and anti-His$_6$, the immediate assumption was that the detected cleavage products of both scFvs were the *C*-terminal GFP domain. A reasonable explanation is that cleavage occurs after translocation, otherwise the domain would not be able to localize

to the periplasm. Cleavage during translocation is not equally likely since it occurs with both pelB and TorA leaders.

DegP is considered to be the primary housekeeping protease of the periplasm (Baneyx and Mujacic 2004). A study of DegP showed that substrates were cleaved at discrete valine/Xaa or isoleucine/Xaa sites, Xaa being any aa (Kolmar, Waller et al. 1996). The study further showed that a chimeric protein known to be an *in vivo* substrate of DegP, was only degraded *in vitro* after reduction of its intramolecular disulfide bond. Hence, Kolmar et al. postulated that since the nonpolar valine and isoleucine is typically buried in the hydrophobic core of proteins, the main substrates of DegP are transiently denatured proteins, unfolded proteins or newly secreted proteins prior to folding and disulphide bond formation. In context of this study, one can therefore speculate whether the scFv entity is degraded by DegP prior to disulfide bond formation, as a result of an inability to form the disulfide bond or if a valine in the proximity of the linker region is part of an unstructured stretch in the folded fusion protein making it available for proteolysis. In addition, many proteases, including DegP, are upregulated as part of stress responses, e. g. over-expression of heterolog protein (Baneyx and Mujacic 2004).

5.2   Over-expression of chaperones

GroEL-GroES

Over-expression of the cytosolic folding chaperones GroEL-GroES may aid formation of a TorA transport-competent fold. In addition, it is possible that the phOx-GFP construct would obtain a conformation more resistant to cleavage. Therefore, scTCR-GFP and phOx-GFP constructs with pelB, TorA1 and TorA2 leaders were expressed in XL-1 Blue with the over-expression of GroEL-GroES.

Over-expression of GroEL-GroES did however not result in any periplasmic protein for the scTCR-GFP constructs. A possible explanation is that the scTCR entity is not able to reach a transport-competent fold in the reducing environment of the cytoplasm, even with GroEL-GroES folding assistance. GroEL-GroES has broad substrate specificity and has been shown to provide folding assistance to proteins as large as 60 kDa (Ellis 2001). It is therefore unlikely that the scTCR entity is a poor substrate for folding assistance.

The phOx-GFP construct with TorA leader was detected in the periplasm as cleaved fusion protein. This indicates that GroEL-GroES did not aid the fusion protein to obtain a fold that prevents the previously observed proteolysis. Since cleavage probably occurs in the unfolded/unstructured linker region, it is reasonable that GroEL-GroES did not impact the cleavage. However, the over-expression of GroEL-GroES abolished the previously detected cleaved protein in the periplasm from expression of phOx-GFP with pelB leader. This result supports that cytosolic folding assistance impedes the translocation of pelB leader constructs.

FkpA

To further investigate if over-expression of chaperones could result in full length periplasmic yields scTCR-GFP or phOx-GFP constructs, the periplasmic chaperone FkpA was chosen. Over-expression of this chaperone has previously been shown to give soluble periplasmic protein of scTCR (Gunnarsen, Lunde et al. submitted). Therefore, FkpA was over-expressed together with scTCR-GFP and phOx-GFP fusions with pelB and TorA2 leader in XL-1 Blue.

The over-expression of FkpA did not change the total expression level of scTCR-GFP fusions with pelB and TorA2 leader and there was still no detection of proteins in the periplasm. Furthermore, the relative fluorescence measured in cell samples and fluorescent imaging properties were not affected by the FkpA over-expression. Since the fusion proteins were not detected in the periplasm, they are not expected to become substrates of the periplasmic FkpA. However, it is possible that translocation occur, but that the intrinsic properties of the polypeptide make it unable to obtain a stable fold despite the oxidizing environment and folding assistance. Hence, the over-expression of FkpA did not change the expression profile or fluorescence properties of the scTCR-GFP fusions.

The over-expression of FkpA did not alter the total expression level of phOx-GFP with pelB leader. However, the total expression level of phOx-GFP with TorA2 leader was reduced. The periplasmic expression level of cleaved fusion protein was reduced with the over-expression of FkpA independent of leader. If FkpA and the translocation pathways utilized by the leader sequences share key componenets, the over-expression of FkpA could result in this reduced transport level.

It is tempting to suggest that the detection of translocated protein is just as dependent on the protein obtaining a stable fold after translocation as it is of the translocation itself. Hence, the ability of the scTCR-GFP fusions to obtain a stable fold in the periplasm is a limiting step in addition to the translocation process itself.

The further analysis of the fluorescent properties showed that the pelB phOx-GFP construct obtained a higher relative fluorescence with FkpA. Since the periplasmic expression level decreased simultaneously, it is possible that the fluorescence originates from the cytosol or that a larger fraction of the periplasmic protein exhibits fluorescence due to the periplasmic folding assistance of FkpA. In contrast, the TorA2 construct experienced a dramatic reduction in relative fluorescence when FkpA was co-expressed. The fluorescence images of the TorA2 construct with and without FkpA reveal that the drop in fluorescence can be cause by the fact that only some cells show fluorescence when FkpA is over-expressed, while almost all the cells show fluorescence without FkpA. The decrease in fluorescence of the TorA construct also correlates with the abovementioned decreased expression level.

5.3 Expression hosts

To further investigate the possibilities of obtaining full length fusion proteins in the periplasm, different expression host strains were tested.

Rosetta Blue (DE3) pLysS

The Rosetta Blue strain contains extra tRNA genes for expression of proteins with codons rarly used in *E. coli* that have been shown to increase expression of heterologous proteins (Bukhtiyarova, Northrop et al. 2004). The scTCR-GFP and phOx-GFP fusions with pelB and TorA2 leader, with and without FkpA, were expressed in Rosetta Blue and XL-1 Blue for comparison.

There was only a minor effect of Rosetta Blue on total and periplasmic expression level of phOx-GFP fusions and in all situations the periplasmic yield was cleaved protein. There seemed to be one exception, the phOx-GFP with TorA2 leader and over-expression of FkpA. Here a great increase in the yield of cleaved protein in the periplasm was observed when

expressed in Rosetta Blue. However, the validation of the FkpA expression showed that this particular construct only gave a minor detectable yield of FkpA in the periplasm.

When over-expressing proteins and tRNAs it is important to consider the stress inflicted on the cell in context of the stability of the protein in interest. The non-fused scTCR is not expressed in the periplasm when using the same stringent expression system that expresses high soluble periplasmic yields of phOx. Therefore, it is reasonable to expect the scTCR-GFP fusions to be less stable than the phOx-GFP fusions, and therefore more sensitive to over-expression of other proteins and tRNAs. Hence, the scTCR-GFP shows greater variation in the effect of Rosetta Blue as an expression host than phOx-GFP.

BL21

To investigate the possibility of obtaining full length periplasmic fusion protein, phOx-GFP with TorA2 leader was expressed in the *E. coli* strain BL21. This strain is deficient in the cytosolic *lon* and periplasmic *ompT* proteases (Terpe 2006), possible contributors to the observed cleavage.

The phOx-GFP with TorA2 showed a decreased total expression level and relative fluorescence when expressed in BL21 compared to XL-1 Blue. In addition, expression in BL21 did not result in any detectable periplasmic yield, not even the cleaved fusion proteins previously observed in XL-1 Blue. Hence, these results indicate that the cells need the quality control provided by the *lon* and *ompT* proteases to maintain a high expression level and to gain periplasmic yield of the fusion protein.

5.4 Expression conditions

To further investigate the possibility of obtaining full length periplasmic fusion proteins, the effects of induction and expression temperature were investigated.

5.4.1 Induction

It has previously been shown that expression without IPTG induction improved the periplasmic yield of the soluble scTCR (Gunnarsen, Lunde et al. submitted). This effect is presumably due to low level transcription of heterologous component only, which alleviates

the metabolic load conferred upon the host during heterologous protein over-expression. Therefore, phOx-GFP with pelB and TorA2 leader, with and without FkpA, were expressed in XL-1 Blue with and without induction.

Induction versus no induction did not significantly affect the total expression level of any of the four constructs. Transport of fusion proteins to the periplasm seemed to be more affected by induction of protein expression. The pelB constructs seemed dependent on induction to be transported, while the TorA constructs were either not affected or negatively affected by induction. When the protein in question is aggregation prone or unstable, an increased transcription and translation level triggered by induction might cause stress responses in the cell resulting in a reduced functional yield. Hence, the different effects of induction probably reflect the variation of an optimal expression level of the phOx-GFP construct targeted to different transport pathways and with or without the over-expression of FkpA. Irrespectively of induction, no full length fusion protein was detected in the periplasm.

Furthermore, induction of pelB constructs gave increased relative fluorescence of the cells. This could indicate that the increased fluorescence is correlated to the periplasmic localization, even though it represents cleaved protein. In contrast, the TorA leader constructs showed a decrease in relative fluorescence when induced. The decreased fluorescence and periplasmic yield of the TorA construct with FkpA upon induction might be caused by the simultaneous over-expression of two proteins, enhanced by induction, and possible overloading of shared key components in the transport pathways.

5.4.2 Temperature

A previous study has reported periplasmic yield of full length scFv-GFP when performing the expression at 16°C for 48 hours (Griep, van Twisk et al. 1999). Therefore, the constructs expressed with and without induction was also expressed at both standard expression conditions (30°C ON) and 16°C for 48 hours.

Without induction the total expression level of all constructs gave a dramatically reduced yield when the expression temperature was lowered to 16°C. However, when induced, the pelB constructs gave reduced yields and while the TorA constructs obtained increased yields at 16°C. Furthermore, the periplasmic yield did not seem to be affected by the expression

temperature and did not result in full length fusion protein. Expression at reduced temperatures is thought to be beneficial for aggregation prone proteins by making them fold slower. However, the overall functions of the cell are also affected by the reduced temperature. Therefore, a putative explanation is that these combined effects result in the opposing effects observed when lowering the expression temperature.

In summary, lowering of the expression temperature and expressing without induction did not resulted in full length periplasmic fusion protein. However, the cleaved fusion protein that was detected is the *C*-terminal GFP entity as it is recognized by anti-GFP and anti-His$_6$.

5.5 Additional detections

To examine if the scFv entity was present in the periplasm, the periplasmic fractions were detected with anti-human lambda L-chain. The detection showed bands at ~30 kDa, consistent with the scFv entity, when phOx-GFP was expressed with a pelB leader. Surprisingly, no bands were visible when expressed with the TorA leader.

A pervious study showed that Tat-transport of a scFv was dependent on an oxidizing cytoplasm (DeLisa, Tullman et al. 2003), presumably by allowing the scFv to reach a transport competent fold by facilitating disulfide bond formation. In addition, another study showed that an *in vivo* substrate of the periplasmic protease DegP was degraded *in vitro* only after reduction of its intramolecular disulfide bond (Kolmar, Waller et al. 1996), presumably due to exposure of cleavage sites when not stably folded with disulfide bonds.

It is reasonable to assume that transport to the periplasm of the scFv with TorA leader could be rescued by the genetic fusion to GFP. Hence, a putative explanation is that the scFv domain was degraded by DegP in the periplasm due to an unstable fold while the GFP domain is stably folded and degradation resistant. In contrast, when the fusion protein is expressed with pelB leader, the protein is transported as a linear polypeptide to the oxidizing environment of the periplasm. Hence, the polypeptide starts its folding in a disulfide bond promoting environment and both the scFv and GFP obtains a stable conformation persisting after cleavage.

Still, no full length fusion protein had been detected in the periplasm. To examine if a only minute amount of full length protein was present, a western blot detected with anti-His$_6$ was developed with very sensitive femto-substrate and over-exposed. Only one of the constructs showed a vague band at ~60 kDa, namely phOx-GFP with pelB leader and FkpA expressed at 30°C without induction. Interestingly, cells expressing this construct with the given conditions barely showed any fluorescence in the microscopic analysis. The relative fluorescence was higher than the autofluorescence of empty cells though, but much lower than what was obtained with other constructs and expression conditions.

5.6 Concluding remarks

This study has showed that the TorA leader provides a higher total expression level of scTCR-GFP than the DsbA and pelB leader. Moreover, only the TorA construct-containing host cells exhibited a fluorescent phenotype. However, none of the leaders provided periplasmic yield of the fusion protein. Neither over-expression of GroEL-GroES nor FkpA, nor expression in Rosetta Blue resulted in periplasmic yield of any full length fusion protein.

We saw that phOx-GFP fusions are translocated to the periplasm with both pelB and TorA leaders. However, they are detected as half-sized proteins, probably cleaved in the linker region connecting the two moieties. Over-expression of GroEL-GroES or FkpA, expression in Rosetta Blue or BL21, expression at 16°C or without induction did not yield full length fusion protein in the periplasm.

The only periplasmic full length fusion proteins detected was phOx-GFP when expressed with pelB leader at 30°C with over-expression of FkpA without induction. However, the minute yield excludes it for use in any downstream application.

In conclusion, the study showed that the expression profiles of the scFvs were not preserved when expressed as fusions to GFP and that the scTCR-GFP fusions were not translocated to periplasm. Hence, further optimalization of expression procedure or construct design is needed.

6 Future perspectives

In order to establish a screening platform based on the constructs described in this study, the expression and/or construct design need to be further improved.

6.1 Expression host

A whole panel of host strains with different properties could be benchmarked against each other, as shown in the table below. In addition to the different strains, each construct can be expressed with and without the over-expression of FkpA. The most "pimped" system would be expression in the Origami B strain pretransformed with the pGro7 with over-expression of FkpA. This would facilitate disulfide bond formation in the cytoplasm together with cytosolic and periplasmic folding assistance in a cell deficient of *lon* and *ompT* proteases to increase the protein stability.

Genotype/ plasmid	*lon/ompT* mut	*trxB/gor* mut	pRARE	pGro7
Phenotype	Less degradation → increased stability	Oxydizing cytosol →disulfide bond formation	Express rare tRNAs	GroEL/ES → cytosolic folding assistance
XL-1 Blue	-	-	-	-
XL-1 Blue + pGro7	-	-	-	+
BL21	+	-	-	-
Rosetta Blue (DE3)pLysS	-	-	+	-
Origami B (DE3)	+	+	-	-
Rosetta-gami B (DE3) pLysS	+	+	+	-
Origami B + pGro7	+	+	-	+

It is possible that periplasmic expression of scTCR-GFP would result in the same cleavage as observed with phOx-GFP. This is based on the sequence similarity of the phOx-GFP and scTCR-GFP linker regions and the presence of both valine and isoleucine which are part of DegP substrate cleavage sites (Kolmar, Waller et al. 1996).

```
          scFv  -   Linker  -  GFP
phOx-GFP:  V  L  G -  A  A  A  - G  S  K G
scTCR-GFP: L  V  I -  A  A  A  - G  S  K G
```

If cleaved periplasmic scTCR-GFP is expressed, a putative solution for both the scTCR and phOx fusions would be to express in a DegP-deficient strain if available.

6.2 Construct design

It would be desirable to change residues in the vicinity of the linker region of the scFv and scTCR constructs to alleviate proteolysis.

6.3 The SsrA peptide

To eliminate background fluorescence of protein in the cytosol, an SsrA peptide could be genetically fused to the *C*-terminal end of the fusion protein that would lead to degradation of cytosolic protein (DeLisa, Samuelson et al. 2002). Hence, all fluorescence from the cell would originate from periplasmic protein rescued from cytosolic degradation by translocation. This is undoubtedly a desired quality in the expression system to create a strong correlation between folded periplasmically expressed protein and fluorescence.

References

Alami, M., I. Luke, et al. (2003). "Differential interactions between a twin-arginine signal peptide and its translocase in Escherichia coli." Mol Cell **12**(4): 937-46.

Arie, J. P., N. Sassoon, et al. (2001). "Chaperone function of FkpA, a heat shock prolyl isomerase, in the periplasm of Escherichia coli." Mol Microbiol **39**(1): 199-210.

Backert, S. and T. F. Meyer (2006). "Type IV secretion systems and their effectors in bacterial pathogenesis." Curr Opin Microbiol **9**(2): 207-17.

Baneyx, F. and M. Mujacic (2004). "Recombinant protein folding and misfolding in Escherichia coli." Nat Biotechnol **22**(11): 1399-408.

Baneyx, F. and J. L. Palumbo (2003). "Improving heterologous protein folding via molecular chaperone and foldase co-expression." Methods Mol Biol **205**: 171-97.

Bendtsen, J. D., H. Nielsen, et al. (2004). "Improved prediction of signal peptides: SignalP 3.0." J Mol Biol **340**(4): 783-95.

Berks, B. C. (1996). "A common export pathway for proteins binding complex redox cofactors?" Mol Microbiol **22**(3): 393-404.

Berks, B. C., T. Palmer, et al. (2005). "Protein targeting by the bacterial twin-arginine translocation (Tat) pathway." Curr Opin Microbiol **8**(2): 174-81.

Blaudeck, N., P. Kreutzenbeck, et al. (2003). "Genetic analysis of pathway specificity during posttranslational protein translocation across the Escherichia coli plasma membrane." J Bacteriol **185**(9): 2811-9.

Bogen, B., B. Malissen, et al. (1986). "Idiotope-specific T cell clones that recognize syngeneic immunoglobulin fragments in the context of class II molecules." Eur J Immunol **16**(11): 1373-8.

Bogen, B., P. A. Ruffini, et al. (2006). "Idiotype-specific immunotherapy in multiple myeloma: suggestions for future directions of research." Haematologica **91**(7): 941-8.

Bolhuis, A., J. E. Mathers, et al. (2001). "TatB and TatC form a functional and structural unit of the twin-arginine translocase from Escherichia coli." J Biol Chem **276**(23): 20213-9.

Bothmann, H. and A. Pluckthun (1998). "Selection for a periplasmic factor improving phage display and functional periplasmic expression." Nat Biotechnol **16**(4): 376-80.

Bothmann, H. and A. Pluckthun (2000). "The periplasmic Escherichia coli peptidylprolyl cis,trans-isomerase FkpA. I. Increased functional expression of antibody fragments with and without cis-prolines." J Biol Chem **275**(22): 17100-5.

Bruser, T. (2007). "The twin-arginine translocation system and its capability for protein secretion in biotechnological protein production." Appl Microbiol Biotechnol **76**(1): 35-45.

Bukhtiyarova, M., K. Northrop, et al. (2004). "Improved expression, purification, and crystallization of p38alpha MAP kinase." Protein Expr Purif **37**(1): 154-61.

Buskiewicz, I., E. Deuerling, et al. (2004). "Trigger factor binds to ribosome-signal-recognition particle (SRP) complexes and is excluded by binding of the SRP receptor." Proc Natl Acad Sci U S A **101**(21): 7902-6.

Carter, P. J. (2006). "Potent antibody therapeutics by design." Nat Rev Immunol **6**(5): 343-57.

Casey, J. L., A. M. Coley, et al. (2000). "Green fluorescent antibodies: novel in vitro tools." Protein Eng **13**(6): 445-52.

Chen, R. and U. Henning (1996). "A periplasmic protein (Skp) of Escherichia coli selectively binds a class of outer membrane proteins." Mol Microbiol **19**(6): 1287-94.

Christie, P. J., K. Atmakuri, et al. (2005). "Biogenesis, architecture, and function of bacterial type IV secretion systems." Annu Rev Microbiol **59**: 451-85.

Cormack, B. P., R. H. Valdivia, et al. (1996). "FACS-optimized mutants of the green fluorescent protein (GFP)." <u>Gene</u> **173**(1 Spec No): 33-8.

Cornelis, G. R. (2006). "The type III secretion injectisome." <u>Nat Rev Microbiol</u> **4**(11): 811-25.

Crameri, A., E. A. Whitehorn, et al. (1996). "Improved green fluorescent protein by molecular evolution using DNA shuffling." <u>Nat Biotechnol</u> **14**(3): 315-9.

de Keyzer, J., C. van der Does, et al. (2003). "The bacterial translocase: a dynamic protein channel complex." <u>Cell Mol Life Sci</u> **60**(10): 2034-52.

de Marco, A. (2009). "Strategies for successful recombinant expression of disulfide bond-dependent proteins in Escherichia coli." <u>Microb Cell Fact</u> **8**: 26.

DeLisa, M. P., P. Samuelson, et al. (2002). "Genetic analysis of the twin arginine translocator secretion pathway in bacteria." <u>J Biol Chem</u> **277**(33): 29825-31.

DeLisa, M. P., D. Tullman, et al. (2003). "Folding quality control in the export of proteins by the bacterial twin-arginine translocation pathway." <u>Proc Natl Acad Sci U S A</u> **100**(10): 6115-20.

Deuerling, E., H. Patzelt, et al. (2003). "Trigger Factor and DnaK possess overlapping substrate pools and binding specificities." <u>Mol Microbiol</u> **47**(5): 1317-28.

Dev, I. K. and P. H. Ray (1990). "Signal peptidases and signal peptide hydrolases." <u>J Bioenerg Biomembr</u> **22**(3): 271-90.

Driessen, A. J. (2001). "SecB, a molecular chaperone with two faces." <u>Trends Microbiol</u> **9**(5): 193-6.

Driessen, A. J., E. H. Manting, et al. (2001). "The structural basis of protein targeting and translocation in bacteria." <u>Nat Struct Biol</u> **8**(6): 492-8.

Economou, A. (1999). "Following the leader: bacterial protein export through the Sec pathway." <u>Trends Microbiol</u> **7**(8): 315-20.

Ellis, R. J. (2001). "Molecular chaperones: inside and outside the Anfinsen cage." <u>Curr Biol</u> **11**(24): R1038-40.

Emanuelsson, O., S. Brunak, et al. (2007). "Locating proteins in the cell using TargetP, SignalP and related tools." <u>Nat Protoc</u> **2**(4): 953-71.

Ewalt, K. L., J. P. Hendrick, et al. (1997). "In vivo observation of polypeptide flux through the bacterial chaperonin system." <u>Cell</u> **90**(3): 491-500.

Gardy, J. L., M. R. Laird, et al. (2005). "PSORTb v.2.0: expanded prediction of bacterial protein subcellular localization and insights gained from comparative proteome analysis." <u>Bioinformatics</u> **21**(5): 617-23.

Graslund, S., P. Nordlund, et al. (2008). "Protein production and purification." <u>Nat Methods</u> **5**(2): 135-46.

Griep, R. A., C. van Twisk, et al. (1999). "Fluobodies: green fluorescent single-chain Fv fusion proteins." <u>J Immunol Methods</u> **230**(1-2): 121-30.

Gunnarsen, K. S., E. Lunde, et al. (submitted). "Periplasmic expression of soluble single chain T cell receptors is rescued by the chaperone FkpA."

Halaby, D. M., A. Poupon, et al. (1999). "The immunoglobulin fold family: sequence analysis and 3D structure comparisons." <u>Protein Eng</u> **12**(7): 563-71.

Hartl, F. U. and M. Hayer-Hartl (2002). "Molecular chaperones in the cytosol: from nascent chain to folded protein." <u>Science</u> **295**(5561): 1852-8.

Henderson, I. R., F. Navarro-Garcia, et al. (2004). "Type V protein secretion pathway: the autotransporter story." <u>Microbiol Mol Biol Rev</u> **68**(4): 692-744.

Hinsley, A. P., N. R. Stanley, et al. (2001). "A naturally occurring bacterial Tat signal peptide lacking one of the 'invariant' arginine residues of the consensus targeting motif." <u>FEBS Lett</u> **497**(1): 45-9.

Jack, R. L., G. Buchanan, et al. (2004). "Coordinating assembly and export of complex bacterial proteins." EMBO J **23**(20): 3962-72.

Johnson, T. L., J. Abendroth, et al. (2006). "Type II secretion: from structure to function." FEMS Microbiol Lett **255**(2): 175-86.

Kipriyanov, S. M., G. Moldenhauer, et al. (1997). "High level production of soluble single chain antibodies in small-scale Escherichia coli cultures." J Immunol Methods **200**(1-2): 69-77.

Kol, S., N. Nouwen, et al. (2008). "Mechanisms of YidC-mediated insertion and assembly of multimeric membrane protein complexes." J Biol Chem **283**(46): 31269-73.

Kolmar, H., P. R. Waller, et al. (1996). "The DegP and DegQ periplasmic endoproteases of Escherichia coli: specificity for cleavage sites and substrate conformation." J Bacteriol **178**(20): 5925-9.

Langdon, R. H., J. Cuccui, et al. (2009). "N-linked glycosylation in bacteria: an unexpected application." Future Microbiol **4**(4): 401-12.

Loset, G. A., E. Lunde, et al. (2007). "Functional phage display of two murine alpha/beta T-cell receptors is strongly dependent on fusion format, mode and periplasmic folding assistance." Protein Eng Des Sel **20**(9): 461-72.

Lu, M., X. G. Gong, et al. (2005). "Cloning, expression, purification, and characterization of LC-1 ScFv with GFP tag." J Zhejiang Univ Sci B **6**(8): 832-7.

Luirink, J., G. von Heijne, et al. (2005). "Biogenesis of inner membrane proteins in Escherichia coli." Annu Rev Microbiol **59**: 329-55.

Makrides, S. C. (1996). "Strategies for achieving high-level expression of genes in Escherichia coli." Microbiol Rev **60**(3): 512-38.

Maynard, J., E. J. Adams, et al. (2005). "High-level bacterial secretion of single-chain alphabeta T-cell receptors." J Immunol Methods **306**(1-2): 51-67.

Missiakas, D., J. M. Betton, et al. (1996). "New components of protein folding in extracytoplasmic compartments of Escherichia coli SurA, FkpA and Skp/OmpH." Mol Microbiol **21**(4): 871-84.

Molloy, P. E., A. K. Sewell, et al. (2005). "Soluble T cell receptors: novel immunotherapies." Curr Opin Pharmacol **5**(4): 438-43.

Murphy, K., P. Travers, et al., Eds. (2008). Immunobiology, Garland Science.

Nakayama, M. and O. Ohara (2003). "A system using convertible vectors for screening soluble recombinant proteins produced in Escherichia coli from randomly fragmented cDNAs." Biochem Biophys Res Commun **312**(3): 825-30.

Natale, P., T. Bruser, et al. (2008). "Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane--distinct translocases and mechanisms." Biochim Biophys Acta **1778**(9): 1735-56.

Nishihara, K., M. Kanemori, et al. (1998). "Chaperone coexpression plasmids: differential and synergistic roles of DnaK-DnaJ-GrpE and GroEL-GroES in assisting folding of an allergen of Japanese cedar pollen, Cryj2, in Escherichia coli." Appl Environ Microbiol **64**(5): 1694-9.

Nishihara, K., M. Kanemori, et al. (2000). "Overexpression of trigger factor prevents aggregation of recombinant proteins in Escherichia coli." Appl Environ Microbiol **66**(3): 884-9.

Oelschlaeger, P., S. Srikant-Iyer, et al. (2002). "Fluorophor-linked immunosorbent assay: a time- and cost-saving method for the characterization of antibody fragments using a fusion protein of a single-chain antibody fragment and enhanced green fluorescent protein." Anal Biochem **309**(1): 27-34.

Omoya, K., Z. Kato, et al. (2004). "Systematic optimization of active protein expression using GFP as a folding reporter." Protein Expr Purif **36**(2): 327-32.

Ormo, M., A. B. Cubitt, et al. (1996). "Crystal structure of the Aequorea victoria green fluorescent protein." Science **273**(5280): 1392-5.

Papanikou, E., S. Karamanou, et al. (2005). "Identification of the preprotein binding domain of SecA." J Biol Chem **280**(52): 43209-17.

Pedelacq, J. D., S. Cabantous, et al. (2006). "Engineering and characterization of a superfolder green fluorescent protein." Nat Biotechnol **24**(1): 79-88.

Pommier, J., V. Mejean, et al. (1998). "TorD, a cytoplasmic chaperone that interacts with the unfolded trimethylamine N-oxide reductase enzyme (TorA) in Escherichia coli." J Biol Chem **273**(26): 16615-20.

Prasher, D. C., V. K. Eckenrode, et al. (1992). "Primary structure of the Aequorea victoria green-fluorescent protein." Gene **111**(2): 229-33.

Robinson, C. and A. Bolhuis (2004). "Tat-dependent protein targeting in prokaryotes and chloroplasts." Biochim Biophys Acta **1694**(1-3): 135-47.

Schafer, U., K. Beck, et al. (1999). "Skp, a molecular chaperone of gram-negative bacteria, is required for the formation of soluble periplasmic intermediates of outer membrane proteins." J Biol Chem **274**(35): 24567-74.

Schierle, C. F., M. Berkmen, et al. (2003). "The DsbA signal sequence directs efficient, cotranslational export of passenger proteins to the Escherichia coli periplasm via the signal recognition particle pathway." J Bacteriol **185**(19): 5706-13.

Schwalbach, G., A. P. Sibler, et al. (2000). "Production of fluorescent single-chain antibody fragments in Escherichia coli." Protein Expr Purif **18**(2): 121-32.

Shaner, N. C., G. H. Patterson, et al. (2007). "Advances in fluorescent protein technology." J Cell Sci **120**(Pt 24): 4247-60.

Shaner, N. C., P. A. Steinbach, et al. (2005). "A guide to choosing fluorescent proteins." Nat Methods **2**(12): 905-9.

Shen, L. M., J. I. Lee, et al. (1991). "Use of site-directed mutagenesis to define the limits of sequence variation tolerated for processing of the M13 procoat protein by the Escherichia coli leader peptidase." Biochemistry **30**(51): 11775-81.

Steiner, D., P. Forrer, et al. (2006). "Signal sequences directing cotranslational translocation expand the range of proteins amenable to phage display." Nat Biotechnol **24**(7): 823-31.

Terpe, K. (2006). "Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems." Appl Microbiol Biotechnol **72**(2): 211-22.

Thammawong, P., W. Kasinrerk, et al. (2006). "Twin-arginine signal peptide attributes effective display of CD147 to filamentous phage." Appl Microbiol Biotechnol **69**(6): 697-703.

Thie, H., T. Schirrmann, et al. (2008). "SRP and Sec pathway leader peptides for antibody phage display and antibody fragment production in E. coli." N Biotechnol **25**(1): 49-54.

Varela-Rohena, A., P. E. Molloy, et al. (2008). "Control of HIV-1 immune escape by CD8 T cells expressing enhanced T-cell receptor." Nat Med **14**(12): 1390-5.

von Heijne, G. (1985). "Signal sequences. The limits of variation." J Mol Biol **184**(1): 99-105.

Waldo, G. S., B. M. Standish, et al. (1999). "Rapid protein-folding assay using green fluorescent protein." Nat Biotechnol **17**(7): 691-5.

Walton, T. A. and M. C. Sousa (2004). "Crystal structure of Skp, a prefoldin-like chaperone that protects soluble and membrane proteins from aggregation." Mol Cell **15**(3): 367-74.

Wang, H. and S. Chong (2003). "Visualization of coupled protein folding and binding in bacteria and purification of the heterodimeric complex." Proc Natl Acad Sci U S A **100**(2): 478-83.

Wu, L., M. Adams, et al. (2008). "lenalidomide enhances natural killer cell and monocyte-mediated antibody-dependent cellular cytotoxicity of rituximab-treated CD20+ tumor cells." Clin Cancer Res **14**(14): 4650-7.

Wulfing, C. and A. Pluckthun (1994). "Correctly folded T-cell receptor fragments in the periplasm of Escherichia coli. Influence of folding catalysts." J Mol Biol **242**(5): 655-69.

Zacharias, D. A., J. D. Violin, et al. (2002). "Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells." Science **296**(5569): 913-6.

Zhang, A., E. J. Cantor, et al. (2005). "Productive interaction of chaperones with substrate protein domains allows correct folding of the downstream GFP domain." Gene **350**(1): 25-31.

Zhang, G., S. Brokx, et al. (2006). "Extracellular accumulation of recombinant proteins fused to the carrier protein YebF in Escherichia coli." Nat Biotechnol **24**(1): 100-4.

Zimmer, M. (2002). "Green fluorescent protein (GFP): applications, structure, and related photophysical behavior." Chem Rev **102**(3): 759-81.