



# Strategies to Enhance Periplasmic Recombinant Protein Production Yields in *Escherichia coli*

Alexandros Karyolaimos and Jan-Willem de Gier\*

Department of Biochemistry and Biophysics, Stockholm University, Stockholm, Sweden

## OPEN ACCESS

### Edited by:

Dong-Yup Lee,  
Sungkyunkwan University, South  
Korea

### Reviewed by:

Dave Siak-Wei Ow,  
Bioprocessing Technology Institute  
(A\*STAR), Singapore  
Georgios Skretas,  
National Hellenic Research  
Foundation, Greece  
Matthew Peter DeLisa,  
Cornell University, United States

### \*Correspondence:

Jan-Willem de Gier  
degier@dbb.su.se

### Specialty section:

This article was submitted to  
Synthetic Biology,  
a section of the journal  
Frontiers in Bioengineering and  
Biotechnology

**Received:** 18 October 2021

**Accepted:** 24 November 2021

**Published:** 14 December 2021

### Citation:

Karyolaimos A and  
de Gier J-W (2021) Strategies to  
Enhance Periplasmic Recombinant  
Protein Production Yields in  
*Escherichia coli*.  
Front. Bioeng. Biotechnol. 9:797334.  
doi: 10.3389/fbioe.2021.797334

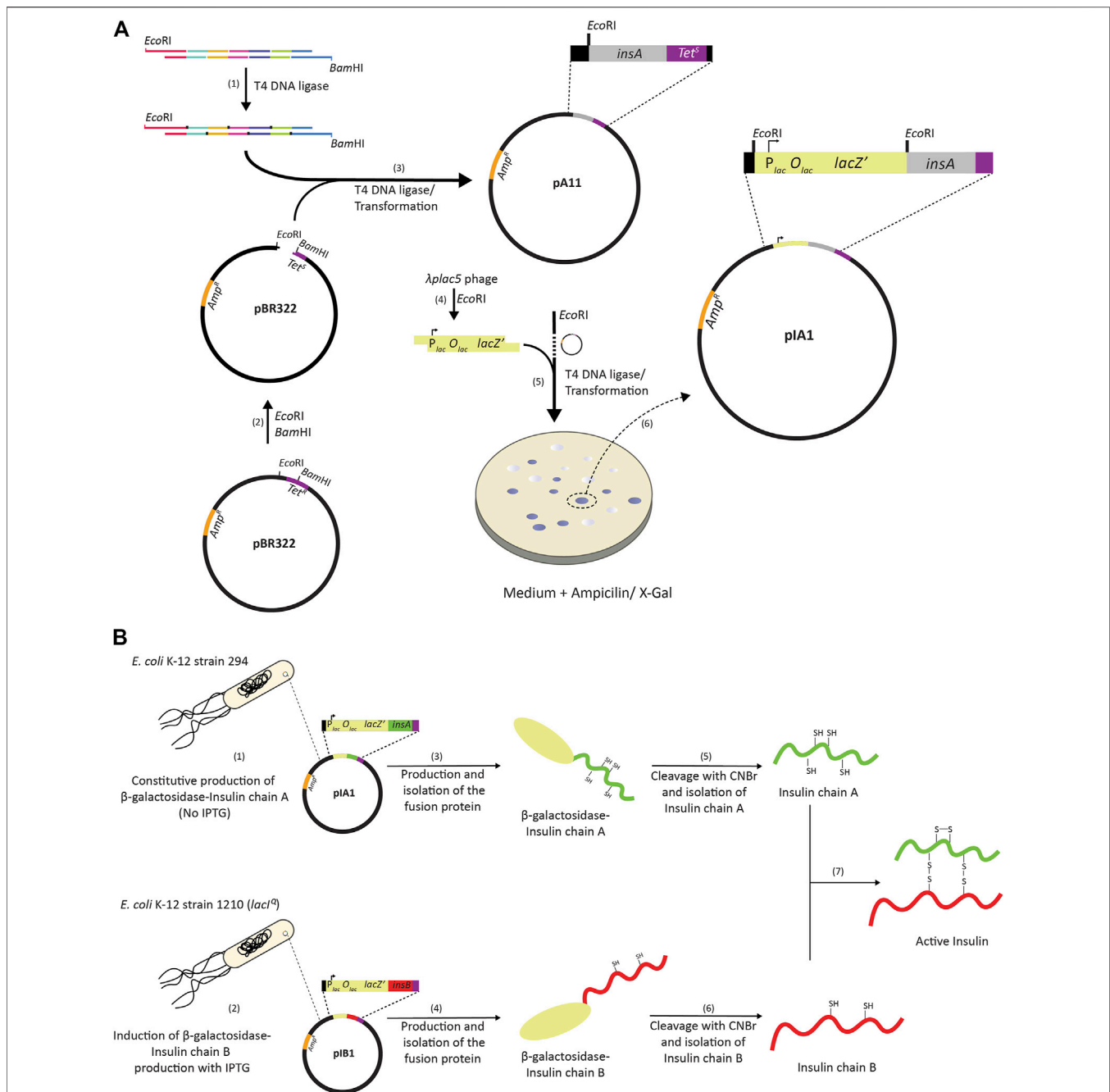
Main reasons to produce recombinant proteins in the periplasm of *E. coli* rather than in its cytoplasm are to -i- enable disulfide bond formation, -ii- facilitate protein isolation, -iii- control the nature of the N-terminus of the mature protein, and -iv- minimize exposure to cytoplasmic proteases. However, hampered protein targeting, translocation and folding as well as protein instability can all negatively affect periplasmic protein production yields. Strategies to enhance periplasmic protein production yields have focused on harmonizing secretory recombinant protein production rates with the capacity of the secretory apparatus by transcriptional and translational tuning, signal peptide selection and engineering, increasing the targeting, translocation and periplasmic folding capacity of the production host, preventing proteolysis, and, finally, the natural and engineered adaptation of the production host to periplasmic protein production. Here, we discuss these strategies using notable examples as a thread.

**Keywords:** *Escherichia coli*, periplasm, recombinant protein, protein production, production optimization

## INTRODUCTION

In 1972, the creation of the first recombinant DNA molecules was reported (Jackson et al., 1972). Not long thereafter, a study describing the “Construction of Biologically Functional Bacterial Plasmids *In Vitro*” was published (Cohen et al., 1973). Together with the development of methodology to chemically synthesize DNA, these two studies paved the way for using *Escherichia coli* as a host for producing recombinant proteins (Itakura et al., 1977). In 1976, the company Genentech—its name is a play of words of “genetic engineering” and “technology”—was founded and it pioneered the use of *E. coli* for the production of therapeutic proteins (Hughes, 2011). The foundation of Genentech is widely considered as the start of recombinant DNA driven biotechnology. Based on a collaborative effort between Genentech and Eli Lilly, human insulin (a.k.a. humulin), became in 1982 the first recombinantly produced therapeutic protein approved by the FDA for human use (Crea et al., 1978; Goeddel et al., 1979). For the production of humulin, the A and B chains of insulin were produced in the cytoplasm of two different *E. coli* strains as C-terminal fusions to  $\beta$ -galactosidase (Figure 1). After partial purification of the chimeric proteins, the insulin chains were released from the fusion proteins by cyanogen treatment. Subsequently, the released A and B chains were purified and joined through air oxidation resulting in functional insulin.

First, recombinant proteins, like above mentioned  $\beta$ -galactosidase insulin fusion proteins, were produced in the cytoplasm of *E. coli* (e.g., (Itakura et al., 1977; Goeddel et al., 1979)). Later on, based on increasing knowledge of protein export in *E. coli*, recombinant proteins were also produced in the periplasm of *E. coli* (e.g., (Fraser and Bruce, 1978; Talmadge et al., 1980)). What are the (potential) benefits of producing a protein in the periplasm of *E. coli* rather than in its cytoplasm?



**FIGURE 1 |** Cloning and production of human insulin using *E. coli*. **(A)** The gene encoding the insulin A chain was constructed using DNA oligos with overlapping parts and T4 DNA ligase **(1)**. *EcoRI* and *BamHI* compatible overhangs were introduced to the 5' end and 3' end of the *insA* gene, respectively. The overhangs were introduced to facilitate cloning of the *insA* gene into the pBR322 vector that had been digested with *EcoRI* and *BamHI* **(2)**. After ligation and transformation **(3)**, positive clones were selected on antibiotic containing media; cells containing the pBR322 vector with the *insA* insert are ampicillin resistant, Amp<sup>R</sup>, and tetracycline sensitive, Tet<sup>S</sup>, since the insertion of *insA* into pBR322 results in the partial deletion of the *tet* gene. Insertion of the insulin chain gene into pBR322 was verified by DNA sequencing. Subsequently, an *EcoRI* fragment from the bacteriophage  $\lambda$ lac5 genome **(4)** containing the *lac* promoter, the *lac* operator and 1,006 codons of the *lacZ* gene (out of the 1,024 codons encoding full length  $\beta$ -galactosidase) was inserted into the *EcoRI* site of the pBR322 vector with the *insA* gene **(5)**. This results in the in-frame fusion of *lacZ* and *insA* enabling the production of a  $\beta$ -galactosidase insulin chain A fusion protein. Plates with medium containing ampicillin and the  $\beta$ -galactosidase substrate 5-bromo-4-chloro-3-indolyl- $\beta$ -D-galactopyranoside (X-Gal) was used to identify blue colonies, *i.e.*, colonies that could produce the  $\beta$ -galactosidase insulin A fusion protein **(6)**. An expression vector for the production of a  $\beta$ -galactosidase insulin chain B fusion was constructed in a similar way. **(B)** The A and B chains of insulin C-terminally fused to  $\beta$ -galactosidase were produced in the cytoplasm using two different *E. coli* strains. *E. coli* K-12 strain 294 was used for the production of the  $\beta$ -galactosidase-InsA fusion **(1)** and *E. coli* K-12 strain D1290 was used for the production of the  $\beta$ -galactosidase-InsB fusion **(2)**. *E. coli* K-12 strain 294 produces the  $\beta$ -galactosidase-InsA fusion constitutively, so no Isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) was used to induce the expression of the gene encoding the fusion. The *E. coli* K-12 strain D1290 contains a mutation in the promoter of the *lacI* gene (*lacI*<sup>q</sup>), which results in higher production of LacI. Here, IPTG had to be added to induce expression of the gene encoding the fusion protein. After partial purification of the fusion proteins **(3, 4)**, the insulin chains were recovered from the fusion proteins by cyanogen treatment **(5, 6)**. Subsequently, the A and B chains were isolated and joined through air oxidation resulting in active insulin **(7)**.

To reach the periplasm, a protein has to be equipped at its N-terminus with a cleavable signal peptide so that it can cross the cytoplasmic membrane (Hegde and Bernstein, 2006; Smets et al., 2019). Upon translocation across the membrane, the signal peptide is cleaved off. This enables to control which amino acid is at the very N-terminus of the mature protein (Smets et al., 2019). Not all proteins naturally contain a methionine at their N-terminus and the cytoplasmic production of a protein in the cytoplasm does not guarantee the presence of an N-terminal methionine due to the action of methionine aminopeptidase (Liao et al., 2004; Frottin et al., 2006). Removal of the N-formyl methionine can be critical for the proper folding, stability and function of a recombinant protein (Endo et al., 2001; Liao et al., 2003).

Protein production in the periplasm enables to isolate the protein from the periplasmic fraction rather than a whole cell lysate. This can greatly facilitate the isolation of a protein since it has to be isolated from a mixture that is less complex (Mergulhao et al., 2005). In this respect it should be mentioned that it is relatively easy to isolate the periplasmic fraction of *E. coli* also at an industrial scale (Mergulhao et al., 2005; Ellis and Humphreys, 2014). Furthermore, periplasmic protein production can also lead to the “spontaneous” release of the protein into the extracellular milieu (Kastenhofer et al., 2021). The release of a protein into the extracellular milieu can also be promoted by actively making the outer membrane more permeable or by fusing the protein to a fusion partner that is naturally secreted into the extracellular milieu (Hsiung et al., 1989; Zhang et al., 2006; Natarajan et al., 2017; Gao et al., 2018)). Extracellular production of a (fusion) protein can even further facilitate protein isolation (Tripathi, 2016). Notably, the costs involved in the isolation of a protein usually make up a large part of the total production costs (Tripathi and Shrivastava, 2019).

Many recombinant proteins, like hormones and antibody fragments, contain disulfide bonds which are essential for stabilizing the fold of a protein (Manta et al., 2019). Disulfide bond formation is not promoted in the reducing environment of the cytoplasm (Manta et al., 2019). However, the oxidative environment of the periplasm along with the presence of the disulfide bond formation (Dsb)-system in the periplasm promotes disulfide bond formation in this compartment, thereby facilitating the production of disulfide bond containing proteins (Manta et al., 2019). Finally, periplasmic protein production can promote the stability of a protein due to minimizing its exposure to cytoplasmic proteases (e.g., (Talmadge and Gilbert, 1982)).

However, the production of a protein in the periplasm is often cumbersome. Usually, a significant fraction of the produced protein accumulates in the unprocessed form in the cytoplasm (Schlegel et al., 2013; Baumgarten et al., 2018; Karyolaimos et al., 2019; Karyolaimos et al., 2020). Periplasmic protein production can also have a negative effect on the fitness of the production host and consequently biomass formation as well as protein production yields (Schlegel et al., 2013; Baumgarten et al., 2018). Furthermore, proteins produced in the periplasm do not always fold properly and can also be prone to degradation

(e.g., (Meerman and Georgiou, 1994; Chen et al., 2004; Ellis et al., 2017)).

In this review, we give an overview of the different strategies used to promote proper targeting, efficient translocation, proper folding and avoid degradation of a protein when attempting to produce it in the periplasm of *E. coli*. We also discuss potential strategies to further enhance protein production yields in the periplasm.

## EXPORT OF PROTEINS FROM THE CYTOPLASM TO THE PERIPLASM

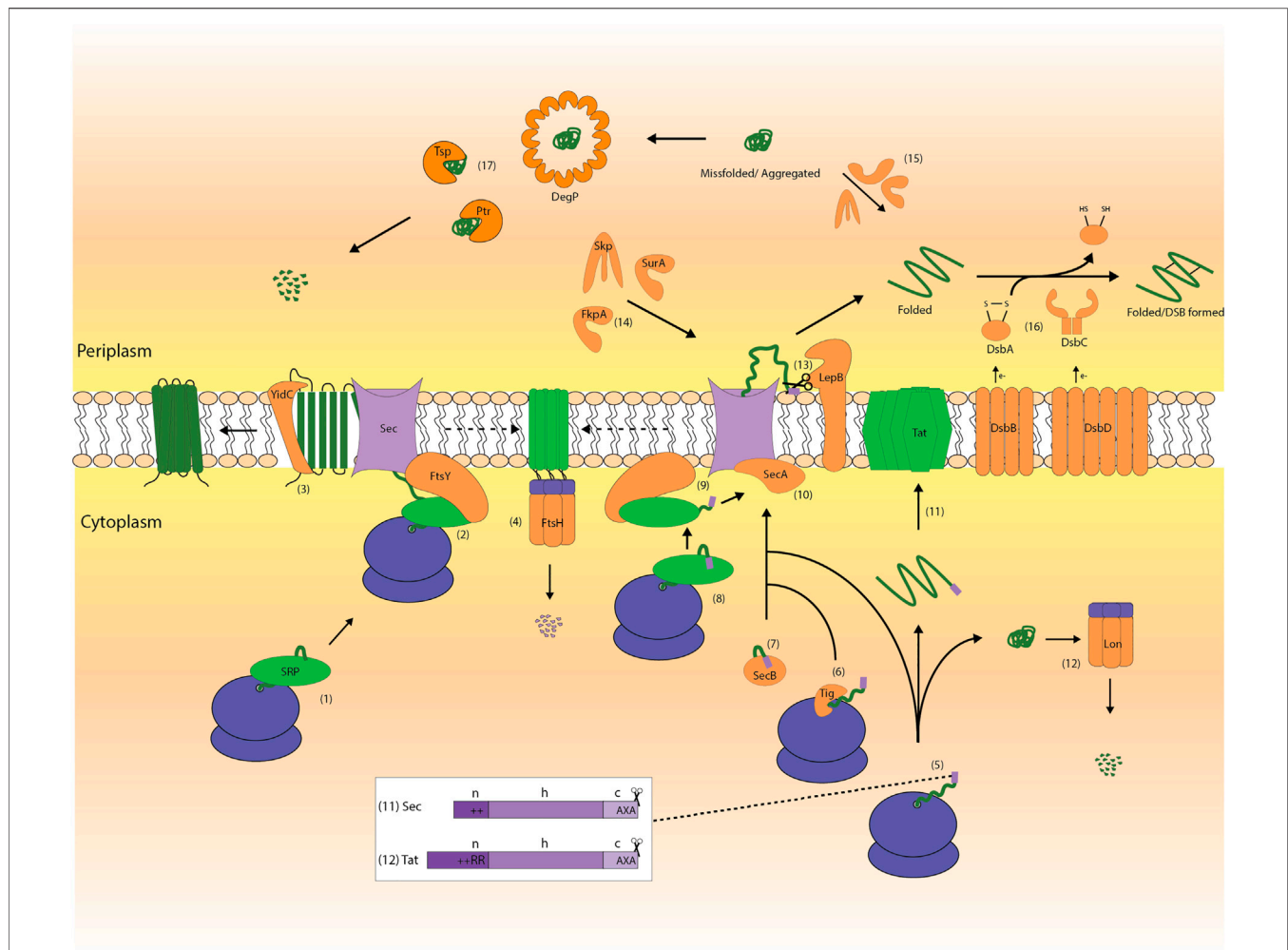
To set the stage for this review, we will first briefly summarize how proteins are exported from the cytoplasm to the periplasm in *E. coli*. A secretory protein is synthesized in the cytoplasm and has at its N-terminus a cleavable signal peptide that targets it to either the Sec-translocon or the Tat-translocon in the cytoplasmic membrane (Frain et al., 2019; Oswald et al., 2021) (**Figure 2**). In *E. coli*, the vast majority of secretory proteins is translocated across the cytoplasmic membrane *via* the Sec-translocon (Smets et al., 2019). During translocation *via* the Sec-translocon, secretory proteins are in an unfolded conformation (De Geyter et al., 2016). Proteins translocated across the cytoplasmic membrane *via* the Tat-translocon fold in the cytoplasm prior to their translocation across the cytoplasmic membrane (Frain et al., 2019; Palmer and Stansfeld, 2020). Upon their translocation the signal peptide is cleaved off from both Sec- and Tat-dependent proteins.

### Protein Targeting

The targeting of proteins to the Sec-translocon can occur either co-translationally *via* the Signal Recognition Particle (SRP)-pathway or post-translationally in a chaperone-dependent or -independent manner (Tsirigotaki et al., 2017; Smets et al., 2019) (**Figure 2**). The SRP-pathway comprises the signal recognition particle (SRP), which is a ribonucleoprotein particle consisting of the 4.5S RNA and the Ffh protein, and its receptor FtsY (Oswald et al., 2021). The cytoplasmic chaperones SecB and Trigger Factor (TF) can assist the post-translational targeting of proteins (Smets et al., 2019; De Geyter et al., 2020). SecB can bind to a subset of secretory proteins and keeps them in a translocation-competent state (Sala et al., 2014). TF, which plays a key role in the folding of cytoplasmic proteins, can also keep secretory proteins in a translocation-competent state and together with SecB aid their export (De Geyter et al., 2020). Targeting of (folded) proteins to the Tat-translocon is mediated by the presence of a signal peptide that contains a highly conserved twin-arginine motif (Sargent et al., 1998; Weiner et al., 1998; Frain et al., 2019; Palmer and Stansfeld, 2020).

### Protein Translocation

The core of the Sec-translocon consists of the membrane proteins SecY and SecE (Oswald et al., 2021) (**Figure 2**). Together they form a protein-conducting channel in the cytoplasmic membrane (Oswald et al., 2021). In the absence of SecE, SecY is degraded by the cytoplasmic membrane protease FtsH (Ito and Akiyama, 2005). The peripheral ATP-dependent motor protein SecA



**FIGURE 2 |** The biogenesis of membrane and secretory proteins in *E. coli*. Membrane and secretory proteins are synthesized in the cytoplasm by the ribosome (**1**, **5**). Most membrane proteins and some secretory proteins are targeted to the Sec-translocon via the SRP-pathway, which comprises the SRP and its receptor FtsY (**1–2**, **8–9**). The membrane protein YidC, which is an insertase/chaperone, can assist the biogenesis of membrane proteins in conjunction with the Sec-translocon as well as an independent entity (**3**). Many secretory proteins are targeted post-translationally to the Sec-translocon in a chaperone-independent (**5**, **10**) or -dependent manner (**6**, **7**, **10**). Chaperones like TF and SecB can facilitate post-translational protein targeting (**6**, **7**). FtsH degrades jammed Sec-translocons (**4**). The TatABC-translocon mediates the translocation of folded proteins with a signal peptide containing the twin arginine motif (**11**). Precursor proteins that get stuck in the cytoplasm can be degraded by cytoplasmic proteases like Lon (**12**). The motor protein SecA pushes secretory proteins through the Sec-translocon channel (**10**). Upon translocation of a protein across the cytoplasmic membrane LepB cleaves off the signal peptide (**13**). Periplasmic folding modulators like SurA, Skp and FkpA catalyze the folding of proteins in the periplasm (**14**, **15**). In the periplasm, the Dsb-system, which consists of DsbA, B, C and D, mediates the formation of disulfide bonds (**16**) and proteins can be degraded by proteases like DegP, Tsp and Ptr (**17**).

pushes secretory proteins through the Sec-translocon channel (Smets et al., 2019; Oswald et al., 2021). The mode of translocation through the Sec-translocon can affect how a protein folds in the periplasm (Kadokura and Beckwith, 2009). The membrane protein YidC, which is an insertase/chaperone, can assist the biogenesis of membrane proteins in conjunction with the Sec-translocon as well as an independent entity (Shanmugaw and Dalbey, 2019). The *E. coli* Tat-translocon comprises multiple copies of the membrane proteins TatA, TatB and TatC (Sargent et al., 1998; Weiner et al., 1998; Frain et al., 2019). The assembly of the Tat-translocon is dynamic and is triggered by a Tat substrate interacting through its signal peptide with the TatBC-receptor complex (Frain et al., 2019). The Tat-

translocon has a quality-control mechanism that prevents translocation of not properly folded proteins (Frain et al., 2019).

Notably, the protein flux through the Tat-translocon is considerably lower than through the Sec-translocon. It takes a few minutes to translocate a Tat-dependent protein *versus* a few seconds for a Sec-dependent protein (Taw et al., 2021). The proton-motive force is required for both efficient Sec- and Tat-mediated protein export (Frain et al., 2019; Oswald et al., 2021).

## Protein Folding and Degradation

Upon translocation of a protein across the cytoplasmic membrane, signal peptidase I (LepB) cleaves off the signal peptide (Paetzel, 2014) (Figure 2). In the periplasm there are

folding modulators that can assist the folding of Sec-dependent proteins (Stull et al., 2018). Proteins that require disulfide bonds are reduced when they are exported from the cytoplasm to the periplasm *via* the Sec-translocon (De Geyter et al., 2016; Smets et al., 2019). In the periplasm, the Dsb-system, which consists of DsbA, B, C and D, mediates the formation of disulfide bonds (Manta et al., 2019). Oxidized DsbA transfers its own disulfide bond to its substrate proteins and is reduced in the process. The membrane protein DsbB re-oxidizes DsbA and it donates the electrons it has received from DsbA to the respiratory chain. DsbC can recognize if the substrate protein is not properly oxidized by DsbA. DsbC either isomerizes substrates that are not properly oxidized to their native state or reduces them thereby allowing DsbA to have another go to introduce a proper disulfide bond. The membrane protein DsbD maintains DsbC in its active reduced state. DsbD receives its electrons from the cytoplasmic thioredoxin TrxA, which receives its electrons from NADPH. Other examples of periplasmic folding modulators are FkpA, which is a periplasmic chaperone/peptidyl-prolyl isomerase (PPIase), SurA, which is a chaperone/PPIase, and the chaperone Skp (Stull et al., 2018). It should be noted that our grasp of the repertoire of periplasmic folding modulators and folding modulator–substrate specificity is still rather poor (Stull et al., 2018). The same holds for periplasmic proteases and periplasmic protease–substrate specificity (Merdanovic et al., 2011).

## Signal Peptides

A signal peptide consists of a positively charged N-terminal region (n-region), an apolar hydrophobic core (h-region) and a more polar C-terminal region (c-region) (Izard and Kendall, 1994; Tsigotaki et al., 2017). The c-region contains an AXA motif (Paetzel, 2014). This motif is recognized by LepB, which cleaves off the signal peptide upon protein translocation across the membrane (Paetzel, 2014). As mentioned before, Tat-dependent signal peptides contain a conserved twin arginine motif in the n-region, at the boundary with the h-region (Frain et al., 2019). The hydrophobicity of the h-region of signal peptides that guide secretory protein to the Sec-translocon plays an important role in the mode of targeting (*e.g.*, (Rusch and Kendall, 1994; Chen et al., 1996; Rusch et al., 2002)). It has been shown that increasing the hydrophobicity of the h-region of a signal peptide can funnel proteins that are normally targeted post-translationally into the co-translational SRP-pathway and *vice versa* (Kim et al., 2001). Furthermore, it has been shown that targeting of an export-defective protein can be rescued by increasing the hydrophobicity of the h-region (Rusch and Kendall, 1994). By funneling the protein into the co-translational SRP-pathway the premature folding of the protein in the cytoplasm is most likely prevented (Rusch and Kendall, 1994). Signal peptides of post-translationally targeted proteins can help to delay the folding of the mature domain in the cytoplasm (Tsigotaki et al., 2018). Recently, it has been shown that post-translational targeting/translocation is not only mediated by the signal peptide, but also by the so-called mature domain targeting sites in the mature part of the preprotein that can also bind to SecA (Chatzi et al., 2017; Smets et al., 2019).

## STRATEGIES TO ENHANCE PERIPLASMIC PROTEIN PRODUCTION YIELDS

Strategies to enhance periplasmic protein production yields have focused on harmonizing protein production rates with the capacity of the secretory apparatus by transcriptional and translational tuning, by signal peptide selection and engineering, by increasing the targeting, translocation and periplasmic folding capacity of the cell, by preventing proteolysis, and by natural and engineered adaptation of the production host. Using notable examples as a thread, we will go through the different strategies used to enhance periplasmic protein production yields. Importantly, if the reader contemplates to use a certain strategy to enhance periplasmic protein production yields, the technical setup used to implement the strategy should be carefully considered.

### Tuning Transcription and Translation

Studies aiming to enhance membrane protein production yields in *E. coli* have played a key role in identifying what can hamper periplasmic protein production. Therefore, first a brief introduction to these studies, in which the BL21(DE3) strain and T7 promoter-based expression vectors were used for protein production, will be given (Studier and Moffatt, 1986). In BL21(DE3), the expression of the gene encoding the target protein is driven by the bacteriophage T7 RNA polymerase (Studier and Moffatt, 1986). This polymerase transcribes approximately 5 times faster than *E. coli* RNA polymerase (Chamberlin et al., 1970; Studier and Moffatt, 1986). The T7 RNA polymerase specifically recognizes the T7 promoter governing the expression of the gene encoding the target protein (Studier and Moffatt, 1986). Expression of the chromosomally localized gene encoding the T7 RNA polymerase is governed by the IPTG-inducible *lacUV5* promoter (Studier and Moffatt, 1986). This promoter is a more powerful variant of the wild-type *lac* promoter and it is not well titratable (Silverstone et al., 1970; Wanner et al., 1977). The idea behind this setup is that the more mRNA encoding the recombinant protein is synthesized, the more protein can be produced. In this respect it is good to mention that the system was developed for the production of soluble proteins in the cytoplasm (Studier and Moffatt, 1986).

The over-production of membrane proteins in BL21(DE3) is usually toxic (Wagner et al., 2007). However, many membrane proteins can be efficiently produced in the BL21(DE3)-derived, C41(DE3) and C43(DE3) strains (Miroux and Walker, 1996). Based on the observation that in the C41(DE3) and C43(DE3) strains the promoter governing the expression of *t7rnsp* is weakened, the BL21(DE3)-derived Lemo21(DE3) strain was constructed (Wagner et al., 2008). In this strain, the activity of the T7 RNA polymerase can be precisely controlled by its natural inhibitor T7 lysozyme (Moffatt and Studier, 1987). The gene encoding the T7 lysozyme is on a plasmid under control of a rhamnose promoter. This promoter is titratable and covers a broad range of expression intensities (Giacalone et al., 2006). The combination of the *lacUV5* promoter governing the expression of the gene encoding the T7 RNA polymerase and the rhamnose

promoter governing the expression of the gene encoding T7 lysozyme, enables to tune the expression intensity of the gene encoding the target membrane protein from a T7 promoter-based expression vector (Wagner et al., 2008). As a consequence, the membrane protein production rate can be harmonized with the capacity of the membrane protein biogenesis apparatus of the cell using transcriptional tuning (Wagner et al., 2008; Schlegel et al., 2012). This results in cells in which the capacity of membrane protein biogenesis apparatus is not saturated. Therefore, these cells can efficiently produce membrane proteins (Wagner et al., 2008; Schlegel et al., 2012).

The Lemo21(DE3) strain was also used to test if saturation of the secretory apparatus can cause problems when attempting to produce proteins in the periplasm (Schlegel et al., 2013). Indeed, periplasmic protein production can also lead to saturation of the secretory apparatus. This results in poor periplasmic protein production yields and accumulation of the precursor target protein in the cytoplasm, poor growth, precursor accumulation of endogenous secretory proteins in the cytoplasm and the induction of the heat shock response, which is indicative for protein misfolding and mistargeting (Schlegel et al., 2013). Harmonizing the production rate of a secretory protein with the capacity of the secretory apparatus leads to more protein produced in the periplasm and strongly reduced negative effects (Schlegel et al., 2013). When the proteome composition of Lemo21(DE3) cells producing a single chain antibody fragment (scFv) under optimized conditions was compared with the proteome composition of Lemo21(DE3) cells with an empty expression vector, there were no notable differences (Baumgarten et al., 2018). Thus, the scFv could be efficiently produced in the periplasm without causing any notable stress or other side effects.

The rhamnose inducible promoter can also be used to directly express a gene encoding a protein (Giacalone et al., 2006; Hjelm et al., 2017). When the rhamnose promoter-mediated production of (cytoplasmic) super folder GFP was monitored in real-time using different amounts of rhamnose in *E. coli* wild-type cells, it became clear that the rhamnose concentration-dependent tunability is due to rhamnose consumption rather than regulating protein production rates (Hjelm et al., 2017). To regulate rhamnose promoter-mediated protein production rates in a rhamnose concentration dependent manner, a strain background in which both rhamnose catabolism as well as rhamnose uptake by the RhaT-transporter are abolished has to be used (Hjelm et al., 2017). Unfortunately, there is no straight forward explanation for this observation. Thus, using the rhamnose promoter in a rhamnose catabolism and *rhaT* deficient strain background enables harmonizing target protein production rates with the capacity of the secretory apparatus in a rhamnose concentration dependent manner (Hjelm et al., 2017). This setup has been used successfully to boost periplasmic protein production yields (Hjelm et al., 2017; Karyolaimos et al., 2019).

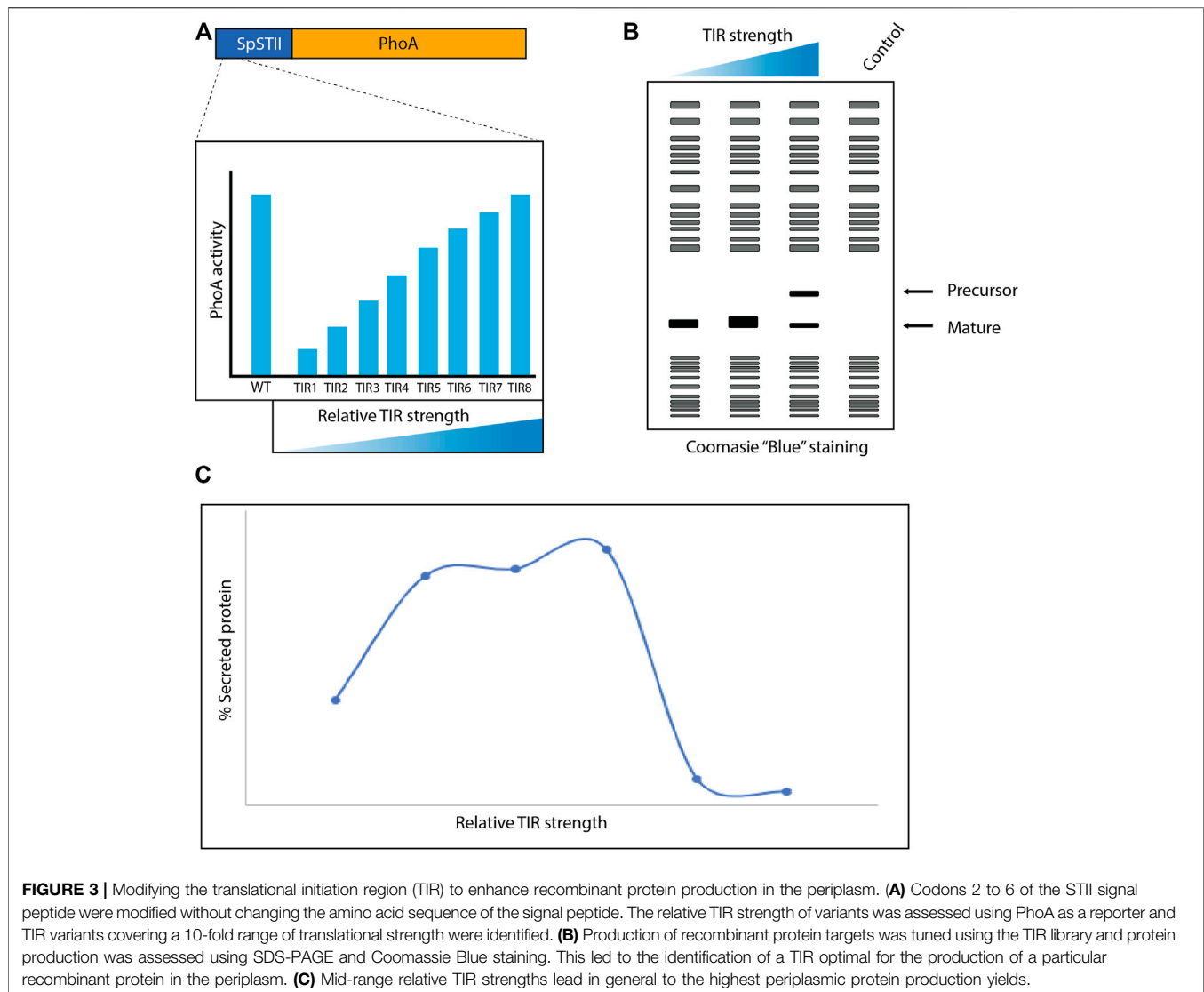
Researchers from Genentech reported in 1996 that the secretion of recombinant proteins into the periplasm of *E. coli* could be enhanced by optimizing translational levels rather than maximizing them (Simmons and Yansura, 1996) (Figure 3). The signal peptide from the *E. coli* heat-stable enterotoxin II (STII) in

combination with the phosphate limitation inducible *phoA* promoter was used to produce a series of recombinant proteins in the periplasm (Chang et al., 1987; Simmons and Yansura, 1996). Notably, STII is targeted post-translationally to the Sec-translocon. To tune translational levels, the translational initiation region (TIR), which covers the immediate upstream part of the ribosome binding site to approximately 20 nucleotides downstream of the initiation codon, was modified (Goodman et al., 2013). More precisely, codons 2 to 6 of the STII signal peptide were modified without changing the amino acid sequence of the signal peptide to create a library of vectors with different translational strengths. Subsequently, the *E. coli* alkaline phosphatase (PhoA) was used as a reporter to characterize the library. This enabled to identify TIR variants covering a 10-fold range of translational strength. These variants were used to enhance periplasmic protein production. Importantly, for each target tested a narrow translational range was required for optimal periplasmic protein production. Interestingly, when an optimal TIR had been identified for a target and the copy number of the expression vector used was increased, periplasmic protein production levels were strongly reduced and precursor material accumulated in the cytoplasm. At the time the authors could only speculate about the underlying mechanisms behind their findings. However, based on what we know now it is obvious that using different TIRs the production rate of a secretory recombinant protein can be harmonized with the capacity of the secretory apparatus, thereby enhancing periplasmic production yields. Increasing the copy number of the expression vector leads to increased protein production and, consequently, the detrimental saturation of the secretory apparatus (Schlegel et al., 2013; Baumgarten et al., 2018).

Taken together, harmonizing secretory protein production rates with the capacity of the secretory apparatus by transcriptional and translational tuning can lead to enhanced periplasmic protein production yields. To regulate secretory protein production rates at the transcriptional level, titratable promoter systems other than the ones described above can also be used (Meyer et al., 2019). To regulate secretory protein production rates at the translational level the ribosome binding site of a TIR can also be modified (Salis et al., 2009; Eichmann et al., 2019). Rather than using TIR libraries, screening approaches can also be used to identify a suitable TIR (Rennig et al., 2018). The T7 RNA polymerase-based RiboTite system, which was recently used to enhance protein production yields in the periplasm, is an example of a system that operates both at the transcriptional and translational level (Horga et al., 2018). In the RiboTite system, transcriptional control is mediated by an IPTG-inducible promoter controlling *t7rnapi* expression (Morra et al., 2016). Translational control in the system is mediated by an orthogonal riboswitch (Morra et al., 2016). This switch sequesters the ribosome binding site in the absence and releases it in the presence of pyrimido-pyrimidine-2,4-diamine.

## Signal Peptide Selection and Engineering

With our current knowledge it is not possible to predict what signal peptide one should use to efficiently produce a protein in the periplasm of *E. coli* (Karyolaimos et al., 2019). Often people

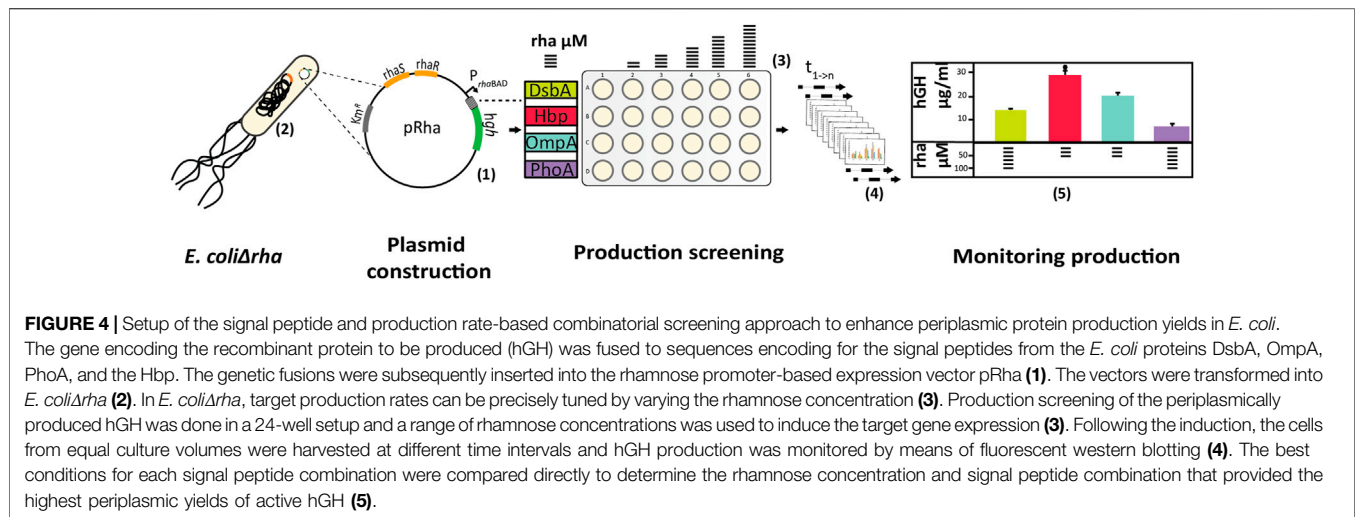


use signal peptides from abundantly present native proteins or signal peptides that have successfully been used before for periplasmic protein production (Kendall et al., 1986; Schierle et al., 2003; Sijbrandi et al., 2003; Schlegel et al., 2013; Schmidt et al., 2016; Hjelm et al., 2017; Baumgarten et al., 2018; Karyolaimos et al., 2019). However, there is no guarantee that these signal peptides will also facilitate the efficient production of any other protein in the periplasm (Karyolaimos et al., 2019). Systematic screening or engineering approaches to identify signal peptides that enhance periplasmic protein production yields are scarce.

To improve the display levels of a series of proteins on filamentous phages, the Plückthun laboratory used a panel of signal peptides that promote co-translational as well as post-translational targeting of phage coat protein fusions to the Sec-translocon (Steiner et al., 2006). Display levels of the proteins of interest were markedly higher when signal peptides promoting co-translational targeting were used and for some targets levels

could be enhanced up to seven hundred fold. Co-translational targeting of the fusions may prevent folding of the protein of interest moieties in the cytoplasm, which would prevent translocation and the subsequent display of the fusions.

Importantly, not only signal peptides can significantly impact periplasmic protein production yields, but as discussed in the previous section also protein production rates. Therefore, our lab set up a combined screen involving four different signal peptides and varying production rates to produce a scFv and human growth hormone (hGH) in the periplasm (Karyolaimos et al., 2019) (Figure 4). The signal peptides from the *E. coli* DsbA, Hbp, OmpA and PhoA proteins were used to guide the two targets to the periplasm. DsbA is, as explained before, a periplasmic thiol: disulfide oxidoreductase and it is generally assumed that the DsbA signal peptide mediates the co-translational targeting of proteins (Schierle et al., 2003). The DsbA signal peptide has been widely and successfully used for the production of recombinant proteins in the periplasm (e.g., (Steiner et al., 2006)). The



Hemoglobin protease (Hbp) is an abundantly produced autotransporter and is co-translationally targeted (Sijbrandi et al., 2003). OmpA is an outer membrane protein and targeted post-translationally in a SecB-dependent manner (Baars et al., 2006). PhoA is a periplasmic enzyme that catalyses the hydrolysis and transphosphorylation of a wide variety of phosphate monoesters and it has been reported that PhoA targeting can occur both post- and co-translationally (Kendall et al., 1986; Kadokura and Beckwith, 2009). The previously discussed rhamnose promoter (see section “Tuning transcription and translation”) and a strain background in which the whole *rha* operon is deleted was used to produce the scFv and hGH equipped with the four different signal peptides at different production rates. Across the screen conditions, the periplasmic production yields of both targets varied a lot. The optimal signal peptide and rhamnose concentration differed for each protein. Interestingly, for scFv the OmpA signal and for hGH the Hbp signal peptide led to the best periplasmic production yields. Thus, the signal peptide from the post-translationally targeted OmpA protein and the signal peptide from a co-translationally targeted protein Hbp protein result in optimal periplasmic production of the scFv and hGH, respectively. There is no straightforward explanation for this observation, but it stresses that signal peptides derived from post-translationally targeted proteins cannot be excluded *a priori* for efficient periplasmic protein production. This study also shows the importance of combinatorial screening approaches for enhancing periplasmic protein production yields.

Rather than trying to identify a suitable signal peptide for the periplasmic production of a protein in the periplasm, such a signal peptide can also be engineered. In a study from Genentech, signal peptide engineering was used to promote the secretion of the heavy chain of a monoclonal antibody (mAb) that was limiting the production of the mAb in the periplasm (Zhou et al., 2016). To improve the secretion of the heavy chain it was produced using different signal peptides at controlled TIR strengths. The use of the DsbA signal peptide resulted in the highest periplasmic production yields compared to the other

signal peptides tested. A close inspection of the different signal peptides used showed that the hydrophobicity of the h-region of the DsbA signal peptide was higher than the hydrophobicity of the h-region of the other signal peptides. This raised the question if the hydrophobicity of the h-region of a signal peptide can play an important role in periplasmic protein production. To explore this the signal peptide of the aforementioned STII protein was used since its n- and c-regions are very similar to the same regions of the DsbA signal peptide. The hydrophobicity of the h-region of the STII signal peptide is lower and increasing the hydrophobicity of its h-region enhanced secretion of the heavy chain and periplasmic mAb production levels. Further increasing the hydrophobicity of the h-region did not result in further enhancing periplasmic production yields. Decreasing the hydrophobicity of the h-region of the DsbA signal resulted in the opposite effect. Thus, in this study promoting co-translational targeting led to enhanced periplasmic protein production.

In several studies it has been shown that the presence of rare codons in the genetic information encoding a signal peptide, especially in the region encoding the n-region, can have a positive impact on both periplasmic protein production yields and on proper protein folding (e.g., (Humphreys et al., 2000; Zalucki et al., 2008; Zalucki et al., 2010; Castiñeiras et al., 2018; Kulmala et al., 2019)). It is tempting to speculate that these rare codons make the TIR/protein translocation kinetics more amenable to efficient periplasmic protein production.

## Increasing the Capacity of the Secretory Apparatus

### Increasing Protein Targeting Capacity

Co-producing the cytoplasmic chaperones GroEL and DnaK, which both have also been implicated in protein secretion, can facilitate periplasmic protein production (Hu et al., 2007; Makino et al., 2011b; Sonoda et al., 2011). Attempts to enhance periplasmic protein production yields by co-producing TF and SecB have not been successful (Puertas et al., 2010; Makino et al., 2011b). TF can hamper protein secretion by blocking access of the



precursor protein to the secretory apparatus (De Geyter et al., 2020). It is possible that increased levels of SecB hamper secretion by retaining a secretory protein in the cytoplasm rather than promoting its secretion. Interestingly, the production of the secreted form of a leech carboxypeptidase inhibitor, targeted by the DsbA signal peptide to the Sec-translocon, is increased in a TF deficient strain background (Puertas et al., 2010). It is possible that in spite of using the DsbA signal peptide, targeting of the leech carboxypeptidase inhibitor occurs only partially co-translationally due to limited SRP-pathway capacity. If this would be the case, the secretion of the target protein could indeed be hampered by TF. This scenario is supported by the observations that co-producing SRP (both Ffh as well as the 4.5S RNA) also increased production of the leech carboxypeptidase inhibitor. Moreover, combining a TF deficient strain background with the co-production of SRP had an additive effect on the production of secreted target protein and strongly reduced the accumulation of the precursor form of the target protein in the cytoplasm (Puertas et al., 2010). Co-production of components of the SRP-targeting pathway, including FtsY, have been widely used to facilitate the co-translational targeting of secretory recombinant proteins (e.g., (Lee et al., 2013)).

### Increasing Protein Translocation Capacity

It has been shown that co-production of the Sec-translocon core components SecY and E can lead to increased periplasmic protein production yields (Makino et al., 2011b). However, when SecY and E were co-produced to increase periplasmic production of human interleukin-6 equipped with an OmpA signal peptide, periplasmic production yields were not increased (Pérez-Pérez et al., 1994). Interestingly, when a mutant version of SecY, *i.e.*, PrlA4, was co-produced rather than wild-type SecY, the periplasmic production yield of human interleukin-6 increased significantly. PrlA4 was isolated in a screen aiming at the isolation of suppressor mutations that could restore the export of LamB equipped with defective signal peptides (Emr et al., 1981). Co-production of PrlA4 without SecE provided a three-fold increase in export of the target protein compared to the 10-fold increase when both SecE and PrlA4 were co-produced. Co-production of SecE with PrlA4 is most likely needed to prevent SecY/PrlA4 degradation by FtsH (Nijeholt et al., 2013). PrlA4 has not only been shown to enable the translocation of proteins with a defective signal peptide or without a signal peptide, but it also has been shown to have a higher binding affinity towards SecA (van der Wolk et al., 1998). This promotes binding of SecA as well as the precursor protein at the translocation site. This may improve the translocation efficiency of proteins the wild-type Sec-translocon has difficulties with translocating (van der Wolk et al., 1998).

The Tat pathway has also been used to produce proteins, including ones with disulfide bonds, in the periplasm of *E. coli* (e.g., (DeLisa et al., 2003; Matos et al., 2014; Alanen et al., 2015; Browning et al., 2017; Montero et al., 2019)). Co-producing the components of the Tat-translocon has been used successfully to increase the Tat-pathway capacity of *E. coli*, thereby enhancing Tat-mediated periplasmic protein production yields (Browning et al., 2017; Montero et al., 2019). To enhance the periplasmic

production of properly folded disulfide bond containing proteins in *E. coli*, co-producing Tat-translocon components has been combined with the CyDisCo (Cytoplasmic Disulfide bond formation in *E. coli*) setup. The CyDisCo setup is based on the cytoplasmic production of a yeast mitochondrial thiol oxidase, Erv1p, and the human protein disulfide isomerase, PDI, both lacking signal peptides (Gaciarz et al., 2016).

Recently, Tat-translocons were isolated enabling high level periplasmic protein production (Taw et al., 2021). To isolate these translocons, cells with the native *tat* operon deleted and expressing a mutant library of *tatABC* operons and at the same time producing a scFv equipped with a Tat-dependent signal peptide and C-terminally fused to TEM- $\beta$  lactamase were put on plates containing a high amount of carbenicillin. This setup enabled to use growth for selecting cells with an increased Tat-export capacity and resulted in the isolation of three “super-secreting” Tat-translocons. These translocons not only mediated the efficient periplasmic production of the target used in the screen, but also of other proteins at levels that are higher than the ones obtained with the wild-type Tat-translocon. All three “super-secreting” Tat-translocons had mutations in the *tatB* and *tatC* genes and no mutations in the *tatA* gene. Using targets that are rejected by the wild-type Tat-translocon it was shown that the quality-control mechanism that prevents translocation of not properly folded proteins in the mutant Tat-translocons appears to be more relaxed. This most likely enables an increased protein flux through these translocons. Importantly, this does not affect the quality of the in the periplasm produced proteins.

### Increasing Signal Peptide Processing Capacity

Co-producing LepB, which cleaves off signal peptides upon protein translocation, can also increase periplasmic protein production yields (van Dijl et al., 1991). It has been shown that increasing the synthesis of protease IV (SppA) by modifying its TIR also results in enhanced periplasmic protein production (Gawin et al., 2020). SppA is a membrane-bound signal peptide peptidase that is required for maintaining proper protein secretion (Novak and Dev, 1988; Dalbey et al., 2012).

### Increasing Periplasmic Protein Folding Capacity

Periplasmic folding modulators have also been co-produced to enhance periplasmic protein production yields. Co-production of one of the or both key players of the Dsb-system, *i.e.*, DsbA and DsbC, has been widely used to enhance the production of disulfide bond containing proteins in the periplasm (e.g., (Joly et al., 1998; Schlapschy et al., 2006; Makino et al., 2011b; Sonoda et al., 2011)). Co-production of e.g., FkpA, SurA and Skp has also helped to enhance periplasmic and extracellular protein production yields (e.g., (Mavrangelos et al., 2001; Schlapschy et al., 2006; Hu et al., 2007; Makino et al., 2011a; Sonoda et al., 2011)). To facilitate the identification of folding modulators enhancing periplasmic protein production yields, expression vectors that enable the co-production of several folding modulators at the same time have been engineered (Schlapschy et al., 2006). It has been shown that culturing setups can determine what folding modulator has to be co-

produced (Ellis et al., 2017). To enhance periplasmic production of an antibody fragment in small shake flasks FkpA had to be co-produced. However, in a high cell density fed-batch culture DsbC had to be co-produced. This observation raises the question if the co-production of a folding modulator can also have indirect positive effects. Indeed, it has been shown that co-producing folding modulators can have a positive effect on the fitness of *E. coli*, which also can help to enhance periplasmic protein production yields (Ow et al., 2010).

Most often two compatible expression vectors have been used for producing the folding modulator and the recombinant protein (e.g., (Joly et al., 1998; Schlapschy et al., 2006; Makino et al., 2011b; Sonoda et al., 2011)). Usually, different promoter systems are used to govern the expression of the genes encoding the folding modulator and the protein. The different promoter systems can also be used for the sequential production of a folding modulator and the protein (Joly et al., 1998; Makino et al., 2011b). Thus, cells can be pre-loaded with a folding modulator before the production of the protein is initiated. This better prepares cells for the production of the protein and lowers the risk of saturating the secretory apparatus. Also artificial operons containing the genes encoding both the folding modulator and the protein have been used (Ellis et al., 2017). This setup requires only one expression vector and one promoter system and is in particular useful for large scale protein production.

## Preventing Proteolysis

An important reason to produce a protein in the periplasm is to minimize its exposure to cytoplasmic proteases (Talmadge and Gilbert, 1982). However, this does not mean that there are no proteases in the cell envelope which can negatively affect periplasmic protein production (Gottesman, 1996). Indeed, it has been shown that the three periplasmic proteases DegP, Tsp and Ptr, can negatively affect protein production yields in the periplasm (Meerman and Georgiou, 1994; Chen et al., 2004). Most work on the identification of proteases affecting periplasmic protein production yields was done at a time when these proteases were the main ones that had been characterized (Gottesman, 1996). Only little is known about their substrates and how they recognize them. Therefore, it cannot be predicted which protease or combination of proteases can negatively affect production yields of a particular protein in the periplasm. This makes that proteases interfering with periplasmic protein production have to be identified by screening protein production yields in protease deficient strains.

The serine endoprotease DegP was identified in a screen aiming at the isolation of mutations preventing the degradation of abnormal periplasmic proteins (Strauch and Beckwith, 1988). DegP may degrade transiently denatured, unfolded proteins and/or newly secreted proteins prior to folding and disulfide bond formation (Strauch et al., 1989; Laskowska et al., 1996; Skorko-Glonek et al., 1999; Soltes et al., 2017). Inactivation of DegP by deleting the *degP* gene often results in improved periplasmic protein production yields (e.g., (Meerman and Georgiou, 1994; Champion et al., 2001; Chen et al., 2004)). DegP can also act as a chaperone (Meltzer et al.,

2009). The protease function of DegP can be inactivated by a simple point mutation without affecting its chaperone activity (Spiess et al., 1999). This mutation is often used to create DegP protease deficient strains that still have DegP chaperone activity (e.g., (Ellis et al., 2017)). Rather than deleting the gene encoding DegP or inactivating its protease function, periplasmic protein production yields have also been improved by lowering DegP levels. This was done by replacing the ribosome binding site in the TIR controlling *degP* expression with a weaker one (Gawin et al., 2020).

Both Tsp and Ptr were first detected when attempting to identify cytoplasmic proteases suggesting that these two periplasmic proteases are abundantly present (Cheng et al., 1979; Swamy and Goldberg, 1981). Tsp can degrade proteins with nonpolar C-termini and based on this observation it was named Tail-specific protease (Keiler and Sauer, 1996). It has been reported that Tsp can degrade antibody light chain fragments (Chen et al., 2004; Ellis et al., 2017). Tsp is also involved in regulating peptidoglycan synthesis by degrading MepS, which is a murein endopeptidase participating in the expansion of the peptidoglycan sacculus during growth and morphogenesis (Singh et al., 2015). Deletion of the gene encoding Tsp impairs growth. Inactivating the *mepS* gene in a Tsp deficient background restores growth (Chen et al., 2004; Ellis et al., 2017). Ptr (Protease III) is a zinc metalloendopeptidase that has been shown to degrade insulin (Swamy and Goldberg, 1981; Becker and Roth, 1992; Ding et al., 1992).

OmpT is an outer membrane protease with specificity for paired basic residues and its catalytic site is facing the bacterial surface (Kramer et al., 2001). OmpT can degrade periplasmically produced proteins after cell lysis (Grodberg and Dunn, 1988; Akiyama and Ito, 1990). Since proteins produced in the periplasm are usually isolated or can leak in the extracellular medium often OmpT deficient production strains are used (Jeong et al., 2009; Studier et al., 2009; Jeong et al., 2015).

Proteases do not only negatively affect periplasmic protein production yields, they can also be required for efficient periplasmic protein production. Secretory proteins can get stuck in the Sec-translocon thereby jamming it (Jiang et al., 2021). The cytoplasmic membrane protease FtsH, which is a processive, ATP-dependent zinc metallopeptidase for both cytoplasmic and cytoplasmic membrane proteins, can clear jammed Sec-translocons (van Stelten et al., 2009). Furthermore, it has been shown that the cytoplasmic Lon protease can clear not properly targeted secretory proteins in the cytoplasm (Kihara et al., 1995; van Stelten et al., 2009). It has been proposed that the cytoplasmic peptidase, PrlC (oligopeptidase A) can remove signal peptides from not properly targeted secretory proteins, thereby setting the stage for their clearance (Conlin et al., 1992).

## Natural and Engineered Adaptation of the Production Host

*E. coli* can respond to stress by changing its proteome composition (e.g., (Wagner et al., 2007; Baumgarten et al., 2018)). It has been shown that if secretory protein production

results in saturation of the secretory apparatus the heat shock response is activated to clear mis-targeted and -folded proteins and increase the protein (re) folding capacity of the cell (e.g., (Wagner et al., 2007; Baumgarten et al., 2018)). Using the Lemo21(DE3) strain it has been shown that harmonizing the production rate of a secretory protein with the capacity of the secretory apparatus can result in cells that have a proteome composition that is similar to the proteome composition of non-producing cells (Baumgarten et al., 2018).

In the section “Tuning transcription and translation”, it was described how hGH was produced using the rhamnose promoter in a *rha* operon deletion strain background at varying concentrations of rhamnose and four different signal peptides. Cells producing hGH in the periplasm at the highest level, for each of the four signal peptides used, were analyzed using proteomics (Karyolaimos et al., 2020). Irrespective of the signal peptide used, the accumulation levels of SecA, LepB and YidC were increased. Thus, enhancing periplasmic hGH production using the rhamnose promoter in a *rha* operon deletion strain background leads to increased Sec-translocon capacity, increased capacity to cleave signal peptides from secretory proteins and an increased capacity of an alternative membrane protein biogenesis pathway, thereby freeing up Sec-translocon capacity for protein secretion. When cells with enhanced periplasmic hGH production yields were harvested and subsequently cultured in the absence of inducer, SecA, LepB, and YidC accumulation levels went down again. This indicates that *E. coli* can adapt its protein secretory apparatus for enhanced recombinant protein production in the periplasm. The gene encoding SecA is in an operon preceded by the gene encoding SecM, which is a secretion monitor (Ito et al., 2018). SecM is encoded by the 5' region of the *secM-secA* mRNA. SecM translation is subject to transient elongation arrest, which is prolonged when the secretion of the nascent SecM is hampered. This results in disrupting the secondary structure of the *secM-secA* mRNA by the stalled ribosome such that the ribosome binding site of *secA* becomes available leading to the synthesis of SecA to relieve the export defect. Unfortunately, we have only limited knowledge on how LepB and YidC accumulation levels are regulated. At any rate, optimizing periplasmic protein production yields can also lead to adaptations that make the *E. coli* cell a more efficient protein factory.

It has been shown that the accumulation levels of phage shock protein A (PspA) can increase upon the production of a protein in the periplasm (Champion et al., 2003). PspA is encoded by the phage shock protein (*psp*) operon (*pspABCDE*) which plays a key role in maintaining membrane integrity and the proton-motive force under various stress conditions (Mitchell and Silhavy, 2019). Interestingly, co-production of PspA can enhance both Sec- and Tat-mediated protein export (Champion et al., 2003; DeLisa et al., 2004). It is tempting to speculate that increased PspA accumulation levels can make *E. coli* more fit for efficient periplasmic protein production by maintaining the integrity of the cytoplasmic membrane and the proton-motive force (Mitchell and Silhavy, 2019).

Finally, to facilitate the adaptation of *E. coli* for the production of a protein in the periplasm an approach called global transcriptional machinery engineering (gTME) has

been used (McKenna et al., 2019). More specifically, to enhance the production of full length antibodies in the periplasm a mutant library of the principal sigma factor RpoD, a.k.a.  $\sigma^{70}$ , was co-produced with the protein. For the sake of clarity, it should be mentioned that the cells still had an intact wild-type *rpoD* gene. Various RpoD variants were isolated that helped to increase full length antibody production in the periplasm. Transcriptome analysis showed that the RpoD variants altered the transcriptomes significantly, but did not really help to reveal why in cells with the RpoD variants production of full length antibodies in the periplasm was enhanced.

## CONCLUDING REMARKS AND FUTURE PERSPECTIVES

Many different strategies have been used to enhance periplasmic protein production yields in *E. coli*. Usually, these are based on changing one parameter or variable at a time. So far, combinatorial screens, i.e., screens changing different parameters or variables simultaneously, have hardly been used (Eichmann et al., 2019; Karyolaimos et al., 2019). New technologies and methodologies facilitating large-scale culturing and monitoring protein production levels will greatly facilitate combinatorial screening using different strains, expression vectors, signal peptides, induction regimes, culturing methods and conditions (Krause et al., 2016; Tripathi and Shrivastava, 2019; Kastenhofer et al., 2021). Tunable and independently controlled promoter setups, like the ones of the recently developed “Marionette” strain/expression vector collection, as well as CRISPR-based gene regulators have tremendous potential to facilitate combinatorial co-production screening (Meyer et al., 2019; Reis et al., 2019; Ho et al., 2020). The “Marionette” promoter setups and CRISPR-based gene regulators enable screening for the impact on periplasmic protein production using a myriad of different expression intensities of the gene encoding the recombinant protein and the gene encoding the co-produced component. Usually, one and sometimes two components are co-produced at a time, whereas co-production of multiple (>2) components may actually be needed to further enhance production yields. With the current DNA synthesis and cloning technologies, it should not be difficult to engineer versatile expression vector library setups for multi-component co-production strategies (Hughes and Ellington, 2017). Aforementioned “Marionette” promoter setups and CRISPR-based gene regulators will enable tunable and independently controlled expression of the genes encoding all actors involved. The co-production of components can also have indirect positive effects on periplasmic protein production yields (Ow et al., 2010). Furthermore, it has been shown that (the co-production of) variants of components involved in the protein biogenesis can also help to improve periplasmic protein production yields (Pérez-Pérez et al., 1994; Taw et al., 2021). Isolation of such variants will require cleverly designed screens or selections. Also hosts can be customized for the production of a particular

recombinant protein in the periplasm using evolutionary and engineering approaches (Wang et al., 2009; Robertson et al., 2021; Wannier et al., 2021).

It should not be forgotten that our knowledge of gene expression, protein translation, protein targeting, translocation, folding and stability is everything but complete. Furthering our knowledge in these areas may also open up new avenues for enhancing periplasmic protein production yields. Studying *E. coli* cells producing proteins in the periplasm using omics approaches, may help to identify components that can positively and also negatively affect periplasmic protein production yields (e.g., (Ceroni et al., 2015; Reis et al., 2019; Karyolaimos et al., 2020; Mateus et al., 2020; Tan et al., 2020; Rugbjerg et al., 2021)). Such studies may extend the repertoire of components that can be co-produced and used for the isolation of variants as well as genes that can be inactivated to enhance periplasmic protein production yields.

Taken together, different strategies have successfully been used to enhance periplasmic protein production yields in *E. coli* and there appears to be ample room for further improvements. The more information becomes available, the easier it may become to develop predictors and other computer-aided tools that can assist in designing strategies to enhance periplasmic protein production yields (Salis et al., 2009; Orfanoudaki et al., 2017; Almagro Armenteros et al., 2019; Reis and Salis, 2020; Bhandari et al., 2021; Cetnar and Salis, 2021).

## REFERENCES

- Akiyama, Y., and Ito, K. (1990). SecY Protein, a Membrane-Embedded Secretion Factor of *E. coli*, Is Cleaved by the *ompT* Protease *In Vitro*. *Biochem. Biophysical Res. Commun.* 167 (2), 711–715. doi:10.1016/0006-291x(90)92083-c
- Alanen, H. I., Walker, K. L., Lourdes Velez Suberbie, M., Matos, C. F. R. O., Bönsch, S., Freedman, R. B., et al. (2015). Efficient export of Human Growth Hormone, Interferon  $\alpha$ 2b and Antibody Fragments to the Periplasm by the *Escherichia coli* Tat Pathway in the Absence of Prior Disulfide Bond Formation. *Biochim. Biophys. Acta (Bba) - Mol. Cel Res.* 1853 (3), 756–763. doi:10.1016/j.bbamcr.2014.12.027
- Almagro Armenteros, J. J., Tsigirig, K. D., Sønderby, C. K., Petersen, T. N., Winther, O., Brunak, S., et al. (2019). SignalP 5.0 Improves Signal Peptide Predictions Using Deep Neural Networks. *Nat. Biotechnol.* 37 (4), 420–423. doi:10.1038/s41587-019-0036-z
- Baars, L., Ytterberg, A. J., Drew, D., Wagner, S., Thilo, C., van Wijk, K. J., et al. (2006). Defining the Role of the *Escherichia coli* Chaperone SecB Using Comparative Proteomics. *J. Biol. Chem.* 281 (15), 10024–10034. doi:10.1074/jbc.M509929200
- Baumgarten, T., Ytterberg, A. J., Zubarev, R. A., and de Gier, J.-W. (2018). Optimizing Recombinant Protein Production in the *Escherichia coli* Periplasm Alleviates Stress. *Appl. Environ. Microbiol.* 84 (12), e00270. doi:10.1128/AEM.00270-18
- Becker, A. B., and Roth, R. A. (1992). An Unusual Active Site Identified in a Family of Zinc Metalloendopeptidases. *Proc. Natl. Acad. Sci.* 89 (9), 3835–3839. doi:10.1073/pnas.89.9.3835
- Bhandari, B. K., Lim, C. S., and Gardner, P. P. (2021). TISIGNER.com: Web Services for Improving Recombinant Protein Production. *Nucleic Acids Res.* 49 (W1), W654–W661. doi:10.1093/nar/gkab175
- Browning, D. F., Richards, K. L., Peswani, A. R., Roobol, J., Busby, S. J. W., and Robinson, C. (2017). *Escherichia coli* "TatExpress" Strains Super-secrete Human Growth Hormone into the Bacterial Periplasm by the Tat Pathway. *Biotechnol. Bioeng.* 114 (12), 2828–2836. doi:10.1002/bit.26434
- Finally, we still can learn a lot from how Genentech capitalized on decades of work on the *lac* operon to produce insulin in *E. coli*, but also from its contributions to the field of periplasmic protein production (e.g., (Goeddel et al., 1979; Chang et al., 1987; Joly et al., 1998; Champion et al., 2001; Champion et al., 2003; Chen et al., 2004; Zhou et al., 2016; McKenna et al., 2019)). The many contributions Genentech has made to this field are based on systematically designed studies that show the benefits of standardization and the importance of considering scalability early on.

## AUTHOR CONTRIBUTIONS

AK and J-WG wrote the paper.

## FUNDING

This work was supported by grants from the Swedish Research Council (2019-04143), the Carl Tryggers Foundation (CTS19:78) and the NOVO NORDISK FOUNDATION (NNF19OC0057673) to J-WG, a Marie Curie Initial Training Network grant (Horizon 2020; ProteinFactory, 642863) to Stockholm University (AK). Funders have supported research in the lab.

- Ceroni, F., Algar, R., Stan, G.-B., and Ellis, T. (2015). Quantifying Cellular Capacity Identifies Gene Expression Designs with Reduced burden. *Nat. Methods* 12 (5), 415–418. doi:10.1038/Nmeth.3339
- Cetnar, D. P., and Salis, H. M. (2021). Systematic Quantification of Sequence and Structural Determinants Controlling mRNA Stability in Bacterial Operons. *ACS Synth. Biol.* 10 (2), 318–332. doi:10.1021/acssynbio.0c00471
- Chamberlin, M., McGrath, J., and Waskell, L. (1970). New RNA Polymerase from *Escherichia coli* Infected with Bacteriophage T7. *Nature* 228 (5268), 227–231. doi:10.1038/228227a0
- Champion, K. M., Nishihara, J. C., Aldor, I. S., Moreno, G. T., Andersen, D., Stults, K. L., et al. (2003). Comparison of the *Escherichia coli* Proteomes for Recombinant Human Growth Hormone Producing and Nonproducing Fermentations. *Proteomics* 3 (7), 1365–1373. doi:10.1002/pmic.200300430
- Champion, K. M., Nishihara, J. C., Joly, J. C., and Arnott, D. (2001). Similarity of the *Escherichia coli* Proteome upon Completion of Different Biopharmaceutical Fermentation Processes. *Proteomics* 1 (9), 1133–1148. doi:10.1002/1615-9861(200109)1:9<1133:Aid-Prot1133>3.0.Co;2-S
- Chang, C. N., Rey, M., Bochner, B., Heyneker, H., and Gray, G. (1987). High-level Secretion of Human Growth Hormone by *Escherichia coli*. *Gene* 55 (2-3), 189–196. doi:10.1016/0378-1119(87)90279-4
- Chatzi, K. E., Sardis, M. F., Tsigirigotaki, A., Koukaki, M., Šoštarić, N., Konijnenberg, A., et al. (2017). Preprotein Mature Domains Contain Translocase Targeting Signals that Are Essential for Secretion. *J. Cel Biol.* 216 (5), 1357–1369. doi:10.1083/jcb.201609022
- Chen, C., Snedecor, B., Nishihara, J. C., Joly, J. C., McFarland, N., Andersen, D. C., et al. (2004). High-level Accumulation of a Recombinant Antibody Fragment in the Periplasm of *Escherichia coli* Requires a Triple-Mutant (degP Prc Spr) Host Strain. *Biotechnol. Bioeng.* 85 (5), 463–474. doi:10.1002/bit.20014
- Chen, H., Kim, J., and Kendall, D. A. (1996). Competition between Functional Signal Peptides Demonstrates Variation in Affinity for the Secretion Pathway. *J. Bacteriol.* 178 (23), 6658–6664. doi:10.1128/jb.178.23.6658-6664.1996
- Cheng, Y. S., Zipser, D., Cheng, C. Y., and Rolseth, S. J. (1979). Isolation and Characterization of Mutations in the Structural Gene for Protease III (Ptr). *J. Bacteriol.* 140 (1), 125–130. doi:10.1128/jb.140.1.125-130.1979

- Cohen, S. N., Chang, A. C. Y., Boyer, H. W., and Helling, R. B. (1973). Construction of Biologically Functional Bacterial Plasmids *In Vitro*. *Proc. Natl. Acad. Sci.* 70 (11), 3240–3244. doi:10.1073/pnas.70.11.3240
- Conlin, C. A., Trun, N. J., Silhavy, T. J., and Miller, C. G. (1992). *Escherichia coli* *prlC* Encodes an Endopeptidase and Is Homologous to the *Salmonella typhimurium* *opdA* Gene. *J. Bacteriol.* 174 (18), 5881–5887. doi:10.1128/jb.174.18.5881-5887.1992
- Crea, R., Kraszewski, A., Hirose, T., and Itakura, K. (1978). Chemical Synthesis of Genes for Human Insulin. *Proc. Natl. Acad. Sci.* 75 (12), 5765–5769. doi:10.1073/pnas.75.12.5765
- Dalbey, R. E., Wang, P., and van Dijk, J. M. (2012). Membrane Proteases in the Bacterial Protein Secretion and Quality Control Pathway. *Microbiol. Mol. Biol. Rev.* 76 (2), 311–330. doi:10.1128/Mmbr.05019-11
- De Geyter, J., Portaliou, A. G., Srinivasu, B., Krishnamurthy, S., Economou, A., and Karamanou, S. (2020). Trigger Factor Is a Bona Fide Secretory Pathway Chaperone that Interacts with SecB and the Translocase. *EMBO Rep.* 21 (6), e49054. doi:10.15252/embr.201949054
- De Geyter, J., Tsirigotaki, A., Orfanoudaki, G., Zorzini, V., Economou, A., and Karamanou, S. (2016). Protein Folding in the Cell Envelope of *Escherichia coli*. *Nat. Microbiol.* 1 (8), 16107. doi:10.1038/nmicrobiol.2016.107
- DeLisa, M. P., Lee, P., Palmer, T., and Georgiou, G. (2004). Phage Shock Protein PspA of *Escherichia coli* Relieves Saturation of Protein export via the Tat Pathway. *J. Bacteriol.* 186 (2), 366–373. doi:10.1128/jb.186.2.366-373.2004
- DeLisa, M. P., Tullman, D., and Georgiou, G. (2003). Folding Quality Control in the export of Proteins by the Bacterial Twin-Arginine Translocation Pathway. *Proc. Natl. Acad. Sci.* 100 (10), 6115–6120. doi:10.1073/pnas.0937838100
- Ding, L., Becker, A. B., Suzuki, A., and Roth, R. A. (1992). Comparison of the Enzymatic and Biochemical Properties of Human Insulin-Degrading Enzyme and *Escherichia coli* Protease III. *J. Biol. Chem.* 267 (4), 2414–2420. doi:10.1016/s0021-9258(18)45895-4
- Eichmann, J., Oberpaul, M., Weidner, T., Gerlach, D., and Czermak, P. (2019). Selection of High Producers from Combinatorial Libraries for the Production of Recombinant Proteins in *Escherichia coli* and *Vibrio Natriegens*. *Front. Bioeng. Biotechnol.* 7, 254. doi:10.3389/fbioe.2019.00254
- Ellis, M., and Humphreys, D. P. (2014). *Bacterial Host Strain Expressing Recombinant DsbC*. US 2014/0141468 A1.
- Ellis, M., Patel, P., Edon, M., Ramage, W., Dickinson, R., and Humphreys, D. P. (2017). Development of a High yielding *E. coli* periplasmic Expression System for the Production of Humanized Fab' Fragments. *Biotechnol. Prog.* 33 (1), 212–220. doi:10.1002/btpr.2393
- Emr, S. D., Hanley-Way, S., and Silhavy, T. J. (1981). Suppressor Mutations that Restore Export of a Protein with a Defective Signal Sequence. *Cell* 23 (1), 79–88. doi:10.1016/0092-8674(81)90272-5
- Endo, S., Yamamoto, Y., Sugawara, T., Nishimura, O., and Fujino, M. (2001). The Additional Methionine Residue at the N-Terminus of Bacterially Expressed Human Interleukin-2 Affects the Interaction between the N- and C-Termini. *Biochemistry* 40 (4), 914–919. doi:10.1021/bi001170r
- Frain, K. M., Robinson, C., and van Dijk, J. M. (2019). Transport of Folded Proteins by the Tat System. *Protein J.* 38 (4), 377–388. doi:10.1007/s10930-019-09859-y
- Fraser, T. H., and Bruce, B. J. (1978). Chicken Ovalbumin Is Synthesized and Secreted by *Escherichia coli*. *Proc. Natl. Acad. Sci.* 75 (12), 5936–5940. doi:10.1073/pnas.75.12.5936
- Frottin, F., Martinez, A., Peynot, P., Mitra, S., Holz, R. C., Giglione, C., et al. (2006). The Proteomics of N-Terminal Methionine Cleavage. *Mol. Cell Proteomics* 5 (12), 2336–2349. doi:10.1074/mcp.M600225-MCP200
- Gaciarz, A., Veijola, J., Uchida, Y., Saaranen, M. J., Wang, C., Hörkö, S., et al. (2016). Systematic Screening of Soluble Expression of Antibody Fragments in the Cytoplasm of *E. coli*. *Microb. Cel Fact* 15, 22. doi:10.1186/s12934-016-0419-5
- Gao, W., Yin, J., Bao, L., Wang, Q., Hou, S., Yue, Y., et al. (2018). Engineering Extracellular Expression Systems in *Escherichia coli* Based on Transcriptome Analysis and Cell Growth State. *ACS Synth. Biol.* 7 (5), 1291–1302. doi:10.1021/acssynbio.7b00400
- Gawin, A., Ertesvåg, H., Hansen, S. A. H., Malmö, J., and Brautaset, T. (2020). Translational Regulation of Periplasmic Folding Assistants and Proteases as a Valuable Strategy to Improve Production of Translocated Recombinant Proteins in *Escherichia coli*. *BMC Biotechnol.* 20 (1), 24. doi:10.1186/s12896-020-00615-0
- Giacalone, M. J., Gentile, A. M., Lovitt, B. T., Berkley, N. L., Gunderson, C. W., and Surber, M. W. (2006). Toxic Protein Expression in *Escherichia coli* Using a Rhamnose-Based Tightly Regulated and Tunable Promoter System. *Biotechniques* 40 (3), 355–364. doi:10.2144/000112112
- Goeddel, D. V., Kleid, D. G., Bolivar, F., Heyneker, H. L., Yansura, D. G., Crea, R., et al. (1979). Expression in *Escherichia coli* of Chemically Synthesized Genes for Human Insulin. *Proc. Natl. Acad. Sci.* 76 (1), 106–110. doi:10.1073/pnas.76.1.106
- Goodman, D. B., Church, G. M., and Kosuri, S. (2013). Causes and Effects of N-Terminal Codon Bias in Bacterial Genes. *Science* 342 (6157), 475–479. doi:10.1126/science.1241934
- Gottesman, S. (1996). Proteases and Their Targets in *Escherichia coli*. *Annu. Rev. Genet.* 30, 465–506. doi:10.1146/annurev.genet.30.1.465
- Grodberg, J., and Dunn, J. J. (1988). *ompT* Encodes the *Escherichia coli* Outer Membrane Protease that Cleaves T7 RNA Polymerase during Purification. *J. Bacteriol.* 170 (3), 1245–1253. doi:10.1128/jb.170.3.1245-1253.1988
- Guerrero Montero, I., Richards, K. L., Jawara, C., Browning, D. F., Peswani, A. R., Labrit, M., et al. (2019). *Escherichia coli* "TatExpress" Strains export Several G/L Human Growth Hormone to the Perioplasm by the Tat Pathway. *Biotechnol. Bioeng.* 116 (12), 3282–3291. doi:10.1002/bit.27147
- Hegde, R. S., and Bernstein, H. D. (2006). The Surprising Complexity of Signal Sequences. *Trends Biochem. Sci.* 31 (10), 563–571. doi:10.1016/j.tibs.2006.08.004
- Hjelm, A., Karyolaimos, A., Zhang, Z., Rujas, E., Vikström, D., Slotboom, D. J., et al. (2017). Tailoring *Escherichia coli* for the L-Rhamnose P<sub>BAD</sub> Promoter-Based Production of Membrane and Secretory Proteins. *ACS Synth. Biol.* 6 (6), 985–994. doi:10.1021/acssynbio.6b00321
- Ho, H. I., Fang, J. R., Cheung, J., and Wang, H. H. (2020). Programmable CRISPR-Cas Transcriptional Activation in Bacteria. *Mol. Syst. Biol.* 16 (7), e9427. doi:10.15252/msb.20199427
- Horga, L. G., Halliwell, S., Castiñeiras, T. S., Wyre, C., Matos, C. F. R. O., Yovcheva, D. S., et al. (2018). Tuning Recombinant Protein Expression to Match Secretion Capacity. *Microb. Cel Fact* 17, 199. doi:10.1186/s12934-018-1047-z
- Hsiung, H. M., Cantrell, A., Luirink, J., Oudega, B., Veros, A. J., and Becker, G. W. (1989). Use of Bacteriocin Release Protein in *E. coli* for Excretion of Human Growth Hormone into the Culture Medium. *Nat. Biotechnol.* 7 (3), 267–271. doi:10.1038/nbt0389-267
- Hu, X., O'Hara, L., White, S., Magner, E., Kane, M., and Gerard Wall, J. (2007). Optimisation of Production of a Domoic Acid-Binding scFv Antibody Fragment in *Escherichia coli* Using Molecular Chaperones and Functional Immobilisation on a Mesoporous Silicate Support. *Protein Expr. Purif.* 52 (1), 194–201. doi:10.1016/j.pep.2006.08.009
- Hughes, R. A., and Ellington, A. D. (2017). Synthetic DNA Synthesis and Assembly: Putting the Synthetic in Synthetic Biology. *Cold Spring Harb Perspect. Biol.* 9 (1), a023812. doi:10.1101/cshperspect.a023812
- Hughes, S. S. (2011). *Genentech: The Beginnings of Biotech*. Chicago Ill, London, UK: University of Chicago Press.
- Humphreys, D. P., Sehdev, M., Chapman, A. P., Ganesh, R., Smith, B. J., King, L. M., et al. (2000). High-Level Periplasmic Expression in *Escherichia coli* Using a Eukaryotic Signal Peptide: Importance of Codon Usage at the 5' End of the Coding Sequence. *Protein Expr. Purif.* 20 (2), 252–264. doi:10.1006/prep.2000.1286
- Itakura, K., Hirose, T., Crea, R., Riggs, A. D., Heyneker, H. L., Bolivar, F., et al. (1977). Expression in *Escherichia coli* of a Chemically Synthesized Gene for the Hormone Somatostatin. *Science* 198 (4321), 1056–1063. doi:10.1126/science.412251
- Ito, K., and Akiyama, Y. (2005). Cellular Functions, Mechanism of Action, and Regulation of FtsH Protease. *Annu. Rev. Microbiol.* 59, 211–231. doi:10.1146/annurev.micro.59.030804.121316
- Ito, K., Mori, H., and Chiba, S. (2018). Monitoring Substrate Enables Real-Time Regulation of a Protein Localization Pathway. *FEMS Microbiol. Lett.* 365 (11), doi:10.1093/femsle/fny109
- Izard, J. W., and Kendall, D. A. (1994). Signal Peptides: Exquisitely Designed Transport Promoters. *Mol. Microbiol.* 13 (5), 765–773. doi:10.1111/j.1365-2958.1994.tb00469.x

- Jackson, D. A., Symons, R. H., and Berg, P. (1972). Biochemical Method for Inserting New Genetic Information into DNA of Simian Virus 40: Circular SV40 DNA Molecules Containing Lambda Phage Genes and the Galactose Operon of *Escherichia coli*. *Proc. Natl. Acad. Sci.* 69 (10), 2904–2909. doi:10.1073/pnas.69.10.2904
- Jeong, H., Barbe, V., Lee, C. H., Vallenet, D., Yu, D. S., Choi, S.-H., et al. (2009). Genome Sequences of *Escherichia coli* B Strains REL606 and BL21(DE3). *J. Mol. Biol.* 394 (4), 644–652. doi:10.1016/j.jmb.2009.09.052
- Jeong, H., Kim, H. J., and Lee, S. J. (2015). Complete Genome Sequence of *Escherichia coli* Strain BL21. *Genome Announc* 3 (2), e00134. doi:10.1128/genomeA.00134-15
- Jiang, C., Wynne, M., and Huber, D. (2021). How Quality Control Systems AID Sec-dependent Protein Translocation. *Front. Mol. Biosci.* 8, 669376. doi:10.3389/fmolb.2021.669376
- Joly, J. C., Leung, W. S., and Swartz, J. R. (1998). Overexpression of *Escherichia coli* Oxidoreductases Increases Recombinant Insulin-like Growth Factor-I Accumulation. *Proc. Natl. Acad. Sci.* 95 (6), 2773–2777. doi:10.1073/pnas.95.6.2773
- Kadokura, H., and Beckwith, J. (2009). Detecting Folding Intermediates of a Protein as it Passes through the Bacterial Translocation Channel. *Cell* 138 (6), 1164–1173. doi:10.1016/j.cell.2009.07.030
- Karyolaimos, A., Ampah-Korsah, H., Hillenaar, T., Mestre Borrás, A., Dolata, K. M., Sievers, S., et al. (2019). Enhancing Recombinant Protein Yields in the *E. coli* Periplasm by Combining Signal Peptide and Production Rate Screening. *Front. Microbiol.* 10, 1511. doi:10.3389/fmicb.2019.01511
- Karyolaimos, A., Dolata, K. M., Antelo-Varela, M., Mestre Borrás, A., Elfageih, R., Sievers, S., et al. (2020). *Escherichia coli* Can Adapt its Protein Translocation Machinery for Enhanced Periplasmic Recombinant Protein Production. *Front. Bioeng. Biotechnol.* 7, 465. doi:10.3389/fbioe.2019.00465
- Kastenhofer, J., Rajamanickam, V., Libiseller-Egger, J., and Spadiut, O. (2021). Monitoring and Control of *E. coli* Cell Integrity. *J. Biotechnol.* 329, 1–12. doi:10.1016/j.jbiotec.2021.01.009
- Keiler, K. C., and Sauer, R. T. (1996). Sequence Determinants of C-Terminal Substrate Recognition by the Tsp Protease. *J. Biol. Chem.* 271 (5), 2589–2593. doi:10.1074/jbc.271.5.2589
- Kendall, D. A., Bock, S. C., and Kaiser, E. T. (1986). Idealization of the Hydrophobic Segment of the Alkaline Phosphatase Signal Peptide. *Nature* 321 (6071), 706–708. doi:10.1038/321706a0
- Kihara, A., Akiyama, Y., and Ito, K. (1995). FtsH Is Required for Proteolytic Elimination of Uncomplexed Forms of SecY, an Essential Protein Translocase Subunit. *Proc. Natl. Acad. Sci.* 92 (10), 4532–4536. doi:10.1073/pnas.92.10.4532
- Kim, J., Rusch, S., Luirink, J., and Kendall, D. A. (2001). Is Ffh Required for export of Secretory Proteins? *FEBS Lett.* 505 (2), 245–248. doi:10.1016/S0014-5793(01)02784-3
- Kramer, R. A., Vandeputte-Rutten, L., de Roon, G. J., Gros, P., Dekker, N., and Egmond, M. R. (2001). Identification of Essential Acidic Residues of Outer Membrane Protease OmpT Supports a Novel Active Site. *FEBS Lett.* 505 (3), 426–430. doi:10.1016/s0014-5793(01)02863-0
- Krause, M., Neubauer, A., and Neubauer, P. (2016). The Fed-Batch Principle for the Molecular Biology Lab: Controlled Nutrient Diets in Ready-Made media Improve Production of Recombinant Proteins in *Escherichia coli*. *Microb. Cel Fact* 15, 110. doi:10.1186/s12934-016-0513-8
- Kulmala, A., Huovinen, T., and Lamminmäki, U. (2019). Improvement of Fab Expression by Screening Combinatorial Synonymous Signal Sequence Libraries. *Microb. Cel Fact* 18 (1), 157. doi:10.1186/s12934-019-1210-1
- Laskowska, E., Kuczyńska-Wisńnik, D., Skórko-Glonek, J., and Taylor, A. (1996). Degradation by Proteases Lon, Clp and HtrA, of *Escherichia coli* Proteins Aggregated *In Vivo* by Heat Shock; HtrA Protease Action *In Vivo* and *In Vitro*. *Mol. Microbiol.* 22 (3), 555–571. doi:10.1046/j.1365-2958.1996.1231493.x
- Lee, Y. J., Kim, H. S., Ryu, A. J., and Jeong, K. J. (2013). Enhanced Production of Full-Length Immunoglobulin G via the Signal Recognition Particle (SRP)-dependent Pathway in *Escherichia coli*. *J. Biotechnol.* 165 (2), 102–108. doi:10.1016/j.jbiotec.2013.03.007
- Liao, Y.-D., Jeng, J.-C., Wang, C.-F., Wang, S.-C., and Chang, S.-T. (2004). Removal of N-Terminal Methionine from Recombinant Proteins by engineered *E. coli* methionine Aminopeptidase. *Protein Sci.* 13 (7), 1802–1810. doi:10.1110/ps.04679104
- Liao, Y.-D., Wang, S. C., Leu, Y. J., Wang, C. F., Chang, S. T., Hong, Y. T., et al. (2003). The Structural Integrity Exerted by N-Terminal Pyroglutamate Is Crucial for the Cytotoxicity of Frog Ribonuclease from *Rana pipiens*. *Nucleic Acids Res.* 31 (18), 5247–5255. doi:10.1093/nar/gkg746
- Lycklama a Nijeholt, J. A., de Keyzer, J., Prabudiansyah, I., and Driessen, A. J. M. (2013). Characterization of the Supporting Role of SecE in Protein Translocation. *FEBS Lett.* 587 (18), 3083–3088. doi:10.1016/j.febslet.2013.07.046
- Makino, T., Skretas, G., and Georgiou, G. (2011a). Strain Engineering for Improved Expression of Recombinant Proteins in Bacteria. *Microb. Cel Fact* 10, 32. doi:10.1186/1475-2859-10-32
- Makino, T., Skretas, G., Kang, T.-H., and Georgiou, G. (2011b). Comprehensive Engineering of *Escherichia coli* for Enhanced Expression of IgG Antibodies. *Metab. Eng.* 13 (2), 241–251. doi:10.1016/j.jymben.2010.11.002
- Manta, B., Boyd, D., and Berkmen, M. (2019). Disulfide Bond Formation in the Periplasm of *Escherichia coli*. *EcoSal Plus* 8 (2), 1. doi:10.1128/ecosalplus.ESP-0012-2018
- Mateus, A., Hevler, J., Bobonis, J., Kurzawa, N., Shah, M., Mitosch, K., et al. (2020). The Functional Proteome Landscape of *Escherichia coli*. *Nature* 588 (7838), 473–478. doi:10.1038/s41586-020-3002-5
- Matos, C. F. R. O., Robinson, C., Alanen, H. I., Prus, P., Uchida, Y., Ruddock, L. W., et al. (2014). Efficient export of Prefolded, Disulfide-Bonded Recombinant Proteins to the Periplasm by the Tat Pathway in *Escherichia coli* CyDisCo Strains. *Biotechnol. Prog.* 30 (2), 281–290. doi:10.1002/btpr.1858
- Mavrougelos, C., Thiel, M., Adamson, P. J., Millard, D. J., Nobbs, S., Zola, H., et al. (2001). Increased Yield and Activity of Soluble Single-Chain Antibody Fragments by Combining High-Level Expression and the Skp Periplasmic Chaperonin. *Protein Expr. Purif.* 23 (2), 289–295. doi:10.1006/prep.2001.1506
- McKenna, R., Lombana, T. N., Yamada, M., Mukhyala, K., and Veeravalli, K. (2019). Engineered Sigma Factors Increase Full-Length Antibody Expression in *Escherichia coli*. *Metab. Eng.* 52, 315–323. doi:10.1016/j.jymben.2018.12.009
- Meerman, H. J., and Georgiou, G. (1994). Construction and Characterization of a Set of *E. coli* Strains Deficient in All Known Loci Affecting the Proteolytic Stability of Secreted Recombinant Proteins. *Nat. Biotechnol.* 12 (11), 1107–1110. doi:10.1038/nbt1194-1107
- Meltzer, M., Hasenbein, S., Mamant, N., Merdanovic, M., Poepsel, S., Hauske, P., et al. (2009). Structure, Function and Regulation of the Conserved Serine Proteases DegP and DegS of *Escherichia coli*. *Res. Microbiol.* 160 (9), 660–666. doi:10.1016/j.resmic.2009.07.012
- Merdanovic, M., Clausen, T., Kaiser, M., Huber, R., and Ehrmann, M. (2011). Protein Quality Control in the Bacterial Periplasm. *Annu. Rev. Microbiol.* 65, 149–168. doi:10.1146/annurev-micro-090110-102925
- Mergulhão, F. J. M., Summers, D. K., and Monteiro, G. A. (2005). Recombinant Protein Secretion in *Escherichia coli*. *Biotechnol. Adv.* 23 (3), 177–202. doi:10.1016/j.biotechadv.2004.11.003
- Meyer, A. J., Segall-Shapiro, T. H., Glassey, E., Zhang, J., and Voigt, C. A. (2019). *Escherichia coli* "Marionette" Strains with 12 Highly Optimized Small-Molecule Sensors. *Nat. Chem. Biol.* 15 (2), 196–204. doi:10.1038/s41589-018-0168-3
- Miroux, B., and Walker, J. E. (1996). Over-production of Proteins in *Escherichia coli*: Mutant Hosts that Allow Synthesis of Some Membrane Proteins and Globular Proteins at High Levels. *J. Mol. Biol.* 260, 289–298. doi:10.1006/jmbi.1996.0399
- Mitchell, A. M., and Silhavy, T. J. (2019). Envelope Stress Responses: Balancing Damage Repair and Toxicity. *Nat. Rev. Microbiol.* 17 (7), 417–428. doi:10.1038/s41579-019-0199-0
- Moffatt, B. A., and Studier, F. W. (1987). T7 Lysozyme Inhibits Transcription by T7 RNA Polymerase. *Cell* 49 (2), 221–227. doi:10.1016/0092-8674(87)90563-0
- Morra, R., Shankar, J., Robinson, C. J., Halliwell, S., Butler, L., Upton, M., et al. (2016). Dual Transcriptional-Translational cascade Permits Cellular Level Tuneable Expression Control. *Nucleic Acids Res.* 44 (3), e21. doi:10.1093/nar/gkv912
- Natarajan, A., Haitjema, C. H., Lee, R., Boock, J. T., and DeLisa, M. P. (2017). An Engineered Survival-Selection Assay for Extracellular Protein Expression Uncovers Hypersecretory Phenotypes in *Escherichia coli*. *ACS Synth. Biol.* 6 (5), 875–883. doi:10.1021/acssynbio.6b00366
- Novak, P., and Dev, I. K. (1988). Degradation of a Signal Peptide by Protease IV and Oligopeptidase A. *J. Bacteriol.* 170 (11), 5067–5075. doi:10.1128/jb.170.11.5067-5075.1988

- Orfanoudaki, G., Markaki, M., Chatzi, K., Tsamardinos, I., and Economou, A. (2017). MatureP: Prediction of Secreted Proteins with Exclusive Information from Their Mature Regions. *Sci. Rep.* 7 (1), 3263. doi:10.1038/s41598-017-03557-4
- Oswald, J., Njenga, R., Natriashvili, A., Sarmah, P., and Koch, H.-G. (2021). The Dynamic SecYEG Translocon. *Front. Mol. Biosci.* 8, 664241. doi:10.3389/fmolb.2021.664241
- Ow, D., Lim, D., Morin Nissom, P., Camattari, A., and Wong, V. (2010). Co-expression of Skp and FkpA Chaperones Improves Cell Viability and Alters the Global Expression of Stress Response Genes during scFvD1.3 Production. *Microb. Cel Fact* 9, 22. doi:10.1186/1475-2859-9-22
- Paetzel, M. (2014). Structure and Mechanism of *Escherichia coli* Type I Signal Peptidase. *Biochim. Biophys. Acta (Bba) - Mol. Cel Res.* 1843 (8), 1497–1508. doi:10.1016/j.bbamcr.2013.12.003
- Palmer, T., and Stansfeld, P. J. (2020). Targeting of Proteins to the Twin-arginine Translocation Pathway. *Mol. Microbiol.* 113 (5), 861–871. doi:10.1111/mmi.14461
- Pérez-Pérez, J., Márquez, G., Barbero, J.-L., and Gutiérrez, J. (1994). Increasing the Efficiency of Protein Export in *Escherichia coli*. *Nat. Biotechnol.* 12 (2), 178–180. doi:10.1038/nbt0294-178
- Puertas, J.-M., Nannenga, B. L., Dornfeld, K. T., Betton, J.-M., and Baneyx, F. (2010). Enhancing the Secretory Yields of Leech Carboxypeptidase Inhibitor in *Escherichia coli*: Influence of Trigger Factor and Signal Recognition Particle. *Protein Expr. Purif.* 74 (1), 122–128. doi:10.1016/j.pep.2010.06.008
- Reis, A. C., Halper, S. M., Vezeau, G. E., Cetnar, D. P., Hossain, A., Clauer, P. R., et al. (2019). Simultaneous Repression of Multiple Bacterial Genes Using Nonrepetitive Extra-long sgRNA Arrays. *Nat. Biotechnol.* 37 (11), 1294–1301. doi:10.1038/s41587-019-0286-9
- Reis, A. C., and Salis, H. M. (2020). An Automated Model Test System for Systematic Development and Improvement of Gene Expression Models. *ACS Synth. Biol.* 9 (11), 3145–3156. doi:10.1021/acssynbio.0c00394
- Rennig, M., Martinez, V., Mirzadeh, K., Dunas, F., Röjsäter, B., Daley, D. O., et al. (2018). TARSyn: Tunable Antibiotic Resistance Devices Enabling Bacterial Synthetic Evolution and Protein Production. *ACS Synth. Biol.* 7 (2), 432–442. doi:10.1021/acssynbio.7b00200
- Robertson, W. E., Funke, L. F. H., de la Torre, D., Fredens, J., Wang, K., and Chin, J. W. (2021). Creating Custom Synthetic Genomes in *Escherichia coli* with REXER and GENESIS. *Nat. Protoc.* 16 (5), 2345–2380. doi:10.1038/s41596-020-00464-3
- Rugbjerg, P., Dyerberg, A. S. B., Quainoo, S., Munck, C., and Sommer, M. O. A. (2021). Short and Long-Read Ultra-deep Sequencing Profiles Emerging Heterogeneity across Five Platform *Escherichia coli* Strains. *Metab. Eng.* 65, 197–206. doi:10.1016/j.mbsen.2020.11.006
- Rusch, S. L., and Kendall, D. A. (1994). Transport of an Export-Defective Protein by a Highly Hydrophobic Signal Peptide. *J. Biol. Chem.* 269 (2), 1243–1248. doi:10.1016/s0021-9258(17)42249-6
- Rusch, S., Mascolo, C., Kebir, M., and Kendall, D. (2002). Juxtaposition of Signal-Peptide Charge and Core Region Hydrophobicity Is Critical for Functional Signal Peptides. *Arch. Microbiol.* 178 (4), 306–310. doi:10.1007/s00203-002-0453-z
- Sala, A., Bordes, P., and Genevaux, P. (2014). Multitasking SecB Chaperones in Bacteria. *Front. Microbiol.* 5, 666. doi:10.3389/fmicb.2014.00666
- Salis, H. M., Mirsky, E. A., and Voigt, C. A. (2009). Automated Design of Synthetic Ribosome Binding Sites to Control Protein Expression. *Nat. Biotechnol.* 27 (10), 946–950. doi:10.1038/nbt.1568
- Sargent, F., Bogsch, E. G., Stanley, N. R., Wexler, M., Robinson, C., Berks, B. C., et al. (1998). Overlapping Functions of Components of a Bacterial Sec-independent Protein export Pathway. *EMBO J.* 17 (13), 3640–3650. doi:10.1093/emboj/17.13.3640
- Schierle, C. F., Berkmen, M., Huber, D., Kumamoto, C., Boyd, D., and Beckwith, J. (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), 5706–5713. doi:10.1128/jb.185.19.5706-5713.2003
- Schlapsch, M., Grimm, S., and Skerra, A. (2006). A System for Concomitant Overexpression of Four Periplasmic Folding Catalysts to Improve Secretory Protein Production in *Escherichia coli*. *Protein Eng. Des. Selection* 19 (8), 385–390. doi:10.1093/protein/gz1018
- Schlegel, S., Löfblom, J., Lee, C., Hjelm, A., Klepsch, M., Strous, M., et al. (2012). Optimizing Membrane Protein Overexpression in the *Escherichia coli* Strain Lemo21(DE3). *J. Mol. Biol.* 423 (4), 648–659. doi:10.1016/j.jmb.2012.07.019
- Schlegel, S., Rujas, E., Ytterberg, A. J., Zubarev, R. A., Luirink, J., and de Gier, J.-W. (2013). Optimizing Heterologous Protein Production in the Periplasm of *E. coli* by Regulating Gene Expression Levels. *Microb. Cel Fact* 12, 24. doi:10.1186/1475-2859-12-24
- Schmidt, A., Kochanowski, K., Vedelaar, S., Ahrné, E., Volkmer, B., Callipo, L., et al. (2016). The Quantitative and Condition-dependent *Escherichia coli* Proteome. *Nat. Biotechnol.* 34 (1), 104–110. doi:10.1038/nbt.3418
- Selas Castiñeiras, T., Williams, S. G., Hitchcock, A., Cole, J. A., Smith, D. C., and Overton, T. W. (2018). Development of a Generic  $\beta$ -lactamase Screening System for Improved Signal Peptides for Periplasmic Targeting of Recombinant Proteins in *Escherichia coli*. *Sci. Rep.* 8, 6986. doi:10.1038/s41598-018-25192-3
- Shanmugaw, S. K., and Dalbey, R. E. (2019). The Conserved Role of YidC in Membrane Protein Biogenesis. *Microbiol. Spectr.* 7 (1), 1. doi:10.1128/microbiolspec.PSIB-0014-2018
- Sijbrandi, R., Urbanus, M. L., ten Hagen-Jongman, C. M., Bernstein, H. D., Oudega, B., Otto, B. R., et al. (2003). Signal Recognition Particle (SRP)-mediated Targeting and Sec-dependent Translocation of an Extracellular *Escherichia coli* Protein. *J. Biol. Chem.* 278 (7), 4654–4659. doi:10.1074/jbc.M211630200
- Silverstone, A. E., Arditti, R. R., and Magasanik, B. (1970). Catalytic-insensitive Revertants of *lac* Promoter Mutants. *Proc. Natl. Acad. Sci.* 66 (3), 773–779. doi:10.1073/pnas.66.3.773
- Simmons, L. C., and Yansura, D. G. (1996). Translational Level Is a Critical Factor for the Secretion of Heterologous Proteins in *Escherichia coli*. *Nat. Biotechnol.* 14 (5), 629–634. doi:10.1038/nbt0596-629
- Singh, S. K., Parveen, S., SaiSree, L., and Reddy, M. (2015). Regulated Proteolysis of a Cross-link-specific Peptidoglycan Hydrolase Contributes to Bacterial Morphogenesis. *Proc. Natl. Acad. Sci. USA* 112 (35), 10956–10961. doi:10.1073/pnas.1507760112
- Skórko-Glonek, J., Zurawa, D., Kuczwar, E., Wozniak, M., Wypych, Z., and Lipinska, B. (1999). The *Escherichia coli* Heat Shock Protease HtrA Participates in Defense against Oxidative Stress. *Mol. Gen. Genet.* 262 (2), 342–350. doi:10.1007/s004380051092
- Smets, D., Loos, M. S., Karamanou, S., and Economou, A. (2019). Protein Transport across the Bacterial Plasma Membrane by the Sec Pathway. *Protein J.* 38 (3), 262–273. doi:10.1007/s10930-019-09841-8
- Soltes, G. R., Martin, N. R., Park, E., Sutterlin, H. A., and Silhavy, T. J. (2017). Distinctive Roles for Periplasmic Proteases in the Maintenance of Essential Outer Membrane Protein Assembly. *J. Bacteriol.* 199 (20), e00418. doi:10.1128/JB.00418-17
- Sonoda, H., Kumada, Y., Katsuda, T., and Yamaji, H. (2011). Effects of Cytoplasmic and Periplasmic Chaperones on Secretory Production of Single-Chain Fv Antibody in *Escherichia coli*. *J. Biosci. Bioeng.* 111 (4), 465–470. doi:10.1016/j.jbiosc.2010.12.015
- Spieß, C., Beil, A., and Ehrmann, M. (1999). A Temperature-dependent Switch from Chaperone to Protease in a Widely Conserved Heat Shock Protein. *Cell* 97 (3), 339–347. doi:10.1016/S0092-8674(00)80743-6
- Steiner, D., Forrer, P., Stumpp, M. T., and Plückthun, A. (2006). Signal Sequences Directing Cotranslational Translocation Expand the Range of Proteins Amenable to Phage Display. *Nat. Biotechnol.* 24 (7), 823–831. doi:10.1038/nbt1218
- Strauch, K. L., and Beckwith, J. (1988). An *Escherichia coli* Mutation Preventing Degradation of Abnormal Periplasmic Proteins. *Proc. Natl. Acad. Sci.* 85 (5), 1576–1580. doi:10.1073/pnas.85.5.1576
- Strauch, K. L., Johnson, K., and Beckwith, J. (1989). Characterization of *degP*, a Gene Required for Proteolysis in the Cell Envelope and Essential for Growth of *Escherichia coli* at High Temperature. *J. Bacteriol.* 171 (5), 2689–2696. doi:10.1128/jb.171.5.2689-2696.1989
- Studier, F. W., Daegelen, P., Lenski, R. E., Maslov, S., and Kim, J. F. (2009). Understanding the Differences between Genome Sequences of *Escherichia coli* B Strains REL606 and BL21(DE3) and Comparison of the *E. coli* B and K-12 Genomes. *J. Mol. Biol.* 394 (4), 653–680. doi:10.1016/j.jmb.2009.09.021
- Studier, F. W., and Moffatt, B. A. (1986). Use of Bacteriophage T7 RNA Polymerase to Direct Selective High-Level Expression of Cloned Genes. *J. Mol. Biol.* 189 (1), 113–130. doi:10.1016/0022-2836(86)90385-2

- Stull, F., Betton, J.-M., and Bardwell, J. C. A. (2018). Periplasmic Chaperones and Prolyl Isomerases. *EcoSal Plus* 8 (1), 1. doi:10.1128/ecosalplus.ESP-0005-2018
- Swamy, K. H. S., and Goldberg, A. L. (1981). *E. coli* Contains Eight Soluble Proteolytic Activities, One Being ATP Dependent. *Nature* 292 (5824), 652–654. doi:10.1038/292652a0
- Talmadge, K., and Gilbert, W. (1982). Cellular Location Affects Protein Stability in *Escherichia coli*. *Proc. Natl. Acad. Sci.* 79 (6), 1830–1833. doi:10.1073/pnas.79.6.1830
- Talmadge, K., Kaufman, J., and Gilbert, W. (1980). Bacteria Mature Preproinsulin to Proinsulin. *Proc. Natl. Acad. Sci.* 77 (7), 3988–3992. doi:10.1073/pnas.77.7.3988
- Tan, J., Sastry, A. V., Fremming, K. S., Björn, S. P., Hoffmeyer, A., Seo, S., et al. (2020). Independent Component Analysis of *E. coli*'s Transcriptome Reveals the Cellular Processes that Respond to Heterologous Gene Expression. *Metab. Eng.* 61, 360–368. doi:10.1016/j.ymben.2020.07.002
- Taw, M. N., Li, M., Kim, D., Rocco, M. A., Waraho-Zhmayev, D., and DeLisa, M. P. (2021). Engineering a Supersecreting Strain of *Escherichia coli* by Directed Coevolution of the Multiprotein Tat Translocation Machinery. *ACS Synth. Biol.* 10 (11), 2947–2958. doi:10.1021/acssynbio.1c00183
- Tripathi, N. K. (2016). Production and Purification of Recombinant Proteins from *Escherichia coli*. *ChemBioEng Rev.* 3 (3), 116–133. doi:10.1002/cben.201600002
- Tripathi, N. K., and Shrivastava, A. (2019). Recent Developments in Bioprocessing of Recombinant Proteins: Expression Hosts and Process Development. *Front. Bioeng. Biotechnol.* 7, 420. doi:10.3389/fbioe.2019.00420
- Tsirigotaki, A., Chatzi, K. E., Koukaki, M., De Geyter, J., Portaliou, A. G., Orfanoudaki, G., et al. (2018). Long-Lived Folding Intermediates Predominate the Targeting-Competent Secretome. *Structure* 26 (5), 695–707. doi:10.1016/j.str.2018.03.006
- Tsirigotaki, A., De Geyter, J., Šoštarić, N., Economou, A., and Karamanou, S. (2017). Protein export through the Bacterial Sec Pathway. *Nat. Rev. Microbiol.* 15 (1), 21–36. doi:10.1038/nrmicro.2016.161
- van der Wolk, J. P., Fekkes, P., Boorsma, A., Huie, J. L., Silhavy, T. J., and Driessen, A. J. (1998). PrlA4 Prevents the Rejection of Signal Sequence Defective Preproteins by Stabilizing the SecA-SecY Interaction during the Initiation of Translocation. *EMBO J.* 17 (18), 3631–3639. doi:10.1093/emboj/17.13.3631
- van Dijk, J. M., de Jong, A., Smith, H., Bron, S., and Venema, G. (1991). Signal Peptidase I Overproduction Results in Increased Efficiencies of export and Maturation of Hybrid Secretory Proteins in *Escherichia coli*. *Mol. Gen. Genet.* 227 (1), 40–48. doi:10.1007/Bf00260704
- van Stelten, J., Silva, F., Belin, D., and Silhavy, T. J. (2009). Effects of Antibiotics and a Proto-Oncogene Homolog on Destruction of Protein Translocator SecY. *Science* 325 (5941), 753–756. doi:10.1126/science.1172221
- Wagner, S., Baars, L., Ytterberg, A. J., Klussmeier, A., Wagner, C. S., Nord, O., et al. (2007). Consequences of Membrane Protein Overexpression in *Escherichia coli*. *Mol. Cell Proteomics* 6 (9), 1527–1550. doi:10.1074/mcp.m600431-mcp200
- Wagner, S., Klepsch, M. M., Schlegel, S., Appel, A., Draheim, R., Tarry, M., et al. (2008). Tuning *Escherichia coli* for Membrane Protein Overexpression. *Proc. Natl. Acad. Sci.* 105 (38), 14371–14376. doi:10.1073/pnas.0804090105
- Wang, H. H., Isaacs, F. J., Carr, P. A., Sun, Z. Z., Xu, G., Forest, C. R., et al. (2009). Programming Cells by Multiplex Genome Engineering and Accelerated Evolution. *Nature* 460 (7257), 894–898. doi:10.1038/nature08187
- Wanner, B. L., Kodaira, R., and Neidhart, F. C. (1977). Physiological Regulation of a Decontrolled *lac* Operon. *J. Bacteriol.* 130 (1), 212–222. doi:10.1128/jb.130.1.212-222.1977
- Wannier, T. M., Ciaccia, P. N., Ellington, A. D., Filsinger, G. T., Isaacs, F. J., Javanmardi, K., et al. (2021). Recombineering and MAGE. *Nat. Rev. Methods Primers* 1 (1), 7. doi:10.1038/s43586-020-00006-x
- Weiner, J. H., Bilous, P. T., Shaw, G. M., Lubitz, S. P., Frost, L., Thomas, G. H., et al. (1998). A Novel and Ubiquitous System for Membrane Targeting and Secretion of Cofactor-Containing Proteins. *Cell* 93 (1), 93–101. doi:10.1016/S0092-8674(00)81149-6
- Zalucki, Y. M., Gittins, K. L., and Jennings, M. P. (2008). Secretory Signal Sequence Non-optimal Codons Are Required for Expression and export of  $\beta$ -lactamase. *Biochem. Biophysical Res. Commun.* 366 (1), 135–141. doi:10.1016/j.bbrc.2007.11.093
- Zalucki, Y. M., Jones, C. E., Ng, P. S. K., Schulz, B. L., and Jennings, M. P. (2010). Signal Sequence Non-optimal Codons Are Required for the Correct Folding of Mature Maltose Binding Protein. *Biochim. Biophys. Acta (Bba) - Biomembranes* 1798 (6), 1244–1249. doi:10.1016/j.bbamem.2010.03.010
- Zhang, G., Brox, S., and Weiner, J. H. (2006). Extracellular Accumulation of Recombinant Proteins Fused to the Carrier Protein YebF in *Escherichia coli*. *Nat. Biotechnol.* 24 (1), 100–104. doi:10.1038/nbt1174
- Zhou, Y., Liu, P., Gan, Y., Sandoval, W., Katakam, A. K., Reichelt, M., et al. (2016). Enhancing Full-Length Antibody Production by Signal Peptide Engineering. *Microb. Cel Fact* 15, 47. doi:10.1186/s12934-016-0445-3

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Karyolaimos and de Gier. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.