# Bacterial signal peptides: structure, optimization, and applications

Esra Erkut[1]*
[1] Department of Biological Sciences, University of Alberta
*Corresponding author: esra@ualberta.ca

## ABSTRACT

Bacterial signal peptides are N-terminal tags that direct proteins for export through one of various transport pathways. These signal peptides are highly important as they are the key determinants of transport, ensuring that the correct protein arrives at the correct pathway. While these peptides consist of three domains with well conserved biochemical properties, there remains a large amount of diversity between the signal sequences for different proteins, transport pathways, and bacterial species. Recent advancements have allowed for the prediction and manipulation of signal sequences to optimize protein export efficiency. This knowledge can then be exploited in the field of recombinant protein production wherein bacterial species are used to produce and secrete proteins of interest. By fusing the protein with an optimized signal peptide, the yield or rate of export can be improved. This review focuses on signal peptides for two primary transport pathways (Sec and Tat) in *Escherichia coli* specifically, with an emphasis on applications and the production of recombinant proteins.

KEY WORDS: **Signal sequence, signal peptide, recombinant proteins, Sec, Tat**

## 1 | INTRODUCTION

For bacterial proteins to be transported to different cellular compartments or secreted from the cell, they must navigate through various transport pathways (Green & Mecsas, 2016). Faithful targeting to these pathways is paramount as it ensures that proteins arrive in the right place at the right time. With more than a dozen secretion systems and a vast exportome portfolio, strict organization must be maintained for bacteria to remain viable and functional. This protein transport specificity is provided by short N-terminal tags dubbed "signal peptides" (Emr et al., 1978; Freudl, 2018). These signal peptides have a tripartite structure with conserved motifs and are typically 16-30 amino acids in length (Freudl, 2018; Peng et al., 2019). Distinct biochemical properties of each domain or motif allow the protein to be transported by one specific pathway while avoiding others. Once the signal peptide has directed the protein to the correct transport pathway, it is cleaved off by signal peptidases (Peng et al., 2019).

This review will outline common features of signal peptides for two main transport pathways—Sec and Tat—in *Escherichia coli* (*E. coli*). The general secretion (Sec) and twin-arginine translocation (Tat) pathways are the primary transport systems in bacteria, with Sec predicted to transport over 90% of all secreted proteins (Georgiou & Segatori, 2005; Green & Mecsas, 2016). In brief, both pathways are composed of multiple protein components and involve transport across the cytoplasmic (inner) membrane (Green & Mecsas, 2016). However, while Sec only transports proteins in an unfolded state, Tat has the unique ability to transport folded proteins.

In addition to discussing key components of signal peptides, this review will also explore practical applications. For example, signal peptides can be exploited in the bacterial production of recombinant proteins (Freudl, 2018). This is highly relevant to many industries such as biopharmaceuticals, food production, and scientific research. By expanding upon the knowledge surrounding signal peptides and optimizing their sequence for maximum efficiency, the yield and purity of recombinant proteins can be improved.
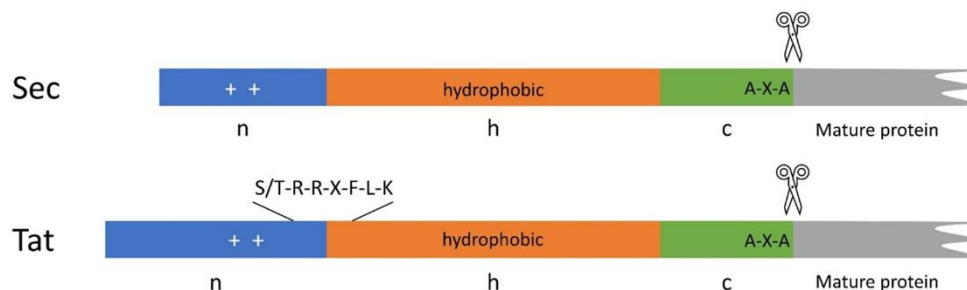
**Figure 1.** Schematic of the general tripartite structure of signal peptides. Both Sec- and Tat-specific signal peptides possess a positively charged n-region, a hydrophobic h-region, and a polar c-region. This polar c-region also contains the A-X-A motif, which is a signal peptidase cleavage site as indicated by the scissors. Tat-specific signal peptides also contain a twin-arginine motif (S/T-R-R-X-F-L-K) at the border between the n- and h-regions.

## 2 | SIGNAL SEQUENCE DIVERSITY — SEC VS. TAT

A specific signal sequence exhibits some variation, both within and between bacterial species. This variation allows for certain proteins to be targeted to different transport pathways for export, thus arriving at different final destinations (Freudl, 2018; Green & Mecsas, 2016). Despite this diversity, there are three typical domains: a positively charged amino-terminal region (n-region), a central hydrophobic region (h-region), and a polar carboxyl-terminal region (c-region) (Freudl, 2018; Peng et al., 2019) (Figure 1). It is within this c-region that the cleavage site for signal peptidases resides, with an A-X-A motif where A is alanine, and X is any amino acid (Rusch & Kendall, 2007) (Figure 1). Due to recurring motifs, bioinformatics tools may be used to predict putative signal sequences *in silico* (Peng et al., 2019). For example, SignalP (Petersen et al., 2011), Phobius (Käll et al., 2007), and PrediSi (Hiller et al., 2004) can be used to predict signal peptides for the Sec pathway, while TatP (Bendtsen et al., 2005), Tatfind (Rose et al., 2002), and PRED-TAT (Bagos et al., 2010) may be used for the Tat pathway. The following subsections will outline common motifs seen in signal peptide sequences, with a focus on *E. coli*. Additionally, differences in the signal peptide that discern Sec vs. Tat transport will be highlighted.

### 2.1 Sec-specific signal peptides

Signal sequences which target proteins to the Sec pathway in *E. coli* follow the general tripartite structure as outlined previously (Rusch & Kendall, 2007) (Figure 1). As evidenced in Table 1, the positively charged n-region typically contains multiple basic amino acid residues such as lysine (K) and arginine (R). Next, the longer h-region consists primarily of hydrophobic amino acid residues (Table 1). Finally, the short c-region contains some polar uncharged amino acid residues such as serine (S) and threonine (T), as well as the signal peptide cleavage motif (A-X-A) (Lüke et al., 2009; Palmer &

Berks, 2012). While the biochemical properties within each domain are well conserved, the individual amino acids themselves exhibit large variation between different proteins (Table 1). Because the amino acid substitutions are often within the same biochemical "family" (e.g., hydrophobic alanine vs. leucine), this diversity is well

**Table 1.** Multiple sequence alignment (MSA) of exemplar Sec signal peptide sequences in *E. coli*. Sample Sec substrates were selected (Crane & Randall, 2017; Molloy et al., 2000) and signal peptide sequences were collected from the UniProt database (UniProt Consortium, 2019), then a ClustalW MSA was performed via MEGA software (Kumar et al., 2018). Colour key for biochemical properties: yellow=hydrophobic; blue=basic; red=acidic; green=hydroxyl, amine; pink=glycine; indigo=proline; brown=sulfhydryl;, aqua=histidine.

| Protein | Signal Peptide Sequence | | |
|---|---|---|---|
| | n-region (positively charged) | h-region (hydrophobic) | c-region (polar, with A-X-A site) |
| 1. LamB | MMITLRKL-PLAV- | AVAAGVMS | AQAMA- |
| 2. MalE | MKIKTGAR-ILALSALTTMMF | | SASALA- |
| 3. OmpA | ----MKKT-AIAI-AVALAGF | | ATVAQA- |
| 4. OmpF | ---MMKRN-ILAV-IVPALLVAGT | | ANA- |
| 5. OmpN | ----MKSK-VLAL-LIPALLAAG | | AAHA- |
| 6. OmpC | ----MKVK-VLSL-LVPALLVAG | | AANA- |
| 7. OmpP | ---MQTK-LLAIMLAAPVVF | | SSQEASA |
| 8. OmpX | ---MKKIA-CLSA-LAAVLAFT | | AGTSVA |
| 9. PhoA | ----MKQS-TIAL-ALLPLLF | | TPVTKA- |
| 10. PhoE | ----MKKS-TLAL-VVMGIVASAS | | VQA- |
| 11. GBP | --MNKKVL-TLSA-VMASMLF | | GAAAHA- |
| 12. TolC | ----MKKLLPILI-GLSLSGF | | SSLSQA- |
| 13. TolB | ----MKQA--LRVAFGFLILWAS | | VLHA- |
| 14. CirA | ---MFRLNPFVRVGLCLSAISC | | AWPVLA |

tolerated and does not affect the overall function of the signal peptide. However, these slight changes can lead to differences in export efficiency, as discussed further in subsection 2.3.

## 2.2 Tat-specific signal peptides

Similar to Sec-specific signal peptides, Tat-specific signal peptides follow the same three-domain pattern (Freudl, 2018) (Figure 1). However, there are a few differences that distinguish Tat signal sequences from those destined for the Sec pathway, as depicted in Table 2. Most notable is the presence of the conserved Tat-specific motif at the boundary between the n- and h-regions: S/T-R-R-X-F-L-K, where X is usually a polar amino acid (Freudl, 2018). The twin arginines (R-R) are invariant in all Tat signal peptides, as this motif plays a key role in the binding of the Tat substrate to the export system (Cristóbal et al., 1999; Palmer & Berks, 2012) (Table 2). Additionally, the +2 and +3 amino acid residues (with respect to the twin arginines) are consistently hydrophobic, as evidenced in Table 2 (Brink et al., 1998; Cristóbal et al., 1999). Another distinction is that Tat signal peptides tend to be less hydrophobic than Sec signal peptides, particularly within the h-region (Cristóbal et al., 1999). This is in part due to comparatively more glycine (G) and less leucine (L) residues in the h-region of Tat signal peptides (Table 2). Additionally, Tat signal peptides are relatively longer than those for the Sec pathway, at average lengths of 38 and 24 amino acids respectively (Cristóbal et al., 1999; Freudl, 2018). A longer n-region in Tat signal peptides contributes to this increased length. The n-region also contains more negatively charged residues than what is seen in Sec signal peptides (Table 2). This may be attributed to the compensatory ability of the large positive charge from the twin-arginine motif (Cristóbal et al., 1999). Finally, Tat signal peptides tend to contain a few positively charged amino acid residues (R, K) within the c-region, termed the "Sec-avoidance" signal (Bogsch et al., 1997; Freudl, 2018) (Table 2). It is important to note that while the twin-arginine motif is ubiquitous in all Tat-specific signal peptides, it alone is insufficient to direct proteins to the Tat pathway (Bogsch et al., 1997; Chaddock et al., 1995; Palmer & Berks, 2012). Instead, a combination of these key distinguishing features is required.

## 2.3 Signal sequence optimization

Specific modifications to the signal sequence can be made to optimize the export of a given protein. For example, a higher charge-to-length ratio in the n-region typically leads to higher secretion efficiency in the Sec pathway (Peng et al., 2019). This was confirmed experimentally in *E. coli* by substituting positively charged residues with either neutral or negatively charged residues in the n-region, thereby reducing the positive charge (Inouye et al., 1982; Nesmeyanova et al., 1997). As a result of these substitutions, the rate of transport decreased and the accumulation of protein aggregates in the cytoplasm increased. Conversely, increasing the net charge of the n-region generally improved export efficiency as demonstrated by Ismail et al. (2011) in *E.*

| Protein | Signal Peptide Sequence | | | |
|---|---|---|---|---|
| | n-region (positively charged) | Tat motif S/T-R-R-X-F-L-K | h-region (hydrophobic) | c-region (polar, with A-X-A site) |
| 1. TorA | `---MNNNDLFQAS` | `RRRFLAQLG` | `-GLTVAGMLGPSLLTP` | `RRATA` |
| 2. TorZ | `---------MTLT` | `RREFIKHS-` | `-GIAAGALVVTSAA-PL` | `PAWA` |
| 3. AmiA | `--MSTFKPLKTLTS` | `RRQVLKAG-` | `---LAALTLSGMS---Q` | `AIA` |
| 4. AmiC | `MS----GSNTAIS` | `RRRLLQGAG` | `----AMWLLSVSQV----` | `SLA` |
| 5. FtsP | `---------MSLS` | `RRQFIQAS-` | `-GIA----LCAGAV-PL` | `KASA` |
| 6. EfeB | `MQYKDENGVNEPS` | `RRRLLKVIG` | `----ALALAGSCPV----` | `AHA` |
| 7. FdnG | `---------MDVS` | `RRQFFKICAGG` | `MAGTTVAALGFA-PKQ` | `ALA` |
| 8. FdoG | `---------MQVS` | `RRQFFKICAGG` | `MAGTTAAALGFA-PSV` | `ALA` |
| 9. YcbK | `---MDK---FDAN` | `RRKLLA-LG` | `-G--VA--LGAAIL-PTP` | `AFA` |
| 10. NrfC | `---------MTWS` | `RRQFLTGVGVL` | `AAVSGTAG-------R` | `VVA` |
| 11. WcaM | `------MPFKKLS` | `RRTFLTASS` | `----ALAFLHTPFA-----` | `RA` |
| 12. YahJ | `-------MKESNS` | `RREFLSQSGKMV` | `TAAALFGTSVP---L` | `AHA` |
| 13. MdoD | `---------MD--` | `RRRFIKGSMAMAAVCG` | `TSGIASLFSQAAFA` | |
| 14. FhuD | `MS-----GLPLIS` | `RRRLLTAMA` | `----LSPLLWQMNT----` | `AHA` |

**Table 2.** Multiple sequence alignment (MSA) of exemplar Tat signal peptide sequences in *E. coli*. Sample Tat substrates were selected (Tullman-Ercek et al., 2007) and signal peptide sequences were collected from the UniProt database (UniProt Consortium, 2019), then a ClustalW MSA was performed via MEGA software (Kumar et al., 2018). Colour key for biochemical properties: yellow = hydrophobic; blue = basic; red = acidic; green = hydroxyl, amine; pink = glycine; indigo = proline; brown = sulfhydryl; aqua = histidine.

*coli*, although this may be protein-specific. This effect is likely due to the electrostatic interactions that are present between the positively charged n-region and negatively charged residues found near the binding groove of the Sec translocase motor protein (SecA) (Chou & Gierasch, 2005; Gelis et al., 2007; Low et al., 2013). Altogether, increasing the charge-to-length ratio of the n-region generally increases transport efficiency for Sec substrates. Manipulating the charge of the n-region may also influence transport efficiency via the Tat pathway (Freudl, 2018). Notably, Li et al. (2006) found that charge distribution in the n-region was more significant than net charge when determining Tat-specific protein secretion rates. This further supports the notion that while there may be general trends, signal peptide optimization is still unique to a given signal sequence and protein.

The h-region can also be modified, as the degree of hydrophobicity appears to be pivotal when determining export efficiency. For example, increasing the hydrophobicity of this region can increase efficiency of export via the Sec pathway, as demonstrated by Chen et al. (1996) in *E. coli*. This is likely because the h-region adopts an α-helical conformation, which is able to form hydrophobic interactions with the SecA motor component of the Sec transport system (Chou & Gierasch, 2005; Gelis et al., 2007; Mori et al., 1997). By increasing the hydrophobicity, and thereby strengthening this interaction, the efficiency of transport may increase as SecA is more readily able to recognize Sec-specific signal peptides (Chou & Gierasch, 2005; Low et al., 2013). Besides affecting the levels of secreted protein, manipulating the hydrophobicity of the h-region can also change the targeted transport pathway. Natural Tat substrates can be redirected to the Sec pathway by increasing the hydrophobicity of the signal peptide (Cristóbal et al., 1999). This aligns with evidence that Sec signal peptides tend to be comparatively more hydrophobic, as discussed in subsection 2.2. Additionally, further increasing the hydrophobicity of a signal peptide destined for the SecB pathway can redirect it to the SRP-dependent pathway, which is an alternate branch of the Sec transport pathway (Bowers et al., 2003; Low et al., 2013). This is due to SRP preferentially binding longer, more hydrophobic α-helices (Low et al., 2013). In general, by manipulating the hydrophobicity of the h-region, both export efficiency and the targeted transport pathway can be altered.

Some current advancements are being made in an attempt to streamline signal peptide optimization. For instance, the generation and screening of "signal peptide libraries" have been performed for a few bacterial species to predict the optimal signal sequence (Brockmeier et al., 2006; Degering et al., 2010; Peng et al., 2019). However, it is important to note that while a given signal peptide may be optimal for one specific protein, it may perform poorly if fused to a different protein. As such, there is no universally optimal signal peptide. These signal peptide libraries and other publicly available databases may be beneficial as initial screening tools, and their utility will increase as they include more species or substrates (Goudenège et al., 2010; Low et al., 2013).

Overall, targeted modifications to the signal peptide can alter the protein's export efficiency. This has beneficial implications for the design and production of recombinant proteins in bacteria, as discussed in the following section.

## 3 | APPLICATIONS

Recombinant proteins such as growth factors and antibodies can be produced in bacteria for use in research, food production, or the pharmaceutical industry (Freudl, 2018). Bacterial systems boast many advantages such as low cost, high yield, short production time, and relatively easy genetic manipulation (Terpe, 2006). Despite drawbacks, such as differing codon biases and a lack of post-translational modification, bacterial systems remain a popular choice for the production of recombinant proteins. Notably, 30% of protein-based pharmaceuticals were produced by *E. coli* in 2009 (Ferrer-Miralles et al., 2009). To produce a desired recombinant protein in bacteria, a cleavable signal peptide can be fused at the N-terminus to allow for export out of the cytoplasm (Terpe, 2006). The specific sequence can be modified to optimize export efficiency, while still bearing in mind the limitations of the endogenous transport machinery (Freudl, 2018).

While recombinant proteins can be expressed and collected from the cytoplasm directly, there are many advantages to exporting the protein to the periplasm prior to collection via the incorporation of a signal peptide (Choi et al., 2000; Guerrero Montero et al., 2019). For example, as the periplasm is the only oxidizing compartment of the cell, it is the only location wherein disulfide bond formation will occur (Guerrero Montero et al., 2019; Pooley et al., 1996). As a result, any recombinant protein with disulfide bonds must be exported to the periplasm to exhibit proper tertiary structure. Additionally, exporting to the periplasm allows for easier downstream processing, as there are fewer

contaminants such as DNA in the periplasm when compared to the cytoplasm (Balasundaram et al., 2009; Guerrero Montero et al., 2019). Furthermore, protein aggregation can be prevented by exporting the recombinant protein out of the cytoplasm (Guerrero Montero et al., 2019). This is significant as cytoplasmic aggregates can be toxic to the bacterial cell, in addition to making downstream processing increasingly difficult. Finally, fewer proteases are present in the periplasm, leading to less turnover of the recombinant protein (Gottesman, 1996; Mergulhão et al., 2005). On the whole, exporting the desired protein to the periplasm via a carefully designed signal peptide is a key step in recombinant protein production.

Currently, many different recombinant proteins are being produced in bacterial systems and exported via the Sec or Tat pathway. For example, human growth hormone (Guerrero Montero et al., 2019), interferon α2b (Alanen et al., 2015), and human antibody fragments (Alanen et al., 2015) can be exported via Tat. Meanwhile, the Sec pathway has been utilized for the export of insulin-like growth factor 1 (Joly et al., 1998), human epidermal growth factor (Wong & Sutherland, 1993), parathyroid hormone (Wong & Sutherland, 1993), and alkaline phosphatase (Choi et al., 2000). In either case, the specific targeting of the recombinant protein to an export pathway is achieved by the fusing of a signal peptide to the N-terminal region. In the following subsections, the production of two key biopharmaceuticals in *E. coli* will be discussed in further depth.

### 3.1 Human growth hormone
One specific example of a biopharmaceutical which can be produced and exported by bacteria is human growth hormone (hGH) (Guerrero Montero et al., 2019). hGH can aid in the treatment of hypopituitarism, obesity, and burn/wound healing (Isaksson et al., 1985). Bacterial systems are a good candidate for producing hGH as it has few disulfide bonds and no glycosylation (Guerrero Montero et al., 2019; Ultsch et al., 1994). However, hGH adopts a complex, folded conformation prior to transport out of the cytoplasm. As such, it is unable to be secreted by the Sec pathway, and the Tat pathway must be used instead. Montero et al. (2019) demonstrated this by fusing the TorA Tat signal peptide to hGH (TorA-hGH). This fusion protein was expressed in *E. coli* "TatExpress" strains, which have elevated levels of Tat proteins to circumvent the slower transport typically associated with the Tat pathway. Mature hGH was found to be readily abundant in the periplasm, with proper disulfide bond formation and correct cleavage of the

TorA signal peptide. Additionally, the yield was quite high at 2.39-5.4 g/L of culture, which far exceeds the yield expectations for bacterial recombinant proteins (typically 0.5-0.8 g/L) (Georgiou & Segatori, 2005). This is likely due to the use of the "TatExpress" strain, allowing for increased export of hGH to the periplasm. The increased export capacity also aids in reducing insoluble hGH aggregates in the cytoplasm, which is a prevalent issue in this field (Guerrero Montero et al., 2019; Patra et al., 2000). Overall, bacterial secretion of TorA-hGH via the Tat pathway appears to be a viable option for the mass production of recombinant hGH.

### 3.2 Full-length monoclonal antibodies
Another specific example of exploiting bacterial signal peptides is the production and secretion of human monoclonal antibodies in *E. coli* (Zhou et al., 2016). Monoclonal antibodies are incredibly valuable to the pharmaceutical industry, as they can be used to treat a wide variety of human diseases such as breast cancer, lung cancer, rheumatoid arthritis, ulcerative colitis, and psoriasis (Spadiut et al., 2014). However, secretion efficiency remains a limitation for monoclonal antibody production in bacteria. In an attempt to circumvent this drawback, Zhou et al. (2016) tested four Sec-specific signal peptides (STII, DsbA, PhoA, MalE) fused to the N-terminus of the heavy chain (HC) antibody component. Of these, the signal peptide from DsbA was deemed to be the most efficient. Zhou et al. (2016) speculated that this was due to the high hydrophobicity of the h-region when compared to the other signal peptides tested. Accordingly, they attempted to further improve secretion efficiency by increasing the hydrophobicity of the h-region for two signal peptides (STII, DsbA). In doing so, Zhou et al. (2016) found that the proportion of mature, secreted HC could be doubled with only a single amino acid substitution in the h-region, wherein a polar amino acid residue was mutated to a more hydrophobic one. Altogether, this demonstrates that the optimization principles discussed in subsection 2.3 can be applied to recombinant protein production in bacterial systems.

## 4 | FUTURE DIRECTIONS
The expanding knowledge of signal peptides and their key features has proven to be incredibly valuable when applied to recombinant protein production in bacterial systems. In the future, natural signal peptides could be modified to improve secretion efficiency, such as increasing the charge of the n-region or increasing the length and hydrophobicity of the h-

region (Low et al., 2013). While there have already been examples of this in the literature (as discussed in subsection 2.3), the optimization of signal peptides is an ongoing process due to their protein- and species-specific nature. As such, each signal peptide must be optimized anew when considering a novel protein or substrate. These optimization efforts may be further aided by the expansion of current signal peptide libraries to include a wider variety of bacterial species and proteins, as well as the development of high-throughput signal peptide screening methods (Goudenège et al., 2010). Another future direction may involve designing signal peptides de novo, as demonstrated by Mhiri et al. (2000) in *Streptomyces*. After comparing the most efficient natural signal peptides, they designed a synthetic signal peptide which then outperformed the natural signal peptide up to six-fold. This remarkable proof of concept opens the door for designing synthetic signal peptides. Finally, much is still unknown about the specific interactions between portions of the signal peptide and the transport machinery at the residue level (Blaudeck et al., 2003; Gelis et al., 2007). As such, improving our understanding of the structural and functional roles of each motif may further advance attempts to optimize natural signal sequences or design them de novo.

## 5 | CONCLUSION

The importance of bacterial signal peptides cannot be understated, given that they appear to be the sole determiner of transport pathway targeting (Bogsch et al., 1997). Through modifying pre-existing signal peptides or designing them de novo, transport efficiency can be optimized (Low et al., 2013). At the same time, identifying conserved regions, such as the twin-arginine motif or the A-X-A signal peptidase site, provides boundaries for regions which should not be perturbed. These optimized signal sequences can then be exploited for the production of desired proteins, such as human hormones or antibodies, in bacterial systems (Freudl, 2018). Overall, the expanding knowledge of signal peptides is a promising step forward when applied to the field of recombinant protein production.

## ACKNOWLEDGEMENTS

## REFERENCES

Alanen, H. I., Walker, K. L., Lourdes Velez Suberbie, M., Matos, C. F. R. O., Bönisch, S., Freedman, R. B., Keshavarz-Moore, E., Ruddock, L. W., & Robinson, C. (2015). Efficient export of human growth hormone, interferon α2b, and antibody fragments to the periplasm by the Escherichia coli Tat pathway in the absence of prior disulfide bond formation. *Biochimica et Biophysica Acta, 1853*(3), 756-763.

Bagos, P. G., Nikolaou, E. P., Liakopoulos, T. D., & Tsirigos, K. D. (2010). Combined prediction of Tat and Sec signal peptides with hidden Markov models. *Bioinformatics (Oxford, England), 26*(22), 2811-2817.

Balasundaram, B., Harrison, S., & Bracewell, D. G. (2009). Advances in product release strategies and impact on bioprocess design. *Trends in Biotechnology, 27*(8), 477-485.

Bendtsen, J. D., Nielsen, H., Widdick, D., Palmer, T., & Brunak, S. (2005). Prediction of twin-arginine signal peptides. *BMC Bioinformatics, 6*(1), 167.

Blaudeck, N., Kreutzenbeck, P., Freudl, R., & Sprenger, G. A. (2003). Genetic analysis of pathway specificity during posttranslational protein translocation across the Escherichia coli plasma membrane. *Journal of Bacteriology, 185*(9), 2811-2819.

Bogsch, E., Brink, S., & Robinson, C. (1997). Pathway specificity for a ΔpH-dependent precursor thylakoid lumen protein is governed by a 'Sec-avoidance' motif in the transfer peptide and a 'Sec-incompatible' mature protein. *The EMBO Journal, 16*(13), 3851-3859.

Bowers, C. W., Lau, F., & Silhavy, T. J. (2003). Secretion of LamB-LacZ by the signal recognition particle pathway of Escherichia coli. *Journal of Bacteriology, 185*(19), 5697-5705.

Brink, S., Bogsch, E. G., Edwards, W. R., Hynds, P. J., & Robinson, C. (1998). Targeting of thylakoid proteins by the delta pH-driven twin-arginine translocation pathway requires a specific signal in the hydrophobic domain in conjunction with the twin-arginine motif. *FEBS Letters, 434*(3), 425-430.

Brockmeier, U., Caspers, M., Freudl, R., Jockwer, A., Noll, T., & Eggert, T. (2006). Systematic screening of all signal peptides from Bacillus subtilis: a powerful strategy in optimizing heterologous protein secretion in Gram-positive bacteria. *Journal of Molecular Biology, 362*(3), 393-402.

Chaddock, A. m., Mant, A., Karnauchov, I., Brink, S., Herrmann, R. G., Klösgen, R. B., & Robinson, C. (1995). A new type of signal peptide: Central role of a twin-arginine motif in transfer signals for the delta pH-dependent thylakoidal protein translocase. *The EMBO Journal, 14*(12), 2715-2722.

Chen, H., Kim, J., & Kendall, D. A. (1996). Competition between functional signal peptides demonstrates variation in affinity for the secretion pathway. *Journal of Bacteriology, 178*(23), 6658-6664.

Choi, J. H., Jeong, K. J., Kim, S. C., & Lee, S. Y. (2000). Efficient secretory production of alkaline phosphatase by high cell density culture of recombinant Escherichia coli using the

Bacillus sp. endoxylanase signal sequence. *Applied Microbiology and Biotechnology*, *53*(6), 640-645.

Chou, Y. T., & Gierasch, L. M. (2005). The conformation of a signal peptide bound by Escherichia coli preprotein translocase SecA. *Journal of Biological Chemistry*, *280*(38), 32753-32760.

Crane, J. M., & Randall, L. L. (2017). The Sec System: Protein Export in *Escherichia coli*. *EcoSal Plus*, *7*(2).

Cristóbal, S., de Gier, J.-W., Nielsen, H., & von Heijne, G. (1999). Competition between Sec- and TAT-dependent protein translocation in Escherichia coli. *The EMBO Journal*, *18*(11), 2982-2990.

Degering, C., Eggert, T., Puls, M., Bongaerts, J., Evers, S., Maurer, K.-H., & Jaeger, K.-E. (2010). Optimization of protease secretion in Bacillus subtilis and Bacillus licheniformis by screening of homologous and heterologous signal peptides. *Applied and Environmental Microbiology*, *76*(19), 6370-6376.

Emr, S. D., Schwartz, M., & Silhavy, T. J. (1978). Mutations altering the cellular localization of the phage lambda receptor, an Escherichia coli outer membrane protein. *Proceedings of the National Academy of Sciences*, *75*(12), 5802-5806.

Ferrer-Miralles, N., Domingo-Espín, J., Corchero, J. L., Vázquez, E., & Villaverde, A. (2009). Microbial factories for recombinant pharmaceuticals. *Microbial Cell Factories*, *8*(1), 17.

Freudl, R. (2018). Signal peptides for recombinant protein secretion in bacterial expression systems. *Microbial Cell Factories*, *17*(1), 52.

Gelis, I., Bonvin, A. M., Keramisanou, D., Koukaki, M., Gouridis, G., Karamanou, S., Economou, A., & Kalodimos, C. G. (2007). Structural basis for signal-sequence recognition by the translocase motor SecA as determined by NMR. *Cell*, *131*(4), 756-769.

Georgiou, G., & Segatori, L. (2005). Preparative expression of secreted proteins in bacteria: Status report and future prospects. *Current Opinion in Biotechnology*, *16*(5), 538-545.

Gottesman, S. (1996). Proteases and their targets in Escherichia coli. *Annual Review of Genetics*, *30*(1), 465-506.

Goudenège, D., Avner, S., Lucchetti-Miganeh, C., & Barloy-Hubler, F. (2010). CoBaltDB: Complete bacterial and archaeal orfeomes subcellular localization database and associated resources. *BMC Microbiology*, *10*(1), 88.

Green, E. R., & Mecsas, J. (2016). Bacterial Secretion Systems: An Overview. *Microbiology spectrum*, *4*(1).

Guerrero Montero, I., Richards, K. L., Jawara, C., Browning, D. F., Peswani, A. R., Labrit, M., Allen, M., Aubry, C., Davé, E., Humphreys, D. P., Busby, S. J. W., & Robinson, C. (2019). Escherichia coli "TatExpress" strains export several g/L human growth hormone to the periplasm by the Tat pathway. *Biotechnology and Bioengineering*, *116*(12), 3282-3291.

Hiller, K., Grote, A., Scheer, M., Münch, R., & Jahn, D. (2004). PrediSi: Prediction of signal peptides and their cleavage positions. *Nucleic Acids Research*, *32*(Web Server issue), W375-W379.

Inouye, S., Soberon, X., Franceschini, T., Nakamura, K., Itakura, K., & Inouye, M. (1982). Role of positive charge on the amino-terminal region of the signal peptide in protein secretion across the membrane. *Proceedings of the National Academy of Sciences*, *79*(11), 3438-3441.

Isaksson, O. G., Edén, S., & Jansson, J. O. (1985). Mode of action of pituitary growth hormone on target cells. *Annual Review of Physiology*, *47*(1), 483-499.

Ismail, N. F., Hamdan, S., Mahadi, N. M., Murad, A. M., Rabu, A., Bakar, F. D. , Klappa, P., & Illias, R. M. (2011). A mutant L-asparaginase II signal peptide improves the secretion of recombinant cyclodextrin glucanotransferase and the viability of Escherichia coli. *Biotechnology Letters*, *33*(5), 999-1005.

Joly, J. C., Leung, W. S., & Swartz, J. R. (1998). Overexpression of Escherichia coli oxidoreductases increases recombinant insulin-like growth factor-I accumulation. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(6), 2773-2777.

Käll, L., Krogh, A., & Sonnhammer, E. L. L. (2007). Advantages of combined transmembrane topology and signal peptide prediction—The Phobius web server. *Nucleic Acids Research*, *35*(Web Server issue), W429-W432.

Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Molecular Biology and Evolution*, *35*(6), 1547-1549.

Li, H., Faury, D., & Morosoli, R. (2006). Impact of amino acid changes in the signal peptide on the secretion of the Tat-dependent xylanase C from Streptomyces lividans. *FEMS Microbiology Letters*, *255*(2), 268-274.

Low, K. O., Muhammad Mahadi, N., & Md Illias, R. (2013). Optimisation of signal peptide for recombinant protein secretion in bacterial hosts. *Applied Microbiology and Biotechnology*, *97*(9), 3811-3826.

Lüke, I., Handford, J. I., Palmer, T., & Sargent, F. (2009). Proteolytic processing of Escherichia coli twin-arginine signal peptides by LepB. *Archives of Microbiology*, *191*(12), 919-925.

Mergulhão, F. J. M., Summers, D. K., & Monteiro, G. A. (2005). Recombinant protein secretion in Escherichia coli. *Biotechnology Advances*, *23*(3), 177-202.

Mhiri, S., Mezghani, M., Mellouli, L., Ben Ali, M., Belguith, K., & Bejar, S. (2000). Efficient synthetic signal peptides for Streptomyces. *Biotechnology Letters*, *22*(16), 1305-1310.

Molloy, M. P., Herbert, B. R., Slade, M. B., Rabilloud, T., Nouwens, A. S., Williams, K. L., & Gooley, A. A. (2000). Proteomic analysis of the Escherichia coli outer membrane. *European Journal of Biochemistry*, *267*(10), 2871-2881.

Mori, H., Araki, M., Hikita, C., Tagaya, M., & Mizushima, S. (1997). The hydrophobic region of signal peptides is involved in the interaction with membrane-bound SecA. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, *1326*(1), 23-36.

Nesmeyanova, M. A., Karamyshev, A. L., Karamysheva, Z. N., Kalinin, A. E., Ksenzenko, V. N., & Kajava, A. V. (1997). Positively charged lysine at the N-terminus of the signal peptide of the Escherichia coli alkaline phosphatase provides the secretion efficiency and is involved in the interaction with anionic phospholipids. *FEBS Letters*, *403*(2), 203–207.

Palmer, T., & Berks, B. C. (2012). The twin-arginine translocation (Tat) protein export pathway. *Nature Reviews. Microbiology*, *10*(7), 483–496.

Patra, A. K., Mukhopadhyay, R., Mukhija, R., Krishnan, A., Garg, L. C., & Panda, A. K. (2000). Optimization of inclusion body solubilization and renaturation of recombinant human growth hormone from Escherichia coli. *Protein Expression and Purification*, *18*(2), 182–192.

Peng, C., Shi, C., Cao, X., Li, Y., Liu, F., & Lu, F. (2019). Factors Influencing Recombinant Protein Secretion Efficiency in Gram-Positive Bacteria: Signal Peptide and Beyond. *Frontiers in Bioengineering and Biotechnology*, *7*, 139.

Petersen, T. N., Brunak, S., von Heijne, G., & Nielsen, H. (2011). SignalP 4.0: Discriminating signal peptides from transmembrane regions. *Nature Methods*, *8*(10), 785–786.

Pooley, H. M., Merchante, R., & Karamata, D. (1996). Overall protein content and induced enzyme components of the periplasm of Bacillus subtilis. *Microbial Drug Resistance*, *2*(1), 9–15.

Rose, R. W., Brüser, T., Kissinger, J. C., & Pohlschröder, M. (2002). Adaptation of protein secretion to extremely high-salt conditions by extensive use of the twin-arginine translocation pathway. *Molecular Microbiology*, *45*(4), 943–950.

Rusch, S. L., & Kendall, D. A. (2007). Interactions That Drive Sec-Dependent Bacterial Protein Transport. *Biochemistry*, *46*(34), 9665–9673.

Spadiut, O., Capone, S., Krainer, F., Glieder, A., & Herwig, C. (2014). Microbials for the production of monoclonal antibodies and antibody fragments. *Trends in Biotechnology*, *32*(1), 54–60.

Terpe, K. (2006). Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. *Applied Microbiology and Biotechnology*, *72*(2), 211–222.

Tullman-Ercek, D., DeLisa, M. P., Kawarasaki, Y., Iranpour, P., Ribnicky, B., Palmer, T., & Georgiou, G. (2007). Export pathway selectivity of Escherichia coli twin arginine translocation signal peptides. *Journal of Biological Chemistry*, *282*(11), 8309–8316.

UniProt Consortium (2019). UniProt: A worldwide hub of protein knowledge. *Nucleic Acids Research*, *47*(D1), D506–D515.

Ultsch, M. H., Somers, W., Kossiakoff, A. A., & de Vos, A. M. (1994). The crystal structure of affinity-matured human growth hormone at 2 Å resolution. *Journal of Molecular Biology*, *236*(1), 286–299.

Wong, W. K. R., & Sutherland, M. L. (1993). *Excretion of heterologous proteins from e. coli* (United States Patent No. US5223407A). U.S. Patent and Trademark Office. https://patents.google.com/patent/US5223407A/en

Zhou, Y., Liu, P., Gan, Y., Sandoval, W., Katakam, A. K., Reichelt, M., Rangell, L., & Reilly, D. (2016). Enhancing full-length antibody production by signal peptide engineering. *Microbial Cell Factories*, *15*(1), 47.

*How to cite this article:*

Erkut, Esra. (2021). Bacterial signal peptides: structure, optimization, and applications. *Eureka University of Alberta Science Undergraduate Research Journal.*, 6(1). https://doi.org/10.29173/eureka28759