Academic Paper

Protein Expression with Signal Peptide Using *Escherichia coli*

Compiled By:

Brilliant Cahya Puspasari (199303082024062002)

FAKULTAS KEDOKTERAN

INSTITUT PERTANIAN BOGOR

BOGOR

2024

Table of Contents

Introduction

As we know that *Escherichia coli* is one of the most widely used hosts for the production of recombinant proteins. *Escherichia coli* could express protein and secret it to the extracellular medium, to the periplasmic production, intracellular production as IBs (insoluble form), and intracellular and soluble production. By dividing into soluble expression and insoluble expression (inclusion body) there are some different things that should be noted. Insoluble expression could be high productivity in culture process, low impurity, low endotoxin level (mostly removed during IBs recovery) but its needs for refolding process. For the soluble expression it could be low productivity in culture process, high impurity, no need to solubilization and refolding but high endotoxin level. Overexpressed proteins are often produced in the form of inclusion bodies, from which biologically active proteins can only be recovered by complicated and costly denaturation and refolding processes. Furthermore, the final yields of these soluble refolded proteins are usually very low, due mainly to protein aggregation resulting from interactions between the hydrophobic regions of the proteins (Choi and Lee 2004). This is because *Escherichia coli* does not have PTM processing. To prevent this condition, strategies for secretory and extracellular production of recombinant proteins using E. coli could be used (Choi and Lee 2004).

A common strategy in engineering proteins is using N-terminal signal peptide to facilitate their secretion to the periplasm (Mirzadeh, Shilling et al. 2020). The signal peptide is a short amino acid sequence at the N-terminus of the protein that directs the protein to the Sec translocon in the cytoplasmic membrane (Karyolaimos, Ampah-Korsah et al. 2019). The signal peptide is cleaved off the protein that passes through the translocon and released into the periplasm. The process need enhances the correct folding of the protein by providing an oxidizing environment (Mirzadeh, Shilling et al. 2020).

The efficiency of the protein secretion affected by various factors, including the sequence of the signal peptide and the translation initiation region (Mirzadeh, Shilling et al. 2020).

Literature Review

2.1 Escherichia coli

Escherichia coli is a type of bacteria that is commonly found in the intestines of humans and other warm-blooded animals. It belongs to the Enterobacteriaceae family and is a Gram-negative, facultatively anaerobic, rod-shaped bacterium. Most strains of E. coli are harmless and play a crucial role in maintaining a healthy gut microbiome. However, some strains can cause serious foodborne illnesses (Bonten, Johnson et al. 2021).

Protein expression in *Escherichia coli* is cornerstone of biotechnology. It enabling the production of recombinant proteins for various applications. The Implementation of N-terminal signal peptides is a common strategy to directs protein to the periplasm (Mirzadeh, Shilling et al. 2020).

2.2 Signal peptides

There are five functions that N-terminal signal peptides that uses in protein expression using *Escherichia coli.* First, N-terminal signal peptides as protein targeting. It guides proteins to the Sec translocon in the cytoplasmic membrane then cleaved off the protein allowing it to fold correctly. Second, N-terminal signal pepties as optimizer for higher yields. Optimizing the N-terminal 5–7 amino acid sequences of signal peptides could enhance the expression and secretion of alkaline pectin lyase in Bacillus subtilis. Third, N-terminal signal peptides as chaperone co-expression. Co-expressing chaperones like GroEL-GroES with enzymes from Bacillus subtilis and Pseudomonas stutzeri significantly increased the production of active enzymes in E. coli (Zhang, Zhen et al. 2024). Fourth, N-terminal signal peptides as Evolutionary Relevance (Kunze and Berger 2015). The last, N-terminal signal peptides as Cleavable Signal Peptides. It can enhance the surface expression of proteins in heterologous cells (Shepard, Natarajan et al. 2013).

Soluble form by *Escherichia coli* does not need refolding process. The expression of *Escherichia coli* in extracellular production needs signal peptide for protein secretion and could be folded. Signal peptide is a short peptide (usually 16-30 amino acids long) present at the N-terminus of the majority of newly synthesized

proteins that are destined towards the secretory pathway. There are many kinds of signal peptide (SPs). Several different secretory pathways have evolved among organisms: 1) to come up with the secretory needs of rapidly growing organisms such as bacteria and yeasts, 2) to offset the low rate of protein synthesis versus the high rate of secretion, and 3) to be able to secrete proteins with different characteristics. To optimize protein secretion could be by choosing the most appropriate SPs or manipulating them with respect to the attributes of the secretory protein. Different types of secretory pathways depending on the N-terminal signal, the related organisms (Owji, Nezafat et al. 2018). The related organisms, and their specific features are summarized in the table below:

Table 1. Secretory pathway relying in N-terminal signal peptide sequence (Owji, Nezafat et al. 2018)

Secretory pathways relying on N-terminal signal sequence (SPs).

Secretory pathway	Species	Specific features	Ref.
Sec-dependent pathway	Present in all taxa Prevail in Gram-negative bacteria	• Conducting post-translational translocation • Having higher capacity for recombinant protein production • Preventing unwanted degradation and folding of proteins • ATP-dependent	(Rusch and Kendall, 2007; Green and Mecsas, 2016; Yuan et al., 2010; Plath et al., 1998; Mordkovich et al., 2015; Papanikou et al., 2007; Hardy and Randall, 1991)
SRP-dependent pathway	Present in all taxa Prevail in eukaryotes	• Conducting co-translational translocation • Appropriate for rapid folding proteins • ATP-dependent	(Yuan et al., 2010; Rapoport et al., 1996) (Luirink and Dobberstein, 1994; Bornemann et al., 2008; Lange et al., 2016)
Tat-dependent pathway	Present in all taxa Prevail in Gram-positive bacteria and archaea	• Carry mature and rapid-folding proteins • Mostly carry SPs with twin-Arg motif • Useful pathway for recombinant protein production in Bacillus • Appropriate for translocation of multimeric and co-factor dependent- enzymes • Dependent on PMF	(Yuan et al., 2010; Palmer and Berks, 2012; van Dijl et al., 2002; Berks et al., 2000; Müller, 2005; Ulfig et al., 2017; Nielsen et al., 1997)
ABC transporter-associated pathway	Present in all taxa Prevail in Gram-positive bacteria	• Translocate proteins not compatible with the previous pathways including lantibiotics and pheromones • Transporting SPs called leader peptides • ATP-dependent	(Fath and Kolter, 1993) (Quentin et al., 1999)
YidC pathway	Present in all taxa	• Responsible for membrane insertion of proteins • Can act in parallel with other pathways	(Yuan et al., 2010; Samuelson et al., 2000; Scotti et al., 2000; Yi and Dalbey, 2005)
SecA2 pathway	Gram-positive specific	• Used in stress condition and virulence of bacteria • Translocate protein through SecYEG secretion system	(Green and Mecsas, 2016; Rigel and Braunstein, 2008; Bensing et al., 2014)
SecA2-SecY2 (aSec) pathway	Gram-positive specific	• Responsible for translocation of large, highly-glycosylated anchor proteins with serine-rich repeats (such as virulent protein) • Translocate proteins through SecY2 channel	(Bensing et al., 2014; Siboo et al., 2008; Mistou et al., 2009)
Injectosome pathway	Gram-positive specific Observed in *Streptococcus pyogens*	• Translocation of Sec-transported toxins between bacterium and the host cell • Functional analog of T3SS	(Madden et al., 2001; Ghosh et al., 2010)
Type II secretion system (T2SS)	Gram-negative specific	• Secrete enzymes and virulence factor out of cell wall • Transport proteins translocated by Sec or Tat pathway • Transport folded proteins	(Korotkov et al., 2012)
Type III secretion system (T3SS) (injectosome) (needle and syringe)	Gram-negative specific	• Transport unfolded proteins, esp virulence factor, in one-step process • Carry uncleavable SPs	(Büttner, 2012)
Type V secretion system (T5SS)	Gram-negative specific	• Transport SecB-translocated virulence factors • Self-secretion of protein harboring a β-barrel domain • Including three pathways, autotransporter, two-partner, and chaperone-usher	(Green and Mecsas, 2016; Leyton et al., 2012)

The appropriate and usually could be used to this production strategy from *Escherichia coli* are listed below.

Table 2. Signal Sequences for recombinant-protein in E.Coli (Choi and Lee 2004).

Table 1 Representative signal sequences used for the secretory production of recombinant proteins in *Escherichia coli*. The signal sequence is composed of N-, H- and C-domains. The N-domains of signal sequences are shown in *bold* while the C-domains are *underlined*

Signal sequences	Amino acid sequences
PelB (pectate lyase B) from *Erwinia carotovora*	**MKY**LLPTAAAGLLLLAAQPAMA
OmpA (outer-membrane protein A)	**MKKT**AIAIAVALAGFATVAQA
StII (heat-stable enterotoxin 2)	**MKK**NIAFLLASMFVFSIATNAYA
Endoxylanase from *Bacillus* sp.	**MFKFKKK**FLVGLTAAFMSISMFSATASA
PhoA (alkaline phosphatase)	**MKQS**TIALALLPLLFTPVTKA
OmpF (outer-membrane protein F)	**MMKR**NILAVIVPALLVAGTANA
PhoE (outer-membrane pore protein E)	**MKKS**TLALVVMGIVASASVQA
MalE (maltose-binding protein)	**MKIKT**GARILALSALTTMMFSASALA
OmpC (outer-membrane protein C)	**MKVK**VLSLLVPALLVAGAANA
Lpp (murein lipoprotein)	**MKATK**LVLGAVILGSTLLAG
LamB (λ receptor protein)	**MMITLRK**LPLAVAVAAGVMSAQAMA
OmpT (protease VII)	**MRAK**LLGIVLTTPIAISSFA
LTB (heat-labile enterotoxin subunit B)	**MNKVK**CYVLFTALLSSLYAHG

There are 3 pathways for secretory system, they are Sec pathway, SRP pathway, and Tat pathway. The Sec-pathway as known as post translational translocation, it is a pathway in which proteins remain unfolded to be sufficiently recognized and translocated via the secretory machinery. Sec machinery consists of SecB (or DnaK), SecA, SecY, SecE, SecG, and SecDF protein. The Sec pathway, which seems to be conserved in all classes of lives, is classified as SecB or SecA2 dependent pathway. The SecB pathways is considered the prevailing pathway in all classes of bacteria and SecA2 is regarded Gram-positive specific. More than 90% of E. Coli proteins follow the Sec pathway. SecB acts as a chaperone, which prevents the protein folding before secretion and subsequently, the pre-protein is guided to SecA. SecA was bound to the pre-protein and ATP, it was inserted into the cell membrane. In the cell membrane, the SecA/pre-protein complex forms a translocation complex with SecY/SecE/SecG. SecY and SecE form a channel conducting protein out of the cell, and SecG activates translocation. The mature protein is released by SecD (or SecDF complex) as the signal peptidase (SPase) cleaves the SP (Owji, Nezafat et al. 2018).

SRP pathway or co-translational translocation is a pathway occurring co-translationally, which prevents any form of cytoplasmic modification. Proteins translocated across the endoplasmic reticulum (ER) membrane utilize the SRP pathway, particularly the ones with more than 100 amino acids (aas) (Owji, Nezafat et al. 2018).

The last pathway is Tat pathway (The twin-Arg transport system). It is a Sec-independent pathway that translocates protein in a mature state. It is from Sps that

carries the twin-Arg motif that found in bacteria, archaea, plant, chloroplast and few plant mitochondria with a higher prevalence in archaea and Gram-positive bacteria (Owji, Nezafat et al. 2018). In contrast to Sec-pathway, Tat system exports folded proteins to the periplasm. In *Escherichia coli*, this system comprises three integral membrane proteins, TatA, TatB, and TatC, which form the TatABC complex that works as a receptor and channel for the export of Tat substrates. Tat system requires the presence of a signal peptide to direct protein export. Although Tat signal peptides share the same general structure with Sec signal peptides (n-region, h-region, and c-region), they have different characteristics: a conserved twin-arginine motif (T/ SRRXFLK, where X is any polar amino acid) at the union of the n- and h-regions, and a less hydrophobic h-region (Santos, Morones-Ramirez et al. 2019).

2.3 Bone Morphogenetic Proteins (BMPs)

Bone Morphogenetic Proteins (BMPs) are a group of growth factors known for their ability to induce the formation of bone and cartilage. They belong to the Transforming Growth Factor-beta (TGF-β) superfamily and play crucial roles in various biological processes (Miyazono, Kamiya et al. 2010). They are essential for bone and cartilage development, as well as for the repair of fractures and cartilage lesions (Martínez, Sacedon et al. 2015). BMPs bind to type II and type I serine-threonine kinase receptors, which then transduce signals through Smad and non-Smad signaling pathways. BMPs are involved in the regulation of cell proliferation, survival, differentiation, and apoptosis. They also play roles in tooth, kidney, skin, hair, muscle, hematopoietic, and neuronal development (Miyazono, Kamiya et al. 2010). BMPs have potential therapeutic applications in treating bone diseases, enhancing spinal fusion, and repairing critical-size bone defects (Martínez, Sacedon et al. 2015).

2.4 Method

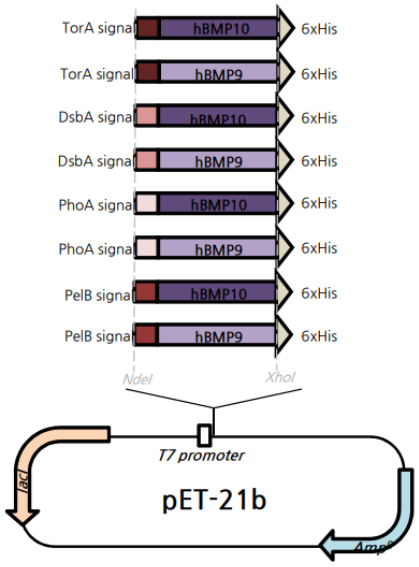
Protein that would be expressed = BMP-9 and BMP-10

Signal peptide	**PelB**	SecB-dependent (SEC) pathway	+ 22 aa
	PhoA	SecB-dependent (SEC) pathway	+ 21 aa
	DsbA	Signal recognition particle (SRP)	+ 19 aa
	TorA	Twin-arginine translocation (TAT)	+ 39 aa

DNA construct would be:

1.	BMP-9 + MOCK CELL	BMP-10 + MOCK CELL
2.	BMP-9 + PelB	BMP-10 + PelB
3.	BMP-9 + PhoA	BMP-10 + PhoA
4.	BMP-9 + DsbA	BMP-10 + DsbA
5.	BMP-9 + TorA	BMP-9 + TorA

Cloning scheme & sequences



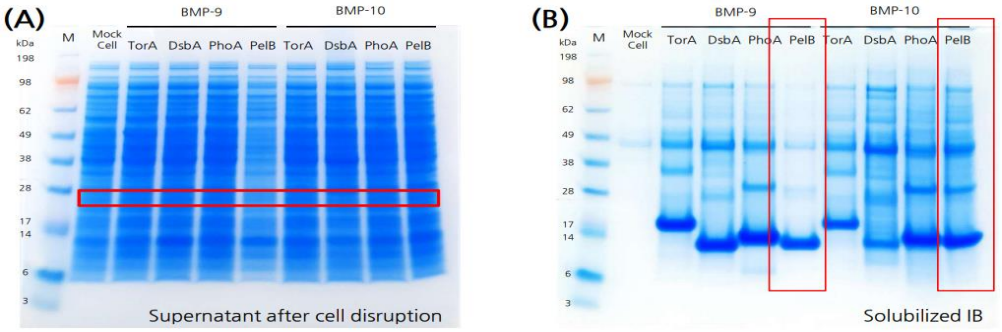
Signal peptide	PelB	SecB-dependent (SEC) pathway	ATGAAATACCTGCTGCCGACCGCTGCTGCTGGTCTGCTG CTCCTCGCTGCCCAGCCGGCGATGGCG
	PhoA	SecB-dependent (SEC) pathway	ATGAAACAAAGCACTATTGCACTGGCACTCTTACCGTTA CTGTTTACCCCTGTGACAAAAGCG
	DsbA	Signal recognition particle (SRP)	ATGAAAAAGATTTGGCTGGCGCTGGCTGGTTTAGTTTTA GCGTTTAGCGCATCGGCG
	TorA	Twin-arginine translocation (TAT)	ATGAACAACAACGACCTGTTCCAGGCGAGCCGTCGTCG TTTCCTGGCGCAACTGGGTGGCCTGACCGTGGCGGGTA TGCTGGGTCCGAGCCTGCTGACCCCGCGTCGTGCGACC GCG
Target gene	hBMP9 (codon optimized)		ATGTCAGCAGGGGCTGGAAGTCACTGTCAAAAAACGA GCCTGCGTGTTAACTTTGAAGATATCGGCTGGGATAGCT GGATCATCGCCCCAAAAGAATACGAGGCGTATGAGTG CAAAGGTGGCTGCTTTTTCCCGTTGGCGGATGACGTGAC CCCGACCAAGCACGCGATTGTTCAGACCTTGGTTCATTT AAAGTTCCCGACTAAGGTGGGTAAAGCATGTTGTGTTCC GACCAAGCTGTCTCCGATTTCCGTGCTGTATAAAGACGA CATGGGTGTGCCGACCCTGAAATACCACTACGAGGGCA TGAGCGTCGCTGAATGCGGTTGCCGCTGA
	hBMP10 (codon optimized)		ATGAATGCTAAAGGAAACTATTGTAAGAGGACACCGTT GTACATTGATTTTAAAGAGATCGGCTGGGACAGCTGGA TCATCGCCCCACCGGGTTATGAGGCGTATGAGTGCCGC GGTGTCTGTAATTACCCGCTCGCTGAACACTTGACCCCG ACGAAGCACGCAATTATCCAAGCGCTGGTGCATCTGAA GAACTCTCAGAAAGCGTCCAAAGCGTGCTGCGTGCCGA CCAAGCTGGAACCGATTAGCATTCTGTACCTGGACAAA GGCGTTGTTACCTATAAATTCAAGTACGAAGGCATGGC AGTTAGCGAATGCGGTTGTCGTTGA

Picture 2.1 Cloning scheme and sequenes

DNA constructs were transformed into *Escherichia coli* BL21(DE3). A single colony was used to prepare a starter culture. After culture, the supernatant of cell disruption and the inclusion body (IB) were analyzed respectively to check the expression/localization.

2.5 Analyze Result

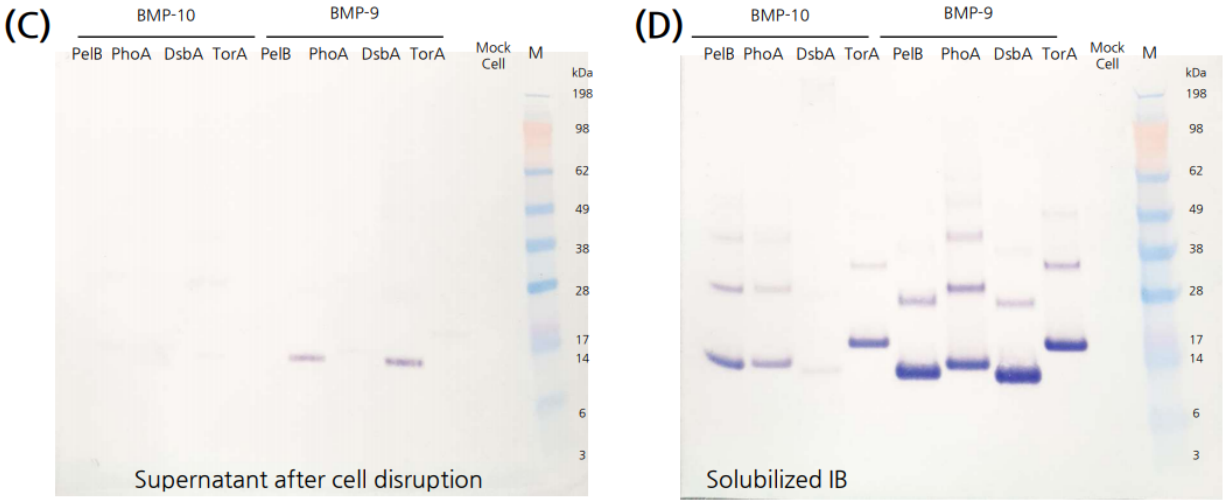
SDS-PAGE



- After collecting the pellet by centrifuging 5 mL of the final culture solution, disrupt the cells with 1 mL of Bugbuster™.
- IBs were dissolved in 1 mL of BMP-2 solubilization buffer (10 M urea + 50 mM DTT)
- After disruption, the supernatant (A) and inclusion body (B) were analyzed to confirm about 26 kDa and 13 kDa bands, respectively.
- BMP-10 with DsbA, which had relatively low cell growth, has low band intensity, but both (A) and (B) confirm a band having the same molecular weight as the target position
- In the case of Solubilized IB, the molecular weight is expected to differ depending on the size of the signal peptide.

Picture 2.2 SDS Page result

Western blot



1st Ab: 6X His monoclonal antibody (clontech, 631212)
2nd Ab: Anti mouse IgG -AP (Sigma, A3562)

- In the result of Western blot, it was confirmed that the target protein was hardly expressed in the periplasm.
- In the case of solubilized IB, it was confirmed that BMP-9/BMP-10 with His tag was expressed. But it was confirmed that a large amount of proteins in other forms than monomers were present.

Picture 2.3 Western Blot result

2.6 Conclusion

The target protein was hardly expressed in the periplasm. It is because protein secretion to the periplasm not always possible. Beside that periplasm proteases can cause proteolysis. The protein that expressed and secreted into periplasm will degrade by periplasm proteases that is why there was no band that showed the target protein (supernatant) in the Western Blot.

A. Suggestion for optimization the soluble form expression

a) Inducers

In addition to lowering the growth temperature, a reduction in transcription rate can also be achieved by lowering the concentration of the induction agent. Decreasing the concentration of IPTG can also enhance the production of soluble protein. When the protein was induced using 0.05 mM IPTG, the protein was soluble and active; however, when the inducer concentration was doubled to 0.1 mM the expressed protein was insoluble and inactive. The most common IPTG concentrations for protein induction range from 0.1 to 1.0 mM, a decrease to even lower levels can effect solubility (Francis and Page 2010).

b) Codon optimization

One of the most common reasons that heterologous proteins fail to express in E. coli is the presence of “rare” codons in the target mRNA. Many proteins, especially human proteins, have mRNA sequences that include codons that are infrequently used E. coli (rare codons). Target genes that contain significant numbers of individual rare codons, or smaller numbers of tandem rare codons, are more likely to experience translational stalling in E. coli, and thus often either completely fail to express, express at very low levels, or are expressed as truncated proteins. These rare codon–rich genes can also be incorrectly translated. The gene is “codon optimized,” i.e., the rare codons are replaced with those that are common to the host. This can be achieved using site-directed mutagenesis (Francis and Page 2010).

c) Strain

Several elements of the E. coli strain used for recombinant protein production have a large impact on the success of soluble expression. E. coli BL21

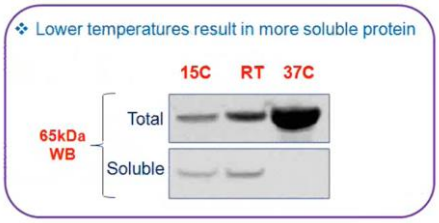
and its derivatives are most frequently used for routine protein expression. These strains are deficient of ompT and lon proteases (Francis and Page 2010).

#	Expression Strain	Rationale
1	BL21(DE3)	Work-horse of recombinant protein expression. Lamba DE3 lysogen. Basal exp high. Non-toxic preferred
2	BL21(DE3) pLysS	T7 lysozyme to repress basal level expression
3	BL21(DE3) pLysE	T7 lysozyme to repress basal level expression. Higher level of repression
4	BL21 Star (DE3)	RNaseE (rne131) mutant. Reduced mRNA degradation. Higher stability of transcripts
5	BL21 Star (DE3) pLysS	As above, coupled with higher level of repression
6	BL21 (DE3) Codon Plus RIPL	Contain extra copies of amino acids R,I,P,L tRNA genes. Supplements for codon deficiencies and bias
7	Origami™ 2(DE3) pLacI	*trxB* and *gor* reductase mutants. Enhance disulfide bonds in cytoplasm. Rare tRNA genes also
8	Origami™ 2(DE3) pLysS	*trxB* and *gor* reductase mutants. Enhance disulfide bonds in cytoplasm. Lower basal repression
9	Origami™ B(DE3) pLacI	Enhance disulfide bonds in cytoplasm and control IPTG entry into cells
10	Origami™ B(DE3) pLysS	As above, coupled with higher level of repression
11	OverExpress C41(DE3)pLysS	Uncharacterized mutation that allows higher production of "difficult" proteins
12	OverExpress C43(DE3) pLysS	Uncharacterized mutation that allows higher production of "difficult" proteins
13	BL21-Gold (DE3)pLysS	Lon and OmpT protease deficient for greater recombinant protein stability. Higher plasmid stability
14	SHuffle™ T7 Express lysY	Enhances disulfide bond formation in cytoplasm, DsbC promotes correction of misoxidized proteins
15	Arctic Express™	Co-expression with chaperonins Cpn60 & Cpn10 for increased yield of soluble proteins
16	Rosetta™ 2(DE3)pLysS	tRNA genes for 7 rare amino acids plus reduced basal expression
17	Rosetta-gami™ B(DE3)pLysS	Above plus increased disulfide bond formation in cytoplasm plus reduced basal expression
18	Tuner™(DE3)pLacI	Lac permease *lacY* mutant allows uniform entry of IPTG into cells, enhances solubility
19	Tuner™(DE3)pLysS	Above function plus reduced basal expression
20	Lemo21(DE3)	Tunable expression by varying lysozyme concentration, regulated by addition of Rhamnose

Picture 2.4 *Escherichia coli* Strain

d) Temperature

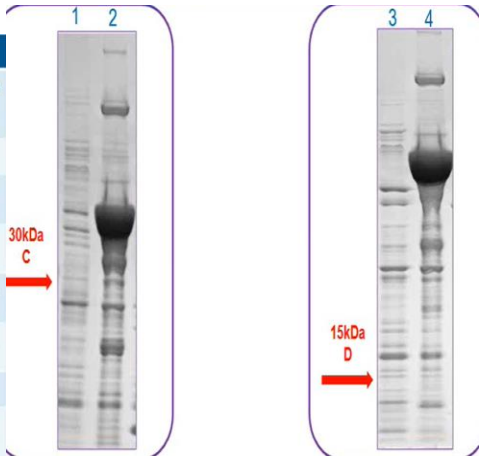
Based on this picture that, one of research had been done in different temperature. It shows that lower temperature result more soluble protein. Lowering the expression temperature routinely improves the solubility of recombinantly expressed proteins. At lower temperatures, cell processes slow down, leading to reduced rates of transcription, translation, and cell division, while also leading to reduced protein aggregation. Moreover, most proteases are less active at lower temperatures, and thus lowering the expression temperature also results in a reduction in the degradation of proteolytically sensitive proteins. The bacterial culture should be cultivated at 37ºC until mid-to-late log phase. The culture is then induced and the recombinant protein synthesized between 15°C to 25°C (Francis and Page 2010).



Picture 2.5 Optimal Temperature

e) Fusion

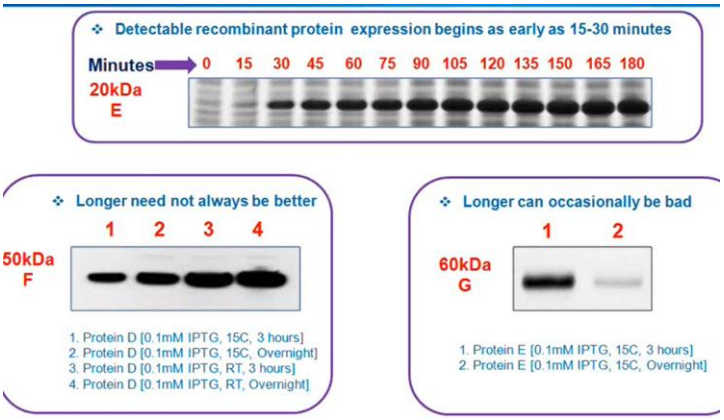
Tag	~MW [kDa]	Pros	Cons
6His	≤1	Universal, small, works with native and denaturing conditions	Usually requires more than 1-step elution. Not entirely innocuous
CBP	4	High specificity	Purity differs greatly
Flag	1	High purity Introduces Enterokinase	Expensive and elution can be a problem occasionally
GST	26	Elution easy, functions as solubilizing tag	Dimerizing nature Leaches occasionally
MBP	40	Solubility enhancer	Need longer contact times, Relatively large tag
Strep	≤1	High specificity, Mild elution conditions	Need longer contact times
NusA	55	Solubility enhancer	Very large tag
Trx	12	Solubility enhancer	Does not work well with larger MW target proteins



Picture 2.6 fusion tags

Fusion tags are proteins or peptides that are genetically fused to the target protein. They are useful because they can improve protein expression, promote folding, increase protein solubility, and facilitate downstream processes such as purification and detection. It is often necessary to test multiple fusion tags to determine which tag results in the highest yields of soluble protein and to also use a combination of tags in order to facilitate both expression and purification. Additionally, the presence of a fusion tag may interfere with the biological activity of the recombinantly expressed protein, and thus, in these cases, it may be important to enzymatically remove the tag after the fusion protein has been purified (Francis and Page 2010).

f) Induction time



Based on this trial (figure on the side)

1. Longer induction time is not always good
2. In some cases it's become important to determine the best induction duration for our target protein
3. Longer induction can result in solubility issues on the target protein

So do the optimization!!!

Picture 2.7 Optimal Induction Time

B. ENHANCING SOLUBILITY BY COEXPRESSION WITH OTHER PROTEINS

Some proteins only express solubly when they are co-expressed with additional biomolecules, typically other proteins, like a binding partner, or molecular chaperones. Larger bacterial and heterologous proteins fold more slowly, and thus require protein chaperones and folding catalysts to prevent aggregation and facilitate folding in E. coli. While molecular chaperones are ubiquitous in the cell, they are rapidly titrated during overexpression of recombinant protein. DnaK-DnaJ-GrpE and GroEL-GroES. The DnaK-DnaJ-GrpE and GroEL-GroES systems are the most extensively characterized molecular chaperones in E. coli. Both groups are folding chaperones that bind to solventexposed hydrophobic domains, preventing nascent peptide aggregation, and through iterative rounds of ATP-driven conformational changes, mediate the folding/unfolding of their substrates (for an extensive review of cytosolic molecular chaperones. Soluble expression of certain proteins is improved by coexpression with DnaK-DnaJ-GrpE, other proteins result in higher yields of soluble protein when coexpressed with GroEL-GroES. molecular chaperone system will improve protein solubility (Francis and Page 2010)

References

1. Bonten, M., et al. (2021). "Epidemiology of Escherichia coli bacteremia: a systematic literature review." Clinical Infectious Diseases **72**(7): 1211-1219.

2. Choi, J. and S. Lee (2004). "Secretory and extracellular production of recombinant proteins using Escherichia coli." Applied microbiology and biotechnology **64**: 625-635.

3. Francis, D. M. and R. Page (2010). "Strategies to optimize protein expression in E. coli." Current protocols in protein science **61**(1): 5.24. 21-25.24. 29.

4. Karyolaimos, A., et al. (2019). "Enhancing recombinant protein yields in the E. coli periplasm by combining signal peptide and production rate screening." Frontiers in microbiology **10**: 1511.

5. Kunze, M. and J. Berger (2015). "The similarity between N-terminal targeting signals for protein import into different organelles and its evolutionary relevance." Frontiers in physiology **6**: 259.

6. Martínez, V. G., et al. (2015). "The BMP pathway participates in human naive CD4+ T cell activation and homeostasis." PloS one **10**(6): e0131453.

7. Mirzadeh, K., et al. (2020). "Increased production of periplasmic proteins in Escherichia coli by directed evolution of the translation initiation region." Microbial cell factories **19**: 1-12.

8. Miyazono, K., et al. (2010). "Bone morphogenetic protein receptors and signal transduction." Journal of biochemistry **147**(1): 35-51.

9. Owji, H., et al. (2018). "A comprehensive review of signal peptides: Structure, roles, and applications." European journal of cell biology **97**(6): 422-441.

10. Santos, B. D., et al. (2019). "Optimizing periplasmic expression in Escherichia coli for the production of recombinant proteins tagged with the small metal-binding protein SmbP." Molecular biotechnology **61**(6): 451-460.

11. Shepard, B. D., et al. (2013). "A cleavable N-terminal signal peptide promotes widespread olfactory receptor surface expression in HEK293T cells." PloS one **8**(7): e68758.

12. Zhang, M., et al. (2024). "N-Terminal Sequences of Signal Peptides Assuming Critical Roles in Expression of Heterologous Proteins in Bacillus subtilis." Microorganisms **12**(7): 1275.