Optimization of a Bacillus sp signal peptide for improved recombinant protein secretion and cell viability in *Escherichia coli*

Is there an optimal signal peptide design?

Kheng Oon Law, Mohd. Anuar Jonet, Noor Faizah Ismail and Rosli Md. Illias*

Department of Bioprocess Engineering; Faculty of Chemical Engineering; Universiti Teknologi Malaysia; Skudai, Johor Malaysia

Keywords: signal peptide, protein secretion, cell lysis, *Escherichia coli*, helix-breaking residue

Recombinant protein production in microbial host cell is an important aspect in modern biotechnology industries. While cytoplasmic expression is widely accepted and practiced, recent advancements in protein secretion understanding and technology have gained much interest among the scientific and industrial communities. Secretory production of recombinant protein provides several advantages over cytoplasmic protein production including N-terminal authenticity, increased protein stability and solubility, ensure correct formation of disulfide bonds, prevents inclusion body formation and facilitates downstream processing.

Bacillus species are well known for their significant capacity in terms of protein secretion. However, heterologous protein production in Bacillus expression systems often suffers from low expression and proteolytic degradation. An alternative approach is the use of *Escherichia coli* expression systems. Despite the ability to overexpress heterologous protein in the cytoplasm, recent findings have shown that *E. coli* is equally capable of high level secretion of target proteins. Various strategies have been devised to achieve or improve protein secretion in *E. coli* and one such approach is the selection and modification of signal peptides. Heterologous proteins can be targeted to the periplasm by fusion to an N-terminal signal peptide and eventually to the extracellular medium under specific circumstances. However, studies have shown that extracellular secretion in *E. coli* may cause cell lysis due to overproduction of recombinant protein. Thus, it becomes a challenge to achieve high level secretory production of recombinant proteins concomitant with a minimum occurrence of cell lysis.

To date, numerous studies have been undertaken to investigate the effect and influence of signal peptides in protein secretion. Although the primary amino acid sequence of signal peptides are not conserved, they share certain common features. Generally, three distinct regions can be found in a typical signal peptide: (1) an N-terminal region (n-region) that contains a number of positively charged amino acids (lysinas and arginines); (2) a central hydrophobic core region (h-region) and (3) a hydrophilic cleavage region (c-region) that contains a sequence motif recognized and cleaved by signal peptidases. Amino acid substitutions in the three functional domains of signal peptide have been an active area of research to elucidate the role of the regions in protein secretion. However, much less is known about the effects of mutations of heterologous signal peptides, particularly from Bacillus sp, on protein secretion in *E. coli*.

In a recent paper, we observed that a signal peptide from Bacillus sp G1 (G1 signal peptide) directs the secretion of cyclodextrin glucanotransferase (CGTase) to the periplasm and eventually, the extracellular space of *E. coli*. Optimization through mutations of G1 signal peptide and addition of chemical additives leads to enhancement in heterologous protein secretion and host cell viability. Our work shows that the G1 signal peptide...
can work well in an E. coli secretion system. This is exemplified by the fact that the G1 signal peptide contributes to increased target protein solubility and stability, enhanced cell viability and has a similar capacity for extracellular CGTase secretion levels, as compared with the widely used PelB signal peptide. The advantage of the G1 signal peptide is clearly shown in Figure 1. SDS-PAGE analysis of the extracellular fraction of the recombinant E. coli cells reveals that a highly specific secretion level of target protein can be achieved by G1 signal peptide (Fig. 1, lane 4). In contrast, the PelB signal peptide resulted in a protein sample contaminated with host native proteins (Fig. 1, lane 5), probably due to cell lysis. A high specific secretion level of target protein is highly desired as it greatly simplifies downstream purification, and ensures high target protein concentration.

Culture conditions are one of the main factors affecting protein expression, translocation and stability and host cell metabolism. In our work, the influence of inducer (IPTG) concentration and post-induction temperature on recombinant CGTase secretion was investigated. As shown in Figure 2A, it seems that the recombinant CGTase translocation (total CGTase activity in periplasm and supernatant) was more favorable at lower levels of IPTG induction. Total CGTase translocation under the induction of 0.01 mM IPTG was 1.3-fold (185.5 U/ml) higher than total CGTase translocation under the IPTG induction of 2.0 mM (80.66 U/ml). A similar observation was found in the secretory production of alkaline phosphatase in E. coli. Lower levels of IPTG induction resulted in an increase of alkaline phosphatase solubility without reduction of expression levels and gave rise to about 90% recovery level of the enzyme in the periplasm. However, a too low IPTG induction (0.005 mM), in our case, resulted in a reduced level of total CGTase translocation (total CGTase activity is 111.98 U/ml), most probably due to a lower protein expression level. Interestingly, a similar secretion pattern was also observed when recombinant CGTase was secreted into the culture medium using the Type I hemolysin transport system in E. coli. Whether this is a common feature for CGTase secretion under the control of T7lac promoter system still needs further clarification.

β-galactosidase is a cytoplasmic enzyme and detection of it in the extracellular medium suggest the occurrence of cell membrane leakage or cell lysis. Surprisingly, the β-galactosidase activity increases with the concentration of IPTG, and vice versa. Based on the observation, a correlation between target protein secretion and cell membrane permeability/cell lysis can be made: a higher recombinant CGTase secretion level, due to reduced IPTG concentration, resulted in a lower cell membrane permeability/cell lysis level. The reason for this observation can be explained from the level of inclusion body formation in the recombinant cells. Villa et al. elegantly demonstrated that the formation of inclusion bodies in the cytoplasm significantly affects membrane protein expression and this increases the cells sensitivity/permeability toward certain nutrients or chemicals. In our work, increased CGTase secretion (by lowering IPTG induction) caused a lower concentration of the recombinant protein in the cytoplasm and inclusion body formation. This induces the cell to express its membrane proteins that in turn strengthen the cell envelope. The hypothesis can be supported by the increased level of extracellular β-galactosidase at elevated temperatures (Fig. 2B). It is known that overexpression of recombinant protein at a higher temperature (i.e., 37°C) can often result in the formation of inclusion bodies than at a lower temperature (i.e., 30/25°C). Therefore, expression of recombinant CGTase at an elevated temperature (from 25–37°C) can result in the increase of inclusion body deposition, which affects the sensitivity/permeability of cell membrane, and eventually, resulted in the leakage or release of cytoplasmic proteins (i.e., β-galactosidase) into the culture medium. Our observations revealed that the integrity of cell membrane and cell viability is closely related to the effectiveness of target protein translocation and inclusion body formation.

It is generally believed that recombinant proteins exported to the periplasm through a Sec-dependent system can be “leaked/released” into the extracellular medium through an unspecified route/pathway. Under the assumption that β-galactosidase and recombinant CGTase traverse through the outer membrane by a similar, unspecific manner, it is interesting to find that their secretion level is inversely proportional to each other, at varied IPTG concentration and post-induction temperature (Fig. 2). In other words, a higher extracellular CGTase activity is accompanied by a lower extracellular β-galactosidase activity and vice versa. Intriguingly, β-galactosidase is a relatively large protein (homo-oligomer with a monomer sequence length of 1,024 residues) compared with CGTase (mature protein sequence length is 674 residues) and any “channel” on the outer membrane that is big enough for β-galactosidase should be equally, if not more, adequate for CGTase to exit. It should be noted that the detection of extracellular β-galactosidase activity suggests the
occurrence of cell lysis and could lead to the "release" of cytoplasmic or periplasmic recombinant protein into the extracellular medium. Indeed, secretory production of recombinant proteins in _E. coli_ often results in a higher level of cell lysis3,11 and can be seen as an obstacle for operation in a large-scale bioreactor. However, we believe that this is not the case for GI signal peptide since it leads to a high specific secretion level of target protein (Fig. 1). We speculate that a lower secretion level of CGTase resulted in an accumulation of intracellular, secretion-competent CGTase species which eventually saturate and overwhelm the Sec translocon. This in turn weakens the integrity of inner membrane and cause rupture and, eventually, cell lysis. A higher secretion level of target protein, however, lowers the concentration of intracellular recombinant protein and minimizes the pressure exerted onto the inner membrane, resulting in a lower level of β-galactosidase activity or cell lysis.

Signal peptide sequence is probably the most important deciding factor for efficient target protein translocation and secretion. With this in mind, modification of the signal peptide is a feasible approach to improve recombinant protein secretion. It is known that signal peptides can differ significantly in sequence and length, with little to no detectable sequence homology. However, signal peptides which are recognized and translocated by the same secretion pathway, in a given organism, should have a minimum level of divergence. Therefore, we believe that there exists a common/universal signature and design of signal peptide for a given organism. The knowledge of this universal design would greatly aid us in the rational design of an optimal signal peptide: With this in mind, a list of signal peptide sequences, derived from _E. coli_ and _B. subtilis_, were collected from Signal Peptide Website (www.signalpeptide.de). The sequences were categorized according to its functional groups (n-, h- and c-region), aligned and presented in the form of sequence logos (Fig. 3). While the result supports the lack of sequence homology between signal peptides, it clearly shows the preference of certain group of residues in specific positions, notably the "MKK" sequence at n-region. However, the introduction of the "MKK" mutation in GI signal peptide failed to improve CGTase secretion.6 Mutation N2K resulted in a reduced periplasmic and extracellular CGTase activity to about 33 and 42% of wild-type level secretion, respectively. Interestingly, further mutation on D3K (mutant MKK carrying double mutation N2K/D3K) restored the total CGTase translocation to wild-type secretion.
level. While positively charged residues are generally found in the n-region of signal peptide, we speculate that the presence of “MKK” sequence is an important element for efficient protein translocation.

Another observation is the presence of helix-breaking residues (i.e., glycine) located at the h-region for both the E. coli and B. subtilis data set. It seems that B. subtilis signal peptide has a higher tendency to possess a helix-breaking residue at the h-region than its E. coli counterpart. It has been shown that helix-breaking residues at h-regions is important for the recognition by signal peptide peptidase (SPPase). Lack of a helix-breaking residue at the h-region prevents its processing by SPPase, which can cause membrane rupture and cell death. Surprisingly, GI signal sequence lacks a helix-breaking residue at the h-region. Substitution with glycine residue at the center of the GI signal peptide h-region resulted in a varied outcome. Mutation S14G (mutant M7) showed a decrease in extracellular CGTase activity (65% of wild-type level) and a wild-type level of periplasmic CGTase activity compared with the wild-type signal peptide; while mutation I16G (mutant M5) resulted in an increase in both the periplasmic (110%) and extracellular CGTase activity (94%) compared with the wild-type signal peptide. The result indicates that both the presence and location of a helix-breaking residue (glycine) in the h-region is an essential feature which increases the preferability of the sec pathway machineries toward signal peptide and eventually ensures efficient protein translocation.

In this work, we showed that Bacillus species signal peptide, particularly GI signal peptide, can work as effectively as, if not better than, its gram negative counterparts (PelB signal peptide) in recombinant protein secretion in E. coli. Importantly, GI signal peptide enables a highly specific target protein secretion while maintaining and improving cell viability. Further optimization of the GI signal peptide by rational design suggests that the “MKK” signature at the n-region and a glycine residue (helix-breaking residue) at the center of the h-region are essential elements to ensure high recombinant protein secretion in E. coli.

This work should aid in the design of an optimal synthetic signal peptide for secretary production of recombinant proteins.

Disclosure of Potential Conflicts of Interest
No potential conflicts of interest were disclosed.

Acknowledgments
This project was supported by the Genomics and Molecular Biology Initiatives Programme of the Malaysia Genome Institute, Ministry of Science, Technology and Innovation Malaysia (Project No. 07-05-MGI-GMB011). K.O.L. is a researcher of Universiti Teknologi Malaysia under the Post-Doctoral Fellowship Scheme.

References
1. Yoon SH, Kim SK, Kim JI. Secretory production of recombinant protein in Escherichia coli. Recent Rev Biotechnol 2010; 4:235-50, PMID:20201800, http://dx.doi.org/10.2174/187220810790069550.
2. Mergulhão FJM, Summers DK, Malmgren G. Recombinant protein secretion in Escherichia coli. Biotechnol Lett 2005; 27:217-22, PMID:15763404; http://dx.doi.org/10.1007/s10529-004-0077-3.
3. Choi JH, Lee SY. Secretary and extracellular production of recombinant proteins using Escherichia coli. Appl Microbiol Biotechnol 2008; 79:809-16, PMID:18346460, http://dx.doi.org/10.1007/s00253-006-0877-5.
4. Yu T, Chen R. Extracellular recombinant protein production from Escherichia coli. Biotechnol Lett 2009; 31:1661-7, PMID:19577705, http://dx.doi.org/10.1007/s10529-008-0189-2.
5. Fu ZB, Ng KL, Lee SY, Wong WK. Cell death caused by hyper-expression of a secretory exoglucanase in Escherichia coli. Protein Expr Purif 2005; 42:627-77, PMID:15882948, http://dx.doi.org/10.1016/j.pep.2005.03.029.
6. van Houte G, Abadie E. Specific-secretion initiation in signal peptide design. Implications for protein secretion in foreign hosts. FEBS Lett 1989; 246:63-6, PMID:2461715, http://dx.doi.org/10.1016/0014-5793(89)80557-8.
7. Ismail NH, Hamdani SN, Muhaidi NM, Majid AM, Rabu A, Bakar FD, et al. A recombinant L-asparaginase II signal peptide improves the secretion of recombinant cyclodextrin glucanotransferase and the viability of Escherichia coli. Biotechnol Lett 2011; 33:999-1005; PMID:21254793. http://dx.doi.org/10.1007/s10522-011-0937-8.

8. Jenar MA, Muhaidi NM, Murad AM, Rabu A, Bakar FD, Rahim RA, et al. Optimization of a heterologous signal peptide by microfluidic microarray for improved secretion of recombinant proteins in Escherichia coli. J Mol Microbiol Biotechnol 2011; 22:45-56. PMID:22454497. http://dx.doi.org/10.1515/1562-7267-2011-22-45.

9. Choi HJ, Hong KJ, Kim SC, Lee SY. Efficient recovery production of alkaline phosphatase by high cell density culture of recombinant Escherichia coli using the Bacillus sp. endoxylanase signal sequence. Appl Microbiol Biotechnol 2006; 73:609-15. PMID:16813199. http://dx.doi.org/10.1007/s00253-006-0338-4.

10. Leow RD, Malaki NM, Rahim RA, Rabu A, Abu Bakar FD, Murad AM, et al. An effective extracellular protein secretion by an ABC transporter system in Escherichia coli: statistical modeling and optimization of cyclodextrin glucanotransferase secretion production. J Ind Microbiol Biotechnol 2011; 38:1987-97. PMID:21394878. http://dx.doi.org/10.1007/s10295-011-0941-9.

11. Villa R, Lepo-M, Ganto-Lafontant P. Components of the E. coli envelope are affected by and can react to protein over-production in the cytoplasm. Microbiol Cell Fact 2009; 8:32; PMID:19500339; http://dx.doi.org/10.1186/1475-2859-8-32.

12. Liu Z, Gu Z, Wang M, Da G, Wu J, Chen J. Delaxed supplementation of glucose enhances extracellular secretion of the recombinant alpha-cyclodextrin glucanotransferase in Escherichia coli. Appl Microbiol Biotechnol 2010; 85:553-61. PMID:19551817. http://dx.doi.org/10.1007/s00253-009-2157-7.

13. Lemberg MK, Bland FA, Wielofen A, Brand VM, Marzoglio B. Intramembrane proteolysis of signal peptides: an essential step in the generation of HLA-E epitopes. J Immunol 2001; 167:6441-6; PMID:11714810.

14. Lemberg MK, Marzoglio B. Requirements for signal peptide peptidase-catalyzed intramembrane proteolysis. Mol Cell 2002; 10:735-44; PMID:12419218; http://dx.doi.org/10.1016/S1097-2765(02)00655-X.

15. Zwizinski C, Wickner W. Purification and characterization of leader (signal) peptidase from Escherichia coli. J Biol Chem 1980; 255:7973-7; PMID:6995457.

16. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 2004; 32:1792-7. PMID:15034147; http://dx.doi.org/10.1093/nar/gkh340.

17. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: a sequence logo generator. Genome Res 2004; 14:1188-90. PMID:15173120; http://dx.doi.org/10.1101/gr.849004.