Korean J. Chem. Eng., **28**(8), 1744-1748 (2011)
DOI: 10.1007/s11814-011-0140-3

Expression of redesigned mussel silk-like protein in *Escherichia coli*

Yun Jung Yang*, Yoo Seong Choi*, Dooyup Jung*, and Hyung Joon Cha*,,†**

*Department of Chemical Engineering and **Ocean Science and Technology Institute,
Pohang University of Science and Technology, Pohang 790-784, Korea
(*Received 12 May 2011 • accepted 31 May 2011*)

Abstract−Silks have been used widely for human beings due to their several extraordinary properties. Until now, the studies on silk proteins have mainly focused on spiders and silkworms. Because silk properties are organism-dependent, novel silk protein types can be found and developed through investigation of new silk-bearing organisms. We noticed that marine mussel has silk-like domains containing many repeats with abundance of glycine and alanine. In the present work, we redesigned mussel-derived silk-like gene sequence which contains alternating repeated and non-repeated regions with optimized codons for *Escherichia coli*. For successful expression of recombinant mussel silk-like protein in *E. coli* cells, we employed several experimental strategies, including use of strong promoter, cold shock expression, and genetic fusions. We observed significant repression on cell growths by even low expression levels of soluble mussel silk-like proteins in cold shock- and glutathione s-transferase (GST) fusion-based systems. Thus, we finally used baculoviral polyhedrin protein as a fusion partner and successfully expressed insoluble mussel silk-like protein with relatively high expression level and without cell growth repression in *E. coli*.

Key words: Silk, Mussel, Genetic Expression, *Escherichia coli*

INTRODUCTION

Silk, an outstanding nature-given material, has been used widely for its fine texture because it rarely causes inflammation and allergic responses. In addition, it has been important for biomaterials due to its biocompatibility and biodegradability as well as its stiffness, strength, extensibility, and toughness [1-3]. The properties of silk proteins are determined by the organisms producing them [4]; each silk-bearing organism, such as spiders, silkworms, and bees, has its unique silk protein property [5,6]. Therefore, it is necessary to find new silk-bearing organisms with characteristic properties of silk proteins. However, efforts to explore new silk-bearing organisms have not yet produced sufficient outcomes.

Byssus, a silky filament of marine mussel has higher resiliency (3-5 times) and density than tendon [7] and is stiffer and stronger than elastin (a typical elastic mammalian protein) and resilin (one of elastic insect proteins) [7,8]. This superiority of mussel byssus might be due to combinational structures. Mussel byssus is composed of three types of proteins: pre-collagen D (distal), pre-collagen P (proximal), and pre-collagen NG (non-gradients) [9-11]. From proximal to distal regions, each collagenous protein makes block copolymer gradually. We noticed that pre-collagen D has a silk-like domain containing rich-glycine and rich-alanine repeats which also exist in silk proteins from spiders and silkworms [11-13].

In the present work, we, for the first time, selected marine mussel as a target organism to develop a new silk protein. We designed a new type of mussel silk-like protein from the original domain and utilized several experimental strategies for successful expression of recombinant mussel silk-like protein in *Escherichia coli*.

†To whom correspondence should be addressed.
E-mail: hjcha@postech.ac.kr

MATERIALS AND METHODS

1. Gene Synthesis and Plasmid Construction

Designed *msp* gene (1,353 bp) encoding mussel silk-like protein was chemically synthesized (Genscript) to include *Nhe*I and *Xho*I restriction endonuclease sites at each terminus. The *msp* gene was cloned into pET23b+ vector (Novagen) and pCOLD DNA (Takara Bio) to construct the recombinant plasmids pET23-MSP and pCOLD-MSP, respectively (Fig. 1). To construct a fusion system, *polh* gene encoding baculoviral polyhedrin was amplified by polymerase chain reaction (PCR) and cloned into the pET23-MSP to make the recombinant plasmid pET23-Polh-MSP (Fig. 1). We also constructed GST fusion plasmid pGEX-MSP by simple insertion of *msp* gene into pGEX-4T-1 (GE Healthcare) (Fig. 1).

2. Culture Condition

Each recombinant plasmid carrying mussel silk-like gene was transformed into *E. coli* BL21 (DE3) (Novagen). The recombinant cells harboring pET23-MSP or pET23-Polh-MSP were grown at 37 °C or 25 °C in 400 mL LB medium containing 50 μg/mL ampicillin. The recombinant cells harboring pCOLD-MSP and pGEX-MSP were grown at 15 °C and 28 °C, respectively, after induction for optimal expression. When each culture was grown to the cell density of 0.8 (at 600 nm, OD_{600}), 0.1 mM isopropyl-β-D-thiogalactopyranoside (IPTG; Sigma) was added for induction of recombinant protein expression. Induction under the pCOLD DNA-based system is practicable with facile low-temperature transition, but IPTG was also used for more tight regulation.

3. Analytical Assays

Cell density (OD_{600}) was measured at 600 nm on a UV/VIS spectrophotometer (Shimadzu). After culturing upon IPTG induction, cells were harvested and disrupted by sonication at 4 °C. Three types of samples (whole cells, soluble supernatant, and insoluble cell lysate)

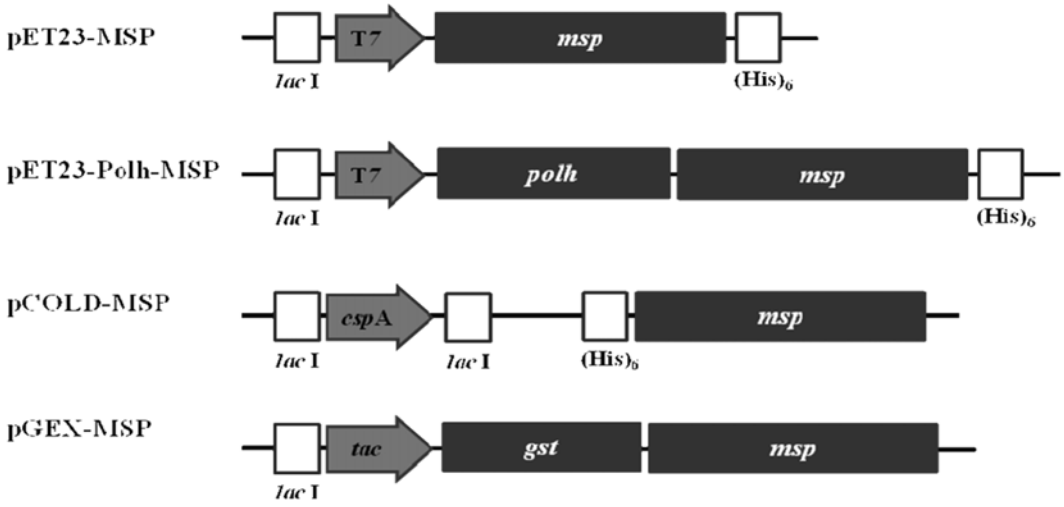

Fig. 1. Gene maps of recombinant plasmid constructs harboring mussel silk-like protein. Abbreviations: *msp*, mussel silk-like protein gene; *polh*, polyhedrin gene; *gst*, GST gene; $(His)_6$, hexahistidine tag; T7, T7 promoter; *csp*A, *csp*A promoter; *tac*, *tac* promoter; *lac*I, *lac*I repressor.

were resuspended in 100 μL of protein sample buffer (0.5 M Tris-HCl, pH 6.8, 10% glycerol, 5% sodium dodecyl sulfate [SDS], 5% β-mercaptoethanol, and 0.25% bromophenol blue) and heated at 95 °C for 5 min. 12% SDS-polyacrylamide discontinuous gel was used for electrophoresis. The gels were visualized by Coomassie blue staining (Bio-Rad). For Western blot analysis, proteins were transferred onto a nitrocellulose membrane (Amersham International) using a semi-dry blotting system (Amersham Pharmacia) and Bjerrum and Schafer-Nielsen transfer buffer (48 mM Tris, 39 mM glycine, 20% methanol) for 1 h at 50 mA and 50 V. Recombinant mussel silk-like protein containing hexahistidine (His_6) tag was detected using monoclonal anti-His_6 antibody (1 : 1,000 v/v) (R&D Systems) and alkaline phosphatase-conjugated anti-mouse IgG (1 : 1,000 v/v) (Sigma) diluted in antibody buffer (0.5% [v/v] Tween 20, pH 7.4, Tris-buffered saline with 1% [w/v] nonfat dry milk). The membrane was then washed and developed colorimetrically with FAST Red TR/Naphtol AS-MX (Sigma). The membrane was scanned and its image was analyzed using Gel-Pro analyzer software (Media Cybernetics).

RESULTS AND DISCUSSION

1. Design of Mussel Silk-like Protein for Expression in *E. coli*

Mussel as a candidate for silk-bearing organism contains three types of collagenous proteins, and one type has silk-like domains containing rich-glycine and rich-alanine repeats, which also exist in spider and silkworm silks [11,12]. Thus, it is predictable that the property of mussel byssus might be similar to that of silk protein. In fact, mussel byssus is stiffer and tougher than the other elastic proteins [8]. The redesigned silk-like protein used in the present work originates from mussel *Mytilus californianus* pre-collagen D protein (86 kDa), which has several types of domains including collagen, histidine-rich domains ($[AH]_8$), acidic part, and silk-like domains. The silk-like domains have sequences of rich-alanine, $(AAA)_n$, and rich-glycine, $(GGX)_n$ (where X is glutamine, serine, leucine, valine, proline, tyrosine, or arginine) [9-11,14].

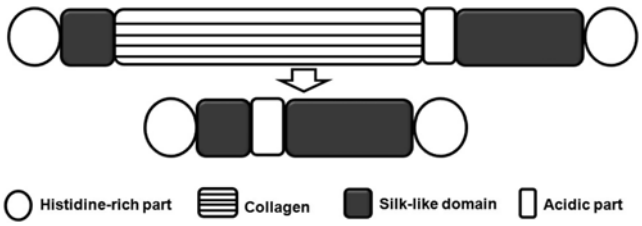

Fig. 2. Schematic diagram for redesign of mussel silk-like protein.

To use two silk-like domains mainly and to combine cross-linking (histidine-rich) and hydrophobic (silk-like domains and acidic part) regions in turn as other silk proteins [15], we eliminated lengthy collagen domain. By deleting the collagen part, a histidine-rich region is located at each terminus, and silk-like domain and acidic part are able to be placed in the middle (Fig. 2). The mussel silk-like protein (447 amino acids and 1,353 base pairs with high GC contents) has about 37 kDa of molecular weight and 10.8 of isoelectric point (pI) (Table 1). From the observation of sequential similarity between two silk-like domains and other silk proteins through FASTA local alignment, we found that the mussel silk-like domain has the highest similarity with one silk type of spider *Nephila clavipes*: the first and second silk-like domains have 78.3% and 71.3% of similarities with minor ampullate, respectively. Through comparison of amino acid compositions, we found that more than 60% of total amino acids are glycine and alanine in the mussel silk-like protein (Table 2), while silkworm (*Bombyx mori*) fibroin and spider (*N. clavipes*) major ampullate have ~73% and ~58% of glycine and alanine contents, respectively [13,16]. Collectively, we knew that the sequence arrangement and composition of mussel silk-like protein are analogous to silks from spider and silkworm in nature.

The gene encoding mussel silk-like protein has highly repeated sequences which might be easily deleted and disturbed in *E. coli* cells, and codon usage of mussel is totally different with *E. coli*. Thus, we designed gene sequences to have optimized *E. coli* codons, minimize same codon usages, and considered mRNA secondary struc-

Table 1. Redesigned mussel silk-like protein with its genetic information

	Histidine-rich part	Silk-like domain 1	Acidic part	Silk-like domain 2	Histidine-rich part	Whole protein
Base pair	204	285	63	726	63	1,353
MW (kDa)	6.77	7.66	2.23	18.58	2.08	37.32
pI	10.36	12.48	3.83	12.48	6.86	10.8
Amino acid	DYGRKYG	AGSASAAARAA	TLVIED	GAGPGGAGPFGGAGPFGGAGAGAGAG	SSVIHG	
	KPSYGEYG	ARAGAVGGFGG	LRTAG	GAGPFGAGPFGAGPFGGVGGAGVGGAG	GGHGG	
	GKRGGGR	GGFGSASAAAA	VESPV	PGGAAGPGGAAGPGGAAGAGGLGGLG	HGGDY	
	VSGAVAHA	ARAAAAGGGL	EAFDA	AGGLGGLGAGGLGGLGAGGLGGQGAG	HKPGY	
	HAHAHASS	GGGFGSASAAA		GLGGLGAGGLGGLGAGLGGGLGGGAG		
	GADGRSR	RAAAAAGGFGG		AAAAAQAAAAANGGLGGGSAAAAARA		
	AHARAVAH	GGFGSASANAA		AAAANAGLGGGAVAAAQAAASAAANS		
	SHSGGGAA	ANAAAGGFGGF		GLGAGAARAAASAAARATVAGAGRGTA		
	HGHPGFP	GGGFGPGF		AAAASAAAQAHAATKAQGGSHAHAAA AAQAAA		

Table 2. Comparison on amino acid distributions of silk proteins from silkworm, spider, and mussel

Amino acid	Silkworm (*Bombyx mori*) fibroin	Spider (*Nephila clavipes*) major ampullate	Mussel (*Mytilus californianus*) silk-like protein
Gly	44.1	37.1	33.3
Ala	29.7	21.2	32.0
Ser	12.4	4.5	5.4
Tyr, Phe	7.5	10.2	4.9
Leu, Ile, Val, Asx, Glx	3.6	11.7	10.7
Thr	1.2	1.7	1.1
Arg	1.5	7.6	3.4
Trp	0.5	-	-
Pro	-	4.5	3.4
His, Cys, Lys	-	1.0	5.1

ture [17]. Finally, the chemically synthesized redesigned mussel silk-like protein gene was cloned into three types of expression vector systems to be expressed (Fig. 1).

2. Expression of Recombinant Mussel Silk-like Protein

First, we used pET23b+ vector system having strong T7 promoter to express recombinant mussel silk-like protein. Through cultures at 37 ºC, we found that the growth of recombinant *E. coli* harboring the pET23-MSP plasmid (triangle) was rather higher than that of the cells harboring the parent pET23b+ vector (circle) (Fig. 3(a)). But, unfortunately, recombinant mussel silk-like protein was not shown on both SDS-PAGE and Western blot analyses (data not shown). It is well known that genetic expressions of silk proteins are generally tricky works [18,19]. Therefore, other experimental strategies were needed for successful expression of recombinant mussel silk-like protein.

When we reduced culture temperature of recombinant *E. coli* harboring the pET23-MSP plasmid to 25 ºC, we found mussel silk-like protein band (~37 kDa) on Western blot membrane even though its band intensity was dim (Fig. 4(a)). However, due to very low expression level, the protein band was not shown in SDS-PAGE analysis (data not shown).

When *E. coli* cells are induced under low temperature environment (15-25 ºC), there are numerous physiological changes in cells, such as decrease of membrane fluidity, decrease in the rates of transportation and secretion, and alternations on mRNA structure, replication, transcription, and translation [20-22]. Low culture temperature can also help target proteins to be expressed as soluble forms [20,23,24]. Thus, to trigger cold shock expression of mussel silk-like protein in *E. coli*, we next used pCOLD DNA vector system. It was reported that expression of heterologous protein that failed in other vector systems became possible with increased expression level and soluble fraction of the protein using pCOLD DNA vector system [20,24]. From the culture of *E. coli* having the recombinant pCOLD-MSP plasmid at 15 ºC, we found that mussel silk-like protein (~37 kDa) was expressed as a soluble form in Western blot analysis (Fig. 4(b)). But, due to low expression, band detection was also impossible on SDS-PAGE gel (data not shown). We also observed significant repression on cell growth (triangle) compared to the cells having the parent vector (circle) (Fig. 3(b)). We suspect that soluble expression of mussel silk-like protein might be a burden to the cells, because insoluble expression using the pET23-MSP plasmid did not repress cell growth (recall Fig. 3(a)).

Fusion protein strategy has solved many protein expression problems through possible changes in structures and functions of target heterologous proteins by forming fusion proteins. For successful application of fusion protein strategy, selection of appropriate fusion

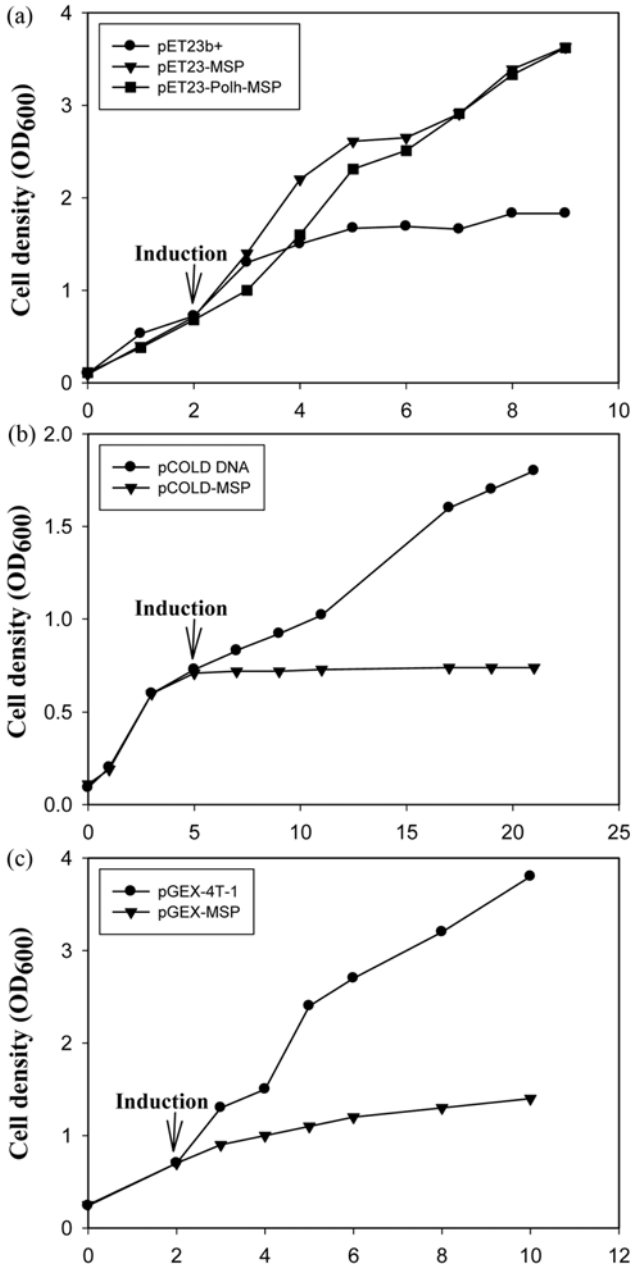

Fig. 3. Time profiles of cell growth under expression of mussel silk-like protein using (a) pET23b+ or/and polyhedrin fusion-, (b) cold shock-, and (c) GST fusion-based systems. Arrow indicates induction point.

partner is generally important [23,25]. We used glutathione s-transferase (GST) fusion system (pGEX-4T-1) because GST protein has the ability to form soluble proteins in the bacterial cytoplasm with correct folding characteristics that affect high level of expression and solubility of target protein [26,27]. However, when we cultured the recombinant *E. coli* cells having the pGEX-MSP, we again observed severe repression on cell growth (Fig. 3(c)). Thus, we confirmed that soluble expression system hampers cell growth significantly. We also failed to detect the band of GST-fused mussel silk-like protein in both SDS-PAGE and Western blot analyses (data not shown).

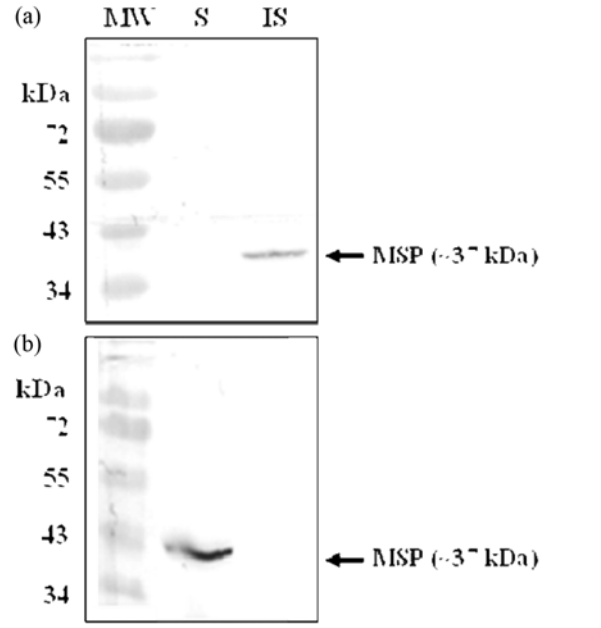

Fig. 4. Western blot analyses for expression of mussel silk-like protein using (a) pET23-MSP and (b) pCOLD-MSP plasmids. Abbreviations: MW, molecular weight marker; MSP, mussel silk-like protein; S, soluble fraction; IS, insoluble fraction.

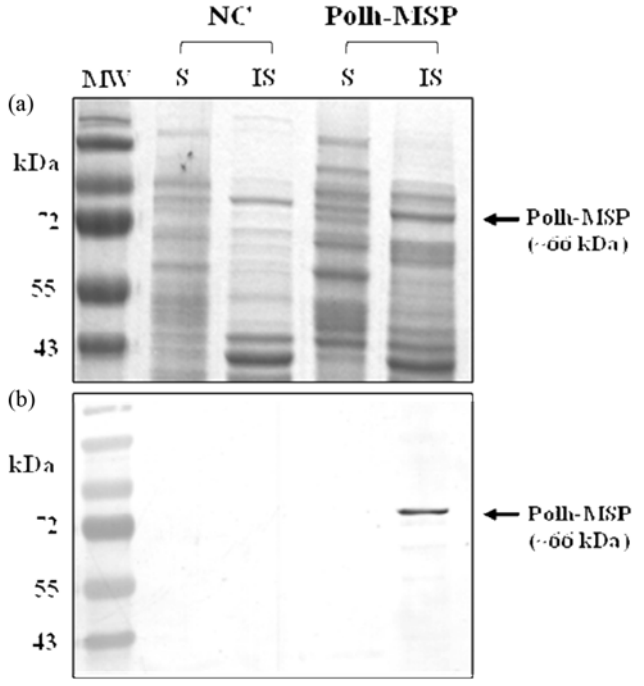

Fig. 5. (a) SDS-PAGE and (b) Western blot analyses for expression of polyhedrin-fused mussel silk-like protein. Abbreviations: MW, molecular weight marker; NC, negative control (parent vector); Polh-MSP, polyhedrin-mussel silk-like protein fusion; S, soluble fraction; IS, insoluble fraction.

Baculoviral polyhedrin is an alkali-soluble component of polyhedra which comes from *Autographa californica* nuclear polyhedrosis virus [28,29]. Fusion with polyhedrin protein induces improved

target protein expression and insoluble inclusion body formation with peculiar alkali-soluble characteristics of polyhedrin in *E. coli* [30-32]. To avoid cell growth repression problem of soluble expression system and increase protein expression level of pET23b+ vector-based system, we finally used polyhedrin fusion strategy: polyhedrin (~29 kDa) was fused at the N-terminus of mussel silk-like protein (Fig. 1). Such growth inhibition as shown in the case of pET23-MSP system was not observed during growth of the cells harboring the pET23-Polh-MSP plasmid (Fig. 3(a)). Previously, we showed that polyhedrin-fusion could solve the cell growth retardation problem [30,33]. Actually, the cells grew well with high cell density in both pET23b+ vector-based systems regardless of polyhedrin fusion (Fig. 39a)). But, there was a definite difference in protein expression between the two. When using pET23-MSP system, the expression level of mussel silk-like protein at the 37 ºC culture was too low for detection on Coomassie blue stained gel (data not shown), but the detection was only possible by sensitive Western blot analysis with very low band intensity (Fig. 4(a)). However, importantly, mussel silk-like protein fused with polyhedrin (~66 kDa) was successfully expressed in inclusion bodies and, more importantly, was detected on even SDS-PAGE gel (Fig. 5(a)) as well as Western blotted membrane in the culture of *E. coli* cells harboring pET23-Polh-MSP plasmid at 37 ºC (Fig. 5(b)). Even though insoluble inclusion body formation requires extra steps for protein function, it also has several advantages such as generally greater levels of foreign protein expression, facile protein isolation in high purity, effective protection of target protein from proteolysis, and efficient production of harmful or toxic proteins [31,32]. Purification of polyhedrin-fused protein in recombinant *E. coli* cells and recovery of sole mussel silk-like protein are currently underway.

CONCLUSIONS

We selected marine mussel as a new silk-bearing organism because we found silk-like domains containing rich-glycine and rich-alanine repeats from mussel pre-collagen D protein. By deleting lengthy collagen domain and having optimized *E. coli* codons, we redesigned mussel silk-like protein sequence for genetic expression in *E. coli*. When we used pET23b+ vector having strong T7 promoter, we could detect expression of mussel silk-like protein at low culture temperature (25 ºC) but could not at 37 ºC. We next used cold shock (15 ºC)-based pCOLD DNA and GST fusion-based pGEX4T-1 vectors to aim for increased soluble expressions. However, there were severe retardations on cell growths and possible detections with only sensitive Western blot analysis due to very limited expressions of soluble mussel silk-like proteins. Therefore, we decided to use baculoviral polyhedrin as an insoluble inclusion body formation-inducing fusion partner along with pET23b+ vector system. Finally, mussel silk-like protein was successfully expressed as an insoluble inclusion body without cell growth inhibition and detected via just SDS-PAGE analysis.

ACKNOWLEDGEMENT

This work was supported by grants from Marine Biotechnology Program funded by the Ministry of Land, Transport and Maritime Affairs, Korea and from Brain Korea 21 Program funded by the Ministry of Education, Science and Technology, Korea.

REFERENCES

1. Y. Cao and B. Wang, *Int. J. Mol. Sci*., **10**, 1514 (2009).
2. J. M. Gosline, M. E. Demont and M. W. Denny, *Endeavour*, **10**, 37 (1986).
3. R. V. Lewis, *Chem. Rev.*, **106**, 3762 (2006).
4. J. Huang, C. Wong Po Foo and D. L. Kaplan, *Polym. Rev.*, **47**, 29 (2007).
5. T. D. Sutherland, J. S. Church, X. Hu, M. G. Huson, D. L. Kaplan and S. Weisman, *PLoS ONE*, **6**, e16489 (2011).
6. S. Weisman, V. S. Haritos, J. S. Church, M. G. Huson, S. T. Mudie, A. J. W. Rodgers, G. J. Dumsday and T. D. Sutherland, *Biomaterials*, **31**, 2695 (2010).
7. J. Gosline, M. Lillie, E. Carrington, P. Guerette, C. Ortlepp and K. Savage, *Phil. Trans. R. Soc. Lond.*, **B 357**, 121 (2002).
8. C. Vendrely and T. Scheibel, *Macromol. Biosci*., **7**, 401 (2007).
9. M. J. Harrington and J. H. Waite, *J. Exp. Biol.*, **210**, 4307 (2007).
10. X.-X. Qin, K. J. Coyne and J. H. Waite, *J. Biol. Chem.*, **272**, 32623 (1997).
11. X.-X. Qin and J. H. Waite, *Proc. Natl. Acad. Sci.*, **95**, 10517 (1998).
12. K. J. Coyne, X.-X. Qin and J. H. Waite, *Science*, **277**, 1830 (1997).
13. S. J. Lombardi and D. L. Kaplan, *J. Arachnol.*, **18**, 297 (1990).
14. A. Rising, M. Widhe, J. Johansson and M. Hedhammar, *Cell. Mol. Life Sci*., **68**, 169 (2011).
15. C. L. Craig and C. Riekel, *Comp. Biochem. Phys., B.*, **133**, 493 (2002).
16. P. Guerette, D. Ginzinger, B. Weber and J. Gosline, *Science*, **272**, 112 (1996).
17. H. G. Menzella, *Microb. Cell Fact.*, **10**, 15 (2011).
18. S. Winkler and D. L. Kaplan, *Rev. Mol. Biotechnol*., **74**, 85 (2000).
19. S. R. Fahnestock and L. A. Bedzyk, *Appl. Microbiol. Biotechnol.*, **47**, 33 (1997).
20. P. E. Vaillancourt, *E. coli Gene Expression Protocols*, **205**, 1 (2003).
21. J. Goldstein, N. S. Pollitt and M. Inouye, *Proc. Natl. Acad. Sci.*, **87**, 283 (1990).
22. H. Tanabe, J. Goldstein, M. Yang and M. Inouye, *J. Bacteriol.*, **174**, 3867 (1992).
23. J. Steczko, G. A. Donoho, J. E. Dixon, T. Sugimoto and B. Axelrod, *Protein Expr. Purif.*, **2**, 221 (1991).
24. M. Tamura, K. Ito, S. Kunihiro, C. Yamasaki and M. Haragauchi, *Protein Expr. Purif.*, **78**, 1 (2011).
25. S. C. Makrides, *Microbiol. Rev.*, **60**, 512 (1996).
26. E. R. LaVallie and J. M. McCoy, *Curr. Opin. Biotechnol.*, **6**, 501 (1995).
27. W. Guo, L. Cao, Z. Jia, G. Wu, T. Li, F. Lu and Z. Lu, *Protein Expr. Purif.*, **77**, 185 (2011).
28. K. A. Harrap, *Virology*, **50**, 124 (1972).
29. G. F. Rohrmann, *J. Gen. Virol*., **67**, 1499 (1986).
30. J. H. Seo, L. Li, J. S. Yeo and H. J. Cha, *Biotechnol. Bioeng.*, **84**, 467 (2003).
31. R. Rudolph and H. Lilie, *FASEB J.*, **10**, 49 (1996).
32. R. L. Quant, M. N. Pearson, G. F. Rohrmann and G. S. Beaudreau, *Appl. Environ. Microbiol.*, **48**, 732 (1984).
33. Q. Wei, Y. S. Kim, J. H. Seo, W. S. Jang, I. H. Lee and H. J. Cha, *Appl. Environ. Microbiol.*, **71**, 5038 (2005).