Biocatalysis and Agricultural Biotechnology 20 (2019) 101180

Contents lists available at ScienceDirect

Biocatalysis and Agricultural Biotechnology

journal homepage: www.elsevier.com/locate/bab

Enhanced recombinant insulin production in transgenic *Escherichia coli* that heterologously expresses carrot heat shock protein 70

Bomin Jang, Yeh-Jin Ahn*

Department of Biotechnology, Sangmyung University, 20 Hongjimun 2-gil, Jongno-gu, Seoul, 03016, South Korea

ARTICLE INFO

Keywords:
Escherichia coli
Heat shock protein
Homologous recombination
Insulin
Molecular chaperone
Recombinant protein

ABSTRACT

Recombinant insulin production was enhanced in transgenic *Escherichia coli* (*E. coli* BL21, DE3) that heterologously expressed carrot (*Daucus carota* L.) heat shock protein 70 (DcHsp70). The gene encoding DcHsp70 with the constitutive bacterial lipoprotein promoter was inserted into the *E. coli* genome using lambda Red-mediated homologous recombination. Nucleotide sequences for the human proinsulin, insulin B and insulin A chains were separately cloned into the pVFT 2S expression vector containing a 6xhistidine and glutathione S-transferase (GST) tag. When induced by isopropyl β-D-1-thiogalactopyranoside treatment, the expression levels of the proinsulin and insulin A and B chain fusion proteins were higher in the transgenic cell lines that heterologously expressed DcHsp70 than in the control cell line. The 6xHis-GST tag was removed from the three versions of the fusion proteins by tobacco etch virus protease. As an active molecular chaperone, DcHsp70 could increase recombinant insulin production when heterologously expressed in *E. coli*. Our results suggest a possible use of plant Hsps as effective molecular chaperones to increase recombinant protein production in *E. coli*.

1. Introduction

Escherichia coli (*E. coli*) is one of the most widely used microorganisms for recombinant protein production. It produces approximately 30% of approved therapeutic proteins, such as insulin, interferon, and hormones (Baeshen et al., 2015). *E. coli* grows rapidly in a well-defined medium. The genetic information of *E. coli* is well understood, and foreign gene insertion is relatively easy using episomal vectors. However, in some cases, *E. coli* accumulates recombinant proteins in inclusion bodies as insoluble aggregates, which requires further renaturation and refolding (Schumann and Ferreira, 2004). Approximately 30% of recombinant proteins are expressed as insoluble aggregates in *E. coli* (Leibly et al., 2012).

To increase the amount and solubility of recombinant proteins produced in *E. coli*, a number of parameters, such as medium composition, host cell lines, expression vectors, fusion tags, and protein induction conditions (induction timing, temperature, duration, and inducer concentration), have been carefully optimized (Papaneophytou and Kontopidis, 2014). Furthermore, *E. coli* can be genetically modified to inhibit genes encoding proteases and/or to introduce genes encoding molecular chaperones, which prevent denaturation of proteins and/or assist correct protein folding (Castiñeiras et al., 2018).

Heat shock proteins (Hsps) are a major class of molecular chaperones that are ubiquitously expressed in all living organisms under stressful conditions by the heat shock response (Richter et al., 2010). Hsps protect proteins from harmful damage caused by environmental stressors. They are classified into five groups based on molecular weight (Hsp100, Hsp90, Hsp70, Hsp60, small Hsp; Vierling, 1991). Each group has unique structural characteristics and functional mechanisms. Bacterial Hsps, such as GroEL/GroES (Hsp60/Hsp10), DnaK/DnaJ/GrpE (Hsp70/Hsp40/cochaperone), ClpB (chaperone), and IbpA/IbpB (small Hsps), were overexpressed in *E. coli* and greatly increased the solubility of recombinant proteins (De Marco et al., 2007).

We developed transgenic *E. coli* cell lines that heterologously express an Hsp from plants. Plants produce the most abundant and diverse Hsps among all organisms (Al-Whaibi, 2011). For example, rice (*Oryza sativa*) has 32 different Hsp70 genes (Sarkar et al., 2013). Considering the sessile nature of plants, abundant and diverse Hsps may confer advanced protection to the organisms under stressful growth conditions.

In this study, we examined whether transgenic *E. coli* cell lines that heterologously express Hsp70 from carrot (*Daucus carota*; DcHsp70) could increase the amount and/or solubility of recombinant human insulin. Hsp70 is one of the most highly conserved molecular chaperones (Richter et al., 2010). It folds and assembles newly synthesized proteins in physiological states. It also plays an important role in

---

* Corresponding author.
E-mail address: yjahn@smu.ac.kr (Y.-J. Ahn).

https://doi.org/10.1016/j.bcab.2019.101180
Received 23 March 2019; Received in revised form 25 May 2019; Accepted 26 May 2019

restoring the folding of denatured proteins under stress conditions (Ghazaei, 2017). Structurally, Hsp70s are characterized by an N-terminal region with an ATPase domain (44 kDa) and a C-terminal region with a peptide binding domain (25 kDa; Bukau and Horwich, 1998). ATP binding to the ATPase domain triggers conformational changes in the C-terminal region opening up the peptide-binding domain (Zhu et al., 1996; Sung et al., 2001). ATP hydrolysis provides energy for substrate refolding. The second round of ATP binding opens up the peptide binding domain, releasing the refolded substrate. By repeating these steps, Hsp70 acts as an efficient ‘folding machine’ to reactivate denatured proteins.

Insulin produced in the endoplasmic reticulum of pancreas β-cells plays an important role in carbohydrate metabolism by promoting glucose uptake into cells (Choi et al., 2013). Proinsulin is transferred to the Golgi apparatus. The C peptide (31 a.a.) in the middle of the proinsulin is then cleaved, resulting in insulin A and B chains (21 and 30 a.a., respectively), followed by disulfide bond formation (Qiao et al., 2003; Dunn, 2005). Human insulin, the first recombinant protein used therapeutically, was produced by Eli Lilly using *E. coli* in 1982. Due to the increase in the number of diabetic patients, the world market for insulin is expected to increase by over 5% annually (Choi et al., 2013). Successful enhancement of recombinant human insulin production will be valuable in the treatment of diabetes mellitus.

2. Materials and methods

2.1. DNA construct to insert the DcHsp70 gene into the E. coli genome

We inserted the *DcHsp70* gene into the *E. coli* genome via homologous recombination using the Quick & Easy *E. coli* gene Deletion Kit (Gene bridges, Heidelberg, Germany). A DNA construct, lipoprotein (Lpp) promoter– DcHsp70 gene - Flippase recombination target (FRT) cassette, which facilitates homologous recombination, was generated by overlap PCR. First, each DNA fragment was generated by regular PCR: the Lpp promoter (GenBank accession no. NC_000913.2) was amplified from *E. coli* K-12 (MG1655) under the PCR conditions of 1 cycle at 98 °C for 30 s, followed by 35 cycles of 98 °C for 10 s, 55 °C for 30 s, and 72 °C for 30 s using Lpp promoter F and R primers (Table 1). DcHsp70 gene (Suppl. Fig. 1) was amplified under the PCR conditions of 1 cycle at 94 °C for 4 min, followed by 35 cycles of 94 °C for 1 min, 56.5 °C for 1 min, 72 °C for 4 min, 1 cycle at 72 °C for 10 min, and 1 cycle at 4 °C for 10 min using DcHsp70 F and R primers (Table 1). The FRT cassette containing the kanamycin resistance gene was amplified under the PCR conditions of 1 cycle at 98 °C for 30 s, followed by 35 cycles of 98 °C for 10 s, 55 °C for 30 s, 72 °C for 1 min 30 s, 1 cycle at 72 °C for 10 min, and 1 cycle at 4 °C for 10 min using FRT cassette F and R primers (Table 1).

The Lpp promoter and *DcHsp70* gene were joined by overlap PCR without primers (1 cycle at 96 °C for 3 min, followed by 35 cycles of 96 °C for 30 s and 72 °C for 2 min 10 s). The FRT cassette was further joined under the PCR conditions of 1 cycle at 96 °C for 3 min, followed by 35 cycles of 96 °C for 30 s, 72 °C for 4 min, 1 cycle at 72 °C for 10 min, and 1 cycle at 4 °C for 10 min.

2.2. Insertion of the DNA construct into the E. coli genome using homologous recombination

To facilitate the homologous recombination of the DNA construct containing the *DcHsp70* gene, the Red/ET plasmid was first inserted into *E. coli* BL21 (DE3) and expressed according to the protocol provided by the kit manufacturer (Gene bridges).

The DNA construct, Lpp promoter - DcHsp70 gene - FRT cassette, was mixed with competent *E. coli* cells (BL21, DE3; 100 µl) containing the Red/ET plasmid. After stabilization on ice for 30 min, the cells were electroshocked (1800 V, 25 µF, 200 Ω) using Gene Pulser X cells (Bio-Rad, Hercules, US). Super optimal SOC broth (1 ml) was then added, and the cells were incubated at 37 °C with agitation at 225 rpm for 3 h, followed by spreading on solid Luria-Bertani (LB) medium containing 15 µg/ml kanamycin. After overnight incubation at 37 °C, colony PCR was performed on surviving colonies using a specific primer set to amplify the DcHsp70 gene (DcHsp70 F and R; Table 1). The DNA sequences were also confirmed (Macrogen, Seoul, Korea).

2.3. Immunoblot analysis to detect DcHsp70 in transgenic E. coli

Wild type (*E. coli* BL21, DE3) and transformed cell lines were cultured in LB broth containing 15 µg/ml kanamycin at 37 °C at 130 rpm overnight. Samples were diluted 1:1000 with fresh LB broth containing 15 µg/ml kanamycin, followed by incubation at 37 °C with agitation at 225 rpm until the $O.D._{600}$ reached 0.6. Samples were then centrifuged at 4 °C and 1360 × *g* for 20 min, and the supernatant was removed. The resulting pellet was dissolved in protein extraction buffer (3 ml; 25 mM Tris-HCl pH 7.5, 300 mM NaCl, and 3 mM β-mercaptoethanol) and ultrasonicated (5 cycles of 420 W and 20 kHz for 10 s and 30 s pause), followed by centrifugation at 20,900 × *g* for 30 min. The protein concentration in the supernatant was quantified by Bradford assay (1976).

Proteins (25 µg) were resolved on a 17% SDS-PAGE gel (Laemmli, 1970). The separated proteins were transferred to polyvinylidene difluoride membranes (GE Healthcare Life Science, Buckinghamshire, UK) at 180 mA overnight. The membrane was incubated in phosphate-buffered saline (PBS)-T buffer (containing 0.1% Tween 20) containing 5% skim milk for 1 h for blocking. The membrane was then incubated with the primary antibody (anti-Hsp70 monoclonal antibody, Abcam, Cambridge, UK; diluted 1:2000 in PBS-T buffer) for 1 h and with the secondary antibody (HRP-conjugated anti-mouse, Amersham Biosciences, Pittsburgh, US; diluted 1:30000 in PBS-T buffer) for 1 h. The detection of DcHsp70 was performed using the ECL Prime System (GE Healthcare Life Science). Chemiluminescent signals were captured by

Table 1
Primers used in this study.

Sets	DNA targets	Direction	Primers (5′-3′)	Product size (bp)
1	Lpp promoter	F	CCGGATCTTCCACAATACCAATCGCAGGCGAGAACATGCGACCCTGTAATATTGCTTT	316
		R	ACCTTTGCTAGCCATTATTAATACCCTCTA	
2	DcHsp70	F	TAGAGGGTATTAATAATGTCGATCATTCCA	1959
		R	GAGAATAGGAACTTCTTAACCAGAAATATC	
3	FRT cassette	F	GAGGAAGTTGACTAAGAAGTTCCTATTCTC	1637
		R	GGTACGCCGGGCTTTGAACTGCCGCTGGAGGGTGAAGTACGCGCGAAGTTCCTATACTTTC	
4	Proinsulin	F	GGGGGGGGATCCATGTTTGTGAACCAACACCTGTGC	258
		R	GGGGGGGAATTCTTAGTTGCAGTAGTTCTCCAGCTG	
5	Insulin B chain	F	GGGGGGGGATCCATGTTTGTGAACCAACACCTGTGC	90
		R	GGGGGGGAATTCTTAGGTCTTGGGTGTGTAGAAGAAGC	
6	Insulin A chain	F	GGGGGGGGATCCATGGGCATTGTGGAACAATGCTGTACC	63
		R	GGGGGGGAATTCTTAGTTGCAGTAGTTCTCCAGCTG	

Notes: Lpp*:* Lipoprotein gene (NC_000913.2); DcHsp70, *Daucus carota* Hsp70 (Suppl. Fig. 1); FRT, Flippase recombinant target; Insulin (NM_000207.2).

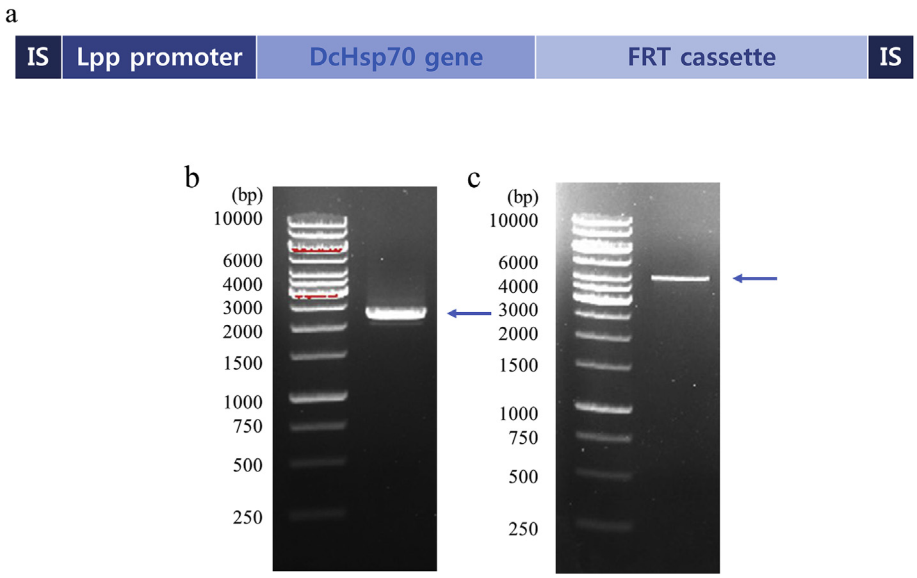

Fig. 1. DNA construct generation for homologous recombination of the DcHsp70 gene. (a) The overall diagram of the DNA construct is shown. The Lpp promoter, DcHsp70 gene, and FRT cassette for homologous recombination were generated by regular PCR and joined stepwise by overlap PCR. (b) The Lpp promoter – DcHsp70 gene (2,275 bp) and (c) Lpp promoter – DcHsp70 gene – FRT cassette (3,912 bp) were successfully amplified from transformed *E. coli*. IS: insertion sequences in the *E. coli* genome (*yddE* pseudogene sequences; GenBank accession no. NC_012971.2); Lpp: Lipoprotein; DcHsp70: *Daucus carota* heat shock protein 70; FRT: Flippase recombination target.

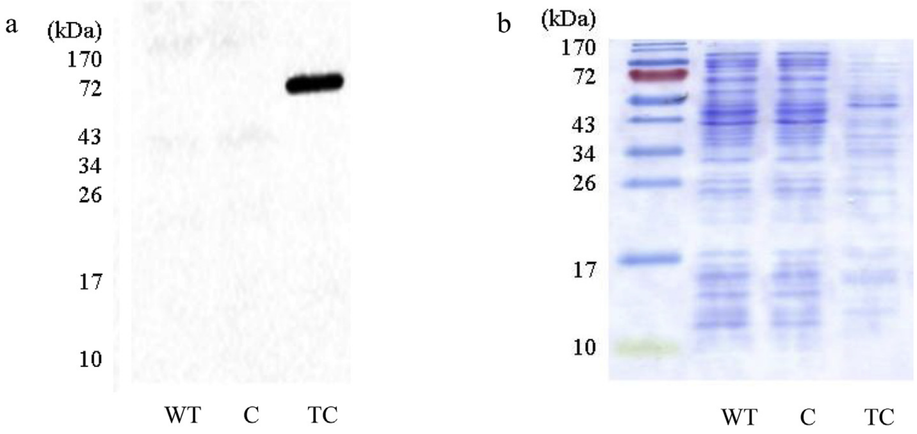

Fig. 2. Heterologous expression of DcHsp70 in transgenic *E. coli.* Proteins were extracted from unmodified *E. coli* BL21, DE3 (WT), *E. coli* containing Red/ET plasmid (C), and transgenic *E. coli* containing the DcHsp70 gene in its genome (TC) and resolved on an SDS-PAGE gel. (a) Immunoblot analysis was performed using an anti-Hsp70 monoclonal antibody. (b) Proteins were visualized using Coomassie blue dye. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

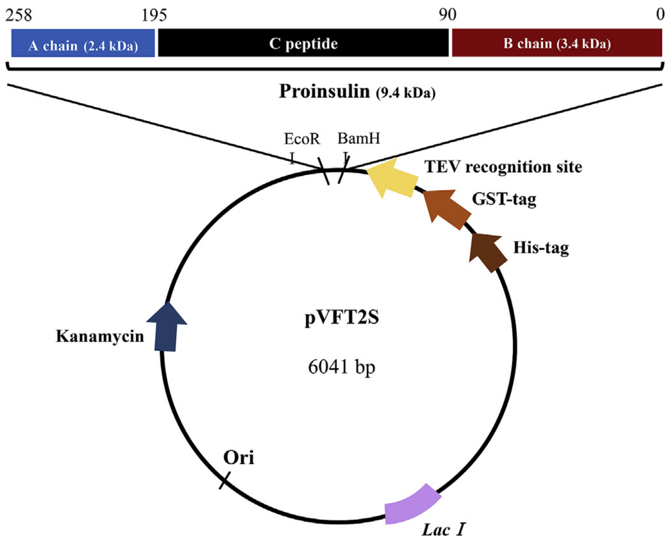

Fig. 3. Cloning of the human insulin gene. cDNAs for human proinsulin and insulin B and A chains were independently inserted into a pVFT2S expression vector containing a 6xHistidine (His)-glutathione S-transferase (GST) tag. A tobacco etch virus protease cleavage site is present to remove the tag after purification. LacI: Lactose inhibitor; Ori: Origin of replication; F1 ori: Ori from F1 phage.

using a ChemiDoc MP Imaging System (Bio-Rad).

2.4. Human insulin gene cloning

Nucleotide sequences for human proinsulin cDNA (GenBank accession no. NM_000207.2; Sino Biological, Chesterbrook, US), insulin B and A chains were amplified by PCR using a set of primers containing restriction enzyme digestion sites, BamHI and EcoRI, at both ends (Table 1; 1 cycle at 96 °C for 1 min, followed by 35 cycles of 96 °C for 20 s, 52 °C for 30 s, 72 °C for 30 s, 1 cycle at 72 °C for 10 min, and 1 cycle at 4 °C for 10 min), followed by double-digestion with the two DNA restriction enzymes. Each DNA fragment was inserted into the expression vector pVFT2S containing a 6xHistidine - GST dual tag and a tobacco etch virus (TEV) protease cleavage site.

Each insert, double-digested proinsulin, and insulin B and A chains was ligated into the pVFT2S expression vector using T4 ligase (Enzynomics, Daejeon, Korea) at 16 °C overnight. The ligation mixture (5 ml) was added to 100 µl of *E. coli* (DH5α) competent cells, followed by incubation on ice for 30 min. The cells were heat-shocked (42 °C) for 2 min to introduce the ligated plasmids and incubated on ice for 10 min. After adding fresh LB medium (1 ml), the cells were cultured at 37 °C for 1 h at 350 rpm and spread on solid LB medium containing 30 µg/ml kanamycin, followed by overnight incubation at 37 °C. Colony PCR was performed on surviving colonies using three sets of primers used in the amplification of proinsulin and insulin B and A chains (Table 1). Each colony was developed into independent cell lines. The DNA sequences

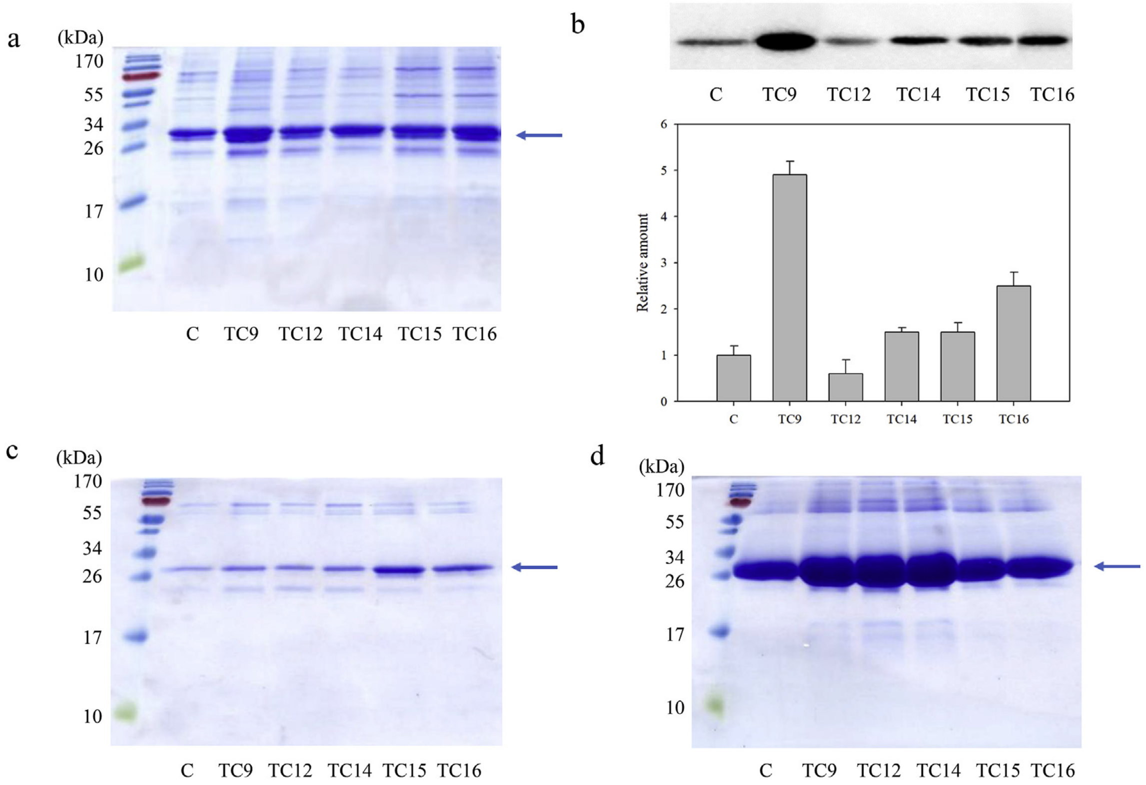

Fig. 4. Expression of recombinant human proinsulin in *E. coli*. Human proinsulin cDNA was inserted into the pVFT2S expression vector and induced by isopropyl β-D-1-thiogalactopyranoside. After protein extraction from *E. coli* cell lines, 6xHis-GST-proinsulin was purified using Ni-NTA His-Bind resin and subjected to (a) SDS-PAGE and (b) immunoblot analysis using an anti-insulin B chain monoclonal antibody. Chemiluminescent signals were quantified, and relative amounts of protein were calculated. Purified 6xHis-GST-proinsulin was centrifuged to separate (c) soluble and (d) insoluble forms. C: *E. coli* BL21, DE3 containing recombinant pVFT2S-proinsulin plasmid; TC: transgenic *E. coli* cell lines that heterologously express DcHsp70 and contain the recombinant pVFT2S-proinsulin plasmid.

were also confirmed (Macrogen).

2.5. *Expression of recombinant human insulin*

Each recombinant plasmid, pVFT2S-proinsulin, pVFT2S-insulin B chain, and pVFT2S-insulin A chain was purified from overnight cultures of transformed *E. coli* DH5α using the DNA-spin Plasmid DNA Purification Kit Protocol (iNtRON biotechnology, Seongnam, Korea) and inserted into *E. coli* BL21, DE3 (unmodified wild type and transgenic cell lines that heterologously express DcHsp70).

Transformed cell lines were cultured in LB medium containing 30 μg/ml kanamycin at 37 °C overnight, diluted 1:1000 with fresh LB medium, and continuously cultured until the $O.D._{600}$ reached 0.6. The cells were then treated with isopropyl β-D-1-thiogalactopyranoside (IPTG, 0.1 mM at 16 °C for 20 h) to induce the recombinant proinsulin, and insulin B and A chains, followed by centrifugation at 1360 × *g* at 4 °C for 20 min. The resulting pellet was dissolved in protein extraction buffer, and proteins were extracted using ultrasonication as described above.

2.6. *Purification of recombinant human insulin*

To purify 6xHis-GST-proinsulin, 6xHis-GST-insulin B chain, and 6xHis-GST-insulin A chain, extracted proteins were incubated with Ni-NTA His-Bind resin (Novagen, Temecula, US) for 1 h. The resin was washed with washing buffer 1 (25 mM Tris-HCl pH 7.5, 200 mM KCl, 10 mM imidazole, 10% glycerol, and 3 mM β-mercaptoethanol) and washing buffer 2 (25 mM Tris-HCl pH 7.5, 500 mM KCl, 50 mM imidazole, 10% glycerol, and 3 mM β-mercaptoethanol) to remove non-specific proteins. Elution buffer (25 mM Tris-HCl pH 7.5, 200 mM KCl, 250 mM imidazole, 10% glycerol, and 3 mM β-mercaptoethanol) was then added to collect purified recombinant proteins.

The purified proteins were further separated by centrifugation at 20,900 × *g* at 4 °C for 1 h to separate soluble fractions. The insoluble pellet was dissolved in elution buffer. Soluble and insoluble proteins (20 μl each) were separated on a 17% SDS-PAGE gel and stained with Coomassie Brilliant Blue R-250 (Bio-Rad).

2.7. *Immunoblot analysis of recombinant human proinsulin and insulin B chain*

Purified proteins were subjected to 17% SDS PAGE and immunoblot analysis using anti-insulin B chain monoclonal antibody (Santa Cruz Biotechnology, Dallas, US) as described above. Chemiluminescent signals were captured and quantified by using a ChemiDoc MP Imaging System (Bio-Rad).

2.8. *Removal of 6xHis-GST-tag*

The 6xHis-GST tag was cleaved from the fusion proteins using TEV protease in the reaction buffer (25 mM Tris-HCl pH 8.0, 150 mM NaCl, and 14 mM β-mercaptoethanol) at 22 °C for 20 h. SDS-PAGE and immunoblot analysis using anti-insulin B chain monoclonal antibody was performed as described above.

3. Results and discussion

3.1. *Heterologous expression of DcHsp70 in E. coli*

To develop transgenic *E. coli* cell lines that heterologously express DcHsp70, the Lpp promoter – DcHsp70 gene – FRT cassette was inserted into the *E. coli* genome by homologous recombination (Fig. 1a).

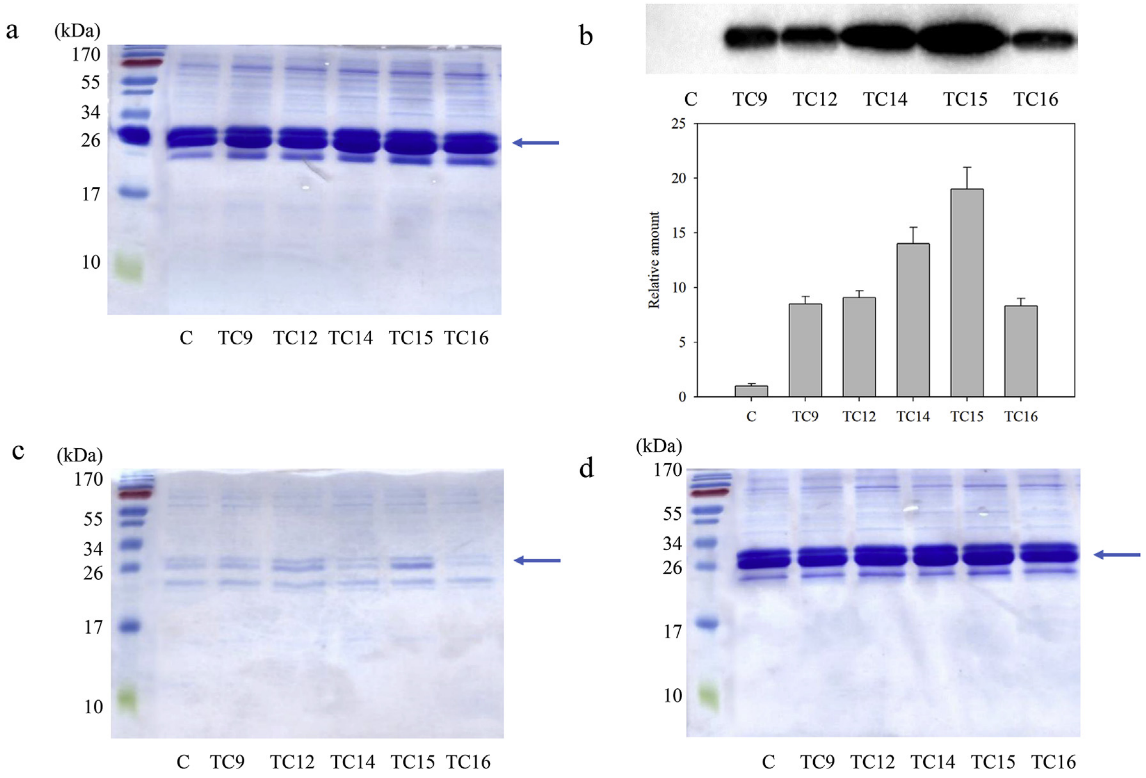

Fig. 5. Expression of recombinant human insulin B chain in *E. coli*. Human insulin B chain cDNA was inserted into the pVFT2S expression vector and induced by isopropyl β-D-1-thiogalactopyranoside. After protein extraction from *E. coli* cell lines, the 6xHis-GST-insulin B chain was purified using Ni-NTA His-Bind resin and subjected to (a) SDS-PAGE and (b) immunoblot analysis using an anti-insulin B chain monoclonal antibody. Chemiluminescent signals were quantified, and relative amounts of protein were calculated. The purified 6xHis-GST-insulin B chain was centrifuged to separate (c) soluble and (d) insoluble forms. C: *E. coli* BL21, DE3 containing recombinant pVFT2S-insulin B chain plasmid; TC: transgenic *E. coli* cell lines that heterologously express DcHsp70 and contain recombinant pVFT2S-insulin B chain plasmid.

The successful insertion of the DNA construct was confirmed by PCR. Two DNA fragments, the Lpp promoter - DcHsp70 gene (2,275 bp) and the Lpp promoter – DcHsp70 gene – FRT cassette (3,912 bp), were successfully amplified from the genomic DNA of the transformed cells (Fig. 1b and c, respectively). The sequences of the DNA construct, including the insertion site of the *E. coli* genome (*yddE* pseudogene, GenBank accession no. NC_012971.2), were also confirmed (data not shown).

To examine whether DcHsp70 was successfully expressed in transgenic *E. coli*, proteins were extracted from unmodified *E. coli* BL21, DE3 (WT), *E. coli* containing the RedE/T plasmid (C) and transformed *E. coli* cell lines containing the Lpp promoter – DcHsp70 gene – FRT cassette construct in the genome (TC). Heterologous expression of DcHsp70 was observed in the transformed *E. coli* by immunoblot analysis using an anti-DcHsp70 monoclonal antibody (Fig. 2a). Control cell lines, WT and C, did not show any signals, suggesting that the protein band observed in the transgenic cell line indicates the presence of DcHsp70 and not the endogenous bacterial Hsp70.

Our results suggest that the bacterial Lpp promoter successfully expressed the heterologous DcHsp70 gene inserted in the *yddE* pseudogene site in the bacterial genome. This system can be used to insert and express heterologous genes and/or extra copies of endogenous bacterial genes. When inserting a foreign gene, genomic insertion by homologous recombination has a number of advantages over using episomal expression vectors, such as avoiding cumbersome DNA restriction enzyme digestion and ligation steps for inserts and vectors. Furthermore, considering the possible instability of plasmids, such as plasmid loss and conformational changes causing truncated forms of the plasmid, genomic insertion appears to be a promising alternative to episomal vectors for recombinant gene expression (Castiñeiras et al., 2018).

Transgenic cell lines that heterologously expressed DcHsp70 showed different total protein profiles compared to control cell lines (Fig. 2b). The major protein bands shown in the SDS-PAGE analysis were changed in all the molecular weight ranges. These results suggest that DcHsp70 directly and/or indirectly affected bacterial protein production. Heterologous protein expression can alter gene expression and biochemical mechanisms of the host organism. Valdez-Cruz et al. (2011) reported that the expression of a foreign protein can alter the levels of various transcripts in *E. coli*. Altered transcript and protein profiles caused by heterologous expression of DcHsp70 in *E. coli* remain to be examined.

3.2. Enhanced expression of recombinant human insulin in transgenic E. coli cell lines

Each cDNA sequence for human proinsulin and insulin B and A chains was amplified by PCR and inserted into a pVFT2S expression vector containing a dual 6xHis-GST tag for purification (Fig. 3). The three recombinant plasmids were independently inserted into *E. coli* BL21, DE3 and the transgenic *E. coli* cell lines expressing DcHsp70. Expression of recombinant proinsulin and the two insulin chains was induced by IPTG treatment, and recombinant proteins were purified using Ni resin. Recombinant proinsulin and insulin B chain were further detected using an anti-insulin B chain monoclonal antibody.

The purified 6xHis-GST-proinsulin (expected size of 34.4 kDa) was detected by SDS-PAGE (Fig. 4a). Immunoblot analysis showed that recombinant proinsulin was present in all *E. coli* cell lines examined. Compared to the proinsulin level in *E. coli* BL21, DE3 ©, four out of five transgenic *E. coli* cell lines expressing DcHsp70 (TC) showed higher levels of recombinant proinsulin (Fig. 4b). Among them, TC9 showed the highest (5-fold) increase compared to the control. The recombinant

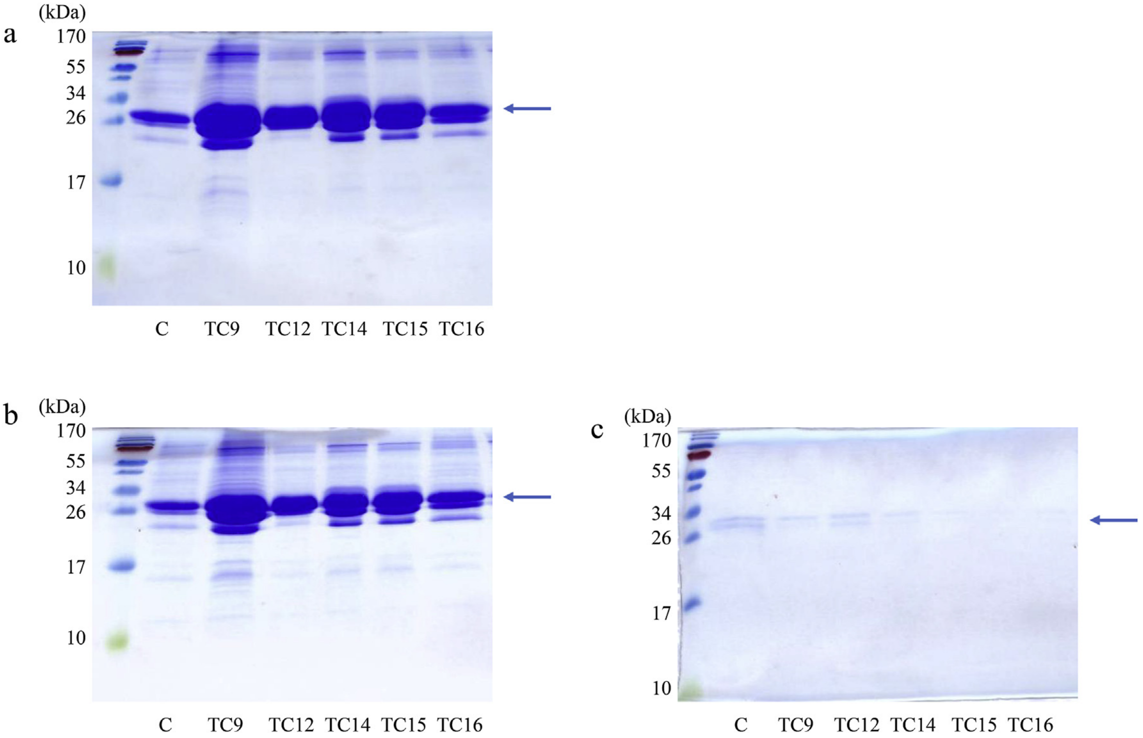

Fig. 6. Expression of recombinant human insulin A chain in *E. coli*. Human insulin A chain cDNA was inserted into the pVFT2S expression vector and induced by isopropyl β-D-1-thiogalactopyranoside. After protein extraction from *E. coli* cell lines, the 6xHis-GST-insulin A chain was purified using Ni-NTA His-Bind resin and subjected to (a) SDS-PAGE. The purified 6xHis-GST-insulin A chain was centrifuged to separate (b) soluble and (c) insoluble forms. C: *E. coli* BL21, DE3 containing recombinant pVFT2S-insulin A chain plasmid; TC: transgenic *E. coli* cell lines that heterologously express DcHsp70 and contain recombinant pVFT2S-insulin A chain plasmid.

6xHis-GST-proinsulin was mostly present in insoluble forms (soluble and insoluble forms: Fig. 4c and d, respectively).

The recombinant 6xHis-GST-insulin B chain (expected size of 28.4 kDa) was purified from the control and transgenic cell lines (Fig. 5a). Immunoblot analysis showed that the fusion protein was detected only in the transgenic cell lines expressing DcHsp70 but was absent in the control (Fig. 5b). Among the five transgenic cell lines examined, TC15 showed the highest level of accumulation. The 6xHis-GST-insulin B chain was also present in insoluble forms (soluble and insoluble forms: Fig. 5c and d, respectively), similar to the recombinant 6xHis-GST-proinsulin.

The recombinant 6xHis-GST-insulin A chain (expected size of 27.4 kDa) was also purified using Ni-resin (Fig. 6a). Unlike the recombinant proinsulin and insulin B chain, which were insoluble after purification, the recombinant insulin A chain was mostly present in soluble forms (soluble and insoluble forms: Fig. 6b and c, respectively). The transgenic cell line TC9 showed the highest level of recombinant insulin A chain among all the *E. coli* cell lines examined.

Our results showed that heterologously expressed DcHsp70 increased the amount of the recombinant proinsulin and insulin B and A chains. However, it did not increase the solubility of the recombinant proinsulin and insulin B chain. It has been previously known that recombinant insulin expressed in *E. coli* tends to form inclusion bodies, insoluble protein aggregates, which should be refolded to regain biological activity (Hwang et al., 2016).

Molecular chaperones were employed to enhance recombinant insulin production. A Hsp from *Caenorhabditis elegans*, *CeHsp17*, prevented the aggregation of misfolded substrates, such as chemically denatured insulin (Zhang et al., 2015). More recently, fusion with the molecular chaperone αB-crystallin increased the expression and proper folding of human insulin (Akbarian and Yousefi, 2018). The results from the present and previous studies suggest that molecular chaperones, such as Hsps, can enhance the amount, folding, and solubility of recombinant human insulin.

3.3. Removal of the 6xHis-GST dual tag from recombinant insulin

The 6xHis-GST purification tag was removed from human proinsulin (Fig. 7a), insulin B chain (Fig. 7b) and insulin A chain (Fig. 7c) by using the TEV protease (27 kDa), which appeared in all three SDS-PAGE analyses. Immunoblot analysis showed that transgenic *E. coli* cell lines produced higher levels of purified recombinant human proinsulin and insulin B chain (Fig. 7d and e, respectively). There was no other protein bands detected by the anti-insulin B chain monoclonal antibody in all the molecular weight ranges, including those of the fusion proteins (34.4 kDa for 6xHis-GST-proinsulin and 28.4 kDa for 6xHis-GST-insulin B chain), suggesting a successful removal of the 6xHis-GST tag. Compared to the levels expressed in the control, transgenic cell lines produced up to 4.5-fold and 9-fold higher levels of proinsulin and insulin B chain, respectively.

Purification tags are essential for recombinant protein production. They often affect the amount and/or solubility of the recombinant fusion protein. GST, a 26 kDa protein from parasitic *Schistosoma japonicum*, has been reported to protect recombinant proteins from degradation and stabilize proteins (Young et al., 2012). When human proinsulin and insulin B and A chain genes were fused to a 6xHis tag, without a GST tag, we did not observe the corresponding recombinant protein expression in SDS-PAGE and immunoblot analyses (data not shown).

4. Conclusions

We showed that Hsp70 from carrot (*Daucus carota* L.) heterologously expressed in *E. coli* could increase recombinant human insulin production. Our results suggest that Hsps from plants can be used to enhance recombinant protein production in *E. coli* crossing species

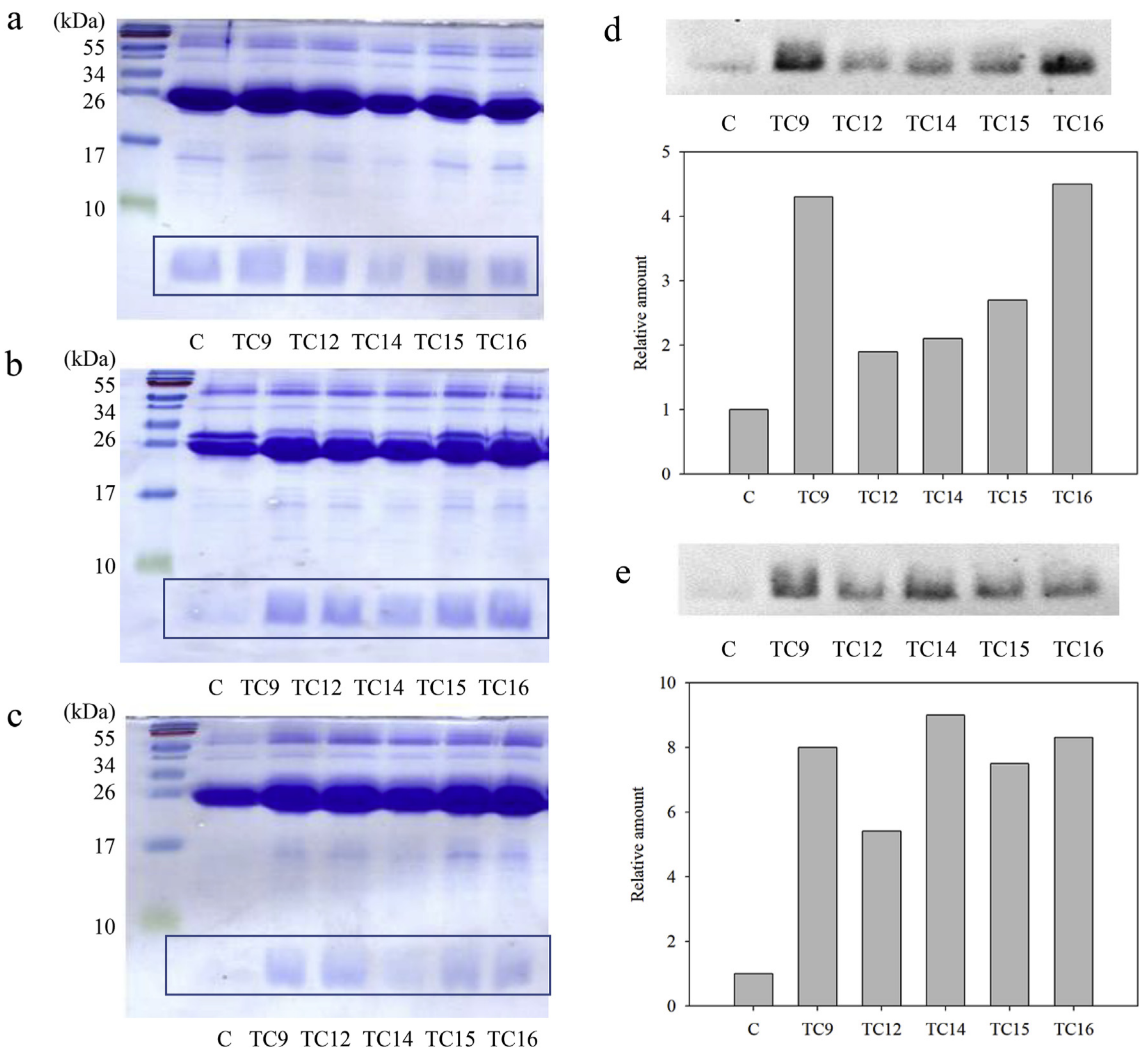

Fig. 7. Tag removal from recombinant proinsulin, insulin B and A chains. The 6xHis-GST tag was cleaved from the fusion proteins using tobacco etch virus protease. (a) Recombinant human proinsulin (9.4 kDa), (b) insulin B chain (3.4 kDa), and (c) insulin A chain (2.4 kDa) are shown by SDS-PAGE. Immunoblot analysis was performed on (d) recombinant proinsulin and (e) insulin B chain using an anti-insulin B chain monoclonal antibody. Chemiluminescent signals were quantified, and relative amounts of protein were calculated. C: *E. coli* BL21, DE3 containing (a) recombinant pVFT2S-proinsulin plasmid, (b) recombinant pVFT2S-insulin B chain plasmid, and (c) recombinant pVFT2S-insulin A chain plasmid; TC: transgenic *E. coli* cell lines that heterologously express DcHsp70 and contain (a) recombinant pVFT2S-proinsulin plasmid, (b) recombinant pVFT2S-insulin B chain plasmid, and (c) recombinant pVFT2S-insulin A chain plasmid.

lines. In the future, it will be useful to select effective molecular chaperones, possibly from plants, that can increase recombinant insulin production in cell factories.

Conflicts of interest

The authors declare no competing interests.

Acknowledgements

This research was supported by a 2019 Research Grant from Sangmyung University.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bcab.2019.101180.

References

Akbarian, M., Yousefi, R., 2018. Human αB crystalline as fusion protein and molecular chaperone increases the expression and folding efficiency of recombinant insulin. PLoS One 13 (10), e0206169.

Al-Whaibi, M.H., 2011. Plant heat-shock proteins: a mini review. J. King Saud Univ. Sci. 23 (2), 139–150.

Baeshen, M.N., Al-Hejin, A.M., Bora, R.S., Ahmed, M.M., Ramadan, H.A., Saini, K.S., Baeshen, N.A., Redwan, E.M., 2015. Production of biopharmaceuticals in *E. coli*: current scenario and future perspectives. J. Microbiol. Biotechnol. 25 (7), 953–962.

Bradford, M.M., 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72, 248–254.

Bukau, B., Horwich, A.L., 1998. The Hsp70 and Hsp60 chaperone machines. Cell 92 (3), 351–366.

Castiñeiras, S.T., Williams, S.G., Hitchcock, A.G., Smith, D.C., 2018. *E. coli* strain engineering for the production of advanced biopharmaceutical products. FEMS Microbiol. Lett. 365 (15). https://doi.org/10.1093/femsle/fny162.

Choi, Y.J., Park, S.H., Kim, J.S., Wi, S.J., Park, K.Y., 2013. Biosynthesis of recombinant human prominiinsulin in *E. coli* and plant systems. J. Plant Biotechnol. 40, 169–177.

De Marco, A., Deuerling, E., Mogk, A., Tomoyasu, T., Bukau, B., 2007. Chaperone-based procedure to increase yields of soluble recombinant proteins produced in *E. coli*. BMC Biotechnol. 7, 32. https://doi.org/10.1186/1472-6750-7-32.

Dunn, M.F., 2005. Zinc-ligand interactions modulate assembly and stability of the insulin hexamer. Biometals 18, 295–303.

Ghazaei, C., 2017. Role and mechanism of the Hsp70 molecular chaperone machines in bacterial pathogens. J. Med. Microbiol. 66 (3), 259–265.

Hwang, H., Kim, K., Lee, S., Kim, C., Min, C., Yun, J., Lee, S., Son, Y., 2016. Recombinant glargine insulin production process using *Escherichia coli*. J. Microbiol. Biotechnol. 26 (10), 1781–1789.

Laemmli, U.K., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227 (5259), 680–685.

Leibly, D.J., Nguyen, T.N., Kao, L.T., Hewitt, S.N., Barrett, L.K., Van Voorhis, W.C., 2012. Stabilizing additives added during cell lysis aid in the solubilization of recombinant proteins. PLoS One 7, e52482 2012.

Papaneophytou, C.P., Kontopidis, G., 2014. Statistical approaches to maximize recombinant protein expression in *Escherichia coli*: a general review. Protein Expr. Purif. 94, 22–32.

Qiao, Z.S., Min, C.Y., Hua, Q.X., Weiss, M.A., Feng, Y.M., 2003. In vitro refolding of human proinsulin. Kinetic intermediates, putative disulfide-forming pathway folding initiation site, and potential role of C-peptide in folding process. J. Biol. Chem. 278, 17800–17809.

Richter, K., Haslbeck, M., Buchner, J., 2010. The heat shock response: life on the verge of death. Mol. Cell 40 (2), 253–266.

Sarkar, N.K., Kundnani, P., Grover, A., 2013. Functional analysis of Hsp70 superfamily proteins of rice (*Oryza sativa*). Cell Stress Chaperones 18 (4), 427–437.

Schumann, W., Ferreira, L.C.S., 2004. Production of recombinant proteins in *Escherichia coli*. Genet. Mol. Biol. 27 (3), 442–453.

Sung, D., Kaplan, F., Guy, C.L., 2001. Plant Hsp70 molecular chaperones: protein structure, gene family, expression and function. Physiol. Plantarum 113, 443–451.

Valdez-Cruz, N.A., Ramírez, O.T., Trujillo-Roldán, M.A., 2011. Molecular responses of *E. coli* caused by heat stress and recombinant protein production during temperature induction. Bioeng. Bugs 2 (2), 105–110.

Vierling, E., 1991. The roles of heat shock proteins in plants. Annu. Rev. Plant Physiol. Plant Mol. Biol. 42, 579–620.

Young, C.L., Britton, Z.T., Robinson, A.S., 2012. Recombinant protein expression and purification: a comprehensive review of affinity tags and microbial applications. Biotechnol. J. 7 (5), 620–634.

Zhang, K., Ezemaduka, A.N., Wang, Z., Hu, H., Shi, X., Lu, X., Fu, X., Chang, Z., Yin, C.C., 2015. A novel mechanism for small heat shock proteins to function as molecular chaperones. Sci. Rep. 5 (8811), 1–8.

Zhu, X., Zhao, X., Burkholder, W.F., Gragerov, A., Ogata, C.M., Gottesman, M.E., Hendrickson, W.A., 1996. Structural analysis of substrate binding by the molecular chaperone DnaK. Science 272, 1606–1614.