

# Enhancement of the Solubility of Proteins Overexpressed in *Escherichia coli* by Heat Shock

Jingqiu Chen<sup>1</sup>, Thomas B. Acton<sup>2,3</sup>, Soumit K. Basu<sup>1</sup>, Gaetano T. Montelione<sup>1,2,3</sup>, and Masayori Inouye<sup>1\*</sup>

<sup>1</sup>Department of Biochemistry, Robert Wood Johnson Medical School, 675 Hoes Lane, Piscataway, NJ 08854, USA

<sup>2</sup>Center for Advanced Biotechnology and Medicine

<sup>3</sup>Department of Molecular Biology and Biochemistry, Rutgers University, 679 Hoes Lane Piscataway, NJ 08854, USA

## Abstract

Protein misfolding resulting in the formation of inclusion bodies is one of the major problems during protein overexpression in *Escherichia coli*. In this paper, we introduce a new method, which is simply to heat shock a cell culture prior to protein induction, allowing effective enhancement of the solubility and thereby the yield of overexpressed proteins in *E. coli*. Using this method, we show that the solubility of the *E. coli* protein KsgA- $\Delta$ N is significantly increased when overexpressed from a T7 promoter. In addition, we also show that the solubility of several *Caenorhabditis elegans* proteins are also enhanced after heat-shock treatment when expressed in *E. coli*. Taken together, these results suggest that the “heat-shock protocol” is a generalizable and useful method for increasing the solubility of many proteins overexpressed in *E. coli*.

## Introduction

Protein overexpression systems, such as the T7 system, have been widely used for protein purification. However, due to improper folding of the overexpressed proteins, they often aggregate to form inclusion bodies, greatly reducing the protein yield. Different methods have been used to solve this problem, such as refolding of proteins after dissolving inclusion bodies in denaturing reagents, enhancing protein folding by culturing cells at low temperature, and expression as fusion proteins to assist protein folding (Guisse *et al.*, 1996; Mukhopadhyay *et al.*, 1997; Kapust *et al.*, 1999). Co-expression of heat-shock proteins with target proteins has also been reported to enhance proper protein folding, since heat-shock proteins, such as GroEL, GroES and DnaK, function as molecular chaperones to prevent protein misfolding (Gragerov *et al.*, 1992; Goenka *et al.*, 2001).

In this paper, we describe a simple method for enhancing soluble protein yields using *heat shock* prior to protein induction. Utilizing this method, protein solubility is significantly increased while induction levels are not affected at all or may even improve. The effectiveness of this method was found to vary widely from protein to protein. Nevertheless, we conclude that it is worthwhile to routinely evaluate this simple method of solubility enhancement before culturing cells on a large scale.

## Result

### Heat Shock can Enhance Protein Solubility without Affecting Induction

The heat shock effect on protein solubility was discovered during the course of purification of a mutant of *E. coli* KsgA, a 16S rRNA methyltransferase (van Buul *et al.*, 1985), in which 10 amino acid residues were deleted from the N-terminus (KsgA- $\Delta$ N). We were comparing the solubility of KsgA- $\Delta$ N from cells grown under different temperature conditions. Interestingly, we found that expression at 42°C resulted in greater protein solubility than at 37°C. To further investigate this phenomenon, we set a series of experiments involving this *heat shock*. As a control experiment, cells were grown in M9 medium at either 30°C or 37°C to 0.4 OD at 600 nm and then isopropyl- $\beta$ -D-thiogalactoside (IPTG) was added to a final concentration of 1 mM for protein induction. Cells were harvested when OD<sub>600</sub> reached 0.9 units (3.5 hr at 30°C and 2 hr at 37°C). For the heat shock experiments, cultures (80 ml each in a 250 ml culture flask) were first incubated at 37°C until OD<sub>600</sub> reached 0.1 units, and then shifted to a 42°C water bath with constant shaking. IPTG (1 mM) was added to each culture when OD<sub>600</sub> reached 0.4 units, and the cultures were then transferred to either 25, 30, 37 or 42°C water baths for further incubation. Cells were harvested after OD<sub>600</sub> reached 0.9 units. As shown in Figure 1A, at 30°C and 37°C without heat shock, only 50% and 40% of the induced protein is soluble, respectively (compare lanes 2 and 4 with lanes 1 and 3). Heat-shock treatment of this same construct significantly improves the yield of soluble KsgA- $\Delta$ N (Figure 1B and Figure 2). The soluble protein recovered in the supernatant fraction after centrifugation was 67, 65, 67, and 57% of the total KsgA- $\Delta$ N expressed at 25°C (lanes 1 and 2, Figure 2B), 30°C (lanes 3 and 4), 37°C (lanes 5 and 6), and 42°C (lanes 7 and 8), respectively. In addition to the observation that heat-shock treatment improved protein solubility, it should also be noted that the overall expression levels are not affected in heat-shock treated cells compared with that of non-heat-shock treated cells. Interestingly, inducing the expression of this gene at 25°C also

\*For correspondence. Email inouye@umdnj.edu; Tel. (732) 235-4115; Fax. (732) 235-4449/4783.

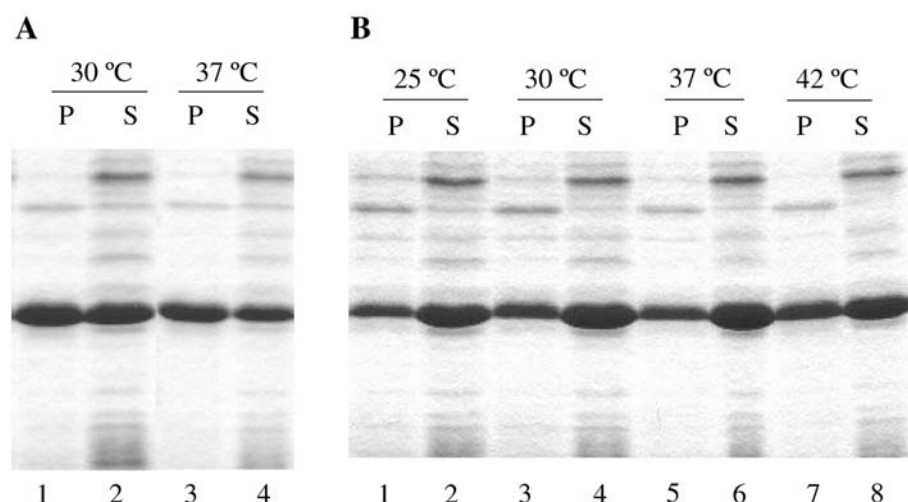


Figure 1. SDS-PAGE analysis of samples from cultures induced at different temperatures, either without heat shock (A), or with heat shock (B). Each sample was centrifuged after sonication, and loaded as insoluble pellet (P) and supernatant (S).

increases the percentage of soluble KsgA- $\Delta$ N and no further solubility improvement is found using heat-shock treatment when inducing at this temperature (data not shown). These data suggest that KsgA- $\Delta$ N is unstable at 37°C and the majority of the protein produced is misfolded resulting in the formation of inclusion bodies. The improved solubility of KsgA- $\Delta$ N observed after heat shock may be associated with the activity of native heat-shock proteins induced under these conditions.

#### The Enhanced Solubility is Maintained

In order to examine whether the duration of induction is critical for protein solubility, we investigated how long this enhanced solubility could be maintained by performing time course analysis after IPTG induction of cultures incubated with and without the heat-shock treatment. 1.5-ml aliquots were taken from both cultures (80 ml) at each time point and the amount of soluble and insoluble KsgA- $\Delta$ N protein was analyzed by SDS-PAGE. As shown in Figure 2 A and B, under the conditions used, protein solubility is enhanced throughout the time course of induction in cultures treated with heat shock. Though the growth curve measurement shows that cell growth is slightly inhibited by IPTG induction (Figure 2C), a phenomenon often observed in T7 systems, the *de novo* synthesis of KsgA- $\Delta$ N is still active at least 100 min after IPTG induction (Figure 2D) without affecting the overall percentage of soluble KsgA- $\Delta$ N in total KsgA- $\Delta$ N produced (Figure 2E), suggesting that the heat-shock factors required for KsgA- $\Delta$ N solubilization are stable at least for approximately 2 hr or even longer. Moreover, the analysis of the total cell lysates shows that the overall level of protein production is not negatively affected by the heat-shock treatment (data not shown).

#### Optimal Timing of the Heat-Shock Protocol

Next, we investigated whether the timing and duration of heat shock is important for its solubilizing effect. To test this, cultures at 37°C were shifted to 42°C at

different OD<sub>600</sub> readings. When heat shock is performed during early log phase (OD<sub>600</sub> of 0.06 units) to mid-log phase (OD<sub>600</sub> of 0.40 units) followed by 2.5 hr IPTG induction at 30°C, protein solubility is significantly enhanced (76% solubility) (compare lanes 1 and 2 of Figure 3A). A similar result (67% solubility) is seen when heat shock is started at a slightly later time point (OD<sub>600</sub> of 0.27 units) (compare lanes 3 and lane 4 of Figure 3A). But when heat shock is carried out from OD<sub>600</sub> of 0.27 units to 0.55 units, protein solubility is reduced (42% solubility) (compare lanes 5 and lane 6 of Figure 3A), indicating that the timing of heat shock is important for higher yields of soluble protein. In our case, we adopted the first heat-shock protocol, heating to 42°C during early to mid log phase (OD<sub>600</sub> = 0.05–0.5 units).

#### Heat-shock Treatment and Eukaryotic Proteins

Clearly, the solubility of KsgA- $\Delta$ N is significantly increased by heat-shock treatment prior to induction of protein expression. The possibility exists that this phenomenon is limited to bacterial proteins that are normally targets for the chaperones induced under high temperature. Alternatively, these chaperones may have the ability to work on a wide variety of proteins from other organisms, including eukaryotes. Therefore, we tested this strategy on ten genes from *Caenorhabditis elegans* to further investigate the ability of this method to help produce proteins with increased solubility. These proteins were originally selected for study as part of a pilot project in structural genomics and their sequences are available from the web site of the Northeast Structural Genomics Consortium ([www.nesg.org](http://www.nesg.org)).

In this strategy, *E. coli* BL21 (DE3) cells were transformed with expression vectors containing the ten different *C. elegans* open reading frames. The expression of which is under the control of a T7 promoter (Studier *et al.*, 1990). Two cultures of each construct (corresponding to a single ORF) were grown from OD<sub>600</sub>

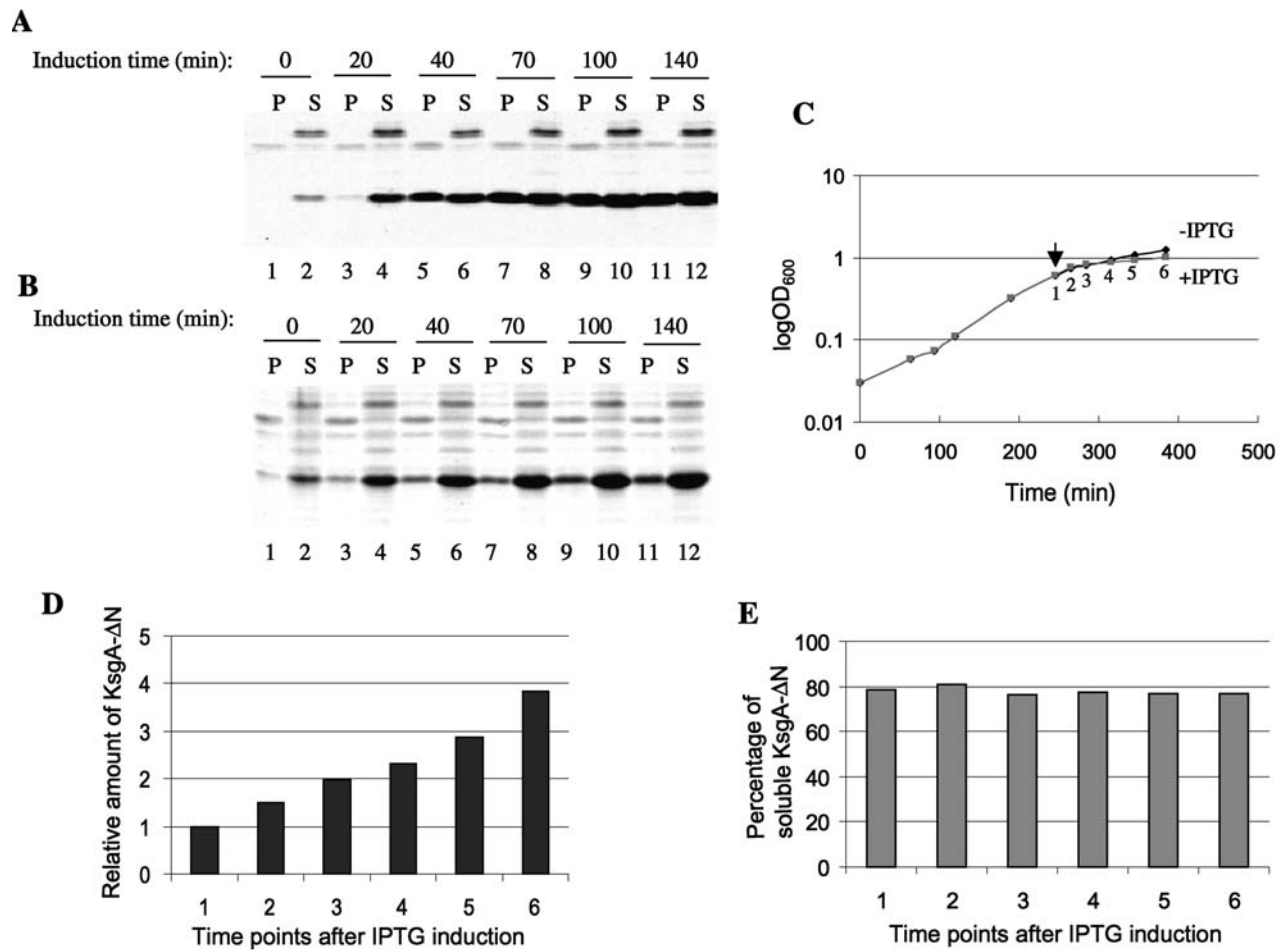


Figure 2. SDS-PAGE analysis of samples taken at different time points after IPTG induction, from non-heat-shocked culture (A) and heat-shocked culture (B). Both cultures were induced at 30°C. Each sample was centrifuged after sonication, and loaded as insoluble pellet (P) and supernatant (S). The growth curve of heat-shocked culture with or without IPTG induction was measured (C). The arrow indicates the point at which IPTG was added. Time points 1 to 6 are corresponding to the time points in panel (A) and (B). The amount of overexpressed KsgA-ΔN was also measured at each time point, and plotted as the relative value to the KsgA-ΔN amount at time point 1 (D). The solubility of KsgA-ΔN of each time point was shown as the percentage of the soluble form in the total KsgA-ΔN induced (E).

of 0.05 units to 0.5 units. The first culture was incubated at 37°C, while the second was incubated at 42°C. Once a culture reached an OD<sub>600</sub> of 0.5 units, it was shifted to 25°C, and allowed to equilibrate with respect to temperature. Expression was then induced with IPTG at a final concentration of 1 mM and incubation continued at 25°C for four hours. The results are shown in Figure 4 and Table 1, the percentage of soluble WR18 protein in the cell lysate doubles after heat-shock treatment (Figure 4A). A similar result is seen with the WR53 protein (Figure 4B). Quantitation of the amount of soluble protein as shown in Table 1 indicates that the solubility increases from a quarter to nearly half of the total induced protein. Two other *C. elegans* open reading frames (WR75 and WR81) also show significant increases in solubility after heat-shock treatment. In both of these cases, without heat-shock treatment most of the expressed protein is found in an insoluble form. However, heat-shock treatment before induction of protein expression results in ~20% soluble protein production (Figure 4C and 4D respectively and Table 1).

Although four proteins have significant increases in solubility after heat treatment, the solubility of the remaining six proteins tested are not affected. The results listed in Table 1 also show that the partially soluble proteins WR21, WR22, and WR78 do not increase in solubility after heat-shock treatment. The three remaining proteins in this study, WR26, WR50 and WR76 are completely insoluble when expressed in these bacterial cells and the application of our heat-shock protocol prior to induction does not enhance their solubility. Taken together these data indicate that although many proteins may have increased solubility with this method, not all proteins can be rescued by the heat-shock protocol. This is not surprising since there are many factors that can contribute to the partitioning of expressed protein between soluble fraction and insoluble inclusion body. It should be noted however, that although many of the proteins were not positively affected, no protein decreased in solubility after heat-shock treatment, nor were protein expression levels detrimentally affected.

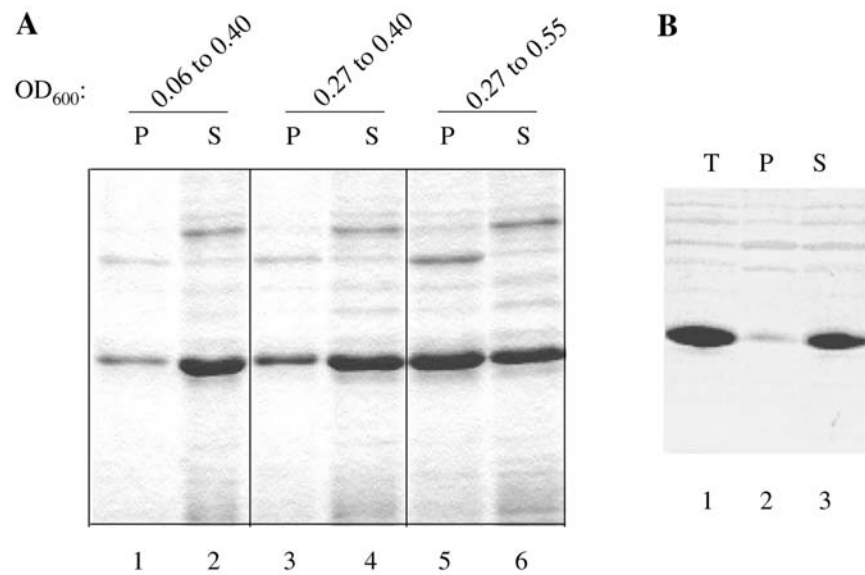


Figure 3. SDS-PAGE analysis of protein solubility when using different timing of heat-shock protocols prior to induction (A). Lanes 1 and 2—heat shock from  $OD_{600}$  0.06 to 0.4; lanes 3 and 4—heat shock from  $OD_{600}$  0.27 to 0.4; lanes 5 and 6—heat shock from  $OD_{600}$  0.27 to 0.55. Each sample was centrifuged after sonication, and loaded as insoluble pellet (P) and supernatant (S). Solubility analysis of a 1.5-ml aliquot from a 1.5-liter culture induction (B). Samples were loaded as total cell lysate (T), insoluble pellet (P) and supernatant (S).

#### Application to a Large-Scale Cell Culture

Since the primary purpose of this method is to increase the yield of properly folded protein in a large-scale culture, the heat-shock protocol was applied to a 1.5-liter culture in a 4-liter flask to express KsgA- $\Delta$ N. In our hands, it takes  $\sim 20$  min for the temperature of a 1.5-liter culture initially at  $37^{\circ}\text{C}$  to reach  $42^{\circ}\text{C}$  in a  $42^{\circ}\text{C}$  water bath. As we previously described, heat shock was initiated at an  $OD_{600}$  of 0.05 units by transferring the culture flask to a  $42^{\circ}\text{C}$  water bath, incubation at this temperature continued for 3 hr until an  $OD_{600}$  of 0.55 units was reached. The cultures were then shifted to  $30^{\circ}\text{C}$  and protein expression was induced by the addition of IPTG at a final concentration of 1 mM. The culture was then incubated for 90 min at this temperature ( $30^{\circ}\text{C}$ ). Over 90% of the KsgA- $\Delta$ N protein induced

under these conditions is soluble as shown in Figure 3B, with a yield of the approximately 100 mg/L soluble protein. Thus, our heat-shock protocol is applicable to large-scale protein production.

#### Discussion

For protein expression, high temperature is always considered to cause protein aggregation and thus is usually unfavorable (Thomas *et al.*, 1996). However, when performed prior to protein overexpression in a carefully controlled manner we have found that heat shock can actually rescue some proteins from going into inclusion bodies. In comparison to other methods, such as proteins refolding, the heat-shock strategy provides

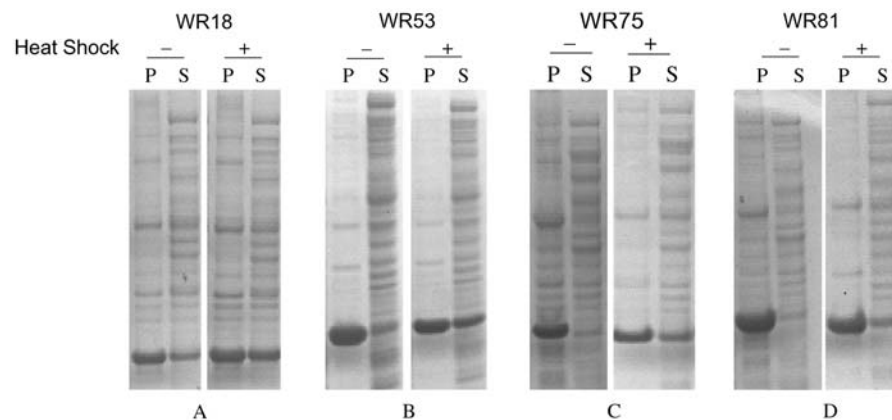


Figure 4. Solubility of several *C. elegans* genes after heat-shock treatment. Increased solubilities of WR18 (A), WR53 (B), WR75 (C) and WR81 (D) resulting from the “heat-shock protocol”. Each sample was centrifuged after sonication, and loaded as insoluble pellet (P) and supernatant (S) in equal amounts.

Table 1. Solubility and effects of heat shock on production of soluble *C. elegans* proteins

Protein	% Sol <sup>a</sup>	% Sol with Heat Shock <sup>a</sup>
WR18	15	30
WR21	80	80
WR22	80	85
WR26	0	0
WR50	2	2
WR53	20	40
WR75	10	20
WR76	5	15
WR78	85	85
WR81	0	20

<sup>a</sup>The relative amount of soluble expressed protein as measured by SDS-PAGE followed by densitometry.

a simple way to achieve increased protein solubility, and works quite well in many cases.

Although the mechanism of increased solubility is not completely clear, there are several attractive possibilities. One likely explanation is that heat-shock proteins or chaperones may play a role, as previously there have been reports of their ability to enhance protein solubility when co-expressed (Gragerov *et al.*, 1992; Goenka *et al.*, 2001). Although we did see the induction of both GroEL and DnaK by heat-shock treatment in two-dimensional gels, we doubt these heat shock induced proteins would have a half-life long enough to maintain the target protein solubility for 2 hr. In addition, it has been previously shown that adding ethanol to the growth media stimulates the expression of heat-shock proteins in a manner similar to that of a 42°C heat shock (Schlesinger *et al.*, 1982; van Bogelen *et al.*, 1987). In fact, this approach has been reported to increase the solubility of an overexpressed protein (Donnelly *et al.*, 2001). However, when we treated the media with 4% EtOH before induction, the solubility of the overexpressed KsgA-ΔN protein was not enhanced (data not shown). This suggests that the mechanism may be more complicated than simple chaperone induction.

The result obtained with the *C. elegans* ORFs reveals that the effect of heat-shock treatment on protein solubility is not limited to only those proteins native to *E. coli*. More importantly, these data suggest that our heat-shock protocol for increasing solubility is generalizable to the production of proteins from other organisms in a soluble form when overexpressed in *E. coli*. Further evidence for the generalization of this method to the proteins of other organisms when overexpressed in *E. coli*, is the result that a gene from *Legionella pneumophila* when expressed using the heat-shock strategy, also showed dramatically increased levels of solubility.

Although this simple “heat-shock protocol” is sometimes very effective in enhancing the solubility of expressed proteins in cell extracts it does not work on every protein that is overexpressed in *E. coli*. Indeed, even the same protein having mutations at different sites can behave quite differently. The KsgA N-terminal

mutant constructs described here are currently the best examples of “heat-shock enhanced” solubilizations. It is noteworthy however, that the solubilities of these same constructs can also be enhanced by lower temperature (25°C) inductions (data not shown). Ultimately, further investigation is needed into the types of proteins rescued by this method, as well as the conditions for solubility enhancement, the result of which should shed some light on the mechanisms at work in this useful strategy.

## Experimental Procedures

### Plasmids and Strains

*Escherichia coli* KsgA-ΔN coding sequences were cloned into pET expression plasmid pET17b. Cell strain BL21 (DE3) *ndk::cam* [*F<sup>-</sup>ompT gal[dcM] [lon] hsdS<sub>B</sub> (r<sub>B</sub><sup>-</sup> m<sub>B</sub><sup>-</sup>); an E. coli B strain*] with DE3, a λ prophage carrying the T7 RNA polymerase gene] was transformed with this plasmid and spread on LB plates with 50 μg/ml ampicillin. Plates were incubated at 37°C overnight. The *Caenorhabditis elegans* genes were PCR amplified from cDNA (Yuji Kohara, National Institute of Genetics, Mishima Japan) and cloned into one of several different modified pET expression plasmids (Novagen). The Genebank accession number of these ORFs are as follows: WR18(B0024.12), WR21(C71341), WR22(D71263), WR26(C50788), WR50(AV189911), WR53(C48848), WR75(F56D1.3), WR76(F59E10.3), WR78(K11H12.1) and WR81(ZK512.8). The corresponding protein sequences are also available at the following web site ([www.nesg.org](http://www.nesg.org)). Expression was carried out in BL21 (DE3) pMgK, which contains a plasmid expressing the genes for three rare tRNAs.

### Cell Cultures

For KsgA-ΔN expression, cells were first inoculated into 5 ml M9 media (1.28% Na<sub>2</sub>HPO<sub>4</sub>·7H<sub>2</sub>O, 0.3% KH<sub>2</sub>PO<sub>4</sub>, 0.05% NaCl, 0.1% NH<sub>4</sub>Cl, 0.04% glucose, 0.2% casamino acids, 20 mM MgSO<sub>4</sub>, 1 μg/ml vitamin B<sub>1</sub>, 0.1 mM CaCl<sub>2</sub>, 50 μg/ml Ampicillin) and incubated shaking at 37°C overnight. These overnight cultures were then diluted 50-fold into 80 ml fresh M9 media (1.28% Na<sub>2</sub>HPO<sub>4</sub>·7H<sub>2</sub>O, 0.3% KH<sub>2</sub>PO<sub>4</sub>, 0.05% NaCl, 0.1% NH<sub>4</sub>Cl, 0.4% glucose, 0.2% casamino acids, 20 mM MgSO<sub>4</sub>, 1 μg/ml vitamin B<sub>1</sub>, 0.1 mM CaCl<sub>2</sub>, 50 μg/ml Ampicillin) for expression. For *C. elegans* proteins, cells were first inoculated into 5 ml LB medium (1% tryptone, 0.5% yeast extract, 1% NaCl, 2% glucose, 50 μg/ml Ampicillin) for an overnight culture at 37°C. The overnight cultures were diluted 50-fold into fresh LB medium. A final concentration of 1 mM IPTG was added during log phase for protein induction.

### Preparation of Cell Lysate and Protein Solubility Analysis

Aliquots of cultures (1.5 ml) were first spun down in 1.5-ml microcentrifuge tubes (15,800 × g, 2 min). Cell pellets were then resuspended in 50 μl buffer A (25 mM Tris-Cl, 5 mM MgCl<sub>2</sub>, 5 mM β-mercaptoethanol), and subjected to sonication (15 sec × 4). These samples were further centrifuged (15,800 × g, 10 min) to separate insoluble inclusion bodies (pellet) and soluble proteins (supernatant). Each supernatant was transferred into a new microcentrifuge tube, and the volume measured by pipette. An equal volume of buffer A was then used to resuspend the corresponding insoluble pellet. For solubility analysis by SDS-PAGE, an equal volume was taken from either the supernatant or insoluble pellet suspension, mixed with 5X loading buffer, and loaded on a 15% SDS-PAGE gel.

### Protein Amount Measurement

Samples were applied to a 15% SDS-PAGE gel, and stained by Coomassie Blue. Gels were scanned after destaining. The amount of KsgA-ΔN in each sample was determined by gel analysis software (written by Yukihiko Yabuta).

### Large Culture Application

A 30-ml overnight culture of BL21/ KsgA-ΔN in M9 medium was prepared as described above, and diluted into 1.5-liter fresh M9 medium in a 4-liter flask. IPTG was added to 1 mM final concentration, inducing protein expression. Upon harvest, a 1.5 ml aliquot was taken and subjected to centrifugation at 15,800 × g for 2 min. The pellet was suspended in 100 μl Buffer A followed by sonication and centrifugation. SDS-PAGE samples were prepared as described above.

#### Acknowledgement

This work has been supported by Grant GM19043 and Grant GM62413-01 from the National Institutes of Health.

#### References

- Donnelly, M.I., Stevens, P.W., Stols L., Su, S.X., Tollaksen, S., Giometti, C., and Joachimiak, A. 2001. Expression of a highly toxic protein, Bax, in *Escherichia coli* by attachment of a leader peptide derived from the GroES cochaperone. *Protein Expr. Purif.* 22: 422–429.
- Goenka, S., and Rao, C.M. 2001. Expression of recombinant  $\zeta$ -crystallin in *Escherichia coli* with the help of GroEL/ES and its purification. *Protein Expr. Purif.* 21: 260–267.
- Gragerov, A., Nudler, E., Komissarova, N., Gaitanaris, G.A., Gottesman, M.E., and Nikiforov, V. 1992. Cooperation of GroEL/GroES and DnaK/DnaJ heat shock proteins in preventing protein misfolding in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA.* 89: 10341–10344.
- Guise, A.D., West, S.M., and Chaudhuri, J.B. 1996. Protein folding in vivo and renaturation of recombinant proteins from inclusion bodies. *Mol. Biotechnol.* 6: 53–64.
- Kapust, R.B., and Waugh, D.S. 1999. *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci.* 8(8): 1668–1674.
- Mukhopadhyay, A. 1997. Inclusion bodies and purification of proteins in biologically active forms. *Adv. Biochem. Eng. Biotechnol.* 56: 61–109.
- Schlesinger, M.J., Ashburner, M., and Tissières, A. 1982. Heat shock, from bacteria to man. Cold spring harbor laboratory.
- Studier F.W., Rosenberg A.H., Dunn J.J., and Dubendorff J.W. 1990. Use of T7 RNA polymerase to direct expression of cloned genes. *Meth. Enz.* 185: 60–89.
- Thomas, J.C., and Baneyx, F. 1996. Protein misfolding and inclusion body formation in recombinant *Escherichia coli* cells overexpression heat-shock proteins. *J. Biol. Chem.* 271: 11141–11147.
- van Bogelen, R.A., Kelley, P., and Neidhardt, F.C. 1987. Differential induction of heat shock, SOS, and oxidation stress regulons and accumulation of nucleotides in *Escherichia coli*. *J. Bacteriol.* 169: 26–32.
- van Buul, C.P.J.J., and VanKnippenberg, P.H. 1985. Nucleotide sequence of the KsgA gene of *Escherichia Coli*: comparison of methyltransferases affecting dimethylation of adenosine in ribosomal RNA. *Gene.* 38: 65–72.