Troubleshooting GST fusion protein expression in *E. coli*

M. Saluta and P. A. Bell
Amersham Pharmacia Biotech, Piscataway, NJ USA

RECOMBINANT PROTEIN EXPRESSION

The production of soluble, intact GST fusion proteins requires optimization of several factors. This discussion focuses on choosing the proper host strain, growth temperature, cell density at time of induction and length of induction for expression of a GST-luciferase fusion protein. Examples demonstrate that for this fusion protein, *E. coli* BL21 grown at 28 °C gave superior results compared with other strains and growth conditions.

Introduction

Protein expression in *E. coli* provides the ability to obtain large amounts of a desired protein which can be used for a wide range of studies including antigen and vaccine production; molecular immunology; and structural, biochemical and cell biology studies. One of the more extensively used systems for the expression and purification of recombinant proteins is the Glutathione-S-Transferase (GST) Gene Fusion System (1). The GST Gene Fusion System provides an integrated system for the expression, purification and detection of glutathione-S-transferase fusion proteins using *E. coli*. Although the production of recombinant fusion proteins in *E. coli* is well established, there are numerous factors which may present obstacles for successful production and purification of full-length, soluble fusion proteins (2). Following are guidelines for troubleshooting a few of these obstacles including choosing the proper host strain for cloning and expression, monitoring growth conditions and optimizing soluble expression.

Cloning and transformation

Expression of inserts cloned into a pGEX vector is under the control of the IPTG-inducible *tac* promoter. All pGEX vectors are also engineered with an internal *lacI*^{q} gene. The LacI^q gene product is a repressor protein that binds to the operator region of the *tac* promoter preventing expression until induction by IPTG, thus maintaining tight control over expression of the insert.

Under induction conditions, pGEX vectors provide the ability for high-level expression of a fusion protein. However, even under non-induced conditions, basal levels of expression ("leaky" expression) from a *lac* promoter located between the 3'-end of the *lacI*^{q} gene and the *tac* promoter may interfere with the ability to isolate clones containing the insert in the proper orientation, an outcome which may persist if the insert is toxic to the host. Basal-level expression, from the *lac* promoter, can be minimized by incorporating a catabolite such as glucose into the growth

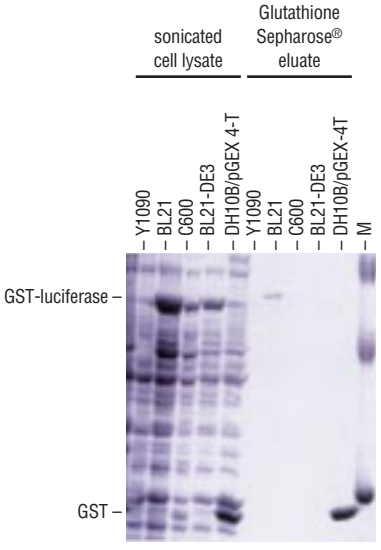

Figure 1. Host strains and fusion protein solubility. Various strains of *E. coli* were transformed with pGEX-4T containing a luciferase gene insert. Cultures were grown, induced and lysed using standard conditions (10). Fusion proteins were purified using Glutathione Sepharose 4B. *E. coli* DH10B transformed with pGEX-4T was used as a control for GST expression. M = molecular weight marker.

medium. Catabolite repression, in the presence of 2% glucose, will affect basal-level expression from the *lac* promoter but will not affect expression from the *tac* promoter upon induction with IPTG. Thus, it is recommended that transformations, stocks and overnight seed cultures should be prepared in the presence of glucose in order to minimize the effects of leaky expression.

Host strains

Although a wide variety of *E. coli* host strains can be used for cloning and expression with the pGEX vectors, there are specially engineered strains which are more suitable for expressing fusion proteins and that may maximize expression of full-length fusion proteins. Strains deficient in known cytoplasmic protease gene products, such as Lon, OmpT, DegP or HtpR may aid in the expression of fusion proteins by minimizing the effects of proteolytic degradation by the host (3–6). An example of such a host is *E. coli* strain BL21, a strain defective in OmpT and Lon protease production. Host strains may in some cases also help to increase the yield of intact soluble fusion proteins. Figure 1 gives an example in which several host strains were tested for their ability to express an intact GST fusion protein in the soluble fraction. For this particular construct the results indicate that of the strains tested, BL21 appears to be the only strain able to express the fusion protein in a soluble, intact form.

Enhancing soluble expression

The high-level expression of foreign fusion proteins in *E. coli* often results in the formation of an insoluble product, termed an inclusion body. An inclusion body refers to the formation of a dense precipitated aggregate in which the majority of expressed fusion protein is complexed with RNA, rendering it impervious to the action of most detergents (7). Although the formation of inclusion bodies can be used as a means to purify an active form of an expressed fusion protein which otherwise may be unstable in the soluble fraction, the steps needed to solubilize and refold the fusion protein can be highly variable and may not always result in high yields of active protein. Alternatively, there are a variety of growth parameters which can be investigated, either solely or in combination with each other, that may provide a greater yield of non-degraded fusion protein in the soluble fraction. These include:

- Lowering the growth temperature to between 20 °C and 30 °C
- Inducing for a shorter period of time
- Inducing at a higher cell density for a shorter period of time
- Increasing aeration

Figures 2 and 3 demonstrate the effect that temperature plays on solubility and degradation of a GST fusion protein. In both cases, the cultures were induced with 0.1 mM IPTG at increasing cell densities for various lengths of time. With growth at 37 °C (Figure 2), the yields of intact soluble fusion protein are compromised at all ranges of cell densities and induction conditions—approximately 50% of the fusion protein is degraded. However, when the growth temperature is reduced to 28 °C (Figure 3) the yield of intact soluble fusion protein is considerably greater for all densities and induction times tested. In general, induction at lower cell densities (A_{600} = 0.5) usually results in greater yields of the fusion protein in a soluble form. Also note that in some cases it may be beneficial to grow the cells up to a high cell density (>1 A_{600} unit) and omit induction by IPTG. Although lower yields may be expected using this approach, more of the fusion protein is obtained in an intact form (last lane of Figure 3).

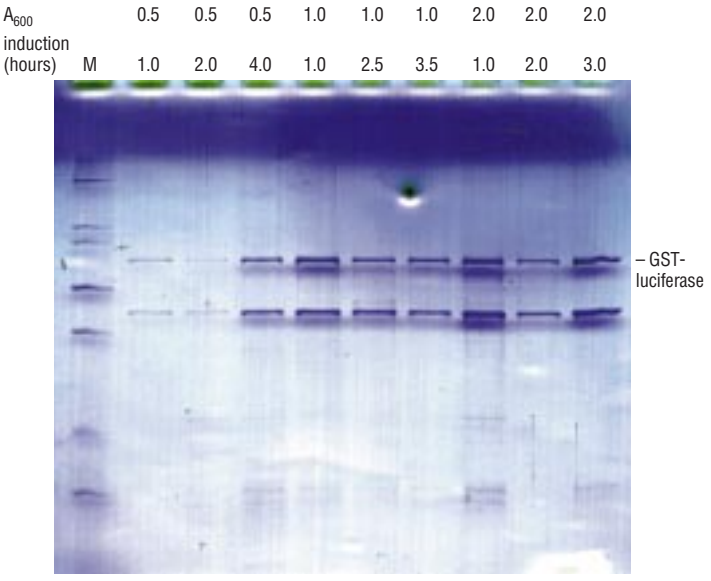

Figure 2. Expression of GST-luciferase fusion protein at 37 °C. *E. coli* BL21 transformed with pGEX-4T-luc were induced with 0.1 mM IPTG at the cell densities and for the times indicated. Cells were lysed by sonication and fusion proteins purified using Glutathione Sepharose 4B (10). M = molecular weight marker.

Figure 3 also shows possible effects which can be observed upon over-sonication. Excessive sonication can result in the co-purification of *E. coli* host proteins and may also lead to denaturation and breakdown of the fusion protein (1). In this figure, the third and fourth lanes (A_{600} = 0.5, induction time = 2.0 hr) contain samples from lysates which were obtained from cultures prepared under identical growth conditions. However, the cells obtained for the third lane were over-sonicated, resulting in a higher degree of breakdown products. One of the

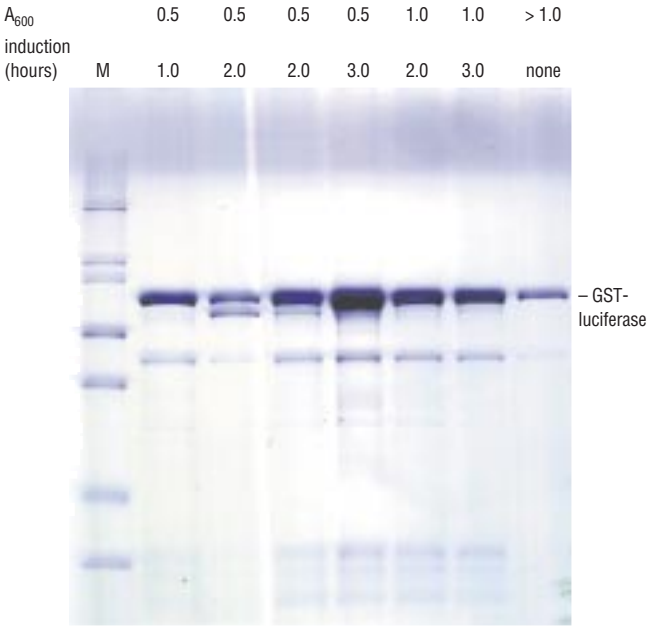

Figure 3. Expression of GST-luciferase fusion protein at 28 °C. *E. coli* BL21 transformed with pGEX-4T-luc were induced with 0.1 mM IPTG at the cell densities and for the times indicated. Cells were lysed by sonication and fusion proteins purified using Glutathione Sepharose 4B (10). M = molecular weight marker.

more common host proteins that co-purifies with GST fusions is DnaK. Separation of DnaK from the GST fusion protein can be achieved by ion exchange chromatography after purification using Glutathione Sepharose™ 4B (9).

Solubilizing fusion proteins

Another approach to enhancing solubilization is by using a combination of Triton™ X-100 and sarkosyl introduced during the lysis stage after cell harvest. This method, described in detail in a 1993 article by Frangione and Neel (8) involves solubilizing the protein in the presence of sarkosyl and EDTA prior to sonication, followed by clarifying the lysate and adding Triton X-100 prior to purification over Glutathione Sepharose. This process may aid in the solubilization of fusion proteins prior to affinity purification. The ratio of sarkosyl:Triton X-100 is empirically determined in order to optimize both solubilization of the protein and binding to Glutathione Sepharose.

Conclusion

In all, successful expression of any foreign fusion protein in *E. coli* depends upon a myriad of events which collectively result in obtaining high yields of a soluble protein whose antigenicity and functionality has been maintained throughout the expression and purification process. The conditions required to obtain optimal yields of a fusion protein will need to be empirically determined for each fusion construct and may require a variety of growth and expression parameters to be investigated.

References

1. Ausubel, F. M. *et al.*, eds. *Current Protocols in Molecular Biology* **Vol. 2**, 16.0.1 (1996).
2. Smith, D. B. and Johnson, K. S., *Gene* **67**, 31-40 (1988).
3. Baker, T. A. *et al., Proc. Natl. Acad. Sci., USA* **81**, 6779-6783 (1984).
4. Strauch, K. L. and Beckwith, J., *Proc. Natl. Acad. Sci., USA* **85**, 1576-1580 (1988).
5. Grodberg, J. and Dunn, J. J., *J. Bacteriol.* **170**, 1245-1253 (1988).
6. Sugimura, K. and Higashi, N., *J. Bacteriol.* **170**, 3650-3654 (1988).
7. Schein, C. H., *Bio/Technology* **7**, 1141 (1989).
8. Frangioni, J. V. and Neel, B. G., *Anal. Biochem.* **210**, 179-187 (1993).
9. Amersham Pharmacia Biotech, *Science Tools* **2**(2), 26 (1997).
10. *GST Gene Fusion System Manual* (code number 18-1123-20), Amersham Pharmacia Biotech, 1997.