Published in final edited form as:

Methods Mol Biol. 2011 ; 681: 259–280. doi:10.1007/978-1-60761-913-0_14.

Purification of proteins fused to glutathione S-transferase

Sandra Harper and David W. Speicher, Ph.D.

The Wistar Institute, Philadelphia, PA, USA

Abstract

This chapter describes the use of glutathione S-transferase (GST) gene fusion proteins as a method for inducible, high-level protein expression and purification from bacterial cell lysates. The protein is expressed in a pGEX vector, with the GST moiety located at the N-terminus followed by the target protein. The use of GST as a fusion tag is desirable because it can act as a chaperone to facilitate protein folding, and frequently the fusion protein can be expressed as a soluble protein rather than in inclusion bodies. Additionally, the GST fusion protein can be affinity purified easily without denaturation or use of mild detergents. The fusion protein is captured by immobilized glutathione and impurities are washed away. The fusion protein then is eluted under mild, non-denaturing conditions using reduced glutathione. If desired, the removal of the GST affinity tag is accomplished by using a site-specific protease recognition sequence located between the GST moiety and the target protein. Purified proteins have been used successfully in immunological studies, structure determinations, vaccine production, protein-protein, and protein-DNA interaction studies and other biochemical analysis.

Keywords

Glutathione S-transferase (GST); pGEX; protein expression; protein purification; thrombin; factor Xa; fusion tags

1. Introduction

Glutathione S-transferase (GST) is a naturally occurring 26 KDa protein found in eukaryotic cells. The gene from the parasitic helminth *Schistosoma japonicum* was used in the development of the pGEX vectors (1). This unit describes the use of a GST affinity tag to aid in the purification of recombinant proteins. The 26KDa GST moiety binds with high affinity to glutathione coupled to a Sepharose matrix. This binding is reversible and the protein can be eluted under mild, non-denaturing conditions by the addition of reduced glutathione to the elution buffer. A specific protease site engineered between the GST moiety and the protein of interest allows removal of the GST moiety from the target recombinant protein. The GST can be removed from the sample by re-chromatography on a glutathione column, and the protein of interest purified to homogeneity by other techniques such as gel filtration or ion exchange. The most commonly used vectors are available from GE Healthcare (and are discussed here), although variations from other suppliers also are available.

²⁹In a batch mode protease cleavage, add an empirically determined amount of enzyme to the fusion protein bound to the resin (see Note 15 a–d). Use 1 ml of PBS-containing enzyme per ml of bed volume. Incubate at room temperature for several hours with gentle agitation. Centrifuge 500 × g for 5 min to sediment the resin. Remove the super which contains the target protein to a separate tube. Wash the glutathione Sepharose with PBS up to 3 times to recover the cleaved fusion protein. Incubate the resin with glutathione buffer to remove GST and residual fusion protein. Use 1 ml glutathione buffer per ml of resin. Centrifuge 500 × g and recover eluted GST. Repeat up to 3 times. Analyze samples by SDS-PAGE.

Proteins produced using the GST fusion system have been used in numerous biological applications. The use of GST fusion proteins has been successful in both NMR and crystallography structure determinations. Generally, the GST tag is removed prior to these determinations, but several crystal structures exist for fused protein (2, 3). Other applications include the study of protein-protein interactions via GST pull-down assays (4). Proteins purified using the GST system also have been used successfully in immunological studies and vaccine production (5). Recently, high-throughput proteomics studies have used GST fusion proteins for directional immobilization in protein microarrays (6) and surface plasmon resonance (7). Successful structure-function studies involving protein-protein and DNA-protein interactions also have been described (8).

Successful purification of GST fusion proteins requires several strategic decisions and may require optimization of methods and conditions for specific proteins. A flow diagram highlighting the basic steps of the vector design, expression and purification processes and some of the key decisions to be made is shown in Figure 1. Each step involves multiple options that are often interrelated and could affect the final product yield and purity. For instance, choice of vector will be influenced by whether the GST moiety ultimately will be cleaved away from the target protein and the desired protease to be used. Something to consider when evaluating whether or not to cleave the GST moiety away is that GST is a homodimeric protein. If it is suspected that oligomeric state will influence properties of the target protein, the GST moiety should be cleaved away. Once a vector has been chosen and successful cloning has occurred, optimization of protein expression conditions such as *E. coli* host cell strain, temperature, concentration of IPTG, and length of induction should be optimized. Sample extraction usually is performed with sonication or French Press; however, some proteins may require more gentle extraction techniques. Several immobilized glutathione Sepharose media are available for use in the purification. Products in lab packs contain bulk resin that can be used in a batch-binding mode or with low-pressure chromatography columns that utilize gravity flow or a peristaltic pump; 96-well filter plates provide a convenient method for high-throughput screening; and prepacked HiTrap and HiPrep columns can be used with a medium-pressure chromatography system (e.g. FPLC), peristaltic pump, or a syringe. Purification media choice largely depends on ultimate use of the protein and the scale of the purification and convenience of the product choices. Detection with SDS-PAGE will give information on sample size and purity, throughout the expression and purification procedure. In cases of low product yield, Western blotting can be used to detect the expressed protein. Cleavage of the target protein from the GST moiety will depend on the vector chosen and can be performed either in solution or while the protein is bound to the column. The enzyme should be inactivated and/or removed from the sample after cleavage. Depending on the sample purity required for ultimate use, the protein should be further purified after removal of the GST moiety using gel filtration, ion exchange, or other purification schemes.

Here we provide a detailed description of the protocols required for successful purification of GST fusion proteins, with an emphasis on maintaining the solubility of the fusion protein and avoiding denaturants or mild detergents. The protocols include basic information on the choice of expression vectors and host *E. coli* cells, detailed procedures for protein expression and purification of expressed protein, enzymatic cleavage of the GST moiety from the target protein, and subsequent steps to purify the target protein to homogeneity (see Note 1).

---

¹Detailed protocols and troubleshooting tips for all aspects of the GST gene fusion protein expression system are available from GE Healthcare in the handbooks: *GST Gene Fusion System*, product number 18-1157-58, and *The Recombinant Protein Handbook*, product number 18-1142-75. The handbooks are available for purchase in hard copy or can be accessed as a PDF on the GE Healthcare website.

2. Materials

Use Milli-Q-purified water or equivalent for the preparation of all buffers

2.1 pGEX gene fusion construction

1. pGEX vector (see Figure 2)
2. A cDNA or PCR product containing the coding sequence for the protein of interest and appropriate adjacent restriction sites that are compatible with the selected pGEX vector
3. Restriction enzymes, agarose, 1X TAE buffer, T4 DNA ligase
4. Competent *E. coli* cells (see Note 2)
5. LB agar plates with 50 µg/ml ampicillin

2.2 Expression of GST fusion protein

1. LB medium per liter: 10 g tryptone, 5 g yeast, 5 g NaCl. Adjust to pH 7.2 with NaOH. Autoclave.
2. Ampicillin, 5 mg/ml stock, filter sterilize
3. Glycerol stock of transformed *E. coli* cells expressing GST fusion protein in pGEX vector
4. Isopropyl-1-thio-β-D-galactopyranoside (IPTG), 100 mM stock

2.3 Affinity purification of GST fusion protein

1. Glutathione Sepharose 4B bulk matrix (GE Healthcare or Sigma) (see Note 3)
2. Column (BioRad 2.5 cm × 10 cm Econo-Column)
3. 10 mM glutathione buffer: 50 mM Tris, 10 mM reduced glutathione, pH 8.0 (make fresh daily)
4. PBS/EDTA/PMSF: 1X PBS, 5 mM EDTA, 0.15 mM PMSF, pH 7.4
5. Pelleted *E. Coli* containing expressed target fusion protein
6. PBS/EDTA: 1X PBS, 5 mM EDTA, pH 7.4
7. Lysis buffer: 50 mM Tris, 50 mM NaCl, 5 mM EDTA, 1 µg/ml leupeptin, 1 µg/ml pepstatin, 0.15 mM PMSF, 1 mM DFP, 1 mM 2-ME, pH 8.0. *Caution: DFP is a very dangerous neurotoxin; see Note 4 and supplied precautions.* (see Note 5).

²The use of JM105 or a similar strain is recommended for cloning and maintenance of the vector. Strains of BL21 and its derivatives, or another protease deficient strain, are recommended for maximal protein expression. Strains deficient in cytoplasmic protease gene products such as lon, ompT, degP or htpR may minimize the effects of proteolytic degradation of overexpressed protein by the host. Do not use a strain of *E. coli* that carries the rec1A allele to propagate pGEX plasmids, as deletions or rearrangements of plasmid DNA have been reported.

³Glutathione Sepharose 4B bulk media is used in this purification protocol. This chromatography media is ideal for screening conditions and allows for easy scale-up. Purifications are performed easily using either gravity flow or a low-pressure chromatography system, or may be used in batch purifications based on centrifugation. GSTrap FF affinity columns are 1 ml or 5 ml prepacked columns that also are available and may be connected in series for scale-up. These columns are for use with a liquid chromatographic system, a pump, or syringe. In addition to these formats, glutathione Sepharose also is available in spin columns and 96-well plate formats for rapid screening of GST fusion protein expression and purification conditions.

⁴DFP is an extremely dangerous neurotoxin. Follow the precautions supplied by the manufacturer and only handle the neat reagent in a chemical fume hood.

2.4 Enzymatic cleavage to remove GST affinity tag

1. Affinity-purified fusion protein
2. Vector-specific enzyme (thrombin, factor Xa, or PreScission protease)
3. 0.15 M PMSF in isopropanol
4. PBS/EDTA/PMSF: 1X PBS, 5 mM EDTA, 0.15 mM PMSF, pH 7.4

2.5 Removal of GST affinity tag after protease cleavage

1. Cleaved fusion protein dialysed into PBS/EDTA/PMSF
2. Glutathione Sepharose column
3. PBS/EDTA/PMSF

2.6 Further purification of cleaved recombinant protein

1. Cleaved fusion protein
2. PBS
3. Concentrator (Amicon Ultra)
4. HPLC system with in-line detector and fraction collector
5. Gel filtration column compatible with the molecular weight range of the sample to be purified

3. Methods

3.1 pGEX gene fusion construction

There are multiple pGEX gene fusion vectors available from GE Healthcare. A shared characteristic of all the vectors include the presence of a *tac* promoter for chemically inducible, high-level protein expression with IPTG. An internal *lacI^q* gene helps to maintain tight control over expression of the insert by binding to the *tac* promoter until induction with IPTG. The available fusion vectors are designed so that the polypeptide of interest can be inserted immediately after the GST gene using a polylinker site (see Figure 2). A protease cleavage site is located between the GST sequence and the polylinker site so that the GST moiety can be enzymatically removed, if desired. Although not depicted in the vector diagram for all the different vectors, the polylinker sites are followed by stop codons in each reading frame. Although expression in *E. coli* is simple, economical, and efficient, it should be noted *E. coli* do not posttranslationally modify most proteins. If posttranslational modifications are required for function, expression in baculovirus (9) and yeast (10) have also been successful.

Choosing an appropriate vector is based on the ultimate use of the protein. If the GST moiety is to be cleaved away from the protein of interest, the pGEX-T series contains a protease cleavage site for thrombin, the pGEX-X series of vectors contain protease cleavage sites for factor Xa, and the pGEX-P series contains a cleavage site for PreScission protease (see Figure 2). When selecting the protease cleavage site, be sure that the protein of interest does not contain an internal recognition sequence for this protease. One of the advantages of using the thrombin recognition site is that it is generally cost effective, as relatively small

---

⁵Addition of protease inhibitors to the lysis buffer will help to prevent proteolytic degradation of the target protein during extraction. However, the exact cocktail of protease inhibitors added to the lysis buffer should be tailored to the characteristics of the target protein. Reducing agents such as 2-ME, DTT, or TCEP at concentrations of 1–10 mM should be added to the buffer as needed. Addition of a nonionic detergent, such as 1% Triton X-100, also may be added to the buffer to aid in extraction.

amounts of thrombin and short incubation times at 37 °C are sufficient to cleave the protein with high efficiency. Factor Xa has very high specificity, but is expensive and generally requires high enzyme-to-substrate ratios for efficient cleavage. PreScission protease is exclusive to GE Healthcare but has several advantageous characteristics, namely it is effective at low temperature (5 °C) and it is also a GST fusion protein, thereby facilitating removal after cleavage. An additional factor to consider when selecting a vector and the restriction sites that will be utilized is the sequence that remains at the N-terminus of the protein after removal of the GST moiety.

In this protocol, a set of oligonucleotides, including the appropriate restriction sites, are used to amplify the target gene by PCR. Detailed descriptions of the preparation of the cDNA inserts and creation of the gene fusion is beyond the scope of this discussion, but rely on standard molecular biology techniques such as those described in (11, 12). After digestion with restriction enzymes, the PCR product is ligated into the pGEX vector and transformed into an *E. coli* host. Several of the transformants should be grown in mini-culture preparations to check for expression of the target protein. Glycerol stocks of the bacterial culture should be prepared for storage at –80 °C. Once positive expression is confirmed, DNA sequencing should be performed to ensure no errors were introduced during the PCR amplification or ligation.

3.1.1 Cloning the GST fusion protein

1. Amplify the target protein sequence by PCR. Be sure to introduce restriction sites at the ends of the target gene fragment that are in-frame and complementary to the vector chosen.
2. Digest the PRC product and the chosen vector with compatible restriction enzymes.
3. Clean digested products by gel electrophoresis followed by a PCR clean-up kit (i.e., Qiagen).
4. Ligate then transform into appropriate, competent *E. coli* cells (see Note 2). Grow transformants overnight at 37 °C on LB agar plates supplemented with ampicillin.
5. Screen several colonies to verify that the insert is in the proper orientation and in the correct reading frame. Transfer several single colonies to separate tubes containing resuspended Ready-to-Go-PCR beads with 10 pmol each pGEX 5' and 3' sequencing primers added. After transferring the colony, streak some of the remaining bacteria on an LB agar plate. This can be used as a source for the mini-protein expressions. Overlay with mineral oil and perform PCR. Analyze 10 µl each by agarose gel electrophoresis.
6. Select several individual colonies that are positive by PCR and grow a separate 5 ml mini-culture of each selected transformant to screen for expression of the target protein (see Note 6).
7. Glycerol stocks may be prepared by mixing equal volumes of log phase bacterial culture and 70% glycerol. Invert several times to mix. Store at –80 °C.
8. Verify the integrity of the target gene sequence by DNA sequence analysis.

3.2 Expression of GST fusion protein

One of the driving forces behind the development of the GST fusion system is that the GST protein accumulates in the cytoplasm (1) and, due to the large size of this affinity tag, it is expected that the solubility of the GST moiety would usually propagate to the fusion partner—thereby eliminating the need for lengthy purification schemes for insoluble proteins that usually involve denaturants. Unfortunately, in practice, it is unpredictable which fusion

proteins will remain soluble and which will accumulate in occlusion bodies. As a general rule of thumb, the larger the protein and the more complex folding it must undergo, the more likely it will accumulate in inclusion bodies. In most cases, where the fusion protein is initially observed primarily in inclusion bodies, it is desirable to modify culture conditions—such as lowering growth temperature, increasing aeration, or altering induction conditions to—attempt to obtain soluble protein, rather than to solubilize inclusion bodies with denaturants and refold the protein after extraction. Hence, before conducting a large-scale protein expression it is recommended to complete a pilot experiment to check solubility (see Note 7). Since the main determinant of successful purification of functional GST fusion proteins in good yields is protein solubility, it is worth the effort to optimize culture conditions prior to large-scale purification experiments (see Note 8).

Yields of GST fusion proteins using expression in *E. coli* are highly variable, often ranging from 10 to 50 mg/liter, but could potentially be much lower in cases where the fusion protein is toxic to the cells or unstable (see Note 9). Since expression levels are typically high, adequate amounts of protein usually can be obtained from several liters or less of bacterial culture grown in shaker flasks. The following protocol describes protein production at 37 °C; however, exact temperature and induction conditions for each construct should be modified to improve the yield of soluble protein. Cell density can be monitored using A₆₀₀ (see Note 10) and protein expression can be analyzed using SDS-PAGE of un-induced and induced samples (see Note 6). Once the culture has been induced with IPTG, do not allow

---

⁶Following is a screening protocol to check protein expression levels of fusion protein in mini-cultures. The presence of an increase in protein levels at the expected molecular weight on an SDS-PAGE gel after induction with IPTG is a good indication of successful protein expression. After induction, a prominent band at the expected molecular weight (molecular weight of target protein + ~ 26 KDa for the GST moiety) should be visible. If it is suspected that the protein is expressed at a low level, verification of correct expression can be accomplished using Western blotting with an antibody specific for either the target protein or GST. If the samples will not immediately be analyzed by SDS-PAGE, they may be stored overnight at 4 °C or at –20 °C for longer-term storage.

- a. Inoculate several isolated colonies into separate 50 ml centrifuge tubes containing 10 ml LB supplemented with 100 µg/ml ampicillin. Grow an overnight culture at 37 °C with shaking.
- b. The next morning, add 0.5 ml overnight culture to 4.5 ml LB with 100 µg/ml ampicillin. Grow 1 h at 37 °C.
- c. Remove 1 ml of un-induced sample. Centrifuge and remove supernatant. Freeze or store on ice until ready to run SDS-PAGE.
- d. Add IPTG to a final concentration of 1 mM and grow for an additional 2–3 h.
- e. Remove 1 ml induced sample. Centrifuge and remove supernatant. Freeze or store on ice until ready to run SDS-PAGE.
- f. Resuspend the cell pellets in 200 µl SDS-PAGE sample buffer; heat 3–5 min at 90 °C.
- g. Analyze 5–10 µl using SDS-PAGE followed by Coomassie blue staining (or 1–2 µl for Western blot).

⁷Samples grown in mini-cultures also can be used to evaluate the solubility of a particular fusion protein (see Note 6 for mini-culture expression protocol). At the end of the induction period, harvest the cells by centrifugation. Resuspend the pellet in 200 µl of cold PBS. Lyse the cells using sonication; keep the sample on ice. Save a small aliquot of the lysate for analysis by SDS-PAGE. Centrifuge the lysed sample for 15 min. Remove the supernatant to a separate tube. Resuspend the pellet in 200 µl PBS. Analyze some of the crude lysate, the supernatant, and the resuspended pellet by SDS-PAGE or Western blotting. If the sample is observed in both the lysate and supernatant fraction, the sample is adequately soluble. If the sample is present primarily in the lysate and resuspended pellet, the sample is most likely in occlusion bodies. In cases where the protein is located in inclusion bodies, explore different expression conditions to shift expression to a soluble form (see Note 8).

⁸If the fusion protein is expressed primarily in inclusion bodies (e.g. >80% insoluble), various strategies may be employed to improve solubility. Often, induction at lower temperature (15–25 °C) is effective at shifting expression from inclusion bodies to the soluble fraction. Cells induced at this lower temperature will grow more slowly and will require overnight induction periods. Use of alternative host strains or modifications to the plasmid construct may be necessary in some cases. If the fusion protein remains primarily in inclusion bodies even after attempts to obtain soluble protein, it may be worthwhile to scale up the production of *E. coli* and purify enough protein from the small amount of soluble protein that is available. In some cases, other fusion protein tags should be explored.

⁹Low protein expression levels could be the result of rare codon usage. In this case, switching to another strain of *E. coli* that contains extra transcripts for rare codon tRNA can improve significantly protein production. Different media formulations or addition of up to 2% glucose also may improve expression (13, 14). If it is suspected that the expressed protein is toxic to the cells (usually they will stop growing after induction with IPTG), try inducing at a later time point for a shorter amount of time. Decreasing the IPTG concentration is another option that should be explored.

the cells to grow for extended periods of time, as cell lysis can occur and release proteases into the cytoplasm that could degrade the fusion protein.

1. Using sterile technique, transfer some of a glycerol culture or isolated colonies grown on a streaked plate to 100 ml LB with 100 µg/ml ampicillin.
2. Incubate the inoculated culture overnight at 37 °C, with shaking at 250–300 rpm.
3. The next morning remove a 1 ml aliquot of cells and check the OD₆₀₀ using a spectrophotometer (see Note 10).
4. Using sterile technique, dilute the overnight starter culture 1:20 into 600 ml fresh LB supplemented with 100 µg/ml ampicillin (see Note 11).
5. Incubate the culture at 37 °C at 250–300 rpm until the OD₆₀₀ is 0.5–0.7 (see Note 10).
6. Remove a 1 ml aliquot of cells to be saved for gel analysis (see Note 6).
7. Add IPTG to a final concentration of 1.0 mM final concentration.
8. Incubate at 37 °C at 250–300 rpm for an additional 3 h (see Note 10), while monitoring the growth at OD₆₀₀. At saturation they will stop dividing. Remove 1 ml aliquot of cells for gel analysis (see Note 6).
9. Harvest cells by centrifugation at 4000 × g for 20 min at 4 °C.
10. Carefully decant the supernatant, leaving ~ 15–50 ml in the centrifuge bottle.
11. Resuspend the cells and transfer to a 50 ml centrifuge tube.
12. Centrifuge 20 min at 4000 × g, 4 °C.
13. Decant the supernatant.
14. Cell pellet may be frozen for up to several months at –80 °C.
15. Analyze un-induced and induced samples by SDS-PAGE to check protein expression levels (see Note 6).

3.3 Affinity purification of GST fusion protein

Soluble GST fusion proteins are purified easily using an immobilized glutathione Sepharose column. There are several options of immobilized glutathione chromatography media available to purify soluble GST fusion proteins from bacterial cell lysates (see Note 3). The protocol described below is an adaptation of the manufacturer's recommendation using glutathione Sepharose 4B poured into a column and using a peristaltic pump to control flow rates. Protease inhibitors and reducing agents should be added to the buffers, as required, to minimize proteolysis of the fusion protein. An exception is that serine protease inhibitors must be removed from the glutathione buffer prior to enzymatic removal of the GST moiety, as they will inhibit enzyme activity (see Note 12). Save a small aliquot from each step of the purification for analysis by SDS-PAGE to monitor the location of the fusion protein throughout the purification (see Notes 13 and 14). A given column or batch of resin should be used exclusively with a single fusion protein to minimize potential cross contamination.

¹⁰Monitor cell growth by reading the optical density at 600 nm (A₆₀₀). An overnight culture should reach an A₆₀₀ ~1–1.2. Cells should be induced at an early phase of the logarithmic growth curve for *E. coli*, A₆₀₀ ~ 0.5 to 0.7. At 37 °C, it will take approximately 2 h for cultures to reach an early log stage of growth. If cells are grown at a lower temperature, the incubation time must be increased, since the cells will grow more slowly at lower temperatures. Generally, the greatest yield of fusion protein will be obtained when the cells are induced at A₆₀₀ = ~ 0.5. After induction with IPTG, a general guideline for incubation time is as follows: 3 h at 37 °C, 5 h at 30 °C, and overnight for 25 °C or lower. Harvest the cells prior to saturation, A₆₀₀ ~ 1.0–1.2.

¹¹To ensure adequate aeration, flasks should only be filled to 20–30% of their capacity.

As an alternative to column purification, a protocol describing batch purifications is described in Note 15. Batch purifications are quick and simple, but frequently the yield and purity of the protein obtained will be somewhat lower than that obtained through chromatographic separations. In order to minimize proteolysis, all steps of the protein purification should be carried out at 4 °C, unless otherwise noted.

1. Resuspend and pour 20 ml glutathione Sepharose 4B resin into a 2.5 × 8 cm column (see Note 16 and ¹⁷).
2. Thoroughly wash the glutathione Sepharose with 5–10 bed volumes (see Note 16) PBS at 1.5 ml/min to remove the ethanol storage solution.
3. Resuspend the pelleted *E. coli* cells in 15 ml cold lysis buffer (cells might be freshly prepared or thawed frozen cell pellets) (see Note 18).
4. Lyse cells by sonication on ice (~10 times for 10 sec each with 1 min rest between bursts to minimize sample heating). Save 50 µl of lysate for gel analysis and transfer remainder to 60 ml centrifuge tube (see Note 19).
5. Centrifuge the lysate at 48,000 × g for 20 min at 4 °C.
6. Decant the supernatant into a clean 50 ml centrifuge tube.

¹²It is important that no serine protease inhibitors be present in the sample prior to cleavage with thrombin or factor Xa. The following protease inhibitors must be removed prior to thrombin or factor Xa cleavage: AEBSF, APMSF, antithrombin III, antipain, α₁-antitrypsin, aprotinin, chymostatin, hirudin, leupeptin, and PMSF. In addition, Pefabloc TH benzamidine is specific for thrombin, and Pefabloc FXa is specific to factor Xa. The following protease inhibitors should be avoided with PreScission Protease: 100 mM ZnCl₂, 100 µM chymostatin, and 4 mM Pefabloc.

¹³There are a number of alternative methods for detection of GST fusion proteins. For proteins that express well at a high level, the simplest method of monitoring the purification is via SDS-PAGE gels stained with Coomassie blue or silver stain. An alternative for proteins that express at low levels is to use Western blotting using an antibody directed at either GST or the target protein. An alternative for small-scale expressions is to monitor for the presence of GST with an ELISA or CDNB enzyme assay.

¹⁴Save a small amount of protein sample from each step of the purification, including lysis, supernatant, pellet, unbound fractions from sample loading, wash steps, and elution steps to be analyzed by SDS-PAGE or Western blotting. After all samples have been analyzed and pooled, discard unwanted fractions.

¹⁵GST proteins also can be purified using a batch purification method. A batch purification method is quick, easy, and requires little equipment; however, the resulting protein may contain more impurities and have a lower yield than a chromatography-based purification. Batch purifications are best utilized when screening purification conditions. The batch purification outlined below is generally performed at room temperature. To minimize the risk of proteolytic degradation, the procedure can be performed at 4 °C by increasing the incubation time 2- to 4-fold.

- a. Lyse cells and obtain soluble fusion protein (see Notes 8 and 21 if sample is in inclusion bodies).
- b. Add 2 ml 50% slurry glutathione Sepharose bulk matrix (1 ml bed volume) per 100 ml bacterial lysate supernatant. Incubate 30 min at room temperature with gentle agitation.
- c. Centrifuge 500 × g for 5 min. Remove the supernatant and save for analysis by SDS-PAGE to determine binding efficiency.
- d. Wash the resin with 10 bed volumes PBS. Gently resuspend, and then centrifuge 500 × g for 5 min. Remove the supernatant. Repeat wash and centrifugation steps for a total of 3 washes of 10 bed volumes each.
- e. Elute the fusion protein by resuspending the resin with 1.0 ml glutathione elution buffer per ml bed volume. Incubate 10 min at room temperature. Centrifuge 500 × g for 5 min. Transfer the fusion protein-containing supernatant to a separate tube. Repeat elution and collection steps a total of 3 times. The supernatants may be pooled or analyzed separately by SDS-PAGE to monitor protein content (see Note 12).

¹⁶A bed volume is equal to one half of the 50% slurry of glutathione Sepharose 4B used for purification. To prepare a 50% slurry of glutathione Sepharose (it is supplied as ~ 75% slurry): Determine the bed volume required for the purification scale. Resuspend the glutathione Sepharose 4B. For each ml of bed volume required, pipet 1.33 ml of the 75% slurry to an appropriate size centrifuge tube. Centrifuge at 500 × g for 5 min. Decant the super. Wash with 10 bed volumes (10 ml per 1.33 ml original slurry) of PBS by gently inverting the tube several times to mix, then centrifuge 500 × g for 5 min. Decant the super. Failure to remove the ethanol storage solution may interfere with subsequent binding steps. Add 1 ml PBS for each 1.33 ml of the original slurry. This results in a 50% slurry.

¹⁷Glutathione Sepharose has an advertised minimum binding capacity of 8 mg/ml. A 20 ml column should be adequate to purify protein from 3 × 600 ml *E. coli* cultures which contain ~ 20–40 mg fusion protein per culture. Both the amount of resin used and the column size can be scaled up or down depending on the amount of protein to be purified.

¹⁸Resuspend the pellet using 25–50 µl wash buffer per ml original culture.

7. Resuspend the pellet in 15 ml PBS buffer using a dounce homogenizer (see Note 18).
8. Run 5–10 μ l each of lysate, supernatant, and resuspended pellet on an SDS-PAGE gel to verify that the fusion protein is in the supernatant fraction (see Note 7). If the fusion protein is located in the pellet fraction, see Note 8 for tips to improve soluble protein expression or Note 20 for methods of extracting the protein from inclusion bodies.
9. Load the soluble fusion protein to the equilibrated glutathione Sepharose column using a flow rate of 0.1 ml/min (see Notes 21 and ²²). Collect fractions and run gels to verify that fusion protein is binding to the column and that the capacity has not

¹⁹Cell disruption is complete when the suspension partially clears and turns a slightly darker color. Avoid frothing during sonication, as this can lead to denaturation of the fusion protein. Avoid oversonication, as this can lead to co-purification of host *E. coli* proteins along with the GST fusion protein. High viscosity due to release of nucleic acids during sonication can be reduced by adding DNase or benzoase to the lysis buffer. Lysozyme up to a concentration of 0.2 mg/ml may be added as an aid to cell lysis. An alternative to sonication is the use of commercially available cell-extraction formulations. However, these products could contain proprietary components that interfere with downstream applications.

²⁰If attempts to shift expression from inclusion bodies to the soluble fraction fail, insoluble GST fusion proteins sometimes can be purified in the presence of denaturants such as urea, followed by refolding. Solubilization using detergents such as sarkosyl (N-laurylsarosine) also has been successfully employed (15, 16). Although the following protocol has been used successfully to denature and re-nature GST fusion proteins, each fusion protein construct is unique and the exact denaturation and re-naturation conditions must be determined empirically. Common denaturants that have been successfully employed include guanidine HCl, urea, Tween 20, CTAB, and SDS (17, 18). The denaturants then must be completely removed to allow proper refolding of the protein. Conditions that should be optimized to facilitate refolding include: type of denaturant, pH, presence of reducing agent, speed of denaturant removal, and protein concentration. Once the protein has been re-natured, verify that the protein has regained its native conformation and function and remove any improperly folded protein.

- a. Pre-equilibrate the glutathione Sepharose column; sonicate the pelleted *E. coli* cells, and separate the lysate supernatant and pellet fractions by centrifugation (see **3.3 Affinity purification of GST fusion protein steps 1–8**). Resuspend the lysate pellet that includes the inclusion bodies, in 15 ml wash buffer. Centrifuge 20 min at $48,000 \times g$ (4 °C). Decant the supernatant and resuspend the washed pellet in 12 ml U-buffer (resuspend in 20 μ l U-buffer per ml of original culture). Incubate 2 h on ice.
- b. Centrifuge 20 min at $48,000 \times g$ (4 °C). Transfer the supernatant that contains the denatured fusion protein to a clean centrifuge tube.
- c. Add Triton X-100 to the supernatant to a final concentration of 1%.
- d. Dialyze the sample into PBS/glycerol for 2–3 h. The volume of the dialysis buffer should be a minimum of 20 times the sample volume.
- e. Dialyze the sample overnight versus PBS with protease inhibitors using a buffer volume that is >100 times the sample volume.
- f. Remove the sample from dialysis and centrifuge 20 min at $4,000 \times g$.
- g. Extracted and re-natured protein from the inclusion bodies now can be applied to glutathione Sepharose columns and purified.

Solutions:

Wash buffer: 50 mM Tris, 5 mM EDTA, 1 μ g/ml leupeptin, 1 μ g/ml pepstatin, 0.15 mM PMSF, pH 8.0

U buffer: 5 M urea, 50 mM Tris, 5 mM EDTA, 5 mM 2-ME, 1 μ g/ml leupeptin, 1 μ g/ml pepstatin, 0.15 mM PMSF, 1 mM DFP pH 8.0

PBS/glycerol: 1X PBS, 20% glycerol, 1% Triton X-100, 5 mM 2-ME, 5 mM EDTA, 5 mM 2-ME, 1 μ g/ml leupeptin, 1 μ g/ml pepstatin, 0.15 mM PMSF, 1 mM DFP pH 7.4

²¹Use of a peristaltic pump is advised to evenly control flow rates. If compression of the resin occurs, the pressure is too high and the flow rate should be reduced.

²²The binding kinetics between glutathione and GST are relatively slow. Therefore it is important to use low flow rates to achieve maximum binding capacity, e.g. 0.1 ml/min for a 2.5 cm I.D. column. Using flow rates that are too fast may decrease the amount of bound fusion protein due to the slow kinetics of association. Washing and elution steps may be performed at faster flow rates to save time (1.5 ml/min and 0.3 ml/min, respectively). For large volume samples, binding is most conveniently performed overnight, with adjustments to flow rates so that the column does not run dry.

been exceeded (see Notes 23). If the fusion protein binds poorly to the resin, see Notes 24 for several possible remedies.

10. Wash the column with 5–10 bed volumes of PBS/EDTA/PMSF using a flow rate of 1.5 ml/min.
11. Wash the column with 10 bed volumes of PBS/EDTA using a flow rate of 1.5 ml/min (see Note 12).
12. Elute the fusion protein with glutathione buffer using a flow rate of 0.3 ml/min. Fractions may be monitored using A_{280} and analysis by SDS-PAGE (see Note 25). Pool fractions that contain the GST fusion protein. Protein may be stored at 4 °C and should be ~90% pure at this stage. If problems are encountered eluting the fusion protein, see Note 26. If high levels of contamination are present, see Note 27 for troubleshooting tips.

3.4 Enzymatic cleavage to remove GST affinity tag

Depending on the vector chosen, the GST affinity tag can be removed with thrombin, factor Xa, or PreScission protease, either in solution or while still bound to the column matrix. Cleavage in solution offers the advantage of more control over optimization of the cleavage conditions such as temperature, enzyme-to-substrate ratio, length of incubation, and buffer conditions. An advantage of on-column cleavage is the high level of purity obtained, but this comes at the expense of generally low yield due to less efficient protease cleavage and decreased control of the digestion conditions. Digestion can be performed in the glutathione buffer used to elute the protein from the affinity matrix provided there are no serine protease inhibitors in this buffer. After incubation, the enzyme can be inhibited using a variety of protease inhibitors or removed using a HiTrap Benzamidine column. Separation of the target protein and the GST moiety can be achieved by re-chromatography on the glutathione Sepharose column (after dialysis to PBS buffer) to remove the GST and any un-cleaved fusion protein. See Note 28 for digestion of GST fusion proteins while bound to the column matrix (this is recommended for use with PreScission protease).

1. Add the appropriate amount of thrombin or factor Xa to the affinity-purified fusion protein and incubate at 37 °C (thrombin) or 25 °C (factor Xa) for the desired length of time (see Note 30).

²³Analysis of the unbound fractions will indicate whether the fusion protein is binding or not. If the protein is not detected in early fractions but appears in later fractions, it indicates that the column capacity has been exceeded. Reduction in protein load or increase in column size will alleviate this condition.

²⁴If the fusion protein binds poorly to the glutathione Sepharose, there are several alternatives to try increasing binding efficiency. Try a milder lysis method. If the method used is too harsh, the protein may become denatured and, therefore, unable to bind to the column. Also, if re-naturing the protein from inclusion bodies, be sure that all denaturants have been removed from the buffer, either through exhaustive dialysis or application of the sample to a desalting column, before applying to the glutathione column. Remove the ethanol storage solution from the glutathione Sepharose; reduce the column with fresh glutathione buffer followed by a wash with PBS immediately prior to sample loading. Increase the amount of resin and/or decrease the flow rate used for loading the sample. Try adding 1–20 mM DTT to the sample. If a problem still persists, clean the column according to the manufacturer's recommendation. If binding still is not restored, try using fresh resin.

²⁵The yield of fusion protein also can be estimated by measuring absorbance at 280 nm (A_{280}). The extinction coefficient for the GST moiety alone is ~1.5, i.e., $1.00 A_{280} \sim 0.6$ mg/ml protein, although the extinction coefficient for the fusion protein will depend partially on the absorbance characteristics of the target protein.

²⁶If there is a problem eluting the protein from the glutathione Sepharose column, try decreasing the flow rate and increasing the volume of elution buffer that is used. Be sure to use fresh reduced glutathione buffer (make it the day that it will be used). Increase the concentration of glutathione up to 40 mM and/or raise the pH of the buffer to pH 8.0–9.0. Add 1–20 mM DTT (or other reducing agent) to the buffer. Addition of a nonionic detergent also may improve solubilization and minimize any aggregation that may be occurring.

²⁷Contamination of the purified fusion protein with *E. coli* host cell proteins is an indication that sonication has been too severe. If degraded fragments of the fusion protein are present, try adding additional protease inhibitors to the lysis buffer. Keep all samples, buffers, and collection tubes cold to minimize proteolysis. If degradation is occurring during protein expression, try inducing the sample late (~0.8 OD₆₀₀) and decrease the length of the induction period. Switching to an alternative host strain also may help.

2. Inactivate the enzyme by adding 0.3 mM PMSF (final concentration) to the sample. To ensure complete inhibition, incubate the sample 15 min at 37 °C for thrombin or 30 min at 25°C for factor Xa (see Note 32).
3. Dialyze the sample against PBS/EDTA/PMSF twice using 2L per dialysis for a minimum for 4 h for each dialysis (see Note 33).
4. Centrifuge the dialyzed sample 20 min at 4,000 × g to remove any precipitated material that may have formed during the digestion or dialysis steps. At this point, the sample can be reapplied to the glutathione Sepharose column to remove the GST moiety and any undigested fusion protein.

3.5 Removal of GST affinity tag after protease cleavage in solution

If the protein was digested in solution, the target protein can be purified further by re-chromatography on the glutathione Sepharose column to remove any un-cleaved fusion protein and the GST moiety. After a dialysis step to remove the glutathione present from the initial isolation, the sample is reapplied to the column and the sample of interest and thrombin is located in the column flow through. Protein purification should be performed at 4 °C to minimize degradation of the target protein.

1. If reusing a glutathione Sepharose column, wash with > 3 bed volumes of glutathione buffer at 1.5 ml/min (see Note 34 and ³⁵).
2. Wash the glutathione Sepharose column with 10 bed volumes PBS at 1.5 ml/min.
3. Load the dialyzed, cleaved, fusion protein onto the column at 0.1 ml/min (see Note 22). Collect fractions for analysis by SDS-PAGE (see Note 36).
4. Wash the column with 2–3 bed volumes PBS/EDTA/PMSF at 1.5 ml/min.

²⁸An alternative to in-solution digestion is protease cleavage of fusion protein while bound to the glutathione Sepharose matrix. The fusion protein is extracted and loaded onto the glutathione Sepharose followed by washing with PBS (see **Affinity purification of GST fusion protein, steps 1–11**). Rather than elute the protein with glutathione buffer, a PBS solution containing the enzyme is loaded onto the column and incubated for several hours at room temperature (4 °C for PreScission protease). The cleaved protein then is washed out of the column with several column volumes of PBS. Residual fusion protein and the GST moiety can be removed from the column by washing the column with reduced glutathione buffer. The amount of enzyme and incubation time must be empirically determined for each fusion protein. Analyze all samples by SDS-PAGE to determine cleavage efficiency and protein purity.

³⁰If using thrombin protease, digestion can be carried out in the glutathione buffer used to elute the protein from the column. If using factor Xa, it is recommended to dialyze the protein into either a Tris buffer or PBS buffer prior to digestion; the glutathione present in the elution buffer can disrupt the disulfide bridges present in factor Xa leading to inefficient digestion of fusion protein. An empirical determination of digestion conditions for each fusion protein must be determined in a pilot digestion experiment. A convenient method is to digest 100 µg of fusion protein over a range of enzyme-to-substrate ratios and vary the incubation time. Typical incubation times range from 2 to 8 hr. Recommended enzyme-to-substrate ratios for thrombin are 1:100, 1:350, 1:1000, and 1:3000 (units of enzyme per µg of fusion protein), and for factor Xa are 1:10, 1:25, 1:50, 1:100, 1:300 (µg enzyme per µg fusion protein). At desired time points, remove 2 µg of protein and stop digestion by adding the sample aliquot to boiling SDS sample buffer. Analyze the samples by SDS-PAGE to determine the optimal cleavage conditions. In addition to enzyme-to-substrate ratio and time, consider altering buffer conditions, such as increasing or decreasing NaCl concentration or adding Ca²⁺ to the buffer (see Note 31 for troubleshooting tips).

³²If inhibition of the enzyme after digestion is complete using serine protease inhibitors is undesirable, the sample may be applied to a HiTrap benzamidine column to remove the enzyme from the sample.

³³If the sample is to be re-chromatographed on the glutathione Sepharose column to remove GST and any undigested fusion protein, it is critical that the reduced glutathione be removed completely from the sample. Glutathione equilibrates very slowly across dialysis membranes; therefore if dialysis tubing with a MWCO <12,000 is used, 3 or more changes of buffer may be required for complete removal. Increased dialysis time (overnight) and a larger volume of buffer also may be required for large sample volumes.

³⁴The same glutathione Sepharose column may be used for both initial isolation of the fusion protein and for repurification after enzymatic cleavage. It is recommended to dedicate a single column to an individual protein construct to avoid potential cross contamination of different recombinant proteins. Columns may be reused multiple times. If binding efficiency decreases over time, clean the column according to the manufacturer's recommendations. If binding activity is not restored, a new column should be used.

³⁵Maximal binding occurs if the column is fully reduced; however, if the glutathione Sepharose column has been washed less than 48 h prior to this step, this step may be skipped.

³⁶The target protein will be in the unbound fraction and the GST and any un-cleaved fusion protein will bind to the column.

5. Elute the bound GST and un-cleaved fusion protein with glutathione buffer at 0.3 ml/min for 5 bed volumes. Collect fractions for analysis by SDS-PAGE.
6. Analyze all unbound and bound fractions by SDS-PAGE. Pool fractions that contain the cleaved target protein and store at 4 °C.

3.6 Further purification of cleaved recombinant protein

Unless a benzamidine column has been used to remove the protease, the cleaved sample contains the inactivated protease. The sample also may contain small amounts of residual GST or the GST fusion protein that did not rebind to the column, as well as aggregates or other minor contaminants such as proteolytic fragments or host cell proteins. While the protein is usually greater than 90% pure at this stage, many applications require an even higher level of purity. Therefore, HPLC gel filtration is recommended as a final polishing step that will isolate properly folded protein from aggregates or other contaminating species. As an alternative to HPLC, ion exchange chromatography also could be performed.

1. Concentrate the cleaved unbound protein to a final volume that is 0.5 to 1% of the column volume to be used for the separation using an ultrafiltration device, such as Amicon Ultra (Millipore), according to the manufacturer's instructions (see Note 37).
2. Remove particulates from the concentrated sample using a 0.22 µm filter unit, or centrifuge the sample at 4,000 × g for 20 min.
3. Inject the concentrated, filtered sample onto an appropriate gel filtration column equilibrated in PBS or other compatible sample buffer.
4. Monitor A₂₈₀ with an online HPLC detector and collect fractions. Analyze the fractions by SDS-PAGE and pool based on sample purity or concentration.
5. Store pooled sample at 0–4 °C or freeze appropriately until further use.

Acknowledgments

This work was supported in part by the National Institutes of Health Grant HL038794 and institutional grants to The Wistar Institute, including an NCI Cancer Core Grant (CA10815) and grants from the Pennsylvania Department of Health.

References

1. Smith DB, Johnson KS. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene*. 1988; 67:31–40. [PubMed: 3047011]
2. Zhan Y, Song X, Zhou GW. Structural analysis of regulatory protein domains using GST-fusion proteins. *Gene*. 2001; 281:1–9. [PubMed: 11750122]
3. Smyth DR, Mrozkiewicz MK, McGrath WJ, Listwan P, Kobe B. Crystal structures of fusion proteins with large-affinity tags. *Protein Sci*. 2003; 12:1313–1322. [PubMed: 12824478]
4. Singh CR, Asano K. Localization and characterization of protein-protein interaction sites. *Methods Enzymol*. 2007; 429:139–161. [PubMed: 17913622]
5. Yip YL, Smith G, Ward RL. Comparison of phage pIII, pVIII and GST as carrier proteins for peptide immunisation in Balb/c mice. *Immunol Lett*. 2001; 79:197–202. [PubMed: 11600198]
6. Saaem I, Papasotiropoulos V, Wang T, Soteropoulos P, Libera M. Hydrogel-based protein nanoarrays. *J Nanosci Nanotechnol*. 2007; 7:2623–2632. [PubMed: 17685276]

---

³⁷Small sample volumes are recommended for good resolution of target protein versus other contaminants. If it is not feasible to concentrate the sample into a small volume, multiple injections of the sample may be performed.

7. Jung JW, Jung SH, Kim HS, Yuk JS, Park JB, Kim YM, Han JA, Kim PH, Ha KS. High-throughput analysis of GST-fusion protein expression and activity-dependent protein interactions on GST-fusion protein arrays with a spectral surface plasmon resonance biosensor. *Proteomics*. 2006; 6:1110–1120. [PubMed: 16402361]
8. Vikis HG, Guan KL. Glutathione-S-transferase-fusion based assays for studying protein-protein interactions. *Methods Mol Biol*. 2004; 261:175–186. [PubMed: 15064458]
9. Davies AH, Jowett JB, Jones IM. Recombinant baculovirus vectors expressing glutathione-S-transferase fusion proteins. *Biotechnology (N Y)*. 1993; 11:933–936. [PubMed: 7763917]
10. Mitchell DA, Marshall TK, Deschenes RJ. Vectors for the inducible overexpression of glutathione S-transferase fusion proteins in yeast. *Yeast*. 1993; 9:715–722. [PubMed: 8368005]
11. Sambrook, J.; aRD. *Molecular Cloning: A Laboratory Manual*. 3. Cold Spring Laboratory Press; Cold Spring Harbor, New York: 2001.
12. Ausubel, FM.; Brent, R.; Kingston, RE.; Moore, DD.; Seidman, JG.; Smith, JA.; Struhl, K., editors. *Current Protocols in Molecular Biology*. John Wiley & Sons; Hoboken, NJ: 2008.
13. Grossman TH, Kawasaki ES, Punreddy SR, Osburne MS. Spontaneous cAMP-dependent derepression of gene expression in stationary phase plays a role in recombinant expression instability. *Gene*. 1998; 209:95–103. [PubMed: 9524234]
14. Pan SH, Malcolm BA. Reduced background expression and improved plasmid stability with pET vectors in BL21 (DE3). *Biotechniques*. 2000; 29:1234–1238. [PubMed: 11126126]
15. Frangioni JV, Neel BG. Solubilization and purification of enzymatically active glutathione S-transferase (pGEX) fusion proteins. *Anal Biochem*. 1993; 210:179–187. [PubMed: 8489015]
16. Grieco F, Hay JM, Hull R. An improved procedure for the purification of protein fused with glutathione S-transferase. *Biotechniques*. 1992; 13:856–858. [PubMed: 1476735]
17. Singh SM, Panda AK. Solubilization and refolding of bacterial inclusion body proteins. *J Biosci Bioeng*. 2005; 99:303–310. [PubMed: 16233795]
18. Singh SM, Eshwari AN, Garg LC, Panda AK. Isolation, solubilization, refolding, and chromatographic purification of human growth hormone from inclusion bodies of *Escherichia coli* cells: a case study. *Methods Mol Biol*. 2005; 308:163–176. [PubMed: 16082034]
19. Guan KL, Dixon JE. Eukaryotic proteins expressed in *Escherichia coli*: an improved thrombin cleavage and purification procedure of fusion proteins with glutathione S-transferase. *Anal Biochem*. 1991; 192:262–267. [PubMed: 1852137]
20. Hakes DJ, Dixon JE. New vectors for high level expression of recombinant proteins in bacteria. *Anal Biochem*. 1992; 202:293–298. [PubMed: 1519755]

Figure 1.

Flow diagram illustrating the key decision-making steps for designing and executing a successful GST fusion protein purification. Selection of the vector depends on whether the GST moiety ultimately will be cleaved away and the enzyme that will be used. The insert should be compatible with the vector chosen. A variety of host cells are available; the use of BL21 strains and their derivatives are recommended. Expression conditions should be optimized to yield the highest level of soluble fusion protein. Exact compositions of lysis and purification buffers depend largely on target protein characteristics and should be modified accordingly. Conditions for removal of the GST moiety also will need to be optimized. Overall, each step involves multiple options that should be tailored to the characteristics of the target protein and intended purpose of the purified protein.

Figure 2.

GST fusion proteins may be constructed using one of 10 different pGEX vectors. Shared characteristics of the vectors include a gene for ampicillin resistance; the *lacI^q* gene product is a repressor protein which binds to the operator region of the *P_{tac}* promoter preventing expression of protein product until induction with IPTG. The gene for the protein of interest can be inserted in the polylinker site immediately after the GST gene. The protease cleavage sites are located between the GST moiety and the protein of interest. Although not shown for all vectors, stop codons in all three reading frames follow the polylinker site. A consideration to be made when selecting the vector is to minimize the number of amino acids remaining after proteolytic cleavage.

(© 2009 General Electric Company – Reproduced by permission of the owner)