Chapter 11

Hexahistidine-Tagged Maltose-Binding Protein as a Fusion Partner for the Production of Soluble Recombinant Proteins in *Escherichia coli*

Brian P. Austin, Sreedevi Nallamsetty, and David S. Waugh

Summary

Insolubility of recombinant proteins in *Escherichia coli* is a major impediment to their production for structural and functional studies. One way to circumvent this problem is to fuse an aggregation-prone protein to a highly soluble partner. *E. coli* maltose-binding protein (MBP) has emerged as one of the most effective solubilizing agents. In this chapter, we describe how to construct combinatorially-tagged His_6MBP fusion proteins by recombinational cloning and how to evaluate their yield and solubility. We also describe a procedure to determine how efficiently a His_6MBP fusion protein is cleaved by tobacco etch virus (TEV) protease in *E. coli* and a method to assess the solubility of the target protein after it has been separated from His_6MBP.

Keywords: Maltose-binding protein; MBP; Inclusion body; Fusion protein; Solubility enhancer; TEV protease; Tobacco etch virus protease; Hexahistidine tag; His-tag; His_6MBP; Gateway cloning; Recombinational cloning

11.1. Introduction

In many cases, the poor solubility of recombinant proteins in heterologous hosts presents a major obstacle to their large-scale production for functional and structural studies *(1)*. Accordingly, researchers have long sought a generally effective means of improving protein solubility. One popular approach has been to fuse an aggregation-prone protein to a highly soluble partner. In many cases, this enables the aggregation-prone protein to be recovered in a soluble and properly folded form. It should be noted, however, that not all highly soluble proteins are equally

Sharon A. Doyle (ed.), *Methods in Molecular Biology: High Throughput Protein Expression and Purification, vol. 498*

Book doi: 10.1007/978-1-59745-196-3

effective solubilizing agents. Several comparative studies of putative solubility-enhancing fusion partners have yielded conflicting results *(e.g., 2–7)*. Although the reasons for these discrepancies remain unclear, there is ample evidence to indicate that *E. coli* maltose-binding protein (MBP) is among the more effective solubility enhancers. Additionally, to our knowledge, MBP is the only solubility-enhancing protein that is also a natural affinity tag. In principle, its affinity for amylose resin can be exploited to facilitate the purification of an MBP fusion protein *(8)*. In practice, however, we and others have observed that amylose affinity chromatography has several noteworthy disadvantages: the resin is fragile and comparatively expensive; some MBP fusion proteins do not bind efficiently to amylose resin; and, even when they do, this technique rarely yields samples of sufficient purity for structural studies *(9,10)*.

To circumvent the disadvantages of amylose affinity chromatography, a hexahistidine tag (His_6) can be added to the N-terminus of MBP and used for immobilized metal affinity chromatography instead *(11,12)*. We have shown that the addition of a His_6 tag to the N-terminus of MBP does not interfere with its ability to promote the solubility of its fusion partners *(11)*. Moreover, the dual His_6MBP tag can be exploited to purify proteins to homogeneity via an entirely generic and potentially automatable process *(11,12)*.

In this chapter, we focus on the utility of MBP (specifically His_6MBP) as a solubility-enhancing fusion partner. We describe in detail a method for constructing His_6MBP fusion protein expression vectors by Gateway recombinational cloning and provide illustrious examples of the most common outcomes. MBP fusion vectors (but not His_6MBP fusion vectors) designed for cloning by conventional methods (i.e., using restriction endonucleases and DNA ligase) can be obtained from New England Biolabs (Beverley, MA, USA) but will not be discussed here.

11.2. Materials

11.2.1. Recombinational Vector Construction

1. The Gateway destination vector pDEST-HisMBP (*see* **Fig. 11.1**), which can be obtained from AddGene (http://www.addgene.org) or the authors.
2. Reagents and thermostable DNA polymerase for PCR amplification (*see* **Note 1**).
3. Synthetic oligodeoxyribonucleotide primers for PCR amplification (*see* **Fig. 11.2**).
4. TE buffer: 10 mM Tris-HCl. pH 8.0, 1 mM EDTA.

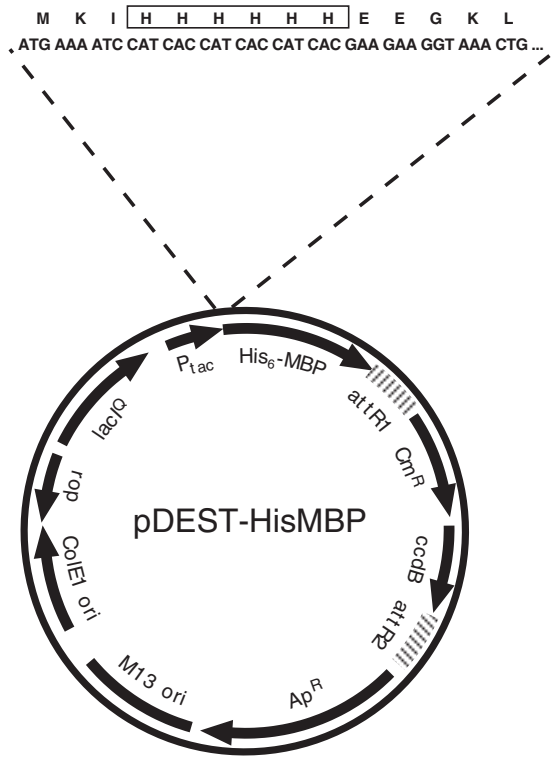

Fig. 11.1. Schematic representation of the Gateway destination vector pDEST-HisMBP. This vector can be recombined with an entry clone that contains an ORF of interest, via the LR reaction, to generate a His_6MBP fusion protein expression vector. The nucleotide and amino acid (single letter code) sequences at the beginning of the open reading frame are shown. Abbreviations: P_{tac} *tac* promoter; *attR1* and *attR2* recombination sites for Gateway cloning, Cm^R chloramphenicol acetyl transferase (chloramphenicol-resistance) gene; *ccdB* gene encoding DNA gyrase poison CcdB, Ap^R β-lactamase (ampicillin-resistance) gene, *M13 ori* origin of replication from bacteriophage M13, *ColE1 ori* origin of replication from ColE1 plasmid, *rop* repressor of primer gene, $lacI^Q$ gene encoding lactose repressor.

5. E-gels and an E-gel base (Qiagen, Valencia, CA, USA) for submarine gel electrophoresis of DNA (*see* **Note 2**).
6. QIAquick gel extraction kit (Qiagen, Valencia, CA, USA) for the extraction of DNA from agarose gels.
7. Chemically competent DB3.1 or "CcdB survival" cells (Invitrogen, Carlsbad, CA, USA) for propagating pDEST-HisMBP and pDONR221.
8. Competent *gyrA*$^+$ cells (e.g., DH5α, MC1061, HB101) (*see* **Note 3**).
9. The Gateway donor vector pDONR221 (Invitrogen, Carlsbad, CA, USA).
10. Gateway BP Clonase II (Invitrogen, Carlsbad, CA, USA).
11. Gateway LR clonase II (Invitrogen, Carlsbad, CA, USA).
12. LB medium and LB agar plates containing ampicillin (100 μg/ml). LB medium: Add 10 g bactotryptone, 5 g bacto yeast extract,

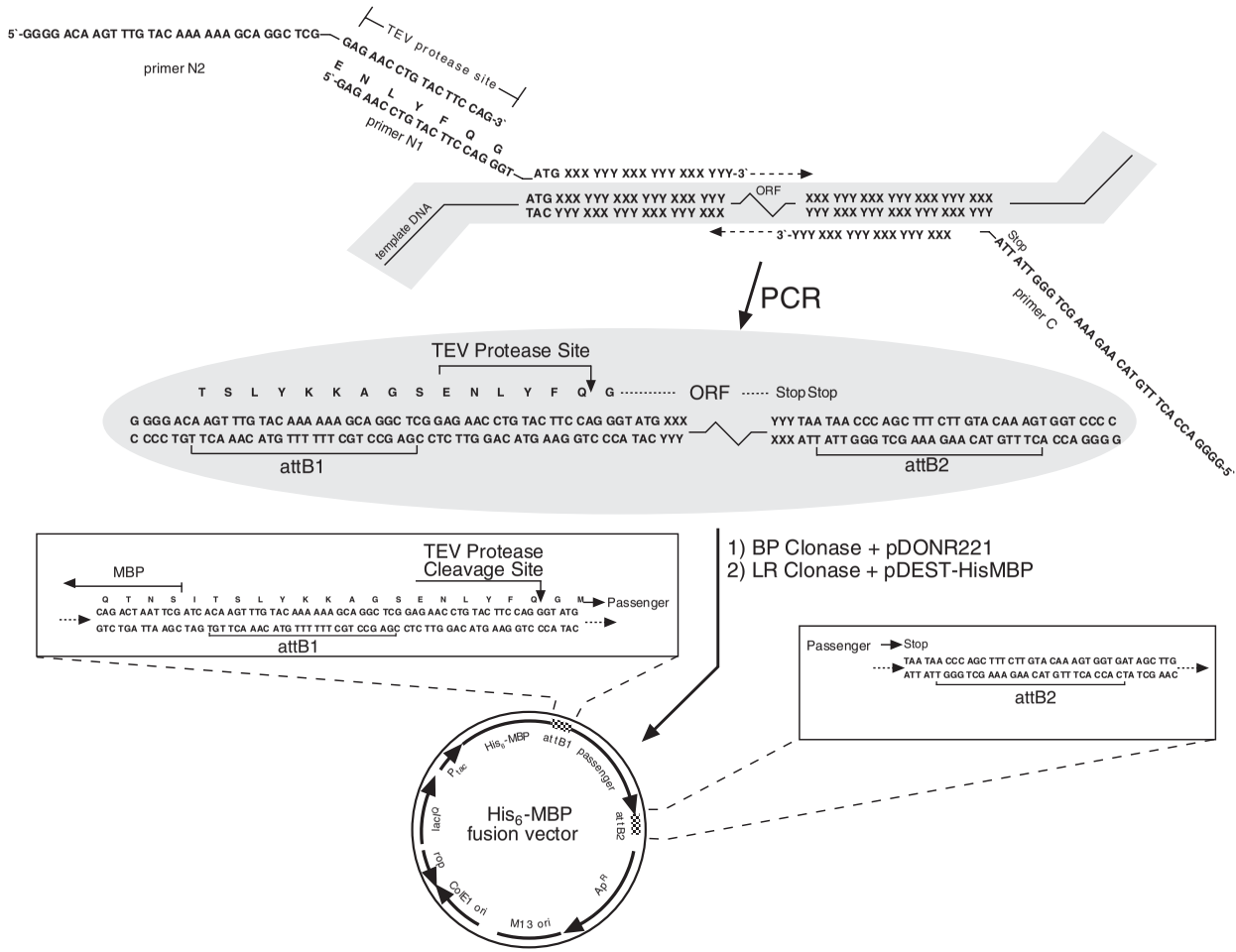

Fig. 11.2. Construction of a His_6MBP fusion vector using PCR and Gateway cloning technology. The ORF of interest is amplified from the template DNA by PCR, using primers N1, N2 and C. Primers N1 and C are designed to base-pair to the 5′ and 3′ ends of the coding region respectively and contain unpaired 5′ extensions as shown. Primer N2 base-pairs with the sequence that is complementary to the unpaired extension of primer N1. The final PCR product is recombined with the pDONR221 vector to generate an entry clone, via the BP reaction. This entry clone is subsequently recombined with pDEST-HisMBP using LR Clonase to yield the final His_6MBP fusion vector.

and 5 g NaCl to 1 L of H_2O and sterilize by autoclaving. For LB agar, also add 12 g of bactoagar before autoclaving. To prepare plates, allow medium to cool until flask or bottle can be held in hands without burning, then add 1 ml ampicillin stock solution (100 mg/ml in H_2O, filter sterilized), mix by gentle swirling, and pour or pipet ca. 30 ml into each sterile petri dish (100 mm dia.).

13. QIAprep Spin Miniprep Kit (Qiagen) for small-scale plasmid DNA isolation (*see* **Note 4**).
14. A microcentrifuge capable of operating at 14,000 rpm.
15. An incubator set at 37°C.

11.2.2. Protein Expression

1. Competent BL21-Pro cells (B&D Clontech, Palo Alto, CA, USA) containing the TEV protease expression vector pRK603 *(13)*. pRK603 can be obtained from AddGene (http://www.addgene.org) (*see* **Notes 5** and **6**).
2. A derivative of pDEST-HisMBP that produces a His_6-MBP fusion protein with a TEV protease recognition site in the linker between MBP and the passenger protein (*see* **Section 11.3.1**).
3. LB agar plates and broth containing both ampicillin (100 µg/ml) and kanamycin (35 µg/ml). *See* **Section 11.2.1 Item 10** for LB broth, LB agar, and ampicillin stock solution recipes. Prepare stock solution of 35 mg/ml kanamycin in H_2O and filter sterilize. Store at 4°C for up to 1 month. Dilute antibiotics 1,000-fold into LB medium or molten LB agar at no more than 50°C.
4. Isopropyl-thio-β-D-galactopyranoside (IPTG), analytical grade. Prepare a stock solution of 200 mM in H_2O and filter sterilize. Store at 4°C.
5. Anhydrotetracycline (ACROS Organics/Fisher Scientific, Springfield, NJ, USA). Prepare a 1,000× stock solution by dissolving in 50% ethanol at 100 µg/ml. Store in a foil-covered tube at −20°C.
6. Shaker/incubator.
7. Sterile baffle-bottom flasks (Bellco Glass, Inc., Vineland, NJ, USA)
8. Cell lysis buffer: 20 mM Tris-HCl, pH 8, 1 mM EDTA.
9. Sonicator (with microtip).
10. 2× SDS-PAGE sample buffer (Invitrogen) and 2-mercaptoethanol (Sigma-Aldrich, St. Louis, MO, USA).
11. SDS-PAGE gel, electrophoresis apparatus, and running buffer (*see* **Note 7**).
12. Gel stain (e.g. Gelcode® Blue from Pierce, Rockford, IL, USA, or PhastGel™ Blue R from Amersham Biosciences, Piscataway, NJ, USA).

13. Spectrophotometer.
14. 1.5 ml microcentrifuge tubes.

11.3. Methods

11.3.1. Construction of His_6MBP Fusion Vectors by Recombinational Cloning

The Gateway recombinational cloning system is based on the site specific recombination reactions that mediate the integration and excision of bacteriophage lambda into and from the *E. coli* chromosome, respectively. For detailed information about this system, consult the technical literature supplied by Invitrogen, Inc. and available online at: http://www.invitrogen.com/Content/Online%20Seminars/gateway/gatewayhome.html.

11.3.1.1. pDEST-HisMBP

To utilize the Gateway system for the production of His_6MBP fusion proteins, one must first construct or obtain a suitable "destination vector". Currently there is no commercial source for such a vector, but one (pDEST-HisMBP) can be obtained from the non-profit distributor of biological reagents AddGene, Inc. (http://www.addgene.org). A schematic diagram of pDEST-HisMBP is shown in **Fig. 11.1**. This plasmid was constructed by inserting an in-frame hexahistidine coding sequence between codons 3 and 4 of MBP in pKM596 *(12)*.

The Gateway cloning cassette in pDEST-HisMBP carries a gene encoding the DNA gyrase poison CcdB, which provides a negative selection against non-recombined destination and donor vectors so that only the desired recombinant is obtained when the end products of the recombinational cloning reaction are transformed into *E. coli* and grown in the presence of ampicillin or kanamycin, respectively. pDEST-HisMBP and other vectors that carry the *ccdB* gene must be propagated in a host strain with a *gyrA* mutation (e.g. *E. coli* DB3.1) or "CcdB survival" cells that are immune to the action of CcdB.

11.3.1.2. Gateway Cloning Protocol

To construct a His_6MBP fusion protein expression vector, one begins by amplifying the open reading frame of interest by PCR using a forward primer containing the TEV protease cleavage site as a 5′ unpaired extension (N1) and a reverse primer containing an attB2 recombination site as a 5′ unpaired extension (C) (*see* **Fig. 11.2**). The remainder of each primer (3′ ends) should consist of approximately 20–25 nucleotides that are complimentary to the ends of the open reading frame. To avoid the need for an excessively long forward primer, the first PCR amplicon is used as the template for a second PCR with a forward primer consisting of the TEV protease recognition site and an attB1 recombination site (N2) and the same reverse primer (C). The PCR

reaction may contain all three primers (*see* **Note 8**). To favor the accumulation of the desired product, the *att*B-containing primers are used at typical concentrations for PCR but the concentration of the gene-specific N-terminal primer (N1) is 10–20 folds lower. The final PCR amplicon is inserted first into the donor vector pDONR221 by recombinational cloning with BP clonase and then into the destination vector pDEST-HisMBP in a second recombinational cloning reaction with LR clonase.

1. The PCR reaction mix is prepared as follows (*see* **Note 9**): 1 µl template DNA (~10 ng µl), 5 µl thermostable DNA polymerase 10× reaction buffer, 1.5 µl dNTP solution (10 mM each), 1.0 µl primer N1 (~30 ng), 3 µl primer N2 (~300 ng), 3.0 µl primer C (~300 ng), 1 µl thermostable DNA polymerase, 64.5 µl H_2O (to 100 µl total volume).
2. The reaction is placed in the PCR thermal cycler with the following program: initial melt for 5 min at 95°C; 35 cycles of 95°C for 30 s, 55°C for 30 s, and 72°C for 60 s (*see* **Note 10**); hold at 4°C.
3. Purification of the PCR amplicon by agarose gel electrophoresis (*see* **Note 2**) is recommended to remove *att*B primer-dimers.
4. To create the His_6MBP fusion vector, the PCR product is recombined first into pDONR221 to yield an entry clone intermediate (BP reaction), and then into pDEST-HisMBP (LR reaction; *see* **Note 11**). This is described in steps 5–13 below.
5. Add to a microcentrifuge tube on ice: 50–100 ng of the PCR product in TE or H_2O, 200 ng of pDONR221 DNA, and enough TE to bring the total volume to 8 µl. Mix well.
6. Thaw BP Clonase II enzyme mix on ice (2 min) and then vortex briefly (2 s) twice (*see* **Note 12**).
7. Add 2 µl of BP Clonase II enzyme mix to the components in (a) and vortex briefly twice.
8. Incubate the reaction at 25°C for at least 4 h (*see* **Note 13**).
9. Add to the reaction: 2 µl of the destination vector (pDEST-HisMBP) at a concentration of 150 ng/µl, and 3 µl of LR Clonase II enzyme mix (*see* **Note 12**). The final reaction volume is 15 µl. Mix by vortexing briefly.
10. Incubate the reaction at room temperature for 3–4 h.
11. Add 2.5 µl of the proteinase K solution and incubate for 10 min at 37°C.
12. Transform 2 µl of the reaction into 50 µl of chemically competent DH5α cells (*see* **Note 3**).
13. Pellet the cells by centrifugation, gently resuspend pellet in 100–200 µl of LB broth and spread on an LB agar plate containing ampicillin (100 µg/ml), the selective marker for

pDEST-HisMBP (*see* **Fig. 11.1**). Incubate the plate at 37°C overnight (*see* **Note 14**).

14. Plasmid DNA is isolated from saturated cultures started from individual ampicillin-resistant colonies and screened by PCR, using the gene-specific primers N1 and C, to confirm that the clones contain the expected gene. Alternatively, plasmids can be purified and screened by conventional restriction digests using appropriate enzymes. We routinely sequence clones that screen positive by either PCR or restriction digest to ensure that there are no PCR-induced mutations.

11.3.2. Protein Expression

To assess the yield and solubility of the fusion protein, the amount of total fusion protein produced in the crude cell extract is directly compared to the soluble fraction by visual inspection of a Coomassie blue stained gel. A parallel experiment is run to determine if the fusion protein is a good substrate for TEV protease and whether or not the cleaved target protein remains soluble after it is released from His_6MBP.

11.3.2.1. Selecting a Host Strain of E. coli

Sometimes delaying the induction of TEV protease until the fusion protein substrate has had time to accumulate in the cells results in greater solubility of the passenger protein after cleavage *(13, 14)*. However, to achieve regulated expression of TEV protease, the *in vivo* processing experiment must be performed in a strain of *E. coli* that produces the Tet repressor, such as BL21-Pro or DH5α–Pro (B&D Clontech, Palo Alto, CA, USA). The Tet repressor blocks the synthesis of TEV protease mRNA and allows the production of the enzyme to be regulated independently of the IPTG-inducible His_6MBP fusion vector. Independent production of TEV protease from the expression vector pRK603 *(13)* is initiated by adding anhydrotetracycline to the cell culture, usually two hours after induction of the fusion protein with IPTG. We prefer using BL21-Pro because of its robust growth characteristics and the fact that it lacks two proteases (Lon and OmpT) that are present in many *E. coli* K12 strains such as DH5α-Pro.

11.3.2.2. Protein Expression

1. Transform competent BL21-Pro or DH5α-Pro cells that already contain pRK603 (*see* **Notes 5** and **6**) with the His_6MBP fusion protein expression vector and spread them on LB agar plates containing ampicillin (100 µg/ml) and kanamycin (35 µg/ml). Incubate the plate overnight at 37°C.
2. Inoculate 2–5 ml of LB medium containing ampicillin (100 µg/ml) and kanamycin (35 µg/ml) in a culture tube or shake-flask with a single colony from the plate. Grow to saturation overnight at 37°C with shaking at 250 rpm.
3. The next morning, inoculate 50 ml of the same medium in a 250 ml baffled-bottom flask with 0.5 ml of the saturated overnight culture.

4. Grow the cells at 37°C with shaking to mid-log phase ($OD_{600\,nm}$ ~ 0.5).
5. Add IPTG (1 mM final concentration) and adjust the temperature to 30°C (*see* **Note 15**).
6. After 2 h, divide the culture into two separate flasks (ca. 20 ml in each). Label one flask "+" and the other "–".
7. Add anhydrotetracycline to the "+" flask (100 ng/ml final concentration).
8. After 2 more hours, measure the $OD_{600\,nm}$ of the cultures (dilute cells 1:10 in LB to obtain an accurate reading). An $OD_{600\,nm}$ of about 3–3.5 is normal, although lower densities are possible. If the density of either culture is much lower than this, it may be necessary to adjust the volume of the samples that are analyzed by SDS-PAGE.
9. Transfer 10 ml of each culture to a 15 ml conical centrifuge tube and pellet the cells by centrifugation (4,000 × *g*) at 4°C.
10. Resuspend the cell pellets in 1 ml of lysis buffer and then transfer the suspensions to a 1.5 ml microcentrifuge tube.
11. Store the cell suspensions at –80°C overnight. Alternatively, the cells can be disrupted immediately by sonication (without freezing and thawing) and the procedure continued without interruption, as described below.

11.3.2.3. Sonication and Sample Preparation

1. Thaw the cell suspensions at room temperature, and then place them on ice.
2. Lyse the cells by sonication (*see* **Note 16**).
3. Prepare samples of the total intracellular protein from the "+" and "–" cultures (T + and T–, respectively) for SDS-PAGE by mixing 50 µl of each sonicated cell suspension with 50 µl of 2× SDS-PAGE sample buffer containing 20% (v/v) 2-mercaptoethanol.
4. Pellet the insoluble cell debris (and proteins) by centrifuging the sonicated cell suspension from each culture at maximum speed in a microcentrifuge for 10 min at 4°C.
5. Prepare samples of the soluble intracellular protein from the "+" and "–" cultures (S + and S–, respectively) for SDS-PAGE by mixing 50 µl of each supernatant from step 4 with 50 µl of 2× SDS-PAGE sample buffer containing 20% (v/v) 2-mercaptoethanol.

11.3.2.4. SDS-PAGE

We typically use pre-cast Tris-Glycine or NuPAGE gradient gels for SDS-PAGE to assess the yield and solubility of MBP fusion proteins (*see* **Note 7**). Of course, the investigator is free to choose any appropriate SDS-PAGE formulation, depending on the protein size and laboratory preference.

1. Heat the T–, T +, S– and S + protein samples at 90°C for about 5 min and then spin them at maximum speed in a microcentrifuge for 5 min.
2. Assemble the gel in the electrophoresis apparatus, fill it with SDS-PAGE running buffer, load the samples (10 µl each), and carry out the electrophoretic separation according to standard lab practices. T and S samples from each culture ("+" and "–") are loaded in adjacent lanes to allow easy assessment of solubility. Molecular weight standards may also be loaded on the gel, if desired.
3. Stain the proteins in the gel with GelCode® Blue reagent, PhastGel™ Blue R, or a suitable alternative.

11.3.2.5. Interpreting the Results

The MBP fusion protein should be readily identifiable in the T- sample after the gel is stained since it will normally be the most abundant protein in the cells. Molecular weight standards can also be used to corroborate the identity of the fusion protein band. If the S- sample contains a similar amount of the fusion protein, this indicates that it is highly soluble in *E. coli*. If little or no fusion protein is observed in the S- sample, then this is an indicator of poor solubility. Of course, a range of intermediate states is also possible.

If the fusion protein is an efficient substrate for TEV protease, then little of it will be present in the T + and S + samples. Instead, one should observe a prominent band at ca. 42 kDa that corresponds to the His_6MBP moiety and another prominent band migrating with the expected mobility of the passenger protein. If the fusion protein is a poor substrate for the protease, then the "+" samples will look similar to the "–" samples.

If the passenger protein is soluble after it is released from His_6MBP, then a similar amount will be present in the T + and S + lanes. At this point, some or all of the passenger protein may precipitate. If a substantial fraction of the passenger protein is insoluble, then troubleshooting may be necessary. Alternatively, an acceptable yield might still be obtained by scaling up cell production.

Examples of several possible outcomes are illustrated in **Fig. 11.3**. In panel A, roughly equal amounts of the His_6MBP-DHFR fusion protein are readily visible in lanes T- and S-, indicating that fusion protein is overexpressed and highly soluble. The results of adding anhydrotetracycline two hours after the induction of the fusion protein with IPTG (to induce the production of TEV protease) are seen in lanes T + and S +. The band migrating at ~42 kDa corresponds to His_6MBP and the smaller band migrating at ~21.5 kDa is the cleaved DHFR. Since nearly equal amounts of DHFR are present in lanes T + and S +, this indicates that virtually all of the DHFR retains its solubility after being cleaved from the His_6MBP. Similarly, the His_6MBP-TIMP fusion protein

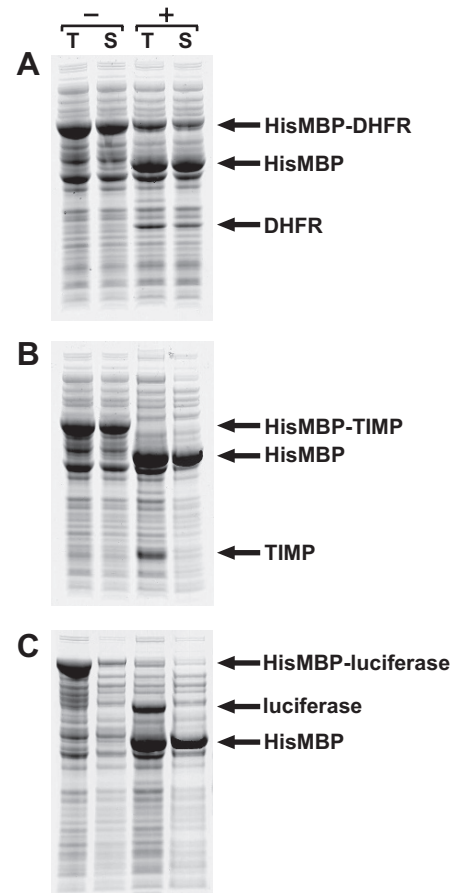

Fig. 11.3. Intracellular processing of His_6MBP fusion proteins by TEV protease. DHFR, TIMP and Luciferase were expressed from derivatives of pDEST-HisMBP in BL21-Pro cells that also contained the TEV protease expression vector pRK603 as described (*see* **Section 11.3.2**). "T" and "S" refer to the total and soluble fractions of the intracellular protein, respectively. All cultures were induced with IPTG to initiate the production of the His_6MBP fusion proteins. Samples marked "+" were induced with anhydrotetracycline to initiate the production of TEV protease 2 h after the addition of IPTG, whereas samples marked "–" were not induced with anhydrotetracyline. The results of this experiment are discussed in the Interpreting the Results section (*see* **Section 11.3.2.5**).

is also highly expressed and soluble (panel B), but in this instance the TIMP precipitates after digestion of the fusion protein *in vivo* with TEV protease. Hence, TIMP is an example of a passenger protein that is only temporarily soluble while it is fused to MBP. The His_6MBP-luciferase fusion protein is included here to illustrate the point that not all aggregation-prone proteins can be made soluble by fusing them to MBP (panel C).

Previous studies have shown that, despite its high solubility in *E. coli*, GST does not have the ability to promote the solubility of its fusion partners *(2, 15)*. Consequently, the solubility of a GST fusion protein is a good indicator of the solubility of a passenger protein in its unfused state. The results obtained when the three passenger proteins utilized here (DHFR, TIMP and luciferase)

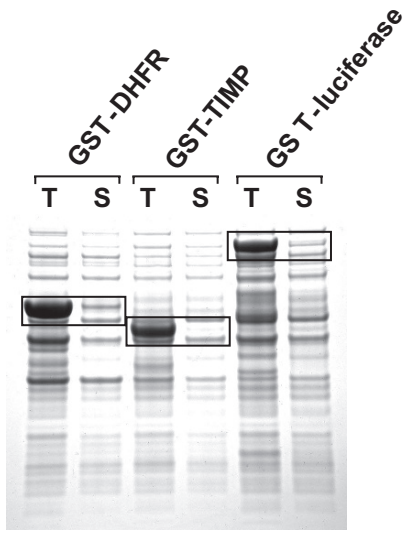

Fig. 11.4. Insolubility of GST fusion proteins. Entry clones encoding the three passenger proteins utilized in this study (DHFR, TIMP and Luciferase) were recombined into the destination vector pDEST3 (Invitrogen), which is designed to produce GST fusion proteins. Samples of the total (T) and soluble (S) intracellular protein from cells expressing the GST fusion proteins were prepared and analyzed by SDS-PAGE. The boxes indicate the positions of the fusion proteins on the Coomassie-stained gel.

were expressed as GST fusion proteins are shown in **Fig. 11.4**. Note that all three of the GST fusion proteins exhibit very poor solubility, indicating that the solubility of DHFR and TIMP was definitely enhanced by fusing them to His_6MBP.

Once it has been established that an aggregation-prone passenger protein can be rendered soluble by fusing it to His_6MBP, that the fusion protein can be cleaved by TEV protease, and that the passenger protein remains soluble after it is released from His_6MBP, then the passenger protein is ready to be purified on a large scale. Detailed instructions for how to accomplish this, using a generic IMAC-based protocol, have been described elsewhere *(12)*.

11.3.2.6. Checking the Biological Activity of the Passenger Protein

Occasionally, a passenger protein may accumulate in a soluble but biologically inactive form after intracellular processing of a His_6MBP fusion protein. Exactly why this occurs is uncertain, but we suspect that fusion to MBP somehow enables certain proteins to form soluble aggregates or evolve into kinetically trapped folding intermediates that are no longer susceptible to aggregation. Therefore, although solubility after intracellular processing is a useful indicator of a passenger protein's folding state in most cases, it is not absolutely trustworthy and can occasionally be misleading. For this reason, we strongly recommend that a biological assay (if available) or biophysical techniques be employed at an early stage to confirm that the passenger protein is in its native conformation.

11.4. Notes

1. We recommend a proofreading polymerase such as *Pfu* Turbo (Stratagene, La Jolla, CA, USA), Platinum *Pfx* (Invitrogen, Carlsbad, CA, USA), or Deep Vent (New England Biolabs, Beverly, MA, USA) to minimize the occurrence of mutations during PCR.
2. We typically purify fragments by electrophoresis using precast E-gels purchased from Invitrogen. However, suitable equipment and reagents for horizontal agarose gel electrophoresis can be purchased from a wide variety of scientific supply companies. DNA fragments are extracted from slices of the ethidium bromide-stained gel using a QIAquick gel extraction kit (Qiagen) in accordance with the instructions supplied with the product.
3. Any *gyrA*$^+$ strain of *E. coli* can be used. We prefer competent DH5α cells (Invitrogen) because they are easy to use and have high transformation efficiencies.
4. We prefer the QIAprep Spin miniprep kit (Qiagen), but similar kits can be obtained from a wide variety of vendors.
5. While any method for the preparation of competent cells can be used (e.g., $CaCl_2$) ***(16)***, we prefer electroporation because of the high transformation efficiency that can be achieved. Detailed protocols for the preparation of electrocompetent cells and electrotransformation procedures can be obtained from the electroporator manufacturers (e.g., Bio-Rad, BTX, Eppendorf). Briefly, the cells are grown in 1 L of LB medium (with antibiotics, if appropriate) to mid-log phase (OD_{600} ~0.5) and then chilled on ice. The cells are pelleted at 4°C, resuspended in 1 L of ice-cold H_2O and then pelleted again. After several such washes with H_2O, the cells are resuspended in 3–4 ml of 10% glycerol, divided into 50 µl aliquots, and then immediately frozen in a dry ice/ethanol bath. The electrocompetent cells are stored at –80°C. Immediately prior to electrotransformation, the cells are thawed on ice and mixed with 10–100 ng of DNA (e.g., a plasmid vector or a Gateway reaction). The mixture is placed into an ice-cold electroporation cuvette and electroporated according to the manufacturers recommendations (e.g., a 1.5 kV pulse in a cuvette with a 1 mm gap). 0.450 ml of SOC medium ***(16)*** is immediately added to the cells and they are allowed to grow at 37°C with shaking (ca. 250 rpm) for 1 h. 5–200 µl of the cells are then spread on an LB agar plate containing the appropriate antibiotic(s).
6. If the open reading frame encoding the passenger protein contains codons that are rarely used in *E. coli* (http://www.doembi.ucla.edu/cgi/cam/racc.html), this can adversely affect

the yield of an MBP fusion protein. In such cases, it is advisable to introduce an additional plasmid into the host cells that carries the cognate tRNA genes for rare codons. The pRIL plasmid (Stratagene, La Jolla, CA, USA) is a derivative of the p15A replicon that carries the *E. coli argU*, *ileY*, and *leuW* genes, which encode the cognate tRNAs for AGG/AGA, AUA and CUA codons, respectively. pRIL is selected for its resistance to chloramphenicol. In addition to the tRNA genes for AGG/AGA, AUA and CUA codons, the pRARE accessory plasmid in the Rosetta™ host strain (Novagen, Madison, WI, USA) also includes tRNAs for the rarely used CCC and GGA codons. Like pRIL, the pRARE plasmid is a chloramphenicol-resistant derivative of the p15A replicon. Both of these tRNA accessory plasmids are compatible with derivatives of pDEST-His-MBP. On the other hand, they are incompatible with the vector pRK603 that we use for intracellular processing experiments (*see* **Section 11.3.2**). Nevertheless, because pRK603 and the tRNA accessory plasmids have different antibiotic resistance markers, it is possible to force cells to maintain both plasmids by simultaneously selecting for kanamycin and chloramphenicol resistance. Alternatively, the kanamycin-resistant TEV protease expression vector pKM586, a pRK603 derivative with the replication machinery of a pSC101 replicon, which can be obtained from the authors, can be stably maintained in conjunction with p15A-type tRNA plasmids.

7. We find it convenient to use pre-cast gels for SDS-PAGE (e.g., 1.0 mm × 10 well, 10–20% Tris-Glycine gradient), running buffer, and electrophoresis supplies from Invitrogen (Carlsbad).
8. Alternatively, the PCR reaction can be performed in two separate steps, using primers N1 and C in the first step and primers N2 and C in the second step. The PCR amplicon from the first step is used as the template for the second PCR. All primers are used at the typical concentrations for PCR in the two-step protocol.
9. The PCR reaction can be modified in numerous ways to optimize results, depending on the nature of the template and primers (*see* **(16)** (Vol. 2, Chapter 8) for more information).
10. PCR cycle conditions can also be varied. For example, the extension time should be increased for especially long genes. A typical rule-of-thumb is to extend for 60 s/kb of DNA.
11. This "one-tube" Gateway protocol bypasses the isolation of an "entry clone" intermediate. However, the entry clone may be useful if the investigator intends to experiment with additional Gateway destination vectors, in which case the BP and LR reactions can be performed sequentially in separate steps; detailed instructions are included with the Gateway

PCR kit. Alternatively, entry clones can easily be regenerated from expression clones via the BP reaction, as described in the instruction manual.

12. Clonase enzyme mixes should be thawed quickly on ice and then returned to the −80°C freezer as soon as possible. It is advisable to prepare multiple aliquots of the enzyme mixes the first time that they are thawed in order to avoid repeated freeze-thaw cycles.
13. At this point, we remove a 5 µl aliquot from the reaction and add it to 0.5 µl of proteinase K solution. After 10 min at 37°C, we transform 2 µl into 50 µl of competent DH5α cells (*see* **Note 3**) and spread 100–200 µl on an LB agar plate containing kanamycin (35 µg/ml), the selective marker for pDONR221. From the number of colonies obtained, it is possible to gauge the success of the BP reaction. Additionally, entry clones can be recovered from these colonies in the event that no transformants are obtained after the subsequent LR reaction.
14. If very few or no ampicillin-resistant transformants are obtained after the LR reaction, the efficiency of the process can be improved by incubating the BP reaction overnight.
15. 30°C is the optimum temperature for TEV protease activity. At 37°C, the protease does not fold properly in *E. coli* and little processing will occur. Reducing the temperature also improves the solubility of some MBP fusion proteins.
16. We routinely break cells in a 1.5 ml microcentrifuge tube on ice with two or three 30 s pulses using a VCX600 sonicator (Sonics & Materials, Newtown, CT, USA) with a microtip at 38% power. The cells are cooled on ice between pulses.

Acknowledgement

This research was supported by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research.

References

1. Chayen, N. E. (2004) Turning protein crystallization from an art into a science. *Curr. Opin. Struct. Biol.* 14, 577–583
2. Kapust, R. B., and Waugh, D. S. (1999) *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of its fusion partners. *Protein Sci.* 8, 1668–1674
3. Shih, Y.-P., Kung, W.-M., Chen, J.-C., Yeh, C.-H., Wang, A. H.-J., and Wang, T.-F. (2002) High-throughput screening of

soluble recombinant proteins. *Protein Sci.* 11, 1714–1719

4. Hammarstrom, M., Hellgren, N., van Den Berg, S., Berglund, H., and Hard, T. (2002) Rapid screening for improved solubility of small human proteins produced as fusion proteins in *Escherichia coli*. *Protein Sci.* 11, 313–321
5. Dyson, M. R., Shadbolt, P. S., Vincent, K. J., Perera, R. L., and McCafferty, J. (2004) Production of soluble mammalian proteins in *Escherichia coli*: identification of protein features that correlate with successful expression. *BMC Biotech.* 4, 32 doi:10.1186/1472–6750/4/32
6. Busso, D., Delagoutte-Busso, B., and Moras, D. (2005) Construction of a set of gateway-based destination vectors for high-throughput cloning and expression screening in *Escherichia coli*. *Anal. Biochem.* 343, 313–321
7. Korf, U., Kohl, T., van der Zandt, H., Zahn, R., Schleeger, S., Ueberle, B., Wandschneiderr, S., Bechtel, S., Schnolzer, M., Ottleben, H., Wiemann, S., and Poustka, A. (2005) Large-scale protein expression for proteomics research. *Proteomics* 5, 3571–3580
8. Riggs, P. (2000) Expression and purification of recombinant proteins by fusion to maltose-binding protein. *Mol. Biotechnol.* 15, 51–63
9. Pryor, K. A., and Leiting, B. (1997) High-level expression of soluble protein in *Escherichia coli* using a His6-tag and maltose binding protein double-affinity fusion system. *Protein Expr. Purif.* 10, 309–319
10. Routzahn, K. M. and Waugh, D. S. (2002) Differential effects of supplementary affinity tags on the solubility of MBP fusion proteins. *J. Struct. Funct. Genomics* 2, 83–92
11. Nallamsetty, S., Austin, B. P., Penrose, K. J., and Waugh, D. S. (2005) Gateway vectors for the production of combinatorially-tagged His6-MBP fusion proteins in the cytoplasm and periplasm of *Escherichia coli*. *Protein Sci.* 14, 2964–2971
12. Tropea, J. E., Cherry, S., Nallamsetty, S., Bignon, C., and Waugh, D. S. (2007) A generic method for the production of recombinant proteins in *Escherichia coli* using a dual hexahistidine-maltose-binding protein affinity tag. *Methods Mol. Biol.* 363, 1–19
13. Kapust, R. B. and Waugh, D. S. (2000) Controlled intracellular processing of fusion proteins by TEV protease. *Protein Expr. Purif.* 19, 312–318
14. Fox, J. D. and Waugh, D. S. (2003) Maltose-binding protein as a solubility enhancer. *Methods Mol. Biol.* 205, 99–117
15. Fox, J. D., Routzahn, K. M., Bucher, M. H., and Waugh, D. S. (2003) Maltodextrin-binding proteins from diverse bacteria and archaea are potent solubility enhancers. *FEBS Lett.* 537, 53–57
16. Sambrook, J. and Russell, D. W. (2001) *Molecular cloning: a laboratory manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY(2002) High-throughput screening of soluble recombinant proteins. *Protein Sci.* 11, 1714–1719