7

Maltose-Binding Protein as a Solubility Enhancer

Jeffrey D. Fox and David S. Waugh

1. Introduction

A major impediment to the production of recombinant proteins in *Escherichia coli* is their tendency to accumulate in the form of insoluble and biologically inactive aggregates known as inclusion bodies. Although it is sometimes possible to convert aggregated material into native, biologically-active protein, this is a time consuming, labor-intensive, costly, and uncertain undertaking ***(1)***. Consequently, many tricks have been employed in an effort to circumvent the formation of inclusion bodies ***(2)***. One approach that shows considerable promise is to exploit the innate ability of certain proteins to enhance the solubility of their fusion partners. Although it was originally thought that virtually any highly soluble protein could function as a general solubilizing agent, this has not turned out to be the case. In a direct comparison with glutathione S-transferase (GST) and thioredoxin, maltose-binding protein (MBP) was decidedly superior at solubilizing a diverse collection of aggregation-prone passenger proteins ***(3)***. Moreover, some of these proteins were able to fold into their biologically active conformations when fused to MBP. It is not entirely clear why MBP is such a spectacular solubilizing agent, but there is some evidence to suggest that it may be able to function as a general molecular chaperone in the context of a fusion protein by temporarily sequestering aggregation-prone folding intermediates of its fusion partners and preventing their self association ***(3–6)***. The ability to promote the solubility of its fusion partners is not an exclusive attribute of MBP (*see* Chapters 8 and 9), but to the best of our knowledge MBP is the only general solubilizing agent that is also a natural affinity tag. Consequently, we consider MBP to be the best "first choice" fusion partner for the production of recombinant proteins in *E. coli.*

From: *Methods in Molecular Biology, vol. 205, E. coli Gene Expression Protocols*
Edited by: P. E. Vaillancourt © Humana Press Inc., Totowa, NJ

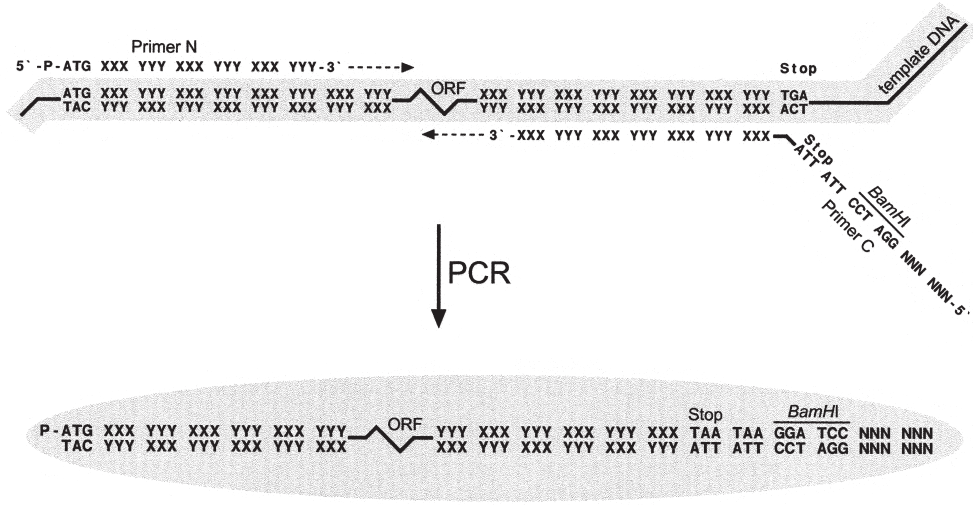


Fig. 1. PCR strategy for conventional cloning into pMAL-c2X. The template DNA is amplified with primers N and C. The primers are designed to base-pair with 20–25 bp of the 5' and 3' ends of the coding region respectively. Primer N is phosphorylated to allow blunt-ligation with the *Xmn*I site of pMAL-c2X. Primer C includes a 5' extension with a *BamH*I site for ligation with the *BamH*I site in pMAL-c2X.

Basic protocols for constructing MBP fusion vectors and for assessing the solubility and folding state of the fusion proteins are described herein. Special attention is given to a rapid and efficient method of generating fusion vectors by recombinational cloning. In addition, a method is described to quickly evaluate the folding state of a passenger protein by intracellular processing of a fusion protein with TEV protease. More detailed descriptions of commercially available MBP fusion vectors and methods for the purification of MBP fusion proteins by amylose affinity chromatography have been presented elsewhere ***(7)***.

2. Materials

2.1. Conventional Vector Construction

1. The desired pMAL vector (New England Biolabs, Beverly, MA, USA).
2. Reagents and thermostable DNA polymerase for PCR amplification (*see* **Note 1**).
3. Appropriate synthetic oligodeoxyribonucleotide primers for PCR amplification (*see* **Fig. 1**).
4. Restriction enzymes and matching reaction buffers for screening putative clones.
5. Tris-Acetate-EDTA (TAE)-agarose, ethidium bromide, and an apparatus 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.

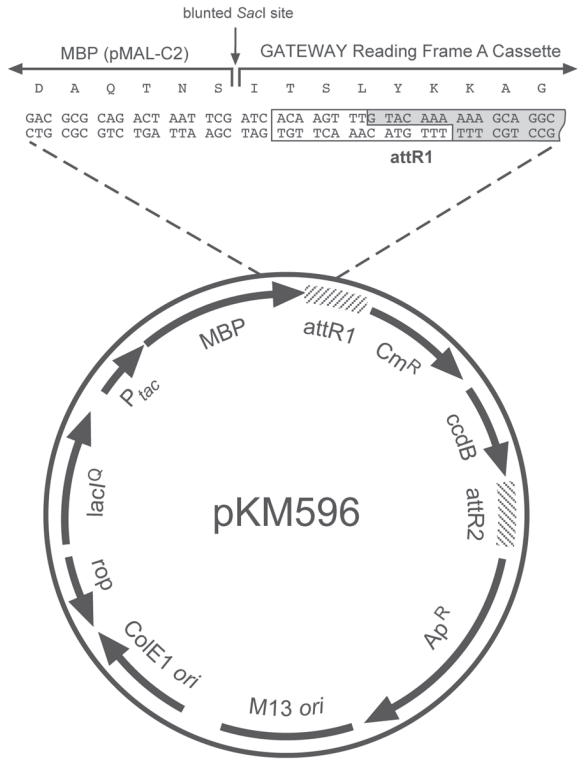
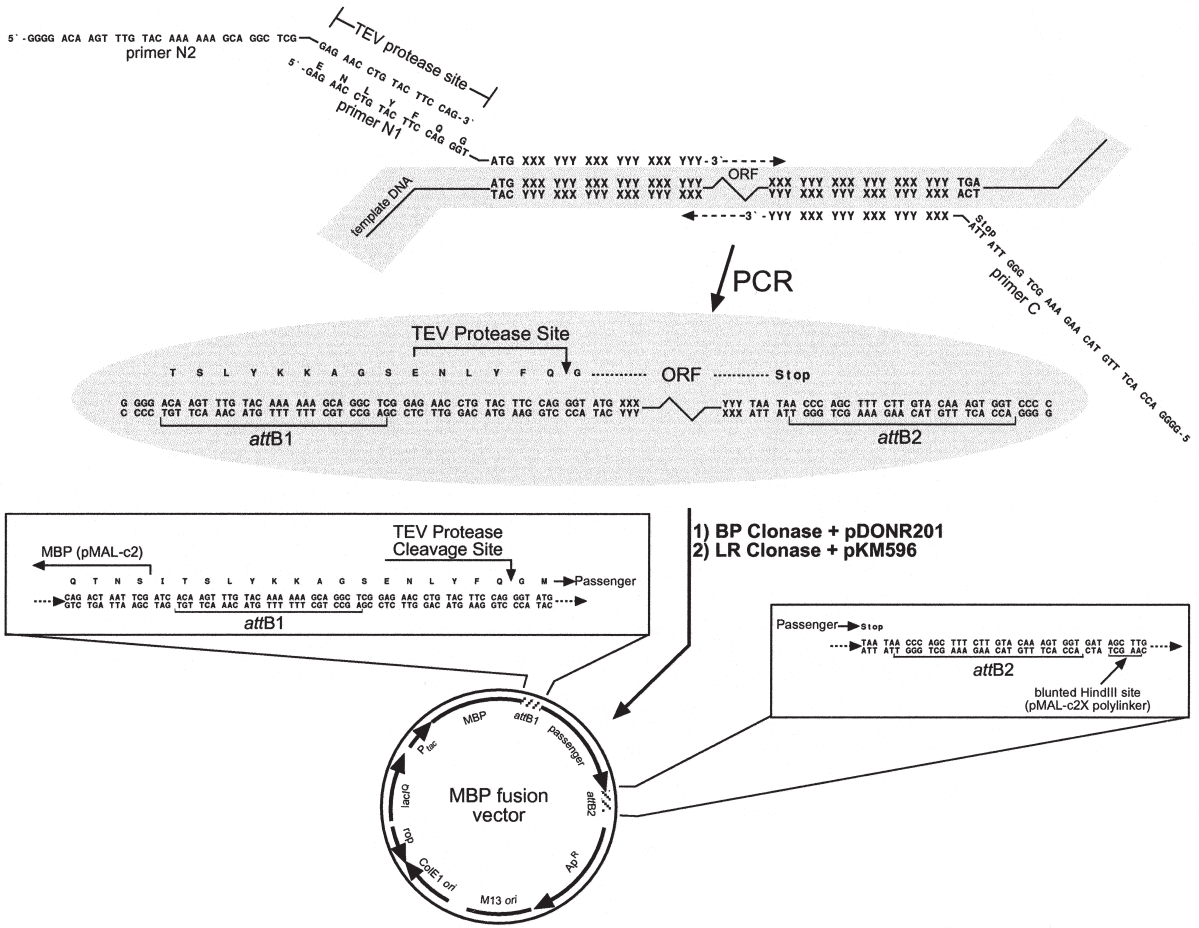


Fig. 2. Schematic representation of the Gateway™ destination vector pKM596. This vector can be recombined with an entry vector that contains an ORF of interest, via the LxR reaction, to generate an MBP fusion protein expression vector.

7. T4 DNA ligase, reaction buffer, and ATP.
8. Competent *E. coli* cells (e.g., DH5α or similar; *see* **Note 3**).
9. Luria-Bertani (LB) medium and LB agar plates containing ampicillin (100 µg/mL). LB medium: Add 10 g bactotryptone, 5 g yeast extract, 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 diameter).
10. Reagents for small-scale plasmid DNA isolation (*see* **Note 4**).

2.2. Recombinational Vector Construction

1. The Gateway™ destination vector pKM596 (*see* **Fig. 2**).
2. Reagents and thermostable DNA polymerase for PCR amplification (*see* **Note 1**).
3. Synthetic oligodeoxyribonucleotide primers for PCR amplification (*see* **Fig. 3**).


5`-GGGG ACA AGT TTG TAC AAA AAA GCA GGC TCG
primer N2
TEV protease site
GAG AAC CTG TAC TTC CAG-3`
E N L Y F Q G
5`-GAG AAC CTG TAC TTC CAG GGT
primer N1
ATG XXX YYY XXX YYY XXX YYY-3`
template DNA
ATG XXX YYY XXX YYY XXX YYY
TAC YYY XXX YYY XXX YYY XXX
ORF
XXX YYY XXX YYY XXX YYY TGA
YYY XXX YYY XXX YYY XXX ACT
5`-YYY XXX YYY XXX YYY XXX-3`
Stop
ATT GGG TCG AAA GAA CAT GTT TCA CCA GGGG-5
primer C
PCR
TEV Protease Site
T S L Y K K A G S E N L Y F Q G ..... ORF ..... Stop
G GGG ACA AGT TTG TAC AAA AAA GCA GGC TCG GAG AAC CTG TAC TTC CAG GGT ATG XXX
C CCC TGT TCA AAC ATG TTT TTT CGT CCG AGC CTC TTG GAC ATG AAG GTC CCA TAC YYY
YYY TAA TAA CCC AGC TTT CTT GTA CAA AGT GGT CCC C
XXX ATT ATT GGG TCG AAA GAA CAT GTT TCA CCA GGG G
attB1
attB2
MBP (pMAL-c2)
TEV Protease Cleavage Site
Q T N S I T S L Y K K A G S E N L Y F Q G M Passenger
CAG ACT AAT TCG ATC ACA AGT TTG TAC AAA AAA GCA GGC TCG GAG AAC CTG TAC TTC CAG GGT ATG
GTC TGA TTA AGC TAG TGT TCA AAC ATG TTT TTT CGT CCG AGC CTC TTG GAC ATG AAG GTC CCA TAC
attB1
1) BP Clonase + pDONR201
2) LR Clonase + pKM596
Passenger
Stop
TAA TAA CCC AGC TTT CTT GTA CAA AGT GGT GAT AGC TTG
ATT ATT GGG TCG AAA GAA CAT GTT TCA CCA CTA TCG AAC
attB2
blunted HindIII site
(pMAL-c2X polylinker)
MBP fusion vector
MBP
attB1
passenger
attB2
Ap R
M13 ori
ColE1 ori
rop
lacIq
P tac

4. Tris-EDTA (TE) buffer: 10 m*M* Tris-HCl, pH 8.0, 1 m*M* ethylenediaminetetraacetic acid (EDTA).
5. TAE-agarose and an apparatus 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. Competent DB3.1 cells (Invitrogen, Carlsbad, CA, USA) for propagating pKM596 and pDONR201 (*see* **Note 3**).
8. Gateway™ PCR Cloning System (Invitrogen, Carlsbad, CA, USA).
9. LB medium and LB agar plates containing ampicillin (100 µg/mL). *See* **Subheading 2.1.9.** for preparation.
10. Reagents for small-scale plasmid DNA isolation (*see* **Note 4**).

2.3. Assessing the Solubility of MBP Fusion Proteins

1. Competent BL21-CodonPlus™-RIL cells (Stratagene, La Jolla, CA, USA) (*see* **Notes 3** and **5**).
2. LB agar plates and broth containing both ampicillin (100 µg/mL) and chloramphenicol (30 µg/mL). See **Subheading 2.1.**, item 9 for LB broth, LB agar, and ampicillin stock solution recipes. Prepare stock solution of 30 mg/mL chloramphenicol in ethanol and store at 4°C for up to 1 mo. Dilute antibiotics 1000-fold into LB medium or molten LB agar.
3. Isopropyl-thio-β-D-galactopyranoside (IPTG). Prepare a stock solution of 200 m*M* in H_2O and filter sterilize. Store at –20°C.
4. Shaker/incubator set at 37°C.
5. 250-mL baffle-bottom flasks (sterile).
6. Cell lysis buffer: 20 m*M* Tris-HCl, pH 8.0, 1 m*M* EDTA.
7. Sonicator (with microtip).
8. 2X Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) sample buffer: 100 m*M* Tris-HCl, pH 6.8, 200 m*M* dithiothreitol (DTT), 4% SDS, 0.2% bromphenol blue, 20% glycerol.
9. SDS-PAGE gel, electrophoresis apparatus, and running buffer (*see* **Note 6**).
10. Gel stain (e.g., Gelcode® Blue from Pierce, Rockford, IL, USA)

Fig. 3. *(opposite page)* Construction of an MBP 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 pDONR201 vector to generate an entry clone via the BxP reaction. This entry clone is subsequently recombined with pKM596 and LxR Clonase to yield the final MBP fusion vector.

2.4. Intracellular Processing of MBP Fusion Proteins by TEV Protease

1. Competent DH5αPRO or BL21PRO cells (Clontech, Palo Alto, CA, USA) containing the TEV protease expression vector pRK603 and the tRNA plasmid pKC1 (*see* **Notes 3**, **5**, and **7**).
2. A derivative of pKM596 (*see* **Subheading 3.3.1.**), or a pMAL vector that produces an MBP fusion protein with a TEV protease recognition site in the linker between domains.
3. LB medium and agar plates containing ampicillin (100 µg/mL), kanamycin (25 µg/mL), and chloramphenicol (30 µg/mL). See **Subheadings 2.1.**, **item 9** and **2.3., item 2.** for preparation. Prepare a stock solution of 25 mg/mL kanamycin in H_2O and store at 4°C for up to 1 mo. Dilute antibiotics 1000-fold into LB medium or molten LB agar.
4. Anhydrotetracycline. Prepare a 1000X stock solution by dissolving in ethanol at 100 µg/mL. Store in a foil-covered tube at –20°C.
5. Other materials as in **Subheading 2.3.**

3. Methods

3.1. Construction of MBP Fusion Vectors by Conventional Techniques

Workers are encouraged to consult the instructions and technical literature available from New England Biolabs related to their MBP fusion product line.

3.1.1. Selecting a pMAL Vector

Before constructing an expression vector, the proper plasmid backbone must be selected. A range of choices is currently available from New England Biolabs. These include pMAL-c2X, pMAL-p2X, pMAL-c2E, pMAL-p2E, pMAL-c2G, and pMAL-p2G; where p or c indicates periplasmic or cytoplasmic localization; and X, E, or G denote the identity of the protease cleavage site that is present in the fusion protein linker. (*See* **Subheading 3.2.** for more information about linkers and proteases.) Many labs also still have in their possession the older vectors pMAL-c2 and/or pMAL-p2 (*see* **Note 8**).

The properties of the passenger protein dictate the proper choice between cytoplasmic and periplasmic expression of an MBP fusion protein. This relates mainly to whether disulfide bonds are expected in the passenger, in which case the more oxidizing environment of the periplasm may be desirable. The methods described in this article pertain specifically to the production of MBP fusion proteins in the cytoplasm. In general, the yield of fusion protein is much greater in the cytoplasm, and purification by amylose affinity chromatography usually removes the majority of contaminating cytoplasmic proteins (*see* **Note 9**).

3.1.2. Assembling an Expression Vector

1. Assuming that pMAL-c2X has been selected (cloning strategies are similar for all six pMAL vectors) (*see* **Note 10**), a suitable restriction fragment encompassing the open reading frame (ORF) of interest must be prepared for ligation with the vector DNA. PCR is by far the most efficient means by which to generate this fragment. For general PCR protocols, *see* **ref. *8***. Typically, oligodeoxyribonucleotide primers are used to amplify the ORF while also extending either or both ends to introduce appropriate restriction site(s) for cloning (*see* **Fig. 1**). If the *Xmn*I site in the pMAL-c2X polylinker is to be used for cloning then the 5' PCR primer (Primer N) must either be phosphorylated or include a properly positioned blunt restriction site (*see* **Note 11**). The 3' extension adds a *Bam*HI site immediately after the stop codon (*see* **Note 12**).
2. The PCR product is digested with the appropriate restriction enzyme(s) (e.g., *Bam*HI in the example in **Fig. 1**) and purified by agarose gel electrophoresis (*see* **Note 2**).
3. pMAL-c2X is digested with *Xmn*I and *Bam*HI followed by gel purification of the large fragment (*see* **Note 2**).
4. The PCR fragment and digested vector backbone are combined and incubated with T4 DNA ligase and ATP (*see* **Note 13**).
5. The products of the ligation reaction are transformed into an appropriate *E. coli* strain (e.g., DH5α; *see* **Note 3**) and then spread on LB agar plates containing 100 μg/mL ampicillin. The plates are incubated overnight at 37°C.
6. Plasmid DNA is isolated from saturated cultures that were inoculated with individual ampicillin-resistant colonies and screened by restriction analysis to identify clones with the desired properties.
7. It is advisable to submit putative clones for sequence analysis to verify the proper construction and lack of PCR-induced mutations.

3.2. Protease Cleavage Sites

In almost every case, the investigator would like to obtain the protein of interest free from its fusion partner and with a minimum of extraneous amino acids. New England Biolabs offers vectors with three different options for protease cleavage: factor Xa, enterokinase, and genenase I ***(7)***. However, we have found that the tobacco etch virus (TEV) protease, which can be purchased from Invitrogen, is superior to the three alternatives offered by New England Biolabs. The major advantage of this protease is its exceptionally high specificity. In contrast to factor Xa, enterokinase, and thrombin, there have never been any documented reports of cleavage by TEV protease at locations other than the designed site in fusion proteins. However, New England Biolabs does not offer a pMAL vector with a TEV protease cleavage site already in the linker. Therefore, to utilize this protease, a recognition site must be introduced by PCR. For an example of a TEV protease site introduced by PCR, *see* **Fig. 3**.

3.3. Construction of MBP Fusion Vectors by Recombinational Cloning

Recombinational cloning can greatly simplify the construction of MBP fusion vectors. Although several different methods for recombinational cloning have been described, we strongly recommend the Gateway™ Cloning System 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, the investigator is encouraged to consult the technical literature supplied by Invitrogen.

3.3.1. Cloning with Gateway™

To utilize the Gateway™ system for the production of MBP fusion proteins, one must first construct or obtain a suitable "destination vector". Currently there are no commercial sources for such vectors. An example of a destination vector that can be used to produce MBP fusion proteins (pKM596) is shown in **Fig. 2**. pKM596 was constructed by replacing the DNA between the *Sac*I and *Hin*dIII restriction sites in the New England Biolabs vector pMAL-c2 with the RfA Gateway™ Cloning Cassette. The Gateway™ cassette consists of two different recombination sites (*att*B1 and *att*B2) separated by DNA encoding two gene products: chloramphenicol acetyl transferase, which confers resistance to chloramphenicol, and the DNA gyrase poison CcdB. The former marker provides a positive selection for the presence of the cassette, which is useful when one is constructing a destination vector. The latter gene product provides a negative selection against the donor vector and various recombination intermediates so that only the desired recombinant is obtained when the end products of the recombinational cloning reaction are transformed into *E. coli*. pKM596 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) that renders the cells immune to the action of CcdB.

The Gateway™ Cloning System has several noteworthy advantages. First, it is much faster and more efficient than conventional cloning techniques that utilize restriction endonucleases and DNA ligase. Second, because it does not rely on restriction endonucleases to generate substrates for ligation, Gateway™ cloning is never complicated by the existence of restriction sites within the ORF of interest that are also used for cloning. In fact, with the exception of the gene-specific primers that are used for PCR amplification, the Gateway protocol is completely generic and therefore readily amenable to automation. Finally, once an ORF has been cloned into a Gateway™ vector, it can easily be transferred by recombinational cloning into a wide variety of destination vectors that are available from Invitrogen. This gives the investigator the flexibility to

experiment with various modes of expression (e.g., different fusion tags) and/or hosts. There is even a destination vector for yeast two-hybrid screening.

3.3.2. An Abbreviated Gateway™ Cloning Protocol

The investigator is encouraged to refer to the detailed protocols in the technical literature from Invitrogen. The easiest way to construct an MBP fusion vector by recombinational cloning is to start with a PCR amplicon wherein the ORF of interest is bracketed by *att*B1 and *att*B2 sites on its N- and C-termini, respectively, which can be generated by amplifying the target ORF with PCR primers that include the appropriate *att*B sites as 5' unpaired extensions (*see* **Fig. 3**). The 3' ends of these PCR primers are chosen so that the primer will be able to form 20–25 bp with the template DNA. So that the passenger protein can be separated from MBP, a target site for TEV protease (or an alternative reagent) is also incorporated between the N-terminus of the ORF and the *att*B1 site in this PCR amplicon. Although it is possible to accomplish this by using a single N-terminal PCR primer for each gene, typically on the order of 75 nucleotides long, we have found that it is convenient to perform the PCR amplification with two N-terminal primers instead, as outlined in **Fig. 3**. Two gene-specific primers (N1 and C) are required for each ORF. The C-terminal primer (C) includes the *att*B2 recombination site as a 5' extension. The 5' extension of the N-terminal primer (N1) includes a recognition site for TEV protease. The PCR product generated by these two primers is subsequently amplified by primers N2 and C to yield the final product. Primer N2 anneals to the TEV protease recognition site and includes the *att*B1 recombination site as a 5' extension. This generic PCR primer can be used to add the *att*B1 site to any amplicon that already contains the TEV protease recognition site at its N-terminal end. The PCR reaction is performed in a single step by adding all three primers to the reaction at once (*see* **Note 14**). 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 20-fold lower.

1. The PCR reaction mix is prepared as follows (*see* **Note 15**): 1 µL template DNA (~10 ng/µL), 10 µL thermostable DNA polymerase 10X buffer, 16 µL dNTP solution (1.25 m*M* each), 2.5 µL primer N1 (~1 µ*M*, or 13 ng/µL for a 40 mer), 2.5 µL primer N2 (~20 µ*M*, or 260 ng/µL for a 40 mer), 2.5 µL primer C (~20 µ*M*, or 260 ng/µL for a 40 mer), 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 (*see* **Note 16**): Initial melt: 94°C, 5 min, 25 cycles of 94°C, 30 s (melting); 55°C, 30 s (annealing); 72°C, 60 s (extension), final extension: 72°C, 7 min, 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 MBP fusion vector, the PCR product is recombined first into pDONR201 to yield an entry clone intermediate (BxP reaction), and then into pKM596 (LxR reaction; *see* **Note 17**).
   a. Add to a microcentrifuge tube on ice: 300 ng of the PCR product in TE, 300 ng of pDONR201 DNA, 4 μL of BxP reaction buffer, and enough Tris-EDTA (TE) or H_2O to bring the total volume to 16 μL. Mix well.
   b. Thaw BP Clonase enzyme mix on ice (2 min) and then vortex briefly (2 s) twice (*see* **Note 18**).
   c. Add 4 μL of BP Clonase enzyme mix to the components in (a.) and vortex briefly twice
   d. Incubate the reaction at room temperature for at least 4 h (*see* **Note 19**).
   e. Add to the reaction: 1 μL of 0.75 *M* NaCl, 3 μL (ca. 450 ng) of the destination vector (pKM596), and 6 μL of LR Clonase enzyme mix (*see* **Note 18**). Mix by vortexing briefly.
   f. Incubate the reaction at room temperature for 3–4 h.
   g. Add 2.5 μL of 10X stop solution and incubate for 10 min at 37°C.
   h. Transform 2 μL of the reaction into 50 μL of competent DH5α cells (*see* **Note 3**).
   i. 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). Incubate the plate at 37°C overnight (*see* **Note 20**).
5. 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 have the expected structure. Alternatively, plasmids can be purified and screened by conventional restriction digests using appropriate enzymes. At this stage, we routinely sequence putative clones to ensure that there are no PCR-induced mutations.

3.4. Assessing the Solubility of MBP Fusion Proteins

The fusion protein is overproduced on a small scale to assess its solubility. The amount of fusion protein in the soluble fraction of the crude cell lysate is compared by SDS-PAGE with the total amount of fusion protein in the cells, and the results are analyzed by visual inspection of the stained gel.

3.4.1. Selecting a Host Strain of E. coli

The pMAL vectors and derivatives of pKM596 can be used in virtually any strain of *E. coli*. However, we prefer BL21 ***(9)*** because of its robust growth characteristics and the fact that it lacks two proteases (Lon and OmpT) present in most *E. coli* K12 strains. To improve the likelihood of obtaining a high yield of MBP fusion protein, we routinely use BL21 cells containing accessory plasmids that overproduce the cognate tRNAs for codons that are rarely used in *E. coli* (e.g., BL21-CodonPlus™-RIL or BL21 cells containing pKC1; *see* **Note 5**).

3.4.2. Pilot Expression Experiment

1. Inoculate 2–5 mL of LB medium containing ampicillin (100 µg/mL) and chloramphenicol (30 µg/mL) in a culture tube or shake-flask with BL21-CodonPlus™-RIL cells harboring an MBP fusion vector. Use a single colony from an LB agar plate containing ampicillin (100 µg/mL) and chloramphenicol (30 µg/mL) as the inoculum. Grow to saturation overnight at 37°C with shaking. *See* **Subheadings 2.1.9.** and **2.3.2.** for the preparation of LB medium and antibiotic stock solutions.
2. The next morning, inoculate 25 mL of the same medium in a 250-mL baffled-bottom flask with 0.25 mL of the saturated overnight culture. Label this flask "+". Also prepare a duplicate culture and label it "–".
3. Grow the cells at 37°C with shaking to mid-log phase (OD_{600nm} ~0.5), and then add IPTG to the "+" flask (1 m*M* final concentration).
4. Continue shaking for 3–4 h at 37°C.
5. Measure the OD_{600nm} of the cultures (dilute cells 1:10 in LB to obtain an accurate reading). An OD_{600nm} of approx 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 (*see* **Subheading 3.4.4.**).
6. Transfer 10 mL of each culture to a 15-mL conical centrifuge tube and pellet the cells by centrifugation.
7. Resuspend the cell pellets in 1 mL of lysis buffer (*see* **Subheading 2.3.6.**) and then transfer the suspensions to a 1.5-mL microcentrifuge tube.
8. Store the cell suspensions at –80°C overnight. Alternatively, the cells can be disrupted immediately by sonication (after freezing and thawing) and the procedure continued without interruption, as described below.

3.4.3. Sonication and Sample Preparation

1. Thaw the cell suspensions at room temperature, then place them on ice.
2. Lyse the cells by sonication (*see* **Note 21**).
3. Prepare samples of the total intracellular proteins from the induced and uninduced cultures (T+ and T–, respectively) for SDS-PAGE by mixing 50 µL of each sonicated cell suspension with 50 µL of 2X SDS-PAGE sample buffer.
4. Pellet the insoluble cell debris (and proteins) by centrifuging the sonicated cell suspension from the "+" culture at maximum speed in a microcentrifuge for 10 min.
5. Prepare a sample of the soluble intracellular proteins (S+) for SDS-PAGE by mixing 50 µL of the supernatant with 50 µL of 2X SDS-PAGE sample buffer.

3.4.4. SDS-PAGE

We typically use precast Tris-glycine SDS-PAGE gels (10–20% gradient) to assess the yield and solubility of MBP fusion proteins (*see* **Note 6**). 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+ and S+ protein samples at 90°C for approx 5 min and then spin them at maximum speed in a microcentrifuge for 5 min.
2. Dilute 10 µL of each sample with enough 1X SDS-PAGE sample buffer to fill the well of the gel.
3. Assemble the gel in the electrophoresis apparatus, fill it with SDS-PAGE running buffer, load the samples, and carry out the electrophoretic separation according to standard lab practices. T+ and S+ samples are loaded in adjacent lanes to allow easy assessment of solubility. Molecular weight standards may also be loaded on the gel, if desired.
4. Stain the proteins in the gel with GelCode® Blue reagent, Coomassie Brilliant Blue, or a suitable alternative.

3.4.5. 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, whereas there will be very little or no fusion protein in the T– (uninduced) sample. 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*. On the other hand, if little or no fusion protein is observed in the S+ sample, then it can be concluded that the protein is poorly soluble. Of course, a range of intermediate states is also possible. Yet, even when the solubility of the MBP fusion protein is relatively poor, an adequate amount of soluble material usually can be obtained by scaling up production.

3.4.6. Improving the Solubility of MBP Fusion Proteins

Not every MBP fusion protein will be highly soluble. However, solubility usually can be increased by reducing the temperature of the culture from 37 to 30°C or even lower during the time that the fusion protein is accumulating in the cells (i.e., after the addition of IPTG). In some cases, the improvement can be quite dramatic. It may also be helpful to reduce the IPTG concentration to a level that will result in partial induction of the fusion protein ***(2)***. The appropriate IPTG concentration must be determined empirically, but is generally in the range of 10–20 µ*M*. Under these conditions, longer induction times (18–24 h) are required to obtain a reasonable yield of fusion protein.

3.5. Determining the Folding State of a Passenger Protein

MBP is an excellent solubilizing agent, but some passenger proteins are unable to fold into their native conformations even after they have been rendered soluble by fusing them to MBP. These proteins evidently exist in a soluble but nonnative form that resists aggregation only as long as they remain fused to MBP. Consequently, it is difficult to assess the folding state of the

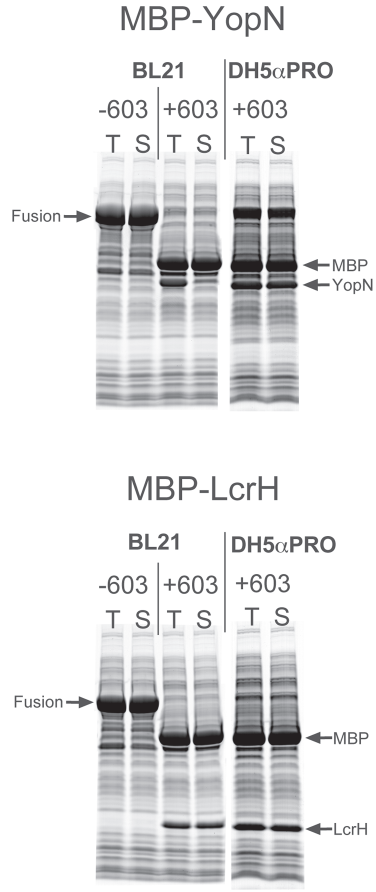


Fig. 4. Intracellular processing of MBP fusion proteins by TEV protease. Two MBP fusion proteins were processed in vivo by TEV protease to illustrate the utility of this method. YopN and LcrH are essential virulence factors from *Yersinia pestis* ***(12)***. They were both expressed from derivatives of pKM596 in *E. coli* strains BL21 and DH5αPRO. In all cases, the cells also contained the tRNA accessory plasmid pKC1 (*see* **Note 5**). TEV protease was produced in vivo by pRK603 (*see* **Note 7**). The production of TEV protease is constitutive in BL21 cells because no Tet repressor is present. However, TEV protease is not produced in DH5αPRO cells until the resident Tet repressor is displaced from the synthetic P_L/tetO promoter/operator by the addition of anhydrotetracycline ***(10)***. Both fusion proteins were processed essentially to completion in BL21 cells. Whereas all of the LcrH was soluble after cleavage, the YopN protein was almost completely insoluble. In DH5αPRO cells, the production of TEV protease was induced 2 h after induction of the fusion proteins with IPTG. Under these circumstances, virtually all of the free YopN protein became soluble. It should be noted, however, that sometimes intracellular processing is less efficient when the induction of TEV protease is delayed for 2 h, as is clearly the case with the MBP-YopN fusion protein.

passenger protein while it is still attached to MBP. In our lab, we have developed a simple method to rapidly ascertain whether a fusion protein will yield a soluble product after cleavage ***(10)***. For this purpose, we use another plasmid vector (pRK603; *see* **Note 7**) to coexpress TEV protease along with the fusion protein substrate. First, IPTG is added to the log phase culture and the fusion protein is allowed to accumulate for a period of time. Then, we stimulate the production of TEV protease by adding anhydrotetracycline to the culture. This protocol must be performed in a strain of *E. coli* that produces the Tet repressor (e.g., DH5αPRO or BL21PRO cells from Clontech); otherwise, the expression of TEV protease will be constitutive. The cells are harvested after the protease has had time to digest the fusion protein, and then samples of the total and soluble protein are prepared and analyzed by SDS-PAGE (*see* **Subheading 3.4.**). If the passenger protein is soluble after intracellular processing, then it is also likely to be soluble after the fusion protein has been purified and processed in vitro. Examples of how this method can be used are illustrated in **Fig. 4**.

3.5.1. Intracellular Processing of MBP Fusion Proteins by TEV Protease

Transform competent DH5αPRO or BL21PRO cells that already contain pRK603 and pKC1 with the MBP fusion protein expression vector (*see* **Note 3**) and spread them on an LB agar plate containing ampicillin (100 μg/mL), chloramphenicol (30 μg/mL), and kanamycin (30 μg/mL). Incubate the plate overnight at 37°C. Inoculate 2–5 mL of LB medium containing ampicillin (100 μg/mL), chloramphenicol (30 μg/mL), and kanamycin (30 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. *See* **Subheadings 2.1.9.**, **2.2.6.**, and **2.3.2.** for the preparation of LB medium and antibiotic stock solutions. The next morning, inoculate 25 mL of the same medium in a 250-mL baffled-bottom flask with 0.25 mL of the saturated overnight culture. Label this flask "+". Also prepare a duplicate culture and label it "–". Grow at 37°C with shaking to mid-log phase (OD_{600nm} ~0.5), and then add IPTG to the "+" flask (1 m*M* final concentration). After 2 h, add anhydrotetracycline to both flasks (100 ng/mL final concentration), and adjust the shaker temperature to 30°C (the optimum temperature for TEV protease cleavage). After 2 more hours, pellet the cells by centrifugation, prepare T–, T+ and S+ samples for SDS-PAGE, and analyze results as described in **Subheadings 3.4.4.** and **3.4.5.** It is advisable also to include a total protein sample from cells producing the same fusion protein in the absence of TEV protease (i.e., the T+ sample prepared in **Subheading 3.4.3.**) on the gel to facilitate interpretation of the results. Examine the gel to determine approximately what fraction of the fusion protein was cleaved and what fraction of the cleaved passenger protein was soluble.

3.5.3. 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 an MBP fusion protein. Exactly how and why this occurs is unclear, but we suspect that fusion to MBP somehow enables certain proteins to 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. For this reason, we strongly recommend that a biological assay be employed (if available) at an early stage to confirm that the passenger protein is in its native conformation.

4. Notes

1. We recommend a proofreading polymerase such as *Pfu* Turbo (Stratagene, La Jolla, CA, USA) or Deep Vent (New England Biolabs, Beverly, MA, USA) to minimize the occurrence of mutations during PCR. This is especially important when attempting to ligate a blunt-ended PCR fragment with a vector fragment produced by digestion with a restriction endonuclease that generates blunt ends, because thermostable polymerases without proofreading activity (e.g., *Taq* polymerase) will add an extra unpaired adenosine residue to the 3' end of the DNA.
2. We typically purify fragments by horizontal electrophoresis in 1% agarose gels run in TAE buffer (40 m*M* Tris-acetate, 1 m*M* EDTA, pH 8). It is advisable to use agarose of the highest possible purity (e.g., Seakem-GTG from FMC BioPolymer, Philadelphia, PA, USA). Equipment for horizontal 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, Valencia, CA, USA) in accordance with the instructions supplied with the product.
3. While any method for the preparation of competent cells can be used (e.g., $CaCl_2$), we prefer electroporation because of the high transformation efficiency that can be achieved. Electrocompetent cells can be purchased from various sources (e.g., Stratagene, Invitrogen, Clontech, Bio-Rad, Novagen). In addition, 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 10% glycerol, then pelleted again. After several such washes with 10% glycerol, 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, a ligation reaction, or a Gateway™ reaction). The mixture is placed into an ice-cold electroporation

cuvet and electroporated according to the manufacturers recommendations (e.g., a 1.8 kV pulse in a cuvet with a 1-mm gap). 1 mL of SOC medium ***(8)*** is immediately added to the cells, and they are allowed to grow at 37°C with shaking for 1 h. 5–200 µL of the cells are then spread on an LB agar plate containing the appropriate antibiotic(s).

4. We prefer the Wizard® miniprep kit (Promega, Madison, WI, USA) or the QIAprep™ Spin miniprep kit (QIAGEN, Valencia, CA, USA), but similar kits can be obtained from a wide variety of vendors.
5. To circumvent the problem of codon bias in *E. coli*, we routinely express proteins in BL21-CodonPlus™-RIL cells (Stratagene). The RIL plasmid 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. These codons are rarely used in *E. coli*, but occur frequently in ORFs from other organisms. Consequently, the yield of some MBP fusion proteins will be significantly greater in cells that harbor RIL, particularly if the target ORF contains tandem runs of rare codons ***(11)***. When this is not the case, RIL can be omitted. RIL is selected for by resistance to chloramphenicol (30 µg/mL). In addition to the tRNA genes for AGG/AGA, AUA, and CUA codons, the accessory plasmid in the recently introduced Rosetta™ host strain (Novagen, Madison, WI, USA) also includes tRNAs for the rarely used CCC and GGA codons. Hence, the Rosetta™ strain may turn out to be even more useful than BL21-CodonPlus™-RIL cells. Like RIL, the Rosetta™ plasmid is a chloramphenicol-resistant derivative of the p15A replicon. For intracellular processing experiments with TEV protease (*see* **Subheading 3.5.1.**), we use pKC1, a derivative of the low copy number plasmid pSC101 that is compatible with the p15A-derived TEV protease expression vector pRK603. pKC1 carries only the *ileX* and *argU* genes and is also selected for with chloramphenicol.
6. We find it convenient to use precast SDS-PAGE gels (e.g., 1.0 mm × 10 well, 10–20% gradient), running buffer, and electrophoresis supplies from Novex, a subsidiary of Invitrogen (Carlsbad, CA, USA).
7. pRK603 is a derivative of the p15A replicon that produces TEV protease when induced by anhydrotetracycline ***(10)***. pRK603 is selected for by its resistance to kanamycin.
8. These older vectors are essentially the same as pMAL-c2X and pMAL-p2X, respectively. However, in the new vectors, an *Nco*I site has been removed from within the MBP coding sequence and an *Nde*I site has been placed immediately at the start of the MBP open reading frame. Also, an *Nde*I site has been removed from another location that existed in the older vectors.
9. Purification of MBP fusion proteins from the periplasm does not rupture the inner membrane and release the contents of the cytosol ***(7)***. Consequently, periplasmic expression often results in higher initial purity of the fusion protein prior to amylose affinity chromatography.

10. When using pMAL-c(or p)2E, the *Kpn*I-digested vector must be filled in to yield a blunt end before ligation with a blunt-ended PCR fragment. Cloning in pMAL-c(or p)2G is like the X vectors, except that *Sna*BI is utilized as the blunt site in the vector (*see* **ref. *7***).
11. Alternatively, one of the other restriction sites in the pMAL polylinker can be used to join the N-terminus of the passenger protein to the C-terminus of MBP (e.g., *Eco*RI), but this would have the effect of adding extra nonnative residues to the linker between MBP and the passenger protein. It is possible that an increase in the length of the linker would affect MBP's ability to promote the solubility of the passenger protein. Moreover, if one intends to exploit a protease cleavage site that is already contained in the linker, the additional residues would end up on the N-terminus of the passenger protein after digestion.
12. If a particular ORF happens to contain a *Bam*HI restriction site, then any of the other sites in the pMAL polylinker may be used instead (*e.g.*, *Eco*RI, *Xba*I, *Sal*I, *Pst*I, or *Hin*dIII).
13. *See* **ref. *8)*** for tips on setting up ligation reactions. A typical reaction contains ~300–400 ng of DNA. The two fragments should be present at approximately equimolar concentrations. The two DNA fragments, 2 µL of 10X ligase buffer, ATP (1 m*M* final concentration), 1 µL of T4 DNA ligase, and H_2O are combined in a total volume of 20 µL. The reaction is incubated at room temperature for several hours or at 16°C overnight.
14. 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 reaction. All primers are used at the typical concentrations for PCR in the two-step protocol.
15. The PCR reaction can be modified in numerous ways to optimize results, depending on the nature of the template and primers. *See* **ref. *8*** (Vol. 2, Chapter 8) for more information.
16. 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.
17. 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 LxR and BxP 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 BxP reaction, as described in the instruction manual.
18. 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.

19. At this point, we remove a 5 µL aliquot from the reaction and add it to 0.5 µL of 10X stop 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 (25 µg/mL). From the number of colonies obtained, it is possible to estimate the percent conversion of the PCR product to entry clone in the BxP reaction. Additionally, entry clones can be recovered from these colonies in the event that no transformants are obtained after the subsequent LxR reaction.
20. If very few or no ampicillin-resistant transformants are obtained after the LxR reaction, the efficiency of the process can be improved by incubating the BxP reaction overnight.
21. We routinely break cells 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.

Acknowledgments

We wish to thank Rachel Kapust and Karen Routzahn for their valuable contributions to the development of these methods and Invitrogen/Life Technologies for granting us early access to the Gateway™ cloning technology.

References

1. Lilie, H., Schwarz, E., and Rudolph, R. (1998) Advances in refolding of proteins produced in *E. coli*. *Curr. Opin. Biotechnol.* **9,** 497–501.
2. Baneyx, F. (1999) In vivo folding of recombinant proteins in *Escherichia coli*, in *Manual of Industrial Microbiology and Biotechnology* (Davies, J. E., Demain, A. L., Cohen, G., et al., eds.), American Society for Microbiology, Washington, D. C., pp. 551–565.
3. Kapust, R. B. and Waugh, D. S. (1999) *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci.* **8,** 1668–1674.
4. Fox, J. D., Kapust, R. B., and Waugh, D. S. (2001) Single amino acid substitutions on the surface of *Escherichia coli* maltose-binding protein can have a profound impact on the solubility of fusion proteins. *Protein Sci.* **10,** 622–630.
5. Richarme, G. and Caldas, T. D. (1997) Chaperone properties of the bacterial periplasmic substrate-binding proteins. *J. Biol. Chem.* **272,** 15,607–15,612.
6. Sachdev, D. and Chirgwin, J. M. (1998) Solubility of proteins isolated from inclusion bodies is enhanced by fusion to maltose-binding protein or thioredoxin. *Protein Expr. Purif.* **12,** 122–132.
7. Riggs, P. (2000) Expression and purification of recombinant proteins by fusion to maltose-binding protein. *Mol. Biotechnol.* **15,** 51–63.
8. Sambrook, J. and Russell, D. W. (2001) *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

9. Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W. (1990) Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* **185,** 60–89.
10. Kapust, R. B. and Waugh, D. S. (2000) Controlled intracellular processing of fusion proteins by TEV protease. *Protein Expr. Purif.* **19,** 312–318.
11. Kane, J. F. (1995) Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coli*. *Curr. Opin. Biotechnol.* **6,** 494–500.
12. Cornelis, G. R., Boland, A., Boyd, A. P., et al. (1998) The virulence plasmid of *Yersinia*, an antihost genome. *Microbiol. Mol. Biol. Rev.* **62,** 1315–1352.