Gateway vectors for the production of combinatorially-tagged His_6-MBP fusion proteins in the cytoplasm and periplasm of *Escherichia coli*

SREEDEVI NALLAMSETTY, BRIAN P. AUSTIN,[1] KERRI J. PENROSE,[1] AND DAVID S. WAUGH

Macromolecular Crystallography Laboratory, Center for Cancer Research, National Cancer Institute at Frederick, Frederick, Maryland 21702, USA

(RECEIVED July 21, 2005; FINAL REVISION September 28, 2005; ACCEPTED September 28, 2005)

Abstract

Many proteins that accumulate in the form of insoluble aggregates when they are overproduced in *Escherichia coli* can be rendered soluble by fusing them to *E. coli* maltose binding protein (MBP), and this will often enable them to fold in to their biologically active conformations. Yet, although it is an excellent solubility enhancer, MBP is not a particularly good affinity tag for protein purification. To compensate for this shortcoming, we have engineered and successfully tested Gateway destination vectors for the production of dual His_6MBP-tagged fusion proteins in the cytoplasm and periplasm of *E. coli*. The MBP moiety improves the yield and solubility of its fusion partners while the hexahistidine tag (His-tag) serves to facilitate their purification. The availability of a vector that targets His_6MBP fusion proteins to the periplasm expands the utility of this dual tagging approach to include proteins that contain disulfide bonds or are toxic in the bacterial cytoplasm.

Keywords: immobilized metal affinity chromatography; IMAC; maltose-binding protein; Gateway cloning; fusion tag; structural genomics; TEV protease; expression vector; fusion protein

A major obstacle 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 (Stevens 2000; Chayen 2004). One means of circumventing this problem that shows considerable promise is to exploit the innate ability of certain proteins to enhance the solubility of their fusion partners. *Escherichia coli* maltose binding protein (MBP) has emerged as an attractive vehicle for the production of recombinant proteins in part because of its ability to enhance the solubility and promote the proper folding of its fusion partners (Kapust and Waugh 1999; Fox et al. 2003). However, MBP fusion proteins do not always bind efficiently to amylose resin (Pryor and Leiting 1997; Routzahn and Waugh 2002); and even when they do, amylose affinity chromatography typically does not produce samples of sufficient purity for structural studies or enzymology.

In an effort to overcome the limitations of amylose affinity chromatography, Routzahn and Waugh incorporated auxiliary affinity tags in various locations within the framework of an MBP fusion protein and analyzed their impact on the ability of MBP to promote the solubility of its fusion partners (Routzahn and Waugh 2002). They found that a biotin acceptor peptide (BAP), but not a polyarginine tag, could be added to the N terminus of MBP without impeding its solubility-enhancing activity. The hexahistidine tag (His-tag), which is widely viewed as the best affinity tag for high-

[1]These authors contributed equally to this work.

Reprint requests to: David S. Waugh, National Cancer Institute at Frederick, P.O. Box B, Frederick, MD 21702, USA; e-mail: waughd@ncifcrf.gov; fax: (301) 846-7148.

Article and publication are at http://www.proteinscience.org/cgi/doi/10.1110/ps.051718605.

throughput protein purification, was also examined as an auxiliary tag in this study, but only in locations where it would remain attached to the passenger protein after endoproteolytic removal of the MBP moiety and could therefore interfere with the structure or activity of the passenger.

In the present study, we show that a His-tag can also be added to the N terminus of MBP without impeding its ability to promote the solubility of its fusion partners. Moreover, we demonstrate that this dual His_6MBP tag forms the basis of an entirely generic method for protein expression and purification in which the MBP moiety serves to improve the yield and solubility of the passenger protein (when required) while the His-tag facilitates its purification to homogeneity, yielding an end product that differs very little or not at all from the sequence of the natural protein. Finally, we show that His_6MBP-tagged proteins can also be exported to the periplasm in *E. coli*, thereby extending the utility of this approach to include target proteins with disulfide bonds.

Results

The addition of a His-tag to the N terminus of MBP does not impede its ability to enhance the solubility of its fusion partners

Although MBP is an excellent solubility enhancer and can also improve the yield of its fusion partners (Kapust and Waugh 1999; Hammarström et al. 2002; Shih et al. 2002; Fox et al. 2003), it is not a particularly good affinity tag for protein purification (Pryor and Leiting 1997; Routzahn and Waugh 2002). On the other hand, the His-tag is very well suited for high-throughput protein purification but does not have an appreciable impact on the yield or solubility of its fusion partners (Waugh 2005). We reasoned that the collective advantages of these two tags could be realized by combining them together.

A previous study demonstrated that when auxiliary affinity tags are incorporated within the context of an MBP fusion protein, both the type of tag and its location can influence the ability of MBP to promote the solubility of its fusion partners (Routzahn and Waugh 2002). Therefore, we first sought to determine what effect the addition of an N-terminal His-tag to MBP would have on its ability to act as a solubility-enhancing fusion partner. To this end, we fused three different passenger proteins to otherwise identical MBPs with or without an N-terminal His-tag. All three of these passenger proteins (human $p16^{INK4}$, *Aquorea victoria* green fluorescent protein, and human papilloma virus E6) are poorly soluble when expressed in an unfused form in *E. coli* (Kapust and Waugh 1999; Fox et al. 2003).

The N-terminally His-tagged MBP fusion proteins were expressed in *E. coli* and their yield and solubility were compared with that of the corresponding MBP fusion proteins with no His-tags under identical conditions (Fig. 1). The results clearly demonstrate that the addition of a His-tag to the N terminus of MBP affects neither the yield nor the solubility of MBP fusion proteins.

A generic method for protein expression and purification using a dual His_6MBP tag

Having established that the dual His_6MBP affinity tag retains the chief advantages of MBP (i.e., its beneficial impact on yield and solubility), we next sought to devise a generic protocol for protein purification that would exploit the advantages of the His-tag. The method we developed consists of three principal steps. First, the His-MBP fusion protein is purified by immobilized metal affinity chromatography (IMAC) on Ni-NTA resin, by virtue of its His-tag. Next, the fusion protein is digested with His-tagged tobacco etch virus (TEV) protease (Kapust et al. 2001) to separate the dual His_6MBP tag from the passenger protein. Finally, the His-tagged TEV protease, the cleaved His_6MBP moiety, and any residual undigested fusion protein are separated from the (now untagged) target protein by a second round of IMAC. Any endogenous proteins that bind nonspecifically to the Ni-NTA resin in the first round of IMAC are also captured in the second round and are thereby separated from the target protein.

To demonstrate the efficacy of this method, we used it to purify IglC, an essential virulence factor from the bacterial pathogen *Francisella tularensis*, from the cyto-

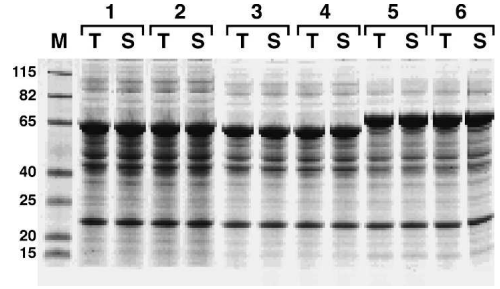

Figure 1. The presence of an N-terminal His-tag on MBP (maltose binding protein) does not reduce the yield or solubility of MBP fusion proteins. MBP (odd numbered lanes) or His_6MBP (even numbered lanes) was fused to human papilloma virus E6 (lanes *1*,*2*), human $p16^{INK4}$ (lanes *3*,*4*), and GFP (lanes *5*,*6*), and the resulting fusion proteins were expressed in *E. coli* BL21(DE3) CodonPlus-RIL cells. Samples of the total (T) and soluble (S) intracellular protein were collected from each culture and analyzed by SDS-PAGE. M, molecular weight standards (kDa).

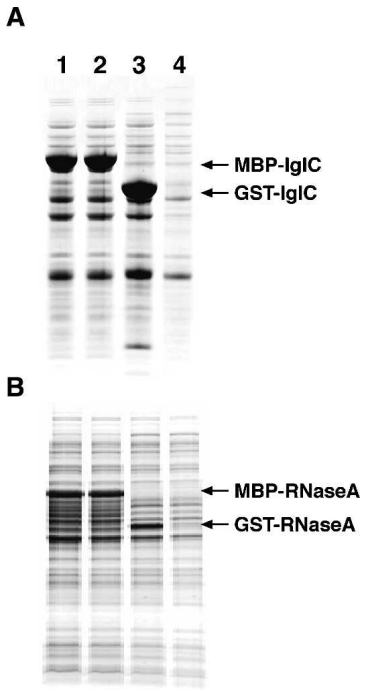

Figure 2. MBP enhances the solubility of IglC and RNaseA. Samples of the total (odd numbered lanes) and soluble (even numbered lanes) intracellular proteins were analyzed by SDS-PAGE. (*A*) His_6MBP–IglC (lanes *1*,*2*) and GST–IglC (lanes *3*,*4*) fusion proteins. (*B*) His_6MBP–RNaseA (lanes *1*,*2*) and GST–RNaseA (lanes *3*,*4*) fusion proteins.

sol of *E. coli*. IglC is poorly soluble when it is produced as a GST fusion protein in *E. coli*, but the His_6MBP–IglC fusion protein is highly soluble (Fig. 2A). Because GST has virtually no ability to promote the solubility of its fusion partners (Kapust and Waugh 1999; Fox et al. 2003), the solubility of a GST fusion protein is a good indicator of the solubility of a passenger protein in its unfused state. Hence, we conclude that a dramatic increase in the solubility of IglC was achieved by fusing it to MBP.

The purity of the protein at various stages of the generic purification process was monitored by SDS-PAGE (Fig. 3A). After the final IMAC step, the IglC was virtually free of detectable impurities (lane 6). Following buffer exchange by gel filtration (lane 8), this preparation of IglC yielded crystals that diffracted X rays to a resolution of 1.8 Å (Fig. 3B). The catalytic domain of *E. coli* Lon protease also yielded diffracting crystals when purified in this manner (Tropea et al. 2005). The observations that IglC remained soluble after it was separated from dual His_6MBP tag and subsequently yielded crystals strongly suggest that it was able to fold into its native conformation after being rendered soluble by fusion to MBP.

Expression, purification, and characterization of His_6MBP–RNaseA fusion proteins

Some proteins are not amenable to overproduction in the cytoplasm of *E. coli*, either because they are toxic in that compartment or because they contain disulfide bonds. In such cases, export to the periplasm offers a possible solution. For proteins with disulfide bonds, the periplasm has the advantage of providing a more favorable redox environment along with a variety of protein disulfide isomerases. Because MBP naturally resides in the periplasm, it can be used as a carrier protein to facilitate the export of its fusion partners (e.g., Giuliani et al. 2001; Huys et al. 2002; Planson et al. 2003; Muramatsu et al. 2005). Seeking to expand the utility of the dual His_6MBP tagging strategy to encompass a wider variety of passenger proteins, we designed and constructed a Gateway destination vector for the production of His_6MBP fusion proteins in the periplasm of *E. coli*, using the natural N-terminal signal peptide of MBP.

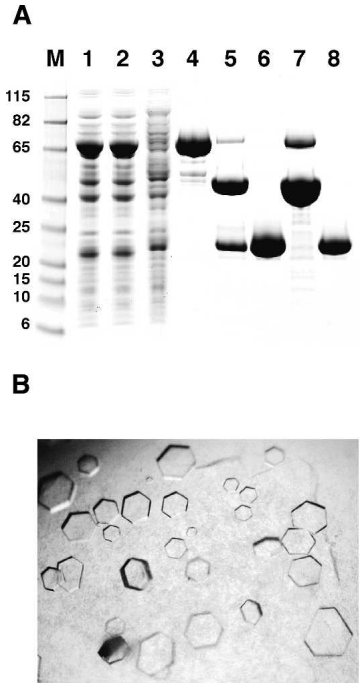

Figure 3. Purification of IglC using a generic, IMAC-based protocol in conjunction with the His_6MBP tag. (*A*) SDS-PAGE of samples at various stages of purification. (Lane *M*) Molecular weight standards (kDa); (lane *1*) total intracellular protein; (lane *2*) soluble intracellular protein; (lane *3*) flow-through (unbound) fraction from first IMAC step; (lane *4*) protein eluted with imidazole from Ni-NTA column during first IMAC step; (lane *5*) products of the TEV protease digest; (lane *6*) flow-through fraction from second IMAC step; (lane *7*) proteins that bound to the Ni-NTA column during the second IMAC step; (lane *8*) IglC after gel filtration. (*B*) Crystals of IglC obtained with Crystal Screen 2 (Hampton Research) condition #26.

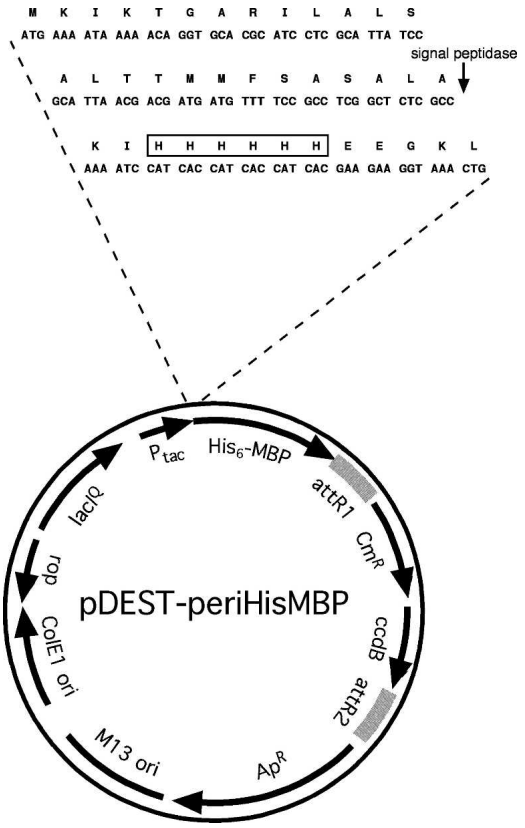

Figure 4. Schematic representation of the periplasmic His_6MBP fusion protein expression vector pDEST-periHisMBP. The nucleotide sequence encoding the signal peptide, the polyhistidine tag, and the first few N-terminal residues of mature MBP, along with the corresponding amino acid sequences in single letter code, are shown *above* the vector schematic.

A schematic diagram of the periplasmic His_6MBP fusion vector, named pDEST-periHisMBP to distinguish it from its cytosolic counterpart pDEST-HisMBP, is shown in Figure 4. pDEST-periHisMBP is identical to pDEST-HisMBP except that codons 2–26 of the pre-MBP open reading frame (shown above the circular map) are absent in the latter plasmid. The His-tag was inserted between codons two and three of the mature MBP polypeptide, two residues after the site where signal peptidase cleaves the leader sequence from wild-type MBP. Because the sequence context in the immediate vicinity of the signal peptidase cleavage site is retained in pDEST-periHisMBP, it was hoped that the removal of the signal peptide would not be affected by the nearby polyhistidine tract.

To ascertain whether pDEST-periHisMBP functions as intended, we used bovine RNaseA as a C-terminal fusion partner. This protein has four intertwined disulfide bonds that are required for enzymatic activity (Klink et al. 2000). When the cytosolic expression vector pDEST-HisMBP was used to produce the dual His_6 MBP-tagged RNaseA fusion protein, it accumulated predominantly in a soluble form (Fig. 2B). In contrast, a GST–RNaseA fusion protein was almost totally insoluble under the same conditions. This indicates that the solubility of RNaseA was enhanced by fusing it to MBP. Even so, after partial purification by IMAC (Fig. 5A), this form of soluble His_6MBP–RNaseA fusion protein exhibited no enzymatic activity (Fig. 6).

When the periplasmic expression vector pDEST-peri-HisMBP was used to produce the dual His_6MBP-tagged RNaseA protein, it was possible to purify a polypeptide of the expected size from osmotic shock fluid by IMAC (Fig. 5B). The N-terminal amino acid sequence of this protein was determined by Edman degradation to be Lys-Ile-His-His-His-His-His-His-Glu-Glu-Gly-Lys, confirming the identity of the fusion protein and demonstrating that the signal peptide was cleaved at the desired site by signal peptidase despite the presence of a nearby polyhistidine

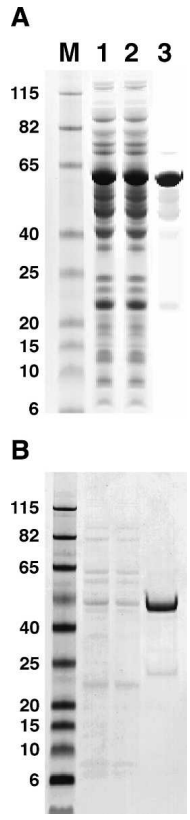

Figure 5. Expression and purification of His_6MBP–RNaseA fusion proteins. (*A*) Cytosolic His_6MBP–RNaseA fusion protein. (Lane *M*) Molecular weight standards (kDa); (lane *1*) total intracellular protein; (lane *2*) soluble intracellular protein; (lane *3*) protein eluted from the Ni-NTA column with imidazole. (*B*) Periplasmic His_6MBP–RNaseA fusion protein. (Lane *M*) Molecular weight standards; (lane *1*) osmotic shock fluid; (lane *2*) flow-through fraction from the Ni-NTA column; (lane *3*) protein eluted from the Ni-NTA column with imidazole.

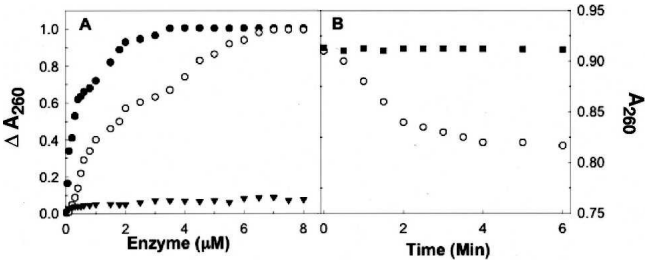

Figure 6. Enzymatic activity of His_6MBP–RNaseA fusion proteins. (○) Periplasmic His_6MBP–RNaseA; (▼) cytosolic His_6MBP–RNaseA; (●) unfused RNaseA standard; (■) no enzyme added. (*A*) The reactions were initiated by adding the indicated amount of enzyme (His_6MBP–RNaseA or unfused RNaseA) to a fixed amount of yeast RNA and recording the change in absorbance at 260 nm after 15 min. (*B*) Time course of absorbance change of RNA at 260 nm. Enzyme (1.5 μg) was added to 1 mL of RNA solution (0.8 mg/mL) at Time 0.

tract. Unlike the His_6MBP–RNaseA fusion protein that was produced in the cytoplasm of *E. coli*, the corresponding fusion protein isolated from the periplasm exhibited ribonuclease activity that was about two-thirds that, on a molar basis, of an RNaseA "standard" that was purchased from a commercial vendor (Fig. 6). Hence, we conclude that at least two-thirds of the periplasmic His_6MBP–RNaseA fusion protein possessed the correct configuration of disulfide bonds within the RNaseA domain. If the N-terminal His_6MBP-tag or C-terminal His-tag impeded the activity of the enzyme to any degree, then the fraction of properly folded RNaseA may have been even greater than this.

Discussion

Among affinity tags that have been characterized to date, only MBP is a potent solubility enhancer (Waugh 2005). In many cases, polypeptides that normally accumulate as inclusion bodies in *E. coli* can be recovered in a properly folded, biologically active form if they are fused to C terminus of MBP (e.g., Pryor and Leiting 1997; Kapust and Waugh 1999; Bach et al. 2001; Nomine et al. 2001; Fox et al. 2003; Goh et al. 2003; Planson et al. 2003). Because insolubility of recombinant proteins appears to be the major obstacle to high-throughput protein expression and purification (Stevens 2000; Chayen 2004), the chaperone-like quality of MBP makes it a particularly valuable fusion partner. In addition, MBP fusion proteins can usually be expressed at high levels in the cytoplasm, and MBP may afford some protection from intracellular proteolysis (Waugh 2005).

Yet, although it has some impressive attributes, MBP does not perform especially well in practice as an affinity tag for protein purification. Some fusion proteins cannot be efficiently absorbed onto amylose resin. Moreover, more contaminants seem to bind to amylose resin than many other affinity matrices, and the resin is prone to degradation by α-amylases. One way to circumvent the problems associated with amylose affinity chromatography while still retaining the other benefits of MBP would be to incorporate one or more additional affinity tags within the framework of an MBP fusion protein and utilize them for affinity purification.

The most widely used affinity tag is the His-tag, which binds to immobilized metal ions. Although it does not exhibit the highest level of specificity, its small size and the robust nature of the resin have contributed to its popularity. Seeking to combine the benefits of both tags, we designed and characterized vectors for the production of dual His_6MBP-tagged proteins in both the cytoplasm and periplasm of *E. coli*.

Using a model fusion protein composed of His_6MBP fused to the N terminus of the IglC virulence factor from *Francisella tularensis*, we demonstrated that IglC could be purified to homogeneity by employing two successive IMAC steps with a TEV protease digest in between them to separate the His_6MBP moiety from the target protein. The application of two successive IMAC steps, instead of just one, is a key feature of this generic protocol for protein purification because it overcomes the principal weakness of IMAC, which is the tendency for endogenous histidine-rich proteins to bind to the Ni-NTA matrix. By employing a second IMAC step after removal of the His_6MBP tag, any endogenous proteins that bind nonspecifically in the first round of IMAC will also do so during the second round and can thereby be separated from the target protein.

Although much higher yields of recombinant proteins are usually obtained in the cytoplasm than in the periplasm (Riggs 2000), the reducing environment and/or chaperone constituents of the cytosol may not be appropriate for the production of all recombinant proteins, particularly those with disulfide bonds. In this regard, another significant but seldom utilized advantage of MBP is that it normally resides in the periplasm of *E. coli* (the DNA sequence encoding the natural signal peptide is deleted for cytoplasmic expression). Hence, MBP can be used as a carrier protein to target proteins to the periplasm. In the present study, we have shown that a His-tag can be incorporated into the secreted form of MBP, just downstream of the signal peptidase processing site, and that such a His_6MBP–RNaseA fusion protein is secreted to the periplasm and processed correctly to generate an N-terminally His-tagged fusion protein. Moreover, we demonstrated that the His_6MBP–RNaseA fusion protein can be purified from osmotic shock fluid by IMAC and that the RNaseA moiety of the fusion protein, which requires four disulfide bonds for enzymatic activity, is catalytically active. Hence, the generic methodology

originally developed for the purification of cytosolic fusion proteins can now be extended to target proteins that are more appropriate for the environment of the periplasm. The compatibility of these His_6MBP fusion vectors with Gateway recombinational cloning methods will further enhance their utility in a high-throughput setting.

Of course no single approach can be expected to solve all protein expression problems. Not all aggregation-prone proteins can be rendered soluble by fusing them to MBP. Moreover, a significant fraction of passenger proteins precipitate after they are separated from MBP by endoproteolytic cleavage of the fusion protein, and the proteolysis step itself is not always very efficient. Still, it seems clear that an aggregation-prone protein stands a better chance of being produced in a soluble, properly folded form if it is fused to MBP than if it is produced in an unfused form or fused to a His-tag alone. For this reason, the cytoplasmic and periplasmic His_6MBP fusion vectors described herein may prove to be one of the most effective means of producing soluble, active recombinant proteins in *E. coli*.

Materials and methods

Gateway destination vectors

Gateway recombinational cloning (Invitrogen) was used to facilitate the construction of fusion protein expression vectors. The construction of a Gateway destination vector for the production of MBP fusion proteins (pKM596) was described previously (Fox and Waugh 2003). The cytoplasmic His_6MBP destination vector (pDEST-HisMBP) was constructed by inserting an in frame hexahistidine sequence between codons three and four of the open reading frame (ORF) encoding MBP in pKM596, as detailed elsewhere (Tropea et al. 2005).

The periplasmic His-MBP expression vector (pDEST-periHisMBP) was assembled as follows. In one PCR, the nucleotide sequence encoding the N-terminal signal peptide of MBP and a portion of the plasmid backbone extending into the *lacI* gene was amplified from the plasmid expression vector pMalP2 (Riggs 2000) with the primers PE-1436 (5′-CTTCGTG ATGGTGATGGTGATGGATTTTGGCGAGAGCCGAG GCGGAAAAC-3′) and PE-1437 (5′-TCTCCCATGAAGAC GGTACGCGACTG-3′). This resulted in the addition of a sequence encoding a hexahistidine tag two codons after the natural signal peptidase cleavage site. In a second PCR, a portion of the ORF encoding MBP and its N-terminal hexahistidine tag was amplified from pDEST-HisMBP with the primers PE-1435 (5′-GTTTTCCGCCTCGGCTCTCGCCAAAATCCAT CACCATCACCATCACGAAG-3′) and PE-1438 (5′-CTCTT ACCTTTCGCTTTCAGTTC-3′). Next, the amplicons from these two PCRs were combined and used as the template for a third PCR with primers PE-1437 and PE-1438. The final PCR amplicon was digested with BstEII and BglII, and then inserted between the BstEII and BglII sites in pDEST-HisMBP to generate pDEST-periHisMBP. The nucleotide sequence of the insert was confirmed experimentally.

Gateway entry clones

The ORF encoding IglC was amplified by PCR, using the following oligodeoxyribonucleotide primers: 5′-GAGAACCTG TACTTCCAGATGATTATGAGTGAGATGATAACAAG-3′ (PE-1656) and 5′-GGGGACCACTTTGTACAAGAAAGCTG GGATTATGCAGCTGCAATATATCCTATTTTA G-3′ (PE-1657). The template for this PCR was genomic DNA from *F. tularensis* obtained from the United States Military Institute of Infectious Diseases (USAMRIID). Next, the resulting PCR amplicon was used as the template for another PCR with primers PE-277 (Evdokimov et al. 2002) and PE-1657. The former primer is designed to anneal to the sequence encoding the TEV protease recognition site and add an attB1 recombination site to the end of the amplicon. The final PCR amplicon was inserted into pDONR201 (Invitrogen) by recombinational cloning to generate the entry clone pSN1751. The nucleotide sequence of the ORF encoding IglC was verified experimentally.

The ORF encoding RNaseA was amplified by PCR, using the following deoxyribonucleotide primers: 5′-GAGAACCTGTAC TTCCAGGGTAAAGAGACAGCAGCCGCAAAGTTTG-3′ (PE-1384) and 5′-ATTAGTGATGATGGTGGTGATGAAC ACTGGCGTCAAAGTGGACAGGAAC-3′ (PE-1385). The template for this PCR reaction was a cDNA clone (pRR16) of bovine RNaseA obtained from Dr. Ronald Raines (DelCardayré et al. 1995). Next, this PCR amplicon was used as the template for another PCR with primers PE-277 and PE-278 (Evdokimov et al. 2002), which are designed to anneal to the sequences encoding the TEV protease recognition site and the His-tag, respectively, and add attB1 and attB2 recombination sites to the ends of the amplicon. The final PCR amplicon was inserted into pDONR201 (Invitrogen) by recombinational cloning to generate the entry clone pSN1430. The nucleotide sequence of the ORF encoding RNaseA was verified experimentally.

Fusion protein expression vectors

The His_6MBP–IglC (pKP1690) and GST–IglC (pSN1752) fusion protein expression vectors were constructed by recombining the IglC ORF from pKP1690 into the destination vectors pDEST-HisMBP and pDEST-15 (Invitrogen), respectively. Cytosolic (pSN1432) and periplasmic (pSN1554) His_6MBP–RNaseA fusion protein expression vectors were constructed by recombining the RNaseA ORF from pSN1430 into the destination vectors pDEST-HisMBP and pDEST-periHisMBP, respectively. The His_6MBP–RNaseA fusion proteins have an identical interdomain linker sequence consisting of the translation product of the attB1 recombination site followed by a TEV protease recognition site (ITSLYKKAGSENLYFQG). The GST–RNaseA fusion protein expression vector (pSN1431) was constructed by recombining the RNaseA ORF from pSN1430 into the destination vector pDEST-15. MBP-p16, MBP-E6, and MBP-GFP fusion protein expression vectors were constructed by recombining the passenger protein ORFs from pre-existing entry clones (Fox et al. 2003) into pKM596. The corresponding His_6MBP fusion protein expression vectors were constructed by recombining the ORFs from these entry clones into pDEST-HisMBP (Fox and Waugh 2003).

Yield and solubility of fusion proteins

The yield and solubility of various GST, MBP, and His_6MBP fusion proteins were assessed by standard procedures, using *E. coli* BL21(DE3) CodonPlus-RIL cells (Stratagene), as de-

scribed (Kapust and Waugh 1999; Nallamsetty et al. 2004). All cultures were grown at 37°C. Samples were analyzed on 10%–20% SDS-polyacrylamide gels (Novex) and visualized by staining with GelCode Blue (Pierce).

Expression and purification of IglC

The His_6-MBP–IglC fusion protein was overproduced in *E. coli* BL21(DE3) CodonPlus-RIL cells (Stratagene). Single antibiotic-resistant colonies of cells transformed with pKP1690 were used to inoculate 100 mL of Luria broth supplemented with 100 μg/mL ampicillin, 30 μg/mL chloramphenicol and 0.2% (D+)-glucose monohydrate (Sigma-Aldrich). These cultures were grown by shaking (250 rev/min) to saturation overnight at 37°C and then diluted 50-fold into several liters of fresh medium. When the cells reached early log phase (OD_{600nm} = 0.3–0.5), the temperature was reduced to 30°C and isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM. Four hours later, the cells were recovered by centrifugation at 5000*g* for 10 min and stored at –80°C.

All chromatography steps were carried out at 4°C. *E. coli* cell paste was suspended in ice-cold 50 mM sodium phosphate (pH 8.0), 150 mM NaCl, 25 mM imidazole (buffer A) containing Complete EDTA-free protease-inhibitor cocktail (Roche Molecular Biochemicals). The cells were lysed with an APV Gaulin Model G1000 homogenizer at 69 MPa and centrifuged at 30,000*g* for 30 min at 4°C. The supernatant was filtered through a 0.45 μm polyethersulfone membrane and then loaded onto three tandem 5 mL HisTrap affinity columns (Amersham Biosciences) equilibrated in buffer A. The column was washed to baseline with buffer A and then eluted with a linear gradient from 25 to 250 mM imidazole in buffer A. Fractions containing recombinant His_6MBP–IglC were pooled and concentrated sixfold using an Amicon YM30 membrane (Millipore). The concentrated sample was then digested overnight with 1 mg of His-tagged TEV protease (Kapust et al. 2001) per 100 mg of fusion protein at 4°C, resulting in free IglC protein. The digest was diluted 1 : 6 with 50 mM sodium phosphate (pH 8.0), 150 mM NaCl to approximate an imidazole concentration of 25 mM and then reapplied to the HisTrap columns equilibrated with buffer A. The sample was eluted isocratically and the flow-through collected and concentrated to a volume of 5 mL. This sample was next applied to a 26/60 HiPrep Sephacryl S100 prep-grade column (Amersham Biosciences) equilibrated with 25 mM Tris-HCl (pH 8.0), 150 mM NaCl, 2 mM TCEP. The peak fractions containing IglC were pooled and concentrated to 10 mg/mL. Aliquots were flash-frozen with liquid nitrogen and stored at –80°C. The final product was judged to be >95% pure by SDS-PAGE (data not shown). The molecular weight of IglC was confirmed by electrospray mass spectrometry.

Expression and purification of cytosolic His_6MBP–RNaseA

E. coli BL21(DE3) CodonPlus-RIL cells (Stratagene) containing the His_6MBP–RNaseA fusion protein expression vector pSN1432 were grown to saturation at 37°C in Luria broth supplemented with 100 μg/mL ampicillin and 30 μg/mL chloramphenicol. The saturated culture was diluted 1 : 50 into 1 L of the same medium and grown in a shake-flask to mid-log phase (A_{600nm} ~ 0.3–0.5). At this point, IPTG was added to a final concentration of 1 mM to initiate production of the His_6MBP–RNaseA fusion protein and the culture was grown for an additional 3 h at 30°C. The cells were then recovered by centrifugation and the cell pellet was stored at –80°C.

E. coli cell paste was suspended in ice-cold 50 mM Tris-HCl (pH 7.4), 150 mM NaCl, and 25 mM imidazole (buffer B). The cells were lysed with an APV Gaulin Model G1000 homogenizer at 69 MPa and centrifuged at 30,000*g* for 30 min at 4°C. The supernatant was filtered through a 0.44 μm polyethersulfone membrane and applied to a 25 mL Ni-NTA Superflow affinity column (Qiagen) equilibrated in buffer B. The column was washed with 10 column volumes of buffer B and then eluted with a linear gradient from 25 to 250 mM imidazole in buffer B. Fractions containing recombinant His_6MBP–RNaseA were pooled and the sample was concentrated using an Amicon YM10 membrane (Millipore). Aliquots were flash-frozen with liquid nitrogen and stored at –80°C until use.

Expression and purification of periplasmic His_6MBP–RNaseA

E. coli X90 cells (Amann et al. 1983) containing the periplasmic His_6MBP–RNaseA fusion protein expression vector pSN1554 were grown to saturation in Luria broth supplemented with 100 μg/mL ampicillin at 37°C. The saturated culture was diluted 1 : 50 in the same medium and grown to mid-log phase (A_{600nm} ~ 0.4–0.5) at 37°C, at which time the temperature was shifted to 25°C and IPTG was added to a final concentration of 0.01 mM to initiate production of the fusion protein. After 4 h at 25°C, the cells were harvested by centrifugation and washed twice with 10 mM Tris-HCl (pH 7.3), 30 mM NaCl.

The osmotic shock procedure of Neu and Chou (1967) was employed to release the His_6MBP–RNaseA fusion protein from the periplasm of *E. coli*. The washed cell pellet was weighed and resuspended in 30 mM Tris-HCl (pH 7.3), 20% sucrose at 20°C, using 80 mL of Tris-sucrose solution per gram of wet cell paste. After resuspension, EDTA was added to a final concentration of 1 mM. The cells were then pelleted by centrifugation at 4°C and suspended in 40 mL of ice-cold water per gram of wet cell paste. The cell suspension was mixed gently for 10 min at 4°C, after which the cells were pelleted again by centrifugation. The supernatant, representing the osmotic shock fluid, was filtered through a 0.2 μm polyethersulfone membrane and applied to a 30 mL Ni-NTA superflow affinity column (Qiagen) equilibrated in buffer B. The column was washed with 10 column volumes of buffer B and eluted with a linear gradient from 25 to 250 mM imidazole in buffer B. Fractions containing the His_6MBP–RNaseA fusion protein were identified by SDS-PAGE, pooled, and then frozen at –80°C until use. A sample of the purified fusion protein was electrophoretically transferred to a polyvinylidene difluoride (PVDF) membrane and subjected to N-terminal amino acid sequencing.

RNaseA activity assay

A "native" bovine RNaseA standard was obtained from Sigma-Aldrich. RNaseA activity was assayed as described (Uchida and Egami 1966), with minor modifications. Varying amounts of enzyme (either a His_6MBP–RNaseA fusion protein or the RNaseA standard) were added to a reaction mixture

consisting of 1% (w/v) yeast RNA (Sigma-Aldrich) in 50 mM Tris-HCl (pH 7.5), 2 mM EDTA to achieve a final volume of 1 mL. The reactions were incubated at 37°C for 15 min and terminated by the addition of 250 μL of ice-cold 25% (v/v) perchloric acid containing 0.75% (w/v) lanthanum nitrate. The quenched reactions were chilled for 20 min on ice and then centrifuged at 15,000*g* for 10 min. One-hundred microliter aliquots of the supernatants were diluted with 2.4 mL of distilled water and then the absorbance of the samples was measured at 260 nm. One unit of RNaseA activity is defined as the amount of enzyme that caused an absorbance increase of 1.0 at 260 nm after 15 min at 37°C.

Acknowledgments

We thank Dr. Ronald Raines for his gift of the bovine RNaseA cDNA clone pRR16 and the Protein Chemistry Laboratory (SAIC-Frederick) for N-terminal amino acid sequencing. This research was supported by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research.

References

Amann, E., Brosius, J., and Ptashne, M. 1983. Vectors bearing a hybrid trp-lac promoter useful for regulated expression of cloned genes in *E. coli*. *Gene* **25:** 167–178.

Bach, H., Mazor, Y., Shaky, S., Shoham-Lev, A., Berdichevsky, Y., Gutnick, D.L., and Benhar, I. 2001. *Escherichia coli* maltose-binding protein as a molecular chaperone for recombinant intracellular cytoplasmic single-chain antibodies. *J. Mol. Biol.* **312:** 79–93.

Chayen, N.E. 2004. Turning protein crystallization from an art into a science. *Curr. Opin. Struct. Biol.* **14:** 577–583.

DelCardayré, S.B., Ribó, M., Yokel, E.M., Quirk, D.J., Rutter, W.J., and Raines, R.T. 1995. Engineering ribonuclease A: Production, purification and characterization of wild-type enzyme and mutants at Gln11. *Protein Eng.* **8:** 261–273.

Evdokimov, A.G., Tropea, J.E., Routzahn, K.M., and Waugh, D.S. 2002. Three-dimensional structure of the type III secretion chaperone SycE from *Yersinia pestis*. *Acta Crystallogr. D Biol. Crystallogr.* **58:** 398–406.

Fox, J.D. and Waugh, D.S. 2003. Maltose-binding protein as a solubility enhancer. *Methods Mol. Biol.* **205:** 99–117.

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:** 53–77.

Giuliani, C.D., Iemma, M.R., Bondioli, A.C., Souza, D.H., Ferreira, L.L., Amaral, A.C., Salvini, T.F., and Selistre-de-Araujo, H.S. 2001. Expression of an active recombinant lysine 49 phospholipase A(2) myotozin as a fusion protein in bacteria. *Toxicon* **39:** 1595–1600.

Goh, L.L., Loke, P., Singh, M., and Sim, T.S. 2003. Soluble expression of a functionally active *Plasmodium falciparum* falcipain-2 fused to maltose-binding protein in *Escherichia coli*. *Prot. Expr. Purif.* **32:** 194–201.

Hammarström, M., Hellgren, N., Van den Berg, S., Berglund, H., and Härd, T. 2002. Rapid screening for improved solubility of small human proteins produced as fusion proteins in *Escherichia coli*. *Protein Sci.* **11:** 313–321.

Huys, I., Dyason, K., Waelkens, E., Verdonck, F., Van Zyl, J., Du Plessis, J., Müller, G.J., Van der Walt, J., Clynen, E., Schoofs, L., et al. 2002. Purification, characterization and biosynthesis of parabutoxin 3, a component of *Parabuthus transvaalicus* venom. *Eur. J. Biochem.* **269:** 1854–1865.

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.

Kapust, R.B., Tözsér, J., Fox, J.D., Anderson, D.E., Cherry, S., Copeland, T.D., and Waugh, D.S. 2001. Tobacco etch virus protease: Mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency. *Protein Eng.* **14:** 993–1000.

Klink, T.A., Woycechowsky, K.J., Taylor, K.M., and Raines, R.T. 2000. Contribution of disulfide bonds to the conformational stability and catalytic activity of ribonuclease A. *Eur. J. Biochem.* **267:** 566–572.

Muramatsu, H., Yoshikawa, K., Hayashi, T., Takasu, S., Kawada, Y., Uchida, K., Sato, S., Takahashi, T., Saga, S., and Ueda, R. 2005. Production and characterization of an active single-chain variable fragment antibody recognizing CD25. *Cancer Lett.* **225:** 225–236.

Nallamsetty, S., Kapust, R.B., Tözsér, J., Cherry, S., Tropea, J.E., Copeland, T.D., and Waugh, D.S. 2004. Efficient site-specific processing of fusion proteins by tobacco vein mottling virus protease in vivo and in vitro. *Protein Expr. Purif.* **38:** 108–115.

Neu, H.C. and Chou, J. 1967. Release of surface enzymes in Enterobacteriaceae by osmotic shock. *J. Bacteriol.* **94:** 1934–1945.

Nomine, Y., Ristriani, T., Laurent, C., Lefevre, J.F., Weiss, E., and Trave, G. 2001. A strategy for optimizing the monodispersity of fusion proteins: Application to purification of recombinant HPV E6 oncoprotein, *Protein Eng.* **14:** 297–305.

Planson, A.G., Guijarro, J.I., Goldberg, M.E., and Chaffotte, A.F. 2003. Assistance of maltose binding protein to the in vivo folding of the disulfide-rich C-terminal fragment from *Plasmodium falciparum* merozoite surface protein 1 in *Escherichia coli*. *Biochemistry* **42:** 13202–13211.

Pryor, K.D. and Leiting, B. 1997. High-level expression of soluble protein in *Escherichia coli* using a His_6-tag and maltose-binding-protein double-affinity fusion system. *Protein Expr. Purif.* **10:** 309–319.

Riggs, P. 2000. Expression and purification of recombinant proteins by fusion to maltose-binding protein. *Mol. Biotechnol.* **15:** 51–63.

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.

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.

Stevens, R.C. 2000. Design of high-throughput methods of protein production for structural biology. *Structure* **8:** R177–R185.

Tropea, J.E., Cherry, S., Nallamsetty, S., Bignon, C., and Waugh, D.S. 2005. A Generic method for the production of recombinant proteins in *Escherichia coli* using a dual His_6-MBP affinity tag. *Methods Mol. Biol.* (in press).

Uchida, T. and Egami, F. 1966. A general ribonuclease assay. In *Proceedings in nucleic acid research* (eds. G.L. Cantoni and D.R. Davis), pp. 46–55. Harper & Row, New York and London.

Waugh, D.S. 2005. Making the most of affinity tags. *Trends Biotechnol.* **23:** 316–320.