Protein Expression and Purification 110 (2015) 159–164

Contents lists available at ScienceDirect

Protein Expression and Purification

journal homepage: www.elsevier.com/locate/yprep

Positional effects of fusion partners on the yield and solubility of MBP fusion proteins

Sreejith Raran-Kurussi, Karina Keefe, David S. Waugh *

Protein Engineering Section, Macromolecular Crystallography Laboratory, Center for Cancer Research, National Cancer Institute, Frederick, MD 21702, USA

ARTICLE INFO

Article history:
Received 3 March 2015
and in revised form 6 March 2015
Available online 14 March 2015

Keywords:
Gateway cloning
Fusion protein
Inclusion bodies
MBP
Solubility enhancer

ABSTRACT

Escherichia coli maltose-binding protein (MBP) is exceptionally effective at promoting the solubility of its fusion partners. However, there are conflicting reports in the literature claiming that (1) MBP is an effective solubility enhancer only when it is joined to the N-terminus of an aggregation-prone passenger protein, and (2) MBP is equally effective when fused to either end of the passenger. Here, we endeavor to resolve this controversy by comparing the solubility of a diverse set of MBP fusion proteins that, unlike those analyzed in previous studies, are identical in every way except for the order of the two domains. The results indicate that fusion proteins with an N-terminal MBP provide an excellent solubility advantage along with more robust expression when compared to analogous fusions in which MBP is the C-terminal fusion partner. We find that only intrinsically soluble passenger proteins (*i.e.*, those not requiring a solubility enhancer) are produced as soluble fusions when they precede MBP. We also report that even subtle differences in inter-domain linker sequences can influence the solubility of fusion proteins.

Published by Elsevier Inc.

Introduction

The post-genomic era has witnessed a huge evolution in the field of protein production. This was mainly a consequence of advances in the field of molecular biology and associated biotechnological applications. Novel methods are continuously being added to the existing pool of protein expression and purification systems. Even so, *Escherichia coli* is still considered to be the powerhouse of protein production. Its main disadvantage, however, is the frequent formation of inclusion bodies during overexpression. Accordingly, much research has focused on overcoming this bottleneck [1].

One particularly effective means of avoiding the formation of insoluble aggregates during protein expression in *E. coli* is to fuse an aggregation-prone protein to a highly soluble partner [2]. One of the most effective solubilizing fusion partners is *E. coli* maltose-binding protein (MBP)[1] [3,4]. The mechanism of solubility enhancement by MBP has been studied in some detail [5–10]. It is thought that MBP functions as a "holdase", maintaining an aggregation-prone passenger protein in a soluble state until it either folds spontaneously or with the assistance of endogenous molecular chaperones. Alternatively, in some cases the fusion protein may persist in the form of a soluble aggregate [11]. It has been proposed that MBP inhibits the formation of insoluble aggregates by transiently binding folding intermediates of an aggregation-prone passenger protein, effectively sequestering it in an intramolecular interaction that impedes the kinetically competing pathway of intermolecular aggregation and precipitation [4]. A corollary of this hypothesis is that MBP must be folded before it can bind to and sequester its passenger protein [6,9]. If so, then as a result of co-translational folding, it follows that MBP should function as a more effective solubilizing agent when it is fused to the N-terminus of a passenger protein than to its C-terminus because in the former case MBP would emerge first from the ribosome and have time to fold before the passenger protein is translated.

Consistent with this model, Sachdev and Chirgwin found that the aspartic proteases pepsinogen and procathepsin D were soluble in *E. coli* when MBP was fused to their N-termini but formed inclusion bodies when the order of the fusion partners was reversed [12]. However, a subsequent study claimed that MBP could function as a solubility enhancer irrespective of which end of the passenger protein it was fused to [13]. In the current work, we have attempted to resolve this controversy by comparing the solubility of a set of MBP fusion proteins in both orientations. In contrast

* Corresponding author. Tel.: +1 301 846 1842; fax: +1 301 846 7148.
E-mail address: waughd@mail.nih.gov (D.S. Waugh).

[1] *Abbreviations used:* MBP, maltose-binding protein; ORFs, open reading frames; GFP, green fluorescent protein; DHFR, dihydrofolate reductase; DUSP14, dual specificity phosphatase 14; TEV, tobacco etch virus.

http://dx.doi.org/10.1016/j.pep.2015.03.004
1046-5928/Published by Elsevier Inc.

to the previous studies, the MBP fusion proteins that we compare here are identical in every respect (*e.g.*, their interdomain linker sequences) except for the order of the two domains.

Materials and methods

Materials

All materials of the highest available purity were purchased from Thermo Fisher Scientific (Rockford, IL, USA), Sigma–Aldrich (St. Louis, MO, USA), American Bioanalytical Inc. (Natick, MA, USA), EMD Millipore (Billerica, MA, USA), or Roche Diagnostics Corp. (Indianapolis, IN, USA), unless otherwise mentioned.

Construction of expression vectors

We used the Gateway multi-site recombinational cloning to assemble the N-terminal and C-terminal MBP fusion protein expression vectors. The appropriate attB sites (attB1, attB2 or attB3) were incorporated into the gene-specific primers (Tables 1 and 2). The N-terminal and C-terminal open reading frames (ORFs) were generated using standard PCR and were inserted into pDONR208 and pDONR209, respectively (Life Technologies Inc., Carlsbad, CA, USA). The ORFs encoding green fluorescent protein (GFP) [4], dihydrofolate reductase (DHFR) [7], dual specificity phosphatase 14 (DUSP14) [14], and tobacco etch virus (TEV) protease [15] were described previously. The MBP ORF was amplified from pDEST566 (Protein Expression Laboratory, Leidos Biomedical Research, Inc., Frederick, MD, USA; Addgene plasmid 11517) without the His_6 tag. Entry clones were sequence verified and subsequently recombined in tandem into the destination vector pDEST527 (PEL, Leidos Biomedical Research; Addgene plasmid 11518) to create the fusion protein expression vectors. For example, the His-GFP-MBP fusion vector was constructed in three steps. First, the GFP ORF was PCR amplified using the forward primer PE-2688 (attB1-GFP) and the reverse primer PE-2689 (GFP-attB3), and the PCR product was recombined into pDONR208 (BP reaction). Second, the MBP ORF was PCR amplified using two partially overlapping forward primers, PE-2690 (attB3-overlap), PE-2691 (overlap-MBP) and a single reverse primer, PE-2692 (MBP-attB2) (Table 1). The overlap region in the forward primers corresponds to amino acids P1–P6 of the TEV protease recognition site. This PCR product was recombined into pDONR209 (BP reaction). Third, inserts from both the N-terminal and C-terminal entry clones were recombined in tandem into pDEST527 (LR reaction), which includes an N-terminal polyhistidine tag in frame with the Gateway cloning cassette. Hence, the final recombination product expressed a tripartite His-GFP-MBP fusion protein with an attB1 site between the His-tag and GFP and an attB3 site between GFP and MBP. All of the fusion protein expression vectors also included a canonical TEV protease recognition site (ENLYFQG) between MBP and the passenger protein or between the passenger protein and MBP (depending on the orientation), except for the vectors encoding TEV protease fusion proteins, which contained an uncleavable recognition site (ENLYFQP) instead [16]. The P1′ proline substitution in the TEV recognition site of the His-TEV-MBP fusion vector was created by site-directed mutagenesis of the C-terminal entry clone using specific mutagenic primers (PE-2728, PE-2729) and a QuikChange Lightning Kit (Agilent Technologies Inc., Santa Clara, CA), while the P1′ proline substitution in the His-MBP-TEV vector was contained in the gene-specific PCR primer (PE-578) used to amplify the TEV ORF for recombination into pDONR209. These mutant entry clones were used in the subsequent LR reactions to give either His-TEV-MBP or His-MBP-TEV. The coding sequence for the catalytic domain of *Yersinia pestis* YopH was PCR amplified using primers PE-2753 (Table 1) and PE-2755 (5′-GGG GAC AAC TTT GTA CAA GAA AGT TGC ATT AGC TAT TTA ATA ATG GTC G-3′) from a full-length YopH clone (pKM835) and recombined into pDONR221 (Life Technologies) to generate an entry clone. This entry clone was subsequently recombined into pDEST527 to create the His-YopH expression vector. Other His-passenger expression vectors used in this study were reported previously [10]. All reactions were carried out as per the manufacturer's protocol.

Construction of "att" site mutants

The complementary mutagenic primers, PE-2765 and 2766 were used to make ΔattB1 mutants of all N-terminal MBP fusions. These primers anneal to the flanking regions of attB1 and loop out the template region (attB1) to be deleted when the primer-template duplex is formed in a QuikChange reaction. The expression vectors prepared by multisite-Gateway cloning were used as the templates. The attB3 to attB1 change in the N-terminal MBP fusion junctions was made using the complementary mutagenic primers PE-2770 and PE-2771 (Table 3) with the ΔattB1 mutants as templates in a second round of QuikChange reactions. These primers anneal to the flanking regions of attB3 and replace the attB3 sequence with attB1.

Since the flanking residues of the target regions were different in His-passenger-MBP fusions, we had to design a unique pair of complementary mutagenic primers for each of them (Table 3). The ΔattB1 mutants were made using complementary mutagenic primers specific for GFP (PE-2787/PE-2788), DHFR (PE-2789/PE-2790), DUSP14 (PE-2791/PE-2792) and TEV protease (PE-2793/PE-2794). The expression vectors prepared by multisite-Gateway cloning were used as the templates in QuikChange reactions as outlined above. Similarly, the attB3 to attB1 changes at the passenger-MBP fusion junctions were made using pairs of complementary mutagenic primers specific for GFP (PE-2795/PE-2796), DHFR (PE-2797/PE-2798), DUSP14 (PE-2799/PE-2800) and TEV protease (PE-2801/PE-2802). These primers anneal to the flanking regions of attB3 and replace the attB3 sequence with attB1. The ΔattB1 mutants were used as templates for the second round of QuikChange reactions.

The QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent Technologies) was used for engineering these desired modifications throughout. The reaction was performed as per the manufacturer's protocol. Mutagenic primer sequences are listed in Table 3. All mutants were confirmed experimentally.

Expression and solubility analysis

E. coli BL21-CodonPlus (DE3)-RIL cells (Agilent Technologies Inc., Santa Clara, CA) were used for all protein expression experiments. Cells harboring one of the protein expression vectors (see above) were grown to mid-log phase ($A_{600} \sim 0.5$) at 37 °C in Luria broth supplemented with 100 μg ml^{-1} ampicillin and 30 μg ml^{-1} choloramphenicol, at which time production of the fusion protein was induced by the addition of IPTG to 1 mM and the temperature was reduced to 30 °C. Four hours later, the cells were pelleted by centrifugation and re-suspended in approximately 0.2 culture volume (corresponding to an A_{600} of 10.0) of lysis buffer (50 mM Tris–HCl pH 8.0, 150 mM NaCl, 1 mM EDTA). The cells were disrupted by sonication. A total protein sample was collected from the cell suspension after sonication, and a soluble protein sample was collected from the supernatant after the insoluble debris was pelleted by centrifugation at 20,000*g*. These samples were subjected to SDS–PAGE and proteins were visualized by staining with Coomassie Brilliant Blue.

Coomassie-stained gels were scanned with an Alpha Innotech AlphaEase FC Imaging System and the pixel densities of the bands