Solubilization and Purification of a Target Protein Fused to a Mutant Maltose-Binding Protein

BACKGROUND

Recombinant proteins have many uses in biotechnology, whenever large amounts of pure protein are needed. Microbial expression systems such as *Escherichia coli* (*E. coli*) and yeast are often the first choice due to their low cost and high yield. When expressing foreign proteins in *E. coli*, it is not uncommon to encounter problems of low levels of expression and/or insolubility of the protein. Even if the protein is expressed well and remains soluble, it must be purified from the myriad of other proteins in the cell extract. To facilitate the expression and purification of a target protein, one method that is in common use is to fuse an affinity tag to the protein. A good affinity tag has properties that facilitate high-level expression when fused to the N-terminus of the target protein, as well as providing a simple one-step affinity purification that allows the target protein to be purified from the expression mileu.

The maltose-binding protein (MBP) of *E. coli* is commonly used as an affinity tag for expression and purification of foreign proteins produced in *E. coli*. The natural role of MBP is to bind maltodextrins at the outer membrane porin and release them to the MalEFK transport apparatus in the inner membrane. Fusion of the C-terminus of MBP to the N-terminus of a target protein permits the expression of the fusion protein in *E. coli* (Fig. 1). MBP and MBP fusions can be purified in a single step by binding to a chromatography matrix containing any of a number of glucose-α1→4-glucose polysaccharides such as amylose, starch or other maltodextrins (US Patent 5,643,758). Many proteins that are soluble in their native host are insoluble when expressed as a recombinant protein. For many of these proteins, fusion to MBP

renders them soluble (Kapust & Waugh, *Protein Sci.* 8:1668-74 (1999)).

The utility of MBP as an affinity tag is tempered by the fact that depending on the protein in a MBP-target protein purification, some fusions don't bind to the affinity matrix as well as others. In addition, the presence of non-ionic detergents such as Triton X100 and Tween 20 can interfere with binding, especially for MBP-target protein fusions that have an inherently lower affinity.

Researchers have used the structure of MBP to make directed mutations in order to alter the binding properties of MBP. The Xray crystal structure of MBP has been reported by Spurlino et al., *J. Biol. Chem.* 266:5202-5219 (1991). MBP consists of two domains, with a cleft between the domains where the polysaccharide binds. The domain that contains the N-terminus is named the N domain, and the domain that contains the C-terminus is named the C domain. Three loops cross between the two domains to form a hinge. Two groups have used the structure to make directed mutations to the region behind the hinges that increase the affinity of MBP for maltose and maltotriose (Marvin et al., *Nature Structural Biology* 8:795-798 (2001); Telmer & Shilton, *Journal of Biol. Chem.* 278:34555-34567 (2003)). However, this approach has an inherent disadvantage, since the modifications to MBP increase the hydrophobicity of the surface of the protein and thus decrease its solubility. This reduces its utility as an affinity tag by increasing its tendency to render a fusion protein insoluble.

SUMMARY

In an embodiment of the invention, a modified MBP fusion protein is provided that is characterized by an MBP amino acid sequence having a mutation wherein the mutation causes the

modified MBP fusion protein to have at least one property selected from (i) an increased affinity for a maltodextrin substrate and (ii) an increased solubility when fused to a target protein having limited solubility. In a further embodiment of the invention, the modified MBP has a mutation located in the hinge region between helices XI and XII or in a region within 10 Å of A313 of the protein. For example, the mutation may be located in the C domain or in a region within 10 Å of S146, at the beginning of β-sheet F of the protein. The modification may be specifically A313V or S146T.

In a further embodiment of the invention, the modified MBP may be fused to a target protein to form a fusion protein and may, for example, have a solubility that is greater than the solubility of an unmodified MBP protein fused to the target protein.

In an embodiment of the invention, the modified MBP may have an amino sequence selected from SEQ ID NO:4 or SEQ ID NO:6 or a DNA sequence encoding the protein selected from SEQ ID NO:3 or SEQ ID NO:5. The DNA may be incorporated into a vector such as a pKLAC1 vector for expression in a host cell such as a *Kluyveromyces* or *E. coli*.

In an embodiment of the invention, a method is provided for purifying a protein that includes: expressing in a host cell such as a *Kluyveromyces* or *E. coli*, a fusion protein that includes any of the modified MBPs described above and a target protein. The modified MBP fusion protein is permitted to bind to a matrix such as a polysaccharide exemplified by a maltodextrin. The fusion protein can then be eluted from the matrix in a selected buffer to obtain the purified protein.

In a further embodiment of the invention, a method is provided for solubilizing a target protein that includes expressing an MBP mutant fused to a target protein so that the fusion protein is solubilized to an extent greater than can be observed in the absence of the mutant MBP and to an extent greater than observed using a non-mutated MBP. Examples of mutations include a A313V mutation and/or an S146T mutation.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 shows a schematic describing the cloning and purification of a target protein by expressing a DNA encoding an MBP fused to a target protein, allowing the fusion protein to selectively bind to amylose, eluting the target protein in a maltose-containing buffer and then recovering the target protein from the purified fusion protein by protease cleavage.

Fig. 2 provides sequences comparing wild-type MBPs with modified MBPs.

Fig. 2A: The DNA sequence (SEQ ID NO:1) encoding wild-type MBP (SEQ ID NO:2) from pMAL-c2X .

Fig. 2B: The DNA sequence (SEQ ID NO:3) encoding the MBP mutant A9 (SEQ ID NO:4). Changes in the modified MBP sequences are indicated in bold.

Fig. 2C: The DNA sequence (SEQ ID NO:5) encoding the MBP mutant G8-1 (SEQ ID NO:6) . Changes in the modified MBP DNA and amino acid sequences are indicated in bold.

Fig. 3 shows a histogram comparing yields of protein using wild-type MBP (pKO1483), or modified MBPs (A9 or G8-1). The yields of MBPs are provided in mg/500 mls of culture.

Fig. 4 shows on an SDS-PAGE gel the purified product of MBP-Klenow from pMAL containing wild-type MBP, MBP A313V and S146T modifications. Plasmids: WT= pIH1062, A313V=pIH1643, S136T=pIH1644. Lanes: CE = crude extract, FT = flow though, and E= eluate. Equal portions of the indicated fraction were loaded in each lane. A significantly higher yield of MBP fusion protein was obtained from A313V mutant MBP than from wild-type MBP (also see Table 1).

Fig. 5. Sequence of pMB50 (SEQ ID NO:7) used for expressing MBP-CBD fusion protein (see Table 2).

Fig. 6. SDS-PAGE gels showing the effect on solubility of fusing wild-type and modified MBPs to DHFR and GAPDH. DHFR fusion plasmids: WT= pIH1616; A313V=pIH1617; S146T=pIH1618; DM=pIH1646. GAPDH fusion plasmids: WT=pIH1625; A313V=pIH1626; S146T=pIH1627; DM=pIH1645. Lanes: T=total cell extract; S=soluble extract; I=resuspension of insoluble material. Equal portions of the indicated fraction were loaded in each lane. The ratio of soluble to insoluble protein was greater for both mutants compared with wild-type MBPs.

Fig. 7 provides the sequence of pKLMF2-PMΔSal (SEQ ID NO:8).

Fig. 8 shows fractions from the affinity purification, using an amylose resin column, of a fusion protein on an SDS-PAGE gel. An enhanced yield of modified MBP fusion protein was produced by expressing MBP (A313V)PMΔSal in *Kluyveromyces lactis* cells carrying an MBP (A313v)-PMΔSal expression cassette.

Lane 1 - NEB Broad range markers, Ipswich, MA;

Lane 2 - crude extract load;

Lane 3 - column flow-through;

Lane 4 - column wash;

Lanes 5, 6 and 7, elution with column buffer + maltose.

DETAILED DESCRIPTION OF THE EMBODIMENTS

Terms that are used herein are discussed below.

"Wild-type" MBP includes the MBP protein produced by expression from a derivative of one of the pMAL-2 plasmids that has a stop codon in the polylinker, for example pKO1483.

"Enhanced solubility of a protein fused to a mutant MBP" is an increase in the amount of soluble protein when compared to that same protein fused to wild-type MBP. Solubility can be expressed as the ratio of soluble protein to the total amount of that protein present before insoluble material is removed, for example by centrifugation.

"Increased affinity of a mutant MBP" or "mutant MBP fusion protein" includes: an increase in the amount of protein that binds to a solid substrate such as a maltodextrin under a defined set of conditions. The efficacy of the affinity purification can be expressed as the ratio of protein that binds to maltodextrin under the specified conditions and is then eluted with a specified buffer to the total amount of that protein applied to the column.

The present embodiments of the invention provide MBP mutants which when fused to a target protein enhance the solubility of the fusion protein during expression *in vivo* and can also improve the affinity of the fusion protein during purification. The present embodiments include mutant MBPs that show increased binding to a polysaccharide media, such as one that includes a maltodextrin, when applied to the media under conditions where wild-type MBP

shows partial binding. The modified MBPs are then eluted from the media using a solution of, for example, a soluble maltodextrin, yielding at least 1.5 to 10-fold more protein when compared to wild-type MBP.

In order to discover these improved mutants of MBP, technical hurdles had to be overcome which include developing techniques which enable a large number of samples to be handled. This required improved methods for breaking up host cells to release solubilized fusion protein where sonication is not practical for large scale purification and lysis buffers could interfere with affinity binding of MBP. It was discovered that by titrating the detergent and the lysozyme, it was possible to identify the appropriate concentration and ratio of these lysis reagents to effectively break up host cells without negatively impacting binding affinity.

In order to screen for mutants with desired binding affinity properties, 96 well microplates were used where each well contained a micro matrix for binding fusion protein and a filter apparatus removed contaminating materials in the filtrate. This made possible rapid screening of large numbers of samples.

The screening methods for obtaining and testing modified mutant MBP proteins as improved tags for purifying proteins are described in the examples. Such modified MBPs that have a higher affinity for a matrix solve the problem associated with wild-type MBP of MBP fusions proteins that bind poorly to a matrix or where binding is disrupted by the presence of non-ionic detergent.

In the examples, two mutations (S146T and A313V) are described with the desired properties of improved solubility and improved binding affinity to a polysaccharide matrix that were isolated using the screening procedure. The S146T mutation is in

the C domain of MBP at the beginning of β-sheet F. When a target protein that is insoluble is fused to MBP containing this mutation, the solubility of the fusion protein is enhanced. The A313V mutation described herein is located in the third hinge region that crosses between the two domains, specifically, in the loop between helices XI and XII. This mutation enhances both the solubility and the affinity of fusion proteins. When expressing foreign proteins in *E. coli*, the protein may be partially or completely expressed in the form of insoluble aggregates. In particular examples, solubility may be increased by 1.1 or greater upwards with an upper limit of total solubility.

All references cited herein, as well as U.S. provisional application number 60/792,133 filed April 14, 2006, are incorporated by reference.

EXAMPLES

Materials

Restriction enzymes, β-agarase, DNA polymerases, T4 ligase, Antarctic phosphatase, Litmus 38, the pMAL Protein Fusion and Purification System including pMAL-c2X and pMAL-c2G, amylose resin (#E8021), anti-MBP monoclonal antibody linked to horse radish peroxidase (#E8038), the USER Friendly Cloning kit, the *K. lactis* Protein Expression Kit including the vector pKLAC1, host strains TB1, ER1992, ER2502, ER2984, NEB 5-alpha, and NEB Turbo, and synthetic oligonucleotides were obtained from New England Biolabs, Inc. (NEB), Ipswich, MA. Unifilter 800 microtiter microplates with filter bottoms were purchased from Whatman, Brentford, England. The Minelute DNA Extraction and Qiaprep Spin kits were purchased from Qiagen, Valencia, CA. Mega 10 was purchased from Dojindo, Gaithersburg, MD. Hen egg white

lysozyme, Coomassie brilliant blue R and acid washed glass beads (425-600 micron) were purchased from Sigma-Aldrich, St. Louis, MO. Sea Plaque GTG low melting temperature agarose was purchased from Cambrex, E. Rutherford, NJ. Disposable polypropylene columns (#732-6008) were purchased from BioRad, Hercules, CA. 10-20% gradient gels were purchased from either Daiichi, Tokyo, Japan or InVitrogen/Novex, Carlsbad, CA. The Complete™ protease inhibitor cocktail was purchased from Roche, Basel, Switzerland. SimplyBlue Safestain was purchased from Invitrogen, Carlsbad, CA. The human dihydrofolate reductase (DHFR) cDNA clone pOTB7-DHFR was purchased from Invitrogen (MGC:857). The GAPDH gene was obtained from pJF931 (Fox et al. *FEBS Lett.* 537:53-57 (2003).

Techniques

The *Serracia marscesens* nuclease was obtained as described in PCT/US05/28739. Minipreps of plasmid DNA were prepared using the Qiaprep Spin kit. Random PCR mutagenesis was carried out as described in Fromant et al. *Analytical Biochemistry* 224, 347-353 (1995). PCR was carried out using Vent® DNA polymerase except as noted. DNA fragments were gel-purified by electrophoresis on 1% Sea Plaque GTG low melting temperature agarose, cutting out the band, and either purifying the DNA using the Minelute DNA Extraction kit, or melting it at 75°C for 5 minutes, cooling to 37°C, and digesting with β-agarOse for 1-2 h. DNA sequencing was performed on Applied Biosystems's (ABI's) automated DNA Sequencer model 3100 ABI, using Big Dye labeled dye-terminator chemistry (ABI, Foster City, CA). SDS-PAGE was carried out according to the instructions of the acrylamide gel provider, and proteins were visualized by staining with Coomassie brilliant blue R except where noted otherewise.

MBP was expressed from either pMal-c2X or pMal-c2G or a derivative of pMal-c2G. The numbering of bases to identify mutations in *malE* refers to the base number in the pMAL-c2X sequence ((Figures 2A-1, 2A-2, 2B-1, 2B-2, 2C-1 and 2C-2 (SEQ ID NOS:1-6)). The pMAL-c2G derivative pSN1578 was created by cleaving the plasmid with BsmI and BsiWI, treating the product with DNA polymerase Klenow fragment plus all four dNTPs, followed by ligation to create a deletion within the *malE* gene.

Site-directed mutagenesis was carried out using a four primer PCR mutagenesis as described in Guan et al. *Nucleic Acid Research*, 33:6225-6234 (2005). MBP and MBP fusion proteins were purified as described in the instructions for the pMAL Protein Fusion and Purification System, except in some cases cells were lysed with a lysozyme/detergent solution instead of sonication.

Large-scale purifications were carried out with crude cell extract prepared from 500 to 1000 mL of culture, and loaded on a 2.5 cm diameter column containing 15 ml of amylose resin (NEB #E8021, Ipswich, MA). Small-scale purifications were carried out with crude extract prepared from 67 ml of culture, and loaded on a disposable polypropylene column containing 1 ml of amylose resin. SDS-polyacrylamide gel electrophoresis was carried out using 10-20% gradient gels. For quantitation of gel bands, gels were dried between cellophane sheets and scanned using a Microtek III scanner Microtek, Carson, CA, and densitometry carried out using Image J (NIH).

Example I: Isolation of Mutants in MBP with Improved Properties

Screening for improved yield after purification:

Random mutagenesis of the *malE* gene from pMAL-c2x was achieved by error-prone PCR using the primers:
oligo 1: 5' GGAGACAUGAATTCAATGAAAATCGAAGAA
(SEQ ID NO:9), and oligo 2:

5' GGGAAAGUAAGCTTAATCCTTCCCTCGATC (SEQ ID NO:10)

PCR fragments were cloned into linearized pNEB208A using the USER Friendly Cloning Kit, following the manufacturer's instructions. Transformants were grown overnight in 1 mL LB + 1 mM IPTG and 100 ug/ml ampicillin, then lysed by adding 0.3 mg/mL lysozyme and 20 units of the *S. marscens* nuclease, incubating for 10 min, then adding 0.1 ml of 2% Tween 20.

The crude extracts were applied to a 50 uL amylose resin column (NEB #E8021, Ipswich, MA) in a Unifilter 800 microplate, and each well was washed with 0.7 ml of 20 mM Tris-Cl, 0.2 M NaCl, 1 mM EDTA, pH 7.4 (column buffer), then with 0.7 mL of 10 mM sodium phosphate, 0.2 M NaCl, 1 mM EDTA, pH 7.2. The protein bound to the amylose resin was then eluted with 0.2 mL of 10 mM maltose, 10 mM sodium phosphate, 0.2 M NaCl, 1 mM EDTA, pH 7.2. The eluate was transferred to an Immulon 2HB microtiter plate (ThermoFisher Scientific, Waltham, MA) and incubated overnight at 4°C. The microtiter wells were then emptied, washed twice with 20 mM Tris-Cl, 150 mM NaCl, pH 7.5 (TBST), then blocked with 0.36 ml TBST + 3% bovine serum albumin for 1 h at 37°C.

The wells were washed twice with TBST, then 0.1 ml of a 1:2000 dilution of anti-MBP monoclonal antibody linked to horse radish peroxidase in TBST + 3% bovine serum albumin was added to each well and the plate incubated at 37°C for 1 h. The wells were emptied, then washed twice with TBST. The wells were developed

with 0.01% o-phenylenediamine, 0.003% hydrogen peroxide in water. The detection reaction was stopped by adding 0.025 mL 4 M H_2SO_4, and wells were assayed spectrophotometrically at 490 nm. Cells were recovered from lysates corresponding to samples that showed higher binding and elution as compared to wild-type MBP. These candidates were grown and retested to confirm the higher binding and elution.

Characterization and separation of mutations obtained after random mutagenesis

Two isolates from a library in USER having increased binding and elution profiles were sequenced (Fig 2). One isolate, G8-1, was found to have a single mis-sense mutation, G1964C, along with a silent mutation.The G1964C mutation corresponds to the amino acid change S146T in MBP. The other isolate, A9 was found to have three mis-sense mutations, A1583G, A2419G and C2465T, along with a silent mutation. The A1583G, A2419G and C2465T mutations correspond to the amino acid changes N19S, K298E and A313V, respectively.

Subcloning into pMal-C2X or pSN1578

Each isolate was amplified by PCR with the following primers: oligo 3:
5'
GACTCATATGAAAATCGAAGAAGGTAAACTGGTAATCTGGATTAACGGC (SEQ ID NO:11) and oligo 4:
5' ATATAAGCTTTCACCTTCCCTCGATCCCGAGGT (SEQ ID NO:12)
The amplified DNA was ethanol precipitated, cut with NdeI and HindIII in NEBuffer 4 (NEB, Ipswich, MA), and gel purified. pSN1578 was cleaved with NdeI and HindIII and the vector backbone was gel purified. The G8-1 and A9 fragments were mixed

with the pSN1578 fragment and ligated, and the ligation was used to transform TB1. A plasmid preparation from each transformant was sequenced and named pIH1596 for G8-1 and pIH1593 for A9. The 3' primer in this experiment has a stop codon in the correct reading frame to prevent malE translation from proceeding into the *lacZα* fragment of pMAL. Thus, these subclones produce a modified MBP that ends after the amino acid sequence ...IEGR encoded by the polylinker. A control plasmid containing a wild-type *malE* gene followed by a stop codon was constructed by cleaving pMAL-c2X in the polylinker between *malE* and *lacZa* with XbaI. The XbaI overhang was filled in using DNA polymerase I, large fragment (Klenow) and all four dNTP's, then the plasmid was recircularized by treatment with T4 ligase. This introduces a stop codon in the same reading frame as *malE,* and this derivative produces an MBP comparable to that produced by G8-1 and A9, except for an 8 residue extension encoded by the polylinker. This control plasmid was named pKO1483. *E. coli* TB1 containing pKO1483, pIH1596 and pIH1593 were grown in a 500 mL culture of LB + 0.1% glucose and 100 ug/ml ampicillin to 2 x 10^8 cells/ml, induced with 0.3 mM IPTG, grown for 2h at 37°C, then harvested. The cells were resuspended in 25 ml column buffer (0.2 mL of 10 mM maltose, 10 mM sodium phosphate, 0.2 M NaCl, 1 mM EDTA, pH 7.2)+ 10 mM β-mercaptoethanol, then lysed by sonication. The extract was clarified by centrifuging at 9000 x g for 30 min, then diluted 1:4 with column buffer and loaded onto a 15 ml column of amylose resin. The column was washed with about 125 mL column buffer, and eluted with column buffer + 10 mM maltose. The yields of MBP were compared among the three strains (Fig. 3). The results confirm that the modified MBPs showed an increased binding to amylose and elution in appropriate buffers.

In order to ascertain which of the three mutation(s) were necessary for increased binding of the A9 variant, the three mutations were subcloned separately into pSN1578, a pMAL-c2G derivative with a deletion internal to the *malE* gene (which allows easy identification of clones which receive an insert). The A1583G and A2419G mutations either had no effect or reduced the yield of MBP in the affinity purification, and were discarded. The C2465T mutation was recreated in isolation by 4 primer site-directed PCR mutagenesis using pMAL-c2X as the first template, with the primers oligo 5 : 5' CTTCAAGGGTCAACCATCCAAACC (SEQ ID NO:13) and oligo 6: 5′ AATACGCGGATCTTTCACCAACTCTTC (SEQ ID NO:14) to create the N-terminal PCR fragment, and with primers oligo 7: 5′ GAAGAGTTGGTGAAAGATCCGCGTATT (SEQ ID NO:15) and oligo 8: 5′ CTGAGAATTCTGAAATCCTTCCCTCGAT (SEQ ID NO:16) to create the C-terminal PCR fragment. The assembly step was carried out with the gel-purified N- and C-terminal fragments as the template and the primers oligo 5 and oligo 8. The final PCR fragment was cut with BlpI and AvaI, gel purified, and ligated to pMAL-c2X that had been cut with BlpI and AvaI and gel purified. The ligation was used to transform TB1, and plasmid was purified from the transformants and sequenced to confirm the C2465T mutation. An isolate was chosen for further study and named pIH1606.

In the construction of pIH1606, the stop codon at the end of MBP was not conserved; this construct expresses MBP fused to the LacZaa fragment. In order to compare the effect of the C2465T mutation to its parent, A9, a stop codon was introduced after the *malE* gene in pIH1606. The plasmid was cleaved with XbaI, filled in with Klenow plus dNTP's, and religated as described above for pKO1483. The C2465T derivative with a stop codon was called pPR1610. Large scale MBP purifications of TB1 bearing this plasmid,

in parallel with pKO1483 and A9, showed that all of the increase in yield of MBP found in A9 could be accounted for by the C2465T mutation. This mutation changes alanine 313 of MBP to a valine (A313V).

In order to be able to compare MBP (S146T) to wild-type MBP and MBP (A313V) in derivatives that have exactly parallel construction, a version of MBP (S146T) was constructed that has the same stop codon as pKO1483 was constructed. An NdeI, BlpI fragment from pIH1596 was purified and subcloned into pKO1493 cut with NdeI and BlpI, creating pIH1619.

Example II: Increased Yield of MBP Fusion Proteins

A: MBP-Klenow

In order to test if the modified MBPs can increase the yield of a fusion protein after purification, the gene encoding the Klenow fragment of *E. coli* DNA polymerase I was cloned into pMAL-c2X, pIH1619 (S146T) and pPR1610 (A313V). The MBP-Klenow fusion was chosen because it has an inherently low affinity for the amylose column, and during the affinity purification some of the MBP-Klenow protein flows though the amylose column without binding. The Klenow portion of DNA polymerase I was PCR'd from the plasmid pPolA, which contains the BglII-HindIII fragment of the polA gene (Genbank ecopolA:206-4127) cloned into pBR322 between the BglII and HindIII sites. PCR was carried out using the primers oligo 9: 5' CCAGAAGTGACGGCAACGGTGATT ((SEQ ID NO:17) and oligo 10: 5' AAGTGCGGCGACGATAGTCATGCCCCGCGC (SEQ ID NO:18).

The PCR fragment was cleaved with HindIII and ligated to pMAL-c2, which had been cleaved with XmnI and HindIII. The resulting construct was named pIH1040. In order to reduce the

chance of PCR errors in the insert, the region from XhoI to HindIII was replaced with the corresponding fragment from pPolA. This construct was checked by sequencing and named pIH1062. The gene encoding the Klenow fragment was cloned into pPR1610 and pIH1619 by PCR using pPolA as a template and the primers oligo 11: 5′ GTGATTTCTTATGACAACTACGTCACCATCCTTGATG (SEQ ID NO:19) and oligo 12: 5′ TTAAGGATCCTTAGTGCGCCTGATCCCAGT (SEQ ID NO:20)

The PCR fragment was cleaved with BamHI, and the vectors were cleaved with XmnI and BamHI, purified, mixed with the PCR fragment, and ligated. The ligation was used to transform NEB Turbo (Ipswich, MA), and transformants were checked for the correct structure by restriction analysis and by expression of the MBP-Klenow fusion and analyzing by SDS-PAGE. The pIH1610-Klenow construct was named pIH1643, and the pIH1619-Klenow construct was named pIH1644.

Affinity purification of MBP-Klenow was performed using cells containing pIH1062, pIH1643 and pIH1644. Crude extract, column flow-through and eluate fractions from each strain were analyzed by SDS-PAGE (Fig. 4). The eluted protein was quantitated by measuring A_{280} and quantitated using the predicted extinction coefficient of the MBP-Klenow protein (Table 1). Cells bearing pIH1643, carrying the MBP-A313V modification, yielded more than twice as much fusion protein as those containing the wild-type MBP or the S146T modified MBP fusion plasmids.

B. MBP-chitin binding domain:

In order to test if the mutant MBP's ability to increase the yield of fusion protein from the affinity purification was a general property, another fusion protein that has an inherently low affinity

for amylose was tested. The MBP fused to the *Bacillus circulans* chitin binding domain (MBP-CBD) is encoded by the plasmid pMB50 (Fig. 5). A good fraction of this MBP-CBD fusion protein tends to flow through an amylose column during the affinity purification, similar to MBP-Klenow. The portion of pPR1610 encoding MBP(A313V) was cleaved from the plasmid using HpaI and SacI, and the fragment was purified. pMB50 was cleaved with the same enzymes, the backbone was also purified, and the two fragments were ligated and transformed into ER2523. The resulting plasmid was named pIH1660.

Affinity purification of MBP-CBD was performed using cells containing pMB50 and pIH1660 harvested from 500 mL cultures. The eluted protein was quantitated by measuring A_{280} and using the predicted extinction coefficient of the MBP-CBD protein (Table 2). Cells bearing pIH1660, carrying the MBP-A313V modification, yielded almost twice as much fusion protein as those containing the wild-type MBP fusion plasmid.

Example III: Solubility Enhancement of MBP Fusion Proteins

Construction of MBP Fusions

In order to test whether the modified MBP's retained the ability to enhance the solubility of the protein fused to MBP, two proteins that tend to be insoluble in *E. coli*, dihydrofolate reductase (DHFR) and glyceraldehyde-3-phosphate dehydrogenase (GAPDH), were cloned into pIH1619 (S146T) and pIH1610 (A313V). As controls, the same two proteins were cloned into pMAL-c2X and a pMAL-c2G derivative containing the mutations in the Telmer et al. modified MBP named "MBP-DM." The vector encoding MBP-DM was constructed as follows: First, a pMAL-c2G derivative with a translationally silent NsiI site at nucleotide 2030 was constructed by

four primer site-directed PCR. The template for both the N- and C-terminal fragments was pMAL-c2X. The primers
oligo 13: 5′ CCATAGCATATGAAAATCGAAGAAG (SEQ ID NO:21)
and oligo 14: 5′ CTTGAATGCATAACCCCCGTCAGCAGC
(SEQ ID NO:22) were used to create the N-terminal fragment, and the primers oligo 15: 5′ GGTTATGCATTCAAGTATGAAAACGGCAAG (SEQ ID NO:23) and oligo 8 were used to create the C-terminal fragment. The PCR fragments were gel purified, then used as a template in the assembly step with the primers oligo 13 and oligo 8. The final PCR fragment was ethanol precipitated, cleaved with NdeI and EcoRI, and gel purified. pMAL- c2X was cleaved with NdeI and EcoRI and gel purified, and the two fragments were mixed and ligated. The ligation was used to transform ER2502, overnight cultures were grown, and miniprep DNA prepared. Several transformants that had acquired the NsiI site and were of the correct size were obtained, but unexpectedly, all of them lacked an EcoRI site. A representative of this set, named pPR1629, was cleaved with NdeI and AvaI, and the *malE* fragment was gel purified. pSN1578 was cleaved with the same two enzymes and the plasmid backbone was gel purified. These two DNA fragments were ligated, the ligation was used to transform ER2502, and plasmid DNA from one transformant was sequenced, and named pPR1633. Second, the gene for the modified MBP called DM by Telmer and Shilton was constructed by four primer PCR site-directed mutagenesis as follows: The template for both the N- and C-terminal fragments was pMAL-c2X. The primers used for the N-terminal fragment were
oligo 16: 5′
TATGCATTCAAATACGGTGACATTAAAGACGTGGGCGTGGAT
(SEQ ID NO:24) and oligo 17: 5′
GCGGCGTTTTCCGCAGTGGCGGCAATACGTGGATCTTTC

(SEQ ID NO:25). The C-terminal fragment was produced with the primers oligo 18:
5' GCGGAAAACGCCGCGAAAGGTGAAATCATGCCGAACATC
(SEQ ID NO:26) and oligo 8. The assembly PCR was performed using the purified N- and C- terminal fragments with oligo 16 and oligo 8 as the primers. The PCR fragment was ethanol precipitated and ligated to Litmus 38 that had been linearized with EcoRV and treated with Antarctic phosphatase. The ligation was used to transform NEB 5-alpha (Ipswich, MA), and plasmid DNA from one transformant was sequenced to confirm the construction and named pPR1638. Plasmid DNAs pPR1633 and pPR1638 were cleaved with NsiI and AvaI, the plasmid backbone from pPR1633 and the '*malE* fragment from pPR1638 were gel purified, ligated, and the ligation was used to transform NEB Turbo. Plasmid DNA from one transformant was selected for sequencing and named pPR1639.

The DHFR gene was PCR'd using pOTB7-DHFR as a template and using the primers
oligo 19: 5' GGATGGTTGGTTCGCTAAACTGCATCGTC (SEQ ID NO:27) and oligo 20: 5' TATTAATCATTCTTCTCATATACTTCAAA (SEQ ID NO:28), then treating the PCR fragment with T4 polymerase and dT for 15 m at room temperature. This treatment produces a short 3' overhang on each end of the fragment, a GG on the upstream end of the DHFR gene, and an A on the downstream end. The vectors pMAL-c2X and pPR1610 were prepared by cleaving with XmnI, then treating with T4 polymerase and dA, which likewise produces a 3' CC overhang at the end of the *malE* region and a T on the *lacZa* end. The vector DNAs were ligated to the DHFR fragment and the ligation was used to transform ER1992. For each vector, a transformant was confirmed by restriction analysis and expression of a fusion protein of the expected size analyzed by SDS-PAGE. The pMAL-c2X-DHFR isolate was named pIH1616, and the pPR1610-

DHFR isolate was named pIH1617. The DHFR insert was then subcloned into pIH1596 as follows: pIH1616 was cleaved with NdeI and BlpI, the digest run on a gel, and the backbone (including the DHFR gene) was cut out and gel purified. pIH1596 was cut with NdeI and BlpI, and the malE' was gel purified. The vector backbone was ligated to the malE' fragment, and the ligations used to transform ER1992. An isolate was confirmed by restriction analysis and expression of a fusion protein of the expected size analyzed by SDS-PAGE, and named pIH1618. The DHFR insert was subloned into pPR1639 by cleaving both pPR1639 and pIH1616 with AvaI and SalI, gel purifying the vector backbone from pPR1639 and the DHFR fragment from pIH1616, and ligating the two fragments. The ligation was used to transform NEB Turbo, and a transformant was confirmed by sequencing and named pIH1646.

The GAPDH gene was subjected to PCR using pJF931 as a template and using the primers
oligo 21: 5' GGATGGTGAAGGTCGGTGTGAACGG (SEQ ID NO:29)
and oligo 22: 5' TATTACTCCTTGGAGGCCATGTAGGCCA
(SEQ ID NO:30), then treating the PCR fragment with T4 polymerase and dT.

The vectors pMAL-c2X, pPR1610 and pIH1619 were prepared by cleaving with XmnI, then treating with T4 polymerase and dA as described above. The vector DNAs were ligated to the GAPDH fragment and the ligation was used to transform ER1992. For each vector, a transformant was confirmed by restriction analysis and expression of a fusion protein of the expected size analyzed by SDS-PAGE. The pMAL-c2X-GAPDH isolate was named pIH1625, the pPR1610-GAPDH isolate was named pIH1626, and the pIH1619-GAPDH isolate was named pIH1627. The GAPDH insert was then subcloned into pPR1639 by cleaving both pPR1639 and pIH1625

with AvaI and SalI, gel purifying the vector backbone from pPR1639 and the GAPDH fragment from pIH1625, and ligating the two fragments. The ligation was used to transform NEB Turbo, and a transformant was confirmed by restriction analysis and expression of a fusion protein of the expected size analyzed by SDS-PAGE, and named pIH1645.

Solubility profile

A 20 ml culture of TB1 containing the eight pMAL plasmids ((pMAL-c2X, pIH1619 (carrying S146T) and pPR1610 (carrying A313V), each plasmid further containing a DNA encoding DHFR or GApDH)) was grown to $2x10^8$/ml in LB amp, induced with 0.3 mM IPTG, incubated for an additional two hours, then cells were harvested by centrifugation at 3000 x g in a microfuge. Each pellet was resuspended in 2 ml of 50 mM Tris-Cl, pH 7.9, 50 mM NaCl, 0.75 mM EDTA, 0.6% Mega 10, 150 ug/ml lysozyme, and 20 Kunitz units/ml *Serracia marscesens* nuclease, and incubated at room temperature for 10 m. The resuspended pellet was designated the total cell extract. A sample of 125 ul was removed and centrifuged for 2 m at 14,000 x g. The supernatant was removed and designated the soluble fraction. The pellet was resupended in 125 ul of the same buffer, and designated the insoluble fraction. A sample (5 ul) of each fraction was run on SDS-PAGE for each strain (Fig. 5). The gels were dried and scanned, and the amount of MBP fusion protein in each lane was quantitated as a ratio of soluble protein to protein present in the cell lysate before centrifugation (Table 2). For both DHFR and GAPDH fusions, the A313V and S146T modified MBPs enhanced the solubility of the fusion protein as compared to wild-type MBP. Fusions made with the DM modified MBP, as expected, showed reduced solubility compared to wild-type MBP.

Example IV: Use of an MBP Mutant in K. lactis

Construction of a *K. lactis* MBP(A313V)-fusion Expression Vector

The gene encoding the mutant maltose-binding protein, MBP(A313V), was amplified by PCR using the forward and reverse primers, respectively:

oligo 23: 5' GCCC**AAGCTT***GCCACC*ATGAAAATCGAAGAAGGT (SEQ ID NO:31) and oligo 24:

5' GCG**CTCGAG***CTTGTCATCGTCATC*CGAGCTCGAATTAGTCTGCGC (SEQ ID NO:32). The forward primer was engineered to contain a *Hind*III restriction enzyme site (bold text) followed by the Kozak consensus sequence (italics) that immediately preceeds the *malE* gene initiation codon (underlined). The reverse primer was engineered to contain a *Xho*I restriction enzyme site (bold text) immediately followed by DNA encoding the proteolytic recognition site of the protease enterokinase (italic and underlined text). No stop codon was incorporated onto the reverse primer to allow for the construction of in-frame C-terminal MBP(A313V) fusion expression cassettes. The *malE* gene was amplified from the plasmid pPR1610 containing the full-length gene, using Phusion polymerase. The amplified gene was cloned into the *Hind*III and *Xho*I restriction sites of the *K. lactis* expression vector pKLAC1 to create a *K. lactis* MBP(A313V)-fusion expression vector (pKlMF2). This cloning strategy results in the replacement of the *K. lactis* α-mating factor pre-pro signal sequence in pKLAC1 with the *malE* gene. Thus the MBP- and MBP(A313V)-fusion proteins will not be directed to the secretory pathway but instead will be retained in the yeast cytosol.

Expression and Purification of an MBP(A313V)-fusion Protein in *K. lactis*

The gene encoding a truncated form of paramyosin (Steel et al., *J. Immunol.* 145:3917-3923 (1990)) was amplified by PCR

using the following primers:
oligo 25: 5'GCG**CTCGAG**AATTCCGCATTCGGTAGTATG (SEQ ID NO:33) and oligo 26:
5'ATAAGAAT**GCGGCCGC***TCA*CGACGTTGTAAAACGACGGCCAGT (SEQ ID NO:34). The forward primer was engineered to contain a XhoI restriction site (bold text). The reverse primer was engineered to contain a NotI restriction site (bold text) immediately upstream of the PMΔSal stop codon (italic text).

The cloning strategy was as follows. The ~750bp paramyosin gene was amplified and purified. The fragment was double-digested with *Xho*I and *Not*I and cloned into the XhoI and NotI restriction sites of the expression vector, pKlMF2 creating an in-frame fusion between the C-terminus and the N-terminus of PMΔSal. The plasmid constructed in this way was named pKLMF2-PMΔSal (Figure 7).

The MBP(A313V)-PMΔSal fusion vector (pKlMF2- PMΔSal) was linearized by *Sac*II restriction digestion and the purified product was transformed into chemically competent *K. lactis* cells using the *K. lactis* Protein Expression Kit according its instructions. Transformant colonies were selected on acetamide plates and clones containing the correctly integrated MBP(A313V)-PMΔSal expression cassette were identified by whole cell PCR as described in the above kit.

A 20 ml YPGalactose medium (1% yeast extract, 2% peptone, 2% galactose) culture was inoculated with a strain of *K. lactis* cells containing a multi-copy integrated MBP(A313V)- PMΔSal expression cassette and grown at 30°C in an incubator with a shaking platform for 4 days. The total cell content of the culture after growth was determined to be approximately 8.2×10^{10} cells (or approximately 4.1×10^{8} cells/ ml). Cells were harvested by centrifugation at

8,000 rpm for 10 minutes at 4^0 C. The cells were washed once in distilled water to remove the excess media.

For the preparation of a lysate, the cells were re-suspended in a 2 ml solution of 10 mg/ ml lyticase in 1M sorbitol and incubated at 37°C for 1 hour. Cells were harvested by gentle centrifugation (7,000 r.p.m.) in a micro-centrifuge for 2 minutes. Cell pellets were resuspended in a total volume of 3 ml ice-cold amylose column buffer (20 mM Tris-Cl ph 7.4, 0.2 M NaCl, 1 mM EDTA) containing a protease inhibitor cocktail (Complete™ protease inhibitor cocktail, Roche). The cell slurry was transferred to a glass tube and placed on ice. An equal volume of acid washed glass beads (425-600 micron, Sigma) were added to the cell slurry and cells were broken by vortexing 10 times, each for a 1 minute duration. Cell lysate was transferred to a new tube and the glass beads washed 4 times with 1 ml column buffer. The cell lysate and washes were pooled and the cellular debris was removed by centrifugation. Cleared cell lysate was diluted to 8 ml with column buffer.

For fusion protein purification, the cell lysate was passed over a 1.6 ml amylose resin column pre-equilibrated with 15 ml of column buffer. The column was washed with 24 ml of column buffer and bound protein was eluted in 0.4 ml fractions with column buffer containing 10 mM maltose. Cell lysate and purified proteins were resolved by SDS-PAGE on a 10-20% Tris-Glycine gradient gel. Proteins were identified by SimplyBlue Safestain. Figure 6 shows that the major eluted protein corresponds to the size expected (68.5 kDa) for an MBP(A313V)-PMΔSal fusion protein.

Table 1. Yield of MBP-Klenow for wild-type and modified MBP's

plasmid	MBP	Yield
pIH1062	WT	0.7 mgs
pIH1643	A313V	2 mgs
pIH1644	S146T	0.7 mgs

Table 2. Yield of MBP-CBD for wild-type and MBP(A313V)

plasmid	MBP	Yield
pMB50	WT	13.2 mgs
pIH1660	A313V	21.8 mgs

Table 3. Solubility of MBP fusion proteins

Plasmid	Fusion protein	% Soluble
pIH1616	MBP-DHFR	36%
pIH1617	MBP(A313V)-DHFR	67%
pIH1618	MBP(S146T)-DHFR	67%
pIH1646	MBP(DM)-DHFR	25%
pIH1625	MBP-GAPDH	42%
pIH1626	MBP(A313V)-GAPDH	59%
pIH1627	MBP(S146T)-GAPDH	53%
pIH1645	MBP(DM)-GAPDH	17%

What is claimed is:

1. A modified maltose-binding protein (MBP) comprising: an MBP amino acid sequence having a mutation wherein the mutation causes the modified MBP when fused to a protein to have an increase in (i) binding affinity for a maltodextrin substrate; and (ii) solubility when the protein has limited solubility compared to the wildtype MBP when fused to the protein.

2. A modified MBP according to claim 1, wherein the mutation is located in the hinge region between helices XI and XII or in a region within 10 Å of A313 of the protein.

3. A modified MBP according to claim 1, wherein the mutation is located in the C domain or in a region within 10 Å of S146, at the beginning of β-sheet F of the protein.

4. A modified MBP according to claim 1, wherein the mutation is A313V.

5. A modified MBP according to claim 1, wherein the mutation is S146T.

6. A modified MBP according to claim 1, fused to a target protein to form a fusion protein.

7. A modified MBP according to claim 6, wherein the fusion protein has a solubility that is greater than the solubility of an unmodified MBP protein fused to the target protein.

8. A modified MBP according to claim 1, comprising SEQ ID NO:4 or SEQ ID NO:6.

9. An isolated DNA encoding a modified MBP according to claim 1, comprising: SEQ ID NO:3 or SEQ ID NO:5.

10. A vector comprising a DNA according to claim 9.

11. A host cell transformed with a vector according to claim 10.

12. A host cell according to claim 11, wherein the transformed host cell is a transformed *E. coli* host cell.

13. A host cell according to claim 11, wherein the transformed host cell is a transformed *Kluyveromyces* host cell.

14. A host cell according to claim 11, wherein the vector is pKlMF2.

15. A method of purifying a protein comprising:

expressing in a host cell, a fusion protein comprising a modified MBP according to claim 1 and a target protein;

permitting the modified MBP fusion protein to bind to a matrix; and

eluting the fusion protein from the matrix in a selected buffer to obtain the purified protein.

16. A method according to claim 15, wherein the matrix is a polysaccharide.

17. A method according to claim 15, wherein the carbohydrate is a maltodextrin.

18. A method according to claim 15, wherein the modified MBP has a mutation selected from A313V and S146T.

19. A method according to claim 15, wherein the modified MBP comprises SEQ ID NO:4.

20. A method according to claim 15, wherein the modified MBP comprises SEQ ID NO:6.

21. A method according to claim 15, wherein the modified MBP is encoded by SEQ ID NO:3 or SEQ ID NO:5.

22. A method according to claim 15, wherein the host cell is *Kluyveromyces* or *E. coli*.

23. A method for solubilizing a protein, comprising: expressing a modified MBP fused to a target protein according to claim 1 so that *in vivo* the fusion protein is solubilized to an extent greater than can be observed in the absence of the mutant MBP and to an extent greater than observed using a non-mutated MBP.

24. A method according to claim 23, wherein the modified MBP has a A313V mutation or an S146T mutation.

25. A method according to claim 24, wherein the modified MBP has an A313V mutation and an S146T mutation.