Engineering of Maltose-Binding Protein to Employ a Poly-Arginine Tag and Improve Protein Purification

Maltose binding protein (MBP), a solubility-enhancing fusion protein, has demonstrated great success in the application of protein purification when fused to a protein of interest. It has already been engineered to include a poly-histidine tag to allow for purification via Immobilized Metal Affinity Chromatography (IMAC). I have engineered this protein to include a poly-arginine tag, which would promote a cation-exchange purification step while maintaining the solubility-enhancing property of the original MBP. However, the inclusion of the poly-arginine tag did not significantly impact the elution profile of MBP during cation-exchange chromatography. This lack of purification suggests that the interactions between the positively charged arginine residues and the negatively charged compounds on the resin were insufficient to affect the salt dependent binding of MBP.

Abby Wachinski - biochemistry
Author Name/Major

Abby Wachinski
Author Signature

Brian G. Fox - biochemistry

Brian G. Fox
Mentor Signature

5-11-07
Date

COVER SHEET

TITLE: _ Engineering of Maltose-Binding Protein to Employ a Poly-Arginine Tag and Improve Protein Purification__

AUTHOR'S NAME:_____Abby Wochinski _____

MAJOR: _____Biochemistry_____

DEPARTMENT: _____Biochemistry_____

MENTOR: _____Professor Brian G. Fox_____

DEPARTMENT:_____Biochemistry_____

MENTOR(2): _____

DEPARTMENT(2):_____

YEAR: _____Senior_____

(The following statement must be included if you want your paper included in the library's electronic repository.)

The author hereby grants to University of Wisconsin-Madison the permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part in any medium now known or hereafter created.

Engineering of Maltose-Binding Protein to Employ a Poly-Arginine Tag and Improve Protein Purification

Abby Wochinski, Paul Blommel, and Brian G. Fox

Fox Lab, New Biochemistry Addition
University of Wisconsin
433 Babcock Dr.
Madison, WI 53706

Abstract

Maltose binding protein (MBP), a solubility-enhancing fusion protein, has demonstrated great success in the application of protein purification when fused to a protein of interest. It has already been engineered to include a poly-histidine tag to allow for purification via Immobilized Metal Affinity Chromatography (IMAC). I have engineered this protein to include a poly-arginine tag, which would promote a cation-exchange purification step while maintaining the solubility-enhancing property of the original MBP. However, the inclusion of the poly-arginine tag did not significantly impact the elution profile of MBP during cation-exchange chromatography. This lack of purification suggests that the interactions between the positively charged arginine residues and the negatively charged compounds on the resin were insufficient to affect the salt dependent binding of MBP.

Introduction

The Center for Eukaryotic Structural Genomics (CESG) concentrates on the determination and analysis of Eukaryotic protein structures. As part of the Protein Structure Initiative, its major goals are to solve the structures of 10,000 proteins over the course of ten years and make the three-dimensional structures of proteins easily obtainable from understanding their corresponding DNA sequences (Protein Structure Initiative, 2006). By analyzing the structure of specific proteins, we can better understand how they function and use this information to explore various issues of biological interest. Because the main objective of CESG focuses on this structural analysis, protein purification is incredibly important. Before any protein can be analyzed, the open reading frame (ORF) coding for it must be cloned, the protein expressed, and

other contaminating proteins removed. In CESG, cloning is done using the Gateway (Wrobel, R. and Zhao, Q. 2004) or Flexi-vector (Blommel, et al. 2005) systems. The protein is over-expressed in *Escherichia coli*, and purification takes place using Immobilized Metal Affinity Chromatography (IMAC). Figure 1 depicts an example of this method of purification. Finally, this purified product is analyzed via crystallography or Nuclear Magnetic Resonance, among other methods.

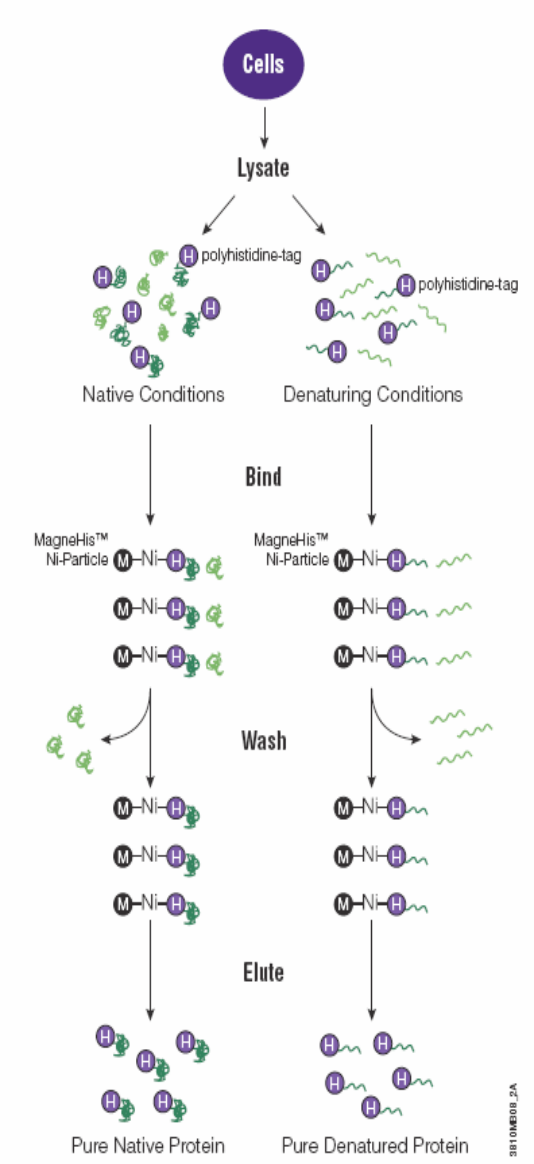


Figure 1: Diagram of Immobilized Metal Affinity Chromatography using a histidine tag and a nickel ion column (Promega Corporation, 2007).

CESG employs several fusion proteins to promote solubility and increase affinity for the charged resins as part of IMAC. Previous research has shown that Maltose Binding Protein (MBP) has a positive effect in assisting the folding of fusion proteins (Planson, A. *et al.* 2003). Through its interaction with other proteins, MBP often induces conformational changes, increasing the overall solubility of the protein of interest. MBP has also demonstrated to be extremely resilient to large amounts of change. In one instance, a sizeable portion of the MBP molecule was removed and replaced with an antibody binding epitope, confirming its ability to

withstand significant amounts of alteration without inhibiting function (Rodseth, L. *et al*. 1990).

The addition of a six or eight histidine tag to the amino or carboxyl terminus has been shown to augment the practical application of MBP by allowing for the use of IMAC protein purification. This type of chromatography, which utilizes immobilized resins, takes advantage of this increased affinity and uses it to aid in the isolation, purification, and identification of specific proteins (Yang. *et al.* 2002). In this case, poly-histidine and MBP tags are fused to the carboxyl or amino terminus of the protein of interest. The histidine residues interact with the positively charged nickel ions bound to the column resin (Figure 2). These interactions attach the tagged protein to the nickel matrix, while impurities continue to flow through the column and are eventually rinsed away. Eventually, the introduction of a highly concentrated base or imidazole solution to the column breaks this bond, allowing the purified protein of interest to be collected.

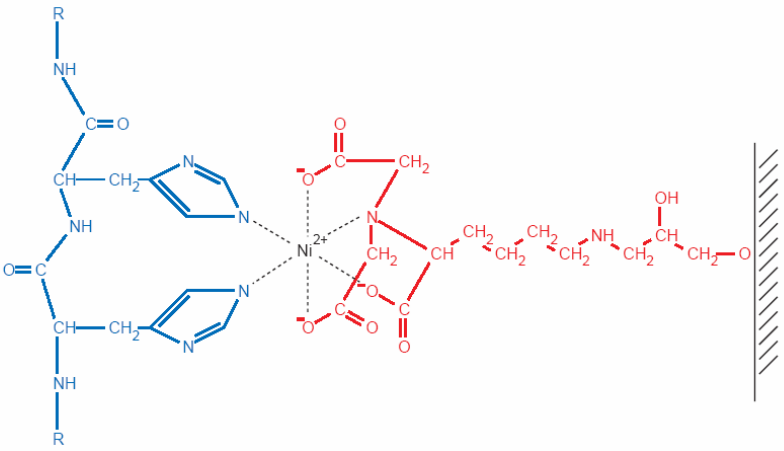


Figure 2: Schematic of the interaction between neighboring residues in the histidine tag and the nickel column (Promega Corporation 2007). The same concept can be applied to the poly-arginine tag, except the interacting residues are arginines, and the anion used for the column can be sulfate or another negatively charged group.

I engineered a protein expression vector that would allow for a cation exchange protein purification step, while maintaining the solubility enhancing ability of the original MBP. The histidine tag on MBP currently interacts with metal ions on a resin (IMAC). However, this new DNA sequence will also contain nucleotides encoding for a poly-

arginine tag made up of several positively charged arginine residues, which would promote binding to a cation-exchange resin. The vector was tested using Green Fluorescent Protein (GFP) to allow for easy analysis using fluorescence measurements. Figure 3 displays two potential mutation schematics.

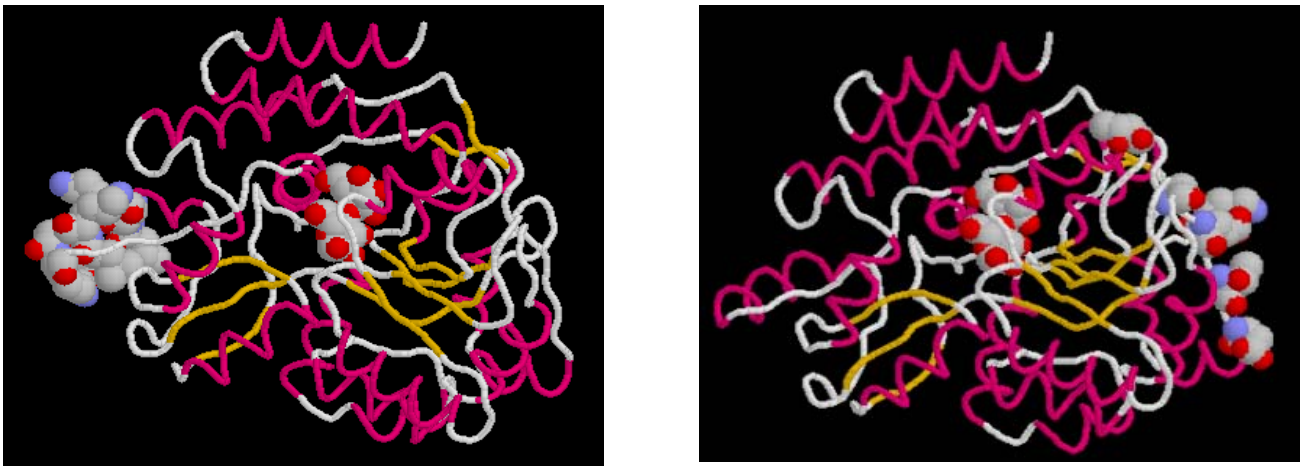


Figure 3: Schematic of potential poly-arginine mutations on MBP. The left image represents the addition of six, sequential arginine residues. The right image depicts six individual point mutations that all project toward the same face of the molecule. Structures were determined from the Protein Data Bank (Spurlino & Quiocho, 2007).

Through the exploitation of affinity chromatography, I will be able to separate these proteins based on their affinities for the negatively charged resin. Normally, purification steps must be separately designed and conducted, taking into consideration the physical and chemical properties of the individual protein. This new mutation would augment the current purification procedures by allowing for another general purification protocol that can be applied to a wide range of proteins. If successful, this protein allow for two general purification steps, utilizing both Immobilized Metal Affinity Chromatography (IMAC) and cation-exchange.

Materials and Methods

Introduction of Mutations via PCR

Using the known sequence for MBP, I designed forward and reverse primers to induce each mutation. I attempted two different approaches, one with a linear mutation and one with several point mutations. See Appendix A for specific primer sequences. Through the use of a two-step polymerase chain reaction (PCR) and DNA polymerase, I annealed these primers to the appropriate sections of the parent strand of the MBP DNA for the vector pVP27-GFP. This PCR program included a 55ºC annealing temperature, a 70ºC extension time for one minute, and a 95ºC denaturation stage. During this process, nucleotides coding for several amino acids were removed and replaced with nucleotides encoding for arginine sequences. Then, the DNA products from this PCR were analyzed via an agarose gel in order to confirm the size of the amplified DNA product. A second PCR procedure was used to piece together the mutated segments.

Restriction Enzyme Digest

I digested the mutated MBP DNA and the expression vector using restriction enzymes *Nsi I* and *Pac I*. The desired DNA fragment and digested expression vector were purified via agarose gel electrophoresis and then extracted using a Qiagen Spin Column Purification Kit.

Ligation and Transformation Reactions

These two fragments were then ligated together using T4 DNA ligase in a 20 µl reaction according to manufacturer specifications and incubated overnight at 4ºC. The products of these ligation reactions were transformed into competent *Escherichia coli* cells on LB media plates containing ampicilin.

Competent Cells and Sequencing

The short growth period of *E. coli* makes them a convenient host for the over-expression of recombinant proteins. Throughout this procedure, several strains of competent *E. coli* cells were used, including Top-10 and JM109. These competent cells have a high success rate in the uptake of new plasmids/DNA. Colonies were picked and grown up overnight in LB media containing ampicilin. After extracting the DNA from the cells via mini-prep using the Qiagen kit, I performed sequencing reactions on each cell using *realMBPmidFor* and *PQEFor*. This PCR reaction included a 58ºC annealing temperature, a 72ºC extension time for ten minutes, and a 95ºC denaturation stage**.** All DNA sequences were determined at the Biotchnology/Genetics building. I compared the obtained sequence data to the original open reading frame for MBP and ascertained the presence of the desired mutations.

Small-Scale Expression

After verifying the presence of the desired mutations, control and mutant plasmids were then transformed into B834-pRARE2 cells to be tested for protein expression. Starter cultures were made up of MDAG5 media and incubated overnight in a shaker at 37°C. Cells from the starter cultures were transferred into auto-induction media and incubated overnight in a shaker at 27°C.

Solubility Testing

I established the solubility of the fusion proteins by screening small amounts of the cultured cell paste. After lysing and sonicating the cells, I collected samples of total protein and samples of the soluble fraction. These fractions were run on an SDS-page gel to compare the amount of MBP in the total versus soluble portions.

Large-Scale Expression

Control and mutant large scale starter cultures were also made up of MDAG5 media and were incubated overnight in a shaker at 37°C. Cultures were inoculated in 500 ml auto-induction media and incubated overnight in a shaker at 27°C ("Protocols" 2004).

Ion-Exchange Chromatography

Cells from the large-scale cultures were resuspended, lysed and sonicated using the same methods as the small-scale cultures. A portion of the soluble fraction was then analyzed via a standard Immobilized Metal Affinity Chromatography (IMAC) bump elution ("Protocols" 2004). Fractions were collected using an AKTA Prime sampler and run on an SDS-page gel to demonstrate the purification made possible by the poly-histidine tag. I also took fluorescence readings of each fraction using a Tecan Ultra 384. The fraction containing the majority of the over-expressed MBP, as indicated by fluorescence measurements, then underwent cation-exchange chromatography using an S sepharose fast flow column. Fractions collected from this portion of the experiment were also analyzed via SDS-page gel and fluorescence.

Results

The procedure outlined for this project was shown to be successful for the creation of both the initial linear arginine mutation, as well as the point mutations. According to the sequence data, all mutations were correctly inserted as designed. However, the MBP-poly-R linear mutation introduced to MBP severely reduced the solubility of the fusion proteins (Figure 4). The over-expressed MBP is indicated by the intense band present in each lane. Even though there was some soluble protein present, the majority of the product was insoluble, as indicated by the decrease in band intensity.

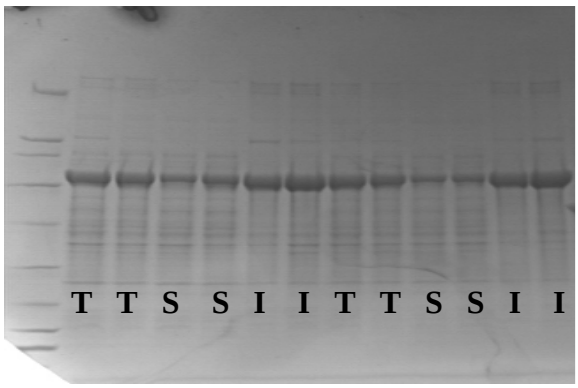
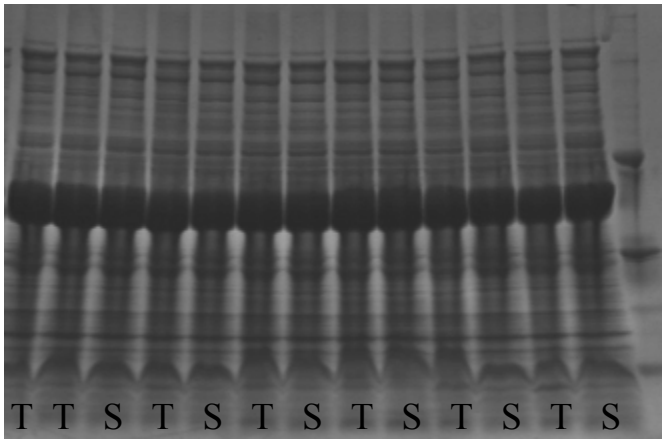


Figure 4: SDS-page gel depicting the results from the solubility screening of the MBP-poly-R linear mutation. T represents total protein; S represents soluble fractions; I represents insoluble fractions. The insoluble band intensity is much stronger than that of the soluble fraction, indicating a loss of protein solubility.

The analysis of the MBP-poly-R comprised of point mutations, I found that there was minimal effect on the tertiary structure of MBP, as protein solubility was retained (Figure 5). Again, the over-expressed MBP is suggested by the intense band present in each lane. The soluble fraction band intensity is just as strong as that of the total protein, indicating a retention of protein solubility.



F*igure 5*: SDS-page gel depicting the results from the solubility screening of the MBP-poly-R comprised of 6 point mutations. T represents total protein; S represents soluble fractions.

From the fluorescence measurements, and the SDS-page analysis, it is clear that the effectiveness of the IMAC purification procedure was retained in the presence of the poly-arginine mutations. The majority of the MBP staying on the column during washing and was the dominant protein found in the eluted fractions (Figure 6).

Figure 6: SDS-page-gel depicting the compositions of the various fractions collected during the IMAC bump elution. S represents total soluble fraction; F represents flow-through fractions; W represents wash fractions; E represents eluted fractions.

After the cation-exchange chromatography using the S sepharose fast flow column, I did not see considerable purification using the cation-exchange method (Figure 7). Fluorescence measurements were approximately the same for the control and mutated MBP fractions. For both the control protein and the mutated protein, the majority of the MBP came off the column during wash steps, and very little stayed on long enough for elution.

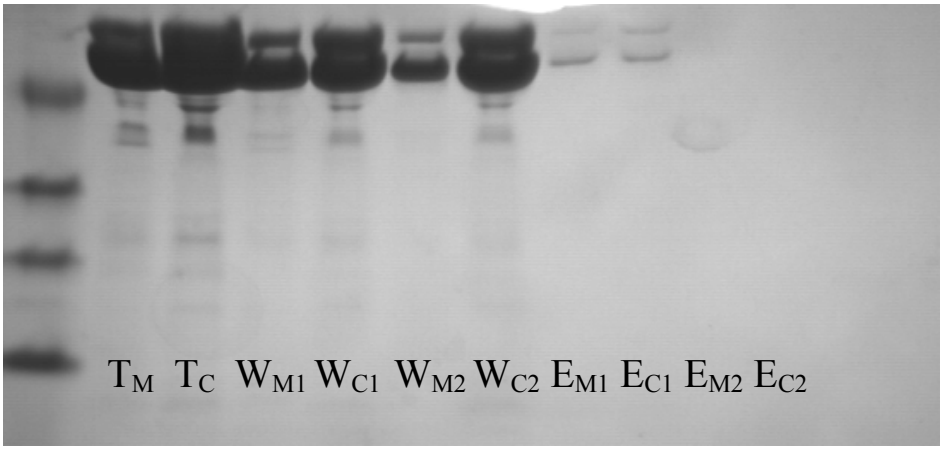


Figure 7: SDS-page gel depicting the compositions of the various fractions collected during the cation-exchange elution. T represents the total fraction eluted from the IMAC purification step; W represents the wash fractions; F represents flow-through fractions; E represents eluted fractions. This gel begins with the marker and alternates between mutant (M) and control (C) fractions.

Discussion

During the first attempt at the mutation, the designed arginine tag was made up of a linear sequence of six arginine residues. However, the number of amino acids removed from the original sequence was not equal to the number of arginines inserted. Because the residues removed were also a significant part of the IV α-helix, it is very likely that this mutation altered the tertiary structure of the protein, ultimately causing the significant decrease in the solubility of maltose binding protein (Spurlino & Quiocho, 2007). Despite previous evidence showing that mutations at this position of MBP did not significantly alter MBP structure of function (Rodseth, L. *et al*. 1990), our attempt using this six, consecutive arginine residues was not tolerated, as indicated by the decrease in protein solubility.

From these results, the mutation schematic was redesigned to place the arginine residues in locations that would minimize the influence of the mutagenesis on the tertiary structure of the protein. As demonstrated by the retention of the fusion protein solubility, performing six independent point mutations had seemingly less of an impact on the physical structure of MBP. Even though the point mutations were not necessarily attached to sequential portions of the protein, the amino acid residues all projected out toward one face of the molecule. Based on the fractions collected via IMAC elution, MBP also retained a functional poly-histidine tag, allowing for protein purification by means of this method. However, cation-exchange chromatography was not significantly enhanced through the incorporation of these mutations. It is possible that the arginine residues were placed too far apart to create a strong interaction and allow for the desired

purification. Perhaps a linear tag is necessary near the N- or C-terminus or in a location similar to that of the histidine tag placement.

Appendix A

Linear Mutations

Forward Primer:

CCTGGGAAGAGATCCCGCGTCGCCGTCGCCGTCGCGGTAAGAGCGCGCTGATGTTC

Reverse Primer:

GAACATCAGCGCGCTCTTACCGCGACGGCGACGGCGACGCGGGATCTCTTCCCAGG

Point Mutations

Mutation 1: Forward Primer: GTTGCGGCACGTGGCGATGG

Reverse Primer: CCATCGCCACGTGCCGCAAC

Mutation 2: Forward Primer: AAATCCGCCCGCGCAAAGCT

Reverse Primer: AGCTTTGCGCGGGCGGATTT

Mutation 3: Forward Primer: GTATTAACCGCGCCAGTCCG

Reverse Primer: CGGACTGGCGCGGTTAATAC

Acknowledgements

I want to thank Professor Brian Fox for giving me the opportunity to spend four years doing research in his lab. The experiences I gained there were invaluable. A big thanks to Paul Blommel, who played a key role in overseeing my project and helping me with data analysis. He was the inspiration behind this project and really helped me grasp a fundamental understanding of the biochemical concepts being applied. I would also like to give a huge thanks to all the graduate and undergraduate students in the Fox lab for answering my questions and making my lab experience everything that it was.

I also want to thank the CALS Honors Program for letting me be part of undergraduate research on one of the most esteemed campuses in the country and the CALS Honors Committee for helping fund my project.

References

Blommel, P.G., Martin, P.A., Wrobel, R.L., Steffen, E., Fox, B.G. November 2005. "High-efficiency single-step production of expression plasmids from cDNA clones using the Flexi Vector cloning system." *Protein Expression and Purification*. 47: 562-70.

Planson, A.G., Guijarro, J.I., Goldberg, M.E., and Chaffotte, A.F. November 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 expressed in Escherichia coli." *Biochemistry* 42(45):13202-11.

Promega Corporation. 2007. "Protein Purification and Analysis Protocols Guide." http://www.promega.com/paguide/chap11.pdf

Protein Structure Initiative. July 2006. *Mission Statement*. http://www.nigms.nih.gov/Initiatives/PSI/Background/MissionStatement.htm

"Protocols." *Center for Eukaryotic Structural Genomics* (CESG). August 2004. http://www.uwstructuralgenomics.org/protocols.htm.

Rodseth, L., Martineau, P., Duplay, P., Hofnung, M., and Quiocho, F. June 1990. "Crystallization of Genetically Engineered Active Maltose-binding Proteins, Including an Immunogenic Viral Epitope Insertion." *Journal of Molecular Biology*. 214(4): 607 – 611.

Spurlino, J. and Quiocho, F. April 2007. "Maltodextrin Binding Protein with Bound Maltose." *RCSB Protein Data Bank*. http://www.rcsb.org/pdb/explore/explore.do?structureId=1ANF

Wrobel, R. and Zhao, Q. November 2004. "Cloning into Entry Vectors." *Center for Eukaryotic Structural Genomics* (CESG). http://www.uwstructuralgenomics.org/protocols.htm.

Yang, H., Viera, C., Fischer, J., and Etzel, M. 2002. "Purification of a Large Protein Using Ion-Exchange Membranes." *Industrial and Engineering Chemistry Research* 41(6): 1597 – 1602.