Single amino acid substitutions on the surface of *Escherichia coli* maltose-binding protein can have a profound impact on the solubility of fusion proteins

JEFFREY D. FOX, RACHEL B. KAPUST, AND DAVID S. WAUGH

Protein Engineering Section, Macromolecular Crystallography Laboratory, National Cancer Institute at Frederick, Frederick, Maryland 21702-1201, USA

(RECEIVED October 24, 2000; FINAL REVISION December 19, 2000; ACCEPTED December 19, 2000)

Abstract

Proteins are commonly fused to *Escherichia coli* maltose-binding protein (MBP) to enhance their yield and facilitate their purification. In addition, the stability and solubility of a passenger protein can often be improved by fusing it to MBP. In a previous comparison with two other highly soluble fusion partners, MBP was decidedly superior at promoting the solubility of a range of aggregation-prone proteins. To explain this observation, we proposed that MBP could function as a general molecular chaperone in the context of a fusion protein by binding to aggregation-prone folding intermediates of passenger proteins and preventing their self-association. The ligand-binding cleft in MBP was considered a likely site for peptide binding because of its hydrophobic nature. We tested this hypothesis by systematically replacing hydrophobic amino acid side chains in and around the cleft with glutamic acid. None of these mutations affected the yield or solubility of MBP in its unfused state. Each MBP was then tested for its ability to promote solubility when fused to three passenger proteins: green fluorescent protein, p16, and E6. Mutations within the maltose-binding cleft (W62E, A63E, Y155E, W230E, and W340E) had little or no effect on the solubility of the fusion proteins. In contrast, three mutations near one end of the cleft (W232E, Y242E, and I317E) dramatically reduced the solubility of the same fusion proteins. The mutations with the most profound effect on solubility were shown to reduce the global stability of MBP.

Keywords: Fusion protein; inclusion bodies; maltose-binding protein; protein folding; solubility; aggregation

Escherichia coli maltose-binding protein (MBP) is often used as a fusion partner for recombinant protein expression, primarily because its affinity for maltodextrins allows for facile purification of fusion proteins by amylose affinity chromatography (Riggs 2000). Proteins such as MBP, glutathione *S*-transferase (GST), ubiquitin, and others have also been observed to increase the yield of their fusion partners in many cases (Butt et al. 1989; Riggs 2000). In addition, MBP is often able to promote the solubility of polypeptides to which it is fused (Pryor and Leiting 1997; Kapust and Waugh 1999). A recent study demonstrated that not all highly soluble proteins are equally effective as solubilizing agents (Kapust and Waugh 1999). In those experiments, MBP consistently enhanced the solubility of aggregation-prone proteins, whereas two other commonly used, highly soluble proteins did not.

Although many investigators have exploited the solubility-enhancing property of MBP, the mechanism by which MBP increases the solubility of its passenger proteins is not understood. One possibility, which seems to be consistent with all of the available experimental evidence (see below), is that MBP possesses chaperone-like qualities. These properties may allow MBP to bind reversibly to folding intermediates of its fusion partners, termed passenger proteins,

Reprint requests to: Dr. David S. Waugh, Protein Engineering Section, Macromolecular Crystallography Laboratory, National Cancer Institute at Frederick, P.O. Box B, Frederick, Maryland 21702-1201, USA; e-mail: waughd@ncifcrf.gov; fax: (301) 846-7148.

Article and publication are at www.proteinscience.org/cgi/doi/10.1110/ps.45201.

and temporarily sequester them in a conformation that prevents their self-association and aggregation. According to this model (see Fig. 4 in Kapust and Waugh 1999), the tether that joins the two interacting partners facilitates iterative cycles of binding and release. By preventing the accumulation of a high concentration of unfolded, unsequestered passenger protein, this process may eventually steer the passenger protein toward its native conformation, provided that it is capable of folding spontaneously in the environment of the bacterial cytoplasm. In fact, proper folding of aggregation-prone passenger proteins fused to MBP has been reported in several cases (Mottershead et al. 1996; Rao and Bodley 1996; Thomas et al. 1996; Kapust and Waugh 1999). Alternatively, kinetically trapped folding intermediates or proteins that aggregate in their native state may be maintained in a soluble form as MBP fusions as a result of the formation of a stable sequestered intermediate. Examples of this behavior have also been noted (Louis et al. 1991; Lee et al. 1993; Saavedra-Alanis et al. 1994; Kishore et al. 1998; Sachdev and Chirgwin 1998b).

The way in which chaperones recognize and interact with their substrates has been the subject of intense investigation. The 70-kD heat shock protein (Hsp70) family of chaperones, exemplified by *E. coli* DnaK, interacts with hydrophobic regions in nascent polypeptides and prevents them from aggregating before the acquisition of native structure (for review, see Feldman and Frydman 2000). The peptide-binding site of DnaK consists of a deep, hydrophobic pocket with a particular affinity for leucine side chains (Zhu et al. 1996; Wang et al. 1998). Substrate recognition by the Hsp60 (or chaperonin) family of chaperones has also been studied in considerable detail (for recent reviews, see Brazil and Horowitz 1999; Feltham and Gierasch 2000). The best-studied member of this family is *E. coli* GroEL, a large multisubunit, ring-shaped protein complex that actively promotes the folding of many proteins through an ATP-dependent cycle (for review, see Ranson et al. 1998; Sigler et al. 1998; Wang and Weissman 1999). Binding to the apical domain in the GroEL complex sequesters the misfolded protein (or folding intermediate) and may subject it to physical unfolding forces as the nucleotide triphosphate and the GroES cochaperonin bind to GroEL and induce conformational changes that release the peptide into the internal cavity of GroEL (Fenton et al. 1994; Shtilerman et al. 1999). After iterative cycles of binding and release, the polypeptide eventually reaches its native conformation. The peptide-binding region of GroEL, which is located in a cleft between two α-helices on the surface of the apical domain, is mostly hydrophobic in nature and capable of binding a wide range of peptide conformations (Buckle et al. 1997; Chatellier et al. 1999; Chen and Sigler 1999; Kobayashi et al. 1999; Tanaka and Fersht 1999). In fact, by itself the apical domain of GroEL, also called the minichaperone, exhibits peptide binding and limited chaperone activities (Zahn et al. 1996; Chatellier et al. 1998).

There is no reason to believe that MBP actively participates in the folding of its fusion partners, but a more passive, antiaggregative role in peptide binding and sequestration, as seen with the GroEL minichaperone and Hsp70, seems plausible. What would the characteristics of such an interaction site be, if it exists? Should one expect hydrophobic character, in accord with peptide-binding sites on chaperones? MBP has an unusually large number of exposed hydrophobic zones on its surface. Moreover, MBP possesses a natural protein-binding site that it uses to interact with other proteins involved in maltose signaling and chemotaxis (Martineau et al. 1990a; Boos and Shuman 1998). In close proximity to this site is a large hydrophobic cleft (Fig. 1) that specifically binds maltodextrins, causing a conformational change that closes the cleft and activates the protein–protein interaction site on the surface (Spurlino et al. 1991; Quiocho et al. 1997).

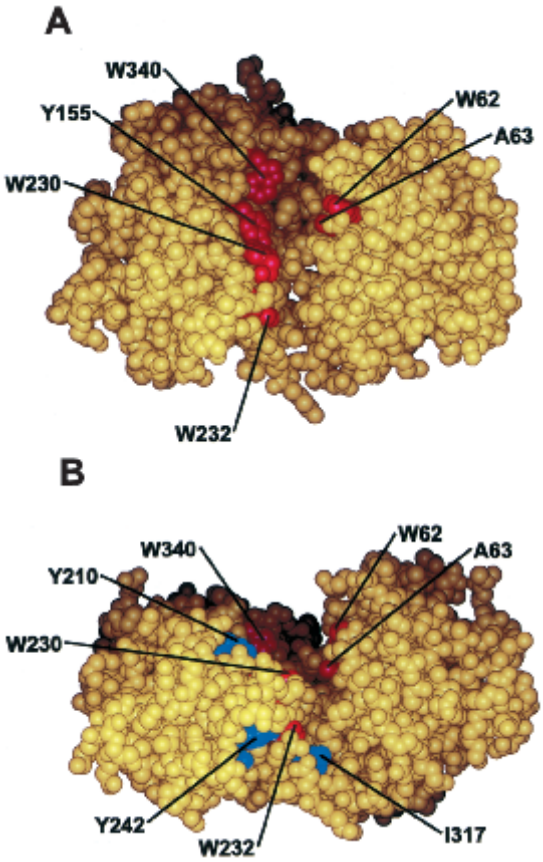

Fig. 1. Space-filling representations of the unliganded *E. coli* MBP crystal structure (PDB code 1OMP) (Sharff et al. 1992). Amino acids that were altered by site-directed mutagenesis are highlighted. (*A*) The amino acids targeted initially are shown in red. (*B*) The view in *A* was rotated back and to the left to show the additional amino acids (in blue) that were targeted for mutagenesis.

There are several reasons to believe that the hydrophobic maltodextrin-binding cleft may be the site where fused polypeptides interact with MBP. Foremost among these is that authentic molecular chaperones (e.g., GroEL and DnaK) use hydrophobic clefts to bind their targets (Buckle et al. 1997; Chatellier et al. 1999; Chen and Sigler 1999; Kobayashi et al. 1999; Tanaka and Fersht 1999). Second, the presence of this cleft distinguishes MBP from other highly soluble proteins that do not function effectively as solubilizing agents (Kapust and Waugh 1999). Third, there is considerable conformational flexibility associated with the cleft (Quiocho et al. 1997); therefore, it could adjust its shape to accommodate a variety of peptides. Finally, if fusion partners interact with MBP by occupying the maltodextrin-binding cleft, this could explain why some MBP fusion proteins do not bind efficiently to amylose resin (Pryor and Leiting 1997).

Fenton and co-workers (1994) successfully used site-directed mutagenesis to identify functionally important amino acid side chains in and around the peptide-binding cleft of GroEL. By systematically changing individual hydrophobic amino acids to bulky charged side chains, the importance of these residues for peptide binding by GroEL was demonstrated (Fenton et al. 1994). Here, we have adopted a similar approach to examine whether hydrophobic side chains in the maltodextrin-binding cleft of MBP (Fig. 1A) are important for its ability to mediate the solubilizing effect. The mutant MBPs were assayed for their ability to enhance the solubility of three passenger proteins that are themselves poorly soluble when overproduced in *E. coli*. Unexpectedly, all but one of the mutations in the cleft had little or no effect on the solubility of the fusion proteins. In light of these results, additional mutations were made near but not directly within the cleft to further define a region on the surface of MBP that is critical for enhancing the solubility of its fusion partners. All of the mutant MBPs were also characterized by equilibrium denaturation methods to assess the effect of individual mutations on the global stability of MBP.

Results

Locations of altered amino acids

To test the hypothesis that the maltodextrin-binding cleft in MBP is involved in promoting the solubility of its fusion partners, we initially targeted six amino acids within the cleft (W62, A63, Y155, W230, W232, and W340) for mutagenesis (Fig. 1A). These positions were chosen because of their aromatic/hydrophobic character and because they span the length and cover both walls of the cleft. Several of these amino acids are integral to the maltodextrin-binding site, but mutations at these positions result in proteins with no significant structural defects, and they can be expressed and purified in soluble form (Martineau et al. 1990b). In accord with the GroEL study (Fenton et al. 1994), the native amino acid at each position was altered to glutamic acid by oligonucleotide-directed mutagenesis. In every case, the entire MBP open reading frame was sequenced to ensure that no unexpected changes were introduced as a consequence of the mutagenic procedure.

Solubility of mutant MBPs in the fused and unfused states

Each mutation was first analyzed in the context of native (unfused) MBP to determine whether the lesion affected the yield or solubility of MBP by itself. The results of these experiments are shown in Figure 2A. The electrophoretic mobility of MBP was altered slightly by some of the mutations, but there was no appreciable difference in yield or solubility between the wild-type and mutant MBPs.

The six cleft mutations were next tested for their impact on the solubility of three different fusion proteins. The passenger proteins used for these experiments were p16 (human cyclin-dependent kinase 4 inhibitor), E6 (human papillomavirus 18 oncoprotein), and GFP (*Aequorea victoria* green fluorescent protein). These proteins are insoluble when they are overproduced in *E. coli* at 37°C but can be solubilized very efficiently by fusing them to MBP (Kapust and Waugh 1999).

p16, a member of the INK4 protein family of tumor suppressors, is composed of four ankryrin (helix–turn–helix)

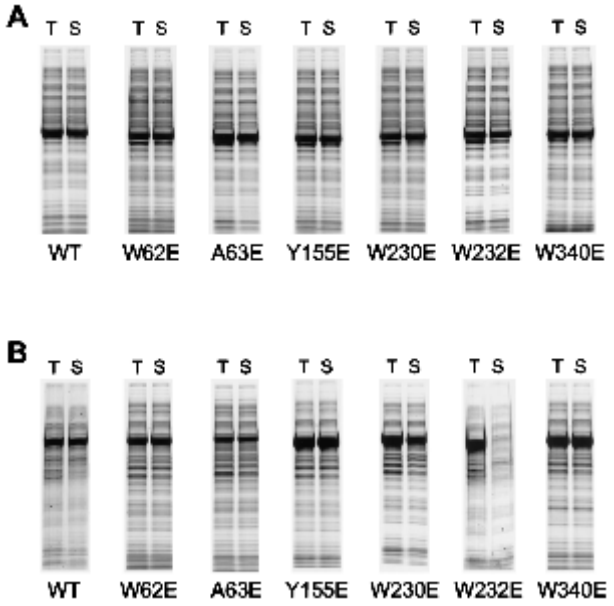

Fig. 2. Expression of MBPs and MBP fusion proteins. (*A*) Expression of unfused MBPs. Each protein was expressed at 37°C in an unfused form and analyzed by SDS-PAGE. (*B*) Expression of MBP-p16 fusion proteins. T, total cellular extract; S, soluble fraction.

repeats that stack together to form a pair of four-helix bundles (Russo et al. 1998). The p16 fold is rather unstable, both in thermodynamic and kinetic terms (Tang et al. 1999). The propensity of GFP to aggregate when expressed at elevated temperatures (e.g., 37°C) is well documented, and several studies have shown that specific mutations can dramatically improve the solubility of the protein under these circumstances (Cormack et al. 1996; Crameri et al. 1996; Siemering et al. 1996; Yang et al. 1996; Kimata et al. 1997). Once folded, however, GFP is exceptionally resistant to denaturation and can be concentrated to at least 30 mg/ml in solution without aggregating (Yang et al. 1996). Relatively little is known about the structure and folding of E6. It is a small nuclear protein that abrogates the activity of proteins whose normal function is to negatively regulate cell growth (Farthing and Vousden 1994). E6 contains two cysteine-rich, Zn-binding domains.

As shown in Figure 2B, the yields of the mutant MBP–p16 fusion proteins were substantial and at least as high as that of the wild-type MBP–p16 fusion protein. Only the MBP(W232E)–p16 fusion protein exhibited a striking solubility defect. With the possible exception of the W230E mutation, which appeared to have a slight impact on the solubility of the MBP–p16 fusion protein, the solubility of the other mutants was virtually the same as that of the wild-type fusion protein. Similar results were obtained when the mutant MBPs were fused to E6 and GFP (Fig. 3A). Thus, only the W232E mutation exhibited the solubility phenotype expected for a mutation that disrupts an interaction between MBP and its fusion partners: it had no effect on the solubility of MBP alone, but dramatically reduced the solubility of several MBP fusion proteins.

Additional site-directed mutants

One possible explanation for these results is that the interaction site overlaps but is not entirely coincident with the maltodextrin-binding cleft of MBP. The W232 position may lie at the edge of the interaction site. To test this hypothesis, we targeted two additional residues (Y242 and I317) for mutagenesis. These solvent-accessible side chains are situated outside of the cleft but adjacent to W232 on the surface of MBP (see Fig. 1B). Although there was no information in the literature to indicate whether these two side chains are critical for the proper folding of MBP, the yields and solubilities of unfused Y242E and I317E MBP were virtually identical to those of wild-type MBP and the other unfused MBP mutants that we characterized (data not shown).

When examined in the context of the p16, E6, and GFP fusion proteins, both the Y242E and I317E mutations exhibited severe solubility phenotypes (Fig. 3B). Like W232E, the Y242E and I317E substitutions dramatically reduced the solubility of all three fusion proteins tested. Therefore, this cluster of solvent-accessible side chains evidently is critical for the ability of MBP to promote the solubility of its fusion partners.

The tyrosyl side chain at position 210 of MBP, which is solvent-exposed and in the vicinity of the cleft (see Fig. 1B), was shown previously to be critical for the interaction of MBP with the MalFGK2 complex; replacement of Tyr210 with a variety of side chains strongly impaired maltose transport (Szmelcman et al. 1997). However, this tyrosyl side chain evidently is not important for the solubilizing activity of MBP, because altering Y210 to alanine had no effect on the solubility of the MBP fusions tested here (data not shown).

Mutations on the surface of MBP affect global stability

Although consistent with the interaction site hypothesis, our results could also be explained if the W232E, Y242E, and I317E substitutions retard the folding of MBP in the cytoplasm. This could affect the solubility of a fusion protein indirectly by reducing the proportion of MBP that is in a state capable of promoting solubility, thereby favoring the kinetically competing pathway of self-association and ag-

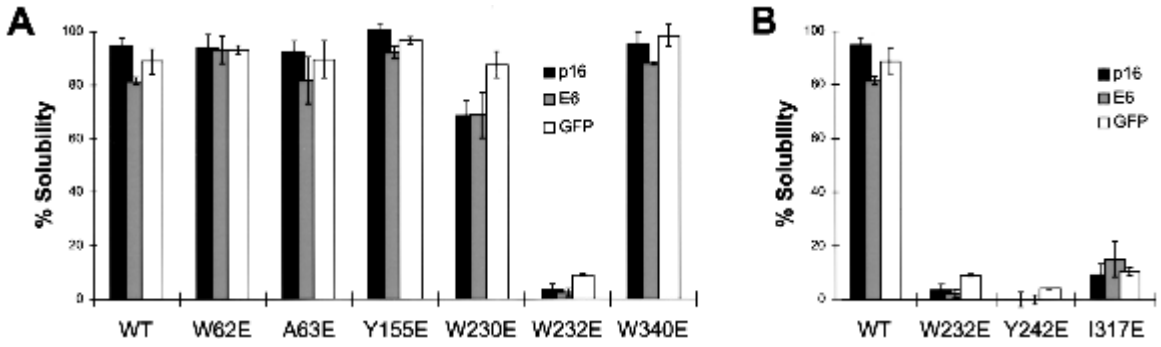

Fig. 3. Solubility of MBP fusion proteins. Fusions between each MBP mutant and three passenger proteins (p16, E6, and GFP) were expressed at 37°C and analyzed by SDS-PAGE. Solubility was estimated by laser scanning densitometry. (*A*) Cleft mutants. (*B*) Additional mutants.

gregation (see Kapust and Waugh 1999). Therefore, the impact of these mutations on solubility could be due to a global rather than local effect. To distinguish between these possibilities, we examined the influence of individual amino acid substitutions on the stability of MBP in equilibrium denaturation experiments monitored by circular dichroism (CD). Each mutant MBP (in the unfused state) was overproduced, purified to homogeneity, and then subjected to equilibrium denaturation at 25°C over a wide range of guanidine hydrochloride (Gd-HCl) concentrations. Measurement of ellipticity at 222 nm served as an indication of the degree of unfolding. Ellipticity data were plotted versus Gd-HCl concentration and fit to a two-state folding model using a nonlinear least-squares algorithm (Santoro and Bolen 1988; Pace and Scholtz 1997). This method of analysis yielded values for the Gd-HCl concentration at the unfolding transition midpoint as well as the free energy of unfolding (extrapolated to zero denaturant), $\Delta G^0_{H_2O}$. The data for each mutant are summarized in Table 1.

All of the mutant MBPs exhibited a sigmoidal, two-state unfolding transition. Wild-type MBP and most of the cleft mutants unfolded with a midpoint Gd-HCl concentration near 1 M, and a $\Delta G^0_{H_2O}$ in the range of 10–14 kcal/mole. These values agree favorably with previously published results for wild-type MBP (Liu et al. 1988; Diamond et al. 1995; Sheshadri et al. 1999). A few of the mutants appeared to show small increases in stability to chemical denaturation relative to wild-type MBP, as indicated by their higher $\Delta G^0_{H_2O}$ values. It is possible that the loss of surface-accessible hydrophobic residues is responsible for this increase in stability. As a group, the mutations that exhibited a significant solubility phenotype are less stable than wild-type MBP. The free energy of unfolding ranges between 5–8 kcal/mole for these mutants (Table 1). Thus, these data are consistent with the notion that the solubility defects we observed may arise from a change in the global stability of MBP rather than from the disruption of a general protein interaction site.

Table 1. *Equilibrium unfolding analysis of wild-type and mutant MBPs*

MBP	C_m (M Gd-HCl)	$\Delta G^\circ_{H_2O}$ (kCal/mol)
WT	1.08	12.9
W62E	1.05	13.3
A63E	1.07	14.7
Y155E	0.95	10.2
Y210A	1.09	14.0
W230E	0.87	9.8
W232E	0.57	5.4
Y242E	0.68	5.7
Y283D	0.82	6.6
I317E	0.76	7.1
W340E	0.89	10.6

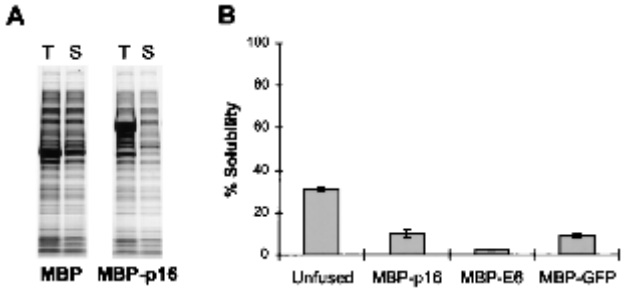

Fig. 4. Analysis of Y283D MBP and mutant fusion proteins. (*A*) Expression of unfused Y283D MBP and Y283D MBP–p16 at 37°C. (*B*) Solubility of Y283D MBP fusion proteins at 37°C.

To corroborate this conclusion, we investigated the solubility phenotype of one additional mutation. Replacing Tyr-283 with aspartic acid retards the folding of MBP in vitro and in vivo (Liu et al. 1988; Chun et al. 1993). Tyr-283 is situated far away from the cleft in MBP. In accord with previous observations (Liu et al. 1988; Chun et al. 1993), our results indicate that Y283D MBP is considerably less stable than the wild-type protein (Table 1). However, unlike the other mutations we studied, Y283D MBP exhibited markedly reduced solubility in its unfused state (Fig. 4). Not surprisingly, when fused to the three passenger proteins (p16, E6, and GFP), the Y283D mutation in MBP also had a severe impact on the solubility of the fusion proteins (Fig. 4). These data are in agreement with the notion that the rate of folding of MBP is crucial for its ability to promote solubility. Interestingly, Y283D MBP is more stable in equilibrium denaturation experiments than some of the other mutants with solubility phenotypes (W232E, Y242E), but its solubility defect seems to be the most drastic of all.

Discussion

The importance of fusion protein solubility

The ability of MBP to promote the solubility of its fusion partners may have important applications in the biotechnology industry, because insolubility of recombinant proteins produced in heterologous expression systems is a very common occurrence (Schein 1989). Although it is sometimes possible to convert aggregated material into properly folded, biologically active protein, this is an arduous, costly, and uncertain undertaking. Consequently, a general means of circumventing the formation of insoluble protein aggregates is potentially of great practical importance.

It is clear that MBP can routinely enhance the solubility of diverse fusion partners (Kapust and Waugh 1999), but it is less certain how often and how efficiently these proteins, once rendered soluble by fusion to MBP, fold into their native, biologically active conformations. Some insight into

this issue can probably be gained through a greater understanding of the mechanism of the solubilizing effect. Elucidating this mechanism may improve our ability to predict when the use of an MBP fusion will succeed for the production of a particular aggregation-prone protein.

Moreover, if the mechanism is known, then it may be possible to manipulate conditions to further promote the solubility and proper folding of a greater number of aggregation-prone proteins. Finally, determining the mechanism of the solubilizing effect may help to further our understanding of fundamental aspects of protein folding and aggregation in vivo.

The mechanism of the solubilizing effect

There are at least four possible models that we can invoke to explain the solubilizing effect of MBP. It should be noted that these models are not necessarily mutually exclusive. One possibility is that the fusion proteins form soluble, micelle-like, multiprotein structures with the hydrophilic MBP domains on the outside and the aggregated passenger proteins occupying the center. Although aggregates of a sort, these particles may not be large enough to be sedimented by centrifugation at low-to-moderate speed. Another possibility is that MBP acts as an anchor to restrict the motion of a slow-folding passenger protein, enabling it to fold in a more entropically favorable environment by reducing the number of possible conformations that can be sampled. A third possibility is that MBP acts as a "chaperone magnet" in vivo by channeling chaperones that normally interact with MBP, such as SecB (Randall et al. 1998), into productive associations with the attached passenger protein. It is also possible that MBP possesses an intrinsic molecular chaperone activity of its own that becomes manifest in the context of a fusion protein. According to this molecular–chaperone model, MBP enhances the solubility of its fusion partners by physically interacting with and sequestering them from self-association during the folding process (Kapust and Waugh 1999). MBP could also inhibit the aggregation of proteins in their native states by the same basic mechanism (i.e., intramolecular association and sequestration).

The micelle, entropic–anchor, and chaperone–magnet models seem inconsistent with at least some published experimental observations. Neither the micelle nor the entropic–anchor model can readily account for the observation that only a subset of highly soluble proteins, such as MBP, are effective solubilizing agents (Kapust and Waugh 1999). Why would not any highly soluble protein be able to form the proposed micelle-like structures with reasonable proficiency? Similarly, any soluble (and folded) fusion partner would be expected to exert a similar entropic effect on the folding of the attached protein. Furthermore, the micelle and chaperone–magnet models are difficult to reconcile with the fact that the solubility of MBP fusion proteins is influenced by the order in which the two domains are translated in vivo; fusion proteins composed of MBP and either procapthepsin D or pepsin were only soluble in cells when MBP was the amino-terminal domain (Sachdev and Chirgwin 1998a). It is unclear why fusion proteins in which the aggregation-prone passengers are fused to the amino terminus of MBP would not be able to form micellar structures, as both termini of MBP are exposed on the surface of the folded structure (Sharff et al. 1992) and are capable of being extended without impeding solubility (Tsao et al. 1996; Kapust and Waugh 1999). Nor is it clear why MBP would be less effective at recruiting chaperones when proteins are fused to its amino terminus than to its carboxyl terminus. The fact that the recovery of some purified proteins after in vitro refolding is enhanced by fusion to MBP (Sachdev and Chirgwin 1998b) and the observation that MBP inhibits the aggregation of certain proteins in vitro (Richarme and Caldas 1997) are also inconsistent with the chaperone–magnet model. Thus, none of these models can account for the existing body of experimental observations.

Chaperone-like activities of MBP

In contrast, all of these observations can be explained if MBP, and not just any soluble protein, can act as a general molecular chaperone in the context of a fusion protein (Kapust and Waugh 1999). According to this model, MBP possesses a site (or sites) with a natural propensity to bind folding intermediates of the attached protein or aggregation-prone proteins in their native states. Because the binding site (or sites) on MBP would not be available until after folding, this model accounts for the observation that the solubility of MBP fusion proteins is influenced by the order in which the two domains are translated in vivo.

A fundamental tenet of the molecular–chaperone model is that MBP engages in a physical interaction with its fusion partners, most likely by way of nonspecific hydrophobic contacts, and that this interaction serves to sequester aggregation-prone folding intermediates and prevent their self-association. If this is true, then it might be possible to destabilize these interactions by altering hydrophobic amino acids on the surface of MBP, which could have the effect of driving the fusion proteins into the kinetically competing, self-association pathway and lead to the formation of insoluble aggregates. This approach succeeded in identifying residues in GroEL that participate in peptide binding (Fenton et al. 1994). However, the results reported here do not support the notion that the maltodextrin-binding cleft in MBP is involved in peptide binding. On the contrary, the fact that none of the cleft mutations had any effect on the solubility of several different fusion proteins suggests that the cleft is not involved in peptide binding. Furthermore, the fact that the Y210A substitution also had no solubility phe-

notype suggests that MBP probably does not use its natural protein-binding site to interact with its fusion partners. Yet, because our results were negative, we cannot definitively rule out either of these possibilities.

On the other hand, these site-directed mutagenesis experiments led to the identification of a cluster of solvent-exposed residues that are absolutely critical for the solubilizing effect. Although they are widely dispersed in the primary sequence and reside in distinct elements of the secondary structure, these three side chains form a nearly contiguous patch on the surface of the folded protein (Fig. 1B). The existence of a solvent-exposed "hot spot" is consistent with the interaction-site hypothesis. Moreover, if this interpretation is correct, we can also conclude that MBP uses the same site to interact with different passenger proteins. However, the correlation we observed between the solubility phenotypes of these mutations and their influence on the global stability of MBP in equilibrium denaturation experiments complicates this interpretation. It is also possible that the W232, Y242, and I317 mutations exert their influence indirectly by retarding the folding of MBP, which, in turn, gives rise to the effects we observed on the solubility of fusion proteins. Our observations regarding destabilization of MBP by the mutations do not allow us to distinguish between these two possibilities.

The fact that these three mutations have a destabilizing effect on MBP does not necessarily contradict the chaperone model. Rather, it merely reinforces the idea that the relatively rapid formation of a properly folded MBP domain is crucial for its ability to promote solubility. Like the Y283D substitution (Chun et al. 1993), we suspect that the mutations we have characterized influence the solubility of MBP fusion proteins by retarding the rate of MBP folding in vivo. Preliminary kinetic refolding experiments support this view (L. Randall, pers. comm.). If we were observing an equilibrium effect, in which the unfolded state of the mutant MBPs was always significantly more populated than in the case of wild-type MBP, then one might expect a lower yield of the unfused mutant MBPs (due to proteolytic degradation) or the accumulation of the protein in inclusion bodies, but neither of these problems occur. Moreover, all of the mutant MBPs behave normally during purification, suggesting that the vast majority of the protein is properly folded. The stability of Y283D MBP, as gauged by equilibrium unfolding experiments, is actually somewhat greater than that of W232E and Y242E MBP, indicating that a rigorous correlation between the global stability and the ability to promote solubility cannot be maintained. Nevertheless, it seems clear that MBP mutants with impaired stability in equilibrium unfolding experiments tend to be poorer solubilizing agents. If, as we suspect, the W232E, W242E, and I317E mutations exert their effect on the solubility of fusion proteins by retarding the folding of MBP, then we have identified a new cluster of solvent-accessible residues that are important for the folding of MBP. Perhaps these residues were not identified in a genetic selection for mutations that retard the folding of MBP (Chun et al. 1993) because the mutagenic specificity of hydroxylamine would not have generated these lesions. Alternatively, unlike Y283D, their effects in vivo may be too subtle to be detected by this method. Further work will be necessary to determine exactly how these mutations influence the folding of MBP and the solubility of MBP fusion proteins.

Materials and methods

Plasmids

All of the plasmid expression vectors used to produce fused and unfused MBP were derived from pMAL-c2 (New England Biolabs, Inc., Beverly, MA). For the analysis of mutations in the context of native (unfused) MBP, pMAL-c2 was modified by PCR to create pDW533, a vector that produces the mature, native form of MBP. Vectors for expression of fusions between wild-type MBP and the three "passenger" proteins (p16, E6, and GFP) were described previously (Kapust and Waugh 1999). These vectors were named pDW385 (MBP-p16), pRK576 (MBP-E6), and pDW520 (MBP-GFP).

Mutagenesis

Single amino acid substitutions in MBP were constructed with the Transformer Site-Directed Mutagenesis Kit (Clontech, Inc., Palo Alto, CA), using pDW533 as the template. All mutants were confirmed by sequencing the entire MBP open reading frame. Appropriate restriction enzymes were used to excise each mutation from the pDW533 derivative and move it into pDW385, pRK576, or pDW520 to yield the mutant fusion constructs.

Protein expression, SDS-PAGE analysis, and densitometry

Protein expression was carried out as described previously (Kapust and Waugh 1999) except that the accessory plasmid pDC952 was not used in this work. Preparation of total and soluble intracellular protein samples, SDS-PAGE analysis, and densitometry of the stained gels were all performed essentially as described (Kapust and Waugh 1999).

Protein purification

Wild-type MBP and mutants that retained the ability to bind maltodextrins (i.e., W232E, Y242E, I317E, and Y283D) were purified by amylose affinity chromatography according to the instructions of the manufacturer (New England Biolabs), with minor modifications. MBP was eluted from the amylose resin with a solution of 20% (w/v) glucose rather than maltose. This modification allows for much easier removal of the sugar by dialysis at a later step (D.E. Anderson, unpubl.). MBP mutants with impaired sugar binding (i.e., W62E, A63E, Y155E, W230E, and W340E), as well as Y210A MBP, were purified by an alternate method. The insoluble material from a 4 M ammonium sulfate cut was removed by centrifugation and the supernatant was applied to a butyl–Sepharose 4

Fast Flow column (Amersham Pharmacia Biotech Inc., Piscataway, NJ). After the column was washed with 1.5 M ammonium sulfate, the MBP was eluted with 25 mM HEPES (pH 8). The fractions containing MBP were pooled, diluted 5-fold with 25 mM HEPES (pH 8), and then applied to a column of Macro Prep ceramic hydroxylapatite (BioRad, Inc., Hercules, CA). The column was washed with 100 mM NaCl, and then the MBP was eluted with 100 mM phosphate in 25 mM HEPES (pH 8). The MBP-containing fractions from this column were diluted 5-fold and applied to a Q-Sepharose Fast Flow column (Amersham Pharmacia Biotech). The column was washed with 10 mM NaCl and then the MBP was eluted with 120 mM NaCl. Fractions from the Q-Sepharose or amylose columns containing purified MBP were concentrated by diafiltration (Centriplus, Amicon, Inc., Beverly, MA) and stored at −80°C until needed.

Equilibrium unfolding of MBP

Proteins were thawed and dialyzed into 20 mM phosphate buffer (pH 7.5) before analysis. Each sample was then diluted to an A_{280} value of ~0.7 in the same buffer. Guanidine hydrochloride (Gd-HCl) denaturation was carried out in a total volume of 40 μL with 20 μL diluted protein, 4 μL buffer (20 mM phosphate at pH 7.5), 0–13 μL 5.42 M Gd-HCl, and distilled water to make up the balance. The concentration of the Gd-HCl stock solution was determined by measuring its refractive index (Pace and Scholtz 1997) using an optical refractometer. The samples were equilibrated for 16–20 h before determining the extent of unfolding by CD spectroscopy. CD measurements were acquired in a CD spectrometer (Model 202, Aviv Instruments, Inc., Lakewood, NJ) using a 1-mm path length at 25°C. Measurement of ellipticity at 222 nm served as an indication of the degree of unfolding. Duplicate samples were each scanned five times to yield average ellipticity values. After plotting the average ellipticity versus Gd-HCl concentration, nonlinear least-squares analysis was used to fit the data to equations describing the two-state unfolding of proteins (Santoro and Bolen 1988; Pace and Scholtz 1997). This analysis yielded values for the free energy of unfolding ($\Delta G^0_{H_2O}$) as well as the Gd-HCl concentration at the transition midpoint (C_m).

Acknowledgments

We are grateful to Scott Cherry for his help with protein purification and to Anne Arthur for expert editorial assistance.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

Boos, W. and Shuman, H. 1998. Maltose/maltodextrin system of *Escherichia coli*: Transport, metabolism, and regulation. *Microbiol. Mol. Biol. Rev.* **62:** 204–229.

Brazil, B.T. and Horowitz, P.M. 1999. The hydrophobic properties of GroEL: A review of ligand effects on the modulation of GroEL hydrophobic surfaces. *Cell Stress Chaperones* **4:** 177–190.

Buckle, A.M., Zahn, R., and Fersht, A.R. 1997. A structural model for GroEL-polypeptide recognition. *Proc. Natl. Acad. Sci.* **94:** 3571–3575.

Butt, T.R., Jonnalagadda, S., Monia, B.P., Sternberg, E.J., Marsh, J.A., Stadel, J.M., Ecker, D.J., and Crooke, S.T. 1989. Ubiquitin fusion augments the yield of cloned gene products in *Escherichia coli*. *Proc. Natl. Acad. Sci.* **86:** 2540–2544.

Chatellier, J., Hill, F., Lund, P.A., and Fersht, A.R. 1998. In vivo activities of GroEL minichaperones. *Proc. Natl. Acad. Sci.* **95:** 9861–9866.

Chatellier, J., Buckle, A.M., and Fersht, A.R. 1999. GroEL recognizes sequential and non-sequential linear structural motifs compatible with extended beta-strands and alpha-helices. *J. Mol. Biol.* **292:** 163–172.

Chen, L. and Sigler, P.B. 1999. The crystal structure of a GroEL/peptide complex: Plasticity as a basis for substrate diversity. *Cell* **99:** 757–768.

Chun, S.Y., Strobel, S., Bassford, P. Jr., and Randall, L.L. 1993. Folding of maltose-binding protein. Evidence for the identity of the rate-determining step in vivo and in vitro. *J. Biol. Chem.* **268:** 20855–20862.

Cormack, B.P., Valdivia, R.H., and Falkow, S. 1996. FACS-optimized mutants of the green fluorescent protein (GFP). *Gene* **173:** 33–38.

Crameri, A., Whitehorn, E.A., Tate, E., and Stemmer, W.P. 1996. Improved green fluorescent protein by molecular evolution using DNA shuffling. *Nature Biotech.* **14:** 315–319.

Diamond, D.L., Strobel, S., Chun, S.Y., and Randall, L.L. 1995. Interaction of SecB with intermediates along the folding pathway of maltose-binding protein. *Protein Sci.* **4:** 1118–1123.

Farthing, A.J. and Vousden, K.H. 1994. Functions of human papillomavirus E6 and E7 oncoproteins. *Trends Microbiol.* **2:** 170–174.

Feldman, D.E. and Frydman, J. 2000. Protein folding in vivo: the importance of molecular chaperones. *Curr. Opin. Struct. Biol.* **10:** 26–33.

Feltham, J.L. and Gierasch, L.M. 2000. GroEL-substrate interactions: Molding the fold, or folding the mold? *Cell* **100:** 193–196.

Fenton, W.A., Kashi, Y., Furtak, K., and Horwich, A.L. 1994. Residues in chaperonin GroEL required for polypeptide binding and release. *Nature (Lond)* **371:** 614–619.

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.

Kimata, Y., Iwaki, M., Lim, C.R., and Kohno, K. 1997. A novel mutation which enhances the fluorescence of green fluorescent protein at high temperatures. *Biochem. Biophys. Res. Commun.* **232:** 69–73.

Kishore, U., Leigh, L.E.A., Eggleton, P., Strong, P., Perdikoulis, M.V., Willis, A.C., and Reid, K.B. 1998. Functional characterization of a recombinant form of the C-terminal, globular head region of the B-chain of human serum complement protein, C1q. *Biochem. J.* **333:** 27–32.

Kobayashi, N., Freund, S.M., Chatellier, J., Zahn, R., and Fersht, A.R. 1999. NMR analysis of the binding of a rhodanese peptide to a minichaperone in solution. *J. Mol. Biol.* **292:** 181–190.

Lee, H.S., Berger, D.K., and Kustu, S. 1993. Activity of purified NIFA, a transcriptional activator of nitrogen fixation genes. *Proc. Natl. Acad. Sci.* **90:** 2266–2270.

Liu, G.P., Topping, T.B., Cover, W.H., and Randall, L.L. 1988. Retardation of folding as a possible means of suppression of a mutation in the leader sequence of an exported protein. *J. Biol. Chem.* **263:** 14790–14793.

Louis, J.M., McDonald, R.A., Nashed, N.T., Wondrak, E.M., Jerina, D.M., Oroszlan, S., and Mora, P.T. 1991. Autoprocessing of the HIV-1 protease using purified wild-type and mutated fusion proteins expressed at high levels in *Escherichia coli*. *Eur. J. Biochem.* **199:** 361–369.

Martineau, P., Saurin, W., Hofnung, M., Spurlino, J.C., and Quiocho, F.A. 1990a. Progress in the identification of interaction sites on the periplasmic maltose binding protein from *E. coli*. *Biochimie* **72:** 397–402.

Martineau, P., Szmelcman, S., Spurlino, J.C., Quiocho, F.A., and Hofnung, M. 1990b. Genetic approach to the role of tryptophan residues in the activities and fluorescence of a bacterial periplasmic maltose-binding protein. *J. Mol. Biol.* **214:** 337–352.

Mottershead, D.G., Polly, P., Lyons, R.J., Sutherland, R.L., and Watts, C.K. 1996. High activity, soluble, bacterially expressed human vitamin D receptor and its ligand binding domain. *J. Cell. Biochem.* **61:** 325–337.

Pace, C.N. and Scholtz, J.M. 1997. Measuring the conformational stability of a protein. In *Protein structure: A practical approach* (ed. T.E. Creighton), pp. 299–321. IRL Press, Oxford.

Pryor, K.D. and Leiting, B. 1997. High-level expression of soluble protein in *Escherichia coli* using a His6-tag and maltose-binding-protein double-affinity fusion system. *Protein Expr. Purif.* **10:** 309–319.

Quiocho, F.A., Spurlino, J.C., and Rodseth, L.E. 1997. Extensive features of tight oligosaccharide binding revealed in high-resolution structures of the maltodextrin transport/chemosensory receptor. *Structure* **5:** 997–1015.

Randall, L.L., Topping, T.B., Smith, V.F., Diamond, D.L., and Hardy, S.J. 1998. SecB: A chaperone from *Escherichia coli*. *Methods Enzymol.* **290:** 444–459.

Ranson, N.A., White, H.E., and Saibil, H.R. 1998. Chaperonins. *Biochem. J.* **333:** 233–242.

Rao, S. and Bodley, J.W. 1996. Expression, purification, and characterization of

the G domain of *Saccharomyces cerevisiae* elongation factor 2. *Protein Expr. Purif.* **8:** 91–96.

Richarme, G. and Caldas, T.D. 1997. Chaperone properties of the bacterial periplasmic substrate-binding proteins. *J. Biol. Chem.* **272:** 15607–15612.

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

Russo, A.A., Tong, L., Lee, J.O., Jeffrey, P.D., and Pavletich, N.P. 1998. Structural basis for inhibition of the cyclin-dependent kinase Cdk6 by the tumour suppressor p16INK4a. *Nature (Lond)* **395:** 237–243.

Saavedra-Alanis, V.M., Rysavy, P., Rosenberg, L.E., and Kalousek, F. 1994. Rat liver mitochondrial processing peptidase. Both alpha- and beta-subunits are required for activity. *J. Biol. Chem.* **269:** 9284–9288.

Sachdev, D. and Chirgwin, J.M. 1998a. Order of fusions between bacterial and mammalian proteins can determine solubility in *Escherichia coli*. *Biochem. Biophys. Res. Commun.* **244:** 933–937.

———. 1998b. Solubility of proteins isolated from inclusion bodies is enhanced by fusion to maltose-binding protein or thioredoxin. *Protein Expr. Purif.* **12:** 122–132.

Santoro, M.M. and Bolen, D.W. 1988. Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants. *Biochemistry* **27:** 8063–8068.

Schein, C.H. 1989. Production of soluble recombinant proteins in bacteria. *Biotechnology (NY)* **7:** 1141–1147.

Sharff, A.J., Rodseth, L.E., Spurlino, J.C., and Quiocho, F.A. 1992. Crystallographic evidence of a large ligand-induced hinge-twist motion between the two domains of the maltodextrin binding protein involved in active transport and chemotaxis. *Biochemistry* **31:** 10657–10663.

Sheshadri, S., Lingaraju, G.M., and Varadarajan, R. 1999. Denaturant mediated unfolding of both native and molten globule states of maltose binding protein are accompanied by large deltaCp's. *Protein Sci.* **8:** 1689–1695.

Shtilerman, M., Lorimer, G.H., and Englander, S.W. 1999. Chaperonin function: folding by forced unfolding. *Science* **284:** 822–825.

Siemering, K.R., Golbik, R., Sever, R., and Haseloff, J. 1996. Mutations that suppress the thermosensitivity of green fluorescent protein. *Curr. Biol.* **6:** 1653–1663.

Sigler, P.B., Xu, Z., Rye, H.S., Burston, S.G., Fenton, W.A., and Horwich, A.L. 1998. Structure and function in GroEL-mediated protein folding. *Annu. Rev. Biochem.* **67:** 581–608.

Spurlino, J.C., Lu, G.Y., and Quiocho, F.A. 1991. The 2.3-Å resolution structure of the maltose- or maltodextrin-binding protein, a primary receptor of bacterial active transport and chemotaxis. *J. Biol. Chem.* **266:** 5202–5219.

Szmelcman, S., Sassoon, N., and Hofnung, M. 1997. Residues in the alpha helix 7 of the bacterial maltose binding protein which are important in interactions with the Mal FGK2 complex. *Protein Sci.* **6:** 628–636.

Tanaka, N. and Fersht, A.R. 1999. Identification of substrate binding site of GroEL minichaperone in solution. *J. Mol. Biol.* **292:** 173–180.

Tang, K.S., Guralnick, B.J., Wang, W.K., Fersht, A.R., and Itzhaki, L.S. 1999. Stability and folding of the tumour suppressor protein p16. *J. Mol. Biol.* **285:** 1869–1886.

Thomas, S., Soriano, S., d'Santos, C., and Banting, G. 1996. Expression of recombinant rat myo-inositol 1,4,5-trisphosphate 3-kinase B suggests a regulatory role for its N-terminus. *Biochem. J.* **319:** 713–716.

Tsao, K.L., DeBarbieri, B., Michel, H., and Waugh, D.S. 1996. A versatile plasmid expression vector for the production of biotinylated proteins by site-specific, enzymatic modification in *Escherichia coli*. *Gene* **169:** 59–64.

Wang, H., Kurochkin, A.V., Pang, Y., Hu, W., Flynn, G.C., and Zuiderweg, E.R. 1998. NMR solution structure of the 21 kDa chaperone protein DnaK substrate binding domain: A preview of chaperone-protein interaction. *Biochemistry* **37:** 7929–7940.

Wang, J.D. and Weissman, J.S. 1999. Thinking outside the box: new insights into the mechanism of GroEL-mediated protein folding. *Nat. Struct. Biol.* **6:** 597–600.

Yang, F., Moss, L.G., and Phillips, G.N. Jr. 1996. The molecular structure of green fluorescent protein. *Nature Biotech* **14:** 1246–1251.

Zahn, R., Buckle, A.M., Perrett, S., Johnson, C.M., Corrales, F.J., Golbik, R., and Fersht, A.R. 1996. Chaperone activity and structure of monomeric polypeptide binding domains of GroEL. *Proc. Natl. Acad. Sci.* **93:** 15024–15029.

Zhu, X., Zhao, X., Burkholder, W.F., Gragerov, A., Ogata, C.M., Gottesman, M.E., and Hendrickson, W.A. 1996. Structural analysis of substrate binding by the molecular chaperone DnaK. *Science* **272:** 1606–1614.