Folding of Maltose-binding Protein

EVIDENCE FOR THE IDENTITY OF THE RATE-DETERMINING STEP IN VIVO AND IN VITRO

(Received for publication, May 10, 1993, and in revised form, June 22, 1993)

Sang-Yeun Chung*, Sharon Strobel†, Philip Bassford, Jr.‡, and Linda L. Randall‡‡

From the Program in Genetics and Cell Biology and the Department of Biochemistry and Biophysics, Washington State University, Pullman, Washington 99164-4680 and the Department of Microbiology and Immunology, University of North Carolina, Chapel Hill, North Carolina 27599-7290

The folding of maltose-binding protein, a periplasmic protein in Escherichia coli, was shown to proceed through the same rate-limiting step whether folding occurred in the cell under physiological conditions or in vitro in the absence of other proteins. Four species of maltose-binding protein containing aminoacyl substitutions identified as decreasing the rate of folding of the protein in vivo were purified, and their denaturant-induced folding transitions were analyzed by monitoring the intrinsic fluorescence of tryptophan. In all four cases the rate of folding in vitro was slower than that of the wild-type maltose-binding protein; thus the same step determines the time of folding in vivo and in vitro. Furthermore, examination of the three-dimensional structure of maltose-binding protein as determined by x-ray crystallography (F. Quiocho, personal communication; Spurillo, J. C., Lu, G.-Y., and Quiocho, F. A. (1991) J. Biol. Chem. 266, 5202–5219) indicates that all 4 of the residues identified as crucial to folding lie in one structural element of the native protein. We conclude that the rate-limiting step both in vivo and in vitro involves formation of this element of structure.

The information governing the intimate folding of polypeptides into compact structures resides in the primary sequence of the aminoacyl residues as is clearly evidenced from the ability of isolated polypeptides to achieve their native conformation spontaneously in vitro (1). Even so in the complex milieu of the cellular cytosol the acquisition of the correct structure often depends upon facilitation by a family of proteins termed chaperones (2). The need for chaperones can be rationalized by the obvious differences between the conditions in vitro and in vivo. The investigator has control over the protein concentration, ionic conditions, and temperature in vitro so that the system can be optimized to favor folding. In the cell where protein concentration is high, the polypeptide has many opportunities to enter nonproductive interactions that would interfere with folding. Two fundamentally different models have been proposed to account for the action of chaperones. First, it is possible that they actively participate in the folding reaction much as enzymes act to catalyze chemical reactions (3–6). In one version of this model the chaperone would have affinity for an intermediate in folding and upon binding would effect a decrease in the activation energy along the folding pathway, thereby increasing the rate constants for the reaction. Another proposal for active participation is that the chaperone binds and releases portions of the polypeptide in an ordered way to direct the folding pathway. A second model is that chaperones do not actively modulate folding but enhance the flux through the proper folding pathway by controlling competing processes that lead to aggregation (7–12). Intermediates along folding pathways often display hydrophobic surfaces and are prone to aggregation at high concentration. Thus a simple way to facilitate folding of a monomeric protein is to bind the intermediate and lower its concentration, thereby decreasing the rate of bimolecular processes such as aggregation with no effect on the rate of unimolecular folding. Higher order reactions such as the formation of oligomers could also be favored if the folding reaction were to proceed on the surface of the chaperone and the polypeptide were released in a state that would oligomerize more rapidly than it would aggregate. It should be noted that folding on the surface does not necessarily require participation of the chaperone in a catalytic sense. The reaction might proceed at the same rate as it would in free solution if the competing aggregation pathway could be blocked.

The difference in these two models is important because of the implications for any attempt to elucidate the pathway of folding. The consensus, based on scrutiny of a tremendous body of data gathered in vitro with many proteins, is that the spontaneous folding observed proceeds via defined intermediates (13, 14). A crucial question is are the pathways and intermediates we can define in vitro using the powerful biophysical tools available to us the pathways utilized in vivo or do chaperones act to direct polypeptides along different pathways?

There is much evidence obtained from studies of chaperone-mediated folding in vitro that the mode of action is one of blockage of aggregation (7–12); however, Hartl and colleagues (6) have shown that changes do occur while polypeptides are tethered to the chaperone. Since the elegant biophysical techniques used in folding studies cannot be applied in vivo, it is difficult to relate the in vitro data to occurrences in the cell. We have taken advantage of the relationship between folding and export of maltose-binding protein to identify alterations of maltose-binding protein that affect the step that determines the rate of folding in vivo, and we demonstrate that the same step in folding is rate-determining in vitro. Examination of the location in the native structure of the residues that are altered indicates that the structure of the transition state in the rate-limiting step is near the native state.
EXPERIMENTAL PROCEDURES

Materials—Urea (ultrapure) was purchased from Boehringer Mannheim, HEPES and Trizma Base from Sigma, and the amylose resin from New England Biolabs. [35S]Methionine (1000 Ci/mmol) was purchased from Du Pont-New England Nuclear.

Bacterial Strains—For ease of discussion, throughout the text we shall designate alleles of malE, the gene encoding maltose-binding protein, and their corresponding protein products according to the aminocyl substitution found in the polypeptide. The first letter following malE represents the aminocyl residue found in the wild-type protein and the number indicates the position of the substitution. If the position of the substitution is in the leader peptide it is given in parentheses as a negative number. The position immediately preceding the first residue of the mature is designated “−1.” The letter following the number represents the aminocyl residue found in place of the wild-type residue. In some cases, we also indicate the name given to the alleles at the time of isolation.

All strains are derivatives of Escherichia coli K-12. An isogenic derivative of MC4100 (F lacU169 araD139 rpsL150 thi 1 flbB5301 deo-7 ptsf2 relA1) (15) carries the double mutation malE−19, malE G19C (hereafter, malE M(−8)R, malE G19C) (16). KM4 and KM19 were provided by M. Manson and are isogenic derivatives of strain MM110 (F− thr−1-Am) leuB6 his−4 thi−1 rpsL136 ara−14 lacY1 met−1 xyI−1 tonA317 tax−7 malT−1 zib−729; (21a) containing the malE54 (malE D55) and malE64 (malE T3451) alleles, respectively (17). Plasmids carrying malE A(−13)E, malE T3451, and malE D55 are derivatives of pBAR43 (18), and plasmids carrying the alleles malE T3451, malE A(−13)E, malE T3451, and the double mutations malE T(−11)K, malE T3451, and malE A(−13)E, malE T3451, are derivatives of pF2 (20). All genes are expressed under the control of lacUV5 promoter. Since the host strain does not contain the lac repressor on a high copy number, plasmid induction was not necessary to achieve high levels of expression. The host strain for the plasmids, BAR1091, a derivative of MC4100, carries a deletion in the malE gene (AmalE312 (20)).

Selection of Suppressors—The plasmid carrying the wild-type malE allele, pmalE, or pmalE-SpeI was mutagenized with hydroxylamine in an attempt to increase the probability of isolating mutants of interest. A sample of the mutagenized plasmid was used to demonstrate that the mutagenesis was effective; the frequency of appearance of colonies with a Mal− phenotype was approximately 1%. Nevertheless, the base changes in the suppressors isolated as described below seem to be spontaneous changes that did not result from the mutagenesis, since they are not characteristic of treatment with hydroxylamine (21). Following mutagenesis, the fragment containing the mature portion of malE was excised from the plasmid using the restriction sites BglII (after 119th codon of the sequence encoding the mature region of maltose-binding protein) and PstI, or SpeI (see Fig. 1) and ligated to the smaller fragment of pmalE A(−13)E digested with BglII and NcoI. Most of the suppressors initially identified carried mutations located outside the mature portion of the gene and affected the level of expression. The search resulted in identification of two mutations in the portion of the gene encoding the mature protein: malE T283D and malE A276G. The bands stained in malE T283D and malE A276G were identified by determining the nucleotide sequence of the entire BglII-NcoI region of the suppressor genes. The sequencing reactions were carried out using the T7 DNA polymerase kit from United States Biochemical Corp. as suggested by the manufacturer. The DNA primers were synthesized using an oligonucleotide synthesizer (Pharmacia Gene Assembler).

Kinetics of Processing of Precursor Maltose-binding Protein—Cells were grown to a density of 2.5 × 10^6 cells/ml in M9 minimal salts medium (21) supplemented with 4 × 10−5% vitamin B_6, 0.4% glycerol, and 0.2% maltose and radiolabeled for 15 s by the addition of 60 Ci/mmol of [35S]methionine and nonradioactive methionine to give a final concentration of 30 nM methionine. At 15 s, nonradioactive methionine was added to give a final concentration of 100 μM, and at the times indicated 10^6 cells were taken directly into trichloroacetic acid to give a final concentration of 5% trichloroacetic acid. The samples were processed for immunoprecipitation with antiserum to maltose-binding protein as described (22). The immunoprecipitated samples were analyzed by SDS, 11% polyacrylamide gel electrophoresis. The gels were soaked in 1 M sodium salicylate (pH 5.8) before they were dried and overlaid with x-ray film for fluorography. The bands of maltose-binding protein at the positions of precursor and mature forms were quantified by densitometry with a Helena Laboratories Quick Scan R + D to calculate the percent of maltose-binding protein that is mature. The data were adjusted for the removal of 3 of the 9 methionines when precursor is matured.

Purification of Maltose-binding Protein Species—The matured forms of wild-type maltose-binding protein, MalE V8G, MalE A276G, and MalE Y283D, were purified from strains harboring plasmids that carried the appropriate malE gene. All four malE alleles encoded a wild-type leader peptide. Mature MalE G19C was purified from a strain carrying the double mutation, malE M(−8)R, malE G19C on the chromosome, after induction by maltose. The purification allowed sufficient export to purify substantial quantities of mature protein. MalE D55 and MalE T3451, expressed from alleles encoding wild-type leaders, were purified from strains carrying malT−1; therefore induction by maltose was not required. The proteins were purified by affinity chromatography using an amylose resin as described (23). The proteins were concentrated by ultrafiltration through an Amicon membrane filter (YM-10) and analyzed by electrophoreses on sodium dodecyl sulfate, 15% polyacrylamide gels to verify the purity. The concentrations of the proteins was determined by the method of Lowry et al. (24) using bovine serum albumin as a standard.

Fluorescence Spectroscopy—the urea-induced unfolding or refolding of the maltose-binding protein species was monitored by the decrease or increase of the intrinsic fluorescence of tryptophan, respectively. Fluorescence intensity was measured using a Shimadzu RF-5000 fluorescence spectrophotometer with an excitation wavelength of 295 nm (slit width, 2 nm) and an emission wavelength of 344 nm (slit width, 5 nm). All fluorescence measurements were made at 25 °C.

Equilibrium Studies of Folding—Native proteins (5 μg) were added to varying concentrations of urea in 3 ml of 10 mM HEPES (pH 7.6 adjusted with KOH), and the fluorescence intensity of each sample was measured after equilibration at 25 °C. The final concentration of urea in each sample was confirmed by measuring the index of refraction. The urea solutions were degassed, filtered, and used within a day.

Kinetic Studies of Folding—Unfolding reactions were initiated by adding approximately 5 μg of native protein to a solution of urea in 10 mM HEPES (pH 7.6) (final volume, 3 ml) held in a stirred cuvette in the chamber of the spectrophotometer. The addition of protein was made manually with a plastic plunger. The reaction was monitored by addition of the sample and the first recording of fluorescence intensity was approximately 3 s. For the series of refolding studies, maltose-binding protein was unfolded by incubation for over 2 h in 6 M urea in 10 mM HEPES (pH 7.6). Folding was initiated by addition of the denatured protein (approximately 5 μg) to a predetermined volume of a solution of the native maltose-binding protein such that the concentration was as indicated for each experiment. The relaxation time to achieve the new equilibrium was extracted from a plot of the log of the change in fluorescence versus time.

![FIG. 1. Structure of plasmid pmalE-SpeI used in isolation of suppressors. Sites for the restriction enzymes used are indicated.](image-url)
Analysis of Folding Data—Equilibrium unfolding data obtained from the changes in the fluorescent intensity were fit to the following two-state model,

\[
K_{\text{app}} \frac{N}{U} = 1 + K_{\text{app}} \quad \text{(Model 1)}
\]

where \(N\) is the native protein, \(U\) is the unfolded protein, and \(K_{\text{app}} = [U]/[N]\). It can be shown that \(F_{\text{app}}\), the apparent fraction of the unfolded form, is as follows.

\[
F_{\text{app}} = \frac{K_{\text{app}}}{1 + K_{\text{app}}} \quad \text{(Eq. 1)}
\]

The equation \(K_{\text{app}} = \exp(-\Delta G^{\text{apo}} + A[\text{urea}]/RT)\) (25, 26) was substituted into the above equation to calculate \(\Delta G^{\text{apo}}\) and \(A\) values, where \(\Delta G^{\text{apo}}\) is the free energy difference in the absence of denaturant, and \(A\) is a parameter that describes the cooperativity of the unfolding transition (27).

For comparison of the wild-type and mutant protein data, the observed changes in fluorescence signals were converted to an apparent fraction of unfolded protein, \(F_{\text{app}}\), such that,

\[
F_{\text{app}} = \frac{(Y_{\text{obs}} - Y_n)}{(Y_U - Y_n)} \quad \text{(Eq. 2)}
\]

where \(Y_{\text{obs}}\) refers to the observed fluorescence intensity at a particular urea concentration, and \(Y_n\) and \(Y_U\) refer to the calculated values for the native and unfolded forms, respectively, at the same denaturant concentration. A linear dependence of \(Y_{\text{obs}}\) on the urea concentration was observed in the pre- and posttransition baseline regions. Linear extrapolations from these baseline regions yielded the concentration independent values for \(Y_n\) and \(Y_U\). Concentration dependent values for \(Y_n\) and \(Y_U\) can be calculated from the following equations,

\[
Y_n = Y_n^0 + M_3[\text{urea}] \quad \text{(Eq. 3)}
\]

\[
Y_U = Y_U^0 + M_4[\text{urea}] \quad \text{(Eq. 4)}
\]

where \(M_3\) and \(M_4\) are the slopes of the pre- or posttransition regions (native and unfolded forms, respectively) (28). The above equations were combined to obtain a single equation for the dependence of the fluorescence intensity on the denaturant concentration.

\[
Y_{\text{obs}} = (Y_n^0 + M_3[\text{urea}]) + \left[\exp(-\Delta G^{\text{apo}} - A[\text{urea}]/RT)\right]/\left[1 + \exp(-\Delta G^{\text{apo}} - A[\text{urea}]/RT)\right] \quad \text{(Eq. 5)}
\]

The observed data were fit to this equation with a nonlinear least squares program NLIN (SAS Institute, Inc., Cary, NC). The midpoint of the unfolding transition curve, \(c_m\), was obtained by solving for the urea concentration when \(F_{\text{app}} = 0.5\).

RESULTS

Identification of Aminoacyl Changes That Slow Folding in Vivo—The scheme that we have used to select for changes in the step that determines the rate of folding in vivo is based on our current understanding of the mechanism of protein export in \(E. coli\). Proteins that are exported to the periplasmic space are synthesized as precursors containing amino-terminal extensions, the leader peptides. These precursors must cross the cytoplasmic membrane before they acquire a stable tertiary structure (22). Following synthesis of the precursors, there is a kinetic partitioning between the pathway of folding and the pathway of export. One of several functions of the leader peptide is to retard folding of the mature moiety allowing the precursor to engage the export apparatus. The leader is also involved in mediating transfer across the membrane. We have shown previously that a decrease in rate along the productive pathway, resulting from a defective leader sequence that mediates slow translocation across the membrane, can be partially overcome by slowing the folding of the precursor and thereby increasing the time during which that polypeptide is competent to enter the export pathway (23, 29). The identification of mutations in \(malE\), the gene encoding maltose-binding protein, that slow folding was accomplished by selection for intragenic suppressors of the export defect caused by a mutational change in the leader peptide (the precursor \(MalE((-13)E)\) carries a leader with the alanyl residue at position -13 changed to a glutamyl residue). This export-defect leader retards folding allowing entry into the export pathway but is very inefficient in promoting transfer through the membrane. Strains producing \(MalE((-13)E)\) contain less than 10% of the normal amount of maltose-binding protein in the periplasm and grow poorly on maltose. Strains carrying suppressors that enhance the ability to utilize maltose can be distinguished from the export-defective parental strain by a difference in color of the colonies grown on MacConkey indicator plates containing maltose. If the level of maltose-binding protein in the periplasm is increased to approximately 20% of wild-type levels, the growth on maltose is normal (30). Therefore suppression can be achieved not only by increasing the efficiency of export but also simply by producing 2- or 3-fold more precursor than normal. In fact, most of the suppressors initially identified were of the class that increase the level of expression of the \(malE\) allele. To identify the mutations of interest for this study, that is those that affect folding, the export kinetics of the candidate strains were examined. The screening yielded two intragenic mutations with alterations in the portion of the gene encoding the mature region, \(malE\) \(Y283D\) (the codon for the residue at position 283 was changed from TAT to GAT) and \(malE\) \(A276G\) (GCC to GGG at the position encoding residue 276). One of these mutations, \(malE\) \(Y283D\), is the same as a mutation isolated previously as \(malE2261\) in a similar search for intragenic suppressors (31). Both mutations \(malE\) \(Y283D\) and \(malE\) \(A276G\) occurred spontaneously; attempts to increase the number of suppressors by mutagenizing the portion of the gene encoding the mature region of the protein were unsuccessful (see “Experimental Procedures”).

Both substitutions improved the export of maltose-binding protein that carried the defective leader (Fig. 24). Newly synthesized polypeptides were labeled by addition of \([35S]\) methionine to exponentially growing cells followed 15 s later by addition of an excess of nonradioactive methionine. Incubation of the culture was continued, and samples were withdrawn at the times indicated. Maltose-binding protein was immunoprecipitated from the cell extracts, and the distribution of radioactivity between the precursor and mature forms was assessed by a combination of sodium dodecyl sulfate-polyacrylamide gel electrophoresis and fluorography. The wild-type precursor maltose-binding protein was rapidly and quantitatively converted to the mature form with a \(t_{1/2}\) of less than 30 s. The precursor species carrying the defective leader, \(MalE((-13)E)\), was very poorly exported. In 10 min less than 10% of the precursor was processed. The presence of either second suppressing mutation which changed a residue in the mature moiety increased the amount of precursor processed so that by 10 min 48% of \(MalE((-13)E)\), \(A276G\) was in the mature form. As published previously (31), 16% of \(MalE((-13)E)\), \(Y283D\) was processed within 10 min. Similar analyses of export kinetics showed that both \(malE\) \(Y283D\) and \(malE\) \(A276G\) could also suppress the defect caused by the leader mutation, \(malE\) \((-11)K\) (data not shown). Neither suppressor could restore export of precursors carrying the substitution \(M(-9)R\) (data not shown). \(MalE\) \((-9)R\) shows no processing of precursor, whereas the other defective leaders support 10–30% of the normal level of processing. This pattern of suppression is consistent with the notion that retardation of folding can increase the probability that at any given time the precursor is competent for export, but it cannot restore export if the translocation step is completely blocked.

\[1\] J. Weiss, personal communication.
binding protein, MalE T345I and MalE D55N, identified in a selection for altered chemotactic behavior (17), were included in the study to demonstrate that not every aminoacyl substitution affects folding. One of these mutations, malE T345I, was also tested for suppression of the leader mutation malE A(-13)E. No suppression of the export defect was observed as assessed by the color of colonies on MacConkey plates containing maltose or by determination of the kinetics of export in vivo (Fig. 2A). As shown below, the substitution has no effect on the rate of folding in vitro. The lack of suppression is consistent with a kinetic partitioning between folding and export in vivo.

**Folding Studies of Purified Proteins by Fluorescence Spectroscopy**—The intrinsic fluorescence of maltose-binding protein, which contains 8 tryptophanyl residues (Fig. 3; Refs. 32 and 33), decreases upon denaturation; thus fluorescence spectroscopy can be used to investigate the stabilities and folding properties of the various species of the protein. The species of maltose-binding protein that folds within the cell is the precursor form. Previous studies using fluorescence spectroscopy with purified proteins have shown that the presence of a leader peptide, whether it be a wild-type or an export-defective species, slows the folding of both maltose-binding protein (23) and ribose-binding protein (29). In addition, for the various species of both polypeptides studied previously, the retardation of folding exerted by the leader is additive to that caused by aminoacyl substitutions in the mature moiety. In the present study, we compared the rate of folding of the precursor form of one of the altered species of maltose-binding protein, MalE A(-13)E, A276G, with that of the wild type (Table I). For the species with two substitutions, the presence of the leader retarded the folding 36-fold relative to the mature form. The wild-type leader retarded folding 34-fold relative to the wild-type mature. Whether we compared the relative rates of folding of the mature forms or of the precursor forms for the two species of maltose-binding protein, the effect of the aminoacyl substitution in the mature moiety was the same: a 14-fold decrease in the rate of folding. Since the simultaneous presence of the leader and the substituted ami-

Table I

| Protein species | \( \tau \) [s⁻¹] |
|-----------------|------------------|
| Precursor MalE wild-type | 1,170 |
| Precursor MalE A(-13)E, A276G | 16,200 |
| Mature MalE wild-type | 34 |
| Mature MalE A276G | 442 |

*The value for the denaturant-dependent folding reaction was extrapolated (see Fig. 5) for comparison with the values of the corresponding observed denaturant-dependent folding of the other species. The observed \( \tau \) at 1.85 M urea was 91 s⁻¹.

The laboratory of P. J. Bassford, Jr. introduced the aminoacyl substitution V8G into malE by site-directed mutagenesis during construction of a restriction site for other purposes. Preliminary observations indicated that export of this altered species was less dependent than was that of wild-type protein on the bacterial chaperone SecB. Such a reduced requirement for SecB, which binds precursors and maintains them in an export-competent state, is consistent with the idea that the polypeptide itself remains in the competent nonnative state longer than does the wild-type protein. If this were true then the aminoacyl substitution should suppress an export defect caused by a mutated leader. To test this idea a construct was created that carried both malE T(-11)K, with two changes, MalE T(-11)K, V8G (§). Data for wild-type maltose-binding protein (○) are taken from A.

The folding properties of two additional species of maltose-

---

* S. Strobel and P. J. Bassford, Jr., personal communication.
The aminoacyl residue demonstrates an additive retardation of folding (476-fold), we can conclude that both the leader and the substitution interfere with folding at the same step in the pathway. Thus we could study the effects of the aminoacyl substitutions on the rate-limiting step in folding using the mature species. Use of the mature species offers considerable advantages. Not only is it difficult, because of problems of aggregation and degradation, to obtain pure precursors in quantities sufficient for spectroscopic studies, but also a more serious problem lies within the inherent slow folding of the precursors. At concentrations of denaturant corresponding to the middle portion of the unfolding transition even the mature forms fold so slowly that experiments are difficult (relaxation times can be as long as 15 h). The addition of the leader would retard folding another 35-fold making the experiments impossible; the relaxation time at the midpoint of the transition would be expected to be on the order of 21 days. Therefore the following in-depth equilibrium and kinetic analyses were carried out with the mature forms.

Equilibrium Unfolding Transitions for the Various Species of Maltose-binding Protein—An equilibrium study of the urea-induced reversible unfolding of the wild-type protein (Fig. 4) shows a smooth sharp transition between folded and unfolded states with a midpoint at 3.5 M urea. The urea-induced unfolding transitions of the six altered species of maltose-binding protein were compared with the transition of the wild type. The four substitutions that suppressed export defects rendered the polypeptide less stable than wild type (Fig. 4 and Table II). Of the two substitutions that affect the chemotactic properties of maltose-binding protein, but are not necessarily expected to alter the rate of folding, one had no effect on stability (MalE D55N) and the other (MalE T345I) shifted the midpoint of the transition to a higher concentration of urea relative to the wild-type protein, indicating increased stability (Fig. 4 and Table II).

Since the object of our study was to determine whether the step that limits the rate of folding in vivo is the same as that which limits the rate in vitro, it was necessary not only to carry out equilibrium studies, which address the stability of proteins, but also to determine directly the effects of the aminoacyl substitutions on folding and unfolding rates.

Kinetic Studies of the Folding Reactions—Purified protein, either in the native (no urea present) or in the unfolded state (6 M urea present), was subjected to a rapid change in conditions that required the protein to achieve a new equilibrium mixture of native and unfolded states. The relaxation time to reach the new equilibrium was determined by monitoring the change in fluorescence with time. A series of experiments was carried out for each species of maltose-binding protein, and the results are displayed in plots of the relaxation time as a function of the final concentration of urea (Figs. 5–7). For all seven proteins studied the unfolding reaction is described by a single kinetic phase that accounts for greater than 90% of the expected change in fluorescence intensity. The rate-limiting step that we monitor in unfolding is from the native state to whatever species follow it. In interpreting the data for refolding, we have assumed that the refolding of all species has as the end point the native folded form. This is clearly demonstrated for the wild-type species, since the relaxation times for unfolding and refolding meet smoothly at the inflexion point (Fig. 5), indicating that the refolding reactions observed are the reverse of the rate-limiting step in unfolding which has as its starting point the native state (14). Kinetic data could not be generated at urea concentrations near the inflexion point for one of the altered species (MalE Y283D) because of artifacts that are likely to result from aggregation of intermediates during the very slow folding phase. However, the relaxation time to reach equilibrium when shifting from no denaturant (i.e. native state) into concentrations of urea so low that folding is favored falls on the line described by refolding from the denatured state; thus, in this case too the folding we observe is the reverse of unfolding (Fig. 7).

Refolding of the wild-type protein is described by three kinetic phases (Fig. 5). A burst phase accounts for approximately 20% of the amplitude. The remaining 80% of the change in the fluorescence intensity is distributed between two refolding phases which we resolve. The relative amplitudes of these phases are 0.6 for the slower phase and 0.4 for the faster when refolding is observed at final concentrations of urea below 1.75 M. In this range, the relaxation times are independent of the concentration of denaturant (r = 20 ± 5, 60 s). Above 1.75 M urea, the relaxation times for refolding increase with increasing concentration of urea, and in this region the refolding phases have relative amplitudes of 0.7 for the slow phase and 0.3 for the fast. Both the slow and fast phases connect smoothly with the corresponding denaturant-independent folding phase as well as with the unfolding phase. The two species that carry alterations unrelated to the competence of the polypeptide for export have no effect on refolding (Fig. 6). One of these species, MalE T345I, displayed a

![Fig. 4. Equilibrium transition curves for unfolding of maltose-binding protein species.](image-url)

**Table II**

| Protein       | ΔC<sub>50</sub> (urea) | ΔC<sub>50</sub> (M, urea) | ΔG<sup>‡</sup> kcal mol<sup>−1</sup> |
|---------------|-------------------------|---------------------------|----------------------------------|
| Wild type     | 2.7 ± 0.3               | 3.50                      |                                  |
| MalE V8G     | 3.6 ± 0.6               | 3.14                      | 1.1                              |
| MalE G19C    | 2.8 ± 0.2               | 2.66                      | 2.3                              |
| MalE A276G   | 3.2 ± 0.3               | 2.98                      | 1.5                              |
| MalE Y283D   | 3.3 ± 0.4               | 2.44                      | 3.2                              |
| MalE T345I   | 2.4 ± 0.4               | 3.79                      | −0.7                             |
| MalE D55N    | 3.1 ± 0.4               | 3.49                      | 0                                |

*The change in stability by mutation is determined as ΔG = (ΔA) - (ΔC<sub>50</sub>) where ΔC<sub>50</sub> is the difference between the value of C<sub>50</sub> for wild-type and altered species and (ΔA) is the average value of A. A positive value for ΔG corresponds to an increase in stability due to the alteration (48).
FIG. 5. The urea dependence of the relaxation times for unfolding and refolding of the wild-type maltose-binding protein. The relaxation times for refolding (open symbols) and unfolding (closed symbols) were obtained as described under “Experimental Procedures.” Each refolding reaction has a slow and a fast phase. The lines are drawn to aid the eye.

decrease in the rate of unfolding relative to the wild type, which would account for the increase in relative stability observed in the equilibrium study.

The four species of maltose-binding protein with aminoacyl substitutions that improved export, and thereby are assumed to slow folding in vivo, refold more slowly than wild type in vitro as well (Fig. 7). In addition, the refolding at each concentration of denaturant is described by two kinetic phases in contrast to the three phases observed in the wild-type species. A burst phase accounts for approximately 20–30% of the amplitude with the recovery of the remainder of the fluorescence intensity described by one relaxation time.

DISCUSSION

There is currently much discussion concerning the relationship of folding pathways delineated in vitro to pathways of folding taken in vivo, which are difficult to study directly (34, 35). In order to compare the in vivo folding pathway of maltose-binding protein with the spontaneous folding assessed by tryptophanyl fluorescence, we have taken advantage of the inherent relation between the rate of folding of the polypeptide in the cytoplasm and its ability to be productively exported to the periplasm of E. coli. The entry of maltose-binding protein into the export pathway is mediated by its interaction with the chaperone SecB (36–38). Although the natural ligand for SecB is the precursor form of maltose-binding protein, which carries an amino-terminal extension, the leader sequence, the selective binding of the precursor is not the result of direct recognition of that leader (39). Rather, the leader slows the folding of the polypeptide, thereby allowing SecB to bind sites that would be rapidly buried in the folded mature form (39-41). The binding to SecB is readily reversible, thus in vivo, there is a kinetic partitioning between productive export and the folding of the precursor. Once folded, the precursor cannot enter the export pathway. The leader sequence has, in addition to its role to retard folding, a second function during the translocation step at the membrane. Altered leader sequences have been identified that effectively retard folding allowing rapid delivery to the membrane but are very inefficient in promoting translocation (23,
According to the model of kinetic partitioning such defects could be partially overcome by alterations in the polypeptide that further reduce the rate of folding. It has been demonstrated that selection for increased export of precursors carrying defective leaders does allow isolation of species of both ribose-binding protein (29) and maltose-binding protein (23) that have altered folding properties. Here in an extended analysis of maltose-binding protein we have examined four species identified as folding slowly in vivo: three species were selected as suppressors of export defects and one species had a substitution introduced for other purposes and was subsequently found to partially overcome export defects. All four species fold more slowly in vitro than does the wild-type protein. We also analyzed the spontaneous folding in vitro of two altered species that were selected in vivo for properties that are not expected to be related to folding. As anticipated the substitutions in these polypeptides did not alter the rate of folding assessed in vitro. We conclude that the rate-determining step in the folding of maltose-binding protein is the same in the pathway taken in the complex milieu of the cell as it is in the spontaneous folding pathway taken by the purified polypeptide. Thus, an analysis of the in vitro folding reaction allows detailed understanding of the step that limits the rate of folding in vivo.

The data demonstrating three kinetic phases in refolding of the wild-type protein, MalE T345I and MalE D55N, can be rationalized by a model for folding that has two parallel pathways (Fig. 8), each with two intermediates I1 and I2 or I1' and I2'. Unfolding is governed by the transition from N to I1 or I1'. Both reactions are proposed to pass through the same transition state to account for the single kinetic phase observed for unfolding. Refolding proceeds through a burst phase followed by the conversion of I1 (or I1') to I2 (or I2'), which is rate-limiting at urea concentrations below 1.75 M. Since these reactions are independent of the concentration of denaturant, they are likely to involve isomerization of the polypeptide backbone (43, 44). Above 1.75 M, the folding rate is determined by the conversion of I1 or I1' to N. The relaxation times of these reactions decrease as the concentration of urea decreases, indicating that these steps involve the folding of the polypeptide. We have proposed that the parallel channels are connected by interconversion at the states U, I1, and I1' because the proportion of the amplitude accounted for by the kinetic phases is dependent on whether the rate of unfolding is limited by I1 to I2 (isomerization) or by I1' to N (folding). We cannot presently eliminate an alternative model in which all four of the intermediate species lie along one pathway. In either case, the model for folding of the altered species is simpler, since one of the kinetic phases disappears and the model collapses to one channel comprising two intermediate states: U ↔ I1 ↔ I2 ↔ N, with U ↔ I1 representing a collapse on the millisecond time scale, I1 ↔ I2 an isomerization reaction, and I2 ↔ N the denaturant-dependent folding reaction.

It is of interest to note that folding studies of the wild-type species carried out in guanidinium chloride or in urea with the addition of KCl can also be accounted for by one channel of folding with two intermediates. Thus it seems that the ionic strength or the chloride ion affects the stability of the two intermediates or of their corresponding transition states.

We shall discuss the equilibrium and kinetic studies of the altered species in terms of this simplified model. The reaction U ↔ I1, which accounts for 20–30% of the total difference in fluorescence between U and N, is over within the time required to initiate the reaction manually. Attempts to resolve the reaction for the wild-type maltose-binding protein using stopped-flow fluorescence spectroscopy indicated that it occurs within 8 ms, the dead-time of the instrument. The intermediate formed is likely to represent a collapsed state that has considerable secondary structure, since analysis of the wild-type species by stopped-flow circular dichroic spectroscopy showed acquisition of substantial ellipticity at 220 nm during the initial burst phase reaction. The reactions that we follow by fluorescence spectroscopy with manual mixing are the much slower reactions from I1, the collapsed state intermediate, to I2 or from I1' to the native state, N. The four aminoacyl substitutions that slow the folding and act to enhance export in vivo do not significantly alter the rate of unfolding. In terms of a reaction coordinate diagram (Fig. 9) describing the reversible folding reaction (for simplicity of discussion we shall omit the isomerization step I1 ↔ I1'), we see that this effect can result from either an increase in stability of I2 relative to the transition state (TS) or a decrease in stability of the transition state that is carried over to the native state. In either case, the activation energy for folding is increased with no effect on the rate of the unfolding reaction. Any increase in the energy level of N without an equal change in the energy level of U would require that the altered protein be less stable than the wild type. Equilibrium studies show that this is true for all of the four altered species. Thus for these species it is likely that the transition state and the corresponding native state are destabilized to the same extent.

Examination of the three-dimensional structure of maltose-binding protein, resolved to 1.7 Å by x-ray crystallography, shows that although the residues that slow the rate of folding are widely spread within the primary sequence (at positions 8, 19, 276, and 283), they are all found in the same element.

3 L. Randall, C. Mann, and C. R. Matthews, unpublished results.
4 C. Mann and C. R. Matthews, personal communication.
5 F. Quiocho, personal communication.
of structure in the native state (Fig. 10). This observation supports the idea that formation of this element of structure is rate-limiting and that the structure of the transition state for this reaction is near-native. Maltose-binding protein is a member of a family of soluble periplasmic binding proteins that have similar structures. They are bilobate α-β proteins with the ligand-binding site in a cleft separating two globular domains (45, 46). The substitutions that limit the rate of folding for maltose-binding protein are in the N domain (also referred to as domain 1). It is of interest that alterations identified by a similar selection scheme as crucial to folding for ribose-binding protein are confined to the corresponding domain of that protein (29, 47). It may be that the rate of folding of all polypeptides in the binding-protein family is limited by acquisition of structure in this domain.

During the initial stages of protein export, one of the functions of the leader peptide is to slow the folding of the precursor to render it competent to engage the export apparatus (40, 41). As we have discussed, the leader has its affect on the same step as do the aminocyl substitutions; thus, the study of folding described here provides insight into a possible mechanism for this retardation of folding. The amino terminus of the mature protein lies near the element of structure defined by the aminocyl substitutions as crucial in folding. In the precursor the leader, which is an extension at the amino terminus, might fold into this element and thereby interfere with the rate-limiting step.

The comparative study of folding in vivo and in vitro presented here cannot help us in determining the exact rate constants of the folding reactions in vivo. Thus this type of analysis does not directly address the question of whether in vivo chaperones act as catalysts by decreasing the activation energy of folding reactions. What can be concluded is that in the case of maltose-binding protein interaction with cellular components does not actively direct the polypeptide along a folding pathway that includes an intermediate for the rate-limiting step different from that present in the spontaneous folding pathway. Both in the complex milieu of the cell and in a purified system the same step determines the rate. Thus studies carried out in vitro can provide insight into the intermediate involved in the acquisition of proper native structure. In addition, in the case of exported proteins such as maltose-binding protein and ribose-binding protein such studies may help to elucidate the structural state that is competent for translocation through the membrane.

Acknowledgments—We thank Florante Quiocio for generously providing the coordinates for the 1.7-Å structure of maltose-binding protein prior to publication. We are grateful to Traci Topping for purification of the wild-type maltose binding protein, to Craig Mann for fitting the data to a two-state model, and to Chankyu Park for help with the computer. We express our deep appreciation to C. Robert Matthews for helpful discussions throughout the entire work.

REFERENCES
1. Epstein, C. J., Goldenberger, R. F., and Anfinsen, C. B. (1983) Cold Spring Harbor Symp. 48, 439-449
2. Eldridge, R. (1987) Nature 328, 378-379
3. Flynn, G. C., Chappell, T. G., and Rothman, J. E. (1989) Science 245, 385-390
4. Osumi, J., Horwich, A. L., Neupert, W., and Hartl, F. U. (1989) Nature 341, 125-130
5. Rothman, J. E. (1989) Cell 59, 591-601
6. Martin, J., Langer, T., Bollinger, J., Schramel, A., Horwich, A. L., and Hartl, F. U. (1991) Nature 352, 36-42
7. Vittam, P. V., Lobben, L., Vriend, G., and Schmitz, H. G. (1990) Biochemistry 29, 5665-5671
8. Buchner, J., Schmidt, M., Schmitz, H., Jaenicke, R., Rudolph, R., Schmid, P. X., and Kiethaber, T. (1991) Biochemistry 30, 1586-1591
9. LaRossa, R. A., and van Dyk, T. K. (1991) Mol. Microbiol. 5, 523-534
10. Fisher, M. T. (1992) Biochemistry 31, 3965-3970
11. Horowitz, J. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 1048-1053
12. Jakob, U., Gaestel, M., Engel, K., and Buchner, J. (1993) J. Biol. Chem. 268, 1517-1520
13. Kim, P. S., and Baldwin, R. L. (1990) Annu. Rev. Biochem. 59, 631-660
14. Matthews, C. R. (1991) Curr. Opin. Struct. Biol. 1, 28-35
15. Casadaban, M. J. (1976) J. Mol. Biol. 104, 541-555
16. Ryan, J. P., Dunck, M. C., Bankaitis, V. A., and Bassford, P. J., Jr. (1986) J. Biol. Chem. 261, 3389-3395
17. Manson, D. M., and Kossmann, H. H. (1986) J. Bacteriol. 165, 34-40
18. Rasmussen, B. A., McGregor, C. H., Ray, P. H., and Bassford, P. J., Jr. (1985) J. Bacteriol. 164, 665-673
19. Bedouelle, H., Bassford, P. J., Jr., Fowler, A. V., Zabin, I., Beckwich, J., and Hofnung, M. (1980) Nature 285, 78-81
20. Fikes, J. D., and Bassford, P. J., Jr. (1987) J. Bacteriol. 169, 2352-2359
21. Miller, J. H. (1972) Experiments in Molecular Genetics, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
22. Randall, L. L., and Hardy, S. J. S. (1986) Cell 46, 921-928
23. Liu, G., Topping, T. B., Cover, W. H., and Randall, L. L. (1988) J. Biol. Chem. 263, 14790-14790
24. Lowry, O. H., Rosebrough, N. J., Farr, A. L., and Randall, R. J. (1951) J. Biol. Chem. 193, 265-275
25. Schellman, J. A. (1978) Biopolymers 17, 1305-1322
26. Pace, C. N. (1986) Methods Enzymol. 131, 266-280
27. Matthews, C. R. (1987) Methods Enzymol. 154, 498-511
28. Cantow, M. M., and Boles, D. W. (1988) Biochemistry 27, 8063-8068
29. Teschke, C. M., Kim, J., Song, T., Park, S., Park, C., and Randall, L. L. (1991) J. Biol. Chem. 266, 11789-11796
30. Mann, M. D., Boos, W. T., Bassford, P. J., Jr., and Rasmussen, B. A. (1985) J. Biol. Chem. 260, 9727-9733
31. Cover, W. H., Ryan, P., Bassford, P. J., Jr., Walsh, K. A., Bollinger, J., and Randall, L. L. (1987) J. Bacteriol. 169, 784-790
32. Martineau, P., Sznitman, S., Spurlino, J. C., Quiocio, F. A., and Hofnung, M. (1986) J. Mol. Biol. 194, 357-352
33. Spurlino, J. C., Lu, G.-Y., and Quiocio, F. A. (1991) J. Biol. Chem. 266, 5219-5221
34. Fecher, G., and Schmid, F. X. (1990) Biochemistry 29, 2205-2212
35. Getting, M. J., and Sambrook, J. (1992) Nature 355, 33-45
36. Kumasato, C. A., and Beckwith, J. (1985) J. Bacteriol. 163, 267-274
37. Collier, D. N., Bankaitis, V. A., Weiss, J. B., and Bassford, P. J., Jr. (1986) Cell 53, 273-283
38. Kumasato, C. A. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 5320-5324
39. Randall, L. L., and Hardy, S. J. S. (1990) Science 245, 860-863
40. Park, S., Liu, G., Topping, T. B., Cover, W. H., and Randall, L. L. (1988) Science 241, 1033-1035
41. Liu, G., Topping, T. B., and Randall, L. L. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 9213-9217
42. Thomasson and Randall, L. L. (1988) J. Bacteriol. 170, 5654-5661
43. Brandt, J. F., Halfonson, H. R., and Brennan, M. (1975) Biochemistry 14, 4953-4963
44. Schmid, F. X., and Baldwin, R. L. (1979) J. Mol. Biol. 133, 285-287
45. Vyas, N. K., Vyas, M. N., and Quiocio, F. A. (1991) J. Biol. Chem. 266, 12525-12527
46. Mowbray, S. L. (1992) J. Mol. Biol. 227, 418-440
47. Mowbray, S. L., and Cole, L. B. (1992) J. Mol. Biol. 225, 155-175
48. Matouschek, A., and Ferahfi, A. R. (1991) Methods Enzymol. 202, 82-112

Fig. 10. Aminocyl residues crucial for folding. A, location within the native structure. B, element of structure enlarged. The coordinates for the 1.7-Å resolution structure were generously provided by F. Quiocio. The native aminocyl residues at the location of changes are represented. The figure was generated by the Insight II program.