Unfolding the maltose-binding protein (MBP) with optical tweezers

We have used optical tweezers to study the folding and unfolding of a single protein, the maltose-binding protein. We will present measurements on the effect of chaperone protein SecB on the (un)folding and we will present molecular dynamics simulations to explain an observed unfolding intermediate.

3.1 Introduction

In *Escherichia coli*, many different proteins are constantly being synthesized by the ribosome. These proteins can be grouped in different classes by their destination: cytosolic proteins remain in the cytosol, membrane proteins should be targeted to either the inner or outer membrane, periplasmic proteins have to be transported to the periplasm, the space between inner and outer membrane, and finally secretory proteins have to be transferred to the outside of the cell.

It is evident that the folding properties of a protein are of importance for the targeting process. Cytosolic proteins have to fold quickly in order to be able to fulfill their task in the cytosol. In the crowded milieu of a cell, folding chaperones such as GroEL [42] and DnaK [43] are often needed for folding. Membrane proteins, periplasmic and secretory proteins, however, should be kept in an unfolded state, in order to pass the narrow pore of the SecYEG translocase (see Chapter 2) to reach their destination. Hence, folding prior to translocation has to be prevented. Another issue arises with the (inner and outer) membrane proteins. These proteins carry large stretches of hydrophobic residues in their sequence, that will eventually fold into transmembrane helices, once the polypeptide has reached its destination. In the cytosol, however, these hydrophobic stretches may cause proteins to aggregate into inclusion bodies, after which the protein cannot be translocated anymore.

In *E. coli*, the folding and aggregation of proteins prior to translocation is prevented by the chaperone protein SecB. During translation of a protein—whether cytosolic or not—SecB binds to certain hydrophobic parts of the newly synthesized polypeptide, preventing the polypeptide from finding its native conformation or from aggregation into inclusion bodies. Next, these non-cytosolic proteins carrying
Figure 3.1: Schematic representation of the trapping geometry that was used in the optical tweezers experiments. Later in this chapter, the nature of the non-covalent links and the procedure to obtain this geometry will be further introduced.

A translocation signal sequence are targeted to the Sec translocase, while cytosolic proteins—without the signal sequence—are released to the cytosol, where they can fold. A more detailed explanation of the mechanism behind this selection will be given later in this chapter.

Previously, bulk assays have been performed to study the effect of SecB-binding to an unfolded polypeptide. Much has been revealed by these studies, but many questions remained unanswered, e.g., on the dynamics of the binding of SecB to a preprotein, and on the effect of SecB-binding to the configuration of an unfolded polypeptide. To answer these questions, we have—for the first time—studied the effect of SecB-binding on the folding and unfolding of a single protein, the maltose-binding protein (MBP).

To unfold (and refold) a protein, we used an optical tweezers setup employing the novel trapping construct that is shown in Figure 3.1. In this geometry, non-covalent links were used to link the C and N terminus of a protein to a polystyrene microsphere and a 920-nm dsDNA linker that was introduced to prevent interactions between the optically trapped microsphere and the protein and to prevent interference of the micropipette microsphere with the trapping beam. This trapping geometry allowed for the repeated unfolding and refolding of a single protein in a controlled way. Moreover, our geometry was much more straightforward and gave a higher yield than a recent other solution [9].

The MBP-unfolding experiments showed that the forced unfolding of MBP often occurs via one or several unfolding intermediates. Repeatedly, we observed one specific unfolding intermediate after a first partial unfolding of MBP at a force of ~16 pN. Further unfolding of MBP to a fully extended polypeptide then occurred at a force of ~25 pN. Unfolding experiments using an engineered protein construct consisting of four in-tandem repeats of MBP confirmed these observations.

For a better interpretation of the observed unfolding intermediates in the forced unfolding of MBP, we undertook steered molecular dynamics (SMD [44]) simulations. These simulations revealed a series of C-terminal surface-exposed α-helices in the structure of MBP that relatively easily detached from the MBP structure (when a force was exerted on C and N terminus in an SMD simulation). This
predicted intermediate agrees remarkably well with the experimentally observed one.

Next, we studied the effect of SecB on the unfolding and refolding of MBP by adding SecB during an optical tweezers unfolding experiment. We saw no effect of SecB binding prior to the complete unfolding of the tethered protein. This was the first direct observation showing that SecB has no affinity to native, stably folded proteins. After the forced unfolding of MBP, we observed that the refolding of the unfolded polypeptide was dramatically prevented by the presence of 0.1 µM of SecB, in an all-or-nothing manner. No single feature pointing at (partial) refolding could be observed. We will discuss the implications of this observation on the SecB-preprotein binding mechanism. Extending an unfolded polypeptide in a SecB-bound state resembled worm-like chain behavior. We will show that this suggests that in SecB-mediated protein translocation by the Sec translocase, only a fraction of the free energy from ATP hydrolysis by SecA is used for the unfolding of a protein.

We start by further introducing chaperone SecB and the protein used in our experiments and simulations, MBP. Then, we will further describe the materials and methods used in our experiments. Next, we will present the results from our experiments and simulations. We will finish this chapter by discussing the implications of our results to protein folding, the binding of SecB to a substrate protein, and to protein translocation in general.

3.1.1 The role of SecB in the targeting of proteins

Figure 3.2 shows, in a schematic way, the role of SecB in the transport of a cytosolic (1) and a periplasmic (2) protein to their respective destination. In the transport pathway of a periplasmic protein (Figure 3.2-2), SecB interacts with a newly synthesized polypeptide during translation. The concentration of SecB in the cytosol is high (~13 µM tetramer [45]) so all periplasmic preproteins will be bound during translation. SecB has no interaction with the protein translocation signal sequence, but binds to peptides in the mature part of the preprotein [46]. Binding of SecB to a polypeptide prevents it from aggregation or folding prior to translocation to the periplasm. The SecB-bound state of a preprotein is termed translocation-competent state. Note that after synthesis of a new protein, there is a competition between folding of the protein and its binding to SecB. This competition is governed by kinetic partitioning. Later in this chapter, we will give a more detailed example of this kinetic partitioning.

After binding of a preprotein to SecB, the preprotein-SecB complex diffuses to the translocase at the inner membrane. There, SecB binds to ATPase SecA with high affinity [47]. Next, also the preprotein signal sequence binds to SecA, tightening the bond between SecA and SecB. Now, the preprotein is released from SecB and delivered to SecA [48]. Upon hydrolysis of ATP by SecA, SecB is released. Next, by repeated ATP hydrolysis of SecA, the preprotein polypeptide is translocated
Figure 3.2: Protein translocation through the E. coli inner membrane shown in a step-by-step manner. The role of SecB is specifically shown.

through the membrane and will fold into its native state in the periplasm. For a review on protein translocation by the Sec translocase, see Driessen et al. [23]

Since the occurrence of high-affinity SecB-binding peptides is the same in both cytosolic and secretory proteins [46], also cytosolic proteins will associate with SecB during translation (Figure 3.2-1) and will subsequently be delivered to the translocase on the inner membrane. The absence of a signal sequence, however, will prevent the preprotein from binding to SecA and translocation cannot start [49]. Eventually, the protein will be released by SecB to the cytoplasm where it can fold into its native state, possibly aided by folding chaperones such as GroEL [42].

3.1.2 Structure of SecB

In Figure 3.3, different representations of the structure of SecB are shown [26]. SecB is a homotetrameric protein with a molecular mass of ~17 kDa that can be regarded as a dimer of dimers. The cartoon representation of Figure 3.3a shows a view on one of the two sides carrying the proposed SecA-binding site [26]. Figure 3.3b shows how channels are located between two monomers. These 70 Å-long channels likely form the peptide binding sites of SecB.

Knoblauch et al. [46] screened 2688 peptides covering sequences of 23 proteins for their SecB binding. This helped defining a binding motif in peptides that bind to SecB. A typical SecB-binding peptide consists of ~9 residues and is enriched in
Figure 3.3: Different representations of the structure of the SecB tetramer (PDB ID: 1FX3 [26]). (a) Cartoon drawing of the top view of SecB showing the β-sheet that is formed by two monomers. This surface contains the SecA-binding site. (b) Surface drawing showing the previous view rotated over 90° around a vertical axis, with the two binding channels indicated. Figures were prepared using VMD [50].

aromatic and basic residues. Acidic residues are disfavored by SecB. In binding of peptides to SecB, hydrophobic interactions contribute most to the binding free energy [51].

3.1.3 The maltose-binding protein (MBP)

In the protein unfolding experiments described in this chapter, the maltose-binding protein (MBP) of *E. coli* was used. MBP is a periplasmic protein that is part of the *E. coli* maltodextrin transport system [52], which belongs to the superfamily of the evolutionarily conserved ABC transport systems [53]. MBP transfers maltooligosaccharides such as maltose from a receptor at the outer membrane to a receptor at the inner membrane, where the maltooligosaccharides are transported to the cytoplasm by ATP-hydrolysis.

MBP is a 370-aa, 40.6 kDa protein containing a maltose binding cleft, surrounded by two globular domains, or lobes. Figure 3.4 shows the crystal structure of MBP [54]. This picture clearly shows that MBP is a bilobate protein. The maltose-binding cleft located in the middle is indicated. The contour length of a fully extended MBP polypeptide chain (i.e., with the secondary structure removed) would amount to 120 nm.

MBP is synthesized in the cytosol as a preprotein (preMBP) carrying a 26-aa signal sequence at the N terminus. MBP has been used as a model protein for bulk protein folding experiments [56, 57, 58], for determining the effect of a signal sequence on the folding of a preprotein [59, 60] and for studying SecB–protein interaction [51, 61, 62, 63, 64, 65]. The pool of knowledge gained from these studies
rendered MBP an ideal protein for our single-molecule optical tweezers study. Titration calorimetry experiments [51] showed that the free energy of stability ($\Delta G$) of the complexes between SecB and both MBP and preMBP amounts to $\sim 42$ kJ/mol or $\sim 17 \, k_B \, T$ at room temperature, equivalent to a dissociation constant $K_d$ of 30 nM for MBP (at 6.5°C) and 27 nM for preMBP at (7.7°C). Moreover, these experiments showed that MBP binds to the SecB tetramer in a 1:1 stoichiometry (i.e., 1 MBP protein binds to 1 SecB tetramer). The unfolding free energy ($\Delta G$) of MBP (at 298 K) was determined by Beena et al. [60] to be 37.2 kJ/mol ($15.0 \, k_B \, T$) for MBP and 30.5 kJ/mol ($12.3 \, k_B \, T$) for preMBP.

After synthesis of a new MBP precursor, note that there is competition between binding to SecB and folding to the native state. This competition is most likely governed by the rates of folding (in s$^{-1}$) and the rate of SecB-binding (in M$^{-1}$ s$^{-1}$) and -unbinding (in s$^{-1}$) of the preprotein. It has been proposed that there is a kinetic partitioning [66] in this competition. One can calculate that it cannot be thermodynamically determined, i.e., by free energy differences between the different states (bound/unbound; folded/unfolded). Thermodynamics dictates that a SecB concentration as high as 4 mM is required at 37°C to have an equilibrium concentration of folded preMBP equal to the concentration of preprotein/SecB complexes, assuming that no new preproteins are synthesized or translocated to the periplasm. In E. coli, the concentration of SecB is more than three orders of magnitude lower than that (~13 µM tetramer [45]) so in the presence of SecB, all MBP precursors would irreversibly fold to their native state and hence cannot be translocated to the periplasm. For the rate of folding of preMBP, a value of 0.25 s$^{-1}$ at 30°C has been estimated [67]. This value is considerably smaller than the proposed binding rate of SecB, 130 s$^{-1}$ [66], assuming that 10% of the 13 µM of SecB is available for binding. Hence, folding is a much slower process than SecB-binding. SecB tetramers will bind to a polypeptide during its folding, thereby preventing complete folding to the native state before reaching the translocase. Cytosolic proteins have a
Materials and methods

3.2 Materials and methods

3.2.1 Experimental configuration

We have used an optical tweezers setup to study the unfolding and refolding of single proteins. In the single-molecule protein unfolding field, most studies employ atomic force microscopy (AFM). In AFM studies, connections between the protein and surfaces are made through non-specific interactions. Compared to an approach using AFM, optical tweezers allow for a more specific tethering of C and N terminus of the protein. Because of the lower spring constant of the optical trap, lower unfolding forces can be detected and lower pulling rates can be reached. This way, we could refold a protein in a controlled way, after first unfolding it.

In single-molecule protein unfolding studies, surface interactions have to be minimized. During unfolding, hydrophobic parts of the polypeptide chain become exposed that might interact with the nearby surface. To avoid this problem, in AFM unfolding studies, the protein of interest is often cloned between several copies of titin immunoglobulin domains [68, 69] or a construct with multiple in-tandem repeats of the protein is used [70, 71] to increase the distance of proteins to a surface during an unfolding measurement. In an optical tweezers study done by Cecconi et al. [9], the C and N terminus of the protein of interest (ribonuclease H) were covalently linked to two 500-bp DNA spacers to prevent interaction of the protein with the surface of one of the two polystyrene microspheres that were used to

Figure 3.5: Linear representation of the sequence of premBp, showing the relative affinity of 13 aa peptides to SecB. Data adapted from Knoblauch et al. [46, fig. 2]. Each gray rectangle represents a 13 aa peptide from MBP. Left of the white line at amino acid position 26, the translocation signal sequence is shown. Four regions with high affinity to SecB are indicated at amino acid position 120, 190, 270 and 360.

higher folding rate (5–10 s⁻¹) and, moreover, are not translocated to the periplasm, hence these proteins are likely to eventually unbind from SecB (unbinding rate ~3 s⁻¹) and fold to the native state where they cannot bind to SecB anymore.

Figure 3.5 shows a linear representation of the sequence of MBP with the relative affinity of 13-aa peptides within MBP to SecB. This data was determined by peptide screening analysis, performed by Knoblauch et al. [46]. It can be seen that, as mentioned before, the signal sequence has a low affinity to SecB. Along the sequence of MBP, four regions can be distinguished (indicated with ovals) with a high occurrence of peptides with a high relative binding affinity to SecB.
exert force on the construct. In our optical tweezers experiments, we used a novel approach using a single DNA linker and non-covalent links between the protein and a microsphere and the DNA linker, respectively. Compared to the covalent links that Cecconi et al. used, this approach allows for connections that are easy to make yet are strong enough to sustain the forces needed to unfold a protein. The streptavidin-biotin connections that we used allow for a high yield and a high resistance.

Our novel trapping geometry for single-molecule protein unfolding experiments is shown schematically in Figure 3.6. Via a 4×c-myc tag at its C terminus, MBP is directly bound to microspheres coated with anti-c-myc antibodies. These microspheres are held by a micropipette. To be able to exert force on the protein, its N terminus is bound to an optically trapped microsphere via a covalently bound biotin, a streptavidin, a dsDNA linker (contour length 920 nm) with covalently bound biotin and digoxigenin molecules, and anti-digoxigenin antibodies. These connections enable a specific yet easy linking of both termini of the protein. This section will describe how the construct is established.

Experiments were performed in a flow cell as the one described in Chapter 1. For the experiments where SecB was added during the experiment, a flow cell was used with an additional input to flow in SecB-containing buffer when needed. In the tweezers setup, syringes that were connected to the left- and rightmost channels of a flow cell were filled with different microsphere suspensions: a suspension containing MBP-coated microspheres (MBP microspheres) and a suspension containing microspheres with dsDNA-tethered streptavidin tetramers (DNA microspheres). The syringe that was connected to the middle channel was filled with the reaction buffer, HEPES/KOH, pH 7.6, 100 mM KCl, 5 mM MgCl₂, supplemented with 0.1% bovine serum albumin (BSA).

Next, the flow from the three syringes was started by applying pressure pulses (see Chapter 1) and tuned such that three separate unmixed flows could be distinguished next to each other. Using manually controlled linear stages, the flow cell was moved such that the optical trap was located in the flow carrying MBP microspheres. A microsphere was trapped and transferred to the micropipette, where it was held through suction. In the same way, also a DNA microsphere was trapped and transferred to the middle of the flow cell. Using the trackball mode of the Labview program, the microspheres could now be touched together to probe for a single DNA/MBP tether (Figure 3.6-1). In the preparation of the DNA microspheres, the DNA-to-microspheres ratio was tuned such that if two microspheres were tethered, it was via a single DNA linker (Figure 3.6-2a). Multiple tethers could be recognized by the occurrence of overstretching transitions at forces higher than ~65 pN (see Appendix A). After a single DNA/MBP construct had been tethered, a folding and unfolding experiment could be started. Using the piezo stage, the micropipette could be moved with respect to the optical trap, thus increasing and decreasing the force exerted on the termini of the protein to drive its unfolding (Figure 3.6-3) or to allow refolding. For an alternative set of experiments, a protein consisting of four in-tandem repeats of MBP was used (Figure 3.6-2b).
Figure 3.6: Schematic view on some of the stages in an MBP unfolding and unfolding experiment. (1) To start an experiment, an anti-c-myc microsphere is pushed against a microsphere carrying streptavidin tetramers that were previously tethered using a dsDNA linker. (2a) After a connection has been made between the streptavidin and the biotin at the C terminus of MBP, the force exerted on the construct can be increased by moving the micropipette away from the optical trap. (2b) Alternatively, a protein consisting of four in-tandem copies of MBP could be used. (3) After unfolding of MBP, the construct consists of a coupled DNA linker and an unfolded polypeptide. Relaxing the construct allows for subsequent refolding of the polypeptide. Picture is not drawn to scale.
3. Unfolding the maltose-binding protein (MBP) with optical tweezers

Figure 3.7: Schematic illustration of the duplication strategy that was followed to make a 4×MBP protein. Two unique restriction sites with complementary cohesive ends (SalI and XhoI) were introduced on both sides of the malE gene (black) that was previously cloned in a plasmid (top left). Next, the plasmid was separately digested with Scal/XhoI (bottom left) and SalI/Scal (bottom right) and the indicated fragments were extracted from gel and ligated (top right), leading to a plasmid with a duplicated malE gene. This duplication strategy was subsequently repeated, resulting in a plasmid with four repeats of malE.

3.2.2 Cloning and protein expression of MBP constructs.

The malE gene of *E. coli* (encoding protein MBP) was cloned into plasmid pET3a, without the sequence corresponding to the translocation signal sequence. Using a polymerase chain reaction (PCR), a cysteine codon was introduced at the end of the gene corresponding to the N terminus of the protein. The sequence for a c-myc tag was introduced at the other end of the gene. Next, three more copies of this c-myc sequence were added by cycles of restriction, gel extraction of fragments and ligation.

For the 4×MBP construct, two unique restriction sites with complementary cohesive ends (SalI and XhoI) were introduced using PCR at both sides of the malE gene. Using the duplication strategy depicted in Figure 3.7, another copy of the malE gene was added, leading to a plasmid with two in-tandem copies of the malE gene. A second application of the duplication strategy led to a plasmid with four in-tandem copies of the malE gene with a single cysteine at one side and four c-myc tags at the other side of the gene.

After cloning, the proteins (MBP or 4×MBP) were expressed in *E. coli*. After lysis of the cells, MBP could be purified using the binding affinity of native MBP to an amylose resin (MBP binds to amylose with its maltose binding pocket). Also 4×MBP was successfully purified using an amylose resin, indicating that a 4×MBP polypeptide indeed folds into native, single-MBP subunits.

Finally, the cysteine at the N terminus of the MBP-constructs was biotinylated using biotin-maleimide. Free biotins were subsequently removed using dialysis.
### Materials and methods

#### Figure 3.8: Immunoblots of purified MBP, 2×MBP and 4×MBP after biotinylation. The purified MBP constructs were positively tested for the presence of a c-myc tag at the C terminus and for a biotin at the N terminus using anti-c-myc and streptavidin-AP (alkaline phosphatase), respectively.

Figure 3.8 shows two immunoblots of purified MBP, 2×MBP and 4×MBP after biotinylation. These blots show that both biotinylation and cloning of the c-myc tag were successful.

### 3.2.3 Optical Tweezers setup

For the folding and unfolding experiments, the optical tweezers setup that was presented in Chapter 1 was used. The laser diode current of the setup was kept at 9 A. By fitting a Lorentzian to the power spectral density (PSD) of the movements of a trapped microsphere (see Chapter 1), the force constant of the optical tweezers and the sensitivity of the QPD were determined every day before doing experiments. On average, the force constant for a 1.88 µm polystyrene microsphere along the x coordinate was 169 pN/µm with a standard deviation of 24 pN/µm. The sensitivity of the QPD was on average ~2.74 V/µm with a standard deviation of 0.24 V/µm. The standard deviation of trap stiffness and sensitivity can be partially explained from the polydispersity of the microsphere suspension. The root mean square (RMS) of the noise in our force measurement was 0.11 pN (measured during 1 s).

During the experiments, microsphere movements were measured by recording the normalized QPD $V_x$ and $V_y$ signals at a frequency of 50 Hz. The analog electronics anti-aliasing filter was set at a filter frequency of 20 Hz. Additionally, the Labview particle tracking algorithm was used to track microspheres at a lower frequency (~5 Hz). For the analysis and for plots, the QPD data was used. The particle tracking data was only used for calibration.

**FU sweep** For the reproducibility of the folding and unfolding experiments, a special folding/unfolding (FU) sweep mode was developed with the following design
Figure 3.9: Schematic demonstration of the folding/unfolding (FU) sweep mode that was introduced for the MBP-unfolding experiments. The top graph shows the piezo x coordinate as a function of time. The bottom graph shows the force on a trapped microsphere in the x direction, measured by the quadrant photodiode (QPD) as a function of time. (i) the piezo x coordinate is increased at a constant rate (µm/s) until a force $F_{\text{max}}$ on the tethered construct is reached. (ii) Next, the force was held at $F_{\text{max}}$ for $\Delta t$ seconds using a force-feedback mode. (iii) Now, the tether length was decreased again by moving back the piezo at a constant rate until position $x_{\text{min}}$ is reached. (iv) At this position, the piezo was kept for $\Delta t$ seconds before again increasing the tether length.

criteria in mind: (1) In forced unfolding experiments, the expected unfolding force (and also refolding) is dependent on the pulling rate (in µm/s). Hence, to keep the probability of unfolding of the protein at a certain force constant, the experiment should be performed at constant pulling rate. (2) After unfolding of a protein, refolding of a tethered polypeptide occurs after it has been held at low force for a certain period. To keep the probability of refolding between relaxing and pulling constant throughout the experiment, the time to allow the refolding should be kept fixed. (3) To avoid breaking of the construct, the exerted force should remain limited.

In Figure 3.9, the FU sweep mode is demonstrated schematically. The FU sweep consists of different phases: (I) Stretch phase: After switching on the sweep mode, the piezo x coordinate is moved at a constant rate (µm/s) such that the force increases. (II) Hold phase: Once the force on the tethered construct reaches a preset force $F_{\text{max}}$, the force is held at this force for $\Delta t$ seconds (generally set at 10 s) to make sure the protein is completely unfolded. (III) Relax phase: Now, the tether length is decreased again by moving back the piezo at a constant rate until position $x_{\text{min}}$ is reached. (IV) Slack phase: At this position, the piezo is kept for $\Delta t$ seconds before again (I) increasing the tether length.

If $F_{\text{max}}$ is set high enough (40 pN), unfolding will occur before the hold phase
is reached. Refolding can occur after again relaxing the construct (i.e., in the relax phase or the slack phase).

In the stretch and relax phases of the FU sweep, the piezo was moved with a rate of 50 nm/s. At the force range where unfolding occurs, this corresponds to a loading rate of $\sim 7$ pN/s (at these forces, the DNA is already at >95% extension and piezo movement will mostly result in movement of the trapped microsphere inside the trap where the force $F$ is proportional to the distance $x$ with trap stiffness $k_x$ ca. 150 pN/µm).

### 3.2.4 Microsphere preparation

Anti-c-myc and anti-digoxigenin (anti-DIG) antibodies (Roche Diagnostics) were covalently coupled to carboxyl-functionalized 1.87 µm-sized polystyrene microspheres (Spherotech) using the crosslinker carbodiimide. A commercially available kit including all needed buffers was used for this crosslinking reaction (Polysciences, cat. no. 19539-1). In this protocol, 100 µg of antibody was coupled to 250 µl 5% w/v microsphere suspension.

Both the anti-c-myc- and the anti-DIG-coated microspheres were blocked using bovine serum albumin (BSA, Sigma-Aldrich). Here, 20 µl of microsphere suspension was diluted in 500 µl 1× HMS (50 mM HEPES-KOH, pH 7.6, 100 mM KCl, 5 mM MgCl$_2$) with 1% w/v BSA. The microspheres were incubated with the BSA for 30 minutes at room temperature in a hand-over-hand mixer. Next, the microspheres were spun down by a table centrifuge for 1 minute at full speed after which the supernatant was removed. This blocking protocol was repeated and the microspheres were resuspended in 20 µl HMS with 0.1% w/v BSA.

Next, the anti-c-myc microspheres were coated with the biotinylated MBP-4×c-myc. First, the MBP was diluted 100–500 times in HMS/0.1% BSA. Next, 0.5 µl of the diluted MBP and 2 µl of anti-c-myc microspheres were diluted in 20 µl of HMS/0.1% BSA. The microspheres were incubated with the MBP for 30 minutes at 4°C in a hand-over-hand mixer. Next, to remove unbound MBP, microspheres were washed in 500 µl HMS/0.1% BSA and, after spinning the microspheres down, resuspended in 400 µl HMS/0.1% BSA.

In the MBP-unfolding experiments, a DNA linker was used with a contour length of 920 nm. This linker was created by performing a polymerase chain reaction (PCR) on pUC19 DNA with a forward primer carrying two biotin groups and a reverse primer carrying two DIG groups. This DNA linker was bound to streptavidin (SA) by mixing ~250 ng of DNA with 0.4 µg/ml streptavidin (Molecular Probes) in 10 µl HMS/0.1% BSA. This is an excess of streptavidin, hence every streptavidin tetramer has at most one DNA linker bound.

Next, the DNA/SA mixture was added to 2 µl anti-DIG microspheres diluted in 10 µl HMS/0.1% BSA. After incubating the DNA and the microspheres for 30 minutes at 4°C in a hand-over-hand mixer, 400 µl of HMS/0.1% BSA was added.
3.2.5 Steered molecular dynamics studies

SMD simulations were performed at AMOLF by Harald Tepper. For the SMD simulations, the crystal structure of MBP without a bound sugar (PDB-ID 1jw4 [54]) was used as a starting structure. Missing atoms were added after analysis of the PDB structure using the SCWRL program [72]. Moreover, some side chains were flipped as suggested by this program. Next, several histidines were replaced by an HSD (neutral histidine with a proton on the N_δ) and hydrogens were added to the protein using psfgen.

Next, the protein and its crystallographic water molecules were solvated in a previously equilibrated water box, extending at least 10 Å from the protein in all directions. All water molecules within 2.4 Å from the atoms in the crystal structure were removed. Subsequently, the system was neutralized by replacing 8 water molecules by Na^+ ions. The final box contained 5378 protein atoms, 13,613 water molecules and 8 counterions (46,217 atoms in total). To relax unphysical contacts between protein and water, the system was energy-minimized for 500 steps, followed by 2000 time steps (4 ps) of canonical MD (298 K) with the protein atoms fixed in space, and 10,000 steps (20 ps) with all atoms free. The pressure at the end of this stage was 700 bar. In a series of consecutive constant pressure simulations (10,000 steps each), the system was brought to a pressure of 1 bar. From this stage, a further extensive equilibration run was launched (T = 298 K, p = 1 bar) that lasted 1 · 10^6 steps (2 ns). After ~0.5 ns, the total energy and volume had become stable. The final box dimensions measured 68.03930 × 76.66410 × 87.2054 Å^3.

For the MD simulations, the program NAMD [73] was used with force field CHARMM22 [74] for the protein and the TIP3P model [75] for the water molecules. The long-range electrostatic interactions were treated with the Particle Mesh Ewald (PME) technique, with a tolerance of 10^{-6} and a Fourier grid size of 72 × 81 × 90. Van der Waals interactions were cut off at 10 Å with smooth interpolation to 0 from 9 Å. Full electrostatics were calculated every 2nd time step. A Langevin thermostat was applied to the non-H atoms, with a damping coefficient of 5 per ps. All bonds to hydrogens were fixed.

To induce unfolding in the simulated protein, two springs (with force constant 5 kcal/mol/Å², or 3.5 · 10^6 pN/μm) were attached to the N of Lys1 and the Cα of Lys370, respectively. In the unfolding simulations, one of the two springs was moved along the vector between the initial positions of the two atoms while the other spring was kept fixed. In our study, we used pulling rates of 0.1 nm/ns (slow) or 1 nm/ns (fast). For analysis purposes, configuration files were written to disk every 1000 time steps and spring forces every time step.
3.3 Results

3.3.1 Forced unfolding of MBP using optical tweezers

To characterize the forced unfolding of MBP, an MBP/DNA construct was tethered between two microspheres, as described in the previous section. By starting the FU piezo sweep mode, an unfolding experiment was started. In Figure 3.10, an example is shown of the force–extension curve resulting from such an experiment showing three consecutive cycles of the FU sweep (stretch→hold→relax→slack, etc.) in increasingly lighter shades of gray.

The triangles (▽) show the force–extension of an extensible worm-like chain (WLC, see Appendix A) corresponding to the dsDNA linker with a persistence length \( P = 53 \text{ nm} \) and an elastic stretch modulus \( S \) of 1200 pN. The crosses (×) show the force–extension of two coupled worm-like chains: an extensible WLC corresponding to DNA with properties as mentioned before and an inextensible WLC corresponding to a polypeptide with contour length \( L = 120 \text{ nm} \). For the polypeptide WLC, a persistence length of 1.5 nm was chosen as
it gave the best fit to the data.

The first stretch curve (black) in Figure 3.10 first follows the DNA WLC curve. At a force of \(~29\) pN, a sudden change in force and extension can be observed, indicating an unfolding event of the tethered protein. By fitting DNA+polypeptide WLC curves with varying polypeptide length, it was determined that the additional contour length of the polypeptide due to this unfolding event amounts to \(~90\) nm (cf. □-curve). When the force is further increased to a force of \(~24\) pN, the protein further unfolds. The total polypeptide contour length due to this unfolding event is 120 nm (cf. ∗-curve). This length corresponds to a fully extended MBP polypeptide chain (taking 3.3 Å per amino acid (9)) indicating that here, the protein has completely unfolded to a conformation with all secondary structure removed. Apparently, in the first unfolding event, a part of the MBP structure corresponding to 90 nm/120 nm×370 amino acids = \(~278\) amino acids has unfolded.

After complete unfolding, the first relax curve (black) follows the DNA+polypeptide WLC curve until the slack phase (10 s waiting time at low force) of the FU sweep is reached, indicating that MBP did not significantly refold during that stage.

The second stretch curve (dark gray) first follows the DNA WLC curve again, indicating that the polypeptide has completely folded in the 10 s waiting time. At \(~7\) pN, the protein unfolds to an unfolding intermediate with a polypeptide contour length of 28 nm (cf. ○-curve), corresponding to \(~85\) amino acids. Eventually, at \(~32\) pN, this intermediate also completely unfolds. After relaxing the construct again, the third stretch curve shows that the polypeptide apparently only partially refolded to the same intermediate that was reached in the second cycle. Similar to the first relax curve, the second and third relax curves showed no refolding.

Next, the unfolding force of 94 unfolding events from 27 unfolded MBP molecules were analyzed. The histogram that is shown in Figure 3.11 gives an overview of the observed unfolding forces. If unfolding occurred via one or several intermediates, the unfolding event leading to the highest additional polypeptide contour length was analyzed for this histogram. In \(>90\)% of the analyzed stretch curves, these events corresponded to the unfolding event with the highest unfolding force. In \(>65\)% of the cases, these events led to the complete unfolding of the protein. The average unfolding force for the analyzed unfolding events was at \(25.4\) pN with standard deviation \(8.6\) pN. The spread in the force can be partially explained from
the stochastic nature of protein unfolding. Additionally, inaccuracy of the trap stiffness parameter contributed to the width of the distribution.

To study whether MBP is folding back to its native state after having unfolded it, the unfolding force as a function of the number of times a protein has been unfolded was studied. Figure 3.12 shows the development of the unfolding force of MBP proteins that were unfolded multiple times. For the graphs in this figure, we used the same dataset as for the unfolding forces histogram in Figure 3.11. It can be observed that the first unfolding event of a molecule occurs at the same force as later unfolding events. No trend up- or downward can be observed, indicating that refolding happens to the same state every time. This state is thus a reproducible starting state before pulling and is most likely the native state of the protein.

Next, the same set of measurements was analyzed for the occurrence of intermediates. All measurements gave clear evidence of intermediates that are stable for seconds under load. An unfolding intermediate with a polypeptide contour length of ~25–30 nm, as in the above example, was observed in all measurements. The mean unfolding force for 17 unfolding events leading to this intermediate was 16.5 pN with standard deviation 11.1 pN.

The refolding of the polypeptide to the native state only occurred in the slack phase of the FU sweep, in all of the analyzed measurements except one. In one of the measurements, however, refolding was observed in the relax phase of the FU sweep. Refolding was visible as an increase of force because of the increased tension due to the shortening of the tether. The event is shown in Figure 3.13.

Several control experiments were performed to confirm that the observed phenomena indeed originated from unfolding of a tethered MBP protein, and not from artifacts stemming from the other components used in the experimental configuration: (1) To check whether the connection between the protein and the micropipette microsphere is via the 4×c-myc tag and the anti-c-myc antibody and not through a specific interaction between the protein and the microsphere, a

![Figure 3.12: Development of the unfolding force of MBP. Every bar represents the unfolding force of the nth unfolding event, averaged over all measurements with a total number of unfolding events nmax before the end of a measurement.](image-url)
control experiment was performed with anti-His\textsubscript{6} microspheres instead of anti-c-myc microspheres. (Microspheres were prepared with anti-His\textsubscript{6} antibodies [Roche Diagnostics] as described in §3.2.4). In this control experiment, the observed tethers broke at very low force, indicating that non-specific protein-antibody interactions are not strong enough to sustain forces needed to unfold a protein. (11) To ensure that the observed drop in force was indeed from unfolding of the MPB protein, a control experiment was performed using biotinylated c-myc epitopes (a 10-aa polypeptide). In this experiment, tethers were observed that could be extended to overstetching of the DNA linker (>60 pN), but without sudden force drops, indicating that the observed force drops were from unfolding of MPB.

### 3.3.2 Forced unfolding of a 4\times MBP construct.

To further investigate the unfolding behavior of MBP, we performed unfolding-and-refolding experiments with the 4\times MBP construct that was described in §3.2.2. For a 4\times MBP construct, the distance between each of the MBP subunits and the polystyrene surface is higher, so less surface effects are expected. The protein was tethered between a DNA spacer and a micropipette microsphere as described in §3.2.1 and an experiment was started by switching on the FU sweep.
Results

Figure 3.14: Forced unfolding of 4×MBP using optical tweezers. A stretch and relax curve are shown. Four major unfolding events can be observed (labeled 1st–4th), corresponding to unfolding of each of the four MBP subunits. Moreover, a hump-like feature can be seen at extension 0.8 µm. WLC curves were fit to estimate contour lengths.

An example of an unfolding measurement of 4×MBP is shown in Figure 3.14. Following the stretch curve of the measurement, the force-extension curve first resembles DNA WLC behavior. At a force of ~10 pN, a rugged force plateau or ‘hump’ can be observed that extends until an additional ~100 nm extension. Now, the force further increases and four major unfolding events (labeled 1st–4th) can be seen at forces between 15 and 21 pN. WLC curves were fit to the observed intermediate in order to estimate contour lengths. These fits are shown in the figure.

The observed ‘hump’ at a force of 10 pN was observed in most of the 4×MBP-unfolding measurements. Careful examination of the hump shows multiple small force drops, resembling unfolding of domains each leading to a several-nm increase of contour length. These events resemble the ~16.5 pN/~25–30 nm unfolding events that were observed in the single-MBP case. A likely explanation for the ‘hump’ is the unfolding of four of these easily unfoldable domains in the MBP subunits that unfold prior to the complete unfolding of the rest of each of the proteins. Figure 3.14 shows a WLC fit to the intermediate that is reached after the hump. This intermediate leads to an additional polypeptide contour length of 112 nm, or four times 28 nm, which fits perfectly in the range of the first intermediate which we found to be between 25
and 30 nm. In the next section we will show molecular dynamics simulations that suggest a molecular mechanism for these low-force unfolding events.

Figure 3.15 shows a zoomed-in representation of a hump feature in a different measurement. After the hump was observed at ~10 pN, the force was decreased again by moving the micropipette. First, one can see that the force–extension curve follows a DNA/polypeptide WLC curve with a polypeptide contour length of 4×28 nm = 112 nm. At a force of 8 pN and lower, a decrease in extension could be observed, that was more than what could be predicted by the WLC model (curve). This can be explained by refolding of the domains that previously unfolded in the hump feature.

The four labeled unfolding events in Figure 3.14 likely represent the consecutive unfolding of the rest of each of the MBP subunits. All four unfolding events led to a similar increase in polypeptide contour length and happen at comparable forces. The unfolding forces for 26 of such unfolding events are plotted in the histogram in Figure 3.16. These unfolding force correspond very well to unfolding forces observed for the single-MBP experiments (see the histogram in Figure 3.11).

Consistent with the single MBP studies, no sudden refolding events can be observed in the subsequent relax curve. This was the case in all other measurements that were performed on this 4×MBP construct.

No refolding as in the single-MBP case was observed after the 10 s waiting time in the slack phase. If the extension was again increased in a second stretch sweep, the contour length of the unfolded polypeptide appeared to have become shorter,
indicating that refolding did happen to some degree. However, often the construct could be extended to overstretching of the DNA linker without any unfolding occurring, possibly indicating that the different MBP subunits had formed one big aggregate that was too stable to unfold.

In several cases, refolding of some of the MBP subunits was observed, however. Figure 3.17 shows an experiment in which refolding was observed multiple times. It shows three cycles of the Fü piezo sweep in increasingly lighter shades of gray. In the first stretch curve, the unfolding of three MBP subunits can be observed. In the next stretch curve, it can be seen that one of the three previously unfolded MBP subunits has refolded: two unfolding events (at 4 pN and at 14 pN) can be observed before the construct reaches its full extension. The third stretch curve shows the same pattern, indicating that, again, two of the four MBP subunits have refolded.

These experiments also confirm that in our novel single-DNA linker approach, the surface has no detectable influence on the unfolding and refolding of a protein. In the 4×MBP case, the chain of MBP subunits has to be pulled away from the surface, before they can be unfolded. In the subsequent unfolding, three of the four MBP subunits will unfold at a distance from the microsphere larger than the size of one MBP subunit (5 nm). Hence the polystyrene surface is too far away from the protein to affect its unfolding. The unfolding forces found in the unfolding of 4×MBP fall well within the distribution of unfolding forces that was found for MBP (see Figure 3.11). Moreover, similar unfolding intermediates can be observed. This strongly suggests that also in the single-MBP case, unfolding happens without significant influences from the nearby polystyrene surface. Moreover, the results summarized in Figure 3.12 indicate that also the folding of a protein can occur unaffected by the nearby surface. Apparently, the anti-c-myc antibody that is connected to the C terminus of the protein provides enough distance to the microsphere surface. Moreover, the constant tension at which the construct is held during a measurement, orients the molecule away from the surface. Hence, the experiments shown in this section support the use of our approach in single-molecule protein unfolding experiments.
3. Unfolding the maltose-binding protein (MBP) with optical tweezers

Figure 3.17: Forced unfolding and refolding of 4xMBP using optical tweezers. Force-extension curves from three consecutive FU cycles are shown. Refolding of two MBP subunits can be observed between the first and the second cycle and between the second and the third.

3.3.3 Steered molecular dynamics simulations on the forced unfolding of MBP

The single-molecule unfolding experiments on the maltose-binding protein that were presented in the previous section showed unfolding intermediates. Some of these intermediates were remarkably strong and were stable for up to seconds under a load of sometimes more than 30 pN. From the visual inspection of the MBP crystal structure (shown in Figure 3.4), no obvious intermediates could be predicted. Also previous bulk unfolding studies on the maltose-binding protein did not point to stable unfolding intermediates. The bilobate MBP structure suggests a stepwise unfolding mechanism, in which both lobes unfold separately. Close examination, however, shows that both lobes are highly entangled and the polypeptide chain crosses the 'bridge' between the two lobes several times. It is thus likely that separate unfolding of a lobe will cause the unfolding of the full protein.

To better explain the observed unfolding intermediates, we undertook steered molecular dynamics (SMD) simulations. In steered molecular dynamics, external forces can be exerted on individual atoms in a molecular-dynamics simulation to drive—on the timescale of MD simulations—rare events such as protein trans-
location [76] and protein unfolding [77, 68]. We performed MD simulations on the unfolding of the MBP structure, while exerting an external force on C and N termini. This was done by attaching virtual springs with a fixed spring constant to both termini and increasing the force by moving one or both springs. Note that to be able to observe an unfolding event within the timescales typical for MD simulations using a high-end, multi-node computer—tens of nanoseconds—one must move the spring with a speed that results in a pulling rate (in pN/s) that is many orders of magnitude higher than in optical tweezers experiments (≈10^{12} pN/s vs ≈10 pN/s). The discrepancy between the sampled timescales in simulation and experiment makes that we have to be cautious with the interpretation of the results. We will show that by pulling on both ends of the molecules and by comparing two different pulling speeds, we can still draw reasonable conclusions about the order in which the different residues will unfold when force is applied to MBP.

In Figure 3.18, a number of snapshots is shown from two SMD simulations that were performed at a high pulling rate (3.5 \cdot 10^{12} pN/s). The four snapshots on the left side show an experiment where the spring that was attached to the C terminus was moved at a speed of 1 nm/\text{ns}. In the four snapshots on the right side, the moving spring was attached to the N terminus. Close examination of the C-terminal structure (see, e.g., Figure 3.20) shows a surface-exposed α-helix. At the N terminus, the first five amino acids constitute a loop domain that is coupled to a buried β-strand. Figure 3.18 shows that, when pulling from the C terminus, the first α-helix can detach from the MBP structure within nanoseconds, while the N-terminal loop is still attached to the surface. When pulling from the N terminus, it is the loop domain that detaches first from the structure. After 3–4 ns, also the C-terminal α-helix detaches, while the first N-terminal β-strand remains buried and the hydrogen-bonded to neighboring β-strands are maintained.

The difference between the two experiments seems remarkable. Because of Newton's third law, one would expect the force exerted on the termini to be equal in magnitude and opposite in direction. Unfolding behavior is then independent of the position of the moving spring. Instead, we do see a difference. This shows that the effect of the force on one of the two termini propagates through the protein at a finite speed. The other protein terminus will only be affected by a force after a certain time.

During the SMD simulations, the forces exerted on the two termini were recorded. Figure 3.19 shows the forces that were exerted on the termini during the experiment that was shown in Figure 3.18. In both graphs, the light gray curves on the background shows the force signal as it was recorded during the simulations. Note that the RMS of this signal—that is mostly caused by collisions with the surrounding water molecules—is enormous (≈100 pN) compared to the forces measured in optical tweezers experiments. On longer timescales, the average effect of this thermal force is nearly zero. After processing the raw force data using a Savitzky-Golay smoothing algorithm, we can analyze the force exerted by the springs. In both graphs, the black curve shows the force exerted on the C terminus, projected on the line connecting the two termini at $t = 0$. The dark gray curve shows the force exerted
Figure 3.18: Steered molecular dynamics (SMD) simulations on the forced unfolding of MBP. Two separate experiments are shown with each four snapshots. On the left side, the spring attached to the C terminus is moved at a speed of 1 nm/ns (see arrow) while the other spring is held fixed. On the right, the spring attached to the N terminus is moved. Figures were prepared using vmd [50].
Results

![Graphs showing forces exerted on C and N terminus](a) and (b).]

**Figure 3.19:** Forces exerted on C and N terminus during the smD experiments that were shown in Figure 3.18. (a) Pulling from C terminus. The light gray signal is the force as it was measured. The dark gray and the black curves are the forces exerted on the N terminus and the C terminus respectively after smoothing the raw (light gray) data using a Savitzky-Golay smoothing algorithm. (b) Pulling from N terminus. The forces shown in Figure 3.19a must be averaged with forces from additional experiments to anneal out viscous effects in the measured forces. In Figure 3.19b, we can see a significant rise in the force exerted on the N terminus starting at ~2 ns. The force rises until the value is around 700 pN. Comparison with the snapshots in Figure 3.18 shows that the rise in force likely corresponds to the moment when the easily unfoldable N-terminal loop has detached and force is mainly used to translate the protein and to detach the C-terminal α-helix at the other side of the protein.

These simulations clearly show that the C-terminal α-helix can be detached from the MBP structure at forces that are several times lower than those needed for the eventual disruption of the N-terminal β-strand from the structure. In fact, we did not see such a disruption.

Because of the finite size of the simulation box with periodic boundary conditions, it was not possible to drive extension of the protein further than what is shown above. Instead, we developed a new technique to avoid problems with the finite size of the simulation box. We performed a series of smD simulations with progressively truncated MBP primary structures. After each smD simulation, we removed those residues from the PDB structure that detached most easily, filled the resulting gaps with water molecules and performed a new pulling smD simulation. Thus, we progressively removed the sequence that is indicated in black in the structure representation of MBP that is shown in Figure 3.20 (corresponding to amino acids 1–5 and 285–370). The carbon atoms of the residues at the positions where the protein was truncated are indicated as white spheres. Between these amino acids, the structure mostly consisted of surface-exposed α helices. The sequence between amino acids Gly327 and Met336 was partially covered by a hairpin-like
Figure 3.20: Cartoon representation of the MBP structure. Amino acids 1–5 and 285–370 are indicated in black. The C atoms of the residues at which the protein was truncated before SMD simulations are shown as white spheres. Figure was prepared using vmd [50].

structure, but in the simulations, this structure was shown to be flexible enough to allow detachment of the sequence shown in black.

These simulations suggest that in the unfolding of MBP, it is amino acids 285–370 that will detach from the protein in a first unfolding event. Earlier bulk mutation studies [58] support the existence of a more stable structure beyond this sequence. These studies showed that in the adjoining α helix (amino acids 269–285) two amino acids (276 and 283) are located that are crucial for folding [58]. Detachment of this α-helix from the MBP structure might lead to global instability of the protein and hence to complete unfolding. Indeed, an unfolding simulation (at high pulling rate) with also amino acids 269–285 removed did lead to global loss of tertiary structure. A simulation at a lower pulling rate should be performed to confirm this. In an extended conformation without secondary structure, detached amino acids 285–370 correspond to a polypeptide contour length of ~28 nm (using 0.33 nm as the contour length per amino acids). In our optical tweezers studies, we observed a predominant unfolding intermediate having a remarkably similar polypeptide contour length. Long equilibration simulations are underway for the different truncated constructs to test their long-term stability.
3.3.4 The effect of chaperone SecB on the forced unfolding of MBP

To study the effect of chaperone protein SecB on the unfolding and the folding of a protein, we used the single-MBP construct that was used in the unfolding-refolding experiments described in §3.3.1. These experiments showed that this protein can repeatedly be unfolded and refolded to its native state, making it an ideal substrate protein for studies of the effect of SecB.

After performing a normal unfolding-refolding experiment to ensure that a single MBP protein was tethered and that it could be unfolded completely and repeatedly refolded to its native state, the construct was held at a force low enough to allow the protein to refold and high enough to extend the DNA to ~80% of its contour length to avoid microsphere-microsphere interactions. Now, buffer with a nearly-physiological concentration of SecB (HMS/0.1% BSA with 0.1 μM SecB added) was introduced via a fourth input in the flow cell. Once the SecB-containing buffer had reached the tethered protein, another FU sweep was started.

In Figure 3.21, an example is shown of an unfolding-refolding experiment after the addition of SecB. The first stretch curve follows the DNA WLC curve, until a first
Figure 3.22: Unfolding of 4×MBP in the presence of SecB. Two consecutive stretch-relax cycles are shown. Stretch curves are shown as uninterrupted lines; relax curves are shown as interrupted lines. Different calculated WLC curves are shown to suggest the contour length of the unfolded polypeptide (in nm) of each of the observed intermediates.

minor unfolding event at ~10 pN. Complete unfolding of the resulting intermediate occurs at ~22 pN, which agrees with our observations described in §3.3.1. This shows that MBP remains in its native state when SecB is added. As in the absence of SecB, the subsequent relax curve follows the DNA+polypeptide WLC curve. Remarkably, the 2nd stretch curve again follows this DNA+polypeptide WLC curve, showing that no refolding of the MBP occurred in the slack phase. Also the subsequent stretch curves showed this same behavior. In all the measurements performed in the presence of SecB, careful analysis of the stretch curves also did not reveal any low-force unfolding events that could be explained from a lowering of the unfolding force due to the effect of SecB.

After a second buffer replacement to remove the SecB, often some of the connections of a tethered construct would break due to the strong flow. Hence, We could not perform a measurement to show the eventual unbinding of SecB from MBP as of yet.

Next, we performed unfolding experiments with the 4×MBP construct with SecB present continuously. Figure 3.22 shows an example of a measurement. This figure shows two consecutive stretch-relax cycles, measured in a buffer containing 0.1 μM
SecB. The first stretch curve shows similar features to the measurements that were discussed in §3.3.2: first the hump at \(\sim 10\) pN, and then four consecutive unfolding events at \(20-25\) pN. Different calculated WLC curves are drawn, showing that the observed measurements can well be described by the simultaneous unfolding of \(4\times85\) amino acids \(4\times28 = 112\) nm in a first low-force event, the hump, followed by the sequential unfolding of the rest of each MBP protein (92 nm) at a higher force.

After relaxing the construct, the second stretch curve shows no refolding of each of the four MBP proteins. The stretch curve follows the earlier relax curve. This is a behavior that is quite different from the situation where SecB was absent. There, the protein construct would show refolding or—more often—irreversible aggregation. Hence, this measurement confirms that indeed, SecB prevents folding of a polypeptide as well as the irreversible aggregation of multiple polypeptides.

The relax curves and the second pull curve in Figure 3.22 show a striking deviation at low force from the theoretical curve of a DNA/polypeptide WLC with a polypeptide contour length corresponding to the fully unfolded \(4\times\text{MBP}\) polypeptide (+). At a force of \(5\) pN, the extension of the construct is shorter than what can be predicted by the worm-like chain model (see Appendix A). This indicates that here, the unfolded polypeptide is compacted more than what can be explained by entropic fluctuations of the polymer. A possible explanation lies in the formation of secondary structure (\(\alpha\) helices) or condensation of (parts of) the polypeptide chain due to hydrophobic interactions between nearby amino acids. The similarity between stretch and relax curves suggests that this is a reversible process; no hysteresis can be observed. We observed similar features in the single-MBP measurements. Close examination of Figure 3.14 (i.e., in the absence of SecB) shows similar features, showing that this effect is not SecB-specific. Comparative analysis should be performed to show whether this extra compaction is enhanced or attenuated by the presence of SecB.

### 3.4 Discussion

We have used optical tweezers to study the effect of *E. coli* chaperone protein SecB on the forced unfolding and refolding of a protein, the maltose-binding protein (MBP). In the unfolding of MBP, we observed unfolding intermediates. We employed steered molecular dynamics (SMD) simulations to find an explanation for one of these intermediates. After the forced unfolding of a protein, it could repeatedly be refolded by decreasing the force exerted on C and N termini. After adding SecB, this refolding was quenched, a result that is consistent with previous bulk experiments.

In this section, we will further discuss our results. We will relate our unfolding experiments to previous single-molecule unfolding experiments performed on other proteins. Furthermore, we will discuss our results on the effect of SecB and will discuss the implications on protein translocation and on protein (un)folding.
### 3. Unfolding the maltose-binding protein (MBP) with optical tweezers

#### 3.4.1 Unfolding forces

In our MBP-unfolding experiments in the absence of SecB, we found an average value of \( \sim 25 \text{ pN} \pm 8.6 \text{ pN} \) for the unfolding force of MBP, at an unfolding rate of \( 7 \text{ pN/s} \). This force was compared to the unfolding force of three other proteins. The results are summarized in Table 3.1. For tenascin and spectrin, forces were measured using atomic force microscopy (AFM); for ribonuclease H, an optical tweezers setup similar to ours was used. The published unfolding forces were extrapolated to a pulling rate of \( 7 \text{ pN/s} \) using the relation \( F \propto \log v_F \) where \( F \) is the unfolding force and \( v_F \) is the pulling rate in pN/s. It can be seen that the unfolding force for MBP lies in the same range as the other unfolding forces summarized in Table 3.1.

In bulk unfolding studies using chemical denaturation, the folding stability of a protein is expressed in the free energy difference \( \Delta G_{D \rightarrow N} \) between the denatured state and the native state. Beena et al. [60] found for MBP a \( \Delta G_{D \rightarrow N} \) equal to \( -37.2 \text{ kJ/mol} \) at room temperature using guanidinium chloride. This corresponds to \( -15.0 \ k_B T \) at room temperature. To directly compare this value to our experiments, we considered the thermodynamics of an unfolding event as we observe it in our measurements.

Figure 3.23a shows a schematic representation of the model that was used in this analysis. The microsphere on the right is held by an optical trap. The total extension \( x_{\text{tot}} \) is increased by moving the micropipette (not shown) that holds the left microsphere (r→1). As a result, the extension \( x \) of the DNA/protein construct increases and \( x_{\text{trap}} \) increases, resulting in a higher force exerted on the construct. If the force \( F \) is high enough, the protein will unfold (1→2). Figure 3.23b shows a force–extension curve \( F(x) \) with a hypothetical unfolding event 1→2. In this graph, states 1 and 2 are as indicated in Figure 3.23a. During the unfolding event, the trapped microsphere moves to a position that is closer to the center of the optical trap potential well. The decrease in potential energy of the trapped microsphere is converted to work performed on the trapped construct (DNA+protein). The work \( W_{\text{tot}} \) performed by the microsphere is equal to:

\[
W_{\text{tot}} = \int_{x_1}^{x_2} F \, dx = \frac{1}{2} (x_2 - x_1)(F_2 + F_1),
\]

(3.1)

<table>
<thead>
<tr>
<th>Protein [ref.]</th>
<th>( F_{\text{unfold}} ) (pN)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MBP</td>
<td>25</td>
</tr>
<tr>
<td>Tenascin [78]</td>
<td>47</td>
</tr>
<tr>
<td>Spectrin [79]</td>
<td>22</td>
</tr>
<tr>
<td>Ribonuclease H [9]</td>
<td>17</td>
</tr>
</tbody>
</table>

Table 3.1: Comparison of unfolding forces for several proteins, extrapolated to a pulling rate of \( 7 \text{ pN/s} \).
Figure 3.23: Schematic view of all the parameters used in the energy analysis that was performed to study the MBP-SecB binding mechanism. (a) Experimental configuration immediately before (1) and after (2) an unfolding event. The micropipette microsphere is shown on the left, the trapped microsphere is shown on the right, with the position of the trap microsphere at zero force indicated as an interrupted circle. ‘r’ shows the relaxed conformation of the construct. (b) Schematic force–extension curve showing an unfolding event and showing the work $W_{\text{total}}$ that is returned to the machine in an unfolding event.

which is graphically equivalent to the gray area indicated in Figure 3.23b.

This work is put in denaturing the protein and in the subsequent extension of the DNA/polypeptide tether from $x_1$ to $x_2$. Hence:

$$W^{\text{tot}} = W^{u} + W^{\text{ext}},$$

with $W^{u}$ equal to the work needed to denature the protein and $W^{\text{ext}}$ equal to the work needed for the subsequent extension of the polypeptide and the DNA from a cumulative extension $x$ equal to $x_1$ to an extension of $x_2$.

At the pulling rate that was used in our experiments, $W^{\text{ext}}$ is approximately reversible (i.e., the work required to extend the DNA/polypeptide construct from $x_1$ to $x_2$ equals the work performed by the construct on the trapped microsphere when relaxing the construct over the same distance.). It can be calculated as follows:

$$W^{\text{ext}} = \Delta G^{\text{ext}}_2 - \Delta G^{\text{ext}}_1,$$

with $\Delta G^{\text{ext}}_2$ equal to the work needed to extend a DNA/polypeptide construct from extension 0 to extension $x_2$ and $\Delta G^{\text{ext}}_1$ the work required to extend the DNA linker
from 0 to $x_t$, respectively. Both $\Delta G_2^{ext}$ and $\Delta G_1^{ext}$ are indicated in Figure 3.23b. They can be calculated by numerically integrating appropriate WLC force–extension relations (see Appendix A).

The unfolding work $W^u$ can be expressed in terms of a reversible part $W_{rev}^u$ and a dissipative part $W_{dis}^u$:

$$W^u = W_{rev}^u + W_{dis}^u.$$

In this equation, the reversible part $W_{rev}^u$ resembles the free energy difference between the native and the denatured state $\Delta G_{N\rightarrow D}$, as determined by bulk reactions. The second law of thermodynamics states that the average dissipative work $\langle W_{dis}^u \rangle \geq 0$, but note that in rare events, $W_{dis}^u$ can be smaller than 0, i.e., when unfolding is partially driven by energy gained from the thermal bath.

Without any force applied, the unfolding rate of MBP is very low ($k_{unfolding} = 0.003$ s$^{-1}$ [60]). In our experiments, we observe unfolding in seconds, indicating that we unfold a protein faster than its internal dynamics. Hence, we expect dissipative effects. Let us now assume an experiment with an infinitesimally small pulling rate $v_F$. Now, we expect a $W_{dis}^u$ that can be close to zero. Without dissipative effects, Eq. 3.2 reduces to:

$$W^{tot} = W_{rev}^u.$$

Hence:

$$W^{tot} - W^{ext} \approx \Delta G_{D\rightarrow N}.$$  

For a quantitative analysis of these relations, we calculated energies $W^{tot}$ and $W^{ext}$ for an unfolding event at a given value of the experimental parameter $x_{tot}$. First we calculated both the force $F$ and all the parameters $x$ that are shown in Figure 3.23a, as a function of $x_{tot}$, using (see Figure 3.23a):

$$x_{d,1} + x_{trap,1} = x_{d,2} + x_{p,2} + x_{trap,2} = x_{tot},$$

with $x_d$, $x_{trap}$, $x_p$, and $x_{tot}$ as indicated in Figure 3.23a, subscripts 1 and 2 denoting the folded and the unfolded state of the system, respectively, and force–extension relations (see Appendix A):

$$\frac{x_d}{L_d} = 1 - \frac{1}{2} \left( \frac{k_B T}{F_d p_d} \right)^{1/2} + \frac{F_d}{S_d},$$

$$\frac{F_p p_p}{k_B T} = \frac{1}{4} \left( 1 - \frac{x_p}{L_p} \right)^{-2} - \frac{1}{4} \frac{x_p}{L_p},$$

$$F_{trap} = k_{trap} \cdot x_{trap},$$

$$F_d = F_p = F_{trap} = F,$$

with contour lengths $L_d = 920$ nm and $L_p = 120$ nm, persistence lengths $p_d = 53$ nm and $p_p = 1$ nm, temperature $T = 298$ K, DNA elastic stretch modulus $S_d = 1200$ pN and trap stiffness $k_{trap} = 150$ pN/s.
Figure 3.24: Calculated thermodynamic parameters of the experiment sketched in Figure 3.23. The work $W^{\text{tot}}$ performed on the tethered construct by the trapped microsphere is shown as a function of $x_{\text{tot}}$. Furthermore, the work $W^{\text{ext}}$ needed to extend the DNA/polypeptide construct from extension $x_1$ to $x_2$ is shown. An additional curve $W^{\text{tot}} - W^{\text{ext}}$ is shown.

The energies $W^{\text{tot}}$ and $W^{\text{ext}}$ could now be calculated using Eq. 3.1 and numerical integration of Eqs 3.8 and 3.9. In Figure 3.24, these energies are shown as a function of $x_{\text{tot}}$, together with their difference, $W^{\text{tot}} - W^{\text{ext}}$. Note that according to Eq. 3.2 this difference equals the total work needed for the unfolding of a protein (i.e., both reversible and dissipative work). $\Delta G_{\text{N-D}} = 15.0 \ k_B T$ [60] is indicated in the graph as a reference. Note that $W^{\text{ext}} < 0$ for $x_{\text{tot}} < 1.03 \ \mu m$. After the unfolding event, the DNA extension has decreased, resulting in an increase of conformational entropy in the DNA that is larger than the decrease of entropy due to extension of the unfolded polypeptide.

At $x_{\text{tot}} = 0.89 \ \mu m$, it can be seen that $W^{\text{tot}} - W^{\text{ext}}$ equals $\Delta G_{\text{N-D}}$. This is the point where Eq. 3.6 holds, i.e., where the free energy increase due to the denaturation of a protein is fully compensated by a cumulative free energy loss of the DNA/polypeptide construct and the microsphere in the optical trap (and vice versa). The system is in equilibrium here. Now, it is interesting to see what the unfolding force would be in such an equilibrium unfolding event. Figure 3.25 shows the relation between $W^{\text{tot}} - W^{\text{ext}}$ and the force $F_i$ at which an unfolding event will take place. In this graph, it can be seen that in the absence of dissipative effects, the expected unfolding force $F_i$ equals 3.7 pN, if one assumes a free energy difference $\Delta G_{\text{D-N}}$ equal to the value found in bulk experiments (15.0 $k_B T$). It is clear that this force is considerably smaller than the value we measured in our
3. Unfolding the maltose-binding protein (MBP) with optical tweezers

![Figure 3.25: The relation between work $W_{\text{tot}} - W_{\text{ext}}$ and the unfolding force $F_1$ in an unfolding experiment as sketched in Figure 3.23.](image)

experiments (~25 pN). This discrepancy can be explained from the pulling rate that we use in our experiments (~7 pN/s). At this finite pulling rate, we expect a dissipation unfolding term $W_{\text{dis}}$ higher than zero. Figure 3.25 shows that at an unfolding force of 20 pN, the difference between total work $W_{\text{tot}}$ and the work $W_{\text{ext}}$ needed to extend the DNA/polypeptide construct is $\sim 230 k_B T$. Hence, the total dissipated energy is $W_{\text{tot}} - W_{\text{ext}} - \Delta G_{D\to N} = 215 k_B T$. This shows that forced protein unfolding is a highly dissipative, non-equilibrium process.

### 3.4.2 Unfolding intermediates

In our protein unfolding experiments, we observed intermediates in the unfolding of MBP. These were found at different levels of unfolding, but there was one predominant one leading to a polypeptide contour length of 25–30 nm. This unfolding intermediate was remarkably stable: further unfolding of the intermediate occurred only after seconds at pulling forces that were sometimes higher than 30 pN. Our measurements on the 4×MBP construct confirmed the existence of this intermediate. Moreover, we performed measurements that showed that refolding from this intermediate to the native state could occur at a force between 5 and 10 pN.

Surprisingly, from bulk experiments, no unfolding intermediates are known for the unfolding of MBP. The structure of MBP (see Figure 3.4) suggests a two-step unfolding where each of the two lobes on both sides of the maltose-binding cleft unfolds independently. Close inspection of the structure reveals, however, that the
polypeptide chain crosses the bridge between the two lobes three times. It is thus unlikely that separate unfolding of these lobes is possible at all.

Unfolding intermediates in single-molecule protein unfolding have been observed before. Dietz and Rief \cite{Dietz2007} saw short-lived (~1 ms) intermediates in the unfolding of the green fluorescent protein (GFP) using AFM and Cecconi et al. \cite{Cecconi2008} saw an intermediate in the folding and refolding of ribonuclease H that had already been observed previously in bulk studies.

We have performed steered molecular dynamics simulations to find an explanation for the observed intermediate(s). The simulations showed a possible explanation for the predominant unfolding intermediate. The last ~85 C-terminal amino acids (285–370) of MBP mainly fold into a series of α-helices that are partially exposed to the surrounding buffer. In the sequence beyond this point, residues are more buried in the structure and moreover, the structure contains two residues at positions 276 and 283, that have been proven crucial for folding \cite{Angarita2005}. Possibly, in our unfolding experiments, the surface-exposed α-helices are ‘peeled’ off in a low-force (~16.5 pN) unfolding event, before the tertiary structure around amino acids 276 and 283 is disrupted, leading to further complete unfolding. The resulting 85 denatured residues would result in an additional contour length of 28 nm, which remarkably well matches the contour length of the predominant intermediate in the optical tweezers studies. Additional SMD simulations will be performed to show that the stability of the MBP structure without amino acids 285–370, is stable enough to explain the intermediate.

In our protein unfolding experiments, we observed features in the relax curves that suggested an additional compaction of the polypeptide at forces below 10 pN. This effect was most prominent for the experiments in which the 4×MBP construct was unfolded. A likely explanation for this effect is the condensation of the polypeptide chain due to interactions between nearby hydrophobic residues, or the early formation of secondary structure.

### 3.4.3 Binding of SecB to MBP

One of the aims of our experiments was to explore, on a single-molecular level, the molecular mechanism of SecB-preprotein binding. We performed unfolding experiments of MBP in the presence of 0.1 μM SecB. These experiments showed that SecB had no effect on the protein before its complete forced unfolding. The force at which MBP unfolds to the denatured state was similar to the unfolding force in the absence of SecB. This result is consistent with literature. Bulk experiments have shown that SecB has no affinity for native, stably folded polypeptides \cite{Bukau2010}.

After the complete unfolding of a protein, both the stretch and the relax force–extension curve of the DNA/polypeptide construct resembled worm-like chain behavior. No features could be seen in the force–extension curve that would point to the (partial) refolding of the protein. Hence, the effect of SecB was of an all-or-nothing kind: no refolding of the protein could be observed upon decreasing the
tension exerted on the SecB-bound polypeptide—not even to a folding intermediate.

In our measurements, we also looked for features that would reveal details of the binding geometry of SecB to a preprotein. Our measurements did not show such features. In Figure 3.26, different models for the binding geometry of SecB to a denatured polypeptide are shown schematically. Two of them show a 1:1 stoichiometry of SecB tetramers to polypeptide, with the four putative preprotein-binding sites binding either locally (b) or to sites throughout the preprotein sequence (a). In Figure 3.26c multiple SecB tetramers bind to a single polypeptide. It is not known whether the parts of the polypeptide that are not bound to SecB can form a primordial secondary or tertiary structure, as is shown in Figure 3.26. Our data suggests that if structure is formed, its stability is insignificant.

To extend the polypeptide in the configuration sketched in Figure 3.26a, several preprotein-binding channels have to unbind from the preprotein, before the protein can be fully extended. In our experiments, we looked for the unbinding of the binding channels from the protein as sudden decreases in force, similar to unfolding.
events. Our measurements do not show such events (see Figs 3.21 and 3.22). We can, however, not exclude the binding mechanism of Figure 3.26a, taking into account the unbinding rate of SecB from MBP, 3 s\(^{-1}\) [66]. Unbinding is a rapid process and is expected to happen within the first second of a stretch sweep in our experiments, where it will go unnoticed due to the compliance of the DNA linker at low force. Experiments at a higher pulling rate should be performed to increase the force at which unbinding is expected.

From bulk studies using MBP, it is known that there is a 1:1 stoichiometry in the binding of SecB tetramers to MBP [51]. More than in the condensed, guanidinium-chloride-denatured state of a polypeptide that is used in these bulk studies, the extended denatured configuration of the MBP polypeptide in our experiments may allow for the binding of multiple SecB tetramers, leading to a mechanism as sketched in Figure 3.26c. Likely binding sites for SecB along the MBP amino acid sequence are indicated in Figure 3.5. To ensure a 1:1 binding of SecB to MBP, unfolding experiments should be performed with an MBP construct that has been bound to SecB prior to tethering the construct between two microspheres, in a bulk reaction in denaturing conditions.

### 3.4.4 Protein translocation context

Here, we will discuss our experiments in the broader perspective of protein translocation. In our experiments on the forced unfolding of MBP, we found an unfolding force of \(\sim 25\) pN at a pulling rate of \(7\) pN/s, a speed that is comparable to the rate of translocation by the Sec translocase.

Our experiments suggest that in the translocation of a polypeptide by ATPase SecA in the presence of chaperone SecB, only a small part of the energy from ATP hydrolysis is used for the unraveling or the unfolding of a preprotein. We performed experiments where a SecB-bound polypeptide was slowly extended using the optical trap. This experiment can be directly compared with the initiation of protein translocation by the Sec translocase. After a SecB-bound preprotein has arrived at the Sec translocase, the preprotein is transferred to ATPase SecA and translocated by \(\sim 40\) aa upon binding of SecB to SecA and a subsequent ATP hydrolysis by SecA [21]. The free energy from hydrolysis of a single molecule of ATP is estimated to be \(50.2\) kJ/mol or \(\sim 19\) \(k_B T\) \textit{in vivo} at 37°C. A central question in protein translocation has been whether this energy is coupled to a 'power stroke' by SecA that drives the unfolding of a preprotein and its translocation by 40 aa. We observed that the force–extension curve of a SecB-bound polypeptide resembles that of an entropy-dominated random polypeptide coil. This shows that to extend a preprotein in a translocation-competent state, only entropic fluctuations have to be pulled out. The free energy that is released by hydrolysis of ATP by SecA is considerably larger than the free energy required to extend a random polypeptide coil by the distance corresponding to a polypeptide of 40 aa (estimated to be \(\sim 0.9\) \(k_B T\) using an analysis similar to that used in §3.4.1). This strongly suggests that the free energy from ATP hydrolysis is mainly used for other purposes than
the unfolding of the preprotein. Here, one can think of, e.g., the induction of conformational changes in the translocase pore. Single-molecule experiments on protein translocation by the Sec translocase, as suggested in Chapter 2 could shed more light on the thermodynamics of SecA hydrolysis.

Previously, Wilcox et al. [81] and Sato et al. [82] performed AFM protein unfolding experiments on mutants of *E. coli* dihydrofolate reductase (DHFR) and the I27 domain of titin, respectively, and studied the relation between the unfolding force and the rate of mitochondrial import of a protein. Mitochondrial import is a process that is comparable to protein translocation by the Sec translocase. It was found that the time needed to import a protein, directly scales with its resistance to mechanical unfolding. Nouwen [83] showed that multiple titin I27 domains—with an unfolding force of ~229 pN at a pulling rate of 135,000 pN/s [81]—can be efficiently translocated by the *E. coli* Sec translocase, when fused with their N terminus to the C terminus of proOmpA (a protein that can readily be translocated; see Chapter 2), even in the absence of SecB.

Remarkably, translocation of MBP—with a considerably lower unfolding force, ~25 pN at a pulling rate of ~7 pN/s—does require SecB to keep the protein in a translocation-competent state. This observation can be explained from the fact that in the Nouwen experiments, translocation of the I27 subunits is preceded by translocation of proOmpA—a membrane protein that cannot fold in its native state before translocation. Moreover, the structure of MBP near the N terminus is very strongly folded; it contains the four residues that have been proven crucial for folding (Val, Gly, Ala276 and Tyr283 [58]). De Cock and Randall [65] even suggested that this structural element could form if MBP is in complex with SecB. Hence, this structure is not easily disrupted, when the pulling is performed only from the N terminus, as is the case in translocation. Our high-pulling-rate SM simulation confirm this notion: to disrupt the interactions between the N-terminal β-strand and the residues in its surroundings, forces are required that are several times higher than those needed to remove the first C-terminal α-helix (see Figure 3.19).

Hence, our study shows the importance of the structure of a protein adjacent to the signal sequence to its translocation efficiency. Translocation experiments using a protein construct with the N terminus of MBP cloned to the C terminus of proOmpA could further show the importance of the N terminal structure in translocation by the Sec translocase. Alternatively, one could use circular permutants [84] of MBP that have their N terminus at a position with less structural rigidity (See Wilcox et al. [81] for AFM protein unfolding studies using circular permutants).