

Dissertation zur Erlangung des Doktorgrades
der Fakultät für Chemie und Pharmazie
der Ludwig-Maximilians-Universität München

Insights into the Mechanism of Action of the Protein
Folding Nanomachine GroEL

Manal Chatila

aus

Beirut, Libanon

2011

Erklärung

Diese Dissertation wurde im Sinne von § 13 Abs. 3 bzw. 4 der Promotionsordnung vom 29. Januar 1998 (in der Fassung der sechsten Änderungssatzung vom 16. August 2010) von Herrn Prof. Ulrich Hartl betreut.

Ehrenwörtliche Versicherung

Diese Dissertation wurde selbständig, ohne unerlaubte Hilfe erarbeitet.
München, 07/07/2011

Manal Chatila

(Unterschrift des Autors / der Autorin)

Dissertation eingereicht am 07/07/2011

1. Gutacher Prof. Ulrich Hartl
2. Gutachter PD Dr. Konstanze Winklhofer

Mündliche Prüfung am 04/08/2011

Acknowledgments

I am first and foremost greatly grateful for my supervisor Dr. Manajit Hayer-Hartl, whose decision to accept me in the GroEL group has changed my life in ways I cannot start to describe, both on the personal and professional levels. I am greatly in debt for Prof. Ulrich Hartl, who not only offered unlimited resources to experiment and learn, but also allowed me to witness great science in-the-making; I will miss those moments during lab meeting when I thought „Wow, that is a great idea!“. Many thanks go to my thesis committee members for taking the time to evaluate my work.

Great thanks go to the colleagues and friends who made my stay in the lab very enjoyable, and more specifically the A1 office: Alex, Giulia, Florian, Paolo, Martin, Julien, Stephan, Roman, and the recent additions who will further carry the GroEL flame Kristina, Amit and Javid. I will greatly miss our FACs (Friday Afternoon Cakes), and always cherish the great friendships I have made.

Huge thanks go to the people without whom my research would not have been possible; Bernhard, Guoxin, Sue, Jyoti and last but not least Kausik.

To the Hartl-fighters, I am thankful for all the late beergarten sessions, the bowling nights, the get-togethers and all the laughs. Those always helped against research blues!

My research as well as my stay in Munich has always been very smooth and pleasant, thanks to the great technical support from Emmanuel, Elizabeth, Silke, Andrea, Evelyn, Romy and especially Nadine. Their willingness to help always brought in a breath of fresh air in the department.

In the end, my deepest thanks and acknowledgments go to my parents for their endless support and their trust in me. Your unlimited love gave me the strength to go on. Special thanks are due to my brother Mohamed, for filling in my blanks. And last, great thanks go to Niclas for being there, and for constantly encouraging and supporting. Without you my sun would not shine every day.

Contents

1. Summary.....	1
2. Introduction.....	2
2.1 Protein folding	2
2.1.1 Protein structure	2
2.1.2 The protein folding problem.....	5
2.1.3 Methods to study protein folding	15
2.2 Protein folding in the cell.....	25
2.3 The cellular chaperone machinery	26
2.3.1 Trigger factor.....	28
2.3.2 The DnaK/J system	29
2.3.3 GroEL/ES	31
2.3.4 Additional <i>E. coli</i> chaperones	41
2.4 Aim of the study.....	43
3. Materials and Methods.....	44
3.1 Materials	44
3.1.1 Chemicals	44
3.1.2 Enzymes	46
3.1.3 Materials.....	46
3.1.4 Instruments	47
3.1.5 Media.....	48
3.1.6 Ampicillin stock solutions.....	49
3.2 Bacterial strains and plasmids.....	49

3.2.1	<i>E.coli</i> strains	49
3.2.2	Plasmids.....	49
3.3	Molecular cloning techniques	50
3.3.1	Preparation and transformation of <i>E. coli</i> competent cells.....	50
3.3.2	Plasmid purification	51
3.3.3	PCR-mediated mutagenesis.....	51
3.3.4	DNA analytical methods	51
3.4	Protein purification	51
3.4.1	GroEL expression and purification	51
3.4.2	GroES expression and purification.....	53
3.4.3	MBP and MBP mutants expression and purification	53
3.5	Protein analytical methods.....	54
3.5.1	Determination of protein concentration	54
3.5.2	SDS-PAGE.....	55
3.6	Biochemical and biophysical methods.....	55
3.6.1	GroEL ATPase assay.....	55
3.6.2	<i>In vitro</i> MBP refolding assay	56
3.6.3	Sample preparation for analysis of disulfide-bond formation in the cysteine mutants	57
3.6.4	Functional characterization of the cysteine mutants by maltose binding.....	58
3.6.5	Unfolding and refolding curves of DM-MBP	58
3.6.6	Tryptophan fluorescence and Circular dichroism wavelength scans	59
3.6.7	Unfolding kinetics by Stopped-flow measurements	59
3.6.8	Static and Dynamic light scattering.....	59

3.7	Single-molecule FRET experiments	60
3.7.1	Thiol-mediated labeling of the cysteine constructs	60
3.7.2	Fluorescence correlation and cross-correlation spectroscopy	61
3.7.3	Single-pair FRET	63
3.8	Hydrogen/deuterium exchange experiments.....	63
3.8.1	Pulse labeling	63
3.8.2	Mass spectrometric analysis.....	64
4.	Results	65
4.1	The spontaneous refolding of DM-MBP is not slowed down by transient aggregation...	65
4.1.1	A model of aggregation-limited DM-MBP refolding is simulated	65
4.1.2	The model of aggregation-limited kinetics is inconsistent with the actual refolding experiments.....	68
4.1.3	FCS and FCCS support the absence of any transient aggregation during DM- MBP refolding	69
4.1.4	Static and dynamic light scattering further support the absence of any aggregation during DM-MBP refolding.....	72
4.2	Spontaneous refolding of DM-MBP is limited by the formation of a kinetically-trapped intermediate.....	73
4.2.1	Denaturation-renaturation curves by Trp fluorescence and CD unveil the kinetically-trapped intermediate.....	73
4.2.2	The kinetically-trapped intermediate is collapsed, but lacks ordered structure	77
4.3	Disulfide-mediated constraints accelerate spontaneous refolding by decreasing the entropic barrier to folding	81
4.3.1	The oxidized mutants fold faster than their reduced counterparts	82
4.3.2	The disulfide bonds had no significant impact on the energy barrier between the native state and the transition state	86

4.3.3	The kinetic trap is largely entropic in nature.....	88
4.4	The chaperonin cage mimics disulfide-mediated constraints	93
4.4.1	SR-EL accelerates the refolding of the cysteine mutants.....	93
4.4.2	Constraining the N and C domains of DM-MBP is comparable to the chaperonin effects.....	95
4.4.3	The disulfide bridges energetically mimic the confinement by the chaperonin	97
4.5	Charge clusters in the chaperonin cage wall are crucial for the folding.....	98
4.5.1	The SR-mutant is unable to reduce the entropic barrier to folding	99
4.5.2	Folding inside the SR-mutant cage mimics spontaneous refolding	100
4.6	Unfolding upon substrate binding seems unessential for rate acceleration	102
4.6.1	FRET analyses reveal more chain compactness in SR- or V263S-bound DM-MBP as compared to EL-bound DM-MBP	102
4.6.2	Crowding in the presence of Ficoll restores folding efficiency of the V263S mutant	105
4.7	A comparison of the folding of DM-MBP cysteine mutants shows no growth advantages conferred by the mutations.....	107
4.7.1	Periplasmic DM-MBP allows for less maltose fermentation than WT-MBP	108
4.7.2	Periplasmic DM-MBP mutants exhibit temperature-dependent phenotype	109
5.	Discussion.....	114
5.1	DM-MBP: A slow folder without detectable aggregation	114
5.2	Basis of slow DM-MBP folding	115
5.3	Mechanism of an active chaperonin	118
5.4	Effects of cavity charges on folding	119
5.5	Substrate unfolding in the folding mechanism	121

5.6	Towards an understanding of periplasmic protein folding <i>in vivo</i>	123
5.7	Biological significance of an active mechanism.....	124
6.	References.....	126
7.	Curriculum Vitae.....	136

1. Summary

GroEL and GroES form a chaperonin nano-cage for single protein molecules to fold in isolation. The folding properties that render a protein chaperonin-dependent are not yet understood. Here, we addressed this question using a double mutant of the maltose-binding protein DM-MBP as a substrate. We find that upon spontaneous refolding, DM-MBP populates a kinetically trapped intermediate that is collapsed but structurally disordered. Introducing two long-range disulfide bonds into DM-MBP reduces the entropic folding barrier of this intermediate and strongly accelerates native state formation. Strikingly, steric confinement of the protein in the chaperonin cage mimics the kinetic effect of constraining disulfides on folding, in a manner mediated by negative charge clusters in the cage wall. These findings suggest that chaperonin dependence correlates with the tendency of proteins to populate entropically stabilized folding intermediates. The capacity to rescue proteins from such folding traps may explain the uniquely essential role of chaperonin cages within the cellular chaperone network.

2. Introduction

Proteins are the working molecules of the cell, and carry out an elaborate program of activities encoded by the genes. They have evolved to acquire many specialized abilities and can be grouped into several broad functional classes: *catalytic* proteins (enzymes), which increase the rates of chemical reactions by several orders of magnitude; *structural* proteins, which provide structural rigidity to the cell; *transport* proteins, which control the flow of material across membranes; *regulatory* proteins, which act as sensors and switches to control protein activity and gene function; *signaling* proteins, including cell-surface receptors and other proteins that transmit external signals to the cell interior; and *motor* proteins, which cause motion. Proteins must adopt a precise three-dimensional structure to fulfill their activities, and they do so by a process called protein folding (Lodish et al., 2003). To carry out this process efficiently, cells make use of a class of proteins called *molecular chaperones*, whose mechanism will be further discussed in this thesis.

2.1 Protein folding

2.1.1 Protein structure

In spite of their diverse functions, proteins are a relatively homogeneous class of molecules, and are built as linear polymers of various combinations of the same 20 amino acids. They differ in the sequence in which the amino-acids are assembled into polymeric chains or what is referred to as the primary structure (Creighton, 1991). Different regions of the sequence form local repeating structures, such as α -helices or β -sheets, which are referred to as secondary structure. The tertiary structure describes the overall structure of the folded polypeptide chain, including

the arrangement of secondary structure elements. The quaternary structure constitutes the spatial arrangement of the various polypeptide chains, loosely referred to as subunits of the protein, in case it contains more than one (Voet et al., 2006).

In the primary structure, the α -carbons of adjacent amino acid residues are separated by three covalent bonds, arranged as C_α -C-N- C_α (Fig. 1). Along with the oxygen atom of the carbonyl group, and the hydrogen atom of the amide group, those 6 atoms lie in a single plane. Evidence for the planarity of the peptide group came as early as 1930, when Pauling and Corey determined the X-ray structures of several amino acids and dipeptides. According to these experiments, the peptide bond has a ~40% double-bond character due to resonance, and the peptide bonds are therefore unable to rotate freely. Rotation is nonetheless allowed around the N- C_α and C_α -C bonds, and the resulting angles are respectively called ϕ (phi) and ψ (psi). By convention, ϕ and ψ are 180° when the polypeptide is in its fully extended conformation, and increase in a clockwise manner when viewed from C_α . In principle, ϕ and ψ can be any value between -180° and 180° , but many angles are excluded because of steric interference between atoms of the polypeptide backbone or the side chains. The conformation of the backbone can thus be described by a pair of angles, ϕ and ψ . G. N. Ramachandran calculated the energy contained in various pairs of ϕ and ψ angles, and found the two most stable combinations, the so-called α and β conformations (Ramachandran et al., 1963). These two pairs of angles are found to almost exclusively occur in native proteins, including the α -helix and β -strand, the most prominent examples of secondary structure.

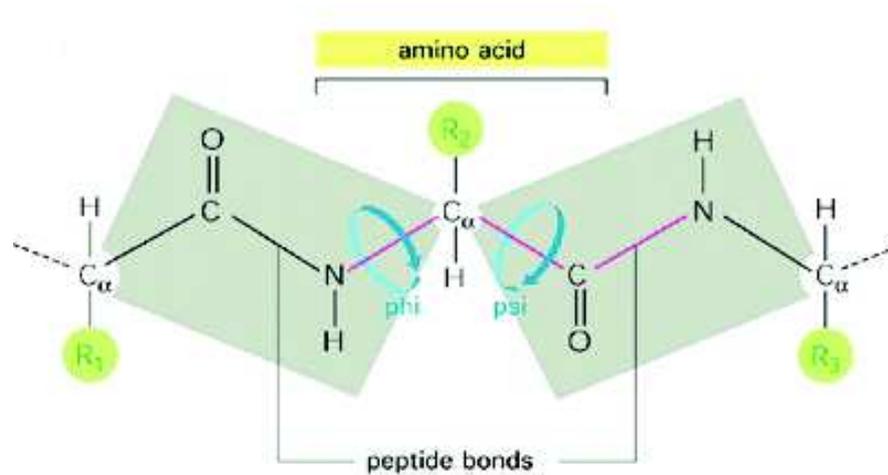


Figure 1: The peptide bond and the torsional angles. Shaded in green are the planar atoms participating in the peptide bond. Drawn in purple are the three bonds contributed to the backbone by one amino acid. Rotation is allowed around the N-C_α and C_α-C bonds, and the resulting shown angles are respectively called φ (phi) and ψ (psi).

At a higher level of structural complexity, the tertiary structure of a protein is that which describes the folding of its secondary structural elements and specifies the positions of each atom in the protein, including those of its side chains. The known protein structures have come to light mainly through X-ray crystallographic or nuclear magnetic resonance (NMR) studies. In fact, the first X-ray structure of a protein was that of sperm whale myoglobin, and was reported in 1958 by John Kendrew and co-workers. Over half a decade later, more than 65 000 protein structures are deposited at the Protein Data Bank (PDB). Of those, around 40 000 are structurally classified in the Structural Classification of Proteins (SCOP) database. The SCOP has been constructed in 1995, due to the exponential growth in the number of resolved tertiary structures, and in an effort to facilitate the understanding and access to this information; this database provides a detailed and comprehensive description of the structural and evolutionary relationships of proteins whose three-dimensional structures have been determined (Murzin et al., 1995). Comparisons of those

reveals that a protein's structure is much better conserved than the amino-acid sequence, giving rise to families of proteins having related folds but varying widely in sequence similarity. Indeed, the preponderance of two major secondary structure conformations (α -helix and β -strand) limits the types of structures observed to four classes (mainly- α , mainly- β , alternating α/β , $\alpha+\beta$) (Orengo, 1994). Within each class are groups of common, or favored, secondary structure arrangements, called folds or motifs, e.g. the β -sandwich and β -barrel arrangements. In some cases, motifs are signatures for specific functions. For instance, the zinc finger motif is commonly found in DNA-binding proteins, whereas the EF hand motif is found in more than a 100 calcium-binding proteins. The presence of the same motif in different proteins with similar functions clearly highlights that these useful combinations of secondary structures have been conserved in evolution (Lodish et al., 2003).

2.1.2 The protein folding problem

The protein folding problem is the question of how a protein's amino acid sequence dictates its three-dimensional atomic structure. The notion of such a folding "problem" emerged shortly after the appearance of the aforementioned myoglobin structure in 1958. This first protein structure had helices that were packed together in unexpected and irregular ways. Since then, the protein folding problem has come to be regarded as three different problems: (a) the folding code: the thermodynamic question of what balance of interatomic forces dictates the structure of the protein, for a given amino acid sequence; (b) protein structure prediction: the computational problem of how to predict a protein's native structure from its amino acid sequence; and (c) the folding process: the kinetics question of what routes or pathways some proteins use to fold so quickly (Dill et al., 2008).

2.1.2.1 The folding code

A major milestone in protein science was the thermodynamic hypothesis of Christian Anfinsen and colleagues (Anfinsen, 1973). From his now-famous experiments on ribonuclease in the 1960s (Fig. 2), Anfinsen postulated that the native structure of a protein is the thermodynamically stable structure; it depends only on the amino acid sequence and on the conditions of solution, and not on the kinetic folding route. Regardless of whether the protein was synthesized on a ribosome, or isolated in a test tube, the native structure attained is the same. And since proteins are only marginally stable at room temperature, it is clear that no type of molecular interaction in reaching the native state could be ignored (Dill, 1990).

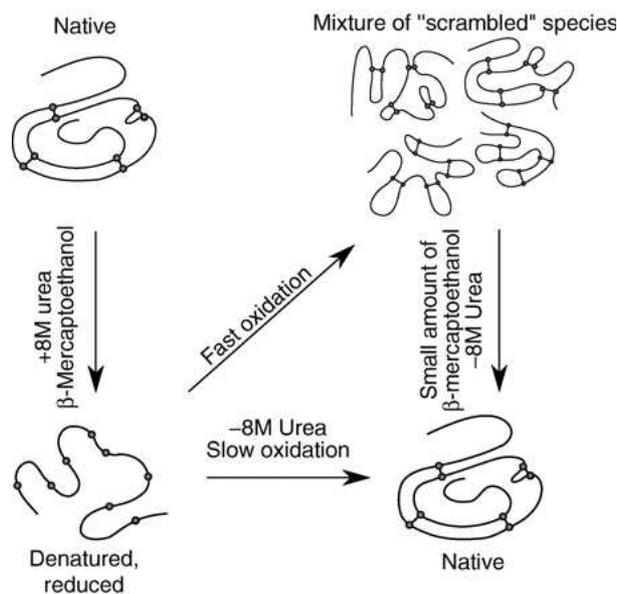


Figure 2: Anfinsen's famous Ribonuclease experiment. Ribonuclease, an enzyme with 124 amino acid residues, contains four disulfide bridges. Treatment of ribonuclease with 8M urea in the presence of the reducing agent β-mercaptoethanol causes a complete unfolding of the ribonuclease and loss of enzymatic activity. Allowing the cysteines to reoxidize under denaturing conditions results in a mixture of "scrambled" species where the eight SH groups randomly pair to form four disulfide bridges. However, when urea is slowly removed by dialysis and a small

amount of β -mercaptoethanol is added, disulfide interchange takes place and the mixture of “scrambled” ribonuclease is eventually converted to a homogeneous product, which is fully enzymatically active and indistinguishable from native ribonuclease.

Prior to the mid-1980s, the protein folding code was seen a sum of many different small interactions-such as hydrogen bonds, ion pairs, van der Waals attractions, and water-mediated hydrophobic interactions. A key idea was that the primary sequence encoded secondary structures, which then encoded tertiary structures (Anfinsen & Scheraga, 1975). However, through statistical mechanical modeling, a different view emerged in the 1980s, namely, that there is a dominant component to the folding code, that it is the hydrophobic interaction, that the folding code is distributed both locally and nonlocally in the sequence, and that a protein’s secondary structure is as much a consequence of the tertiary structure as a cause of it (Dill et al., 2008).

There are still current debates in the field regarding the contributions of each of those forces. As far as ion-pairing is concerned, structural studies conducted by Barlow and Thornton provided strong evidence that it is not the dominant driving force in protein folding; they observed that ion pairs are not highly conserved in evolution, and that the number of ion pairs in proteins is small (Barlow & Thornton, 1983). Nevertheless, genome-wide and structural comparisons of thermophilic and mesophilic proteins indicate that salt bridges may significantly contribute to the enhanced thermal stability of proteins from thermophilic organisms (Szilagyi et al., 2006).

Also, the role of hydrogen bonds in folding and stability is still a matter of debate; based on the argument that intramolecular hydrogen bonds are energetically equivalent to hydrogen bonds

between the protein and the solvent, among other lines of evidence, the role of hydrogen bonding in driving the folding was considered minor (Dill, 1990). However, a number of mutational studies suggested that hydrogen bonds do contribute to protein stability, and that this contribution is as important as the hydrophobic effect (Pace et al., 1996).

The important role of the hydrophobic interaction in driving the folding is nevertheless uncontested in the field. It used to be explained as a primarily entropic effect arising from the rearrangement of hydrogen bonds between solvent molecules around an apolar solute. This hydration process is energetically unfavorable, and therefore drives apolar solutes together, thereby decreasing their solvent-exposed surface area (Makhatadze & Privalov, 1995). Today, the hydrophobic effect is usually viewed as a combined effect of hydration and van der Waals interactions between solute molecules (Creighton, 1993).

2.1.2.2 Protein structure prediction

A long-standing goal of computational biology has been to devise a computer algorithm that takes, as input, an amino acid sequence and gives, as output, the three-dimensional native structure of a protein. A main motivation is to make drug discovery faster and more efficient by replacing slow expensive structural biology experiments with fast cheap computer simulations. A major milestone in computer-based native structure prediction was the invention of CASP (Critical Assessment of Techniques for Structure Prediction) by John Moult in 1994. An experiment in the sociology of science, CASP is a community-wide blind test to predict unknown protein structures, given only the amino acid sequence. Web servers and software packages are now able to predict structures of small globular proteins to within 2-6 Å of the experimental structure (Dill et al., 2007). Alongside the bioinformatics-based approach, a small

scientific community currently aims to use purely physics-based methods, without knowledge derived from databases (such as statistical energy functions or secondary structure predictors), to explore native structures and folding processes. A milestone was the development by Pande and colleagues of Folding@home, a distributed grid computing system running on the screensavers of volunteer computers worldwide. They have for instance simulated the folding kinetics of villin, and were able to reach high resolution structures (Dill et al., 2007).

2.1.2.3 Folding kinetics and the emerging “New View”

In an effort to explain how proteins reach their native state, Cyrus Levinthal proposed that proteins must fold by specific “folding pathways”. His argument, which was later named “Levinthal’s paradox”, rests on the fact that there are too many possible conformations for proteins to find the “needle” (native structure) in the “haystack” (conformational space) by random searching, and in biologically-relevant time scales. To circumvent the Levinthal paradox, it was envisaged that proteins could fold by defined pathways and mechanisms that removed the need to search all possible conformations (Fig. 3).

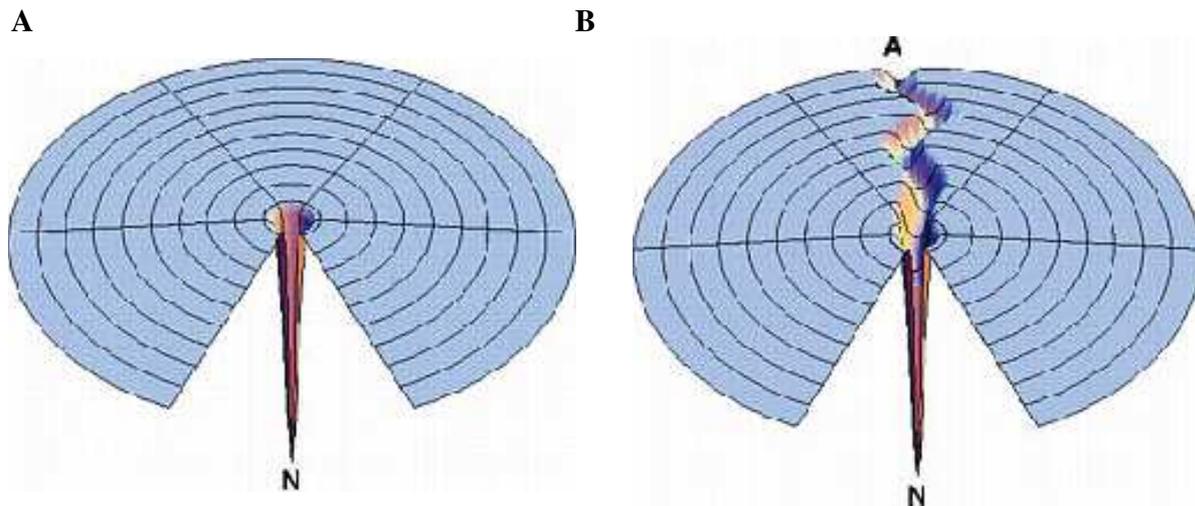


Figure 3: The Levinthal golf course landscape and its solution. A) Levinthal’s argument has its roots in such hypothetical golf-course depictions of the protein folding. B) The pathway solution is assumed to lead from a denatured conformation A to the native conformation S, so conformational searching is more directed and folding is faster.

At one extreme, the *framework model* proposed that the stepwise formation of structure would simplify the task. In other words, rapid formation of local secondary structure, which functions as a scaffold, would be followed by the acquisition of tertiary structure. Two mechanisms were proposed: *diffusion-collision*, which generally involves formation of secondary structures, followed by their diffusion, collision, and coalescence to form tertiary structure; and *nucleation*, in which a nucleus is slowly formed, followed by the rapid propagation of structure. At the other extreme, the *hydrophobic-collapse* mechanism proposes that the initial steps in folding involve hydrophobic collapse. Acquisition of secondary structure and the correct packing interactions are then formed in a confined volume (Daggett & Fersht, 2003). The many models are not mutually exclusive, and try to grasp different aspects of folding, based on the experimental results. A newer model, named “nucleation-condensation” is an attempt to unite features of both the framework and the hydrophobic collapse models. In fact, those two are seen as extreme

manifestations of the nucleation-condensation model, when either the secondary structure, or tertiary interactions, respectively, become over stabilized (Daggett & Fersht, 2003). These models are nevertheless only retrospective summaries of experimental data, according to Dill and colleagues. They do not constitute a general principle which would be used to predict folding routes for any amino acid sequence. One possible folding principle is the Zipping and Assembly (Z&A) model. In Z&A, local structuring happens first in independent peptide fragment sites along the chain, then those structures either grow (zip) or coalesce (assemble) with other structures, along pathways involving topologically local contacts (Voelz & Dill, 2007). According to the Z&A mechanism, proteins can fold quickly because they don't search all their degrees of freedom at the same time. This principle is in accord with Plaxco's observation that the fastest folders are proteins having predominantly local contacts in their native states (Plaxco, et al., 1998).

The classical view of the kinetics of protein folding is based on simple phenomenological kinetics models, where raw data are single or multi-exponential time decays of optical properties that monitor changes in the protein structure after a jump to folding or unfolding conditions. Folding kinetics is described by simple mass action models, such as $D \rightarrow I \rightarrow N$ (on-path intermediate I between the denatured state D and native state N) or $X \rightarrow D \rightarrow N$ (off-path intermediate X), where the symbol I or X represents macrostates that are invoked for the purpose of curve-fitting experimental kinetics data. These classical experiments probe only the average behavior of proteins, and are unable to solve much of the atomic detail (Dill & Chan, 1997).

In contrast, the new view, exemplified by funnel models, aims to explain the kinetics in terms of underlying physical forces. The new view has emerged from advances in both

experiment and theory. The main experimental advances have been those that give detailed structural information, such as HD-exchange and mass spectrometry, but also those which explore very early events in folding such as fast laser-triggered methods. The theoretical advances contributing to the new view are due to a class of statistical mechanics models with highly-simplified, most often lattice-based, representations of chain geometries and interactions, analyzed by analytical methods and computer simulations (Dill & Chan, 1997). Although they lack atomic details, these models nevertheless go beyond the classical ones, by including chain connectivity, flexibility, excluded volume, and sequence-dependent intrachain interactions (Dill & Chan, 1997). The new view replaces the pathway concept of sequential events with the funnel concept of parallel events, and considers folding as a diffusion-like process, where the motion of individual chains is asynchronous, and where all chains ultimately find their way to the native structure. Figure 4 shows how funnels resolve Levinthal's paradox, and eliminate the need for intermediates along the folding pathway.

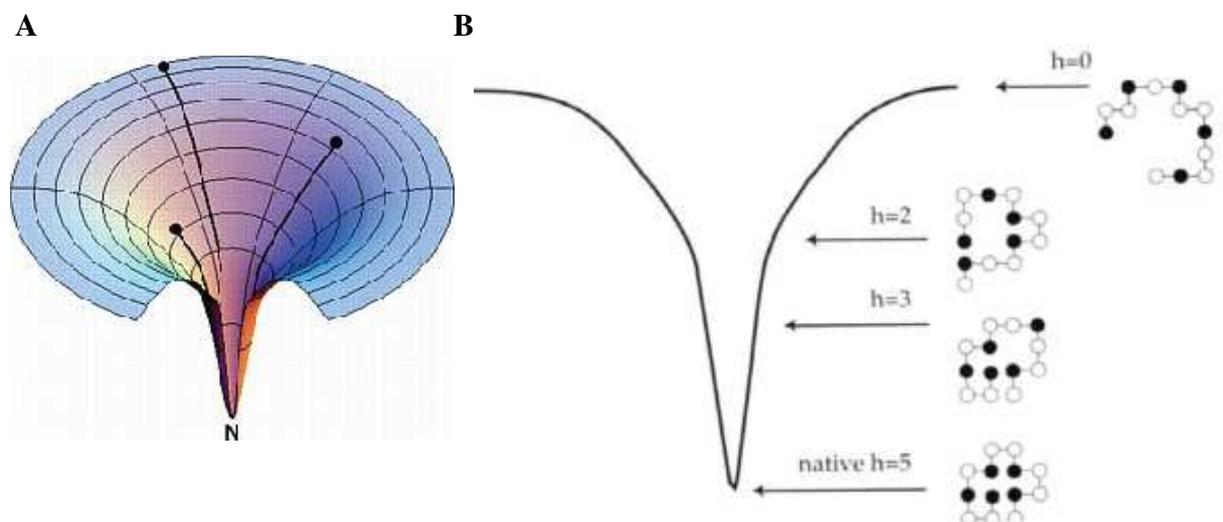


Figure 4: Idealized funnel landscape. A) As the chain forms increasing number of intrachain contacts, and lowers its internal free energy, its conformational freedom is also reduced. B) Slice through A). In the lattice model, black

and white beads represent hydrophobic and polar monomers, respectively, and h is the number of hydrophobic contacts. Exact enumeration studies show that there are many open conformations, fewer compact conformations, and only one having $h=5$. An ensemble of molecules can reach the global minimum, satisfying Anfinsen's hypothesis, and do so quickly, satisfying Levinthal's concern, without following specific pathways.

More specifically, the new view uses the language of energy landscapes. An energy landscape is just the free energy of each conformation as a function of the degrees of freedom, such as dihedral bond angles along the peptide backbone. The vertical axis represents the "internal free energy" of a given chain configuration, including the sum of hydrogen bonds, ion-pairs, torsional angle energies, hydrophobic and solvation free energies, or in other words everything except the conformational entropy. The many lateral axes represent the conformational coordinates. The high dimensionality of this representation reflects the many degrees of freedom of a protein chain (Fig. 5). To specify a conformation, many coordinates, for example the dihedral angles are needed. Each conformation is then represented by a point in this multidimensional energy surface, and geometrically-close conformations are also close to one another on the landscape. Many features are present in those landscapes: for instance, hills correspond to high energy conformations (such as those with buried polar groups in hydrophobic core), and valleys represent more favorable conformations (Dill & Chan, 1997).

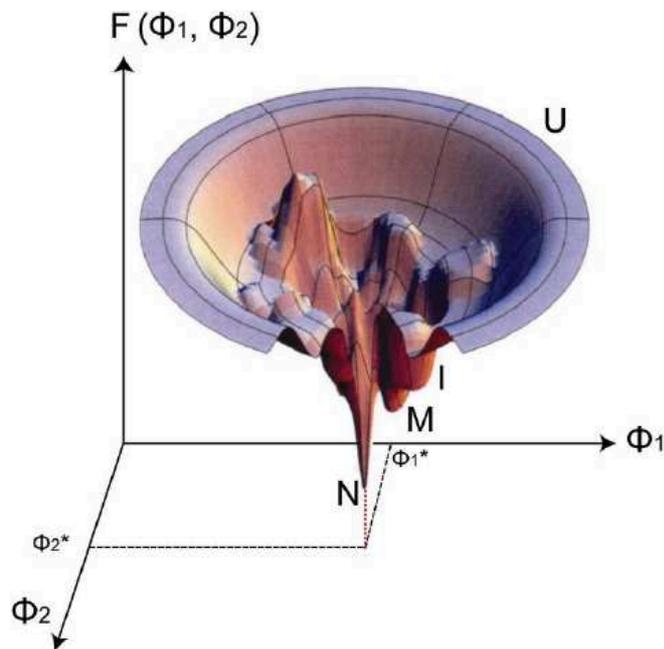


Figure 5: Schematic representation of the free energy landscape of protein folding. The vertical axis of the funnel represents the internal free energy, $F(\Phi_1, \Phi_2, \dots)$ of a given chain conformation as a function of the degrees of freedom, Φ_1, Φ_2 . A multitude of unfolded conformations (U) on the top of the funnel proceeds through a number of local energy minima (I) towards the native conformation (N) with the lowest free energy. In some local minima, misfolded species (M) may be trapped irreversibly.

Nevertheless, the folding funnel concept is not the complete answer to Levinthal's puzzle. According to Dill, a complete solution would entail instructing a computer program to find a native state more efficiently than Monte Carlo or molecular dynamics, and for that, one needs to know the microscopic folding routes. Thus, the previously mentioned Z&A mechanism in conjunction with the folding funnel concept provide a plausible answer to Levinthal's kinetic protein folding problem and shows why proteins don't need supercomputers to guide them to their native structures (Dill et al., 2007). Indeed, recent tests show that the Z&A mechanism speeds up conformational searching sufficiently that physics-only models can now find approximately correct folds for chain lengths up to around 100 monomers (Ozkan et al., 2007).

2.1.3 Methods to study protein folding

Characterization of all the different species populated during the protein folding process is an important aspect of unveiling the all-atom detail folding mechanism. Substantial advances have been made towards this goal for a few small proteins. This has been enabled by the development of experimental approaches with faster timescales of measurement, and enhanced sensitivity, together with improvements in computing power and new theoretical tools (Table 1). Today, the battery of biophysical methods allows researchers to monitor transitions from picosecond to second (or longer) timescales, and to identify and structurally assess species populated to as little as 0.5% (Bartlett & Radford, 2009). Below is a brief description of the methods used in this work.

Technique	Timescale	Information content
Intrinsic tryptophan fluorescence	ns ^a	Environment of tryptophan (through measurement of intensity and λ_{max})
Far UV CD	s ^a	Secondary-structure content
Near UV CD	s ^a	Packing of aromatic residues
Raman spectroscopy	s ^a	Solvent accessibility, conformation of aromatic residues
Infrared spectroscopy	ns ^a	Secondary-structure content
ANS (1-anilino-8-naphthalene sulfonic acid) binding	s ^a	Exposure of aromatic surface area
FRET	ps ^a	Molecular ruler, dependent on the distance between two fluorophores (r^{-6} dependence assuming free rotation of the dyes)
FCS	ps ^a	Diffusion time (and hence size and shape)
Anisotropy	s ^a	Correlation time measurements provide information about shape and size of molecule
Small-angle X-ray scattering	s ^a	Radius of gyration
Absorbance	ns ^a	Environment of chromophore
Real-time NMR	>min	Structural information via chemical shifts and measurement of NOEs
Native-state hydrogen exchange	h	Global stability, detection of metastable states
Pulsed H/D exchange by NMR	ms	Hydrogen exchange protection of folding intermediates on a per-residue basis
Pulsed H/D exchange by ESI-MS	ms	Hydrogen exchange protection of folding populations
NMR relaxation methods	ms	Nonrandom structure in denatured states and conformational exchange between different species
Protein engineering	Depends on probe used	Role of an individual residue in determining the rate of folding and stability of a species of interest

Table 1: Experimental techniques which have been applied to the study of protein folding (adapted from Barlett & Radford, 2009). ^a Time scales depend on the methods used to initiate the folding: temperature jump (ns), pressure jump (ms), ultra-rapid mixing (ms), stopped flow (ms) or manual mixing (s).

2.1.3.1 Tryptophan fluorescence

More specifically, in this work, we have used Tryptophan fluorescence as a probe of native state formation. Most proteins contain intrinsically-fluorescent amino acid residues: tryptophan, tyrosine, and phenylalanine. Tryptophan is by far the most useful of the three. Due to its aromatic character, is often (although not always) found fully or partially buried in the

hydrophobic core of protein interiors, at the interface between two protein domains or subdomains, or at the subunit interface in oligomeric protein systems. Upon disruption of the protein's tertiary or quaternary structure, these side chains become more exposed to solvent. Tryptophan's emission energy is highly sensitive to the polarity (and dynamics) of the environment; residues that are fully or partially buried in the relatively hydrophobic interior or interfaces of proteins will exhibit blue-shifted emission anywhere from 309 nm to 335 nm, whereas residues exposed to water (or upon unfolding) exhibit a red shift with a maximum at 355 nm (Royer, 2006). It is not possible to predict what the effect of solvent exposure will be upon the quantum yield (total intensity) of emission, due to the difficulty in predicting the quantum yield of tryptophan in proteins. If no major quenching or energy transfer to prosthetic groups (heme groups, NADH, or other tryptophan residues) occurs in the folded state, then one can predict that solvent quenching resulting from tryptophan exposure upon unfolding will lead to an overall decrease in the intensity of emission (Royer, 2006). In the case of Maltose-binding protein, the model substrate used in this work, 8 tryptophans are scattered throughout both domains of the protein, and refolding rates were obtained by monitoring the recovery of fluorescence intensity in a kinetics experiment.

2.1.3.2 Circular dichroism (CD)

Plane polarised light can be viewed as being made up of 2 circularly polarised components of equal magnitude, one rotating counter-clockwise (left-handed, L) and the other clockwise (right handed, R). Circular dichroism (CD) refers to the differential absorption of these 2 components. If, after passage through the sample being examined, the L and R components are not absorbed or are absorbed to equal extents, the recombination of L and R would regenerate

radiation polarised in the original plane. However, if L and R are absorbed to different extents, the resulting radiation would be said to possess elliptical polarization. A CD signal will be observed when a chromophore is chiral (optically active). CD signals only arise where absorption of radiation occurs, and thus spectral bands are easily assigned to distinct structural features of a molecule. An advantage of the CD technique in studies of proteins is that complementary structural information can be obtained from a number of spectral regions. Absorption in this region (240 nm and below), principally due to the peptide bond, allows for determination of the different types of regular secondary structure found in proteins, since these give rise to characteristic CD spectra in the far UV (Kelly et al., 2005).

2.1.3.3 Förster resonance energy transfer (FRET)

FRET occurs through a non-radiative long-range dipole-dipole interaction of a pair of chromophores over distances in the 10-100 Å range. Upon absorption of a photon, a donor chromophore may lose excess energy in a number of ways: emission of a photon (donor fluorescence), non-radiative relaxation or non-radiative energy transfer. Transfer of the energy from the donor to the acceptor results in the donor returning to the ground state, and the acceptor entering an excited state. Relaxation from this state can then result in fluorescence from the acceptor.

According to Förster's theory, and verified experimentally, the rate of energy transfer is given by the equation:

$$k_{\text{FRET}} = (1/\tau_{\text{D}}) \times (R_0/R)^6$$

where R_0 is the Förster critical distance, τ_D is the donor lifetime in the absence of the acceptor, and R is the distance separating the donor and acceptor chromophores. The Förster critical distance R_0 is defined as the acceptor-donor separation radius for which the transfer rate equals the rate of donor decay (de-excitation) in the absence of acceptor. In other words, when the donor and acceptor radius R equals the Förster distance, then the transfer efficiency is 50 percent. The value of R_0 (in nanometers) may be calculated from the following expression:

$$R_0 = 2.11 \times 10^{-2} \times (\kappa^2 \times J(\lambda) \times \eta^4 \times Q_D)^{1/6}$$

in which κ -squared is a factor describing the relative orientation in space between the transition dipoles of the donor and acceptor, $J(\lambda)$ is the overlap integral in the region of the donor emission and acceptor absorbance spectra (with the wavelength expressed in nanometers), η represents the refractive index of the medium, and Q_D is the quantum yield of the donor. In ensemble measurements, the transfer efficiency is typically measured using the relative fluorescence intensity of the donor, in the absence F_D and presence F_{DA} of acceptor, according to the following expression:

$$E = 1 - (F_{DA}/F_D)$$

In contrast to ensemble studies, the investigation of individual molecules promises direct access to information on microscopic pathways. The ultimate goal of single-molecule FRET studies is the time-resolved observation of individual protein-folding events (Schuler & Eaton, 2008). Single-pair (sp)FRET measurements combined with pulsed interleaved excitation (PIE) (Kapanidis et al., 2004; Müller et al., 2005) allow for the selective analysis of donor-acceptor double-labeled molecules by exciting the donor and acceptor dye alternately on a timescale much

faster than the diffusion of the particles through the probe volume. Only molecules that have both an active donor and acceptor are included in the FRET analysis. For spFRET measurements, the intensities of the donor and acceptor fluorescence are used to determine FRET efficiencies which are empirically determined using the following expression.

$$f_E = F_{AD} / (\alpha F_{DA} + F_{AD})$$

where α is the detection-correction factor between the green and red channels, F_{DA} is the fluorescent intensity of the donor in the presence of the acceptor, F_{AD} is the fluorescent intensity of the acceptor in the presence of donor.

2.1.3.4 Fluorescence correlation and cross-correlation spectroscopy (FCS and FCCS)

FCS is one of the many different modes of high-resolution spatial and temporal analysis of extremely dilute biomolecules. It is typically performed by monitoring fluorescence signal from molecules, which are freely diffusing in femtoliter volumes. Excitation is usually accomplished with a laser focused to a diffraction-limited spot. A confocal pinhole is used to reject signal from outside the desired volume.

The average number of fluorophores in the volume is determined by the bulk concentration and remains constant in a stationary experiment. However, random diffusion of the fluorophores results in time-dependent changes, called fluctuations, in the number of fluorophores in the volume. Fluctuations in occupancy number in this small volume result in intensity fluctuations in the fluorescence intensity. The theory for FCS provides the basis for extracting molecular information from the molecules by analysis of the rates and amplitudes of the intensity

fluctuations. It is based on Poisson statistics, and for random discrete events, the number of fluorophores in the volume can be described by:

$$P(n,N) = (N^n/n!) \times e^{-N},$$

where $P(n,N)$ is the probability of n fluorophores being present in the volume when the average number of molecules in the volume is N .

The changes in occupation number of the volume will result in changes in the intensity. If the diffusion is slow the fluorophores will move slowly into and out of the volume. If the diffusion is faster the occupation number and intensity will change more rapidly. This time-dependent intensity is analyzed statistically to determine the amplitude and frequency distribution of fluctuations. The intensity at a given time $F(t)$ is compared with the intensity at a slightly later time $F(t + \tau)$. If diffusion is slow, $F(t)$ and $F(t + \tau)$ are likely to be similar. If diffusion is fast, $F(t)$ and $F(t + \tau)$ are likely to be different. In an FCS instrument a dedicated correlation board is usually used to calculate the correlation between $F(t)$ and $F(t + \tau)$ for a range of delay times τ . This results in an autocorrelation function $G(\tau)$ that contains information on the diffusion coefficient and occupation number of the observed volume (Lakowicz, 2006).

The weak dependence of the diffusion coefficients on molecular weight makes it difficult to use FCS to measure binding reactions unless there is a large change in molecular weight. Formation of dimers is near the resolution limit for FCS using the diffusion time to distinguish two species. The addition of two-color excitation and detection to FCS changes the form of the correlation functions and provides new applications of FCS. In a typical experiment to detect dimer formation, the target protein is labeled with either the green (G) or red (R) fluorophores. A

sample made by mixing these could be observed with an FCS instrument configured for two-color measurements, and separate detectors for the R and G signals. Different time-dependent fluctuations will be observed in each channel. If a G-labelled protein diffuses into the volume there is a burst of photons in the G channel, and similarly for an R-labelled protein. If dimers do form, then a significant fraction would be RG-labelled, and upon diffusion into the volume, will be detected by the burst of photons observed in both channels. This shows that doubly labeled molecule can be detected by cross-correlation between the two channels. The method is called FCCS (Lakowicz, 2006).

2.1.3.5 Hydrogen-deuterium exchange by Mass Spectrometry (HD-MS)

Hydrogen isotope exchange experiments have been used since the 1960s to gain information about the structural stability of protein conformers and an understanding of the mechanisms involved in protein mobility. Amide hydrogen isotope exchange studies provide information on protein folding and unfolding to and from the native state, on local fluctuations of the native state, and on ligand-protein interactions. Methods to monitor hydrogen exchange have included tritium counting, infrared spectroscopy, and nuclear magnetic resonance (Maier & Deinzer, 2005).

There are essentially two competing processes during which hydrogen isotope exchange takes place. One involves internal or local fluctuations from the native state, and the other involves subglobal or global unfolding. In either case, hydrogen bonds are broken and the amide hydrogens are exposed to solvent so that exchange can occur. More specifically, the exchangeable hydrogens are the polar side-chain hydrogens bound to heteroatoms (N, O, and S), the N- and C-terminal hydrogens, and the backbone peptide amide. In proteins, the polar side-

chain hydrogens exchange 10^3 - 10^6 times faster than peptide amide hydrogens at pH 7 and these are usually too high to be readily determined. Therefore, most of the hydrogen exchange studies focus on determining the exchange characteristics of peptide amide hydrogens. Fortunately, peptide amide hydrogens represent excellent structural probes that report on surface accessibility, secondary structural bonding network, and conformational stability. MS-based approaches use the fact that the difference in mass between hydrogen and deuterium is 1 Da, and, thus, changes in the number of deuteriums incorporated in a molecule can be directly related to mass shifts (Maier & Deinzer, 2005).

2.1.3.6 Static and Dynamic light scattering (SLS and DLS)

The basic SLS experiment measures the average intensity of light scattered by a protein solution of defined concentration in excess of that scattered by background sources (solvent, straylight, etc.). As the term implies, SLS can be used to obtain the so-called “static” parameters of the protein such as weight-average molecular weight. SLS of a macromolecular solution is described by the classical Rayleigh relationship

$$\Delta LS = (I_{\theta}/I_0)_{\text{solution}} - (I_{\theta}/I_0)_{\text{buffer}} = K (dn/dc)^2 M_w C,$$

In which:

I_{θ}/I_0 is the ratio of the intensities of the light scattered at angle θ and the incident light. This angle θ is between the direction of the incident and scattered light. ΔLS is the excess of light scattered at a given θ by the solution containing the scattering protein compared to the light scattered by the buffer. M_w is the weight-averaged molecular mass. C is the concentration in mg/ml. dn/dc is the specific refractive index increment of the protein, which relates changes in the refractive

index of the solution to the protein concentration. K is a constant, which depends on the refractive index of the solution without macromolecule (n), the wavelength of the light used (λ), the angle between the incident and scattered light (θ) and the distance between the scattering molecule and the detector (r); N is Avogadro's number.

$$K = (2\pi^2 n^2 / \lambda^4 N) \{ (1 + \cos^2 \theta) / r^2 \}$$

The above relation holds only if the proteins are small compared to the wavelength of the light in order not to cause an angular dependency of the light scattering. This condition is satisfied by globular proteins with molecular masses up to several million Dalton (Slotboom et al., 2008).

In DLS, fluctuations about the average intensity are measured rather than the absolute average intensity as required for SLS. The scattered intensity fluctuations are primarily due to the random diffusive motion of the protein molecules into and out of a focused laser beam. Just like in FCS, the fluctuations are directly related to the rate of diffusion of the molecule through the solvent, and this can be described by a correlation function. Since globular proteins are often assumed to be spherical in shape, the Stokes-Einstein relation can be used to estimate the apparent hydrodynamic radius R_h :

$R_h = kT / 6\pi\eta D(c)$, where k is Boltzmann constant, T is the absolute temperature, η is the solution viscosity, and $D(c)$ is the concentration-dependent translational diffusion coefficient of the protein, which is obtained from the correlation function (Wilson, 2003).

2.2 Protein folding in the cell

An additional challenge in the study of protein folding is to translate how the insights gained from biophysical studies of folding *in vitro* relate to the physiological process of folding in the cell, or in other words, taking into account a new layer of complexity arising from the crowded milieu in which the nascent proteins find themselves, and in which chaperone networks collaborate to assist folding of proteins, thereby shaping their energy landscapes.

Folding in the cell occurs in the presence of 300-400 g l⁻¹ of macromolecules. Under these conditions, although the concentration of each protein ranges from nM to μM (Ando & Skolnick, 2010), the distance between neighboring proteins is comparable to the size of the proteins. Crowding therefore has a tremendous effect on the thermodynamics and kinetics of processes, and it has indeed been shown to increase the propensity of non-folded proteins to aggregate due to excluded volume effect (Ellis & Minton, 2006).

An additional threat to the non-folded protein is imposed by the process of translation, whereby the risk of misfolding and aggregation is potentially increased, because incomplete polypeptide chains cannot fold into stable native conformations. For instance, the exit channel of the large ribosomal subunit, which is ~100 Å long, prevents the C-terminal 40-60 residues of the chain from participating in long-range interactions. As a consequence, productive folding can occur only after a complete protein or at least a domain (~50-300 amino acids) has emerged from the ribosome, and the nascent chains are exposed in partially folded, aggregation-sensitive states for prolonged periods of time, given the slow translation rate. The fact that proteins are synthesized by polysomes introduces yet another risk of aggregation by the process of translation (Hartl & Hayer-Hartl, 2009). Nonetheless, polysomes were shown to be arranged in such a way

as to maximize the distance between nascent chains on adjacent ribosomes, thereby reducing the probability of intermolecular interactions (Brandt et al., 2009).

Cells have evolved cunning schemes over the course of evolution to enable proteins to fold correctly, and avoid aggregation.

2.3 The cellular chaperone machinery

A molecular chaperone is defined as any protein which interacts, stabilizes or helps a non-native protein to acquire its native conformation but is not present in the final functional structure. Chaperones are found to be major players in many cellular activities, such as de novo folding, refolding of stress-denatured proteins, oligomeric assembly, intracellular protein transport and assistance in proteolytic degradation. Many chaperones are known as stress proteins or heat-shock proteins, as they are upregulated by cells under conditions of conformational stress in which the concentration of aggregation-prone folding intermediates increases. They are usually classified according to their molecular weight (Hsp40, Hsp60, Hsp70, Hsp90, Hsp100 and the so-called small Hsp proteins) (Hartl & Hayer-Hartl, 2009).

The chaperone pathways and networks acting in protein folding in the cytosol follow general organizational rules. Across all domains of life, two major principles of chaperone action, represented by (a) machinery that functions in stabilizing nascent polypeptides on ribosomes and initiating folding and (b) components that act downstream in completing the folding process (Fig. 6). The first category of factors includes chaperones that bind directly to the large ribosomal subunit in close proximity to the polypeptide exit site, such as bacterial Trigger factor, whereas the second category includes chaperones that do not bind directly to the ribosome and

mediate co- or post-translational folding through ATP-regulated binding, such as Hsp70 and Hsp60 (Hartl & Hayer-Hartl, 2009). Below is a description of the paradigmatic chaperones in the bacterial cytosol.

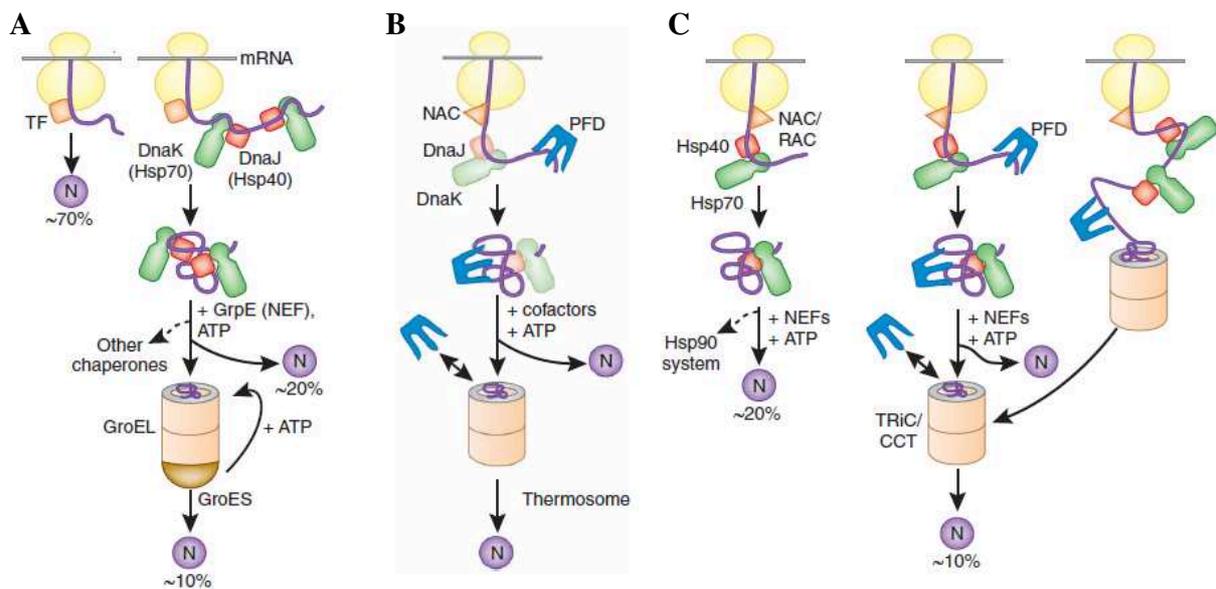


Figure 6: Model of the chaperone-assisted protein folding in the cytosol. A) Bacteria. Nascent chains generally interact with Trigger factor (TF), and most small proteins (~70% of total) may fold rapidly upon synthesis without further assistance. Longer chains interact subsequently with DnaK and DnaJ (Hsp70 system) and fold upon one or several cycles of ATP-dependent binding and release (~20% of total). About 10% of chains transit the chaperonin system (GroEL and GroES) for folding. N stands for native protein. (B) Archaea. PFD, prefoldin; NAC, nascent chain-associated complex. (C) Eukarya. Like TF, NAC probably interacts generally with nascent chains, but the role of NAC in folding is not yet clear. About 20% of chains reach their native states in a reaction assisted by RAC (ribosome-associated complex), Hsp70 and Hsp40. A fraction of these must be transferred to Hsp90 for folding. About 10% of chains are co- or post-translationally passed on to the chaperonin TRiC/CCT in a reaction mediated by Hsp70 and PFD, both of which interact directly with TRiC/CCT. PFD recognizes the nascent chains of certain TRiC substrates, including actin and tubulins. (Adapted from Hartl & Hayer-Hartl, 2009).

2.3.1 Trigger factor

Trigger factor (TF) is an abundant protein of ~50kDa size, which consists of an N-terminal ribosome-binding domain, a peptidylprolyl isomerase (PPIase) domain and a C-terminal domain (Ferbitz et al., 2004) (Fig. 7). It exists as a monomer when bound to the ribosome and as a dimer when free in the cytosol. The PPIase domain, which can catalyze prolyl cis-trans isomerization *in vitro*, functions as an auxiliary chaperone site independent of proline residues (Kaiser et al., 2006). With its two arm-like protrusions, the C domain is the major binding region for hydrophobic nascent chain segments (Lakshmipathy et al., 2007). The N domain provides binding contacts to ribosomal proteins L23 and L29 next to the polypeptide exit site, and TF stays on the ribosome for about 10-15 s (Ferbitz et al., 2004; Kaiser et al., 2006). Ribosome binding is thought to activate TF for nascent chain interaction by causing a conformational opening. Chain release from TF is ATP-independent, and is probably driven by the tendency of the bound polypeptide to bury hydrophobic regions. Nascent chains that interact only weakly may begin to fold co-translationally. However, when the nascent chain exposes strongly hydrophobic segments, Trigger factor leaves the ribosome but remains associated with the elongating chain, explaining how TF delays the folding of certain multidomain proteins relative to translation. The eventual dissociation of Trigger factor facilitates folding or polypeptide transfer to downstream chaperones such as DnaK (reviewed in Hartl & Hayer-Hartl, 2009).

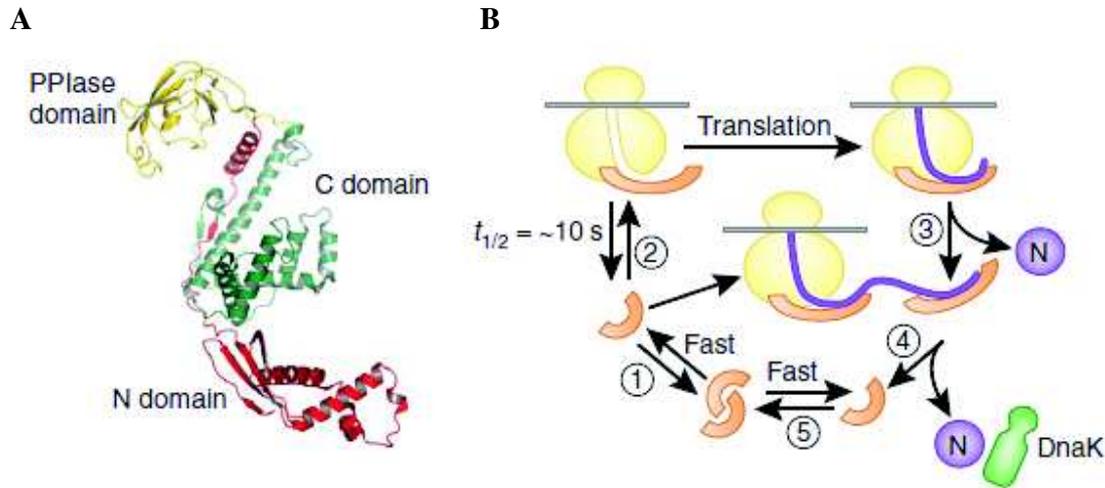


Figure 7: Structure and function of Trigger factor. A) The domain structure of TF (PDB 1W26). B) Model for the TF reaction cycle. (1) Free TF is in rapid equilibrium between monomeric and dimeric states. (2) TF monomer binds to nontranslating ribosomes with a K_d of $\sim 1 \mu\text{M}$. Ribosome binding causes a conformational expansion of TF and may activate TF for interaction with nascent chains. (3) Nascent chains that interact weakly with TF may begin to compact co-translationally in the vicinity of TF. Release of TF from the nascent chain coincides with TF dissociation from the ribosome and allows completion of folding to native state (N). (4) Structurally more complex proteins may interact strongly with TF. TF remains bound to the nascent chain after dissociating from the ribosome and a new TF molecule can enter at the ribosome. Eventual chain dissociation from TF facilitates transfer to DnaK or folding. (5) Released TF enters the monomer-dimer pool. (Adapted from Hartl & Hayer-Hartl, 2009).

2.3.2 The DnaK/J system

The DnaK/J system belongs to the Hsp70 family of proteins. These Hsps are involved in a plethora of folding processes, including protein trafficking and assistance in the proteolytic degradation of terminally misfolded proteins. They occur both as constitutively expressed and stress-inducible forms, and generally collaborate with chaperones of the Hsp40 (DnaJ) family and nucleotide-exchange factors (NEFs) in the ATP-regulated binding and release of non-native proteins. Their role in *de novo* folding begins by binding to nascent chains, without directly

interacting with the ribosome (with a few exceptions). DnaK consists of an N-terminal ATPase domain of ~ 40kDa and a C-terminal peptide-binding domain (PBD) of ~ 25kDa (Zhu et al., 1996) (Fig. 8). The latter domain consists of a β -sandwich subdomain, which recognizes extended ~7-residue segments that are enriched in hydrophobic amino acids (Rüdiger et al., 1997), and as more recently suggested, regions with substantial tertiary structure (Schlecht et al., 2011). Additionally, the PBD contains an α -helical lid, which closes upon ATP hydrolysis in the ATPase domain, thereby increasing the affinity of DnaK to the peptide. Hydrolysis of ATP to ADP is strongly accelerated by Hsp40, and is mediated via the J domain, which is present in all Hsp40s (Mayer et al., 2000). Hsp40s also interact directly with unfolded polypeptides and can recruit Hsp70 to protein substrates. Following ATP-hydrolysis, the GrpE NEF binds to DnaK ATPase domain and catalyzes ADP-ATP exchange, which results in lid opening and substrate release, thereby completing the reaction cycle (reviewed in Hartl & Hayer-Hartl, 2009). Substrates that are unable to fold at this stage are either recaptured by DnaK, or transferred to downstream chaperones called chaperonins.

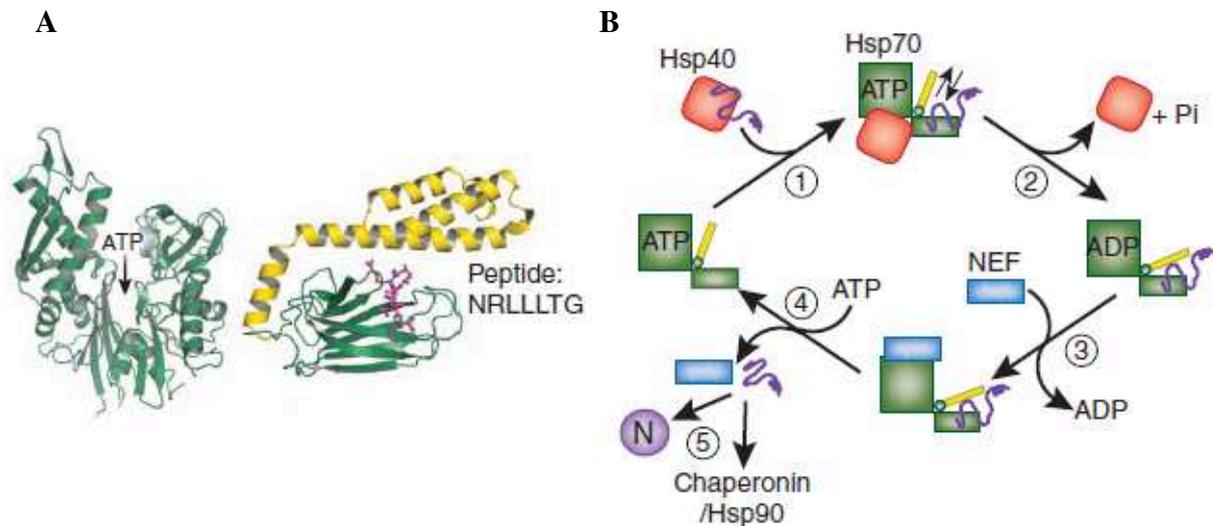


Figure 8: DnaK system and reaction cycle. A) Structures of the ATPase domain (PDB 1DKG) and the peptide-binding domain (PDB 1DKZ) of Hsp70 shown representatively for *E. coli* DnaK. The α -helical lid of the peptide binding domain is shown in yellow and the extended peptide substrate as a ball-and-stick model in pink. ATP indicates the position of the nucleotide-binding site. The amino acid sequence of the peptide is indicated. B) Hsp70 reaction cycle. (1) Hsp40-mediated delivery of substrate to ATP-bound Hsp70. (2) Hydrolysis of ATP to ADP, accelerated by Hsp40, results in closing of the α -helical lid and tight binding of substrate by Hsp70. Hsp40 dissociates from Hsp70. (3) Dissociation of ADP catalyzed by NEF. (4) Opening of the α -helical lid, induced by ATP binding, results in substrate release. (5) Released substrate either folds to native state (N), or is transferred to downstream chaperones or rebinds to Hsp70. (Hartl & Hayer-Hartl, 2009).

2.3.3 GroEL/ES

Chaperonins are large, multimeric complexes ~1MDa in size, consisting of a double-ring structure with central cavities. They are divided in two groups: Group I, also called Hsp60, are found in bacteria (GroEL), and organelles of bacterial origins. They are made up of two heptameric rings, and collaborate with Hsp10 (GroES in bacteria) which forms a lid that closes

the ring opening. Group II chaperonins are found in eukaryotic cytosol (TRiC/CCT), and archaea (Thermosome). The rings are made up of 8 to 9 subunits, and possess a built-in lid, in contrast to group I chaperonins; they are therefore independent of Hsp10s. Just like DnaK, the chaperonin cycle is also regulated by ATP. Nevertheless, chaperonins are unique in that they encapsulate proteins and allow their folding to occur inside a specialized nanocompartment, unimpaired by aggregation.

The discovery of chaperonin, more specifically the GroEL/ES system, was a serendipitous event in the 1970s. Georgopoulos and colleagues had isolated *E. coli* temperature-sensitive mutants which were unable to support bacteriophage growth (Georgopoulos et al., 1973). It turned out later on that the mutations in question were in the GroE operon, whose genes products, EL and ES, are essential for bacterial growth at all temperatures (Fayet et al., 1989). Because of the GroE chaperone machine's effect on phage morphogenesis and the fact that Hsp60, the eukaryotic GroEL homologue, is also involved in correct assembly of the large oligomeric Rubisco protein in plants, GroE was thought to be uniquely involved in macromolecular protein assembly. It wasn't until 1989 that the concept began to unravel, when different groups working on bacterial GroEL and mitochondrial Hsp60 published their findings. Cheng et al. showed that Hsp60 is essential for assembly of proteins imported into yeast mitochondria (Cheng et al., 1989), whereas Goloubinoff et al. showed that GroEL is necessary for prokaryotic Rubisco assembly in *E. coli* (Goloubinoff et al., 1989). The important breakthrough was when Ostermann et al. showed for the first time that Hsp60 functions in the actual folding of authentic mitochondrial proteins in an ATP-dependent manner (Ostermann et al., 1989). Shortly after, it was demonstrated that EL/ES assists the refolding of prokaryotic

Rubisco in a purified system (Goloubinoff et al., 1989). The ensuing years witnessed an explosion in the knowledge of mechanistic details of this machinery.

2.3.3.1 Structural consideration

The structure of the EL/ES complex in the absence of nucleotide has been solved in 1994 (Braig et al., 1994). GroEL is shown to be a porous thick-walled cylinder, with a substantial central channel. The structure reveals two heptameric rings stacked back to back forming an extensive interface with one another, across a flat equatorial plane (Fig. 9). The subunits are arranged with nearly exactly 7-fold rotational symmetry. Each subunit is ~57kDa, and consists of three domains: the equatorial domain, that contains an ATP binding domain and most of the inter and intra ring contacts; the apical domain, which forms the opening of the central cavity and contains the substrate binding site that binds non-native polypeptides through hydrophobic interactions, and an intermediate hinge-like domain connecting these two domains. The central channel of GroEL functions as two cavities, one in each ring, that are separated from each other by the crystallographically disordered 23-amino-acid C-terminal segments of the seven subunits, which inhibit substrate escape from the equatorial domain (Weissman et al., 1996). Throughout the 1990s, it was becoming clear that there is a physiological requirement for GroES in the folding of substrate proteins and that the folding reaction in the presence of GroES and ATP is qualitatively different from one in their absence.

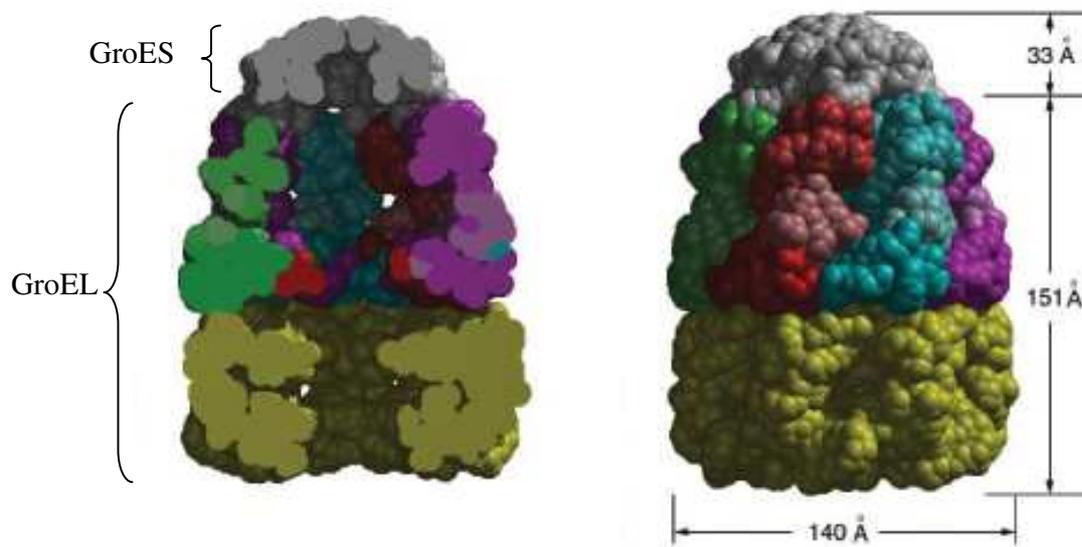


Figure 9: Dimensions of the GroEL/ES complex. Shown is a slice through a space filling model of the GroEL/ES/ADP complex, in addition to the dimensions of the complex. In the presence of nucleotide, the GroES caps one end of the GroEL, thereby providing a space where substrate protein can fold in isolation. The trans ring is colored uniformly green, GroES is uniformly gray, whereas different subunits of the cis ring are colored individually. (Sigler et al., 1998).

In 1997, the crystal structure of the ADP-bound EL/ES complex was solved (Xu et al., 1997), and revealed the *en bloc* movements which the GroES-bound ring, termed the cis ring, undergoes upon nucleotide binding (Fig. 10). First, the intermediate domain swings down towards the equatorial domain and the central cavity, pivoting approximately 25° around Pro137 and Gly410. This movement locks the nucleotide binding sites and generates new interactions with the bound nucleotide and the equatorial domain. Second, the apical domain swings up 60° and twists around the long axis of the domain about 90° , forming new interactions with nearby apical domains and leading to an interaction with the mobile loop of GroES. The pivot point of the apical domain's movement is again a slender link, in this case a pair of glycine residues (Gly

192 and Gly 375) between the intermediate and apical domains. Eventually, the domain rearrangements result in burying hydrophobic residues and hence changing the environment inside the GroEL-GroES cavity to hydrophilic. In addition the volume of the cavity is approximately enlarged by two-fold (Chen et al., 1994); it becomes about $175\,000\text{ \AA}^3$ and is able to accommodate single partially folded polypeptides up to ~60 kDa.

Similar to GroEL, GroES forms a heptameric ring of ~10 kDa subunits. Each GroES subunit is folded into a single domain, which contains nine β -strands with one exceptionally long β -hairpin loop, the so-called mobile loop. The GroEL-GroES contact is mediated through this mobile loop (Landry et al., 1993). It was shown to contain 16 amino acids which appear highly mobile in the uncomplexed GroES but become more structured upon interacting with GroEL. Because the hydrophobic binding regions of GroEL for polypeptide overlap with those for GroES (Fenton et al., 1994), GroES binding leads to the displacement of the bound polypeptide from the apical domains into the central cavity which is then permissive for folding (Langer et al., 1992; Weissman et al., 1996).

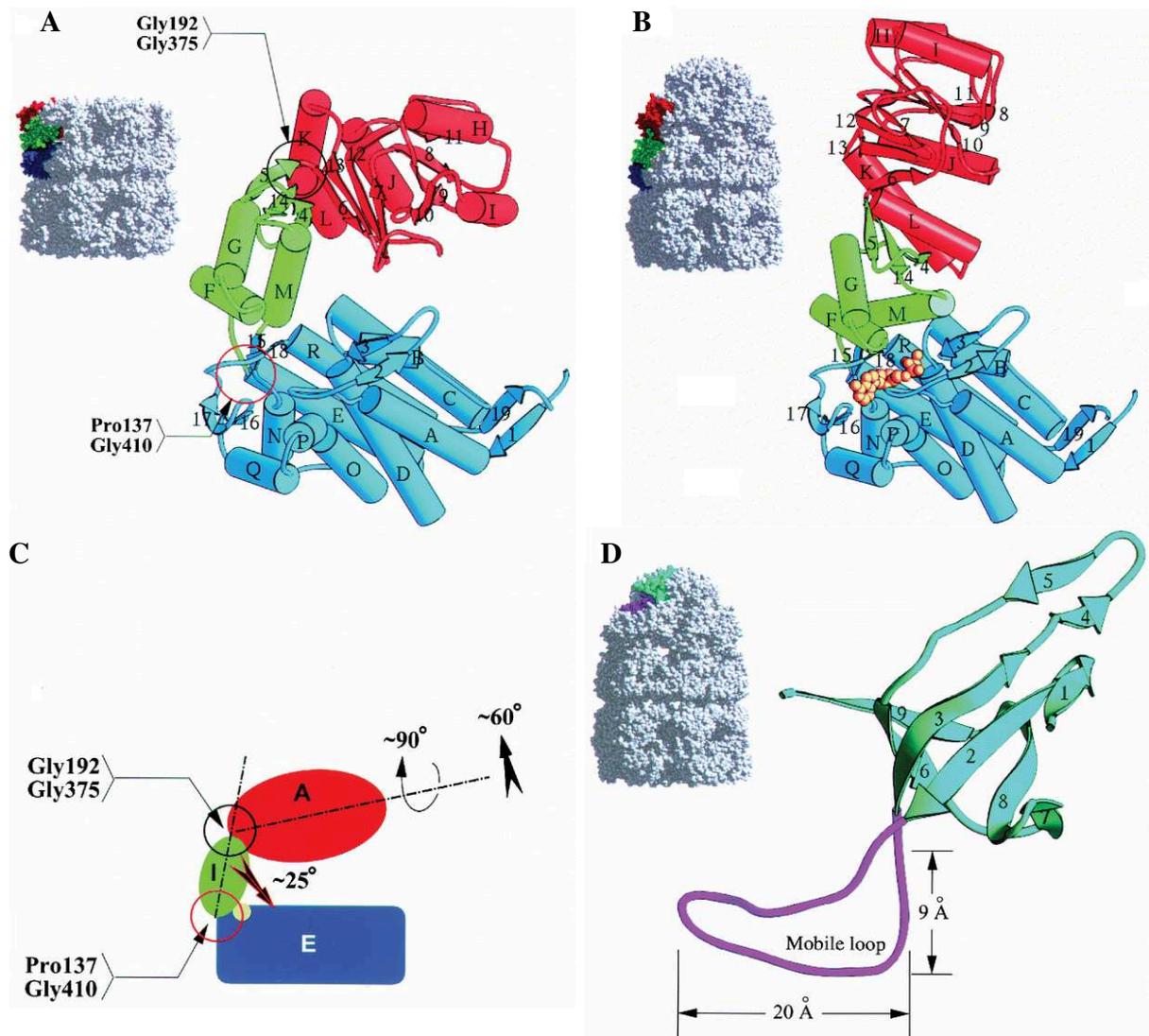


Figure 10: General architecture and domain movements in the EL/ES system. The upper panels show ribbon diagrams of an individual subunit of unliganded (A) and liganded (B) GroEL. The orientation of the representative subunit is the same as the colored subunit in the nearby space-filling model. The equatorial, intermediate, and apical domains are blue, green, and red, respectively. ADP in the right-hand structure is a yellow space-filling model. (C) Schematic representation of GroEL showing diagrammatically the en bloc movements that occur around the pivot points at the ends of the intermediate domain. Domains are colored as in the upper panels, and the small yellow

circle on the top of the equatorial domain represents the nucleotide. (D) The left panel shows a space filling model of a side view of a GroEL-GroES complex. The GroES subunit is colored individually. The right panel shows a side view of a single GroES subunit in a ribbon drawing. Mobile loop as well as the GroEL interacting loop is indicated. (Xu *et al.*, 1997).

2.3.3.2 Mechanism and proposed modes of action

Significant refinements to the above-mentioned conformational changes occurring during the chaperonin cycle have been made (Chaudhry *et al.*, 2003a; Cliff *et al.*, 2006a; Ranson *et al.*, 2001; Ranson *et al.*, 2006). Although the complete movie of such a refolding reaction is not available at the moment, the different provided snapshots are starting to align and provide the peak moments of that movie.

It is generally accepted that the initial polypeptide acceptor state *in vivo* and in a cycling reaction is an asymmetric, ADP-containing complex (Rye *et al.*, 1997; Xu *et al.*, 1997), although more recent experiments suggest that ATP-containing complexes could also function as an acceptor state (H. Rye, conference communication). ATP and non-native polypeptide bind rapidly to the open ring of the ADP asymmetric complex, with ATP producing downward rigid body movement of the intermediate domain via contacts with the nucleotide pocket (Chaudhry *et al.*, 2003) and a small elevation of $\sim 10^\circ$ and counterclockwise twist of $\sim 25^\circ$ of the apical domains (Ranson *et al.*, 2001), the latter of which directly enable GroES binding. Such binding induces a further $\sim 50^\circ$ elevation and $\sim 115^\circ$ clockwise twisting of the apical domains, thereby removing hydrophobic peptide-binding surfaces from the central cavity and ejecting polypeptide into the cavity. This state is the one which has been crystallographically observed previously (Xu *et al.*, 1997). Nevertheless, there are still some matters of debate, in addition to unresolved structures which are hypothesized to exist along the folding reaction. For instance, the order of arrival of

peptide and ATP is not settled (Horwich & Fenton, 2009), and awaits experimental setups where one could monitor this order in an *in vivo* context. It is also unclear how the apical domains twist from the -25° (the counterclockwise twist reached upon ATP binding) to the final 90° clockwise twist evident in the GroEL-ES-ADP complex. Additionally, there are uncertainties with regards to the binding of ES to an open EL-ATP-polypeptide ring; the mobile loop of ES, which associates to the same hydrophobic surfaces as does the substrate, could effectively compete with polypeptides for the apical domains, leading to premature escape. Some fluorescent studies have suggested a brief phase when the apical surfaces might be occupied with both substrate and mobile loop of ES, which would then ensure proper substrate encapsulation rather than premature escape (Cliff et al., 2006b; Taniguchi et al., 2004). The step would then be followed by the large twist and elevation of the apical domain. Nevertheless, this structure has not been observed directly.

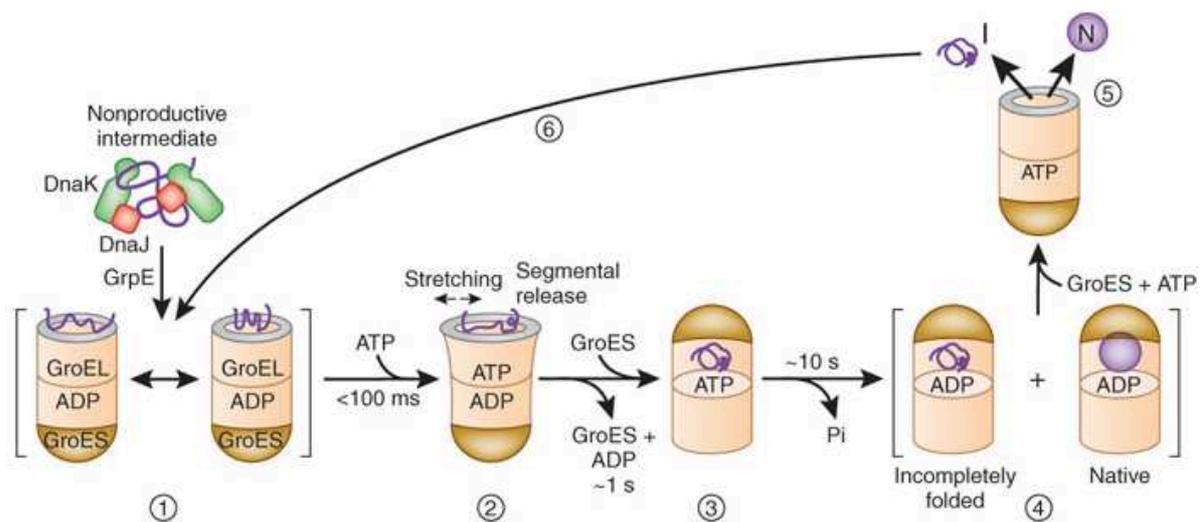


Figure 11: Working model summarizing the conformational changes in a substrate protein upon transfer from Hsp70 system to GroEL and during GroEL-GroES-mediated folding. (1) Substrate protein may be delivered to GroEL by DnaK-DnaJ in a non-aggregated, but kinetically trapped, state. Upon binding to GroEL it

undergoes local unfolding to an ensemble of expanded and more compact conformations. (2) ATP-dependent domain movement of the apical GroEL domains result in stretching of tightly bound regions of substrate and in release and partial compaction of less stably bound regions. (3) Compaction is completed upon substrate encapsulation by GroES. (4) Folding in the chaperonin cage. (5) Substrate release upon GroES dissociation. (6) Rebinding of incompletely folded states. N, native state; I, folding intermediate. (Hartl & Hayer-Hartl, 2009).

This folding-active cis complex is the longest-lived state in the folding cycle. ATP hydrolysis, occurring with a half-time of ~10 s, primes the cis complex for discharge of its ligands by weakening affinity of GroEL for GroES (Rye et al., 1997). The reaction perpetuates by binding of ATP to the trans ring which sends an allosteric signal to the opposite ring, leading to ejection of its ligands: GroES, substrate polypeptide, and ADP (Rye et al., 1997) (Fig. 11). ATP binding has been shown to be positively cooperative within rings, and negatively cooperative between rings, overall described as a “nested” behavior (Yifrach & Horovitz, 1995). The mechanism for positive cooperativity has been suggested to lie at the level of the movement of the intermediate domains upon ATP binding. More specifically, the rotation of the intermediate domain disrupts a salt-bridge between that domain and the apical domain of the nearby subunit (E386-R197), and leads to the formation of a new salt-bridge with the equatorial domain of the neighbouring subunit (E386-K80) (Ranson et al., 2001). The mechanism of negative cooperativity has been suggested to lie at the level of the so-called helix D (residues 89-109) of the equatorial domain, which connects the nucleotide binding sites of the two rings via an inter-ring contact (Ranson et al., 2001). ATP (and polypeptide) binding to the trans ring discharges GroES and polypeptide from the old cis ring and leads to a new round of cis encapsulation and folding in this opposite ring. It is still not very clear how the binding of ATP to the trans ring leads to the discharge. Based on a EM structure of the slow ATP hydrolyzing mutant of EL, the EL-D398A complexed with ADP and ES (to which ATP was later on added),

it was proposed that ATP binding in trans of the ES-ADP-bound ring leads to the previously observed 25° counterclockwise twist of the apical domain, while nevertheless still maintaining the salt bridge with the intermediate domain of the nearby subunit. This would then lead to a strained conformation which can only be relieved by ejection of the load on the other ring (Ranson et al., 2001).

Given that the 10 s or so of folding time in the cis complex is not sufficient for most molecules of stringent, GroEL-GroES-dependent substrate proteins to fold to native form; typically, only a few percent reach the native state. The other molecules are discharged in non-native states that, *in vitro*, are rebound by other GroEL complexes for another attempt at folding. In the cell, such non-native forms can also partition to other chaperones or the degradative apparatus, their fate depending on the relative concentrations and affinities of the various components (Horwich & Fenton, 2009). It has also been argued that under the crowding conditions prevailing in the cytoplasm, substrates could complete folding without partitioning between different EL complexes because the “leakiness” of the chaperonin under physiological conditions is adjusted not to allow such partitioning (Martin & Hartl, 1997).

It had initially been suggested that EL/ES accelerates folding of proteins merely by preventing aggregation; John Ellis coined the term “Anfinsen cage” in 1993, to describe the model where proteins fold at infinite dilution inside the cis chamber, without the complications of folding in bulk solution (Saibil et al., 1993). Another model, named “forced unfolding”, suggests that EL exerts an unfolding action on the substrate via multiple rounds of binding and release, and thereby removes the proteins from kinetically trapped states (Shtilerman et al., 1999). A more recent model entitled “steric confinement” proposes that the cavity actively

remodels the energy landscape of the folding and destabilizes trapped intermediates, by providing spatial confinement, in addition to a wall with a mixture of polar and hydrophobic properties (Brinker et al., 2001; Tang et al., 2006).

2.3.3.3 Substrates of EL

The GroEL proteome was thoroughly characterized in 2005 (Kerner et al., 2005). Approximately 250 of the ~2400 cytosolic *E. coli* proteins were shown to interact with GroEL, and were assigned to three different classes depending on their GroEL dependence; class I and II substrates are only partially chaperonin dependent, whereas class III substrates are obligate ones. Approximately 85 substrates were assigned to class III, and these were shown to occupy 75-80% of the chaperonin capacity. Additionally, those obligate substrates are between 20 and 50 kDa, and have complex α/β and $\alpha+\beta$ domain topologies. Such proteins are stabilized by many long-ranged contacts, and are predicted to have a marked propensity to populate kinetic intermediates during folding. More specifically, ~45% of class III proteins share the $(\beta\alpha)_8$ TIM barrel domains, and it is thought that GroEL may have helped in evolving the TIM-barrel fold into one of the most versatile structural platforms for implementation of enzymatic function. The finding that thirteen of the proteins assigned to class III are essential *E. coli* proteins explains the fact that the EL/ES system is essential under all growth conditions.

2.3.4 Additional *E. coli* chaperones

Additional chaperones operate in the bacterial cytosol. One of these molecular chaperones, Hsp90, is an abundant protein in the cytosol of eukaryotes and bacteria, where it is called HtpG. Unlike the eukaryotic Hsp90, the prokaryotic counterpart seems to lack established client

proteins, as the deletion of the *htpG* gene does not lead to striking phenotypes (Buchner, 2010). It was recently shown that cyanobacterial Hsp90 is an important player in the assembly and stability of phycobilisomes, large protein complexes involved in light harvesting (Sato et al., 2010).

Bacteria also possess the small heat shock proteins IbpA and IbpB, which were initially identified as proteins associated with inclusion bodies, and later shown to be present in aggregates formed by heat shock. Deletion of the *ibpA* and *ibpB* genes results in decreased viability manifested during prolonged growth at extreme temperatures and correlates with the increase in intracellular protein aggregation. IbpA, in contrast to IbpB which forms polydisperse oligomers (Jiao et al., 2005), forms fibrillar structures that are atypical for sHsp (Ratajczak et al., 2010). In both cases, exposure to high temperatures stimulates dissociation into smaller molecules and the chaperone activity.

Bacteria additionally contain Hsp100/Clp proteins of the AAA+ (ATPase Associated with various cellular Activities) superfamily. Those are ATP-dependent chaperones that transfer proteins into the proteolytic chamber of an associated self-compartmentalized protease. These AAA+ proteins, for example, ClpA or ClpX, are assembled into hexameric rings with a narrow pore in the middle; they use the energy generated by ATP hydrolysis to unfold protein substrates and to translocate the unfolded polypeptide chain through the central pore, into the chamber of an associated barrel-like protease complex, such as ClpP or ClpQ, where the peptide bonds are hydrolysed. The Hsp100/Clp also cooperates with the Hsp70 system, to act as a chaperone system that can disaggregate and refold protein aggregates (reviewed in Kirstein et al., 2009).

2.4 Aim of the study

This study focuses on one model substrate of the chaperonin system, namely DM-MBP, which folds spontaneously slow, and exhibits rate acceleration in the presence of the chaperonin system. In the light of conflicting literature (Tang et al., 2006; Apetri and Horwich, 2008), one aim of the study is to show the cause of the slow spontaneous folding of DM-MBP, which turns out to be the population of a kinetically-trapped intermediate with a large entropic folding barrier. Another aim of the study is to clarify the mechanism of action of the chaperonin. Our data strongly support the “steric confinement” model, whereby rate acceleration of folding is mediated both by charges lining the cavity, as well as geometric confinement.

3. Materials and Methods

3.1 Materials

3.1.1 Chemicals

Acetic acid	Merck
Adenosine 5'-(β,γ -imido)triphosphate tetralithium salt (AMP-PNP)	Sigma-Aldrich
Adenosine triphosphate, disodium salt (ATP)	Sigma-Aldrich
Agarose (SeaKem LE)	Cambrex Bio Science
Anti-mouse IgG secondary antibody	New England Biolabs
Ammonium persulfate (APS)	Sigma-Aldrich
Ampicillin	Merck
Amylose resin	New England Biolabs
Atto-532 C5 maleimide	Atto-Tec
Atto-647N C5 maleimide	Atto-Tec
Bacto agar	Difco
Bacto trypton	Difco
Bacto yeast extract	Difco
Bovine Serum Albumin (BSA)	Sigma-Aldrich
Bromophenol blue	Sigma-Aldrich
Calcium chloride	Merck

Complete EDTA-free protease inhibitor	Roche
Coomassie brilliant blue R-250	Roth
Dextran 40	Sigma-Aldrich
Dimethylsulfoxide (DMSO)	Merck
Dithiothreitol (DTT)	Roche
ECL™ detection kit	Amersham Pharmacia Biotech
Ethanol	Merck
Ethidium bromide	BioRad
Ethylenediaminetetraaceticacid –sodium salt (EDTA)	Merck
Ficoll 70	Sigma-Aldrich
Formaldehyde	Sigma-Aldrich
Glucose	Sigma-Aldrich
Glycerol	Merck
Glycine	Roth
Guanidium hydrochloride (GuHCl)	Sigma-Aldrich
HEPES	Sigma-Aldrich
Hydrochloric acid (37%)	Merck
Isopropyl-β-D-thiogalactopyranoside (IPTG)	BioMol
Magnesium chloride	Merck
Maltose	Sigma-Aldrich
β-mercaptoethanol	Sigma-Aldrich
Methanol	Merck
Mouse anti-MBP antibody	New England Biolabs

Phenyl-methyl-sulfonyl-fluoride (PMSF)	Sigma-Aldrich
Polyacrylamide/bisacrylamide solution 30 % (30 : 0.8)	Roth
Potassium hydroxide	Sigma-Aldrich
Silver nitrate	Sigma-Aldrich
Sodium chloride	Merck
Sodium dodecylsulfate (SDS)	Sigma-Aldrich
Sodium hydroxide	Sigma-Aldrich
Sucrose	Merck
N, N, N', N'-Tetramethylethylenediamine (TEMED)	Sigma-Aldrich
Tris-base	Sigma-Aldrich
Triton X-100	Sigma-Aldrich
Tween-20	Calbiochem

3.1.2 Enzymes

Benzonase	Merck
Lysozyme	Sigma-Aldrich
Pfu DNA polymerase	Stratagene
Restriction enzymes	New England Biolabs
Shrimp Alkaline Phosphatase	Roche
T4 DNA ligase	New England Biolabs

3.1.3 Materials

Centricon 10 kDa cut-off	Amicon
--------------------------	--------

Centricon 30 kDa cut-off	Amicon
Microcon 10 kDa cut-off	Amicon
Microcon 30 kDa cut-off	Amicon
Nitrocellulose transfer membrane	Whatman Schleicher & Schuell
Sterile filter 0.22 μm	Millipore
Sterile filter 0.45 μm	Millipore

3.1.4 Instruments

AIDA gel imaging software version 2.31	Raytest
ÄKTA Explorer 100	Amersham Pharmacia Biotech
Balance AG285, PB602	Mettler Toledo
Centrifuges: Avanti J-25, Avanti J20 XP, J-6B, GS-6R	Beckmann
Centrifuges 5415C and 5417R	Eppendorf
Chromatography columns (HiPrep Desalting, MonoQ, HiTrap Heparin, Sephacryl S200/S300, Superdex 200, Superose 6, Sephadex G25 (NAP-5, NAP-10); chromatography resins: Q-Sepharose, DE52, Source 30 Q, Source 30 S)	Amersham Pharmacia Biotech
Deionization system MilliQ plus PF	Millipore
DAWN EOS multi-angle light scattering	Wyatt Technology
Electrophoresis chambers MiniProtean 3	Bio-Rad
Electrophoresis power supply Power PAC 300	Bio-Rad
Fluorescence spectrometer Fluorolog 3	HORIBA Jobin Yvon
FPLC systems	Amersham Pharmacia Biotech

EmulsiFlex high pressure homogenizer	Avestin
Gilson Pipetman (2, 10, 20, 100, 200, 1000 µl)	Abimed
HDMS ESI-QToF mass spectrometer	Waters Synapt
Incubators Innova 4430	New Brunswick Scientific
Luminescent Image Analyzer LAS-3000	FUJIFILM
Mini Trans-Blot Electrophoretic Transfer Cell	Bio-Rad
Microtime 200 confocal microscope	PicoQuant
PCR-Thermocycler T3	Biometra
pH meter Accumet Basic	Fisher Scientific
SMART system	Amersham Pharmacia Biotech
Sonicator Ultrasonic Processor XL	Misonix Inc.
Spectrophotometer DU 640 UV/VIS	Beckmann
Spectrophotometer LS50	Perkin-Elmer
Spectropolarimeter J-715	Jasco
Stopped-Flow Reaction Analyser Sx.18MV	Photo Physics
SZX10 Stereomicroscope	Olympus
Thermomixer Comfort	Eppendorf
Vortex	Ikamag
Water bath	Bioblock Scientific

3.1.5 Media

LB medium: 10 g/l tryptone, 5 g/l yeast extract, 5 g/l NaCl, (+ 15 g/l agar for solid medium). Adjusted to pH 7.0 with NaOH (Sambrook *et al.*, 1989).

SOC medium:	20 g/l tryptone, 5 g/l yeast extract, 0.5 g/l NaCl, 0.186 g/l KCl, 0.95 g/l MgCl ₂ . Twenty ml of filter sterilized 1M glucose were added after autoclaving (Sambrook <i>et al.</i> , 1989).
MacConkey medium:	40 g/l of MacConkey base was autoclaved. Maltose was then added to a final concentration of 1% (solid medium).

3.1.6 Ampicillin stock solutions

Ampicillin was prepared as a stock of 100 mg/ml in 50% ethanol, filter-sterilized and stored in -20°C. This stock was diluted a 1000x in growth media.

3.2 Bacterial strains and plasmids

3.2.1 *E. coli* strains

DH5α F'	F' / <i>endA1 hsdR17</i> (rk-, mk+) <i>glnV44 thi-1 recA1 gyrA</i> (Nal ^r) <i>relA1Δ</i> (<i>lacIZYA-argF</i>) <i>U169 deoR</i> (<i>φ80dlacΔ</i> (<i>lacZ</i>) <i>M15</i>)
BL21(DE3) Gold (Stratagene)	B strain, F- <i>dcm+</i> <i>Hte ompT hsdS</i> (r _B - m _B -) <i>gal l</i> (DE3) <i>endA Tet^r</i>
Pop6499	MC4100 <i>malT^c ΔmalE444 recA1 srl::Tn10</i>

3.2.2 Plasmids

GroEL, the chaperonin charge mutant KKK2 and MBP mutants (SM-MBP (Y283D), DM-MBP (V8G-Y283D) were constructed in a pCH vector backbone by Yun Chi Tang (Tang et al., 2006). DM-MBP (A52C-P298C) (Sharma et al., 2008), DM-MBP (N18C-D296C), DM-MBP (D184C-K362C) and DM-MBP (N18C-D296C-D184C-K362C) mutants were generated by site-directed mutagenesis in a pCH vector on the background of DM-MBP. GroES was constructed

in a pET11a vector inserted via the NdeI and BamHI sites (Brinker et al., 2001). GroEL-V263S was constructed by site-directed mutagenesis in a pET11a vector backbone containing GroEL. All the MBP mutant constructs with leader sequence were constructed by site-directed mutagenesis in a pPD1 vector (containing leader sequence + MBPwt), which was a kind gift from Dr. Betton (Duplay et al., 1984).

3.3 Molecular cloning techniques

3.3.1 Preparation and transformation of *E. coli* competent cells

For preparation of chemically-competent *E. coli* cells, a single colony was used to inoculate 500 ml LB medium (including antibiotic, if applicable) and grown to an optical density (OD₆₀₀) of 0.25-0.5 at 37°C. The cells were then chilled on ice for 15 min and harvested at 5000 x g for 10 min at 4°C. The cell pellet was washed with 80 ml ice-cold Ca/glycerol buffer (10 mM PIPES, 60 mM CaCl₂, 15 % glycerol; pH 7.0, adjusted with NaOH, and filter-sterilized) once and incubated with additional 80 ml Ca/glycerol buffer on ice for 30 min. Finally, the cells were pelleted and resuspended in 6 ml of Ca/glycerol buffer. 100 µl aliquots were frozen in liquid nitrogen and stored at -80°C.

For transformation, ~50 µl competent cells were mixed with 0.05-0.2 µg plasmid DNA or 1-5 µl ligation reaction and incubated on ice for 15 min. The cells were heat-shocked at 42°C for 45-90 s and subsequently placed on ice for 2 min. 1 ml of LB medium was added and the cells were shaken at 37°C for 1 h. The cell suspension was then plated on selective plates and incubated at 37°C, until colonies had developed (typically 10-16 h).

3.3.2 Plasmid purification

LB medium containing the appropriate antibiotic was inoculated with a single *E. coli* colony harboring the DNA plasmid of interest and shaken 8-16 h at 37°C. Plasmids were isolated using the Wizard *Plus* SV Minipreps DNA purification System from Promega.

3.3.3 PCR-mediated mutagenesis

PCR (polymerase chain reaction)-mediated site-directed mutagenesis was performed according to the instructions of the QuickChange Site-directed Mutagenesis kit by Stratagene.

3.3.4 DNA analytical methods

DNA concentrations were measured by UV absorption spectroscopy at $\lambda=260$ nm on a Nanodrop 1000 (Thermo Scientific). Agarose gel electrophoresis was performed in TAE buffer (40 mM Tris, 1 mM EDTA, 20 mM acetic acid) and 1-2 % TAE-agarose gels, supplemented with 1 $\mu\text{g/ml}$ ethidium bromide, at 4-6 V/cm. DNA sequencing was performed by the Microchemistry Core Facility at the Max Planck Institute of Biochemistry.

3.4 Protein purification

3.4.1 GroEL expression and purification

GroEL was purified with modifications to the protocol described by Hayer-Hartl *et al.* (1996). *E. coli* BL21 (DE3) Gold cells harboring the plasmid pCH-GroEL were grown in 6 l LB medium containing 100 mg/l ampicillin at 37°C to OD_{600} of ~0.5. The induction was then

proceeded by the addition of 1 mM (final concentration) of IPTG to the culture for 5-6 h. After harvesting the cultures by centrifugation for 30 min at 2500 x g, cells were resuspended in 100 ml lysis buffer (50 mM Tris-HCl pH 7.5, 50 mM NaCl, 1 mM DTT, 1 mM EDTA) and Complete protease inhibitor (1 tablet/25 ml). The suspension was further treated with lysozyme (~0.5 mg/ml) and benzonase (~500 units) for 60 min at 4°C. Lysis was achieved by homogenization of the cell suspension in an EmulsiFlex C5 device kept on ice. Cell debris was removed by ultracentrifugation for 60 min at 40,000 x g, 4°C and the supernatant subsequently passed through 0.2 µm filter. The supernatant was applied to a 400 ml DE52 column attached to an ÄKTA Explorer chromatography system. After washing with two column volumes of the lysis buffer, the protein was eluted using a NaCl gradient from 50 mM to 600 mM in five column volumes. The GroEL containing fractions were collected and dialyzed in 5 l lysis buffer overnight at 4°C. The desalted supernatant pool was applied onto a 20 ml MonoQ column, equilibrated in 30 mM Tris-HCl pH 7.5, 1 mM DTT, 1 mM EDTA and GroEL eluted with a NaCl gradient from 0 to 0.5 M. GroEL containing fractions were collected and dialyzed against 30 mM Tris-HCl pH 7.5, 30 mM NaCl, 1 mM DTT and 1 mM EDTA. The sample was then applied to a 4 x 5 ml Heparin Sepharose column (HiTrap Heparin) and eluted with 30 mM Tris-HCl pH 7.5 with a NaCl gradient from 0 to 0.5 M NaCl. GroEL-eluted flow through was collected and concentrated to less than 5 ml in 50 kDa cut-off Centriprep concentrators. Finally the concentrated sample was applied to a Sephacryl S 300 (XK 26/60) size exclusion column equilibrated in 30 mM Tris-HCl pH 7.5, 30 mM NaCl, 1 mM DTT, 1 mM EDTA and 10% glycerol. GroEL oligomer (approximate size 800 kDa) fractions were collected and concentrated to ~35 mg/ml (equivalent to 44 µM of GroEL oligomer). Protein concentration was determined

based on extinction coefficient of GroEL ($\epsilon_{280}=126,800 \text{ M}^{-1}\text{cm}^{-1}$). And aliquots were flash frozen in liquid N_2 and stored at -80°C . Total yield of GroEL was typically $\sim 600 \text{ mg}$.

3.4.2 GroES expression and purification

The expression and purification of GroES was similar as GroEL as described above, including the induction, lysis and centrifugation procedures. The supernatant was applied to a 400 ml DE52 column attached to an ÄKTA Explorer chromatography system. After washing with two column volumes of the above buffer, the protein was eluted using a NaCl gradient from 50 mM to 500 mM in five column volumes. GroES containing fractions were collected and dialyzed in 5 l lysis buffer overnight at 4°C . The desalted pool was applied into a 20 ml MonoQ column. Proteins were eluted in 30 mM Tris-HCl pH 7.5, 1 mM DTT, 1 mM EDTA and a NaCl gradient from 0 to 0.5 M. GroES containing fractions were collected and concentrated to less than 5 ml in 10 kDa cut-off Centriprep concentrators. The concentrated sample was finally applied to a Sephacryl S 200 (XK 26/60) size exclusion column equilibrated in 30 mM Tris-HCl pH 7.5, 30 mM NaCl, 1 mM DTT, 1 mM EDTA and 10% glycerol. GroES oligomer (approximate size 70 kDa) fractions were collected and concentrated to $\sim 15 \text{ mg/ml}$ based on extinction coefficient of GroES ($\epsilon_{280}=8,943 \text{ M}^{-1}\text{cm}^{-1}$). Aliquots were frozen in liquid N_2 and stored at -80°C . Total yield of GroES was typically $\sim 400 \text{ mg}$.

3.4.3 MBP and MBP mutants expression and purification

MBP and MBP mutants were purified using an amylose affinity column (New England Biolab). *E. coli* BL21 (DE3) Gold cells harboring the plasmid pCH-MBP wild-type and pCH-MBP mutants were grown in 6 l LB medium containing 100 mg/l ampicillin at 37°C to OD_{600} of

~0.1. The induction was then continued by adding 0.1 mM final concentration of IPTG to the culture for 12-16 h at 25°C. After harvesting the cultures by centrifugation for 30 min at 2500 x g, cells were resuspended in 100 ml amylose buffer (20 mM Tris-HCl pH 7.5, 200 mM NaCl, 1 mM DTT, and 1 mM EDTA) containing Complete protease inhibitor (1 tablet/25 ml). The same lysis and centrifugation conditions as used for GroEL purification were applied here. The supernatant was next dialyzed in amylose buffer to remove cellular maltose and slowly loaded on to a 100 ml amylose column. After washing with 12 column volumes of amylose buffer, MBP was eluted with amylose buffer containing 10 mM maltose. Fractions containing MBP were collected and dialyzed in 5 l amylose low salt buffer (20 mM Tris-HCl pH 7.5, 50 mM NaCl, 1 mM DTT, and 1mM EDTA) overnight at 4°C. MBP was concentrated to 10 mg/ml, and aliquots were frozen in liquid N₂ and stored at -80°C. Typical yield of WT-MBP from 6 l culture was ~ 500 mg, while the SM-MBP, DM-MBP gave yields of ~200 mg and ~100 mg respectively, due to largely partition of the expressed protein in the insoluble fractions. Protein concentration was determined using extinction coefficient of $\epsilon_{280}=64,720 \text{ M}^{-1}\text{cm}^{-1}$ of WT-MBP, $\epsilon_{280}=63,440 \text{ M}^{-1}\text{cm}^{-1}$ of SM-MBP and DM-MBP.

3.5 Protein analytical methods

3.5.1 Determination of protein concentration

Protein concentrations were determined spectrophotometrically by A_{280} (in 6 M GuHCl), based on the theoretical extinction coefficient of the respective protein at $\lambda=280$ nm as calculated by the ProtParam tool at the ExPASy proteomics server (<http://www.expasy.org>). Molar concentrations of chaperones are expressed for the native state oligomer, while the GroEL substrates are presented as monomer.

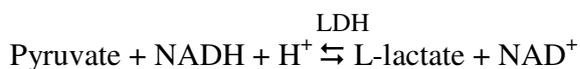
3.5.2 SDS-PAGE

SDS-Polyacrylamide gels were prepared according to Maniatis's Molecular Cloning Laboratory Manual. SDS-PAGE was performed using a discontinuous buffer system (Laemmli, 1970) in BioRad Mini-Protean 3 electrophoresis chambers employing a constant current of 30 mA/gel in 50 mM Tris-Base, 380 mM glycine, 0.1% SDS (pH 8.3). Protein samples were prepared for SDS-PAGE by mixing with 5x Laemmli buffer (Laemmli, 1970) (final concentration of 1x Laemmli buffer: 60 mM Tris-HCl, pH 6.8, 1% SDS, 10% glycerol, 0.01% Bromophenol blue, 0.1 mM β -mercaptoethanol) and boiling samples at 95°C for 3-5 min before loading onto a gel. After electrophoresis, gels were stained with Coomassie blue staining solution (0.1% Coomassie brilliant blue R-250, 40% ethanol, 7% acetic acid) for 1 h or longer and destained in 20% ethanol, 7% acetic acid.

3.6 Biochemical and biophysical methods

3.6.1 GroEL ATPase assay

A coupled enzymatic assay was used in order to assay ATPase activity of the chaperonin (Kreuzer & Jongeneel, 1983). In this assay, hydrolysis of ATP to ADP is linked to the oxidation of NADH to NAD⁺ by the combined action of Pyruvate kinase (PK) and L-lactate dehydrogenase (LDH):



Since NADH absorbs strongly at 340 nm but NAD⁺ does not, the decrease in absorbance at 340 nm can be followed for 10 min, and converted into ATPase activity, whereby 1 molecule of NADH oxidized to NAD⁺ corresponds to the production of 1 molecule of ADP by GroEL: the change in OD₃₄₀ per minute was determined from the straight part of the kinetics, and converted to μM by dividing by the extinction coefficient at 340 nm for NADH ($3.3 \times 10^{-3} \mu\text{M}^{-1} \text{cm}^{-1}$). This value was divided by the concentration of GroEL in μM to give the ATPase rate per GroEL per minute. Briefly, Phosphoenolpyruvate, PK, LDH and ATP with or without 0.5 μM GroES were incubated in buffer B (50 mM BisTris-NaOH pH 7.5, 100 mM KCl, 10 mM MgCl₂) with final concentrations of 2 mM, 25U/mL, 0.5 mM and 1 mM, respectively. The mixture was incubated for ~3 mins at 25°C to allow for the mopping up of unwanted ADP by PK. The kinetics was started by adding GroEL to a final concentration of ~0.2 μM . Final reaction volume was 50 μL .

3.6.2 *In vitro* MBP refolding assay

Generally, DM-MBP and its cysteine mutants (25 μM) (reduced or oxidized) were denatured in 6 M GuHCl with or without 5 mM DTT and refolded at 25°C upon 100-fold dilution into buffer A (20 mM Tris, pH 7.5, 200 mM KCl, 5 mM Mg(OAc)₂) or into buffer B (20 mM Tris, pH 7.5, 20 mM KCl, 5 mM Mg(OAc)₂) in the absence or presence of chaperonins at the concentrations indicated in the figure legends. Refolding experiments were also carried out at different final GuHCl concentrations, at different TMAO concentrations, and at different temperatures when indicated. Refolding was monitored (295 nm excitation, 345 nm emission) by following the increase in intrinsic Trp fluorescence on a Fluorolog spectrofluorometer (FL3-22, Spex) for DM-MBP concentrations of 10 nM to 1.5 μM (Tang et al., 2006), taking advantage of

the absence of Trp residues in GroEL, SR-EL, and GroES (Martin et al., 1991). Excitation (0.5 or 1 nm) and emission slit widths (5 or 10 nm), as well as shutter speed were adjusted at different protein concentrations to avoid photobleaching. Refolding of DM-MBP cysteine mutants under reducing conditions was performed in the presence of 5 mM DTT. Oxidized cysteine mutants were prepared by removing DTT from the native protein by Micro-Biospin 6 columns (Biorad) (equilibrated with buffer A or B), and oxidizing it for 1 hr at 25°C in the presence of 50 μ M CuCl₂. In experiments where the effect of oxidation was measured after dilution of the reduced protein from denaturant, DTT was first removed from the denatured protein by passing through Micro-Biospin 6 column, equilibrated with degassed and nitrogen saturated denaturing buffer (6 M GuHCl). The denatured protein was then diluted 100-fold into degassed and nitrogen saturated buffer A or B, followed by addition of the oxidizing agent (50 μ M CuCl₂) within 5 s.

3.6.3 Sample preparation for analysis of disulfide-bond formation in the cysteine mutants

To analyze the speed of disulfide bond formation, the charge state distribution of reduced and oxidized DM-MBP cysteine mutants was analyzed by LC-MS. For this purpose, 100 μ M denatured DM-MBP cysteine mutants (in 6 M GuHCl/10 μ M DTT) was diluted 50-fold into buffer B, followed by the addition of CuCl₂ 30 s later, and then inhibition of refolding and disulfide bond formation 5 s later by addition of 6M GuHCl pH 2.5 (final GuHCl 4.8 M). Samples were immediately subjected to LC-MS as described below. Quenching of disulfide bond formation was efficient because no oxidized protein was detected when CuCl₂ was added after 6M GuHCl pH 2.5.

3.6.4 Functional characterization of the cysteine mutants by maltose binding

To analyze the ability to bind maltose, purified DM-MBP and its cysteine mutants (25 μ M) (reduced or oxidized) were either diluted 50-fold into buffer B (native proteins, N) or denatured in 3 M GuHCl and refolded for 90 min upon 50-fold dilution into buffer B (final 500 nM MBP) (refolded proteins, R). The samples were added to 100 μ l amylose beads, equilibrated in buffer B, and incubated for 20 min at 25°C with gentle rocking. After a quick spin on a microfuge, the supernatant was removed and the beads washed with 500 μ l buffer B. The bound protein was eluted with buffer B containing 50 mM maltose. Eluates were analyzed by 12.5% SDS-PAGE and Coomassie staining.

3.6.5 Unfolding and refolding curves of DM-MBP

Refolding experiments were carried out by 50-fold dilution of DM-MBP proteins from buffer A/3M GuHCl into buffer A containing ~60 mM to ~2M GuHCl. Unfolding experiments were carried out by 50-fold dilution of native DM-MBP proteins into GuHCl-containing buffer A. Samples were incubated for 12 hr at 25°C. The fraction of folded protein was determined at 25°C by monitoring intrinsic Trp fluorescence at 345 nm and Far-UV circular dichroism (CD) at 220 nm (Jasco J-715 spectropolarimeter equipped with Peltier-thermostat at 25°C using 0.1 or 0.2 cm cuvettes). The percentages of secondary structure were determined using Contin software (Sreerama & Woody, 2004).

3.6.6 Typtophan fluorescence and Circular dichroism wavelength scans

Trp fluorescence of native, denatured and 0.5 M GuHCl intermediate of DM-MBP (250 nM each) was recorded between 310 and 410 nm on a FL3-22 Spex, using an excitation wavelength of 295 nm. CD wavelength scans of native, denatured and 0.5M GuHCl intermediate of DM-MBP (2 μ M) was measured at 25°C on a Jasco J-715, using 0.1 cm cuvettes, and a scanning speed of 50 nm/min.

3.6.7 Unfolding kinetics by Stopped-flow measurements

The rate of unfolding of WT-MBP, DM-MBP, and the reduced and oxidized DM-MBP cysteine mutants was monitored at various concentrations of GuHCl in buffer A, via stopped-flow mixing experiments using an Applied Photophysics SX.18MV instrument with a 1:1 mixing ratio at 25°C. The dead time of the instrument was \sim 3 ms. The decrease in fluorescence at 345 nm with an excitation at 295 nm was followed. The rate obtained is an average of 10-20 independent measurements. The final protein concentration was 500 nM. Unfolding rates in the absence of denaturant were determined by extrapolation.

3.6.8 Static and Dynamic light scattering

Protein samples (1 μ M) were analyzed using static and dynamic light scattering by auto-injection of the sample into the apparatus at a flow rate of 0.2 ml/min (system buffer: 20 mM Tris-HCl pH 7.5, 200 mM KCl, 5 mM Mg(OAc)₂, 10 mM maltose, 50 mM GuHCl) at 25°C online with DAWN EOS multi-angle light scattering (Wyatt Technology, 690 nm laser) and variable wavelength UV absorbance (set at 280 nm; Agilent 1100 series) detectors (Wyatt,

1993). The protein sample (900 μ l) reaches the detectors 5 min after injection by the autosampler (Agilent 1100 series). DM-MBP (60 μ M) denatured in 3 M GuHCl was diluted 60-fold into buffer (20 mM Tris, pH 7.5, 200 mM KCl, 5 mM Mg(OAc)₂, 10 mM maltose) and injected 1 to 60 min after initiating refolding. Accordingly light scattering begins to be recorded at 6 to 65 min after initiating refolding. UV absorbance showed that the protein concentration was similar for each measurement at the peak of the scattering signal. Native DM-MPB and Rubisco were mixed at various molar ratios with regard to protomer and adjusted to a final total protein concentration of 1 μ M. Protein mass and the hydrodynamic radii were calculated using the ASTRA software (Wyatt Technology) with the dn/dc value for protein set to 0.185 ml/g. Bovine serum albumin was used as the calibration standard.

3.7 Single-molecule FRET experiments

3.7.1 Thiol-mediated labeling of the cysteine constructs

The two-cysteine protein DM-MBP (A52C-P298C) (from now on abbreviated as DM52-298) could be labelled specifically as position 298, because it shows a differential accessibility in presence or absence of maltose. MBP Cys mutants (100 μ M) were labeled in PBS buffer (30 mM PO₄⁻, 150 mM NaCl, pH 7.8) with 500 mM malto-triose (Sigma) for 2 hr at 25°C in the presence of a 1.1 fold molar excess of the fluorophore Atto 532 maleimide, (ATTO-TEC, Inc.). Maltose and unbound fluorophore was removed by dialysing in PBS buffer overnight at 4°C. The second step labeling was carried out at similar conditions in the absence of maltose with 2 fold molar excess of Atto 647N maleimide (ATTO-TEC, Inc.). Any unbound dye was removed by using 10kDa centricons. The coupling efficiency measured by the absorption of MBP ($\epsilon_{280} = 69 \text{ mM}^{-1} \text{ cm}^{-1}$), and Atto 532 ($\epsilon_{532} = 115 \text{ mM}^{-1} \text{ cm}^{-1}$) was more than 90%.

3.7.2 Fluorescence correlation and cross-correlation spectroscopy

For correlation experiments, DM52-298 was fluorescently labeled either with Atto532 maleimide or Atto647N maleimide (ATTO-TEC) as explained above. Fluorescence correlation spectroscopy (FCS) measurements using pulsed interleaved excitation (PIE) (Müller et al., 2005) were performed either on a confocal system based on an inverted microscope (Nikon TE2000) or on a Microtime 200 instrument (Pico-Quant GmbH) to analyze the oligomeric state of DM-MBP during spontaneous refolding. Diffusion times of 10 nM DM52-298, labeled with Atto532 at both positions, through the confocal volume (1 femtoliter) with or without 1 or 2 μ M unlabeled DM-MBP(52-298) were measured by FCS. Atto532 labeled DM52-298 bound to GroEL was used as a control. Measurements were taken at 20°C within the first 800 s after initiating refolding by 100-fold dilution from 3 M GuHCl into buffer A. The FCS curves were fitted using Origin 8.0 (OriginLab, Northampton) with the following model:

$$y = \frac{\gamma}{N} \left(1 + \frac{\tau}{\tau_D}\right)^{-1} \left(1 + \frac{1}{\rho^2} \frac{\tau}{\tau_D}\right)^{-1/2} \times \left(1 + A e^{-(\tau/\tau_A)^\beta}\right) + y_0$$

for freely diffusing particles with a characteristic diffusion time τ_D through the confocal volume characterized by the structure parameter $\rho = W_0/Z_0$, where W_0 and Z_0 are related to the radial and axial dimensions of the confocal spot. N describes the average number of particles in the confocal volume ($V_{\text{eff}} = (\pi/2)^{3/2} W_0^2 Z_0$) and the factor γ corrects for geometrical effects (in the case of a three dimensional Gaussian, $\gamma = 2^{-3/2}$). The stretched exponential with width β compensates for afterpulsing of the detectors. Timescales of afterpulsing τ_A were typically on the order of 1

μs . y_0 corrects for a non-zero baseline due to slow variations during the experiment and was typically on the order of $\pm 10^{-3}$. The diffusion (D) given by

$$D = \frac{w_o^2}{4\tau_D}$$

Fluorescence cross-correlation spectroscopy (FCCS) measurements were performed under refolding conditions as above with a 1:1 mixture of Atto532 labeled DM52-298 and Atto647N labeled DM52-298 (labeled at position 52) at final concentration of ~ 5 nM each. For the Microtime, two pulsed lasers (LDH-P-FA-530 (green), LDH-P-C-640B (red), PicoQuant GmbH) were alternated for excitation. The excitation light was focused on the sample by UPlanSApo 60x/1.20 (OLYMPUS) water immersion objective. The laser power at the objective was $30 \mu\text{W}$ for the FCS and FCCS experiments. The fluorescence emission was focused on the confocal pinhole (diameter $50 \mu\text{m}$) and subsequently split by a dichroic mirror (600 dcxr, Semrock). Scattered laser light was blocked by using emission filters. The fluorescent light was detected by two avalanche photodiodes (Micro Photon Devices srl) coupled to a time-correlated single-photon counting device (HydraHarp 400, PicoQuant GmbH) operated in time-tagged time resolved mode (TTTR), which was synchronized to the same 40 MHz master clock used for driving the alternated laser pulses. The cross-correlation functions were calculated by correlating the photons produced by the green excitation laser and detected in the green channel with the photons produced by the red excitation laser and detected in the red channel. In this way, spectral crosstalk can be removed, making FCCS much more sensitive to weak interactions. A double-labeled DNA oligomer (40 base pairs), labeled with Atto532 in one strand and Atto647 in the

other at a 22 base-pair separation (IBA GmbH, Germany), was used as a positive control for cross-correlation and the free dyes were measured as a negative control.

3.7.3 Single-pair FRET

Single-pair FRET (SpFRET) measurements were performed on either the home-built system or the Microtime 200 using PIE (Müller et al., 2005; Sharma et al., 2008). The concentration of double-labeled DM52-298 in the sample was diluted to ~100 pM to ensure that the probability of having more than one particle in the probe volume at the same time is small. Hence, for each particle detected, the probability of a second particle being in the volume simultaneously is < 1%. For each experiment, at least 500 particles were measured and the experiments were repeated with different protein preparations to verify the reproducibility of the results.

3.8 Hydrogen/deuterium exchange experiments

3.8.1 Pulse labeling

MBP protein (100 μ M) was unfolded in 3 M GuHCl. Samples were diluted 50- fold and incubated for at least 12 hr at final GuHCl concentrations of 60 mM to 3 M in buffer B. Pulse H/D exchange was performed by diluting MBP protein (2 pmol/ μ l; 50 μ l) 1:10 into deuterated 20 mM Tris DCl, 20 mM KCl, 5 mM Mg(OAc)₂, D₂O, pH 7.5 (pH meter reading uncorrected for isotope effect; Connelly et al., 1993) at 25°C. MBP samples containing GuHCl were labeled with buffer containing an equivalent concentration of GuDCl. The exchange reaction was quenched after 10 s to pH ~2.5 by addition of 200 μ l ice-cold 600 mM sodium phosphate buffer, pH 2.4. Quenched samples (700 μ l) were immediately concentrated to 100 μ l using Vivaspin spin

columns at $\sim 0^{\circ}\text{C}$ for 5 min. Samples were flash frozen in liquid nitrogen and stored at -80°C and analyzed by MS within 2 days as described in the following section.

3.8.2 Mass spectrometric analysis

Frozen MBP samples (75 μl) were thawed rapidly within 1 min and immediately injected into a custom built cooled microflow HPLC system (designed by John. R. Engen, Barnett Institute, Boston) at a flow rate of 60 $\mu\text{l}/\text{min}$. Proteins were desalted for 2 min on a 1 mm x 50 mm Vydac C-4 column and directly eluted into the mass spectrometer with an 8 min gradient of 40%-75% acetonitrile. The mobile phase contained 0.1% formic acid. The C-4 analytical column, as well as the injection and switching valves were maintained at $\sim 0^{\circ}\text{C}$ by placing them in a cooled housing. The mobile phases were precooled through stainless steel loops within a thermoelectric cooling device. Mass spectral analyses were carried out on a Waters Synapt HDMS ESI-QToF mass spectrometer. Capillary voltage was set to 3.2 kV. The ESI source and desolvation temperatures were 50°C and 175°C , respectively, with a desolvation gas flow of 600 l/hr and a cone gas flow of 50 l/hr. Mass spectra were acquired using a 1 s scan time. All QToF data were collected in ESI (+) and V mode. Protein mass spectra were corrected online using a solution of myoglobin (4 pmol/ μl) as lock mass. Masses were calculated by deconvoluting multiple charge states of combined protein spectra using MassLynx software and the MaxEnt1 algorithm (Wales & Engen, 2006).

4. Results

4.1 The spontaneous refolding of DM-MBP is not slowed down by transient aggregation

In order to address the mechanism by which GroEL accelerates the folding of substrate proteins, we have chosen the model substrate DM-MBP. DM-MBP does not experience irreversible aggregation as most GroEL substrates do, since its refolding occurs with ~90% yield, and experiences ~10 fold acceleration in the presence of EL/ES system. It has been suggested that the prevention of transient aggregates which would form during the spontaneous refolding of DM-MBP may be the cause of this rate acceleration.

4.1.1 A model of aggregation-limited DM-MBP refolding is simulated

We have used Berkeley Madonna to simulate a model in which the spontaneous refolding of DM-MBP would be limited by the formation of the smallest reversible aggregate, namely a dimer, as shown in figure 12. By setting k_f value to the rate of folding inside GroEL ($\sim 0.004\text{s}^{-1}$), and varying the k_a and k_d values, we simulated the various possible $t_{1/2}$ values of the refolding reaction with 10 nM initial concentration of DM-MBP. Notably, it was the K_D value (ratio of k_d/k_a) which was determining as to the $t_{1/2}$ values. We found that the experimentally observed $t_{1/2}$ value of DM-MBP spontaneous refolding (~ 1000 sec, under low salt conditions) was obtained in the simulation only when using a K_D value of 2.5 nM for dimer formation (or aggregation). Accordingly, we simulated the folding rate of DM-MBP at various concentrations, and noted that it was decreasing with increasing concentrations. Another prediction of our model was that a

dimeric state is populated for about 200 sec upon initiation of the refolding to about 80%. In the following sections, we test the predictions of this model by performing DM-MBP refolding at increasing concentrations, and by monitoring dimeric or or larger aggregate species by FCS, FCCS and light scattering.

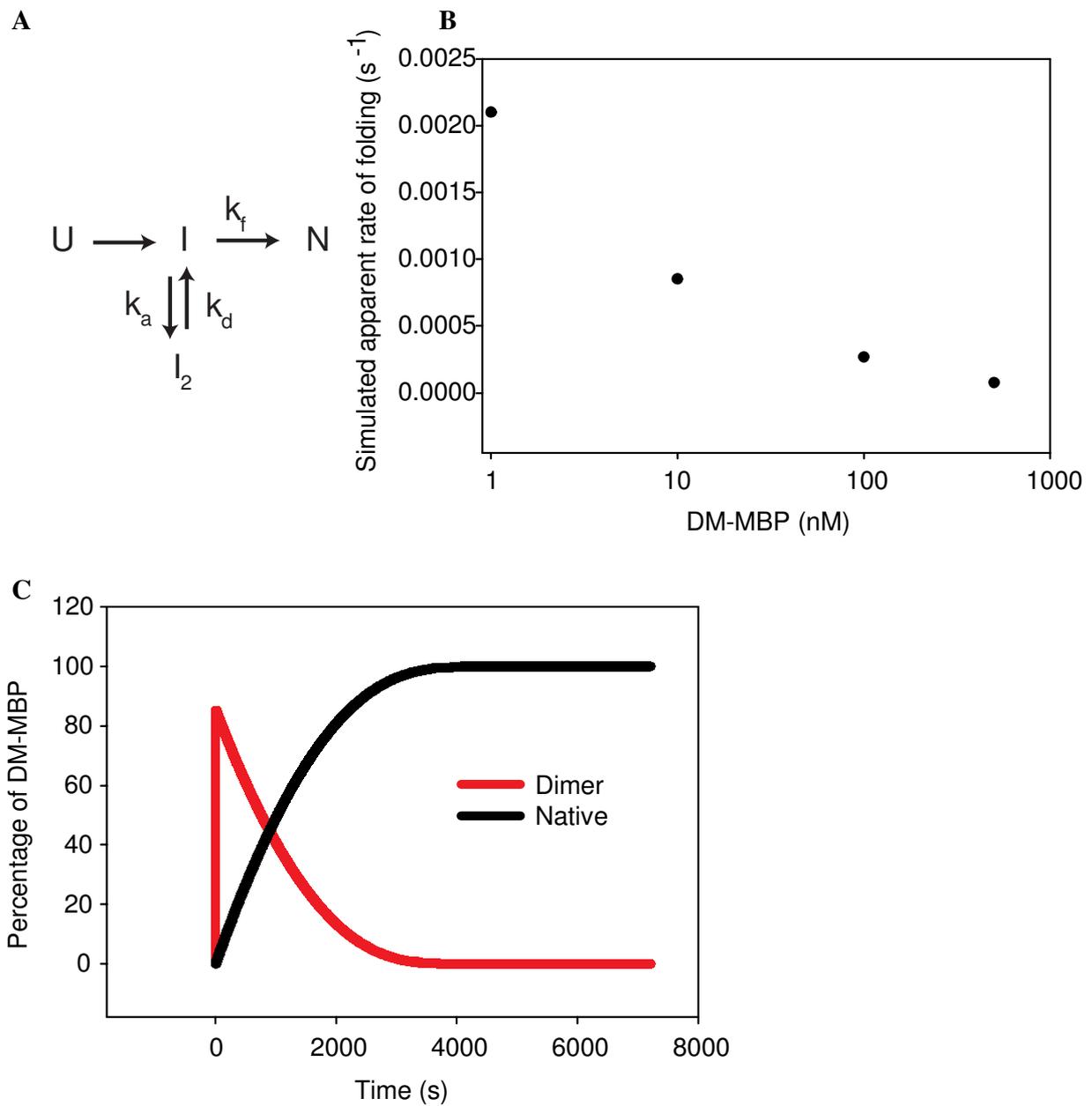


Figure 12: Model for reversible aggregation-limited refolding of DM-MBP and predictions. A) Simplified scheme using a DM-MBP dimer as the smallest possible reversible aggregate. B) Using Berkeley Madonna, we simulated the model in A), using k_f value of $0.004s^{-1}$, a k_d/k_a value of $2.5nM$, and starting with $10 nM$ DM-MBP. The model predicted a decreasing apparent rate of refolding with increasing DM-MBP. C) Prediction of $\sim 75\%$ DM-MBP in the dimeric state in the first ~ 250 sec of refolding.

4.1.2 The model of aggregation-limited kinetics is inconsistent with the actual refolding experiments

We monitored the refolding of DM-MBP by tryptophan fluorescence, over a wide concentration range (from 10 to 1500 nM; Fig. 13), and found that the apparent rate constant did not decrease with increasing concentrations, which is inconsistent with the above model. In contrast, the rate of refolding was independent of the concentration at which the refolding was carried out. These experiments provide a compelling argument against the notion that the spontaneous refolding of DM-MBP is limited by reversible aggregation. Additionally, the protein experienced rate acceleration in the presence of the EL/ES system, implying that this cannot be merely due to aggregation prevention.

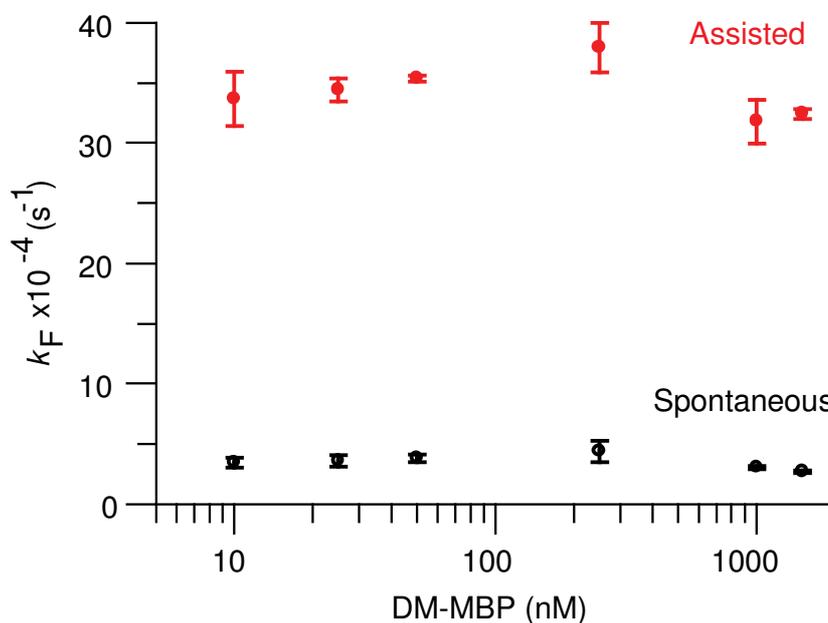


Figure 13: Rates of spontaneous and GroEL/ES-mediated refolding of DM-MBP at different DM-MBP concentrations (10 nM-1.5 μM). DM-MBP was denatured in 6 M GuHCl and diluted 100-fold to 60mM GuHCl into buffer A (spontaneous) or buffer containing 0.5-4 μM GroEL and 1-8 μM GroES (assisted) at 25°C. Assisted

refolding was initiated by the addition of 5 mM ATP. Refolding was monitored by Trp fluorescence. Standard deviations from three independent measurements.

4.1.3 FCS and FCCS support the absence of any transient aggregation during DM-MBP refolding

In order to assay aggregate formation more directly, we analyzed DM52-298 states by FCS and FCCS. The rate of spontaneous folding of the labeled protein was similar to unlabeled DM-MBP and was accelerated ~6-fold by GroEL/ES (Sharma et al., 2008). To detect higher-order aggregates, FCS measurements were performed with DM52-298, labeled with Atto532, and the decay of the autocorrelation function as a measure of the average diffusion time of particles through the probe volume was monitored. The diffusion coefficient of refolding DM-MBP (10 nM) ($\sim 54 \mu\text{m}^2/\text{s}$) (Fig. 14), measured during the first 800 s upon dilution from denaturant, was similar to that of native DM-MBP ($\sim 58 \mu\text{m}^2/\text{s}$). Importantly, the diffusion rate of the refolding, labeled DM-MBP remained unchanged in the presence of excess (1 or 2 μM) unlabeled, refolding DM-MBP, excluding the formation of large aggregates (Fig 14). GroEL-bound DM-MBP (~ 800 kDa) was used as a control.

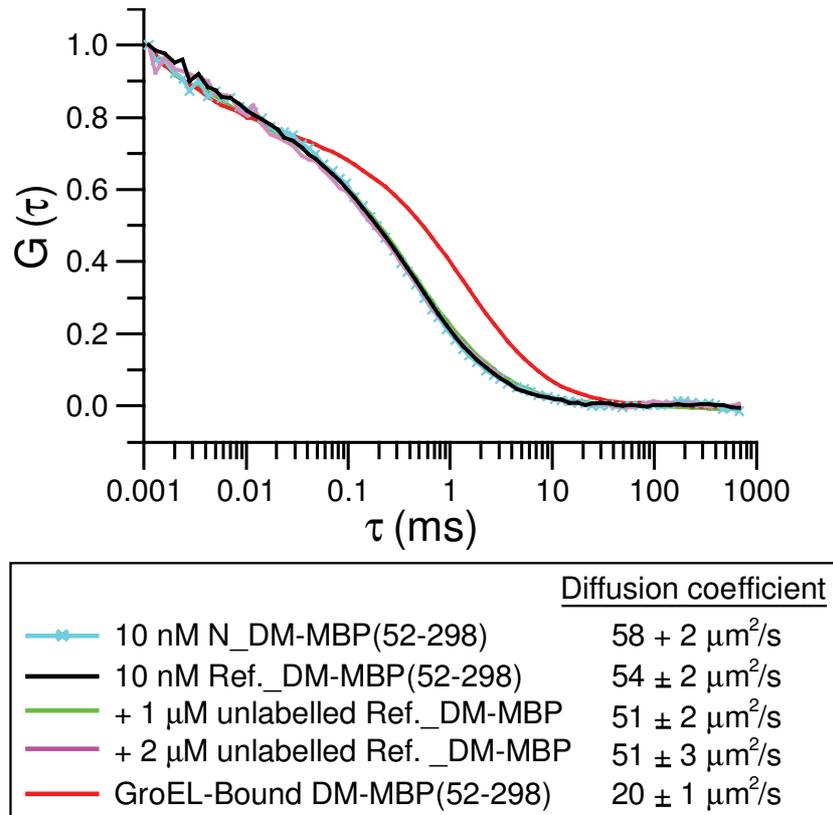


Figure 14: Fluorescence correlation spectroscopy (FCS) to measure the diffusion rates of spontaneously refolding DM-MBP. Normalized fluorescence autocorrelation amplitudes $G(\tau)$ are shown. Diffusion times were measured at 20°C during the first 800 s of refolding with 10 nM Atto532-labeled DM-MBP(52-298) in the absence or presence of 1 or 2 μM unlabeled, refolding DM-MBP(52-298) (final GuHCl 30 μM in buffer A). Native and GroEL-bound labeled DM-MBP(52-298) were used as controls. Diffusion coefficients are given as averages \pm standard deviation (SD) from three independent experiments (performed by Guoxin Jiang).

To test whether small aggregates, including dimers, were formed during refolding, we performed FCCS experiments. DM-MBP labeled with either Atto532 or Atto647N was unfolded as a 1:1 mixture, and refolding initiated at a final concentration of ~ 10 nM. Assuming aggregation-limited folding kinetics (see model in Fig. 12), aggregates would be expected to be populated to $\sim 75\%$ during the first 250 s of refolding, corresponding to $\sim 35\%$ of particles containing both labels in case of exclusive dimer formation. However, dimeric or multimeric

species were undetectable during refolding, based on the absence of a cross-correlation signal (Fig. 15). A mixture of the free dyes served as a negative control and a double-labeled DNA sample as a positive control. These experiments clearly demonstrate the absence of reversible aggregation during the spontaneous refolding of DM-MBP.

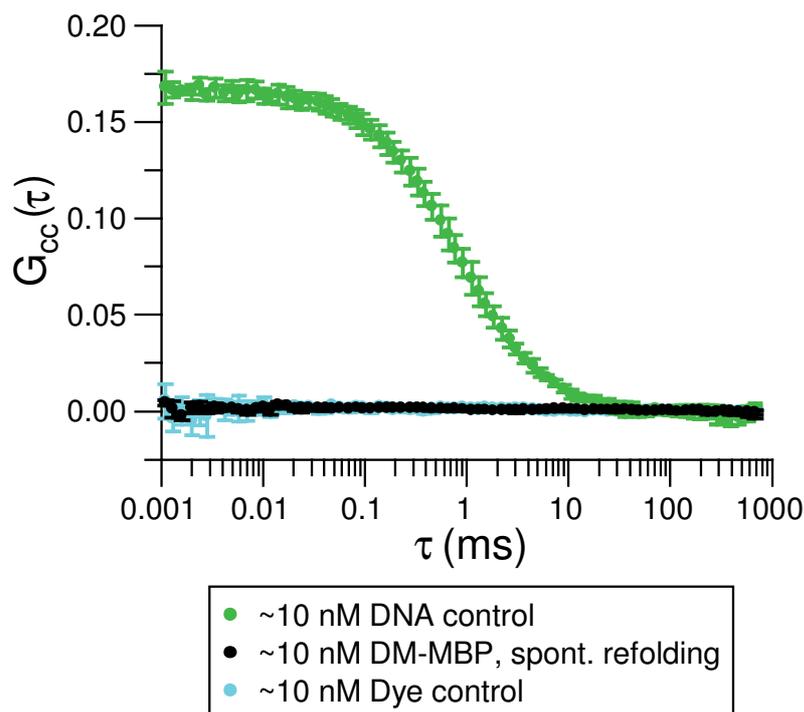


Figure 15: Fluorescence cross-correlation spectroscopy (FCCS) of a 1:1 mixture of DM-MBP(52-298) labeled at position 52 with Atto532 or Atto647N. Labeled DM-MBP molecules were denatured and diluted as in figure 3 to a final concentration of ~5 nM each. FCCS was recorded with pulsed interleaved excitation (PIE) (Müller et al., 2005) within 120 s of initiating refolding. Approximately 10 nM of DNA (40 base pair) labeled with Atto532 and Atto647 spaced 22 base pairs apart was used as a positive control and the same fluorophores freely diffusing in solution served as a negative control (performed by Guoxin Jiang).

4.1.4 Static and dynamic light scattering further support the absence of any aggregation during DM-MBP refolding

We have also assayed the presence of aggregates by measuring light scattering at a 90° angle. The molar mass and the hydrodynamic radius of native DM-MBP, dimeric Rubisco and various molar ratios of the two proteins were extracted from those measurements and are as indicated in figure 16. Thus, ~15% DM-MBP dimers (~80 kDa) as the smallest possible aggregate would have resulted in a detectable increase in light scattering. When collecting light scattering measurements on refolding DM-MBP at various times during the refolding, we clearly see that there is no increase in the amount of scattered light, again arguing against the presence of aggregates.

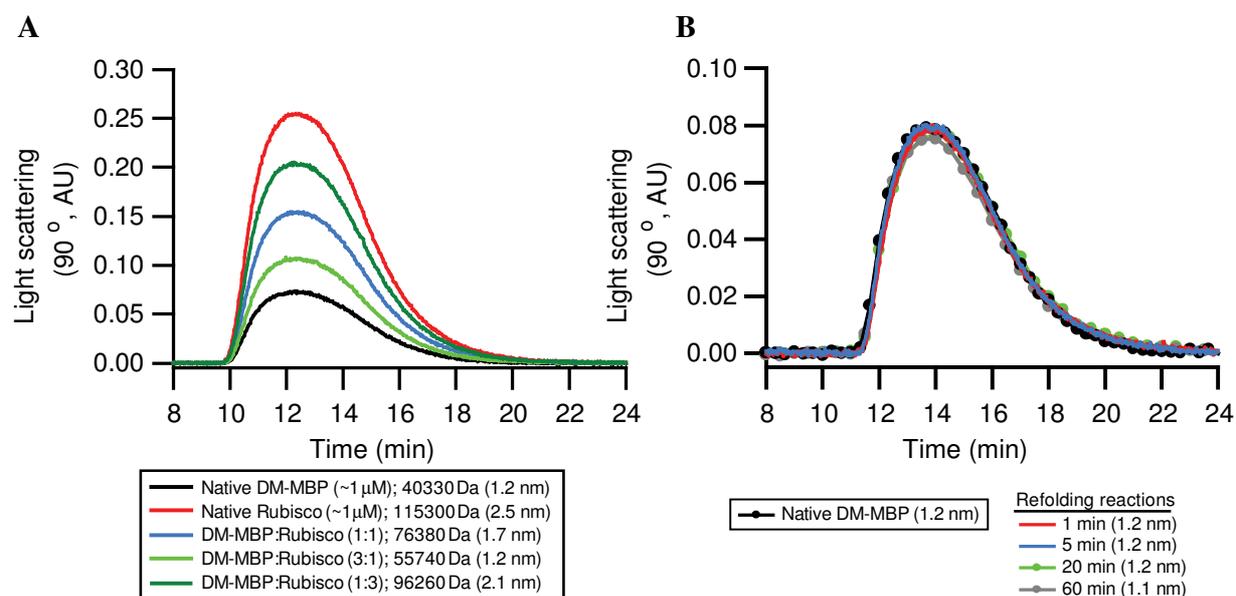


Figure 16: Light scattering analysis of DM-MBP. A) Light scattering measurements collected at an angle of 90° are shown. Molar mass and the hydrodynamic radius of native DM-MBP, dimeric Rubisco and various molar ratios of the two proteins (protomer concentrations) are indicated. Final total concentrations are adjusted to ~1 μM. Using dimeric bacterial Rubisco (~55 kDa monomer) as a control protein, the presence of ~15% DM-MBP dimers (~80

kDa) as the smallest possible aggregate would have resulted in a detectable increase in light scattering. B) Light scattering measurements collected at an angle of 90° for native DM-MBP ($\sim 1 \mu\text{M}$) and of DM-MBP spontaneously refolding at various times after the initiation of the refolding reaction are shown. Denatured DM-MBP in 3MGuHCl was diluted 60-fold to a final concentration of $\sim 1 \mu\text{M}$. Numbers in parenthesis refer to the hydrodynamic radii measured for the respective samples. The light scattering signal, the hydrodynamic radius and molar mass of the refolding DM-MBP remained equivalent to that of the native protein (performed by Manajit Hayer-Hartl).

Having excluded aggregation as a cause of slow spontaneous DM-MBP refolding, it became clear that the rate acceleration experienced in the presence of the EL/ES system was not due to aggregation prevention. In the next section of the results, we investigate the cause of the slow folding of DM-MBP.

4.2 Spontaneous refolding of DM-MBP is limited by the formation of a kinetically-trapped intermediate

Since aggregation was not playing any role in slowing down the refolding process, it seemed likely that a trapped intermediate was being formed during the refolding. We defined and characterized this intermediate by various methods, as described below.

4.2.1 Denaturation-renaturation curves by Trp fluorescence and CD unveil the kinetically-trapped intermediate

To define the intermediate species, we first recorded the denaturation-renaturation curves by Trp and CD measurements after 12 hr of incubation in varying concentrations of denaturant. A prominent hysteresis effect was detected with an intermediate state being populated at 0.5-0.8 M GuHCl during the refolding phase, although the native state was thermodynamically stable up to

~0.7M GuHCl (Fig. 17). Thus, this intermediate is kinetically trapped during refolding at intermediate denaturant concentration. The hysteresis effect was also observed with urea as denaturant in the absence of chloride salt (Fig. 18). Next, we took spectroscopic measurements immediately upon dilution of the unfolded DM-MBP protein into refolding conditions, and we characterized the transient refolding intermediate by extrapolating the signals of manual-mixing kinetic refolding reactions at various denaturant concentrations to time-point zero. The different zero-time point values at different denaturant concentrations were then plotted and compared to the values of the kinetically trapped-intermediate. This analysis suggested that the intermediate from which the refolding reaction starts has similar, if not identical properties to the kinetically-trapped intermediate at 0.5-0.8M GuHCl (Fig. 17).

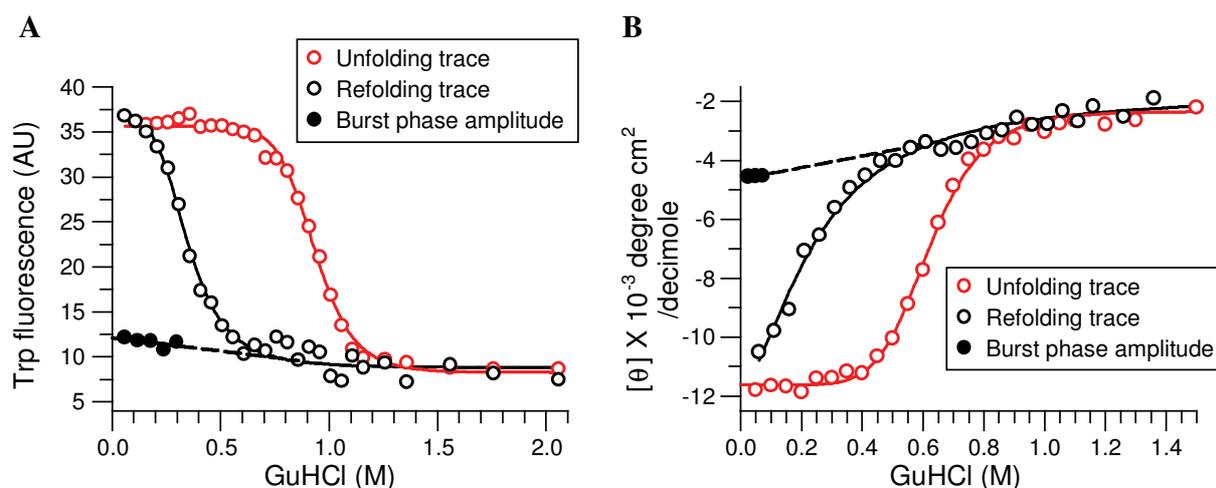


Figure 17: Detection of a kinetically trapped refolding intermediate of DM-MBP. GuHCl-dependent unfolding and refolding of DM-MBP (1 μM) was monitored by Trp fluorescence at 345 nm (A) and circular dichroism at 220 nm (B). Unfolding trace: native DM-MBP was incubated for 12 hr in buffer A containing ~60 mM to ~2 M GuHCl at 25°C. Refolding trace: DM-MBP (50 μM) was unfolded in 3 M GuHCl and then diluted 50-fold into buffer A containing different concentrations of GuHCl, followed by incubation for 12 hr at 25°C. Burst phase amplitudes were determined immediately on dilution from 3MGuHCl. Representative data from at least two independent experiments.

Interestingly, the hysteresis was less pronounced with a single mutant of MBP containing only the Y283D substitution (SM-MBP), while no hysteresis was observed with wild-type MBP (WT-MBP) (Fig. 18), consistent with the fact that the relative order of folding rates of these proteins is DM-MBP < SM-MBP < WT-MBP (Tang et al., 2006). The same results held true when performing the same experiment using urea as a denaturant.

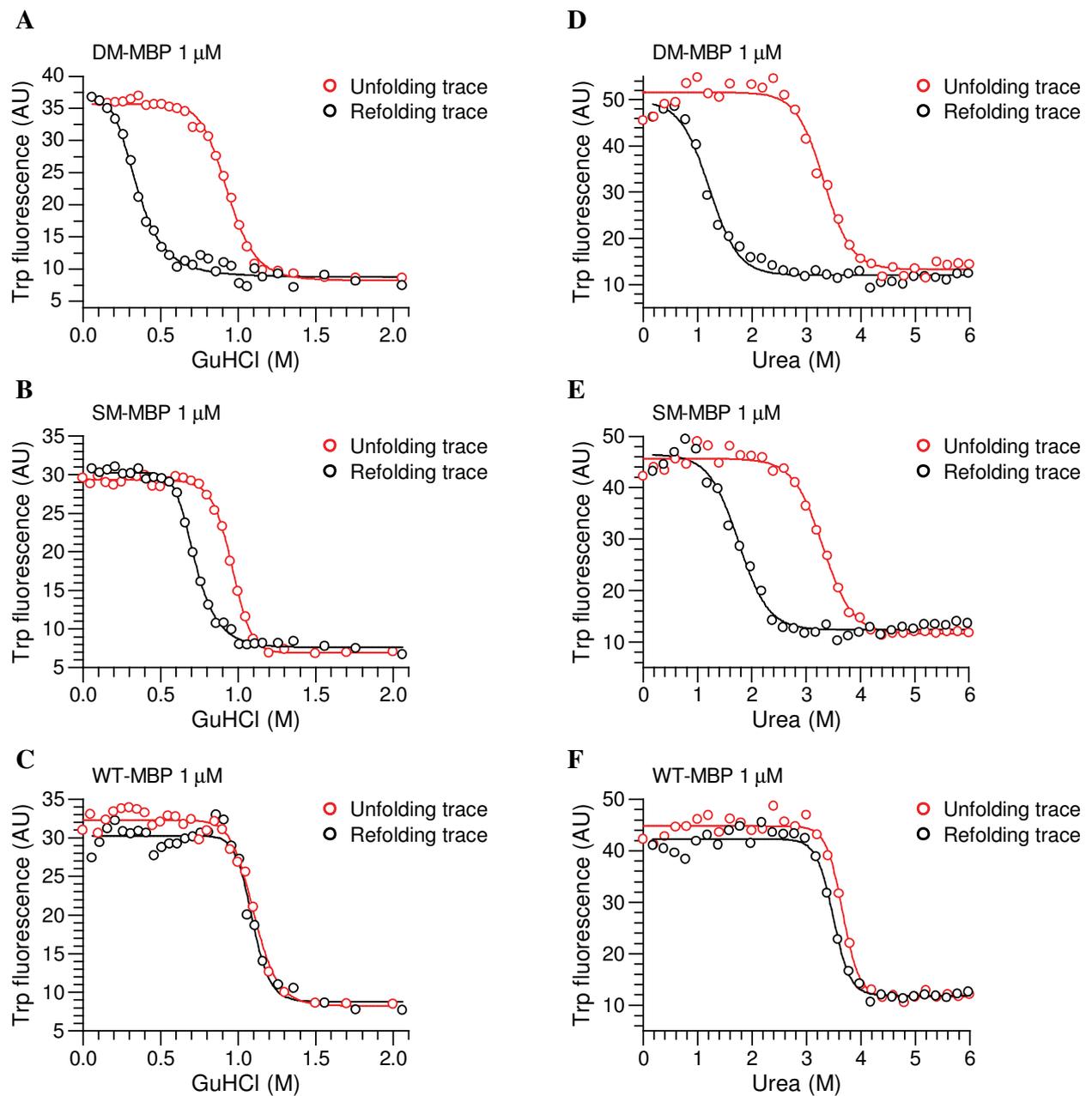


Figure 18: Detection of hysteresis with SM-MBP, but not WT-MBP. (A-C) GuHCl-dependent unfolding and refolding of MBP. Unfolding and refolding of DM-MBP, SM-MBP and WT-MBP (1 μ M each) was monitored by Trp fluorescence with excitation at 295 nm and emission at 345 nm. Unfolding trace: native MBP was incubated for 12 hr with increasing concentrations of GuHCl in buffer A at 25°C. Refolding trace: MBP (100 μ M) was unfolded in 6 M GuHCl and then diluted 100-fold into buffer A containing increasing concentrations of GuHCl, followed by

incubation for 12 hr at 25°C. Representative measurements from at least two independent experiments. (D-F) Urea-dependent unfolding and refolding of MBP. Same procedure as with GuHCl.

4.2.2 The kinetically-trapped intermediate is collapsed, but lacks ordered structure

Single-pair FRET (spFRET) measurements in solution were next performed to analyze the compactness of the intermediate at 0.5 M GuHCl. DM52-298 was double-labeled with Atto532 (position 52) as the fluorescence donor and Atto647N (position 298) as the acceptor (Sharma et al., 2008). These two positions are ~ 33 Å apart in the native structure. The fully unfolded protein in 3 M GuHCl and the native protein showed narrow distributions of FRET efficiencies (f_E) with peak values of ~ 0.08 and ~ 0.84 , respectively (Fig. 19) (Sharma et al., 2008). In comparison, the kinetically trapped refolding intermediate displayed a broader distribution centering on a f_E of ~ 0.80 , suggesting that this species has an average compaction similar to that of the native protein but with a greater variability in structure, as demonstrated by the broader intramolecular distance distribution.

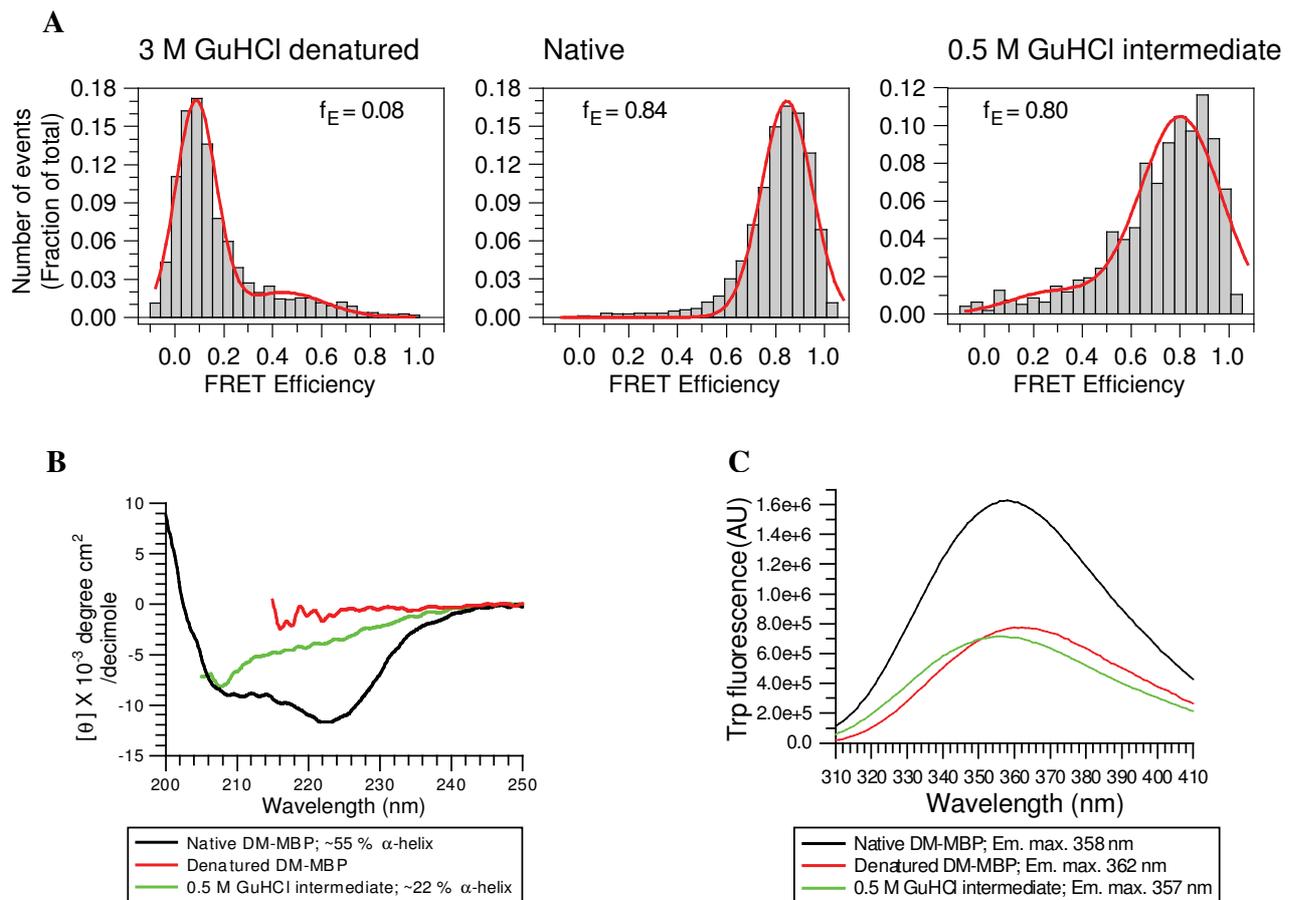


Figure 19: Characterization of the kinetically trapped DM-MBP intermediate. A) SpFRET analysis in solution of DM-MBP(52-298) double-labeled with Atto532 and Atto647N. Denatured protein in 3 M GuHCl, native protein, and kinetically trapped folding intermediate in 0.5 M GuHCl are shown. Final protein concentration ~100 pM in buffer A containing GuHCl as indicated. Peak values of a Gaussian fit to the FRET efficiency distributions (f_E) are indicated. The shoulder of the f_E peaks of the denatured and the intermediate states were fitted with a second Gaussian to allow the correct determination of the peak f_E . Representative histograms from at least three independent measurements are shown. (B) Circular dichroism wavelength scans. CD wavelength scans of native, denatured and 0.5 M GuHCl intermediate of DM-MBP (2 μ M) measured at 25°C with 0.1 cm cuvettes. (C) Trp fluorescence wavelength scans. Trp fluorescence of native, denatured and 0.5 M GuHCl intermediate of DM-MBP (250 nM each) at 25°C. Emission maxima are indicated.

The kinetically trapped refolding intermediate of DM-MBP at 0.5 M GuHCl had molten globule-like properties. It contained only ~22% α -helical structure by CD and had a Trp fluorescence intensity as low as the unfolded protein, suggesting the absence of ordered tertiary structure (Fig. 19B-C). Pulsed hydrogen-deuterium (H/D) exchange measurements, coupled to mass spectrometry (MS), were next performed to further characterize this conformational ensemble (Fig. 20). The protonated and fully deuterated proteins were used as reference. Incorporation of deuterium by the kinetically trapped folding intermediate, generated by dilution from 3 M to 0.5 M GuHCl and incubation for 12 hr, corresponded to ~310 exchangeable amides (when corrected for ~10% back exchange and extrapolated to 100% D₂O during exchange) (Fig. 20B). This was indistinguishable from the exchange properties of the fully denatured protein in 3 M GuHCl (~310 exchangeable amides) (Fig. 20B), demonstrating the absence of stable secondary structure in the kinetically trapped intermediate. In contrast, WT-MBP under the same conditions displayed a structural stability similar to that of the native state at 60 mM GuHCl (~75 rapidly exchangeable amides) (Fig. 20C), consistent with the absence of hysteresis between the unfolding and refolding curves (Fig. 18). Furthermore, pulsed H/D exchange of DM-MBP under refolding conditions (60 mM GuHCl) (Fig. 20D) showed that an intermediate with exchange properties comparable to the kinetically trapped state at 0.5M GuHCl was populated for more than 5 min during refolding, converting slowly to the native state (Fig. 20E and 20B). A similar intermediate state was only transiently populated by WT-MBP (Fig. 20F). Thus, the H/D exchange measurements together with the spectroscopic analysis demonstrate that DM-MBP populates under refolding conditions a kinetically trapped intermediate that is collapsed but lacks ordered structure. Structure formation within this dynamic intermediate appears to be rate limiting for folding, suggesting the existence of a significant entropic folding barrier.

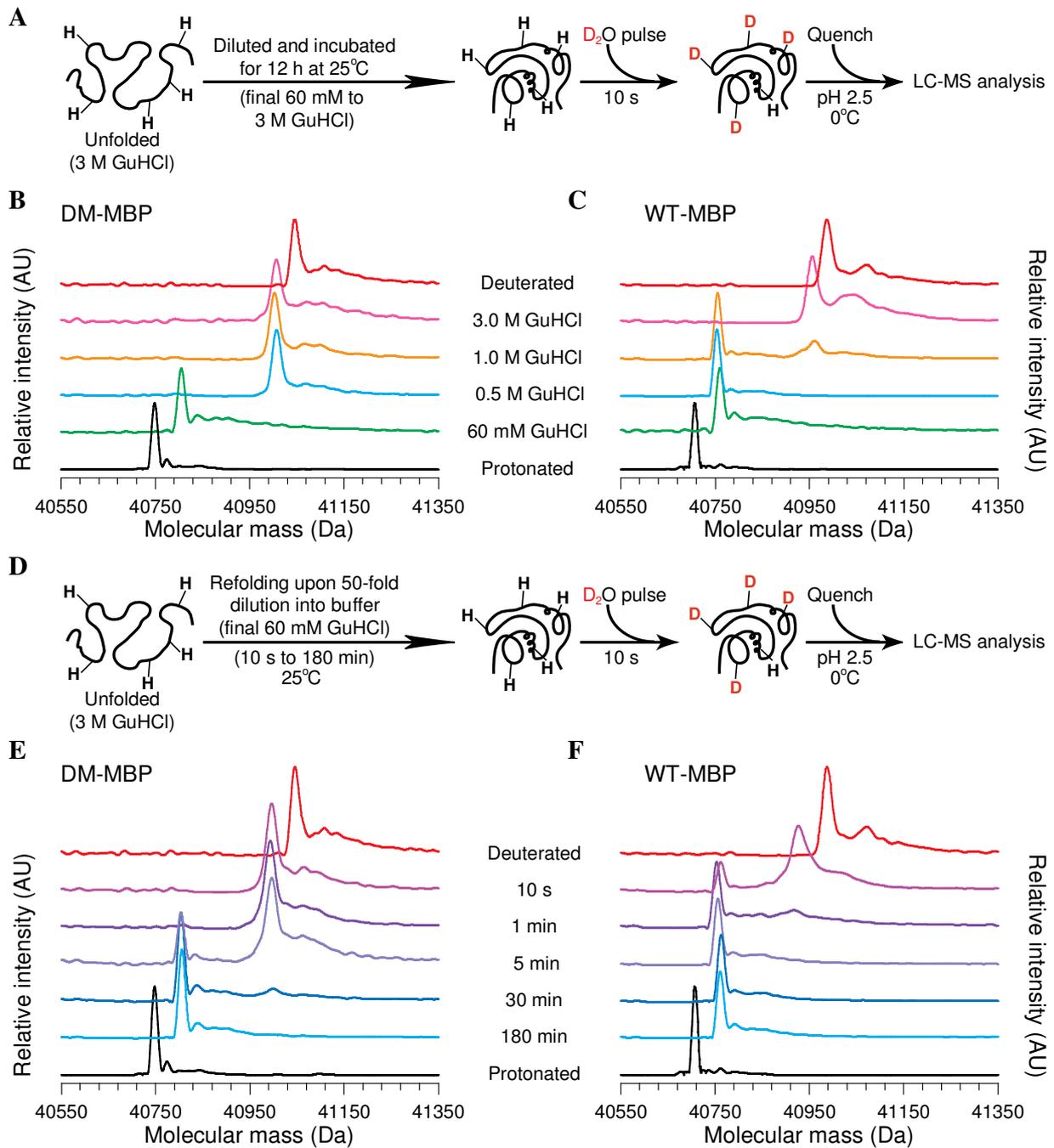


Figure 20: Dynamic Nature of Kinetically Trapped Refolding Intermediate Characterized by H/D Exchange.

(A-C) Pulsed H/D exchange after incubation in different denaturant concentrations. Schematic representation of the experiment (A), deconvoluted mass spectra of DMMBP (B), and WT-MBP (C) are shown. Global H/D exchange patterns as a function of denaturant monitored by ESI-QToF MS are shown. Proteins were diluted from 3 M GuHCl

into buffer B with the final GuHCl concentrations indicated. After incubation for at least 12 hr, samples were subjected to a 10 s deuterium pulse. The native protonated and deuterated proteins are shown as reference. (D-F) Pulsed H/D exchange under refolding conditions. Schematic representation of the experiment (D), deconvoluted mass spectra of DM-MBP (E), and WT-MBP (F) are shown. Global H/D exchange pattern as a function of refolding time is indicated. Refolding was initiated by dilution of denatured protein to a final concentration of 60 mM GuHCl in buffer B (performed by Bernhard Poschner).

4.3 Disulfide-mediated constraints accelerate spontaneous refolding by decreasing the entropic barrier to folding

Long-range disulfide bonds may serve to configurationally constrain flexible regions in folding intermediates, thereby entropically destabilizing them relative to the transition state and accelerating folding (Camacho & Thirumalai, 1995; Mamathambika & Bardwell, 2008) (Fig. 21A). So in order to probe the entropic component of the folding barrier of DM-MBP, we engineered double-cysteine mutants with the cysteines having appropriate positions and orientations for disulfide bond formation in the native state, while being far apart in the amino acid sequence (Fig. 21B).

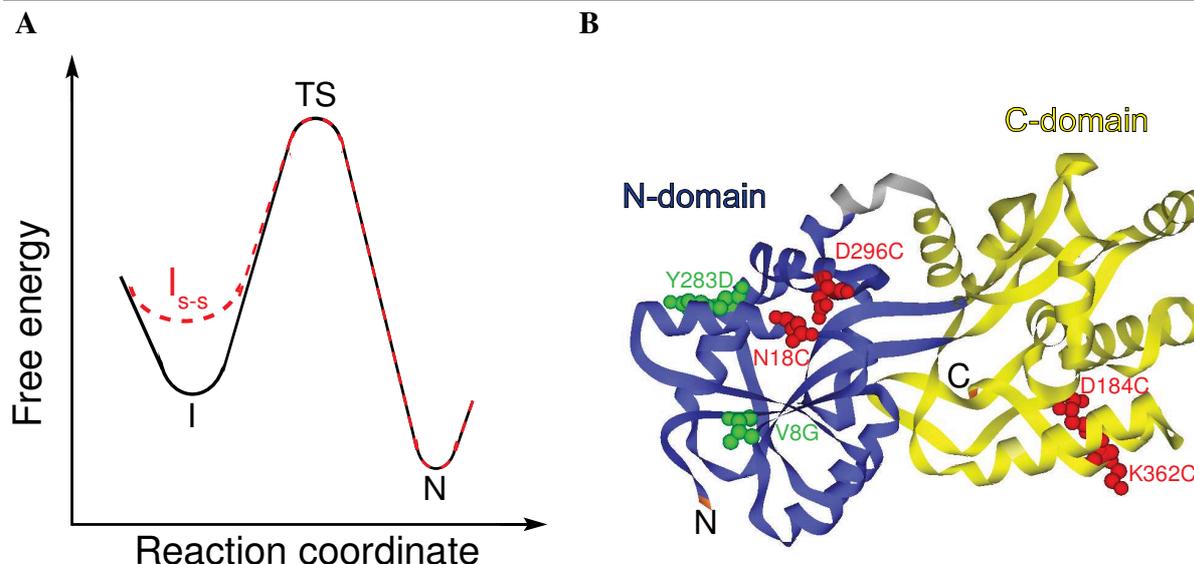


Figure 21: Engineered long-range disulfides in DM-MBP predicted to destabilize the trapped intermediate.

(A) Free energy diagram illustrating the entropic destabilization of the kinetically trapped refolding intermediate (I) by disulfide-mediated structural constraints (I-s) relative to the folding transition state (TS), resulting in accelerated conversion to the native state (N). (B) Ribbon diagram of the structure of MBP, indicating the positions of DM-MBP mutations V8G and Y283D (green) and the cysteine pair mutations N18C-D296C and D184C-K362C (red). The two discontinuous domains are shown in blue (N-domain) and yellow (C-domain).

4.3.1 The oxidized mutants fold faster than their reduced counterparts

We were able to distinguish between the reduced and oxidized states of the mutants by measuring charge state distribution by LC-MS (Zhang et al., 2001). Accordingly, both the N- and C-domain mutants, DM-MBP (18C-296C) and DM-MBP (184C-362C) (Fig. 22A-B) were found to readily form the disulfide bond upon oxidation with CuCl_2 . We next determined whether the oxidized mutants retained their native state, and bound maltose. Indeed, both mutants bound maltose as efficiently as DM-MBP in the native state (Fig. 22C).

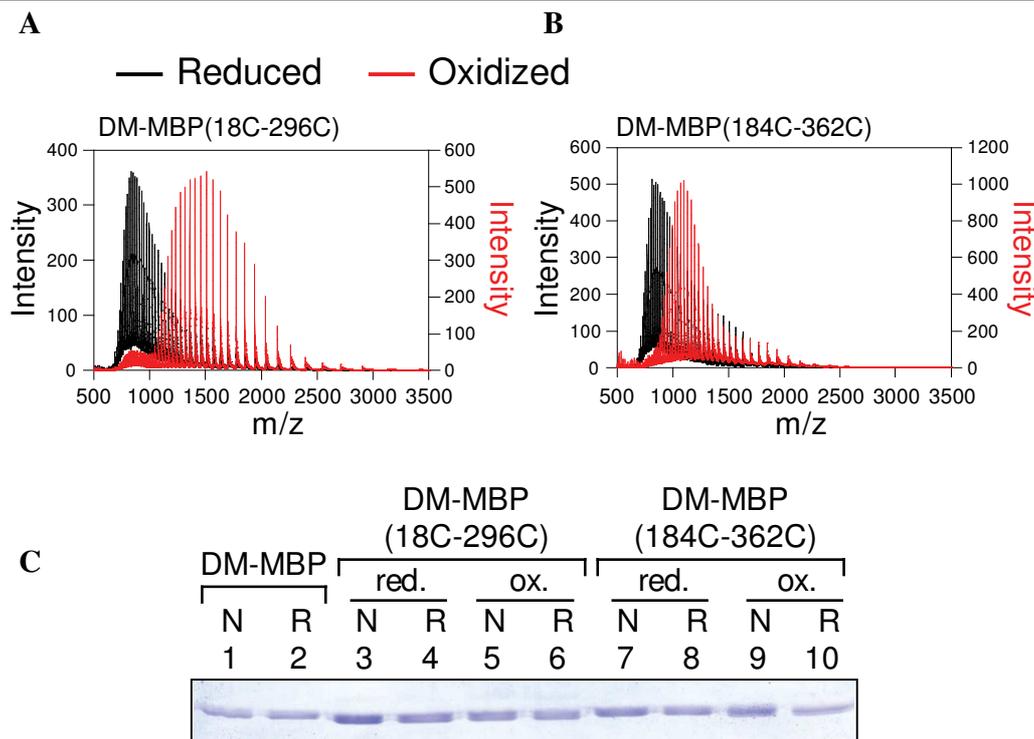


Figure 22: Characterization of Cysteine Pair Mutants of DM-MBP. (A and B) Charge state distribution measured by LC-MS. DM-MBP(18C-296C) and DM-MBP(184C-362C), were analyzed under reducing and oxidizing conditions on a Waters Synapt HDMS ESI-QToF mass spectrometer. The distribution of the charge state centers distinguishes between the reduced (black) and disulfide-bonded proteins (red) (Zhang et al., 2001) (Performed by Bernhard Poschner). (C) Affinity of maltose binding. Native (N) and spontaneously refolded (R) DM-MBP, as well as reduced (red.) and oxidized (ox.) DM-MBP(18C-296C) and DM-MBP(184C-362C) (final 500 nM) were incubated with amylose beads for 20 min at 25°C. Bound protein was eluted by addition of buffer B/ 50 mM maltose and samples were analyzed by 12% SDS-PAGE and Coomassie staining.

Next, we carried out the refoldings of these proteins in buffer A, both in the reduced and oxidized states, and found that the oxidized proteins refolded ~5-fold faster compared to the reduced counterparts. Since DM-MBP undergoes collapse within milliseconds of initiating its refolding (Sharma et al., 2008), we wanted to determine whether it was necessary to constrain the unfolded state versus the collapsed state to get the acceleration effect. Towards this end, we started the refolding using the reduced protein, and added CuCl_2 within 5 sec after having started

the refolding. We obtained the same refolding rates as when starting with the oxidized unfolded protein, so constraining the unfolded state is not necessary for the rate acceleration (Figs 23A-B). Thus, the disulfides could be forming either in the collapsed intermediate or in the transition state. In an effort to discriminate between the 2 possibilities, we determined the speed of disulfide bond formation, namely by diluting the unfolded reduced proteins either into buffer B or buffer B/GuHCl for 30 s, subjecting it to CuCl_2 for 5 s, and quenching it using 6 M GuHCl/pH 2.5 (Fig. 23C-E). By measuring the charge state distribution by LC-MS, it was clear that the proteins could not form disulfides when diluted into 6 M GuHCl, but did get oxidized with the short 5-second pulse of CuCl_2 . In other words, the disulfides formed within seconds of adding CuCl_2 to the collapsed protein and at a much faster rate than the refolding reaction; the disulfides thus form in the collapsed intermediate prior to the transiently populated transition state. Notably, the presence of the bridge was necessary for rate acceleration, as when the refolding was initiated from the oxidized protein and DTT was added within a few seconds, the refolding reverted to the slow rate of the reduced protein (Fig. 23A-B). We next attempted to determine whether the cause of the observed rate acceleration was indeed a decrease in the hypothesized entropic barrier to folding. Generally, the rate acceleration could either be attributed to lowering the energy barrier between the native and transition states, or between the intermediate and transition state.

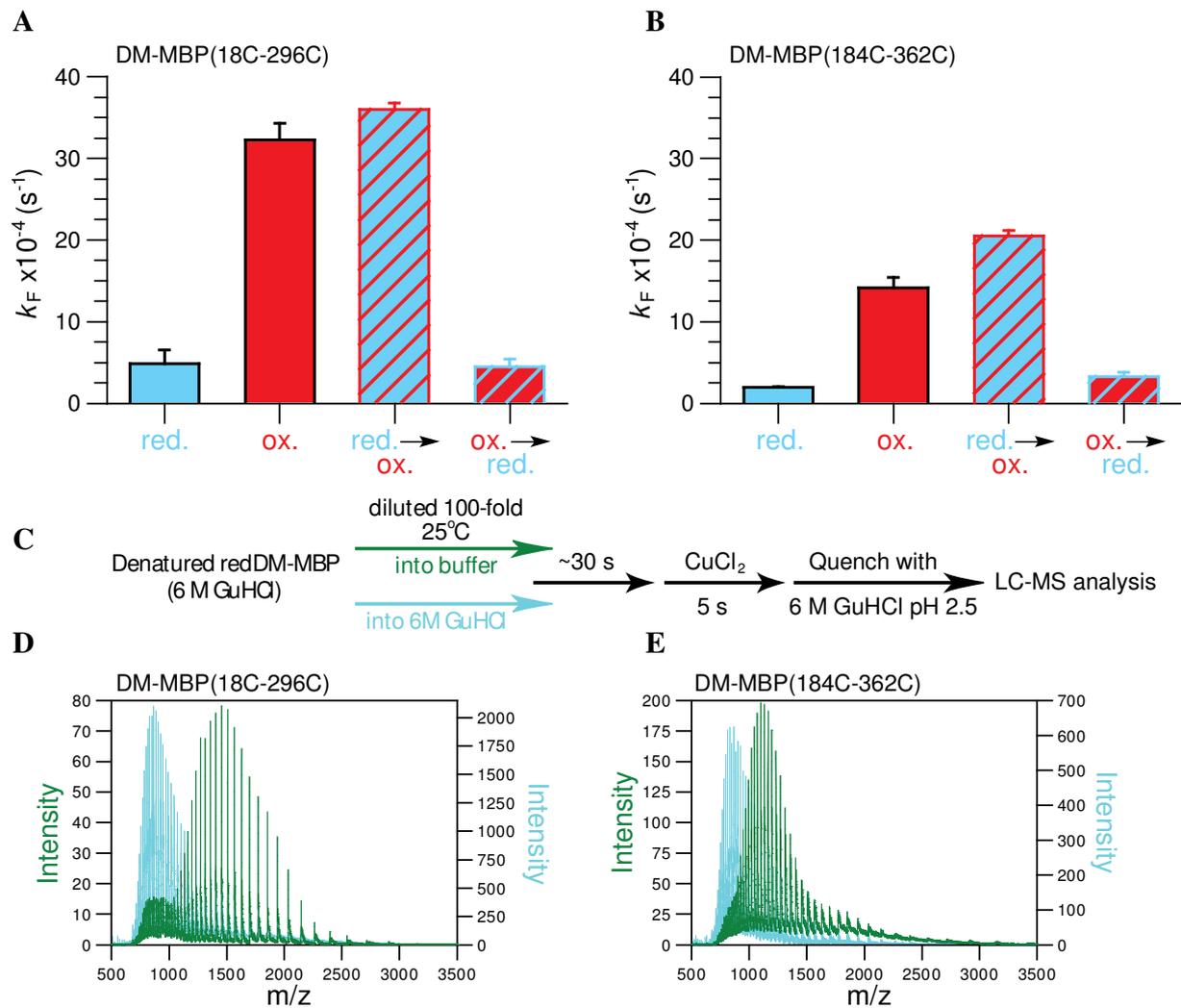


Figure 23: Disulfide bridge formation in the collapsed state accelerates the spontaneous refolding of DM-MBP. (A and B) Rates of spontaneous refolding of DM-MBP(18C-296C) (A) and DM-MBP(184C-362C) (B), 250 nM each in buffer A (final GuHCl 60 mM). Folding of reduced protein (red.) and oxidized protein (ox.), upon oxidizing the reduced protein 5 s after initiation of refolding (red./ox.) or upon reducing the oxidized protein 5 s after initiation of refolding (ox./red.). Refolding was followed by monitoring Trp fluorescence. Standard deviation from three independent measurements. (C-E) Rapid disulfide-bond formation upon protein collapse. (C) Schematic representation of the experiment. Reduced (red.) DM-MBP(18C-296C) (D) and DM-MBP (184C-362C) (E) (final $\sim 2 \mu\text{M}$) were unfolded in 6 M GuHCl and diluted 50-fold into either buffer B (green) or buffer B containing 6MGuHCl (blue), followed by the addition of CuCl_2 after 30 s. Refolding was stopped 5 s later by addition of 6 M GuHCl, pH 2.5, followed by LC-MS analysis.

4.3.2 The disulfide bonds had no significant impact on the energy barrier between the native state and the transition state

To estimate the possible effect of the disulfide bonds on the energy barrier between the transition state and the native state, we measured the rate of unfolding of the reduced and oxidized proteins. In a two-state system, the unfolding free energy of activation ($\Delta G_U^{E_a}$) can be calculated from the measured rate of unfolding, extrapolated to zero denaturant (k_U), according to the Eyring (or Arrhenius) equation:

$$k_U = \frac{k_B T}{h} e^{\frac{-\Delta G_U^{E_a}}{RT}} \quad (1)$$

where k_B is Boltzmann's constant ($1.381 \times 10^{-23} \text{ JK}^{-1}$), h is Planck's constant ($6.626 \times 10^{-34} \text{ Js}$), R is the gas constant ($8.31 \text{ JK}^{-1} \text{ mol}^{-1}$) and T is temperature in K. We found that the rate of unfolding was essentially unaffected by the disulfide bonds (Fig. 24A-B; Table 1), which meant that the unfolding energy barrier was unaffected by the introduction of the disulfide bonds. In addition, the oxidized and reduced native proteins showed similar stability toward denaturant (Figs. 24C-D and 26A-B), which meant that the disulfides had no effect on the stability of the transition state. In conclusion, introducing disulfide bonds did not change the energy barrier from the native state to the transition state, and the only other option is that the disulfides are acting on the energy barrier between the intermediate and the transition state. Furthermore, since the free energy levels of the transition state are unchanged, the disulfides are then mainly increasing the free energy of the intermediate state, or in other words destabilizing the intermediate, decreasing the energy barrier of folding, and hence accelerating the folding.

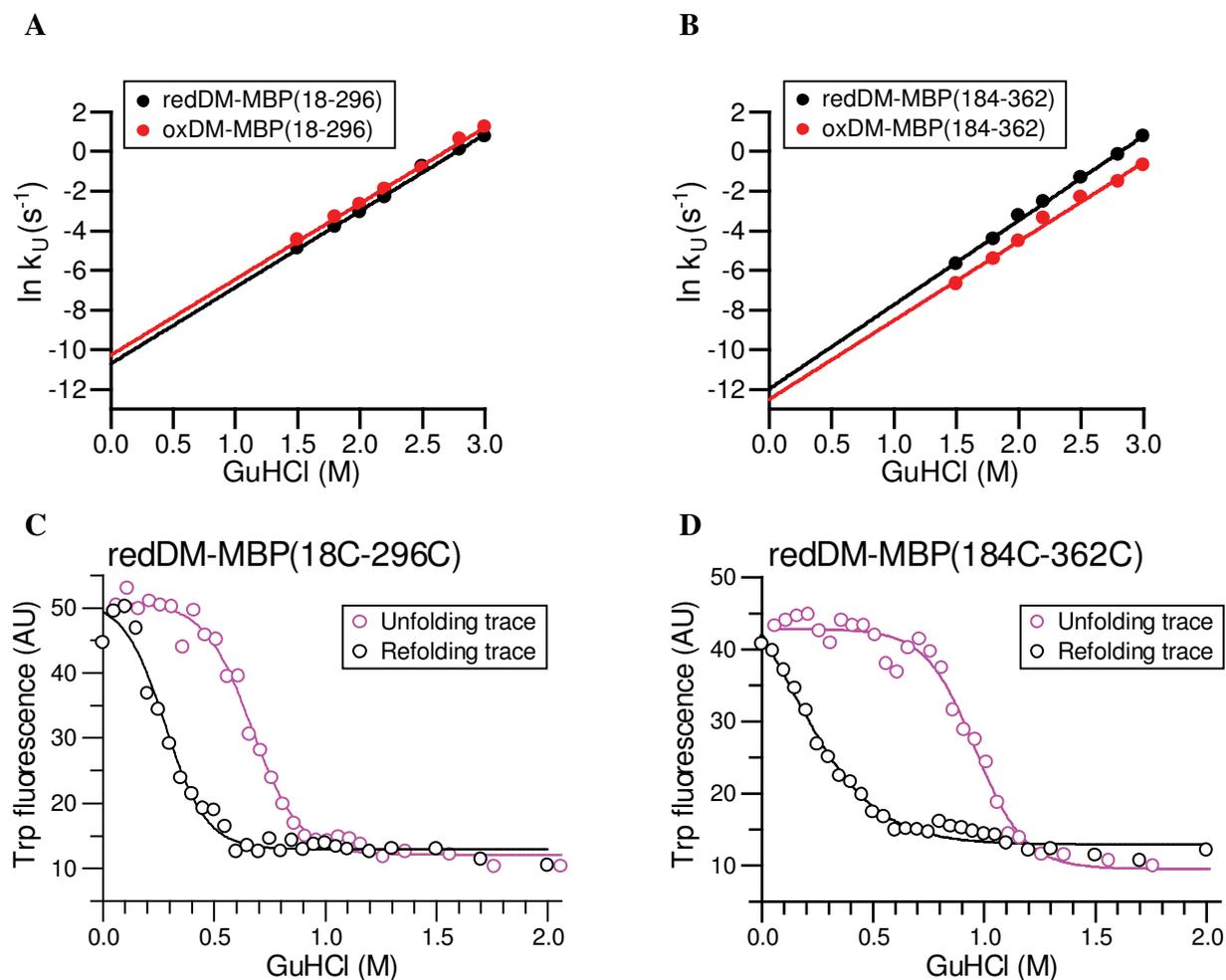


Figure 24: Disulfides do not affect the unfolding rates or the stabilities of the native states. (A and B) Kinetics of unfolding. The rate of unfolding of reduced (red.) and oxidized (ox.) DM-MBP cysteine mutants at different concentrations of GuHCl in buffer A was monitored in stopped-flow mixing experiments at 25°C by following the decrease in Trp fluorescence at 345 nm. The final protein concentration was 500 nM. Unfolding rates in the absence of denaturant were determined by extrapolation. (C and D) GuHCl-dependent unfolding and refolding. Unfolding and refolding of reduced (red.) DM-MBP(18C-298C) (C) and redDM-MBP(184C-362C) (D) was monitored by Trp fluorescence after incubation at different GuHCl concentrations for 12 hr at 25°C.

	k_U (s ⁻¹) *	$\Delta G_U^{E_a}$ (kJ mol ⁻¹)
WT-MBP	3.0×10^{-6}	104.5
DM-MBP	4.2×10^{-6}	103.7
redDM-MBP(18-296)	2.2×10^{-5}	98.5
oxDM-MBP(18-296)	3.4×10^{-5}	98.5
redDM-MBP(184-362)	6.2×10^{-6}	102.7
oxDM-MBP(184-362)	3.7×10^{-6}	104.0
redDM-MBP(4C)	1.7×10^{-5}	100.3
oxDM-MBP(4C)	1.3×10^{-5}	100.8

* The apparent unfolding rate in buffer A.

Table 2: Comparison of unfolding rate constants (k_U) and activation energy of unfolding ($\Delta G_U^{E_a}$) for WT-MBP, DM-MBP and reduced (red.) and oxidized (ox.) DM-MBP cysteine mutants at 25°C.

4.3.3 The kinetic trap is largely entropic in nature

Based on the apparent two-state behavior of the folding reaction, in which the kinetically trapped intermediate and the native state are the main populated species (Figs. 20 and 21A), we used the temperature dependence of the folding rate to provide an approximate estimate of the enthalpic and entropic contributions to the folding energy barrier. According to transition state theory, the rate constant of folding, k_F , of a two-state reaction is defined by the Eyring (or Arrhenius) equation:

$$k_F = \frac{k_B T}{h} e^{\frac{-\Delta G_F^{E_a}}{RT}} \quad (2)$$

where $\Delta G_F^{E_a}$ is the activation energy of folding. Using the definition of the free energy:

$$\Delta G_F^{E_a} = \Delta H_F - T\Delta S_F \quad (3)$$

the Eyring equation can be rewritten as

$$\ln(k_F) = \left(\ln\left(\frac{k_B T}{h}\right) + \frac{\Delta S_F}{R} \right) - \frac{\Delta H_F}{R} \left(\frac{1}{T} \right) \quad (4)$$

As $\ln(T)$ depends weakly on $1/T$ in the range measured, equation (4) is approximately linear with the slope equal to ΔH_F and an intercept that depends on ΔS_F . Hence, the enthalpy and entropy contribution to the barrier between the intermediate state (I) and the transition state (TS) can be extracted from the graph.

As pointed out earlier, the data on the flexibility and lack of structure of the intermediate state from which the refolding starts pointed to the presence of an entropic barrier to folding. Indeed, when performing the refolding at different temperature, the rate of refolding of DM-MBP and the reduced cysteine-mutants shows little temperature dependence, consistent with a largely entropic barrier to folding (Fig. 25). Calculations of the enthalpy and entropy values are reflective of a small contribution of enthalpy to the barrier (Table 2). For rate acceleration to occur, the disulfides must be reducing this entropic barrier; Assuming that the cysteines of DM-MBP(18C-296C) or DM-MBP(184C-362C) are proximal only in a certain subset of conformations of the collapsed state, disulfide formation would shift the distribution to more ordered conformations, in essence decreasing the chain entropy of the folding intermediate and

destabilizing it with respect to the transition state. Indeed, when the disulfides are introduced, the refolding acquires positive temperature dependence, reflective of an increased enthalpic contribution, and a reduced entropic contribution to the folding barrier (Table 2).

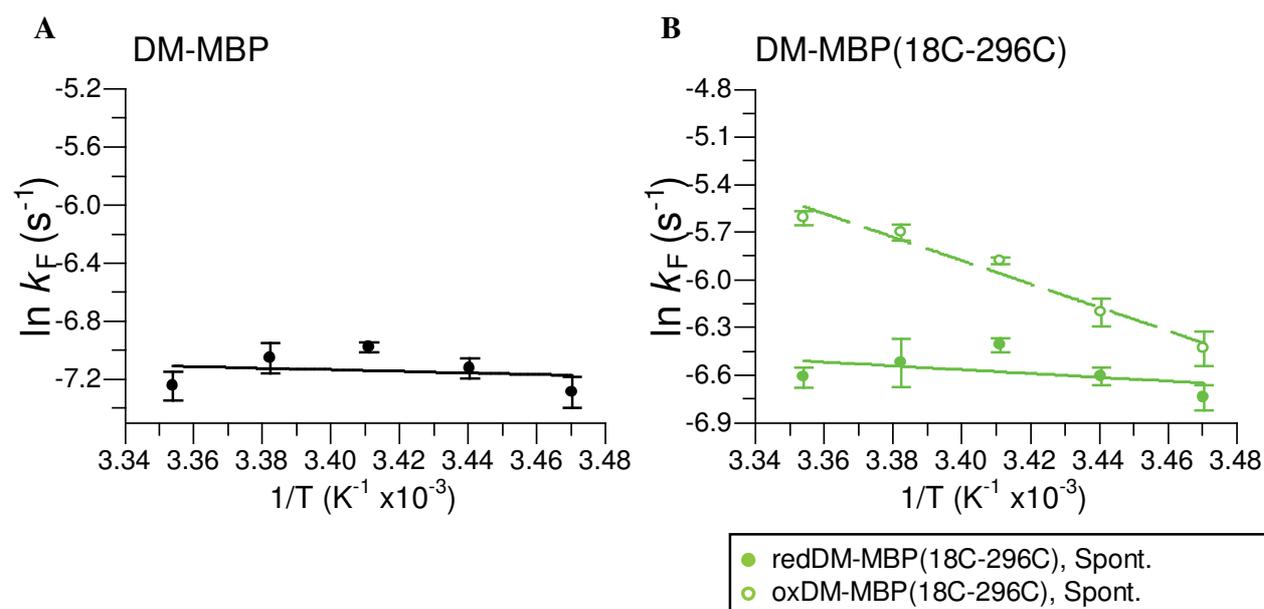


Figure 25: Introducing disulfides increases the enthalpic contribution to the folding barrier. (A and B) Arrhenius plots of the spontaneous refolding of DM-MBP (A), and DM-MBP(18C-296C) both reduced and oxidized (B). Refoldings were performed at 250nM DM-MBP in buffer B. Temperature was varied from 15-25°C. Standard deviation from three independent experiments.

	$k_F \times 10^{-4} \text{ (s}^{-1}\text{)}^*$	$\Delta G_F^{\ddagger} \text{ (kJ mol}^{-1}\text{)}$	$\Delta H_F \text{ (kJ mol}^{-1}\text{)}$	$T\Delta S_F \text{ (kJ mol}^{-1}\text{)}$
DM-MBP, Spont.	7.1±0.7	90.9	4.54	-86
DM-MBP, SR-EL/ES-ass.	40.9±3.1	86.6	92.2	5.4
DM-MBP, SR-KKK2/ES-ass.	7.4±0.3	90.8	12.1	-79
redDM-MBP(18C-296C), Spont.	13.4±0.9	89.4	9.96	-79
oxDM-MBP(18C-296C), Spont.	36.6±1.6	86.9	61.4	-26
redDM-MBP(18C-296C), SR-EL/ES-ass.	53.1±4.2	86.0	61.8	-24
oxDM-MBP(18C-296C), SR-EL/ES-ass.	61.8±0.9	85.6	58.4	-27
redDM-MBP(18C-296C), SR-KKK2/ES-ass.	13.0±0.9	89.4	6.4	-83
oxDM-MBP(18C-296C), SR-KKK2/ES-ass.	53.3±2.1	86.0	87.5	1.5
redDM-MBP(4C), Spont.	10.5±0.4	90.0	23.1	-67
oxDM-MBP(4C), Spont.	55.5±1.0	85.9	69.3	-17
redDM-MBP(4C), SR-EL/ES-ass.	49.2±2.8	86.2	84.4	-1.8
oxDM-MBP(4C), SR-EL/ES-ass.	59.3±0.8	85.7	64.2	-22
redDM-MBP(4C), KKK2/ES-ass.	10.5±0.4	90.0	17.7	-72
oxDM-MBP(4C), KKK2/ES-ass.	62.8±0.9	85.5	51.9	-34

* The apparent refolding rate in buffer B (3 independent experiments).

Table 3: Comparison of folding rate constants (k_F), activation energy of folding (ΔG_F^{\ddagger}) and the entropy barrier ($T\Delta S_F$) for WT-MBP, DM-MBP and reduced (red.) and oxidized (ox.) DM-MBP cysteine mutants at 25°C.

Moreover, the impact of disulfide bonds on folding may differ dependent on the exact regions of the protein that are restricted. Indeed, oxDM-MBP(18C-296C) refolded faster than oxDM-MBP(184C-362C) (Fig. 23A-B). This difference in kinetics correlated with the absence of hysteresis in the unfolding-refolding curves of oxDM-MBP(18C-296C), whereas oxDM-MBP(184C-362C) preserved the hysteresis effect (Fig. 26A-B). In addition, H/D exchange

measurements confirmed that 12 hours after diluting the unfolded protein into 0.5 M GuHCl, a significant fraction of oxDM-MBP(18C-296C) was in the native state, whereas oxDM-MBP(184C-362C) was mostly unfolded (Fig. 26C).

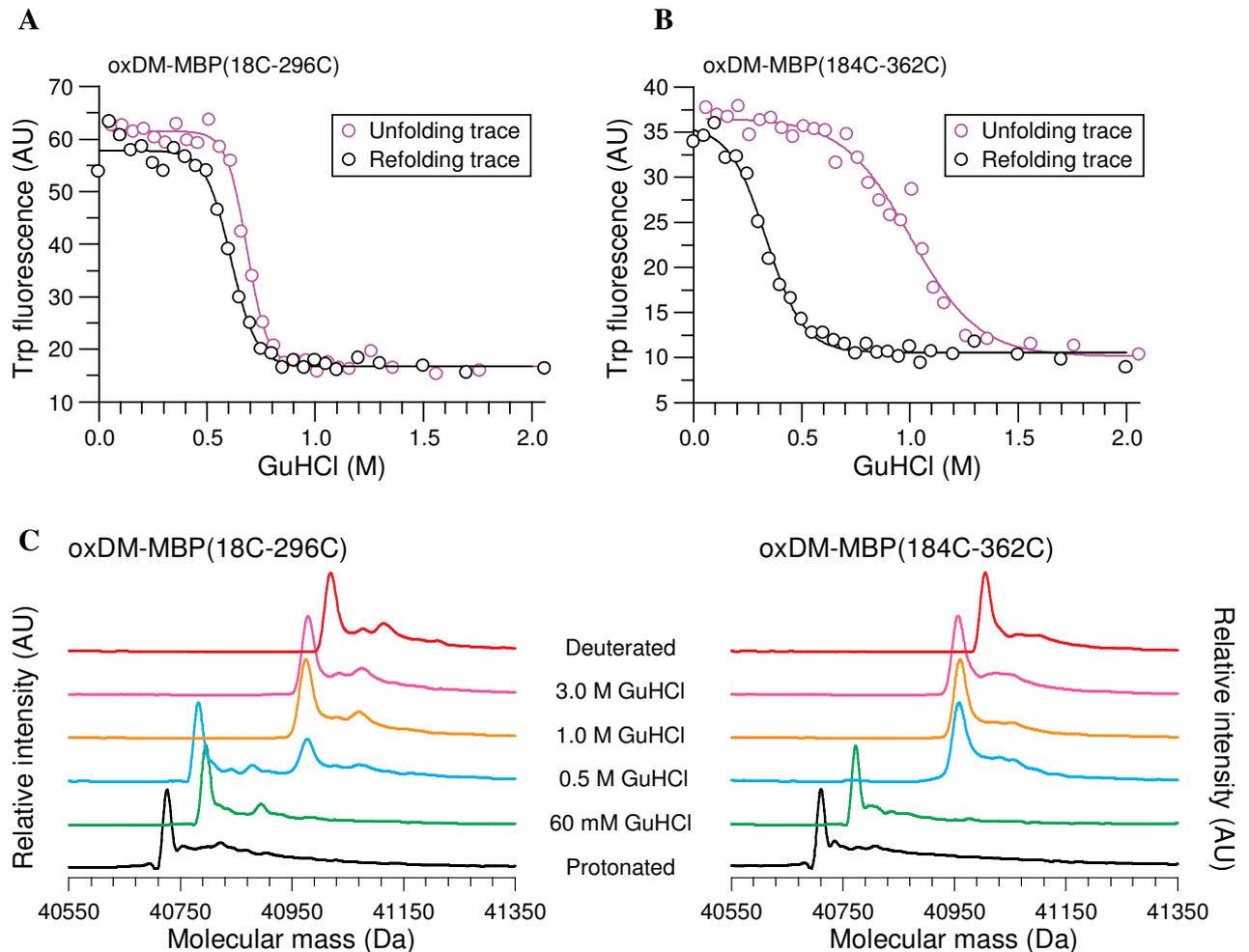


Figure 26: The disulfide bridge in DM-MBP(18C-296C) eliminates the hysteresis, but not in DM-MBP(184C-362C). (A and B) GuHCl-dependent unfolding and refolding of oxidized (ox) DM-MBP(18C-298C) (A) and oxDM-MBP(184C-362C) (B) were monitored by Trp fluorescence. (C) Characterization of the kinetically trapped refolding intermediate of DM-MBP cysteine mutants by H/D exchange. Pulsed H/D exchange after incubation in different denaturant concentrations. Deconvoluted mass spectra of oxidized (ox.) DM-MBP(18C-296C) and oxDM-

MBP(184C-362C) as a function of denaturant monitored by ESI-QToF mass spectrometry. Proteins were diluted from 3 M GuHCl into buffer B to the final GuHCl concentrations indicated. After incubation for 12 hr samples were subjected to a 10 s deuterium pulse. The native protonated and deuterated proteins are shown as reference.

4.4 The chaperonin cage mimics disulfide-mediated constraints

Having determined that the barrier to DM-MBP refolding was largely entropic in nature, it seemed plausible that GroEL was accelerating the refolding of DM-MBP by decreasing this barrier, in a way that is similar to the disulfide-mediated constraints. To test this hypothesis, we have used the single-ring version of GroEL, or SR-EL, which undergoes one single round of ATP hydrolysis upon GroES binding, resulting in stable substrate encapsulation in low-salt buffer (Fig. 27A). Hence, one would be able to measure the refolding rates without the interference from cycling (Hayer-Hartl et al., 1996; Weissman et al., 1996).

4.4.1 SR-EL accelerates the refolding of the cysteine mutants

Notably, under these low-salt conditions, the spontaneous rate of refolding of DM-MBP and its mutants is faster, an effect which we attribute to the stabilization of the kinetically-trapped intermediate by higher salt concentrations. In low-salt buffer, SR-EL nevertheless accelerated the refolding of both oxidized mutants to a similar extent; ~1.5 times beyond their spontaneous refolding rate (Fig. 27B-C). Importantly, SR-EL accelerated both the reduced and oxidized proteins to a similar extent, suggesting that the chaperonin effects and the disulfide effects were not additive (Fig. 27B-C). In other words, the chaperonin could be accelerating the refolding similarly to the disulfides, namely by reducing the entropic barrier to folding. Notably, the finding that the chaperonin accelerated the refolding of both the oxidized proteins beyond their

spontaneous rates would be consistent with the chaperonin cavity exerting a global confinement effect, and the disulfide bond-mediated constraints acting more locally. In order to validate this hypothesis, we have constructed the tetracysteine DM-MBP(4C) variant, which combines both N- and C-domain mutants.

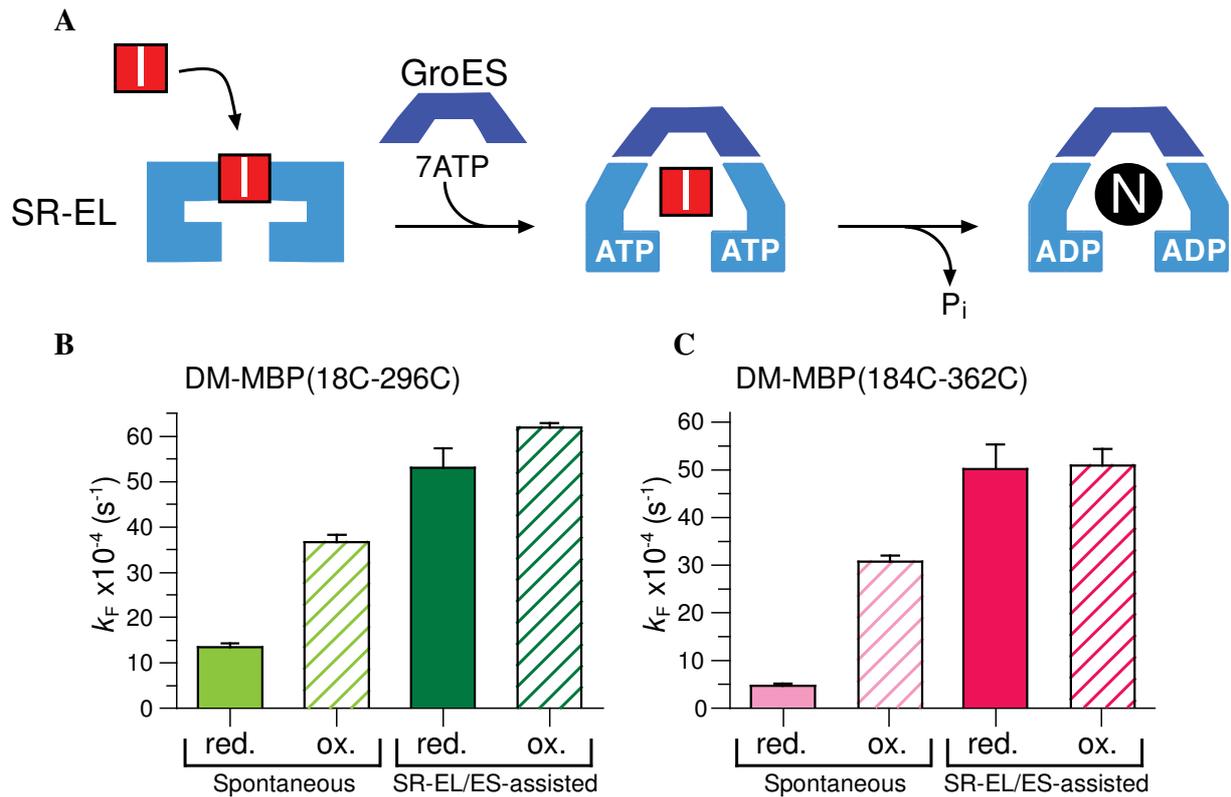


Figure 27: Chaperonin-assisted folding is insensitive to disulfide formation. (A) Protein folding in the SR-EL/ES cage upon a single round of encapsulation of folding intermediate (I). Due to the lack of a second ring, the native protein (N) remains encapsulated. (B-C) Rates of spontaneous and assisted refolding of DM-MBP(18C-296C) (B) and DM-MBP(184C-362C) (C) in the reduced (red.) and oxidized states (ox.). Refolding was performed at 250 nM DM-MBP in buffer B. Standard deviation from three independent measurements.

4.4.2 Constraining the N and C domains of DM-MBP is comparable to the chaperonin effects

DM-MBP(4C) was able to form disulfides in the presence of CuCl_2 , as determined by LC-MS (Fig. 28A), and bound maltose as efficiently as DM-MBP (Fig. 28B).

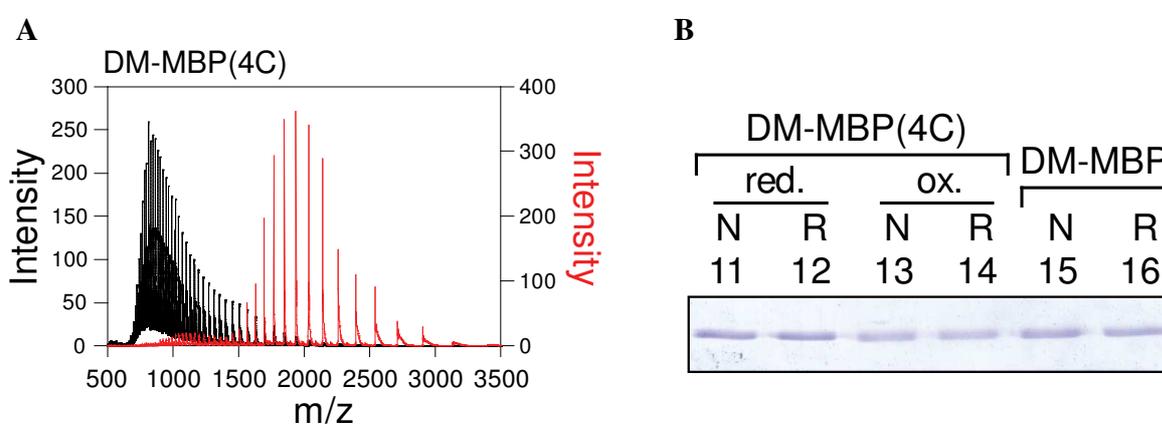


Figure 28: The reduced and oxidized versions of DM-MBP(4C) are functional at maltose binding. (A) Charge state distribution of reduced and oxidized DM-MBP(4C) measured by LC/MS. (B) Affinity of maltose binding of reduced, oxidized DM-MBP(4C) and DM-MBP.

Strikingly, the spontaneous refolding of oxDM-MBP(4C) was fully accelerated to the rate of chaperonin-assisted folding, and this was reflected in the absence of hysteresis effect in the guanidinium denaturation-renaturation curves (Fig. 29B-C). The redDM-MBP(4C) on the other hand, preserved the hysteresis effect in the guanidinium melts, just like DM-MBP (Fig. 29A). The native state stability was only slightly destabilized by the presence of the disulfide bonds, as shown by the unfolding melts of both red and oxDM-MBP(4C) (Fig. 29A-B). The unfolding rate was nevertheless unchanged for both proteins, suggesting that the tetra-cysteines in DM-

MBP(4C) are acting similarly to the disulfides in the double-cysteine mutants (Fig. 29D). Furthermore, as in the case of the single disulfide bonds, the rate of assisted refolding was the same for the reduced and oxidized proteins, and this again highlighted the redundancy of the disulfide effect and the chaperonin effect (Fig. 29C). All together, the results suggest that the effect of constraining both domains is comparable to the global compaction exerted by the chaperonin.

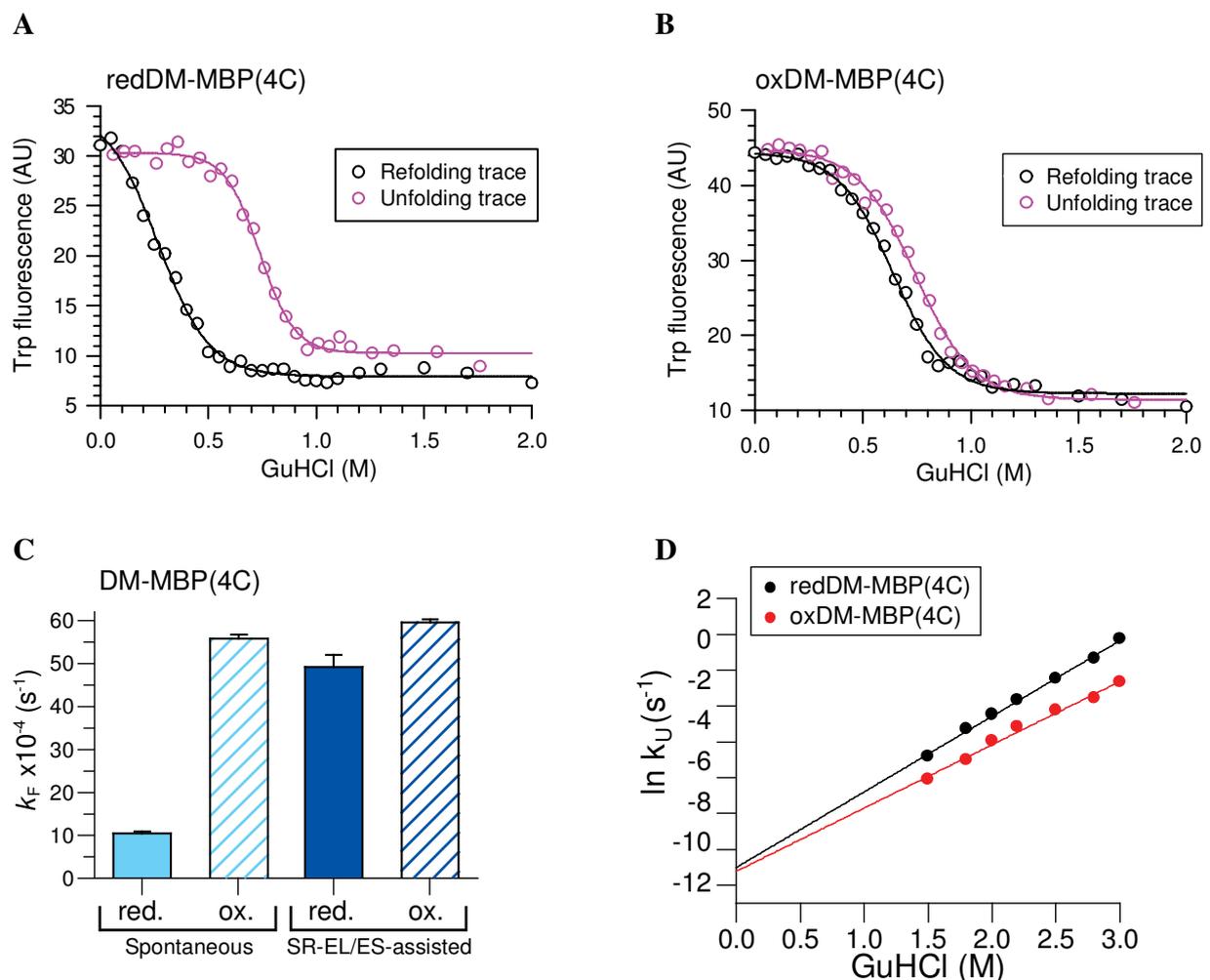


Figure 29: Disulfide bridges in both N and C domains of DM-MBP accelerate folding similarly to the chaperonin cage. (A and B) GuHCl-dependent unfolding and refolding of oxidized (ox) DM-MBP(4C) (A) and reduced (red) DM-MBP(4C) (B) were monitored by Trp fluorescence. (C) Rates of spontaneous and assisted

refolding of DM-MBP(4C) in the reduced (red.) and oxidized states (ox.). Refolding was performed at 250 nM DM-MBP in buffer B. Standard deviation from three independent measurements. (D) Kinetics of unfolding of DM-MBP(4C). The rate of unfolding of reduced (red.) and oxidized (ox.) DM-MBP cysteine mutant at different concentrations of GuHCl in buffer A was monitored in stopped-flow mixing experiments at 25°C by following the decrease in Trp fluorescence at 345 nm. The final protein concentration was 500 nM. Unfolding rates in the absence of denaturant were determined by extrapolation.

4.4.3 The disulfide bridges energetically mimic the confinement by the chaperonin

To investigate whether long-range disulfide bonds energetically mimic the confinement by the chaperonin cage, we analyzed the temperature dependence of the folding rates. The spontaneous refolding of redDM-MBP(4C) proved to be relatively insensitive to temperature variation between 15°C and 25°C (Fig. 30C), consistent with a large entropic folding barrier (Bicout & Szabo, 2000). As expected, estimations of enthalpic and entropic contributions to the activation barrier for the spontaneous folding of the redDM-MBP(4C) showed that the barrier was mostly entropic in nature (Table 2). In contrast, the folding rate of oxDM-MBP(4C) showed a pronounced, positive temperature dependence (Fig. 30C), indicating that the activation barrier has gained a significant enthalpic component and the entropic contribution is largely reduced (Table 2). Strikingly, the SR-EL/ES-assisted folding of DM-MBP and the all the reduced cysteine mutants displayed a similar positive temperature dependence (Fig. 30A-C), suggesting a strongly reduced entropic folding barrier (Table 2). Thus, the spontaneous folding of the oxidized proteins and the chaperonin-assisted folding of the reduced proteins have similar rate-limiting steps, consistent with a common mechanism of accelerated folding by entropic confinement. As for the enthalpic component of the activation barrier, it may be reflecting side-

chain friction during folding, resulting either from disulfide-mediated constraints or steric restrictions imposed by the chaperonin cage.

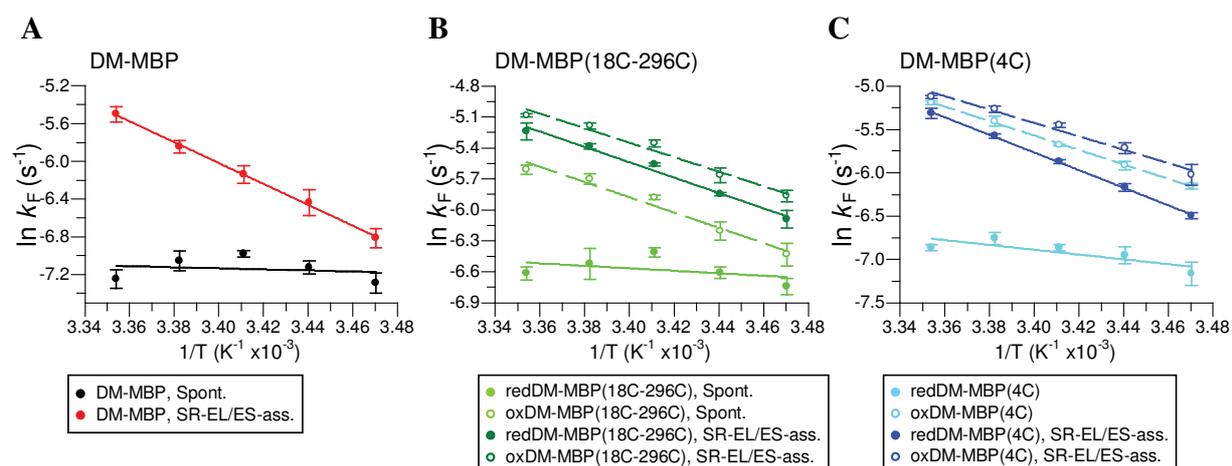


Figure 30: The chaperonin-assisted and disulfide mediated refoldings show a reduced entropic component.

(A-C) Arrhenius plots of spontaneous and SR-EL/ES assisted refolding of DM-MBP (A), red and oxDM(18C-296C) (B), and red and oxDM(4C). Refoldings were performed at 250nM DM-MBP in buffer B. Temperature was varied from 15-25°C. Standard deviation from three independent experiments.

4.5 Charge clusters in the chaperonin cage wall are crucial for the folding

The wall of the GroEL cis-cavity has a net charge of -42 (189 negatively and 147 positively charged amino acid residues). A number of negative charges (residues E252, D253, E255, D359, D361, and E363), all in the apical domain, cluster in two circular layers (Fig. 31A). Most of these residues (E252, D253, E255, and E363) are highly conserved among GroEL homologs, although they have no apparent role in the basic GroEL functions of substrate and GroES binding. In the following section, we try to underpin the role of those charges in the folding of

DM-MBP, by taking advantage of a mutant of SR-EL, SR-KKK2 (Tang et al., 2006), which has a cavity net charge of zero due to the mutations of D359, D361, and E363 into lysines.

4.5.1 The SR-mutant is unable to reduce the entropic barrier to folding

As shown previously in Tang et al. (2006), SR-KKK2 binds and encapsulates DM-MBP as efficiently as SR-EL but is unable to accelerate its folding, whereas the rate of WT-MBP folding is essentially unaffected. This suggests that the KKK2 mutant may have lost the ability to reduce the entropic barrier of DM-MBP folding, in which case the rate-limiting step of folding inside the mutant chaperonin may be similar to that of spontaneous folding. To test this possibility, we first recorded the temperature dependence of SR-KKK2/ES-assisted folding. (Note that SR-EL and SR-KKK2 have similar ATPase rates and undergo only a single round of ATP hydrolysis upon GroES binding). In contrast to SR-EL/ES-mediated folding, the SR-KKK2/ES-assisted folding of DM-MBP was temperature independent, similar to spontaneous folding (Fig. 31B). Introducing disulfide bonds restored the positive temperature dependence of folding, as shown for oxDM-MBP(18C-296C) (Fig. 31B). Thus, SR-KKK2 seems to have lost the ability to entropically destabilize the DM-MBP folding intermediate (Table 2).

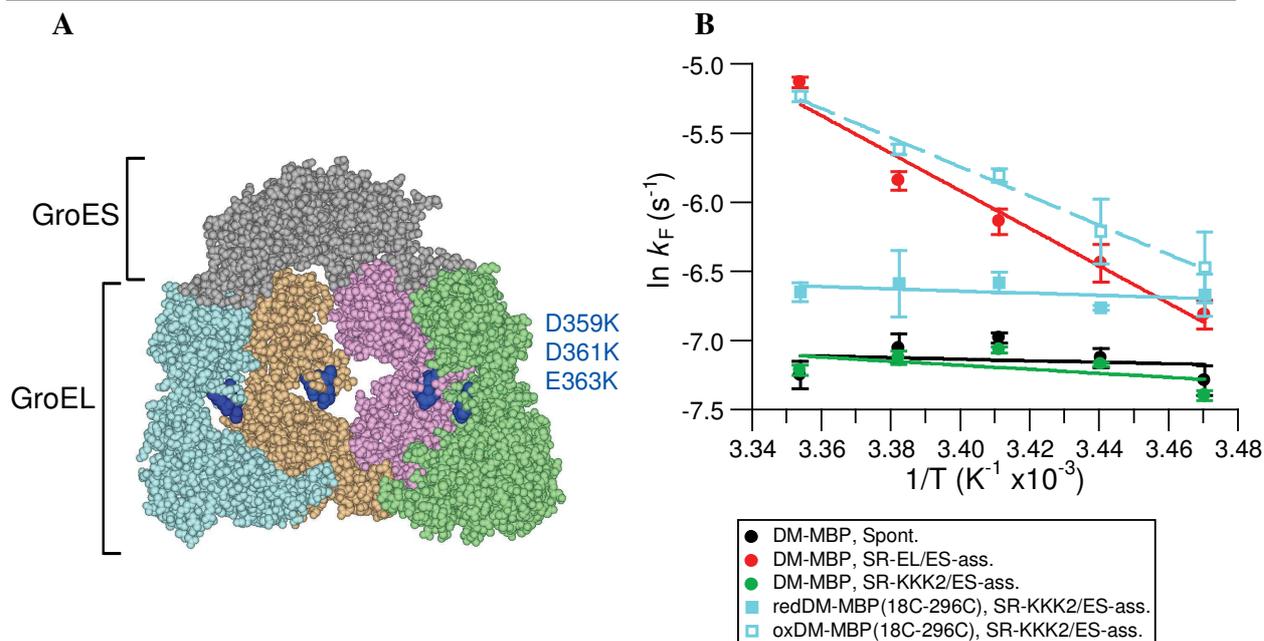


Figure 31: SR-mutant is unable to decrease the entropic barrier to the folding of DM-MBP. (A) Space-filling model of four subunits of the GroEL/ES-(ADP)₇ complex (Xu et al., 1997; PDB 1AON, DS ViewerPro) offering a view into the chaperonin folding cage. The negatively charged residues mutated in SR-KKK2 (D359K; D361K; E363K) are highlighted in blue. (B) Temperature dependence of spontaneous and assisted folding by SR-EL/ES or SR-KKK2/ES of DM-MBP and reduced and oxidized DM-MBP(18C-296C). Refolding was measured previously. Standard deviation from three independent measurements.

4.5.2 Folding inside the SR-mutant cage mimics spontaneous refolding

To corroborate the previous conclusion, we analyzed the folding reactions under various physical conditions, including the effect of low denaturant, which may increase the flexibility of the kinetically trapped intermediate, thereby decelerating folding. Indeed, the rate of spontaneous and SR-KKK2/ES-assisted folding of DM-MBP decreased with increasing GuHCl (15-60 mM) (Fig. 32A). In contrast, the SR-EL/ES-assisted folding rate was independent of denaturant (Fig. 32A), consistent with the ability of the wild-type chaperonin to promote structure formation in the intermediate. Next, we tested the influence of trimethylamine N-oxide (TMAO) on folding,

an osmolyte known to reduce structural flexibility in proteins (Qu & Bolen, 2003), presumably via the enhancement of water structure (Zou et al., 2002). The rate of spontaneous and SR-KKK2/ES-assisted folding showed a positive dependence on TMAO concentration (Fig. 32B), whereas the SR-EL/ES-assisted folding remained essentially unchanged (Fig. 32B). This suggests that the wild-type chaperonin mimics the effect of TMAO, consistent with the proposal that the charged lining of the cage may promote protein compaction by an ordering effect on water structure (England & Pande, 2008). Together, these results provide evidence that by removing its negative net charge, the chaperonin cavity is converted from an active to a largely passive folding environment.

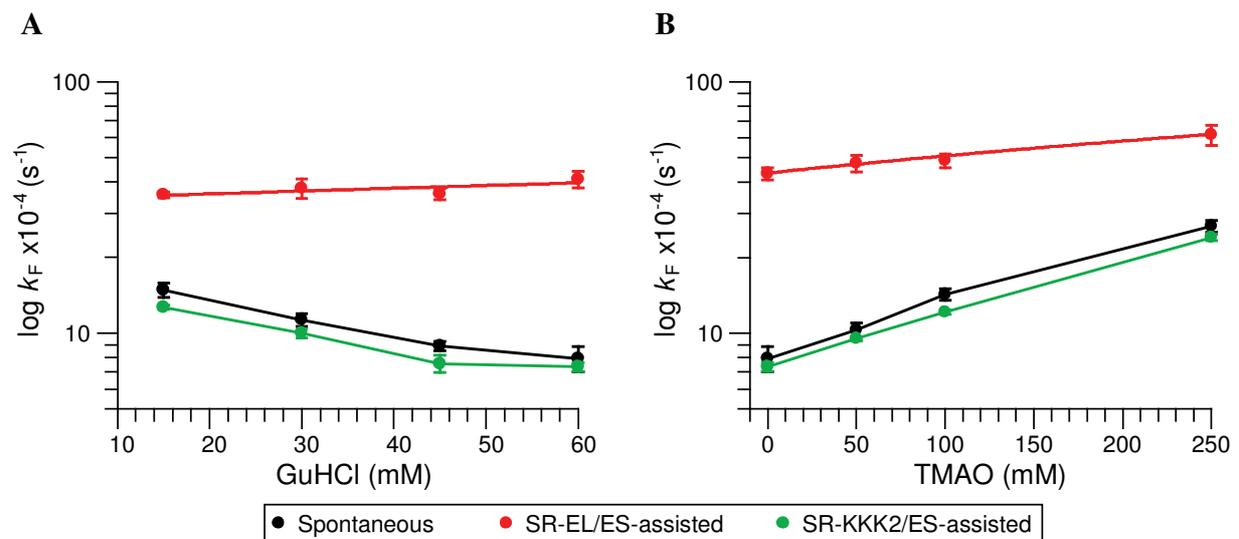


Figure 32: Folding inside the SR-mutant mimics the spontaneous folding. (A and B) Rates of spontaneous and assisted refolding of DM-MBP at varying concentrations of GuHCl (A) and TMAO (B) are shown. Refolding was performed as above, either spontaneous or assisted by SR-EL/ES or SR-KKK2/ES as indicated. Standard deviation from three independent measurements.

4.6 Unfolding upon substrate binding seems unessential for rate acceleration

One of the mechanisms proposed to explain how GroEL accelerates the folding of proteins is “iterative annealing” (Shtilerman et al., 1999; Thirumalai & Lorimer, 2001). According to this mechanism, GroEL is thought to unfold kinetically-trapped proteins, which are unable to fold due to some non-native stabilizing interactions, thereby positioning them at a higher energy level, and affording them a new chance to refold to the native state. In the case of DM-MBP, there appears to be no such trapped state that is stabilized by non-native interaction, since the kinetically-trapped intermediate was shown to be flexible, and devoid of any structure. In the section below, we provide further experiments which support the idea that unfolding is not essential for rate acceleration, at least in the case of DM-MBP.

4.6.1 FRET analyses reveal more chain compactness in SR- or V263S-bound DM-MBP as compared to EL-bound DM-MBP

We have measured FRET efficiencies along vectors 52-298 and 175-298, for both SR- and EL-bound DM-MBP. EL-bound proteins showed bimodal distribution alongside vector 52-298 with more than one-third of the proteins in an expanded conformation ($f_E \sim 0.05$), whereas those bound to SR showed a smaller fraction ($\sim 25\%$) in the expanded conformation ($f_E \sim 0.05$) (Fig. 33A). The majority of the SR-bound proteins centered around an $f_E \sim 0.65$ (Fig. 33A). As for vector 175-298, both EL- and SR-bound proteins showed a large population in a compact state (Fig. 33B). When comparing the refolding in GroEL and SR, little difference is observed (Fig.

33C). Hence, it is plausible that unfolding by GroEL is not really a necessary event to accelerate the folding.

We have also created the V263S mutant of GroEL, which was first reported in 2000 by Farr et al. as a binding deficient mutant. This mutation lies in the center of the peptide binding surface of the apical domain on an α -helix and faces upward toward an overlying α -helix that forms the inlet to the central cavity when GroEL is in an open acceptor state. It was shown to abolish both GroES and substrate binding, and complementation experiments showed that the mutant was unable to rescue the growth of EL-deficient *E. coli* (Farr et al., 2000). We nevertheless found that this mutant retained certain affinity for substrate protein such as DM-MBP, as shown by the inability of DM-MBP to refold in the presence of this mutant. It seemed plausible that the binding pattern of DM-MBP to this mutant would be different than to wt GroEL. We therefore determined FRET efficiencies of V263S-bound DM-MBP between positions 52 and 298, as well as 175-298 (Fig. 33A-B). A significant fraction of DM-MBP, larger than in the case of EL, along this vector was in a collapsed form. We could nevertheless not rule out a correlation between the size of the high f_E population and the folding acceleration as in the case of SR, since the refolding in the presence of this mutant was drastically slowed down. It seemed though that this decrease in the folding rate might have been due to the early escape of DM-MBP into bulk solution, due to its decreased affinity to V263S. Indeed, when the refolding experiment was conducted at an 80 fold excess of V263S mutant, the refolding rate was restored (Fig. 33C). This experiment highlights the importance of the hydrophobic character of the apical domains in maintaining a strong enough affinity to the substrate, and preventing its premature release before binding of GroES.

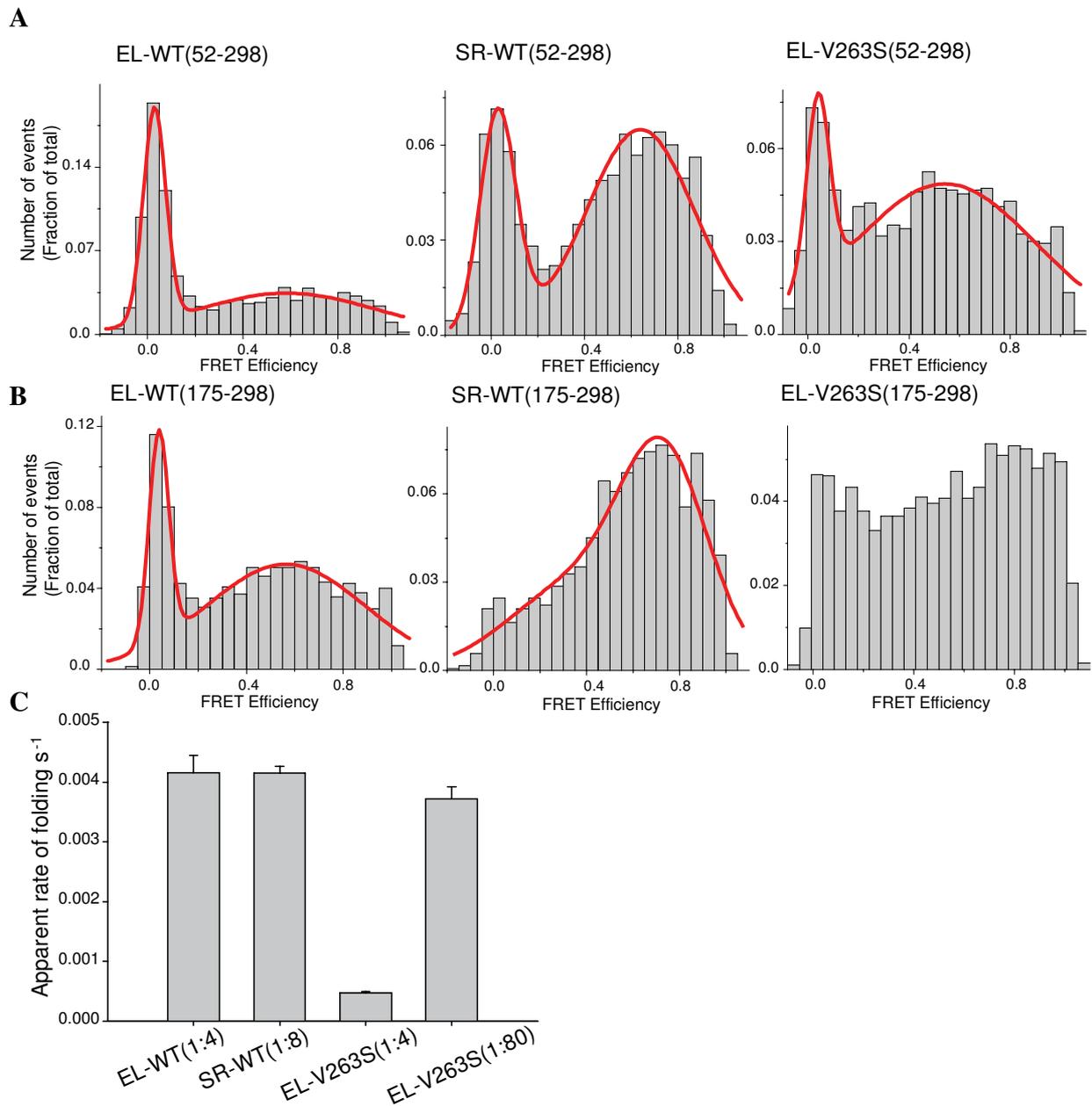


Figure 33: FRET distributions of chaperonin-bound DM-MBP and folding rates. (A and B) SpFRET histograms of GroEL-, SR-, and EL-V263S-bound DM-MBP along distance vectors 52-298 and 175-298. GuHCl-denatured double-labeled DM-MBP (3 nM) was diluted 50-fold (60 pM final concentration) into low-salt or high-salt buffer containing 3 μ M GroEL. Peak values of a Gaussian fit to the FRET efficiency distributions (f_E) are indicated in the text. Representative histograms of two independent measurements are shown. (C) Refolding rates of DM-MBP in the presence of various chaperonins. Refoldings were performed as described previously, by

monitoring the Trp fluorescence of 250 nM DM-MBP. The ratio of DM:EL(or SR) are indicated in the figure. The ES:EL ratio was always 2:1. Shown are averages and standard deviations from three independent experiments.

4.6.2 Crowding in the presence of Ficoll restores folding efficiency of the V263S mutant

We next performed FRET measurements in the presence of 16% Ficoll70, a crowding reagent used to mimic the cellular crowding conditions. It seemed that crowding was greatly enhancing the population of stretched DM-MBP bound to SR, as shown in Fig. 34A for both distance vectors 52-298 (from ~25% to ~50%) and 175-298 (from very little to ~25%) (Compare to Fig. 33A-B). If the size of the population in an extended conformation were indeed playing a role in rate acceleration, one would expect to see a significant increase in the folding rate in the presence of Ficoll. We nevertheless observe that SR accelerates the refolding to a similar extent in the presence and absence of Ficoll, again suggesting that stretching of DM-MBP does not enhance the rate acceleration, and is merely a result of the multi-domain binding to various apical domains (Fig. 34B). One could nevertheless argue that the fraction of substrates in the unfolded state could be in fast equilibrium with the compact fraction of substrates, and therefore play a role in rate acceleration. Binding of GroES takes around 1 sec to complete, and the compact population could have unfolded during that time, given a fast equilibrium (Sharma et al., 2008).

Interestingly, we observe that the rate acceleration afforded by a 80 fold excess of EL-V263S in high-salt buffer was now obtained with only 4 fold excess of the mutant in the presence of Ficoll (Fig. 34B). In other words, the EL-V263S in Ficoll was as efficient as GroEL in accelerating the refolding of DM-MBP in high-salt buffer, probably due to an increased

association constant of substrate to EL-V263S. In fact, association constants under crowded conditions relative to those in dilute solution are predicted to increase by as much as 2-3 orders of magnitude (Zimmerman & Minton, 1993). Again, this highlights the importance of an interaction of the substrate with the apical domains that is prolonged enough for GroES to bind and encapsulate the protein, and the crowding differences between folding *in vitro* and *in vivo*.

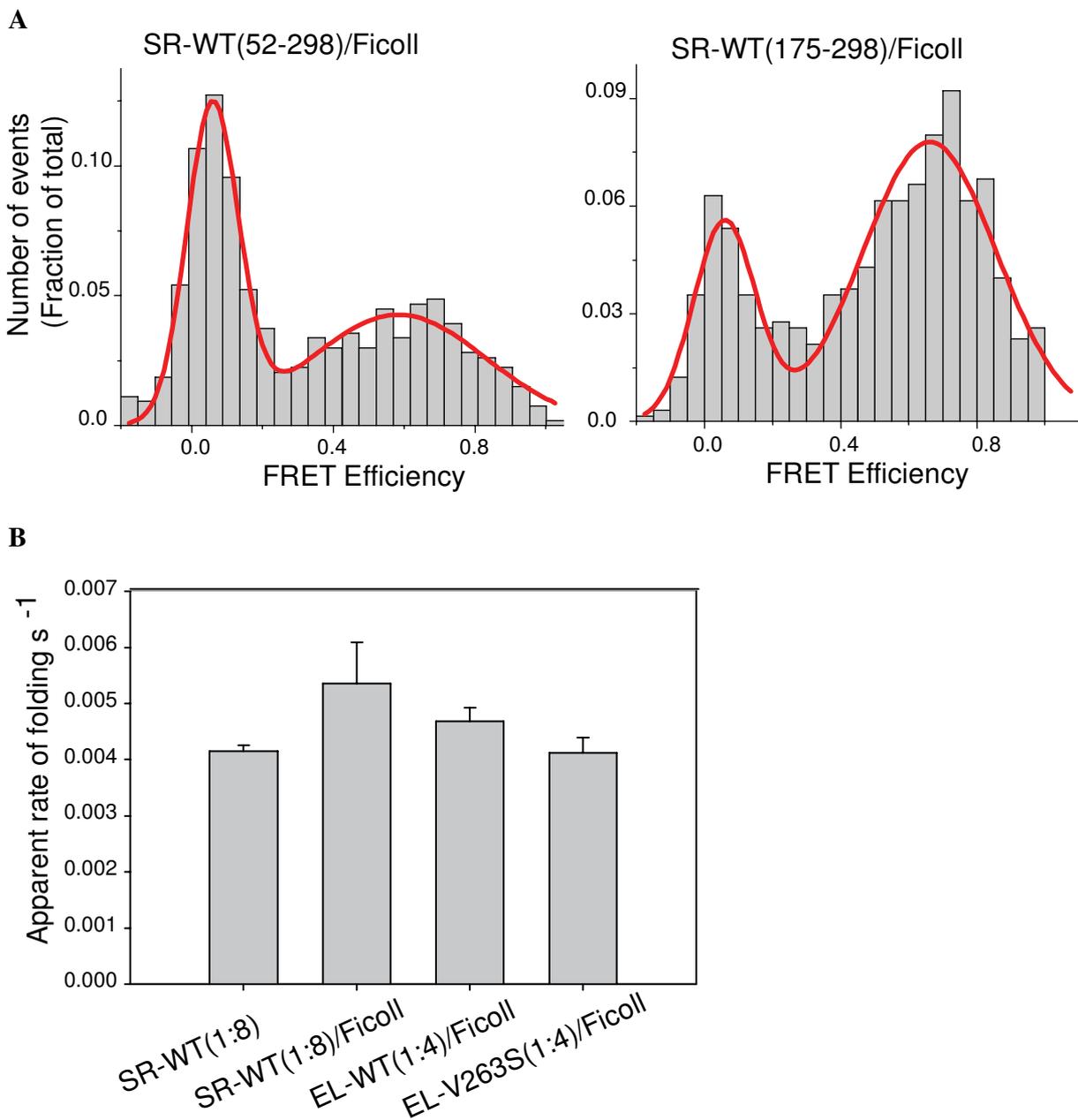


Figure 34: Effect of Ficoll70 on DM-MBP refolding and SpFRET distributions. (A) SpFRET distributions of SR-bound DM-MBP. The protein was diluted into low-salt buffer containing 3 μ M SR, and ficoll70 was only later on added to avoid aggregation (final concentration of 16%). (B) Effect of Ficoll on refolding. Refolding was carried out as previously described. Care was taken to avoid aggregation, by adding ficoll after binding DM-MBP to the chaperonin. The ratio of DM:EL(or SR) are indicated in the figure. The ES:EL ratio was always 2:1. Shown are averages and standard deviations from three independent experiments.

Such increases in the association constants under crowding conditions could result in an increased propensity of folding protein molecules to aggregate (Ellis & Hartl, 1996). The chaperone network is thus of great importance for the cell to protect itself from the devastating consequences of aggregation. It is nevertheless notable that chaperonins are absent from oxidative environments of the cell, such as the periplasm. We hypothesize that this conspicuous absence could be related to the redundancy of the role of GroEL and the periplasmic disulfide isomerases in accelerating folding through entropic confinement.

4.7 A comparison of the folding of DM-MBP cysteine mutants shows no growth advantages conferred by the mutations

We envisaged that if the disulfides were acting similarly to GroEL in an *in vivo* context, i.e accelerating the folding by entropic confinement, we would be able to see a rate acceleration of DM-MBP(18C-296C) upon formation of the disulfide bond during translocation to the periplasm. In order to test this, we have engineered the same cysteine mutants as previously, but on a background of DM-MBP which contains the leader sequence targeting it to the periplasm (plasmid PD1 encoding WT-MBP along with its leader sequence was a kind gift from Dr. Betton). As an assay of rate acceleration, we have monitored the amount of maltose fermentation

of an otherwise MBP-deficient *E. coli* strain (pop6499) upon transformation with plasmids encoding the various mutants. Notably, this strain carries the *maltC* allele, whose presence ensures constitutive expression of the maltose operons, even in the absence of maltose. After plating the cells on Mac Conkey agar, the color of the colonies shifts to red upon maltose fermentation due to the release of acids. We assume that if the proteins don't fold fast enough in the periplasm, they would aggregate and/or degrade, and lead to less fermentation of maltose.

4.7.1 Periplasmic DM-MBP allows for less maltose fermentation than WT-MBP

As a control, cells were transformed with the plasmid lacking the leader sequence, and these do not exhibit any maltose fermentation phenotype (Fig. 35). When observing colonies transformed with WT-MBP, they are almost fully red. In contrast, those transformed with DM-MBP are mainly either white or white with a little red surface in the middle (Fig. 35). This would suggest that DM-MBP aggregates and/or is degraded in the periplasm and is therefore unable to shuttle maltose. It also shows that "older cells" accumulate enough folded DM-MBP with time, and are able to ferment maltose; hence the red surfaces in the middle of some colonies.

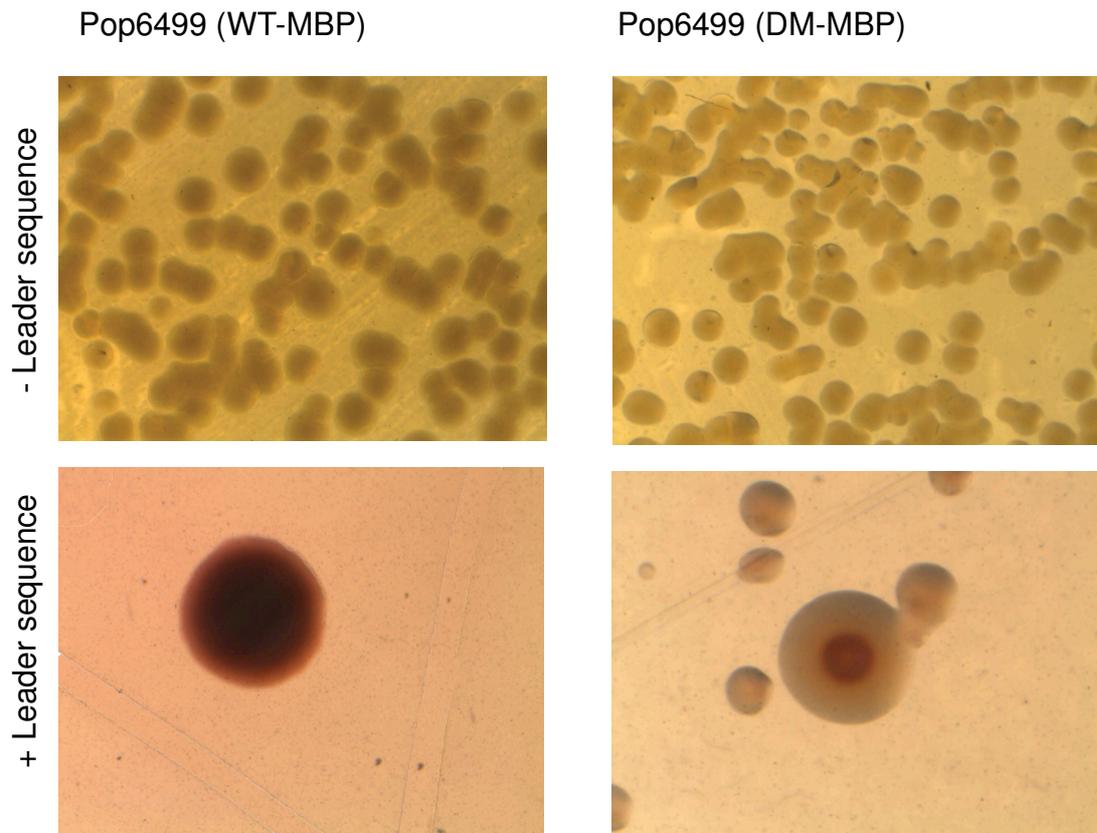


Figure 35: Maltose fermentation by Pop6499 expressing WT-MBP and DM-MBP. Cells were transformed with the indicated plasmids, plated on 0.4% maltose-containing MacConkey plates, and grown at 37°C overnight. Shown are representative colonies. Images were acquired on an Olympus stereomicroscope.

4.7.2 Periplasmic DM-MBP mutants exhibit temperature-dependent phenotype

We next examined the color of Pop6499 cells which are transfected with either plasmid encoding for DM-MBP, or the three cysteine mutants, and grown at 37°C. Figure 36 shows three representative colonies randomly chosen from each the DM-MBP and the DM-MBP(184C-362C) cysteine mutant plate, and monitored over time. No particular color differences were noted within a day of the plating, which suggested that the disulfide bridge in the C-domain of

DM-MBP did not lead to a faster refolding of the protein in the periplasm, or at least not to such an extent that is detectable by this method (Fig. 36A). Similar results were obtained with the other cysteine mutants. When comparing the sizes of the colonies expressing the different DM-MBP mutants (Fig. 36B), it was obvious that those colonies expressing the periplasmic cysteine mutants exhibited a smaller size as compared to those expressing periplasmic WT-MBP, and cytoplasmic DM-MBP, suggesting that those proteins might be causing periplasmic stress that hinders cell division.

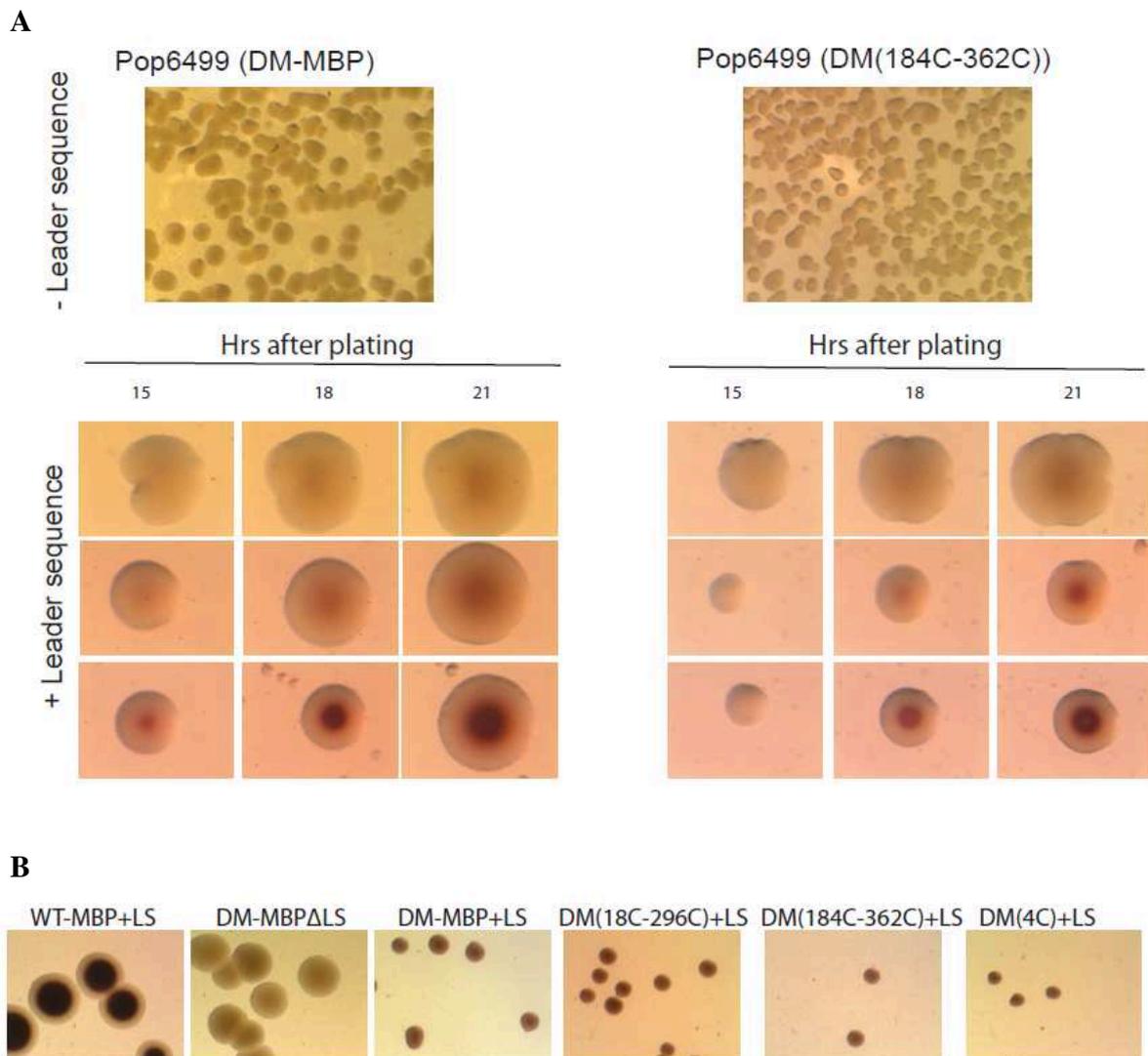


Figure 36: Maltose fermentation by various DM-MBP mutants. (A and B) Cells were transformed with the indicated plasmids, plated on 0.4% maltose-containing MacConkey plates, and grown at 37°C overnight. Three clones were photographed at various time points, and show the increase in maltose fermentation, as evident from the gradual appearance of the red spot. The other 2 cysteine mutants, namely the DM(18C-296C) and DM(4C) showed similar results (data not shown). LS stands for leader sequence. In B, magnification was kept constant for the various images.

Interestingly, as shown in figure 37, when those plates were incubated for an extra 24 hours at 37°C, the shape of the colonies expressing the various DM-MBP variants was no longer circular but rather amorphous, almost looking like a tumor, with bulging growths. This could be due to the appearance of colonies within the initial “primary” colony, which outgrow the primary colony because they have somehow acquired a growth advantage. In fact, all the “outgrowths” were yellowish in color indicating that they do not ferment maltose (or to a lesser extent), and one could hypothesize that they have a lower rate of translocation of the MBP mutants, which create periplasmic stress. The colonies expressing WT-MBP were still circular, but no longer red, which is probably due to the exhaustion of maltose from the plates, and the conversion to peptones as a carbon-source. Another interesting observation was made at a growth temperature of 30°C. When the colonies expressing the various cysteine mutants were grown for an extra day at 30°C, the level of maltose fermentation was similar to the case of WT-MBP, and the colonies retained their circular shape. This could be explained by the propensity of the various mutants to cause periplasmic stress at 37°C, but not at 30°C, and that either by aggregating in the periplasm, or by occupying the chaperone machinery that deals with misfolded periplasmic proteins. The experiment was also carried out at 25°C, which was the maximal temperature at which the *in vitro* work was carried out, but results were similar to the one conducted at 30°C.

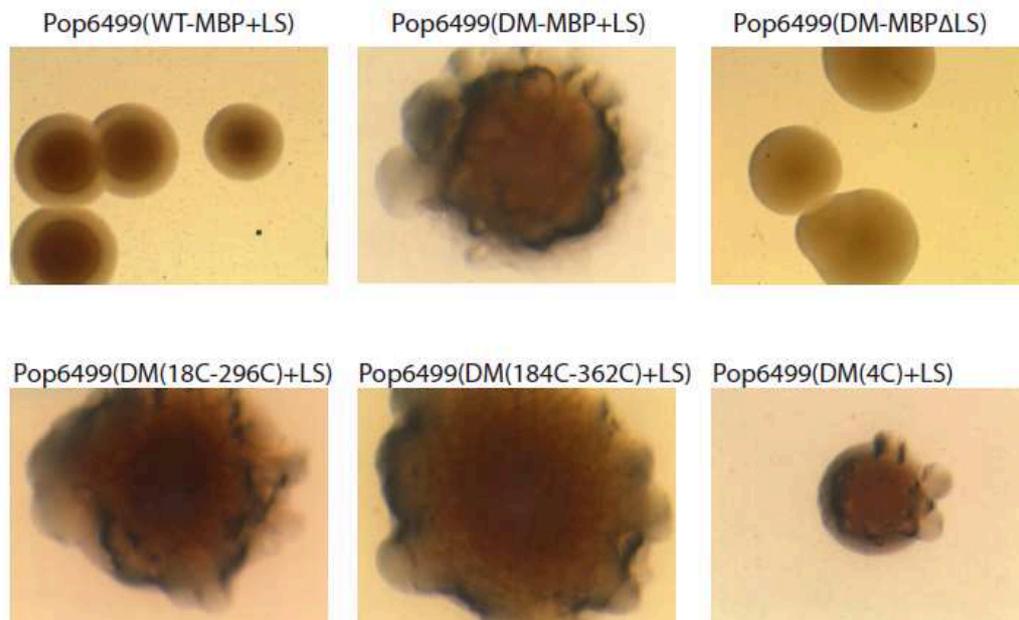
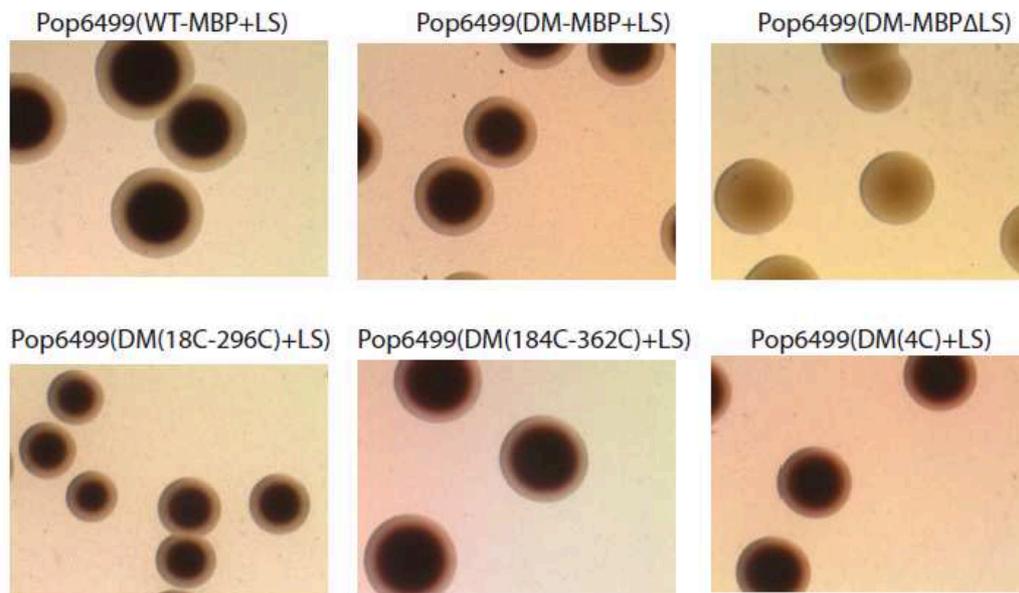
A Growth temperature: 37°C**B** Growth temperature: 30°C

Figure 37: Comparison of growth at 30 and 37°C 48 hours post-plating. (A and B) Cells were transformed with the corresponding plasmids, and grown for 2 days at 30 and 37°C as indicated, on MacConkey plates containing 0.4% maltose.

To test whether the MBP cysteine mutants were aggregating in the periplasm, we have carried out cellular fractionation 480 minutes after the start of a growth curve experiment, and did Western Blot using anti-MBP antibody. Surprisingly, we found that none of the periplasmic mutants was detectable; except for cytosolic DM-MBP (lacking the leader sequence), and WT-MBP(+LS), we could not detect any mutant in the whole cell lysate. Additionally, DM-MBP(Δ LS) was not detected in the periplasm, but rather in the cytoplasmic fraction, as expected. As for WT-MBP(+LS), most of it was detected in the periplasmic fraction. It would be interesting to determine whether the various mutants are more stable, and hence detectable by Western blot at 30°C. If so, then this would explain why the MBP mutants enabled maltose fermentation to a similar extent as the WT-MBP at 30°C. A more conclusive experiment would need to be carried out in MacConkey liquid medium.

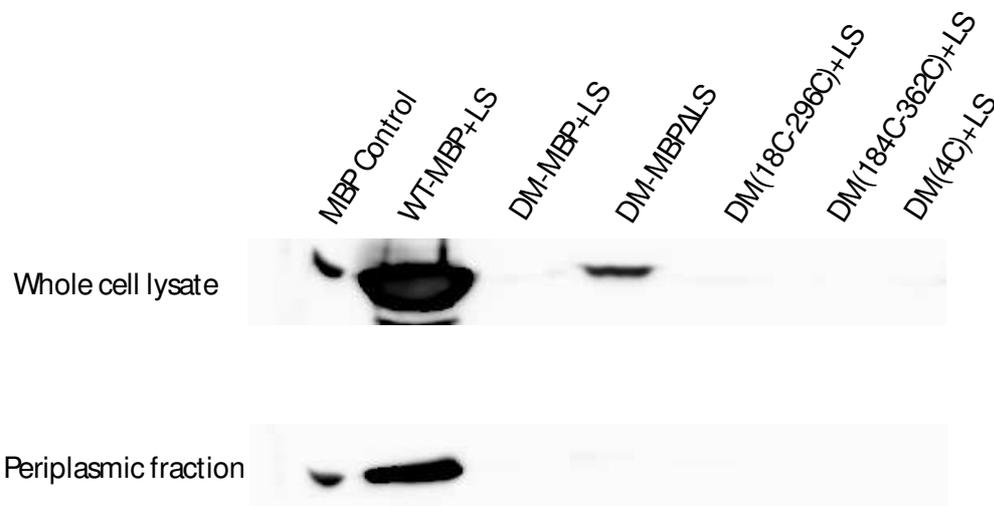


Figure 39: Fractionation and analysis of MBP-mutant localization by Western blot. Cells were harvested 8 hours after the start of the growth curves in LB+0.4% maltose, fractionated according to the protocol in Arié et al. (2006), and blotted against mouse anti-MBP antibody.

5. Discussion

In the present work, we have shown that DM-MBP is not prone to reversible aggregation in the range of concentrations used. We have also shown that DM-MBP faces mainly an entropic barrier while folding to the native state and that introducing disulfide bridges which enforce native-like contacts in the N and C domains decreased this barrier. We hypothesized that GroEL may be acting in a similar mechanism, by destabilizing the intermediate from which the refolding starts and thereby decreasing the entropic barrier. Indeed, we have shown that the effects of disulfides and folding by chaperonin were redundant, and that the rate-limiting steps of the refolding in both cases were similar, as shown by their Arrhenius plots. Additionally, we have demonstrated a critical role for charges on the chaperonin wall in accelerating the refolding. We proposed a role for those charges in ordering water molecules inside the cage, which could increase the hydrophobic interaction. The findings presented in this work highlight the uniqueness of the chaperonins in the cellular chaperone network.

5.1 DM-MBP: A slow folder without detectable aggregation

It has been proposed by Horwich and colleagues that GroEL accelerates the refolding of DM-MBP by merely preventing transient aggregation (Apetri & Horwich, 2008). We have carried out refolding experiments at different concentrations of DM-MBP, and found the rates to be independent of concentration. We also failed to detect multimers by FCS, FCCS or light scattering, and therefore rule out transient aggregation as a cause of the slow refolding of DM-MBP. We proposed that since the protein is not aggregating, GroEL/ES must be actively

accelerating the folding. Another hypothesis we drew from this data concerns the basis of slow DM-MBP folding; the protein must be populating a kinetically-trapped intermediate.

5.2 Basis of slow DM-MBP folding

By carrying out unfolding and refolding at different denaturant concentrations, we were able to show that DM-MBP populates a kinetically-trapped intermediate between 0.5-0.8 M denaturant, as shown by the hysteresis in unfolding/refolding curves, measured both by CD and Trp fluorescence. We have shown that this intermediate is similar, if not identical to the intermediate populated immediately upon dilution of the protein from denaturant into refolding conditions. Additionally, we have shown that the intermediate is collapsed but has a broader intramolecular distance distribution than the native state, as determined by spFRET. Moreover, it is structurally disordered and dynamic based on its CD, Trp fluorescence, and H/D exchange properties. The lack of ordered structure in this intermediate suggested that it is not stabilized by any native or non-native state contacts, and that it faces an entropic barrier to folding. We have indeed found a negligible dependence of the folding of DM-MBP on temperature, which indicated a small enthalpic and a substantial entropic component of the folding barrier. We have further tested this finding by introducing configurational constraints mediated by cysteine residues in positions that are juxtaposed in the native structure, but far apart along the amino acid sequence. The main effect of such cross-links is to destabilize the disordered state by decreasing its conformational entropy (Creighton, 1993). As such, constraining disulfides could accelerate the rate of folding when it is limited by conformational entropy. After having excluded potential effects of the disulfides on the stability of native and transitions states, we concluded that the refolding acceleration afforded by these was indeed due to effects on the intermediate state.

Notably, acceleration of folding was observed not only when the unfolded protein had already been oxidized, but also when the disulfide bonds were allowed to form only after the collapse reaction, which happens within milliseconds upon dilution of the unfolded protein into refolding conditions. This suggested that the respective residues are transiently proximal in the kinetically trapped intermediate. Constraining already proximal residues should accelerate the folding only if substantial flexibility exists around these regions. Introducing the disulfides restored the temperature dependence of the folding rate of DM-MBP, reflecting an increased enthalpic component, and a reduced entropic component of the folding barrier.

Additionally, we have found that the rate acceleration afforded by the disulfides depended on their location. The N-domain mutant experienced faster refolding as compared to the C-domain mutant. This might in fact be due to the distance along the primary structure, separating the 2 cysteines in each case. The farther apart in the primary structure the residues that are cross-linked, the greater the decrease in conformational entropy (Creighton, 1993). Indeed, the cysteines in the N-domain mutant, which experiences the larger rate acceleration, are separated by 277 amino acids, whereas those in the C-domain are separated by 177 amino acids. The folding of MBP was also studied by Randall and colleagues in the 1990s (Chun et al., 1993). These authors concluded that the rate-limiting step of spontaneous MBP is the same, *in vivo* and *in vitro*, based on studies on slow-folding mutants of MBP. Since all those mutants were found in the same element of structure in the N-domain, they also proposed that formation of this element of structure is rate-limiting. One could then argue that enforcing native-like contacts in the N-domain by introducing the disulfides would lead to a larger effect on the rate acceleration, as we see by comparing refoldings of both mutants.

A significant correlation between the topological complexity of proteins and their rates of folding has been shown to exist (Plaxco et al., 1998). As a quantification of topological complexity, Plaxco and colleagues have used relative contact order, (CO), which reflects the relative importance of local and non-local contacts to a protein's native structure. Relative contact order is the average sequence distance between all pairs of contacting residues normalized by the total sequence length. The larger the contact order of a protein, the more topologically complex it is, and the slower the folding rate. Indeed, the native states of topologically complex proteins are stabilized by long-range contacts, whose formation can be more effectively counteracted by conformational entropy than local contacts, leading to slower kinetics. Plaxco and colleague have also shown a correlation between transition state placement and contact order; proteins characterized by large contact orders tend to exhibit more well-ordered transition states, presumably because more of the polypeptide must be ordered to form the requisite number of favorable contacts. Accordingly, these proteins fold slower than proteins that fold via poorly packed transition states (i.e. proteins that surmount the rate-limiting step with smaller chain entropy losses). These correlations represented qualitative experimental evidence for the contribution of chain entropy loss to the free energy barrier of folding. Our findings show indeed that the topologically complex DM-MBP populates a kinetically trapped intermediate during folding, which faces a large entropic barrier and folds slowly. By introducing the disulfides, long-range native-like contacts were enforced, and the conformational entropic barrier to folding was overcome, to various extents, as seen by the temperature dependencies.

5.3 Mechanism of an active chaperonin

In order to evaluate the effects of a global confinement which we hypothesized take effect inside GroEL, we created the tetracysteine mutant, which has both its N and C domains constrained in a native-like fashion, simulating a global confinement inside the cage. Indeed we found that this protein exhibits rate acceleration to the same extent afforded by GroEL. We have moreover shown similar positive temperature dependencies of the folding in either case, and that strongly supports the assumption of a common, underlying principle in reducing the entropic activation barrier of the folding reaction. Additionally, we have shown that global confinement in the chaperonin cage accelerated folding more efficiently than the single disulfide bonds but did not enhance the faster folding rate obtained by combining the two disulfides. In other words, the accelerating effects of disulfide bonds and of chaperonin on folding were non-additive, which is indicative of the redundancy of disulfide-mediated constraints and confinement by the chaperonin in modulating the thermodynamic parameters of the folding reaction. The results presented in this work argue that the chaperonin accelerates the folding by limiting the entropic space in which the protein floats before it reaches the native state, similarly to the entropic constraints afforded by the long-range disulfide bonds (Fig. 35).

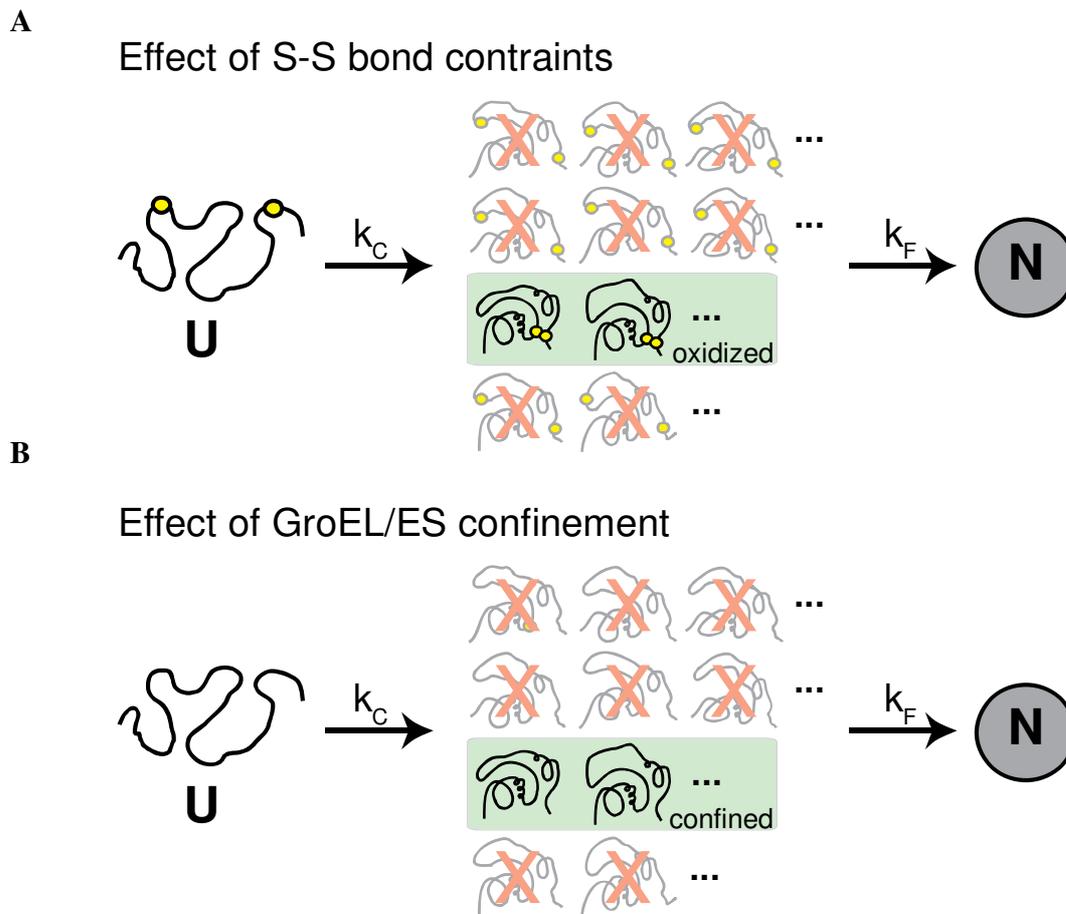


Figure 35: Schematic representations of refolding of DM-MBP with disulfide bridge-mediated restriction of conformational flexibility (A) and confinement by chaperonin (B). Upon dilution from denaturant, unfolded DM-MBP (U) undergoes rapid collapse (k_C) to an ensemble of intermediate states that must cross an entropic barrier for folding (k_F) to the native state (N). Introducing long-range disulfide bonds reduces this barrier by conformationally restricting the ensemble of kinetically trapped states to more ordered states, resulting in accelerated folding. Confinement of DM-MBP inside the chaperonin cage mimics the effect of disulfide bonds by eliminating more disordered states.

5.4 Effects of cavity charges on folding

We found that the ability of GroEL to accelerate folding was markedly dependent on the net negative charge clusters at the inner surface of the apical domains. Our lab had previously shown

that the folding of DM-MBP is impaired inside the SR-KKK2 mutant, which has a net cavity charge of zero. In this work, we further explain the mechanism behind this rate deceleration, by performing the refolding inside the KKK-2 mutant at different temperatures, and showing that it has lost its ability to entropically destabilize the substrate. The folding was in fact independent of temperature, indicating a large entropic barrier. Notably, folding of the oxidized proteins inside SR-KKK2 was not decelerated, but the rates were rather similar to folding inside the wild-type chaperonin. This might be attributed to the geometric confinement afforded by the mutant cage to the N-terminal mutant. We hypothesize that since DM-MBP's mobility inside the mutant cage is substantially decreased, it could be interacting with the walls of the uncharged cage longer than in the wild type situation. The oxidized N-terminal domain mutant however, is expected to bury its hydrophobic residues faster than DM-MBP, and would therefore interact less with the walls of the mutant cage. As such, it could be destabilized merely by the passive geometric confinement afforded by that cage, and reach the faster folding rate of the wt cage. Our findings which suggest that the KKK-2 mutant is providing DM-MBP with a passive folding environment are consistent with recent theoretical considerations, emphasizing the importance of a charged surface in inducing ordered water structure, with the resulting increase in the density of water facilitating folding by enhancing the hydrophobic effect and thus promoting global protein compaction (England & Pande, 2008; Lucent et al., 2007). Importantly, this change in solvent behavior can only take effect when the folding protein is brought into close proximity to the cavity wall. In accordance with theory and simulation (Baumketner et al., 2003; Hayer-Hartl & Minton, 2006), decreasing the size of the chaperonin cage has been shown to accelerate the folding of smaller GroEL substrates (Tang et al., 2008), suggesting that charge effects from the

cavity wall and geometric confinement act in concert to smooth the folding energy landscape (Fig. 36).

Folding funnel

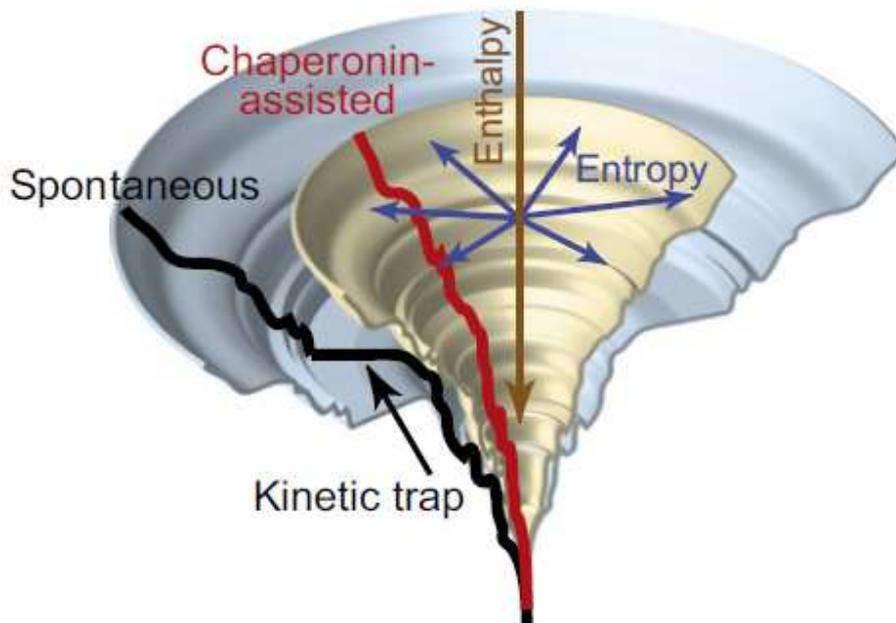


Figure 36: Model for accelerated folding by entropic destabilization of kinetically-trapped refolding intermediate. The hypothetical folding funnel of DM-MBP is shown in light blue with the vertical axis denoting enthalpy (arrow to lower enthalpy) and the radial plane denoting entropy of the folding intermediates (arrow to higher entropy). The large flat area in the folding funnel denotes the kinetically trapped intermediate that is separated from the native state through an entropic barrier. This is modified inside the chaperonin cage (light gray funnel), resulting in a more downhill funnel without the presence of prominent isoenthalpic regions representing entropically stabilized intermediate states.

5.5 Substrate unfolding in the folding mechanism

The cage-mediated acceleration of folding by GroEL/ES described in this work is mechanistically independent of repeated cycles of substrate binding and release from the

hydrophobic apical domains of GroEL, as shown by the rate acceleration afforded by the non-cycling single chaperonin. As mentioned in the introduction, substrate cycling has been proposed to accelerate folding by iteratively unfolding kinetically trapped intermediates stabilized by non-native interactions, allowing repartitioning to a productive folding pathway upon release (iterative annealing) (Shtilerman et al., 1999; Thirumalai & Lorimer, 2001). Indeed, in the context where ATP is added after substrate, ATP-dependent apical domain movements could cause local structural expansion (Lin et al., 2008; Sharma et al., 2008), but in the case of DM-MBP and Rubisco, such “forced” unfolding was dispensable for folding acceleration (Brinker et al., 2001; Sharma et al., 2008). Additionally, unfolding was observed upon initial binding of the substrate to chaperonin, and was proposed to play a role in accelerating the refolding of Rubisco (Lin et al., 2008). Nevertheless, the experimental evidence for this was only obtained in an artificial setting, namely by extending the incubation time of SR-bound protein complex before initiating the refolding. For DM-MBP, initial stretching was observed, but its functionality in accelerating the folding could not be demonstrated (Sharma et al., 2008). Our finding that the kinetically trapped folding intermediate of DM-MBP is highly disordered and thus unlikely to contain strong non-native contacts therefore provides indirect evidence that the initial unfolding is dispensable. Additionally, we have monitored FRET distributions of DM-MBP bound to SR, in the presence and absence of the crowding agent Ficoll, and found a substantial increase in the fraction of substrate in the extended state. We nevertheless could not detect a corresponding increase in the folding rate, which one would expect if unfolding upon binding had been a crucial aspect of rate acceleration.

5.6 Towards an understanding of periplasmic protein folding *in vivo*

To confirm our *in vitro* results, we undertook experiments to determine whether the various MBP mutants do experience rate acceleration when placed in an oxidizing environment. However, we were unable to detect any growth advantages conferred by any of those mutants as compared to the WT-MBP, but rather the contrary. Upon growth on 0.4% maltose-containing MacConkey plates at 37°C, all the mutants exhibited smaller-sized colonies as compared to either periplasmic WT-MBP, or cytosolic DM-MBP, which seems to suggest that the mutant proteins cause periplasmic stress; they could therefore be used to study the heat stress response in the periplasm, which is not as well characterized as the cytoplasmic one. Surprisingly, we were unable to detect any of the mutants by Western blot at 37°C, and one might argue that they were either degraded, or were not expressed to start with. Nevertheless, all clones were able to ferment maltose to various extents upon plating on MacConkey agar at 30 and 37°C, which argues against the latter possibility. A more quantitative approach might be to clone the mutants between an N-terminal signal recognition particle (SRP)-dependent signal peptide and a C-terminal selectable marker, TEM- β -lactamase, as was done by Mansell et al. in 2010. The DsbA signal peptide would enable cotranslational translocation across the cytoplasmic membrane via the SRP-dependent pathway, and nascent polypeptides would enter the periplasm where they either fold properly, conferring Ampicillin resistance to cells, or misfold, leading to eventual aggregation and/or proteolysis (Mansell et al., 2010).

Given that the responses that the periplasmic DM-MBP cysteine mutants elicited at various temperatures were identical to that of DM-MBP, we were unable to confirm our initial

hypothesis on improved folding in the presence of disulfides; this could in fact be due to the mutants not being recognized by the disulfide bond formation (DSB) system in the periplasm. Indeed, the precise mechanism of substrate binding is not clear (Ito & Inaba, 2008), and it might therefore be the case that the mere presence of cysteines is not enough. Nevertheless, it seems plausible that these mutants could be used to study the periplasmic stress response, and it would be interesting to test whether any of the periplasmic chaperones or proteases is upregulated in response to their expression at 37°C.

5.7 Biological significance of an active mechanism

The GroEL/ES system is essential under all growth conditions and is normally utilized by ~10% of newly synthesized cytosolic proteins, including ~80 proteins that are predicted to be chaperonin dependent for folding (Ewalt et al., 1997; Houry et al., 1999; Kerner et al., 2005). Like DM-MBP, these proteins have complex alpha and beta domain topologies and are thought to populate kinetically trapped folding intermediates (Kerner et al., 2005). In view of the fact that cells contain multiple, partially redundant chaperone systems for aggregation prevention, the ability to actively promote the folding of such intermediates would explain the uniquely essential role of the chaperonin cages. On the other hand, the conspicuous absence of chaperonins from oxidizing compartments correlates with the role of disulfide bond formation in providing an alternative mechanism to lower entropic folding barriers.

Another essential aspect of the biological significance of chaperonins is their role as capacitors of evolution. It was recently shown that GroEL/ES was able to rescue stability-impaired mutants by increasing the fraction of soluble, functional enzyme, and therefore had marked effects on evolutionary rates, and allowed divergence of new function (Tokuriki &

Tawfik, 2009). Moreover, a comparative study relying on 446 sequenced proteobacterial genomes has shown that greater GroEL-dependency entails increased evolutionary rates with GroEL obligatory proteins (class III) evolving on average up to 15% faster than GroEL partially dependent proteins (class II) and 35% faster than GroEL independent proteins (class I) (Bogumil & Dagan, 2010).

In conclusion, we propose that the capacity of the chaperonin cage to overcome entropic barriers in the folding energy landscape serves to promote the folding of a restricted set of obligate substrates and may also endow GroEL/ES with the ability to buffer mutations that kinetically disable the folding pathways of otherwise chaperonin-independent proteins.

6. References

Abe, J., Shibata, Y., Fujisue, M., & Hizukuri, S. (1996). Expression of periplasmic alpha-amylase of *Xanthomonas campestris* K-11151 in *Escherichia coli* and its action on maltose. *Microbiology (Reading, England)*, *142* (Pt 6, 1505-12. Retrieved May 16, 2011, from <http://www.ncbi.nlm.nih.gov/pubmed/8704990>.

Ando, T., & Skolnick, J. (2010). Crowding and hydrodynamic interactions likely dominate *in vivo* macromolecular motion. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(43). doi: 10.1073/pnas.1011354107.

Anfinsen, C. B. (1973). Principles that govern the folding of protein chains. *Science (New York, N.Y.)*, *181*(96), 223-30. The American Association for the Advancement of Science. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/4124164>.

Anfinsen, C. B., & Scheraga, A. (1975). Experimental and theoretical aspects of protein folding. *Advances in Protein Chemistry*, *29*, 205-300.

Apetri, A. C., & Horwich, A.L. (2008). Chaperonin chamber accelerates protein folding through passive action of preventing aggregation. *Proceedings of the National Academy of Sciences*, *105*(45), 17351. National Acad Sciences. Retrieved March 10, 2011, from <http://www.pnas.org/content/105/45/17351.full>.

Arié, J.-P., Miot, M., Sassoon, N., & Betton, J.-M. (2006). Formation of active inclusion bodies in the periplasm of *Escherichia coli*. *Molecular microbiology*, *62*(2), 427-37. doi: 10.1111/j.1365-2958.2006.05394.x.

Barlow, D. J., & Thornton, J. M. (1983). Ion-pairs in proteins. *Journal of molecular biology*, *168*(4), 867-85. Retrieved February 1, 2011, from <http://www.ncbi.nlm.nih.gov/pubmed/6887253>.

Bartlett, A. I., & Radford, S. E. (2009). An expanding arsenal of experimental methods yields an explosion of insights into protein folding mechanisms. *Nature structural & molecular biology*, *16*(6), 582-8. doi: 10.1038/nsmb.1592.

Baumketner, A., Jewett, A., & Shea, J. E. (2003). Effects of confinement in chaperonin assisted protein folding: rate enhancement by decreasing the roughness of the folding energy landscape. *Journal of Molecular Biology*, *332*(3), 701-713. Retrieved from <http://linkinghub.elsevier.com/retrieve/pii/S002228360300929X>.

Bicout, D. J., & Szabo, A. (2000). Entropic barriers, transition states, funnels, and exponential protein folding kinetics: a simple model. *Protein Science*, *9*(3), 452-465. Cold Spring Harbor Laboratory Press. Retrieved from

- <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2144570&tool=pmcentrez&rendertype=abstract>.
- Bochkareva, E. S., Lissin, N. M., & Girshovich, A. S. (1988). Transient association of newly synthesized unfolded proteins with the heat-shock GroEL protein. *Nature*, *336*(6196), 254-7. doi: 10.1038/336254a0.
- Bogumil, D., & Dagan, T. (2010). Chaperonin-dependent accelerated substitution rates in prokaryotes. *Genome biology and evolution*, *2*(Class I), 602-608. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/20660111>.
- Braig, K., Otwinowski, Z., Hegde, R., Boisvert, D. C., Joachimiak, A., Horwich, A L, et al. (1994). The crystal structure of the bacterial chaperonin GroEL at 2.8 Å. *Nature*, *371*(6498), 578-86. doi: 10.1038/371578a0.
- Brandt, F., Etchells, S. A., Ortiz, J. O., Elcock, A. H., Hartl, F Ulrich, & Baumeister, W. (2009). The native 3D organization of bacterial polysomes. *Cell*, *136*(2), 261-271. Elsevier Inc. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/19167328>.
- Brinker, A., Pfeifer, G., Kerner, M J, Naylor, D J, Hartl, F U, & Hayer-Hartl, M. (2001). Dual function of protein confinement in chaperonin-assisted protein folding. *Cell*, *107*(2), 223-33. Retrieved February 28, 2011, from <http://www.ncbi.nlm.nih.gov/pubmed/11672529>.
- Buchner, J. (2010). Bacterial Hsp90--desperately seeking clients. *Molecular microbiology*, *76*(3), 540-4. doi: 10.1111/j.1365-2958.2010.07140.x.
- Camacho, C. J., & Thirumalai, D. (1995). Theoretical predictions of folding pathways by using the proximity rule, with applications to bovine pancreatic trypsin inhibitor. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(5), 1277-1281. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=42502&tool=pmcentrez&rendertype=abstract>.
- Chaudhry, C., Farr, George W, Todd, M. J., Rye, Hays S, Brunger, A. T., Adams, P. D., et al. (2003a). Role of the γ -phosphate of ATP in triggering protein folding by GroEL±GroES: function, structure and energetics. *EMBO Journal*, *22*(19), 4877-4887.
- Chaudhry, C., Farr, George W, Todd, M. J., Rye, Hays S, Brunger, A. T., Adams, P. D., et al. (2003b). Role of the γ -phosphate of ATP in triggering protein folding by GroEL–GroES: function, structure and energetics. *the The European Molecular Biology Organization Journal*, *22*(19), 4877-4887. Oxford University Press. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=204461&tool=pmcentrez&rendertype=abstract>.
- Chen, S., Roseman, A. M., Hunter, A. S., Wood, S. P., Burston, S G, Ranson, N A, et al. (1994). Location of a folding protein and shape changes in GroEL-GroES complexes imaged by cryo-electron microscopy. *Nature*, *371*(6494), 261-4. doi: 10.1038/371261a0.

- Cheng, M. Y., Hartl, F U, Martin, J., Pollock, R. A., Kalousek, F., Neupert, W., et al. (1989). Mitochondrial heat-shock protein hsp60 is essential for assembly of proteins imported into yeast mitochondria. *Nature*, 337(6208), 620-625. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/2645524>.
- Chun, S. Y., Strobel, S., Bassford, P., & Randall, L. L. (1993). Folding of maltose-binding protein. Evidence for the identity of the rate-determining step *in vivo* and *in vitro*. *The Journal of Biological Chemistry*, 268(28), 20855-20862. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8407916>.
- Cliff, M. J., Limpkin, C., Cameron, A., Burston, Steven G, & Clarke, Anthony R. (2006a). Elucidation of steps in the capture of a protein substrate for efficient encapsulation by GroE. *The Journal of biological chemistry*, 281(30), 21266-75. doi: 10.1074/jbc.M601605200.
- Cliff, M. J., Limpkin, C., Cameron, A., Burston, Steven G, & Clarke, Anthony R. (2006b). Elucidation of steps in the capture of a protein substrate for efficient encapsulation by GroE. *The Journal of biological chemistry*, 281(30), 21266-75. doi: 10.1074/jbc.M601605200.
- Creighton, T. E. (1993). *Proteins: Structures and Molecular Properties* (Second.). W.H. Freeman. Retrieved March 11, 2011, from <http://www.amazon.com/Proteins-Structures-Properties-Thomas-Creighton/dp/0716723174>.
- Daggett, Valerie, & Fersht, A. (2003). The present view of the mechanism of protein folding. *Nature reviews. Molecular cell biology*, 4(6), 497-502. Nature Publishing Group. doi: 10.1038/nrm1126.
- Dill, K A, & Chan, H. S. (1997). From Levinthal to pathways to funnels. *Nature structural biology*, 4(1), 10-9. New York, NY: Nature Pub. Co., c1994-c2003. doi: 10.1038/nsb0197-10.
- Dill, Ken A. (1990). Dominant forces in protein folding. *Biochemistry*, 29(31), 7133-7155. doi: 10.1021/bi00483a001.
- Dill, K. a, Ozkan, S. B., Shell, M. S., & Weikl, T. R. (2008). The protein folding problem. *Annual review of biophysics*, 37, 289-316. doi: 10.1146/annurev.biophys.37.092707.153558.
- Dill, K. a, Ozkan, S. B., Weikl, T. R., Chodera, J. D., & Voelz, V. a. (2007). The protein folding problem: when will it be solved? *Current opinion in structural biology*, 17(3), 342-6. doi: 10.1016/j.sbi.2007.06.001.
- Duplay, P., Bedouelle, H., Fowler, A., Zabin, I., Saurin, W., & Hofnung, M. (1984). Sequences of the malE gene and of its product, the maltose-binding protein of *Escherichia coli* K12. *The Journal of Biological Chemistry*, 259(16), 10606-10613. ASBMB. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/6088507>.

- Ellis, R J, & Hartl, F U. (1996). Protein folding in the cell: competing models of chaperonin function. *The FASEB journal official publication of the Federation of American Societies for Experimental Biology*, 10(1), 20-26. Retrieved from <http://www.fasebj.org/cgi/reprint/10/1/20>.
- Ellis, R John, & Minton, Allen P. (2006). Protein aggregation in crowded environments. *Biological chemistry*, 387(5), 485-97. doi: 10.1515/BC.2006.064.
- England, J. L., & Pande, V. S. (2008). Potential for modulation of the hydrophobic effect inside chaperonins. *Biophysical Journal*, 95(7), 3391-3399. The Biophysical Society. Retrieved from <http://arxiv.org/abs/0802.0522>.
- Farr, George W, Furtak, Krystyna, Rowland, M. B., Ranson, Neil A, Saibil, Helen R, Kirchhausen, T., et al. (2000). Multivalent Binding of Nonnative Substrate Proteins by the Chaperonin GroEL. *Cell*, 100, 561-573.
- Fayet, O., Ziegelhoffer, T., & Georgopoulos, C. (1989). The groES and groEL heat shock gene products of *Escherichia coli* are essential for bacterial growth at all temperatures. *Journal Of Bacteriology*, 171(3), 1379-1385. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=2563997.
- Fenton, W A, Kashi, Y., Furtak, K, & Horwich, A L. (1994). Residues in chaperonin GroEL required for polypeptide binding and release. *Nature*, 371(6498), 614-9. doi: 10.1038/371614a0.
- Ferbitz, L., Maier, Timm, Patzelt, H., Bukau, Bernd, Deuerling, E., & Ban, N. (2004). Trigger factor in complex with the ribosome forms a molecular cradle for nascent proteins. *Nature*, 431(7008), 590-6. doi: 10.1038/nature02899.
- Georgopoulos, C. P., Hendrix, R. W., Casjens, S. R., & Kaiser, A. D. (1973). Host participation in bacteriophage lambda head assembly. *Journal of molecular biology*, 76(1), 45-60. Retrieved February 18, 2011, from <http://www.ncbi.nlm.nih.gov/pubmed/4578100>.
- Goloubinoff, P., Christeller, J. T., Gatenby, A. A., & Lorimer, G. H. (1989). Reconstitution of active dimeric ribulose biphosphate carboxylase from an unfoleded state depends on two chaperonin proteins and Mg-ATP. *Nature*, 342(6252), 884-9. doi: 10.1038/342884a0.
- Goloubinoff, P., Gatenby, A. A., & Lorimer, G. H. (1989). GroE heat-shock proteins promote assembly of foreign prokaryotic ribulose biphosphate carboxylase oligomers in *Escherichia coli*. *Nature*, 337(6202), 44-47. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/2562907>.
- Hartl, F Ulrich, & Hayer-Hartl, Manajit. (2009). Converging concepts of protein folding *in vitro* and *in vivo*. *Nature structural & molecular biology*, 16(6), 574-81. doi: 10.1038/nsmb.1591.
- Hayer-Hartl, M. K., Weber, F., & Hartl, F U. (1996). Mechanism of chaperonin action: GroES binding and release can drive GroEL-mediated protein folding in the absence of ATP hydrolysis. *the The European Molecular Biology Organization Journal*, 15(22), 6111-6121.

- Hayer-Hartl, Manajit, & Minton, Allen P. (2006). A simple semiempirical model for the effect of molecular confinement upon the rate of protein folding. *Biochemistry*, 45(44), 13356-13360. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/17073456>.
- Horwich, Arthur L., & Fenton, Wayne A. (2009). Chaperonin-mediated protein folding: using a central cavity to kinetically assist polypeptide chain folding. *Quarterly reviews of biophysics*, 42(2), 83-116. doi: 10.1017/S0033583509004764.
- Ito, K., & Inaba, K. (2008). The disulfide bond formation (Dsb) system. *Current opinion in structural biology*, 18(4), 450-8. doi: 10.1016/j.sbi.2008.02.002.
- Jiao, W., Qian, M., Li, P., Zhao, L., & Chang, Z. (2005). The essential role of the flexible termini in the temperature-responsiveness of the oligomeric state and chaperone-like activity for the polydisperse small heat shock protein IbpB from *Escherichia coli*. *Journal of molecular biology*, 347(4), 871-84. doi: 10.1016/j.jmb.2005.01.029.
- Kaiser, C. M., Chang, H.-C., Agashe, V. R., Lakshmipathy, S. K., Etchells, S. A., Hayer-Hartl, Manajit, et al. (2006). Real-time observation of trigger factor function on translating ribosomes. *Nature*, 444(7118), 455-60. Nature Publishing Group. doi: 10.1038/nature05225.
- Kapanidis, A. N., Lee, N. K., Laurence, T. A., Doose, S., Margeat, E., & Weiss, S. (2004). Fluorescence-aided molecule sorting: analysis of structure and interactions by alternating-laser excitation of single molecules. *Proceedings of the National Academy of Sciences of the United States of America*, 101(24), 8936-8941. National Academy of Sciences. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=428450&tool=pmcentrez&rendertype=abstract>.
- Kelly, S. M., Jess, T. J., & Price, N. C. (2005). How to study proteins by circular dichroism. *Biochimica et biophysica acta*, 1751(2), 119-39. doi: 10.1016/j.bbapap.2005.06.005.
- Kerner, Michael J, Naylor, Dean J, Ishihama, Y., Maier, Tobias, Chang, H.-C., Stines, A. P., et al. (2005). Proteome-wide analysis of chaperonin-dependent protein folding in *Escherichia coli*. *Cell*, 122(2), 209-220. Elsevier. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/16051146>.
- Kirstein, J., Molière, N., Dougan, D. A., & Turgay, K. (2009). Adapting the machine: adaptor proteins for Hsp100/Clp and AAA+ proteases. *Nature reviews. Microbiology*, 7(8), 589-99. Nature Publishing Group. doi: 10.1038/nrmicro2185.
- Kreuzer, K. N., & Jongeneel, C. V. (1983). *Escherichia coli* phage T4 topoisomerase. *Methods in Enzymology*, 100, 144-160.
- Lakowicz, J. R. (2006). *Principles of Fluorescence Spectroscopy* (Third.). Springer. Retrieved February 10, 2011, from <http://www.amazon.com/Principles-Fluorescence-Spectroscopy-Joseph-Lakowicz/dp/0387312781>.

- Lakshmipathy, S. K., Tomic, S., Kaiser, C. M., Chang, H.-C., Genevoux, P., Georgopoulos, Costa, et al. (2007). Identification of nascent chain interaction sites on trigger factor. *The Journal of biological chemistry*, 282(16), 12186-93. doi: 10.1074/jbc.M609871200.
- Landry, S. J., Zeilstra-Ryalls, J., Fayet, O., Georgopoulos, C., & Gierasch, L. M. (1993). Characterization of a functionally important mobile domain of GroES. *Nature*, 364(6434), 255-8. doi: 10.1038/364255a0.
- Lin, Z., Madan, D., & Rye, Hays S. (2008). GroEL stimulates protein folding through forced unfolding. *Nature Structural & Molecular Biology*, 15(3), 303-311. doi: 10.1038/nsmb.1394.
- Lodish, H., Berk, A., Matsudaira, P., Kaiser, C. A., Krieger, M., Scott, M. P., et al. (2003). *Molecular Cell Biology* (Fifth.). W. H. Freeman. Retrieved February 10, 2011, from <http://www.amazon.com/Molecular-Cell-Biology-Harvey-Lodish/dp/0716743663>.
- Lucent, D., Vishal, V., & Pande, V. S. (2007). Protein folding under confinement: A role for solvent. *Proceedings of the National Academy of Sciences of the United States of America*, 104(25), 10430-10434. National Academy of Sciences. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/17563390>.
- Maier, C. S., & Deinzer, M. L. (2005). Protein conformations, interactions, and H/D exchange. *Methods in enzymology*, 402(05), 312-60. doi: 10.1016/S0076-6879(05)02010-0.
- Makhatadze, G. I., & Privalov, P. L. (1995). Energetics of protein structure. *Advances in protein chemistry*, 47, 307-425. Elsevier. Retrieved February 3, 2011, from <http://www.ncbi.nlm.nih.gov/pubmed/8561051>.
- Mamathambika, B. S., & Bardwell, J. C. (2008). Disulfide-linked protein folding pathways. *Annual Review of Cell and Developmental Biology*, 24(1), 211-235. Annual Reviews. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/18588487>.
- Mansell, T. J., Linderman, S. W., Fisher, A. C., & DeLisa, M. P. (2010). A rapid protein folding assay for the bacterial periplasm. *Protein science : a publication of the Protein Society*, 19(5), 1079-90. doi: 10.1002/pro.388.
- Martin, J., & Hartl, F U. (1997). The effect of macromolecular crowding on chaperonin-mediated protein folding. *Proceedings of the National Academy of Sciences of the United States of America*, 94(4), 1107-12. Retrieved February 25, 2011, from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=19752&tool=pmcentrez&rendertype=abstract>.
- Martin, J., Langer, T., Boteva, R., Schramel, A., Horwich, A L, & Hartl, F U. (1991). Chaperonin-mediated protein folding at the surface of groEL through a "molten globule"-like intermediate. *Nature*, 352(6330), 36-42.

- Mayer, M P, Rüdiger, S., & Bukau, B. (2000). Molecular basis for interactions of the DnaK chaperone with substrates. *Biological chemistry*, 381(9-10), 877-85. Walter de Gruyter und Co., Genthiner Str. 13 Berlin 10785 Germany. doi: 10.1515/BC.2000.109.
- Murzin, a G., Brenner, S. E., Hubbard, T., & Chothia, C. (1995). SCOP: a structural classification of proteins database for the investigation of sequences and structures. *Journal of molecular biology*, 247(4), 536-40. doi: 10.1006/jmbi.1995.0159.
- Müller, B. K., Zaychikov, E., Bräuchle, C., & Lamb, D. C. (2005). Pulsed interleaved excitation. *Biophysical Journal*, 89(5), 3508-3522. Biophysical Society. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1366845&tool=pmcentrez&rendertype=abstract>.
- Orengo, C. (1994). Classification of protein folds. *Current Opinion in Structural Biology*, 4(3), 429-440. doi: 10.1016/S0959-440X(94)90113-9.
- Ostermann, J., Horwich, A L, Neupert, W., & Hartl, F U. (1989). Protein folding in mitochondria requires complex formation with hsp60 and ATP hydrolysis. *Nature*, 341(6238), 125-130.
- Ozkan, S. B., Wu, G. A., Chodera, J. D., & Dill, Ken A. (2007). Protein folding by zipping and assembly. *Proceedings of the National Academy of Sciences of the United States of America*, 104(29), 11987-11992. National Academy of Sciences. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/17620603>.
- Pace, C. N., Shirley, B. A., McNutt, M., & Gajiwala, K. (1996). Forces contributing to the conformational stability of proteins. *The FASEB journal official publication of the Federation of American Societies for Experimental Biology*, 10(1), 75-83. FASEB. Retrieved from <http://www.fasebj.org/cgi/content/abstract/10/1/75>.
- Plaxco, K. W., Simons, K. T., & Baker, D. (1998). Contact order, transition state placement and the refolding rates of single domain proteins. *Journal of molecular biology*, 277(4), 985-94. Elsevier. doi: 10.1006/jmbi.1998.1645.
- Qu, Y., & Bolen, D. W. (2003). Hydrogen exchange kinetics of RNase A and the urea:TMAO paradigm. *Biochemistry*, 42(19), 5837-5849. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12741842>.
- Ramachandran, G. N., Ramakrishnan, C., & Sasisekharan, V. (1963). Stereochemistry of polypeptide chain configurations. *Journal of molecular biology*, 7(1), 95-9. Academic Press Inc. (London) Ltd. Retrieved February 1, 2011, from <http://www.ncbi.nlm.nih.gov/pubmed/13990617>.
- Ranson, N a, Farr, G W, Roseman, a M., Gowen, B., Fenton, W. a, Horwich, a L., et al. (2001). ATP-bound states of GroEL captured by cryo-electron microscopy. *Cell*, 107(7), 869-79. Retrieved February 23, 2011, from <http://www.ncbi.nlm.nih.gov/pubmed/11779463>.

- Ranson, Neil a, Clare, D. K., Farr, George W, Houldershaw, D., Horwich, Arthur L, & Saibil, Helen R. (2006). Allosteric signaling of ATP hydrolysis in GroEL-GroES complexes. *Nature structural & molecular biology*, 13(2), 147-52. doi: 10.1038/nsmb1046.
- Ratajczak, E., Stróżecka, J., Matuszewska, M., Zietkiewicz, S., Kuczyńska-Wiśnik, D., Laskowska, E., et al. (2010). IbpA the small heat shock protein from *Escherichia coli* forms fibrils in the absence of its cochaperone IbpB. *FEBS letters*, 584(11), 2253-7. doi: 10.1016/j.febslet.2010.04.060.
- Royer, C. a. (2006). Probing protein folding and conformational transitions with fluorescence. *Chemical reviews*, 106(5), 1769-84. doi: 10.1021/cr0404390.
- Rüdiger, S., Buchberger, A., & Bukau, B. (1997). Interaction of Hsp70 chaperones with substrates. *Nature Structural Biology*, 4(5), 342-349. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=9145101.
- Saibil, H R, Zheng, D., Roseman, A. M., Hunter, A. S., Watson, G. M., Chen, S., et al. (1993). ATP induces large quaternary rearrangements in a cage-like chaperonin structure. *Current biology : CB*, 3(5), 265-73. Retrieved February 28, 2011, from <http://www.ncbi.nlm.nih.gov/pubmed/15335746>.
- Sato, T., Minagawa, S., Kojima, E., Okamoto, N., & Nakamoto, H. (2010). HtpG, the prokaryotic homologue of Hsp90, stabilizes a phycobilisome protein in the cyanobacterium *Synechococcus elongatus* PCC 7942. *Molecular microbiology*, 76(3), 576-89. doi: 10.1111/j.1365-2958.2010.07139.x.
- Schlecht, R., Erbse, A. H., Bukau, Bernd, & Mayer, Matthias P. (2011). Mechanics of Hsp70 chaperones enables differential interaction with client proteins. *Nature structural & molecular biology*, 18(3), 345-351. Nature Publishing Group. doi: 10.1038/nsmb.2006.
- Schneider, E., Freundlieb, S., Tapio, S., & Boos, W. (1992). Molecular characterization of the MalT-dependent periplasmic alpha-amylase of *Escherichia coli* encoded by malS. *The Journal of biological chemistry*, 267(8), 5148-54. Retrieved May 16, 2011, from <http://www.ncbi.nlm.nih.gov/pubmed/1544897>.
- Schuler, B., & Eaton, W. a. (2008). Protein folding studied by single-molecule FRET. *Current opinion in structural biology*, 18(1), 16-26. doi: 10.1016/j.sbi.2007.12.003.
- Sharma, S., Chakraborty, K., Müller, B. K., Astola, N., Tang, Y.-chi, Lamb, D. C., et al. (2008). Monitoring protein conformation along the pathway of chaperonin-assisted folding. *Cell*, 133(1), 142-53. doi: 10.1016/j.cell.2008.01.048.
- Shtilerman, M., Lorimer, G. H., & Englander, S. W. (1999). Chaperonin function: folding by forced unfolding. *Science*, 284(5415), 822-825. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10221918>.

- Sigler, P B, Xu, Z., Rye, H S, Burston, S G, Fenton, W A, & Horwich, A L. (1998). Structure and function in GroEL-mediated protein folding. *Annual Review of Biochemistry*, 67, 581-608. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9759498>.
- Slotboom, D. J., Duurkens, R. H., Olieman, K., & Erkens, G. B. (2008). Static light scattering to characterize membrane proteins in detergent solution. *Methods (San Diego, Calif.)*, 46(2), 73-82. Elsevier Inc. doi: 10.1016/j.ymeth.2008.06.012.
- Sreerama, N., & Woody, R. W. (2004). Computation and analysis of protein circular dichroism spectra. *Methods in Enzymology*, 383(1), 318-351. Elsevier. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/15063656>.
- Tang, Y.-C., Chang, H.-C., Roeben, A., Wischnewski, D., Wischnewski, N., Kerner, Michael J, et al. (2006). Structural features of the GroEL-GroES nano-cage required for rapid folding of encapsulated protein. *Cell*, 125(5), 903-14. doi: 10.1016/j.cell.2006.04.027.
- Taniguchi, M., Yoshimi, T., Hongo, K., Mizobata, T., & Kawata, Y. (2004). Stopped-flow fluorescence analysis of the conformational changes in the GroEL apical domain: relationships between movements in the apical domain and the quaternary structure of GroEL. *The Journal of biological chemistry*, 279(16), 16368-76. doi: 10.1074/jbc.M311806200.
- Thirumalai, D., & Lorimer, G. H. (2001). Chaperonin-mediated protein folding. *Annual Review of Biophysics and Biomolecular Structure*, 30, 245-269. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11340060>.
- Tokuriki, N., & Tawfik, D. S. (2009). Chaperonin overexpression promotes genetic variation and enzyme evolution. *Nature*, 459(7247), 668-673. Nature Publishing Group. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/19494908>.
- Voelz, V. A., & Dill, Ken A. (2007). Exploring Zipping and Assembly as a Protein Folding Principle. *Bioinformatics*, 888(December 2006), 877-888. doi: 10.1002/prot.
- Voet, D., Voet, J. G., & Pratt, C. W. (2006). *Voet, Voet, Pratt: Fundamentals of Biochemistry: Life at the Molecular Level, 2nd Edition - Student Companion Site* (Second.). Retrieved February 10, 2011, from <http://bcs.wiley.com/he-bcs/Books?action=index&itemId=0471214957&itemTypeId=BKS&bcsId=2261>.
- Wales, T. E., & Engen, J. R. (2006). Hydrogen exchange mass spectrometry for the analysis of protein dynamics. *Mass Spectrometry Reviews*, 25(1), 158-170. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/16208684>.
- Weissman, J. S., Rye, H S, Fenton, W A, & And A L Horwich, J. M. B. (1996). Characterization of the active intermediate of a GroEL-GroES-mediated protein folding reaction. *Cell*, 84(3), 481-490.

- William Wilson, W. (2003). Light scattering as a diagnostic for protein crystal growth—A practical approach. *Journal of Structural Biology*, 142(1), 56-65. doi: 10.1016/S1047-8477(03)00038-8.
- Xu, Z., Horwich, A L., & Sigler, P B. (1997). The crystal structure of the asymmetric GroEL-GroES-(ADP)₇ chaperonin complex. *Nature*, 388(6644), 741-50. doi: 10.1038/41944.
- Yifrach, O., & Horovitz, A. (1995). Nested cooperativity in the ATPase activity of the oligomeric chaperonin GroEL. *Biochemistry*, 34(16), 5303-5308.
- Zhang, Y. H., Yan, X., Maier, C. S., Schimerlik, M. I., & Deinzer, M. L. (2001). Structural comparison of recombinant human macrophage colony stimulating factor beta and a partially reduced derivative using hydrogen deuterium exchange and electrospray ionization mass spectrometry. *Protein Science*, 10(11), 2336-2345. Cold Spring Harbor Laboratory Press.
- Zhu, X., Zhao, X., Burkholder, W. F., Gragerov, A., Ogata, C. M., Gottesman, M. E., et al. (1996). Structural analysis of substrate binding by the molecular chaperone DnaK. *Science (New York, N.Y.)*, 272(5268), 1606-14. AAAS. Retrieved February 17, 2011, from <http://www.ncbi.nlm.nih.gov/pubmed/8658133>.
- Zimmerman, S. B., & Minton, A P. (1993). Macromolecular crowding: biochemical, biophysical, and physiological consequences. *Annual Review of Biophysics and Biomolecular Structure*, 22(1), 27-65. Annual Reviews 4139 El Camino Way, P.O. Box 10139, Palo Alto, CA 94303-0139, USA. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/7688609>.
- Zou, Q., Bennion, B. J., Daggett, V, & Murphy, K. P. (2002). The molecular mechanism of stabilization of proteins by TMAO and its ability to counteract the effects of urea. *Journal of the American Chemical Society*, 124(7), 1192-1202. American Chemical Society. doi: 10.1021/ja004206b.

7. Curriculum Vitae

Personal Information

Name	Manal Chatila
Date of Birth	April 15, 1983
Place of Birth	Beirut, Lebanon
Nationality	Lebanese

Education

2007-Present	PhD at the Max Planck Institute of Biochemistry-Martinsried, Germany
2004-2006	Master's degree in Biology at the American University of Beirut-Beirut, Lebanon
2001-2004	Bachelor's degree in Biology at the American University of Beirut-Beirut, Lebanon
1986-2001	French and Lebanese Baccalaureates at Collège Protestant Français-Beirut, Lebanon

Publications

Chakraborty K., Chatila M., Sinha J., Shi Q., Poschner B., Sikor M., Jiang G., Lamb D.C., Hartl F.U., and Hayer-Hartl M. (2010). Chaperonin-catalyzed rescue of kinetically trapped states in protein folding. *Cell* 142, 112-122.

Chatila M. and Gali-Muhtasib H. (2007). Quinones as anti-angiogenic agents. *New Research on Angiogenesis Inhibitors*. Nova Science Publishers Inc., Chapter V.

El-Najjar N., Chatila M., Moukadem H., Vuorela H., Ocker M., Gandesiri M., Schneider-Stock R., and Gali-Muhtasib H. (2010). Reactive oxygen species mediate thymoquinone-induced apoptosis and activate ERK and JNK signaling. *Apoptosis* 15, 183-295.