

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

Chaperonin-Catalyzed Rescue of Kinetically
Trapped States in Protein Folding

Jyoti Sinha

aus

Patna

Bihar/India

2010

Erklärung

Diese Dissertation wurde im Sinne von § 13 Abs. 3 bzw. 4 der Promotionsordnung vom 29. Januar 1998 (in der Fassung der vierten Änderungssatzung vom 26. November 2004) von Herrn Prof. Dr. F. Ulrich Hartl betreut

Ehrenwörtliche Versicherung

Diese Dissertation wurde selbständig, ohne unerlaubte Hilfe erarbeitet.

München, den

Dissertation eingereicht am 18.11.2010

1. Gutacher Dr. F. Ulrich Hartl

2. Gutachter PD. Dr. Konstanze Winklhofer

Mündliche Prüfung am 14.12.2010

Acknowledgements

I am grateful to Dr. Manajit Hayer-Hartl and Prof. Dr. F. Ulrich Hartl for giving me the opportunity to learn and interact with some of the best minds. I would like to take this opportunity to thank them for their encouragement, faith, continual support and giving me the freedom to learn from my mistakes. Their ideas and advice have been crucial for development of this project and me as a person.

I would like to thank PD. Dr. Konstanze Winklhofer for being a co-referee of my thesis and my thesis advisory committee members Prof. Dr. Dr. Walter Neupert and Prof. Dr. Alexander Buchberger.

I would also like to thank our very competent and helpful support system Andrea, Silke and Evelyn who were always there with kind words and solutions to all my administrative and other problems making my stay in Munich smooth.

I am also grateful to Elisabeth, Bernd, Dirk, Nadine, and Emmanuel for all their logistical support and help. I am indebted to all my colleagues in the department of cellular biochemistry and specially office A-1 for making the working and learning a pleasurable experience. In particular, help and suggestions provided by Alex Hastie has been valuable to my learning curve.

I sincerely thank Dr. Yun-Chi Tang for her enthusiasm in mentoring me during initial phases of my PhD and Eva and Romy for their excellent technical assistance.

This work essentially is a result of the team work together with Manal and Kausik, with whom I have learned more about scientific temper than just technicalities. Many thanks to Bernhard Poschner, Guoxin Jiang and Martin Sikor for helpful suggestions during the course of this work.

My heartfelt thanks to the 'Bhartiye Junta' here Pratibha, Shruti, Bhumi, Juhi, Pankaj, Partho Rashmi, Sathish, Madhu, Anoop, Ranga, Naga, Kirti, Rochelle, Aarathi, Krishna, Venky, Niti and Rajat for endless parties and all the fun.

My deepest thanks to Vivek for his enormous support in this journey. And at last, what actually stands first is the unconditional love and encouragement of my parents and brothers (Deepak and Prakash) whose unwavering belief in me made it all happen.

1. Summary	1
2. Introduction	3
2.1. Protein Folding	3
2.1.1. Protein Structure	3
2.1.2. Dominant Forces in Protein Folding	6
2.1.3. Mechanism of Protein Folding	7
2.1.4. Classical View versus New View to Protein Folding Kinetics	8
2.1.5. Disulfide Bonds Mediated Protein Folding	16
2.2. Methods for Studying Protein Folding	17
2.2.1. Fluorescence	19
2.2.2. FRET	21
2.2.3. Single Molecule FRET	24
2.3. Protein Folding <i>In Vivo</i>	26
2.3.1. The Chaperone Network in the Cytosol	30
2.3.2. Ribosome-Associated Chaperone : Trigger Factor	32
2.3.3. The Bacterial Hsp70 System	34
2.3.4. The Chaperonins	36
2.4. Group I Chaperonin System: GroEL and GroES	37
2.4.1. Architecture of GroEL and GroES	37
2.4.2. Polypeptide, Nucleotide and GroES Binding	40
2.4.3. Mechanism of GroEL/ES mediated Protein Folding	44
2.4.4. Substrates of GroEL	48
3. Aim of the Study	49
4. Materials and Methods	50
4.1. Materials	50
4.1.1. Chemicals	50
4.1.2. Enzymes	53
4.1.3. Materials	53
4.1.4. Instruments	53
4.1.5. Media	55
4.1.6. Antibiotic Stock Solutions	55
4.2. Bacterial Strains and Plasmids	55
4.2.1. <i>E.coli</i> strains	55
4.2.2. Plasmids	55
4.3. Molecular Cloning Methods	58

4.3.1.	Preparation and Transformation of <i>E. coli</i> Competent Cells.....	58
4.3.2.	Plasmid Purification	59
4.3.3.	PCR Amplification	59
4.3.4.	DNA Restriction and Ligation.....	60
4.3.5.	DNA Analytical Methods.....	61
4.4.	Protein Purification.....	61
4.4.1.	GroEL Expression and Purification.....	61
4.4.2.	GroES Expression and Purification.....	62
4.4.3.	MBP and MBP Mutants Expression and Purification	63
4.4.4.	MetF Expression and Purification	64
4.4.5.	Rhodanese Preparation	64
4.5.	Protein Analytical Methods	65
4.5.1.	Determination of Protein Concentration.....	65
4.5.2.	SDS-PAGE	65
4.5.3.	Western-Blotting	66
4.6.	GroEL Functional Activity Assays.....	67
4.6.1.	ATPase Assay.....	67
4.6.2.	Aggregation Prevention Assay of Denatured Rhodanese.....	67
4.7.	In vitro Protein Refolding and Activity Assays.....	68
4.7.1.	MBP Refolding.....	68
4.7.2.	MetF Refolding.....	69
4.7.3.	Rhodanese Refolding.....	69
4.7.4.	Rubisco Refolding	70
4.8.	Biochemical and Biophysical Methods	71
4.8.1.	Thiol-Mediated Labeling of the Cys Constructs	71
4.8.2.	Characterization of Cysteine Pair Mutants of DM-MBP	71
4.8.3.	Denaturant Dependent Unfolding and Refolding of MBP	72
4.8.4.	Circular Dichroism Measurements.....	72
4.8.5.	Unfolding Kinetics by Stopped-Flow Measurements	73
4.8.6.	Trptophan Fluorescence Wavelength Scan	73
4.8.7.	Proteinase K Protection of GroEL/ES Substrate Complex	73
4.8.8.	Maltose Binding Experiments	74
4.8.9.	Static and Dynamic Light Scattering.....	74
4.9.	Single-Molecule Experiments	75
4.9.1.	Fluorescence Correlation and Cross-correlation Spectroscopy.....	75
4.9.2.	Single-Pair FRET	77
4.10.	Mass Spectrometric Analysis	77

5. Results	79
5.1. Model Substrate – Maltose Binding Protein.....	79
5.2. Chaperonin-Assisted Refolding is Independent of Aggregation Prevention.....	80
5.2.1. Simulating the Effect of Aggregation on DM-MBP Refolding	82
5.2.2. Formation of Reversible Aggregates is not the Rate Limiting Step in.....	83
Chaperonin-Assisted DM-MBP Folding	83
5.2.3. Aggregation Limits Yield but not Rate of Rubisco Refolding	88
5.3. Slow Conversion of Trapped Intermediate Limits DM-MBP Refolding	91
5.3.1. Folding of DM-MBP Starts from a Monomeric Intermediate State.....	91
5.3.2. Kinetically Trapped Intermediate Populated during DM-MBP	95
Folding is Collapsed and Lacks Ordered Structure	95
5.3.3. Kinetically Trapped Intermediate during DM-MBP Folding is Dynamic	97
5.4. Disulfide-Mediated Constraints Accelerate Spontaneous	100
Refolding	100
5.5. Chaperonin Cage Mimics Disulfide- Mediated Constraints.....	109
5.6. Charge Clusters in the Cage Wall Plays an Active Role in Promoting Folding...119	
5.7. Charges in the Cavity Optimize the Folding Capacity of Chaperonin	126
6. Discussion.....	134
6.1. Passive versus Active Chaperonin Mechanism	134
6.2. Basis of Slow DM-MBP Folding	135
6.3. Disulfides-Mediated Folding of DM-MBP	136
6.4. Chaperonin-Assisted versus Disulfide-Mediated Folding of DM-MBP	137
6.5. Effect of Chaperonin Cavity on the Folding of DM-MBP.....	139
6.6. Effect of the Mildly Hydrophobic C-Terminal GGM Repeat on Folding of GroEL Substrates	141
6.7. Effect of ATPase rate and cycling on folding of different GroEL substrates	142
6.8. Biological Relevance of Chaperonin-Mediated Folding	143
7. References	145
A. Appendix	156
A.1. Supplementary data.....	156
A.2. Abbreviations.....	164
A.3. Curriculum Vitae	166

1. Summary

Proteins perform most biological processes in cells. To become biologically active, most proteins need to fold into specific three dimensional structures. It is known that the folding information is coded in the primary structure of the protein. However, the underlying mechanism utilized by proteins to avoid sampling the extraordinarily large number of possible conformations during their folding process at a biologically relevant timescale is not yet fully understood. Molecular chaperones have evolved to assist the folding of newly synthesized and denatured proteins in attaining their native state in the crowded cellular environment.

GroEL, the Hsp-60 of *E.coli*, and its cofactor, GroES, forming a specialized nano-compartment, is one of the most studied chaperone systems. The GroEL/ES system has been proposed to be a passive cage providing an isolated environment for single protein molecules to fold, unimpaired by aggregation. However, an active mechanism in promoting folding appears to operate in addition, based on the demonstration that GroEL/ES can substantially enhance the folding rate for proteins like bacterial RuBisCo. Recent experiments have also shown ~10-fold acceleration of folding for a double mutant of maltose binding protein (DM-MBP). The mechanism by which GroEL/ES accelerates folding of its substrate protein remains unclear.

Here we have explored the mode of chaperonin action in accelerating folding of its substrate proteins. We have used DM-MBP as a model substrate to understand the folding pathway of this protein and analyzed how GroEL/ES catalyzes its conversion to the native state. We also performed a series of experiments in which the charge properties of the GroEL central cavity was altered, and the effect on the folding rate and yield of substrate proteins (Rhodanese, MetF and DM-MBP) was measured.

By using different spectroscopic techniques, we show that DM-MBP folding is not limited by the formation of reversible aggregates but rather populates a kinetically trapped intermediate. We have been able to characterize this trapped folding intermediate by Sp-FRET (Single-Pair-Fluorescence Resonance Energy Transfer) and H/D (Hydrogen-Deuterium) exchange experiments and found it to be collapsed but structurally disordered.

We show that the DM-MBP folding rate can be accelerated by configurationally constraining the molecule by introducing long-range disulfide bonds in the kinetically trapped intermediate. Interestingly, steric confinement of the unfolded DM-MBP in the chaperonin cage closely mimics the effect of constraints imposed by the disulfide bonds on folding kinetics, in a manner mediated by negative charge clusters of the cage wall. We were also able to show that the effect of disulfide bonds and chaperonin in accelerating folding are nonadditive.

The inner wall of the GroEL/ES cavity has a net negative charge of 42. This suggested that electrostatic interactions may also influence the folding rate. By mutating the flexible GroEL C-terminal repeat sequences (13 amino acids), the charge properties of the cavity were varied. Strikingly, we found that the wild-type GroEL cavity is optimal for rate acceleration of folding for the substrates tested, although the cavity environment can be optimized for folding of a particular substrate like MetF. In summary, this study suggests that protein confinement in the chaperonin cage has the capacity to reduce entropic folding barriers, thereby promoting the formation of native contacts. This function may define the uniquely essential role of GroEL/ES cage in protein folding.

2. Introduction

2.1. Protein Folding

Proteins are linear polymers of amino acids folded into specific three dimensional structures. In general, the genetic code specifies 20 different amino acids. Proteins constitute more than half of the cell's dry weight and perform most biological processes in the cell. In all kingdoms of life, proteins play a pivotal role in all cellular processes, such as maintenance of cellular structure and inter- and intra-cellular communication, metabolism and transport. In order to perform these diverse functions, proteins must adopt specific three-dimensional structures, which are very diverse depending on the kind of function performed. Protein structure ranges from fibrillar (structural proteins), globular (many metabolic enzymes) to channels traversing membranes in energy producing and transport systems. The process by which a protein acquires its unique three-dimensional structure is called protein folding.

2.1.1. Protein Structure

Protein structures are described at four hierarchical levels: *Primary structure* is the linear amino acid sequence of a polypeptide chain. *Secondary structures* are local structures formed by different regions of the sequence. The most common secondary structures are α -helices and β -sheets (Pauling and Corey, 1951b). *The tertiary structure* is formed by packing such secondary structure elements and describes the overall structure of the polypeptide. Quaternary structure refers to the spatial arrangement of subunits in an assembly of two or more polypeptide chains. Many proteins are organized in a modular fashion, which are referred to as domains and are structurally and functionally distinct units (Doolittle, 1995; Orengo et al., 1994). Linking two or more domains has facilitated the evolution of

polypeptides with novel functions (Kummerfeld and Teichmann, 2005). Multi-domain proteins occur in all kingdoms of life, although they are more abundant in eukaryotes than in prokaryotes (65% vs 40% respectively) (Ekman et al., 2005). In the 1930s and 1940s, Linus Pauling and Robert Corey determined the X-ray structure of several amino acids and dipeptides in an effort to elucidate the structural constraints on the conformations of a polypeptide chain. These studies indicated that the peptide group has a rigid, planar structure, which is a consequence of resonance interactions that gives the peptide ~40% double-bond character. In the primary structure, the α carbons of adjacent amino acid residues are separated by three covalent bonds, arranged as $C\alpha - C - N - C\alpha$. The six atoms of the peptide group lie in a single plane, with the oxygen atom of the carboxyl group and the hydrogen of the amide nitrogen group (Figure 2.1).

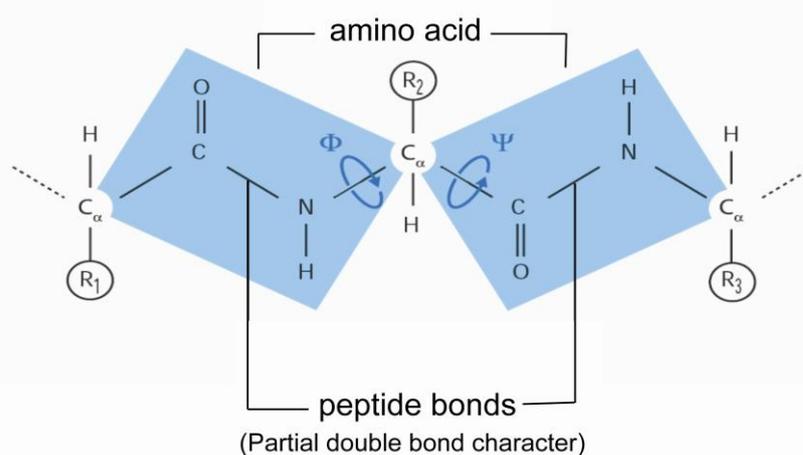


Figure 2.1: The torsional degrees of freedom in a peptide. The peptide bond is planar (blue shading) and does not allow rotation. The rotation about the $C\alpha - N$ bond and $C\alpha - C$ bond are represented by ϕ and ψ , respectively. R indicates the side chain residues of the corresponding amino acid.

A polypeptide's backbone conformation can be specified by the torsion angles (dihedral angles) about $C\alpha - N$ bonds (ϕ) and $C\alpha - C$ bonds (ψ) which are the only allowed rotation in the polypeptide backbone. The peptide $C - N$ bonds are unable to rotate freely because of their partial double bond character. These angles ϕ and ψ , are both defined as 180°

when the polypeptide chain is in its planar, fully extended conformation. There are several steric constraints on the torsion angles, ϕ and ψ , of a polypeptide backbone that limit its conformational range. The sterically allowed values of ϕ and ψ can be determined by calculating the distances between the atoms of a tripeptide at all values of ϕ and ψ for the central polypeptide. Sterically forbidden conformations are those in which any non-bonding inter-atomic distance is less than its corresponding Vanderwaals distance. The best known description of conformational restriction in proteins is provided by the ‘Ramachandran Plot’ (Ramachandran and Sasisekharan, 1968) (Figure 2.2).

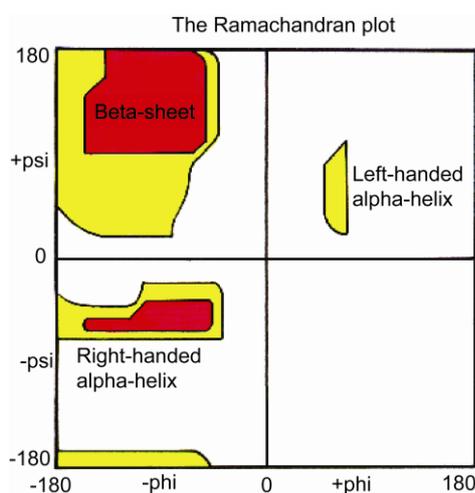


Figure 2.2: Ramachandran Plot: Only three small regions of conformational space are available to the polypeptide chain. The fully allowed regions are shown in red where as partially allowed regions are shown in yellow. Secondary structure elements like right handed α -helix and β -sheet occupy the fully allowed regions. The plots for actual proteins may have many points which do fall in forbidden regions. However, these points would be allowed if the peptide bonds are twisted by a few degrees.

Figure 2.2 indicates that ~75% of the Ramachandran plot (most combinations of ϕ and ψ) are conformationally inaccessible to a polypeptide chain. Only three small regions of the conformational map are physically accessible to a polypeptide chain which includes all of the common types of secondary structures found in proteins (α -helices, β -strands and turns). G. N. Ramachandran calculated the energy contained in various pairs of ψ and ϕ angles and found two most stable pairs, the so called α and β conformations (Ramachandran and Sasisekharan,

1968). These two pairs of angles are found to almost exclusively occur in folded proteins, including the two most prominent examples of secondary structure: α -helix and β -sheet.

The α -helix and the β -sheet elements keep the main chain in an unstrained conformation, and allow hydrogen-bonding potential of the main-chain N-H and C=O groups. If a polypeptide chain is twisted by the same amount about each of its C α atoms, it assumes a helical conformation. The α -helix is stabilized by hydrogen bonds between the carbonyl oxygen of the amino acid residue at the position n in the polypeptide chain with amide group, NH, of the residue $n + 4$. The first α -helix was described in the protein α -keratin (an abundant protein of the skin) and its derivatives, found in hair, nails and horns (Pauling and Corey, 1951a). Generally, about one-fourth of all amino acid residues in polypeptides are found in α -helices. In β -sheet conformation, hydrogen bonding occurs between neighboring polypeptide chains rather than within the same chain as in α -helices (found in the protein Fibroin; major constituent of silk) (Pauling and Corey, 1951b). The occurrence of β -sheets is often correlated with high hydrophobicities of the involved amino acid residues. Turns involve a 180° change in the direction of the polypeptide chain and are stabilized by a hydrogen bond between the carbonyl oxygen of the residue at the position n with the amide group, NH, of the residue $n + 3$ (Fersht et al., 1985).

2.1.2. Dominant Forces in Protein Folding

Protein folding occurs as result of many small driving forces such as hydrogen bonds, ion pairs, Vanderwaal's attractions, and water-mediated hydrophobic interactions. An early idea was that the primary sequence encoded secondary structures, which then encoded tertiary structures (Anfinsen, 1973). However, through statistical mechanical modeling, a different view emerged which demonstrates that there is a dominant component to the folding code and folding code is distributed both locally and globally in the protein sequence. Also,

the protein's secondary structure is as much a consequence of the tertiary structure as a cause of it (Dill, 1990; Dill, 1999). Because native proteins are only 5-10 kcal/mol more stable than their denatured states, no type of intermolecular forces can be neglected in folding and structure prediction (Yang et al., 2007). Hydrogen-bonding interactions are important as all possible hydrogen bonding interactions are satisfied in native structures. Hydrogen bonds among backbone amide and carbonyl groups are key components of all secondary structures, and studies of mutations in different solvents estimate their strengths to be around 1- 4 kcal/mol (Byrne et al., 1995) or stronger (Deechongkit et al., 2004). Similarly, tight packing in proteins implies that Vanderwaal's interactions are also important.

The role of hydrophobic interactions in protein folding also becomes important when taking into account the following evidences: (a) Proteins have a hydrophobic core, implying non-polar amino acids are driven to be sequestered from water. (b) Model compound studies show 1-2 kcal/mol for transferring a hydrophobic side chain from water into oil-like media (Wolfenden, 2007). (c) Proteins are readily denatured in non-polar solvents and (d) Sequences that are jumbled and retain only their correct hydrophobic and polar patterning sometimes fold into their expected native states (Cordes et al., 1996; Hecht et al., 2004), in the absence of efforts to design packing, charges, or hydrogen bonding. Instead of a single dominant force there are many small driving forces which drive protein folding, of which hydrogen bonding, Vanderwaal's interactions and hydrophobic forces play an important role.

2.1.3. Mechanism of Protein Folding

Understanding the folding problem requires answering the simple question: How the primary sequence of amino acids in a protein chain determines its 3D folded conformation in space? The notion of a folding "problem" first emerged around 1960, with the appearance of the first atomic-resolution protein structure of myoglobin which had helices packed together in an irregular manner in contrast to the crystalline regularity of α -helices anticipated by Linus

Pauling and colleagues (Pauling and Corey, 1951a, b). Since then the protein folding problem has come to be regarded as three inter-related problems : (a) the folding code : the thermodynamic question of the balance of interatomic forces that 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 kinetic question regarding the routes or pathways that proteins follow during folding.

2.1.4. Classical View versus New View to Protein Folding Kinetics

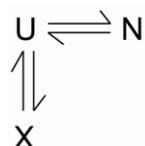
In the early 1960s Christian Anfinsen and colleagues showed that denatured Ribonuclease A folded spontaneously to its native state *in vitro* upon dilution from the denaturant, as measured by its ability to regain its enzymatic activity (Haber and Anfinsen, 1962). From this experiment he postulated that the native structure of the protein is the thermodynamically stable structure and folding depends only on the amino acid sequence of the protein and on the conditions of the solution, and not on the kinetic folding route. But in the late 1960s Cyrus Levinthal made the argument, later termed as 'Levinthal Paradox', that there are too many possible conformations for proteins to find the 'needle' (the native structure) in the 'haystack' (conformational space) by random searching. Levinthal performed mathematical calculations regarding the time that would be required for a protein to adopt its native structure if the folding process were a completely random process. If only the two most stable backbone conformations, α and β , are considered, a hypothetical protein of 150 aa in length can adopt approximately 10,300 different conformations. Taking into account that the fastest possible rate for conformational changes is approximately 10^{11}s^{-1} , it would take this hypothetical protein more than 10^{11} years to reach its native structure (Adesnik and Levinthal, 1969; Dinner et al., 2000). The discrepancy between the estimated time for random folding and the observed

fast folding of proteins is called the ‘Levinthal Paradox’. To better understand the folding process, Levinthal concluded that proteins must fold using specific ‘folding pathways’.

Levinthal framed the puzzle as if the two goals – achieving the global minimum (under thermodynamic control) and doing it so quickly (under kinetic control) – were mutually exclusive. Thermodynamic control meant that a protein reaches its global minimum in energy and that folding is pathway independent (that is, the native structure is determined only by the final native conditions) but it takes a long time because it requires an extensive search. Kinetic control meant that folding happens quickly (on a biological time scale) as it is pathway dependent. This led to a search for folding pathways and the emergence of the concept of folding intermediates in which local folded elements are stabilized and determine further folding of the polypeptide (Baldwin, 1996; Baldwin and Rose, 1999; Privalov, 1996). These intermediates would greatly reduce the number of possible conformations during folding, defining steps of a folding pathway through the random folding space, and thus allow protein folding to take place on a biologically relevant time scale. Indeed, folding intermediates were observed for various model proteins, such as Apomyoglobin, Ribonuclease A, Barstar and Lysozyme (Jamin and Baldwin, 1996; Radford et al., 1992; Udgaonkar and Baldwin, 1990; Wildegger and Kiefhaber, 1997). This is the ‘Classical View’ of protein folding which states that ‘The search for the native state through the vastness of conformational space flows through predetermined pathways defined by discrete intermediates and barriers’ (Baldwin, 1999; Creighton, 1974; Kim and Baldwin, 1990; Matthews, 1993). The classical view is based on simple phenomenological kinetics model. The raw data are single or multiple exponential time decays of optical properties that monitor changes in the protein structure after a jump to folding or unfolding conditions. Mass action models are used to fit the time constants and amplitudes by postulating macroscopic states, such as unfolded, native and various

intermediates. Reaction schemes for the three most important classical models are described below:

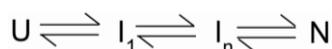
Off-pathway model



On-pathway model



Sequential model



U represents the fully unfolded denatured state, N the native state, and X or I represent intermediate states that have properties between U and N. Arrows define the pathways connecting each state in the mass-action kinetics law. Models are chosen based on which one gives the best fit to the experimental rates and amplitudes. The classical experiments generally probe the ensemble behavior of the protein, and are not able to resolve the atomic detail. One of the major challenges to the classical view of protein folding is how folding can be both pathway independent (thermodynamic control) and dependent (kinetic control) at the same time (Levinthal Dichotomy).

The ‘New View’ of protein folding recognizes that the solution to Levinthal’s paradox is ‘funnels’, not ‘tunnels’. It replaces the ‘Folding pathway’ concept of sequential events with the ‘Energy landscape’ and ‘folding funnel’ concept of parallel events (Abkevich et al., 1994; Bryngelson et al., 1995; Chan and Dill, 1994; Socci et al., 1996; Wolynes et al., 1995). The two goals of reaching a global energy minimum and doing so quickly are not mutually exclusive. The ‘New View’ is based on statistical mechanics models (lattice based representations of chain geometries and interactions) and analyzed by analytical methods and computer simulations. It also lacks atomic details but includes microscopic properties of

proteins like chain connectivity, flexibility and sequence dependent intra-chain interactions (Chan and Dill, 1996). It attempts to capture the ensemble nature of chain conformations and recognizes unfolded (U), intermediate (I) or native (N) states as ensembles of individual chain conformations. Hence, statistical mechanical models help to relate the microscopic chain dynamics to the macroscopic experimental observables. The landscape perspective readily explains the process of reaching a global minimum in free energy (satisfying Anfinsen's experiment) and doing so quickly (satisfying Levinthal concerns) by multiple folding routes on funnel-like energy landscapes (Leopold et al., 1992). Instead of viewing folding as a process in which all chains perform essentially the same sequence of events to reach the native state, the New View envisions folding as representing the ensemble average of a process that is microscopically more heterogeneous. In this process, each individual protein molecule may follow its own trajectory, but eventually reach the same point at the bottom of the funnel, the native state.

According to the principles of thermodynamics, if a system has n degrees of freedom $\Phi = [\Phi_1 + \Phi_2, \dots, \Phi_n]$, the stable state of the system can be found by determining the set of values $\Phi^* = [\Phi_1^* + \Phi_2^*, \dots, \Phi_n^*]$ that gives the minimum value of free energy function $F(\Phi) = F(\Phi_1 + \Phi_2, \dots, + \Phi_n)$, when explored over all possible values of Φ . Such functions $F(\Phi)$ are called *energy landscapes*. A microstate is a single point on an energy landscape and has free energy $F_{\text{micro}} = F(\Phi)$, which is also called the *internal free energy* (Dill and Chan, 1997). A macro state has free energy $F_{\text{macro}} = F(\xi)$, where ξ represents particular ensemble of microscopic conformations (Figure 2.3).

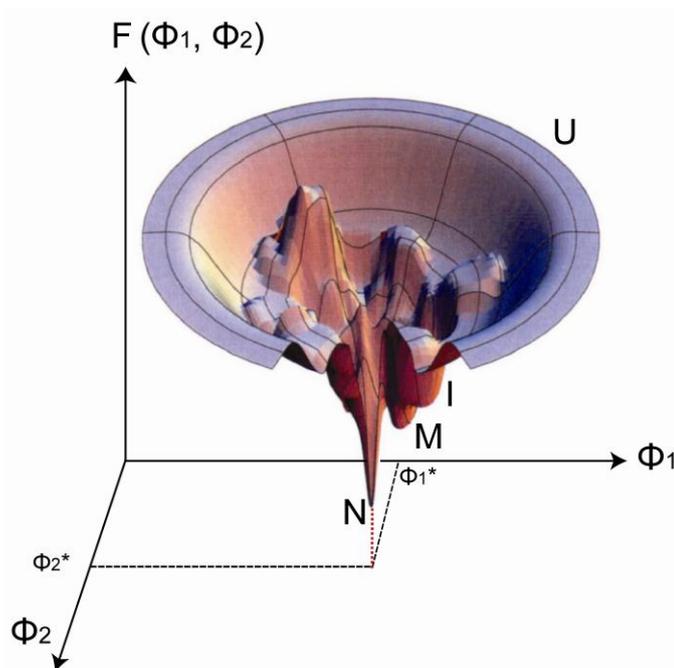


Figure 2.3: Schematic Representation of the Free Energy Landscape of Protein Folding:

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, \dots such as the backbone and side-chain bond angles. Each conformation is represented as a point on the landscape. Hills correspond to high energy conformations (for example, burying polar groups in hydrophobic cores, or unfavorable $\Phi \Psi$ angles) and valleys are configurations that are more favorable than others nearby. 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.

In other words, an energy landscape is the free energy of each conformation as a function of the degrees of freedom, such as the dihedral bond angles along the peptide backbone and side chains (Dill and Chan, 1997). The vertical axis of funnel represents the *internal free energy* of a given chain conformation: the sum of all the intra-chain enthalpies and solvent interactions (hydrogen bonds, torsion angle energies, monomer contacts, etc). Internal free energy is not the macroscopic free energy that would be measured in a folding experiment, because it describes only a single chain conformation, and not the ensemble average over all chain conformations. The many lateral axes represent the conformational coordinates. The high dimensionality of this representation reflects the many degrees of

freedom of a protein chain. Each conformation is represented by a point on the multidimensional energy surface. Conformations that are similar geometrically are close to one another on energy landscapes and can interconvert into one another. The native state of a protein, defined as the conformation with the lowest free energy, is thus the lowest point in the energy landscape, or the bottom of the funnel. Denatured protein usually resembles a random coil in which local interactions dominate the conformational behavior, giving rise to a highly heterogeneous state and this forms the top of the funnel. During folding, the protein follows a route from the top of the funnel, representing a disordered denatured state to the bottom of the funnel. If the folding polypeptide cannot escape a local minimum, it becomes kinetically trapped and eventually misfolds in an off-pathway reaction.

The actual pathway along which the protein folds is dependent on the physical environment in which the folding reaction takes place. An implication of using the internal free energy as the vertical coordinate is that the temperature or denaturant can stretch or compress the landscape in the vertical direction (Abkevich et al., 1994). The solvent environment *in vivo* and *in vitro* modulates the stability of local minima on the free energy landscape. Generally, this environmental modulation relative to simple aqueous solvent is small (a few RT), but resulting effects can be dramatic. A seemingly small modification of sequence or environment can cause a protein to unfold or aggregate, fold to a new state, or accelerate folding dramatically, as seen for the engineered downhill folders.

The new view proposes that multi-dimensional energy landscapes of proteins can have a much broader array of shapes that involve hills valleys, ridges, channels, moats, varying slopes etc. Bryngelson and Wolynes first explored the bumpiness of protein folding landscapes in a simplified spin-glass model (Bryngelson and Wolynes, 1987; Bryngelson and Wolynes, 1989). Leopold, Montal and Onuchi (Leopold et al., 1992) first described in some detail how the shape of a folding funnel depends on amino acid sequence, by computer enumeration of

conformations in lattice hetero-polymer models. Two-state fast folding kinetics is described by funnel-shaped landscapes with no significant kinetic traps. Slow multi-exponential folding is represented by bumpy or rugged landscapes (Figure 2.3). Two other energy landscapes are shown below describing different rate limiting steps of folding (Figure 2.4).

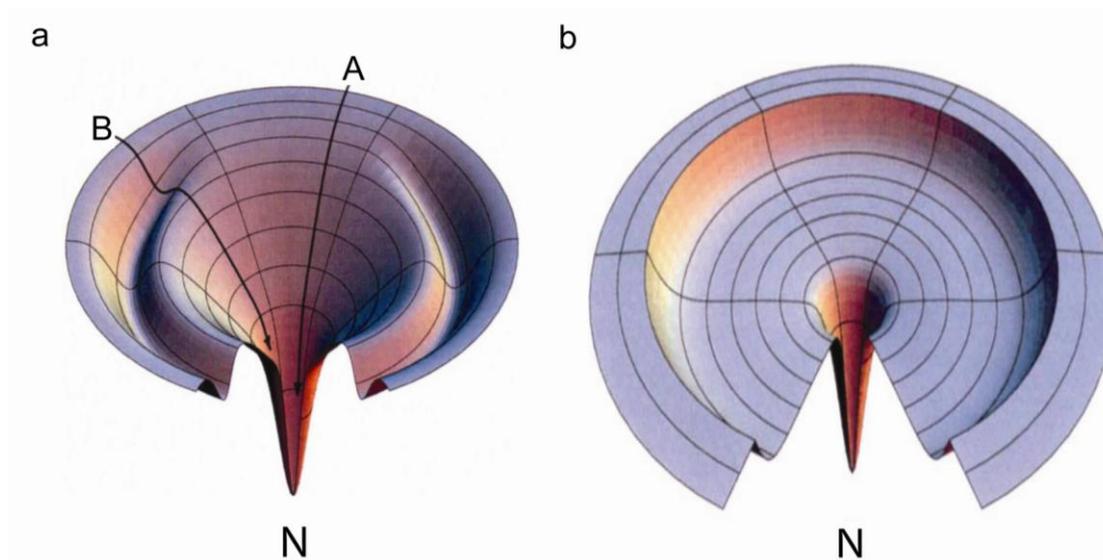


Figure 2.4: Different Folding Scenarios: (a) Moat landscape, where protein could have a fast folding throughway process A, in parallel with a slow-folding process B, which involves kinetic trap. (b) Champagne glass landscape, where conformational entropy causes free energy barriers to folding.

New view recognizes on-pathway or off-pathway intermediates explained in mass action models just as a difference in the distribution of energy traps on the energy landscapes. Here, energy means internal free energy of a conformation - the vertical scale on the energy landscape (Figure 2.3). Such intermediates correspond to ‘misfolds’ (Baldwin, 1996; Dobson et al., 1994; Radford et al., 1992; Sosnick et al., 1996; Sosnick et al., 1994) resulting from formation of low-energy nonnative contacts or steric hinderance, either of which could stall further progress (Chan and Dill, 1994). In classical chemical kinetics, ‘transition state’ is a term that describes a rate-limiting step corresponding to a particular ‘bottleneck’ conformation on a specific reaction pathway. According to the New View, transition states are not specific structures but they are ensembles of conformations forming a kinetic bottleneck to folding

(Chen and Matthews, 1994; Sosnick et al., 1996). It is fundamentally a concept about rates, not about specific structures. The kinetic bottle-neck for folding does not necessarily describe specific conformations or particular structures of the chain, although the new view does not preclude them (Bryngelson et al., 1995).

Figure 2.3 describes a folding reaction which is limited by formation of kinetic traps. Folding may proceed in two or more kinetic phases, often with a fast collapse to a compact ensemble followed by slow reconfiguration of kinetically trapped compact non-native conformation into the native structure. In this case, the transition state is the ensemble of lookout-point conformations that have been opened apart, relative to the compact trapped states from which they originated (Abkevich et al., 1994; Camacho and Thirumalai, 1993; Chan and Dill, 1994; Fersht, 1995; Sali et al., 1994). Hence, folding transition state can be many different chain conformations and not some specific structures. Fig 2.4a shows a moat landscape, indicating a funnel like throughway pathway for the A routes and obligatory kinetic traps for the B route which is referred to as Kinetic Partitioning (Dill et al., 1995). Figure 2.4b shows a champagne glass landscape where the bottleneck or rate limiting step to folding is the aimless wandering on the flat plateau as the chain tries to find its way downhill (Dill and Chan, 1997). Folding may be slowed down by conformational entropy barriers, which are more like plains on energy landscape and not basins. Free energy barriers occur when folding is slow due to extensive conformational search.

The central experimental result for protein folding with two state kinetics are chevron plots of denaturant effects on rates, and Arrhenius plots of temperature effects on rates. In the energy landscape picture, changing the solvent or temperature towards more native-like conditions has the effect of stretching down the energy surface. The statistical mechanical models can rationalize non-Arrhenius rate laws (Chan and Dill, 1994; Dill et al., 1995), chevron plots (Chan and Dill, 1994), mutational effects (Shortle et al., 1992), kinetic traps,

barriers (Bryngelson et al., 1995; Camacho and Thirumalai, 1993; Dill et al., 1995) and the relationship of equilibrium properties and fluctuations kinetics (Chan and Dill, 1996; Klimov and Thirumalai, 1996; Sali et al., 1994).

2.1.5 Disulfide Bond-Mediated Protein Folding

The focus of protein folding has been the identification and characterization of the initial, final and intermediate conformational states as well as the determination of the steps by which they are inter-converted. Most protein folding intermediates are only transiently present, making their isolation and characterization difficult by commonly used spectroscopic techniques. However, most secretory proteins have an important covalent modification: disulfide bonds. Disulfide bonds are one of the few post translational covalent modifications that occur during protein folding. Disulfide bond formation in proteins is required not only for folding but also for stability and function. Failure to form the correct disulfide bonds is likely to cause protein aggregation and subsequent degradation by cellular proteases.

Disulfide bonds are formed because of the reduction-oxidation chemistry of the covalent interaction between two thiol groups. The relatively slow kinetics of formation of the disulfide bond and the availability of thiol trapping reagents that rapidly quench disulfide bond formation have facilitated the isolation, purification, and characterization of folding intermediates. These trapped intermediates have been used to determine the pathways of several disulfide rich proteins *in vitro* (Park et al., 2007). Knowledge of these disulfide-linked folding pathways has furthered our understanding of protein structure-function relationships.

Most disulfide bonds serve to stabilize protein structure. It is generally accepted that protein disulfide bonds stabilize the native conformation of a protein by destabilizing the denatured form; i.e., they decrease the entropy of the unfolded form, making it less favorable compared with the folded form (Thornton et al., 1981). According to theoretical studies, the

increase in the stability of the native structure due to the formation of a particular disulfide bond is directly proportional to the number of residues between the linked cysteines: the larger the number of residues between the disulfide, the greater is the stability imparted to the native structure (Pace et al., 1988). The kinetics of protein folding are greatly affected by the location of the disulfide bond relative to the folding nucleus. Disulfide bonds introduced in or near the folding nucleus accelerate protein folding, whereas disulfide bonds introduced elsewhere can decelerate folding by up to three orders of magnitude (Abkevich and Shakhnovich, 2000). The formation of disulfide bonds is thermodynamically coupled to the process of protein folding. In general, most disulfide bonds stabilize proteins and affect the rate of protein folding.

However, minor population of disulfide bonds also serves a functional role. Functional disulfides can be further classified into catalytic disulfides and allosteric disulfides. Catalytic disulfides are typically found at the active site of enzymes that mediate thiol-disulfide exchange (oxidoreductases). These dithiols/disulfides are transferred to a protein substrate, resulting in the formation, reduction, or isomerization of disulfide bonds. Allosteric disulfides regulate function in a nonenzymatic way by mediating changes in the protein structure (Hogg, 2003; Schmidt et al., 2006).

2.2. Methods for Studying Protein Folding

Major advances have been made in understanding the protein folding mechanism since the statistical mechanical model came into the picture. These have been derived from wealth of new and elegant experimental approaches (Table 1), combined with theoretical methods. The key to discerning the nature of folding mechanisms is to combine the results from different techniques so that different aspects of folding can be probed and the results combined into a common picture of folding process (Dobson et al., 1994).

Table 1. Experimental techniques to investigate protein folding

Technique	Time scale	Structural parameter probed
1. Fluorescence	ns-s	
a. Intrinsic fluorescence		Environment of Trp and Tyr
b. ANS binding		Exposure of hydrophobic surface area
c. Substrate binding		Formation of native contacts
d. FRET		Inter-residue distance
Anisotropy		Correlation time/ mobility
2. Circular Dichroism	ns-s	
a. Far UV		Secondary structure formation
b. Near UV		Tertiary structure formation
3. Small-angle X-ray scattering	\geq ms	Dimension and shape of polypeptide
4. Absorbance (near UV)	ns-s	Environment of aromatic residues
5. FTIR	ns-s	Secondary structure formation
6. NMR		
a. Real time	ms-s	Environment of individual residues
b. Dynamic NMR		Lineshape analysis provides folding-unfolding rate close to equilibrium
7. Hydrogen exchange (HX)		
a. Native state	min	Global stability and metastable states
b. Pulsed HX ESI MS	ms-s	Folding population
c. Pulsed HX NMR	ms-s	Hydrogen bond formation in specific residues
8. Atomic force spectroscopy (AFM)	s	Unfolding force and rate constants of single molecules

The table was modified from (Brockwell et al., 2000; Radford, 2000).

Abbreviations: ANS: 1-anilino naphthalene sulphonic acid; ESI MS: electrospray ionization mass spectrometry; FRET: fluorescence resonance energy transfer; FTIR; fourier transform infra-red.

2.2.1. Fluorescence

Fluorescence spectroscopy is more sensitive to the environment of the chromophore than absorption or CD spectroscopy providing more flexibility to assess biochemical information in living cells. During the approximate lifetime (10^{-9} s) of the fluorophores several kinds of processes such as protonation or deprotonation, solvent cage relaxation, local conformational changes and any process coupled to translational or rotational motion of the molecule may occur (Steinberg, 1971). Therefore, fluorescence is used to monitor conformational changes within a molecule. Fluorescence life-time and quantum yield are the most important characteristics of a fluorophore.

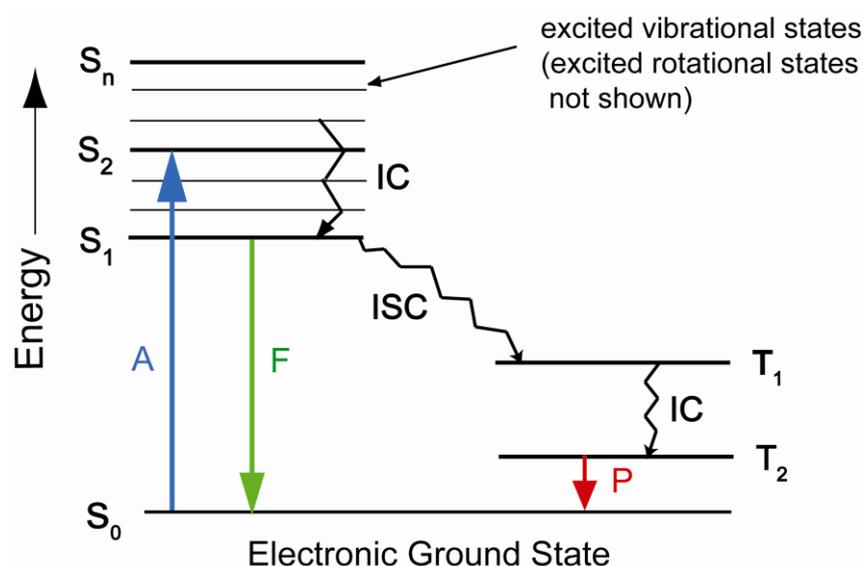


Figure 2.5: Jablonski Diagram. It shows a number of possible routes by which an excited molecule can return to its ground state. Upon absorption of light, electrons are excited to higher energy singlet states (S_n) which returns to ground state (S_0) by spontaneous emission of photon with higher wavelength. This phenomenon is termed as *Fluorescence* (F). Two other processes compete with fluorescence: Internal conversion (IC), a non-radiative decay where the energy is transferred to the environment and inter system crossing (ISC) where an electron undergoes a transition from a singlet to a triplet state. The process of decay from the excited triplet state to the ground state occurs at significantly slower timescales and higher wavelength as compared to fluorescence. This process is termed as *Phosphorescence* (P). Adapted from ‘Principles of Fluorescence Spectroscopy’ by Lakowicz, J.R.

In fluorescence spectroscopy, the information available from the photon is the number of photons (*intensity information*), the position in space where the photon was detected (*image information*), the energy of the photon (*spectral information*), its polarization (*orientational information*), and the delay between excitation and fluorescence emission (*Lifetime information*). The processes that occur between the absorption and emission of light are schematically illustrated by the Jablonski diagram (Figure 2.5). Prior to excitation, the electronic configuration of the molecule is described as being in the electronic ground state (S_0). Upon absorbing a photon of light, electrons are excited to higher energy state (S_n), a process that takes only a few femto-seconds, a time too short for significant displacement of nuclei (Frank-Condon principle). By spontaneous emission of a photon, the molecule with singlet spin is able to return to the ground state. This event is termed fluorescence. Fluorescence typically occurs at lower energies or longer wavelengths, which is called Stokes shift. The lifetime of the excited state is the average time the molecule spends in the excited state before returning to the ground state. Generally, fluorescence lifetimes are near 10 ns. This determines the time available for the fluorophore in the excited state to interact with other molecules or diffuse in its environment. Following light absorption, several other processes compete with fluorescence, briefly described as follow:

- 1) Internal conversion: the process in which excitation energy is lost by collision with solvent or by dissipation through internal vibrational modes. Internal conversion is generally complete prior to emission as fluorescence lifetime is 10^{-8} s as compared to internal conversion which generally occurs within 10^{-12} s. In general, this process increases with increase in temperature which becomes critical while monitoring thermally induced macromolecular conformational changes.
- 2) De-excitation: results from interactions of solute molecules capable of quenching the excited state.

- 3) Inter-system crossing: a process that converts excited singlet spin into triplet state. Although forbidden, this could still occur if the spin orbit interaction is strong enough. Triplet state converts to the ground state either by phosphorescence (emission of photon) or by internal conversion. Phosphorescence occurs at longer wavelength as compared to fluorescence as triplet state is lower in energy than singlet state.

Fluorescence quantum yield is the number of emitted photons relative to the number of absorbed photons. The quantum yield could be close to unity if the radiationless decay is much smaller than the rate of radiative decay. Energy yield of fluorescence is always less than unity because of Stokes losses.

2.2.2. FRET

Fluorescence or Förster Resonance Energy Transfer (FRET) is the radiationless transfer of energy of an excited fluorophore (Donor, D) to a second fluorophore (Acceptor, A). It is an electro-dynamic phenomenon which occurs between a donor molecule in excited state and acceptor molecule in ground state (Figure 2.6a). This mechanism was first elucidated by Theodor Förster. The donor molecules typically emit at shorter wavelengths that overlap with the absorption spectrum of the acceptor (Figure 2.6b). The rate of energy transfer depends on this spectral overlap, the quantum yield of the donor, the relative orientation of the donor and acceptor transition dipoles, and the distance between the donor and acceptor molecules (Stryer, 1978). The energy transfer occurs due to long range dipole-dipole coupling and not by emission or re-absorption of photons. Consequently, FRET contains molecular information independent of solvent relaxation effects, excited state interactions, fluorescence quenching, or anisotropy, except for their effect on the spectral properties of the donor or acceptor.

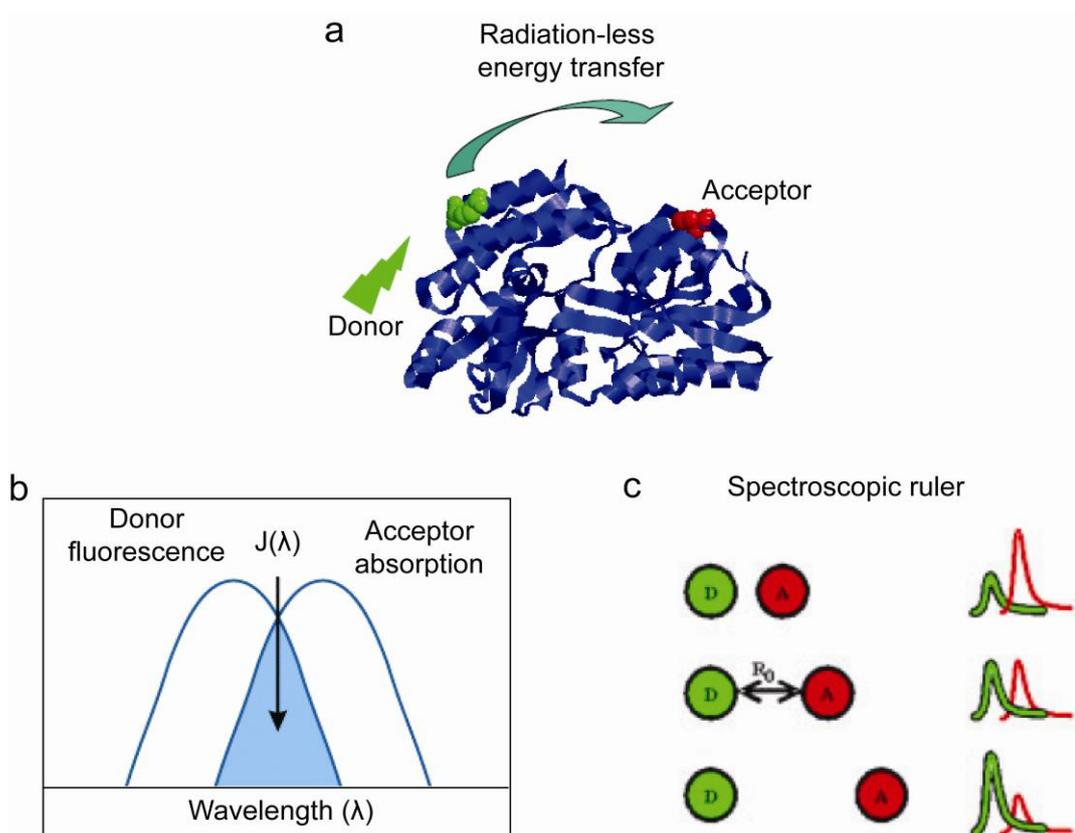


Figure 2.6: Schematic Representation of Basic Concept of FRET. (a) Spectral overlap between donor and acceptor fluorophores. (b) The radiation-less energy transfer from donor to acceptor upon selective excitation. (c) Spectroscopic ruler used to investigate molecular dynamics.

The most common application of FRET is to measure the distances between two sites in a range of 10 - 80 Å on a macromolecule by measuring extent of energy transfer between donor and acceptor (Figure 2.6c). Ultra-sensitive fluorescence methods allow to investigate the dynamics and interactions of biomolecules and also global structural alterations with high accuracy even at the level of single fluorophore. Such ultra-sensitive methods include fluorescence correlation spectroscopy (FCS), fluorescence cross-correlation spectroscopy (FCCS), burst analysis, single molecule studies and single-pair fluorescence resonance energy transfer (SpFRET) experiments.

The distance at which FRET is 50% efficient is called the Förster distance, which is typically in the range of 20-60 Å. The rate of energy transfer from a donor (D) to an acceptor (A) $K_T(R)$ is given by

$$K_T = \left(\frac{1}{\tau_d}\right) \left(\frac{R_0}{R}\right)^6 \quad (1)$$

Where τ_D is the lifetime of the donor fluorophore without an acceptor, R is the distance between the dyes and R_0 is the Förster radius. Förster radius is given by

$$R_0^6 = 8.8 \cdot 10^{-28} \kappa^2 n^{-4} \phi_D J \text{ in } [\text{Å}]^6 \quad (2)$$

where κ^2 is the orientation factor, n is the refractive index of the solvent, Φ_D the quantum yield of the donor without acceptor and J is the spectral overlap integral.

The FRET efficiency is defined as the ratio of transferred photons to photons absorbed by the donor. This is typically measured using the relative fluorescence intensity of the donor, in the absence (F_D) and presence (F_{DA}) of acceptor.

$$E = 1 - \left(\frac{F_{DA}}{F_D}\right) \quad (3)$$

The transfer efficiency can also be calculated from the lifetimes under these respective conditions (τ_D and τ_{DA}):

$$E = 1 - \left(\frac{\tau_{DA}}{\tau_D}\right) \quad (4)$$

This expression allows the Förster distance to be calculated from the spectral properties of the donor and acceptor and the donor quantum yield.

$$E = \frac{R_0^6}{R_0^6 + r^6} \quad (5)$$

The transfer efficiency is strongly dependent on distance, and is proportional to r^{-6} when the D-A distance is near R_0 .

A major limiting step in the determination of molecular distances by FRET is the orientation factor for the dipole-dipole coupling because it cannot be determined by any currently available technique. Therefore, distances calculated from energy transfer data are not unique except where an appropriate average value of orientation factor can be applied.

Resonance energy transfer is also used to study macromolecular systems in which a single D-A is not present like unfolded proteins or proteins while they are folding, where there is a distribution of such distances. Such systems are best studied with time-resolved measurements (Bagshaw and Cherny, 2006). Alternatively, energy transfer can be measured between identical chromophores that have a limited Stokes shift which is referred to as homotransfer (Kalinin and Johansson, 2004).

2.2.3. Single Molecule FRET

Protein folding is a process characterized by a large degree of conformational heterogeneity. In such cases, classical experimental methods give only mean values, averaged over large ensembles of molecules. Such methods fail to elucidate microscopic distributions of conformations, trajectories, or sequence of events underlying molecular mechanisms. This could be avoided by studying single molecules. The correlation of the behavior of a single molecule to that of an ensemble of molecules is given by statistical thermodynamics. This proposition indicates that the temporal average of a certain property of a single molecule is equal to the ensemble average of the associated property at a single time point. A particularly

versatile method is fluorescence detection. In combination with FRET, distances and conformational dynamics can be investigated in single molecules. Different methods have been developed to utilize various combinations of the available information. Additional information is available when multiple excitation sources are used like alternating laser excitation (ALEX) (Lee et al., 2005). Here, two excitation sources are interleaved on a timescale between 25 and 300 μ s, switching between both excitation sources on a time scale faster than the diffusion of the particle in the probe volume.

Pulsed interleaved excitation (PIE) (Muller et al., 2005) could be used to perform more accurate FRET measurements by including molecules containing only active donor and acceptor in the analysis. PIE is the use of two or more pulsed excitation sources, alternated with sufficient delay that all the emitted photons from one laser pulse are detected before the next pulse of a different color arrives. The difference between ALEX and PIE is that the alternation between green and red excitation occurs faster than the rate of photon detection. The experimental setup of PIE is based on two-channel confocal microscope. Generally, one laser source is delayed by 100 ns with respect to the other. The photons are detected using time-correlated single-photon counting (TCSPC). The data acquisition card and excitation sources are synchronized such that the excitation source responsible for generating the detected photon is encoded into the arrival time of the photon. The use of two, subnanosecond pulsed lasers allow PIE to include lifetime information of the fluorophores. The faster timescale of the interleaved excitation allows FCS experiments to be performed with sub-microsecond resolution in addition to all the other possibilities of ALEX. In addition, PIE-FCCS increases the sensitivity of FCCS by removing any residual cross-talk from the cross correlation function. Even for the spFRET measurements, the measured fluorescence is averaged over the same period in time, and is not affected by significant diffusion of the particle between alternating excitation pulses.

For spFRET, measurements, the intensities of the donor and acceptor fluorescence are used to determine FRET efficiencies which are empirically determined using the following equation:

$$f_E = \frac{F_{AD}}{\alpha F_{DA} + F_{AD}} \quad (6)$$

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.

2.3. Protein Folding *In Vivo*

It is well known that information required for a protein to attain its native three-dimensional structure resides in its primary sequence (Anfinsen, 1973; Dobson and Karplus, 1999). Irrespective of whether a protein folds in an intact cell, *in vivo*, or in the test tube, *in vitro* essentially the same native structure is obtained. However, the macromolecular crowded environment of the cell imposes different constraints to folding which can lead to misfolding and aggregation and hence, drastically reduce the efficiency of the folding process.

The effective protein concentration in *E.coli* cells has been estimated to be as high as 300-400 mg/ml (Ellis, 2001; Zimmerman and Minton, 1993). The high concentration of proteins, nucleic acids and other macromolecules makes the cellular environment highly crowded posing another obstacle for efficient protein folding. Crowding gives rise to excluded volume effects, which can result in increased affinities between interacting macromolecules by up to 10-100 fold (Minton, 2000; van den Berg et al., 1999). As a result, the inter-molecular binding constants between partially folded states are increased, leading to an increased probability of aggregation during folding. In addition to the highly crowded cellular environment, extremes of pH and temperature can also lead to aggregation and misfolding (Shortle, 1996).

Folding of the protein in the cell is linked to its biosynthesis, which is a vectorial process (N-terminus to C-terminus). Additionally, the exit tunnel of the large ribosomal subunit, which is around ~100 Å long and ~20 Å wide (Ban, 2000), largely precludes the folding beyond the formation of α -helical elements (Kramer et al., 2009; Lu and Deutsch, 2005) and thus prevents the C-terminal 40-60 residues of the chain from participating in long range interactions. As a result the information for the folding process becomes available sequentially and not in the form of complete protein molecule as is the case for *in vitro*

refolding. Taking into account the slow translational speed (~15-75 s for a 300 amino-acid protein), nascent chains are exposed in partially folded, aggregation sensitive states for a prolonged period of time. Moreover, non-native intra-chain contacts formed during translation could block folding upon completion of synthesis. This incomplete availability of structural information during translation could affect the process of attaining the final native structure. Folding of multi-domain proteins in eukaryotes becomes even more complicated. In the absence of interacting domains, a folded domain may expose its hydrophobic residues which are otherwise buried in native oligomeric state making them highly aggregation prone. Also post-translational modifications in eukaryotes may affect the folding process.

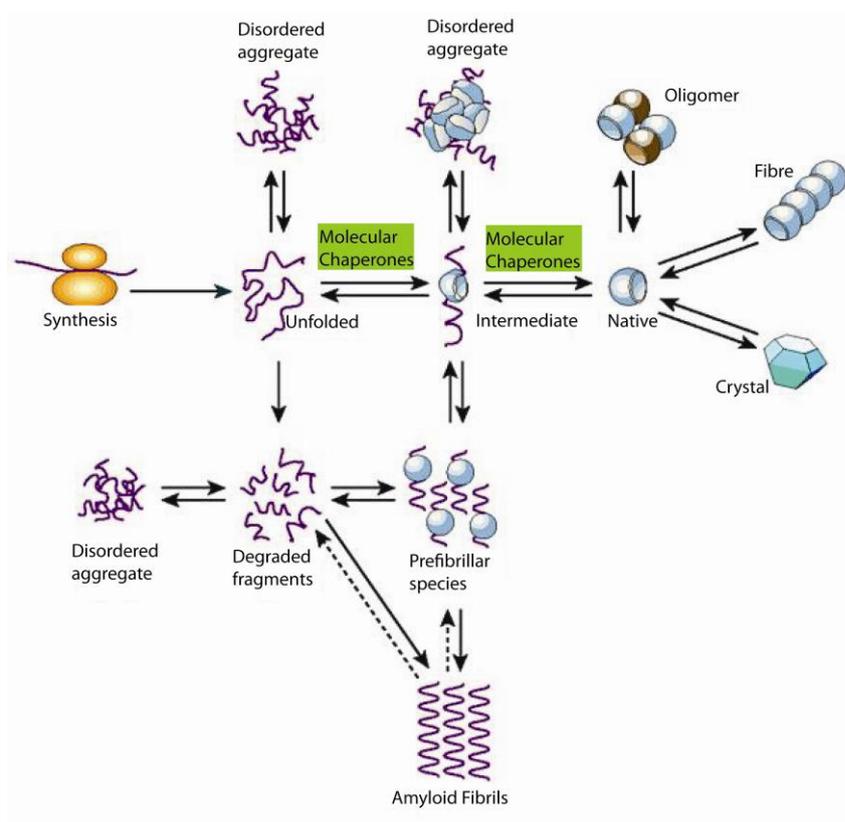


Figure 2.7: Protein Folding *In Vivo*: A unified view representing possible fates of a polypeptide upon synthesis from ribosomes. *De novo* polypeptides reach their native structure, often through one or more partially folded intermediates, generally assisted by molecular chaperones. Non-native polypeptides may aggregate as a side reaction of productive folding, in the crowded cellular environment. An amyloid fibril is just one form of aggregate, but it is unique in having highly organized but ‘misfolded’ structure. Other assemblies, including functional oligomers and natural

protein fibres contain natively folded molecules like the protein crystals produced *in vitro* for X-ray diffraction studies. The populations and interconversions of the various states are determined by their relative thermodynamic and kinetic stabilities under given cellular conditions. These transitions are also regulated by molecular chaperones, proteolytic enzymes and other factors. Failure of such regulatory mechanism is the major factor in the onset and development of misfolding diseases. Adapted from (Dobson, 2003).

Misfolded proteins *in vivo* are recruited into aggregates. This leads not only to inactivation of the affected proteins, but can also result in severe cellular dysfunction, causing a number of human diseases (Barral et al., 2004). Several cellular machineries have evolved to counteract the tendency of non-native polypeptide chains to aggregate inside the cell. One group of proteins involved in this protection are the so called ‘Molecular Chaperones’ (Figure 2.7). The term molecular chaperone was originally coined to describe the function of nucleoplasmin, a nuclear protein that facilitates proper assembly of chromatin by preventing improper interactions between histones and DNA (Ellis and Hartl, 1996). Later on, this term was generalized to include a range of functionally related, but diverse proteins that assist the folding and assembly of other proteins. Many of these are known as stress proteins or heat-shock proteins, as they are up-regulated during stress when concentration of aggregation prone folding intermediates increase. Chaperones are usually named according to their molecular weight (Hsp40, Hsp70, Hsp60, Hsp90, Hsp100 and small heat-shock proteins) and they all have the capacity to prevent aggregation. Certain members of Hsp100 family in bacteria and fungi have been shown to actively dissociate aggregates for subsequent folding or degradation (Weibezahn et al., 2005).

Our current understanding of molecular chaperones is that they transiently bind and stabilize unstable conformations of other proteins, and through regulated binding and release cycles (which may or may not be ATP dependent), facilitate their correct folding. They can play roles in folding (following *de novo* synthesis, during translocation across membranes or stress induced denaturation), oligomeric assembly, interaction with other cellular partners,

switching between active and inactive conformations, intracellular transport, or proteolytic degradation, either singly or with the help of cofactors (Fink, 1999). Molecular chaperones, generally, do not contribute steric information to the folding process unlike folding catalysts such as peptidyl-prolyl-isomerases (PPIases) and protein disulfide isomerases (PDIs) (Freedman et al., 1994; Schmid, 1993).

Notably a number of essential proteins have kinetically frustrated folding pathways and in order to overcome kinetic folding barriers they need assistance of chaperones. Mutations in proteins can often disrupt the protein's ability to fold and chaperones may help to buffer such deleterious mutations (Maisnier-Patin et al., 2005). This buffering function is thought to be crucial in the evolution of new protein functions and phenotypic traits (Rutherford and Lindquist, 1998).

2.3.1. The Chaperone Network in the Cytosol

Protein folding is so challenging *in vivo* that molecular chaperones have evolved and are distributed ubiquitously across the three kingdoms of life. Multiplicity of intracellular chaperones necessitates a high degree of functional coordination to ensure efficient folding of many proteins (Vabulas and Hartl, 2005). As summarized in Figure 2.8, in all three kingdoms of life, there are two major principles of chaperone action represented by (i) the machinery that functions in stabilizing nascent polypeptides on ribosomes and initiating folding (Weibezahn et al., 2005) and (ii) components that act downstream in completing the folding process (Albanèse et al., 2006; Langer et al., 1992). Both systems cooperate in coherent pathways.

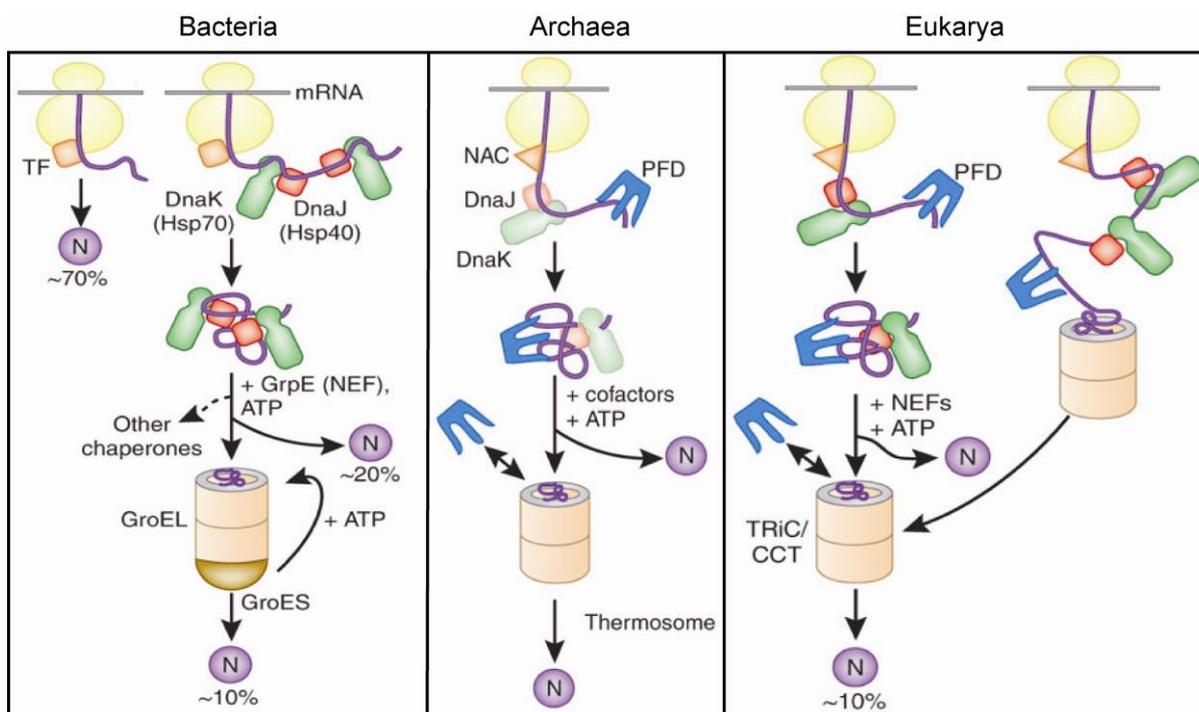


Figure 2.8: Model for *De Novo* Protein Folding Assisted by Chaperone in the Cytosol. In bacteria, most small proteins (~70% of total) may fold rapidly upon synthesis without further assistance upon release from the ribosome and ribosome bound Trigger factor (TF). Longer chains interact consequently with DnaK/DnaJ and fold upon one or several rounds of ATP-dependent binding and release (~20% of total). About 10% of protein gets transferred to the chaperonin system (GroEL/ES) to reach to its native state (N). In archaea, only some species contain DnaK/DnaJ. In eukaryotes, nascent chain associated complex (NAC) probably interacts with nascent chains. About 20% of polypeptides reach their native state in a reaction assisted by RAC (Ribosome associated complex), Hsp70 and Hsp40. A subset of Hsp70 substrates is transferred to Hsp90. About 10% of proteins are co- or post-translationally transferred to the eukaryotic chaperonin TRiC/CCT in a reaction mediated by Hsp70 and prefoldin (PFD). PFD recognizes the nascent chains of certain TRiC substrates, including actin and tubulin (adapted from (Hartl and Hayer-Hartl, 2009)).

A nascent polypeptide chain interacts with ribosomes-bound chaperones positioned in close proximity to the polypeptide exit site. This category includes Trigger Factor in bacteria (Figure 2.8), a specialized Hsp70 system called RAC (Ribosome associated complex) in eukaryotes (figure 2.8) and NAC (Nascent chain associated complex) in archaea and eukaryotes (Chang et al., 2007; Kramer et al., 2009). The next class of chaperones are members of Hsp70 family which mediate co- or post translational folding through ATP-

regulated binding cycles. Downstream from the Hsp70 class of chaperones are the chaperonins, which are large, cylindrical complexes and function by enclosing protein molecules in a nano-compartment, so that folding can occur unimpaired by aggregation. It is interesting to note that the eukaryotic chaperonin interacts directly with Hsp70 (Cuéllar et al., 2008), and thus can be recruited to nascent chains during protein synthesis (McCallum et al., 2000). This functional coupling between different chaperones facilitates co-translational folding and avoids the partitioning of non-native protein in the bulk cytosol (Siegers et al., 1999; Thulasiraman et al., 1999). The cytosolic chaperones of bacteria have been investigated most extensively and are best understood.

2.3.2. Ribosome-Associated Chaperone : Trigger Factor

Nascent polypeptide chains exposing hydrophobic regions interact with the ribosome-associated chaperones during exit from the ribosomal exit tunnel. Trigger factor (TF) is the bacterial component of the ribosome-associated chaperones which has been shown to interact with the aggregation prone nascent polypeptide chain by cross linking studies (Hesterkamp et al., 1996). It is highly abundant *E.coli* protein of ~50 kDa, consisting of three discontinuous domains: N-terminal ribosome binding domain, a peptidyl-prolyl isomerase (PPIase) domain and a C-terminal domain positioned between the N and PPIase domain (Ferbitz et al., 2004) (Figure 2.9a).

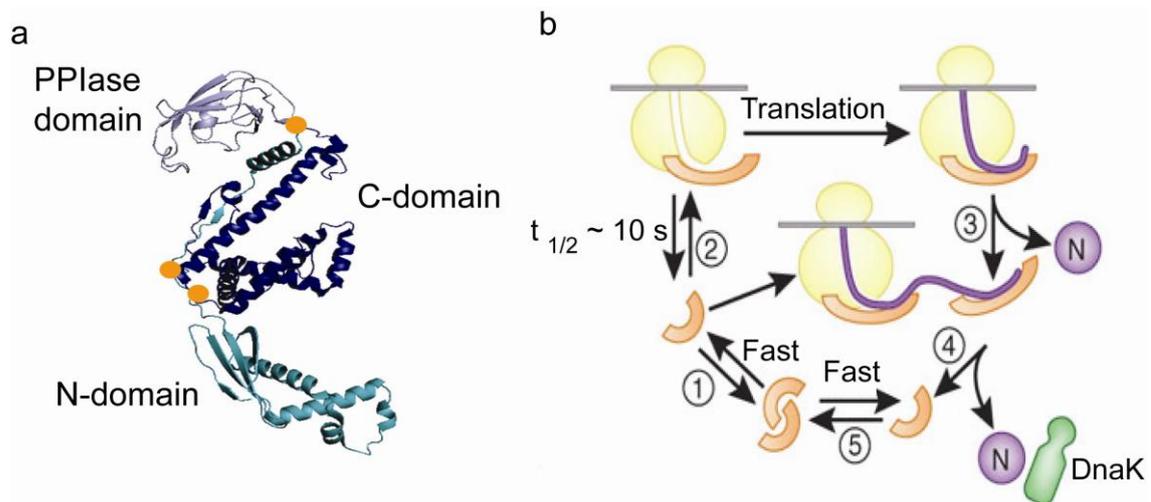


Figure 2.9: Structure and Reaction Cycle of Trigger factor. (a) The domain structure of trigger factor showing PPIase domain(aa 150-245) which displays catalytic activity as peptidyl-prolyl *cis/trans* isomerase (PPIase) and belong to FKBP (FK506 binding protein) family of proteases, C-domain (aa 246-432) positioned between PPIase and N-domain and N-domain (aa 1-149) which is necessary for ribosome binding (Ferbitz et al., 2004). (b) A model of the TF reaction cycle. (1) Free Trigger Factor is in rapid equilibrium between monomeric and dimeric states. (2) TF monomer binds to translating ribosomes causing conformational expansion of TF which activates it for interaction with nascent chains. (3) Release of TF from the nascent chain coincides with TF dissociation from the ribosome and allows completion of folding to the native state. (4) Structurally more complex protein 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 then enters the monomer-dimer pool (adapted from (Hartl and Hayer-Hartl, 2009).

The N domain of TF binds to ribosomal protein L23 and L29 next to the polypeptide exit site (Merz et al., 2006), with a mean residence time of ~10-15 s (Kaiser et al., 2006a). TF is a monomer when bound to the ribosome and a dimer when free in the cytosol. The C-domain, containing two arm-like protrusions is the major binding region for the hydrophobic nascent chain segments (Lakshmipathy et al., 2007; Merz et al., 2008). TF function is not regulated by ATP (Hesterkamp and Bukau, 1996) but rather it scans the emerging nascent polypeptide chain for hydrophobic regions, shielding them from aggregation (Figure 2.9b). It has been shown using fluorescence spectroscopy that ribosome binding causes a

conformational opening of TF, and presumably activating it for nascent chain interaction (Baram et al., 2005; Kaiser et al., 2006b; Schlünzen et al., 2005). The biological significance of the PPIase domain in protein folding, which can catalyze prolyl *cis-trans* isomerisation *in vitro* is still unclear (Genevaux et al., 2004; Kramer et al., 2004).

2.3.3 The Bacterial Hsp70 System

The Hsp70 system is the most diverse group of molecular chaperones which occur both as constitutively expressed and stress-inducible forms (Chang et al., 2007). They are present in the cytosol of eubacteria, eukarya and some archaea, as well as within eukaryotic organelles, like mitochondria and endoplasmic reticulum. In addition to their role in *de novo* folding, they are also important players in protein trafficking and proteolytic degradation of terminally misfolded proteins (Bukau et al., 2006). Hsp70s generally work together with chaperones of Hsp40 (DnaJ) family and nucleotide-exchange factors (NEFs) in the ATP regulated binding and release of non-native proteins (Mayer et al., 2000).

Hsp70 chaperones are monomeric protein of ~70 kDa size. *E.coli* DnaK/J is the best studied Hsp70 system. The crystal structures of individual domains of Hsp70s have been solved. It is comprised of two functional domains: a 45 kDa amino-terminal ATPase domain (Harrison, 1997) and a 25 kDa carboxy-terminal polypeptide binding domain (Zhu et al., 1996). The substrate binding C-domain is further divided into β sandwich subdomain with a peptide-binding cleft and an α -helical latch like segment (Zhu et al., 1996) (Figure 2.10a). Recently a full length structure of bovine Hsc70 has been determined (Jiang et al., 2005) which allows further understanding of the inter-domain interaction and functional cycle of Hsp70.

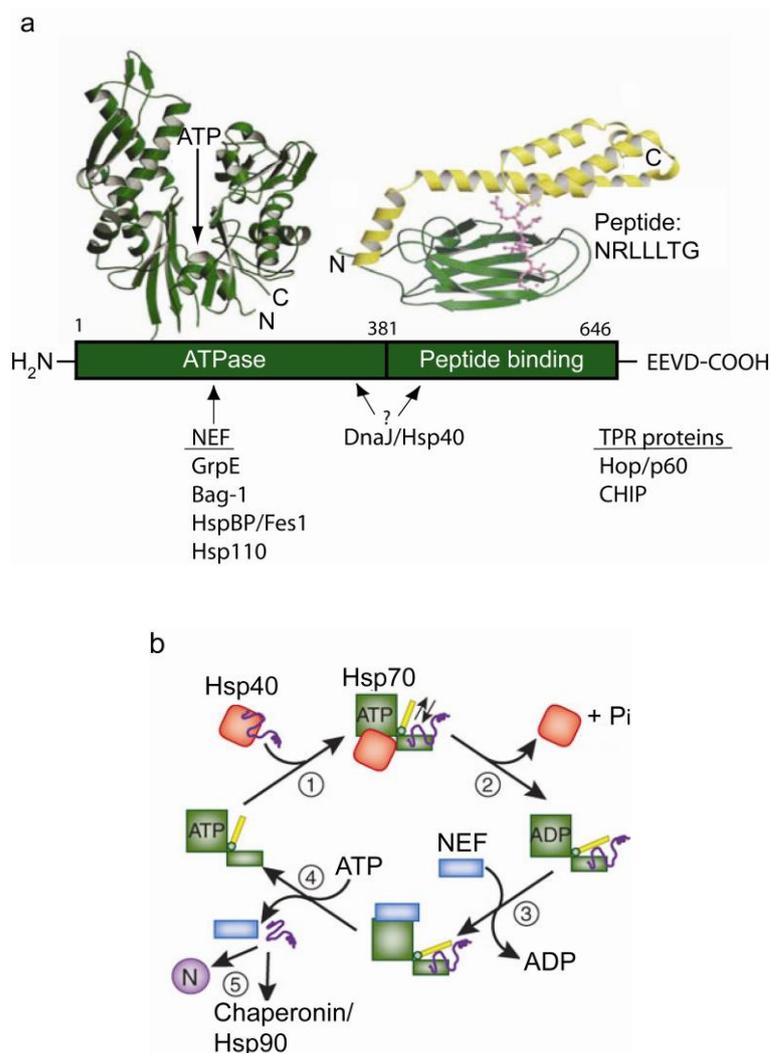


Figure 2.10: Structure and Reaction Cycle of Hsp70 System. (a) Structures of the ATPase domain (Harrison et al., 1997) and the peptide binding domain of Hsp70 (Zhu et al., 1996) shown for *E. coli* DnaK. The α -helical lid of the peptide binding domain is shown in yellow and the extended peptide substrate (amino acid sequence shown) as a ball- and -stick model in pink. ATP indicates the position of the nucleotide-binding site. Eukaryotic and prokaryotic co-factors that interact with Hsp70 are shown schematically. Residue numbers refer to human Hsp70. Only the Hsp70 proteins of the eukaryotic cytosol have the COOH-termina sequence EEVD, which is involved in binding of tetratricopeptide repeat (TPR) cofactors. (b) Reaction cycle of Hsp70. (1) substrate protein is transferred to ATP-bound Hsp70 by Hsp40. (2) Hsp40 accelerates ATP hydrolysis resulting in closure of α -helical lid and tight binding of substrate by Hsp70. Hsp40 dissociates from Hsp70. (3) NEF catalyzes the dissociation of ADP. (4) ATP binding induces opening of α -helical lid leading to release of substrate. (5) Released substrate either fold to a native state (N), is transferred to downstream chaperones or rebinds to Hsp70.

As shown in Fig 2.10a, the β sandwich domain recognizes extended (~7 amino acids residue) hydrophobic segment (Rudiger et al., 1997). The α -helical lid and the conformational change in the β -sandwich domain regulate the affinity state for polypeptide in an ATP-dependent manner (Mayer et al., 2000; Pellicchia et al., 2000). In the ATP-bound state, the lid adopts an open conformation resulting in low affinity for polypeptides i.e it binds and release substrates rapidly (low affinity state). ATP hydrolysis which is strongly accelerated by its co-chaperone DnaJ (Hsp40) leads to lid closure and stable polypeptide binding (high affinity state) (Figure 2.10b). Interaction of DnaK with substrates is mediated by the J-domain of DnaJ proteins (Mayer et al., 2000). DnaJ can also directly interact with unfolded polypeptides and recruit DnaK to protein substrates. The nucleotide exchange factor, GrpE, promotes the release of ADP from DnaK. The substrate and GrpE dissociates from DnaK upon subsequent ATP binding to DnaK (Schmid et al., 1994) and has the option of either folding, rebinding to DnaK and DnaJ, or being transferred to another chaperone system, like chaperonins, for final folding (Figure 2.10b).

Hsp70 mediated folding and prevention of aggregation may be explained by its ability to shield the exposed hydrophobic regions in polypeptides protecting it from aggregation thereby reducing the concentration of aggregation-prone species (Kinetic partitioning). Binding to Hsp70 may result in conformational remodeling (Schiene-Fischer et al., 2002), perhaps removing the kinetic barrier to fast folding.

2.3.4. The Chaperonins

Chaperonins are conserved class of large double ring complexes (Braig et al., 1994; Hartl, 1996). There are two groups of chaperonin, Group I and II, which are related in topology and share ~40% sequence similarity (Hartl, 2002; Maeder et al., 2005). Group I chaperonins (also called Hsp60s) found in bacteria (GroEL), mitochondria and chloroplasts are

tetradecamer of ~800 kDa with a seven-fold symmetry. They co-operate with heptameric ring shaped co-chaperone GroES (Hsp10) that form the lid of the folding cage in which polypeptide can be encapsulated during refolding (Mayhew et al., 1996; Weissman et al., 1996). Like Hsp70, substrate binding by chaperonin is regulated by ATP, but, unlike the Hsp70s, chaperonin can promote folding by multiple cycles of protein encapsulation in a sequestered environment. Our detailed understanding of the chaperonin is better defined for the Group I chaperonin of *E.coli* GroEL (explained later in the section).

2.4 Group I Chaperonin System: GroEL and GroES

GroEL was first identified by the observation that certain temperature sensitive mutations in the GroE operon were unable to support the growth of bacteriophage lambda and T4 (Georgopoulos et al., 1973). Later on it was found that GroEL and GroES are essential for *E.coli* growth at all temperatures (Fayet et al., 1989). Interestingly, the *in vitro* refolding efficiency of bacterial Rubisco was shown to be significantly improved by GroEL and GroES (Goloubinoff et al., 1989). Also mitochondrial Hsp60 was found to play a role in folding of proteins that were imported to mitochondria (Ostermann et al., 1989).

2.4.1 Architecture of GroEL and GroES

The structure of the GroEL/GroES complex has been extensively studied by electron microscopy (Braig et al., 1993; Langer et al., 1992; Saibil et al., 1993) and X-ray crystallography (Braig et al., 1994; Xu et al., 1997). GroEL is a homotetradecamer composed of two seven-membered rings of ~57 kDa subunits arranged in seven-fold rotational symmetry. Two rings are stacked back to back with a cavity in each ring.

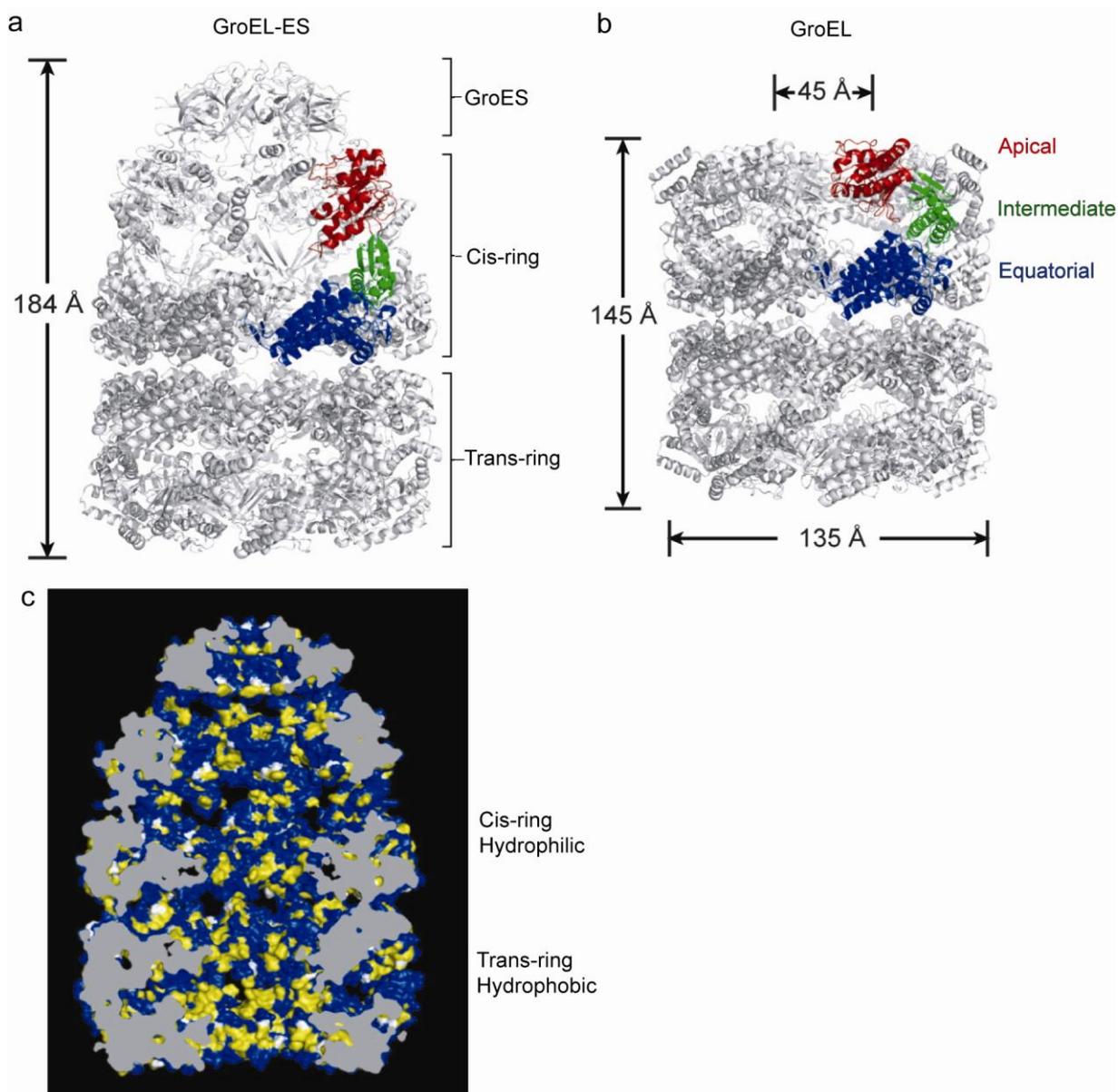


Figure 2.11: Architecture of Group I Chaperonins: (a, b) Crystallographic model of the *E.coli* GroEL and GroEL/GroES representing type I chaperonin in ADP-bound state. One of subunit has been colored to reveal the domain structure. The equatorial ATPase domain (blue) is linked to the substrate-binding apical domain (red) by a flexible hinge-like intermediate domain (green). GroES binds asymmetrically forming *cis*-GroEL ring while leaving the *trans*-GroEL ring open. (c) Space filling model to show a cutaway view to reveal the wall character of the central cavities (hydrophilic, blue; hydrophobic, yellow). All solvent excluded surfaces at the subunit interface are grey. The surface of the *trans*-GroEL is considerably more hydrophobic than the surface of *cis* GroEL (Xu et al., 1997).

The GroEL subunit (547 amino acid residues) folds into three distinctive domains: Apical, Intermediate and Equatorial domain (Figure 2.11).

- 1) The 'apical' domain contains the substrate and GroES binding sites. This domain is composed of an orthogonal β -sheet structure flanked at its inside and outside by α -helices. The apical domain is the least resolved of the three GroEL domains in crystal structure, suggesting an intrinsic flexibility that appears to be functionally necessary to accommodate binding of a large variety of different polypeptides.
- 2) A small 'intermediate' domain forms a hinge-like covalent connection between the apical and the equatorial domain
- 3) A well-ordered, highly α -helical, 'equatorial' domain provides the major contacts between subunits both within and between rings. It has been shown that the equatorial domain forms an important route for allosteric communication (Aharoni and Horovitz, 1996; Horovitz et al., 1994). It is the site for ATP binding and hydrolysis.

The crystal structures of GroEL resolve only 524 amino acid residues with the last 23 C-terminal residues unresolved in each subunit, indicating that GroEL C-terminal is flexible and disordered. Last 13 amino acid residues of the unresolved C-terminal tail contains highly conserved GGM repeats partly accounting for their disordered nature (Braig et al., 1994). Electron microscopy has indicated an axial mass at the equatorial level of each ring (Chen et al., 1994; Elad et al., 2007; Saibil et al., 1993). This block is apparently functional, as a polypeptide cannot escape through the equatorial segments of a single-ring mutant of GroEL (Rye et al., 1997; Weissman et al., 1996). Thus entry and exit of polypeptide appears to be restricted to the apical end of each ring.

GroEL works in conjunction with its co-chaperone GroES which is a heptameric ring of 10 kDa subunits (Hunt et al., 1996; Mande et al., 1996). GroES bind to GroEL in a nucleotide dependent manner (Chandrasekhar et al., 1986; Langer et al., 1992; Saibil et al., 1993; Schmid et al., 1994). Each GroES subunit has a core β -barrel structure consisting of two β -hairpin loops, one arching upward and inward forming the top of the dome, and second loop

interacts with the GroEL subunit, probably in a 1:1 ratio with residues in the apical domain (Fenton et al., 1996)

2.4.2. Polypeptide, Nucleotide and GroES Binding

GroEL binds to non-native substrate polypeptides that expose hydrophobic surfaces (Viitanen et al., 1992). In general, GroEL bound polypeptides appear to be in collapsed and loosely structured states, possessing conformation somewhere between fully unfolded and fully native state (Hayer-Hartl et al., 1994; Martin et al., 1991; Zahn and Pluckthun, 1994; Zahn et al., 1994b). Although the GroEL apical domain can recognize a range of local structures, including α -helices (Landry and Gierasch, 1991; Landry et al., 1992; Preuss et al., 1999; Wang et al., 1999), extended strands (Buckle et al., 1997) and β -hairpins (Chen and Sigler, 1999), the most important contacts appear to be hydrophobic. Structure guided extensive mutational analysis identified a set of hydrophobic residues in the apical domain (eight hydrophobic and one serine on helices H and I and a long loop), lining the wall of the central cavity, that are required for polypeptide binding (Braig et al., 1994; Fenton et al., 1994). However, some of these residues are also involved in binding the mobile loop of GroES as seen in the crystal structure of the GroEL/ES complex (Chaudhry et al., 2003; Xu et al., 1997). Recent cryo-EM studies have also shown non-native polypeptide to interact with residues positioned deep within the cavity, contacting 5-7 surrounding GroEL sub-units of an open GroEL ring (Elad et al., 2007; Kanno et al., 2009).

Given the atomic features of GroES and electron microscopy images of GroEL/ES binary complexes, it is evident that in GroEL/ES complexes, the interior cavity of GroES becomes continuous with that of GroEL cavity. Earlier NMR work showed that a mobile loop in GroES became ordered upon binding to GroEL (Landry et al., 1993). Correspondingly, this mobile loop region fail to resolve in the crystal structures of unbound GroES (Hunt et al.,

1996). Electron microscopy studies of the site of contact suggests that GroES mobile loops at least in part bind to and compete for the same surface of GroEL like non native substrate proteins (Chen et al., 1994; Roseman et al., 1996). Mutational findings suggest that the same cavity facing residues which are critical for polypeptide binding are also important for GroES binding (Fenton et al., 1994). Such apparent competition by GroES for the peptide binding sites results in the release of polypeptide in the so called '*cis-cavity*' which is capped by GroES, resulting in initiation of folding (Weissman et al., 1995a; Weissman et al., 1996).

Details of the ATP binding site and residues essential for hydrolysis became evident with the 2.8 Å structure of GroEL complexed with 14 ATPγS (Boisvert et al., 1996). The ATP binding site of each subunit lies in the top surface of the equatorial domain facing the cavity. In the absence of GroES stabilized conformational changes, the binding site is comprised of a set of highly conserved loops, including the phosphate-binding loop sequence 87-91 (GDGTT) between helices C and D (Kim et al., 1994). Mutation in Asp87 excludes the binding of ATP and completely abolishes ATPase activity. Mutations in the intermediate domain just beyond the lower hinge region, adjacent to the ATP binding pocket (residues 150, 151, 152, 405 and 406), abolishes ATPase activity similar to residue 383 near the upper hinge region (Fenton et al., 1994).

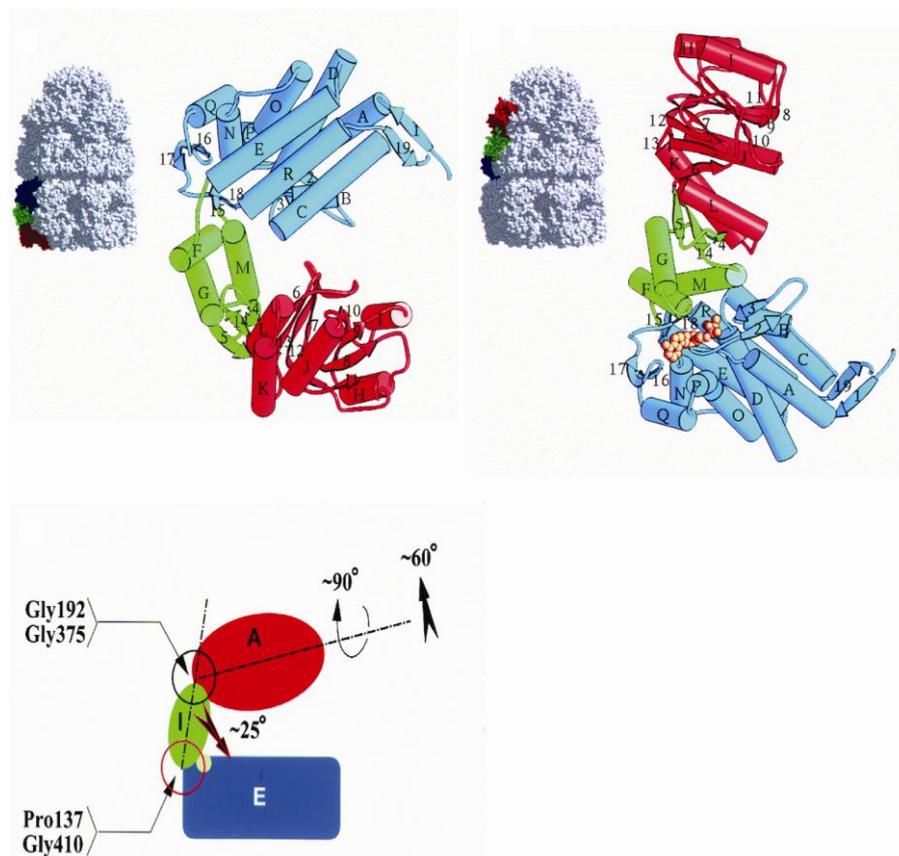


Figure 2.12: The Domain Movements within Individual Subunits of GroEL Ring: Ribbon diagram of an individual subunit of (a) *trans* and (b) *cis* GroEL. The orientation of the representative subunit is the same as the colored subunit in the respective space filling model. The equatorial, intermediate and apical domains are blue, green and red respectively. The yellow circle on the top of the equatorial domain represents nucleotide (c) Schematic representation of GroEL showing the en block movements that occur around the pivot points at the ends of the intermediate domain (adapted from (Xu et al., 1997).

Each ring of GroEL is an allosteric unit that binds and hydrolyzes seven molecules of ATP with positive intra-ring cooperativity (Gray and Fersht, 1991; Kafri et al., 2001) and negative inter-ring cooperativity (Kafri et al., 2001; Yifrach and Horovitz, 1994, 1995). At the structural level, binding of ATP in an open ring produces an initial allosterically mediated movement, a mild elevation and counterclockwise twist of the polypeptide binding apical domains (Ranson et al., 2001) optimizing the positioning of GroES binding sites which result in tight association of GroES ($>4 \times 10^7 \text{ M}^{-1}\text{s}^{-1}$) to the same ring (Bochkareva et al., 1992; Burston et al., 1995; Jackson et al., 1993) forming an ATP-bound *cis* complex. This

association is accompanied further by large rigid body movement of the apical domains that elevates them and turns them clockwise (Ranson et al., 1995; Xu et al., 1997) (Figure 2.12). Eventually, the domain rearrangements result in burying hydrophobic residues and hence changing the environment inside the GroEL/ES cavity to hydrophilic and driving the release of polypeptide in less than 1 s into the hydrophilic *cis* cavity. In addition the volume of the cavity is also approximately enlarged by two-fold (Hayer-Hartl et al., 1996; Mayhew et al., 1996; Roseman et al., 1996; Weissman et al., 1996; Weissman et al., 1995b). This hydrophilic cage can accommodate a single partially folded polypeptide up to ~60 kDa (Sigler et al., 1998; Viitanen et al., 1992). Hydrolysis of *cis*-bound ATP generates an ADP-bound *cis* complex that is stable until a disassembly signal is transmitted from the opposite *trans*-ring after ATP and substrate polypeptide binding (Rye et al., 1997; Rye et al., 1999; Todd et al., 1994).

In contrast to the cooperative binding of ATP, however, GroEL bind ADP in a non-cooperative manner and with 10-fold lower affinity (Cliff et al., 1999; Horovitz et al., 2001). GroES binding to GroEL is nucleotide dependent (ADP or ATP) (Burston et al., 1995; Todd et al., 1994). The extent of ligand induced conformational changes in GroEL follows the order ATP > AMP-PNP > ADP, implying highly stereo-explicit interactions with the β - γ phosphoanhydride of the triphosphate moiety. The same order is observed in the apparent rates and affinities of GroES binding to the different GroEL nucleotide complexes (Martin et al., 1991). But the productive folding of GroEL/ES dependent substrates like rhodanese, MDH, or Rubisco, occurs only in the presence of ATP and not ADP (Goloubinoff et al., 1989; Martin et al., 1991; Ranson et al., 1995).

2.4.3. Mechanism of GroEL/ES mediated Protein Folding

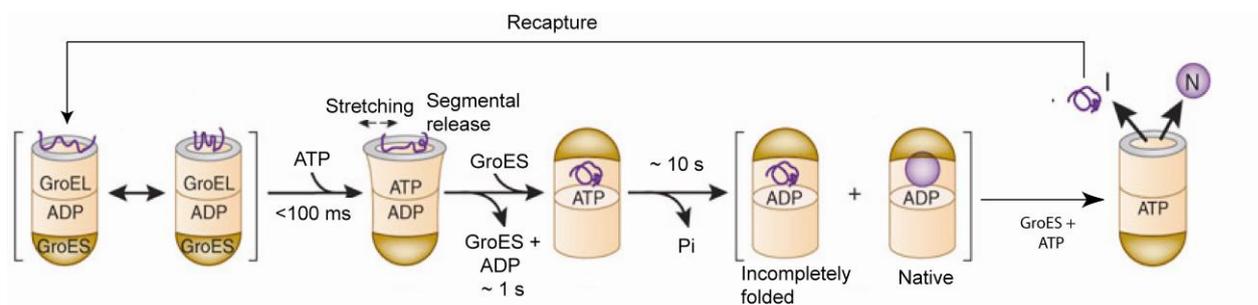


Figure 2.13: Schematic Representation of GroEL/GroES Assisted Folding: Unfolded substrate polypeptide binds to the *trans* ring of GroEL. ATP-dependent domain movement of the apical GroEL domains result in stretching of tightly bound regions of the substrate and in release and partial compaction of less stably bound regions of substrate. GroES binds after ATP binding leading to substrate encapsulation. Folding occurs in the chaperonin cage in the time scale ~ 10 s regulated by ATP hydrolysis rate. Upon binding of GroES and ATP in the other ring, GroES dissociates leading to substrate release. Incompletely folded substrate again binds to GroEL for the next folding attempt. Adapted from (Hartl and Hayer-Hartl, 2009).

The reaction cycle of GroEL/ES system have been investigated extensively (Hartl and Hayer-Hartl, 2009; Horwich et al., 2009). Under physiological conditions, GroEL is in the bullet form, the acceptor state of the GroEL for the protein (Rye et al., 1999). *Cis* ring of GroEL bullet has GroES and ADP bound and the *trans*-ring is in the substrate acceptor state. Unfolded polypeptide and ATP binds to the *trans*-ring of GroEL which triggers specific conformational changes resulting in the binding of GroES and subsequent release of bound polypeptide in the *cis* hydrophilic cavity. An encapsulated polypeptide of ~ 50 kDa is free to fold in this environment for 10-15 s at 25°C , the time needed for cooperative ATP hydrolysis. Subsequent ATP and GroES binding to the opposite GroEL ring causes the dissociation of GroES and release of substrate. Non-native protein is rapidly recaptured by GroEL for another folding attempt (Figure 2.13). Proteins that exceed the size limit of the chaperonin cage either use the Hsp70 system for folding (Agashe et al., 2004; Kerner et al., 2005) or may reach their

native state through binding and release from GroEL without encapsulation (*trans*-folding) (Chaudhuri et al., 2001), which is still an open question.

Several possible mechanisms have been proposed to explain the rate acceleration of folding of substrate proteins in presence of GroEL/ES system. Broadly, these mechanisms may be classified as active or passive. In a passive folding mechanism, GroEL is thought to have no direct effect on the conformation of a non-native protein, serving simply to block inhibitory aggregation and free energy of ATP hydrolysis is used only to assemble or disassemble an isolation chamber. An active model, by contrast, involves the direct modification of a substrate protein's accessible conformational space while folding. Still, GroEL mechanism of protein folding is debated.

Passive binding of an aggregation-prone intermediate by GroEL can block aggregation, but captured proteins must eventually be released back into the free solution in order to complete the final steps of folding or oligomer assembly. Folding takes place at infinite dilution, with each GroEL-associated protein physically isolated from other aggregation-prone proteins (Ellis, 1994a, b; Saibil et al., 1993). A GroEL ring is imagined to shift between high and low affinity states for a non-native substrate protein which is driven by ATP hydrolysis and GroES binding (Badcoe et al., 1991; Martin et al., 1991). This model of GroEL-mediated protein folding is known as the Anfinsen cage model (Ellis, 1994b; Saibil et al., 1993). Essentially, the enclosed cavity is envisioned to be near ideal, isolated environment where folding could proceed unhindered, propelled only by the intrinsic thermodynamic drive encoded by the protein amino acid sequence. But the purely passive model ignores the possibility that dominant kinetic traps could be misfolded states that have little or no possibility of accessing the native state.

One efficient way to bypass local energy minima is unfolding of substrate protein when bound to GroEL. Two general mechanisms, one thermodynamic and one catalytic, have been suggested to explain how GroEL could induce unfolding of substrate proteins. In the thermodynamic partitioning model (Zahn et al., 1994a; Zahn et al., 1994b), GroEL preferentially binds less folded conformations within an ensemble of non-native states and thereby shifts a pre-existing, intrinsic equilibrium towards less folded states without unfolding. In contrast, catalytic unfolding model suggests that GroEL could catalytically drive unfolding by lowering the free energy barriers that separate different folded states from one another. However, a direct connection between stimulated folding and the disruption of intra-molecular structure in a non native and stringent protein has been difficult to establish.

An alternate mechanism for unfolding was proposed by Lorimer and colleagues who suggested that substrate protein unfolding could be directly linked to the ATP-driven structural rearrangements of the GroEL ring itself (Lin et al., 2008; Shtilerman, 1999) which was termed as 'Forced Unfolding model' or 'Iterative Annealing model'. It suggests that the ATP-driven elevation and rotation of the GroEL apical domains that are necessary for GroES binding might apply a mechanical strain to a non-native protein bound across multiple GroEL apical domains. The forced protein unfolding in consecutive rounds of chaperonin binding promotes the reversal of kinetically trapped states. Consistent with such a mechanism, proteins undergo a conformational expansion on initial binding to GroEL and upon subsequent ATP binding (Lin et al., 2008; Sharma et al., 2008). A single ring mutant of GroEL, SR-EL, exists as a single, seven subunit ring that possess normal ATP binding, hydrolysis and GroES binding like GroEL (Weissman et al., 1995a; Weissman et al., 1996). Since SR-EL lacks the signal from the *trans* ring, it can only complete single round of binding and encapsulation of the substrate polypeptide. Remarkably, SR-EL is fully capable of driving productive folding of stringent substrate proteins (Brinker et al., 2001; Rye et al., 1997; Tang et al., 2006; Weissman

et al., 1996) without unfolding cycles. Also, many hydrogen-deuterium exchange experiments and FRET measurements failed to detect evidence for forced unfolding upon ATP and GroES binding to a substrate protein bound to SR-EL (Chen et al., 2001; Lin and Rye, 2004; Park et al., 2005). Also there is no evidence for the extensive and steady conformational progression predicted by continuous annealing.

On the other hand, the confinement model of protein folding suggests that the physical properties of the cavity are likely to have additional effects on folding beyond prevention of aggregation. Spatial restriction of a non-native protein may narrow a protein's folding funnel, physically excluding large regions of conformational space and confining the subsequent search for the native state to a smaller range of states. Effectively, both smoothening and narrowing the folding funnel, in theory, can result in accelerated folding. Recently, confinement of the substrate polypeptide has been shown to have direct influence on rate of protein folding (Brinker et al., 2001; Tang et al., 2006). Mutational analysis has been done to show that the polar residues of the cavity wall are crucial for rapid folding (Brinker et al., 2001; Lin and Rye, 2004; Tang et al., 2006). According to molecular dynamics simulations, these polar residues are expected to promote folding by accumulating ordered water molecules in their vicinity, thereby generating a local environment in which a substrate polypeptide is forced to bury exposed hydrophobic residues more effectively (England et al., 2008). Dynamic interaction between non-native folding intermediates and a weakly hydrophobic cavity wall has been suggested to accelerate folding by lowering the free energy barriers between different states (Betancourt and Thirumalai, 1999; Chan and Dill, 1996). This enhancement could be due to stabilization of more hydrophobic and less folded conformational transition states, providing an additional smoothening of the free energy landscape. Alternately, the GroEL/ES cavity has been suggested to increase ruggedness of a broad and slowly crossed transition state

ensemble, opening folding routes that are not significantly populated in bulk solution (Jewett et al., 2004).

2.4.4. Substrates of GroEL

In Vivo ~10-15% of *E. coli* cytosolic proteins have been shown to interact with GroEL (Ewalt et al., 1997; Houry et al., 1999). But *in vitro* GroEL is able to bind a large and diverse variety of non-native proteins, including some heterologous substrates like mitochondrial malate dehydrogenase, *R. rubrum* rubisco and bovine rhodanese. In a high-throughput proteomic analysis, approximately 250 proteins were identified that were found to functionally interact with GroEL. Most of the proteins are between 20 kDa and 50 kDa and have complex α/β or $\alpha+\beta$ domain topologies, with a distinct enrichment of the $(\beta/\alpha)_8$ TIM barrel fold (Kerner et al., 2005). The identified GroEL substrates have been grouped in three classes based on their increasing requirement for GroEL.

Class I proteins (e.g. enolase, 46 kDa; glyceraldehydes-3-phosphate dehydrogenase, 35 kDa) are largely independent of chaperone interaction but their folding yield can be increased by chaperones. Consequently, they exhibit low aggregation propensity upon dilution from denaturant.

Class II proteins (e.g. glutamate decarboxylase α , 53 kDa) do not refold spontaneously under standard conditions, due to rapid aggregation. They can utilize either the Hsp70 (DnaK/DnaJ/GrpE) or the GroEL/ES system for folding. Thus it is likely that DnaK and GroEL share a number of substrates mainly in the preferred size range of GroEL.

Class III substrates (e.g. 5, 10 methylenetetrahydrofolate reductase, MetF, 33 kD; DAPA, 31 kDa) are fully dependent on the GroEL/ES system for folding. The Hsp70 system can bind Class III proteins and prevent their aggregation, but folding is only achieved upon transfer to GroEL.

3. Aim of the Study

The GroEL/ES chaperonin system is structurally and mechanistically one of the best characterized molecular machines. The basic mechanism of GroEL/ES action involves encapsulation of a single molecule of non-native protein in the specialized nano-compartment, thereby allowing folding to occur unimpaired by aggregation (Mayhew et al., 1996; Weissman et al., 1996). In view of the fact that only a limited number of proteins are GroEL/ES dependent, it has been suggested that the physical environment of the chaperonin cage, in addition to providing a sequestered folding space, may actively rescue proteins from kinetic folding traps, thereby accelerating their folding speed (Hartl and Hayer-Hartl, 2009; Jewett and Shea, 2010; Brinker et al., 2001). Exploring the properties of the trapped folding intermediate from which the chaperonin catalyzes folding is essential for understanding the mechanism underlying this active process.

The aim of this study was to elucidate the basic mechanism of chaperonin action in accelerating folding. We performed a series of experiments to show that encapsulation in the chaperonin cage accelerates folding not by blocking transient aggregation, but by mediating the progression of a collapsed, yet disordered folding intermediate towards the native state. Cysteine mutants of DM-MBP were engineered to probe the similarity between the rate limiting step of folding inside the chaperonin cage and the effect of constraining the intermediate by introducing long-range disulfide bonds. In addition, the effect of the negatively charged cavity wall was also investigated for its impact in facilitating protein folding.

4. Materials and Methods

4.1. Materials

4.1.1. Chemicals

L-Amino acids	Sigma-Aldrich
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
Alexa Fluor 488 C ₅ maleimide	Molecular Probes
Ammonium persulfate (APS)	Sigma-Aldrich
Ampicillin	Merck
Amylose resin	New England Biolabs
Arabinose	Sigma-Aldrich
Atto532 C ₅ maleimide	Atto-Tech
Atto647-C ₅ maleimide	Atto-Tech
Bacto-agar	Difco
Bacto-tryptone	Difco
Bacto yeast extract	Difco
Bovine Serum Albumin (BSA)	Sigma-Aldrich
Bromophenol blue	Sigma-Aldrich
Calcium chloride	Merck
CDTA (trans-1,2-diaminocyclohexane-N,N,N',N'-tetracetic acid)	Sigma-Aldrich

Cholramphenicol	Sigma-Aldrich
Complete EDTA-free protease inhibitor	Roche
Coomassie brilliant blue R-250	Roth
Dextran 40	Sigma-Aldrich
Dimethyl sulfoxide (DMSO)	Merck
DTSSP [3,3'-dithiobis(sulfo-succinimidyl propionate)]	PIERCE
Dithiothreitol (DTT)	Roche
ECL TM detection kit	Amersham Pharmacia Biotech
Ethanol	Merck
Ethidium bromide	Biorad
Ethylenediaminetetraacetic acid-sodium salt (EDTA)	Merck
Ferric nitrate	Sigma-Aldrich
Flavin adenine dinucleotide (FAD)	Sigma-Aldrich
Ficoll 70	Sigma-Aldrich
Formaldehyde	Sigma-Aldrich
Glucose	Sigma-Aldrich
Glycerol	Merck
Glycine	Roth
Guanidinium hydrochloride (GuHCl)	Sigma-Aldrich
HEPES	Sigma-Aldrich
Hydrochloric acid (37%)	Merck
Imidazole	Merck
Isopropyl- β -D-thiogalactopyranoside (IPTG)	BioMol
Kanamycin	Sigma-Aldrich
Magnesium chloride	Merck
Maltose	Sigma-Aldrich

Menadione	Sigma-Aldrich
β -mercaptoethanol	Sigma-Aldrich
Methanol	Merck
Nickel-NTA agarose beads	Qiagen
Nicotinamide adenine dinucleotide (NADH)	Roche
Phosphoenol pyruvate (PEP)	Sigma-Aldrich
PIPES	Sigma-Aldrich
Phenyl-methyl-sulfonyl-fluoride (PMSF)	Sigma-Aldrich
Polyacrylamide/bisacrylamide solution 30% (30 : 0.8)	Roth
Polyethylene glycol 2000 (PEG)	Merck
Potassium cyanide	Sigma-Aldrich
Potassium hydroxide	Sigma-Aldrich
D-ribulose 1,5-diphosphate (RuDP)	Sigma-Aldrich
Silver nitrate	Sigma-Aldrich
Sodium chloride	Merck
Sodium [C ₁₄] bicarbonate	Amersham Pharmacia Biotech
Sodium dodecylsulfate (SDS)	Sigma-Aldrich
Sodium hydroxide	Sigma-Aldrich
Sodium thiosulfate	Merck
Spectinomycin	Sigma-Aldrich
Sucrose	Merck
N,N,N',N' Tetramethylethylenediamine (TEMED)	Sigma-Aldrich
Tris(2-carboxyethyl)phosphine hydrochloride (TCEP-HCl)	PIERCE
Tris-base	Sigma-Aldrich
Triton X-100	Sigma-Aldrich
Tween-20	Calbiochem

4.1.2 Enzymes

Apyrase	Sigma-Aldrich
Benzonase	Merck
Lysozyme	Sigma-Aldrich
<i>Pfu</i> DNA polymerase	Stratagene
Pyruvate kinase/Lactate dehydrogenase	Sigma-Aldrich
Restriction enzymes	New England Biolabs
Rhodanese	Sigma-Aldrich
Shrimp Alkaline Phosphatase	Roche
T ₄ DNA ligase	New England Biolabs
Vent DNA polymerase	New England Biolabs

4.1.3. Materials

Centricon 10 kDa cut off	Amicon
Centricon 50 kDa cut off	Amicon
Microcon 100 kDa cut off	Amicon
Nitrocellulose transfer membrane	Whatman Schleicher and Schuell
Sterile filter 0.22 µm	Millipore
Sterile filter 0.45 µm	Millipore

4.1.4. 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 J-20XP, J-6B, GS-6R	Beckman
Centrifuges 5415C and 5417R	Eppendorf

Chromatographic 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
Electrophoresis chamber 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
Gene Pulser II electroporation system	Bio-Rad
Gilson Pipetman (2, 10, 20, 100, 200, 1000 µl)	Eppendorf
Incubators Innova 4430	New Brunswick Scientific
Luminescent Image Analyzer LAS-3000	FUJIFILM
Mini Trans-Blot Electrophoretic Transfer Cell	Bio-Rad
PCR-Thermocycler T3	Biometra
pH meter Accumetic Basic	Fisher Scientific
Sonicator Ultrasonic Processor XL	Misonix Inc.
Spectrophotometer DU 640 UV/VIS	Beckmann
Spectrophotometer LS50	Perkin-Elmer
Synergy HT UV/VIS/Fluorescence/Luminescence Plate Reader	Bio-Tek
Sx.18MV Stopped-Flow Reaction Analyser	Photo Physics
UV/VIS Spectrometer V-560	Jasco
Thermomixer Comfort	Eppendorf
Vortex	Ikamag
Water bath	Bioblock Scientific

4.1.5. 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 MgCl₂. After autoclave, add 20 ml of filter sterilized 1M glucose (Sambrook et al., 1989)

4.1.6. Antibiotic Stock Solutions

Antibiotic additives to growth media were prepared as 1000x stock solutions and filter-sterilized before usage:

Ampicillin: 100 g/l

Chloramphenicol: 25 g/l

4.2. Bacterial Strains and Plasmids

4.2.1. *E.coli* strains

DH5 α F' F'/*endA1 hsdR17* (r_k⁻, m_k⁺) *glnV44 thi-1 recA1 gyrA* (NA1^r) *relA1 Δ (lacIZYA-argF) U169 deoR* (ϕ 80*dlac* Δ (*lacZ*)M15)

BL21(DE3) Gold B strain, F-*dcm*+ *The ompT hsdS* (r_B-m_B-) gal 1 (DE3) *endA Tet^R*
(Stratagene)

4.2.2. Plasmids

GroEL and all chaperonin mutants were constructed in pCH vector backbone (Chang et al., 2005) inserted via *NdeI* and *NheI* sites. Synthetic oligonucleotides encoding wild-type or

mutant C-terminal extensions of GroEL were introduced into the pCH-ELAC or SR-ELAC plasmids between the *NheI* and *HindIII* sites. EL-KKK2, SR-EL KKK2, ELGGA, SR-EL GGA were constructed in a pCH vector backbone by Yun Chi Tang (Tang et al., 2006). MBP cysteine mutants were generated by site-directed-mutagenesis in pCH vector on the background of the DM-MBP. SM-MBP (Y283D), DM-MBP (V8G, Y283D), WT-MBP were generated by Yun Chi Tang (Tang et al., 2006) and DM-MBP A52C/P298C, DM-MBP K175C/P298C were made by Shruti Sharma (Sharma et al., 2008). GroES was constructed in a pET11a vector inserted via *NdeI* and *BamHI* sites (Brinker et al., 2001).

Plasmid	Promotor/Origin	Selection Marker
GroEL and GroES		
pCH-EL	T7/ColE1	Ampicillin
pCH-SR	T7/ColE1	Ampicillin
pET11a-ES	T7/ColE1	Ampicillin
pCH-EL GGD	T7/ColE1	Ampicillin
pCH-EL GGG	T7/ColE1	Ampicillin
pCH-EL GGI	T7/ColE1	Ampicillin
pCH-EL GGK	T7/ColE1	Ampicillin
pCH-EL GGQ	T7/ColE1	Ampicillin
pCH-EL GGS	T7/ColE1	Ampicillin
pCH-EL GGY	T7/ColE1	Ampicillin
pCH-SR GGD	T7/ColE1	Ampicillin
pCH-SR GGG	T7/ColE1	Ampicillin
pCH-SR GGI	T7/ColE1	Ampicillin
pCH-SR GGK	T7/ColE1	Ampicillin
pCH-SR GGQ	T7/ColE1	Ampicillin

pCH-SR GGS	T7/ColE1	Ampicillin
pCH-SR GGY	T7/ColE1	Ampicillin
DM-MBP mutants		
pCH-DM-MBP N18C/D296C	T7/ColE1	Ampicillin
pCH-DM-MBP D184C/K362C	T7/ColE1	Ampicillin
pCH-DM-MBP N18C/D296C/D184C/K362C	T7/ColE1	Ampicillin
pCH-DM-MBP A21C/V37C	T7/ColE1	Ampicillin
pCH-DM-MBP E22C/A292C	T7/ColE1	Ampicillin
pCH-DM-MBP D65C/M330C	T7/ColE1	Ampicillin
pCH-DM-MBP A109C/L299C	T7/ColE1	Ampicillin
pCH-DM-MBP D164C/K251C	T7/ColE1	Ampicillin
pCH-DM-MBP D184C/Q365C	T7/ColE1	Ampicillin
pCH-DM-MBP A206C/A351C	T7/ColE1	Ampicillin
pCH-DM-MBP N218C/S238C	T7/ColE1	Ampicillin
pCH-DM-MBP V244C/R316C	T7/ColE1	Ampicillin
pCH-DM-MBP P254C/K326C	T7/ColE1	Ampicillin
pCH-DM-MBP K189C/D358C	T7/ColE1	Ampicillin
pCH-DM-MBP K14C/K6C	T7/ColE1	Ampicillin
pCH-DM-MBP D15C/K362C	T7/ColE1	Ampicillin

The following positions were not active upon mutation and could not be purified via affinity purification: pCH-DM-MBP A21C/V37C, pCH-DM-MBP E22C/A292C, pCH-DM-MBP D65C/M330C, pCH-DM-MBP A109C/L299C, pCH-DM-MBP D164C/K251C, pCH-DM-MBP D184C/Q365C, pCH-DM-MBP A206C/A351C, pCH-DM-MBP N218C/S238C, pCH-DM-MBPV244C/R316C, pCH-DM-MBPP254C/K326C, pCH-DM-MBP K189C/D358C, pCH-DM-MBP K14C/K6C, , pCH-DM-MBP D15C/K362C.

GroEL/SR-EL GGQ was not used for further studies.

4.3. Molecular Cloning Methods

4.3.1. 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 (OD_{600}) 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 grown (typically 10-16 h).

Alternatively, electroporation was applied to improve the transformation efficiency. Electrocompetent cells were prepared as follows: 500 ml bacterial culture was grown to an optical density (OD_{600}) of 0.8 in LB medium at 37°C. The cells were washed carefully with 250 ml ice-cold sterilized water for two times and finally the cells were pelleted and resuspended in 2 ml of ice-cold 10% glycerol. 40 µl aliquots were frozen in liquid nitrogen and stored at -80°C. For electroporation transformation, competent cells (40 µl) were mixed with 1-2 µl plasmid DNA (or ligation product) and transferred into a 0.2 cm Gene Pulser cuvette. The electroporation was done at 2.5 kV, 25 µFD and 200Ω settings with a Gene Puser

II electroporation device. The transformed cells were allowed to recover in 1 ml of SOC medium with 225 rpm shaking at 37°C for 1 h. The cell suspension was then plated on selective plates and incubated until colonies had developed (Dower *et al.*, 1988).

4.3.2. Plasmid Purification

LB medium containing the appropriate antibiotic was inoculated with a single *E. coli* colony harboring the DNA plasmid of interest and shaken for 8–16 h at 37°C. Plasmids were isolated using the QIAprep Spin Miniprep Kit or QIAGEN Plasmid Midi Kit (Qiagen) according to the manufacturer's instructions.

4.3.3. PCR Amplification

PCR (polymerase chain reaction)-mediated amplification of DNA was performed according to a standard protocol with minor modifications:

DNA Template	10-20 ng (plasmid DNA) 250 ng or less (bacterial genomic DNA)
Primers	20 pmole each
dNTPs	200 μ M each
Pfu DNA Polymerase	2.5U
Polymerase buffer	1X
Additives	3-6% DMSO if GC content was > 50%
Final Volume	50 μ l

Cycling conditions (35 cycles):

Initial denaturation	94°C, 5 min
Cycle denaturation	94°C, 30-60 s
Annealing	~55°C, 30-60 s
Extension	72°C, duration dependent on template length 1kbp/min
Final extension	72°C, 10 min
Stored at 4°C or -20°C	

PCR products were further purified using the QIAquick PCR purification and gel extraction kits (Qiagen) according to the manufacturer's instructions.

4.3.4. DNA Restriction and Ligation

DNA restriction was performed according to the manufacturer's instructions (New England Biolabs) of the respective enzymes. Typically, a 50 µl reaction contained 1-2 µl of each restriction enzyme and 0.5-2 µg purified PCR product or 1-5 µg plasmid DNA in the appropriate reaction buffer. Digested vector DNA was dephosphorylated with shrimp alkaline phosphatase.

For ligation, 100-200 ng (~1-2 µl) dephosphorylated vector DNA, 100-200 ng (~5-10 µl) DNA insert and 1 µl (100 U) T₄ ligase were incubated in ligase buffer at 25°C for 1 h or, for increased efficiency, at 16°C overnight. The ligation product was transformed into competent *E. coli* DH5α cells as described.

4.3.5. DNA Analytical Methods

DNA concentrations were measured by UV absorption spectroscopy at $\lambda = 260$ nm. A solution of 50 $\mu\text{g/ml}$ of double stranded DNA in H_2O exhibits approximately $A_{260 \text{ nm}} = 1$.

Agarose gel electrophoresis was performed in TAE buffer (40 mM Tris, pH 7.5, 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 Medigenomix GmbH (Martinsried, Germany) or Sequiserve (Vaterstetten, Germany).

4.4. Protein Purification

GroEL D87K, RuBisCo and RuBisco mutant K168E from *R. rubrum* were obtained from the Hartl laboratory collection of purified proteins.

4.4.1. 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 flowthrough 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 μ M 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 N₂ and stored at -80 °C. Total yield of GroEL was typically ~600 mg.

4.4.2. 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 ~15 mg/ml based on extinction coefficient of GroES ($\epsilon_{280}=8,943 \text{ M}^{-1}\text{cm}^{-1}$). Aliquots were frozen in liquid N₂ and stored at -80°C. Total yield of GroES was typically ~400 mg.

4.4.3. 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 2 l LB medium containing 100 mg/l ampicillin at 37°C to OD₆₀₀ of ~0.1. The induction was then carried out by adding 0.2 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 (30 mM Phosphate buffer, pH 7.5, 150 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 3 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 (30 mM Phosphate buffer, 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 MBP from 2 l culture was ~100 mg.

4.4.4. MetF Expression and Purification

MetF was expressed at 30°C in *E. coli* cells harboring elevated levels of GroEL and GroES (Kerner *et al.*, 2005). Cells were harvested and resuspended in running buffer (50 mM Tris-HCl, pH 7.3, 300 mM NaCl) containing Complete EDTA-free protease inhibitor (1 tablet/25 ml). Lysis was achieved by homogenization of the cell suspension in an EmulsiFlex high pressure homogenizer device and the lysate was cleared by centrifugation at 40,000 x *g* for 1 h, 4°C. The supernatant was applied to a 5 ml HiTrap metal chelating column pre-charged with Ni²⁺. The column was washed with a gradient of 10 to 50 mM imidazole in running buffer (for over 10 column volumes) and the proteins were eluted with 250 mM imidazole in running buffer. Fractions containing MetF were dialyzed in 20 mM MOPS-KOH, pH 7.4, 200 mM NaCl. MetF concentration was determined based on the absorption of bound FAD at 447 nm ($\epsilon_{280}=14,300 \text{ M}^{-1}\text{cm}^{-1}$) (Sheppard *et al.*, 1999). Proteins were aliquoted, flash-frozen in liquid N₂ and stored at -80 °C.

4.4.5. Rhodanese Preparation

Bovine mitochondrial rhodanese was purchased from Sigma (R-1756). 11.5 mg rhodanese was dissolved in 600 µl rhodanese buffer (20 mM MOPS-NaOH, pH 7.4, 50 mM NaCl). Impurities were removed by ultracentrifugation for 20 min at 4°C and ~20,000 x *g*. The supernatant was transferred to a new vial and the protein concentration was determined using extinction coefficient ($\epsilon_{280}=59,840 \text{ M}^{-1}\text{cm}^{-1}$). Proteins were aliquoted and stored at -80°C.

4.5. Protein Analytical Methods

4.5.1. 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 (Gill and von Hippel, 1989) 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.

4.5.2. SDS-PAGE

SDS-Polyacrylamide gels were prepared as follows:

Chemicals	Stacking gel		Separating gel	
	4%	10%	12%	15%
30% Acrylamide (0.8% bis)	6.5 ml	16.7 ml	20 ml	25 ml
0.5 M Tris, pH 6.8	12.5 ml	–	–	–
1.5M Tris, pH 8.8	–	12.5 ml	12.5 ml	12.5 ml
10% SDS	0.5 ml	0.5 ml	0.5 ml	0.5 ml
2 M Sucrose	12.5 ml	12.5 ml	12.5 ml	12.5 ml
H ₂ O (Upto 50 ml)	30.5 ml	7.8 ml	4.5 ml	–
TEMED	50 μ l	25 μ l	25 μ l	25 μ l
10% APS	500 μ l	500 μ l	500 μ l	500 μ l

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, pH6.8, 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.

4.5.3. Western-Blotting

Proteins were separated by SDS-PAGE and transferred to nitrocellulose membranes in a Mini Trans-Blot Electrophoretic Transfer Cell (Bio-Rad) in 25 mM Tris, 192 mM glycine, 20% methanol, pH 8.4 at constant current of 150 mA/gel for 1 h (Towbin *et al.*, 1979). Nitrocellulose membranes were blocked in 5% skim milk powder in TBST (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.05% Tween 20) for 1 h. The membranes were then incubated with a 1:2000 – 1:10,000 dilution of primary antibody serum in TBST for 2 h and extensively washed in TBST before incubation with a 1:5000 (for anti-mouse IgG) or 1:10,000 (for anti-rabbit IgG) dilution of secondary antibody in TBST (anti-rabbit IgG and anti-mouse IgG, whole molecule – horseradish peroxidase conjugate, Sigma-Aldrich). After extensive washing, protein bands were detected by incubating the membranes with ECL chemiluminescence solution and exposure to X-ray film (High performance chemiluminescence film) or a Luminescent Image Analyzer LAS-3000 system.

4.6. GroEL Functional Activity Assays

4.6.1. ATPase Assay

A coupled ATP regenerating enzyme system was employed following a previously described method (Todd and Lorimer, 1995; Yifrach and Horovitz, 1994). Buffer (20 mM Tris-HCl, pH 7.5, 200 mM KCl, 5 mM MgCl₂) containing 1.1 mM phosphoenolpyruvate, 700-1000 U/ml Pyruvate Kinase/Lactate Dehydrogenase mixture (Sigma), and 0.25 mM NADH was incubated with 1 mM ATP for 2 min at 25°C. The reaction was initiated by the addition of GroEL (0.2 μM oligomer). When ATPase assay was performed in the presence of GroEL and GroES, then GroES (0.4 μM oligomer) was added before adding ATP to the reaction mixture. The kinetics of the ATPase activities of GroEL was measured by following a decrease in the absorbance of NADH at 340 nm for 10 minutes. ATPase rate was calculated by determining $\Delta A_{340}/\Delta t$.

4.6.2. Aggregation Prevention Assay of Denatured Rhodanese

Rhodanese (25 μM) was denatured for 1 h at 25°C in denaturing buffer (6 M GuHCl, 20 mM Tris-HCl, pH 7.5, 20 mM KCl, 5 mM MgCl₂, 5 mM DTT) and diluted 100-fold into buffer A (20 mM Tris-HCl, pH 7.5, 200 mM KCl, 5 mM Mg(OAc)₂) alone or into buffer A containing 0.25 μM GroEL (oligomers). Aggregation was monitored by light scattering at 320 nm (Weber *et al.*, 1998). Values indicating aggregation in buffer A after 10 min is set to 100%.

4.7. *In vitro* Protein Refolding and Activity Assays

Protein refolding reactions containing chaperones (when present) were carried out with the following molar concentration ratios of chaperones to substrate, unless ratio is indicated:

1 substrate (monomer): 2 GroEL (tetradecamer) : 4 GroES (heptamer)

1 substrate (monomer): 4 SR-EL (tetradecamer) : 8 GroES (heptamer).

4.7.1. MBP Refolding

Generally, DM-MBP and its cysteine mutants (25 μ M) oxidized or reduced were denatured in 6 M GuHCl, 20 mM Tris-HCl, pH 7.5, 20 mM KCl/ 200 mM KCl, 5 mM Mg(OAc)₂ or 9M Urea, 100 mM HEPES, pH 7.2, 20 mM KOAc/ 200 mM KOAc, 5 mM Mg(OAc)₂ with or without 5 mM DTT for 1 h at 25°C and refolded upon 100-fold dilution into high salt buffer A (20 mM Tris-HCl, pH 7.5, 200 mM KCl, 5 mM Mg(OAc)₂) or low salt buffer B (20 mM Tris-HCl, pH 7.5, 20 mM KCl, 5 mM Mg(OAc)₂) or chloride free buffer C (100 mM HEPES, pH 7.2, 20 mM KOAc, 5 mM Mg(OAc)₂) in the absence or presence of chaperones at the concentration indicated in the figure legends. GroEL/ES assisted refolding was initiated at 25°C by the addition of 5 mM ATP. Refolding experiments were also carried out at different final GuHCl concentrations (15-60 mM), different TMAO concentrations (0-250 mM), different temperatures (15-25°C) and at different KCl concentrations (0-400 mM KCl) wherever indicated. Refolding was monitored (295 nm excitation, 345 nm emission) by following the increase in intrinsic tryptophan fluorescence on a Flouorolog spectrofluorometer (FL3-22, Spex) for DM-MBP concentrations of 10 nM to 1.5 μ M (Tang et al.,2006), taking advantage of the absence of Trp residues in GroEL, SR-EL and GroES (Martin et al.,1991).

Refolding of DM-MBP cysteine mutants under reducing conditions was performed in presence of 5 mM DTT. Oxidized cysteine mutants were prepared by removing DTT from

native protein by Micro-Biospin 6 columns (Biorad) equilibrated with buffer A or B and oxidizing them for 1 hr at 25°C in the presence of 50 μM CuCl_2 . In experiments where the effect of oxidation was measured after dilution of the reduced protein from denaturant, DTT was removed from the denatured protein by passing through Micro-Biospin 6 columns, 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 (5 μM CuCl_2) within 5 s.

4.7.2. MetF Refolding

MetF (50 μM) was denatured with 6 M GuHCl in buffer A containing 10 mM DTT for 1 h at 25°C and refolded upon 100-fold dilution into buffer A containing 50 μM FAD and 1 g/L BSA at 25°C in the absence or presence of chaperones. A 2-fold molar excess of GroES over GroEL was added and refolding initiated with 5 mM ATP. At the times indicated, aliquots were taken and reactions stopped with 40 mM CDTA. MetF activity was measured at 25°C using an NADH-menadione oxidoreductase assay as described previously (Sheppard et al., 1999).

4.7.3. Rhodanese Refolding

Rhodanese (50 μM) was denatured for 1 h at 25°C in denaturing buffer (6 M GuHCl, 20 mM Tris-HCl, pH 7.5, 20 mM KCl, 5 mM MgCl_2 , 5 mM DTT) and refolded upon 100-fold dilution into buffer A or B supplemented with chaperones as indicated in the figure legends. A 2-fold molar excess of GroES over GroEL was added and refolding initiated with 5 mM ATP. At the times indicated, further refolding was stopped with 50 mM CDTA and rhodanese activity measured at 460 nm at 25°C, as previously described (Hayer-Hartl et al., 1996; Martin et al., 1991).

4.7.4. Rubisco Refolding

Spontaneous and GroEL-assisted refolding of Rubisco at final Rubisco concentrations of 10 nM, 25 nM, 50 nM and 100 nM at 9.5°C and 15°C were performed as described previously (Brinker et al., 2001) with minor modifications. Typically, 1 µM, 2.5 µM, 5 µM, 10 µM Rubisco were denatured in 6 M GuHCl, 20 mM Tris-HCl (pH 7.8), 1 mM EDTA, 50 mM NaHCO₃, 200 mM βME (or 10 mM DTT), and 10% glycerol for 1 hr at 25°C respectively, and diluted 100-fold into spontaneous refolding buffer 20 mM Tris-HCl, pH 7.8, 250 mM KCl, 5 mM Mg(OAc)₂, 1mg/ml BSA and 200 mM β-ME (or 10 mM DTT) or into assisted refolding buffer (20 mM Tris-HCl, pH 7.5, 20 mM KCl, 5 mM Mg(OAc)₂, 1mg/ml BSA and 200 mM β-ME (or 10 mM DTT)). Note that the presence of BSA was without effect on the rate of folding but improved the yield, presumably by preventing the absorption of Rubisco to tube walls. For the spontaneous reaction, 50-fold molar excess of Rubisco mutant K168E was added to accelerate assembly. At the times indicated, aliquots (40 µl) were withdrawn and rapidly mixed with 10-fold molar excess of GroEL mutant D87K to trap the denatured Rubisco. Assisted refolding was carried out in the presence of 2 or 10-fold molar excess of GroEL and initiated by adding 2-fold molar excess of GroES over GroEL and 5 mM ATP. Aliquots (40 µl) were rapidly mixed with solution containing 7.5 mM CDTA and apyrase (10 U) to stop GroEL action and Rubisco folding.

Rubisco enzyme activity was determined after incubation at 25°C for 1.5 hr (for 10 nM and 25 nM Rubisco) or 1 hr (for 50 nM and 100 nM Rubisco) as described by Goloubinoff et al. (1989) with modifications. 50 mM Mg(OAc)₂ was added to the enzyme assay to compensate for the presence of CDTA. To determine the yield of refolding in these experiments, native Rubisco was treated by the same procedure, except that GuHCl was omitted.

4.8. Biochemical and Biophysical Methods

4.8.1. Thiol-Mediated Labeling of the Cys Constructs

Double cysteine constructs of DM-MBP in which one of the positions was 298 could be labeled specifically as this position shows a differential accessibility in presence or absence of maltose. MBP cysteine mutants (100 μ M) were labeled in PBS buffer (30 mM Phosphate buffer, pH 7.8, 150 mM NaCl) with 500 μ M malto- triose (Sigma) for 2 hrs at 25°C in the presence of 1.1 fold molar excess of the flourophores Atto532 maleimide, (ATTO-TEC, Inc). Unbound flourophores was removed by keeping it for dialysis in 3L of PBS buffer pH 7.8 for 12- 16 hours at 4°C and also by using 10 kDa centricon till dye fluorescence is not visible in flowthrough. Second step labeling was carried out at similar conditions with 2 fold molar excess of Atto647N maleimide, (ATTO-TEC, Inc). Excess dye was removed again using 10 kDa centricon. The coupling efficiency measured by the absorption of MBP ($\epsilon_{280} = 69 \text{ mM}^{-1}\text{cm}^{-1}$), and Atto 532 ($\epsilon_{280} 532 = 115 \text{ mM}^{-1}\text{cm}^{-1}$) was more than 90%.

4.8.2. Characterization of Cysteine Pair Mutants of DM-MBP

Oxidized and reduced states of DM-MBP cysteine mutants were differentiated by distribution of charge states when subjected to LC-MS analysis. To analyse the speed of disulfide bond formation, the charge state distribution of reduced and oxidized DM-MBP, 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 CuCl_2 after 30 s and inhibition of refolding and disulfide formation 5 s later by addition of 6 M GuHCl pH 2.5 (final GuHCl 4.8 M). Samples were immediately subjected to LC-MS as described later in the section. Quenching of disulfide bond formation was efficient because no oxidized protein was detected when CuCl_2 was added after 6 M GuHCl pH 2.5.

Oxidized and reduced states of cysteine mutants of DM-MBP also showed a different migration pattern on 12% SDS gel with coomassie staining and western blotting against DM-MBP.

4.8.3. Denaturant Dependent Unfolding and Refolding of MBP

Unfolding and refolding of WT-MBP, SM-MBP, DM-MBP and its cysteine mutants (Oxidized or reduced) was done at different concentrations of denaturant (GuHCl or Urea) as indicated in the figure legends. To follow the unfolding at different denaturant concentrations, native MBP and MBP mutants were incubated for 12 hrs with either increasing concentrations of GuHCl (~60 mM - 2 M) in buffer A or buffer B with or without 5 mM DTT or increasing concentrations of Urea (~0.2 - 6 M) in different buffers (buffer C, buffer D (100 mM HEPES, pH 7.2; 200 mM KOAc, 5 mMg(OAc)₂), buffer E (100 mM HEPES, pH 7.2; 20 mM KCl, 5mMg(OAc)₂), buffer F (100 mM HEPES, pH 7.2; 200 mM KCl, 5 mMg(OAc)₂). For monitoring refolding, MBP and MBP mutants (100 μM) were either unfolded in 6M GuHCl and then diluted 100 fold in buffer A or unfolded in 9 M Urea and then diluted 100 fold in buffer C, D, E or F containing increasing concentration of denaturant (GuHCl or Urea), followed by incubation for 12 hrs at 25°C. The fraction of folded protein was determined at 25°C by monitoring intrinsic Trp fluorescence with excitation at 295 nm and emission at 345 nm.

4.8.4. Circular Dichroism Measurements

CD wavelength scans of native, denatured and 0.5 M GuHCl intermediate state of DM-MBP (2 μM) were measured at 25°C in 0.1 cm cuvettes with scan speed of 50 nm/min in Jasco J-715 spectropolarimeter equipped with Peltier-thermostat. The percentage of secondary structure was determined using Contin software (Sreerama and Woody, 2004). Denaturant dependent unfolding and refolding of DM-MBP was carried out like Trp fluorescence

measurement. The fraction of folded protein was determined at 25°C by monitoring Far-UV circular dichroism (CD) at 220 nm (Jasco J-715 spectropolarimeter equipped with Peltier-thermostat) using 0.2 cm cuvettes for 1 μ M protein.

4.8.5. Unfolding Kinetics by Stopped-Flow Measurements

All stopped-flow experiments were done by using an Applied Photo Physics SX.18MV Stopped-Flow Reaction Analyzer with a 1:1 mixing ratio at 25°C. The rate of unfolding of WT-MBP, DM-MBP and reduced (red) and oxidized (ox) DM-MBP cysteine mutants at various concentrations of GuHCl (1.5-3 M) was monitored by following the decrease in Trptophan (Trp) fluorescence at 345 nm with an excitation at 295 nm. The dead time of the instrument was ~3 ms. 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.

4.8.6. Trptophan Fluorescence Wavelength Scan

Wavelength scan (310 – 410 nm) for Trptophan (Trp) fluorescence (excitation 295 nm and emission 345 nm) was performed for native, denatured and 0.5 M GuHCl intermediate of WT-MBP and DM-MBP (250 nM) in buffer B at 25°C on a Fluorolog 3 Spectrofluorometer (Spex).

4.8.7. Proteinase K Protection of GroEL/ES Substrate Complex

Rhodanese, DM-MBP or RuBisCo (25 μ M each) was denatured in GuHCl as described and diluted 100-fold into buffer A or B in the presence of a 2- or 4-fold molar excess of GroEL or SR-EL, respectively, at 25°C. Treatment with proteinase K (2 μ g/ml) was followed for 0-20 min. At the indicated times, sample was removed and further proteolysis was stopped with 1.5 mM PMSF. Protease protection of substrate protein by GroEL/ES

complex was determined by immunoblotting and quantification by Aida Image Analyzer v. 3.52.

4.8.8. Maltose Binding Experiments

To analyze the ability to bind maltose, purified DM-MBP and its cysteine mutants (25 μM) (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 μl 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 μl 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.

4.8.9. Static and Dynamic Light Scattering

Protein samples (1 μM) were analyzed using static and dynamic light scattering by auto-injection 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 μl) reaches the detectors 5 min after injection by the autosampler (Agilent 1100 series). DM-MBP (60 μM) 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-MBP and Rubisco were mixed at various molar ratios with respect to protomer and adjusted to a final total protein concentration

of 1 μM . 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.

4.9. Single-Molecule Experiments

4.9.1. Fluorescence Correlation and Cross-correlation Spectroscopy

For correlation experiments, a DM-MBP double cysteine mutant (DM-MBP(52-298)) was fluorescently labeled either with Atto532 maleimide or Atto647N maleimide (ATTO-TECH) as described before.

Fluorescence correlation spectroscopy (FCS) measurements using pulsed interleaved excitation (PIE) (Muller et al., 2005) were performed either on a confocal system based on an inverted microscope (Nikon TE 2000) or on a Microtime 200 instrument (PicoQuant GmbH) to analyze the oligomeric state of DM-MBP during spontaneous refolding. Diffusion times of 10 nM DM-MBP(52-298), labeled with Atto532 at both positions, through the confocal volume (1 femtoliter) with or without 1 or 2 μM unlabeled DM-MBP(52-298) were measured by FCS. Atto532-labelled DM-MBP(52-298) bound to GroEL was used as a control. Measurements were taken at 20 $^{\circ}\text{C}$ 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:

$$y = \frac{\tau}{N} \left(1 + \frac{\tau}{\tau_D}\right)^{-1} \left(1 + \frac{1}{\rho^2} \frac{\tau}{\tau_D}\right)^{-1/2} \times \left(1 + A e^{-\left(\frac{\tau}{\tau_A}\right)^\beta}\right) + y_0 \quad (7)$$

for freely diffusing particles with a characteristic diffusion time τ_D through the confocal volume characterized by the structure parameter X 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 ($V_{\text{eff}} = \left(\frac{\pi}{2}\right)^{3/2} w_0^2 z_0$) 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 after pulsing of the detectors. Time scales of afterpulsing τ_A were typically in 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) is given by the equation,

$$D = \frac{w_0^2}{4\tau_D} \quad (8)$$

Fluorescence cross-correlation spectroscopy (FCCS) measurements were performed under refolding conditions as above with a 1:1 mixture of Atto532-labelled DM-MBP (52-298) and Atto647N-labelled DM-MBP(52-298) at final concentrations 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 μW for the FCS and FCCS experiments. The fluorescence emission was focused on the confocal pinhole (diameter 50 μm) and subsequently split by dichroic mirror (600 dcmr, 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 red excitation laser and detected in the red channel. In this way, spectral crosstalk can be removed, making FCCS much more sensitive to the weak interactions. A double-labelled DNA oligomer (40 base-pairs), labeled with Atto532 in one strand and Atto647N 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.

4.9.2. Single-Pair FRET

Single-pair FRET (SpFRET) measurements were performed on either the home-built system or the Microtime 200 using pulsed interleaved excitation (PIE) (Muller et al., 2005). DM-MBP double cysteine mutant (DM-MBP(52-298)) was labelled with Atto532 maleimide (position 52) and Atto647N maleimide (position 298) as described before (Sharma et al., 2008). The concentration of double-labeled protein 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 having more than one particle in the probe 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.

4.10. Mass Spectrometric Analysis

Frozen samples (75 μ l) from H/D exchange pulse labeling were thawed rapidly within 1 min and immediately injected into a custom built HPLC system (designed by John. R. Engen, Bernett Institute, Boston) at a flow rate of 40 μ l/min. Proteins were desalted for 2 min on a 1 mm \times 50 mm Vydac C-4 column and directly eluted into the mass spectrometer with a 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 ~0°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/h and a cone gas flow of 50 l/h. 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/ μ l) as LockMass. Masses were calculated by deconvoluting multiple charge states of combined protein spectra using MassLynx software and MaxEnt1 algorithm (Wales and Engen, 2006).

5. Results

5.1. Model Substrate – Maltose Binding Protein

It is difficult to compare spontaneous and chaperonin-assisted refolding rates of obligate GroEL substrates owing to their high propensity to aggregate upon *in vitro* refolding (Kerner et al., 2005). Maltose binding protein (MBP) fits the criteria of being a model substrate protein as it has been reported that GroEL/ES can increase the folding speed of a mutant form of MBP (Sparrer et al., 1997; Tang et al., 2006). MBP is a monomeric 41 kDa periplasmic protein involved in maltose transport in *E coli*, folds robustly in the cytosol when expressed without its N-terminal export sequence.

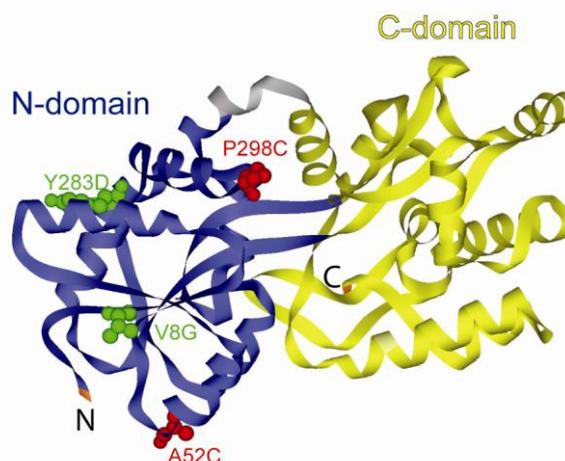


Figure 5.1: Ribbon Diagram of the Structure of MBP (Spurlino et al., 1991); pdb 1OMP; DS viewer-Pro), indicating the positions of DM-MBP mutations V8G and Y283D (green), and the cysteine mutations A52C and P298C (red) used for fluorescent labeling. The two discontinuous domains are shown in blue (N-domain) and yellow (C-domain).

MBP consists of two globular domains, discontinuous in sequence, containing secondary structural $\beta\alpha\beta$ units with the binding site for maltose located in a cleft between the domains (Figure 5.1) (Spurlino et al., 1991). Formation of native contacts within the N domain is thought to be the rate limiting step in folding (Chun et al., 1993). Several slow folding

mutants of MBP, with the mutations V8G and Y283D in a strand and loop segment of the N-domain have been reported like the single mutant Y283D (SM-MBP) and the double mutant V8G/Y283D (DM-MBP) (Wang et al., 1998). MBP possesses eight tryptophans distributed over both domains. Their fluorescence is reduced upon unfolding, and the recovery of fluorescence can be used as a measure of folding (Chun et al., 1993) both in the absence and presence of GroEL/ES (GroEL and GroES lack tryptophan residues). Importantly, the spontaneous refolding of DM-MBP is slow ($t_{1/2}$ ~30 min at 25°C), occurring with ~90% yield, but is accelerated ~10 fold in the presence of GroEL/ES system. In contrast to GroEL/ES, the bacterial Hsp70 chaperone system (DnaK, DnaJ, GrpE) strongly retards the folding of DM-MBP (Tang et al., 2006). Very similar properties are described for obligate GroEL substrates which are highly aggregation prone (Kerner et al., 2005). Thus, DM-MBP can be considered a model substrate to study GroEL mediated refolding.

5.2. Chaperonin-Assisted Refolding is Independent of Aggregation Prevention

The spontaneous refolding of DM-MBP is slow ($t_{1/2}$ ~30 min at 25°C), occurring with ~90% yield, but is accelerated up to 10-fold in the presence of GroEL/ES system (under standard refolding conditions of 200 mM KCl and 60 mM GuHCl) (Figure 5.2) (Tang et al., 2006). Spontaneous refolding is chloride sensitive (Apetri and Horwich, 2008) and is ~2-fold faster at 20 mM KCl or in the absence GuHCl. In contrast, the chaperonin-assisted folding is salt insensitive (Figure 5.2).

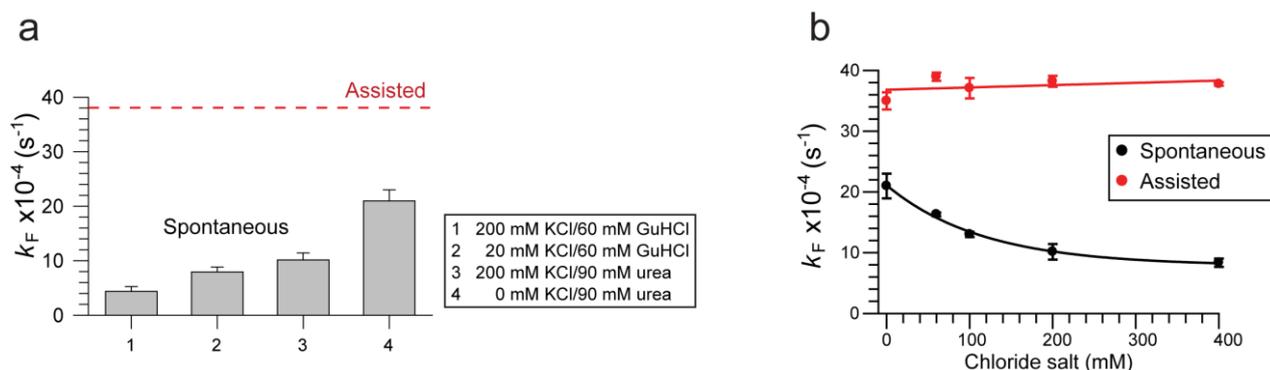


Figure 5.2: Effect of Chloride Salt on Spontaneous and Assisted Refolding of DM-MBP.

(a) Rates of spontaneous refolding of DM-MBP at 25°C under varying conditions of chloride salt in presence of residual concentrations of GuHCl and urea, as indicated. The rates of chaperonin-assisted refolding, which is independent of chloride salt or the specific denaturant, are shown as a reference. Standard deviation from at least three independent experiments are shown. **(b)** Rates of spontaneous and assisted refolding of DM-MBP at varying concentrations of chloride salt in the presence of urea is shown. DM-MBP (25 μM) was denatured in 9 M urea and diluted 100-fold into buffer (100 mM HEPES pH 7.2, 20 mM KOAc, 5 mM $\text{Mg}(\text{OAc})_2$) containing chloride salt (KCl) concentrations indicated either alone (spontaneous) or with 1 μM GroEL/2 μM GroES (assisted) at 25°C. Assisted refolding was initiated by the addition of 2 mM ATP. Refolding was monitored by Trp fluorescence. Standard deviation is shown for three independent experiments (DM-MBP refolding experiments were done together with Manal Chatila)

The rate enhancement of folding by GroEL/ES could reflect an active role of the chaperonin in catalyzing DM-MBP refolding. Alternatively, the chaperonin may accelerate folding in a passive manner by preventing the formation of reversible aggregates that would reduce the folding rate but not the yield (Apetri and Horwich, 2008). To have a closer look, we first simulated the effect of reversible aggregation on DM-MBP refolding rate and then experimentally measured DM-MBP refolding rate at different concentration.

5.2.1. Simulating the Effect of Aggregation on DM-MBP Refolding

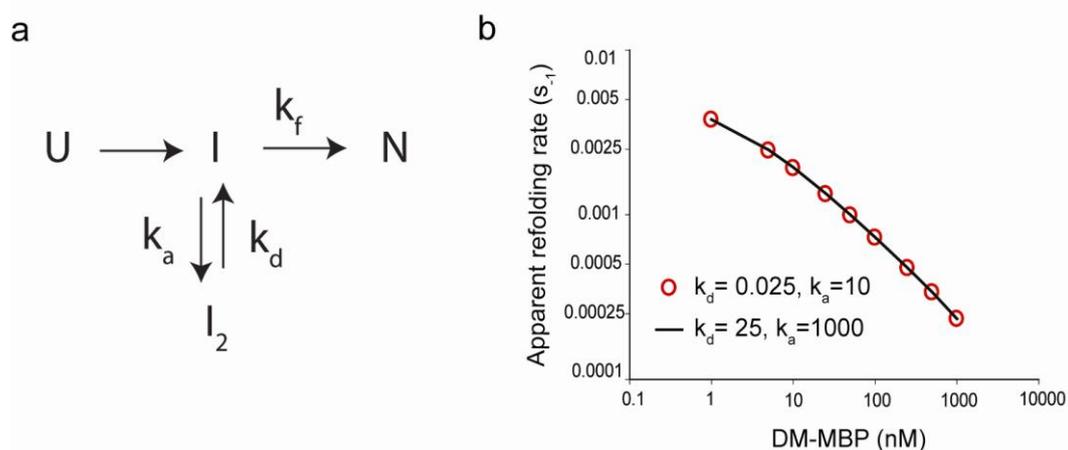


Figure 5.3: Simulation of Spontaneous Refolding Limited by Reversible Aggregation. (a)

A simple model of reversible aggregation involving the formation of a dimeric species (I_2) was used to simulate folding kinetics limited by the reversible formation of non-native oligomers. The rates of formation and dissociation of the dimeric species is given by k_a and k_d , respectively. The folding rate k_f , unimpaired by aggregation, is given by the refolding rate of protein inside the GroEL/ES cavity, $\sim 0.007 \text{ s}^{-1}$ ($t_{1/2} \sim 1000 \text{ s}$) as measured for DM-MBP (52-298) labeled with Atto532 at position 52 (Sharma et al., 2008) (b) Simulated refolding rates at different concentrations of DM-MBP based on model described in (a). Varying k_a and k_d showed that apparent refolding rate is dependent only on the equilibrium dissociation constant (K_D) of the dimer and not on the individual values of k_a and k_d .

Aggregation may affect the rate and/or the yield of folding, depending on whether the reaction is limited by the formation of reversible or irreversible aggregates. Formation of reversible aggregates would reduce the apparent folding rate but not the yield of folding. On the other hand, formation of irreversible aggregates would only reduce the folding yield without compromising the folding rate. These two rate limiting step may also occur in combination. To simulate the possibility that DM-MBP refolding is limited by aggregation, we applied a simple scheme for the formation of reversible off-pathway aggregate (minimally a non-native dimer I_2) as the rate limiting step of DM-MBP refolding (Figure 5.3a). The spontaneous refolding rate of the protein, K_f , unaffected by aggregation, is given by the

refolding rate of the protein inside the GroEL/ES cavity ($\sim 0.007 \text{ s}^{-1}$, $t_{1/2} \sim 1000 \text{ s}$), as measured for DM-MBP(52-298) labeled with Atto-532 at position 52 (Sharma et al., 2008) (Figure 5.1). Varying the rate of formation (k_a) and dissociation (k_d) of the dimeric aggregate species (I_2) showed that the apparent refolding rate was dependent only on the equilibrium dissociation constant (K_D) of the dimer rather than on the individual values of k_a and k_d (Figure 5.3b). At a K_D of 0.4 nM, the rate of refolding matched the experimentally observed rate of spontaneous refolding of DM-MBP(52-298) labeled with Atto532 at position 52 (0.0007 s^{-1} , $t_{1/2} \sim 1000 \text{ s}$) (Sharma et al., 2008). The apparent rate of refolding was simulated at different concentrations of DM-MBP and was found to increase with decreasing concentrations of DM-MBP (Figure 5.3b). Thus, if the folding rate of DM-MBP is actually limited by formation of reversible dimers (or higher order multimers), it should be sensitive to the concentration of DM-MBP.

5.2.2 Formation of Reversible Aggregates is not the Rate Limiting Step in Chaperonin-Assisted DM-MBP Folding

To have a closer look into the chaperonin assisted folding mechanism and verify experimentally if DM-MBP refolding is limited by formation of reversible aggregates, we measured DM-MBP refolding rate over a wide range of concentrations (10 nM to 1.5 μM) by monitoring intrinsic tryptophan (Trp) fluorescence. These experiments were performed under our standard refolding conditions (200 mM KCl and 60 mM GuHCl).

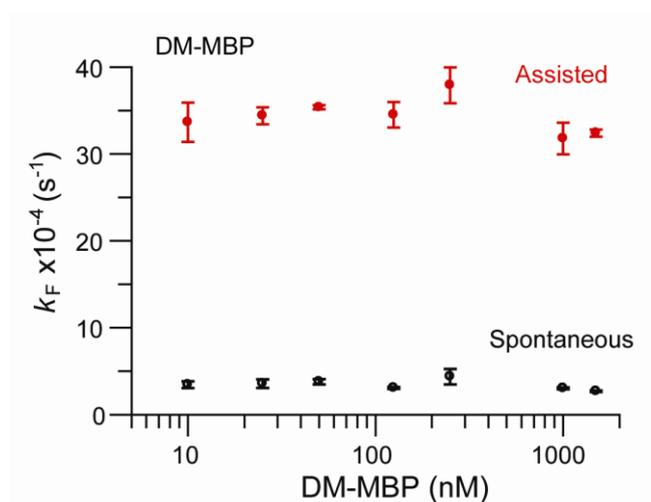


Figure 5.4: Rates of Spontaneous and GroEL/ES-Assisted Refolding of DM-MBP at Different DM-MBP Concentrations (10 nM-1.5 μ M). DM-MBP was denatured in 6 M GuHCl and diluted 100-fold to 60 mM GuHCl into buffer A (spontaneous) or buffer containing 0.5-4 μ M GroEL and 1-8 μ M GroES (assisted) at 25°C. Assisted refolding was initiated by the addition of 5 mM ATP. Refolding was monitored by Trp fluorescence. Standard deviations of 3 independent measurements are shown.

We found that DM-MBP refolding rate is independent of concentration over the whole concentration range tested (Figure 5.4). Note that the experiments had to be performed with a very small excitation slit-width and a measurement interval of 120 s for spontaneous refolding, with excitation only during measurement time, to avoid photobleaching during measurements. At higher excitation slit widths, at low DM-MBP concentrations, photobleaching caused an apparent increase in the refolding rate. The chaperonin-assisted refolding rate, as expected, was independent of DM-MBP concentration and was accelerated \sim 10-fold as compared to spontaneous folding of DM-MBP. These data argue against transient aggregation as a rate limiting step during spontaneous refolding of DM-MBP.

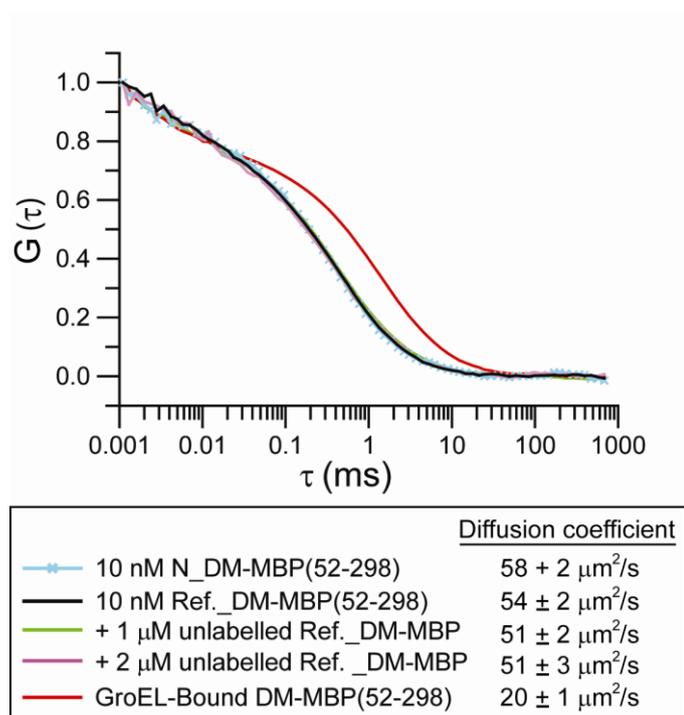


Figure 5.5: Fluorescence Correlation Spectroscopy (FCS). FCS was used to measure the diffusion rates of spontaneously refolding DM-MBP. Normalized fluorescence autocorrelation amplitudes $G(\tau)$ are shown. Diffusion times were measured at 20°C during the first 800 s of refolding with 10 nM Atto532 labeled DM-MBP(52-298) in the absence or presence of 1 or 2 μM unlabeled, refolding DM-MBP(52-298) (final GuHCl 30 mM in buffer A). Native and GroEL-bound labeled DM-MBP(52-298) were used as controls. Diffusion coefficients are given as averages \pm SD from 3 independent experiments (FCS and FCCS experiments were done together with Guoxin Jiang)

To be certain that DM-MBP refolding is not limited by aggregation, we next performed fluorescence correlation (FCS) and cross-correlation (FCCS) spectroscopy with fluorescently-labeled DM-MBP to look for formation of aggregates in a more direct way. For this purpose, we have used a double cysteine mutant, DM-MBP(52-298), labeled with Atto532 (Figure 5.1). The rate of spontaneous folding of the labeled protein was similar to the unlabeled DM-MBP (0.0007 s^{-1} , $t_{1/2} \sim 1000 \text{ s}$), and was accelerated ~ 6 -fold by GroEL/ES (Sharma et al., 2008). FCS measurements were performed by monitoring the decay of the autocorrelation function as a measure of the average diffusion time of particles through the probe volume. The diffusion coefficient of refolding DM-MBP (10 nM) ($54 \mu\text{m}^2$), measured

during the first 800 s of refolding upon dilution from the denaturant, was similar to that of native DM-MBP ($\sim 58 \mu\text{m}^2$). Importantly, the diffusion rate of refolding, DM-MBP remained unchanged in the presence of excess (1 or 2 μM) unlabeled, refolding DM-MBP (Figure 5.5). This excludes the formation of large aggregates during spontaneous refolding of DM-MBP. GroEL bound DM-MBP ($\sim 800 \text{ kDa}$) was used as a control.

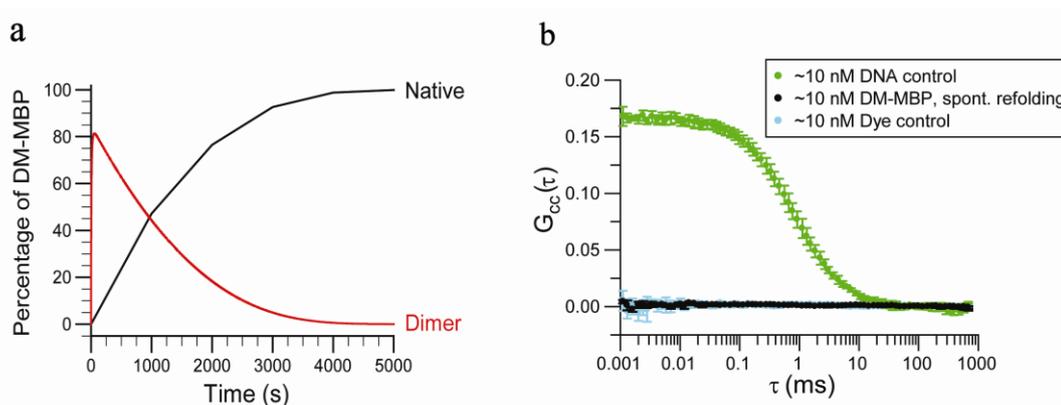


Figure 5.6: Fluorescence Cross-Correlation Spectroscopy (FCCS). (a) Numerical kinetic simulations of refolding using the model in Figure 5.3 were performed while keeping the concentration of DM-MBP fixed at 4 nM. (b) Fluorescence cross-correlation spectroscopy (FCCS) of a 1:1 mixture of DM-MBP(52- 298) labeled at position 52 with Atto532 or Atto647N. Labeled DM-MBP molecules were denatured and diluted as in Figure 5.5 to a final concentration of $\sim 5 \text{ nM}$ each. FCCS was recorded with pulsed interleaved excitation (PIE) (Muller, 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 negative control.

Numerical simulations also suggested that reversible aggregation would result in formation of stable dimeric or multimeric aggregate species that would decay as the same rate as the formation of the native state. Given the very slow rate of formation of the native state for DM-MBP, this species should represent a large fraction of the folding molecule population. To monitor the formation of the smallest aggregate (dimer in this case), we performed FCCS experiments. A cysteine mutant DM-MBP(52-298) labeled at both positions 52 and 298 with either Atto532 or Atto647N were mixed 1:1 and unfolded. Refolding was initiated at a final concentration of $\sim 5 \text{ nM}$. Assuming aggregation limited folding kinetics,

aggregates are expected to be populated to $\sim 75\%$ during the first 250 s of refolding, corresponding to $\sim 35\%$ of molecules containing both labels in case of exclusive dimer formation (Figure 5.6a). However, no dimeric or multimeric species were detected in the FCCS measurements, indicating the absence of aggregate formation during DM-MBP refolding (Figure 5.6b). As a positive control, double labeled DNA was used, which exhibited nearly quantitative cross-correlation. A mixture of free dyes served as a negative control.

We also analyzed DM-MBP refolding by static light scattering measurements. The light scattering signal measured and the calculated mass of the particles corresponded to that of folded DM-MBP and remained unchanged during the course of folding reaction (Figure 5.7).

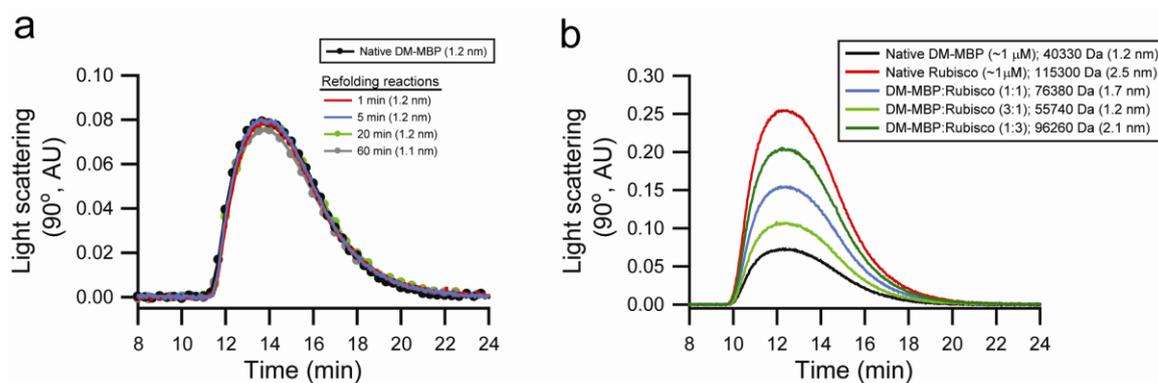


Figure 5.7: Light Scattering Analysis of DM-MBP. (a) Light scattering measurements collected at an angle of 90° for native DM-MBP ($\sim 1 \mu\text{M}$) and of DM-MBP spontaneously refolding at various times after the initiation of the refolding reaction are shown. Denatured DM-MBP in 3 M GuHCl was diluted 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. (b) 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 were adjusted to $\sim 1 \mu\text{M}$ (Light scattering experiments were done by Manajit Hayer-Hartl).

Using dimeric bacterial Rubisco (~50 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 (Figure 5.7b). These measurements failed to reproduce the presence of highly scattering aggregates during DM-MBP refolding as reported by Apetri and Horwich (Apetri and Horwich, 2008). However, the observed scattering signal did not decay within time scale of refolding. This would be expected if dissociation of aggregate limited the rate of DM-MBP refolding, and thus is inconsistent with the proposed presence of reversible DM-MBP aggregates as the cause of slow folding of DM-MBP.

5.2.3 Aggregation Limits Yield but not Rate of Rubisco Refolding

We also explored the relationship between the folding rate and aggregate formation for bacterial Rubisco, a GroEL substrate that is highly aggregation-prone. We measured the refolding rate of Rubisco at 9.5°C at different Rubisco concentration in a range of 10 nM to 100 nM. Spontaneous refolding occurred with ~80% yield and GroEL/ES accelerated the refolding rate by ~3-fold. Interestingly, we found both spontaneous and assisted refolding rate of Rubisco to be concentration independent in the range tested (Figure 5.8a, b).

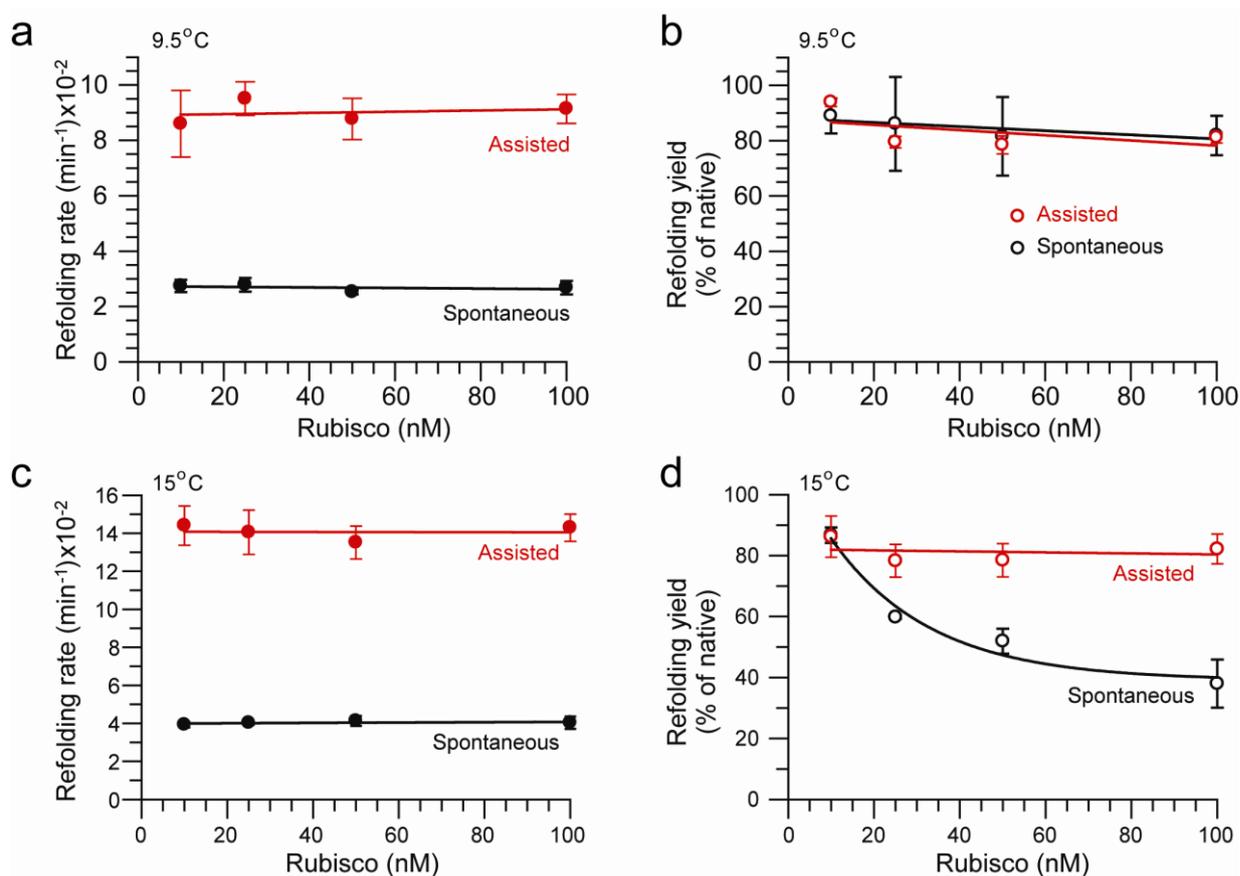


Figure 5.8: Spontaneous and Assisted Refolding of Rubisco (a) Rates and (b) Yields of spontaneous and GroEL/GroES-mediated refolding of bacterial Rubisco at final concentrations of 10 nM to 100 nM were measured. Rubisco was denatured in 6 M GuHCl and diluted 100-fold into buffer alone (spontaneous) or buffer containing 2 to 10-fold molar excess of GroEL and 2-fold molar excess of GroES over GroEL (assisted) at $\sim 9.5^\circ\text{C}$ (a, b) or $\sim 15^\circ\text{C}$ (c, d). Assisted refolding was initiated by the addition of 5 mM ATP. Refolding was assayed by recording enzymatic activity. Yields are expressed as percent of native control. Standard deviations from 3 independent measurements are shown (RuBisco refolding was done by Qiaoyun Shi).

However, at $\sim 15^\circ\text{C}$, both spontaneous and assisted refolding rate was found to be independent of concentration with rate acceleration in presence of GroEL/ES (Figure 5.8c), but the yield of spontaneous refolding decreased substantially with increasing concentration (Figure 5.8d). This is consistent with the pronounced temperature-dependence of aggregation of Rubisco (Brinker et al., 2001).

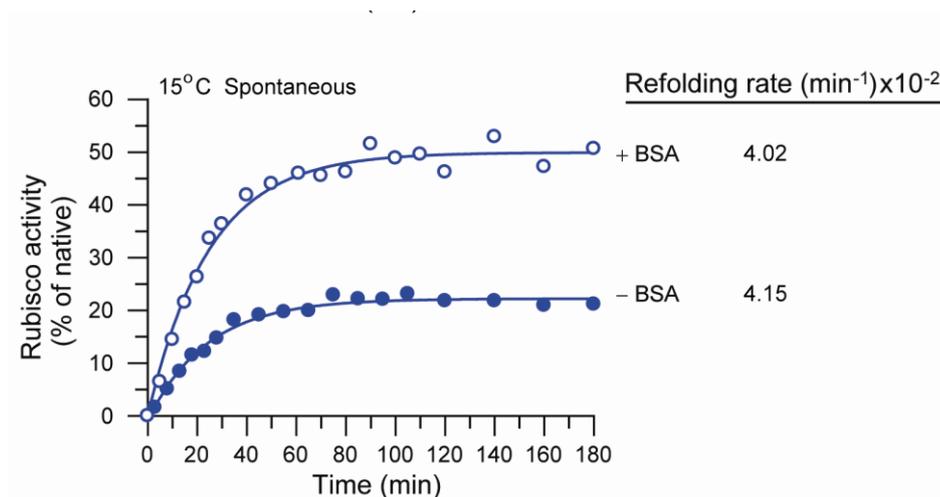


Figure 5.9: Spontaneous Refolding of RuBisco in Presence of BSA. Spontaneous refolding of 100 nM Rubisco was performed in the absence or presence of BSA (~15 μ M) at 15°C. Activities are expressed as a percentage of the activity of 100 nM native Rubisco. Presence of BSA in the refolding buffer, to prevent Rubisco adsorption to the tube walls, did not affect the refolding rate but the folding yield is reduced significantly, suggesting the formation of irreversible aggregates.

Furthermore, it was recently suggested that BSA which is routinely used in refolding reactions to limit non-specific protein loss, modulates Rubisco aggregation, perhaps by favoring the formation of reversible aggregates during spontaneous refolding (Apetri and Horwich, 2008). Upon repeating the refolding reactions with and without BSA, we observed that the apparent rate of spontaneous refolding remained unchanged, but the folding yield was significantly reduced in the absence of BSA (Figure 5.9), supporting our conclusion that aggregate formation and/or adsorption to tube walls is irreversible at the time-scale of folding. We conclude that under the assay conditions, Rubisco aggregation is irreversible and hence does not limit the rate of refolding but the yield. Hence, aggregation prevention by chaperonin explains the increase in the yield but not in the rate of Rubisco refolding.

In summary, for both DM-MBP and Rubisco, the observed rate acceleration for folding in presence of GroEL/ES cannot be attributed to prevention of aggregation by a solely passive cage mechanism. It must involve an active role of the chaperonin in modulating the intrinsic properties of these proteins.

5.3. Slow Conversion of Trapped Intermediate Limits DM-MBP Refolding

5.3.1. Folding of DM-MBP Starts from a Monomeric Intermediate State

Having excluded the formation of reversible aggregates as the cause of slow spontaneous folding of DM-MBP, it seemed plausible that DM-MBP refolding is limited by formation of the kinetically trapped intermediate. To have a better understanding of the folding pathway of DM-MBP, we performed isothermal equilibrium GuHCl denaturation experiments for DM-MBP (final concentration 1 μ M) after 12 h incubation in varying concentrations of the denaturant, using intrinsic tryptophan (Trp) and circular dichroism (CD) as a probe of structure formation. There was a prominent hysteresis in the denaturation-renaturation profile indicating the presence of a kinetic intermediate, stably populated at 0.5-0.8 M GuHCl. This indicated that even though the native structure was thermodynamically stable till \sim 1 M GuHCl, it was kinetically disfavored during the refolding phase, indicating towards the presence of a kinetically trapped intermediate during the refolding phase (Figure 5.10).

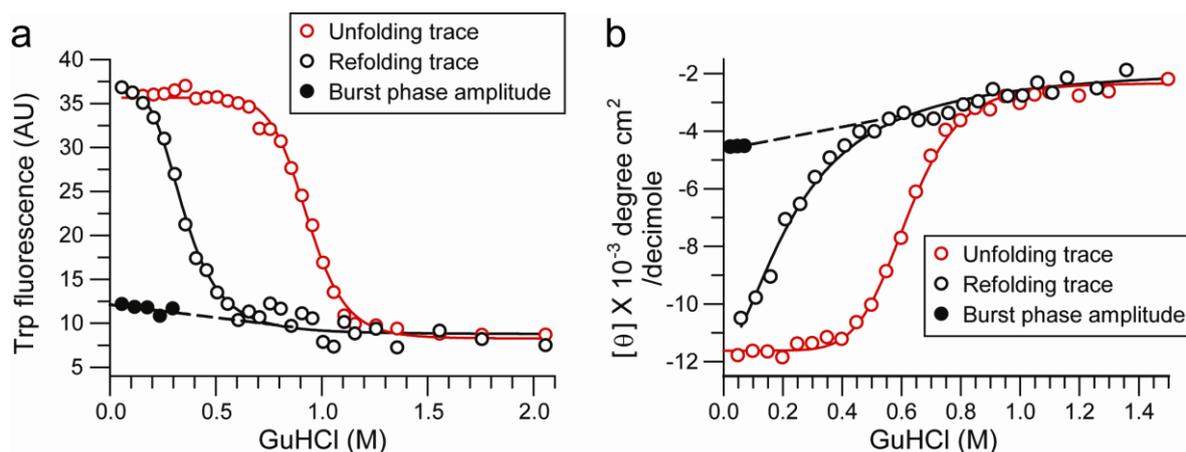


Figure 5.10: GuHCl-Dependent Unfolding and Refolding of DM-MBP. Unfolding and refolding of DM-MBP at a final concentration of 1 μ M in different GuHCl concentration was monitored by (a) Trp fluorescence at 345 nm and (b) circular dichroism at 220 nm. *Unfolding trace:* Native DM-MBP was incubated for 12 h in buffer A containing ~60 mM to ~2 M GuHCl at 25°C. *Refolding trace:* DM-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 h at 25°C. Burst phase amplitudes were determined immediately on dilution from 6 M GuHCl. Representative measurements from at least 2 independent experiments are shown.

Next, we sought to investigate whether the intermediate that is populated at 0.5-0.8 M GuHCl is similar to the kinetic intermediate that is formed during the standard refolding conditions (~60 mM GuHCl). For this, we started the refolding reaction after dilution from denaturant (final GuHCl concentration 60 mM) and analyzed the burst phase Trp fluorescence and CD signals (Figure 5.10a, b). The signals were extrapolated to zero point. Spectroscopic showed that the populated intermediate (at 0.5-0.8 M GuHCl) does represent the starting point of the folding reaction.

WT-MBP and a single mutant of MBP containing only the Y283D substitution (SM-MBP) were also investigated by the isothermal equilibrium GuHCl experiments to verify the presence of kinetically trapped intermediate as the reason for the slow folding of DM-MBP. Refolding rates of the MBP variants follow the following order DM-MBP < SM-MBP < WT-

MBP (Tang et al., 2006). Interestingly, hysteresis was less pronounced with SM-MBP as compared to DM-MBP, while no hysteresis was observed with WT-MBP (Figure 5.11 a, b, c), reflecting the relative order of folding rates of these proteins.

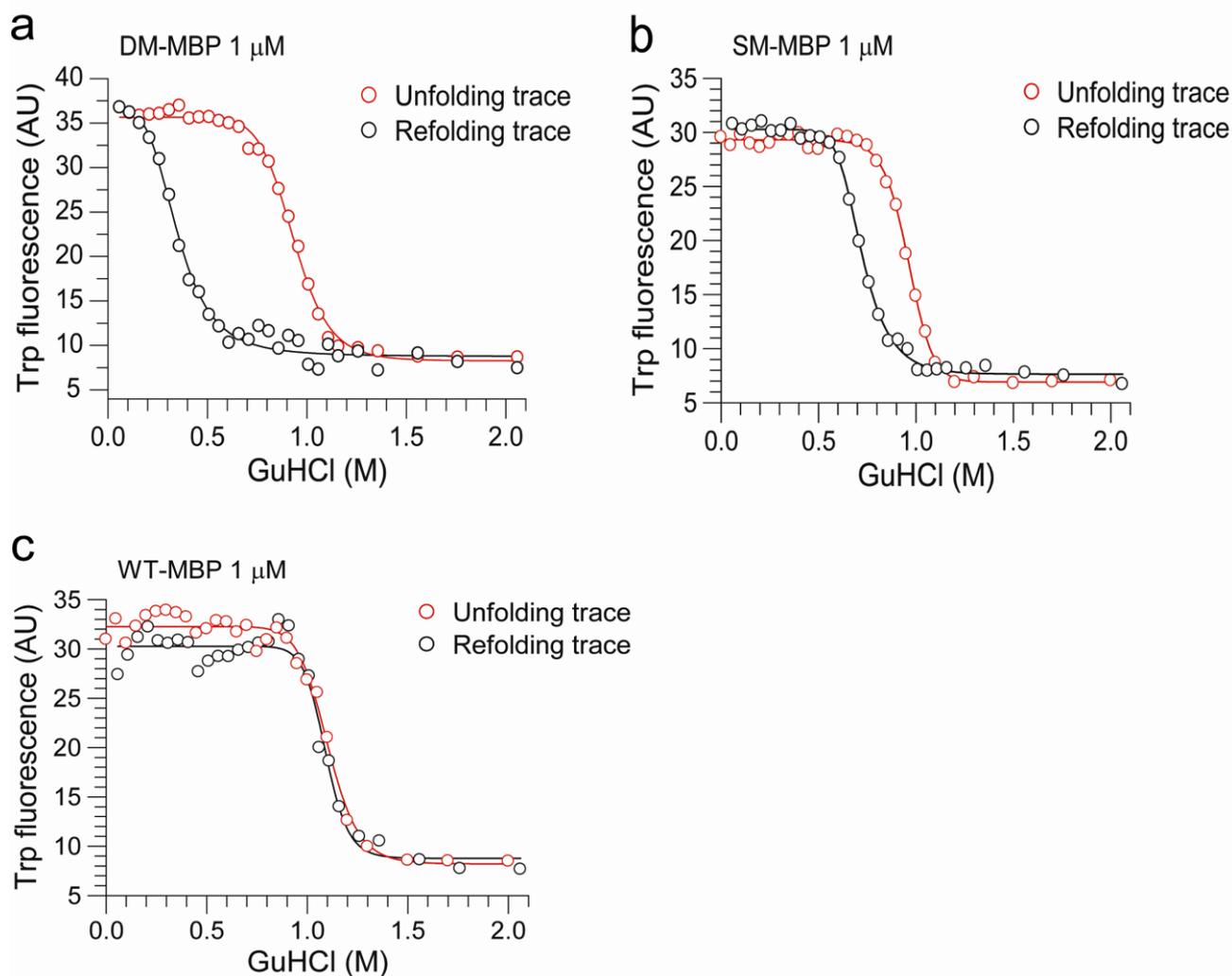


Figure 5.11: GuHCl Dependent Unfolding and Refolding of MBP mutants. Unfolding and refolding of (a) DM-MBP, (b) SM-MBP and (c) WT-MBP (1 μ M each) was monitored by Trp fluorescence with excitation at 295 nm and emission at 345 nm. *Unfolding trace:* Native MBP and MBP mutants were incubated for 12 h with increasing concentrations of GuHCl in buffer A at 25°C. *Refolding trace:* MBP and MBP mutants (100 μ M) were unfolded in 6 M GuHCl and then diluted 100-fold into buffer A containing increasing concentrations of GuHCl, followed by incubation for 12 h at 25°C. Representative measurements from at least 2 independent experiments are shown.

We also performed similar experiments with DM-MBP, SM-MBP and WT-MBP using urea as the denaturant in the absence of chloride salt and observed similar trend in the hysteresis profile (Fig 5.12a, b, c).

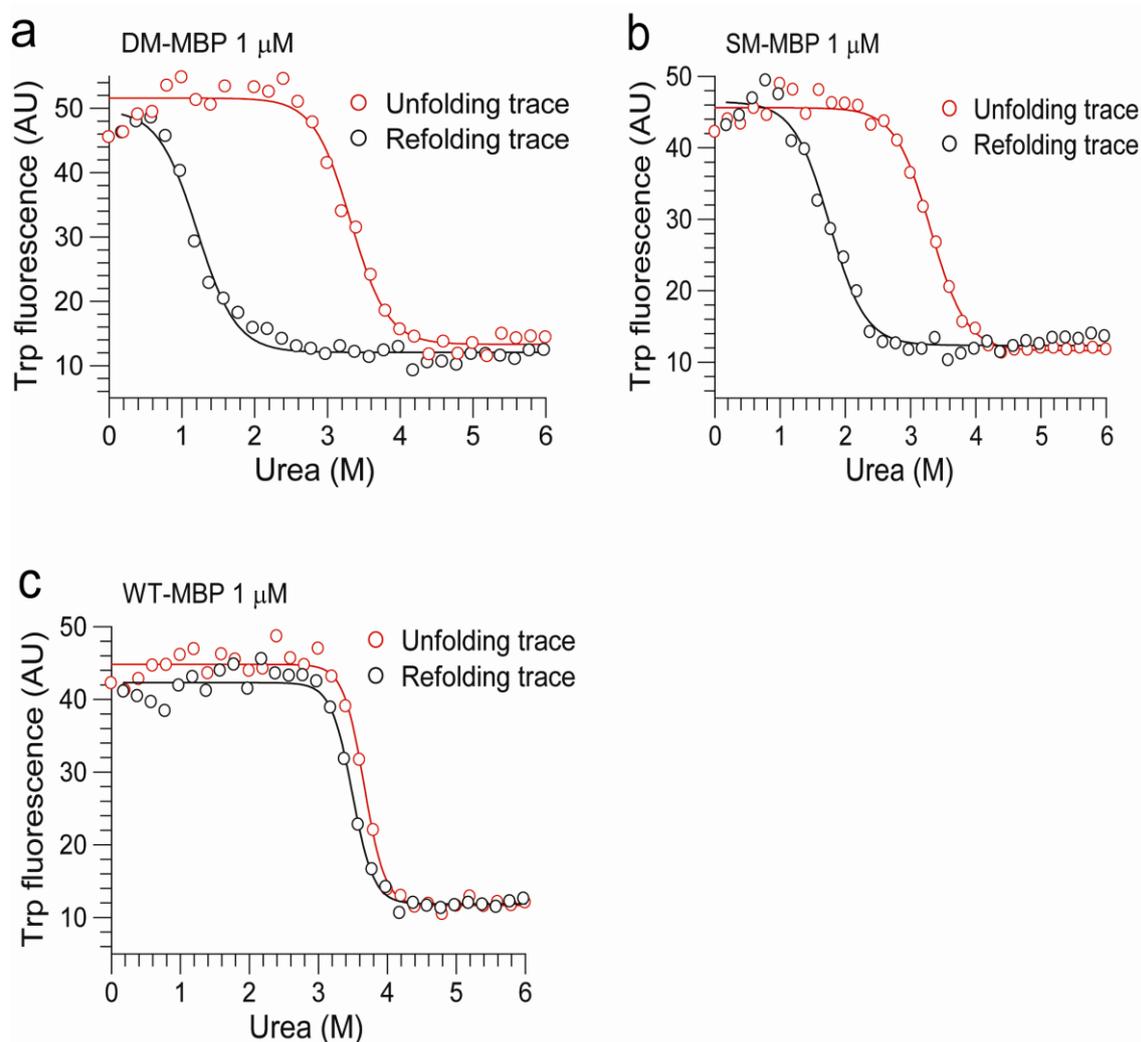


Figure 5.12: Urea Dependent Unfolding and Refolding of MBP Mutants. Unfolding and refolding of (a) DM-MBP, (b) SM-MBP and (c) WT-MBP (1 μ M each) was monitored as described before. *Unfolding trace*: Native MBP and MBP mutants were incubated for 12 h with increasing concentrations of Urea in buffer D (100 mM HEPES, pH 7.2; 200 mM KOAc, 5 mM Mg(OAc)₂) at 25°C. *Refolding trace*: MBP and MBP mutants (100 μ M) were unfolded in 9 M Urea and then diluted 100-fold into buffer D (100 mM HEPES, pH 7.2; 200 mM KOAc, 5 mM Mg(OAc)₂) containing increasing concentrations of Urea, followed by incubation for 12 h at 25°C. Representative measurements from at least 2 independent experiments are shown.

On the other hand, hysteresis effect was also found to be independent of protein concentration, as tested for 0.25 and 2 μM DM-MBP (Figure 5.13a, b). This suggests that the observed hysteresis effect is independent of DM-MBP concentration and denaturant used.

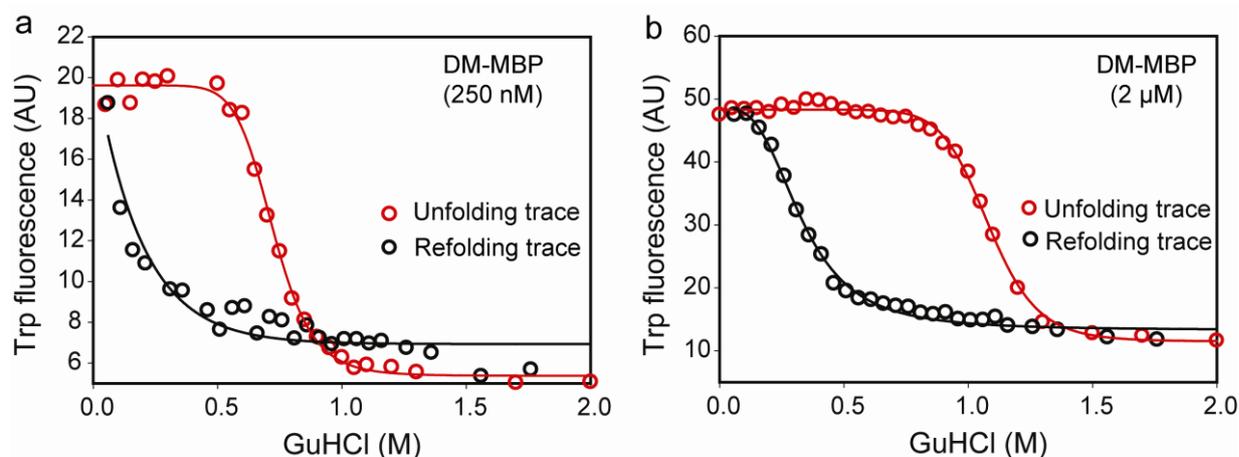


Figure 5.13: GuHCl Dependent Unfolding and Refolding at Different DM-MBP Concentration. Unfolding and refolding of (a) 0.25 μM , and (b) 2 μM , DM-MBP was performed at different concentrations of GuHCl as described in Figure 5.11 a.

5.3.2. Kinetically Trapped Intermediate Populated during DM-MBP

Folding is Collapsed and Lacks Ordered Structure

To investigate the structure of this intermediate populated at $\sim 0.5\text{M}$ GuHCl, we performed CD spectroscopy and found that it has only $\sim 22\%$ α -helical structure (Figure 5.14a). Also Trp fluorescence intensity was as low as the unfolded protein, suggesting the absence of ordered tertiary structure in the intermediate state populated at $\sim 0.5\text{-}0.8\text{ M}$ GuHCl (Figure 5.14b).

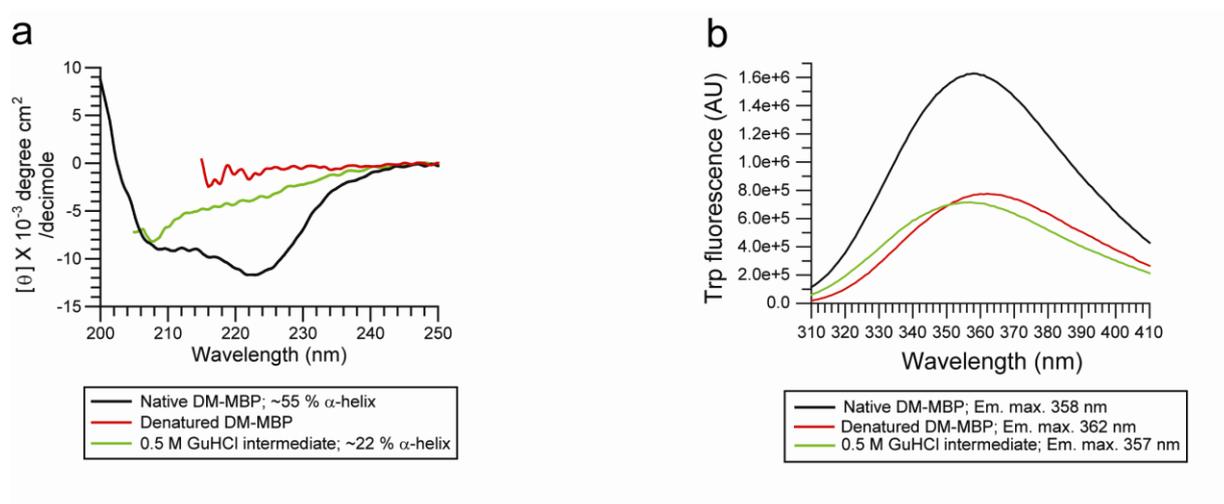


Figure 5.14: Circular Dichroism and Tryptophan Wavelength Scan. (a) CD wavelength scans of native, denatured and 0.5 M GuHCl intermediate of DM-MBP (2 μM) were measured at 25°C with 0.1 cm cuvettes. (b) Tryptophan fluorescence scan of native, denatured and 0.5 M GuHCl intermediate of DM-MBP (250 nM each) was performed with an excitation at 295 at 25°C. Emission maxima are indicated.

Single-pair FRET (spFRET) measurements in solution were next performed to analyze the compactness of the intermediate at 0.5 M GuHCl. DM-MBP(52-298) was used, and labeled with Atto532 (Donor) preferentially at position 52 in presence of maltose and Atto647N (Acceptor) at position 298 after maltose was removed from the labeling reaction. Site specific labeling was performed as described before (Sharma et al., 2008). These two positions are $\sim 33 \text{ \AA}$ apart in the native structure (Figure 5.1). The fully unfolded protein in 3 M GuHCl exhibits low FRET efficiency values with distributions of FRET efficiencies (f_E) at peak value of ~ 0.08 . Similarly, native protein showed distributions of FRET efficiencies at peak value of ~ 0.84 indicating the compact nature of the native state (Sharma et al., 2008). Interestingly, the kinetic trapped intermediate state ($\sim 0.5 \text{ M GuHCl}$), even though is not folded exhibited the FRET efficiency distribution centering at f_E of ~ 0.8 , similar to what was observed for native state. Also, this intermediate displayed a broader distribution of FRET efficiencies as compared to either native or denatured state (Figure 5.15).

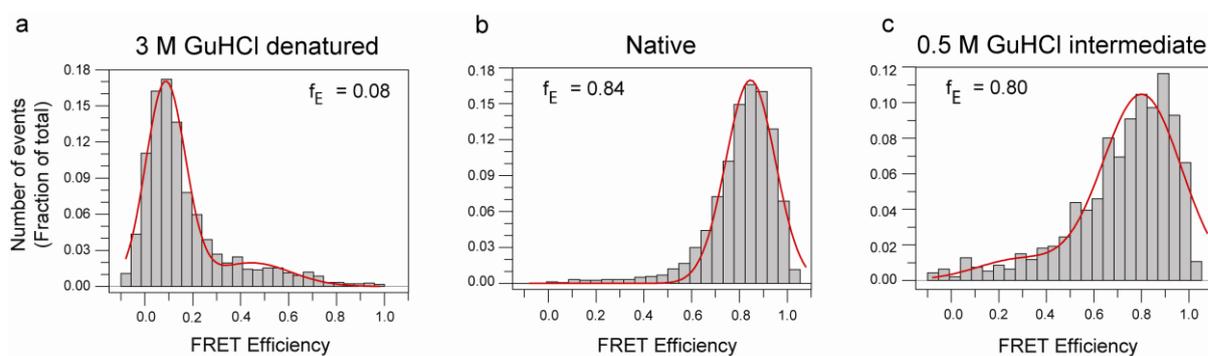


Figure 5.15: SpFRET Analysis in Solution of DM-MBP(52-298). DM-MBP(52-298) double-labeled with Atto532 at position 52 and Atto647N at position 298 was (a) denatured in 3 M GuHCl/Buffer A (Denatured protein), (b) diluted in buffer A (Native protein), and (c) incubated in 0.5 M GuHCl/Buffer A at 25°C for 12 h (Trapped intermediate). Final protein concentration was ~100 pM. Peak values of a Gaussian fit to the FRET efficiency distributions (f_E) are indicated. The shoulder of the f_E peaks in (a) and (c) were fitted with a second Gaussian to allow the correct determination of the peak f_E . Representative histograms of at least 3 independent measurements are shown.

This suggests that the kinetically trapped intermediate has an average compaction similar to that of the native protein but with a greater variability in the structure, as demonstrated by the broader intramolecular distance distribution.

5.3.3 Kinetically Trapped Intermediate during DM-MBP Folding is Dynamic

Pulsed hydrogen-deuterium (H/D) exchange measurements, coupled to mass spectrometry were next performed to further characterize the conformational ensemble representing the kinetically trapped intermediate during DM-MBP refolding (Fig 5.16a). The protonated and fully deuterated proteins were used as reference. DM-MBP fully denatured in 3 M GuHCl exhibited ~310 exchangeable amides which was very similar to the fully deuterated protein. Incorporation of deuterium by the kinetically trapped folding intermediate (at ~0.5 M GuCl), also corresponded to ~310 exchangeable amides which was indistinguishable from the exchange properties of the denatured protein (Figure 5.16b).

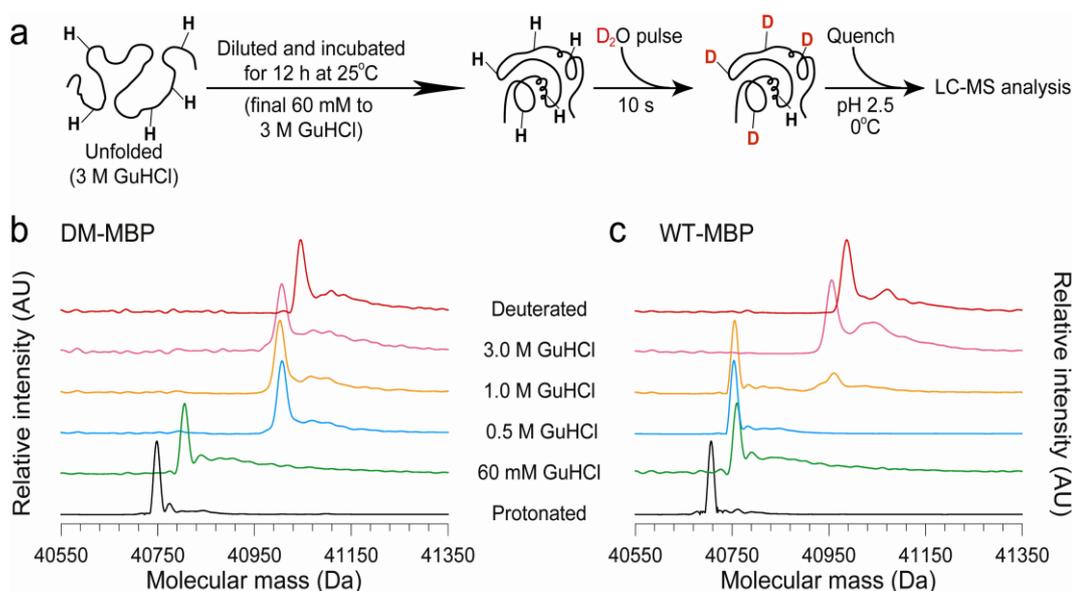


Figure 5.16: Pulsed H/D Exchange at Different GuHCl Concentration.

(a) Schematic representation of the H/D exchange experiment is shown. Unfolded protein at 3 M GuHCl was diluted into buffer B with varying concentrations of GuHCl (60 mM to 3 M) and incubated for 12 h at 25°C. Samples were subjected to D₂O pulse for 10 s and the exchange reaction was quenched by lowering to pH 2.5 at 0°C. Deconvoluted mass spectra of (b) DM-MBP and (c) WT-MBP are shown. Global H/D exchange patterns as a function of denaturant was monitored by ESI-QToF MS. The native protonated and deuterated proteins are shown as reference.

This demonstrates the absence of stable secondary structure in the kinetically trapped intermediate. However, WT-MBP (with a faster refolding rate as compared to DM-MBP) under the same conditions displayed a structural stability similar to that of native state at 60 mM GuHCl (~75 exchangeable amides) (Figure 5.16c) consistent with the absence of hysteresis between the unfolding and refolding curves (Figure 5.11c). Next, pulsed H/D exchange measurements of DM-MBP under refolding conditions (60 mM GuHCl) were performed at different time points during the course of refolding (Figure 5.17a).

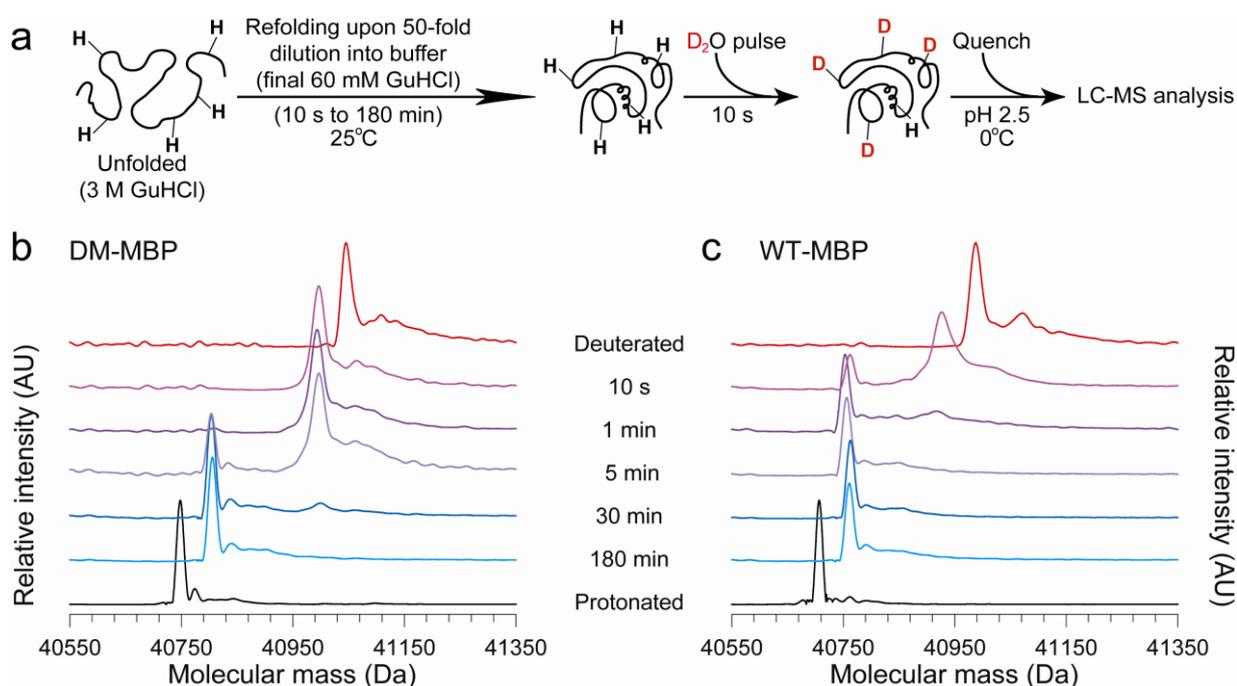


Figure 5.17: Pulsed H/D Exchange Under Refolding Conditions.

(a) Schematic representation of the experiment is shown. Proteins were diluted from 3 M GuHCl into buffer B to a final concentration of 60 mM to initiate refolding. Samples were subjected to D₂O pulse for 10 s. Reaction was quenched at different time points (10 s to 180 min) by lowering to pH 2.5 at 0°C. Deconvoluted mass spectra of (b) DM-MBP and (c) WT-MBP are shown. Global H/D exchange pattern as a function of refolding time, is indicated.

We found that the kinetically trapped species at 0.5 M GuHCl was populated for more than 5 min during refolding, converting slowly to the native state (Figures 5.17b). A similar intermediate was only transiently populated by WT-MBP and converted to the native state in less than 1 min (Figure 5.17c). This is consistent with the fast folding rate of WT-MBP as compared to that of DM-MBP. Thus H/D exchange measurements together with the spectroscopic analysis demonstrate that DM-MBP populates a kinetically trapped intermediate under refolding conditions 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.

5.4. Disulfide-Mediated Constraints Accelerate Spontaneous Refolding

If there is an entropic barrier for DM-MBP folding, then long-range disulfide bonds may serve to configurationally constrain flexible regions in folding intermediates. This would eventually lead to entropic destabilization of the intermediate state relative to the transition state resulting in accelerated folding (Figure 5.18) (Camacho and Thirumalai, 1993; Mamathambika and Bardwell, 2008).

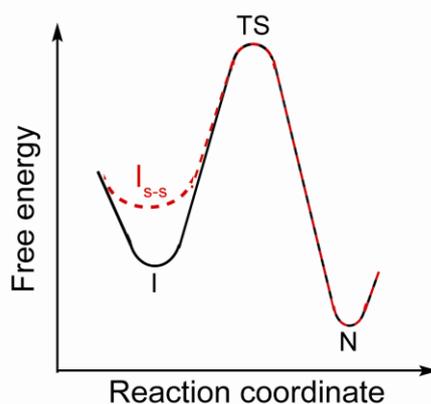


Figure 5.18: Free Energy Diagram Showing the Reaction Coordinate of DM-MBP Folding. Free energy diagram illustrating the entropic destabilization of the kinetically trapped refolding intermediate (I) by disulfide-mediated structural constraints (I_{s-s}) relative to the folding transition state (TS). This explains how disulfide-mediated constraints could result in rate acceleration of DM-MBP folding by lowering the activation energy barrier of folding by destabilizing the intermediate state.

To probe the entropic confinement of the folding barrier of DM-MBP, we engineered a series of 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. Structure of DM-MBP showing cysteine pair mutations, N18C- D296C and D184C- K362C in red is shown in Figure 5.19.

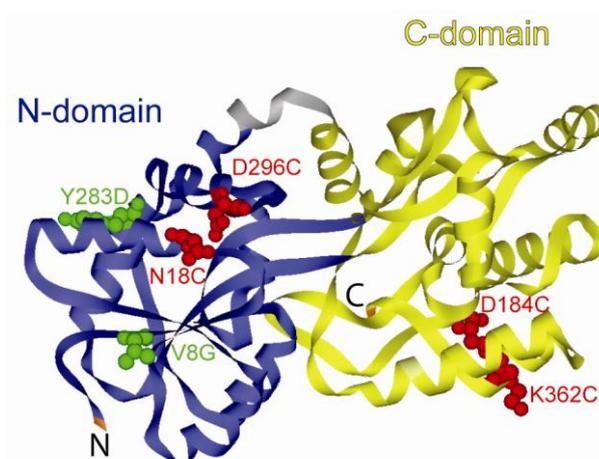


Figure 5.19: Structure of MBP Showing Cysteine Pair Mutations. Ribbon diagram of the structure of MBP, indicating the positions of DM-MBP mutations V8G and Y283D (green), and the cysteine mutations N18C-D296C and D184C-K362C (red) are shown. The two discontinuous domains are shown in blue (N-domain) and yellow (C-domain).

DM-MBP cysteine mutants were purified in the reduced state in presence of DTT. Proteins were oxidized to form disulfide bonds in presence of 5 μM CuCl_2 , after removal of DTT. Formation of oxidized proteins was confirmed by LC-MS analysis (Figure 5.20).

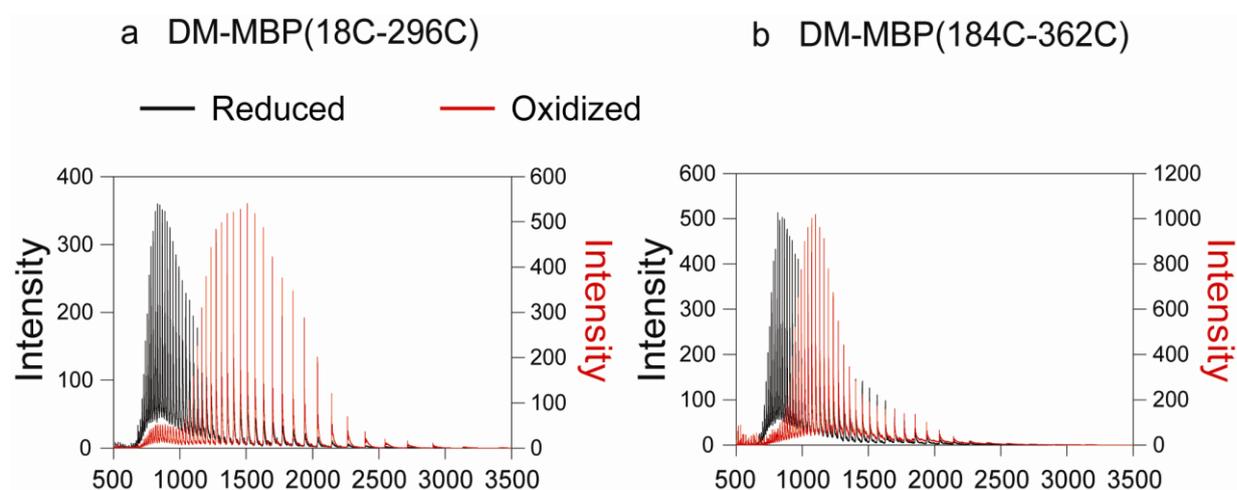


Figure 5.20: Characterization of Cysteine Pair Mutants of DM-MBP. Charge state distribution was measured by LC-MS. (a) DM-MBP(18C-296C) and (b) DM-MBP(184C-362C) were analyzed under reducing and oxidizing conditions on a Waters Synapt HDMS ESI-QToF mass spectrometer. The distribution of the charge state centers distinguishes between the reduced (black) and disulfide-bonded proteins (red) (Zhang et al., 2001).

After characterizing the cysteine pair mutants of DM-MBP in both oxidized and reduced state, we next performed refolding experiments by monitoring trp fluorescence as described before for DM-MBP. Indeed, the oxidized (ox) proteins refolded ~5-fold faster as compared to the reduced (red) proteins (Figure 5.21). This indicated that constraining native like contacts facilitate the formation of native state for DM-MBP.

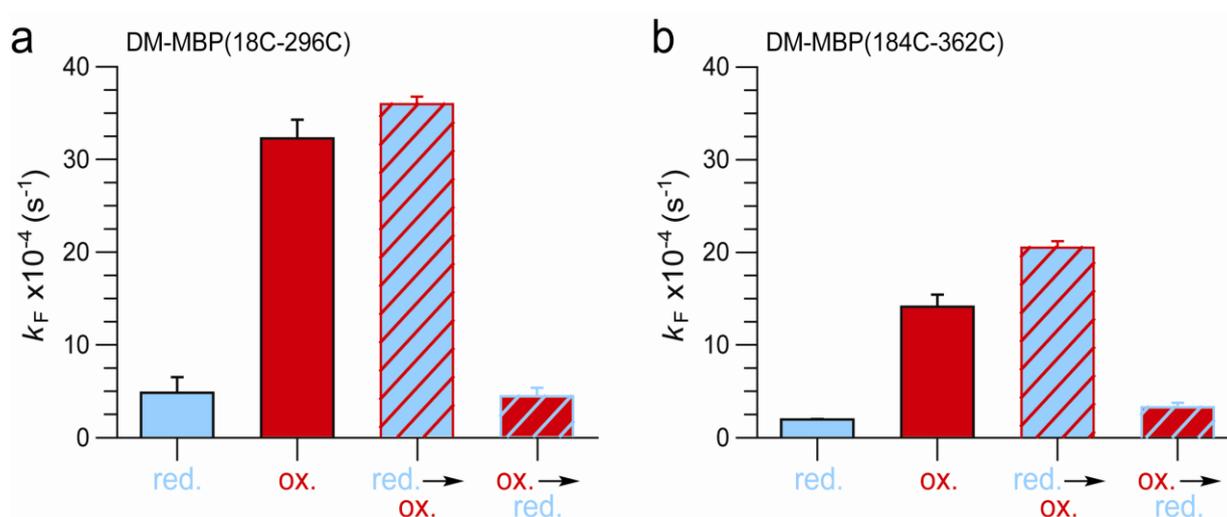


Figure 5.21: Spontaneous Refolding of Cysteine Mutants of DM-MBP.

Refolding of (a) DM-MBP(18C-296C) and (b) DM-MBP(184C-362C) was initiated after 100-fold dilution into buffer A from respective denatured proteins in 6 M GuHCl (Final GuHCl 60 mM; Protein 250 nM). Refolding was followed by monitoring Trp fluorescence. Folding of reduced protein (red.), oxidized protein (ox.), upon oxidizing the reduced protein 5 s after initiation of refolding (red. \rightarrow ox.) by adding 5 μ M CuCl_2 , or upon reducing the oxidized protein 5 s after initiation of refolding (ox. \rightarrow red.) by adding 5 mM DTT was performed. Standard deviation of 3 independent measurements are shown.

Next, we wanted to determine whether constraining the unfolded state was necessary for this effect. We took advantage of the finding that DM-MBP undergoes conformational collapse within milliseconds after initiation of spontaneous folding (Sharma et al., 2008). We initiated the refolding of reduced protein after dilution from the denaturant and added CuCl_2 (oxidizing agent) after 5 s of initiation of folding. Interestingly, folding was accelerated to the

same extent as for the oxidized proteins (Figure 5.21). Reduced unfolded state, upon dilution from denaturant collapses to a compact intermediate state within milliseconds, which is capable of forming the disulfides within seconds. To further substantiate the importance of disulfides formed in the folding intermediate, we initiated the refolding from oxidized unfolded protein and reduced the disulfides with 5 mM DTT after 5 s of initiation of refolding (Figure 5.21). The refolding rate was identical to that of the reduced protein implying that entropic constraint in the intermediate state is necessary and sufficient to accelerate the spontaneous folding reaction. Refolding yields for oxidized and reduced forms of DM-MBP(18C-296C) and DM-MBP(184C-362C) were between 80-100% (Figure 5.22).

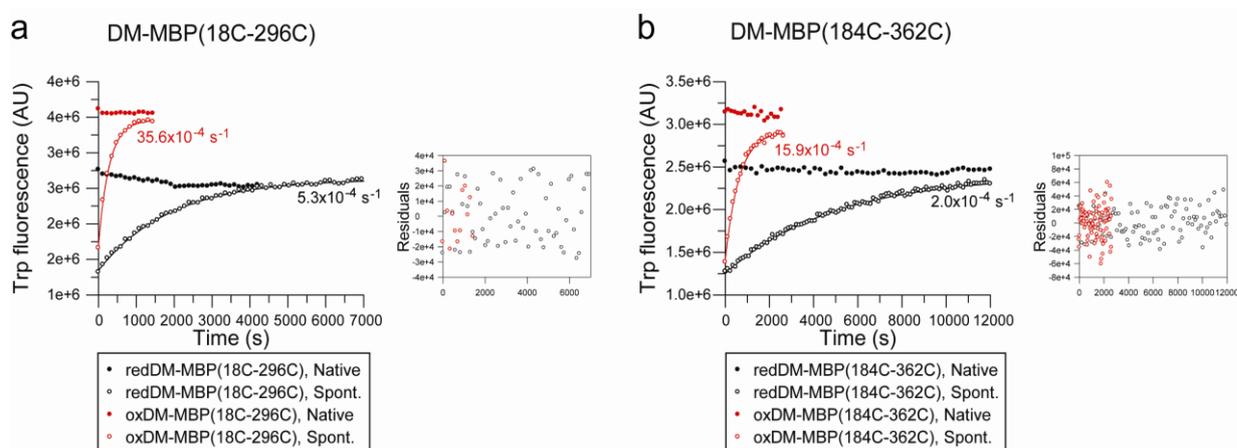


Figure 5.22: Single Exponential Rates and Yield of Refolding of DM-MBP Cysteine Mutants. Representative raw data of rates and yields of spontaneous refolding of (a) DM-MBP(18C-296C) and (b) DM-MBP(184C-362C) at final concentration of 250 nM are shown. Reduced (red.) or oxidized (ox.) DM-MBP cysteine mutants were denatured in 6 M GuHCl and diluted 100-fold into buffer A at 25°C. Refolding was monitored by Trp fluorescence at 345 nm. Note that the Trp fluorescence of the reduced proteins (black curves) is quenched relative to the oxidized proteins (red curves). The apparent rates of folding fitted to a single exponential. The residuals of the fits are shown in the inserts. Rates are indicated next to the respective folding curves. The fluorescence intensities of native protein controls are shown. Refolding yields were between 80-100%.

We then checked for the disulfide bond formation rate during spontaneous folding of cysteine mutants of DM-MBP, and found that disulfide bond formation occurred within seconds of addition of oxidizing agent in the refolding reaction (Fig 5.23)

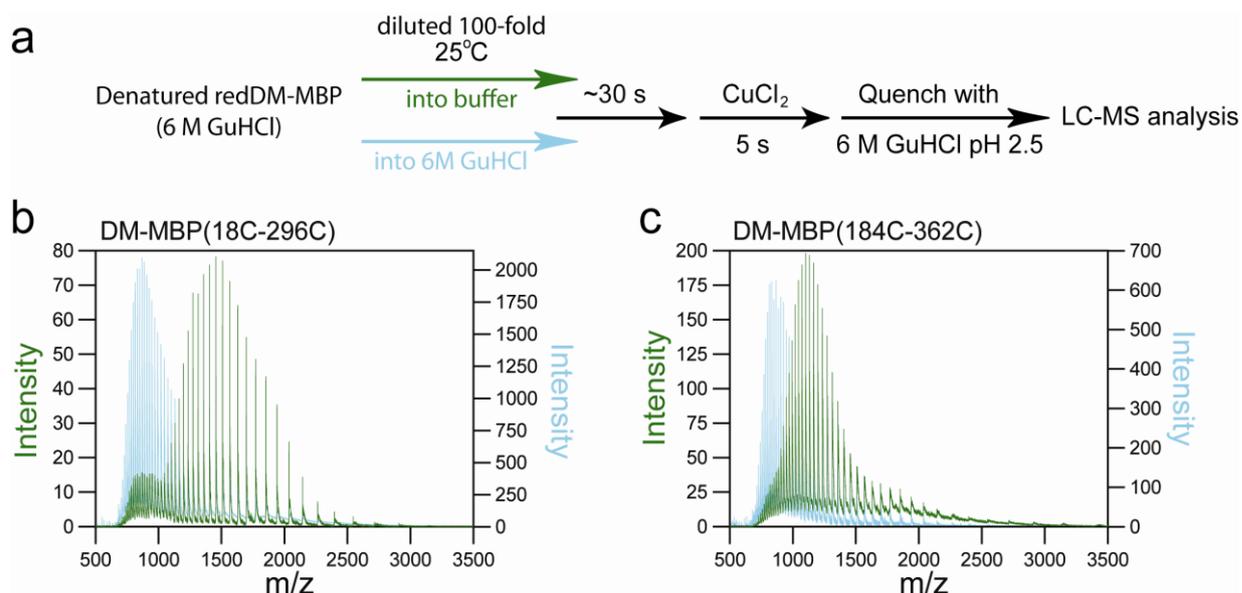


Figure 5.23: Rapid Disulfide-Bond Formation Upon Protein Collapse.

(a) Schematic representation of the experiment is shown. Unfolded protein in 6 M GuHCl was diluted 100-fold in buffer A followed by addition of CuCl₂ after 30 s. The disulfide bond formation was quenched by addition of 6 M GuHCl at pH 2.5. (b) Reduced (red.) DM-MBP(18C-296C) and (c) DM-MBP(184C-362C) (final ~2 μM) were unfolded in 6 M GuHCl and diluted 100-fold into buffer B, is shown as green trace. Reduced unfolded protein was desalted to remove DTT in presence of 6 M GuHCl. CuCl₂ was added after 5 s of initiation of refolding for oxidation of cysteines. Disulfide bond formation was quenched by buffer B containing 6 M GuHCl after 30 s of addition of CuCl₂ shown in blue. After quenching samples were subjected to LC-MS analysis.

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 bond formation would shift the distribution to more ordered conformations, thus decreasing the entropy of the folding intermediate and destabilizing it with respect to the transition state (Figure 5.18). This also indicates that the impact of disulfide bonds on folding may differ depending on the exact regions of protein that are constricted. Constraining already proximal residues can only affect the folding rate if flexibility around the region chosen for disulfide engineering, retards the formation of native state. Indeed, we found that oxDM-MBP(18C-296C) (N-domain mutant) refolded faster than oxDM-MBP(184C-362C) (C-domain mutant) (Figure 5.21a, b). This difference in kinetics correlated with the absence of hysteresis in the

unfolding-refolding curves of oxDM-MBP(18C-296C), while oxDM-MBP(184C-362C) preserved the hysteresis effect (Figures 5.24a, b). It is important to note that the reduced proteins without any disulfides preserved the hysteresis like DM-MBP (Fig 5.24c, d).

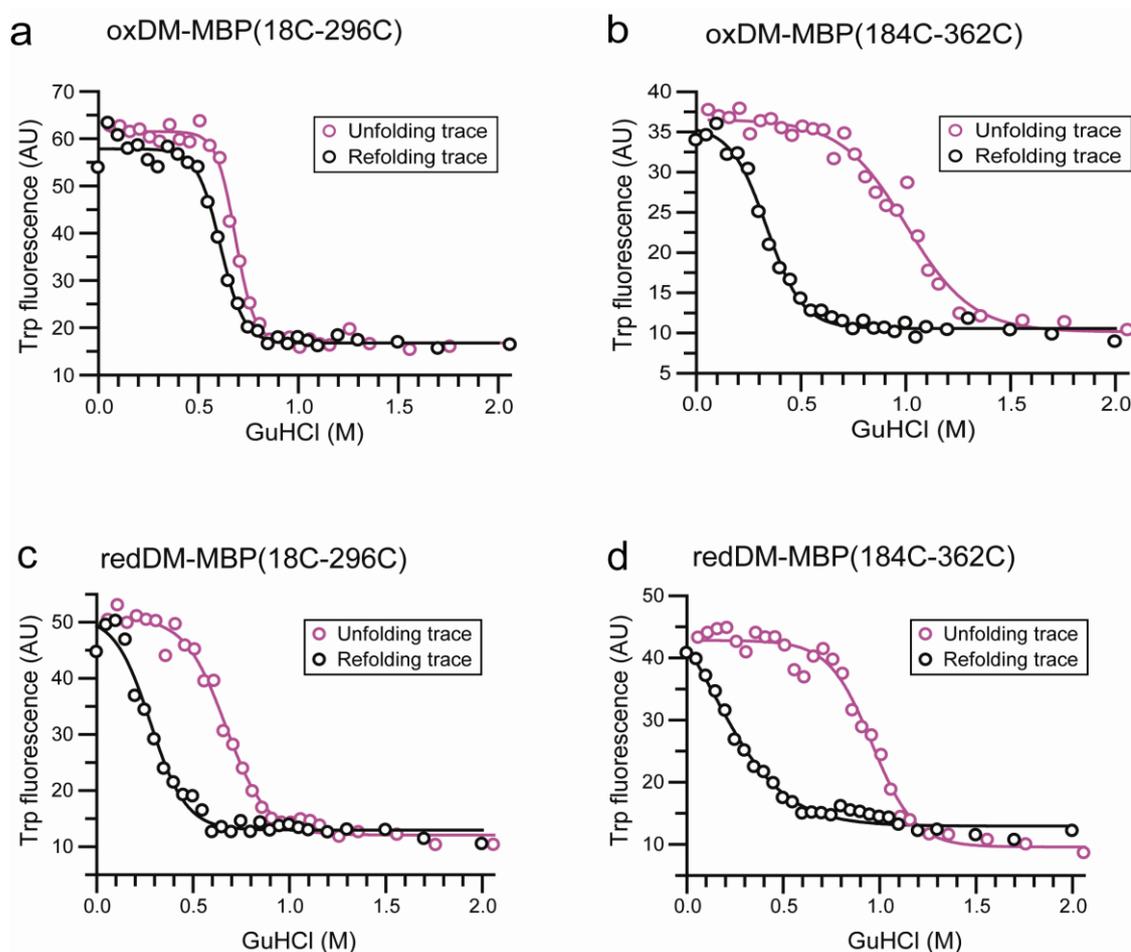


Figure 5.24: GuHCl-Dependent Unfolding and Refolding of Cysteine Mutants of DM-MBP. Unfolding and refolding of oxidized (ox) and reduced (red.), (a and c) DM-MBP(18C-298C) and (b and d) DM-MBP(184C-362C) was monitored by Trp fluorescence after incubation at different GuHCl concentrations for 12 h at 25°C as described before for Figure 5.10 a.

Also H/D exchange measurements confirmed that at 0.5 M GuHCl, oxDM-MBP(18C-296C) was more structured than oxDM-MBP(184C-362C) which is consistent with the faster refolding rate of oxDM-MBP(18C-296C) as compared to oxDM-MBP(184C-362C) (Figure 5.25a, b).

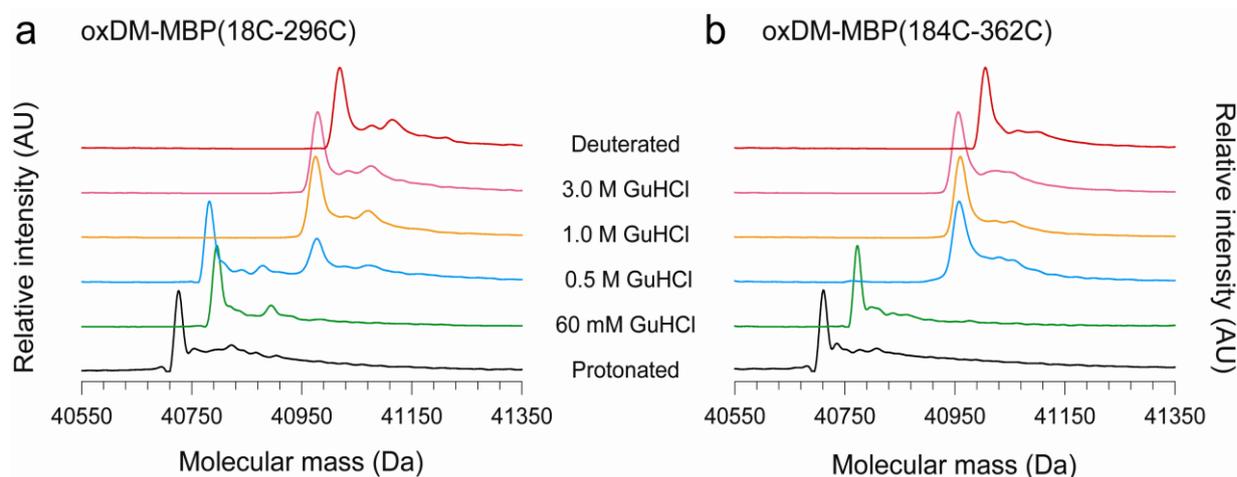


Figure 5.25: Characterization of the Kinetically Trapped Refolding Intermediate of DM-MBP Cysteine Mutants by H/D exchange. Pulsed H/D exchange experiments were performed after incubation in different denaturant concentrations. Deconvoluted mass spectra of the oxidized (a) oxDM-MBP(18C-296C) and (b) oxDM-MBP(184C-362C) is shown 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 as indicated. After incubation for 12 h samples were subjected to a 10 s deuterium pulse. The native protonated and deuterated proteins are shown as reference.

On the other hand, disulfide bonds can also stabilize or destabilize the native state (Betz, 1993). In this case, if the transition state is stabilized with respect to the intermediate state because of disulfide bonds, it may also result in faster folding (Figure 5.18). To estimate the possible effect of the disulfide bonds on the stability of the native state, we measured the rate of unfolding of the reduced and oxidized proteins. Also unfolding rate of WT-MBP and DM-MBP was measured to look for any relation between faster refolding rate and unfolding kinetics. We found that unfolding rate was essentially unaffected by the disulfide bonds or mutation (V8G, Y283D) which makes WT-MBP a slow folding mutant (Figure 5.26 and Table 2).

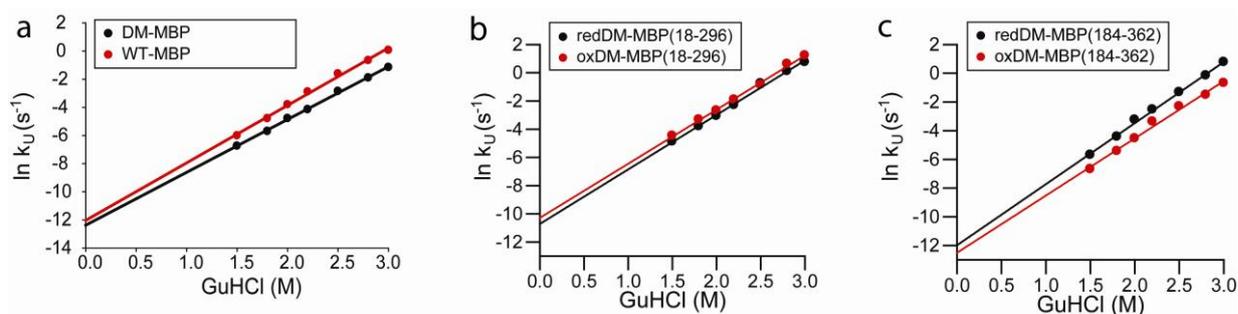


Figure 5.26: Kinetics of Unfolding. The rate of unfolding of (a) WT-MBP and DM-MBP and reduced (red) and oxidized (ox) (b) DM-MBP(18C-296C) and (c) DM-MBP(184C-362C) at different concentrations of GuHCl in buffer A was monitored in stopped-flow mixing experiments at 25°C by following the decrease in Trp fluorescence at 345 nm. The final protein concentration was 500 nM. Unfolding rates in the absence of denaturant were determined by extrapolation.

In a two-state system, the unfolding free energy of activation (ΔG_U^{Ea}) can be calculated from the measured rate of unfolding, extrapolated to zero denaturant (k_U), according to the Eyring (or Arrhenius) equation:

$$k_U = \frac{k_B T}{h} e^{\frac{-\Delta G_U^{Ea}}{RT}} \quad (9)$$

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.

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

	k_U (s ⁻¹) *	ΔG_U^{Ea} (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

Unfolding kinetics of the reduced and the oxidized cysteine mutants of DM-MBP indicate that transition state is not significantly affected by the formation of disulfide bonds, consistent with the observation that the oxidized and reduced proteins showed similar stability towards denaturant (Figures 5.24). These results suggest that introducing disulfide bonds did not stabilize the transition state, rather the observed rate acceleration is due to destabilization of the intermediate state. Thus, it is plausible that the faster folding rate of the oxidized proteins is mainly due to a reduction of the energy barrier from the intermediate to the transition state due to destabilization of the intermediate state (Figure 5.18).

5.5. Chaperonin Cage Mimics Disulfide-Mediated Constraints

Disulfide-mediated constraints accelerated the spontaneous refolding rate of DM-MBP by reducing the entropic barrier to folding. It seemed possible that confinement in the chaperonin cage may enhance the folding speed by a similar mechanism. Alternatively, the chaperonin cage may accelerate refolding by a completely different mechanism, in which case the presence of disulfides may have an additive effect. To test the effect of disulfide bond formation on the folding within the chaperonin cage, we used SR-EL, a single ring mutant of GroEL (Weissman et al., 1996). SR-EL undergoes only a single round of ATP hydrolysis upon GroES binding, resulting in stable substrate encapsulation in low salt buffer B (Hayer-Hartl et al., 1996). Since it lacks the signal from the other ring, it is very good tool to study folding inside the chaperonin cage. For SR-assisted refolding low salt conditions were used as SR-EL is not stable under high salt conditions. As described before (Figure 5.2), salt (KCl) has an effect on spontaneous refolding rate of DM-MBP whereas chaperonin assisted refolding rate is unaffected. Under low salt conditions (20 mM KCl/60 mM GuHCl), spontaneous refolding rate of DM-MBP and reduced (red) and oxidized (ox) DM-MBP(18C-296C) and DM-MBP(184C-362C) is ~2-fold faster than standard high salt conditions (200mM KCl/60 mM GuHCl) (Figure 5.27).

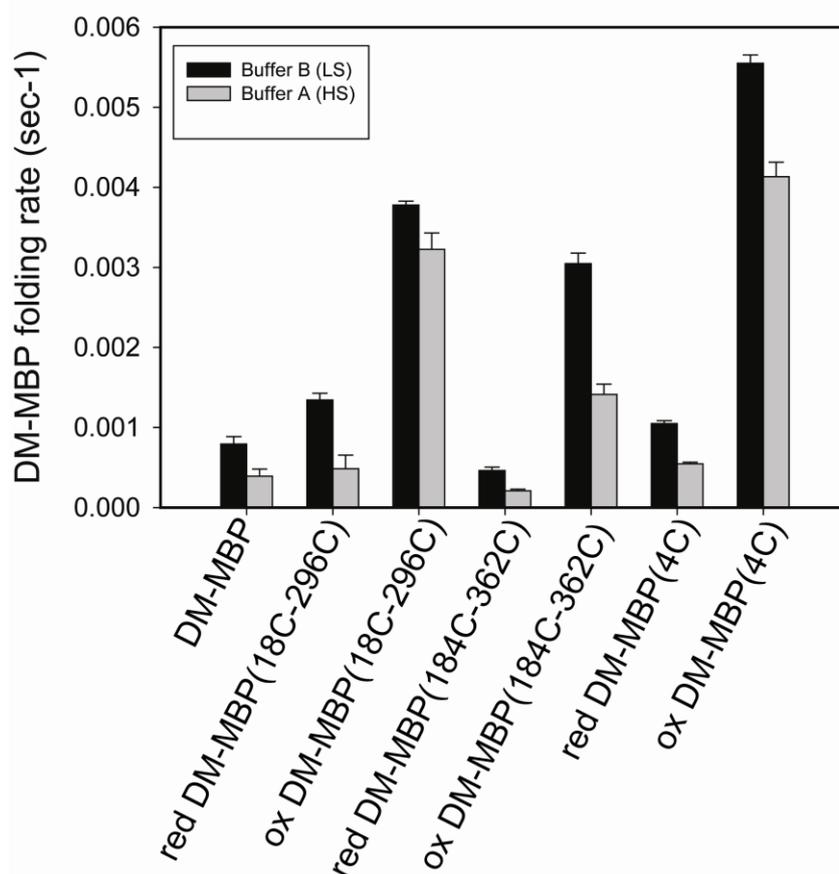


Figure 5.27: Effect of Salt on Refolding Rate of DM-MBP and DM-MBP Cysteine Mutants. 25 μ M (a) DM-MBP, (b) DM-MBP(18C-296C) and (c) DM-MBP(184C-362C) was denatured in 6 M GuHCl and spontaneous refolding was initiated by 100-fold dilution (final 60 mM GuHCl) in high salt buffer A (20 mM Tris, pH 7.5, 200 mM KCl, 5 mM Mg(OAc)₂) (Grey bar) or low salt buffer B (20 mM Tris, pH 7.5, 20 mM KCl, 5 mM Mg(OAc)₂) (Black bar) at 25°C. Refolding was monitored by Trp fluorescence. Standard deviations of 3 independent measurements are shown.

Under low salt conditions, oxDM-MBP(18C-296C) and oxDM-MBP(184C-362C) refolded at similar rates. Interestingly, SR-EL/ES accelerated the folding of both oxidized proteins ~1.5 times as compared to their spontaneous refolding rates. Also, SR-EL/ES refolding rates for both oxidized and reduced protein were similar (Figure 5.28a, b).

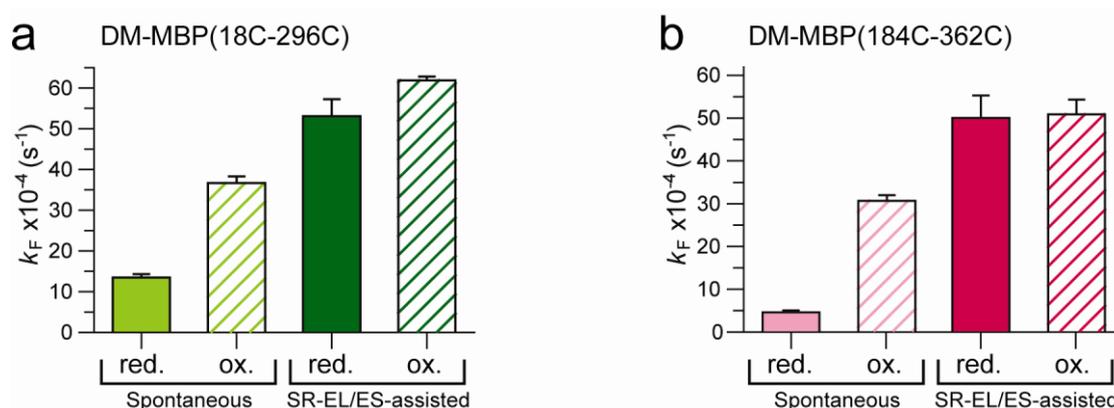


Figure 5.28: Chaperonin-Assisted Folding of DM-MBP Cysteine Mutants.

Rates of spontaneous and assisted refolding of (a) DM-MBP(18C-296C) and (b) DM-MBP(184C-362C) in the reduced (red.) and oxidized (ox.) state. 25 μM unfolded protein in 6 M GuHCl was diluted either in buffer B alone (Spontaneous refolding) or in buffer B containing 1 μM SR-EL and 2 μM GroES (Assisted refolding). Refolding was initiated by addition of 5 mM ATP. Standard deviation of three independent measurements are shown.

This indicated that SR-EL/ES assisted refolding rate is independent of the constraints imposed by disulfides. The chaperonin chamber provides an environment that favors the formation of native-like contacts, rendering the folding process insensitive to the disulfide-mediated entropic confinement.

The finding that SR-EL/ES accelerated the folding of the oxidized proteins beyond their spontaneous folding rates would be consistent with the chaperonin cavity exerting a global confinement effect, whereas disulfide-mediated structural constraints act more locally either in the N-domain or the C-domain. To examine this possibility, we constructed a DM-MBP mutant which could form disulfide bond in both the N- and C-domain simultaneously, DM-MBP(4C) with cysteines at positions 18, 296 (N-domain) and 184, 362 (C-domain). DM-MBP(4C) mutant formed disulfide bonds upon oxidation in presence of CuCl_2 which was confirmed by LC-MS analysis (Figure 5.29a) and different migration pattern on 12% SDS gel (Figure 5.29b, c) like N- and C-domain mutants of DM-MBP. Oxidized and reduced forms of

DM-MBP(4C) was native as they bound maltose as efficiently as DM-MBP, DM-MBP (18C-296C) and DM-MBP (184C-362C) in the native state (Figure 5.29d).

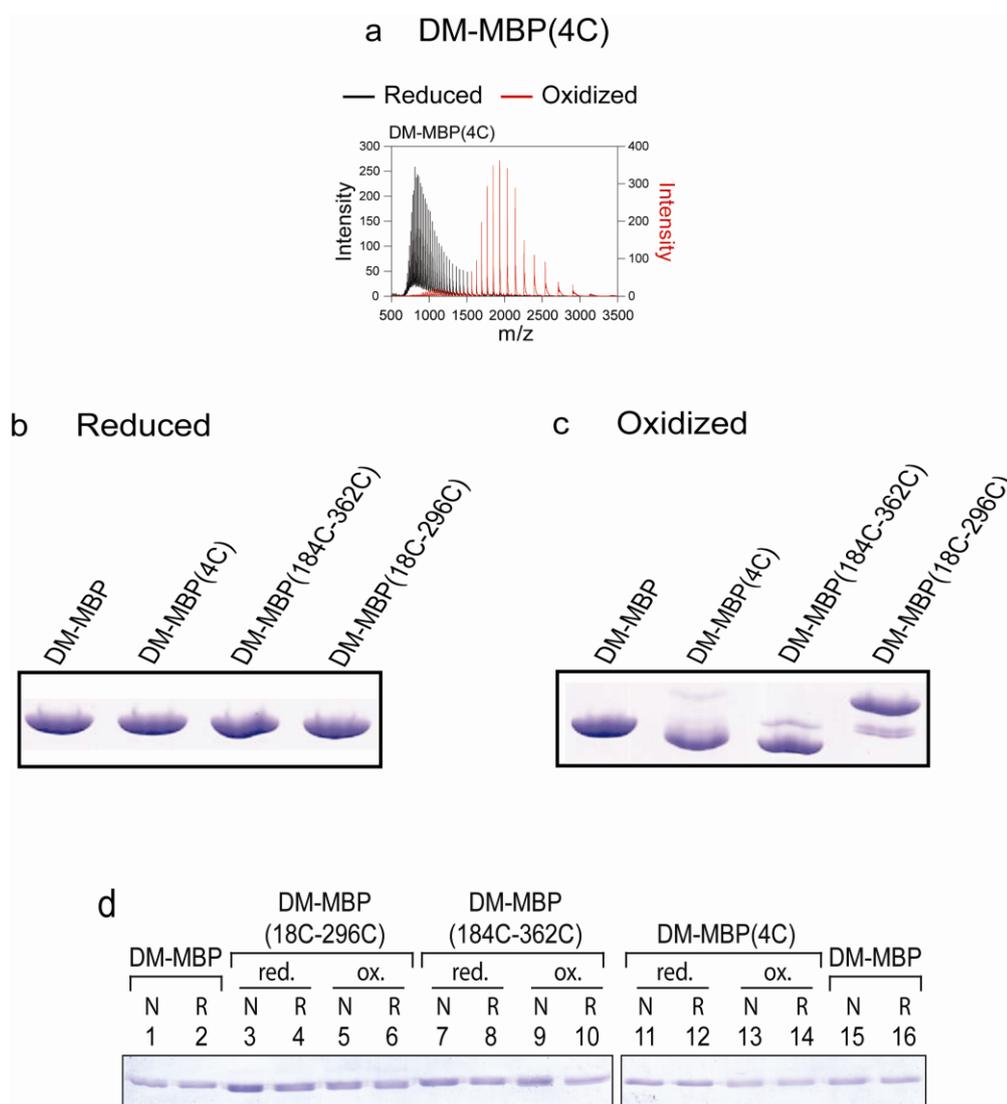


Figure 5.29: Characterization of DM-MBP(4C). (a) Charge state distribution of DM-MBP(4C) was measured by LC-MS under reducing and oxidizing conditions on a Waters Synapt HDMS ESI-QToF mass spectrometer. The distribution of the charge state centers distinguishes between the reduced (black) and disulfide-bonded proteins (red) (Zhang et al., 2001). (b) Oxidized and (c) reduced forms of DM-MBP(4C) shows different migration pattern on 12% SDS gel stained with coomassie. (d) Native (N) and spontaneously refolded (R) DM-MBP, as well as reduced (red.) and oxidized (ox.) DM-MBP(4C) (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.

After complete characterization of reduced (red.) and (ox.) state of DM-MBP(4C) mutant, refolding rate was measured. Surprisingly, the spontaneous refolding rate of oxDM-MBP(4C) was accelerated to the rate of chaperonin-assisted folding rate. Similar to the single disulfide bonds, the rate of assisted folding was the same for the oxidized and reduced protein (Figure 5.30). For DM-MBP(4C), experiment where refolding was initiated from reduced unfolded state and oxidizing agent was added after 5 s was not done because of the possibility of wrong intramolecular disulfide bond formation.

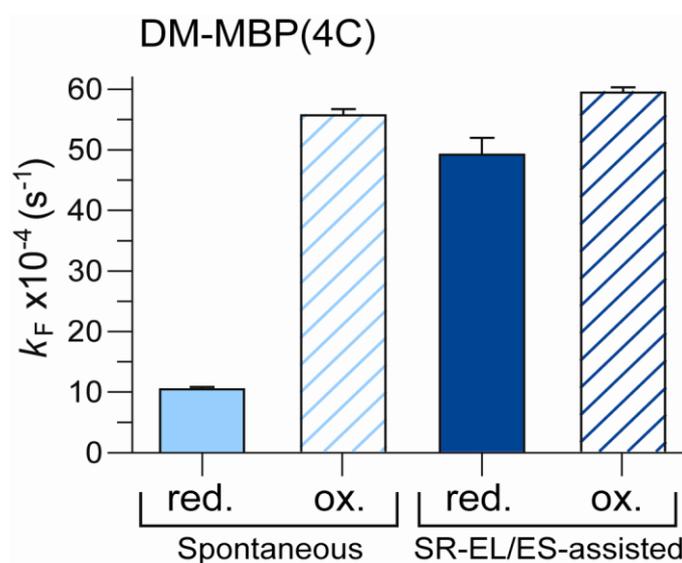


Figure 5.30: Rates of Spontaneous and Assisted Refolding of DM-MBP(4C).

25 μM DM-MBP(4C) was unfolded in 6 M GuHCl and diluted 100-fold in buffer B alone (Spontaneous refolding) or in buffer containing 1 μM SR-EL and 2 μM GroES. Refolding was initiated by addition of 5 mM ATP (Assisted refolding). Final concentration of protein was 250 nM. Refolding was followed by monitoring Trp fluorescence. Standard deviation of three independent measurements are shown.

The faster spontaneous refolding rate of oxDM-MBP(4C) as compared to redDM-MBP(4C) is also reflected in the hysteresis profile where oxDM-MBP(4C) does not show any hysteresis whereas redDM-MBP(4C) preserves the hysteresis (Figure 5.31a, b).

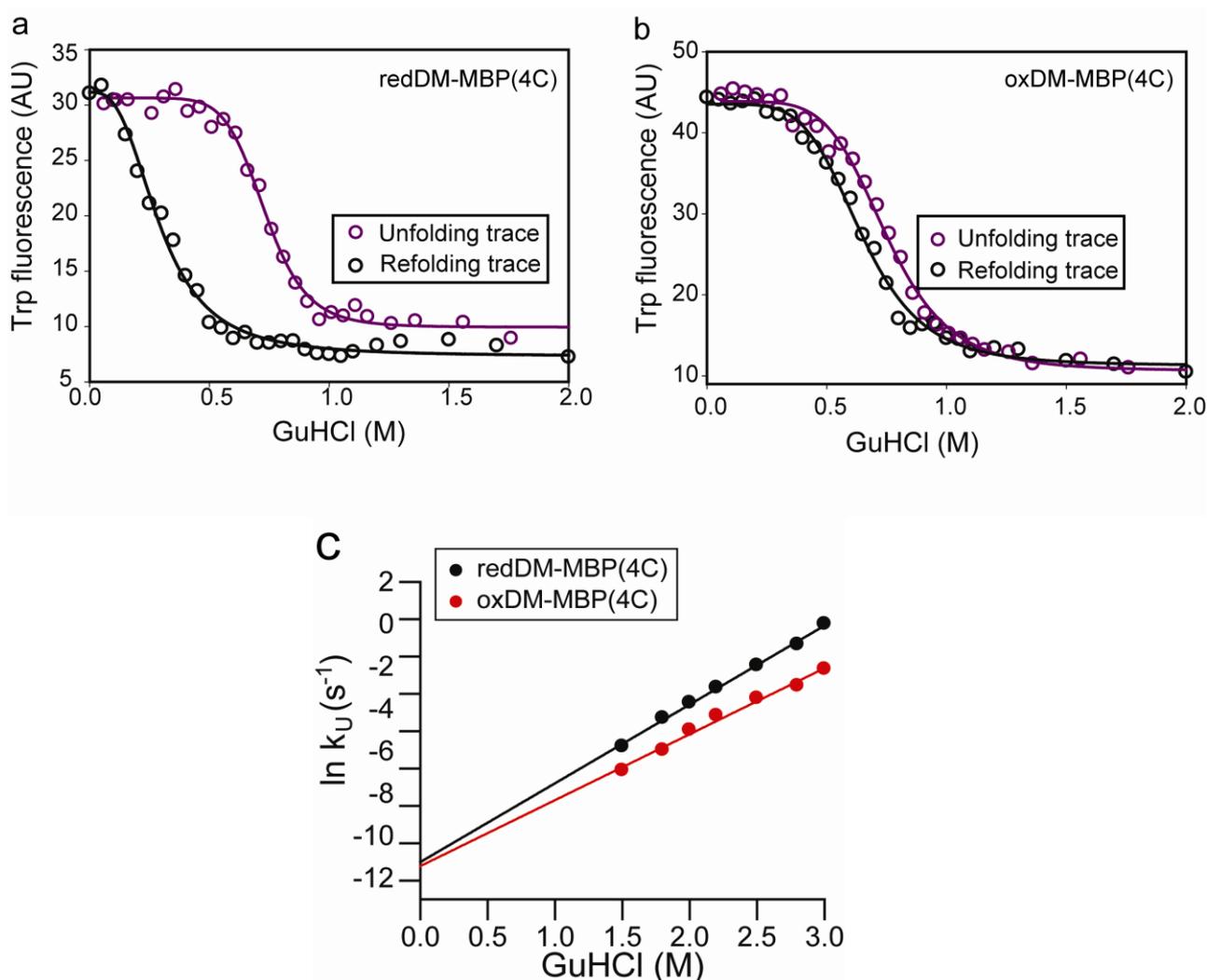


Figure 5.31: GuHCl-Dependent Unfolding and Refolding and Unfolding Kinetics of DM-MBP(4C). Unfolding and refolding of (a) oxidized (ox) and (b) reduced (red.) DM-MBP(4C) was monitored by Trp fluorescence after incubation at different GuHCl concentrations for 12 h at 25°C as described before for Figure 4.8a. (c) The rate of unfolding of oxidized (ox) and reduced (red) DM-MBP(4C) 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.

Unfolding rate was also measured for reduced and oxidized DM-MBP(4C) (Figure 5.31c) (Table 2). We found that like N- and C-domain cysteine mutants, DM-MBP(4C) mutant also does not exhibit a difference in the unfolding rate for oxidized and reduced proteins indicating that the destabilization of the intermediate by disulfide bonds is the reason for faster rate acceleration observed. This suggests that the effect of constraining both the

domains of DM-MBP is constraining the protein even more than N- or C-domain disulfide bonds separately and is comparable to the global confinement effect exerted by the chaperonin.

Next, we wanted to investigate whether rate acceleration caused by long-range disulfide bonds energetically mimic the rate acceleration due to confinement by the chaperonin cage. Since the temperature dependence of reaction rate is sensitive to the activation energy barrier, it can reflect the similarity in rate limiting step of the reaction. The spontaneous refolding of DM-MBP, redDM-MBP(18C-296C) and redDM-MBP(4C) was insensitive to temperature variation between 15°C and 25°C (Figure 5.32), consistent with large entropic folding barrier (Bicout and Szabo, 2000). SR-EL/ES-assisted refolding rate increased strongly with increasing temperature (Figure 5.32), indicating that the rate-limiting step of folding inside the SR-EL/ES cavity differs from that of spontaneous folding. Temperature dependence for folding oxDM-MBP(18C-296C) and oxDM-MBP(4C) was found to be similar to SR-EL/ES-assisted refolding indicating the similarity in the rate limiting step between SR-EL/ES-assisted refolding and disulfide-mediated refolding.

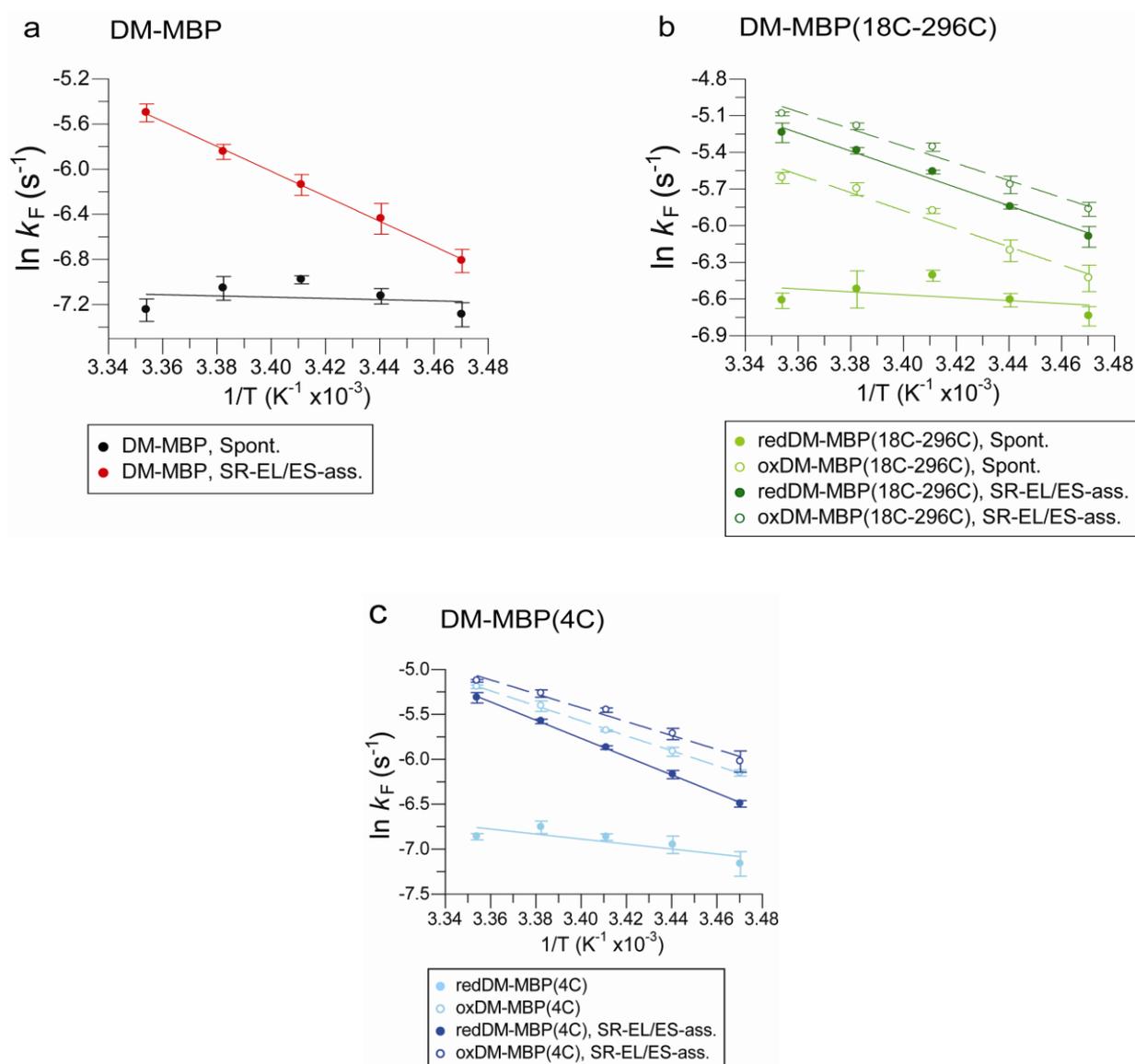


Figure 5.32: Spontaneous and Assisted Refolding of DM-MBP and its Cysteine Mutants at Different Temperatures. Rates of spontaneous and SR-EL/ES-assisted refolding of (a) DM-MBP, reduced and oxidized (b) DM-MBP(18C-296C) and (c) DM-MBP(4C) at varying temperature was measured by following trp fluorescence. Refolding was performed at 250 nM DM-MBP in buffer B as described in Figure 5.26. Temperature was varied from 15°C-25°C. Standard deviation of three independent measurements are shown.

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 (Figures 5.18), we used the temperature dependence of folding rate to approximately estimate the enthalpic and entropic contributions to the folding energy barrier (Table 3).

According to the transition state theory, the rate constant of folding, k_F , of a two state reaction is defined by the Eyring (or Arrhenius) equation (10):

$$k_F = \frac{k_B T}{h} e^{\frac{-\Delta G_F^{Ea}}{RT}} \quad (10)$$

where ΔG_{RT}^{Ea} is the activation energy of folding, k_B is the Boltzmann's constant (1.381×10^{-23} JK⁻¹), and h is Planck's constant (6.626×10^{-34} Js), R is the gas constant (8.31 JK⁻¹ mol⁻¹) and T is temperature in Kelvin.

Using the definition of the free energy:

$$-\Delta G_F^{Ea} = \Delta H_F - T\Delta S_F \quad (11)$$

the Eyring equation can be rewritten as

$$\ln(k_F) = \ln\left(\frac{k_B T}{h}\right) + \frac{\Delta S_F}{R} - \frac{\Delta H_F}{R} \left(\frac{1}{T}\right) \quad (12)$$

As $\ln(T)$ depends weakly on $1/T$ in the range measured, equation (12) 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. The results are given in Table 3.

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

	$k_F \times 10^{-4} (s^{-1})^*$	$\Delta G_F^{Ea} (KJ mol^{-1})$	$\Delta H_F (KJ mol^{-1})$	$T\Delta S_F (KJ mol^{-1})$
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), SR-KKK2/ES-ass.	10.5±0.4	90.0	17.7	-72
oxDM-MBP(4C), SR-KKK2/ES-ass.	62.8±0.9	85.5	51.9	-34

*Apparent refolding rate in buffer B (three independent experiments).

As expected, the activation energy barrier for spontaneous folding of reduced proteins was mostly entropic in nature. In contrast, the folding rate of the oxidized cysteine mutants showed a pronounced, positive temperature dependence (Figure 5.32), indicating that the activation energy has gained a significant enthalpic component and the entropic contribution is largely reduced (Table 3). The enthalpic component may result from side-chain friction during folding of the constrained protein. Strikingly, the SR-EL/ES-assisted folding of DM-MBP and the reduced cysteine mutants displayed a similar positive temperature dependence (Figure 5.32 a-c), suggesting a strongly reduced entropic barrier (Table 3). Thus, the spontaneous folding of the oxidized proteins and the assisted folding of the reduced proteins in the chaperonin cage have similar rate limiting steps, consistent with a common mechanism of accelerated folding by entropic confinement. The enthalpic component of the activation energy barrier may reflect side-chain friction during folding, resulting either from disulfide-mediated constraints or steric restrictions imposed by the chaperonin cage.

5.6. Charge Clusters in the Cage Wall Play an Active Role in Promoting Folding

The wall of the GroEL/ES cage is highly charged, exhibiting a net charge of minus 42 (147 positively and 189 negatively charged residues) (Xu et al., 1997). Most of these residues are highly conserved among GroEL homologs, although they have no apparent role in either substrate or GroES binding (Brocchieri and Karlin, 2000; Stan et al., 2003). We were interested in determining whether these charge properties are important in promoting DM-MBP folding or mere geometric confinement within the chaperonin cavity is sufficient to favor the formation of native-like contacts. To this end, we took advantage of a mutant of SR-EL, SR-KKK2 (Tang et al., 2006), which has a cavity net charge of zero due to mutation of

three conserved, negatively charged residues per subunit (D359, D361, E363) to lysines (Figure 5.33).

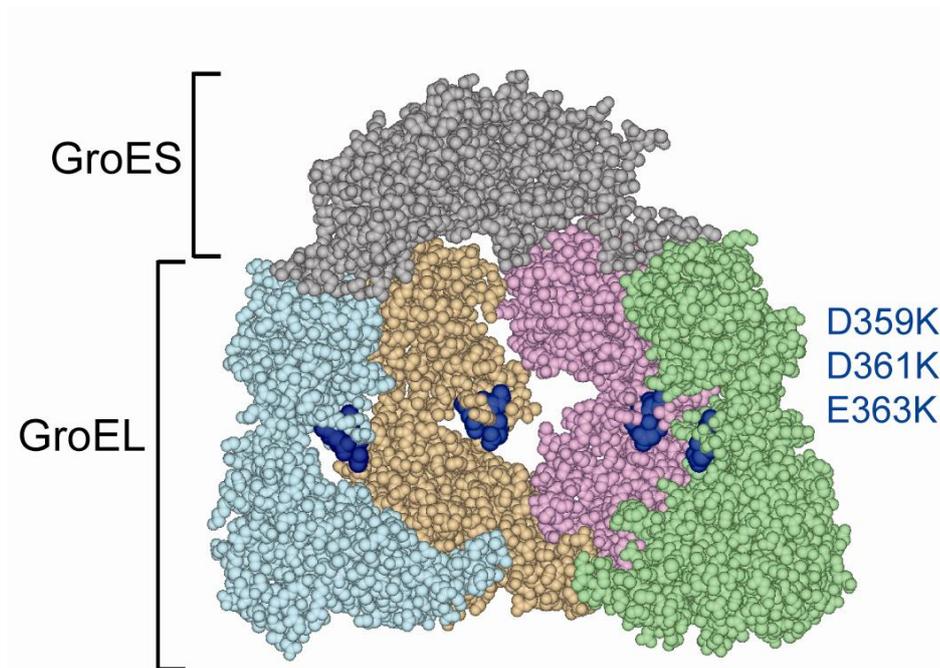


Figure 5.33: Space-filling Model of GroEL/ES-(ADP)₇ Complex. Structure of GroEL/ES complex is shown (Xu et al., 1997 pdb 1AON, DS ViewerPro, offering a view into the chaperonin folding cage. The negatively charged residues mutated in SR-KKK2 (D359K; D361K; E363K) are highlighted in blue.

SR-KKK2 is unable to accelerate the folding of DM-MBP like SR-WT, while the rate of WT-MBP folding is essentially unaffected (Tang et al., 2006) indicating that the mutant SR-KKK2 does not confer any inhibitory effect on folding. Although it has been shown that SR-KKK2 binds and encapsulates DM-MBP as efficiently as SR-EL (Tang et al., 2006) (Figure 5.34).

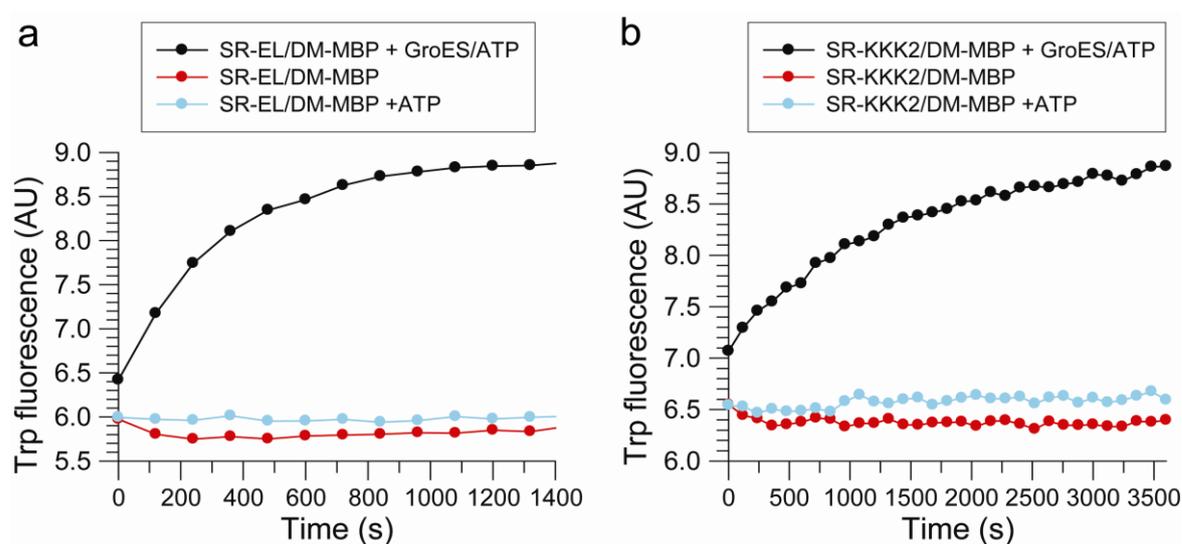


Figure 5.34: Binding Affinity of the SR-KKK2 Mutant for DM-MBP. Refolding of GuHCl-denatured DM-MBP (25 μ M) at 25°C upon 100-fold dilution into buffer B containing 1.0 μ M SR-EL or SR-KKK2 (red); 1.0 μ M SR-EL/2 mM ATP or SR-KKK2/2 mM ATP (blue); 1.0 μ M SR-EL/2 μ M GroES/2 mM ATP or SR-KKK2/2 μ M GroES/2 mM ATP (black). Inhibition of spontaneous DM-MBP refolding by SR-KKK2 (and SR-EL) in the presence of ATP but absence of GroES indicates that SR-KKK2 has a similar affinity for DM-MBP as SR-EL (or GroEL) (Tang et al., 2006; Tang et al., 2008).

This suggests that the SR-KKK2 mutant may have lost the ability to reduce the entropic barrier of DM-MBP folding and hence cannot accelerate the folding rate like SR-WT. If this is the case, then rate limiting step of folding inside the mutant chaperonin may be similar to that of the spontaneous folding.

To address this possibility, we measured the temperature dependence of SR-KKK2/ES-assisted folding. (Note that SR-EL and SR-KKK2 have similar ATPase rate and undergo only a single round of ATP hydrolysis upon GroES binding). In contrast to the SR-EL/ES-assisted folding, SR-KKK2/ES-assisted folding of DM-MBP was temperature independent, similar to the spontaneous folding of DM-MBP (Figure 5.35) (Table 3).

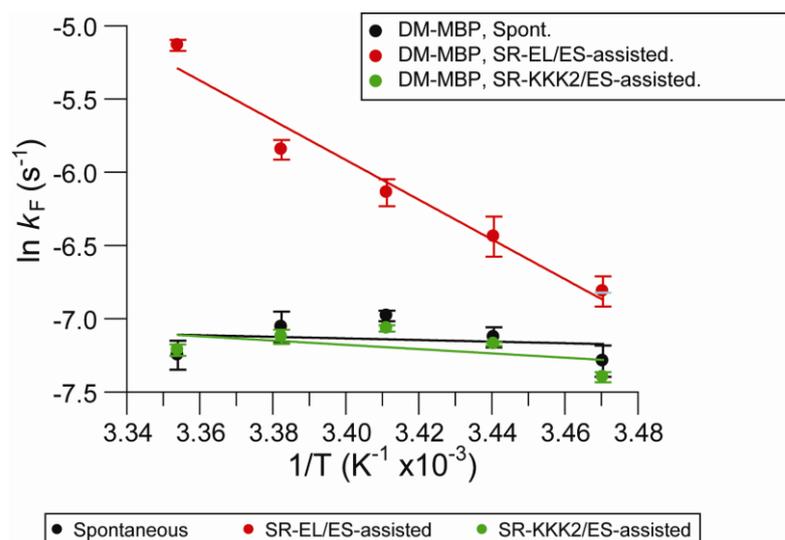


Figure 5.35: Temperature Dependence of Spontaneous and Assisted Folding of DM-MBP. Spontaneous and assisted refolding of DM-MBP was performed in presence of SR-EL/ES or SR-KKK2/ES as indicated. 25 μ M DM-MBP was unfolded in 6M GuHCl and diluted 100-fold in buffer B alone (Spontaneous refolding), or buffer containing 1 μ M SR-EL/SR-KKK2 and 2 μ M GroES. Reaction was initiated by addition of 5 mM ATP (Assisted refolding). Refolding was monitored by following Trp fluorescence at different temperatures (15-25°C). Standard deviation of three independent measurements are shown.

To further check for the similarity in rate limiting steps for spontaneous and SR-KKK2/ES assisted refolding, we performed folding experiments under different physical conditions. Effect of increasing the denaturant concentration which may stabilize the kinetically trapped intermediate, and hence decelerate folding rate, was tested. Indeed, the rate of spontaneous and SR-KKK2/ES-assisted folding of DM-MBP decreased with increasing concentration of GuHCl (15-60 mM) (Figure 5.36a). Since GuHCl dependence of the folding rate is a measure of disorder in the folding intermediate, this finding also suggests that the SR-KKK2 cavity is unable to induce structure formation in DM-MBP during folding. In contrast, the SR-EL/ES-assisted refolding rate was independent of denaturant (Figure 5.36a), consistent with the ability of the wild-type chaperonin to promote structure formation in the intermediate.

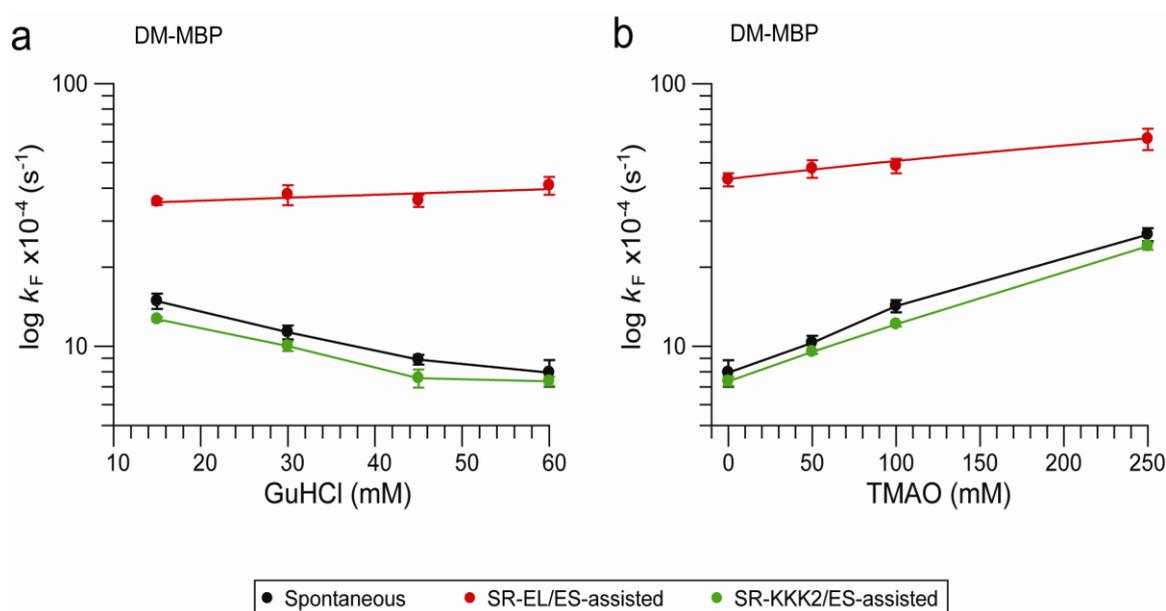


Figure 5.36: Spontaneous and Assisted Refolding of DM-MBP in Different Conditions.

Spontaneous and assisted refolding of DM-MBP was performed in presence of SR-EL/SR-KKK2 as indicated. (a) 50 μM DM-MBP denatured in 3 M GuHCl was 200-fold diluted in buffer B alone (Spontaneous refolding), or buffer containing 1 μM SR-EL/SR-KKK2 and 2 μM GroES. Reaction was initiated by addition of 5 mM ATP (Assisted refolding). Refolding was monitored by following Trp fluorescence at 25°C. Both spontaneous and assisted refolding was performed in buffer B with increasing concentration of GuHCl (15-60 mM) (b) Refolding was performed as described in Figure 5.34 at 25°C at varying concentrations of TMAO (0-250 mM). Standard deviation of three independent measurements are shown.

Next, we tested the influence of trimethylamine *N*-oxide (TMAO) on folding, an osmolyte known to reduce structural flexibility in proteins (Qu et al., 2003), presumably via the enhancement of water structure (Zou et al., 2002). Interestingly, the rate of spontaneous folding and SR-KKK2/ES-assisted folding showed a positive dependence on TMAO concentration, while the SR-EL/ES-assisted folding remained essentially unchanged (Figure 5.36b). Taken together, these experiments suggests that rate limiting step of folding inside the mutant chaperonin SR-KKK2 is similar to that of the spontaneous folding whereas folding inside the SR-EL/ES cavity differs from the spontaneous folding of DM-MBP. Due to altered charge property inside the cavity, it seems possible that SR-KKK2 has lost (or significantly

reduced) the ability to form native like contacts (or entropically destabilize the DM-MBP folding intermediates).

To test this possibility, we measured the refolding rate of oxDM-MBP(18C-296C) which facilitates formation of native like contacts as described before (Figure 5.21a), in presence of SR-KKK2/ES. Notably, SR-KKK2/ES was able to accelerate the refolding rate of oxDM-MBP(18C-296C) to the extent of SR-EL/ES where as folding of redDM-MBP(18C-296C) inside SR-KKK2/ES cavity was not accelerated (Figure 5.36). This suggests that the passive geometric confinement contributes to accelerating the folding rate of DM-MBP only when native-like contact formation is favored by the charged cavity wall or by constraints imposed by disulfide bonds.

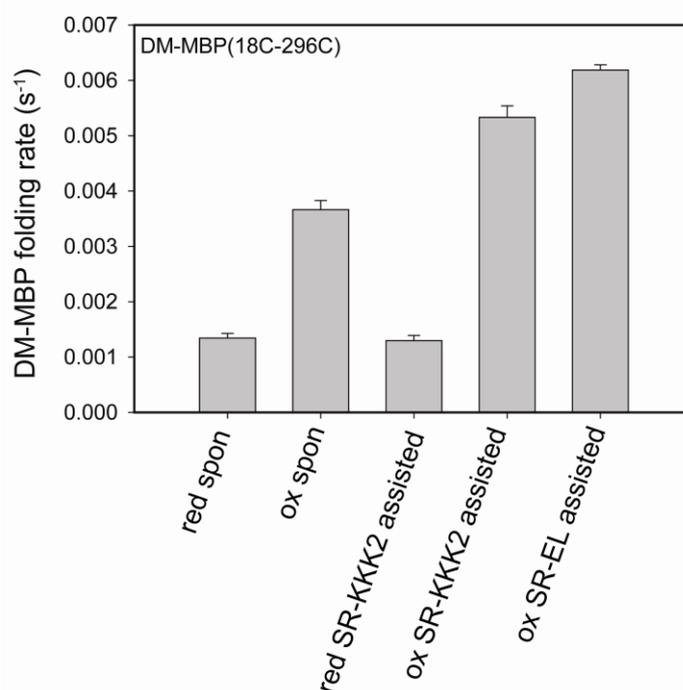


Figure 5.37: Spontaneous and Assisted Refolding Rate of DM-MBP(18C-296C).

Spontaneous and assisted refolding rate of oxidized (ox.) and reduced (red.) DM-MBP(18C-296C) at 25°C was measured in presence of SR-EL/SR-KKK2 as indicated. Refolding was performed as described in Figure 5.28.

Also, introducing disulfide bonds restored the positive temperature dependence of folding, as shown for spontaneous and SR-KKK2 assisted folding of oxDM-MBP(18C-296C) (Figure 5.38a). Similar trend was observed when refolding rate of oxDM-MBP(4C) was analyzed at different temperatures with SR-KKK2/ES (Figure 5.38b).

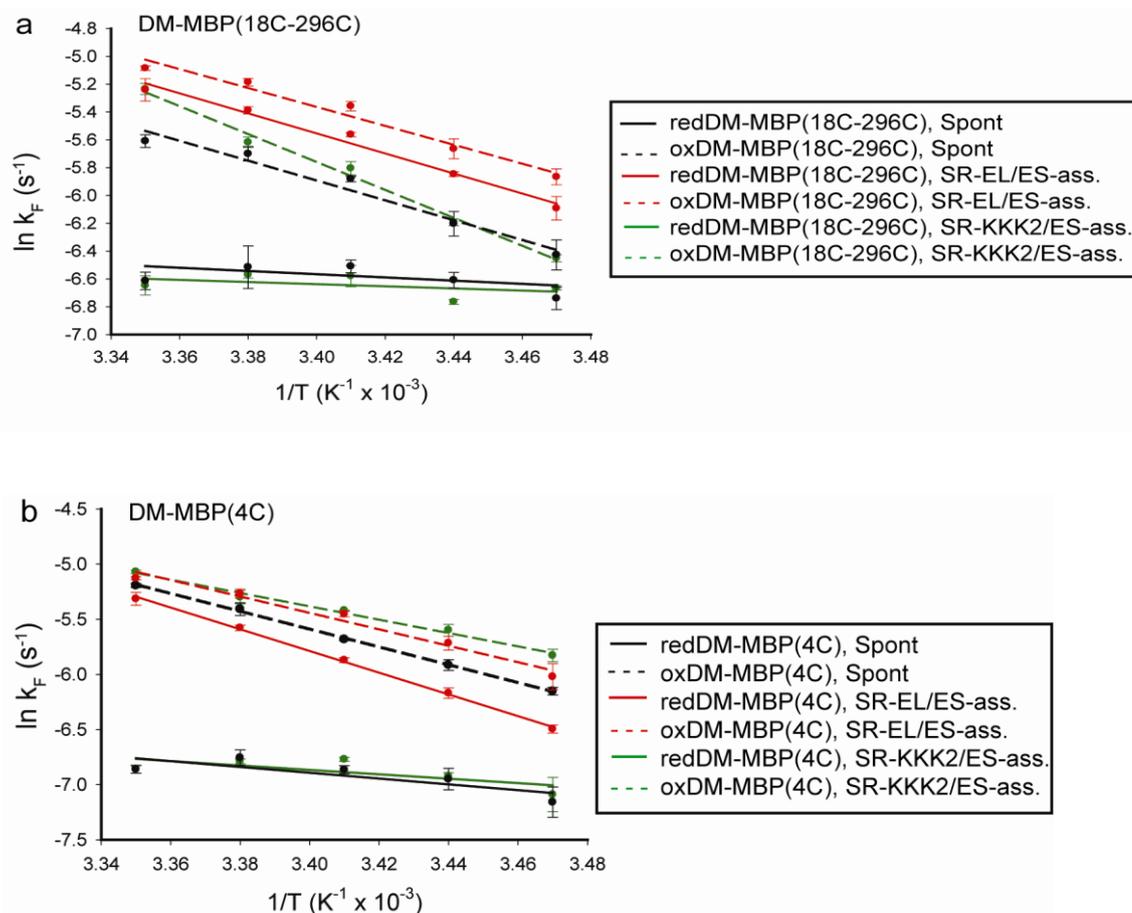


Figure 5.38: Temperature Dependence of Spontaneous and Assisted Folding of DM-MBP Cysteine mutants. Spontaneous and assisted refolding of (a) DM-MBP(18C-296C) and (b) DM-MBP(4C) was performed in presence of SR-EL/ES or SR-KKK2/ES as indicated. Refolding was performed at different temperatures (15-25°C), as described in Figure 5.33. Standard deviation of three independent measurements are shown.

This clearly demonstrates that the passive geometric confinement (like SR-KKK2/ES) can accelerate the folding rate of DM-MBP only when native-like contact formation is favored as mimicked by the formation of disulfide bonds. Charges lining the cavity plays an important role in enforcing formation of native like contacts thus reducing the entropic barrier to folding.

This is consistent with the proposal that charged lining of the cavity may promote protein compaction by an ordering effect on water structure (England and Pande, 2008). Together, these results provide evidence that by removing its net negative charge, the chaperonin cavity is converted from an active to a largely passive folding environment.

5.7. Charges in the Cavity Optimize the Folding Capacity of Chaperonin

Having observed the effect of the charge properties of chaperonin cavity on folding of DM-MBP, we wanted to know how charges influence the folding of other known substrate proteins. To this end we took advantage of the flexible C-terminal tail of GroEL which consists of 4 Gly-Gly-Met (GGM) repeats and ends with an additional Met residue. These [GGM]₄M sequences extend from the equatorial domain into the GroEL cavity but are not resolved in the crystal structure (Braig et al., 1994). It has been shown previously that GroEL/ES cavity size (changed by deletion or addition of GGM repeats) plays an important role in protein folding (Tang et al., 2006). For our study, we engineered a series of mutants where we systematically mutated [GGM]₄M to [GGX]₄X where X represents either a hydrophobic, charged or neutral amino acid. This would change the net charge of the cavity without affecting the cavity volume significantly (Table 4). GroEL-ΔC is the GroEL mutant with the sequence [GGM]₄M deleted.

Table 4: (GGM)₄M → (GGX)₄X

Amino acid residues	Mass (Daltons)	Vanderwaal's volume (Å ³)	Cis-cavity volume (Å ³)	% change as compared to GroEL-WT	Hydrophobicity*
Met (M)WT	131.19	124	161.133	0	1.9
ΔC	-	-	168.161	+4.36	-
Aliphatic side chain					
Gly (G)	57.05	48	163.793	+1.65	-0.4
Ala (A)	71.09	67	163.128	+1.24	1.8
Ile (I)	113.16	124	161.133	0	4.5
Aromatic side chain					
Tyr (Y)	163.18	141	160.538	-0.37	-1.3
Polar neutral side chain					
Ser (S)	87.08	73	162.918	+1.11	-0.8
Acidic side chain					
Asp (D)	115.09	91	162.288	+0.72	-3.5
Basic side chain					
Lys (K)	128.17	137	160.748	-0.24	-3.9

Volume of *cis*-cavity without C-terminal tail is 175.000 Å³

* (Kyte and Doolittle, 1982).

Mutant chaperonins were overexpressed and purified from bacterial strain BL21(DE3).

We verified that purified GroEL mutants formed tetradecamers similar to GroEL-WT (Wild type) by size exclusion chromatography and blue native gels.

Next to verify if chaperonin mutants bind unfolded substrate protein with affinity similar to GroEL-WT, we performed rhodanese aggregation prevention assay. Bovine mitochondrial rhodanese is a monomeric protein (33 kDa) comprised of two domains which catalyzes the formation of thiocyanide from thiosulfate and cyanide. It is highly aggregation prone protein and it has been shown that the folding of rhodanese on import into yeast mitochondria or expression in *E.coli* cytosol is dependent on Hsp60/Hsp10 or GroEL/ES respectively (Ewalt et al., 1997; Rospert et al., 1996). Upon dilution from the denaturant into buffer alone, rhodanese rapidly aggregates which is reflected by increase in the absorbance at 320 nm. When denatured rhodanese is diluted into buffer containing 2-fold excess of GroEL-WT, no aggregation is observed (Figure 5.39). GroEL tail mutants bound unfolded rhodanese with affinity similar to GroEL-WT except for the mutant GroEL-GGD (Figure 5.39a) which had a lower binding affinity for rhodanese as compared to GroEL-WT. No aggregation was observed when 4-fold excess of GroEL-GGD over denatured rhodanese (Rhodanese: GroEL-GGD :: 1 : 4) was used (Figure 5.39b).

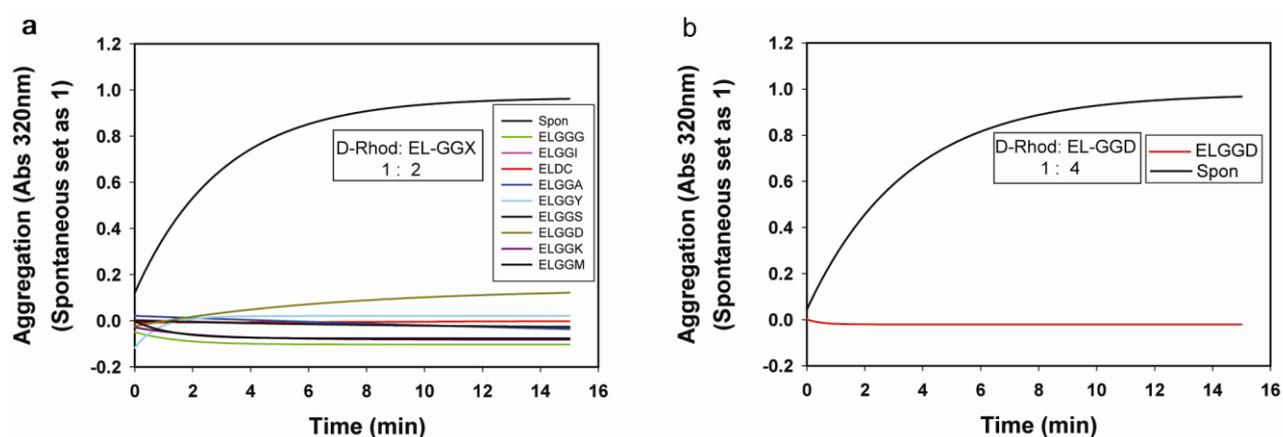


Figure 5.39: Inhibition of Rhodanese Aggregation by GroEL Tail Mutants.

Unfolded Rhodanese (25 μ M) in 6 M GuHCl was diluted 100-fold dilution in buffer A alone (Spontaneous reaction) or in the presence of (a) wild-type GroEL or GroEL tail mutants (0.5 μ M each) as indicated and (b) 1 μ M GroEL-GGD. Aggregation was monitored by following absorbance at 320 nm on spectrophotometer. Spontaneous experiment after 10 min was set to 1.

In the rhodanese refolding experiment described later, increased amount of GroEL-GGD mutant over denatured rhodanese (Rhodanese: GroEL-GGD :: 1 : 4) was used due to its low binding affinity with Rhodanese.

Next, we measured the rates of ATP hydrolysis for GroEL-WT and GroEL tail mutantsr GroEL tail mutants in the absence or presence of GroES using a coupled ATP regenerating enzyme system. Except for the mutant GroEL-GGD which exhibited drastic decrease in ATP hydrolysis rate in presence of GroES, all GroEL tail mutants showed ~50% decrease in ATPase rate in presence of GroES like GroEL-WT (Figure 5.40). However, the absolute rates of ATP hydrolysis varied significantly for the GroEL tail mutants. This suggests that GroEL C-terminal tail mutants affect ATP binding and/or hydrolysis by GroEL directly or indirectly.

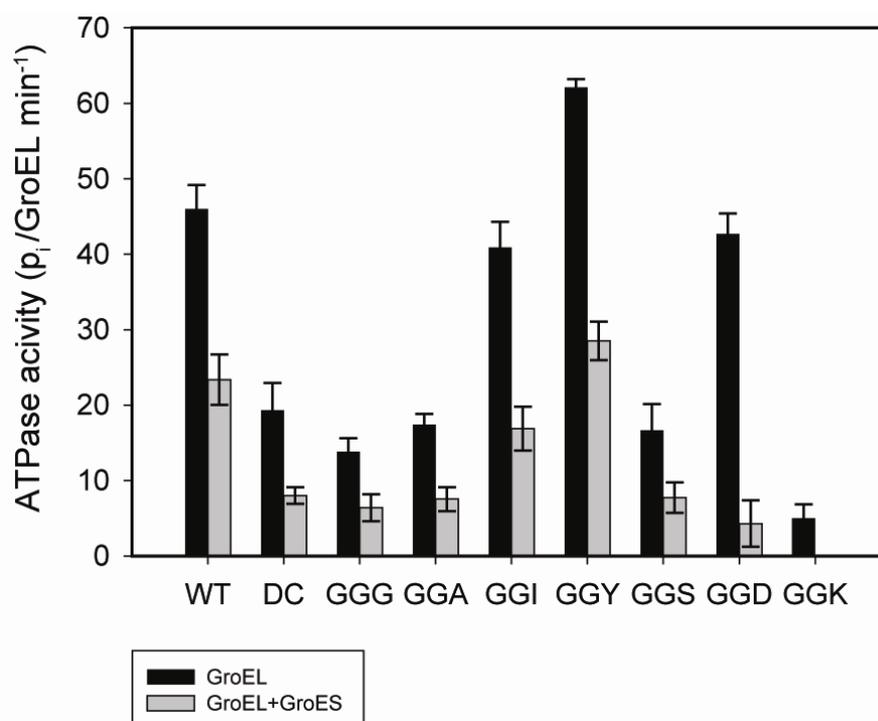


Figure 5.40: ATPase Activity of GroEL Tail Mutants. ATP hydrolysis rate was measured by following decrease in absorbance at 340 nm in presence of either GroEL alone (black) or GroEL and GroES (Grey). ATPase rates of GroEL are indicated as number of ATP hydrolyzed per GroEL tetradecamer per minute. Standard deviation of three independent measurements are shown.

Having established that the GroEL tail mutants can bind substrate protein and GroES, we next investigated the refolding of various substrate proteins. The refolding rate of 33 kDa rhodanese was significantly reduced for the mutants GroEL-GGY and GroEL-GGI, while only the yield of rhodanese folding was drastically affected in presence of GroEL-GGG. Rhodanese refolding rate and yield was not significantly affected in presence of other GroEL tail mutants with different net cavity charge (Figure 5.41).

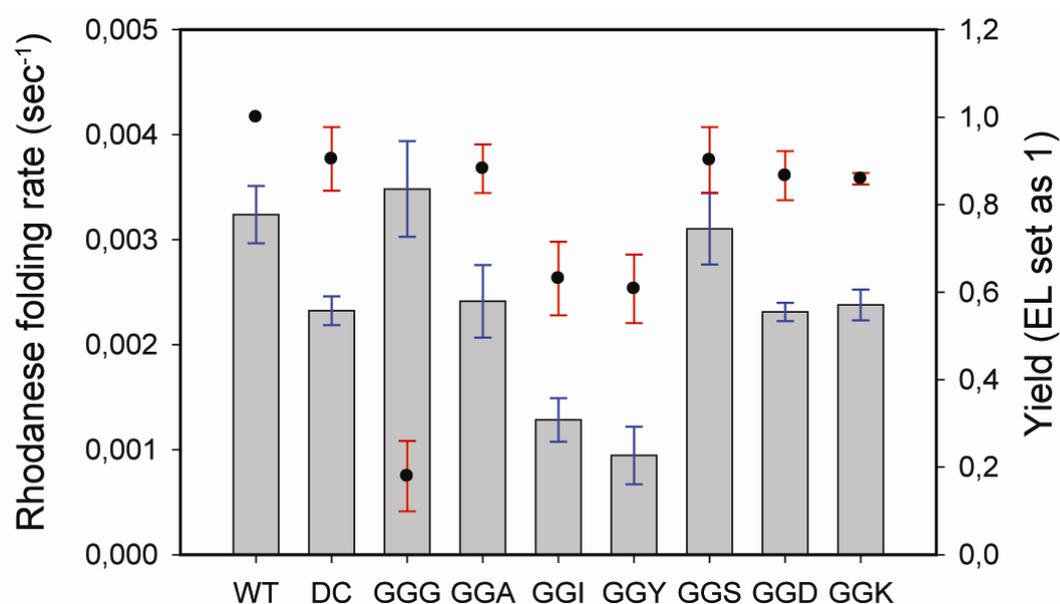


Figure 5.41: Rhodanese Refolding in Presence of GroEL Tail Mutants. Refolding of rhodanese was performed in presence of GroEL tail mutants. 25 μ M rhodanese was denatured in 6 M GuHCl and 100-fold diluted in buffer A with 0.5 μ M GroEL. 1 μ M GroEL-GGD was used in the refolding reaction due to low binding affinity of this mutant with rhodanese. GroES was present in 2-fold excess of the chaperonin. Refolding was initiated by addition of 5 mM ATP at 25°C. Enzyme activity was measured at different time points by taking absorbance at 460 nm. Standard deviation of three independent measurements are shown. Bars indicate rates and circle indicates yield of refolding.

It is possible that the increased hydrophobicity of the cavity as compared to GroEL-WT hinders rhodanese refolding.

Next, we performed refolding experiments using 42 kDa DM-MBP as the substrate protein. Interestingly, the refolding rate of DM-MBP in presence of GroEL-GGI and GroEL-

GGY was similar to that of GroEL-WT, while the refolding of this substrate was significantly decreased in presence of other GroEL tail mutants. These findings suggest that the smaller 33 kDa rhodanese is not affected by changes in the charge properties of the cavity, while the 42 kDa DM-MBP folding is sensitive to the electrostatic property of the cavity. Also the overall hydrophobicity similar to wild-type GroEL cavity is optimal for folding of DM-MBP (Figure 5.42).

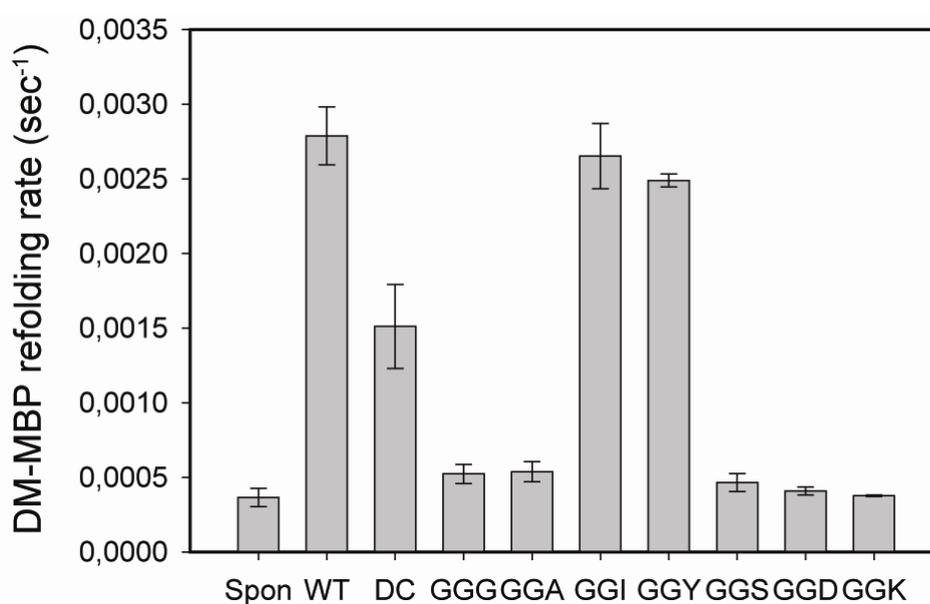


Figure 5.42: DM-MBP Refolding in Presence of GroEL Tail Mutants.

DM-MBP refolding was performed in presence of GroEL tail mutants. 25 μ M DM-MBP was denatured in 6 M GuHCl and 100-fold diluted in buffer A with 0.5 μ M GroEL. GroES was present in 2-fold excess of chaperonin. Refolding was initiated by addition of 5 mM ATP at 25°C and followed by monitoring Trp fluorescence with excitation at 295 nm and emission at 345 nm. Standard deviation of three independent measurements are shown.

Notably, the chaperonin-assisted refolding rates of both rhodanese and DM-MBP show no direct correlation with the ATP hydrolysis rate of the various tail mutants, suggesting that the assisted refolding of these monomeric proteins is independent of the repeated cycles of GroES binding.

The third substrate that we investigated was methylenetetrahydrofolate reductase (MetF) which catalyzes the reduction of 5,10-methylenetetrahydrofolate to 5-methyltetrahydrofolate during methionine biosynthesis (Hatch et al., 1961). It is purified from *E.coli* as tetramer with identical subunits of 33 kDa each (Guenther et al., 1999). It is one of the class III substrates of GroEL (Brinker et al., 2001) which cannot fold without the assistance of GroEL/ES system. Interestingly, when we checked the refolding rate of MetF in presence of GroEL tail mutants, we observed ~5-fold increase in refolding rate of MetF in presence of GroEL-GGK (Figure 5.43).

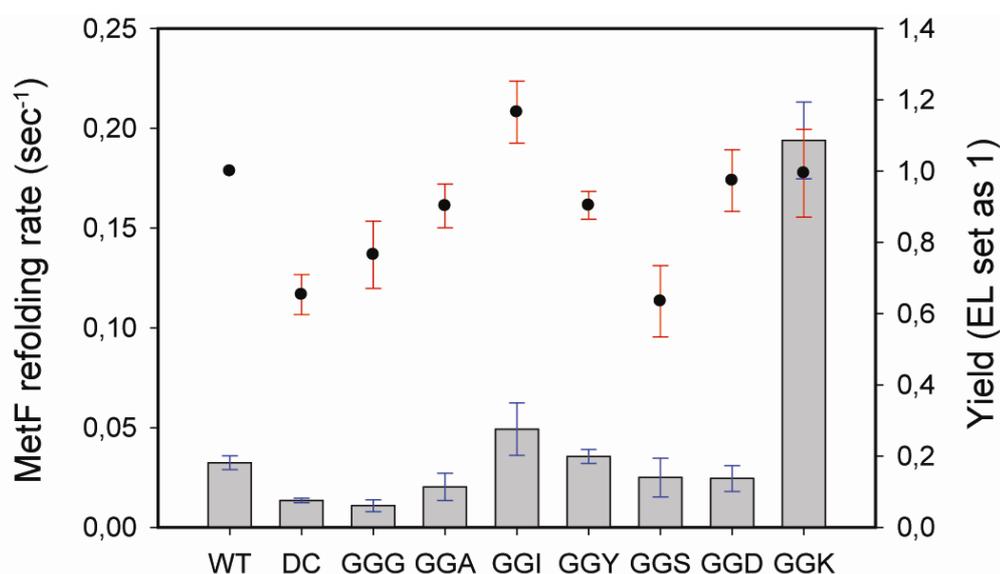


Figure 5.43: MetF Refolding in Presence of GroEL Tail Mutants. MetF refolding was performed in presence of GroEL tail mutants. 25 μ M MetF was denatured in 6 M GuHCl and 100-fold diluted in buffer A with 0.5 μ M GroEL (GroEL-assisted refolding). GroES was present in 2-fold excess of chaperonin. Refolding was initiated by addition of 5 mM ATP at 25°C. Enzyme activity was measured at different time points by taking absorbance at 343 nm. Standard deviation of three independent measurements are shown. Bars indicate rates and circle indicates yield of refolding.

This clearly indicates that GroEL-GGK alters the cavity environment in a way that accelerates the formation of native state of MetF.

These findings strongly suggest that the net charge or the fine balance between the overall hydrophobic and hydrophilic character of the GroEL cavity is of profound significance to promote folding of certain substrate proteins. While the charge effects on specific protein may vary, the overall environment of the GroEL cavity is optimal for folding of diverse set of substrate proteins. Also, the ATPase rate seems to have no direct effect on the folding of substrate proteins tested, consistent with the fact that SR-EL which undergoes just one round of ATP hydrolysis can refold substrate with a similar rate and yield like GroEL-WT (Weissman et al., 1996).

6. Discussion

Protein folding inside the cell is highly dependent on assistance by molecular chaperones. The chaperonin system of *E.coli*, GroE system, in particular has been the subject of intensive study. Previous works have helped to identify the natural substrate spectrum of the GroEL/ES system and revealed insight into the degree of chaperone dependence of the identified substrate proteins (Houry et al., 1999; Kerner et al., 2005). The mechanism of the folding cycle has been elucidated and it has been shown that unfolded proteins get encapsulated inside the GroEL/ES cavity and fold unimpaired by aggregation. However, recent experimental and theoretical studies indicated that the GroEL/ES nano-cage is not a passive cage but the physical environment of the chaperonin cage can alter the folding energy landscape, resulting in accelerated folding of some proteins. Three structural features of the chaperonin cage have been suggested as the major contributors to this capacity: i) geometric confinement exerted on the encapsulated protein inside the limited volume of the cage, ii) the mildly hydrophobic C-terminal tail of GroEL at the bottom of the cage and iii) clusters of negatively charged amino acid residues exposed in the *cis*-cavity wall (Takagi et al., 2003; Tang et al., 2006; Zhou, 2004).

6.1. Passive versus Active Chaperonin Mechanism

In the present study, we have employed a range of biophysical techniques to rule out transient aggregation as the cause of slow spontaneous folding of the model substrate DM-MBP. DM-MBP resembles the authentic (Class III) substrates of GroEL with regard to fold topology and inability to interact productively with the Hsp70 system (DnaK/DnaJ) (Kerner et al., 2005; Tang et al., 2006). We found that refolding rate of DM-MBP was independent of concentration over a wide range (10 nM to 1.5 μ M). FCS, FCCS and light scattering experiments showed the absence of dimers or multimers during DM-MBP refolding. Based on

the above observations, we conclude that DM-MBP refolding *in vitro* is not limited by the formation of reversible aggregates (Figures 5.3-5.6). Moreover, we also demonstrate that in case of an obligate substrate of GroEL, Rubisco, the yield of the spontaneous folding is affected due to the formation of irreversible aggregates but not the rate of folding. Thus, the observed rate acceleration of DM-MBP and Rubisco folding indicates an active mechanism of GroEL in promoting folding which may act synergistically with the capacity of nano-cage to prevent aggregation by substrate encapsulation (Brinker et al., 2001).

6.2. Basis of Slow DM-MBP Folding

The folding of large, topologically complex proteins is often slow. The native states of such proteins are stabilized significantly by long range contacts. It has been shown that formation of such long range contacts are more effectively counteracted by the conformational entropy than that of local contacts (Plaxco et al., 1998). In GuHCl isothermal denaturation experiments, we found a prominent hysteresis in the unfolding/refolding curves. This indicates the presence of a kinetically trapped intermediate during DM-MBP folding. Unfolding/refolding curves show that this intermediate is populated around 0.5 M GuHCl. Also, burst phase analysis clearly indicates that the intermediate populated around 0.5 M GuHCl is the same as the intermediate formed after dilution from the denaturant during refolding reaction (Figure 5.10). We conclude that the basis of slow folding of DM-MBP is the presence of kinetically trapped intermediate that is separated from the native state by a significant free energy barrier. The intermediate in the presence of 0.5 M GuHCl is collapsed but has a broader intra-molecular distance distribution than the native state as determined by spFRET. CD spectroscopy and Trp fluorescence shows that this intermediate has little secondary structure and lacks tertiary interactions. Also, H/D exchange measurements confirmed the dynamic nature of this kinetically trapped intermediate (Figure 5.10-5.17). The lack of ordered structure suggests that the kinetically trapped intermediate is not significantly

stabilized by native or non-native side chain packing, thus imposing a large entropic barrier to folding. Temperature dependence of the refolding rate of DM-MBP also suggests that the barrier imposed to folding, because of the presence of kinetically trapped intermediate, has a large entropic component (Figure 5.31 and Table 3), forming the basis of slow folding of DM-MBP.

6.3. Disulfides-Mediated Folding of DM-MBP

The hypothesis that the refolding rate of DM-MBP is limited by the presence of an entropic barrier between the intermediate and the native state was further tested by introducing disulfide bonds in DM-MBP. The disulfide bonds between the residues that are juxtaposed in the native structure, but far apart along the sequence (Figure 5.19) were introduced to configurationally constrain DM-MBP. The two domains of DM-MBP are discontinuous in sequence, suggesting that they are structurally interdependent. Formation of native contacts in the N-domain, carrying the two mutations V8G and Y283D has been shown to be rate limiting for folding (Chun et al., 1993). Indeed, introducing a long-range disulfide bond in the N- or C-domain accelerated folding by several-fold (Figure 5.21 and figure 5.28). In order to rule out that the observed rate acceleration of DM-MBP folding is due to stabilization of the native state and not due to the destabilization of intermediate state, we compared the unfolding rate of the reduced and oxidized protein. We found no significant difference in the unfolding rate of reduced and oxidized protein indicating that indeed, the rate acceleration is due to the destabilization of the intermediate state and not due to stabilization of the native state (Figure 5.24, figure 5.26 and table 2).

DM-MBP has been shown to collapse within milliseconds, after dilution from the denaturant (Sharma et al., 2008). This provided us with a unique opportunity to determine if disulfide bonds are formed in the intermediate state and if this is sufficient to enhance the

folding rate. Notably, acceleration of folding was observed when disulfide bonds were allowed to form only after the initial collapse reaction (Figure 5.21). This suggested that the respective residues are transiently proximal in the kinetically trapped intermediate and constraining this intermediate is necessary and sufficient to cause the rate acceleration to folding. Constraining already proximal residues should accelerate folding only if substantial flexibility exists around these regions. This is consistent with the view that for proteins with complex topology, reducing chain entropy by stabilizing long-range, native contacts may accelerate the search for the favorable energetic interactions that define transition states (Bartlett and Radford, 2009; Plaxco et al., 1998; Vendruscolo et al., 2003; Wallin et al., 2006). Introducing disulfide bonds rendered the folding rate of DM-MBP temperature dependent, reflecting an increased enthalpic component and a strongly reduced entropic component of the folding barrier (Table 3).

6.4. Chaperonin-Assisted versus Disulfide-Mediated Folding of DM-MBP

Refolding rates of either reduced or oxidized DM-MBP were accelerated to a similar extent inside the chaperonin cage indicating that the accelerating effects of the disulfide bonds and of the chaperonin cage were non-additive (Figure 5.28). Our results argue that the chaperonin cage limits the conformational entropy of the kinetically trapped folding intermediate of DM-MBP in a way resembling the entropic constraints afforded by long-range disulfide bonds (Figure 5.44). Confinement in the chaperonin cage accelerated folding more effectively than the single disulfide bonds.

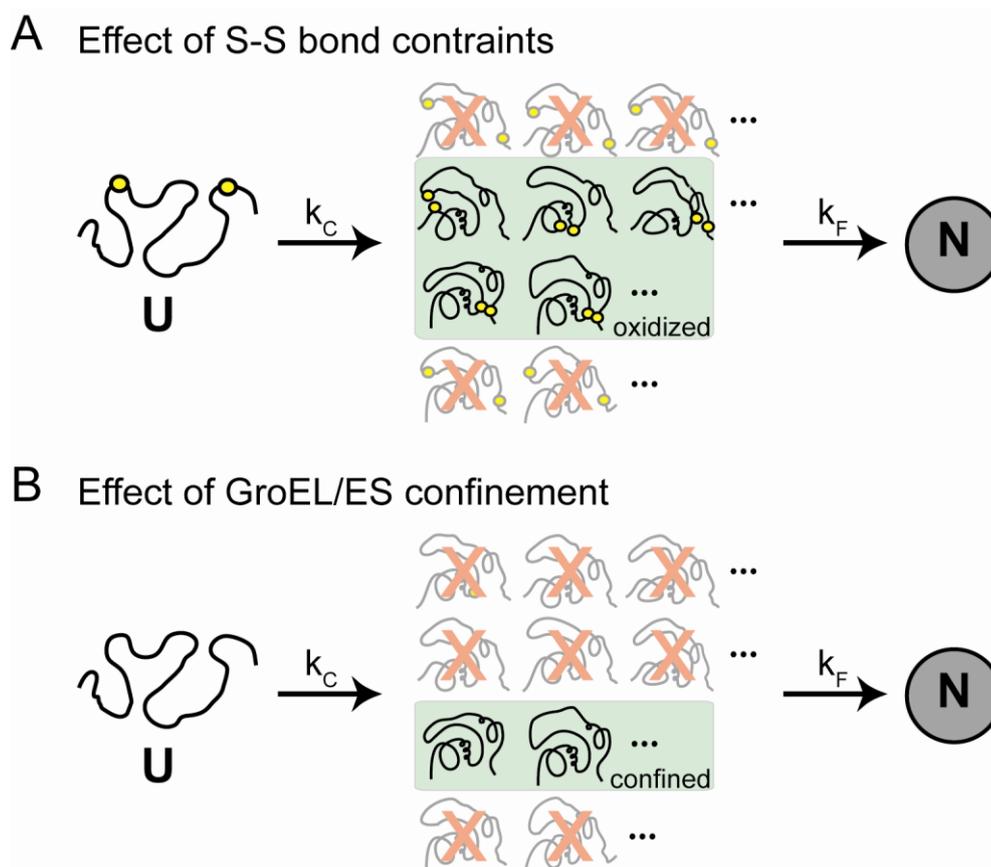


Figure 5.44: Model for Accelerated Folding by Entropic Destabilization of Trapped Folding Intermediate. (a) Schematic representation of refolding of DM-MBP with disulfide bridge-mediated restriction of conformational flexibility and (b) confinement by chaperonin. Upon dilution from denaturant, unfolded 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). The presence of disulfide bonds in the N- and C-domains reduces this barrier by conformationally restricting the ensemble of 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.

This suggests that the chaperonin cage exerts a global confinement effect leading to enhanced rate acceleration as compared to the rate acceleration caused by the local constraint imposed by single disulfide bonds. This lead us to make a mutant of DM-MBP, DM-MBP(4C), where we could constrain both the domains of DM-MBP by forming disulfide bonds simultaneously, in a way resembling the global confinement effect exerted by the chaperonin cage. Indeed, the refolding rate of DM-MBP with disulfide bonds in both the N-

and C-domain was similar to the rate acceleration in the chaperonin cage. Notably, global confinement in the chaperonin cage did not enhance the folding rate obtained by combining the two disulfides. This suggested that the disulfide-mediated constraints and confinement by the chaperonin modulates the thermodynamic parameters of the folding reaction in a similar way. Furthermore, the finding that both the chaperonin-assisted and the disulfide-mediated folding has a positive temperature dependence of the folding rate strongly supports the assumption of a common, underlying principle in reducing the entropic activation energy barrier of the folding reaction.

6.5. Effect of Chaperonin Cavity on the Folding of DM-MBP

The ability of GroEL/ES to accelerate folding is markedly dependent on the negative charge character of the cavity wall. Removal of the net-negative charge rendered the cavity unable to accelerate folding of DM-MBP but did not affect the folding of WT-MBP indicating that altering the charge properties of the cavity do not have an inhibitory effect, rather it lost the ability to accelerate folding of DM-MBP (Tang et al., 2006). This suggested that the mutant with altered charge properties of the cavity provides DM-MBP with a more passive environment. We checked the similarity in the rate limiting step of DM-MBP folding spontaneously and within the mutant chaperonin cavity in various conditions. TMAO, an osmolyte that induces secondary structure formation accelerates refolding rate of DM-MBP folding both spontaneously and within the mutant chaperonin cage in a similar way. Also, GuHCl which is supposed to increase the flexibility of the intermediate state, and thus reduce the refolding rate, decelerated the refolding rate of DM-MBP folding spontaneously and within the mutant chaperonin cavity in a similar way. Notably, the folding of DM-MBP inside the wild-type chaperonin cavity remains unaffected in presence of either TMAO or GuHCl. These findings clearly indicate that the mutant chaperonin cavity with altered charge

properties as compared to the wild type chaperonin cavity with net-negative charge has lost its ability to accelerate folding of DM-MBP, and behaves like a passive cage.

Interestingly, refolding of the oxidized protein with disulfide bonds was accelerated inside the mutant chaperonin cage. This indicates that the mutant chaperonin cage has lost its ability to form native like contacts and hence is unable to accelerate folding of DM-MBP. When these native-like contacts are introduced in form of disulfide bonds, DM-MBP folding rate is accelerated to a similar extent to refolding inside the wild-type chaperonin cage. These findings are consistent with recent theoretical considerations that the charge surface may induce 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 and 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 and Minton, 2006), decreasing the size of the chaperonin cage has been shown to accelerate the folding of smaller GroEL substrates (Tang et al., 2008; Tang et al., 2006), suggesting that the charge effects from the cavity wall and the geometric confinement act in concert to smooth the folding energy landscape (Figure 5.45).

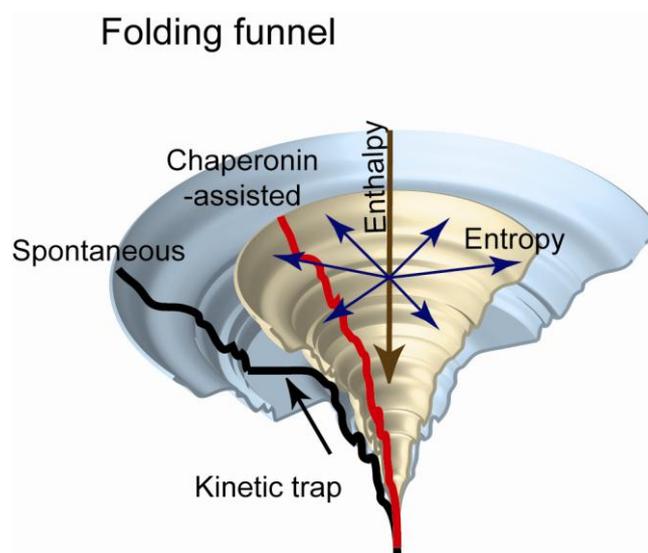


Figure 5.45: The folding Landscape of DM-MBP. The hypothetical folding funnel of DM-MBP is shown in light blue with the z-axis denoting enthalpy and the x-y plane denoting entropy of the folding intermediates. The large flat area in the folding funnel denotes the kinetically trapped species that is separated from the native state through an entropic barrier. This is modified inside the chaperonin cage (light brown funnel) resulting in a more downhill funnel without the presence of prominent isoenthalpic regions representing entropically stabilized intermediate state.

6.6. Effect of the Mildly Hydrophobic C-Terminal GGM Repeat on Folding of GroEL Substrates

Many bacterial GroEL sequences have a highly conserved and mildly hydrophobic motif located at the C-terminus. This motif, although variable in length and exact amino acid sequence, consists mainly of glycine and methionine residues. The GGM repeats are also observed in other ATPase protein families, like Hsp 70 (Brocchieri and Karlin, 1998) or RecA (Brendel et al., 1997). The significance of these repetitive elements is unknown. In other proteins like keratins, glycine rich motifs are organized into loops in which the hydrophobic residues (methionine) and the flexible structure behaves like a molecular spring. However, the role of the C-terminal GGM repeats in GroEL mediated folding is unknown. There have been many studies, including molecular dynamics simulations which suggest a role of a moderately hydrophobic cavity in causing significant rate acceleration of folding (Jewett et al., 2004). The

mildly hydrophobic GGM repeat may fulfill such a role. Our results indicate that the GGM repeat is optimal for folding of a diverse array of GroEL substrates explaining the conserved nature of these residues among different species. Notably, the GroEL cavity may be optimized further for folding of a specific substrate protein like MetF which has significantly faster folding rate inside the GroEL-GGK cavity as compared to the wild-type GroEL (GGM repeat). This, however, would be at the expense of folding of other substrate proteins

6.7. Effect of ATPase rate and cycling on folding of different GroEL substrates

Substrate cycling has been proposed to promote folding by iteratively unfolding kinetically trapped intermediates, stabilized by non-native interactions, allowing re-partitioning to a productive folding pathway upon release (Shtilerman, 1999; Thirumalai and Lorimer, 2001). Indeed, ATP dependent apical domain movements can cause local structure expansion (Lin et al., 2008; Sharma et al., 2008), but in case of DM-MBP and RuBisco, such ‘forced unfolding’ was dispensable for folding acceleration (Brinker et al., 2001; Sharma et al., 2008). This may be readily explained by our finding that the kinetically trapped folding intermediate of DM-MBP is highly disordered and thus unlikely to contain strong non-native contacts. Consequently, further unfolding would not circumvent the formation of the folding trap. Furthermore, different tail mutants with mutations in the C-terminal repeat were found to hydrolyze ATP with different rates but there was no trend correlating the ATP hydrolysis rates with refolding rates of the various substrate proteins like DM-MBP, Rhodanese or MetF. Interestingly, GroEL-GGK which showed higher rate acceleration for MetF refolding as compared to the GroEL-WT has very slow ATPase rate. These results argue that ATPase rate and iterative annealing of the unfolded protein might not play an important role in rate acceleration caused by GroEL, at least for the substrate proteins tested. This is also consistent

with the data that the SR-EL mutant with one cycle of ATP hydrolysis is able to accelerate the refolding to a similar extent as GroEL (Weissman et al., 1996).

Taken together with the entropic destabilization of the intermediate inside the chaperonin cage, our data support a model in which the bimodal character (hydrophilic *cis*-cavity and hydrophobic C-terminal tail) of the cavity facilitates the re-configuration of the folding intermediates within the confined cage. Smoothing of folding energy landscape is achieved in single round of encapsulation by sequestering the protein a confined space with an optimized hydrophobic property.

6.8. Biological Relevance of Chaperonin-Mediated Folding

The GroEL/ES system is essential under all growth conditions, and ~85 *E. coli* cytoplasmic proteins are predicted to be strictly dependent on GroEL for folding, including 13 essential proteins (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 stable folding intermediates (Kerner et al., 2005). Interestingly, while most of these proteins are aggregation-prone, several of them appear to form soluble intermediates upon translation *in vitro* in the absence of GroEL (Niwa et al., 2009), suggesting that their folding may depend on entropic confinement by chaperonin. In view of the fact that cells contain multiple, partially redundant chaperone systems for aggregation prevention, this function would explain the uniquely essential role of the chaperonin cages in protein folding. On the other hand, the conspicuous absence of chaperonins from oxidizing cellular compartments could correlate with the capacity of disulfide bond formation to lower entropic barriers during folding.

In addition to assisting the folding of a restricted set of wild-type proteins, GroEL/ES is thought to buffer mutations that render otherwise chaperonin-independent proteins folding-

defective. This function may greatly expand the number of foldable protein variants and may have facilitated structural protein evolution (Tokuriki and Tawfik, 2009). As illustrated by the example of DM-MBP, such mutant proteins may be similar to their wild-type counterparts in terms of thermodynamic stability of the native state but may be kinetically folding defective and hence dependent on the chaperonin to reach their active conformation at a biologically relevant time scale. Significant structural deviations may be tolerated when additional, specialized forms of GroEL are expressed to allow adaptation of an organism to specific growth conditions. Interestingly, many pathogenic bacteria and bacteria residing in extreme conditions like high temperature or high salt conditions express two forms of GroEL. For example, mycobacteria express two forms of GroEL, of which GroEL 1 lacks the C-terminal repeat and instead has an 18 amino acid histidine-rich sequence which appears to be critical for folding of the proteins required for bacterial biofilm formation (Ojha et al., 2005). Overall the surface properties of the chaperonin cavity represent an evolutionary compromise that helps the bacterial cell to produce functional proteins fast enough to survive in the competitive microbial world.

7. References

- Abkevich, V.I., Gutin, A.M., and Shakhnovich, E.I. (1994). Free energy landscape for protein folding kinetics: Intermediates, traps, and multiple pathways in theory and lattice model simulations. *The Journal of Chemical Physics* *101*, 6052.
- Abkevich, V.I., and Shakhnovich, E.I. (2000). What can disulfide bonds tell us about protein energetics, function and folding: simulations and bioinformatics analysis. *J Mol Biol* *300*, 975-985.
- Adesnik, M., and Levinthal, C. (1969). Synthesis and maturation of ribosomal RNA in *Escherichia coli*. *J Mol Biol* *46*, 281-303.
- Agashe, V.R., Guha, S., Chang, H.C., Genevoux, P., Hayer-Hartl, M., Stemp, M., Georgopoulos, C., Hartl, F.U., and Barral, J.M. (2004). Function of trigger factor and DnaK in multidomain protein folding: increase in yield at the expense of folding speed. *Cell* *117*, 199-209.
- Aharoni, a., and Horovitz, a. (1996). Inter-ring communication is disrupted in the GroEL mutant Arg13 --> Gly; Ala126 --> Val with known crystal structure. *Journal of molecular biology* *258*, 732-735.
- Albanèse, V., Yam, A.Y.-W., Baughman, J., Parnot, C., and Frydman, J. (2006). Systems analyses reveal two chaperone networks with distinct functions in eukaryotic cells. *Cell* *124*, 75-88.
- Anfinsen, C.B. (1973). Principles that govern the folding of protein chains. *Science* *181*, 223-230.
- Apetri, A.C., and Horwich, A.L. (2008). Chaperonin chamber accelerates protein folding through passive action of preventing aggregation. *Proc Natl Acad Sci U S A* *105*, 17351-17355.
- Badcoe, I.G., Smith, C.J., Wood, S., Halsall, D.J., Holbrook, J.J., Lund, P., and Clarke, A.R. (1991). Binding of a chaperonin to the folding intermediates of lactate dehydrogenase. *Biochemistry* *30*, 9195-9200.
- Bagshaw, C.R., and Cherny, D. (2006). Blinking fluorophores: what do they tell us about protein dynamics? *Biochem Soc Trans* *34*, 979-982.
- Baldwin, R.L. (1996). Commentary Why is protein folding so fast ? *93*, 2627-2628.
- Baldwin, R.L. (1999). Protein folding from 1961 to 1982. *Nature structural biology* *6*, 814-817.
- Baldwin, R.L., and Rose, G.D. (1999). Is protein folding hierarchic? II. Folding intermediates and transition states. *Trends in biochemical sciences* *24*, 77-83.
- Ban, N. (2000). The Complete Atomic Structure of the Large Ribosomal Subunit at 2.4 Å Resolution. *Science* *289*, 905-920.
- Baram, D., Pyetan, E., Sittner, A., Auerbach-Nevo, T., Bashan, A., and Yonath, A. (2005). Structure of trigger factor binding domain in biologically homologous complex with eubacterial ribosome reveals its chaperone action. *Proceedings of the National Academy of Sciences of the United States of America* *102*, 12017-12022.
- Barral, J.M., Broadley, S.A., Schaffar, G., and Hartl, F.U. (2004). Roles of molecular chaperones in protein misfolding diseases. *Semin Cell Dev Biol* *15*, 17-29.
- Bartlett, A.I., and Radford, S.E. (2009). An expanding arsenal of experimental methods yields an explosion of insights into protein folding mechanisms. *Nat Struct Mol Biol* *16*, 582-588.
- Betancourt, M.R., and Thirumalai, D. (1999). Exploring the kinetic requirements for enhancement of protein folding rates in the GroEL cavity. *J Mol Biol* *287*, 627-644.
- Bochkareva, E.S., Lissin, N.M., Flynn, G.C., Rothman, J.E., and Girshovich, A.S. (1992). Positive cooperativity in the functioning of molecular chaperone GroEL. *J Biol Chem* *267*, 6796-6800.

- Boisvert, D.C., Wang, J., Otwinowski, Z., Horwich, A.L., and Sigler, P.B. (1996). The 2.4 Å crystal structure of the bacterial chaperonin GroEL complexed with ATP gamma S. *Nature structural biology* 3, 170-177.
- Braig, K., Otwinowski, Z., Hegde, R., Boisvert, D.C., Joachimiak, A., Horwich, A.L., and Sigler, P.B. (1994). The crystal structure of the bacterial chaperonin GroEL at 2.8 Å. *Nature* 371, 578-586.
- Braig, K., Simon, M., Furuya, F., Hainfeld, J.F., and Horwich, A.L. (1993). A polypeptide bound by the chaperonin groEL is localized within a central cavity. *Proc Natl Acad Sci U S A* 90, 3978-3982.
- Brendel, V., Brocchieri, L., Sandler, S.J., Clark, A.J., and Karlin, S. (1997). Evolutionary comparisons of RecA-like proteins across all major kingdoms of living organisms. *J Mol Evol* 44, 528-541.
- Brinker, A., Pfeifer, G., Kerner, M.J., Naylor, D.J., Hartl, F.U., and Hayer-Hartl, M. (2001). Dual function of protein confinement in chaperonin-assisted protein folding. *Cell* 107, 223-233.
- Brocchieri, L., and Karlin, S. (1998). A symmetric-iterated multiple alignment of protein sequences. *J Mol Biol* 276, 249-264.
- Brockwell, D.J., Smith, D.A., and Radford, S.E. (2000). Protein folding mechanisms: new methods and emerging ideas. *Curr Opin Struct Biol* 10, 16-25.
- Bryngelson, J.D., Onuchic, J.N., Socci, N.D., and Wolynes, P.G. (1995). Funnels, pathways, and the energy landscape of protein folding: a synthesis. *Proteins* 21, 167-195.
- Bryngelson, J.D., and Wolynes, P.G. (1987). and the statistical mechanics of protein folding. *Biophysics* 84, 7524-7528.
- Bryngelson, J.D., and Wolynes, P.G. (1989). Intermediates and barrier crossing in a random energy model (with applications to protein folding). *The Journal of Physical Chemistry* 93, 6902-6915.
- Buckle, a.M., Zahn, R., and Fersht, a.R. (1997). A structural model for GroEL-polypeptide recognition. *Proceedings of the National Academy of Sciences of the United States of America* 94, 3571-3575.
- Bukau, B., Weissman, J., and Horwich, A. (2006). Molecular chaperones and protein quality control. *Cell* 125, 443-451.
- Burston, S.G., Ranson, N.A., and Clarke, A.R. (1995). The origins and consequences of asymmetry in the chaperonin reaction cycle. *J Mol Biol* 249, 138-152.
- Byrne, M.P., Manuel, R.L., Lowe, L.G., and Stites, W.E. (1995). Energetic contribution of side chain hydrogen bonding to the stability of staphylococcal nuclease. *Biochemistry* 34, 13949-13960.
- Camacho, C.J., and Thirumalai, D. (1993). Kinetics and thermodynamics of folding in model proteins. *Proceedings of the National Academy of Sciences of the United States of America* 90, 6369-6372.
- Chan, H.S., and Dill, K.a. (1994). Transition states and folding dynamics of proteins and heteropolymers. *The Journal of Chemical Physics* 100, 9238.
- Chan, H.S., and Dill, K.a. (1996). Comparing folding codes for proteins and polymers. *Proteins* 24, 335-344.
- Chandrasekhar, G.N., Tilly, K., Woolford, C., Hendrix, R., and Georgopoulos, C. (1986). Purification and properties of the groES morphogenetic protein of *Escherichia coli*. *The Journal of biological chemistry* 261, 12414-12419.
- Chang, H.C., Kaiser, C.M., Hartl, F.U., and Barral, J.M. (2005). De novo folding of GFP fusion proteins: high efficiency in eukaryotes but not in bacteria. *J Mol Biol* 353, 397-409.
- Chang, H.C., Tang, Y.C., Hayer-Hartl, M., and Hartl, F.U. (2007). SnapShot: molecular chaperones, Part I. *Cell* 128, 212.

- Chaudhry, C., Farr, G.W., Todd, M.J., Rye, H.S., Brunger, A.T., Adams, P.D., Horwich, A.L., and Sigler, P.B. (2003). Role of the gamma-phosphate of ATP in triggering protein folding by GroEL-GroES: function, structure and energetics. *EMBO J* 22, 4877-4887.
- Chaudhuri, T.K., Farr, G.W., Fenton, W.A., Rospert, S., and Horwich, A.L. (2001). GroEL/GroES-mediated folding of a protein too large to be encapsulated. *Cell* 107, 235-246.
- Chen, J., Walter, S., Horwich, A.L., and Smith, D.L. (2001). Folding of malate dehydrogenase inside the GroEL-GroES cavity. *Nat Struct Biol* 8, 721-728.
- Chen, L., and Sigler, P.B. (1999). The crystal structure of a GroEL/peptide complex: plasticity as a basis for substrate diversity. *Cell* 99, 757-768.
- Chen, S., Roseman, A.M., Hunter, A.S., Wood, S.P., Burston, S.G., Ranson, N.A., Clarke, A.R., and Saibil, H.R. (1994). Location of a folding protein and shape changes in GroEL-GroES complexes imaged by cryo-electron microscopy. *Nature* 371, 261-264.
- Chen, X., and Matthews, C.R. (1994). Thermodynamic properties of the transition state for the rate-limiting step in the folding of the alpha subunit of tryptophan synthase. *Biochemistry* 33, 6356-6362.
- Chun, S.Y., Strobel, S., Bassford, P., Jr., and Randall, L.L. (1993). Folding of maltose-binding protein. Evidence for the identity of the rate-determining step in vivo and in vitro. *J Biol Chem* 268, 20855-20862.
- Cliff, M.J., Kad, N.M., Hay, N., Lund, P.A., Webb, M.R., Burston, S.G., and Clarke, A.R. (1999). A kinetic analysis of the nucleotide-induced allosteric transitions of GroEL. *J Mol Biol* 293, 667-684.
- Cordes, M.H., Davidson, a.R., and Sauer, R.T. (1996). Sequence space, folding and protein design. *Current opinion in structural biology* 6, 3-10.
- Creighton, T.E. (1974). Intermediates in the refolding of reduced pancreatic trypsin inhibitor. *J Mol Biol* 87, 579-602.
- Cuéllar, J., Martín-Benito, J., Scheres, S.H.W., Sousa, R., Moro, F., López-Viñas, E., Gómez-Puertas, P., Muga, A., Carrascosa, J.L., and Valpuesta, J.M. (2008). The structure of CCT-Hsc70 NBD suggests a mechanism for Hsp70 delivery of substrates to the chaperonin. *Nature structural & molecular biology* 15, 858-864.
- Deechongkit, S., Dawson, P.E., and Kelly, J.W. (2004). Toward assessing the position-dependent contributions of backbone hydrogen bonding to beta-sheet folding thermodynamics employing amide-to-ester perturbations. *Journal of the American Chemical Society* 126, 16762-16771.
- Dill, K.A. (1990). *Perspectives in Biochemistry*. Society 29.
- Dill, K.a. (1999). Polymer principles and protein folding. *Protein science : a publication of the Protein Society* 8, 1166-1180.
- Dill, K.a., Bromberg, S., Yue, K., Fiebig, K.M., Yee, D.P., Thomas, P.D., and Chan, H.S. (1995). Principles of protein folding--a perspective from simple exact models. *Protein science : a publication of the Protein Society* 4, 561-602.
- Dill, K.A., and Chan, H.S. (1997). From Levinthal to pathways to funnels. *Nat Struct Biol* 4, 10-19.
- Dinner, a.R., Sali, a., Smith, L.J., Dobson, C.M., and Karplus, M. (2000). Understanding protein folding via free-energy surfaces from theory and experiment. *Trends in biochemical sciences* 25, 331-339.
- Dobson, C.M. (2003). Protein folding and misfolding. *Nature* 426, 884-890.
- Dobson, C.M., Evans, P.a., and Radford, S.E. (1994). Understanding how proteins fold: the lysozyme story so far. *Trends in biochemical sciences* 19, 31-37.
- Doolittle, R.F. (1995). The origins and evolution of eukaryotic proteins. *Philos Trans R Soc Lond B Biol Sci* 349, 235-240.

- Ekman, D., Bjorklund, A.K., Frey-Skott, J., and Elofsson, A. (2005). Multi-domain proteins in the three kingdoms of life: orphan domains and other unassigned regions. *J Mol Biol* 348, 231-243.
- Elad, N., Farr, G.W., Clare, D.K., Orlova, E.V., Horwich, A.L., and Saibil, H.R. (2007). Topologies of a substrate protein bound to the chaperonin GroEL. *Mol Cell* 26, 415-426.
- Ellis, R.J. (1994a). Chaperoning nascent proteins. *Nature* 370, 96-97.
- Ellis, R.J. (1994b). Molecular chaperones. Opening and closing the Anfinsen cage. *Curr Biol* 4, 633-635.
- Ellis, R.J. (2001). Macromolecular crowding: an important but neglected aspect of the intracellular environment. *Current opinion in structural biology* 11, 114-119.
- Ellis, R.J., and Hartl, F.U. (1996). Protein folding in the cell: competing models of chaperonin function. *FASEB J* 10, 20-26.
- England, J.L., Lucent, D., and Pande, V.S. (2008). A role for confined water in chaperonin function. *J Am Chem Soc* 130, 11838-11839.
- England, J.L., and Pande, V.S. (2008). Potential for modulation of the hydrophobic effect inside chaperonins. *Biophys J* 95, 3391-3399.
- Ewalt, K.L., Hendrick, J.P., Houry, W.A., and Hartl, F.U. (1997). In vivo observation of polypeptide flux through the bacterial chaperonin system. *Cell* 90, 491-500.
- Fayet, O., Ziegelhoffer, T., and 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, 1379-1385.
- Fenton, W.A., Kashi, Y., Furtak, K., and Horwich, A.L. (1994). Residues in chaperonin GroEL required for polypeptide binding and release. *Nature* 371, 614-619.
- Fenton, W.A., Weissman, J.S., and Horwich, A.L. (1996). Putting a lid on protein folding: structure and function of the co-chaperonin, GroES. *Chem Biol* 3, 157-161.
- Ferbitz, L., Maier, T., Patzelt, H., Bukau, B., Deuerling, E., and Ban, N. (2004). Trigger factor in complex with the ribosome forms a molecular cradle for nascent proteins. *Nature* 431, 590-596.
- Fersht, a.R. (1995). Characterizing transition states in protein folding: an essential step in the puzzle. *Current opinion in structural biology* 5, 79-84.
- Fersht, A.R., Shi, J.P., Knill-Jones, J., Lowe, D.M., Wilkinson, A.J., Blow, D.M., Brick, P., Carter, P., Waye, M.M., and Winter, G. (1985). Hydrogen bonding and biological specificity analysed by protein engineering. *Nature* 314, 235-238.
- Fink, a.L. (1999). Chaperone-mediated protein folding. *Physiological reviews* 79, 425-449.
- Freedman, R.B., Hirst, T.R., and Tuite, M.F. (1994). Protein disulphide isomerase: building bridges in protein folding. *Trends in biochemical sciences* 19, 331-336.
- Genevaux, P., Keppel, F., Schwager, F., Langendijk-Genevaux, P.S., Hartl, F.U., and Georgopoulos, C. (2004). In vivo analysis of the overlapping functions of DnaK and trigger factor. *EMBO Rep* 5, 195-200.
- Georgopoulos, C.P., Hendrix, R.W., Casjens, S.R., and Kaiser, a.D. (1973). Host participation in bacteriophage lambda head assembly. *Journal of molecular biology* 76, 45-60.
- Goloubinoff, P., Christeller, J.T., Gatenby, A.A., and 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, 884-889.
- Gray, T.E., and Fersht, A.R. (1991). Cooperativity in ATP hydrolysis by GroEL is increased by GroES. *FEBS Lett* 292, 254-258.

- Guenther, B.D., Sheppard, C.A., Tran, P., Rozen, R., Matthews, R.G., and Ludwig, M.L. (1999). The structure and properties of methylenetetrahydrofolate reductase from *Escherichia coli* suggest how folate ameliorates human hyperhomocysteinemia. *Nature structural biology* 6, 359-365.
- Harrison, C.J. (1997). Crystal Structure of the Nucleotide Exchange Factor GrpE Bound to the ATPase Domain of the Molecular Chaperone DnaK. *Science* 276, 431-435.
- Hartl, F.U. (1996). Molecular chaperones in cellular protein folding. *Nature* 381, 571-579.
- Hartl, F.U., and Hayer-Hartl, M. (2009). Converging concepts of protein folding in vitro and in vivo. *Nat Struct Mol Biol* 16, 574-581.
- Hatch, F.T., Larrabee, A.R., Cathou, R.E., and Buchanan, J.M. (1961). Enzymatic synthesis of the methyl group of methionine. I. Identification of the enzymes and cofactors involved in the system isolated from *Escherichia coli*. *J Biol Chem* 236, 1095-1101.
- Hayer-Hartl, M.K., Ewbank, J.J., Creighton, T.E., and Hartl, F.U. (1994). Conformational specificity of the chaperonin GroEL for the compact folding intermediates of alpha-lactalbumin. *EMBO J* 13, 3192-3202.
- Hayer-Hartl, M.K., Weber, F., and Hartl, F.U. (1996). Mechanism of chaperonin action: GroES binding and release can drive GroEL-mediated protein folding in the absence of ATP hydrolysis. *EMBO J* 15, 6111-6121.
- Hecht, M.H., Das, A., Go, A., Bradley, L.H., and Wei, Y. (2004). De novo proteins from designed combinatorial libraries. *PoLAR*, 1711-1723.
- Hesterkamp, T., and Bukau, B. (1996). Identification of the prolyl isomerase domain of *Escherichia coli* trigger factor. *FEBS letters* 385, 67-71.
- Hesterkamp, T., Hauser, S., Lütcke, H., and Bukau, B. (1996). *Escherichia coli* trigger factor is a prolyl isomerase that associates with nascent polypeptide chains. *Proceedings of the National Academy of Sciences of the United States of America* 93, 4437-4441.
- Horovitz, A., Bochkareva, E.S., Yifrach, O., and Girshovich, A.S. (1994). Prediction of an inter-residue interaction in the chaperonin GroEL from multiple sequence alignment is confirmed by double-mutant cycle analysis. *J Mol Biol* 238, 133-138.
- Horovitz, A., Fridmann, Y., Kafri, G., and Yifrach, O. (2001). Review: allostery in chaperonins. *J Struct Biol* 135, 104-114.
- Houry, W.A., Frishman, D., Eckerskorn, C., Lottspeich, F., and Hartl, F.U. (1999). Identification of in vivo substrates of the chaperonin GroEL. *Nature* 402, 147-154.
- Hunt, J.F., Weaver, A.J., Landry, S.J., Gierasch, L., and Deisenhofer, J. (1996). The crystal structure of the GroES co-chaperonin at 2.8 Å resolution. *Nature* 379, 37-45.
- Jackson, G.S., Staniforth, R.A., Halsall, D.J., Atkinson, T., Holbrook, J.J., Clarke, A.R., and Burston, S.G. (1993). Binding and hydrolysis of nucleotides in the chaperonin catalytic cycle: implications for the mechanism of assisted protein folding. *Biochemistry* 32, 2554-2563.
- Jamin, M., and Baldwin, R.L. (1996). Refolding and unfolding kinetics of the equilibrium folding intermediate of apomyoglobin. *Nat Struct Biol* 3, 613-618.
- Jewett, A.I., Baumketner, A., and Shea, J.E. (2004). Accelerated folding in the weak hydrophobic environment of a chaperonin cavity: creation of an alternate fast folding pathway. *Proc Natl Acad Sci U S A* 101, 13192-13197.
- Jiang, J., Prasad, K., Lafer, E.M., and Sousa, R. (2005). Structural basis of interdomain communication in the Hsc70 chaperone. *Molecular cell* 20, 513-524.
- Kafri, G., Willison, K.R., and Horovitz, A. (2001). Nested allosteric interactions in the cytoplasmic chaperonin containing TCP-1. *Protein Sci* 10, 445-449.

- Kaiser, C.M., Chang, H.-C., Agashe, V.R., Lakshmipathy, S.K., Etchells, S.A., Hayer-Hartl, M., Hartl, F.U., and Barral, J.M. (2006a). Real-time observation of trigger factor function on translating ribosomes. *Nature* *444*, 455-460.
- Kaiser, C.M., Chang, H.C., Agashe, V.R., Lakshmipathy, S.K., Etchells, S.A., Hayer-Hartl, M., Hartl, F.U., and Barral, J.M. (2006b). Real-time observation of trigger factor function on translating ribosomes. *Nature* *444*, 455-460.
- Kalinin, S., and Johansson, L.B. (2004). Utility and considerations of donor-donor energy migration as a fluorescence method for exploring protein structure-function. *J Fluoresc* *14*, 681-691.
- Kanno, R., Koike-Takeshita, A., Yokoyama, K., Taguchi, H., and Mitsuoka, K. (2009). Cryo-EM structure of the native GroEL-GroES complex from *thermus thermophilus* encapsulating substrate inside the cavity. *Structure* *17*, 287-293.
- Kerner, M.J., Naylor, D.J., Ishihama, Y., Maier, T., Chang, H.C., Stines, A.P., Georgopoulos, C., Frishman, D., Hayer-Hartl, M., Mann, M., *et al.* (2005). Proteome-wide analysis of chaperonin-dependent protein folding in *Escherichia coli*. *Cell* *122*, 209-220.
- Kim, P.S., and Baldwin, R.L. (1990). Intermediates in the folding reactions of small proteins. *Annual review of biochemistry* *59*, 631-660.
- Kim, S., Willison, K.R., and Horwich, A.L. (1994). Cytosolic chaperonin subunits have a conserved ATPase domain but diverged polypeptide-binding domains. *Trends Biochem Sci* *19*, 543-548.
- Klimov, D., and Thirumalai, D. (1996). Criterion that determines the foldability of proteins. *Physical review letters* *76*, 4070-4073.
- Kramer, G., Boehringer, D., Ban, N., and Bukau, B. (2009). The ribosome as a platform for co-translational processing, folding and targeting of newly synthesized proteins. *Nature structural & molecular biology* *16*, 589-597.
- Kramer, G., Rutkowska, A., Wegrzyn, R.D., Patzelt, H., Kurz, T.A., Merz, F., Rauch, T., Deuerling, E., and Bukau, B. (2004). Functional Dissection of *Escherichia coli* Trigger Factor : Unraveling the Function of Individual Domains. *Society* *186*, 3777-3784.
- Kummerfeld, S.K., and Teichmann, S.A. (2005). Relative rates of gene fusion and fission in multi-domain proteins. *Trends Genet* *21*, 25-30.
- Kyte, J., and Doolittle, R.F. (1982). A simple method for displaying the hydrophobic character of a protein. *J Mol Biol* *157*, 105-132.
- Lakshmipathy, S.K., Tomic, S., Kaiser, C.M., Chang, H.C., Genevaux, P., Georgopoulos, C., Barral, J.M., Johnson, A.E., Hartl, F.U., and Etchells, S.A. (2007). Identification of nascent chain interaction sites on trigger factor. *J Biol Chem* *282*, 12186-12193.
- Landry, S.J., and Gierasch, L.M. (1991). Accelerated Publications. *Proteins* *30*, 7359-7362.
- Landry, S.J., Jordan, R., McMacken, R., and Gierasch, L.M. (1992). Different conformations for the same polypeptide bound to chaperones DnaK and GroEL. *Nature* *355*, 455-457.
- Landry, S.J., Zeilstra-Ryalls, J., Fayet, O., Georgopoulos, C., and Gierasch, L.M. (1993). Characterization of a functionally important mobile domain of GroES. *Nature* *364*, 255-258.
- Langer, T., Pfeifer, G., Martin, J., Baumeister, W., and Hartl, F.U. (1992). Chaperonin-mediated protein folding: GroES binds to one end of the GroEL cylinder, which accommodates the protein substrate within its central cavity. *EMBO J* *11*, 4757-4765.
- Lee, N.K., Kapanidis, A.N., Wang, Y., Michalet, X., Mukhopadhyay, J., Ebright, R.H., and Weiss, S. (2005). Accurate FRET measurements within single diffusing biomolecules using alternating-laser excitation. *Biophys J* *88*, 2939-2953.

- Leopold, P.E., Montal, M., and Onuchic, J.N. (1992). Protein folding funnels: a kinetic approach to the sequence-structure relationship. *Proceedings of the National Academy of Sciences of the United States of America* 89, 8721-8725.
- Lin, Z., Madan, D., and Rye, H.S. (2008). GroEL stimulates protein folding through forced unfolding. *Nat Struct Mol Biol* 15, 303-311.
- Lin, Z., and Rye, H.S. (2004). Expansion and compression of a protein folding intermediate by GroEL. *Mol Cell* 16, 23-34.
- Lu, J., and Deutsch, C. (2005). Folding zones inside the ribosomal exit tunnel. *Nature structural & molecular biology* 12, 1123-1129.
- Lucent, D., Vishal, V., and Pande, V.S. (2007). Protein folding under confinement: a role for solvent. *Proc Natl Acad Sci U S A* 104, 10430-10434.
- Maisnier-Patin, S., Roth, J.R., Fredriksson, A., Nyström, T., Berg, O.G., and Andersson, D.I. (2005). Genomic buffering mitigates the effects of deleterious mutations in bacteria. *Nature genetics* 37, 1376-1379.
- Mamathambika, B.S., and Bardwell, J.C. (2008). Disulfide-linked protein folding pathways. *Annu Rev Cell Dev Biol* 24, 211-235.
- Mande, S.C., Mehra, V., Bloom, B.R., and Hol, W.G. (1996). Structure of the heat shock protein chaperonin-10 of *Mycobacterium leprae*. *Science (New York, NY)* 271, 203-207.
- Martin, J., Langer, T., Boteva, R., Schramel, A., Horwich, A.L., and Hartl, F.U. (1991). Chaperonin-mediated protein folding at the surface of groEL through a 'molten globule'-like intermediate. *Nature* 352, 36-42.
- Matthews, C.R. (1993). Pathways of protein folding. *Annual review of biochemistry* 62, 653-683.
- Mayer, M.P., Rüdiger, S., and Bukau, B. (2000). Molecular basis for interactions of the DnaK chaperone with substrates. *Biological chemistry* 381, 877-885.
- Mayhew, M., da Silva, A.C., Martin, J., Erdjument-Bromage, H., Tempst, P., and Hartl, F.U. (1996). Protein folding in the central cavity of the GroEL-GroES chaperonin complex. *Nature* 379, 420-426.
- McCallum, C.D., Do, H., Johnson, a.E., and Frydman, J. (2000). The interaction of the chaperonin tailless complex polypeptide 1 (TCP1) ring complex (TRiC) with ribosome-bound nascent chains examined using photo-cross-linking. *The Journal of cell biology* 149, 591-602.
- Merz, F., Boehringer, D., Schaffitzel, C., Preissler, S., Hoffmann, A., Maier, T., Rutkowska, A., Lozza, J., Ban, N., Bukau, B., *et al.* (2008). Molecular mechanism and structure of Trigger Factor bound to the translating ribosome. *The EMBO journal* 27, 1622-1632.
- Merz, F., Hoffmann, A., Rutkowska, A., Zachmann-Brand, B., Bukau, B., and Deuerling, E. (2006). The C-terminal domain of *Escherichia coli* trigger factor represents the central module of its chaperone activity. *The Journal of biological chemistry* 281, 31963-31971.
- Minton, a.P. (2000). Protein folding: Thickening the broth. *Current biology : CB* 10, R97-99.
- Muller, B.K., Zaychikov, E., Brauchle, C., and Lamb, D.C. (2005). Pulsed interleaved excitation. *Biophys J* 89, 3508-3522.
- Niwa, T., Ueda, T., and Taguchi, H. (2009). [Comprehensive aggregation analysis of whole *E. coli* proteins for understanding of protein aggregation]. *Tanpakushitsu Kakusan Koso* 54, 1870-1875.
- Ojha, A., Anand, M., Bhatt, A., Kremer, L., Jacobs, W.R., Jr., and Hatfull, G.F. (2005). GroEL1: a dedicated chaperone involved in mycolic acid biosynthesis during biofilm formation in mycobacteria. *Cell* 123, 861-873.
- Orengo, C.A., Jones, D.T., and Thornton, J.M. (1994). Protein superfamilies and domain superfolds. *Nature* 372, 631-634.

- Ostermann, J., Horwich, A.L., Neupert, W., and Hartl, F.U. (1989). Protein folding in mitochondria requires complex formation with hsp60 and ATP hydrolysis. *Nature* *341*, 125-130.
- Pace, C.N., Grimsley, G.R., Thomson, J.A., and Barnett, B.J. (1988). Conformational stability and activity of ribonuclease T1 with zero, one, and two intact disulfide bonds. *J Biol Chem* *263*, 11820-11825.
- Park, E.S., Fenton, W.A., and Horwich, A.L. (2005). No evidence for a forced-unfolding mechanism during ATP/GroES binding to substrate-bound GroEL: no observable protection of metastable Rubisco intermediate or GroEL-bound Rubisco from tritium exchange. *FEBS Lett* *579*, 1183-1186.
- Park, E.S., Fenton, W.A., and Horwich, A.L. (2007). Disulfide formation as a probe of folding in GroEL-GroES reveals correct formation of long-range bonds and editing of incorrect short-range ones. *Proc Natl Acad Sci U S A* *104*, 2145-2150.
- Pauling, L., and Corey, R.B. (1951a). The polypeptide-chain configuration in hemoglobin and other globular proteins. *Proc Natl Acad Sci U S A* *37*, 282-285.
- Pauling, L., and Corey, R.B. (1951b). The structure of fibrous proteins of the collagen-gelatin group. *Proc Natl Acad Sci U S A* *37*, 272-281.
- Pellecchia, M., Montgomery, D.L., Stevens, S.Y., {Vander Kooi}, C.W., Feng, H.P., Gierasch, L.M., and Zuiderweg, E.R. (2000). Structural insights into substrate binding by the molecular chaperone DnaK. *Nature structural biology* *7*, 298-303.
- Plaxco, K.W., Simons, K.T., and Baker, D. (1998). Contact order, transition state placement and the refolding rates of single domain proteins. *J Mol Biol* *277*, 985-994.
- Preuss, M., Hutchinson, J.P., and Miller, a.D. (1999). Secondary structure forming propensity coupled with amphiphilicity is an optimal motif in a peptide or protein for association with chaperonin 60 (GroEL). *Biochemistry* *38*, 10272-10286.
- Privalov, P.L. (1996). Intermediate states in protein folding. *Journal of molecular biology* *258*, 707-725.
- Radford, S.E. (2000). Protein folding: progress made and promises ahead. *Trends Biochem Sci* *25*, 611-618.
- Radford, S.E., Dobson, C.M., and Evans, P.A. (1992). The folding of hen lysozyme involves partially structured intermediates and multiple pathways. *Nature* *358*, 302-307.
- Ramachandran, G.N., and Sasisekharan, V. (1968). Conformation of polypeptides and proteins. *Adv Protein Chem* *23*, 283-438.
- Ranson, N.A., Dunster, N.J., Burston, S.G., and Clarke, A.R. (1995). Chaperonins can catalyse the reversal of early aggregation steps when a protein misfolds. *J Mol Biol* *250*, 581-586.
- Ranson, N.A., Farr, G.W., Roseman, A.M., Gowen, B., Fenton, W.A., Horwich, A.L., and Saibil, H.R. (2001). ATP-bound states of GroEL captured by cryo-electron microscopy. *Cell* *107*, 869-879.
- Roseman, A.M., Chen, S., White, H., Braig, K., and Saibil, H.R. (1996). The chaperonin ATPase cycle: mechanism of allosteric switching and movements of substrate-binding domains in GroEL. *Cell* *87*, 241-251.
- Rospert, S., Looser, R., Dubaquié, Y., Matouschek, A., Glick, B.S., and Schatz, G. (1996). Hsp60-independent protein folding in the matrix of yeast mitochondria. *EMBO J* *15*, 764-774.
- Rudiger, S., Buchberger, A., and Bukau, B. (1997). Interaction of Hsp70 chaperones with substrates. *Nat Struct Biol* *4*, 342-349.
- Rutherford, S.L., and Lindquist, S. (1998). Hsp90 as a capacitor for morphological evolution. *Nature* *396*, 336-342.
- Rye, H.S., Burston, S.G., Fenton, W.A., Beechem, J.M., Xu, Z., Sigler, P.B., and Horwich, A.L. (1997). Distinct actions of cis and trans ATP within the double ring of the chaperonin GroEL. *Nature* *388*, 792-798.

- Rye, H.S., Roseman, A.M., Chen, S., Furtak, K., Fenton, W.A., Saibil, H.R., and Horwich, A.L. (1999). GroEL-GroES cycling: ATP and nonnative polypeptide direct alternation of folding-active rings. *Cell* 97, 325-338.
- Saibil, H.R., Zheng, D., Roseman, A.M., Hunter, A.S., Watson, G.M., Chen, S., Auf Der Mauer, A., O'Hara, B.P., Wood, S.P., Mann, N.H., *et al.* (1993). ATP induces large quaternary rearrangements in a cage-like chaperonin structure. *Curr Biol* 3, 265-273.
- Sali, A., Shakhnovich, E., and Karplus, M. (1994). How does a protein fold? *Nature* 369, 248-251.
- Schiene-Fischer, C., Habazettl, J., Schmid, F.X., and Fischer, G. (2002). The hsp70 chaperone DnaK is a secondary amide peptide bond cis-trans isomerase. *Nature structural biology* 9, 419-424.
- Schlünzen, F., Wilson, D.N., Tian, P., Harms, J.M., McInnes, S.J., Hansen, H.a.S., Albrecht, R., Buerger, J., Wilbanks, S.M., and Fucini, P. (2005). The binding mode of the trigger factor on the ribosome: implications for protein folding and SRP interaction. *Structure (London, England : 1993)* 13, 1685-1694.
- Schmid, D., Baici, a., Gehring, H., and Christen, P. (1994). Kinetics of molecular chaperone action. *Science (New York, NY)* 263, 971-973.
- Schmid, F.X. (1993). Prolyl isomerase: enzymatic catalysis of slow protein-folding reactions. *Annual review of biophysics and biomolecular structure* 22, 123-142.
- Sharma, S., Chakraborty, K., Muller, B.K., Astola, N., Tang, Y.C., Lamb, D.C., Hayer-Hartl, M., and Hartl, F.U. (2008). Monitoring protein conformation along the pathway of chaperonin-assisted folding. *Cell* 133, 142-153.
- Sheppard, C.A., Trimmer, E.E., and Matthews, R.G. (1999). Purification and properties of NADH-dependent 5, 10-methylenetetrahydrofolate reductase (MetF) from *Escherichia coli*. *J Bacteriol* 181, 718-725.
- Shortle, D. (1996). The denatured state (the other half of the folding equation) and its role in protein stability. *The FASEB journal : official publication of the Federation of American Societies for Experimental Biology* 10, 27-34.
- Shortle, D., Chan, H.S., and Dill, K.a. (1992). Modeling the effects of mutations on the denatured states of proteins. *Protein science : a publication of the Protein Society* 1, 201-215.
- Shtilerman, M. (1999). Chaperonin Function: Folding by Forced Unfolding. *Science* 284, 822-825.
- Siegers, K., Waldmann, T., Leroux, M.R., Grein, K., Shevchenko, A., Schiebel, E., and Hartl, F.U. (1999). Compartmentation of protein folding in vivo: sequestration of non-native polypeptide by the chaperonin-GimC system. *EMBO J* 18, 75-84.
- Sigler, P.B., Xu, Z., Rye, H.S., Burston, S.G., Fenton, W.A., and Horwich, A.L. (1998). Structure and function in GroEL-mediated protein folding. *Annu Rev Biochem* 67, 581-608.
- Socci, N.D., Onuchic, J.N., and Wolynes, P.G. (1996). Diffusive dynamics of the reaction coordinate for protein folding funnels. *The Journal of Chemical Physics* 104, 5860.
- Sosnick, T.R., Mayne, L., and Englander, S.W. (1996). Molecular collapse: the rate-limiting step in two-state cytochrome c folding. *Proteins* 24, 413-426.
- Sosnick, T.R., Mayne, L., Hiller, R., and Englander, S.W. (1994). The barriers in protein folding. *Nat Struct Biol* 1, 149-156.
- Sparrer, H., Rutkat, K., and Buchner, J. (1997). Catalysis of protein folding by symmetric chaperone complexes. *Proc Natl Acad Sci U S A* 94, 1096-1100.
- Spurlino, J.C., Lu, G.Y., and Quioco, F.A. (1991). The 2.3-A resolution structure of the maltose- or maltodextrin-binding protein, a primary receptor of bacterial active transport and chemotaxis. *J Biol Chem* 266, 5202-5219.
- Steinberg, I.Z. (1971). Long-range nonradiative transfer of electronic excitation energy in proteins and polypeptides. *Annu Rev Biochem* 40, 83-114.
- Stryer, L. (1978). Fluorescence energy transfer as a spectroscopic ruler. *Annu Rev Biochem* 47, 819-846.

- Takagi, F., Koga, N., and Takada, S. (2003). How protein thermodynamics and folding mechanisms are altered by the chaperonin cage: molecular simulations. *Proc Natl Acad Sci U S A* *100*, 11367-11372.
- Tang, Y.C., Chang, H.C., Chakraborty, K., Hartl, F.U., and Hayer-Hartl, M. (2008). Essential role of the chaperonin folding compartment in vivo. *EMBO J* *27*, 1458-1468.
- Tang, Y.C., Chang, H.C., Roeben, A., Wischnewski, D., Wischnewski, N., Kerner, M.J., Hartl, F.U., and Hayer-Hartl, M. (2006). Structural features of the GroEL-GroES nano-cage required for rapid folding of encapsulated protein. *Cell* *125*, 903-914.
- Thirumalai, D., and Lorimer, G.H. (2001). Chaperonin-mediated protein folding. *Annu Rev Biophys Biomol Struct* *30*, 245-269.
- Thulasiraman, V., Yang, C.F., and Frydman, J. (1999). In vivo newly translated polypeptides are sequestered in a protected folding environment. *The EMBO journal* *18*, 85-95.
- Todd, M.J., Viitanen, P.V., and Lorimer, G.H. (1994). Dynamics of the chaperonin ATPase cycle: implications for facilitated protein folding. *Science* *265*, 659-666.
- Tokuriki, N., and Tawfik, D.S. (2009). Chaperonin overexpression promotes genetic variation and enzyme evolution. *Nature* *459*, 668-673.
- Udgaonkar, J.B., and Baldwin, R.L. (1990). Early folding intermediate of ribonuclease A. *Proceedings of the National Academy of Sciences of the United States of America* *87*, 8197-8201.
- Vabulas, R.M., and Hartl, F.U. (2005). Protein synthesis upon acute nutrient restriction relies on proteasome function. *Science* *310*, 1960-1963.
- van den Berg, B., Ellis, R.J., and Dobson, C.M. (1999). Effects of macromolecular crowding on protein folding and aggregation. *The EMBO journal* *18*, 6927-6933.
- Vendruscolo, M., Paci, E., Karplus, M., and Dobson, C.M. (2003). Structures and relative free energies of partially folded states of proteins. *Proc Natl Acad Sci U S A* *100*, 14817-14821.
- Viitanen, P.V., Gatenby, a.a., and Lorimer, G.H. (1992). Purified chaperonin 60 (groEL) interacts with the nonnative states of a multitude of Escherichia coli proteins. *Protein science : a publication of the Protein Society* *1*, 363-369.
- Wales, T.E., and Engen, J.R. (2006). Hydrogen exchange mass spectrometry for the analysis of protein dynamics. *Mass Spectrom Rev* *25*, 158-170.
- Wallin, S., Hammarstrom, L., Blart, E., and Odobel, F. (2006). Electronic interactions and energy transfer in oligothiophene-linked bis-porphyrins. *Photochem Photobiol Sci* *5*, 828-834.
- Wang, J.D., Michelitsch, M.D., and Weissman, J.S. (1998). GroEL-GroES-mediated protein folding requires an intact central cavity. *Proc Natl Acad Sci U S A* *95*, 12163-12168.
- Wang, Z., Feng, H.P., Landry, S.J., Maxwell, J., and Gierasch, L.M. (1999). Basis of substrate binding by the chaperonin GroEL. *Biochemistry* *38*, 12537-12546.
- Weibezahn, J., Schlieker, C., Tessarz, P., Mogk, A., and Bukau, B. (2005). Novel insights into the mechanism of chaperone-assisted protein disaggregation. *Biological chemistry* *386*, 739-744.
- Weissman, J.S., Hohl, C.M., Kovalenko, O., Kashi, Y., Chen, S., Braig, K., Saibil, H.R., Fenton, W.A., and Horwich, A.L. (1995a). Mechanism of GroEL action: productive release of polypeptide from a sequestered position under GroES. *Cell* *83*, 577-587.
- Weissman, J.S., Rye, H.S., Fenton, W.A., Beechem, J.M., and Horwich, A.L. (1996). Characterization of the active intermediate of a GroEL-GroES-mediated protein folding reaction. *Cell* *84*, 481-490.

- Weissman, J.S., Sigler, P.B., and Horwich, A.L. (1995b). From the cradle to the grave: ring complexes in the life of a protein. *Science* 268, 523-524.
- Wildegger, G., and Kiefhaber, T. (1997). Three-state model for lysozyme folding: triangular folding mechanism with an energetically trapped intermediate. *Journal of molecular biology* 270, 294-304.
- Wolfenden, R. (2007). Experimental measures of amino acid hydrophobicity and the topology of transmembrane and globular proteins. *The Journal of general physiology* 129, 357-362.
- Wolynes, P.G., Onuchic, J.N., and Thirumalai, D. (1995). Enat _ 1.0. *Science* 267.
- Xu, Z., Horwich, A.L., and Sigler, P.B. (1997). The crystal structure of the asymmetric GroEL-GroES-(ADP)₇ chaperonin complex. *Nature* 388, 741-750.
- Yang, J.S., Chen, W.W., Skolnick, J., and Shakhnovich, E.I. (2007). All-atom ab initio folding of a diverse set of proteins. *Structure (London, England : 1993)* 15, 53-63.
- Yifrach, O., and Horovitz, A. (1994). Two lines of allosteric communication in the oligomeric chaperonin GroEL are revealed by the single mutation Arg196-->Ala. *J Mol Biol* 243, 397-401.
- Yifrach, O., and Horovitz, A. (1995). Nested cooperativity in the ATPase activity of the oligomeric chaperonin GroEL. *Biochemistry* 34, 5303-5308.
- Zahn, R., Axmann, S.E., Rucknagel, K.P., Jaeger, E., Laminet, A.A., and Pluckthun, A. (1994a). Thermodynamic partitioning model for hydrophobic binding of polypeptides by GroEL. I. GroEL recognizes the signal sequences of beta-lactamase precursor. *J Mol Biol* 242, 150-164.
- Zahn, R., and Pluckthun, A. (1994). Thermodynamic partitioning model for hydrophobic binding of polypeptides by GroEL. II. GroEL recognizes thermally unfolded mature beta-lactamase. *J Mol Biol* 242, 165-174.
- Zahn, R., Spitzfaden, C., Ottiger, M., Wuthrich, K., and Pluckthun, A. (1994b). Destabilization of the complete protein secondary structure on binding to the chaperone GroEL. *Nature* 368, 261-265.
- Zhang, Y.H., Yan, X., Maier, C.S., Schimerlik, M.I., and 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 Sci* 10, 2336-2345.
- Zhou, H.X. (2004). Protein folding and binding in confined spaces and in crowded solutions. *J Mol Recognit* 17, 368-375.
- Zhu, X., Zhao, X., Burkholder, W.F., Gragerov, a., Ogata, C.M., Gottesman, M.E., and Hendrickson, W.a. (1996). Structural analysis of substrate binding by the molecular chaperone DnaK. *Science (New York, NY)* 272, 1606-1614.
- Zimmerman, S.B., and Minton, A.P. (1993). CROWDING : Biochemical ,. *Molecules*, 27-65.

A. Appendix

A.1. Supplementary data

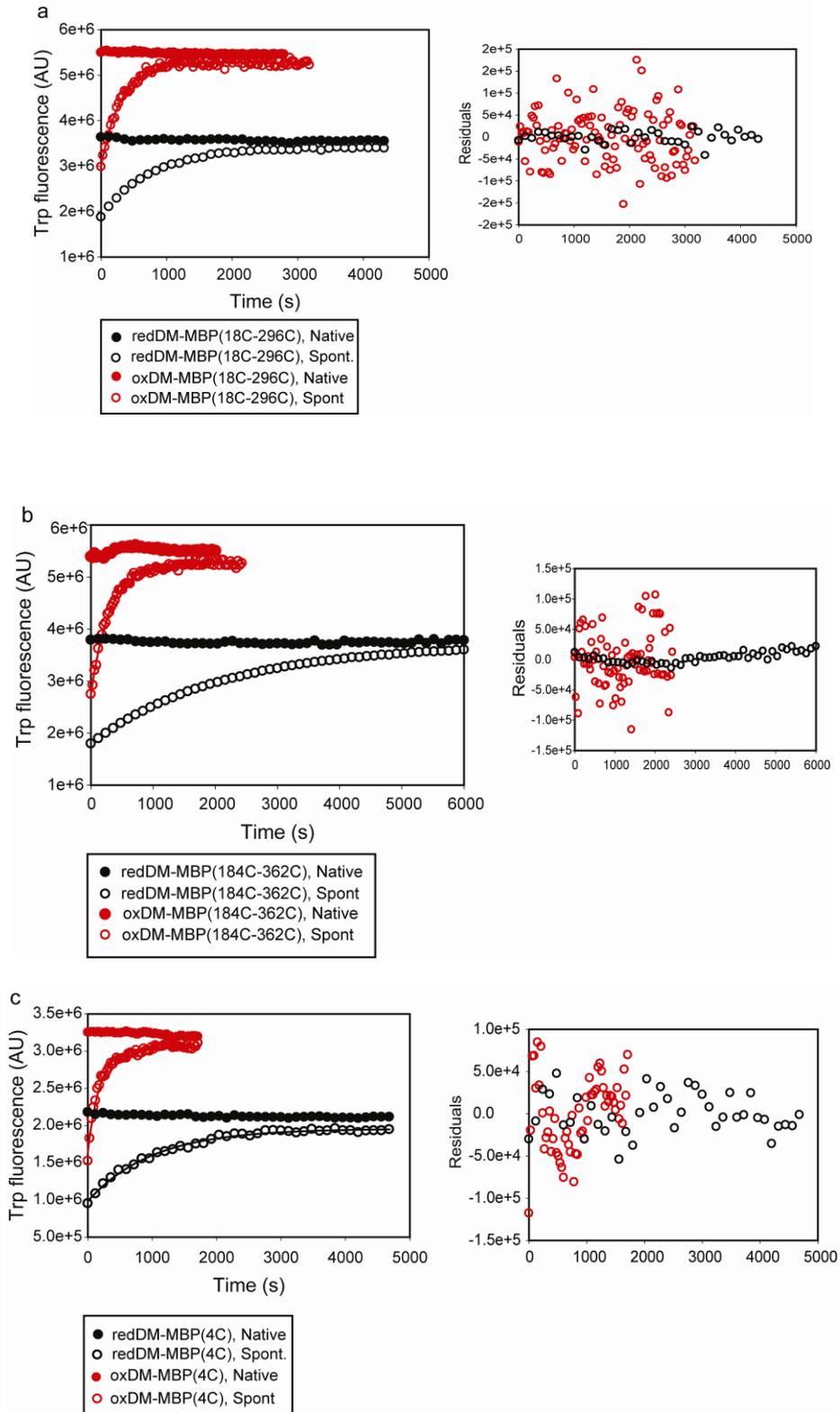
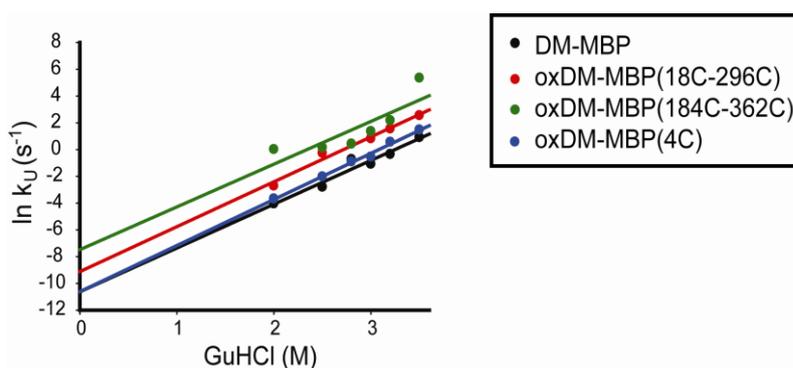


Figure A.1.1: Single Exponential Rates and Yield of Refolding of DM-MBP Cysteine Mutants.

Representative raw data of rates and yields of spontaneous refolding of (a) DM-MBP(18C-296C), (b) DM-MBP(184C-362C) and (c) DM-MBP(4C) at final concentration of 250 nM are shown. Reduced (red.) or oxidized (ox.) DM-MBP cysteine mutants were denatured in 6 M GuHCl and diluted 100-fold into buffer B at 25°C. Refolding was monitored by Trp fluorescence at 345 nm. Note that the Trp fluorescence of the reduced proteins (black curves) is quenched relative to the oxidized proteins (red curves). The apparent rates of folding fitted to a single exponential. The residuals of the fits are shown in the inserts. The fluorescence intensities of native protein controls are shown. Refolding yields were between 80-100%.

**Figure A.1.2: Kinetics of Unfolding.**

The rate of unfolding of DM-MBP and reduced oxidized (ox) DM-MBP(18C-296C), DM-MBP(184C-362C) and DM-MBP(4C) at different concentrations of GuHCl in buffer B 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.

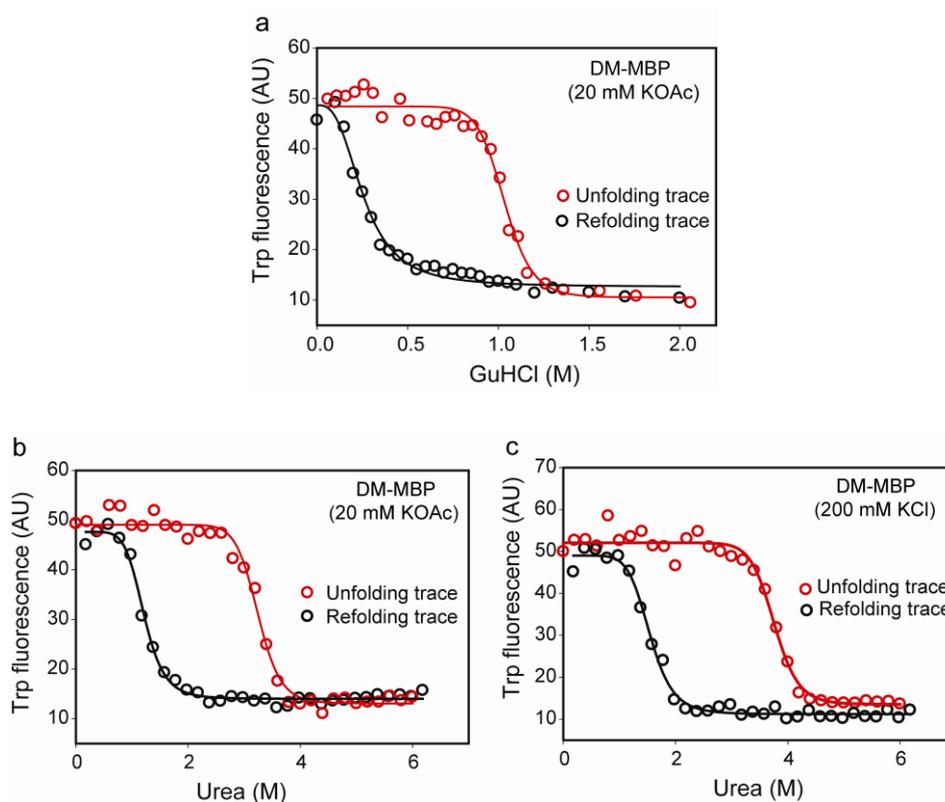


Figure A.1.3: Denaturant-Dependent Unfolding and Refolding of DM-MBP.

Unfolding and refolding of DM-MBP at a final concentration of 1 μM in different denaturant concentrations was monitored by Trp fluorescence at 345 nm. *Unfolding trace:* Native DM-MBP was incubated for 12 h in (a) buffer (20 mM Tris, pH 7.5, 20 mM KOAc, 5 mM $\text{Mg}(\text{OAc})_2$ containing ~60 mM to ~2 M GuHCl (b) buffer (100mM HEPES, pH 7.2, 20 mM KOAc, 5 mM $\text{Mg}(\text{OAc})_2$) containing ~120 -6 M Urea and (c) buffer (100 mM pH 7.2, 20 mM KOAc, 5 mM mM $\text{Mg}(\text{OAc})_2$) containing ~120 -6 M Urea at 25°C. *Refolding trace:* DM-MBP (100 μM) was unfolded in (a) 6 M GuHCl (b.c) 9 M Urea and then diluted 100-fold in respective buffers containing increasing concentrations of GuHCl or urea as indicated followed by incubation for 12 h at 25°C.

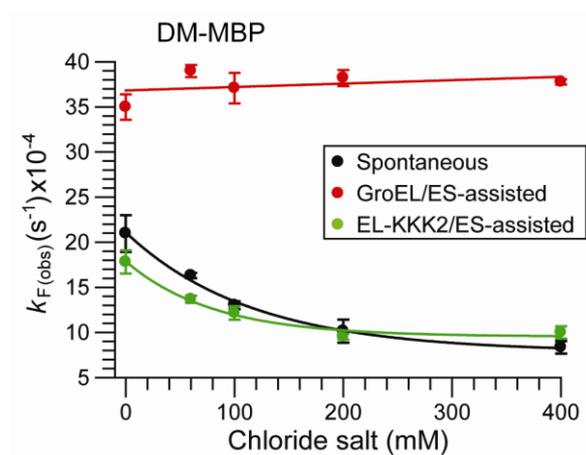
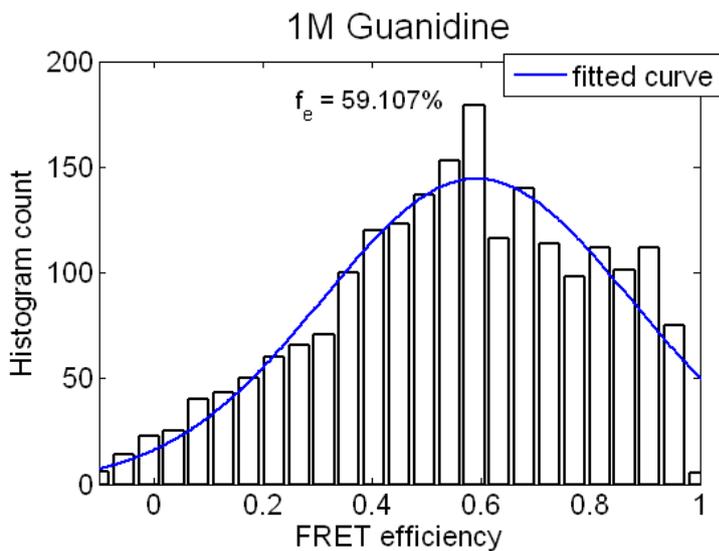
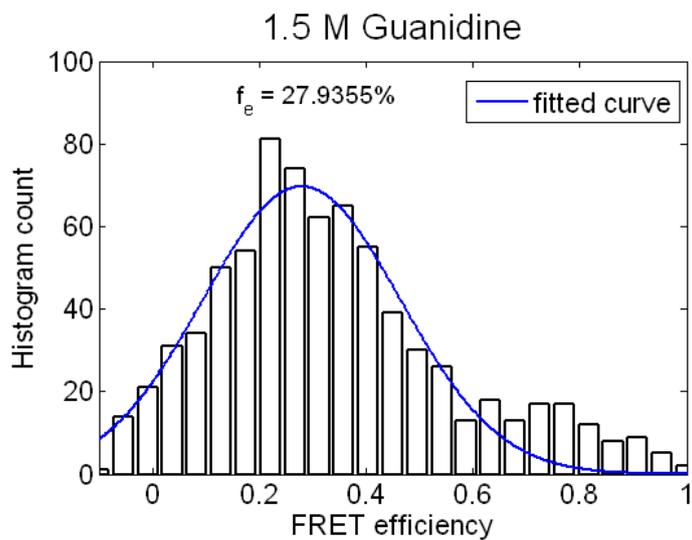
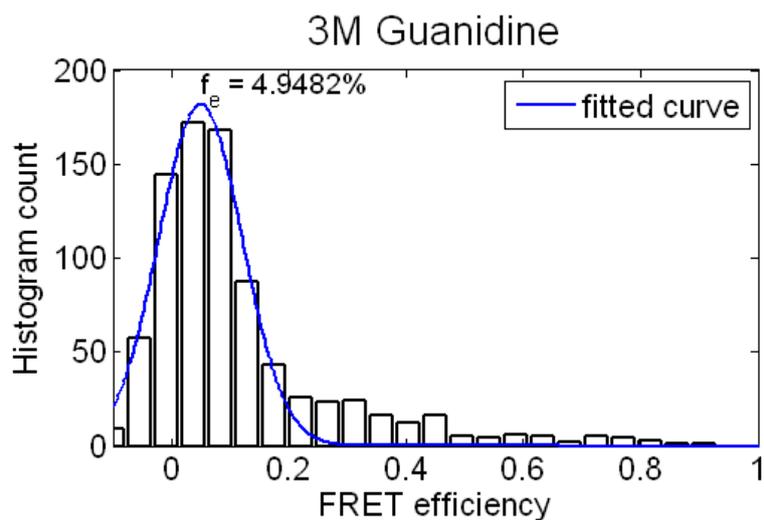


Figure A.1.4: Effect of Chloride Salt on Spontaneous and Assisted Refolding of DM-MBP.

Rates of spontaneous and assisted refolding of DM-MBP at varying concentrations of chloride salt. DM-MBP (25 μM) was denatured in 9 M urea and diluted 100-fold into buffer (100 mM Hepes pH 7.2, 20 mM KOAc, 5 mM Mg(OAc)₂) containing the chloride salt (KCl) concentrations indicated either in the absence of GroEL (spontaneous), with 1 μM GroEL/2 μM GroES or with 1 μM ELKKK2/2 μM GroES at 25 °C (assisted). Assisted refolding was initiated by the addition of 2 mM ATP. Refolding was monitored by Trp fluorescence. Standard deviation from 3 independent experiments are shown.



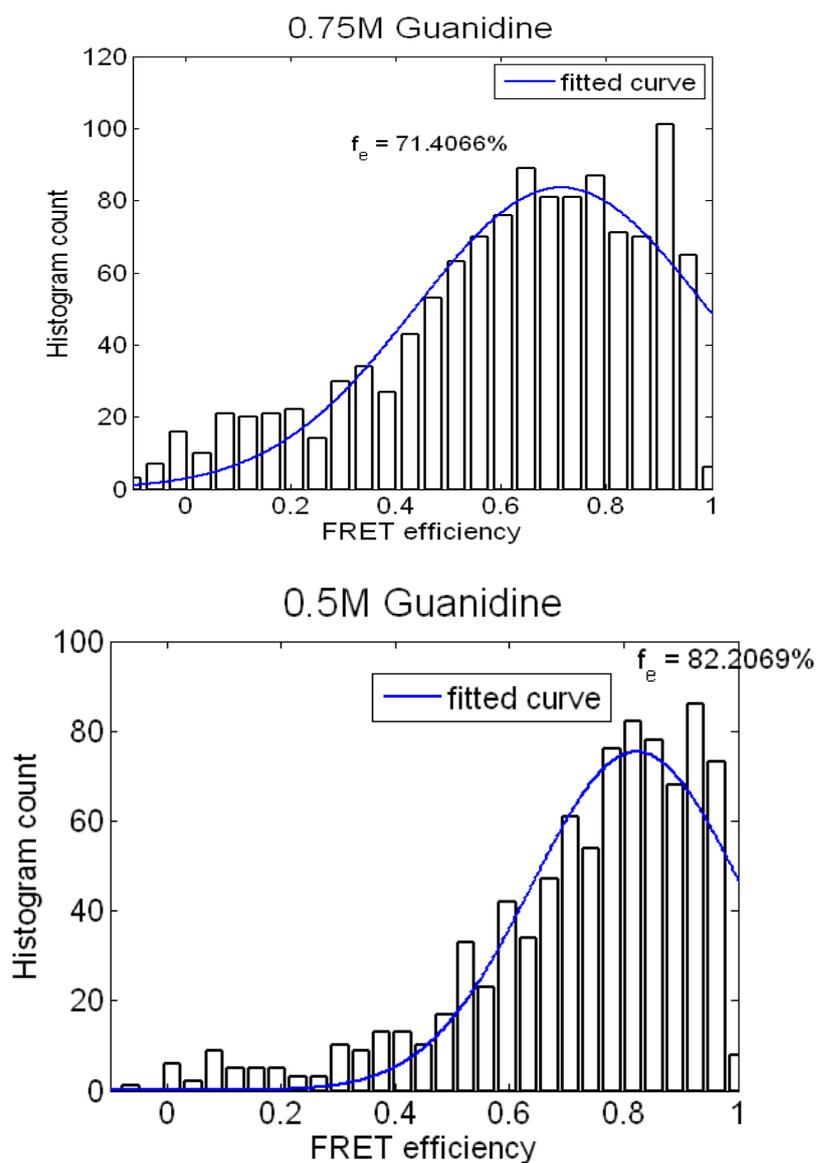


Figure A.1.5: SpFRET Analysis in Solution of DM-MBP(52-298).

DM-MBP(52-298) double-labeled with Atto532 at position 52 and Atto647N at position 298 was denatured in 3 M GuHCl/Buffer A and diluted in different concentration of GuHCl as indicated. Final protein concentration was ~ 100 pM. Peak values of a Gaussian fit to the FRET efficiency distributions (f_E) are indicated. Shift in the peak values of f_E with increasing concentrations of GuHCl from 0.5 M (Intermediate state) to 3 M GuHCl (Denatured state) corresponds with the Trp fluorescence intensity in the hysteresis profile of DM-MBP (Figure 5.8 and figure 5.13).

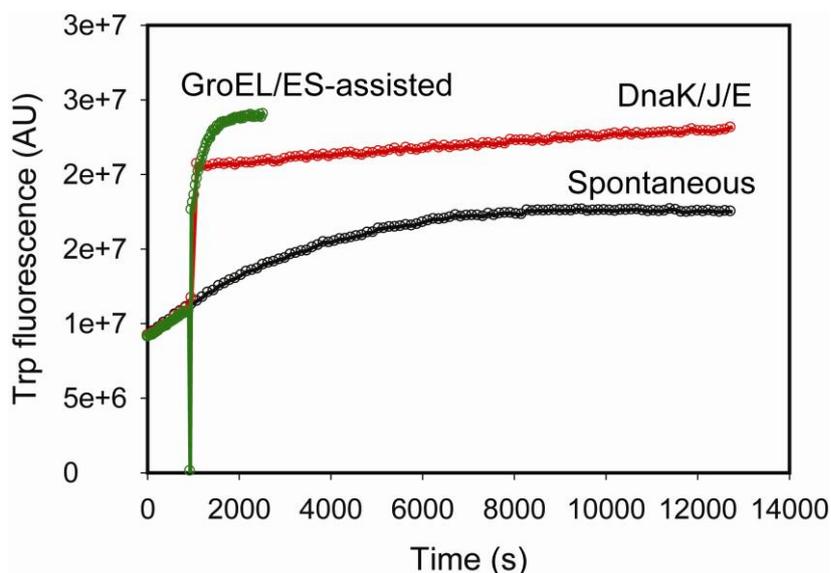


Figure A.1.6: Recognition of the Intermediate State by GroEL and DnaK.

25 μ M DM-MBP was denatured in 6 M GuHCl and diluted 100-fold in buffer A (Spontaneous refolding). After 15 min of spontaneous refolding either 0.5 μ M GroEL/1 μ M GroES/5 mM ATP or 1.25 μ M DnaK/0.625 DnaJ/ 1.25 GrpE/ 5 mM ATP was added to check for the binding of the intermediate state to GroEL or DnaK.

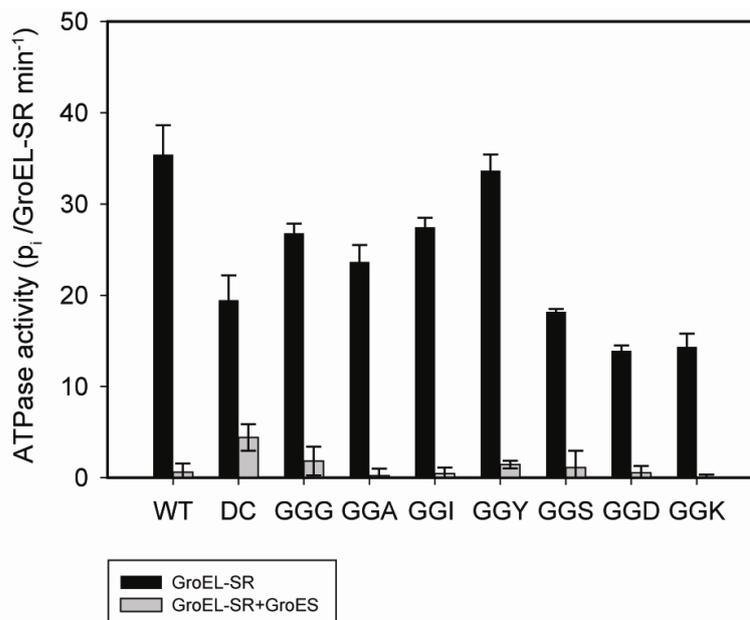


Figure A.1.7: ATPase Activity of SR-EL Tail Mutants.

ATP hydrolysis rate was measured by following decrease in absorbance at 340 nm in presence of either SR-EL alone (black) or SR-EL and GroES (Grey). ATPase rates of SR-EL are indicated as number of ATP hydrolyzed per SR-EL heptamer per minute. Standard deviation of three independent measurements are shown.

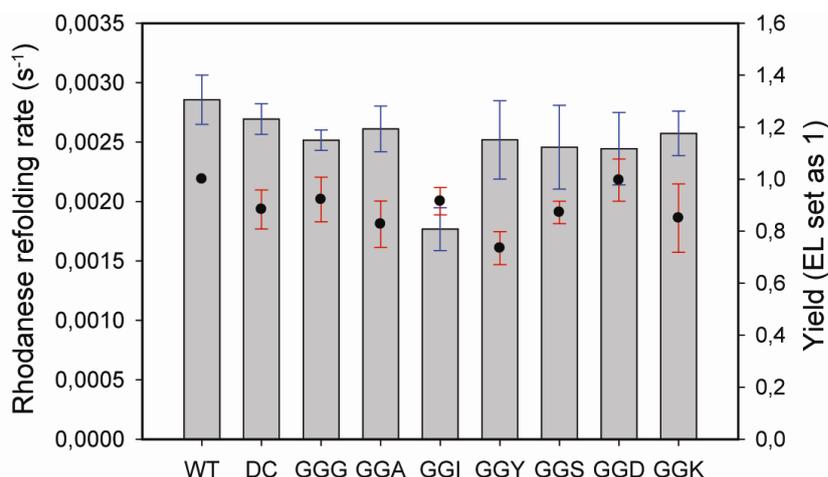


Figure A.1.8: Rhodanese Refolding in Presence of SR-EL Tail Mutants.

Refolding of rhodanese was performed in presence of SR-EL tail mutants. 25 μM rhodanese was denatured in 6 M GuHCl and 100-fold diluted in buffer A with 1 μM SR-EL. 2 μM SR-ELGGD was used in the refolding reaction due to low binding affinity of this mutant with rhodanese. GroES was present in 2-fold excess of the chaperonin. Refolding was initiated by addition of 5 mM ATP at 25°C. Enzyme activity was measured at different time points by taking absorbance at 460 nm. Standard deviation of three independent measurements are shown. Bars indicate rates and circle indicates yield of refolding.

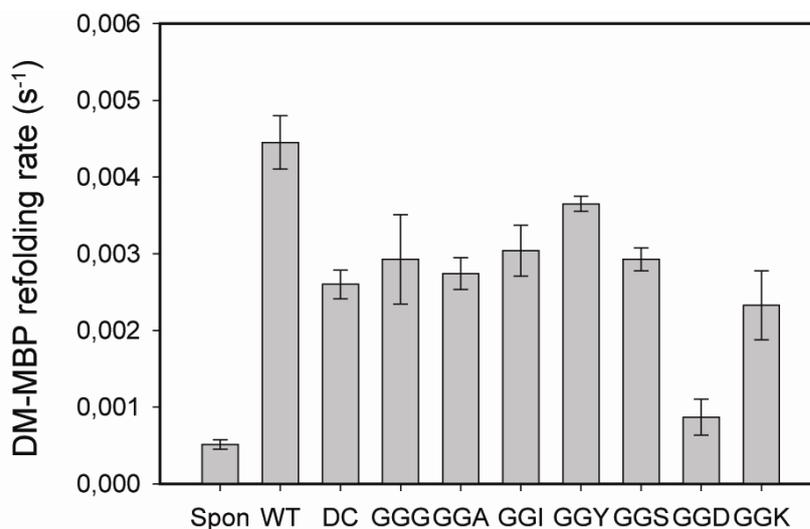


Figure A.1.9: DM-MBP Refolding in Presence of SR-EL Tail Mutants.

DM-MBP refolding was performed in presence of SR-EL tail mutants. 25 μM DM-MBP was denatured in 6 M GuHCl and 100-fold diluted in buffer A with 1 μM GroEL. GroES was present in 2-fold excess of chaperonin. Refolding was initiated by addition of 5 mM ATP at 25°C and followed by monitoring Trp fluorescence with excitation at 295 nm and emission at 345 nm. Standard deviation of three independent measurements are shown.

A.2. Abbreviations

Units are expressed according to the international system of units (SI), including outside units accepted for use with the SI. Amino acids are abbreviated as their one or three letter symbols.

Protein names are abbreviated according to their SWISS-PROT database entries.

ADP	adenosine 5'-diphosphate
Amp	ampicillin
AMP-PNP	adenosine 5'-(β,γ -imido)triphosphate tetralithium salt
APS	ammonium peroxodisulfate
ATP	adenosine 5'-triphosphate
BSA	albumin bovine serum
CAM	chloramphenicol
CDTA	<i>trans</i> -1,2-diaminocyclohexane- <i>N,N,N',N'</i> -tetraacetic acid
DNA	deoxyribonucleic acid
DnaJ	bacterial Hsp40 chaperone
DnaK	bacterial Hsp70 chaperone
DTT	dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	ethylenediaminetetraacetic acid
g	acceleration of gravity, 9.81 m/s ²
GuHCl	guanidinium hydrochloride
GFP	Green fluorescent protein
GroEL	bacterial Hsp60 chaperonin
GroES	bacterial Hsp10 cochaperonin
GrpE	bacterial nucleotide exchange factor of DnaK

h	hour
IPTG	isopropyl- β -D-1-thiogalactopyranoside
Kan	kanamycin
LB	Luria Bertani
MBP	Maltose binding protein
METF	5,10-methylenetetrahydrofolate reductase
MOPS	3-(N-morpholino)propanesulfonic acid
NAC	nascent chain-associate complex
NADPH	β -nicotinamide adenine dinucleotide 2'-phosphate
OAc	acetate
OD	optical density
PAGE	PolyAcrylamide Gel Electrophoresis
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank. Repository for processing and distribution of 3-D structure data of proteins and nucleic acids. http://www.rcsb.org/pdb/
PPIase	prolyl- <i>cis/trans</i> isomerase
RAC	ribosome-associate complex
<i>S. cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
SDS	sodiumdodecylsulfate
TEMED	<i>N,N,N',N'</i> -tetramethylethylenediamine
TF	trigger factor
TRiC	Tailless complex polypeptide ring complex
Tris HCl	tris(hydroxymethyl)aminomethane hydrochloride

A.3. Curriculum Vitae

Name	Jyoti Sinha
Birth Date	26, February, 1982
Birth Place	Patna, Bihar
Nationality	India

Dissertation

Oct. 2006- Present	Ph.D dissertation under the supervision of Prof. Dr. F. Ulrich Hartl in the Department of Cellular Biochemistry, Max Planck Institute of Biochemistry, Martinsried: <i>Chaperonin-Catalyzed rescue of kinetically trapped states in protein folding.</i>
--------------------	--

Education

May 2005 – May 2006	Masters thesis at School of Biotechnology, Jawaharlal Nehru University, New Delhi, Dr. Rakesh Bhatnagar: Cloning, Expression and Purification of GroEL from <i>Bacillus anthracis</i>
May 2004 – May 2006	Masters of Science at School of Biotechnology, Jawaharlal Nehru University, New Delhi
May 2001 – April 2004	Bachelors in Science (Chemistry Honors) at Faculty of Science, Banaras Hindu University, Varanasi.

Higher Education

June 1998	High School, Government Intermediate College, Dalmianagar, Rohtas, Bihar.
-----------	---