

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

Unfolding and Compaction in
Chaperonin-assisted Protein Folding
followed by Single molecule and Ensemble
FRET

Shruti Sharma
aus
Saharsa, Bihar, India

2007

Erklärung

Diese Dissertation wurde im Sinne von §13 Abs. 3 bzw. 4 der Promotionsordnung vom 29. Januar 1998 von Herrn Prof. Dr. F.U. Hartl betreut.

Ehrenwörtliche Versicherung

Diese Dissertation wurde selbständig, ohne unerlaubte Hilfe erarbeitet.

München, den

Dissertation eingereicht am 04.09.2007

1. Gutachter: Prof. Dr. F.U. Hartl

2. Gutachter: Dr. Don Lamb

Mündliche Prüfung am 09.10.2007

Courage to give
Courage to think different,
Courage to invent,
Courage to discover the impossible,
Courage to travel into an unexplored path,
Courage to share the knowledge,
Courage to remove the pain,
Courage to reach the unreached,
Courage to combat the problem and Succeed,
Are the unique Qualities of youth.

- A.P.J ABDUL KALAM
XI PRESIDENT OF INDIA

Acknowledgements

I am grateful to Dr. Manajit Hayer-Hartl and Prof. Dr. F. Ulrich Hartl for accepting me as a student and giving me the opportunity to learn from and interact with some of the best minds that there are. For their faith, support and allowing me the freedom to make and learn from my mistakes.

I would like to thank Dr. Don C. Lamb our collaborator and co-referee of my thesis, for his time, patience and insightful discussions during the course of this work; without which it would not have been possible for me to complete my thesis in time.

I would also like to thank Barbara K. Müller for making day to day working smooth and fun, for not minding late evening schedules and for keeping up with our ever increasing demands for more measurement time on her instrument.

I would also take this opportunity to thank our very competent and helpful support system i.e Andrea and Silke who have been there, ready with a kind word and solutions to all our administrative and other problems. I am also grateful to Elisabeth, Bernd, Dirk, Nadine and Emmanuel for all their logistical support and help.

Contribution of Dr. Raina Boteva and Chi has been invaluable to my learning curve. I would also like to I would also thank all my colleagues (Angela, Michi, Tobias, Giulia, Alice, Sathish, Jose, Christian, Kaiser, Leslie, Niclas, Alex and last but not the least Jyoti) in the department of cellular biochemistry for making working and learning pleasurable. I am especially indebted to all the *Friday-evening-discussions* with all these people at some point during my doctoral work. I would also like to thank Bhavna, Rinki and Anu for going through this phase with me literally and figuratively across all the distances.

This work essentially is a result of the team work together with Nagore and Kausik, with whom I have learned more about science and scientific temper than just technicalities.

To Bhumi for being the cause and effect in the scheme of things. Last but not the least to my parents and family for their belief and love and for all the *V Discipline*.

Contents

1	Summary	5
2	Introduction	7
2.1	Protein Folding	7
2.1.1	Protein Structure	8
2.1.2	Protein Folding: the Determinants of Structure	8
2.1.3	Levinthal Paradox and Folding Mechanisms	9
2.2	Protein Folding <i>in Vivo</i>	14
2.2.1	Ribosome-Associated Chaperones	17
2.2.2	Hsp70 Chaperones	18
2.2.3	Hsp60 and Hsp10: The Chaperonins	18
2.2.4	Additional Chaperone Systems	19
2.2.5	Cellular Chaperone Networks	20
2.3	GroE Chaperone System	21
2.3.1	Structural Information of GroEL-GroES Machinery	21
2.3.2	Effect of Nucleotides	23
2.3.3	Substrates of GroEL-GroES	24
2.3.4	Maltose Binding Protein - Model Substrate	25
2.3.5	Conformational Properties of Chaperonin Bound Substrates	26
2.3.6	GroEL-GroES Reaction Cycle in Protein Folding	29
2.4	Fluorescence	31
2.4.1	FRET	32
2.4.2	Single Molecule FRET	36
2.4.3	Anisotropy	38
2.5	Aim of the Study	42
3	Materials and Methods	43
3.1	Materials	43
3.1.1	Chemicals	43
3.1.2	Enzymes	45
3.1.3	Materials	45

3.1.4	Instruments	45
3.1.5	Media	46
3.1.6	Antibiotic Stock Solutions	46
3.2	Plasmids	46
3.3	Molecular Cloning Methods	49
3.3.1	Preparation and Transformation of <i>E. coli</i> competent cells	49
3.3.2	Plasmid Purification	49
3.3.3	PCR Amplification	50
3.4	DNA Restriction and Ligation	50
3.5	DNA Analytical Methods	50
3.6	Protein Purification	51
3.6.1	GroEL Expression and Purification	51
3.6.2	GroES Expression and Purification	51
3.6.3	Expression of MBP and MBP Mutants and Purification	52
3.6.4	Thiol-Mediated Labelling of the Cys Constructs	52
3.6.5	Thiol-Mediated Labelling of GroEL and GroES	53
3.7	Protein analytical methods	53
3.7.1	Determination of Protein Concentration	53
3.7.2	SDS-PAGE	53
3.8	Biochemical and Biophysical Methods	54
3.8.1	MBP Refolding	54
3.8.2	Stopped-Flow Experiments	54
3.8.3	Preparation of GroEL-ADP-GroES "Bullet"	55
3.8.4	Equilibrium Unfolding of MBP	55
3.8.5	Sample for Single Molecule Spectroscopy of Chaperonin-Substrate Complex	55
3.8.6	Sample Preparation for Single Molecule Spectroscopy of GroEL- Cycling	55
3.8.7	Calculation of FRET Efficiencies Using Burst Analysis	56
3.8.8	Determination of the Forster Radius	57
4	Results	59
4.1	Sp-FRET Distribution of Double labelled DM-MBP	59
4.2	Conformational Heterogeneity of GroEL Bound Protein	62
4.3	Diffusion Coefficients of Different States of DM-MBP	65
4.4	Anisotropy Decay for Various DM-MBP Labelled Proteins	67
4.5	Polyproline Fusion Protein - A Spectroscopic Ruler.	67
4.6	Conformational Dynamics of GroEL Bound Substrate Upon Nucleotide Binding	69

4.7	Effect of GroES Binding on Chaperonin-Substrate Complex	71
4.8	Stopped-Flow Ensemble FRET Analysis of Nucleotide and GroES Binding	73
4.9	GroEL Cycling and the Bound-State	76
4.10	Binding and Segmental Release of Substrate by GroEL	80
4.11	GroEL-GroES assisted and Spontaneous Folding Pathways	83
5	Discussion	89
5.1	GroE and Protein Folding	89
5.2	The Chaperonin DM-MBP Bound State at Single-Molecule Resolution.	90
5.3	Bimodal Distribution and Allosteric Regulation	90
5.4	GroEL Binding Induced Local Unfolding	91
5.5	Bound State and the Rate of Folding	91
5.6	Nucleotide Induced Transient Unfolding and Subsequent Compaction Upon GroES Binding	92
5.7	Step-Wise Release of the Substrate Protein from GroEL	93
5.8	Rate Acceleration by GroEL via a Different Folding Pathway	94
5.9	The GroEL-Substrate Complex Returns to the Initial Bound State During GroEL Cycling.	95
5.10	GroEL Dependence of a Substrate Protein	97
A	Appendix	99
A.1	Sp-FRET Measurement of DM-MBP Cys Constructs	99
A.2	Abbreviations	118
A.3	Curriculum Vitae	120

Chapter 1

Summary

To become biologically active, most proteins need to fold into precise three dimensional structures. It has been well established that all the folding information is contained within the primary structure of a protein. However, the mechanisms utilized by proteins to avoid sampling the extraordinarily large amount of possible conformations during their folding process are just beginning to be understood. Molecular chaperones assist the folding of newly synthesized and denatured proteins in acquiring their native state in the crowded intracellular environment.

As a nascent chain leaves the ribosome, it is captured first by the upstream chaperones and then possibly transferred to the downstream chaperonins. GroEL-GroES, the Hsp-60 of *E.coli*, is one of the best studied chaperone systems. An appreciable amount of data is available providing information regarding its structure and function. GroEL encapsulates the substrate into the central cavity where folding occurs unimpaired by aggregation and unwanted inter-molecular interactions. Nevertheless, many important aspects of the GroEL mechanism remain to be addressed. Some of the open questions we have addressed in this study include: In what conformation does a substrate protein bind to the apical domains of GroEL; how is it that GroEL is able to accelerate the rate of folding of certain proteins, and how do the conformational properties of the substrate change as it undergoes repeated cycling.

By using ensemble FRET and Sp-FRET (Single Pair-Fluorescence Resonance Energy Transfer), we have probed the conformation of the model substrate DM-MBP (Double Mutant Maltose Binding Protein) during different stages of the functional cycle of GroEL. With Sp-FRET coupled to PIE (Pulsed Interleaved Excitation), we have been able to explore the heterogeneity of the GroEL bound substrate protein and observed a bimodal conformational distribution. One of the two populations is as compact as the native state, whereas the other is as extended as the unfolded protein in denaturant. This unfolding is a local phenomena and can also be observed when the substrate is

transferred from DnaK/J system (bacterial Hsp70) to GroEL, indicating the possibility of the existence of this conformational heterogeneity *in vivo* as the protein follows the cellular chaperone pathway.

Subsequent to GroEL binding, there is a transient expansion of the protein upon binding of ATP to GroEL, followed by compaction when GroES triggers the encapsulation of the protein inside the chaperonin cage. This transient expansion is however found not to be a necessary event for the rate acceleration of DM-MBP folding, since ADP-AlFx (transition state analogue of ATP hydrolysis) results in a much slower rate of expansion, which does not cause a change in the folding rate.

Anisotropy measurements, probing the freedom of motion of different regions of the GroEL bound protein, revealed that there is a segmental release of the substrate protein from the GroEL surface upon binding of ATP and GroES. As a consequence, the hydrophobic collapse of the protein upon encapsulation by GroES follows a step-wise mechanism. In this process, less hydrophobic regions are released upon binding of ATP, prior to more hydrophobic ones which are released only by GroES binding. Thus, the order of Hydrophobic collapse is reversed as compared to spontaneous folding possibly resulting in conformationally different folding intermediates.

Evidence that the folding pathway inside the cage differs from that of spontaneous folding was obtained by observing the effect of external perturbations (e.g. mutations in substrate protein and use of different solvent conditions) on the rate of spontaneous and GroEL assisted folding reactions. These two folding reactions respond differently to the introduced perturbations. Kinetic data obtained from ensemble FRET measurements suggest that the conformation of refolding intermediate is altered by the GroEL cavity, which leads to a folding pathway that is different from the spontaneous refolding pathway.

In summary, this study revealed significant novel aspects of the GroEL folding mechanism and provided insights into the basis of rate acceleration of the substrate protein by the chaperonin. This work may thus contribute to advance our fundamental knowledge of the chaperonin system and the basic mechanism of protein folding.

Chapter 2

Introduction

2.1 Protein Folding

Proteins constitute more than half of a cell's dry weight. Proteins determine the shape and the structure of a cell and also serve as the main instruments of molecular recognition and catalysis. In computer terminology, the DNA and mRNA molecules represent the software instructions that a cell receives from its parent. Proteins and catalytic RNA molecules are the hardware or machinery that executes the program stored in the memory.

The process by which a protein acquires its unique three-dimensional structure is called protein folding. The importance of protein folding has been recognized for many years. Over half a century ago, Linus Pauling discovered two quite simple regular arrangements of amino acids; the α helix and the β sheet that are found in almost every protein (Pauling and Corey, 1951a; Pauling and Corey, 1951c). The pioneering discovery of *in vitro* protein folding by Anfinsen demonstrated that formation of the native protein from the unfolded state is a spontaneous process determined by the global free energy minimum, and is encoded in the primary sequence of the protein.

Proteins are linear polymers of up to 20 different amino acids (aa). All amino acids contain an amino - and a carboxyl-group covalently linked to a central C_{α} -atom. In addition, a functional group (or a hydrogen atom) is coupled to the C_{α} -atom, which is different for each amino acid. In a condensation reaction, amino acids can polymerize to polypeptides (or proteins) via the formation of a planar peptide bond. The peptide bonds form the backbone of the protein. In all organisms, proteins play a pivotal role in all cellular processes such as maintenance of cellular structure and integrity, inter- and intra-cellular communication, metabolism and transport. To fulfill these functions, proteins must adopt specific three-dimensional structures, which are very diverse, including

fibrillar shapes of structural proteins, globular structures of some metabolic enzymes and channels traversing membranes in energy producing systems.

2.1.1 Protein Structure

Four hierarchical levels of protein structure are distinguished: The *primary structure* is the sequence of the different amino acids and the *secondary structure* refers to common repeating structures found in many proteins. The two most common secondary structure elements are the α -helix and the β -pleated sheet (Pauling and Corey, 1951b). The *tertiary structure* describes the overall structure of the folded polypeptide chain, including the arrangements of secondary structural elements. The *quaternary structure* constitutes the arrangement of subunits in an assembly of two or more polypeptide chains. Many proteins are organized in a modular fashion. These modules are referred to as domains, which are typically 100 to 300 aa in length and are structurally and functionally distinct units (Doolittle, 1995; Orengo et al., 1994). Linking two or more domains on the primary structure level 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).

2.1.2 Protein Folding: the Determinants of Structure

Pioneering experiments on protein folding were performed in the late 1960's by Anfinsen and co-workers. Purified, denatured Ribonuclease A was shown to be able to regain its native state in solution upon removal of denaturant, as measured by its enzymatic activity, in a spontaneous and unaided fashion (Anfinsen, 1972; Anfinsen, 1973; Taniuchi and Anfinsen, 1969). These experiments demonstrated that the information defining the tertiary structure of a protein is contained within its amino acid sequence. The native tertiary structure is the polypeptide conformation with the lowest free energy in physiological conditions. The spontaneous folding process was found to be reversible and occurred on a biologically relevant time scale (Schechter et al., 1970). Usually, folding does not involve the formation or breakage of covalent bonds. Rather, the polypeptide is conformationally flexible through rotation around single bonds. The natively folded structure is stabilized by a multitude of weak interactions, such as hydrogen bonds and hydrophobic and ionic interactions. Considering only the protein backbone, rotation can occur around the axis formed by the C_α and the carboxyl carbon as well as around the axis formed by the C_α and the amino nitrogen, described by the angles ψ and ϕ , respectively (Fig.2.1). No rotation can occur around the axis formed by the peptide bond due

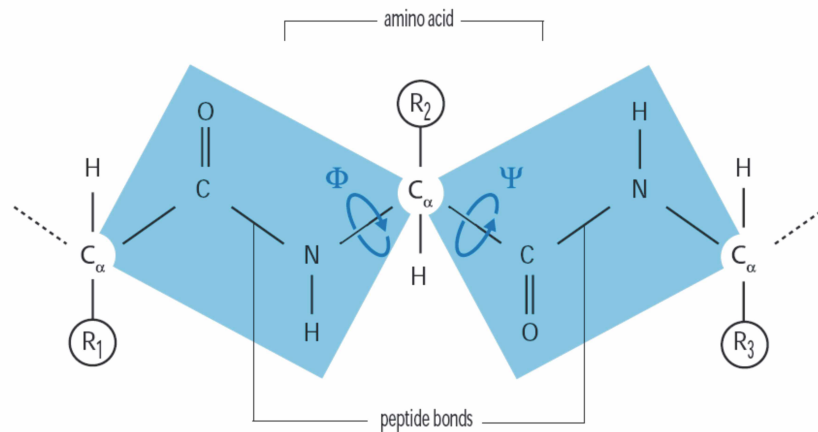


Figure 2.1: Steric Limitations of the Peptide Bond Angles in a Protein.. Each amino acid contributes three bonds to the polypeptide backbone. The peptide bond is planar (blue shading) and does not allow rotation. The $N-C_\alpha$ and $C_\alpha-C$ bonds, however, allow rotation - their angles are called ϕ and ψ . R is used to indicate the side chain residues of the corresponding amino acid.

to its planar nature. Thus, the conformation of the backbone can be described by a pair of angles, ψ and ϕ , for each peptide bond. It was found that two of the numerous possible pairs of ψ and ϕ occur with a high frequency in natively folded proteins (Ramachandran et al., 1963). These two pairs of angles are referred to as the α and the β conformation.

2.1.3 Levinthal Paradox and Folding Mechanisms

There are at least 100 fundamentally different folds adopted by natural protein domains and many variants within these. The question of how individual proteins efficiently and reliably achieve their native state following synthesis is one of the most intriguing problems in structural biology. The fundamental question is how a sequence codes for the fold. Two features of proteins make this question particularly intriguing. Firstly, since the main chains of all proteins have an identical composition, how do the side chains dictate the overall fold. Secondly, since the number of possible conformations of a polypeptide chain is astronomically large, how does a given polypeptide find its native structure in a finite time? The latter problem is known as the Levinthal paradox (Levinthal et al., 1962; Plocke et al., 1962).

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 approx. 2^{150} different

conformations. Taking into account that the fastest possible rate for conformational changes is approx. $10^{11}s^{-1}$, it would take the hypothetical protein more than 10^{11} years to reach its native structure (Dinner et al., 2000; Adesnik and Levinthal, 1969). The discrepancy between the estimated time for random folding and the observed fast folding of proteins is called the Levinthal paradox.

These considerations led to the conclusion that a protein cannot fold by sampling the entire available conformational space randomly. Rather, the efficient folding must proceed through transient 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 (Agashe et al., 1995; Jamin and Baldwin, 1996; Radford et al., 1992b; Udgaonkar and Baldwin, 1990; Wildegger and Kiefhaber, 1997).

Three mechanisms were put forth to provide insight into the different modes by which a protein reaches its native state, namely (Fersht, 1997):

- The **framework model** proposed that local elements of native local secondary structure could form independently of tertiary structure. These elements would diffuse until they collide, come together successfully to give the tertiary structure (Kim and Baldwin, 1982; Kim and Baldwin, 1990; Ptitsyn and Rashin, 1975).
- The **hydrophobic collapse model** dictates that a protein rapidly undergoes a collapse around its hydrophobic side-chains and then rearranges from the restricted mobility and sterically restricted conformations in which secondary structures are directed by native like tertiary interactions (Baldwin, 1989; Schellman, 1955).
- The **nucleation model** states that some residues in a sequence would form native secondary structure that would act as nucleus from which the native structure would propagate and the tertiary structure would form as a necessary consequence of its secondary structure (Scheraga et al., 1984).

The observations made with chymotrypsin inhibitor 2 led to the formulation of a **nucleation condensation mechanism**, which unifies the different models. It suggests that both hydrophobic interactions and secondary structural elements stabilize the transition state (Fersht, 1997). Another example of a two-state folding mechanism is provided by the bacterial immunity protein Im9 (Ferguson et al., 1999). Interestingly, the analysis of

the structurally highly homologous proteins Im7 and Im9 revealed that small alterations in amino acid sequence can lead to a change in the folding mechanism. Whereas Im9 folds mainly via a two state-mechanism, stabilized folding intermediates were observed for Im7 (Ferguson et al., 1999; Paci et al., 2004). This demonstrated that the folding kinetics depend on the exact polypeptide sequence rather than on the overall stability or tertiary structure of the native protein.

Recent experiments involving sophisticated techniques of stopped-flow coupled with CD and Fluorescence detection, laser induced temperature jump kinetics, Nuclear Magnetic Resonance Spectroscopy (NMR) or Mass spectrometry coupled with H-D exchange paint a more complex picture. In combination with these different approaches, ϕ -value analysis (Fersht et al., 1992) and molecular dynamics simulations (Ferguson and Fersht, 2003; Vendruscolo and Dobson, 2005) have provided detailed insights into folding mechanism in which the behavior of different proteins is quite distinct. This can vary from being two state folders to those involving well-defined intermediates. However, most of the analysis is still limited to relatively small proteins of typically less than 100 aa. Detailed kinetic analysis, monitoring the folding of several small proteins by various techniques, has led to the description of protein folding in terms of an "energy landscape" or "folding funnel" (Zwanzig et al., 1992; Radford et al., 1992a; Dill and Chan, 1997; Paci et al., 2004)(Fig.2.2).

The energy landscape is a three- or multi-dimensional surface representing the free energy of a polypeptide molecule as a function of two or more variables, e.g. conformational parameters. An energy funnel is represented by Enthalpy on its Y-axis and Entropy on its X-axis. Each point on the energy surface represents a conformation of the polypeptide and the corresponding free energy. 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. On the other hand, the denatured protein usually resembles a random coil in which local interactions dominate the conformational behavior, which consequently gives rise to a highly heterogeneous state and forms the top of the funnel. The effect of chaotropic denaturants such as GuHCl or urea is to solvate the hydrocarbons more readily. Hence, the more solvent exposed the state, the greater the reduction in its free energy at a given concentration of denaturant. During folding, the protein follows a route from the top of the funnel, representing a disordered, denatured state, to the bottom of the funnel. Populated intermediates on the way from the unfolded to the native protein are local minima in the energy landscape (Radford, 2000; Schultz, 2000; Troullier et al., 2000). 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 a protein folds is dependent on the physical

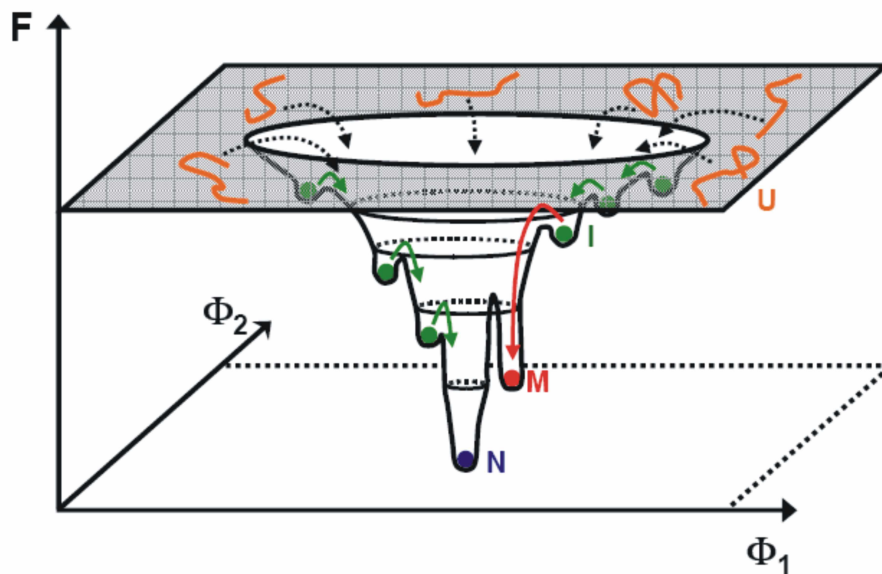


Figure 2.2: Schematic Representation of the Free Energy Landscape of a Folding Protein.

The free energy (E) is displayed as a function of two reaction coordinates ϕ_1 and ϕ_2 which describe the conformation of the folding protein. A multitude of unfolded conformations (U , orange) on the top enter the funnel that contains an almost indefinite number of local energy minima, representing folding intermediates (I , green). The folding protein proceeds then through local minima towards the native conformation (N , blue), which has the lowest free energy. In some local minima, misfolded species may be trapped irreversibly (M , red).

environment in which the folding reaction takes place.

The solvent environment *in vitro* and *in vivo* modulates the stability of local minima and of inter conversion barriers connecting local minima on the free energy landscape. Usually, this environmental modulation relative to simple aqueous solvent is small (a few RT), but the 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. Protein folding is generally much less favored thermodynamically (protein function often requires proteins to be flexible and at the brink of stability), yet folding is fast at room temperature. At its simplest, protein folding can be defined as a directed conformational search which locates the kinetically accessible state of lowest free energy.

One of the most important sources of energetic frustration in the folding pathway is enforced by protein function. Proteins evolve for function, not just for thermodynamic and kinetic *foldability*, and the sequence requirement for function can be incompatible with efficient folding. Function can affect folding in many ways. For example charged or

polar residues and water pockets in the protein core may be required for binding of substrates and prosthetic groups, reducing the core's hydrophobicity, a major driving force for folding. This suggests that modification of a sequence to more secondary/tertiary-structure-friendly side chains could speed up folding at the cost of its function. Some proteins that are constrained by their sequence, topology, size, and function; simply cannot fold by themselves and are instead prone to misfolding and aggregation.

2.2 Protein Folding *in Vivo*

The primacy of the amino acid sequence in determining the final native structure is well established; however, certain properties of the cellular environment are expected to favor protein misfolding and aggregation and thereby, drastically reduce the efficiency of the folding process. High concentrations of solutes and extremes of pH and temperature can also lead to aggregation and misfolding (Shortle, 1996).

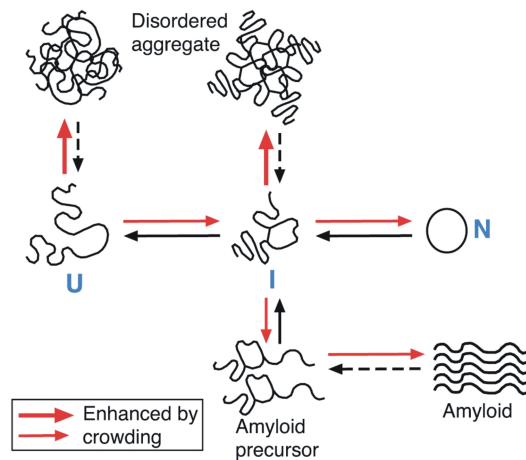


Figure 2.3: Protein Folding *in vivo* Aggregation of nonnative protein chains as a side-reaction of productive folding in the crowded environment of the cell. Enhancement of aggregation and chain compaction by macromolecular crowding (red arrows). U, unfolded protein chain released from ribosome; I, partially folded intermediate; N, native, folded protein. Crowding is predicted to enhance the formation of amyloid fibrils, but this effect has not yet been demonstrated experimentally (from (Hartl, 2002)).

There can be numerous reasons for proteins to be misfolded and thus be recruited into aggregates *in vivo* (Fig.2.3). Firstly, the folding of the protein in the cell is linked to its biosynthesis, which is an obligatorily vectorial process (N terminus to C terminus). As a result, the information for the folding process becomes available sequentially and not all at once as in the case of *in vitro* refolding. This incomplete availability of structural information during translation may not hinder the secondary structure formation but it greatly limits the formation of native tertiary structure. Secondly, the approx. 30 odd C-terminal residues in the exit tunnel of the ribosome are topologically constrained and cannot participate in folding reaction until they are released.

In multi-domain proteins, a complete domain can fold in principle as the rate of domain folding *in vitro* is generally much faster (ranging from seconds or less) than the rate of biosynthesis (4 – 20 amino acids per second). It has also been noted in Eubacteria

that the rapid rate of translation (approx. 5 – 10 times faster than in higher eucaryotes) (Dobson and Karplus, 1999) would favor a post translational mechanism for folding at least for those proteins that fold slowly (Netzer and Hartl, 1997). Even a folded domain, however, may be unstable because of the absence of interacting domains and these partially folded unstable structures often expose hydrophobic residues that are otherwise buried in the native states. This makes these intermediates highly aggregation prone. The close proximity of such nascent chains due to the translation on polysomes could also result in intermolecular entanglement. The highly crowded nature of the cellular milieu (approx. 300 g/l of protein and other macromolecules) presents another obstacle for efficient protein folding (Zimmerman and Minton, 1993).

Crowding can cause excluded volume effects, which can result in increased intermolecular association of unfolded polypeptides (Ellis and Hartl, 1996). This endangers even those polypeptides, which have escaped misfolding during biosynthesis. The cellular machinery of chaperones is designed to effectively counteract the tendency of non-native polypeptide chains to aggregate under the conditions prevailing in the cell. 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 interaction between histones and DNA (Ellis and Hartl, 1996). This specific usage was later generalized to include a range of functionally related, but diverse, proteins that assist the folding and assembly of other proteins.

Initially, the emphasis was on mediation of protein assembly but subsequent functional studies with the mitochondrial chaperonin Hsp60 and eubacterial homologue GroEL, showed the primary role to be in facilitating protein folding (Goloubinoff et al., 1989). These studies provided the experimental basis for a deviation from the view that protein folding is generally a spontaneous process, in strong support of the chaperone concept. Our current definition of a *molecular chaperone* is that of a protein which transiently binds to and stabilizes an unstable conformer of another protein, and through regulated binding and release (which may or may not be ATP dependent), facilitates its correct fate *in vivo*. Its function can be folding (following *de novo* synthesis, transit across a membrane, or stress induced denaturation), oligomeric assembly, interaction with other cellular components, switching between active and inactive conformations, intracellular transport, or proteolytic degradation, either singly or with the help of co-factors (Fink, 1999).

During translation, the majority of polypeptide chains may depend on molecular chaperones for protection (Hartl, 2002). Once they are released from the ribosome into the cytosol, chaperones assistance may be restricted to slow folding proteins. Chap-

erones do not contribute steric information to the folding process, which distinguishes them from the folding catalysts such as peptidyl-prolyl-isomerases (PPIases) and protein disulfide isomerases (PDIs). These proteins catalyze specific reactions such as prolyl isomerisation and disulfide bond formation (Schmid, 1993; Freedman et al., 1994; Puig et al., 1994).

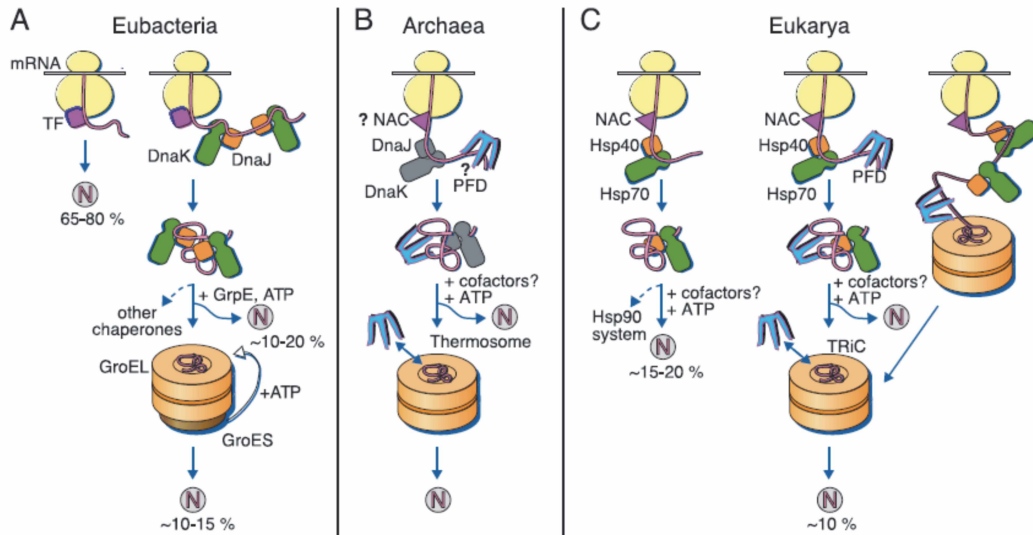


Figure 2.4: Model for de novo Protein Folding Assisted by a Network of Molecular Chaperones in the Cytosol of Bacteria, Archaea and Eukarya. N: Natively folded protein, TF: trigger factor, NAC: nascent chain-associated complex, PFD: prefoldin. (A) Many proteins in the bacterial cytosol fold without further assistance upon release from the ribosome and ribosome-bound TF. DnaK assists the remainder of proteins in folding, and can transfer substrates to the chaperonin system (GroEL/GroES). (B) Only some archaea contain DnaK/DnaJ. Interaction of PFD with nascent chains and existence of NAC is not experimentally confirmed. (C) In the example of the mammalian cytosol, NAC probably interacts with nascent polypeptide chains together with Hsp70 and Hsp40. The majority of proteins can fold upon release from these factors. A subset of Hsp70 substrates is transferred to the Hsp90 system. Furthermore, PFD interacts with nascent chains and transfers these to TRiC, the eukaryotic chaperonin (Hartl, 2002).

This problem of folding *in vivo* is so deeply entrenched in living systems that molecular chaperones have evolved and are distributed ubiquitously across all the three kingdoms of life. They act in the cell at all temperatures but the levels of many are greatly upregulated under stress conditions (Fig.2.4). Therefore, molecular chaperones are also known as heat shock proteins (Hsps). Their respective molecular weight determines their names, e.g. Hsp104, Hsp70, Hsp40, Hsp60, Hsp10. More than 20 different families of chaperone are currently known that have been conserved in evolution. Approximately one-fourth of these are stress inducible and many of them are essential under normal conditions of growth. The prominent chaperone families include the Hsp70s, Hsp40s

(DnaJ), the chaperonins, Hsp90 and the small heat shock proteins. Cells sometimes need to protect their proteins against many different stress conditions such as the denaturing influence of heat etc. with the help of these chaperones or heat shock proteins. These assist other proteins in not just folding but also in maintaining their folded states. Some proteins never fold in cells at all except with the assistance of chaperone molecules that either isolate individual proteins so that their folding is not interrupted by interactions with other proteins or help to unfold misfolded proteins, giving them a second chance to refold properly.

In general, chaperones do not actively fold their substrate proteins; they rather create a local environment favoring productive protein folding over functionally non-productive side reactions. Binding and release of substrate polypeptides by chaperones is often achieved by ATP-driven conformational changes, allowing multiple rounds of binding and rebinding between substrate and chaperone machinery, until a native structure is achieved (Fink, 1999).

2.2.1 Ribosome-Associated Chaperones

Polypeptides are generated and released into the cytosol sequentially from the ribosome and therefore expose large unstructured and hydrophobic regions during their synthesis. In order to prevent aggregation of partly completed polypeptides, ribosome associated chaperones reversibly bind to aggregation-prone nascent polypeptide chains at the ribosomal exit tunnel (Hartl, 2002). The first chaperone that interacts with a nascent chain during their synthesis at the ribosome is trigger factor (TF). It is associated with the ribosome itself. The 48 kDa *E. coli* protein binds to a docking site at protein L23 of the large ribosomal subunit (Kramer et al., 2004a). TF is thought to scan the nascent polypeptide as it emerges from the ribosomal exit tunnel for hydrophobic regions and binds to these regions as they are encountered. The TF reaction is not ATPase driven (Hesterkamp and Bukau, 1996). TF also exhibits peptidyl prolyl cis-trans isomerase activity, but the biological relevance of this activity for protein folding is still unclear, since it is not essential for the function of TF *in vivo* (Genevaux et al., 2004; Kramer et al., 2004b). It has been shown using fluorescence spectroscopy that TF interacts with ribosomes and translating polypeptides in a dynamic reaction cycle involving all three functional domains. Binding to the ribosome conformationally activates TF for nascent-chain association. Activated TF departs from the ribosome but may remain bound to the elongating polypeptide thus preventing aggregation and unspecific inter-molecular interactions (Kaiser et al., 2006).

2.2.2 Hsp70 Chaperones

The Hsp70 system constitutes a central part of the molecular chaperone arsenal of the cell. The common mode of DnaK action, the *E. coli* homologue of Hsp70, appears to be binding to short, extended hydrophobic peptide sequences in the substrate proteins with an ATP-regulated and ligand induced change in affinity for binding and release (Liberek et al., 1991). By shielding exposed hydrophobic surfaces, Hsp70 chaperones prevent further folding and aggregation of bound substrate proteins for the time they are bound. Native proteins do not usually expose such hydrophobic fragments and are thus not recognized by DnaK. DnaK is active as a monomer of 70 kDa and is comprised of two functional domains: a 45 kDa amino-terminal ATPase domain and a 25 kDa carboxy-terminal polypeptide binding domain whose structures have both been solved by X-ray crystallography independently (Harrison et al., 1997; Zhu et al., 1996) but not in an intact DnaK molecule. Communication between the two domains in the functional cycle results in efficient binding and release of substrate polypeptides. Recently, structure of a C-terminal truncated but a functionally intact Hsp70 was resolved that provides insights into the allosteric mechanisms of the Hsp70 chaperone system. A mutational analysis of the observed interdomain interface and the immediately adjacent inter-domain linker identifies inter-domain interactions critical for chaperone function and supports an allosteric mechanism in which the inter-domain linker invades and disrupts the inter-domain interface when ATP binds (Jiang et al., 2005).

2.2.3 Hsp60 and Hsp10: The Chaperonins

The chaperonins constitute a conserved class of essential gene products encoded in the genome of almost every organism sequenced to date, distributed among eukaryotes, archaea and prokaryotic organisms (Fayet et al., 1989; Knapp et al., 1994; Ostermann et al., 1989). Chaperonins are large, multimeric, nearly 1 MDa complexes with a double-ring structure, forming two central cavities. They are divided into two groups, which are related in topology but do not share close sequence similarity. Group I chaperonins occur in the bacterial cytosol (GroEL) and in eukaryotic organelles of bacterial endosymbiotic origin (Cpn60 in chloroplasts, and Hsp60 or Cpn60 in mitochondria). They have a seven-fold symmetry. Group I chaperonins function in cooperation with cofactors of the Hsp10 family (GroES in bacteria, Hsp10 or Cpn10 in mitochondria and chloroplasts). Group II chaperonins occur in archaea and the eukaryotic cytosol. The archaeal chaperonin is called thermosome (Maeder et al., 2005; Hartl, 2002) and the eukaryotic homolog is called either TRiC (TCP1 Ring Complex) or CCT (Chaperonin Containing T-complex protein 1) (Stoldt et al., 1996). Group II chaperonins do not interact with Hsp10-like cofactors, but the function provided by this factor is thought to be directly embedded into the structure of group II chaperonins themselves (Macario et al., 2004). TRiC has

an eight-fold symmetry. The detailed introduction to structure and function of chaperonins (explained later in the section) is limited to the class I chaperonin homologues of *E. coli*, GroEL and GroES.

2.2.4 Additional Chaperone Systems

In addition to the cytosolic chaperone systems described above, a large number of other cellular factors assist in the folding of newly-synthesized or stress-denatured proteins, often in co-operation with the Hsp70 system or the chaperonins. Compartments other than the cytosol, such as the endoplasmic reticulum in eukaryotes or the periplasm in bacteria, harbor their own unique inventory of specialized molecular chaperones. These will not be discussed further here.

Small heat-shock proteins, which are often found associated with inclusion bodies in *E. coli*, are stress-inducible molecular chaperones that bind unfolded proteins, preventing their aggregation and facilitating their refolding by ATP-dependent chaperones (Ehrnsperger et al., 1997; Studer et al., 2002; van Montfort et al., 2001). The eubacterial proteins IbpA and IbpB are members of this class of proteins. Small heat shock proteins of the eukaryotic cytosol include Hsp12 and Hsp42 as well as the mammalian α -crystallins.

The Clp (Hsp100) proteins constitute a subfamily of AAA proteins (ATPases Associated with various cellular Activities). They participate in the re-solubilization of aggregated proteins in cooperation with the DnaK chaperone system or the protein degradation machinery (Glover and Lindquist, 1998; Mogk and Bukau, 2004; Schirmer et al., 1996; Weibezahn et al., 2005). Prefoldin, also called Gim (genes involved in microtubule biogenesis) complex or GimC (Vainberg et al., 1998), is a hetero-oligomeric complex of two α and four β subunits that is found in the cytosol of eukaryotes and archaea. Its structure resembles the shape of a jellyfish with six α -helical coiled-coil structures protruding from a β -barrel body (Siegert et al., 2000). GimC is able to stabilize non-native proteins, both co- and post-translationally, and transfer them to the chaperonin (Leroux et al., 1999).

Eukaryotic Hsp90 is a member of the ATP-dependent, homo-dimeric Hsp90 family of chaperones (Young et al., 2004). Similar to the Hsp70-system, the ATPase activity of Hsp90 regulates substrate binding and release, but the biochemical mechanism of Hsp90 is not yet fully understood. In contrast to chaperones of the Hsp70 class, it appears to act at a later stage of the folding process on more compact folding intermediates of a variety of proteins, including transcription factors, regulatory kinases and other signaling and structural proteins (Pratt and Toft, 2003; Richter and Buchner, 2001; Young et al.,

2001). Hsp90 often works together with the cytosolic Hsp70 family member Hsc70, forming the Hsc70. Hsp90 machinery is regulated by diverse co-factor proteins. A feature often found in these co-factors is a tetratricopeptide repeat (TPR) clamp domain that mediates interaction with Hsp90, Hsc70, or both. An example of such a co-chaperone is the mammalian Hsp-organizing protein (HOP), which has two independent TPR clamp domains and co-ordinates the action of Hsc70 and Hsp90 (Young et al., 2004; Young and Hartl, 2003). The wide variety of co-factor proteins allows the Hsc70-Hsp90 system to operate in diverse functional contexts (Young and Hartl, 2003).

2.2.5 Cellular Chaperone Networks

It is highly probable that the multiplicity of intracellular chaperones necessitates a high degree of functional coordination. One proposed model called the **pathway model** dictates that the Hsp70 and Hsp40 chaperones- the "upstream" chaperones (and the trigger factor in *E.coli*) bind co-translationally to the majority of newly synthesized proteins and then relay a subset of those (the slow folders) to the chaperonins or other "downstream" chaperone system, which in turn complete the folding process and releases the protein in the cytosol. The proteins unable to reach the native state may be recaptured by Hsp70s and targeted for degradation.

Another alternative model is called the **network model**, which allows all chaperones to compete for substrate proteins at all times. The non-native states of substrates are frequently released into the cytosol and captured by any chaperone based on law of mass action and individual affinities.

Chaperones share the ability to transiently associate with non-native conformers of proteins by recognizing exposed hydrophobic patches. There are, however, differences with respect to the molecular mechanism of substrate recognition. E.g. Hsp70, in functional cooperation with DnaJ co-chaperones, binds to short stretches of consecutive hydrophobic residues in a polypeptide and Hsp60 recognizes hydrophobic patches in secondary structure elements. The small heat shock proteins (sHsps) form oligomers with an average size of 12 to 42 subunits. Each oligomer can bind several protein substrates, up to one molecule per subunit, and thus serves as a very efficient scaffold for misfolded or unfolded substrates

2.3 GroE Chaperone System

2.3.1 Structural Information of GroEL-GroES Machinery

GroEL is made up of two oligomeric rings stacked back to back with a cavity at each end that provides a protective environment for protein folding. Each ring is made of seven identical subunits of 60kDa. Each subunit consists of three domains: *the equatorial domain*, that contains an ATP binding domain and most of the inter and intra ring contacts; *the apical domain*, which forms the opening of the central cavity and contains the substrate binding site that binds non native polypeptides through hydrophobic interactions and an intermediate *hinge-like domain* connecting these two domains. The N terminus begins with the equatorial domain and then continues through the intermediate domain. The central part of the sequence forms the apical domain and subsequently the chain returns down through the intermediate domain. Both N and C termini (a total of 30 residues are not resolved in the crystal structure) face the central channel, which is continuous through the 14-mer in the crystal structure (Fig.2.5).

However, the missing 30 residues per subunit appear to form a central constriction to

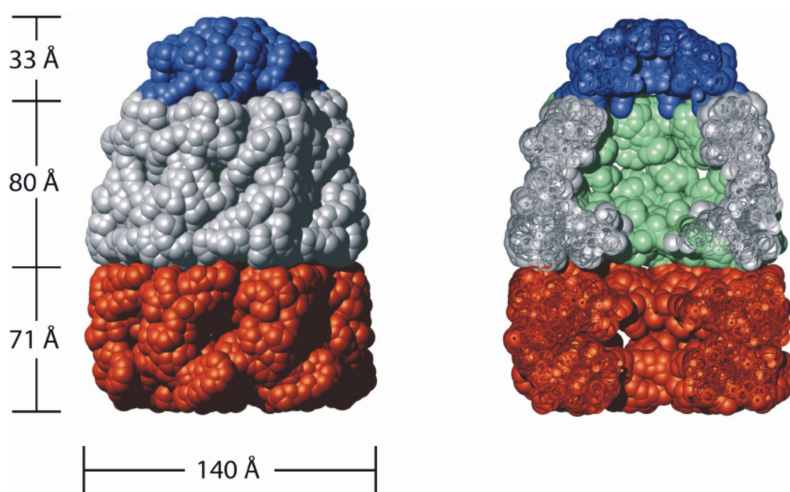


Figure 2.5: Asymmetric Structure of the GroEL/GroES Complex. Space-filling models of GroEL/GroES (PDB 1AON; (Xu et al., 1997) with 6Å Van der Waals spheres around C_{α} atoms. The two rings of GroEL are red and grey, GroES is shown in blue. Outside view (left) and inside view (right) of GroEL/GroES, generated by slicing the structure with a vertical plane through the heptameric symmetry axis. To indicate the interior of the cis cavity, subunits located at the back of GroEL are colored in green.

the channel as seen in cryo EM images (Chen et al., 1994) (Elad et al., 2007). Mutation of Lys4Glu (sequence beginning at Met1) completely excludes oligomer assembly (Horovitz et al., 1993b), suggesting an important structural role for this apparently disordered re-

gion. The equatorial domain, containing N and C terminal parts of the chain, forms the backbone of the oligomeric structure. It accounts for most of the inter and intra-subunit contacts, the interaction between adjacent subunits, and also the only contacts between the 2 rings in the 14-mer, an important route of allosteric communication (Aharoni and Horovitz, 1996; Horovitz et al., 1994).

Much of the apical domain surface is not symmetric and closely spaced with neighboring domains. Thus, there is little steric hinderance to its movement, both locally in some parts and in the overall orientation of the domain. This part of the structure deviates most from 7-fold symmetry in the crystal.

The helix bundle in the equatorial domain contains the ATP binding site. This pocket is bordered by a highly conserved sequence motif containing Asp87, GDGTT, and lined by other stretches of highly conserved residues (Fenton et al., 1994). Mutations in Asp87 excludes the binding of ATP and completely abolishes ATPase activity.

The ATP-binding pocket is adjacent to the lower hinge region and mutations in the intermediate domain just beyond this hinge region also abolish ATPase activity (residues 150, 151, 152, 405 and 406), as does mutation of residue 383, near the upper hinge region (Fenton et al., 1994). These strategic locations of key residues around the hinge regions strongly suggest that hinge movements are involved in the hydrolysis mechanism.

The extensive mutagenesis analysis identified a set of hydrophobic residues on the apical domain, along the inner channel. This region appears to be flexible in the GroEL structure. Single amino acid substitutions, at some of these residues (e.g. residues 199, 203, 204, 234, 237, 259, 263, and 264), abolish substrate binding (Fenton et al., 1994). The ring of 7 binding sites is easily accessible from outside, being very near the surface of the cylinder, but the inward-facing orientation may protect GroEL from self-aggregation. Mutation of residue 152, near the lower hinge region, also has a strong effect on substrate binding (Fenton et al., 1994). The ability of monomeric cpn60, missing 78 N-terminal residues, to bind substrates and partially promote their folding is consistent with the apical location of substrate binding sites (Horovitz et al., 1993b; Horovitz et al., 1993a).

GroEL works in conjunction with its co-chaperonin GroES which functions as a lid . GroES consists of seven identical 10 kDa subunits arranged in the form of a ring (Braig et al., 1994). Each subunit consists of a barrel region that forms most of the contacts around the ring and a hairpin pointing slightly upwards and towards the center of the ring. This region is loosely packed, with little inter-subunit contact, and forms the roof of the dome-like structure of the oligomer. Earlier NMR work showed that a mobile

domain in GroES became ordered upon binding to GroEL (Shewmaker et al., 2001). Upon being capped by GroES, the cavity undergoes an increase in size and the physical properties of the walls change from hydrophobic to hydrophilic in character. The mobile region of GroES (residues 17 – 32) not seen on the X-ray structure of GroES, is likely to form at least part of the binding contact. When in complex with GroEL, the mobile loop of GroES is probably in contact with GroEL, because its accessibility to trypsin is reduced in the complex and a synthetic peptide with the loop sequence binds to GroEL (Landry et al., 1993). Many of the mutations that preclude substrate binding also interfere with GroES binding, suggesting that the binding sites for GroES and substrate overlap (Fenton et al., 1994). In addition, the GroES binding surface appears to extend further over the top surface of the apical domain.

2.3.2 Effect of Nucleotides

ATP binding precedes the binding of GroES in the GroEL reaction cycle and results in a 5-10 degree clock-wise turning motion of the apical GroEL domains (Chen et al., 1994). ATP binding results in a decrease in the affinity for substrate of GroEL and ATP hydrolysis drives the key functional cycle of GroEL, regulating between the substrate acceptor and release states. In the ADP-bound form or without nucleotide, GroEL has a high affinity for the unfolded substrate, and vice-versa in the ATP-bound form (Staniforth et al., 1994). Alternation between these states has been proposed as the basis of the assisted folding mechanism. These two states have different conformations in cryo EM images, with an opening of the apical domains in the presence of ATP. The apical domains open out, elongating the cylinder and widening the binding cavity. ATP binding makes the hinge regions more flexible. Non-hydrolysable analogues of ATP appear to cause the same type of cavity opening as ATP but ADP causes a more subtle change (Langer et al., 1992). ATP is shown to induce asymmetry of the GroEL 14-mer (Bochkareva et al., 1994).

There is a positive cooperativity exhibited in ATP binding between intra-ring subunits and a negative cooperativity between inter-ring subunits of GroEL (Yifrach and Horovitz, 1996). How these rigid body movements affects conformation of the substrate bound to the GroEL and how this in turn contribute to accelerated folding of the substrate protein is not well correlated. Although there are a lot of studies exploring the effect of nucleotide on the conformational changes of GroEL and GroEL substrate interactions, very little is actually available which mechanistically correlates these two aspects into the folding scheme.

2.3.3 Substrates of GroEL-GroES

10 - 15 % of all cytosolic proteins in *E. coli* were observed to transit through GroEL *in vivo* (Ewalt et al., 1997; Houry et al., 1999). Despite this relatively low number, which is based on the nature of the molecular chaperone network inside the cell (described before), GroEL is able to bind to a large and diverse variety of non-native proteins *in vitro* (Viitanen et al., 1992). In a high-throughput approach approximately 250 proteins were

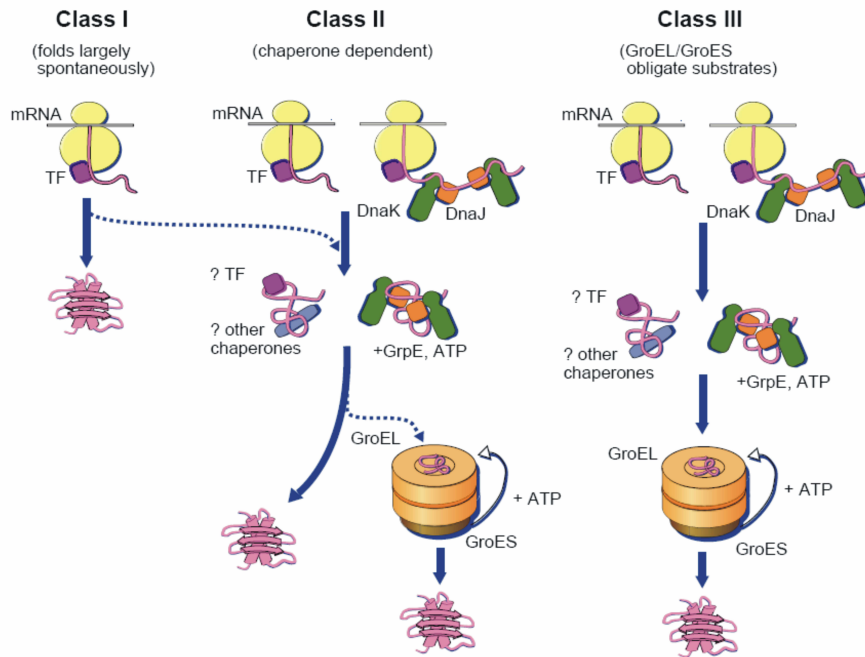


Figure 2.6: Chaperone usage of nascent polypeptides in *E. coli* upon synthesis on the ribosome. This model distinguishes three classes of chaperonin dependence. Modified from (Hartl and Hayer-Hartl, 2002).

identified that were found to functionally interact with GroEL (Kerner et al., 2005). The selected proteins could be grouped in three classes based on their increasing requirement for GroEL (Fig.2.6).

- Class I substrates exhibit a low propensity to aggregate upon dilution from denaturant and, consequently, only a partial chaperone requirement for refolding *in vitro* e.g. Enolase (46 kDa) and glyceraldehyde-3-phosphate dehydrogenase (35 kDa).
- In contrast, Class II proteins do not refold spontaneously under standard conditions due to their rapid aggregation. They include glutamate decarboxylase α (53 kDa) and galactitol-1-phosphate 5-dehydrogenase (37 kDa). The presence of both GroEL and GroES, but not GroEL alone, was shown to be necessary in assisting the refolding of these proteins at 37°C. Although GroES was shown to be not absolutely

required for refolding at 25°C (Kerner et al., 2005), suggesting that these proteins are not obligate GroEL/GroES substrates. The DnaK system was as efficient in mediating refolding at 37°C. Thus, it is likely that DnaK and GroEL share a number of substrates mainly in the preferred size range of GroEL (up to 60 kDa), whereas larger proteins may generally be more adapted for folding by the DnaK system.

- Class III proteins are found to be stringently chaperonin dependent, e.g. 5,10-methylenetetrahydrofolate reductase (MetF-33 kDa), S-adenosyl methionine synthetase (MetK-42 kDa), and DAPA (31 kDa). While METK and METF failed to refold spontaneously under a variety of conditions known to reduce aggregation, slow but efficient spontaneous refolding was observed for DAPA in the presence of 0.5 M arginine.

2.3.4 Maltose Binding Protein - Model Substrate

Proteins with an obligate GroEL dependence typically aggregate upon in vitro refolding (Kerner et al., 2005), and thus it is difficult to compare their spontaneous and chaperonin-assisted folding rates. To avoid this complication, suitability of maltose binding protein (MBP) as a model substrate was established based on previous reports that GroEL/GroES can increase the folding speed of a mutant form of MBP (Sparrer and Buchner, 1997) has been well established (Tang et al., 2006). MBP is a monomeric 41 kDa periplasmic protein that folds robustly in the cytosol when expressed without its cleavable N-terminal export sequence.

It is composed of two globular domains formed by discontinuous sequence elements consisting of secondary structural $\beta\alpha\beta$ units with the binding site for maltose located in a cleft between the domains (Fig.2.7)(Spurlino et al., 1991).

Several slow-folding mutants of MBP are known, for instance, the single mutant Y283D (SM-MBP) and the double mutant V8G/Y283D (DM-MBP) (Wang et al., 1998). Mutations V8G and Y283D are located in close proximity in a strand and loop segment, respectively, of the N-domain. Formation of native contacts within the N-domain is hypothesized to be rate-limiting for folding and is further slowed by these mutations (Chun et al., 1993). MBP possesses eight tryptophans distributed over both domains (Fig.2.7). Their fluorescence signal is reduced about 5-fold upon unfolding, and the recovery of fluorescence can be used as a measure of folding (Chun et al., 1993) both in the presence and absence of GroEL/GroES, which lack tryptophan residues. Importantly, GroEL-GroES is shown to accelerate the rate of refolding of SM-MBP by about 3-fold and that of DM-MBP about 13-fold compared to spontaneous folding (Tang et al., 2006; Sparrer and Buchner, 1997).

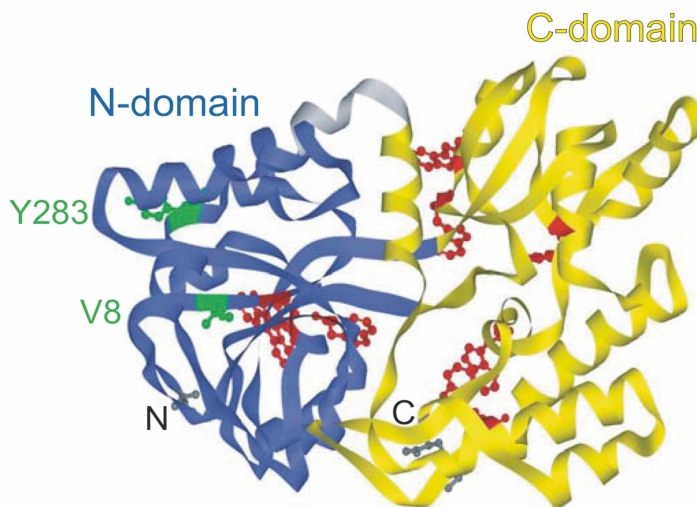


Figure 2.7: Ribbon diagram of the structure of MBP (Spurlino et al., 1991; pdb 1OMP; DS Viewer- Pro), indicating the positions of mutated amino acids (green). The two discontinuous domains are shown in blue and yellow, respectively; the eight tryptophans are shown in red. (Tang et al, 2006).

In contrast to GroEL/GroES, the bacterial Hsp70 chaperone system, consisting of DnaK (Hsp70), DnaJ, GrpE, and ATP, strongly retards the folding of SM-MBP and DM-MBP. Very similar properties were recently described for several highly aggregation sensitive, authentic GroEL substrates (Kerner et al., 2005).

Folding rates for this protein and yields are essentially concentration-independent between 50 nM and 1 μ M for wt-MBP (Ganesh et al., 2001) and for the two mutant proteins (Tang et al., 2006), arguing against reversible aggregation as the cause of slow spontaneous folding of mutant MBP.

Single event of encapsulation of mutant MBP in the GroEL-GroES cage is sufficient for accelerated folding, as seen from refolding experiments that are carried out with the non-cycling single-ring mutant of GroEL (SR-EL), which binds and encapsulates unfolded protein in a GroES - and ATP-dependent reaction but does not release GroES (Hayer-Hartl et al., 1996; Weissman et al., 1996). SR-EL/GroES in the presence of ATP fully reproduces the rate acceleration of SM-MBP and DM-MBP folding observed with the cycling GroEL/GroES system. Thus, DM-MBP can be considered to be a GroEL dependent substrate as is also shown *in vivo*.

2.3.5 Conformational Properties of Chaperonin Bound Substrates

Molecular chaperones are able to discriminate between folded and unfolded proteins. One of the most obvious features of unfolded proteins is the exposure of hydrophobic

residues unlike the native or folded state. Very little is known about the substrate binding specificity of chaperonins. Employing a GroEL molecule comprised of two rings, each produced as a single continuous polypeptide (Farr et al., 2000), it was possible to determine more closely the effects of mutations that had previously been observed to abolish polypeptide binding (Fenton et al., 1994). Three consecutive wild-type apical domains were found to be necessary to facilitate stable binding of the substrates rhodanese and malate dehydrogenase *in vitro*. This arrangement was also found to be necessary to allow cell viability.

Moreover in another study folding of 30% of cytosolic proteins was found to be affected in a temperature sensitive groEL mutant strain (Horwich et al., 1993). GroEL does not recognize short peptides or long peptides in extended conformations with high affinity. Conformational analysis by NMR has shown that a 13mer peptide derived from rhodanese was weakly bound and stabilized in an α -helical conformation (Landry et al., 1992; Landry and Gierasch, 1991). Since GroEL also binds β -sheet structures it seems likely that it binds structural motifs other than secondary structure elements, possibly the exposure of hydrophobic residues or surfaces.

Negative charges are found to be largely missing from peptides that display high affinity. Despite the promiscuity in substrate binding there appears to be some underlying features pertinent to substrate chaperonin interaction. It has been proposed using FCS measurements (Pack et al., 2000) the apo-cyt c a positively charged substrate, has a much stronger affinity for GroEL than do denatured pepsin and reduced Lactalbumin (rLA), which are negatively charged substrates. Moreover, substrate binding is influenced by cations in the solution and divalent cations are more effective than monovalent cations depending on substrates.

This enhancement of binding due to positively charged residues could indicate the importance of ionic interactions at some stage in the binding process, or might simply reflect the fact that positively charged (as well as hydrophobic) residues can interact with aromatic side-chains (Burley and Petsko, 1986), such as in the phenylalanine and tyrosine residues involved in polypeptide binding on the inner surface of the apical domain of GroEL (Fenton et al., 1994). Such interactions are known as cation- π interaction. It is a non-covalent molecular interaction between the electron-rich π orbitals of an aromatic ring with adjacent cation. Such interactions are relatively strong, being roughly equivalent in energy to that of a hydrogen bond (Gallivan and Dougherty, 1999).

The main promiscuity of substrate binding by GroEL has been attributed to the plasticity of its hydrophobic binding sites, which results in their association with a multitude

of substrates (Chen and Sigler, 1999). A high affinity peptide ('strongly binding peptide', SBP) was identified and crystallized in combination with GroEL apical domains and the full tetradecameric complex. The binding site for this peptide was found to be a flexible hydrophobic groove on the apical domains of GroEL, lining the openings of the cavity. The site was overlapping with the region that accommodates the mobile loop of GroES when bound to GroEL (Chen and Sigler, 1999). Other peptides which have been characterized in their binding to GroEL correspond to amphiphilic α -helical regions of rhodanese (Hlodan et al., 1995).

A number of studies have indirectly attempted to define the conformation of the substrate protein while bound by chaperonin. The binding of substrate protein to GroEL is usually performed by rapidly diluting the protein from denaturant to the GroEL containing buffer. Upon dilution into the buffer protein undergo a collapse and form secondary structure elements on a micro - to millisecond timescale. However, the binding of substrate to GroEL happens on a much slower time scale (seconds), suggesting that the protein conformation recognized by chaperonin is likely to contain secondary structure.

Substrate proteins for which various conformational parameters have been analyzed while bound to GroEL include rhodanese (Martin et al., 1991), Dihydro-folate-reductase (DHFR) (Martin et al., 1991), RuBisCo (van der Vies et al., 1992), β -glucosidase, Pre- β -lactamase, and α -lactalbumin (Hayer-Hartl et al., 1994). Common features of these chaperonin bound proteins are high protease sensitivity, tryptophan fluorescence, intermediate to that between the native and the completely unfolded state, partial exposure of tryptophan residues to the solvent as seen from the emission maxima and significant adsorption of the hydrophobic fluorescent dye anilino-naphthalene sulfonate (ANS).

On the basis of these observations it has been proposed that the conformation of a chaperonin bound polypeptide resembles that of the so-called molten globule, a compact folding intermediate that contains secondary structure but lacks stable tertiary interactions. The problem with such optical probes is the uncertainty in interpreting the signal. The fluorescence of tryptophan may be shifted in wavelength or enhanced by interaction with the hydrophobic sites on the GroEL, and ANS binding sites may be formed by pockets in the GroEL: substrate complex rather than those exclusively in the bound protein. A major goal of physical experiments on chaperonins is to understand what happens to the conformational properties of folding proteins during their encounters with the binding surface of GroEL. This information then may lead to the insights regarding how protein and nucleotide binding energies are used to more efficient folding of the substrates in terms of rate and yield.

A crystal structure of GroEL with a complete bound substrate protein bound remains to be obtained. Problems associated with this are the large conformational variety of even a single type of bound protein, and the heptameric rotational symmetry of GroEL, which would decrease the abundance of even a single state of a bound substrate to approx. 15 % in a crystal. Techniques for following spontaneous folding are relatively better worked upon and reasonably well developed such as time resolved CD spectroscopy, amide hydrogen-deuterium exchange coupled to mass spectroscopy and Nuclear Magnetic Resonance. All of these techniques become impossible or extremely difficult when chaperonins are introduced into the equation.

2.3.6 GroEL-GroES Reaction Cycle in Protein Folding

GroEL facilitates protein folding by undergoing rigid body movements that are coordinated in space and time by complex allosteric regulations by ATP binding and hydrolysis. This serves as a signal for substrate and GroES binding and subsequent release of the substrate into solution, as well as a clock for the time of encapsulation of the substrate inside the cavity.

Three affinity states of GroEL for substrate binding can be distinguished: the nu-

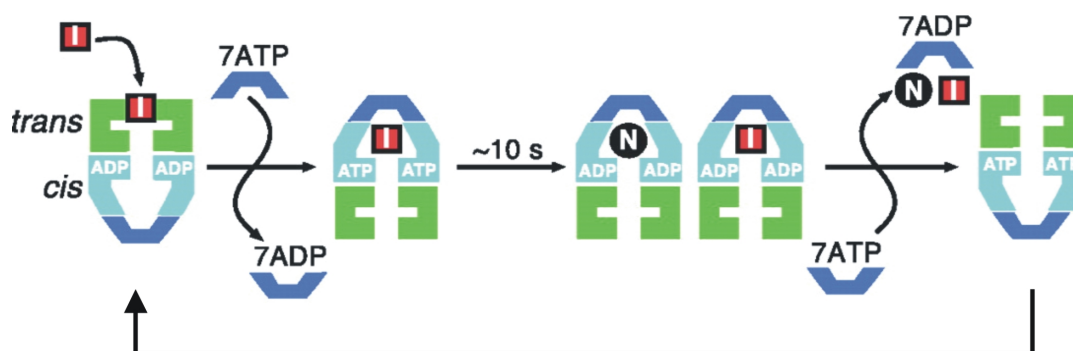


Figure 2.8: Reaction Cycle of GroEL-GroES Simplified reaction of protein folding in the GroEL-GroES cage. I, folding intermediate bound by the apical domains of GroEL; N, native protein folded inside the cage. For a typical GroEL substrate, multiple rounds of chaperonin action are required for folding; both I and N accumulate after a single reaction cycle and exit the cage upon GroES dissociation. I is then rapidly re-bound by GroEL.

cleotide free form, which appears to have the highest affinity for the unfolded protein; ATP bound form, having lowest affinity for the unfolded protein and the ADP bound form, which associates with the substrate protein with an affinity closer to that of the nucleotide free state than the ATP bound state. These forms probably represent different conformational states of GroEL that are inter-convertible by nucleotide exchange

and ATP hydrolysis (Fig.2.8). The following sequence of steps can be proposed for the GroEL-GroES reaction cycle:

1. Under physiological conditions, GroEL is in the closed form the physiological acceptor state of the GroEL for the protein (Rye et al., 1999). The "GroEL closed" is ADP bound GroEL which is in a complex with GroES and the trans ring in the substrate acceptor state.
2. Unfolded protein binds to the trans ring of GroEL and triggers ADP dissociation in the cis ring, which in turn results in the release of GroES.
3. The removal of ADP allows binding of ATP which was regulated by the negative co-operativity between the two rings. ATP binding reduces substrate affinity for the bound substrate protein.
4. ATP binding triggers conformational changes which enables GroES binding to GroEL in the ATP state and may cover the ring that contains the bound substrate. Cooperative ATP hydrolysis releases the substrate protein for folding in the ring cavity.
5. GroES binding becomes stabilized in the regained ADP state and partially folded protein may re associate for another round of interaction.

The underlying principle of GroEL mediated protein-folding hinges on alternating cycles of binding and release of substrate protein and GroES until the protein has reached native state and has buried its hydrophobic patches and thus lost its affinity for GroEL. Association of GroES with only one of the GroEL rings imparts the asymmetry required for this process and ensures cooperative inter-ring ATP hydrolysis.

2.4 Fluorescence

Fluorescence generally is much more sensitive to the environment of the chromophore than absorption or CD spectroscopy. During the approximate lifetime (10^{-9} sec) of the fluorophore all 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 lifetime and quantum yield are the most important characteristics of a fluorophore.

The quantum yield is the fraction of excited singlets that decay by fluorescence or the number of emitted photons relative to the number of absorbed photons. Although the quantum yield can be close to unity if the radiationless decay rate is much smaller than the rate of radiative decay; the energy yield of fluorescence is always less than unity because of Stokes' losses (Fig.2.9).

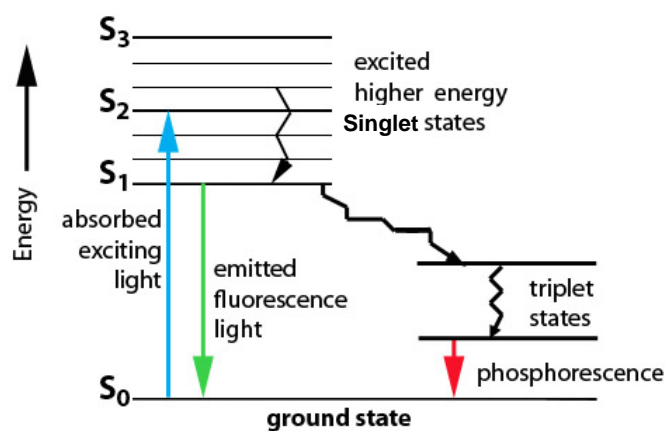


Figure 2.9: Jablonski Diagram It shows a number of possible routes by which an excited molecule can return to its ground state. A rapid return (in ns) results in fluorescence and a delayed return (in μ s) results in phosphorescence where relaxation happens via the triplet state.

The lifetime of the excited state is defined by the average time the molecule spends in the excited state prior to return to the ground state. Certain other factors also affect the intensity of fluorescence (although these need not necessarily directly related to intensity) such as:

1. Internal conversion: the process in which excitation energy is lost by collision with solvent or by dissipation through internal vibrational modes. In general, the process increases with increase in temperature. This should be particularly noted when using this method to monitor thermally induced macromolecular conformational changes.
2. Deexcitation: resulting from collisions or interactions with solute molecules capable of quenching the excited state.
3. Intersystem crossing: the process in which the nominally forbidden spin exchange converts an excited singlet into an excited triplet state, which in turn converts to the ground state either by phosphorescence (emission of a photon) or by internal conversion. The triplet state is generally lower in energy than the excited singlet. Hence phosphorescence occurs at longer wavelengths and can be easily resolved from fluorescence.

A number of fluorescent molecules have a property of being strongly quenched in aqueous environment but are strongly enhanced (over 20 fold) in a non-polar or a rigid environment. Fluorescence lifetime determines the time available for the fluorophore in its excited state to interact with other molecules or diffuse in its environment.

2.4.1 FRET

Analysis of structure, dynamics, and interactions of biomolecules is fundamental for understanding molecular mechanisms. Fluorescence resonance energy transfer (*FRET*) is a powerful method that can perform such an analysis.

FRET is a radiationless transfer of excitation energy from a donor to an acceptor governed by a long-range dipole-dipole interaction. This mechanism was first elucidated by Theodore Förster (Stryer, 1978). FRET is a process which does not involve emission and re-absorption of photons. The theory of energy transfer is based on the concept of a fluorophore as an oscillating dipole, which can exchange energy with another dipole with a similar frequency (Fig.2.11). 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. FRET offers an experimental approach for determination of molecular distances in the range of 10 to 80 Angstroms through measurement of efficiency of transfer between a donor and an acceptor located at two specific sites. FRET is also a sensitive technique for detection of global structural alterations. The donor molecule typically emit at shorter wavelengths, which overlap with the absorption spectrum of the acceptor. The rate of the energy transfer depends on the following parameters:

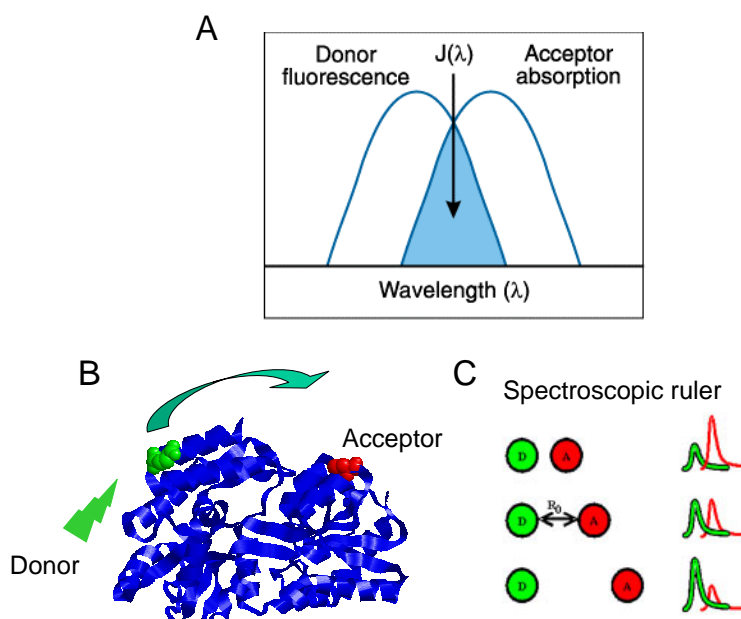


Figure 2.10: Pictorial representation of the basic concept of FRET A) The prerequisite spectral overlap between the donor and acceptor molecules. B) The radiation-less transfer from donor to acceptor upon selective excitation. C) The usage of this technique as a molecular ruler to investigate molecular dynamics.

1. The extent of **spectral overlap** of the emissions spectrum of the donor with the absorption spectrum of the acceptor. The overlap integral $J(\lambda)$ expresses the degree of spectral overlap between a chosen donor and acceptor pair.

$$J(\lambda) = \int_0^{\infty} F_D(\lambda) \epsilon_A(\lambda) \lambda^4 d\lambda = \frac{\int_0^{\infty} F_D(\lambda) \epsilon_A(\lambda) \lambda^4 d\lambda}{\int_0^{\infty} F_D(\lambda) d\lambda}$$

(F_D) is the fluorescence of donor, ϵ_A is extinction coefficient of the acceptor expressed in the units of $M^{-1}cm^{-1}$, λ is the wavelength in centimeters. Therefore, the overlap integral calculated is in units of $M^{-1}cm^3$.

2. The **quantum yield** of the donor. This parameter has already been discussed before in the chapter dealing with fluorescence in general. It can be calculated for a dye of interest under a particular experimental condition by comparing it to a known standard sample eg. Fluorescein or Quinine Hemi-sulfate.

$$Q = \frac{F_t}{F_s} \times \frac{A_s}{A_t}$$

F is the fluorescence of the test (t) or the standard (s) and A is the absorbance of the same test (t) or standard (s) at the excitation wavelength.

3. The relative **orientation** of the donor and acceptor transition dipoles; and the distance between the donor and acceptor.

The rate of transfer of energy is given by the expression:

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

Where R_0 is the Förster Radius at which 50% of the excitation energy is transferred to the acceptor (50% transfer efficiency). If the wavelength is in centimeters and $J(\lambda)$ is in the units of $M^{-1}cm^3$, then Forster radius is given by:

$$R_0 = 9.78 \times 10^3 \sqrt[6]{\kappa^2 \eta^{-4} Q_D J(\lambda)}$$

The efficiency of energy transfer (E) is the fraction of photons absorbed by the donor that are transferred to the acceptor. This is typically measured using the relative fluorescence intensity of the donor, in the absence (F_D) and presence (F_{DA}) of acceptor.

$$E = 1 - \frac{F_{DA}}{F_D}$$

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

$$E = 1 - \frac{\tau_{DA}}{\tau_D}$$

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

$$E = \frac{R_0^6}{R_0^6 + r^6}$$

This expression shows that the transfer efficiency is strongly dependent on distance when the D–A distance is near R_0 .

Alternatively, energy transfer can be measured between identical chromophores that have a limited Stokes shift and is referred to as homotransfer (Kalinin and Johansson, 2004). The theoretical analysis of homotransfer is intrinsically complex since it can only be detected by depolarization experiments and is observed as a decrease in the anisotropy of a fluorophore at higher concentration.

A major uncertainty in the determination of molecular distances by FRET is in the orientation factor for dipole-dipole coupling. This parameter cannot be determined by any current solution technique, therefore the distances calculated from energy transfer data usually are not unique except for cases where an appropriate average value of orientation factor can be applied.

Early studies were based on the assumption that both donor and acceptor dipoles randomize rapidly (dynamic averaging) and sample all orientations (isotropic condition) during the short interval when energy transfer occurs. Under these conditions, $\kappa^2 = \frac{2}{3}$. Frequently the fluorophore can be limited by the surrounding macromolecular structure. A variation of κ^2 from 4 to $\frac{2}{3}$ results in 35% error in the distance calculated but a variation from $\frac{2}{3}$ to 0.01 results in a twofold decrease in R (Dale et al., 1979). There can be many ways to minimize the uncertainty in orientation factor for using FRET for relevant estimations of molecular distances rather than just qualitative data namely:

1. Using polarized emission data to define the mobility of donor and acceptor bound to a macromolecular substrate and to estimate a range of the orientation factor.
2. Choosing donor and acceptor fluorophores that have mixed polarizations and exhibit small limiting polarization properties, and statistical interpretations of energy transfer data to define the limits for the donor acceptor distance and the most probable distance.

Resonance transfer energy is also used to study macromolecular systems in which a single D–A distance is not present such as unfolded proteins or membranes or protein while they are folding, where there is a distribution of such distances. Such systems are better studied with time resolved measurements (Bagshaw and Cherny, 2006).

Presence of a distribution has profound impact on the time resolved decays of the donor. For the native protein, the single D–A distance results in a single transfer rate for all donors. Hence, the decay time of the donor is shortened, and there is only one transfer rate as a result of which the donor decay remains a single exponential. It is this assumption of a single distance which allows calculation of the distance in a steady state scenario. A range of D–A distances result in a frequency response which is spread out along the frequency axis, and one which is no longer a single exponential. The goal of

most distance distribution studies is to recover the D-A probability distribution from the non-exponential decays of the donor. The information content of such data are limited.

2.4.2 Single Molecule FRET

Protein folding is a process characterized by a large degree of conformational heterogeneity. In such cases, classical experimental methods yield only mean values, averaged over large ensembles of molecules. The microscopic distributions of conformations, trajectories, or sequences of events often remain unknown, and with them the underlying molecular mechanisms. Signal averaging can be avoided by observing individual molecules. A particularly versatile method is highly sensitive fluorescence detection. In combination with Förster resonance energy transfer, distances and conformational dynamics can be investigated in single molecules. Fluorescence has become a powerful tool for investigating the dynamics of biological systems and bio-molecules. The availability of high sensitivity photo-detectors and small probe volumes obtainable with visible light have contributed to the development of ultra-sensitive fluorescence spectroscopy and microscopy methods. Ultra-sensitive fluorescence methods allow one to investigate the interactions and dynamics of bio-molecules with high accuracy even on the level of single fluorophores. Such ultra-sensitive methods include fluorescence correlation spectroscopy (FCS) and fluorescence cross-correlation spectroscopy (FCCS), burst analysis, single-molecule studies, single-pair fluorescence resonance energy transfer (spFRET) experiments, and single virus tracing.

In ultra-sensitive fluorescence measurements, it is important to maximize the information retrievable with each photon. The more information that is recorded during a measurement, the more potential exists in the analysis. 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*).

Different methods have been developed that utilize various combinations of the available information. Additional information is available when multiple excitation sources are used. Alternating laser excitation (ALEX) was introduced by Kapanidis and colleagues (Lee et al., 2005). Two excitation sources were interleaved on a timescale between 25 and 3000 μ s, switching between both excitation sources on a timescale faster than diffusion of the particle through the probe volume.

With pulsed interleaved excitation (PIE), the alternation timescale is pushed to the

nanosecond regime. For the experimental conditions used for this study, many excitation pulses from all excitation sources illuminate the sample between the detection of each photon. Hence, on the microsecond timescale, the measurements with the various excitation sources can be considered to be simultaneous. 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. PIE-FCCS increases the sensitivity of FCCS by removing any residual cross talk from the cross-correlation function (CCF). As in ALEX, PIE can also be used to determine the labelling stoichiometry in spFRET and perform more accurate FRET measurements by including only samples containing an active donor and acceptor in the analysis.

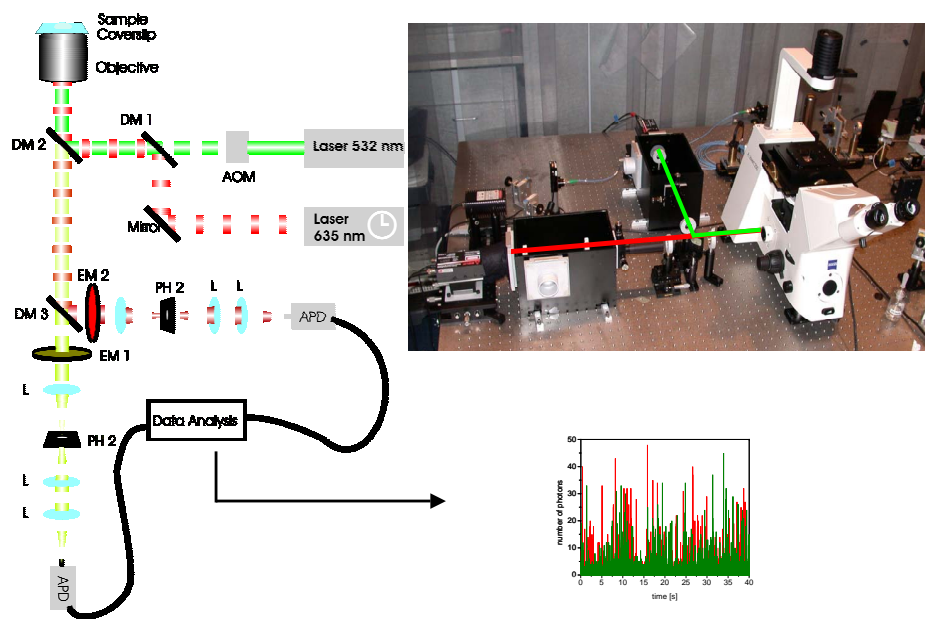


Figure 2.11: Single Molecule Instrumentation and Setup. Schematic and photographic representation of the dual-color confocal microscope with pulsed interleaved excitation sources. In the diagram, AOM refers to the acousto-optic modulator, DM to the dichroic mirrors, EM to the emission filters, PH to the pinholes, L to the lenses, and APD to the avalanche photodiodes. The signal read out is the excitation pulse train as measured by a photodiode. The green and red excitation pulses are colored accordingly (Muller et al., 2005).

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. In the system developed by Muller *et al*, The experimental setup is based on a two-channel confocal microscope. Two excitation sources were used: a pulsed laser diode at 635 nm and a continuous-wave, frequency-doubled laser 532 nm, which was pulsed by an acousto-optic modulator. The lasers were synchronized by a master clock and one source was delayed by 100 ns with respect to the other. The repetition rate used for the measurements discussed was 5 MHz. The photons are detected using time-correlated single-photon counting (TCSPC). The data acquisition card and excitation sources are synchronized with the master clock 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 would allow additional capabilities of PIE by including lifetime information of the fluorophores. From the arrival time, the excitation source for each photon is known and this additional information can be utilized in the data analysis.

The data analysis was performed with the in-house software routines written in PVWave (Visual Numerics, Houston, TX) by Mueller *et al*. The difference between ALEX and PIE is that the alternation between green and red excitation occurs faster than the rate of photon detection. In this work, the repetition rate of the excitation was 5 MHz, much higher than the typical count rates of 100 kHz. Hence, the relevant fluorescence information is collected simultaneously. The fast switching rate gives us the ability to perform auto- and cross-correlation analysis on the collected data with submicrosecond time resolution. 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 expression.

$$f_E = \frac{F_{AD}}{\alpha F_{DA} + F_{AD}}$$

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

2.4.3 Anisotropy

Anisotropy (the opposite of isotropy) is the property of being directionally dependent. Something which is anisotropic may appear different or have different characteristics in

different directions. Fluorescence anisotropy decay belongs to the general class of relaxation methods which monitor the time dependence of the transition of the system from a biased to a random arrangement.

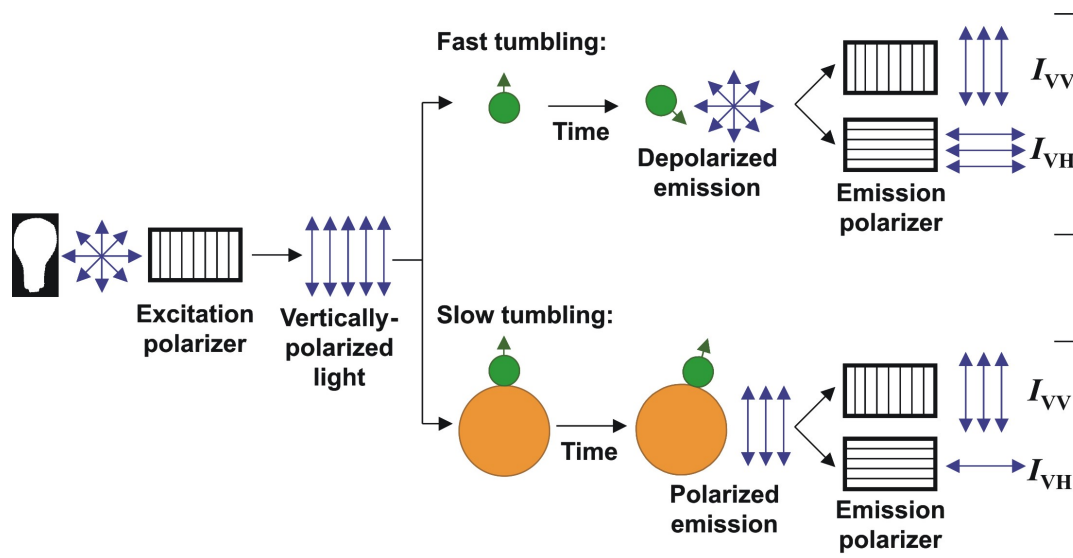


Figure 2.12: The Basic Principle of Anisotropy. The Intensity of fluorescence is monitored first parallel (VV), then perpendicular (VH) to the excitation which is then used to calculate anisotropy which reports the conformational mobility of the probe under study.

The origin of this phenomenon is based on the existence of transition moments for absorption and emission, which lie along specific directions within the fluorophore. In homogenous solutions ground state fluorophores are randomly oriented. Upon excitation with polarized light only those fluorophores get preferentially excited whose absorption transition moment lie along the electric vector of the incident light. Anisotropy (r) can be measured by the following expression.

$$r = \frac{I_{\parallel} - I_{\perp}}{I_{\parallel} + 2I_{\perp}}$$

Where I_{\parallel} and I_{\perp} are parallel and perpendicular components of the monitored fluorescence. By measuring the intensity of emitted light polarized perpendicular to the stimulating radiation, and comparing it with the intensity of light emitted parallel to the stimulating radiation over time, we obtain a measure of how easily the molecules change their orientation. The measure is an anisotropy decay curve, which tells us what proportion of the emitted light is polarized in the same sense as the stimulating light at various times after stimulation. The goal of the intensity measurements is usually to measure a signal proportional to the total intensity (I_t), not one proportional to I_{\perp} or I_{\parallel} or else incorrect decay times are recovered. With the use of polarizers, the measured

intensity can be made proportional to the total intensity, irrespective of the degree of polarization of the sample, by orienting the excitation polarizer in the vertical position and the emission polarizer 54.7° from the vertical. In previous publications one frequently encounters the term polarization which is given by:

$$P = \frac{I_{\parallel} - I_{\perp}}{I_{\parallel} + I_{\perp}}$$

Anisotropy has a substantial advantage in many applications because firstly, it is independent of the total intensity of the sample. Secondly, in multicomponent mixtures of substances with equal fluorescence intensity but variable anisotropy (A_i) the total anisotropy (A_t) is equal to

$$A_t = \sum_i \chi_i A_i$$

where (χ_i) is the mole fraction of the i th component. However, polarization and anisotropy can be inter converted by the relationship.

$$P = \frac{3r}{2 + r}$$

$$r = \frac{2P}{3 - P}$$

The anisotropy measurement reveals the average angular displacement of the fluorophore that occurs between absorption and subsequent emission of the photon. This angular displacement is dependent on the rate and extent of rotational diffusion of the fluorophore during the lifetime of the excited state. These motions in turn depend on the viscosity of the medium, size and shape of the particle. Therefore anisotropy measurements have been used to quantify protein denaturation, its association with other macromolecules or in this case the internal dynamics of the protein.

Steady state anisotropies are measured using continuous illumination and represent an average of the anisotropy decay over intensity decay. Measurement of steady state anisotropy decay is simple, but its interpretation usually depends on an assumed form for the anisotropy decay, which is not directly observed in the experiment. Additional information is available if one measures the time-dependent anisotropy, that is, the values of $r(t)$ following pulsed excitation. The form of the anisotropy decay depends on the size, shape, and flexibility of the labelled molecule. It is valuable to understand the factors that can affect anisotropy decays.

Shape: For a spherical molecule, the anisotropy is expected to decay with a single rotational correlation time (θ). The most common interpretation of the correlation time is in terms of overall rotational correlation time of the protein. But the measured values of θ can also result in multi-exponential anisotropy decay curves. These can be due to non-spherical fluorophores or proteins. In such cases the correlation times are determined by the rates of rotation of the molecule about various molecular axes. Thus, anisotropy can also be used to estimate the shapes of the molecules.

Segmental Flexibility: In addition to shape, anisotropy decay is also affected by the segmental flexibility of the molecule. For instance, tryptophan anisotropy decays for proteins frequently display correlation times that are too short to be due to rotational diffusion of the protein. These short correlation time components are often due to independent motion of the tryptophan residue within the protein. Such measurements are widely used to understand the internal dynamics of proteins.

HomoFRET: Anisotropy decays can also be affected by resonance energy transfer between molecules of the same type of fluorophore, that is, depolarisation due to homotransfer.

2.5 Aim of the Study

The GroE chaperonin system is structurally and mechanistically one of the most extensively studied molecular machines. But despite the large body of data available, the conformational properties of the substrate protein as it transits the GroEL/GroES cycle, have remained largely unclear.

Using single molecule and ensemble FRET methods, the aim of this study was to address the following questions concerning the conformation of a model protein (DM-MBP) during its Chaperonin (GroEL/GroES) assisted folding.

1. *Bound State* What is the conformation of the protein when it is bound to GroEL in comparison to the unfolded state in free solution? Is there heterogeneity in the bound state or does the bound protein populate a defined conformation? What is the degree of expansion if observed of the bound protein and how is the nature of the bound state different between the GroEL and SR-EL chaperonins?
2. *Expansion or Compaction* As the protein transits the GroEL/GroES cycle, how does the conformation of the bound protein protein changes as a result of various events like nucleotide or GroES binding to GroEL?
3. *GroEL cycling and the Bound State* Does the protein revert to its initial conformation during chaperonin cycling?
4. *Segmental mobility* Is there a pattern of release for the substrate protein from the GroEL surface upon binding of nucleotide and or GroES? Does it have a correlation with the secondary structure elements or hydrophobicity of the different regions of the protein?
5. *Spontaneous and GroEL assisted folding pathway:* how does GroEL accelerate the rate of folding of certain proteins? Does it interact preferentially with the native state or the transition state of the substrate with out changing the pathway of folding or does it re-route the pathway resulting in intermediates that are different from those of the spontaneous refolding pathway.

Chapter 3

Materials and Methods

3.1 Materials

3.1.1 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 C5 maleimide	Molecular Probes
Alexa Fluor 594 C5 maleimide	Molecular Probes
Ammonium persulfate (APS)	Sigma-Aldrich
Ampicillin	Merck
Amylose resin	New England Biolabs
Atto 532 C5 maleimide	Atto-Tec
Atto 647-N C5 maleimide	Atto-Tec
Bacto agar	Difco
Bacto trypton	Difco
Bacto yeast extract	Difco
Bovine Serum Albumin (BSA)	Sigma-Aldrich
Bromophenol blue	Sigma-Aldrich
Calcium chloride	Merck
CDTA (trans-1,2-diaminocyclohexane - N, N, N', N'-tetracetic acid)	Sigma-Aldrich
Chloramphenicol	Sigma-Aldrich
Complete EDTA-free protease inhibitor	Roche

Coomassie brilliant blue R-250	Roth
Dimethylsulfoxide (DMSO)	Merck
Dithiothreitol (DTT)	Roche
<i>ECLTM</i> detection kit	Amersham Pharmacia Biotech
Ethanol	Merck
Ethidium bromide	BioRad
Ethylenediaminetetraaceticacid-sodium salt(EDTA)	Merck
Glucose	Sigma-Aldrich
Glycerol	Merck
Glycine	Roth
Guanidium hydrochloride (GuHCl)	Sigma-Aldrich
HEPES	Sigma-Aldrich
Hydrochloric acid (37%)	Merck
Isopropyl- β -D-thiogalactopyranoside (IPTG)	BioMol
Magnesium chloride	Merck
Maltose	Sigma-Aldrich
β -mercaptoethanol	Sigma-Aldrich
Methanol	Merck
Phenyl-methyl-sulfonyl-fluoride (PMSF)	Sigma-Aldrich
Polyacrylamide/bisacrylamide solution 30 % (30 : 0.8)	Roth
Potassium hydroxide	Sigma-Aldrich
Sodium chloride	Merck
Sodium dodecylsulfate (SDS)	Sigma-Aldrich
Sodium hydroxide	Sigma-Aldrich
Sodium thiosulfate	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

3.1.2 Enzymes

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

3.1.3 Materials

Centricon 10 kDa cut-off	Amicon
Microcon 10 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

3.1.4 Instruments

AIDA gel imaging software version 2.31	Raytest
AKTA Explorer 100	Amersham Pharmacia Biotech
Balance AG285, PB602	Mettler Toledo
Centrifuges: Avanti J-25, Avanti J20 XP, J-6B, GS-6R	Beckmann
Centrifuges 5415C and 5417R	Eppendorf
Chromatography columns (HiPrep Desalting, MonoQ, HiTrap Heparin, Sephacryl S200/S300, Superdex 200, Superose 6, Sephadex G25 (NAP-5, NAP-10, NAP-25); chromatography resins: Q-Sepharose, DE52, Source 30 Q, Source 30 S)	Amersham Pharmacia Biotech
Deionization system MilliQ plus PF	Millipore
Electrophoresis chambers MiniProtean 3	Bio-Rad
Electrophoresis power supply Power PAC 300	Bio-Rad
Fluorescence spectrometer Fluorolog 3	HORIBA Jobin Yvon
FPLC systems	Amersham Pharmacia Biotech
EmulsiFlex high pressure homogenizer	Avestin
Gilson Pipetman (2, 10, 20, 100, 200, 1000 μl)	Abimed

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 Accumet Basic	Fisher Scientific
SMART system	Amersham Pharmacia Biotech
Sonicator Ultrasonic Processor XL	Misonix Inc.
Spectrophotometer DU 640 UV/VIS	Beckmann
Sx.18MV Stopped-Flow Reaction analyser	Photo Physics
Synergy HT UV/VIS/fluorescence/luminescence plate reader	Bio-Tek
UV/VIS Spectrometer V-560	Jasco
Thermomixer Comfort	Eppendorf
Vortex	Ikamag
Water bath	Bioblock Scientific

3.1.5 Media

- **LB medium** 10 g/l tryptone, 5 g/l yeast extract, 5 g/l NaCl, (and 15 g/l agar for solid medium). Adjusted to pH 7.0 with NaOH (Sambrook et al., 1989).
- **SOC medium** 20 g/l tryptone, 5 g/l yeast extract, 0.5 g/l NaCl, 0.186 g/l KCl, 0.95 g/l $MgCl_2$. After autoclave, add 20 ml of filter sterilized 1M glucose (Sambrook et al., 1989).

3.1.6 Antibiotic Stock Solutions

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

Ampicilin:	100 g/l
Chloramphenicol:	25 g/l

3.2 Plasmids

GroEL and all chaperonin size and charge mutants were constructed in a pCH vector backbone by Yun Chi Tang (Tang et al., 2006) and were kindly given for use in this study. All the MBP mutants (cys-derivatives) were generated by site-directed mutagenesis in pCH vector on the background of the DM-MBP. Poly-Proline and DM-MBP fusion

construct was made by introducing synthetic oligonucleotides into the pCH DM-MBP stop to Gly conversion mutation. GroES was constructed in a pET11a vector inserted via the NdeI and BamHI sites (Brinker et al., 2001).

The sequences of all final constructs were confirmed by DNA sequencing (Sequiserve and Medigenomix GmbH). All new DNA constructs used in this work are listed in the tabulated form.

Plasmid	Promoter/Origin	Selection marker
PCH-MBP-DM N TERMINAL CYS	T7/ColEI	Ampicillin
PCH-MBP-DM A52C	T7/ColEI	Ampicillin
PCH-MBP-DM C TERMINAL CYS	T7/ColEI	Ampicillin
PCH-MBP-DM A52C/N TERMINAL CYS	T7/ColEI	Ampicillin
PCH-MBP-DM A52C/C TERMINAL CYS	T7/ColEI	Ampicillin
PCH-MBP-DM N and C TERMINAL CYS	T7/ColEI	Ampicillin
PCH-MBP-DM S114C/N TERMINAL CYS	T7/ColEI	Ampicillin
PCH-MBP-DM S263C/N TERMINAL CYS	T7/ColEI	Ampicillin
PCH-MBP-DM A52C/S114C	T7/ColEI	Ampicillin
PCH-MBP-DM A52C/S263C	T7/ColEI	Ampicillin
PCH-MBP-DM A52C/S303C	T7/ColEI	Ampicillin
PCH-MBP-DM A52C/D95C	T7/ColEI	Ampicillin
PCH-MBP-DM D65C/K175C	T7/ColEI	Ampicillin
PCH-MBP-DM D30C/P298C	T7/ColEI	Ampicillin
PCH-MBP-DM S73C/P298C	T7/ColEI	Ampicillin
PCH-MBP-DM P298C/T345C	T7/ColEI	Ampicillin
PCH-MBP-DM A52C/A312C	T7/ColEI	Ampicillin
PCH-MBP-DM A52C/P298C	T7/ColEI	Ampicillin
PCH-MBP-DM A134C/P298C	T7/ColEI	Ampicillin
PCH-MBP-DM K175C/P298C	T7/ColEI	Ampicillin
PCH-MBP-DM S-(Pro)10-C	T7/ColEI	Ampicillin
PCH-MBP-DM C-(Pro)10-C	T7/ColEI	Ampicillin
PCH-MBP-DM D30C/S73C	T7/ColEI	Ampicillin
PCH-MBP-DM D30C/A134C	T7/ColEI	Ampicillin
PCH-MBP-DM D30C7K175C	T7/ColEI	Ampicillin
PCH-MBP-DM D30C/A312C	T7/ColEI	Ampicillin
PCH-MBP-DM D30C/T345C	T7/ColEI	Ampicillin
PCH-MBP-DM A21C	T7/ColEI	Ampicillin
PCH-MBP-DM K34C	T7/ColEI	Ampicillin

Plasmid	Promoter/Origin	Selection marker
PCH-MBP-DM K42C	T7/ColEI	Ampicillin
PCH-MBP-DM K46C	T7/ColEI	Ampicillin
PCH-MBP-DM K88C	T7/ColEI	Ampicillin
PCH-MBP-DM K127	T7/ColEI	Ampicillin
PCH-MBP-DM A141C	T7/ColEI	Ampicillin
PCH-MBP-DM K170C	T7/ColEI	Ampicillin
PCH-MBP-DM K179C	T7/ColEI	Ampicillin
PCH-MBP-DM A190C	T7/ColEI	Ampicillin
PCH-MBP-DM K202C	T7/ColEI	Ampicillin
PCH-MBP-DM A269C	T7/ColEI	Ampicillin
PCH-MBP-DM K277C	T7/ColEI	Ampicillin
PCH-MBP-DM K295C	T7/ColEI	Ampicillin
PCH-MBP-DM K326C	T7/ColEI	Ampicillin
PCH-MBP-DM K362C	T7/ColEI	Ampicillin
PCH-MBP-DM L113C	T7/ColEI	Ampicillin
PCH-MBP-DM I161S	T7/ColEI	Ampicillin
PCH-MBP-DM P298C	T7/ColEI	Ampicillin
PCH-MBP-DM A312C	T7/ColEI	Ampicillin
PCH-MBP-DM D30C	T7/ColEI	Ampicillin
PCH-MBP-DM L147C	T7/ColEI	Ampicillin
PCH-MBP-DM A52C/W230C	T7/ColEI	Ampicillin
PCH-MBP-DM A52C/M321C	T7/ColEI	Ampicillin
PCH-MBP-DM A52/S233C	T7/ColEI	Ampicillin
PCH-MBP-DM A52/A141C	T7/ColEI	Ampicillin
PCH-MBP-DM K34C/K141C	T7/ColEI	Ampicillin
PCH-MBP-DM K34C/A190C	T7/ColEI	Ampicillin
PCH-MBP-DM K34C/A202C	T7/ColEI	Ampicillin
PCH-MBP-DM K34C/K362C	T7/ColEI	Ampicillin

- The following positions were not active upon mutation and could not be purified via affinity purification: S114C, W230C (reported to have reduced affinity to maltose), and S233C (reported to have reduced affinity to maltose).
- The following positions were refolding defective after labelling : S263C, A303C, S73C and A134C.
- The following positions were not amenable for labelling: the N-terminal and the C-terminal Cys residues.

3.3 Molecular Cloning Methods

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

For preparation of chemically-competent *E. coli* cells, a single colony was used to inoculate 500 ml LB medium (including antibiotic, if applicable) and grown to an optical density (OD₆₀₀) of 0.25 – 0.5 at 37°C. The cells were then chilled on ice for 15 min and harvested at 5000 × 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 *CaCl*₂, 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, approx. 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 90 s and subsequently placed on ice for 2 min. 1 ml of LB medium was added and the cells were shaken at 37°C for 1 h. The cell suspension was then plated on selective plates and incubated at 37°C, until colonies had developed (typically 10 – 16 h).

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 (OD₆₀₀) 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 same as in the previous case.

3.3.2 Plasmid Purification

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

3.3.3 PCR Amplification

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

DNA Template:	50 ng (plasmid DNA)
Primers:	20 pmol each
dNTPs:	200 μ M each
Pfu DNA Polymerase:	2.5 U
Polymerase buffer:	1 x
Additives:	3 – 6 % DMSO if GC content was more than 50 %
Final volume:	50 μ l
Cycling conditions	35 cycles
Initial denaturation:	94°C, 5 min
Cycle denaturation:	94°C, 30 – 60 s
Annealing:	approx. 55°C, 30 – 60 s
Extension:	72°C, duration dependent on template length: 1 kbp/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.

3.4 DNA Restriction and Ligation

DNA restriction was performed according to the manufacturer's instructions of the respective enzymes. Typically, a 50 μ l reaction contained 1 – 2 μ l of each restriction enzyme and 0.5 – 2 μ g purified PCR product or 1 – 5 μ g plasmid DNA in the appropriate reaction buffer. Digested vector DNA was dephosphorylated with shrimp alkaline phosphatase. For ligation, 100 – 200 ng (approx. 1 – 2 μ l) dephosphorylated vector DNA, 100 – 200 ng (approx. 5 – 10 μ l) DNA insert and 1 μ l (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* DH5 α cells as described.

3.5 DNA Analytical Methods

DNA concentrations were measured by UV absorption spectroscopy at $\lambda = 260$ nm. A solution of 50 μ g/ml of double stranded DNA in H_2O exhibits approximately A_{260} nm = 1. Agarose gel electrophoresis was performed in TAE buffer (40 mM Tris, 1 mM EDTA, 20 mM acetic acid) and 1 – 2% TAE-agarose gels, supplemented with 1 μ g/ml ethidium bromide, at 4 – 6 V/cm. DNA sequencing was performed by Medigenomix GmbH (Martinsried, Germany) or Sequiserve (Vaterstetten, Germany).

3.6 Protein Purification

3.6.1 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 about 0.5. The induction was then proceeded by adding 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 in 25 ml). The suspension was further treated with lysozyme (1 mg/ml) and benzonase (approx.500 units). Lysis was achieved by homogenization of the cell suspension in an EmulsiFlex C5 device kept on ice. Cell debris were removed by ultracentrifugation for 60 min at 40,000 x g, 4°C and with subsequent filtration (0.2 μ m) for clearance of the supernatant. The supernatant was applied to a 400 ml DE52 column attached to an AKTA Explorer chromatography system. After washing with two column volumes of the lysis buffer, a NaCl gradient was elevated to 0.6 M in five column volumes. The GroEL 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, eluted in 30 mM Tris-HCl pH 7.5, 1 mM DTT, 1 mM EDTA with a NaCl gradient from 0 to 0.5 M. 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-containing flow through was collected and concentrated to less than 5 ml in Centriprep concentrators. Eventually 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 approx. 35 mg/ml (equivalent to 44 μ M of GroEL oligomer). And aliquots were flash frozen in liquid N_2 and stored at -80°C. Total yield of GroEL was typically approx 600 mg.

3.6.2 GroES Expression and Purification

The expression and purification of GroES was similar as GroEL as described above, including the induction, lysis and spin conditions. The supernatant was applied to a 400 ml DE52 column attached to an AKTA Explorer chromatography system. After washing with two column volumes of the above buffer, a NaCl gradient was increased upto 0.5 M in five column volumes. And GroEL containing fractions were collected and dialyzed in

5 l lysis buffer overnight at 4 °C. The deslating 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 Centriprep concentrators. Eventually the concentrated sample was 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 approx. 15 mg/ml. Aliquots were frozen in liquid N₂ and stored at -80°C. Total yield of GroES was typically about 400 mg.

3.6.3 Expression of MBP and MBP Mutants 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 2l LB medium containing 100 mg/l ampicillin at 37°C to OD_{600} of approx. 0.9. The induction was then proceeded 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 (50 mM Tris-HCl pH 7.5, 20 mM NaCl, 1 mM DTT, and 1 mM EDTA) and Complete protease inhibitor (1 tablet for 100 ml). The same lysis and spin conditions as GroEL purification were applied. The supernatant was further dialyzed in amylose buffer to remove cellular maltose and slowly loaded on to a 50 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 (20 mM Tris-HCl pH 7.5, 50 mM NaCl, 1 mM DTT, and 1mM EDTA) overnight at 4°C then concentrated to 10 mg/ml. Aliquots were frozen in liquid N₂ and stored at -80 °C. Typical yield of MBP from 2 l culture was about 100 mg.

3.6.4 Thiol-Mediated Labelling of the Cys Constructs

Double cys constructs in which one of the positions was 298 could be labelled specifically as this position shows a differential accessability in presence or absence of Maltose. MBP Cys mutants (100 μ M) were labelled in buffer La - PBS pH 7.8, 500mM malto-triose (Sigma). for 2 hr at 25°C in the presence of a 1.1 fold molar excess of the fluorophore Atto 532 maleimide, (ATTO -TEC,Inc.). Unbound fluorophore was removed using micro NAP columns (Amersham Biosciences) equilibrated in PBS pH 7.8. Before the second step labelling, the protein was denatured in Buffer-Lb 3.0 M GuHCl in PBS pH 7.8 and incubated with Activated Thiol Sepharose (Amersham Biosciences) for 2 hours and the singly labelled species were eluted with buffer Lb and 50mM DTT. Second step labeling

was carried out at similar conditions with 2 fold molar excess of Atto 647N maleimide, (ATTO -TEC,Inc.) The coupling efficiency measured by the absorption of MBP ($\epsilon_{280} = 69mM^{-1}cm^{-1}$), and Atto 532 ($\epsilon_{532} = 115mM^{-1}cm^{-1}$) was more than 90%.

For stochastic labelling, the protein was labelled in buffer Lb with equimolar concentration of both the dyes mixed together. The reaction was carried out for 2 hours at 25°C and the protein was desalted to remove any unbound dye using the micro NAP columns; and concentration was determined using the above mentioned extinction coefficients.

3.6.5 Thiol-Mediated Labelling of GroEL and GroES

GroEL E315C and GroES 73C were used for thiol mediated labelling. GroEL or GroES were labelled in buffer La - PBS pH 7.8, upon incubation with the dye with 1 : 1 Molar ratio for 2 hrs on ice. The excess unbound fluorophore was removed using micro NAP columns (Amersham Biosciences) equilibrated in Buffer LS.

3.7 Protein analytical methods

3.7.1 Determination of Protein Concentration

Protein concentrations were determined spectrophotometrically by A280 (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 as oligomers and for DM-MBP as monomer.

3.7.2 SDS-PAGE

SDS-Polyacrylamide gels were prepared as follows: 4 % stacking gels and 10 %, 12 % or 15 % separating gels were prepared according the following recipe:

% acrylamide (AA)	12	15	4
30 % AA, 0.8%bis-AA	8 ml	10 ml	1.3 ml
1.5 M Tris-HCl, pH 8.8	5 ml	5 ml	-
0.5 M Tris-HCl, pH 6.8	-	-	2.5 ml
10 % SDS	0.2 ml	0.2 ml	0.1 ml
Water	1.8 ml	-	6.1 ml
TEMED	10 μ l	10 μ l	5 μ l
APS	50 μ l	50 μ l	100 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.

3.8 Biochemical and Biophysical Methods

Protein refolding reactions for ensemble studies containing chaperones (when present) were carried out with the following molar concentration ratios of chaperones to substrate: 1 substrate (monomer) : 2 GroEL (tetradecamer) : 4 GroES (heptamer), unless otherwise specified.

3.8.1 MBP Refolding

MBP-DM and its derivative cys mutants (12.5 μ M) were denatured in 3 M GuHCl, 5 mM Tris-HCl, pH 7.5, 20 mM KCl (Buffer B) at 25°C and refolded upon 50-fold dilution into low salt buffer - Buffer A (20 mM Tris-HCl, pH 7.5, 20 mM KCl, 5 mM Mg(OAc)₂) in the absence or presence of chaperones. GroEL/GroES assisted refolding was initiated at 25°C by the addition of 5 mM ATP. Intrinsic tryptophan fluorescence was monitored on a Fluorolog 3 Spectrofluorometer (Spex) with an excitation wavelength of 295 nm (slit width 2 nm) and an emission wavelength of 345 nm (slit width 5 nm) (Chun et al., 1993).

3.8.2 Stopped-Flow Experiments

All stopped-flow experiments were done by using an Applied Photo Physics SX.18MV stopped-flow reaction analyzer instrument, with a 1 : 1 mixing ratio at 25°C. The excitation wavelength used was 515 nm with a cut-off filter of wavelength 530 nm. For anisotropy measurements dual channel measurements were done using photomultiplier tubes (PMTs) measuring the parallel and the perpendicular measurements simultaneously.

3.8.3 Preparation of GroEL-ADP-GroES "Bullet"

1mM ADP was added to buffer A containing 10 μ M of GroEL and 15 μ M of GroES. The excess ADP was removed using micro bio-spin columns with a cut-off of 40KDa (Biorad) equilibrated in buffer A. The obtained solution was diluted to 1 μ M GroEL. The integrity of the bullet was tested in presence and absence of ATP by observing FRET between Atto-532 labelled GroEL and Atto-647N labelled GroES.

3.8.4 Equilibrium Unfolding of MBP

The equilibrium unfolding of MBP (0.25 μ M) was performed at 25°C in buffer A containing 0 – 1.5 M GuHCl and followed by intrinsic tryptophan fluorescence at 345 nm. Equilibrium unfolding curves fit a two state transition.

3.8.5 Sample for Single Molecule Spectroscopy of Chaperonin-Substrate Complex

Substrate and GroEL/SR-EL Complexes were formed by diluting double labelled denatured MBP-DM in buffer B (3 μ M) 100 fold into buffer A containing GroEL/EL-SR (3 μ M). These complexes were then analyzed for their diffusion time through the probe volume by fluorescence correlation spectroscopy.

This solution was then diluted in buffer A 500 times for spFRET by burst analysis. The picomolar(pM) concentrations of labelled protein (the final concentration of chaperonin was 3 μ M) ensure that the probability of more than one molecule in the probe volume at the same time is negligible.

3.8.6 Sample Preparation for Single Molecule Spectroscopy of GroEL-Cycling

After initiating GroEL/GroES/ATP mediated refolding of DM-MBP(52 – 298) and DM-MBP (30 – 312), the reaction was terminated using EDTA and apyrase at 90sec and 300sec. Native and unbound protein that may have been formed during the course of the reaction was removed using amylose resin which would bind native protein and was passed through a 100kD centricon to remove any DM-MBP that remains unbound to GroEL (Fig 3.1). Sample was also prepared to check the distribution of GroEL bound substrate when GroEL binds to the substrate already undergoing spontaneous refolding. GroEL was added to an aliquot of spontaneously refolding double labelled DM-MBP(52 – 298) at similar time points and the sample was processed in a similar fashion to remove and any native and unbound protein.

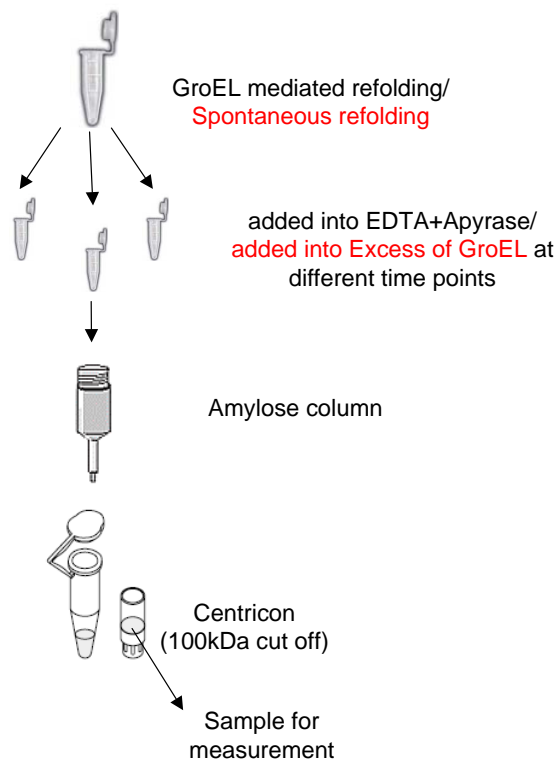


Figure 3.1: Schematic Representation of the Method. Sample Preparation for Investigating the Bound State upon GroEL Cycling and Spontaneous Refolding.

3.8.7 Calculation of FRET Efficiencies Using Burst Analysis

The FRET efficiencies were determined using burst analysis (Deniz et al., 1999). The single molecule diffusing through the confocal volume emits a burst of photons, which were counted as significant, when the number of photons in a 2ms time bin was larger than a threshold of 17 photons, after the green excitation. For ensuring the presence of an active acceptor, additional 12 photons must be detected after the red excitation. FRET efficiencies were calculated as discussed in the introduction using the equation:

$$f_E = \frac{F_{AD}}{\alpha F_{DA} + F_{AD}}$$

where $\alpha = (\phi_A \eta_A / \phi_D \eta_D)$ corrects for differences in the quantum yields (ϕ) and the detection efficiencies (η) of the donor and acceptor, respectively. This factor was deter-

mined as described in (Lee et al., 2005) to be 0.9 for Atto532 and Atto647N labelled to MBP-DM. Average values of crosstalk and direct excitation were subtracted from the fluorescence signals.

3.8.8 Determination of the Forster Radius

The Forster radius R_0 for the FRET-pair Atto532/647N labelled to MBP-DM was calculated with the program PhotChemCad using $n=1.33$ for the index of refraction, $\kappa^2 = 2/3$ for the orientation factor and $\epsilon_{Atto647} = 1.5 \times 10^5 \text{ l/mol cm}$ for the absorption coefficient. The absorption-/emission-spectra were taken of the bound fluorophores, the quantum yield of the donor was determined by comparing the quantum yield of the free dye in water (0.9 given by Atto-Tec) and the bound fluorophore to be $\phi_{532} = 0.4$.

Chapter 4

Results

4.1 Sp-FRET Distribution of Double labelled DM-MBP

In order to study the conformational heterogeneity of unfolded protein when bound to GroEL or SR-EL, we used DM-MBP which is known to be an *in vivo* substrate of GroEL and *in vitro* exhibits acceleration in refolding rate in presence of the GroE system (Tang et al., 2006). We constructed a series of double cysteine mutants on the background of DM-MBP to facilitate thiol specific double labelling of the protein (Fig.4.1A). The double labelling of these proteins for single molecule measurements were carried out stoichiometrically as PIE-spFRET¹ technique is capable of analyzing only double-labelled molecules from a pool of heterogeneously labelled samples. The cysteins engineered at various positions e.g. 30, 52, 175, 298, and 312 did not affect GroEL/ES/ATP assisted folding rate of the proteins as examined by Trp fluorescence (Tang et al., 2006), which reports the fraction of refolded protein.

As Trp fluorescence is highly quenched in the double labelled proteins, the fraction of folded proteins formed was quantified by performing the refolding reaction in presence or absence of maltose and monitoring the fluorescence of Atto-532. DM-MBP(52–298) and DM-MBP(175 – 298) were labelled preferentially at position 52 and 175 with Atto-532 respectively in presence of maltose and 298 was labelled with Atto-647N after maltose was removed from the labelling reaction. Position 298 in the double labelled protein is sensitive to maltose binding and hence this can be used to estimate the fraction of native protein. Maltose-bound double-labelled protein has about 2.5-fold higher fluorescence compared to maltose-free native protein. This may be due to change in the environment of the acceptor from hydrophilic to hydrophobic brought upon by binding of maltose or by its quenching. The ratio of the fluorescence intensities of the refolding in presence

¹Sp-FRET measurements were done together with Barbara K. Mueller.

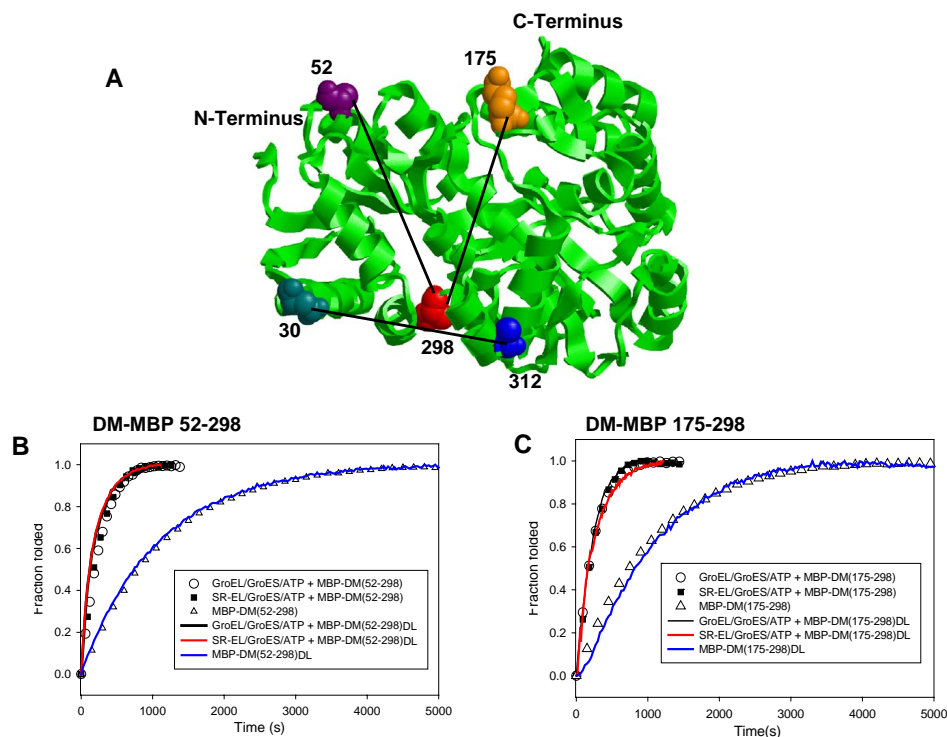


Figure 4.1: Representation of the Cys-Mutants Used and Their Folding Behaviour. Ribbon diagram of the structure of MBP (spurolino et al, 1991; pdb 1OMP; PyMol) with the positions for engineered cysteine into the protein shown as space fill structures. Refolding of DM-MBP(52 – 298) (B) and DM-MBP(175 – 298) (C) was initiated upon 50-fold dilution of the proteins ($12.5\mu\text{M}$) unfolded in 3M GuHCl at 25°C into buffer A. Chaperonin assisted rates were obtained by dilution of the unfolded proteins in buffer A containing $3\mu\text{M}$ of chaperonin, $6\mu\text{M}$ of GroES and 2mM of ATP.

and absence of maltose was used to quantitate the fraction of protein capable of binding maltose (Fig.4.1B and C).

Protein	Spontaneous Fold- ing Rates($\times 10^3$)	GroEL/ES Fold- ing Rates($\times 10^3$)	SR-EL/ES Fold- ing Rates($\times 10^3$)
DM-MBP(52 – 298)	0.66 ± 0.04	3.8 ± 0.05	3.9 ± 0.08
DM-MBP(52 – 298)SL	0.71 ± 0.05	3.94 ± 0.08	4.6 ± 0.14
DM-MBP(52 – 298)DL	0.77 ± 0.06	5.27 ± 0.12	5.3 ± 0.11
DM-MBP(175 – 298)	0.7 ± 0.05	4.2 ± 0.03	4.6 ± 0.12
DM-MBP(175 – 298)SL	0.8 ± 0.06	4.7 ± 0.1	4 ± 0.16
DM-MBP(175 – 298)DL	0.73 ± 0.07	4.2 ± 0.07	4.1 ± 0.05

Rate of formation of native protein was similar for both labelled and unlabelled proteins indicating that the introduction of label does not change the spontaneous as

well as the chaperonin assisted folding reaction. All the refolded proteins reached the native state as monitored by their ability to bind maltose (Fig.4.2).

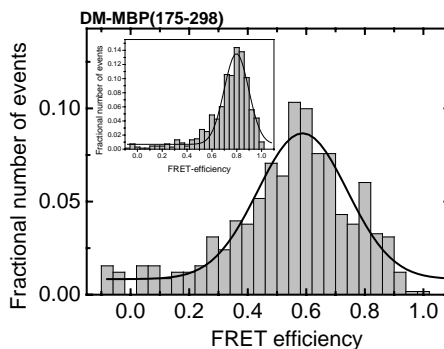


Figure 4.2: Single Molecule Distribution of the Maltose Bound and Maltose Free Form of DM-MBP(175 – 298). Sp-FRET distribution of the native state of DM-MBP(175 – 298) without maltose (inset) changes upon binding of maltose as shown by a shift in the peak FRET efficiency.

The distance distributions obtained for the native state for these constructs as well as others (described in appendix) correspond well with the distances calculated from the crystal structure (Tabulated below) of MBP (1OMP, (Spurlino et al., 1991)) acting as a control for the validity of the Sp-FRET method for this protein. The Förster radius of the dye pair when coupled to protein was found to vary between 45 – 50Å.

Construct	Distance (Crystal Structure) (in Å)	Distance Refolded (FRET) (in Å)	Distance Unfolded (FRET) (in Å)
52-298	33	35	67
175-298	32	34	61
30-175	40	41	52
30-134	61	61	60
52-312	39	40	74
73-298	32	32	65
134-298	37	38	58

The refolded proteins also show distributions with high E_{FRET} identical to that of native state indicating that the labelled proteins are able to reach the native state upon refolding.

The distance distributions obtained after spontaneous and GroEL/GroES-assisted refolding were indistinguishable (Fig.4.3 B and C) and were in good agreement with the distances calculated from the MBP crystal structure (as seen from the Table above).

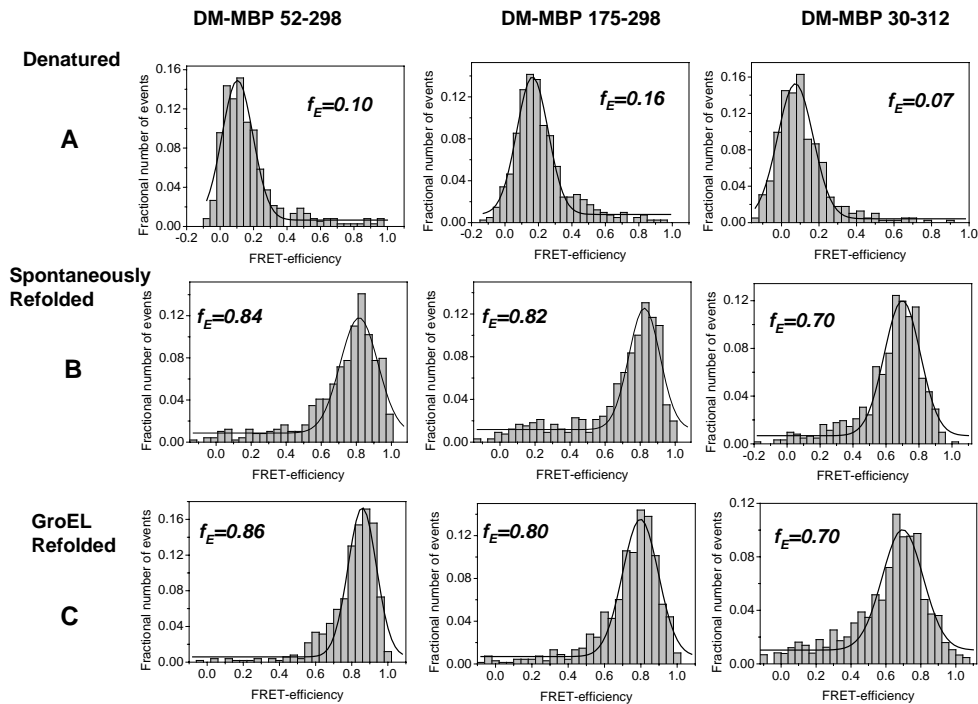


Figure 4.3: Sp-FRET Measurements of Different Cys Variants of Double Labeled DM-MBP Sp-FRET measurements of double-labelled DM-MBP(52 – 298) (left), DM-MBP(175 – 298)(center) and DM-MBP(30 – 312)(right). FRET efficiency distribution is shown for the protein unfolded in 3M GuHCl (A) ,spontaneously refolded protein (B) and GroEL refolded protein (C).

Fully denatured DM-MBP in solution exhibited similarly low E_{FRET} values for all three cysteine pairs measured (Fig.4.3A).

This denatured conformation was compared with that of the protein bound to GroEL upon dilution from denaturant.

4.2 Conformational Heterogeneity of GroEL Bound Protein

Denatured protein was diluted into GroEL containing buffer. Interestingly, a bimodal E_{FRET} distribution was observed for the GroEL-bound protein (Fig.4.4 A). In the case of DM-MBP(52 – 298), a broad distribution of molecules (approx. 60 % of total) centered around a high E_{FRET} of approx. 0.7, comparable in compactness to the native state, whereas another population of molecules (about 40% of total) exhibited a more narrow distribution with an E_{FRET} of 0.06, more expanded than the average conformational

distribution of the fully denatured protein (Fig.4.4, left).

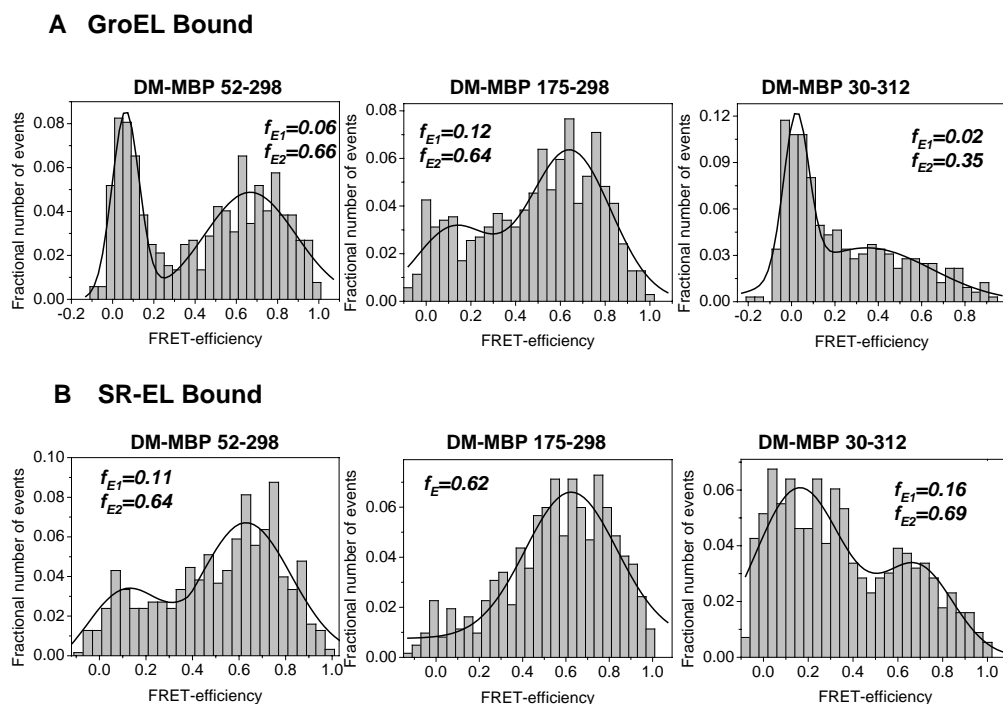


Figure 4.4: Sp-FRET Measurements of Chaperonin bound DM-MBP) Sp-FRET measurements of double-labelled DM-MBP(52–298) (left), DM-MBP(175–298) (center) and DM-MBP(30–312) (right). GuHCl-denatured double labelled protein (3nM) at 25°C was diluted 50-fold into buffer A containing 3μM of GroEL (A) or SR-EL (B) to obtain the chaperonin bound distributions.

Qualitatively similar results were obtained for DM-MBP(30 – 312), but in this case a greater fraction of GroEL-bound molecules (about 65% of total) populated the expanded conformation (Fig.4.4, right). DM-MBP(175 – 298) on the other hand exhibited a very broad E_{FRET} distribution when bound to GroEL, with only a small fraction of molecules being as expanded as in denaturant (Fig.4.4, center). To rule out the possibility that there are unbound unfolded species in the sample which exhibits the low FRET efficiency distribution, the measurements were performed with two to three times higher concentrations of GroEL or SR-EL which resulted in similar distributions.

More compact conformational distribution were generally observed when the proteins were bound to the GroEL single-ring mutant, SR-EL (Fig.4.4). This effect correlated with a lower affinity of SR-EL for unfolded DM-MBP compared to wild-type (WT) GroEL. With SR-EL, FRET efficiencies show a single distribution having a peak with

high FRET efficiency (62%) (Fig.4.4B). Since SR-EL has a lower binding affinity for DM-MBP than GroEL (shown in later section), absence of the expanded state in the SR-EL-DM-MBP complex rules out the possibility that the open state seen with GroEL is an artifact of unbound unfolded species. These experiments were also reproduced by ensemble steady state FRET measurements which showed that FRET efficiency of the protein bound to GroEL has lower efficiency than the one bound to SR-EL.

The GroEL and SR-EL bound distributions differ from each other where the distribution of SR-EL bound form is broader and shifted towards higher FRET efficiency state than GroEL bound form. DM-MBP(175C-298C) on the contrary exhibits very similar distributions when bound to GroEL (Fig.4.4A, center) or SR-EL (Fig.4.4B, center) and the expanded conformation is nearly absent. The difference in the proportion of expanded and compact conformations for the three different distances monitored suggest that the opening up observed for DM-MBP is a local phenomenon and that there is considerable heterogeneity in either the GroEL or SR-EL bound state representing an ensemble of dynamic states. Importantly, ensemble FRET measurements upon dilution of DM-MBP from denaturant demonstrated that the molecule collapses to native-like compactness within milliseconds from FRET efficiency of approx. 10% (Fig.4.5 red). Subsequent addition of GroEL resulted in a rapid expansion with kinetics similar to binding, as shown for DM-MBP(52 – 298) (Fig.4.5, black).

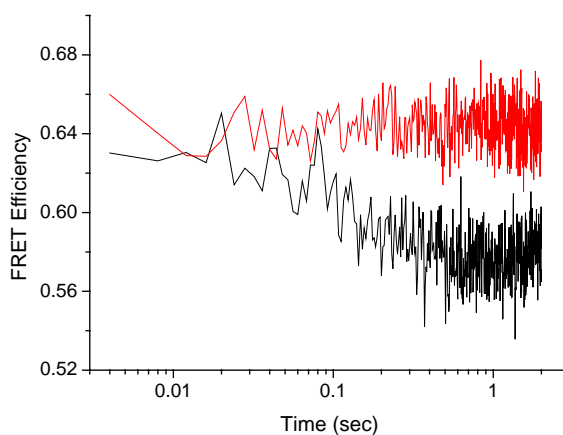


Figure 4.5: Binding Kinetics of the Refolding Protein to GroEL as Observed by Stopped-Flow Analysis. Refolding of DM-MBP(52 – 298) was initiated by 25 fold dilution of DM-MBP(52 – 298), denatured in 1.5M GuHCl, into refolding buffer in a stopped-flow apparatus and is followed by change in FRET efficiency. To check the kinetics of GroEL binding and expansion of a refolding protein, spontaneously refolding DM-MBP(52 – 298) was mixed with buffer only (red) and buffer containing GroEL(black).

4.3 Diffusion Coefficients of Different States of DM-MBP

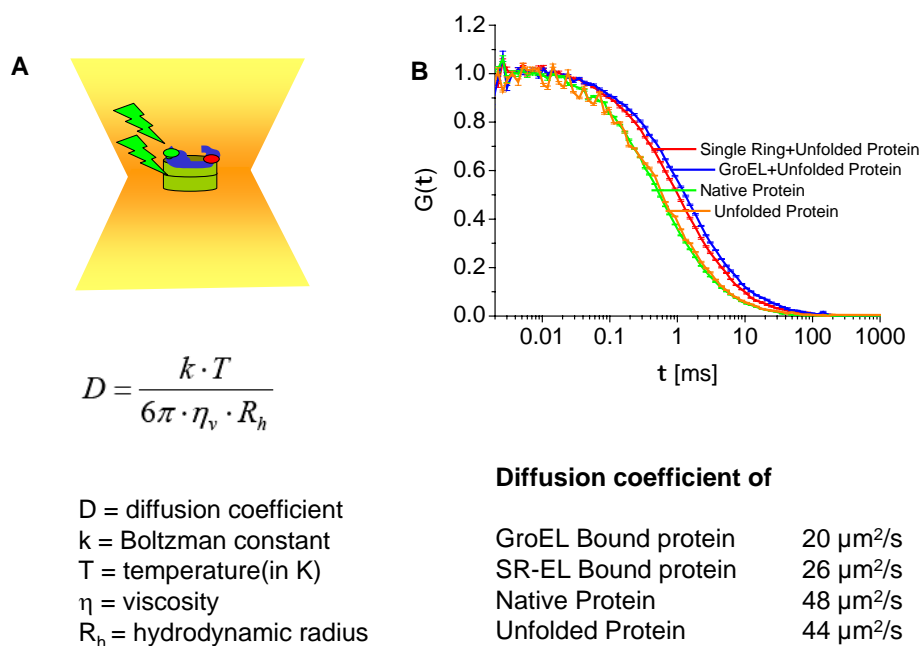


Figure 4.6: Autocorrelation Curve of the Various States of the Labelled Protein for Comparing Diffusion Coefficients. A) Diagrammatic representation of the Fluorescence Correlation Spectroscopy of the Chaperonin-Substrate Complexes. B) The autocorrelation data for Atto-532 labelled DM-MBP in the various states were obtained by fluorescence correlation spectroscopy. The residence times follow a good correlation with the hydrodynamic radii of DM-MBP or DM-MBP/chaperonin complex, with GroEL bound protein exhibiting the highest residence time and the native protein the lowest.

Using Fluorescence Correlation Spectroscopy (FCS), the diffusion coefficient of pre-formed complexes of DM-MBP-GroEL and DM-MBP-SREL as well as for free native and unfolded DM-MBP could be determined. All these curves could be fit to a single component model suggesting that in the measurements with GroEL/SR-EL, most of the observable molecules are bound to GroEL/SR-EL(Fig.4.6).

To rule out that the bimodal distribution observed in the FRET analysis for the GroEL-substrate complex is due to two substrate proteins bound on each side of the GroEL double ring, Fluorescence Cross-Correlation Spectroscopy (FCCS) was performed

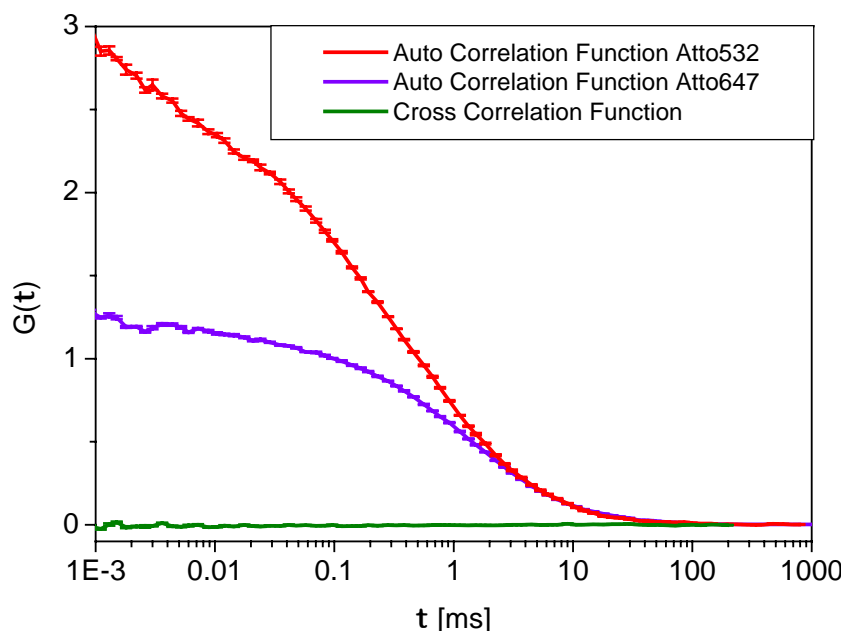


Figure 4.7: The autocorrelations and the cross-correlation of the dyes as observed in FCCS. Auto and cross correlation curve of Single labelled DM-MBP(52C). Two single labelled proteins, 1.5nM DM-MBP(52C) labelled with Atto-532 (DM-MBP(52C)-Atto532) and 1.5nM DM-MBP(52C) labeled with Atto-647N (DM-MBP(52C)-Atto647N) were unfolded together in 3M GuHCl. This was then diluted 50 fold in buffer A containing 3 μ M GroEL.

on a mixture of unfolded donor labelled DM-MBP (52C) and acceptor labelled DM-MBP (52C) with GroEL. If there are two differently labelled proteins each bound onto either side of the tetra-decamer, they will diffuse simultaneously through the probe volume and a cross correlation between the signal of the two dyes should be observed. The absence of cross-correlation between the two dyes indicates that the proteins did not co-migrate during diffusion, ruling out the possibility of two substrate molecules binding to one molecule of GroEL (Fig.4.7).

Inherent heterogeneity in GroEL preparations has been reported (Panda and Horowitz, 2002) and to rule out the possibility of heterogeneity in preparation being the reason for obtaining a bimodal distribution with GroEL, experiments were repeated with different batches of GroEL preparation which exhibited identical distributions.

4.4 Anisotropy Decay for Various DM-MBP Labelled Proteins

It is known that the relative orientation of the donor and the acceptor dyes plays an important role in FRET (Dale, 1979). To judge the conformational flexibility of the dyes, time resolved anisotropy decays were obtained for all the donor and acceptor labelled proteins in complex with GroEL (Tabulated 4.8) which exhibit fast phase of anisotropy decay compared to the excited state lifetime of the dye at all the positions. This suggests that there is considerable flexibility of the dye molecules during its excited state lifetime and thus rules out the possibility that a biased orientation of acceptor or donor resulted in the bimodal distributions or influences the measured FRET efficiencies.

Position	Dye	ro	α_f	τ_f (ns)	α_s	τ_s (ns)
30	Atto-532	0.40	0.09	0.8	0.1	7.5
30	Atto-647N	0.40	0.06	0.15	0.04	7.5
52	Atto-532	0.37	0.07	0.67	0.11	5.38
175	Atto-532	0.36	0.044	0.86	0.11	7.6
298	Atto-647N	0.40	0.043	0.28	0.045	4
312	Atto-532	0.40	0.1	1	0.09	7
312	Atto-647N	0.40	0.03	0.2	0.07	20

Figure 4.8: Anisotropy values at the different engineered Csy positions for Atto-532 and Atto-647N. The table reports fundamental anisotropy (ro), fast (τ_f) and slow (τ_s) phase of anisotropy decay and their respective amplitudes (α_f and α_s) at the different positions.

4.5 Polyproline Fusion Protein - A Spectroscopic Ruler.

Control experiments were performed with a 10 residue poly-proline ruler at the C-terminus of DM-MBP. Two cysteines were positioned at the N-terminus and C-terminus of the poly-proline stretch and were labelled with the same donor and acceptor dyes (Fig.4.9). The two positions were labelled stoichiometrically by addition of both Atto-532 and Atto-647N at 1 : 1 molar concentrations. PIE-SpFRET was used to analyze only double labelled molecules in the heterogeneous population. Both the samples (GroEL or SR-EL) show a single distribution with similar peak position and width, indicating that

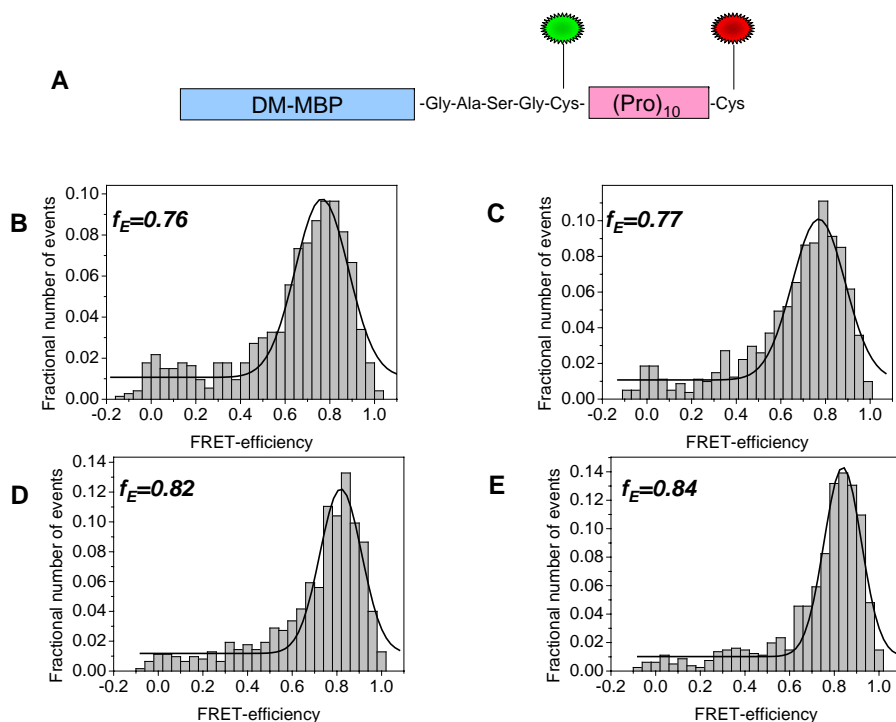


Figure 4.9: Control Experiments with Sp-FRET Using 10 Residue Polyproline Ruler. A) The construct of DM-MBP-Cys-(Pro)₁₀-Cys was designed as shown schematically. The preformed polyproline-chaperonin complex was mixed with ATP or ADP with the final concentration of 2mM of nucleotide, 3 μ M of chaperonin and 60pM of the substrate protein in buffer A. 3nM of the labelled protein (DL-DM-MBP-polyPro) was unfolded in buffer B and then diluted 50 fold in buffer A containing 3 μ M GroEL(B) or 3 μ M SR-EL(C). 2mM ATP (D) or 2mM ATP and 3 μ M of GroES (E) was added to SR-EL/ (DL-DM-MBP-polyPro).

GroEL or SR-EL as such does not affect the distribution. The peak efficiency did not show any significant difference in refolding buffer. To rule out the possibility that either addition of ATP or the encapsulation of the protein causes a significant change in the properties of the dyes, measurements were also done under those conditions. Near identical distribution confirmed that efficiency measured for the different states are relevant.

4.6 Conformational Dynamics of GroEL Bound Substrate Upon Nucleotide Binding

It has been earlier proposed that GroEL brings about unfolding of substrate upon nucleotide binding, facilitating removal of kinetic traps and thus resulting in more efficient refolding (Shtilerman et al., 1999). To test this hypothesis, the bound state conformation of DM-MBP cysteine mutants double labelled with FRET pairs were used to probe the conformational changes upon addition of nucleotides. DM-MBP(52 – 298) GroEL complexes were performed in the presence of ATP to investigate these effects. With 2mM ATP, the bimodal distribution of DM-MBP (52 – 298) is shifted to a single distribution at high FRET efficiency (Fig.4.10B). Similarly compact state is also obtained when ADP (Fig.4.10C) is used instead of ATP. Since GroEL hydrolyses ATP much faster than the

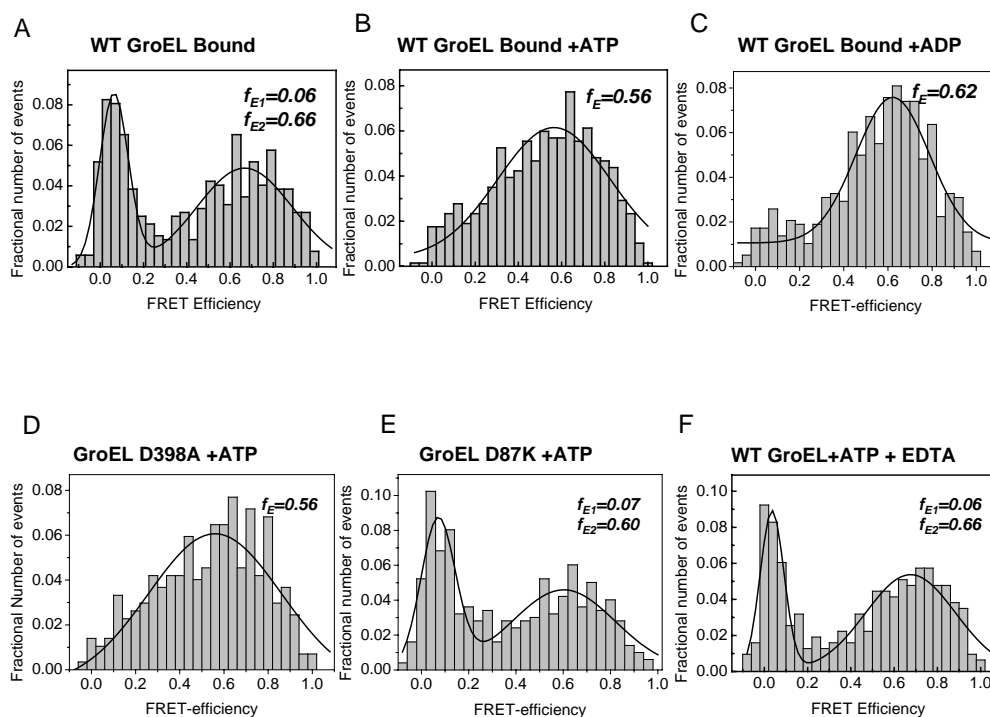


Figure 4.10: Single Molecule FRET Distribution of Chaperonin Bound Protein Complex in Presence of Nucleotide. The preformed substrate-chaperonin complex was mixed with ATP or ADP with the final concentration of 2mM of nucleotide $3\mu\text{M}$ of chaperonin and 60pM of the substrate protein in buffer A. Substrate-capture-complex of DM-MBP(52 – 298) and WT GroEL(A) with ATP (B) or ADP (C). DM-MBP(52 – 298)complexed with GroEL D398A (D) or GroEL D87K (E); or with WT GroEL,ATP and EDTA (F)

time-scale of the measurements, an ATP hydrolysis defective mutant of GroEL (GroEL-

D398A) was used to delineate the effect of ATP binding from hydrolysis and it was found to be identical to that of WT-GroEL (Fig.4.10 D and A respectively). Therefore WT-GroEL was used for further experiments. Control experiments were performed with an ATP binding deficient mutant of GroEL, GroEL (D87K), which does not exhibit similar shift in conformation upon addition of ATP or ADP (Fig.4.10F) proving that ATP binding is essential for the shift in conformation. Since GroEL-bound DM-MBP (52 – 298) shows similar distributions when bound to GroEL or GroEL (D398A) in the presence of ATP, and in the presence of ADP, it is concluded that ATP or ADP binding and not hydrolysis is important for this shift in conformation.

This shift is reversible upon dilution of nucleotide or by addition of EDTA indicating that the structure of the substrate protein is dynamic and interconvertible between the expanded and the compact conformation. Since SR-EL in nucleotide-free state exhibits only the more compact state, no significant difference could be observed upon binding of ATP or ADP.

DM-MBP(30 – 312), which shows higher percentage of extended state when bound to GroEL as compared to that from SR-EL. This may be because of lower affinity of SR-EL for DM-MBP which in turn may be due to different orientation of the apical domains. The proportion of this expanded state also decreases, upon addition of ATP and ADP (Fig.4.11, left). In case of SR-EL/DM-MBP(175 – 298) there is no distinct difference as compared to that of GroEL bound distribution. Interestingly, it is observed that the distributions are different in ADP and ATP bound states in both GroEL and SR-EL, indicating that GroEL has distinct conformations when bound to ATP or ADP and this is reflected in the subtle differences in conformation of the bound protein. This conformational difference may result from the load presented by the substrate protein on the apical domain of GroEL where the binding energy of the γ -phosphate group is essential for effecting a conformational change in presence of the substrate protein (Motojima et al., 2004).

DM-MBP(175–298) which exhibits compact distribution in GroEL and SR-EL bound form does not undergo any significant change in distribution upon addition of ATP or ADP(Fig.4.11, right). The fact that the conformational distribution of DM-MBP can be regulated by ATP concentration suggests that the two distributions result as a function of GroEL conformation, which is known to be regulated by ATP or ADP. The binding affinity of GroEL for unfolded protein is higher in nucleotide-free state than in the nucleotide bound state which in turn is known to be modulated by ATP or ADP induced change in hydrophobicity of substrate binding site (Kim et al., 2005). This would indicate that the expanded state of the unfolded protein results due to the binding of the unfolded

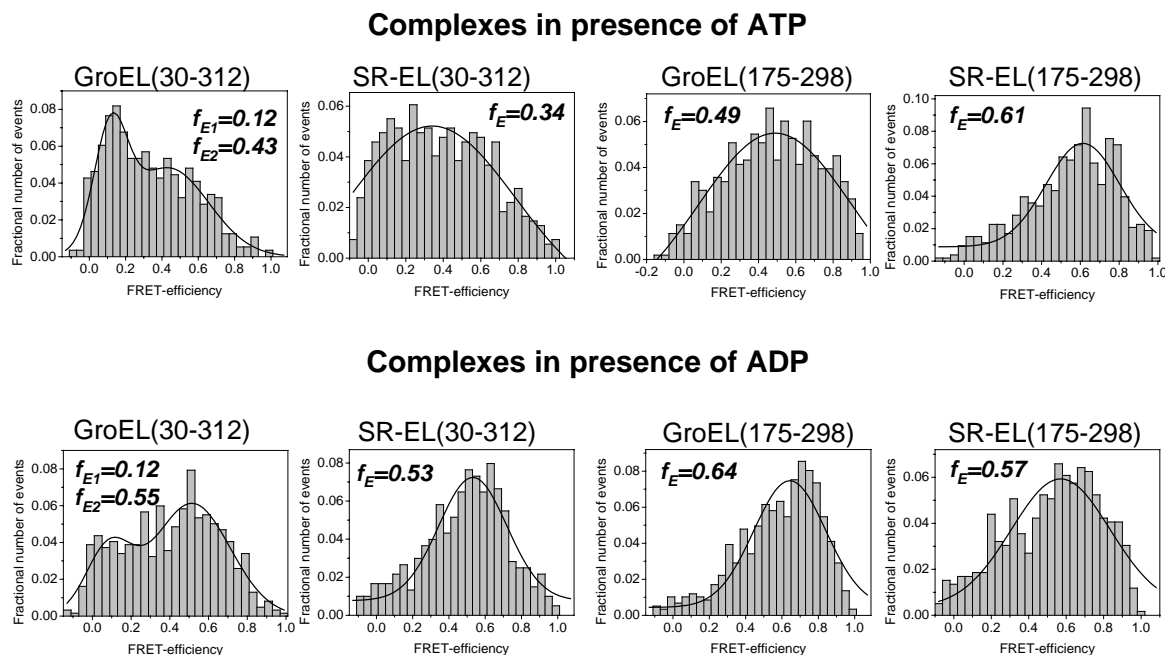


Figure 4.11: Single molecule FRET Distribution of Chaperonin Bound Protein Complex in Presence of ATP and ADP. The preformed substrate-chaperonin complex was mixed with ATP or ADP with the final concentration of 2mM of nucleotide, $3\mu\text{M}$ of chaperonin and 60pM of the substrate protein in buffer A.

protein to high-affinity binding state of GroEL which is absent in SR-EL or nucleotide bound state of GroEL. It is possible that even though we do not observe any expansion event in nucleotide bound state of the chaperonin, expansion is facilitated upon GroES binding to the GroEL-substrate-nucleotide complex.

4.7 Effect of GroES Binding on Chaperonin-Substrate Complex

In order to investigate whether GroES binding to preformed SR-EL/DM-MBP/nucleotide complex results in expansion of the protein, ADP and ATP were used to form the complex. It was observed that the SR-EL/DM-MBP complex in the presence of ADP is able to bind GroES but not refold the protein. This maybe due to a partial release of DM-MBP into the cavity forming a SR-EL/DM-MBP/GroES unproductive ternary complex in the presence of ADP as observed with ensemble FRET studies (see later section). These ternary complexes with ADP and ATP were used to observe if during the course of GroES binding, there is any unfolding event on the apical domain. We saw no signifi-

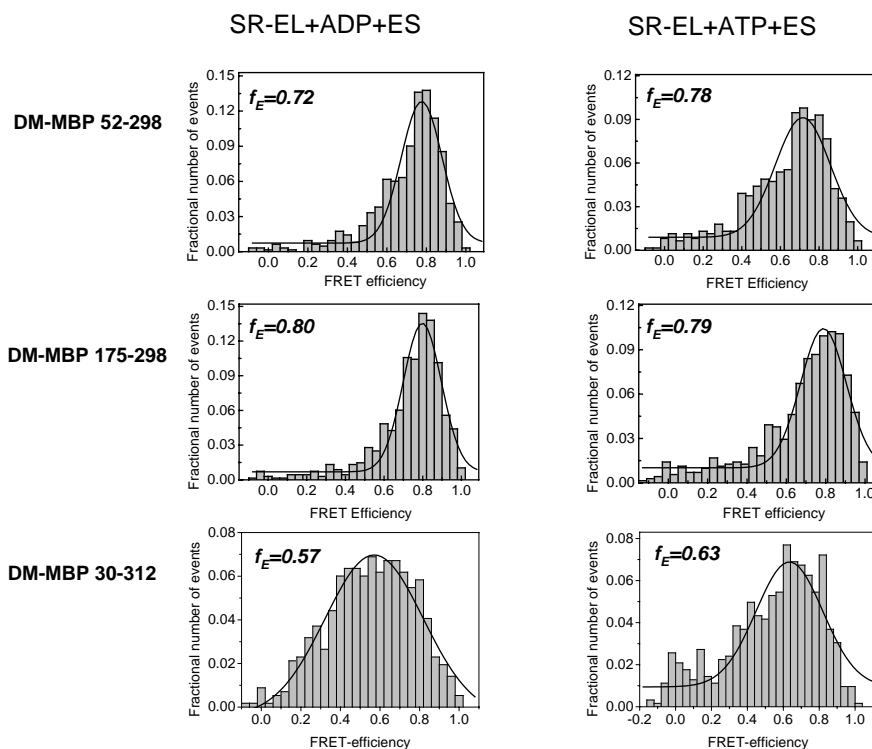


Figure 4.12: Single Molecule FRET Distribution of the Ternary Complex Between SR-EL Substrate Protein and GroES in Presence of ATP or ADP. The preformed substrate-chaperonin complex was mixed with ATP or ADP with the final concentration of 2mM of nucleotide, 3 μ M of chaperonin and 60pM of the substrate protein and 6 μ M of GroES in buffer A.

cant difference in the distribution of FRET efficiency of the SR-EL/DM-MBP/nucleotide complex in case of the different constructs observed (Fig.4.12). Although, upon GroES binding it was observed that the protein undergoes significant compaction.(Motojima et al., 2004).

In the case of DM-MBP, we see a progressive collapse of the bound state onto the GroEL and no subsequent opening up observable by this method and with this time resolution, after GroES binding. Therefore, we monitored the fast kinetics of this process by following ensemble FRET on a stopped-flow apparatus. This technique was exploited to resolve processes in the order of milliseconds which is much faster than what is observable in the Single-Molecule experimental set-up, which has a dead time of around 50sec.

4.8 Stopped-Flow Ensemble FRET Analysis of Nucleotide and GroES Binding

ATP binding precedes the binding of GroES in the GroEL reaction cycle and is known to cause a 5 – 10 degree clock-wise turning motion of the apical GroEL domains (Chen et al., 1994). Moreover, steady-state single molecule FRET experiments showed that DM-MBP assumed native-like compactness in the presence of nucleotide, and upon encapsulation by GroES (Fig.4.12). However, due to the limited time resolution, these measurements could not detect a transient expansion of DM-MBP caused by GroES binding. Therefore we, performed ensemble FRET measurements upon stop-flow² mixing of DM-MBP/SR-EL complexes with ATP or ATP/GroES.

Interestingly, for both DM-MBP(52 – 298) (Fig.4.13, black) and DM-MBP(175 – 298) (not shown), addition of ATP alone resulted in a rapid, transient expansion of the bound protein that was followed by compaction. No such expansion was detected upon addition of ADP (Fig.4.13 red).

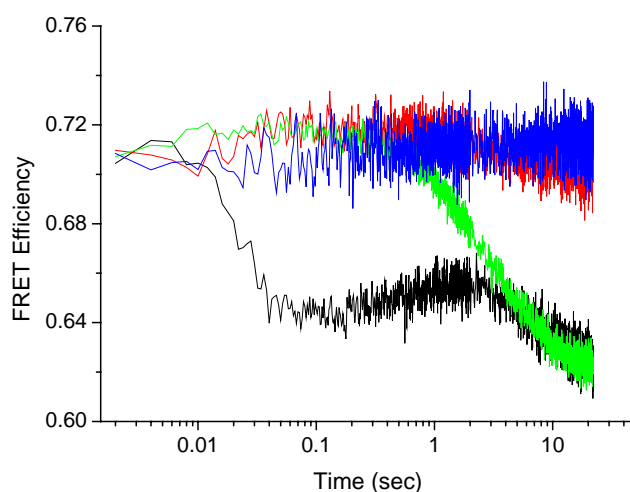


Figure 4.13: Conformational Dynamics of the GroEL Bound Protein in Presence of Different Nucleotides. 100nM DM-MBP(52 – 298)/SR-EL complex was prepared by diluting 5 μ M of unfolded DM-MBP(52 – 298) in buffer containing 1 μ M of SR-EL. This was taken in one syringe of the stopped flow apparatus and mixed with an equal volume of buffer(blue), ADP(red), ATP(black) or ADP-AlFx(green) to obtain the conformational change in DM-MBP upon binding of nucleotide to SR-EL. These efficiencies plotted are not corrected for presence of inactive acceptors in the double-labelled protein.

²Stopped-flow experiments were done together with Kausik Chakraborty.

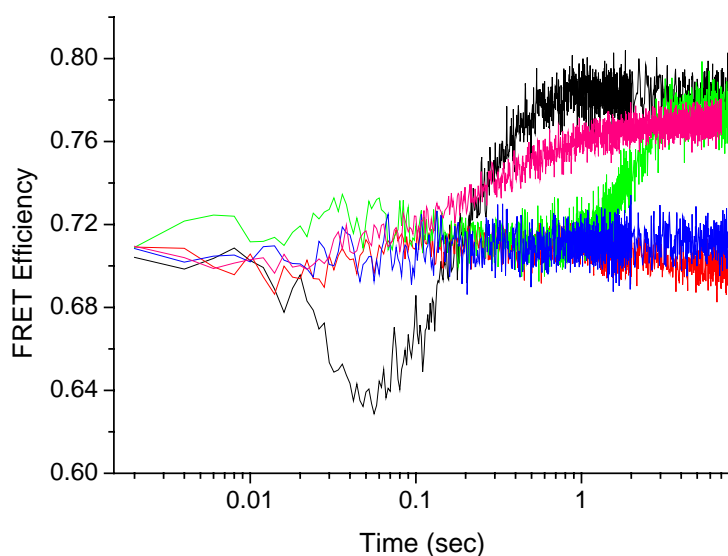


Figure 4.14: Conformational dynamics of the GroEL bound protein in presence of different nucleotides and GroES. 100nM of DM-MBP(52–298)/SR-EL complex was mixed with equal volume of buffer(blue), ADP and GroES(red), ATP and GroES (black) and ADP-AIFx and GroES(green) to obtain the conformational change of DM-MBP upon encapsulation. To obtain the effect of GroES binding to preformed DM-MBP(52–298)/SR-EL /ATP complex, the DM-MBP(52–298)/SR-EL/ATP complex was preformed and mixed with equal volume of GroES (pink).

GroES binding to GroEL is associated with a dramatic 60 degree clock-wise turn and an upwards motion of the GroEL apical domains, which may cause further stretching of the substrate (Chen et al., 1994). However, GroES binding to DM-MBP/GroEL/ATP complexes resulted in a rapid compaction of the bound protein ($t_{1/2}$ approx. 500 msec), concomitantly with its release into the GroEL cavity (Fig.4.14, pink). On the other hand, when GroES was mixed with DM-MBP/GroEL complexes and GroES binding was triggered by addition of ATP, transient expansion of the protein followed by compaction was again observed (Fig.4.14, black). These changes in efficiency can be taken as the corresponding change in conformation of the protein itself and not an indirect effect because of the change in the spectral properties of the dyes during the process, as measurements with poly-proline construct showed no interference in determining FRET efficiency of the bound and encapsulated protein in case of GroEL or SR-EL.

Thus, in the case of DM-MBP, local unfolding triggered by ATP-mediated apical domain movement is followed by a compaction, which is completed upon GroES-mediated displacement of bound protein into the chaperonin cage. To address the question whether ATP-mediated structural expansion of DM-MBP is a mechanistic requirement for sub-

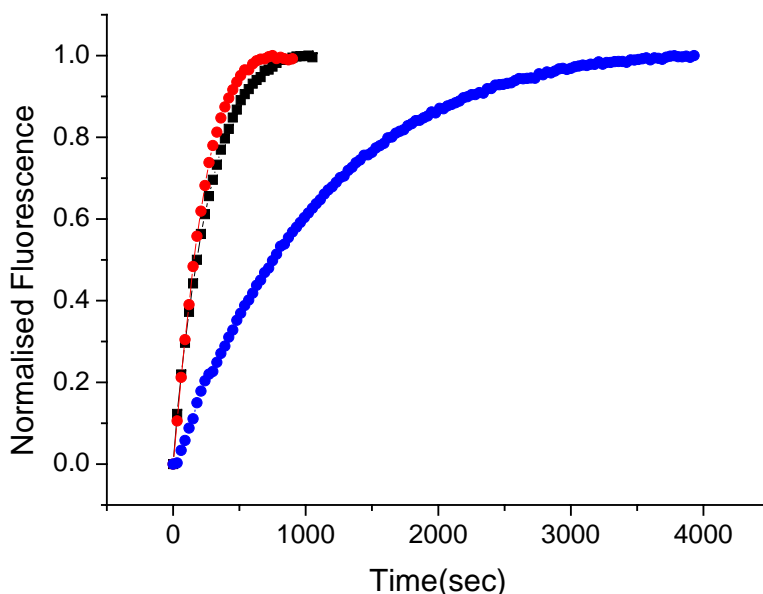


Figure 4.15: SR-EL and GroES Mediated Refolding Kinetics of DM-MBP in the Presence of ATP and ADP-AIFx. Spontaneous refolding (blue) was initiated by diluting the unfolded protein in buffer A; SR-EL mediated refolding of DM-MBP was initiated in presence of SR-EL/GroES/ATP (red) or SR-EL/GroES/ADP-AIFx (black) and monitored by Trp fluorescence of DM-MBP.

sequent accelerated folding, we took advantage of the finding that a mixture of ADP and AIFx, (Inobe et al., 2003) (made in situ by mixing NaF and $KAl(SO_4)_2$) mimics the transition state of ATP hydrolysis (Chaudhry et al., 2004). ADP-AIFx produced a similar degree of expansion of GroEL-bound DM-MBP as ATP but with approx. 50-fold slower kinetics (Fig.4.13, green).

As a consequence, simultaneous addition of ADP/AIFx and GroES circumvented the transient unfolding step, resulting in immediate compaction (Fig.4.14, green). Importantly, under these conditions folding of DM-MBP within the GroES-enclosed cage of SR-EL occurred with similar accelerated kinetics as in the presence of ATP (Fig.4.15). Similarly, when ATP and GroES was added to SR-EL bound protein in the presence of 750mM $(NH_4)_2SO_4$, the transient unfolding before the encapsulation triggered compaction is not seen (Fig.4.16). But even under these circumstances, the rate of refolding of DM-MBP is the same as that carried out under usual buffer conditions. Thus, ATP-mediated structural expansion is not a requirement for subsequent rapid folding, at least in the case of DM-MBP.

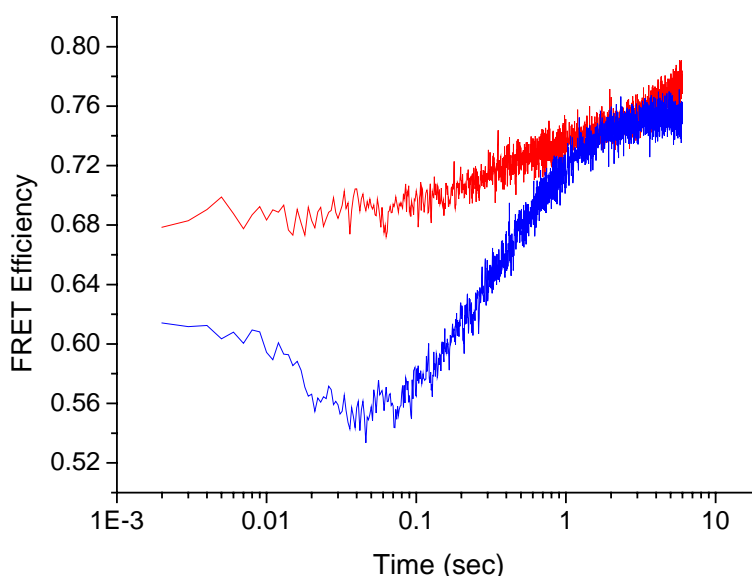


Figure 4.16: Conformational Dynamics of GroEL Bound Protein in the Presence of $(NH_4)_2SO_4$ Induced with ATP and GroES. SR-EL/GroES/ATP mediated refolding of DM-MBP was initiated in the absence (blue) or presence of $(NH_4)_2SO_4$ (red) as monitored with FRET .

4.9 GroEL Cycling and the Bound-State

It has also been proposed that iterative annealing of substrate protein might be able to facilitate the refolding of substrate protein (Todd et al., 1996). To see how GroEL cycling and rebinding affects the distribution of the bound state, GroEL assisted refolding was initiated with ATP and then stopped at specific time points by addition of apyrase and EDTA as described earlier (Brinker et al., 2001). DM-MBP-GroEL complexes were purified as described in Materials and Methods and analyzed by SpFRET. They were seen to have the same distribution at different time points after the initiation of the refolding reaction (Fig.4.17). The small increase in the fraction of compact conformation compared to binding to nucleotide free GroEL may be attributed to the presence of small amount of ADP even after the treatment but the sample definitely does not contain a higher proportion of expanded conformation.

More importantly, no observable change in conformational distribution is observed between reactions stopped at 90 s and 300s. Considering that substrate protein would have undergone 3-fold more cycling events at the end of 300s that at the end of 90sec, would indicate that distribution of conformation of the bound state does not depend on the number of times the protein has been through the GroEL cycle.

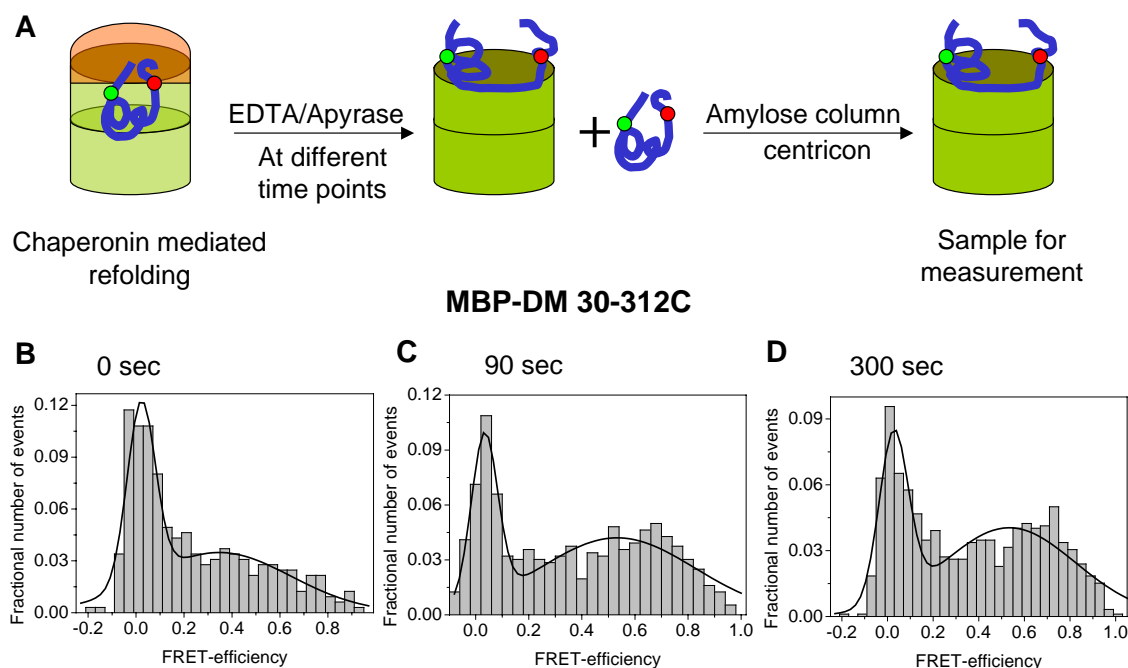


Figure 4.17: Sp-FRET Distribution of GroEL DM-MBP(30–312) Complexes as a Function of Time. (A) Graphical representation of the process of sample preparation. Native and unbound protein that may be formed during the course of reaction was removed using an amylose resin column which would bind native protein and 100kD centricron to remove any DM-MBP that remains unbound to GroEL. (B) SpFRET distributions obtained as mentioned earlier. After initiating GroEL/GroES/ATP mediated refolding of DM-MBP(30–312), the reaction was terminated using EDTA and apyrase at (C) 90sec and (D) 300sec.

Thus, during successive chaperonin cycles, not-yet folded protein may revert to the locally expanded conformational ensemble of the initial bound state. To test whether unfolding upon re-binding is sufficiently rapid to precede encapsulation by GroES, a collapsed DM-MBP(52–298) folding intermediate was mixed with the asymmetric GroEL/GroES/ADP complex, the natural substrate acceptor state in the reaction cycle (Rye et al., 1999). A conformational expansion was observed by FRET to occur at a rate comparable to unfolding upon binding to unliganded GroEL (Fig.4.18). Thus, transient unfolding may occur upon substrate re-binding during successive chaperonin cycles, possibly re-setting kinetically trapped intermediates to a higher position in the energy landscape. However, in the case of DM-MBP, this process does not contribute to the 10-fold rate acceleration of folding, because fully accelerated folding occurs with a single round of encapsulation in SR-EL/GroES, independent of iterative binding cycles. This would suggest that for DM-MBP iterative annealing does not lead to any further conformational change of GroEL bound DM-MBP.

In an *in vivo* scenario, GroEL is a down stream chaperone and in the view of the

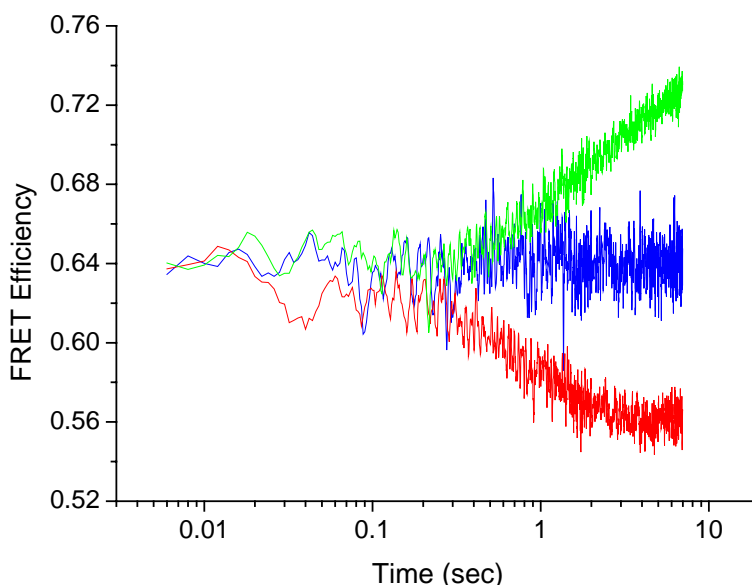


Figure 4.18: Kinetics of the Conformational Rearrangement of DM-MBP (52 – 298) Upon binding "GroEL-Bullet" To obtain the kinetics of the conformational rearrangement of DM-MBP(52 – 298) upon binding to GroEL in a more physiological situation, $1\mu\text{M}$ GroEL/GroES/ADP bullets were reconstituted and was mixed with 100nM of DM-MBP(52 – 298) in the presence (green) or absence (red) of ATP or with buffer alone(blue).

"Pathway Model" of chaperone function in the cell, the substrate most likely is handed down from some of the up-stream chaperones such as DnaK/J system. Therefore, to investigate whether under such circumstances the substrate undergoes any expansion upon binding to GroEL, E_{FRET} of the DM-MBP(52-298) was measured when bound to DnaK/J system in the presence of ATP, then GroEL and GrpE were added to facilitate transfer to GroEL and the E_{FRET} measured and compared to that of the spontaneous intermediate, bound to DnaK/J, bound to GroEL, and from DnaK/J transferred to GroEL (Fig.4.19). It can be seen that upon transfer from the Hsp-70 system, the substrate protein undergoes expansion upon binding GroEL, indicating that this expansion may be important for some substrate proteins to be pulled out from kinetically trapped conformation.

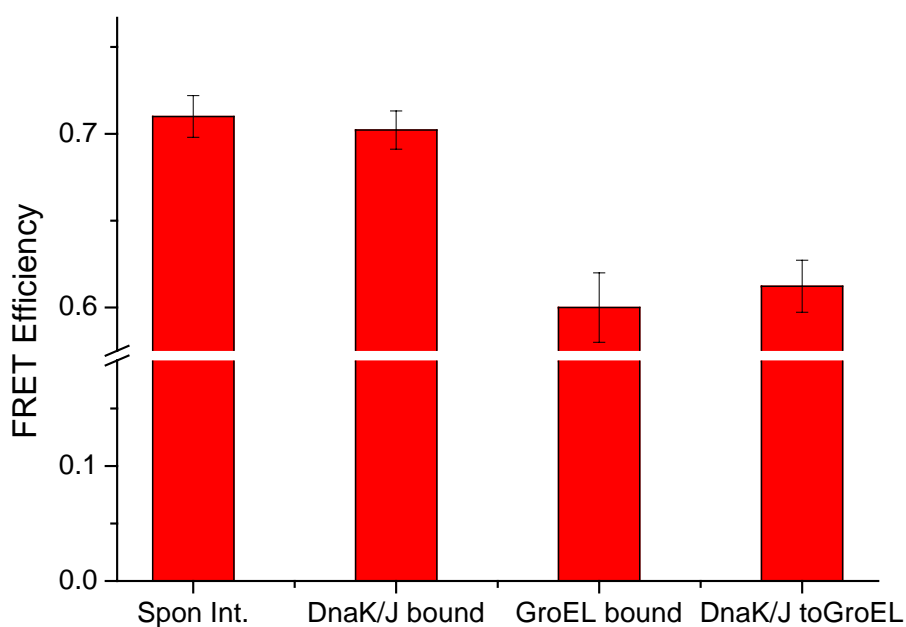


Figure 4.19: FRET efficiencies of DM-MBP(52 – 298) when bound to GroEL or DnaK/J system or a folding intermediate in the spontaneous folding pathway. To obtain the steady state E_{FRET} of DM-MBP(52 – 298), denatured protein was diluted into Buffer A alone or with GroEL ($3\mu\text{M}$) or DnaK/J ($1.25\mu\text{M}$ and $0.6\mu\text{M}$ respectively). To facilitate transfer from DnaK/J to GroEL $1.25\mu\text{M}$ of GrpE was added along with GroEL.

4.10 Binding and Segmental Release of Substrate by GroEL

Steady-state anisotropy measurements³ with the Atto532 dye attached to single cysteines positioned throughout the DM-MBP sequence were performed to obtain insight into how GroEL binding restricts the mobility of the protein. Since the lifetime of Atto532 is on the order of 4 ns or less, the steady-state anisotropy of labelled DM-MBP is affected most by the restriction of segmental mobility of the protein resulting from binding to GroEL, which has a rotational correlation time orders of magnitude larger than the lifetime of the dye. Segmental mobility differed substantially for different regions of the GroEL-bound protein (Fig.4.20). A flexible N-terminal segment (first 40 residues) is

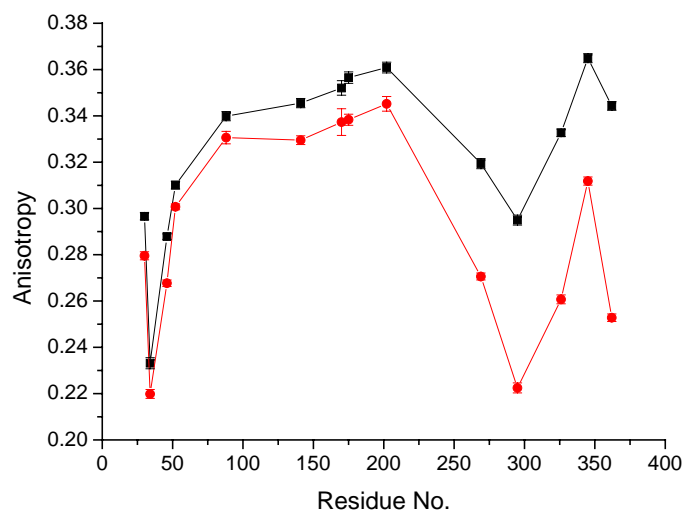


Figure 4.20: Segmental Mobility of Different Regions of the Labelled DM-MBP When Bound to GroEL. The steady-state anisotropy of the single cysteine mutants of DM-MBP labelled with (Atto-532) were measured when the protein was bound to GroEL (filled black Squares with black line) and with addition of ATP (filled red circles with red line) to the GroEL bound complex.

followed by a highly immobile region (next 200 residues). This segment also contains the most hydrophobic regions of DM-MBP, possibly constituting the hydrophobic core of the folded protein. In contrast, the following 30 residues are more mobile, followed by a mobility-restricted C-terminal segment.

Binding of ATP to GroEL is known to decrease the affinity of the chaperonin towards unfolded proteins. In order to verify the same for DM-MBP, equilibrium dissociation constants for the binding of DM-MBP to GroEL, GroEL/ATP and SR-EL were obtained (Fig.4.21). Fraction of DM-MBP bound to chaperonin was obtained by the amplitude of

³anisotropy experiments were done by Kausik Chakraborty

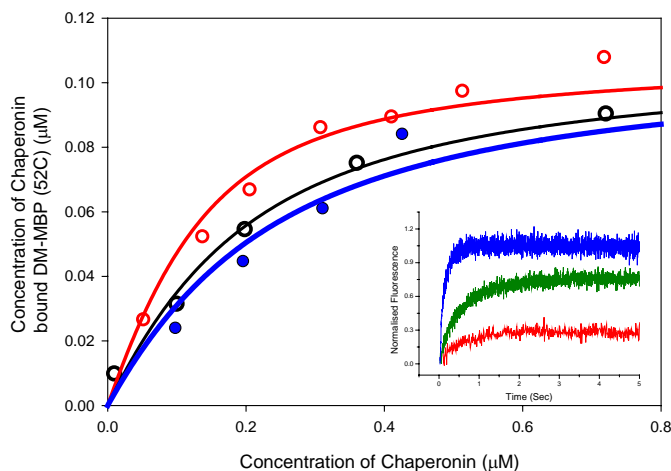


Figure 4.21: Determination of K_D of the chaperonin for the substrate protein. 100nM DM-MBP(52C) was taken in a syringe and then mixed with different concentrations of GroEL (red), SR-EL (blue) and GroEL with ATP (black). The bound proportion of the protein was estimated by fitting the curve and taking the amplitude change as a function of concentration of the substrate-chaperonin complex formed. The proportion of bound protein was estimated from the amplitude change (inset) of the labelled protein when mixed with different concentration of chaperonin.

fluorescence change of DM-MBP(52C) labelled with Atto-532 upon binding to GroEL. Equilibrium dissociation constant was obtained by quantifying the fluorescence amplitude change as a function of chaperonin concentration, and fitting the resultant graph to the equation:

$$B = \frac{(c + x + K_D) - \sqrt{(c + x + K_D)^2 - 4 \times x \times c}}{2}$$

where B is the concentration of the bound labelled DM-MBP to the chaperonin, c is the total concentration of chaperonin, x is the total concentration of DM-MBP labelled with Atto-532 and K_D is the equilibrium dissociation constant. The equilibrium dissociation constant as expected was lower for GroEL (60 nM) than for GroEL/ATP (130nM) and SR-EL (170nM). This shows that the binding affinity of GroEL for unfolded protein is significantly decreased in presence of ATP. In line with this thought, we investigated whether ATP dependent conformational changes in GroEL modulate the structure of the bound substrate in a global manner. We have seen that ATP induced structural changes in the chaperonin manifests itself in structural rearrangements in the substrate protein in a position specific manner.

Therefore, we wanted to probe whether the release of the substrate protein from the apical domain is concerted or it follows a segmental release in a position specific manner.

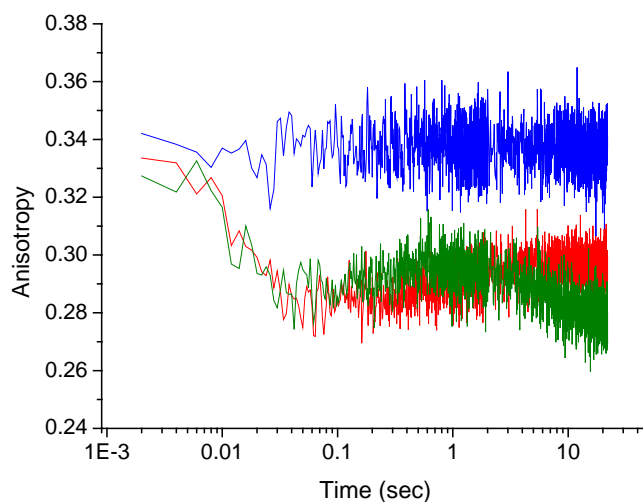


Figure 4.22: Segmental Release of the Protein as Studied by Steady-State Anisotropy using a Stopped-Flow Apparatus. 100nM DM-MBP(345)/SR-EL complex (B) was prepared by diluting $5\mu\text{M}$ of unfolded protein in buffer containing $1\mu\text{M}$ of SR-EL. This sample was loaded in one syringe of the stopped flow apparatus and mixed with an equal volume of buffer (blue), ATP(red), or ATP and GroES (green) to obtain the conformational change in DM-MBP upon binding of nucleotide to SR-EL.

Interestingly, addition of ATP had only little effect on the mobility of the hydrophobic regions, but strongly increased the mobility of the more flexible, loosely bound segments 5.1.

Stop-flow mixing experiments with DM-MBP labelled at position 345 in the C-domain showed that this segmental release occurred rapidly within 100 milliseconds upon ATP addition and was not accelerated by the simultaneous presence of GroES (Fig.4.22). In contrast, the tightly bound segment in the N-domain was only released upon addition of GroES (Fig.4.23). Significantly, this occurred at a time scale slower than release of the C-domain segment. Thus, hydrophobic collapse of the protein upon encapsulation by GroES follows a step-wise mechanism in which less hydrophobic regions are released from GroEL prior to more hydrophobic ones, reversing the order of burial of hydrophobic segments during spontaneous folding. Indeed, a rapid, indiscriminate collapse was observed for DM-MBP upon dilution from denaturant, as indicated by similar rates of compaction for the various FRET pairs.

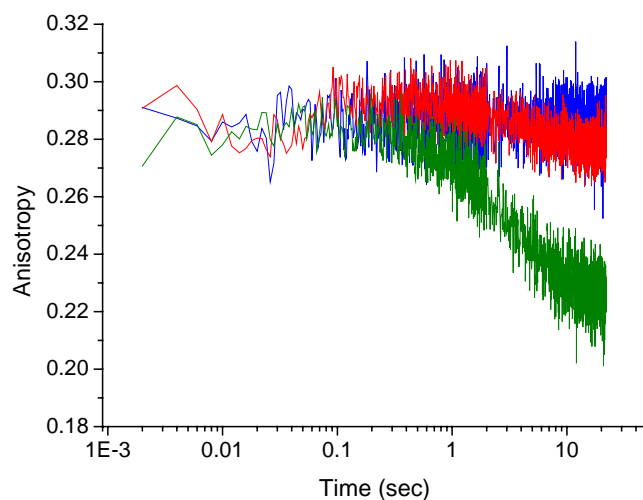


Figure 4.23: Segmental Release of the Protein as Studied by Steady-State Anisotropy using a Stopped-Flow Apparatus. 100nM DM-MBP(52)/SR-EL complex was prepared by diluting 5 μ M of unfolded protein in buffer containing 1 μ M of SR-EL. This sample was loaded in one syringe of the stopped flow apparatus and mixed with an equal volume of buffer(blue), ATP(red), or ATP and GroES(green) to obtain the conformational change in DM-MBP upon binding of nucleotide to SR-EL.

4.11 GroEL-GroES assisted and Spontaneous Folding Pathways

FRET efficiencies were calculated for 3 pairs of positions to estimate the compaction and conformation of the protein in general while it is inside the GroEL cavity. One of the positions (298), was kept fixed as it could be protected from labelling upon addition of maltose, and the other positions (52, 345 and 175) were chosen so that the spontaneous refolding rates obtained are not affected upon mutagenesis to cysteine or upon labelling of the cysteines. Double labelling was performed as described in materials and methods and double labelling was confirmed by life-time analysis of the labelled protein which exhibited less than 5% component from single labelled species.

The FRET efficiencies for the different FRET pairs of the folding intermediate were obtained by extrapolating the refolding curve to the initial time point. It is observed that the protein is highly compact with some positions being as compact as the native state within 10 seconds of initiating the refolding reaction. Similar compaction is also seen for spontaneous folding indicating that DM-MBP undergoes a fast hydrophobic collapse resulting in a compact state both in spontaneous and GroEL/SR-EL assisted folding. The efficiencies for the intermediate obtained with DM-MBP(52 – 298) (Fig.4.24A) FRET

pair were similar for spontaneous refolding and GroEL/ES/ATP mediated folding, indicating that confinement of the protein in the GroEL cavity does not change the structure of the intermediate in this case significantly. For DM-MBP(345 – 298) we see a difference (Fig.4.24B) in the efficiencies obtained for the SR-EL/ES/ATP as compared to the spontaneous refolding. For this distance, the refolding intermediate formed in SR-EL is slightly more expanded than the refolding intermediate formed spontaneously but the native states have similar distances. SR-EL/ES/ATP mediated refolding for these FRET pairs are identical to that of GroEL/ES/ATP mediated refolding. The efficiency of the

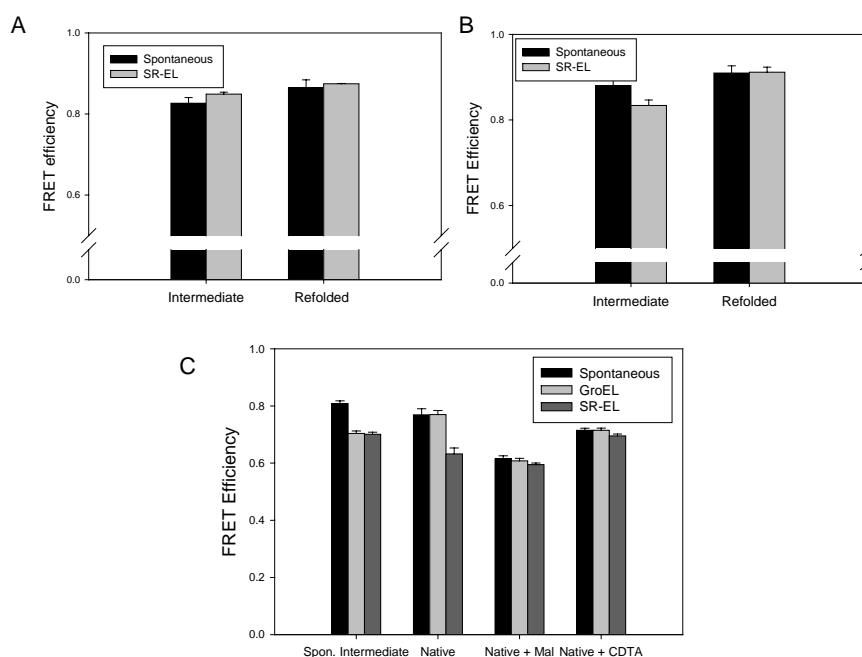


Figure 4.24: Ensemble FRET Efficiency of the Intermediate and the Native State Formed in Case of Spontaneous Refolding, GroEL and SR-EL Mediated Refolding. Refolding of GuHCl-denatured DM-MBP in buffer B ($12.5\mu\text{M}$) at 25°C upon 50-fold dilution into reactions containing either buffer A alone (spontaneous, black), buffer A with $1.0\mu\text{M}$ SR-EL (grey) or $0.5\mu\text{M}$ GroEL (light grey) and $1.0\mu\text{M}$ GroES, 2 mM ATP, for DM-MBP(52 – 298)(A), DM-MBP(298 – 345)(B) and DM-MBP(175 – 298)(C).

intermediate of DM-MBP(175 – 298) (Fig.4.24C) pair shows a more expanded conformation for the intermediate in the chaperonin cavity compared to the spontaneous refolding intermediate. Though this might not provide a realistic estimate of the distance or the dimension of the intermediate due to a change in R_o , it indicates that the conformation of the intermediate is different inside the cavity. Maltose binding was monitored to confirm that the protein reached the native state in case of both GroEL mediated and spontaneous folding as shown previously.

The FRET efficiency for the intermediate as well as the refolded protein shows a difference in FRET efficiency in SR-EL as compared to spontaneous reaction or GroEL refolded protein. This could be attributed to the fact that the refolded protein is trapped inside the SR-EL cavity with an altered conformation, and the efficiency reaches the native state value when CDTA is added to release the protein from the SR-EL cavity (Fig.4.24C). This is also confirmed by SpFRET.

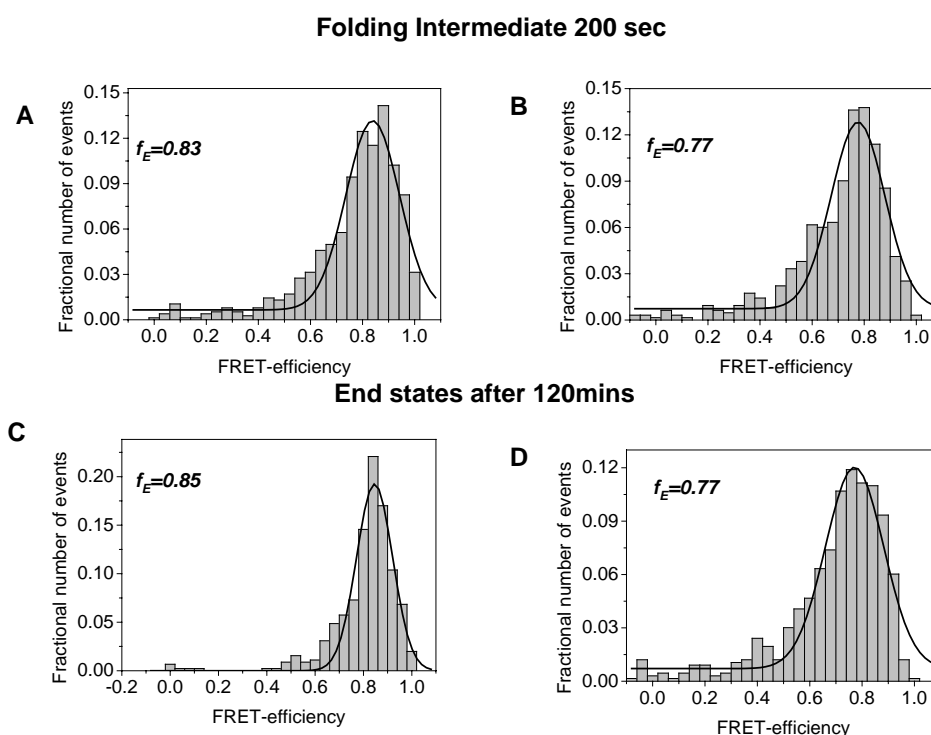


Figure 4.25: Sp-FRET Efficiency Distribution of the Intermediate and the Native State Formed in Case of Spontaneous Refolding and SR-EL Mediated Refolding. Sp-FRET distribution of the refolding intermediate of DM-MBP(175 – 298) formed during the first 200 seconds of spontaneous (A) and SR-EL/GroES/ATP mediated (B) refolding reaction were obtained. The Sp-FRET distributions of the spontaneously refolded protein (C) and SR-EL/GroES/ATP refolded protein trapped in the SR-EL cavity (D) were also obtained.

The E_{FRET} values from single-molecule measurements for the folding intermediates at 200 seconds after initiating folding, when 10% (spontaneous) and 40% (assisted) of total DM-MBP have reached native state (Fig.4.25 A and B), were overall similar for spontaneous and chaperonin mediated refolding, but the SR-EL-enclosed protein was slightly more expanded than the free form. Similarly, the end-state of folding reached

inside the chaperonin cavity was somewhat more expanded than the free native state (Fig.4.25 C and D).

Though the conformation of the folded protein inside the SR-EL cavity is altered, it is observed that upon addition of maltose there is a change in efficiency and it reaches the same value obtained for maltose-bound native protein, indicating that the folded form is able to bind maltose and is in fact the native state with local conformational differences with the native protein in solution. This raises the possibility that the GroEL cavity has the ability to effect local conformational changes upon the intermediate as well as the native state.

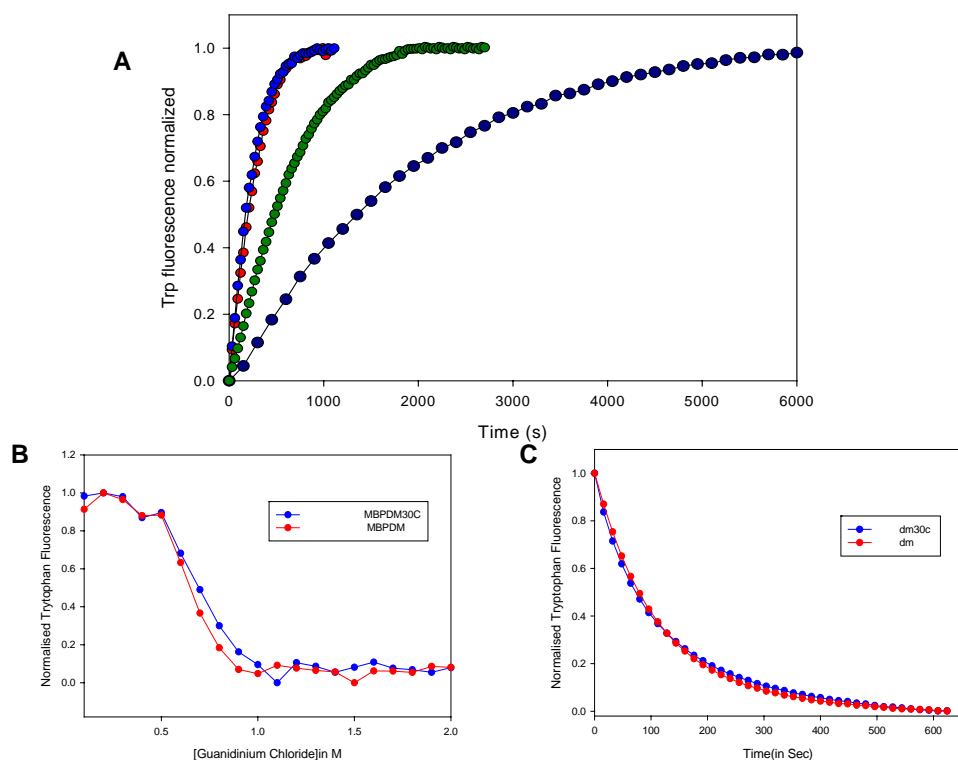


Figure 4.26: Rate of Spontaneous and GroEL Assisted Refolding of DM-MBP and DM-MBP(30–298). Effect of mutations D30C and P298C on the rate of refolding (A) as seen in spontaneous refolding of DM-MBP(30 – 298) (green) as compared to DM-MBP (dark-blue) and SR-EL mediated refolding of DM-MBP(30 – 298) (royal-blue) and DM-MBP (red). (B) Isothermal chemical denaturation curves to compare the thermodynamic stability of DM-MBP and DM-MBP(30 – 298). 250nM of either protein was incubated in buffer A containing different concentrations of GdnHCl for 4 hours at 25°C before obtaining fluorescence readings. Fluorescence was measured at 350nm (slit width 5nm) with excitation at 295nm (slit width 2nm) (C). Unfolding of DM-MBP and DM-MBP(30 – 298) at 1.5M GdnHCl as observed by Trp fluorescence.

If the folding pathway is altered in the GroEL cavity compared to the pathway during spontaneous refolding the rates would be hypothesized to behave differentially to rate perturbing mutations or external conditions. The mutations D30C and P298C on DM-MBP was found to accelerate the spontaneous refolding rate by approx. 2.5 fold (Fig.4.26A) without a change in the stability of the protein (Fig.4.26 B) or the unfolding rate (Fig.4.26C). There is a slight shift in C_m though the ΔG values obtained from the fits (4.7kcal /mole for DM-MBP and 4.1 kcal/mole for DM-MBP(30 – 298)), indicating that rate enhancement of refolding of DM-MBP(30 – 298) compared to DM-MBP does not result from increased stability of the protein. DM-MBP and DM-MBP(30 – 298) exhibit identical unfolding rates, which indicates that the faster rate of folding of DM-MBP(30 – 298) is not resultant from a slower unfolding rate of the protein. Identical rates of unfolding of DM-MBP(30 – 298) and DM-MBP are also obtained for other concentrations of GdnHCl.

This indicates that D30 and P298 play a role in decelerating the spontaneous refolding

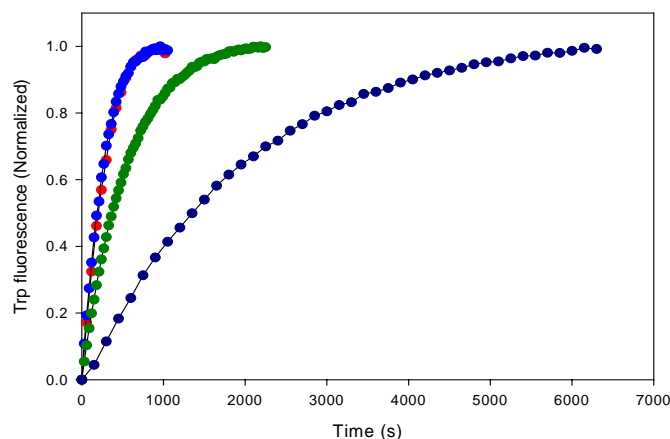


Figure 4.27: Effect of D_2O on the Refolding Rate of Spontaneous Refolding and SR-EL Mediated Refolding. Spontaneous refolding of DM-MBP in the absence (dark-blue) and presence (green) of D_2O and SR-EL mediated refolding of DM-MBP in the absence (red) and presence (royal-blue) of D_2O .

process. In contrast to spontaneous refolding, the SR-EL mediated refolding rate of the protein is not affected by this mutation. This suggests these residues play a rate-retarding role during the spontaneous folding process whereas it does not play a discernible role during the folding of the protein in the GroEL cavity. This can be explained only if the pathways are different with respect to the local structure formed around these residues in the folding intermediate or in the rate limiting transition state.

It has been shown in previous studies that D_2O can increase the hydrophobic effect and hence perturb folding rates. In the case of DM-MBP, D_2O , accelerated the spontaneous refolding rate by approx. 3 fold when compared to H_2O (Fig.4.27). This may be due to an increase in the rate of hydrophobic collapse in D_2O . In contrast to spontaneous refolding, the GroEL mediated refolding rate is not changed in the presence of D_2O . The fact that the spontaneous folding pathway can be affected by D_2O whereas the GroEL mediated folding pathway is insensitive to the same indicates that the pathways are different.

This result together with the fact that the conformation of the intermediate, as monitored by DM-MBP(175 – 298), is different in GroEL mediated folding indicates that GroEL may affect these changes by mediating a controlled collapse of the locally expanded substrate protein and confining folding intermediates in a hydrophilic cage. It seems likely that the changes in the structural features of the intermediate prevails in the refolded protein as long as the protein is encapsulated in the GroEL cage indicating that the cavity surface has the potential to remodel the structure of a refolding protein through confinement and/or surface interactions.

Chapter 5

Discussion

5.1 GroE and Protein Folding

Protein folding of newly synthesized proteins in the cell is generally dependent on the assistance of molecular chaperones. The chaperone systems of *E.coli*, in particular the GroE system, have been the subject of intensive study. Previous works have helped to identify the natural substrate spectrum of the GroEL-GroES system and provide quantitative information on substrate interaction as well as revealed direct insight into the degree of chaperone dependence of the identified substrate proteins (Houry et al., 1999; Kerner et al., 2005).

Mechanistically, the folding cycle has also been attempted to be elucidated and it has been shown that folding follows upon enclosure of non-native protein in the GroEL-GroES complex, which is essential for folding to proceed unimpaired by aggregation. Additionally, folding inside the cage can be significantly faster than folding in free solution, independently of ATP-driven cycles of GroES binding and release (Brinker et al., 2001).

The GroEL/GroES nano-cage allows a single protein molecule to fold in isolation. This reaction has been compared to spontaneous folding at infinite dilution. However, recent experimental and theoretical studies indicated that it is not a passive cage of infinite dilution but actively modulates the folding by exerting confinement on the folding protein inside the limited volume of the cage; through its mildly hydrophobic, interactive C-terminal tails at the bottom of the cage; and by the clusters of negatively charged amino acid residues exposed on the cavity wall (Tang et al., 2006).

5.2 The Chaperonin DM-MBP Bound State at Single-Molecule Resolution.

The expanded structure observed in the GroEL bound state for the protein double labelled at position, 52C and 298C clearly shows that, even under solution conditions that strongly favor folding, the protein can be significantly opened up when bound to the apical domains of GroEL. Though it has been previously reported that proteins might show an expansion when bound to GroEL (Lin and Rye, 2004), this study clearly showed that the opening up on the GroEL surface may have been underestimated when probed by ensemble measurements.

The Chaperonin-Substrate complex has been speculated to be a heterogenous distribution as the NMR based study was not able to resolve a number of peaks (Horst et al., 2005) and hence was not able to give exhaustive structural insight into the bound state conformation. Recent studies using Cryo-Electron Microscopy and Malate Dehydrogenase as the model substrates have attempted to investigate the nature of the bound substrate when bound to GroEL (Elad et al., 2007), but lacks the information of subsequent changes, if any, in the protein conformation upon nucleotide and GroES binding. Moreover, as the sample itself is frozen, it limits the information regarding the dynamics of the bound state.

For the first time, in this study, we have been able to explore the GroEL bound substrate conformation at single molecule resolution, which gives us an insight into the heterogeneity which has been otherwise compromised in previous studies. The substrate protein bound to GroEL shows a bimodal distribution of conformations. One of the populations is as expanded if not more than the unfolded protein in denaturant and the other population is compact similar to native like distributions. At the same time, this is not a global phenomena and different regions of the protein experience varying degrees of unfolding upon binding to GroEL.

5.3 Bimodal Distribution and Allosteric Regulation

It is also seen that the expanded conformation seen on the GroEL surface is not discernible during the initial phase of spontaneous refolding of the protein, indicating that the surface of GroEL with its multivalent binding capability has the unique ability to present an unfolding environment even under conditions that strongly favor folding.

The most expanded state of the unfolded protein results from binding to the nucleotide-

free, high affinity state of GroEL. Thus, the bimodal distribution in the conformation of bound protein observed in the absence of nucleotide suggests that normally an equilibrium exists between high and low-affinity GroEL states, the latter of which is stabilized, rather than induced, by nucleotide binding. This is consistent with the theory of nested cooperativity (Yifrach and Horovitz, 1995). According to this theory, the subunits within a ring are in equilibrium between a T state, which has a high affinity for the unfolded protein substrate but low affinity for ATP, and an R state, which has the converse properties. Allosteric regulation is mediated by the protein and the nucleotide in opposite directions of equilibrium between the two states of GroEL (Poso et al., 2004).

5.4 GroEL Binding Induced Local Unfolding

The energy landscape of the kinetically populated folding intermediate of DM-MBP is altered by GroEL. Upon dilution of the unfolded protein into refolding buffer, distinct native like distributions of FRET efficiencies are obtained for different FRET pairs in the folding intermediate. The observation that there are unimodal distributions with different FRET pairs indicates that there is a local energy minimum on the folding pathway of DM-MBP. This is changed upon the binding of the intermediate to GroEL. The presence of bimodal distribution indicates that there are multiple minima when the protein is bound to GroEL.

Increase in the width of the distributions upon binding of GroEL indicates that the minima are shallow. The increase in the heterogeneity of the distribution upon GroEL binding also indicates that the bound state of DM-MBP has higher conformational entropy than the collapsed state of the protein formed at the start of spontaneous refolding. The restructuring of the energy landscape by GroEL can, in principle, cause the protein to be removed from a kinetically trapped energy minimum, resulting in an accelerated refolding process.

Expansion of protein by nucleotide-free GroEL may be of significant importance for proteins that might need only the unfolding action of the GroEL apical domain to facilitate their folding.

5.5 Bound State and the Rate of Folding

In the case of DM-MBP, we see that in SR-EL the expanded conformation is either absent or present in a lower fraction as compared to GroEL, even though SR-EL, is observed to accelerate the refolding of DM-MBP to the same extent as GroEL. This indicates that

the percentage of expanded conformation does not correlate with the rate acceleration of refolding of DM-MBP. Although the presence of the more opened-up state on GroEL may provide an opportunity for some of the substrate proteins to be stretched out of their kinetically trapped conformations, this observation negates the possibility of the presence of this stretched out bound state being a necessary step for the rate enhancement of folding in case of DM-MBP.

A crucial difference between GroEL and SR-EL in their functional capabilities *in vitro* might be in their differential ability to unfold proteins that need unfolding as a necessary event in order to initiate their refolding process. It will be of importance to see whether proteins that are unable to refold in the presence of SR-EL *in vitro* and does so in the presence of GroEL gain from the ability of GroEL to unfold the protein.

5.6 Nucleotide Induced Transient Unfolding and Subsequent Compaction Upon GroES Binding

It has earlier been hypothesized that the binding of ATP to GroEL causes forced unfolding of the protein (Shtilerman et al., 1999). In this work, it is shown that DM-MBP bound to GroEL/EL-SR undergoes transient expansion upon binding of ATP. This happens on the time-scale of ATP binding to GroEL/EL-SR indicating that ATP binding and not hydrolysis is essential for the unfolding of the protein.

This unfolding is followed by subsequent compaction, as monitored by SpFRET to yield more compact conformational states. Subsequent GroES binding to the complex results in compaction of the protein without any further expansion indicating that GroES binding does not cause any forced unfolding event. Under conditions where the nucleotide induced expansion of DM-MBP is slow (in the presence of ADP:AlFx) or absent (in the presence of 750mM $(NH_4)_2SO_4$; which offers a highly hydrophobic environment) the rate of refolding is not altered.

Under such conditions, SR-EL is nonetheless able to accelerate the refolding of DM-MBP to the same extent as that of ATP induced transient unfolding suggesting that ATP induced expansion of DM-MBP is not an essential event for the rate acceleration of refolding of DM-MBP. It may be essential for some other substrate proteins. It was also observed by SpFRET experiments that within the first 200 seconds of the spontaneous refolding or EL-SR/GroES/ATP mediated refolding reaction, DM-MBP reaches near native compaction with a unimodal distribution. The absence of expanded states under

these conditions indicates the absence of expanded intermediates in the spontaneous or chaperonin assisted folding pathway of DM-MBP.

5.7 Step-Wise Release of the Substrate Protein from GroEL

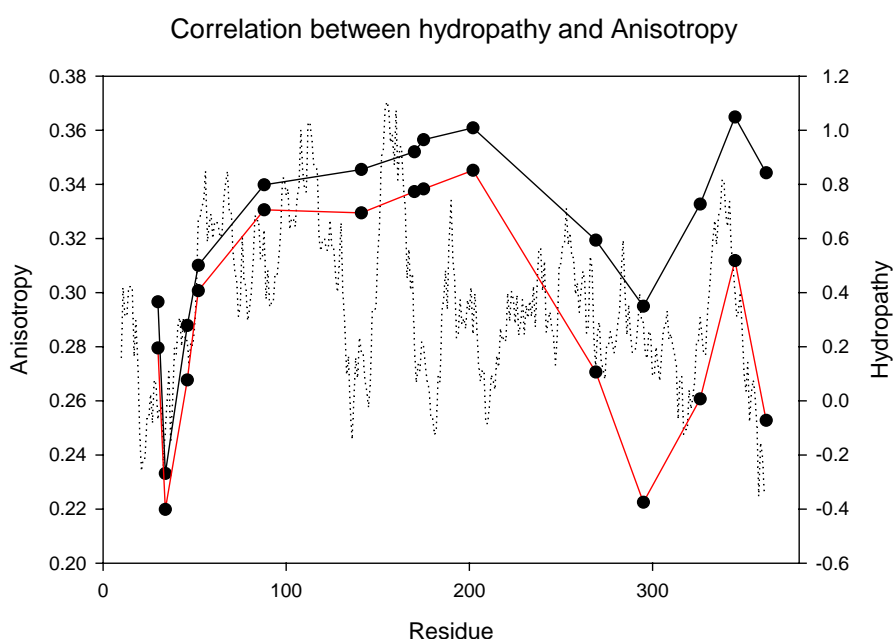


Figure 5.1: Segmental Mobility of Different Regions of the Labelled DM-MBP When Bound to GroEL Correlated With Hydrophobicity. The steady-state anisotropy of the single cysteine mutants of DM-MBP labelled with (Atto-532) were measured when the protein was bound to GroEL (filled black Squares with black line) and with addition of ATP (filled red circles with red line) to the GroEL bound complex. The hydropathy index of DM-MBP, calculated from the sequence using Abraham and Leo scale of hydrophobicity with a sequence window of 19 residues is also shown (dashed line).

The release of the unfolded protein from the GroEL apical domain upon addition of ATP and GroES appears to occur in a stepwise manner. With regions of lower hydrophobicity being released in the first phase, when GroEL binds ATP, followed by the release of other regions upon binding of GroES to the GroEL/ATP complex. The binding of GroEL apical domain to substrate protein is believed to be mainly driven by hydrophobic interactions and it is also well known that nucleotide binding to GroEL induces conformation change in the apical domain in the chaperonin resulting in a decreased

hydrophobic surface area (Kim et al., 2005). Therefore, the release as measure by the anisotropy correlates well with the hydrophathy plot of DM-MBP calculated from the sequence using Abraham and Leo scale of hydrophobicity with a window of 19 residues (Fig.5.1).

The observation that the release of the substrate protein happens in a stepwise manner is easily explainable with the reduction of hydrophobic area on GroEL apical domain upon ATP binding and release of all but the highly hydrophobic regions of the protein which are released only after further conformational changes/or steric competition upon GroES binding. The biphasic release has the potential to alter the folding pathway of DM-MBP inside the GroEL cavity. Since it is well known that the initial collapse of most of the proteins are driven by hydrophobicity (Dill and Chan, 1997), sequestration of the most hydrophobic regions of the protein by GroEL and releasing the weaker hydrophobic segments before the segments with higher hydrophobicity may cause the protein to collapse differently from the spontaneous refolding event, leading to a different intermediate formation in the GroEL cavity.

5.8 Rate Acceleration by GroEL via a Different Folding Pathway

GroEL/GroES/ATP mediated protein folding may be faster inside the cavity if GroEL specifically interacts with native like elements in the transition state of refolding and, hence, stabilizes the GroEL-ES-DM-MBP complex* (where DM-MBP* is the transition state of refolding of DM-MBP). In this mechanism, the folding pathway is essentially the same without any change in structural features of the intermediate or transition state induced by GroEL. This would result in a similar behavior of the spontaneous and GroEL/ES/ATP assisted folding pathways to external perturbations like mutation or solvent conditions which is ruled out by the present study. Kinetic data of folding obtained from ensemble FRET measurements suggest that the conformation of the refolding intermediate is altered by the GroEL cavity, which leads to a folding pathway that is different from the spontaneous refolding pathway.

The difference in conformation of the GroEL mediated folding intermediate from the folding intermediate formed during spontaneous folding is also retained as a difference in conformation between the refolded protein trapped in the SR-EL cavity and the spontaneously or GroEL/GroES/ATP refolded protein. This indicates that GroEL is able to change the structural properties of the intermediate as well as the refolded state which

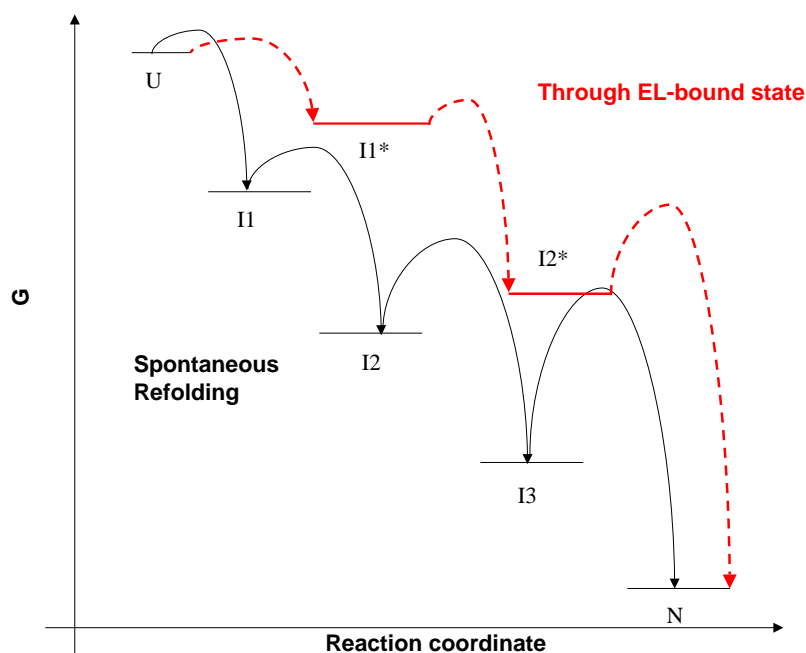


Figure 5.2: Schematic Representation of How GroEL-GroES May be Helping to Increase the Folding Rate. GroEL/ES restructures the intermediates and re-routes them through a different pathway. The free energy barriers for kinetically populated intermediates are thus decreased resulting in global smoothing of the free energy landscape.

might lead to a different folding pathway which contains fewer kinetic traps or with a less rugged energy landscape (Fig.5.2).

This is also supported by the fact that refolding in GroEL is less sensitive to one of the mutations, D30C, P298C, or D_2O containing buffer that accelerate the spontaneous folding rate indicating a difference in folding pathway. This is the first report of an alternate folding pathway in the GroEL cavity which should facilitate future experiments to investigate this change as a function of confinement and surface properties of the GroEL cavity.

5.9 The GroEL-Substrate Complex Returns to the Initial Bound State During GroEL Cycling.

It has been previously hypothesized that iterative annealing of the unfolded protein is essential for accelerated protein folding in the GroEL cavity (Todd et al., 1996). Here we show that iterative annealing or cycling does not affect the starting conformational prop-

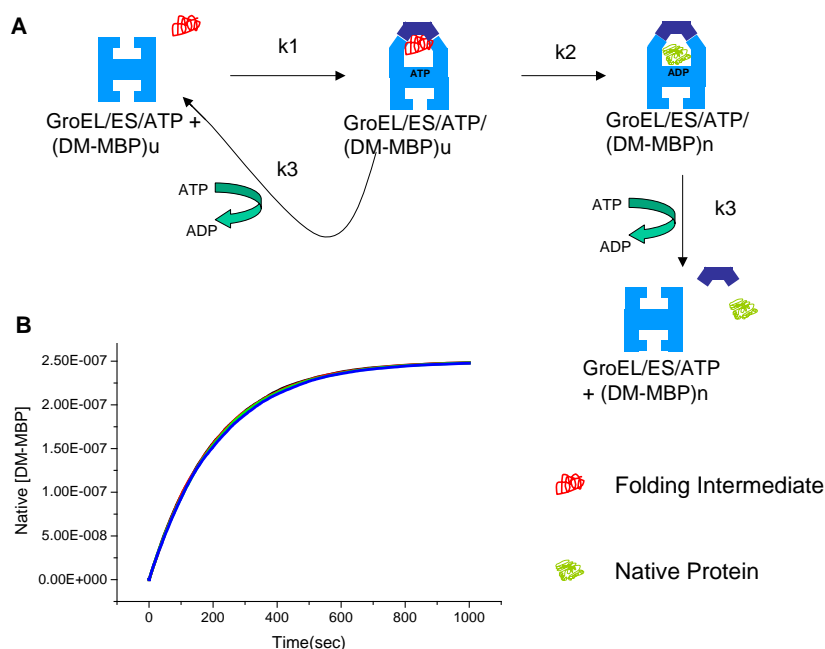


Figure 5.3: ATPase Rate and Cycling does not Affect the Rate of Refolding of DM-MBP. Representation of the different rates involved in the functional GroEL cycle. The rate of cycling does not affect the rate of refolding as seen by mathematical simulation (B) when the ATPase rate is changed from 5 sec to 40 sec in the super-imposable red, blue and green curve.

erties of the GroEL bound substrate protein. Upon iterative re-binding of the substrate protein to GroEL apical domain, the unfolded protein returns to similar conformational distribution irrespective of the number of cycling events.

Since the protein returns to the original expanded state at the end of each cycle, it was necessary to verify by numerical simulations whether cycling could affect the actual refolding rate of a GroEL assisted refolding reaction. To obtain estimates of refolding rates under conditions involving GroEL/ES cycling, numerical simulations¹ were performed with the kinetic scheme as shown in (Fig5.3A). The rate for the binding and encapsulation of DM-MBP under the conditions used for the reactions (1 μ M of GroEL, 2 μ M of GroES, 2mM ATP, and 250nM of denatured DM-MBP) is limited by the binding of GroEL to the unfolded protein with a rate of approx. $4 \times 10^6 M^{-1} s^{-1}$ (k1), which is followed by fast binding of GroES to result in substrate encapsulation. The intrinsic rate of refolding of DM-MBP in GroEL was fixed at $5 \times 10^{-3} s^{-1}$ (k2) and the ATPase induced cycling rate was varied from $0.2 s^{-1}$ to $0.025 s^{-1}$ (k3) to obtain the effect

¹This work was performed by Kausik Chakraborty.

of rate of cycling on the final refolding rate (Fig 5.3B). It was observed that the refolding rate of the DM-MBP under these conditions was still $5 \times 10^{-3} \text{s}^{-1}$ and is not affected by the fact, that at the end of a cycle, the not yet native protein rebinds to GroEL to form an expanded conformation. It was also observed that the refolding rate is not affected by the cycling rate of the GroEL-GroES system, indicating that the ATPase rate might not play a significant role in determining the rate of a refolding reaction, which is corroborative with the data that SR-EL with one cycle of ATP hydrolysis is able to accelerate the refolding to a similar extent as GroEL.

5.10 GroEL Dependence of a Substrate Protein

There have been many studies aiming to find any structural correlation in the native state to the regions interacting with GroEL within a protein e.g. (Stan et al., 2006) where it has been shown that GroEL interacts with the hydrophobic stretches of a protein which are buried and not exposed in the native state. This is also consistent with the fact that GroEL is shown to not interact with a protein in its native state. There are also studies that have aimed at identifying GroEL substrates by sequence alignments to the mobile loop region of GroES (Chaudhuri and Gupta, 2005) based on the rationale that natural substrate proteins should contain binding motifs similar in sequence to the mobile loop peptide of GroES that displaces the binding motif during the chaperonin cycle. These parameters, if substantiated, can be used to extrapolate it to an unknown set to predict the GroEL dependence of a particular protein. Certain folds or structural motifs have also been shown to be enriched as compared to the natural protein occurrence in the proteome, for instance TIM barrels (Kerner et al., 2005).

We know that Maltose Binding Protein is not a natural substrate of GroEL as it is a preplasmic protein and it translocated by the Sec pathway (Collier et al., 1988). Nevertheless the cytosolic variant of DM-MBP is a GroEL dependent Class III substrate (Tang et al., 2006). The obvious difference between the two variants of MBP are the folding rate and GroEL dependent rate acceleration, therefore it may be speculated that GroEL dependence of a protein may also be dependent on its refolding pathway. As it has been shown that the GroEL binding kinetics is in the order of seconds and the initial collapse or compaction of a fast folding protein is in the the order of ms or even less, the protein in such a case would not get a chance to interact with GroEL. In the case of Maltose Binding Protein, two point mutations decelerate the rate of folding and the resultant DM-MBP is hence dependent on GroEL.

Appendix A

Appendix

A.1 Sp-FRET Measurement of DM-MBP Cys Constructs

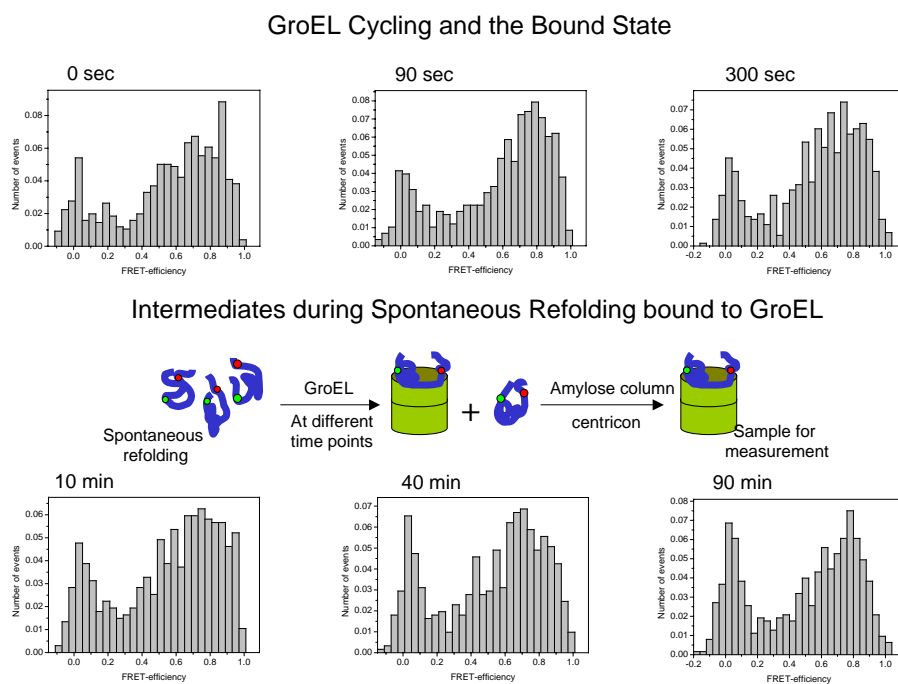


Figure A.1: Sp-FRET measurements of the complex of GroEL complexed with double-labelled MBP-DM(52–298). After initiating GroEL/GroES/ATP mediated (above) and spontaneous refolding (below) of DM-MBP(52–298), the reaction was terminated using EDTA and apyrase at specific below mentioned time points. Measurements were taken after removing unbound and native protein.

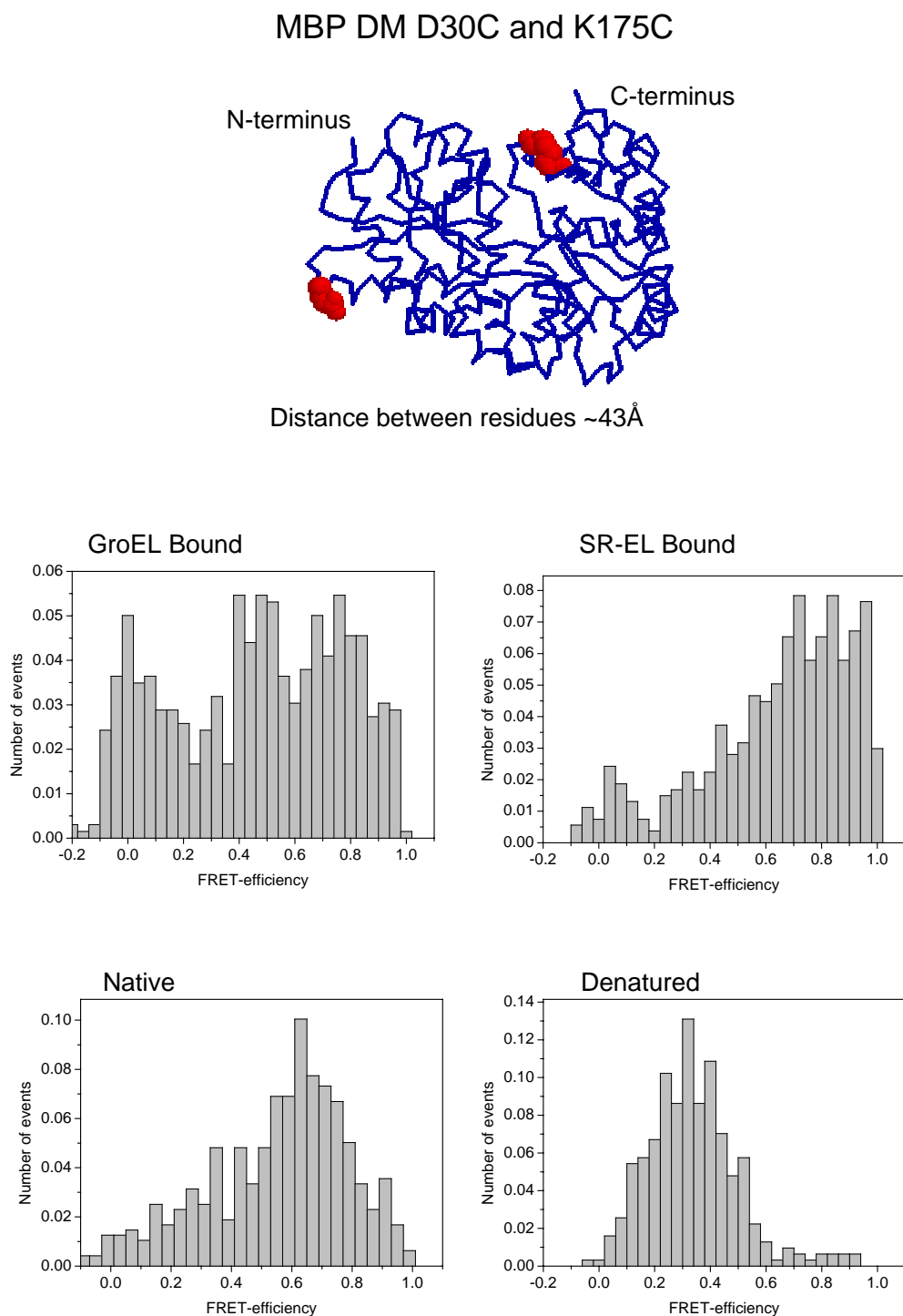


Figure A.2: Sp-FRET measurements of double-labelled MBP-DM(30–175) . FRET efficiency distribution is shown for the protein bound to GroEL, SR-EL, spontaneously refolded protein, unfolded in 3M GuHCl.

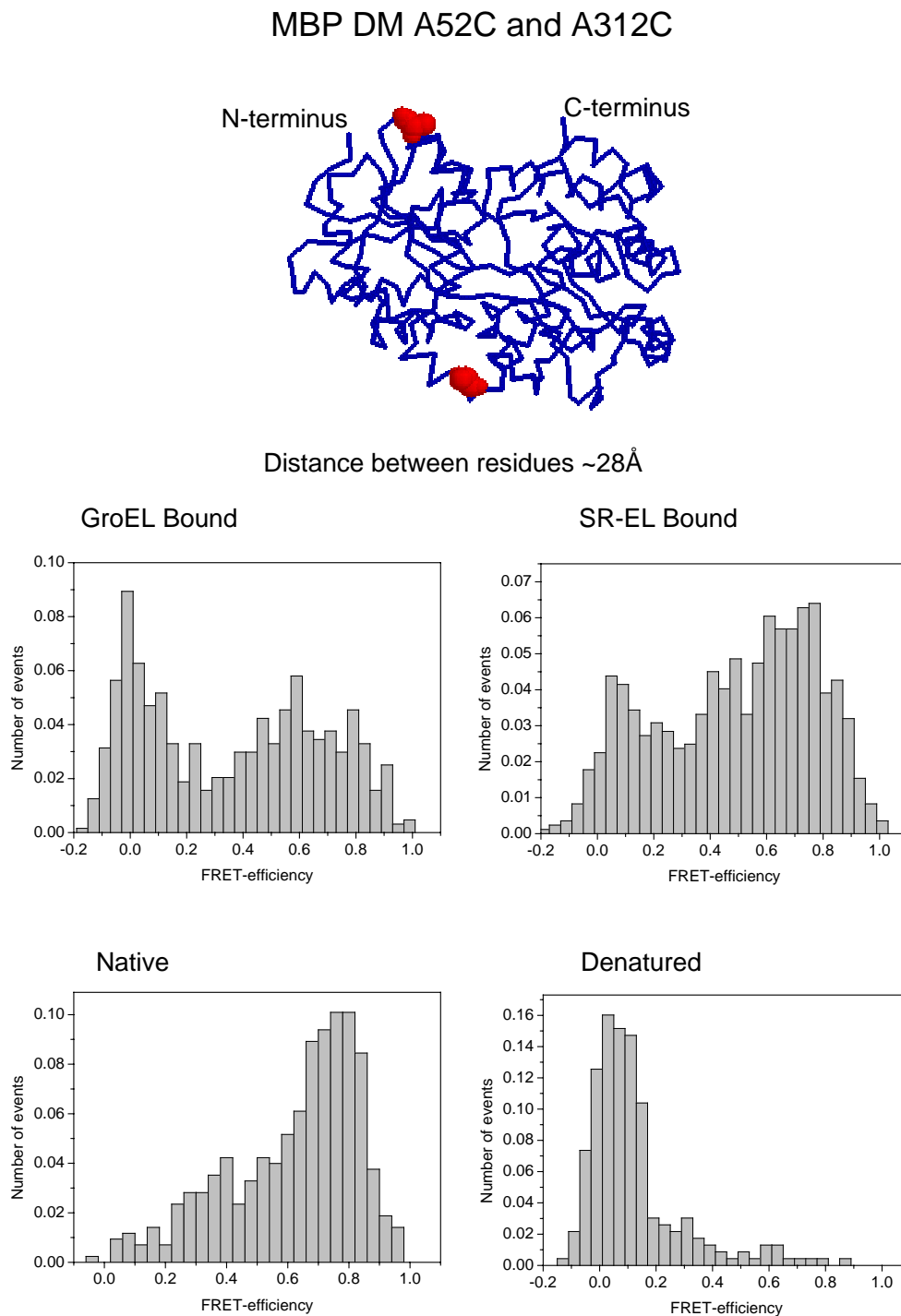


Figure A.3: Sp-FRET measurements of double-labelled MBP-DM(52–312). FRET efficiency distribution is shown for the protein bound to GroEL, SR-EL, spontaneously refolded protein, unfolded in 3M GuHCl.

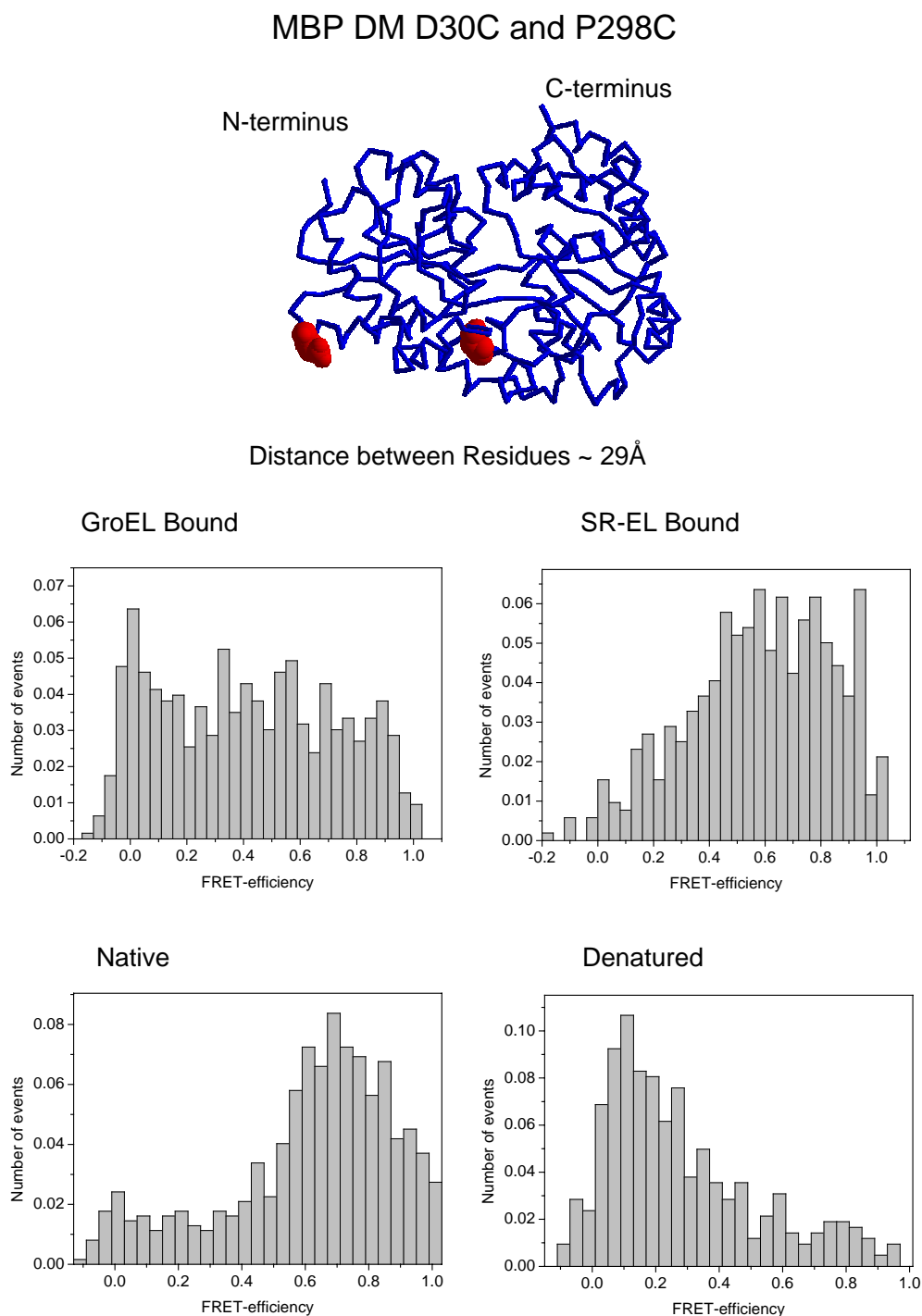


Figure A.4: Sp-FRET measurements of double-labelled MBP-DM(30–298) . FRET efficiency distribution is shown for the protein bound to GroEL, SR-EL, spontaneously refolded protein, unfolded in 3M GuHCl.

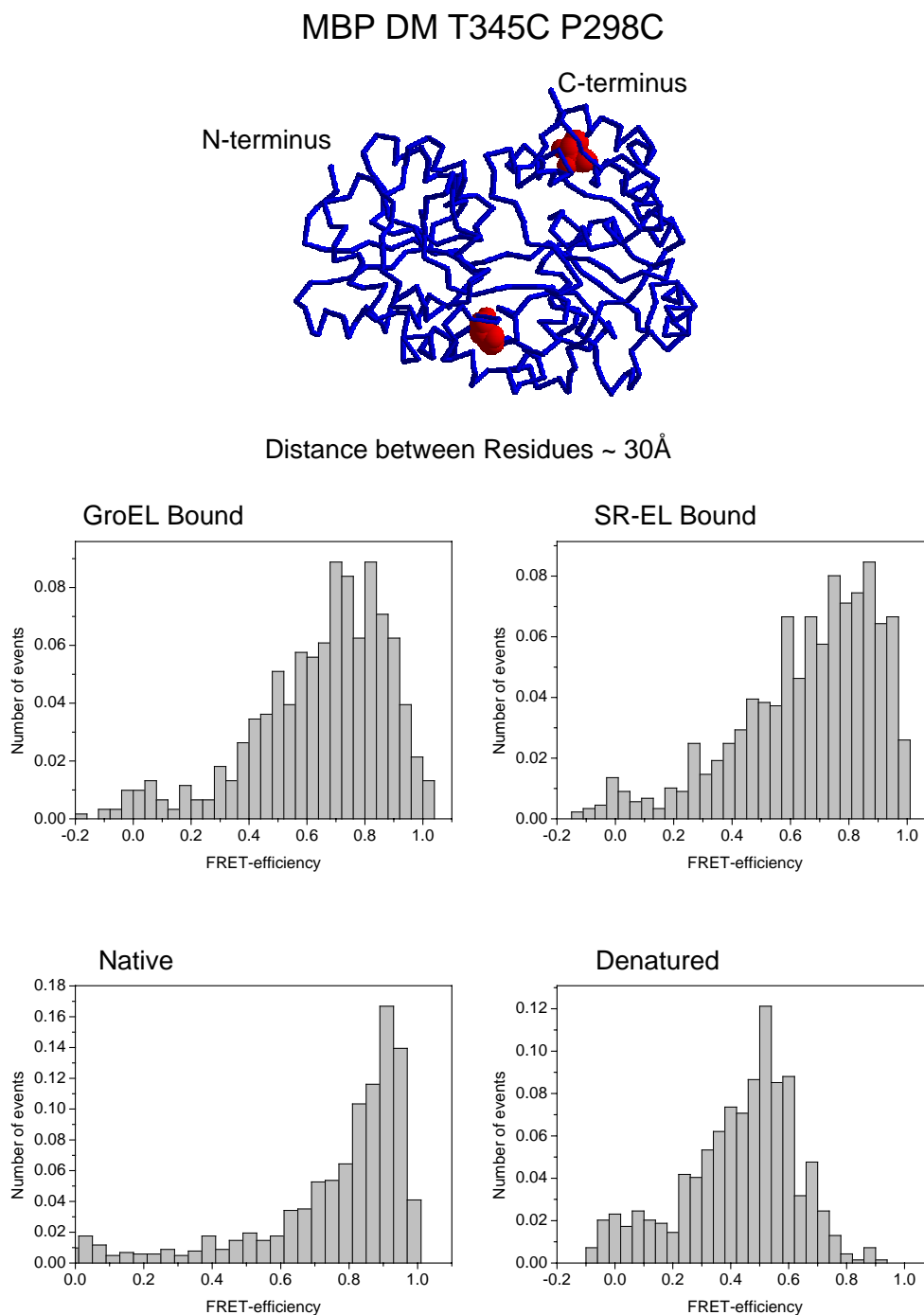


Figure A.5: Sp-FRET measurements of double-labelled MBP-DM(298 – 345) . FRET efficiency distribution is shown for the protein bound to GroEL, SR-EL, spontaneously refolded protein, unfolded in 3M GuHCl.

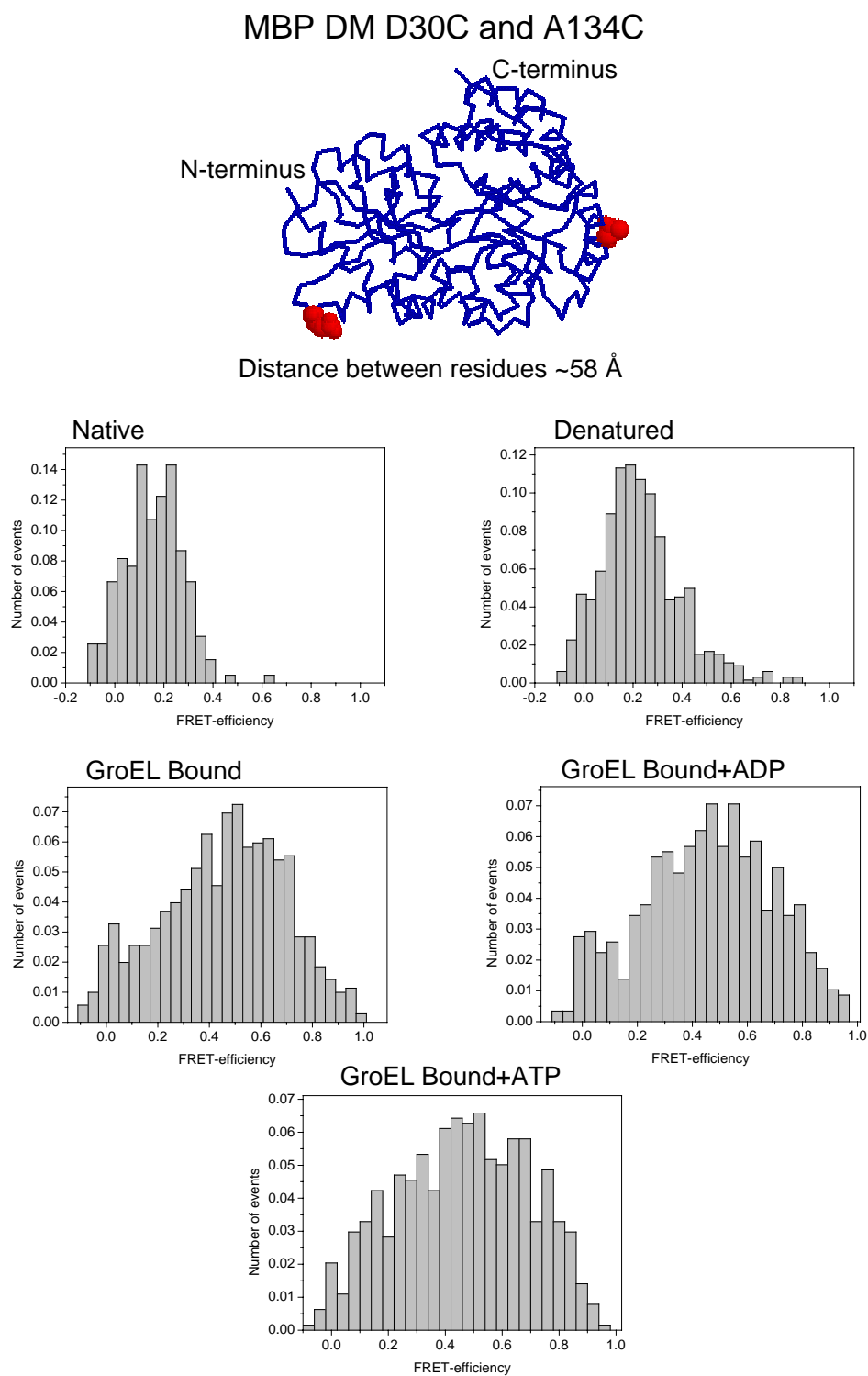


Figure A.6: Sp-FRET measurements of double-labelled MBP-DM(30–134). FRET efficiency distribution is shown for spontaneously refolded protein, unfolded in 3M GuHCl, the protein bound to GroEL, with ADP and with ATP,.

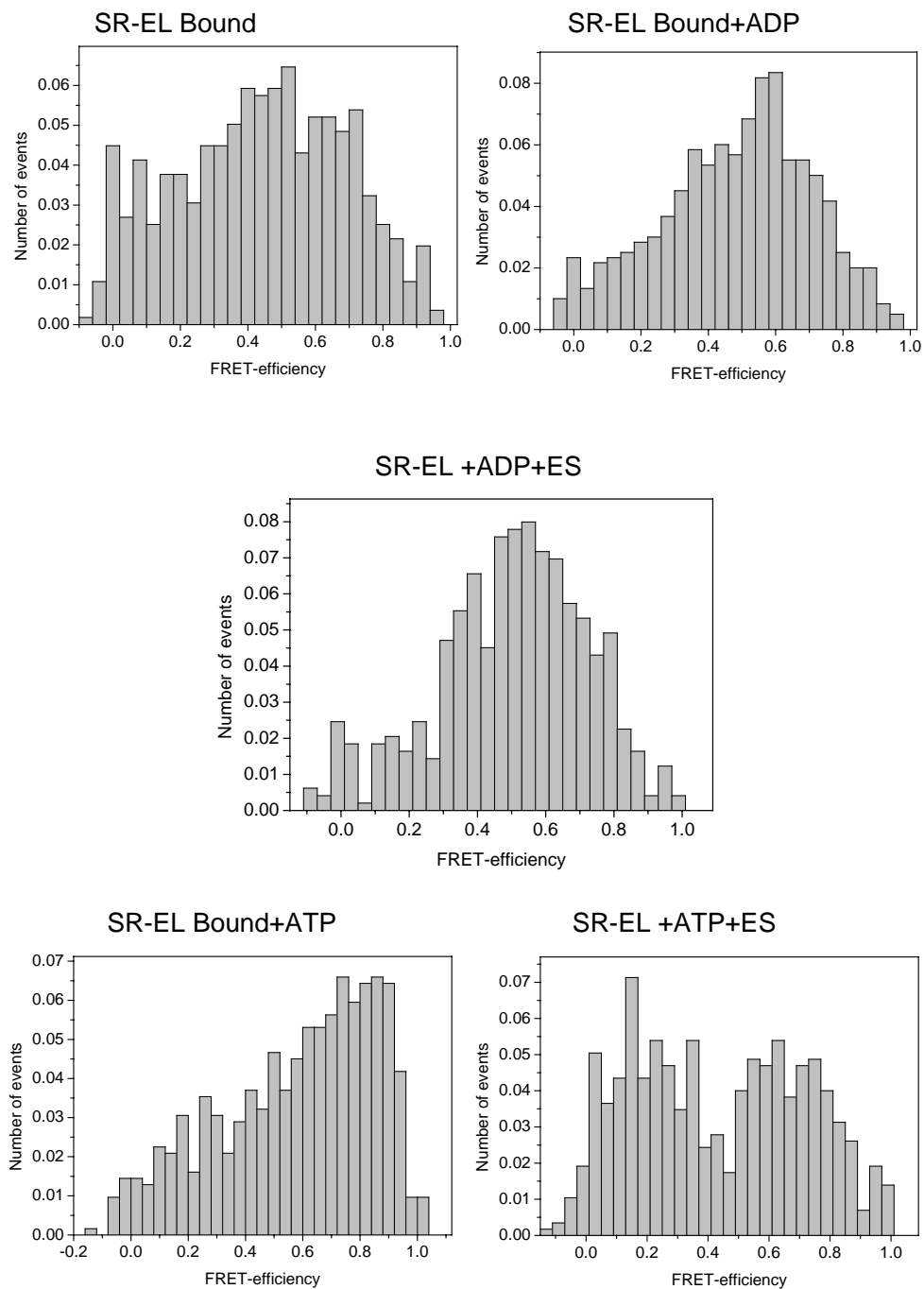


Figure A.7: Sp-FRET measurements of double-labelled MBP-DM(30–134). FRET efficiency distribution is shown for the protein bound to SR-EL, with ADP with ADP and GroES, with ATP, and with ATP and GroES.

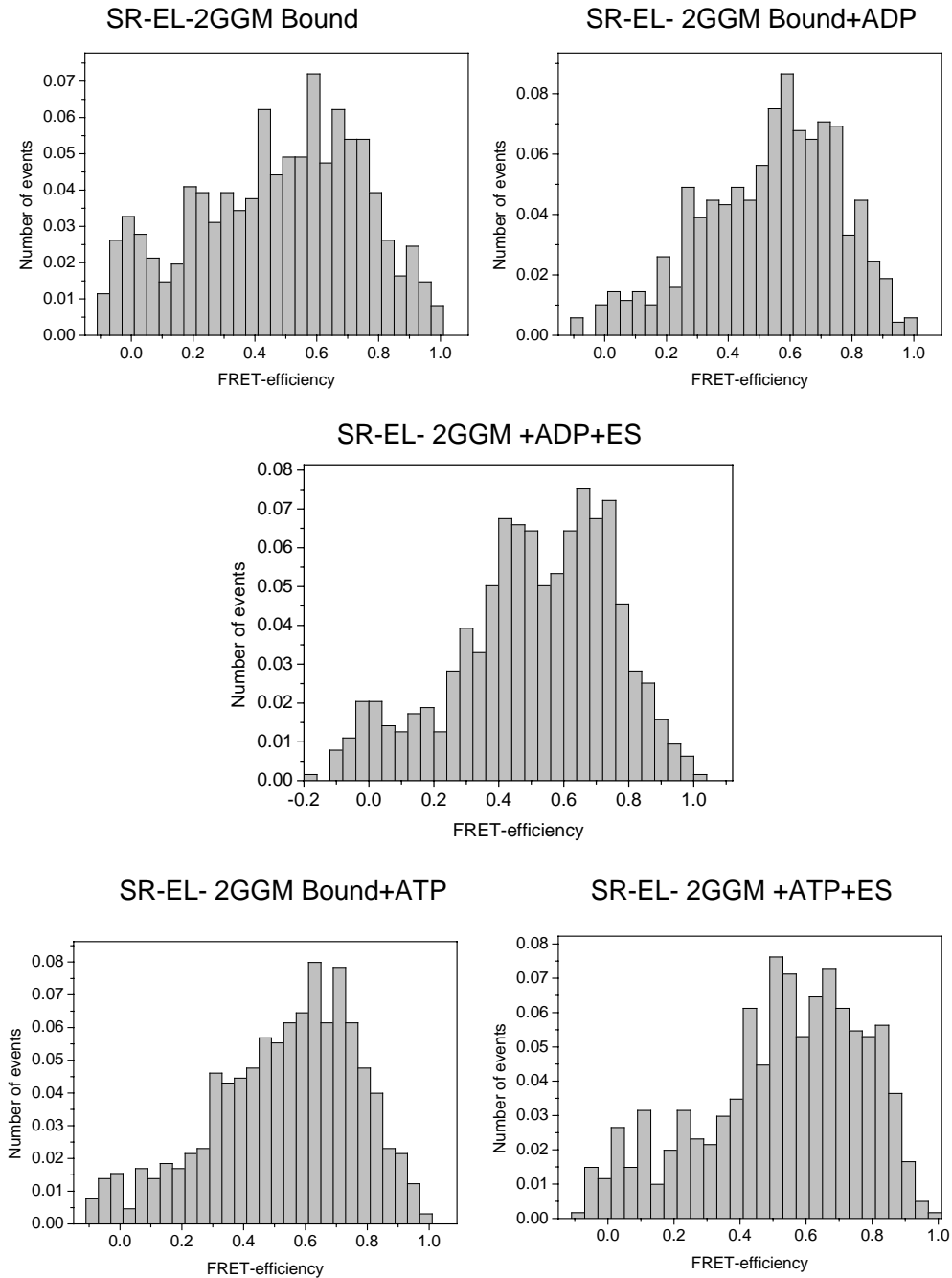


Figure A.8: Sp-FRET measurements of double-labelled MBP-DM(30–134). FRET efficiency distribution is shown for the protein bound to SR-EL-2GGM, with ADP with ADP and GroES, with ATP, and with ATP and GroES.

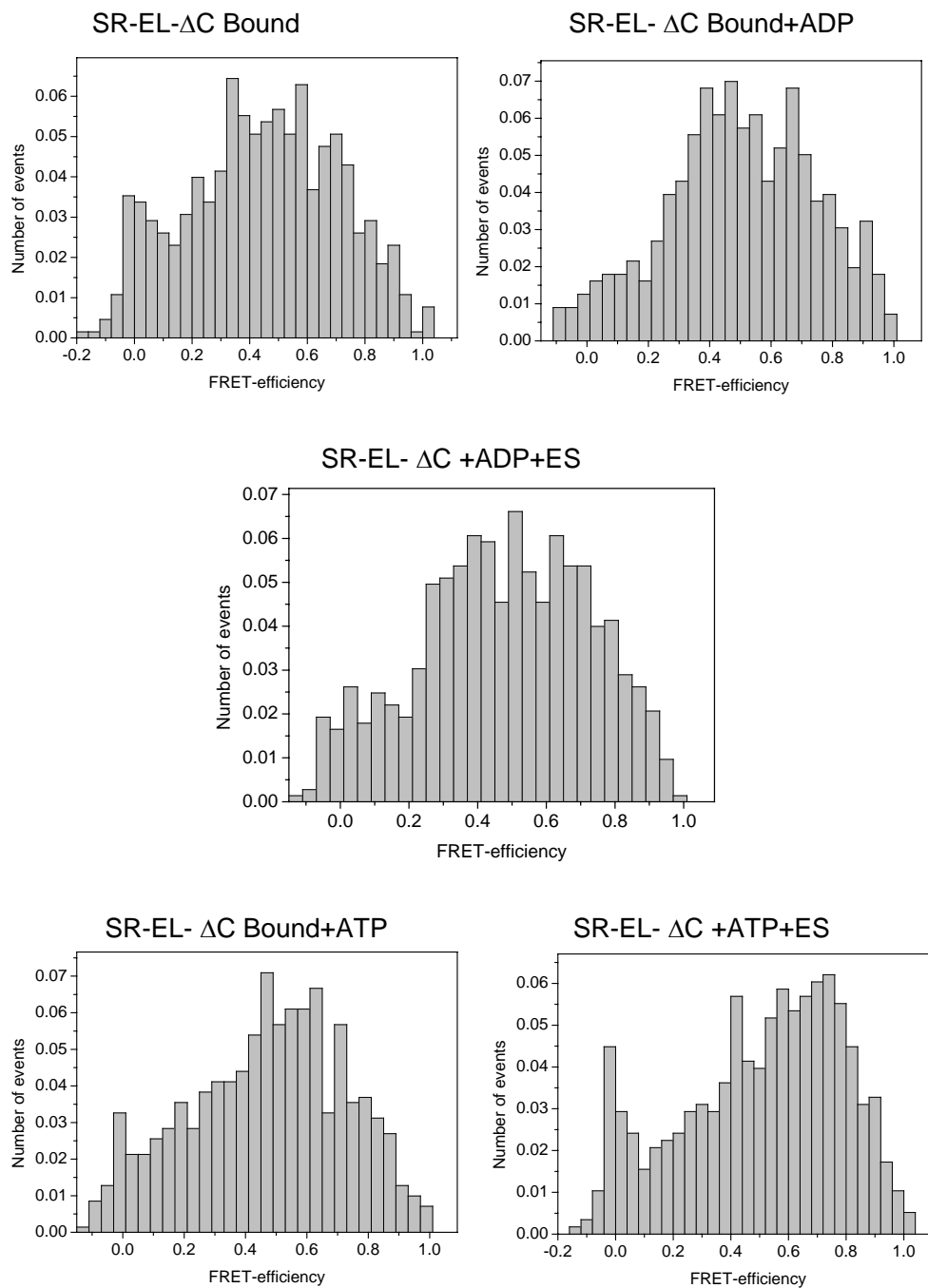


Figure A.9: Sp-FRET measurements of double-labelled MBP-DM(30–134). FRET efficiency distribution is shown for the protein bound to SR-EL- δ C, with ADP with ADP and GroES, with ATP, and with ATP and GroES.

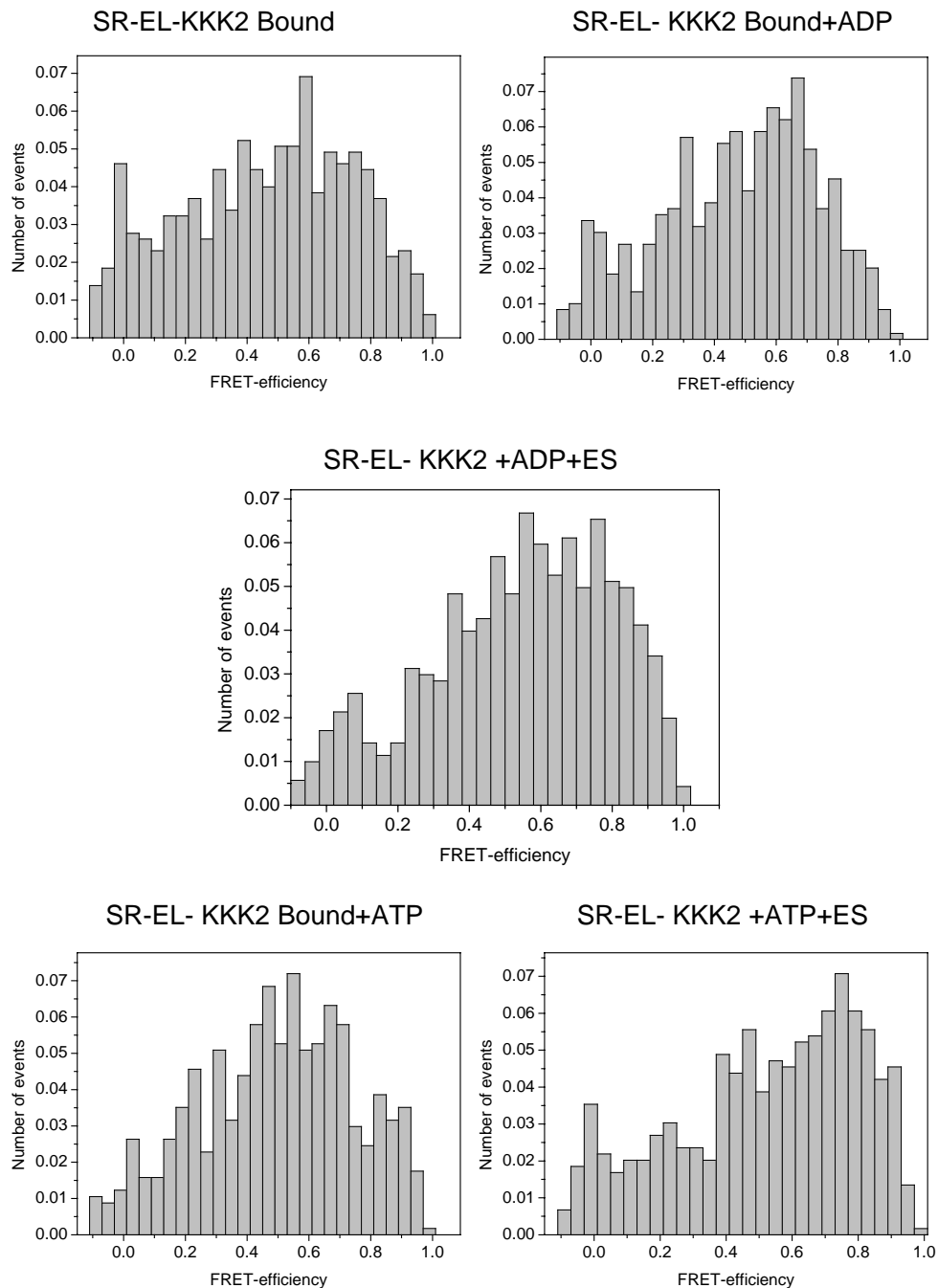


Figure A.10: Sp-FRET measurements of double-labelled MBP-DM(30 – 134) . FRET efficiency distribution is shown for the protein bound to SR-EL-KKK22, with ADP with ADP and GroES, with ATP, and with ATP and GroES.

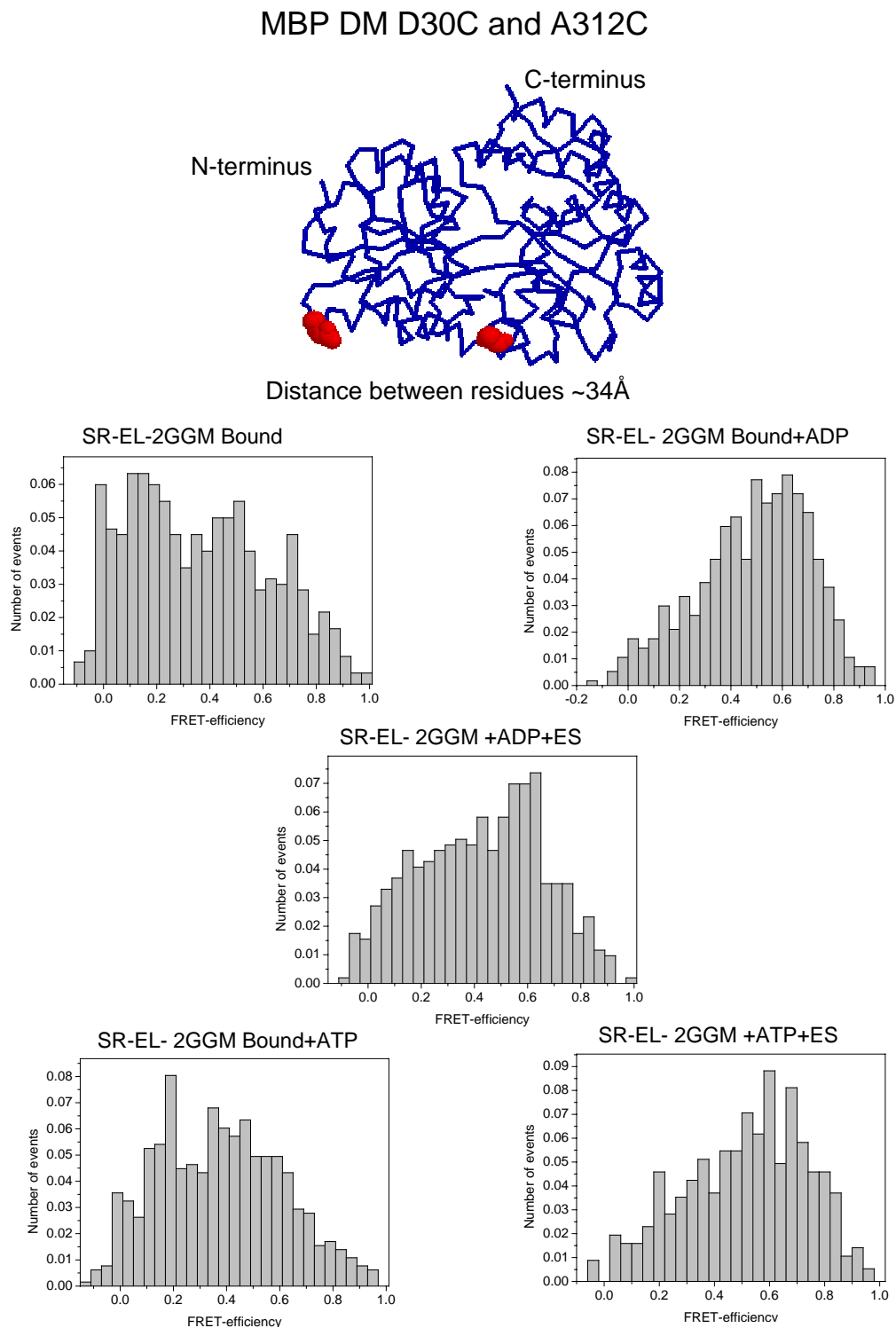


Figure A.11: Sp-FRET measurements of double-labelled MBP-DM(30 – 312) . FRET efficiency distribution is shown for the protein bound to SR-EL2GGM, with ADP with ADP and GroES, with ATP, and with ATP and GroES.

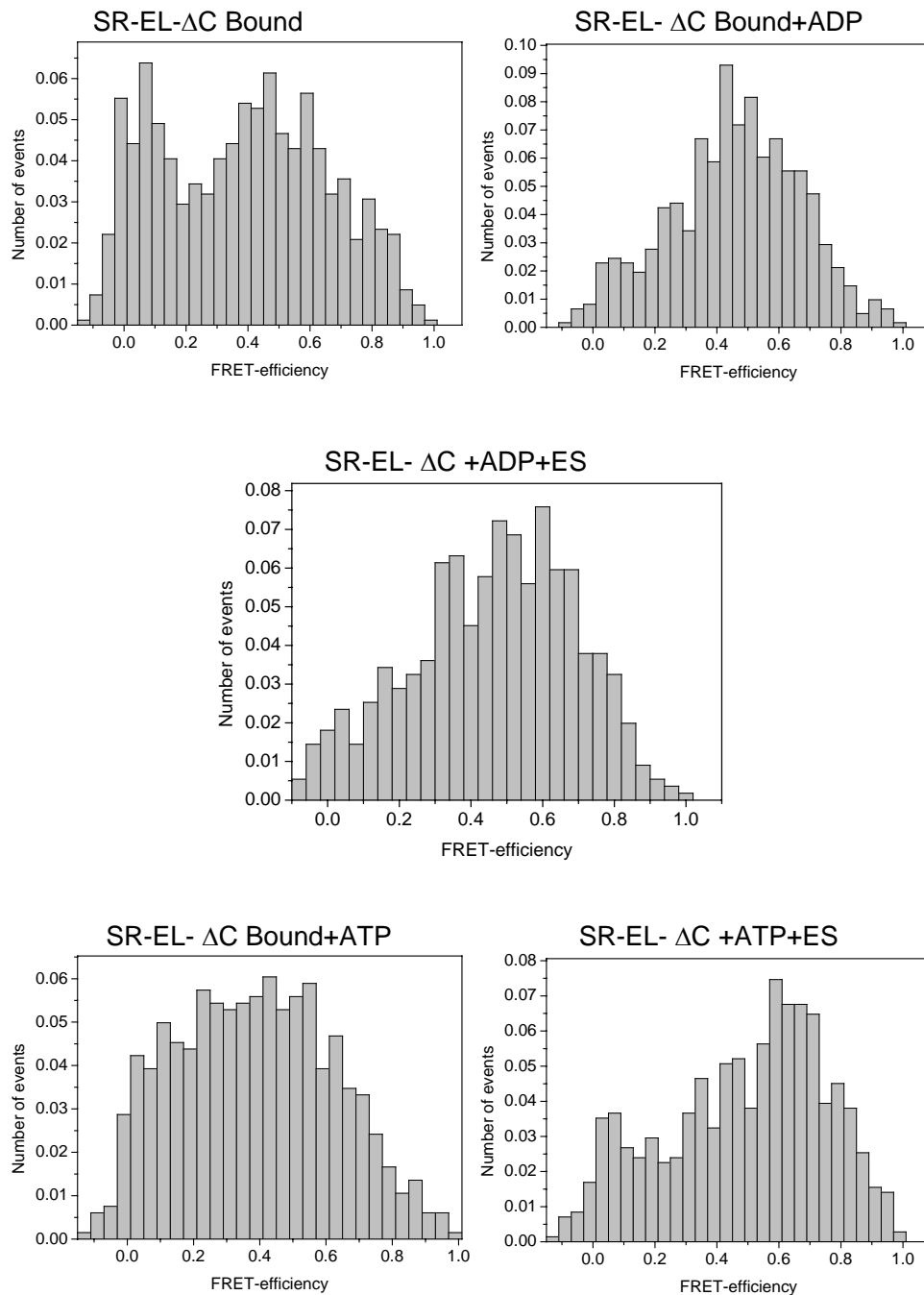


Figure A.12: Sp-FRET measurements of double-labelled MBP-DM(30 – 312) . FRET efficiency distribution is shown for the protein bound to SR-EL- δ C, with ADP with ADP and GroES, with ATP, and with ATP and GroES.

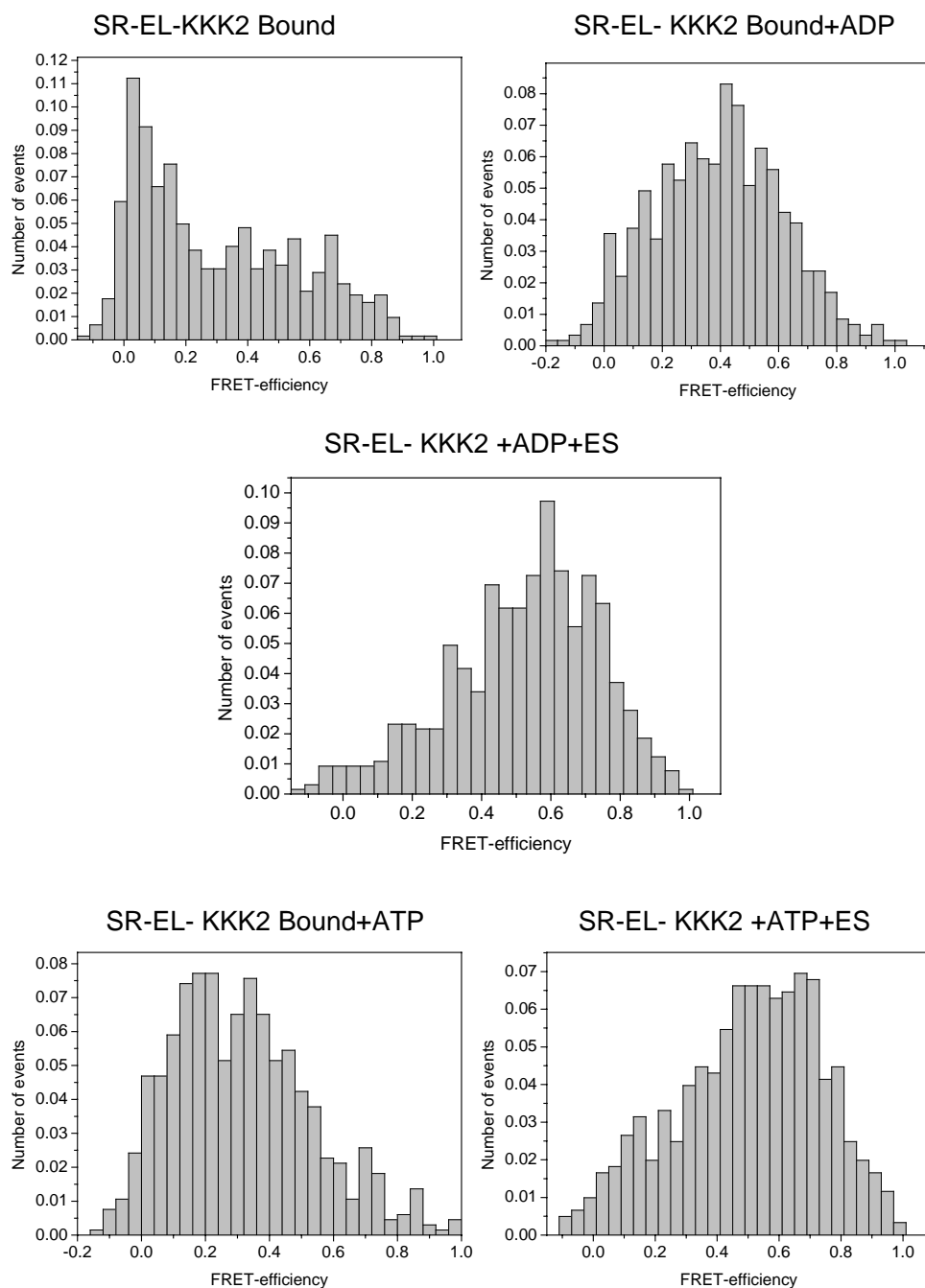


Figure A.13: Sp-FRET measurements of double-labelled MBP-DM(30 – 312) . FRET efficiency distribution is shown for the protein bound to SR-EL-KKK2, with ADP with ADP and GroES, with ATP, and with ATP and GroES.

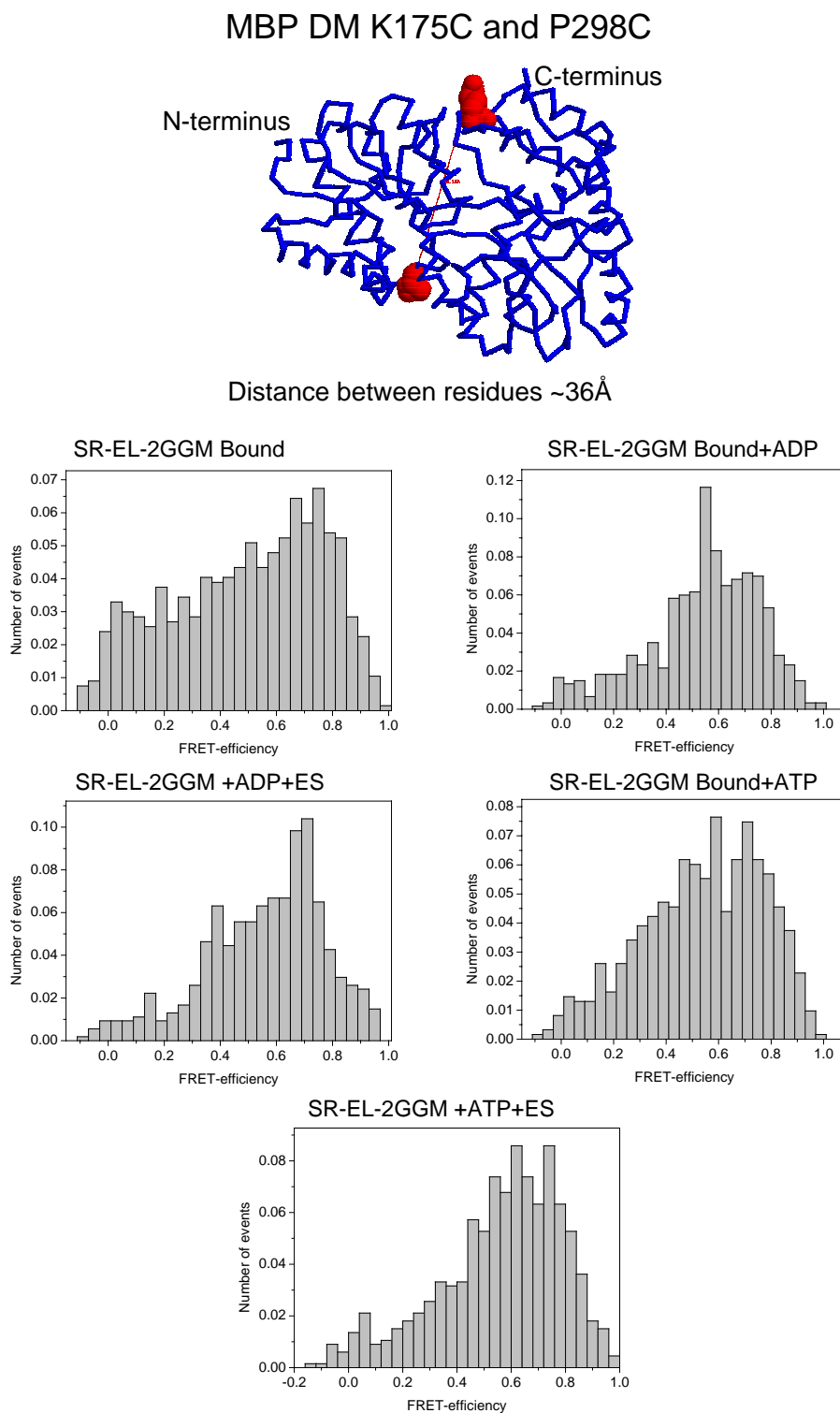


Figure A.14: Sp-FRET measurements of double-labelled MBP-DM(175 – 298) . FRET efficiency distribution is shown for the protein bound to SR-EL2GGM, with ADP with ADP and GroES, with ATP, and with ATP and GroES.

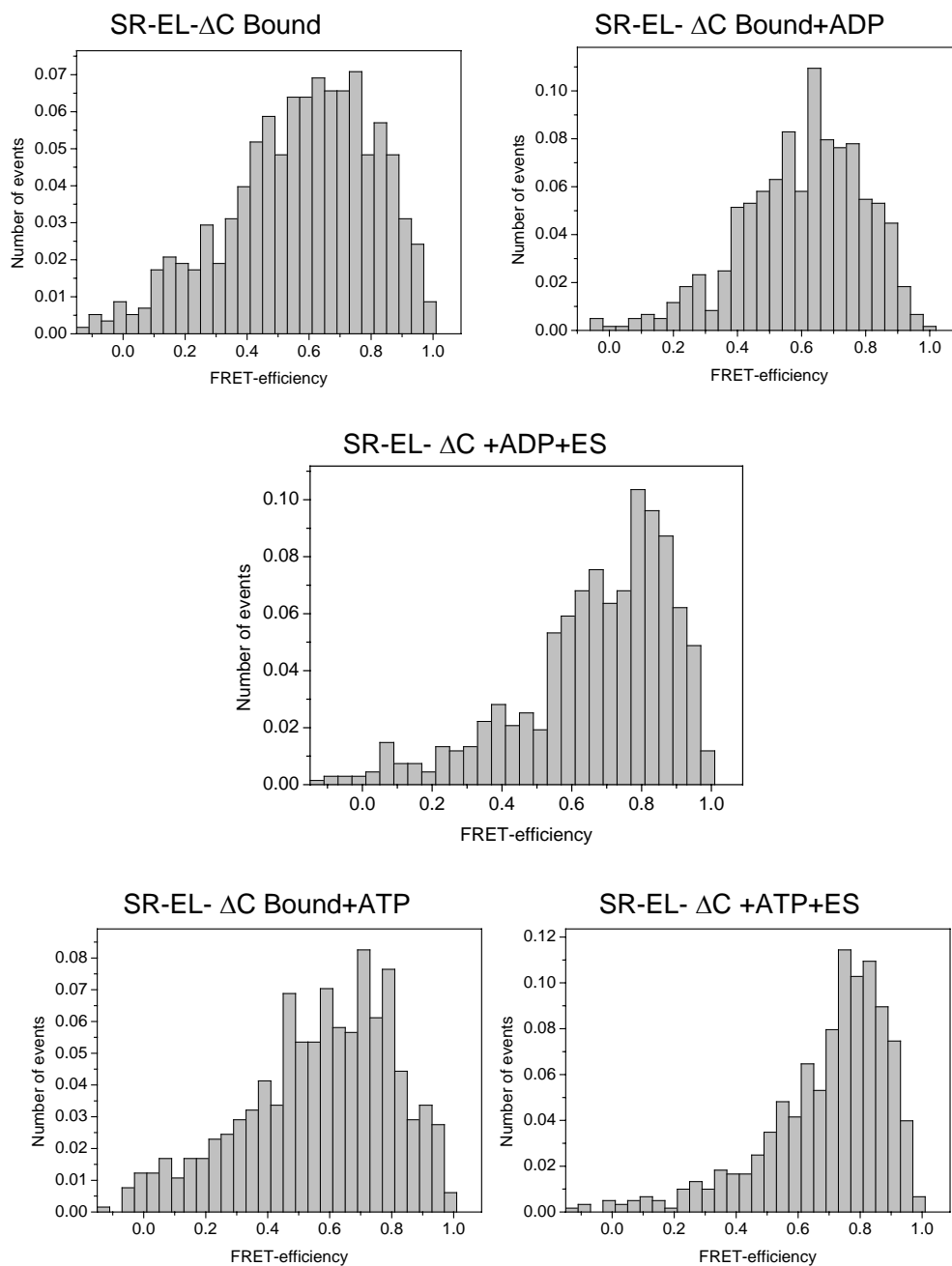


Figure A.15: Sp-FRET measurements of double-labelled MBP-DM(175 – 298) . FRET efficiency distribution is shown for the protein bound to SR-EL- δ C, with ADP with ADP and GroES, with ATP, and with ATP and GroES.

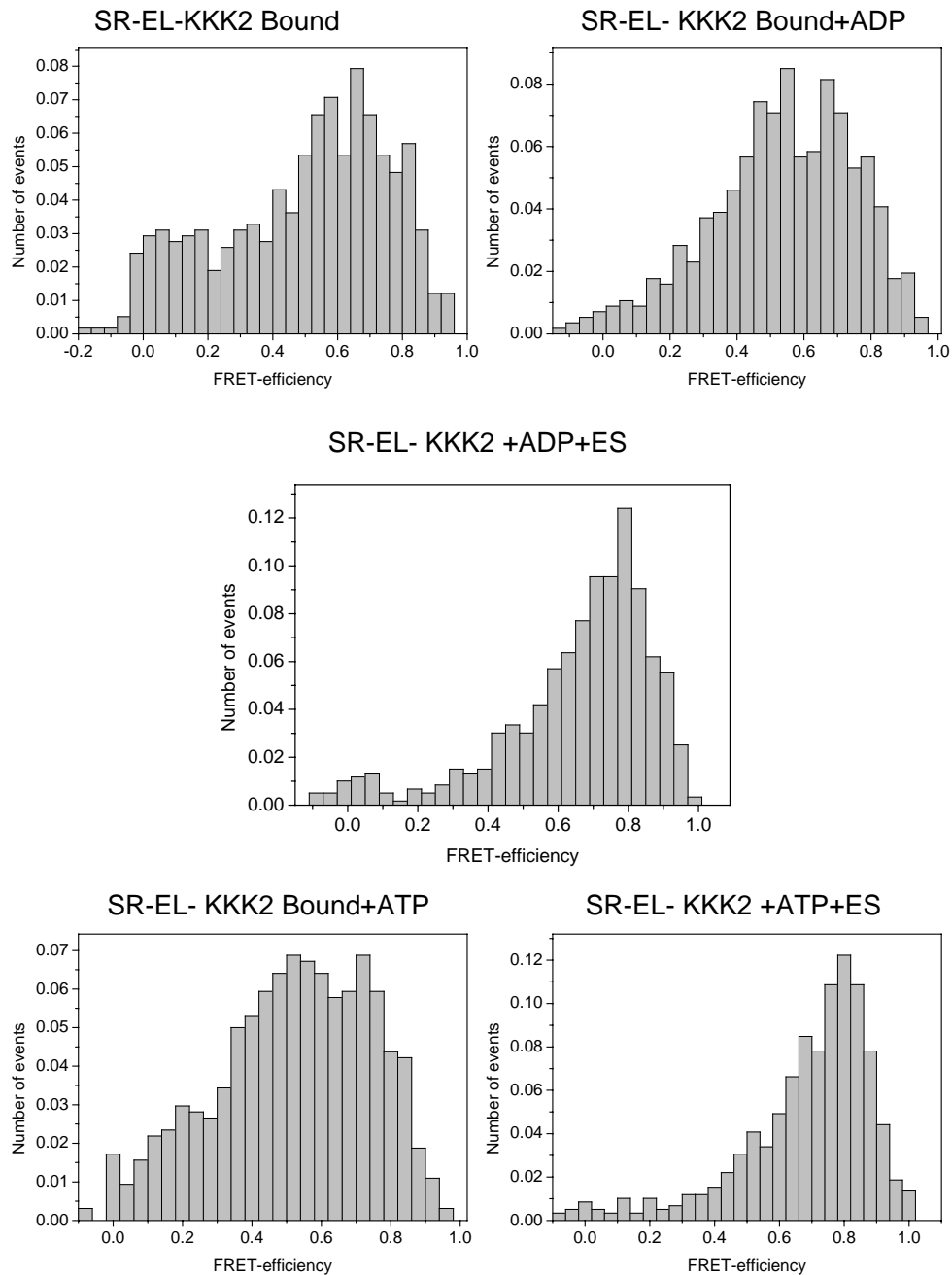


Figure A.16: Sp-FRET measurements of double-labelled MBP-DM(175 – 298) . FRET efficiency distribution is shown for the protein bound to SR-EL-KKK2, with ADP with ADP and GroES, with ATP, and with ATP and GroES.

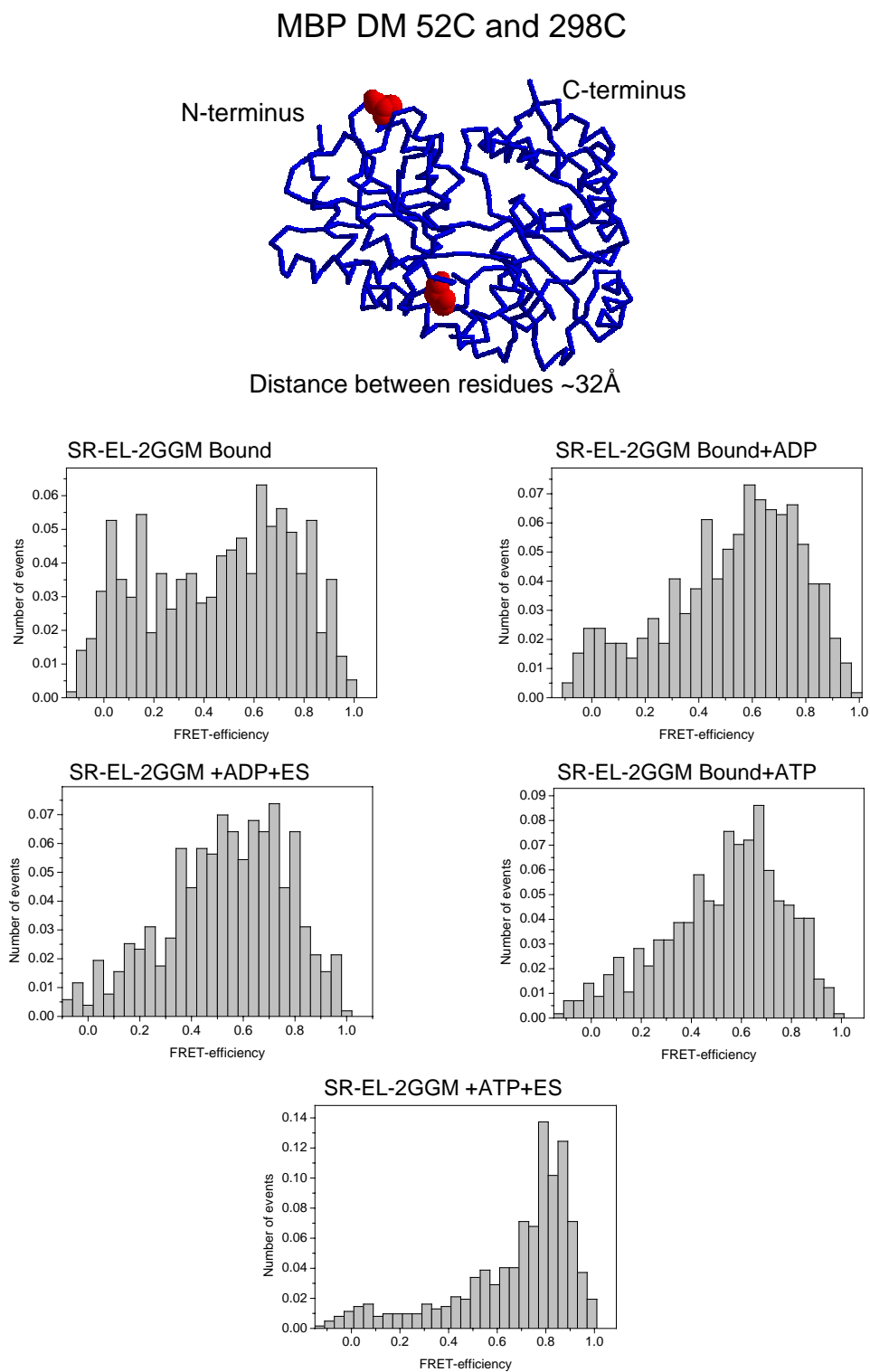


Figure A.17: Sp-FRET measurements of double-labelled MBP-DM(52 – 298) . FRET efficiency distribution is shown for the protein bound to SR-EL2GGM, with ADP with ADP and GroES, with ATP, and with ATP and GroES.

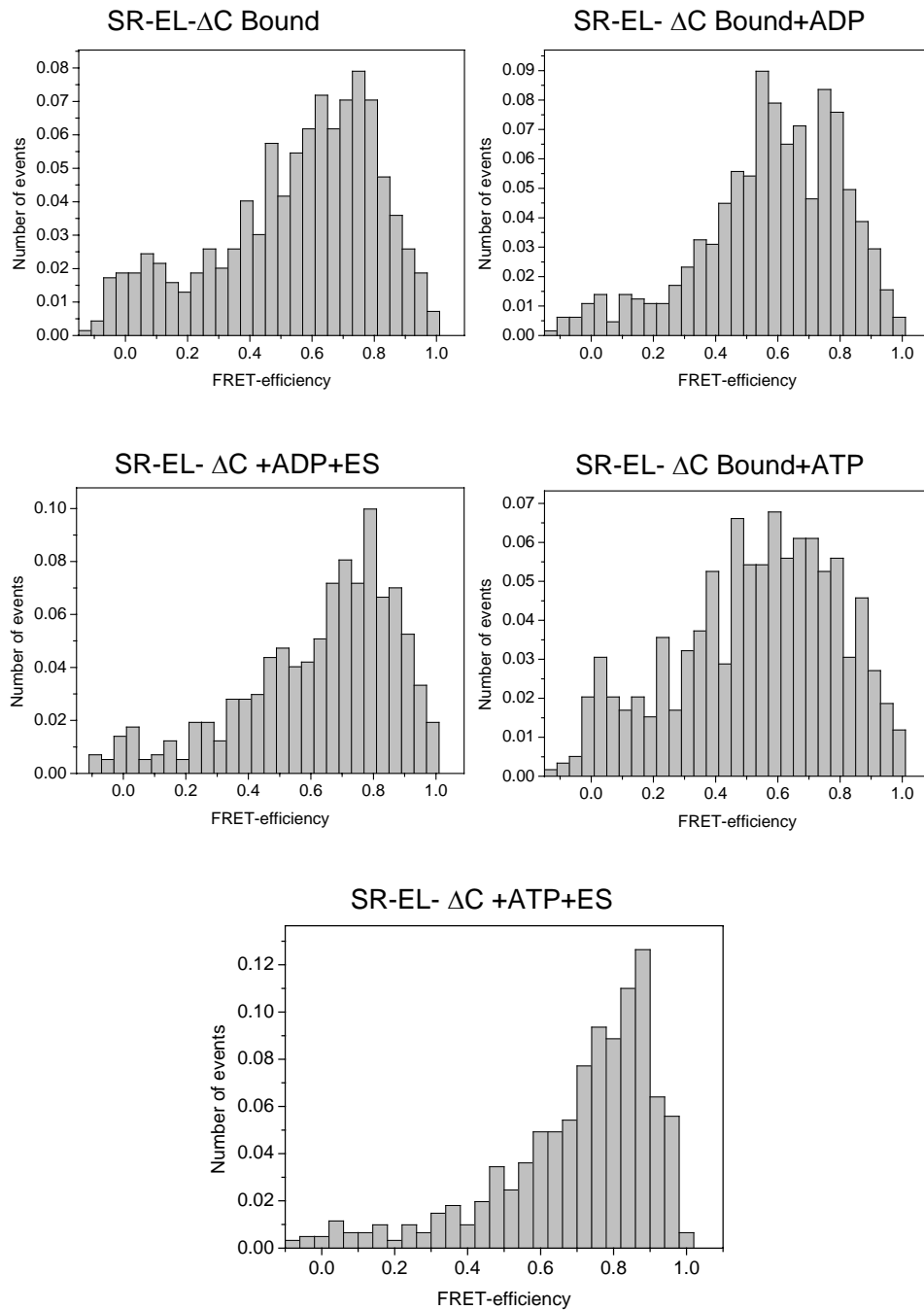


Figure A.18: Sp-FRET measurements of double-labelled MBP-DM(52 – 298) . FRET efficiency distribution is shown for the protein bound to SR-EL- δ C, with ADP with ADP and GroES, with ATP, and with ATP and GroES.

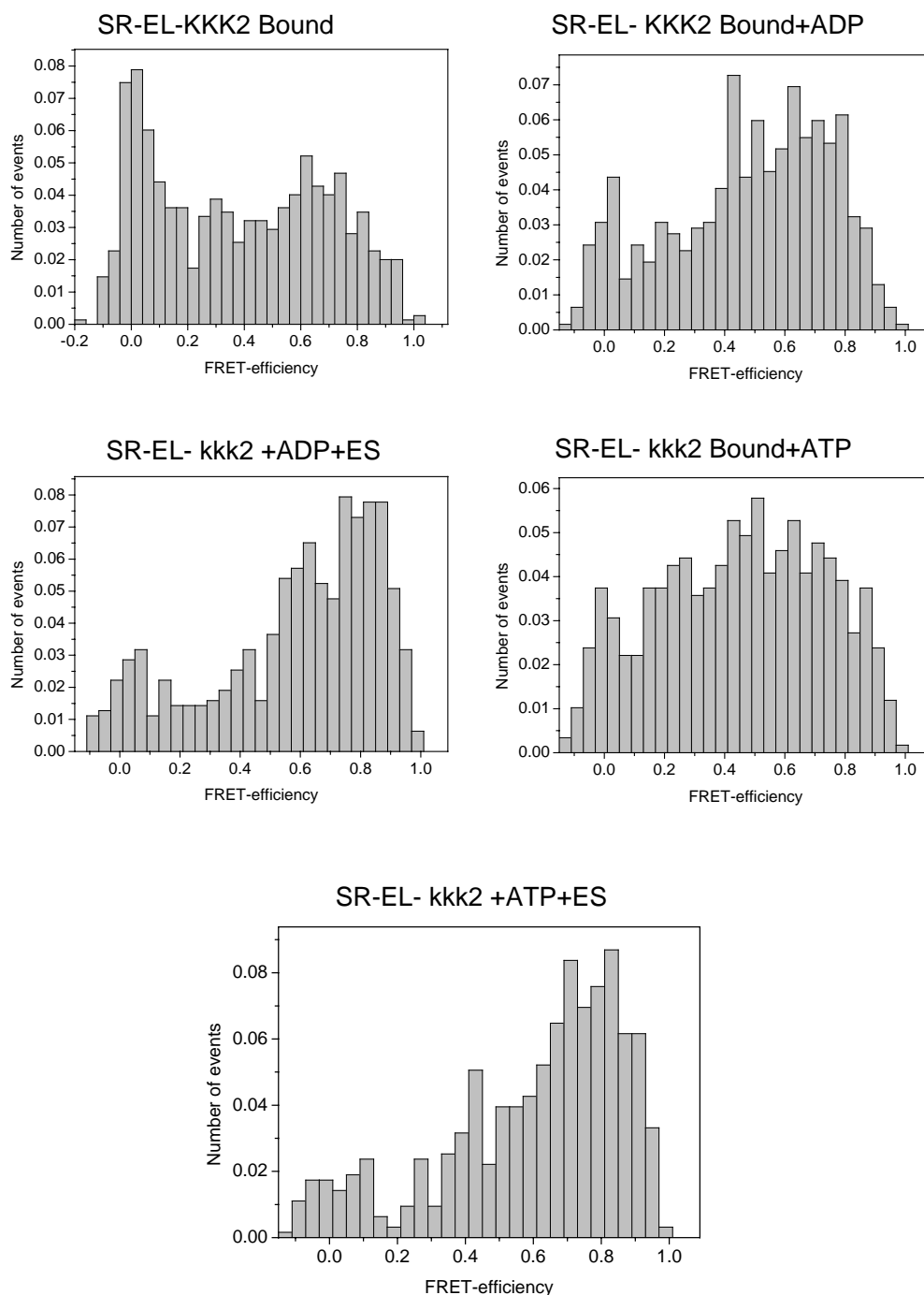


Figure A.19: Sp-FRET measurements of double-labelled MBP-DM(52 – 298) . FRET efficiency distribution is shown for the protein bound to SR-EL-KKK2, with ADP with ADP and GroES, with ATP, and with ATP and GroES.

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 with their one or three letter symbols.

Protein names are abbreviated according to their SWISSPROT database entries.

ADP	adenosine 5'-diphosphate
Amp	ampicillin
AMP-PNP	Adenosine 5' - (β , γ -imido)triphosphate
APS	ammonium peroxodisulfate
ATP	adenosine 5'-triphosphate
BLAST	Basic Local Alignment Search Tool
BSA	bovine serum albumin
Cam	chloramphenicol
CDTA	<i>trans</i> - 1, 2-diaminocyclohexane- <i>N</i> , <i>N</i> , <i>N'</i> , <i>N'</i> -tetraacetic acid
ClpB	chaperone clpB
D	donor
DM	Double Mutant
DNA	deoxyribonucleic acid
DnaJ	bacterial Hsp40 chaperone
DnaK	bacterial Hsp70 chaperone
DTT	dithiothreitol
<i>E.coli</i>	<i>Escherichiacoli</i>
EDTA	ethylenediaminetetraacetic acid
FCCS	Fluorescence Cross Correlation Spectroscopy
FCS	Fluorescence Correlation Spectroscopy
FPLC	Fast Protein Liquid Chromatography
FRET	Fluorescence Resonance Energy Transfer
g	acceleration of gravity, $9.81m/s^2$
GuHCl	guanidinium hydrochloride
GroEL	bacterial Hsp60 chaperonin
GroES	bacterial Hsp10 co-chaperonin
GrpE	bacterial nucleotide exchange factor of DnaK
hr	hour
Hz	Hertz

IPTG	isopropyl- β -D-1-thiogalactopyranoside
LB	Luria Bertani
MBP	Maltose Binding Protein
MOPS	3-(<i>N</i> -morpholino)propanesulfonic acid
OAc	acetate
OD	optical density
PAGE	PolyAcrylamide Gel Electrophoresis
PBS	phosphate buffered saline
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank. http://www.rcsb.org/pdb/
PIE	Pulsed Interleaved Excitation
SDS	sodiumdodecylsulfate
Sp-FRET	Single Pair FRET
TEMED	<i>N, N, N', N'</i> -tetramethylethylenediamine
TF	trigger factor
TFA	trifluoroacetic acid
Tris HCl	tris(hydroxymethyl)aminomethane hydrochloride

A.3 Curriculum Vitae

Name	Shruti Sharma
Date of Birth	25 July 1980
Place of Birth	Saharsa, Bihar, India
Gender	Female
Family Status	Single
Nationality	Indian

Scientific Experience

Sep. 2003 - today	PhD thesis in the department of "Cellular Biochemistry", Prof. Dr. F. Ulrich Hartl, at Max-Planck-Institut für Biochemie, Martinsried; titled : <i>Unfolding and Compaction in Chaperonin-assisted Protein Folding followed by Single molecule and Ensemble FRET.</i>
-------------------	---

University Education

Jan2003 -Jul2003	Masters thesis at Institute of Nuclear Medicine and Allied Sciences, Dr. S. Chandana; titled: <i>comparative protein profiling in Sf9 and BMG-1 cell lines under radiation stress.</i>
Aug 2001 - Jul 2003	Master of Science in Biomedical Sciences, from Ambedkar Center for Biomedical Research (ACBR), Delhi University, Delhi.
Jul 1998 - Jul 2001	Bachelor of Science in Microbiology, from Delhi University

Higher Education

Jun 1998	High School, Air Force Golden Jubilee Institute, Delhi.
----------	---

Bibliography

- Adesnik, M. and Levinthal, C. (1969). *J Mol Biol* 46 (2), 281–303.
- Agashe, V. R., Shastry, M. C., and Udgaonkar, J. B. (1995). *Nature* 377 (6551), 754–757.
- Aharoni, A. and Horovitz, A. (1996). *J Mol Biol* 258 (5), 732–735.
- Anfinsen, C. B. (1972). *Biochem J* 128 (4), 737–749.
- Anfinsen, C. B. (1973). *Science* 181 (96), 223–230.
- Bagshaw, C. R. and Cherny, D. (2006). *Biochem Soc Trans* 34 (Pt 5), 979–982.
- Baldwin, R. L. (1989). *Trends Biochem Sci* 14 (7), 291–294.
- Baldwin, R. L. (1996). *Proc Natl Acad Sci U S A* 93 (7), 2627–2628.
- Baldwin, R. L. and Rose, G. D. (1999). *Trends Biochem Sci* 24 (2), 77–83.
- Bochkareva, E. S., Horovitz, A., and Girshovich, A. S. (1994). *J Biol Chem* 269 (1), 44–46.
- Braig, K., Otwinowski, Z., Hegde, R., Boisvert, D. C., Joachimiak, A., Horwich, A. L., and Sigler, P. B. (1994). *Nature* 371 (6498), 578–586.
- Brinker, A., Pfeifer, G., Kerner, M. J., Naylor, D. J., Hartl, F. U., and Hayer-Hartl, M. (2001). *Cell* 107 (2), 223–233.
- Chaudhry, C., Horwich, A. L., Brunger, A. T., and Adams, P. D. (2004). *J Mol Biol* 342 (1), 229–245.
- Chaudhuri, T. K. and Gupta, P. (2005). *Cell Stress Chaperones* 10 (1), 24–36.
- Chen, L. and Sigler, P. B. (1999). *Cell* 99 (7), 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). *Nature* 371 (6494), 261–264.

- Chun, S. Y., Strobel, S., Bassford, P. J., and Randall, L. L. (1993). *J Biol Chem* 268 (28), 20855–20862.
- Collier, D. N., Bankaitis, V. A., Weiss, J. B., and Bassford, P. J. J. (1988). *Cell* 53 (2), 273–283.
- Dale, R. E., Eisinger, J., and Blumberg, W. E. (1979). *Biophys J* 26 (2), 161–193.
- Deniz, A. A., Dahan, M., Grunwell, J. R., Ha, T., Faulhaber, A. E., Chemla, D. S., Weiss, S., and Schultz, P. G. (1999). *Proc Natl Acad Sci U S A* 96 (7), 3670–3675.
- Dill, K. A. and Chan, H. S. (1997). *Nat Struct Biol* 4 (1), 10–19.
- Dinner, A. R., Sali, A., Smith, L. J., Dobson, C. M., and Karplus, M. (2000). *Trends Biochem Sci* 25 (7), 331–339.
- Dobson, C. M. and Karplus, M. (1999). *Curr Opin Struct Biol* 9 (1), 92–101.
- Doolittle, R. F. (1995). *Philos Trans R Soc Lond B Biol Sci* 349 (1329), 235–240.
- Ehrnsperger, M., Graber, S., Gaestel, M., and Buchner, J. (1997). *Embo J* 16 (2), 221–229.
- Ekman, D., Bjorklund, A. K., Frey-Skott, J., and Elofsson, A. (2005). *J Mol Biol* 348 (1), 231–243.
- Elad, N., Farr, G. W., Clare, D. K., Orlova, E. V., Horwich, A. L., and Saibil, H. R. (2007). *Mol Cell* 26 (3), 415–426.
- Ellis, R. J. and Hartl, F. U. (1996). *Faseb J* 10 (1), 20–26.
- Ewalt, K. L., Hendrick, J. P., Houry, W. A., and Hartl, F. U. (1997). *Cell* 90 (3), 491–500.
- Farr, G. W., Furtak, K., Rowland, M. B., Ranson, N. A., Saibil, H. R., Kirchhausen, T., and Horwich, A. L. (2000). *Cell* 100 (5), 561–573.
- Fayet, O., Ziegelhoffer, T., and Georgopoulos, C. (1989). *J Bacteriol* 171 (3), 1379–1385.
- Fenton, W. A., Kashi, Y., Furtak, K., and Horwich, A. L. (1994). *Nature* 371 (6498), 614–619.
- Ferguson, N., Capaldi, A. P., James, R., Kleanthous, C., and Radford, S. E. (1999). *J Mol Biol* 286 (5), 1597–1608.
- Ferguson, N. and Fersht, A. R. (2003). *Curr Opin Struct Biol* 13 (1), 75–81.

- Fersht, A. R. (1997). *Curr Opin Struct Biol* 7 (1), 3–9.
- Fersht, A. R., Matouschek, A., and Serrano, L. (1992). *J Mol Biol* 224 (3), 771–782.
- Fink, A. L. (1999). *Physiol Rev* 79 (2), 425–449.
- Freedman, R. B., Hirst, T. R., and Tuite, M. F. (1994). *Trends Biochem Sci* 19 (8), 331–336.
- Gallivan, J. P. and Dougherty, D. A. (1999). *Proc Natl Acad Sci U S A* 96 (17), 9459–9464.
- Genevaux, P., Keppel, F., Schwager, F., Langendijk-Genevaux, P. S., Hartl, F. U., and Georgopoulos, C. (2004). *EMBO Rep* 5 (2), 195–200.
- Glover, J. R. and Lindquist, S. (1998). *Cell* 94 (1), 73–82.
- Goloubinoff, P., Gatenby, A. A., and Lorimer, G. H. (1989). *Nature* 337 (6202), 44–47.
- Harrison, C. J., Hayer-Hartl, M., Di Liberto, M., Hartl, F., and Kuriyan, J. (1997). *Science* 276 (5311), 431–435.
- Hartl, F. U. (2002). *Biol Chem* 383 (10), 1481.
- Hayer-Hartl, M. K., Ewbank, J. J., Creighton, T. E., and Hartl, F. U. (1994). *Embo J* 13 (13), 3192–3202.
- Hayer-Hartl, M. K., Weber, F., and Hartl, F. U. (1996). *Embo J* 15 (22), 6111–6121.
- Hesterkamp, T. and Bukau, B. (1996). *FEBS Lett* 389 (1), 32–34.
- Hlodan, R., Tempst, P., and Hartl, F. U. (1995). *Nat Struct Biol* 2 (7), 587–595.
- Horovitz, A., Bochkareva, E. S., and Girshovich, A. S. (1993a). *J Biol Chem* 268 (14), 9957–9959.
- Horovitz, A., Bochkareva, E. S., Kovalenko, O., and Girshovich, A. S. (1993b). *J Mol Biol* 231 (1), 58–64.
- Horovitz, A., Bochkareva, E. S., Yifrach, O., and Girshovich, A. S. (1994). *J Mol Biol* 238 (2), 133–138.
- Horst, R., Bertelsen, E. B., Fiaux, J., Wider, G., Horwich, A. L., and Wuthrich, K. (2005). *PNAS* 102 (36), 12748–12753.
- Horwich, A. L., Low, K. B., Fenton, W. A., Hirshfield, I. N., and Furtak, K. (1993). *Cell* 74 (5), 909–917.

- Houry, W. A., Frishman, D., Eckerskorn, C., Lottspeich, F., and Hartl, F. U. (1999). *Nature* 402 (6758), 147–154.
- Inobe, T., Kikushima, K., Makio, T., Arai, M., and Kuwajima, K. (2003). *J Mol Biol* 329 (1), 121–134.
- Jamin, M. and Baldwin, R. L. (1996). *Nat Struct Biol* 3 (7), 613–618.
- Jiang, J., Prasad, K., Lafer, E. M., and Sousa, R. (2005). *Mol Cell* 20 (4), 513–524.
- 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. (2006). *Nature* 444 (7118), 455–460.
- Kalinin, S. and Johansson, L. B. (2004). *J Fluoresc* 14 (6), 681–691.
- 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., and Hartl, F. U. (2005). *Cell* 122 (2), 209–220.
- Kim, P. S. and Baldwin, R. L. (1982). *Annu Rev Biochem* 51, 459–489.
- Kim, P. S. and Baldwin, R. L. (1990). *Annu Rev Biochem* 59, 631–660.
- Kim, S. Y., Semyonov, A. N., Twieg, R. J., Horwich, A. L., Frydman, J., and Moerner, W. E. (2005). *J Phys Chem B Condens Matter Mater Surf Interfaces Biophys* 109 (51), 24517–24525.
- Knapp, S., Schmidt-Krey, I., Hebert, H., Bergman, T., Jornvall, H., and Ladenstein, R. (1994). *J Mol Biol* 242 (4), 397–407.
- Kramer, G., Patzelt, H., Rauch, T., Kurz, T. A., Vorderwulbecke, S., Bukau, B., and Deuerling, E. (2004a). *J Biol Chem* 279 (14), 14165–14170.
- Kramer, G., Rutkowska, A., Wegrzyn, R. D., Patzelt, H., Kurz, T. A., Merz, F., Rauch, T., Vorderwulbecke, S., Deuerling, E., and Bukau, B. (2004b). *J Bacteriol* 186 (12), 3777–3784.
- Laemmli, U. K. (1970). *Nature* 227 (5259), 680–685.
- Landry, S. J. and Gierasch, L. M. (1991). *Biochemistry* 30 (30), 7359–7362.
- Landry, S. J., Jordan, R., McMacken, R., and Gierasch, L. M. (1992). *Nature* 355 (6359), 455–457.
- Landry, S. J., Zeilstra-Ryalls, J., Fayet, O., Georgopoulos, C., and Gierasch, L. M. (1993). *Nature* 364 (6434), 255–258.

- Langer, T., Lu, C., Echols, H., Flanagan, J., Hayer, M. K., and Hartl, F. U. (1992). *Nature* 356 (6371), 683–689.
- Lee, N. K., Kapanidis, A. N., Wang, Y., Michalet, X., Mukhopadhyay, J., Ebright, R. H., and Weiss, S. (2005). *Biophys J* 88 (4), 2939–2953.
- Leroux, M. R., Fandrich, M., Klunker, D., Siegers, K., Lupas, A. N., Brown, J. R., Schiebel, E., Dobson, C. M., and Hartl, F. U. (1999). *Embo J* 18 (23), 6730–6743.
- Levinthal, C., Signer, E. R., and Fetherolf, K. (1962). *Proc Natl Acad Sci U S A* 48, 1230–1237.
- Liberek, K., Skowrya, D., Zylicz, M., Johnson, C., and Georgopoulos, C. (1991). *J Biol Chem* 266 (22), 14491–14496.
- Lin, Z. and Rye, H. S. (2004). *Mol Cell* 16 (1), 23–34.
- Macario, A. J., Malz, M., and Conway Macario, E. d. (2004). *Front Biosci* 9, 1318–1332.
- Maeder, D. L., Macario, A. J., and Macario, E. C. d. (2005). *J Mol Evol* 60 (3), 409–416.
- Martin, J., Langer, T., Boteva, R., Schramel, A., Horwich, A. L., and Hartl, F. U. (1991). *Nature* 352 (6330), 36–42.
- Mogk, A. and Bukau, B. (2004). *Curr Biol* 14 (2), R78–80.
- Motojima, F., Chaudhry, C., Fenton, W. A., Farr, G. W., and Horwich, A. L. (2004). *Proc Natl Acad Sci U S A* 101 (42), 15005–15012.
- Muller, B. K., Zaychikov, E., Brauchle, C., and Lamb, D. C. (2005). *Biophys J* 89 (5), 3508–3522.
- Netzer, W. J. and Hartl, F. U. (1997). *Nature* 388 (6640), 343–349.
- Orengo, C. A., Jones, D. T., and Thornton, J. M. (1994). *Nature* 372 (6507), 631–634.
- Ostermann, J., Horwich, A. L., Neupert, W., and Hartl, F. U. (1989). *Nature* 341 (6238), 125–130.
- Paci, E., Friel, C. T., Lindorff-Larsen, K., Radford, S. E., Karplus, M., and Vendruscolo, M. (2004). *Proteins* 54 (3), 513–525.
- Pack, C. G., Aoki, K., Taguchi, H., Yoshida, M., Kinjo, M., and Tamura, M. (2000). *Biochem Biophys Res Commun* 267 (1), 300–304.
- Panda, M. and Horowitz, P. M. (2002). *Biochemistry* 41 (6), 1869–1876.

- Pauling, L. and Corey, R. B. (1951a). *Proc Natl Acad Sci U S A* 37 (5), 282–285.
- Pauling, L. and Corey, R. B. (1951b). *Proc Natl Acad Sci U S A* 37 (5), 272–281.
- Pauling, L. and Corey, R. B. (1951c). *Proc Natl Acad Sci U S A* 37 (5), 241–250.
- Plocke, D. J., Levinthal, C., and Vallee, B. L. (1962). *Biochemistry* 1, 373–378.
- Poso, D., Clarke, A. R., and Burston, S. G. (2004). *J Mol Biol* 338 (5), 969–977.
- Pratt, W. B. and Toft, D. O. (2003). *Exp Biol Med (Maywood)* 228 (2), 111–133.
- Privalov, P. L. (1996). *J Mol Biol* 258 (5), 707–725.
- Ptitsyn, O. B. and Rashin, A. A. (1975). *Biophys Chem* 3 (1), 1–20.
- Puig, A., Lyles, M. M., Noiva, R., and Gilbert, H. F. (1994). *J Biol Chem* 269 (29), 19128–19135.
- Radford, S. E. (2000). *Trends Biochem Sci* 25 (12), 611–618.
- Radford, S. E., Buck, M., Topping, K. D., Dobson, C. M., and Evans, P. A. (1992a). *Proteins* 14 (2), 237–248.
- Radford, S. E., Dobson, C. M., and Evans, P. A. (1992b). *Nature* 358 (6384), 302–307.
- Ramachandran, G. N., Ramakrishnan, C., and Sasisekharan, V. (1963). *J Mol Biol* 7, 95–99.
- Richter, K. and Buchner, J. (2001). *J Cell Physiol* 188 (3), 281–290.
- Rye, H. S., Roseman, A. M., Chen, S., Furtak, K., Fenton, W. A., Saibil, H. R., and Horwich, A. L. (1999). *Cell* 97 (3), 325–338.
- Schechter, A. N., Chen, R. F., and Anfinsen, C. B. (1970). *Science* 167 (919), 886–887.
- Schellman, J. A. (1955). *C R Trav Lab Carlsberg [Chim]* 29 (14-15), 223–229.
- Scheraga, H. A., Konishi, Y., and Ooi, T. (1984). *Adv Biophys* 18, 21–41.
- Schirmer, E. C., Glover, J. R., Singer, M. A., and Lindquist, S. (1996). *Trends Biochem Sci* 21 (8), 289–296.
- Schmid, F. X. (1993). *Annu Rev Biophys Biomol Struct* 22, 123–142.
- Schultz, C. P. (2000). *Nat Struct Biol* 7 (1), 7–10.
- Shewmaker, F., Maskos, K., Simmerling, C., and Landry, S. J. (2001). *J Biol Chem* 276 (33), 31257–31264.

- Shortle, D. (1996). *Faseb J* 10 (1), 27–34.
- Shtilerman, M., Lorimer, G. H., and Englander, S. W. (1999). *Science* 284 (5415), 822–825.
- Siegert, R., Leroux, M. R., Scheufler, C., Hartl, F. U., and Moarefi, I. (2000). *Cell* 103 (4), 621–632.
- Sparrer, H. and Buchner, J. (1997). *J Biol Chem* 272 (22), 14080–14086.
- Spurlino, J. C., Lu, G. Y., and Quioco, F. A. (1991). *J Biol Chem* 266 (8), 5202–5219.
- Stan, G., Brooks, B. R., Lorimer, G. H., and Thirumalai, D. (2006). *Proc Natl Acad Sci U S A* 103 (12), 4433–4438.
- Staniforth, R. A., Burston, S. G., Atkinson, T., and Clarke, A. R. (1994). *Biochem J* 300 (Pt 3), 651–658.
- Steinberg, I. Z. (1971). *Annu Rev Biochem* 40, 83–114.
- Stoldt, V., Rademacher, F., Kehren, V., Ernst, J. F., Pearce, D. A., and Sherman, F. (1996). *Yeast* 12 (6), 523–529.
- Stryer, L. (1978). *Annu Rev Biochem* 47, 819–846.
- Studer, S., Obrist, M., Lentze, N., and Narberhaus, F. (2002). *Eur J Biochem* 269 (14), 3578–3586.
- Tang, Y. C., Chang, H. C., Roeben, A., Wischnewski, D., Wischnewski, N., Kerner, M. J., Hartl, F. U., and Hayer-Hartl, M. (2006). *Cell* 125 (5), 903–914.
- Taniuchi, H. and Anfinsen, C. B. (1969). *J Biol Chem* 244 (14), 3864–3875.
- Todd, M. J., Lorimer, G. H., and Thirumalai, D. (1996). *Proc Natl Acad Sci U S A* 93 (9), 4030–4035.
- Troullier, A., Reinstadler, D., Dupont, Y., Naumann, D., and Forge, V. (2000). *Nat Struct Biol* 7 (1), 78–86.
- Udgaonkar, J. B. and Baldwin, R. L. (1990). *Proc Natl Acad Sci U S A* 87 (21), 8197–8201.
- Vainberg, I. E., Lewis, S. A., Rommelaere, H., Ampe, C., Vandekerckhove, J., Klein, H. L., and Cowan, N. J. (1998). *Cell* 93 (5), 863–873.
- van der Vies, S. M., Viitanen, P. V., Gatenby, A. A., Lorimer, G. H., and Jaenicke, R. (1992). *Biochemistry* 31 (14), 3635–3644.

- van Montfort, R. L., Basha, E., Friedrich, K. L., Slingsby, C., and Vierling, E. (2001). *Nat Struct Biol* 8 (12), 1025–1030.
- Vendruscolo, M. and Dobson, C. M. (2005). *Philos Transact A Math Phys Eng Sci* 363 (1827), 433–50; discussion 450–2.
- Viitanen, P. V., Gatenby, A. A., and Lorimer, G. H. (1992). *Protein Sci* 1 (3), 363–369.
- Wang, J. D., Michelitsch, M. D., and Weissman, J. S. (1998). *Proc Natl Acad Sci U S A* 95 (21), 12163–12168.
- Weibezahn, J., Schlieker, C., Tessarz, P., Mogk, A., and Bukau, B. (2005). *Biol Chem* 386 (8), 739–744.
- Weissman, J. S., Rye, H. S., Fenton, W. A., Beechem, J. M., and Horwich, A. L. (1996). *Cell* 84 (3), 481–490.
- Wildegger, G. and Kiefhaber, T. (1997). *J Mol Biol* 270 (2), 294–304.
- Xu, Z., Horwich, A. L., and Sigler, P. B. (1997). *Nature* 388 (6644), 741–750.
- Yifrach, O. and Horovitz, A. (1995). *Biochemistry* 34 (16), 5303–5308.
- Yifrach, O. and Horovitz, A. (1996). *J Mol Biol* 255 (3), 356–361.
- Young, J. C., Agashe, V. R., Siegers, K., and Hartl, F. U. (2004). *Nat Rev Mol Cell Biol* 5 (10), 781–791.
- Young, J. C. and Hartl, F. U. (2003). *Cell* 113 (1), 1–2.
- Young, J. C., Moarefi, I., and Hartl, F. U. (2001). *J Cell Biol* 154 (2), 267–273.
- Zhu, X., Zhao, X., Burkholder, W. F., Gragerov, A., Ogata, C. M., Gottesman, M. E., and Hendrickson, W. A. (1996). *Science* 272 (5268), 1606–1614.
- Zimmerman, S. B. and Minton, A. P. (1993). *Annu Rev Biophys Biomol Struct* 22, 27–65.
- Zwanzig, R., Szabo, A., and Bagchi, B. (1992). *Proc Natl Acad Sci U S A* 89 (1), 20–22.