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

A single molecule study of the GroEL active cage mechanism

Amit Jean Gupta
aus
Ha'il, Saudi Arabien

2015

Erklärung

Diese Dissertation wurde im Sinne von §7 der Promotionsordnung vom 28. November 2011 von Herrn Prof. Dr. F. Ulrich Hartl betreut.

Eidesstattliche Versicherung

Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfe erarbeitet.

München, 18.06.2015 Amit Jean Gupta

Dissertation eingereicht am	04.08.2015
1. Gutachter	Prof. Dr. F. Ulrich Hartl
2. Gutachter	Prof. Dr. Walter Neupert
Mündliche Prüfung am	19.10.2015

The important thing is not to stop questioning. Curiosity has its own reason for existing. One cannot help but be in awe when he contemplates the mysteries of eternity, of life, of the marvelous structure of reality. It is enough if one tries merely to comprehend a little of this mystery every day. Never lose a holy curiosity.

Albert Einstein (1879-1955)

Acknowledgement

Foremost I would like to thank my supervisors Dr. Manajit Hayer-Hartl and Prof. Dr. Franz Ulrich Hartl for their support, help, advice and mentoring.

I would like to express my cordial thanks to Dr. Shubhasis Haldar, Goran Miličić and Yan Xiao for their collaboration on exploring the nature of GroEL assisted protein folding on the level of single molecules.

I thank Dr. Andreas Bracher, Prof. Dr. John Engen, Dr. Florian Georgescauld and Kristina Popova for their fruitful collaboration on investigating the assisted folding of TIM-barrel proteins by GroEL.

I thankfully acknowledge the great professional help of the Hartl-lab technical Staff: Emmanuel Burghardt, Nadine Wischnewski, Romy Lange and Anastasia Jungclaus.

I also thank the staff of the MPIB core facility, especially Prof. Dr. Dr. hc. Luis Moroder and Hans Musiol for the synthesis of β-semialdehyde, and Elisabeth Weyher-Stingl for support with mass spectrometric analysis of purified proteins and CD spectroscopy.

I thank the administrative Staff of the Hartl-department, Evelyn Frey-Roysten, Silke Leuze-Bütün and Darija Pompino for their tireless and dedicated help.

I thank all Hartl-lab members for interesting discussions on all aspects of life, especially those of scientific nature.

I am truly grateful to Dr. David Balchin and Dr. Christoph Klingner for critical discussion of this thesis.

I thank the Hartl-Lab Manhattan Group (David Balchin, Timm Hassemer, Thomas Hauser, Christian Löw, Tobias Neudegger and Leonhard Popilka) for friendship, moral support and their scientific expertise.

Finally, my deepest thanks go to my family and especially my wife Maike. Without their love, patience and constant support this work would not exist.

Contents

List of abbreviations

aa	amino acid
ADP	Adenosine 5’-diphosphate
ATP	Adenosine 5’-triphosphate
Bis-ANS	4,4’-Dianilino-1,1’-binaphthyl-5,5’-disulfonic acid dipotassium salt
CD	Circular dichroism
Cys	Cysteine
DapA	Dihydrodipicolinate synthase
dcFCCS	Dual color fluorescence cross-correlation spectroscopy
DM-MBP	Double mutant maltose binding protein
DNA	Deoxyribonucleic acid
DOL	Degree of labeling
DTT	Dithiothreitol
E. coli	*Escherichia coli*
FCS	Fluorescence correlation spectroscopy
FRET	Förster resonance energy transfer
GFP	Green fluorescent protein
GuHCl	Guanidine hydrochloride
Hepes	2-[4-(2-hydroxythyl)piperazin-1-yl]ethanesulfonic acid
Hsp	Heat shock protein
IPTG	Isopropyl β-D-thiogalactopyranoside

LB	Lysogenic broth
MWCO	Molecular weight cutoff
NADH	Nicotinamide adenine dinucleotide
PCR	Polymerase chain reaction
PEP	Phosphoenolpyruvate
PET	Photoinduced electron transfer
Pyk	Pyruvate kinase
RuBisCO	Ribulose 1,5-biphosphate carboxylase/oxygenase
SDS-PAGE	Sodiumdodecylsulfate polyacrylamid gelelectrophoresis
smFRET	Single molecule FRET
SR	Single ring
Tris	2-Amino-2-hydroxymethyl-propane-1,3-diol
Trp	Tryptophan

1 Summary

The cylindrical chaperonin GroEL and its lid-shaped cofactor GroES of *Escherichia coli* perform an essential role in assisting protein folding by transiently encapsulating non-native substrate in an ATP-regulated mechanism. It remains controversial whether the chaperonin system functions solely as an infinite dilution chamber, preventing off-pathway aggregation, or actively enhances folding kinetics by modulating the folding energy landscape. Here we developed single-molecule approaches to distinguish between passive and active chaperonin mechanisms. Using low protein concentrations to exclude aggregation, in combination with highly sensitive spectroscopic methods, such as single-molecule Förster resonance energy transfer (FRET) and fluorescence correlation spectroscopy (FCS), we measured the spontaneous and GroEL/ES-assisted folding of double-mutant maltose binding protein (DM-MBP), and a natural GroEL substrate - dihydrodipicolinate synthase (DapA). We show that both proteins form highly flexible, kinetically trapped folding intermediates, when folding in free solution and do not engage in inter-molecular interactions, such as aggregation, at sufficiently low concentration. We find that in the absence of aggregation, GroEL/ES accelerates folding of DM-MBP up to 8-fold over the spontaneous folding rate. The folding of DapA could be measured at physiological temperature and was found to be ~130-fold accelerated by GroEL/ES. As accelerated folding was independent of repetitive cycles of protein binding and release from GroEL, we demonstrate that iterative annealing does not significantly contribute to chaperonin assisted substrate folding. With a single molecule FRET based approach, we show that a given substrate molecule spends most of the time (~80%) during the GroEL reaction cycle inside the GroEL central cavity, in line with the inner GroEL cage being the active principle in folding catalysis. Moreover, photoinduced electron transfer experiments on DM-MBP provided direct experimental evidence that the confining environment of the chaperonin cage restricts polypeptide chain dynamics. This effect is mainly mediated by the net-negatively charged wall of the GroEL/ES cavity, as shown using the GroEL mutant EL(KKK2) in which the net-negative charge is removed.

Taken together, we were able to develop novel approaches, based on single molecule spectroscopy and making use of GroEL as a single molecule sorting machine, to measure GroEL substrate folding

rates at sub-nanomolar concentrations. We also, for the first time, provide direct experimental evidence of conformational restriction of an encapsulated polypeptide in a chaperonin cage. Our findings suggest that global encapsulation inside the GroEL/ES cavity, not iterative cycles of annealing and forced unfolding, can accelerate substrate folding by reduction of an entropic energy barrier to the folded state, in strong support of an active chaperonin mechanism. Accelerated folding is biologically significant as it adjusts folding rates relative to the rate of protein synthesis.

2 Introduction

2.1 Proteins and protein structures

Proteins are major components of almost all biological processes and acquire a defined three dimensional structure, which is in most cases inherently linked to their function. Considering protein structures, one usually distinguishes four different levels of structural organization. The linear sequence of amino acids in a protein is called the primary structure. Patterns of hydrogen bonds between the main chain N-H and C=O groups form local structural elements, the secondary structure. The most common secondary structure elements are α-helices, β-strands and turns. The secondary structure is packed into one or several globular units (domains), shaping the tertiary structure. As many proteins consist of more than one individual polypeptide chain, the spatial organization of the individual subunits comprises the quaternary structure of a protein.

Chemically, the C_α atoms of two adjacent amino acids are separated by three covalent bonds (C_α – C – N – C_α). All atoms, and in addition the oxygen atom of the carbonyl group and the hydrogen atom of the amide group, constitute a single two-dimensional plane. Rotation is possible around N – C_α and C_α – C bonds with the torsion angles φ (N – C_α) and ψ (C_α – C) (Fig. 2.1). The energetically limited possible combinations of torsion angles (for all amino acids except glycine), and the sequence of torsion angles within a polypeptide chain, define the secondary structure elements as denoted by the well known Ramachandran diagram (Ramachandran and Shasisekharan, 1968). Since proteins *in vivo* are synthesized on ribosomes in a vectorial manner, i.e. one amino acid being attached after another in a single direction, one of the fundamental questions in biochemistry is how proteins fold from a linear chain to a complex three dimensional structure with several levels of structural complexity.

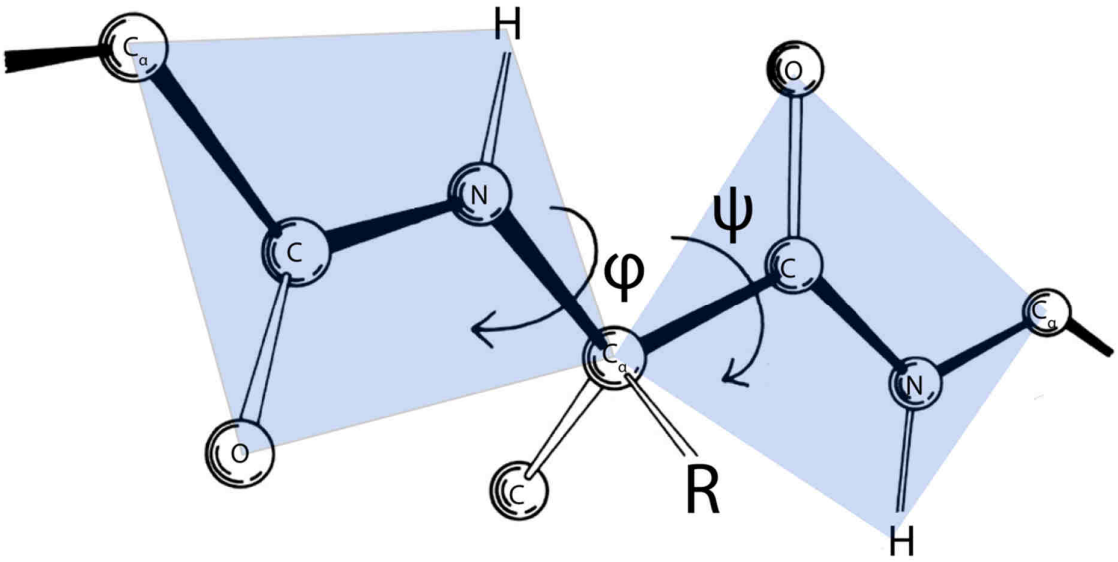

Figure 2.1Torsion angles in the protein backbone

Simplified illustration of a protein backbone with indicated torsion angles between two peptide bond planes (blue). The rotation around the peptide bond itself (C – N) is restricted. (Modified from Jane Shelby Richardson, Duke University http://kinemage.biochem.duke.edu/teaching/anatax/html/anatax.1b.html).

2.2 Protein folding

Pioneering experiments by Christian Anfinsen on the refolding of small proteins suggested that the information for the three dimensional structure of a protein is encoded in the primary sequence, and that the native state of a protein is usually the thermodynamically most stable state. Thus, the native state has a lower free energy than the unfolded state, making folding energetically favorable (Anfinsen, 1973). After decades of protein folding research, it is now clear that folding from an unfolded ensemble to the native state is energetically mainly driven by the burial of hydrophobic residues in the protein core, accompanied by a gain of entropy in the solvent. Additional factors that are usually considered are hydrogen bonding between residues and to the solvent, formation of salt bridges, covalent bonds (disulfides), van der Waals contacts between atoms, hydrophobic interactions, and importantly, the entropy of the protein chain itself.

In an unfolded protein chain, there is little restriction for free rotation. Considering a 100 residue polypeptide with 198 Φ and Ψ angles, even if every angle could only adopt two different values, there would be 2^{198} possible conformers. Even at very fast sampling rates at picosecond timescales, folding would be an astronomically slow process, if folding to the lowest energy state would be a randomized

trial and error process. Interestingly, however, as shown by Anfinsen and others, protein folding is usually a fast process that can occur within only milliseconds to seconds, at least for small single domain proteins. How exactly proteins fold on a fast timescale without having to randomly sample all possible conformers is a fundamental biological question and has become infamously known as the Levinthal Paradox (Levinthal, 1969).

A number of models were proposed to explain how proteins fold efficiently without having to sample a large number of conformers. An apparent solution is folding pathways that proceed through partially stabilized intermediate states with local, correctly folded structural elements, thereby reducing the amount of available conformers (Baldwin, 1996; Baldwin and Rose, 1999a, 1999b; Levinthal, 1968, 1969). Most prominently, the initial hydrophobic collapse of an unfolded polypeptide chain to a molten globule intermediate (Kauzmann, 1959; Tanford, 1962), and the early formation of hydrogen bonds in the protein backbone (Pace et al., 1996; Teufel et al., 2011), provide a plausible theoretical framework to explain the rapid search for the native state. It is likely that future advances in computational folding simulation will provide a more detailed description of the folding process.

An important approach to energetically describe the global folding process is the progression of a protein on several downhill routes, on a funnel-shaped, three dimensional, potential energy surface (Dill and Chan, 1997; Hartl et al., 2011). On such a three dimensional surface, each point represents a different conformational state with a respective free energy (Fig. 2.2). Usually the folding energy landscape is rugged, due to the presence of local energy minima. Such kinetic traps can slow the overall process of folding for mainly two reasons. Either the protein contains many long-range contacts in the native state and upon collapse adopts a globular collapsed state with large conformational entropy. In such a case the rate-limiting step of folding is the search for critical native contacts. On the other hand, proteins can populate misfolded intermediates by acquiring stabilizing non-native contacts, which must be broken to return to a productive folding pathway (Hartl et al., 2011).

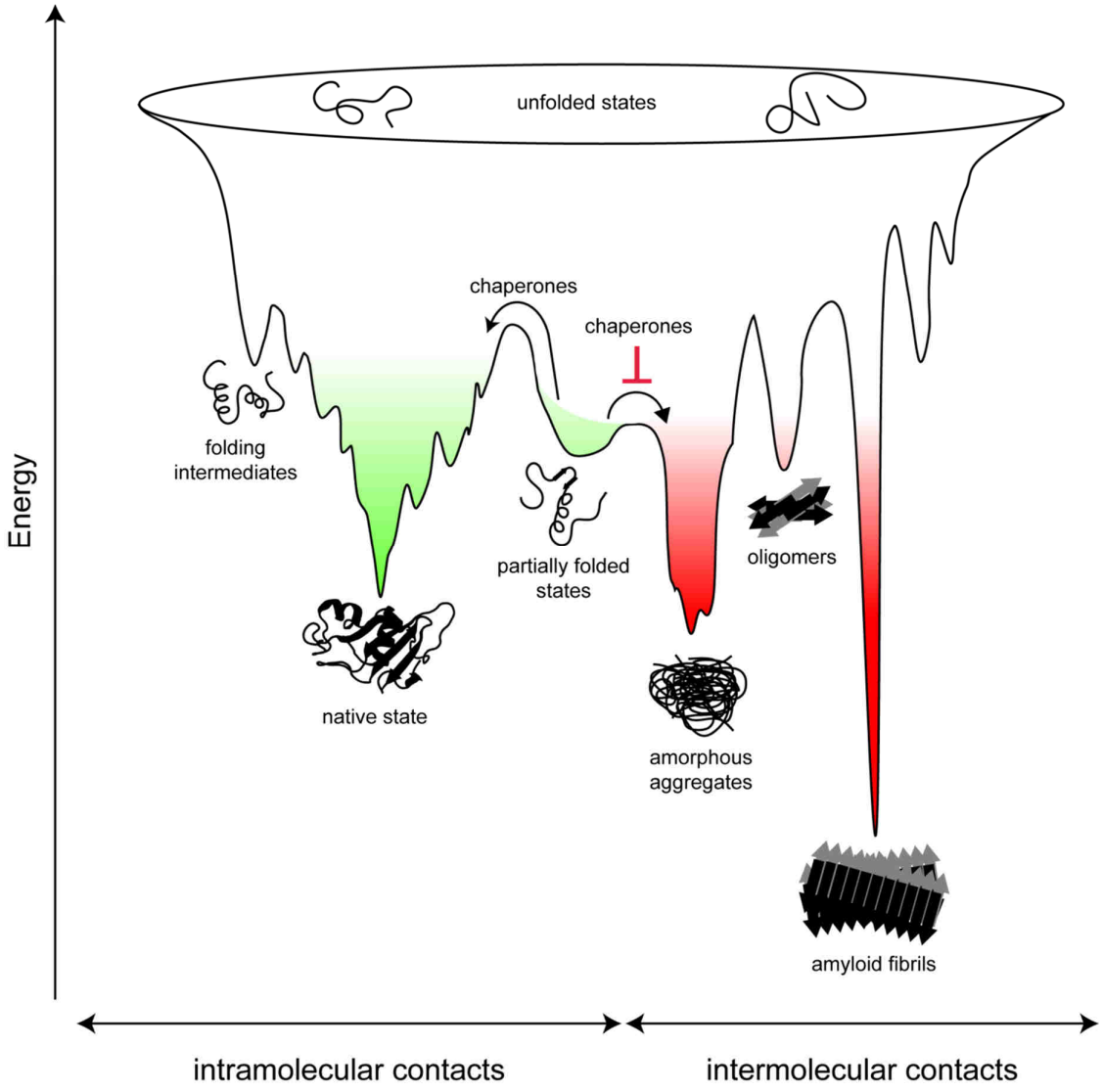

Figure 2.2 Illustration of a funnel shaped folding energy landscape

Schematic illustration of a funnel shaped folding energy landscape. Proteins that fold from an unfolded ensemble to the native state can proceed through local energy minima, kinetic traps, on their downhill path. Protein folding is governed by the formation of native intramolecular contacts. In case several proteins fold in the same space, such as the cytosol, intermolecular contacts can occur. The folding energy landscape can in such a case overlap with that of intermolecular aggregation. Aggregates can occur as small oligomers or as amorphous or fibrillar structures, large and stable protein deposits. Chaperones can interact with intermediate states and either prevent their aggregation or assist their productive folding. Figure was adapted and modified from (Hartl et al., 2011). The structure of barnase (PDB 2KF4) was used to render a model of a native protein structure with VMD.

The presence of long-lived intermediates, that expose hydrophobic side chains and non-structured protein backbone during folding, can lead to significant aggregation in a concentration dependent

manner (Eichner et al., 2011). Aggregates are often amorphous and driven by hydrophobic interactions. However, some proteins can also form ordered, amyloidogenic aggregates with cross-β-structres that have a high thermodynamic stability. Aggregates can have an even lower free energy than the native state, making disaggregation unfavorable and one of the major complications in protein folding. Furthermore, the pathological aggregation of certain proteins can lead to cell death and subsequently to neurodegenerative diseases such as morbus Huntigton or Alzheimer's disease (Chiti and Dobson, 2006; Dobson et al., 1998; Hartl et al., 2011; Kim et al., 2013).

2.2.1 Protein folding in the cell

Inside the cell, macromolecules occupy a substantial fraction of the total volume (Zimmerman and Minton, 1993). At concentrations of ~350 g/L (Zimmerman and Trach, 1991), the crowded cellular environment complicates protein folding, especially considering that during folding hydrophobic residues are exposed to the solvent. Furthermore, the excluded volume effect in highly crowded environments not only increases the affinity for favorable molecular interactions, but also, in the case of folding proteins, the aggregation propensity. Also, many proteins of considerable length and complexity fail to fold *in vitro* and are often subject to misfolding and aggregation (Hartl et al., 2011). As a further complication, when proteins are synthesized at an average rate of 20 aa per second (in *E. coli*), they emerge from the ribosomal exit tunnel as elongated chains. Therefore the protein chain must be maintained in an unfolded state until a sufficient amount of the sequence is translated before folding can be initiated. This is especially true for proteins with many long-range contacts, which is the case for more complex topologies. In addition, cellular stresses, such as elevated temperature or changes in pH, can induce protein misfolding and aggregation (Krishna et al., 2004). To cope with the problem of efficient folding of large proteins with complex topologies in a highly crowded environment, cells evolved an array of molecular chaperones, helper proteins that assist in de novo folding, refolding, assembly or transport.

2.3 Single molecule fluorescence research in protein folding

Usually, protein folding experiments in the presence and absence of chaperones are carried out at an ensemble level. The observation volume that is investigated is usually in the range of 1 µL to 1 mL and contains billions of molecules. The experimental readout, spectral information, enzymatic activity etc. is usually averaged over all protein molecules in the solution. The advantage of such ensemble approaches is a high signal to noise ratio (Mashaghi et al., 2014). On the other hand, information about molecular heterogeneity remains hidden, especially when complex structural rearrangements are investigated, as it is often the case when proteins interact with chaperones. Moreover, the heterogeneity

of intermediate states populated during folding cannot be revealed by ensemble measurements. Furthermore, ensemble experiments on proteins are usually carried out at high concentrations (nM to mM). This has the advantage of working at physiologically relevant concentrations, as they occur inside the cell. However, unwanted molecular interactions such as concentration dependent aggregation can cause a strong bias in folding experiments. Due to great technological advances during the last decades, it has become possible to investigate single molecules mainly by two distinct approaches, fluorescence spectroscopy and force spectroscopy. Isolation of single molecules is often either achieved by immobilization on a surface, by the use of optical tweezers, or by working at highly dilute concentrations in solution (<100 pM), using confocal laser spectroscopy. With immobilization, it is possible to investigate a single molecule over an extended period of time, while the immobilization itself might cause a defect in conformational flexibility. In the latter case it is possible to investigate molecules in solution, while the observation time is limited by the diffusion of molecules through the small observation volume (~1 fL), which usually occurs on the timescale of micro- to milliseconds. In addition, confocal single molecule spectroscopy requires working at low concentrations that are often unphysiological.

Single molecule fluorescence spectroscopy is especially attractive to investigate processes that occur with a high molecular heterogeneity or are accompanied by unproductive side reactions at higher concentrations. Both apply to protein folding, which explains the large number of published studies on protein folding using single molecule fluorescence methods; most notably, single molecule Förster resonance energy transfer (smFRET), fluorescence correlation and cross-correlation spectroscopy (FCS and FCCS), and photoinduced electron transfer (PET). All methods usually involve chemical modification of the protein of interest with fluorescent dyes, which are often attached to cysteine or lysine residues by maleimides or NHS-esters respectively. The presence of hydrophobic fluorophores can sometimes have a negative impact on protein structure, function, and folding. Important controls are therefore necessary to confirm the structural integrity of a protein after labeling.

FRET is based on the radiationless energy transfer from an excited donor molecule to a suitable acceptor molecule in close proximity. FRET is strongly distance dependent ($E \sim r^{-6}$) and sensitive to distance changes in the range of 2-10 nm, depending on the fluorophore pair (Fig. 2.3). Therefore large conformational changes and dynamics can be measured either by changes in fluorescence intensities of donor and acceptor, or changes in the fluorescence lifetime of the donor fluorophore. Such measurements can be performed on single molecules and are additionally simplified by elaborate excitation techniques, such as pulsed interleaved excitation (PIE) (Mashaghi et al., 2014; Müller et al., 2005). On the other hand, FRET is limited to measuring one distance, i.e. between donor and acceptor, and does not provide global information about protein conformational changes. In addition, smFRET

measured with confocal spectroscopy is based on the presence of one single molecule in the observation volume at a time, which can only be achieved by working at high dilutions, thereby reducing the statistical probability of a second molecule being present in the focal spot. Therefore macromolecular complexes are often inaccessible to single molecule FRET measurements. Single molecule FRET has proved to be particularly useful for the investigation of intrinsically disordered proteins, as these proteins are usually highly dynamic prone to large structural transitions (Mukhopadhyay et al., 2007).

Florescence correlation spectroscopy (FCS) is based on the analysis of fluctuations in the measured fluorescence signal, when a fluorescent molecule traverses the confocal observation volume (Haustein and Schwille, 2004) (Fig. 2.3). The observed fluctuations can be further processed to reveal information about diffusion speed, particle concentrations, rotational motion, molecular sizes and many other molecular processes (Kim et al., 2007). In protein folding, FCS is often used to measure changes in molecular size, when a protein is chemically denatured or diluted from denaturant (Haldar et al., 2010). A modification of FCS that uses cross correlation of the fluctuation in two different spectral channels, dual color florescence cross correlation spectroscopy (dcFCCS), can be used to investigate bimolecular interactions with high specificity and an excellent signal to noise ratio (Bacia and Schwille, 2007). As FCS is based on measuring fluctuations in the obtained fluorescence signal, there is a limitation for measuring FCS at high concentrations (minor fluctuations on a constantly high fluorescence signal) and also at low concentrations (too few diffusion events to obtain statistically relevant information within reasonable measurement times). Particle concentrations are ideal for FCS at an average of one molecule inside the focal observation volume.

Photoinduced electron transfer (PET) in combination with FCS is an emerging technique to investigate molecular processes at fast timescales (ns to µs) (Sauer and Neuweiler, 2014) (Fig. 2.3). PET is based on the fluorescence quenching of an oxazine fluorophore by van der Waals contact with, for example, a Trp residue. Such contacts result in direct transfer of electrons and subsequent quenching of the fluorophore. The transition of the fluorophore between bright and dark states can be observed as an additional exponential decay of the auto-correlation curve in FCS measurements. The relaxation time of the PET-induced signal depends on the timescale of contact formation between dye and Trp, making PET-FCS an ideal method to monitor structural rearrangements at fast timescales. Furthermore, in contrast to FRET, PET-FCS is sensitive to conformational changes <2 nm. PET-FCS was used to monitor loop closure events in peptides during early events in protein folding (Teufel et al., 2011), and structural plasticity in unfolded proteins (Neuweiler et al., 2007). Moreover, PET-FCS was used to resolve the structural transition between two distinct conformational states (Frank et al., 2010).

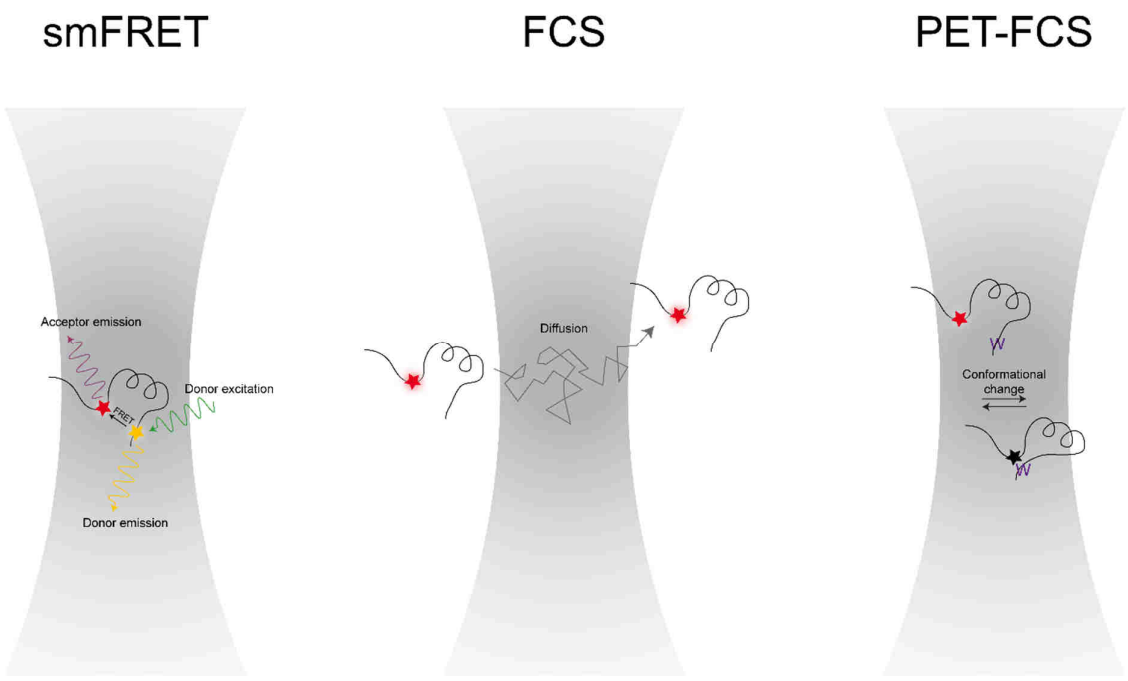

Figure 2.3 Single molecule fluorescence spectroscopy

Principle of the three confocal microscopy based spectroscopic methods that are most commonly used for single molecule investigation of chaperones. In smFRET, the distance dependent energy transfer between donor and acceptor fluorophore is determined by measuring the fluorescence intensities of both fluorophores upon donor excitation. In a PIE setup, acceptor excitation is used as an additional control. FCS can be used to measure all processes that induce fluorescence fluctuations in the focal spot. Most prominently, diffusion coefficients are extracted based on diffusion time and the size of the confocal observation volume. PET-FCS uses the quenching of an oxazine fluorophore in van-der-Waals distance to Trp residues to analyze conformational changes on short timescales (ns-µs). The signal fluctuations induced by changing from a fluorescent to a dark conformer can be analyzed by FCS.

Taken together, confocal single molecule fluorescence spectroscopy provides an array of useful tools to investigate structural transitions in proteins. Methods like smFRET, FCS or PET-FCS are particularly suited to investigating protein folding processes.

2.4 Molecular chaperones

Chaperones were first discovered as proteins with elevated cellular expression levels upon heat stress and therefore dubbed heat shock proteins (HSPs) (Tissières et al., 1974). It was later discovered that chaperones are crucial for the folding and/or assembly of certain client proteins (Cheng et al., 1989; Goloubinoff et al., 1989; Ostermann et al., 1989). Molecular chaperones are now known to be key players in the maintenance of cellular proteostasis and form a network inside the cell, guiding newly

synthesized substrate proteins to the finally folded native state (Fig. 2.4). As a general feature, many chaperones assist client folding by the recognition of exposed hydrophobic residues and subsequent ATP regulated cycles of binding and release (Mayer, 2010). The general organization of chaperone pathways is conserved throughout evolution (Kim et al., 2013) (Fig. 2.4). In all domains of life, ribosome bound chaperones, such as trigger factor (TF) and the nascent chain associated complex (NAC), are the first chaperones encountered by a nascent polypeptide. A second, non-ribosome-bound set of chaperones (the Hsp70 system) can interact already co-translationally. Folding then occurs either co-translationally or post-translationally. Some proteins require further assistance from downstream chaperone systems, such as Hsp90 or the chaperonins.

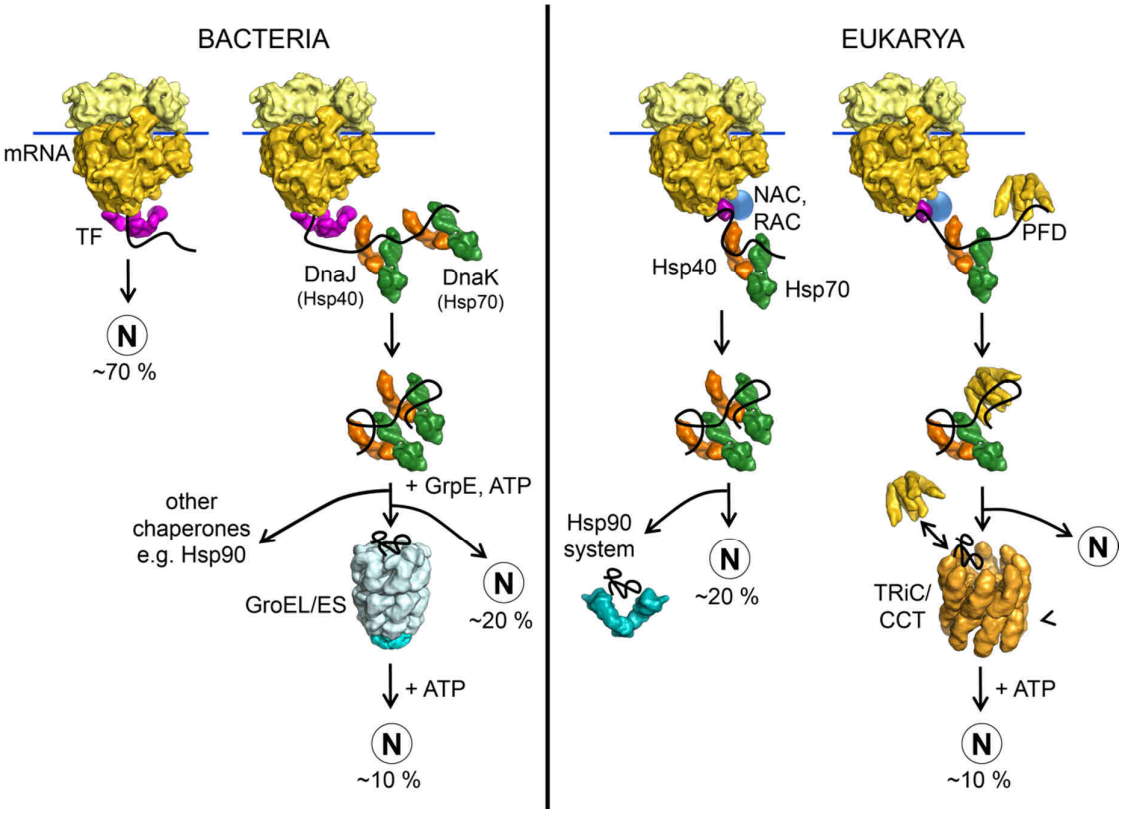

Figure 2.4 Cytosolic chaperone networks

The cytosolic chaperone networks in bacteria and eukarya are evolutionary conserved. Ribosome bound chaperones initially recognize emerging polypeptides. Hsp70 functions as a second chaperone system for longer nascent chains, and facilitates co-and post-translational folding. Hsp70 also cooperates with other downstream chaperones, such as Hsp90 or the chaperonins, in effective folding of the cellular proteome. The respective interacting fraction for a given chaperone is indicated in percent of the whole proteome. N = native state, TF = trigger factor, PFD = prefoldin. Figure was adapted from (Hartl et al., 2011) and further modified.

In general, two major functional principles of chaperone function are known. On the one hand client proteins go through repetitive cycles of binding to the respective chaperone, followed by release into free solution. Hsp70 and Hsp90 are prominent examples. On the other hand, client proteins can also be released into evolved cages, folding competent micro-compartments, as in case of Hsp60.

2.5 Ribosome associated chaperones

As soon as they emerge from the ribosomal exit tunnel, newly synthesized proteins encounter a first set of ribosome associated chaperones, e.g. trigger factor (TF) in bacteria and the nascent chain associated complex (NAC) in eukarya (Hartl et al., 2011). As translation occurs linearly, while folding is a 3-dimensional process, early chaperone action on nascent polypeptides, by shielding hydrophobic residues, is necessary to prevent premature folding and unfavorable interactions. In *E. coli*, TF is associated with the large ribosomal subunit, close to the ribosomal exit tunnel (Merz et al., 2008), and binds to nascent chains of ~100 residues length upon encountering the first hydrophobic segments of the emerging polypeptide chain (Oh et al., 2011). Release of trigger factor from nascent chains is ATP independent and allows folding or transfer to downstream chaperones. Deletion of TF in *E. coli* is only lethal upon co-deletion of DnaK, the major Hsp70, and vice versa, indicating that these chaperone systems are partially redundant (Calloni et al., 2012; Genevaux et al., 2004).

2.6 The HSP70 machinery

The Hsp70 class of chaperones is one of the most versatile in that it is involved not only in de novo folding of client substrates but also in membrane translocation, protein degradation and transport processes. In *E. coli* the major cytosolic Hsp70, DnaK, was described as a central hub of the cytosolic chaperone network (Calloni et al., 2012). In its functional cycle, DnaK cooperates with two co-factors. DnaJ, an Hsp40 protein, and GrpE, a nucleotide exchange factor (NEF).

Hsp70 proteins consist of two functional domains connected by a hydrophobic linker region: an N-terminal nucleotide binding domain (NBD) and a C-terminal substrate binding domain (SBD) (Bukau and Horwich, 1998) (Fig. 2.5). The NBD, by hydrolyzing ATP, regulates functional substrate cycling. The SBD consists of a β-sandwich domain with the substrate binding site, connected to a lid-like α-helical domain. The SBD binds 5-7 residue peptide stretches in substrate proteins by hydrogen bonding and van der Waals interaction with hydrophobic side chains. The substrate binding sites are usually enriched in hydrophobic residues, flanked by positively charged amino acids (Bukau and Horwich, 1998; Mayer, 2010).

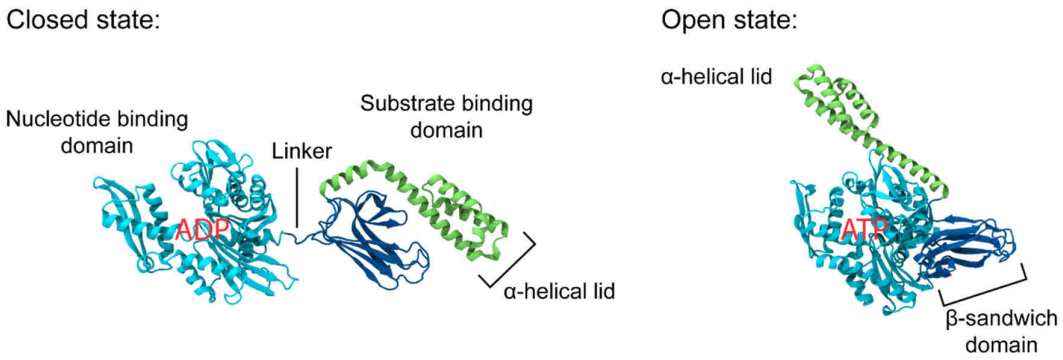

Figure 2.5 Structure of Hsp70

Hsp70 consists of two domains connected by a conserved linker, a nucleotide binding domain (NBD, cyan) and a substrate binding domain (SBD, green and dark blue). In absence of nucleotide or presence of ADP, Hsp70 adopts a closed conformation (PDB 2KHO (Bertelsen et al., 2009)). The α-helical lid-domain (green) is in close proximity to the β-sandwich-domain (dark blue). In the presence of ATP, Hsp70 adopts an open conformer (PDB 4B9Q (Kityk et al., 2012)) with the lid-domain contacting the nucleotide binding domain (cyan). Figure was adapted from (Kim et al., 2013) and rendered with VMD. VMD is developed with NIH support by the Theoretical and Computational Biophysics group at the Beckman Institute, University of Illinois at Urbana-Champaign.

During functional cycling, ATP binding to the NBD triggers opening of the peptide binding pocket by conformational rearrangement, resulting in attachment of the inter-domain linker and the α-helical lid to the NBD (Kityk et al., 2012; Zhuravleva and Gierasch, 2011). Hydrolysis of ATP to ADP then results in detachment of the lid-domain from the NBD and subsequent closure of the SBD (Bertelsen et al., 2009; Mapa et al., 2010) (Fig. 2.6). Hsp40 molecules (a large family of Hsp70 co-chaperones) are chaperones in their own right, and act in recruiting the Hsp70 system to client proteins, thereby providing a scaffold which dictates Hsp70 substrate specificity. Many Hsp40 proteins, amongst them DnaJ, bind to the Hsp70 NBD as well as the hydrophobic linker segment and thereby strongly stimulate ATP hydrolysis in the NBD, which tightens the interaction between Hsp70 and substrate in the closed conformation (Hartl and Hayer-Hartl, 2009; De Los Rios and Barducci, 2014). Association of a NEF induces exchange of ADP to ATP, completing the functional cycle. By going through consecutive cycles of high and low substrate affinity, i.e. binding and release of hydrophobic segments, Hsp70 has a strong impact on the folding of its client proteins (Kim et al., 2013).

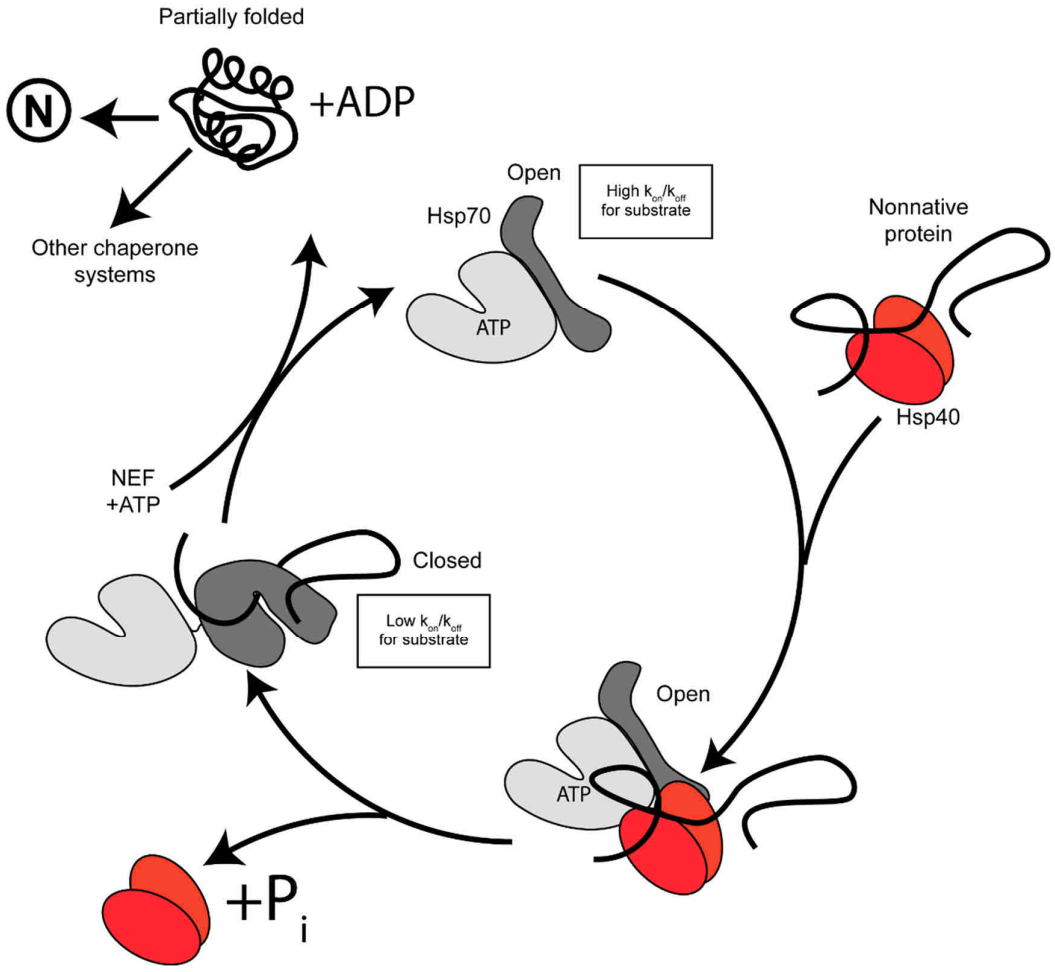

Figure 2.6 Functional cycle of Hsp70

Functional cycle of Hsp70. ATP binding to the Hsp70 NBD stabilizes the open state, facilitating substrate binding to the SBD. Substrates can be recruited to Hsp70 by Hsp40 co-chaperones. Hsp40 stimulates Hsp70 ATP hydrolysis, resulting in a conformational rearrangement mainly of the α-helical lid-domain, which closes over the bound substrate peptide. NEFs can stimulate ADP release from the Hsp70 NBD. Subsequent ATP binding induces substrate release and completes the functional reaction cycle. Figure was adapted from (Kim et al., 2013).

The important structural rearrangements and the heterogeneity of conformational changes that are otherwise difficult to study, have made the Hsp70 system an ideal target for single molecule FRET based studies (Böcking et al., 2011; Kellner et al., 2014; Mapa et al., 2010; Marcinowski et al., 2011; Sikor et al., 2013). For example, single molecule FRET distributions revealed previously undiscovered conformational heterogeneity of Hsp70 in the ADP bound state (Mapa et al., 2010).

Importantly, the *E. coli* Hsp70 DnaK is also prominently involved in stabilizing substrates for subsequent folding by the downstream GroEL/ES system (Calloni et al., 2012).

2.7 The chaperonin machinery

The chaperonins form large double ring assemblies with 7-9 ~60 kDa subunits per ring. They are unique in that they provide protein folding nano-cages for substrate proteins to fold in isolation, unimpaired by aggregation (Kim et al., 2013). Chaperonins are divided into two distantly related groups: group I and group II (Horwich et al., 2007; Tang et al., 2007).

2.7.1 Group I chaperonins

The group I chaperonins consist of heptameric rings and are present in the bacterial cytosol (GroEL) as well as in organelles of endosymbiotic origin, i.e. in the mitochondrial matrix (Hsp60) and in the chloroplast stroma (Cpn60) (Horwich et al., 2007). To form an enclosed cage for substrate encapsulation, group I chaperonins depend on the presence of a lid-like cofactor (GroES in bacteria, Hsp10 in mitochondria and Cpn10/Cpn20 in chloroplasts) (Kim et al., 2013). The functional reaction cycle of group I chaperonins involves closing and opening of a central cavity by cycles of association and dissociation of the respective co-factor in an ATP dependent manner (Hartl et al., 2011). In the closed state, an encapsulated substrate protein can fold inside the central cavity.

2.7.2 Group II chaperonins

The group II chaperonins consist of two octa- or nonameric rings and occur in the eukaryotic cytosol (TRiC/CCT) and in archaea (thermosome). Often, group II chaperonins are hetero-oligomers. The eukaryotic, hexadecameric TRiC/CCT for example consists of 8 different subunits (Leitner et al., 2012). In contrast to group I chaperonins, group II chaperonins contain a built-in lid in the form of an apical protrusion, replacing an additional cofactor. Opening and closing of the central cavity also involves an ATP dependent mechanism. In group II chaperonins, ATP hydrolysis triggers closing of the cage by structural rearrangements in an iris-like fashion (Meyer et al., 2003).

2.8 GroEL – the most widely studied chaperonin

The *E. coli* cytosolic chaperonin GroEL is the most intensely studied group I chaperonin, with ~2700 entries in PubMed for the search term “GroEL” (2015). The cylindrical GroEL homotetradecamer consists of two heptameric rings of identical ~57 kDa subunits stacked back to back in a staggered conformation, with one subunit in one ring interdigitating between two subunits in the opposing ring (Braig et al., 1994) (Fig. 2.7).

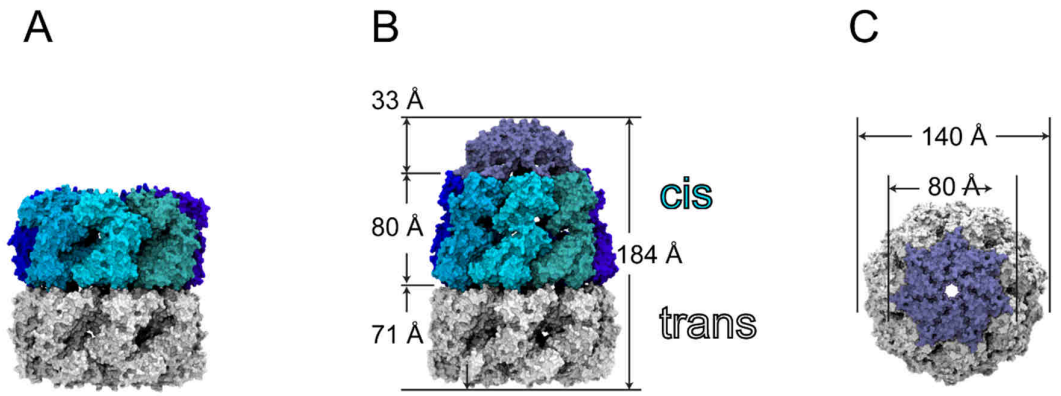

Figure 2.7 Structure and dimensions of apo GroEL and the GroEL/ES complex

(A) Apo conformation of GroEL in absence of nucleotide (PDB 1XCK (Bartolucci et al., 2005)), with one ring colored in grey and the subunits in the other ring colored in shades of blue. **(B)** Structure and dimensions of the ADP/GroES bound GroEL complex (PDB 1AON (Xu et al., 1997)). The subunits of the GroES-liganded cis-ring are colored in different shades of blue, while GroES itself is colored in purple. The unliganded GroEL trans-ring is colored in grey. **(C)** Top view of the GroEL/ES complex as shown in (B). GroEL is colored in grey and GroES in purple. All structured were rendered with VMD. Dimensions from (Xu et al., 1997).

The individual subunit is composed of three functional domains: the equatorial domain, the apical domain, and the intermediate hinge domain (Fig. 2.9). The equatorial domain harbors the nucleotide binding site and contains the major interfaces for inter-ring and intra-ring contacts to neighboring subunits. The intermediate domain serves as a linker, transmitting structural changes from equatorial to apical domains. The apical domains of GroEL line the entrance to the central cavity and contain hydrophobic segments for recognition and binding of non-native substrate proteins, as well as GroES, a lid-shaped heptameric co-factor of ~10 kDa subunits. The 23 C-terminal residues of GroEL protrude from the equatorial domains into the central cavity. These segments are largely unstructured and terminate in four Gly-Gly-Met repeats. The role of these flexible C-termini is still unclear. They were suggested to assist in protein folding itself, in efficient encapsulation of substrate protein, and in providing a barrier between the two GroEL cavities (Chen et al., 2013; Dalton et al., 2015; Ishino et al., 2015; Tang et al., 2006).

Each subunit of GroES consists of a single domain, containing nine β-strands, and one highly flexible, 16 amino acid loop, which forms the GroEL binding motif (Landry et al., 1993). This loop region is largely unstructured in unliganded GroES but forms a stable hairpin structure upon binding to GroEL (Shewmaker et al., 2001; Xu et al., 1997).

2.8.1 The GroEL reaction cycle

The reaction cycle of GroEL-assisted protein folding is strongly linked to the GroEL ATPase function. The binding of ATP to GroEL occurs with an intra-ring positive cooperativity and an inter-ring negative cooperativity. Due to this allosteric regulation of the GroEL ATPase cycle, GroEL acts like a two-stroke engine, with only one ring operating at a time, and the two rings operating in an alternating fashion (Horovitz and Willison, 2005) (Fig. 2.8). Binding of ATP to GroEL results in large structural rearrangements, priming the apical domains for binding of GroES. Binding of GroES to the apical domains then results in displacement of the substrate protein into the now hydrophilic cage, and leads to formation of an asymmetric complex, with the GroES bound ring being called cis-ring and the unliganded ring called trans-ring. ATP hydrolysis in the cis-ring takes ~10 sec at 25°C in absence of substrate (Tang et al., 2006) and ~2.2 sec at 37°C in presence of substrate (Gupta et al., 2014), giving the encapsulated substrate time for folding in isolation. The functional cycle is then completed by binding of ATP and GroES to the trans-ring, inducing release of GroES, substrate and ADP from the former cis-ring. Substrate proteins that could not fold during encapsulation are rapidly recaptured and can undergo subsequent folding attempts.

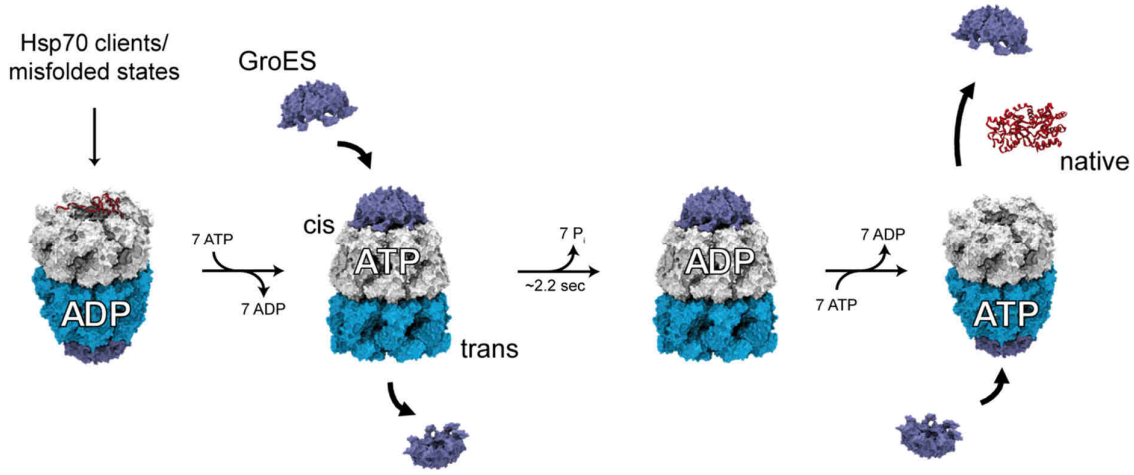

Figure 2.8 The GroEL reaction cycle

The conventional GroEL reaction cycle is initiated by binding of a non-native substrate protein to the apical domains of the trans ring. Non-native substrates are delivered to GroEL by the upstream Hsp70 system, or by TF. Binding of 7 ATP molecules and GroES displace the substrate into the central cavity. The protein is encapsulated for at least the duration of ATP hydrolysis, ~2.2 sec at 37°C in presence of substrate (Gupta et al., 2014). Binding of 7ATP and GroES to the opposing ring triggers release of ADP, GroES and substrate. In case the substrate could not fold to the native state, it is rapidly recaptured and undergoes subsequent rounds of folding. Substrate is colored in red, structures (PDB 1AON (Xu et al., 1997), 1OMP (Sharff et al., 1992)) were rendered in VMD.

The structure of GroEL in different nucleotide-bound states and in the presence and absence of GroES was extensively studied by electron microscopy (Braig et al., 1993; Langer et al., 1992; Saibil et al., 1991) and crystallography (Braig et al., 1994; Xu et al., 1997). Together, the available structural information results in a detailed model of structural rearrangements during the GroEL/ES reaction cycle.

Substrate binding to GroEL occurs with a high affinity in the apo-state, while the affinity is reduced in the nucleotide bound state. Binding to the apical domains is mainly mediated by the GroEL helices H (residues 233-243) and I (residues 255-267) (Fenton et al., 1994). These two helices expose hydrophobic residues to the central cavity, creating a hydrophobic interaction surface for substrate proteins. Substrate binding usually occurs to multiple subunits within one ring (Elad et al., 2007; Horwich et al., 2007; Kim et al., 2013) in a molten-globule like conformation, lacking stable tertiary elements (Hartl, 1996; Hillger et al., 2008; Horst et al., 2005; Sharma et al., 2008). Interestingly, the binding of collapsed states to GroEL is accompanied by an overall structural expansion, with a further expansion occurring in the context of the structural transitions of GroEL upon nucleotide binding (Hofmann et al., 2010; Lin et al., 2013; Sharma et al., 2008).

The important conformational rearrangement of GroEL subunits from the trans- to the cis-state is initiated by the cooperative binding of ATP to one GroEL ring. In the apo-state, the GroEL subunits are in equilibrium between a T state (low affinity for ATP) and an R state (high affinity for ATP). Binding of ATP with positive cooperativity within one ring stabilizes all subunits in the R state. Importantly, negative cooperativity between the two rings locks the trans-ring subunits in the T state, ensuring complex asymmetry.

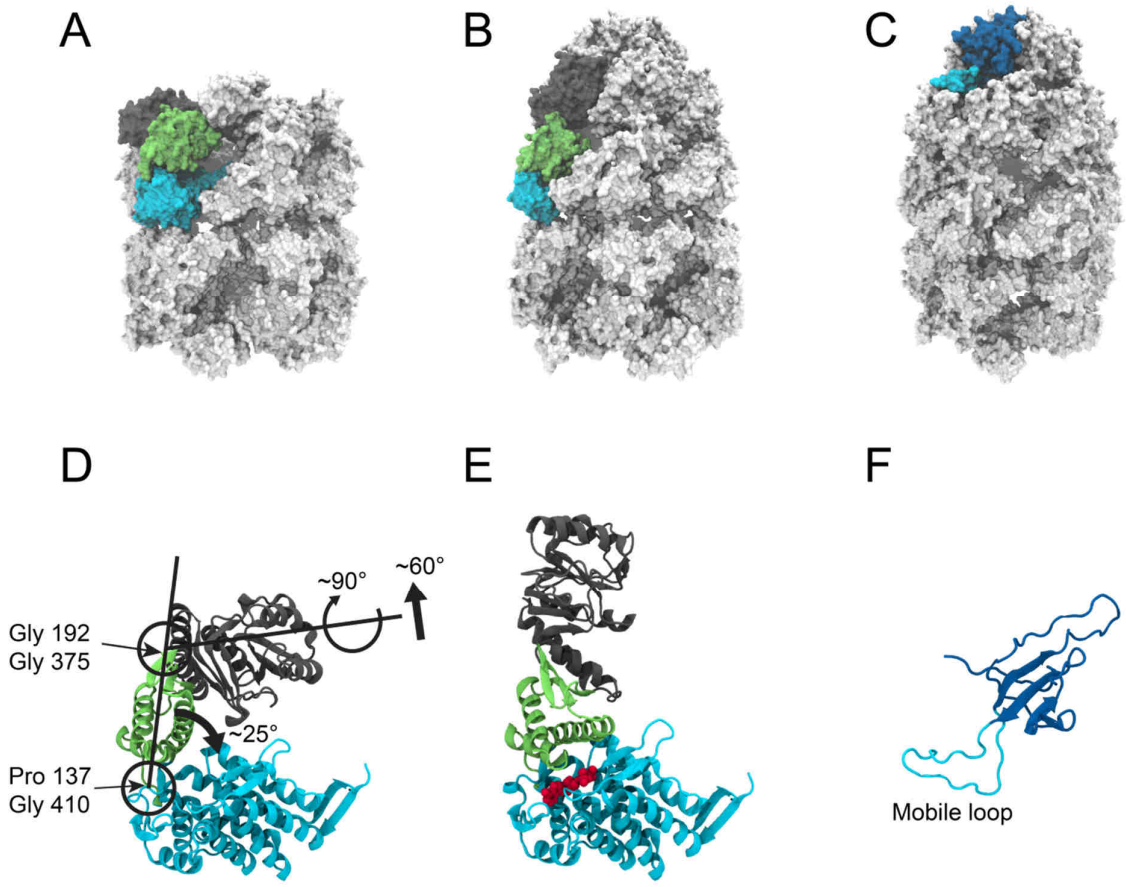

Figure 2.9 Structural properties of GroEL and GroES

(A) Domain architecture of one GroEL subunit in the apo-state (PDB 1XCK (Bartolucci et al., 2005)). Equatorial domain is colored in cyan, intermediate domain in green and apical domain in dark grey. All remaining GroEL subunits are colored in light grey **(B)** Domain architecture of one GroEL subunit in the GroES/ADP bound GroEL complex (PDB 1AON (Xu et al., 1997)), colored as in (A). **(C)** Illustration of one GroES subunit in the GroES/ADP bound GroEL complex (PDB 1AON). The GroES mobile loop is colored cyan and the remainder of the same GroES subunit in blue. **(D)** Cartoon model of one GroEL subunit in the apo-state, illustrating domain movements upon nucleotide binding, as well as the important pivot points at the rim of the intermediate domain. Coloring as in (A). **(E)** Cartoon model of one GroEL subunit in the GroES/ADP bound GroEL complex as in (B). ADP is colored in red. **(F)** Cartoon model of one GroES subunit in the GroES/ADP bound GroEL complex as in (C). All figures were rendered with VMD. Domain movements adapted from (Xu et al., 1997).

ATP binding initiates a cascade of structural changes in the GroEL cis ring. First, after nucleotide binding, the intermediate domain swings down towards the equatorial domain, pivoting ~25° around Pro137 and Gly410, which form a link to the equatorial domain (Fig. 2.9). This movement locks the nucleotide binding sites and in addition establishes numerous new intra- and inter-domain interactions. Second, the apical domain swings up ~60° relative to the horizontal plane and twists ~90° around the

vertical axis, pivoting around Gly192 and Gly375, which link intermediate and apical domains. This second motion primes the GroEL subunit for an interaction with the GroES mobile loop (Xu et al., 1997) (Fig. 2.9). Binding of GroES occurs simultaneously with a step-wise release of the substrate into the central cavity, ensuring efficient substrate encapsulation (Chen et al., 2013; Lin et al., 2013; Sharma et al., 2008). As the GroEL substrate binding site overlaps with the binding site for GroES (Fenton et al., 1994), the high affinity binding of GroES to the GroEL apical domains, displaces the substrate into the central GroEL cavity, which upon closure provides a folding-permissive environment.

The structure of the enclosed cis-cavity is markedly different from the former trans-cavity (Fig. 2.10). The mentioned helices H and I no longer contribute to the inner cage surface. In general, the structural rearrangements not only change the inner cage wall from an overall hydrophobic to an overall hydrophilic character (Fig. 2.10 A), they also approximately double the cage volume to ~175.000 $Å^3$ (Chen et al., 1994). Most notably, the cis-cavity is highly negatively charged (net charge -42), due to the presence of two ring-like charge clusters. The residues D359, D361 and E363 of all seven subunits constitute one cluster, and the residues E252, D253 and E255 constitute the other cluster (Tang et al., 2006) (Fig. 2.10 B). Most of these residues are conserved amongst GroEL homologues. Interestingly, mutation of the residues 259, 361 and 363 to Lysine (KKK2 mutant) results in impaired chaperonin assisted folding of certain substrates, suggesting that the negative charges play an important role in substrate folding (Chakraborty et al., 2010; Tang et al., 2006).

The enlarged hydrophilic cavity can accommodate polypeptides up to ~60 kDa. The successful folding of a substrate in most cases critically depends on global encapsulation within the cis-cavity. Larger proteins either use the Hsp70 system (Agashe et al., 2004; Calloni et al., 2012; Kerner et al., 2005) or in rare cases undergo GroEL assisted trans-like folding without encapsulation by GroES (Horwich et al., 2007).

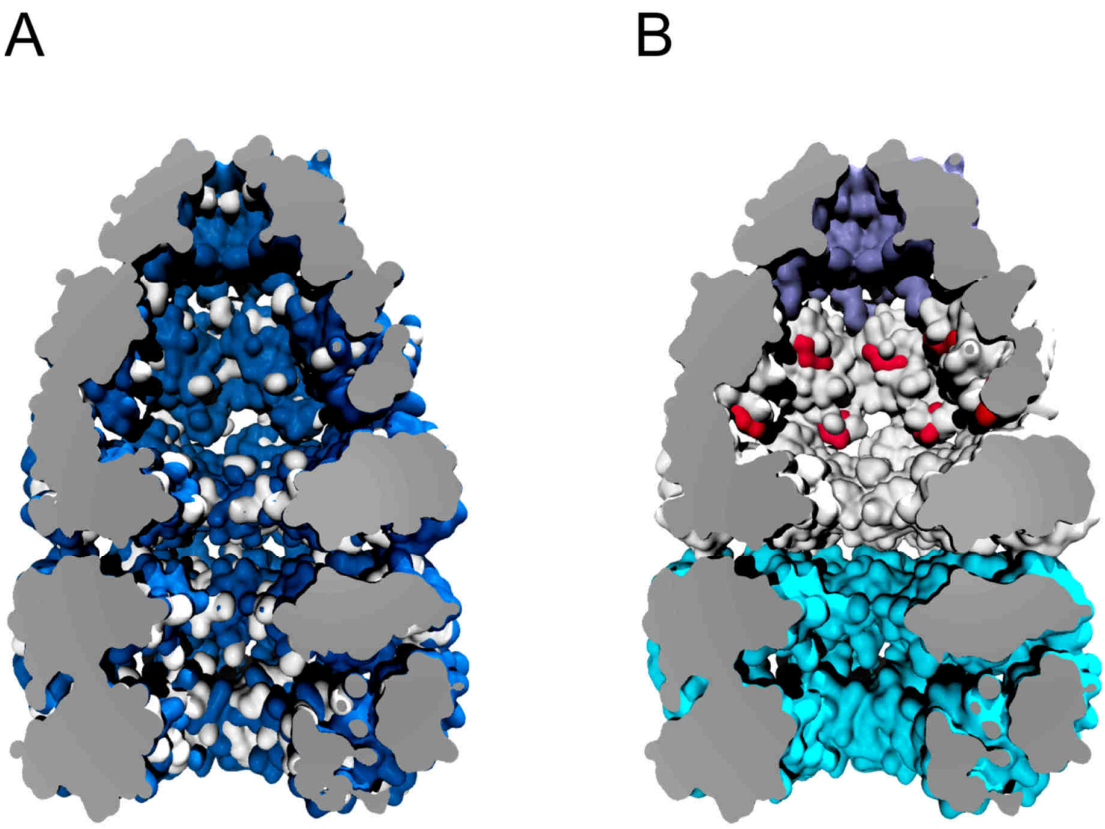

Figure 2.10 Structural features of the GroEL central cavity

(A) Cross-section of the GroEL/ES/ADP complex (PDB 1AON (Xu et al., 1997)) with hydrophobic residues colored in white and hydrophilic residues colored in blue. Illustration of the marked differences between cis- and trans-cavity. **(B)** Cross-section as in (A) with the trans-ring colored in cyan, the cis-ring colored in grey and GroES colored in purple. The important negatively charged clusters (upper E252, D253,E255 and lower D359, D361, E363) are colored in red.

Upon completion of ATP hydrolysis in one GroEL ring, the ADP bound state is conformationally distinct from the ATP bound state. The negative allostery between the two rings is reduced, allowing cooperative binding of ATP to the trans-ring, which in turn induces release of ADP and GroES from the former cis-ring. GroES binding to the trans-ring results in formation of a new cis-ring and completes the functional reaction cycle.

2.8.2 GroEL substrates

The interactome of GroEL was determined by a quantitative proteomics approach, wherein ~250 of the ~2400 cytosolic *E. coli* proteins were identified as interactors of GroEL (Kerner et al., 2005). All identified GroEL interactors were grouped into three different classes. Proteins of class I interact with GroEL but do not require GroEL for folding. In fact, less than 1% of class I proteins fold by assistance

of GroEL. Proteins of class II require chaperone assistance for efficient folding at 37°C but readily fold spontaneously at 25°C. In addition, class II proteins can fold via the Hsp70 system and therefore only partially require GroEL for folding. The ~84 class III proteins are however obligate GroEL substrates, and occupy 75%-80% of the total cellular chaperonin capacity. Upon depletion of GroEL, class III proteins either aggregate or are depleted from the cellular proteome. Notably, in an *in vitro* translation system, lacking other chaperones, GroEL reduced the aggregation of 776 of 800 tested aggregation prone proteins (Niwa et al., 2012), suggesting that substrate specificity in the chaperone network is mainly mediated by chaperones upstream of GroEL. As class III contains 14 essential proteins, GroEL is pivotal for cell viability. Interestingly, class III is enriched in proteins with a $(\beta\alpha)_8$-TIM barrel fold, a common structural scaffold for diverse enzymatic functions (Richard et al., 2014). Also, class III proteins show lower net-charges as well as larger molecular weights than the average of the cytosolic proteome (Kerner et al., 2005).

Subsequent analysis of GroEL requirement for protein solubility revealed an enrichment of metabolic enzymes amongst obligate GroEL interactors (Fujiwara et al., 2010), suggesting that from an evolutionary perspective, GroEL might have buffered structurally destabilizing mutations in enzymes, with on the other hand improved enzymatic activities, by either preventing aggregation or promoting folding of these enzymes.

2.8.2.1 DM-MBP - a GroEL model substrate

The *E. coli* maltose binding protein (MBP) is a monomeric, ~41 kDa periplasmic protein. Upon deletion of its targeting sequence MBP folds robustly and rapidly in the cytosol. MBP is often used as a fusion protein to either mediate solubility of its fusion partners or as an affinity-tag for amylose affinity purification (Raran-Kurussi and Waugh, 2012). MBP consists of βαβ-secondary structure elements, forming two globular domains that are discontinuous in sequence, with the maltose binding site located in a cleft between the two domains (Spurlino et al., 1991). MBP contains 8 Trp residues that are spaced throughout the sequence. The intrinsic Trp fluorescence of MBP is ~5 fold reduced upon unfolding. The fluorescence recovery upon renaturation can be used as a convenient readout for folding kinetics. A number of destabilizing mutants of MBP were described that show slowed folding kinetics (Chun et al., 1993; Wang et al., 1998). Most notably a double mutant (V8G, Y283D) was shown to strongly interact with GroEL and to show accelerated folding in presence of GroEL/ES/ATP (Tang et al., 2006). The two mutations are located in close proximity in the N-domain. Native contact formation in the N-domain was shown to be rate-limiting for MBP folding (Chun et al., 1993). DM-MBP has been frequently used as a GroEL model substrate (Chakraborty et al., 2010; Sharma et al., 2008; Tang et al., 2006, 2008). Upon dilution from denaturing conditions, DM-MBP was shown to form a collapsed kinetically trapped

intermediate state within only milliseconds (Sharma et al., 2008). This collapsed state has a high structural flexibility and shows little to no secondary structure formation (Chakraborty et al., 2010). Binding to GroEL leads to local structural expansion. ATP binding to GroEL induces a transient further expansion of some sequence elements upon addition of ATP, while some moderately hydrophobic segments show increased mobility. Upon encapsulation inside the GroEL central cavity, DM-MBP adopts a compact conformation and can readily fold to the native state (Sharma et al., 2008).

2.8.2.2 DapA – a class III TIM-barrel GroEL substrate

DapA is an essential *E. coli* protein that catalyses the condensation of L-aspartate-β-smialdehyde and pyruvate to dihydrodipicolinic acid, a metabolite for lysine and peptidoglycane biosynthesis. DapA is a tetrameric enzyme, consisting of 31.2 kDa subunits. Each monomer consists of an N-terminal $(\beta\alpha)_8$-TIM barrel domain with the central active site, and an α-helical C-terminal extension, which contributes to the tetrameric interfaces (Dobson et al., 2005). DapA was initially identified as an obligate GroEL substrate when a deficiency in cell wall biosynthesis was observed upon depletion of GroE (McLennan and Masters, 1998). DapA was later confirmed by proteomics to be an obligate class III GroEL substrate. In the same study it was also shown that refolding of His-tagged DapA could be accelerated ~10 fold in presence of GroEL/ES/ATP (Kerner et al., 2005). As DapA folds inside the GroEL cavity, subunit refolding is followed by assembly to dimers and tetramers to reach the final native state and hence full enzymatic activity (Reboul et al., 2012).

2.9 Mechanisms of chaperonin assisted protein folding

The refolding of GroEL substrate proteins can be studied in free solution, i.e. absence of chaperonin (spontaneous folding) or in the presence of GroEL (assisted folding). Effective assisted in-cage folding requires the co-chaperone GroES as well as ATP. It has been reported previously that in presence of the GroEL/ES chaperonin system an apparent acceleration of refolding can be observed for a subset of substrate proteins. Folding rate acceleration has been demonstrated for Ribulose-1,5-bisphosphate carboxylase/-oxygenase (RuBisCO) of *R. rubrum* (Brinker et al., 2001; Chakraborty et al., 2010; Lin et al., 2008; Weaver and Rye, 2014), *E. coli* DapA (Kerner et al., 2005), green fluorescent protein (GFP) (Tang et al., 2008), the knot containing proteins YibK and YbeA (Mallam and Jackson, 2011), and for mutants of MBP (Apetri and Horwich, 2008; Beissinger et al., 1999; Chakraborty et al., 2010; Tang et al., 2006, 2008; Tyagi et al., 2011). The underlying mechanism of chaperonin catalyzed refolding, especially of DM-MBP, has however been a matter of intensive debate during the last decade (Apetri and Horwich, 2008; Brinker et al., 2001; Chakraborty et al., 2010; Horwich et al., 2009; Tang et al.,

2006, 2008; Tyagi et al., 2011; Yang et al., 2013). Three different models have been proposed to explain the apparent acceleration of substrate refolding in the presence of chaperonin.

2.9.1 The passive cage model

The “passive cage” (also “Anfinsen cage”) model suggests that GroEL provides an infinite dilution chamber, in which folding occurs at the same rate as in free solution (Ellis, 1994; Ellis and Hartl, 1996; Horwich et al., 2009). The model implies that GroEL/ES-dependent proteins fold at biologically relevant timescale as long as aggregation is prevented. The observed apparent rate acceleration in presence of chaperonin in this model is explained not by an active folding acceleration in presence of GroEL but rather by a slowed spontaneous folding reaction, due to the presence of transient substrate aggregation as a rate-limiting side reaction. In presence of GroEL however, aggregation cannot occur, due to encapsulation of the substrate inside the GroEL central cavity. Encapsulation results in infinite dilution of the substrate, which during folding exposes hydrophobic residues. Therefore, inside GroEL, an encapsulated substrate molecule is shielded from unproductively interacting with other substrate molecules and cannot form protein aggregates. In the passive cage model the substrate would undergo an Anfinsen type refolding reaction, while the absence of transient formation of aggregates eliminates the rate limiting side reaction. Thus, resulting in an apparent acceleration of protein folding.

2.9.2 The active cage model

In contrast, the “active cage” model states that, besides preventing aggregation, the physical environment of the GroEL cage modulates the folding energy landscape, resulting in an accelerated refolding of certain substrates. This is attributed to an effect of steric confinement that limits the conformational space to be explored during folding (Baumketner et al., 2003; Brinker et al., 2001; Chakraborty et al., 2010; Hayer-Hartl and Minton, 2006; Lucent et al., 2007; Tang et al., 2006, 2008). The confinement of a substrate might favor the formation of long-range contacts in the transition state, which would effectively smoothen the folding energy landscape by avoiding kinetically trapped states. Importantly, the conformational restriction exerted by engineered internal disulfide bonds on a flexible folding intermediate can mimic the effect of spatial confinement in the GroEL/ES cage (Chakraborty et al., 2010).

Furthermore, the net negative charge of the GroEL cis-cavity contributes significantly to the folding acceleration of some substrate proteins (Chakraborty et al., 2010; Tang et al., 2006, 2008). It was suggested from simulations that the charge clusters have an effect on the order of water structure inside the EL cage (England et al., 2008). As the unavailability of solvent molecules for hydrogen bonding thermodynamically favors compaction of hydrophobic residues, such ordering of water molecules might enhance substrate folding by strengthening the hydrophobic effect. An additional factor, that might

contribute to accelerated folding of substrate molecules are the GroEL flexible C-termini. The presence of a mildly hydrophobic interaction surface inside the cis-cavity might help in structural rearrangements of the encapsulated substrate (Tang et al., 2006; Weaver and Rye, 2014).

Taken together, the active cage model implies that cells contain a set of proteins with kinetically frustrated folding pathways that require folding catalysis to reach their native state at biologically relevant speed, i.e. faster than the rate of protein synthesis (~20 aa per second in *E. coli*). It was suggested that steric confinement is especially effective in overcoming energy barriers with a large entropic component (Chakraborty et al., 2010).

2.9.3 The iterative annealing model

Finally, the iterative annealing model posits that GroEL actively modulates the substrate refolding by iterative cycles of substrate annealing to GroEL, forced unfolding and subsequent release with substrate refolding occurring either inside or outside the GroEL cage (Corsepius and Lorimer, 2013; Thirumalai and Lorimer, 2001; Yang et al., 2013). The working principle of GroEL in this model is the active unfolding of kinetically trapped states that can then partition between productive and unproductive folding trajectories. Forced unfolding would therefore play a major role, with substrate encapsulation being a mere byproduct of the unfolding reaction (Yang et al., 2013). Interestingly, substrate folding acceleration was not only observed during the GroEL cycling reaction, but also inside the SREL cavity (Chakraborty et al., 2010; Tang et al., 2006). SREL is a single ring variant of GroEL that is capped by GroES upon ATP binding, which results in formation of a near infinitely stable cis-complex due to the lack of negative allosteric signaling from the trans-ring. In such a case the substrate undergoes only one single round of forced unfolding, suggesting that forced unfolding only to a minor extent (if at all) contributes to the acceleration of substrate folding. Importantly, 100% folding yields were observed during SREL assisted refolding (Chakraborty et al., 2010; Tang et al., 2006), suggesting that substrate proteins are not prone to misfold during encapsulation, making additional rounds of forced unfolding obsolete.

2.9.4 Single molecule fluorescence research on GroEL

Due to its multimeric nature, single molecule experiments on GroEL that require working at low concentrations (e.g. in confocal spectroscopy), usually investigate a labeled substrate molecule, with the unlabeled GroEL cage being present in excess at near physiological concentrations. To circumvent this obstacle and to investigate GroEL itself, GroEL molecules are usually attached to surfaces and monitored as single immobile molecules. Although it was shown that GroEL substrate complexes can be observed on glass surfaces without prior particle immobilization (Yamasaki et al., 1999), many single molecule experiments on GroEL itself are carried out using particle immobilization and total internal

reflection (TIRF) microscopy. Such immobilization based approaches were used to investigate binding and release of GroES (Sameshima et al., 2010; Taguchi et al., 2001; Ueno et al., 2004), as well as online monitoring of single in-cage folding events (Takei et al., 2012; Ueno et al., 2004).

As GroEL is sensitive to chemical modification, in solution approaches are to be favored over immobilization based approaches. Most in solution experiments on GroEL were carried out measuring single molecule FRET on labeled substrate molecules. In a pioneering study, the conformational dynamics of DM-MBP along the chaperone pathway was analyzed with a combination of single molecule FRET and ensemble stopped flow approaches (Sharma et al., 2008). DM-MBP transitions from an expanded denatured state to a collapsed Hsp70 bound state, followed by local expansion upon transfer to GroEL. The expansion is furthered upon apical domain movement. Release of DM-MBP into the chaperonin cage was observed upon binding of GroES, followed by conformational compaction and folding to the native state. Conformational heterogeneity of the substrate upon binding to GroEL and conformational effects of GroEL on the substrate during cycling were also observed for mitochondrial rhodanese and the von Hippel-Lindau tumor suppressor protein by smFRET (Hillger et al., 2008). Even in-cage folding rates could be measured by a combination of smFRET and microfluidic mixing (Hofmann et al., 2010). Not surprisingly, rhodanese in this study globally folded at the same rate inside GroEL as in bulk solution, as rhodanese, a comparatively small substrate, was already shown to benefit from encapsulation, only when the size of the chaperonin cage was reduced by a mutagenic extension of the C-terminus (Tang et al., 2006). Interestingly however, it was shown that the two domains of rhodanese experience different effects upon encapsulation inside the chaperonin cage, supporting the view that the effect of encapsulation strongly depends on the nature of the encapsulated substrate. Later, smFRET was used to determine the structural flexibility of the DM-MBP folding intermediate that can be stabilized in presence of 0.5 M GuHCl (Chakraborty et al., 2010). The structural flexibility of folding intermediate states appears to be a hallmark of GroEL substrates that are accelerated in folding upon encapsulation inside the chaperonin cage.

In a PET-FCS approach, in which labeled GroEL switches between a fluorescent conformer in the T-state and a quenched conformer in the R-state, single molecule experiments could demonstrate the importance of tuned ADP release rates from GroEL for efficient substrate binding to the T-state trans-ring (Frank et al., 2010).

Another, yet to be explored option is to investigate single GroEL complexes in solution without prior immobilization using an anti-brownian electrokinetic (ABEL) trap, previously used to determine nucleotide stoichiometries in TRiC/CCT complexes (Jiang et al., 2011).

2.10 Aim of this study

GroEL is the most widely investigated member of the chaperonin family of molecular chaperones. The observation of accelerated substrate folding in the presence of GroEL and its co-factor GroES has sparked a long-standing debate as to the mechanism by which GroEL assists the folding of a subset of its substrate molecules. Three different models were proposed to explain the apparent acceleration of folding rates in the presence of chaperonin. The passive- or Anfinsen-cage model (Horwich et al., 2009), the active cage model (Chakraborty et al., 2010), and the iterative annealing model (Thirumalai and Lorimer, 2001). The aim of this study was to use the well-studied GroEL model substrate DM-MBP and a natural TIM-barrel substrate DapA to thoroughly investigate the validity of all three models and to investigate the influence of the GroEL central cavity on the encapsulated folding substrate.

3 Materials and Methods

3.1 Materials

3.1.1 Chemicals

Unless indicated otherwise, chemicals and reagents used were of pro analysis (p.A.), ACS quality or comparable assay grade, and were mostly purchased from Sigma-Aldrich (St. Louis, USA).

Table 3.1 Chemicals

Supplier	Chemical/Reagent
Atto-Tec GmbH (Siegen, Germany)	Atto 647N maleimide
	Atto 532 maleimide
	Atto 655 maleimide
BD (Franklin Lakes, USA)	Bacto™ Tryptone
	Bacto™ Yeast Extract
Biomol GmbH (Hamburg, Germany)	2-[4-(2-hydroxythyl)piperazin-1-yl]ethanesulfonic acid (Hepes)
BioRad (Hercules, USA)	Bradford reagent
Carl Roth (Karlsruhe, Germany)	Ampicillin sodium salt
	Potassium chloride
Dyomics GmbH (Jena, Germany)	Dy 530 maleimide

Hoffmann-La Roche (Basel, Switzerland)	Adenosine-5'-triphosphate disodium salt (ATP)
	cOmplete Protease Inhibitor Cocktail Tablets
Life technologies (Carlsbad, USA)	Alexa 647 maleimide
Merck Millipore (Billerica, USA)	Guanidine hydrochloride, ≥99%
	Magnesium chloride hexahydrate ($MgCl_2$)
	Sodium hydroxide pellets (NaOH)
Metabion Int. AG (Martinsried, Germany)	Oligonucleotides
MBIP Microchemistry Core Facility	L-aspartate-beta-semialdehyde
New England Biolabs (Ipswich, USA)	Amylose resin, restriction enzymes
Serva Electrophoresis GmbH (Heidelberg, Germany)	Acrylamide/Bis Solution, 37.5:1 (30% w/v)
	Dodecylsulfate sodium salt in pellets
	Serva Blue R
Sigma-Aldrich (St. Louis, USA)	2-Amino-2-hydroxymethyl-propane-1,3-diol (Tris)
	2-mercaptoethanol
	3',3'',5',5''-Tetrabromophenolsulfonephthalein (Bromphenol Blue)
	4,4'-Dianilino-1,1'-binaphthyl-5,5'-disulfonic acid dipotassium salt (Bis-ANS)
	Ammonium persulfate (APS)
	α-Ketopropionic acid (Pyruvic acid)
	β-Nicotinamide adenine dinucleotide, reduced dipotassium salt
	D-(+)-Maltose Monohydrate
	Guanidine Hydrochloride Solution, 8M in H_2O
	Magnesium acetate tetrahydrate

	N,N,N',N'-Tetramethylethylenediamine (TEMED)
	o-Aminobenzaldehyde (ABA)
	Phosphoenolpyruvic acid monopotassium salt
	Urea ReagentPlus®, ≥99.5%, pellets
Stratagene (Cedar Creek, USA)	QuickChange mutagenesis kit
Thermo Fisher Scientific (Waltham, USA)	Dithiothreitol (DTT)
	Isopropyl β-D-1-thiogalactopyranoside (IPTG)
VWR (Radnor, USA)	Hydrochloric acid, 37%
	Sodium chloride (NaCl)
	Methanol

3.1.2 Proteins, enzymes and kits

Table 3.2 Proteins, enzymes and kits

Supplier	Material
New England Biolabs (Ipswich, USA)	DpnI restriction enzyme
Promega (Madison, USA)	MiniPrep DNA plasmid purification kit
	Pfu DNA Polymerase
Sigma Aldrich (St. Louis, USA)	Apyrase from potatoes
	Benzonase Nuclease
	Bovine Serum Albumin (BSA)
	Lysozyme
	Pyruvate kinase/lactate dehydrogenase mix

3.1.3 Instruments

Table 3.3 Instruments

Supplier	Material
Beckman Coulter (Pasadena, USA)	Benchtop centrifuge GS-6
	High capacity centrifuge J6-MI
	Ultracentrifuge Optima L-90K
	Ultracentrifuge rotor type 45 Ti
Biometra (Göttingen, Germany)	PCR thermocycler
Drummond Scientific (Broomall, USA)	Pipet-Aid pipet controller
Eppendorf (Hamburg, Germany)	Benchtop centrifuges 5415D and 5417R
	Research plus pipette (2.5 µL, 10 µL, 20 µL, 100 µL, 200 µL, 1 mL)
	Thermomixer comfort
Fujifilm (Tokio, Japan)	FLA-2000 Fluorescence Imager
GE Healthcare (München, Germany)	Äkta Explorer, Äkta Purifier, chromatography columns (S30Q, MonoQ, Sephacryl S200, Heparin)
Horiba Yvon	FluoroLog 3 spectrofluorometer
Ibidi (Martinsried, Germany)	µ-slide 8 well chambered microscope coverslip
Jasco (Gross-Umstadt, Germany)	J-715 Spectropolarimeter
	V-560 Spectrophotometer
Mettler-Toledo (Greifensee, Switzerland)	Balances AG285 and PB602
Milipore (Bedford, USA)	Amicon centrifuge filter units, steritop filter units

Misonix (Farmingdale, USA)	Sonicator 3000
New Brunswick Scientific	Innova 44 incubator
Olympus (Tokio, Japan)	IX71 microscope body, microscope objective (60X W, NA 1.2)
PicoQuant (Berlin, Germany)	MicroTime 200 time resolved, confocal fluorescence microscope.
Carl Roth (Karlsruhe, Germany)	ZelluTrans dialysis membrane
Scientific Industries	Vortex Mixer Genie 2
Thermo Fisher Scientific (Waltham, USA)	NanoDrop 1000
WTW (Weilheim, Germany)	pH meter

3.1.4 Buffers and media

All buffers used for protein purification, storage, or for in-vitro experiments were filtered using SteriTop filter units (0.2 µm) and sonicated. The buffers used for protein purification are indicated in the respective section.

Table 3.4 Buffers nomenclature for experimental results

Buffer name	Composition	Mainly used for
MBP refolding buffer	20 mM Tris/HCl pH 7.5, 200 mM KCl, 5 mM $Mg(C_2H_3O_2)_2$	Experiments with DM-MBP alone and with GroEL/ES
MBP LS refolding buffer	20 mM Tris/HCl pH 7.5, 20 mM KCl, 5 mM $Mg(C_2H_3O_2)_2$	Experiments with DM-MBP alone and with GroEL/ES under low salt condition
SREL buffer	50 mM Hepes/NaOH pH 7.5, 20 mM KCl, 10 mM $MgCl_2$	Experiments with SREL

DapA refolding buffer	20 mM Tris/HCl pH 7.5, 100 mM KCl, 10 mM $MgCl_2$, 10 mM pyruvate	Experiments with DapA alone and with GroEL/ES

Lysogeny broth medium (Bertani, 1951)

10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl. The pH was found to be ~7.0 and not adjusted with NaOH to avoid pH variation during cultivation.

Comassie gel staining solution

40% ethanol, 8% acetic acid, 0.1% (w/v) Serva Coomassie Blue R-250

Comassie gel destaining solution

10% ethanol, 7% acetic acid

SDS-PAGE running buffer

15 g/L Tris, 72 g/L Glycine, 5 g/L SDS

Sample buffer SDS PAGE

62.5 mM Tris-HCl pH 6.8, 2.5% SDS, 0.002% Bromphenol Blue, 5% 2-mercaptoethanol, 10% glycerol

3.1.5 Strains and plasmids

The *E. coli* strains DH5α and BL21 (DE3) Gold (Stratagene) were used for cloning and protein expression, respectively.

The *E. coli* genes *groEL* and *groES* were cloned into pET11a using BamH1 and NdeI restriction sites. The SREL (R452E, E461A, S463A, V464A) (Weissman et al., 1995) and EL(KKK2) (D359K, D361K, E363K) (Tang et al., 2006) variants of GroEL were constructed by site directed mutagenesis using QuickChange (Stratagene).

The gene encoding MBP (Wang et al., 1998) was cloned into a pCH vector backbone using NdeI and NheI restriction sites. The double mutant (V8G, Y283D) and all cysteine variants (MBP A312C, DM-MBP A312C and DM-MBP D30C A312C) were constructed by sited directed mutagenesis using QuickChange (Stratagene).

The gene *dapA* from *E.coli* was cloned into the vector pUC19 (pT7-DapA; pT7-EcNanA). The DapA mutant, DapA (293C) (in which the three surface cysteines were replaced with serines, C20S, C141S, C218S, and an additional cysteine added to the C-terminus (293C), was generated by

QuickChange mutagenesis (Stratagene) of the wild-type *dapA* gene. The authenticity of each construct was confirmed by DNA sequencing.

3.2 Molecular biology methods

3.2.1 Transformation of competent *E. coli* cells

30 ng DNA were added to 50 µL of a suspension of competent *E. coli* cells. The suspension was incubated on ice for 10 min. A 90 sec heat shock at 42°C was used for efficient transformation, followed by addition of 950 µL LB medium and incubation at 37°C for 1 h with constant shaking. The cells were subsequently pelleted by centrifugation for 1 min at 25°C. The pellet was resuspended in 50 µL LB medium and plated on pre-heated, antibiotic containing LB plates. LB plates were incubated at 37°C over night.

3.2.2 Site directed mutagenesis

Single amino acid mutations in proteins were achieved by site directed mutagenesis on DNA plasmids. For multiple mutations, iterative steps of site directed mutagenesis were performed. Primers were designed to carry the desired mutation and to self complementary align with plasmid DNA. PCR of the full length plasmid DNA was achieved using Pfu DNA polymerase in an automated PCR thermocycler.

PCR protocol:

Temperature (°C)	Time (min)	Cycles
95	2	1
95	0.5	
55	1	18
68	10	
68	15	1
4	∞	1

Subsequent to PCR, 0.5 µL Dpn I restriction enzyme was added. The reaction mix was incubated for 1 h at 37°C to ensure efficient digestion of methylated template DNA. The remaining plasmid DNA was transformed into competent *E. coli* (DH5α) cells.

3.3 Protein biochemistry methods

3.3.1 Purification of GroEL

Buffer A (200 mM Tris/HCl pH 7.5, 10 mM DTT)

Buffer B (20 mM Tris/HCl pH 7.5, 20 mM NaCl, 1 mM EDTA, 1 mM DTT)

Buffer C (20 mM Tris/HCl pH 7.5, 1 M NaCl, 1 mM EDTA, 1 mM DTT)

Buffer D (20 mM Mops/NaOH pH 7.2, 100 mM NaCl, 20% methanol, 1 mM DTT)

Buffer E (20 mM Mops-NaOH pH 7.2, 100 mM NaCl, 1 mM DTT)

GroEL and variants were purified from *E. coli* BL21 gold strain as described previously (Hayer-Hartl et al., 1994) with some modifications. Cells were grown at 37°C to an OD of 0.45. Protein expression was induced upon addition of 1 mM IPTG (Roth). The cells were furthermore grown for 4 h at 37°C, harvested by centrifugation (Beckman J6-MI, 3200 x g, 45 min, 4°C) and subsequently frozen in liquid nitrogen as a suspension in buffer A.

Thawed cells were lysed by incubation for 1 h at 4°C in the presence of complete protease inhibitor (Roche), 1 mg/mL of 42.35 U/mg Lysozyme (Sigma) and 10 U/mL Benzonase (Novagen) and subsequent sonication with a tip sonicator (Misonix Sonicator 3000, power output 7 in pulse mode, 10 30 sec pulses interrupted by 90 sec pause), while the suspension was cooled on ice to prevent protein precipitation.

After removal of cell debris and membranes by ultracentrifugation (Beckman L-90K, Ti45 rotor, 200 000 x g, 45 min, 4°C), the supernatant was fractionated by chromatography on Source 30Q (Pharmacia Biotech) equilibrated in buffer B with a gradient to buffer C.

Fractions containing GroEL were pooled, adjusted to buffer B, and next applied to a MonoQ HR 16/10 column (Pharmacia Biotech), equilibrated in buffer B with a gradient to buffer C.

GroEL containing fractions were again pooled, adjusted to buffer B and applied to a Heparin column (GE Healthcare), equilibrated in buffer B with a gradient to buffer C.

GroEL containing fractions were subsequently pooled and subjected to Sephacryl S300 HiPrep 26/60 (Amersham Biosciences) gel filtration chromatography in buffer D.

Fractions containing pure GroEL were adjusted to buffer E, concentrated at 4°C using Vivaspin (MWCO 30 kDa; GE Healthcare) and supplemented with 5 % glycerol, flash-frozen in liquid nitrogen and stored at -80°C. After every chromatography step the purity of proteins was controlled by SDS-PAGE. Finally all GroEL purifications were controlled for ATPase activity in presence and absence of GroES (Poso et al., 2004), rhodanese aggregation prevention (Weber and Hayer-Hartl, 2000) and DM-MBP refolding (Tang et al., 2006).

3.3.2 Purification of GroES

Buffer A (200 mM Tris/HCl pH 7.5, 10 mM DTT)
Buffer B (20 mM Tris/HCl pH 7.5, 20 mM NaCl, 1 mM EDTA)
Buffer C (20 mM Tris/HCl pH 7.5, 1 M NaCl, 1 mM EDTA)
Buffer D (20 mM Imidazol pH 5.8, 10 mM NaCl)
Buffer E (20 mM Imidazol pH 5.8, 1 mM NaCl)
Buffer F (20 mM Tris-HCl pH 7.5, 100 mM NaCl, 1 mM EDTA)

GroES was purified from *E. coli* BL21 gold strain. Cells were grown at 37°C to an OD of 0.45. Protein expression was induced upon addition of 1 mM IPTG (Roth). The cells were furthermore grown for 4 h at 37°C, harvested by centrifugation (Beckman J6-MI, 3200 x g, 45 min, 4°C) and subsequently frozen in liquid nitrogen as a suspension in buffer A.

Thawed cells were lysed by incubation for 1 h at 4°C in the presence of complete protease inhibitor (Roche), 1 mg/mL of 42.35 U/mg Lysozyme (Sigma) and 10 U/mL Benzonase (Novagen) and subsequent sonication with a tip sonicator (Misonix Sonicator 3000, power output 7 in pulse mode, 10 30 sec pulses interrupted by 90 sec pause), while the suspension was cooled on ice to prevent protein precipitation.

After removal of cell debris and membranes by ultracentrifugation (Beckman L-90K, Ti45 rotor, 200 000 x g, 45 min, 4°C), the supernatant was fractionated by chromatography on DEAE sepharose (GE Healthcare), equilibrated in buffer B with a gradient to buffer C.

Fractions containing GroES were pooled and adjusted to buffer D, and next applied to a MonoQ HR 16/10 column (GE Healthcare), equilibrated in buffer D with a gradient to buffer E.

GroES containing fractions were again pooled and subjected to Superdex 200 HiPrep 26/60 (GE Healthcare) gel filtration chromatography in buffer F.

Fractions containing pure GroES were concentrated at 4^{o}C using Vivaspin (MWCO 10 kDa; GE Healthcare) and supplemented with 5 % glycerol, flash-frozen in liquid nitrogen and stored at -80^{o}C. After every chromatography step the purity of proteins was controlled by SDS-PAGE. Finally all GroES purifications were controlled for ATPase activity originating from impurities, efficient inhibition of GroEL ATPase activity (Poso et al., 2004) and DM-MBP refolding (Tang et al., 2006).

3.3.3 Purification of WT-MBP, DM-MBP and cysteine variants

Buffer A (200 mM Tris/HCl pH 7.5, 10 mM DTT)
Buffer B (20 mM Tris-HCl, 100 mM NaCl, 10 mM DTT)

MBP and mutants of MBP were purified from soluble material after over expression in *E. coli* BL21 gold strain essentially as described previously (Sharma et al., 2008; Tang et al., 2006). Cells were grown at 37°C to an OD of 0.45. Protein expression was induced upon addition of 1 mM IPTG (Roth). The cells were furthermore grown for 4 h at 37°C, harvested by centrifugation (Beckman J6-MI, 3200 x g, 45 min, 4°C) and subsequently frozen in liquid nitrogen as a suspension in buffer A.

Thawed cells were lysed by incubation for 1 h at 4°C in the presence of complete protease inhibitor (Roche), 1 mg/mL of 42.35 U/mg Lysozyme (Sigma) and 10 U/mL Benzonase (Novagen) and subsequent sonication with a tip sonicator (Misonix Sonicator 3000, power output 7 in pulse mode, 10 30 sec pulses interrupted by 90 sec pause), while the suspension was cooled on ice to prevent protein precipitation.

After removal of cell debris and membranes by ultracentrifugation (Beckman L-90K, Ti45 rotor, 200 000 x g, 45 min, 4°C), the supernatant was dialyzed over night against 100 fold excess of buffer B to remove cellular maltose.

Subsequently the supernatant was applied to an Amylose column (NEB) equilibrated in buffer B. After washing with 5 column volumes of buffer B, pure MBP was eluted with buffer B containing 5 mM maltose.

Fractions containing MBP were pooled and subjected to Sephacryl S300 HiPrep 26/60 (Amersham Biosciences) gel filtration chromatography in buffer B, 1 mM DTT.

Fractions containing pure MBP were then concentrated at 4^{o}C using Vivaspin (MWCO 3 kDa; GE Healthcare) and supplemented with 5 % glycerol, flash-frozen in liquid nitrogen and stored at -80^{o}C. After every chromatography step the protein purity was controlled by SDS-PAGE.

3.3.4 Purification of DapA

Buffer A (200 mM Tris/HCl pH 7.5, 10 mM DTT)
Buffer B (20 mM Tris-HCl pH 7.5, 100 mM NaCl, 10 mM DTT)
Buffer C (20 mM Tris-HCl pH 7.5, 1 M NaCl, 10 mM DTT)
Buffer D (20 mM Tris-HCl pH 7.5, 2 M $(NH_4)_2SO_4$, 10 mM DTT)
Buffer E (20 mM Tris-HCl pH 7.5, 10 mM DTT)

DapA and DapA(293C) were purified from *E. coli* BL21 gold strain as described previously (Laber et al., 1992) with modifications. Cells were grown at 30°C to an OD of 0.45. Protein expression was induced upon addition of 1 mM IPTG (Roth). The cells were furthermore grown for 4 h at 30°C, harvested by centrifugation (Beckman J6-MI, 3200 x g, 45 min, 4°C) and subsequently frozen in liquid nitrogen as a suspension in buffer A.

Thawed cells were lysed by incubation for 1 h at 4°C in the presence of complete protease inhibitor (Roche), 1 mg/mL of 42.35 U/mg Lysozyme (Sigma) and 10 U/mL Benzonase (Novagen) and subsequent sonication with a tip sonicator (Misonix Sonicator 3000, power output 7 in pulse mode, 10 30 sec pulses interrupted by 90 sec pause), while the suspension was cooled on ice to prevent protein precipitation.

After removal of cell debris and membranes by ultracentrifugation (Beckman L-90K, Ti45 rotor, 200 000 x g, 45 min, 4°C), the supernatant was fractionated by chromatography on Source 30Q (Pharmacia Biotech), equilibrated in buffer B with a gradient to buffer C.

Fractions containing DapA were pooled and adjusted to buffer D and any precipitated protein was removed by centrifugation (3200 x g, 30 min, 4°C). The supernatant was next applied to a phenyl-Sepharose CL-4B column (GE Healthcare), equilibrated in buffer D and proteins eluted with a gradient to buffer E.

Fractions containing DapA were pooled, dialyzed against buffer B and applied to a MonoQ HR 16/10 column (Pharmacia Biotech), equilibrated in buffer B and proteins were eluted with a gradient to buffer C.

DapA containing fractions were pooled and subjected to Sephacryl S300 HiPrep 26/60 (Amersham Biosciences) gel filtration chromatography in buffer B. Fractions containing DapA were concentrated at 4^{o}C using Vivaspin (MWCO 10 kDa; GE Healthcare) and supplemented with 5 % glycerol, flash-frozen in liquid nitrogen and stored at -80^{o}C. After every chromatography step the purity of proteins was controlled by SDS-PAGE and the activity was assessed by enzymatic assay.

3.3.5 MBP maleimide labeling

Buffer A (20 mM Tris-HCl, 100 mM NaCl, 10 mM DTT)
Buffer B (20 mM Tris-HCl, 100 mM NaCl)

For single molecule experiments, genetically engineered cysteine mutants of MBP and DM-MBP have been constructed, taking advantage of the lack of Cys residues in the MBP sequence. Newly introduced, surface exposed cysteine residues were modified with fluorescent probes (AttoTec) using maleimide chemistry.

The purified protein in buffer A was first buffer exchanged on a NAP5 column (Amersham Biosciences) equilibrated in buffer B, immediately mixed with a 1.2 molar excess of dye molecules and incubated for 30 min at 20°C in a dark environment. In case of double labeling of DM30/312, the protein was incubated with a 3 fold excess of a 1:1 mixture of donor and acceptor dye. The labeling reaction was quenched by addition of 10 mM DTT. Free dye was removed by binding the labeled protein to an Amylose column (NEB) followed by extensive washing with buffer A. Efficient binding to amylose

resin confirmed the native structure of the labeled protein. Labeled DM-MBP was then eluted using buffer A including 5 mM maltose. Subsequently the buffer was exchanged on a NAP5 column (Amersham Biosciences) to buffer A without maltose. The protein was concentrated using Vivaspin concentrators (MWCO 3 kDa, GE Healthcare). For long term storage 5 % glycerol was added, the protein was flash frozen in liquid nitrogen and stored at -80°C.

The degree of labeling (DOL) was controlled with absorption spectroscopy (MBP: ε_{280} = 64860 M^{-1} cm^{-1}; Atto532: ε_{max} = 115000 M^{-1} cm^{-1} cf_{280} = 0.11; Atto647N: ε_{max} = 150000 M^{-1} cm^{-1} cf_{280} = 0.05; Atto655: ε_{max} = 125000 M^{-1} cm^{-1} cf_{280} = 0.08), using the following equation,

$$\mathrm{DOL} = \frac{A_{dye} \times \varepsilon_{dye}}{\left(-\left(A_{dye} \times cf_{280}\right) + A_{280}\right) \times \varepsilon_{protein}}$$

and found to be >90 %. The absence of free dye in the sample was confirmed by fluorescence correlation spectroscopy (FCS).

3.3.6 DapA maleimide labeling

Buffer A (20 mM Tris-HCl pH 7.5, 100 mM NaCl, 10 mM DTT)
Buffer B (20 mM Tris-HCl pH 7.5, 100 mM NaCl)

For single molecule experiments, a genetically engineered mutant of DapA was constructed in which the three surface exposed cysteines were replaced by serine (C20S, C141S, C218S) and an additional cysteine was added to the C-terminus of the protein in position 293 (DapA(293C)). DapA(293C) was labeled with either Alexa647 (Invitrogen) or Dy530 (Dyomics) using maleimide chemistry.

The purified protein in buffer A, was first buffer exchanged on a NAP5 column (Amersham Biosciences), equilibrated in buffer B and immediately mixed with a 1.2 molar excess of dye molecules and incubated for 30 min at 20°C in a dark environment. The labeling reaction was quenched by addition of 10 mM DTT. Free dye was removed using a NAP5 column equilibrated in buffer A and concentrated using Vivaspin (MWCO 10 kDa, GE Healthcare). The degree of labeling (DOL) was controlled with absorption spectroscopy (DapA: ε_{280} = 12950 M^{-1} cm^{-1}; Alexa647: ε_{max} = 265000 M^{-1} cm^{-1} cf_{280} = 0.023; Dy530: ε_{max} = 100000 M^{-1} cm^{-1} cf_{280} = 0.15) using the following equation,

$$\mathrm{DOL} = \frac{A_{dye} \times \varepsilon_{dye}}{\left(-\left(A_{dye} \times cf_{280}\right) + A_{280}\right) \times \varepsilon_{protein}}$$

and found to be >90 %. The absence of free dye in the sample was confirmed by fluorescence correlation spectroscopy (FCS).

3.3.7 ATPase assay

The ATPase activity of 0.2 µM GroEL or EL(KKK2) or 0.1 µM SREL or SR(KKK2) was measured in MBP refolding buffer (20 mM Tris/HCl pH 7.5, 200 mM KCl, 5 mM $Mg(C_2H_3O_2)_2$) at 20°C in absence or presence of 0.4 µM GroES or 0.4 µM GroES with increasing concentration of denatured DM-MBP (diluted 200-fold from 6 M GuHCl). Control reactions received equivalent amounts of GuHCl (30 mM final). The hydrolysis of ATP to ADP was followed photometrically using a NADH coupled enzymatic assay (2 mM phosphoenolpyruvate, 30/20 U mL^{-1} pyruvate kinase/lactate dehydrogenase, 0.5 mM NADH , 1 mM ATP) at 20°C in a spectrophotometer (Jasco) essentially as described previously (Poso et al., 2004).

3.3.8 Analysis of protein encapsulation

Encapsulation experiments were performed in SREL buffer (50 mM Hepes/NaOH pH 7.5, 20 mM KCl, 10 mM $MgCl_2$). DM-MBP(Atto655) was unfolded in 10 M urea/10 mM DTT for 1 h at 50°C. The denatured protein was diluted 200-fold (final protein concentration 30 nM) into refolding buffer containing 1 µM SREL or SR(KKK2). The reaction was incubated for 5 min at room temperature. Refolding was started by addition of 3 µM GroES and 1 mM ATP at 20°C. The reaction was separated on a Superdex 200 PC3.2/30 gel filtration column (Amersham Biosciences), equilibrated in SREL buffer/50 mM urea/1 mM ATP, either immediately, or after 30 to 60 min incubation at 20°C with or without dissociation of the SREL/ES complex by the addition of 50 mM CDTA/70 mM GuHCl/200 mM KCl. Fractions were collected, analyzed by 15% SDS-PAGE, Coomassie staining and fluorescence imaging (FujiFilm FLA3000), and quantified by densitometry.

3.3.9 Refolding of DapA followed by enzymatic activity

DapA was unfolded in 7.2 M GuHCl. Refolding was induced upon 100- to 200-fold dilution into DapA refolding buffer (spontaneous folding) or DapA refolding buffer containing 2 µM GroEL (assisted folding). GroEL assisted refolding was started by addition of 4 µM GroES and 5 mM ATP. Spontaneous refolding was stopped as indicated in the figure legends by addition of GroEL or GroEL D87K. Assisted folding was stopped by addition of either 50 mM CDTA or apyrase. Enzymatic activity was measured as described previously (Kerner et al., 2005) after 1 h during which productive assembly of native tetramer was allowed to proceed. The photometrically obtained enzymatic activity data was normalized to a native control and fit with a single exponential rate.

3.4 Biophysical methods

3.4.1 Fluorimetric DM-MBP folding rate measurement

Spontaneous and GroEL assisted refolding rates of DM-MBP and variants at a concentration of 100 nM were measured on a Fluorolog F3-22 spectrofluorimeter (Horiba), equipped with a Peltier thermostat, maintaining a constant temperature of 20°C. DM-MBP variants were unfolded in either 6 M GuHCl, 10 mM DTT for 1 h at 20°C or 10 M urea, 10 mM DTT for 1 h at 50°C. Refolding was induced upon 200-fold dilution into refolding buffer for spontaneous refolding. In case of assisted refolding, DM-MBP was diluted 200-fold into refolding buffer, containing either 2 µM GroEL or 1 µM SREL. Refolding was induced upon addition of 4 µM GroES and 5 mM ATP.

Refolding was followed by increase of intrinsic tryptophan fluorescence (excitation: 295 nm, emission: 345 nm), taking advantage of the absence of Trp residues in GroEL and GroES. For double labeled DM30/312, the presence of two fluorophores resulted not only in a notable decrease of spontaneous folding rate but also in strong bleaching of tryptophan fluorescence. However, fluorescence of the donor dye was significantly different between unfolded and native state. Therefore rate measurements of 100 nM double labeled DM30/312 were performed using an excitation wavelength of 532 nm and an emission wavelength of 550 nm. Photobleaching was carefully avoided by limiting the excitation slit width to 2 nm, with the emission slit width being set to 8 nm. Fluorescence signal was collected for an integration time of 100 ms every 30 seconds. Else, excitation light was blocked from the sample with an automated shutter.

3.4.2 Fluorescence correlation spectroscopy

FCS measurements using pulsed interleaved excitation (PIE) (Müller et al., 2005) were performed on a Microtime 200 inverse time-resolved fluorescence microscope (PicoQuant), which was maintained at a constant temperature of 20°C. For excitation of red and green absorbing dyes, picosecond pulsed diode lasers at 640 nm (LDH-PC-640B) and at 530 nm (LDH-P-FA-530) were used, respectively. Each laser had a laser power of 60 µW measured before the major dichroic. The lasers were pulsed with a rate of $26.\overline{6}$ mHz. The excitation light was guided through a water immersion objective (60 × 1.2 NA, Olympus) into the sample cuvette (Ibidi). The emitted fluorescence was separated from excitation light by a dichroic mirror (Z532/635RPC), guided through a pinhole (75 µm) and in case of cross correlation split according to wavelength by a beamsplitter (600 DCXR) onto photon avalanche diodes (SPADs) (PDM series, MPD). The emission light was cleaned up by emission bandpass filters (HQ 690/70 and HQ 580/70, Chromas) in front of the respective detector. Detection was performed using time correlated single photon counting, making it possible to correlate any given photon with the excitation source. In case of auto correlation measurements, after-pulsing artifacts were removed using fluorescence lifetime

filters (Symphotime, PicoQuant) (Enderlein et al., 2005). Recorded fluorescence traces were either auto correlated or cross correlated. The general function to express auto-correlation or dual color cross-correlation of fluorescence fluctuations is:

$$G_{GR}(\tau) = \frac{\langle \delta F_G(t) \cdot \delta F_R(t+\tau) \rangle}{\langle F_G(t) \rangle \cdot \langle F_R(t) \rangle}$$

where δF_G and δF_R denote the fluctuation of the signal of green and red fluorescence at the time points t and $t + \tau$. In case of auto correlation of a single color, G = R.

The amplitude of the correlated data in autocorrelation is inversely proportional to the concentration of particles and in case of cross correlation is directly proportional to the concentration of double labeled particles:

$$\langle C_{RG} \rangle = \frac{G_{cc}(0)}{G_R(0) \cdot G_G(0) \cdot V_{eff}}$$

The average concentration of double labeled particles is directly proportional to the amplitude of the cross correlation function $G_{cc}(0)$ and inversely proportional to the amplitudes of the auto correlation functions for red $G_R(0)$ and green $G_G(0)$ labeled particles. V_{eff} denotes the volume of the focal spot.

3.4.3 FCS based folding rate measurement at 100 pM protein concentration

Refolding kinetics of spontaneous and assisted refolding were measured for 100 pM Alexa647 labeled DapA(293C) as well as 100 pM DM-MBP (double labeled). Spontaneous folding reactions were stopped after different times by addition of 2 µM GroEL. For assisted refolding, the unfolded protein was diluted to 100 pM final concentration into buffer containing 2 µM GroEL. Refolding was initiated by addition of 4 µM GroES and 5 mM ATP and stopped after different times by addition of Apyrase (Sigma). By stopping the folding reaction, not-yet folded protein will be bound by GroEL whereas folded protein remains free in solution. The significant size difference of folded DapA or DM-MBP and protein in complex with GroEL (~830 kDa), results in different diffusion rates, which can be monitored by FCS. FCS measurements were performed within 10 min after stopping the reaction for DapA and within 1 h after stopping for DM-MBP.

The auto-correlation data was fitted with the following one triplet one diffusion equation using the Symphotime software (Picoquant):

$$G(\tau) = \left[1 - T + T \times e^{\left(-\frac{\tau}{\tau_T}\right)}\right] \times \left[\rho \times \left(1 + \frac{\tau}{\tau_D}\right)^{-1} \times \left(1 + \frac{\tau}{\tau_D \times \kappa^2}\right)^{-1/2}\right]$$

The mean diffusion time τ_D of particles through the focal spot is described by the structural parameter $\kappa = z_0/\omega_0$ where z_0 and ω_0 denote the axial and radial dimensions of the confocal volume, respectively. The amplitude of the correlation function is denoted by ρ. The first term is used to compensate for fast dynamics arising from dye photophysics such as triplet blinking with the amplitude T on the timescale τ_T (Widengren et al., 1995). The diffusion coefficients were calculated using the following equation

$$D = \frac{\left(V_{eff} \times \pi^{-3/2} \times \kappa^{-1}\right)^{2/3}}{4 \times \tau_D}$$

by calibrating the confocal volume V_{eff} with Atto655 dye, for which accurate diffusion parameters have been published (Müller et al., 2008). To analyze refolding kinetics the mean diffusion time, reflecting a shift of molecules from GroEL-bound to free, was plotted against the refolding time and fitted with a single exponential rate.

3.4.4 Dual color fluorescence cross correlation spectroscopy (dcFCCS)

Dual color FCCS (dcFCCS) was employed to demonstrate the absence of inter-molecular association during spontaneous refolding of DM-MBP and DapA at 100 pM. DM-MBP(312C) was either labeled with Atto532 or Atto657N as described above. DapA(293C) was labeled with either Alexa647 or Dy530 as described above.

For DM-MBP the labeled proteins were denatured in 6 M GuHCl, 10 mM DTT for 1 h at 20°C. For DapA the labeled proteins were denatured at in 7.2 M GuHCl, 10 mM DTT for 1 h at 20°C. Refolding was induced for by 200 fold dilution into buffer to a final concentration of 50 pM of each labeled species. FCCS was recorded with PIE (Müller et al., 2005) during refolding at 20°C.

As a positive control in case of DM-MBP, 5 pM of DM30/312, double labeled with Atto532 and Atto647N was mixed with 50 pM of each of the single labeled DM-MBP (312C) species, to mimic the presence of a dimeric species and to demonstrate the high sensitivity of this approach. In case of DapA, the two labeled and unfolded molecule populations were mixed 1:1 at a concentration of 100 nM each and allowed to refold and assemble (note that folded DapA will not assemble at 100 pM). The assembled tetramer was then diluted to a final particle concentration of 100 pM for dcFCCS analysis.

For DM-MBP a mixture of 50 pM native DM-MBP(312C)-Atto532 and 50 pM native DM-MBP(312C)-Atto647N was used as a negative control. A 1:1 mixture of Dy530 and Alexa647 dyes, again at 50 pM concentration each, was used as a negative control in case of DapA.

FCCS measurements at 37°C were performed using a temperature controlled sample holder (ibidi). The sample holder was heated so that at constant temperature of 37°C was observed inside the measurement cuvette. Temperature stability was monitored during the time course of the experiment.

3.4.5 Single molecule FRET-based refolding rate measurements

For folding rate measurements at single molecule conditions, a novel smFRET based assay was developed. A double cysteine variant of DM-MBP (DM30/312) was randomly labeled with Atto532 and Atto647N, a commonly used FRET pair with a Förster radius of 52 Å (Sharma et al., 2008). DM30/312 was then unfolded in 6 M GuHCl, 10 mM DTT for 1 h at 20°C.

Spontaneous refolding was induced upon 200 fold dilution into buffer. Refolding at 100 pM was allowed to proceed at 20°C and was stopped at different time points by addition of 2 µM GroEL. All non native conformers of DM-MBP were recognized by GroEL and by stretching on the GroEL apical domains, converted to a low FRET population. All natively folded protein molecules were compact, therefore not recognized by GroEL and showed high FRET efficiencies. Hence GroEL is not only protein of interest but also acts as a sorting machine for different states of single molecules.

For assisted refolding reactions, unfolded DM30/312 was diluted into buffer, containing 2 µM GroEL. Refolding was then started by addition of 4 µM GroES and 5 mM ATP. Refolding was rapidly stopped by addition of apyrase (Sigma). By depletion of ATP, GroEL reverts to the Apo state and as in the spontaneous folding reaction acts as a single molecule sorting machine. After stopping, the refolding mix was transferred from the reaction tube to the measurement cuvette inside the microscope.

Measurements were performed on a MicroTime200 instrument in two color mode with PIE as described under fluorescence correlation spectroscopy. Data was analyzed using a burst intensity approach (Deniz et al., 1999; Zander et al., 1996) in SymphoTime (PicoQuant). A single molecule diffusing through the confocal observation volume results in a burst in fluorescence intensity. A burst was considered as an evaluable event, if it contained more than 25 photons in a 1 ms time window. In addition a threshold of 15 photons following red excitation was used to check for the presence of a functional acceptor fluorophore. FRET efficiencies were calculated from fluorescence intensities of Donor I_D and Acceptor I_A fluorophore by the equation:

$$E = \frac{I_A}{I_A + \gamma I_D}$$

Where $\gamma = (\Phi_A \eta_A / \Phi_D \eta_D)$ denotes a correction factor for differences in quantum yields (Φ) and detection efficiencies (η) (Lee et al., 2005; Sharma et al., 2008) and has been found to be 0.9 for the used FRET pair. Average intensity values of spectral crosstalk and direct excitation of acceptor fluorophores by the green laser have been subtracted.

The resulting FRET efficiency histograms have further been analyzed using Origin (OriginLabs). To quantify the fraction of native molecules, the integrated area of the histogram corresponding to native molecules was divided by the total integrated area of the histogram. This fraction was plotted against refolding time and then fitted with a single exponential function, yielding the rate of folding. Importantly, if the disappearance of the peak area corresponding to GroEL-bound molecules was analyzed, the same result was obtained. For each experiment (i.e. time point) a minimum of 1000 particles was analyzed. All experiments were done at least in triplicate to ensure reproducibility.

In addition to steady state smFRET, we performed smFRET measurements on double labeled DM-MBP during the first minute of GroEL and SREL assisted refolding. In such cases the experiment was conducted multiple times, always considering only particles that were observed during the first minute after initiation of assisted refolding by addition of ATP. The experiment was repeated until FRET was measured for a minimum of 1000 particles. Gaussian fitting allowed extraction of low and high FRET particle fractions.

3.4.6 PET-FCS

PET-FCS (Neuweiler et al., 2009; Sauer and Neuweiler, 2014; Teufel et al., 2011) was used as an approach to assess conformational dynamics in DM-MBP refolding. Atto655 labeled DM-MBP (312C) was unfolded in 6 M GuHCl, 10 mM DTT for 1 h at 20°C. Refolding was induced upon 200-fold dilution of the protein into refolding buffer at 20°C. FCS measurements were started immediately. In order to resolve fast dynamics in the microsecond timescale, fluorescence was recorded on two detectors simultaneously. Cross correlation of the signals allowed removal of detector after pulsing. The correlated data was fitted with the following one exponential one diffusion term equation, with the exponential term describing amplitude F and rate τ_r of photoinduced electron transfer.

$$G(\tau) = \left[1 - F + F \times e^{\left(-\frac{\tau}{\tau_r}\right)}\right] \times \left[\rho \times \left(1 + \frac{\tau}{\tau_D}\right)^{-1} \times \left(1 + \frac{\tau}{\tau_D \times \kappa^2}\right)^{-1/2}\right]$$

The data was fitted either in Origin (OriginLabs) or in SymphoTime (PicoQuant). For relaxation rate extraction only the first or last 30 seconds of a two hours refolding experiment were considered. For folding rate extraction the measurement was subdivided into time-windows of two minutes and extracted values for F were plotted against refolding time. This data was fitted with a single exponential function in Origin (OriginLabs) to give refolding rates.

3.4.7 ANS fluorescence

4,4′-Dianilino-1,1′-binaphthyl-5,5′-disulfonic acid (bis-ANS) stock solution was prepared in DapA refolding buffer/10% methanol (v/v) and adjusted to 2 mM based on the absorption of bis-ANS using the extinction coefficient ε_{385} of 16790 M^{-1} cm^{-1}. Spontaneous refolding at 10°C was performed in DapA refolding buffer at a final concentration of 200 nM. After different times of refolding, bis-ANS was added to a final concentration of 1 µM and fluorescence spectra were recorded immediately. Spectra of the native and unfolded protein were recorded as a reference. The experiments were performed using a FluoroLog 3 spectrofluorometer (Horiba), with the excitation wavelength set to 390 nm (2 nm slitwidth). The emission spectra were recorded from 405-600 nm (2 nm slitwidth) at a rate of 1 nm s^{-1}, using a sampling rate of 0.1 s. A Peltier-thermostat was used to maintain 10°C during the measurement.

Values of the fluorescence maximum at 485 nm were plotted against refolding time. For normalization, the bis-ANS fluorescence measured immediately upon dilution of unfolded protein into bis-ANS containing buffer D was set to 1. The data was fitted with a single exponential function.

3.4.8 CD spectroscopy

Far UV circular dichroism (CD) spectroscopy was performed on a J-715 spectrapolarimeter (Jasco) equipped with Peltier-thermostat at 10°C using 0.1 cm cuvettes.

Spectral acquisition of secondary structure elements of DapA as well as kinetic refolding of DapA was performed at a final DapA concentration of 2 µM in DapA refolding buffer at 10°C. To follow DapA refolding kinetically, recovery of CD signal at 225 nm during renaturation was observed.

4 Results

4.1 Substrate refolding can be strongly accelerated by GroEL/ES

We used the GroEL model substrate DM-MBP to investigate the validity of all three proposed models for how chaperonins promote substrate refolding. DM-MBP has been shown to fold at different rates in the presence and absence of the GroEL/ES chaperonin system (Chakraborty et al., 2010; Sparrer et al., 1996; Tang et al., 2006). DM-MBP carries two destabilizing mutations, V8G and Y283D, which strongly delay the rate limiting folding of the MBP N-domain (Chun et al., 1993) (Fig. 4.1 A). Slow refolding of DM-MBP was attributed to formation of a kinetically trapped folding intermediate (KTI) (Chakraborty et al., 2010). Further, DM-MBP is an ideal subject for spectroscopic investigation, due to its low aggregation propensity, high Trp content (8 Trp residues spaced throughout the sequence) and the lack of intrinsic Cys residues, facilitating site specific introduction of fluorescent probes at engineered Cys residues.

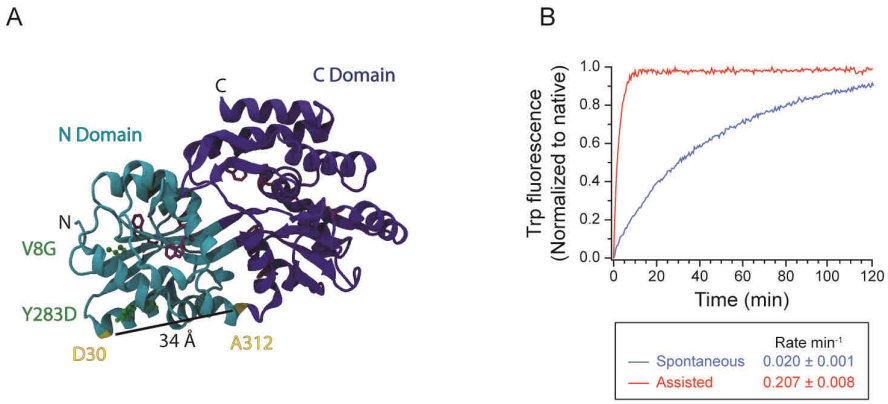

Figure 4.1 Substrate refolding can be accelerated in presence of chaperonin

(A) Structure of DM-MBP (PDB: 1OMP). The C- and N-domains are shown in dark blue and cyan, respectively. The two destabilizing mutations V8G and Y283D are shown in green. Residues D30 and A312 (shown in yellow) were either individually or together mutated to Cys for site specific labeling. Trp residues are shown in

purple. **(B)** Spontaneous (blue) and assisted (red) refolding of DM-MBP analyzed by Trp fluorescence at a final concentration of 100 nM in MBP refolding buffer at 20°C. DM-MBP was denatured in 6 M GuHCl, 10 mM DTT. In the case of assisted refolding 2 µM GroEL and 4 µM GroES were used. Refolding was started by addition of 5 mM ATP. Representative curves are shown. Rates were extracted by single exponential fit and are given as arithmetic mean ± s.d. from at least 3 independent experiments.

As described previously, we confirmed that the spontaneous refolding of chemically denatured DM-MBP in MBP refolding buffer (20 mM Tris/HCl, pH 7.5, 200 mM KCl, 5 mM $Mg(C_2H_3O_2)_2$) occurs as a two-state reaction with a slow rate of 0.02 min^{-1}, $t_{½}$ ~35 min, to full yield at 20°C (Apetri and Horwich, 2008; Tang et al., 2006) and is ~10-fold accelerated in presence of GroEL, GroES and ATP (rate: 0.21 min^{-1}, $t_{½}$ ~3 min) (Chakraborty et al., 2010; Tang et al., 2006) (Fig. 4.1 B).

4.2 Slow spontaneous refolding is not rate limited by transient aggregation

In the light of the ongoing controversy regarding the mechanism of assisted refolding of DM-MBP, our first goal was to decisively distinguish between an active and a passive mechanism of chaperonin function.

4.2.1 Refolding DM-MBP does not oligomerize at low concentration

The passive cage model of chaperonin function (Anfinsen cage model) describes GroEL as an inert cage in which the folding pathway of the substrate is unchanged compared to folding in free solution. The apparent folding rate acceleration in the presence of GroEL is attributed to unproductive reversible aggregation of DM-MBP during spontaneous folding (Apetri and Horwich, 2008; Tyagi et al., 2011).

In order to establish conditions of spontaneous refolding in which transient aggregation is excluded, we resorted to single molecule fluorescence methods. Fluorescence cross correlation spectroscopy (FCCS) can detect the interaction of two spectrally different labeled particles with very high resolution (Bacia and Schwille, 2007). The amplitude of the cross correlation signal correlates with the presence of double-labeled particles, i.e. co-diffusing complexes consisting of two single-labeled species. To test for aggregation of refolding DM-MBP molecules, we labeled two different populations of DM-MBP (312C) with either Atto647N or Atto532 maleimide. First, we mixed the two labeled populations in the native state at a concentration of 50 pM each. As expected, no cross correlation signal was observed (Fig. 4.2 A, purple), as native, soluble DM-MBP is not expected to aggregate or oligomerize. To investigate the oligomeric state of DM-MBP under refolding conditions, the differently labeled DM-MBP molecules in the mixture were unfolded in 6 M GuHCl and refolded by dilution from denaturant to 100 pM final protein concentration. Importantly, no cross correlation signal was observed during

folding (Fig. 4.2 A, blue). Under these conditions, DM-MBP therefore forms a monomeric intermediate state during refolding and does not form transient aggregates. The addition of only 5 pM native double labeled DM30/312 (mimicking the presence of dimeric aggregates) to the refolding mixture of single labeled DM-MBP (312C)-Atto647N and DM-MBP (312C)-Atto532, resulted in an observable cross correlation signal, demonstrating the high sensitivity of this approach (Fig. 4.2 A, black).

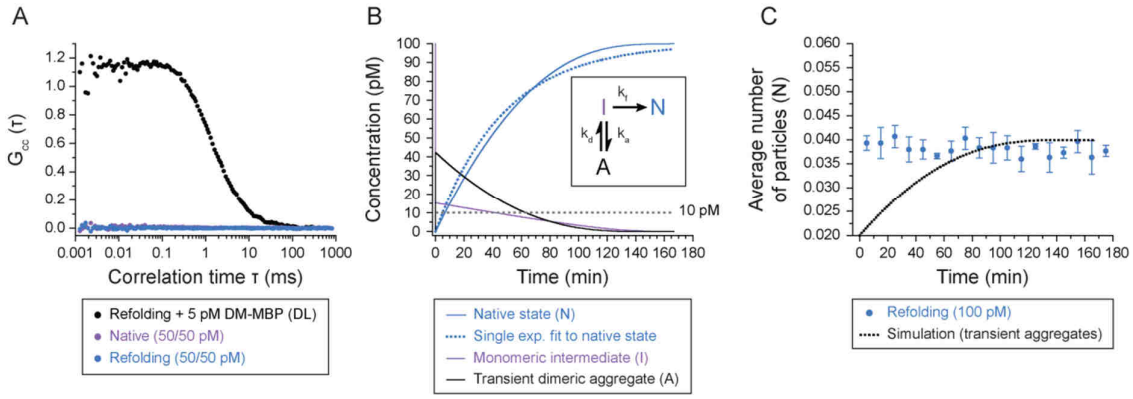

Figure 4.2 DM-MBP does not form transient aggregates during refolding

(A) Absence of dcFCCS signal G_{cc} (τ) during spontaneous refolding of DM-MBP. A 1:1 mixture of DM-MBP (312C) labeled with either Atto532 or Atto647N was denatured in 6 M GuHCl, 10 mM DTT and diluted 200-fold into MBP refolding buffer to a final concentration of 50 pM each. FCCS was recorded within the first 10 minutes of refolding (blue). As a positive control a final concentration of 5 pM double labeled protein DM-MBP (DL) was added to the refolding mix, to simulate the presence of an oligomeric (dimeric) species (black). Native single labeled proteins, again at 50 pM concentration each, were used as a negative control (purple). **(B)** *In silico* kinetic simulation of the Anfinsen cage model including an off-pathway transient dimerization reaction (insert). The concentration of DM-MBP was fixed to 100 pM. Variation of the equilibrium dissociation constant for the formation of dimeric aggregates (A, black) from monomeric intermediates (I, purple) resulted in apparently slower formation of native protein (N, blue). The formation of native protein could not be fitted to a first order reaction (dotted blue line). **(C)** Average number of Atto647N labeled DM-MBP (312C) particles inside the confocal observation volume during the course of spontaneous refolding. DM-MBP (312C) labeled with Atto647N was denatured in 6 M GuHCl, 10 mM DTT and diluted 200-fold into MBP refolding buffer. FCS was recorded for 3 h. Analysis was then performed for time windows of 1 min. The average number of particles N was extracted from the amplitude ρ of the fit to the auto correlated data. . Arithmetic mean ± s.d. from at least 3 independent experiments is shown. The dotted line shows the simulated increase in the number of particles, assuming transient dimeric aggregation of DM-MBP during refolding, as described in (Fig. 4.2 B).

In order to achieve a more detailed understanding of the Anfinsen cage model, we performed a kinetic simulation using Berkley Madonna (University of Berkeley). We used a kinetic model in which a monomeric intermediate folds to the native state with a rate k_f. We assumed that k_f should be fixed to the fastest rate that can be possibly observed, i.e. in presence of GroEL, when aggregation is completely prevented, k_f ~0.2 min^{-1}. In this simulation the monomeric intermediate state can also form the smallest

possible aggregate, a dimer, with a rate k_a. Dissociation of this aggregate is determined by a rate k_d. We then increased the affinity for dimer-formation until the rate of formation of native particles k_{obs} occurred at the slow rate of spontaneous folding. As expected, the formation of native protein did not follow single exponential behavior, due to the presence of the rate limiting and concentration dependent side reaction (Fig. 4.2 B). Most importantly, the simulated data showed that in order to reduce the observed rate of spontaneous refolding k_{obs} ~6 fold at 100 pM protein concentration, the affinity for a transient dimer would have to be in the picomolar range and such a dimer should be present with an abundance of more than 10 pM during the first 30 min of the refolding reaction, i.e. double the concentration of the described FCCS positive control. Taken together, the combination of theoretical simulation as well as experimental FCCS data clearly excludes the existence of transient aggregates in the spontaneous refolding pathway of DM-MBP at 100 pM protein concentration.

Analysis by conventional fluorescence correlation spectroscopy (FCS) further confirmed the monomeric nature of DM-MBP during refolding. In single color FCS the amplitude of the correlation function $G(0)$ is inversely proportional to the concentration of diffusing particles. If transient aggregation occurs, a gradual increase in the concentration of diffusing molecules should be observed, corresponding to the disassembly rate (k_d) of the transiently aggregated particles. DM-MBP (312C) labeled with Atto647N was spontaneously refolded at 100 pM concentration. During refolding, analysis of the amplitude of the auto correlated signal revealed that the number of particles inside the observation volume is constant over the full refolding time, which again rules out the presence of transiently aggregated material during the refolding of DM-MBP at single molecule level (Fig. 4.2 C).

4.2.2 smFRET can be used to assess folding rates at low concentrations

Having decisively ruled out the presence of transient oligomeric species during DM-MBP refolding at a final protein concentration of 100 pM, we wanted to use these conditions to test the passive cage model for DM-MBP folding (Apetri and Horwich, 2008; Tyagi et al., 2011). If the spontaneous folding rate is indeed limited by a disaggregation reaction at higher concentrations but not at infinite dilution, then at 100 pM the spontaneous and assisted folding rates should be equal. This observation would support the passive cage model. If however at 100 pM, i.e. in absence of transient aggregation, the folding of DM-MBP is accelerated by GroEL/ES, the acceleration is likely a result of a modulation of the substrate folding pathway inside the GroEL central cavity.

In order to determine protein folding rates at high dilution, we developed a novel approach to study the folding of double-labeled DM-MBP by single molecule FRET. As described previously (Sharma et al., 2008), a mutant of DM-MBP (D30C, A312C) double-labeled with Atto532 and Atto647N, DM-MBP (DL), shows distinct FRET efficiency (f_E) distributions in the native state and when bound to GroEL (Fig. 4.3 A and B). The native protein is a compact conformer with a single population of an

average high f_E of 0.72. In contrast, the GroEL bound state of DM-MBP shows a higher heterogeneity with ~40% of all molecules populating an expanded conformer with a low f_E of 0.06 and the remainder of the molecules populating a broadened distribution with a mean f_E of 0.38.

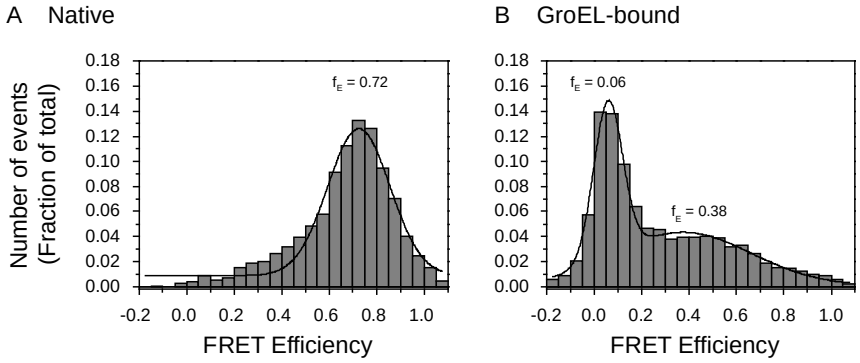

Figure 4.3 DM-MBP (DL) FRET efficiency distributions for native and GroEL-bound conformers

(A) smFRET histogram of native DM-MBP (DL) at 100 pM concentration and 20°C in MBP refolding buffer. **(B)** smFRET histogram of GroEL bound DM-MBP (DL) at 100 pM concentration and 20°C in MBP refolding buffer. DM-MBP (DL) was denatured in 6 M GuHCl, 10 mM DTT and diluted 200-fold into MBP refolding buffer containing 2 µM GroEL. The sample was immediately subjected to smFRET analysis for one hour.

Taking advantage of the ability of GroEL to recognize and bind substrate folding intermediates, we used GroEL to stop the spontaneous refolding reaction at different times. Non-native DM-MBP rapidly binds to the GroEL apical domains, and is thus reverted to the low FRET state. The assisted refolding reaction could be stopped by the addition of the enzyme apyrase, resulting in rapid hydrolysis of ATP to ADP and AMP. GroEL is thereby reverted to its apo state, and can thus act as a single molecule sorting machine which separates native protein from not yet folded material.

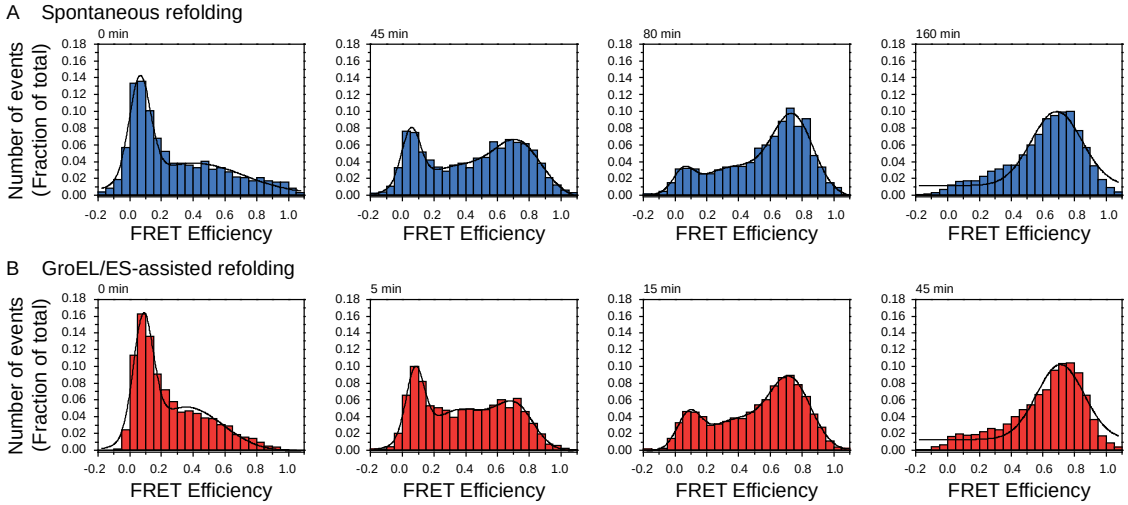

Figure 4.4 DM-MBP refolding followed at single molecule level using smFRET

(A) and **(B)** From left to right, smFRET analysis of representative, consecutive kinetic points during spontaneous (blue) and assisted (red) refolding of DM-MBP (DL) at 20°C and 100 pM DM-MBP concentration. DM-MBP (DL) was denatured in 6 M GuHCl, 10 mM DTT and diluted 200-fold into MBP refolding buffer (spontaneous) or MBP refolding buffer containing 2 µM GroEL (assisted). Assisted refolding was started by addition of 4 µM GroES and 5 mM ATP. The refolding reaction was stopped at the indicated time points by addition of 2 µM GroEL or 5 U apyrase. The given numbers indicate refolding times in minutes. Representatives of at least three independent experiments are shown.

As expected, we observed that in both cases over time the high FRET population corresponding to native protein increased, while the low FRET population corresponding to GroEL bound material concomitantly decreased (Fig. 4.4 A and B).

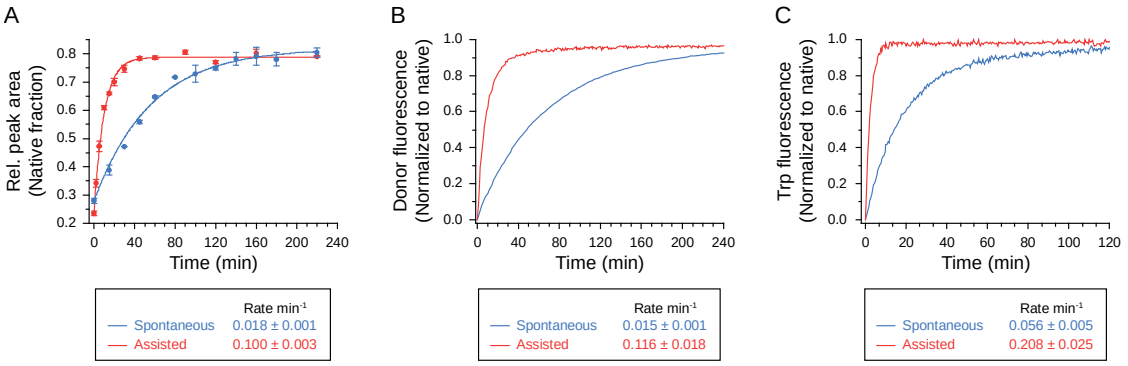

Figure 4.5 DM-MBP refolding is accelerated by GroEL/ES at single molecule level

(A) Quantification of smFRET data shown in Fig 4.4 A and B obtained at 100 pM protein concentration. The relative area of the high f_E peak, corresponding to native DM-MBP (DL), was quantified for spontaneous (blue) as well as assisted (red) refolding. The data was fitted with a single exponential function for folding rate

extraction. Arithmetic mean ± s.d. is shown. **(B)** Spontaneous (blue) and assisted (red) refolding kinetics of DM-MBP (DL), measured by conventional fluorescence spectroscopy in an ensemble approach at 100 nM final protein concentration in MBP refolding buffer at 20°C. Refolding was monitored at donor excitation and emission wavelengths of 532 nm and 550 nm respectively. Representative curves of at least three individual repeats are shown. **(C)** Spontaneous (blue) and assisted (red) refolding kinetics of unlabeled DM-MBP (D30C, A312C), measured by conventional fluorescence spectroscopy in an ensemble approach at 100 nM final protein concentration in MBP refolding buffer at 20°C. Refolding was monitored at tryptophan excitation and emission wavelengths of 295 nm and 345 nm respectively. Representative curves of at least three individual repeats are shown.

Quantification of the corresponding FRET peak areas over time enabled us to measure protein folding rates at a final concentration of 100 pM DM-MBP. Strikingly, we found the spontaneous refolding rate to be 0.02 min^{-1}, and therefore ~5.6 fold slower than the assisted refolding rate (0.1 min^{-1}) even at picomolar concentrations (Fig. 4.5 A). To validate our findings we measured the folding rate of labeled DM30/312 by following the increase in donor fluorescence on an ensemble level at a protein concentration of 100 nM (Fig. 4.5 B). We observed a rapid initial decrease in donor fluorescence upon dilution from denaturant, due to collapse of the protein chain and FRET (data not shown). This initial decrease was followed by a gradual increase, apparently due to changes in the chemical environment of the donor fluorophore upon folding. The observed rates for spontaneous (0.02 min^{-1}) and for assisted refolding (0.12 min^{-1}) were in good agreement with the single molecule data and again showed a significant acceleration of protein folding in presence of chaperonin. We also followed the increase in tryptophan fluorescence for the unlabeled protein during refolding at a protein concentration of 100 nM (Fig. 4.5 C). The observed rates for spontaneous (0.06 min^{-1}) and for assisted folding (0.21 min^{-1}) showed a ~3.7 fold acceleration of protein folding in presence of chaperonin. Covering a broad concentration range and measuring protein folding at highly dilute concentrations, we therefore unequivocally ruled out transient aggregation as a cause for the observed rate acceleration in chaperonin-mediated folding. These observations constitute strong evidence for an active cage mechanism of GroEL assisted protein folding, in which the physical and chemical properties of the GroEL cavity likely impact the kinetic energy barriers for productive substrate refolding.

4.2.3 A novel FCS-based approach to investigate GroEL substrate refolding

Not only does GroEL binding shift DM30/312 to a low FRET state but it also slows down the apparent diffusion rate of labeled DM30/312 (Fig. 4.6 A). The diffusion rate of proteins can be analyzed by FCS using the auto correlated signal of only one dye, and therefore also single labeled particles. As expected, we observed a significant difference in the diffusion speed of GroEL-bound DM-MBP (~49 $\mu m^2 s^{-1}$) and free DM-MBP (~160 $\mu m^2 s^{-1}$).

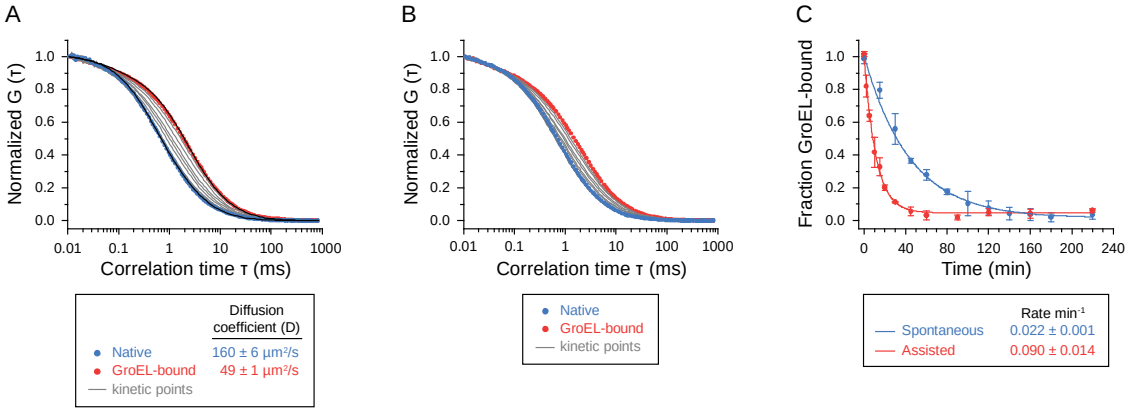

Figure 4.6 FCS can be used to assess GroEL substrate refolding at single molecule level

(A) Representative auto correlation curves of Atto647N fluorescence for GroEL bound (red) and spontaneously refolded (blue) DM-MBP (DL) as well as kinetic points taken during spontaneous refolding (grey) at 100 pM. Experiments were performed as in Fig. 4.4 A. For starting and final timepoints diffusion coefficients were calculated as arithmetic mean ± s.d. from three independent experiments. **(B)** Representative auto correlation curves of Atto647N fluorescence for GroEL bound (red) and chaperonin refolded (blue) DM-MBP (DL) as well as kinetic points taken during assisted refolding (grey) at 100 pM. Experiments were performed as in Fig. 4.4 B. **(C)** Refolding kinetics showing spontaneous and assisted refolding of DM-MBP (DL) as measured by the mean particle diffusion time through the confocal observation volume. The mean diffusion time was extracted from auto correlation data of Atto647N fluorescence as shown in (A) and (B) and converted into the fraction of GroEL bound material, which was plotted versus refolding time. Single exponential fitting was used to extract the rate of folding. Arithmetic mean ± s.d. of three independent experiments is shown.

We analyzed the data for the average diffusion speed of Atto647N-labeled DM-MBP molecules during spontaneous (Fig. 4.6 A) and assisted (Fig. 4.6 B) refolding. From the time dependent decrease in diffusion time, we were able to extract refolding rates (Fig. 4.6 C), which were in good agreement with the smFRET data. Thereby we established a simplified approach to study the refolding of a GroEL dependent substrate protein which only carries a single fluorescent label, at highly dilute concentrations, using FCS. It is self-evident that this approach could potentially be extended in a generalized manner to study folding rates of highly aggregation prone proteins with an affinity to chaperonin complexes.

4.3 Encapsulation by GroEL reduces substrate flexibility

Having decisively ruled out the passive cage model as a working mechanism for accelerated folding in presence of chaperonin, we wanted to further elucidate how GroEL modifies kinetically trapped folding intermediates. It has been suggested previously that DM-MBP forms a highly dynamic folding intermediate with a high entropic barrier to the native state, and that this entropic barrier is overcome during encapsulation inside the GroEL cavity (Brinker et al., 2001; Chakraborty et al., 2010; Tang et

al., 2006). To test whether a decrease in chain entropy correlates with the folding of a flexible intermediate state to the native conformation, and whether this process is accelerated in the chaperonin cage, we used PET-FCS as a method. In PET-FCS the fluorescence of an oxazine dye (Atto655) is quenched in close proximity of a Trp residue by direct transfer of an electron (Neuweiler et al., 2009; Sauer and Neuweiler, 2014; Teufel et al., 2011). As Atto655 does not show a significant amount of triplet state formation or blinking, it is well suited to follow dynamics in timescales from nanoseconds to milliseconds (Sauer and Neuweiler, 2014) and has been used as a PET-FCS probe for early events and dynamics in protein folding (Neuweiler et al., 2009; Teufel et al., 2011).

MBP contains 8 Trp residues spaced throughout the sequence (Fig. 4.1 A) that could serve as potential dynamic quenchers for Atto655 in case a flexible structural state is formed, i.e. in case van der Waals contact between dye and Trp residues is possible by intra chain contact formation. The closest Trp residue (W232) in the crystal structure of the DM-MBP native state (PDB: 1OMP) is 15.8 Å distant from position 312. As PET requires van der Waals contact, the DM-MBP native state does not show a PET induced correlated signal, and can be well approximated by a simple diffusion model (data not shown).

In contrast, correlation data obtained during the first minute of refolding, when most DM-MBP molecules populate the intermediate state (Chakraborty et al., 2010), shows fast fluctuations of the correlated signal at a microsecond timescale. As a result, the obtained data cannot be fitted to a simple single component diffusion model (Fig. 4.7 A). An additional exponential term needs to be added to the fitting equation (Fig. 4.7 B). The relaxation time τ_R of this component gives a measure for chain motion and was found to be 40 ± 3 µs, while its amplitude F is proportional to the abundance of particles in the intermediate state. The fast fluctuation in the µs timescale demonstrates the highly dynamic nature of this intermediate and correspondingly high chain entropy. The addition of a second exponential term (one diffusion two exponentials) did not further improve the quality of the fit (data not shown), indicating that a one diffusion one exponential model is most appropriate.

As a control, WT-MBP (312C) labeled with Atto655 showed no PET-induced fluctuation during refolding, as WT-MBP folds significantly faster ($t_{1/2}$ ~23 s) and does not significantly populate the flexible intermediate state (Chakraborty et al., 2010). Accordingly, the resulting auto-correlation data can be well fitted using a one diffusion model (Fig. 4.7 C). Addition of an exponential term could not improve the fit further (Fig. 4.7 D). This clearly demonstrates that a kinetically trapped, flexible intermediate is only observed by PET-FCS in the case of double mutant MBP, which has been previously shown by hydrogen deuterium exchange measurements as well as equilibrium unfolding and refolding trajectories (Chakraborty et al., 2010).

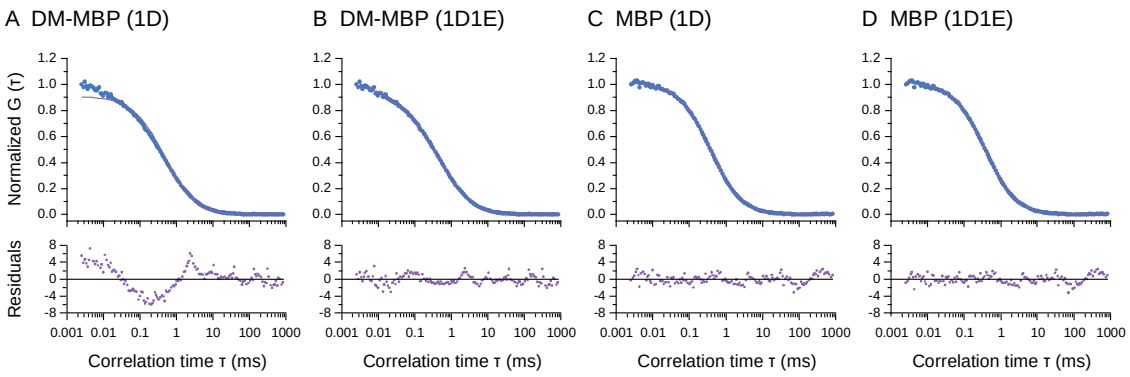

Figure 4.7 Conformational dynamics of a DM-MBP folding intermediate can be assessed by PET-FCS

(A) PET-FCS measurement for DM-MBP folding intermediate. DM-MBP (312C) labeled with Atto655 was denatured in 6 M GuHCl, 10 mM DTT and diluted 200-fold into MBP refolding buffer to a final concentration of 1 nM. PET-FCS was recorded during the first minute of spontaneous refolding. The resulting auto correlated data was fitted with a simple one diffusion model containing an amplitude ρ and a diffusion time τ_D. **(B)** The data described in (A) was fitted with a one diffusion one exponential model, with the additional exponential term containing an amplitude F and a relaxation time τ_R. **(C)** PET-FCS measurement for MBP folding intermediate. MBP (312C) labeled with Atto655 was denatured in 6 M GuHCl, 10 mM DTT and diluted 200-fold into MBP refolding buffer to a final concentration of 1 nM. PET-FCS was recorded during the first minute of spontaneous refolding. The resulting auto correlated data was fitted with a simple one diffusion model containing an amplitude ρ and a diffusion time τ_D. **(D)** The data described in (C) was fitted with a one diffusion one exponential model, with the additional exponential term containing an amplitude F and a relaxation time τ_R. One representative measurement of at least three individual experiments is shown. Fit residuals are shown in all cases to demonstrate the quality of the regression.

When we allowed DM-MBP (312C) to refold for 2 h, we found that the auto correlation curve of the refolded state did not contain significant fluctuation for short correlation times (Fig. 4.8 A). This indicates that DM-MBP refolded to the compact native state. As expected, we found that the amplitude of the PET-signal measured at different time points during refolding at 1 nM protein concentration, i.e. the fraction of particles populating the dynamic intermediate state, decreased over time with the refolding rate as established by Trp fluorescence at higher concentrations (Fig. 4.8 B and 4.8 C). The folding rates measured by the two different approaches (PET-FCS and Trp fluorescence) were in excellent agreement (Fig. 4.9 A).

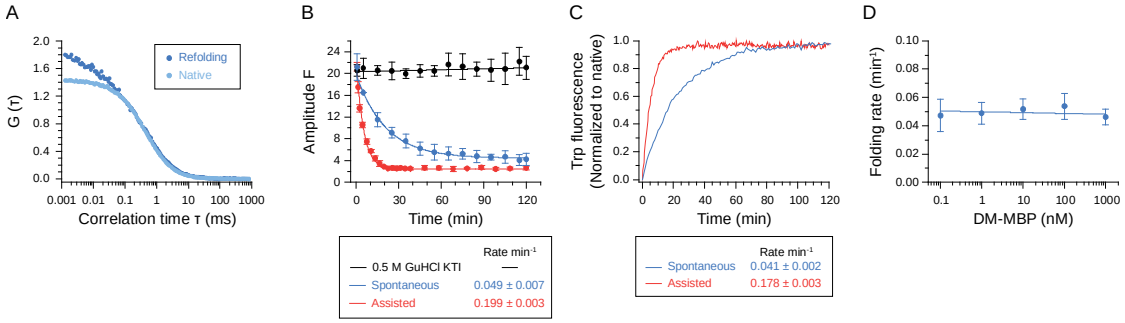

Figure 4.8 PET FCS based refolding rate measurements

(A) PET disappears upon refolding of DM-MBP. DM-MBP (312C) labeled with Atto655 was denatured in 6 M GuHCl, 10 mM DTT and diluted 200-fold into MBP refolding buffer to a final concentration of 1 nM. PET-FCS was recorded during the first minute (refolding) and after completion (native) of spontaneous refolding. Representative data of at least three individual experiments is shown. **(B)** Kinetic evaluation of the change in amplitude F during refolding. DM-MBP (312C) labeled with Atto655 was denatured as in (A) and diluted 200-fold into MBP refolding buffer (spontaneous) or MBP refolding buffer containing 2 µM GroEL (assisted). Assisted refolding was started by addition of 4 µM GroES and 5 mM ATP. FCS recording was started immediately and continued for two hours. In addition denatured DM-MBP (312C) was diluted into MBP refolding buffer containing 0.5 M GuHCl to stabilize the kinetically trapped intermediate state (KTI). Time windows of 2 min for early time points and 10 min for late time points (GroEL/ES-assisted) or time windows of 10 min for early and late time points (spontaneous refolding and KTI) were correlated and fitted with a one diffusion one exponential function. Refolding rates were extracted by single exponential fits to plots of the amplitude of the exponential component F versus refolding time. Arithmetic mean ± s.d. from at least 3 independent measurements is shown. **(C)** Refolding of unlabeled DM-MBP (312C) measured by increase in Trp fluorescence. DM-MBP (312C) was denatured as in (A) and diluted 200-fold to a final concentration of 100 nM into MBP refolding buffer (spontaneous) or MBP refolding buffer containing 2 µM GroEL (assisted). Assisted refolding was started by addition of 4 µM GroES and 5 mM ATP. Representative curves of at least three independent experiments are shown. **(D)** DM-MBP folding rate is concentration independent. The refolding rate of DM-MBP (312C) labeled with Atto655 was measured as described in (B) following the time dependent decrease in F upon dilution into MBP refolding buffer to a final concentration of 100 pM or 1 nM. For concentrations higher than 1 nM unlabeled denatured DM-MBP (312C) was added to the refolding mix. Rates are shown as arithmetic mean ± s.d. of at least three independent experiments.

Thus we established a novel approach to assess protein folding rates of DM-MBP by direct observation of chain dynamics. This clearly demonstrates that folding of DM-MBP is rate limited by the slow conversion of a highly flexible intermediate to the compact native state, crossing a kinetic barrier with a large entropic component. Strikingly, the decrease in the amplitude of the fast fluctuating component is accelerated ~4-fold in the presence of GroEL/ES, indicating that GroEL overcomes the entropic barrier and thereby enhances the folding speed of DM-MBP. It has been previously shown that the kinetically trapped intermediate formed by DM-MBP is a stable and predominant conformer in the presence of 0.5 M GuHCl (Chakraborty et al., 2010). In agreement, the amplitude of the PET signal was identical to the amplitude observed for the folding intermediate during the early phase of refolding and

was stable over the course of two hours (Fig. 4.8 B, black). Notably, the spontaneous refolding rate of DM-MBP measured by PET-FCS was concentration independent over four orders of magnitude, again supporting the active cage model and refuting transient aggregation as a cause for observed rate acceleration by chaperonin (Fig. 4.8 D).

4.3.1 Evidence for conformational restriction of encapsulated substrate

To further investigate conformational flexibility of the folding polypeptide chain, we extracted the relaxation time of the observed PET signal under different experimental conditions. We were especially interested in the impact on substrate conformational flexibility exerted by the GroEL central cavity. The relaxation time of the PET signal, which is a direct measure of the kinetics of chain motion (Neuweiler et al., 2009; Sauer and Neuweiler, 2014; Teufel et al., 2011), was 40 ± 3 µs during the first minute of spontaneous folding (Fig. 4.9 B). Consistent with this observation, the relaxation time found for the KTI in 0.5 M GuHCl was 34 ± 10 µs (Fig. 4.9 B). The τ_R observed for the GroEL bound protein was 59 ±10 µs, indicating conformational restriction by interaction with the GroEL apical domains (Fig. 4.9 B). Most interestingly, the relaxation time during the first minute of GroEL assisted refolding (≤ 20% of molecules folded) was increased ~2.5 fold to 96 ± 5 µs, correlating with significantly lower chain motility (Fig. 4.9 B).

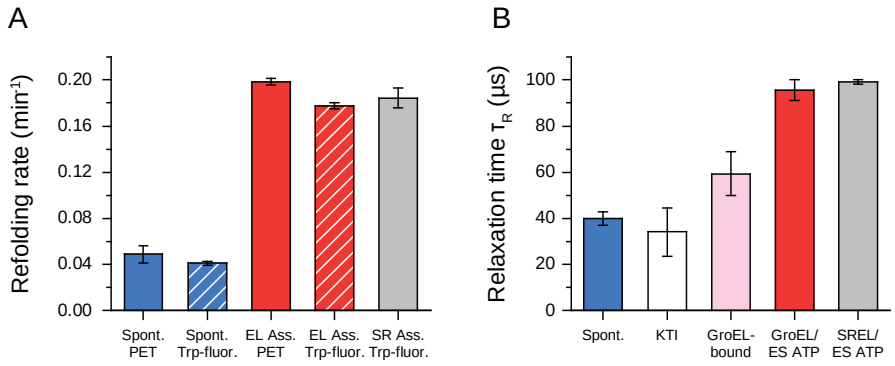

Figure 4.9 Relaxation times and folding rates of DM-MBP in different conditions

(A) Refolding rate of DM-MBP (312C) followed either by PET-FCS at 100 pM protein concentration as in Fig. 4.8 B or by tryptophan fluorescence at 100 nM concentration as in Fig. 4.8 C. Folding rates were measured in MBP refolding buffer or SREL buffer (50 mM Hepes/NaOH pH7.5, 20 mM KCl, 10 mM $MgCl_2$) for experiments with SREL. Rates are shown as arithmetic mean ± s.d. of at least three independent experiments. **(B)** Relaxation times observed for 1 nM denatured, Atto655 labeled DM-MBP (312C) during the first minute of spontaneous (spont.) or assisted (GroEL/ES ATP) refolding as well as after dilution into MBP refolding buffer containing either 0.5 M GuHCl (KTI) or 2 µM GroEL (GroEL-bound). Data for SREL assisted refolding was obtained in SREL buffer during the first minute of refolding. Relaxation times were extracted as fit

parameters from one diffusion one exponential fits to autocorrelation data obtained for the described experiments. Relaxation times are shown as arithmetic mean ± s.d. of at least three independent experiments. PET related data for this figure was obtained in collaboration with Dr. Shubhasis Haldar.

To investigate this further, we used the non-cycling single ring variant of GroEL (SREL) (Weissman et al., 1996) that forms a stable complex with GroES and undergoes a single round of ATP hydrolysis due to a lack of allosteric signaling from the trans-ring. SREL allows substrates to be studied in the encapsulated state, without the further complication of repetitive cycles of binding and release. However, the SREL/GroES complex is salt sensitive (Hayer-Hartl et al., 1996; Motojima et al., 2012). Thus all experiments with SREL were performed in a low salt SREL buffer (50 mM Hepes/NaOH pH7.5, 20 mM KCl, 10 mM $MgCl_2$) with Urea denatured DM-MBP. First, we confirmed stable complex formation and efficient encapsulation of substrate by a series of size exclusion experiments. In order to achieve accurate and sensitive detection of DM-MBP in gel filtration, we used DM-MBP(312C) labeled with Atto655. We performed gel filtration chromatography of a preformed mixture of SR1, GroES and fluorescently labeled DM-MBP(312C), and subsequently assessed the amount of encapsulated DM-MBP by fluorescence imaging and coomassie staining.

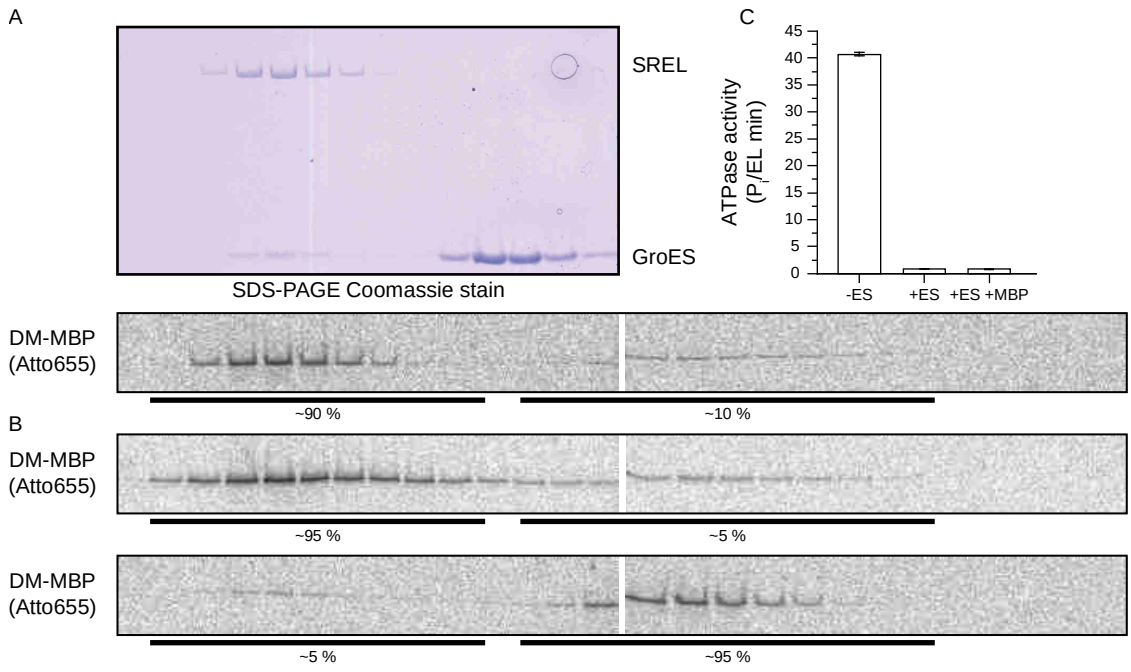

Figure 4.10 SREL forms a stable cis-like complex with GroES that stably encapsulates DM-MBP

(A) SDS-PAGE of size exclusion fractionation of a preformed SREL, GroES, DM-MBP complex. DM-MBP (Atto655) was unfolded in 10 M urea, 10 mM DTT for 1 h at 50°C and diluted 200-fold into SREL buffer at 20°C containing 1 µM SREL. After 5 min encapsulation of DM-MBP was initiated by addition of 4 µM GroES

and 1 mM ATP. The reaction mix was applied to a Superdex 200 gel filtration column equilibrated in SREL buffer, 50 mM urea, 1 mM ATP. Fractions of 50 µL were collected over a period of 30 min. Top panel shows Coomassie stained SDS PAGE of collected fractions. Bottom panel shows fluorescence scan of the same SDS-PAGE including densitometric quantification of free and complexed DM-MBP. **(B)** Size exclusion experiment to establish long term complex stability. The preformed SREL, DM-MBP and GroES complex was incubated for 30 min at 20°C, followed by size exclusion chromatography without further additions (top panel) or subsequent to addition of 50 mM CDTA, 200 mM KCl, 70 mM GuHCl to induce dissociation of the SREL GroES complex and release encapsulated DM-MBP (bottom panel). The collected fractions were analyzed by SDS-PAGE as described in (A). **(C)** GroES significantly inhibits SREL ATPase activity. ATPase activity of 100 nM SREL in SREL buffer was measured at 20°C by a coupled enzymatic assay following the photometric conversion of NADH, in absence and presence of 400 nM GroES or 400 nM GroES and 1 µM denatured DM-MBP. The rate of NADH consumption was converted into the production of phosphate by one SREL 7-mer per minute. Data for figures A and B kindly provided by Goran Milićič.

SREL and GroES co-eluted, indicating stable complex formation (Fig. 4.10 A, top). ~90% of Atto655-labeled DM-MBP (312C) co-eluted with SREL and GroES, corresponding to encapsulated protein. The remaining ~10% of DM-MBP eluted at lower molecular weight fractions, corresponding to free protein (Fig. 4.10 A, bottom). DM-MBP was stably encapsulated for at least 30 minutes (Fig. 4.10 B, top) but was efficiently released upon complex dissociation by addition of Mg^{2+}-chelator (50 mM CDTA), GuHCl (70 mM) and high salt (200 mM KCl) (Fig. 4.10 A, bottom). In addition, we found that GroES efficiently inhibited the ATPase activity of SREL (41 P_i/EL min) reducing it to a basal level of 0.8 P_i/EL min (Fig. 4.10 C). Strong inhibition of SREL ATPase activity by GroES indicates formation of a stable complex that undergoes only a single round of ATP hydrolysis. Note that also the presence of an excess of denatured substrate protein did not affect efficient GroES binding (Fig. 4.10 C).

When we measured the refolding rate of DM-MBP(312C) by Trp fluorescence, we observed that SREL/ES assisted refolding resulted in a similar acceleration of the refolding rate as provided by the GroEL double ring / GroES system (Fig 4.9 A). Note that in-cage refolding of DM-MBP, in contrast to spontaneous refolding, is salt insensitive (Chakraborty et al., 2010 and this work). We also found that full refolding yields were obtained in case of SREL-assisted refolding (data not shown). Therefore, a single round of encapsulation is sufficient, not only to result in full yield but also to catalyze folding. This clearly shows that the GroEL/ES system is an active system, with an active cage providing a scaffold for accelerated folding of substrate proteins.

When we measured PET-FCS for stably encapsulated Atto655-labeled DM-MBP during the first minute of folding inside SREL, we found a τ_R of 99 ±1 µs, identical to the value obtained for the cycling GroEL WT system under the same low salt SREL buffer condition (Fig. 4.9 B). These findings indicate an important reduction of chain entropy inside the GroEL cavity, even compared to the GroEL bound state. In other words, the encapsulated substrate experiences a reduction in conformational freedom. As described earlier, steric confinement inside the cage could result in reduced chain entropy and therefore

a more rapid conversion of the highly dynamic intermediate to the native state. In PET-FCS this effect is not only shown by a rapid decrease in the amplitude of the fast fluctuating component, but also in an increase in the characteristic decay time of this fast fluctuating component. Therefore, PET-FCS does not only allow us to follow folding kinetics of DM-MBP inside the GroEL cavity, but also the difference in conformational space that can be explored by the folding polypeptide chain. Finally, it is interesting to note that an increase in relaxation time and therefore decrease in conformational flexibility correlates well with an increase in refolding rate (Fig. 4.9).

4.4 DM-MBP refolding but not aggregation is salt dependent

It has previously been shown that spontaneous folding of DM-MBP can be chloride salt dependent, whereby a decrease in salt concentration results in an apparent increase in the folding rate (Apetri and Horwich, 2008; Chakraborty et al., 2010; Tyagi et al., 2011). This effect could have two different underlying mechanisms. Either the folding energy landscape changes due to changes in the electrochemical environment of the folding protein (Chakraborty et al., 2010) or a decrease in ionic strength of the solvent is concomitant with a decrease in the aggregation propensity of the folding intermediate (Apetri and Horwich, 2008; Tyagi et al., 2011). While a change in the folding energy landscape would be in line with an active cage model of GroEL function, an impact on aggregation propensity would be consistent with the passive cage model. To discriminate between the two effects we decided to measure the spontaneous refolding rate of DM-MBP at low salt concentration (20 mM KCl), using single molecule concentrations to exclude the effect of aggregation. We used the already described smFRET-based folding assay to measure refolding of DM-MBP in MBP LS refolding buffer (20 mM KCl) (Fig. 4.11 A). Interestingly, we found that DM-MBP folds at an increased rate of 0.055 min^{-1} at a concentration of 20 mM KCl (Fig. 4.11 B), a ~3-fold acceleration of the folding rate obtained at 200 mM KCl. Chloride salt therefore decreases the rate of folding by modulation of the intrinsic folding properties of DM-MBP (Chakraborty et al., 2010), rather than influencing it's aggregation. To investigate this salt effect further, we employed PET-FCS to test for a salt dependent change in chain entropy of the DM-MBP folding intermediate. We observed the formation of a folding intermediate to a similar extent at both 20 mM and 200 mM KCl, as judged by the amplitude F of the observed PET signal (Fig. 4.11 C). The time dependent decrease in F, as already described, correlates with the folding rate of DM-MBP. Folding rates measured by PET-FCS at 100 pM (Fig. 4.11 C) and by Trp fluorescence at 100 nM (Fig. 4.11 D) showed a similar ~2-fold rate acceleration at a 10-fold decrease in salt concentration for labeled and unlabeled DM-MBP (312C), respectively. Interestingly, the relaxation time of the observed PET signal was ~2.3-fold increased with a decrease in salt concentration, corresponding to a decrease in protein chain flexibility (Fig. 4.11 E) and comparable to

the increase in folding rate. Measuring the DM-MBP (312C) refolding rate at 100 nM by Trp fluorescence, we also confirmed a previously described result (Chakraborty et al., 2010), that the folding rate of DM-MBP in presence of GroEL/ES/ATP is salt independent, in striking contrast to spontaneous folding (Fig. 4.11 D). These findings suggest that the electrochemical environment of the GroEL cage renders the folding of DM-MBP salt insensitive. Taken together with the observation that the electrochemical environment can significantly alter the folding pathway of DM-MBP, it is likely that in addition to the steric confinement effect the highly negatively charged inner cage wall (42 net negative charges) of GroEL strongly impacts the folding trajectory of the DM-MBP intermediate state. This model has been suggested previously (Tang et al., 2006), but had remained controversial (Motojima et al., 2012).

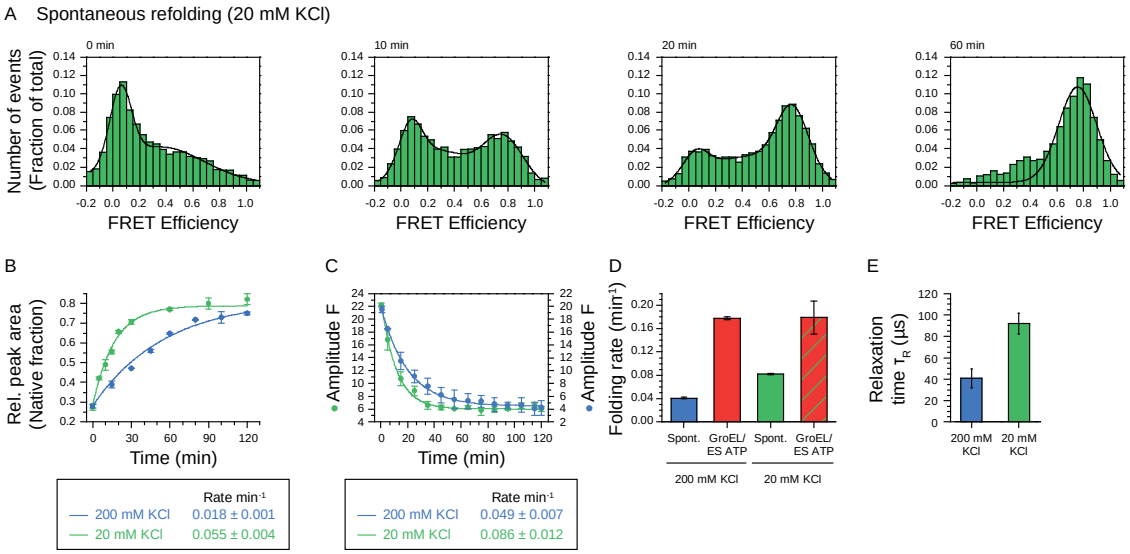

Figure 4.11 DM-MBP refolding but not aggregation is salt dependent

(A) DM-MBP (DL) spontaneous refolding under low salt conditions measured by smFRET. DM-MBP (DL) was denatured in 6 M GuHCl, 10 mM DTT and diluted 200-fold at 20°C into MBP LS refolding buffer (20 mM Tris/HCl pH 7.5, 20 mM KCl, 5 mM $Mg(C_2H_3O_2)_2$) to a final protein concentration of 100 pM. Refolding was stopped at different kinetic points by addition of 2 µM GroEL. Samples were then subjected to smFRET analysis. Representative histograms of at least three independent experiments for characteristic kinetic points are shown. **(B)** DM-MBP spontaneous refolding is accelerated by a decrease in salt concentration. Refolding data obtained from smFRET measurements as shown in (A) was analysed kinetically and compared to spontaneous refolding at physiological salt concentration of 200 mM KCl as shown in Fig. 4.4 A. Data is shown as arithmetic mean ± s.d. of at least three independent experiments. **(C)** Refolding rate of DM-MBP (312C) labeled with Atto655 measured by PET-FCS. Unfolded DM-MBP (312C) was refolded spontaneously at 100 pM and 20°C by dilution into either MBP refolding buffer (200 mM KCl) or MBP LS refolding buffer (20 mM KCl). Refolding was followed by time dependent decrease of the fit parameter F as described in Fig. 4.8 B. **(D)** Comparison of spontaneous and assisted folding rates measured by tryptophan fluorescence at

different salt concentrations (20 mM and 200 mM KCl). Refolding of unlabeled DM-MBP (312C) was measured following the increase in Trp fluorescence upon 200-fold dilution from denaturant into buffer at 20°C and to a final protein concentration of 100 nM. **(E)** Comparison of conformational flexibility of DM-MBP folding intermediate at different salt concentrations (20 mM and 200 mM KCl). DM-MBP (312C) labeled with Atto655 was denatured in 6 M GuHCl, 10 mM DTT and diluted into buffer containing either 20 mM or 200 mM KCl. FCS recording was started immediately for 1 min. The autocorrelated data was fitted with a one diffusion one exponential model to extract the relaxation time τ_R of the conformational flexibility. Values for τ_R are given as arithmetic mean ± s.d. of at least three independent experiments. Data for figures C, D and E was obtained in collaboration with Dr. Shubhasis Haldar.

4.5 Assisted substrate folding occurs inside the GroEL cage

The finding that the substrate chain entropy is strongly modified during folding in the presence of GroEL/ES as compared to folding in free solution, suggested an important role of substrate confinement in the GroEL central cavity. In addition, chain entropy was reduced to the same extent during both cycling and stable encapsulation (Fig. 4.9 B), indicating that the substrate protein spends most of its time in the encapsulated state during folding. Since the iterative annealing model (Sparrer et al., 1996; Yang et al., 2013), in contrast to the active cage model (Brinker et al., 2001; Chakraborty et al., 2010; Tang et al., 2006, 2008), does not assign a functional relevance to substrate encapsulation, we decided to quantitatively measure the time a given substrate molecule spends inside the cage and in bulk solution during a single round of the chaperonin cycle.

We first tested whether substrate folding occurs predominantly inside the cage (in-cage folding) or outside the cage (out-of-cage folding). Therefore we measured the diffusion time of DM-MBP (DL) during the first minute of GroEL/ES assisted refolding. We found that in the first minute after starting the folding reaction with ATP, the average diffusion time of DM-MBP (DL) was in good agreement with the diffusion time found for GroEL-bound denatured DM-MBP (DL) and could be well discriminated from freely diffusing DM-MBP (DL) during the first minute of spontaneous refolding (Fig. 4.12 F). Diffusion time measurements therefore indicate that during the first minute of assisted refolding the majority of DM-MBP is in complex with a chaperonin molecule.

It is, however, not clear whether DM-MBP is mostly in the unfolded state, bound to the GroEL apical domains, or encapsulated inside the cavity formed by GroEL and GroES. We have already established that DM-MBP (DL), when bound to the GroEL apical domains, shows a FRET efficiency distribution corresponding to stretched conformations (Fig. 4.3 B and Fig. 4.12 A). In contrast, when in free solution DM-MBP (DL) shows a compact conformation with a high FRET efficiency distribution (Fig. 4.3 A). Based on previous observations, especially conformational confinement as observed by PET-FCS, we reasoned that when encapsulated inside the GroEL central cavity, DM-MBP would adopt

a compact conformation and should also show a high FRET-efficiency distribution. In fact, when we refolded urea denatured DM-MBP (DL) in presence of SREL/ES/ATP under conditions where we previously ensured efficient and stable encapsulation (Fig. 4.10), we found that DM-MBP adopts a compact conformation with an average FRET-efficiency of $f_E = 0.66$ (Fig. 4.12 E) during the first minute of refolding. Note that DM-MBP(DL) showed very similar FRET efficiency distributions when bound to SREL and GroEL, with ~34-40% of molecules in a highly stretched conformation corresponding to $f_E = 0.06$ in the case of GroEL, and $f_E = 0.1$ in the case of SREL. The remainder of the molecules showed an intermediate f_E of 0.38 and 0.44 respectively (Fig. 4.12 A and D).

Having established that DM-MBP (DL) adopts a compact conformation with a high f_E when stably encapsulated in SREL, we tested the conformational state of DM-MBP(DL) during cycling conditions. When we measured single molecule FRET during the first minute of GroEL assisted refolding in presence of GroES and ATP, we observed a bimodal FRET-efficiency distribution with most of the molecules being in a compact state with a high f_E of 0.65 and the remainder of molecules being in a stretched conformation with a f_E of 0.12 (Fig. 4.12 B). We therefore reasoned, that while a part of the molecules are bound to the GroEL apical domains during refolding, the majority of molecules is encapsulated inside GroEL, or reached the native state and is therefore not GroEL associated. For absolute quantification, we assessed the amount of folded material by stopping the assisted folding reaction after one minute by addition of apyrase (Fig. 4.12 C). Using the established FRET peak quantification approach, we found that the fraction of folded molecules was ~12%. Taking this into consideration, we calculated that during the first minute of assisted refolding ~82% of GroEL associated molecules were encapsulated and a small amount of ~12% were in the GroEL bound state.

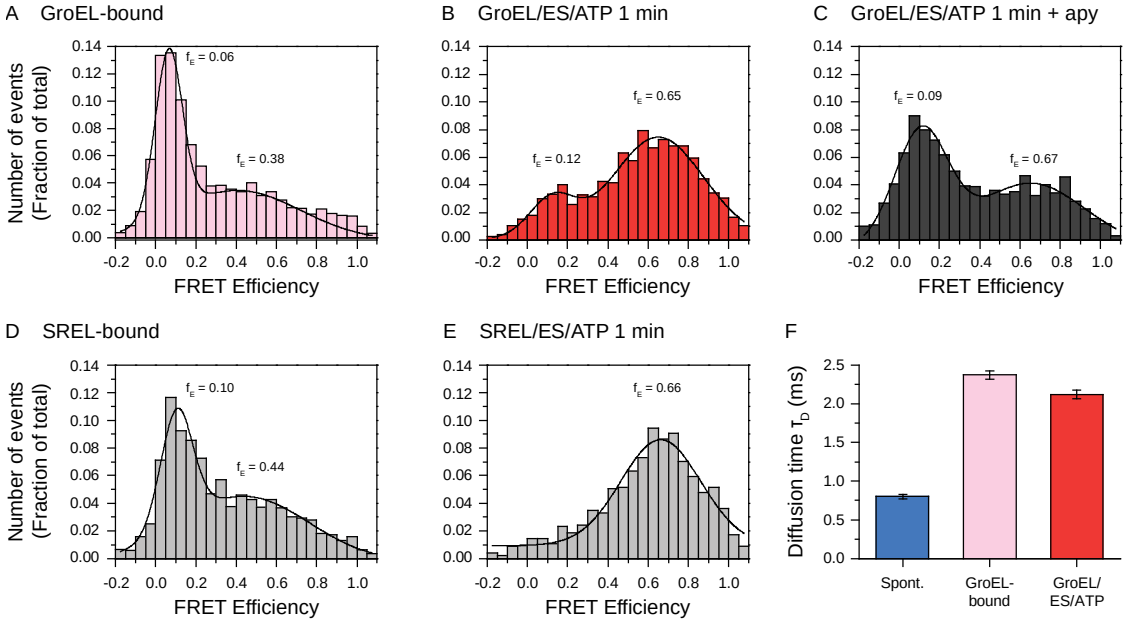

Figure 4.12 Substrate refolding occurs inside the GroEL central cavity

(A) smFRET histogram of GroEL bound DM-MBP (DL). DM-MBP (DL) was denatured in 6 M GuHCl, 10 mM DTT and diluted 200-fold into MBP refolding buffer containing 2 µM GroEL to a final protein concentration of 100 pM. The sample was immediately subjected to smFRET analysis. One representative histogram of three independent experiments is shown. **(B)** smFRET distribution of DM-MBP (DL) during the first minute of GroEL assisted refolding. DM-MBP (DL) was denatured in 6 M GuHCl, 10 mM DTT and diluted 200-fold into MBP refolding buffer containing 2 µM GroEL. The folding reaction was initiated by addition of 4 µM GroES and 5 mM ATP. smFRET was recorded during the first minute of refolding. The experiment was repeated so as to acquire statistically relevant data from at least 1000 particles. The FRET efficiency values for all particles were histogrammed and analyzed by Gaussian fit. **(C)** Refolding of DM-MBP (DL) was started as in (B) but stopped after 1 minute by addition of 10 U apyrase. smFRET was subsequently measured for 30 min. One representative histogram of three independent experiments is shown. **(D)** smFRET histogram of SREL bound DM-MBP (DL). DM-MBP (DL) was denatured in 10 M Urea, 10 mM DTT and diluted 200-fold into SREL buffer containing 1 µM SREL to a final protein concentration of 100 pM. The sample was immediately subjected to smFRET analysis. One representative histogram of three independent experiments is shown. **(E)** smFRET distribution of DM-MBP (DL) during the first minute of SREL assisted refolding. DM-MBP (DL) was denatured in 10 M Urea, 10 mM DTT and diluted 200-fold into SREL buffer containing 1 µM SREL. The folding reaction was initiated by addition of 4 µM GroES and 5 mM ATP. smFRET was recorded during the first minute of refolding. The experiment was repeated so as to acquire statistically relevant data from at least 1000 particles. The FRET efficiency values for all particles were histogrammed and analyzed by Gaussian fit. **(F)** Average diffusion time of DM-MBP (DL) measured at 100 pM in MBP refolding buffer for 1 min. DM-MBP (DL) was denatured in 6 M GuHCl, 10 mM DTT and diluted 200-fold into MBP refolding buffer (spont.) or MBP refolding buffer containing 2 µM GroEL (GroEL-bound). The assisted folding reaction was also initiated by addition of 4 µM GroES and 5 mM ATP (GroEL/ES/ATP). FCS was recorded during the first minute of refolding. Average diffusion times were determined by fitting of the autocorrelated data for Atto647N fluorescence and are represented as arithmetic mean ± s.d. of at least three independent experiments.

Next, we set out to quantify the amount of time one given substrate molecule spends in the encapsulated state as compared to being bound to the GroEL apical domains. To this end, we measured the initial ATPase activity of GroEL in presence of varying amounts of non-native substrate protein (DM-MBP). GroEL hydrolyzed ATP at a rate of ~53 ATP min^{-1} at 20°C. The hydrolysis rate decreased to ~21 ATP min^{-1} in presence of GroES (Chandrasekhar et al., 1986) (Fig 4.13 A). For constant concentrations of GroEL (0.2 µM) and GroES (0.4 µM) we found an increase in ATPase activity with increasing substrate concentration (Fig 4.13 B). It has been reported that substrate protein can stimulate the GroEL ATPase function (Martin et al., 1991) by triggering release of ADP and GroES from the trans ring (Hayer-Hartl et al., 1995; Martin et al., 1993; Ye and Lorimer, 2013). At a substrate concentration of 0.8 µM, corresponding to a 4-fold excess of substrate over GroEL, we found the ATPase activity reaching a saturated ~3-fold stimulation (Fig. 4.13 A and B) to ~59 ATP min^{-1}. Accordingly, the time it takes for one GroEL complex to hydrolyze 7 ATP molecules (the GroEL hemi-cycle) at 20°C was ~7 seconds in presence of saturating amounts of non-native substrate (1 µM) (Fig. 4.13 C). As established by stopped flow mixing studies, binding of non-native DM-MBP to GroEL is complete after ~0.3 s and substrate encapsulation upon binding of GroES after ~0.5 s (Sharma et al., 2008). Therefore, the substrate spends ~1 s in the GroEL bound state and ~6 s inside the central GroEL/ES cavity, corresponding to ~14% and ~86% of the hemi cycle duration, respectively. These values are in excellent agreement with those obtained in smFRET measurements (Fig. 4.12). We also measured the hemi-cycle length at a physiological temperature of 37°C and found it to be ~2 seconds (Fig. 4.13 C). This change in hemi-cycle length corresponds to a Q10 temperature coefficient of ~2 and therefore shows the Arrhenius-like temperature dependence of the GroEL chaperonin cycle, suggesting that all steps of the chaperonin mechanism undergo similar temperature dependent acceleration.

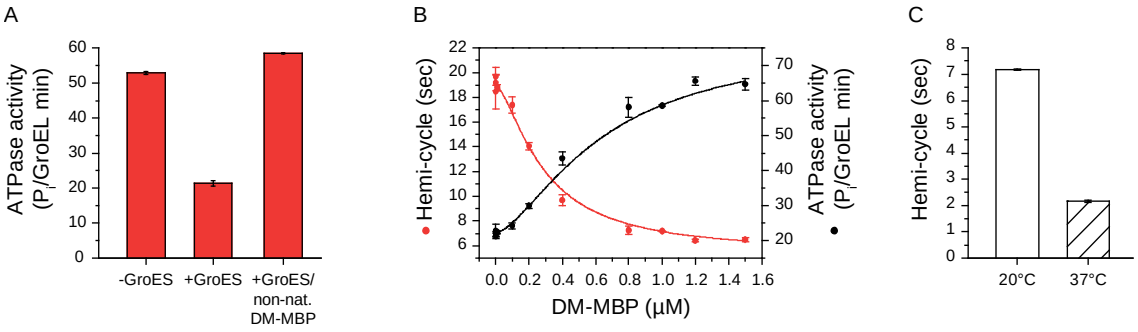

Figure 4.13 Presence of substrate stimulates GroEL ATPase activity

(A) ATPase avtivity of GroEL alone and in presence of GroES and denatured DM-MBP. ATPase activity of GroEL (0.2 µM) was measured in MBP refolding buffer at 20°C in absence or presence of GroES (0.4 µM) or GroES (0.4 µM)/non-native DM-MPB (1 µM). Rates are represented as arithmetic mean ± s.d. of at least three

independent experiments. **(B)** Substrate dependence of GroEL ATPase activity. ATPase of 0.2 µM GroEL was measured in presence of 0.4 µM GroES as a function of non-native substrate (DM-MBP) concentration. Rates are represented as arithmetic mean ± s.d. of at least three independent experiments. A sigmoidal fit was applied to guide the eye. **(C)** Average GroEL hemi-cycle duration ± s.d. in presence of substrate. The duration of a hemi-cycle was defined as the time needed for hydrolysis of seven ATP molecules per GroEL 14-mer and calculated from ATPase rate measurements. ATPase rates were obtained from three individual experiments for 0.2 µM GroEL in presence of 0.4 µM GroES and 1 µM non-native substrate (DM-MBP) at 20°C and 37°C. All ATPase activities were measured photometrically using a NADH coupled enzymatic assay.

4.6 GroEL cage charges strongly impact assisted refolding

The inner surface of the GroEL cage in the GroES bound cis conformation has a high negative net charge of -42. Two patches of three negatively charged amino acids in each GroEL subunit (E252, D253, E255 and D359, D361, E363) form two ring like charge clusters along the cis-cavity wall (Tang et al., 2006). Although these residues have a high conservation score amongst GroEL homologues, they do not play an important role in binding of substrate or GroES. It has therefore been suggested, and demonstrated, that these charge clusters play an important role in substrate refolding by potentially altering the chemical microenvironment inside the GroEL cavity (Sharma et al., 2008; Tang et al., 2006). In addition, it was shown that GroEL substrate proteins are enriched for negative charges as compared to the bulk proteome, suggesting a potential effect on protein folding by charge-charge repulsion (Kerner et al., 2005). Further, computational studies have suggested an impact of charge clusters in GroEL on structuring of water molecules inside the cis-cavity (England et al., 2008). It has been shown that inversion of one charge cluster in SREL (D359K, D361K, E363K) results in a net neutrally charged mutant SR(KKK2) that cannot accelerate refolding of DM-MBP and RuBisCO, but has no impact on refolding of Rhodanese (Tang et al., 2006). In addition, DM-MBP and RuBisCO show enhanced folding kinetics in presence of wild-type SREL/ES, while Rhodanese does not (Brinker et al., 2001; Tang et al., 2006). Taken together, these findings suggest an important role of negative charges in folding rate acceleration. Here, we further analyzed the influence of the GroEL cage net negative charge on the refolding of an encapsulated substrate protein using the SR(KKK2) and EL(KKK2) mutants.

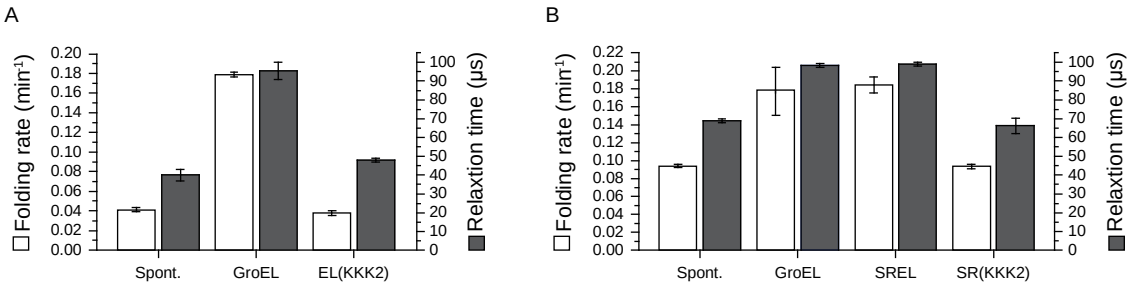

Figure 4.14 Effect of GroEL cavity surface charges on DM-MBP refolding kinetics and conformational dynamics

(A) EL(KKK2) does not accelerate DM-MBP refolding. Unlabeled DM-MBP (312C) was denatured in 6 M GuHCl, 10 mM DTT for 1 h at 20°C and diluted 200-fold at 20°C into MBP refolding buffer (200 mM KCl) to a final protein concentration of 100 nM for spontaneous refolding. For assisted refolding, denatured DM-MBP was diluted into MBP refolding buffer containing either 2 µM GroEL WT or 2 µM EL(KKK2). Assisted refolding was started by addition of 4 µM GroES and 5 mM ATP. Refolding kinetics were followed by Trp fluorescence increase and average refolding rates were extracted by single exponential fitting of three independent repeats. Relaxation time measurements during the first minute of spontaneous and assisted refolding, were performed by initiating the refolding of denatured, Atto655 labeled DM-MBP (312C) as described for PET-FCS based rate measurements at a final concentration of 1 nM. Relaxation times were extracted from data obtained during the first minute of refolding. **(B)** SR(KKK2) does not accelerate DM-MBP refolding. Unlabeled DM-MBP (312C) was denatured in 10 M Urea, 10 mM DTT for 1 h at 50°C and diluted 200-fold at 20°C into SREL buffer (20 mM KCl) to a final protein concentration of 100 nM for spontaneous refolding. For assisted refolding, denatured DM-MBP was diluted into SREL buffer containing either 2 µM GroEL WT, 1 µM SREL or 1 µM SR(KKK2). Assisted refolding was started by addition of 4 µM GroES and 5 mM ATP. Refolding kinetics were followed by Trp fluorescence increase and average refolding rates were extracted by single exponential fitting of three independent repeats. Relaxation time measurements during the first minute of spontaneous and assisted refolding, were performed by initiating the refolding of denatured, Atto655 labeled DM-MBP (312C) in SREL buffer as described for PET-FCS based rate measurements at a final concentration of 1 nM. Relaxation times were extracted from data obtained during the first minute of refolding.

At physiological salt concentration, GroEL accelerated the refolding of DM-MBP(312C) by ~4.5-fold (Fig 4.14 A). In contrast, no rate acceleration was observed with EL(KKK2)/ES (Fig. 4.14 A). Accordingly, EL(KKK2) did not restrict DM-MBP chain dynamics as measured during the first minute of folding by PET-FCS (Fig. 4.14 A). DM-MBP(DL), when bound to SR(KKK2) or EL(KKK2), displayed the same conformational properties as when bound to GroEL, as demonstrated by smFRET measurements (Fig. 4.4 B and Fig. 4.15 A and D). Moreover, during the first minute of folding with EL(KKK2)/ES/ATP, the diffusion time of DM-MBP(DL) was identical to that of the EL(KKK2)-bound protein (Fig. 4.15 F), indicating that essentially all substrate protein was chaperonin associated. The fraction of bound and encapsulated substrate determined from smFRET histograms that were recorded during the first minute of folding, was ~16% and ~84%, respectively, close to the values obtained with GroEL/ES (Fig. 4.15 B). The ATPase activity of EL(KKK2) was similar to that of GroEL and was

efficiently inhibited by GroES (Fig. 4.16 C). However, unlike GroEL, excess non-native DM-MBP had only a minor effect in stimulating the ATPase activity of EL(KKK2)/ES. These results suggested that the charge properties of the cis-cavity wall may on the one hand entropically stabilize encapsulated substrate protein, and on the other hand couple the presence of substrate to the ATPase activity of GroEL.

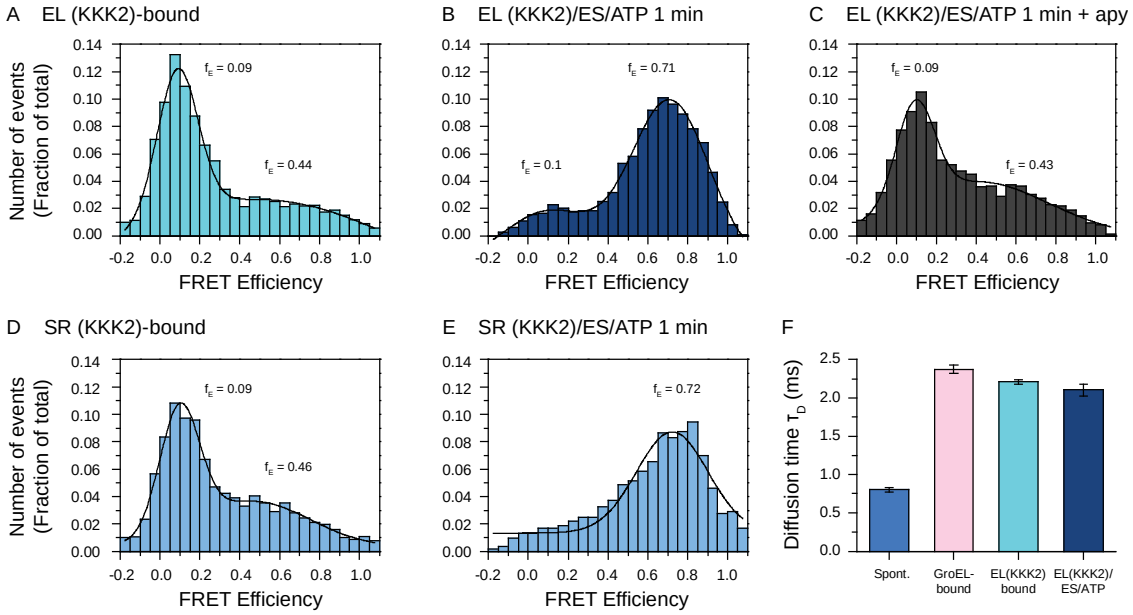

Figure 4.15 Slow assisted refolding by EL(KKK2) occurs inside the central cavity

(A) smFRET histogram of EL(KKK2) bound DM-MBP (DL). DM-MBP (DL) was denatured in 6 M GuHCl, 10 mM DTT and diluted 200-fold into MBP refolding buffer containing 2 µM EL(KKK2) to a final protein concentration of 100 pM. The sample was immediately subjected to smFRET analysis. One representative histogram of three independent experiments is shown. **(B)** smFRET distribution of DM-MBP (DL) during the first minute of EL(KKK2) assisted refolding. DM-MBP (DL) was denatured in 6 M GuHCl, 10 mM DTT and diluted 200-fold into MBP refolding buffer containing 2 µM EL(KKK2). The folding reaction was initiated by addition of 4 µM GroES and 5 mM ATP. smFRET was recorded during the first minute of refolding. The experiment was repeated so as to acquire statistically relevant data from at least 1000 particles. The FRET efficiency values for all particles were histogrammed and analyzed by Gaussian fit. **(C)** Refolding of DM-MBP (DL) was started as in (B) but stopped after 1 minute by addition of 10 U apyrase. smFRET was subsequently measured for 30 min. One representative histogram of three independent experiments is shown. **(D)** smFRET histogram of SR(KKK2) bound DM-MBP (DL). DM-MBP (DL) was denatured in 10 M Urea, 10 mM DTT and diluted 200-fold into SREL buffer containing 1 µM SR(KKK2) to a final protein concentration of 100 pM. The sample was immediately subjected to smFRET analysis. One representative histogram of three independent experiments is shown. **(E)** smFRET distribution of DM-MBP (DL) during the first minute of SR(KKK2) assisted refolding. DM-MBP (DL) was denatured in 10 M Urea, 10 mM DTT and diluted 200-fold into SREL buffer containing 1 µM SR(KKK2). The folding reaction was initiated by addition of 4 µM GroES and 5 mM ATP. smFRET was recorded during the first minute of refolding. The experiment was repeated so as to acquire statistically relevant data from at least 1000 particles. The FRET efficiency values for all particles were

histogrammed and analyzed by Gaussian fit. **(F)** Average diffusion time of DM-MBP (DL) measured at 100 pM in MBP refolding buffer for 1 min. DM-MBP (DL) was denatured in 6 M GuHCl, 10 mM DTT and diluted 200-fold into MBP refolding buffer (spont.) or MBP refolding buffer containing either 2 µM GroEL (GroEL-bound) or 2 µM EL(KKK2) (EL(KKK2) bound). The assisted folding reaction was also initiated by addition of 4 µM GroES and 5 mM ATP (GroEL/ES/ATP). FCS was recorded during the first minute of refolding. Average diffusion times were determined by fitting of the autocorrelated data for Atto647N fluorescence and are represented as arithmetic mean ± s.d. of at least three independent experiments.

To separately investigate the effect of GroEL negative cage charges on a folding substrate under non-cycling conditions, we next used SR(KKK2) to analyze the chain dynamics and folding kinetics of DM-MBP during folding when stably encapsulated. We established, again by size exclusion chromatography, that GroES-mediated substrate encapsulation by SR(KKK2) at low salt (Fig 4.16 A and B) was as efficient as with SREL (Fig. 4.10 A and B). In addition, using single molecule FRET, we established that during the first minute of encapsulation in SR(KKK2), DM-MBP(DL) populated compact conformations, as observed with SREL/ES (Fig. 4.15 E).

GroEL/ES and SREL/ES mediated the refolding of DM-MBP(312C) at essentially the same accelerated rate (measured at low salt) (Fig. 4.14 B), whereas folding by SR(KKK2)/ES was not accelerated beyond the spontaneous rate. Interestingly, while the DM-MBP folding rates in presence of GroEL/ES and SREL/ES are salt independent (Fig. 4.11 and Fig. 4.14), DM-MBP in presence of the KKK2 mutant displays a similar salt-dependence of the folding rate as during spontaneous renaturation (Fig. 4.11 and Fig. 4.14) (Chakraborty et al., 2010). Importantly, DM-MBP(Atto655) when stably encapsulated by SR(KKK2)/ES displayed significantly higher chain dynamics (τ_R 66 ± 4 µs) as compared to SREL/ES (τ_R 99 ± 1 µs) (Fig. 4.14 B). Together, these findings indicate that the net-negative charge of the GroEL cis-cavity plays a critical role in conformational restriction of dynamic folding intermediates of the encapsulated substrate, thereby accelerating their conversion to the native state.

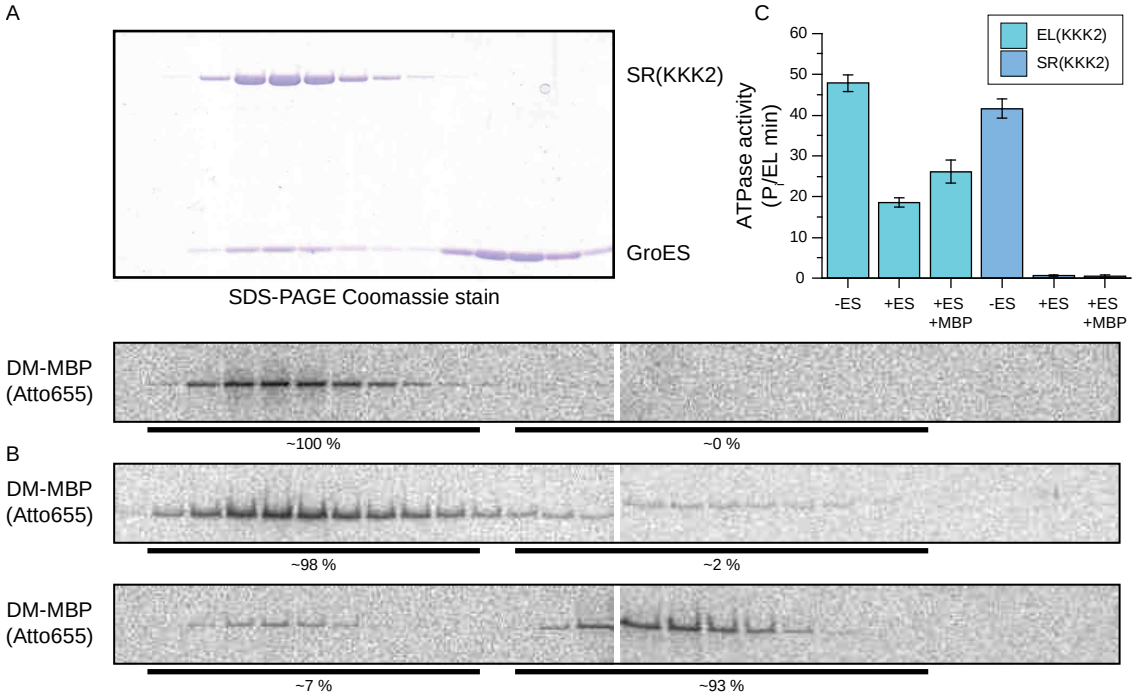

Figure 4.16 SR(KKK2) forms a stable cis-like complex with GroES that stably encapsulates DM-MBP

(A) SDS-PAGE of size exclusion fractionation of a preformed SR(KKK2), GroES, DM-MBP complex. DM-MBP (Atto655) was unfolded in 10 M urea, 10 mM DTT for 1 h at 50°C and diluted 200-fold into SREL buffer at 20°C containing 1 µM SR(KKK2). After 5 min encapsulation of DM-MBP was initiated by addition of 4 µM GroES and 1 mM ATP. The reaction mix was applied to a Superdex 200 gel filtration column equilibrated in SREL buffer, 50 mM urea, 1 mM ATP. Fractions of 50 µL were collected over a period of 30 min. Top panel shows Coomassie stained SDS PAGE of collected fractions. Bottom panel shows fluorescence scan of the same SDS-PAGE including densitometric quantification of free and complexed DM-MBP. **(B)** Size exclusion experiment to establish long term complex stability. The preformed SR(KKK2), DM-MBP and GroES complex was incubated for 30 min at 20°C, followed by size exclusion chromatography without further additions (top panel) or subsequent to addition of 50 mM CDTA, 200 mM KCL, 70 mM GuHCl to induce dissociation of the SR(KKK2) GroES complex and release encapsulated DM-MBP (bottom panel). The collected fractions were analyzed by SDS-PAGE as described in (A). **(C)** GroES efficiently binds EL(KKK2) and SR(KKK2) as established by ATPase activity. ATPase activity of 200 nM EL(KKK2) or 100 nM SR(KKK2) in MBP refolding buffer or SREL buffer was measured at 20°C by a coupled enzymatic assay following the photometric conversion of NADH, in absence and presence of 400 nM GroES or 400 nM GroES and 1 µM denatured DM-MBP. The rate of NADH consumption was converted into the production of phosphate by one GroEL 14-mer or one SREL 7-mer per minute. Data for figures A and B kindly provided by Goran Milićič.

4.7 DapA: A natural substrate of GroEL/ES

While DM-MBP is a frequently used GroEL model substrate, a large scale investigation of folding rate acceleration for natural and obligate GroEL substrates (class III substrates) (Fujiwara et al., 2010; Kerner et al., 2005) remains elusive. One important obstacle is the high aggregation propensity of many class III proteins. In order to compare folding in the absence and presence of GroEL, permissive conditions for spontaneous folding need to be identified, i.e. conditions under which aggregation is limited and folding to the native state is energetically favorable. Here, we identified the homotetrameric class III enzyme Dihydrodipicolinate synthase (DapA, 31.2 kDa) as a protein for which efficient spontaneous refolding to high yield at observable rate can be followed by enzymatic activity at temperatures ≤ 25°C. Similar to many class III proteins, DapA contains an N-terminal $(\beta\alpha)_8$ triose phosphate isomerase (TIM)-barrel fold (Fujiwara et al., 2010; Kerner et al., 2005) with an alpha helical C-terminal extension (Fig. 4.17 B). DapA is therefore an appropriate natural substrate protein with which to further investigate the mechanism of GroEL assisted protein folding.

4.7.1 DapA refolding is accelerated in the presence of GroEL/ES

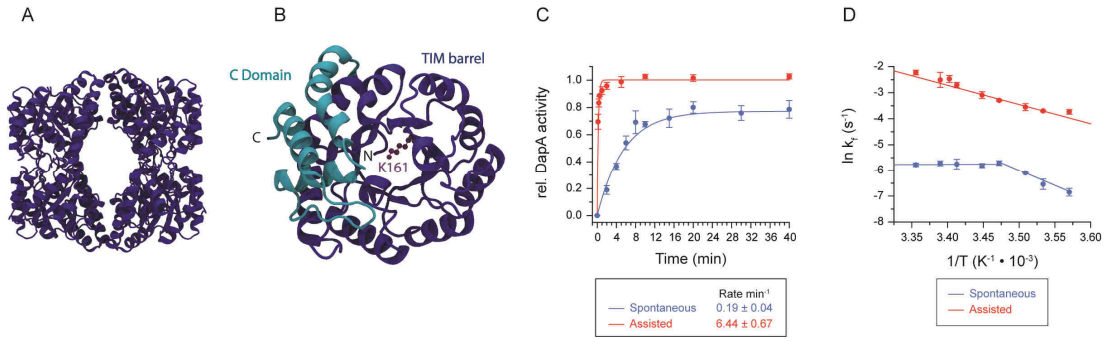

Figure 4.17 Acceleration of DapA refolding by GroEL/ES/ATP

(A) Structure of the DapA tetramer (PDB: 1DHP). **(B)** Magnification of one DapA subunit showing the TIM barrel domain in dark blue and the C-terminal α-helical domain in cyan as well as the active site lysine K161 in purple. **(C)** Spontaneous (blue) and assisted (red) subunit refolding of DapA analyzed by enzymatic activity at a final concentration of 200 nM at 25°C. DapA was denatured in 7.2 M GuHCl, 10 mM DTT and diluted 100-fold into DapA refolding buffer. In case of assisted refolding 2 µM GroEL and 4 µM GroES were used and refolding was started by addition of 5 mM ATP. Spontaneous refolding was stopped by addition of 0.8 µM GroEL D87K. Assisted refolding was stopped by addition of 50 mM CDTA. Refolding reactions were incubated for 1 h at 25°C to allow efficient assembly of native folded subunits. Arithmetic mean ± s.d. from at least 3 independent experiments is shown. **(D)** Arrhenius plot of spontaneous (blue) and assisted (red) refolding measured as in (C) for temperatures between 7°C and 25°C. Data for Figures (C) and (D) was kindly provided by Kristina Popova.

Unlike monomeric DM-MBP, DapA refolding requires folding of single subunits and subsequent tetrameric assembly (Fig. 4.17 A). When DapA was chemically denatured in 7.2 M GuHCl, 10 mM DTT and refolded by 100-fold dilution into DapA refolding buffer (20 mM Tris-HCl, pH 7.5, 100 mM KCl, 10 mM $MgCl_2$, 10 mM pyruvate) at 25°C, efficient refolding to a yield of ~75% was observed with a rate of ~0.2 min^{-1} ($t_{½}$ ~3.6 min) (Fig. 4.17 C). However, when refolding of DapA was performed in presence of GroEL/ES and ATP, 100% yield was obtained, as well as a strong ~30-fold acceleration of the refolding rate to ~6.0 min^{-1} ($t_{½}$ ~7 s), indicating high refolding yields after only a few chaperonin cycles. With decreasing temperature the rate of GroEL assisted DapA refolding decreased, showing conventional Arrhenius behavior (Fig. 4.17 D, red). In contrast, spontaneous refolding was temperature independent between 25°C and 15°C (Fig. 4.17 D, blue), indicating a kinetically trapped folding intermediate with a high entropic barrier to the native state (Bicout and Szabo, 2000; Dobson et al., 1998; Matagne et al., 2000), as described for DM-MBP (Chakraborty et al., 2010). At temperatures below 15°C the rate of spontaneous folding decreased with a constant slope, indicating the contribution of an enthalpic component to the transition state at lower temperatures (Dobson et al., 1998; Oliveberg et al., 1995). We decided to use a variety of biophysical approaches to further investigate the formation of a potential folding intermediate formed by DapA during spontaneous folding.

4.7.2 DapA forms a kinetically trapped folding intermediate

In order to better characterize the DapA folding intermediate we tested the bis-ANS binding capacity of DapA during refolding. Upon a change in the dielectric nature of the solvent, e.g. binding to hydrophobic regions in proteins, bis-ANS shows a significant increase in fluorescence at 485 nm (Hawe et al., 2008). While denatured and native DapA bound only minor amounts of bis-ANS, we observed strong bis-ANS fluorescence when denatured DapA was diluted into DapA refolding buffer at 10°C, containing the fluorescent probe (Fig. 4.18 A). We concluded that DapA forms a molten globule like intermediate state with little tertiary structure that allows binding of bis-ANS to exposed hydrophobic regions. Interestingly, the binding of bis-ANS decreased with a similar rate (0.07 min^{-1}) (Fig. 4.18 B) as was observed for recovery of enzymatic activity at 10°C (0.09 min^{-1}) (Fig. 4.17 D). Therefore, refolding of DapA to the native state correlates with the disappearance of the bis-ANS-binding intermediate state.

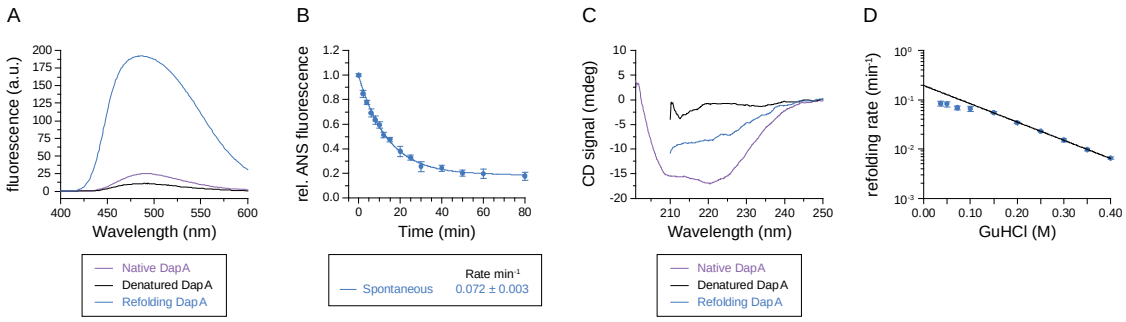

Figure 4.18 DapA forms a kinetically trapped folding imtermediate

(A) bis-ANS fluorescence spectra upon binding to different conformational states of DapA. bis-ANS fluorescence spectra in presence of the denatured state of DapA (in 7.2 M GuHCl, black), the native state (purple) and the intermediate state (blue), formed upon 100-fold dilution from denaturant, were recorded with an excitation wavelength of 390 nm. Baseline spectra of ANS in buffer or GuHCl were subtracted. Exemplary curves are shown. **(B)** Time dependent decrease of bis-ANS binding to DapA during refolding. DapA was denatured in 7.2 M GuHCl. Refolding was initiated by 100-fold dilution into DapA refolding buffer to a final concentration of 200 nM. After different kinetic points, bis-ANS was added to a final concentration of 1 µM and fluorescence spectra were recorded immediately. Fluorescence at 485 nm was used for quantification. The fluorescence observed immediately after mixing of denatured DapA into bis-ANS containing buffer was set as 1. The resulting data was fitted to a single exponential rate. An average of three independent experiments is shown. **(C)** Secondary structure of different DapA conformational states. Far-UV CD spectrum of DapA (purple) was recorded at a concentration of 2 µM in DapA refolding buffer at 10°C. DapA was unfolded in 7.2 M GuHCl, 10 mM DTT for 1 h at 20°C. The CD spectrum of denatured DapA (black) was recorded in presence of 7.2 M GuHCl. Denatured DapA was diluted 100-fold into DapA refolding buffer at 10°C. A CD spectrum of the initial folding intermediate (blue) was recorded immediately. Buffer and GuHCl spectra were subtracted. Due to a high GuHCl background signal, denatured and intermediate state could not be measured beyond 210 nm. **(D)** Spontaneous refolding of DapA at different concentrations of residual GuHCl. DapA was denatured in 7.2 M GuHCl, 10 mM DTT for 1 h at 20°C. Refolding was induced upon dilution into DapA refolding buffer to different final GuHCl concentrations. Refolding of 2 µM DapA was followed by the decrease in CD signal at 225 nm. The resulting data was fitted with a single exponential rate equation to extract folding rates. The resulting refolding rates were plotted against the GuHCl concentration. Data obtained between 150 mM and 400 mM GuHCl was approximated with a linear fit to guide the eye.

Furthermore, we measured CD spectra of DapA immediately after dilution from denaturant into DapA refolding buffer at 10°C (Fig. 4.18 C). We observed the initial formation of only ~20% of secondary structure elements, indicating high flexibility of the kinetically trapped folding intermediate. Stopped-flow CD experiments suggested that the formation of the initial secondary structure occurred within the dead-time of the instrument (low ms time-range, data not shown) and therefore likely reflects backbone collapse resulting in formation of dynamic α-helices, an early event described in protein folding studies (Teufel et al., 2011). As acquisition of the remaining secondary structure content kinetically reflected folding to the native state, we used time resolved CD-spectroscopy to measure refolding of DapA at

varying concentrations of chaotrope (GuHCl) (Fig. 4.18 D). The refolding arm of the resulting chevron plot showed a characteristic rollover at low GuHCl concentrations, characteristic of the rate-limiting formation of a kinetically trapped folding-intermediate (Kaya and Chan, 2003). To further exclude the transient aggregation of a DapA folding intermediate as a cause for slow spontaneous refolding, we again resorted to single molecule fluorescence techniques.

4.7.3 DapA does not form transient aggregates during refolding at single molecule level

In order to investigate folding of DapA under single molecule conditions, where aggregation is unlikely (Mukhopadhyay et al., 2007), we first established a mutant in which all surface accessible cysteine residues in DapA were replaced by serine (C20S, C141S, C218S) and an additional cysteine was attached C-terminally at position 293. We fluorescently labeled this DapA variant (DapA (293C)) with two different fluorophores, Alexa647 and Dy530 using maleimide chemistry. First the unlabeled mutant was tested to be enzymatically active and able to refold to the native state. We observed efficient spontaneous refolding at 20°C with a rate of 0.04 min^{-1} (Fig. 4.19 A, blue). Assisted refolding of DapA(293C) was accelerated ~21-fold to a rate of 0.76 min^{-1} (Fig. 4.19 A, red). For the Alexa labeled protein, spontaneous and assisted refolding occurred at 0.014 min^{-1} and 0.89 min^{-1}, respectively (Fig. 4.20 D). Therefore the mutation of the intrinsic DapA Cys residues resulted in a reduction of refolding kinetics. The enzymatic activity and secondary structure of the protein (as measured by CD spectroscopy) was preserved, as well as the important rate acceleration of refolding in presence of chaperonin. The influence of fluorescent labeling, when compared to the mutagenesis, was marginal. Note also that Dy530 labeled DapA could spontaneously refold to the enzymatically active state with a high yield (data not shown).

We then used the two differently labeled populations of DapA (293C) to perform similar FCCS experiments as described for DM-MBP (Fig. 4.19 B). As a negative control we used the two fluorophores instead of native protein, to avoid artifacts from subunit mixing at low concentrations. As expected, no cross correlation signal was observed (Fig. 4.19 B, purple). When both DapA variants were unfolded together and diluted into DapA refolding buffer to a final concentration of 100 pM, again no cross correlation signal was observed (Fig. 4.19 B, blue). This absence of cross correlation signal clearly shows that DapA not only forms a monomeric folding intermediate but is also unable to assemble to the tetrameric state at such low concentration. As a positive control we refolded the differently labeled populations together at 200 nM final protein concentration (100 nM each) in presence of GroEL/ES/ATP. The refolded and assembled tetramer was then diluted to 100 pM (concentration of monomers) for cross correlation analysis. Interestingly, significant cross correlation signal was observed (Fig. 4.19 B, black), demonstrating that the tetrameric state is kinetically stable and does not readily disassemble upon dilution.

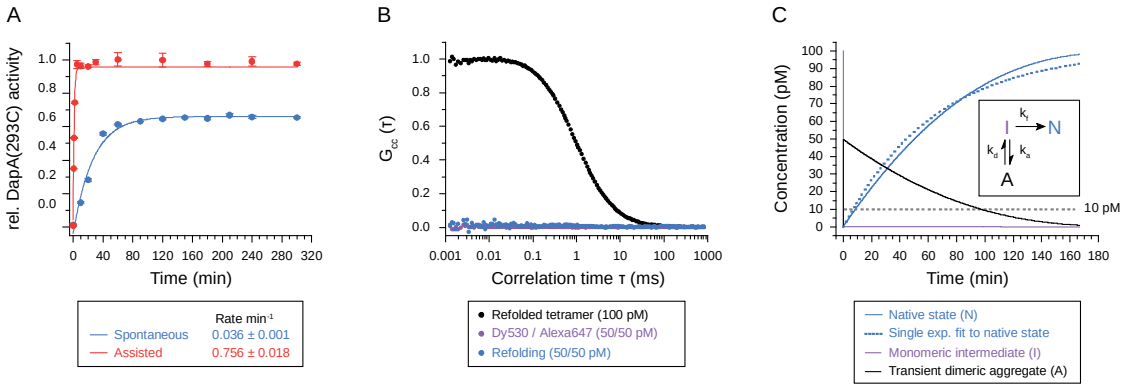

Figure 4.19 DapA does not form transient aggregates during refolding

(A) Spontaneous (blue) and assisted (red) subunit refolding of unlabeled DapA (293C) analyzed by enzymatic activity at a final concentration of 200 nM in DapA refolding buffer at 25°C. Unlabeled DapA (293C) was denatured in 7.2 M GuHCl, 10 mM DTT and diluted 200-fold into DapA refolding buffer. In case of assisted refolding 2 µM GroEL and 4 µM GroES were used and refolding was started by addition of 5 mM ATP. Spontaneous refolding was stopped by addition of 2 µM GroEL. Assisted refolding was stopped by addition of 50 mM CDTA. Refolding reactions were incubated for 1 h at 25°C to allow efficient assembly. Arithmetic mean ± s.d. from at least 3 independent experiments is shown **(B)** Absence of dcFCCS signal G_{cc} (τ) during spontaneous refolding of DapA. A 1:1 mixture of DapA (293C) labeled with either Dy530 or Alexa647 was denatured in 7.2 M GuHCl, 10 mM DTT and diluted 200-fold into DapA refolding buffer to a final concentration of 50 pM each. FCCS was recorded with pulsed interleaved excitation within the first 30 minutes of refolding (blue). As a positive control a 1:1 mixture of differently labeled DapA was refolded in presence of GroEL, GroES and ATP. The refolded and assembled protein was diluted to 100 pM for FCCS analysis (black). A mix of dye molecules again at 50 pM concentration each was used as a negative control (purple). **(C)** *In silico* kinetic simulation of the Anfinsen cage model including an off-pathway transient dimerization reaction (insert). The concentration of DapA was fixed to 100 pM. Variation of the equilibrium dissociation constant for the formation of dimeric aggregates (A, black) from monomeric intermediates (I, purple) resulted in apparently slower formation of native subunits (N, blue). The formation of native protein could not be fitted to a first order reaction (dotted blue line).

We also performed a kinetic simulation for a transient aggregation containing model of spontaneous DapA refolding using Berkeley Madonna (Chakraborty et al., 2010). The simulation was based on the same model as described for DM-MBP (Fig. 4.19 C). Due to the marked ~30 fold acceleration of folding in presence of GroEL, the simulation showed that the affinity of a transient, non-native, dimeric aggregate would have to be in the femtomolar range to explain the observed slow spontaneous refolding. Taken together, experimental data and simulation unequivocally rule out the existence of transient multimeric DapA aggregates during refolding at 100 pM to an extent that could explain rate acceleration by a passive, aggregation preventing mechanism of GroEL.

4.7.4 The rate of DapA subunit refolding is concentration independent

Having demonstrated the reproducibility and sensitivity of an FCS-based approach to measure folding rates at single molecule level using DM-MBP (Fig 4.6), we took advantage of this method to measure refolding for an authentic GroEL substrate. Initial experiments showed very different diffusion coefficients for GroEL-bound DapA (49 ±1 µm^2 s^{-1}) and spontaneously refolded DapA monomer (102 ±2 µm^2 s^{-1}).

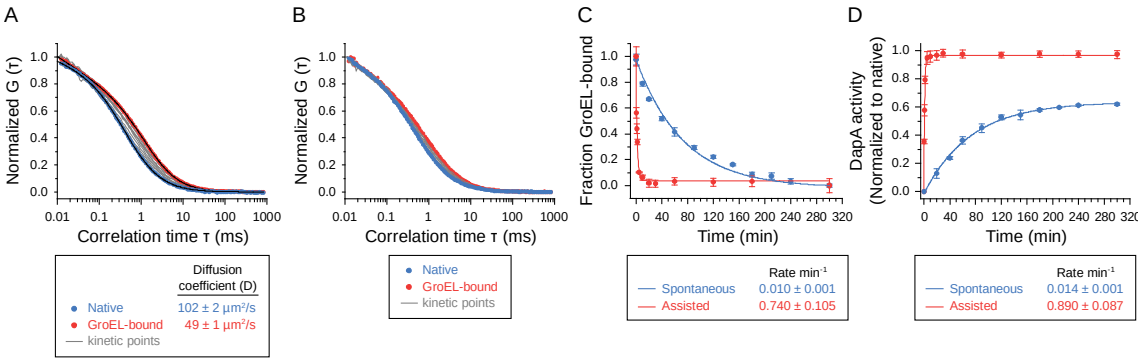

Figure 4.20 DapA subunit refolding is concentration independent

(A) Representative auto correlation curves of Alexa647 fluorescence for GroEL bound (red) and spontaneously refolded (blue) DapA-Alexa as well as kinetic points taken during spontaneous refolding (grey). DapA (293C) labeled with Alexa 647 was denatured in 7.2 M GuHCl, 10 mM DTT for 1 h at 20°C. Refolding was initiated at 20°C by 200-fold dilution into DapA refolding buffer. Spontaneous refolding was stopped at different kinetic points by addition of 2 µM GroEL. FCS was subsequently recorded for 10 min. For starting and final time points diffusion coefficients were calculated as arithmetic mean ± s.d. from three independent experiments. **(B)** Representative auto correlation curves of Alexa647 fluorescence for GroEL bound (red) and chaperonin refolded (blue) DapA-Alexa as well as kinetic points taken during assisted refolding (grey). DapA (293C) labeled with Alexa 647 was denatured in 7.2 M GuHCl, 10 mM DTT for 1 h at 20°C. Unfolded DapA was diluted 200-fold into DapA refolding buffer containing 2 µM GroEL. Assisted refolding was initiated by addition of 4 µM GroES and 5 mM ATP. Assisted refolding was stopped at different kinetic points by addition of 10 U apyrase. FCS was subsequently recorded for 10 min. **(C)** Refolding kinetics showing spontaneous and assisted refolding of DapA-Alexa as measured by the mean particle diffusion time through the confocal observation volume. The mean diffusion time was extracted from auto correlation data of Alexa647 fluorescence as shown in (A) and (B) and converted into the fraction of GroEL bound material, which was plotted versus refolding time. Single exponential fitting was used to extract the rate of folding. Arithmetic mean ± s.d. of three independent experiments is shown. **(D)** Spontaneous (blue) and assisted (red) subunit refolding of Alexa647 labeled DapA (293C) analyzed by enzymatic activity at a final concentration of 200 nM in DapA refolding buffer at 20°C. DapA was denatured in 7.2 M GuHCl, 10 mM DTT and diluted 200-fold into DapA refolding buffer. In case of assisted refolding 2 µM GroEL and 4 µM GroES were used and refolding was started by addition of 5 mM ATP. Spontaneous refolding was stopped by addition of 2 µM GroEL. Assisted refolding was stopped by addition of 10 U apyrase. Refolding reactions were incubated for 1 h at 25°C to allow efficient assembly. Arithmetic mean ± s.d. from at least 3 independent experiments is shown.

Following this observation, we unfolded DapA-Alexa in 7.2 M GuHCl and diluted it 100-fold into DapA refolding buffer to a final concentration of 100 pM. In order to distinguish between folded material and non-native protein, we added 2 µM GroEL at different time points. We again observed that during refolding, labeled DapA shifted to an average faster diffusion time (Fig. 4.20 A and B) as more particles reached the native state. While free DapA monomers showed an average diffusion coefficient of 102 ± 2 $\mu m^2 s^{-1}$, GroEL bound DapA showed an average diffusion coefficient of 49 ± 1 $\mu m^2 s^{-1}$. As we have demonstrated in cross correlation measurements, protein assembly upon refolding does not occur at such low concentration (Fig. 4.19 B), and refolding of DapA at 100 pM therefore results in folded monomeric subunits that are not GroEL associated. The average spontaneous refolding rate obtained by FCS was ~0.01 min^{-1} (Fig. 4.20 C). In accordance we performed assisted refolding of GroEL-bound DapA-Alexa at 100 pM by addition of 4 µM GroES and 5 mM ATP. The refolding reaction was stopped at different time points by addition of apyrase. Again, we observed a gradual shift from slow to fast diffusion rates. Analysis of the average diffusion rate for different kinetic points revealed a refolding rate of ~0.74 min^{-1}, and therefore an important folding acceleration of ~30-fold. The extracted rate constants were in excellent agreement to the rate constants obtained for the labeled protein by enzymatic activity at 200 nM (Fig. 4.20 D). Importantly, the spontaneous refolding rate of DapA thus appears to be constant over a concentration range of 3 orders of magnitude, and is therefore not limited by aggregation. This finding clearly rules out a passive cage mechanism for folding acceleration of DapA in the presence of the chaperonin system.

4.7.5 GroEL accelerates DapA refolding up to 130-fold at physiological temperature

Having confirmed the active cage model for DapA assisted folding, we decided to use the established single molecule based approaches to investigate DapA refolding at the physiological temperature of 37°C.

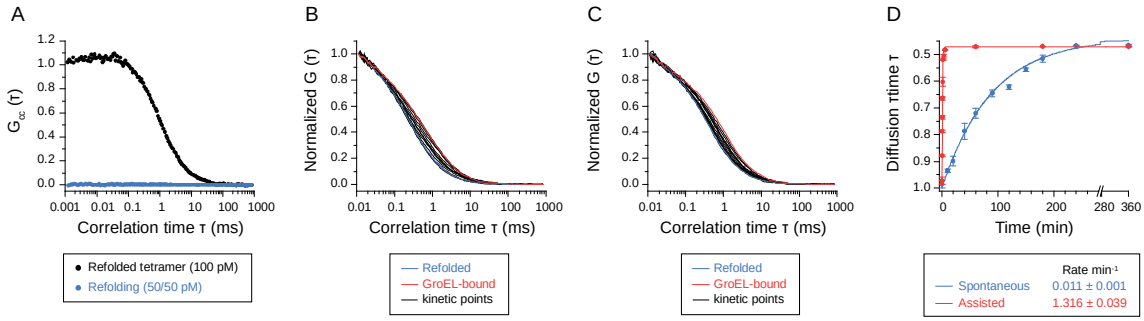

Figure 4.21 DapA refolding at physiological temperature

(A) Absence of dcFCCS signal G_{cc} (τ) during spontaneous refolding of DapA. A 1:1 mixture of DapA (293C) labeled with either Dy530 or Alexa647 was denatured in 7.2 M GuHCl, 10 mM DTT and diluted 200-fold into DapA refolding buffer (preincubated in at 37°C) to a final concentration of 50 pM each. FCCS was recorded in a temperature controlled cuvette with pulsed interleaved excitation within the first 30 minutes of refolding (blue). As a positive control a 1:1 mixture of differently labeled DapA was refolded in presence of GroEL, GroES and ATP also at 37°C. The refolded and assembled protein was diluted to 100 pM for FCCS analysis (black). **(B)** Representative auto correlation curves of Alexa647 fluorescence for GroEL bound (red) and spontaneously refolded (blue) DapA-Alexa as well as kinetic points taken during spontaneous refolding at 37°C (black). DapA (293C) labeled with Alexa 647 was denatured in 7.2 M GuHCl, 10 mM DTT for 1 h at 20°C. Refolding was initiated at 37°C by 200-fold dilution into DapA refolding buffer. Spontaneous refolding was stopped at different kinetic points by addition of 2 µM GroEL. FCS was subsequently recorded at room temperature for 10 min. **(C)** Representative auto correlation curves of Alexa647 fluorescence for GroEL bound (red) and chaperonin refolded (blue) DapA-Alexa as well as kinetic points taken during assisted refolding (black). DapA (293C) labeled with Alexa 647 was denatured in 7.2 M GuHCl, 10 mM DTT for 1 h at 20°C. Unfolded DapA was diluted 100-fold into DapA refolding buffer containing 2 µM GroEL. Assisted refolding 37°C was initiated by addition of 4 µM GroES and 5 mM ATP. Assisted refolding was stopped at different kinetic points by addition of 10 U apyrase. FCS was subsequently recorded for 10 min at room temperature. **(D)** Refolding kinetics showing spontaneous and assisted refolding of DapA-Alexa at 37°C, as measured by the mean particle diffusion time through the confocal observation volume. The mean diffusion time was extracted from auto correlation data of Alexa647 fluorescence as shown in (B) and (C) and plotted versus refolding time. Single exponential fitting was used to extract the rate of folding. Arithmetic mean ± s.d. of three independent experiments is shown.

At the high concentrations needed for enzymatic assays, DapA aggregated substantially at temperatures >25°C. At 100 pM however, we did not observe the formation of aggregated material even at 37°C, as tested by cross correlation of refolding, labeled DapA(293C) in a temperature controlled cuvette (Fig. 4.21 A). Therefore, we performed FCS based refolding assays at 37°C to measure the rate of both spontaneous and GroEL/ES assisted refolding. The Arrhenius behavior of DapA refolding (Fig. 4.17 D) showed that the spontaneous folding rate of DapA is temperature independent, whereas the assisted folding rate showed a Q_{10} temperature coefficient of ~2. As expected, we observed a strong acceleration of DapA refolding by GroEL/ES at 37°C. The assisted rate obtained at 37°C (1.32 min^{-1}) was ~130 times faster than the slow rate of spontaneous refolding (0.01 min^{-1}) (Fig. 4.21 D) The observation of a strong acceleration of substrate refolding by GroEL/ES at 37°C indicates the importance of the GroEL assisted refolding of an otherwise trapped intermediate state at physiological conditions, as well as the strong potential of GroEL as an active foldase in *E.coli*.

5 Discussion

5.1 Active versus passive cage model

The observation of accelerated folding of substrate proteins in the presence of GroEL/ES initiated a long standing debate as to how chaperonins could promote refolding of their cognate substrates (Ambrose et al., 2015; Apetri and Horwich, 2008; Brinker et al., 2001; Chakraborty et al., 2010; Horwich et al., 2009; Tang et al., 2006, 2008; Tyagi et al., 2011; Yang et al., 2013). Here, we set out to investigate the three prominent models of chaperonin assisted protein folding, and to test their validity using advanced spectroscopic methodology.

It has been proposed that accelerated folding of DM-MBP by GroEL/ES is the result of GroEL preventing reversible aggregation, a process that would otherwise slow the rate of spontaneous folding (Ambrose et al., 2015; Apetri and Horwich, 2008; Tyagi et al., 2011). In such a passive cage model, in-cage folding would occur at the same rate as spontaneous folding at infinite dilution, and GroEL would function solely as an anti-aggregation device (Horwich et al., 2009). To test this hypothesis, we investigated the oligomeric state of DM-MBP during refolding at single molecule level. In dcFCCS measurements we found DM-MBP to be monomeric during refolding and showed that even 10 pM of aggregated material would be resolvable by this approach. *In silico* kinetic simulation of the passive cage model would predict aggregated material to be present at much higher amounts than 10 pM. In addition, we could demonstrate with FCS that the average concentration of particles in the refolding mix is stable over time, inconsistent with a model including a reversible aggregation reaction.

Having established that DM-MBP is monomeric during refolding at 100 pM, we devised a novel single molecule FRET based approach to assess spontaneous and assisted refolding rates of DM-MBP at such low concentration. Based on the previous finding (Sharma et al., 2008) that a FRET labeled mutant of DM-MBP (D30C/A312C) shows distinct FRET spectra when bound to GroEL as compared to in free solution, we devised an approach to quantitatively analyze the amount of native and non-native molecules in solution. Addition of GroEL during spontaneous folding of DM-MBP, or apyrase during assisted folding, resulted in a rapid molecular sorting of native and non-native DM-MBP. Specifically

non-native molecules were efficiently recognized by GroEL and bound to the apical GroEL domains in a stretched (low-FRET) population. Note that native molecules would not be recognized by GroEL and remain in solution as a compact (high-FRET) conformer. Using the quantitative information from FRET histograms, we evaluated the number of native and non-native molecules at different kinetic points during refolding. We found that the folding rates measured by this approach at 100 pM are in excellent agreement with folding rates measured by ensemble fluorescence at 100 nM, i.e. three order of magnitude higher concentration. Importantly, adding GroEL/ES resulted in a ~5-fold acceleration over the spontaneous folding rate, in the absence of aggregation. Taken together these findings not only demonstrate folding catalysis under conditions where aggregation is excluded, they also show concentration independence of spontaneous folding of DM-MBP, inconsistent with the occurrence of transient aggregation (Apetri and Horwich, 2008; Chakraborty et al., 2010; Tyagi et al., 2011).

In a simplified approach based on diffusion measurements by FCS, also at 100 pM, we could confirm the folding rate measurements obtained by smFRET and in addition establish a universal strategy to measure refolding rates of GroEL substrate proteins at single molecule level without facing the potential bottleneck of finding a suitable FRET pair. Binding of non-native labeled protein to GroEL significantly shifts the diffusion rate of the labeled molecule to that of a high molecular weight complex. In this case DM-MBP, with a diffusion coefficient of ~160 $\mu m^2/s$ in solution, shifted to ~49 $\mu m^2/s$ when bound to GroEL. As substrate protein refolds to the native state, fewer molecules interact with GroEL and diffusion of labeled molecules occurs on average at a faster rate. The addition of GroEL at different kinetic points of spontaneous refolding, or apyrase during assisted refolding, resulted in a molecular sorting of native and non-native molecules. The obtained rates that were extracted by plotting the average diffusion times of molecules against the refolding time resulted in reproducible folding rates that were in excellent agreement to the data obtained by single molecule FRET analysis.

Importantly, explaining folding acceleration under the passive cage model requires two competing reactions: folding of a monomeric intermediate to the native state and transient aggregation of such an intermediate. Notably for the model to be consistent, the rate of aggregation must be slower than the rate of folding. Therefore any folding rate measurement should not be possible to be approximated by a single exponential fit, i.e. a two state reaction model (Sabelko et al., 1999). Interestingly, all measurements of spontaneous folding of DM-MBP obtained here and previously, could be fitted with a single exponential function (Apetri and Horwich, 2008; Chakraborty et al., 2010; Tang et al., 2006; Tyagi et al., 2011).

Furthermore, light scattering data, used previously to support transient aggregation of DM-MBP (Apetri and Horwich, 2008; Tyagi et al., 2011), is inconsistent with the passive cage model. If the light scattering signal shows formation of transient aggregates, it should decrease with a rate (k_d) that is faster

than the rate of folding (k_f). However, the observed scattering signal is constant over time and can therefore only show permanently aggregated material that would not further participate in productive folding and therefore not alter the rate of folding.

All in all our data strongly support an active cage model of chaperonin function for GroEL, with in cage folding occurring at a faster rate than spontaneous folding in free solution, at least for a subset of substrate proteins with a high entropic energy barrier to the native state, as proposed previously (Chakraborty et al., 2010).

5.2 Direct experimental evidence for conformational restriction

Substrate encapsulation according to the active cage model is not only necessary to prevent premature aggregation of unfolded protein chains, but also accelerates the folding process, adjusting it to the relative speed of translation and thereby preventing accumulation of unfolded or misfolded molecules in the bacterial cytosol (Brinker et al., 2001). The underlying mechanism was proposed to be linked to steric confinement and entropic destabilization of an otherwise flexible folding intermediate inside the GroEL cage (Chakraborty et al., 2010; Tang et al., 2006, 2008). Two separate structural features of the GroEL cage might be involved in conformational restriction of an encapsulated substrate, the cage volume in comparison to the size of the encapsulated substrate (i.e. the steric confinement proper) (Baumketner et al., 2003; Hayer-Hartl and Minton, 2006; Lucent et al., 2009; Sirur and Best, 2013; Tang et al., 2006, 2008) and the highly negatively charged cage wall that was proposed to increase the hydrophobic effect by ordering water structure (England and Pande, 2008; England et al., 2008; Tang et al., 2006, 2008).

To directly investigate the effect of encapsulation on a folding substrate molecule, we used a combination of photoinduced electron transfer (PET) and FCS to monitor substrate flexibility inside GroEL and in free solution, in real time (Neuweiler et al., 2009; Sauer and Neuweiler, 2014; Teufel et al., 2011). Interestingly, the quenching of the oxazine fluorophore by direct contact with Trp residues along the protein chain occurred at a rate of 44 ± 3 µs. Note that intra-chain diffusion processes as they would occur in a Gaussian-like unfolded protein chain usually occur at a timescale of ns (Krieger et al., 2003; Neuweiler et al., 2007).

We found that the rate of disappearance of the PET signal, i.e. the disappearance of the observed intermediate state, was in good agreement with the rate of folding obtained by conventional ensemble fluorescence spectroscopy. Thereby, we observed that PET-FCS allows assessment of accurate folding rates, by following the disappearance of the quenched conformer of the Atto655-labeled intermediate state. In conclusion, we established an approach that can correlate structural flexibility with folding of a protein.

Interestingly, we observed only moderate conformational restriction of the flexible folding intermediate upon binding to the apical domains of GroEL. Conformational entropy was however significantly reduced upon stable encapsulation inside the GroEL central cavity. In fact the relaxation time of the PET signal was ~2.5 fold reduced from 44 ± 3 µs in free solution to 99 ± 1 µs inside the GroEL cage. This indicates that the number of possible unfolded-like conformers of refolding DM-MBP was reduced by GroEL, decreasing intrinsic chain entropy and thereby improving the rate of native contact formation. Taken together, by direct observation of chain entropy, we have established direct evidence that the folding of DM-MBP is, as suggested previously (Chakraborty et al., 2010), rate limited by an entropic barrier and GroEL overcomes the entropic barrier by conformational restriction during encapsulation, enhancing the folding speed of DM-MBP. Interestingly, the relaxation time for DM-MBP was reduced to a similar extent during cycling with GroEL/ES as well as upon stable encapsulation in SREL. This indicates that the substrate protein spends the majority of time in the encapsulated state during the GroEL reaction cycle. In addition, gel filtration experiments showed clearly that substrate encapsulation by SREL was stable and did not reveal significant "escape" from the enclosed cavity, in contrast to a recent report (Motojima and Yoshida, 2010).

Since it has been shown that spontaneous DM-MBP refolding is salt-sensitive (Apetri and Horwich, 2008; Chakraborty et al., 2010; Tyagi et al., 2011), we decided to investigate the mechanism of salt influence on DM-MBP refolding using PET-FCS. We found, that under our conditions a 10-fold decrease in KCl concentration from 200 mM to 20 mM resulted in ~2-fold faster refolding of DM-MBP (312C). When we measured PET-FCS of Atto655 labeled DM-MBP (312C) under these conditions, we found the same rate acceleration at 1 nM protein concentration. In addition we performed smFRET-based refolding measurements at 100 pM that confirmed the rates obtained at higher concentrations. These observations rule out a concentration-dependent effect of chloride salt on transient oligomerization. In contrast, we found a direct influence of the KCl concentration on conformational flexibility of the DM-MBP intermediate state. We hypothesize that the difference in ionic strength of the refolding buffer directly impacts the electrostatic properties of the refolding protein (Song et al., 2007) and especially the hydrophobic effect that strongly depends on solvent structure (England and Pande, 2008; Kalra et al., 2001). Thereby charge repulsion/attraction between side-chains and hydrogen bonding of surrounding water molecules could be affected, such that the entropic folding barrier that otherwise frustrates the DM-MBP folding energy landscape might be reduced.

5.3 Substrate folding occurs inside the GroEL cage

The iterative annealing model of chaperonin function posits that iterative ATP dependent cycles of substrate binding, forced unfolding, and release by the GroEL apical domains followed by folding either

inside or outside the GroEL cage serve to unfold kinetically trapped intermediate states, affording them another chance to partition between productive and non-productive folding trajectories (Thirumalai and Lorimer, 2001; Yang et al., 2013).

In an experiment where we used SREL to measure assisted refolding rates, we were able to show that a single round of encapsulation inside the central GroEL cavity is sufficient to achieve full rate catalysis as well as full yield. Therefore iterative cycling of GroEL, despite being an intrinsic property of the system, is not a necessity to achieve productive folding of DM-MBP. However, we cannot rule out the existence of substrates that misfold during encapsulation and thus would benefit from iterative annealing and forced unfolding. Also, we note that in the presence of SREL, a single round of forced unfolding occurs and might contribute to accelerate folding inside the central cavity. Stretching of unfolded or misfolded substrate proteins, i.e. forced unfolding, has been described here and elsewhere (Kim et al., 2010; Lin et al., 2008; Sharma et al., 2008). Also, upon binding of GroES to the apical domains of GroEL, the substrate protein is sequentially released into the central cavity, with the less hydrophobic sequences being released first (Sharma et al., 2008). This might alter the intrinsic mechanism of folding and thereby help to avoid misfolded unproductive conformers.

In stark contrast to an active cage model, however, the iterative annealing model does not assign a specific function to GroEL for aggregation prevention or folding catalysis. Encapsulation in the iterative annealing model is not an active principle and it is further proposed that substrate folding may occur equally inside or outside the GroEL cage (Yang et al., 2013).

As our PET-FCS measurements initially suggested, using diffusion time measurements and single molecule FRET analysis, we found the substrate protein to be encapsulated ~80% of the time during a cycling reaction. The remainder of the cycle time the substrate is mostly bound to the GroEL apical domains. We measured the duration of one GroEL cycle at substrate saturation and found the time needed to hydrolyze 7 ATP molecules (the GroEL hemi cycle) to be ~7 s at 20°C. Considering that binding, stretching and release occur within less than 1 second (~14% of the cycle time), as established previously (Sharma et al., 2008), a given substrate would spend ~6 seconds (~86% of the cycle time) in the encapsulated state, values in excellent agreement with the smFRET data. Accordingly, the amount of substrate folding out of cage is insignificant, especially with re-binding of non-native substrate to GroEL being highly efficient and occurring at a fast rate (~0.3 s or less) (Sharma et al., 2008). Moreover, we did not observe a reduction in yield or rate even at a large excess of GroEL over substrate molecules, as in single molecule experiments (100 pM substrate, 2 µM GroEL).

It was argued that at 37°C the amount of time a substrate spends inside the GroEL cage becomes insignificant (Yang et al., 2013). We measured the GroEL hemi cycle duration at 37°C at substrate saturation and found it to be ~2 s. Assuming that all steps of the GroEL cycle undergo a similar

temperature dependent rate acceleration, the substrate would nevertheless spend the majority of time in the encapsulated state. In the light of many GroEL substrates undergoing highly temperature dependent aggregation reactions in absence of chaperonin (Calloni et al., 2012; Fujiwara et al., 2010; Kerner et al., 2005), the encapsulation of substrates in a nano-cage is an ideal mechanism to prevent irreversible aggregation. In addition, it was shown that when efficient recapture of substrate by GroEL is prevented under non-permissive conditions, such as high temperature, substrate refolding stops immediately (Brinker et al., 2001). Refolding of substrate proteins by GroEL/ES under physiological conditions must therefore occur in-cage.

5.4 The role of the net negatively charged GroEL cage wall

One of the most interesting features of the GroEL cage wall is its high negative net charge of -42. We used the GroEL mutants EL(KKK2) and SR(KKK2) (Tang et al., 2006), in which the net charge is reduced to 0, to investigate the influence of negative charges on substrate chain dynamics with PET-FCS. Note that substrate encapsulation inside SR(KKK2) was as efficient as in SREL. Most interestingly, the GroEL charge mutants were not able to restrict substrate chain entropy as we observed in case of GroEL WT. This effect correlated with the inability to accelerate refolding of DM-MBP over the spontaneous folding rate. Thus EL(KKK2) is a passive cage with respect to DM-MBP as a substrate. As SR(KKK2) showed the same effect, even though it also unfolds the substrate protein once, clearly annealing, forced unfolding and release do not contribute significantly to acceleration of substrate refolding. Note that irrespective of their inability to accelerate DM-MBP refolding, substrate refolding occurred in-cage, as shown by diffusion and smFRET measurements. Interestingly, the ATPase activity of EL(KKK2) was not stimulated in presence of substrate, indicating that the negative charges play a role in sensing the presence of substrate and linking it to the GroEL ATPase cycle. The exact mechanism how a highly charged cage wall affects protein folding remains to be investigated. It is possible, that Coulombian repulsion contributes to steric confinement, reducing entropic freedom of a net-negatively charged encapsulated chain. It was however suggested from *in silico* simulations, that the negative charges rather have an ordering effect on water molecules inside the chaperonin cage, making them unavailable for hydrogen bonding with the substrate (England et al., 2008). Thereby the substrate would face an increased hydrophobic effect, which in turn results in conformational restriction. The recent assessment of water dynamics inside the GroEL cage (Franck et al., 2014) was performed not only in absence of substrate protein, but also with the spin label being located on the surface of GroES, distant from the important charge clusters. The contribution of water structure inside the GroEL cavity to substrate folding therefore remains to be experimentally investigated.

It was shown that a deletion of the GroEL C-terminal tails, potentially decreasing the steric confinement effect due to an increased cage volume, results in slowed in-cage refolding of DM-MBP (Tang et al., 2006). Our data obtained with GroEL charge mutants suggests that the net-negatively charged GroEL cage wall is even more important for accelerated folding of DM-MBP than the steric confinement proper. Nevertheless, steric confinement, charge repulsion and ordered water structure, would likely all result in a restriction of conformational flexibility of the encapsulated substrate. To separately investigate the contribution of all factors requires further experimental studies.

Taken together, our results again indicate that the inner cage of GroEL is not only its most important feature but that it has evolved in a way to ensure folding on biologically relevant timescales with high yields. We propose that the steric confinement proper and the high negative net-charge of the cage wall cooperatively promote protein folding. The contribution of both factors strongly depends on the physiochemical properties of the substrate, especially on size and net-charge.

5.5 DapA as a natural substrate of GroEL

To further test the active cage model for a natural substrate of GroEL, we used DapA, a tetrameric TIM-barrel enzyme involved in cell wall and lysine biosynthesis (Dobson et al., 2005). GroEL interaction studies revealed a strong enrichment of TIM barrel-fold proteins amongst obligate GroEL substrates (Fujiwara et al., 2010; Kerner et al., 2005). What distinguishes GroEL substrates from other proteins and how GroEL promotes their folding is an unsolved problem (Azia et al., 2012; Gershenson and Gierasch, 2011; Jewett and Shea, 2010). Therefore DapA is an ideal substrate to investigate the active cage model on a physiologically relevant complex fold that contains many long-range contacts.

DapA refolding is strongly accelerated (~30 fold at 25°C) in the presence of GroEL/ES, as revealed by measuring refolding kinetics by recovery of enzymatic activity. The contribution of tetrameric assembly was excluded from activity assays by stopping the spontaneous refolding reaction by addition of GroEL or trap GroEL (D87K) and by chelating magnesium in assisted folding reactions. The refolding mix was given sufficient time for productive assembly before enzymatic activity was measured. We established that DapA, like DM-MBP, forms a kinetically trapped intermediate state with a high entropic barrier to the folded state and little initial secondary structure. The intermediate only slowly buried hydrophobic residues during spontaneous folding and showed a characteristic chevron rollover. It is likely that all proteins that are significantly accelerated in folding inside the GroEL cage form a flexible, kinetically trapped intermediate state with a high entropic energy barrier. The lack of stable secondary structure and high chain entropy might be a hallmark of such substrates. It is reasonable to assume that the formation of long range contacts within the primary structure of the protein are a

limiting factor during spontaneous folding and that conformational restriction overcomes this hurdle of efficient folding to the native conformation.

To again rule out transient aggregation as a cause for slow spontaneous folding, we created a DapA mutant with a single C-terminal cysteine residue for fluorescent labeling. The unlabeled mutant as well as the labeled protein showed enzymatic activity, and both proteins were able to productively refold after chemical denaturation. In dcFCCS measurements we could establish that at a concentration of 100 pM DapA was monomeric during refolding and was not assembly competent. When we measured spontaneous and assisted DapA refolding rates at 100 pM, using the FCS based approach established with DM-MBP, we found a significant ~75-fold rate acceleration. The measured rates were in excellent agreement with rates measured by enzymatic activity at 200 nM, ruling out concentration dependent transient aggregation as a cause for apparently slow spontaneous folding. GroEL therefore actively promotes folding of DapA. As working at low concentrations is permissive even at physiological temperature, we established the absence of DapA aggregation at 37°C and subsequently measured refolding rates by FCS. GroEL was able to accelerate DapA refolding ~130-fold at 37°C. At 37°C DapA could be refolded by GroEL within only a few chaperonin cycles. Folding is therefore faster than the rate of protein synthesis (~20 amino acids per second), demonstrating the physiological significance of an active cage to adjust the rate of folding to the rate of protein synthesis and cell division and thereby preventing accumulation of misfolded protein. This ensures a robust ability of the cell to adapt to environmental changes by rapid production of native protein.

In parallel to the study presented here, the folding pathways of DapA in solution and inside the GroEL were investigated by hydrogen deuterium exchange (HDX) experiments coupled to mass spectrometry (Georgescauld et al., 2014). Notably, in solution DapA folded with a high degree of cooperativity, i.e. most structural elements of the TIM-barrel domain acquired solvent protection simultaneously. In contrast, inside the GroEL/ES folding of the TIM-barrel domain proceeded in a sequential manner, consistent with a reduced entropic energy barrier.

5.6 Conclusion

In this study we provide strong evidence for an active chaperonin mechanism. We show that transient aggregation and therefore a passive cage model cannot explain the accelerated folding of DM-MBP or the natural GroEL substrate DapA. Furthermore, we provide direct experimental evidence for conformational restriction of dynamic folding intermediates inside the GroEL/ES cavity. We show that the negatively charged cage wall of GroEL plays an important role in accelerated substrate folding. Our study shows clearly that global encapsulation of a substrate inside the GroEL nano-cage is the working principle of an active chaperonin mechanism. As substrates spend the major amount of time in the

encapsulated state, out-of cage folding can be excluded as a working hypothesis for GroEL assisted protein folding. Our experimental data further suggests that iterative annealing and forced unfolding, if at all, contribute only to a minor extent to accelerated substrate refolding. The important ~130-fold rate acceleration of DapA assisted refolding at physiological temperature demonstrates the pivotal importance of GroEL mediated folding catalysis for substrates with frustrated folding pathways. It is likely that GroEL from an evolutionary perspective buffered the accumulation of destabilizing mutations in its substrates, maintaining substrate folding rates to be faster than the rate of protein synthesis.

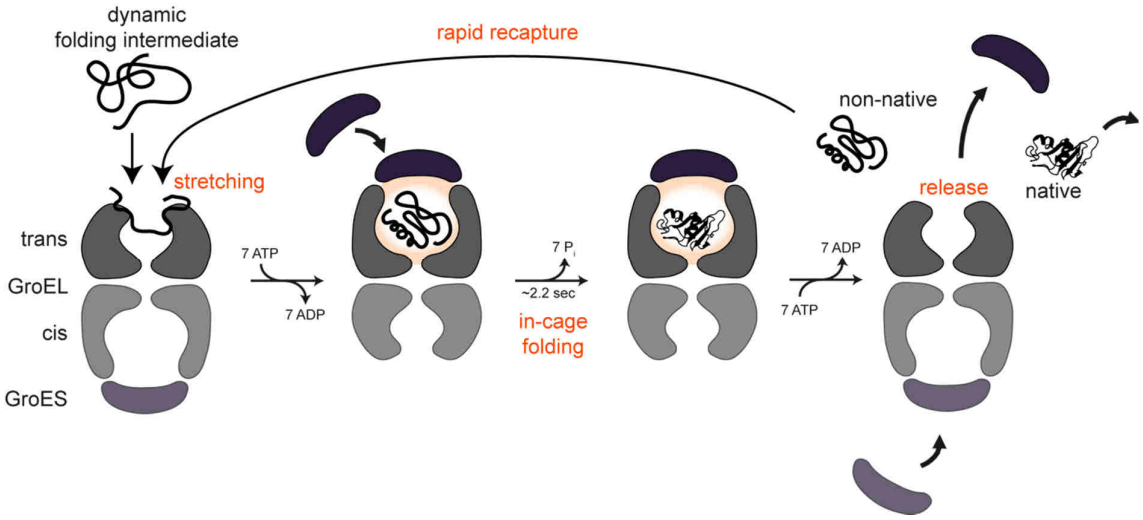

Figure 5.1 Proposed model of chaperonin assisted protein folding

The GroEL reaction cycle is initiated by binding of a non-native substrate protein that exposes hydrophobic side-chains to the apical domains of the trans ring. Binding of 7 ATP molecules and GroES sequentially displace the substrate into the central cavity. The protein is encapsulated for at least the duration of ATP hydrolysis, ~2.2 sec at 37°C in presence of substrate (Gupta et al., 2014). The GroEL cage actively promotes substrate folding by reducing the entropic energy barrier of the transition state. Binding of 7ATP and GroES to the opposing ring triggers release of ADP, GroES and substrate. In case the substrate could not fold to the native state, it is rapidly recaptured and undergoes subsequent rounds of folding. Native substrate is released into the cytosol.

Application of Ockham's razor to more complex models containing unproductive side reactions or out of cage folding, leaves a simplified active cage model in which especially the GroEL cage wall evolved to assist folding of proteins on a biologically relevant timescale, involving substrate encapsulation as an active principle. Further experimental research is required to separately investigate the contribution of the steric confinement proper, the negatively charged cage wall and the GroEL C-tails to substrate folding catalysis. It is likely that all factors contribute to a different extent, depending on the encapsulated substrate molecule. Furthermore, it remains to be seen if the active cage model does not only apply to GroEL but also to other chaperonins, especially in higher organisms. Interestingly, folding

acceleration in the presence of an archaeal group II chaperonin was reported for acid denatured GFP (Nakagawa et al., 2014; Yamamoto et al., 2014). For GroEL, the active cage model remains the most convincing and elegant explanation for chaperonin assisted protein folding.

6 References

Agashe, V.R., Guha, S., Chang, H.C., Genevaux, P., Hayer-Hartl, M., Stemp, M., Georgopoulos, C., Hartl, F.U., and Barral, J.M. (2004). Function of trigger factor and DnaK in multidomain protein folding: Increase in yield at the expense of folding speed. Cell *117*, 199–209.

Ambrose, A.J., Fenton, W., Mason, D.J., Chapman, E., and Horwich, A.L. (2015). Unfolded DapA forms aggregates when diluted into free solution, confounding comparison with folding by the GroEL/GroES chaperonin system. FEBS Lett. 8–10.

Anfinsen, C.B. (1973). Principles that govern the folding of protein chains. Science *181*, 223–230.

Apetri, A.C., and Horwich, A.L. (2008). Chaperonin chamber accelerates protein folding through passive action of preventing aggregation. Proc. Natl. Acad. Sci. U. S. A. *105*, 17351–17355.

Azia, A., Unger, R., and Horovitz, A. (2012). What distinguishes GroEL substrates from other Escherichia coli proteins? FEBS J. *279*, 543–550.

Bacia, K., and Schwille, P. (2007). Practical guidelines for dual-color fluorescence cross-correlation spectroscopy. Nat. Protoc. *2*, 2842–2856.

Baldwin, R.L. (1996). Why is protein folding so fast? Proc. Natl. Acad. Sci. U. S. A. *93*, 2627–2628.

Baldwin, R.L., and Rose, G.D. (1999a). Is protein folding hierarchic? I. Local structure and peptide folding. Trends Biochem. Sci. *24*, 26–33.

Baldwin, R.L., and Rose, G.D. (1999b). Is protein folding hierarchic? II. Folding intermediates and transition states. Trends Biochem. Sci. *24*, 77–83.

Bartolucci, C., Lamba, D., Grazulis, S., Manakova, E., and Heumann, H. (2005). Crystal structure of wild-type chaperonin GroEL. J. Mol. Biol. *354*, 940–951.

Baumketner, A., Jewett, A., and Shea, J.E.E. (2003). Effects of Confinement in Chaperonin Assisted Protein Folding: Rate Enhancement by Decreasing the Roughness of the Folding Energy Landscape. J. Mol. Biol. *332*, 701–713.

Beissinger, M., Rutkat, K., and Buchner, J. (1999). Catalysis, commitment and encapsulation during GroE-mediated folding. J. Mol. Biol. *289*, 1075–1092.

Bertani, G. (1951). Studies on lysogenesis. I. The mode of phage liberation by lysogenic Escherichia coli. J. Bacteriol. *62*, 293–300.

Bertelsen, E.B., Chang, L., Gestwicki, J.E., and Zuiderweg, E.R.P. (2009). Solution conformation of wild-type E. coli Hsp70 (DnaK) chaperone complexed with ADP and substrate. Proc. Natl. Acad. Sci. U. S. A. *106*, 8471–8476.

Bicout, D.J., and Szabo, A. (2000). Entropic barriers, transition states, funnels, and exponential protein folding kinetics: a simple model. Protein Sci. *9*, 452–465.

Böcking, T., Aguet, F., Harrison, S.C., and Kirchhausen, T. (2011). Single-molecule analysis of a molecular disassemblase reveals the mechanism of Hsc70-driven clathrin uncoating. Nat. Struct. Mol. Biol. *18*, 295–301.

Braig, K., Simon, M., Furuya, F., Hainfeld, J.F., and Horwich, a L. (1993). A polypeptide bound by the chaperonin groEL is localized within a central cavity. Proc. Natl. Acad. Sci. U. S. A. *90*, 3978–3982.

Braig, K., Otwinowski, Z., Hegde, R., Boisvert, D.C., Joachimiak, A., Horwich, A.L., and Sigler, P.B. (1994). The crystal structure of the bacterial chaperonin GroEL at 2.8 A. Nature *371*, 578–586.

Brinker, A., Pfeifer, G., Kerner, M.J., Naylor, D.J., Hartl, F.U.U., and Hayer-Hartl, M. (2001). Dual function of protein confinement in chaperonin-assisted protein folding. Cell *107*, 223–233.

Bukau, B., and Horwich, A.L. (1998). The Hsp70 and Hsp60 chaperone machines. Cell *92*, 351–366.

Calloni, G., Chen, T., Schermann, S.M., Chang, H.-C., Genevaux, P., Agostini, F., Tartaglia, G.G., Hayer-Hartl, M., and Hartl, F.U. (2012). DnaK functions as a central hub in the E. coli chaperone network. Cell Rep. *1*, 251–264.

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

Chandrasekhar, G.N., Tilly, K., Woolford, C., Hendrix, R., and Georgopoulos, C. (1986). Purification and properties of the groES morphogenetic protein of Escherichia coli. J. Biol. Chem. *261*, 12414–12419.

Chen, D.-H., Madan, D., Weaver, J., Lin, Z., Schröder, G.F., Chiu, W., and Rye, H.S. (2013). Visualizing GroEL/ES in the act of encapsulating a folding protein. Cell *153*, 1354–1365.

Chen, S., Roseman, A.M., Hunter, A.S., Wood, S.P., Burston, S.G., Ranson, N.A., Clarke, A.R., and Saibil, H.R. (1994). Location of a folding protein and shape changes in GroEL-GroES complexes imaged by cryo-electron microscopy. Nature *371*, 261–264.

Cheng, M.Y., Hartl, F.U., Martin, J., Pollock, R.A., Kalousek, F., Neupert, W., Hallberg, E.M., Hallberg, R.L., and Horwich, A.L. (1989). Mitochondrial heat-shock protein hsp60 is essential for assembly of proteins imported into yeast mitochondria. Nature *337*, 620–625.

Chiti, F., and Dobson, C.M. (2006). Protein misfolding, functional amyloid, and human disease. Annu. Rev. Biochem. *75*, 333–366.

Chun, S.Y., Strobel, S., Bassford, P., and Randall, L.L. (1993). Folding of maltose-binding protein. Evidence for the identity of the rate-determining step in vivo and in vitro. J. Biol. Chem. *268*, 20855–20862.

Corsepius, N.C., and Lorimer, G.H. (2013). Measuring how much work the chaperone GroEL can do. Proc. Natl. Acad. Sci. U. S. A. *110*, E2451–E2459.

Dalton, K.M., Frydman, J., and Pande, V.S. (2015). The Dynamic Conformational Cycle of the Group I Chaperonin C-Termini Revealed via Molecular Dynamics Simulation. PLoS One *10*, e0117724.

Deniz, A.A., Dahan, M., Grunwell, J.R., Ha, T., Faulhaber, A.E., Chemla, D.S., Weiss, S., and Schultz, P.G. (1999). Single-pair fluorescence resonance energy transfer on freely diffusing molecules: observation of Förster distance dependence and subpopulations. Proc. Natl. Acad. Sci. U. S. A. *96*, 3670–3675.

Dill, K.A., and Chan, H.S. (1997). From Levinthal to pathways to funnels. Nat. Struct. Biol. *4*, 10–19.

Dobson, C.M., Šali, A., and Karplus, M. (1998). Protein Folding: A Perspective from Theory and Experiment. Angew. Chemie Int. Ed. *37*, 868–893.

Dobson, R.C.J., Griffin, M.D.W., Jameson, G.B., and Gerrard, J.A. (2005). The crystal structures of native and (S)-lysine-bound dihydrodipicolinate synthase from Escherichia coli with improved resolution show new features of biological significance. Acta Crystallogr. D. Biol. Crystallogr. *61*, 1116–1124.

Eichner, T., Kalverda, A.P., Thompson, G.S., Homans, S.W., and Radford, S.E. (2011). Conformational Conversion during Amyloid Formation at Atomic Resolution. Mol. Cell *41*, 161–172.

Elad, N., Farr, G.W., Clare, D.K., Orlova, E. V., Horwich, A.L., and Saibil, H.R. (2007). Topologies of a Substrate Protein Bound to the Chaperonin GroEL. Mol. Cell *26*, 415–426.

Ellis, R.J. (1994). Molecular chaperones. Opening and closing the Anfinsen cage. Curr. Biol. *4*, 633–635.

Ellis, R.J., and Hartl, F.U. (1996). Protein folding in the cell: competing models of chaperonin function. FASEB J. *10*, 20–26.

Enderlein, J., Gregor, I., Patra, D., and Fitter, J. (2005). Statistical analysis of diffusion coefficient determination by fluorescence correlation spectroscopy. J. Fluoresc. *15*, 415–422.

England, J.L., and Pande, V.S. (2008). Potential for modulation of the hydrophobic effect inside chaperonins. Biophys. J. *95*, 3391–3399.

England, J.L., Lucent, D., and Pande, V.S. (2008). A role for confined water in chaperonin function. J. Am. Chem. Soc. *130*, 11838–11839.

Fenton, W.A., Kashi, Y., Furtak, K., and Horwich, A.L. (1994). Residues in chaperonin GroEL required for polypeptide binding and release. Nature *371*, 614–619.

Franck, J.M., Sokolovski, M., Kessler, N., Matalon, E., Gordon-Grossman, M., Han, S.I., Goldfarb, D., and Horovitz, A. (2014). Probing water density and dynamics in the chaperonin GroEL cavity. J. Am. Chem. Soc. *136*, 9396–9403.

Frank, G.A., Goomanovsky, M., Davidi, A., Ziv, G., Horovitz, A., and Haran, G. (2010). Out-of-equilibrium conformational cycling of GroEL under saturating ATP concentrations. Proc. Natl. Acad. Sci. U. S. A. *107*, 6270–6274.

Fujiwara, K., Ishihama, Y., Nakahigashi, K., Soga, T., and Taguchi, H. (2010). A systematic survey of in vivo obligate chaperonin-dependent substrates. EMBO J. *29*, 1552–1564.

Genevaux, P., Keppel, F., Schwager, F., Langendijk-Genevaux, P.S., Hartl, F.U., and Georgopoulos, C. (2004). In vivo analysis of the overlapping functions of DnaK and trigger factor. EMBO Rep. *5*, 195–200.

Georgescauld, F., Popova, K., Gupta, A.J., Bracher, A., Engen, J.R., Hayer-Hartl, M., and Hartl, F.U. (2014). GroEL/ES chaperonin modulates the mechanism and accelerates the rate of TIM-barrel domain folding. Cell *157*, 922–934.

Gershenson, A., and Gierasch, L.M. (2011). Protein folding in the cell: challenges and progress. Curr. Opin. Struct. Biol. *21*, 32–41.

Goloubinoff, P., Gatenby, A.A., and Lorimer, G.H. (1989). GroE heat-shock proteins promote assembly of foreign prokaryotic ribulose bisphosphate carboxylase oligomers in Escherichia coli. Nature *337*, 44–47.

Gupta, A.J., Haldar, S., Miličić, G., Hartl, F.U., and Hayer-Hartl, M. (2014). Active cage mechanism of chaperonin-assisted protein folding demonstrated at single-molecule level. J. Mol. Biol. *426*, 2739–2754.

Haldar, S., Mitra, S., and Chattopadhyay, K. (2010). Role of protein stabilizers on the conformation of the unfolded state of cytochrome c and its early folding kinetics: investigation at single molecular resolution. J. Biol. Chem. *285*, 25314–25323.

Hartl, F.U. (1996). Molecular chaperones in cellular protein folding. Nature *381*, 571–579.

Hartl, F.U., and Hayer-Hartl, M. (2009). Converging concepts of protein folding in vitro and in vivo. Nat. Struct. Mol. Biol. *16*, 574–581.

Hartl, F.U., Bracher, A., and Hayer-Hartl, M. (2011). Molecular chaperones in protein folding and proteostasis. Nature *475*, 324–332.

Haustein, E., and Schwille, P. (2004). Single-molecule spectroscopic methods. Curr. Opin. Struct. Biol. *14*, 531–540.

Hawe, A., Sutter, M., and Jiskoot, W. (2008). Extrinsic fluorescent dyes as tools for protein characterization. Pharm. Res. *25*, 1487–1499.

Hayer-Hartl, M., and Minton, A.P. (2006). A simple semiempirical model for the effect of molecular confinement upon the rate of protein folding. Biochemistry *45*, 13356–13360.

Hayer-Hartl, M.K., Ewbank, J.J., Creighton, T.E., and Hartl, F.U. (1994). Conformational specificity of the chaperonin GroEL for the compact folding intermediates of alpha-lactalbumin. EMBO J. *13*, 3192–3202.

Hayer-Hartl, M.K., Martin, J., and Hartl, F.U. (1995). Asymmetrical interaction of GroEL and GroES in the ATPase cycle of assisted protein folding. Science *269*, 836–841.

Hayer-Hartl, M.K., Weber, F., and Hartl, F.U. (1996). Mechanism of chaperonin action: GroES binding and release can drive GroEL-mediated protein folding in the absence of ATP hydrolysis. EMBO J. *15*, 6111–6121.

Hillger, F., Hänni, D., Nettels, D., Geister, S., Grandin, M., Textor, M., and Schuler, B. (2008). Probing protein-chaperone interactions with single-molecule fluorescence spectroscopy. Angew. Chemie - Int. Ed. *47*, 6184–6188.

Hofmann, H., Hillger, F., Pfeil, S.H., Hoffmann, A., Streich, D., Haenni, D., Nettels, D., Lipman, E.A., and Schuler, B. (2010). Single-molecule spectroscopy of protein folding in a chaperonin cage. Proc. Natl. Acad. Sci. U. S. A. *107*, 11793–11798.

Horovitz, A., and Willison, K.R. (2005). Allosteric regulation of chaperonins. Curr. Opin. Struct. Biol. *15*, 646–651.

Horst, R., Bertelsen, E.B., Fiaux, J., Wider, G., Horwich, A.L., and Wüthrich, K. (2005). Direct NMR observation of a substrate protein bound to the chaperonin GroEL. Proc. Natl. Acad. Sci. U. S. A. *102*, 12748–12753.

Horwich, A.L., Fenton, W.A., Chapman, E., and Farr, G.W. (2007). Two families of chaperonin: physiology and mechanism. Annu. Rev. Cell Dev. Biol. *23*, 115–145.

Horwich, A.L., Apetri, A.C., and Fenton, W.A. (2009). The GroEL/GroES cis cavity as a passive anti-aggregation device. FEBS Lett. *583*, 2654–2662.

Ishino, S., Kawata, Y., Taguchi, H., Kajimura, N., Matsuzaki, K., and Hoshino, M. (2015). Effects of C-terminal truncation of chaperonin GroEL on the yield of in-cage folding of the green fluorescent protein. J. Biol. Chem. jbc.M114.633636.

Jewett, A.I., and Shea, J.-E. (2010). Reconciling theories of chaperonin accelerated folding with experimental evidence. Cell. Mol. Life Sci. *67*, 255–276.

Jiang, Y., Douglas, N.R., Conley, N.R., Miller, E.J., Frydman, J., and Moerner, W.E. (2011). Sensing cooperativity in ATP hydrolysis for single multisubunit enzymes in solution. Proc. Natl. Acad. Sci. U. S. A. *108*.

Kalra, A., Tugcu, N., Cramer, S.M., and Garde, S. (2001). Salting-In and Salting-Out of Hydrophobic Solutes in Aqueous Salt Solutions Salting-In and Salting-Out of Hydrophobic Solutes in Aqueous Salt Solutions. Society *105*, 6380–6386.

Kauzmann, W. (1959). Some Factors in the Interpretation of Protein Denaturation. Adv. Protein Chem. *14*, 1–63.

Kaya, H., and Chan, H.S. (2003). Origins of chevron rollovers in non-two-state protein folding kinetics. Phys. Rev. Lett. *90*, 258104.

Kellner, R., Hofmann, H., Barducci, A., Wunderlich, B., Nettels, D., and Schuler, B. (2014). Single-molecule spectroscopy reveals chaperone-mediated expansion of substrate protein. Proc. Natl. Acad. Sci. U. S. A.

Kerner, M.J., Naylor, D.J., Ishihama, Y., Maier, T., Chang, H.-C., Stines, A.P., Georgopoulos, C., Frishman, D., Hayer-Hartl, M., Mann, M., et al. (2005). Proteome-wide analysis of chaperonin-dependent protein folding in Escherichia coli. Cell *122*, 209–220.

Kim, S.A., Heinze, K.G., and Schwille, P. (2007). Fluorescence correlation spectroscopy in living cells. Nat. Methods *4*, 963–973.

Kim, S.Y., Miller, E.J., Frydman, J., and Moerner, W.E. (2010). Action of the chaperonin GroEL/ES on a non-native substrate observed with single-molecule FRET. J. Mol. Biol. *401*, 553–563.

Kim, Y.E., Hipp, M.S., Bracher, A., Hayer-Hartl, M., and Hartl, F.U. (2013). Molecular chaperone functions in protein folding and proteostasis. Annu. Rev. Biochem. *82*, 323–355.

Kityk, R., Kopp, J., Sinning, I., and Mayer, M.P. (2012). Structure and Dynamics of the ATP-Bound Open Conformation of Hsp70 Chaperones. Mol. Cell *48*, 863–874.

Krieger, F., Fierz, B., Bieri, O., Drewello, M., and Kiefhaber, T. (2003). Dynamics of unfolded polypeptide chains as model for the earliest steps in protein folding. J. Mol. Biol. *332*, 265–274.

Krishna, M.M.G., Lin, Y., and Walter Englander, S. (2004). Protein misfolding: Optional barriers, misfolded intermediates, and pathway heterogeneity. J. Mol. Biol. *343*, 1095–1109.

Laber, B., Gomis-Rüth, F.X., Romão, M.J., and Huber, R. (1992). Escherichia coli dihydrodipicolinate synthase. Identification of the active site and crystallization. Biochem. J. *288 (Pt 2*, 691–695.

Landry, S.J., Zeilstra-Ryalls, J., Fayet, O., Georgopoulos, C., and Gierasch, L.M. (1993). Characterization of a functionally important mobile domain of GroES. Nature *364*, 255–258.

Langer, T., Pfeifer, G., Martin, J., Baumeister, W., and Hartl, F.U. (1992). Chaperonin-mediated protein folding: GroES binds to one end of the GroEL cylinder, which accommodates the protein substrate within its central cavity. EMBO J. *11*, 4757–4765.

Lee, N.K., Kapanidis, A.N., Wang, Y., Michalet, X., Mukhopadhyay, J., Ebright, R.H., and Weiss, S. (2005). Accurate FRET measurements within single diffusing biomolecules using alternating-laser excitation. Biophys. J. *88*, 2939–2953.

Leitner, A., Joachimiak, L.A., Bracher, A., Mönkemeyer, L., Walzthoeni, T., Chen, B., Pechmann, S., Holmes, S., Cong, Y., Ma, B., et al. (2012). The molecular architecture of the eukaryotic chaperonin TRiC/CCT. Structure *20*, 814–825.

Levinthal, C. (1968). Are there pathways for protein folding? J. Chim. Phys. Physico-Chimie Biol. *65*, 44–45.

Levinthal, C. (1969). How to Fold Graciously. In Mossbauer Spectroscopy in Biological Systems: Proceedings of a Meeting Held at Allerton House, Monticello, Illinois, J.T.P. Debrunnder, and E. Munck, eds. (University of Illinois Press), pp. 22–24.

Lin, Z., Madan, D., and Rye, H.S. (2008). GroEL stimulates protein folding through forced unfolding. Nat. Struct. Mol. Biol. *15*, 303–311.

Lin, Z., Puchalla, J., Shoup, D., and Rye, H.S. (2013). Repetitive protein unfolding by the trans ring of the GroEL-GroES chaperonin complex stimulates folding. J. Biol. Chem. *288*, 30944–30955.

De Los Rios, P., and Barducci, A. (2014). Hsp70 chaperones are non-equilibrium machines that achieve ultra-affinity by energy consumption. Elife *2014*.

Lucent, D., Vishal, V., and Pande, V.S. (2007). Protein folding under confinement: a role for solvent. Proc. Natl. Acad. Sci. U. S. A. *104*, 10430–10434.

Lucent, D., England, J., and Pande, V. (2009). Inside the chaperonin toolbox: theoretical and computational models for chaperonin mechanism. Phys. Biol. *6*, 015003.

Mallam, A.L., and Jackson, S.E. (2011). Knot formation in newly translated proteins is spontaneous and accelerated by chaperonins. Nat. Chem. Biol. *8*, 147–153.

Mapa, K., Sikor, M., Kudryavtsev, V., Waegemann, K., Kalinin, S., Seidel, C.A.M., Neupert, W., Lamb, D.C., and Mokranjac, D. (2010). The Conformational Dynamics of the Mitochondrial Hsp70 Chaperone. Mol. Cell *38*, 89–100.

Marcinowski, M., Höller, M., Feige, M.J., Baerend, D., Lamb, D.C., and Buchner, J. (2011). Substrate discrimination of the chaperone BiP by autonomous and cochaperone-regulated conformational transitions. Nat. Struct. Mol. Biol. *18*, 150–158.

Martin, J., Langer, T., Boteva, R., Schramel, A., Horwich, A.L., and Hartl, F.U. (1991). Chaperonin-mediated protein folding at the surface of groEL through a “molten globule”-like intermediate. Nature *352*, 36–42.

Martin, J., Mayhew, M., Langer, T., and Hartl, F.U. (1993). The reaction cycle of GroEL and GroES in chaperonin-assisted protein folding. Nature *366*, 228–233.

Mashaghi, A., Kramer, G., Lamb, D.C., Mayer, M.P., and Tans, S.J. (2014). Chaperone action at the single-molecule level. Chem. Rev. *114*, 660–676.

Matagne, A., Jamin, M., Chung, E.W., Robinson, C. V, Radford, S.E., and Dobson, C.M. (2000). Thermal unfolding of an intermediate is associated with non-Arrhenius kinetics in the folding of hen lysozyme. J. Mol. Biol. *297*, 193–210.

Mayer, M.P. (2010). Gymnastics of molecular chaperones. Mol. Cell *39*, 321–331.

McLennan, N., and Masters, M. (1998). GroE is vital for cell-wall synthesis. Nature *392*, 139.

Merz, F., Boehringer, D., Schaffitzel, C., Preissler, S., Hoffmann, A., Maier, T., Rutkowska, A., Lozza, J., Ban, N., Bukau, B., et al. (2008). Molecular mechanism and structure of Trigger Factor bound to the translating ribosome. EMBO J. *27*, 1622–1632.

Meyer, A.S., Gillespie, J.R., Walther, D., Millet, I.S., Doniach, S., and Frydman, J. (2003). Closing the folding chamber of the eukaryotic chaperonin requires the transition state of ATP hydrolysis. Cell *113*, 369–381.

Motojima, F., and Yoshida, M. (2010). Polypeptide in the chaperonin cage partly protrudes out and then folds inside or escapes outside. EMBO J. *29*, 4008–4019.

Motojima, F., Motojima-Miyazaki, Y., and Yoshida, M. (2012). Revisiting the contribution of negative charges on the chaperonin cage wall to the acceleration of protein folding. Proc. Natl. Acad. Sci. U. S. A. *109*, 15740–15745.

Mukhopadhyay, S., Krishnan, R., Lemke, E.A., Lindquist, S., and Deniz, A.A. (2007). A natively unfolded yeast prion monomer adopts an ensemble of collapsed and rapidly fluctuating structures. Proc. Natl. Acad. Sci. U. S. A. *104*, 2649–2654.

Müller, B.K., Zaychikov, E., Bräuchle, C., and Lamb, D.C. (2005). Pulsed interleaved excitation. Biophys. J. *89*, 3508–3522.

Müller, C.B., Loman, A., Pacheco, V., Koberling, F., Willbold, D., Richtering, W., and Enderlein, J. (2008). Precise measurement of diffusion by multi-color dual-focus fluorescence correlation spectroscopy. EPL (Europhysics Lett. *83*, 46001.

Nakagawa, A., Moriya, K., Arita, M., Yamamoto, Y., Kitamura, K., Ishiguro, N., Kanzaki, T., Oka, T., Makabe, K., Kuwajima, K., et al. (2014). Dissection of the ATP-dependent conformational change cycle of a group II chaperonin. J. Mol. Biol. *426*, 447–459.

Neuweiler, H., Löllmann, M., Doose, S., and Sauer, M. (2007). Dynamics of Unfolded Polypeptide Chains in Crowded Environment Studied by Fluorescence Correlation Spectroscopy. J. Mol. Biol. *365*, 856–869.

Neuweiler, H., Johnson, C.M., and Fersht, A.R. (2009). Direct observation of ultrafast folding and denatured state dynamics in single protein molecules. Proc. Natl. Acad. Sci. U. S. A. *106*, 18569–18574.

Niwa, T., Kanamori, T., Ueda, T., and Taguchi, H. (2012). Global analysis of chaperone effects using a reconstituted cell-free translation system. Proc. Natl. Acad. Sci. *109*, 8937–8942.

Oh, E., Becker, A.H., Sandikci, A., Huber, D., Chaba, R., Gloge, F., Nichols, R.J., Typas, A., Gross, C.A., Kramer, G., et al. (2011). Selective ribosome profiling reveals the cotranslational chaperone action of trigger factor in vivo. Cell *147*, 1295–1308.

Oliveberg, M., Tan, Y.J., and Fersht, A.R. (1995). Negative activation enthalpies in the kinetics of protein folding. Proc. Natl. Acad. Sci. U. S. A. *92*, 8926–8929.

Ostermann, J., Horwich, A.L., Neupert, W., and Hartl, F.U. (1989). Protein folding in mitochondria requires complex formation with hsp60 and ATP hydrolysis. Nature *341*, 125–130.

Pace, C.N., Shirley, B.A., McNutt, M., and Gajiwala, K. (1996). Forces contributing to the conformational stability of proteins. FASEB J. *10*, 75–83.

Poso, D., Clarke, A.R., and Burston, S.G. (2004). A kinetic analysis of the nucleotide-induced allosteric transitions in a single-ring mutant of GroEL. J. Mol. Biol. *338*, 969–977.

Ramachandran, G.N., and Sasisekharan, V. (1968). Conformation of polypeptides and proteins. Adv. Protein Chem. *23*, 283–438.

Ramachandran, G.N., Ramakrishnan, C., and Sasisekharan, V. (1963). Stereochemistry of polypeptide chain configurations. J. Mol. Biol. *7*, 95–99.

Raran-Kurussi, S., and Waugh, D.S. (2012). The Ability to Enhance the Solubility of Its Fusion Partners Is an Intrinsic Property of Maltose-Binding Protein but Their Folding Is Either Spontaneous or Chaperone-Mediated. PLoS One *7*.

Reboul, C.F., Porebski, B.T., Griffin, M.D.W., Dobson, R.C.J., Perugini, M.A., Gerrard, J.A., and Buckle, A.M. (2012). Structural and dynamic requirements for optimal activity of the essential bacterial enzyme dihydrodipicolinate synthase. PLoS Comput. Biol. *8*.

Richard, J.P., Zhai, X., and Malabanan, M.M. (2014). Reflections on the catalytic power of a TIM-barrel. Bioorg. Chem. *57*, 206–212.

Sabelko, J., Ervin, J., and Gruebele, M. (1999). Observation of strange kinetics in protein folding. Proc. Natl. Acad. Sci. U. S. A. *96*, 6031–6036.

Saibil, H., Dong, Z., Wood, S., and auf der Mauer, A. (1991). Binding of chaperonins. Nature *353*, 25–26.

Sameshima, T., Iizuka, R., Ueno, T., Wada, J., Aoki, M., Shimamoto, N., Ohdomari, I., Tanii, T., and Funatsu, T. (2010). Single-molecule study on the decay process of the football-shaped GroEL-GroES complex using zero-mode waveguides. J. Biol. Chem. *285*, 23159–23164.

Sauer, M., and Neuweiler, H. (2014). PET-FCS: probing rapid structural fluctuations of proteins and nucleic acids by single-molecule fluorescence quenching. Methods Mol. Biol. *1076*, 597–615.

Sharff, A.J., Rodseth, L.E., Spurlino, J.C., and Quiocho, F.A. (1992). Crystallographic evidence of a large ligand-induced hinge-twist motion between the two domains of the maltodextrin binding protein involved in active transport and chemotaxis. Biochemistry *31*, 10657–10663.

Sharma, S., Chakraborty, K., Müller, B.K., Astola, N., Tang, Y.-C., Lamb, D.C., Hayer-Hartl, M., and Hartl, F.U. (2008). Monitoring protein conformation along the pathway of chaperonin-assisted folding. Cell *133*, 142–153.

Shewmaker, F., Maskos, K., Simmerling, C., and Landry, S.J. (2001). The Disordered Mobile Loop of GroES Folds into a Defined β-Hairpin upon Binding GroEL. J. Biol. Chem. *276*, 31257–31264.

Sikor, M., Mapa, K., von Voithenberg, L.V., Mokranjac, D., and Lamb, D.C. (2013). Real-time observation of the conformational dynamics of mitochondrial Hsp70 by spFRET. EMBO J. *32*, 1639–1649.

Sirur, A., and Best, R.B. (2013). Effects of interactions with the GroEL cavity on protein folding rates. Biophys. J. *104*, 1098–1106.

Song, B., Cho, J.H., and Raleigh, D.P. (2007). Ionic-strength-dependent effects in protein folding: Analysis of rate equilibrium free-energy relationships and their interpretation. Biochemistry *46*, 14206–14214.

Sparrer, H., Lilie, H., and Buchner, J. (1996). Dynamics of the GroEL-protein complex: effects of nucleotides and folding mutants. J. Mol. Biol. *258*, 74–87.

Spurlino, J.C., Lu, G.Y., and Quiocho, F.A. (1991). The 2.3-A resolution structure of the maltose- or maltodextrin-binding protein, a primary receptor of bacterial active transport and chemotaxis. J. Biol. Chem. *266*, 5202–5219.

Taguchi, H., Ueno, T., Tadakuma, H., Yoshida, M., and Funatsu, T. (2001). Single-molecule observation of protein-protein interactions in the chaperonin system. Nat. Biotechnol. *19*, 861–865.

Takei, Y., Iizuka, R., Ueno, T., and Funatsu, T. (2012). Single-molecule observation of protein folding in symmetric GroEL-(GroES)2 complexes. J. Biol. Chem. *287*, 41118–41125.

Tanford, C. (1962). Contribution of hydrophobic interactions to the stability of the globular conformation of proteins. J. Am. Chem. Soc. *84*, 4240.

Tang, Y.C., Chang, H.C., Hayer-Hartl, M., and Hartl, F.U. (2007). SnapShot: Molecular Chaperones, Part II. Cell *128*.

Tang, Y.-C., Chang, H.-C., Roeben, A., Wischnewski, D., Wischnewski, N., Kerner, M.J., Hartl, F.U., and Hayer-Hartl, M. (2006). Structural features of the GroEL-GroES nano-cage required for rapid folding of encapsulated protein. Cell *125*, 903–914.

Tang, Y.-C., Chang, H.-C., Chakraborty, K., Hartl, F.U., and Hayer-Hartl, M. (2008). Essential role of the chaperonin folding compartment in vivo. EMBO J. *27*, 1458–1468.

Teufel, D.P., Johnson, C.M., Lum, J.K., and Neuweiler, H. (2011). Backbone-driven collapse in unfolded protein chains. J. Mol. Biol. *409*, 250–262.

Thirumalai, D., and Lorimer, G.H. (2001). Chaperonin-mediated protein folding. Annu. Rev. Biophys. Biomol. Struct. *30*, 245–269.

Tissières, A., Mitchell, H.K., and Tracy, U.M. (1974). Protein synthesis in salivary glands of Drosophila melanogaster: relation to chromosome puffs. J. Mol. Biol. *84*, 389–398.

Tyagi, N.K., Fenton, W.A., Deniz, A.A., and Horwich, A.L. (2011). Double mutant MBP refolds at same rate in free solution as inside the GroEL/GroES chaperonin chamber when aggregation in free solution is prevented. FEBS Lett. *585*, 1969–1972.

Ueno, T., Taguchi, H., Tadakuma, H., Yoshida, M., and Funatsu, T. (2004). GroEL mediates protein folding with a two successive timer mechanism. Mol. Cell *14*, 423–434.

Wang, J.D., Michelitsch, M.D., and Weissman, J.S. (1998). GroEL-GroES-mediated protein folding requires an intact central cavity. Proc. Natl. Acad. Sci. U. S. A. *95*, 12163–12168.

Weaver, J., and Rye, H.S. (2014). The C-terminal tails of the bacterial chaperonin GroEL stimulate protein folding by directly altering the conformation of a substrate protein. J. Biol. Chem. *289*, 23219–23232.

Weber, F., and Hayer-Hartl, M. (2000). Prevention of rhodanese aggregation by the chaperonin GroEL. Methods Mol. Biol. *140*, 111–115.

Weissman, J.S., Hohl, C.M., Kovalenko, O., Kashi, Y., Chen, S., Braig, K., Saibil, H.R., Fenton, W.A., and Horwich, A.L. (1995). Mechanism of GroEL action: productive release of polypeptide from a sequestered position under GroES. Cell *83*, 577–587.

Weissman, J.S., Rye, H.S., Fenton, W.A., Beechem, J.M., and Horwich, A.L. (1996). Characterization of the active intermediate of a GroEL-GroES-mediated protein folding reaction. Cell *84*, 481–490.

Widengren, J., Mets, U., and Rigler, R. (1995). Fluorescence correlation spectroscopy of triplet states in solution: a theoretical and experimental study. J. Phys. Chem. *99*, 13368–13379.

Xu, Z., Horwich, A.L., and Sigler, P.B. (1997). The crystal structure of the asymmetric GroEL-GroES-(ADP)7 chaperonin complex. Nature *388*, 741–750.

Yamamoto, Y.Y., Abe, Y., Moriya, K., Arita, M., Noguchi, K., Ishii, N., Sekiguchi, H., Sasaki, Y.C., and Yohda, M. (2014). Inter-ring communication is dispensable in the reaction cycle of group II chaperonins. J. Mol. Biol. *426*, 2667–2678.

Yamasaki, R., Hoshino, M., Wazawa, T., Ishii, Y., Yanagida, T., Kawata, Y., Higurashi, T., Sakai, K., Nagai, J., and Goto, Y. (1999). Single molecular observation of the interaction of GroEL with substrate proteins. J. Mol. Biol. *292*, 965–972.

Yang, D., Ye, X., and Lorimer, G.H. (2013). Symmetric GroEL:GroES2 complexes are the protein-folding functional form of the chaperonin nanomachine. Proc. Natl. Acad. Sci. U. S. A. *110*, E4298–E4305.

Ye, X., and Lorimer, G.H. (2013). Substrate protein switches GroE chaperonins from asymmetric to symmetric cycling by catalyzing nucleotide exchange. Proc. Natl. Acad. Sci. U. S. A. *110*, E4289–E4297.

Zander, C., Sauer, M., Drexhage, K.H., Ko, D.-S., Schulz, A., Wolfrum, J., Brand, L., Eggeling, C., and Seidel, C.A.M. (1996). Detection and characterization of single molecules in aqueous solution. Appl. Phys. B Laser Opt. *63*, 517–523.

Zhuravleva, A., and Gierasch, L.M. (2011). Allosteric signal transmission in the nucleotide-binding domain of 70-kDa heat shock protein (Hsp70) molecular chaperones. Proc. Natl. Acad. Sci. U. S. A. *108*, 6987–6992.

Zimmerman, S.B., and Minton, A.P. (1993). Macromolecular crowding: biochemical, biophysical, and physiological consequences. Annu. Rev. Biophys. Biomol. Struct. *22*, 27–65.

Zimmerman, S.B., and Trach, S.O. (1991). Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of Escherichia coli. J. Mol. Biol. *222*, 599–620.