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

Single-Molecule Spectroscopy: Investigations of Protein Folding to Multi-Laboratory Consistencies on Proteins

Ganesh Namdeo Agam

aus
Kothale, India

2021

Erklärung

Diese Dissertation wurde im Sinne von §7 der Promotionsordnung vom 28. November 2011 von Herrn Prof. Don C. Lamb, Ph.D., betreut.

Eidesstattliche Versicherung

Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfe erarbeitet.

München, 13.07.2021

Ganesh Namdeo Agam

Dissertation eingereicht am 22.04.2021

1. Gutachter: Prof. Don C. Lamb

2. Gutachter: Prof. Dr. Johannes Buchner

Mündliche Prüfung am 21.06.2021

Dedicated to my beloved brother Ramesh Namdeo Agam (2 October 1993- 6 May 2021)

Abstract

The investigation of complex biological processes has been challenging and require a variety of sophisticated tools to interpret the underlying processes. The study of the folding process in proteins is one of the focuses of this thesis work. To this end, both spontaneous and chaperone-assisted folding mechanisms were investigated. Single-molecule fluorescence spectroscopy has been extensively applied to the study of biomolecular bindings, conformational changes, and their dynamics due to its high sensitivity, time resolution, and its ability to differentiate between homogenous and heterogenous populations. Specifically, single-molecule Förster Resonance Energy Transfer (smFRET) studies on protein folding have elucidated the basic mechanisms of spontaneous protein folding, and properties of the chaperone-substrate interactions. The possibility to measure at low concentrations making it possible to avoid the aggregation, which is difficult to avoid in ensemble experiments.

To investigate the spontaneous folding mechanisms in large multi-domain proteins, two-color smFRET studies were carried out on a slowly folding version of the two-domain Maltose-binding protein (MBP). Three-color smFRET, an extension of typical two-color smFRET to three-colors, was applied on specifically labeled MBP to visualize the co-ordination between the domains as they fold. Chaperone-substrate interactions are crucial to process the substrates and thus enable them to carry out their physiological function. Cavity confinement effect of GroEL/ES, a bacterial Hsp60 on MBP folding landscape was demonstrated. Another substrate protein, p53-DNA-binding domain was probed concerning the combined action of Hsp70 and Hsp90 chaperone on its folding. To conclude the thesis work, a smFRET comparison study on proteins involving 16 laboratories was undertaken to assess the accuracy and precision of smFRET measurements as well as to determine a detection limit for dynamic motions in proteins.

Table of Contents

1 Introduction

The biological sciences have seen tremendous development in the last century. Most biological experiments were performed in the context of whole tissue or an organism until the introduction of Molecular Biology. Molecular Biology enabled the investigation of specific biomolecules in a test tube or the context of a biological environment like a cell or an organism. This approach gave the possibility to understand life on the molecular level.

Proteins are the working horse that carries out most of the processes in cells. The variety of tasks are achieved by the characteristic structures of proteins. As there have been predictions that the human genome can synthesize 20-25,000 different proteins, the fundamental understanding of how these proteins attain their structure is important. Proteins attain their functional structure by the process of folding after they are biologically synthesized by ribosomes. Understanding the protein folding phenomenon, in general, can increase our understanding of protein function and can provide insights into disease conditions that are accompanied by the loss of protein function. Molecular chaperones, a subset of the family of heat shock proteins (Hsp), play key roles in assisting in the folding of large, multi-domain, metastable, proteins with complex topologies, in the assembly of oligomeric proteins as well as, in the quality control of the proteome. A major part of this thesis characterized the domain-wise folding of a large, slow folding version of a two-domain model protein, the Maltose Binding Protein (MBP) and effect of GroEL/ES, a bacterial Hsp60, cavity on its folding landscape. A second major part of the thesis dealt with interaction between Hsp70 and Hsp90 chaperones and p53 as their substrate in the presence of co-chaperones.

Folding is the largest conformation change that takes place in a protein. An investigation of the folding process and dissection of conformational changes associated with the chaperone-substrate interactions, demands a methodology having a resolution relative to the molecular sizes, i.e. in the nanometer regime. Förster Resonance Energy Transfer (FRET) is a non-radiative energy transfer from an electronically excited molecule to another molecule in close proximity (usually <10 nanometers) via dipole-dipole coupling. Being sensitive enough to detect on a single molecule level, FRET, in combination with the single-molecule (sm) approach (smFRET), can be employed to analyze the heterogeneity present in the population. Additionally, with access to the time scales down to microseconds, smFRET is a powerful tool to probe the dynamics and kinetics of the associated conformations and binding interactions relevant for biomolecular function in real-time. Furthermore, adding another color to the conventional two-color smFRET experiments extends the methodology to monitor multiple distances simultaneously. Utilizing three-color smFRET to follow MBP folding case revealed the effect of the slowly folding domain on the global folding of the protein, demonstrating co-operativity as a fundamental aspect of the folding mechanism in discontinuous domain topology. Underlying dynamic intermediate state responsible for the slow folding was observed in both the domains and inter-domain interface of a protein. The folding sequence in the domains was demonstrated with two-color and three-color smFRET. Of note, the same intermediate was observed when MBP was bound to the GroEL, in addition to the completely

unfolded conformation. Effect of GroEL/ES cavity was illustrated where compaction in the entropically trapped refolding intermediate of DM-MBP resulted in the accelerated folding by shifting the dynamic equilibrium towards folded conformation.

Another chaperone-substrate interaction the study involved a structured, DNA-binding domain (DBD) of tumor suppressor protein p53 which was carried out using smFRET as the primary tool. Owing to its structural metastability, p53-DBD was known as a substrate for both Hsp70 and Hsp90 chaperone systems. SmFRET confirmed that the Hsp70 chaperone system keeps p53-DBD in an unfolded conformation at physiological temperatures. Interestingly, when the Hsp90 system was introduced to the Hsp70 bound DBD, an increment in the natively folded fraction of DBD was observed. Thus, the study provided direct evidence that the Hsp70 and Hsp90 chaperones can control the p53-DBD activity by processing the p53-DBD conformation based on the cellular need mainly by modulating the relative concentrations of both the Hsps.

To conclude the thesis work, a comparative study to assess the accuracy and precision of smFRET measurements on proteins is presented. This study involved measurements reporting the FRET efficiencies on the same labeled proteins across 16 laboratories. The study is a follow-up of a similar smFRET community-wide report, which dealt with the smFRET measurements on DNAs. We found a similar accuracy and precision for protein measurements as for DNA. However, most of the proteins, except structural proteins, show the presence of motions in their structure, which is important for their function, a clear distinctive case from most of the DNA structures except DNA-DNA and DNA-protein interactions involved in the processes of genome transfer, expression, maintenance, and higher-order organization of the genomic DNA. To this end, we asked how reliably one can detect these motions in proteins employing smFRET. SmFRET is one of the most suitable approaches to study protein dynamics ranging from microsecond to millisecond time regime. Hence, we included a second protein system with known microsecond flexibility in its structure as compared to the first quasi-static protein system. We assessed the ability to detect the dynamics across 8 laboratories specialized in these types of analyses. We found a clear consensus in being able to detect the dynamics in the one protein system whereas the other system was found to be static by the participating laboratories. Moreover, two important key conclusions emerged from the study; the first was the choice of the method employed to determine the detection correction factor, the most crucial factor that corrects for the efficiency in detecting donor photons with respect to the acceptor photons. Hence, governs the value of FRET efficiency and this correction heavily depends on the labeled sample, during which care should be taken to rule out any dye artifacts present in the sample. The second conclusion was that one should at least use two different dye pairs while evaluating the dynamic nature of the system under study.

The thesis is divided into the following sections. First, a general introduction to smFRET, spontaneous protein folding, and various chaperone systems are presented. The introduction is followed by the materials section and a description of the methods involved. The first two result sections are detailed in the form of two manuscripts, a spontaneous folding study on MBP and conformational processing of p53 by the Hsp70 and Hsp90 chaperones. As the previous results depend on the reliability of the smFRET experiments, the final results section discusses the

comparative investigation of smFRET measurements on proteins for precision, accuracy in determining the FRET efficiencies, distances, and detectability of dynamics.

2 Fluorescence

Biological phenomena are complex and dynamic due to the involvement of many kinds of biomolecules, compartments, physiological conditions etc. Any *in vitro* or *in vivo* measured biological activity is a manifestation of the interaction of biomolecules. Biological structures provide a means of understanding the mechanism of their action with access to the underlying biomolecular recognitions involved. Due to the inherent complexity of biological processes, chemical and physical approaches have contributed to the underlying principles governing biology, by in-depth analyses treating the biological systems from different perspectives, especially with the advancements of the various tools from physics. Hence, birth of the field of Biophysics was obvious.

Fluorescence, the term was coined by George Stokes in 1852 with his observation of quinine sulphate solution properties. Fluorescence gives information regarding various processes that occur during relaxation of a molecule to the ground state from the laser induced electronic excited state. Separating fluorescence from scattered light and excitation light with optical filters improved its sensitivity. There were efforts to use Raman scattering in quasi-elastic light scattering (QELS) method to characterize the sample but it's use was limited to large molecules and highly concentrated solutions (Elson & Magde, 1974).

Fluorescence has the advantage of high sensitivity and excellent signal-to-noise ratio. In addition to fluorescence intensity, the absorption spectra, fluorescence emission spectra, fluorescence anisotropy, fluorescence lifetime can provide a lot of information regarding the surrounding environment, fluorescent probe and the molecule attached to it. Many improvements in the field of florescence has taken place upon invention of the microscopes and improved photon and imaging detectors. Fluorescence spectroscopy has benefited by the invention of confocal microscope where spatial filters like pin-hole that greatly decreased the observation volume to ~femtolitres making possible to work with one or few fluorescent molecules in solution (Minsky, 1957). Methods like fluorescence correlation spectroscopy (FCS) were developed to quantitatively analyze the diffusion and thermodynamics of biomolecular processes and were first performed with the use of ethidium bromide bound to DNA and on freely diffusing fluorescent dyes (Magde, Elson, & Webb, 1972). Due to the limitation of slow electronics (on the order of a few milliseconds), it was only possible to study binding/unbinding and diffusion kinetics. Later, Total internal reflection microscopy has become another parallel approach to decrease the observation volume by the use of evanescent field with a depth of 100-200 nm further to study individual molecules immobilized on the surface for time scale of a few seconds/minutes (Axelrod, 1981)(Takashl, Harada, Tokunaga, Salto, & Yanagida, 1995).

Förster explained the resonance energy transfer (FRET) theory that the electronically excited state of a molecule can transfer energy to another molecule with a red shifted absorption spectrum (Forster, 1946). Owing to its sensitivity within ~10 nm (on the order of biomolecular size), FRET was applied in an ensemble assays that includes conformational changes, binding/unbinding interactions, folding etc. and dubbed as a "molecular ruler" (Stryer & Haugland, 1967), (Clegg et al., 1992). FRET was measured as a donor fluorescence decays in

an ensemble assays too, to interpret as energy transfer process (Grinvald, Haas, & Steinberg, 1972).

Utilization of fluorescence microscopy widened upon the discovery of gene encodable proteinogenic fluorescent probes (green fluorescent protein-GFP), usable in the context of cells. Their further use in improving the cellular spatial resolution in nanometer scale circumventing the diffraction limit, has won two Nobel prizes in the past two decades (Discovery of GFP- Chemistry 2008 and Super-resolution microscopy- Chemistry 2014). Super-resolution techniques like Stochastic optical reconstruction microscopy (STORM), photo-activated localization microscopy (PALM) and stimulated emission depletion (STED) microscopy revolutionized cell microscopy providing an optical image with a resolution of ~50-200 nm. This has made it possible to study cellular processes with much improved resolution. Spatio-temporal resolution of cellular processes can be enhanced using fluorescence approaches to measure the dynamics of proteins fused with fluorescent proteins. With STED, intracellular dynamic processes can be extracted but is limited due to the high laser powers needed and control over cellular concentrations is often tricky. In these cases, approaches such as Raster Image Correlation Spectroscopy (RICS) and FLIM-FRET (Fluorescence Lifetime Imaging-FRET) may be useful (Digman et al., 2005) (Wallrabe & Periasamy, 2005)(Schrimpf, Lemmens, et al., 2018) (Gregor, Chizhik, Karedla, & Enderlein, 2019).

2.1 Biophysics of single molecules with fluorescence spectroscopy

Need and discovery

One of the key advancements in Biophysics was the structural determination of biomolecules. X-ray crystallography and single-particle cryo-electron microscopy (cryo-EM) provide atomic-level structural details which represents thermodynamically stable static structure (Kendrew et al., 1958) (Y. Cheng, 2015) (Scheres, 2013). NMR, small-angle x-ray scattering (SAXS), and small-angle neutron-scattering (SANS), provide an ensemble of structures (Kovermann, Rogne, & Wolf-Watz, 2016) (Blanchet & Svergun, 2013) (Engelman & Moore, 1975).

Biomolecules move on the time scales from microseconds to seconds/minutes that are relevant for their functions. NMR and electron spin resonance (EPR) have access to dynamics over a wide range of timescales (~picoseconds-minutes) with specialized experiments (Kovermann et al., 2016) (Nesmelov & Thomas, 2010). Cross-linking mass-spectroscopy (MS), pulsed hydrogen-deuterium exchange mass-spectroscopy (HDX-MS) methods and molecular dynamics simulations have been used to interpret the structural dynamics with amino acid resolution (Konermann, Pan, & Liu, 2011) . Recently developed X-ray free electron lasers (XFEL) made possible to find more structural insights by performing femto-second resolved X-ray crystallography and time-resolved cry-EM towards the reaction co-ordinate of a undergoing reaction (Boutet et al., 2012). Recent time-resolved cryo-EM approaches are helping to get structural details of biomolecules at short times as they move and function (Frank, 2017).

To avoid the averaging of biological processes, single molecule studies have captured a lot attention since they show various steps each molecule undergoes for their function. Identification of the heterogeneity in the populations have paved way to understand the

complexity of the biological systems in real time. Single molecule tools can detect different states/conformations using either mechanical manipulation of each biomolecule with highly sensitive (~piconewton) force spectroscopies like optical tweezers, magnetic tweezers, and atomic force microscope etc. or with fluorescence by isolating and observing them individually within the small volume of a confocal microscope (~0.5-5 μm) or by attaching them on the surface and measuring them using TIRF microscopy (Mehta, Rief, Spudich, Smith, & Simmons, 1999) (Joo, Balci, Ishitsuka, Buranachai, & Ha, 2008). Methods like ABEL (anti-Brownian electro-kinetic trap) trap makes it possible to observe single molecules within a confocal microscope in solutions without immobilization (Cohen & Moerner, 2006).

Single molecule fluorescence spectroscopic methods have an advantage over force spectroscopy methods as they do not perturb the biomolecular structure or a stability by utilizing external force to change the energy surface of a molecule and are not sensitive to the fluorophore attachment points. But on the other hand, force spectroscopy accesses information over whole molecule whereas fluorescence only provides information from the labeling positions. This has limited single molecule fluorescence studies to mostly hypothesis driven investigations based on the available structures/models which allows determination of labeling positions to address specific questions. However, this is not always the case. There is a recent surge studies on unstructured/disordered regions/proteins that are not crystallizable where single molecule spectroscopy along with NMR are preferred for studying their conformational dynamics (Ferreon, Ferreon, Wright, & Deniz, 2013) (A. Borgia et al., 2018).

The very first report on single molecule absorption spectroscopy came in 1989 in solids (Moerner & Kador ', 1989) and the first single molecule fluorescence spectroscopy in 1990 on free dyes in solutions as bursts of photons in a single detection channel exploiting the fluorescence intensity and fluorescence lifetime (Shera ', Seitzinger, Davis, Keller, & Las, 1990). Single molecule reports with two colors and the possibility to measure FRET via fluorescence intensities with specifically attached labels came in 1996 using surface attached biomolecules (Ha et al., 1996) and for free molecules in solution in 1999 (Ashok A. Deniz et al., 1999). Burst integrated fluorescence lifetime changes and conformational dynamics with one fluorescence label were used to investigate the biomolecular motions (C. Eggeling, Fries, Brand, Günther, & Seidel, 1998).

Current state of the art: Fluorescence (Cross-) Correlation Spectroscopy

Improvement in detection electronics sensitive to single photons and fast data acquisition (avalanche photo-diodes, time-correlated single photon counting and charged coupled devices, and their combination with pulsed excitation sources) made it possible to access fast time scales down to nanoseconds. In terms of temporal resolution from single molecules, spectroscopy has benefited a lot from the amalgamation of fast time scales, analysis tools and dye labeling methods to study concentrations, translational and rotational diffusion, photophysical properties, binding, unbinding and interactions, conformations. These processes are possible with the utilization of methods like FCS, FRET beyond two colors spanning wide range of time scales.

Now a days, its routine to study molecular diffusional, rotational, co-diffusional properties and dye properties like triplet or darks states on picoseconds-microseconds time-scales with FCS

and FCCS (fluorescence cross-correlation on differently labeled molecule) (Haustein & Schwille, 2004). Studying fast dynamics of conformational changes (nanoseconds-microseconds) with FCS (other specialized approaches may include FRET-FCS, nanosecond-FCS), PET (Photo-induced Electron Transfer) and relatively slow dynamics (microseconds-milliseconds) with FRET experiments is possible (Nettels & Schuler, 2007), (Neuweiler & Sauer, 2004) (T. Torres & Levitus, 2007) (Kalinin, Sisamakis, Magennis, Felekyan, & Seidel, 2010).

Current state of the art: Single molecule FRET

Currently, smFRET experiments can be performed on custom built as well as commercial confocal set-ups on labeled molecules in solutions with picomolar concentrations to measure each molecule at a time. In smFRET experiments, we typically measure FRET efficiency, stoichiometry, steady state as well as time resolved fluorescence on single molecules (Kudryavtsev et al., 2012). As all photons arriving from a single burst are averaged together, these measured parameters are integrated over the diffusion time. Information from these integrated values on many single molecules can be extracted to interpret the slow dynamics spanning over diffusion time-scales, where FRET efficiency histogram broadenings beyond photon detection shot noise (photon distribution analysis, burst variance analysis) and raw photon numbers (dynamic photon distribution analysis) can be used (Antonik, Felekyan, Gaiduk, & Seidel, 2006) (Torella, Holden, Santoso, Hohlbein, & Kapanidis, 2011) (Kalinin, Sisamakis, et al., 2010). For extracting the dynamics faster than the diffusion time, burst-wise fluorescence anisotropy and lifetime (FRET efficiency versus intensity averaged donor lifetime) can be employed as they have access to the fast time scales and hence sensitive to the processes in the regime of microseconds to milliseconds (Kalinin, Sisamakis, et al., 2010). These analytical approaches can be important in describing the heterogeneities and their intermixing accurately over a wide range of timescales. Selecting out molecules undergoing FRET and performing correlation analysis on them in case of Filtered-FCS or species-FCS, has allowed to quantitatively analyze the conformational inter-conversions on fast time scales (<milliseconds) (Felekyan, Kalinin, Sanabria, Valeri, & Seidel, 2012). The above mentioned methods can be adapted to study any biomolecular system with respect to dynamics and heterogeneity (Sanabria et al., 2020), (Barth et al., 2018). Notably, these methods are well suitable for highly dynamic systems like domains attached via a flexible linker, dynamic processes of protein folding and for intrinsically disordered proteins. Computer modeling has helped to put the wealth of information available from these advanced single molecule experiments and analysis for structural interpretations (Dimura et al., 2016).

Although control over labeling positions gives high specificity, it suffers from the limitation of measuring only one distance at a time. Recent improvements in labeling methods to specifically attach multiple labels (unnatural amino acid incorporations, click chemistries, sortase ligations) and statistical analysis to treat photon noise more precisely have made it possible to perform fluorescence experiments beyond two colors (T. C. Lee, Moran, Cistrone, Dawson, & Deniz, 2018) (Mao, Hart, Schink, & Pollok, 2004) (Abhishek Chatterjee, Sun, Furman, Xiao, & Schultz, 2013) (Kolb, Finn, & Sharpless, 2001). This has opened the multi-dimensional view to the biomolecular processes and watching co-ordination between the motions involving large

sizes and multiple interaction partners using three- and even with four- color FRET. Most reported multi-color FRET studies are limits to the special case of high-affinity and on complexes attached to the surface (Ratzke, Hellenkamp, & Hugel, 2014), (Kopperger et al., 2018) (Hoi Sung Chung et al., 2017) (Yoo, Louis, Gopich, & Chung, 2018).

Current challenges of single molecule FRET

Potential roadblocks to making smFRET experiments as routine tool in molecular interactions are the available spectral range and specific labeling strategies for more than two labels. Another potential issue is the low micromolar affinity ranges of most biomolecular interactions. Typically, for measuring intra-molecular FRET, the labeled biomolecule is kept at very low (picomolar) concentration and unlabeled interaction partner is kept at micromolar concentrations. But for intermolecular multi-color FRET (as well as for FCCS) measured in solution, the concentration of the labeled partner is limited to picomolar to nanomolar range. Fast microfluidic dilution after forming multi-labeled complexes could solve the hurdle for measuring multi-color FRET in low affinity complexes (Lemke et al., 2009) (Pfeil, Wickersham, Hoffmann, & Lipman, 2009) (Gambin et al., 2011). Limiting the observation volume with zero mode waveguides can also be another possibility for measuring low affinity complexes (J. De Torres, Ghenuche, Moparthi, Grigoriev, & Wenger, 2015). Confining labeled molecules with their interaction partner in a vesicle has also worked well to measure low affinity complexes. Recently, a STED-based approach has been applied to measure sm FRET in the nanomolar regime (N. Kim et al., 2018). In parallel improvements in metal-enhanced fluorescence emission based nano-antennas may also allow measurements of high concentrations regime experiments (Acuna et al., 2012) (Punj et al., 2015). Future directions towards making good fluorescent labels in the far-red spectrum would certainly be useful with the employment of multiple labels by using unnatural amino acid incorporation, sortase ligations and click chemistries for both confocal based and TIRF based smFRET modalities in general.

In vitro to in vivo single molecule studies

Fluorescence spectroscopies are applied in general under controlled *in vitro* conditions as a bottom-up approach to reconstitute the cellular environment. However, doing spectroscopy in cells is also possible on single molecule level. Previously, FCS has been applied in cells to get the cellular dynamics of processes like signaling of calcium sensing and STED-FCS was used to show protein diffusion in membranes (S. A. Kim, Heinze, & Schwille, 2007) (Clausen et al., 2015).

Interestingly, smFRET was performed in cells after *in vitro* labeled proteins were electroporated, but performing smFRET in cells is a challenging task with the limitations of intracellular labeling approaches and limited control over the concentrations (König et al., 2015). A recent quantitative approach used for measuring accurate FRET in cells is encouraging, where FRET was measured between the fluorescent proteins applying the corrections similar to smFRET spectroscopy (Coullomb et al., 2020).

2.2 Theory of fluorescence

2.2.1. Electronic transitions that can occur during fluorescence

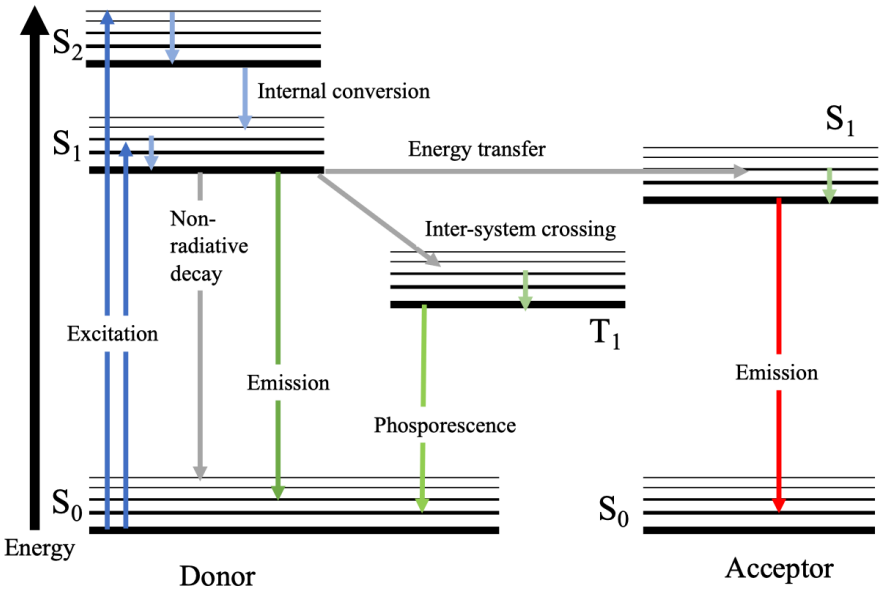

Figure 2.1 Jablonski diagram for electronic transitions.

Upon absorption of a photon, a singlet electron goes from the ground state S_0 into the excited electronic state (S_1, S_2 or higher). From the excited states, the electron relaxes to the low energy level of the S_1 state before relaxing to the ground state. In each state, vibrational relaxation occur that brings the electron to the lowest vibrational level. Return to the electronic ground state can occur with the emission of a photon. Fluorescence of the electron can be de-excited via a non-radiative route without emission of a photon. From the excited state the electron can convert slowly into a forbidden triplet (T_1) state via a spin inversion process. This state relaxes to the ground state very slowly (seconds-minutes). Another competing process called energy transfer can take place via dipole-dipole coupling when another fluorophore is in the vicinity (<10 nanometers) and meets the criteria for energy transfer.

Typically, a Jablonski diagram, as shown in **Figure 2.1** is used to illustrate the different electronic transitions occurring in a fluorescent molecule. When a photon interacts with an electron in the electronic ground state S_0, of the molecule, it excites the electron to a higher energy level instantaneously (~femtosecond timescale). Each electronic state has different vibrational energy levels. After excitation, the electron can end up in a number of vibrational levels of the S_1 state (or higher S states). Due to inherent instability of the excited state, the electron loses its energy within ~nanoseconds returning to the original ground state with emission of energy in the form of a photon. If the electron is excited to the even higher energy levels (S_2, S_3, ..etc.), it relaxes to quickly to the lowest energy level of the S_1 state. This results in the same emission spectrum irrespective of the wavelength of excitation (Kasha's Rule). An electronic transition between the same vibrational levels of ground state and excited state takes place spontaneously (pico-nanoseconds) because of the fast movements of electrons (owing to their much lighter weight than the nucleus) (Frank-Condon Principle). This leads to the similar transition probabilities for excitation events for an electron from the any vibrational level of

the ground state to similar vibrational level in the excited state as well as from any vibrational level of the excited state to similar vibrational level in the ground state during the emission. Thus, both the absorption and emission spectrum mirrors each other (Mirror Image Rule) **(Figure 2.2).** The energy lost via vibrational relaxations is dissipated in the form of heat exchange with the solvent. Hence, emitted photons have a lower energy (i.e. a higher wavelength) compared to the excitation energy (lower wavelength). This difference (generally on the order of ~20 nm) in the wavelengths is called the Stoke's Shift **(Figure 2.2)**.

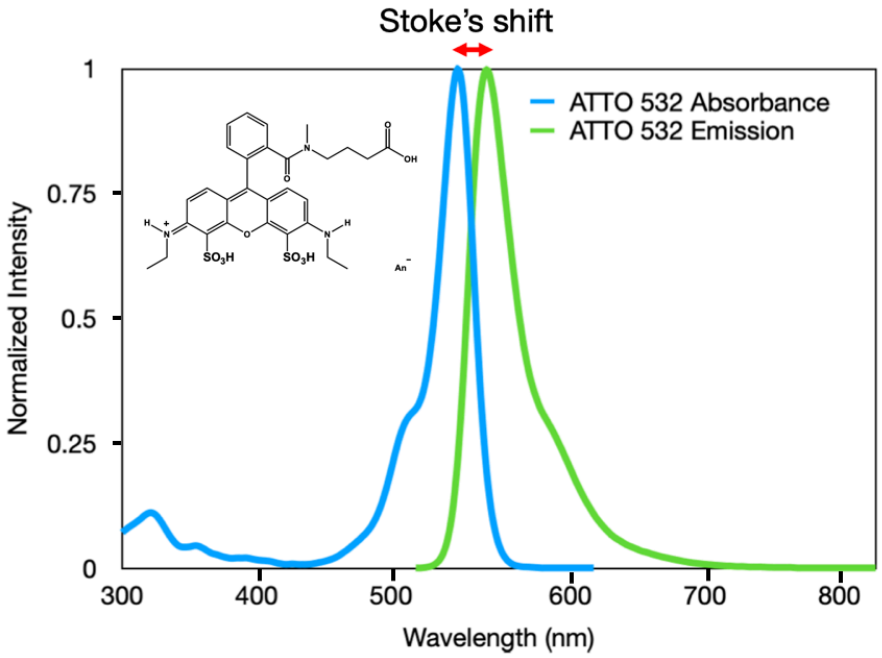

Figure 2.2 The absorption and emission spectrum of the Atto532 fluorophore.

An absorption spectrum of Atto532 fluorophore shown in blue and the emission spectrum is shown in green. Structure of the carboxy derivative of Atto532 is shown on the upper left corner of the plot. Source of the data and the structure: ATTO-TECH.

Sometimes, with a low probability, the excited electron transition take place to a triplet state, T where electron changes its spin. This transition is fast but the rate is slow (~few µs) as this transition is forbidden. This inter-system crossing leads to a further lowering in the excited electronic state of an electron. From the triplet state, return to the ground state is a forbidden transition because of the spin change. The relaxation of the triplet state can be on the seconds to minutes time scale. This prolonged luminescence is known as Phosphorescence.

Other competing de-excitation processes for the excited electron could be non-radiative in nature like transitions to the ground state without photon emission by losing the energy in the form of heat, or by transferring the energy to another molecule via coupling of transition dipole-moments (FRET).

An ideal fluorophore is the one that has a high absorption probability, and very slow non-radiative decay rates (k_{nr}) in comparison to the inherent fluorescence lifetime (k_r). Generally, a term called the Quantum yield (ϕ) is used to access the fraction of photons emitted to excitations and can be given as

$$\phi = \frac{k_r}{k_r + k_{nr}} \tag{2.1}$$

This quantity is closely related to another important aspect of a fluorophore, the measured fluorescence lifetime (τ) that depicts the amount of time electron spends on average in the excited state. It is given as the inverse of all de-excitation processes.

$$\tau = \frac{1}{k_r + k_{nr} + \cdots} \tag{2.2}$$

These two quantities are critical in terms any fluorescence measurement that needs to be thoroughly analyzed to assess the overall results.

For a free fluorophore, the most important factor that influences the fluorescence property of the fluorophore might be the solvent. Other factors that might influence these quantities can be temperature, pH, specific ions, or the near-field created in metal-enhanced fluorescence with nano-antennas. For fluorophores attached to the biomolecule of interest the attached biomolecule can have a large influence on the fluorescence property of the dye. Proteins can have a variety of different chemical properties and their dynamic motions can easily influence these quantities. This may sometimes lead to quenching of a fluorophore. Nevertheless, in some cases, this process can be used as a readout for a site-specific probe to characterize the local environment. For example, in case of photo-electron transfer (PET), where tryptophan is positioned in a close proximity (less than 10 Å) of Atto655 fluorophore to induce quenching as a readout of chain dynamics (Neuweiler & Sauer, 2004).

2.2.2 Förster Resonance Energy Transfer

One of the process that allows an excited electron to return to the ground state is by transferring its energy to another molecular by the process of dipole-dipole coupling interactions. First correct theoretical description of FRET devised by Förster and hence his name is used now a days to refer to the process as *Förster Resonance Energy Transfer* (FRET) (Forster, 1946). The fraction of the energy transfer rate (k_{FRET}) of all the decay rates is defined as the FRET efficiency (*E*).

$$E = \frac{k_{FRET}}{k_{FRET} + k_r + k_{nr}} \tag{2.3}$$

The rate at which an excited molecule (donor) transfers its energy to another molecule (acceptor) mainly depends on the distance between them (*R*) in nanometer regime. The transition dipole moments between the donor and acceptor should have a certain overlap to meet a criterion for a resonance, which explains the dependence of the energy transfer rate to the sixth power of distance between them.

$$k_{FRET} = \frac{1}{\tau_D}\left(\frac{R_0}{R}\right)^6 \tag{2.4}$$

$$E = \frac{1}{1 + \left(\frac{R}{R_0}\right)^6} \tag{2.5}$$

where τ_D is the fluorescence lifetime of a donor in the absence of an acceptor and R_0 is the Förster radius, a distance at which the FRET efficiency E is 50%. For the frequently used dye pair Atto532 and Alexa647, $R_0 = 62$ Å. The FRET efficiency is highly dependent on the distance between the fluorophores ($0.5\ R_0 - 2\ R_0$) as it appears in the equation 2.5 to sixth power (**Figure 2.3**).

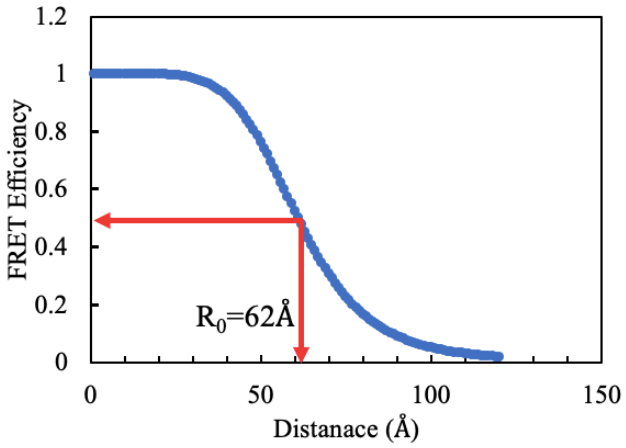

Figure 2.3 The distance dependency of the FRET efficiency assuming a $R_0 = 62$ Å for the FRET pair Atto532-Alexa647. Red arrows indicates the Förster radius (in case of 62 Å) at 0.5 FRET efficiency.

The R_0 depends upon the quantum yield of the donor without an acceptor (ϕ_D), relative orientation of the emitting donor dipole and the absorbing acceptor dipole, commonly described as orientation factor (K^2), the overlap integral of the donor emission spectrum and acceptor absorption spectrum (J) and the refractive index of the medium (n).

$$R_0^6 = \frac{9000(\ln 10)\phi_D K^2 J}{128\pi^5 n^4 N_A} \tag{2.6}$$

where N_A is the Avogadro's constant.

The orientation factor (K^2) depends on the angles of the donor emission dipole moment (θ_D) and the acceptor absorption dipole moments (θ_A) relative to vector joining the donor-acceptor (R_{DA}) and the angle between the planes defined by the transition dipole moments of the donor and the acceptor molecules and the distance vector between them (the planes are shown in the green and red respectively). **(Figure 2.4)**

$$K^2 = (\cos\theta_{DA} - 3\cos\theta_D \cos\theta_A)^2 \tag{2.7}$$

K^2 is one of the most critical quantities controlling the FRET process as it varies from 0 to 4. The precise value cannot be experimentally calculated as the dipole orientations are not accessible. This difficulty can be circumvented by assuming the unhindered rotation of both the donor and acceptor fluorophores, giving a theoretical calculated value of 2/3. This assumption can be experimentally verified by ruling out any hinderance to the rotation of the fluorophores. This information is made available with polarization dependent measurements.

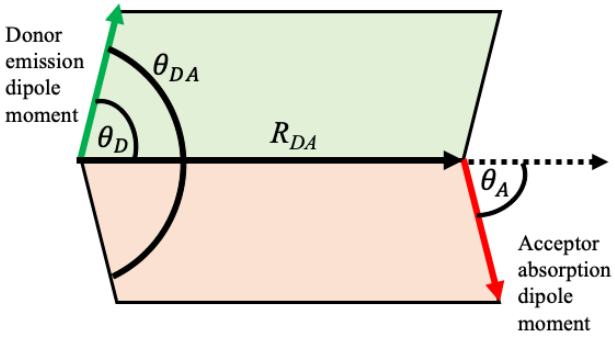

Figure 2.4 A schematic show the angles need to calculate the orientation factor.

θ_D and θ_A are donor emission dipole moment and the acceptor absorption dipole moments respectively. Vectors connecting both the dipole moments are shown in black arrow. θ_{DA} is the angle between the planes defined by the transition dipole moments of the donor and the acceptor molecules and the distance vector between them.

Another important quantity, that controls the distance dependency is the overlap integral J and can be tuned by the choice in fluorophores. The overlap integral is calculated from the emission spectrum of the donor and the absorption spectrum of the acceptor as

$$J = \int f_D(\lambda)\epsilon_A(\lambda)\lambda^4 d\lambda \tag{2.8}$$

where $f_D(\lambda)$ is the donor emission spectrum normalized to the area and $\epsilon_A(\lambda)$ is the extinction of an acceptor over a spectrum. A normalized overlap integral is shown in **Figure 2.5.**

Fluorescence anisotropy

The polarization of emitted light depends on the orientation and rotational freedom of the fluorophore. Hence, it can be used to extract rotational diffusion parameters about the fluorophore and the surrounding environment. By using linearly polarized light, the maximum absorption is induced in fluorophores having transition dipole moments aligned in the same direction. As fluorophore rotates, it emits the photons in all the directions and depolarization of an incident light takes place. This depolarization depends on the size, shape of both the fluorophore and the molecule attached to it as well as on the viscosity of a medium. Typically, anisotropy is measured by exciting a fluorophore with linearly polarized light and measuring the intensity in a plane parallel ($I_{\parallel}$) to and perpendicular ($I_{\perp}$) to the polarization plane. The

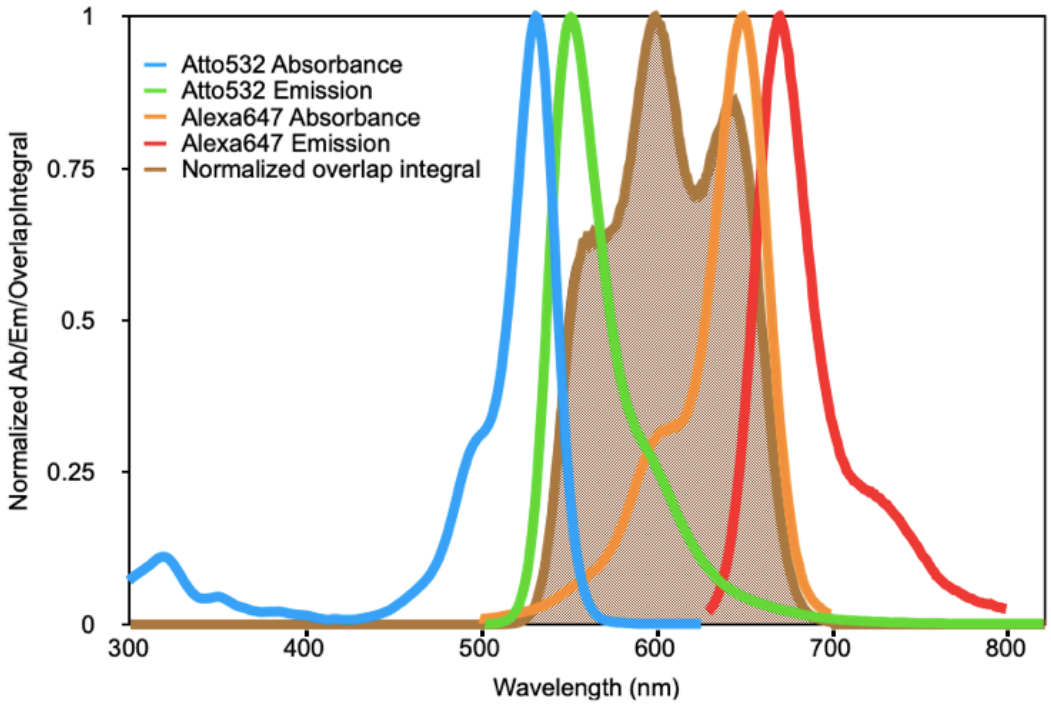

Figure 2.5 The overlap integral calculated for Atto532 and Alexa647 dyes.

The absorption and emission spectrum for Atto532 is shown in blue and green respectively. Likewise, absorption and emission spectrum for Alexa647 is shown in orange and red. The normalized calculated overlap integral is shown in shaded brown color.

fluorescence anisotropy (r) is calculated as

$$r = \frac{I_{\parallel} - I_{\perp}}{I_{\parallel} + 2\, I_{\perp}} \tag{2.9}$$

The anisotropy can be computed as steady state (integrated over time) or time resolved anisotropy. Equation 2.9 can be applied to time resolved decay collected after pulsed excitation. Time-resolved anisotropy information can be used to assess the rotational diffusion coefficient of a fluorophore and the molecule it is attached to. The free rotation of a fluorophore can be checked to validate the assumption made regarding the orientation factor i.e. $K^2 = 2/3$ for the given FRET measurement using both steady state as well as time-resolved anisotropy calculations.

3 Protein folding

3.1 Basic structural organization in proteins

Protein, the name originated from the Greek word “proteios” meaning “primary”, is certainly an apt name for their essential role in almost all biological functions. They are the most complex amongst biomolecules in terms of their chemical and structural composition. Most of the proteins need to attain a defined three-dimensional structure to carry out a variety of functions, with the exception of an increasing number of discovery of intrinsically disordered/unstructured proteins or regions/domains of proteins. Typically, each specific function is defined by a specific structure. Nevertheless, evolution has achieved the feat of making billions of proteins by using basic hierarchical orders of secondary structures.

Amino acids are the building blocks for proteins. There are so far 22 known amino acids (R-C_αH(NH2)-COOH, R is the side chain specific for every amino acid). The side chain group imparts the amino acid a unique property. These groups can be hydrophilic with polar nature and hydrophobic with non-polar nature. Hydrophilic groups mostly lie on the surface of a globular protein and hydrophobic and bulky amino acid side chains are mostly buried in the core. Each triplet codon defines a specific amino acid. The codon sequence defines the primary sequence of amino acids unique to each protein. Amino acids are covalently attached via peptide bond formation (CO-NH) between the carboxylic acid group and an amine group of the subsequent amino acid. Commonly found secondary structural elements in proteins are α-helix and a β-sheets. These structures are defined by the arrangement of the hydrogen bonds formed between the backbone carbonyl oxygen and the hydrogen of an amino group. In an α-helix, hydrogen bonds are formed between every third or fourth amino acid residue, whereas β-sheets are stabilized by the hydrogen bonds network via parallel or anti-parallel arrangement of a protein chain. Peptide bond has a partial double bond nature which restricts the rotational freedom between the C-N but C_α-N (ϕ) and C_α-C (ψ) bonds can rotate freely (**Figure 3.1**). Rotation of the peptide bond plane is further restricted by the side chain group except glycine. There are certain preferences for a given amino acid side chain groups to be accommodated sterically in the secondary structural units (Ramachandran, Ramakrishnan, & Sasisekharan, 1963). A 2D plot of the preferred angles of rotation found in secondary structures are referred to as a Ramachandran Plot. Other secondary structures that help in packaging of the typical secondary structures are β-turns (found on the ends of β-sheets) and loops (mostly found at the ends of mixed secondary structures). The locally packed protein chain forms globular domains and arrangement of these domains leads to the tertiary structure of the protein. Another higher level of organization exists in proteins that are multimeric in nature. Multiple chains with tertiary structures form a unique assembly, important for multi-functional biological processes referred to as the quaternary structures.

3.2 Protein folding mechanism

One of the early key questions of the central dogma of molecular biology is: How is the information encoded in DNA translated into functional proteins via folding? With the advancements of protein sequencing techniques, initial obvious predictions were made in the context of the primary sequence. Seminal experiments on Ribonuclease by Christian Anfinsen gave birth to protein folding research (Sela, White, & Anfinsen, 1957). Further, *in vitro* folding

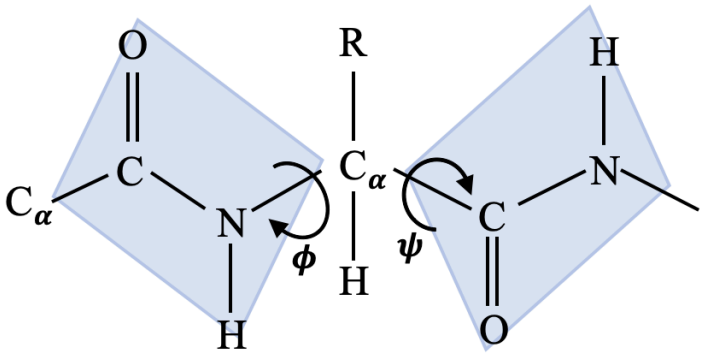

Figure 3.1 Planar geometry and dihedral angles of peptide bond.
A peptide bond C(O)-N(H) is sterically hindered due to delocalized electrons. Free rotation is possible between C_α-N (ϕ) and C_α-C (ψ) bonds.

studies on various proteins led Anfinsen to come up with a "thermodynamic hypothesis" of protein folding. It states that the amino acid sequence guides the folding of a polypeptide chain to the lowest thermodynamically stable conformation under physiological conditions (Anfinsen, 1973). For most proteins, the native conformation is the functional form present in cells. The next intuitive question was; How does the primary sequence defines the three-dimensional structure in proteins?

There have been many studies dedicated to understand the energetics of the folding process. Parallel protein denaturation studies (Kauzmann, 1959) and free energy calculations from amino acid transfer experiments (Tanford, 1962) predicted the importance of hydrophobic core of globular proteins. Overall folding process can be considered as hydrophobic collapse leading to a formation of a protein core accompanied by the loss of water constraints surrounding the hydrophobic side chains. This entropy gain of water thermodynamically drives the folding process. Chemically speaking, it became clear that, the hydrogen bonds between backbone-backbone, backbone-side chains and protein-water, salt bridges between the side chains and van der Waals packing forces are critical for protein folding into a structure with marginal stability at physiological conditions. In addition, the long-range disulfide bonds are important to keep the folded structures stable.

Different mechanisms were put forward to explain the folding process itself. Availability of crystal structures (Kendrew et al., 1958) gave a strong hint that the hydrophobic force was the main driving force for folding. One of the first attempts to get insights into the complexity of the protein folding process was provided by Levinthal. According to Levinthal, a protein with a 100 amino acids long would have 2^{198} conformations to sample randomly (with 198 pair of ϕ and ψ angles, and assuming only 2 conformations for each ϕ and ψ angle). If one conformation is sampled at every picosecond, it can take astronomical amount of time to sample these conformations randomly to find the correct native conformation (Levinthal, 1969). But from in vitro experiments it was known that most small proteins fold within a few minutes, and even few proteins could fold within few tens or hundreds of microseconds. This indicated that the protein folding must be guided rather than occur via a random conformational search. Experimental work that measured the folding kinetics of small proteins gave strong evidence that the proteins folds via hydrophobic collapse and subsequent formation of

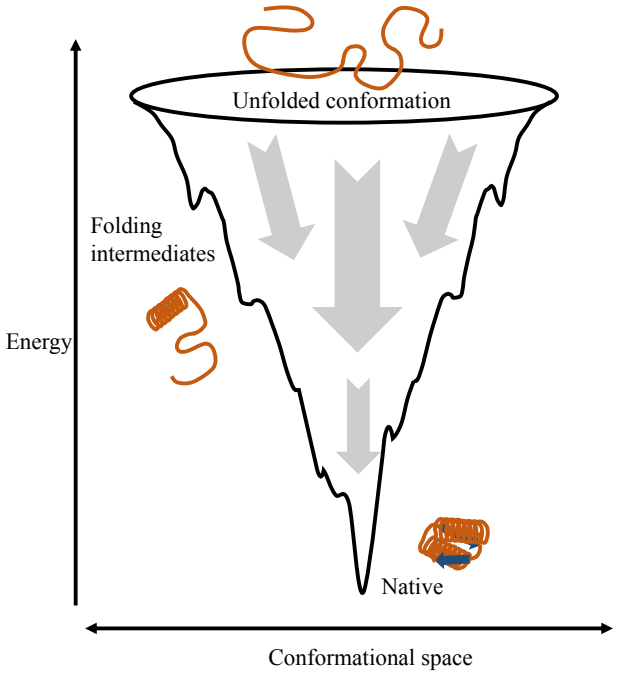

Figure 1.2 Energy landscape of Protein folding.
The presentation of a folding process as a funnel for energy landscape. A unfolded polypeptide chain has the highest energy state at the top of the funnel. The chain starts to form local structures as folding progresses, with the interactions leading to more stable conformations of lower energy. The funnel leads this progresses downhill (shown with arrows) to the native folded conformation. Folding intermediates may appear with local minima's in the energy surface. The lowest energy state is the natively folded conformation of a protein at the bottom of the funnel. Adapted from (Bryngelson, Onuchic, Socci, & Wolynes, 1995), (Wolynes, Onuchic, & Thirumalai, 1995).

intermediates on the path to folding as proposed with the Framework model (Udgaonkar & Baldwin, 1988), (Roder, Elöve, & Englander, 1988), (Agashe & Udgaonkar, 1995). As these models heavily depended on the model protein, further theoretical work to account for the complexity of the pathways to achieve the native state starting with the unfolded state as the starting point, led to the representation of the energy landscape of the protein folding process as a funnel (Bryngelson et al., 1995) (Wolynes et al., 1995). **Figure 3.2** describes the folding funnel where, the unfolded protein chain can exist in an ensemble of structures. As the global folding process depends on the formation of local contacts, secondary structures evolve leading to a lowering of the energy and a concomitant downhill progression. Distinct folding intermediates may be populated as folding progresses downhill and finally, the lowest thermodynamic energy state the native state is reached at the bottom of the funnel. Roughness on the surface of the funnel represents the complexity in forming the contacts present in the native structure. This representation can explain the complexity of parallel folding pathways, their intermixing, and can be generally applied to any given protein folding landscape (Wolynes et al., 1995). Due to the existence of parallel folding pathways, non-native interactions for example internal friction arising from salt-bridge formations, compete for the productive downhill. This leads to increased transition times for crossing the energy barrier

necessary to achieve the native state (Hoi Sung Chung, Piana-Agostinetti, Shaw, & Eaton, 2015). An additional complexity has been added to the protein folding problem with the discovery of an increasing number of unstructured regions/proteins that first fold upon binding to another biomolecule (Dyson & Wright, 2002).

There are many advantages of understanding the protein folding problem. One huge advantage is the possibility to understand various diseases out there on the protein level. Moreover, if one understands how primary sequence defines the three-dimensional structure of a protein, then one can predict the structure from the primary sequence itself. Additionally, one can design a new protein with a novel amino acid sequence that forms a particular structure for a desired function (Kuhlman et al., 2003). Recent successes by computational tools like DeepMind, Rosetta etc. using energy minimization methods to predict structures and even de novo designing of new proteins has shown immense potential in understanding the basic folding process in the near future (P.-S. Huang, Boyken, & Baker, 2016).

3.3 Protein folding in large and multi-domain proteins

Secondary elements and small proteins fold quickly within a few to a few tens of microseconds as they have a smooth energy landscape. Notably, the folding study of a β-hairpin was one of the first experimental evidence to see the fast folding and formation of a defined secondary structure (Blanco, Rivas, & Serrano, 1994). In contrast, owing to their long-range contacts and the presence of individually folding domains, large multi-domain proteins have a rugged folding landscape. The more long range interactions protein has, the more likely it is to find a local energy minimum because of the formation of unwanted contacts in the collapsed state that need to break before the native like structure. Hence, folding timescales can be on the order of milliseconds to minutes. The presence of partially folded/mis-folded states arising from a large number of hydrophobic contacts may lead to the kinetic traps with large local minima along the folding pathway (**Figure 3.3)** (Freire, Murphy, Galisteo, Sanchez-Ruiz, & Privalov, 1992) (Andrews, Capraro, Sulkowska, Onuchic, & Jennings, 2013) (Fitter, 2009) (M. B. Borgia et al., 2011) (Pirchi et al., 2011) (Walters, Mayne, Hinshaw, Sosnick, & Englander, 2013) (Roder, Elove, & Englander, 1988). From the large amount of genomic data available, bioinformatics analyses that predicted more than ~67% of proteins have two or more domains (Teichmann, Parkt, & Chothia, 1998). Folding studies on large and multi-domain proteins are still scarce and further investigation are needed. Similar complications may be present for metastable, unstructured and for multimeric proteins.

3.4 *In vivo* Protein folding and molecular chaperones

Most of what is known about protein folding comes from *in vitro* studies of chemical or temperature mediated unfolding and refolding experiments. Proteins are synthesized on ribosomes in a vectorial manner from the N-terminus part to the C-terminal end. Hence, folding of the growing polypeptide may be challenging for the N terminal regions as they lack the context of the whole protein while being synthesized. The ribosome exit tunnel may have some influence on this process (Bhushan et al., 2010). Small proteins are thought to fold spontaneously in cells as their synthesis is quick and they do not have to depend on other large part of the protein to fold. Proteins with multiple domains may fold domain-wise as they exit the ribosome (Holtkamp et al., 2015). However, multi-domain arrangements with discontinuity in sequence can have exaggerated situation as N terminal part needs to wait for the C-terminal

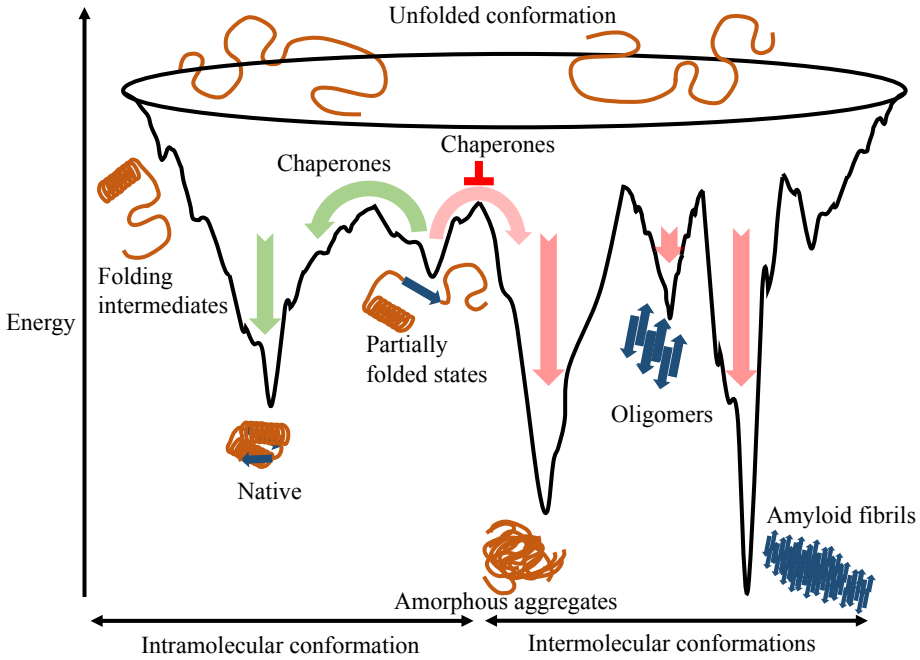

Figure 3.3 Energy landscape of a large, multi-domain protein for folding and aggregation
Small proteins fold downhill along the funnel with the presence of only a few folding intermediates to the native state. The folding landscape is simple, mostly governed by intramolecular interactions. Large, multi-domain proteins show the presence of partially folded/mis-folded states, which generally have large local minima representing kinetically trapped state on the folding pathway. These partially folded/mis-folded states may form aggregates due to the presence of multi-molecular interactions such as in the cellular environment. Aggregates are mostly irreversible as their energy is lower than the native state. Chaperones may play important roles in protein folding by assisting partially folded structures to fold or preventing them from forming aggregates in the cellular context where multiple interactions need to be avoided. Green arrows indicate the productive folding pathways and the red arrow highlight unproductive pathways leading to aggregation. Adapted from (Jahn & Radford, 2005)**,** (Hartl, Bracher, & Hayer-Hartl, 2011).

region to fold co-operatively (Kaiser, Goldman, Chodera, Tinoco, & Bustamante, 2011). Furthermore, given the crowded environment of the cell, large complex and multi-domain proteins can have problems with aggregation because of multi-molecular interactions in the long-lived partially folded/mis-folded regions of hydrophobic residues (Wright, Teichmann, Clarke, & Dobson, 2005) (Liu, Chen, & Kaiser, 2019)(Vishwas R. Agashe et al., 2004). This scenario can be visualized in the energy landscape as an even lower energy state of terminally aggregated, mis-folded protein in comparison to the native state (**Figure 3.3**). The propensity to aggregate gets crucial under stress conditions like elevated temperature. In eukaryotic cells, another layer of complexity can arise for folding in the endoplasmic reticulum due to oxidative environment.

Different levels of aggregation can lead to various disease conditions as well as in aging. Formed amorphous aggregates may lead to the formation of cross-β-structures in fibrillar aggregates when not removed in cells. These stacked ordered amyloidogenic structures have the highest thermodynamic stability and are the reason for various neuropathologies like

Alzheimer's, Huntigton disease (Chiti & Dobson, 2017), (Y. E. Kim, Hipp, Bracher, Hayer-Hartl, & Hartl, 2013).
Evolution has dealt with this situation of misfolding and aggregation by evolving an array of proteins called Molecular chaperones, a subset of heat shock proteins (Hartl et al., 2011). Molecular chaperones are considered to be a major part of protein quality control mechanisms by ensuring that the proteins do not mis-fold and aggregate during polypeptide synthesis on the ribosome. Chaperones also helps in the process of *de novo* protein folding in the cytosol, protein translocation, assembly, transport, refolding or degradation of denatured protein after aggregate dissolution.

3.5 Molecular chaperones

Molecular chaperones are proteins that directly interact with nascent proteins to help them attain the final native structure or assembly without itself being a part of the final structure. In other words, a folding/assembly catalyst. The proteins on which they act are referred to as substrates or clients and they may have unfolded polypeptide or partially folded/mis-folded properties as well as near-native or a metastable characteristics. Substrates or clients may have a stretch or stretches of hydrophobic side chains or partially folded/mis-folded structures. Chaperones are generally present at a basal concentration in cells. Their expression is induced after heat shock or other proteostatic stresses to maintain the structural integrity of various cellular proteins. Hence, they are part of heat shock protein family (Hsp). The major groups of chaperones are small Hsp (sHsp), Hsp40, Hsp60, Hsp70, Hsp90 and Hsp100, named according to their molecular weights. They are thought to have co-evolve with their clients and hence their divergence has been driven by the properties of their client. Likewise, organelle/compartment specific functions of the Hsps may represent the type of client pool they are to protect from mis-folding/aggregation. There are characteristic differences in how chaperone mechanisms take place. Most abundant mechanism involves the interaction to the hydrophobic residues of the clients and the client folding occurs free in solution. Hsp70 chaperone works on this mechanism. Distinctively, other chaperone bind to their clients and isolates them from the rest of the cytoplasm in a closed environment. This mechanism is fundamental to GroEL/ES chaperonins. Another set of chaperones takes care of aggregated clients to refold or degrade them. This is the working mechanism of small Hsps. Multiple rounds of chaperone-substrate binding and release cycles maybe required to achieve the final folding with high yield of folded protein in cellular milieu.

3.5.1 Discovery of molecular chaperones and proteostasis networks

The effect of heat shock was first studied in 1974 as a pattern of swellings (puffs) observed in the chromosomes of in salivary glands from Drosophila (Tissiéres, Mitchell, & Tracy, 1974). Later, the term "Molecular chaperones" was coined in the context of DNA chaperone nucleoplasmins (Laskey, Honda, Mills, & Finch, 1978). Bacteriophage λ replication studies in *E. coli* have already found the GroEL, DnaJ, DnaK genes playing an important role in λ-phage head assembly and DNA replication (Georgopoulos & Hohnt, 1978) (Yochem et al., 1978). Interestingly, with the discovery of an increasing number of evolutionarily conserved Hsps like Hsp90s and Hsp70s in bacteria, yeast, Drosophila, plants and humans (and in various organelles including endoplasmic reticulum, cytoplasm, and mitochondria etc.), their possible

functions were being discussed in the context of protein catalyzed protein folding and protein assembly (Pelham, 1986), (Ingolia, Slater, & Craig, 1982). The first direct evidence that "chaperonins", a class of molecular chaperones having a unique cylindrical folding chamber may have function in the folding and assembly of multimeric proteins came from measurements on mitochondrial matrix proteins and chloroplast proteins (M. Y. Cheng et al., 1989),(Ostermann, Horwich, Neupert, & Hartl, 1989) (Goloubinoff, Gatenby, & Lorimer, 1989), (Hemmingsen et al., 1988). Since then, various fundamental aspects of the molecular chaperone function such as the ATP hydrolysis cycle, substrate specificities and structural implications were studied for each class of Hsps in detail (Hartl et al., 2011)(Mayer, 2010) (Rosenzweig, Nillegoda, Mayer, & Bukau, 2019)(Schopf, Biebl, & Buchner, 2017a) (Richter, Haslbeck, & Buchner, 2010) (Nillegoda, Wentink, & Bukau, 2018). The analysis on available genomic and sequencing data from higher organisms has revealed a whole network of protein quality control (proteostasis network, PN) important for the maintenance of cellular proteome (Labbadia & Morimoto, 2015). Molecular chaperones form a central part of this network. They function in the various cellular processes like translation on ribosomes, translocation across endoplasmic reticulum (ER), mitochondria, and chloroplast. They are major players in folding, assembly, and maturation in cytoplasm, ER, mitochondria, chloroplast, and nucleus. Moreover, protein sorting, transport, signaling and proteasomal degradation also require chaperones (Preissler & Deuerling, 2012), (Baker, Tatsuta, & Langer, 2011), (Brodsky, 2012). Other members of the proteostasis network integrate the folding and assembly of proteins with quality maintenance by forming, dissolving or degrading toxic aggregates (Morimoto, 2008) (Nillegoda et al., 2018). In addition to folding/mis-folding diseases like prions, and neurodegenerative diseases, these phenomena have important implications in the cellular health and fate by their involvement in aging, cancer, various stress signaling (oxidative/reductive stress, UV etc), inflammation and infections (Labbadia & Morimoto, 2015) (Taylor & Dillin, 2011).

3.5.2 Cytoplasmic chaperones

One of the major functions of molecular chaperones is that they help *de novo* synthesized polypeptide chains (also called substrates or clients) to attain their final functional conformation on a physiological time scale. This process is regulated by dynamic binding and release of the exposed hydrophobic core of the client proteins to the Hsps. A further layer of regulation is achieved by ATP hydrolysis, the involvement of different co-chaperones or via the direct transfer of clients in between the Hsps. Typical molecular chaperones found in the cytoplasm of bacterial and eukaryotic cells cells are shown in **Figure 3.4** (Balchin, Hayer-Hartl, & Hartl, 2016). They are shown as they interact with the client protein in the order of their progression from synthesis to their final folding. Cytoplasm hosts diverse client pool and hence, it has the most divergence in Hsps, their co-chaperones and mechanisms.

Owing to the N- to C-terminal directionality of protein synthesis and the slow translation rate with respect to the time necessary for the collapse of individual domains, N-terminal parts of the translating polypeptide need to be protected from misfolding and/or aggregation processes (Kramer, Boehringer, Ban, & Bukau, 2009). Recent insights from folding studies on translating

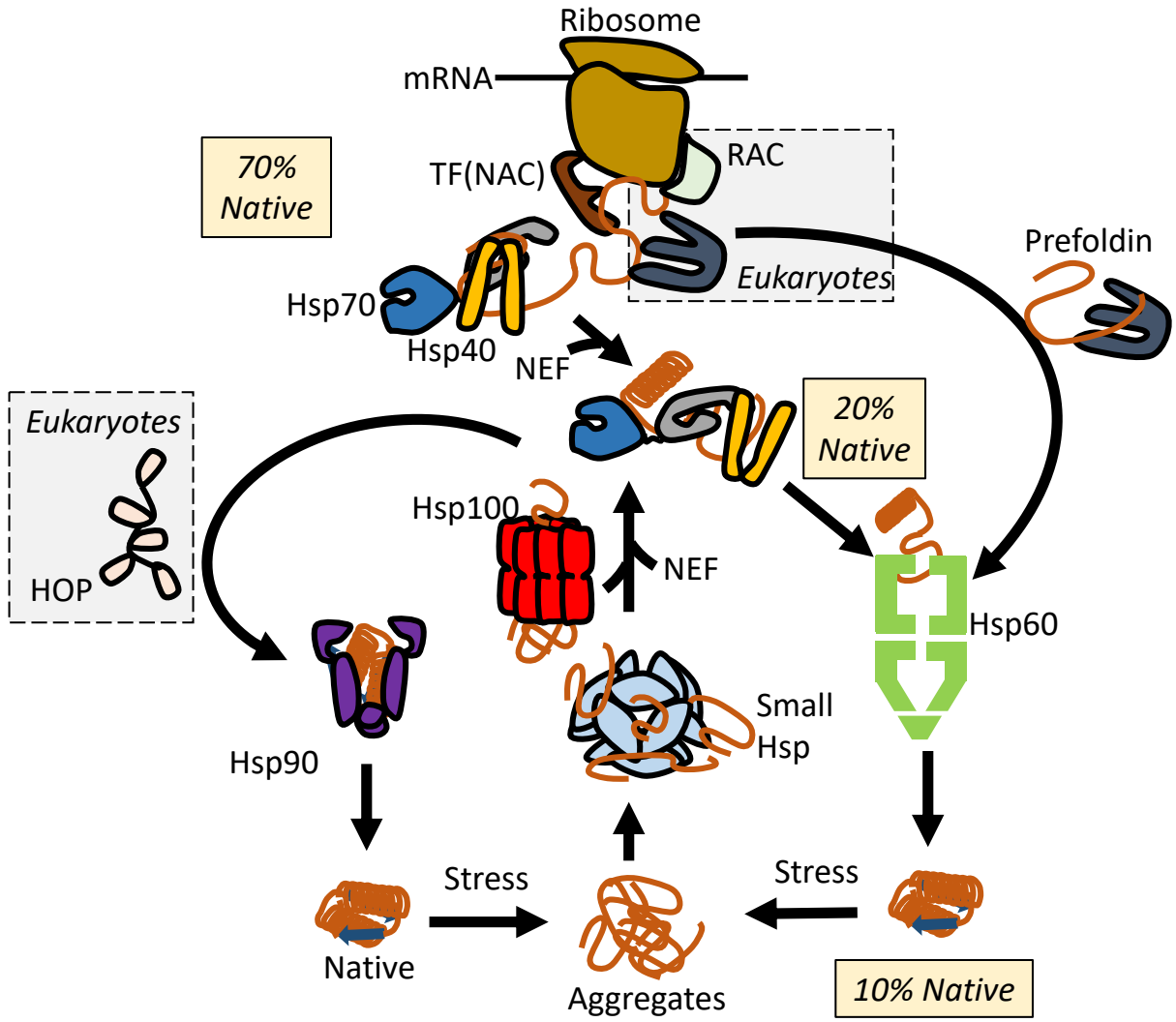

Figure 3.4 Commonly conserved molecular chaperone mechanisms found in the cytosol.

A representation of evolutionary conserved molecular chaperones residing in bacteria and eukaryotic cells. Due to added proteome complexity, eukaryotic cells have evolved additional chaperone functionalities and adaptor proteins to physically connect chaperones for efficient client transfer. Trigger factor (in bacteria) and the nascent chain associated complex/NAC (in eukaryotic cells) constitute the first chaperones encountered while the polypeptide is still being synthesized on the ribosomes. The ribosome associated complex/RAC is an extra ribosome associated specialized Hsp40/Hsp70 system and prefoldin form a direct connection to Hsp60 in eukaryotic cells, both of them are not found in bacteria. At this point, almost 70% of all proteins attain their native structure. The remaining proteins are mostly transferred to the Hsp40/Hsp70 system, where they undergo multiple rounds of chaperoning with the help of ATP hydrolysis and NEF. Here, another ~20% of all proteins fold. Those proteins that could not yet fold (~10%), are chaperoned by Hsp60. Certain proteins are metastable and needs the chaperoning from Hsp90s after Hsp40/Hsp70 pre-chaperoning. Eukaryotic cells have an adaptor protein HOP that connects Hsp70 and Hsp90 chaperones. To process aggregates or stress denatured proteins, small Hsps form heterogenous large multimeric substrate-chaperone complexes that are accessed further by Hsp100 and NEFs and may enter into the chaperone system via Hsp40/Hsp70 processing. In describing the folded protein fractions along the downstream pathways of chaperones, recently explored intrinsically disordered proteins comprising ~30% of the total proteome,

are not considered in the above description. The Figure is adapted from (Balchin et al., 2016) (Rosenzweig et al., 2019), and (Richter et al., 2010).

ribosomes interestingly showed that the polypeptide exit tunnel of the ribosome plays a role. The exit tunnel keeps the growing chain mostly in an unfolded form for a stretch of ~30-40 aa on the C-terminal end, although small secondary structures such as α-helices can form (Bhushan et al., 2010). Most of the small single domain proteins (~100 aa) fold as soon as chain translation is over (Oh et al., 2011). In bacteria, co-translational folding has implications for the translation of operonic genes. They assemble into multi-protein complexes as they are synthesized on the ribosome (Chang, Kaiser, Hartl, & Barral, 2005). Interestingly, eukaryotic cells predicted to have ~67 % more than two domains in proteins (Teichmann et al., 1998). Hence the process of domain folding on ribosomes (each domain having an average size of ~100-150 aa) may play an important role to assist domain-wise folding in multi-domain proteins (Kaiser, Goldman, & Chodera, 2011) (Holtkamp et al., 2015).

As an additional layer of support for the productive folding of translating chains, a Trigger factor bound next to the exit tunnel of the ribosome interacts non-specifically to the emerging hydrophobics patches shielding them from aggregation or mis-folding, in an ATP independent manner. Shielding leads to the folding of single domain larger in size as well as two-domain protein with a discontinuous topology. In both the cases, N-terminal part of the polypeptide chain waits for the synthesis of the C-terminal part for folding (Liu, Chen, et al., 2019)(Vishwas R. Agashe et al., 2004) (Kaiser et al., 2006). Thus, larger multi-domain protein may need the assistance of the Hsp40/70 chaperones (Oh et al., 2011). The role of Trigger factor Trigger factor is taken up by the the nascent chain associated-complex (NAC) and ribosome associated-complex (RAC) systems in eukaryotic cells with a specialization in functions, though they are structurally distinct from Trigger factor (Deuerling & Bukau, 2004). RAC are specialized Hsp40/Hsp70s that assists the co-translational folding, while NAC is able to bind to the small nascent chains (Deuerling, Gamerdinger, & Kreft, 2019). Moreover, NAC has been seen to help the ribosome to find the ER via a signal recognition particle (SRP) dependent pathway. This process is involved in maintenance of the ER protein synthesis pathway. NAC may also help to route the translation of membrane proteins of mitochondria and chloroplasts (Preissler & Deuerling, 2012).

In bacteria, the translating nascent chain on the ribosome first encounters Trigger factor chaperone. Trigger factor may be accompanied by the next chaperone system comprising DnaK (Hsp70), DnaJ (Hsp40) and a nucleotide exchange factor (NEF) GrpE, the most abundant Hsp40/Hsp70 systems found in the bacterial cytoplasm (Liberek, Galitski, Zylicz, & Georgopoulos, 1992). Further downstream chaperones in the bacterial cytoplasm include either HtpG, an Hsp90 (Sorger & Pelham, 1987), or GroEL/ES, a Hsp60 chaperonin (Hartl, 1996).

In the eukaryotic cytoplasm, counterpart of Trigger factor is replaced by both RAC and NAC (Deuerling & Bukau, 2004). Similar to Trigger factor, these complexes interact with the translating nascent chains. Various Hsp40/Hsp70 and NEF counterparts exist in eukaryotic cells as further downstream. Two Hsp90 isoforms and a TRiC/CCT a eukaryotic chaperonin, acts downstream to Hsp40/70 system in the eukaryotic cytoplasm. Interestingly, to circumvent the cytoplasmic proteome complexity present in eukaryotic cells, Prefoldin acts as an additional

evolutionary member present which transfers nascent chains directly to the eukaryotic downstream chaperone TRiC/CCT (Frydman et al., 1992) (Vainberg et al., 1998). This pathway is predominant in Archaea as they lack Hsp70s. Moreover, to route the folding of clients to Hsp90 from Hsp40/70 system, an additional component, Hsp70/Hsp90 organizing protein (HOP), is present in eukaryotic cells (B. D. Johnson, Schumacher, Ross, & Toft, 1998).

These above mentioned well-studied molecular chaperones generally interact with the folding competent clients but other classes of Hsps exist in the cytoplasm that interact with severely misfolded and aggregated proteins such as small Hsps and Hsp100s. Small Hsps are a small, ATP independent and promiscuous class of Hsps. Being an ATP independent chaperone, their chaperoning activity uniquely depends on the ability to form oligomers. They tend to change the oligomeric state and dynamic behavior in the presence of aggregated proteins (Richter et al., 2010). As they mostly bind to severe aggregates, they co-operate with other chaperones like Hsp100 dis-aggregases to dissolve aggregates and with the Hsp40/Hsp70 system to refold the dissolved proteins to their native state. A prominent member of this class is αB-crystallin, found in the eye lens of vertebrates (Nillegoda et al., 2018). ClpB is a well-studied Hsp100 member in *E. coli.* Clps (Caseinolytic peptidases) are also associated with ATP driven protease activities to route severe aggregates for degradation (Richter et al., 2010). In eukaryotic cells, a whole new network of Hsp100 exist to deal with the complex proteome. They typically act as disaggregases by dissolving the misfolded protein aggregates in association with the Hsp40/Hsp70 system (Nillegoda et al., 2018). They also help in degrading aggregates by recruiting the 26 S proteasomes.

3.5.3 ER chaperones

The bacterial periplasm and eukaryotic ER has an oxidizing environment in contrast to the reducing environment in the cytoplasm. The oxidative environment is needed to catalyze the disulfide bridges in the proteins that reside in the membrane or are to be exported or to be displayed on the extracellular side as receptors. Hence, proteins studied in the context of ER are important in terms of immune response and infectious diseases. These proteins are generally synthesized on the rough ER and helped by the highly abundant ER resident Hsp70 system called, the Immunoglobulin heavy chain binding protein (BiP) (Haas & Wabl, 1983) (Munro & Pelham, 1986). The translocation machinery of the ER play a pivotal role for the translocation of the proteins undergoing translation into the ER coordinated with BiP (Rapoport, Li, & Park, 2017). Other Hsps include Hsp90 and high molecular weight Hsp70s (Sorger & Pelham, 1987). ER Hsps are glucose regulated proteins (GRP), and hence, often referred to as Grps instead of Hsp's . BiP has its own array of NEFs and Hsp40 (ER resident DnaJ:ERdj) co-chaperones to assist the variety of client protein folding. The BiP associated protein (BAP) is a well-studied NEF. A less well-studied class of higher molecular weight Hsp70s, Grp170 exists in the ER and is often implicated in various ER diseases. It acts as a NEF as well as an Hsp70, but its precise functional mechanisms are still unclear (Behnke, Feige, & Hendershot, 2015). Other specialized chaperones exist in the ER like Protein disulfide isomerase that catalyzes the disulfide bond formation in proteins. Various quality control mechanism exists to in the ER that are important to relieve stressed conditions and misfolded proteins accumulated in ER. Unfolded protein response (UPR) pathway senses the presence of

unfolded/misfolded proteins in ER lumen and sends a stress signal to the rest of the cellular components. ER associated protein degradation (ERAD) pathway helps to route the misfolded proteins to degradation. Signals for both the pathways are mediated via transmembrane proteins (Walter & Ron, 2011)(Hwang & Qi, 2018) (Faye et al., 2019) (Brodsky, 2012).

3.5.4 Organellar chaperones of Mitochondria, chloroplasts, bacterial periplasm and the eukaryotic nucleus.

Membrane bound organelles like Mitochondria and chloroplasts have specialized, compartment specific Hsp40/70, Hsp60, Hsp90 and other Hsps (Baker et al., 2011). Endosymbiotic theory may explain their similarity to bacterial chaperones. Historically, these compartments have a special importance for the discovery of the Hsp60 chaperonin. They were part of these discoveries because their proteins are encoded in genomic nuclear DNA and thus need to be imported into the organelle after their cytoplasmic synthesis (Ostermann et al., 1989). The co-ordination between the translocon (TIM and TOM complexes) and membrane bounds Hsp70s is crucial to unfolding, and translocation of the import of proteins (Lill & Neupert, 1996). Ssc1 is a well-studied mitochondrial Hsp70. Chloroplast has cpn60 as a Hsp60 chaperonin (Viitanen et al., 1995).

The ER, the bacterial periplasm and the inter-membrane space of mitochondria and chloroplasts share the same property of an oxidizing environment. A well-known chaperone system in the bacterial periplasm is the SecYEG translocon and the SecB chaperone, that are involved aiding protein complexes that need to be exported or need to be displayed to the extracellular environment (Goemans, Denoncin, & Collet, 2014). Other chaperones in the periplasm include SurA, Skp and the recently investigated Spy to assist in oxidative folding (Quan et al., 2011). Proteases like Lon, and DegP are responsible for the quality control mechanisms in this specialized environment (Goemans et al., 2014).

Eukaryotic nuclei are known to import Hsp70s and sHsps (Hsp26) from the cytoplasm under stress conditions. Specially, the nucleolus is known to accumulate imported Hsp70s (Gallagher, Oeser, Abraham, Kaganovich, & Gardner, 2014) (Frottin et al., 2019).

3.5.5 The Hsp70 chaperone system

This class of chaperones is the most versatile, redundant, and abundantly present in all cellular compartments (e.g. cytoplasm, ER, mitochondria, chloroplast, nucleus) of all the kingdoms of life with the exception of Archaea (Boorstein, Ziegelhoffer, & Craig, 1994) (Rosenzweig et al., 2019). Due to their promiscuous nature of binding to 5-7 amino-acid hydrophobic stretches (Gragerov, Zeng, Zhao, Burkholder, & Gottesman, 1994) (Rüdiger, Germeroth, Schneider-Mergener, & Bukau, 1997), they are at the center of protein folding and proteostasis networks (Calloni et al., 2012) (Liberek et al., 1992).

Multiples of Hsp70 molecules can bind to the same client. There can be multiple isoforms present in each compartment. Typically, there is a constitutively expressed Hsp70 at a basal level (e.g. in the cytosol, Hsc70) and another isoform is overexpressed under different stress conditions (e.g. in the cytosol, Hsp70).

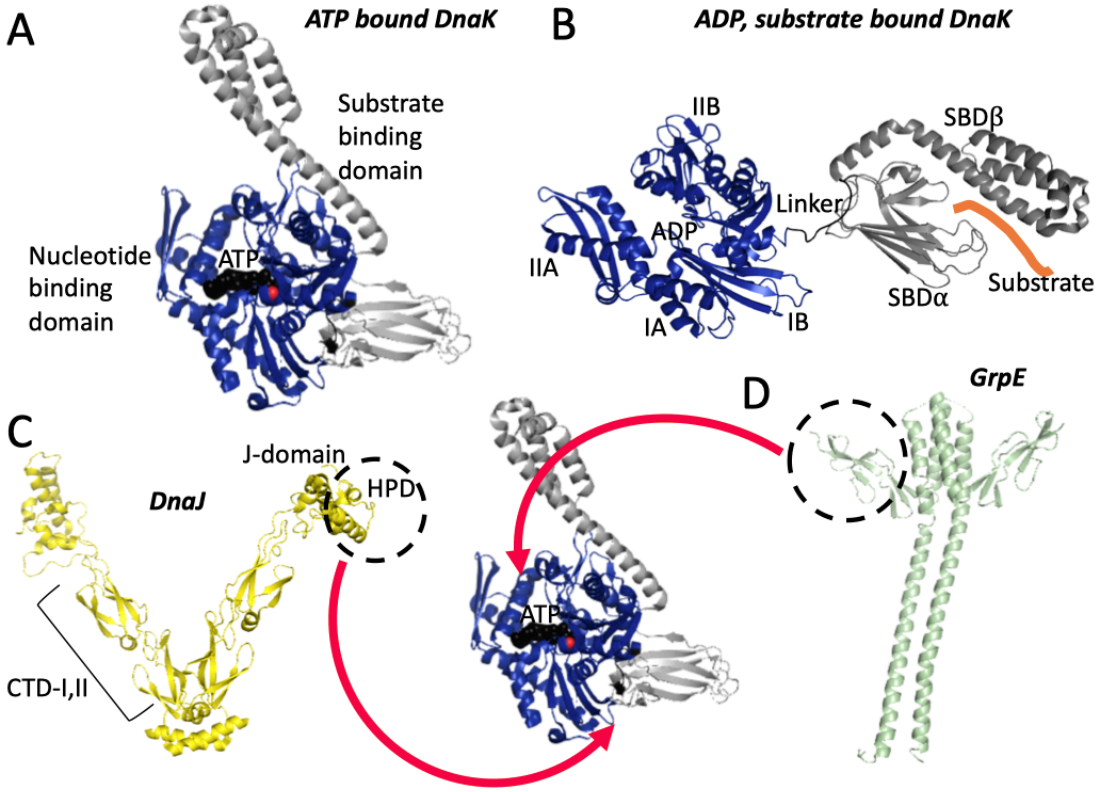

Figure 3.5 Structures of Hsp70 and important interaction partners DnaJ and GrpE.

A) The *E. coli* Hsp70 DnaK crystal structure bound to ATP. The NBD is shown in blue, the SBD in grey and ATP as black spheres. Mg^{2+} is highlighted as a red sphere. DnaK.ATP:4B9Q (Kityk, Kopp, Sinning, & Mayer, 2012) **B)** The crystal structure of DnaK bound to ADP and a substrate peptide. DnaK.ADP.peptide:2KHO (Bertelsena, Chang, Gestwicki, & Zuiderweg, 2009). NBD lobes IA, IB, IIA, and IIB as well as SBD sub-domains SBDα and SBDβ are shown. **C-D)** *E. coli* Hsp40 DnaJ and NEF GrpE crystal structures are shown in **C** and **D** respectively. DnaJ:4J8O (Barends et al., 2013), GrpE:1DKG (Harrison, Hayer-Hartl, Di Liberto, Hartl, & Kuriyan, 1997). Both are shown in their functional dimer conformation. DnaJ's HPD motif interacts with the interdomain region of DnaK and the wedge created by GrpE's β-sheet domain with the NBD (highlighted with dashed circles) (Kityk, Kopp, & Mayer, 2018) (Harrison et al., 1997).

Hsp70 Structure and conformations

Figure 3.5 A and B shows the highly conserved structure of DnaK, a Hsp70 from *E. coli*. Mainly, Hsp70 has two-domains, the N-terminal nucleotide binding domain (NBD) and C-terminal substrate binding domain (SBD) connected via a short linker. NBD, with a size of ~40 kDa, has four sub-domains with actin-like folds (IA, IB, IIA, IIB) that forms two lobes creating a deep groove for binding a nucleotide (Flaherty, Wilbanks, DeLuca-Flaherty, & McKay, 1994) . ~30 kDa SBD comprises an α-helical lid and β-sandwich sub-domain. The β-sandwich sub-domain has a groove to hold the hydrophobic stretch of the client, while the lid stabilizes the binding (Bertelsena et al., 2009). The linker provides flexibility between the domains. An additional C-terminal unstructured extension has an $(Aspartate)_2$-Valine-Glutamate (EEVD) motif that is important for interaction with other chaperones like Hsp90s via the tetratricopeptide repeat domain (TPR). Structure in **Figure 3.5 A** represents the ATP bound state. ATP binds to the NBD while the SBD lid is docked onto the NBD and is wide open, and

has a low affinity for substrate because of high association and dissociation rates. In this state, the linker is constricted between the domains (Kityk et al., 2012). On the other hand, in the ADP-bound state, the SBD lid is now closed over the β-sandwich sub-domain holding the hydrophobic stretch of a substrate. The decrease in the rate of dissociation leads to the increase in an affinity for the substrates. Moreover, the linker is stretched and the inter-domain distance increases when ADP and/or substrates are bound (**Figure 3.5 B**) (Bertelsena et al., 2009) .

Hsp70 cochaperones: Hsp40 chaperones and nucleotide exchange factors

Additional functional diversity in Hsp70 chaperones is achieved from bacteria to higher organism with the presence of multitudes of co-chaperones (Hsp40s) and NEFs (Kampinga & Craig, 2010) (Bracher & Verghese, 2015). Hsp40s stimulates the ATP hydrolysis rate of Hsp70s and NEFs assist in replacing the ATP with ADP, thereby both the cochaperones play a central role in the Hsp70 function.

Hsp40s are found ubiquitously in all life forms and bind non-specifically to hydrophobic residues. Three main classes of Hsp40s (Class A, B and C), as seen in **Figure 3.5 C** for bacterial Hsp40 DnaJ, are constructed from the characteristic helical hairpin J-domain at the N-terminus (hence Hsp40s are also called JDPs for J-domain proteins), a glycine/phenylalanine (G/F) rich flexible region and multiple domains in the C-terminal regions (CTDs) (Kelley, 1998) (Barends et al., 2013). A histidine-proline-aspartate (HPD) motif at the junction of J-domain helices interacts with both the Hsp70 linker and β-sandwich sub-domain of Hsp70 SBD, to stimulate Hsp70 ATP hydrolysis (Figure 1.9 C) (Kityk et al., 2018). Each CTD have a β-sandwich fold. Presence of two CTDs creates two polypeptide binding sites and interestingly one of the CTDs may interact with the C-terminal unstructured tail of Hsp70-SBD to help client transfer (Jiang, Rossi, & Kalodimos, 2019). The C-terminal domain helps for the dimerization of two Hsp40 protein chains. Although JDPs mainly acts as co-chaperones for Hsp70, they may act as chaperones on their own by keeping unfolded proteins in a non-aggregating form. Well-studied JDPs are DnaJ (in *E. coli*), Ydj1 (in yeast *Saccharomyces cerevisiae*) and Hdj1 (in human) (Kampinga & Craig, 2010).

Figure 3.5 D shows the structure of GrpE dimer, a bacterial NEF. Most NEFs function by opening up the nucleotide binding grove of the NBD, mainly by tilting the outer IB-IIB lobe and allowing ADP to dissociate. Interestingly, the mechanism by which NEFs function is conserved (which is by directly interacting with the NBD), although eukaryotic cells have NEFs with structural diversity. Bcl-2-associated anthonogene (BAG) domain family NEFs are common in the eukaryotic cytosol, for example BAG1 from yeast *Saccharomyces cerevisiae*. Moreover, the mechanistic divergence has occurred in NEFs during the course of evolution, where an extra domain is dedicated to expand the function by modulating the substrate-SBD binding. Examples of cytosolic NEFs in eukaryotic cells are HspBP1 and Hsp110. BAP, an ER resident NEF, has a similar mechanism to that of cytosolic HspBP1 (Bracher & Verghese, 2015) (Rosam et al., 2018) .

Conformational regulation of Hsp70 chaperone cycle

Figure 3.6 summarizes the Hsp70 functional cycle. ATP binding to the NBD domain opens the SBD lid. ATP hydrolysis leads to the closure of the SBD lid onto the substrates. The

intrinsic ATP hydrolysis rate of Hsp70s is very low (1 ATP per 6-40 min), even when bound to substrates (McCarty, Buchberger, Reinstein, & Bukau, 1995). Thus, Hsp70 alone is ineffective in stably binding substrates. Hsp40 co-chaperones typically bind to the unfolded polypeptides and transfers them to Hsp70s by direct interactions (Jiang et al., 2019). The binding of J-domains to the interdomain linker of Hsp70s stabilizes the domain dynamics acting as latch (Kityk et al., 2018) (Mapa et al., 2010) (Marcinowski et al., 2011). Via an allosteric cascade, the ATP hydrolysis rate increases (>1000-fold) which leads to a further closing of the SBD lid onto the substrate, which is bound to the β-sandwich sub-domain of the SBD (Kityk, Vogel, Schlecht, Bukau, & Mayer, 2015). This closed conformation of the SBD traps the unfolded polypeptide and hence increases the affinity of the substrate (10-50-fold higher affinity). Hydrolyzed nucleotides are dissociated from the NBD with the help of NEFs. NBD has a high affinity for ATP and hence the cycle begins with the ATP bound NBD that leads to opening of the SBD lid. This results in the release of bound peptide. Released peptide may either fold in solution or be transferred to the downstream chaperones. Bound substrate can be kept in an unfolded or partially unfolded conformation by ADP bound Hsp70. Low association and dissociation rates decrease the propensity for aggregation and allows effective folding after release.

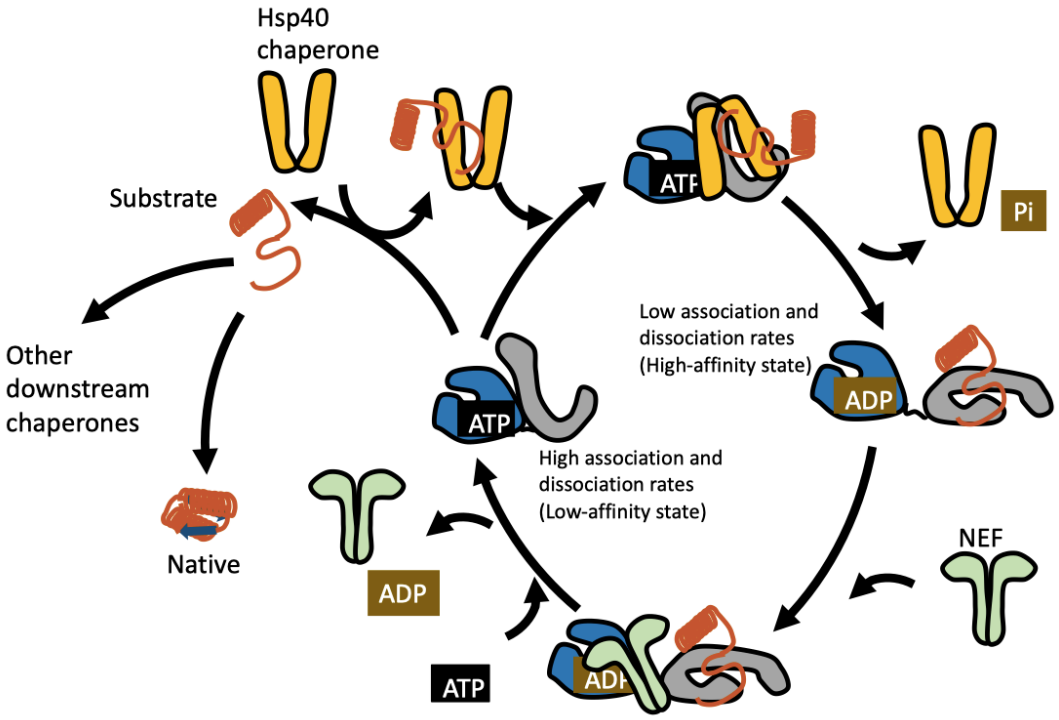

Figure 3.6 Hsp70 chaperone cycle.

ATP binding to Hsp70-NBD opens the SBD-lid (low affinity state). Partially folded protein chains are recognized by Hsp40 chaperones. By directly interacting with ATP bound Hsp70, Hsp40 transfers the substrate to Hsp70-SBD and accelerates ATP hydrolysis. ATP hydrolysis closes the lid of the SBD onto the substrate (high affinity state). NEF binding to the ADP-bound NBD of Hsp70 releases the ADP. ATP binds with high affinity to the NBD-Hsp70 after the previous ADP nucleotide has been released. ATP binding induces substrate release and the Hsp70 cycle starts over again with the low affinity state.

Released substrate may fold to a native state, be passed on to the downstream chaperones or again enter into the Hsp70 chaperone cycle. Adapted from (Rosenzweig et al., 2019) (McCarty et al., 1995).

E. coli has DnaK as the primary Hsp70, DnaJ as corresponding Hsp40 and GrpE acts as a NEF. Eukaryotic cells have expanded the functional diversity with 13 homologs of Hsp70, several Hsp40s with different classes of J-domain proteins and several NEFs. Compartment specific isoforms of Hsp70s exists with their own specific Hsp40 and NEFs.

3.5.6 The Hsp90 chaperone system

Hsp90 chaperones are unique that they mostly interact with metastable clients (i.e. client with unstable folded structures close to that of the native state) which are part of various signaling cascades. Clients includes ~60% of all kinases and steroid receptors, ~30% of all E3-ligases and ~7% of all transcription factors. Moreover, Hso90s have the most structurally divers co-chaperones among the Hsps specific to the signaling "hubs", each cochaperone being specific for a particular client (Taipale et al., 2012) (Taipale et al., 2014). Thus, Hsp90 has wide effects on cellular processes like stress response regulation, DNA repair mechanisms, neuronal signaling, development, and the immune response pathways. In the ligand-receptor signaling, they function by stabilization of native folds in receptor proteins, allowing stable ligand binding. The characteristic of Hsp90 to bind metastable clients has led to its ability to buffer mutations in the clients, although these mutations may lead to different phenotypes in the absence of Hsp90. Thus, Hsp90 is thought to be a factor in the process of evolution (Rutherford & Lindquist, 1998). In *E. coli*, HtpG (heat temperature protein G) is the only Hsp90. In the cytosol of Yeast and human cells there are constitutively expressed versions of the Hsp90 family (Hsc82 and Hsp90β respectively) and heat induced versions (Hsp82 and Hsp90α respectively) (B. Chen, Zhong, & Monteiro, 2006). Likewise, the ER, mitochondria and chloroplasts have their cognate Hsp90s (J. L. Johnson, 2012).

Structure and conformations

Structurally, Hsp90 functions as a homodimer. Each monomer is divided into the N-terminal domain (NTD), Middle domain (MD) and C-terminal domain (CTD) (**Figure 3.7 A**). ATP binds to the cleft in NTD (Prodromou et al., 1997), the surface the MD has interacts with clients and plays an important role in ATP hydrolysis. NTD and MD are connected via a charged linker. Hsp90 monomers are dimerized by the CTD (Harris, Shiau, & Agard, 2004). Additionally, the CTD has a Methionine-(Aspartate)$_2$-Valine-Glutamate (MEEVD) motif. In the absence of ATP, Hsp90 has an open "V" shaped conformation (Cunningham, Krukenberg, & Agard, 2008) (**Figure 3.7 A**). This class of Hsps have a very low affinity for ATP (K_D of ~400 μM), in contrast to Hsp70s (K_D ~10 nM) but have similarly low ATP hydrolysis rate (1 ATP per 10 min) (Scheibel et al., 1997). **Figure 3.7** describes the chaperone cycle of Hsp90. After binding of ATP to the NTD, large conformational change occurs in MD and CTD with at least two intermediate states, the closed 1 and closed 2 states (Cunningham et al., 2008). In the closed 1 state just NTDs are dimerized. In the closed 2 state, dimerization occurs over the MD region (**Figure 3.7 B**). These large structural rearrangements in the MD and NTD hold

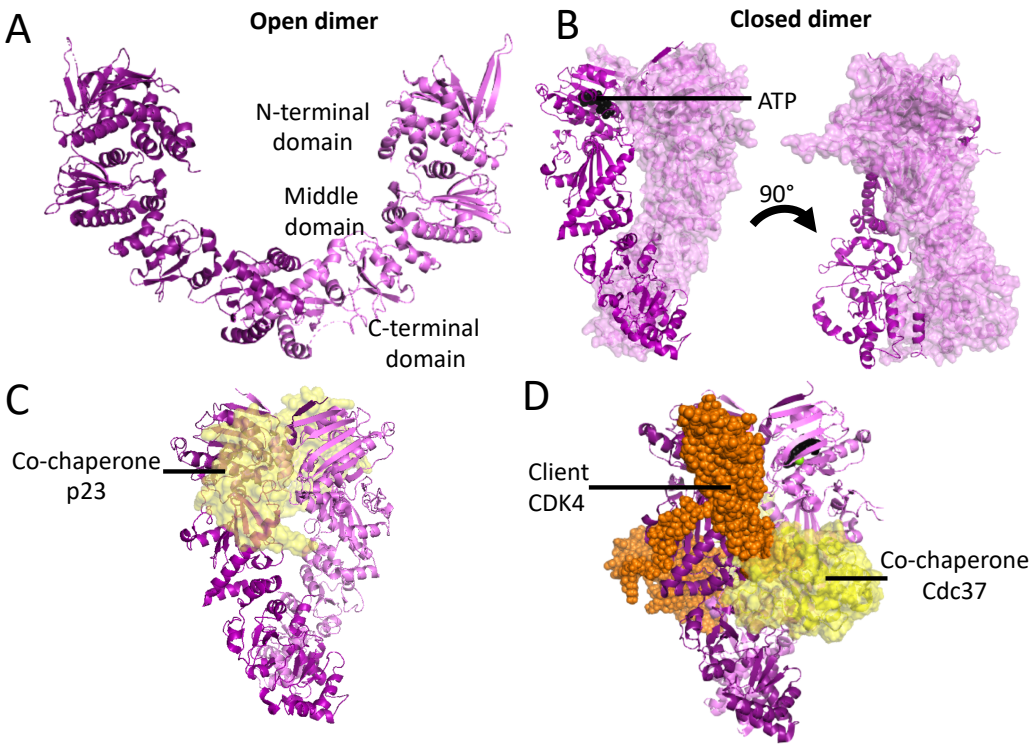

Figure 3.7 Structural insights into the Hsp90 chaperone system.

A) An open dimeric *E. coli* Hsp90 HtpG crystal structure. Each monomer has an N-terminal domain that binds a nucleotide, a middle domain with multiple promiscuous client binding sites and a dimerizing C-terminal domain. Htpg:2IOQ (Cunningham et al., 2008) **B-C)** Closed dimeric crystal structures of yeast Hsp90. **B)** bound to ATP and **C)** bound to the co-chaperone p23(yellow), where predominantly interacts with N-terminal domain. Hsp82.ATP.p23:2CG9 (Ali et al., 2006)**. D)** A cryo-EM structure of the human Hsp90 bound to the co-chaperone Cdc37 (yellow) and a client (orange). Hsp90β.Cyclin dependent kinase 4.Cdc37:5FWL (Verba et al., 2016)**.**

clients and helps stabilize them as the cycle progress. As is typical for the split ATPase family of proteins, ATP hydrolysis progresses via NTD and MD interactions. These large conformational changes in the NTD and MD are the major rate limiting steps in the Hsp90 chaperoning cycle (Meyer et al., 2003). After ATP hydrolysis, ADP is immediately released without the need of any NEF.

Conformational regulation

Complexity is added to the chaperone cycle by the cochaperones and/or clients that influence the time spent in each stage of the cycle, after ATP binding (Taipale et al., 2014) (Schopf, Biebl, & Buchner, 2017b). The details of the chaperone cycle are summarized in the **Figure 3.8**. Mainly, NTD and MD domain-interacting co-chaperones inhibits the ATPase activity by restricting the closed 1 or 2 intermediate state progression and, hence, the client has more time to mature. Typical eukaryotic co-chaperones like p23 act via prolonging the lifetime of the closed 2 stage. **Figure 3.7 C** shows how the p23 cochaperone binds to the MD-NTD region of the Hsp90 and stabilizes the closed 2 state, which hinders further progression of the cycle (Ali

et al., 2006). An activator co-chaperone, Aha1, in contrast, accelerates the last closed 2 stage. From the client's point-of-view, the type of structural stability and cooperativity involved in client folding determines its specific need for client processing (Taipale et al., 2012) (Schopf et al., 2017b). For example, kinases are mainly processed by Cdc37 (**Figure 3.7 D**) (Verba et al., 2016).

The CTD's MEEVD motif mediates interactions, with certain Tetratricopeptide repeat (TPR) containing co-chaperones, which stabilizes the open conformation and thus slows down closing of the dimers. It is important to note, that HOP possesses three TPR domains in higher eukaryotic cells (known as Sti1 in yeast), which interacts simultaneously with the MEEVD motif of Hsp90 and the EEVD motif from Hsp70 to mediate efficient client exchange (B. D. Johnson et al., 1998) (Kirschke, Goswami, Southworth, Griffin, & Agard, 2014) (Röhl et al., 2015). Chemical modifications may decrease the speed of different stages within the conformational cycle directly via the chemical modification itself or by recruiting the co-chaperones specified by the modification (Röhl, Rohrberg, & Buchner, 2013).

Yeast and human cells have similar classes of co-chaperones and may regulate the Hsp90 chaperone cycle with high similarity. There is no single co-chaperone reported yet for *E. coli* Hsp90 chaperone HtpG (B. Chen et al., 2006).

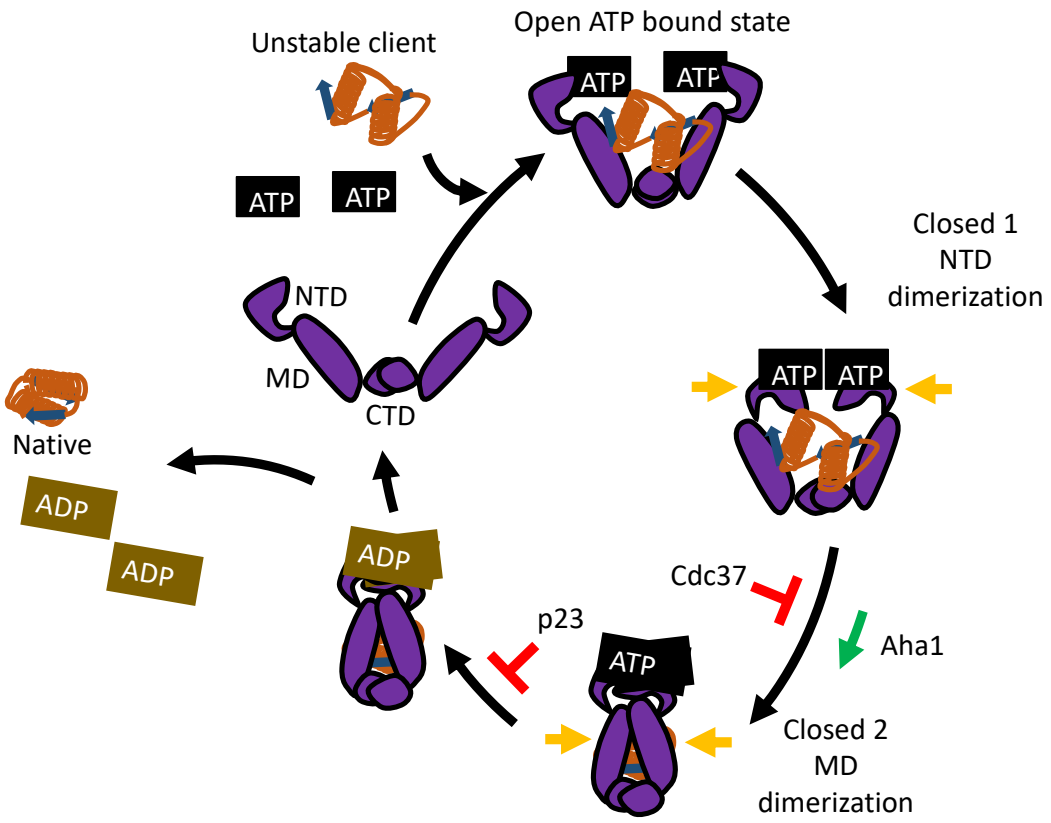

Figure 3.8 The Hsp90 chaperone cycle.

When in the apo state, Hsp90 dimers are wide open and linked by the C-terminal dimerization domains. ATP binding to the NTDs leads to a closed dimer conformation. This process is accompanied by the

capturing of the unstable clients in the binding cleft created at the dimerization interface. The closed dimer has two intermediates, closed state 1 involves the closure of NTDs, whereas the MDs close in the closed state 2. These two transitions define the time spent by the clients on the Hsp90 binding surface, which mainly spans MD and some parts of the CTD and NTD. As clients mostly have unstable structures, their stabilization is controlled via progression to later stages of the cycle. For example, the well-known Cdc37 co-chaperone blocks the conformational cycle at the Closed 1 state by binding to the MD region. The Aha1 co-chaperone, on the contrary, helps to rapidly attain the closed 2 state. The closed 2 state can be further stabilized by co-chaperones like p23. ATP binding leads to the closed 1 and 2 state transitions while hydrolysis leads to opening of the dimers and hence to the release of clients and co-chaperones. Adapted from (Schopf, Biebl, & Buchner, 2017).

3.5.7 The Hsp60 Chaperonin system

This class of Hsps has a characteristic cylindrical chamber made from multiple subunits that isolate the folding substrate from the rest of the cytoplasm to avoid aggregation. The main feature of their substrates is that they are prone to aggregation, which is most likely due to the presence of the complex folds/topology or they have a kinetically trapped state on their folding pathway (Kerner et al., 2005). Hence, Hsp60s are mostly terminal chaperones in the folding process and ~10% of all proteins are either folded directly aided by Hsp60s or are transferred to Hsp60s from other upstream chaperones (Kerner et al., 2005) (Niwa, Fujiwara, & Taguchi, 2016) (Fujiwara, Ishihama, Nakahigashi, Soga, & Taguchi, 2010). They are essential in all kingdoms of life. Typically, each chaperonin is a ~1 MDa complex with two stacked rings made from 7-9 subunits. This class of chaperones is generally divided into Group I and II chaperonins depending on their overall structural organization. Group I chaperonins are primitive, having the chamber wall and lid encoded by separate genes, and are found in bacteria (GroEL), mitochondria (Hsp60) and chloroplasts (Cpn60). Whereas, Group II chaperonins have a in-built lid with the chamber wall subunits and are found in archaea (thermosome) and in eukaryotic cells (TRiC/CCT)(Hartl et al., 2011).

Structure and conformations

Figure 3.9 summarizes the structural features of the *E. coli* chaperonin GroEL/ES. GroEL has a long 14.5 nm cylindrical chamber made from two back-to-back stacked rings. Depending on the conformation of the ring, the co-chaperone GroES can capped the open ends. Each ring has an inner diameter of 4.5 nm and an outer diameter of 14 nm and is made from seven 57 kDa subunits, arranged in a circle (Hsp60) (Braig et al., 1994). A GroES (Hsp10) lid also has seven subunits, each of ~10 kDa, arranged to form a cap-like structure (**Figure 3.9A**). Each GroEL subunit comprises an apical at the N-terminal end, intermediate hinge and an equatorial domain at the C-terminal end. Apical domains forming the opening of each ring have hydrophobic residues that binds to the exposed hydrophobic patches of a substrate protein (S. Chen et al., 1994) (Fenton, Kashi, Furtak, & Horwich, 1994). Equatorial domains possess an ATP binding site, C-terminal loops that are implicated in the substrate interactions in the cavity and forms a connection between the two rings important for allosteric communication between the rings (**Figure 3.9 B**). (Xu, Horwich, & Sigler, 1997).

Conformational regulation

The chaperonin cycle has two forms of the rings, cis and trans. The cis-ring is active, ADP bound, with substrate encapsulated, and GroES capped. The trans-ring is a nucleotide free, GroES uncapped state. The trans-ring's apical domains have substrate proteins bound via

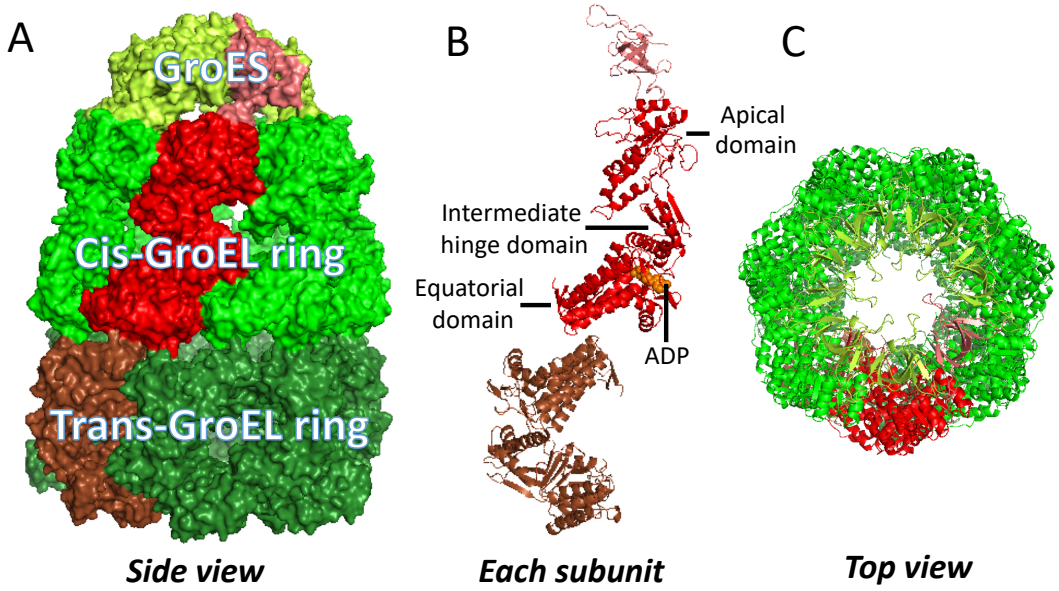

Figure 3.9 Structure of the *E. Coli* Hsp60, GroEL and its co-chaperone GroES

A) X-ray crystal structure of the GroEL/ES, Hsp60/Hsp10 chaperonin system from *E. coli*. Generally, chaperonins have two rings and each ring is capped by a co-chaperone GroES. For GroEL, each ring is made up of seven subunits stacked together side-by -side to form a cavity. A GroEL ring bound to ADP and capped with GroES froms a cis-cavity and is capable of enclosing and allows release of the client protein inside the cavity. The trans-ring is nucleotide free. One subunit of GroEL and GroES is highlighted. $GroEL_2GroES.ADP$:1AON (Xu et al., 1997) **B)** A ribbon schematic for the X-ray crystal structure of $GroEL_2GroES$. Only one subunit of $GroEL_2GroES$ is shown. Each GroEL subunit has a N-terminal apical, middle intermediate and C-terminal equatorial domain. The later domain binds to a nucleotide. The ADP bound cis conformation is shown. Cis-ring subunits are stretched to form a cavity. GroES binds to the apical domain of the GroEL capping the cavity. The negative allostery of cis and trans ring is coupled with ATP hydrolysis. **C**) A top view of a ribbon schematic of the GroEL/ES chaperonin structure shown in **A.** The central cavity and capping by the GroES are visible.

hydrophobic interactions (Fenton, Kashi, Furtak, & Horwich, 1994). As seen in **Figure 3.10**, as soon as seven ATP's occupy all seven equatorial domains, the whole trans-ring is subsequently capped by GroES on the apical domains encapsulating the substrate in the cis-cavity via an ATP mediated cascade of conformational changes transferred from the equatorial to the apical domain. This changes the trans-ring to the cis-ring conformation. The intermediate linker domain gets tilted and, apical domain gets stretched. These structural changes increase the cavity size by ~two-fold and the amino acids that line the inner cavity changes in nature from hydrophobic to hydrophilic, leading to the release of the bound substrate into the cavity (Saibil, Fenton, Clare, & Horwich, 2013). As an upper limit size, a ~60 kDa protein can fit in the cavity. ATP hydrolysis occur over ~2-10 seconds and gives the encapsulated protein the opportunity to fold in this time (Tang et al., 2006a). The transition from a trans to cis ring is allosterically coupled with the other ring (Yifrach & Horovitz, 2000). ATP binding induces the trans to cis-ring transition, the other ring previously being in the cis conformation turns into

trans-ring with release of ADP. The ADP release is accompanied by the release of GroES and the encapsulated protein from the cavity. This negative co-operativity between the rings ensures that each ring is on a different conformation (Yifrach & Horovitz, 2000). The released substrate protein may either have formed native contacts or still has partially folded structures that need to undergo another chaperone cycle.

Folding mechanism

There can be several mechanisms at work in assisting the protein to fold. In addition to isolating the protein from from cytosol, aggregation is prevented by acting as an "Anfinsen's cage".

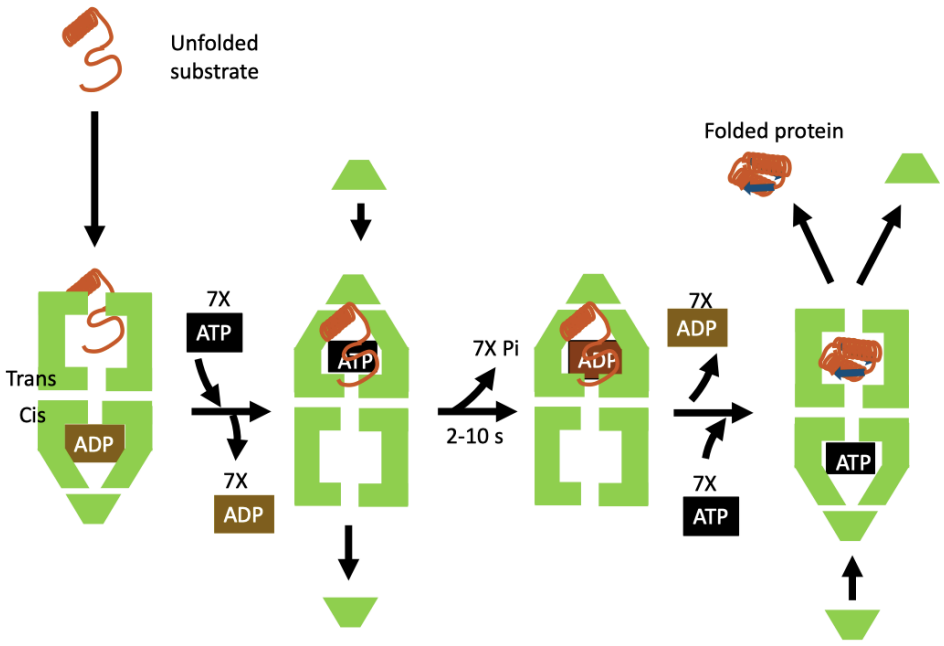

Figure 3.10 The GroEL/ES chaperone cycle.

GroEL trans-ring binds the substrates via hydrophobic interactions between the apical domain and unfolded substrate. Subsequent ATP binding and GroES capping of the trans-ring closes the cavity and encapsulates the substrate. The ring subunit rearranges into the cis-conformation leading to an increase in the cavity size. These changes are accompanied by a rearrangement of the surface exposed amino acids altering the cavity from a hydrophobic to more hydrophilic surface, allowing the substrate to fold inside the cavity until the complete ATP hydrolysis occurs. With the negative allostery between the rings, the GroES lid releases from the ADP bound ring, and the substrate to be released. The other ring goes into a cis-conformation with the start of another substrate folding cycle. When released substrate is not yet folded it may need to go through another cycle of chaperoning. Adapted from (Hartl et al., 2011)

Confinement in the cavity stabilizes the chain entropy of a folding substrate protein (J. D. Wang, Michelitsch, & Weissman, 1998b) (Apetri & Horwich, 2008) (Chakraborty et al., 2010) (Gupta, Haldar, Miličić, Hartl, & Hayer-Hartl, 2014) (Georgescauld et al., 2014). Additionally, unfolding by apical domains and then segmental release into the cavity is also thought to assist the folding, especially in case of a substrate having kinetically trapped intermediate or a mis-folded conformation on the folding pathway (Shtilerman, Lorimer, & Englander, 1999) (Lin, Madan, & Rye, 2008) (Sharma et al., 2008). Multiple rounds of the annealing and folding cycle may help to improve the folding of aggregation prone and kinetically trapped intermediates

(Weissman, Kashi, Fenton, & Horwich, 1994). C-terminal loops in the equatorial domain protruding towards the cavity are known to play a role in the folding process but the mechanism is still not clear (Tang et al., 2006a). Moreover, the mechanism of allostery between the rings in the functional cycle of GroEL/ES is still an open question. There is still a debate as to whether GroEL/ES exists predominantly in the asymmetric/bullet ($GroEL_2$/GroES) form or whether symmetric cis-form/football is possible ($GroEL_2$/$GroES_2$) (H Sparrer, Rutkat, & Buchner, 1997) (Haldar et al., 2015).

4 Materials and Methods

4.1 Materials

4.1.1 Proteins, chemicals/reagents and kits

Proteins

All the MBP protein mutants used in this thesis work were purified in our lab using known protocols explained in the later sections of the Biochemistry methods. To carry out the interaction studies of MBP with chaperones, GroEL/ES proteins were kindly provided by Prof. F. Ulrich Hartl, Max Planck Institute of Biochemistry, Munich, Germany.

For the other projects, our collaborators kindly provided us the proteins needed for the measurements. For the p53 project, we received p53-DNA binding domain (p53-DBD) labeled protein, yeast Hsp40 ydj1, and human Hsp70, Hop, Bag-1, and Hsp90β proteins from Dr. Vinay Dahiya from the group of Prof. Johannes Buchner, Technische Universität München, Garching, Germany. The U2AF65 protein was used to compare the precision and accuracy of FRET measurements as a dynamic system across the participating labs. It was generously provided by Dr. Hyun-Seo Kang from the group of Prof. Michael Sattler, Technische Universität München, Garching, Germany. To investigate the smFRET based conformational dynamics of σ^{32} protein, four different mutants were kindly supplied by Dr. Roman Kityk from the group of Prof. Matthias Mayer, Ruprecht-Karls Universität Heidelberg, Germany. Continuous provision of various mutants of DnaK, DnaJ, and GrpE first from Dr. Romain Kityk and later from Lukas Rohland of Prof. Mayer's lab was highly appreciated.

U2AF and σ^{32} proteins were further labeled with the desired dye-pairs for smFRET measurements in our laboratory.

Chemicals/reagents

Chemicals/reagents	**Company**
2-(4-(2-hydroxyethyl)-1-piperazinyl)- ethansulfonic acid (Hepes)	Sigma-Aldrich
2-Amino-2-hydroxymethyl-propane-1,3-diol (Tris)	Sigma-Aldrich
2-mercaptoethanol	Sigma-Aldrich
Acrylamide/Bis-acrylamide 37.5:1(30% w/v)	CarlRoth
Adenosine triphosphate	Sigma-Aldrich
Agar-Agar	Sigma-Aldrich

Agarose	Sigma-Aldrich
Alexa-488 maleimide	Thermo Fischer Scientific
Alexa-568 maleimide	Thermo Fischer Scientific
Alexa-647 maleimide	Thermo Fischer Scientific
Ammonium per sulfate	Sigma-Aldrich
Ampicillin sodium salt	Sigma-Aldrich
Amylose resin	New England Biolab
Atto-488-azide	Atto-tech
Atto-532 maleimide	Atto-tech
Atto-565 maleimide	Atto-tech
Atto-643 maleimide	Atto-tech
Atto-647N maleimide	Atto-tech
Bovine Serum Albumin	New England Biolab
Chloramphenicol	Sigma-Aldrich
Coomassie Brilliant Blue R-250	Thermo Fischer Scientific
Copper sulfate ($CuSO_4$)	Merck
Deoxyribonucleotides	Metabion and Biomers
Dithiothreitol (DTT)	Sigma-Aldrich
Glycerol	Sigma-Aldrich
Guanidinium hydrochloride (GuHCl)	Sigma-Aldrich
Isopropyl β- d-1-thiogalactopyranoside (IPTG)	Sigma-Aldrich
Luria broth base	Sigma-Aldrich
Magnesium chloride ($MgCl_2$)	Sigma-Aldrich
Maltose monohydrate	Sigma-Aldrich

N-Propargyl-L-Lysine (PrK)	SiChem
N,N,N',N'-Tetramethylethylenediamine (TEMED)	CarlRoth
Ni-NTA super flow	Thermo Fischer Scientific
Nickel sulfate-hexahydrate	-
Pfu DNA polymerase	Thermo Fischer Scientific
Phenylmethylsulphonyl fluoride (PMSF)	Sigma-Aldrich
Phosphate buffer saline (PBS) pellets	Sigma-Aldrich
Potassium chloride (KCl)	Sigma-Aldrich
Protease inhibitor cocktail pellets	Sigma-Aldrich
Ribonucleotides	IBA Lifesciences
RNA polymerase	New England Biolab
Sodium ascorbate	Sigma-Aldrich
Sodium azide	Sigma-Aldrich
Sodium chloride (NaCl)	Sigma-Aldrich
Sodium dodecyl sulfate (SDS)	Sigma-Aldrich
Tetracycline hydrochloride	Sigma-Aldrich
Tris-((1-benzyl-1-H-1,2,3-triazo-4-yl)-methyl)-amin (TBTA)	Sigma-Aldrich
Tryptone	Sigma-Aldrich
Urea	Sigma-Aldrich
Water	New England Biolab

Kits

Kits	**Company**
BirA500: BirA Biotin-protein ligase standard reaction kit	Avidity
Phusion site-directed mutagenesis kit	Thermo Fischer Scientific
Pierce BCA protein assay kit	Thermo Fischer Scientific
Plasmid Miniprep kit	QIAGEN
Taq DNA polymerase PCR kit	QIAGEN

4.1.2 Buffer solutions, and media

All the buffers were filter sterilized using a 0.4 µm filter and the media were autoclaved.

Buffer/Media	**Use**	**Components**
Annealing buffer	dsDNA annealing	10 mM Tris, 1 mM EDTA, 50 mM NaCl, pH 7.0
De-staining buffer	De-staining SDS-PAGE gels	10% Acetic acid, 40% Ethanol
SDS-PAGE Gel loading buffer	A sample buffer for SDS-PAGE	100mM Tris, 4% SDS (w/v), 0.2% Bromo Phenol Blue (w/v), 20% Glycerol (v/v), 200 mM DTT
HKM buffer	p53- measurements	50 mM Hepes, 150 mM KCl, 10 mM $MgCl_2$, pH 7.5
Luria Bertani broth	Culturing of *E. coli*	15.5 g of Luria broth base (10 g/L Tryptone, 5 g/L Yeast extract, 0.5 g/L NaCl) for 1 liter
Cell lysis buffer	Cell lysis for protein purification	PBS, 10% Glycerol, 2mM DTT, 1mM PMSF/PIC
MBP refolding buffer	MBP refolding assay	20 mM Tris, pH 7.5, 20 mM KCl
MBP unfolding buffer	Complete unfolding of MBP	20 mM Tris, pH 7.5, 20 mM KCl, 3 M GuHCl, 5 mM DTT
RRM buffer	U2AF65 labeling, storage and measurements	20 mM KH_2PO_4 and K_2HPO_4, 50 mM NaCl, pH 6.5

SDS-PAGE running buffer	Running a SDS-PAGE gel	25 mM Tris base, 250 mM Glycine, 0.1 % SDS
σ^{32}-buffer	σ^{32} storage, labeling and measurements	10 mM Hepes, 100 mM KCl, pH 7.9, 10% glycerol
Staining buffer	SDS-PAGE gel staining	0.1% Coomassie Brilliant Blue R-250, 10% Acetic acid, 40% Ethanol
Tris EDTA (TE) buffer	DNA storage and measurements	10 m M Tris, 1mM EDTA, pH 8.0
Tris acetate EDTA (TAE) buffer	Agarose gel electrophoresis	4.85 g Tris, 1.14 mL glacial acetic acid, 20 mL 0.5 M EDTA (pH 8.0) make up the volume to 1 Liter with ddH_20
LB-agar plates	Growing *E. coli* colonies after plating	15.5 g/L of Luria broth base in 15 g/L agar

4.1.3 Plasmids, constructs and bacterial hosts

The bacterial host system of *E. coli* was used for all the cloning, plasmid propagation and protein expression. Specifically, we used commercially available TOP10 cells based on *E. coli* DH5α strain often used for cloning and plasmid amplification. Manually prepared competent cells based on the *E. coli* DH5α strain were sometimes used. For protein expression and purification purposes, *E. coli* BL21 DE3 or BL21 AI strains were transformed with expression plasmids.

Most of the cloning and purification was involved in the MBP folding project. For this purpose, pCH-series (pET-22b) based plasmid encoding DM-MBP (V8G, Y283D) was generously provided by Prof. F.Ulrich Hartl and Prof. Manajit Hayer-Hartl (MPI of Biochemistry, Martinsried, Germany).

pET-22b DM-MBP plasmid was first amplified in TOP10 cells. For the cloning and mutagenesis purposes, *E. coli* TOP10 or DH5α strains were used as mentioned above. The pET-22b DM-MBP construct was further mutated to single cysteine (A52C, P298C) or double cysteine mutants for the domain constructs (NTD: A52C-P298C, CTD: K175C-P298C, N-C interface: A52C-K175C). All the three DM-MBP domain constructs were back mutated to MBP with the reversing both the double mutations from V8G, Y283D to G8V, D283Y. These served as a wt-MBP control for smFRET experiments. All these mutants were generated using site-directed mutagenesis (Thermo Scientific-Phusion Site-Directed Mutagenesis Kit). Mutagenesis steps were performed one at a time and the mutation was confirmed after sequencing. A52TAG-K175CP298C mutant was generated for three-color FRET experiments on DM-MBP. This was created by mutating the codon for alanine at position 52 to the amber

stop codon (TAG) under the background of the DM-MBP CTD construct. This enabled us to incorporate an unnatural amino acid (UAA) at 52 position.

All single and double cysteine DM-MBP mutants and wt-MBP double cysteine mutant proteins were expressed in *E. coli* BL21-AI (L-(+)-arabinose controlled T7 RNA polymerase expressing strain).

To enable unnatural amino acid incorporation the pEvol-tRNA Pyl PylRS WT plasmid was used, originally cloned by Prof. Edward Lemke, Johannes Gutenberg University, Mainz, Germany. This pyrrolysine based plasmid was used for expressing tRNA and amino acyl-tRNA transferase enzyme for Propargyl-L-Lysine needed for an unnatural amino acid incorporation. A list of all the plasmids and primers used is given in the Appendix.

4.2 Molecular biology methods

4.2.1 Bacterial transformation of plasmid DNA

Various *E. coli* strains were transformed either with template plasmid or a PCR amplified ligated plasmid. Transformation in the *E. coli* DH5α strain was utilized for generating enough plasmid for mutagenesis purposes. *E. coli* BL21DE3 or Bl21 AI strains were transformed or co-transformed with the expression plasmids.

The cells used for the transformation were commercially available *E. coli* DH5α TOP10 cells or a manually prepared competent cells of *E. coli* DH5α and *E. coli* BL21 DE3 or *E. coli* BL21 AI. The manual prepared cells were made competent chemically by the use of $CaCl_2$. The basic principle of the method is to make bacterial membranes fragile and thereby form pores by the use of salts at cold temperatures. First, the overnight grown *E. coli* DH5α, *E. coli* BL21 DE3 or *E. coli* BL21 AI cells were gently centrifuged at 6000 rpm for 5 min. All the *E. coli* growing steps were done at 37°C and 200 rpm. All the centrifuging bottles, small tubes and $CaCl_2$ solutions were sterilized and kept throughout on ice. Centrifuged pellets were resuspended in 0.1 M $CaCl_2$ solution and kept for 30 min on ice. Afterwards, at the end of the centrifugation step, the pellet was resuspended in small volumes and snap frozen with liquid nitrogen.

For manually prepared competent cells, the transformation efficiency was checked using a standard plasmid with known amount. The cells were streaked on the agar plates with an antibiotic selecting for the plasmid bearing cell.

To carry out the transformation, competent cells were first thawed on ice and all the steps were performed with a gentle care. First cells are mixed with ~10-20 ng of plasmid DNA and the tube was tapped few times. The mixed suspension was kept on ice for 30 min to make a charged contact between the $CaCl_2$ and negatively charged DNA to be formed on the cell membrane. Afterwards, a heat shock of 42°C was given to the cells for 45 s. Heat shock opens the pores on the outer membrane allowing the bacteria to take up the DNA electrostatically stabilized on the surface. Placing these heat shocked bacteria on ice again for at least 5 min allows the pores to close and trap the DNA inside the cell. Bacterial cells were then plated on LB-agar plates containing the antibiotic against which the plasmid provides the resistance. Transformed cells for plasmids bearing Ampicillin resistance were plated directly on a prewarmed plate while cells transformed with a plasmid containing other selection genes, were first grown for 1 hr and then plated on a LB agar plate. Transformed cells form a colony from an individual cell,

which can be observed after overnight growth and marks a successful transformation. All the transformants in DH5α, BL21 DE3 or BL21AI cells for all the constructs were stored in a -80 freezer in 50% glycerol in LB-medium.

4.2.2 Plasmid preparation

The template plasmids were first prepared in large quantities before mutating them. To check for successful cloning, grown colonies after transformation were used to prepare larger amounts of the plasmids that are then sequenced. For both the purposes, *E. coli* DH5α strains were used.

Plasmids were prepared using the Plasmid Miniprep kit from QIAGEN. The kit uses the standard methods of lysis and precipitation of the genomic DNA to due to its big size whereas the plasmid DNA remains soluble and can be isolated in a next step after running through a positively charged column. Cells were first grown overnight in 10-20 mL LB medium containing the respective antibiotic. This culture is harvested from a single colony picked from the agar plate. According to the protocol, grown cells are pelleted for 3 min at 10,000 rpm and resuspended in P1 buffer. This buffer is isotonic such that bacterial cells are resuspended uniformly which enhances the efficiency of a plasmid preparation. As a next step, the cells were lysed in P2 solution containing NaOH by slowly inverting the tube a few times. The reaction is allowed to proceed for a maximum of 5 min to avoid breaks in the genomic DNA that would interference in purification of the plasmid DNA. This step is followed by addition of the N3 solution which neutralizes the basic pH and allows the genomic DNA to aggregate. The aggregates can be easily separated by centrifugation step for 10 min at 13,000 rpm. Plasmid DNA remains in the solution and is purified by passing the solution through a column that binds plasmid DNA. After washing steps with 100% ethanol, plasmid DNA is eluted with TE buffer or nuclease free water. Plasmid DNA is quantified via absorption spectroscopy at 260 nm using the NanoDrop spectrophotometer (ND-1000, Thermo Fischer Scientific, USA). Typical yields were between 200-400 ng/µL for 10-20 mL of grown cells.

4.2.3 Site-Directed Mutagenesis

Site specific mutations are created especially for either testing a hypothesis with respect to a protein system or introducing cysteines to facilitate fluorophore labeling. Point mutations were achieved by the use of the Phusion Site-directed mutagenesis kit from Thermo Fischer Scientific. Primers were designed with a length of ~25-30 nt and the point mutation to be introduced was given in the center surrounded by at least 10-15 nt perfectly matching on each side. Each primer pair was amplified on a plasmid template in opposite directions to generate many copies of plasmids with the introduced mutation. Primers were synthesized from Metabion or Biomers and PAGE purified. Melting temperatures for both forward and revers primers were matched and optimized to be close to the annealing temperatures of 71°C used for the PCR. To set up a mutagenesis reaction, template plasmid DNA was kept at 10 ng in a 50 µL of a reaction volume. Phusion buffer was added from the stock solution to a final concentration of 1x concentration. Mixture of dNTPs was added keeping the concertation of each NTP at 200 µM. Both primers were added to a final concentration of 0.5 µM each. Finally, 0.02 enzyme Units/µL Phusion Hot start DNA polymerase enzyme was added. The reaction volume was filled to 50 µL by adding water.

This reaction was followed by PCR. For PCR, the template DNA was first denatured at 98°C for 30 s. Then, 25 cycles of 10 s denaturation at 98°C, 30 s annealing at ~71°C and 2 min extension at 72°C were carried out. The final extension was kept at 72°C for 5 min.

Mutagenesis PCR amplified plasmid DNA was confirmed by its detection on an agarose gel. For this purpose, 1% agarose gel is prepared in TAE buffer. A 1kb ladder was used to compare the size of an amplified plasmid. A template plasmid was also loaded for comparison. The sample was mixed with a 6X DNA gel loading buffer. Detection of DNA was done under UV-illumination after staining the agarose gel with SyBr gold solution.

After successful amplification, 5 µL of the amplified reaction volume was ligated in a ligation buffer provided in the Phusion Site-directed mutagenesis kit. T4 DNA ligase was preadded in the buffer system. The ligation reaction was incubated at least for 10 min at RT. After ligation, the ligation product was transformed into *E. coli* DH5*α* cells.

4.2.4 Confirmation of a desired mutation

After transformation of ligated products, colonies were observed the next day on LB-agar plates containing the selection antibiotics. At least 3 random colonies were selected and grown overnight for their plasmid generation. To check for incorporation of the desired mutation in the clones, the plasmid region was amplified with a sequencing primer spanning the mutation sequence. The following PCR protocol was adopted for carrying out amplification of the plasmids with the *Taq* DNA polymerase PCR kit from QIAGEN.

The PCR reaction mixture was similar to the Mutagenesis PCR reaction except the fact that a total of 2.5 enzyme Units of *Taq* DNA polymerase was added per reaction volume of 100 µL. The thermal cycler PCR protocol was as follows. Initial denaturation was performed for 3 min at 94°C followed by 35 rounds of amplification steps. Each amplification step involved a 1 min denaturation at 95°C, 1 min annealing between 55-68°C and 1 min extension at 72°C. Final extension was for 10 min at 72°C.

The amplified product was checked again with an agarose gel and sent for sequencing for further confirmation. Sanger sequencing was ordered either from the Genomics Service Unit of Biology Department of LMU or from GATC Biotech. For the Genomics Service Unit, 7 µL of amplified DNA sample with a concentration of 20-50 ng was needed in 10 mM Tris/Cl, pH 8.0 without containing EDTA. The sequencing primers were required to be between 18-25 nt in length with a melting temperature between 52-60°C and a concentration of between 2-10 pmol. For GATC Biotech, an online portal was used to submit the entries with barcoded sequencing reactions. Samples were submitted in 20 µL volumes with a 10-50 ng/µL concentration. Sequencing primer conditions were similar to the Genomics Service Unit. Both the sequencing facilities provided the results in *.ab1*or *.seq* format, which were readable with Chromas Lite software. Upon confirmation of a specific mutation, the protein clone was transformed in *E. coli* BL21 DE3 or BL21 AI cells and subsequently expressed.

4.3 Biochemistry methods

4.3.1 Determination of protein concentrations

Protein concentrations for both purified proteins or proteins provided by collaborators were determined using an absorption spectroscopy on a Nano-drop spectrophotometer. Concentration of the protein, c, can be determined from the Beer-Lambert relationship:

$$A_{280} = \varepsilon c l \tag{4.1}$$

where A_{280} is the absorption over a path length of l and ε is the absorption extinction coefficient of a particular protein at 280 nm.

For very accurate measurements and for proteins having low absorbance at 280 nm, the BCA protein estimation kit was occasionally used with the provided standard of BSA.

4.3.2 Sodium dodecyl sulphate-Poly acryl amide gel electrophoresis (SDS-PAGE)

SDS-PAGE was used routinely to check the quality and presence of proteins after their induced expression, purification, and labeling and also their stability. SDS-PAGE uses SDS as a denaturing agent, which decorates the denatured polypeptide with a uniform negative charge per amino acid. The protein is allowed to migrate through a gel matrix of specific pore size made from the acryl-amide polymer. Hence irrespective of its native charge, protein mobility in an electric field solely depends on its size. To differentiate the proteins of different sizes, the pore size is controlled by adjusting the proportion of a matrix monomer acryl amide and a cross-linker bis-acryl-amide. The sample was prepared in gel loading buffer with a few µg of protein, determined from the known concentration of protein . The sample was boiled for 15 min.

Usually proteins were detected on using a 12% separating gel. A typical recipe for a 12% SDS-PAGE gel is as follows. Separating gel: 3.75 mL of water, 2.8 mL of 1.5 M Tris-SDS, 4.7 mL of 30% Acryl/Bisacrylamide, 13 µL TEMED and 45 µL 20% APS. Stacking gel: 2.9 mL of water, 1.2 mL of 0.5 M Tris-SDS, 0.8 mL of 30% Acryl/Bisacrylamide, 6 µL of TEMED and 30 µL of 20% APS.

For calibration a pre-stained protein marker of 10-180 kDa (#26617, Thermo Scientific) was used. Gels were run in SDS-PAGE running buffer with a BIO-RAD apparatus. For staining and de-staining, specific buffers were used as indicated.

4.3.3 Protein expression and purification

For all MBP protein expressions, stored transformed cells were freshly streaked on an LB-agar plate containing suitable antibiotics. For the best results, a fresh transformation was performed. All the cells were grown at 37°C and 200 rpm shaking conditions in LB-medium. For MBP selection, 100 µM ampicillin was incorporated in all the growth media. Expression checks were done for all purifications, where at least 3 colonies from a plate were selected and grown overnight in 10 mL LB-media. On the next day, a fresh inoculum of the cells in 10 mL of media was done from the overnight preculture at a 1:100 dilution of cells. After growing cells for ~1 hr, one verified that the O.D. was less than 0.6, and then the culture was induced for at least another 4 hr. Cells were pelleted and analyzed using SDS-PAGE. Expression profiles were

checked for the induced cells compared to the uninduced control. The best performing clones were selected for the large-scale expressions.

MBP proteins without unnatural amino acid incorporation were expressed in *E. coli* BL21 DE3 cells. Their plasmid was based on the pET-22b backbone and thus, BL21 DE3 are suitable hosts for expression (Tang et al., 2006b). For purification purpose, expression was done in 1 L of culture. From the 50 mL overnight grown pre-inoculum, 5 mL of cells were added to the 1 L LB-media containing ampicillin. Cells were grown for ~1.5 hr to reach an O.D. of 0.5-0.6. 1mM IPTG was added at his point to start the induction. After induction of protein expression for 4 hr, cells were harvested in big centrifugal bottles by centrifuging them at a speed of 6000 rpm at 4°C for 30 min. Harvested cells were resuspended in 15 mL Lysis buffer. For sonification, a with 50% duty cycle and output of 3 on a sonicator instrument was used to disrupt the cells on ice. Upon complete lysis of the cells, the solution turns clear and pale yellow in color after about 3-4 cycles of 10-15 min. Cell debris was separated with centrifugation at 16,000 rpm for 30 min at 4°C. The supernatant was collected for further purification steps. As MBP (i.e., maltose binding protein) binds maltose, we used an amylose resin to purify MBP proteins in a single-step affinity-chromatography approach. The supernatant was mixed with ~1.5 mL of amylose resin beads and allowed to mix overnight at 8°C in the fridge with a constant slow stirring. On the next day, the mixture was allowed to flow through a PD-10 purification column (~10 mL). Amylose resin beads stayed at the bottom of the column by the use of a filter while the liquid was allowed to flow through. After 4 washing steps, each with 50 mL lysis buffer, MBP proteins were eluted with a lysis buffer containing 100 mM Maltose in 5 steps of 1 mL each. Purity was checked for eluted proteins via SDS-PAGE. Eluted volumes were pooled together and concentrated with centrifugal filters with a 10 kDa cut-off filter. Proteins were washed for at least 4 times (each ~10 mL) with lysis buffer to remove maltose before snap freezing and storing them at -80 after adding 15% glycerol to the solutions.

For MBP proteins containing the amber stop codon (TAG) at position 52, unnatural amino acid incorporation was enabled by the use of an orthogonal translation system of Pyrrolysine (Pyl) WT tRNA and Pyrrolysine-tRNA synthetase (PylRS). To this end, pEvol-tRNA Pyl PylRS WT plasmid containing the above orthogonal translation system was co-transformed with a MBP protein expression plasmid with the amber mutation in *E. coli* Bl21 AI cells (A. Chatterjee, Xiao, & Schultz, 2012). Dually transformed cells were selected by the combination of two antibiotics, 50 µM of ampicillin for MBP and 33 µM of chloramphenicol for pEvol-tRNA Pyl PylRS WT selection. Cells were precultured, inoculated in 1 L of media as for MBP proteins. Expression of MBP with an unnatural amino acid (UAA), in our case N-Propargyl-L-Lysine (PrK) was achieved by adding 1 mM PrK in LB-media 40 min prior to induction. The PrK stock was prepared in 0.1 M NaOH. In *E. coli* Bl21 AI cells, MBP biosynthesis was induced via 1 mM IPTG and to facilitate the UAA incorporation, 0.02% arabinose was used to induce PylRS expression. The yield was low compared to the other MBP protein expressions due to the low rate of successful UAA incorporation at the TAG stop codon. Purification was performed following similar steps as for the other MBP proteins with cell harvesting after 4 hr of induction, cell lysis, and amylose resin affinity purification.

4.3.4 MBP refolding assay

For understanding the processes involved in MBP refolding, various approaches were applied to analyzed the MBP refolding. For example, increase in the tryptophan fluorescence was taken as a measure of folding progression. Folding in the NTD, CTD and N-C interface were analyzed with smFRET.

For MBP refolding experiments, 500 nM-2 μM MBP protein was first unfolded in MBP unfolding buffer containing 3 M GuHCl and 5 mM DTT at 50°C for at least 30 min. Complete unfolding was verified using equilibrium measurements of tryptophan fluorescence as well as by smFRET measurements. To induce refolding, unfolded MBP was diluted from unfolding buffer into refolding buffer with at least a 30-fold dilution to have maximum of 0.1 M GuHCl in the refolding buffer. Lower concentrations didn't change the kinetics significantly. Overall, with tryptophan fluorescence the kinetics of the refolding was found to be 22 min for DM-MBP, while is similar to the rate obtained for different domain constructs analyzed with smFRET separately as well as in previous studies (Tang et al., 2006b) (Sharma et al., 2008) (Chakraborty et al., 2010).

Kinetic tryptophan fluorescence measurements were performed at 20°C on our FLS1000 Photoluminescence Spectrometer (Edinburgh Instruments). A data point was collected every 60 s and for measured for 2 s. Intrinsic tryptophan fluorescence was induced using excitation at 290 nm with a slit width of 2 nm and the fluorescence emission was detected at 345 nm with a 5 nm slit width. Photobleaching was avoided by adjusting the slit widths, measurement time intervals and acquisition times. For equilibrium tryptophan measurements, steady state tryphophan signal as recorded.

To analyze the kinetics of MBP refolding with smFRET, refolding measurements in 0.1 M GuHCl were analyzed every 5-10 min for 2 hr. The fraction of folded protein at each data point was extracted by quantifying the fraction of molecules having FRET efficiency similar to the native protein. Th fraction of folded protein was plotted against the time and the data fit a mono-exponential model to obtain the folding rate.

4.4 Fluorophore labeling methods

4.4.1 Fluorophore labeling using Cysteine-Maleimide chemistry

All the purified MBP proteins containing single or double cysteine mutations were labeled using cysteine -maleimide chemistry to enable coupling of fluorophore at the desired position. Typically, 50-100 μM MBP protein concentration was used to achieve efficient labeling. The sulfhydryl groups of MBP cysteine mutants were reduced with 10 mM DTT in phosphate buffered saline (PBS) at RT for 20 min. Excess DTT was removed by at least 4 washing steps, each step with a 10-fold dilution, using 10 kDa cut-off Amicon centrifugal filters (Merck-Millipore) with deoxygenated PBS containing 50 μM tris(2-carboxyethyl) phosphine (TCEP). This basal amount of TCEP keeps cysteines in the reduced state. The maleimide-fluorophore conjugate was added in ~3-fold molar excess to the washed protein solution and the reaction was incubated at RT for 3 hr in the dark. To carry out two-color labelling, equimolar maleimide fluorophore conjugates were added simultaneously in the reaction. To avoid excess of donor-only labels for smFRET experiments, in the first step, protein can be labeled with ~0.7

stoichiometric amount of acceptor, followed by the addition of ~2-fold excess of donor-maleimide conjugate. The remaining free fluorophores were removed by washing with MBP refolding buffer containing 1 mM DTT in centrifugal filters. Labelling was either checked by SDS-PAGE or using FCS with more precision to access whether any free dye remained. FCCS was applied to confirm double-labelling. Labeling reactions were quantitated by calculating the degree of labelling (DOL) by absorption spectroscopy as:

$$DOL = \frac{A_F \times \varepsilon_P}{(A_{280} - A_F \times CF_{280}) \times \varepsilon_F} \quad (4.2)$$

where A_F is the absorption at a λ_{max} of a fluorophore. ε_P is the extinction coefficient of a protein at 280 nm whereas ε_F is a fluorophore extinction coefficient at its λ_{max}. A_{280} is the absorbance of a protein at 280 nm to judge the overall protein concentration and CF_{280} is the correction factor for fluorophore ansorption at 280 nm. DOLs of the single labeling reactions were ~70-80% for maleimide-cysteine chemistry. Accordingly, for double labeling, overall 50-65% DOL was achieved. The total fraction of the protein with both donor and acceptor labeles was 25-32.5%, including the labels on the different positions.

4.4.2 Fluorophore labeling using Click chemistry

Click chemistry has the advantage that it does not require a cysteine replacement in a protein as it can be applied when indigenous cysteines are important for protein function. Given the advantage, disadvantage are the low yield and use of harsh labeling conditions that might be a challenge for keeping the protein stable over the course of labeling. Being one of the efficient labeling reactions, click chemistries are beneficial. We carried out the copper catalyzed azide-alkyne cyclo-addition (CuAAC) reaction between fluorophore-azide and protein alkyne conjugates. This reaction requires the reducing agent $CuS0_4$, TBTA as a ligand to stabilize the copper, TCEP to keep the reducing environment and sodium ascorbate which starts the reaction by reducing the copper. These reaction conditions were first optimized for MBP stability. After optimization, we used 200 µM $CuS0_4$, 200 µM TBTA, 50 µM TCEP, and 200 µM sodium ascorbate for labeling ~50 µM of A52PrK MBP protein (alkyne-protein) with 3-fold excess of Atto488-azide (Tyagi & Lemke, 2013). The reaction was carried out at RT for 3 hrs in the dark under mild shaking condition. Unreacted dye was removed by washing with the MBP refolding buffer in centrifugal filters. The efficiency of the reaction was determined to be more than 90%.

4.4.3 Specific labeling of the Maltose Binding Protein with three fluorophores

Specific labeling of the MBP-A52PrK-K175C-P298C protein was desired to perform three-color smFRET refolding measurements. To achieve the specific labeling of all three fluorophores, one position was labeled with specific click-chemistry and the other two positions were labeled with cysteine-maleimide chemistry where one of them was preferred under the specific conditions.

In the first step, the alkyne group of A52PrK was specifically conjugated to the Atto488-azide with CuAAC. ~120 µM of MBP-A52PrK-K175C-P298C protein was used in this case to start with a high concentration of protein added, to achieve a high yield. The reaction was carried out as mentioned above for 3 hrs in the dark. The excess of Atto488-azide dye was then

removed. As a second step, ~70 µM of 52PrK-atto488 labelled MBP-A52PrK-K175C-P298C was reduced by 10 mM DTT and washed with de-oxygenated PBS containing 500 mM maltose. It has been shown that maltose binding in the inter-domain cleft of MBP results in the burial of domain-interface including P298 position (Sharma et al., 2008)(Mapa, Tiwari, Kumar, Jayaraj, & Maiti, 2012a). 2-fold excess of Alexa-647 was added in the reaction mixture and preferential labeling was done only for 1 hr to specifically label position K175C. For the last third labeling step, excess Alexa647 dye and maltose were removed with a centrifugal filter, 3-fold excess of Atto565 maleimide was mixed with the remaining protein. This enabled coupling of Atto565 to P298C, which is now available for labeling in the absence of maltose. Coupling of each dye after each labelling step was monitored by measuring the absorption of both the respective fluorophores and the protein at 280 nm.

4.5 Single-molecule methods

The application of single-molecule methods is seminal in interpreting biomolecular processes in an unprecedented way. Intensity fluctuations on single or double-labeled molecules can be correlated to analyze the underlying processes on different relaxation timescales. This analysis is performed with concentration on the picomolar to nanomolar range. Fluorescence fluctuations in in a single fluorophore allows to study translational motions by covalent attachment to the interest of molecule, rotational motions, photophysical properties etc. When two spectrally different fluorophores are used and coupled specifically to the molecule under study, the distance change between the labeled positions in molecule and the associated conformations molecule undergoes can be analyzed. The various processes that can be studied includes protein folding, inherent flexibility in protein structure, and binding interactions, etc. SmFRET has the main advantage of having the ability to resolve different populations that are generally averaged out in ensemble assays. To measure FRET at the single molecule level, very low concentrations, on the order of a few tens of picomolar of labeled particles, are used to minimize multi-molecular events that complicate the final calculation of the parameters. Typically, smFRET is either measured on the diffusing molecules with a confocal microscope or on the immobilized molecules via total internal reflection fluorescence microscopy (referred to as smTIRF). The diffusion-based approach catches the fast fluctuations in the structure present on the order of the diffusion times of a few milliseconds. On the contrary, immobilization provides a way to observe conformational changes over an extended period of a few minutes but misses the nanosecond to sub-millisecond dynamics. The limit of the observation time window is set by the acquisition method. While diffusion smFRET is measured with photodiodes with a time resolution of a few tens of picoseconds for detected photons, smTIRF is measured with the cameras hence ~5-10 ms is typically the count but can go faster upto 0.25 ms.

First, a brief introduction to confocal microscopy and the combined implementation of a pulsed interleaved excitation scheme with multi-parameter fluorescence detection is given. Various approaches and parameters extracted from these implementations are briefly explained. Later, smFRET with two and three-fluorophores as well as Fluorescence Correlation Spectroscopy, used in this thesis work, are discussed.

4.5.1 Single molecule FRET measurements in solution

To measure the fluorescence signal from diffusing molecules in the solution, confocal microscopes are typically employed. In confocal microscopy, a pinhole used to select a certain focal plane of the laser spot is selected (Marvin Minsky, 1957). This has the advantage of a small observation volume (~fL) where single molecules can be observed. The lateral resolution (d) obtained with a confocal microscope is given by:

$$d = 0.61\frac{\lambda}{NA} \tag{4.3}$$

where λ is the wavelength of light used, $NA = n \times \sin(\theta)$, the numerical aperture that defines the light collection efficiency of an objective where θ is half the conical angle created by an objective and n is the index of refraction.

The realization of confocal microscopes for measuring single molecules labeled with organic fluorophores, typically uses lasers to induce the fluorescence. Technically, donor excitation is sufficient to assess the FRET efficiency in single-molecule measurements and single bursts can be analyzed using donor fluorescence lifetime decays as well (C. Zander et al., 1996) (Ashok A. Deniz et al., 1999). For smFRET, the FRET efficiency calculation includes the number of detected acceptor photons. The assumption is made that the acceptor is active all the time during a burst. Hovever, this is not be true whether acceptor abruptly photobleaches or blinks during a burst. To circumvent these situations and to check the performance of an acceptor dye, a second laser, directly exciting the acceptor was introduced (Kapanidis et al., 2004). This implementation was done with continuous-wave laser sources that were alternatively switched between the donor and acceptor excitation cycles on the µs timescales, hence the name Alternating laser excitation (µsALEX). Later, an important extension using the alternation of nanosecond pulsed laser sources accessed the fluorescence lifetime which helped to perform smFRET measurements more accurately. This extension is called pulsed interleaved excitation (PIE), or nsALEX (Müller, Zaychikov, Bräuchle, & Lamb, 2005).

4.5.2 Pulsed interleaved excitation (PIE)

When combining PIE with time-correlated single-photon counting (TCSPC), each photon is stamped with a macro-time and a micro-time which are analyzed to calculate various parameters explained later in the chapter (**Figure 4.1)**. The optimal pulse rate depends on the fluorescence lifetime of the used fluorophores. As a rule of thumb, an interval of 4-5 times the fluorescence lifetime of the unquenched state should be used.

4.5.3 Multi-parameter fluorescence detection with Pulsed interleaved excitation (MFD-PIE)

During the single molecule detection, the wealth of information is enhanced by measuring polarization of the detected photon in addition to the wavelength. This approach was implemented in a multi-parameter fluorescence detection (MFD) (C Eggeling et al., 2001). PIE can provide the information on the fast times of nanoseconds as well as the information on the acceptor, thereby its combination with MFD makes possible to calculate multiple parameters

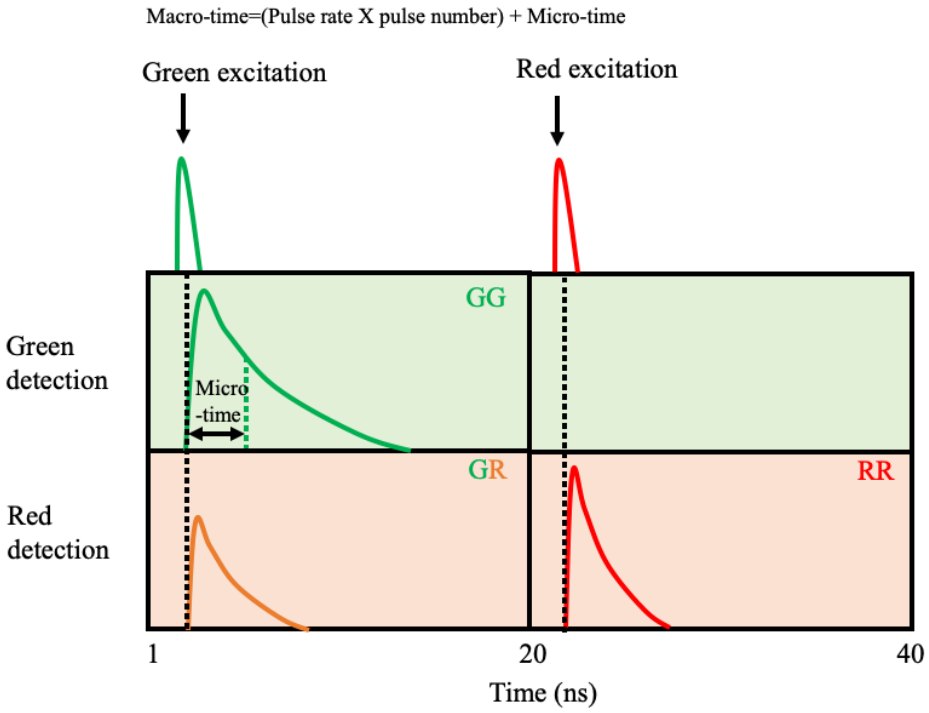

Figure 4.1: Pulsed interleaved excitation (PIE). Green and red lasers are pulsed at a rate of 20 MHz rate. After each excitation, photons are detected in the separate green and red detection. Upon green excitation, the green photons are detected in GG channel and the red photons are detected in GR channel. Red photons after red excitation are detected in the RR channel. The micro-time is the time between photon arrival and the preceding laser pulse. The macro-time is the total measurement time.

like FRET efficiency, labeling stoichiometry, fluorescence anisotropy and lifetime of an acceptor, in addition to the typical information available with MFD like fluorescence anisotropy and lifetime of a donor (Kudryavtsev et al., 2012). A typical confocal set-up combining both the utilities is depicted in **Figure 4.2**.After exciting the sample with a plane polarized laser, fluorescence is separated by a polychroic mirror before detection. Collected fluorescence is then passed through the pinhole towards the detection system. The detection system consists of avalanche photodiodes, which detect arriving photons after separating them first for their polarization via polarizing beam splitter and later for their wavelengths by dichroic mirrors.

All the two-color smFRET measurements performed in this thesis work were carried out on a MFD-PIE set-up. More specifically, a set-up was built around a Nikon Eclipse inverted microscope. For green excitation, frequency-doubled 1064 nm laser (PicoTA, Toptica) with an excitation wavelength at 532 nm and for red excitation a pulsed laser LDH-D-C640 operating at 640 nm from PicoQuant were used. Both the lasers were synchronized and operated at 26.7 MHz with a delay of ~18 ns with respect to each other. The power for both the lasers was set to ~100 µW. A 60x water-immersion objective from Nikon (Plan Apo IR 60x1.27 WI) was used to collect the fluorescence. A pinhole of 75 µm dimeter made it possible to employ confocal detection and afterwards was separated into parallel and perpendicular light using a

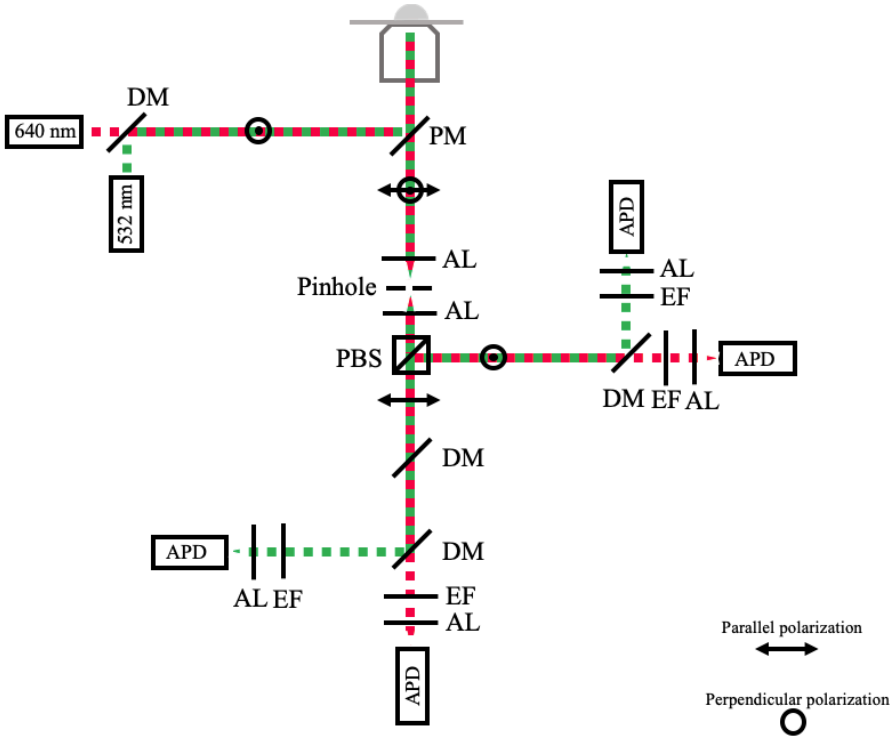

Figure 4.2: Schematics of a two-color PIE-MFD set-up. 532 nm: nanosecond pulsed green laser, 640 nm: nanosecond pulsed red laser, DM: dichroic mirror, PM: polychroic mirror, AL: achromatic lens, PBS: polarizing beam splitter, EF: emission filter, APD: avalanche photodiode.

polarizing beam splitter PBS3 from Thorlabs. For each polarization pathway, the green light is separated from red by a DualLine z532/635 dichroic mirror from AHF Analysetechnik. Emission filters for green (Brightline 582/75 from Semrock) and red (HQ700/75 from Chroma) were used before being detected on single-photon counting (SPC 154) avalanche photodiodes (APD) from Perkin-Elmer. Furthermore, they are synchronized with the TCSPC cards from Becker and Hickl (Kudryavtsev et al., 2012).

All the analysis performed in this work was carried out in a custom-written software in Matlab called PIE analysis with Matlab (PAM) and is available at https://www.gitlab.com/PAM-PIE/PAM (Schrimpf, Barth, Hendrix, & Lamb, 2018). All the modules used in this work are written by Dr. Anders Barth.

To analyze diffusion-based smFRET measurements, single-molecule events are typically identified from the burst of photons detected, when a labeled molecule diffuses through the laser spot. The process of analyzing these identified single-molecule bursts is called Burst Analysis (C. Zander et al., 1996).

First, the bursts are identified using a threshold either set by the inter-photon time (Lee filter,) or by setting the number of photons detected in a specific time window (Nir et al., 2006b). The inter-photon time that distinguishes the region of the photons as a part of a burst can be adjusted from inter-photon times of a background counts. In the latter case, a threshold of a minimum of 50-100 photons per burst and 5-10 photons per time window of 500 µs is generally applied. An all photon burst search (APBS), where burst selection is based on the minimum number of photons in any detection channel. Another parallel approach is the dual-channel burst search

(DCBS), where applies thresholds to the photons detected in both the detection channels and hence is only sensitive to doubly-labeled molecule only.

An APBS is helpful for identifying various labeling stoichiometries, to distinguish between singly and doubly labeled species, as well as other minor populations arising from bleaching or blinking events of one of the dyes. Even higher-order stoichiometries can be detected with the APBS approach.

After correcting for background in all the channels, various parameters are derived for each burst (Kudryavtsev et al., 2012) (N. K. Lee et al., 2005) (C. Zander et al., 1996) (Schaffer et al., 1999). The background correction is applied to the total number of detected photons (N) from the average background counts ($\langle B \rangle$) as:

$$F = N - \langle B \rangle \times t_{burst} \tag{4.4}$$

where, t_{burst} is the time duration for each burst. The background correction is applied for each detection channel to compute the final signal, F.

The most important parameters are the FRET efficiency and Stoichiometry. The uncorrected FRET efficiency also called Proximity Ratio (PR) is given by the fluorescence signal detected in the acceptor channel after donor excitation over the total signal after donor excitation:

$$PR = \frac{F_{DA}}{F_{DD} + F_{DA}} \tag{4.5}$$

where F_{XY} is the fluorescence signal in Y channel after X excitation.

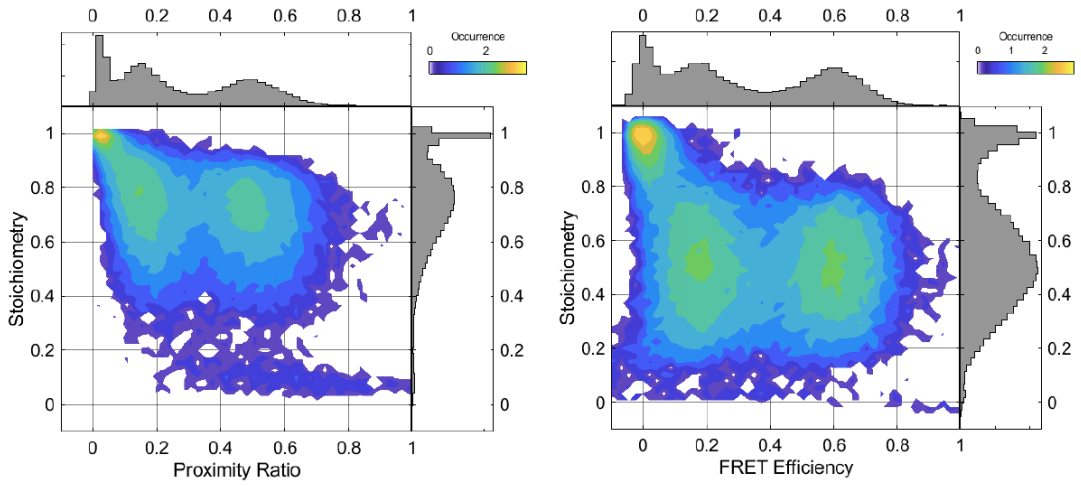

Figure 4.3: A comparison of uncorrected and corrected stoichiometry versus FRET efficiency plots for a measurement on a 40 base-pair long dsDNA labeled with Atto532 and Atto647N, where the labels were separated by 10 and 15 base-pairs. Left: Stoichiometry versus proximity ratio (uncorrected FRET efficiency) for all the measured molecules using an All photon burst search. Note the cross-talk and direct excitation of an acceptor before the corrections, visible with donor-only and acceptor-only sepcies. Right: Background, cross-talk, direct acceptor, and γ corrected stoichiometry versus FRET efficiency plot.

For stoichiometry (S), the total signal after donor excitation is normalized to the total signal after both donor and acceptor excitations as:

$$S = \frac{F_{DD} + F_{DA}}{F_{DD} + F_{DA} + F_{AA}} \tag{4.6}$$

Uncorrected data is shown in Figure 4.3, left panel.

Various correction factors are applied to account for the set-up and the performance of the used fluorophores under the experimental conditions.

One of the correction factors, α accounts for the spectral cross-talk in the acceptor channel after donor excitation. Hence, by definition, it is the signal detected in the acceptor channel after donor excitation for the donor-only population when no acceptor signal is present when using ALEX or PIE, given by:

$$\alpha = \frac{F_{DA}^{D-only}}{F_{DD}^{D-only}} \tag{1.7}$$

Another necessary correction is to account for direct excitation of the acceptor after donor excitation, δ. This correction factor is calculated by normalizing the signal in the acceptor channel after donor excitation over the signal detected in the acceptor channel after acceptor excitation for acceptor only molecules:

$$\delta = \frac{F_{DA}^{A-only}}{F_{AA}^{A-only}} \tag{4.8}$$

The last important factor is the detection correction factor, γ. As the name suggests, it corrects for the difference in the sensitivity of detection between donor and acceptor fluorophores including quantum yield differences between them. It can be written as:

$$\gamma = \frac{\eta_A . \phi_A}{\eta_D . \phi_D} \tag{4.9}$$

where η is the detection efficiency and ϕ is the quantum yield. This factor is the most challenging to correct for and the best approach depends on the specific sample. It is not easy to quantify quantum yields from the single-molecule experiments and may not remain the same for the different FRET efficiency states. The detection efficiency is defined by the dichroic filters and optical components used for each detection channel. A simple approach is to assume that the quantum yields of both donor and acceptor do not change for all the FRET efficiency states. In this case, the stoichiometry for double-labeled molecules with different FRET efficiencies is set to the same value to correct for the differences in the detection efficiencies (N. K. Lee et al., 2005). This correction is either applied to the background, cross-talk and direct acceptor excitation corrected data or on the raw photon data (Coullomb et al., 2020).

Laser excitation powers may differ and can be corrected for by setting the stoichiometry value to 0.5 after correcting for γ. Another approach is based on the donor fluorescence lifetime information when the FRET efficiency states are static, the correct γ should result in a linear inverse relationship between donor lifetime and FRET efficiency according to:

$$E = 1 - \frac{\tau_{D(A)}}{\tau_{D(0)}} \tag{4.10}$$

In this case, care has to be taken that appropriate corrections are made if the acceptor lifetime changes between the different FRET efficiency states.

Finally, after applying all the corrections, the corrected stoichiometry (S) and FRET efficiency (E) are given by:

$$S = \frac{\gamma F_{DD} - \alpha F_{DD} + F_{DA}}{\gamma F_{DD} + F_{DA} - \alpha F_{DD} - \delta F_{AA} + F_{AA}} \tag{4.11}$$

$$E = \frac{F_{DA} - \alpha F_{DD} - \delta F_{AA}}{\gamma F_{DD} + F_{DA} - \alpha F_{DD} - \delta F_{AA}} \tag{4.12}$$

All the above-mentioned correction factors are applied to all the bursts including, donor-only, double-labeled, and acceptor-only molecules. Corrected data is plotted in Figure 4.3, right panel. As FRET efficiency calculations only makes sense for the donor-acceptor labeled molecules, to select the dual-color labeled species, a DCBS search is one option. Hence, at the beginning when sorting bursts only double-labeled molecules are selected. However, the DCBS procedure may include molecules having phot-bleaching and blinking events. These events can be further filtered out with the use of $T_{GG} - T_{GR}$ and $T_{GX} - T_{RR}$ filters (Christian Eggeling et al., 2006) (Kudryavtsev et al., 2012), where T_{XY} represents the mean arrival time of the photons after X excitation detected in Y channel. Both the filters remove all the events leading to the acceptor bleaching or blinking, where the $T_{GG} - T_{GR}$ filter removes these events after green excitation and $T_{GX} - T_{RR}$ filter after both excitations. A $T_{GG} - T_{GR}$ filter also removes the dynamic events and hence the $T_{GX} - T_{RR}$ filter is preferred.

When APBS is performed, the ALEX-2CDE filter is used to sort out double-labeled molecules. This filter is calculated based on the photon densities for the photons that arrived in both the detectors simultaneously. Lower thresholds <15 are generally applied to select double-labeled species only. As a general approach, stoichiometry thresholds between 0.3-0.7 can also be applied for corrected stoichiometries in addition to the above approaches. See Figure 4.4.

4.5.3.1 Burst-wise fluorescence lifetime and dynamics

Additional important parameters obtained with MFD-PIE are the fluorescence lifetime and anisotropy. These parameters provide the information regarding performance of fluorophores and can be helpful to interpret the underlying FRET efficiency or a stoichiometry variation in the case observed.

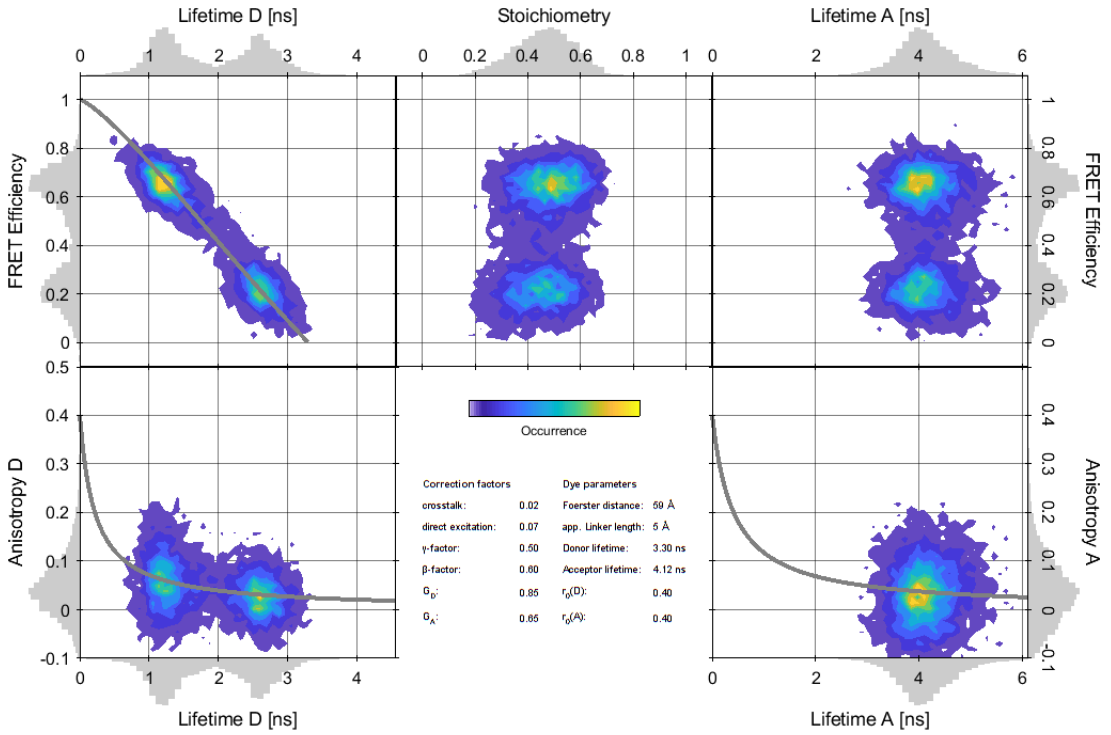

Figure 4.4: MFD-PIE plots for the measurements on dsDNA labeled with Atto532 and Atto647N. The data was filtered with the ALEX-2CDE filter with a value of 8. **Upper plots:** Left plot shows the FRET efficiency versus donor lifetime plot. The static-FRET line is shown in grey. The middle plot is a FRET efficiency versus stoichiometry plot. The right plot shows the FRET efficiency versus the acceptor lifetime plot. **Lower plots:** Left panel shows steady-state donor anisotropy versus donor lifetime plot. A fit to the Perrin equation is shown as a grey line. All the applied correction factors and used parameters are mentioned in the middle panel. The right panel depicts the steady-state acceptor anisotropy versus acceptor lifetime. A fit to the Perrin equation is shown as a grey line.

The fluorescence lifetime for each burst is challenging to estimate accurately due to the low photon numbers found in single-molecule experiments which are on the order of 50-100 photons. Fluorescence lifetime is generally a convolution for the instrument response function (IRF) with an exponential function as:

$$I(t) = IRF(t) \otimes exp\left(-\frac{t}{\tau}\right) \tag{4.13}$$

where $I(t)$ is the measured fluorescence intensity decay and $\otimes$ is the convolution operator. To increase the accuracy of a lifetime with low photon statistics, a maximum likelihood estimation is typically used to estimate the lifetime more accurately as shown previously (C. Zander et al., 1996). The donor fluorescence lifetime can be very useful when dynamics are present in the system. When the system under study has two distinct FRET states with donor lifetimes τ_1 and τ_2 , the intensity-derived average lifetime, $\langle\tau\rangle$ is biased to weigh long lifetimes because the number of photons in low-FRET is higher than for the high-FRET state. This bias can be seen as a change in the ideal relationship between τ and FRET efficiency (E) from the eq. (7) as follows (Kalinin, Valeri, Antonik, Felekyan, & Seidel, 2010):

$$E = 1 - \frac{\tau_1 . \tau_2}{\tau_{D(0)}[\tau_1 + \tau_2 - \langle\tau\rangle]} \tag{4.14}$$

4.5.3.2 Burst-wise fluorescence anisotropy

Burst-wise steady-state anisotropy information carries information on the rotational flexibility of dyes and the molecules to which the dyes are attached. To calculate the anisotropy accurately, differences in the detection efficiencies in the parallel ($\eta_{\parallel}$) and perpendicular ($\eta_{\perp}$) channel needs to be corrected for:

$$G = \frac{\eta_{\perp}}{\eta_{\parallel}} \tag{4.15}$$

where G is the grating factor. For this purpose, fast-rotating molecules like dyes are used to measure the counts in both the parallel and perpendicular channels to determine the detection efficiencies. Afterwards, the steady-state anisotropy is correlated to the lifetime of a dye by using the Perrin equation:

$$r = \frac{r_0}{1 + \frac{\tau}{\rho}} \tag{4.16}$$

where r_0 is the fundamental anisotropy and ρ is the rotational correlation time. For free dyes, it is on the order of ~0.2 ns and for larger molecules, it shifts to ~1-3 ns depending on the size. For an exemplary plot see Figure 4.4. If the dye is stuck to the attached molecule and cannot rotate freely, longer rotational correlation times are observed.

4.5.4 Three-color smFRET measurements in solution

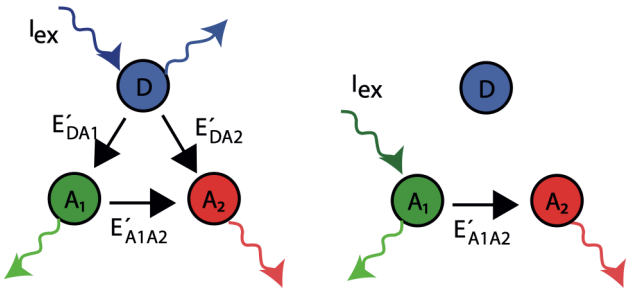

Figure 4.5: A scheme for three-color smFRET. Adapted from (Barth, Voith Von Voithenberg, & Lamb, 2019a).

4.5.4.1 Three-color FRET

Three-color FRET has an advantage over two-color FRET is that three distances can be simultaneously analyzed at a time in a single molecule. In Ttis way, one can investigate the correlated motions in a molecule, which is common in biological systems. To measure three-color FRET, one of the approaches is shown in Figure 4.5. This approach is developed and discussed in the study Barth, Voith Von Voithenberg, & Lamb, 2019. The equations and derivations below are adapted from this study. A_1 and A_2 form a two-color FRET system where

A_1 acts as a donor for A_2 and A_2 as an acceptor. With the introduction of a donor (D), both A_1 and A_2 now act as acceptors for the donor. Hence, the original designation as A_1.

4.5.4.2 Overview of the calculations of three FRET efficiencies

The FRET efficiency between A_1 and A_2 is calculated as for two-color FRET:

$$E_{A1A2} = \frac{F_{A2}^{A1\,ex}}{F_{A1}^{A1\,ex} + F_{A2}^{A1\,ex}} \tag{4.17}$$

where $F_Y^{X\,ex}$ is the background corrected fluorescence signal after X excitation and detection in Y channel.

Additionally, FRET efficiencies for DA_1 and DA_2 are also obtained from the given three-color scheme. While deriving the FRET efficiency DA_1, care has been taken to consider the loss of A_1 signal after D excitation due to energy transfer from A_1 to A_2 known with FRET efficiency E_{A1A2}. For calculating the FRET efficiency DA_2, the increase in the A_2 signal due to energy transfer from D via A_1 is accounted for as:

$$E_{DA1} = \frac{F_{A1}^{D\,ex}}{F_D^{D\,ex}(1 - E_{A1A2}) + F_{A1}^{D\,ex}} \tag{4.18}$$

$$E_{DA2} = \frac{F_{A2}^{D\,ex} - E_{A1A2}(F_{A1}^{D\,ex} + F_{A2}^{D\,ex})}{F_D^{D\,ex} + F_{A2}^{D\,ex} - E_{A1A2}(F_D^{D\,ex} + F_{A1}^{D\,ex} + F_{A2}^{D\,ex})} \tag{4.19}$$

Three-color FRET measurements were employed to study the domain-wise folding in DM-MBP. For this purpose, specifically triply-labeled DM-MBP was measured under unfolding and refolding conditions. The MFD-PIE analysis was extended to three-color FRET as previously explained (Barth et al., 2019a). An APBS algorithm was used to select single-molecule events including three-color events with some modifications to the two-color FRET selection criterion. At least 30 photons per sliding window of 500 µs and a total of 100 photons per burst threshold was applied in this case.

4.5.4.3 The Three-color FRET set-up

For measurements, a three-color confocal single-molecule set-up equipped with PIE and MFD was developed as described previously (Figure 4.6) (Barth, Voith Von Voithenberg, & Lamb, 2019). Three pulsed lasers with a pulse frequency of 16.7 MHz and a ~20 ns delay between each pulse were used (PicoQuant, Germany; LDH-D-C-485, LDH-D-TA-560, LDH-D-C-640). Lasers were synchronized with a laser driver (PicoQuant, Germany; Sepia II). The microscope was equipped with a 60x water immersion objective with 1.27 N.A. (Nikon, Germany; Plan Apo IR 60x 1.27 WI). Laser powers of ~120 µW for blue, ~75 µW for green, and ~35 µW for red, measured before the objective were typically applied. A polychroic mirror (AHF Analysentechnik; zt405/488/561/633, Germany) separated the laser light from the emitted fluorescence which was passed through a 50 µm pinhole. After the pinhole, the emitted light is separated first for its parallel and perpendicular polarizations with a polarizing beam splitter

(Thorlabs, Germany) and then spectrally for blue, green, and red detection by two dichroic mirrors for green and red color (AHF Analysentechnik; BS560, 640DCXR). The blue, green, and red light is further cleaned by emission filters (AHF Analysentechnik; ET525/50, ET607/36, ET670/30). The blue light is detected using count-blue detectors (LaserComponents, 2x COUNT-100B), whereas green and red are detected on APD's (Perkin Elmer, 4x SPCM-AQR14). A HydraHarp TCSPC module is used to achieve single-photon detection (PicoQuant; HydraHarp400).

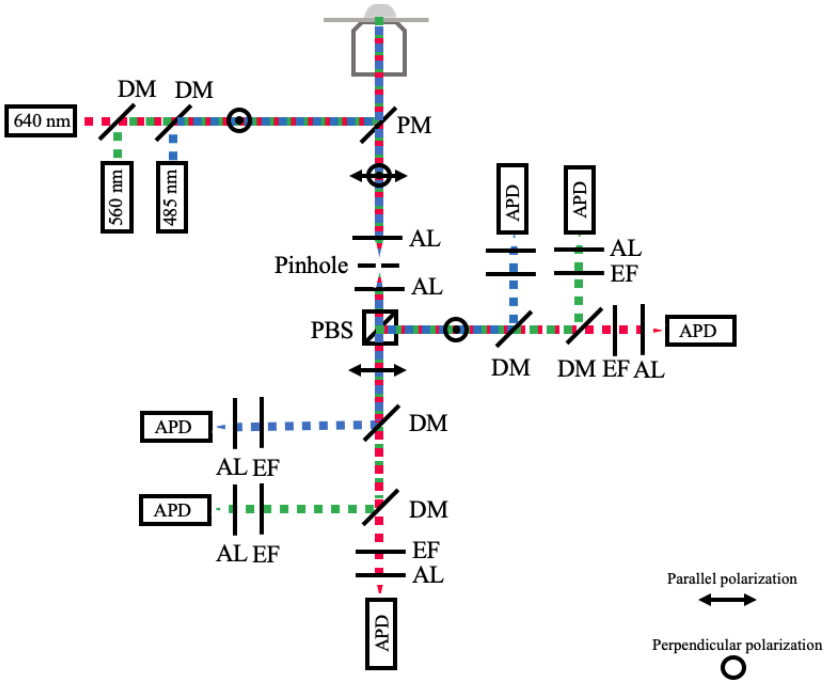

Figure 4.6: Schematics of a three-color PIE-MFD set-up. 485 nm: nanosecond pulsed blue laser, 560 nm: nanosecond pulsed green laser, 640 nm: nanosecond pulsed red laser, DM: dichroic mirror, PM: polychroic mirror, AL: achromatic lens, PBS: polarizing beam splitter, EF: emission filter, APD: avalanche photodiode.

4.5.5 Fluorescence correlation and cross-correlation spectroscopy

4.5.5.1 Fluorescence correlation spectroscopy

When fluorescent molecule diffuses through a confocal spot , the resultant intensity fluctuations can be correlated on time scales ranging from picoseconds to milliseconds, and each timescale reports on a specific process. The lower range is typically limited by the laser pulse frequency.

Of the different time scales, the correlation time in the picosecond range provides the information about the number of emitters in the observation volume. According to the antibunching phenomenon, as the molecule diffuses the photons emitted from a single emitter has separate arrival times, this requires a single fluorescent particle to be measured at a time. The processes analyzed in the nanosecond regime are polymer chain dynamics, and rotational dynamics. In the microsecond regime, generally photophysical properties for example the triplet states, of the fluorophores show up. Photo-induced electron transfer (PET) uses the

fluorescence fluctuations in this time regime to probe site-specific contact formations in a molecule, where another amino acid was replaced by tryptophan to induce the artificial electron acceptor from an excited fluorophore having no visible triplet state (for instance Atto655).

Diffusion term

The most widely used correlation time scale for biomolecular interactions at nanomolar regime is on the order of the diffusion time. This is typically 100s of microseconds on a normal FCS setup, we have increased the volume on the MFD setups where diffusion times are on the order of few milliseconds. In this case, diffusion properties and hence binding, unbinding, or even co-diffusion can be investigated by labeling one of the interaction partners, preferably low molecular weight species, or with second dye in the case of a co-diffusive study. Large complexes or aggregated species having a slower diffusion can also be studied using FCS but on the upper time scale of a few tens of milliseconds limited by the confocal volume dimensions.

Conceptually, fluctuations in fluorescence intensity, $\delta F = F(t) - \langle F \rangle$ at time *t* is correlated with the intensity fluctuations after lag times $t+\tau$ as follows:

$$G(\tau) = \frac{\langle \delta F(t).\, \delta F(t+\tau) \rangle}{\langle F(t) \rangle^2} \tag{4.20}$$

Th analytical form of the autocorrelation function, $\mathrm{G}(\tau)$ can be obtained by approximating the confocal volume by a 3D Gaussian function:

$$G(\tau) = \frac{\gamma}{N} \frac{1}{\left(1 + \frac{\tau}{\tau_D}\right)} \frac{1}{\sqrt{\left(1 + \frac{\tau}{\tau_D} \frac{1}{\rho^2}\right)}} + y_0 \tag{4.21}$$

where *N* is the average number of particles in the observation volume, defined by the structure parameter $\rho = w_z/w_r$, where w_z and w_r are the axial and the radial dimensions of confocal volume. τ_D is the diffusion time, the time particle takes to transverse the volume. γ is a geometric factor that corrects for the spread of the volume. y_0 is a baseline offset.

τ_D is related to the diffusion coefficient *D* as:

$$\tau_D = {w_r}^2/(4D) \tag{4.22}$$

At time lag 0, the above equation is used to calculate the average number of labeled species, *N* which is given as:

$$G(0) = \frac{\gamma}{N} \tag{4.23}$$

Triplet term

Often, fluorophores undergo a blinking event due to the conversion into the triplet state. This process is analyzed separately in addition to the diffusion term assuming the unaltered diffusion properties of a molecule:

$$G_{T_{triplet}}(\tau) = 1 + \frac{T_{triplet}}{1 - T_{triplet}} e^{-\frac{t}{\tau_{T_{triplet}}}} \tag{4.24}$$

where $T_{triplet}$ is the fraction of molecules in the triplet state and $\tau_{T_{triplet}}$ is the triplet time.

4.5.5.2 Fluorescence cross-correlation spectroscopy

The above mentioned correlation is done with a single color and single detection channel. For a co-diffusive study of two particles labeled with different fluorophores, A and B when they are assessed for the interaction between them, the correlation analysis can be extended to two channels, using two excitations and two detection channels. The cross-correlation function is given by (Schwille, Meyer-Almes, & Rigler, 1997):

$$G_{CC}(\tau) = \frac{\langle \delta F_A(t) . \delta F_B(t+\tau) \rangle}{\langle F_{AT}(t) \rangle \langle F_{BT}(t) \rangle} \tag{4.25}$$

The resulting cross-correlation function G_{CC} at time lag 0 provides information about the fraction of interacting molecules with both labels, N_{AB} as:

$$G_{cc}(0) = \frac{\gamma \times N_{AB}}{N_{AT} \times N_{BT}} \tag{4.26}$$

$N_{AT} = N_A + N_{AB}$ and $N_{BT} = N_B + N_{AB}$ are the total number of molecular species with each label.

One of the advantages of a method used in this work, PIE, is that the alternating pulsed lasers on nanosecond regime can be used to eliminate spectral cross-talk. Hence, by employing PIE, the cross-correlation signal becomes very sensitive to the tiny fraction of interactions between differently labeled species.

5 Summary of the papers

Two of the papers resulting from this thesis work are summarized in this chapter. Both papers dealt with protein folding studies employing single-molecule FRET as the primary tool. The first paper presents the dissection of the domain-wise folding pathway of the maltose-binding protein, a two-domain protein with a discontinuous domain topology. The study utilizes both two- and three-color single-molecule FRET to investigate the dynamics of folding as well as the folding order of the two domains. The second paper highlights the opposite roles of the Hsp70 and Hsp90 chaperone systems on the folding dynamics of the p53-DNA binding domain. p53 is transcription factor which functions as a tumor suppressor protein, and a guardian of the genome in eukaryotic cells. It was known to interact with both the Hsp70 and Hsp90 chaperone systems previously, here, we studied the effect of individual and both the chaperone systems on p53 conformation.

5.1 Paper 1: A dynamic intermediate state limits the folding rate of a discontinuous two-domain protein

5.1.1 Aim

Protein folding upon ribosomal synthesis is a critical process for physiological functions of most of the proteins. Since the inception of protein folding studies, mostly small proteins possessing a single domain were investigated for studying the folding principles (Sela et al., 1957) (Batey, Nickson, & Clarke, 2008). However, almost 70% of all proteins contain two or more domains (ref). To extend our understanding of the protein folding mechanisms that evolution has exploited for multiple domain proteins, we chose to study a two-domain model protein, the maltose binding protein (MBP). MBP is of particular interest as it contains a discontinuous chain geometry. For the studies, we employed DM-MBP (double-mutant MBP), a slow folding version of wt-MBP with V8G and Y283D point mutations in N-terminal domain (Tang et al., 2006b). Wt-MBP folds within a minute, whereas DM-MBP folds on the order of half an hour. It has been previously shown that that DM-MBP folding is limited by a long-lived folding intermediate. (Sharma et al., 2008). The intermediate was shown to possess an entropic barrier for folding, that GroEL/ES (a bacterial Hsp60 chaperonin) reduces to accelerate its folding (Chakraborty et al., 2010). Due to its ability to measure distances on the nanometer scale on individual proteins, single-molecule FRET has become one of the tools to study the folding process (Forster, 1946) (Ha et al., 1996) (Schuler, Lipman, & Eaton, 2002). Specifically, the ability to decipher the real time dynamics has made single molecule FRET a preferred method of choice to study the dynamics as the folding process (Kellner et al., 2014) (Dahiya et al., 2019) (Pirchi et al., 2011).

With the possibility to perform single molecule three-color FRET, where three distances are measured simultaneously (Gambin & Deniz, 2010), (Ratzke et al., 2014), (Barth, Voith Von Voithenberg, & Lamb, 2019), we asked what the folding order is of the N- and C-terminal domains (NTD and CTD) during DM-MBP folding. To achieve this, we specifically labeled DM-MBP with three fluorophores such that FRET efficiency between the blue and green fluorophores (FRET efficiency BG) reported on the conformation of the NTD, the green and red fluorophores (FRET efficiency GR) on the conformation of the CTD, and blue and red (FRET efficiency BR) dyes reported on the domain interface (**Figure 5.1A**). We complemented the study with two-color single molecule FRET experiments on the individual domains and investigated the underlying conformational states during domain folding using the donor fluorescence lifetime and the photon distribution analysis (PDA) (Kalinin, Valeri, et al., 2010). Furthermore, to obtain structural details on the folding events, we performed Molecular dynamics (MD) simulations on the wt-MBP and DM-MBP to simulate the unfolding process.

5.1.2 Results and discussion

Two-color single molecule FRET measurements on the individual domains during the unfolding and refolding of DM-MBP revealed the presence of intermediate state in all the three investigated NTD, CTD and domain interface constructs (**Figure 5.1B**). A dynamic analysis using FRET efficiency versus donor fluorescence lifetime provided evidence for fast conformational sampling on the order of the diffusion timescale (milliseconds) between the near native and unfolded conformations during the refolding in all the co-ordinates

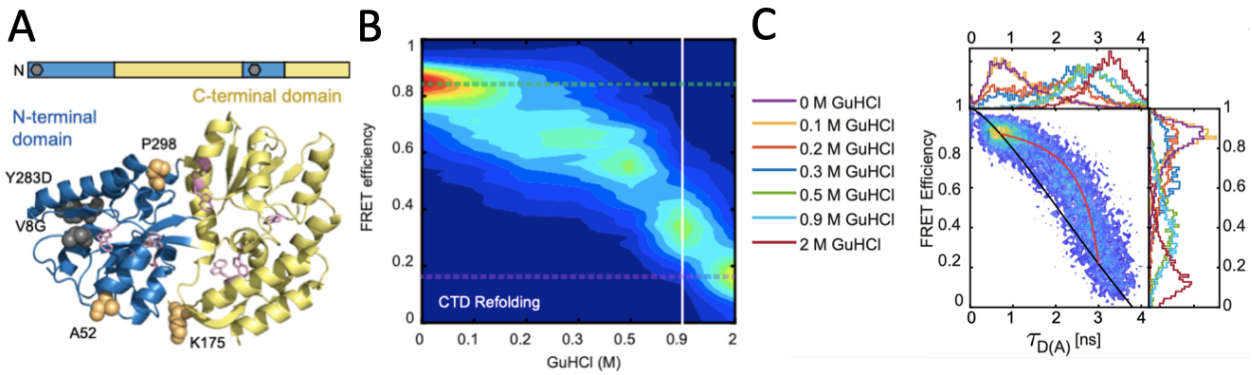

Figure 5.1 DM-MBP refolding shows the presence of an intermediate folding state and sub-millisecond dynamics. **A)** A ribbon schematic of the X-ray crystal structure of MBP (PDB ID: 1OMP). The N- (NTD) and C-terminal domain (CTD) are shown in blue and yellow respectively. The upper graph highlights the discontinuity in the MBP sequence for the NTD and CTD where the positions of the mutations V8G and Y283D are depicted as grey hexagons. The residues involved in the two folding mutations and the three labeling positions A52, K175, P298 for coupling the fluorescent dyes are indicated via a space filling model in dark grey and yellow respectively. Eight tryptophan side-chains are highlighted as a stick model in pink. **B)** 2D-plots of the FRET efficiency versus denaturant concentration for the CTD are presented in a waterfall scheme for equilibrium refolding measurements. For denaturant concentrations between ~ 0.2 and 0.5 GuHCl, an intermediate state with a FRET efficiency of 0.6 is visible. **C)** 2D-plots of FRET efficiency vs donor lifetime in the presence of an acceptor ($\tau_{D(A)}$) (E-τ plot) for NTD equilibrium refolding measurements. For comparison, all measurements are superimposed. The 1-D projections of the different GuHCl measurements, color coded according to the denaturant concentrations, are shown above and to the right of the 2D plots. The ideal relationship between FRET efficiency and donor lifetime is shown in black (static-FRET line) for a molecule that remains in a single FRET state during the observation time. For molecules that dynamically interconvert on the micro- to millisecond time-scales between the folded and unfolded FRET states, a deviation in the E-τ plot is observed, as shown in red (dynamic-FRET line).

(**Figure 5.1C**). Using a dynamic PDA analysis, we could provide estimates for the interconversion rates, which were on the order of 0.12-2.19 ms^{-1}. From dynamic PDA, we observed that the NTD folded first followed by the CTD although NTD folding is slowed down by the folding mutations whereas the CTD possesses fast dynamics during folding (**Figure 5.2A**). We anticipated that the disruption of the hydrophobic interactions in NTD, which create an entropic barrier, to be the cause for the delayed folding. Using three-color single molecule FRET, we could verify the folding order of the domains, where first NTD takes on the native conformation and later followed by the CTD (**Figure 5.2B-C**). MD simulations of MBP unfolding gave a hint on the formation of a folding nucleus in NTD as the first event in the folding of wt-MBP. The same event that takes place in DM-MBP on the slower time scale. Finally, we investigated the influence of GroEL/ES on the previously known accelerated folding on DM-MBP, verifying that the cavity of GroEL/ES restricts the conformational space

and hence entropic search of DM-MBP. The protein still samples unfolded states, but the equilibrium is shifted to more near native state inside the chaperonin cavity.

5.1.3 Conclusions and future perspectives

The current study reiterates the folding complexity present in multi-domain proteins. Although the folding order of DM-MBP domains remains same as in wt-MBP, this study provides a glance into how entropic barriers may create a problem in folding between domains when their boundaries are not distinct (Walters et al., 2013) (Han, Batey, Nickson, Teichmann, & Clarke, 2007) (Inanami, Terada, & Sasai, 2014) (Y. Wang, Chu, Suo, Wang, & Wang, 2012) (Arviv & Levy, 2012). Chaperone substrates often have complex

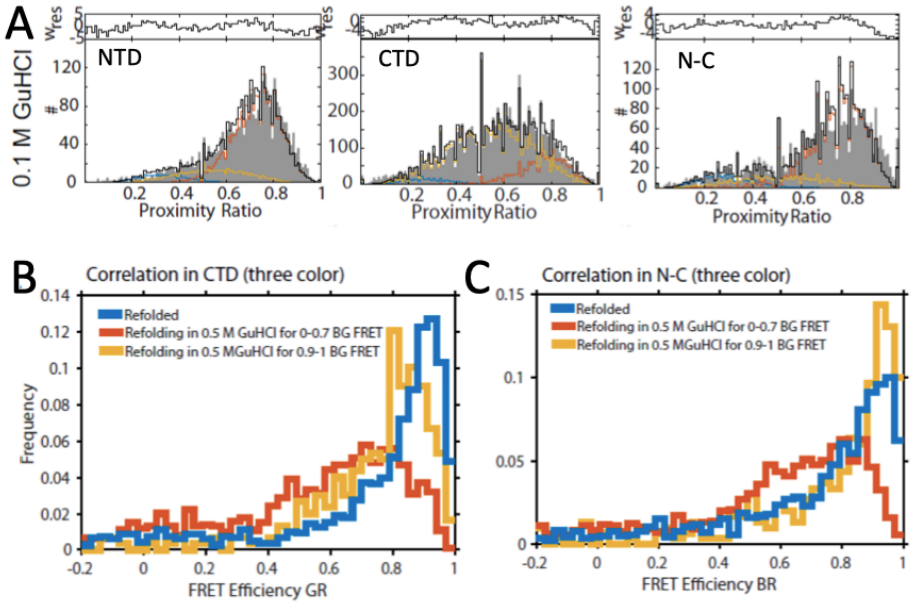

Figure 5.2. The NTD folds first followed by the CTD during DM-MBP refolding.

A) A dynamic PDA analysis of NTD refolding (left panel), CTD refolding (middle panel) and for the N-C interface (right panel) in 0.1 M GuHCl. In grey are the proximity ratio (PR) histograms with 1 ms binning. The fit from dynamic PDA is highlighted as a black outline in the histogram. In the dynamic PDA fit, the folded population is highlighted in red, the unfolded population in blue, and yellow represents the contribution of the interconverting species to the PR. **B-C)** A comparison of the three color smFRET histograms for molecules in the intermediate state ($0< E_{BG} < 0.7$, orange) and in the native-like conformation ($E_{BG} > 0.9$, yellow), as determined from the FRET efficiency in the NTD. The three color-BR smFRET histograms for the CTD (**B**) and the three-color-BR for the N-C interface (**C**) are shown. Measurements of triple-labelled DM-MBP refolding in 0.5 M GuHCl were used for the analysis.

topologies and may lead to aggregation under crowded cellular environment (M. B. Borgia et al., 2011), (Mashaghi et al., 2013) (Cerminara, Schöne, Ritter, Gabba, & Fitter, 2020). The measurements on DM-MBP when encapsulated in the GroEL/ES cavity clearly indicates how chaperonin confinement accelerates folding of multi-domain proteins by limiting the

conformational space available (Vishwas R. Agashe et al., 2004) (Liu, Maciuba, & Kaiser, 2019) (Imamoglu, Balchin, Hayer-Hartl, & Hartl, 2020). Encapsulation also slows down aggregation as encapsulated proteins cannot aggregate. Moreover, multi-color FRET can be employed to investigate complex transitions in different regions of large proteins as well as the effect of the chaperones on folding (Ratzke et al., 2014), (Barth, Voith Von Voithenberg, & Lamb, 2019) (Osváth, Köhler, Závodszky, & Fidy, 2005) (Li, Terakawa, Wang, & Takada, 2012).

5.2 Paper 2: Coordinated Conformational Processing of the Tumor Suppressor Protein p53 by the Hsp70 and Hsp90 Chaperone Machineries

5.2.1 Aim

Cells maintain their genome integrity under various stress conditions with the help of stress sensors that act as transcription factors. p53, a "guardian of the genome", is a key player in deciding cell fate in response to genome damage by activating various cellular pathways to repair the damaged DNA or to undergo apoptosis. Its dysfunction leads to various cancerous conditions due to the accumulation of mutations and unchecked growth. Hence, p53 is also known as a tumor suppressor protein. Most of the cancerous mutations in p53 are found in its DNA-binding domain (DBD), the only structured domain in p53 (Joerger & Fersht, 2007). Previous in vitro as well in vivo studies have shown that, due its intrinsic instability at physiological temperatures, p53 interacts with Hsp70 and Hsp90 systems under normal as well as different stressed conditions (Hagn et al., 2011) (Walerych et al., 2009). The cellular activity of p53 was found to be dependent on the cellular levels of these chaperones (Walerych et al., 2009). However, it was not known how these chaperones modulate the conformation of p53 and hence modulate its function.

In the current study, we aimed to study the individual and combined effects of the Hsp70 and Hsp90 chaperones on the DNA binding function of p53. We employed ensemble measurements of time-resolved anisotropy and analytical ultracentrifugation (AUC) together with single molecule FRET measurements on the conformation of p53-DBD.

5.2.2 Results and discussion

The binding of the p53-DBD to promoter DNA was investigated using ensemble time-resolved anisotropy measurements to monitor the function of p53-DBD. Cy5-labeled p53-DBD can bind to promoter DNA, which gives rise to an increase in anisotropy. Similar measurements with the Hsp40/Hsp70 chaperone system at physiological temperatures of 37°C resulted in a dissociation of DNA bound p53-DBD but was not observed at 25°C. Interestingly, addition of Hsp90, HOP (Hsp70-Hsp90 adapter protein) and a nucleotide exchange factor, BAG1, resulted in restoration of the DNA binding activity of p53-DBD. To support these findings, we devised a p53-DBD FRET sensor to probe the p53 conformations directly using single molecule FRET. Upon probing the Hsp70 bound conformation of p53-DBD at 37°C, we found an increase in the unfolded population in p53-DBD (as opposed to what we observed at 25°C). This supported the finding that Hsp70 may bind and unfold the p53-DBD at physiological temperatures leading to a loss of its DNA binding activity. Further FRET experiments in the presence of Hsp90,

HOP and BAG1 led to the complete conversion of Hsp70 bound p53-DBD to the native conformation of p53-DBD. We also observed that the above interactions were highly dynamic on the diffusion timescale of milliseconds. The overall study highlighted a possible explanation for the interplay of Hsp70 and Hsp90 on the conformation of p53-DBD and on its function under cellular conditions. Similar findings were verified for full-length p53 by using ensemble anisotropy and AUC experiments.

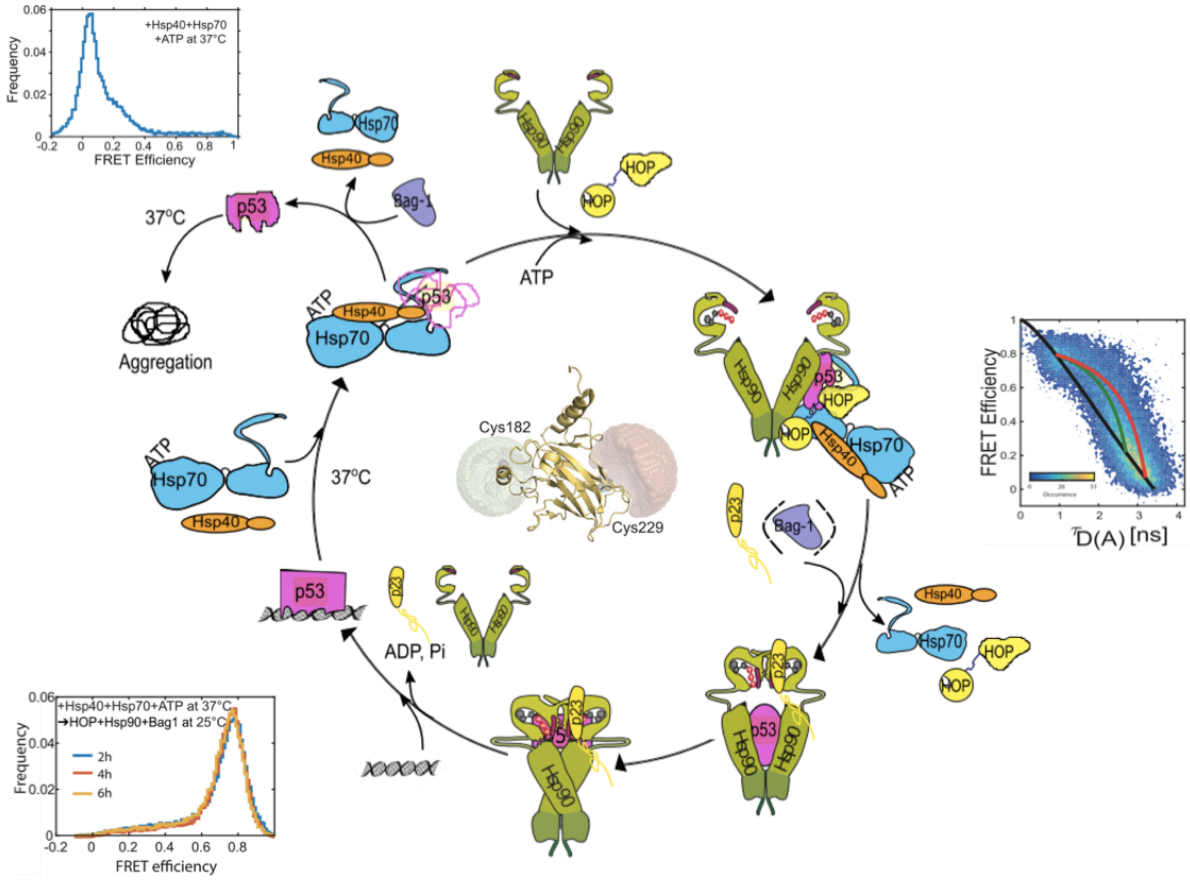

Figure 5.3: Chaperoning of p53 by the Hsp70 and Hsp90 chaperone systems.

Under normal cellular conditions, p53 is unstable and readily bound by the Hsp70 chaperone system keeping p53 in an unfolded conformation. Upon stress, when the function of p53 is required, the subsequent effect of the Hsp90 chaperone system, Hsp90, HOP and BAG1, leads to the release of p53 from Hsp70 allowing it to folding to its native state. The influence of both chaperone systems were simultaneously probed using both ensemble measurements and single molecule FRET studies. Anisotropy measurements verified the ability of p53 to fold and bind a cy5-labeled promoter DNA and AUC experiments confirmed the formation of the p53-DNA complexes. Single molecule FRET experiments were performed on the p53-DBD FRET sensor, shown in the middle of the schematic. The conformation of the p53-DBD bound to Hsp70 was found to be completely unfolded with a low FRET efficiency, whereas the action of Hsp90, HOP and BAG1 allowed p53-DBD to return to the native-like structure with a high FRET efficiency. The simultaneous action of the Hsp70 and Hsp90 chaperone systems revealed

a highly dynamic p53-DBD over the diffusion time of milliseconds as shown in the FRET efficiency versus donor lifetime plots. Figure adapted from Dahiya et al., 2019 (Dahiya et al., 2019).

5.2.3 Conclusions and future perspectives

This study of the interactions between p53 and both the Hsp70 and Hsp90 chaperone systems highlights the complex interplay of these systems, which have opposing effects on the function of p53. We could show that the amount of functional p53 can be controlled, according to the requirements of the cell, by varying the chaperone concentrations. Effects of other proteins regulating the p53 activity, like Mdm2 which keeps the p53 in an inactive form as well as the CHIP, a chaperone associated ubiquitin ligase that directs p53 to 26 S proteasomal degradation, can be probed with the approaches used in this study.

5.3 Additional published papers

Determining the Stoichiometry of Small Protein Oligomers using Steady State Fluorescence Anisotropy.
Heckmeier PJ, **Agam G**, Teese MG, Hoyer M, Stehle R, Lamb DC, Langosch D
Biophys. J. 2020, 119:99-114

Bap (Sil1) regulates the molecular chaperone BiP by coupling release of nucleotide and substrate.
Rosam M, Krader D, Nickels C, Hochmair J, Back KC, **Agam G**, Barth A, Zeymer C, Hendrix J, Schneider M, Antes I, Reinstein J, Lamb DC, Buchner J.
Nat. Struct. Mol. Biol. 2018, 25:90-100

6 Accuracy and precision of single-molecule FRET measurements on proteins

Single molecule FRET is a one of the methods of choice for detecting various structural states in the biomolecules beyond standard structural biology tools like X-ray crystallography and Cryo-Electron microscopy, which often reports on the stable and most prominent states. In addition to identifying the transiently populated states, smFRET provides information regarding the underlying time scales of these fluctuations in the structures spanning nanoseconds to seconds. These states and their dynamics may impart the unique functional consequences under physiological concentrations in cells, may remain mostly undetected by other methods. With its important contribution to the dynamic structural field, smFRET reports are increasing in number since it was first introduced. In spite of the wealth of the information available from these reports, a comparison of reported FRET efficiencies is challenging given the variety of set-ups and analysis methods used. To check the reproducibility of the smFRET experiments across labs, a worldwide study on dsDNA measurements was carried out (Hellenkamp et al., 2018) . To fill the gap between the dsDNA and proteins as real systems investigated with smFRET, a similar study to assess the accuracy and precision of FRET measurements on proteins was initiated by Christian Gebhardt and Prof. Thorben Cordes, Faculty of Biology, Ludwig Maximilian's University, Munich, Germany. They initiated the study with the static protein system of MalE, a Maltose-binding protein. We joined the effort with the introduction of second protein system of U2AF65, a U2 auxiliary factor of 65 kDa from a pre-mRNA splicing machinery that was shown previously to be highly dynamic (Voith von Voithenberg et al., 2016). Both the protein systems were chosen such that their confirmation is tunable after binding to its ligand. Dr. Anders Barth, Milana Popara, Julian Folz and Prof. Claus A.M. Seidel from Heinrich Heine University, Düsseldorf, Germany joined the efforts to analyze the limits on detecting the dynamics across the labs. The U2AF65 protein was provided by Dr. Hyun-Seo Kang, Prof. Michael Sattler, Technical University Munich, Garching, Germany.

Following 19 labs participated in the current protein study (order in alphabetically):

- Thorben Cordes, Christian Gebhardt, Rebecca Mächtel, Niels Zijstra, Jonathan Schneider, Ludwig-Maximilians-Universität, Germany
- Thorben Cordes , Marijn de Boer, University of Groningen, Groningen, The Netherlands
- Timothy Craggs, Ben Ambrose, University of Sheffield, UK
- Dina Grohmann, Kevin Kramm, University of Regensburg, Germany
- Taekjip Ha, Jaba Mitra, Johns Hopkins University, USA
- Jelle Hendrix, Carine Jackers, Hasselt University, KU Leuven, Belgium
- Christian Hübner, Verena Hirschfeld, University of Lübeck, Germany
- Thorsten Hugel, Johann Thurn, University of Freiburg, Germany
- Achillefs Kapanidis, Nicole Robb, Dominik Kammerer, University of Oxford, UK
- Jens Michaelis, Anna Sefer, Ulm University, Germany

- Don C. Lamb, Ganesh Agam, Ludwig-Maximilians-Universität, Germany
- Edward Lemke, Piau Siong Tan, Johannes Gutenberg University, Mainz, Germany.
- Eitan Lerner, Hebrew University of Jerusalem, Israel
- Emmanuel Margeat, Robert Quast, Université de Montpellier, France
- John van der Noort, Kristen Martens, Leiden University, Leiden, The Netherlands
- Claus A.M. Seidel, Anders Barth, Milana Popara, Julian Folz, Heinrich-Heine-Universität Düsseldorf, Germany
- Michael Schlierf, Andreas Hartmann, Neharika Chamachi, Georg Krainer, TU Dresden, Germany
- Philip Tinnefeld, Tim Schröder, Ludwig-Maximilians-Universität, Germany
- Shimon Weiss, Gabriel Moya, University of California, Los Angeles, USA

6.1 Introduction

SmFRET is a preferred approach to unravel structural diversity underlying the biomolecular behavior in single proteins complexes. However, a comparison and validation of results on reported FRET efficiencies and distance changes by various laboratories can be challenging because of vast variety in experimental set-ups and implementations in analysis software. Even the FRET efficiency and distances reported in this thesis on various chaperone and substrates systems investigated depend on the accuracy and the precision of a smFRET measurements. In general, since its inception, an increasing number of reports mention the smFRET technique as just one of the methods used in their studies. Thus, it is important to know how accurate and precise smFRET measurements can be across the labs.

6.1.1 Previous dsDNA study

The mentioned challenges were partly resolved in the first blind study of FRET measurements on DNA standards across 20 labs worldwide (Hellenkamp et al., 2018). The study involved the measurement of two dsDNA standards where donor and acceptor fluorophores, were separated by 15 and 23 bp (**Figure 6.1 A**). Reported FRET efficiencies were compared across the labs. The comparison yielded a standard deviation of less than ±0.05 in FRET efficiency, which is remarkable and attributed to a common agreed on detailed correction procedure provided to the labs when reporting on the FRET efficiency (**Figure 6.1 B**). When translating this accuracy in FRET efficiency into distances, distances were accurate to within ±6 Å (**Figure 6.1 C**).

6.1.2 The need to assess the consistency for proteins

The results were encouraging and transcended the power of standardized measurement and correction procedures in the field. DNA being uniform in structure and chemically less complex than proteins provides the easiest case for the such a comparison, when same DNA sequence was kept near the same dye for previous benchmark study on DNA. Easy synthesis of DNA makes possible to test the deviations in the measurements, for example, from the dye photo-physics reasonably quickly. On the other hand, in addition to the proof of principles in the method development and optimization made on DNA strands, measurements on more complex DNA structures like DNA origami, DNA- and RNA-protein complexes can provide great insights into the molecular conformations as well. To make a case different than DNA or DNA interactions, we investigated the FRET efficiencies exclusively on protein systems.

Proteins are the work horses carrying out most of all the basic physiological processes in cells. Biomolecular studies often deal with investigating their functional states. Although it comes with its own set of challenges, smFRET is one of the methods of choices for these types of investigations on proteins. Their purification, their stochastic labeling, diversity in chemical composition and stability makes the measurements of FRET efficiencies more difficult.

Fluorophore attachments may have an adverse effect on protein function or vice-e-versa the protein may have adverse effects on fluorophore properties. These effects may arise due to transient or permanent sticking between the fluorophore and certain side chains of amino acids

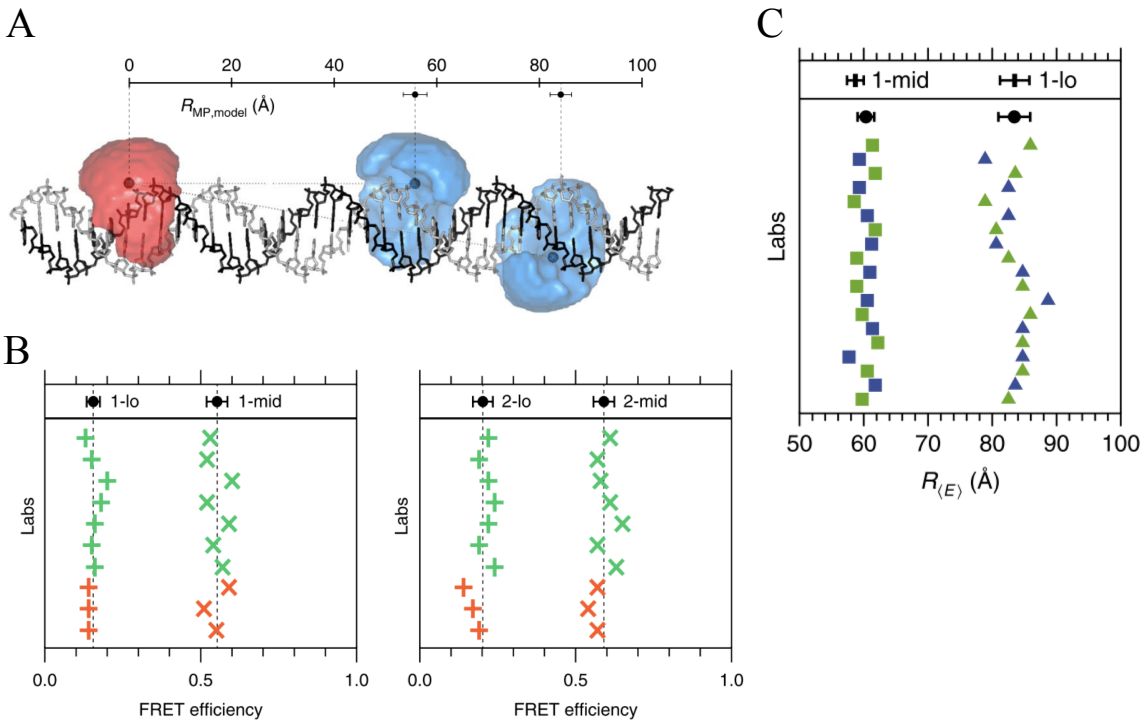

Figure 6.1: Accuracy and precision from the previous world-wide study on DNA. **(A)** A schematic of the dsDNA sample depicting the attachment points for donor (blue) and acceptor (red) for two measured distances of 15- and 23-bp. Accessible volume surfaces derived by the FRET positioning and screening software (Kalinin et al., 2012) for the respective dyes are shown and their mean positions are highlighted by darker spheres. **(B)** A comparison of the reported FRET efficiencies by the labs from diffusion measurements. lo- refers to the 23 bp and mid- refers to 15 bp distance. 1- refers to the dye pair Atto550-Atto647N and 2- refers to the Atto550-Alexa647. The green and red colors indicate different approaches for correcting γ: green being a global and red being the species-based approach. Note, the accuracy in FRET efficiency of less than ±0.05 as shown by the standard deviation in upper portion of the plot with its mean as a dash line. **(C)** A comparison of derived distances from the values in **(B).** Green points indicate the confocal and blue indicate TIRF-based measurements. The upper mean ± s.d. is from the static model and the lower mean ± s.d. is from the experiments. A s.d. of ±6 Å was found from the experiments on upper limit for larger distance. The figure panels are taken from Hellenkamp et al., 2018.

in the protein. Sticking of the dye changes its rotational behavior as well as its propensity for quenching or blinking propensity in dyes. Dyes may induce mal-function in a protein to which it is attached either by directly interacting with interaction partner or ligand of an its attached protein or indirectly by inducing changes in conformational space of a protein to which it is attached. Sample preparation, measurements and analysis gets more complicated when the protein is dynamic. Dynamic systems possess additional challenges, such as keeping them stable during storage due to the inherent structural flexibility and unpredicted behavior than the fluorophore may have with the different conformations. The dye itself may also change the timescale of the fluctuations in proteins. One of the advantages of smFRET is that it can extract kinetics information on the fluctuations in protein conformation which can span several orders of magnitude in times. Thus, assessment and comparison of the FRET measurements on proteins having different levels of dynamic complexities (static and fast dynamic) is worth investigating to validate the versatility of the method and to keep up with the pace of the challenges when its use is increasing in translational and biomedical applications.

6.2 Introduction of the investigated systems

6.2.1 Static system of MalE

To begin the study on proteins, a stable and robust protein system, MalE, was chosen. MalE is a maltose-binding protein, which forms a periplasmic part of the *E. coli* ATP binding cassette transporter operon $MalFGK_2$. It is involved in the uptake of maltose from the medium. MalE has two structured domains connected via beta sheet linkers (Sharff, Rodseth, Spurlino, & Quiocho, 1992) (**Figure 6.2 A)**. Maltose is a small, specific ligand for which different substrate-binding proteins have evolved that do not require large fluctuations or conformational changes in the system to bind to ligand (De Boer et al., 2019). Fairly rigid linkers allow the protein to sample conformations between the domains in an apo form, most populated conformation being the open one (**Figure 6.2 B**). One of the beta sheet linker allow the domains to close and, is stabilized by maltose binding where the domains are locked in close proximity on a ~100 ms time scales under high ligand concentrations as previously shown (Spurlino, Lu, & Quiocho, 1991) (E. Kim et al., 2013) (De Boer et al., 2019) (**Figure 6.2 C**). This simple conformational change upon ligand binding was analyzed with the three pairs of distances by coupling the fluorophores. These three pair of the labeling positions were chosen such that the distance increases in one of them, as it is placed on the same interface of the domains where maltose binding takes place (MalE-1: K29C-S352C), a second one decreases, as it lies on the opposite interface (Mal-2: D87C-A186C) and the last one does not change at all, as they are placed on the same domain (Mal-3: A134C-A186C) upon ligand binding (**Figure 6.2 B-C**). These mutants were designed to detect these simple FRET efficiency changes controllable by a small ligand across the participating labs to benchmark the accuracy and precision in FRET efficiencies when measured on proteins.

These three pair of positions were mutated to the cysteines to enable the coupling of the donor-acceptor dye-pair of Alexa546-Alexa647 via cysteine-maleimide chemistry. The FRET efficiency histograms correctly predicted the behavior as expected for all the three mutants, where for MalE-1, holo exhibited higher FRET than its apo form, MalE-2 shows the opposite behavior with a decrease in FRET efficiency in holo compared to apo and no change was

apparent for MalE-3 upon maltose binding. Although, MalE-3 mutant is not very sensitive to conformational change due to its high FRET efficiency. **(Figure 6.2 D-F)**. The stability and functionality of the samples were checked before sending the three mutants along with the buffers with and without maltose to the 19 participating labs. The designs of the MalE mutants and its purifications, labeling, stability and functionality assays were carried out by Christian Gebhardt in the laboratory of Prof. Thorben Cordes. The samples were sent on ice packs with instructions to follow while correcting the data.

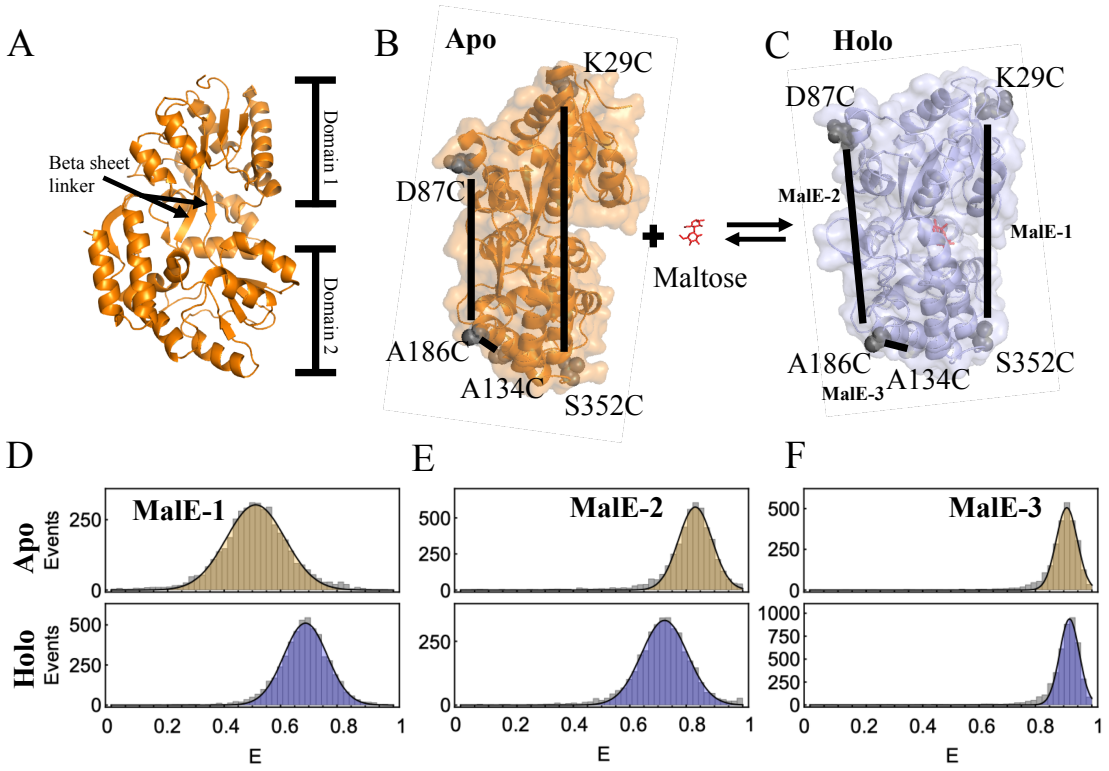

Figure 6.2: The Maltose binding protein (MBP). **(A)** Ribbon schematics of MBP structure in the apo state (PDB:1OMP). (**B)** A second view of the apo structure as is shown with a semi-transparent orange surface presentation in addition to the ribbon schematic. The three distances measured are shown by black lines. The attachment amino acids are shown as black spheres. **(C)** The holo structure is shown in semi-transparent blue surface presentation in addition to the ribbon schematic. (PDB:1ANF). After binding to maltose (red), the relative changes in the distances are compared with apo. Three positions were chosen such that, on maltose binding, the distance will decrease (Mal-1: K29C-S352C), increase (Mal-2: D87C-A186C) or remain unchanged (Mal-3: A134C-A186C). (**D-F)** FRET efficiency histograms for MalE-1 in apo (orange) and holo state (blue) in **(D)**, for MalE-2 in **(E)** and for MalE-3 in **(F).** MalEs were labeled with Alexa546-Alexa647 and 1 mM maltose concertation was used to measure the holo state**.** Color scheme of orange for apo and blue for holo is maintained throughout unless specified, otherwise. Figure schematic adopted from the initial draft of a figure from Christian Gebhardt.

6.2.2 Dynamic protein system of U2AF65

A second protein system that we introduced assesses the capacity of smFRET in detecting ad differentiating fast dynamics, was a U2AF65, a U2 auxiliary factor of 65 kDa. U2AF65 is part of the spliceosome complex in humans and has two structured domains namely the RNA recognition motif 1 (RRM1) and RRM2, that are connected via an unstructured linker (MacKereth et al., 2011) **(Figure 6.3 A)**. Each RRM recognizes a stretch of a few pyrimidines (py). U2AF65 recognizes the 3'-splice sites from mRNAs rich in poly-py and the unstructured linker provides the high flexibility needed to efficiently reorient the domains to accommodate long stretches of py-tracts as long as 13 pyrimidines (Voith von Voithenberg et al., 2016). To probe the free and RNA bound conformations of U2AF65 by smFRET, we used the previously tested fluorophore attachment sites namely, L187C and G326C, each on separate RRM domains **(Figure 6.3 B)**. Thus, a free U2AF65 protein is highly flexible on the sub-millisecond time-scales and the dynamics between the domains is stabilized upon binding to the py-tract **(Figure 6.3 C)**. We introduced the system to have this two-state mechanism of dynamic and stable conformation tunable via RNA ligand binds as a good system to verify the presence of the fast dynamics across the labs.

For the Atto532-Atto643 as a donor-acceptor dye pair used, the FRET efficiencies for apo conformations showed that the free RRMs are, on average fairly compact with a FRET

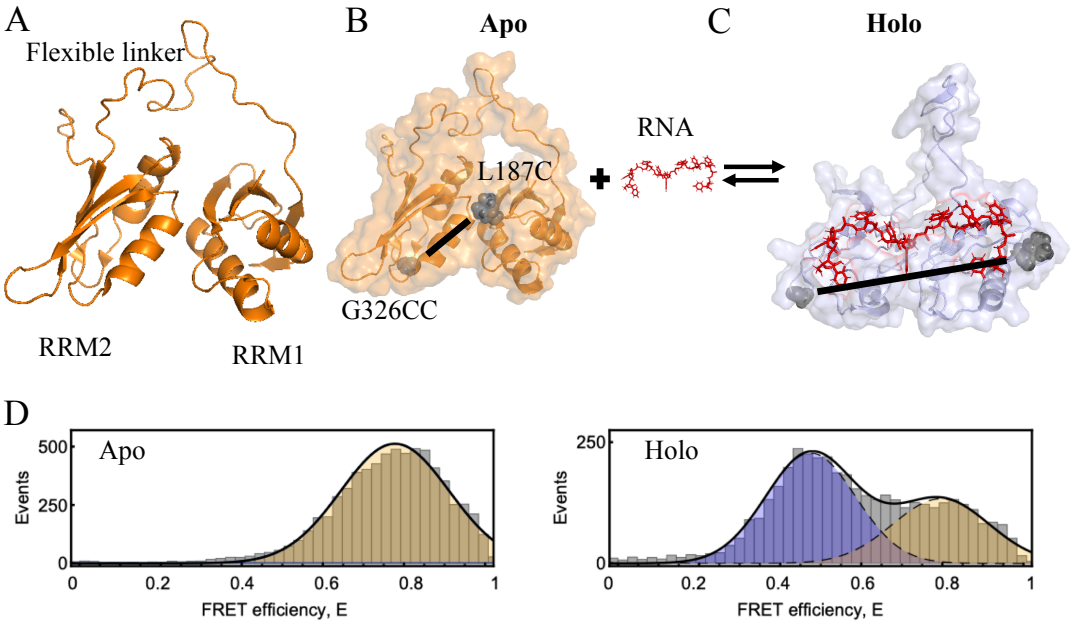

Figure 6.3: U2AF65. **(A)** A schematic NMR structure of U2AF65 showing the two RNA recognition motifs (RRM1,2) connected by a flexible linker in an apo closed state (PDB: 2YH0). **(B)** A second view of the apo structure as is shown with a semi-transparent orange surface presentation in addition to the ribbon schematic. The distance measured is shown by black lines. The attachment amino acids are shown as black spheres **(C)** U9-bound holo state of U2AF65 shows stabilization of open conformation. U9 is shown in stick presentation in red. (PDB:2YH1) **(D)** The FRET efficiency histograms obtained for apo (orange) and holo (blue) states (with 5 µM U9 RNA) after labeling with the Atto532-Atto643 dye pair. A fraction of apo in the holo measurement is shown in orange to be consistent with apo-like conformation.

efficiency of 0.75, compared to the U9-bound conformation (FRET efficiency of 0.45), similar to our earlier study **(Figure 6.3 D)** (Voith von Voithenberg et al., 2016). The apo state has high a FRET efficiency, it has a broad width, which underlies the conformational heterogeneity due to its known detached ensemble of conformations (J. R. Huang et al., 2014). The holo-state with 5 µM U9 RNA, had some residual apo state, rendering it either RNA free or represents a defective small oligomeric protein species. U9, being an RNA is sensitive to the degradation and may lead to some RNA free protein. Nevertheless, we found a similar binding affinity (K_D of ~1.3 µM) for U9 from analyzing the FRET efficiency histograms from RNA titration to what we measured previously confirming the proper functioning of the system **(Figure65.4)** (Voith von Voithenberg et al., 2016). With a K_D of ~1.3 µM, the 5 µM concentration used in the holo measurements would only lead to binding of ~85% RNA to U2AF65 and explains the remaining apo species visible in the FRET efficiency histograms.

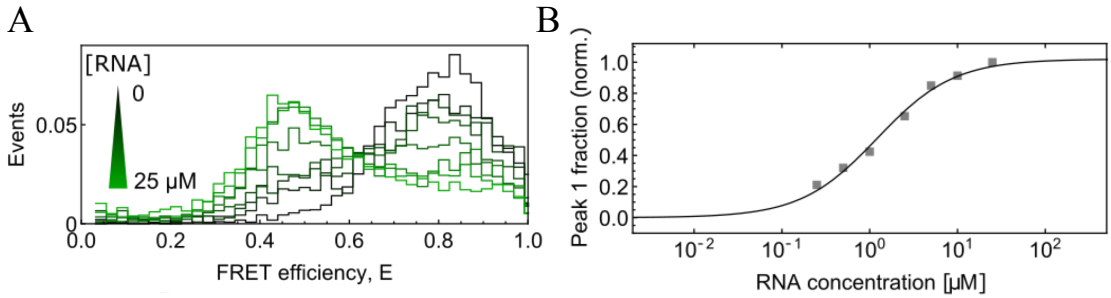

Figure 6.4: Determination of the dissociation constant for labeled U2AF65 and U9. **(A) A** U9 RNA titration measurement for U2AF65. The U9 RNA concentrations used were 0, 0, 0.25, 0.5, 1.0, 2.5, 5, 10 and 25 µM. The FRET efficiency histograms for lower to higher RNA concentrations are shown from dark to light green. **(B)** To determine the dissociation constant (K_D) from the measurements shown in the panel **(A)**, area between FRET efficiencies 0.1-0.6 (peak 1 fraction) were taken as a fraction for apo conformation and were plotted against the increasing U9 RNA concentration. Th resulting curve was fit with a kinetic model $AB = \frac{K_D+A+B-\sqrt{(K_D+A+B)^2-4\,A\,B}}{2}$. A is total, free and a bound fraction of the U2AF65 protein concentration, similarly B is the total U9 RNA concentration and AB is the concentration of a formed complex. The K_D was found to be 1.3 µM.

6.3 Results

6.3.1 Precision and consistency in the reported FRET efficiencies on proteins

For MalE, all the three mutants were sent to 19 labs, out of which, 16 labs could send us the results. The labs were asked to report on the FRET efficiency values after correcting the data according to the step-wise correction procedures recommended in Hellenkamp et. al., 2018 (Hellenkamp et al., 2018). Specifically, groups were asked to perform a global γ correction and to report the FRET efficiency values after Gaussian fitting for the populations in a 2-D E-S histogram. Moreover, we also asked them to report a statement about both the systems, whether its dynamic or not, topic of the discussion in the later part of the results. Some of the

labs performed measurements with µs-ALEX and the rest of them used the ns-ALEX approach of PIE and MFD.

Figure 6.5 summarizes the FRET efficiency results reported by the labs for the MalE constructs. All the 16 labs reported the proper functioning of the MalE system for all the three mutants. Expected trend upon maltose binding for MalE-1 (low to high: from 0.49 to 0.66 FRET efficiency), MalE-2 (from high to low: 0.83 to 0.71 FRET efficiency), and MalE-3 (no change from the FRET efficiency of 0.91) was seen. The reported FRET efficiency values for all the three constructs had a maximum standard deviation of ±0.059 for the MalE-1 apo construct and as low as ±0.019 for MalE-3 apo construct. This could be due to the fact that, MalE-3 construct is not very sensitive to the deviations at very high FRET efficiency. The trend in the standard deviation was apparent, with decrease in standard deviations from low to high FRET efficiencies hinting on the shot-noise as a cause for the distribution widths. FRET efficiency is calculated as number of acceptor photons detected over the total number of photons detected. Uncertainty in photon detection has an inherent shot noise which results into the broadening of FRET efficiency histograms. As the FRET efficiency increases, the detected number of acceptor photons increases. This results into a narrow histogram for high FRET efficiency compared to the low FRET efficiency histogram that are constructed from low number of detected acceptor photons. This trend is visible in **Figure 6.2** as well.

In the case of the U2AF65 measurements, only 8 labs participated because of the system complexity and the dynamic analysis procedures involved as a part of reporting results. The consistency of the FRET efficiencies reported by the labs on the U2AF protein measurements are displayed in **Figure 6.6, upper panels**. As expected, for the apo protein, the FRET efficiency distributions were broad with a standard deviation of ±0.049. A similar deviation of ±0.053 was found for the apo fraction in holo measurement and the RNA-bound fraction had a standard deviation of ±0.041. When comparing the individual histograms **(Figure 6.6, lower panels)**, two clear outliers for FRET efficiency values reported for apo were evident with the reported FRET efficiency values being higher than expected. The same labs reported consistently higher FRET efficiency values for the holo measurements as well (clearly seen in the apo species). Interestingly, one extra outlier was found reporting higher fraction for apo species in the holo measurement. A change in the fraction of holo protein can be explained by a decrease in U9 RNA concentration due to degradation (common to RNA), an increase in non-functional protein fraction or a change in the experimental conditions (e.g. solvent conditions or temperature). **(Figure 6.4)**.

While shipping the U2AF65 protein, we shock froze the protein and shipped it on dry ice. This treatment might have caused an increase in non-functional protein. Additionally, measurements conditions were not specified to incorporate the maximum variability in measuring the dynamic sample. We found that temperature had a significant effect on the U2AF65-U9 complex stability **(Figure 6.7)**. A slight increase in the temperature performing the measurement can lead to a high RNA free protein fraction.

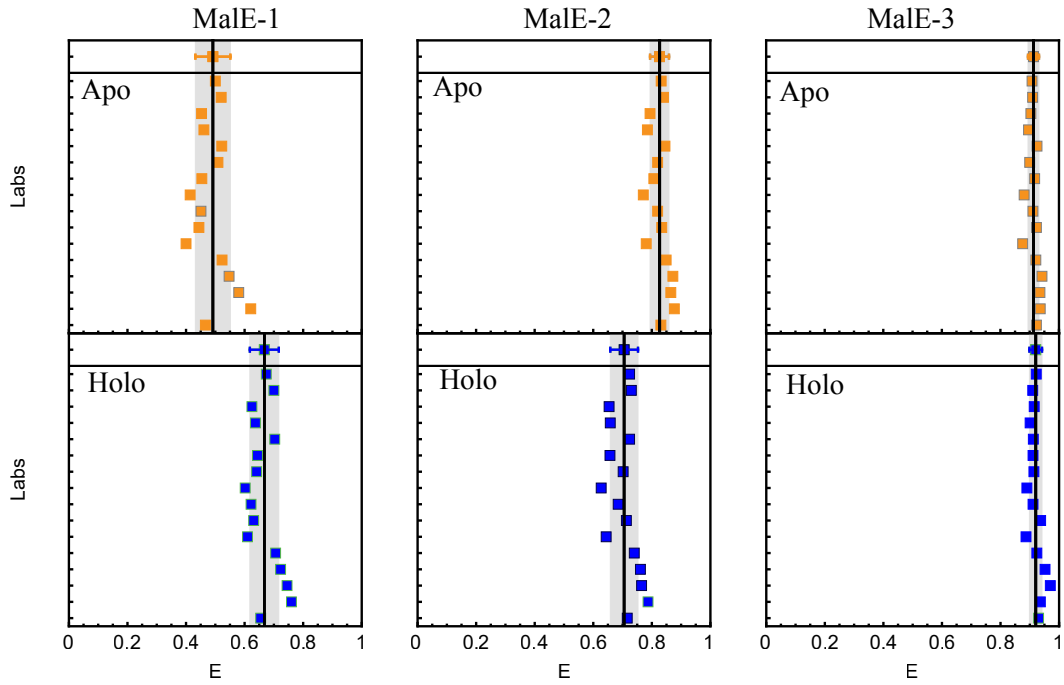

Figure 6.5: Consistencies in the reported FRET efficiency values for the MalE system from 16 labs. The apo results are shown in orange and the reported values for holo measurements are in blue. The mean value is shown on each plot as an upper separate panel. The standard deviations from all the values obtained are shown around the mean values as error bars in the upper panel and as a grey shaded area in the lower panels. The sequence of the labs plotted in the figure is randomized and is not according to alphabetical order as mentioned in the beginning. This random sequence is kept constant throughout the study for consistency. The plotted data was made available by Christian Gebhardt and Prof. Torben Cordes.

6.3.2 Reanalysis of the collected data from 8 labs.

During the analysis, systematic error in determining the various correction factors could lead to the observed inconsistencies. Out of these, detection correction factor, γ is one of the difficult corrections to determine. In case of U2AF, due to presence of dynamics, its sensitivity to the solvent and experimental conditions as well as the presence of apo species (having different stoichiometry) in holo, made the γ-fcotor determination more challenging. To find the reasons for higher inconsistencies in the U2AF system than for MalE, we asked 8 groups to send us their raw data and reperform the analysis by a single person. On a closer inspection of the raw data from the groups, specially the two outliers, we found that one of the groups has measured more of the apo species in the holo measurement. Before reperforming the analysis on the collected data, we first investigated the apo species we measured in the current sample and compared it to a previous study. We found the apo species to have either quenched or have a different acceptor: donor stoichiometry of N_A:N_A+1 (N_A is the number of acceptors) leading to the higher stoichiometry of ~0.6 (**Figure 6.8 A)** (Voith von Voithenberg et al., 2016). Another explanation could be fluorophore aggregation under the protein labeling conditions.

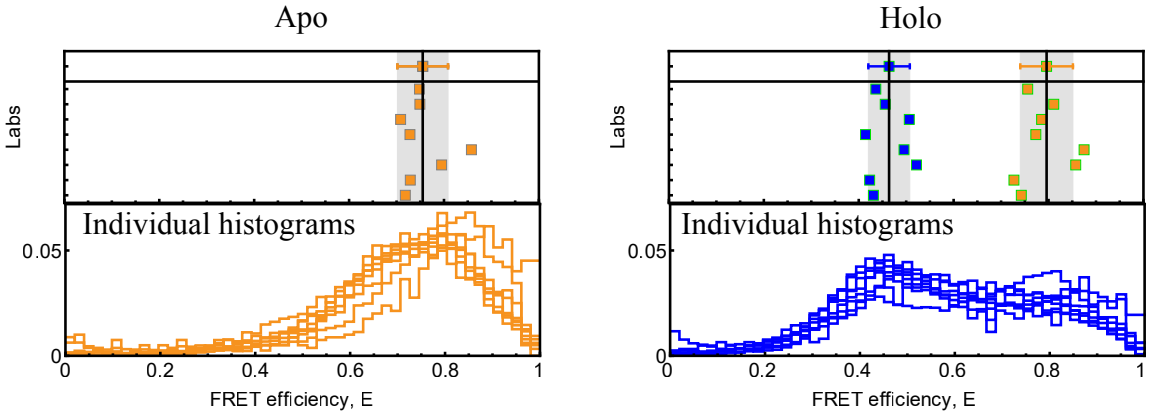

Figure 6.6: Consistency in the reported FRET efficiency values for the U2AF65 system from 8 labs, the apo results are on the left panels (in orange) and reported values for the holo measurements are shown in the right panels (in blue). In the upper panels, the mean and standard deviations are displayed. The mean value is shown on each plot as an upper separate panel. The standard deviation from all values obtained are shown as error bars around the mean values and as a grey shaded area in the lower panels. The lower panels show the histograms from the individual labs.

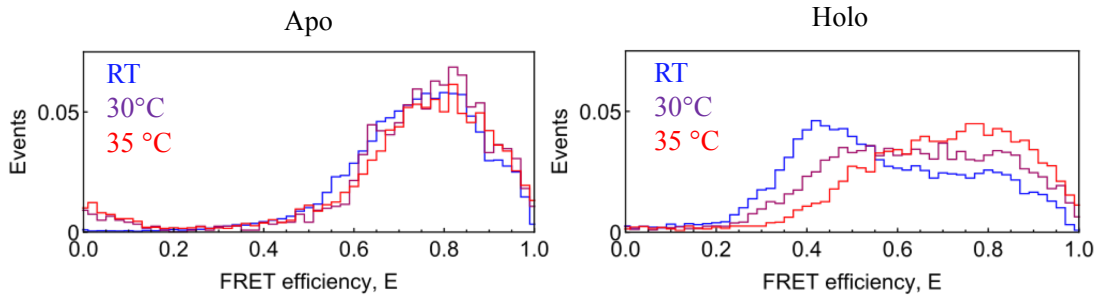

Figure 6.7: Temperature dependent variations in apo and holo measurements of U2AF65. The apo measurement does not show much variation except increase in the low FRET efficiency fraction partly due to an increase in multi-molecules because of temperature induced denatured or degradation of protein. In contrast, in holo measurement, as temperature increase, U9-bound fraction decrease significantly as well as its FRET efficiency peak shifts from 0.44 to 0.5.

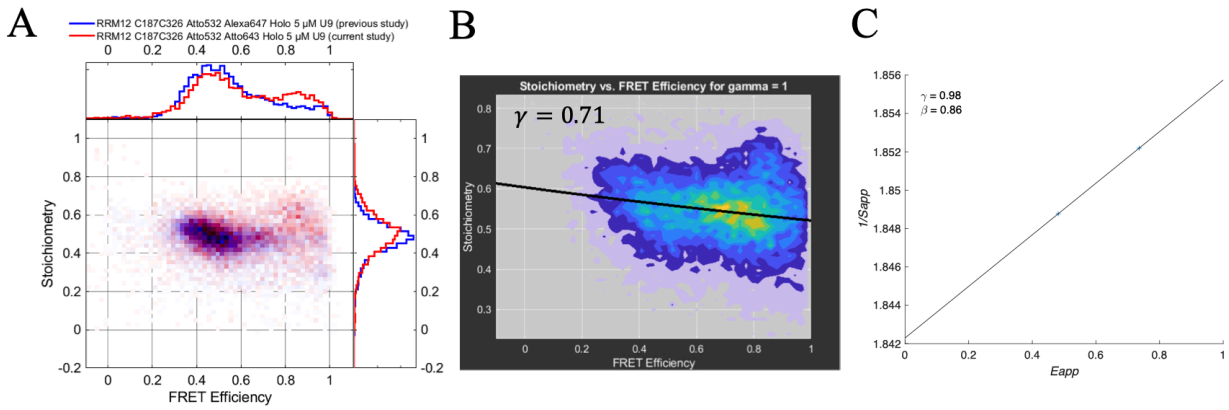

Figure 6.8: A comparison of a previous and current sample preparation and its effect on the γ determination. **(A)** Comparison of the holo measurement of U2AF65 from the current (red) and previous study (blue). In addition to the known holo population at 0.45 FRET efficiency, an extra population was observed with a FRET efficiency of ~0.8 and a stoichiometry of ~0.6. **(B)** γ-factor determined from all the burst from apo and holo measurements including the ~0.6 stoichiometry population by fitting a line to $1/S_{PR}$ vs E_{PR}. A γ of 0.71 was found in this case. **(C)** The γ-factor determined for the population means of main apo and holo populations ignoring the ~0.6 stoichiometry population. The fit is a line to the $1/S_{PR}$ vs E_{PR} for population means.

During reanalysis all the analysis steps were performed according to the agreed upon procedures same as groups were asked (Hellenkamp et al., 2018). To test the effect of this population with 0.6 stoichiometry on the analysis, we followed different approaches while calculating the γ factor in reanalyzing the data. We rationalized, when taking all the bursts to calculate γ factor by fitting a line to $1/S_{PR}$ vs E_{PR} (Lee et al, 2005), depending on the abundance of the 0.6 stoichiometry population across the labs, γ-factor might be estimated differently. As the γ factor calculation assumes equal brightness for the all bursts which is not the case for the 0.6 stoichiometry population, the γ estimation might be underestimated with the increase in the fraction of this population (**Figure 6.8 B**). To account for this fact, we approached γ determination by fitting a line to $1/S_{PR}$ vs E_{PR} only for the population means. For apo, we had only one population but for holo, we ignored the 0.6 stoichiometry population and took population means for holo and apo species **(Figure 6.8 C)**. This is line with a fact that different stoichiometries cannot have the same γ. Indeed, when we estimated the γ factors for all the 8 collected data sets, we consistently found lower γ values as compared to the reported values (Table 5.1 and 5.2). For one of the outliers, the difference between the reported and the reanalyzed γ was the maximum, indicating the highest effect of correct determination of γ in this case after reanalysis. The same group has consistently reported FRET values higher than the expected.

Table 6.1: Correction factors reported by 8 labs on U2AF65 measurements.

Lab	α	β	δ	γ
1	0.02	0.78	0.06	0.59
2	0.06	-	0.23	1.1
3	0.04	0.64	0.05	0.80
4	0.03	1.05	0.09	0.83
7	0.04	0.62	0.06	1.24
8	0.03	0.91	0.05	0.73
11	0.05	-	0.02	0.64
14	0.05	-	0.09	0.64

α is the correction for cross-talk of donor photons in the acceptor channel, β is the ratio of acceptor laser intensity to the donor laser intensity, δ corrects for the direct excitation of an acceptor and γ is the detection correction factor. Numbering of the labs is maintained as given for the MalE study.

Table 5.2: Correction factors obtained after reanalysis of collected U2AF65 datasets from 8 different labs.

Lab	α	β	δ	γ
1	0.02	0.72	0.06	0.73
2	0.04	0.22	0.20	1.26
3	0.03	0.77	0.06	0.98
4	0.03	1.15	0.08	0.98
7	0.03	1.22	0.05	1.00
8 *	0.03	0.86	0.05	0.98
11	0.03	1.21	0.07	0.65
14	0.03	1.00	0.05	0.92

* Determined γ is shown in **Figure 5.8 C**

When we plotted the consistencies in the FRET efficiencies with the correction factors obtained after reanalysis, we observed a better consistency and less variation at least in the apo population **(Figure 6.9 A-B)**. Variation in the holo remained similar, most likely due to the complexity of the apo species determining the γ-factor in the holo measurement. We did the same reanalysis on the collected data from the same 8 labs on the MalE system as a control. Not to our surprise, there was no big impact of the reanalysis on this dataset. **Figure 6.9 C-D** shows the results for the apo state of MalE as an example. We attribute these observations to the more challenging system of U2AF65 where populations display fast dynamics, an RNA-based ligand, population impurities and its sensitivity to the measurements conditions like temperature.

6.3.3 Detection of dynamics and its consistency across the labs

One of the main differences between this study and the previous DNA study was the use of proteins that can undergo conformational dynamics on the seconds to micro-second timescales. Our chosen systems were also known to have different dynamic time-scales where MalE was known to undergo slow sub-second fluctuations (~100 ms) and U2AF65 on the sub-millisecond timescales (De Boer et al., 2019) (Voith von Voithenberg et al., 2016). To see how consistently groups could recognize the dynamics on these different timescales, we had asked the groups to report on their analysis about dynamics. Various approaches are available for analyzing the fluctuations of the single-molecules and quantitating the dynamics on different time-scales. We focused on two approaches, namely the Burst-variance analysis (BVA) and donor lifetime-FRET efficiency (E-τ analysis), which are widely used by the participating groups detecting the dynamics on slow and fast time-scales respectively. BVA provides information on

deviations of the PR (proximity ratio or E^*) beyond that of shot-noise within a burst by separating the burst observation time-windows into certain defined number of photons, here 5 photons. If molecular conformation fluctuates within the time-windows of a burst, BVA detects the deviation in the proximity ratio and hence is sensitive to slow dynamics during a burst (Torella et al., 2011). The E-τ approach is sensitive to slow and fast fluctuations as the donor lifetime samples these fluctuations on the nanosecond timescale. The typical inverse relationship of donor lifetime versus FET efficiency is considered for a static system. When a molecule undergoes conformational fluctuations during a burst, the intensity based FRET efficiency calculation is a species weighted averaged while the donor lifetimes are biased towards low FRET states as their calculation is based on a photon weighted average. The dynamics is observed as a deviation in the donor lifetime for low FRET state from the standard relationship (Kalinin, Valeri, et al., 2010).

Before getting into the details of the qualitative and quantitative dynamic analysis reported by the groups, we first analyzed and characterized both systems. MalE was known to undergo slow fluctuations mediated by maltose on the time scales of ~100 ms as previously observed (De Boer et al., 2019). With a BVA analysis, all the MalE mutants both under apo and holo conditions seems to be static over the burst duration of a few milliseconds as there was no deviation in the calculated variance for the bursts beyond what is expected for shot-noise limited bursts **(Figure 6.10 A-C)**. The shot-noise limited standard deviation is shown as a hemi-circle, the highest being at a proximity ratio of 0.5. Interestingly, when we analyzed these mutants with E-τ analysis **(Figure 6.11 A-C)**, a deviation from the static-line was observed for both apo and holo states for MalE-1, while MalE-2 and MalE-3 were found to be static as they fall close to the static-line. Here, a clear distinction between static and dynamic is ambiguous because of the lack of information on the limits of sensitivity detecting the dynamics, which we analyzed later in the study quantitatively. These analyses on MalE mutants indicated that BVA is sensitive to the slow fluctuations showing MalE as a static molecule, whereas the E-τ analysis showed that MalE-1 has fast dynamic fluctuations and MalE-2, and MalE-3 are static systems. As a part of later investigations on dynamic nature of MalE, it was found that, in case of MalE-1 the dye sticking was detected as a deviation in E-τ analysis.

The U2AF65 dynamic analysis was challenging. We had expected a clear two-state mechanistic tunability of its dynamics by RNA-ligand binding (MacKereth et al., 2011), but we observed a slightly different behavior where ligand tuned the system less efficiently giving rise to intermixing of the states which made the overall analysis and interpretations difficult. Nevertheless, when we analyzed U2AF65 with BVA, the burst variance of the apo population indicated the presence of dynamics with an observed increase in the standard deviation in the bursts (**Figure 6.12 A)**. Although, the holo conformation of U2AF65 upon binding to RNA had less detectable deviation in BVA than apo, we had the apo species in the holo measurement, which was found to be dynamic consistent with the apo conformation (**Figure 6.12 B)**. As RNA-binding with a K_D of 1.3 μM, k_{off} of 0.76 s^{-1} μM^{-1} and k_{on} of 1 s^{-1}, takes place slower than the time-scale of diffusion, the high variance in the bursts bridging the apo and holo conformations is most likely due to apo species in the holo measurement. These observations indicated that the U2AF apo has sub-millisecond dynamic component but the holo conformation becomes stable after binding to RNA (MacKereth et al., 2011).

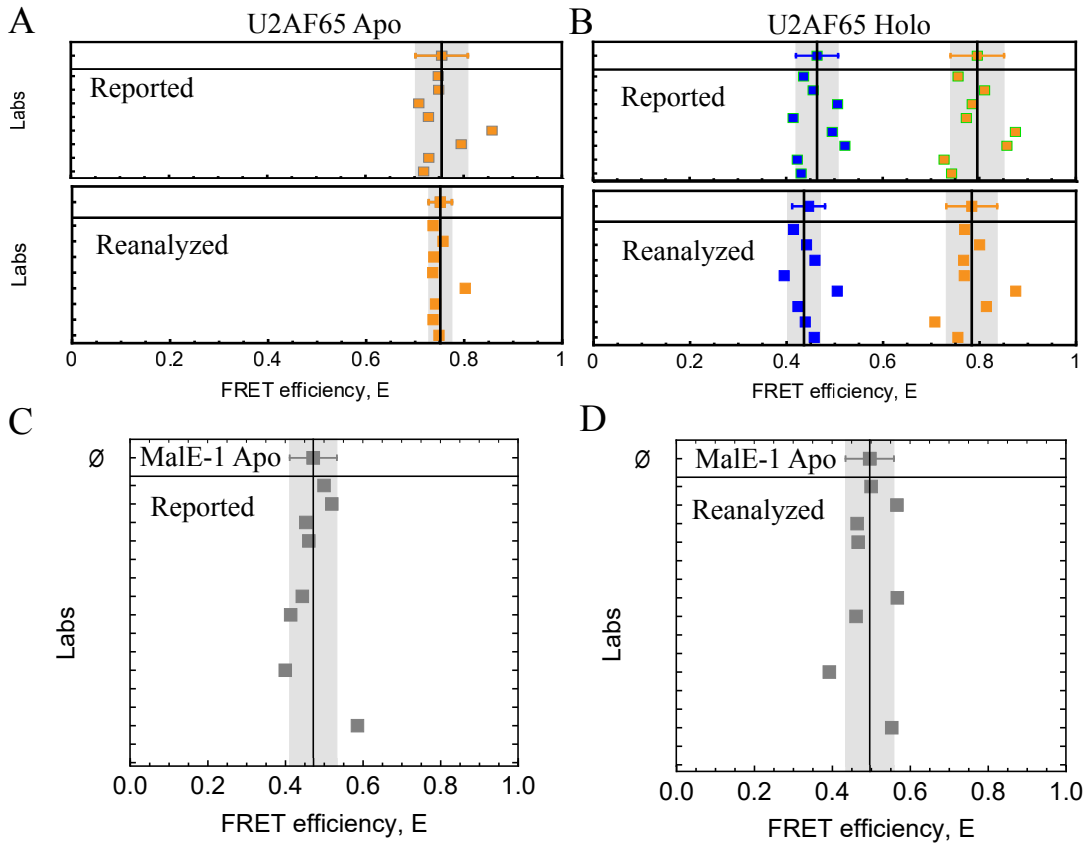

Figure 6.9: The consistency in FRET efficiency values for both U2AF65 and MalE-1 apo system from 8 labs after reanalysis. **(A-B)** U2AF65 reported FRET efficiencies (upper panel) are compared with the reanalyzed (Lower panel) for apo **(A)** and for holo **(B)**. **(C-D)** Reanalysis was performed on all MalE measurements from 8 labs but only shown here for MalE-1 apo as an example. the lab numbering is maintained in the case of MalE as for Figure 5.5 to allow a direct comparison. Reported **(C)** and reanalyzed **(D)** FRET efficiency values are plotted. The mean value is shown on each plot in an upper separate panel. The standard deviation from all the values obtained is shown as error bars around the mean values as the upper panels and the grey shaded region in the lower panels.

The *E*-τ plot in **Figure 6.12 C** for U2AF65 apo has a clear shift, in fact the highest shift observed so far, indicating the fast fluctuations on the sub-millisecond timescales in the system as expected for the ligand free protein (Voith von Voithenberg et al., 2016). The *E*-τ analysis on the holo conformation of U2AF65 is shown in **Figure 6.12 D.** The U9-stablized conformation comes closer to the static-line as expected, but the apo species in the holo measurement has the similar dynamic shift as for apo. The *E*-τ analysis which is sensitive to fast fluctuations highlights that U2AF apo is a highly dynamic family of states on the sub-millisecond timescale which upon binding to U9-RNA ligand becomes static (J. R. Huang et al., 2014). A U2AF65 BVA analysis shows a slight deviation due to these fast dynamics, but it is close to the boundary of the sensitivity BVA can detect.

On the basis of the above analyses on MalE and U2AF65, we compared the reported results from the groups. The overview of the qualitative results on both systems are summarized in Table 5.3. Groups submitted a statement on whether the system is dynamic or not.

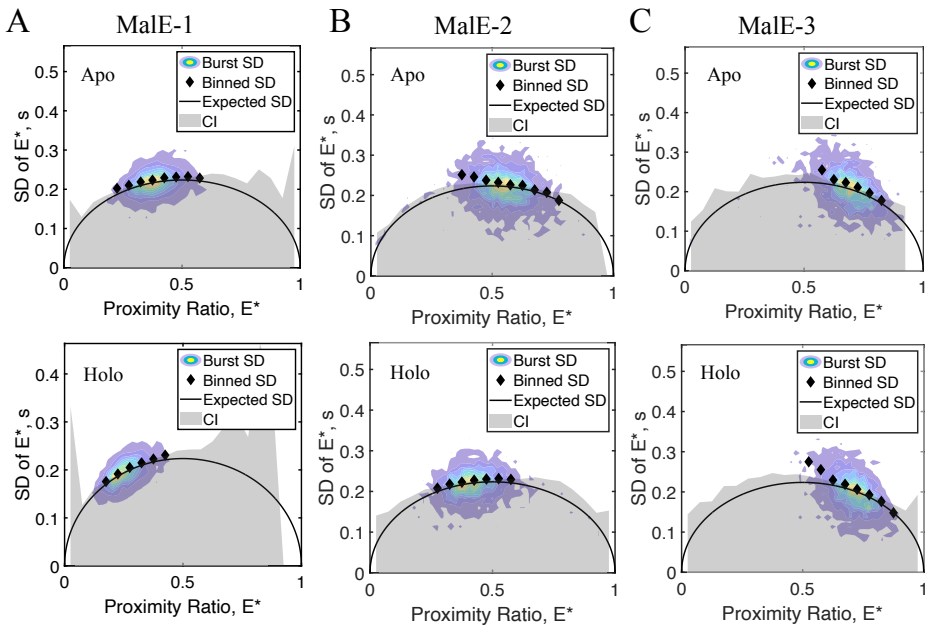

Figure 6.10: The Burst variance analysis (BVA) of MalE. **(A-C)** A BVA of MalE-1**(A)**, MalE-2 **(B)**, and MalE-3 **(C)** is shown. The Upper panels are for apo proteins and and lower panels for holo conditions. BVA analyzes the deviations in the burst beyond the shot-noise limit, which is calculated as $\sigma_{E^*} = \sqrt{\frac{E^*(1-E^*)}{n}}$ where σ_{E^*} is the expected standard deviation (black semi-circle) and n=5. The burst-wise variance is shown in the contour plot. Each population is binned with E^* of 0.05, variance in each bin is shown as black diamonds. Calculated 99% confidence intervals for shot-noise limited variance is shown in grey.

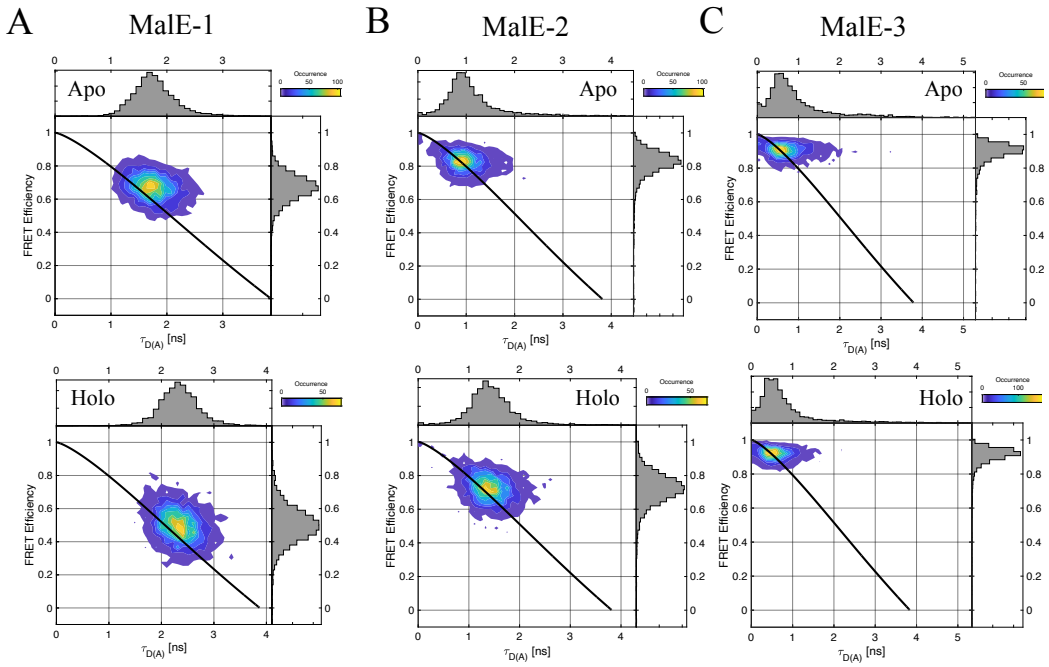

Figure 6.11: Dynamic analysis of MalE mutants with FRET efficiency versus donor lifetime analysis (E-τ analysis). **(A-C)** MalE-1 E-τ is shown in **(A)**, for MalE-2 in **(B),** and for MalE-3 in **(C)**. The upper panels shows the results for apo and the lower panels under holo conditions. The E-τ analysis is plotted as FRET efficiency vs donor lifetime in the presence of an acceptor ($\tau_{D(A)}$) where the ideal relationship between them is given by , $E = 1 - \frac{-0.01771+0.6185\,\langle\tau_{D(A)}\rangle+0.2185\,\langle\tau_{D(A)}\rangle^2+0.03\,\langle\tau_{D(A)}\rangle^3}{\langle\tau_{D(0)}\rangle}$ which is plotted as a black line. $\tau_{D(0)}$ is the donor lifetime in the absence of an acceptor. Deviation from this line to the longer donor lifetimes to the right is indicative of dynamics.

Some groups specialized in more advanced approaches like the photon distribution analysis (PDA) and filtered FCS (fFCS) also reported their feedback with these additional analyses. It can be clearly seen that most groups with three exceptions, reported MalE as static on the sub-millisecond time regime, whereas U2AF65 was detected correctly as dynamic by all 8 participating labs. The reason that one of the groups (#2) detected dynamics in all the MalE mutants as was based upon the slight deviation observed on the E-τ plots. This lead to calculation of limits in both BVA and E-τ analysis methods which one can say whether a system is dynamic or static. As a

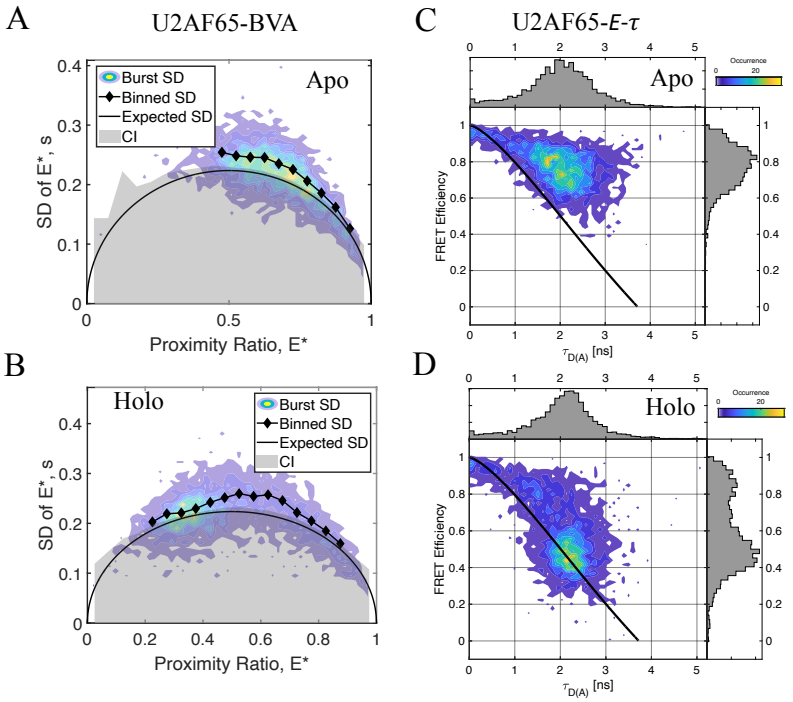

Figure 6.12: Analysis of U2AF65 mutants with the burst variance analysis (BVA) and a FRET efficiency analysis versus donor lifetime (E-τ analysis). **(A-B)** U2AF65 apo BVA is shown in **(A)** and for holo in (**B)**. BVA analyzes the deviations in the burst beyond the shot-noise which is calculated as $\sigma_{E^*} = \sqrt{\frac{E^*(1-E^*)}{n}}$ where σ_{E^*} is the expected standard deviation (black semi-circle) and n=5. The burst-wise variance is shown in the contour plot. Each population is binned with an E* of 0.05, the variance in each bin is shown as black diamonds. The calculated 99% confidence intervals for the shot-noise limited variance is shown in grey. **(C-D)** E-τ plot for U2AF65 apo is plotted in **(C)** and holo in **(D)**. The E-τ analysis is plotted as FRET efficiency vs donor lifetime in the presence of an acceptor ($\tau_{D(A)}$) where the ideal relationship between them is given by $E = 1 -$

$\frac{-0.01973+0.5597\,\langle\tau_{D(A)}\rangle+0.2665\,\langle\tau_{D(A)}\rangle^2+0.03798\,\langle\tau_{D(A)}\rangle^3}{\langle\tau_{D(0)}\rangle}$ where $\tau_{D(0)}$ is donor lifetime in the absence of acceptor is plotted as black line. A deviation from this line to the longer donor lifetimes is indicative of dynamics.

next step in finding the consistencies in quantitative analysis of dynamics, we compared the quantitative analysis on the U2AF65 system provided by 5 out of 8 labs. We asked the labs

Table 6.3: Reported dynamic results by the participating groups.

	Method	**Sub-millisecond dynamics Sub-ms dynamics in**	
		MalE-1/2/3	**U2AF65 (apo / holo)**
#1	BVA, E-τ	no / no / no	yes / yes
#2	E-τ, fFCS, PDA	yes / yes / yes	yes / yes
#3	BVA, E-τ	no / no / no	yes / yes
#4	BVA, E-τ	no / no / no	yes / yes
#5	n/a	n/a	n/a
#6	E-τ, fFCS	no / no / -	yes / yes
#7	BVA	no / - / yes	yes / yes
#8	BVA	no / no / no	yes / yes
#9	BVA, E-τ	no / no / no	n/a
#10	BVA	no / no / no	n/a
#11	BVA, E-τ	no / no / no	yes / yes
#12	BVA	no / no / no	n/a
#13	fFCS	no / no / yes	n/a
#14	BVA, E-τ	no / no / no	n/a
#15	n/a	n/a	n/a
#16	BVA	no / no / no	n/a

n.a.-not applicable. These Labs didn't report the analysis either because of the absence of suitable analysis tool or they didn't measure the protein.

that volunteered to perform a kinetic analysis only on the dynamic system of U2AF65 as MalE was known to have static behavior. Previous observations hinted on the complexities involved in interpreting and finding the kinetics of dynamics in the U2AF65 system. Additionally, there

can be a dependence of the model and the method employed to find the kinetic rates and relaxation time-scales of the interconversions. While performing the kinetic analysis, we asked the groups to report the various information which that can help to compare the analysis in a comprehensive way. First, we asked the groups to perform a donor fluorescence lifetime decay analysis with a bi-exponential model on both apo and holo measurements of U2AF65 (C. Eggeling et al., 1998). This analysis helped to define the states with specific donor lifetimes involved in the underlying interconversions independent of the kinetic model. Next, for the apo state, we asked the groups to employ correlation analysis methods as we had observed fast fluctuations in apo. The correlation analysis methods available are FRET-FCS (T. Torres & Levitus, 2007) and filtered-FCS (Felekyan et al., 2012). These methods are used to find the relaxation timescales covering a wide range of timescales but are often found in the microsecond regime. FRET-FCS correlates the anti-correlated signal in the FRET and donor detection channels to derive the time-scales of fluctuations underlying changes in the FRET efficiencies (T. Torres & Levitus, 2007). On the other hand, filtered-FCS extracts the relaxation times in their conformations for the defined species as states, where they are often defined by selecting the high and low FRET values for different species. Species are defined by increasing the contrast between them by taking into account all the information available from lifetime, color, and anisotropies (Felekyan et al., 2012). In the case of holo enzyme, we asked the groups to perform a dynamic photon distribution analysis (dPDA) assuming a two-state kinetic model. Dynamic PDA solves the proximity ratio histograms analytically by fitting the rates for interconverting species between the static states during the diffusion (Kalinin, Valeri, et al., 2010). To catch the kinetics, the molecular bursts are sliced into different timebin. Thus, the dPDA method is sensitive to the dynamics occurring on the timescale of diffusion.

Again, before comparing the results from different labs, we first delved into the description of kinetic analysis of U2AF65 system ourselves in detail. First, we performed a fluorescence decay analysis of the donor intensities with a bi-exponential fit for both apo and holo **(Figure 6.13 A)**.

The extracted donor lifetime components corresponded to the distances of ~38 Å and ~59-65 Å distances. The relative amplitudes for these populations change with more molecules in the extended in the holo measurements. Furthermore, with apo being highly dynamic as predicted from the E-τ plots, we employed the filtered-FCS method. We found the relaxation times of 100-200 µs for apo **(Figure 6.13 B)**. This explains the behavior in the E-τ plots where the high and low FRET conformations are faster than the diffusion by orders of magnitude and of only one averaged species is obtained. The donor lifetime, however, is biased to high lifetimes. From the E-τ plots and previous observations, it was known that, depending on the used dye-pair used for the measurements, varied proportions of the very high FRET state (~0.95) were detected **(Figure 6.12 C-D)** (Sánchez-Rico, Voith von Voithenberg, Warner, Lamb, & Sattler, 2017). When applying dPDA to the apo state, we described the main apo conformation (0.75 FRET efficiency) with a static population as PDA considers the bursts averaged over the diffusion time. A relaxation time, τ_R [$(1/(k_{12}+k_{21})$] of 1.5 ms was found for these interconversions with forward (k_{12}) and backward rates (k_{12}) of 0.57 ms^{-1} and 0.09 ms^{-1} for very high FRETstate to the high FRET **(Figure 6.13 C)**. To analyze the holo conformation, we applied dPDA where we treated the RNA-bound conformation (FRET efficiency of 0.44) static

and similarly described the apo-sepcies as a mixed population **(Figure 6.13 D)**. We found a similar relaxation time for the high FRET state to the very high FRET state of 1.8 ms with a fairly good fit. This suggested that the apo species is not significantly interacting with the RNA-bound conformation and can be fully described as apo alone. Considering all the above analyses, we describe U2AF65 model as depicted in **Figure 6.14**, where Apo conformation shows ensemble of fast exchanging conformations on the 100-200 µs (obtained by filtered-FCS), which exchanges slowly on a few ms timescales with apo closed conformation (obtained with dPDA) and RNA stabilizes the holo conformation where domains are far apart than the apo closed conformation.

With the above analyses, when we compared the rates from the 5 labs summarized in the Table 5, all the labs reported correctly the two-state behavior of the system. Three labs who analyzed the apo state with correlation methods reported two relaxation times with a mean±standard deviations of 297.7±88.6 µs and 8.1±3.7 µs. The slower dynamic rate has a value close to range we obtained with filtered FCS between 100-200 µs. In the case of the holo state, the considered distances of 37-38 Å and 56-60 Å by the groups agreed well with our analysis. Furthermore, the holo analysis being complicated by the apo species, we found a fair agreement with the reported slow relaxation times between 1.25-1.8 ms, but two groups reported fast dynamics most likely due to the higher fraction of apo species. As these fractions were known to change according to the measurement temperature conditions.

Overall, the above results were encouraging in terms of the qualitative analysis of the dynamics present in the proteins at different time regimes especially given the different methods, set-ups and configurations in data acquisitions, analyses employed by the groups evident with the collected different data formats and the observed correction factors. Another aspect that we learned about with respect to quantifying dynamics is that at least two different dye-pairs should be used to ensure that the dynamics measured is an intrinsic property of the protein and not dependent on the dye used. We found slightly different relaxation time-scales (0.46 ms) for slow exchange between the very high FRET state and high FRET state in apo for Alexa dye combinations. Interestingly, fast dynamics in the apo state does not change significantly indicating that the kinetics is inherent to the protein for the Atto dye-pair **(Figure 6.15)**.

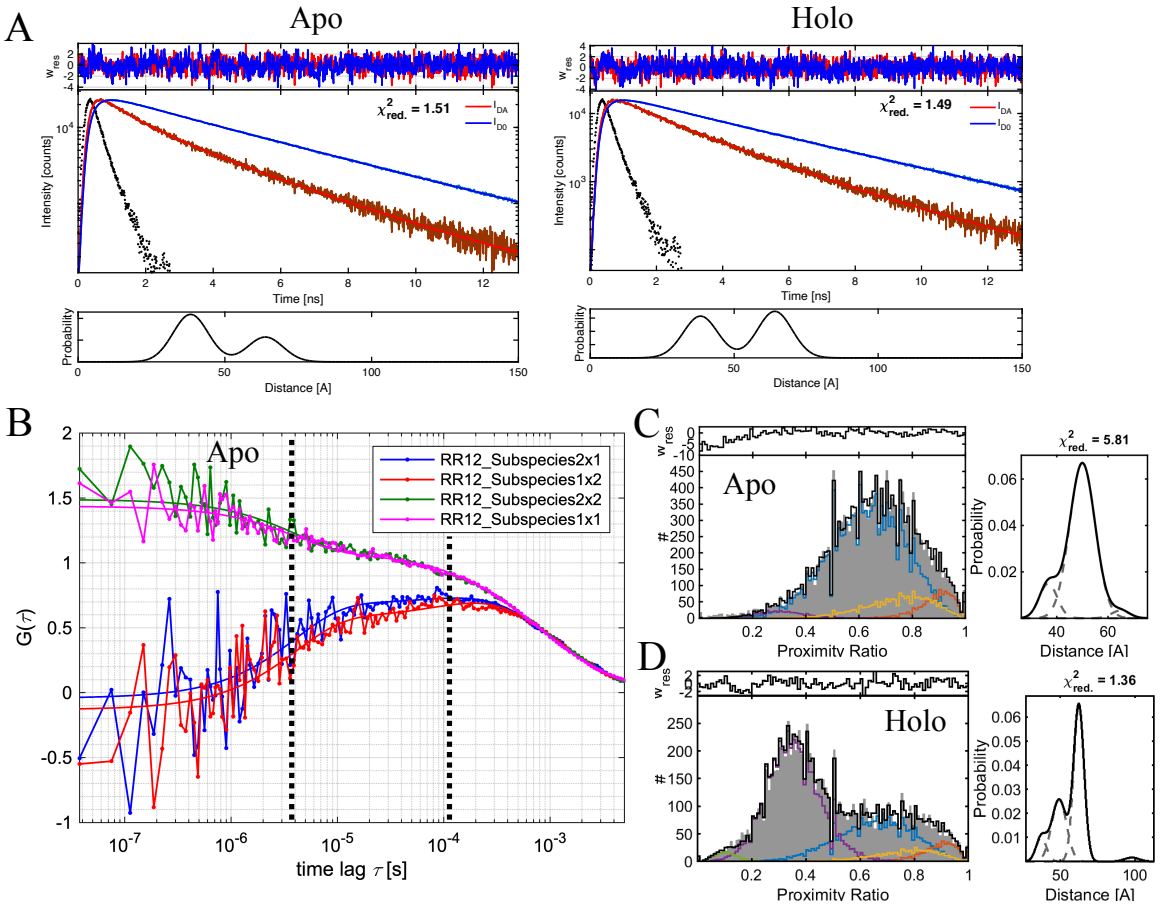

Figure 6.13: Quantitative dynamic analysis of U2AF65 labeled with Atto532-Atto643. **(A)** The donor lifetime decays fitted to a two-Gaussian distance distribution using the maximum entropy method. In blue is the donor decay without any acceptor and in red, in the presence of an acceptor. The dotted black line is the instrument response function. Apo is shown on the left and holo to the right. The lower panels present the distance distribution underlying the fit. **(B)** Filtered FCS correlation curves of apo-state and the corresponding global fit. Subspecies 1 and 2 are defined using the burst showing the highest and the lowest FRET values. The green and magenta are the respective autocorrelation functions of subspecies 1 and 2, while blue and red curves are depicting the cross-correlation function between them. The extracted relaxation times are shown as dotted vertical lines to guide the eye at ~3 µs and ~110 µs. **(C-D)** A dynamic PDA analysis. The proximity ratio histograms were binned in 0.5 ms, 1 ms, and 1.5 ms bins and were fitted to the two-state kinetic model to quantify the interconversion rates between very high FRET (orange) and, the apo state (blue). Only the 1 ms bin is shown here. The dynamic interconverting population is shown in yellow. Apo is shown in **(C)** and holo in **(D).** A small fraction of the holo state (purple) is needed to improve the fit for Apo. An additional green population was needed in the holo histogram to account for occasional acceptor bleaching events. The distance distributions extracted from the global PDA fit are shown to the right.

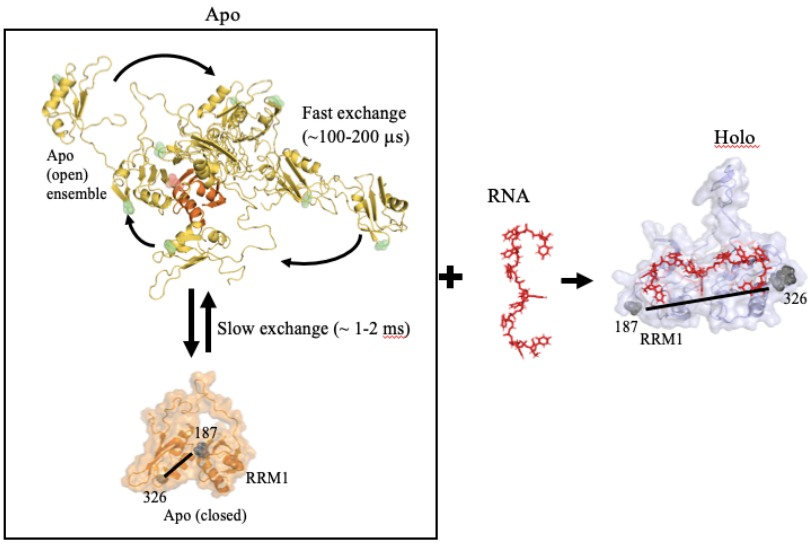

Figure 6.14: A model describing the U2AF65 dynamics. U2AF Apo open conformational ensemble show fast exchange on the time scales of 100-200 µs as evident in filtered-FCS curves (**Figure 6.13B**), which exchanges slowly to apo closed state as shown by dynamic PDA (**Figure 6.13C**). Upon binding to RNA, U2AF is stabilized in holo conformation.

Table 6.4: Quantitative description of the kinetic analysis of U2AF65 by 5 labs labeled with Atto532-Atto643.

Lab	Method	Apo	Holo
#1	Filtered-FCS	τ_{R1}=90-200 µs, τ_{R2} =3-10 µs	-
	Dynamic PDA (R_1=38 Å, R_2=59 Å)	k_{12}=0.43 ms^{-1}, k_{21}=0.07 ms^{-1} (τ_R =2 ms)	k_{12} =0.52 ms^{-1}, k_{21}=0.14 ms^{-1} (τ_R =1.42 ms)
#2	Filtered-FCS	τ_{R1} =320 µs, τ_{R2} =6 µs	τ_{R1} =320 µs, τ_{R2} =6 µs
	FRET-FCS	τ_{R1} =32 µs, τ_{R2}= 2 µs	τ_{R1} =321 µs, τ_{R2} =10 µs
	Dynamic PDA (R_1=37 Å, R_2=60 Å)	-	k_{12}=1.81 ms-1, k_{21}=2.18 ms^{-1} (τ_R =250 µs)
#8	Dynamic PDA (R_1=38 Å, R_2=56 Å)	-	k_{12}=0.47 ms^{-1}, k_{21}=0.15 ms^{-1} (τ_R =1.6 ms)
#11	Dynamic PDA (R_1=38 Å, R_2=56 Å)	-	k_{12}=0.63 ms^{-1}, k_{21}=0.83ms^{-1} (τ_R =680 µs)
#14	Filtered-FCS	τ_{R1} =370 µs, τ_{R2} =12 µs	τ_{R1} =617 µs, τ_{R2} =15 µs
	Dynamics PDA	k_{12}=2.3 ms^{-1}, k_{21} =4.6 ms^{-1}	k_{12}=0.47 ms^{-1}, k_{21} =0.15 ms^{-1}

	(R_1=37 Å, R_2=60 Å	(τ_R =144 µs)	(τ_R =1.6 ms)

R_1 and R_2 are the distances for compact and extended conformations derived from the dPDA fit. k_{12} and k_{21} are the forward and backward rates for the R_1 and R_2 distances extracted. τ_R is the calculated relaxation time from the k_{12} and k_{21} rates as $(1/(k_{12}+k_{21})$.

τ_{R1} and τ_{R2} are the relaxation times found using the correlation analyses.

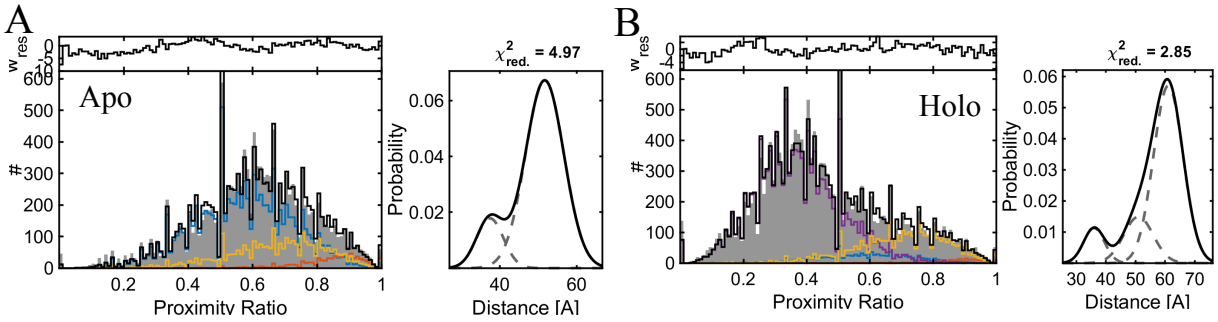

Figure 6.15: Dynamic PDA analysis of U2AF65 labeled with Alexa546-Alexa647. Proximity ratio histograms binned in 0.5 ms, 1 ms, and 1.5 ms time bins were fitted to the two-state kinetic model to quantify the interconversion rates between very high FRET (orange) and apo state (blue). Only the 1.5 ms bins are shown here. The dynamic interconverting population is shown in yellow. Apo is shown in **(A)** and holo in **(B)** panel. The distance distributions extracted from the global PDA fit is shown to the right.

6.3.4 Conclusion and outlook

The previous and current study of accuracy and consistency in smFRET measurements from dsDNA to proteins has been a remarkable journey in terms of establishing a reliable method for studying dynamic structural biology. Hellenkamp et. al. provided the first stepping stone by revealing the consistencies as good as ±0.05 in FRET efficiencies across 20 labs. Since its establishment, many proteins have been analyzed with smFRET. This study contributes by pushing the methodology in a broader sense to increase reliability on protein studies with smFRET. Handling proteins has its own challenges but provides valuable information about the working and functional principles controlling the basic cellular processes.

With the 16 participating groups, this study on proteins established a similar consistency in the FRET efficiencies on proteins as observed for dsDNA. Extending this finding to the consistencies in dynamic detection and analyses made this study a lot more challenging but more useful in terms of studying and extending the reliability of dynamic information provided by the smFRET based analytical approaches and its importance. Ultimately, the study of real protein systems is highly rewarding as an end application to link the dynamics and their functions. The study comprised two protein systems with different time-scales on the conformational motions as they undergo as they perform their function. The static system of MalE provided the tunability of biomolecules on slower time-scales fairly undetectable as dynamic in smFRET. On the other hand, the U2AF65 system provided the prototype of a system representing dynamic systems existing in biomolecules.

Complications in the measurements on the dynamic system was found to be particularly challenging in terms of sample preparations, measurement conditions (like solvent and temperature), ligand concentrations, and storage. The combined influence of these variables can make determination of the proper correction factors difficult. These kinds of effects may differ system to system and need to be conceived each system individually. In the case of U2AF65, the variations needed to be considered by calculating the detection correction factor differently. Nevertheless, labs that participated in the dynamic aspect of the study, were correctly able to distinguish the MalE as a static system and U2AF65 as a dynamic molecule. This result provides the strong support for the view of smFRET as a method to detect consistently dynamics even with various set-up configurations, measurements modes and analysis platforms. Furthermore, to compare the quantitative dynamic analyses across labs often depends on the model chosen for the system under study. Hence, model selection has a general bias in which direction efforts needed to be put forward. Model selection becomes challenging when studies are performed with only one dye-pair. This demands the usage of at least two dye-pairs to choose the analytical model correctly while reporting on the dynamics.

The studies are being performed to determine the limitations of the various approaches used for detecting and quantifying dynamic while considering all the aspects of dye photophysics, which makes things complicated due to rotation, blinking, bleaching, and varied linker lengths. Additional incorporation of other dye-labeling positions in the system provides new perspectives with respect to the distance changes of the protein in space and in time.

7 Summary and conclusion

The thesis work outlines the use of the smFRET methodology for spontaneous and chaperone-assisted protein folding, as well as the consistencies of smFRET measurements performed on proteins with participation from 16 labs.

For spontaneous folding, the folding order of two domains in DM-MBP was investigated extending the use of two-color FRET to three-color FRET. The study also comprised the effect of GroEL/ES chaperonin on conformational dynamics elucidated in this thesis work extending the chaperone-assisted protein folding mechanism. Furthermore, the effect of the chaperone systems, Hsp70 and Hsp90, was illustrated in the context of p53 folding at physiological temperatures. Lastly, a FRET comparison study on proteins was carried out to assess the accuracy and precision of the smFRET across 16 labs, with a focus on consistencies in the determination of FRET efficiency and detecting the dynamics.

Although folding is an indispensable process for most of the proteins, folding investigations have been limited to small and single-domain proteins. Hence, we studied MBP, a two-domain protein possessing discontinuity in its domains, in the context of its dynamics between the domains as protein folds. Particularly, we used DM-MBP, a slow folding mutant previously known to have an entropic trap in its folding pathway that it resolves by the GroEL chaperonin cavity. Separate two-color smFRET measurements on the individual domains led to an observation that, although the slow folding domain, NTD undergoes fast dynamic interconversions between the unfolded and native conformations, CTD domain waits until NTD folds. CTD achieves the folding due to a virtue of its sub-millisecond dynamic fluctuations between the unfolded and near-native conformation. The same observation was confirmed via three-color smFRET measurements on the same protein molecule concerning the folding order of the domains. Although the order of domain folding is conserved from wt-MBP to the slow folding version, DM-MBP acquires an entropic barrier due to the mutations in the folding nucleus, making it to sample conformations fast on sub-milliseconds. The observed folding mechanism can be important in proteins possessing domains with a significant structural overlap in their final structures, typical for enzymes and regulatory proteins. GroEL/ES chaperonin cavity is known to decrease the configurational entropy to stabilize the native state. On this line, we verified the effect of the GroEL/ES cavity on the dynamics present in DM-MBP folding. Interestingly, DM-MBP shows similar conformational fluctuations when encapsulated in the cavity as to spontaneous folding in solution, but equilibrium is shifted to more native-like conformation. The study entailed the presence of dynamics in the conformations when multi-domain protein with a discontinuous arrangement undergoes folding that is limited by an entropic barrier and also illustrates the power of multi-color FRET on reporting coordinated motions in biomolecules. On this line of work, future folding studies can be devised on large, multi-domain proteins with complex topology in both spontaneous as well as chaperone-assisted folding contexts.

Another study involving chaperone-assisted protein folding aspect, where we probed the effect of Hsp70 and Hsp90 chaperones on p53-DBD conformations directly by smFRET. Decision on cellular fate is important to control tumor and cancer instances, and their progression. p53 is known as tumor suppressor protein and a guardian of the genome, attributed to its function as a transcriptional factor to upregulate the signaling cascade required to repair DNA damage. Interestingly, owing to p53 instability at cellular temperatures of 37°C, its activity was known to depend heavily on its interactions with Hsp70 and Hsp90 chaperone systems. Our study unraveled a complex and dynamic interplay where under normal cellular conditions, Hsp70 keeps p53 in an inactive state by stabilizing the unfolded state, and Hsp90 together with other cochaperones shifts the equilibrium to a folded state of p53 when the cell encounters the stress. The opposing actions of these chaperone systems can be achieved via their relative cellular concentrations or modulating the chaperone activities. The study explains the previous observations made regarding the Hsp70 and Hsp90 interactions with p53 under varying cellular conditions.

SmFRET is a heavily used tool in similar biological investigations as used in this thesis work. Access to fast timescales of nanoseconds to microseconds, easy implementation, flexibility in both confocal and TIRF modalities, and ease to test the hypothesis are a few of the several advantages of the smFRET approach. The previous study demonstrated smFRET as an accurate and precise method in a research community. The study involved 20 labs to assess the accuracy and precision in reporting the FRET efficiencies on the same DNA samples. To expand the scope to a variety of biomolecules, we followed up the study on protein samples with a focus on investigating the accuracy in detecting the dynamics across the labs. Focusing on the dynamics, the study was performed on a static system of MBP and a dynamic system of U2AF. Reported FRET efficiency values were as accurate as on DNAs, reiterating the previous study. Protein systems have many challenges such as difficulty in handling and storing them because their less stable nature, as well as their composition makes them difficult to label and avoid any unexpected interactions with dyes. These and many other things led us to come across a few new aspects during this study. Correction of the detection factor was not straightforward in the case of a U2AF system because of the presence of fast dynamics, owing to the less efficient tuning of ligand-induced conformational change than expected, and RNA as a ligand. Additional challenges emerged because of the change in the stoichiometry for one of the population most likely due to either protein-induced dye quenching or multi-molecular associations. Moreover, the focus being on the discovery of dynamics, most of the labs correctly reported the absence and presence of dynamics in the case of MBP and U2AF respectively. Certain dye behaviors when attached to biomolecules may lead to an interpretation of the presence of dynamics and also methods used for detecting dynamics have their own set of limitations. To this end, we are still investigating the limits to detect the dynamics comparing different dye pairs and various approaches to quantify the dynamics. Overall, the study represents smFRET as a reliable tool to detect and quantify dynamics in molecular motions.

7 Summary and conclusion

Abbreviations

2CDE	Two channel kernel density estimation
A	Adenine
A	Alanine
ADP	Adenosine diphosphate
ALEX	Alternating laser excitation
APBS	All photon burst search
APD	Avalanche photodiode
ATP	Adenosine triphosphate
BAP	BiP associated protein
BCA	Bicinchoninic acid
BSA	Bovine
BVA	Burst variance analysis
C	Cysteine
CF	Correction factor
CTD	C-terminus domain
CuAAC	Copper catalyzed Azide-Alkyl cyclo-addition
Cy	Cyanine
DBD	DNA-binding domain
DCBS	Dual color burst search
DM-MBP	Double mutant-Maltose binding protein
DNA	Deoxyribose nucleic acid
DOL	Degree of labeling
dPDA	Dynamic photon distribution analysis

DTT	Dithiothreitol
E. coli	Escherichia coli
EDTA	Ethyl-diamine tetra acetic acid
EF	Emisssion filter
EM	Electron microscope
EPR	Electron paramagnetic resonance
ER	endoplasmic reticulum
ERAD	ER associated degradation
FCCS	Fluorescence cross-correlation spectroscopy
FCS	Fluorescence correlation spectroscopy
fFCS	Filtered FCS
fL	Femto-litre
FLIM	Fluorescence Lifetime imaging
FRET	Förster resonance energy transfer
GRP	Glucose regulated protein
GuHCl	Guanidinium hydrochloride
HDX	Hydrogen-deuterium exchange
Hsp	Heat shock protein
IPTG	Isopropyl β- d-1-thiogalactopyranoside
IRF	Instrument response function
K_D	Dissociation constant
LB	Luria bertani
MD	Middle domain
MEEVD	Methionine-Glutamate-Glutamate-Valine-Aspartate
MEEVD	Methionine-Glutamate-Glutamate-Valine-Aspartate

MFD	Multi-parameter fluorescence detection
N-C	NTD-CTD
NA	Numerical aperture
NBD	Nucleotide binding domain
NMR	Nuclear magnetic resonance
NTD	N-terminus domain
NTD	N-terminus domain
PALM	Photo activated localization microscopy
PAM	PIE analysis in Matlab
PBS	Phosphate buffer saline
PET	Photo electron transfer
PIC	Protease cocktail inhibitor
PMSF	Phenylmethylsulfonyl fluoride
RAC	Ribosome associated complex
RICS	Raster image correlation spectroscopy
RNA	Ribonucleic acid
RRM	RNA recognizing motif
SAXS	Small angle X-ray scattering
SBD	Substrate binding domain
sHsp	Small heat shock protein
smFRET	Single molecule FRET
SRP	Signal recognizing protein
STED	Stimulate demission depletion
STORM	Stochastic reconstruction microscopy
TAE	Tris Acetate EDTA

TBTA	Tris(benzyltriazolyolylmethyl)amine
TCEP	Tris(2-carboxyethyl)phosphine
TCSPC	Time correlated single photon counting
TE	Tris EDTA
TIRF	Total Internal Reflection microscopy
TPR	Tetratricopeptide repeat
U2AF	U2 auxiliary factor
UAA	Unnatural amino acid
UPR	Unfolded protein response
WT	wild-type
XFEL	X-ray free electron laser

Bibliography

Acuna, G. P., Möller, F. M., Holzmeister, P., Beater, S., Lalkens, B., & Tinnefeld, P. (2012). Fluorescence enhancement at docking sites of DNA-directed self-assembled nanoantennas. *Science*. https://doi.org/10.1126/science.1228638

Agashe, Vishwas R., Guha, S., Chang, H. C., Genevaux, P., Hayer-Hartl, M., Stemp, M., … 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*(2), 199–209. https://doi.org/10.1016/S0092-8674(04)00299-5

Agashe, Vlshwas R., & Udgaonkar, M. C. R. S. B. J. B. (1995). *Initial hydrophobic collapse in the folding of barstar*. *Nature* (Vol. 377). https://doi.org/10.1038/246170a0

Ali, M. M. U., Mark Roe, S., Vaughan, C. K., Meyer, P., Panaretou, B., Piper, P. W., … Pearl, L. H. (2006). Crystal structure of an Hsp90-nucleotide-p23/Sba1 closed chaperone complex. *Nature*. https://doi.org/10.1038/nature04716

Andrews, B. T., Capraro, D. T., Sulkowska, J. I., Onuchic, J. N., & Jennings, P. A. (2013). Hysteresis as a marker for complex, overlapping landscapes in proteins. *Journal of Physical Chemistry Letters*, *4*(1), 180–188. https://doi.org/10.1021/jz301893w

Anfinsen. (1973). ’ CIE : NCES Folding of Protein Chains. *Science*, *181*(4096), 223–230.

Antonik, M., Felekyan, S., Gaiduk, A., & Seidel, C. A. M. (2006). Separating structural heterogeneities from stochastic variations in fluorescence resonance energy transfer distributions via photon distribution analysis. *Journal of Physical Chemistry B*, *110*(13), 6970–6978. https://doi.org/10.1021/jp057257+

Apetri, A. C., & Horwich, A. L. (2008). Chaperonin chamber accelerates protein folding through passive action of preventing aggregation. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(45), 17351–17355. https://doi.org/10.1073/pnas.0809794105

Arai, M., Iwakura, M., Matthews, C. R., & Bilsel, O. (2011). Microsecond subdomain folding in dihydrofolate reductase. *Journal of Molecular Biology*, *410*(2), 329–342. https://doi.org/10.1016/j.jmb.2011.04.057

Arviv, O., & Levy, Y. (2012). Folding of multidomain proteins: Biophysical consequences of tethering even in apparently independent folding. *Proteins: Structure, Function and Bioinformatics*, *80*(12), 2780–2798. https://doi.org/10.1002/prot.24161

Axelrod, D. (1981). *Cell-Substrate Contacts Illuminated by Total Internal Reflection Fluorescence RAPID COMMUNICATIONS. THE JOURNAL OF CELL BIOLOGY* (Vol. 89).

Baker, M. J., Tatsuta, T., & Langer, T. (2011). Quality control of mitochondrial proteostasis. *Cold Spring Harbor Perspectives in Biology*. https://doi.org/10.1101/cshperspect.a007559

Balchin, D., Hayer-Hartl, M., & Hartl, F. U. (2016, July 1). In vivo aspects of protein folding and quality control. *Science*. American Association for the Advancement of Science. https://doi.org/10.1126/science.aac4354

Bandyopadhyay, A., Saxena, K., Kasturia, N., Dalal, V., Bhatt, N., Rajkumar, A., … Chakraborty, K. (2012). Chemical chaperones assist intracellular folding to buffer mutational variations. *Nature Chemical Biology*, *8*(3), 238–245.

https://doi.org/10.1038/nchembio.768

Barends, T. R. M., Brosi, R. W. W., Steinmetz, A., Scherer, A., Hartmann, E., Eschenbach, J., … Reinstein, J. (2013). Combining crystallography and EPR: Crystal and solution structures of the multidomain cochaperone DnaJ. *Acta Crystallographica Section D: Biological Crystallography*. https://doi.org/10.1107/S0907444913010640

Barth, A., Hendrix, J., Fried, D., Barak, Y., Bayer, E. A., & Lamb, D. C. (2018). Dynamic interactions of type I cohesin modules fine-tune the structure of the cellulosome of Clostridium thermocellum. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(48), E11274–E11283. https://doi.org/10.1073/pnas.1809283115

Barth, A., Voith Von Voithenberg, L., & Lamb, D. C. (2019a). Quantitative Single-Molecule Three-Color Förster Resonance Energy Transfer by Photon Distribution Analysis. *Journal of Physical Chemistry B*, *123*(32), 6901–6916. research-article. https://doi.org/10.1021/acs.jpcb.9b02967

Barth, A., Voith Von Voithenberg, L., & Lamb, D. C. (2019b). SI-Quantitative Single-Molecule Three-Color Förster Resonance Energy Transfer by Photon Distribution Analysis. *Journal of Physical Chemistry B*, *123*(32), 6901–6916. https://doi.org/10.1021/acs.jpcb.9b02967

Batey, S., Nickson, A. A., & Clarke, J. (2008). Studying the folding of multidomain proteins. *HFSP Journal*, *2*(6), 365–377. https://doi.org/10.2976/1.2991513

Bechtluft, P., van Leeuwen, R. G. H., Tyreman, M., Tomkiewicz, D., Nouwen, N., Tepper, H. L., … Tans, S. J. (2007). Direct observation of chaperone-induced changes in a protein folding pathway. *Science (New York, N.Y.)*, *318*(5855), 1458–1461. https://doi.org/10.1126/science.1144972

Behnke, J., Feige, M. J., & Hendershot, L. M. (2015, April 10). BiP and Its Nucleotide Exchange Factors Grp170 and Sil1: Mechanisms of Action and Biological Functions. *Journal of Molecular Biology*. Academic Press. https://doi.org/10.1016/j.jmb.2015.02.011

Bertelsena, E. B., Chang, L., Gestwicki, J. E., & Zuiderweg, E. R. P. (2009). *Solution conformation of wild-type E. coli Hsp70 (DnaK) chaperone complexed with ADP and substrate*. *Proceedings of the National Academy of Sciences of the United States of America* (Vol. 106). https://doi.org/10.1073/pnas.0903503106

Bertz, M., & Rief, M. (2008). Mechanical Unfoldons as Building Blocks of Maltose-binding Protein. *Journal of Molecular Biology*, *378*(2), 447–458. https://doi.org/10.1016/j.jmb.2008.02.025

Best, R. B., Hummer, G., & Eaton, W. A. (2013). Native contacts determine protein folding mechanisms in atomistic simulations. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(44), 17874–17879. https://doi.org/10.1073/pnas.1311599110

Bhushan, S., Gartmann, M., Halic, M., Armache, J. P., Jarasch, A., Mielke, T., … Beckmann, R. (2010). α-Helical nascent polypeptide chains visualized within distinct regions of the ribosomal exit tunnel. *Nature Structural and Molecular Biology*. https://doi.org/10.1038/nsmb.1756

Blanchet, C. E., & Svergun, D. I. (2013). Small-angle X-ray scattering on biological macromolecules and nanocomposites in solution. *Annual Review of Physical Chemistry*.

https://doi.org/10.1146/annurev-physchem-040412-110132

Blanco, F. J., Rivas, G., & Serrano, L. (1994). A short linear peptide that folds into a native stable β-hairpin in aqueous solution. *Nature Structural & Molecular Biology*, *1*(9), 584–590. https://doi.org/10.1038/nsb0994-584

Boorstein, W. R., Ziegelhoffer, T., & Craig, E. A. (1994). Molecular evolution of the HSP70 multigene family. *Journal of Molecular Evolution*. https://doi.org/10.1007/BF00175490

Borgia, A., Borgia, M. B., Bugge, K., Kissling, V. M., Heidarsson, P. O., Fernandes, C. B., … Schuler, B. (2018). Extreme disorder in an ultrahigh-affinity protein complex. *Nature*, *555*(7694), 61–66. https://doi.org/10.1038/nature25762

Borgia, M. B., Borgia, A., Best, R. B., Steward, A., Nettels, D., Wunderlich, B., … Clarke, J. (2011). Single-molecule fluorescence reveals sequence-specific misfolding in multidomain proteins. *Nature*, *474*(7353), 662–665. https://doi.org/10.1038/nature10099

Boutet, S., Lomb, L., Williams, G. J., Barends, T. R. M., Aquila, A., Doak, R. B., … Schlichting, I. (2012). High-Resolution Protein Structure Determination by Serial Femtosecond Crystallography. *Science*. https://doi.org/10.1126/science.1217737

Bracher, A., & Verghese, J. (2015). The nucleotide exchange factors of Hsp70 molecular chaperones. *Frontiers in Molecular Biosciences*, *2*(APR), 10. https://doi.org/10.3389/fmolb.2015.00010

Braig, K., Otwinowski, Z., Hegde, R., Boisvert, D. C., Joachimiak, A., Horwich, A. L., & Sigler, P. B. (1994). The crystal structure of the bacterial chaperonln GroEL at 2.8 Å. *Nature*. https://doi.org/10.1038/371578a0

Brodsky, J. L. (2012). Cleaning Up: ER-associated degradation to the rescue. *Cell*. https://doi.org/10.1016/j.cell.2012.11.012

Bryngelson, J. D., Onuchic, J. N., Socci, N. D., & Wolynes, P. G. (1995). Funnels, pathways, and the energy landscape of protein folding: A synthesis. *Proteins: Structure, Function, and Bioinformatics*, *21*(3), 167–195. https://doi.org/10.1002/prot.340210302

C. Zander, Sauer, M., Drexhage, K. H., Ko, D.-S., Schulz, A., Wolfrum, J., … Seide, C. A. M. (1996). Detection and characterization of single molecules in aqueous solution. *Applied Physics B*, *63*, 517–523.

Calloni, G., Chen, T., Schermann, S. M., Chang, H. C., Genevaux, P., Agostini, F., … Hartl, F. U. (2012). DnaK Functions as a Central Hub in the E. coli Chaperone Network. *Cell Reports*. https://doi.org/10.1016/j.celrep.2011.12.007

Case, D. A., Cerutti, D. S., Cheatham, T. E. I., Darden, T. A., Duke, R. E., Giese, T. J., … Kollman, P. A. (2017). *Amber 2017. University of California, San Francisco*.

Cerminara, M., Schöne, A., Ritter, I., Gabba, M., & Fitter, J. (2020). Mapping Multiple Distances in a Multidomain Protein for the Identification of Folding Intermediates. *Biophysical Journal*, *118*(3), 688–697. https://doi.org/10.1016/j.bpj.2019.12.006

Chakraborty, K., Chatila, M., Sinha, J., Shi, Q., Poschner, B. C., Sikor, M., … Hayer-Hartl, M. (2010). Chaperonin-Catalyzed Rescue of Kinetically Trapped States in Protein Folding. *Cell*, *142*(1), 112–122. https://doi.org/10.1016/j.cell.2010.05.027

Chang, H. C., Kaiser, C. M., Hartl, F. U., & Barral, J. M. (2005). De novo folding of GFP fusion proteins: High efficiency in eukaryotes but not in bacteria. *Journal of Molecular Biology*, *353*(2), 397–409. https://doi.org/10.1016/j.jmb.2005.08.052

Chatterjee, A., Xiao, H., & Schultz, P. G. (2012). Evolution of multiple, mutually orthogonal prolyl-tRNA synthetase/tRNA pairs for unnatural amino acid mutagenesis in Escherichia coli. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.1212454109

Chatterjee, Abhishek, Sun, S. B., Furman, J. L., Xiao, H., & Schultz, P. G. (2013). A versatile platform for single- and multiple-unnatural amino acid mutagenesis in escherichia coli. *Biochemistry*. https://doi.org/10.1021/bi4000244

Chen, B., Zhong, D., & Monteiro, A. (2006). Comparative genomics and evolution of the HSP90 family of genes across all kingdoms of organisms. *BMC Genomics*. https://doi.org/10.1186/1471-2164-7-156

Chen, S., Roseman, A. M., Hunter, A. S., Wood, S. P., Burston, S. G., Ranson, N. A., … Saibil, H. R. (1994). Location of a folding protein and shape changes in GroEL-GroES complexes imaged by cryo-electron microscopy. *Nature*. https://doi.org/10.1038/371261a0

Cheng, M. Y., Hartl, F. U., Martin, J., Pollock, R. A., Kalousek, F., Neuper, W., … Horwich, A. L. (1989). Mitochondrial heat-shock protein hsp60 is essential for assembly of proteins imported into yeast mitochondria. *Nature*, *337*(6208), 620–625. https://doi.org/10.1038/337620a0

Cheng, Y. (2015). Single-particle Cryo-EM at crystallographic resolution. *Cell*. https://doi.org/10.1016/j.cell.2015.03.049

Chiti, F., & Dobson, C. M. (2017). Protein misfolding, amyloid formation, and human disease: A summary of progress over the last decade. *Annual Review of Biochemistry*. https://doi.org/10.1146/annurev-biochem-061516-045115

Chothia, C., Gough, J., Vogel, C., & Teichmann, S. A. (2003). Evolution of the protein repertoire. *Science*, *300*(5626), 1701–1703. https://doi.org/10.1126/science.1085371

Chun, S. Y., Strobel, S., Bassford, P., & Randall, L. L. (1993). Folding of maltose-binding protein. Evidence for the identity of the rate- determining step in vivo and in vitro. *Journal of Biological Chemistry*, *268*(28), 20855–20862.

Chung, H. S., McHale, K., Louis, J. M., & Eaton, W. a. (2012). Single-Molecule Fluorescence Experiments Determine Protein Folding Transition Path Times. *Science*, *335*(6071), 981–984. https://doi.org/10.1126/science.1215768

Chung, Hoi Sung, Meng, F., Kim, J. Y., McHale, K., Gopich, I. V., & Louis, J. M. (2017). Oligomerization of the tetramerization domain of p53 probed by two- and three-color single-molecule FRET. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.1700357114

Chung, Hoi Sung, Piana-Agostinetti, S., Shaw, D. E., & Eaton, W. A. (2015). Structural origin of slow diffusion in protein folding. *Science*, *349*(6255), 1504–1510. https://doi.org/10.1126/science.aab1369

Clamme, J.-P., & Deniz, A. A. (2005). Three-color single-molecule fluorescence resonance energy transfer. *Chemphyschem : A European Journal of Chemical Physics and Physical Chemistry*, *6*(1), 74–77. https://doi.org/10.1002/cphc.200400261

Clausen, M. P., Sezgin, E., Bernardino de la Serna, J., Waithe, D., Lagerholm, B. C., & Eggeling, C. (2015). A straightforward approach for gated STED-FCS to investigate lipid membrane dynamics. *Methods*. https://doi.org/10.1016/j.ymeth.2015.06.017

Clegg, R. M., Zechel, A., Carlberg, C., Diekmann, S., Murchie, A. I. H., & Lilley, D. M. J. (1992). Fluorescence Resonance Energy Transfer Analysis of the Structure of the Four-Way DNA Junction. *Biochemistry*, *31*(20), 4846–4856. https://doi.org/10.1021/bi00135a016

Cohen, A. E., & Moerner, W. E. (2006). Suppressing Brownian motion of individual biomolecules in solution. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.0509976103

Collinet, B., Hervé, M., Pecorari, F., Minard, P., Eder, O., & Desmadril, M. (2000). Functionally accepted insertions of proteins within protein domains. *Journal of Biological Chemistry*, *275*(23), 17428–17433. https://doi.org/10.1074/jbc.M000666200

Coullomb, A., Bidan, C. M., Qian, C., Wehnekamp, F., Oddou, C., Albigès-Rizo, C., … Dupont, A. (2020). QuanTI-FRET: a framework for quantitative FRET measurements in living cells. *Scientific Reports*. https://doi.org/10.1038/s41598-020-62924-w

Cunningham, C. N., Krukenberg, K. A., & Agard, D. A. (2008). Intra- and intermonomer interactions are required to synergistically facilitate ATP hydrolysis in Hsp90. *Journal of Biological Chemistry*. https://doi.org/10.1074/jbc.M800046200

Dahiya, V., Agam, G., Lawatscheck, J., Rutz, D. A., Lamb, D. C., & Buchner, J. (2019). Coordinated Conformational Processing of the Tumor Suppressor Protein p53 by the Hsp70 and Hsp90 Chaperone Machineries. *Molecular Cell*, *74*(4), 816-830.e7. https://doi.org/10.1016/j.molcel.2019.03.026

De Boer, M., Gouridis, G., Vietrov, R., Begg, S. L., Schuurman-Wolters, G. K., Husada, F., … Cordes, T. (2019). Conformational and dynamic plasticity in substrate-binding proteins underlies selective transport in ABC importers. *ELife*. https://doi.org/10.7554/eLife.44652

De Torres, J., Ghenuche, P., Moparthi, S. B., Grigoriev, V., & Wenger, J. (2015). FRET enhancement in aluminum zero-mode waveguides. *ChemPhysChem*. https://doi.org/10.1002/cphc.201402651

Deniz, A. A., Laurence, T. A., Beligere, G. S., Dahan, M., Martin, A. B., Chemla, D. S., … Weiss, S. (2000). Single-molecule protein folding: Diffusion fluorescence resonance energy transfer studies of the denaturation of chymotrypsin inhibitor 2. *Proceedings of the National Academy of Sciences*, *97*(10), 5179–5184. https://doi.org/10.1073/pnas.090104997

Deniz, Ashok A., Dahan, M., Grunwell, J. R., Ha, T., Faulhaber, A. E., Chemla, D. S., … Schultz, P. G. (1999). Single-pair fluorescence resonance energy transfer on freely diffusing molecules: Observation of Förster distance dependence and subpopulations. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(7), 3670–3675. https://doi.org/10.1073/pnas.96.7.3670

Deuerling, E., & Bukau, B. (2004). Chaperone-assisted folding of newly synthesized proteins in the cytosol. *Critical Reviews in Biochemistry and Molecular Biology*. https://doi.org/10.1080/10409230490892496

Deuerling, E., Gamerdinger, M., & Kreft, S. G. (2019). Chaperone interactions at the ribosome. *Cold Spring Harbor Perspectives in Biology*. https://doi.org/10.1101/cshperspect.a033977

Digman, M. A., Brown, C. M., Sengupta, P., Wiseman, P. W., Horwitz, A. R., & Gratton, E. (2005). Measuring fast dynamics in solutions and cells with a laser scanning microscope.

Biophysical Journal. https://doi.org/10.1529/biophysj.105.062836

Dimura, M., Peulen, T. O., Hanke, C. A., Prakash, A., Gohlke, H., & Seidel, C. A. (2016). Quantitative FRET studies and integrative modeling unravel the structure and dynamics of biomolecular systems. *Current Opinion in Structural Biology*, *40*, 163–185. https://doi.org/10.1016/j.sbi.2016.11.012

Dinner, A. R., & Karplus, M. (1999). Is protein unfolding the reverse of protein folding? A lattice simulation analysis. *Journal of Molecular Biology*, *292*(2), 403–419. https://doi.org/10.1006/jmbi.1999.3051

Dyson, H. J., & Wright, P. E. (2002). Coupling of folding and binding for unstructured proteins. *Current Opinion in Structural Biology*. https://doi.org/10.1016/S0959-440X(02)00289-0

Eggeling, C., Fries, J. R., Brand, L., Günther, R., & Seidel, C. A. M. (1998). Monitoring conformational dynamics of a single molecule by selective fluorescence spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(4), 1556–1561. https://doi.org/10.1073/pnas.95.4.1556

Eggeling, C, Berger, S., Brand, L., Fries, J. R., Schaffer, J., Volkmer, A., & Seidel, C. A. M. (2001). *Data registration and selective single-molecule analysis using multi-parameter fluorescence detection*. *Journal of Biotechnology* (Vol. 86). Retrieved from www.elsevier.com/locate/jbiotec

Eggeling, Christian, Widengren, J., Brand, L., Schaffer, J., Felekyan, S., & Seidel, C. A. M. (2006). Analysis of photobleaching in single-molecule multicolor excitation and Förster resonance energy transfer measurements. *Journal of Physical Chemistry A*. https://doi.org/10.1021/jp054581w

Elson, E. L., & Magde, D. (1974). Fluorescence correlation spectroscopy. I. Conceptual basis and theory. *Biopolymers*, *13*(1), 1–27. https://doi.org/10.1002/bip.1974.360130102

Engelman, D. M., & Moore, P. B. (1975). Determination of quaternary structure by small angle neutron scattering. *Annual Review of Biophysics and Bioengineering*. https://doi.org/10.1146/annurev.bb.04.060175.001251

Faye, K., Pobre, R., Poet, G. J., Linda, X., Hendershot, M., & Allewell, N. M. (2019). The endoplasmic reticulum (ER) chaperone BiP is a master regulator of ER functions: Getting by with a little help from ERdj friends Downloaded from. *REVIEWS 2098 J. Biol. Chem*, *294*(6), 2098–2108. https://doi.org/10.1074/jbc.REV118.002804

Felekyan, S., Kalinin, S., Sanabria, H., Valeri, A., & Seidel, C. A. M. (2012). Filtered FCS : Species Auto- and Cross-Correlation Functions Highlight Binding and Dynamics in Biomolecules **, 1036–1053. https://doi.org/10.1002/cphc.201100897

Fenton, W. A., Kashi, Y., Furtak, K., & Horwich, A. L. (1994). Residues in chaperonin GroEL required for polypeptide binding and release. *Nature*. https://doi.org/10.1038/371614a0

Ferreon, A. C. M., Ferreon, J. C., Wright, P. E., & Deniz, A. A. (2013). Modulation of allostery by protein intrinsic disorder. *Nature*, *498*(7454), 390–394. https://doi.org/10.1038/nature12294

Fitter, J. (2009). The perspectives of studying multi-domain protein folding. *Cellular and Molecular Life Sciences*, *66*(10), 1672–1681. https://doi.org/10.1007/s00018-009-8771-9

Flaherty, K. M., Wilbanks, S. M., DeLuca-Flaherty, C., & McKay, D. B. (1994). Structural basis of the 70-kilodalton heat shock cognate protein ATP hydrolytic activity. II. Structure of the active site with ADP or ATP bound to wild type and mutant ATPase fragment.

Journal of Biological Chemistry. https://doi.org/10.2210/pdb1nga/pdb

Forster, T. (1946). Energiewanderung und Fluoreszenz. *Die Naturwissenschaften*, *33*(6), 166–175. https://doi.org/10.1007/BF00585226

Frank, J. (2017). Time-resolved cryo-electron microscopy: Recent progress. *Journal of Structural Biology*. https://doi.org/10.1016/j.jsb.2017.06.005

Freire, E., Murphy, K. P., Galisteo, M. L., Sanchez-Ruiz, J. M., & Privalov, P. L. (1992). The Molecular Basis of Cooperativity in Protein Folding. Thermodynamic Dissection of Interdomain Interactions in Phosphoglycerate Kinase. *Biochemistry*. https://doi.org/10.1021/bi00116a034

Frottin, F., Schueder, F., Tiwary, S., Gupta, R., Körner, R., Schlichthaerle, T., … Hipp, M. S. (2019). The nucleolus functions as a phase-separated protein quality control compartment. *Science*, *365*(6451), 342–347. https://doi.org/10.1126/science.aaw9157

Frydman, J., Nimmesgern, E., Erdjument-Bromage, H., Wall, J. S., Tempst, P., & Hartl, F. U. (1992). Function in protein folding of TRiC, a cytosolic ring complex containing TCP-1 and structurally related subunits. *The EMBO Journal*, *11*(13), 4767–4778. https://doi.org/10.1002/j.1460-2075.1992.tb05582.x

Fujiwara, K., Ishihama, Y., Nakahigashi, K., Soga, T., & Taguchi, H. (2010). A systematic survey of in vivo obligate chaperonin-dependent substrates. *EMBO Journal*. https://doi.org/10.1038/emboj.2010.52

Gallagher, P. S., Oeser, M. L., Abraham, A. C., Kaganovich, D., & Gardner, R. G. (2014). Cellular maintenance of nuclear protein homeostasis. *Cellular and Molecular Life Sciences*. https://doi.org/10.1007/s00018-013-1530-y

Gambin, Y., & Deniz, A. A. (2010). Multicolor single-molecule FRET to explore protein folding and binding. *Molecular BioSystems*. https://doi.org/10.1039/c003024d

Gambin, Y., Vandelinder, V., Ferreon, A. C. M., Lemke, E. A., Groisman, A., & Deniz, A. A. (2011). Visualizing a one-way protein encounter complex by ultrafast single-molecule mixing. *Nature Methods*. https://doi.org/10.1038/nmeth.1568

Gansen, A., Valeri, A., Hauger, F., Felekyan, S., Kalinin, S., Tóth, K., … Seidel, C. A. M. (2009). Nucleosome disassembly intermediates characterized by single-molecule FRET. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(36), 15308–15313. https://doi.org/10.1073/pnas.0903005106

Georgescauld, F., Popova, K., Gupta, A. J., Bracher, A., Engen, J. R., Hayer-Hartl, M., & Hartl, F. U. (2014). GroEL/ES chaperonin modulates the mechanism and accelerates the rate of TIM-barrel domain folding. *Cell*. https://doi.org/10.1016/j.cell.2014.03.038

Georgopoulos, C. P., & Hohnt, B. (1978). *Identification of a host protein necessary for bacteriophage morphogenesis (the groE gene product) (transducing bacteriophage A/sodium dodecyl sulfate gel electrophoresis/in vitro recombination)*. *Biochemistry* (Vol. 75).

Goemans, C., Denoncin, K., & Collet, J. F. (2014). Folding mechanisms of periplasmic proteins. *Biochimica et Biophysica Acta - Molecular Cell Research*. https://doi.org/10.1016/j.bbamcr.2013.10.014

Goloubinoff, P., Gatenby, A. A., & Lorimer, G. H. (1989). GroE heat-shock proteins promote assembly of foreign prokaryotic ribulose bisphosphate carboxylase oligomers in Escherichia coli. *Nature*. https://doi.org/10.1038/337044a0

Gragerov, A., Zeng, L., Zhao, X., Burkholder, W., & Gottesman, M. E. (1994). Specificity of DnaK-peptide binding. *Journal of Molecular Biology*. https://doi.org/10.1006/jmbi.1994.1043

Gregor, I., Chizhik, A., Karedla, N., & Enderlein, J. (2019). Metal-induced energy transfer. *Nanophotonics*. https://doi.org/10.1515/nanoph-2019-0201

Grinvald, A., Haas, E., & Steinberg, I. Z. (1972). *Evaluation of the Distribution of Distances Between Energy Donors and Acceptors by Fluorescence Decay (energy transfer/fluorescence/decay/conformation/polymers)* (Vol. 69).

Gupta, A. J., Haldar, S., Miličić, G., Hartl, F. U., & Hayer-Hartl, M. (2014). Active cage mechanism of chaperonin-assisted protein folding demonstrated at single-molecule level. *Journal of Molecular Biology*, *426*(15), 2739–2754. https://doi.org/10.1016/j.jmb.2014.04.018

Ha, T., Enderle, T., Ogletree, D. F., Chemla, D. S., Selvin, P. R., & Weiss, S. (1996). Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor. *Proceedings of the National Academy of Sciences*, *93*(13), 6264–6268. https://doi.org/10.1073/pnas.93.13.6264

Haas, I. G., & Wabl, M. (1983). Immunoglobulin heavy chain binding protein. *Nature*. https://doi.org/10.1038/306387a0

Hagn, F., Lagleder, S., Retzlaff, M., Rohrberg, J., Demmer, O., Richter, K., … Kessler, H. (2011). Structural analysis of the interaction between Hsp90 and the tumor suppressor protein p53. https://doi.org/10.1038/nsmb.2114

Haldar, S., Gupta, A. J., Yan, X., Miličić, G., Hartl, F. U., & Hayer-Hartl, M. (2015). Chaperonin-assisted protein folding: Relative population of asymmetric and symmetric GroEL:GroES complexes. *Journal of Molecular Biology*, *427*(12), 2244–2255. https://doi.org/10.1016/j.jmb.2015.04.009

Han, J.-H., Batey, S., Nickson, A. a, Teichmann, S. a, & Clarke, J. (2007). The folding and evolution of multidomain proteins. *Nature Reviews. Molecular Cell Biology*, *8*(4), 319–330. https://doi.org/10.1038/nrm2144

Harris, S. F., Shiau, A. K., & Agard, D. A. (2004). The crystal structure of the carboxy-terminal dimerization domain of htpG, the Escherichia coli Hsp90, reveals a potential substrate binding site. *Structure*. https://doi.org/10.1016/j.str.2004.03.020

Harrison, C. J., Hayer-Hartl, M., Di Liberto, M., Hartl, F. U., & Kuriyan, J. (1997). Crystal structure of the nucleotide exchange factor GrpE bound to the ATPase domain of the molecular chaperone DnaK. *Science*. https://doi.org/10.1126/science.276.5311.431

Hartl, F. U. (1996). *Molecular chaperones in cellular protein folding*. *Nature* (Vol. 381). https://doi.org/10.1038/381571a0

Hartl, F. U., Bracher, A., & Hayer-Hartl, M. (2011). Molecular chaperones in protein folding and proteostasis. *Nature*, *475*(7356), 324–332. https://doi.org/10.1038/nature10317

Haustein, E., & Schwille, P. (2004). Single-molecule spectroscopic methods. *Current Opinion in Structural Biology*. https://doi.org/10.1016/j.sbi.2004.09.004

Hellenkamp, B., Schmid, S., Doroshenko, O., Opanasyuk, O., Kühnemuth, R., Rezaei Adariani, S., … Hugel, T. (2018). Precision and accuracy of single-molecule FRET measurements—a multi-laboratory benchmark study. *Nature Methods*, *15*(9), 669–676. https://doi.org/10.1038/s41592-018-0085-0

Hemmingsen, S. M., Woolford, C., Van Der Vies, S. M., Tilly, K., Dennis, D. T., Georgopoulos, C. P., … Ellis, R. J. (1988). *Homologous plant and bacterial proteins chaperone oligomeric protein assembly*. *Nature* (Vol. 333). https://doi.org/10.1038/333330a0

Hofmann, H., Hillger, F., Pfeil, S. H., Hoffmann, A., Streich, D., Haenni, D., … Schuler, B. (2010). Single-molecule spectroscopy of protein folding in a chaperonin cage. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(26), 11793–11798. https://doi.org/10.1073/pnas.1002356107

Hohng, S., Joo, C., & Ha, T. (2004). Single-molecule three-color FRET. *Biophysical Journal*, *87*(2), 1328–1337. https://doi.org/10.1529/biophysj.104.043935

Holtkamp, W., Kokic, G., Jäger, M., Mittelstaet, J., Komar, A. A., & Rodnina, M. V. (2015). Cotranslational protein folding on the ribosome monitored in real time. *Science*, *350*(6264), 1104–1107. https://doi.org/10.1126/science.aad0344

Huang, C., Rossi, P., Saio, T., & Kalodimos, C. G. (2016). Structural basis for the antifolding activity of a molecular chaperone. *Nature*, *537*(7619), 202–206. https://doi.org/10.1038/nature18965

Huang, J. R., Warner, L. R., Sanchez, C., Gabel, F., Madl, T., Mackereth, C. D., … Blackledge, M. (2014). Transient electrostatic interactions dominate the conformational equilibrium sampled by multidomain splicing factor U2AF65: A combined NMR and SAXS study. *Journal of the American Chemical Society*, *136*(19), 7068–7076. https://doi.org/10.1021/ja502030n

Huang, P.-S., Boyken, S. E., & Baker, D. (2016). The coming of age of de novo protein design. *Nature*, *537*(7620), 320–327. https://doi.org/10.1038/nature19946

Hwang, J., & Qi, L. (2018). Quality Control in the Endoplasmic Reticulum: Crosstalk between ERAD and UPR pathways. *Trends in Biochemical Sciences*. https://doi.org/10.1016/j.tibs.2018.06.005

Imamoglu, R., Balchin, D., Hayer-Hartl, M., & Hartl, F. U. (2020). Bacterial Hsp70 resolves misfolded states and accelerates productive folding of a multi-domain protein. *Nature Communications*, *11*(1), 365. https://doi.org/10.1038/s41467-019-14245-4

Inanami, T., Terada, T. P., & Sasai, M. (2014). Folding pathway of a multidomain protein depends on its topology of domain connectivity. *Proceedings of the National Academy of Sciences*, *111*(45), 15969–15974. https://doi.org/10.1073/pnas.1406244111

Ingolia, T. D., Slater, M. R., & Craig, E. A. (1982). *Saccharomyces cerevisiae Contains a Complex Multigene Family Related to the Major Heat Shock-Inducible Gene of Drosophila*. *MOLECULAR AND CELLULAR BIOLOGY* (Vol. 2). Retrieved from http://mcb.asm.org/

Jäger, M., Michalet, X., & Weiss, S. (2005). Protein-protein interactions as a tool for site-specific labeling of proteins. *Protein Science*, *14*(8), 2059–2068. https://doi.org/10.1110/ps.051384705

Jahn, T. R., & Radford, S. E. (2005). The Yin and Yang of protein folding. *FEBS Journal*, *272*(23), 5962–5970. https://doi.org/10.1111/j.1742-4658.2005.05021.x

Jiang, Y., Rossi, P., & Kalodimos, C. G. (2019). Structural basis for client recognition and activity of Hsp40 chaperones. *Science*, *365*(6459), 1313–1319. https://doi.org/10.1126/science.aax1280

Joerger, A. C., & Fersht, A. R. (2007). Structure-function-rescue: the diverse nature of common p53 cancer mutants. *Oncogene*, *26*, 2226–2242. https://doi.org/10.1038/sj.onc.1210291

Johnson, B. D., Schumacher, R. J., Ross, E. D., & Toft, D. O. (1998). Hop modulates hsp70/hsp90 interactions in protein folding. *Journal of Biological Chemistry*. https://doi.org/10.1074/jbc.273.6.3679

Johnson, J. L. (2012). Evolution and function of diverse Hsp90 homologs and cochaperone proteins. *Biochimica et Biophysica Acta - Molecular Cell Research*. https://doi.org/10.1016/j.bbamcr.2011.09.020

Jones, S., Stewart, M., Michie, A., Swindells, M. B., Orengo, C., & Thornton, J. M. (1998). Domain assignment for protein structures using a consensus approach : Characterization and analysis.

Joo, C., Balci, H., Ishitsuka, Y., Buranachai, C., & Ha, T. (2008). Advances in single-molecule fluorescence methods for molecular biology. *Annual Review of Biochemistry*. https://doi.org/10.1146/annurev.biochem.77.070606.101543

Kaiser, C. M., Chang, H. C., Agashe, V. R., Lakshmipathy, S. K., Etchells, S. A., Hayer-Hartl, M., … Barral, J. M. (2006). Real-time observation of trigger factor function on translating ribosomes. *Nature*. https://doi.org/10.1038/nature05225

Kaiser, C. M., Goldman, D. H., & Chodera, J. D. (2011). Nascent Protein Folding. *Science (New York, N.Y.)*, *334*(December), 1723–1727.

Kaiser, C. M., Goldman, D. H., Chodera, J. D., Tinoco, I., & Bustamante, C. (2011). The ribosome modulates nascent protein folding. *Science*. https://doi.org/10.1126/science.1209740

Kalinin, S., Peulen, T., Sindbert, S., Rothwell, P. J., Berger, S., Restle, T., … Seidel, C. a M. (2012). A toolkit and benchmark study for FRET-restrained high-precision structural modeling. *Nature Methods*, *9*(12), 1218–1227. https://doi.org/10.1038/NMETH.2222

Kalinin, S., Sisamakis, E., Magennis, S. W., Felekyan, S., & Seidel, C. A. M. (2010). On the origin of broadening of single-molecule FRET efficiency distributions beyond shot noise limits. *Journal of Physical Chemistry B*, *114*(18), 6197–6206. https://doi.org/10.1021/jp100025v

Kalinin, S., Valeri, A., Antonik, M., Felekyan, S., & Seidel, C. A. M. (2010). Detection of Structural Dynamics by FRET: A Photon Distribution and Fluorescence Lifetime Analysis of Systems with Multiple States. *Journal of Physical Chemistry B*, *114*(23), 7983–7995. https://doi.org/10.1021/jp102156t

Kampinga, H. H., & Craig, E. A. (2010). The HSP70 chaperone machinery: J proteins as drivers of functional specificity. *Nature Reviews Molecular Cell Biology*. https://doi.org/10.1038/nrm2941

Kapanidis, A. N., Lee, N. K., Laurence, T. A., Doose, S., Margeat, E., & Weiss, S. (2004). Fluorescence-aided molecule sorting: Analysis of structure and interactions by alternating-laser excitation of single molecules. *Proceedings of the National Academy of Sciences*, *101*(24), 8936–8941. https://doi.org/10.1073/pnas.0401690101

Kauzmann, W. (1959). Some Factors in the Interpretation of Protein Denaturation. *Advances in Protein Chemistry*. https://doi.org/10.1016/S0065-3233(08)60608-7

Kelley, W. L. (1998). The J-domain family and the recruitment of chaperone power. *Trends in Biochemical Sciences*. https://doi.org/10.1016/S0968-0004(98)01215-8

Kellner, R., Hofmann, H., Barducci, A., Wunderlich, B., Nettels, D., & Schuler, B. (2014). Single-molecule spectroscopy reveals chaperone-mediated expansion of substrate protein. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.1407086111

Kendrew, J. C., Bodo, G., Dintzis, H. M., Parrish, R. G., Wyckoff, H., & Phillips, D. C. (1958). A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature*. https://doi.org/10.1038/181662a0

Kerner, M. J., Naylor, D. J., Ishihama, Y., Maier, T., Chang, H. C., Stines, A. P., … Hartl, F. U. (2005). Proteome-wide analysis of chaperonin-dependent protein folding in Escherichia coli. *Cell*. https://doi.org/10.1016/j.cell.2005.05.028

Kim, E., Lee, S., Jeon, A., Choi, J. M., Lee, H.-S., Hohng, S., & Kim, H.-S. (2013). A single-molecule dissection of ligand binding to a protein with intrinsic dynamics. *Nature Chemical Biology*, *9*(5), 313–318. https://doi.org/10.1038/nchembio.1213

Kim, N., Kwon, J., Lim, Y., Kang, J., Bae, S., & Kim, S. K. (2018). Incorporation of STED technique into single-molecule spectroscopy to break the concentration limit of diffusing molecules in single-molecule detection. *Chemical Communications*. https://doi.org/10.1039/c8cc05726e

Kim, S. A., Heinze, K. G., & Schwille, P. (2007). Fluorescence correlation spectroscopy in living cells. *NATURE METHODS* |, *4*(11), 963. https://doi.org/10.1038/NMETH1104

Kim, Y. E., Hipp, M. S., Bracher, A., Hayer-Hartl, M., & Hartl, F. U. (2013). *Molecular chaperone functions in protein folding and proteostasis. Annual review of biochemistry* (Vol. 82). https://doi.org/10.1146/annurev-biochem-060208-092442

Kirschke, E., Goswami, D., Southworth, D., Griffin, P. R., & Agard, D. A. (2014). Glucocorticoid receptor function regulated by coordinated action of the Hsp90 and Hsp70 chaperone cycles. *Cell*, *157*(7), 1685–1697. https://doi.org/10.1016/j.cell.2014.04.038

Kityk, R., Kopp, J., & Mayer, M. P. (2018). Molecular Mechanism of J-Domain-Triggered ATP Hydrolysis by Hsp70 Chaperones. *Molecular Cell*, *69*(2), 227-237.e4. https://doi.org/10.1016/j.molcel.2017.12.003

Kityk, R., Kopp, J., Sinning, I., & Mayer, M. P. (2012). Structure and Dynamics of the ATP-Bound Open Conformation of Hsp70 Chaperones. *Molecular Cell*. https://doi.org/10.1016/j.molcel.2012.09.023

Kityk, R., Vogel, M., Schlecht, R., Bukau, B., & Mayer, M. P. (2015). Pathways of allosteric regulation in Hsp70 chaperones. *Nature Communications*, *6*, 1–11. https://doi.org/10.1038/ncomms9308

Kolb, H. C., Finn, M. G., & Sharpless, K. B. (2001). Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angewandte Chemie - International Edition*. https://doi.org/10.1002/1521-3773(20010601)40:11<2004::AID-ANIE2004>3.0.CO;2-5

Konermann, L., Pan, J., & Liu, Y. H. (2011). Hydrogen exchange mass spectrometry for studying protein structure and dynamics. *Chemical Society Reviews*. https://doi.org/10.1039/c0cs00113a

König, I., Zarrine-Afsar, A., Aznauryan, M., Soranno, A., Wunderlich, B., Dingfelder, F., … Schuler, B. (2015). Single-molecule spectroscopy of protein conformational dynamics in live eukaryotic cells. *Nature Methods*. https://doi.org/10.1038/nmeth.3475

Kopperger, E., List, J., Madhira, S., Rothfischer, F., Lamb, D. C., & Simmel, F. C. (2018). A

self-assembled nanoscale robotic arm controlled by electric fields. *Science*. https://doi.org/10.1126/science.aao4284

Kovermann, M., Rogne, P., & Wolf-Watz, M. (2016). Protein dynamics and function from solution state NMR spectroscopy. *Quarterly Reviews of Biophysics*. https://doi.org/10.1017/S0033583516000019

Kramer, G., Boehringer, D., Ban, N., & Bukau, B. (2009). The ribosome as a platform for co-translational processing, folding and targeting of newly synthesized proteins. *Nature Structural and Molecular Biology*. https://doi.org/10.1038/nsmb.1614

Kramers, H. A. (1940). *Brownian motion in a field of force and the diffusion model of chemical reactions*. *Physica* (Vol. 7). https://doi.org/10.1016/S0031-8914(40)90098-2

Kudryavtsev, V., Sikor, M., Kalinin, S., Mokranjac, D., Seidel, C. a M., & Lamb, D. C. (2012). Combining MFD and PIE for accurate single-pair f??rster resonance energy transfer measurements. *ChemPhysChem*, *13*(4), 1060–1078. https://doi.org/10.1002/cphc.201100822

Kuhlman, B., Dantas, G., Ireton, G. C., Varani, G., Stoddard, B. L., & Baker, D. (2003). Design of a Novel Globular Protein Fold with Atomic-Level Accuracy. *Science*. https://doi.org/10.1126/science.1089427

Labbadia, J., & Morimoto, R. I. (2015). The biology of proteostasis in aging and disease. *Annual Review of Biochemistry*. https://doi.org/10.1146/annurev-biochem-060614-033955

Laskey, R. A., Honda, B. M., Mills, A. D., & Finch, J. T. (1978). Nucleosomes are assembled by an acidic protein which binds histones and transfers them to DNA. *Nature*. https://doi.org/10.1038/275416a0

Lazaridis, T., & Karplus, M. (1997). "New view" of protein folding reconciled with the old through multiple unfolding simulations. *Science*, *278*(5345), 1928–1931. https://doi.org/10.1126/science.278.5345.1928

Lee, N. K., Kapanidis, A. N., Wang, Y., Michalet, X., Mukhopadhyay, J., Ebright, R. H., & Weiss, S. (2005). Accurate FRET measurements within single diffusing biomolecules using alternating-laser excitation. *Biophysical Journal*, *88*(4), 2939–2953. https://doi.org/10.1529/biophysj.104.054114

Lee, T. C., Moran, C. R., Cistrone, P. A., Dawson, P. E., & Deniz, A. A. (2018). Site-Specific Three-Color Labeling of α-Synuclein via Conjugation to Uniquely Reactive Cysteines during Assembly by Native Chemical Ligation. *Cell Chemical Biology*, 1–5. https://doi.org/10.1016/j.chembiol.2018.03.009

Lemke, E. A., Gambin, Y., Vandelinder, V., Brustad, E. M., Liu, H. W., Schultz, P. G., … Deniz, A. A. (2009). Microfluidic device for single-molecule experiments with enhanced photostability. *Journal of the American Chemical Society*, *131*(38), 13610–13612. https://doi.org/10.1021/ja9027023

Levinthal, C. (1969). How to fold graciously? Retrieved October 8, 2020, from https://www.cc.gatech.edu/~turk/bio_sim/articles/proteins_levinthal_1969.pdf

Li, W., Terakawa, T., Wang, W., & Takada, S. (2012). Energy landscape and multiroute folding of topologically complex proteins adenylate kinase and 2ouf-knot. *Proceedings of the National Academy of Sciences*, *109*(44), 17789–17794. https://doi.org/10.1073/pnas.1201807109

Liberek, K., Galitski, T. P., Zylicz, M., & Georgopoulos, C. (1992). The DnaK chaperone modulates the heat shock response of Escherichia coli by binding to the sigma 32 transcription factor. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(8), 3516–3520. https://doi.org/10.1073/pnas.89.8.3516

Lill, R., & Neupert, W. (1996). Mechanisms of protein import across the mitochondrial outer membrane. *Trends in Cell Biology*. https://doi.org/10.1016/0962-8924(96)81015-4

Lin, Z., Madan, D., & Rye, H. S. (2008). GroEL stimulates protein folding through forced unfolding. *Nature Structural and Molecular Biology*, *15*(3), 303–311. https://doi.org/10.1038/nsmb.1394

Liu, K., Chen, X., & Kaiser, C. M. (2019). Energetic dependencies dictate folding mechanism in a complex protein. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(51), 25641–25648. https://doi.org/10.1073/pnas.1914366116

Liu, K., Maciuba, K., & Kaiser, C. M. (2019). The Ribosome Cooperates with a Chaperone to Guide Multi-domain Protein Folding. *Molecular Cell*, *74*(2), 310-319.e7. https://doi.org/10.1016/j.molcel.2019.01.043

MacKereth, C. D., Madl, T., Bonnal, S., Simon, B., Zanier, K., Gasch, A., … Sattler, M. (2011). Multi-domain conformational selection underlies pre-mRNA splicing regulation by U2AF. *Nature*, *475*(7356), 408–413. https://doi.org/10.1038/nature10171

Magde, D., Elson, E., & Webb, W. W. (1972). Thermodynamic fluctuations in a reacting system measurement by fluorescence correlation spectroscopy. *Physical Review Letters*, *29*(11), 705–708. https://doi.org/10.1103/PhysRevLett.29.705

Mao, H., Hart, S. A., Schink, A., & Pollok, B. A. (2004). Sortase-Mediated Protein Ligation: A New Method for Protein Engineering. *Journal of the American Chemical Society*. https://doi.org/10.1021/ja039915e

Mapa, K., Sikor, M., Kudryavtsev, V., Waegemann, K., Kalinin, S., Seidel, C. A. M., … Mokranjac, D. (2010). The Conformational Dynamics of the Mitochondrial Hsp70 Chaperone. *Molecular Cell*, *38*(1), 89–100. https://doi.org/10.1016/j.molcel.2010.03.010

Mapa, K., Tiwari, S., Kumar, V., Jayaraj, G. G., & Maiti, S. (2012a). Information encoded in Non-native states drives substrate-chaperone pairing. *Structure*, *20*(9), 1562–1573. https://doi.org/10.1016/j.str.2012.06.014

Mapa, K., Tiwari, S., Kumar, V., Jayaraj, G. G., & Maiti, S. (2012b). Information encoded in Non-native states drives substrate-chaperone pairing. *Structure*, *20*(9), 1562–1573. https://doi.org/10.1016/j.str.2012.06.014

Marcinowski, M., Höller, M., Feige, M. J., Baerend, D., Lamb, D. C., & Buchner, J. (2011). Substrate discrimination of the chaperone BiP by autonomous and cochaperone-regulated conformational transitions. *Nature Publishing Group*. https://doi.org/10.1038/nsmb.1970

Marvin Minsky. (1957). Microscopy Apparatus. US Patent US3013467A. Filed Nov. 7, 1957, granted Dec. 19, 1961. *Patent*. Retrieved from https://patents.google.com/patent/US3013467A/en

Mashaghi, A., Bezrukavnikov, S., Minde, D. P., Wentink, A. S., Kityk, R., Zachmann-Brand, B., … Tans, S. J. (2016). Alternative modes of client binding enable functional plasticity of Hsp70. *Nature*, 1–16. https://doi.org/10.1038/nature20137

Mashaghi, A., Kramer, G., Bechtluft, P., Zachmann-Brand, B., Driessen, A. J. M., Bukau, B., & Tans, S. J. (2013). Reshaping of the conformational search of a protein by the chaperone

trigger factor. *Nature*, *500*(7460), 98–101. https://doi.org/10.1038/nature12293

Mayer, M. P. (2010). Gymnastics of molecular chaperones. *Molecular Cell*, *39*(3), 321–331. https://doi.org/10.1016/j.molcel.2010.07.012

McCarty, J. S., Buchberger, A., Reinstein, J., & Bukau, B. (1995). The Role of ATP in the Functional Cycle of the DnaK Chaperone System. *Journal of Molecular Biology*, *249*(1), 126–137. https://doi.org/10.1006/jmbi.1995.0284

Mehta, A. D., Rief, M., Spudich, J. A., Smith, D. A., & Simmons, R. M. (1999). Single-molecule biomechanics with optical methods. *Science*. https://doi.org/10.1126/science.283.5408.1689

Meyer, P., Prodromou, C., Hu, B., Vaughan, C., Roe, S. M., Panaretou, B., … Pearl, L. H. (2003). Structural and functional analysis of the middle segment of Hsp90: Implications for ATP hydrolysis and client protein and cochaperone interactions. *Molecular Cell*. https://doi.org/10.1016/S1097-2765(03)00065-0

Minsky, M. (1957). Microscopy Apparatus. *US Patent 3013467*.

Moerner, W. E., & Kador ', L. (1989). *Optical Detection and Spectroscopy of Single Molecules in a Solid* (Vol. 62).

Moon, C. P., Kwon, S., & Fleming, K. G. (2011). Overcoming hysteresis to attain reversible equilibrium folding for outer membrane phospholipase A in phospholipid bilayers. *Journal of Molecular Biology*, *413*(2), 484–494. https://doi.org/10.1016/j.jmb.2011.08.041

Morimoto, R. I. (2008). Proteotoxic stress and inducible chaperone networks in neurodegenerative disease and aging. *Genes and Development*, *22*(11), 1427–1438. https://doi.org/10.1101/gad.1657108

Müller, B. K., Zaychikov, E., Bräuchle, C., & Lamb, D. C. (2005). Pulsed interleaved excitation. *Biophysical Journal*, *89*(5), 3508–3522. https://doi.org/10.1529/biophysj.105.064766

Munro, S., & Pelham, H. R. B. (1986). An hsp70-like protein in the ER: Identity with the 78 kd glucose-regulated protein and immunoglobulin heavy chain binding protein. *Cell*. https://doi.org/10.1016/0092-8674(86)90746-4

Nam, K. L., Kapanidis, A. N., Hye, R. K., Korlann, Y., Sam, O. H., Kim, Y., … Weiss, S. (2007). Three-color alternating-laser excitation of single molecules: Monitoring multiple interactions and distances. *Biophysical Journal*, *92*(1), 303–312. https://doi.org/10.1529/biophysj.106.093211

Nesmelov, Y. E., & Thomas, D. D. (2010). Protein structural dynamics revealed by site-directed spin labeling and multifrequency EPR. *Biophysical Reviews*. https://doi.org/10.1007/s12551-010-0032-5

Nettels, D., & Schuler, B. (2007). Subpopulation-resolved photon statistics of single-molecule energy transfer dynamics. *IEEE Journal on Selected Topics in Quantum Electronics*. https://doi.org/10.1109/JSTQE.2007.902848

Neuweiler, H., & Sauer, M. (2004). Studying conformational fluctuations in single biomolecules using electron transfer reactions. In *Biophotonics Micro- and Nano-Imaging*. https://doi.org/10.1117/12.552213

Nillegoda, N. B., Wentink, A. S., & Bukau, B. (2018). Protein Disaggregation in Multicellular Organisms. *Trends in Biochemical Sciences*. https://doi.org/10.1016/j.tibs.2018.02.003

Nir, E., Michalet, X., Hamadani, K. M., Laurence, T. A., Neuhauser, D., Kovchegov, Y., & Weiss, S. (2006a). Shot-noise limited single-molecule FRET histograms: Comparison between theory and experiments. *Journal of Physical Chemistry B*, *110*(44), 22103–22124. https://doi.org/10.1021/jp063483n

Nir, E., Michalet, X., Hamadani, K. M., Laurence, T. A., Neuhauser, D., Kovchegov, Y., & Weiss, S. (2006b). Shot-Noise Limited Single-Molecule FRET Histograms: Comparison between Theory and Experiments †. https://doi.org/10.1021/jp063483n

Niwa, T., Fujiwara, K., & Taguchi, H. (2016). Identification of novel in vivo obligate GroEL/ES substrates based on data from a cell-free proteomics approach. *FEBS Letters*. https://doi.org/10.1002/1873-3468.12036

Oh, E., Becker, A. H., Sandikci, A., Huber, D., Chaba, R., Gloge, F., … Bukau, B. (2011). Selective ribosome profiling reveals the cotranslational chaperone action of trigger factor in vivo. *Cell*, *147*(6), 1295–1308. https://doi.org/10.1016/j.cell.2011.10.044

Onuchic, J. N., & Wolynes, P. G. (2004). Theory of protein folding. *Current Opinion in Structural Biology*. https://doi.org/10.1016/j.sbi.2004.01.009

Ostermann, J., Horwich, A. L., Neupert, W., & Hartl, F.-U. (1989). *Protein folding in mitochondria requires complex formation with hsp60 and A TP hydrolysis*.

Osváth, S., Köhler, G., Závodszky, P., & Fidy, J. (2005). Asymmetric effect of domain interactions on the kinetics of folding in yeast phosphoglycerate kinase. *Protein Science : A Publication of the Protein Society*, *14*(6), 1609–1616. https://doi.org/10.1110/ps.051359905

Pelham, H. R. B. (1986). *Speculations on the functions of the major heat shock and glucose-regulated proteins*. *Cell* (Vol. 46). https://doi.org/10.1016/0092-8674(86)90693-8

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, *25*(13), 1605–1612. https://doi.org/10.1002/jcc.20084

Pfeil, S. H., Wickersham, C. E., Hoffmann, A., & Lipman, E. A. (2009). A microfluidic mixing system for single-molecule measurements. *Review of Scientific Instruments*. https://doi.org/10.1063/1.3125643

Pirchi, M., Ziv, G., Riven, I., Cohen, S. S., Zohar, N., Barak, Y., & Haran, G. (2011). Single-molecule fluorescence spectroscopy maps the folding landscape of a large protein. *Nature Communications*, *2*, 493. https://doi.org/10.1038/ncomms1504

Preissler, S., & Deuerling, E. (2012). Ribosome-associated chaperones as key players in proteostasis. *Trends in Biochemical Sciences*. https://doi.org/10.1016/j.tibs.2012.03.002

Prodromou, C., Roe, S. M., O'Brien, R., Ladbury, J. E., Piper, P. W., & Pearl, L. H. (1997). Identification and structural characterization of the ATP/ADP-binding site in the Hsp90 molecular chaperone. *Cell*. https://doi.org/10.1016/S0092-8674(00)80314-1

Punj, D., Regmi, R., Devilez, A., Plauchu, R., Moparthi, S. B., Stout, B., … Wenger, J. (2015). Self-Assembled Nanoparticle Dimer Antennas for Plasmonic-Enhanced Single-Molecule Fluorescence Detection at Micromolar Concentrations. *ACS Photonics*. https://doi.org/10.1021/acsphotonics.5b00152

Quan, S., Koldewey, P., Tapley, T., Kirsch, N., Ruane, K. M., Pfizenmaier, J., … Bardwell, J. C. A. (2011). Genetic selection designed to stabilize proteins uncovers a chaperone called

Spy. *Nature Structural and Molecular Biology*. https://doi.org/10.1038/nsmb.2016

Raffy, S., Sassoon, N., Hofnung, M., & Betton, J. M. (1998). Tertiary structure-dependence of misfolding substitutions in loops of the maltose-binding protein. *Protein Science : A Publication of the Protein Society*, *7*(10), 2136–2142. https://doi.org/10.1002/pro.5560071010

Ramachandran, G. N., Ramakrishnan, C., & Sasisekharan, V. (1963). Stereochemistry of polypeptide chain configurations. *Journal of Molecular Biology*. https://doi.org/10.1016/S0022-2836(63)80023-6

Rapoport, T. A., Li, L., & Park, E. (2017). Structural and mechanistic insights into protein translocation. *Annual Review of Cell and Developmental Biology*. https://doi.org/10.1146/annurev-cellbio-100616-060439

Ratzke, C., Hellenkamp, B., & Hugel, T. (2014). Four-colour FRET reveals directionality in the Hsp90 multicomponent machinery. *Nature Communications*, *5*(1), 4192. https://doi.org/10.1038/ncomms5192

Richter, K., Haslbeck, M., & Buchner, J. (2010). The Heat Shock Response: Life on the Verge of Death. *Molecular Cell*, *40*(2), 253–266. https://doi.org/10.1016/j.molcel.2010.10.006

Roder, H., Elove, G. A., & Englander, S. W. (1988). *Structural characterization of folding intermediates in cytochrome c by ff-exchange labelling and proton NMR*.

Roder, H., Elöve, G. A., & Englander, S. W. (1988). Structural characterization of folding intermediates in cytochrome c by H-exchange labelling and proton NMR. *Nature*, *335*(6192), 700–704. https://doi.org/10.1038/335700a0

Roe, D. R., & Cheatham, T. E. (2013). PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. https://doi.org/10.1021/ct400341p

Röhl, A., Rohrberg, J., & Buchner, J. (2013). The chaperone Hsp90: Changing partners for demanding clients. *Trends in Biochemical Sciences*. https://doi.org/10.1016/j.tibs.2013.02.003

Röhl, A., Wengler, D., Madl, T., Lagleder, S., Tippel, F., Herrmann, M., … Buchner, J. (2015). Hsp90 regulates the dynamics of its cochaperone Sti1 and the transfer of Hsp70 between modules. https://doi.org/10.1038/ncomms7655

Rosam, M., Krader, D., Nickels, C., Hochmair, J., Back, K. C., Agam, G., … Buchner, J. (2018). Bap (Sil1) regulates the molecular chaperone BiP by coupling release of nucleotide and substrate. *Nature Structural and Molecular Biology*, *25*(1), 90–100. https://doi.org/10.1038/s41594-017-0012-6

Rosenzweig, R., Nillegoda, N. B., Mayer, M. P., & Bukau, B. (2019). The Hsp70 chaperone network. *Nature Reviews Molecular Cell Biology*, *20*(11), 665–680. https://doi.org/10.1038/s41580-019-0133-3

Rüdiger, S., Germeroth, L., Schneider-Mergener, J., & Bukau, B. (1997). Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries. *EMBO Journal*. https://doi.org/10.1093/emboj/16.7.1501

Rutherford, S. L., & Lindquist, S. (1998). Hsp90 as a capacitor for morphological evolution. *Nature*. https://doi.org/10.1038/24550

Saibil, H. R., Fenton, W. A., Clare, D. K., & Horwich, A. L. (2013). Structure and allostery of the chaperonin GroEL. *Journal of Molecular Biology*. https://doi.org/10.1016/j.jmb.2012.11.028

Sanabria, H., Rodnin, D., Hemmen, K., Peulen, T. O., Felekyan, S., Fleissner, M. R., … Seidel, C. A. M. (2020). Resolving dynamics and function of transient states in single enzyme molecules. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-020-14886-w

Sánchez-Rico, C., Voith von Voithenberg, L., Warner, L., Lamb, D. C., & Sattler, M. (2017). Effects of Fluorophore Attachment on Protein Conformation and Dynamics Studied by spFRET and NMR Spectroscopy. *Chemistry - A European Journal*, *23*(57), 14267–14277. https://doi.org/10.1002/chem.201702423

Schaffer, J., Volkmer, A., Eggeling, C., Subramaniam, V., Striker, G., & Seidel, C. A. M. (1999). Identification of Single Molecules in Aqueous Solution by Time-Resolved Fluorescence Anisotropy. *Journal of Physical Chemistry A*, *103*(3). https://doi.org/10.1021/jp9833597

Scheibel, T., Neuhofen, S., Weikl, T., Mayr, C., Reinstein, J., Vogel, P. D., & Buchner, J. (1997). ATP-binding properties of human Hsp90. *Journal of Biological Chemistry*. https://doi.org/10.1074/jbc.272.30.18608

Scheres, S. H. W. (2013). Single-particle processing in RELION. *Manuals*.

Schopf, F. H., Biebl, M. M., & Buchner, J. (2017a). The HSP90 chaperone machinery. *Nature Reviews Molecular Cell Biology*, *18*(6), 345–360. https://doi.org/10.1038/nrm.2017.20

Schopf, F. H., Biebl, M. M., & Buchner, J. (2017b). The HSP90 chaperone machinery. *Nature Reviews Molecular Cell Biology*, *18*(6), 345–360. https://doi.org/10.1038/nrm.2017.20

Schrimpf, W., Barth, A., Hendrix, J., & Lamb, D. C. (2018). PAM: A Framework for Integrated Analysis of Imaging, Single-Molecule, and Ensemble Fluorescence Data. *Biophysical Journal*, *114*(7), 1518–1528. https://doi.org/10.1016/j.bpj.2018.02.035

Schrimpf, W., Lemmens, V., Smisdom, N., Ameloot, M., Lamb, D. C., & Hendrix, J. (2018). Crosstalk-free multicolor RICS using spectral weighting. *Methods*, *140–141*, 97–111. https://doi.org/10.1016/j.ymeth.2018.01.022

Schuler, B., Lipman, E. A., & Eaton, W. A. (2002). Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy. *Nature*, *419*(6908), 743–747. https://doi.org/10.1038/nature01060

Schwille, P., Meyer-Almes, F.-J., & Rigler, R. (1997). *Dual-Color Fluorescence Cross-Correlation Spectroscopy for Multicomponent Diffusional Analysis in Solution*. *Biophysical Journal* (Vol. 72). https://doi.org/10.1016/S0006-3495(97)78833-7

Sela, M., White, F. H., & Anfinsen, C. B. (1957). Reductive cleavage of disulfide bridges in ribonuclease. *Science*, *125*(3250), 691–692. https://doi.org/10.1126/science.125.3250.691

Selmke, B., Borbat, P. P., Nickolaus, C., Varadarajan, R., Freed, J. H., & Trommer, W. E. (2018). Open and Closed Form of Maltose Binding Protein in Its Native and Molten Globule State as Studied by EPR Spectroscopy. *Biochemistry*, acs.biochem.8b00322. https://doi.org/10.1021/acs.biochem.8b00322

Shank, E. A., Cecconi, C., Dill, J. W., Marqusee, S., & Bustamante, C. (2010). The folding cooperativity of a protein is controlled by its chain topology. *Nature*, *465*(7298), 637–640. https://doi.org/10.1038/nature09021

Sharff, A. J., Rodseth, L. E., Spurlino, J. C., & 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*(44), 10657–10663. https://doi.org/10.1021/bi00159a003

Sharma, S., Chakraborty, K., Müller, B. K., Astola, N., Tang, Y. C., Lamb, D. C., … Hartl, F. U. (2008). Monitoring Protein Conformation along the Pathway of Chaperonin-Assisted Folding. *Cell*, *133*(1), 142–153. https://doi.org/10.1016/j.cell.2008.01.048

Shera ', E. B., Seitzinger, N. K., Davis, L. M., Keller, R. A., & Las, S. A. S. (1990). *Detection of single fluorescent molecules*. *CHEMICAL PHYSICS LETTERS* (Vol. 174).

Shtilerman, M., Lorimer, G. H., & Englander, S. W. (1999). Chaperonin function: Folding by forced unfolding. *Science*, *284*(5415), 822–825. https://doi.org/10.1126/science.284.5415.822

Soranno, A., Buchli, B., Nettels, D., Cheng, R. R., Müller-Späth, S., Pfeil, S. H., … Schuler, B. (2012). Quantifying internal friction in unfolded and intrinsically disordered proteins with single-molecule spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(44), 17800–17806. https://doi.org/10.1073/pnas.1117368109

Soranno, A., Holla, A., Dingfelder, F., Nettels, D., Makarov, D. E., & Schuler, B. (2017). Integrated view of internal friction in unfolded proteins from single-molecule FRET, contact quenching, theory, and simulations. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(10), E1833–E1839. https://doi.org/10.1073/pnas.1616672114

Sorger, P. K., & Pelham, H. R. B. (1987). The glucose-regulated protein grp94 is related to heat shock protein hsp90. *Journal of Molecular Biology*. https://doi.org/10.1016/0022-2836(87)90380-9

Sparrer, H, Rutkat, K., & Buchner, J. (1997). Catalysis of protein folding by symmetric chaperone complexes. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(4), 1096–1100. https://doi.org/10.1073/pnas.94.4.1096

Sparrer, Helmut, Lilie, H., & Buchner, J. (1996). Dynamics of the GroEL – Protein Complex : Effects of Nucleotides and Folding Mutants, 74–87.

Spurlino, J. C., Lu, G. Y., & 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. *Journal of Biological Chemistry*, *266*(8), 5202–5219. https://doi.org/2002054

Stryer, L., & Haugland, R. P. (1967). Energy transfer: a spectroscopic ruler. *Proceedings of the National Academy of Sciences of the United States of America*, *58*(2), 719–726. https://doi.org/10.1073/pnas.58.2.719

Taipale, M., Krykbaeva, I., Koeva, M., Kayatekin, C., Westover, K. D., Karras, G. I., & Lindquist, S. (2012). Quantitative analysis of Hsp90-client interactions reveals principles of substrate recognition. *Cell*. https://doi.org/10.1016/j.cell.2012.06.047

Taipale, M., Tucker, G., Peng, J., Krykbaeva, I., Lin, Z. Y., Larsen, B., … Lindquist, S. (2014). A quantitative chaperone interaction network reveals the architecture of cellular protein homeostasis pathways. *Cell*. https://doi.org/10.1016/j.cell.2014.05.039

Takashl, F., Harada, Y., Tokunaga, M., Salto, K., & Yanagida, T. (1995). *Imaging of single fluorescent molecules and individual ATP turnovers by single myosin molecules in aqueous solution*. *Nature* (Vol. 374). https://doi.org/10.1038/255242a0

Tanford, C. (1962). Contribution of Hydrophobic Interactions to the Stability of the Globular

Conformation of Proteins. *Journal of the American Chemical Society*. https://doi.org/10.1021/ja00881a009

Tang, Y. C., Chang, H. C., Roeben, A., Wischnewski, D., Wischnewski, N., Kerner, M. J., … Hayer-Hartl, M. (2006a). Structural Features of the GroEL-GroES Nano-Cage Required for Rapid Folding of Encapsulated Protein. *Cell*, *125*(5), 903–914. https://doi.org/10.1016/j.cell.2006.04.027

Tang, Y. C., Chang, H. C., Roeben, A., Wischnewski, D., Wischnewski, N., Kerner, M. J., … Hayer-Hartl, M. (2006b). Structural Features of the GroEL-GroES Nano-Cage Required for Rapid Folding of Encapsulated Protein. *Cell*, *125*(5), 903–914. https://doi.org/10.1016/j.cell.2006.04.027

Taylor, R. C., & Dillin, A. (2011). Aging as an event of proteostasis collapse. *Cold Spring Harbor Perspectives in Biology*. https://doi.org/10.1101/cshperspect.a004440

Teichmann, S. A., Parkt, J., & Chothia, C. (1998). Structural assignments to the Mycoplasma genitalium proteins show extensive gene duplications and domain rearrangements. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(25), 14658–14663. https://doi.org/10.1073/pnas.95.25.14658

Thirumalai, D., O'Brien, E. P., Morrison, G., & Hyeon, C. (2010). Theoretical Perspectives on Protein Folding. https://doi.org/10.1146/annurev-biophys-051309-103835

Tissiéres, A., Mitchell, H. K., & Tracy, U. M. (1974). *Protein synthesis in salivary glands of Drosophila melanogaster: Relation to chromosome puffs*. *Journal of Molecular Biology* (Vol. 84). https://doi.org/10.1016/0022-2836(74)90447-1

Tomov, T. E., Tsukanov, R., Masoud, R., Liber, M., Plavner, N., & Nir, E. (2012). Disentangling subpopulations in single-molecule FRET and ALEX experiments with photon distribution analysis. *Biophysical Journal*, *102*(5), 1163–1173. https://doi.org/10.1016/j.bpj.2011.11.4025

Toofanny, R. D., & Daggett, V. (2012). Understanding protein unfolding from molecular simulations. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, *2*(3), 405–423. https://doi.org/10.1002/wcms.1088

Torella, J. P., Holden, S. J., Santoso, Y., Hohlbein, J., & Kapanidis, A. N. (2011). Identifying molecular dynamics in single-molecule fret experiments with burst variance analysis. *Biophysical Journal*, *100*(6), 1568–1577. https://doi.org/10.1016/j.bpj.2011.01.066

Torres, T., & Levitus, M. (2007). Measuring Conformational Dynamics: A New FCS-FRET Approach. *The Journal of Physical Chemistry B*, *111*(25), 7392–7400. https://doi.org/10.1021/jp070659s

Tyagi, S., & Lemke, E. A. (2013). Genetically Encoded Click Chemistry for Single-Molecule FRET of Proteins. *Methods in Cell Biology*. https://doi.org/10.1016/B978-0-12-407239-8.00009-4

Udgaonkar, J. B., & Baldwin, R. L. (1988). NMR evidence for an early framework intermediate on the folding pathway of ribonuclease A. *Nature*. https://doi.org/10.1038/335694a0

Vainberg, I. E., Lewis, S. A., Rommelaere, H., Ampe, C., Vandekerckhove, J., Klein, H. L., & Cowan, N. J. (1998). Prefoldin, a chaperone that delivers unfolded proteins to cytosolic chaperonin. *Cell*. https://doi.org/10.1016/S0092-8674(00)81446-4

Verba, K. A., Wang, R. Y.-R., Arakawa, A., Liu, Y., Shirouzu, M., Yokoyama, S., & David A. Agard. (2016). Atomic structure ofHsp90-Cdc37-Cdk4 reveals that Hsp90 traps and

stabilizes an unfolded kinase. *Science*, *352*(6293), 1542–1547. https://doi.org/10.1126/science.aaf7419

Viitanen, P. V., Schmidt, M., Buchner, J., Suzuki, T., Vierling, E., Dickson, R., … Soll, J. (1995). Functional characterization of the higher plant chloroplast chaperonins. *Journal of Biological Chemistry*. https://doi.org/10.1074/jbc.270.30.18158

Vogel, C., Bashton, M., Kerrison, N. D., Chothia, C., & Teichmann, S. A. (2004). Structure, function and evolution of multidomain proteins. *Current Opinion in Structural Biology*, *14*(2), 208–216. https://doi.org/10.1016/j.sbi.2004.03.011

Voith von Voithenberg, L., Sánchez-Rico, C., Kang, H. S., Madl, T., Zanier, K., Barth, A., … Lamb, D. C. (2016). Recognition of the 3′ splice site RNA by the U2AF heterodimer involves a dynamic population shift. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(46), E7169–E7175. https://doi.org/10.1073/pnas.1605873113

Walerych, D., Olszewski, M. B., Gutkowska, M., Helwak, A., Zylicz, M., & Zylicz, A. (2009). Hsp70 molecular chaperones are required to support p53 tumor suppressor activity under stress conditions. *Oncogene*, *28*, 4284–4294. https://doi.org/10.1038/onc.2009.281

Wallrabe, H., & Periasamy, A. (2005). Imaging protein molecules using FRET and FLIM microscopy. *Current Opinion in Biotechnology*. https://doi.org/10.1016/j.copbio.2004.12.002

Walser, R., Mark, A. E., & Van Gunstere, W. F. (2000). On the temperature and pressure dependence of a range of properties of a type of water model commonly used in high-temperature protein unfolding simulations. *Biophysical Journal*, *78*(6), 2752–2760. https://doi.org/10.1016/S0006-3495(00)76820-2

Walter, P., & Ron, D. (2011). The unfolded protein response: From stress pathway to homeostatic regulation. *Science*. https://doi.org/10.1126/science.1209038

Walters, B. T., Mayne, L., Hinshaw, J. R., Sosnick, T. R., & Englander, S. W. (2013). Folding of a large protein at high structural resolution. *Proceedings of the National Academy of Sciences*, *110*(47), 18898–18903. https://doi.org/10.1073/pnas.1319482110

Wang, J. D., Michelitsch, M. D., & Weissman, J. S. (1998a). GroEL-GroES-mediated protein folding requires an intact central cavity. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(21), 12163–12168. https://doi.org/10.1073/pnas.95.21.12163

Wang, J. D., Michelitsch, M. D., & Weissman, J. S. (1998b). GroEL-GroES-mediated protein folding requires an intact central cavity. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.95.21.12163

Wang, Y., Chu, X., Suo, Z., Wang, E., & Wang, J. (2012). Multidomain protein solves the folding problem by multifunnel combined landscape: Theoretical investigation of a Y-family DNA polymerase. *Journal of the American Chemical Society*, *134*(33), 13755–13764. https://doi.org/10.1021/ja3045663

Weissman, J. S., Kashi, Y., Fenton, W. A., & Horwich, A. L. (1994). GroEL-mediated protein folding proceeds by multiple rounds of binding and release of nonnative forms. *Cell*. https://doi.org/10.1016/0092-8674(94)90533-9

Wetlaufer, D. B. (1973). Nucleation, Rapid Folding, and Globular Intrachain Regions in Proteins. *Proceedings of the National Academy of Sciences*, *70*(3), 697–701.

https://doi.org/10.1073/pnas.70.3.697

Wolynes, P. G., Onuchic, J. N., & Thirumalai, D. (1995). Navigating the Folding Routes. *Science*, *267*(March).

Wright, C. F., Teichmann, S. A., Clarke, J., & Dobson, C. M. (2005). The importance of sequence diversity in the aggregation and evolution of proteins. *Nature*. https://doi.org/10.1038/nature04195

Xu, Z., Horwich, A. L., & Sigler, P. B. (1997). *The crystal structure of the asymmetric GroEL–GroES–(ADP)7 chaperonin complex*. *Nature* (Vol. 388). Retrieved from www.nature.com

Ye, X., Mayne, L., Kan, Z., & Englander, S. W. (2018). Folding of maltose binding protein outside of and in GroEL. *Proceedings of the National Academy of Sciences*, (33), 201716168. https://doi.org/10.1073/pnas.1716168115

Yifrach, O., & Horovitz, A. (2000). Coupling between protein folding and allostery in the GroE chaperonin system. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.040449997

Yochem, J., Uchida, H., Sunshine, M., Saito, H., Georgopoulos, C. P., & Feiss, M. (1978). Genetic analysis of two genes, dnaJ and dnaK, necessary for Escherichia coli and bacteriophage lambda DNA replication. *MGG Molecular & General Genetics*. https://doi.org/10.1007/BF00267593

Yoo, J., Louis, J. M., Gopich, I. V., & Chung, H. S. (2018). Three-Color Single-Molecule FRET and Fluorescence Lifetime Analysis of Fast Protein Folding. *Journal of Physical Chemistry B*. https://doi.org/10.1021/acs.jpcb.8b07768

Appendix

Plasmids used in the course of this the work. This list includes all the original plasmids provided and cloned in this work that were used for expression purposes.

Plasmid	Donor organism	Mutations	Purpose	Provide by	Antibiotic resistance
pET-22b-DM-MBP-A52C-P298C	*E. coli*	V8G, Y283D, A52C, P298C	Folding double mutant of Maltose binding protein with cysteine mutations as attachment positions for labeling	Prof. F. Ulrich Hartl, Max Planck Institute of Biochemistry.	Ampicillin
pET-22b-DM-MBP-K175C-P298C	*E. coli*	V8G, Y283D, K175C, P298C	Folding double mutant of Maltose binding protein with cysteine mutations as attachment positions for labeling	Prof. F. Ulrich Hartl, Max Planck Institute of Biochemistry	Ampicillin
pET-22b-DM-MBP-P298C-T345C	*E. coli*	V8G, Y283D, P298C, T345C	Folding double mutant of Maltose binding protein with cysteine mutations as attachment positions for labeling	Prof. F. Ulrich Hartl, Max Planck Institute of Biochemistry	Ampicillin
pET-22b-DM-MBP-A134C-P298C	*E. coli*	V8G, Y283D, A134C, P298C	Folding double mutant of Maltose binding protein with cysteine mutations as attachment positions for labeling	Prof. F. Ulrich Hartl, Max Planck Institute of Biochemistry	Ampicillin

pET-22b-DM-MBP-A52C	*E. coli*	V8G, Y283D, A52C	Folding double mutant of Maltose binding protein with cysteine mutation as attachment positions for labeling control and FCCS	Cloned in this work	Ampicillin
pET-22b-DM-MBP-P298C	*E. coli*	V8G, Y283D, P298C	Folding double mutant of Maltose binding protein with cysteine mutation to check for specific labeling	Cloned in this work	Ampicillin
pET-22b-DM-MBP-A52C-K175C	*E. coli*	V8G, Y283D, A52C, K175C	Folding double mutant of Maltose binding protein with cysteine mutations as attachment positions for labeling	Cloned in this work	Ampicillin
pET-22b-wt-MBP-K175C-P298C	*E. coli*	K175C, P298C	Wt-Maltose binding protein with cysteine mutations as attachment positions for labeling	Cloned in this work	Ampicillin
pET-22b-wt-MBP-A52C-P298C	*E. coli*	A52C, P298C	Wt-Maltose binding protein with cysteine mutations as attachment positions for labeling	Cloned in this work	Ampicillin

pET-22b-DM-MBP-A52tag-K175C-P298C	*E. coli*	V8G, Y283D, A52tag, K175C, P298C	Folding double mutant of Maltose binding protein with cysteine mutations as attachment positions for labeling and amber stop codon at 52 position for un-natural amino acid incorporation	Cloned in this work	Ampicillin
pEvol-tRNA Pyl PylRS WT	*M. jannaschii*	-	Pyrrolysine based tRNA synthetase and Propargyl-L-Lysine-tRNA-transferase, an orthogonal translation system	Prof. Edward Lemke, Johannes Gutenberg University, Mainz, Germany	Chloramphenicol

Oligonucleotides used in the work are listed below. DNAs are listed using capital letters and RNAs in small letters. p=phosphate. All the oligonucleotides were PAGE purified quality. Bold letters indicate the desired change in the codon after mutagenesis.

Oligonucleotides	**Sequence 5'->3'**
MBP-52TAG fw	p-A CAG GTT GCG **TAG** ACT GGC GAT G
MBP-52TAG rev	p-GGG AAT TTC TCT TCC AGT TTA TCC GGA T
MBP seq-1 fw	p-GTG ATG TCG GCG ATA TAG GC
MBP seq-1 rev	p-CCT CAA GAC CCG TTT AGA GG
MBP-C298P fw	p-AT AAA GAC AAA **CCG** CTG GGT GCC
MBP-C298P rev	p-TAA CCG CTT CCA GAC CTT CAT CAG T
MBP-K175C fw	p-G TAT GAA AAC GGC **TGT** TAC GAC ATT AAA G

MBP-K175C rev	p-TTG AAC GCA TAA CCC CCG T
MBP-G8V fw	p-GAA GGT AAA CTG **GTG** ATC TGG ATT AAC
MBP- G8V rev	p-TTC GAT TTT CAT ATG TAT ATC TCC TTC TTA AAG
MBP-D283Y-fw	p-TCC TCG AA AAC **TAT** CTG CTG ACT G
MBP- D283Y-rev	p-ACT CTT TCG CCA GCT CTT TGT T
MBP seq-2 fw	p-GAA CTG AAA GCG AAA GGT AAG AGC
MBP seq-2 rev	p-CGA AAA GTG CCA CCT AAA TTG TAA GC
MBP-C175K fw	p-G TAT GAA AAC GGC **AAG** TAC GAC ATT AAA G
MBP-C175K rev	p-TTG AAC GCA TAA CCC CCG T
U9	uuuuuuuuu

Acknowledgements

It is hard to find something which interests for years. I'm fortunate to find something like that in the last few years. Any journey is not possible alone. There are many pillars of this journey too.

To begin with, Don, I have to admit that you have influenced the most my last years. Thank you for entertaining me for years in the lab and allowing me to undertake research work and learn many more things beyond it. Be it your humor, patience, thoroughness in almost everything, fairness, and trying to do best for everyone has always amazed me and had set a good example as a supervisor. I will always be grateful for your efforts to improve my writing, especially during thesis corrections. I cannot thank you enough!

Next, I want to thank Johannes Buchner for being always interested and excited to carry out studies on new chaperone systems. Thanks for all your support, views, and fruitful collaboration as well as for being my second thesis examiner. Further, I want to thank Philip Tinnefeld, who made the combined meeting stimulating and for providing a great atmosphere within intergroup activities and being always supportive. Again, thanks for being part of my thesis committee. I'm indebted to Thorben Cordes for an exciting, enduring, and challenging collaborative work on FRET comparison study on proteins. I look forward to seeing how in the end our work takes shape. And thanks for your support by becoming my thesis committee member. Last but not the least, I want to convey my sincere thanks to F. Ulrich Hartl for long-standing discussions on MBP folding studies and chaperone mechanisms. Thank you for providing always a supply for GroEL/ES proteins and also for becoming a part of my thesis committee. I should not forget to thank the support from SFB1035 to have continued funding to withstand these years as well as Prof. Dr. Christoph Braüchle for supporting us on the first floor of Building E in many ways. To help me to get started at Munich and for introducing me to the field of protein folding, I will always be grateful to Dr. Kausik Chakraborty and Dr. Koyeli Mapa.

I was fortunate to interact, work, and get support from almost four generations of AK Lamb lab colleagues. From the previous generation people, special thanks to Lena for providing a smooth start for biochemistry as well as for being a good friend. Waldi, Sushi, Jens, Viola, Fabian Wehnekamp, and Alvaro made the lab and overall environment friendly. Thanks, Maria and Fabian Wehnekamp for sharing an office for the initial years. I remember we had a quite wide spectrum of cold-heat sensitivities in our office. It was nice to collaborate with you Daniela for microfluidic measurements. It was delightful to have complaint sessions of Waldi for a change, Sushi for sharing Indian aura, and Jens for your mischievous smile. Ellen, though from the neighboring lab it always felt like you were integral part of our lab, thanks for making kitchen breaks fun and for being a good friend.

There were some moments of relief to share struggles that we went through together or alone at times with my generation colleagues including Anders, Bässem, Chen, Evelyn, Ivo, Kira, Maria, Nader, and Philipp. It was a great time spent at lunchtimes on campus, several dinners, celebrations, and festivals, swimming together. It was always a pleasure to discuss with you Anders, be it my biochemical excitement or your technical expertise. I learned a lot about single

molecules from you. Chen, you will be always mysterious person. Thanks Ivo for always being there to talk, meet over the weekends. Thanks Kira for sharing biolab for most times as well as for starting the second Ph.D. defence wave. Thanks Nader for keeping our computers up and running and your all-time energy. Getting to know a talented pool of new generation colleagues like Frank, Simon, Vanessa, Adrian was great. It was always a pleasure seeing Vanessa for your excitement for various lab activities and digging in the non-chaperone story of Ssc1. Thanks Frank for cherishing the last years in the office in building B, Simon for timely expertise in data analysis, and most importantly for providing us tasty coffee. Thanks Evelyn for showing the calm and composure needed at times with us. Adrian, I have never seen anyone as a responsible person as you. The newest generation of Pooyeh, Fabian will certainly take the legacy far. A very special thanks to Moritz for wonderful explanations of German culture, especially bureaucracies.

Thanks to my beloved friends Ashish, Ajay, Disha, Vipul, Kaurab, Sanjay, Dipak, Bhimsen, Kannan, Satyam, Raj, Gopal, Rupa, Sharang, Muenster team (Pratibha, Deblina, Sana, Bhavin) for keeping sanity for the last decade.

Last but not the least, मला पंख दिल्याबद्दल आई आणि आन्ना यांचे खूप खूप आभार. आजी, तुमची काळजी घेण्याची संधी न मिळाल्याबद्दल मी नेहमी दु: खी असेन. I will be always and will continue thanking Rani Tai, Ramesh, Adi daji, Srushti and Kartik for providing shoulders many times throughout the years. I should not forget to mention the support I got from my in-law family includig Radha Suresh, Kumaraswamy Suresh and Nani.

And finally, my sincere thanks to my better half Mega, for your endurance and tolerance towards me. What's life without you! Let's make this life a great experience together.

Attached papers

Paper 1: A dynamic intermediate state limits the folding rate of a discontinuous two-domain protein

A dynamic intermediate state limits the folding rate of a discontinuous two-domain protein

Ganesh Agam[1], Anders Barth[1,†,] and Don C. Lamb[1]*

[1] Department of Chemistry, Center for NanoScience, Nanosystems Initiative Munich (NIM), and Center for Integrated Protein Science Munich (CiPSM), Ludwig-Maximilians University Munich, Munich, Germany.

[†] Current address: Department of Bionanoscience, Kavli Institute of Nanoscience Delft, Delft University of Technology, 2629, HZ, Delft, The Netherlands.

*** corresponding author: d.lamb@lmu.de**

ABSTRACT

Protein folding is an indispensable process for the majority of proteins after their synthesis from ribosomes in the cell. Most *in vitro* protein folding studies have focused on small, single-domain proteins, which account for approximately one-third of all proteins. Hence, it is also important to understand the folding process of large, multi-domain proteins, especially when the domains are discontinuous. To study the co-dependent folding of two globular discontinuous domains, we choose the Maltose binding protein (MBP) as a model system. In particular, we studied a mutant of MBP that folds slowly and contains two mutations in the N-terminal domain. Despite the wide use of the double mutant MBP (DM-MBP) in studies of the cellular folding machinery, a systematic study of its domain-wise folding is still missing. Here, using two- and three-color single-molecule Förster resonance energy transfer (smFRET) experiments, we study the refolding of both the domains and the interaction between the domains of DM-MBP. Initial two-color smFRET measurements of the N-terminal domain (NTD) reveal the presence of a folding intermediate, also known from previous studies. The same folding intermediate is observed in measurements monitoring the C- terminal domain (CTD) and the NTD-CTD (N-C) interface. The refolding intermediate is not due to the aggregation in the refolding reaction and is dynamic on the sub-millisecond timescale. The dynamics governs DM-MBP folding by shifting the equilibrium between native and unfolded conformation depending on the denaturant concentrations. Quantitative analysis on underlying dynamic interconversions revealed a delay in NTD folding imposed by the entropic barrier being the primary cause for slow DM-MBP folding. Moreover, highly dynamic CTD folds after NTD completes the folding. Using three-color smFRET, we could show the NTD folds first and CTD later in a same protein. Molecular dynamic simulations for temperature-induced unfolding on WT- and DM-MBP identify a folding nucleus in the NTD, which is rich in hydrophobic residues, and explains why the two mutations in this folding nucleus slow down the folding kinetics. In the presence of the bacterial Hsp60 chaperonin system GroEL/ES, we observe that DM-MBP is still dynamic within the chaperonin cavity but the chaperonin limits the conformational space that the substrate explores. Hence, confinement aids DM-MBP in overcoming the entropically limited folding barrier**.** The study reports on the subtle tuning and co-dependency for protein folding between two-domains with a discontinuous arrangement.

INTRODUCTION

The three-dimensional structure of a protein is defined by its primary amino acid sequence (Anfinsen, 1973). After its synthesis from the ribosome as a polypeptide chain, most of the proteins, with the exception of intrinsically disordered proteins, fold into their native structure either spontaneously or with the help of chaperones. Though the ultimate aim is to understand the folding process *in vivo* and studies of folding on ribosomes are becoming technically possible (Holtkamp et al., 2015)*, in vitro* unfolding-refolding studies allow precise control of the folding/unfolding conditions and investigation of the process in isolation (Sela et al., 1957). Most folding studies have focused on simple, small proteins (upper limit of ~100 amino acids) with a few exceptions (Batey et al., 2008). However, it is estimated that ~67% of the eukaryotic proteins have more than a single domain (Teichmann et al., 1998) (Chothia, Gough, Vogel, & Teichmann, 2003)(Han et al., 2007), with the majority being two-domain proteins (Jones et al., 1998). Two-domain proteins typically show continuous domain topology, ensuring the efficient folding of each domain independently. Continuous-domain proteins comprise about ~72% of all two-domain proteins (Jones et al., 1998). The remaining ~28% of two-domain proteins have discontinuous domains with the insertion of at least one region from one domain within the other. Discontinuous domains make the folding process energetically inefficient as the domains are co-dependent for assuming the final fold (Wetlaufer, 1973), (Collinet et al., 2000). Most enzymes and regulatory proteins exhibit this type of complex topology (Shank, Cecconi, Dill, Marqusee, & Bustamante, 2010) (Shank 2010). Despite their abundance in catalysis and in allosteric processes, little is known about their folding process. Most of the small proteins generally fold spontaneously within a short time (microseconds to milliseconds) (H. S. Chung, McHale, Louis, & Eaton, 2012). Different theories have been put forward to explain the process of folding (Wolynes et al., 1995) (Thirumalai, O'Brien, Morrison, & Hyeon, 2010). Large proteins having multiple domains fold relatively slow (seconds to minutes) and exhibit a complex multi-phase folding process (Arai, Iwakura, Matthews, & Bilsel, 2011), (Wolynes et al., 1995), (Onuchic & Wolynes, 2004). In the case of two-domain proteins, especially with a discontinuous domain topology, the main questions are whether the domains fold sequentially or simultaneously in a cooperative way.

Here, we investigated folding of the Maltose binding protein (MBP). MBP is a monomeric ~42 kDa protein having two discontinuous domains, namely the N- and C-terminal domains (NTD and CTD). Both domains contain α-β-α secondary structural motifs although the β-sheets extend over both the domains. The maltose-binding pocket is positioned between the domains (Spurlino, Lu, & Quiocho, 1991). The folding of WT-MBP and several folding-defective mutants have been studied previously (Chun, Strobel, Bassford, & Randall, 1993). One of them, containing the mutations V8G and Y283D in the NTD, folds with a significantly slower rate (half-life, $t_{1/2}$ ~30 min) as compared to WT-MBP ($t_{1/2}$ ~25 s) (Tang et al., 2006b). Interestingly, this double-mutant MBP (DM-MBP) shows intermediate states during the folding process (Sharma et al., 2008), (Chakraborty et al., 2010). The structure and domain-wise conformation of these intermediates is currently unknown but would be of high interest due to the discontinuous domain arrangement of DM-MBP (Wetlaufer, 1973) (Collinet et al., 2000).

Single-molecule Förster resonance energy transfer (smFRET) methods have been instrumental in understanding the conformational heterogeneity and dynamic properties of biomolecules. In two-color smFRET, one fluorophore (donor) transfers its energy non-radiatively to another fluorophore (acceptor), whereby the efficiency of energy transfer strongly

depends on the distance between them in the range of 2-10 nm (Forster, 1946) (Ha et al., 1996). To avoid artifacts due to surface immobilization, most smFRET-based folding studies have been performed on freely diffusing molecules except small domain proteins (A. A. Deniz et al., 2000), (Schuler et al., 2002), (H. S. Chung et al., 2012). Folding is the largest conformational change a protein can undergo. It is thought to be a cooperative process involving all parts of the protein (Wolynes et al., 1995). Compared to ensemble folding experiments, smFRET can investigate the heterogeneity of the populations present within the folding process (M. B. Borgia et al., 2011). SmFRET folding studies have led to insights into spontaneous folding, folding intermediates, and the regulation of folding pathways by various chaperone systems (Schuler et al., 2002), (Sharma et al., 2008), (Hofmann et al., 2010), (Schuler et al., 2002), (Kellner et al., 2014) (Dahiya et al., 2019). However, it remains challenging to study the folding pathway of a large protein with multiple domains by two-color smFRET (Pirchi et al., 2011). Multi-color smFRET has the unique advantage of monitoring multiple distances simultaneously (Hohng, Joo, & Ha, 2004), (Clamme & Deniz, 2005), (Nam et al., 2007). By positioning multiple fluorophores specifically on each domain, one can study the folding of individual domains in the context of the folding pathway of the whole protein. Recent advancements in diffusion- and surface-based multi-color smFRET single-molecule experiments have been applied to unravel the inter or intra-molecular correlated motions in biomolecules, (Gambin & Deniz, 2010), (Ratzke et al., 2014), (Barth, Voith Von Voithenberg, & Lamb, 2019).

In the current study, we characterize the known folding intermediate of DM-MBP and its importance in the overall folding. Furthermore, we ask the question whether both DM-MBP domains fold simultaneously or in a particular order. We address this by performing diffusion-based two- and three-color smFRET measurements of the NTD, CTD and of the N-C interface of DM-MBP. We characterized the previously reported intermediate state in the folding of DM-MBP under non-aggregating conditions (Chakraborty et al., 2010). Separate two-color smFRET measurements of the NTD, CTD and the N-C interface at different denaturant conditions revealed the unique conformation of the intermediate that is present in all the three probed coordinates. Moreover, plots of the FRET efficiency versus donor fluorescence lifetime show that the intermediate state is dynamic on the sub-millisecond timescale and fluctuates between native and unfolded conformations of the protein. Quantitative information on the microscopic rates of the conformational transitions governing the intermediate state were extracted employing a dynamic photon distribution analysis (PDA). Comparing the macroscopic and the microscopic rates of refolding confirmed that the intermediate state limits the folding mainly due to an entropic barrier. Analysis of the individual domains illustrated the delayed folding in the NTD being the primary cause for overall slow folding imposed by the double mutations. As soon as the NTD folds, folding of highly dynamic CTD precedes. Three-color smFRET allows to address the coordination of the conformations for three different positions at the same time. For three-color smFRET, the labeling positions were chosen such that the three inter-dye distances report on the folding of the NTD, CTD and N-C domain interface. Equilibrium three-color smFRET measurements provided evidence that the intermediate state is present in all the three coordinates simultaneously. Furthermore, probing three distances simultaneously shows that the native state of the two domains is assumed in a sequential way where NTD folds first followed by CTD , consistent with the results from the dynamic PDA on individual domains. Using molecular dynamics (MD) simulations, we investigate the temperature-induced unfolding of WT and DM-MBP and provide insights into the effect of the point mutations on the folding pathway. MD simulations are consistent with

the finding that the DM-MBP folding has to cross an entropic barrier, which delays folding. We also investigated the effect of GroEL\ES chaperonin, a bacterial Hsp60 chaperone and Hsp10 co-chaperone system, on the DM-MBP folding. It was previously shown that GroEL\ES accelerates the DM-MBP folding by 8-13 fold. We found that DM-MBP is still dynamic when bound to GroEL and when encapsulated within the cavity of GroEL\ES. Interestingly, when bound to GroEL alone, the equilibrium of the dynamics of DM-MBP is slightly shifted towards the unfolded conformation whereas, in co-operation with GroES, cavity confinement leads to a compaction of the dynamic folding intermediate state towards the native state. Hence, by limiting the conformational space explored by DM-MBP, the chaperonin helps DM-MBP to overcome the entropic barrier during folding. The proposed mechanism suggests one pathway for the folding of discontinuous multi-domains on physiological timescales when the folding of the individual domains is interdependent and how chaperonins accelerate their folding by modulating the folding landscape.

RESULTS

Characterization of a previously known hysteresis in the folding of DM-MBP

We first characterized the denaturation/refolding of DM-MBP using intrinsic tryptophan fluorescence at the ensemble level, taking advantage of the eight tryptophan residues present in DM-MBP (Figure 1A) (PDB ID:1OMP) (Sharff et al., 1992). Burial of solvent exposed hydrophobic tryptophan residues inside the protein core upon folding results in an increase in the tryptophan fluorescence. When DM-MBP is completely unfolded in 3 M guanidine-hydrochloride (GuHCl), the tryptophan fluorescence is reduced by ~2.5-fold as compared to the native protein. We first performed equilibrium measurements of unfolding and refolding by titrating with different GuHCl concentrations. The DM-MBP concentration for these experiments was kept at 40 nM, lower than the previous studies to avoid aggregation (Figure 1B). In the case of unfolding, when native DM-MBP was denatured with increasing concentrations of denaturant, DM-MBP starts to unfold above ~1.2 M GuHCl. Complete unfolding was observed above ~1.6 M GuHCl. The unfolding experiments display a two-state unfolding transition, which is typical for a globular protein. Strikingly, the protein only refolds from the denatured state starting at a concentration of ~0.4 M GuHCl, significantly lower than the concentration required for the beginning of unfolding (1.2 M). The hysteresis in formation of the native state with respect to unfolding under similar denaturing conditions has been attributed to kinetic traps present in the folding landscape of the protein (Andrews et al., 2013). The same observation was made by Chakraborty et al. 2010, who also proposed a kinetically trapped state in DM-MBP. Notably, DM-MBP completely refolds to its native state below a final concentration of ~0.1 M GuHCl in refolding buffer (Figure 1B). Hence, for all the ensemble and single molecule refolding kinetic measurements performed throughout this study, we kept the final GuHCl concentrations below 0.1 M. Next, we measured the kinetics of refolding. Denatured DM-MBP in 3 M GuHCl was allowed to refold after a 75-fold dilution at a final concentration of 40 nM DM-MBP and 40 mM GuHCl. The refolding half-time ($t_{1/2}$) for DM-MBP was found to be ~22.6 min with a refolding rate of 0.03 min^{-1} (Figure 1C) (Table 1).

The delayed refolding giving rise to the hysteresis effect might be caused by inter-molecular interactions (i.e. aggregation) of the unfolded peptide chain due to the high degree of exposed hydrophobic residues. To exclude this possibility, we used nanomolar concentrations of protein for the tryptophan fluorescence assays. To verify that there is no aggregation at the nanomolar concentrations used for these measurements, we performed fluorescence cross-

correlation spectroscopy (FCCS) experiments. FCCS, in combination with pulsed interleaved excitation (PIE) (Müller et al., 2005) is highly sensitive to the presence of any interaction between molecules labeled with different fluorophores, giving rise to a cross-correlation signal. As a positive control, we measured double-stranded DNA labeled with both Atto532 and Atto647N, which gives rise to a significant cross-correlation signal (green curve in Figure 1D). A single-cysteine mutant of DM-MBP (A52C) labeled with either Atto532 or Alexa647 sample (500 - 1000 nM) were unfolded in 3 M GuHCl for 30 min at 50°C, mixed in equal amounts and diluted to allow refolding in 0.1 M GuHCl at a final protein concentration of ~2 nM. No cross-correlation amplitude was detected, indicating that no aggregation occurred in the refolding over the course of 30 minutes (red and black curves in Figure 1D). As a negative control, a mixture of freely-diffusing dyes was also measured and showed no cross-correlation signal (cyan curve in Figure 1D). The observed hysteresis effect and refolding rate of DM-MBP are in good agreement with previous reports (Chakraborty et al., 2010) (Tang et al., 2006b). In addition, we could verify that the previously observed hysteresis is not induced by aggregation.

smFRET identifies a unique intermediate population as a kinetically trapped state

To gain further insights into the structural properties of the trapped state of DM-MBP during refolding, we performed single molecule two-color smFRET measurements using multiparameter fluorescence combined with PIE (MFD-PIE) (Kudryavtsev et al., 2012). By using picomolar (pM) concentrations of labeled molecules in combination with the small observation volume (~fL) of a confocal microscope, bursts of fluorescence signal are observed for single molecules diffusing through the observation spot on the order of several milliseconds. Using MFD-PIE, the fluorescence signal obtained for single molecule events contains information about the labelling stoichiometry, FRET efficiency, fluorescence lifetime and anisotropy. The type of information that can be collected is shown in Supplementary Figure 1 (see Materials and Methods). Since the two folding mutations are in the NTD of MBP, we first investigated the refolding of the NTD. To probe the conformation of the NTD, a double-cysteine mutant of DM-MBP (52C-298C) was stochastically labeled with the fluorophores Atto532 (donor) and Alexa647 (acceptor) using cysteine-maleimide chemistry (Figure 1A) (Materials and Methods). The Förster distance, the inter-dye separation at which the FRET efficiency reduces to 50%, was determined to be ~62 Å for Atto532 and Alexa647 (Voith von Voithenberg et al., 2016) (Table 2). The two-color smFRET measurements exhibited a FRET efficiency of ~0.85 for the native NTD (Figure 2A, first row of FRET efficiency histograms). From a photon distribution analysis (PDA) of the smFRET efficiency histogram, we calculated an inter-dye distance distribution with an average separation of 48.3 Å. From the crystal structure (PDB ID:1OMP), an accessible volume (AV) calculation for the given dyes yielded a distance of 51.3 Å (Table 3) (Supplementary Figure 2A) (Materials and Methods) (Antonik et al., 2006) (Kalinin et al., 2012). Thus, the experimentally determined distance is indicative of a folded NTD. When we measured the NTD conformation under denaturing conditions in 2 M GuHCl, a FRET efficiency of ~0.1 was obtained (Figure 2A, last row of FRET efficiency histograms). For this unfolded state, the PDA analysis yielded an average separation of ~83.8 Å with a broad width (17 Å) (Table 3). This broad distance distribution is typical for unfolded polypeptide chains existing in an ensemble of conformations (Thirumalai et al., 2010). The above FRET efficiency values and donor-acceptor separations for both the native and unfolded state agrees well with previous smFRET measurements using fluorophores with similar spectroscopic properties attached at the same labeling positions (Sharma et al., 2008).

Next, we performed equilibrium unfolding and refolding measurements on the NTD at different GuHCl concentrations (Figure 2A), similar to the unfolding and refolding measurements performed using tryptophan fluorescence (Figure 1C). Gaussian fits to the obtained FRET efficiency distributions identified one intermediate FRET population with ~0.6 FRET efficiency, in addition to the native and completely unfolded states. Note that, as the denaturant concentration decreases under the denaturing conditions from 2 M to 0.9 M, the FRET efficiency of the unfolded state changes from ~0.1 to ~0.3 due to compaction of the denatured polypeptide at low GuHCl concentrations. To trace the origin of the intermediate population, the obtained average FRET efficiencies were compared with the tryptophan fluorescence measurements. Interestingly, during the refolding measurements, there is a significant contribution of the intermediate FRET population to the smFRET histograms for measurements collected within the hysteresis range of 0.4-1.4 M GuHCl (Figure 2B). The above analysis gives a glimpse into the folding process: starting from the completely unfolded state of ~0.08 FRET efficiency, there is a gradual hydrophobic collapse (depending on the GuHCl concentration in the buffer), and then the protein starts to refold into the native state. During refolding from the collapse state to the native state, a peak is observed in the smFRET histogram with an intermediate FRET efficiency of 0.6.

To analyze the refolding protein under non-equilibrium conditions, we performed a kinetic analysis of a smFRET measurement where refolding was initiated in denatured DM-MBP by a 75-fold dilution of the denaturant. Indeed, during the first 2000 s of the measurement, the refolding NTD had a FRET efficiency of 0.6 in the beginning and completely folds to a native state of 0.85 FRET efficiency (Figure 2C). From a quantitative analysis of the FRET efficiency histograms, we obtained a $t_{1/2}$ of 29.5 min, a similar rate as to what was obtained for refolding of an unlabeled protein (Figure 2D) (Table 1). This indicates that the fluorophores do not interfere with the refolding process of DM-MBP. In summary, the above equilibrium and kinetic smFRET experiments probing NTD has identified an intermediate state unique to the refolding reaction.

The intermediate state is also present in the CTD and N-C interface

Having characterized the intermediate state in the NTD, we investigated whether refolding of the CTD and formation of the N-C interface also exhibit the hysteresis. To monitor the CTD conformation, we made use of the previously characterized double-cysteine mutant of DM-MBP (175C-298C) (Figure 1A) (Sharma et al., 2008). Equilibrium unfolding and refolding curves were recorded using two-color smFRET measurements on the CTD labeled with the same dye-pair as for the NTD (Atto532 and Alexa647, Figure 3A). The native and completely unfolded states of the CTD have similar donor-acceptor separations as for the native and completely unfolded state of NTD, respectively (Supplementary Figure 2B) (Table 3). A plot of the FRET efficiency versus the titrated GuHCl concentration for folding and unfolding show a similar trend as for the NTD where unfolding behaves as a two-state system and refolding has a similar intermediate population with a FRET efficiency of 0.6 (Figure 3A). Kinetic experiments on CTD refolding yielded a $t_{1/2}$ of 24.9 min (Figure 3B). A similar observation was made for equilibrium and kinetic measurements on the N-C interface (52C-175C) (Figure 3C,D). For the N-C interface, the intermediate population is more compact with a FRET efficiency of 0.85. The structure of the N-C interface is well defined in the native state, having a small width of 2.8 Å for a inter-dye distance of 44.8 Å as determined by PDA. The AV simulations predict an inter-dye separation of 40.0 Å (Supplementary Figure 2C and Table 3). The unfolded state has a distance of 74.9 Å, comparable to the NTD and CTD unfolded

states (Table 3). The kinetics of refolding are similar to the other mutants and to the unlabeled protein, which confirms the functionality of the both the labeled CTD and N-C interface constructs (Figure 3B and 3D). Due to the similar FRET efficiencies of the folded and the kinetically trapped state in the N-C interface construct, a kinetic analysis based on the FRET efficiency histograms is not possible. Hence, we made use of the observation that the fluorescence lifetime of Alexa647 is different in the unfolded and the refolded protein (Supplementary Figure 3A). Due to steric restriction, the cis-trans isomerization of Alexa647 is hindered in the unfolded state, typical for cyanine-based dyes. The dye hence exhibits a higher lifetime of ~1.7 ns for the unfolded state, but has a lifetime of ~1.2 ns for the folded protein, close to that of free dye in water (Supplementary Figure 3A-B and Table S1). Despite the different acceptor lifetimes, the measured anisotropy for Alexa647 remained below ~0.25 for all constructs (Table S3). To ensure that the observed intermediate was not an artifact arising from the acceptor, we measured refolding with a different acceptor, Atto647N, where the same intermediate state was measured (Supplementary Figure 3C). Appropriate corrections were made to account for differences in the acceptor quantum yield for the different populations for all constructs labeled with Alexa647 (Supplementary Figure 3D and Table 2).

The above smFRET measurements on the NTD, CTD and N-C interface all exhibit the presence of an intermediate population during folding. To explore the origin of this universal intermediate in DM-MBP folding, we carried out two-color smFRET experiments on wt-MBP folding on one of the domains, in this case the CTD. As wt-MBP folds within ~25 s, it should not have any intermediate population visible in the 0.4-1.4 M GuHCl refolding titrations. Indeed, both unfolding and refolding curves have the same transition between the unfolded and native state, without any visible intermediate population (Supplementary Figure 4). This confirms that the observed intermediate state is unique to DM-MBP and absent in the wildtype protein.

The Intermediate state is a dynamically averaged state undergoing sub-millisecond dynamics

Next, we investigated whether the intermediate state is structured or dynamic. Here, we made use of the fluorescence lifetime of the donor fluorophore, which is available in MFD-PIE experiments (Supplementary Figure 1). In the absence of dynamics, there is a linear relationship between the FRET efficiency and the donor fluorescence lifetime (static FRET-line, black line in upper panels of Figure 4A-C, Materials and Methods). When a molecule interconverts between different conformations (with different FRET efficiencies that correspond to a specific donor fluorescence lifetime) during the observation time, a deviation from the static FRET-line is observed (Materials and Methods) (Kalinin, Valeri, et al., 2010). We performed a lifetime analysis of two-color measurements for all the three constructs (Table 4, Figure 4A-C). In all cases, the intermediate states deviate from the static FRET-line, indicating the presence of conformational dynamics. The observed deviation can be described by a single dynamic-FRET line (red line in upper panels of Figure 4A-C) connecting the unfolded and native states. This indicates that the intermediate state is not a distinct structural state, but originates from dynamic switching between the unfolded and folded states during the observation time (~ms, Figure 4D). The intermediate state is thus a dynamically averaged state where conformational fluctuations of the molecule much faster than the burst duration lead to a species-averaged FRET efficiency in the intensity-based calculations. To determine the FRET efficiencies of the end conformations, we fit the donor fluorescence lifetime data from all detected bursts to a biexponential model function (Supplementary Figure 5). The

determined FRET efficiencies were compared with the intensity averaged FRET efficiency values and are plotted as white boxes in the lower panels of Figure 4A-C (Table 4). Note that the equilibrium between the unfolded and folded states shifts with denaturant concentrations and, hence, a corresponding FRET efficiency averaging takes place. Moreover, changing the acceptor dye does not affect the dynamics present in the system (Supplementary Figure 6).

Interestingly, when comparing the refolding curves for all the three constructs, we observed that the refolding reaction performed in 0.1 M GuHCl where DM-MBP completely refolds, NTD and N-C interface significantly showed folded fraction whereas CTD still has substantial fraction in the intermediate state (lower panels of Figure 4A-C). We looked into more detail in this direction in the following section finding the kinetics of the underlying transitions.

Dynamic Photon distribution analysis quantifies the folding order of the domains

To quantify the microscopic rates for the conformational transitions during refolding, we employed dynamic PDA (Kalinin, Valeri, et al., 2010). This analysis routine quantifies the rates at which subpopulations interconvert during the burst duration of a few milliseconds. The distance distributions are modeled on the raw photon counts to quantify the broadening beyond shot-noise. The robustness of the fit is increased by generating FRET efficiency histograms (or proximity ratio histograms) of different time resolutions (or binning the photons) on the order of milliseconds (Material and Methods) and globally fitting the histograms. The distances, widths and kinetic rates can be extracted from the analysis to determine the microscopic rates between the states, where states were defined by donor lifetimes (Table 4).

Table 5 summarizes the microscopic rates for conformational fluctuations between the natively folded (F) and unfolded state (U). For all three constructs, the unfolding rate ($k_{F\to U}$) increases with increasing GuHCl concentration (Figure 5A-C) (Supplementary Figure 7). This is consistent with a loss of the natively stabilized contacts by the increasing amount of GuHCl in solution. In contrast, there are striking differences in the folding rates ($k_{U\to F}$) of the three constructs. For the CTD and N-C interface, the transition rates remained unchanged between 2-3 ms^{-1} for the CTD and between 3.3 and 3.9 ms^{-1} for the N-C interface throughout all the refolding curves (Supplementary Figure 7) with the exception of refolding of the N-C interface in 0.1 M GuHCl (Table 5). As GuHCl destabilizes native contacts, the refolding rates are expected to be independent of GuHCl concentration. The reduction in the refolding kinetics of the N-C interface at 0.1 GuHCl is most likely due to the slower structural rearrangements necessary during the final folding step. For the NTD, a more complicated pattern is observed. The unfolding rates ($k_{F\to U}$) are between 2-3 ms^{-1} for lower GuHCl concentrations and then slow down to 0.6 - 0.9 ms^{-1} at higher GuHCl concentrations. One possible explanation is that the hydrophobic core in the NTD defined by the residues 8-21, 22-24 and 45-63 is needed to stabilize the folded state (Ye, Mayne, Kan, & Englander, 2018). Disruption of the hydrophobic core by GuHCl also decreases the refolding transition monitored with the NTD FRET construct. The similar relaxation times for the NTD and the N-C interface indicates cooperativity in the folding process.

With the given dynamic PDA analysis, we conclude that the CTD folds and unfolds quickly, waiting for the NTD to fold to stabilize the structure. Once the NTD finds the correct structure, both the CTD and N-C interface can lock in place. The combined microscopic rates of 0.75-0.84 ms^{-1} for the transitions in the NTD and N-C interface and 3.4 ms^{-1} in CTD are much faster than the macroscopic folding rate of 0.03 min^{-1} (3×10^{-7} ms^{-1}) for DM-MBP (Table 1). This provides direct evidence for the rate-limiting step in DM-MBP folding being the native-like

contact formation. From Kramer's equation, the difference in Gibbs free energy (ΔG_f) required to cross the folding barrier can be calculated directly from the folding times (Kramers, 1940), (Schuler et al., 2002),

$$\tau_f \approx 2\pi\tau_0 exp\left(\frac{\Delta G_f}{RT}\right) \quad (2)$$

where τ_f is the folding time, R is the gas constant and T is the temperature. Here, an assumption is made for harmonic frequencies to be uniform at the surface of an unfolded well and the top of the barrier. τ_0 is the reconfiguration time in the unfolded well and hence defines the pre-exponential factor. τ_0 is assumed to be $\tau_0 \approx 100$ ns based on what has been observed in other proteins (Soranno et al., 2012) (Soranno et al., 2017), the above assumption and from the folding time, $\tau_f \sim 2 \times 10^6$ ms (33.3 minutes) for DM-MBP folding, a ΔG_f of 51.75 41.2 /mol is required to achieve folding. When compared to the wt-MBP case, with the folding rate from ensemble measurement of 1.8 min^{-1} ($\tau_f \sim 33.3 \times 10^3$ ms), a Gibbs free energy for folding of 41.20 kJ/mol is observed (Figure 5D). The additional 10.55 kJ/mol Gibbs free energy difference in the DM-MBP folding is unlikely to be enthalpic in nature and could be explained by the extra dynamics due to the absence of the hydrophobic core in DM-MBP. To verify that the barrier in DM-MBP folding is entropic in nature, trimethylamine N-oxide (TMAO) was incorporated in the assay buffer when studying DM-MBP refolding. TMAO is known to act as a chemical chaperone that confines the configuration space available to the protein by stabilizing the solvent shell around the protein (Bandyopadhyay et al., 2012). The confinement of DM-MBP by TMAO, similar to the cavity of GroEL/ES (Chakraborty et al., 2010), accelerated the formation of the native-like structure as observed for refolding of the NTD in 0.2 M GuHCl, where DM-MBP has significant population of the flexibility driven-trapped intermediate state without TMAO (Figure 5E). These finding are consistent with the notion that the folding intermediate has indeed an entropic trap rather the local kinetic trap as TMAO would stabilize the kinetic trap further.

The native and unfolded state of DM-MBP can be probed with three-color FRET

Two-color FRET allowed us to probe the folding of the different domains separately. To obtain a coherent global view of DM-MBP folding, we applied three-color smFRET. Multi-color smFRET has the potential to visualize the correlative motions in different parts of biomolecules in real time (Gambin & Deniz, 2010), (Ratzke et al., 2014),(Barth, Voith Von Voithenberg, & Lamb, 2019). Using three labels on the protein, we can probe the folding of the two domains and the inter-domain interface simultaneously using the three FRET efficiencies between the blue, green and red fluorophores (Figure 6A) (Supplementary Figure 8A). For three-color FRET, we used the dyes Atto488, Atto565 and Alexa647, which were chosen to maximize the use of visible spectrum with a distance sensitivity indicated by their respective Förster distances (~50-70 Å) (Table 2). From a three-color smFRET measurement, we obtain the three FRET efficiencies between the different dye pairs: blue and green (BG), blue and red (BR) and green and red (GR), allowing us to address the three intramolecular distances in one MFD-PIE smFRET experiment (Materials and Methods, Supplementary Figure 8B-E) (Barth, Voith Von Voithenberg, & Lamb, 2019). For the three-color smFRET experiment, it is important to label the three fluorophores site-specifically. To this end, we incorporated an unnatural amino acid at one position, while ensuring selective labeling of the two cysteines using their differential solvent accessibility when maltose is bound (Supplementary Figure 9) (Materials and Methods). The triple-labeled DM-MBP was confirmed to be able to fold by assaying the

kinetics of Alexa647 lifetime change upon refolding, which was found to be similar to the unlabeled protein (Supplementary Figure 8F) (Table 1 and S2).

First, we measured the native, denatured and refolded state of DM-MBP. According to the labeling positions, the FRET efficiency between the blue and green dyes (BG) reports on the NTD conformation, while the FRET efficiency between the green and red dyes (GR) corresponds to the CTD conformation and the FRET efficiency between the blue and red dyes (BR) monitors the N-C interface (Figure A). The three-color smFRET measurement of triple-labeled DM-MBP in the native state showed a high FRET peak at ~0.9 for all the three FRET efficiencies. Refolded DM-MBP showed the same FRET efficiencies as for the native state for all three FRET pairs, confirming the correct refolding of the triple-labeled DM-MBP (Figure 6B-D). When DM-MBP was denatured in 3 M GuHCl, the FRET efficiency BG and BR was found to be centered at a value of ~0, while the FRET efficiency GR was at ~0.35 (Figure 6 B-D). To confirm the specific labeling of the dyes and to validate the results with the dyes used for the three-color FRET experiments, we additionally measured double-cysteine mutants probing NTD, CTD and N-C interface separately with two-color smFRET measurements (Supplementary Figure 10A-C) using the respective dye-pairs.

Three-color FRET on DM-MBP revealed the correlated domain-wise folding of a globally trapped state towards the native state

As before, we performed equilibrium unfolding and refolding measurements using triple-labeled DM-MBP. Consistent with the results of the two-color smFRET experiments (Figures 2-3), a two-state unfolding was observed (Figure 6E). Next, we probed the equilibrium refolding, obtaining a similar trend as compared to the two-color FRET experiments (Figure 2-3 and Figure 6F). For the histogram of the FRET efficiency between the blue and green dye, an intermediate population with E_{BG} = 0.7 was evident at concentrations of 0.1-0.5 M GuHCl, in agreement with the intermediate state seen for the NTD (Figure 2) and the two-color FRET controls using the blue-green dye pair (Supplementary Figure 10D-E, left panels). Similarly, an intermediate state was present for the CTD and N-C interface with E_{GR} ~0.8 and E_{BG}~0.7, respectively, which was confirmed by two-color FRET controls with the green-red and blue-red dye pairs, respectively (Supplementary Figure 10D-E, middle and right panels). In summary, the three-color FRET measurements confirmed the global intermediate state spanning both the domains and inter-domain interface.

A question that could not be addressed by the two-color FRET experiments is whether transition from the folding intermediate to the native state occurs independently for the NTD and CTD, or whether the two domains fold cooperatively in the same protein molecule. Dynamic PDA on two-color smFRET gave an evidence for the order of folding but three-color smFRET can resolve these transitions in the same molecule. This information is made available from the three-color smFRET experiment by selecting molecules in the intermediate state for one domain and examining the conformation of the other domain (Figure 7A). We used a concentration of 0.5 M GuHCl where the transition occurs from intermediate to the native state (Figure 6F). Molecules were classified according to their conformation detected in the NTD with molecules in the intermediate state were selected using an upper threshold of E_{BG} ~0.7 and in the native state using a lower threshold of E_{BG} ~0.9 (Figure 6F and Supplementary Figure 10E). The conformation of the CTD and N-C interface were probed using the FRET efficiencies E_{GR} and E_{BR}. Molecules found in the intermediate state of the NTD also exhibited the intermediate state for the CTD and N-C interface. On the other hand, the majority of molecules that showed a folded NTD had also reached the native state N-C

interface but interestingly CTD still has not folded to the native state (Figure 7A and 7B). Thus, the correlated information available with three-color FRET suggests that the intermediate state transitions to the native state in a sequential fashion where NTD and N-C interface folds first and CTD folding follows later.

Molecular Dynamic simulations of MBP and DM-MBP unfolding

To obtain structural insights into the folding of DM-MBP, we performed all-atom molecular dynamics (MD) simulations. It is currently not possible to simulate the folding trajectory starting from the denatured state on a realistic time scale. However, unfolding can be induced by increasing the temperature, giving insights into the folding pathways of proteins (Lazaridis & Karplus, 1997). Assuming microscopic reversibility of the folding and unfolding processes, the temperature-induced unfolding would corresponds to the reversed folding pathway (Dinner & Karplus, 1999),(Toofanny & Daggett, 2012). In the case of WT-MBP, which folds within a minute, it was shown that the NTD folds before the CTD (Walters et al., 2013). Here, for DM-MBP, we found the same folding order but delayed on the timescale of ~20-30 min (Figure 1C) (Tang et al., 2006b). To investigate the origin of these differences, we simulated the temperature-induced unfolding of WT-MBP and DM-MBP with all atom MD simulations starting from the native conformation (PDB: 1OMP). Simulations were carried out for 2 µs time at 450 K after equilibrating the protein for 2 µs at 400 K. Indeed, the unfolding trajectory for WT-MBP shows that the majority of the secondary structures in the CTD unfolds first, while, for the NTD, parts of the secondary structure are preserved until the end of the simulation (Figure 7C). These regions include two alpha-helices (15-25 amino acid residues and 280-286 amino acid residues) and two-anti-parallel beta sheets (Figure 7F). This suggests that, when WT-MBP folds, these two helices and two- beta sheets are the first structural elements in the NTD and the rest of the CTD folds later. This folding nucleus is disrupted by the double mutations introduced by the DM-MBP. The root-mean-square-deviation (RMSD) of the backbone atoms for WT-MBP unfolding shows step-wise changes, implying step-wise unfolding events (Figure 7E, left panel). We ran two repeats for WT-MBP for unfolding simulations directly at 450 K for 2 µs without the initial 2 µs equilibration run at 400 K (Supplementary Figure 11A-B), which showed the same trend of unfolding events for the NTD and CTD. These results are in agreement with the folding order measured in previous studies on WT-MBP (Walters et al., 2013), (Ye et al., 2018). Strikingly, for DM-MBP, one global unfolding event was observed, and no major secondary structure was preserved at the end of the simulation (Figure 7D). In addition, the RMSD for DM-MBP unfolding also indicates one major unfolding event within 500 ns of the temperature shift from 400 K to 450 K (Figure 7E, right panel). Repeating the simulations run directly at 450 K also lead to complete unfolding in a single step within 300 ns (Supplementary Figure 11C-D). This clearly suggests a decreased stability as the loss of native contacts was higher and faster in DM-MBP.

Thus, results from MD simulations on wt-MBP correctly captured the folding order showing that the NTD folds first followed by the CTD. This shows that the order of folding is conserved from wt- to the double mutant version but with a significant delay in the latter case. However, while the NTD folds faster in wt-MBP, the CTD fluctuates quickly between the folded and unfolded conformations in DM-MBP and can lock into the folded state once the NTD domain finishes folding.

GroEL cavity and GroEL\ES confinement modulates the dynamics in the folding landscape of DM-MBP

Molecular chaperones are key players in the process of protein folding. They help nascent chains as well as denatured and aggregated proteins to fold in their final native structure (Balchin et al., 2016). The bacterial chaperonin, GroEL\ES, is one of the most widely studied system to understand the mechanism of chaperone assisted protein folding. Previously, it has been shown that the GroEL\ES cavity has effect in accelerating DM-MBP spontaneous folding by 8-13 fold (Tang et al., 2006b) (Sharma et al., 2008). In the current study, we have elucidated the underlying fast dynamics of DM-MBP, where the protein fluctuates between near a native state and the unfolded configuration. We asked the question how the GroEL\ES cavity influences the folding dynamics in DM-MBP. Here, we analyzed the initial 5 minutes of DM-MBP folding of the NTD spontaneously as well as in the cavity of GroE\ES. During spontaneous folding, a broad FRET efficiency distribution was observed that exhibits dynamic fluctuations between near native and unfolded state (Figure 8A). When DM-MBP is bound to GroEL in the absence of GroES and ATP, the majority of the DM-MBP is still dynamic between the same of similar states but with an equilibrium shifted towards the unfolded state. In addition, a new static conformation is observed where DM-MBP is further stretched as observed for unfolded DM-MBP conformation (Figure 8B), as observed previously by Sharma et al. (Sharma et al., 2008). Strikingly, a dynamic intermediate is still observed within the cavity of GroEL\ES with similar endpoints observed in the fluorescence lifetime distribution, but the equilibrium is significantly shifted to the compact, high FRET native-like conformation. This suggests that GroEL\ES induces native like conformation in DM-MBP within 5 minutes as previously found, but this study reveals that DM-MBP is still dynamic in the GroEL\ES cavity. However, the chaperonin cavity limits the conformational landscape of the protein, decreasing the entropy of finding the correct native-like structure and thus increasing the folding rate (Figure 8C), as also seen in Sharma et al.

DISCUSSION

Spontaneous folding of MBP has been studied for over three decades to investigate different facets of the protein folding process (Spurlino et al., 1991),(Chun et al., 1993),(Raffy, Sassoon, Hofnung, & Betton, 1998),(Bertz & Rief, 2008),(Walters et al., 2013),(Selmke et al., 2018). MBP and its slow-folding-mutants, SM-MBP (Y283D) and DM-MBP (V8G and Y283D) have been extensively used as model substrates for chaperone assisted folding studies (H Sparrer et al., 1997),(J. D. Wang, Michelitsch, & Weissman, 1998a),(Tang et al., 2006b). Importantly, the unfolded states and folding intermediates of WT-MBP and MBP folding mutants show similar features under *in vivo* and *in vitro* conditions, as they both serve as substrate models for different chaperone systems (Mapa, Tiwari, Kumar, Jayaraj, & Maiti, 2012b), (Bandyopadhyay et al., 2012) (C. Huang, Rossi, Saio, & Kalodimos, 2016). Some of the *E. coli* chaperone systems studied with WT-MBP and DM-MBP are SecB (part of translocation machinery)(Bechtluft et al., 2007), trigger factor (ribosome-associated chaperone)(Mashaghi et al., 2013), DnaK/DnaJ/GrpE (cytoplasmic Hsp40/Hsp70 chaperone system)(Mashaghi et al., 2016) and GroEL/ES chaperonins (Hsp60/Hsp10 chaperonins)(Sharma et al., 2008),(Chakraborty et al., 2010),(Gupta et al., 2014),(Ye et al., 2018).

In three-color smFRET, three spectrally distinct fluorescent dyes are attached to the specific sites on a protein or on multiple interacting partners to obtain three FRET efficiencies at a time. In general, three separately measured FRET efficiencies can also serve the purpose to get the same information, but three-color smFRET has an advantage to have simultaneous information of three distance vectors (Clamme & Deniz, 2005), (Gambin & Deniz, 2010).

Importantly, simultaneous information gives access to observe the degree of co-ordination among them. This information is useful, especially for biological systems (Ratzke et al., 2014), (Barth, Voith Von Voithenberg, & Lamb, 2019). Thus, three-color smFRET has opened the possibility to study folding dynamics in multiple domains of large proteins.

In this report, we have investigated the DM-MBP domain-wise folding by considering each domain and the domain interface separately with both by two-color and three-color smFRET. Two-color smFRET provided the information about the conformational heterogeneity with clearly separated unfolded, intermediate and native states in all the three co-ordinates during DM-MBP refolding. We showed that the previously identified intermediate state is indeed not a kinetically-trapped state, but originates from rapid folding-unfolding dynamics on the micro-millisecond time-scale. Additionally, strong evidence provided by dynamic PDA confirms the belief that NTD folds first and CTD folds later. Though, the evidence for co-ordination between all the three vectors was not possible with two-color smFRET analysis in the same molecule; uniquely, three-color smFRET filled this gap by providing the unpreceded details about the correlative motions in domains in DM-MBP folding.

Since the inception of first studies on SM-MBP and DM-MBP as a substrate in the context of GroEL/ES chaperonins, there has been significant interest to understand the effect of mutations on spontaneous folding of MBP (Helmut Sparrer, Lilie, & Buchner, 1996), (J. D. Wang et al., 1998a). WT-MBP folds within a minute ($t_{1/2}$ ~25 s) while SM-MBP ($t_{1/2}$ ~3 min) and DM-MBP ($t_{1/2}$ ~30 min) mutants folding is delayed (Tang et al., 2006b). Subsequent studies revealed with advent of single molecule FRET methods, that DM-MBP stays in a long-lived intermediate stat refolding is delayed after diluting out the denaturant, which has been attributed to the population of long-lived intermediate states. Characteristic properties shown by the hypothesized long-lived intermediate state, e.g. the exposition of hydrophobic patches, presumably explained its strong interaction with DnaK/J/E and GroEL/ES chaperones (Sharma et al., 2008) (Mapa et al., 2012b). The long-lived intermediate has been thought to be the origin of the hysteresis in the unfolding-refolding curve (Figure 1C) (Chakraborty et al., 2010). Often, *In vitro* folding studies on complex systems can result into aggregation showing hysteresis effect (Moon, Kwon, & Fleming, 2011). We ruled out this possibility in case of hysteresis present in DM-MBP refolding by confirming absence of any reversible aggregates with FCCS (Figure 1D). In this study we have characterized the hysteresis in DM-MBP, with both two and three-color smFRET. Certainly, the cause of the of hysteresis can be attributed to the properties of the conformations displayed by DM-MBP refolding molecules between 0.4-~1.4 M GuHCl concentrations, as seen previously as well (Figure 2-3, Figure 6) (Chakraborty et al., 2010). It has been speculated that the double mutations present in DM-MBP, delay the formation of nucleation core in NTD, a rate limiting step in overall folding of a protein (Chun et al., 1993). NTD core encompasses the residues which are far apart in sequence, but juxta positioned in the native state. Disruption of this core as a result of two point-mutations with less hydrophobic side chains (V8G, Y283D), led to increased flexibility of the segment containing two helices and two beta sheets (Figure 7C-F) (Ye et al., 2018). Constraining this flexible region by introducing disulfide bridge tethering the distant segments found to remove the hysteresis by decreasing the entropic barrier to form native-like contacts. Furthermore, usage of a small molecule known to reduce structure flexibility in protein loops enhances the folding of DM-MBP (Figure 5E) (Bandyopadhyay et al., 2012). This indicates that DM-MBP encounters an entropic barrier leading to a local minimum in the folding funnel, which the protein has to overcome to attain the native state (Figure 4D) (Figure 5D). The resulted barrier

can be overcome e.g. with the help of a chaperonin cage that decreases the entropic barrier by confinement (Gupta et al., 2014), (Tang et al., 2006b), (Chakraborty et al., 2010).

A hydrogen exchange-mass spectroscopy (HX-MS) study on WT-MBP revealed that the NTD folds first ($t_{1/2}$ ~1 s) while the CTD folds later ($t_{1/2}$ ~40 s) (Walters et al., 2013). Subsequent HX-MS study with one of the mutant V8G in MBP demonstrated the delay in NTD folding ($t_{1/2}$ ~20 s) and no effect on CTD folding ($t_{1/2}$ ~40 s) (Ye et al., 2018), implying that the V8G mutation doesn't impart the dependency of the CTD on the NTD folding. In the current study, we confirmed the dependency between the folding of the NTD and CTD in the presence of the additional Y283D mutation, limiting the overall folding of DM-MBP ($t_{1/2}$ ~30 min), as shown with the analysis of the folding kinetics by dynamic PDA (Figure 5A-C) (Table 5). Remarkably, three-color smFRET measurements allowed us to follow the folding of both the domains and domain interface simultaneously. This provided strong evidence for implicated intermediate state, spanning both the domains and inter-domain interface, dependent folding of the NTD and CTD in the same molecule (Figure 7A-B). The correlated information obtained from three-color smFRET provided direct evidence for the order of the domain folding, showing that the folding of the NTD is necessary for the CTD to fold. This demonstrates how certain mutation affects the domain-wise folding as well as the extend of domain dependency on overall folding as a result of sequence discontinuity in domains imposing a significant entropic barrier of 10.55 kJ/mol (Figure 5D). Furthermore, it recapitulates the evolution of cooperative domain folding having discontinuity in amino acid sequence (Vogel, Bashton, Kerrison, Chothia, & Teichmann, 2004) (Han et al., 2007) (Inanami et al., 2014) (Y. Wang et al., 2012) (Arviv & Levy, 2012), having implications in preventing aggregation between the domains and subsequent misfolding stress (M. B. Borgia et al., 2011), (Mashaghi et al., 2013) (Cerminara et al., 2020). Remarkably, to achieve the task of complex multi-domain folding, cells might have to depend occasionally on ribosomes and chaperones to attain native structures in physiological time-scales (Holtkamp et al., 2015) (Vishwas R. Agashe et al., 2004) (Liu, Maciuba, et al., 2019) (Imamoglu et al., 2020). This also supports the notion of co-evolved chaperone-substrate networks. We found that DM-MBP bound to GroEL alone and in the cavity of GroEL\ES is still mostly dynamic, but the environment reshapes the underlying dynamics and explored conformational space. The hydrophobic lining of GroEL captures the unfolded substrate and shifts the equilibrium conformation slightly towards the unfolded state whereas the hydrophilic cavity of GroEL\ES limits the conformational space explored by the protein and allows it to move downhill in the folding funnel. This shows the capability of chaperones to tune the folding landscape of a substrate protein and is one mechanism to overcome entropic barriers in multi-domain proteins.

In summary, the presented study illustrates how the folding of complex multi-domain proteins can be studied using two- and three-color smFRET. Using this workflow, the generality of the folding pathways can be tested using well studied systems like adenylate kinase and phosphoglycerate kinase with complex domain topologies (Osváth et al., 2005) (Li et al., 2012). Certainly, this is possible with improved spectral choices for multi-color smFRET and robust labeling approaches to achieve specific labeling of more than two fluorophores. Eventually, it would be interesting to see how known interacting chaperone systems of DM-MBP like DnaK/J/E and GroEL/ES modulates the domain-wise folding of DM-MBP with the feasibility to perform three-color smFRET measurements. This can provide the difference in general folding mechanisms adopted by these two folding systems which might be applicable to speculate about folding pathways of their specific substrates.

MATERIALS AND METHODS

MBP constructs, protein expression and purification

Plasmid encoding DM-MBP (V8G, Y283D) was a generous gift from F. Ulrich Hartl and Manajit Hayer-Hartl (MPI of Biochemistry, Martinsried, Germany). The DM-MBP plasmid backbone is pCH-series based, (Chang et al., 2005), enabling IPTG inducible expression. Single cysteine (A52C, P298C), double cysteine (A52C-P298C, K175C-P298C, A52C-K175C) DM-MBP mutants were generated using site-directed mutagenesis (Thermo Scientific-Phusion Site-Directed Mutagenesis Kit). The A52TAG-K175C-P298C mutant was generated by mutating the codon for alanine at position 52 to the amber stop codon (TAG) to incorporate the unnatural amino acid (UAA) N-Propargyl-L-Lysine (PrK).

All single and double cysteine DM-MBP mutant proteins were expressed in *E. coli* BL21-AI (L-(+)-arabinose controlled T7 RNA polymerase expressing strain) at 30° C for 4 hr, 200 RPM with the addition of both 0.2% L-(+)-arabinose and 0.5 mM IPTG.

UAA incorporation requires the co-expression of the used orthogonal translation system of $tRNA^{Pyl}$ and $PylRS^{WT}$ with the expression of the given protein. For this purpose, we used pEvol-$tRNA^{Pyl}PylRS^{WT}$ (A. Chatterjee et al., 2012). *E. coli* BL21-AI cells harboring both pCH-A52TAG-K175C-P298C-DM-MBP and pEvol-$tRNA^{Pyl}PylRS^{WT}$ plasmids were grown for 1 hr at 30° C before the addition of 1mM of PrK (SiChem GmbH) in the media. PrK was prepared in 0.1 M NaOH solution. Later, DM-MBP-A52PrK-K175C-P298C expression was achieved similar to the other mutants with L- (+)-arabinose and IPTG.

All DM-MBP proteins were purified with an amylose column (New England Biolabs) as described previously (Sharma et al., 2008). Proteins were quantified spectrophotometrically at 280 nm.

Tryptophan fluorescence measurements

All the tryptophan fluorescence measurements were performed at 20 °C on FLS1000 Photoluminescence Spectrometer (Edinburgh Instruments). To investigate the kinetics of DM-MBP refolding, intrinsic tryptophan fluorescence measurements were performed on refolding DM-MBP. Fluorescence intensity was measured for 2 s at intervals of 60 s. 3 µM DM-MBP was denatured with 3 M GuHCl in buffer A (20 mM Tris, pH 7.5, 20 mM KCl) and allowed to refold after 75-fold dilution in buffer A. Intrinsic tryptophan fluorescence was excited at 290 nm with a slit width of 2 nm and detected at 345 nm with a 5 nm slit width. Photobleaching was avoided by adjusting the slit widths, measurement time intervals and acquisition times.

For both unfolding and refolding curves, steady-state tryptophan fluorescence was measured after 20 hr. For the unfolding curve, ~40 nM native DM-MBP was incubated at 22 °C in buffer A containing 0.2 M, 0.4 M, 0.6 M, 0.8 M, 1.2 M, 1.6 M and 2 M GuHCl. For the refolding curve, 2 µM DM-MBP was denatured in 3 M GuHCl/10 mM DTT at 50° C for 1 h in buffer A and incubated at 22 °C after 50-fold dilution in 0 M, 0.2 M, 0.4 M, 0.6 M, 0.8 M, 1.2 M, 1.6 M and 2 M GuHCl prepared in buffer A.

The Boltzmann function for unfolding and refolding titrations

The unfolding and refolding titrations from the tryptophan measurements were fit using a Boltzmann function:

$$y = \frac{A_1 - A_2}{1 + e^{(I - I_0)/\Delta I}} + A_2$$

where A_1 is the final and A_2 is the initial data point respectively, I is the center of the transition and ΔI is the increment between data points.

For equilibrium unfolding and refolding experiments with smFRET, an intermediate population was observed. In these cases, the titrations were fit using a double Boltzmann function given by:

$$y = y0 + \frac{f_1}{1 + e^{(I - I1_0)/\Delta I_1}} + \frac{(1 - f_1)}{1 + e^{(I - I2_0)/\Delta I_2}}$$

where f_1 is the amplitude of the first transition, $I1_0$ and $I2_0$ are the centers of the transitions and ΔI_1 and ΔI_2 are the respective increments between data points.

Fluorophore labelling of MBP

Cysteine-maleimide labelling with one and two fluorophores

All cysteine-maleimide couplings were performed according to the manufacturer's instructions (Atto-Tec) with a few modifications. The single-cysteine mutant, A52C was labelled either with Atto532- (Atto-Tec) or Alexa647-maleimide (Invitrogen). All double-cysteine mutants, A52C-P298C, K175C-P298C and A52C-K175C were stochastically labelled with Atto532- and Alexa647-maleimides. To analyze two-color controls of three-color FRET analysis, the double-cysteine mutant A52C-P298C was stochastically labelled with Atto488- and Atto565-maleimides, K175C-P298C with Atto565- and Alexa647-maleimides and, A52C-K175C with Atto488- and Alexa647-maleimides.

Briefly,the sulfhydryl groups of ~50 µM DM-MBP cysteine mutants were reduced with 10 mM DTT in phosphate buffered saline solution (PBS) at room temperature (RT) for 20 min. Excess DTT was removed by washing the protein in a 10 kDa cut-off Amicon centrifugal filter (Merck-Millipore) with de-oxygenated PBS containing 50 µM tris(2-carboxyethyl) phosphine (TCEP). Approximately a 3-fold molar excess of the maleimide-fluorophore conjugate was added to the washed protein solution and the reaction was carried out at RT for 3 hr in the dark. For double labelling, the fluorophores were labeled stochastically by adding an equimolar mixture of both maleimide-fluorophore conjugates simultaneously in the reaction. Unreacted maleimide-fluorophores were washed out with buffer A containing 1 mM DTT using a centrifugal filter. Successful labeled was verified using FCS. FCCS was used to confirm the double-labelling. The degree of labelling was quantified spectrophotometrically. Labeled of DM-MBP with the above mentioned cysteine-maleimide chemistries did not affect the folding rates (Table 1).

Specific labelling with three fluorophores

DM-MBP-A52PrK-K175C-P298C protein was specifically labelled with three fluorophores for the three-color smFRET experiments. In the first step, A52PrK with an alkyne group was specifically conjugated to the azide moiety of Atto488-azide (Atto-tec) via copper-catalyzed alkyne-azide cycloaddition, one type of click chemistry reaction (A. Chatterjee et al., 2012). ~120 µM of DM-MBP-A52PrK-K175C-P298C protein was allowed to react with 3-fold molar excess of Atto488-azide, in the presence of 200 µM $CuSO_4$, 50 µM TCEP, 200 µM TBTA and freshly prepared 200 µM sodium ascorbate in PBS at RT for ~3 hr, in the dark, under mild shaking conditions (Tyagi & Lemke, 2013). Unreacted dye was removed by washing with PBS using centrifugal filters.

K175C and P298C were labelled with Alexa647 and Atto565, respectively, using cysteine-maleimide chemistry. It has been shown that maltose binding in the inter-domain cleft of MBP buries some residues at the domain-interface including 298 position (Sharma et al., 2008) (Mapa et al., 2012a). Consistent with this observation, we labelled the Alexa647-maleimide specifically to the cysteine at 175 position by labeling in the presence of maltose, which blocks the competing cysteine of 298 position (Sharma et al., 2008), (Mapa et al., 2012a) (Jäger, Michalet, & Weiss, 2005). For the second step of labelling, we took ~70 µM of 52PrK-atto488 labelled DM-MBP-A52PrK-K175C-P298C, reduced with 10 mM DTT addition and washed with de-oxygenated PBS containing 500 mM Maltose. The cysteine-maleimide reaction was performed by the addition of 2-fold molar excess of Alexa647-maleimide and allowed to react for 1 hr at RT to minimize the possibility of mis-labelling the maltose blocked cysteine at position 298 (Supplementary Figure 9). As a third and final step, ~30 µM of 52PrK-Atto488-175C-Alexa647 labelled DM-MBP-A52PrK-K175C-P298C protein was washed with PBS to remove the excess of unlabeled Alexa647 dye and maltose. The washed protein was then labelled with 3-fold molar excess of Atto565-maleimide to the only available cysteine at position 298. Coupling of each dye after each labelling step was monitored by measuring the absorption of the respective fluorophores at their respective wavelengths of maximum absorption and of the protein at 280 nm. The overall degree of labeling for all three labels and was estimated to be ~15% by absorption spectroscopy. However, on the lower limit ~2 % triple-labeled molecules were finally used to analyze the FRET histograms after filtering for photobleaching and blinking events (Supplementary Figure 8B-C). Covalent attachment of the three-fluorophores did not have any significant influence on the folding (Figure 6B-D, Supplementary Figure 8F).

Single molecule measurements and data analysis

Two-color and three-color FRET measurements

All the FRET measurements were performed on custom -build confocal set-ups as described below. 50-100 pM of double- or triple-labelled DM-MBP proteins were measured to minimize the possibility of having more than one molecule in the confocal volume at a time. Before starting a smFRET measurement on the glass slide, the surface was passivated with 1 mg/ml BSA. SmFRET measurements of native MBP proteins were performed after serially diluting the labelled proteins in buffer A (20 mM Tris, pH 7.5, 20 mM KCl). To measure the FRET efficiency in the completely unfolded state, first ~500 nM labelled protein was denatured in buffer A containing 3 M GuHCl, 10 mM DTT at 50° C for 1 h and later was measured after serially diluting the protein concentration to picomolar 50-100 pM in 2 or 3 M GuHCl prepared in buffer A. SmFRET experiments performed in 6 M GuHCl showed no additional changes in the conformation of the denatured protein.

For all unfolding and refolding FRET measurements, 0.001% tween-20 was added to the buffer to prevent unfolded and refolded molecules from sticking to the surface. To measure the FRET efficiency under unfolding conditions, labelled native protein was diluted to 50-100 pM in buffer A containing either 0 M, 0.1 M,0.2 M,0.3 M,0.5 M,0.9 M or 1 M GuHCl. Refolding FRET measurements were performed by first denaturing 500-1000 nM protein in 3 M GuHCl/10 mM DTT at 50°C for 1 hr followed by serial dilutions in 3 M GuHCl and a final 50-fold dilution in buffer A/GuHCl to obtain 50-100 pM labeled protein in buffer with the desired GuHCl concentration.

Two-color setup

Fluorescence correlation and cross-correlation spectroscopy and for two-color smFRET measurements on DM-MBP proteins labelled with Atto532 and Alexa647 were performed on a home-built confocal microscope capable of multi-parameter fluorescence detection (MFD) combined with pulsed interleaved excitation (PIE, MFD-PIE). PIE was implemented using a 532-nm green laser (Toptica; PicoTA 530) to excite the donor and a 640-nm red laser (PicoQuant; LDH-D-C-640) for directly exciting the acceptor molecules. The lasers were synchronized to a repetition rate of ~26.7 MHz with a delay of ~18 ns between each pulse. Laser excitation powers were set to ~100 µW before the objective for both the lasers. A 60x water immersion objective (Nikon; Plan Apo IR 60x1.27 Water Immersion) was used to collect the emitted fluorescence and focused onto a 75µm diameter pinhole for confocal detection. To implement MFD, the fluorescence signal collected after the pinhole was separated into parallel and perpendicular polarized light by a polarizing beam-splitter (Thorlabs; PBS3) and then spectrally separated for green and red fluorescence by a dichroic mirror (AHF Analysetechnik; Dual Line z532/635,). Finally, before fluorescence detection using four single-photon-counting (SPC) avalanche photodiodes (APD) (Perkin-Elmer), both the green and red fluorescence spectra were cleaned with emission filters (green: Semrock, Bright line 582/75; red: Chroma, HQ700/75 M). SPC cards (Becker and Hickl; SPC 154) were synchronized with the laser drivers to record the arrival times of the photons.

Fluorescence Correlation and Cross-correlation Spectroscopy

Fluorescence correlation spectroscopy (FCS) and fluorescence cross-correlation spectroscopy (FCCS) experiments were performed with PIE capabilities to investigate aggregation of DM-MBP during refolding (Müller et al., 2005). For these experiments, a single-cysteine mutant of DM-MBP (A52C) labeled with either Atto532 maleimide or Alexa647 maleimide dye was used and measured with the two-color MFD-PIE setup as described above. FCCS detects the coincidence of fluctuations in both the green and red detection channels and hence sensitively detector whether oligomers have formed during. For these experiments, 500 nM Atto532-DM-MBP and 500 nM Alexa647-DM-MBP were denatured in 3 M GuHCl at 50°C for 1 hr. The samples were then mixed at RT and refolding was initiated by diluting the solution to a final labeled-protein concentration of 2 nM (1 nM of Atto532-DM-MBP and 1 nM of Alexa647-DM-MBP) in buffer A. FCCS experiments were performed during the initial 2 minutes after dillution and after 30 to monitor the time dependent oligomer formation over the entire refolding process. A 40 base-pair dsDNA labeled with both Atto532 and Atto647 on separate strands was used as a positive control. Free dyes were measured as a negative control (Figure 1D).

A 3D Gaussian confocal volume was assumed yielding the following analytical equation for the auto- and cross-correlation functions and was used to analyze the FCS and FCCS data:

$$G(\tau) = \frac{\gamma}{N}\frac{1}{\left(1+\frac{\tau}{\tau_D}\right)}\frac{1}{\sqrt{\left(1+\frac{\tau}{\tau_D}\frac{1}{\rho^2}\right)}} + y_0$$

where, $\gamma = 2^{-3/2}$ is the geometric factor, to correct for confocal shape, *N* is the average number of diffusing molecules in the probe volume with a diffusion time of $\tau_D = \frac{w_0^2}{4D}$, where D is the diffusion coefficient, ρ is the structure parameter defined as w_0/z_0, where w_0 and z_0 are the axial and radial dimensions from the center of the point-spread-function to the position where the intensity has decayed to $\frac{1}{e^2}$ and y_0 is a baseline to compensate for a potential offset in the correlation functions. For the FCCS analysis, the amplitudes of the green and red

autocorrelation functions can be used to determine the total number of diffusing particles containing a green label and a red label:

$$N_{GT} = N_G + N_{GR} \text{ and } N_{RT} = N_R + N_{GR}$$

where N_G and N_R represent the number of diffusing particles containing only a green and a red label respectively. The number of double-labeled molecules (N_{GR}) was determined from the amplitude of the cross-correlation function, which is given by:

$$G_{CC}(0) = \frac{\gamma \times N_{GR}}{N_{GT} \times N_{RT}}$$

Two-color MFD-PIE analysis

For two-color smFRET measurements, ~50-100 pM double labeled MBP sample was measured. PIE-MFD data analysis was performed to calculate correct two-color FRET efficiency, burst-wise lifetimes and anisotropy for single molecule events (Supplementary Figure 1) as described previously (Zander et al., 1996) (Kudryavtsev et al., 2012). Single molecule bursts in two-color FRET measurements were distinguished from the background with a use of photon burst search algorithm by applying a threshold of at least 5 photons for sliding time window of 500 µs and a total of 50 photons per burst (Nir et al., 2006a). The burst-wise fluorescence lifetime was estimated from the florescence decay by reconvolution with the instrument response function. The fluorescence anisotropy was fitted with respect to the fluorescence lifetime using Perrin equation as shown previously (Schaffer et al., 1999):

$$r = \frac{r_0}{1+\frac{\tau}{\rho}},$$

where, r is steady state anisotropy, r_0 is the fundamental anisotropy, ρ is rotational correlation time and τ is a fluorescence lifetime. Molecules labelled with both donor and acceptor dyes showing a stoichiometry of ~0.5 were selected for further analysis. An ALEX-2CDE filter with an upper value of 12 was used to filter out photobleaching and blinking events (Tomov et al., 2012). After correcting for background, crosstalk of green fluorescence in red channel (α), direct excitation of acceptor by donor excitation laser (δ), and differences in detection efficiencies and quantum yields of the dyes (γ) were estimated. The corrected labelling stoichiometry (S) and FRET efficiency for all the bursts was calculated as:

$$\left[S = \frac{\gamma F_{GG} + F_{GR} - \alpha F_{GG} - \delta F_{RR}}{\gamma F_{GG} + F_{GR} - \alpha F_{GG} - \delta F_{RR} + F_{RR}}\right]$$

$$\left[E = \frac{F_{GR} - \alpha F_{GG} - \delta F_{RR}}{\gamma F_{GG} + F_{GR} - \alpha F_{GG} - \delta F_{RR}}\right]$$

where, F_{GG} and F_{GR} are the fluorescence signal in donor and acceptor channel after donor excitation respectively and F_{RR} is the fluorescence signal in red channel after acceptor excitation. For the determined correction factors, see Table 1.

Ideally, when the conformational of the protein is static while the molecule transits the laser spot, the FRET efficiency (*E*) is related to the fluorescence lifetime of the donor in presence of an acceptor as:

$$\left[E = 1 - \frac{\tau_{D(A)}}{\tau_{D(0)}}\right]$$

where $\tau_{D(A)}$ is the fluorescence lifetime of donor in presence of an acceptor and $\tau_{D(0)}$ is the fluorescence lifetime of donor in absence of an acceptor. This relationship changes slightly when the inter-dye separation becomes comparable to the relative linker lengths, where linker flexibility dominates. An accurate relationship can be derived for a specific set of fluorophores as has been described previously (Gansen et al., 2009). For the dye pair Atto532-Alexa647, this relationship is given by the following third-order polynomial:

$$\left[E = 1 - \frac{-0.0178 + 0.6226\,\langle\tau_{D(A)}\rangle + 0.2188\,\langle\tau_{D(A)}\rangle^2 + 0.0312\,\langle\tau_{D(A)}\rangle^3}{\langle\tau_{D(0)}\rangle}\right]$$

When dynamics are present between two states with their respective donor fluorescence lifetimes τ_1 and τ_2, the relationship of intensity averaged FRET efficiency (*E*) to the donor lifetime changes to (Kalinin, Valeri, et al., 2010)Ö

$$E = 1 - \frac{\tau_1.\tau_2}{\tau_{D(0)}[\tau_1 + \tau_2 - \langle\tau\rangle]}$$

where, the average donor fluorescence lifetime $\langle\tau\rangle$ is calculated from the total florescence signal over a burst.

Fluorescence lifetime analysis

To determine the FRET efficiencies of the different FRET states undergo dynamic transitions, a fluorescence lifetime analysis was performed. The photons from all bursts were sum together and fit to a biexponential function convoluted with the instrument response function and a scattering component was included as an additional species in the fit. From the fluorescence lifetime, the FRET efficiency and distances were calculated using:

$$\left(E = 1 - \frac{\tau_{D(A)}}{\tau_{D(0)}}\right) \text{ and } \left(E = \frac{1}{1+\left(\frac{R}{R_0}\right)^6}\right)$$

where $\tau_{D(A)}$ is the fluorescence lifetime of donor in presence of an acceptor, $\tau_{D(0)}$ is the fluorescence lifetime of donor in absence of an acceptor (~3.6 ns for Atto532), *R* is the distance between the dyes and R_0 is the Förster radius (62 Å for Atto532-Alexa647 dye pair). The quality of the fit was evaluated using the reduced χ^2 value, χ^2_{red}.

Three-color setup

Both triply-labelled DM-MBP with Atto488-Atto565-Alexa647 and doubly-labelled DM-MBP with Atto488-Atto565, Atto565-Alexa647 and Atto488-Alexa647 dye pairs were measured on a three-color confocal single molecule setup equipped with MFD-PIE as previously described (Barth, Voith Von Voithenberg, & Lamb, 2019). Briefly, PIE experiments were performed with three pulsed lasers having ~20 ns delay between each pulse (PicoQuant, Germany; LDH-D-C-485, LDH-D-TA-560, LDH-D-C-640). The pulse frequency of 16.7 MHz and its synchronization was achieved using a laser driver (PicoQuant, Germany; Sepia II). A 60x water immersion objective with 1.27 N.A. (Nikon, Germany; Plan Apo IR 60x 1.27 WI) was

used to focus the lasers into the sample with a power measured before the objective of ~120 µW for blue, ~75 µW for green and ~35 µW for red laser. The emitted fluorescence was collected by the same objective and separated from the laser excitation using a polychroic mirror (AHF Analysentechnik; zt405/488/561/633, Germany) and passed through 50 µm pinhole for defining the confocal volume. Light coming through the pinhole was first separated for its parallel and perpendicular polarization with a polarizing beam splitter (Thorlabs, Germany). Afterwards, the light in each polarization channel was separated into blue, green and red spectral regions using two dichroic mirrors (AHF Analysentechnik; BS560, 640DCXR). The blue, green and red detection channels were spectrally clean using emission filters (AHF Analysentechnik; ET525/50, ET607/36, ET670/30) before the fluorescence was detected on APD's (LaserComponents, 2x COUNT-100B; Perkin Elmer, 4x SPCM-AQR14). The timing of the detected photons was synchronized with lasers pulses using a TCSPC module (PicoQuant; HydraHarp400).

Three-color MFD-PIE analysis

In a three-color FRET system, the blue dye acts as a donor (D), the green dye acts as a first acceptor (A_1) and the red dye as a second acceptor (A_2). Additionally, the green dye can also act as a donor for the red dye (Supplementary Figure 8A). Extending MFD-PIE to three-colors makes it possible to detect photons in green and red channels after blue excitation as well as in the green and red channels after green excitation. This enables one to calculate all three stoichiometries (S_{BG}, S_{BR} and S_{GR}) and FRET efficiencies (E_{BG}, E_{BR} and E_{GR}) for blue-green, blue-red and green-red dye-pairs respectively. The latter case of green-red is similar to the typical two-color MFD-PIE scheme.

For three-color FRET measurements, ~50-100 pM triple labelled DM-MBP was measured on a passivated glass surface. An all-photon burst search algorithm was used to detect the single molecule events from the background and required at least 30 photons per sliding window of 500 µs and a total of 100 photons per burst. A typical burst-wise MFD-PIE analysis inclusive of stoichiometry, FRET efficiency, fluorescence lifetime and anisotropy, was extended to three-color MFD-PIE on the selected bursts as explained previously (Supplementary Figure 8) (Barth et al., 2019a).

Briefly, the three stoichiometries (S_{BG}, S_{BR} and S_{GR}) were calculated as follows:

$$\left[S_{BG} = \frac{F_{BB} + F_{BG} + F_{BR}}{F_{BB} + F_{BG} + +F_{BR} + F_{GG} + +F_{GR}}\right]$$

$$\left[S_{BR} = \frac{F_{BB} + F_{BG} + F_{BR}}{F_{BB} + F_{BG} + +F_{BR} + F_{RR}}\right]$$

$$\left[S_{GR} = \frac{F_{GG} + F_{GR}}{F_{GG} + F_{GR} + F_{RR}}\right]$$

where F_{XY} represents the detected fluorescence signal in the *Y* channel after exciting the *X* dye. For the three-color analysis, triple labeled molecules were sorted by applying the ALEX-2CDE-filter for all the three stoichiometries with maximum value of 15 for both blue-green, blue-red and 20 for green-red dye-pairs. Typical values of S_{BG}, S_{BR} and S_{GR} for triple-labeled DM-MBP molecules with Ato488, Atto565 and Alexa647 dyes were found to be ~0.2, ~0.15 and ~0.5 respectively (Supplementary Figure 8B-C).

The corrected three FRET efficiencies for selected triple-labeled molecules were derived as detailed in Barth, Voith Von Voithenberg, & Lamb, 2019 and can be written as:

$$\left[E_{GR} = \frac{F_{GR} - \alpha_{GR}\, F_{GG} - \delta_{GR}\, F_{RR}}{\gamma_{GR}\, F_{GG} + F_{GR} - \alpha_{GR}\, F_{GG} - \delta_{GR}\, F_{RR}}\right]$$

$$\left[E_{BG} = \frac{F_{BG}^{cor.}}{\gamma_{BG}\, F_{BG}\,(1 - E_{GR}) + F_{BG}^{cor.}}\right]$$

$$\left[E_{BR} = \frac{F_{BR}^{cor.} - E_{GR}\left(\gamma_{GR}\, F_{BG}^{cor.} + F_{BR}^{cor.}\right)}{\gamma_{BR}\, F_{BB} + F_{BR}^{cor.} - E_{GR}\left(\gamma_{BR}\, F_{BB} + \gamma_{GR}\, F_{BG}^{cor.} + F_{BR}^{cor.}\right)}\right]$$

Where, the intermediate correction terms $F_{BG}^{cor.}$ and $F_{BR}^{cor.}$ are defined as:

$$[F_{BG}^{cor.} = F_{BG} - \alpha_{BG}\, F_{BB} - \delta_{BG}\, F_{GG}]$$

$$[F_{BR}^{cor.} = F_{BR} - \alpha_{BR}\, F_{BB} - \delta_{BR}\, F_{RR} - \alpha_{GR}(F_{BG} - \alpha_{BG}\, F_{BB}) - \delta_{BG}\, E_{GR}\,(1 - E_{GR})^{-1}\, F_{GG}]$$

The respective crosstalk, direct excitation and detection correction factors are depicted as α_{XY}, δ_{XY} and γ_{XY} for signal in channel *Y* after exciting the dye *X*.

Dynamic photon distribution analysis

The raw photon signal carries important information about the kinetic heterogeneity of the system. For the purpose of computing the interconversion rates between two states in a robust way, first the proximity ratio (*PR*) [($PR = F_{GR}/(F_{GR}+F_{GG})$] collected during the burst was sliced into 0.5, 1 and 1.5 ms time bins to capture the influence of the kinetics. A global analysis of all the three-time bins was performed to extract the rates after solving it analytically. To take care of the broadening due to photon detection noise, a constant width (σ) for a static state was assumed to scale with the inter-dye distance (R) as $\sigma = 0.07R$ (Kalinin, Sisamakis, et al., 2010). The states were defined using the donor fluorescence lifetimes of the double-labeled molecules (See Table 4). Additional states were incorporated to account for impurities and donor only mixing in a sample typically at low proximity ratios. The analysis was applied to extract the rates between the unfolded and folded states in DM-MBP during refolding in various denaturant concentrations and are reported in Table 5.

Data analysis software

All the fluorescence correlation, burst analysis, fluorescence lifetime and photon distribution analysis were performed with the open-source PIE Analysis with MATLAB (PAM) software, a custom-written software in MATLAB (The MathWorks) (Schrimpf, Barth, et al., 2018).

All-atom molecular dynamics simulations

To obtain atomistic insights into the structural changes of DM-MBP during folding, we performed all-atom molecular dynamics (MD) simulations on both wt- and DM-MBP. For simulations on wt-MBP, the crystal structure with Protein Data Bank (PDB) ID 1OMP (Sharff et al., 1992) was used. For DM-MBP, the amino acids at postion 8 and 283 were exchanged with glycine and aspartate (V8G, Y283D) respectively to create the DM-MBP structure in PyMOL/Chimera (The PyMOL Molecular Graphics system, version 2.0 Schrodinger)/ (Pettersen et al., 2004). The AMBER16 MD package with the ff14SB force field was used for the simulations (Case et al., 2017). TIP3P water model was used to solvate the MBP molecule in a box of octagonal geometry. Care was taken to exclude the vaporization effects at high temperature, as explained earlier (Walser, Mark, & Van Gunstere, 2000) . A distance of 3 nm between the polypeptide chain and the box wall was maintained not to have a boundary effects in simulations. The charge of the system was neutralized by the addition of sodium ions. A small excess of sodium chloride was added, corresponding to a concentration of ~4 mM.

Stability of the simulation box was verified at high temperatures at 400/450 K during the initial 500 ps of equilibration using 2 fs steps. MD runs were performed using the NPT ensemble with volume scaling enabled for both the thermostat and a barostat. A Nvidia GTX 1080 Ti GPU was used to carry out the simulations, which typically ran with an average of 50 ns per day. The unfolding trajectory was analyzed using various AmberTools (Roe & Cheatham, 2013). The progress of unfolding was evaluated based on the secondary structure assignment by the DSSP algorithm (Definition of Secondary Structure of Proteins) (Kabsch & Sander, 1983) over the span of the unfolding trajectory at time steps of 5 ns. Furthermore, the fraction of native contacts (Q(X)) was used to evaluate the overall unfolding in the protein (Best, Hummer, & Eaton, 2013).

We performed simulations at both 400 K and 450 K. Simulations at 400 K were run for 2 μs to check the stability of the system at high temperature. Later, simulations were continued for another 2 μs at 450 K. In addition, separate simulations were carried out at 450 K for 1 μs.

Accessible volume calculations

To model the FRET distances in the labeled DM-MBP for comparison with the experimentally determined distances for the various labeling positions (PDB ID: 1OMP), we performed geometric accessible volume (AV) calculations with the use of the FRET positioning and screening (FPS) software (Kalinin et al., 2012). The input parameters used for simulating the dye with the AV1 model were: 20 Å (dye linker-length), 4.5 Å (dye width), and 3.5 Å (dye radius).

Acknowledgements

We gratefully acknowledge the financial support of the SFB 1035 (German Research Foundation DFG, Sonderforschungsbereich 1035, Projektnummer 201302640, project A11 to D.C.L.) and of the Ludwigs-Maximillians-Universität München through the Center for NanoScience (CeNS) and LMUinnovativ BioImaging Network (BIN). We thank Prof. F. Ulrich Hartl from the Max Planck Institute of Biochemistry, Martinsreid, for kindly providing the GroEL,GroES proteins and for valuable discussions.

Author contributions

G.A. purified and labeled proteins, performed all experiments, analyzed single-molecule data and prepared figures. A.B. implemented the three-color smFRET and performed and analyzed molecular dynamics simulations. G.A. wrote the initial version of the manuscript and all authors contributed to the writing of the final version of the manuscript. D.C.L. designed the research.

REFERENCES

Agashe, V. R., Guha, S., Chang, H. C., Genevaux, P., Hayer-Hartl, M., Stemp, M., … 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*(2), 199–209. https://doi.org/10.1016/S0092-8674(04)00299-5

Andrews, B. T., Capraro, D. T., Sulkowska, J. I., Onuchic, J. N., & Jennings, P. A. (2013). Hysteresis as a marker for complex, overlapping landscapes in proteins. *Journal of Physical Chemistry Letters*, *4*(1), 180–188. https://doi.org/10.1021/jz301893w

Anfinsen. (1973). ' CIE : NCES Folding of Protein Chains. *Science*, *181*(4096), 223–230.

Antonik, M., Felekyan, S., Gaiduk, A., & Seidel, C. A. M. (2006). Separating structural heterogeneities from stochastic variations in fluorescence resonance energy transfer distributions via photon distribution analysis. *Journal of Physical Chemistry B*, *110*(13), 6970–6978. https://doi.org/10.1021/jp057257+

Arai, M., Iwakura, M., Matthews, C. R., & Bilsel, O. (2011). Microsecond subdomain folding in dihydrofolate reductase. *Journal of Molecular Biology*, *410*(2), 329–342. https://doi.org/10.1016/j.jmb.2011.04.057

Arviv, O., & Levy, Y. (2012). Folding of multidomain proteins: Biophysical consequences of tethering even in apparently independent folding. *Proteins: Structure, Function and Bioinformatics*, *80*(12), 2780–2798. https://doi.org/10.1002/prot.24161

Balchin, D., Hayer-Hartl, M., & Hartl, F. U. (2016, July 1). In vivo aspects of protein folding and quality control. *Science*. American Association for the Advancement of Science. https://doi.org/10.1126/science.aac4354

Bandyopadhyay, A., Saxena, K., Kasturia, N., Dalal, V., Bhatt, N., Rajkumar, A., … Chakraborty, K. (2012). Chemical chaperones assist intracellular folding to buffer mutational variations. *Nature Chemical Biology*, *8*(3), 238–245. https://doi.org/10.1038/nchembio.768

Barth, A., Voith Von Voithenberg, L., & Lamb, D. C. (2019a). Quantitative Single-Molecule Three-Color Förster Resonance Energy Transfer by Photon Distribution Analysis. *Journal of Physical Chemistry B*, *123*(32), 6901–6916. research-article. https://doi.org/10.1021/acs.jpcb.9b02967

Barth, A., Voith Von Voithenberg, L., & Lamb, D. C. (2019b). SI-Quantitative Single-Molecule Three-Color Förster Resonance Energy Transfer by Photon Distribution Analysis. *Journal of Physical Chemistry B*, *123*(32), 6901–6916. https://doi.org/10.1021/acs.jpcb.9b02967

Batey, S., Nickson, A. A., & Clarke, J. (2008). Studying the folding of multidomain proteins. *HFSP Journal*, *2*(6), 365–377. https://doi.org/10.2976/1.2991513

Bechtluft, P., van Leeuwen, R. G. H., Tyreman, M., Tomkiewicz, D., Nouwen, N., Tepper, H. L., … Tans, S. J. (2007). Direct observation of chaperone-induced changes in a protein folding pathway. *Science (New York, N.Y.)*, *318*(5855), 1458–1461. https://doi.org/10.1126/science.1144972

Bertz, M., & Rief, M. (2008). Mechanical Unfoldons as Building Blocks of Maltose-binding Protein. *Journal of Molecular Biology*, *378*(2), 447–458. https://doi.org/10.1016/j.jmb.2008.02.025

Best, R. B., Hummer, G., & Eaton, W. A. (2013). Native contacts determine protein folding mechanisms in atomistic simulations. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(44), 17874–17879. https://doi.org/10.1073/pnas.1311599110

Borgia, M. B., Borgia, A., Best, R. B., Steward, A., Nettels, D., Wunderlich, B., … Clarke, J. (2011). Single-molecule fluorescence reveals sequence-specific misfolding in multidomain proteins. *Nature*, *474*(7353), 662–665. https://doi.org/10.1038/nature10099

C. Zander, Sauer, M., Drexhage, K. H., Ko, D.-S., Schulz, A., Wolfrum, J., … Seide, C. A. M. (1996). Detection and characterization of single molecules in aqueous solution. *Applied Physics B*, *63*, 517–523.

Case, D. A., Cerutti, D. S., Cheatham, T. E. I., Darden, T. A., Duke, R. E., Giese, T. J., … Kollman, P. A. (2017). *Amber 2017*. *University of California, San Francisco*.

Cerminara, M., Schöne, A., Ritter, I., Gabba, M., & Fitter, J. (2020). Mapping Multiple Distances in a Multidomain Protein for the Identification of Folding Intermediates. *Biophysical Journal*, *118*(3), 688–697. https://doi.org/10.1016/j.bpj.2019.12.006

Chakraborty, K., Chatila, M., Sinha, J., Shi, Q., Poschner, B. C., Sikor, M., … Hayer-Hartl, M. (2010). Chaperonin-Catalyzed Rescue of Kinetically Trapped States in Protein Folding. *Cell*, *142*(1), 112–122. https://doi.org/10.1016/j.cell.2010.05.027

Chang, H. C., Kaiser, C. M., Hartl, F. U., & Barral, J. M. (2005). De novo folding of GFP fusion proteins: High efficiency in eukaryotes but not in bacteria. *Journal of Molecular Biology*, *353*(2), 397–409. https://doi.org/10.1016/j.jmb.2005.08.052

Chatterjee, A., Xiao, H., & Schultz, P. G. (2012). Evolution of multiple, mutually orthogonal prolyl-tRNA synthetase/tRNA pairs for unnatural amino acid mutagenesis in Escherichia coli. *Proceedings of the National Academy of Sciences*.

https://doi.org/10.1073/pnas.1212454109

Chothia, C., Gough, J., Vogel, C., & Teichmann, S. A. (2003). Evolution of the protein repertoire. *Science*, *300*(5626), 1701–1703. https://doi.org/10.1126/science.1085371

Chun, S. Y., Strobel, S., Bassford, P., & Randall, L. L. (1993). Folding of maltose-binding protein. Evidence for the identity of the rate- determining step in vivo and in vitro. *Journal of Biological Chemistry*, *268*(28), 20855–20862.

Chung, H. S., McHale, K., Louis, J. M., & Eaton, W. a. (2012). Single-Molecule Fluorescence Experiments Determine Protein Folding Transition Path Times. *Science*, *335*(6071), 981–984. https://doi.org/10.1126/science.1215768

Clamme, J.-P., & Deniz, A. A. (2005). Three-color single-molecule fluorescence resonance energy transfer. *Chemphyschem : A European Journal of Chemical Physics and Physical Chemistry*, *6*(1), 74–77. https://doi.org/10.1002/cphc.200400261

Collinet, B., Hervé, M., Pecorari, F., Minard, P., Eder, O., & Desmadril, M. (2000). Functionally accepted insertions of proteins within protein domains. *Journal of Biological Chemistry*, *275*(23), 17428–17433. https://doi.org/10.1074/jbc.M000666200

Dahiya, V., Agam, G., Lawatscheck, J., Rutz, D. A., Lamb, D. C., & Buchner, J. (2019). Coordinated Conformational Processing of the Tumor Suppressor Protein p53 by the Hsp70 and Hsp90 Chaperone Machineries. *Molecular Cell*, *74*(4), 816-830.e7. https://doi.org/10.1016/j.molcel.2019.03.026

Deniz, A. A., Laurence, T. A., Beligere, G. S., Dahan, M., Martin, A. B., Chemla, D. S., … Weiss, S. (2000). Single-molecule protein folding: Diffusion fluorescence resonance energy transfer studies of the denaturation of chymotrypsin inhibitor 2. *Proceedings of the National Academy of Sciences*, *97*(10), 5179–5184. https://doi.org/10.1073/pnas.090104997

Dinner, A. R., & Karplus, M. (1999). Is protein unfolding the reverse of protein folding? A lattice simulation analysis. *Journal of Molecular Biology*, *292*(2), 403–419. https://doi.org/10.1006/jmbi.1999.3051

Forster, T. (1946). Energiewanderung und Fluoreszenz. *Die Naturwissenschaften*, *33*(6), 166–175. https://doi.org/10.1007/BF00585226

Gambin, Y., & Deniz, A. A. (2010). Multicolor single-molecule FRET to explore protein folding and binding. *Molecular BioSystems*. https://doi.org/10.1039/c003024d

Gansen, A., Valeri, A., Hauger, F., Felekyan, S., Kalinin, S., Tóth, K., … Seidel, C. A. M. (2009). Nucleosome disassembly intermediates characterized by single-molecule FRET.

Proceedings of the National Academy of Sciences of the United States of America, *106*(36), 15308–15313. https://doi.org/10.1073/pnas.0903005106

Gupta, A. J., Haldar, S., Miličić, G., Hartl, F. U., & Hayer-Hartl, M. (2014). Active cage mechanism of chaperonin-assisted protein folding demonstrated at single-molecule level. *Journal of Molecular Biology*, *426*(15), 2739–2754. https://doi.org/10.1016/j.jmb.2014.04.018

Ha, T., Enderle, T., Ogletree, D. F., Chemla, D. S., Selvin, P. R., & Weiss, S. (1996). Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor. *Proceedings of the National Academy of Sciences*, *93*(13), 6264–6268. https://doi.org/10.1073/pnas.93.13.6264

Han, J.-H., Batey, S., Nickson, A. a, Teichmann, S. a, & Clarke, J. (2007). The folding and evolution of multidomain proteins. *Nature Reviews. Molecular Cell Biology*, *8*(4), 319–330. https://doi.org/10.1038/nrm2144

Hofmann, H., Hillger, F., Pfeil, S. H., Hoffmann, A., Streich, D., Haenni, D., … Schuler, B. (2010). Single-molecule spectroscopy of protein folding in a chaperonin cage. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(26), 11793–11798. https://doi.org/10.1073/pnas.1002356107

Hohng, S., Joo, C., & Ha, T. (2004). Single-molecule three-color FRET. *Biophysical Journal*, *87*(2), 1328–1337. https://doi.org/10.1529/biophysj.104.043935

Holtkamp, W., Kokic, G., Jäger, M., Mittelstaet, J., Komar, A. A., & Rodnina, M. V. (2015). Cotranslational protein folding on the ribosome monitored in real time. *Science*, *350*(6264), 1104–1107. https://doi.org/10.1126/science.aad0344

Huang, C., Rossi, P., Saio, T., & Kalodimos, C. G. (2016). Structural basis for the antifolding activity of a molecular chaperone. *Nature*, *537*(7619), 202–206. https://doi.org/10.1038/nature18965

Imamoglu, R., Balchin, D., Hayer-Hartl, M., & Hartl, F. U. (2020). Bacterial Hsp70 resolves misfolded states and accelerates productive folding of a multi-domain protein. *Nature Communications*, *11*(1), 365. https://doi.org/10.1038/s41467-019-14245-4

Inanami, T., Terada, T. P., & Sasai, M. (2014). Folding pathway of a multidomain protein depends on its topology of domain connectivity. *Proceedings of the National Academy of Sciences*, *111*(45), 15969–15974. https://doi.org/10.1073/pnas.1406244111

Jäger, M., Michalet, X., & Weiss, S. (2005). Protein-protein interactions as a tool for site-specific labeling of proteins. *Protein Science*, *14*(8), 2059–2068.

https://doi.org/10.1110/ps.051384705

Jones, S., Stewart, M., Michie, A., Swindells, M. B., Orengo, C., & Thornton, J. M. (1998). Domain assignment for protein structures using a consensus approach : Characterization and analysis.

Kalinin, S., Peulen, T., Sindbert, S., Rothwell, P. J., Berger, S., Restle, T., … Seidel, C. a M. (2012). A toolkit and benchmark study for FRET-restrained high-precision structural modeling. *Nature Methods*, *9*(12), 1218–1227. https://doi.org/10.1038/NMETH.2222

Kalinin, S., Sisamakis, E., Magennis, S. W., Felekyan, S., & Seidel, C. A. M. (2010). On the origin of broadening of single-molecule FRET efficiency distributions beyond shot noise limits. *Journal of Physical Chemistry B*, *114*(18), 6197–6206. https://doi.org/10.1021/jp100025v

Kalinin, S., Valeri, A., Antonik, M., Felekyan, S., & Seidel, C. A. M. (2010). Detection of Structural Dynamics by FRET: A Photon Distribution and Fluorescence Lifetime Analysis of Systems with Multiple States. *Journal of Physical Chemistry B*, *114*(23), 7983–7995. https://doi.org/10.1021/jp102156t

Kellner, R., Hofmann, H., Barducci, A., Wunderlich, B., Nettels, D., & Schuler, B. (2014). Single-molecule spectroscopy reveals chaperone-mediated expansion of substrate protein. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.1407086111

Kramers, H. A. (1940). *Brownian motion in a field of force and the diffusion model of chemical reactions*. *Physica* (Vol. 7). https://doi.org/10.1016/S0031-8914(40)90098-2

Kudryavtsev, V., Sikor, M., Kalinin, S., Mokranjac, D., Seidel, C. a M., & Lamb, D. C. (2012). Combining MFD and PIE for accurate single-pair f??rster resonance energy transfer measurements. *ChemPhysChem*, *13*(4), 1060–1078. https://doi.org/10.1002/cphc.201100822

Lazaridis, T., & Karplus, M. (1997). “New view” of protein folding reconciled with the old through multiple unfolding simulations. *Science*, *278*(5345), 1928–1931. https://doi.org/10.1126/science.278.5345.1928

Li, W., Terakawa, T., Wang, W., & Takada, S. (2012). Energy landscape and multiroute folding of topologically complex proteins adenylate kinase and 2ouf-knot. *Proceedings of the National Academy of Sciences*, *109*(44), 17789–17794. https://doi.org/10.1073/pnas.1201807109

Liu, K., Maciuba, K., & Kaiser, C. M. (2019). The Ribosome Cooperates with a Chaperone to

Guide Multi-domain Protein Folding. *Molecular Cell*, *74*(2), 310-319.e7. https://doi.org/10.1016/j.molcel.2019.01.043

Mapa, K., Tiwari, S., Kumar, V., Jayaraj, G. G., & Maiti, S. (2012a). Information encoded in Non-native states drives substrate-chaperone pairing. *Structure*, *20*(9), 1562–1573. https://doi.org/10.1016/j.str.2012.06.014

Mapa, K., Tiwari, S., Kumar, V., Jayaraj, G. G., & Maiti, S. (2012b). Information encoded in Non-native states drives substrate-chaperone pairing. *Structure*, *20*(9), 1562–1573. https://doi.org/10.1016/j.str.2012.06.014

Mashaghi, A., Bezrukavnikov, S., Minde, D. P., Wentink, A. S., Kityk, R., Zachmann-Brand, B., … Tans, S. J. (2016). Alternative modes of client binding enable functional plasticity of Hsp70. *Nature*, 1–16. https://doi.org/10.1038/nature20137

Mashaghi, A., Kramer, G., Bechtluft, P., Zachmann-Brand, B., Driessen, A. J. M., Bukau, B., & Tans, S. J. (2013). Reshaping of the conformational search of a protein by the chaperone trigger factor. *Nature*, *500*(7460), 98–101. https://doi.org/10.1038/nature12293

Moon, C. P., Kwon, S., & Fleming, K. G. (2011). Overcoming hysteresis to attain reversible equilibrium folding for outer membrane phospholipase A in phospholipid bilayers. *Journal of Molecular Biology*, *413*(2), 484–494. https://doi.org/10.1016/j.jmb.2011.08.041

Müller, B. K., Zaychikov, E., Bräuchle, C., & Lamb, D. C. (2005). Pulsed interleaved excitation. *Biophysical Journal*, *89*(5), 3508–3522. https://doi.org/10.1529/biophysj.105.064766

Nam, K. L., Kapanidis, A. N., Hye, R. K., Korlann, Y., Sam, O. H., Kim, Y., … Weiss, S. (2007). Three-color alternating-laser excitation of single molecules: Monitoring multiple interactions and distances. *Biophysical Journal*, *92*(1), 303–312. https://doi.org/10.1529/biophysj.106.093211

Nir, E., Michalet, X., Hamadani, K. M., Laurence, T. A., Neuhauser, D., Kovchegov, Y., & Weiss, S. (2006). Shot-noise limited single-molecule FRET histograms: Comparison between theory and experiments. *Journal of Physical Chemistry B*, *110*(44), 22103–22124. https://doi.org/10.1021/jp063483n

Onuchic, J. N., & Wolynes, P. G. (2004). Theory of protein folding. *Current Opinion in Structural Biology*. https://doi.org/10.1016/j.sbi.2004.01.009

Osváth, S., Köhler, G., Závodszky, P., & Fidy, J. (2005). Asymmetric effect of domain interactions on the kinetics of folding in yeast phosphoglycerate kinase. *Protein Science : A Publication of the Protein Society*, *14*(6), 1609–1616.

https://doi.org/10.1110/ps.051359905

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, *25*(13), 1605–1612. https://doi.org/10.1002/jcc.20084

Pirchi, M., Ziv, G., Riven, I., Cohen, S. S., Zohar, N., Barak, Y., & Haran, G. (2011). Single-molecule fluorescence spectroscopy maps the folding landscape of a large protein. *Nature Communications*, *2*, 493. https://doi.org/10.1038/ncomms1504

Raffy, S., Sassoon, N., Hofnung, M., & Betton, J. M. (1998). Tertiary structure-dependence of misfolding substitutions in loops of the maltose-binding protein. *Protein Science : A Publication of the Protein Society*, *7*(10), 2136–2142. https://doi.org/10.1002/pro.5560071010

Ratzke, C., Hellenkamp, B., & Hugel, T. (2014). Four-colour FRET reveals directionality in the Hsp90 multicomponent machinery. *Nature Communications*, *5*(1), 4192. https://doi.org/10.1038/ncomms5192

Roe, D. R., & Cheatham, T. E. (2013). PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. https://doi.org/10.1021/ct400341p

Schaffer, J., Volkmer, A., Eggeling, C., Subramaniam, V., Striker, G., & Seidel, C. A. M. (1999). Identification of Single Molecules in Aqueous Solution by Time-Resolved Fluorescence Anisotropy. *Journal of Physical Chemistry A*, *103*(3). https://doi.org/10.1021/jp9833597

Schrimpf, W., Barth, A., Hendrix, J., & Lamb, D. C. (2018). PAM: A Framework for Integrated Analysis of Imaging, Single-Molecule, and Ensemble Fluorescence Data. *Biophysical Journal*, *114*(7), 1518–1528. https://doi.org/10.1016/j.bpj.2018.02.035

Schuler, B., Lipman, E. A., & Eaton, W. A. (2002). Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy. *Nature*, *419*(6908), 743–747. https://doi.org/10.1038/nature01060

Sela, M., White, F. H., & Anfinsen, C. B. (1957). Reductive cleavage of disulfide bridges in ribonuclease. *Science*, *125*(3250), 691–692. https://doi.org/10.1126/science.125.3250.691

Selmke, B., Borbat, P. P., Nickolaus, C., Varadarajan, R., Freed, J. H., & Trommer, W. E. (2018). Open and Closed Form of Maltose Binding Protein in Its Native and Molten Globule State as Studied by EPR Spectroscopy. *Biochemistry*, acs.biochem.8b00322. https://doi.org/10.1021/acs.biochem.8b00322

Shank, E. A., Cecconi, C., Dill, J. W., Marqusee, S., & Bustamante, C. (2010). The folding cooperativity of a protein is controlled by its chain topology. *Nature*, *465*(7298), 637–640. https://doi.org/10.1038/nature09021

Sharff, A. J., Rodseth, L. E., Spurlino, J. C., & 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*(44), 10657–10663. https://doi.org/10.1021/bi00159a003

Sharma, S., Chakraborty, K., Müller, B. K., Astola, N., Tang, Y. C., Lamb, D. C., … Hartl, F. U. (2008). Monitoring Protein Conformation along the Pathway of Chaperonin-Assisted Folding. *Cell*, *133*(1), 142–153. https://doi.org/10.1016/j.cell.2008.01.048

Soranno, A., Buchli, B., Nettels, D., Cheng, R. R., Müller-Späth, S., Pfeil, S. H., … Schuler, B. (2012). Quantifying internal friction in unfolded and intrinsically disordered proteins with single-molecule spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(44), 17800–17806. https://doi.org/10.1073/pnas.1117368109

Soranno, A., Holla, A., Dingfelder, F., Nettels, D., Makarov, D. E., & Schuler, B. (2017). Integrated view of internal friction in unfolded proteins from single-molecule FRET, contact quenching, theory, and simulations. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(10), E1833–E1839. https://doi.org/10.1073/pnas.1616672114

Sparrer, H, Rutkat, K., & Buchner, J. (1997). Catalysis of protein folding by symmetric chaperone complexes. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(4), 1096–1100. https://doi.org/10.1073/pnas.94.4.1096

Sparrer, Helmut, Lilie, H., & Buchner, J. (1996). Dynamics of the GroEL – Protein Complex : Effects of Nucleotides and Folding Mutants, 74–87.

Spurlino, J. C., Lu, G. Y., & 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. *Journal of Biological Chemistry*, *266*(8), 5202–5219. https://doi.org/2002054

Tang, Y. C., Chang, H. C., Roeben, A., Wischnewski, D., Wischnewski, N., Kerner, M. J., … Hayer-Hartl, M. (2006). Structural Features of the GroEL-GroES Nano-Cage Required for Rapid Folding of Encapsulated Protein. *Cell*, *125*(5), 903–914. https://doi.org/10.1016/j.cell.2006.04.027

Teichmann, S. A., Parkt, J., & Chothia, C. (1998). Structural assignments to the Mycoplasma

genitalium proteins show extensive gene duplications and domain rearrangements. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(25), 14658–14663. https://doi.org/10.1073/pnas.95.25.14658

Thirumalai, D., O'Brien, E. P., Morrison, G., & Hyeon, C. (2010). Theoretical Perspectives on Protein Folding. https://doi.org/10.1146/annurev-biophys-051309-103835

Tomov, T. E., Tsukanov, R., Masoud, R., Liber, M., Plavner, N., & Nir, E. (2012). Disentangling subpopulations in single-molecule FRET and ALEX experiments with photon distribution analysis. *Biophysical Journal*, *102*(5), 1163–1173. https://doi.org/10.1016/j.bpj.2011.11.4025

Toofanny, R. D., & Daggett, V. (2012). Understanding protein unfolding from molecular simulations. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, *2*(3), 405–423. https://doi.org/10.1002/wcms.1088

Tyagi, S., & Lemke, E. A. (2013). Genetically Encoded Click Chemistry for Single-Molecule FRET of Proteins. *Methods in Cell Biology*. https://doi.org/10.1016/B978-0-12-407239-8.00009-4

Vogel, C., Bashton, M., Kerrison, N. D., Chothia, C., & Teichmann, S. A. (2004). Structure, function and evolution of multidomain proteins. *Current Opinion in Structural Biology*, *14*(2), 208–216. https://doi.org/10.1016/j.sbi.2004.03.011

Voith von Voithenberg, L., Sánchez-Rico, C., Kang, H. S., Madl, T., Zanier, K., Barth, A., … Lamb, D. C. (2016). Recognition of the 3′ splice site RNA by the U2AF heterodimer involves a dynamic population shift. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(46), E7169–E7175. https://doi.org/10.1073/pnas.1605873113

Walser, R., Mark, A. E., & Van Gunstere, W. F. (2000). On the temperature and pressure dependence of a range of properties of a type of water model commonly used in high-temperature protein unfolding simulations. *Biophysical Journal*, *78*(6), 2752–2760. https://doi.org/10.1016/S0006-3495(00)76820-2

Walters, B. T., Mayne, L., Hinshaw, J. R., Sosnick, T. R., & Englander, S. W. (2013). Folding of a large protein at high structural resolution. *Proceedings of the National Academy of Sciences*, *110*(47), 18898–18903. https://doi.org/10.1073/pnas.1319482110

Wang, J. D., Michelitsch, M. D., & Weissman, J. S. (1998). GroEL-GroES-mediated protein folding requires an intact central cavity. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(21), 12163–12168. https://doi.org/10.1073/pnas.95.21.12163

Wang, Y., Chu, X., Suo, Z., Wang, E., & Wang, J. (2012). Multidomain protein solves the folding problem by multifunnel combined landscape: Theoretical investigation of a Y-family DNA polymerase. *Journal of the American Chemical Society*, *134*(33), 13755–13764. https://doi.org/10.1021/ja3045663

Wetlaufer, D. B. (1973). Nucleation, Rapid Folding, and Globular Intrachain Regions in Proteins. *Proceedings of the National Academy of Sciences*, *70*(3), 697–701. https://doi.org/10.1073/pnas.70.3.697

Wolynes, P. G., Onuchic, J. N., & Thirumalai, D. (1995). Navigating the Folding Routes. *Science*, *267*(March).

Ye, X., Mayne, L., Kan, Z., & Englander, S. W. (2018). Folding of maltose binding protein outside of and in GroEL. *Proceedings of the National Academy of Sciences*, (33), 201716168. https://doi.org/10.1073/pnas.1716168115

Table 1: Refolding rates and half-life ($t_{1/2}$) of refolding for various DM-MBP mutants. All the refolding measurements used for measuring the refolding rates were carried out in GuHCl concentrations below 0.1 M and the data was fitted with a mono-exponential function.

	Refolding rate (min^{-1})	**$t_{1/2}$ of refolding (min)**
DM-MBP[a]	0.030	22.6
NTD (52C-298C) Atto532-Alexa647[b]	0.024	29.5
NTD (52C-298C) Atto532-Atto647N[b]	0.031	22.2
NTD (52C-298C) Atto488-Atto565 [b]	0.032	21.5
CTD (175C-298C) Atto532-Alexa647 [b]	0.029	24.9
CTD (175C-298C) Atto565-Alexa647**[c]	0.019	35.4
N-C interface (52C-175C) Atto532-Alexa647*[c]	0.025	27.2
N-C interface (52C-175C) Atto488-Alexa647***[b]	0.050	13.8
DM-MBP (52PrK-175C-298C) Atto488-Atto565-Alexa647*[c]	0.031	22.3

[a] Rate was estimated from tryptophan fluorescence measurements.

[b] Rate was estimated from FRET histogram by calculating the increase in the folded fraction relative to the folded fraction for refolded protein.

[c] Rate was estimated by calculating the increase in the fraction of low acceptor lifetime relative to refolded protein (Alexa647 as an acceptor). See Table S1 and Table S2.

*FRET histograms could not be used for these constructs to determine the refolding rate because of low contrast between the intermediate and folded state due to either the used dye pairs (Förster distance of 70 Å in the case of 175C-298C Atto565-Alexa647), the close proximity of the attached dyes (in the case of N-C interface 52C-175C) or the broad FRET histogram distributions measured in three-color smFRET (in the case of DM-MBP 52PrK-175C-298C).

** Although the FRET histogram for the refolded protein looked identical to the native protein (Supplementary Figure 8B), estimated rate from acceptor lifetime is lower compared to the expected rate. We attribute this effect to the different local orientation of the attached Alexa647 fluorophore at one of the positions given the stochastic labeling employed. Most likely, at position 298, Alexa647 has a different environment. At position 175, Alexa647 behaves as expected for a folded protein as in the case of three-color labeled MBP. See Table S2.

***Because of the low contrast between the intermediate and folded state, specific for the 52C-175C construct, the rate is overestimated in this case. See Right panel of Supplementary Figure 10C.

Table 2: Förster distance and correction factors used for various combinations of the dye-pairs in this study. *Indicates the corrected γ used in the case of Alexa647 lifetime changes, see Supplementary Figure 3G.

	Förster distance, R_0 (Å)	Detection correction factor, γ	Spectral cross-talk, α	Direct acceptor excitation, δ
Atto532-Aexa647 (Two color)	62	0.50/0.37*	0.03	0.07
Atto532-Atto647N (Two color)	59	0.59	0.02	0.06
Atto488-Atto565 (Three-color /Two-color)	63	0.40/0.44	0.07	0.05
Atto488-Alexa647 (Three-color /Two-color)	53	0.20/0.25	0.01	0.01
Atto565-Alexa647 (Three-color /Two-color)	70	0.43/0.50	0.14	0.13

Table 3: Comparison of distances from accessible volume (AV) simulations and two-color FRET experiments performed with all the three double cysteine domain mutants of DM-MBP and labeled with the Atto532-Alexa647 dye-pair. Distances derived from AV simulations are the FRET averaged distances ($<R_{DA}>_E$) and extracted using the FPS software as explained in the materials and methods (Kalinin et al., 2012). Experimental distances were estimated using the photon distribution analysis (PDA) (Antonik et al., 2006). PDA fitting was applied on the *PR* histogram obtained from burst-wise binned raw photon counts. The Monto Carlo method was used to simulate the Gaussian distance distributions for the D-A separation with one or two populations in the *PR* histogram. d1 and d2 are the two derived inter-dye distances for the peaks of the two populations with their respective fractions $f_1$and f_2 and widths (standard deviations) σ_1 and σ_2.

	A-V (Å)	**Experimental**					
		d1(Å)	**σ_1(Å)**	**f_1**	**d2(Å)**	**σ_2(Å)**	**f_2**
Native-NTD (52C-298C)	51.3	48.3	3.45	0.87	55.4	10	0.13
Native-CTD (175C-298C)	54.5	49.6	3.6	0.80	57.6	10	0.20
Native-N-C interface (52C-175C)	40.0	44.8	2.8	1.00	-		-
Denatured-NTD (52C-298C)	-	-	-	-	83.8	17	1.00
Denatured-CTD (175C-298C)	-	-	-	-	80.9	13	1.00
Denatured-N-C interface (52C-175C)	-	49.1	4.5	0.07	74.9	5.0	0.93

Table 4: Donor (Atto532) fluorescence lifetimes (τ_1 and τ_2), their fractions (f_1 and f_2), the corresponding calculated FRET efficiencies (E_1 and E_2) and the distance estimations (*d1* and *d2*) from a biexponential fit of the fluorescence decay for different double cysteine mutants of DM-MBP labeled with Atto532-Alexa647 (see Material and Methods). For comparison, the lifetime fractions from native state and under denaturing condition are given. Refolding was performed at the mentioned concentrations of GuHCl. All fits are shown as Supplementary figure 5.

NTD (52C-298C)		τ_1 (ns)	f_1	E_1	d1 (Å)	τ_2 (ns)	f_2	E_2	d2 (Å)	χ^2_{red}
Native		0.55	0.84	0.84	46.6	2.06	0.16	0.42	65.2	1.50
Refolding in GuHCl (M)	0.1	0.58	0.79	0.83	47.2	2.14	0.21	0.40	66.1	1.61
	0.2	0.81	0.51	0.77	50.5	2.79	0.49	0.22	76.2	1.58
	0.3	0.51	0.58	0.85	45.9	2.59	0.42	0.28	72.5	1.20
	0.5	0.68	0.30	0.80	48.6	2.97	0.70	0.17	80.3	1.03
	0.9	0.47	0.38	0.86	45.3	2.92	0.62	0.18	79.2	1.07
	2	0.85	0.22	0.76	51.0	3.24	0.78	0.09	89.5	1.60
CTD (175C-298C)		τ_1 (ns)	f_1	E_1	d1 (Å)	τ_2 (ns)	f_2	E_2	d2 (Å)	
Native		0.53	0.76	0.85	46.3	1.95	0.24	0.45	63.7	1.38
Refolding in GuHCl (M)	0.1	0.47	0.66	0.86	45.2	2.36	0.34	0.34	69.1	2.03
	0.2	0.53	0.64	0.85	46.3	2.90	0.36	0.19	78.7	1.75
	0.3	0.63	0.61	0.82	47.9	3.01	0.39	0.16	81.4	1.65
	0.5	0.58	0.51	0.83	47.1	3.18	0.49	0.11	86.9	1.94
	0.9	0.56	0.45	0.84	46.8	3.15	0.55	0.12	85.8	1.95
	2	0.91	0.24	0.74	51.8	3.14	0.76	0.12	85.7	1.68
N-C (52C-175C)		τ_1 (ns)	f_1	E_1	d1 (Å)	τ_2 (ns)	f_2	E_2	d2 (Å)	
Native		0.13	0.92	0.96	36.1	2.29	0.08	0.36	68.1	1.21
Refolding	0.1	0.49	0.74	0.86	45.7	2.38	0.26	0.33	69.3	1.63

in GuHCl (M)	0.2	0.62	0.69	0.82	47.7	2.62	0.31	0.27	73.0	1.85
	0.3	0.54	0.59	0.84	46.5	2.61	0.41	0.27	72.9	1.74
	0.5	0.80	0.56	0.77	50.3	2.58	0.44	0.28	72.4	1.62
	0.9	0.35	0.59	0.90	42.8	2.49	0.41	0.30	71.0	1.44
	2	0.65	0.23	0.81	48.3	2.95	0.77	0.17	79.9	1.71

Table 5: Results from the dynamic photon distribution analysis for all the three two-color DM-MBP constructs labeled with Atto532-Alexa647. Data for refolding measurements performed in 0.2, 0.3, 0.5, and 0.9 M GuHCl were analyzed. By fixing the two inter-dye distances for the folded and unfolded states ($\boldsymbol{R_F}$ and $\boldsymbol{R_U}$ respectively) estimated from the lifetime analysis (Table 4), the rates for interconversion ($\boldsymbol{k_{U->F}}$ and $\boldsymbol{k_{F->U}}$) between the states were extracted. Widths were kept at a constant fraction of the distance ***R*** (0.07 ***R***) (Kalinin, Sisamakis, et al., 2010). The relaxation time, $\boldsymbol{\tau_R}$, was calculated as 1/($\boldsymbol{k_{U->F}}$ + $\boldsymbol{k_{F->U}}$). See Material and Methods for a description of the analysis. * In the marked cases, distances used in the dynamic PDA analysis were extracted from the FRET efficiencies measured in the intermediate GuHCl titrations or from the native state because the fluorescence lifetime values at very high and very low FRET efficiencies can be difficult to determine correctly (e.g. due to donor and acceptor quenching and/or donor only/acceptor blinking events). Errors are the 95% confidence intervals obtained from the Jacobian fit. See Supplementary Figure 7.

NTD (52C-298C)		$\boldsymbol{R_F}$ **(Å)**	$\boldsymbol{k_{F\to U}}$ **(ms^{-1})**	$\boldsymbol{R_U}$ **(Å)**	$\boldsymbol{k_{U\to F}}$ **(ms^{-1})**	$\boldsymbol{\tau_R}$ **(μs)**
Refolding in GuHCl (M)	0.1	47.2	0.12±0.03	66.1	0.72±0.19	1190±287
	0.2	45.9*	1.03±0.10	76.2	3.23±0.27	234±15
	0.3	45.9	1.38±0.12	72.5	2.13±0.22	284±19
	0.5	48.6	2.08±0.21	80.3	0.89±0.10	336±26
	0.9	45.3	1.91±0.17	79.2	0.63±0.07	393±28
CTD (175C-298C)		$\boldsymbol{R_F}$ **(Å)**	$\boldsymbol{k_{F\to U}}$ **(ms^{-1})**	$\boldsymbol{R_U}$ **(Å)**	$\boldsymbol{k_{U\to F}}$ **(ms^{-1})**	$\boldsymbol{\tau_R}$ **(μs)**
Refolding	0.1	45.2	1.19±0.09	69.1	2.19±0.17	295±16

in GuHCl (M)	0.2	46.3	1.02±0.11	78.7	1.98±0.37	333±43
	0.3	47.9	1.17±0.11	81.4	2.76±0.22	254±15
	0.5	47.1	2.31±0.22	81.4*	3.00±0.28	188±12
	0.9	46.8	3.95±0.29	81.4*	1.60±0.12	180±10
N-C (52C-175C)		R_F (Å)	$k_{F\rightarrow U}$ (ms^{-1})	R_U (Å)	$k_{U\rightarrow F}$ (ms^{-1})	τ_R (μs)
Refolding in GuHCl (M)	0.1	44.8*	0.13±0.04	69.3	0.62±0.16	1333±307
	0.2	44.8*	0.39±0.05	73.0	3.34±0.59	268±43
	0.3	46.5	0.77±0.07	72.9	3.85±0.33	216±16
	0.5	44.8*	0.55±0.08	72.4	3.75±0.49	232±27
	0.9	46.5*	1.70±0.23	71.0	3.51±0.55	191±22

Main Figures

Figure 1

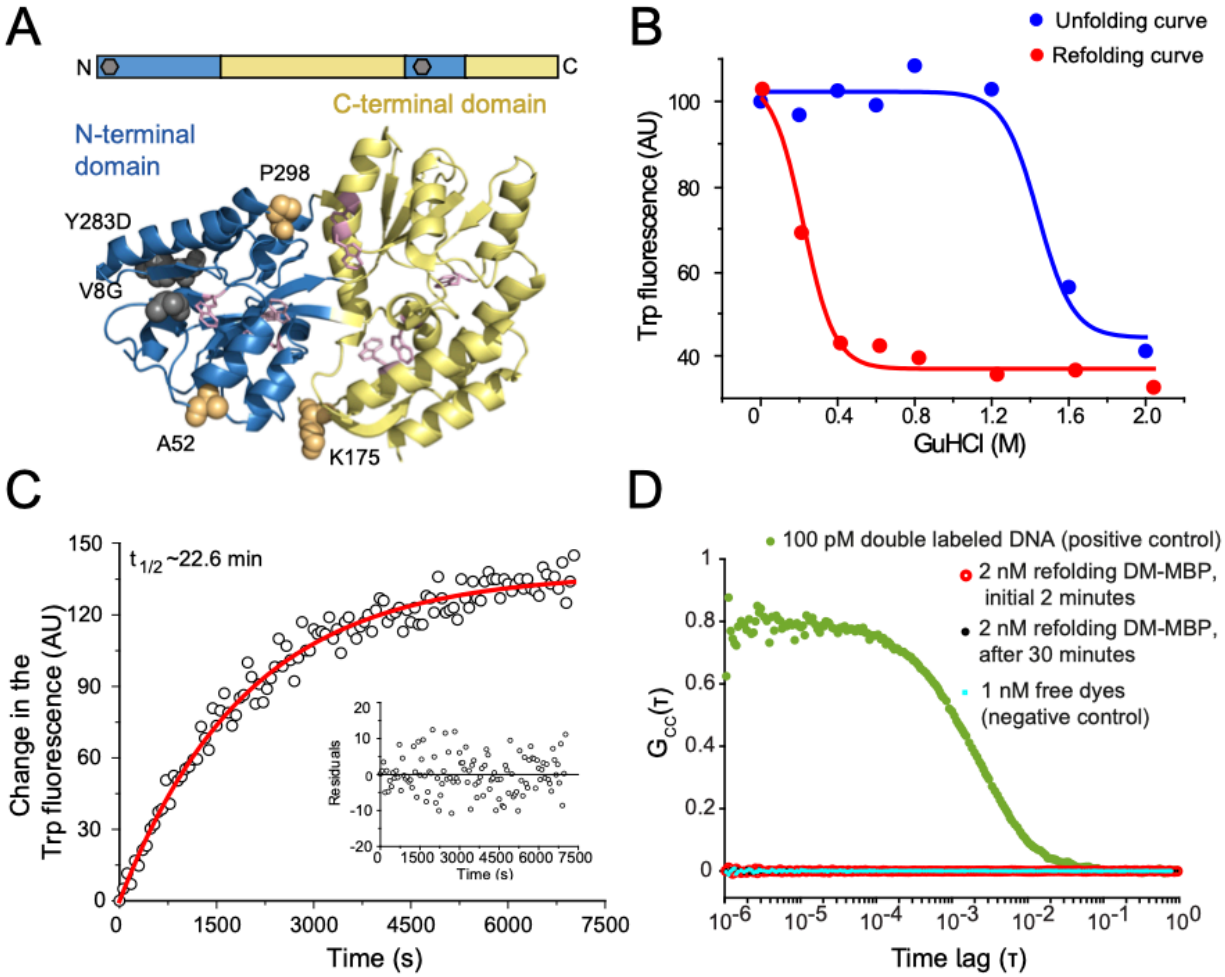

Figure 1. DM-MBP unfolding / refolding

(A) A ribbon structure of the maltose binding protein (MBP) (PDB ID:1OMP) showing the N- (NTD) and C-terminal domain (CTD) in blue and yellow respectively. The upper schematic represents the discontinuity in MBP sequence for the NTD and CTD where the positions of the mutations V8G and Y283D are depicted as grey hexagons. The residues involved in the two folding mutations and the three labeling positions A52, K175, P298 for coupling the fluorescent dyes are indicated via a space filling model in dark grey and orange respectively. Eight tryptophan side-chains are highlighted as a stick model in pink.

(**B)** Equilibrium unfolding and refolding experiments in different GuHCl concentrations of DM-MBP measured with tryptophan fluorescence. To record the unfolding curve (blue), steady state tryptophan fluorescence of ~40 nM native DM-MBP was measured after 20 hr in the respective GuHCl concentration at 22°C. For the refolding curve (red), 2 µM of DM-MBP was first denatured in 3 M GuHCl for 1 hr at 50°C, diluted 50-fold and incubated for 3-4 hrs at the indicated final GuHCl concentrations for before measuring the steady state tryptophan fluorescence. Data points from a single titration are shown although the data is representative of two independent titrations. Each curve was fitted with a Boltzmann function (see Materials and Methods).

(**C)** The kinetics of DM-MBP refolding monitored by the increase in tryptophan fluorescence. The initial fluorescence at time t=0 was subtracted from the subsequent data points. 3 µM DM-MBP was denatured in 3 M GuHCl for 1 hr at 50°C before being diluted 75-fold in Buffer A to start the refolding reaction (at t=0, the final concentrations were ~40 nM of DM-MBP and 40 mM of GuHCl). Data were fitted with a single exponential function yielding a R^2 value of 0.979. The residuals of the fit are shown in the inset. The presented curve is a single measurement representative of two independent measurements. The half-life ($t_{1/2}$) for the refolding process is 22.6 min.

(**D)** Fluorescence cross-correlation spectroscopy (FCCS) measurements of a mixture of 1 nM Atto532 labeled DM-MBP (A52C) and 1 nM Alexa647 labeled DM-MBP undergoing refolding during the initial 2 min and after 30 min (red and black curves, respectively). 500-1000 nM DM-MBP (A52C) was first denatured in 3 M GuHCl. The sample was then diluted first serially in 3 M GuHCl before the final dilution to achieve the nanomolar concentrations of the labeled protein. The FCCS amplitude was used to analyze the presence of small oligomers in the sample. Atto532 and Atto647N labeled DNA served as a positive control for a cross-correlation signal (green curve). Freely diffusing Atto532 and Atto655 dyes were measured as a negative control with no cross-correlation amplitude visible (cyan curve).

Figure 2

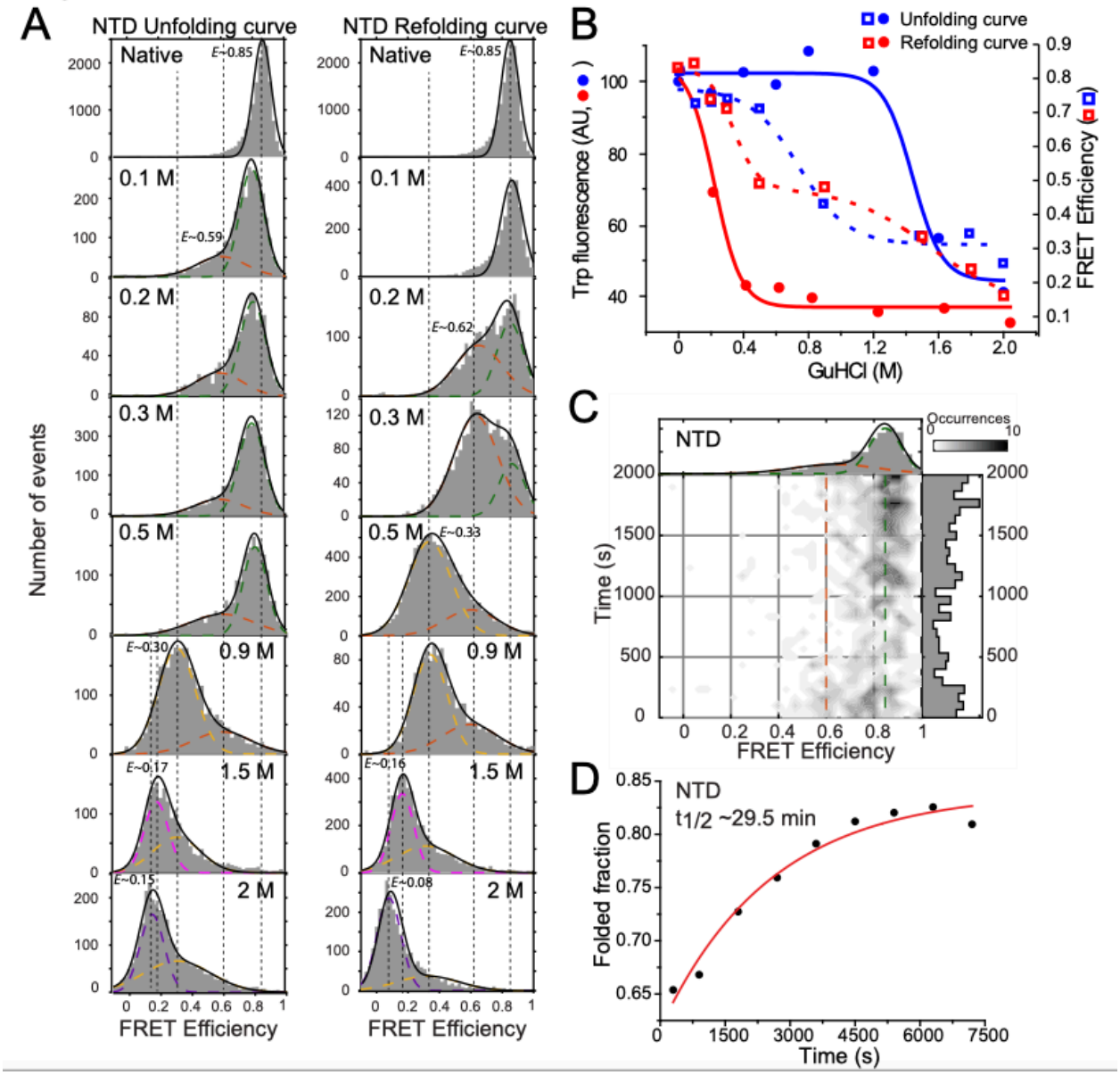

Figure 2. Equilibrium unfolding-refolding two-color smFRET measurements on the NTD reveals a refolding intermediate population

(A) SmFRET histograms and corresponding Gaussian fits for DM-MBP unfolding and refolding titrations at the indicated denaturant concentrations. Each underlying population is highlighted with a dotted line. The native or folded state is shown in green, intermediate states are shown in orange and in yellow, and the completely unfolded state is shown in violet (2 M GuHCl). The equilibrium unfolding curve is obtained by performing two-color smFRET measurements on native DM-MBP (52-298) in buffer containing different GuHCl concentrations. To measure the refolding curve, 500-1000 nM DM-MBP (52-298) was first denatured in 3 M GuHCl for 1 hr at 50°C. The sample was then diluted first serially in 3 M GuHCl at room temperature before the final 50- fold dilution to achieve the picomolar concentrations of labeled protein in buffer or buffer/GuHCl mixture at the indicated final GuHCl concentrations and refolding commences immediately. Protein refolds completely within an hour when refolding was performed in 0.1 M GuHCl. The histogram does not change in the other measurements throughout the course of the measurement times of 3-4 hr.

(B) To dissect out the heterogeneity in the ensemble curves measured in Figure 1B, average FRET efficiencies obtained from (A) were compared to the equilibrium unfolding and refolding curve measured with tryptophan fluorescence at the same denaturant titrations. The tryptophan curves (circles) and average FRET efficiencies (squares) for unfolding and refolding titrations are depicted in blue and red respectively. Ensemble curves are taken from Figure (1B). Average FRET efficiency refolding and unfolding curves were fitted with a single and double Boltzmann function respectively. The unfolding curve is in line with a two-state model for a single transition. On the other hand, refolding has two transitions, first a collapse and then a slow folding to the native state.

(C) Kinetics of NTD refolding was analyzed for the initial 2000 s. Refolding was induced after 30-fold dilution of 3 M GuHCl denatured NTD. A Gaussian fit to the FRET histogram shows the presence of an intermediate population with a FRET efficiency of 0.6 in addition to the folded fraction at 0.85 FRET efficiency.

(D) The refolding rate [half-life ($t_{1/2}$)] for the reaction in **(C)** was calculated by assaying the increase in the folded fraction (0.85 FRET efficiency) as a function of time. The red curve is a mono-exponential fit to the data (shown in black). Also see Table 1.

Figure 3

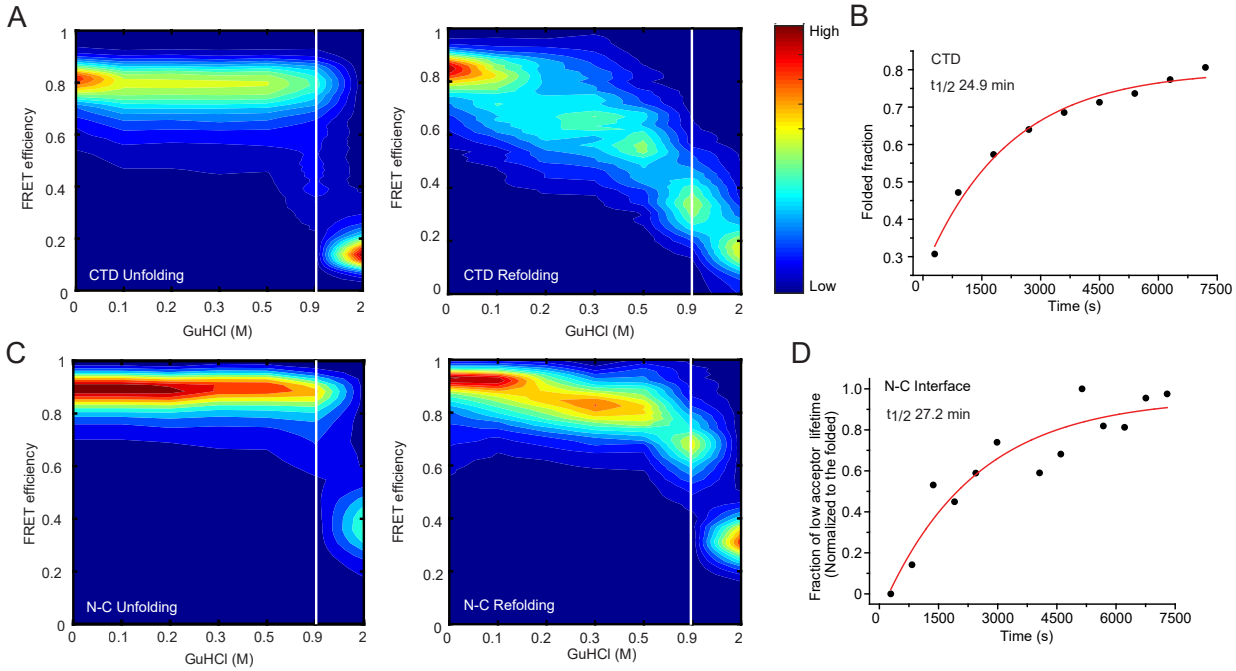

Figure 3. Evidence for an intermediate state under kinetic and equilibrium refolding conditions

(A-B) Equilibrium and kinetic analyses of CTD unfolding and refolding. **(A)** 2D-plots of FRET efficiency versus denaturant concentration are presented in a waterfall scheme for equilibrium unfolding (left) and refolding (right) measurements. For denaturant concentrations between ~ 0.2 and 0.5 GuHCl, an intermediate state with a FRET efficiency of 0.6 is visible. **(B)** A kinetic analysis of FRET histograms for CTD refolding. Refolding was induced after 30-fold dilution of 3 M GuHCl denatured protein. The red curve is a mono-exponential fit to the data (black). The half-life ($t_{1/2}$) for the refolding process is 24.9 min (see Table 1).

(C-D) Equilibrium and kinetic analyses of N-C interface. **(C)** 2D-plots of FRET efficiency versus denaturant concentration are presented in a waterfall scheme for equilibrium unfolding (left) and refolding (right) measurements. Similar to the NTD and CTD, the N-C interface also shows an intermediate state that is unique to the refolding process. **(D)** A kinetic analysis for N-C interface refolding was assayed from smFRET experiments. Refolding was induced after 30-fold dilution of 3 M GuHCl denatured protein. Refolding was monitored via the increase in the low acceptor lifetime (see Supplementary Figure 3). The red curve is a mono-exponential fit to the data (black).The half-life ($t_{1/2}$) for the refolding process is 27.2 min (see Table 1 and Table S2).

Figure 4

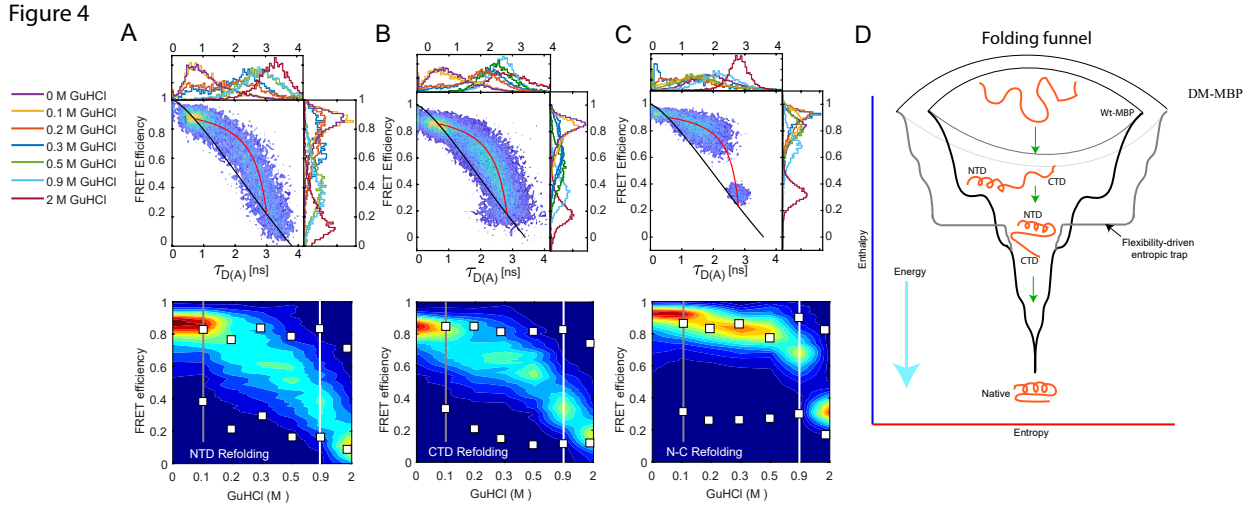

Figure 4. FRET efficiency versus donor lifetime analysis shows that the conformational search is the cause for then entropic trap

(A-C) Upper panels: 2D-plots of FRET efficiency vs donor lifetime in presence of an acceptor ($\tau_{D(A)}$) (E-τ plot), for NTD equilibrium refolding measurements shown in the right panel of Figure 2A (A), for CTD equilibrium refolding measurements shown in the right panel Figure 3A (B) and for N-C interface experiments shown in the right panel of Figure 3C (C). For comparison, all measurements are superimposed. The 1-D projections of the different GuHCl measurements, color coded according to the legend, are shown above and to the right of the 2D plots. The ideal relationship between FRET efficiency and donor lifetime is shown in black (static-FRET line) for a molecule that remains in a single FRET state during the observation time. Dynamic interconversion on micro- to millisecond time-scales between the folded and unfolded FRET states leads to a deviation in the E-τ plot, as shown in red (dynamic-FRET line). See Materials and Methods. For simplicity, the dynamic FRET red line is only shown for the 0.9 M GuHCl refolding experiment.

Lower panels: A waterfall scheme comparing the burst averaged FRET efficiency versus denaturant concentration for NTD refolding (A), CTD refolding (B) and formation of the N-C interface (C). The burst-averaged FRET efficiencies were compared with the FRET efficiencies obtained from the two donor lifetime components determined from a fit to the photons detected from all bursts. For each denaturant concentration, the donor lifetime derived FRET efficiencies represent two conformational states which are interconverting during refolding, shown as filled squares in white. A white line separates the 0.9 M GuHCl measurement from the higher denaturant concentrations, as there is a continuous distance increase in the unfolded state between 1 to 2 M GuHCl (see Table 4). Note that the donor lifetime derived FRET efficiency values for the unfolded state continuously decreases with increasing denaturant concentration. Grey line indicates the folded conformation of NTD and N-C interface, whereas CTD is still yet to be fold when refolding was performed in 0.1 M GuHCl.

(D) A schematic of the folding funnel describing the origin of entropic barrier resulting in intermediate state during the refolding process. Vertical axis represents the enthalpic energy and horizontal axis represents the configurational entropy. Dark funnel is for Wt-MBP and grey funnel for DM-MBP. In case of Wt-MBP, NTD folds faster with a guided by the hydrophobic interaction, which then is followed by CTD folding. In case of DM-MBP, due to the disrupted hydrophobic nucleus, NTD folding is limited by the high configurational entropy leading to slow folding. This step represents the higher degree of native contacts needed to stabilize the NTD

folding than typically required by Wt-MBP to compensate for the loss of binding energy due to double mutations. As soon as the NTD finds its folding competent conformation in DM-MBP, CTD folds faster and follows the rest of the folding funnel as for Wt-MBP. Presence of a flexibility in the DM-MBP refolding experimentally demonstrated with FRET efficiency versus donor lifetime analysis highlights the extended region in the horizontal plane creating an entropic barrier towards the folded native state.

Figure 5

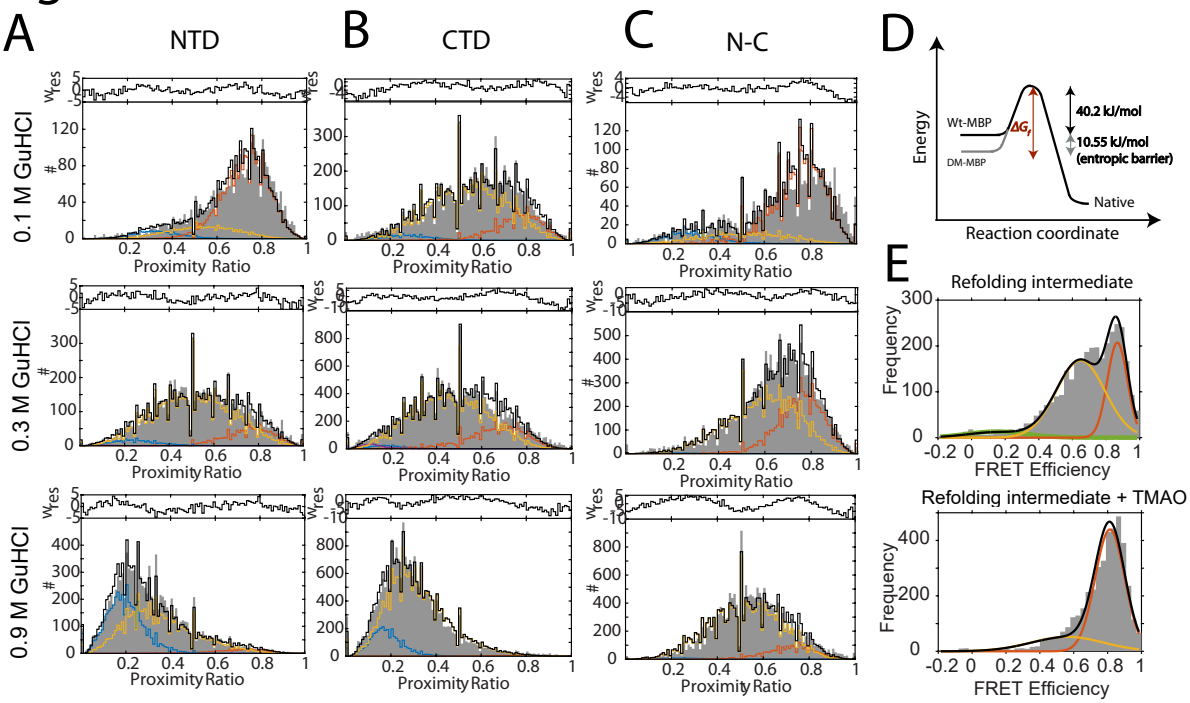

Figure 5. Quantification of the energy barrier in an entropic trap

(A-C) A dynamic PDA analysis for NTD refolding (**A**), CTD refolding (**B**) and for the N-C interface (**C**) in 0.1 M GuHCl (upper panels), 0.3 M GuHCl (middle panels), and 0.9 M GuHCl (lower panels). The initial 2000 s of the refolding measurements were analyzed in the case of refolding in 0.1 M GuHCl for all the three constructs. In grey are the proximity ratio (*PR*) histograms with 1 ms binning. The dynamic PDA fitting is highlighted as a black outline in the histogram. In the dynamic PDA fit, the folded population is highlighted in red, the unfolded population in blue, and yellow represents the contribution of the interconverting species to the *PR*. Photons were binned into histograms with 0.5, 1 and 1.5 ms binning and were fitted globally to quantify the transition rates. See Supplementary Figure 6B-D and Table 5 for all the titrations.

(D) A schematic showing the Gibbs free energy versus reaction coordinate, where an additional energy barrier of 10.55 kJ/mol is imposed by the flexible-entropically trapped intermediate on DM-MBP folding.

(E) Upper panel: A smFRET histogram for NTD refolding in 0.2 M GuHCl, where a significant fraction of the intermediate population is observed.

Lower panel: A smFRET histogram for NTD refolding in 0.2 M GuHCl where 500 mM trimethylamine-N-oxide has been added. The trimethylamine-N-oxide acts as a molecular chaperone by confining the conformational space available to the protein through stabilization of the solvent shell around the protein. In the case of DM-MBP, this allows the protein to overcome the entropic barrier, leading to faster folding and an increase in the native conformation in 0.2 M GuHCl.

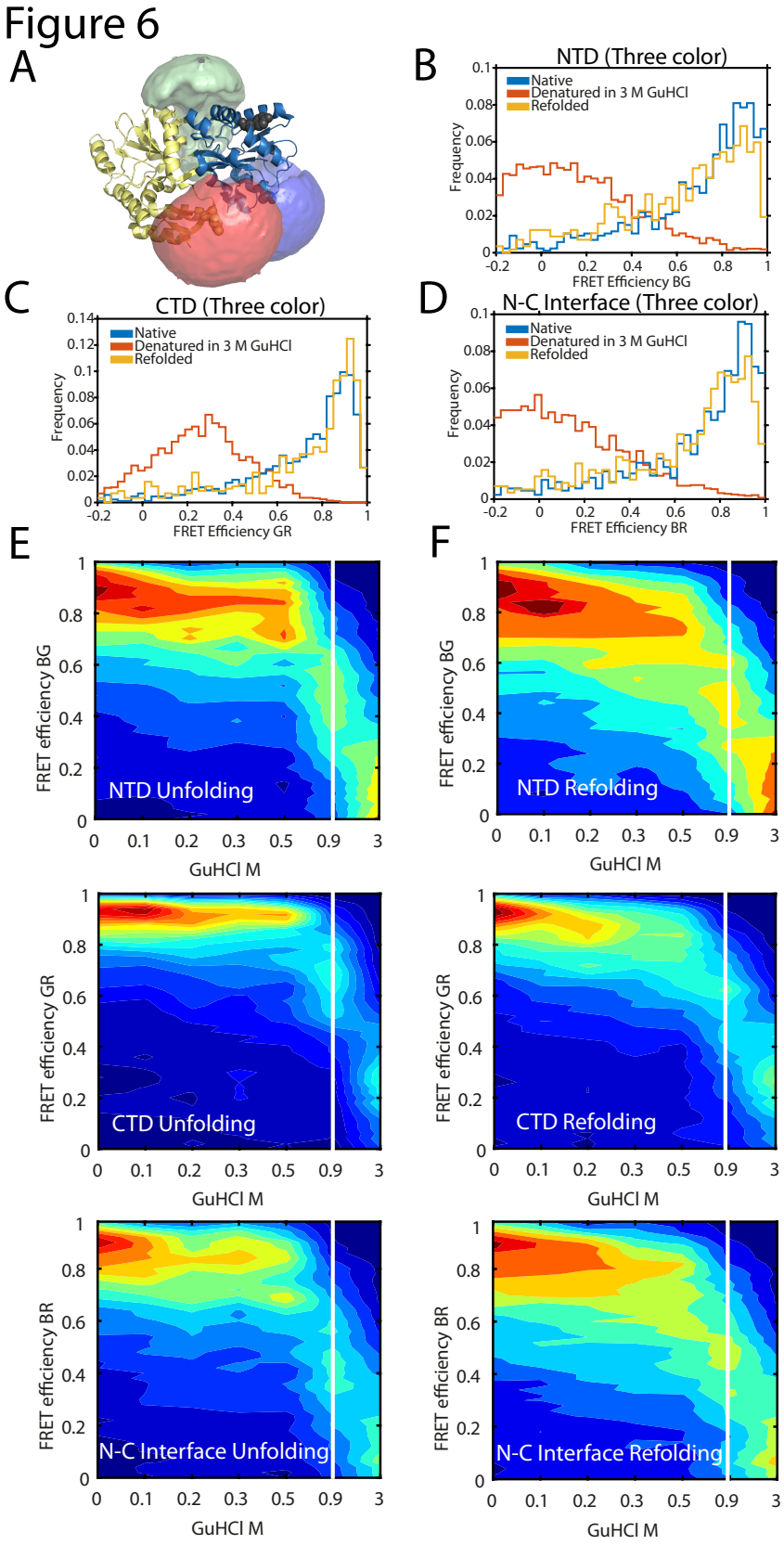

Figure 6. Three-color smFRET demonstrates the co-existence of an intermediate population

(A) Accessible volume calculations for triply-labeled DM-MBP on the MBP structure (PDB ID:1OMP) at the labeling positions A52, K175 and P298 for Atto488, Atto565 and Alexa647 dyes respectively.
(B)-(D) SmFRET histograms from the three-color FRET experiments for the three FRET pairs GR, monitoring the NTD (**B**), BG monitoring the CTD (**C**) and BR for the N-C interface (**D**). SmFRET histograms are compared for the native conformation (blue), denatured state (in 3 M GuHCl in orange) and refolded conformation (yellow).
(E)-(F) 2D-plot of FRET efficiency versus GuHCl concentration, presented as waterfall plots to visualize the conformational changes during equilibrium unfolding (**E)** and refolding **(F)** of triple-labeled DM-MBP. The upper panels show three color-BG FRET histograms indicative of the NTD, middle panels show three color-GR FRET histograms indicative of CTD and lower panel show three color-BR FRET histograms for N-C interface. For the equilibrium unfolding experiments, three-color FRET measurements were performed on native triple-labeled DM-MBP diluted into different GuHCl concentrations. For the refolding measurements, first, DM-MBP was denatured in 3 M GuHCl for 1 hr at 50°C. The sample was allowed to refold by diluting the sample to indicated final GuHCl concentrations. The white line separates the 0.9 M GuHCl measurement from the higher denaturant concentrations, where the average separation in the unfolded state continues increase from 1 to 2 M GuHCl.

Figure 7

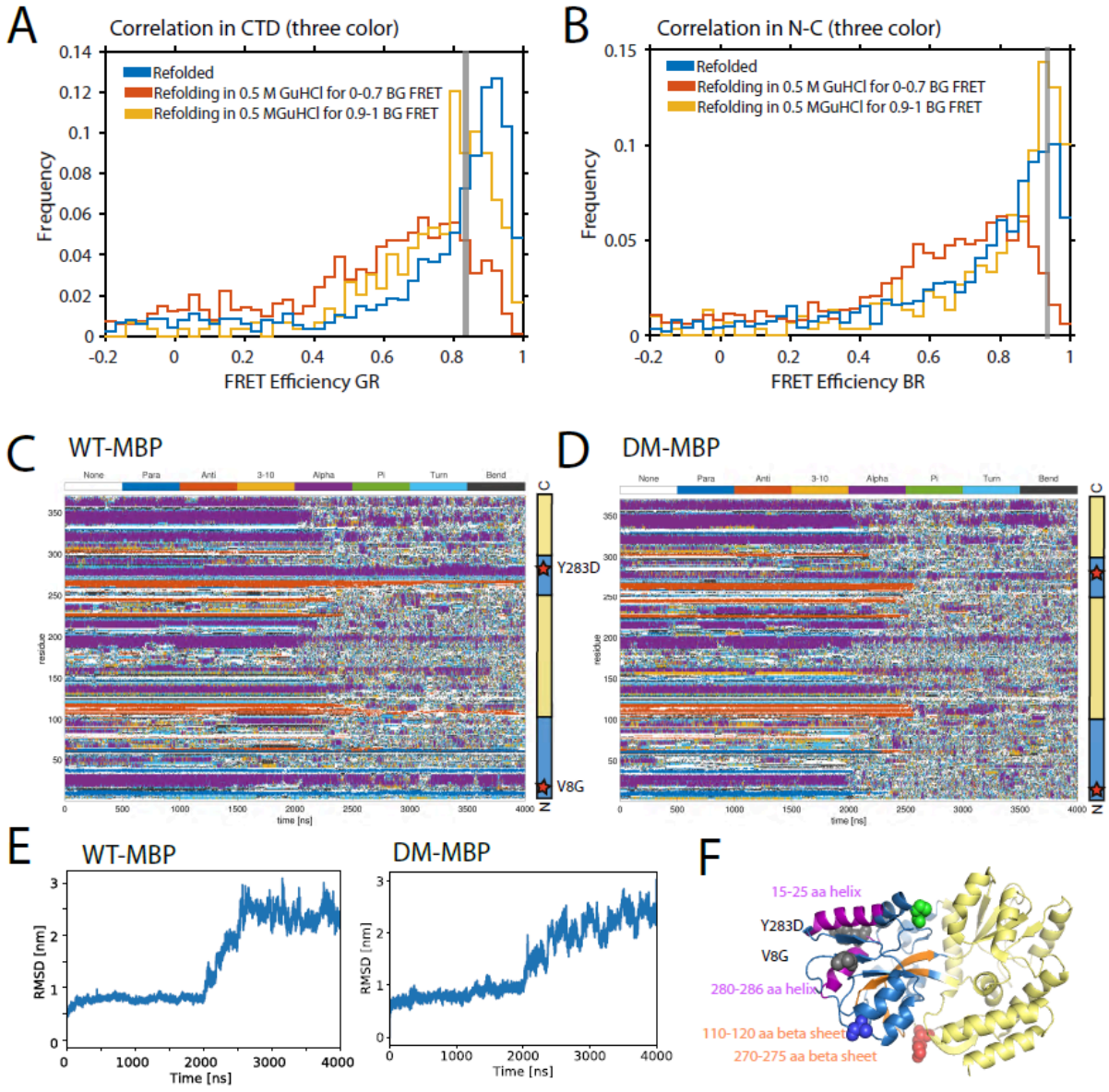

Figure 7. Correlative refolding analysis from three-color smFRET and unfolding MD simulations

(**A-B)** A comparison of the three color smFRET histograms for molecules in the intermediate state (0< E_{BG} < 0.7, orange) and in the native-like conformation (E_{BG} > 0.9, yellow), as determined from the FRET efficiency in the NTD. The three color-BR smFRET histograms for the CTD (**A**) and the three-color-BR for the N-C interface (**B**) are shown. Measurements of triple-labelled DM-MPB refolding in 0.5 M GuHCl were used for the analysis. All the three FRET efficiencies are available simultaneously from the same molecule. For comparison, the respective smFRET histograms for the refolded state are shown in blue. Grey line shows that the CTD is not folded yet while N-C interface has folded when molecules were selected for folded state for NTD.

(C-D) DSSP (Definition of Secondary Structure of Proteins) based secondary structure annotation plot for MD simulations based on PDB ID:1OMP were performed on WT-MBP (**C**)

and DM-MBP (**D**) during temperature-induced unfolding. The plot highlights the various secondary structures elements present in amino acid sequence from N to C terminus as a function of simulation time. The simulation was performed at 400 K for the initial 2 μs to reach equilibrium and then temperature was than increased to 450 K for another 2 μs to induce unfolding. Color scheme: random coil is shown in white, parallel beta sheets in blue, anti-parallel beta sheets in orange, 3-10 helices in yellow, alpha helices in purple, Pi helices in green, beta turns in sky blue and bends in black. Predominantly found alpha helices and anti-parallel beta sheets secondary structures can be seen throughout the MBP structure. The locations of the double mutations present in DM-MBP, V8G and Y283D, are depicted in right panel in the MBP sequence.
(E) RMSD (Root mean square deviation) plot for MD simulations performed during temperature induced unfolding from panel C on WT-MBP (left panel) and from panel D on DM-MBP for (right panel)**.** RMSD is calculated for backbone atoms over a period of MD simulation with respect to the initial structure.
(F) Secondary structures are highlighted on the MBP structure that are preserved in the MD simulations performed on WT-MBP after heating at 450 K as shown in panel C. Alpha helices are shown in purple, anti-parallel beta sheets are shown in orange. Double mutations in dark grey and three dye labels are shown in their respective color codes used for three-color smFRET.

Figure 8

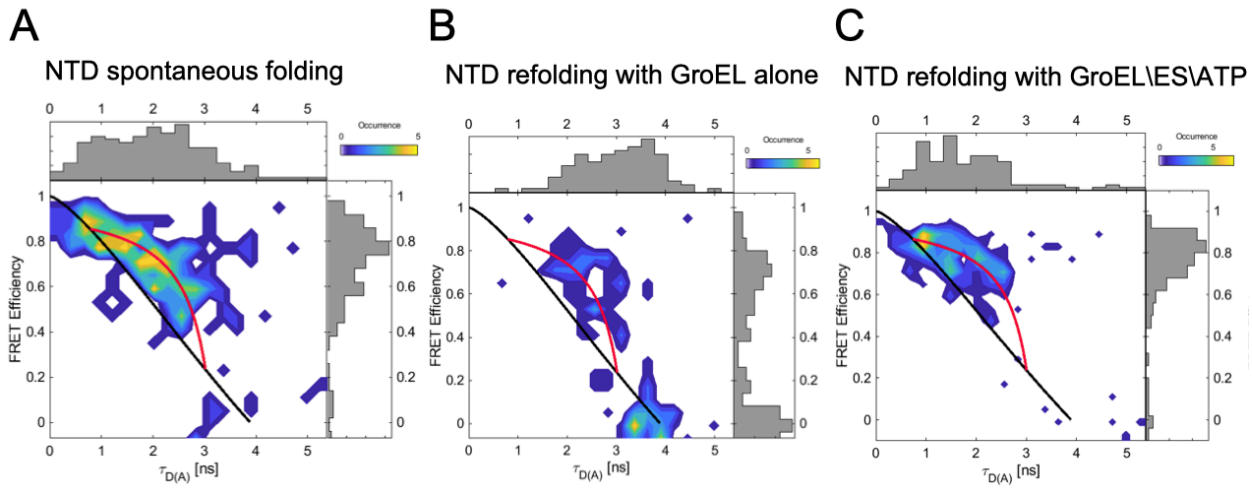

Figure 8. The influence of GroEL and GroEL\ES\ATP mediated confinement on the folding landscape of DM-MBP

(A-C) 2D FRET efficiency vs donor lifetime ($\tau_{D(A)}$) histograms (E-τ plot) of the NTD during refolding in 0.1 M GuHCl. The refolding of the NTD is shown for (A) spontaneous folding, (B) in presence of 3 μM GroEL, and (C) upon the addition of 3 μM GroEL, 6 μM GroES and 2 mM ATP. Only the initial 5 minutes of all the three measurements are shown. For GroEL and GroEL\ES\ATP measurements, the added components were incorporated in the dilution buffer. The 1-D projections are shown above and to the right of the 2D plots. The ideal relationship between FRET efficiency and donor lifetime is shown in black (static-FRET line) for a molecule that remains in a single FRET state during the observation time. Dynamic interconversion on micro- to millisecond time-scales between the folded and unfolded FRET states leads to a deviation in the E-τ plot, as shown in red (dynamic-FRET line). See Materials and Methods for details.

Paper 2: Coordinated Conformational Processing of the Tumor Suppressor Protein p53 by the Hsp70 and Hsp90 Chaperone Machineries

Please cite this article in press as: Dahiya et al., Coordinated Conformational Processing of the Tumor Suppressor Protein p53 by the Hsp70 and Hsp90 Chaperone Machineries, Molecular Cell (2019), https://doi.org/10.1016/j.molcel.2019.03.026

Molecular Cell
Article

CellPress

Coordinated Conformational Processing of the Tumor Suppressor Protein p53 by the Hsp70 and Hsp90 Chaperone Machineries

Vinay Dahiya,[1] Ganesh Agam,[2] Jannis Lawatscheck,[1] Daniel Andreas Rutz,[1] Don C. Lamb,[2,*] and Johannes Buchner[1,3,*]
[1]Center for Integrated Protein Science Munich CIPSM at the Department Chemie, Technische Universität München, Garching, Germany
[2]Department of Chemistry, Center for NanoScience, Nanosystems Initiative Munich (NIM) and Center for Integrated Protein Science Munich (CiPSM), Ludwig Maximilians University Munich, Munich, Germany
[3]Lead Contact
*Correspondence: d.lamb@lmu.de (D.C.L.), johannes.buchner@tum.de (J.B.)
https://doi.org/10.1016/j.molcel.2019.03.026

SUMMARY

p53, the guardian of the genome, requires chaperoning by Hsp70 and Hsp90. However, how the two chaperone machineries affect p53 conformation and regulate its function remains elusive. We found that Hsp70, together with Hsp40, unfolds p53 in an ATP-dependent reaction. This unfolded state of p53 is susceptible to aggregation after release induced by the nucleotide exchange factor Bag-1. However, when Hsp90 and the adaptor protein Hop are present, p53 is transferred from Hsp70 to Hsp90, allowing restoration of the native state upon ATP hydrolysis. Our results suggest that the p53 conformation is constantly remodeled by the two major chaperone machineries. This connects p53 activity to stress, and the levels of free molecular chaperones are important factors regulating p53 activity. Together, our findings reveal an intricate interplay and cooperation of Hsp70 and Hsp90 in regulating the conformation of a client.

INTRODUCTION

The tumor suppressor protein p53 is a sequence-specific transcription factor and central player in the regulation of cell fate (Kastenhuber and Lowe, 2017). p53 consists of four functional domains (Römer et al., 2006). The DNA binding domain (p53-DBD) harbors the majority of the mutations that have been found in more than 50% of human cancers (Joerger and Fersht, 2007; Olivier et al., 2010). Biochemical and structural studies of the p53-DBD suggest that p53 is dynamic and conformationally unstable (Cañadillas et al., 2006; Hainaut et al., 1995). This intrinsic conformational flexibility allows it to interact with the molecular chaperones Hsp70/Hsp40 and Hsp90 (Hagn et al., 2011; Hainaut and Milner, 1992; Rüdiger et al., 2002; Zylicz et al., 2001). Hsp90 is an essential ATP-dependent molecular chaperone required for the activation and stabilization of many regulatory proteins, referred to as ''clients'' (Taipale et al., 2012). Upon ATP binding and hydrolysis, Hsp90 undergoes conformational rearrangements important for client activation (Graf et al., 2009; Li et al., 2012; Pullen and Bolon, 2011). The clients are often handed over from Hsp70 to Hsp90 for their final maturation (Smith et al., 1992; Wegele et al., 2006). Hsp70 is an abundant heat shock protein whose general role is the stabilization of unfolded or partially folded states of proteins (Kellner et al., 2014; Sharma et al., 2010). Thus, Hsp70 prevents aggregation of unfolded proteins or assists with the transfer to downstream chaperones such as Hsp90 (Clerico et al., 2015; Frydman, 2001; Karagöz et al., 2014; Kirschke et al., 2014). Hence, Hsp70 and Hsp90 act synergistically to promote structure formation in proteins. In this process of client maturation, Hsp70 and Hsp90 collaborate with a multitude of co-chaperones that regulate their activities (Echtenkamp and Freeman, 2012; Mayer and Bukau, 2005; Schopf et al., 2017).

Both wild-type and mutant forms of p53 have been shown to interact with Hsp90 (Blagosklonny et al., 1996; Müller et al., 2004; Sepehrnia et al., 1996; Walerych et al., 2004; Whitesell et al., 1998). The extent of chaperone dependence of p53 seems to be dictated by stress signals like heat shock and oncogenic mutations that destabilize its conformation (Römer et al., 2006). Wild-type p53 undergoes transient interactions with Hsp90, whereas p53 mutants display extended interactions with the chaperone machinery (Muller et al., 2008; Whitesell and Lindquist, 2005). Several *in vitro* reports have demonstrated that p53 is able to maintain DNA binding activity without the aid of chaperones at room temperatures but requires Hsp90 for its stabilization under physiological and heat shock temperatures (Müller et al., 2004; Wang and Chen, 2003). In cells, the cooperation between Hsp90 and Hsp70 increases under heat shock to maintain p53 activity (Walerych et al., 2009). The studies done so far establish that wild-type p53 associates with Hsp70 and Hsp90 under normal and stressed conditions and that they are important in keeping p53 active under stress. However, how the two molecular chaperone machineries affect the conformation of p53 has remained largely unknown. Here we reconstituted the chaperoning cycle of the Hsp70 and Hsp90 system for p53 *in vitro* and found opposing effects of the two machineries on p53 activity.

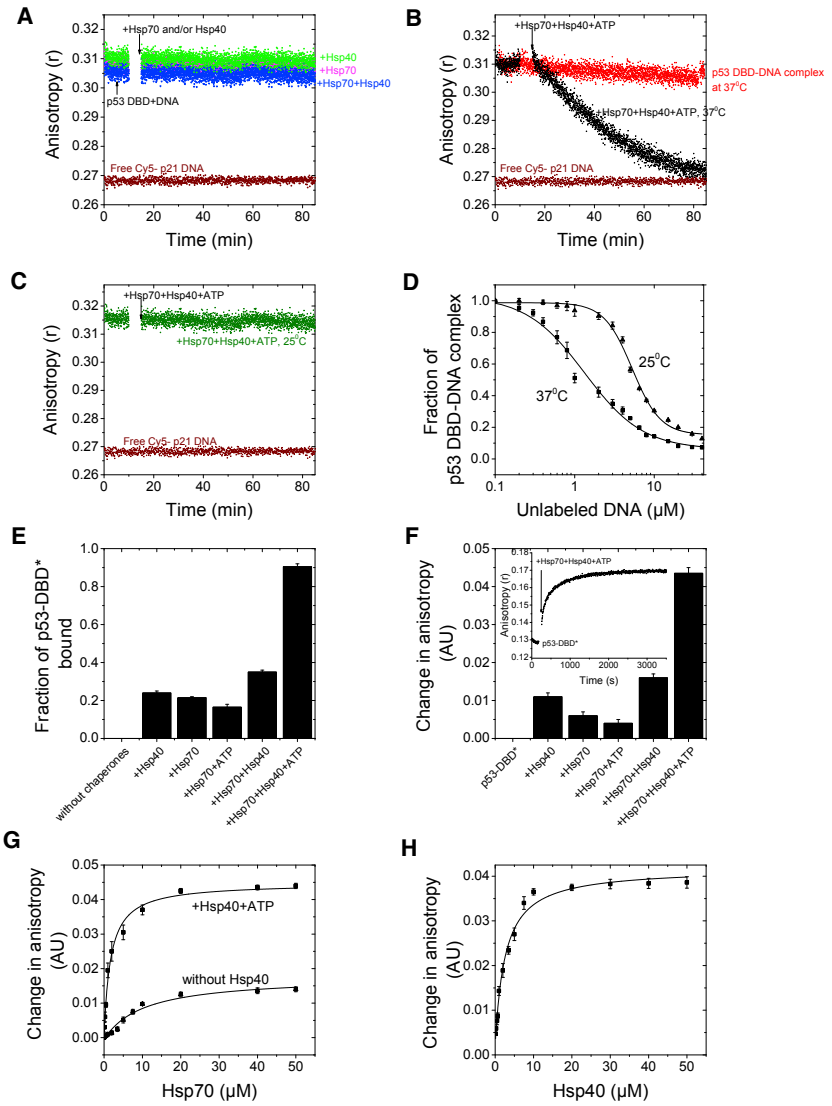

Figure 1. The p53-DBD Interacts with Hsp70-Hsp40 and Is Dissociated from DNA in the Presence of ATP

(A) The dissociation kinetics of the p53-DBD-DNA complex were monitored over time after addition of 2 μM Hsp40 (green), 15 μM Hsp70 (pink), or 2 μM Hsp40 and 15 μM Hsp70 (blue) at 37°C. Anisotropy of free Cy5 labeled p21 DNA is shown in brown.

(B) DNA dissociation kinetics of the p53-DBD in the presence of 2 μM Hsp40, 15 μM Hsp70, and 5 mM ATP at 37°C (black); the anisotropy of the p53-DBD-DNA complex at 37°C is shown in red. See also Figure S1A.

(C) p53-DBD is not dissociated from DNA by Hsp40, Hsp70, and ATP at 25°C (olive). The anisotropy of free Cy5 labeled p21 DNA is shown in brown for comparison.

(D) Dissociation of the p53-DBD from Cy5-labeled DNA initiated by increasing concentrations of unlabeled DNA at 37°C (squares) or 25°C (triangles). Fractions of p53-DBD-DNA complexes at each unlabeled DNA concentration were calculated from the change in anisotropy of free Cy5 DNA, setting the anisotropy change of the Cy5-DNA-p53-DBD complex in the absence of any unlabeled DNA as 100%. The data were fitted to a Hill plot to obtain the IC_{50} values of unlabeled DNA at different temperatures. Mean values and SD are from three separate experiments.

(E) The fraction of the p53-DBD* bound to Hsp70 and Hsp40 is calculated from AUC sedimentation velocity curves; see Figure S1C. The experiment was conducted in triplicates, with SD shown.

(F) Change in anisotropy of Atto-488-labeled p53-DBD (p53-DBD*) upon addition of Hsp70, Hsp40, and ATP as indicated. The inset shows the kinetics of increase in anisotropy when Hsp70, Hsp40, and ATP are added to the p53-DBD* at 25°C. The mean of three independent experiments with SD is plotted.

(G) Binding of the p53-DBD to Hsp70 (triangles) and Hsp70/Hsp40 (squares) in the presence of ATP, monitored by the change in anisotropy of IAEDANS-labeled p53-DBD* at 25°C. See also Figures S1D and S1E.

(H) Binding of Hsp40 to the p53-DBD*.

The data in (G) and (H) were fitted to a one-site binding model, and the average values with SD are shown from three different experiments.

See also Figures S1F and S1G.

RESULTS

Hsp70 and Hsp40 Inactivate p53 in an ATP-Dependent Reaction

Although it has been known that p53 interacts with the Hsp70 and Hsp90 machineries *in vivo* (Hainaut and Milner, 1992; Müller et al., 2004; Walerych et al., 2004, 2009; Wang and Chen, 2003; Zylicz et al., 2001), the structural consequences for p53 have remained enigmatic. We used an anisotropy-based assay to test DNA binding to the p53-DBD in the presence or absence of the chaperones (Figure 1A). Addition of Hsp70 and/or Hsp40 to a p53-DBD-DNA complex had no effect (Figure 1A). However, upon inclusion of ATP with Hsp70-Hsp40, we observed a decrease in anisotropy consistent with release of the labeled DNA (Figure 1B). DNA dissociation of the p53-DBD by Hsp70 is strictly dependent on ATP hydrolysis, as the non-hydrolysable ATP analog AMP-PNP and ADP were ineffective in dissociating the p53-DBD from DNA (Figure S1A). Interestingly, Hsp70 could displace the p53-DBD from the DNA only at 37°C (Figure 1B) but not at 25°C (Figure 1C). The half-maximal inhibitory concentration of unlabeled DNA (IC_{50}) required for dissociation of p53-DBD-DNA complexes was found to be 4.5 times higher at 25°C (IC_{50}, 5.5 μM) than at 37°C (IC_{50}, 1.2 μM) (Figure 1D). This implies that, at 37°C, the p53-DBD conformation is less stable, rendering p53-DBD-DNA complexes weaker, which enables Hsp70 to dissociate them. To monitor the interaction of Hsp70 with the p53-DBD, we used labeled p53-DBD (p53-DBD*) in combination with anisotropy or analytical ultracentrifugation

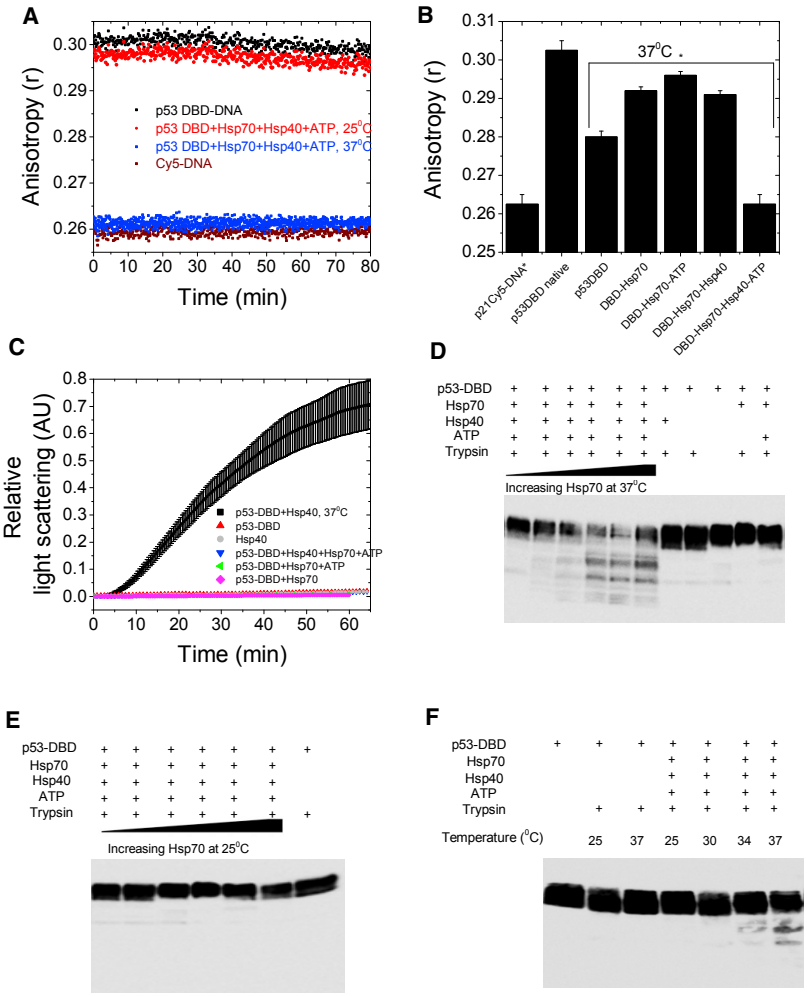

Figure 2. Hsp70-Hsp40-ATP Abolishes DNA Binding of the p53-DBD by Unfolding It at 37°C

(A) Association kinetics of Cy5-labeled p21 DNA to the p53-DBD (black) and in the presence of Hsp70, Hsp40, and ATP. The complex was formed at 25°C (red) or 37°C (blue).

(B) Steady-state anisotropy values of DNA at 25°C in the presence of the p53-DBD and when the p53-DBD was incubated with Hsp40, Hsp70, and ATP in different combinations at 37°C for 1 h. Average values and SD are plotted from three independent measurements. See also Figure S1I.

(C) Kinetics of light scattering monitored at 37°C for the p53-DBD (red); Hsp40 (gray); p53-DBD with Hsp40 (black); p53-DBD with Hsp70 (pink); p53-DBD with Hsp70 and ATP (green); or p53-DBD with Hsp70, Hsp40, and ATP (blue). The mean and SD from three separate measurements are depicted for the Hsp40- and p53-DBD-containing sample.

(D) Limited proteolysis of the p53-DBD was performed in the presence of trypsin at 37°C. p53-DBD fragments were detected by western blot.

(E) Limited proteolysis assay as in (D) at 25°C.

(F) Temperature titrations showing increased sensitivity of p53-DBD when the complexes with Hsp70, Hsp40, and ATP were formed at increasing temperatures.

(AUC) coupled to fluorescence detection. The presence of the fluorophore did not affect the functionality of the p53-DBD, as confirmed by its DNA binding activity (Figure S1B). We tested the interaction both at 37°C and at 25°C by AUC. p53-DBD* alone sediments at 2.7 S, and, upon addition of Hsp70/Hsp40/ATP, the value increases to 14.5 S at both temperatures, indicating complex formation (Figure S1C). Binding of the p53-DBD* to Hsp70 was strongly enhanced in the presence of Hsp40 and ATP, as evidenced by the complete disappearance of the free p53-DBD* peak (Figure S1C). Based on the AUC sedimentation analysis shown in Figure S1C, the bound fraction of p53-DBD* was calculated as described before (Lorenz et al., 2014) for each condition (Figure 1E). It can be seen that, in the presence of Hsp70/Hsp40/ATP, 90% of the p53-DBD* was found in complex with the chaperones. In the absence of Hsp40 and/or ATP, only 20% of the p53-DBD* was bound (Figure 1E). Similarly, a significant change in the anisotropy of the p53-DBD*, indicative of an interaction, occurred only upon addition of Hsp70, Hsp40, and ATP (Figure 1F). This is supported by the shift of the S value of the p53-DBD*-Hsp40 complex from 12 S to 14.5 S upon addition of Hsp70 and ATP (Figure S1C). Both by anisotropy and AUC, the K_D for the interaction of the p53-DBD* and Hsp70 was found to be 2 μM in the presence of Hsp40 and ATP at 25°C (Figures 1G, S1D, and S1E). Interestingly, Hsp40 alone binds to the p53-DBD* with similar affinity (K_D 3 μM) as the complete Hsp70-Hsp40 system (Figures 1H, S1F, and S1G). Of note, more than one Hsp40 molecule binds to the p53-DBD*, as evident from high sedimentation values of p53-DBD*-Hsp40 complexes (~12 S; Figure S1F) in the absence of aggregation because no increase in light scattering was observed in these samples (Figure S1H). Taken together, these data suggest that Hsp40 forms complexes with the p53-DBD and primes it for Hsp70 interaction. The combined action of Hsp70 and Hsp40 results in ATP-dependent loss of DNA binding of the p53-DBD at 37°C.

To further decode the effect of the Hsp70-Hsp40 system on p53-DBD conformation, we assayed DNA binding of the p53-DBD after forming its complexes with Hsp70-Hsp40 in the presence of ATP at 25°C and 37°C. We observed complete inhibition of DNA binding activity of the p53-DBD only for complexes formed at 37°C but not at 25°C (Figure 2A). As expected, complete inactivation of DNA binding of the p53-DBD by Hsp70 occurs only in the presence of Hsp40 and ATP (Figure 2B), where stable complexes are generated. Partial inactivation of the p53-DBD alone was also observed upon incubation at 37°C (Figure 2B). This implies that the p53-DBD is prone to inactivation at physiological temperatures. Interestingly, an increase in anisotropy was observed for p53-DBD samples incubated at 37°C

with Hsp70 with or without ATP (Figure 2B). This implies that Hsp70 is either increasing the pool of active p53-DBD or that the p53-DBD binds to DNA in complex with Hsp70. To distinguish between these two possibilities, an electrophoretic mobility shift assay (EMSA) was performed. This showed that the presence of Hsp70 prevents irreversible inactivation of the p53-DBD and increases the DNA-binding active fraction of the p53-DBD (Figure S1I). No significant up-shift of p53-DBD complexes was observed in the presence of Hsp70/Hsp40, implying that the chaperones do not bind to DNA along with p53 (Figure S1I).

Hsp70-Hsp40 Increases the Population of Unfolded p53-DBD

Complete inhibition of the DNA binding activity of p53-DBD at 37°C in the presence of Hsp70-Hsp40-ATP could be due to two reasons. Either complex formation of the p53-DBD with Hsp70-Hsp40 prevents DNA binding sterically, or the chaperone system traps unfolding p53, and, hence, DNA-binding is lost. If the loss in DNA binding goes along with loss in structure and stability of the p53-DBD, this may lead to aggregation. Light scattering experiments monitoring the p53-DBD at 37°C showed that the protein does not aggregate alone or in the presence of Hsp70-Hsp40 and ATP (Figure 2C). Surprisingly, Hsp40 was found to trigger aggregation of the p53-DBD, in contrast to Hsp70, which prevented aggregation at 37°C (Figure 2C). This indicates that Hsp40's interaction with the p53-DBD results in the destabilization of its structure necessary for stable interaction of the p53-DBD with Hsp70. To test the conformational status of the p53-DBD in complex with Hsp70 and Hsp40, limited proteolysis assays were performed, with the complexes formed at 25°C and 37°C. These experiments demonstrated that the p53-DBD in the Hsp70-Hsp40-ATP complex is sensitive to proteolysis at 37°C (Figure 2D) but not at 25°C (Figure 2E). Further, increased sensitivity of the p53-DBD to proteolysis was found with increasing temperatures (Figure 2F). Hsp70 alone, in the absence of Hsp40, does not have an effect on proteolysis of the p53-DBD (Figure 2D), which is in line with the EMSA results, where Hsp70 was found to increase the pool of active p53-DBD. Hsp40 does not affect the proteolysis sensitivity of p53 at 37°C, possibly because aggregates formed in the presence of Hsp40 are resistant to trypsin (Figure 2D).

Unfolding of the p53-DBD induced by Hsp70/40 was further analyzed by single-pair fluorescence resonance energy transfer (spFRET) experiments with multi-parameter fluorescence detection and pulsed interleaved excitation (MFD-PIE) (Kudryavtsev et al., 2012). Based on the accessibility and the distance between the naturally exposed cysteines in the p53-DBD (Huang et al., 2009; Scotcher et al., 2011), an intra-molecular FRET construct of the p53-DBD was designed using the naturally occurring cysteines at positions 182 and 229 (Figure 3A) and mutating the two other accessible cysteine residues to serines (C124S and C277S). The labeled p53 FRET sensor was shown to bind DNA and interact with the Hsp70-Hsp40 system, similar to the wild-type protein (Figures S2A and S2B). As seen in the spFRET histogram in Figure 3B, the native p53-DBD C182C229 sensor shows a clear FRET peak with a FRET efficiency of ~0.8 at room temperature. We used photon distribution analysis (PDA) (Kudryavtsev et al., 2012) to delineate multiple populations and the width of the population distribution and to gain insight into conformational flexibility. The PDA showed that the native p53-DBD C182C229 (Table 1) has a well-defined structure with an inter-dye separation of ~49 Å and a width of $\sigma = 3.6$ Å (Figure 3B). This is in reasonable agreement with the expected distance from the accessible volume calculations of ~55 Å for the p53-DBD structure (Kalinin et al., 2012). To probe the flexibility, we analyzed the FRET efficiency and donor lifetime for each burst (STAR Methods; Kalinin et al., 2010). The native DBD C182C229 shows a small hint of dynamics, as highlighted by the dynamic FRET line (Figure 3B, right, red). For comparison, the ideal FRET efficiency versus donor lifetime relationship in the absence of dynamics is shown as a black line (Figure 3B), which corresponds to the majority of the measured bursts. Hence, the p53-DBD is conformationally labile, which makes it a candidate for chaperone interactions.

Next we investigated the conformation of p53-DBD C182C229 in complex with Hsp40, Hsp70, and ATP at 25°C. We observed a small fraction of molecules in a low FRET state with a FRET efficiency of ~0.1, representing the unfolded population of DBD C182C229 (Figure 3C). At higher temperatures, the population of the unfolded p53 DBD fraction increased (Figure 3C). Above 33°C, the majority of molecules displayed a low FRET efficiency value (Figure 3C). Again, we observed dynamics of the p53-DBD between the folded and unfolded states on the millisecond timescale (the dynamic FRET lines are shown in red in Figure 3C for 25°C and in red and green for 30°C). With Hsp40, Hsp70, and ATP at 37°C (Figure 3D), the majority of molecules were in a low FRET state with a donor-acceptor separation of ~95 Å in the DBD and a high conformational flexibility ($\sigma \gtrsim 10$ Å). A minor population was present with a separation of ~73 Å (FRET efficiency of ~0.2) (Table 1). The conformation of the p53-DBD C182C229 sensor alone analyzed at 37°C did not show the presence of any unfolded species (Figure 3D). As the temperature increases, the DBD structure becomes more flexible, as seen by an increase in the width of the folded population as well as in the dynamic transitions between the folded and unfolded states observed in the FRET efficiency versus donor lifetime plots (Figure S2C). Binding of Hsp40 to the p53-DBD triggers aggregation at higher temperatures (Figure 2C); hence, to determine Hsp40 effects on the p53-DBD structure in spFRET, we formed complexes of the p53-DBD C182C229 sensor with Hsp40 at 30°C. Under these conditions, Hsp40 does not exhibit a major unfolding effect on the p53-DBD (Figure S2D).

The distance of 95 Å observed for the p53-DBD FRET sensor upon unfolding by Hsp70-Hsp40-ATP (Table 1) is much larger than expected for the denatured state. To verify this, we performed spFRET measurements with the completely unfolded protein, p53-DBD C182C229, in 3 M GdmCl. We found that, in the unfolded state, the major population has an average FRET efficiency of ~0.5 (Figure S2E) corresponding to an inter-dye distance, per definition, of the Förster radius (STAR Methods), which is ~62 Å for the fluorophores Atto 532 and Alexa 647. This is in excellent agreement with the theoretically calculated end-to-end distance between the labeling positions C182 and C229 in a randomly unfolded protein using the freely joined chain (FJC) model (Lapidus et al., 2002) (STAR Methods). This leads to

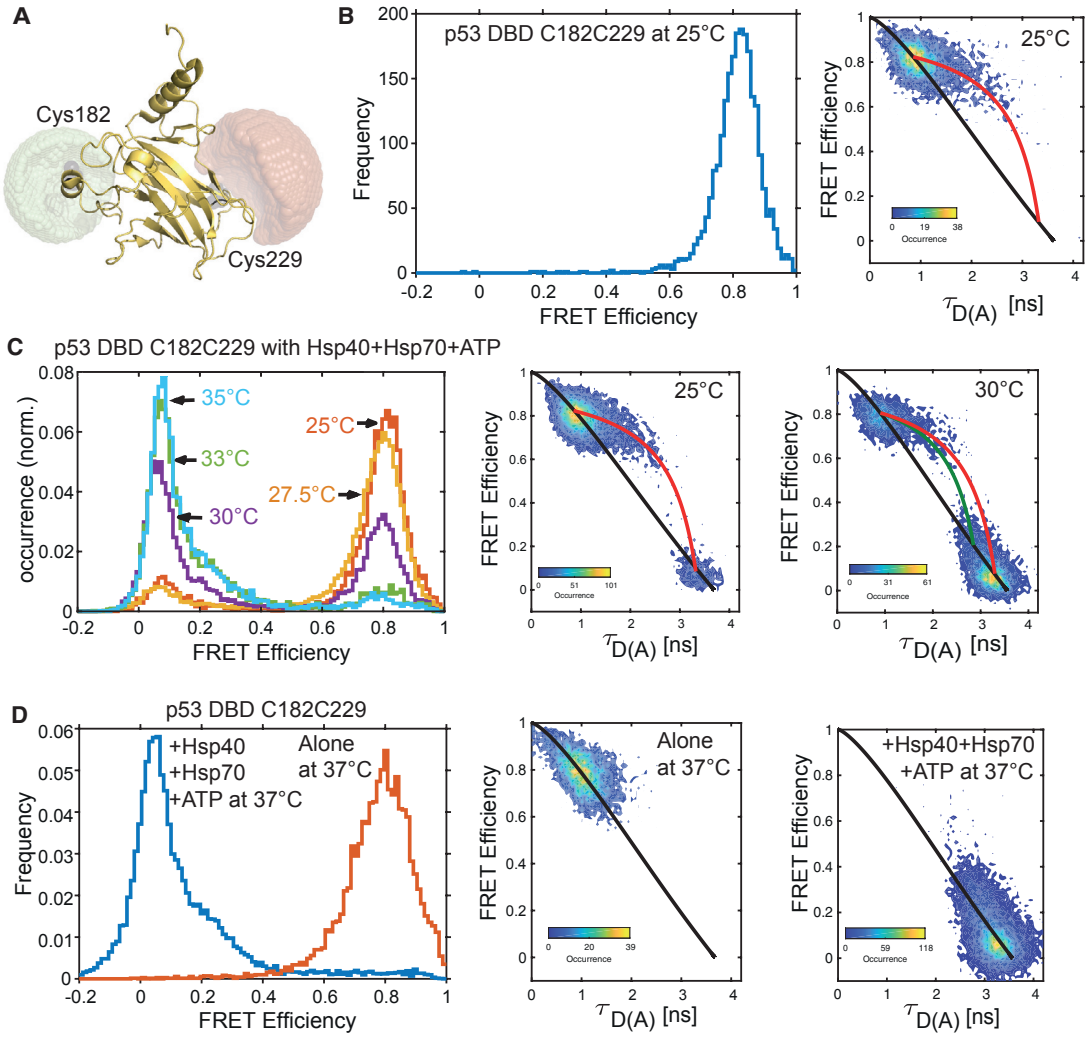

Figure 3. Hsp70-Bound Conformation and Dynamics of p53-DBD Probed by spFRET

(A) Accessible volume calculations at the labeling positions C182 and C229 on the p53-DBD structure (PDB ID: 1UOL). Dyes are labeled stochastically. See also Figures S2A and S2B.

(B) SpFRET histogram showing the native p53 DBD C182C229 sensor (left) and a 2D histogram for FRET efficiency versus donor lifetime in the presence of the acceptor ($\tau_{D(A)}$, right). The black line (static FRET line) describes the ideal relationship between FRET efficiency and donor lifetime when populations are from static molecules. The red line (dynamic FRET line) describes the molecules undergoing dynamic transition between populations within microseconds to milliseconds. See also Figures S2C.

(C) SpFRET histograms showing the p53-DBD C182C229 sensor conformation when bound with Hsp40 and Hsp70 in the presence of ATP at 25°C (orange line), 27.5°C (yellow line), 30°C (purple line), 33°C (green line), and 35°C (light blue line) for 1 h, as indicated (left). Also shown is a 2D histogram for FRET efficiency versus donor lifetime in the presence of the acceptor ($\tau_{D(A)}$) for the p53-DBD FRET sensor incubated with Hsp70-Hsp40-ATP at 25°C (center) and 30°C (right), respectively. The red and green lines are the dynamic FRET lines. See also Figures S2D and S7 and Table 1.

(D) SpFRET histograms showing the p53-DBD C182C229 conformation alone at 37°C (orange line) and when bound with Hsp40 and Hsp70 in the presence of ATP at 37°C (blue line) (left). A 2D histogram for FRET efficiency versus donor lifetime in the presence of the acceptor ($\tau_{D(A)}$) is shown for both conditions (center and right).

See also Figures S2E and S2F.

the conclusion that, in the presence of Hsp40, Hsp70, and ATP, p53 is not in a random conformation but in a highly elongated structure stabilized by its interactions with Hsp70. The separation of 95 Å is over half of the contour length for a polypeptide chain of 47 amino acids. Using fluorescence correlation spectroscopy (FCS) (STAR Methods), we quantified the diffusion coefficient of p53-DBD C182-229 alone and in complex with Hsp40-Hsp70-ATP at 37°C (Figure S2F). The p53-DBD in the presence of Hsp40-Hsp70-ATP shows a much slower diffusion coefficient (D = 29 $\mu m^2/s$) compared with p53-DBD alone

Table 1. PDA Results for the p53-DBD Sensor under Different Conditions, Measured with spFRET

p53 DBD Sensor	Subpopulation 1 d(Å) σ (Å) f(%)	Subpopulation 2 d(Å) σ (Å) f(%)	Subpopulation 3 d(Å) σ (Å)f(%)	Donor Only f(%)	$\chi^2_{red.}$
Alone at 25°C	48.8 3.6 100	–	–	–	1.63
Alone at 30°C	49.4 2.9 100	–	–	–	3.72
Alone at 37°C	50.1 4.9 100	–	–	–	4.18
+Hsp40 at 30°C	49.4 3.1 100	–	–	–	4.38
+Hsp40+Hsp70+ATP at 25°C	49.3 3.1 91.2	73.8 5.1 2.1	99.0 16.0 6.2	0.5	2.79
+Hsp40+Hsp70+ATP at 27.5°C	50.4 3.1 88.8	72.0 8.0 4.0	98.0 15.5 7.0	0.2	4.55
+Hsp40+Hsp70+ATP at 30°C	50.5 3.5 59.2	77.5 9.3 16.2	106.2 15.8 23.7	0.9	3.30
+Hsp40+Hsp70+ATP at 33°C	51.0 4.4 16.8	79.5 9.5 30.5	104.0 17.7 51.5	1.2	3.58
+Hsp40+Hsp70+ATP at 35°C	51.0 4.2 11.0	73.2 7.9 20.1	97.1 14.2 66.0	2.9	3.91
+Hsp40+Hsp70+ATP at 37°C	53.8 6.8 9.4	73.2 4.6 25.2	94.9 13.8 65.0	0.4	2.56
+Hsp40+Hsp70+ATP at 37°C → Bag1	51.8 4.1 84.1	71.1 8.8 12.5	99.8 5.3 2.9	0.5	4.66
+Hsp40+Hsp70+ATP at 37°C → Hop+Hsp90+Bag1	51.1 2.9 86.6	60.0 11.1 13.4	–	–	5.42
+Hsp40+Hsp70+ATP at 37°C → Hop+Hsp90	51.3 4.9 46.0	70.0 10.8 47.0	99.0 4.8 7.0	–	2.58
+Hsp40+Hsp70+ATP at 37°C → Hop	50.1 3.6 19.7	74.0 9.5 64.3	99.2 5.4 14.4	1.6	3.30
+Hsp40+Hsp70+ATP at 37°C → Hsp90	50.2 6.8 12.1	73.0 9.8 84.4	99.3 5.4 3.5	–	2.16

The fraction of molecules is given by *f*, the distance by *d*, and the SD by σ. Occasionally, during a burst, the acceptor fluorophore goes into a dark, non-FRET state, leading to a small number of time intervals that need to be described as a "donor-only" population denoted by D only. Photons from each burst were used to build a proximity ratio histogram. The resulting histogram was then fitted using a Monte Carlo approach for simulating the burst-wise histogram using one, two, or three populations with a Gaussian distribution in donor-acceptor separation. The corresponding distance and width were calculated for each population. See also Figures S7 and S8.

(D = 81 $\mu m^2/s$), indicating binding of several chaperone molecules to the p53-DBD, which is consistent with a recently published molecular dynamics study (Kellner et al., 2014). In summary, the combined action of Hsp40, Hsp70, and ATP results in a complete shift of the p53-DBD toward an unfolded structure at physiological temperature. Thus, the activity of the Hsp70/40 system would render p53 inactive in the cell. However, it is well established that Hsp90 plays a major role in stabilizing p53 *in vivo*.

Hsp90 and Hop Relieve Hsp70 Inhibition of the p53-DBD in an ATP-Dependent Process

Because Hsp70 facilitates client delivery to Hsp90 (Alvira et al., 2014; Kirschke et al., 2014; Morán Luengo et al., 2018), and Hsp90 is important for maintaining p53 activity (Müller et al., 2004; Walerych et al., 2004), we tested whether Hsp90 can restore the DNA binding activity of the Hsp70-bound p53-DBD. In this context, the adaptor protein Hop, which binds to both Hsp70 and Hsp90 simultaneously, had been shown to promote client transfer from Hsp70 to Hsp90 (Johnson et al., 1998). We monitored the recovery of DNA binding of the Hsp70-40-p53-DBD complex formed at 37°C, as indicated in the experimental scheme (Figure 4A), after adding Hop and/or Hsp90. Addition of Hsp90 or Hop alone had a slight effect on recovery of DNA binding (Figure 4B), with 18%–20% of the p53-DBD molecules reaching the DNA-binding state. This effect could not be attributed to crowding because BSA added in similar amounts as Hop or Hsp90 did not result in any recovery of p53 function (Figure 4B). It is been shown that the Hsp70 and Hsp90 machineries can interact directly for client transfer (Genest et al., 2019). This might account for the slight effects of Hsp90 on the recovery of DNA binding activity of p53-DBD. When we added Hsp90 and Hop together, an ~50% regain of DNA binding activity was observed (Figure 4C), with a faster rate compared with Hsp90 alone (Figure 4B). Direct visualization of complexes by AUC using labeled p53-DBD* showed that addition of Hsp90 alone at 25°C to the Hsp70-40-p53-DBD* complex (Figure 4D, red, 12.2 S) resulted in a minimum release of p53-DBD* from the complex (Figure 4E), with no shift in S value indicating no or minimum incorporation of Hsp90 into the Hsp70-Hsp40-p53-DBD* complex (Figure 4D, olive). When Hop and Hsp90 were added together to the Hsp70-Hsp40-p53-DBD* complex, a shift in S value from 12.2 S to 14.2 S was observed (Figure 4D, pink), suggesting the presence of Hop and Hsp90 in the complex and resulting in the release of ~45% of p53-DBD*, as derived from the bound fraction calculated from the free p53-DBD* peak (Figure 4E). This correlates with the ~50% DNA binding recovery seen in Figure 4B in the presence of Hop and Hsp90.

Active Hsp90 is absolutely necessary for the p53-DBD to reach the native state because the Hsp90 inhibitor radicicol, which binds to the nucleotide binding pocket of Hsp90 (Roe et al., 1999), completely abolished DNA binding (Figure 4F). We used the ATP binding- and hydrolysis-defective mutants of Hsp90, D88A and E42A (Boczek et al., 2015), to further explore the role of ATP in Hsp90-triggered p53-DBD reactivation. No DNA binding recovery was obtained for either of the Hsp90 mutants (Figure 4F), implying that both ATP binding and hydrolysis by Hsp90 are required for reversing Hsp70 inhibition.

To test whether ATP hydrolysis is constantly needed, we added apyrase, an enzyme that rapidly hydrolyses ATP. We found that, in the presence of apyrase, the DNA binding activity of the p53-DBD was completely abolished (Figure 4G).

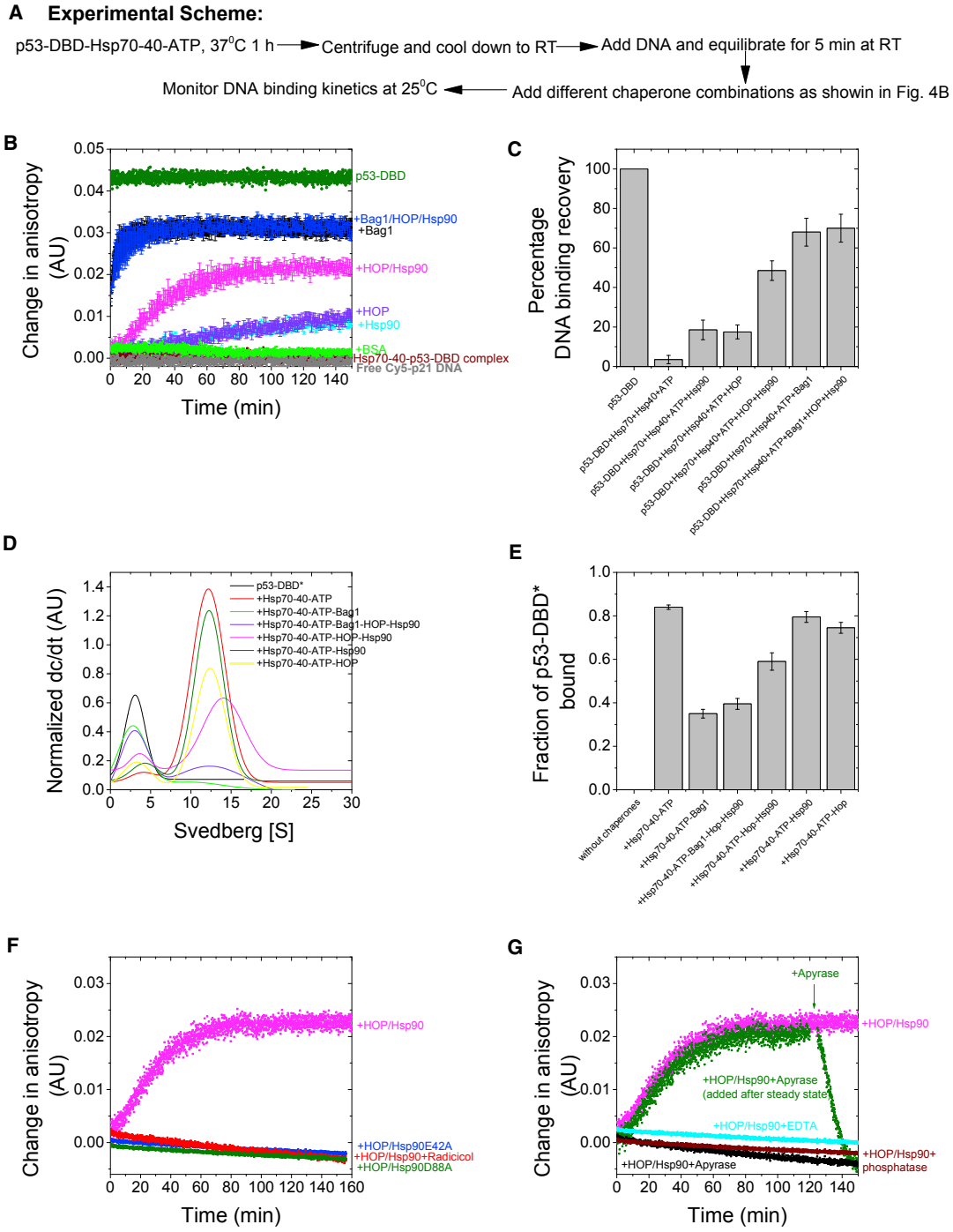

Figure 4. Reactivation of p53-DBD from Hsp70 Inhibition by Bag-1, Hop, and Hsp90

(A) Experimental scheme.

(B) Fluorescence anisotropy kinetics showing the change in anisotropy of Cy5-labeled p21 DNA (gray) when Hop (violet), Hsp90 (cyan), Hsp90 and HOP (pink), Bag-1 (black), or Bag-1 plus Hsp90 and Hop (blue) were added to the Hsp70-Hsp40-p53-DBD complex (brown). The change in anisotropy of the native p53-DBD-DNA complex is shown in olive. As a control, 20 μM BSA was added to the Hsp70-Hsp40-p53-DBD complex (green). The mean and SD from 3 separate experiments are plotted. See also Figures S4A–S4C.

(legend continued on next page)

Interestingly, when apyrase was added after the Hop and Hsp90-mediated DNA binding recovery of p53-DBD had reached a steady state, a rapid decrease in DNA binding activity could be observed (Figure 4G). Use of other ATP hydrolysis quenchers (phosphatase and EDTA) showed similar effects as apyrase on inhibition of p53 function (Figure 4G). This indicates that, under steady-state conditions, continuous rounds of ATP hydrolysis cycles by Hsp70 and Hsp90 are required.

The p53-DBD Forms a Multi-chaperone Complex with Hsp70, Hsp40, HOP, and Hsp90 Sensitive to Bag-1 Dissociation

To further define the role of Hop, *in vitro* pull-downs of the GFP-tagged p53-DBD were performed. The GFP-tagged p53-DBD behaved similarly as the untagged p53-DBD (Figures S3A and S3B). Hsp90 could be efficiently incorporated into the Hsp70-40-p53-DBD complex only in the presence of Hop (Figures S3C and S3D). Interestingly, addition of Hop and Hsp90 to the Hsp70-40-p53 complex released some Hsp70 and Hsp40 from the p53-DBD, with a major effect on the presence of Hsp40 (Figures S3C and S3D). This suggests that Hop primarily evicts Hsp40 from the complexes, enabling Hsp90 to take over the p53-DBD. We separated the complexes via SEC and found that addition of Hop and Hsp90 to the p53-DBD-Hsp70-Hsp40 complex resulted in elution of a separate peak with simultaneous reduction of the p53-DBD-Hsp70-Hsp40 complex peak (Figure S3E). The SDS-PAGE of the corresponding fractions of this peak (Figures S3E and S3F, top, red curve) showed that Hsp70, Hop, and Hsp90 could be located in the same fractions as the p53-DBD and Hsp40 (Figure S3F, center and bottom), implying that Hsp40 is part of the complex. This was also confirmed using fluorescently labeled Hsp40* in AUC experiments (Figure S3G). Hsp40* alone sediments at 5 S, and the S value increases to 16 S after addition of Hsp70/ATP and the p53-DBD and further to 18.5 S upon inclusion of Hop and Hsp90, indicating complex formation between Hsp40*, Hsp70, the p53-DBD, Hop and Hsp90 (Figure S3G). Together, these findings suggest the presence of an intermediate complex of Hsp70-Hsp40-p53-DBD-Hop-Hsp90, in which the p53-DBD is handed over from Hsp70 to Hsp90.

The reactivation of the p53-DBD initiated by Hop and Hsp90 exhibited a lag phase followed by slow kinetics (0.03 min^{-1}), and full recovery of DNA binding was not obtained (Figure 4B). Hence, we tested the effects of other Hsp90 co-chaperones like p23 (Cox and Johnson, 2011; Freeman et al., 1996), PPIases (Cyp40 or FKBP51 and 52) (Chen et al., 1998), or Aha1 (Panaretou et al., 2002; Retzlaff et al., 2010). None of them had a significant effect on the kinetics or the total recovery of p53-DBD DNA binding activity in addition to Hop and Hsp90 (Figure S4A). We reasoned that inefficient release of the p53-DBD from the Hsp70-Hsp40 complex may cause incomplete DNA binding recovery by Hop and Hsp90. To test this, we added Bag-1, a nucleotide exchange factor (NEF) shown previously to regulate interactions of Hsp70 with p53 (King et al., 2001). We found that p53-DBD and Bag-1 do not interact in AUC experiments. Thus, Bag-1 does not directly affect p53. At 25°C, Bag-1 induced substantial release of the p53-DBD from the Hsp70-40 complex (Figure 4B), resulting in ~70% recovery (Figure 4C). Furthermore, by AUC, we showed directly that Bag-1 releases the p53-DBD* from the Hsp70-Hsp40 complex, as evidenced by the increase in the free p53-DBD* peak (Figures 4D, green, and 4E). In the pull-down experiments, Bag-1 evicted both Hsp70 and Hsp40 from the complex (Figures S3C and S3D).

Supplementing the Hop/Hsp90-mediated reaction with Bag-1 increased the DNA binding recovery from 50% to 70% (Figure 4C). Furthermore, Bag-1 increased the rate of recovery from 0.03 min^{-1} to 0.33 min^{-1}, accelerating the Hsp90-mediated p53 reactivation 10-fold (Figure 4B). AUC analysis showed that Bag-1 is important for recovering the p53-DBD from the stalled Hsp70-40-Hop-Hsp90 complex by dissociating this intermediate complex and releasing the p53-DBD* (Figures 4D, violet, and 4E). Pull-downs revealed pronounced displacement of Hsp40 besides a reduction in Hsp70 and Hop levels (Figures S3C and S3D) when Bag-1, Hop, and Hsp90 were added together. This observation further supports the presence of a Hsp70-40-p53-DBD-Hop-Hsp90 complex, in agreement with the notion that Bag-1 does not dissociate Hsp40-p53 or Hsp90-Hop-p53 complexes but acts only on Hsp70-containing complexes as a nucleotide exchange factor of Hsp70.

Even in the presence of all components (Hsp70, Hsp40, Bag-1, Hop, and Hsp90), the p53-DBD was still present in complexes (Figure 4E), and, hence, complete recovery of DNA binding was not observed (Figures 4B and 4C). As in the absence of Bag-1, in its presence, Hsp90 co-chaperones had no effects either (Figure S4B). Other possible reasons for the incomplete DNA binding recovery could be aggregation of the p53-DBD during the assay or its rebinding to Hsp70/40. Aggregation was ruled out because no increase in light scattering was observed at 25°C (Figure S4C). Therefore, what we see is the sum of the effects of Hsp70 and Hsp90, which both constantly act on the p53 population and regulate its conformation.

Hsp90 and Hop Promote the Folded Conformation of the p53-DBD

To directly probe the p53-DBD conformation in various complexes, we performed spFRET experiments with the DBD

(C) Bar chart comparing the regain of DNA binding activity by the p53-DBD under the chaperone combinations tested in (B). The percentages are calculated from the final change in anisotropy obtained in (B), with the change of anisotropy for the native p53-DBD set as 100%. The mean and SD of three independent experiments are plotted.

(D) AUC analysis of the complexes formed between the p53-DBD and different chaperones as indicated. See also Figure S3.

(E) Fraction of the p53-DBD* bound under different conditions, calculated from the sedimentation velocity curves in (D). The values represent mean with SD from three independent experiments.

(F) DNA binding recovery of Hsp70-40-bound p53-DBD monitored after addition of Hop and Hsp90 (pink), in the presence of 100 μM radicicol (red), with Hsp90 E42A (blue), or with Hsp90 D88A (olive).

(G) The DNA binding recovery of Hsp70-40-bound p53-DBD monitored by a change in anisotropy of the labeled DNA after addition of Hop and Hsp90 (pink), in the presence of phosphatase (brown) or 10 mM EDTA (cyan), or apyrase added from the start (black) or under steady-state conditions (olive).

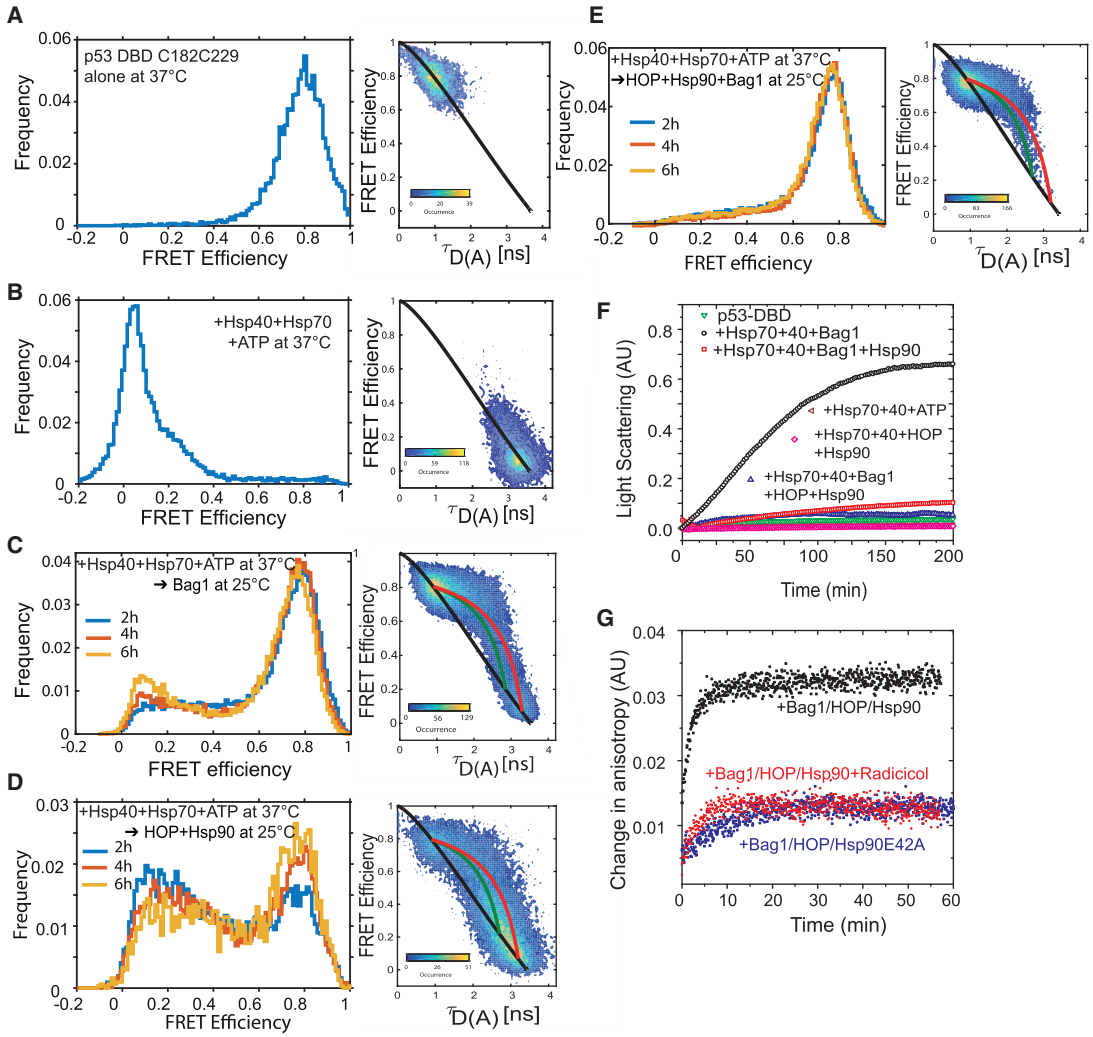

Figure 5. Hsp90 Chaperone p53-DBD Conformation

(A–E) SpFRET histograms and donor lifetime plots for the p53-DBD C182C229 sensor when Bag-1/Hop/Hsp90, Hop/Hsp90, or Bag-1 was added to the preformed complex of p53-DBD-Hsp40-Hsp70 at the start of the measurement. Reactions with Bag-1 alone, Hop+Hsp90, and Bag-1+Hop+Hsp90 were followed for 6 h, and the FRET efficiency histogram (averaged over a 1-h interval) was determined every 2 h. The 2-h time point (i.e., 1–2 h) is shown in blue, 3–4 h in orange, and 5–6 h in yellow. A 2D histogram for FRET efficiency versus donor lifetime in the presence of the acceptor ($\tau_{D(A)}$) showing the conformational dynamics of the p53-DBD sensor under different conditions is also included. See also Figure S8.

(A) p53-DBD C182C229 measured alone at 37°C.

(B) p53-DBD C182C229 with Hsp40, Hsp70, and ATP at 37°C.

(C) Bag-1 added to (B). See also Figure S4F.

(D) Hop and Hsp90 added to (B). See also Figures S4D and S4E.

(E) Bag-1, Hop, and Hsp90 added to (B).

(F) Change in light scattering with time at 37°C; reactions were generated as in Figure 4A, except that DNA was absent. The kinetics of light scattering at 350 nm were recorded after addition of Bag-1 (black); Bag-1 and Hsp90 (red); Bag-1, Hop, and Hsp90 (blue); and Hop and Hsp90 (pink) to the Hsp70-Hsp40-bound p53-DBD (brown). The p53-DBD alone is shown in green.

(G) DNA-binding recovery kinetics of Hsp70-40-bound p53-DBD monitored after adding Bag-1, Hop, and Hsp90 in the absence (black) and presence of 100 μM radicicol (red) or in the presence of Bag-1, Hop, and Hsp90E42A (blue).

C182C229 sensor. As shown above, the p53 DBD is completely unfolded at 37°C in the presence of Hsp40, Hsp70, and ATP (Figures 3D and 5B). When Bag-1 was added to this preformed complex at 25°C, most of the p53 molecules could attain the native conformation with a FRET efficiency of ~0.8 (Figure 5C, left). Two hours after addition of Bag-1, a small population of molecules was found to be in the unfolded conformation with FRET efficiencies distributed over a broad range from 0 to 0.4

(Figure 5C, left). This conformational heterogeneity might originate from the presence of both Hsp70-Hsp40-bound and Bag-1-released conformations of the p53-DBD and account for the incomplete DNA binding recovery shown in Figure 4B. Interestingly, an increase in the lowest FRET population was observed over a period of 6 h (Figure 5C, left). This suggests rebinding of Hsp40 and Hsp70 to p53-DBD molecules that were previously released by Bag-1 from Hsp70-Hsp40, shifting the population of the p53-DBD to the unfolded states. Again, the FRET efficiency versus donor lifetime plots clearly indicate dynamic transitions between low FRET and high FRET states (Figure 5C, right). Hence, the p53-DBD is very dynamic, undergoing constant fluctuations between folded and unfolded states.

To gain further structural insights into the Hop/Hsp90- and Bag-1-mediated regain of p53-DBD activity, we performed spFRET measurements after addition of Hop/Hsp90 at 25°C with and without Bag-1 to the preformed Hsp70-40-p53-DBD C182C229 complex at 37°C (Figures 5D and 5E). Hop/Hsp90 addition led to regain of the native structure for 50% of the molecules (Figure 5D, left). The other half was still mostly unfolded. Over time, an increase in the high FRET population was observed, indicating folding of the previously Hsp70-Hsp40-bound p53-DBD C182C229 by Hsp90 (Figure 5D, left). spFRET experiments with the Hsp70-Hsp40-bound p53-DBD C182C229 in the presence of Hop and Hsp90 alone showed that, in both cases, ~10%–20% of the DBD population reached the native conformation, whereas the rest of the molecules stayed in the unfolded conformations (Figures S4D and S4E). In addition, no further change in the FRET populations was observed with time (Figures S4D and S4E). SpFRET directly shows that Hop is needed for efficient transfer of the p53-DBD from Hsp70 to Hsp90, which then chaperones its refolding to the native state. Again, dynamic transitions are observed between the folded and unfolded states (Figure 5D, right). The population of the native state observed in the case of Hop and Hsp90 corresponds well with the DNA binding recovery of the p53-DBD observed in anisotropy experiments (Figures 4B and 4C). Nearly half of the p53-DBD molecules are in non-native conformations, displaying a broad range of FRET efficiencies from 0–0.4 in the presence of Hop/Hsp90 (Figure 5D). These populations comprise the p53-DBD bound to the Hsp70-Hsp40-Hop-Hsp90 complexes, as seen by AUC, where ~60% of the p53-DBD was still present in complexes after addition of Hop and Hsp90 to Hsp70-40-p53-DBD* (Figures 4D and 4E).

After addition of Bag-1 to the measurements above with Hop/Hsp90 (Figure 5E), virtually all p53-DBD C182C229 molecules attained the native conformation, with a small fraction of molecules still displaying dynamic transitions to the unfolded states on the millisecond timescale (Figure 5E, right). Inclusion of Bag-1 together with Hop and Hsp90 significantly shifted the unfolded populations seen with Hop/Hsp90 (Figure 5D) to the folded population (Figure 5E). The fraction of the folded molecules obtained by the addition of Bag-1/Hop/Hsp90 to Hsp70-bound p53-DBD C182C229 correlates with the fraction of DNA-binding p53-DBD molecules observed in the anisotropy experiments (Figure 4C). Strikingly, in contrast to the Bag-1-alone measurement (Figure 5C, left), there was no change in the FRET populations over a period of 6 h when Bag-1 was added along with Hop and Hsp90 (Figure 5E, left). This directly shows that the conformation released from Hsp90 is folded and does not rebind to Hsp70. We investigated the conformation using limited proteolysis but could not detect changes in the protease sensitivity of the p53-DBD between conformations released in the presence or absence of Hsp90 (Figure S4F). It is possible that the conformational differences are not significantly pronounced to be differentiated by trypsin digestion. We reasoned that, if the Bag-1-released conformation is non-native, then it should be prone to aggregation at higher temperatures. To test this idea, we first formed the Hsp70-Hsp40-p53-DBD complex and then added Bag-1 and monitored light scattering at 37°C. Interestingly, addition of Bag-1 resulted in an increase in light scattering, which was not observed in the presence of Hop and Hsp90 (Figure 5F). These results clearly indicate that the conformation of p53-DBD released from Hsp70 by Bag-1 is different from the p53-DBD conformation released directly from Hsp90. Taken together, Bag-1 and Hop/Hsp90 together allow the p53-DBD to refold to the native conformation.

The ATPase activity of Hsp90 is essential for the recovery of DNA binding by p53-DBD (Figure 4F). To test whether Hsp90 acts as a passive release factor for Hsp70 or as an active chaperone, we took advantage of the Bag-1-mediated release of p53-DBD from Hsp70. By adding Bag-1 along with Hop and Hsp90, we could decouple the release function from the chaperone function of Hsp90. When we performed the experiment in the presence of the Hsp90 E42A mutant, which is deficient in ATPase activity, the recovery of DNA binding was inefficient, even in the presence of Bag-1 (Figure 5G). Radicicol inhibition of Hsp90 also produced a similar outcome (Figure 5G). This implies that p53-DBD reactivation by Hsp90 is not just limited by its release from Hsp70 but also requires active chaperoning by Hsp90.

Full-Length p53 Recapitulates the Behavior of the p53-DBD

To test the physiological relevance of our findings for the p53-DBD, we investigated the behavior of full-length p53 (p53-FL) toward Hsp70 and Hsp90 interactions. Identical to the behavior of the p53-DBD, p53-FL was inactive in DNA binding in the presence of Hsp70-40 and ATP at 37°C (Figure 6A). Furthermore, like for the p53-DBD, Hop and Hsp90 were able to restore the DNA binding activity of p53-FL at 25°C to 50% (Figure 6B), and the ATPase activity of Hsp90 was also found to be essential for p53-FL reactivation (Figures 6C and 6D). Unexpectedly, however, upon addition of Bag-1 to the Hsp70-40-p53-FL complex, no recovery of DNA binding was obtained, and this was also true when Bag-1, Hop, and Hsp90 were added together (Figure 6B). We further analyzed the complexes of Atto-488-labeled p53-FL* with different chaperones in AUC. Labeling did not compromise p53-FL DNA binding activity (Figure S5A). AUC analysis revealed that Bag-1 was efficient in releasing p53-FL* from the Hsp70-40 complex (Figures 6E and 6F), which means that the inability of p53-FL to bind DNA after Bag-1 addition is not due to inefficient release of p53-FL from the Hsp70 complex. Similar to the DBD, we observed aggregation of p53-FL at 37°C after Bag-1 addition to the Hsp70-40-p53-FL complex (Figure S5B). This indicates that the conformation of p53-FL released by Bag-1 was unable to fold spontaneously and aggregated, resulting in a loss of

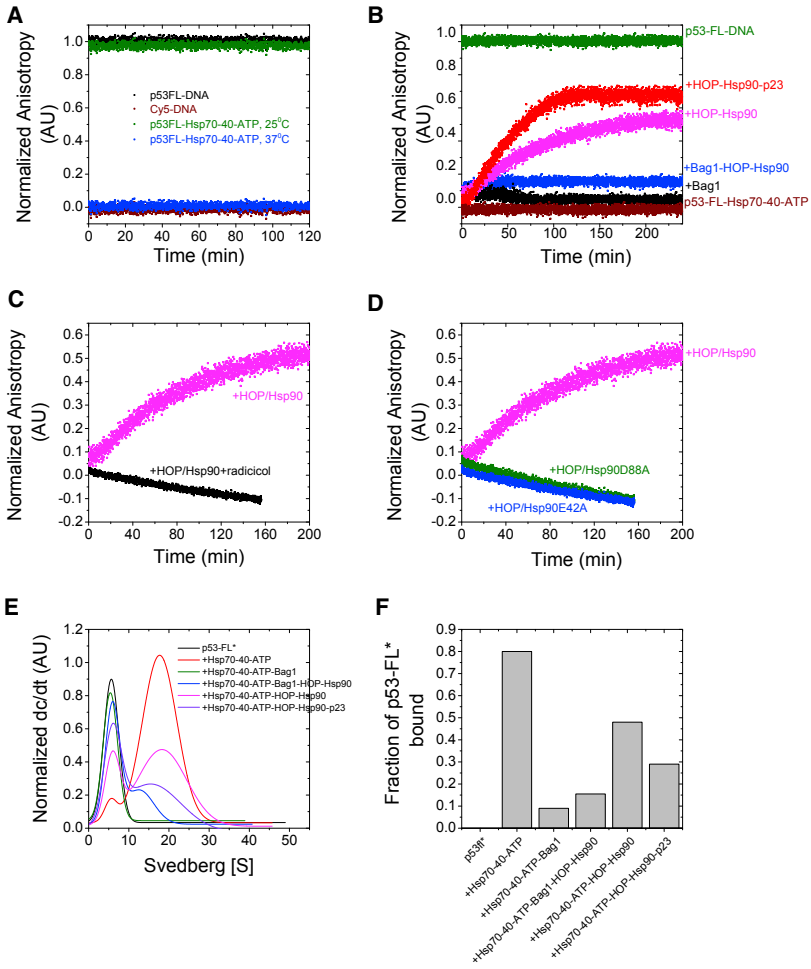

Figure 6. Full-Length p53 Is Chaperoned by Hsp70 and Hsp90, Similar to the DBD

(A) Effects of chaperones on DNA binding of p53-FL-DNA, measured by anisotropy. p53-FL (black) was incubated with Hsp70, Hsp40, and ATP at 37°C (blue) or 25°C (olive). The anisotropy of labeled DNA is shown in brown.

(B) DNA binding recovery experiments carried out with p53-FL, similar to the experimental scheme in Figure 4A. Shown are fluorescence anisotropy kinetics showing relative anisotropy of Cy5-labeled p21 DNA when Bag-1 (black), Bag-1 plus Hsp90 and Hop (blue), Hsp90 and Hop (pink), and Hop, Hsp90, and p23 (red) were added to the Hsp70-Hsp40-p53-FL complex (brown). The anisotropy of p53-FL and DNA is shown in olive and set to 1. The other values are normalized to this value. See also Figure S5B.

(C) DNA binding of p53-FL in the presence of Hop and Hsp90 (pink) and 100 μM radicicol (black).

(D) Effects of the Hsp90 ATPase mutants E42A (blue) and D88A (olive) on the DNA binding recovery of Hsp70-bound p53-FL in the presence of Hop.

(E) AUC analysis of complexes formed between Atto-488-labeled p53-FL* and different chaperone combinations, as indicated. See also Figure S5A.

(F) Fraction of p53-FL* bound under different conditions, calculated from the sedimentation velocity curves in (E).

recovery of DNA binding. Strikingly, when we tested other Hsp90 co-chaperones, we found that the late-acting Hsp90 co-chaperone p23 (Freeman et al., 2000) enhanced the Hop- and Hsp90- mediated DNA binding recovery of Hsp70-bound p53-FL. Inclusion of p23 along with Hop and Hsp90 resulted in ~70% DNA binding recovery with accelerated kinetics (Figure 6B). This suggests that, for p53-FL, p23 releases p53-FL* from the Hsp70-40-Hop-Hsp90 complex (Figures 6E and 6F).

DISCUSSION

p53 is dynamic and exists in equilibrium between different conformers (Cañadillas et al., 2006; Hainaut et al., 1995). Stresses like heat shock affect the active conformation of p53, and complexes with Hsp70, Hsp40, and Hsp90 are increasingly observed. In this study, we reconstituted interactions of p53 with the Hsp70 and Hsp90 chaperone systems *in vitro*. The prevailing notion has been that Hsp70 and Hsp90 are important for *de novo* folding of p53 or its stabilization under stress conditions (Fourie et al., 1997; Walerych et al., 2009; Wang and Chen, 2003; Wawrzynow et al., 2018; Zylicz et al., 2001). Our results suggest a different picture, in which Hsp70 and Hsp90 constantly control the conformation of p53 in an ATP-dependent manner. We observed formation of large complexes of the p53-DBD with Hsp40 and Hsp70. The finding that Hsp70-Hsp40 forms stable complexes with the structured DNA binding domain of p53 even under cycling conditions (i.e., in the presence of ATP), was surprising. The presence of several Hsp70 binding sites in the DBD of p53 (Fourie et al., 1997) may contribute to the apparent stable interaction. The Hsp70-Hsp40-p53 complexes formed at 25°C are reversible and able to bind DNA; however, complexes formed at 37°C are not (Figure 2A). These results imply that Hsp70 keeps transcription factors like p53 in an inactive, non-DNA binding state in the absence of stress signals by forming stable complexes. In these complexes, p53 is globally unfolded by Hsp70 and Hsp40. SpFRET revealed that the p53-DBD is inherently dynamic and that higher temperatures increase these dynamics, providing a rational explanation for why p53 interactions with Hsp70 and Hsp90 increase under stress conditions. Interestingly, Hsp70/40/ATP is able to dissociate the p53-DBD from DNA at 37°C but not at 25°C. Hsp70 may trap the unfolded segments of p53 at physiological temperature and promote its dissociation from DNA. This suggests that destabilization of the p53-DBD and trapping non-native states by Hsp70/40 at elevated temperatures together result in its complete unfolding. In this context, when the supply of Hsp70 becomes limited in the cell during stress, p53 can be released from the chaperone control and become active. Additional

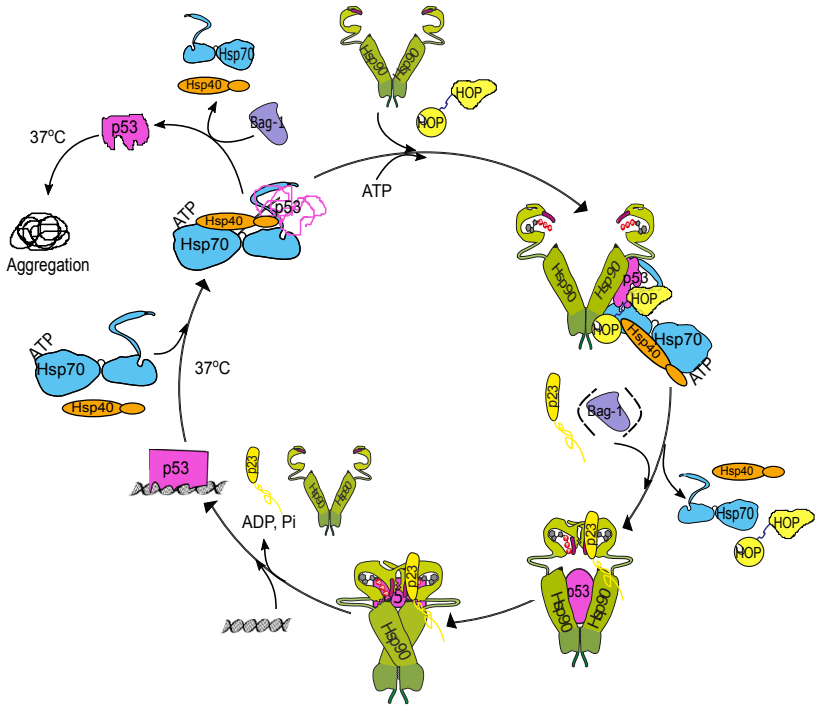

Figure 7. Model of Chaperoning of p53 by Hsp70 and Hsp90

Under physiological conditions, Hsp70, together with its co-chaperone Hsp40 and ATP, binds to p53 and unfolds it completely. In these complexes, p53 is incapable of binding DNA unless it can be dissociated from Hsp70. Bag-1 is able to release p53 from the Hsp70-Hsp40 complex but in a conformational state that is aggregation-prone. Hsp90 channels p53 on the productive folding pathway. Hsp90 is efficiently incorporated by Hop into a multi-chaperone complex with p53, Hsp70, and Hsp40. For p53-FL, p23 is able to dissociate this complex, allowing efficient transfer of p53-FL to Hsp90. For the p53-DBD, Bag-1 (in brackets) assists with the transfer to Hsp90, evicting primarily Hsp40, Hsp70, and, to some degree, Hop from the multi-chaperone complex. Upon ATP hydrolysis and in the presence of DNA, p53 is released from Hsp90 in a folded and DNA-binding active conformation. Cycling of p53 between Hsp70 and Hsp90 maintains its DNA binding activity under physiological or stress conditions and may also affect the duration of p53-mediated transcription.

factors and post-translational modifications (PTMs) may further modulate the Hsp70-p53 interaction. By forming stable complexes with the p53-DBD at higher temperatures, Hsp70 will also protect p53 under conditions where it is prone to unfold and aggregate.

Bag-1 and Hsp90 act in the p53-DBD reactivation pathway to achieve its refolding from the Hsp70-inhibited state. In the presence of Hsp90, Hop, and Bag-1, a complete shift toward a folded population is observed using spFRET. In the absence of Bag-1, FRET populations are distributed over a wide range between unfolded and folded states. At present, we do not know whether the folded states we observe in spFRET in the presence of Hsp90 are still bound to Hsp90 or released. In this context, earlier studies have suggested that the chaperone action of Hsp90 results in partial unfolding of p53 (Park et al., 2011; Rüdiger et al., 2002; Walerych et al., 2010). However, both our functional assay and spFRET analysis show that the conformation of p53-DBD released in the presence of Hsp90 does not aggregate and does not readily rebind Hsp70, suggesting that Hsp90 promotes folding of the p53-DBD, as also observed by Boysen et al. (2019) in this issue of *Molecular Cell*. This, together with our observation that ATP hydrolysis by Hsp90 is needed for p53-DBD reactivation, demonstrates that active chaperoning of Hsp90 is required. Interestingly, considerable conformational dynamics are seen in spFRET between Hsp70-bound and Hsp90-bound or released states of the p53-DBD. This implies that the conformation of the p53-DBD constantly fluctuates between unfolded and folded states and that both chaperone machineries are constantly working on the p53-DBD because the overall process is strictly dependent on ATP supply. Hence, the relative concentrations or vacant Hsp70 and Hsp90 present at a particular time in the cell will dictate which conformation prevails.

For p53-FL, the same pattern of Hsp70 and Hsp90 effects are seen. Strikingly, however, in the case of p53-FL, Bag-1 releases the protein in a conformation unable to bind DNA. Thus, p53-FL stringently requires Hsp90 to regain its native conformation, whereas the p53-DBD could refold spontaneously after release. Unexpectedly, in the presence of Hop and Hsp90, Bag-1 was also insufficient to support the regain of DNA binding activity of the p53-FL. This has implications for p53 homeostasis in cells, where the decision between degradation and folding involves competition between Hop and Bag-1, which can simultaneously associate with Hsp70 and the proteasome (Höhfeld et al., 2001). A possible reason for this could be competition between Bag-1 and Hop for Hsp70 compromising Hop binding, especially for p53-FL, and, hence, resulting in inefficient transfer of p53 to Hsp90. In pull-down experiments with the p53-DBD, the amount of Hop was reduced upon Bag-1 addition, but Hop could still bind to Hsp70 in the presence of Bag-1 and contribute to efficient Hsp90 transfer. The situation is different for p53-FL where, potentially because of its size and complexity, Hop binding in the presence of Bag-1 is disfavored. Interestingly, only in the context of p53-FL, the Hsp90 co-chaperone p23 leads to an increase in DNA binding of p53-FL, performing a similar role as Bag-1 for the p53-DBD.

In summary, Hsp70 and Hsp90 act together in a relay system, with opposing effects on p53 conformation and with co-chaperones regulating different steps of the pathway. In our model (Figure 7), Hsp70/40/ATP generate and stably bind an unfolded conformation of p53. From this complex, a partially folded p53 is transferred to Hsp90. Whether p53 achieves its native conformation when still bound to Hsp90 or in solution after release from

Hsp90 cannot be experimentally resolved with the current approaches. Nevertheless, the interaction with Hsp90 plays an important role in activation of p53 by allowing it to reach to its functionally competent state, which is otherwise not accessible. The availability of chaperones and other factors, such as Mdm2 and CHIP (Wawrzynow et al., 2018), as well as PTMs may further modulate this scheme and influence the decision between folding and degradation.

The opposing effects Hsp70 and Hsp90 have on p53 conformation could also be exploited *in vivo* for regulating the DNA binding activity of p53. It is tempting to speculate that, under physiological conditions, the ability of Hsp70 to unfold and release p53 from DNA is important to tightly control the duration of p53-induced transcription. Under stress conditions, Hsp70 becomes involved in additional processes and, as a result, p53-dependent transcription would increase. Upon return of normal conditions, Hsp70 can again remove p53 from DNA. The opposing action of Hsp70 and Hsp90 on p53 conformation and activity would be beneficial in the cellular context because this would maintain p53 in a repressed but DNA binding-competent state. This allows p53 to quickly respond to changes in cellular conditions without being irreversibly denatured under physiological or stressed conditions, where it is unstable. Thus, cycling p53 through these chaperone machineries enables it to constantly remodel its structure and to link its activity status to the general stress state of the cell.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- METHOD DETAILS
  - Protein Expression and Purification
  - Fluorescence Anisotropy Measurements
  - Analytical Ultracentrifugation (AUC)
  - p53-DBD-GFP *In Vitro* Pull-Down
  - Size Exclusion Chromatography
  - Limited Proteolysis
  - Aggregation Assay
  - Electrophoretic Mobility Shift Assay
  - p53 Labeling
  - Single-Pair FRET Experiments
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.molcel.2019.03.026.

ACKNOWLEDGMENTS

We thank Lena Voith von Voithenberg and Daniela Krader for assistance with spFRET and Anders Barth for support with the Monte Carlo burst analysis. We gratefully acknowledge financial support from the Alexander von Humboldt Foundation (to V.D.) and the Deutsche Forschungsgemeinschaft (DFG) (SFB 1035, Project A3, A11 to J.B. and D.C.L.). D.C.L. also acknowledges support from the Center for NanoScience and the BioImaging Network.

AUTHOR CONTRIBUTIONS

V.D. designed experiments, generated the constructs, performed purification and fluorescence labeling, and conducted all biochemical experiments. J.L. and D.A.R. performed the AUC experiments. G.A. performed and analyzed the spFRET experiments. V.D. analyzed the results of the biochemical experiments and performed AUC data analysis. J.B. and D.C.L. designed experiments, oversaw the project, and analyzed the results. V.D., G.A., J.B., and D.C.L. wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: August 2, 2018
Revised: December 17, 2018
Accepted: March 22, 2019
Published: April 23, 2019

REFERENCES

Alvira, S., Cuéllar, J., Röhl, A., Yamamoto, S., Itoh, H., Alfonso, C., Rivas, G., Buchner, J., and Valpuesta, J.M. (2014). Structural characterization of the substrate transfer mechanism in Hsp70/Hsp90 folding machinery mediated by Hop. Nat. Commun. *5*, 5484.

Blagosklonny, M.V., Toretsky, J., Bohen, S., and Neckers, L. (1996). Mutant conformation of p53 translated in vitro or in vivo requires functional HSP90. Proc. Natl. Acad. Sci. USA *93*, 8379–8383.

Boczek, E.E., Reefschläger, L.G., Dehling, M., Struller, T.J., Häusler, E., Seidl, A., Kaila, V.R., and Buchner, J. (2015). Conformational processing of oncogenic v-Src kinase by the molecular chaperone Hsp90. Proc. Natl. Acad. Sci. USA *112*, E3189–E3198.

Boysen, M., Kityk, R., and Mayer, M.P. (2019). Hsp70- and Hsp90-mediated regulation of the conformation of p53 DNA binding domain and p53 cancer variants. Molecular Cell *74*. Published online April 23, 2019. https://doi.org/10.1016/j.molcel.2019.03.026.

Cañadillas, J.M.P., Tidow, H., Freund, S.M.V., Rutherford, T.J., Ang, H.C., and Fersht, A.R. (2006). Solution structure of p53 core domain: structural basis for its instability. Proc. Natl. Acad. Sci. USA *103*, 2109–2114.

Chen, S., Sullivan, W.P., Toft, D.O., and Smith, D.F. (1998). Differential interactions of p23 and the TPR-containing proteins Hop, Cyp40, FKBP52 and FKBP51 with Hsp90 mutants. Cell Stress Chaperones *3*, 118–129.

Clerico, E.M., Tilitsky, J.M., Meng, W., and Gierasch, L.M. (2015). How hsp70 molecular machines interact with their substrates to mediate diverse physiological functions. J. Mol. Biol. *427*, 1575–1588.

Cox, M.B., and Johnson, J.L. (2011). The role of p23, Hop, immunophilins, and other co-chaperones in regulating Hsp90 function. Methods Mol. Biol. *787*, 45–66.

Dehner, A., Klein, C., Hansen, S., Müller, L., Buchner, J., Schwaiger, M., and Kessler, H. (2005). Cooperative binding of p53 to DNA: regulation by protein-protein interactions through a double salt bridge. Angew. Chem. Int. Ed. Engl. *44*, 5247–5251.

Echtenkamp, F.J., and Freeman, B.C. (2012). Expanding the cellular molecular chaperone network through the ubiquitous cochaperones. Biochim. Biophys. Acta *1823*, 668–673.

Fourie, A.M., Hupp, T.R., Lane, D.P., Sang, B.-C., Barbosa, M.S., Sambrook, J.F., and Gething, M.-J.H. (1997). HSP70 binding sites in the tumor suppressor protein p53. J. Biol. Chem. *272*, 19471–19479.

Freeman, B.C., Toft, D.O., and Morimoto, R.I. (1996). Molecular chaperone machines: chaperone activities of the cyclophilin Cyp-40 and the steroid aporeceptor-associated protein p23. Science *274*, 1718–1720.

Freeman, B.C., Felts, S.J., Toft, D.O., and Yamamoto, K.R. (2000). The p23 molecular chaperones act at a late step in intracellular receptor action to differentially affect ligand efficacies. Genes Dev. *14*, 422–434.

Frydman, J. (2001). Folding of newly translated proteins in vivo: the role of molecular chaperones. Annu. Rev. Biochem. *70*, 603–647.

Genest, O., Wickner, S., and Doyle, S.M. (2019). Hsp90 and Hsp70 chaperones: Collaborators in protein remodeling. J. Biol. Chem. *294*, 2109–2120.

Graf, C., Stankiewicz, M., Kramer, G., and Mayer, M.P. (2009). Spatially and kinetically resolved changes in the conformational dynamics of the Hsp90 chaperone machine. EMBO J. *28*, 602–613.

Hagn, F., Lagleder, S., Retzlaff, M., Rohrberg, J., Demmer, O., Richter, K., Buchner, J., and Kessler, H. (2011). Structural analysis of the interaction between Hsp90 and the tumor suppressor protein p53. Nat. Struct. Mol. Biol. *18*, 1086–1093.

Hainaut, P., and Milner, J. (1992). Interaction of heat-shock protein 70 with p53 translated in vitro: evidence for interaction with dimeric p53 and for a role in the regulation of p53 conformation. EMBO J. *11*, 3513–3520.

Hainaut, P., Rolley, N., Davies, M., and Milner, J. (1995). Modulation by copper of p53 conformation and sequence-specific DNA binding: role for Cu(II)/Cu(I) redox mechanism. Oncogene *10*, 27–32.

Hayes, D.B., and Stafford, W.F. (2010). SEDVIEW, real-time sedimentation analysis. Macromol. Biosci. *10*, 731–735.

Höhfeld, J., Cyr, D.M., and Patterson, C. (2001). From the cradle to the grave: molecular chaperones that may choose between folding and degradation. EMBO Rep. *2*, 885–890.

Huang, F., Rajagopalan, S., Settanni, G., Marsh, R.J., Armoogum, D.A., Nicolaou, N., Bain, A.J., Lerner, E., Haas, E., Ying, L., and Fersht, A.R. (2009). Multiple conformations of full-length p53 detected with single-molecule fluorescence resonance energy transfer. Proc. Natl. Acad. Sci. USA *106*, 20758–20763.

Joerger, A.C., and Fersht, A.R. (2007). Structure-function-rescue: the diverse nature of common p53 cancer mutants. Oncogene *26*, 2226–2242.

Johnson, B.D., Schumacher, R.J., Ross, E.D., and Toft, D.O. (1998). Hop modulates Hsp70/Hsp90 interactions in protein folding. J. Biol. Chem. *273*, 3679–3686.

Kalinin, S., Valeri, A., Antonik, M., Felekyan, S., and Seidel, C.A.M. (2010). Detection of structural dynamics by FRET: a photon distribution and fluorescence lifetime analysis of systems with multiple states. J. Phys. Chem. B *114*, 7983–7995.

Kalinin, S., Peulen, T., Sindbert, S., Rothwell, P.J., Berger, S., Restle, T., Goody, R.S., Gohlke, H., and Seidel, C.A. (2012). A toolkit and benchmark study for FRET-restrained high-precision structural modeling. Nat. Methods *9*, 1218–1225.

Karagöz, G.E., Duarte, A.M., Akoury, E., Ippel, H., Biernat, J., Morán Luengo, T., Radli, M., Didenko, T., Nordhues, B.A., Veprintsev, D.B., et al. (2014). Hsp90-Tau complex reveals molecular basis for specificity in chaperone action. Cell *156*, 963–974.

Kastenhuber, E.R., and Lowe, S.W. (2017). Putting p53 in Context. Cell *170*, 1062–1078.

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. USA *111*, 13355–13360.

King, F.W., Wawrzynow, A., Höhfeld, J., and Zylicz, M. (2001). Co-chaperones Bag-1, Hop and Hsp40 regulate Hsc70 and Hsp90 interactions with wild-type or mutant p53. EMBO J. *20*, 6297–6305.

Kirschke, E., Goswami, D., Southworth, D., Griffin, P.R., and Agard, D.A. (2014). Glucocorticoid receptor function regulated by coordinated action of the Hsp90 and Hsp70 chaperone cycles. Cell *157*, 1685–1697.

Kudryavtsev, V., Sikor, M., Kalinin, S., Mokranjac, D., Seidel, C.A., and Lamb, D.C. (2012). Combining MFD and PIE for accurate single-pair Förster resonance energy transfer measurements. ChemPhysChem *13*, 1060–1078.

Lapidus, L.J., Steinbach, P.J., Eaton, W.A., Szabo, A., and Hofrichter, J. (2002). Effects of Chain Stiffness on the Dynamics of Loop Formation in Polypeptides. Appendix: Testing a 1-Dimensional Diffusion Model for Peptide Dynamics. J. Phys. Chem. B *106*, 11628–11640.

Li, J., Soroka, J., and Buchner, J. (2012). The Hsp90 chaperone machinery: conformational dynamics and regulation by co-chaperones. Biochim. Biophys. Acta *1823*, 624–635.

Lorenz, O.R., Freiburger, L., Rutz, D.A., Krause, M., Zierer, B.K., Alvira, S., Cuéllar, J., Valpuesta, J.M., Madl, T., Sattler, M., and Buchner, J. (2014). Modulation of the Hsp90 chaperone cycle by a stringent client protein. Mol. Cell *53*, 941–953.

Mayer, M.P., and Bukau, B. (2005). Hsp70 chaperones: cellular functions and molecular mechanism. Cell. Mol. Life Sci. *62*, 670–684.

Morán Luengo, T., Kityk, R., Mayer, M.P., and Rüdiger, S.G.D. (2018). Hsp90 Breaks the Deadlock of the Hsp70 Chaperone System. Mol. Cell *70*, 545–552.e9.

Müller, L., Schaupp, A., Walerych, D., Wegele, H., and Buchner, J. (2004). Hsp90 regulates the activity of wild type p53 under physiological and elevated temperatures. J. Biol. Chem. *279*, 48846–48854.

Muller, P., Hrstka, R., Coomber, D., Lane, D.P., and Vojtesek, B. (2008). Chaperone-dependent stabilization and degradation of p53 mutants. Oncogene *27*, 3371–3383.

Olivier, M., Hollstein, M., and Hainaut, P. (2010). TP53 mutations in human cancers: origins, consequences, and clinical use. Cold Spring Harb. Perspect. Biol. *2*, a001008.

Panaretou, B., Siligardi, G., Meyer, P., Maloney, A., Sullivan, J.K., Singh, S., Millson, S.H., Clarke, P.A., Naaby-Hansen, S., Stein, R., et al. (2002). Activation of the ATPase activity of hsp90 by the stress-regulated cochaperone aha1. Mol. Cell *10*, 1307–1318.

Park, S.J., Borin, B.N., Martinez-Yamout, M.A., and Dyson, H.J. (2011). The client protein p53 adopts a molten globule-like state in the presence of Hsp90. Nat. Struct. Mol. Biol. *18*, 537–541.

Pullen, L., and Bolon, D.N. (2011). Enforced N-domain proximity stimulates Hsp90 ATPase activity and is compatible with function in vivo. J. Biol. Chem. *286*, 11091–11098.

Retzlaff, M., Hagn, F., Mitschke, L., Hessling, M., Gugel, F., Kessler, H., Richter, K., and Buchner, J. (2010). Asymmetric activation of the hsp90 dimer by its cochaperone aha1. Mol. Cell *37*, 344–354.

Roe, S.M., Prodromou, C., O'Brien, R., Ladbury, J.E., Piper, P.W., and Pearl, L.H. (1999). Structural basis for inhibition of the Hsp90 molecular chaperone by the antitumor antibiotics radicicol and geldanamycin. J. Med. Chem. *42*, 260–266.

Römer, L., Klein, C., Dehner, A., Kessler, H., and Buchner, J. (2006). p53–a natural cancer killer: structural insights and therapeutic concepts. Angew. Chem. Int. Ed. Engl. *45*, 6440–6460.

Rüdiger, S., Freund, S.M.V., Veprintsev, D.B., and Fersht, A.R. (2002). CRINEPT-TROSY NMR reveals p53 core domain bound in an unfolded form to the chaperone Hsp90. Proc. Natl. Acad. Sci. USA *99*, 11085–11090.

Schopf, F.H., Biebl, M.M., and Buchner, J. (2017). The HSP90 chaperone machinery. Nat. Rev. Mol. Cell Biol. *18*, 345–360.

Schrimpf, W., Barth, A., Hendrix, J., and Lamb, D.C. (2018). PAM: A Framework for Integrated Analysis of Imaging, Single-Molecule, and Ensemble Fluorescence Data. Biophys. J. *114*, 1518–1528.

Scotcher, J., Clarke, D.J., Weidt, S.K., Mackay, C.L., Hupp, T.R., Sadler, P.J., and Langridge-Smith, P.R. (2011). Identification of two reactive cysteine residues in the tumor suppressor protein p53 using top-down FTICR mass spectrometry. J. Am. Soc. Mass Spectrom. *22*, 888–897.

Sepehrnia, B., Paz, I.B., Dasgupta, G., and Momand, J. (1996). Heat shock protein 84 forms a complex with mutant p53 protein predominantly within a cytoplasmic compartment of the cell. J. Biol. Chem. *271*, 15084–15090.

Sharma, S.K., De los Rios, P., Christen, P., Lustig, A., and Goloubinoff, P. (2010). The kinetic parameters and energy cost of the Hsp70 chaperone as a polypeptide unfoldase. Nat. Chem. Biol. *6*, 914–920.

Smith, D.F., Stensgard, B.A., Welch, W.J., and Toft, D.O. (1992). Assembly of progesterone receptor with heat shock proteins and receptor activation are ATP mediated events. J. Biol. Chem. *267*, 1350–1356.

Taipale, M., Krykbaeva, I., Koeva, M., Kayatekin, C., Westover, K.D., Karras, G.I., and Lindquist, S. (2012). Quantitative analysis of HSP90-client interactions reveals principles of substrate recognition. Cell *150*, 987–1001.

Tomov, T.E., Tsukanov, R., Masoud, R., Liber, M., Plavner, N., and Nir, E. (2012). Disentangling subpopulations in single-molecule FRET and ALEX experiments with photon distribution analysis. Biophys. J. *102*, 1163–1173.

Walerych, D., Kudla, G., Gutkowska, M., Wawrzynow, B., Muller, L., King, F.W., Helwak, A., Boros, J., Zylicz, A., and Zylicz, M. (2004). Hsp90 chaperones wild-type p53 tumor suppressor protein. J. Biol. Chem. *279*, 48836–48845.

Walerych, D., Olszewski, M.B., Gutkowska, M., Helwak, A., Zylicz, M., and Zylicz, A. (2009). Hsp70 molecular chaperones are required to support p53 tumor suppressor activity under stress conditions. Oncogene *28*, 4284–4294.

Walerych, D., Gutkowska, M., Klejman, M.P., Wawrzynow, B., Tracz, Z., Wiech, M., Zylicz, M., and Zylicz, A. (2010). ATP binding to Hsp90 is sufficient for effective chaperoning of p53 protein. J. Biol. Chem. *285*, 32020–32028.

Wang, C., and Chen, J. (2003). Phosphorylation and hsp90 binding mediate heat shock stabilization of p53. J. Biol. Chem. *278*, 2066–2071.

Wawrzynow, B., Zylicz, A., and Zylicz, M. (2018). Chaperoning the guardian of the genome. The two-faced role of molecular chaperones in p53 tumor suppressor action. Biochim. Biophys. Acta Rev. Cancer *1869*, 161–174.

Wegele, H., Wandinger, S.K., Schmid, A.B., Reinstein, J., and Buchner, J. (2006). Substrate transfer from the chaperone Hsp70 to Hsp90. J. Mol. Biol. *356*, 802–811.

Whitesell, L., and Lindquist, S.L. (2005). HSP90 and the chaperoning of cancer. Nat. Rev. Cancer *5*, 761–772.

Whitesell, L., Sutphin, P.D., Pulcini, E.J., Martinez, J.D., and Cook, P.H. (1998). The physical association of multiple molecular chaperone proteins with mutant p53 is altered by geldanamycin, an hsp90-binding agent. Mol. Cell. Biol. *18*, 1517–1524.

Zylicz, M., King, F.W., and Wawrzynow, A. (2001). Hsp70 interactions with the p53 tumour suppressor protein. EMBO J. *20*, 4634–4638.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-p53 (Ab-3) (Mutant) Mouse mAb (PAb240)	Merck Millipore	Catalog #: OP29-200UG; RRID: AB_565797
Rabbit polyclonal anti-ydj1 antibody	Pineda Research, Berlin	N/A
Bacterial and Virus Strains		
BL21-CodonPlus (DE3)-RIL	Agilent	Catalog #: 230245
Biological Samples		N/A
Chemicals, Peptides, and Recombinant Proteins		
Atto-488 maleimide	ATTO-TEC	Catalog #: AD 488-45
Atto-488 NHS	ATTO-TEC	Catalog #: AD 488-31
Atto-532 maleimide	ATTO-TEC	Catalog #: AD 532-45
Atto-655-maleimide	ATTO-TEC	Catalog #: AD 532-41
IAEDANS	Thermo Fisher Scientific	Catalog #: I14
Alexa Fluor 647 C_2 Maleimide	Thermo Fisher Scientific	*Catalog #: A20347*
ATP	Roche	Catalog #: 10519987001
Apyrase	Sigma-Aldrich	Catalog #: A7646-500UN
Phosphatase	New England Biolabs	Catalog #: M0289S
Radicicol	Sigma-Aldrich	Catalog #: R2146
Protease Inhibitor G	SERVA electrophoresis	Catalog #: 39101
Protease Inhibitor HP	SERVA electrophoresis	Catalog #: 39106
Benzamidine	Sigma-Aldrich	Catalog #: B-6506
PMSF	Sigma-Aldrich	Catalog #: P7626-25G
GFP-Trap_A resin	Chromotek	Catalog #: gta-20
SERVA*Gel* TG PRiME 8 - 16%, precast gels	SERVA electrophoresis	Catalog #: 43279
Sequencing Grade Modified Trypsin	Promega	Catalog #: V511A
Bovine Serum Albumin	New England Biolabs	B9000S
Critical Commercial Assays		
Pierce BCA Protein Assay Kit	Thermo Fisher Scientific	Catalog #: 23225
Deposited Data		
Raw image data	This work	https://doi.org/10.17632/fgt8rw96tt.1
Experimental Models: Organisms/Strains		
BL21-CodonPlus (DE3)-RIL	Agilent	Catalog #: 230245
Oligonucleotides		
p21 promoter sequence fwd: 5′-Cy5-AGCTTAGACATGCCTAGACATGCCTA	This work	N/A
p21 promoter sequence rev: 3′–ATC TGTACGGATCTGTACGGATTCG-A-5′	This work	N/A
Recombinant DNA		
pET28-p53-DBD	This work	N/A
pET28-p53-FL	Retzlaff M., TU München	N/A
pET28-p53-DBD C182C229	This work	N/A
pET28-p53-DBD-GFP	This work	N/A
pET28a-hHsp90β	Tippel F., TU München	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
pET28-hHsp90-E42A	Boczek E., TU München	N/A
pET28-hHsp90-D88A	Boczek E., TU München	N/A
pET28-Hop	Richter K., TU München	N/A
pET23-p23	D. Toft	N/A
pETSumo-hHsp70	Kreihuber E., TU München	N/A
pETSumo-ydj1	Schmid A., TU München	N/A
pET28-Bag-1	This work, Invitrogen GeneArt (Thermo Fisher Scientific)	N/A
Software and Algorithms		
OriginPro 9.1	OriginLab	N/A
SedView	Hayes and Stafford, 2010	N/A
PAM	https://www.cup.uni-muenchen.de/pc/lamb/software/pam.html https://gitlab.com/PAM-PIE/PAM	N/A

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Prof. Johannes Buchner (johannes.buchner@tum.de).

METHOD DETAILS

Protein Expression and Purification

Human Hsp90β wild-type and mutants E42A and D88A containing6X-His tag were expressed in *E.coli* at 37°C for 4h using 1 mM IPTG (SERVA). The cells were harvested and re-suspended in Ni-NTA Buffer A (50 mM Na_2HPO_4 and 300 mM NaCl pH 7.8) containing protease inhibitors G and HP (SERVA) and Dnase I. The cells were then lysed with a French press at 1.8 kbar. The lysate was centrifuged at 20,000 rpm for 45 min at 4°C and the supernatant was collected and loaded on a 5 mL His-Trap FF or HP column (GE Healthcare) pre-equilibrated in Ni-NTA buffer A. The proteins were eluted in 60% Ni-NTA buffer B (50 mM Na_2HPO_4, 300 mM NaCl, 500 mM Imidazole, pH 7.8). The elute fractions were pooled and diluted 1:10 in Resource Q buffer A (40 mM HEPES pH 7.5, 20 mM KCl, 1 mM EDTA, 1 mM DTT) to reduce the salt concentration, and loaded on to a Resource Q ion exchange column. Elution of the bound proteins was done using a continuous salt gradient of Resource Q buffer B (40 mM HEPES pH 7.5, 1 M KCl, 1 mM EDTA, 1 mM DTT). The peak fractions were run on a SDS-PAGE and the relatively pure protein fractions were pooled and concentrated to 5 mL using a centricon of molecular weight cut off, MWCO, of 30 kDa. The last step of purification was size exclusion chromatography in buffer: 40 mM HEPES pH 7.5, 150 mM KCl, 5 mM $MgCl_2$, 1 mM tris(2-carboxyethyl)phosphine (TCEP). The concentrated protein was then loaded on a Superdex 200 pg HiLoad 16/60 column (GE Healthcare). The peak protein fractions were analyzed by SDS-PAGE followed by Coomassie staining and pure protein fractions were again pooled and concentrated to 1 ml. Concentration of Hsp90 was estimated using the molar extinction coefficient of 57,760 $M^{-1}cm^{-1}$ calculated from the respective amino acid sequence. Protein was then flash-frozen in liquid nitrogen and stored at −80°C.

Human Hsp70 and yeast ydj1 were expressed in a pET28b vector having a cleavable N-terminal 6X His tag and a SUMO solubility tag. The expression was started at an O.D. of 0.6 with 1 mM IPTG. Hsp70 and ydj1 were expressed at 30°C for overnight and 4 h respectively. The cells were harvested and re-suspended in Ni-NTA buffer A without ATP (40 mM HEPES pH 7.5, 150 mM KCl, 350 mM NaCl, 20 mM $MgCl_2$, 5% glycerin, 10 mM Imidazole, 2 mM β-mercaptoethanol, 0.5 mM ATP) for Hsp70 and in 40 mM NaH_2PO_4, 500 mM NaCl, 20 mM Imidazole, 2 mM β-ME, 10% glycerin for ydj-1. The supernatant after cell lysis was loaded on 5 mL His-Trap FF column pre-equilibrated in Ni-NTA Buffer A. The protein was eluted with 100%Ni-NTA buffer B (Ni-NTA buffer A with 300 mM Imidazole). The eluted fractions were pooled, diluted with cold water 1:5 times and dialyzed overnight against gel filtration buffer (40 mM HEPES pH 7.5, 150 mM KCl, 5 mM $MgCl_2$, 1 mM TCEP) at 4°C after addition of SUMO protease to the pooled fractions for cleaving off the SUMO tag. The cleaved SUMO tag and any uncut protein were removed by running a second Ni-NTA column and collecting the flow-through fractions. These fractions were pooled and concentrated to 10 mL with a centricon of MWCO 30 kDa for Hsp70 and 10 kDa for ydj1. The concentrate was loaded on the gel filtration column Superdex200 pg HiLoad 26/60 for Hsp70 and Superdex75 for ydj1. The peak fractions were run on a SDS-PAGE to assess protein purity and pure fractions were pooled, concentrated and flash-frozen in liquid nitrogen and stored at −80°C.

Human Hop and p23 were expressed and purified without any tag in *E.coli*. The first step was ion exchange chromatography (IEX). Lysate was loaded on self-packed 35 mL of Q-Sepharose (GE Healthcare) column pre-equilibrated in IEX buffer A (25 mM HEPES,

5 mM $MgCl_2$, 1 mM DTT, pH 7.6). The protein was eluted with a linear gradient of IEX buffer B (IEX buffer A + 500 mM KCl) over 12 column volumes. The fractions were checked on the SDS-PAGE gel for the presence of Hop or p23, pooled and dialyzed against 10 mM potassium phosphate buffer pH 7.5. After dialysis, protein was loaded on the hxdroxyapatite column equilibrated in dialysis buffer and then protein was eluted with 400 mM potassium phosphate buffer, pH 7.5 in a linear gradient. Relatively pure fractions were pooled and concentrated to 10 mL and loaded on size exclusion column Superdex75 pg HiLoad 26/60 equilibrated in 40 mM HEPES pH 7.5, 150 mM KCl, 5 mM $MgCl_2$, 1 mM TCEP. The pure protein fractions as judged by SDS-PAGE were pooled and concentrated and stored at −80°C after flash freezing in liquid nitrogen.

The human gene for Bag-1 was synthesized from Invitrogen GeneArt (Thermo Fisher Scientific) in a pET28b vector having an N-terminal thrombin cleavable 6X His tag. Bag-1 was purified by Ni-NTA His trap FF column (GE Healthcare), anion exchange chromatography on ResQ column (GE Healthcare), second Ni-NTA run after thrombin digestion overnight at 4°C followed by size exclusion chromatography using Superdex75 16/60 column (GE Healthcare). The buffers and purification procedure for Ni-NTA and ResQ column were the same as for Hsp90 described before. Bag-1 was stored at −80°C in 40 mM HEPES pH 7.4, 150 mM KCl, 5 mM $MgCl_2$, 1 mM TCEP. Since Bag-1 does not have any tryptophan residues, protein concentration was estimated using a Pierce BCA Protein Assay Kit (Thermo Fisher Scientific).

All p53 constructs (p53-DBD, p53-FL, p53-DBD C182 C229) were expressed and purified without any tag. p53 expression was induced at an OD_{600} = 0.6-0.8 with 1 mM IPTG and 100-200 μM $ZnSO_4$. Cells were shifted to 20°C for expression overnight, harvested and re-suspended in buffer A (50 mM KH_2PO_4/K_2HPO_4 pH 7.5, 50 mM KCl and 1 mM TCEP) containing protease inhibitor mix G, 1 mM benzamidine (Sigma), Dnase I and lysed. The cleared lysate was applied on a 5 mL Heparin HP column (GE Healthcare) equilibrated in buffer A containing 5% glycerol and 1 mM PMSF (Sigma) and eluted using a linear gradient of KCl (0-1 M KCl) in 15 column volumes. Fractions were checked on the gel and relatively pure fractions were pooled. The pooled fractions were diluted 1:10 times in 20 mM potassium phosphate buffer pH 7.5 and loaded on to 5 mL Q and SP-Sepharose FF prepacked columns (GE healthcare) connected in series equilibrated in 50 mM KH_2PO4/K_2HPO4 pH 7.5, 20 mM KCl, 1 mM DTT. The proteins were eluted from the SP-Sepharose column for DBD and Q-Sepharose column for p53-FL with a linear gradient of 0%–70% buffer B (0-1 M KCl). The pure fractions were pooled, concentrated and loaded on a Superdex 75 SEC 16/60 column for the p53-DBD or Superdex 200 for p53-FL equilibrated in 50 mM potassium phosphate pH 7.5, 100 mM KCl, 1 mM TCEP. Concentration of pure protein was estimated from the molar extinction coefficients and stored at −80°C.

The GFP tagged p53-DBD contained a GFP tag at the C terminus of p53-DBD in the pET28b vector having a cleavable N-terminal 6X His tag and a SUMO solubility tag. p53-DBD-GFP was purified with Ni-NTA, second Ni-NTA after digestion of the SUMO tag with SUMO protease overnight at 4°C, followed by gel filtration chromatography using Superdex75 16/60 column (GE Healthcare). The buffers used for Ni-NTA were, buffer A: 50 mM KH_2PO4/K_2HPO4 pH 7.5, 100 mM KCl, 1 mM TCEP and buffer B contained 500 mM imidazole in buffer A. Purified protein was stored at −80°C in 50 mM potassium phosphate (pH 7.5), 100 mM KCl, 1 mM TCEP.

FKBP51, FKBP52, Cyp40 were a kind gift from Franziska Tippel while human Aha1 was provided by Daniel Rutz.

Fluorescence Anisotropy Measurements

For the determination of dissociation constants for p53-DBD binding to Hsp70, Hsp40 fluorescence anisotropy measurements were carried out using either Atto-488 maleimide (ATTO-TEC) or IAEDANS (Invitrogen) labeled p53-DBD. Increasing concentrations of the respective chaperones were added to 500 nM of labeled p53-DBD* in a total of 150 μL reaction volume. For Hsp70 titrations, increasing amounts of Hsp70 (500 nM-50 μM) were added to 500 nM p53-DBD* in the presence and absence of 2 μM Hsp40 and 5 mM ATP (Roche). Measurements were done in a JASCO-8500 fluorescence spectrometer with polarizers at 25°C in 20 mM HEPES pH 7.4, 50 mM KCl, 5 mM $MgCl_2$, 1 mM TCEP buffer. Excitation and emission wavelengths were set at 500 nm and 522 nm for Atto-488 and 337 nm and 487 nm for IAEDANS respectively. The final steady state values of anisotropy were averaged over 50 s and used to calculate the change in anisotropy of the labeled p53-DBD* upon each titration. These values were then plotted against chaperone concentrations and the data was fitted to a one site binding model in OriginPro 9.1 to determine the K_D for Hsp70 and Hsp40 binding.

For DNA binding assays, 60-100 nM Cy5 labeled double stranded p21 DNA was used. Single stranded 26 bp complementary oligonucleotides (Eurofins Genomics) containing p53 consensus DNA binding p21 promoter sequences 5′-Cy5-AGCTT**AGACATGCCTAGACATGCCT**A-3′ and 3′–A**TCTGTACGGATCTGTACGGA**TTCG-A-5′ were annealed. To determine the K_D for p53-DBD or p53-FL binding to DNA, increasing amounts of p53 was added and the change in anisotropy of labeled DNA was measured at 20°C in 50 mM KH_2PO_4/K_2HPO_4 (pH 7.4), 50 mM KCl, 1 mM TCEP buffer using excitation and emission wavelengths of 650 nm and 663 nm respectively. Data was fit to a simple cooperative binding model (Dehner et al., 2005) in OriginPro 9.1.

For monitoring DNA dissociation kinetics in Figures 1 and S1A, 10 μM Hsp70, 2 μM Hsp40 and 5 mM ATP or AMP-PNP or ADP was added when indicated to pre-formed p53-DBD-DNA complexes. 1.5 μM p53-DBD and 100 nM Cy-5-p21-DNA was used. For the DNA association kinetics in Figure 2A, assay conditions were 1.5 μM p53-DBD, 2 μM Hsp40, 10 μM Hsp70 and 5 mM ATP and the incubation time was 2 h at 25°C or 1-1.5 h at 37°C respectively before monitoring the kinetics at 25°C after DNA addition.

For DNA binding recovery kinetics measurements, 1.5 μM of p53-DBD or p53-FL was incubated with 2 μM Hsp40, 10 μM Hsp70 and 5 mM ATP at 37°C for 1-1.5 h respectively in 20 mM HEPES pH 7.4, 50 mM KCl, 5 mM $MgCl_2$, 1 mM TCEP buffer. The samples were centrifuged at 14,000 g for 5 min and cooled down to room temperature, followed by incubation with 100 nM Cy5 ds p21 DNA for 5 min. DNA binding recovery was initiated at 25°C by the addition of the respective chaperones: 20 μM Bag-1 and/or 20μM Hsp90 and/or 20 μM Hop in various combinations as shown in Figure 4 and the anisotropy kinetics of Cy5-DNA was monitored. The

anisotropy values at each time point in each case were subtracted from the corresponding values of Cy-5 DNA in the case of the Hsp70-40-p53 complex and plotted as a change in anisotropy against time. In the case of apyrase (Sigma) or phosphatase (NEB), 1 unit of the enzymes was added to the Hop and Hsp90 mediated reaction. Radicicol (Sigma) was added at a final concentration of 100 μM. For p53-FL, the measurement buffer used was 40 mM HEPES pH 7.5, 150 mM KCl, 5 mM $MgCl_2$ and 1 mM TCEP.

Analytical Ultracentrifugation (AUC)

Analytical Ultracentrifugation experiments were performed with a ProteomLab Beckman XL-A centrifuge (Beckman Coulter) equipped with an AVIV fluorescence detection system in 20 mM HEPES pH 7.4, 50 mM KCl, 5 mM $MgCl_2$, 1 mM TCEP buffer unless stated otherwise. p53-DBD or p53-FL was randomly labeled with Atto-488 on cysteine residues following the protocol of the manufacturer. For titration experiments with Hsp40, increasing concentrations of Hsp40 (200 nM-15 μM) were added to 700 nM of p53-DBD*. For Hsp70 titrations, increasing amounts of Hsp70 (500 nM-15 μM) were added to 700 nM p53-DBD* containing 2 μM Hsp40 and 5 mM ATP. The reactions were incubated at 25°C for 2 h before measurements in AUC.

For AUC experiments to investigate multi-chaperone complex assemblies, 700 nM of p53-DBD* or p53-FL* was incubated with 2 μM Hsp40, 10 μM Hsp70 and 5 mM ATP at 37°C for 1 h respectively. The samples were then centrifuged at 14,000 g for 5 min, followed by additions of 20 μM of Bag-1, Hsp90, Hop and p23 in various combinations as denoted in Figures 4 and 6. The reactions were further incubated at 25°C for 2 h to allow steady state to reach before measuring in the AUC. For p53-FL, the measurement buffer used was 40 mM HEPES pH 7.5, 150 mM KCl, 5 mM $MgCl_2$ and 1 mM TCEP. The reactions were prepared in a total of 350 μL reaction volume. For experiments with labeled Hsp40, Hsp40 was randomly labeled with Atto-488 on lysine residues following the protocol of the manufacturer. 1 μM of Hsp40* was incubated with 1 μM p53-DBD, 10 μM Hsp70, 5 mM ATP at 37°C for 1 h. For reaction with Hop and Hsp90, 20 μM of each was added and the reaction was allowed to equilibrate at RT for 2 h before measuring in the AUC. Experiments were carried out at 42,000 rpm and 20°C in an eight-hole Ti-50 Beckman-Coulter rotor. Data analysis to generate normalized dc/dt plots was carried out using the SedView program (Hayes and Stafford, 2010) and OriginPro 9.1. The fraction of labeled protein bound (f) was calculated from sedimentation curves as follows:

$$f = \frac{\frac{dc}{dt}_{Norm,\ free\ only} - \frac{dc}{dt}_{Norm,\ free\ w/\ complex}}{\frac{dc}{dt}_{Norm,\ free\ only}}$$

where, $dc/dt_{Norm,\ free\ only}$ means the normalized dc/dt value of free labeled protein in the absence of any addition; $dc/dt_{Norm,\ free\ w/\ complex}$ represents the normalized dc/dt value of free labeled protein left after complex formation in each case. The maximum value of dc/dt was used from each fitted curve.

p53-DBD-GFP *In Vitro* Pull-Down

In vitro pull-down experiments to identify the components in the chaperone complexes were performed using GFP tagged p53-DBD. 5 μM of p53-DBD-GFP was incubated with 15 μM Hsp70, 3 μM Hsp40, 5 mM ATP at 37°C for 1 h. The reactions were centrifuged at high speed to remove any precipitate followed by the addition of 15 μM Bag-1, and/or 15 μM Hop, 15 μM Hsp90 as indicated in Figures S3C and 3D in a total reaction volume of 150 μL. The reactions were further incubated for 2 h at room temperature and then cooled on ice before adding pre-washed 20 μL GFP-Trap_A resin (gta-20, Chromotek). The reactions were then incubated at 4°C for 45 min-1 h with end to end mixing, followed by washing and elution as described on manufacturer's site (Chromotek). Elution was done directly in 2X SDS sample buffer before running on an 8%–16% precast SDS-PAGE gel (SERVA). For enhanced detection of ydj1 (Hsp40), a western blot was performed and Hsp40 was detected using an anti-ydj1 antibody (Pineda). 20 mM HEPES pH 7.4, 50 mM KCl, 5 mM $MgCl_2$, 0.05% tween-20 was used as binding and washing buffer. Protein amounts were quantitated by densitometry using the 1D gel analysis tool in the ImageQuant TL software (GE Healthcare). Mean values from five separate experiments were calculated for each of the proteins under each of the conditions used in Figure S3D. For each of the proteins, the highest value was set to 1.

Size Exclusion Chromatography

To identify the different complexes formed by p53-DBD and multiple chaperones, SEC coupled with SDS-PAGE was employed. 10 μM p53-DBD was incubated with 3 μM Hsp40, 15 μM Hsp70 and 5 mM ATP at 37°C for 1 h followed by addition of 20 μM each of Hsp90 and Hop respectively and equilibrating the reaction further at RT for 2 h. Complexes were then separated at a flow rate of 0.5 ml/min on a Superdex 200 increase 10/300 (GE healthcare) column attached to a Shimadzu HPLC system with a fraction collector. The column was equilibrated in the reaction buffer (20 mM HEPES pH 7.4, 50 mM KCl, 5 mM $MgCl_2$, 1 mM TCEP, 0.05% NaN_3) and 100 μL of sample was loaded. Collected fractions were run on an SDS-PAGE. p53-DBD and Hsp40 were detected on western blot using an anti- p53 antibody PAb240 (Merck Millipore) and ydj1 antibody (Pineda).

Limited Proteolysis

0.5 μM p53-DBD was incubated with 2 μM Hsp40, 10 μM Hsp70 and 5 mM ATP at different temperatures (25-37°C) for 1 h in 20 mM HEPES pH 7.4, 50 mM KCl, 5 mM $MgCl_2$, 1 mM TCEP. For Hsp70 concentration titrations, 0-15 μM Hsp70 was used keeping Hsp40

and ATP concentration constant and incubated at 37°C or 25°C for 1 h. Trypsin (Promega, sequencing grade) was added at 1:20 (w/w) to p53-DBD and the reactions were incubated on ice for 15 min. Reactions were quenched directly by adding 5X SDS loading buffer and boiling the samples. 200 ng of digested reactions were separated on an SDS-PAGE followed by western blot. p53-DBD was detected with the p53 antibody PAb240. For Hsp70 alone samples, 10 μM Hsp70 was used. For reactions with Bag-1 and Bag-1-Hop-Hsp90 in Figure S4F, 20 μM of each was added as indicated to the Hsp70-Hsp40-p53-DBD-ATP complex and incubated at 25°C for 1 h followed by addition of trypsin.

Aggregation Assay

Aggregation kinetics was monitored on a Cary 100 UV-visible spectrophotometer (Varian). 1.5 μM of p53-DBD or p53-FL was incubated with 2 μM Hsp40, 10 μM Hsp70, 5 mM ATP at 37°C for 1 h in 20 mM HEPES pH 7.4, 50 mM KCl, 5 mM $MgCl_2$, 1mM TCEP for DBD and 40 mM HEPES pH 7.5, 150 mM KCl, 5 mM $MgCl_2$ and 1 mM TCEP for p53-FL. Light scattering was monitored at 350 nm with time at 37°C after addition of 10 μM Bag1 and/or 10 μM HOP, Hsp90 as given in Figures 5F and S5B.

Electrophoretic Mobility Shift Assay

The DNA binding activity of p53-DBD in the presence and absence of Hsp70, Hsp40, ATP at 37°C as indicated in Figure S1I was analyzed by EMSA. 125 ng p53-DBD was incubated alone and with 1250 ng of Hsp70 with or without 2 mM ATP; 1250 ng of Hsp70, 250 ng of Hsp40 and 2 mM ATP; 250 ng Hsp40 at 37°C for 15 min. 5 μL of above reactions was added to 20 μL reaction mix containing 20 ng Cy5 labeled ds p21 DNA and pET28 vector as competitor DNA in DNA binding buffer (40 mM HEPES, 50 mM KCl, 10 mM $MgCl_2$, 0.1% Triton X-100, 1 mg/ml BSA, 15% (v/v) glycerol, 5 mM DTT). The reactions were incubated on ice for 10 min before separating on a 4% native polyacrylamide gel at 4°C. The gels were scanned in a Typhoon 9200 Fluorescence imager using 650 nm and 670 nm excitation and emission filters respectively.

p53 Labeling

For anisotropy and AUC measurements, p53-DBD or p53-FL was randomly labeled on cysteine residues with ATTO-488 maleimide or IAEDANS fluorescent dyes according to the manufacturer's protocol with a few modifications. Briefly, p53-DBD or p53-FL was diluted in 50 mM KH_2PO_4/K_2HPO_4 (pH 7.4), 100 mM KCl, 1 mM TCEP buffer such that the final p53 concentration was 60-70 μM. Labeling was performed on ice. The dyes were added at 2.5-3 fold molar excess of the protein and the reactions were incubated at 4°C overnight. The free dye was removed using a Superdex 75 10/300 column (GE Healthcare) or dialysis in the case of p53-DBD and using a PD-10 column (GE Healthcare) for p53-FL. The degree of labeling was determined according to the formula given on the manufacturer's site. DNA binding of labeled p53 was monitored as described for unlabeled p53.

Single-Pair FRET Experiments

Construction of p53-DBD FRET sensor

p53-DBD has 10 cysteines out of which six are buried (Cys-135, Cys-141, Cys-275) or involved in Zn^{2+} coordination (Cys-176, Cys-238, Cys-242). One cysteine is partly exposed (Cys-124) and three are fully exposed (Cys-182, Cys-229, Cys-277) to the solvent (Huang et al., 2009; Scotcher et al., 2011). The naturally exposed cysteines Cys-182 and Cys-229 were used for setting up an intra-molecular FRET system, p53-DBD C182C229. To avoid mislabeling, the remaining exposed (C277) and partly exposed (C124) cysteines were mutated to serines. The respective genes were synthesized and cloned by GeneArt (Thermo Fisher Scientific) in the pET28b vector at Nco I and Xho I sites. The FRET mutant was purified like the wild-type p53-DBD. We used Atto-532 (ATTO-TEC) as the donor and Alexa-647 (Invitrogen) as the acceptor dye for the FRET pair. The p53-DBD FRET mutant was labeled by the addition of both the maleimide dyes at a 5-6 fold molar excess of the protein and in equal molar ratios to one another respectively following the protocol as for p53-DBD wild-type. Functionality of the Atto-532 labeled p53-DBD FRET sensor was tested by measuring the change in fluorescence anisotropy of Cy5 labeled p21 DNA and the K_D was determined as described for the wild-type p53-DBD.

The K_D for DNA binding dropped slightly from 2 μM for the labeled DBD FRET sensor as compared to 0.7 μM for the wild-type p53-DBD (see Figure S2A). Further, the interaction of the labeled DBD FRET sensor with the Hsp40-Hsp70 system was also not significantly altered. Binding of the Atto532/Alexa647 labeled p53-DBD C182C229 sensor to Hsp70 in presence of Hsp40 and ATP was monitored by fluorescence anisotropy as described for the wt p53-DBD. The K_D of Hsp70 binding dropped only by a factor of three, to 6 μM for the labeled DBD FRET sensor versus 2 μM for the wild-type p53-DBD (see Figure S2B). The DNA and Hsp70 binding control experiments indicated that mutations and dyes did not have significant effects on p53-DBD structure and functionality.

Accessible Volume Calculations

The accessible volumes of the fluorophores at position C182 and C229 on p53-DBD is shown in Figure 3A (PDB ID: 1UOL). FRET-restrained positioning and screening (FPS) software as described by Kalinin et al. (2010) was used to get accessible volumes and respective approximation of distance for linker length (20 Å), linker width (4.5 Å) and fluorophore radius (3.5 Å).

Sample preparation

10 nM p53-DBD FRET sensor labeled with Atto 532 and Alexa 647 were incubated with 3 μM Hsp40 or 3 μM Hsp40, 20 μM Hsp70 and 5 mM ATP at different temperatures as mentioned in the results for 1 h in 20 mM HEPES, pH 7.4, 50 mM KCl, 5 mM $MgCl_2$, 1 mM

TCEP. Approximately 70-100 pM p53-DBD was measured at RT with MFD-PIE, keeping the added components at the concentrations used for incubation. For FCS measurements, 10 nM of p53-DBD FRET sensor was used. For measurements with Bag-1, Hop and Hsp90; 20 μM Bag-1 or 20 μM Hop/Hsp90 or 20 μM Bag-1/Hop/Hsp90 were added just before the measurement to Hsp70-Hsp40-p53-DBD C182C229 complex formed at 37°C for 1 h.

For measuring FRET efficiency in the denatured state, the p53-DBD FRET sensor was unfolded in 3 M GdmCl for 2 h at 37°C. 70-100 pM denatured p53-DBD was measured in 3 M GdmCl at RT.

SpFRET Measurements

All solution based spFRET measurements were done on a custom-build confocal microscope. Protein molecules diffuse through the laser focus generally within five milliseconds. Approximately 70-100 pM of labeled sample was measured to minimize multiple molecules being in the focus of laser at a time and thus ensure single molecule resolution. Pulsed interleaved excitation (PIE) was used with a 532-nm (PicoTA 530; Toptica) and a 640-nm (LDH-D-C-640; PicoQuant) laser to excite the donor and acceptor fluorophores respectively. The lasers were synchronized at a repetition rate of 26.7 MHz with a 18ns delay between them. A laser power of 100 μW was measured before the objective for both the lasers. Fluorescence was collected with a 60x water immersion objective (Plan Apo IR 60x1.27 WI; Nikon) and focused through a pinhole with 75μm diameter. Multi-Parameter Fluorescence Detection (MFD) was used for detection where the collected fluorescence was first separated via polarization with a polarizing beam-splitter (PBS3; Thorlabs) and later using a dichroic mirror (DualLine z532/635, AHF Analysetechnik) to separate the green and red fluorescence signal. Emission filters (for green-Brightline 582/75, Semrock; for red-HQ700/75 M, Chroma) were used to clean the florescence before being detected on four single-photon-counting (SPC) avalanche photodiodes (APD) (Perkin-Elmer). Synchronized SPC cards (SPC 154; Becker and Hickl) were used to record the photons and data was analyzed with PAM, a custom-written software in MATLAB (The MathWorks) (Schrimpf et al., 2018).

MFD-PIE analysis

In MFD-PIE, all the photons from each molecule were used to calculate multiple parameters including FRET efficiency, stoichiometry, fluorescence lifetime and fluorescence anisotropy (See also Figure S6). Initially, an all photon burst search with a criterion of a minimum of 100 photons with a count rate of at least 5 photons in a sliding time window of 500 μs was used to distinguish the burst from individual molecule. Molecules with a stoichiometry (S) of ~0.5 were selected for further analysis. Stoichiometry is calculated as:

$$\left[S = \frac{F_{GG} + F_{GR}}{F_{GG} + F_{GR} + F_{RR}}\right]$$

where F_{GG} and F_{GR} are the fluorescence intensities of the donor and acceptor channels after donor excitation, respectively and F_{RR} is the fluorescence intensity of the acceptor channel after acceptor excitation.

An ALEX-2CDE filter was used to remove photo-bleaching and blinking events (Tomov et al., 2012). The FRET efficiencies were calculated after correcting for spectral cross talk of green fluorescence in the red channel ($\alpha \sim 0.03$) and direct excitation of the acceptor ($\delta \sim 0.07$). Differences in quantum yields of the fluorophores and detection efficiencies are taken care by the γ-factor (0.5). The corrected FRET efficiency is defined as

$$\left[E = \frac{F_{GR} - \alpha F_{GG} - \delta F_{RR}}{\gamma F_{GG} + F_{GR} - \alpha F_{GG} - \delta F_{RR}}\right]$$

The burst-wise lifetime and anisotropy was also calculated. The ideal relationship between FRET efficiency and fluorescence lifetime of the donor is described in the absence of conformational dynamics as:

$$\left[E = 1 - \frac{\tau_{D(A)}}{\tau_{D(0)}}\right]$$

where $\tau_{D(A)}$ is the fluorescence lifetime of donor in presence of an acceptor and $\tau_{D(0)}$ is the fluorescence lifetime of donor in absence of an acceptor.

When a molecule undergoes conformational changes within a burst, the average donor lifetime is dominated by the low FRET efficiency state (where more donor photons are detected). This leads to a deviation in lifetime from the expected FRET efficiency versus donor lifetime relationship for a static FRET state indicating the presence of dynamics. The new relationship is given by Kalinin et al. (2010):

$$\left[E = \frac{\tau_1 . \tau_2}{\tau_{D(0)}[\tau_1 + \tau_2 - \tau]}\right]$$

where τ_1 and τ_2 are the donor lifetimes of two different FRET states and τ is the fluorescence weighted average lifetime of a single burst.

Förster radius calculation

The distance between the donor and acceptor (R) can be related to the FRET efficiency using:

$$E = \frac{1}{1 + \left(\frac{R}{R_0}\right)^6}$$

where R_0 is the Förster radius, that is the distance at which half of the energy is transferred to the acceptor molecule. The Förster radius depends on the particular dye-pair and is calculated using the following equation:

$$R_0 = \frac{9000(\ln 10)\phi_{D(0)}\kappa^2 J}{128\pi^5 n^4 N_A}$$

where $\phi_{D(0)}$ is the quantum yield of the donor in the absence of the acceptor, κ is the relative orientation of dipoles of donor emission and acceptor absorption, J is the overlap integral of the donor emission spectrum and acceptor absorption spectrum, n is the refractive index of the medium and N_A is the Avogadro's constant.

With the above calculations, the Förster radius for the dye-pair used (Atto 532 and alexa 647) was found to be 62 Å.

The Freely Joined Chain (FJC) model

We compared the distance between the dye pairs of the p53 DBD FRET sensor in the unfolded state (i.e., in 3 M GdmCl) with what would be expected. Using the Freely Joined Chain (FJC) model and treating the unfolded protein as a polymer, the average end-to-end distance is given by:

$$\left\langle \vec{R}^2 \right\rangle = n.l^2,$$

where $\vec{R}$ is the total end-to-end distance vector, n is the total number of units and l is the length of each unit. Using a persistent length of 0.6-0.7 nm in the unfolded form (Lapidus et al., 2002), a separation of 47 amino acids between C182 and C229 in the DBD FRET sensor results in an end-to-end distance of ~50 Å. After adding the length of the linkers used to attach the dyes to the protein (adding ~5 Å distance for each dye), the separation of the fluorophores is ~60 Å.

Fluorescence Correlation Spectroscopy

Fluorescence correlation spectroscopy (FCS) measurements were performed using MFD-PIE to measure the diffusion times of 10 nM p53-DBD alone and in the formed complex with Hsp40-Hsp70-ATP at 37°C. The diffusion time for free dye was measured as a control. The FCS curves were fitted with the following model function for 3D diffusion assuming a 3D Gaussian probe volume using our custom written software PAM (Schrimpf et al., 2018):

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

where, N is the average number of freely diffusing molecules with specific diffusion times τ_D, detected in the confocal volume with a structure parameter ρ, ($\rho = w_0/z_0$, where w_0 is radial and z_0 axial dimensions of the confocal volume). The geometric factor, $\gamma = 2^{-3/2}$, corrects for the shape of the confocal spot assuming a three-dimensional Gaussian shape. y_0 is a baseline offset and corrects for the slow variations in the experiment typically on second timescales. The diffusion coefficient (D) is related to diffusion time (τ_D) via:

$$D = \frac{w_0^2}{4\tau_D}$$

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical parameters including the exact value of n (replicates) and dispersion and precision measures (mean ± s.d.) are reported in the Figures and Figure Legends. For the quantification of bands in the pull downs, $p < 0.05$ was considered significant. Data analysis was performed in OriginPro 9.1 and in the custom written PAM software. Software used for specific methods are also mentioned in the respective STAR Methods section.

DATA AND SOFTWARE AVAILABILITY

The custom software package *PAM* was used for the analysis of the single-pair FRET experiments. The code is freely distributed as open-source software under https://www.cup.uni-muenchen.de/pc/lamb/software/pam.html and hosted at https://gitlab.com/PAM-PIE/PAM. Details are given in Schrimpf et al. (2018). The raw image data is available at https://doi.org/10.17632/fgt8rw96tt.1