Heat shock proteins studied by optical tweezers

Heat shock proteins studied by optical tweezers

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. ir. K.C.A.M. Luyben,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op donderdag 10 december 2015 om 10:00 uur

door

Sergey Vladimirovich BEZRUKAVNIKOV

Master of Science, Moscow State University
geboren te Kaluga, Rusland

Dit proefschrift is goedgekeurd door de

promotor: prof. dr. ir. S.J. Tans

Samenstelling promotiecommissie:

Rector Magnificus,	voorzitter
Prof. dr. ir. S.J. Tans,	Technische Universiteit Delft
Onafhankelijke leden:	
Prof. dr. A. Engel	Technische Universiteit Delft
Prof. dr. P. Goloubinoff	Université de Lausanne, Zwitserland
Prof. dr. P. Senet,	Université de Bourgogne, Frankrijk
Prof. dr. H.H. Kampinga,	Rijksuniversiteit Groningen
Prof. dr. M.M.A.E. Claessens	Universiteit Twente
Dr. M.E. Aubin-Tam,	Technische Universiteit Delft
Prof. I.T. Young,	Technische Universiteit Delft, reservelid

FOM Institute AMOLF

The work described in this thesis was performed at the FOM Institute AMOLF, Science Park 104, 1098XG Amsterdam, the Netherlands. This work is part of the research program of the Foundation for Fundamental Research on Matter (FOM), which is financially supported by the Netherlands Organization for Scientific Research (NWO).

Nederlandse titel: Temperatuurstress eiwitten gestudeerd met optical tweezers

Printed by Gildeprint - Enschede

Cover design: S. Bezrukavnikov with contribution from M. Avellaneda and J. Langeveld

ISBN 978-90-77209-95-0

An electronic version of this dissertation is available at http://www.amolf.nl/ and at http://repository.tudelft.nl/. Printed copies can be obtained by request via email to library@amolf.nl

To my wife Inna

Contents

1

Introduction

Chaperone-assisted protein folding has already been studied extensively with a range of techniques. This chapter begins with a brief description of historical landmarks of the protein folding science. It is followed by an overview of conventional techniques used to study protein folding. However, it has been proven difficult to address certain crucial questions with these techniques, such as the state of the protein client upon interaction with the chaperone and precise mechanistic details of this interaction. I motivate the use of single molecule techniques to tackle these problems and provide a quick description of those that were applied to protein folding. Finally, the scope of the thesis is presented in this chapter.

1

1.1. Basic concepts of protein folding

1.1.1. Important milestones of protein science

Proteins were first described as a distinct class of molecules by Mulder and named by Berzelius in 1838 [1]. Since that time, structure and function of these remarkable molecules are thoroughly investigated.

In 1902, Franz Hofmeister and Emil Fischer nearly simultaneously proposed at the same conference that proteins are linear chains of amino acids. James Sumner discovered a catalytic function of the proteins in 1926 by showing that enzyme urease is a pure protein [2]. Shortly after that, Hsien Wu first suggested that the protein denaturation, a process that results in the loss of solubility and enzymatic activity by the protein, is a pure conformational change, namely protein unfolding [3]. The idea of a protein having a specific configuration that may fold and unfold (it was also shown by that time that the protein denaturation is reversible [4] was further supported by the work of Alfred Mirsky and Linus Pauling in 1936, suggesting that the protein chain is coiled in a specific way that accounts for the function and is lost upon denaturation [5]. However, the details of these structural features were predicted only 15 years later by Linus Pauling and his coworkers that reported an occurence of local interactions resulting in formation of α-helices and ß-sheets in protein chains based on the structures of amino acids and peptides [6]. First fully sequenced protein was reported by Sanger in 1953 [7–10] with the following conclusion made from those results: "It would thus seem that no general conclusions can be drawn from these results concerning the general principles which govern the arrangement of amino-acid residues in protein chains. In fact, it would seem more probable that there are no such principles, but that each protein has its own unique arrangement; an arrangement which endows it with its particular properties and specificities and fits it for the function that it performs in nature." [8] In other words, it was suggested that the protein function is determined by the sequence, a hypothesis that had a major effect on the growing field of molecular biology. However, the absence of direct experimental clues about the protein structure was limiting further research in this direction. First 3D-structures of proteins (Fig. 1.1) were determined by Kendrew in 1958 and Perutz in 1960 [11, 12], followed by the work of Phillips on lysozyme [13] in 1965 where a first high-resolution (2Å) protein structure was obtained. Contrary to the expectations of regular crystalline-like structure by protein scientists, supported by earlier Pauling anticipations, it featured helices that were packed together in a peculiar irregular way. It led to a conclusion that "…it is more complicated than has been predicated by any theory of protein structure. Though the detailed principles of construction do not yet emerge, we may hope that they will do so at a later stage of the analysis." [11]

The puzzling complexity of the protein structure immediately resulted in a notion of a "protein folding problem" [14, 15]. It featured two interconnected questions: 1) How does the protein sequence determine its 3D native structure and what are the driving forces? 2) What are the pathways of folding and unfolding and can they be predicted?

A landmark work of Christian Anfinsen on ribonuclease refolding in 1961 showed

Figure 1.1: **First reported protein structure.** 3D model of the myoglobin structure solved by Kendrew et al. Reprinted with permission from Dill 2012.

that the native protein structure is encrypted solely in its amino acid composition [16], a principle that was further developed by him into Anfinsen's dogma [17]. It stated that postulating that the native structure corresponds to the free energy minimum of the protein-solvent system. Furthermore, proteins were assumed to have only one stable native structure. Among many other important consequences of this postulate, it assumes that (at least for relatively simple and small proteins) the protein structure is independent of whether it had been folded in the cell or refolded in the test tube in isolation, allowing for in vitro studies of protein folding.

1.1.2. Driving forces of protein folding

According to Anfinsen's dogma, the reason why two proteins with different sequences acquire different native structures is rooted in the physico-chemical properties of amino acid sidechains. Since protein stability is generally quite subtle at room temperature (less than 1/10 kT per residue or 5-20 kcal/mol per unfolding [18, 19] even small interactions, when taken together, may significantly contribute to the stability, hence it is of major importance to understand them and carefully account for them when studying folding process. Below I list the most important forces leading to the unique structural features of proteins, referring to the review by Dill [20] for further details.

1. Hydrogen bonding is recognized as a special case of electrostatic interaction when a hydrogen atom is shared between two electronegative atoms. It often occurs between the carbonyl C=O and amide NH groups of peptide

backbone in proteins underlying the formation of protein secondary structure [6]. Hydrogen bond energy is in the range of 2 to 10 kcal/mol [21].

2. Van der Waals interaction is a weak attraction force that occurs between the molecules in close proximity due to dipole-dipole interaction of their electron clouds. The energy of a single pair ranges from 0.5 to 1 kcal/mol [22]. Tight packing in proteins implies that the cumulative contribution of van der Waals interactions is important [23].

3. Electrostatic interactions. Amino acid side chains determine the total electrical charge of a protein. When the protein is highly charged (at extremes of pH for example), it leads to the nonspecific repulsions that, however, have no contribution near the isoelectric point. Furthermore, when two charged side chains are in close proximity, it may lead to specific attraction or repulsion, stabilizing or destabilizing the folded protein state. Ion pairing (attraction of two oppositely charged sidechains) contributes 1-3 kcal/mol per ion pair to protein stability [24].

4. Hydrophobic effect is driven by the energetic preference of nonpolar side chains in the protein for non-water environment. The solvation of nonpolar side chains results in increased free energy of unfolded protein state, and thus drives the formation of folded state where the nonpolar residues are packed deep inside the protein and shielded from water. Model compound studies display the transfer energy of a single nonpolar side chain from water into hydrophobic medium being 1-2 kcal/mol [25].

Evidence exists that the major driving force of protein folding is hydrophobic [14]. Both native and amyloid-like fibril protein structures were successfully mimicked by copying their non-polar and polar patterns in synthetic proteins [26, 27]. Even though secondary structure is formed around hydrogen bonds, it is significantly stabilized by the chain compactness that arises from hydrophobic collapse [14]. Many proteins can form secondary structure depending on the surrounding [28].

1.1.3. Folding pathways and the energy landscape concept

Protein folding is a transition from a disorder to an order. It was already assumed in the 1930s by Anson and Mirsky that it could occur as all-or-none reaction, in which the protein is either in the fully native or fully unfolded state [30]. Strictly speaking, each protein state is a set of microstates, which are transiently populated in certain conditions [31]. Since two neighboring amino acids can bind each other with three different bond angles, an unfolded chain possesses a huge number of degrees of freedom, and thus from a mechanistic point of view it can fold in multitude of different ways. However, many proteins are able to find a “right”, native solution in milliseconds. In 1969, during a conference presentation, Cyrus Levinthal postulated his famous paradox, asking how proteins can fold so fast, taking into account an enormous number of possible conformations [32]. The paradox can be easily illustrated by considering the energetic aspect of

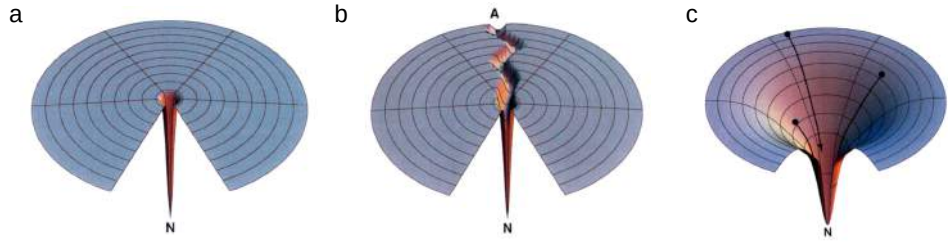


Figure 1.2: **Energy landscape concept of protein folding.** Drawings of protein energy landsdcapes representing different scenarios of protein folding: random search of an unfolded protein chain **(a)**, folding via specified pathway **(b)**, and folding funnel **(c)**. Reprinted by permission from Macmillan Publishers Ltd: Nature Structural Biology [29], ©1997.

the search process that is how the system's free energy depends on the protein conformation. This concept is now called an energy landscape approach. Strictly speaking, protein's energy landscape is a hypersurface of protein-solvent "internal free energy" [33] in the coordinates describing the position of every residue in the protein chain. Kinetically stable protein macrostates appear as valleys at the energy landscape, separated by the hills that are energy barriers (similar to transition states in small-molecules reactions) [31]. Levinthal's argument assumed all the unfolded conformations having about the same free energy and thus the energy landscape looking like a flat field with a tiny deep hole in the middle corresponding to the native folded conformation that is the minimum of free energy (Fig. 1.2a). Naturally, a random conformational search for the native state at such a landscape (that may be compared to playing golf blindfolded) would take an eternity. Different mechanisms were proposed to simplify the search and split it to many small search problems that may be solved stepwise. Levinthal suggested a mechanism characterized by initial local contacts guiding the propagation of tertiary structure without kinetically stable intermediates [32]. Ptitsyn suggested that protein folding could occur hierarchically with the formation of either secondary (framework model) [34] or tertiary (hydrophobic collapse model) [35] structure followed by the subsequent formation of native contacts. First experimental observations also demonstrated the presence of kinetically stable folding intermediates [36], hence it became largely assumed that a protein folds to the native state through a certain sequence of metastable intermediate states that are local minima of the energy landscape (see Fig. 1.2b for the energy landscape corresponding to this assumption). The presence of the intermediate states was assumed to restrict a number of conformations to be searched from and thus to reduce the folding time to experimental range. However, it was also invoked that proteins may have a number of parallel folding pathways [37]. In the late eighties Wolynes and Dill came up with a novel approach to protein folding studies by using the models from statistical mechanics to account for the microscopic properties of protein chains that effectively determine the energy landscape of the system [38, 39]. These studies revealed that more compact and closer to na-

tive protein macrostates have indeed fewer conformations compared to unfolded macrostates, but, however, despite an unfavorable contribution of chain configurational entropy, the system's free energy decreases as protein folds. It hence indicated that the protein folding energy landscape has a shape of a funnel instead of the golf course (Fig. 1.2c), since the lateral area of the energy landscape at a given depth represents the number of conformations having the given free energy. The resolution of the Levinthal paradox lied in a fact that the conformational search is energetically biased against locally unfavorable configurations and thus only a limited number of conformations is visited during folding, as demonstrated by Zwanzig et al. [40] using mathematical model. The local roughness and bumps on the funnel reflected the kinetic traps that lead to multi-state folding (Fig. 1.3a), and smoother funnels were assumed to illustrate two-state folding (Fig. 1.3b).

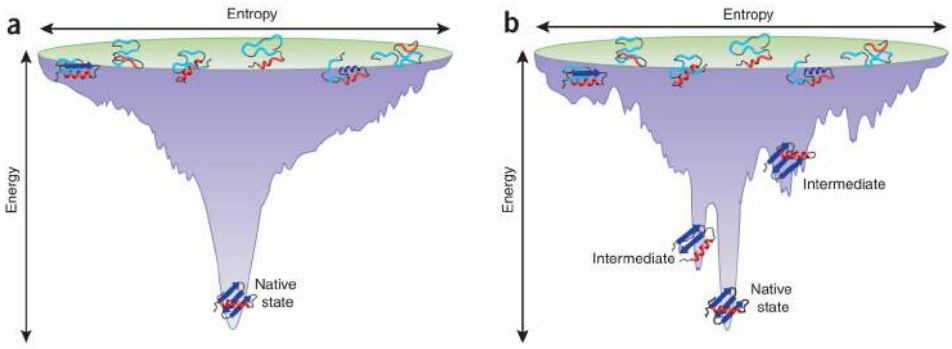


Figure 1.3: **Energy landscape determines a folding pathway.** Cartoons of protein energy landsdcapes for a two-state folding protein **(a)** and for a multi-state protein **(b)**. Reprinted by permission from Macmillan Publishers Ltd: Nature Structural Biology [41], ©2009.

Three previously proposed folding mechanisms were addressed in 1990s by the discovery that proteins could fold without the intermediates [42] and that secondary and tertiary structure could form in parallel as protein folds [43]. These discoveries led to the synthesis of both framework and hydrophobic collapse models, resulting in the notion of the nucleation-condensation mechanism (Fig. 1.4a) that invokes the collapse of the protein to a native state around the diffuse nucleus driven by both long-range hydrophobic interaction and local contacts [44]. This mechanism has been observed in folding of barnase, a small multidomain protein, where long-range interactions trigger the formation of secondary structure, which in turn aids in the stabilization of native tertiary conformation (Fig. 1.4b) [45].

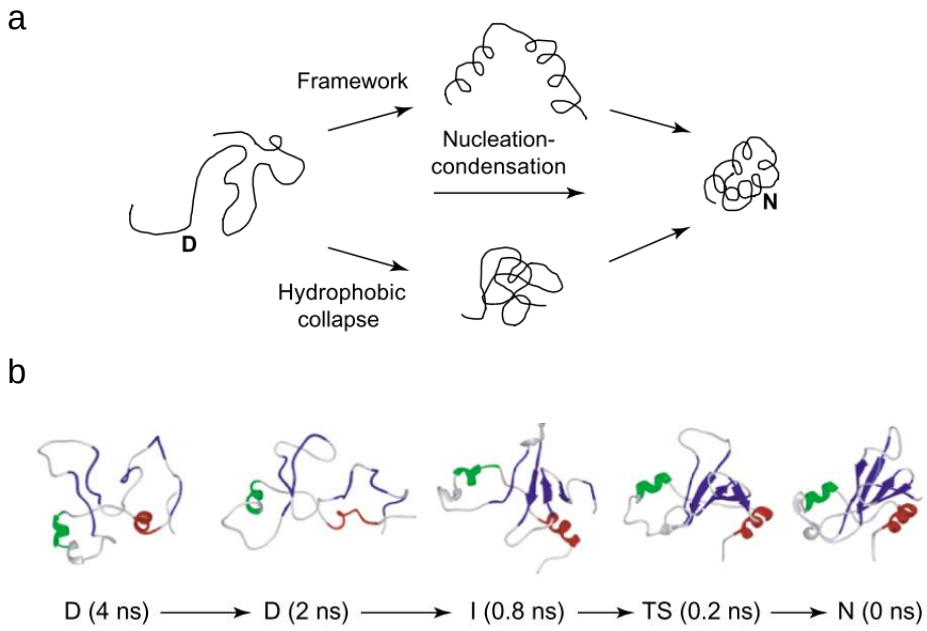


Figure 1.4: **Nucleation-condensation folding mechanism. a,** Schematic comparison of the framework and the hydrophobic collapse protein models. The nucleation-condensation mechanism is a synthesis of these two models. **b,** Snapshots of molecular dynamics simulation of barnase temperature-driven unfolding presented in reverse order. Reprinted from Trends in Biochemical Sciences, vol. 28, no. 1, pp. 18–25., ©2003, with permission from Elsevier.

1.2. Experimental protein folding studies: advances and challenges

1.2.1. Novel techniques for protein folding studies

The previously mentioned advances in understanding of protein folding would be largely impossible to achieve without the experimental advances that provided more detailed data on protein folding. When first protein structures were solved and the "protein folding problem" was noted, an experimental arsenal of protein scientists that allowed studying it was quite limited. Protein denaturation in the presence of different physical and chemical factors, such as pH, salts, detergents, urea, and temperature, was typically performed with the extent of denaturation measured by monitoring the optical properties of the medium or quantifying the precipitation [46, 47]. Time dependence of denatured fraction was fitted to single or multiple exponentials to extract kinetic information such as rate of denaturation. These techniques were unable to follow the folding mech-

anism with high structural resolution and did not detect transiently populated species. Thus they were not suitable for sampling a multitude of conformations that are accessible for a folding protein according to the funnel concept. However, detailed characterization of all non-native protein conformations is important for complete understanding of folding mechanisms. This experimental challenge facilitated the boom of new techniques to monitor protein folding, which are described in details in many excellent reviews [41, 48–50], hence here I will just give a short overview of most common and influential techniques used to elucidate new details of protein folding.

1. Protein engineering. It involves introducing mutations in different positions throughout the sequence and measuring the folding kinetics and global stability of the mutants [51]. It allows quite detailed description of protein folding transition states, while monitoring the mutation-dependent changes of folding kinetics (FI-values analysis) [52].

2. X-ray crystallography. It characterizes X-ray diffraction on protein crystals to acquire structural information. It is successfully applied since the 60s to get static snapshots of stable native protein states [13]. Furthermore, the technique has been recently extended to acquire picosecond time-resolved X-ray diffraction to study the ultrafast dynamics of ligand release [53]. The disadvantage of X-ray crystallography is that the formation of protein crystal is prerequisite for structure determination, and only a limited number of proteins are capable of that.

3. NMR spectroscopy. It exploits the magnetic resonances of atomic nuclei in the sample placed into external magnetic field [54]. It is one of a few techniques available to study the structures of partially folded, denatured and natively disordered proteins [55]. The resonance frequencies of nuclei in proteins are sensitive to each other and to environment, allowing for determination of inter-atomic distances and protein structure [56]. Furthermore, relaxation dispersion NMR spectroscopy provides an atomic resolution picture of millisecond time scale transitions in proteins [57].

4. H/D exchange. In this method, hydrogen exchange between polar groups of protein and solvent is quantified as a measure of protein structure. It is achieved by replacing hydrogens in protein to deuterium and following the kinetics of its exchange to hydrogen from water using 2D NMR [58]. Pulsed H/D exchange allows extracting structural and kinetic information on folding intermediates by applying short hydrogen-labeling pulses. Native H/D exchange captures metastable protein states that are visited transiently from the native state.

5. Infrared spectroscopy (FTIR). It monitors the relative position and orientation of molecular groups in a protein by shining the infrared light on it and characterizing its infrared absorption spectrum. Location of characteristic absorption bands in protein infrared spectrum is sensitive to its secondary

structure content, since the chemical bonds that typically absorb the radiation and vibrate in proteins are involved in hydrogen bonding [59]. It is extended to non-equilbrium processes using, for example, stopped-flow mixing [60] or laser light [61] to trigger conformational transitions. Transient 2D FTIR allowed observing real-time hydrogen bond dynamics on picosecond timescale [61].

6. Raman spectroscopy represents another variant of vibrational spectroscopy. In this method, deep UV radiation is used to illuminate the sample, and changes of light properties (frequency shifts) are used to obtain information about the sample composition. Incident light at frequencies falling within a specific electronic absorption band drives a resonance enhancement of particular vibrations up to 108 fold, leading to very high detection sensitivities [62]. For example, C_α-H band in proteins can be monitored to determine angle distributions for peptide bonds that provide an estimate for protein conformation. Raman spectra are sensitive to both secondary and tertiary structure. It has also been used in combination with the laser-induced T-jump technique to study the kinetics of protein folding [63, 64]. In comparison to IR measurements, which are often limited to experiments in D_2O because of background water IR absorption, usage of Raman spectroscopy can greatly reduce solvent interference [65].

7. Circular dichroism (CD) is an excellent method for determination of the secondary structure of proteins, their folding properties and proteins interactions. Proteins absorb left-hand and right-hand circularly polarized light differently with different structural elements having characteristic CD spectra [66, 67]. Conformational changes in proteins can be followed on the submicrosecond timescale as denatured proteins are placed in refolding conditions using stopped-flow CD [68]. As aforementioned spectroscopic methods, CD has low structural resolution compared to X-ray crystallography or NMR as it does not provide residue-specific information, but it is much less demanding and can be performed in a wide range of conditions [69]. As NMR, CD is also particularly valuable in studies of peptides and disordered proteins [70].

8. Fluorescence-based methods. Fluorescence correlation spectroscopy exploits tiny fluctuations of fluorescent signal of a labeled molecule of interest as it enters the small area where the fluorescence is excited [71]. From the autocorrelation of fluorescent signal, the diffusion coefficient can be calculated, and changes in protein conformational state can be extracted [72, 73]. Fluorescent resonance energy transfer (FRET) measurements detect the effect of a donor fluorescent label on an acceptor label that depends on the distance between two labels [74]. Signals from labels on the same molecule provide a wealth of distance information as it folds (act as a “spectroscopic ruler” [75]).

To summarize, current experimental advances has significantly moved the boundaries of explorable in protein folding. It is now possible in certain cases

to observe protein structures at atomic resolution and follow folding processes occurring as fast as at femtosecond timescale.

1.2.2. Bulk and single molecule studies

Conventional bulk assays to study protein folding usually have a large number of biomolecules in the sample providing a signal corresponding to their conformational changes. This signal is averaged upon collection. Transient intermediate protein states appear in the signal as fluctuations. Individual proteins do not fold in synchrony, and thus the signal from the states visited rarely and transiently is cancelled out upon averaging. Hence it becomes unrealistic to detect transient intermediate states as protein folds with bulk assays. On the other hand, if the signal is not averaged over population, but measured from a single protein, intermediate states could readily identified (Fig. 1.5). It is of particular importance for protein folding studies, because these non-native conformations are often key for understanding of protein structure and function. Furthermore, the existence of functional proteins having no single, well-defined native structure attracts significant attention since recently [76, 77]. They lack significant structural elements in their native conformations and are highly dynamic, hence being particularly challenging to study with the bulk methods, where this functional dynamics is often not observed due to ensemble averaging. It is now known that as much as 30% of human proteins possess significant disordered regions [78], thus this problem is of high importance for protein folding field.

Single molecule assays allow to detect individual members of heterogeneous population, quantitatively analyze their subpopulations, and follow their time trajectories (kinetic information) without perturbing the system to push it out of equilibrium. Controlling a number of molecules under study is also beneficial to avoid intermolecular interactions affecting the total signal. For example, well-known effect of protein intermolecular aggregation (when unfolded proteins with hydrophobic cores exposed stick to each other due to hydrophobic interactions) slows down folding process and cannot be discriminated from intramolecular misfolding in functional bulk assays [49]. Folding of single proteins has been studied with two main groups of techniques: fluorescence methods and force spectroscopy. Fluorescence methods involve the introduction of fluorescent labels into specific positions in the protein of interest. The labels are excited with a laser beam and the fluorescent signals are monitored. Changes in fluorescent signals reflect different labels in one or different molecules approaching and affecting each other, resolving molecular interactions and conformations. The signal generated by a pair of fluorophores belonging to the same molecule carries information about the state of a single molecule under study [79]. In force spectroscopy, the protein of interest is immobilized by attaching its termini to various surfaces (e.g. of coverslip, micrometer-sized beads, or cantilever) and then stretched to force it to partially folded or unfolded states. During this process, force acting on the protein and its extension can be monitored with resolution up to sub-picoNewton and sub-nanometer level, allowing one to observe transient intermediate states visited during protein folding and obtain information about

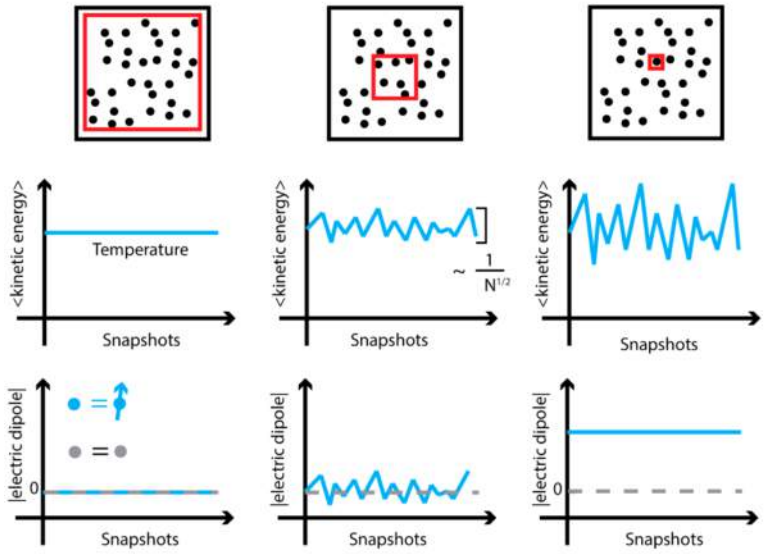


Figure 1.5: **Differences of the results measured in bulk and at single-molecule scale.** Illustration of two intensive properties (top graphs - kinetic energy, bottom graphs - electric dipole moment per molecule) measured in bulk, at mesoscale, and at the single-molecule scale as illustrated by the boxes above. The graphs of dipole moment feature grey lines representing the results for a solution of non-polar molecules, and blue line - for polar molecules. Bulk dipole moment measurements cannot discriminate between these two solutions, but single-molecule methods do it easily. Reprinted with permission from [49]. ©2013 American Chemical Society.

their structure and stability. Next, the protein can be relaxed and stretched again after a certain waiting period to probe its refolding as a function of time. Furthermore, force spectroscopy provides an access to areas of protein energy landscape that are not accessible in normal conditions by lowering the free energy of partially folded or unfolded conformations and allowing for spontaneous fluctuations between different states depending on applied force [80].

1.2.3. Fluorescence methods

Fluorescence resonance energy transfer (FRET)

FRET is a process of energy transfer from an excited donor molecule to an acceptor molecule [81] that depends on the distance between them. Usually, the light intensities emitted by the donor and acceptor or the donor fluorescence duration (that is reduced upon energy transfer to the acceptor) are measured and the FRET efficiency is calculated from these values [82], but more elaborate multiple excitation (for example, ALEX [83]) and multiparameter detection schemes [84] are used for more accurate characterization of different fluorescent species and their stoichiometry. Direct excitation of the acceptor also helps to discrim-

inate between the signals coming from low-FRET population and ”donor-only” population formed, for example, because of incomplete labeling [83].

Single molecule FRET detection (smFRET) ia achieved by diluting the sample to allow freely diffusing labeled molecules enter the excited volume one at a time. Confocal microscopy is used to reduce the excited volume and lower a background signal this way. However, the diffusion time in such an approach limits the observation time to a few milliseconds, providing a researcher only with a snapshot of a process to be studied (e.g., protein folding trajectory) [85]. To observe processes with FRET for extended periods of time, molecules are immobilized on a surface and the excitation schemes based on total internal reflection (TIR) are used instead the confocal excitation [86].

A wide range of biological questions has been addressed using FRET. Examples include conformational transitions in DNA, RNA, proteins, ribosomes, and their interactions with other molecules [82]. A single FRET pair provides distance information in a range of 2-10 nm, a value that is well suited to reflect conformational transitions in proteins. However, this value is not always enough, and a one-dimensional distance reporter does not always completely reflect a complicated folding transition. In order to map the transition in more detail, more labels could be added as implemented in three-color schemes that allowed to study the conformational dynamics of DNA Holliday junction [87]. Three-color ALEX has been recently used to simultaneuosly measure three distances between labels on DNA and monitor the translocation of RNA polymerase on DNA [88].

This powerful technique has already proven itself as a major tool in molecular biology; however, it still has certain drawbacks and limitations. First, smFRET requires perturbing the molecule under study with at least two external dye molecules that quite often are challenging to introduce site-specifically. Second, its distance range does not allow to follow the folding transition of many proteins completely, since conformational changes of hundreds nm can regularly occur upon it.

Photoinduced electron transfer (PET)

Similarly to FRET, electron transfer is a radiationless energy transfer process [89]. However, the distance range of PET differs from the one of FRET: it occurs when the molecules are in Angstrom proximity, thus providing the information that is inaccessible for FRET. The rate of electron transfer can be estimated by monitoring the donor fluorescence quenching in the close proximity of acceptor. The technique was used to monitor the conformational transitions in proteins as they fold or unfold [90].

Fluorescence correlation spectroscopy (FCS)

In this technique, fluorescence fluctuations of labeled molecules are analyzed to detect the events occuring at timescales faster than the signal count rate [91]. In the most basic form, the molecule of interest has a single label that is excited with a single wavelength. Temporal autocorrelation of its signal allows extracting structural and dynamic information about the labeled molecule: its local concentration, diffusion coefficient, number and occupancy of conformational states

[92]. Conformational changes of the molecule can be followed via changes of diffusion coefficient as demonstrated for protein unfolding in denaturant [93]. If the molecule has multiple labels, their cross-correlation function can also be calculated to study the correlated variations of their signals that may be useful for FRET data analysis [94].

1.2.4. Force spectroscopy

Importantly for protein folding studies, the application of force allows to explore the protein energy landscape in more detail. When force is applied, it lowers transition barriers and favors more unfolded states that may not be accessible in equilibrium. These data provide the information about relative populations of certain states and rate constants of transitions between them.

Optical tweezers

Optical tweezers exploit a potential well that forms near the focus of a focused laser beam to trap and manipulate micrometer-sized particles [95]. A molecule of interest is usually linked to the beads and can be stretched and relaxed by manipulating them. Force acting on the bead in the trap and its position can be extracted from the collected laser light [96]. Optical tweezers can apply forces from less than 1pN up to several hundred pN and detect subnanometer length changes that matches well the stability and dimensions of typical biomolecules [97, 98]. Biomolecules are held together by non-covalent bonds that exert comparable cumulative force, hence optical tweezers are perfectly suitable to probe the interactions between and within biomolecules and follow subsequent structural transitions. One more advantage of optical tweezers is that the positions of the traps may be alternated at kilohertz frequency using piezo-driven mechanical, acoustic or electrical systems [99]. It allows to create multiple independent traps by time-sharing between different positions [100] or to combine optical tweezers with fluorescence detection by temporal separation of trapping and fluorescence excitation.[1] Among drawbacks of optical tweezers is a necessity of grabbing both ends of biomolecule under study that may perturb the native structure or function of the molecule, and is challenging to do if its termini are buried inside. Also, optical tweezers, as most of other single molecule force applying instruments so far, are relatively costly and rather complex, requiring solid expertise in physics for construction and often even for operation.

Magnetic tweezers (MTs)

This technique involves usage of permanent magnets to exert stretching force on the molecule of interest via paramagnetic bead. The other end of molecule is typically immobilized on the surface. Magnetic tweezers can apply forces in the wide range from 10 fN to 100 pN. However, the spatial and temporal resolution of MTs are limited to 2-10 nm and 10^{-2} - 10^{-4} s correspondingly, which are lower

[1] It is otherwise not possible to overlap the trapping beam and fluorescence path because of strong background in fluorescent signal coming from trapping laser and strongly enhanced bleaching of fluorophores.

than those of optical tweezers [101]. Furthermore, magnetic tweezers allow applying torque to biomolecules and, after certain modification, measuring it [102]. It provides ability to measure torsional stiffness and control supercoiling, hence making it a very useful technique for various DNA studies, such as DNA-enzyme interaction [103, 104] and chromatin packing [105]. However, this technique has also been applied to study mechanical properties of proteins [106].

Atomic force microscopes (AFM)

AFM exploit mechanical force to pull on surface-immobilized proteins by moving the cantilever tip that acts as the other connection point. Applied force lies in a range from 10 pN to a few nN [107] and is measured by monitoring the cantilever deflection upon stretching. AFM is used extensively for protein folding studies [108]. AFM are particularly useful when forces higher than 100 pN are required, since this force range is generally not accessible by other spectroscopic techniques. Low force resolution is a limitation of AFM, together with the drift [109].

Nanopores

Protein translocation through nanometer-sized pores occurs regularly in the cell when proteins need to be transported to certain compartments separated by membranes. Protein transport through nanopores in the laboratory has shed some light on the mechanisms underlying protein translocation [110]. Furthermore, the protein passing the pore can be characterized by the drop of current through the pore caused by the displacement of electrolyte from the pore. This signal provides information about the folding state of the protein, and depends on the amino acid composition and charge of the protein chain [111]. Both protein pores and channels positioned in lipid membranes and synthetic nanopores made in silicon-based membranes were utilized to study proteins [112–114].

Novel “simple” single molecule force spectroscopy

As already discussed in previous sections, precise manipulation of single molecules has already provided significant insights in a wide range of areas. However, widespread usage of these techniques is still hindered by relatively high cost and increased complexity in construction and operation compared to more traditional bulk techniques. Recently, several approaches have been demonstrated which allow faster, cheaper and simpler single molecule measurements.

Wong et al. [115] have employed centrifugal force to perform a parallel single-molecule study of the antigen-antibody unbinding kinetics. Multiple molecules were tethered between the coverslip surface and beads, and were stretched by rotating the sample in the centrifuge. Extension and rupture of tethers resulted in beads displacement that can be observed on the microscope image.

Yet another example of simpler single molecule assay is an acoustic manipulation device developed by Sitters et al. [116] This method employs acoustic standing wave generated by a piezo element to stretch multiple molecules individually tethered between the coverslip and the beads. The device performance

is demonstrated in parallel DNA overstretching and antigen-antibody bond rupture studies.

These techniques have quite some potential for protein folding measurements due to their simplicity, wide force range, less stringent requirements for the beads properties and the absence of oxidative damage of the sample. However, the distance resolution of these methods is limited by the beads tracking routine (typically, about 5 nm), and thus is lower than of optical tweezers, but in the same range with magnetic tweezers method.

1.3. Chaperone-assisted protein folding

1.3.1. Chaperones definition and classification

Almost all native structures of globular proteins reported to date have dry hydrophobic core with residues tightly packed together. However, when proteins are synthesized at the ribosome, they are extended and their hydrophobic residues are still solvated. Protein secondary structure forms quite fast, but completing the tertiary structure is rather slow and sometimes may last for minutes [117]. Proteins, before they acquire their native fold, diffuse freely in the cellular environment, which is very crowded with lots of biomolecules interacting with each other. Hence proteins are at risk of formation of undesired interactions that hinder the folding yield. For example, if protein chains in non-native states are in close proximity, their hydrophobic fragments may stick to each other, increasing the chance of improper folding and intermolecular aggregation. To account for such a scenario, cells possess specialized proteins that assist other proteins to fold properly. These proteins are called molecular chaperones.

The term ”molecular chaperones” has been used first by Laskey et al. in 1978 to describe a protein shielding histones [118], and has been later repurposed to proteins preventing unwanted protein-protein interactions. It was believed in the 70’s that protein folding is solely thermodynamically driven and that the native protein structure is always accessible in the cell, in accordance with Anfinsen’s dogma. However, after his landmark work on ribonuclease, it was shown both in vitro and in vivo that many other proteins do not refold that well spontaneously and form insoluble aggregates instead [119]. On the other hand, it became clear that a class of proteins called ”heat shock proteins” is capable of protection of cell under stress [120]. These findings were summarized in a model by Pelham [121], where he proposed a heat shock induced 70 kDa protein called Hsp70 to bind and rescue aggregating proteins in a cycle coupled with ATP hydrolysis [122].

A wide range of molecular chaperones maintains proteome quality and integrity in the cell during the complete life cycle. Chaperones are active in de novo protein folding, prevention of protein aggregation upon various stresses, solubilization of protein aggregates, protein transport between different cellular compartments. ATP-independent chaperones, such as small heat shock proteins (sHsps) and SecB are generally thought to act as ”holdases”, which bind protein clients that were unfolded upon various stresses and prevent their aggregation until the stress is gone. More complex chaperones, such as chaperonins (Hsp60) and

Hsp 70, can act as "foldases" by utilizing the energy of ATP hydrolysis to maintain the cycle of client binding and release, promoting its native folding. However, this classification is not exclusive, since all classes of chaperones are involved in a multitude of processes. For example, ATP-independent chaperone trigger factor promotes native protein folding by shaping its energy landscape through interaction with partial folds [123].

Chaperones of different classes often collaborate, acting on the same protein at different folding stages. Trigger factor is associated with ribosomes and already acts on the emerging protein chain cotranslationally. For many proteins, such assistance is enough to fold successfully, but less efficient folders are transferred downstream to Hsp70 or Hsp60, which guide them towards a native structure [124]. Upon stress conditions, sHsps bind unfolded proteins and coaggregate with them, promoting subsequent disaggregation and refolding by other chaperones such as Hsp70 and Hsp100 [125].

1.3.2. Hsp70

The heat shock protein 70 kDa (Hsp70) system is the central hub of chaperone activity in stressed and unstressed cells of all domains of life [126]. It contributes to a wide range of folding processes, such as folding of newly synthesized polypeptides and maintenance of the conformational state of existing proteins [127]. Hsp70s are highly conserved from bacteria to human, so the bacterial Hsp70 called DnaK can be considered as a model system for the human Hsp70 chaperone. Being a key player in a number of cellular stress responses, it is overexpressed to 13% of the total cellular protein content upon heat stress [128]. In the absence of the stress DnaK also performs critical chaperone functions and represents about 1% of total proteins [129]. It interacts with a broad array of protein substrates, which is estimated to comprise about 25% of the cytosolic proteome [126].

In the framework of the canonical model, DnaK improves folding by cyclically clamping and releasing peptide chains. According to the allosteric mechanism underlying DnaK chaperone activity, ATP binding to the N-terminal nucleotide binding domain (NBD) switches between substrate affinity modes in the C-terminal substrate binding domain (SBD), and, the other way around, protein substrate binding to SBD to stimulate ATP hydrolysis in NBD [130]. Client binding occurs through the association of an extended hydrophobic peptide chain with the substrate binding groove located in the ß-subdomain of SBD. The α-helical lid subdomain of SBD secures the bound substrate by closing tightly over the groove. Clamping by a closed lid domain provides protection from aggregation, while release upon lid opening allows for folding or transfer to other chaperones (Fig. 1.6a-b). This ATP-driven cycle is mediated by the co-chaperones DnaJ and GrpE.

This model is corroborated by solved structures of the peptide-chaperone complex [131], and is consistent with the observation that DnaK can destabilize proteins that are folded [132], misfolded [133], and aggregated [134]. However, whether DnaK indeed assists folding by binding and releasing peptides is unclear, given the inherent difficulty to study non-native client states in complexes with

DnaK.

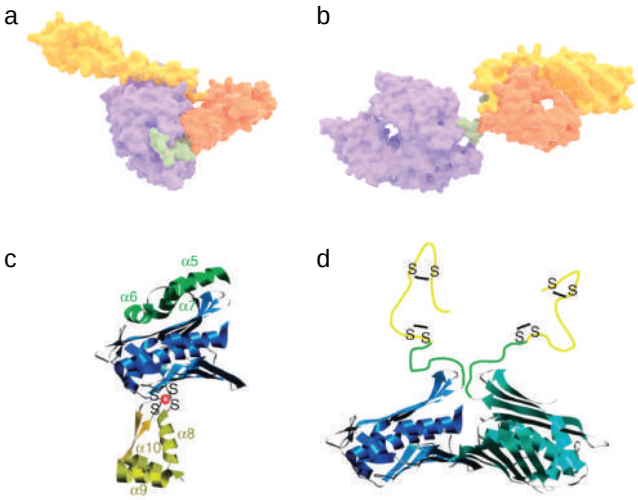


Figure 1.6: **Chaperones structures and functional conformational changes. a-b,** Structure of *E. coli* DnaK with domain boundaries indicated in different colors: nucleotide binding domain (purple, 1-383), hydrophobic linker (green, 384-396), substrate binding domain- ß-sheet subdomain with groove (orange, 397-502), substrate binding domain – α-helical lid subdomain (yellow, 503-602). **a,** DnaK bound to peptide substrate and ADP nucleotide (in ADP state, PDB code: 2KHO). **b,** DnaK in ATP bound open conformation (PDB code: 4B9Q). **c-d,** Domain structure of *E. coli* Hsp33 color-coded as follows: N-terminal domain is shown in blue, linker domain - in green, C-terminal domain - in yellow. **c,** Structural model of Hsp33 monomer based on the crystal structure of the *B. subtilis* Hsp33 dimer. **d,** Schematic model of active Hsp33 dimer. Figures **(a-b)** are made by Mario Avellaneda. **(c-d)** are reprinted with permission from [49]. ©2013 American Chemical Society.

1.3.3. Hsp33

Hsp33 is a redox-regulated chaperone that is identified in most prokaryotic species and also in some eukaryotic parasites [135]. Hsp33 protects the cell from oxidative stress conditions leading to massive protein unfolding by binding unfolding proteins and preventing their aggregation. To specifically respond to stress conditions, it features regulatory switches sensing the oxidation and activating the chaperone, while in non-stress conditions it mostly stays inactive. Hsp33 has 292 residues and consists of a compactly folded N-terminal domain that is believed to harbor a client binding site, and a C-terminal oxidation sensing domain, with a flexible linker connecting two domains. Oxidation sensing domain has four conserved cysteine residues that coordinate a zinc ion in non-stress conditions. During oxidative stress, cysteines become oxidized and form intramolecular disulphide bonds, whereas zinc ion is released. Formation of disulphide bonds triggers unfolding of both zinc-binding and linker domain and causes dimerization of Hsp33 (Fig. 1.6c-d). This partially disordered state of Hsp33 is characterized by high chaperone activity [136]. Activated Hsp33 has high affinity to secondary

structure regions in client proteins, that is suggested to allow chaperone to effectively protect early unfolding intermediates upon stress and prevent massive aggregation. When the stress is over, Hsp33 transfers client proteins to Hsp70 system that is supposed to facilitate their further refolding [137]. However, little is known so far about the precise structure of chaperone-substrate complexes and the role of conformational changes of Hsp33 in interaction with substrates. It remains to be determined whether the chaperone uses these rearrangements only as a stress sensor to fine-tune its activity or employs the released energy to trigger conformational changes in the client and return it to the productive folding pathway.

1.3.4. Small heat shock proteins (sHsps)

The small heat shock proteins comprise a widespread class of chaperones. They are usually overexpressed upon heat stress, but also often act at non-stress conditions. These proteins are present in all living organisms except for certain bacterial species [138], binding unfolded protein chains in ATP-independent manner and hence preventing their aggregation and facilitating further refolding by partner chaperones. Small heat shock proteins monomers have 15-35 kDa molecular weights and feature a C-terminal α-crystalline domain, consisting of 80-90 residues that form a ß-sheet [139]. Otherwise, their structures are quite diverse. sHsps are known to form oligomers of various sizes having different client binding preferences to achieve truly promiscuous substrate binding [140], a property that is essential to protect the proteome upon heat stress. Oligomerization occurs via α-crystalline domain [141, 142]. Notably, sHsps often undergo posttranslational modifications that further diversifies their properties and function, primarily in N-terminal domain that has a number of phosphorylation sites [143]. In humans, sHsps family is named HspB and features 10 members. For all HspBs, some extent of chaperone activity was detected in vitro, but it remains unclear how physiologically relevant this activity is. Due to their high heterogeneity, little is known about how exactly they interact with client proteins and whether their activity is limited to ”holdase” mode. HspB chaperones are involved in the development of numerous diseases; however, mechanisms underlying this contribution remain elusive.

1.3.5. Experimental advances in chaperone-assisted folding

Hence, despite a significant amount of knowledge that has been acquired in past decades, many important questions about mechanisms underlying chaperone activity remain unadressed. These unknowns are quite challenging to investigate using ensemble techniques, since chaperones often interact with their clients transiently and in a very heterogeneous manner, depending on the conformational states of client and chaperone, energy input and the presence of numerous cofactors. Since proteins are not synchronized in the population, signal from these heterogeneous interactions is often beyond the detection limit in the population-averaged signal. On the other hand, single molecule techniques

are capable of studying individual protein molecules in real time, which suggests that the single-molecule scale is optimal for chaperone-assisted protein folding studies. In recent years, quite some progress has been made in this direction and many detailed studies on the role of chaperones in protein folding processes have been published. Here I will briefly mention some of them to provide a reader with an overview, and for those looking for more detail I refer to an excellent review by Mashaghi et al. [49] At the single-molecule scale, chaperone-assisted protein folding (similarly to non-assisted protein folding) was mostly studied with fluorescence and force spectroscopy methods.

Fluorescence methods

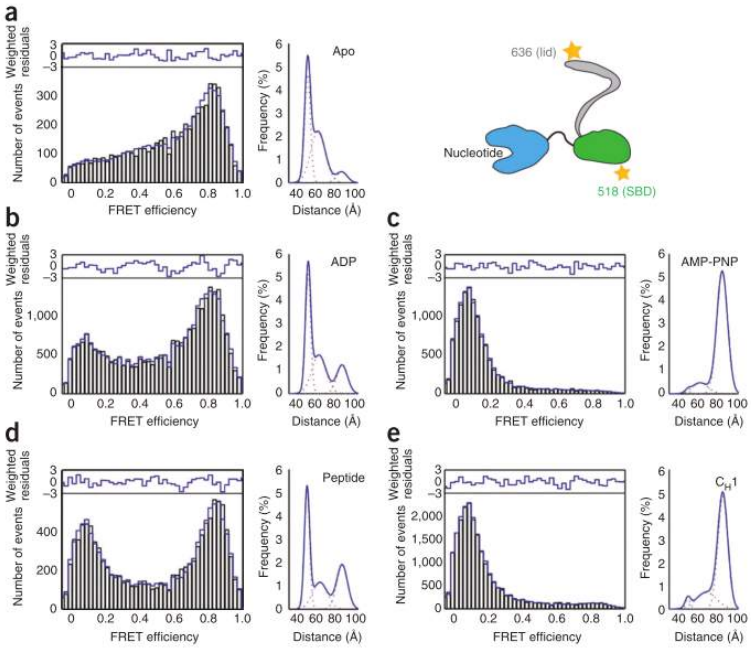


Figure 1.7: **Single-pair FRET analysis of BiP chaperone.** BiP was double-labeled as shown in the cartoon to monitor the distance between the NBD and the SBD. FRET efficiences were measured and their histograms plotted in the absence of nucleotides **(a),** in the presence of ADP **(b),** in the presence of AMP-PNP **(c),** in the presence of the HTFPAVL peptide **(d),** in the presence of C_H1 protein **(e).** Reprinted by permission from Macmillan Publishers Ltd: Nature Structural & Molecular Biology [144], ©2011.

Single-molecule FRET has been used to follow conformational transitions in client protein during chaperone-assisted folding. Sharma et al. studied the acceleration of folding of MBP double mutant (DM-MBP), which folds slowly in the isolation, by GroEL/GroES chaperonin system [145]. The authors introduced two labels in DM-MBP and followed the conformations of the client upon refolding

in the isolation and in the presence of the chaperone. To verify the stoichometry of the chaperone-client complex, they used pulsed interleaved excitation (PIE) in the combination with fluorescent cross-correlation spectroscopy (FCCS) and concluded that GroEL forms the complex with one client. In the isolation, DM-MBP rapidly compacted to a stable intermediate state after dilution from denaturant. However, in the presence of GroEL DM-MBP states were distributed heterogeneously with a fraction of expanded conformations. To evaluate possible biological relevance of GroEL-mediated protein unfolding, the authors followed a functional cycle of the chaperonin by adding ATP and the cofactor GroES. They performed FRET measurements at various positions in DM-MBP sequence and revealed a segmental release of DM-MBP from GroEL upon ATP addition, followed by a compaction indicating an encapsulation of the client protein into GroEL folding cavity.

Conformational transitions in the chaperone itself underlying its functional cycle have also been analyzed using FRET by Marcinowski et al. [144] They studied a mammalian Hsp70 chaperone BiP that assists folding in the endoplasmic reticulum. As other Hsp70 chaperones, BiP is believed to undergo ATP-dependent conformational changes mediating its substrate affinity. To monitor the domain distances within individual BiP molecules upon binding of nucleotides, cofactors and client proteins, the authors developed a spFRET assay by incorporating and labeling cysteines at different positions within nucleotide-binding domain (NBD), substrate-binding domain (SBD) and α-helical lid. NBD and SBD were tightly coupled in the presence of ATP, but the distance between them increased and the distance distribution broadened in the presence of ADP or in the absence of nucleotide, consistently with previously published reports on Hsp70 mechanics. Furthermore, previously unknown effect of the substrate state on the communication between the domains in BiP was also reported: NBD and SBD were coupled more tightly in the presence of a protein substrate than with the peptide. The lid conformation was also significantly affected by the substrate. When a peptide was bound to BiP, major fraction of chaperone molecules populated a closed lid conformation, however, the lid was mostly open upon binding of a client protein (Fig. 1.7). By performing these experiments when an Hsp40 cofactor of BiP ERdJ3 was also present in the system, the authors concluded that these conformational changes shaping the affinity of the chaperone to different substrates are regulated by both the nucleotide and the cofactor. It provided important insights in the understanding of Hsp70 functional cycle.

A bacterial homolog of BiP, DnaK, was investigated in the study by Kellner et al. [146], DnaK substrate protein rhodanese was labeled at different positions and its conformations were monitored by smFRET. Upon denaturation, rhodanese populated a compacted intermediate state. When a bacterial Hsp40, DnaJ, was added to previuosly denatured rhodanese, it resulted in a broadening of transfer efficiency histograms thus suggesting that DnaJ blocks rhodanese refolding by holding it in a heterogeneous ensemble of nonnative conformations that are substantially different from those of denatured rhodanese in the absence of chaperones. The addition of DnaK and ATP to preformed DnaJ-rhodanese complexes led to surpris-

ingly large changes in transfer efficiency histograms (Fig. 1.8). A pronounced shift of the distributions to lower transfer efficiencies was observed, indicating that DnaK binding led to the formation of highly expanded conformations of rhodanese. To obtain a more detailed structural picture of underlying chaperone action, coarse-grained molecular simulations of the chaperone-substrate complex were performed. Both the experiments and the simulations supported the picture of a substrate protein whose chain is almost saturated with DnaK, and thus the volume exclusion between the bulky DnaK molecules underlies the observed expansion. The resulting disruption of intramolecular contacts in the substrate protein may represent an efficient way of rescuing misfolded proteins from stable non-pathway states that otherwise would prevent their native folding.

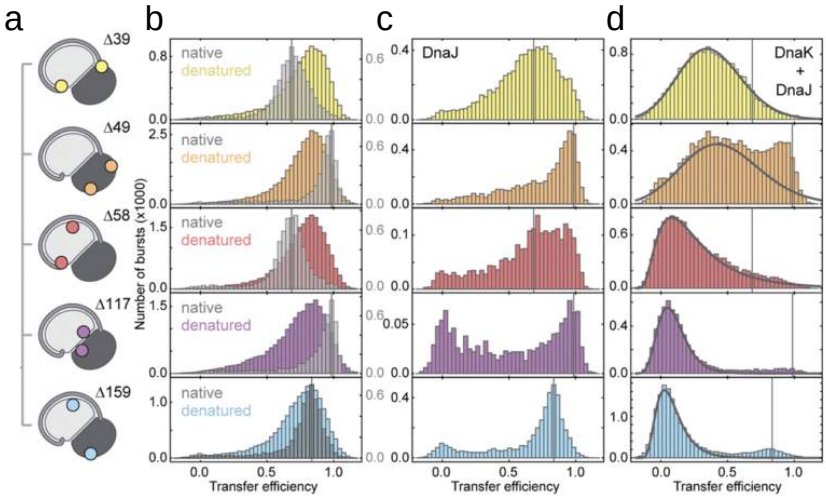


Figure 1.8: **Expansion of rhodanese upon interaction with Hsp70. a,** Cartoon of rhodanese structure with the fluorescent labels positions shown. **b,** FRET efficiencies histograms of native rhodanese with various labels compared with the same results from denatured rhodanese. **c,** FRET efficiencies histograms of denatured rhodanese in complex with DnaJ. **d,** FRET efficiencies histograms of denatured rhodanese in the presence of DnaK and DnaJ. Reprinted from [146].

Force spectroscopy methods

In mechanical manipulation approaches, substrate proteins of chaperones are usually held via termini between two surfaces (either of cantilever and sample chamber in case of AFM or of microparticles in case of optical tweezers) and are stretched and relaxed, when both force acting on them and resulting extension are recorded, to study their unfolding and refolding pathways in the absence and in the presence of chaperones. More rarely, chaperones themselves are immobilized and stretched either to study their conformational changes upon interaction with substrates and cofactors or to directly measure functional forces that they apply to their substrates.

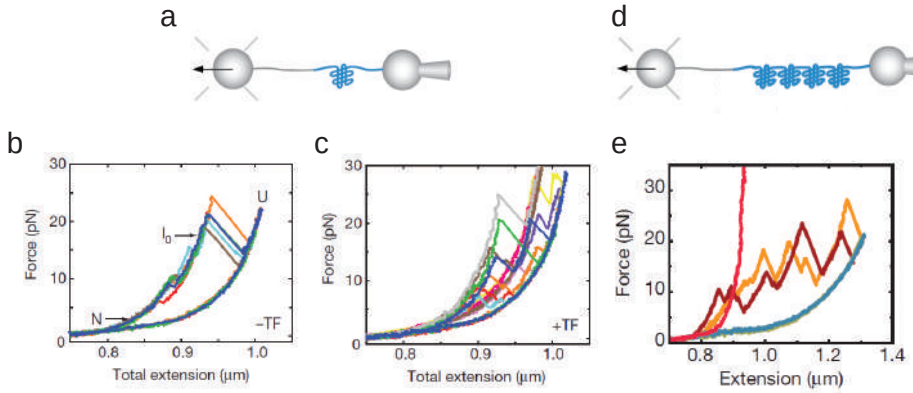


Figure 1.9: **TF chaperone directly shapes the protein folding pathway. a,** Scheme of the experiment presented in the subfigures **b-c**. **b,** Stretching-relaxation experiments with MBP in the absence of TF. MBP repeatedly unfolds in two steps as indicated by the force-extension data. **c,** Stretching-relaxation experiments in the presence of TF, indicating abundant stable partial folds. **d,** Scheme of the experiment presented in the subfigure **e**. **e,** Stretching-relaxation experiments with 4MBP. In the absence of TF (red curve), it forms tight aggregated structure after being unfolded (orange) and relaxed (blue). In the presence of TF (dark red curve), 4MBP unfolds with native-like characteristics and does not aggregate. Adapted by permission from Macmillan Publishers Ltd: Nature [123], ©2013.

Mashaghi et al. used optical tweezers to reveal the chaperone-driven modulation of a folding pathway to promote the native refolding of a protein client [123]. The authors studied the effects of trigger factor (TF), an ATP-independent ribosome-associated bacterial chaperone, on MBP. When MBP was stretched and relaxed by optical tweezers in the presence of TF in experimental buffer, intermediate protein lengths were observed more frequently and for longer times than without the chaperone (Fig. 1.9a-c). It indicated the stabilization of partial folds of MBP by TF. To test an assumption that complex multi-domain proteins are more prone to undesired interactions between domains, and thus may benefit more from the effect of TF, the authors performed stretching experiments with an engineered four-domain repeat of MBP (4MBP) that refolds poorly and has a very strong tendency to misfold after being unfolded in the isolation. Misfolded 4MBP is characterized by protein length as compact as of the native 4MBP structure and by unfolding forces that are often higher than maximum force available in this assay. In the presence of TF, tight misfolding occurs only rarely, and the protein was able to refold in the native-like fashion, as indicated by a characteristic unfolding pattern (Fig. 1.9d-e). The study indicates that the TF can promote native folding of local protein fragments and at the same time suppress long-range interactions. It demonstrated an ability of a chaperone to directly affect the conformational search of a protein and suggested an importance of this interaction mode for nascent protein folding.

Jahn et al. probed mechanical stability of Hsp90 chaperone to reveal subtle yet unknown structural details and link them to the function of the chaperone

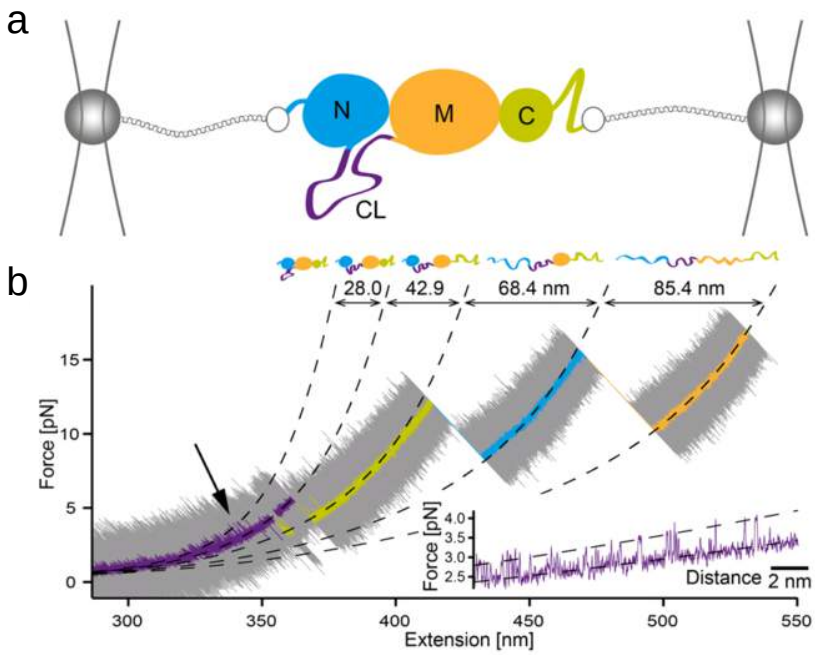


Figure 1.10: **Unfolding of Hsp90 by optical tweezers. a,** Scheme of the experiment. **b,** Stretching-relaxation experiments with Hsp90. Low-force fluctuations indicate the interaction of charged linker of Hsp90 with one of its domains. Reprinted from [147].

[147]. Its three domains are connected via linkers, one of which is charged, that allow the rearrangement of domains during the Hsp90 chaperone cycle. As demonstrated by low-force fluctuations of size similar to the length of the charged linker (CL) in the optical tweezers data before unfolding of the domains sets in, the CL interacted with one of the domains (Fig. 1.10). This interaction was abolished by replacement of the CL by a neutral sequence or by the deletion of N-domain indicating the sequence-specific interaction of the CL with the N-domain. The authors performed smFRET experiments to study the importance of the CL for switching of Hsp90 between open and closed dimer conformations that is considered essential for the function of the chaperone. The distribution of FRET efficiencies of multiple Hsp90 molecules is shifted towards higher values in case of Hsp90 mutants with CL substitution compared to WT, indicating that the mutation of the CL favors the closed conformation of Hsp90. To investigate the implications of these findings for the chaperone function, the authors studied the viability of S. cerevisiae carrying two different CL substitution mutants of Hsp90. Both mutations resulted in growth defects and lower stress tolerance, whereas the substitution of the whole linker had more deleterious effects than the mutation of the linker excluding the region that is expected to dock on the N-domain. This study indicated

the significance of the CL for the chaperone cycle of Hsp90.

In the study by Scholl, Yang and Marszalek, the authors unraveled the mechanism that a chaperone rabbit reticulocyte lysate (RRL) employs to assist its client protein, firefly luciferase, refolding [148]. AFM was used to mechanically induce either partial or complete luciferase unfolding and follow its subsequent refolding. In the absence of the chaperone, luciferase could robustly refold only if partially unfolded, however, it failed to refold and formed a misfolded state after complete unfolding. The presence of RRL allowed luciferase to refold qiuckly, suggesting that the chaperone inhibits non-native contacts in the protein and hence allows it to recover its native stucture after unfolding and relaxation.

1.3.6. Outlook

About 50 years ago, the protein folding problem was born and it became one of major challenges in modern science. Since that time, tremendous progress has been made in understanding how protein chains fold to their native functional state. Furthermore, working on this problem led to a development of a multitude of new experimental techniques, powerful supercomputers, communal databases. Advances in protein folding laid a foundation for new research areas, such as understanding of folding-related diseases and synthesis of designer protein-based materials. However, our understanding of protein folding is far from complete. We cannot consistently predict native protein structure based on the sequence, we cannot model in detail the protein-ligand interaction, we do not have an accurate atomistic picture of folding pathway for arbitrary sequence, we do not know how protein folding is affected by the macromolecular crowding in the cell. Importantly, little is known about the mechanisms underlying folding-related diseases and how to fight them.

Chaperone-assisted folding poses even more challenges, since in this case more factors (chaperones, cofactors, energy input) are involved in the reaction leading to an overall increase of its complexity. Conventional techniques provided a wealth of information on structures of chaperones, their functional cycles and yields of natively folded proteins in the presence of chaperones. However, these techniques are not well suited for mechanistic studies of chaperone-client interactions that determine a folding pathway of the client protein. On the other hand, single-molecule techniques have already been successfully used to directly characterize folding pathways and how they are shaped by chaperones. Furthermore, single-molecule techniques can be used to probe energetic contributions of chaperones to protein folding. Protein energy landscape has already been characterized for certain proteins in the isolation [149, 150], and it would be of great interest to measure the energy flux of ATP-dependent chaperones.

1.4. Scope of this thesis

Central goal of this thesis is to improve the understanding of physical principles underlying protein folding process and the role of chaperones in it. The first part of the thesis focuses on the basic concepts of protein folding and tech-

niques that were used to study it. My approach to protein folding studies is to use optical tweezers. To control and monitor a protein state with high spatial and temporal resolution, a custom assay should be designed and built. In chapter 2 I describe in detail the construction and assembly process of a dual-trap optical tweezers setup. Furthermore, I provide an alignment manual for those who are interested in building similar optical setups and highlight the possibilities of its further development for chaperone-assisted protein folding experiments.

Chapter 3 is dedicated to designing, building and testing of a temperature control unit for optical tweezers. It features a heating chamber around the sample that allows stable and uniform increase of sample temperature by heating the air around objectives and sample chamber. Such a construction minimizes differential drift of the optical components and allows heating while performing a single-molecule stretching experiment. It improves versatility of optical tweezers by providing an opportunity for various heat-shock related studies, as well as for investigation of role of temperature in the mechanics of biomolecules.

Chapter 4 is an in-depth single-molecule study of the DnaK chaperone system action on a protein substrate. This chaperone system is known to interact with the protein in a so-called chaperone cycle that features energy input by ATP hydrolysis and is shaped by co-chaperones. It involves a multitude of transient low-populated states that are beyond the detection limit of most bulk assays, therefore this chaperone system is perfectly suited to be studied with optical tweezers. By studying the effects of individual members of the family on single protein substrate in a certain state, we found a new mode of DnaK-substrate interaction. To our knowledge, it is a first direct proof that DnaK is capable of interacting with partially folded protein substrate via its folded part. Experiments with various mutants of DnaK provide mechanistic insights on its interaction with substrates. Based on our data, we propose a new model of DnaK-assisted de novo protein folding and suggest further implications for a range of DnaK functions in the cells.

Chapters 5 and 6 contain preliminary reports on the mechanisms of chaperone activity of two ATP-independent chaperones, traditionally classified as holdases. However, our data suggest the presence of additional interaction modes, which may possibly be interconnected with the new mode reported for Hsp70 in chapter 4. In chapter 5, I focus on a redox-regulated bacterial chaperone Hsp33 is distinguished from other known classes of chaperones by its mode of functional regulation. Oxidizing conditions cause disulfide bonds to form in Hsp33, a process that leads to its partial unfolding and activation of its chaperone function. Chapter 6 deals with the effect of human small heat shock protein HspB6 on folding and aggregation of an intrinsically disordered protein α-synuclein that is implicated in a range of diseases.

Chaperone systems are known to participate actively in stress response processes in the cells, both as stress sensors and stress buffers. However, other biomolecules can also cooperate with, complement and even replace chaperones in stress protection. In chapter 7, I study the effect of disaccharide trehalose on DNA mechanical stability with optical tweezers. Trehalose is known to protect

a range of biomolecules against various stresses, as well as destabilize certain molecules at certain conditions. The destabilization of DNA by trehalose was observed in our experiments. Using a thermodynamic model, we propose a physical mechanism underlying this effect. Molecular dynamics simulations allow us to get the atomistic details of trehalose-DNA interaction.

References

[1] H. Hartley, *Origin of the Word 'Protein',* Nature **168**, 244 (1951).

[2] J. Sumner, *The isolation and crystallization of the enzyme urease preliminary paper,* J. Biol. Chem. (1926).

[3] H. Wu, *Studies on denaturation of proteins. XIII. A theory of denaturation.* Adv. Protein Chem. **46**, 6 (1995).

[4] A. Mirsky and M. Anson, *Protein coagulation and its reversal: the reversal of the coagulation of hemoglobin,* J. Gen. Physiol. , 133 (1929).

[5] A. Mirsky and L. Pauling, *On the structure of native, denatured, and coagulated proteins,* PNAS **22**, 439 (1936).

[6] L. Pauling and R. Corey, *Configurations of polypeptide chains with favored orientations around single bonds: two new pleated sheets,* PNAS **37** (1951).

[7] F. Sanger and E. Thompson, *The amino-acid sequence in the glycyl chain of insulin. 1. The identification of lower peptides from partial hydrolysates,* Biochem. J. **53** (1953).

[8] F. Sanger and E. Thompson, *The amino-acid sequence in the glycyl chain of insulin. 2. The investigation of peptides from enzymic hydrolysates,* Biochem. J. , 366 (1953).

[9] F. Sanger and H. Tuppy, *The amino-acid sequence in the phenylalanyl chain of insulin. 1. The identification of lower peptides from partial hydrolysates,* Biochem. J. (1951).

[10] F. Sanger and H. Tuppy, *The amino-acid sequence in the phenylalanyl chain of insulin. 2. The investigation of peptides from enzymic hydrolysates,* Biochem. J. **49** (1951).

[11] J. Kendrew, G. Bodo, and H. Dintzis, *A three-dimensional model of the myoglobin molecule obtained by x-ray analysis,* Nature (1958).

[12] M. Perutz, M. Rossmann, and A. Cullis, *Structure of hæmoglobin: a three-dimensional Fourier synthesis at 5.5-Å. resolution, obtained by X-ray analysis,* Nature (1960).

[13] *Structure of Hen Egg-White Lysozyme: A Three-dimensional Fourier Synthesis at 2 ÅResolution,* Nature **206**, 757 (1965).

[14] K. a. Dill, S. B. Ozkan, M. S. Shell, and T. R. Weikl, *The protein folding problem.* Annu. Rev. Biophys. **37**, 289 (2008).

[15] A. Fersht, *From the first protein structures to our current knowledge of protein folding: delights and scepticisms,* Nat. Rev. Mol. Cell Biol. (2008).

[16] C. Anfinsen and E. Haber, *The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain,* PNAS **47**, 1309 (1961).

[17] C. Anfinsen, *Principles that govern the folding of protein chains,* Science **181** (1973).

[18] C. N. Pace, *The stability of globular proteins.* CRC Crit. Rev. Biochem. **3**, 1 (1975).

[19] P. L. Privalov, *Stability of proteins: small globular proteins.* Adv. Protein Chem. **33**, 167 (1979).

[20] K. Dill, *Dominant forces in protein folding,* Biochemistry **29** (1990).

[21] L. Pauling, *The Nature of the Chemical Bond and the Structure of Molecules and Crystals: An Introduction to Modern Structural Chemistry* (Cornell University Press, 1960).

[22] J. M. Berg, J. L. Tymoczko, and L. Stryer, *Biochemistry* (W H Freeman, 2002).

[23] J. Chen and W. E. Stites, *Packing is a key selection factor in the evolution of protein hydrophobic cores,* Biochemistry **40**, 15280 (2001).

[24] A. R. Fersht, *Conformational equilibria in a-and δ-chymotrypsin. The energetics and importance of the salt bridge.* J. Mol. Biol. **64**, 497 (1972).

[25] R. Wolfenden, *Experimental measures of amino acid hydrophobicity and the topology of transmembrane and globular proteins.* J. Gen. Physiol. **129**, 357 (2007).

[26] S. Kamtekar, J. M. Schiffer, H. Xiong, J. M. Babik, and M. H. Hecht, *Protein design by binary patterning of polar and nonpolar amino acids.* Science **262**, 1680 (1993).

[27] C. Wurth, W. Kim, and M. H. Hecht, *Combinatorial approaches to probe the sequence determinants of protein aggregation and amyloidogenicity.* Protein Pept. Lett. **13**, 279 (2006).

[28] D. L. Minor and P. S. Kim, *Context-dependent secondary structure formation of a designed protein sequence.* Nature **380**, 730 (1996).

[29] K. Dill and H. Chan, *From Levinthal to pathways to funnels,* Nat. Struct. Biol. (1997).

[30] M. Anson and A. Mirsky, *The equilibrium between active native trypsin and inactive denatured trypsin,* J. Gen. Physiol. , 393 (1934).

[31] R. Lumry and R. Biltonen, *Validity of the "two-state" hypothesis for conformational transitions of proteins.* Biopolymers **4**, 917 (1966).

[32] C. Levinthal, *How to Fold Graciously,* in *Mossbauer Spectrosc. Biol. Syst. Proc. a Meet. held Allert. House, Monticello, Illinois* (1969) pp. 22 – 24.

[33] H. S. Chan and K. a. Dill, *Protein folding in the landscape perspective: chevron plots and non-Arrhenius kinetics.* Proteins **30**, 2 (1998).

[34] O. B. Ptitsyn, *Stages in the mechanism of self-organization of protein molecules.* Dokl. Akad. Nauk SSSR **210**, 1213 (1973).

[35] D. a. Dolgikh, R. I. Gilmanshin, E. V. Brazhnikov, V. E. Bychkova, G. V. Semisotnov, Venyaminov SYu, and O. B. Ptitsyn, *Alpha-Lactalbumin: compact state with fluctuating tertiary structure?* FEBS Lett. **136**, 311 (1981).

[36] C. Tanford, *Isothermal unfolding of globular proteins in aqueous urea solutions,* J. Am. Chem. Soc. **86**, 2050 (1964).

[37] S. Harrison and R. Durbin, *Is there a single pathway for the folding of a polypeptide chain?* PNAS **82**, 4028 (1985).

[38] J. D. Bryngelson and P. G. Wolynes, *Spin glasses and the statistical mechanics of protein folding.* Proc. Natl. Acad. Sci. U. S. A. **84**, 7524 (1987).

[39] K. Dill, *Theory for the folding and stability of globular proteins,* Biochemistry , 1501 (1985).

[40] R. Zwanzig, A. Szabo, and B. Bagchi, *Levinthal's paradox.* Proc. Natl. Acad. Sci. **89**, 20 (1992).

[41] A. I. Bartlett and S. E. Radford, *An expanding arsenal of experimental methods yields an explosion of insights into protein folding mechanisms.* Nat. Struct. Mol. Biol. **16**, 582 (2009).

[42] S. E. Jackson and A. R. Fersht, *Folding of chymotrypsin inhibitor 2. 1. Evidence for a two-state transition,* Biochemistry **30**, 10428 (1991).

[43] L. S. Itzhaki, D. E. Otzen, and a. R. Fersht, *The structure of the transition state for folding of chymotrypsin inhibitor 2 analysed by protein engineering methods: evidence for a nucleation-condensation mechanism for protein folding.* J. Mol. Biol. **254**, 260 (1995).

[44] A. Fersht, *Optimization of rates of protein folding: the nucleation-condensation mechanism and its implications,* PNAS **92**, 10869 (1995).

[45] A. Li and V. Daggett, *Molecular dynamics simulation of the unfolding of barnase: characterization of the major intermediate.* J. Mol. Biol. **275**, 677 (1998).

[46] F. Hopkins, *Denaturation of Proteins by Urea and Related Substances,* Nature **126** (1930).

[47] A. Fischer, *Heat Denaturation of Proteins as a Chain Reaction,* Nature **137** (1936).

[48] C. M. Dobson, *Experimental investigation of protein folding and misfolding.* Methods **34**, 4 (2004).

[49] A. Mashaghi, G. Kramer, D. C. Lamb, M. P. Mayer, and S. J. Tans, *Chaperone action at the single-molecule level.* Chemical reviews **114**, 660 (2014).

[50] L. Giehm, N. Lorenzen, and D. E. Otzen, *Assays for α-synuclein aggregation.* Methods **53**, 295 (2011).

[51] A. Zarrine-Afsar and A. R. Davidson, *The analysis of protein folding kinetic data produced in protein engineering experiments.* Methods **34**, 41 (2004).

[52] A. Matouschek, J. Kellis, L. Serrano, and A. Fersht, *Mapping the transition state and pathway of protein folding by protein engineering,* Nature (1989).

[53] F. Schotte, M. Lim, T. Jackson, and A. Smirnov, *Watching a protein as it functions with 150-ps time-resolved X-ray crystallography,* Science **300**, 1944 (2003).

[54] H. Dyson and P. Wright, *Elucidation of the protein folding landscape by NMR,* Methods Enzymol. **394** (2005).

[55] J. Baum, C. Dobson, P. Evans, and C. Hanley, *Characterization of a partly folded protein by NMR methods: studies on the molten globule state of guinea pig. alpha.-lactalbumin,* Biochemistry , 7 (1989).

[56] M. Williamson, T. Havel, and K. Wüthrich, *Solution conformation of proteinase inhibitor IIA from bull seminal plasma by 1 H nuclear magnetic resonance and distance geometry,* J. Mol. Biol. , 295 (1985).

[57] D. M. Korzhnev and L. E. Kay, *Probing invisible, low-populated States of protein molecules by relaxation dispersion NMR spectroscopy: an application to protein folding.* Acc. Chem. Res. **41**, 442 (2008).

[58] M. M. G. Krishna, L. Hoang, Y. Lin, and S. W. Englander, *Hydrogen exchange methods to study protein folding.* Methods **34**, 51 (2004).

[59] A. Barth, *Infrared spectroscopy of proteins,* Biochim. Biophys. Acta - Bioenerg. **1767**, 1073 (2007).

[60] H. Fabian and D. Naumann, *Methods to study protein folding by stopped-flow FT-IR*. Methods **34**, 28 (2004).

[61] C. Kolano, J. Helbing, M. Kozinski, W. Sander, and P. Hamm, *Watching hydrogen-bond dynamics in a beta-turn by transient two-dimensional infrared spectroscopy*. Nature **444**, 469 (2006).

[62] S. Oladepo, K. Xiong, Z. Hong, and S. Asher, *Elucidating Peptide and Protein Structure and Dynamics: UV Resonance Raman Spectroscopy,* J. Phys. Chem. Lett. **2**, 334 (2011).

[63] I. K. Lednev, A. S. Karnoup, M. C. Sparrow, and S. A. Asher, *α-Helix Peptide Folding and Unfolding Activation Barriers: A Nanosecond UV Resonance Raman Study,* Journal of the American Chemical Society **121**, 8074 (1999).

[64] C. Y. Huang, G. Balakrishnan, and T. G. Spiro, *Early events in apomyoglobin unfolding probed by laser T-jump/UV resonance raman spectroscopy,* Biochemistry **44**, 15734 (2005).

[65] A. L. Serrano, M. M. Waegele, and F. Gai, *Spectroscopic studies of protein folding: Linear and nonlinear methods,* Protein Sci. **21**, 157 (2012).

[66] N. J. Greenfield, *Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions*. Nat. Protoc. **1**, 2527 (2006).

[67] N. J. Greenfield, *Using circular dichroism spectra to estimate protein secondary structure*. Nat. Protoc. **1**, 2876 (2006).

[68] M. Pflumm, J. Luchins, and S. Beychok, *Stopped-flow circular dichroism*. Methods Enzymol. **130**, 519 (1986).

[69] S. M. Kelly, T. J. Jess, and N. C. Price, *How to study proteins by circular dichroism,* Biochim. Biophys. Acta - Proteins Proteomics **1751**, 119 (2005).

[70] J. Hope, M. Shearman, H. Baxter, a. Chong, S. Kelly, and N. Price, *Cytotoxicity of Prion Protein Peptide (PrP) Differs in Mechanism from the Cytotoxic Activity of the Alzheimer's Disease Amyloid Peptide, AB25-35,* Neurodegeneration **5**, 1 (1996).

[71] E. Haustein and P. Schwille, *Fluorescence correlation spectroscopy: novel variations of an established technique*. Annu. Rev. Biophys. Biomol. Struct. **36**, 151 (2007).

[72] K. Chattopadhyay, E. L. Elson, and C. Frieden, *The kinetics of conformational fluctuations in an unfolded protein measured by fluorescence methods*. Proc. Natl. Acad. Sci. U. S. A. **102**, 2385 (2005).

[73] E. Bismuto, E. D. Maggio, S. Pleus, M. Sikor, C. Rocker, G. U. Nienhaus, and D. C. Lamb, *Molecular dynamics simulation of the acidic compact state of apomyoglobin from yellowfin tuna,* Proteins Struct. Funct. Bioinforma. **74**, 273 (2009).

[74] J. Buchner, *Protein folding handbook, Volume 3* (Wiley-VCH, 2005).

[75] L. Stryer and R. P. Haugland, *Energy transfer: a spectroscopic ruler.* Proc. Natl. Acad. Sci. U. S. A. **58**, 719 (1967).

[76] V. N. Uversky, *Intrinsically disordered proteins and their (disordered) proteomes in neurodegenerative disorders.* Front. Aging Neurosci. **7**, 18 (2015).

[77] V. N. Uversky, *Proteins without unique 3D structures: biotechnological applications of intrinsically unstable/disordered proteins.* Biotechnol. J. **10**, 356 (2015).

[78] a. K. Dunker, J. D. Lawson, C. J. Brown, R. M. Williams, P. Romero, J. S. Oh, C. J. Oldfield, A. M. Campen, C. M. Ratliff, K. W. Hipps, J. Ausio, M. S. Nissen, R. Reeves, C. Kang, C. R. Kissinger, R. W. Bailey, M. D. Griswold, W. Chiu, E. C. Garner, and Z. Obradovic, *Intrinsically disordered protein,* J. Mol. Graph. Model. **19**, 26 (2001).

[79] X. Michalet, S. Weiss, and M. Jäger, *Single-Molecule Fluorescence Studies of Protein Folding and Conformational Dynamics,* Chem. Rev. **106**, 1785 (2006).

[80] M. T. Woodside and S. M. Block, *Reconstructing folding energy landscapes by single-molecule force spectroscopy.* Annual review of biophysics **43**, 19 (2014).

[81] T. Förster, *Zwischenmolekulare Energiewanderung und Fluoreszenz,* Ann. Phys. **437**, 55 (1948).

[82] R. Roy, S. Hohng, and T. Ha, *A practical guide to single-molecule FRET.* Nat. Methods **5**, 507 (2008).

[83] A. N. Kapanidis, N. K. Lee, T. a. Laurence, S. Doose, E. Margeat, and S. Weiss, *Fluorescence-aided molecule sorting: analysis of structure and interactions by alternating-laser excitation of single molecules.* Proc. Natl. Acad. Sci. U. S. A. **101**, 8936 (2004).

[84] J. Widengren, V. Kudryavtsev, M. Antonik, S. Berger, M. Gerken, and C. a. Seidel, *Single-molecule detection and identification of multiple species by multiparameter fluorescence detection 1,* Anal.Chem. **78**, 2039 (2006).

[85] B. Schuler, E. Lipman, and W. Eaton, *Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy,* Nature , 743 (2002).

[86] E. V. Amirgoulova, J. Groll, C. D. Heyes, T. Ameringer, C. Röcker, M. Möller, and G. U. Nienhaus, *Biofunctionalized polymer surfaces exhibiting minimal interaction towards immobilized proteins,* ChemPhysChem **5**, 552 (2004).

[87] S. a. McKinney, A.-C. Déclais, D. M. J. Lilley, and T. Ha, *Structural dynamics of individual Holliday junctions.* Nat. Struct. Biol. **10**, 93 (2003).

[88] N. K. Lee, A. N. Kapanidis, H. R. Koh, Y. Korlann, S. O. Ho, Y. Kim, N. Gassman, S. K. Kim, and S. Weiss, *Three-color alternating-laser excitation of single molecules: monitoring multiple interactions and distances.* Biophys. J. **92**, 303 (2007).

[89] S. Doose, H. Neuweiler, and M. Sauer, *Fluorescence quenching by photoinduced electron transfer: A reporter for conformational dynamics of macromolecules,* ChemPhysChem **10**, 1389 (2009).

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

[91] E. L. Elson and D. Magde, *Fluorescence correlation spectroscopy. I. Conceptual basis and theory,* Biopolymers **13**, 1 (1974).

[92] S. Charier, A. Meglio, D. Alcor, E. Cogné-Laage, J. F. Allemand, L. Jullien, and A. Lemarchand, *Reactant concentrations from fluorescence correlation spectroscopy with tailored fluorescent probes. An example of local calibration-free pH measurement,* J. Am. Chem. Soc. **127**, 15491 (2005).

[93] K. Chattopadhyay, S. Saffarian, E. L. Elson, and C. Frieden, *Measuring unfolding of proteins in the presence of denaturant using fluorescence correlation spectroscopy.* Biophys. J. **88**, 1413 (2005).

[94] C. Eggeling, P. Kask, D. Winkler, and S. Jäger, *Rapid analysis of Forster resonance energy transfer by two-color global fluorescence correlation spectroscopy: trypsin proteinase reaction.* Biophys. J. **89**, 605 (2005).

[95] A. Ashkin, J. M. Dziedzic, J. E. Bjorkholm, and S. Chu, *Observation of a single-beam gradient force optical trap for dielectric particles,* Opt. Lett. **11**, 288 (1986).

[96] F. Gittes and C. F. Schmidt, *Interference model for back-focal-plane displacement detection in optical tweezers.* Opt. Lett. **23**, 7 (1998).

[97] E. a. Abbondanzieri, W. J. Greenleaf, J. W. Shaevitz, R. Landick, and S. M. Block, *Direct observation of base-pair stepping by RNA polymerase.* Nature **438**, 460 (2005).

[98] X. Zhang, L. Ma, and Y. Zhang, *High-resolution optical tweezers for single-molecule manipulation.* Yale J. Biol. Med. **86**, 367 (2013).

[99] K. Neuman and S. Block, *Optical trapping,* Rev. Sci. Instrum. **75**, 2787 (2004).

[100] K. Visscher and S. M. Block, *Versatile optical traps with feedback control.* Methods Enzymol. **298**, 460 (1998).

[101] I. D. Vilfan, J. Lipfert, D. A. Koster, S. G. Lemay, and N. H. Dekker, *Magnetic Tweezers for Single-Molecule Experiments,* in *Handb. Single-Molecule Biophys.* (2009) pp. 371–395.

[102] J. Lipfert, J. W. J. Kerssemakers, T. Jager, and N. H. Dekker, *Magnetic torque tweezers: measuring torsional stiffness in DNA and RecA-DNA filaments.* Nat. Methods **7**, 977 (2010).

[103] A. Revyakin, C. Liu, R. H. Ebright, and T. R. Strick, *Abortive initiation and productive initiation by RNA polymerase involve DNA scrunching.* Science **314**, 1139 (2006).

[104] N. H. Dekker, V. V. Rybenkov, M. Duguet, N. J. Crisona, N. R. Cozzarelli, D. Bensimon, and V. Croquette, *The mechanism of type IA topoisomerases.* Proc. Natl. Acad. Sci. U. S. A. **99**, 12126 (2002).

[105] P. Gupta, J. Zlatanova, and M. Tomschik, *Nucleosome assembly depends on the torsion in the DNA molecule: a magnetic tweezers study.* Biophys. J. **97**, 3150 (2009).

[106] H. Chen, H. Fu, X. Zhu, P. Cong, F. Nakamura, and J. Yan, *Improved high-force magnetic tweezers for stretching and refolding of proteins and short DNA,* Biophys. J. **100**, 517 (2011).

[107] K. C. Neuman and A. Nagy, *Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy.* Nat. Methods **5**, 491 (2008).

[108] R. Garcia and E. T. Herruzo, *The emergence of multifrequency force microscopy.* Nat. Nanotechnol. **7**, 217 (2012).

[109] G. Zoldák and M. Rief, *Force as a single molecule probe of multidimensional protein energy landscapes.* Curr. Opin. Struct. Biol. **23**, 48 (2013).

[110] W. Wickner and R. Schekman, *Protein translocation across biological membranes.* Science **310**, 1452 (2005).

[111] D. S. Talaga and J. Li, *Single-molecule protein unfolding in solid state nanopores,* J. Am. Chem. Soc. **131**, 9287 (2009).

[112] A. Han, G. Schürmann, G. Mondin, R. A. Bitterli, N. G. Hegelbach, N. F. de Rooij, and U. Staufer, *Sensing protein molecules using nanofabricated pores,* Appl. Phys. Lett. **88**, 093901 (2006).

[113] G. Oukhaled, J. Mathé, A.-L. Biance, L. Bacri, J.-M. Betton, D. Lairez, J. Pelta, and L. Auvray, *Unfolding of Proteins and Long Transient Conformations Detected by Single Nanopore Recording*, Phys. Rev. Lett. **98**, 158101 (2007).

[114] J. W. F. Robertson, C. G. Rodrigues, V. M. Stanford, K. A. Rubinson, O. V. Krasilnikov, and J. J. Kasianowicz, *Single-molecule mass spectrometry in solution using a solitary nanopore*. Proc. Natl. Acad. Sci. U. S. A. **104**, 8207 (2007).

[115] K. Halvorsen and W. P. Wong, *Massively parallel single-molecule manipulation using centrifugal force*. Biophysical journal **98**, L53 (2010).

[116] G. Sitters, D. Kamsma, G. Thalhammer, M. Ritsch-Marte, E. J. G. Peterman, and G. J. L. Wuite, *Acoustic force spectroscopy*. Nat. Methods **12** (2015), 10.1038/nmeth.3183.

[117] T. E. Creighton, *Toward a better understanding of protein folding pathways*. Proc. Natl. Acad. Sci. **85**, 5082 (1988).

[118] R. A. Laskey, B. M. Honda, A. D. Mills, and J. T. Finch, *Nucleosomes are assembled by an acidic protein which binds histones and transfers them to DNA,* Nature **275**, 416 (1978).

[119] D. C. Williams, R. M. Van Frank, W. L. Muth, and J. P. Burnett, *Cytoplasmic inclusion bodies in Escherichia coli producing biosynthetic human insulin proteins*. Science **215**, 687 (1982).

[120] S. Munro and H. R. Pelham, *Use of peptide tagging to detect proteins expressed from cloned genes: deletion mapping functional domains of Drosophila hsp 70*. EMBO J. **3**, 3087 (1984).

[121] S. Munro and H. R. Pelham, *An Hsp70-like protein in the ER: identity with the 78 kd glucose-regulated protein and immunoglobulin heavy chain binding protein*. Cell **46**, 291 (1986).

[122] A. L. Horwich, *Molecular chaperones in cellular protein folding: the birth of a field*. Cell **157**, 285 (2014).

[123] A. Mashaghi, G. Kramer, P. Bechtluft, B. Zachmann-Brand, A. J. M. Driessen, B. Bukau, and S. J. Tans, *Reshaping of the conformational search of a protein by the chaperone trigger factor*. Nature **500**, 98 (2013).

[124] T. Langer, C. Lu, H. Echols, J. Flanagan, M. K. Hayer, and F. U. Hartl, *Successive action of DnaK, DnaJ and GroEL along the pathway of chaperone-mediated protein folding*. Nature **356**, 683 (1992).

[125] J. Tyedmers, A. Mogk, and B. Bukau, *Cellular strategies for controlling protein aggregation*. Nat. Rev. Mol. Cell Biol. **11**, 777 (2010).

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

[127] M. P. Mayer and B. Bukau, *Hsp70 chaperones: cellular functions and molecular mechanism*. Cell. Mol. Life Sci. **62**, 670 (2005).

[128] S. L. Herendeen, R. A. VanBogelen, and F. C. Neidhardt, *Levels of major proteins of Escherichia coli during growth at different temperatures*. J. Bacteriol. **139**, 185 (1979).

[129] J. Gamer, H. Bujard, and B. Bukau, *Physical interaction between heat shock proteins DnaK, DnaJ, and GrpE and the bacterial heat shock transcription factor sigma 32*. Cell **69**, 833 (1992).

[130] A. Zhuravleva, E. M. Clerico, and L. M. Gierasch, *An interdomain energetic tug-of-war creates the allosterically active state in Hsp70 molecular chaperones*. Cell **151**, 1296 (2012).

[131] X. Zhu, X. Zhao, W. F. Burkholder, A. Gragerov, C. M. Ogata, M. E. Gottesman, and W. A. Hendrickson, *Structural analysis of substrate binding by the molecular chaperone DnaK*. Science **272**, 1606 (1996).

[132] F. Rodriguez, F. Arsène-Ploetze, W. Rist, S. Rüdiger, J. Schneider-Mergener, M. P. Mayer, and B. Bukau, *Molecular basis for regulation of the heat shock transcription factor sigma32 by the DnaK and DnaJ chaperones*. Mol. Cell **32**, 347 (2008).

[133] S. K. Sharma, P. De los Rios, P. Christen, A. Lustig, and P. Goloubinoff, *The kinetic parameters and energy cost of the Hsp70 chaperone as a polypeptide unfoldase*. Nat. Chem. Biol. **6**, 914 (2010).

[134] J. R. Glover and S. Lindquist, *Hsp104, Hsp70, and Hsp40: a novel chaperone system that rescues previously aggregated proteins*. Cell **94**, 73 (1998).

[135] U. Jakob, W. Muse, M. Eser, and J. C. Bardwell, *Chaperone activity with a redox switch*. Cell **96**, 341 (1999).

[136] C. Kumsta and U. Jakob, *Redox-regulated chaperones*. Biochemistry **48**, 4666 (2009).

[137] D. Reichmann, Y. Xu, C. M. Cremers, M. Ilbert, R. Mittelman, M. C. Fitzgerald, and U. Jakob, *Order out of disorder: working cycle of an intrinsically unfolded chaperone*. Cell **148**, 947 (2012).

[138] E. V. Mymrikov, A. S. Seit-Nebi, and N. B. Gusev, *Large potentials of small heat shock proteins*. Physiol. Rev. **91**, 1123 (2011).

[139] C. Garrido, C. Paul, R. Seigneuric, and H. H. Kampinga, *The small heat shock proteins family: the long forgotten chaperones*. Int. J. Biochem. Cell Biol. **44**, 1588 (2012).

[140] J. L. P. Benesch, M. Ayoub, C. V. Robinson, and J. A. Aquilina, *Small heat shock protein activity is regulated by variable oligomeric substructure.* J. Biol. Chem. **283**, 28513 (2008).

[141] K. K. Kim, R. Kim, and S. H. Kim, *Crystal structure of a small heat-shock protein.* Nature **394**, 595 (1998).

[142] R. L. van Montfort, E. Basha, K. L. Friedrich, C. Slingsby, and E. Vierling, *Crystal structure and assembly of a eukaryotic small heat shock protein.* Nat. Struct. Biol. **8**, 1025 (2001).

[143] H. S. McHaourab, J. A. Godar, and P. L. Stewart, *Structure and mechanism of protein stability sensors: chaperone activity of small heat shock proteins.* Biochemistry **48**, 3828 (2009).

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

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

[146] R. Kellner, H. Hofmann, A. Barducci, B. Wunderlich, D. Nettels, and B. Schuler, *Single-molecule spectroscopy reveals chaperone-mediated expansion of substrate protein,* Proc. Natl. Acad. Sci. **111**, 13355 (2014).

[147] M. Jahn, A. Rehn, B. Pelz, B. Hellenkamp, K. Richter, M. Rief, J. Buchner, and T. Hugel, *The charged linker of the molecular chaperone Hsp90 modulates domain contacts and biological function,* Proc. Natl. Acad. Sci. **111**, 17881 (2014).

[148] Z. N. Scholl, W. Yang, and P. E. Marszalek, *Chaperones rescue luciferase folding by separating its domains.* The Journal of biological chemistry **289**, 28607 (2014).

[149] C. Cecconi, E. a. Shank, C. Bustamante, and S. Marqusee, *Direct observation of the three-state folding of a single protein molecule.* Science **309**, 2057 (2005).

[150] J. P. Junker, F. Ziegler, and M. Rief, *Ligand-dependent equilibrium fluctuations of single calmodulin molecules.* Science **323**, 633 (2009).

2

Dual-trap optical tweezers design, construction and alignment

In last 20 years, optical tweezers were demonstrated to be a great tool to study physical and biological systems at different scales of matter organization: cells and tissues as well as single molecules. However, commercial manufacturing of these systems is not fully set up yet, and in-house assembly of optical tweezers remains quite a challenging task, especially for biology labs. This chapter presents a brief overview of physics underlying optical traps formation and particles manipulation, followed by thorough description of the setup design dictated by the nature of our assay. Furthermore, the chapter features a step-by-step construction and alignment manual, including common pitfalls, which is dedicated to help beginners to design and build optical tweezers setup from scratch.

2.1. Introduction

2.1.1. Basic principles of optical trapping

Existence of radiation pressure had been first suggested about 400 years ago by Johannes Kepler, but this effect was applied to manipulate matter only recently. Optical tweezers were invented in 1986 at AT&T Bell Laboratories by Ashkin, Chu, and coworkers [1]. They used a single tightly focused laser beam to trap a dielectric particle in three dimensions. At that time, optical tweezers entered the lab to further develop later in an important manipulation and measurement tool widely used in physics, biology, medicine and materials science. Optical tweezers employ the ability of a focused light beam to attract a transparent dielectric particle to the maximum of light intensity due to the radiation pressure that light exerts on the particle.

For particles which are much smaller than the light wavelength ($d \ll \lambda$), this effect is described in the context of Rayleigh scattering. The particle is polarized by the light and can be treated as a point dipole. The electromagnetic field exerts force on it, which can be derived from the Lorentz force law and Maxwell equations, and attracts it to the laser focus (region of highest energy density). Every photon carries a momentum, which changes when the light is scattered on the particle. Net force $\vec{F}$ acting on the particle is determined by the difference in momentum flux between the photons entering $\vec{S_{in}}$ and leaving $\vec{S_{in}}$ the particle:

$$\vec{F} = \frac{n}{c} \iint (\vec{S_{in}} - \vec{S_{in}}) \mathrm{d}A \tag{2.1}$$

where n is the refractive index, c is the speed of light, and dA is the element of area normal to $\vec{S}$. In this case, it is a scattering force pushing the particle along the beam axis. When the attractive and pushing forces are equilibrated, the particle is stably trapped.

For particles which are much larger than the light wavelength ($d \gg \lambda$), optical trapping can be intuitively explained by geometrical optics. When photons are reflected from the particle surface, it results in pushing the particle along the beam direction (Fig. 2.1a), i.e., forward from the laser focus. However, when photons are refracted by the particle located after the laser focus, the Snell's law implies (only in the case of the particle having higher refractive index than surrounding medium) that the divergent laser beam becomes more convergent. It increases the momentum components along the beam axis, and thus, leads to a force acting on the particle in the direction of the beam focus (Fig 2.1b).

In the case of our optical tweezers setup, neither of the regimes discussed above is, strictly speaking, applicable, since we operate with a laser wavelength of 1064 nm and use particles that are 2 μm in diameter. However, for the trapping of spherical objects in optical traps of standard shape, the geometrical optics framework remains helpful due to its simplicity. This approach suggests using tightly focused beams for effective trapping, because in this case the attractive forces will be high enough to balance the pushing forces. It can be achieved by using high numerical aperture (NA) microscope lenses to focus the laser beam.

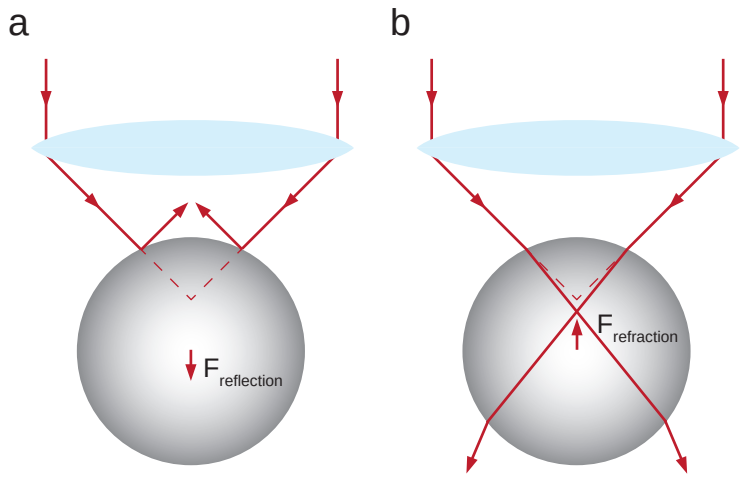


Figure 2.1: **Basic principles of optical trapping. a,** Force acting on a trapped particle when photons are reflected from its surface. **b,** Force acting on a trapped particle when photons are refracted by it. When these forces are equilibrated, an optical trap is produced.

The particles should have higher refractive index than trapping medium to be successfully trapped. We use polystyrene ($n = 1.55$) beads in water ($n = 1.33$).

One is able to precisely manipulate the trapped particle by moving the focus of the laser beam. Laser light scattering on the bead contains information the displacement of the bead from the trap center in the transverse plane. It can be read out by collecting the transmitted light by the means of condenser and imaging it on a position sensitive photodetector.

2.1.2. New optical tweezers design: motivation and reasoning

When one decides to work with optical tweezers, first step is obviously to decide where and how to get one. Since optical tweezers increasingly become more and more popular, some commercial solutions are available on the market. Companies like Thorlabs, Nikon Instruments, Zeiss, JPK, etc. offer their preassembled setups to the customers. Purchasing the tweezers saves a lot of time and allows biology labs, which are usually not proficient in laser optics, to get the setup and use it. However, most of commercial solutions are limited to basic functionality and lack certain flexibility. For example, if one wants to achieve non-conventional trapping regime or combine trapping with various imaging modes, it is still more straightforward to build the desired setup in house. Nevertheless, more and more advanced solutions become available on the market in recent time. Thus it is recommended to thoroughly check commercially available options before starting with building. When constructing the tweezers setup, it is possible to use a commercial microscope as a framework. It ensures the presence

of preassembled imaging pathway of good quality, although it implies spatial constraints in the system, which may prevent implementation of certain additional functionality. Furthermore, a combination of optical tweezers and commercial microscope on the same optical table often involves using tall upfront-standing components to direct the beams perpendicular to the optical table to enter the microscope system. These tall components are more prone to vibrations and thus may introduce additional noise into the signal. Thus, depending on the desired design, some labs refrain from using the microscope and construct their setups from the scratch on the optical table, usually remaining close to its surface.

In our lab, we had a single optical tweezers setup built without using a commercial microscope about ten years ago. Both an expansion of the group and an opportunity to have a more advanced setup motivated us to build a second one. Our old setup features a single optical trap that is capable of trapping one bead at a time, and the other bead is held on a micropipette by means of suction. A connection between the beads is formed and stretched by manipulating the micropipette, which is attached to the nanopositioning piezo stage. This construction is relatively simple. However, it results in noise from the optical table and piezo stage getting into the signal that leads to a decrease of its quality.

After careful consideration, in the new setup we decided to go for a design similar to the one published by Carlos Bustamante lab [2]. It features two optical traps generated from a single laser source that is splitted in two by polarization. This way, external noise in both traps is synchronized since it has the same source, and by detecting only differential fluctuations the resolution can be significantly enhanced [3], achieving subnanometer sensitivity. Another new feature planned in the new setup is the switchable confocal/widefield fluorescence imaging that, in combination with labeled chaperones, is expected to provide insight in the process of chaperones binding and unbinding, when at the same time their effect on protein client conformation is monitored by conventional optical tweezers functionality.

2.2. Experimental design: plan of the setup

Our experimental setup is a custom-made dual trap optical tweezers apparatus with differential detection designed similarly to the setup described by Bustamante et al. [2] Components (refer to Fig. 2.2 for a schematic) are positioned on an optical table (VH-3660W-OPT, Newport) that is carefully isolated from external vibrations. All optics and optomechanical components used to build the setup are listed in Extended Data Table 2.1.

2.2.1. Trapping laser, beam expansion and steering

For trapping, we use a 5 W ytterbium fiber laser (YLM-5-LP-SC, IPG Photonics) that generates a linearly polarized beam ($\lambda = 1064$ nm, 1.6 mm wide). The laser beam first passes a system of two half-wave plates (HW1, HW2) and a polarizing beam splitter cube (BS1) that controls the relative power of the two optical traps to be created. The first half-wave plate controls the total laser power that

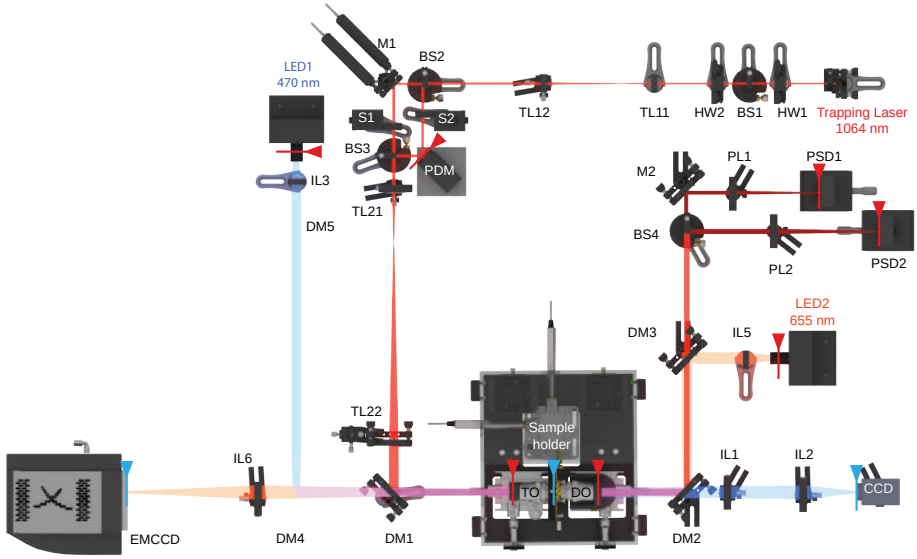


Figure 2.2: **The optical tweezers setup drawing.** The trapping laser beam is shown in red, two imaging beams are shown in blue and orange. Refer to Extended Data Table 2.1 for the components. The image is made by Mario Avellaneda.

enters the system by rotating the initial laser polarization and controlling the fraction of power that is directed into a beam dump by the beam splitter. The second half-wave plate controls the relative power of both traps, determined by the direction of the polarization vector of the laser beam prior to its splitting. Next, the laser beam gets expanded 2.5 times with a custom telescope (TL11, TL12) and split in two orthogonally polarized beams by a beam splitter cube (BS2). One of these beams can be finely laterally steered by moving a mirror attached to a two-axis piezo-driven mirror actuator (PDM; Nano-MTA2XHS, Mad City Labs), which allows for precise linear movement of the optical trap in the sample plane. From now on this beam and the corresponding optical trap will be reffered to as ”steerable”. The second beam can also be laterally steered, but less accurately, using a stepper motorized mirror mount (M1; KS1-ZST, Thorlabs), for the sake of alignment simplicity. This beam will be referred to as ”stationary”. Then the beams are recombined immediately by an identical beam splitter (BS3) and further expanded 3 times with a second telescope (TL21, TL22).

2.2.2. Optical traps formation, imaging and detection

Optical traps are formed with a water immersion microscope objective (TO; NA = 1.2, CFI Plan Apo VC 60x WI, Nikon) inside a glass flowcell mounted in a custom-made sample holder, which is accommodated on a three-axis precision translation stage (M-561-D-XYZ, Newport). The translation stage is operated

remotely using miniature DC servo motors (CONEX-TRA12CC, Newport). The flowcell is made of glass according to our design by Micronit Microfluidics (Enschede, Netherlands). Alternatively, one can fabricate a hand-made flowcell by sandwiching two layers of Nescofilm with cut channels between two coverslips (#2, Marienfeld). Refer to section 2.3 for more details on flowcells design and fabrication. The TO is also used to image the sample plane that is illuminated with blue light by LED1 ($\lambda = 470$ nm, Philips Lumileds) on CCD camera (XCG-V60E, Sony). An identical microscope objective (DO) collects the laser light that is further used for the detection and forms an additional (with another magnification) image of the sample plane using red LED2 ($\lambda = 655$ nm, Philips Lumileds) on EMCCD camera (iXon Ultra 897, Andor). Both objectives are attached to identical three-axis precision translation stages (M-561-D-XYZ-LH, Newport) but have different holders. The holder of the detection objective is equipped with a home-made magnetic mount to allow replacing it with high repeatability between experiments. Trapping and imaging beams are coupled and separated using dichroic mirrors (DM1, DM2). The back focal plane of the detection objective is imaged on individual position-sensitive detectors (DL100-7PCBA3, Pacific Silicon Sensor) for each trapping beam. It is achieved by re-separating the beams by polarization (using a beam splitter BS4) after the detection objective. Then the individual beams are imaged on the position-sensitive detectors (PSDs) using biconvex lenses (PL1, PL2). Individual detection paths for both laser beams allows for accurate position detection of trapped particles in both traps.

2.2.3. Software and data acquisition

High-frequency data acquisition and processing are performed using a LabView-programmed PXI FPGA board (NI PXI-7852R, National Instruments) according to Fig. 2.3. Four voltages from each PSD (V_x^{diff}, V_x^{sum}, V_y^{diff}, V_y^{sum}) determining a centroid position on the detector are analog filtered (8th order Butterworth filter) prior to being sent to analog input channels of the FPGA. Two analog outputs of the FPGA board are used to control the PDM position (V_x, V_y), with these values also used further in the software as a reference of the PDM position. Digital outputs of the FPGA control the state of the shutters and also send commands to pressure distribution system (see Section 2.3), while digital inputs of FPGA accept information about current pressure in reservoirs of the pressure distribution system. Next, the data from the FPGA gets digitized, undergoes digital filtering (this step is optional as analog filtering is usually used), further averaged if required, and then transferred to DMA stream of PC via PCI bus.

Other equipment communicates to the PC directly. The motors driving the sample holder translation stage and the mirror M1 are controlled via USB, the trapping laser - via RS-232 interface, and the CCD camera - via GigE interface. Most data processing and overall control of the setup are performed with custom-written C# program called Foldometer (Fig. 2.4). The measurement data contains a timestamp, the data from the FPGA (PDM and PSDs values), and the bead tracking data calculated based on the CCD image using the template location routine from National Instruments (IMAQ Match Pattern 3). The data can be saved with

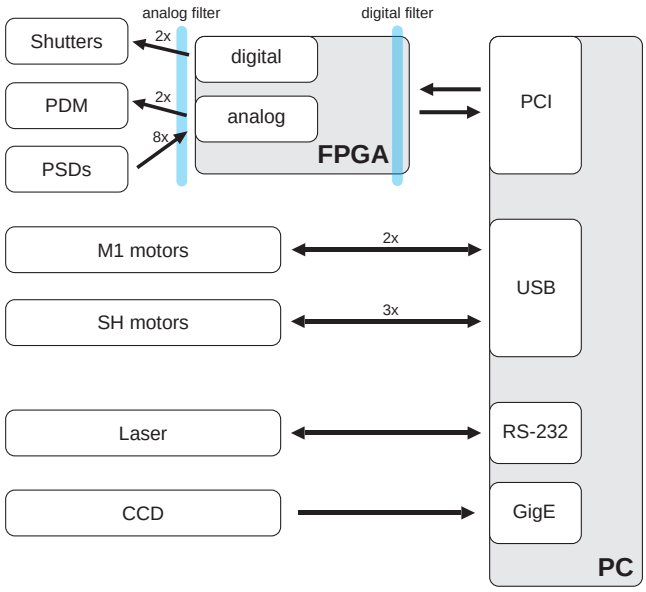


Figure 2.3: **Electronics scheme.** The signals are shown with arrows.

the frequency up to 100 kHz. The temporal resolution of bead tracking routine is technically limited to the CCD frame rate of 90 Hz.

2.2.4. Calibration and determination of force and extension

If a trapped microsphere remains close to its equilibrium position in the trap, the force response of the trap can be regarded as that of a Hookean spring, and the force exerted on the microsphere along the X axis can be written as $F_x = -k_x(x - x_0)$. For the other two coordinates, the same relation holds. Hence, in order to calculate the force acting on the particle, one should know its displacement from the trap center and the trap stiffness. If one wants to plot the force acting on the tether between two microparticles versus its extension, the distance between the beads should also be monitored.

We typically use the PSDs to monitor the positions of trapped beads in relation to the centers of the traps and the position sensor of the PDM to calculate the distance between two beads. Alternatively, the CCD image may also be used to monitor the distance between the beads with the lower frequency (90 Hz vs 100 kHz). To extract the coordinates of microspheres from the camera image, we used the aforementioned template recognition algorithm. To get the coordinates in units of length from these multiple sources, the readouts of PSDs and PDM and camera image should be calibrated prior to the measurement.

The trap stiffness is also determined from the data collected by PSDs. To

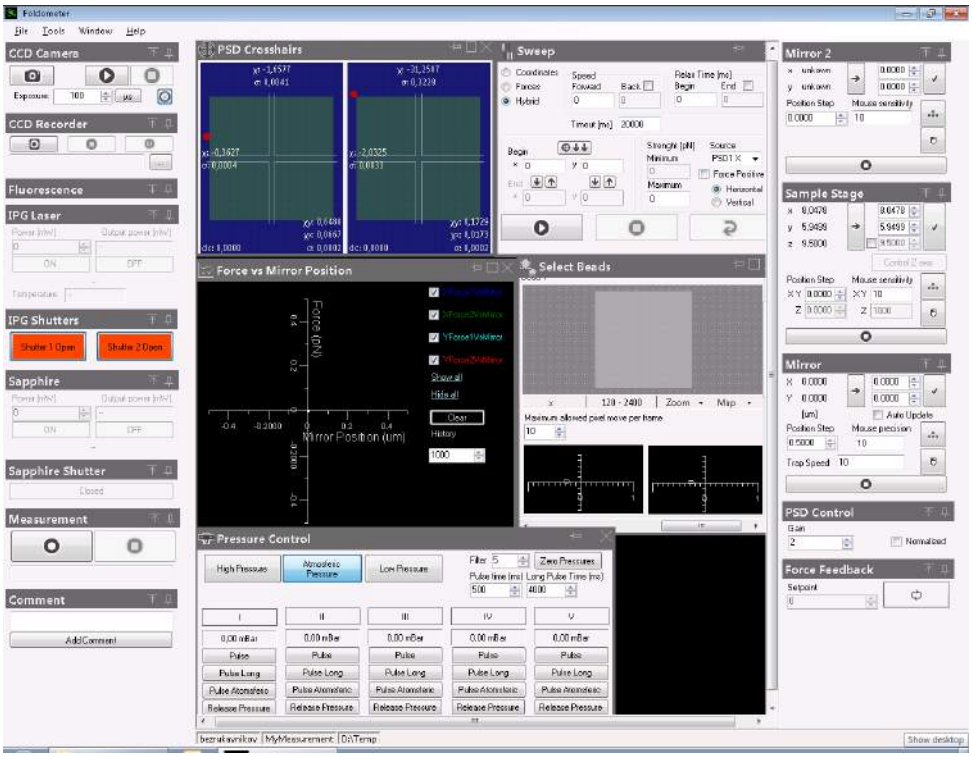


Figure 2.4: **Foldometer software screen.**

calculate it, several methods may be used [4]. We utilize the power spectrum method [5, 6], which employs Brownian fluctuations of a trapped bead to calibrate the PSDs and extract the trap stiffness.

Camera image calibration

The camera image is calibrated with a known distance reference such as stage micrometer (R1L3S2P, Thorlabs). The micrometer is mounted in the sample holder and its scale is imaged on the CCD. It yields a calibration factor $p = \frac{d}{N}$, where d is the known distance read from the micrometer and N is the number of pixels along this distance.

Mirror sensor calibration

The PDM is calibrated to obtain the relation between PDM position sensor output in arbitrary units and actual trap-to-trap distance (without the tether between two beads this value equals to bead-to-bead distance) in the sample plane in units of length. It is done by moving the steerable trap around and measuring its displacement in the sample plane using the calibrated camera image.

PSDs position and force calculation

We perform this calibration routine using the power spectrum method as described previously in Flyvbjerg et al. [5] The bead in a harmonic trapping potential fluctuates according to Langevin equation:

$$\dot{x}(t) + 2\pi f_c x(t) = \sqrt{2D}\xi(t) \tag{2.2}$$

where $f_c \equiv \frac{k}{2\pi\gamma}$, $\langle\xi(t)\rangle = 0$, $\langle\xi(t)\xi(t')\rangle = \sigma(t-t')$, $D = \frac{k_b T}{\gamma}$, f_c - corner frequency, γ - drag coefficient, D - diffusion coefficient. From Stokes' law drag coefficient $\gamma = 6\pi\eta R$, where η is the fluid viscosity and R is the bead radius.

Fourier transform of $x(t)$ and $\xi(t)$:

$$\tilde{x}_n = \int_{\frac{-T_{msr}}{2}}^{\frac{T_{msr}}{2}} \mathrm{e}^{2\pi f_n t} x(t)\mathrm{d}t \tag{2.3}$$

where $f_n = \frac{n}{T_{msr}}$, and T_{msr} is the measurement time. We put it into the equation 2.2:

$$\tilde{x}_n = \frac{\tilde{\xi}_n\sqrt{2D}}{2\pi(f_c - if_n)} \tag{2.4}$$

The power spectrum is defined as:

$$P_n^{ex} \equiv \frac{|\tilde{x}_n|^2}{T_{msr}} = \frac{\frac{D}{2\pi^2 T_{msr}}|\tilde{\xi}_n|^2}{f_c^2 + f_n^2}, n > 0 \tag{2.5}$$

Expected value of the power spectrum is a Lorentzian:

$$P_n \equiv \langle P_n^{ex}\rangle = \frac{D}{2\pi^2(f_c^2 + f_n^2)} \tag{2.6}$$

By fitting the Lorentzian to experimental data (Fig. 2.5a) we obtain the corner frequency f_c (measured in Hz) and the diffusion coefficient D (measured in $\frac{\mathrm{V}^2}{\mathrm{s}}$). The fit is done using the Matlab program published by Hansen et al. [7], which employs a Levenberg-Marquardt algorithm. In order to judge the quality of a fit, the program computes the reduced χ^2 value and the backing of the fit. By equating the fitted value of D to the theoretical $D = \frac{k_b T}{\gamma}$ we determine the conversion factor (so-called sensitivity of the PSD) from voltages to units of physical length: $D^{th} = \alpha^2 D^{ex}$, which means that $x = \alpha x^{volt}$. From $f_c \equiv \frac{k}{2\pi\gamma}$ we obtain the trap stiffness. The same routine holds for the Y axis.

Such calibration method is straightforward to implement and easy to use. It allows the calibration with same microparticles that are used for the experiments and can be done just before or just after the measurement. It provides good precision in determination of PSD sensitivity and trap stiffness ($< 1\%$). However, the accuracy of this method is limited, since one needs to know several experimental parameters, such as flowcell temperature, fluid viscosity and bead radius, to calculate required calibration factors. Temperature and viscosity may vary

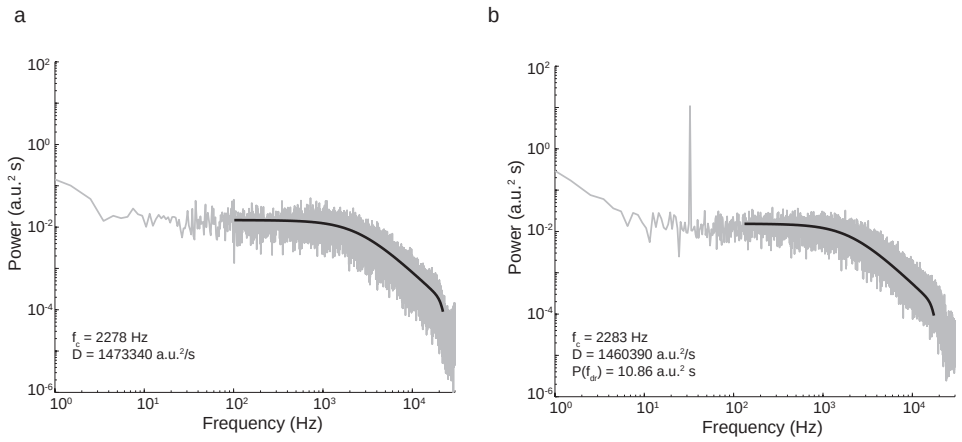


Figure 2.5: **Power spectra fitting. a,** The power spectrum of Brownian fluctuations of a bead in the stationary optical trap. The experimental data is shown in grey, the fitted Lorentzian - in black. Acquired corner frequency and diffusion coefficient are stated below. **b,** The power spectrum of Brownian fluctuations of a bead in the sinusoidally moving optical trap. A peak in the data corresponds to the driving frequency of the trap (32 Hz).

significantly during the experiment (for example, due to laser-induced sample heating), and bead radius may differ significantly between two trapped beads. One may try to limit the calibration inaccuracy by measuring the temperature inside the flowcell with a thermocouple and discarding the beads that are visually off the standard size. However, local heating effects and direct absorption of laser light by the thermocouple wire still do not allow for exact determination of temperature.

A different approach to this problem has been reported by Tolic-Nørelykke et al. [6] If the trap or the flowcell is moved sinusoidally while the position fluctuations of a trapped bead are recorded, the drag coefficient can be measured directly from the calibration. This routine can be applied to the steerable trap of our tweezers setup.

The trap is moved sinusoidally relative to the flow cell with a frequency f_{dr} and amplitude A. The position of the stage as a function of time is

$$x_{dr}(t) = A\sin(2\pi f_{dr}t) \tag{2.7}$$

Langevin equation in this case:

$$[\dot{x}(t) - V_{dr}(t)] + 2\pi f_c x(t) = \sqrt{2D}\xi(t) \tag{2.8}$$

Fourier transform of $x(t)$ and $\xi(t)$:

$$\hat{x}(f) = \int_{-\infty}^{\infty} \mathrm{e}^{i2\pi f t} x(t)\mathrm{d}t \tag{2.9}$$

Power spectral density (expected value) of the bead positions:

$$P(f) \equiv \frac{2\langle|\hat{x}(f)|^2\rangle}{T_{msr}} = P_T(f) + P_{resp}(f) \underset{T_{msr}\to\infty}{\longrightarrow} \frac{D}{\pi^2(f^2+f_c^2)} + \frac{\frac{1}{2}A^2}{1+\frac{f_c^2}{f_{dr}^2}}\sigma(f-f_{dr}) \tag{2.10}$$

The spectrum consists of a Lorentzian (first term) plus a delta-function spike (second term) at the frequency with which the trap was driven (Fig. ??b).

Experimentally, positions are measured in volts. Assuming linearity, $x(t) = \alpha x^{volt}$. We determine this calibration factor from the height of spike:

$$\alpha = \sqrt{\frac{W_{th}}{W_{ex}}} \tag{2.11}$$

where W_{ex} is the experimentally determined power of spike (measured in V^2) and W_{th} is the same quantity measured in m^2.

$$W_{th} = \int_0^{f_{Nyq}} P_{resp}(f)\mathrm{d}f = \frac{\frac{1}{2}A^2}{1+\frac{f_c^2}{f_{dr}^2}} \tag{2.12}$$

If $T_{msr} = \frac{N}{f_{dr}}$, (N - integer) then the spike consists of a single datum and consequently

$$W_{ex} = (P^{volt}(f_{dr}) - P_T^{volt}(f_{dr}))\delta f \tag{2.13}$$

where $P_T^{volt}(f_{dr})$ is the PSD of thermal background at f_{dr} and $\Delta f = \frac{1}{T_{msr}}$. To determine the value of $P_T^{volt}(f_{dr})$ we can use the theoretical expression of $P_T^{volt}(f)$ after it has been fitted to the experimental Lorentzian. To determine the trap stiffness k we use the definition of f_c and Einstein's relation:

$$k = \frac{2\pi f_c k_b T}{\alpha^2 D^{volt}} \tag{2.14}$$

We also calculate the experimental result for the drag coefficient:

$$\gamma = \frac{k_b T}{\alpha^2 D^{volt}} \tag{2.15}$$

It allows more precise force calculation; however, the temperature is still uncertain, which may result in a systematic error in the trap stiffness determination.

When the traps are calibrated, one could calculate the force acting on a biomolecule tethered between the beads and its extension (Fig. 2.6). Assuming a one-dimensional process for simplicity, consider the slow tether length change of ΔX. It can be measured as $\Delta X = \Delta(X_{PDM} - (X_{PSD1} - X_{PSD2}))$. As a result, the first bead experiences an additional force from the biomolecule equal to $k_{mol}\Delta L$, and the second bead - a force of $-k_{mol}\Delta L$. Since the beads are in equilibrium, the optical traps act on them with the equal oppositely directed force, which can

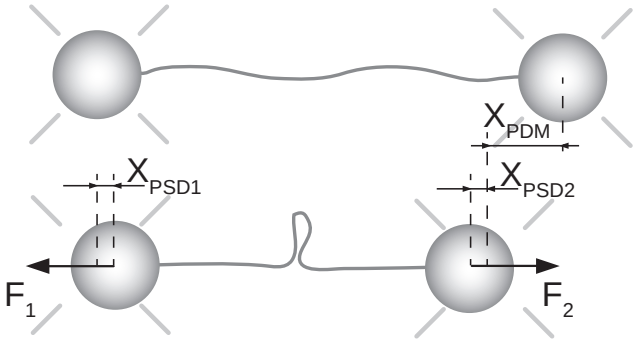


Figure 2.6: **Force and extension in optical tweezers.** The tether extension and the force acting on it are calculated from the position of the PDM and displacements of the beads in the traps.

be measured as $F_1 = -F_2 = k_{x1}\Delta X_{PSD1} = -k_{x2}\Delta X_{PSD2}$. Thus one can use either PSD1 or PSD2 to calculate the force acting on the tether.

As pointed out by Moffitt et al. [3], the signal-to-noise ratio of dual-trap optical tweezers is maximized if the deflections of both microspheres are monitored simultaneously and are used to construct an optimal coordinate. For traps of identical stiffness and identical beads, the optimal coordinate reduces to the difference coordinate: $X_- = X_{PSD1} - X_{PSD2}$, which we use to calculate the extension.

2.3. Microfluidics

For our single-molecule optical tweezers experiments robust and fast optical trapping of beads of desired type is essential. To achieve this, we developed and built a flow system that controls the buffer conditions and the flow speeds at different stages of an experiment. Optical trapping and manipulation occurs in the flowcell, which has a certain number of inlets and outlets. The inlets are connected to barrels containing different types of experimental buffers. The flow speed is controlled independently for each inlet by controlling the air pressure in the corresponding barrel with a home-built pressure distribution system. The pressure system features five reservoirs (according to the maximum number of inlets in our flowcells), connected to the barrels (Fig. 2.7). The pressure in these reservoirs is individually controlled by temporarily connecting them to either a compressor (P_{high}), a vacuum pump (P_{low}), or atmosphere via a raw-cut tube (P_{atm}). It is achieved using a system of solenoid valves as shown in Fig. 2.7.

2.3.1. Custom-made flowcell

In Fig. 2.8a, an example of a hand-made flowcell is shown. It is positioned in a sample holder that is used to place the flowcell on the sample stage and to connect the inlets of the flowcell with the outlets of the barrels. The flowcell is assembled from two identical coverslips (#2, 24x60 mm, Marienfeld), one of

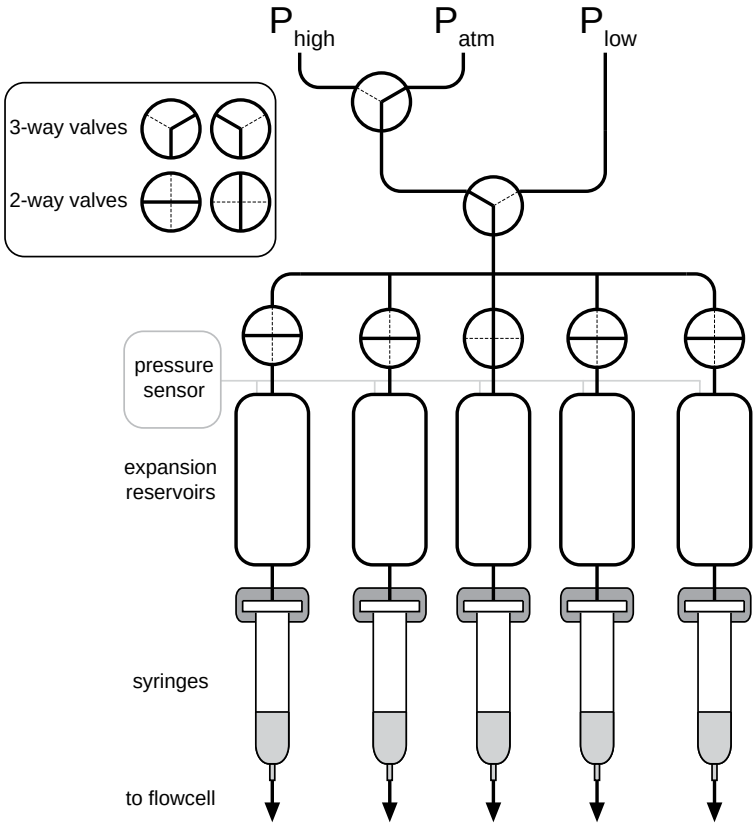


Figure 2.7: **Pressure system scheme.**

which has 10 holes of 1 mm diameter made by sandblasting, and two Nescofilm molds with channels cut in them (Fig. 2.8b-c). The molds were cut by hand or using a home-made shape puncher. The molds were sandwiched between two clean coverslips and positioned on a hotplate preheated to 250°C. The flowcell was heated while applying gentle pressure until Nescofilm melted and sealed the two glasses tightly.

Such a construction also allows for placing a micropipette inside the flowcell that may be useful for certain assays (for example, three-bead assays are quite useful for motor proteins studies). A bead from one of the traps can be transferred to a tip of the micropipette and held there by suction. Micropipettes are made by using a custom pipette-pulling apparatus. A glass fiber (OD 80 µm, ID 40 µm, King Precision Glass) is placed into the apparatus with its two ends fixed and its middle positioned inside a loop of metal wire. A weight is attached to the bottom end of the fiber. When a current is applied to the wire, it results in local heating and melting of the fiber, and its bottom part is pulled down by the weight. It results

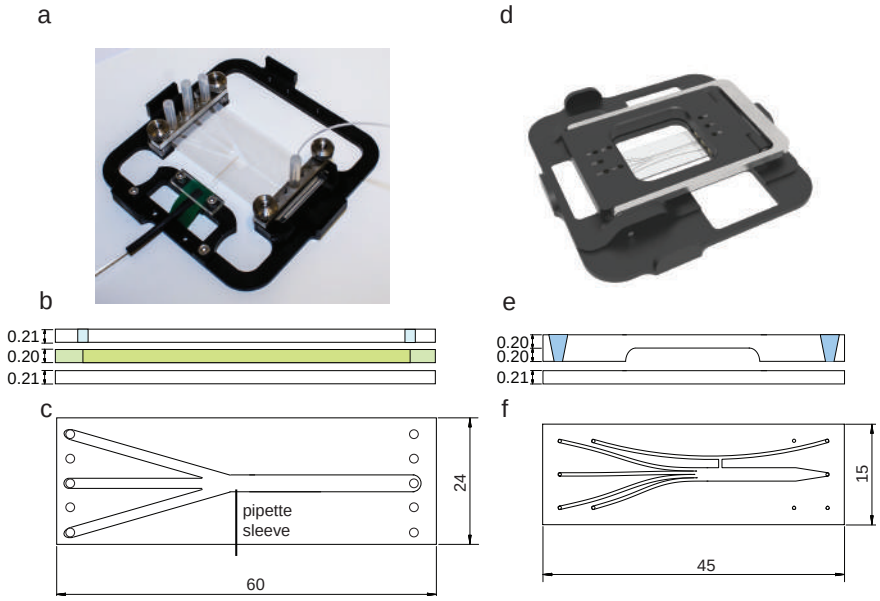


Figure 2.8: **Two types of flowcells. a,** Photograph of the hand-made flowcell in the holder. **b,** Cross-section of the hand-made flowcell. Nescofilm is shown in green, glass - in white, holes in the glass - in blue. **c,** Drawing of the hand-made flowcell. The dimensions are in millimeters. **d,** 3D model of the company-made flowcell. **e,** Cross-section of the company-made flowcell. **f,** Drawing of the company-made flowcell. The drawing and the 3D model of the company-made flowcell are made by Henk-Jan Boluijt.

in the formation of a strongly tapered micropipette with the tip size of about 1 um that is small enough to hold a bead. The pipette is connected to polyethylene tubing using heat-shrink tubing exposed to a hot air gun. The pipette is guided in the flowcell using a stereo microscope through a metal tube (OD 164 µm, ID 100 µm, World Precision Instruments) placed between two Nescofilm layers prior to the flowcell assembly. Such flowcells are cost-effective, although are quite laborious to assemble and leakage-prone. Since they are very thin, they also often break upon tightening of screws in the holder. Hand cut walls of the channels are relatively rugged that may result in flow irregularities and turbulences inside the flowcell.

2.3.2. Company-made reusable flowcell

We have overcome the disadvantages of custom flowcells by designing and ordering flowcells made exclusively of glass from Micronit Microfluidics (Enschede, NL). Fig 2.8d shows the glass flowcell in the sample holder, which was designed together with Micronit and made in AMOLF. The desired channel pattern is wet etched and 8 holes (5 inlets and 3 outlets) are powder blasted in a layer of glass (400 µm). Next, a second layer of glass (210 µm) is bonded to it, forming the

flowcell (Fig 2.8e). Such a construction is robust and almost leak-free. It can be reused after extensive cleaning, saves time previously spent on flowcell assembly, and does not break in the carefully designed holder where overtightening is prevented by the construction features. More importantly, the channels of such flowcell have much smoother walls that prevent flow perturbations and, in combination with the pressure distribution system, allows for precise temporal and spatial control over the chemical environment of the optical traps and the tethered molecule. Such an approach may be utilized to establish a certain trapping order of different types of beads, or in a setup that combines optical trapping with a fluorescence assay to reduce background fluorescence [8]. For the experiments presented in this thesis, the following channels pattern was etched in the flowcell (Fig. 2.8f): 4 inlets are merged in the middle of the flowcell into a joint channel that leads a central outlet, and a remaining inlet is connected to a corresponding outlet forming a channel parallel to a joint one. Central joint channel and a side channel are connected with a "bridge". Such pattern is inspired by a design described by Biebricher et al. [9] and is planned to be used for the future fluorescence studies (see subsection 2.5.1).

2.4. Step-by-step alignment manual

1. Define a Cartesian coordinate system. In this section, I will always refer to the system where XZ plane is parallel to the plane of the optical table with Z axis being parallel to the beam direction.

2. Check how the optical table is damped. Make sure it is floating properly, check the mechanical vibrations with accelerometer. Check whether the table surface is leveled (use a spirit level).

3. Mark preliminary positions of the components, check whether all distance constraints (mostly coming from the conditions for planes conjugation) are met. For complete list of the optical components and optomechanical equipment we used refer to Extended Data Table 2.1.

2.4.1. List of distance constraints in our setup

To expand a collimated laser beam by a factor of F, we construct a telescope by placing two plano-convex lenses of focal lengths f_1 and $f_2 = f_1 F$ in such a way that the distance between them equals $f_1 + f_2$.

The conjugate plane of a given plane, P, is a plane P' such that points on P are imaged at P′ [10]. In dual-trap optical tweezers design, as in a conventional optical microscope, two sets of conjugate focal planes are defined along the beam path:

- Beam steering and detection planes: PDM, back focal planes of TO and DO, PSD1, PSD2. Also, to achieve proper Kohler illumination, LED1 (2) planes should be conjugate to back focal plane of TO (DO).

- Imaging planes: CCD, sample plane.

See Fig. 2.2 for the positions of conjugate planes. Beam steering and detection planes (also LED plane) are marked in red, imaging planes – in blue. Derived from basic optics, here are the relations to calculate the distances between conjugate planes:

$$D_2 = \frac{f_{22}}{f_{21}}(f_{21} + f_{22} - \frac{f_{22}}{f_{21}}D_1) \tag{2.16}$$

where D_2 denotes the distance from TL22 to back focal plane of TO, f_{22} – focal distance of TL22, f_{21} – focal distance of TL21, D_1 – distance from PDM to TL21 or from M1 to TL21.

$$D_4 = \frac{f_{l1}D_3}{D_3 - f_{l1}} \tag{2.17}$$

where D_4 denotes the distance from PL1 to PSD1, and D_3 - the distance from PL1 to the back focal plane of DO (same logic applies to the position of PL2). It is convenient to place the components, whose position is critical, on translation stages to achieve more accurate conjugation. For our design, it is recommended to position telescope lenses (TL11-22) and PSDs1-2 on stages (or their combination) that allow 2-axis translation (XZ), and mirror M1 on a stage that allows 1-axis translation (Z). We used miniature dovetail stages (DT12XY/M, Thorlabs) to achieve it.

To locate the back focal plane of an objective, one may use the method described in Smith et al. [11]

2.4.2. Initial laser output tests

4. Mount the laser output collimator in the holder, clamp it stably. Choose the holder that allows angular adjustment of the collimator position (kinematic V-mount, KM100V/M, Thorlabs).

5. Clamp the laser fiber to prevent its bending (bending more than specified maximum angle may perturb the laser beam). We used a magnetic arm (2704101, Horex) with hand-made adapter for the fiber.

Perform the alignment procedures at low laser power (< 100 mW) to prevent optical damage to the components (and to the developers), unless otherwise specified.

It is highly recommended to test the laser properties at the initial phase of setup assembly, since exceptional laser characteristics are key to building stable and sensitive optical tweezers. Even if the laser is purchased from trusted manufacturer, independent testing is necessary to reveal potential problems, since the performance may vary between individual laser systems, and damage to the laser may occur after its shipping.

6. First, test its output power stability. Set the laser to the power it will be operated at, direct the beam into the power meter (Nova II, Ophir Photonics), and monitor the power fluctuations overnight. Our IPG laser displayed the power fluctuations within 1%, according to the manufacturer's specs.

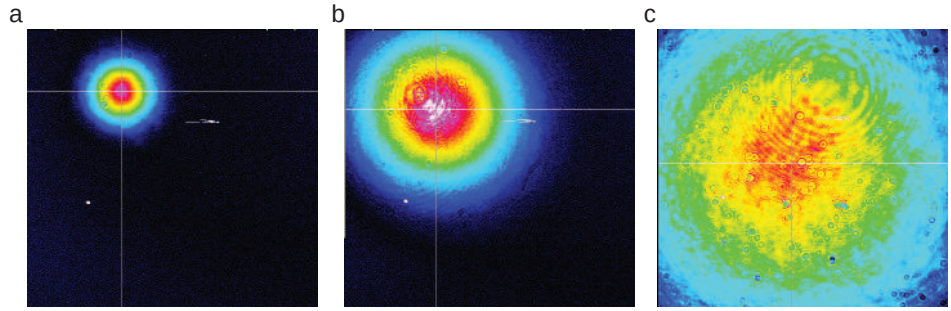


Figure 2.9: **Trapping beam profiles. a,** The beam imaged with the beam profiler as it is emitted from the laser output collimator. **b,** The beam after the first telescope. **c,** The beam after the second telescope before entering the TO.

Note that the laser power stability may vary depending on the power. Tend to operate the laser at the power that is recommended by manufacturer, since a laser like ours typically runs most stably at powers well above its threshold.

7. Second, test the laser beam shape. If the laser wavelength does not belong to the visible spectrum (that is the case for the majority of lasers used for optical trapping), one needs to use specialized sensors to visualize the beam. Most basic one is the infrared sensor card. We use the cards from various vendors (i.e. F-IRC4, Newport, and VRC4, Thorlabs), but none of them displays the beam shape optimally for our application. They are either not sensitive enough to display lower power beams of mm diameter, or have good sensitivity but display beams with incorrect shape. Thus to measure the laser beam intensity profile one needs more advanced sensors, that is a commercial beam profiler. We use the WinCamD from Dataray. One may also use just a conventional CCD camera that has good sensitivity in near-infrared. However, a great care must be taken to avoid the damage to the camera sensor. Always protect the sensor with strong neutral density filters and display focused beams only if you are absolutely sure that the laser power is very low.

8. Image the laser beam with the beam profiler at several distances from the output collimator (Fig. 2.9). Compare the beam overall shape and width with the specifications, make sure that the beam is properly collimated by comparing the beam images from different distances. A more precise way to test the trapping beam for the collimation is to use a shear-plate interferometer (SI, Thorlabs). Make sure that the plate used fits to the dimensions of the beam to be tested (we used plates suitable for the beams of 1-3 mm and of 2.5-5 mm, SI035P and SI050P, Thorlabs). Since the trapping beam is not in the visible domain, the interference pattern should be imaged on the CCD camera. If fringes in the pattern are perfectly horizontal, it

indicates that the beam directed into the interferometer is collimated (Fig. 2.10).

9. Third, test the beam pointing stability. It can be done by utilizing a position-sensitive photodetector (PSD) to track the laser spot over time.

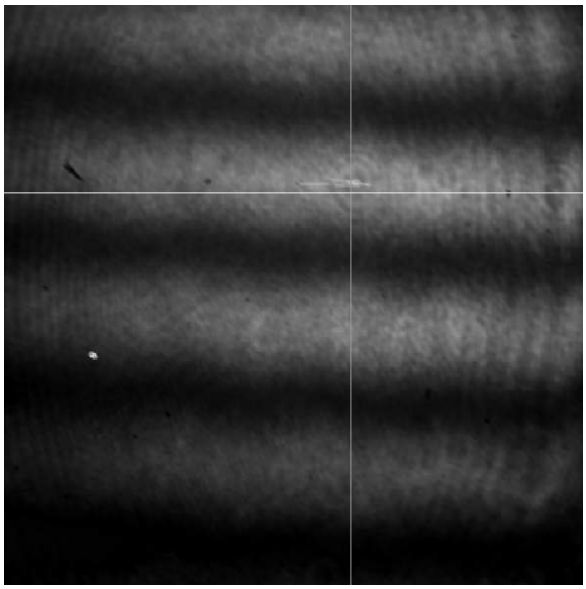


Figure 2.10: **The interference pattern of the trapping beam observed on a shear-plate interferometer.** It indicates that the trapping beam is collimated.

2.4.3. Setting up the trapping laser path

When one is done with testing the laser beam properties, compared it with the specs provided by the manufacturer, and is satisfied with the results, one can proceed with the alignment.

10. Adjust the output collimator holder position to align the beam along a row of holes on the optical table for further convenience.

11. Adjust the beam angle by turning the knobs on the holder to make it parallel to the optical table surface. Always keep the beam at constant height. It simplifies the alignment process greatly. To control the angle precisely, use the following routine (also further in the alignment procedure). Position a straight reference (a long ruler works satisfactory for this purpose) along a row of holes. Two screws can be used to align the ruler. Place the beam profiler adjacent to the ruler at the fixed height to image the beam. Slide the profiler along the ruler without changing its height, covering a substantial distance (~1 m). If the beam does not change its position on the sensor,

it is sufficiently parallel to the table surface and to the selected row of holes. Some beam profilers have a built-in centroid determination routine that can be employed to track the beam center more precisely.

12. If neither the beam profiler nor CCD camera are available, one can use an alternative (more laborious) routine to adjust the beam angle. Position two irises with their centers at desired beam height on the optical path, at substantial distance from each other. Center the beam on the first iris. Close both irises, leaving the smallest possible aperture. Check whether the beam passes through the second iris. If not, adjust the beam angle. Iterate until the beam passes through both irises, manifesting that the beam is leveled.

13. To have a convenient reference for revealing a cause of possible future misalignment, place pairs of irises along every straight part of optical path. Thus, if misalignment occurs, first iris that blocks the beam when all irises are closed pinpoints the component responsible for misalignment.

14. Place the mirror M1 at the specified position. Align the beam along a row of holes on the table. Position the dichroic mirror DM1, adjust the beam angle as described previously.

15. Place the objectives mounts (we use 3-axis precision translation stages M-561D-XYZ-LH from Newport with custom-made objective brackets) and sample mount (M-561D-XYZ, Newport with custom-made sample holder) on the table. Do not place the objectives in for now. Place the pellicle beamsplitter (BP145B3, Thorlabs) inside the objective bracket. Center an infrared alignment disk with a through hole (VRC2D1, Thorlabs) along the optical path before the objective bracket. Make sure that a beam reflected back by the pellicle overlaps perfectly with the incident beam. It indicates that the bracket back plane (as well as a back focal plane of the objective) is perpendicular to the beam.

16. Place a threaded infrared alignment disk (VRC2SM1, Thorlabs) in the trapping objective holder. Overlap a target on the disk with the center of the incident beam by adjusting holder position. Repeat for the detection objective holder.

17. Position a dichroic mirror DM2 and a mirror M2. Adjust the beam angle appropriately. Mark the current beam position on the wall. It provides one with a useful reference point for the further alignment.

2.4.4. Alignment of a bright-field microscope

This stage involves a construction of a bright-field microscope and testing its ability to produce a stable high-quality image on CCD camera. The alignment of the additional imaging path (using LED2 and EMCCD) can be performed accordingly.

18. Coarsely position the components involved in the bright-field beam path (LED1, IL1-3, DM4, BF, CCD) according to the drawing. Align LED beam along a row of holes in the optical table.

19. Place both objectives in their holders. Be careful to prevent any direct contact between front elements of objectives. It may cause damage of lenses.

20. Adjust lateral position of detection objective (DO) to make sure that its principal axis overlaps with the one of trapping objective (TO). Do this with the LED off. DO is aligned when the z-translation of its stage produces symmetric changes in the shape of the TO front face seen through the back aperture of DO (by eye).

21. Adjust the LED and IL4 such that the beam passes through the center of DO back aperture. The beam path can be checked using plain paper.

22. Adjust the LED and/or IL4 z-position to focus the LED surface image approximately at the back aperture of DO. It may be easier to do if either LED or IL4 are mounted on a 1-axis translation stage (such as MT1/M, Thorlabs).

23. Apply a layer of an immersion fluid, which has the same refractive index as water ($n = 1.33$), between objectives. It becomes critical to use an immersion fluid that is viscous enough since the principal axis of the objectives is horizontal. Otherwise the fluid may drip down that leads to a loss of layer integrity. We use Immersol W 2010 from Carl Zeiss Microscopy.

24. Check the bright-field beam path after DO. Adjust the position of DM4 to center it on the beam and change its angle to keep the reflected beam parallel to the table.

25. Center IL3 on the beam. Position the CCD at the focal point of IL3 (at 200 mm distance in our case) and center the bright spot on the camera sensor.

26. Change the axial position of DO and TO. Make sure that the bright-field beam expands and shrinks symmetrically with no movement of the center.

27. Mount the flowcell with a micropipette. Visualize the micropipette on the CCD camera by adjusting the axial position of both objectives. Displace both objectives slightly in lateral direction to achieve uniform and symmetric micropipette imaging. Adjust the CCD exposure time and LED power to avoid image saturation. Alternatively one may use beads adsorbed to a coverslip for a reference image. Check the image stability to make sure that there is no systematic drift of optical components that distorts the image over time.

2.4.5. Alignment of a single trap

An outline of the trapping beam path is constructed already, and at this stage the trapping beam should be modified and directed in a way to form the stable optical trap, and the corresponding detection beam path should be set up.

28. Place a half-wave plate HW1 in a precision rotation mount. Place a polarizing beam splitter BS1 in a prism mount that allows its angular adjustment (we use Ultima Gimbal Prism Mounts UGP-1 from Newport). Use such mounts for all beam splitters, since their angular orientation is critical for precise beams steering and needs to be adjusted finely. Make sure that the transmitted beam position on the irises did not change. Direct the laser beam reflected by BS1 into a beam dump.

29. Since our trapping laser beam is linearly polarized, its polarization vector rotates upon passing through a half-wave plate. Find and note the orientation of HW1 that corresponds to a maximum transmission by BS1. It may be further changed to get rid of an excess of laser power, if the laser stability dictates to operate the laser at powers higher than required for optical trapping.

30. Position TL11 on the table, center it on the beam. Adjust its position to re-center the beam on the irises again.

31. Position TL12 on the table, center it on the beam. Make sure that the beam position on the irises that are further on the beam path did not change.

32. Image the beam with the beam profiler close after TL12, and far away from it. Compare the beam width with expected value (~3.5 mm at our setup, Fig. 2.9b), test its collimation with the shear-plate interferometer. If the beam is not collimated, fine-tune the distance between lenses until the beam is collimated.

33. Place TL22 at the specified distance from TO with DO removed. Adjust lateral position of the lens to direct the beam to the mark on the wall, and adjust its distance to the TO to minimize the transmitted beam size. It indicates that front focus of TL22 and back focus of TO overlap, which is one of the conditions for planes conjugation.

34. Remove TO. Place TL21 according to the $4f$ arrangement. Adjust its lateral position to center the beam on irises I3-6, and on the wall mark. Translate TL21 along the beam direction to collimate the beam. Test the collimation with the interferometer. Compare resulting beam width with the expected value. We expanded our trapping beam to the final width of 11 mm (Fig. 2.9c). This way it slightly overfills the back aperture of TO (~9 mm), that is essential for generation of strong optical trap, since peripheral rays contribute the most to the pulling force.

35. Check that $4f$ arrangement that involves M1, PDM, TL21, TL22, and TO is still met (a piezo-driven mirror PDM is not mounted yet, use its expected position). If there is an error, adjust the position of PDM and M1. If angular adjustment of M1 is required, re-center the beam on previously placed irises.

36. Mount both TO and DO. Apply a layer of immersion fluid between them. Adjust the lateral position of DO to overlap the transmitted beam with the wall mark, and its axial position to collimate the beam.

37. Place a small drop of immersion fluid on TO. Mount the flowcell in the holder. Slowly bring it in contact with the immersion fluid. Adjust the flowcell position slightly along the beam direction to focus the beam on the glass-immersion interface. A back-reflected light from this interface will be observed on the CCD as so-called airy disks, which are concentric rings formed around focused beam. Note that some components, such as, for example, bandpass filters that one may have between TO and CCD can block trapping laser light. Make sure that the back reflected light reaches CCD. A Gaussian beam of proper width that is well aligned into TO should provide clean and symmetric airy disk pattern that expands and contracts symmetrically when the beam is focused or defocused on the interface (Fig. 2.11). Any defects or asymmetries in the pattern manifest the presence of optical aberrations in the beam. See the protocol by Lee et al. [12] for detailed guidance on expected types of aberrations and their corresponding airy patterns.

38. Change the lateral position of CCD to center the airy pattern in the field of view. Adjust the angles of M1 and DM1 to make the airy pattern more symmetric as it contracts and expands.

39. Iterate the step 38 until the airy pattern is centered and circularly symmetric.

40. Prepare the beads solution (~0.001%) for testing the optical trap and infuse it into the flowcell. Check the bright-field image of the flowcell, repeat the steps 20-27 if the illumination is uneven or distorted.

41. Focus the laser beam in the middle of the flowcell, at approximately equal distance from both glass walls, by adjusting the flowcell axial position. Observing the airy disks at glass-immersion interfaces helps in locating the flowcell center.

42. Set the laser power to about 50 mW in the focus. Move the flowcell around laterally or open a slow flow and try trapping a bead. The bead should stay localized stably in the beam focus.

43. If a bead is trapped, but its image is defocused, it indicates that the imaging plane and the trapping plane do not overlap. Make sure that the imaging plane is positioned in the middle of the flowcell by obtaining a sharp and focused image of an object inside the flowcell (such as the micropipette). It can be achieved by adjusting the positions of CCD and L4. If the planes still do not overlap, adjust the positions of TL11 and TL12 along the beam to fine-tune the trapping plane position. Do not change the positions of TL21

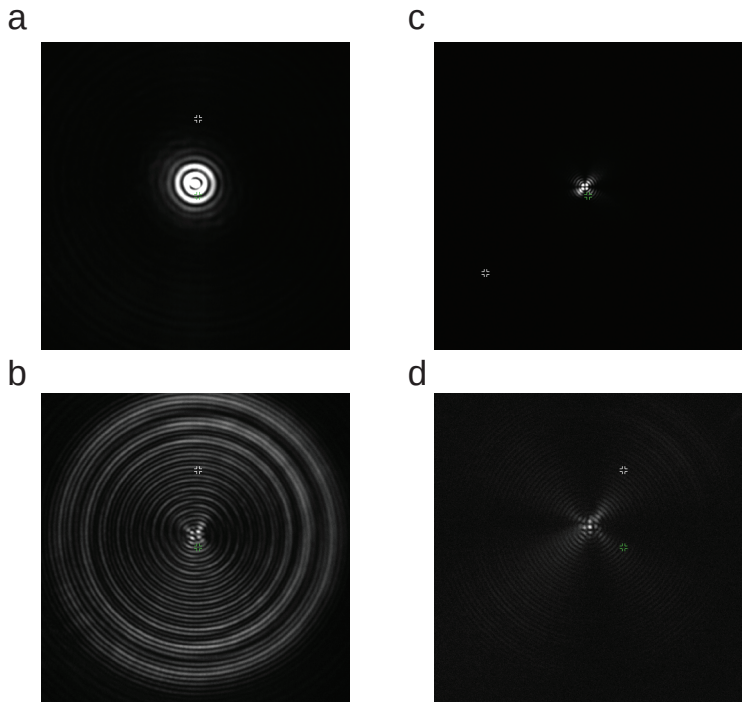


Figure 2.11: **Airy disk patterns by the trapping beam observed on the glass-water interface. a-b,** Compacted and expanded patterns of the stationary trap. **c-d,** Compacted and expanded patterns of the steerable trap. Different intensities of these images arise from polarization-dependent transmission of optical components between the TO and EMCCD and do not manifest a difference in powers of the optical traps.

and TL22 since it will affect the conjugation of steerable mirrors and back focal plane of TO required for the effective beam steering.

44. If beads cannot be trapped, it may happen due to one of the following reasons: incorrect beam shape and power loss due to the clipping along the optical path, underfilled TO aperture, wrong laser power settings (either too low or too high), aberrations in the trapping beam due to misalignment, or not perfect collimation of the trapping beam before entering the TO. Check for the presence of these problems and eliminate them to improve the trapping.

45. Trapping may also be significantly affected by the beads concentration in the flowcell. If there are too many of them, it may result in trapping of multiple beads at the same time, or a trapped bead may be quickly kicked out of the trap by freely diffusing ones. However, if there are too few, trapping may also get less efficient.

46. As a quick test of the trap strength, move the flowcell in lateral and axial

directions with set speed. One may calculate the critical force at which bead escapes the trap by estimating the drag force acting on the bead due to its movement in relation to the surrounding fluid.

47. Rotate M1 to move the trap around. The bead should stay in focus at all positions. If it does not, it may be an indication of imperfections in conjugation leading to the movement of the laser spot on the TO back aperture. In this case, one can measure the involved critical distances as accurately as possible and try to adjust the distances between optical elements correspondingly.

48. If the optical trap is capable to hold beads laterally and axially, the detection system can be aligned. Place PL1 at the specified position and center it on the beam. Make sure to attenuate the laser power with a neutral density filter NDF1 before placing the PSD (our PSDs have optical damage threshold of 1.5 W/cm^2 that can be easily exceeded if the laser beam is focused on the PSD).

49. Position the beam profiler in place of the PSD. Visualize the beam with the trapped bead. The pattern on the profiler is an image of the back focal plane of DO. Lateral displacements of a trapped bead from the trap center result in changes of the pattern. They can be captured by tracking its image with the PSD.

50. Position the PSD1. Check whether the laser spot falls on the sensor and can be detected. Rotate M1 and make sure that the laser spot on the PSD1 does not move. If it moves, translate the PSD axially to minimize this movement.

51. To calibrate the trap, sample the Brownian fluctuations of a trapped bead. Analyze their power spectrum to obtain a corner frequency that can be used to determine trap stiffness (see subsection 2.2.4). One can use a MATLAB program written by Hansen et al. [7] to analyze the power spectrum. To give a reader a feeling of calibration results, we calibrated the optical trap at 2 W output laser power that corresponds to 250 mW per trap in the laser focus. We trapped polystyrene beads of 2.1 μm in diameter. The calibration yielded a corner frequency of 2278 Hz in x direction and 2503 Hz in y direction. Consequently, the corresponding trap stiffnesses are $k_x = 283$ $^{pN}/_{\mu m}$ and $k_y = 311$ $^{pN}/_{\mu m}$. If the difference between the values is significant, or if the power spectrum is largely distorted, it may be either due to imperfections in the beam shape or due to misalignment in the optical path. The power spectrum also yields a calibration factor converting PSD readout to the physical displacement of the bead in the trap.

52. Analyze the power spectrum to pinpoint noise sources in the system. Try to reduce the noise according to the subsection 2.4.10.

2.4.6. Alignment of the steerable trap

To align a dual-trap optical tweezers setup well and easy, one should use the previously aligned single trap as a reference. Creation of a second optical trap involves splitting the trapping beam by polarization, and using two orthogonally polarized beams for independent trapping and detection. One should reduce the differential optical path for two beams: overlap them immediately after steering and split again just before the detection.

53. Remove TO and DO. Position HW2 on the table and center it on the beam.
54. Position BS2 on the table and center it on the beam. Make sure that the transmitted beam did not move at following irises.
55. The relative power of two traps depends on the angular orientation of HW2. Rotate HW2 and measure the power of transmitted (stationary) and reflected (steerable) beams after BS2. Achieve equal power; note the current position of HW2.
56. Adjust the angle of BS2 to direct the steerable beam parallel to the stationary one. Check their mutual orientation at a long distance.
57. Position PDM and BS3 on the table. Align PDM and BS3 position such that both beams coarsely overlap at I3-6. Achieve a complete overlap by iterative adjustments of angular positions of BS2 and BS3.
58. Mount individual motorized shutters S1 and S2 for both traps.
59. Mount TO, flowcell, and DO. Image the airy patterns coming from both traps. Note that due to different transmission of optical components between TO and CCD for different polarizations the airy patterns coming from different traps may have different intensities. It does not necessarily mean misalignment of one of the traps.
60. If the airy pattern of the steerable trap is not in the field of view, slightly adjust either BS2 or BS3 to overlap it with the pattern of stationary trap.
61. Trap beads in both traps. Test their strength, stability, and location according to the section 2.4.5. Try to bring two beads close to each other and touch them together to test for any detrimental position-dependent effects. One or both beads may escape the traps when brought close due to interference between beams that deteriorates traps quality. Such interference may be typically minimized by careful alignment.
62. Mount BS4, PL2, and PSD2. Set up the detection path for the steerable trap as it has been done for the stationary trap in the previous section. Make sure that the angular movement of PDM does not result in the movement of a laser spot on PSD2.
63. Record the Brownian fluctuations of beads in both traps, analyze the power spectra and calibrate the traps.

64. Calibrate the PDM sensor and the CCD image as described in the subsection 2.2.4.

2.4.7. Crosstalk elimination

In our setup we split single trapping beam in two using polarization optics, thus they are separated well until their polarizations are not affected. However, high-NA microscope objectives change the polarization of transmitted laser beam, leading to a so-called crosstalk in detection system. It can be observed as a parasitic signal in one of the traps depending on the position of the other. A method of spatial filtering described by Ott et al. [13] can be used to eliminate the detection crosstalk. A pinhole can be placed in front of the stationary trap-associated PSD in the plane that is conjugate to the trapping plane to transmit only stationary beam and cut the parasitic signal from the steerable beam.

2.4.8. Measurement regimes

Force ramp

Our typical optical tweezers measurement is conducted in a "force ramp" regime. In this regime, the distance between the beads increases monotonously in the stretching phase, and decreases monotonously in the relaxation phase. Unless otherwise specified, the standard stretching and relaxation speed is 50 $^{nm}/_{s}$. When the tether is fully extended (its contour length is roughly equal to its extension), the constant speed of traps separation results constant loading rate on the tether. For DNA tether, forces > 5 pN are enough to fully extend it. In force ramp protein unfolding experiments, stretched protein is out of equilibrium conditions. In order to extract the parameters of the unfolding reaction without applied force from the force spectroscopy data, the Bell model [14] or the adaptation of Kramers' diffusion model [15] is often used. Dynamic force spectroscopy measurements also allow obtaining certain structural information about the stretched protein structure. In this technique, protein unfolding experiments are performed with varied loading rate, and mean unfolding force is plotted versus the loading rate. Then the data is fitted to the Bell equation, and the parameters such as the transition state coordinate and the unfolding rate at zero force are extracted. Dependence of the transition state coordinate on the loading rate is employed to discriminate, for example, between the state having stable tertiary structure and the molten globule state, which is more sensitive to the applied force [16].

Force clamp

In addition to other advantages of force spectroscopy, it also offers an opportunity to explore different regions of protein energy landscape. Applied force lowers the transition barriers between states and the energies of non-native conformations, allowing the protein to visit these states more often. To follow the transitions between these states in equilibrium, force clamp measurement regime is often used. In this regime, the force acting on the tether is kept constant, and spontaneous fluctuations between different states (depending on the applied force)

are monitored to extract, for example, folding reaction rates. Since the applied force is extension-dependent, feedback systems are used to compensate for the change of the extension occuring upon unfolding/refolding. In our setup, a PID force feedback loop is utilized to control the PDM position in response to the tether extension change.

2.4.9. Power estimates

Measure the laser power at the exit of the laser collimator and before the TO. The ratio shows the power loss along the optical path. Adjust an iris aperture to the size that corresponds to the diameter of the TO back aperture (~1 cm for our objective) and measure the transmitted laser power. It provides an estimate of transmission loss due to clipping on back aperture of TO. We measured the transmission of 79% that is in line with geometrical considerations.

Mount TO and DO. Measure the transmitted laser power after DO. Since the objectives are identical, the square root of power loss after both objectives equals to the power loss of a single objective. For our objectives, we estimated a transmission of ~44% at 1064 nm wavelength. Make sure that the laser power in the sample plane is enough to trap beads. We have 250 mW per trap that corresponds to initial (emanating from the laser output collimator) power of 2 W. If the power is too high, it may result in the loss of trapping stability due to the domination of light pressure forces over refraction forces. If it is required to focus the laser beam on any optical element or detector, be absolutely sure that the laser power density does not exceed the damage threshold of the equipment.

2.4.10. Noise considerations

To achieve sufficient spatial resolution in optical tweezers, it is necessary to carefully isolate the setup from all sources of environmental noise. Air flows in the experimental room induce excessive laser beam pointing fluctuations and should be avoided. First, the room should be temperature-controlled to minimize temperature gradients. Second, any sources of temperature instability and air flow, such as electronics containing fans and even a setup operator should be ideally located in another room. If such radical measures are not possible, it is recommended to cover the setup to screen the beam path from air currents. Covering also helps to keep the setup dust-free and less prone to accidental misalignment. Temperature gradients do not only result in air convection, but also introduce differential drift of the optical components due to temperature-dependent material expansion. Optical tweezers are very sensitive to the drift as it degrades the resolution and results in misalignment of the setup. It highlights the importance of accurate control over the temperature in the experimental room.

It is necessary to locate optical tweezers on a properly damped optical table to isolate the setup from external mechanical fluctuations. Make sure that fluctuations are not transferred into the system via a rigid piece of equipment touching both a frame of the table and the breadboard. Furthermore, it is recommended to choose a room with additionally stabilized foundation. Level of mechanical

fluctuations in the room and on the table can be measured using an accelerometer to test whether the room is suitable for sensitive experiments and whether the setup is well isolated. To assess how the different sources of noise affect the measured signal and pinpoint most detrimental ones, one can measure the power spectra of signals coming from different pieces of equipment (Fig. 2.12). More detailed guidelines for the noise minimization can be found in the protocol by Bustamante et al. [17]

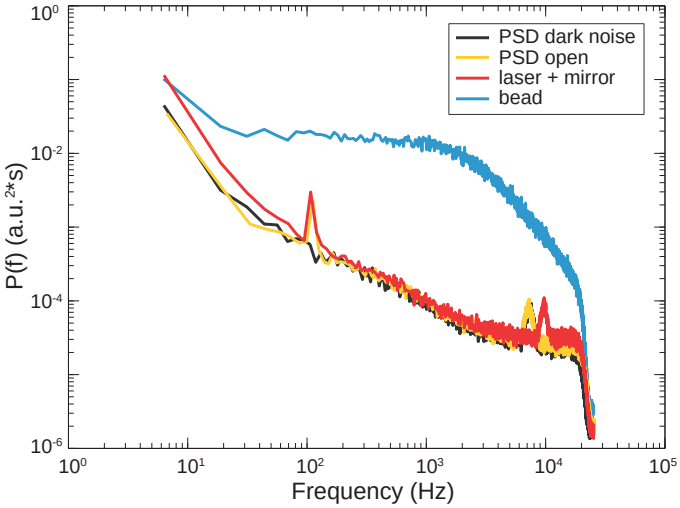


Figure 2.12: **Power spectra of signal fluctuations from various setup components.** "PSD dark noise" signal is recorded with closed PSD, "PSD open" – with PSD open to the ambient light, but with laser off, "laser + mirror" – with the signal from the steerable trap imaged on the PSD, "bead" – when a bead was trapped in the steerable trap.

2.5. Outlook

2.5.1. Chaperone-assisted protein folding: see a chaperone acting

In the protein folding experiments presented in this thesis, a chaperone bound to a protein client cannot be visualized directly. The presence of the chaperone around the client is guaranteed by buffer exchange, and the interaction between them is demonstrated by stretching experiments. However, a range of questions remains not addressed in such assay. For example, what happens to the

chaperone when its effect on the client is not observed anymore: does the chaperone leave the client or stays bound but in an inactive state? How does applied force affect chaperone binding? Does the bound chaperone allow native protein refolding? How many chaperone molecules are bound to the client at different conformational states? Where are they bound?

An ability to observe chaperone molecules interacting with the client, together with single molecule force spectroscopy, can provide enough insight to resolve many of these questions. A combination of optical tweezers with fluorescence microscopy has been applied successfully to study DNA-protein interaction and DNA mechanics [18, 19] and is thus potentially useful for chaperone-assisted protein folding experiments. Various fluorescent imaging techniques have been implemented in previous studies, such as widefield and confocal fluorescence, FRET, TIRF, and super-resolution microscopy (STED) [8, 20–23]. A general requirement for these solutions is a careful spatial [19, 24] or temporal [23] separation of trapping and imaging regions that imposes certain limitations on assay geometry. Otherwise the trapping beam may affect fluorescence detection due to its much higher intensity than fluorescence emission, and also enhance fluorophore bleaching. Widefield fluorescence microscopy and ensemble FRET are relatively simple to implement, but typically suffer from background noise, which needs to be especially strongly reduced to detect weaker signal in single molecule assays. It can be achieved by extensive washing of unbound fluorophores [8]. Alternatively, the background noise can be lowered by using localized emission techniques such as TIRF or confocal [21, 25].

We have planned an upgrade of new tweezers setup with switchable widefield and confocal modes of fluorescent imaging combined with optical trapping (Fig. 2.13). The excitation light is emitted from a laser of 532 nm (Coherence Compass 315M). Many fluorophores are compatible with this wavelength, such as Alexa Fluor 532 or Cyanine-5. Their emission is imaged on the EMCCD camera (iXon Ultra 897, Andor). The collimated laser source is guided with the help of two mirrors (M3 and M4) to the first objective after being expanded 10 times (beam expander). A 2D galvomirror (GVS202, Thorlabs) is added to allow scanning of the excitation beam if required. The optical system following this galvomirror is flexibly designed to allow three different kinds of fluorescence microscopy. It consists of two lenses (TL31, $f = 150$ mm and TL32, $f = 450$ mm respectively) forming a x3 Keplerian telescope, and two dichroic mirrors (DM4 and DM5). The first lens can be mechanically removed by means of a flip mount adapter (FM90, Thorlabs). DM5 is a shortpass dichroic mirror at 500 nm (NC440545, Chroma), letting the blue LED illumination enter the optical path. At the position of DM5, two dichroic mirrors can be set, leading to two different methods of fluorescence detection, as explained below. Besides, this mirror can be translated along the optical axis with a manual stage (DT12/M, Thorlabs). Depending on the arrangement of only TL31 and DM4, one can switch between the widefield and confocal imaging.

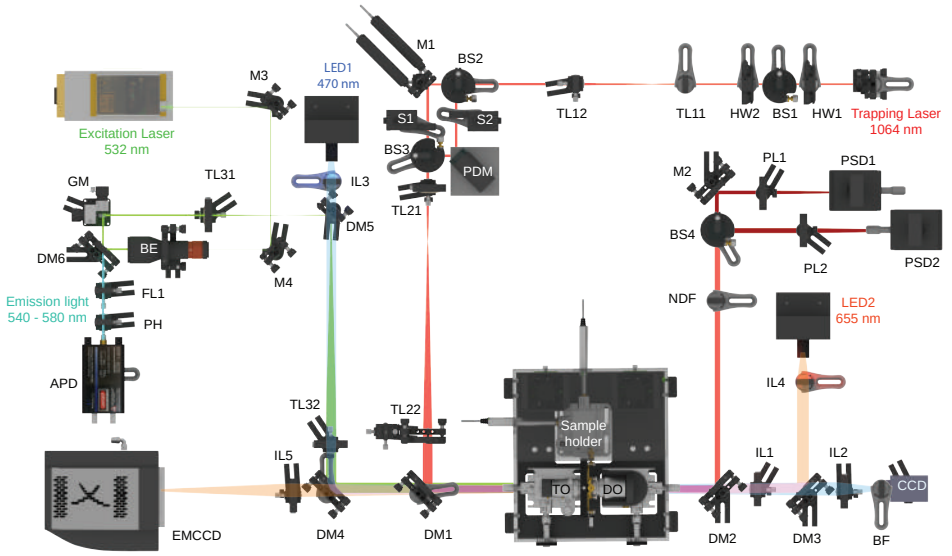


Figure 2.13: **Drawing of the optical tweezers setup combined with the fluorescence imaging.** Red - trapping beam, blue, orange - imaging beams, green - excitation beam, cyan - emission light. The image is made by Mario Avellaneda.

Widefield imaging

In this configuration, TL31 is absent and DM4 is a longpass dichroic mirror with a 532 nm threshold (ZT532rdc, Chroma). The excitation laser is focused on the back focal plane of the TO by means of TL32, providing bright-field illumination on the sample. The dichroic mirror allows the emission light to reach the EMCCD camera, reflecting the excitation light.

Confocal imaging

By adding the lens TL31, the beam becomes collimated after the telescope, and the plane of the galvomirror and the back focal plane of the TO are conjugated. The expansion provided by the telescope also ensures that the beam overfills the entrance aperture of the objective, creating a diffraction-limited spot in the sample. The galvomirror is then scanned to excite the sample locally. Depending on which mirror is at DM4, two modes of detection are possible: descanning or digital confocal. In the first mode, DM4 is a longpass dichroic mirror with 612 nm threshold, and both the excitation and the emission light are reflected. Therefore, the galvomirror can descan the emitted light and guide it to an avalanche photodiode (APD) after passing through a 50 nm pinhole (PH). Synchronizing the galvomirror movement and the photon counts the image can be reconstructed pixel by pixel with the aid of software. In the digital confocal mode, a longpass mirror with 532 nm threshold is used, and the emitted light will be directed to the EMCCD. In this case, a more complicated digital analysis based on point spread functions can be

performed. At the moment, this detection mode has not been fully developed, but in the future it can provide an additional method for fluorescence microscopy in our setup.

Furthermore, total internal reflection fluorescence (TIRF) is compatible with either excitation and detection mode (widefield illumination, descanning or digital confocal). In this configuration, the first lens of the telescope TL31 is absent, and DM4 is displaced in order to change the entrance point of the beam into the objective.

In both modes, the control over the composition of the experimental buffer is critical to allow effective interaction of labeled chaperones with the protein client on the beads and to suppress the background signal from unbound chaperones. We will employ a microfluidic scheme similar to the one presented by Biebricher et al. [9] for that. A laminar non-mixing flow is established in the flowcell to first allow the interaction of dye with the tether and then perform force-extension experiments and collect fluorescence emission in the dye-free "measurement channel" (Fig. 2.14). Our company-made flowcell design (Fig. 2.8d-f) is optimized for this regime and allows stable turbulence-free laminar flow.

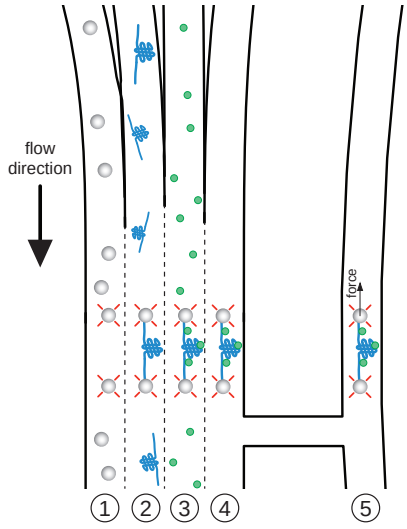


Figure 2.14: **Scheme of the laminar-flow fluorescence single-molecule force spectroscopy assay.** The main channel of the flowcell features four laminar flows, and the optical traps move through the flows sequentially as indicated by the numbers below. First beads are trapped, then the protein construct is trapped in situ, then it is exposed to fluorescently labeled chaperones and allowed to interact. Next the tethered complex is transferred into the isolated channel to prevent mixing, flow is stopped and the force spectroscopy measurements are performed while collecting the fluorescence signal.

2.5.2. Heat shock proteins in heat shock conditions

Chaperones have been discovered because they were massively overexpressed in the cell upon heat shock (37-42°C). Indeed, certain chaperones, in addition to important functions at non-stress temperatures, are essential for cell viability under heat shock conditions. Furthermore, some chaperones, such as many small heat shock proteins, are only active at elevated temperatures. Hence it seems logical to extend the chaperone-assisted protein folding studies to the heat shock conditions to be able to get mechanistic insight into the function of such chaperones and to understand whether chaperones that are also active at room temperature perform differently upon heat shock.

In order to achieve it, the optical tweezers apparatus should be equipped with a heating module that allows heating the sample to desired temperatures. Efforts have been made by several groups to combine optical tweezers with temperature control. However, it still remains a challenging task, since the performance of trapping unit is severely affected by temperature changes. We have implemented a heating unit, which is based on heating of the air around the sample, in the new optical tweezers setup and tested its performance. Chapter 3 is dedicated to the detailed description of our temperature control method.

2.6. Extended data

Table 2.1: List of optics and optomechanical components used in the construction of the setup.

Stage	Drawing	Essential parts	Manufacturer & model
Trapping laser control	Trapping laser	Trapping laser	YLM-5-LP-SC, IPG
		Fiber clamp magnetic arm	2704101, Horex
		V-clamp mount	KM100V/M, Thorlabs
	HW1	High-precision rotation mount	PRM05/M, Thorlabs
		Half-wave plate	WPH05M-1064, Thorlabs
	HW2	High-precision rotation mount	PRM05/M, Thorlabs
		Half-wave plate	WPH05M-1064, Thorlabs
	BS1	Laser line polarizing cube beamsplitter	05BC16PC.9, Newport
		12.7 mm cube adapter kit	UGP-KIT-1, Newport
		Ultima gimbal prism mount	UGP-1, Newport
		Beam dump	BT600/M, Thorlabs

Continued on next page

Table 2.1: List of optics and optomechanical components used in the construction of the setup.

Stage	Drawing	Essential parts	Manufacturer & model
Telescope 1	TL11	Telescope lens $f = 50$ mm	LA1131-C, Thorlabs
		Mount for 25.4 mm optics	LMR1/M, Thorlabs
	TL12	Telescope lens $f = 125$ mm	LA1986-C, Thorlabs
		Mount for 25.4 mm optics	LMR1/M, Thorlabs
		Iris diaphragm	ID25/M, Thorlabs
		Iris diaphragm	ID25/M, Thorlabs
Beam steering	BS2	Laser line polarizing cube beamsplitter	05BC16PC.9, Newport
		12.7 mm cube adapter kit	UGP-KIT-1, Newport
		Ultima gimbal prism mount	UGP-1, Newport
	BS3	Laser line polarizing cube beamsplitter	05BC16PC.9, Newport
		12.7 mm cube adapter kit	UGP-KIT-1, Newport
		Ultima gimbal prism mount	UGP-1, Newport
	M1	Mirror	NB1-K14, Thorlabs
		Stepper motorized mirror mount	KS1-ZST, Thorlabs
	PDM	Piezo-driven mirror actuator	Nano-MTA2XHS, Mad City Labs
		Mirror	NB1-K14, Thorlabs
		Mounting block	Mad City Labs
	S1	Beam shutter	SH05, Thorlabs
	S2	Beam shutter	SH05, Thorlabs
	DM1	Dichroic mirror (SP 805 nm)	DMSP805L, Thorlabs
		Mirror kinematic mount	KM200, Thorlabs
	DM2	Dichroic mirror (SP 620 nm)	620dcspxr, Chroma
		Kinematic rectangular mount	KM100C, Thorlabs
		Iris diaphragm	ID25/M, Thorlabs

Continued on next page

Table 2.1: List of optics and optomechanical components used in the construction of the setup.

Stage	Drawing	Essential parts	Manufacturer & model
		Iris diaphragm	ID25/M, Thorlabs
Telescope 2	TL21	Telescope lens $f = 100$ mm	LA1050-C, Thorlabs
		Mount for 50.8 mm optics	LMR2/M, Thorlabs
		XY dovetail translation stage	DT12XY/M, Thorlabs
	TL22	Telescope lens $f = 300$ mm	LA1484-C, Thorlabs
		Mirror kinematic mount	KM200, Thorlabs
		XYZ dovetail translation stage	DT12XYZ/M, Thorlabs
		Iris diaphragm	ID25/M, Thorlabs
		Iris diaphragm	ID25/M, Thorlabs
Objectives & sample	FO	Water immersion 60x1.2 NA microscope objective	CFI Plan Apo VC 60x WI, Nikon
		XYZ translation stage (LH)	M-561D-XYZ-LH, Newport
		Manual actuator (x3)	SM-13, Newport
		Custom built objective holder	
	BO	Water immersion 60x1.2 NA microscope objective	CFI Plan Apo VC 60x WI, Nikon
		XYZ translation stage (LH)	M-561D-XYZ-LH, Newport
		Manual actuator (x3)	SM-13, Newport
		Custom built objective holder	
	SC	Custom built sample holder	
		XYZ translation stage (RH)	M-561D-XYZ, Newport
		Motorized actuator (x3)	Conex-TRA12CC, Newport
	LED1	Blue diode (470 nm wavelength)	LXML-PB01-0040, Philips Lumileds

Continued on next page

Table 2.1: List of optics and optomechanical components used in the construction of the setup.

Stage	Drawing	Essential parts	Manufacturer & model
		Custom built diode mount	
	LED2	Red diode (660 nm wavelength)	LXM3-PD01, Philips Lumileds
		Custom built diode mount	
	IL1	Bi-convex lens $f = 250$ mm	LA1461-C, Thorlabs
		Mount for 25.4 mm optics	LMR1/M, Thorlabs
	IL2	Bi-convex lens $f = 50$ mm	LB1471-A, Thorlabs
		Mount for 25.4 mm optics	LMR1/M, Thorlabs
	IL3	Telescope lens $f = 200$ mm	LA1979-A, Thorlabs
		Mount for 50.8 mm optics	LMR2/M, Thorlabs
	IL4	Bi-convex lens $f = 100$ mm	LB1630-A, Thorlabs
		Mount for 50.8 mm optics	LMR2/M, Thorlabs
	IL4	Bi-convex lens $f = 200$ mm	LB1199-A, Thorlabs
		Mount for 50.8 mm optics	LMR2/M, Thorlabs
	DM3	Dichroic mirror (LP 800 nm)	800dcxru, Chroma
		Kinematic rectangular mount	KM100C, Thorlabs
	DM4	Dichroic mirror (LP 612 nm)	T612lpxr, Chroma
		Kinematic rectangular mount	KM100C, Thorlabs
	CCD	CCD camera	XCG-V60E, Sony
		Custom made mount	
	EMCCD	EMCCD camera	iXon 897 Ultra, Andor
		Custom made mount	
	BF	Bandpass filter	FB450-40, Thorlabs
		Mount for 25.4 mm optics	LMR1/M, Thorlabs
	NDF	Ø25 mm NIR absorptive ND filter	NENIR20A, Thorlabs
		Mount for 25.4 mm optics	LMR1/M, Thorlabs

Continued on next page

Detection

Table 2.1: List of optics and optomechanical components used in the construction of the setup.

Stage	Drawing	Essential parts	Manufacturer & model
	BS4	Laser line polarizing cube beamsplitter	10BC16PC.9, Newport
		Ultima gimbal prism mount	UGP-1, Newport
	M2	Mirror	BB2-E03, Thorlabs
		Mirror kinematic mount	KM200, Thorlabs
	PL1	Bi-convex lens $f = 150$ mm	LA1417-C, Thorlabs
		Mount for 50.8 mm optics	LMR2/M, Thorlabs
	PL2	Bi-convex lens $f = 150$ mm	LA1417-C, Thorlabs
		Mount for 50.8 mm optics	LMR2/M, Thorlabs
	PSD1	Position-sensitive photodetector	DL100-7PCBA3, Pacific Silicon Detectors
		Custom made mount	
	PSD2	Position-sensitive photodetector	DL100-7PCBA3, Pacific Silicon Detectors
		Custom made mount	
Misc		Ø25 mm pedestal pillar posts	Thorlabs
		Ø25 mm pillar post spacers	Thorlabs
		Ø25 mm pillar post extensions	Thorlabs
		Clamping forks	PS-F, Newport
		Ø12.7 mm swiveling magnetic base post holders	Thorlabs
		Slip-on post collars	R2/M, Thorlabs
		Ø12.7 mm stainless steel optical posts	Thorlabs
		Infrared alignment cards and disks	Thorlabs

References

[1] A. Ashkin, J. M. Dziedzic, J. E. Bjorkholm, and S. Chu, *Observation of a single-beam gradient force optical trap for dielectric particles,* Opt. Lett. **11**, 288 (1986).

[2] C. Bustamante, Y. R. Chemla, and J. R. Moffitt, *High-resolution dual-trap optical tweezers with differential detection: instrument design.* Cold Spring Harb. Protoc. **2009**, pdb.ip73 (2009).

[3] J. R. Moffitt, Y. R. Chemla, D. Izhaky, and C. Bustamante, *Differential detection of dual traps improves the spatial resolution of optical tweezers.* Proc. Natl. Acad. Sci. U. S. A. **103**, 9006 (2006).

[4] K. Neuman and S. Block, *Optical trapping,* Rev. Sci. Instrum. **75**, 2787 (2004).

[5] K. Berg-Sø rensen and H. Flyvbjerg, *Power spectrum analysis for optical tweezers,* Rev. Sci. Instrum. **75**, 594 (2004).

[6] S. Tolić-Nørrelykke and E. Schäffer, *Calibration of optical tweezers with positional detection in the back focal plane,* Rev. Sci. Instrum. **77**, 103101 (2006), arXiv:0603037v2 [arXiv:physics] .

[7] P. M. Hansen, I. M. Tolic-Nørrelykke, H. Flyvbjerg, and K. Berg-Sørensen, Comput. Phys. Commun. , 572arXiv:0603037 .

[8] P. Gross, G. Farge, E. J. G. Peterman, and G. J. L. Wuite, *Methods Enzymol.*, 1st ed., Vol. 475 (Elsevier Inc., 2010) pp. 427–53.

[9] A. Biebricher, S. Hirano, J. H. Enzlin, N. Wiechens, W. W. Streicher, D. Huttner, L. H.-C. Wang, E. a. Nigg, T. Owen-Hughes, Y. Liu, E. Peterman, G. J. L. Wuite, and I. D. Hickson, *PICH: a DNA translocase specially adapted for processing anaphase bridge DNA.* Mol. Cell **51**, 691 (2013).

[10] W. Smith, *Modern Optical Engineering, 4th Ed.* (McGraw Hill Professional, 2007) p. 764.

[11] S. B. Smith, Y. Cui, and C. Bustamante, *Optical-trap force transducer that operates by direct measurement of light momentum.* Methods Enzymol. **361**, 134 (2003).

[12] W. M. Lee, P. J. Reece, R. F. Marchington, N. K. Metzger, and K. Dholakia, *Construction and calibration of an optical trap on a fluorescence optical microscope.* Nat. Protoc. **2**, 3226 (2007).

[13] D. Ott, S. N. S. Reihani, and L. B. Oddershede, *Crosstalk elimination in the detection of dual-beam optical tweezers by spatial filtering,* Rev. Sci. Instrum. **85** (2014), 10.1063/1.4878261.

[14] G. I. Bell, *Models for the specific adhesion of cells to cells.* Science **200**, 618 (1978).

[15] J. Shillcock and U. Seifert, *Escape from a metastable well under a time-ramped force,* Physical Review E **57**, 7301 (1998).

[16] M. Schlierf and M. Rief, *Single-molecule unfolding force distributions reveal a funnel-shaped energy landscape.* Biophys. J. **90**, L33 (2006).

[17] C. Bustamante, Y. R. Chemla, and J. R. Moffitt, *High-Resolution Dual-Trap Optical Tweezers with Differential Detection: Managing Environmental Noise,* Cold Spring Harb. Protoc. **2009**, pdb.ip72 (2009).

[18] I. Heller, T. P. Hoekstra, G. a. King, E. J. G. Peterman, and G. J. L. Wuite, *Optical tweezers analysis of DNA-protein complexes.* Chem. Rev. **114**, 3087 (2014).

[19] S. Hohng, R. Zhou, M. K. Nahas, J. Yu, K. Schulten, D. M. J. Lilley, and T. Ha, *Fluorescence-force spectroscopy maps two-dimensional reaction landscape of the holliday junction.* Science **318**, 279 (2007).

[20] S. Lee and S. Hohng, *An optical trap combined with three-color FRET,* J. Am. Chem. Soc. **135**, 18260 (2013).

[21] M. I. Snijder-Van As, B. Rieger, B. Joosten, V. Subramaniam, C. G. Figdor, and J. S. Kanger, *A hybrid total internal reflection fluorescence and optical tweezers microscope to study cell adhesion and membrane protein dynamics of single living cells.* J. Microsc. **233**, 84 (2009).

[22] I. Heller, G. Sitters, O. D. Broekmans, G. Farge, C. Menges, W. Wende, S. W. Hell, E. J. G. Peterman, and G. J. L. Wuite, *STED nanoscopy combined with optical tweezers reveals protein dynamics on densely covered DNA.* Nat. Methods **10**, 910 (2013).

[23] M. J. Comstock, T. Ha, and Y. R. Chemla, *Ultrahigh-resolution optical trap with single-fluorophore sensitivity.* Nat. Methods **8**, 335 (2011).

[24] M. D. Brenner, R. Zhou, and T. Ha, *Forcing a connection: impacts of single-molecule force spectroscopy on in vivo tension sensing.* Biopolymers **95**, 332 (2011).

[25] G. Sirinakis, Y. Ren, Y. Gao, Z. Xi, and Y. Zhang, *Combined versatile high-resolution optical tweezers and single-molecule fluorescence microscopy.* Rev. Sci. Instrum. **83**, 093708 (2012).

3

A low-drift heating method for optical tweezers

Temperature is an important parameter in optical tweezers experiments. Efforts have been made to introduce temperature control into the confined environment of optical tweezers setups, but it still remains challenging. Here we present a novel heating method for optical tweezers based on heating of the air around the sample chamber that is easy to implement. Using this method, we achieved temperature control of the sample with 0.5°C precision. Heating of the air introduces a slight increase of beam pointing fluctuations that is minimized by physically shielding the beam. Dual-trap optical tweezers are well isolated from the differential drift that is further minimized by the uniform heating of both the objectives and the sample chamber. It allows one to experiment with the same biomolecule before, during and after the temperature upshift. We performed single-molecule DNA stretching experiments at various temperatures up to 45°C to demonstrate the setup performance. The DNA overstretching transition force is lowered upon temperature increase consistent with previous reports.

3.1. Temperature control methods for optical tweezers

Optical tweezers were invented by Ashkin et al. in the early 1970s [1] and for the last three decades they have been used extensively in physics and biology [2, 3]. They have been used successfully to manipulate small objects ranging from atoms [1] and various nanostructures [4] to micrometer-sized particles [5], living microbes [6] and individual cells in animals [7]. The technique allows one to apply forces ranging from sub-picoNewton up to hundreds of picoNewtons and to detect displacements with sub-nanometer precision that fit well to the force and displacement scale of numerous natural processes. Optical tweezers are therefore a useful tool to address a wide range of scientific puzzles. In the physical sciences, optical tweezers provide important insights in the fields of colloid physics, physics of condensed matter and quantum information science by providing non-invasive access to the organization of mesoscopic and microscopic matter [3]. In biology, optical tweezers are applied to study folding dynamics of nucleic acids [8, 9] and proteins [10, 11], mechanics of molecular motors [12–14] at the single-molecule level, mechanical properties of single biopolymers [15–18] and their networks [19], and bacterial motility [6] and cell mechanics [20] at the single-cell level.

Temperature is an important experimental parameter in many processes [21]. Examples include temperature dependence of reaction rates[1] [22], thermomechanical studies on DNA and proteins [23–27] and ligand binding studies [28, 29]. Furthermore, certain biomolecules, such as molecular chaperones [30, 31] and DNA-binding proteins [32] are temperature-activated. These set a need for precise control over the sample temperature during the experiment.

Efforts have been made to combine optical tweezers experiments with temperature control. Several groups used an additional near-infrared laser to heat the water locally around the trapped particle [33, 34]. Laser-induced local heating is very fast and simple to implement and control. However, the disadvantages of this method are the water convection caused by heat gradients, which creates an additional drag force on the trapped particle due to excessive water flows, and the limited heating range. Additionally, the laser used for trapping (usually near-infrared) is a heat source itself that may introduce an observable temperature gradient inside the sample chamber as well as instrumental drift [35]. An alternative method is to provide a delocalized heat source to the sample to avoid parasitic convection-related forces. Williams et al. [23] placed the sample between copper plates brought in thermal contact with a thermoelectric cooler. This method has wide temperature range (limited only by properties of the sample and the optics) and allows for both heating and cooling. However, the presence of heating elements required the usage of objectives with longer working distance that reduces the spatial resolution. Mao et al. [33] adjusted the method for shorter working distance objectives by making copper jackets with internal fluid circula-

[1] According to Arrhenius law, the reaction rate depends on temperature exponentially, hence even small temperature change significantly affects the reaction kinetics.

tion that fit tightly around the objectives. With this approach the local heating is minimized but the fabrication of the jackets is relatively sophisticated, and fluid circulation adds to the noise in the signal. This temperature control method has also been used by Mahamdeh and Schäffer [36] to develop an objective temperature feedback system based on heating foils wrapped around the objectives. It is capable of keeping the temperature of the objectives constant with milliKelvin precision to compensate for parasitic heating by trapping laser. However, they did not utilize this system for heating. Therefore, most noted problems that one is faced with whilst controlling the temperature in optical tweezers experiment are the following: 1) Drift of the optical elements caused by their thermal expansion that leads to misalignment of the tweezers setup and degradation of its trapping and imaging properties [35]. 2) Local temperature gradients [34, 35]. 3) Limited space for implementing thermal elements and sensors.

Here we present a novel easy-to-implement method to control sample temperature in optical tweezers, and demonstrate its performance in a custom breadboard-assembled dual-trap optical tweezers setup. The method, which is based on heating of the air around the objectives, allows us to heat the sample and control its temperature with 0.5°C precision. Like any dual-trap optical tweezers, our proposed setup is well isolated from the differential drift of any kind due to the fact that both traps are generated from the same source [37]. The drift is further minimized by the uniform heating of the sample and both the trapping and detection objectives. It allows one to experiment with the same biomolecule before, during and after the heating process. To further test the capabilities of our setup, we performed single-molecule DNA stretching experiments at various temperatures.

3.2. Methods

3.2.1. Optical tweezers setup with temperature control unit

Our experimental setup is a custom-made dual trap optical tweezers apparatus with differential detection (see Fig. 3.1a). The instrumentation configuration used in this chapter is identical to the configuration used in chapter 2 and is described in detail in section 2.2. In addition, to control the sample temperature, we developed a temperature control unit that allows for quick and uniform heating and passive cooling of both objectives and sample by heating air around the system (Fig. 3.1b). It features a thermally insulated plastic box around the sample chamber containing both objectives and two heaters (FLH030, Phannenberg) positioned on the back wall of the box. Temperature in this heating chamber is monitored with a Pt1000 sensor (Pt100 Insertion thermometer, Jumo), and the power to the heaters is adjusted accordingly by a PID controller (3216CCVHRRXXR, Eurotherm). Two low-speed fans equipped with custom outlets are positioned above the heaters to ensure proper air mixing and a uniform temperature within the plastic box. Temperatures of the sample and of the trapping objective are monitored independently by placing a thermocouple (COCO-003-BW, Type T, Omega Engineering) inside the flowcell close to the traps and fixing it on the trapping objective.

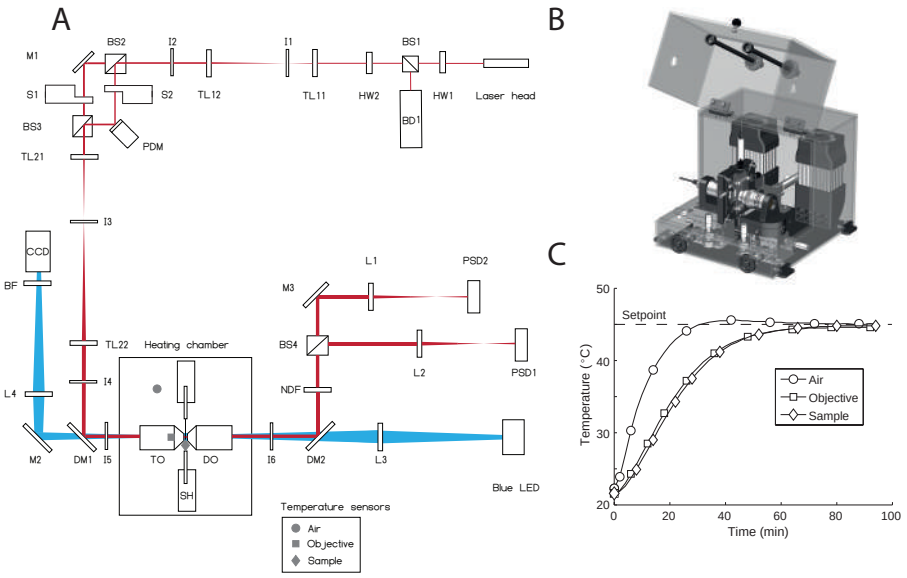


Figure 3.1: **Experimental setup. a,** Schematic drawing of the dual-trap optical tweezers setup equipped with a heating unit. Grey shapes indicate an approximate location of temperature sensors inside the heating chamber. The near-infrared trapping laser beam is shown in red. The LED light used for imaging in blue. Refer to chapter 2 for detailed description of the setup. **b,** Image of the heating chamber. Two heaters (grey) are mounted on the back wall of the chamber with the fans above and custom outlets below (both black). **c,** Heating profile of the chamber. Temperature readouts of the sensors are plotted versus temperature. Setpoint temperature is set to 45°C.

3.2.2. DNA stretching experiments

Anti-digoxigenin (Anti-Dig) antibody-coated polystyrene beads (2.1 µm diameter, DIGP-20-2, Spherotech) were coated with DNA (2553 base pairs) as described in ref. [38]. Neutravidin-coated beads (2.1 µm diameter, NVP-20-5, Spherotech) were stored until use. A DNA tether (immobilized DNA molecule between two optically trapped beads) was assembled inside the flow chamber in HMK buffer (for details see ref. [38]).

DNA stretching experiments were performed at various temperatures using the optical tweezers setup. Deflections of trapped beads and the position of the steerable trap were collected at 50 kHz. These signals were filtered using an analog 8th order Butterworth filter at 20 kHz, and further averaged 256 times by an analog averaging filter to roughly 195 Hz prior to saving the data. Data for calibrations were acquired at 100 kHz and filtered at 20 kHz.

3.3. Results

3.3.1. Performance of temperature control unit

All experiments in this study were performed in a temperature controlled room. Room temperature was measured to be $T_0 = 21.2^\circ$C. To assess the characteristics of the heating chamber, we used a temperature setpoint of 45°C, switched the fans on, and sampled the readouts from the temperature sensors as the system is heated. The trapping objective and sample reached the setpoint temperature synchronously after 85 minutes since heating started (Fig. 3.1c). The air sensor temperature first displayed a slight overshoot of 2°C and then reached the setpoint simultaneously with the objective and the sample. Once heated, temperature inside the heating chamber remained stable for hours (data not shown). However, when the fans were switched off longer than 40 minutes, a detectable temperature gradient grew in the heating chamber. Next we probed the effects of the trapping laser on the temperature of the system. Two traps were overlapped spatially in the sample plane, and the sample temperature was measured by placing the sample-associated thermocouple at different distances from the laser focus. However, the direct measurement of temperatures close to the laser focus is not possible using the thermocouple, since it strongly absorbs the laser radiation that results in additional heating and false results. After 3.5 hours of keeping the laser on at 500 mW power in laser focus (250 mW per trap, corresponds to the total power of 2 W emanating from the laser output collimator), the sample temperature increased and became equal to 27.2°C at the minimal distance from the focus accessible for the thermocouple (40 μm). Interestingly, the temperature of the trapping objective also increased by 6°C, suggesting that the measured sample heating occurred not exclusively because of absorption of laser light by water but also due to a thermal exchange with the objective acting as a heatsink. To summarize, the laser power affects the sample temperature significantly within the relevant power range and should be accounted for. However, contribution of the trapping laser to the sample temperature depends only on laser power and mechanical characteristics of the setup. Thus the extent of additional heating can be pre-calibrated, and the temperature setpoint can be adjusted accordingly. An accuracy of our heating method is limited to an accuracy of laser-caused heating estimation. It can be increased by minimizing the trapping laser light absorption by the experimental medium (see Discussion). Since the reservoirs containing the experimental buffers were not preheated, an opening of fluid flow in the flowcell resulted in a transient decrease of the sample temperature by 1°C. However, such decrease is irrelevant at our experimental conditions. The sample temperature stabilizes within 30 s after stopping the flow due to the intensive thermal exchange between the objectives and the sample.

3.3.2. Effects of heating on data quality

To see whether our heating method adds noise to the system and affects data quality, we analyzed the Brownian fluctuations of the trapped beads at elevated temperatures and compared them to the measurements when the heating is off.

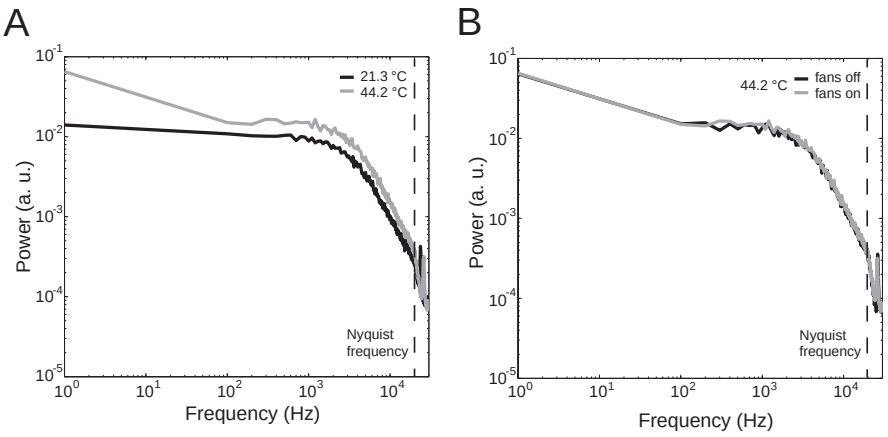


Figure 3.2: **Noise characteristics of the heating unit.** Power spectra of Brownian fluctuations of trapped beads are measured at various conditions. **a,** Power spectra at room (black) and elevated (grey) temperatures indicate increased fluctuations at elevated temperature, especially at low frequencies, that are attributed to the warm air convection. **b,** Power spectra measured at elevated temperature with the fans on (grey) and off (black) do not show additional increase of fluctuations due to the fans.

Power spectra at elevated temperature display a slight increase of the bead fluctuations at all measured frequencies, with a more apparent increase at low frequencies (Fig. 3.2a). Such effect at low frequencies can be attributed to air fluctuations due to extensive mixing and convection. To reduce these effects, we made a beam guide that features two tubes guiding the incident laser beam from an inlet of the heating chamber to the trapping objective and the collected beam from the detection objective to the outlet of the chamber. Such simple approach lowers the low-frequency fluctuations at elevated temperatures (Extended Data Fig. 3.1), however, they are still higher compared to the low-frequency fluctuations at room temperature (Fig. 3.1a). Notably, keeping the fans on during the measurement did not result in additional fluctuations (Fig. 3.1b), indicating that mechanical fluctuations (that are attributed to the movement of the fans and are transmitted into the signal through the optical table) are not responsible for increased noise. It suggests that the noise is increased due to convection of the warm air, which is still present even if the fans are used to mix the air extensively. The convection affects the refractive index of the air that is temperature dependent, and, consequently, increases the beam pointing fluctuations.

3.3.3. Drift estimates with DNA stretching

Any rise in temperature is known to affect the performance of the optical tweezers setup severely by causing thermal expansion of the optical components [33, 36]. To estimate the amount of drift that the heating introduces into actual measurements, we stretched individual DNA molecules as the chamber is heated. A

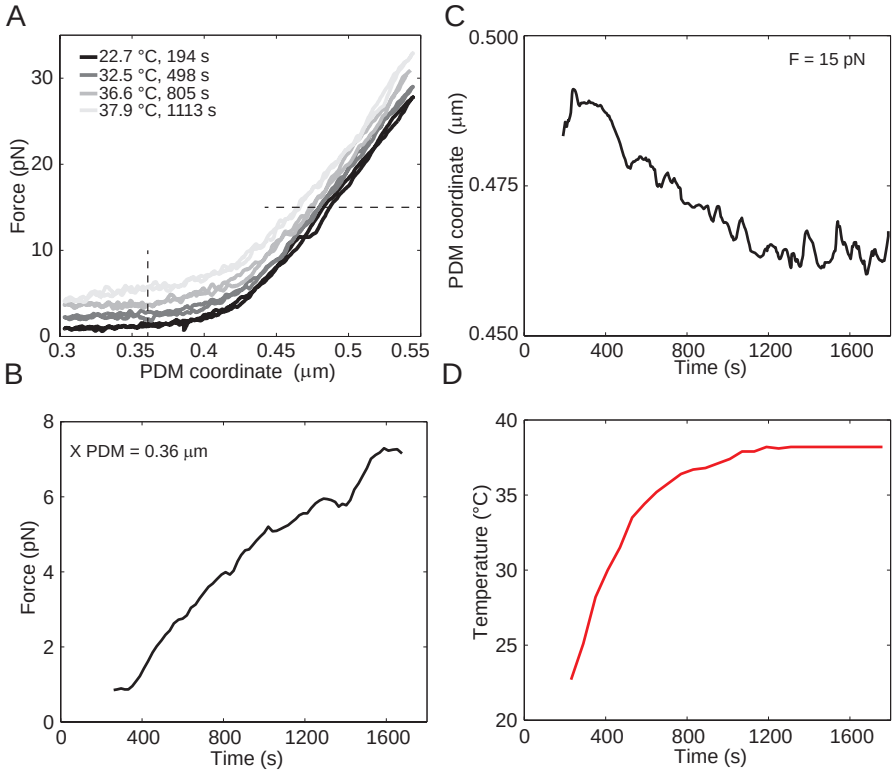


Figure 3.3: **DNA stretching as the system is heated.** Initial temperature is 21.2°C. Temperature setpoint is at 37°C. **a,** Force acting on DNA is plotted versus the coordinate of steerable mirror. Time from the start of the heating process is color-coded: earlier curves are shown in darker shades (see the figure legend). Dashed lines denote the cooordinates, at which changes with time were plotted. **b,** The change of force with time during heating at constant PDM coordinate. **c,** The change of PDM coordinate with time during heating at constant force.**d,**Temperature increase during the heating.

DNA tether was assembled at room temperature and stretched and relaxed constantly by moving the steerable trap between two fixed positions upon heating to 37°C (Fig. 3.3a). Notably, we were able to keep following the beads movement and measuring on the tether during the heating process. This indicates that realignment of the optics, reported to be necessary to continue the measurements after heating in the previous attempt [33], is not required. However, the measurements are not completely drift-free. For a fixed position of the steerable trap, the force acting on the DNA increased upon heating indicating an increased distance between traps due to the differential drift (Fig. 3.3b). Importantly, the amount of the drift does not depend on the position of the steerable trap, thus the thermal drift can easily be corrected for. At low force regime, the DNA mechanics does not change significantly upon heating (see the next section and

Fig. 3.4a), thus it can be assumed that the force increase comes solely from the drift, and this value can be used to compensate for the drift along the complete stretching trajectory.

3.3.4. DNA stretching experiments

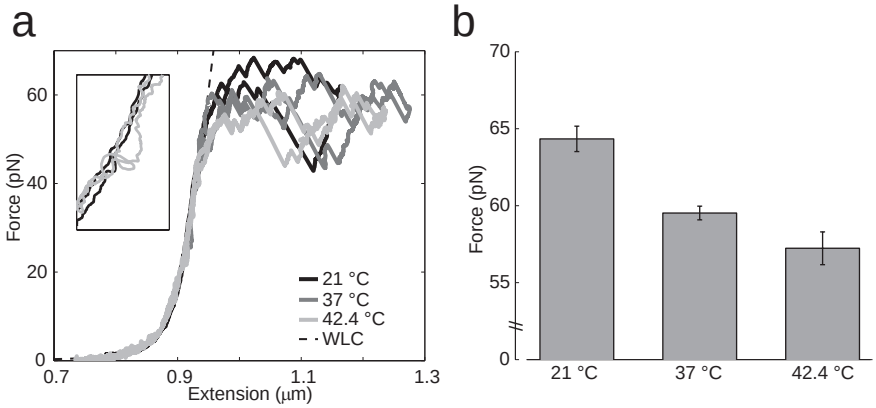


Figure 3.4: **DNA overstretching at various temperatures. a,** Force-extension curves of DNA stretching. The stretching curves are displayed as solid lines. The dashed line indicates a theoretical compliance of the DNA molecule calculated in accordance with the worm-like chain model (WLC).**b,** DNA overstretching force at different temperatures. The errorbars indicate the standard deviations of the forces from various measurements (number of events $N = 7$ at 21°C, $N = 3$ at 37°C, $N = 5$ at 42.4°C).

We further tested our method by overstretching individual double-stranded (ds)DNA molecules at room and elevated temperatures. A typical force-extension curve corresponding to stretching at room temperature is presented in Fig. 3.4a (black curve). The force rose as the DNA helix was extended near its contour length, and the dsDNA elasticity up to 65 pN could be described by the discrete worm-like chain (WLC) model proposed by Manghi et al. [39] (see [38] for details of fitting procedure) As the force reached about 65 pN, a so-called DNA overstretching transition occurred as reported previously, resulting in a 70% increase of DNA length [16, 40]. At the overstretching force, dsDNA may undergo different structural changes depending on experimental conditions. For the torsionally unconstrained DNA, two types of transitions can occur: unpeeling of strands from the ends of the molecule, that involves disruption of base pairs and yields ssDNA; and the so-called "B-to-S" transition, where the DNA helix unwinds with the base pairs remaining intact yielding another dsDNA structure called S-DNA [18, 41]. In our experimental conditions, the loss of basepairing interactions occurs, but we do not have an evidence for the complete absence of S-DNA formation [38]. The transition occurred in bursts observed at the overstretching plateau as a saw-tooth pattern that is consistent with the force-induced loss of basepairing interactions [42]. The overstretching force for each stretching curve was defined by averaging

over the plateau (extension between 1 and 1.1 μm). Relaxation curves, when the DNA extension was decreased, exhibited significant hysteresis [42]. Consecutive stretching after a 5 second waiting period when the DNA was relaxed yielded again similar curves that can be fitted to WLC, followed by overstretching. It indicates that the single-stranded fragments of DNA reannealed and DNA regenerated to the double-stranded conformation. When the DNA was stretched at elevated temperatures (Fig 3.4a), the magnitude of the overstretching force decreased by 7 pN (at 43°C), consistent with previously published results [33, 43]. The amount of the DNA that remained single-stranded upon relaxation (estimated from the hysteresis area of relaxation curves) increased with the temperature (Fig 3.4b), again consistent with previous reports [33]. Since the pH of the HMK buffer changes with temperature (pH = 7.5 at 20°C, but pH = 7.2 at 45°C), it may also affect the DNA overstretching force. However, a change of 1 pN per pH unit in this pH range has been reported by Williams et al. with a DNA similar to ours [43], hence it does not explain the observed force decrease.

3.4. Discussion

The heating method presented here is easy to implement and can be readily adjusted for almost any experimental layout. The synchronous heating of the trapping objective and the sample allowed us to avoid the laborious process of placing thermocouple inside the flowcell. In order to probe the sample temperature, one could measure the objective temperature that is much easier to achieve. The presence of additional heat sources, such as the near-infrared trapping laser, does not affect the application of this method, since it only introduces an offset into the sample temperature and can be easily corrected for by choosing a lower setpoint for the temperature controller. In order to decrease the local heating effects of near-infrared trapping laser, one could select a trapping laser with different wavelength (for example, 830 nm, which has 4.7 times less absorbance in water than 1064 nm) [44]. Alternatively, one may perform the trapping in D2O that does not absorb infrared light and thus does not heat [33], but may affect the stability of biomolecules under study [45]. Heating of air around the sample chamber and the objectives allowed us to heat them from room temperature (21.2°C) to 45°C in 85 minutes and to 37°C in 25 minutes (data not shown). It is slower than the timescale of many biological processes (e.g. protein folding), hence it does not allow exploration of the dynamics of those induced solely by temperature change. However, non-equilibrium processes still can be studied at various temperatures by applying force to an object of study. To make the temperature shift faster, one could incorporate more heating elements or increase the power of existing ones. However, air being a very bad heat conductor is the limiting factor for the heating speed in such design, thus we do not expect a major speed increase by this improvement. One could also work on better thermal insulation to avoid heat leakages. In our design, the side and top walls of heating chamber are covered with insulation foam, but the bottom of the chamber is not insulated, meaning that the heat may dissipate through the optical table. Free surface on the table can be easily insulated, however, the

bases of the optical elements such as objectives are difficult to insulate without affecting their mechanical stability.

A slight degradation of the image quality of the trapped beads occurs upon heating (data not shown), which may affect the performance of particle tracking algorithms used to extract the beads positions from the microscope image. Our experiments, however, did not suffer from it, since we used this image only for reference purposes. We combined the signals from position-sensitive detectors and the steerable mirror internal position sensor to calculate the tether extension. Uniform heating of both objectives and the sample chamber minimizes the differential drift of the optical components and allows experiments at room and elevated temperatures without realignment of the optics. Furthermore, our low-drift heating method allows one to experiment with the same biomolecule before, during, and after the heating process that may be of interest for various applications.

However, heating the air around the sample does result in increased fluctuations observed in the data (Fig. 3.2a). We also note that the DNA force-extension curves look slightly noisier at the elevated temperatures. They feature distinct low-frequency "loops" (Fig. 3.4a, inset) that we never observed in the experiments at room temperature. We believe that these loops originate from the fluctuations of hot air inside the heating chamber. We suppressed these effects by switching the fans off during the measurements and by shielding the laser beam from the air fluctuations using the beam guide. If further suppression is required, it can be done by housing the optics in helium atmosphere [46].

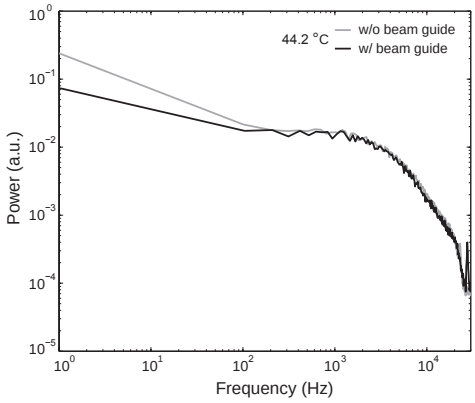
3.5. Conclusions

The method presented here provides a simple, non-invasive, low-cost and low-drift heating for optical tweezers. It is capable of stable and uniform heating of the sample with the precision of 0.5°C. To the best of our knowledge, this heating method is the first that allows one to follow the transitions in the biomolecule during the heating process in real time. Such an upgrade makes the traditional optical tweezers a powerful tool to investigate the mechanisms underlying heat-activated biological reactions.

3.6. Acknowledgements

This work was supported by the research programme of the Foundation for Fundamental Research on Matter (FOM), which is part of the Netherlands Organization for Scientific Research (NWO). We are grateful to Dr. Andrea Candelli and Dr. Andreas Biebricher for sharing their insightful ideas on dual-trap optical tweezers construction. We also thank Dr. Nicholas Kurniawan for critical reading of the manuscript.

3.7. Extended data



Extended Data Figure 3.1: **The beam guide suppresses the fluctuations in the measured signal.** Power spectra of trapped bead fluctuations are measured at elevated temperature in the presence (black) and in the absence (grey) of the beam guide. It indicates that the beam guide is capable of suppressing low-frequency fluctuations that are apparent at elevated temperature.

References

[1] A. Ashkin, *Acceleration and Trapping of Particles by Radiation Pressure*, Phys. Rev. Lett. **24**, 156 (1970).

[2] F. M. Fazal and S. M. Block, *Optical tweezers study life under tension,* Nat Phot. **5**, 318 (2011).

[3] D. G. Grier, *A revolution in optical manipulation,* Nat Phot. **424**, 810 (2003).

[4] O. M. Maragò, P. H. Jones, P. G. Gucciardi, G. Volpe, and A. C. Ferrari, *Optical trapping and manipulation of nanostructures.* Nat. Nanotechnol. **8**, 807 (2013).

[5] J. R. Moffitt, Y. R. Chemla, S. B. Smith, and C. Bustamante, *Recent advances in optical tweezers.* Annu. Rev. Biochem. **77**, 205 (2008).

[6] T. L. Min, P. J. Mears, L. M. Chubiz, C. V. Rao, I. Golding, and Y. R. Chemla, *High-resolution, long-term characterization of bacterial motility using optical tweezers,* Nat Meth **6**, 831 (2009).

[7] M.-C. Zhong, X.-B. Wei, J.-H. Zhou, Z.-Q. Wang, and Y.-M. Li, *Trapping red blood cells in living animals using optical tweezers.* Nat. Commun. **4**, 1768 (2013).

[8] M. T. Woodside, P. C. Anthony, W. M. Behnke-Parks, K. Larizadeh, D. Herschlag, and S. M. Block, *Direct measurement of the full, sequence-dependent folding landscape of a nucleic acid.* Science **314**, 1001 (2006).

[9] J. Liphardt, B. Onoa, S. B. Smith, I. Tinoco, and C. Bustamante, *Reversible unfolding of single RNA molecules by mechanical force*. Science **292**, 733 (2001).

[10] C. Cecconi, E. a. Shank, C. Bustamante, and S. Marqusee, *Direct observation of the three-state folding of a single protein molecule*. Science **309**, 2057 (2005).

[11] A. Mashaghi, G. Kramer, P. Bechtluft, B. Zachmann-Brand, A. J. M. Driessen, B. Bukau, and S. J. Tans, *Reshaping of the conformational search of a protein by the chaperone trigger factor*. Nature **500**, 98 (2013).

[12] K. Svoboda, C. F. Schmidt, B. J. Schnapp, and S. M. Block, *Direct observation of kinesin stepping by optical trapping interferometry*, Nature **365**, 721 (1993).

[13] R. A. Maillard, G. Chistol, M. Sen, M. Righini, J. Tan, C. M. Kaiser, C. Hodges, A. Martin, and C. Bustamante, *ClpX(P) generates mechanical force to unfold and translocate its protein substrates*. Cell **145**, 459 (2011).

[14] M.-E. Aubin-Tam, A. O. Olivares, R. T. Sauer, T. a. Baker, and M. J. Lang, *Single-molecule protein unfolding and translocation by an ATP-fueled proteolytic machine*. Cell **145**, 257 (2011).

[15] M. D. Wang, H. Yin, R. Landick, J. Gelles, and S. M. Block, *Stretching DNA with optical tweezers*. Biophys. J. **72**, 1335 (1997).

[16] S. B. Smith, Y. Cui, and C. Bustamante, *Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules*. Science **271**, 795 (1996).

[17] T. Perkins, D. Smith, and S. Chu, *Direct observation of tube-like motion of a single polymer chain*, Science **264**, 819 (1994).

[18] G. A. King, P. Gross, U. Bockelmann, M. Modesti, G. J. Wuite, and E. J. Peterman, *Revealing the competition between peeled ssDNA, melting bubbles, and S-DNA during DNA overstretching using fluorescence microscopy*, Proc Natl Acad Sci U S A **110**, 3859 (2013).

[19] I. K. Piechocka, R. G. Bacabac, M. Potters, F. C. Mackintosh, and G. H. Koenderink, *Structural hierarchy governs fibrin gel mechanics*. Biophys. J. **98**, 2281 (2010).

[20] H. Zhang and K.-K. Liu, *Optical tweezers for single cells*, J. R. Soc. Interface **5**, 671 (2008).

[21] H. G. Garcia, J. Kondev, N. Orme, J. A. Theriot, and R. Phillips, *Thermodynamics of biological processes*. Methods in enzymology **492**, 27 (2011).

[22] R. M. Daniel and M. J. Danson, *A new understanding of how temperature affects the catalytic activity of enzymes.* Trends Biochem. Sci. **35**, 584 (2010).

[23] M. C. Williams, J. R. Wenner, I. Rouzina, and V. a. Bloomfield, *Entropy and heat capacity of DNA melting from temperature dependence of single molecule stretching.* Biophys. J. **80**, 1932 (2001).

[24] S. Geggier, A. Kotlyar, and A. Vologodskii, *Temperature dependence of DNA persistence length.* Nucleic Acids Res. **39**, 1419 (2011).

[25] E. J. Guinn, W. S. Kontur, O. V. Tsodikov, I. Shkel, and M. T. Record, *Probing the protein-folding mechanism using denaturant and temperature effects on rate constants.* PNAS **110**, 16784 (2013).

[26] G. Somero, *Proteins and temperature,* Annu. Rev. Physiol. (1995).

[27] D. Nettels and S. Müller-Späth, *Single-molecule spectroscopy of the temperature-induced collapse of unfolded proteins,* Proc. … **106**, 20740 (2009).

[28] R. Perozzo, G. Folkers, and L. Scapozza, *Thermodynamics of Protein-Ligand Interactions: History, Presence, and Future Aspects,* J. Recept. Signal Transduct. **24**, 1 (2004).

[29] K. Datta, A. J. Wowor, A. J. Richard, and V. J. LiCata, *Temperature dependence and thermodynamics of Klenow polymerase binding to primed-template DNA.* Biophys. J. **90**, 1739 (2006).

[30] S. M. Leung, G. Senisterra, K. P. Ritchie, S. E. Sadis, J. R. Lepock, and L. E. Hightower, *Thermal activation of the bovine Hsc70 molecular chaperone at physiological temperatures: physical evidence of a molecular thermometer.* Cell stress & chaperones **1**, 78 (1996).

[31] M. Haslbeck, S. Walke, T. Stromer, M. Ehrnsperger, H. E. White, S. Chen, H. R. Saibil, and J. Buchner, *Hsp26: a temperature-regulated chaperone.* EMBO J. **18**, 6744 (1999).

[32] V. Zimarino and C. Wu, *Induction of sequence-specific binding of Drosophila heat shock activator protein without protein synthesis,* Nature **327**, 727 (1987).

[33] H. Mao, J. R. Arias-Gonzalez, S. B. Smith, I. Tinoco, and C. Bustamante, *Temperature control methods in a laser tweezers system.* Biophys. J. **89**, 1308 (2005).

[34] D. Braun and A. Libchaber, *Trapping of DNA by thermophoretic depletion and convection.* Phys. Rev. Lett. **89**, 188103 (2002).

[35] C. Bustamante, Y. R. Chemla, and J. R. Moffitt, *High-Resolution Dual-Trap Optical Tweezers with Differential Detection: Managing Environmental Noise,* Cold Spring Harb. Protoc. **2009**, pdb.ip72 (2009).

[36] M. Mahamdeh and E. Schaffer, *Optical tweezers with millikelvin precision of temperature-controlled objectives and base-pair resolution,* Opt Express **17**, 17190 (2009).

[37] J. R. Moffitt, Y. R. Chemla, D. Izhaky, and C. Bustamante, *Differential detection of dual traps improves the spatial resolution of optical tweezers.* Proc. Natl. Acad. Sci. U. S. A. **103**, 9006 (2006).

[38] S. Bezrukavnikov, A. Mashaghi, R. J. van Wijk, C. Gu, L. J. Yang, Y. Q. Gao, and S. J. Tans, *Trehalose facilitates DNA melting: a single-molecule optical tweezers study.* Soft matter **10**, 7269 (2014).

[39] M. Manghi, N. Destainville, and J. Palmeri, *Mesoscopic models for DNA stretching under force: New results and comparison with experiments.* Eur. Phys. J. E. Soft Matter **35**, 110 (2012).

[40] P. Cluzel, A. Lebrun, and C. Heller, *DNA: an extensible molecule,* Science **271**, 792 (1996).

[41] X. Zhang, H. Chen, S. Le, I. Rouzina, P. S. Doyle, and J. Yan, *Revealing the competition between peeled ssDNA, melting bubbles, and S-DNA during DNA overstretching by single-molecule calorimetry.* Proc. Natl. Acad. Sci. U. S. A. **110**, 3865 (2013).

[42] P. Gross, N. Laurens, L. B. Oddershede, U. Bockelmann, E. J. G. Peterman, and G. J. L. Wuite, *Quantifying how DNA stretches, melts and changes twist under tension,* Nat. Phys. **7**, 731 (2011).

[43] M. C. Williams, J. R. Wenner, I. Rouzina, and V. A. Bloomfield, *Effect of pH on the overstretching transition of double-stranded DNA: evidence of force-induced DNA melting,* Biophys J **80**, 874 (2001).

[44] K. F. Palmer and D. Williams, *Optical properties of water in the near infrared,* J. Opt. Soc. Am. **64**, 1107 (1974).

[45] Y. M. Efimova, S. Haemers, B. Wierczinski, W. Norde, and A. A. Van Well, *Stability of globular proteins in H2O and D2O,* Biopolymers **85**, 264 (2007).

[46] E. a. Abbondanzieri, W. J. Greenleaf, J. W. Shaevitz, R. Landick, and S. M. Block, *Direct observation of base-pair stepping by RNA polymerase.* Nature **438**, 460 (2005).

4

Functional plasticity in Hsp70 enabled by diverse modes of client binding

Hsp70 system is a central hub of chaperone activity in all domains of life. Hsp70 performs a plethora of tasks, including folding assistance, protection against aggregation, protein trafficking, and enzyme activity regulation [1–5], and interacts with non-folded chains, as well as near-native, misfolded, and aggregated proteins [6–11]. It is believed that Hsp70 achieves its many physiological roles by binding peptide segments that extend from these different protein conformers within a groove that can be covered by an ATP-driven helical lid [12–16]. However, it has been difficult to test directly. Moreover, recent indications of diverse lid conformations in Hsp70-substrate complexes raise the possibility of additional interaction mechanisms [16–19]. Addressing these issues is technically challenging, given the conformational dynamics of both chaperone and client, the transient nature of their interaction, and the involvement of co-chaperones and the ATP hydrolysis cycle [20]. Using optical tweezers, we show that bacterial Hsp70 (DnaK) binds not only extended peptide segments, but also folded protein structures. Our data reveal that the Hsp70 groove and lid have distinct functions: truncating the lid abolishes structure binding, while mutating the groove most severely impairs peptide-binding. This diversity of binding modes has important consequences: 1) Hsp70 not only can destabilize but also stabilize folded structures, in a nucleotide-regulated manner. 2) Like Hsp90 and GroEL, Hsp70 can affect late stages of folding. 3) Hsp70 can suppress aggregation by protecting partially folded structures as well as unfolded protein chains. Overall, the findings indicate an extension of the Hsp70 canonical model that potentially affects all the physiological roles of the Hsp70 system.

4.1. Methods

4.1.1. Expression and purification of MBP and 4MBP

N-terminally biotinylated MBP and 4MBP C-terminally fused with 4 Myc tag sequences were produced in *E. coli* as hybrid proteins consisting of an N-terminal Ulp1-cleavable N-terminal His10-SUMO sequence followed by an AviTag sequence (Avidity, LCC, Aurora, Colorado, USA), facilitating in vivo biotinylation and four consecutive C-terminal Myc-tag sequences. Proteins were purified from E. coli BL21 cells harboring pBirAcm encoding the biotin ligase (Avidity, LCC, Aurora, Colorado, USA). For over-expression cells over-night cultures were diluted 1:100 in fresh LB medium supplemented with 20 mg/l Biotin, 20 mg/l Kanamycin, 10 mg/l Chloramphenicol, 0.2% glucose and incubated under vigorous shaking at 30°C. Expression was induced at OD600= 0.6 by addition of 1 mM IPTG for 3 h. Cells were chilled, harvested by centrifugation, flash-frozen in liquid nitrogen and stored at -70°C. Cell pellets were resuspended in ice-cold buffer A (20 mM Tris-HCl pH 7.5, 0.2 M NaCl, 1% Triton X-100, 1 mM PMSF) and lysed using a French Pressure Cell. The lysate was cleared from cell debris by centrifugation at 35.000 g for 30 min and incubated with Ni-IDA matrix (Protino; Macherey-Nagel, Düren, Germany) for 30 min at 4°C. The matrix was washed extensively with buffer A and bound hybrid proteins were eluted in buffer A containing 250 mM imidazole. The eluate was supplemented with His6-Ulp1 protease and dialyzed overnight at 4°C in buffer D (20 mM Tris-HCl pH 7.5, 0.2 M NaCl). Following dialysis coupled with Ulp1 digestion, His6-Ulp1 protease and the His10-SUMO fragment were removed by incubation with Ni-IDA matrix. The unbound MBP or 4 MBP that remained in the unbound fraction was then loaded on Amylose resin (New England Biolabs) previously equilibrated in buffer D, washed with cold buffer D and bound proteins were eluted in buffer D supplemented with 20 mM maltose. Elution fractions were dialyzed tree times for 2 hours at 4°C in 100-fold excess volume of buffer S (20 mM Tris-HCl pH 7.5, 0.2 M NaCl, 1 mM EDTA). 4 MBP purifications in addition were subjected to size-exclusion chromatography using a HiLoad 16/600 Superdex prep grade column. Purified proteins were concentrated using Vivaspin centrifugal concentrators, aliquoted, flash frozen in liquid nitrogen and stored at -70°C.

Biotinylated proteins were also produced as follows. Purification of MBP and 4MBP containing one N-terminal cystein was performed as described [21]. For biotinylation the purified proteins were dialyzed in buffer B (20 mM Tris-HCl pH 7.5, 0.2 M NaCl, 1 mM EDTA) and incubated on ice for 15 min in buffer B containing 5 mM TCEP to reduce disulfide bonds. A ten-fold molar excess of Maleimide-PEG11-Biotin (Thermo Scientific) dissolved in DMSO was added and biotinylation was performed for 2 hours at 25 °C. Proteins were concentrated and subjected to size-exclusion chromatography using a HiLoad 16/600 Superdex prep grade column. Purified fractions were analyzed by SDS-PAGE, pooled, aliquoted, flash frozen in liquid nitrogen and stored at -70°C.

4.1.2. WT and mutant chaperones expression and purification

All proteins were produced in *E. coli* from expression plasmids. Wild-type DnaK, DnaK(2-538) (lid-less), and DnaK-V436F (groove-less) were produced in a ΔdnaK strain and purified by ammonium sulfate precipitation, anion exchange chromatography (DEAE-sepharose), ATP-agarose affinity chromatography, gel filtration and strong anion exchange chromatography (Resource Q) as described earlier[22]. DnaJ was purified by chromatography on cation exchange and hydroxyapatite material. GrpE was expressed in a ΔdnaK strain and purified by chromatography on DEAE-Sepharose, hydroxyapatite and Superdex 200 as described [23].

4.1.3. Luciferase activity protection assay

Luciferase (80 nM in 40 mM $^{HEPES}/_{KOH}$ pH 7.5, 50 mM potassium acetate, 5 mM magnesium acetate, 2 mM DTT) was incubated in the absence or presence of ATP (2 mM), ADP (3.2 μM), DnaK wild-type or mutant proteins (800 nM), DnaJ (160 nM) and GrpE (400 nM) at 37°C. Luciferase activity was determined by diluting 1 μl aliquots into 124 μl assay buffer (100 mM potassium phosphate (pH 7.8), 25 mM glycylglycine (pH 7.4), 100 mM potassium acetate, 15 mM magnesium acetate, 5 mM ATP), mixing with 125 μl luciferin (80 μM in assay buffer) and monitoring bioluminescence for 5 s without delay using a Biolumat (Berthold).

4.1.4. Optical tweezers assay

Anti-digoxigenin (anti-Dig) coated beads (DIGP-20-2, diameter 2.1 μm) were purchased from Spherotech and stored at 4°C until use. 2553 base pairs DNA handles were prepared and functionalized with digoxigenin (Dig) and biotin at 5'-ends of both strands as described [21]. DNA-coated microspheres were made by mixing ~70 ng of biotin–dsDNA–Dig and 4 μl Dig-coated beads in 10 ml HMK buffer (50 mM HEPES, pH 7.5, 5 mM MgCl2, 100 mM KCl). After a 30 minute incubation on a rotary mixer (4°C), 1 μl of neutravidin solution (1% w/v) was added to the mixture and incubated at 4°C for 10 more minutes. Next, the unbound neutravidin was washed away by centrifuging, and the beads were dissolved in 400 μl HMK buffer for use in optical tweezers experiments. Carboxylated polystyrene beads (CP-20-10, diameter 2.1 μm, Spherotech) were covalently attached to anti-Myc antibody (Roche Diagnostics) via carbodiimide reaction (Poly-Link Protein Coupling Kit, Polysciences Inc.). Briefly, the beads were washed and then mixed with freshly prepared 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide and the antibody, and mixture was incubated for 3 hours. MBP-coated beads were made by mixing 5 μl of 50 μM solution of either 1MBP or 4MBP and 2 μl Myc-coated beads in 10 μl HMK buffer and incubating the mixture for 30 minutes on a rotary mixer at 4°C. Subsequently, the beads were dissolved in 400 μl HMK buffer for use in optical tweezers experiments.

Optical tweezers assays, buffer contents: Fig. 4.1b,c: no chaperones; Fig. 4.1e,f: 100 nM DnaK, 100 nM DnaJ, 50 nM GrpE, 1 mM ATP; Fig. 4.2b,c: no chaperones; Fig. 4.2d-g: 1 μM DnaK, 1 mM ADP loading buffer; Fig. 4.2h,i: 100 nM DnaK, 1 mM ATP; Fig. 4.3a,b: 1 μM DnaK lid-less, 1 mM ADP loading buffer;

Fig. 4.3c,d: 1 µM DnaK groove-less, 1 mM ADP loading buffer; Extended Data Fig. 4.5c-f, k: 1 µM DnaK, 1 mM ADP loading buffer; Extended Data Fig. 4.5b: 100 nM DnaK, 1 mM ADP loading buffer; Extended Data Fig. 4.5g-h: 1 µM DnaK T199A, 1 mM ATP; Extended Data Fig. 4.5i-j: 1 µM DnaK, 1 mM ADP purified by HPLC; Extended Data Fig. 4.7b: 100 nM DnaK, 1 mM ATP; Extended Data Fig. 4.7c: 100 nM DnaK lid-less, 1 mM ATP; Extended Data Fig. 4.7d: 100 nM DnaK groove-less, 1 mM ATP; Extended Data Fig. 4.8b: 1 µM DnaJ; Extended Data Fig. 4.8c: 100 nM DnaJ; Extended Data Fig. 4.8d: 1 µM GrpE. Number of events (high force): N = 38 (Fig. 4.1b,c), 55 (Fig. 4.1e,f), 52 (Extended Data Fig. 4.7b), 131 (Extended Data Fig. 4.7c), 96 (Extended Data Fig. 4.7d). Number of events (low force): N = 50 (Fig. 4.2b,c), 35 (Fig. 4.2d,e), 15 (Fig. 4.2f,g), 24 (Fig. 4.2h,i), 15 (Fig. 4.3a,b), 21 (Fig. 4.3c,d), 10 (Extended Data Fig. 4.5c,d), 11 (Extended Data Fig. 4.5e,f), 18 (Extended Data Fig. 4.5g,h), 4 (Extended Data Fig. 4.8b), 9 (Extended Data Fig. 4.8c), 10 (Extended Data Fig. 4.8d). ADP loading buffer was prepared by dissolving ADP sodium salt (A2754, Sigma Aldrich) in HMK buffer. The ATP fraction in this buffer was ~5%, as measured by HPLC.

Stretching experiments were performed at room temperature using a custom made optical tweezers setup [21]. MBP-coated bead was trapped in the optical trap and then transferred to a micropipette tip, and, subsequently, DNA-coated bead was trapped. Next, two beads were brought in close contact, allowing a tether (DNA-MBP linker) between the beads to form. When the micropipette is moved away from the trap, the tether experiences a stretching force, leading to its extension and deflection of the trapped bead from the trap center. The deflection is monitored by collecting the transmitted laser light and directing it to quadrant photodiode where it is recorded at 50 Hz. The data is filtered with 5th order Butterworth filter at 20 Hz prior to saving. Trap stiffness and sensitivity were measured to be 169 ± 24 pN/µm and 2.74 ± 0.24 pN/µm correspondingly. A nanopositioning piezo stage is used to move the flow chamber and micropipette at a speed of 50 nm/s that corresponds to a pulling rate on the tethered MBP construct of ~5 pN/s at unfolding.

4.2. Results

To study the DnaK system at the single-molecule level, we first used an engineered substrate that is prone to misfolding *in vitro:* four Maltose Binding Proteins linked head-to-tail (4MBP, Fig. 4.1a). Mechanical stretching of 4MBP in the absence of chaperone [21, 24] first produced a gradual unfolding transition (Fig. 4.1b, N ⇒ 4), followed by the distinct unfolding events of the four remaining core structures (Fig. 4.1b, 4 ⇒ 3 ⇒ 2 ⇒ 1 ⇒ U). Relaxation to low forces and waiting for 5 s provided an opportunity to refold. However, while subsequent stretching showed a compact structure had formed, these structures typically failed to unfold (termed tight misfold, Fig. 4.1b) or unfolded in steps larger than one MBP core (termed weak misfold). Isolated 4MBP thus formed misfolded mini-aggregates consisting of more than one MBP repeat (Fig. 4.1c). Next, we performed these experiments in a buffer containing ATP and the DnaK system (DnaK, DnaJ, GrpE, Fig. 4.1d, Extended Data Fig. 4.1, 4.2, 4.3, 4.4). Unfold-

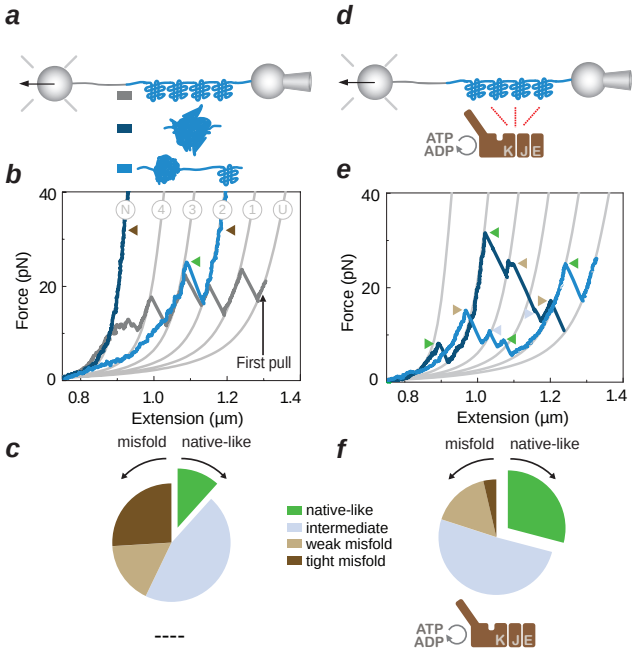


Figure 4.1: **DnaK chaperone system simultaneously suppresses aggregation and promotes refolding. a,** Schematic diagram of the setup. Protein construct of 4 MBP monomers (4MBP, blue) is attached to a DNA linker (grey) and tethered between two beads [21]. **b,** 4MBP in the absence of chaperones. Thick grey line: First stretching curve showing native 4MBP unfolding. Blue: subsequent stretching curves corresponding to different types of misfolded states. Triangles: types of events or states (see panel c and f and main text). Thin grey lines: theoretical compliance of DNA-protein construct in states ranging from native (N) to fully unfolded (U). Numbers indicate the number of folded core structures. **c,** Corresponding event fractions ($N = 38$). Intermediate: unfolding events with unfolded length shorter than one MBP core. **d,** Schematic diagram. **e,** Stretching curves of 4MBP in the presence of the DnaK system and ATP. **f,** Corresponding event fractions ($N = 55$).

ing and relaxation now rarely produced tight misfolds, though weak misfolds still formed (Fig. 4.1e, f). Moreover, the data showed more unfolding steps corresponding to one MBP core ($p < 0.05$, Fig. 4.1e, green arrows), indicating the native-like refolding of core structures. These results show that the DnaK system can promote the correct refolding of a misfolding- and aggregation-prone protein, consistent with previous bulk studies of Hsp70 [2, 5, 14]. However, they do not reveal how the various parts of the DnaK system affect client conformations during the different stages of folding.

To understand the underlying mechanisms we focused on the simpler case of single MBP to preclude aggregation and omitted the co-chaperones (Fig. 4.2a). Different protocols were followed. First, we unfolded MBP, which occurred in a two-step manner via the core state [21, 24] (Fig. 4.2b), and then performed cycles of relaxation, waiting at 0 pN for 5 s, and stretching again, all in the absence of

chaperones. Stretching indicated the chain either had remained unfolded, or had refolded to a compact structure that unfolded in the same two-step manner (Fig. 4.2b). The refolding probability p_r was about 0.85 (Fig. 4.2c), which increased for longer waiting periods [21, 24]. Next, cycling was paused at 0 pN, and a buffer with DnaK was flown in and DnaK was allowed to interact. We used a 1mM ADP buffer with small amounts (~5%) of ATP (termed ADP loading buffer) in order to stall the ATP cycle and to promote the loading of DnaK onto the substrate. Cycling was subsequently resumed for the same molecule. Surprisingly, the compact protein structures now typically failed to unfold for multiple stretching-relaxation cycles until the tether broke. The structures sustained high forces ($F_u > 40$ pN) well above the native unfolding force ($\langle F_u \rangle = 22 \pm 5$ pN), and sometimes even exceeded the effective maximum for our assay ($F_u = 65$ pN) [25] (Fig. 4.2d, e, j, Extended Data Fig. 4.5b). Notably, this scenario is unlike the 4MBP aggregation (Fig. 4.1b), as aggregation partners are not available here. Instead, the data suggested that in the presence of ADP, DnaK bound protein structures and stabilized them against forced unfolding.

Previous work had suggested DnaK rather stabilizes the unfolded state of substrates (Extended Data Fig. 4.2, 4.3, 4.4) [26]. To try to promote this binding mode, we modified the protocol. MBP was unfolded, kept at a moderate force to prevent refolding, and incubated for minutes with DnaK and ADP loading buffer. In subsequent relaxation-stretching cycles, MBP now remained unfolded (p_r ~0.05; Fig. 4.2f and 4.2g), indicating that DnaK had bound and stabilized the unfolded chain. Next, we reduced the DnaK incubation time from minutes to seconds by first unfolding native MBP, and then immediately subjecting it to relaxation and stretching (Extended Data Fig. 4.5c,d). Note that the first unfolding curve was similar as without chaperone ($\langle F_u \rangle = 23 \pm 4$ pN), which indicated a lack of interaction between DnaK and native MBP. Importantly, unfolded chains now indeed refolded in subsequent cycles. Moreover, the refolded structures typically unfolded at high forces ($F_u > 40$ pN), until they were locked into a compact state that could not be unfolded (Extended Data Fig. 4.5d). Consistently, such a compact and locked state was achieved in earlier cycles for longer waiting times at 0 pN (Extended Data Fig. 4.5e,f). These data suggested a kinetic partitioning: unfolded chains can be bound by DnaK [27], which blocks their refolding, unless the chains fold first (folding time is about 1 s for MBP) [21], and are then stabilized by DnaK.

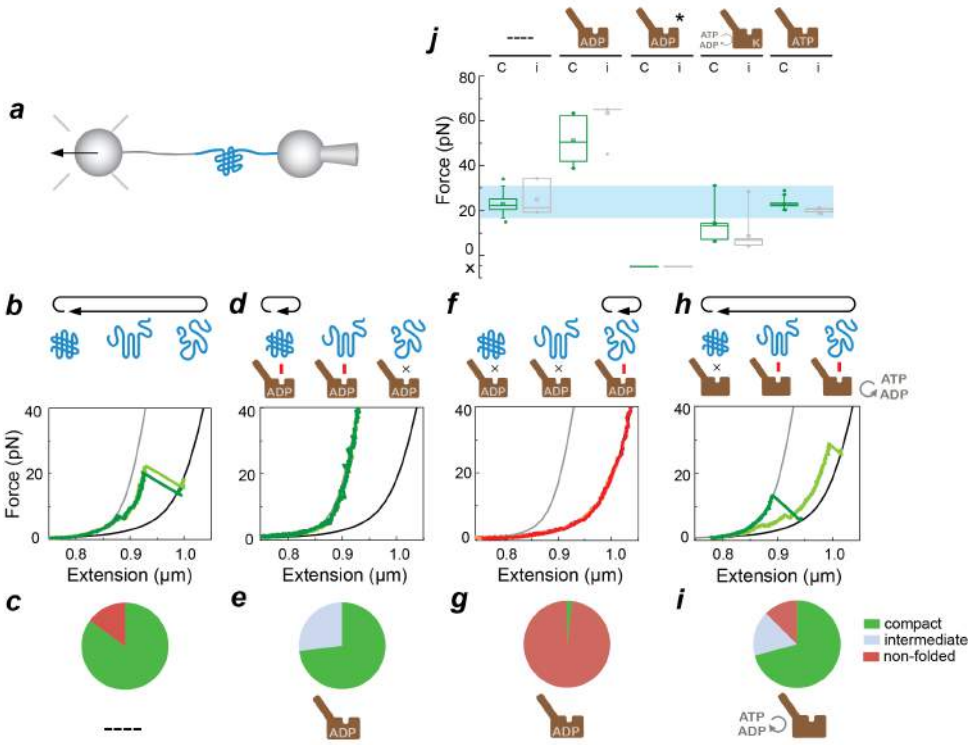


Figure 4.2: **DnaK with ADP binds folded structures and stabilizes them against forced unfolding. a,** Schematic diagram for single MBP experiments. **b-i,** Stretching experiments on: MBP in isolation **(b-c;** $N = 50$), stretched and refolded MBP that was incubated with DnaK and ADP loading buffer **(d-e;** $N = 35$), stretched and kept unfolded MBP that was incubated with DnaK and ADP loading buffer **(f-g;** $N = 15$), MBP with DnaK and ATP **(h-i;** $N = 24$). Top: diagrams of compact, (various) partially folded, and fully unfolded states, and observed transitions between them. Red bar indicates chaperone-protein interaction, cross no interaction. Middle: MBP stretching curves. Grey and black lines: theoretical compliance of folded and unfolded MBP. Bottom: corresponding refolding probability during the 5 s. time window at 0 pN, to compact structure or structure of intermediate length. **j,** Unfolding forces, corresponding to data in panels **b, d, f,** and **h.** Last column is data for DnaK mutant T199A, which traps DnaK in the ATP state in an ATP buffer ($N = 18$). Star indicates MBP was kept unfolded when incubated (see panel **f**), 'c' the core state, 'i' a state with intermediate measured length; in between core state and unfolded state. Cross-label at force axis: folded structures were not observed. Blue band: the force range (between first and third quartiles) of MBP core unfolding without chaperones. Whiskers indicate the 90% and 10% extrema.

Interestingly, the observed stabilization by DnaK could occur also during the ATP cycle. In the presence of DnaK and 1 mM ATP, we found that refolded structures unfolded frequently via partially folded structures smaller than the core state (44% of the traces, mean state lifetime during stretching $\langle\tau\rangle = 1.1 \pm 0.6$ s, Fig. 4.2h and i). For example, the light green curve in Fig. 4.2h displayed a partially folded structure beyond 5 pN that was stable for τ ~4 s, and then unfolded fully at 30 pN. This observation suggested the folded structure was stabilized by DnaK, and then unfolded when DnaK was released, as it progressed through the ATP cycle. Consistently, the occurrence of transiently stable partial folds was higher than without chaperones ($p < 0.01$, 8% of the traces, $\langle\tau\rangle$ ~0.5 s), but lower than with DnaK and ADP loading buffer ($p < 0.01$, all traces, τ up to 150 s during multiple stretching-relaxation cycles, Fig. 4.2d, Extended Data Fig. 4.5).). In addition, we found that stabilization of folded structures by DnaK was not observed for a DnaK mutant (T199A) that is trapped in the open ATP state (Fig. 4.2j, Extended Data Fig. 4.5g,h), but was observed a buffer with HPLC purified ADP and no ATP (Fig. 4.2j, Extended Data Fig. 4.5i,j), and for native MBP that was only partially unfolded (Extended Data Fig. 4.5j). Overall, these results suggested that DnaK can stabilize folded structures in a binding reaction modulated by the ATP cycle.

Which structural elements of the Hsp70 substrate-binding domain are conducive to stabilizing folded structures? The helical lid may play a role, as implied by the observation that ATP affects both folded-structure binding (Fig. 4.2d, h) and lid conformation [17], but this evidence is indirect. To address this issue more directly, we investigated two DnaK variants (Extended Data Fig. 4.6): one with a truncated lid (DnaK 2-538; termed lid-truncated), and one carrying a mutation that introduces a bulky side group within the binding groove (DnaK-V436F; termed groove-mutated) [27]. With lid-truncated DnaK and ADP loading buffer (Fig. 4.3a,b), relaxed unfolded chains refolded less readily ($p_r \sim 0.47$) than without chaperone ($p < 0.01, p_r \sim 0.85$). Hence, the lid is not required to bind extended polypeptide segments and suppress refolding. The lid does contribute to the efficiency of extended polypeptide binding, as indicated by further reduction in the refolding probability for WT DnaK ($p < 0.01, p_r \sim 0.05$, Fig. 4.2f, g). More importantly, refolded chains typically unfolded in native-like fashion (Fig. 4.3a) and lacked high-force stabilization seen for WT DnaK (Fig. 4.2d, 4.3e). The lid-truncation thus abolished the stabilization of folded structures. In contrast, with groove-mutated DnaK and ADP loading buffer (Fig. 4.3c and d), relaxed unfolded chains refolded frequently to a compact state ($p_r \sim 0.75$, Fig. 4.3d), consistent with impaired binding of extended peptides [27]. Refolded chains unfolded often via partial folds (52% of the traces), at forces of a few pN up to 65 pN (Fig. 4.3c and 4.3e). Thus, the groove mutation does not abolish the stabilization of folded structures. Overall, these data show that the DnaK lid is essential to binding folded structures, while the groove is essential to binding exposed polypeptides (Fig. 4.3f).

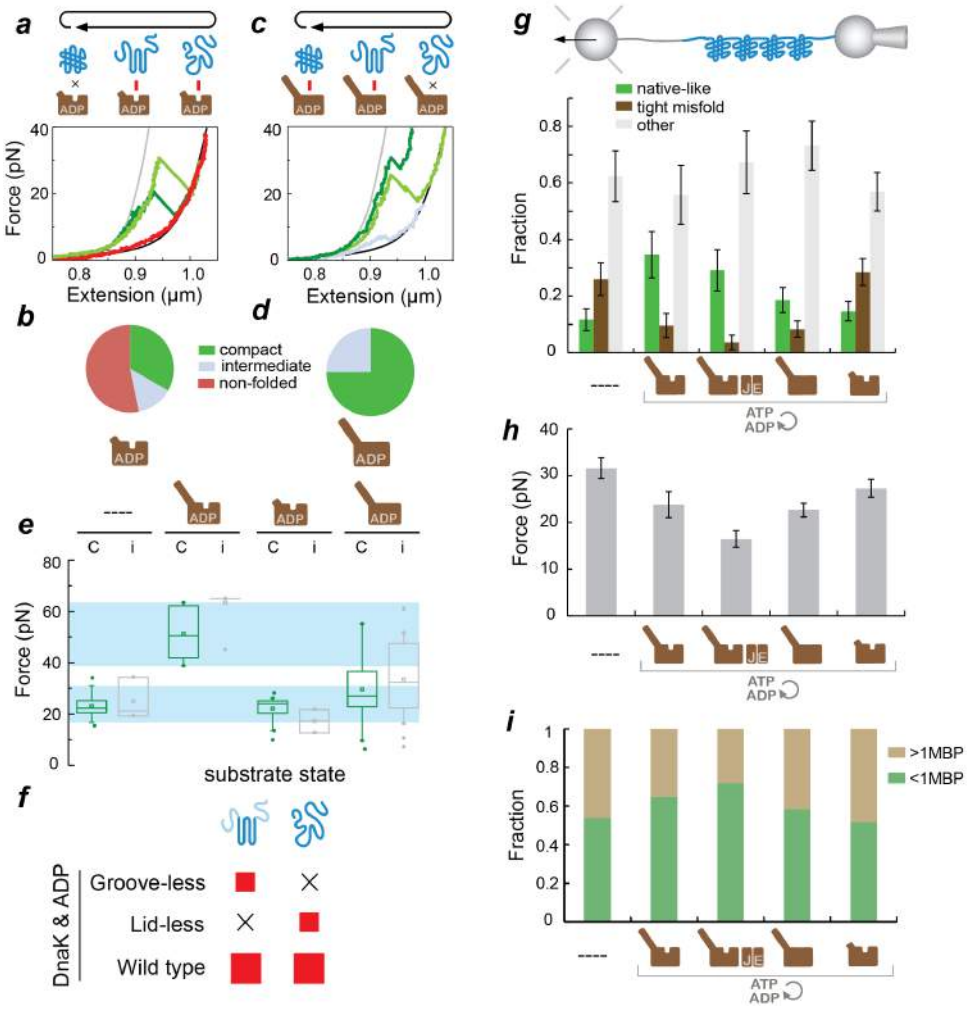


Figure 4.3: **DnaK lid binds folded structures and suppresses aggregation. a-d,** Stretching curves of MBP with ADP loading buffer, and lid-less DnaK (DnaK(2-538)) **(a-b;** $N = 15$) or groove-less DnaK (DnaK V436F) **(c-d;** $N = 21$). Top: diagrams of protein states and observed transitions between them. Middle: MBP stretching curves. Bottom: refolding probability during 5 s time window at 0 pN. **e,** Unfolding forces with lid-less and groove-less DnaK. As a reference, we indicate the unfolding force ranges of MBP in isolation and with DnaK and ADP. 'c' indicates the core state, 'i' a state with intermediate measured length, in between core state and unfolded state. **f,** Interaction of DnaK variants with folded and unfolded protein structures. Crosses denote the absence of interaction; red squares the presence of interaction, and their size qualitatively indicates the strength of the effect. **g,** Statistics for native-like refolding and tight misfolding in 4MBP with different DnaK variants. Indicated is the type of unfolding event, as a fraction of the total number of unfolding events. **h,** Unfolding force of non-native structures formed during 4MBP refolding.Whiskers indicate the 90% and 10% extrema.Whiskers indicate the 90% and 10% extrema. **i,** Fractions of structures formed during 4MBP refolding that unfold in steps smaller (green) and larger (brown) than one MBP core (92 nm).

We surmised that binding partially folded structures could be important to suppressing aggregation [9, 10, 28] (Fig. 4.1e and 4.1f). If so, groove-less DnaK should retain some of this capability, as it binds folded structures (Fig. 4.3c). We found that compared to without chaperones, groove-less DnaK and 1 mM ATP indeed decreased tight aggregation of 4MBP ($p < 0.01$). The aggregated structures that did form, unfolded at lower forces (Fig. 4.3h,$p < 0.01$). These findings are consistent with ideas that DnaK can destabilize misfolded structures [4]. The results may appear paradoxical: how can DnaK both stabilize (Fig. 4.2d) and destabilize folded structures (Fig. 4.3h)? The ATP-driven DnaK cycle offers a possible rationale: DnaK-mediated stabilization of small partial folds is only transient, while it can at the same time suppress the formation of larger (mis)folds that are more permanently stable (Fig. 4a). In this scenario, 4MBP should refold into smaller structures with the full DnaK system, and hence unfold in smaller steps, which we indeed find to be the case (Fig. 4.3i, $p < 0.05$). Note that DnaK may also be able to recognize the difference between native and non-native partial folds, and directly destabilize the latter [4]. Overall, the data indicate that the formation and stability of aggregates can be suppressed by the transient stabilization of (partially) folded structures by DnaK.

Other components of the DnaK system may also contribute to limiting aggregation. With the full DnaK system present, tight aggregation was less frequent than for DnaK and ATP, though the difference was not large ($p < 0.01$, Fig. 4.3g). As was expected given its nucleotide exchange role [7], the co-chaperone GrpE did not display any substrate interactions by itself (Extended Data Fig. 4.8e). In contrast, DnaJ alone completely abolished aggregation (Extended Data Fig. 4.8b), but blocked refolding as well (Extended Data Fig. 4.8d). DnaJ clearly was not such a dominant polypeptide holdase within the full DnaK system (see Fig. 4.1e and 4.1f). Hence, these findings support previous observations that ATP, GrpE and DnaK can promote DnaJ release [29].

Finally, the observed stabilization of protein structure led us to speculate that DnaK could stabilize protein function under heat stress [5]. To test this idea, we monitored the enzymatic activity of Luciferase in a bulk assay during a temperature increase to 37°C. In the absence of chaperones the activity decreased to less than 10% in 45 minutes, consistent with thermally induced unfolding (Extended Data Fig. 4.9). In the presence of DnaK and ADP, this decrease was substantially suppressed, whereas the ADP buffer alone did not affect the activity decrease. In the presence of ATP, DnaK was also found to limit the activity decrease more than ATP alone (Extended Data Fig. 4.9). Taken together, the data suggest that DnaK can stabilize protein structure and maintain enzymatic activity when ATP is limited.

4.3. Discussion

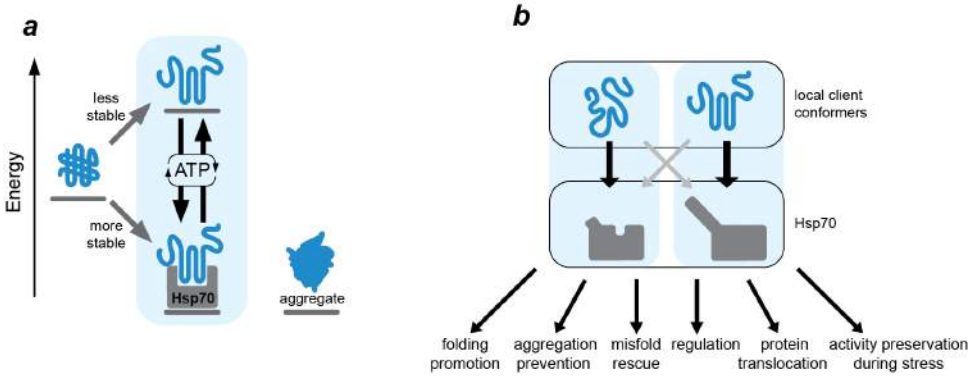


Figure 4.4: **Model of DnaK chaperone action. a,** Effect on substrate stability. DnaK binding can stabilize (partially) folded states, which are less stable than fully folded states, in an ATP-dependent manner. In turn, this DnaK-binding can suppress the formation of large and more permanently stable aggregates (right), and hence yield structures that are effectively less stable than such aggregates. Modulation of steps in the ATP cycle can alter the DnaK on- and off-rates, and hence the stabilizing effect of DnaK. For instance, high ADP levels and low ATP levels can trap substrates in compact stable states that resist unfolding, whereas conversely, high ATP levels can yield structures that unfold at lower forces than observed without chaperone. Alternatively, the steps in the ATP cycle may be modulated in a more specific manner by cofactors. **b,** The binding of DnaK to both unfolded and folded local conformers can generally affect most of its physiological roles (see main text).

From these experiments on DnaK, a picture emerges of an Hsp70 functional repertoire that is broader than previously assumed (Fig. 4.4b):

1. Hsp70 binds extended peptides [4, 16, 17, 28], but can also directly bind folded structures. The Hsp70 lid is essential for the latter, and could adopt open conformations that have been reported recently [16–18, 20].

2. In the canonical model, protein chains fold autonomously after Hsp70 release, and can subsequently be transferred to other chaperones acting in later stages, such as Hsp90 and GroEL [11, 15]. Our findings suggest Hsp70 can also act late in folding [19], and potentially not fully release protein substrates until near native.

3. Aggregation suppression by Hsp70 is achieved by shielding not only extended peptide segments [30], but also partially folded structures.

4. Hsp70 is known to destabilize (mis)folded protein structures [4]. Our data indicates Hsp70 can also play an opposing role, and stabilize folded protein structures. These contrasting effects can be controlled by the ATP cycle and the co-chaperones that influence it.

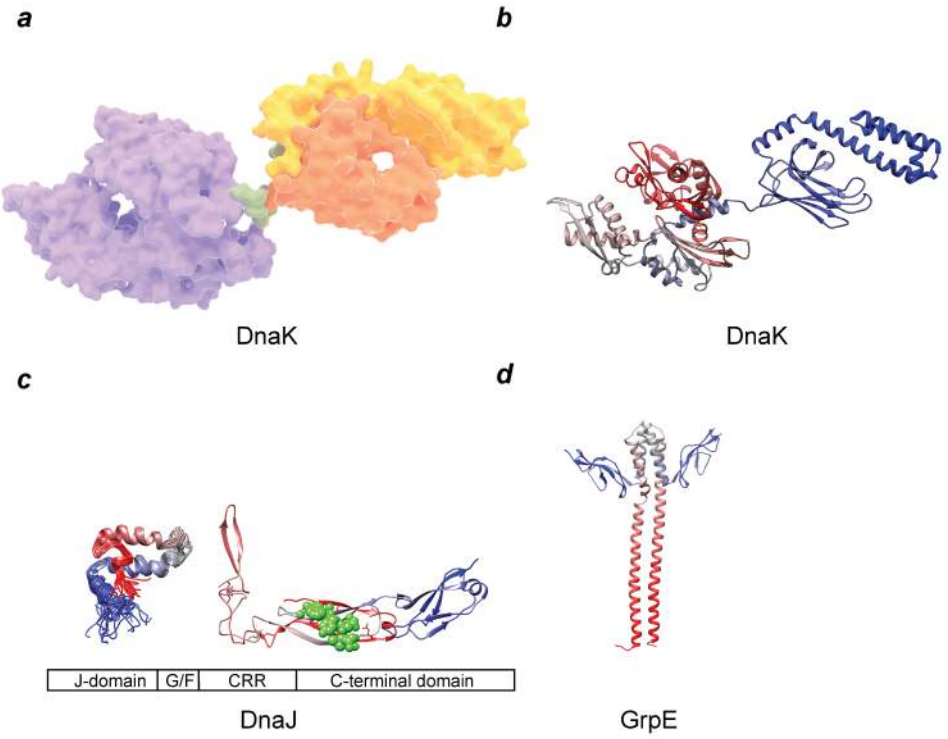
5. The findings suggest Hsp70 can preserve enzymatic functions during stress at low energy costs, which is in line with the reduced energy availability in episodes of stress.

These Hsp70 functions could have important consequences for our understanding of normal cellular physiology as well as pathophysiology. Hsp70 may transfer clients to other chaperones [9, 10] without having to (locally) unfold them, and contribute to an evolutionary capacitor role by stabilizing substrate proteins [31]. Co-factors such as HIP [32] that stabilize the ADP state could play a role in regulating client binding and stabilization. Finally, Hsp70 could control signaling pathways [7] by interacting directly with the folded states of sigma factors, kinases, and receptors, in addition to their unfolded segments [1, 3].

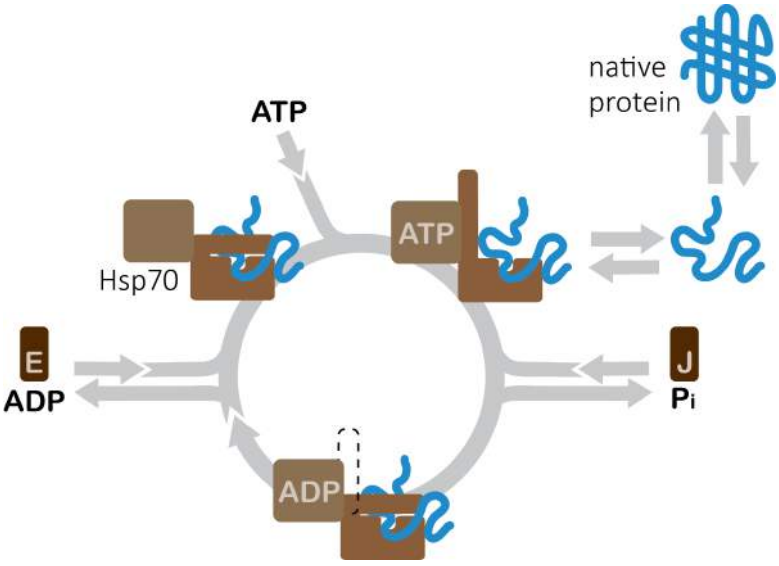
4.4. Acknowledgements

Work in the laboratory of S.J.T. is part of the research programme of the Stichting voor Fundamenteel Onderzoek der Materie (FOM), which is financially supported by the Nederlandse Organisatie voor Wetenschappelijke Onderzoek (NWO). We thank M. Avellaneda for help with preparing protein structure illustrations, and T. Shimizu, E. Garnett and F. Huber for critical reading of the manuscript.

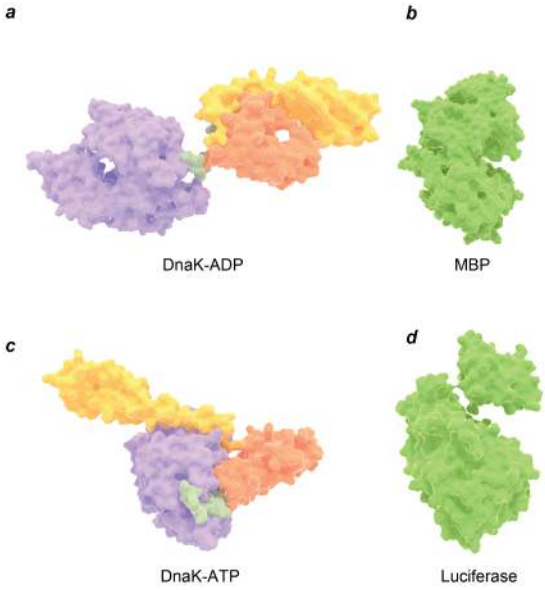
4.5. Extended data



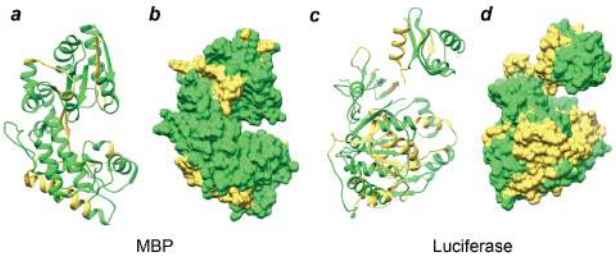
Extended Data Figure 4.1: **Domain architectures of DnaK, DnaJ and GrpE. a,** Structure of *E. coli* DnaK bound to peptide substrate and ADP nucleotide (PDB 2KHO, domain boundaries indicated in different colors: Nucleotide binding domain (purple, 1-383), hydrophobic linker (green, 384-396), Substrate binding domain - β-sheet subdomain with groove (orange, 397-502), Substrate binding domain – α-helical lid subdomain (yellow, 503-602) [33]. **b,** same structure in N-to-C terminal color gradient from red (N-terminus) over dark grey to blue (C-terminus). **c,** DnaJ domain structure containing J-domain, Glycine/Phenylalanine-rich domain (G/F), Cysteine-Rich Region/Zn finger (CRR) and C-terminal domain; partial NMR structures of J domain (PDB ID: 1XBL) and crystal structure comprising cysteine-rich domain (CRR) and peptide binding domain of the yeast homolog with co-crystalized peptide substrate highlighted in green spheres (PDB ID: 1NLT) [34, 35]. **d,** GrpE structure, (PDB ID : 1DKG) [36]



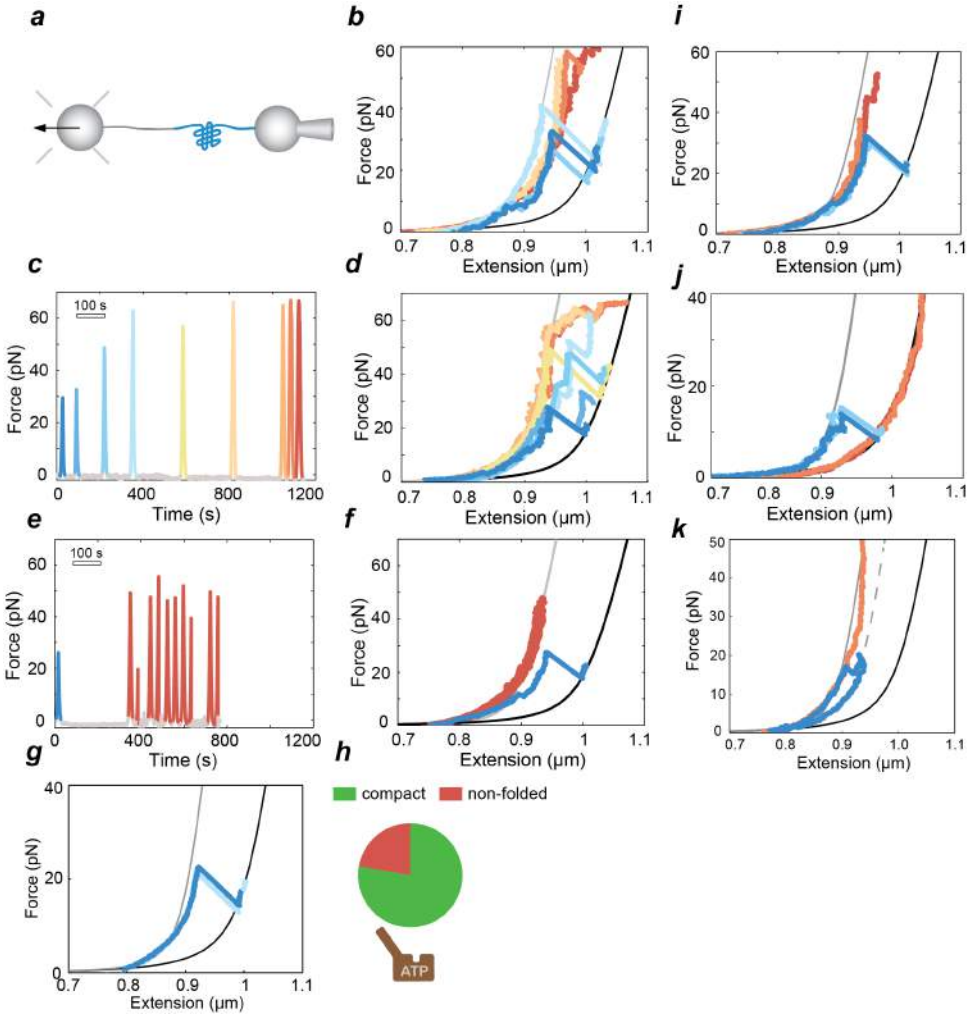
Extended Data Figure 4.2: **Canonical nucleotide-driven cycle of Hsp70-client interactions.** Hsp70 is known to interact with a large range of protein conformations from unfolded nascent chains over near-native proteins to aggregates and misfolded states. Non-native protein conformers are captured via short extended peptide segments by the substrate-binding groove of Hsp70. Hsp40 (DnaJ) then accelerates the ATP hydrolysis reaction together with bound substrate. Hsp70-ADP locks the substrate under the closed helical lid of the substrate-binding domain in the 'high-affinity' state. Nucleotide Exchange Factor (GrpE) subsequently accelerates the release of ADP and either spontaneous fluctuations of the lid or new ATP molecules facilitate the release of the substrate to regenerate Hsp70-ATP for another round of the cycle (adapted from ref. [37]).



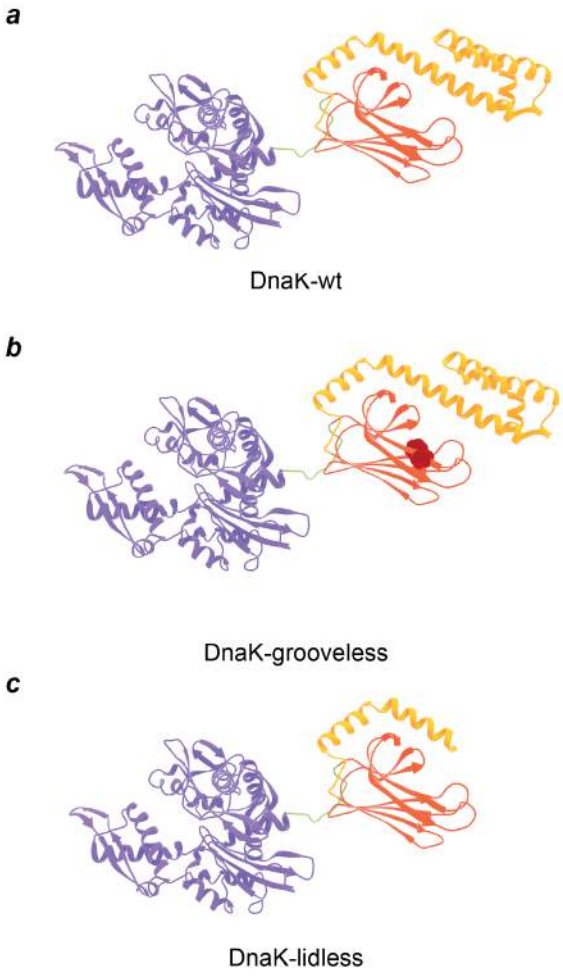
Extended Data Figure 4.3: **Structures of open and closed DnaK conformations, as well as MBP and luciferase. a,** DnaK in ADP state (PDB code: 2KHO). **b,** apo MBP (PDB code: 2MV0). **c,** DnaK in ATP bound open conformation (PDB code: 4B9Q). **d,** Firefly luciferase (PDB code: 1LCI)16,38. [16, 38]



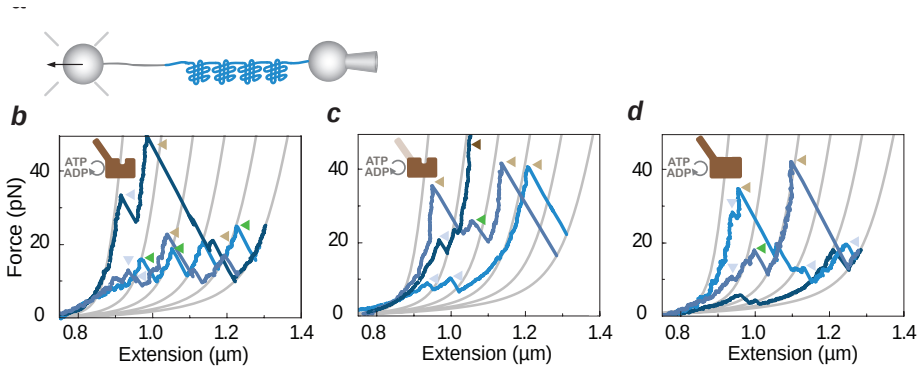
Extended Data Figure 4.4: **Structural maps of MBP and luciferase peptide segments predicted to tightly bind DnaK. a, b** Peptide-library trained predictions of DnaK-interacting peptide segments in the unfolded chain of MBP mapped in yellow, using the algorithm introduced by Rudiger et. al. (PDB code: 2MV0) [13]. **c,d,** Idem for firefly luciferase (PDB code: 1LCI) [13].



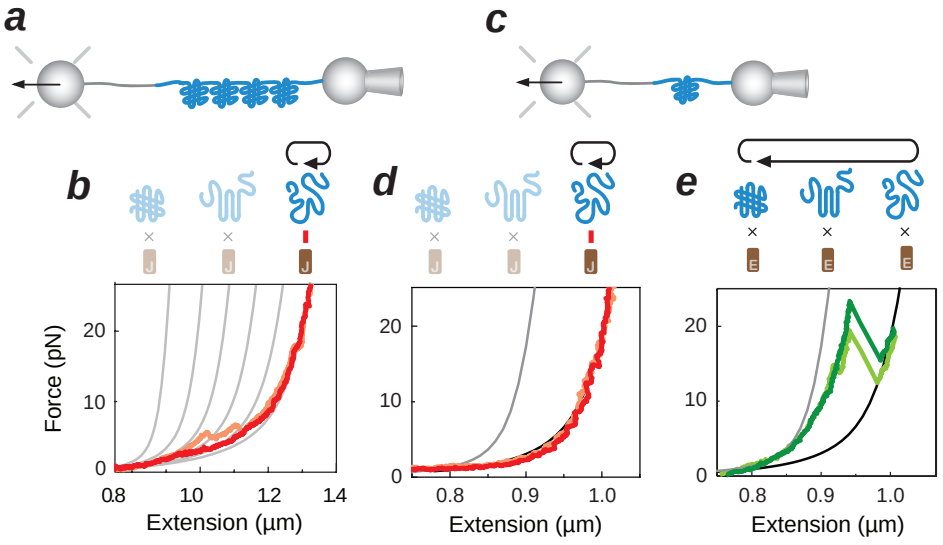
Extended Data Figure 4.5: **a,** Schematic diagram of the set-up. DnaK and ADP were present in the experimental buffer from the start of the experiment. **b,** Force-extension curves of MBP in the presence of 100 nM DnaK and 1 mM ADP loading buffer. Blue: first stretching curves, orange and red: stretching curves after refolding and 3 min waiting at low force. **c,** Force acting on MBP is plotted versus time. Pulls (single stretching-relaxation cycles) on the same MBP molecule are followed by the increasing waiting periods at 0 pN, in the presence of 1 μM DnaK and 1 mM ADP loading buffer. **d,** Force-extension curves of MBP indicate increasing protection of partial folds against force (corresponding to panel **c**). Time of pull from the start of experiment is color-coded: earlier pulls are shown in blue, latter – in red. In other experiment, after first stretching and relaxation, the protein was held relaxed for 5 min, and then was stretched and relaxed again. **e,** Force on MBP versus time. Identical buffer conditions as panels **c** and **d**. **f,** Force-extension curves of MBP indicate the stabilization of compact state of MBP (corresponding to panel **e**). First pull is highlighted in blue, subsequent pulls – in red. **g-h,** MBP stretching and relaxation experiments in the presence of 1 μM DnaK T199A and 1 mM ATP. **h,** Refolding probability during 5 s time window at 0 pN. **i-j,** MBP stretching and relaxation experiments in the presence of 1 μM DnaK with 1 mM ADP purified by HPLC. First stretching curves are shown in blue, stretching curves denoted in shades of red were acquired after 3 min waiting at low force. **k,** In the first stretching curve (blue), only C-terminal fragment of MBP was unfolded, and then force was immediately reduced to low force preventing further unfolding. After 3 min waiting at low force, subsequent stretching showed (orange) that resists forced unfolding up to 50 pN. Experiments performed in 1 μM DnaK with 1 mM ADP loading buffer.



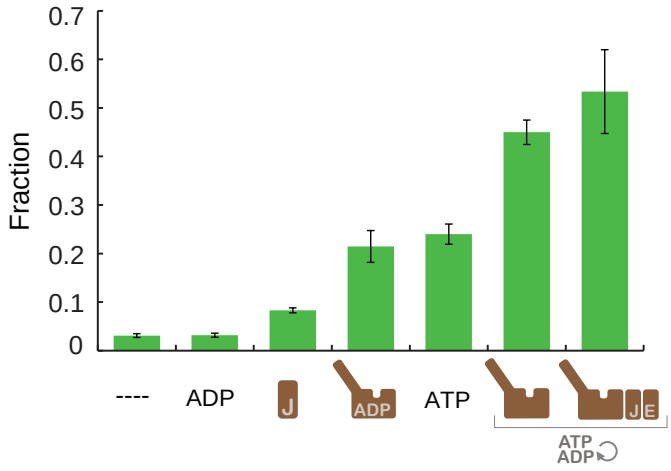
Extended Data Figure 4.6: **DnaK mutant structures. a,** DnaK wild-type structure (PDB: 2KHO) **b,** DnaK V436F (termed groove-mutated) structure model highlighting in red spheres the location of the point mutation hindering access to its groove, resulting in a 17.3-fold reduction in peptide-binding affinity [27] (V436F substitution modeled onto PDB code: 2KH0). **c,** DnaK(2-538) (termed lid-truncated) structure (lid deletion modeled using 2KHO). The part of the lid that is still present in DnaK(2-538) interacts with the NCB domain in the ATP bound open conformation, which might be important for the mechanics of DnaK.



Extended Data Figure 4.7: **4MBP refolding in the presence of ATP and wild type or mutants of DnaK. a,** Scheme of the experiment. Stretching curves of 4MBP in the presence of ATP wild-type DnaK **(b),** lid-less DnaK **(c),** and groove-less DnaK **(d).**



Extended Data Figure 4.8: **MBP interaction with cochaperone DnaJ and nucleotide exchange factor GrpE. a-b,** 4MBP stretching in the presence of DnaJ. **a,** Scheme of the experiment. **b,** A diagram of protein states and stretching curves. **c,** Scheme of the 1MBP experiment. **d,** MBP stretching in the presence of DnaJ. **e,** MBP stretching in the presence of GrpE.



Extended Data Figure 4.9: **DnaK preserves enzyme activity.** Bulk luciferase protection functional assay monitoring luciferase activity at 37°C in the absence or presence of nucleotides and chaperones as indicated. The active fraction of luciferase after 45 min from the start of temperature upshift is shown. The experiment is performed in triplicates; error bars indicate the standard error of the mean.

References

[1] F. Rodriguez, F. Arsène-Ploetze, W. Rist, S. Rüdiger, J. Schneider-Mergener, M. P. Mayer, and B. Bukau, *Molecular basis for regulation of the heat shock transcription factor sigma32 by the DnaK and DnaJ chaperones*. Mol. Cell **32**, 347 (2008).

[2] D. Skowyra, C. Georgopoulos, and M. Zylicz, *The E. coli dnaK gene product, the hsp70 homolog, can reactivate heat-inactivated RNA polymerase in an ATP hydrolysis-dependent manner*. Cell **62**, 939 (1990).

[3] E. Kirschke, D. Goswami, D. Southworth, P. R. Griffin, and D. A. Agard, *Glucocorticoid receptor function regulated by coordinated action of the Hsp90 and Hsp70 chaperone cycles*. Cell **157**, 1685 (2014).

[4] S. K. Sharma, P. De los Rios, P. Christen, A. Lustig, and P. Goloubinoff, *The kinetic parameters and energy cost of the Hsp70 chaperone as a polypeptide unfoldase*. Nat. Chem. Biol. **6**, 914 (2010).

[5] H. Schröder, T. Langer, F. U. Hartl, and B. Bukau, *DnaK, DnaJ and GrpE form a cellular chaperone machinery capable of repairing heat-induced protein damage*. EMBO J. **12**, 4137 (1993).

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

[7] M. P. Mayer and B. Bukau, *Hsp70 chaperones: cellular functions and molecular mechanism*. Cell. Mol. Life Sci. **62**, 670 (2005).

[8] J. Frydman, E. Nimmesgern, K. Ohtsuka, and F. U. Hartl, *Folding of nascent polypeptide chains in a high molecular mass assembly with molecular chaperones*. Nature **370**, 111 (1994).

[9] F. U. Hartl and M. Hayer-Hartl, *Molecular chaperones in the cytosol: from nascent chain to folded protein*. Science **295**, 1852 (2002).

[10] J. Tyedmers, A. Mogk, and B. Bukau, *Cellular strategies for controlling protein aggregation*. Nat. Rev. Mol. Cell Biol. **11**, 777 (2010).

[11] J. M. Warrick, H. Y. Chan, G. L. Gray-Board, Y. Chai, H. L. Paulson, and N. M. Bonini, *Suppression of polyglutamine-mediated neurodegeneration in Drosophila by the molecular chaperone HSP70*. Nat. Genet. **23**, 425 (1999).

[12] X. Zhu, X. Zhao, W. F. Burkholder, A. Gragerov, C. M. Ogata, M. E. Gottesman, and W. A. Hendrickson, *Structural analysis of substrate binding by the molecular chaperone DnaK*. Science **272**, 1606 (1996).

[13] S. Rüdiger, L. Germeroth, J. Schneider-Mergener, and B. Bukau, *Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries,* EMBO J. **16**, 1501 (1997).

[14] a. Szabo, T. Langer, H. Schröder, J. Flanagan, B. Bukau, and F. U. Hartl, *The ATP hydrolysis-dependent reaction cycle of the Escherichia coli Hsp70 system DnaK, DnaJ, and GrpE.* Proc. Natl. Acad. Sci. U. S. A. **91**, 10345 (1994).

[15] R. Qi, E. B. Sarbeng, Q. Liu, K. Q. Le, X. Xu, H. Xu, J. Yang, J. L. Wong, C. Vorvis, W. a. Hendrickson, L. Zhou, and Q. Liu, *Allosteric opening of the polypeptide-binding site when an Hsp70 binds ATP.* Nat. Struct. Mol. Biol. **20**, 900 (2013).

[16] R. Kityk, J. Kopp, I. Sinning, and M. P. Mayer, *Structure and dynamics of the ATP-bound open conformation of Hsp70 chaperones.* Mol. Cell **48**, 863 (2012).

[17] R. Schlecht, A. H. Erbse, B. Bukau, and M. P. Mayer, *Mechanics of Hsp70 chaperones enables differential interaction with client proteins.* Nat. Struct. Mol. Biol. **18**, 345 (2011).

[18] R. G. Smock, M. E. Blackburn, and L. M. Gierasch, *Conserved, disordered C terminus of DnaK enhances cellular survival upon stress and DnaK in vitro chaperone activity.* J. Biol. Chem. **286**, 31821 (2011).

[19] A. Buchberger, H. Schröder, T. Hesterkamp, H. J. Schönfeld, and B. Bukau, *Substrate shuttling between the DnaK and GroEL systems indicates a chaperone network promoting protein folding.* J. Mol. Biol. **261**, 328 (1996).

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

[21] P. Bechtluft, R. G. H. van Leeuwen, M. Tyreman, D. Tomkiewicz, N. Nouwen, H. L. Tepper, A. J. M. Driessen, and S. J. Tans, *Direct observation of chaperone-induced changes in a protein folding pathway.* Science **318**, 1458 (2007).

[22] A. Buchberger, A. Valencia, R. McMacken, C. Sander, and B. Bukau, *The chaperone function of DnaK requires the coupling of ATPase activity with substrate binding through residue E171.* EMBO J. **13**, 1687 (1994).

[23] L. Packschies, H. Theyssen, A. Buchberger, B. Bukau, R. S. Goody, and J. Reinstein, *GrpE accelerates nucleotide exchange of the molecular chaperone DnaK with an associative displacement mechanism,* Biochemistry **36**, 3417 (1997).

[24] A. Mashaghi, G. Kramer, P. Bechtluft, B. Zachmann-Brand, A. J. M. Driessen, B. Bukau, and S. J. Tans, *Reshaping of the conformational search of a protein by the chaperone trigger factor.* Nature **500**, 98 (2013).

[25] M. D. Wang, H. Yin, R. Landick, J. Gelles, and S. M. Block, *Stretching DNA with optical tweezers*. Biophys. J. **72**, 1335 (1997).

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

[27] M. P. Mayer, H. Schröder, S. Rüdiger, K. Paal, T. Laufen, and B. Bukau, *Multistep mechanism of substrate binding determines chaperone activity of Hsp70*. Nat. Struct. Biol. **7**, 586 (2000).

[28] R. Kellner, H. Hofmann, A. Barducci, B. Wunderlich, D. Nettels, and B. Schuler, *Single-molecule spectroscopy reveals chaperone-mediated expansion of substrate protein,* Proc. Natl. Acad. Sci. **111**, 13355 (2014).

[29] J. Gamer, G. Multhaup, T. Tomoyasu, J. S. McCarty, S. Rüdiger, H. J. Schönfeld, C. Schirra, H. Bujard, and B. Bukau, *A cycle of binding and release of the DnaK, DnaJ and GrpE chaperones regulates activity of the Escherichia coli heat shock transcription factor sigma32*. EMBO J. **15**, 607 (1996).

[30] J. M. Nũnes, M. Mayer-Hartl, F. U. Hartl, and D. J. Müller, *Action of the Hsp70 chaperone system observed with single proteins*. Nat. Commun. **6**, 6307 (2015).

[31] S. L. Rutherford and S. Lindquist, *Hsp90 as a capacitor for morphological evolution*. Nature **396**, 336 (1998).

[32] Z. Li, F. U. Hartl, and A. Bracher, *Structure and function of Hip, an attenuator of the Hsp70 chaperone cycle*. Nat. Struct. Mol. Biol. **20**, 929 (2013).

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

[34] M. Pellecchia, T. Szyperski, D. Wall, C. Georgopoulos, and K. Wüthrich, *NMR structure of the J-domain and the Gly/Phe-rich region of the Escherichia coli DnaJ chaperone*. J. Mol. Biol. **260**, 236 (1996).

[35] J. Li, X. Qian, and B. Sha, *The Crystal Structure of the Yeast Hsp40 Ydj1 Complexed with Its Peptide Substrate,* Structure **11**, 1475 (2003).

[36] C. J. Harrison, *Crystal Structure of the Nucleotide Exchange Factor GrpE Bound to the ATPase Domain of the Molecular Chaperone DnaK,* Science **276**, 431 (1997).

[37] M. P. Mayer, *Gymnastics of molecular chaperones*. Mol. Cell **39**, 321 (2010).

[38] E. Conti, N. P. Franks, and P. Brick, *Crystal structure of firefly luciferase throws light on a superfamily of adenylate-forming enzymes,* Structure **4**, 287 (1996).

5

Binding flexibility of internally disordered chaperone Hsp33 studied with optical tweezers

Hsp33 is a prokaryotic redox-regulated chaperone that undergoes activation at oxidative stress. Activated Hsp33 has been shown to interact with unfolded polypeptides and prevent them from aggregation. However, recent indications of Hsp33 preferred binding to early unfolding intermediates and its possible involvement in protein degradation pathways raise the possibility of additional interaction mechanisms. Overall, the molecular and mechanical basis of chaperone-client complex is not understood. Here we investigate how Hsp33 interacts with folded and unfolded proteins using optical tweezers. We demonstrated that it does not only prevent aggregation, but allows native-like refolding of an aggregation-prone protein. In addition to the interaction with non-folded protein stretches, Hsp33 was found to interact with folded protein structures during refolding and interfere with stable protein core formation. The ability of Hsp33 to interact with folded structure could be related to possessing intrinsically disordered regions that may allow for such remarkable binding flexibility.

5.1. Introduction

All living organisms experience a variety of stresses during the lifetime. A multitude of stressing factors, such as temperature, pH, salt concentration, exists that organisms need to cope with to survive. Reactive oxygen species (ROS) are produced normally at the many stages of aerobic metabolism [1]. However, they are dangerous for the cell, because they cause oxidative modifications of various cell components (DNA, lipids, proteins) [2]. In the case of proteins, ROS can drive their aggregation and lower the activity of enzymes [3]. Severe oxidative stress in vivo during aging typically results in massive protein unfolding and loss of function [4].

In order to protect the proteome in stress conditions, cells are equipped with a range of molecular chaperones, which are activated specifically upon certain stress. After being activated, they promiscuously bind non-folded protein stretches independently of ATP presence and prevent their aggregation during stress conditions (such chaperones are generally referred to as "holdases") [5, 6]. As the conditions are back to normal, the holdases transfer their protein client to other chaperones, which assist protein native refolding in an ATP-dependent manner ("foldases") [7]. Various regulatory switching mechanisms are employed by the holdases to sense the environmental change and tune the affinity for a protein client [8-10]. They are often only expressed upon stress, whereas foldases are typically constitutively expressed being occasionally upregulated under stress conditions [5].

In this study, we investigate an oxidation-activated chaperone Hsp33 from *E. Coli*. It features 292 amino acids and is compactly folded in non-stress conditions (Fig. 5.1). Its expression is upregulated upon various stress conditions leading to protein unfolding [11]. Hsp33 is activated in oxidative stress conditions leading to massive protein unfolding, such as peroxide stress at slightly elevated temperatures, or hypochlorite treatment without temperature upshift [12, 13]. The activation occurs via extensive conformational changes in C-terminal redox-switch domain that acts as a dual stress sensor [14]. It features the tetra-cysteine zinc-coordinating center and an alpha-helical linker region [15]. At physiological conditions, both subdomains are folded and packed on the core N-terminal domain, and Hsp33 does not bind to proteins. The Zn-coordinating center unfolds when cysteines become oxidized, and the linker domain unfolds at denaturing conditions. Hence, most of the C-terminal domain becomes destabilized upon severe oxidative stress or combined heat shock and oxidation. Unfolding of C-terminal domain leads to the dimerization of Hsp33 and to full activation of its chaperone functions [14]. Recent findings suggest that Hsp33 binds protein clients via the destabilized linker region that regains stability upon binding the client [10]. When the stress is over, the release of the protein client is regulated by Hsp70 chaperone system. Hsp33 is classified as a holdase, and, as other holdases, is believed to interact exclusively with non-folded protein stretches. However, recent indications of Hsp33 preferred binding to early unfolding intermediates and its possible involvement in protein degradation pathways raise the possibility of additional interaction mechanisms [10, 16]. Overall, little is known about the

mechanical properties of the chaperone-client complex.

We utilize single molecule force spectroscopy in optical tweezers to study mechanistic details of Hsp33 interaction with protein client at the single molecule scale. Forced unfolding and subsequent refolding of maltose-binding protein (MBP) and its aggregation-prone engineered repeat (4MBP) were followed in the isolation and in the presence of a constitutively active Hsp33 mutant Y12E[1]. We found that Hsp33 binds unfolded polypeptides as well as folded structures. The binding of Hsp33 to unfolded chains suppresses folding and prevents aggregation. Native-like structures do not interact with the chaperone as expected, but the protein structures that fold in the presence of Hsp33 are destabilized by binding to Hsp33.

5.2. Methods

5.2.1. Expression and purification of Hsp33 and Hsp33 Y12E

It was performed in the group of Prof. Ursula Jakob in the University of Michigan (Ann Arbor, USA) as previously described [17].

5.2.2. Expression and purification of biotinylated MBP/4MBP

MBP and 4MBP were produced in *E. coli* and purified as previously described in chapter 4, section 4.1.1.

5.2.3. Optical tweezers assay

The experiments described in this chapter were performed using a custom made optical tweezers setup [18]. The process of experiment preparation, incubation of materials, tether assembly and data acqusition has been described extensively in chapter 4, section 4.1.4.

5.3. Results

To probe the aggregation prevention capability of Hsp33 Y12E (is called Hsp33 from now on for simplicity) at the single-molecule scale, we used a protein client that is prone to aggregation in vitro: four MBP molecules linked head-to-tail (Fig. 5.1a) [18]. First, we subjected 4MBP to repeated cycles of stretching and relaxation without any chaperone ($N = 97$), separated by 5 s waiting period when the protein is fully relaxed (at 0 pN). Stretching of natively folded 4MBP (first stretching curve) produced a gradual unfolding transition followed by four distinct unfolding steps (Fig. 5.1b) indicating a sequential opening of core structures (80-104 nm) of four MBP molecules. Then we relaxed the protein to low forces providing it with an opportunity to refold, and stretched it again. Subsequent stretching showed that a compact structure had been formed (Fig. 5.1c), however, it typically failed to unfold or unfolded in steps larger than one MBP domain, indicating

[1] Activated wild type Hsp33 is active in a low-salt buffer, which degrades the properties of linkages in the optical tweezers assay. Hence we focused on the mutant which usage did not interfere with assay performance.

the formation of interdomain interactions (termed misfold). Overall, we found that a major part of protein chain was involved in misfolding interactions (40%), and only very small fraction (3%) displayed native-like refolding (Fig. 5.1f). Thus, isolated 4MBP refolded into misfolded mini-aggregates consisting of more than one MBP repeat.

Next, these experiments were performed in the buffer containing Hsp33 Y12E. In the presence of chaperone, natively folded 4MBP followed the same unfolding pathway in the first pull (Fig. 5.1b) while the relaxed unfolded polypeptide chain at zero force refolded differently in a concentration-dependent manner (Fig. 5.1d,e). With 0.5 μM of Hsp33 Y12E (Fig. 5.1d), the fraction of misfolded structures lowered significantly (24%). Furthermore, the probability of native-like refolding increased (14%) ((Fig. 5.1d,f). This was also accompanied with the slight decrease of the total length of polypeptide chain that refolded (46% vs 51%). When the chaperone concentration was increased to 5μM, misfolding was completely suppressed (0.03%, Fig. 5.1e), however, most of the protein chain stayed unfolded (86%). It indicates that Hsp33 interacts with 4MBP and prevents its aggregation when bound to non-folded 4MBP chain. Moreover, native-like fraction of the protein length increased to 8%. It suggests that Hsp33 can possibly promote the native-like refolding of an aggregation-prone protein by preventing its aggregation without the collaboration with Hsp70 system.

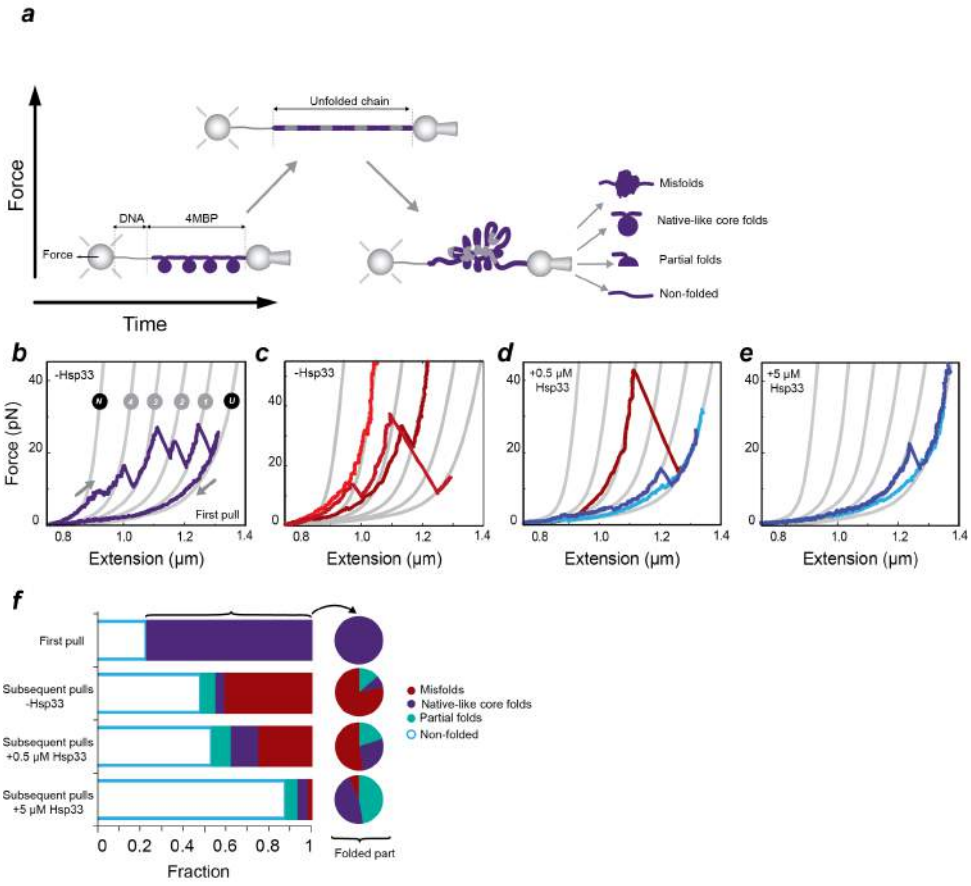


Figure 5.1: **Hsp33 suppresses misfolding between domains and promotes native-like folds. a,** Schematic diagram of the set-up. A construct repeat of four maltose binding proteins (4MBP) is tethered by means of a DNA handle between two beads: one is attached to the piezo-driven micropipette, the other is held by an optical trap that allows force detection. The DNA-protein tether can be stretched by displacing the pipette bead, and this results in the protein unfolding. When the beads are brought together, the applied force decreases and the polypeptide chain relaxes. Keeping the construct at zero force for 5 s gives the chain another chance to refold. **b,** First stretching (upward arrow)-relaxation (downward arrow) curve, showing the unfolding pattern of natively-folded 4MBP. Gray lines represent the theoretical WLC characterizing the DNA-protein construct from natively-folded (N) to the unfolded (U) state. When 4MBP is stretched, after C-terminal unfolding (N ⇒ 4), four native-like core unfolding events (4 ⇒ 3 ⇒ 2 ⇒ 1 ⇒ U) are observed. After relaxation and waiting at zero force for 5 s in the absence of chaperone, second or subsequent stretching curves **(c)** show no unfolding (bright red) that indicates tight misfolding interactions between domains. **d,** In the presence of Hsp33 (0.5 μM), second or subsequent stretchings show characteristics of native interactions (purple) and weak non-native interactions (red), but no tight ones. Sometimes there are fully unfolded chains (blue). **e,** When the concentration of Hsp33 was increased (5 μM), misfolded structures vanished and the polypeptide chains either stayed fully unfolded (blue) or displayed characteristics of native interactions (purple line). **f,** Fractions of structures formed in 4MBP refolding. Indicated fractions in bar charts are the average length fractions of the protein chain that is involved in one of the four states. Pie charts represent relative fractions of different folded structures within folded part.

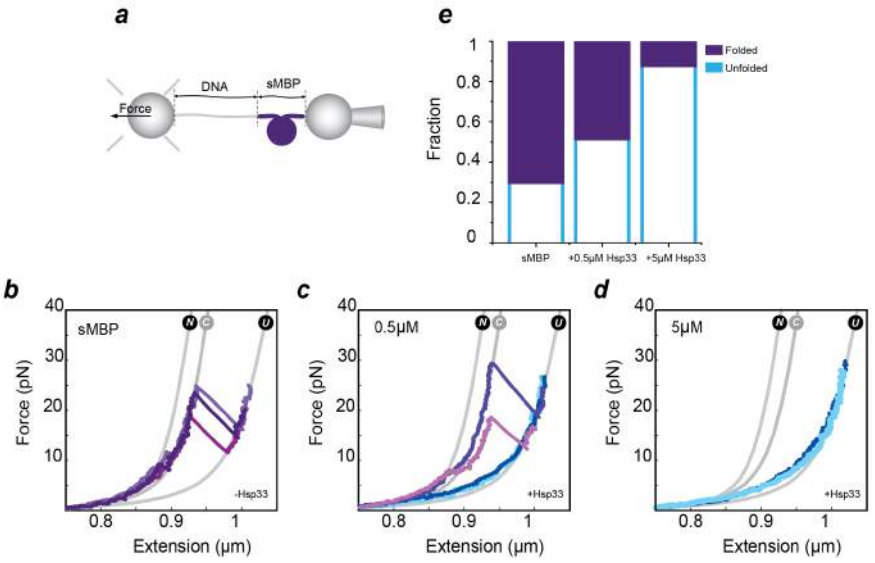


Figure 5.2: **Hsp33 prevents folding. a,** Schematic diagram of a single MBP (sMBP) construct tethered between two beads. **b** Stretching-relaxation experiments in the absence of Hsp33 show frequent native-like refolding of sMBP. The native-like structure unfolds via two transitions, corresponding to C-terminal unfolding (N ⇒ C) and core unfolding (C ⇒ U) respectively. **c,** In the presence of Hsp33 (0.5 μM), stretching-relaxation experiments indicate that the refolding of sMBP is partially prevented. **d,** By increasing the concentration of Hsp33 (5 μM) refolding could be almost completely prohibited. **e,** Refolding fractions in the absence and presence of Hsp33.

To obtain more detailed mechanistic understanding of chaperone-client interaction, we focused on the simpler case of single MBP molecule (Fig. 5.2a) that can refold spontaneously and does not aggregate ($N = 138$). In the absence of the chaperone, MBP unfolding results in a simple pattern involving two transitions (Fig. 5.2b). When the force reached about 10 pN, the total measured protein contour length increased by 28 nm (N ⇒ C), indicating the unfolding of a the unfolding of a C-terminal α-helix. The remaining "core" structure unfolded completely at about 25 pN, yielding a step size of 92 nm (C ⇒ U). When the protein chain was relaxed to low force and allowed to refold for 5 s, it refolded in 70% of the cases (Fig. 5.2e) and yielded similar unfolding pattern in subsequent stretching. In the presence of Hsp33, the first stretching curve was unchanged but refolding of unfolded polypeptide chains was suppressed in the subsequent pulls (Fig. 5.2e). Consistently with 4MBP results, the level of suppression depends on the chaperone concentration. When 5 uM of Hsp33 Y12E were added to the buffer, MBP refolding was largely prevented (12%, Fig. 5.2e), and in most of the cases (88%) the protein stayed unfolded indicating stable interaction of the chaperone with non-folded protein chain (Fig. 5.2d). When Hsp33 Y12E concentration was lowered to 0.5 uM, native-like refolding fraction of MBP increased (52%, Fig. 5.2c,e) but remained lower than w/o Hsp33, indicating the effect of the chaperone being less dominant than at higher concentration.

5.4. Discussion and conclusion

It had been suggested that, once bound, Hsp33 does not leave the client until it is handed out to the Hsp70 chaperone system [10], a feature that is also characteristic for small heat shock proteins (sHsps). Hence, in the absence of the functional Hsp70 system, Hsp33 is believed to stay bound to the unfolded protein preventing its refolding. However, our data demonstrate native-like core refolding of one domain in 4MBP after it was blocked in the fully unfolded state by Hsp33. It suggests that the chaperone-client complex disrupts followed by spontaneous MBP refolding. Alternatively, it may be that Hsp33 Y12E did not interact with this domain, but allowed it to refold spontaneously by blocking the neighboring domains and lowering the aggregation risk. We note that Hsp33 unbinding from the client in the absence of Hsp70 may not occur in ensemble studies, because in our assay the disruption of stabilizing interaction between Hsp33 and 4MBP may had been triggered by stretching of the protein client (for example, unwinding of an α-helix).

To test the consistency between 4MBP and MBP experiments and check whether the ability of native-like refolding within 4MBP construct just comes from blocking misfolding partners by Hsp33 or not, we modified a simple model (Extended Data Subsection 5.5.1) based on the statistical thermodynamics. In this model, a fully unfolded 4MBP is represented by a one-dimensional flexible lattice that features four binding sites (one site for each domain). The sites are allowed to interact with each other with an equal probability (α) and this interaction results in misfolding, whereas non-interacting sites could fold or stay unfolded (Extended Data Fig. 5.1). Using 4MBP stretching-relaxation experiments, we calculated the αvalue (Extended Data Subsection 5.5.1). In the presence of Hsp33, chaperone molecules bind non-specifically and non-cooperatively to the lattice and occupy the sites, in a way that each lattice site could be occupied only with one chaperone molecule (1:1 stoichiometric ratio) [19-21]. Occupied sites are assumed to be blocked in the non-folded state, while unoccupied domains could either (natively, partially or mis-) fold or stay unfolded (Extended Data Fig. 5.2). Based on these assumptions and using the experimentally obtained refolding probabilities of MBP (in the presence of chaperone), we modeled how 4MBP would interact with Hsp33 and compared the predicted behavior with the experimental observations (Extended Data Tables 5.2,5.3). For 5 μM Hsp33 condition, the prediction from the model matches experimentally observed unfolded and misfolded fractions of polypeptide chain. For 0.5 μM Hsp33 condition, the model predicts less folded fraction than what we have observed experimentally (Extended Data Table 5.2). Interestingly, the folded fraction increased on account of native-like folding fraction rather than misfolding (Extended Data Table 5.3). This may suggest that Hsp33 does something more than stabilizing the polypeptide chain in the unfolded state.

Recent findings suggest that Hsp33 prefers binding to fragments of protein secondary structure in early unfolding intermediates [10]. We found that the refolding probability of MBP decreased in the presence of Hsp33 Y12E indicating that it also interacts with non-folded protein stretches. We note that we can-

not exclude the possibility of the chaperone binding to the residual secondary structure fragments, however, our data clearly indicates that Hsp33 interacts with almost completely unfolded protein structures. Consistently, most of 4MBP stretching curves, after they had been relaxed in the presence of Hsp33, did not display aggregation and stayed fully unfolded. It indicates that Hsp33 is bound to non-folded 4MBP and stays bound during relaxation and subsequent stretching, preventing 4MBP aggregation.

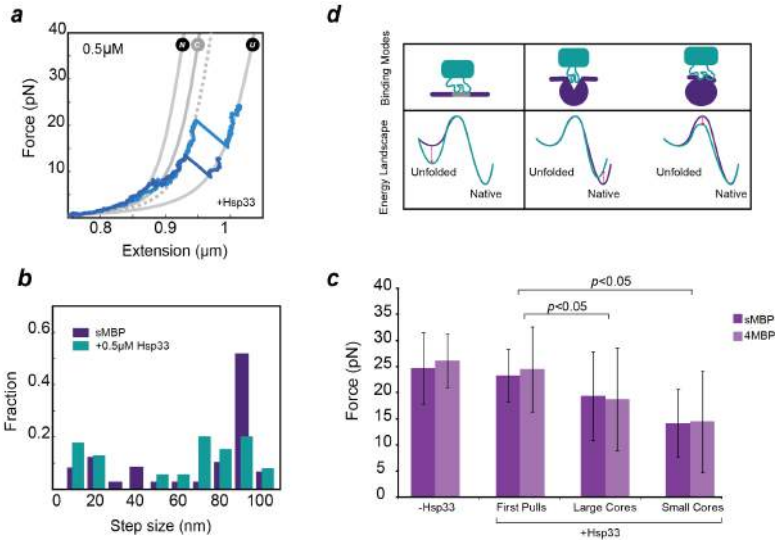


Figure 5.3: **Hsp33 binds and destabilizes folded structures. a,** The force-extension data suggest some of the structures that formed in the presence of Hsp33 have a smaller core-size (dashed line), as compared to the native MBP core (C line). This could be a consequence of interaction between the folded structure and Hsp33. **b,** Length distribution of unfolding steps with (green) and without (purple) Hsp33, as derived from the force-extension data. The number of times a certain step length is visited is indicated as a fraction of the total number of all visited steps. Without Hsp33, there is a clear sharp peak at ~ 90 nm (corresponds to the MBP native core), which becomes significantly shorter and wider with Hsp33 (0.5 μM). **c,** Core unfolding forces of sMBP (dark purple) and 4MBP (light purple) constructs. In the presence of Hsp33, the cores withstand roughly the same force in first pulls, but they become less stable in the subsequent pulls and get unfolded at lower forces, as compared to no chaperone condition. This is statistically significant ($p < 0.05$) for both large (80-100 nm) and small (60-80 nm) cores. **d,** Cartoon of folding landscape reshaping by Hsp33. Left: binding of Hsp33 to the unfolded state and stabilization of this state can be seen as deepening the energy valley of the unfolded state. Right: the formation of smaller cores and the destabilization of folded structures by Hsp33 could be interpreted as either formation of a new valley with an energy higher than the native state, or as lowering the energy barrier between the unfolded and native state.

Further analysis of the structures formed in the presence of Hsp33 likely helps the understanding of the interaction between the chaperone and the substrate. In the presence of 0.5 μM Hsp33, the folded molecules at the first glance seemed to be native-like. They mostly were as compact as native-like folded single MBP, indicated by following the same theoretical WLC curve at low force (Fig. 5.3a,

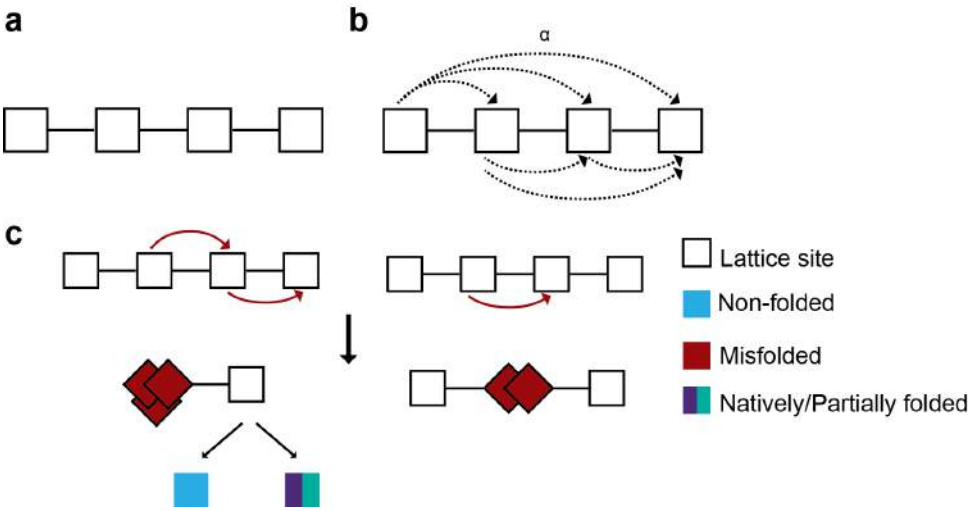
N line), and they unfolded via two distinct transitions (Fig. 5.3a). However, fine analysis of the unfolding step sizes (Fig. 5.3b) and unfolding forces (Fig. 5.3e) revealed that the presence of Hsp33 affected the properties of refolded MBP protein core. In the absence of the chaperone, properly refolded MBP core unfolded at 24.8 ± 6.9 pN and yielded $\sim 92 \pm 12$ nm contour length. However, with 0.5 μM of Hsp33, the protein length involved in the MBP core formation became shorter on average (80 ± 10 nm, Fig. 5.3a,b). Since the refolding occured, it indicates that Hsp33 did not bind unfolded MBP in this case, but suggests that Hsp33 can also bind to partially folded structures during MBP refolding and interfere with the core formation. Since the unfolding pattern of natively folded MBP (in the first stretching cycle on a fresh MBP molecule) is not affected by Hsp33 (Fig. 5.3c), we do not have any evidence for the interaction of Hsp33 with natively folded protein core. However, after MBP was allowed to refold in the presence of the chaperone, it in some cases entered the subsequent stretching cycle in the state as compact as fully folded MBP. It suggests that Hsp33 binds to MBP during the relaxation, because otherwise it would refold natively and the chaperone-protein interaction would not be possible anymore.

Moreover, the unfolding force of MBP cores of both normal (80-104 nm) and smaller (60-80 nm) size decreased in the presence of Hsp33 to 18.4 ± 8.9 and 16.8 ± 5.7 pN respectively (Fig. 5.3c). It indicates that the chaperone is capable of affecting the energy landscape of a protein client (Fig. 5.3d). An ability to destabilize misfolded and aggregated protein has been reported for ATP-dependent chaperones, such as Hsp70 and Hsp110, which can potentially exploit the energy of ATP-hydrolysis to overcome energy barriers and forcefully unfold stable aggregates (reviewed in [22]). However, Hsp33 acts in an ATP-independent manner. Transient binding and unbinding of Hsp33 to the protein client may offer a possible explanation of the destabilization. The chaperone may transiently stabilize unstable folding intermediates by binding the partial fold directly, which results in effective destabilization after the chaperone unbinding. Alternatively, Hsp33 may be bound to a non-folded protein fragment close to the protein core that may also be enough to interfere with its refolding and make it less mechanically stable.

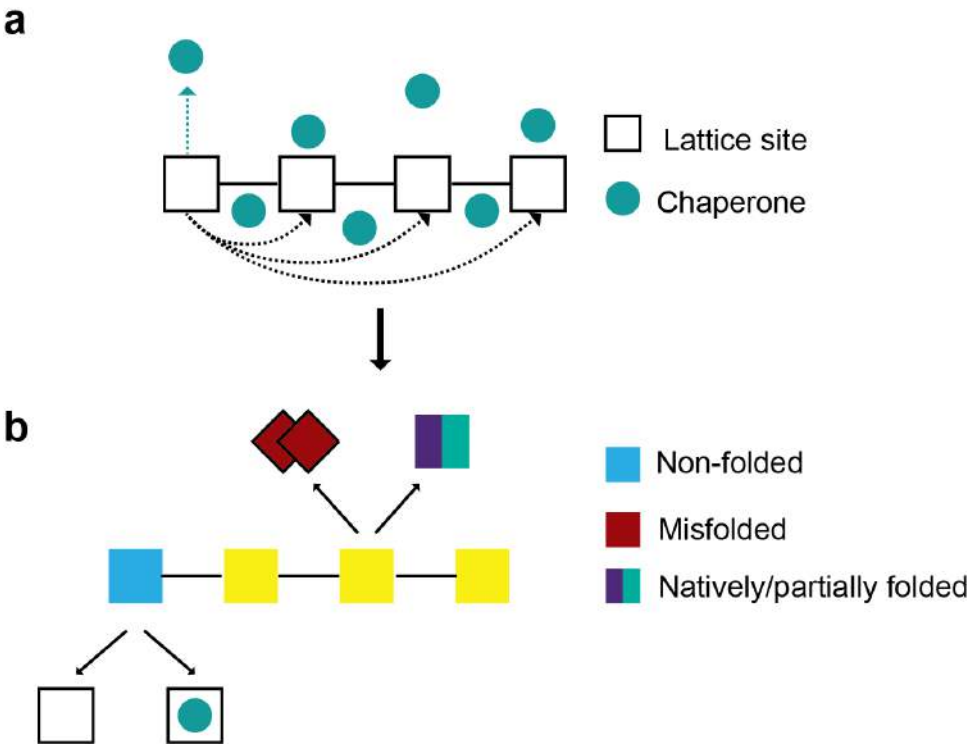
Overall, our findings indicate that Hsp33 function is broader than "holdase" activity as it is traditionally defined. In addition to the prevention of protein aggregation by blocking them in the fully unfolded state, Hsp33 allows some refolding of a client protein while present in sub-stoichiometric concentrations. It suggests that the clients of Hsp33 may be delivered to the "foldase" chaperones in a state closer to native, than previously assumed. The destabilization of the MBP core by Hsp33 indicates that the chaperone may bind not only non-folded protein fragments, but also partially folded protein states, consistently with previous findings by Reichmann et al. [10] This study also indicated that Hsp33 uses its own disordered linker region, which is also disordered in Hsp33 Y12E, to interact with the clients. Intrinsically disordered substrate binding site can provide large flexibility of binding and allow to avoid sequence specificity. Hence, it can possibly be the key for the remarkable diversity of the binding modes of Hsp33

that we observed. It is consistent with the binding specificity of other disordered proteins whose binding partners are thought to act as stabilizing scaffolds [23]. More research is clearly necessary to elucidate further the mechanisms by which Hsp33 interacts with their target proteins. In the next step, our results should be complemented with bulk assays and optical tweezers experiments with activated wild type (WT) Hsp33.

5.5. Extended data



Extended Data Figure 5.1: **Schematic representation of 4MBP model. a,** A fully unfolded 4MBP has been considered as a lattice featuring 4 sites. **b,** All the sites could interact with each other with the same probability equal to α. **c,** Forming a connection between two domains leads to misfolding while non-interacting sites could either natively/partially fold or stay unfolded.



Extended Data Figure 5.2: **Schematic representation of 4MBP model in the presence of chaperone. a,** Fully unfolded 4MBP is considered as a lattice having 4 binding sites. Each site may interact with a chaperone or with other non-occupied domain of 4MBP. **b,** Non-folded domain can either interact with Hsp33, which blocks it in the unfolded state, or natively (partially) fold on the isolation. When two domain interact to each other, they form a msifolding interaction.

5.5.1. Statistical model

In this model a fully unfolded 4MBP is considered to be a one-dimensional flexible lattice that features four chaperone binding sites (Extended Data Fig. 5.1a). All sites are allowed to interact with each other with an equal probability (Extended Data Fig. 5.1b). If an interaction forms between two sites, they enter the misfolding interaction (Extended Data Fig. 5.1c). The size of misfolded structure indicates how many sites (domains) have been involved in misfolding. If the site is not in interaction with other sites, it is observed either as an unfolded stretch or a natively/partially folded structure (Extended Data Fig. 5.1c). αrepresents the probability of interaction between two sites and P_i stands for the probability of getting i misfolded sites. Experimental results were used to calculate the value of α. Extended Data Table 5.1 shows the experimentally measured vs calculated values for different i.

To calculate the MBP-Hsp33 complex dissociation constant, we considered a simple first-order reaction between the protein substrate (S) and the chaperone

(C). Assuming the fraction of bound chaperones was equilibrated at zero force, the occupancy number could be calculated as follows:

$$\langle n \rangle = \frac{[\text{Hsp33.MBP}]}{[\text{MBP}_{\text{total}}]} = \frac{[\text{Hsp33}]}{K_D + [\text{Hsp33}]} = \frac{\overline{C}}{\overline{C} + K_D} \tag{5.1}$$

We used the experimentally obtained refolding probability (P_f) of MBP in the presence of different concentrations of Hsp33 (namely, 0.5 μM and 5 μM) to calculate the MBP-Hsp33 complex dissociation constant (K_D). The main assumption is that when Hsp33 is bound to the polypeptide chain, it does not refold. But when the polypeptide chain is free, it could either fold or stay unfolded (consistent with no-chaperone experiment observations). K_D derived from both datasets matches closely (with 0.5 μM Hsp33 $K_D = 1.4$ μM and with 5 μM Hsp33 $K_D = 1.03$ μM), and it is consistent with previous reports for some client proteins [10, 19].

In the presence of the chaperone, the protein lattice sites have two possibilities for interaction: extra-lattice interaction with chaperone molecules and intra-lattice interaction with other sites (Extended Data Fig. 5.2a). Chaperone molecules bind non-specifically and non-cooperatively to the lattice and occupy the sites in a way that each site could be occupied only with one chaperone molecule (1:1 stoichiometric ratio). If a site is occupied by a chaperone molecule, it could not further interact with other sites and would be represented as an unfolded domain, while when it is unoccupied it could either interact (form misfolded structures) or not interact (fold natively/partially or stay unfolded) with other sites (Extended Data Fig. 5.2b). At any given concentration of Hsp33, using the refolding probability (P_f) of MBP one could predict the possibility of getting i unfolded sites (P_i) within the 4MBP lattice. Extended Data Table 5.2 demonstrates the experimentally measured vs calculated values for different P_i s at different concentrations of Hsp33. For both conditions, the folded fraction is bigger than what was expected, although in the presence of 0.5 μM Hsp33, this difference is more significant. Now combining the P_i values and the probability of getting interaction between two sites (α), we could calculate the probability of getting i misfolded sites in the presence of the chaperone (P_i) and see if the bigger folded fraction resulted in the formation of more misfolded structures (Extended Data Table 5.3). We used the experimentally measured P_i values for this comparison to avoid underestimating the misfolded fraction.

Table 5.1: Probabilities of misfolding within 4MBP construct.

State	Misfolded domains	Calculated	Measured
P_0	0	-	0.15
P_2	2	0.33	0.28
P_3	3	0.28	0.33
P_4	4	0.249	0.23

Table 5.2: Probabilities of having non-folded domains within 4MBP construct in the presence of Hsp33.

State	Non-folded domains	5µM		0.5µM	
		Calculated	Measured	Calculated	Measured
P_0	0	0.0002	0	0.07	0.19
P_1	1	0.0061	0.05	0.27	0.5
P_2	2	0.067	0.19	0.374	0.17
P_3	3	0.327	0.37	0.23	0.06
P_4	4	0.6	0.39	0.053	0.08

Table 5.3: Probabilities of having misfolded domains within 4MBP construct in the presence of Hsp33.

State	Misfolded domains	5µM		0.5µM	
		Calculated	Measured	Calculated	Measured
P_0	0	0.87	0.88	0.44	0.395
P_2	2	0.12	0.12	0.35	0.24
P_3	3	0.01	0	0.16	0.365
P_4	4	0	0	0.05	0

References

[1] M. Schrader and H. D. Fahimi, *Peroxisomes and oxidative stress,* Biochim. Biophys. Acta - Mol. Cell Res. **1763**, 1755 (2006).

[2] J. A. Imlay, *Pathways of oxidative damage.* Annu. Rev. Microbiol. **57**, 395 (2003).

[3] B. S. Berlett, *Protein Oxidation in Aging, Disease, and Oxidative Stress,* J. Biol. Chem. **272**, 20313 (1997).

[4] T. C. Squier, *Oxidative stress and protein aggregation during biological aging,* Exp. Gerontol. **36**, 1539 (2001).

[5] K. Richter, M. Haslbeck, and J. Buchner, *The Heat Shock Response: Life on the Verge of Death,* Mol. Cell **40**, 253 (2010).

[6] M. Haslbeck, T. Franzmann, D. Weinfurtner, and J. Buchner, *Some like it hot: the structure and function of small heat-shock proteins.* Nat. Struct. Mol. Biol. **12**, 842 (2005).

[7] A. Mogk, C. Schlieker, K. L. Friedrich, H. J. Schönfeld, E. Vierling, and B. Bukau, *Refolding of substrates bound to small Hsps relies on a disaggregation reaction mediated most efficiently by ClpB/DnaK,* J. Biol. Chem. **278**, 31033 (2003).

[8] M. Haslbeck, S. Walke, T. Stromer, M. Ehrnsperger, H. E. White, S. Chen, H. R. Saibil, and J. Buchner, *Hsp26: a temperature-regulated chaperone.* EMBO J. **18**, 6744 (1999).

[9] A. Müller, S. Langklotz, N. Lupilova, K. Kuhlmann, J. E. Bandow, and L. I. O. Leichert, *Activation of RidA chaperone function by N-chlorination,* Nat. Commun. **5**, 5804 (2014).

[10] D. Reichmann, Y. Xu, C. M. Cremers, M. Ilbert, R. Mittelman, M. C. Fitzgerald, and U. Jakob, *Order out of disorder: working cycle of an intrinsically unfolded chaperone.* Cell **148**, 947 (2012).

[11] S. Chuang and F. Blattner, *Characterization of twentysix new heat shock genes of Escherichia coli,* J. Bacteriol. **175**, 5242 (1993).

[12] J. Winter, K. Linke, A. Jatzek, and U. Jakob, *Severe oxidative stress causes inactivation of DnaK and activation of the redox-regulated chaperone Hsp33.* Mol. Cell **17**, 381 (2005).

[13] J. Winter, M. Ilbert, P. C. F. Graf, D. Ozcelik, and U. Jakob, *Bleach activates a redox-regulated chaperone by oxidative protein unfolding.* Cell **135**, 691 (2008).

[14] M. Ilbert, J. Horst, S. Ahrens, J. Winter, P. C. F. Graf, H. Lilie, and U. Jakob, *The redox-switch domain of Hsp33 functions as dual stress sensor.* Nat. Struct. Mol. Biol. **14**, 556 (2007).

[15] I. Janda, Y. Devedjiev, U. Derewenda, Z. Dauter, J. Bielnicki, D. R. Cooper, P. C. F. Graf, A. Joachimiak, U. Jakob, and Z. S. Derewenda, *The crystal structure of the reduced, Zn2+-bound form of the B. subtilis Hsp33 chaperone and its implications for the activation mechanism.* Structure **12**, 1901 (2004).

[16] N. Bruel, M.-P. Castanié-Cornet, A.-M. Cirinesi, G. Koningstein, C. Georgopoulos, J. Luirink, and P. Genevaux, *Hsp33 controls elongation factor-Tu stability and allows Escherichia coli growth in the absence of the major DnaK and trigger factor chaperones.* J. Biol. Chem. **287**, 44435 (2012).

[17] C. M. Cremers, D. Reichmann, J. Hausmann, M. Ilbert, and U. Jakob, *Unfolding of metastable linker region is at the core of Hsp33 activation as a redox-regulated chaperone,* J. Biol. Chem. **285**, 11243 (2010).

[18] P. Bechtluft, R. G. H. van Leeuwen, M. Tyreman, D. Tomkiewicz, N. Nouwen, H. L. Tepper, A. J. M. Driessen, and S. J. Tans, *Direct observation of chaperone-induced changes in a protein folding pathway.* Science **318**, 1458 (2007).

[19] J. Graumann, H. Lilie, X. Tang, K. a. Tucker, J. H. Hoffmann, J. Vijayalakshmi, M. Saper, J. C. Bardwell, and U. Jakob, *Activation of the redox-regulated molecular chaperone Hsp33-a two-step mechanism.* Structure **9**, 377 (2001).

[20] J. H. Hoffmann, K. Linke, P. C. F. Graf, H. Lilie, and U. Jakob, *Identification of a redox-regulated chaperone network.* EMBO J. **23**, 160 (2004).

[21] P. C. F. Graf, M. Martinez-Yamout, S. VanHaerents, H. Lilie, H. J. Dyson, and U. Jakob, *Activation of the Redox-regulated Chaperone Hsp33 by Domain Unfolding*, J. Biol. Chem. **279**, 20529 (2004).

[22] R. U. H. Mattoo and P. Goloubinoff, *Molecular chaperones are nanomachines that catalytically unfold misfolded and alternatively folded proteins*, Cell. Mol. Life Sci. **71**, 3311 (2014).

[23] P. Tompa and P. Csermely, *The role of structural disorder in the function of RNA and protein chaperones*. FASEB J. **18**, 1169 (2004).

6

Small heat shock protein HspB6 facilitates the compaction of α-synuclein

An aggregation of α-synuclein (aSyn) results in a range of neurodegenerative diseases in vertebrates, including humans. The small heat shock proteins (sHsps) comprise a widespread family of chaperones that are known to interact with aggregation-prone proteins in order to prevent their aggregation. Even though this interaction is not understood yet, it is apparent that having a mechanistic picture of how chaperones interfere with the aSyn aggregation would be very helpful for therapeutic applications. Here we investigate the effect of human sHsp HspB6 on the aggregation of engineered aSyn repeats at the single molecule scale, using optical tweezers. Surprisingly, we found that HspB6 stimulates the formation of compact structures in aSyn tetramer and stabilizes these structures against forced unfolding.

6.1. Introduction

6.1.1. Protein aggregation and alpha-synuclein

A few decades ago protein aggregation was mostly considered as an annoying side process initiated by improper handling of proteins in a test tube. However, it attracted a great interest after discovery of numerous diseases characterized, among other factors, by accumulation of protein aggregates in tissues of a patient. These diseases include a range of neurodegenerative disorders such as Alzheimer's and Parkinson's, prion diseases, type II diabetes, and others. Many of these diseases are becoming more and more widespread due to longer average life expectancy or lifestyle changes. Hence, there is an urgent medical need to understand what causes protein aggregation and find ways to prevent and reverse it. Furthermore, protein aggregation process is not only relevant in pathophysiology, but also in healthy conditions, since cells exploit protein aggregates for certain functions [1–3]. Understanding of structural features of protein aggregates and how they form are also potentially useful for the design of novel nanomaterials with unique properties [4].

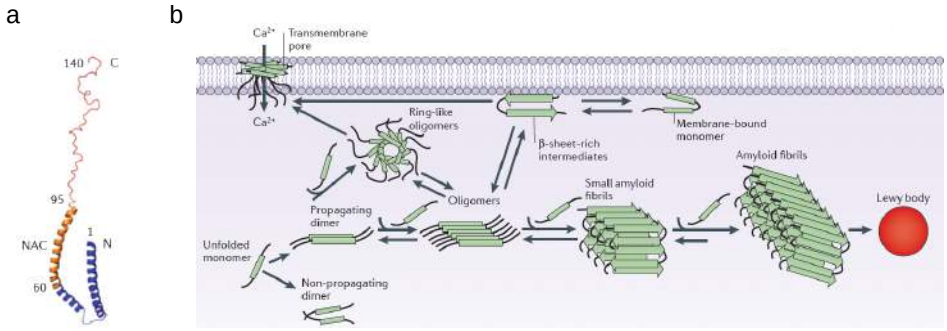


Figure 6.1: **α-synuclein structure, propagation and accumulation. a,** Structure of micelle-bound αSyn monomer (Protein Data Bank ID: 1XQ8) [5]. The N-terminal region, the NAC region, and the C-terminal region are shown in blue, orange and red colour, respectively. Number refer to amino acids flanking corresponding regions. b, Schematic representation of an interplay between different αSyn structures. Adapted by permission from Macmillan Publishers Ltd: Nature Reviews Neuroscience [6], ©2013.

Major advances have been already made in the understanding of structural properties of aggregates and in the kinetics of aggregate formation. Protein aggregates are often amorphous chunks consisting of disordered monomers of mixed sequences, however, most of the disease-associated protein show remarkable aggregation behavior, where small oligomeric protein states arrange into beta-protofibrils, which subsequently form amyloid fibrils. The fibrils are relatively well characterized [7]. They are highly ordered arrangements of mostly unfolded or partially folded protein stretches of primarily one sequence that adopt a fibrillar structure of few nanometers in diameter, but often reaching micrometers

in length. Core of each protofibril has a beta-sheet structure that is hydrogen bonded with the parallel protofibril [8]. Such beta-rich structure provides the fibril with remarkable thermodynamic stability that often exceeds the stability of native protein structure [9, 10]. Fibril architecture is generic for different primary protein sequences [8, 11, 12].

At early stages of aggregation, protein often forms oligomer structures of various sizes that are believed to act as precursors of the fibril formation. For some proteins, it has been demonstrated that they are capable of slow compaction resembling the beta-sheet formation [13]. However, the dynamic process of fibril formation remains poorly understood due to extreme heterogeneity of prefibrillar species. Their structural features and aggregation pathways are challenging to identify in ensemble studies. Apart from the possible key structural role in the fibril formation, oligomeric aggregates are widely suspected to have negative effect on cell viability and serve as toxic agents in the aggregation-associated diseases [14, 15]. Hence, having a thorough understanding of them is important for effective drug development.

α-synuclein (αSyn) is an amyloidogenic protein mainly expressed in the neural tissue of vertebrates, including humans [16]. All homologs of αSyn are quite small (127-140 residues) and highly charged [17]. Human αSyn has 140 amino acids and features three distinct fragments (Fig. 6.1a): lysine-rich N-terminus, which regulates interaction with membranes, acidic C-terminal tail, which is involved in interaction with metals and proteins, and central highly hydrophobic non-amyloid-beta (NAC) fragment, which is essential for αSyn aggregation [6]. Monomeric αSyn can adopt a whole zoo of conformations. In the test tube, native αSyn lacks any stable secondary or tertiary structure [18, 19]. However, upon association to the cell membrane, αSyn becomes partially alpha-helical [20]. Furthermore, a stable tetramer αSyn state resisting aggregation has been observed at physiological conditions [21, 22]; however, there are still debates over this finding [23]. In physiological state, αSyn is essential for normal development of cognitive functions in mice [24]. At the presynaptic terminal, αSyn plays an important role in the synaptic function by regulating neurotransmitter release and vesicle trafficking [6]. αSyn also acts as a molecular chaperone assisting in folding and refolding of synaptic SNARE proteins [25].

An aggregation of αSyn results in the deposition of inclusions mainly consisting of aggregated αSyn in neuronal cell cytoplasm, which are called Lewy bodies (Fig. 6.1b). Together with degradation of dopaminergic neurons, they are markers of Lewy body diseases [26]. It still remains rather unclear how exactly protein aggregation results in disease symptoms. Lewy bodies are considered less toxic by themselves. They are assumed to reflect a part of protective mechanism, where toxic species are sequestered, inactivated and deposited in a less toxic state [6]. Instead of fibrils, αSyn oligomeric aggregates are widely suspected as toxic agent in the Lewy body diseases. Both in vivo and in vitro, αSyn can form various oligomeric structures [27] that are possibly involved in amyloid fibril formation. They exist in equilibrium with monomeric αSyn and can slowly convert to fibrils [13]. Furthermore, round pore-like αSyn oligomers have toxic effects on

the cells, since they can cause membrane leakage [28]. Hence, a thorough understanding of the relationship between different oligomeric states and the kinetics of their interconversions is essential to potentially develop the drugs against neurodegenerative diseases.

6.1.2. Small heat shock proteins

The small heat shock proteins comprise a class of chaperones that are overexpressed upon heat stress, but also often act at non-stress conditions. In humans, sHsps family is named HspB and features 10 members. For all HspBs, some extent of chaperone activity was detected in vitro, but it remains unclear how physiologically relevant this activity is. HspB chaperones are involved in the development of numerous diseases; however, mechanisms underlying this contribution remain elusive.

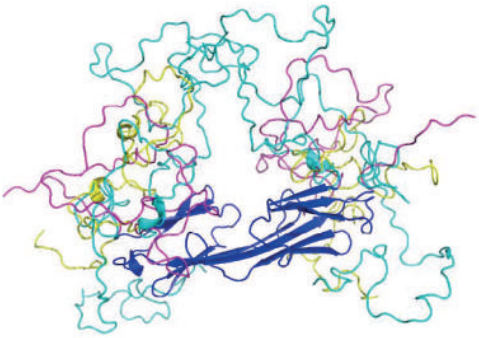


Figure 6.2: **Cartoon representation of the functional HspB6 dimer structure.** Ensemble of the structures is represented. Crystallographic core is coloured in blue, modeled extensions are coloured in magenta, cyan, and yellow. Reprinted from Journal of Structural Biology, vol. 185, no. 3, pp. 342–354., ©2014, with permission from Elsevier.

In general, sHsps are believed to act in an ATP-independent manner at an early stage of the stress response by recognizing partially unfolded proteins to prevent formation of toxic aggregates [29, 30]. Previous work demonstrated that sHsps have no capacity to promote refolding of bound substrates [31]. Failure of HspB chaperones function is implicated in the development of numerous diseases, such as cataract, cardiac myopathies, and neurodegerative diseases [32, 33]. In vivo evidence exists for an ability of HspBs to interact with and modulate the structure and stability of the cytoskeleton [34]. Due to the fact that they exist as large and dynamic oligomer assemblies that constantly interconvert in the cell depending on their posttranslational state, little is known about how exactly they interact with client proteins. Apart from the holdase activity, it has been suggested that sHsps can use transient low-affinity interactions to redirect unfolded-like protein intermediates back on the folding pathway to allow their spontaneous refolding

[35]. Moreover, it has also been demonstrated that sHsps are capable of collaboration with other chaperones, such as Hsp70 system together with Hsp110, to promote depolymerization of preformed αSyn fibrils [36].

HspB6 is an ubiquitous chaperone with the highest level of expression in muscles [37, 38]. It has rather small molecular weight and features C-terminal alpha-crystalline domain, however, its expression is not heat-inducible. As other family members, it has chaperone-like activity, or, in other words, prevents aggregation of (partially) unfolded proteins [39]. An atypical feature of HspB6 is that it does not assemble in big oligomers, but forms only dimers in the solution that represent a minimal active chaperoning unit [39] (Fig. 6.2).

6.1.3. Single-molecule methods

Even though an interaction of chaperones with aggregation-prone proteins is not completely understood yet, it is already apparent that a complete mechanistic picture of their interference with protein aggregation pathways would be very useful for various therapeutic applications. However, this task remains highly challenging due to extreme heterogeneity of precursors involved in the aggregation process, diversity of aggregation pathways, and transient nature of many key states and chaperone-protein interactions. Single molecule methods, which allow the characterization of intermediate species and reconstruction of reaction pathways while observing the state of a single protein-chaperone complex in real time, have thus a high potential being applied to this problem.

In this study, we used single molecule force spectroscopy in optical tweezers to study the structural details, stability and kinetics of aggregation of small αSyn oligomers, and to investigate the effect of various HspB chaperones on this process. We focused first on HspB6, since it retains chaperone activity at room temperature and does not form oligomers of various sizes in the solution that makes the chaperone-protein interaction more homogeneous and better controlled. Both folding pathway of αSyn monomers and aggregation of small αSyn oligomers have been studied previously by force spectroscopy that confirmed mostly unfolded state of αSyn as a monomer and formation of stably misfolded states in oligomers [40–43]. Numerous studies have investigated the interaction between αSyn and molecular chaperones (reviewed in [44]). However, the αSyn aggregation prevention by sHsps has never been studied at the single molecule scale, to the best of our knowledge. We used optical tweezers to grab a single αSyn molecule or an engineered repeat of four αSyn molecules linked head-to-tail and submitted them to cycles of stretching and relaxation. When relaxed, a protein has a chance to form a compacted structure that will be probed in subsequent stretching cycle. We performed these experiments in the isolation and in the presence of HspB6, and compared the extent of aggregation and the stability of compacted structures against force. We found that HspB does not promote the fully unfolded protein state to prevent its aggregation, similarly to other chaperones like SecB [45], but favours the compaction of αSyn and stabilizes interactions in the protein chain against force unfolding. Interestingly, long-range interactions (on a length scale bigger than the contour length of αSyn monomer) are partic-

ularly promoted by HspB6. These findings may shed new light on mechanistic aspects of aggregation prevention by sHsps in vivo.

6.2. Methods

6.2.1. Beads functionalization and sample preparation

We used commercially available polystyrene beads covalently coated with anti-digoxigenin (Anti-Dig) antibody (DIGP-20-2), and beads coated with neutravidin (NVP-20-5) from Spherotech Inc. We received hybrid dsDNA-αSyn construct (αSyn chimeras) with modified ends to allow easy and specific linking to the beads from the group of Prof. Michael Woodside (University of Alberta, Edmonton, Canada). It features either monomeric αSyn or engineered tetramer of αSyn, and was prepared by Woodside lab as described previously in [40].

First, both types of beads were washed 5 times in phoshpate-buffered saline with addition of 0.2% Tween (from now on termed PBT buffer). Next, 2 ul of Anti-Dig beads at concentration of ∼150 pM were mixed with 5 µl of 150 pM αSyn chimeras solution and incubated for 1 h on a rotary mixer at 4°C. After the incubation, the Anti-Dig beads were dissolved in 500 µl of MOPS buffer (pH = 7.5, 50 mM MOPS, 200 mM KCl) containing an oxygen scavenging mix (8 mU/µl glucose oxidase, 20 mU/µl catalase, and 0.01% w/v glucose). Neutravidin-coated (NTV) beads were stored at 4°C until use and were dissolved in the same buffer containing the oxygen scavengers cocktail.

6.2.2. Force spectroscopy

Stretching experiments were performed as described in chapter 4, section 4.1.4. Anti-Dig and αSyn coated bead was held on the micropipette, and the NTV-coated bead was held in the optical trap.

6.2.3. STEM imaging

αSyn was expressed in Escherichia coli strain BL21(DE3) using the pT7-7 expression plasmid and purified in the presence of 1 mM DTT as previously reported [46]. To prepare amyloid fibrils, 100 mM of αSyn monomer was placed in PBS buffer in eppendorf tubes and kept under constant shaking at 37°C at 1000 rpm in a Thermomixer. Subsequently, the resulting mature fibrils were filtered to remove non-aggregated monomers and diluted with MilliQ water to a final concentration of 10 mM. αSyn fibrils were allowed to interact with 1:1 ratio of HspB6 for 1 hour and then imaged using scanning transmission electron microscopy (STEM). 5 µl of 20 µM αSyn fibrils (in the presence and in the absence of HspB6) were adsorbed on 300 mesh formvar coated copper grids for 5 minutes and then washed 5 times with water. The grid was dried at 37°C and then transferred under vacuum into the STEM setup. Prior to acquiring the images, condenser stigmators were carefully adjusted to give a circular beam profile when the beam was viewed on the grids, and the beam was carefully centered and spread to produce uniform illumination over the field of view. The images were acquired using a FEI Verios 460 microscope operating at 20 kV electron beam energy.

6.3. αSyn tetramer aggregation in the isolation and in the presence of HspB6

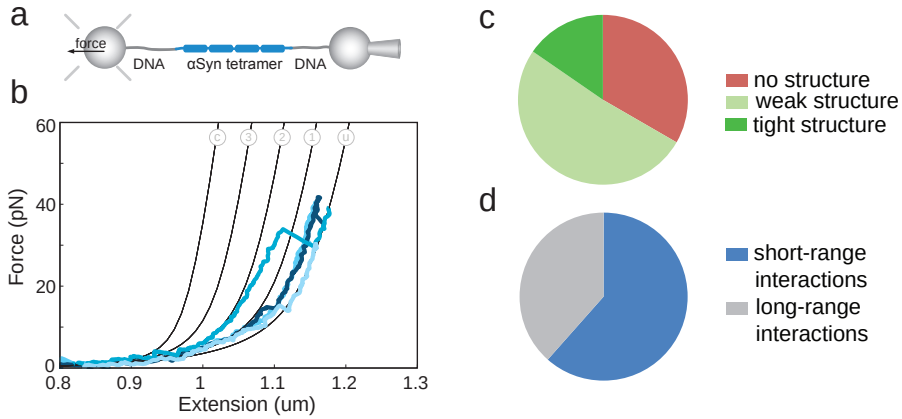


Figure 6.3: **Stretching-relaxation experiments with αSyn tetramer in the isolation. a,** Schematic drawing of the experimental assay. **b,** Force-extension data of αSyn tetramer. Subsequent stretching curves are shown in blue. Grey lines: theoretical compliance of DNA-protein construct in the states from compact (c) to fully unfolded (u). **c,** Events fraction during subsequent αSyn stretching. **d,** Fractions of short-range (within one αSyn domain, 55 nm) and long-range interactions in all compaction events during αSyn stretching.

We subjected αSyn tetramer to cycles of stretching and relaxation separated by a 5s waiting period when the protein is fully relaxed, in the absence of any chaperone (Fig. 6.3a). Typical force-extension stretching curves are shown in Fig. 6.3b. Although αSyn monomer is mostly unstructured, tetrameric αSyn demonstrated a wide range of compacted structures. After the protein was completely stretched out and relaxed for 5 s, some compaction was observed in 59% of cases (N = 32). Most of these structures did not sustain subsequent stretching (77% of all compaction events, $N = 26$). However, a minor fraction (23%) of compact protein structures were stable over multiple stretching-relaxation cycles and sustained forces over 60 pN (Fig. 6.3c). It indicated the formation of stable misfolded states. In order to estimate the protein length involved in the compaction, we measured the anticipated contour length change of αSyn upon unfolding by fitting the force-extension data to the WLCs with different αSyn folded fractions. In most of the compaction events (62%, $N = 26$), the involved protein length did not exceed the length of αSyn monomer, indicating the prevalence of short-range interactions (Fig. 6.3d).

Next, we performed stretching-relaxation experiments on αSyn tetramer in the presence of 5 μM of HspB6 in the experimental buffer (Fig. 6.4a). With the chaperone, the force-extension data changed significantly, indicating the interaction between the chaperone and the protein (Fig. 6.4b). Tight compacted structures formed more often (47% of compaction events, $N = 36$), however, the

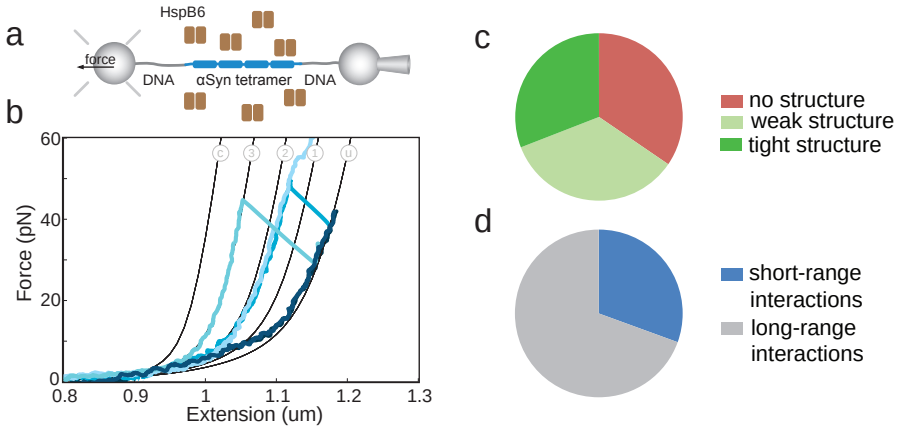


Figure 6.4: **Stretching-relaxation experiments with αSyn tetramer in the presence of HspB6. a,** Schematic drawing of the experimental assay. **b,** Force-extension data of αSyn tetramer. Subsequent stretching curves are shown in blue, they indicate increased compaction in the presence of HspB6. **c,** Events fraction during subsequent αSyn stretching. **d,** Fractions of short-range and long-range interactions in all compaction events during αSyn stretching.

total fraction of curves that became compacted did not change (Fig. 6.4c). In the presence of HspB6, partially compacted αSyn is more resistant to applied force than in the isolation (Fig. 6.5a). Apparently, HspB6 is capable of binding αSyn and stabilizing its partially compact states against forced unfolding. Shorter protein lengths are more populated in the presence of HspB6, indicating that the chaperone favours the compaction of αSyn (Fig. 6.5c). Moreover, longer fragments of αSyn are involved in the structure formation in the presence of HspB6 (Fig. 6.5b). The fraction of unfolding events smaller than αSyn monomer decreased to only 31% (Fig. 6.4d). It suggests that HspB6 stimulates αSyn compaction by promoting long-range interactions in the protein client.

To get more insight in the nature of short-range and long-range interactions in αSyn tetramer, we compared the mechanical stability of all unfolding events in a scatter plot (Fig. 6.5d). Most of the short-range interactions (smaller that αSyn monomer, blue-shaded area) disrupted at forces lower than 45 pN and were slightly stabilized by the chaperone (Fig. 6.5e). However, the stability of long-range interactions in αSyn depends more significantly on the presence of HspB6. In the absence of the chaperone, they sustain much lower forces, than with HspB6 (Fig. 6.5e). They are also more frequent, when HspB6 is present (Fig. 6.3d, 6.4d). These results indicate that HspB6 favours long-range interactions in αSyn and can stabilize them against applied force. However, they do not reveal the mechanism of HspB6-αSyn interactions underlying the unexpected effect of αSyn compaction.

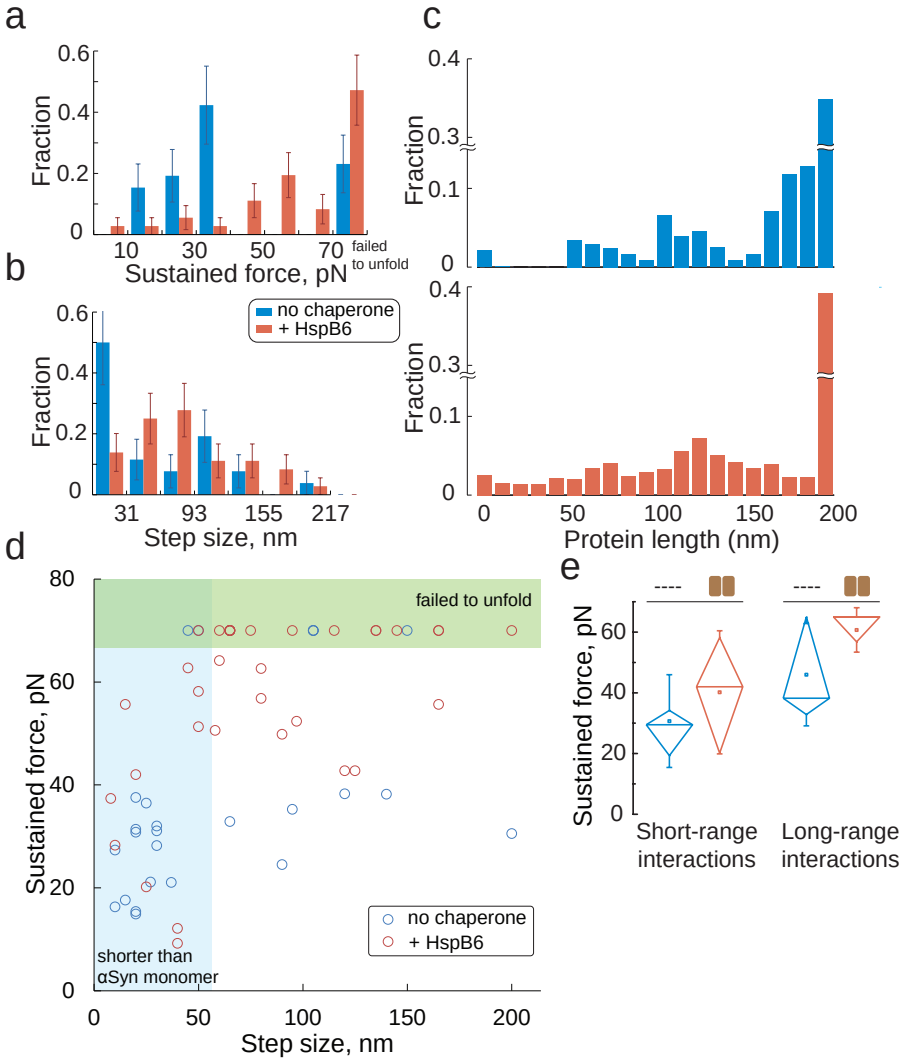


Figure 6.5: **Effect of HspB6 on αSyn folding pathways. a,** Force sustained by compacted αSyn structures distribution affected by HspB6. **b,** Step sizes statistics in the presence and in the absence of the chaperone. **c,** Absolute protein length histograms. Shorter protein lengths are favoured by HspB6. **d,** Scatter force-step size plot of structures observed in the isolation and with HspB6. **e,** Box-and-whisker plot of forces sustained by short-range and long-range interactions. Long-range interactions are more stabilized by HspB6.

6.4. Discussion and conclusion

Previous work suggests that HspB6 rather interacts with non-folded protein stretches, preventing their intermolecular aggregation upon heat shock [39, 47]. According to this opinion, our data look totally counterintuitive: HspB6 does not promote the fully unfolded protein state, but favours its compaction. However, mechanistic details of this interaction, stoichiometry of the chaperone-client complex, and reaction rates remain elusive so far. The effect of HspB6 observed in our data might be explained considering these mechanisms (Fig. 6.6):

1. HspB6 binding to two distant residues in a non-folded protein chain and clamping them.
2. HspB6 binding a non-folded protein fragment with additional contacts to neighboring protein stretches resulting in effective shortening of protein contour length that can be attributed to protein compaction in our assay.
3. HspB6 binding directly to structured fragments of the protein client and stabilizing partially folded protein states.

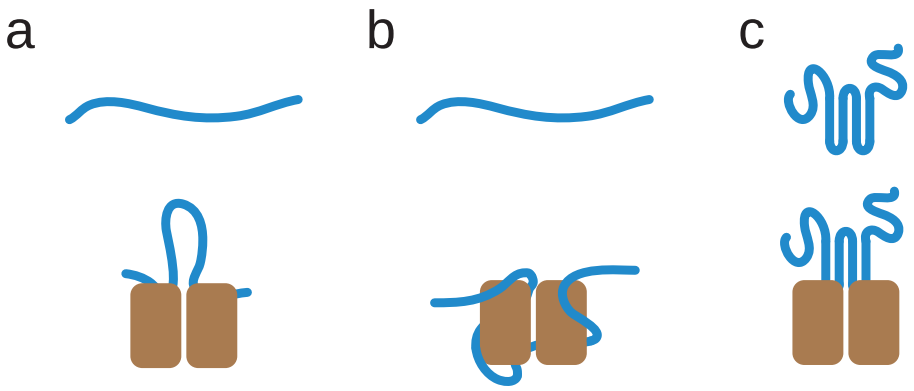


Figure 6.6: **Speculations for possible mechanism of HspB6 interaction with a client protein. a,** HspB6 binding and clamping of a non-folded chain that results in loop formation. **b,** HspB6 binding of a non-folded chain that wraps around a chaperone. **c,** Direct binding of Hsp B6 to folded protein structures.

To test these hypotheses, it is important to estimate the preference of HspB6 to structured or unstructured protein fragments. We performed a complementary bulk experiment to probe the interaction between HspB6 and αSyn fibrils, which are considered a model system for folded protein structure. The fibrils were pre-formed and purified from non-aggregated αSyn monomer. Next, they were allowed to interact with HspB6 and the sample was imaged using the scaning transmission electron microscopy (STEM; Fig. 6.7a,b). The images clearly demonstrated that the chaperones are able to interact with the mature αSyn fibril along

its complete length. It is consistent with our single molecule observations and indicates that HspB6 is capable of interaction with folded protein structure, since the fibrils were already formed when were exposed to the chaperone. It favours the third mechanism from discussed above.

Stretching-relaxation experiments with αSyn monomer can further address this question and shed more light on mechanistic aspects of this interaction, because in this case the aggregation will be omitted and the protein client would have much less structured fragments than αSyn tetramer. To test HspB6 binding to folded protein structures, stretching experiments with MBP, which refolds quickly [45], can be performed. Alternatively, the preference can be tested with αSyn tetramer by slightly modifying the protocol of the experiments with presented in the previous section. If the protein is exposed to the chaperones when it is not completely relaxed, but is kept under small force (5-10 pN), it is enough to prevent its spontaneous compaction and allow the chaperone to interact exclusively with non-folded protein. Such experiment will also reveal whether HspB6 binds non-folded protein fragments.

The hypothetical ability of HspB6 to bind folded protein structures would be consistent with recent suggestions of downstream ATP-dependent chaperones (foldases) binding folded proteins [48, 49]. Taking into account the collaboration of sHsps with foldases in aggregation prevention and subsequent refolding, it would suggest that the interaction with folded proteins is important for the function of the global chaperone network in the cell.

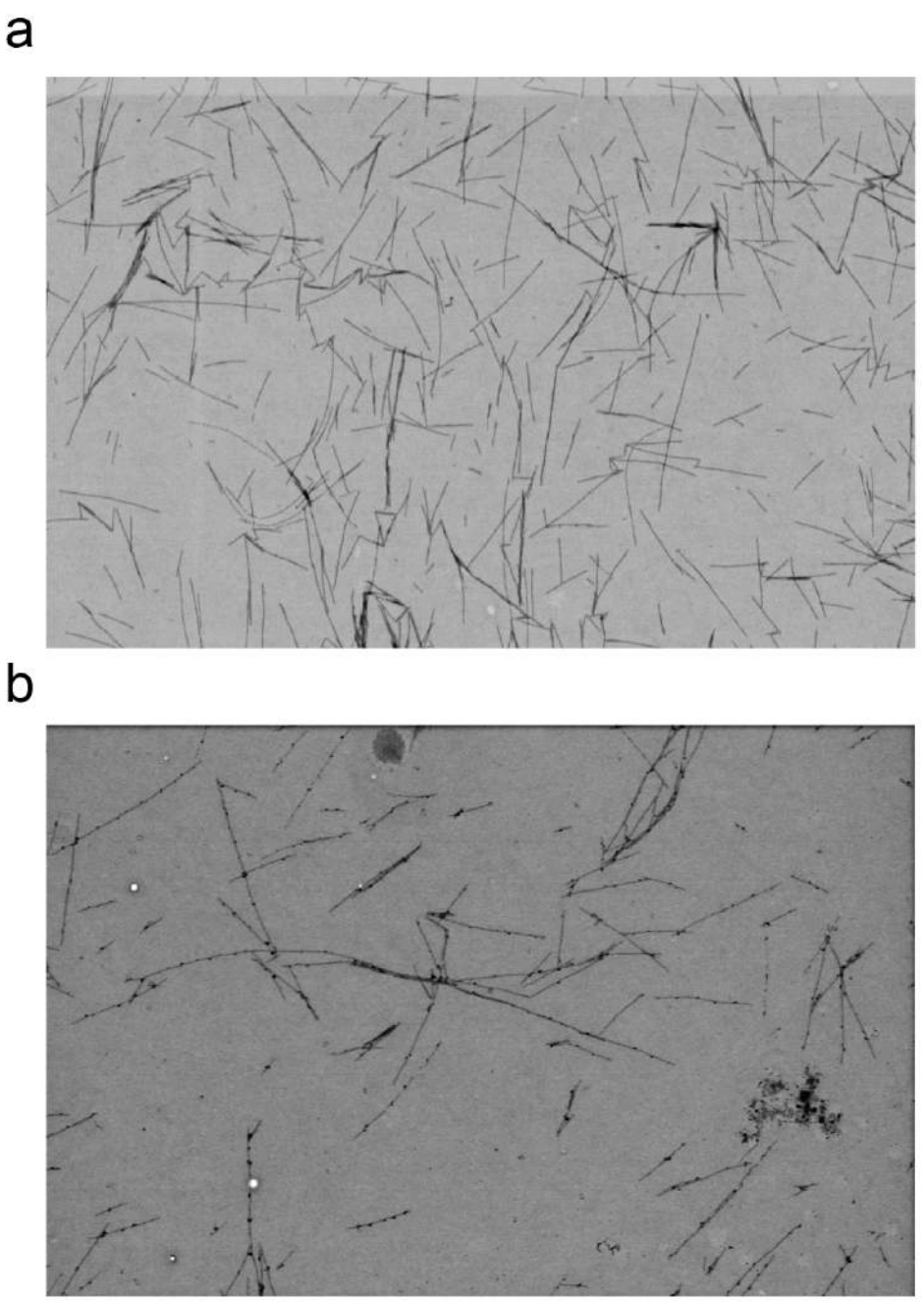


Figure 6.7: **STEM imaging of the interaction between HspB6 and mature αSyn fibrils. a,** Image of the fibrils without the chaperone. Fibrils look smooth and do not feature any irregularities. **b,** Image of the fibrils that were allowed to interact with HspB6 after they were formed. Dot-like structures along the fiber length are HspB6 dimers. Their size is consistent with the structures observed when HspB6 was imaged without the fibrils (data not shown). Magnification of both images is 12500x.

References

[1] M. P. Badtke, N. D. Hammer, and M. R. Chapman, *Functional amyloids signal their arrival*. Sci. Signal. **2**, pe43 (2009).

[2] J. M. Kenney, D. Knight, M. J. Wise, and F. Vollrath, *Amyloidogenic nature of spider silk*, Eur. J. Biochem. **269**, 4159 (2002).

[3] M. R. Chapman, L. S. Robinson, J. S. Pinkner, R. Roth, M. r. Hammar, S. Normark, and S. J. Hultgren, *Role of Escherichia coli Curli Operons in Directing Amyloid Fiber Formation*, Science **295**, 851 (2010).

[4] I. Cherny and E. Gazit, *Amyloids: Not only pathological agents but also ordered nanomaterials*, Angew. Chemie - Int. Ed. **47**, 4062 (2008).

[5] T. S. Ulmer, A. Bax, N. B. Cole, and R. L. Nussbaum, *Structure and dynamics of micelle-bound human alpha-synuclein*. J. Biol. Chem. **280**, 9595 (2005).

[6] H. a. Lashuel, C. R. Overk, A. Oueslati, and E. Masliah, *The many faces of α-synuclein: from structure and toxicity to therapeutic target*. Nat. Rev. Neurosci. **14**, 38 (2013).

[7] T. P. J. Knowles, M. Vendruscolo, and C. M. Dobson, *The amyloid state and its association with protein misfolding diseases*. Nat. Rev. Mol. Cell Biol. **15**, 384 (2014).

[8] A. W. P. Fitzpatrick, G. T. Debelouchina, M. J. Bayro, D. K. Clare, M. a. Caporini, V. S. Bajaj, C. P. Jaroniec, L. Wang, V. Ladizhansky, S. a. Müller, C. E. MacPhee, C. a. Waudby, H. R. Mott, A. De Simone, T. P. J. Knowles, H. R. Saibil, M. Vendruscolo, E. V. Orlova, R. G. Griffin, and C. M. Dobson, *Atomic structure and hierarchical assembly of a cross-β amyloid fibril*. Proc. Natl. Acad. Sci. U. S. A. **110**, 5468 (2013).

[9] A. J. Baldwin, T. P. J. Knowles, G. G. Tartaglia, A. W. Fitzpatrick, G. L. Devlin, S. L. Shammas, C. A. Waudby, M. F. Mossuto, S. Meehan, S. L. Gras, J. Christodoulou, S. J. Anthony-cahill, P. D. Barker, M. Vendruscolo, and C. M. Dobson, *Metastability of native proteins and the phenomenon of amyloid formation*, J. Am. Chem. Soc. , 14160 (2011).

[10] E. Gazit, *The "correctly folded" state of proteins: Is it a metastable state?* Angew. Chemie - Int. Ed. **41**, 257 (2002).

[11] T. Lührs, C. Ritter, M. Adrian, D. Riek-Loher, B. Bohrmann, H. Döbeli, D. Schubert, and R. Riek, *3D structure of Alzheimer's amyloid-beta(1-42) fibrils*. Proc. Natl. Acad. Sci. U. S. A. **102**, 17342 (2005).

[12] C. Wasmer, A. Lange, H. Van Melckebeke, A. B. Siemer, R. Riek, and B. H. Meier, *Amyloid fibrils of the HET-s(218-289) prion form a beta solenoid with a triangular hydrophobic core*. Science **319**, 1523 (2008).

[13] D. M. Walsh, I. Klyubin, J. V. Fadeeva, W. K. Cullen, R. Anwyl, M. S. Wolfe, M. J. Rowan, and D. J. Selkoe, *Naturally secreted oligomers of amyloid beta protein potently inhibit hippocampal long-term potentiation in vivo.* Nature **416**, 535 (2002).

[14] H. A. Lashuel and P. T. Lansbury, *Are amyloid diseases caused by protein aggregates that mimic bacterial pore-forming toxins?* Q. Rev. Biophys. **39**, 167 (2006).

[15] N. Cremades, S. I. a. Cohen, E. Deas, A. Y. Abramov, A. Y. Chen, A. Orte, M. Sandal, R. W. Clarke, P. Dunne, F. a. Aprile, C. W. Bertoncini, N. W. Wood, T. P. J. Knowles, C. M. Dobson, and D. Klenerman, *Direct observation of the interconversion of normal and toxic forms of α-synuclein,* Cell **149**, 1048 (2012).

[16] C. Lavedan, *The synuclein family,* Genome Res. **8**, 871 (1998).

[17] M. R. Cookson, *The biochemistry of Parkinson's disease.* Annu. Rev. Biochem. **74**, 29 (2005).

[18] V. N. Uversky, *A protein-chameleon: conformational plasticity of alpha-synuclein, a disordered protein involved in neurodegenerative disorders.* J. Biomol. Struct. Dyn. **21**, 211 (2003).

[19] P. H. Weinreb, W. Zhen, A. W. Poon, K. a. Conway, and P. T. Lansbury, *NACP, a protein implicated in Alzheimer's disease and learning, is natively unfolded,* Biochemistry **35**, 13709 (1996).

[20] W. S. Davidson, a. Jonas, D. F. Clayton, and J. M. George, *Stabilization of alpha-synuclein secondary structure upon binding to synthetic membranes.* J. Biol. Chem. **273**, 9443 (1998).

[21] T. Bartels, J. G. Choi, and D. J. Selkoe, *α-Synuclein occurs physiologically as a helically folded tetramer that resists aggregation.* Nature **477**, 107 (2011).

[22] W. Wang, I. Perovic, J. Chittuluru, a. Kaganovich, L. T. T. Nguyen, J. Liao, J. R. Auclair, D. Johnson, a. Landeru, a. K. Simorellis, S. Ju, M. R. Cookson, F. J. Asturias, J. N. Agar, B. N. Webb, C. Kang, D. Ringe, G. a. Petsko, T. C. Pochapsky, and Q. Q. Hoang, *A soluble -synuclein construct forms a dynamic tetramer,* Proc. Natl. Acad. Sci. **108**, 17797 (2011).

[23] B. Fauvet, M. K. Mbefo, M. B. Fares, C. Desobry, S. Michael, M. T. Ardah, E. Tsika, P. Coune, M. Prudent, N. Lion, D. Eliezer, D. J. Moore, B. Schneider, P. Aebischer, O. M. El-Agnaf, E. Masliah, and H. a. Lashuel, *α-Synuclein in central nervous system and from erythrocytes, mammalian cells, and Escherichia coli exists predominantly as disordered monomer,* J. Biol. Chem. **287**, 15345 (2012).

[24] V. S. Kokhan, M. A. Afanasyeva, and G. I. Van'kin, *α-Synuclein knockout mice have cognitive impairments.* Behav. Brain Res. **231**, 226 (2012).

[25] S. Chandra, G. Gallardo, R. Fernández-Chacón, O. M. Schlüter, and T. C. Südhof, *α-Synuclein cooperates with CSPα in preventing neurodegeneration,* Cell **123**, 383 (2005).

[26] I. G. McKeith, D. Galasko, K. Kosaka, E. K. Perry, D. W. Dickson, L. A. Hansen, D. P. Salmon, J. Lowe, S. S. Mirra, E. J. Byrne, G. Lennox, N. P. Quinn, J. A. Edwardson, P. G. Ince, C. Bergeron, A. Burns, B. L. Miller, S. Lovestone, D. Collerton, E. N. Jansen, C. Ballard, R. A. de Vos, G. K. Wilcock, K. A. Jellinger, and R. H. Perry, *Consensus guidelines for the clinical and pathologic diagnosis of dementia with Lewy bodies (DLB): report of the consortium on DLB international workshop.* Neurology **47**, 1113 (1996).

[27] I. Horvath, C. F. Weise, E. K. Andersson, E. Chorell, M. Sellstedt, C. Bengtsson, A. Olofsson, S. J. Hultgren, M. Chapman, M. Wolf-Watz, F. Almqvist, and P. Wittung-Stafshede, *Mechanisms of protein oligomerization: Inhibitor of functional amyloids templates a-synuclein fibrillation,* J. Am. Chem. Soc. **134**, 3439 (2012).

[28] K. Furukawa, M. Matsuzaki-Kobayashi, T. Hasegawa, A. Kikuchi, N. Sugeno, Y. Itoyama, Y. Wang, P. J. Yao, I. Bushlin, and A. Takeda, *Plasma membrane ion permeability induced by mutant a-synuclein contributes to the degeneration of neural cells,* J. Neurochem. **97**, 1071 (2006).

[29] M. Haslbeck, T. Franzmann, D. Weinfurtner, and J. Buchner, *Some like it hot: the structure and function of small heat-shock proteins.* Nat. Struct. Mol. Biol. **12**, 842 (2005).

[30] B. Bukau, J. Weissman, and A. Horwich, *Molecular chaperones and protein quality control.* Cell **125**, 443 (2006).

[31] M. Ehrnsperger, S. Gräber, M. Gaestel, and J. Buchner, *Binding of non-native protein to Hsp25 during heat shock creates a reservoir of folding intermediates for reactivation,* EMBO J. **16**, 221 (1997).

[32] L. E. Goldstein, J. A. Muffat, R. A. Cherny, R. D. Moir, M. H. Ericsson, X. Huang, C. Mavros, J. A. Coccia, K. Y. Faget, K. A. Fitch, C. L. Masters, R. E. Tanzi, L. T. Chylack, and A. I. Bush, *Cytosolic beta-amyloid deposition and supranuclear cataracts in lenses from people with Alzheimer's disease.* Lancet (London, England) **361**, 1258 (2003).

[33] N. S. Rajasekaran, P. Connell, E. S. Christians, L. J. Yan, R. P. Taylor, A. Orosz, X. Q. Zhang, T. J. Stevenson, R. M. Peshock, J. a. Leopold, W. H. Barry, J. Loscalzo, S. J. Odelberg, and I. J. Benjamin, *Human αB-Crystallin Mutation Causes Oxido-Reductive Stress and Protein Aggregation Cardiomyopathy in Mice,* Cell **130**, 427 (2007).

[34] C. M. Dreiza, C. M. Brophy, P. Komalavilas, E. J. Furnish, L. Joshi, M. a. Pallero, J. E. Murphy-Ullrich, M. von Rechenberg, Y.-s. J. Ho, B. Richardson,

N. Xu, Y. Zhen, J. M. Peltier, and A. Panitch, *Transducible heat shock protein 20 (HSP20) phosphopeptide alters cytoskeletal dynamics.* FASEB J. **19**, 261 (2005).

[35] M. Kulig and H. Ecroyd, *The small heat-shock protein αB-crystallin uses different mechanisms of chaperone action to prevent the amorphous versus fibrillar aggregation of α-lactalbumin.* Biochem. J. **448**, 343 (2012).

[36] M. L. Duennwald, A. Echeverria, and J. Shorter, *Small heat shock proteins potentiate amyloid dissolution by protein disaggregases from yeast and humans.* PLoS Biol. **10**, e1001346 (2012).

[37] B. E. Cross, H. M. O'Dea, and D. J. MacPhee, *Expression of small heat shock-related protein 20 (HSP20) in rat myometrium is markedly decreased during late pregnancy and labour,* Reproduction **133**, 807 (2007).

[38] K. a. Huey, J. S. Thresher, C. M. Brophy, and R. R. Roy, *Inactivity-induced modulation of Hsp20 and Hsp25 content in rat hindlimb muscles,* Muscle and Nerve **30**, 95 (2004).

[39] S. D. Weeks, E. V. Baranova, M. Heirbaut, S. Beelen, A. V. Shkumatov, N. B. Gusev, and S. V. Strelkov, *Molecular structure and dynamics of the dimeric human small heat shock protein HSPB6,* J. Struct. Biol. **185**, 342 (2014).

[40] K. Neupane, A. Solanki, I. Sosova, M. Belov, and M. T. Woodside, *Diverse metastable structures formed by small oligomers of α-synuclein probed by force spectroscopy.* PLoS One **9**, e86495 (2014).

[41] M. Sandal, F. Valle, I. Tessari, S. Mammi, E. Bergantino, F. Musiani, M. Brucale, L. Bubacco, and B. Samorì, *Conformational equilibria in monomeric α-synuclein at the single-molecule level,* PLoS Biol. **6**, 0099 (2008).

[42] A. V. Krasnoslobodtsev, I. L. Volkov, J. M. Asiago, J. Hindupur, J.-C. Rochet, and Y. L. Lyubchenko, *α-Synuclein misfolding assessed with single molecule AFM force spectroscopy: effect of pathogenic mutations.* Biochemistry **52**, 7377 (2013).

[43] J. Yu, S. Malkova, and Y. L. Lyubchenko, *α-Synuclein Misfolding: Single Molecule AFM Force Spectroscopy Study,* J. Mol. Biol. **384**, 992 (2008).

[44] D. Cox, J. a. Carver, and H. Ecroyd, *Preventing α-synuclein aggregation: The role of the small heat-shock molecular chaperone proteins,* Biochim. Biophys. Acta - Mol. Basis Dis. **1842**, 1830 (2014).

[45] P. Bechtluft, R. G. H. van Leeuwen, M. Tyreman, D. Tomkiewicz, N. Nouwen, H. L. Tepper, A. J. M. Driessen, and S. J. Tans, *Direct observation of chaperone-induced changes in a protein folding pathway.* Science **318**, 1458 (2007).

[46] M. E. van Raaij, I. M. Segers-Nolten, and V. Subramaniam, *Quantitative Morphological Analysis Reveals Ultrastructural Diversity of Amyloid Fibrils from α-Synuclein Mutants,* Biophysical Journal **91**, L96 (2006).

[47] O. V. Bukach, A. S. Seit-Nebi, S. B. Marston, and N. B. Gusev, *Some properties of human small heat shock protein Hsp20 (HspB6).* European journal of biochemistry / FEBS **271**, 291 (2004).

[48] R. Schlecht, A. H. Erbse, B. Bukau, and M. P. Mayer, *Mechanics of Hsp70 chaperones enables differential interaction with client proteins.* Nat. Struct. Mol. Biol. **18**, 345 (2011).

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

7

Trehalose facilitates DNA melting: a single-molecule optical tweezers study

Using optical tweezers, here we show that the overstreching transition force of double stranded DNA (dsDNA) is lowered significantly by the addition of the disaccharide trehalose as well as certain polyol osmolytes. This effect is found to depend linearly on the logarithm of the trehalose concentration. We propose an entropic driving mechanism for the experimentally observed destabilization of dsDNA that is rooted in the higher affinity of the DNA bases for trehalose than for water, which promotes base exposure and DNA melting. Molecular dynamics simulation reveals the direct interaction of trehalose with nucleobases. Experiments with other osmolytes confirm that the extent of dsDNA destabilization is governed by the ratio between polar and apolar fractions of an osmolyte.

7.1. Introduction

The disaccharide trehalose (Figure 7.1a) is a metabolite found in a wide range of organisms from different domains of life. It has been found that trehalose is synthesized specifically upon various stresses [1] and is thought to protect various cellular components such as DNA, proteins and membranes from loss of activity caused by heat, cold, dessication, dehydration, and oxidation [2, 3]. *In vitro* experiments showed that trehalose can either stabilize or destabilize biomolecules depending on the substrate and the concentration of trehalose. Stabilization was observed with double-stranded DNA (dsDNA) in dry state, that is, at very high concentration of trehalose [3] and with folded proteins [4, 5]. Unfolded proteins were reported to be stabilized by trehalose [6, 7]. On the other hand, trehalose can destabilize folded proteins [8] and dsDNA in solution [9]. The extent of DNA destabilization due to trehalose depends on DNA length and on percent guanine-cytosine (GC) content as concluded from systematic thermal experiments on short duplexes [10].

While the evidence for the influence of trehalose on proteins and nucleic acids is compelling, our understanding of the underlying mechanism is incomplete. Various mechanisms have been considered to explain the effect of trehalose on biomolecules [1], including physically shielding parts of biomolecules and removing the hydration waters by replacing (direct interaction with biomolecule) or sequestering them (no direct interaction with biomolecule). Trehalose has been suggested to form hydrogen bonds (H-bonds) with the backbone phosphates and to dehydrate dsDNA molecules [3]. In these scenarios the repulsive electrostatic interactions between the phosphate groups may cause DNA melting [3, 11]. One could also imagine a scenario in which trehalose competes with the complementary nucleobase in forming H-bonds and thereby hindering base-pairing and shielding single-stranded DNA (ssDNA).

Understanding the effect of trehalose and other osmolytes on the physicochemical properties of biomolecules is relevant not only for the biological context but also for technological applications [12]. Trehalose is commonly used in pharmaceutical industries and is contained in many therapeutic, cosmetic and food products [13]. It is also used in chemical industries for example to increase the yield in mass spectrometry [14] and DNA amplification. Trehalose increases the yield of polymerase chain reaction (PCR) by lowering the melting temperature of DNA [15].

Here we address this problem by mechanically inducing melting of single dsDNA molecules with optical tweezers. Mechanical properties of biomolecules can be studied in real time using single molecule techniques such as optical tweezers. This method provides sub-piconewton resolution for force and sub-nanometer resolution in length scale. Optical tweezers have been extensively used to study the melting transition of single DNA molecules as well as the interaction of DNA with proteins and drugs [16–20] and are thus well suited to investigate the effect of trehalose on the melting of DNA. We use optical tweezers to grab the single DNA molecule at its ends and to induce its melting mechanically by stretching with different concentrations of trehalose added to the reaction buffer. We show that

trehalose lowers the overstretching "transition"[1] force of dsDNA. To explain the experimentally observed destabilization by trehalose, we looked at the thermodynamics of the system by adapting the model of DNA melting from Williams et al. [21] We also performed optical tweezers experiments on the effect of other molecules, such as: 1-propanol, 1-hexanol and glycerol, on DNA mechanics depending on the solubility of the additive. To obtain structural insights into the trehalose-DNA interaction we performed molecular dynamics (MD) simulations of ssDNA and dsDNA in the trehalose solution.

7.2. Methods

7.2.1. Functionalization of polystyrene beads

Carboxylated polystyrene beads (diameter ~ 1.8 µm, Spherotech) were covalently attached to anti-digoxigenin (Anti-Dig) antibody (Roche Diagnostics) via carbodiimide reaction (PolyLink Protein Coupling Kit, Polysciences Inc.). Briefly, the beads were washed and then mixed with freshly prepared 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide and the antibody, and mixture was incubated for 3 hours. Streptavidin-coated beads (diameter ~ 1.8 µm, Spherotech) were stored at 4 °C until use. 2553 base pairs DNA handles were prepared and functionalized with digoxigenin (Dig) and biotin at 5'-ends of both strands as described in [22]. DNA-coated microspheres were made by mixing ~70 ng of biotin-dsDNA-Dig and 1 µl Dig-coated beads in 10 µl HMK (50 mM HEPES, pH 7.6, 5 mM MgCl2, 100 mM KCl) buffer. After a 30 minute incubation on a rotary mixer (4 °C), the beads were dissolved in 400 µl HMK buffer for use in optical tweezers experiments. Trehalose solution was made by dissolving trehalose dihydrate (Sigma Aldrich) in HMK buffer.

7.2.2. Force spectroscopy

Pulling experiments were performed at room temperature using a custom made optical tweezers apparatus (for details see [22, 23]). Forces acting on the trapped bead lead to deflections of the laser beam, which were recorded with a quadrant photodiode at 50 Hz. Trap stiffness and sensitivity were measured to be 169±24 pN μm^{-1} and 2.74±0.24 V μm^{-1} respectively. The presence of trehalose barely affected the trap stiffness within the range of concentrations used (up to 0.1 M). A piezo stage (Physik Instrumente) was used to move the flow chamber and micropipette at a speed of 50 nm s^{-1}.

A tether (dsDNA construct) is immobilized between two beads. Digoxigenin-dsDNA-biotin is immobilized between an anti-digoxigenin coated bead and a streptavidin coated bead via 5'-ends of its two strands. These connections enable easy and specific linking of both termini to the two types of beads used.

Table 7.1: Parameters for individual MD simulations. DNA sequences used in the simulations:
dsAT: 5′-ATATATATATATATATATATATATATAT-3′,
dsGC: 5′-GCGCGCGCGCGCGCGCGCGCGCGCGC-3′,
ds50%GC: 5′-GAGATGCTAACCCTGATCGCTGATTCCTTGGAC-3′,
ssAT: 5′-ATATATATATATATATATATATATATAT-3′,
ssGC: 5′-GCGCGCGCGCGCGCGCGCGCGCGCGC-3′.

DNA Sequence	Solvent	Trehalose	Water	Na^+	Duration (ns)
dsAT	Pure water	–	11673	58	95
	0.5 M trehalose	108	11565	58	45
dsGC	Pure water	–	8417	58	108
	0.5 M trehalose	78	8339	58	70
ds50%GC	Pure water	–	9929	64	156
	0.5 M trehalose	92	9837	64	70
ssAT	Pure water	–	8941	29	168
	0.5 M trehalose	83	8912	29	66
ssGC	Pure water	–	8947	29	145
	0.5 M trehalose	83	8918	29	65

7.2.3. Thermodynamic analysis

We adapted the model from Williams et al. [21] to understand the physical origin of dsDNA destabilization induced by trehalose. According to this model, the overstretching transition of dsDNA can be seen as force-induced melting which implies that the transition force should be dependent on the temperature. In the force (F)-temperature (T) space, at the boundary between the dsDNA regime and the ssDNA regime[2] the difference in total free energy between the two states is zero [21]:

$$\Delta G_{tot}(F,T) = 0 \tag{7.1}$$

Thus, the melting transition does not change the free energy of the system and the two states co-exist at the boundary. We assume that the total transition free energy can be written as the sum of the temperature-dependent transition free energy in the absence of force $\Delta G(T)$ and a force-dependent term $\Delta\Phi(F)$:

$$\Delta G_{tot}(F,T) = \Delta G(T) + \Delta\Phi(F) \tag{7.2}$$

$$\Delta G_{tot}(F,T) = \Delta S(T_m^0)(T_m^0 - T) - \frac{\Delta C_p}{2}\frac{(T-T_m^0)^2}{T_m^0} + \Delta\Phi(F) \tag{7.3}$$

where $\Delta S(T)$ denotes the total transition entropy of the system, ΔC_p is the change in the heat capacity of the system during the transition, $\Delta\Phi(F)$ is a one-dimensional

[1]This is not a true thermodynamic transition, but we used this term since it is widely used in the field. The exact term should be ”crossover”.

[2]Strictly speaking, state of DNA after transition depends on a variety of factors. For more details see Discussion. It does not affect the validity of method we used, thus we will refer to overstretched DNA as ssDNA for simplicity.

thermodynamic potential analogous to the Gibbs free energy and T_m^0 is the temperature at which the melting transition occurs in the absence of force. At room temperature, stretching of the dsDNA from its two ends using optical tweezers will result in melting at a certain force which depends on the concentration of trehalose. At the melting transition we have:

$$\Delta S(T_m^0)(T - T_m^0) + \frac{\Delta C_p}{2}\frac{(T - T_m^0)^2}{T_m^0} = \Delta\Phi(F) \tag{7.4}$$

The right side of the identity can be derived directly from integration of the force-extension curve fits.

Force-extension optical tweezers data for dsDNA behaviour was fitted to the generalized discrete worm-like chain model (WLC) proposed by Manghi et al.[24] According to this model, the following relation between extension and force acting on the molecule is valid for pure dsDNA (equation 47 in Manghi et al.[24]):

$$\begin{aligned}\frac{z}{a_{Ds}N} = (1 + \frac{F}{\tilde{E}_{Ds}} - \frac{1}{2\alpha_{Ds}})\varphi_{Ds} + \gamma(1 - \frac{1}{2\alpha_{Ss}})\varphi_{Ss} \\ + \frac{\langle\sigma_i\sigma_{i+1}\rangle - 1}{4} \times (\frac{1}{2\alpha_{Ds}}\frac{\kappa_{Ds} - \kappa_{DsSs}}{\kappa_{DsSs} + F/2 + \alpha_{Ds}} \\ + \frac{\gamma}{2\alpha_{Ss}}\frac{\kappa_{Ss} - \kappa_{DsSs}}{\kappa_{DsSs} + \gamma F/2 + \alpha_{Ss}})\end{aligned} \tag{7.5}$$

where indices 'Ds' and 'Ss' are used, respectively, for dsDNA and ssDNA states. F denotes force, z - extension, N - number of residues in DNA, a - monomer size, $\tilde{E}$ - adimensional stretching modulus, $\gamma = \frac{a_{Ss}}{a_{Ds}}$, $\alpha_{Ds} = \sqrt{\kappa_{Ds}F + (\frac{F}{2})^2}$, $\alpha_{Ss} = \sqrt{\kappa_{Ss}\gamma F + (\frac{\gamma F}{2})^2}$, κ - adimensional bending modulus for dsDNA and ssDNA states. $\varphi(F)$ denotes the fraction of base pairs in certain state of DNA. σ_i is an Ising variable that denotes the state (broken or unbroken) of each basepair in DNA. $\langle\sigma_i\sigma_{i+1}\rangle$ is the two-point correlation function. Here we used a_{Ds} = 0.34 nm, a_{Ss} = 0.62 nm, $\tilde{E}_{Ds}$ =101, γ = 1.829, κ_{Ds} = 147, κ_{Ss} = 1.28.

The part of the data after overstretching transition force was fitted to the discrete Marko-Siggia WLC interpolation with non-linear elasticity [24]. It describes the elastic response of ssDNA with this relation (equation 16 in Manghi et al.[24]):

$$\begin{aligned}\frac{aF}{k_BT} = \frac{z}{L}(1 + U_{nl}(F))(3\frac{1 - u(\kappa)}{1 + u(\kappa)} - \frac{1}{\sqrt{1 + 4\kappa^2}}) \\ + \sqrt{\frac{1}{[1 - z(1 + U_{nl}(F)/L]^2} + 4\kappa^2} - \sqrt{1 + 4\kappa^2})\end{aligned} \tag{7.6}$$

where F denotes force, z - extension, L - contour length, κ - adimensional bending modulus, and a - monomer size. $u(\kappa)$ is the Langevin function defined as $u(\kappa) = \coth\kappa - 1/\kappa$, and $U_{nl}(F)$ is a non-linear stretching polynomial. Here we used L = 920 nm, κ = 1.28, a = 0.665 nm.

By directly integrating the WLC curves, one can find the force-dependent free energy change $\Delta\Phi(F)$ associated with the melting transition of the DNA:

$$\Delta\Phi(F) = \int_0^F [z_{ss}(F') - z_{ds}(F')]\, dF' \tag{7.7}$$

Integrations for the Gibbs free energy were performed over a 0 – 100 pN interval to find $\Delta\Phi$(F). Using a melting temperature of T_m = 100 °C from base-stacking calculations [19], room temperature of T = 21 °C and $\Delta S(T_m)$ = 24.7 cal/mol K bp, one can thus numerically solve the overstretching force associated with each value of ΔC_p for different trehalose concentrations using equation 3.

7.2.4. MD simulations

Five different molecules, which are 30-mer AT (both ds and ssDNA), 30-mer GC (both ds and ssDNA) and 33-mer dsDNA with GC-content of 50% were studied in the molecular dynamics simulation. They were all solvated in either pure water or 0.5M trehalose aqueous solutions, namely, resulting in an overall of ten simulation systems. All MD simulations were conducted using AMBER11 suite of programs [25]. The SPCE model [26] was used to describe the explicit water solvent. Nucleic acid and trehalose force field parameters were taken from the AMBER ff10 parameter set [27] and GLYCAM06 [28], respectively. In each system, the canonical B form of DNA was used as the initial structure and then immersed into a cubic box containing water and trehalose. The detailed parameters of the simulation systems are listed in Table 7.1.

For each system, the simulation procedure included the energy minimization, heating up and the equilibration using the NPT ensemble. First, the system was minimized through 500 steps of steepest descent minimization and the following 500 steps of conjugate gradient minimization with DNA being fixed using harmonic restraints. Then the restraints on DNA were released and the system was minimized using 1000 steps of the steepest descent and then 1500 steps of conjugate gradient minimization. In order to relax the system further (especially the solvent molecules), it was heated to 360 K first and equilibrated at this temperature for 300 ps, then it was cooled down to 300 K. Finally, production runs were performed and the temperature of all systems was maintained at 300 K using the Langevin dynamics with a friction coefficient of 5 ps^{-1} . The pressure of the system was adjusted to 1 atm by Berendsen weak-coupling [29] algorithm with relaxation time constants of 2 ps.

In all simulations, the SHAKE algorithm [30] was used in order to restraint all covalent bonds involving hydrogen. Therefore, all dynamics employed an integral time step of 2 fs. Periodic boundary conditions were used and the particle-mesh Ewald(PME) [31] with a direct space cutoff of 10 $\AA$ was utilized to treat long-range electrostatic interactions.

The figures of DNA-trehalose interaction were made in PyMOL [32].

7.3. Results

7.3.1. Single molecule mechanics of dsDNA

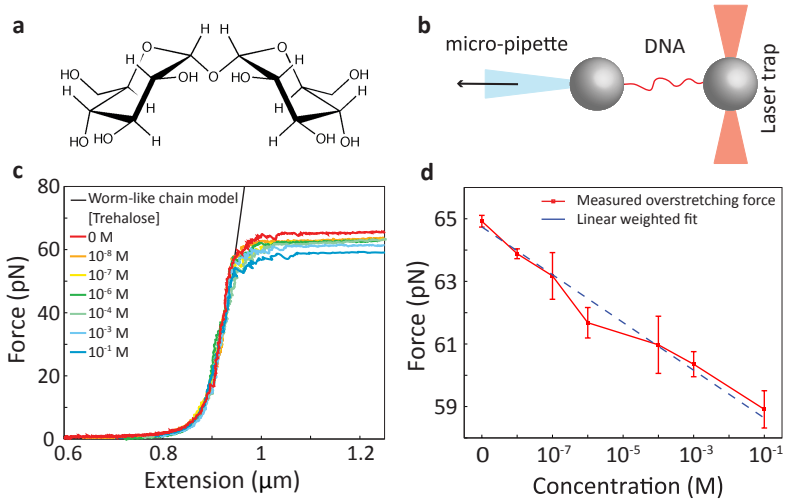


Figure 7.1: **(a)** Chemical structure of trehalose. **(b)** A schematic drawing of the optical tweezers assay. **(c)** A force applied to DNA tether is plotted versus its extension. In the absence of trehalose the overstretching transition occurs at about 65 pN. The overstretching force decreases upon addition of trehalose. The black line is a worm-like chain model for the double stranded DNA. **(d)** Overstretching force as a function of the log of trehalose concentration. The dashed line is a linear weighted fit ($R^2 = 0.9718$). The errorbars indicate the standard deviations of the forces from various measurements.

We used optical tweezers to measure the mechanical response of dsDNA (2553 base pairs, GC content ~ 51%) to applied forces. In these experiments, the DNA was grabbed at its ends by immobilizing the end of one strand on each side to micron-sized spheres, thus leaving the DNA torsionally unconstrained and end-opened (Figure 7.1b). A typical force-extension stretching curve is presented in Figure 7.1c (red curve). In the stretching curve, when the force on the DNA increases, overstretching occurred at about 65 pN as reported previously [33, 34]. During the overstretching transition the length of the DNA increased, while force barely changed, indicating alteration of structure. We note that the force during the overstretching was approximated to be constant. The overstretching force for each stretching curve was defined as an average force value for the transition region between the fitting curves of dsDNA and ssDNA.

The relaxation curves, when the beads were brought close together again, exhibited significant hysteresis [19] (Extended Data Figure ??). A consecutive stretching after 5 s waiting time at zero force resulted again in a similar curve followed by overstretching which indicates that the single-stranded fragments of DNA annealed to regenerate the double stranded conformation.

7.3.2. DNA stretching in the presence of trehalose

The trehalose experiments were performed in three steps. First the DNA was stretched and relaxed multiple times in the absence of trehalose in HMK buffer. Next, HMK buffer containing trehalose was flown into the sample chamber. Finally, the DNA was stretched and relaxed multiple times to assess the effect of trehalose. Without the disaccharide, the DNA compliance at low forces is in line with a worm-like chain model [35], and at force about 65 pN the extended DNA molecule undergoes the overstretching transition. When trehalose is added to the system, DNA behavior does not change at low forces. However, the magnitude of the overstretching force was significantly affected by the presence of trehalose (Figure 7.1c), and decreased by nearly 5 pN over the range of trehalose concentrations studied (up to 100 mM). Results from experiments (Figure 7.1d) show that the overstretching force decreases almost linearly with the log of the concentration of trehalose. However, relaxation curves of DNA did not exhibit more hysteresis in the presence of trehalose (data not shown).

7.3.3. Effect of other amphiphiles on dsDNA mechanics

While it has been shown that different osmolytes have different effects on the stability of the substrate [6], a detailed mechanism has not yet been determined. For instance, it is unclear how specific structural features of the osmolyte give rise to the observed effects. In order to investigate the importance of the chemical structure of osmolytes on the effect they have on dsDNA mechanics, we have performed pulling experiments under similar conditions, adding solutes that exhibit different ratios of hydrophobic and hydrophillic groups (Figure 7.2a-b).

For the same concentration of osmolyte (1 mM), we observed almost no effect on the overstretching force with 1-propanol and 1-hexanol, while a decrease in order of 4 pN was detected with glycerol (Figure 7.2c).

7.3.4. MD simulations

The optical tweezers experiments show the destabilizing effect of trehalose on dsDNA molecule but the details of this interaction remain unclear. To find out if trehalose interacts directly with the DNA, we performed molecular dynamics simulations on ssDNA and dsDNA with 30-mer AT, 30-mer GC and 33-mer with 50%GC in the presence of 0.5 M trehalose (Figure 7.3). We determined the number of hydrogen bonds between trehalose and DNA during the simulation for the different DNA molecules, and found that trehalose interacted in majority with the backbone in the case of dsDNA whereas for ssDNA trehalose interacted in majority with its bases (Extended Data Figure 7.2), leading to significant elongation of the ssDNA (Figure 7.3b). As compared to ssDNA in the absence of trehalose we observed $\sim 30\%$ and $\sim 50\%$ increase in the average end-to-end distances for AT and GC ssDNA molecules in the presence of 0.5 M trehalose. In contrast, for dsDNA a reduction in the end-to-end distances was observed. Notably, during the simulation trehalose formed more hydrogen bonds with GC DNA, than with AT DNA (Extended Data Figure 7.3), that is consistent with more pronounced effect of trehalose on GC DNA [10].

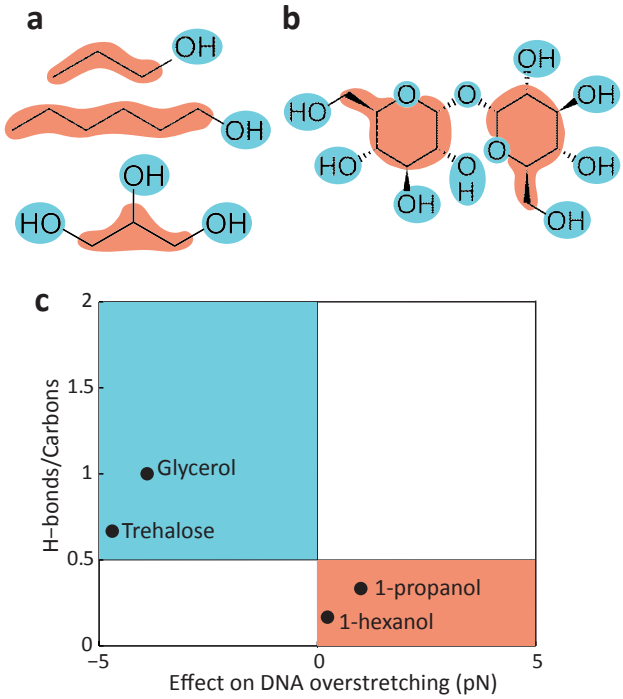


Figure 7.2: Effect on dsDNA overstretching force for various amphiphiles. **(a)** From top to bottom: 1-propanol, 1-hexanol, glycerol. **(b)** Trehalose. Hydrophilic parts of the molecules are highlighted in blue and hydrophobic parts in orange. **(c)** Diagram showing the relationship between the effect of each molecule on dsDNA overstretching force (x axis) versus the ratio of the amount of possible hydrogen bonds to the number of carbons present in the molecule (H-bonds/Carbons, y axis).

We also analyzed the effect of trehalose on the surrounding aqueous environment by calculating the radial distribution function (RDF) for water molecules and DNA in the presence and in the absence of trehalose in the simulation box (Extended Data Figure 7.4). Addition of trehalose lowers RDF without affecting its shape. It suggests that trehalose replaces some of the waters in the vicinity of DNA. It is in line with direct interaction between trehalose and nucleobases via hydrogen bonds.

7.4. Discussion

7.4.1. Physical mechanism of trehalose induced dsDNA destabilization

At the overstretching force, dsDNA undergo structural changes that depend on the experimental conditions. For the torsionally unconstrained DNA, two types of transitions can occur: unpeeling of strands from the ends of the molecule,

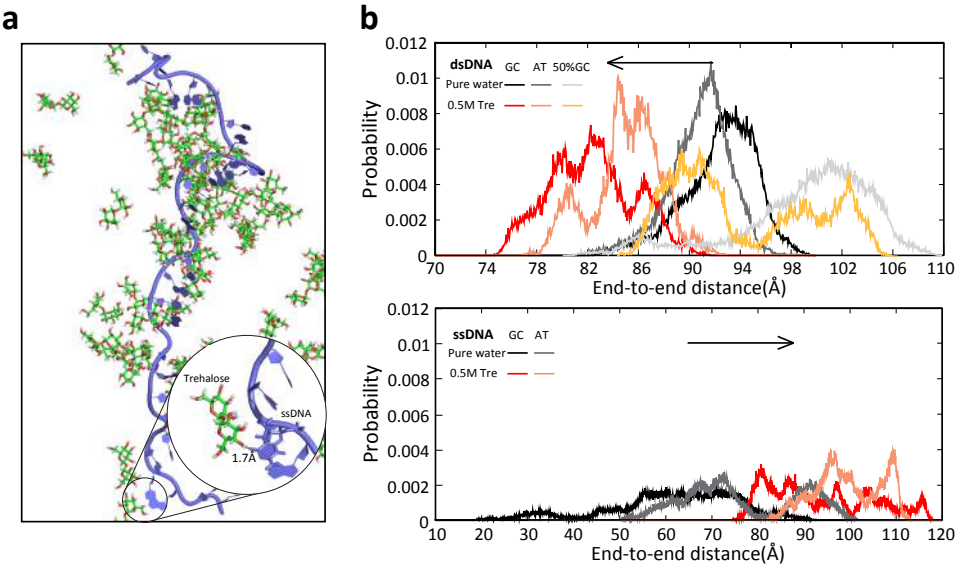


Figure 7.3: Results from molecular dynamics simulation of trehalose-DNA system. (a) Direct interaction of trehalose with nucleobases in ssDNA. (b) The effect of trehalose on end-to-end distance of ss and ds DNA. This distance is defined as the distance between midpoints of C3 in the furan rings of the first and the last base pair. The distribution of end-to-end distances is plotted for dsDNA (top) and for ssDNA (bottom) without trehalose (black trace for GC, dark grey for AT and light grey for 50%GC DNA) and in the presence of 0.5M trehalose (red trace for GC, orange for AT and yellow for 50%GC DNA). While trehalose reduces the average end-to-end distance of dsDNA, it enhances the end-to-end distance of ssDNA. In both cases the effect on GC DNA is more pronounced than AT DNA.

that involves disruption of base pairs and yields ssDNA; and the so-called "B-to-S" transition[3], where the DNA helix unwinds with the base pairs remaining intact [36, 37], yielding another dsDNA structure called S-DNA.

As an indicator of the occurrence of certain DNA structures at the overstretching we considered the presence or absence of hysteresis in the relaxation curve. The B-to-S transition is non-hysteretic, while the unpeeling transition exhibits significant hysteresis in the relaxation curve [38]. The presence of hysteresis in our data (Extended Data Figure 7.1) suggests that base pairing is indeed disrupted during the overstretching. With our DNA sequence and within our experimental conditions, we never observed a stretching curve without the hysteresis in relaxation phase. This indicates the transition is not exclusively B-to-S. More likely, we observe a mixture of unpeeled DNA and S-DNA that coexist in one DNA molecule. This hypothesis is also supported by our fits of DNA stretching curves after overstretching transition (see equation 6), which yield bending modulus κ = 1.28 that is equal to and monomer size a = 0.665 nm that is slightly less than the values reported by Manghi et al.[24] (κ = 1.28, a = 0.7 nm) for ssDNA. The competition

[3]B-DNA is the name for the most common dsDNA structure under the conditions found in cells.

between different overstretching mechanisms has been demonstrated recently experimentally with an arbitrary DNA sequence in conditions similar to ours by King et al. [36]

The DNA overstretching occurs at applied force that varies negligibly during the transition and at constant temperature, and it has already been characterized as force-induced melting process [39]. From a thermodynamics point of view, during the melting the Gibbs free energy does not change while the enthalpy and the entropy of the material do increase. Melting occurs when the Gibbs free energy of the ssDNA becomes lower than that of the dsDNA. The temperature at which this occurs depends on the external force and the buffer composition. At zero force the melting temperature ($T_m{}^0$) is well above the ambient temperature and dsDNA is stable. At a force of 65 pN (overstretching force), the melting temperature ($T_m{}^{Fos}$) equals the ambient temperature and the transition occurs. Trehalose has been reported as a potent PCR enhancer, which indicates that it lowers $T_m{}^0$ [9]. Under similar buffer conditions as ours in bulk assays, the zero-force melting temperature of dsDNA molecules with similar GC content was found to decrease by 4.4 °C per molar concentration of trehalose [10]. The extent of this change was found to depend strongly on the GC content but only minimally on the DNA length [10]. Over the range of trehalose concentrations assayed here (from 0 to 100 mM), these results would imply a decrease of the order of 1 °C in the melting temperature ($T_m{}^0$) and in $T_m{}^{Fos}$ in our experiments.

A previous experimental analysis of the dependence of the overstretching force on temperature in optical tweezers assays [39] can also be compared with our data. Given the decreases in overstretching force observed here, this previous work suggests a much larger decrease of the melting temperature of about 15 °C. This indicates that the 1 °C change in the melting temperature afforded by enthalpic and entropic changes associated with breaking the base pairing alone cannot account for the significant change in overstretching forces measured in Figure 7.1d. Another aspect of the stability of molecules is their interaction with the solvent. DNA melting results in the exposure of the nucleobases which due to the interactions with their hydrophobic surfaces have an effect on the interfacial water[40, 41]. This transition affects the degrees of freedom of the water, and hence the heat capacity of the entire system, which is the energy needed to increase the temperature (kinetic energy per degree of freedom) of the system by one degree. We note that the hydrophobicity of nucleobases is a complicated concept. Bases can be considered hydrophilic because they can form hydrogen bonds and are water soluble. However, they also possess considerable hydrophobic surfaces that provide significant contribution to DNA duplex formation [42].

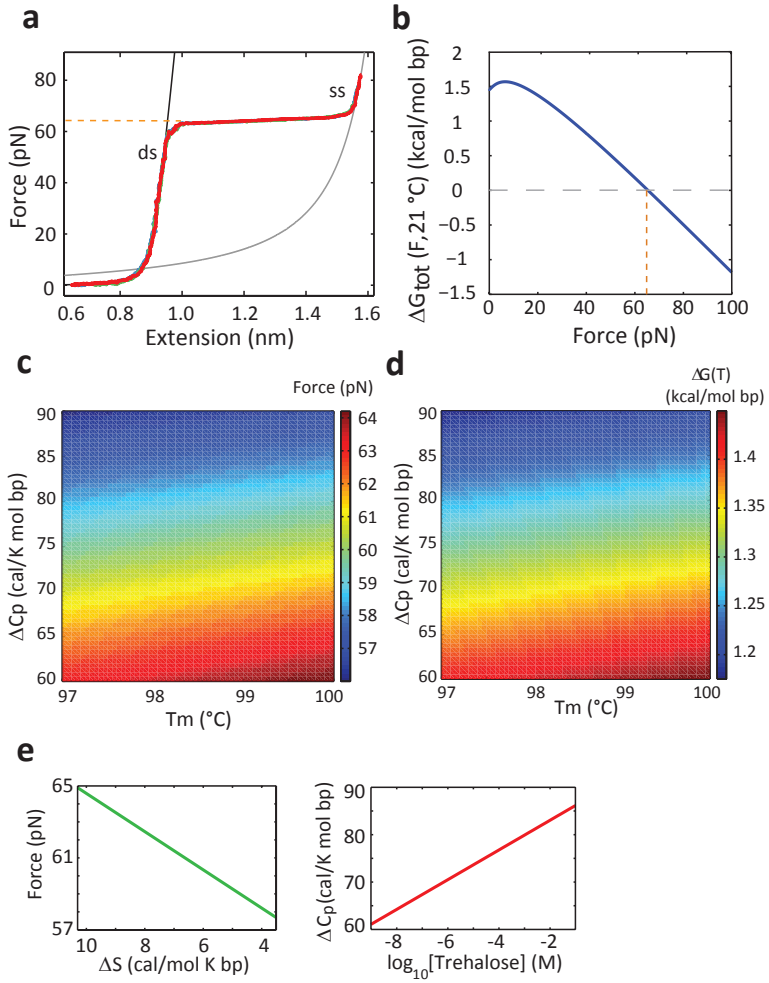


Figure 7.4: Physical mechanism of trehalose induced destabilization of double stranded DNA. **(a)** A typical force-extension plot is shown. Black line is the WLC model for dsDNA, grey line - for ssDNA. The dashed line indicates the force at which the DNA melting from dsDNA to ssDNA occurs at room temperature (21°C) in the absence of trehalose. **(b)** The total free energy difference (ΔG_{tot}(F,T)), as a function of force at room temperature, T_m = 100°C and at ΔC_p = 60 cal/mol K bp. This value of ΔC_p corresponds to the DNA in the absence of trehalose as calculated in [?]. **(c)** Color map of the force required for melting as a function of T_m and ΔC_p (equations 4 and 5). **(d)** Color map of the temperature-dependent transition free energy as a function of T_m and ΔC_p. **(e)** Left panel: overstretching force of DNA measured for different concentrations of trehalose is plotted versus calculated total transition entropy from dsDNA to ssDNA in solution (ΔS). Right panel: DNA heat capacity change upon melting (ΔC_p) versus the concentration of trehalose. For the thermodynamic model to numerically fit the measured overstretching forces (Fig. 1D), it would have to produce a linear increase in ΔC_p with the log of the trehalose concentration, as indicated here. ΔC_p goes down linearly by increasing ΔS i.e. ΔS is proportional to the log of [Trehalose].

The contribution of the surrounding aqueous environment to heat capacity effects for DNA melting has already been underscored in the literature: Rouzina et al. [43] argue that the entropic change associated with changes in heat capacity upon melting ΔC_p of DNA can be a factor 2 – 4 times larger than the entropic change associated with only breaking the base pairing $\Delta S_{bp} \approx 25$ cal/mol K. Therefore the surrounding environment significantly contributes to the total entropy of the system (ΔS), an effect that changes with varying solution conditions. They also reason that an increase in hydrophobicity can greatly increase this heat capacity. We note that this effect could not be observed in PCR experiments or other experiments in which the system is observed in thermostatic conditions and not during heating and cooling.

7.4.2. Trehalose provides a hydrophobic environment for DNA

DNA basepairing in aqueous buffer is enthalpically driven by establishing the hydrogen bonds between the bases. Entropy also has a known effect on the stability of double stranded DNA. For example the chain conformational entropy reduces upon transition from ssDNA to dsDNA [44]. The reduction of entropy has been attributed to a reduced configuration space for hydrogen or alternatively to an increase of many-body correlations [45].The arrangement of water molecules at the interface with DNA, e.g. in the grooves [40, 46], will also have an entropic contribution. The water arrangement entropy could be influenced by the change in the composition of aqueous buffer, for example by addition of amphiphilic molecules [41] and solutes [47].

Trehalose possesses both hydrophobic and hydrophilic moieties, being a moderate amphiphile [48]. Hydrophobicity on the molecular scale is widely believed to have an entropic origin [47, 49, 50], particularly at room temperature. To test if an increase in hydrophobicity due to trehalose could account for the changes in the overstretching force, we adapted the model from Williams et al. [39]. It assumes that DNA melting occurs during the overstretching transition and that the total transition free energy can be written as the sum of a temperature-dependent term in the absence of force and a force-dependent term. Using this model, we can estimate what changes in heat capacity – and the consequent change in the system's entropy – would explain the measured DNA overstretching forces. Although the DNA used and the buffer conditions in Williams et al.[39] are different from ours, a comparative assessment can still be made about the effects of increasing trehalose concentration on the overstretching force.

We used the discrete worm-like chain (WLC) to fit data of the dsDNA and ssDNA. (Figure 7.4a). By directly integrating the WLC curves, one can find the force-dependent free energy change $\Delta\Phi(F)$ associated with the melting transition of the DNA as shown in equation 7. Thus one can numerically solve the overstretching force associated with each value of ΔC_p for different trehalose concentrations using equation 4.

To illustrate, ΔG_{tot}(F,T) is plotted in Figure 7.4b at room temperature T = 21°C, ΔC_p = 60 cal/mol K bp and $\Delta S(T_m^0)$ = 24.7 cal/mol K bp. The orange dashed

line marks the transition of the DNA molecule favouring the double-stranded state at lower forces and the single-stranded state at higher forces. Changes in T_m over a reasonable range (of order 1°C, see [39]) show little effect on the temperature-dependent part of the transition free energy ΔG(T) (Figure 7.4d) and hence the force required for the melting transition (Figure 7.4c). The left side of the equation 4 is a linear function of ΔC_p. The dependence of the overstretching force and consequently $\Delta\Phi(F)$, on the trehalose concentration, implies that ΔC_p has to be dependent on the concentration of trehalose as well. Changes in C_p over a reasonable range (see [43]) do result in significant adjustments of ΔG(T) (Figure 7.4d). Moreover, the model provides a good quantitative fit to the measured overstretching force (Figure 7.1d), for a linear dependence of ΔC_p on the log[Trehalose] (Figure 7.4e).

A change in ΔC_p as a function of trehalose concentration indicates a change in the number of degrees of freedom of the system and consequently the system entropy, that depends on heat capacity change as follows [**?**]:

$$\Delta S = \Delta S(T_m) + \Delta C_p \times \ln \frac{T}{T_m} \tag{7.8}$$

Here, an increasing ΔC_p corresponds to decreasing transition entropy (see equation 8; note that for $T < T_m$, $\ln \frac{T}{T_m}$ is negative). These findings indicate that a dependence of C_p on trehalose concentration could explain the observed effect of trehalose on the overstretching force (Figure 7.1d).

These thermodynamic considerations provide physical insight into the trehalose-mediated DNA melting process. However, multiple underlying molecular (chemical) mechanisms can be envisioned. Water-trehalose interactions (competition for water with DNA) leading to trehalose-mediated sequestering of water molecules could promote a dehydration of DNA molecule, that is assumed to destabilize DNA duplex[11], while providing a relatively hydrophobic environment for the DNA bases. Alternatively, due to amphiphilic properties of trehalose [48], its direct interaction with the DNA bases (either via H-bonds or van der Waals interactions) leading to replacement of water by trehalose in the vicinity of the DNA could provide more hydrophobic interface for bases, and at the same time provide more hydrophilic interface for water. With the results discussed so far we cannot sketch an atomistic picture of the process, which we aimed to address by investigating some of the molecular motifs found in polyol osmolytes and by using MD simulations.

7.4.3. The ratio between polar and apolar fractions of an osmolyte determines the effect on the DNA

To explain the more prominent dsDNA destabilization caused by glycerol and trehalose and the lack of effect with 1-propanol and 1-hexanol we consider the solubility of the additives. Water is structured around the osmolyte and it results in the heat capacity change ΔC_p, which is positive for apolar solutes and negative for polar solutes [51]. The amount of polar and apolar groups in the additive can be seen as an indicator of a (de-)stabilization effect on DNA double helix. As a rough

estimate of polar and apolar fraction we have considered the ratio of the amount of possible hydrogen bonds to the number of carbons present in each molecule (Figure 7.2a-b).

We find that this ratio correlates well with the overstretching force (Figure 7.2c). In contrast, the total number of polar or apolar groups do not correlate well with the overstretching force. For example, glycerol (which reduces DNA overstretching force) and 1-propanol (which has no effect on it) contain the same number of apolar groups. Moreover, trehalose and glycerol contain a very different number of polar groups but reduce the overstretching force to similar extent. These data suggest that neither the hydrophilicity nor the hydrophobicity of the molecule determines the effect, but the ratio between the parts is important. The data suggest that molecules with the particular ratio, such as trehalose (and glycerol), are more effective at providing a bridging interface between the hydrophobic surface area of the DNA bases and the surrounding water, lowering the transition energy from dsDNA to ssDNA and decreasing the overstretching force.

7.4.4. Trehalose binds the nucleobases

The results from MD simulations show that trehalose interacts directly with the bases of ssDNA in aqueous solution. It leads to elongation of ssDNA, that indicates the stiffening of the polymer upon trehalose addition. In the case of dsDNA, trehalose causes the increase in flexibilty of the polymer, that is in agreement with experimentally observed dsDNA destabilization. From these data, a simple mechanism is proposed: thermal fluctuations (particularly when the transition barrier between ssDNA and dsDNA is suppressed in the presence of mechanical force) disrupt the base-pairing in the dsDNA leading to exposure of the bases (base flipping). Trehalose then interacts with the bases and increases the likelihood of the flipped-out state. This can explain the experimental observations of the reductions in overstretching forces of dsDNA. The more prominent effect in the case of CG is in agreement with the larger hydrophobic surface areas of guanine and cytosine, particularly the stacking areas [52, 53].

7.5. Conclusion

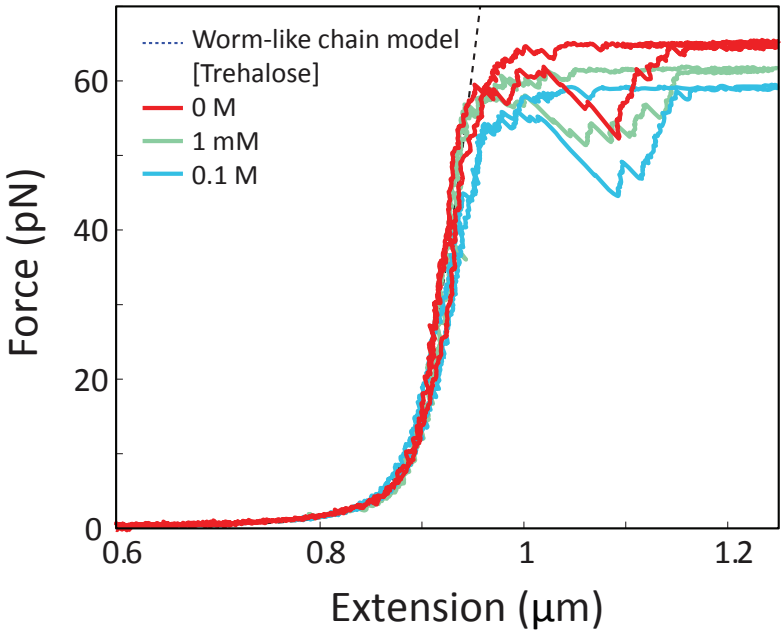
Our single molecule optical tweezers study, together with the thermodynamic analysis and molecular dynamics simulations provide a general mechanism that explains the influence of trehalose on biomolecules. The data indicate that – due to the presence of hydrophobic moieties – trehalose interacts more with hydrophobic parts of biomolecules than water and favours the exposure of hydrophobic fragments. Furthermore its amphiphilic character allows trehalose to bridge between hydrophobic moieties and surrounding water resulting in an entropically favored interface.

7.6. Acknowledgements

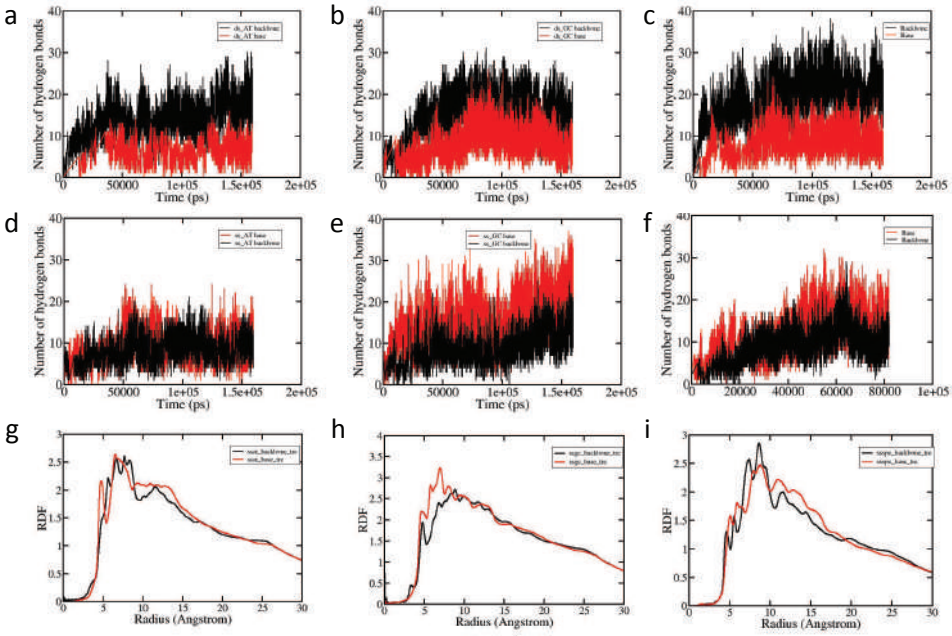
We are thankful to Dr. Yves Rezus and Magdalena P. López for critical reading of the manuscript. This work was supported by the research programme of

the Foundation for Fundamental Research on Matter (FOM), which is part of the Netherlands Organisation for Scientific Research (NWO) (to S.J.T.). This work was also supported by National Key Basic Research Special Funds [grant number 2012CB917304 to Y.Q.G.] and National Natural Science Foundation of China [grant number 21125311 to Y.Q.G.].

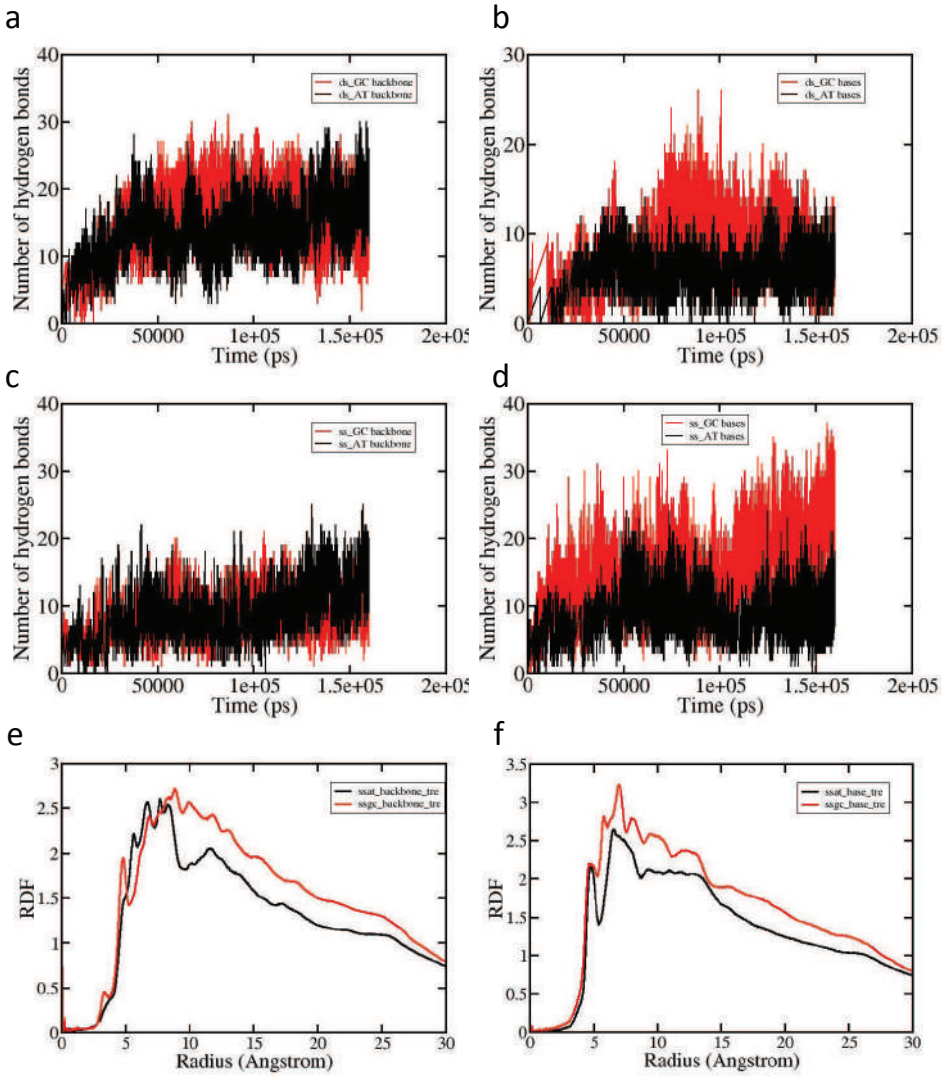
7.7. Extended data



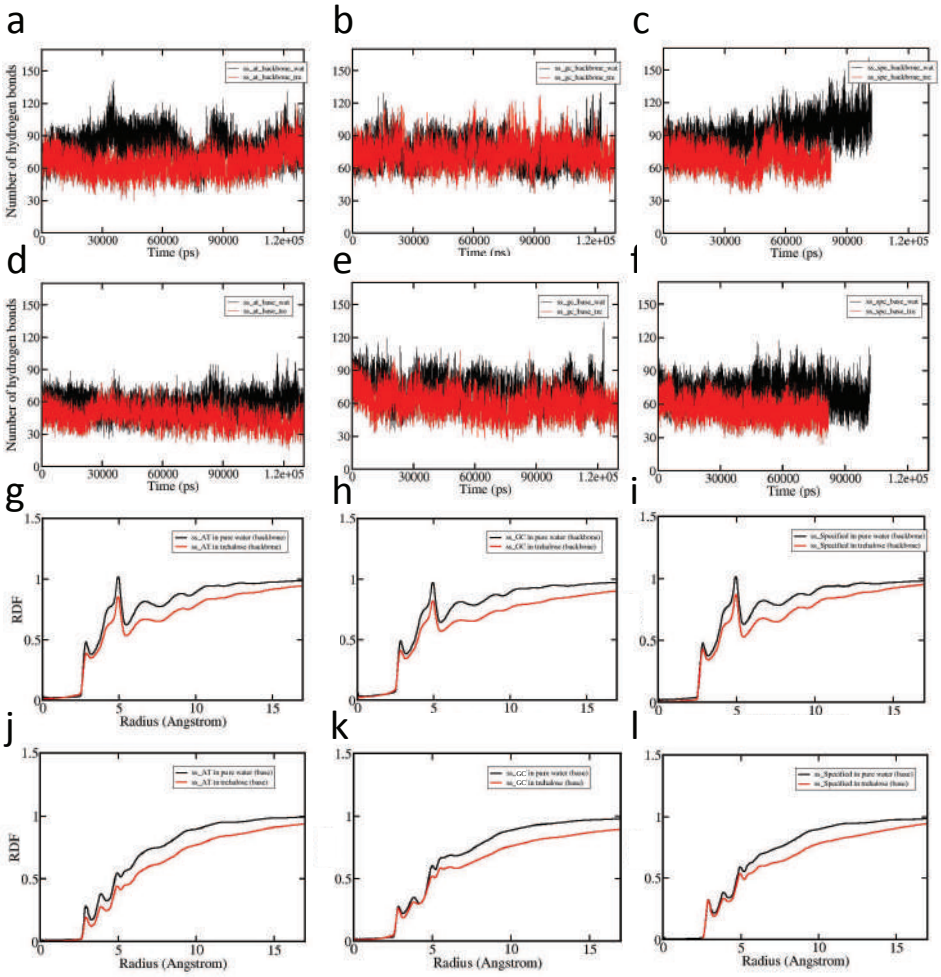
Extended Data Figure 7.1: Stretching and relaxation curves of dsDNA. The mechanical response of dsDNA was measured in the absence and in the presence of trehalose. In both conditions DNA exhibits significant hysteresis upon relaxation.



Extended Data Figure 7.2: Base-to-backbone preference of trehalose compared for dsDNA and ssDNA. **(a-f)** Number of hydrogen bonds between DNA and trehalose determined from MD simulations is plotted versus simulation time. The number of hydrogen bonds between DNA backbone and trehalose is shown in black, between DNA nucleobases and trehalose – in red. The results are shown for the following DNA molecules: **(a)** dsAT, **(b)** dsGC, **(c)** ds50%GC, **(d)** ssAT, **(e)** ssGC, **(f)** ss50%GC. **(g-i)** Radial distribution function (RDF) is calculated for ssDNA and trehalose from MD simulations results. Here and in the following figures RDFs calculated are averages of RDFs for every residue. DNA backbone is defined as O3' atom of phosphodiester, DNA bases are defined as C4 atom of guanine, N3 atom of cytosine, C4 atom of adenine and N3 atom of thymine. Water is defined as its O atom. RDF for ssDNA backbone and trehalose is shown in black, for ssDNA nucleobases and trehalose - in red. The results are displayed for the following DNA sequences: **(g)** AT, **(h)** GC, **(i)** 50%GC.



Extended Data Figure 7.3: Abundance of trehalose-DNA interaction is compared for AT-DNA and GC-DNA. **(a-d)** Number of hydrogen bonds between DNA and trehalose determined from MD simulations is plotted versus simulation time. The number of hydrogen bonds between GC-DNA and trehalose is shown in black, between AT-DNA and trehalose – in red. **(a)** and **(b)** display the number of hydrogen bonds between trehalose and dsDNA: **(a)** between trehalose and backbone, **(b)** between trehalose and nucleobases. **(c)** and **(d)** display the number of hydrogen bonds between trehalose and ssDNA: **(c)** between trehalose and backbone, **(d)** between trehalose and nucleobases. **(e-f)** Radial distribution function (RDF) for AT-DNA and trehalose (shown in black) is compared with RDF for GC-DNA and trehalose (in red): **(e)** for DNA backbone, and **(f)** for DNA bases.



Extended Data Figure 7.4: Effect of trehalose on DNA-water interface. **(a-f)** Number of hydrogen bonds between ssDNA and water determined from MD simulations is plotted versus simulation time. The number of hydrogen bonds between DNA and water in the absence of trehalose is shown in black, between DNA and water when trehalose is present – in red. **(a)** shows the number of hydrogen bonds between the backbone of ssAT and water, **(b)** - backbone of ssGC, **(c)** - backbone of ss50%GC, **(d)** - bases of ssAT, **(b)** - bases of ssGC, **(c)** - bases of ss50%GC. **(g-l)** Radial distribution function (RDF) for ssDNA and water molecules is plotted when trehalose was absent (black) and when it was present (red) for the following molecules and parts of DNA: **(g)** - backbone of ssAT, **(h)** - backbone of ssGC, **(i)** - backbone of ss50%GC, **(j)** - bases of ssAT, **(k)** - bases of ssGC, **(l)** - bases of ss50%GC.

References

[1] N. K. Jain and I. Roy, *Effect of trehalose on protein structure*. Protein Sci. **18**, 24 (2009).

[2] A. D. Elbein, Y. T. Pan, I. Pastuszak, and D. Carroll, *New insights on trehalose: a multifunctional molecule*. Glycobiology **13**, 17R (2003).

[3] B. Zhu, T. Furuki, T. Okuda, and M. Sakurai, *Natural DNA mixed with trehalose persists in B-form double-stranding even in the dry state*. J. Phys. Chem. B **111**, 5542 (2007).

[4] J. K. Kaushik and R. Bhat, *Why is trehalose an exceptional protein stabilizer? An analysis of the thermal stability of proteins in the presence of the compatible osmolyte trehalose*. J. Biol. Chem. **278**, 26458 (2003).

[5] E. P. Melo, L. Chen, J. M. S. Cabral, P. Fojan, S. B. Petersen, and D. E. Otzen, *Trehalose favors a cutinase compact intermediate off-folding pathway*. Biochemistry **42**, 7611 (2003).

[6] M. a. Singer and S. Lindquist, *Multiple effects of trehalose on protein folding in vitro and in vivo*. Mol. Cell **1**, 639 (1998).

[7] M. Tanaka, Y. Machida, S. Niu, T. Ikeda, N. R. Jana, H. Doi, M. Kurosawa, M. Nekooki, and N. Nukina, *Trehalose alleviates polyglutamine-mediated pathology in a mouse model of Huntington disease*. Nat. Med. **10**, 148 (2004).

[8] S. Habib, M. a. Khan, and H. Younus, *Thermal destabilization of stem bromelain by trehalose*. Protein J. **26**, 117 (2007).

[9] A.-N. Spiess, N. Mueller, and R. Ivell, *Trehalose is a potent PCR enhancer: lowering of DNA melting temperature and thermal stabilization of taq polymerase by the disaccharide trehalose*. Clin. Chem. **50**, 1256 (2004).

[10] J. L. Hart, Z. M. Harris, and S. M. Testa, *Analyzing and predicting the thermodynamic effects of the metabolite trehalose on nucleic acids*. Biopolymers **93**, 1085 (2010).

[11] P. Del Vecchio, D. Esposito, L. Ricchi, and G. Barone, *The effects of polyols on the thermal stability of calf thymus DNA*. Int. J. Biol. Macromol. **24**, 361 (1999).

[12] A. Mashaghi and A. Katan, *A physicist's view of DNA*, De Physicus **24**, 59 (2013), arXiv:1311.2545 .

[13] S. Ohtake and Y. J. Wang, *Trehalose: Current use and future applications, J. Pharm. Sci.*, **100**, 2020 (2011).

[14] Y.-P. Kim, M.-Y. Hong, J. Kim, E. Oh, H. K. Shon, D. W. Moon, H.-S. Kim, and T. G. Lee, *Quantitative analysis of surface-immobilized protein by TOF-SIMS: effects of protein orientation and trehalose additive*. Anal. Chem. **79**, 1377 (2007).

[15] Y. Mizuno, P. Carninci, Y. Okazaki, M. Tateno, J. Kawai, H. Amanuma, M. Muramatsu, and Y. Hayashizaki, *Increased specificity of reverse transcription priming by trehalose and oligo-blockers allows high-efficiency window separation of mRNA display*, Nucleic Acids Res. **27**, 1345 (1999).

[16] M. D. Wang, H. Yin, R. Landick, J. Gelles, and S. M. Block, *Stretching DNA with optical tweezers*. Biophys. J. **72**, 1335 (1997).

[17] C. Bustamante, S. B. Smith, J. Liphardt, and D. Smith, *Single-molecule studies of DNA mechanics*. Curr. Opin. Struct. Biol. **10**, 279 (2000).

[18] I. Tessmer, C. G. Baumann, G. M. Skinner, J. E. Molloy, J. G. Hoggett, S. J. B. Tendler, and S. Allen, *Mode of drug binding to DNA determined by optical tweezers force spectroscopy*, J. Mod. Opt. **50**, 1627 (2003).

[19] P. Gross, N. Laurens, L. B. Oddershede, U. Bockelmann, E. J. G. Peterman, and G. J. L. Wuite, *Quantifying how DNA stretches, melts and changes twist under tension*, Nat. Phys. **7**, 731 (2011).

[20] A. Candelli, G. J. Wuite, and E. J. Peterman, *Combining optical trapping, fluorescence microscopy and micro-fluidics for single molecule studies of DNA-protein interactions*, Phys Chem Chem Phys **13**, 7263 (2011).

[21] M. C. Williams, J. R. Wenner, I. Rouzina, and V. a. Bloomfield, *Entropy and heat capacity of DNA melting from temperature dependence of single molecule stretching*. Biophys. J. **80**, 1932 (2001).

[22] P. Bechtluft, R. G. H. van Leeuwen, M. Tyreman, D. Tomkiewicz, N. Nouwen, H. L. Tepper, A. J. M. Driessen, and S. J. Tans, *Direct observation of chaperone-induced changes in a protein folding pathway*. Science **318**, 1458 (2007).

[23] A. Mashaghi, P. J. Vach, and S. J. Tans, *Noise reduction by signal combination in Fourier space applied to drift correction in optical tweezers*. Rev. Sci. Instrum. **82**, 115103 (2011).

[24] M. Manghi, N. Destainville, and J. Palmeri, *Mesoscopic models for DNA stretching under force: New results and comparison with experiments*. Eur. Phys. J. E. Soft Matter **35**, 110 (2012).

[25] D. A. Case, T. A. Darden, T. E. Cheatham III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. C. Walker, W. Zhang, K. M. Merz, *et al.*, *AMBER 11*. (University of California, San Francisco, 2010).

[26] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma, *The missing term in effective pair potentials*, J. Phys. Chem. **91**, 6269 (1987).

[27] A. Pérez, I. Marchán, D. Svozil, J. Sponer, T. E. Cheatham, C. A. Laughton, and M. Orozco, *Refinement of the AMBER force field for nucleic acids: improving the description of alpha/gamma conformers*. Biophys. J. **92**, 3817 (2007).

[28] K. N. Kirschner, A. B. Yongye, S. M. Tschampel, J. González-Outeiriño, C. R. Daniels, B. L. Foley, and R. J. Woods, *GLYCAM06: a generalizable biomolecular force field. Carbohydrates.* J. Comput. Chem. **29**, 622 (2008).

[29] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, *Molecular dynamics with coupling to an external bath,* J. Chem. Phys. **81**, 3684 (1984).

[30] J.-P. Ryckaert, G. Ciccotti, and H. J. Berendsen, *Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes,* J. Comput. Phys. **23**, 327 (1977).

[31] T. Darden, D. York, and L. Pedersen, *Particle mesh Ewald: An N□log(N) method for Ewald sums in large systems,* J. Chem. Phys. **98**, 10089 (1993).

[32] *The PyMOL Molecular Graphics System, Version 1.3.*, Schrödinger, LLC (2010).

[33] S. B. Smith, Y. Cui, and C. Bustamante, *Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules.* Science **271**, 795 (1996).

[34] P. Cluzel, A. Lebrun, and C. Heller, *DNA: an extensible molecule,* Science **271**, 792 (1996).

[35] J. F. Marko and E. D. Siggia, *Stretching DNA,* Macromolecules **28**, 8759 (1995).

[36] G. a. King, P. Gross, U. Bockelmann, M. Modesti, G. J. L. Wuite, and E. J. G. Peterman, *Revealing the competition between peeled ssDNA, melting bubbles, and S-DNA during DNA overstretching using fluorescence microscopy.* Proc. Natl. Acad. Sci. U. S. A. **110**, 3859 (2013).

[37] X. Zhang, H. Chen, S. Le, I. Rouzina, P. S. Doyle, and J. Yan, *Revealing the competition between peeled ssDNA, melting bubbles, and S-DNA during DNA overstretching by single-molecule calorimetry.* Proc. Natl. Acad. Sci. U. S. A. **110**, 3865 (2013).

[38] D. H. Paik and T. T. Perkins, *Overstretching DNA at 65 pN does not require peeling from free ends or nicks.* J. Am. Chem. Soc. **133**, 3219 (2011).

[39] M. C. Williams, J. R. Wenner, I. Rouzina, and V. A. Bloomfield, *Effect of pH on the overstretching transition of double-stranded DNA: evidence of force-induced DNA melting,* Biophys J **80**, 874 (2001).

[40] S. K. Pal, L. Zhao, and A. H. Zewail, *Water at DNA surfaces: ultrafast dynamics in minor groove recognition.* Proc. Natl. Acad. Sci. U. S. A. **100**, 8113 (2003).

[41] R. S. Dias, L. M. Magno, A. J. M. Valente, D. Das, P. K. Das, S. Maiti, M. G. Miguel, and B. Lindman, *Interaction between DNA and cationic surfactants: effect of DNA conformation and surfactant headgroup*. J. Phys. Chem. B **112**, 14446 (2008).

[42] P. Yakovchuk, E. Protozanova, and M. D. Frank-Kamenetskii, *Base-stacking and base-pairing contributions into thermal stability of the DNA double helix*. Nucleic Acids Res. **34**, 564 (2006).

[43] I. Rouzina and V. A. Bloomfield, *Heat capacity effects on the melting of DNA. 1. General aspects*. Biophys. J. **77**, 3242 (1999).

[44] D. T. Haynie, *Biological Thermodynamics*, 1st ed. (Cambridge University Press, 2001) pp. 160–163.

[45] A. Godec and F. Merzel, *Physical origin underlying the entropy loss upon hydrophobic hydration*. J. Am. Chem. Soc. **134**, 17574 (2012).

[46] D. Biswal, B. Jana, S. Pal, and B. Bagchi, *Dynamical transition of water in the grooves of DNA duplex at low temperature*. J. Phys. Chem. B **113**, 4394 (2009).

[47] S. Garde and A. J. Patel, *Unraveling the hydrophobic effect, one molecule at a time*. Proc. Natl. Acad. Sci. U. S. A. **108**, 16491 (2011).

[48] Y. Koga, K. Nishikawa, and P. Westh, *Relative hydrophobicity/hydrophilicity of fructose, glucose, sucrose, and trehalose as probed by 1-propanol: a differential approach in solution thermodynamics*. J. Phys. Chem. B **111**, 13943 (2007).

[49] D. Chandler, *Interfaces and the driving force of hydrophobic assembly*. Nature **437**, 640 (2005).

[50] F. H. Stillinger, *Water revisited*. Science **209**, 451 (1980).

[51] K. R. Gallagher and K. A. Sharp, *A new angle on heat capacity changes in hydrophobic solvation*. J. Am. Chem. Soc. **125**, 9853 (2003).

[52] S. J. Sowerby, C. a. Cohn, W. M. Heckl, and N. G. Holm, *Differential adsorption of nucleic acid bases: Relevance to the origin of life*. Proc. Natl. Acad. Sci. U. S. A. **98**, 820 (2001).

[53] K. M. Guckian, B. A. Schweitzer, R. X.-F. Ren, C. J. Sheils, D. C. Tahmassebi, and E. T. Kool, *Factors Contributing to Aromatic Stacking in Water: Evaluation in the Context of DNA*. J. Am. Chem. Soc. **122**, 2213 (2000).

8

Conclusion

8.1. Challenges in chaperone-assisted protein folding studies

Protein folding science is a relatively young interdisciplinary field, and it is striking to look back and trace its booming growth. From the first protein structures snapshots solved by X-ray crystallography and modeled on the laboratory table, it quickly developed into a realm of ultrafast dynamic studies of folding process, all-atom molecular simulations, and investigations of protein folding in living cells. In chapter 1 of this thesis, I give a quick overview of the history and briefly mention the landmarking advances that would not be achievable without contributions of scientists from different backgrounds. A theory of native conformational search and thermodynamic principles underlying protein folding process were driven forward by mathematicians and statistical physicists, a multitude of successful experimental techniques that pushed forward temporal and spatial resolution of protein folding observations was developed by physicists, computer science allowed for in-depth modeling and visualization of protein folding. Complicated cellular machinery that synthesizes proteins and promotes their correct folding has been discovered and characterized by cellular biologists. However, our understanding of protein folding is far from complete. Conventional experimental techniques often suffer from several limitations that do not allow to explore the folding pathway of a protein in full details. Due to the ensemble averaging, transient low-populated intermediate states of the folding pathways cannot be detected. Functional in vitro assays and in vivo cell viability assays do not discriminate non-folded, misfolded and aggregated proteins. On the other hand, single-molecule techniques are very well suited to observe and probe rare non-native protein states. This property is crucial for answering long-standing questions in the protein folding field, especially in the chaperone-assisted protein folding. Conformational transitions and interconversions of chaperones themselves are relatively well studied by the conventional assays. However, it is always a

challenge for them to focus on a protein client instead of a chaperone and track in details how exactly the protein is shaped by the chaperone, where the chaperone binds to the client and whether the chaperone guides the client towards the native structure along the folding pathway or just prevents its aggregation. Central goal of this thesis is to improve the understanding of physical principles underlying protein folding process and the role of chaperones in it by using the single molecule force spectroscopy by optical tweezers.

8.2. Optical tweezers in single-molecule biophysics

Optical tweezers have already been applied successfully to some of these questions by our research group. The chaperone action has been observed at the single molecule scale for the first time in 2007 , and has been demonstrated along the folding pathway in 2013 . This powerful technique can definitely contribute much more not only to chaperone studies, but also to many other areas of biology, physics and material science. However, the complexity of building, alignment and operation of the setup remains a limiting factor for further spreading of the technique in more and more labs. In chapter 2, I describe an optical tweezers setup I built and provide a detailed step-by-step manual for those planning to build their own optical tweezers. I hope that it will lower the entry barrier into the optical tweezers field and will allow non-experts in optics to build the setup in less time and cost and with less frustration and effort. In that chapter, I also outline certain improvements for the optical tweezers I built. The combination of single molecule force spectroscopy with confocal fluorescence aims to enhance the outcome of the chaperone-assisted protein folding studies and to allow observing chaperones acting on a single protein client for the first time. Implementation of sample temperature control in optical tweezers aims to increase the versatility of this powerful technique to make it applicable for even wider range of question in physics and biology, where temperature is an important experimental parameter.

The sample temperature control has been built and tested, and it has been described in chapter 3. The method, which is based on heating of air around the sample, is easy to implement, low-cost and can be easily adjusted to different systems. It allows to heat the sample up non-invasively and does not deteriorate the optical tweezers performance significantly by introducing drifts. I demonstrated the capabilities of the method by performing overstretching of the same DNA molecule at different temperatures that has not been achieved before using similar methods.

8.3. Heat shock proteins studied with optical tweezers

Three chapters of the thesis (4-6) are dedicated to investigations of the effect of heat shock proteins (chaperones) on protein folding on the single molecule scale. Chaperones perform a plethora of tasks in the cell, including folding assistance, protection against aggregation, protein trafficking and enzyme activity

regulation. They can interact with non-folded protein chains, misfolded and aggregated proteins, and even near-native protein structures. It is believed that chaperones achieve their physiological roles by binding non-folded peptide segments that extend from these different protein conformers.However, it has been difficult to test directly how chaperones interact with protein clients and affect their structures, given the conformational dynamics of both chaperone and its client. Optical tweezers are very well suited to address these questions, since they allow for direct observation of the protein state in the absence and in the presence of chaperones. Furthermore, they can be used to shape the energy landscape of the client protein by applying certain force and increasing the population of a desired non-native folding state.

Using optical tweezers, I studied three different chaperone systems: bacterial heat shock proteins Hsp70 and Hsp33, and human chaperone HspB6. Hsp70 is traditionally classified as a foldase from the physiological point of view, meaning that it supposed to assist protein folding by interacting with it. Two other chaperones are classified as holdases that indicates that they were found to bind non-folded protein chains and prevent not only their aggregation, but also native refolding. However, our single molecule mechanistic studies demonstrated that functional repertoire of all these chaperones is broader than previously assumed, and that they are not limited to the interaction with non-folded protein chains.

In chapter 4, we show that Hsp70 binds not only extended peptide segments, but also folded protein structures, and it can stabilize them against applied force. Hence, Hsp70 can not only destabilize but also stabilize folded proteins in a nucleotide-regulated manner. This finding suggests that it may preserve enzymatic function during stress at low energy costs. Furthermore, an ability of Hsp70 to interact with folded protein structures indicates that it may act later in native folding than previously assumed. It is believed that proteins either refold autonomously when they are released from Hsp70 or are transfered to later acting chaperones, such as GroEL or Hps90. However, our findings indicate that Hsp70 may guide a protein along the folding pathway and stay bound to it until the protein is almost natively folded. The interaction with partially folded protein structures is potentially also very relevant for aggregation prevention. It has been suggested recently that partially folded proteins also present aggregation risk, which can be lowered effectively by Hsp70 binding to them. Overall, these mechanistic findings suggest that Hsp70 may have more functions and handle them more effectively. It could have important consequences for our understanding of cell physiology and pathophysiology.

Chapters 5 and 6 are dedicated to studies of Hsp33 and HspB6, which are considered to protect proteins upon stress by binding to unfolded proteins and shielding them from aggregation. When the stress is over, Hsp33 is believed to hand over the clients to Hsp70 that assists their further refolding. We demonstrated that Hsp33 can prevent protein aggregation and allow its native refolding to some extent when present iin sub-stoichiometric concentrations even without the collaboration with Hsp70 system. Importantly, we have found that Hsp33 affects the properties of protein native core structure (both size and stability),

indicating that it is capable of binding to it. The Hsp33 disordered linker region may be responsible for the chaperone-client interaction as suggested by previous studies . HspB6 has been found by us to unexpectedly favour the compact state of an aggregation-prone protein and stabilize it against applied force, instead of preventing its aggregation. Our findings suggest that function of these chaperones is broader than holdase activity as it is traditionally defined. It could indicate that their clients are delivered to downstream chaperones (such as Hsp70) in a state closer to native when the stress is over. Together with the results presented in chapter 4, it may mean that whole chaperone network of the cell may be more versatile and flexible, interacting with more protein states than previously assumed. It may suggest the shift of complete chaperone-assisted folding paradigm from an image of chaperones being ”paramedics” that serve at pathological conditions and at the very beginning of native protein folding pathway to a consideration of chaperones as molecules that guide native folding from a start to an end, and are also very important for overall normal cell physiology. Findings presented in this thesis could serve as a mechanistic basis for explanation of regulatory function of chaperones that have been reported to affect enzyme activity.

Summary

Since the first protein structure reported, the understanding of how proteins fold so fast and efficiently into such peculiar shapes has quickly become one of the major problems of biology. A lot of knowledge has been acquired since that time, and even though the protein folding problem is not fully solved yet, it led to the rapid development of a new interdisciplinary field, namely protein folding science. Growing demands of protein scientists resulted in a booming development of new experimental techniques, which allowed to observe protein folding in the test tube.

However, protein folding in the living cell features a new level of complexity, due to highly crowded cellular environment, translation process, and various posttranslational modifications that proteins often experience. In order to overcome this problem, cells are equipped with specialized proteins that help other proteins to fold efficiently, facilitate the refolding of unfolded and misfolded proteins, prevent their association in non-functional aggregates and even dissolve the aggregates that already formed. These proteins are called molecular chaperones. Mechanisms, which are exploited by the chaperones to maintain the proteome integrity in normal conditions and in the variety of stresses, are quite poorly understood, despite their significant relevance for a range of diseases (including neurodegenerative diseases and diabetes).

Understanding the mechanistic details of these molecular machines is limited by serious experimental challenges. Interactions of chaperones with substrate proteins are often very heterogeneous and transient, which limits the applicability of biochemical ensemble techniques to this problem. In this thesis, I used single molecule force spectroscopy in optical tweezers to investigate the effect of various chaperone systems on protein folding pathways. This technique is perfectly suited to answer the following questions, which were rarely addressed in the literature previously: What happens to protein substrate as chaperone binds to it? Which protein states can chaperones bind? How can a chaperone perform such diverse functions by just binding and releasing a protein?

Chapter 1 presents a reader with a short historical overview of protein folding field and describes main experimental advances and challenges in the area. It introduces key open questions and explains how single molecule techniques can be exploited to address them. Next, it highlights the chaperone classes studied in this thesis and provides a few examples of insightful single molecule studies on them.

It is followed by chapter 2, which focuses on the technical side of this thesis. In this chapter, I describe design principles underlying the building and assembly of optical tweezers setup. Furthermore, it features advice on necessary components

and detailed alignment manual for those interested in building their own optical tweezers setup.

Chapter 3 presents another technical development, namely a low-drift heating unit for optical tweezers. Function of certain chaperones is linked to a heat stress that motivates our interest in being able to heat the sample. However, heating in optical tweezers is not quite straightforward due to drift of optical components and limited space around the sample. In this chapter, I present a novel heating unit that heats the air around the sample and the objectives uniformly in order to minimize the drift. Using this method, the mechanics of one molecule was probed at various temperatures and also during the heating for the first time.

Chapter 4 deals with the molecular mechanism of Hsp70 chaperone system. The Hsp70 system is a central hub of chaperone activity in all forms of life. It performs a plethora of diverse functions, and it is believed that the binding and release of non-folded peptide stretches in a substrate protein in ATP- and cochaperone-regulated manner underlies them all. However, observing directly how the chaperone interacts with protein substrates at the different stages of folding was proven difficult with conventional techniques, since these interactions are regulated by cochaperones and ATP hydrolysis, and are often transient. We used optical tweezers to systematically study the effect of the complete Hsp70 chaperone system and its individual components on the folding pathway of maltose binding protein (MBP). We demonstrate that bacterial Hsp70 (DnaK) binds not only extended non-folded fragments of a substrate protein, but also folded protein structures. Based on these data, we suggest an extension to the Hsp70 canonical action model that potentially has implications for all the Hsp70 functions.

Chapters 5 and 6 are dedicated to two eukaryotic chaperones, which are classified as ”holdases”. They are believed to bind non-folded protein fragments promiscuously with certain affinity that is not regulated by ATP. In both chapters, we report unexpected chaperone-substrate interaction modes. More experimental data from biochemical assays are necessary to obtain further details on the nature of these interactions. Chapter 5 is a study of an intrinsically disordered chaperone Hsp33 that is activated by oxidation. We demonstrated for the first time that Hsp33 can bind protein structures as they fold and interfere with their complete refolding, while allowing near-native conformations. We also built a simple thermodynamic model that is consistent with our experimental data. In chapter 6, I focus on the effect of chaperones on a process that underlies a range of medical challenges, namely α-synuclein (αSyn) aggregation. In humans, it leads to Lewy bodies diseases, which are a serious threat for ageing population. We studied the effect of a chaperone HspB6 on the aggregation of an engineered repeat construct of four αSyn monomers. Surprisingly, our results indicate that HspB6, instead of preventing aggregation, binds compact structures of αSyn and stabilizes them against forced unfolding. Our data are supported by the electron microscopy study, which demonstrated that HspB6 is capable of binding to mature αSyn fibrils.

In chapter 7, I study another stress response mechanism present in the organ-

isms from different domains of life. The disaccharide trehalose has been found to either stabilize or destabilize different biomolecules, depending on the substrate and the concentration of trehalose. This property is thought to be exploited by organisms in the variety of stresses, and is also used in the industry, however, the complete understanding of its mechanism is still lacking. We demonstrated that trehalose lowers the overstretching transition force of double stranded DNA and proposed an entropic driving mechanism for this effect. Molecular dynamics simulation revealed the direct interaction of trehalose with nucleobases of the DNA.

Samenvatting

Sinds de ontdekking van de eerste eiwitstructuren is het uitvinden hoe eiwitten zo snel en efficiënt in zulke eigenaardige vormen kunnen vouwen in hoog tempo één van de belangrijkste vraagstukken in de biologie geworden. Er is sindsdien veel kennis bijgekomen, en hoewel het vraagstuk over eiwitvouwing nog niet volledig is opgelost, heeft het geleid tot een snel ontstaan van een nieuw interdisciplinair veld, genaamd de wetenschap der eiwitvouwing. Groeiende interesse naar wetenschappers gespecialiseerd in eiwitten heeft geleid tot een bloeiende ontwikkeling van nieuwe experimentele technieken, die het mogelijk maken om eiwitvouwing te observeren in een reageerbuis.

Eiwitvouwing in de levende cel brengt echter een nieuw niveau aan complexiteit met zich mee. Dit komt door de dichtbevolkte cellulaire omgeving, het translatie proces en de verschillende posttranslationele modificaties waar eiwitten vaak mee te maken hebben. Om dit probleem te overwinnen zijn cellen vaak uitgerust met gespecialiseerde eiwitten die andere eiwitten helpen om efficiënt te vouwen, het mogelijk maken om verkeerd gevouwen en ongevouwen eiwitten opnieuw te vouwen, het verenigen in onbruikbare aggregaten voorkomt en zelfs de al gevormde aggregaten oplost. Deze eiwitten heten moleculaire chaperonne-eiwitten. De mechanismen die chaperonne-eiwitten hanteren om de integriteit van het proteoom te behouden onder normale omstandigheden en verschilende stress, zijn heel slecht begrepen, ondanks hun grote belang voor een scala aan ziektes (inclusief neurodegeneratieve ziektes en diabetes).

Het begrijpen van de mechanistische details van deze moleculaire machines is beperkt door serieuze experimentele uitdagingen. Interacties van chaperonne-eiwitten met substraat-eiwitten zijn vaak heterogeen en tijdelijk. Dit limiteert de toepasbaarheid van biochemische ensemble technieken op dit probleem. In dit proefschrift heb ik single molecule kracht spectroscopie in optical tweezers gebruikt om het effect van verschillende chaperonne-eiwit systemen op eiwitvouwingspaden te onderzoeken. Deze techniek is uitermate geschikt om de volgende vragen te beantwoorden, welke nauwelijks eerder in de literatuur onder de loep zijn genomen: Wat gebeurt er met een substraat-eiwit als een chaperonne-eiwit hier aan bindt? Aan welke eiwittoestanden kan een chaperonne-eiwit binden? Hoe kan een chaperonne-eiwit zulke verschillende taken vervullen door alleen maar een eiwit te binden en los te laten.

Hoofdstuk 1 presenteert de lezer met een kort historisch overzicht van het eiwitvouwingsveld en beschrijft de belangrijkste experimentele ontwikkelingen en uitdagingen in dit gebied. Het leidt belangrijke open vragen in en legt uit hoe single molecule technieken uitgebuit kunnen worden om ze te beantwoorden. Vervolgens licht het de in dit proefschrift bestudeerde klassen van chaperonne-eiwitten uit en presenteert een aantal voorbeelden van inzichtelijke single mole-

cule studies daarvan.

Hierop volgt hoofdstuk 2, welke zich richt op de technische kant van dit proefschrift. In dit hoofdstuk beschrijf ik de ontwerpprincipes die ten grondslag liggen aan het bouwen en assembleren van een optical tweezers opstelling. Daarnaast bevat het advies over de noodzakelijke onderdelen en een gedetailleerde handleiding voor de uitlijning van de opstelling voor degenen die geïnteresseerd zijn in het bouwen van hun eigen optical tweezers opstelling.

Hoofdstuk 3 presenteert nog een andere technische ontwikkeling, namelijk die voor een driftarme verwarmingsunit voor optical tweezers. Het functioneren van sommige chaperonne-eiwitten is gekoppeld aan een temperatuurstress die ons de motivatie geeft om onze sample te kunnen verwarmen. Maar verwarming in optical tweezers is niet zo vanzelfsprekend vanwege drift van de optische componenten en beperkte ruimte rondom het sample. In dit hoofdstuk presenteer ik een nieuw bedachte verwarmingsunit die de lucht rondom het sample en de objectieven gelijkmatig verwarmt. Gebruikmakend van deze methode zijn de mechanica van één molecuul onderzocht op verschillende temperaturen en ook voor het eerst gedurende verwarming.

Hoofdstuk 4 behandelt het moleculaire mechanisme van het Hsp70 chaperonne-eiwitten systeem. Het Hsp70 systeem is een centraal punt van chaperonne-eiwit activiteit in alle vormen van het leven. Het vervult een overvloed aan verschillende functies en het wordt geloofd dat het binden en loslaten van ongevouwen peptide regio's in een substraat-eiwit, gereguleerd door ATP en cochaperonne-eiwitten, hieraan ten grondslag ligt. Het is echter moeilijk gebleken om met conventionele technieken direct te kunnen observeren hoe het chaperonne-eiwit interactie heeft met substraat-eiwitten gedurende verschillende stadia van het vouwen. Dit is omdat deze interacties zijn gereguleerd door cochaperonne-eiwitten en ATP hydrolyse. Ook zijn ze vaak maar tijdelijk. We hebben gebruik gemaakt van optical tweezers om het effect van het complete Hsp70 chaperonne-eiwitten systeem en de losse onderdelen op het eiwitvouwingspad van het maltose bindend eiwit (MBP) te bestuderen. We demonstreren dat bacterieel Hsp70 (DnaK) niet alleen verlengde ongevouwen fragmenten van een substraat-eiwit bindt, maar ook gevouwen eiwit structuren. Gebaseerd op deze data, suggereren wij een uitbreiding van het Hsp70 kanonieke actie model, welke potentieel implicaties heeft voor alle Hsp70 functies.

Hoofdstukken 5 en 6 zijn toegewijd aan twee eukaryotische chaperonne-eiwitten, welke geclassificeerd zijn als "holdases". Er wordt geloofd dat ze willekeurig binden aan ongevouwen eiwit fragmenten met een vaste affiniteit, welke niet is gereguleerd door ATP. In beide hoofdstukken melden we onverwachte chaperonne-substraat interactie wijzes. Meer experimentele data van biochemische analyses zijn noodzakelijk om meer details te verwerven over de aard van deze interacties. Hoofdstuk 5 onderzoekt een intrinsiek wanordelijk chaperonne-eiwit Hsp33 die geactiveerd wordt door oxidatie. We laten voor de eerste keer zien dat Hsp33 kan binden aan eiwit structuren terwijl ze vouwen en hun volledige hervouwing kan verhinderen, terwijl het nagenoeg natieve conformaties toestaat. We hebben ook een simpel thermodynamisch model gebouwd welke overeenkomt met onze

experimentele data. In hoofdstuk 6 richt ik me op het effect van chaperonne-eiwitten op een proces dat ten grondslag ligt aan een scala van medische uitdagingen, namelijk α-synuclein (αSyn) aggregatie. In mensen leidt dit tot Lewy-body ziekten, welke een grote bedreiging vormen voor de vergrijzende bevolking. We hebben het effect van een chaperonne-eiwit HspB6 bestudeerd op de aggregatie van een eiwit construct samengesteld uit 4 herhaalde αSyn monomeren. Onverwacht wijzen onze resultaten erop dat HspB6, in plaats van het voorkomen van aggregatie, compacte structuren van αSyn bindt en ze stabiliseert tegen geforceerd ontvouwen. Onze data is ondersteund door een onderzoek met een elektronen microscoop, dat liet zien dat HspB6 bekwaam is om volwassen αSyn fibrillen te binden.

In hoofdstuk 7 bestudeer ik nog een andere stressresponse mechanisme, welke aanwezig is in organismes van verschillende domeinen in het leven. Het is gebleken dat de dissacharide trehalose verschillende biomoleculen zowel kan stabiliseren als destabiliseren, afhankelijk van het substraat en de concentratie trehalose. Men denkt dat deze eigenschap door organismes wordt gebruikt onder een verscheidenheid aan spanningen. Deze eigenschap wordt ook gebruikt in de industrie. Het volledig begrip van het achterliggende mechanisme ontbreekt nog. We laten zien dat trehalose de kracht voor de overstrekkingstransitie van dubbelstrengs DNA verlaagt en stellen een entropisch aandrijvingsmechanisme voor dit effect voor. Moleculaire dynamica simulaties lieten zien dat er een directe interactie bestaat tussen trehalose en de nucleobasen van het DNA.

Acknowledgements

When 4.5 years, which I spent at AMOLF working on my PhD project, passed at full tilt, there is now time to look back and understand what it was like. The life of a PhD candidate is not always a piece of cake, but I have truly enjoyed this experience. Most importantly, it is due to bright, cheerful, helpful and understanding people, which surrounded me at AMOLF and around it. I am happy to take the opportunity now and thank all of them for their constant support, since without them this book would clearly be not possible.

First of all, big thanks go to **Sander**, my promotor. He took a risk of hiring a person without solid experimental experience, and, under his clever supervision (which did not limit a vital for any researcher freedom to explore interesting things), I was able to grow effectively and deliver the results. Very importantly, he was always open to discuss the fresh results and roadmaps of our projects. His contagious excitement about science provided me with additional motivation to keep on going, despite long days and constant complications with getting the data in optical tweezers. **Sander**, thank you for this great adventure! I will clearly miss long chaperones discussions at my new job.

During these years at AMOLF, I was happy to be part of a great optical tweezers team. Driven by a common motivation to unravel the puzzles of chaperone-assisted protein folding and a daily goal to squeeze more data out of the machine, everybody was always happy to help with doing an experiment or interpreting the results. **Alireza,** thank you for teaching me how to handle the proteins, perform the experiments and analyze the results, but also for taking time to support me constantly. You (together with **Peter**) saved me a lot of frustration during my first steps into the realm of experimentalists and you always were a role model for me of true devotion to science. **Fatemeh** and **Arif,** together we survived those six months of frustration and proteins troubleshooting, thanks for keeping a positive attitude while making our way back to working experiment. It definitely made all of us stronger! It was always a pleasure to work in the lab with you and think of new possible projects. **David,** your knowledge of proteins and biochemistry is as broad as a lab bench you need to keep all your reagents! Thanks for your energetic improvement of our assays and your contribution to DnaK story. **Roeland,** you came to the group twice and you were always bringing something new. Thank you for teaming up with me on numerous projects, for your computational support, and for great time in California! **Mario,** a man of many talents, brought quite some fresh ideas on the Foldometer building and took over this project from me. Special thanks for making the DnaK story and my thesis much more beautiful! Success with the animations and your Russian studies! **Mohsin,** I did not have time to work with you a lot, but I am sure that you will make a great contribution to the group. Keep up the good work! I would also like to thank members of the single

cells and chemotaxis team, past and present: **Marjon, Philippe, Sarah, Rutger, Dmitry, Noreen, Martijn, Katja, Sebastian,** for their insightful input and great discussions during group meetings. **Vanda,** thank you for your countless help in the lab, for lots of great stories, and for being patient and understanding when your pipettes were missing!

During my PhD, I was fortunate to work with leading figures in the chaperones and protein folding fields. I take this opportunity to thank all the collaborators: **Bernd Bukau, Matthias Mayer** and **Guenter Kramer** from Heidelberg, **Ursula Jakob** from Michigan, **Michael Woodside** from Alberta, **Rachel Klevit** from Seattle, **Yi Qin Gao** from Beijing. It was a pleasure to work with you and I am immensely thankful to all of you for your critical input and help with shaping the projects, for your important experimental contributions, and for sharing proteins and chaperones with us. I am especially thankful to **Vinod Subramaniam** for his kindness, support and interest in my work. It was great to be a part of the αSyn working group!

I also thank all AMOLFers for their kindness, constant interest to science, and great support. Thanks to my nice office-mates: **Marjon, Peter, Corianne, Arif, Katja, Poulomi, Merlijn**, for keeping the local atmosphere warm and welcoming. In particular I am thankful to the technical support facility. **Duncan, Marco, Erik, Sjoerd, Henk-Jan, Hinco, Marko** - without your contributions I would never be able to finish the Foldometer!

I thank **Ron** and his group for epic volleyball tournaments, strong competition spirit and great celebrations of our numerous victories. AMOLF has definitely become less sportive when you moved to Maastricht!

Thanks to all my friends, old and new, for keeping me sane and sociable and for making sure that I do not spend too much time in the lab. **Мама,** спасибо тебе огромное за постоянную поддержку во всех начинаниях, понимание, любовь и заботу. Твоя помощь действительно неоценима. My brother **Alex,** thanks for setting a good example for me. Without you doing it earlier, I would have probably never gone to study to Moscow and then abroad. Most importantly, I am thankful to my wife **Inna.** You like no other know how hard it sometimes was for me to keep going, and only with you on my side and due to your positive attitude, constant support, love and understanding I could have done it and can do much more.

Curriculum Vitæ

Sergey Vladimirovich Bezrukavnikov

23-02-1988	Born in Kaluga, USSR.
2000–2005	Gymnasium №9, Kaluga
2005–2008	Undergraduate in Physics Moscow State University, Russia
2008–2011	Master's degree in Biophysics Moscow State University	Russian Cardiovascular Research Center *Master's project:* Modeling of substrate binding to mitochondrial ATP–ADP translocase *Supervisors:* Prof. dr. A.N. Tikhonov Dr. M.K. Aliev
2011–2015	PhD research in Biophysics group FOM Institute AMOLF, Amsterdam, The Netherlands *Thesis:* Heat shock proteins studied by optical tweezers *Promotor:* Prof. dr. ir. S.J. Tans

List of Publications

1. **S. Bezrukavnikov**, R. J. van Wijk, A. Mashaghi, H.-J. Boluijt, E. Claij, and S. J. Tans. *A low-drift heating method for optical tweezers*. Manuscript in preparation.
   Chapter 3

2. A. Mashaghi*, **S. Bezrukavnikov***, D.-P. Minde, R. Kityk, B. Zachmann-Brand, M. P. Mayer, G. Kramer, B. Bukau, and S. J. Tans. *Functional plasticity in Hsp70 enabled by diverse modes of client binding*. Manuscript in revision.
   Chapter 4

3. F. Moayed, **S. Bezrukavnikov**, C. Cremers, U. Jakob, and S. J. Tans. *Binding flexibility of internally disordered chaperone Hsp33 studied with optical tweezers*. Manuscript in preparation.
   Chapter 5

4. **S. Bezrukavnikov**, F. Moayed, K. Neupane, A. Iyer, M. Woodside, V. Subramaniam, R. E. Klevit, and S. J. Tans. *Small heat shock protein HspB6 facilitates the compaction of a-synuclein*. Manuscript in preparation.
   Chapter 6

5. **S. Bezrukavnikov***, A. Mashaghi*, R. J. van Wijk, C. Gu, L. J. Yang, Y. Q. Gao, and S. J. Tans. *Trehalose facilitates DNA melting: a single-molecule optical tweezers study*. Soft Matter 10, 7269-7277 (2014).
   Chapter 7

*Equal contribution