Chaperone-mediated protein rescue

A single-molecule study

Mario Jesús Avellaneda Sarrió

Chaperone-mediated protein rescue

A single-molecule study

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus prof. dr. ir. T. H. J. J. Hagen
chair of the Board for Doctorates
to be defended publicly on
Thursday 28 November 2019 at 10:00 o'clock

by

Mario Jesús Avellaneda Sarrió

Master of Science in Nanobiophysics
Technische Universität Dresden, Germany
born in Albacete, Spain.

This dissertation has been approved by the promotor.

Composition of the doctoral committee:

Rector Magnificus,	chairperson
Prof. dr. ir. S. J. Tans,	Delft University of Technology, promotor

Independent members:

Prof. dr. C. Dekker,	Delft University of Technology,
Prof. dr. H. H. Kampinga,	University of Groningen,
Prof. dr. H. Saibil,	Birkbeck University of London, United Kingdom
Prof. dr. M. Schlierf,	Technische Universitat Dresden, Germany
Dr. R. Rosenzweig,	Weizmann Institute, Israel
Dr. A. Mogk,	University of Heidelberg, Germany
Prof. dr. N. H. Dekker,	Delft University of Technology, reserve member

HUMAN BEINGS MAKE LIFE SO INTERESTING. DO YOU KNOW, THAT IN A UNIVERSE SO FULL OF WONDERS, THEY HAVE MANAGED TO INVENT BOREDOM

Sir Terry Pratchett, Hoghfather

CONTENTS

SUMMARY

The interaction between proteins is central not only to this thesis, but to most processes in the cell. After millions of years of evolution, the accomplished variety, complexity and beauty of the proteomic network is astonishing. When one realizes that these interplays rely on the delicate process of protein folding, a very special sort of protein interaction comes into play: that between molecular chaperones and their clients. Chaperones are specialized proteins crucial to protein folding. They are thought to guide polypeptides through their conformational search from synthesis, preventing alternative hazardous pathways, and to rescue proteins from misfolded and aggregated states.

Despite tremendous efforts spanning several decades, studying the details of how chaperones achieve their critical task remains challenging, due to their highly dynamic, heterogeneous and promiscuous nature. Recently, single-molecule techniques have provided invaluable insight into those mechanisms. **Chapter 1** compiles some of the most important milestones achieved by these approaches in the chaperone field. However, even these powerful techniques have their limitations, and ingenious new strategies need to be devised in order to address the intricacies of chaperone action. **Chapter 2** describes the development of one such novel approach, based on the combination of optical tweezers force-sensing and single-molecule fluorescence imaging, and enabled by a new protein-DNA tethering strategy. It allows the measurement of conformational changes of individual proteins within a complex, while simultaneously monitoring the composition of the latter. This method has enabled all the studies of this thesis and opened the door for further ground-breaking research. In the Chapter, we also show the efficiency and robustness of the approach, and demonstrate its potential with two key proof-of-principle applications. In particular, we study the dynamic binding and dissociation of the chaperone *Trigger Factor* to its protein substrate in the different folding states, providing the first direct observations of the reciprocal dependence between chaperone binding and client conformation. We find that Trigger Factor binds to unfolded substrates for up to tens of seconds, preventing their refolding.

Arguably the most archetypal molecular chaperone is the one that was first discovered back in the 70s. The chaperonin *GroEL* displays a hollow barrel-like structure, with an internal chamber that can be sealed by its co-chaperonin *GroES*. The box-like structure of GroEL-ES has fascinated for decades: polypeptides go in entangled, and somehow come out fully folded, but even the physical principle it employs is unknown. We address this central problem in **Chapter 3**, where we demonstrate that the folding of individual substrates is actively accelerated by GroEL, in a process that differs from mere aggregation suppression. We find that the GroEL cavity exerts attractive forces on its polypeptide substrates, driving their compaction and triggering tertiary structure formation even in slow-folding proteins. Surprisingly, this en-

hanced collapse is aided by GroES binding to the apical domains, but does not strictly require it. We hypothesize that polypeptide collapse is a general folding determinant that can be controlled in cells by GroEL-ES and other chaperones.

Despite the robust guidance of proteins through their proper folding pathway provided by many chaperones, a mere prophylactic approach is not sufficient. Often, proteins can elude the chaperone radar towards misfolded states and, ultimately, aggregation. Chaperones that do not prevent, but actively reverse those hazardous states are therefore crucial to maintain homeostasis. *ClpB* belongs to this category, and is key for the dissociation and reactivation of aggregated proteins in bacteria. However, the mechanism employed to achieve its task has remained unresolved, as the process has never been directly observed. In **Chapter 4** we present the first study to directly follow disaggregase activity in real time. We find that polypeptide loops are translocated processively by ClpB, and refold while emerging from the translocation channel. This process is remarkably fast and powerful, reflecting the challenges of disaggregation. Using a novel super-resolution technique that combines the optical tweezers and confocal fluorescence signals, we also observe that ClpB can translocate both arms of the loop, switching to a single-arm when encountering obstacles to avoid pore jamming. Together, our results depict ClpB as a perfectly equipped machine to deal with aggregates of various topologies and stabilities. We conjecture that polypeptide loop extrusion is the general extraction principle underlying all Hsp100 disaggregases.

A different yet unique sort of protein interaction is presented in **Chapter 5**. Here we study the effect of *catch bonds* on protein networks. While the lifetime of normal or slip bonds monotonically decreases when subjected to external forces, catchbonds become stronger under low forces. Notably, a wide range of cytoskeletal crosslinkers display catch bonding properties, suggesting an important role in providing biological matter of its extraordinary properties. However, it is challenging to identify the real effect of catch bonding on network mechanics, as many other cross-linker properties can be confounded. Here we used a novel single-molecule approach to characterize wild-type α-actinin 4 and its mutant as an ideal catch/slip bond pair. We then used a rheological assay to show that, although individual catch bonds are weaker than their slip counterpart, resulting actin networks are stronger. We hypothesize that the reduced strength of catch bonds results in a higher mobility that allows to redistribute loads in areas presenting tension inhomogeneities, thus preventing fracture and network failure.

Finally, in **Chapter 6** a series of preliminary and future experiments are described. These include novel applications of our approach to study other challenging aspects of protein folding and chaperone action, and new scientific questions emerging from and complementing the results of this thesis.

SAMENVATTING

Na miljoenen jaren van evolutie is de ontstane diversiteit, complexiteit en schoonheid van het proteomische netwerk verbazingwekkend. Het eiwit interactie netwerk staat centraal bij vele processen in de cel, maar ook in dit proefschrift. Wanneer men zich realiseert dat deze interacties afhankelijk zijn van het eiwitvouwingsproces, komt er en zeer speciale soort eiwitinteractie in beeld: die tussen moleculaire chaperonnes en hun substraten. Chaperonnes zijn gespecialiseerde eiwitten die cruciaal zijn voor het vouwen van specifieke eiwitten. Ze geleiden deze eiwitten in hun conformationele zoektocht, en verhinderen zo gevaarlijke alternatieve paden. Daarnaast zijn ze instaat om aggregaten en verkeerd gevouwen eiwitten te redden.

Ondanks enorme inspanningen de afgelopen decennia, blijft het een uitdaging om gedetailleerd te bestuderen hoe chaperonnes hun taak uitvoeren. Dit komt door de zeer dynamische en heterogene aard van chaperonnes en de vele interactie partners. Recentelijk hebben technologische ontwikkelingen het mogelijk gemaakt om interacties tussen enkele moleculen te observeren, deze enkel-molecuul technieken leverden inzicht in de chaperonne mechanismen van onschatbare waarde. **Hoofdstuk 1** bevat enkele van de belangrijkste mijlpalen in het chaperonne veld die bereikt zijn met enkel-molecuul methoden. Echter hebben zelfs deze krachtige technieken hun beperkingen en kunnen niet alle aspecten van de chaperonne werking bestudeerd worden, hiervoor moesten er nieuwe ingenieuze methodes worden bedacht om deze op te helderen. **Hoofdstuk 2** beschrijft de ontwikkeling van een dergelijke nieuwe methode, gebaseerd op een combinatie van gelijktijdige krachtmeting met een optisch pincet en de detectie van fluorescentie van een enkel-molecuul, mede mogelijk gemaakt door een nieuwe aanpak voor het verkrijgen van een eiwit met handgrepen van DNA. Deze nieuwe methode maakt het mogelijk om de conformatie veranderingen van individuele eiwitten te meten terwijl de samenstelling van een complex geobserveerd kan worden en vormt de basis voor alle studies in dit proefschrift en opent de deur voor verder baanbrekend onderzoek. Daarnaast laten we in dit hoofdstuk ook de efficiëntie en robuustheid van deze methodezien en demonstreren we de potentie met twee belangrijke proof-of-principle toepassingen. In het bijzonder bestuderen we de dynamische binding en dissociatie van de chaperonne Trigger Factor met zijn eiwitsubstraat in verschillende vouwtoestanden, wat de eerste directe waarnemingen van de wederzijdse afhankelijkheid tussen chaperonnebinding en eiwit vouwtoestand is. We vonden dat *Trigger Factor* zich tot tientallen seconden bindt aan ongevouwen substraten, waardoor deze niet meer kunnen vouwen.

Wellicht de meest archetypische moleculaire chaperonne is degene die voor het eerst werd ontdekt in de jaren '70. De chaperonne *GroEL* vertoont een holle tonvormige structuur, met een interne kamer die kan worden afgesloten door zijn cochaperonne *GroES*. De doosachtige structuur en vouwingsmechanismen van GroEL-

ES fascineren wetenschappers al decennia: een polypeptide gaat verstrengeld GroEL in en komt op de een of andere manier volledig gevouwen eruit. Zelfs het fysische principe dat het GroEL-ES complex gebruikt is onbekend. We behandelen dit centrale probleem in **Hoofdstuk 3**, waar we aantonen dat het vouwen van individuele substraten actief wordt versneld door GroEL in een proces dat verschilt van meer aggregatie-onderdrukking. We vinden dat de GRoEL-holte aantrekkingskrachten uitoefent op de polypeptidesubstraten, waardoor hun compactie wordt aangedreven en tertiaire structuurvorming wordt veroorzaakt, zelfs in langzaam vouwende eiwitten. Verrassend genoeg wordt deze verbeterde compactie geholpen door GroES die zich bindt aan de apicale domeinen, maar dit niet strikt vereist. We veronderstellen dat de polypeptide compactie een algemene vouwmechanismeis die in cellen kan worden uitgevoerd door GroEL-ES en andere chaperonnes.

Ondanks de robuuste begeleiding van eiwitten via een correcte vouwweg die door een corresponderende chaperonne wordt geboden, volstaat een loutere profylactische aanpak niet. Vaak kunnen eiwitten de chaperonne-radar ontwijken en verkeerd vouwen om uiteindelijk te aggregeren. Chaperonnes die deze gevaarlijke toestanden niet voorkomen, maar actief omkeren, zijn daarom cruciaal om de homeostase te behouden. *ClpB* behoort tot deze categorie en is de sleutel voor de dissociatie en reactivering van geaggregeerde eiwitten in bacteriën. Het desaggregatie mechanisme van ClpB was tot dusver onbekend, omdat het proces nooit direct is waargenomen. In **Hoofdstuk** 4 presenteren we de eerste studie om de desaggregaseactiviteit in rechtstreeks te volgen. We vinden wanneer ClpB eenmaal gebonden is aan zijn substraat, het meerdere polypeptidelussen kan transloceren zonder opnieuw te moeten binden, welke en opnieuw worden opgevouwen terwijl ze uit het ClpB translocatiekanaal komen. Dit proces is opmerkelijk snel en krachtig en weerspiegelt de uitdagingen van desaggregatie. Met behulp van een nieuwe superresolutietechniek die de optische pincetten en confocale fluorescentiesignalen combineert, zien we ook dat ClpB beide armen van de lus kan verplaatsen en overschakel op een enkele arm bij het tegenkomen van obstakels om een blokkade van het translocatiekanaal te voorkomen. Uit onze resultaten blijkt ClpB de perfect uitgeruste machine voor het oplossen aggregaten van verschillende topologieën en stabiliteiten. We vermoeden dat de translocatie van polypeptidelussen het algemene extractieprincipe is dat ten grondslag ligt aan alle Hsp100-desaggregasen.

Een andere maar unieke soort eiwitinteractie wordt gepresenteerd in **Hoofdstuk 5**. Hier bestuderen we het effect van zogenaamde *catch bonds* op eiwitnetwerken. Terwijl de levensduur van normale of slip bonds monotoon afneemt bij blootstelling aan uitwendige krachten, worden catch bonds sterker onder lage krachten. Opmerkelijk is dat een breed scala aan cross linkers van het cytoskelet eigenschappen vertoont van catch bonds, wat een belangrijke rol suggereert bij het verschaffen van biologische materie van zijn buitengewone eigenschappen. In dit hoofdstuk hebben we een nieuwe enkel-molecuul methode gebruikt om wild-type α-actinine 4 en zijn mutant te karakteriseren als een catch/slip bond paar. Vervolgens hebben we een reologische test gebruikt om aan te tonen dat, hoewel individuele catch bonds zwakker zijn dan hun tegenhanger, de actine-netwerken sterker zijn. We veronderstellen dat de verminderde sterkte van catch bonds resulteert in een hogere mobiliteit die het

mogelijk maakt krachten opnieuw te verdelen in gebieden die heterogeniteit van de spanning vertonen, waardoor breuk- en netwerkfalen worden voorkomen.

Ten slotte wordt in **Hoofdstuk 6** een reeks voorlopige en toekomstige experimenten beschreven. Deze omvatten nieuwe toepassingen van onze methodes voor het bestuderen van andere aspecten van eiwitvouwing en chaperonnes en de nieuwe wetenschappelijke vragen die voortkomen uit- en een aanvulling vormen op de resultaten van dit proefschrift.

1

KEEP A CLOSE EYE ON YOU

MOLECULAR CHAPERONES AT THE SINGLE-MOLECULE LEVEL

Protein folding is well known to be supervised by a dedicated class of proteins called chaperones. However, the core mode of action of these molecular machines has remained elusive due to several reasons including the promiscuous nature of the interactions between chaperones and their many clients, as well as the dynamics and heterogeneity of chaperone conformations and the folding process itself. While troublesome for traditional bulk techniques, these properties make an excellent case for the use of single-molecule approaches. In this review, we will discuss how force spectroscopy, fluorescence microscopy, FCS and FRET methods are starting to zoom in on this intriguing and diverse molecular toolbox that is of direct importance for protein quality control in cells, as well as numerous degenerative conditions that depend on it.

1.1. INTRODUCTION

In order to perform their role within cells, proteins typically interact with a limited set of the proteome. Histidine kinases recognize specific response regulators in order to transmit detected signals, while kinesins move along microtubules to deliver neuro-transmitters in axons.

Molecular chaperones are a notable exception to this general rule. This class of proteins is involved in assisting a wide range of proteins throughout their life cycle. As soon as a newly synthesized polypeptide emerges from the ribosomal tunnel, chaperones bind and protect it against aggregation with other polypeptides and proteins, promote proper folding into a functional structure, and pass it on to other chaperone systems [1, 2]. However, the function of chaperones is not limited to de-novo folding. They act to disrupt already formed aggregates [3], help the formation of multiprotein complexes [4], regulate the activity of large numbers of receptors and kinases [5], and are involved in a range of other tasks. As such, chaperones are implicated in many normal cellular processes such as the cell cycle and apoptosis, but also in numerous medical conditions ranging from cancer to neurodegeneration diseases [6, 7].

Most chaperones are constitutively present but overexpressed at high temperatures, as well as under oxidative stress [8], deviating pH, and various other conditions [9]. Many chaperones owe their name to this effect, and since their discovery in 1974 are hence referred to as heat shock proteins followed by their molecular weight (e.g. Hsp70) [10]. Some chaperones undergo important structural changes triggered by the ATP hydrolysis cycle. For instance, Hsp60, known in bacteria as GroEL, is a barrel-like structure that can accommodate (mis)folded proteins and is closed-off by the GroES cap. Also, Hsp70 is known to bind exposed polypeptides in a groove that can be covered by a helical lid. Chaperones have therefore been referred to as 'folding machines', though this term does not do justice to their far wider range of cellular functions.

Despite the huge amount of knowledge acquired in the last decades, it is striking that many of the most basic questions remain unresolved to this day. For instance, it is still debated whether chaperones can directly guide and promote folding beyond suppressing aggregation. Merely detecting whether chaperones interact with partially folded chains along their folding pathway is already a challenge. When they do appear to promote folding, the physical principle is obscure, and may range from affecting chain entropy to recognizing key transition states of the folding protein [11]. The list of open questions is endless: interaction sites on chaperone and client protein are often unknown, as is the interplay between ATP hydrolysis, chaperone and client conformational changes. The reason for these gaps in our knowledge is clear: conformational changes and folding transitions are hard to measure in ensemble measurements. Indeed, in bulk refolding assays it is even a challenge to distinguish (reversible) aggregation from intrinsically delayed folding. Other complicating factors are the transient nature of chaperone-client complexes, the conformational dynamics of the chaperone, and the involvement of numerous co-factors. These technical challenges can in principle be addressed by zooming into single client-chaperone complexes.

The past decade has witnessed a rapid development of novel single-molecule approaches that are now beginning to address these crucial questions. Diverse tech-

niques have been employed, ranging from single-molecule FRET to optical tweezers and atomic force microscopy. Here we discuss a number of example studies that reflect these efforts – without aiming to systematically cover this field – and mention complementary bulk approaches where appropriate. We have organized the studied chaperone systems by their complexity, starting with the ATP independent chaperones trigger factor and SecB, and then moving to the ATP dependent chaperone classes Hsp70, Hsp90, and GroEL. This exciting first look at the action of chaperones at the single-molecule level is revealing a range of unexpected mechanisms, and first answers to big open questions.

1.2. Trigger factor, a cradle for nascent chains

The chaperone trigger factor (TF) is the first protein that most newly synthesized proteins interact with in bacteria [12]. This dragon-shaped [13] protein (Fig. 1.1A) associates with the ribosome with its tail bound close to the ribosome exit site, and the body and arms forming a cradle that receives the nascent chain when it emerges from the exit tunnel. TF can leave the ribosome while bound to the nascent chain [14] and suppress their aggregation [15]. With the latter function, TF exhibits functional overlap with the chaperones DnaK, GroEL [16, 17] and SecB [12, 18], which do not directly bind the ribosome.
A number of key questions have remained difficult to address with approaches used so far. Specifically, it is difficult to obtain structural information on client-chaperone complexes owing to the conformational dynamics of the unfolded polypeptide clients. We also lack information on how TF affects these conformational dynamics, and the process of folding into active proteins with tertiary structure. These questions, which go to the heart of chaperone functions, are now beginning to be addressed by single-molecule methods, as well as by computational approaches and NMR spectroscopy. Here we discuss a few of these recent studies, and contrast them to findings on another independent chaperone, SecB.

Structural data of TF-substrate complexes has been lacking until recently, due to the transient nature of the underlying interactions and conformations. Pushing the envelope of the size of protein systems addressable by NMR spectroscopy, Saio et al. [19] studied the interaction between TF and alkaline phosphatase PhoA. PhoA is a periplasmic protein that remains in an unfolded state under reducing conditions. The data indicated four substrate binding-sites on TF for unfolded PhoA: three on the body and arms and one on the head, all highly enriched in nonpolar residues. They found the same binding sites, and an additional one on one of the TF arms, for an unfolded fragment of maltose binding protein (MBP, see Fig. 1.1A) and the trans membrane region of OmpA, suggestion some commonality in where substrates bind to TF. The authors also probed the interaction sites on the substrate PhoA, and found that they are –in addition to non-polar residues - rich in aromatic residues. In contrast, hydrophobic stretches lacking aromatic residues seemed to have low affinity for TF. The binding sites were promiscuous: each site could bind some or even all of the TF sites with relatively low affinities. An encounter between unfolded PhoA and TF should thus result in a dynamic search for the combination of TF-PhoA binding sites
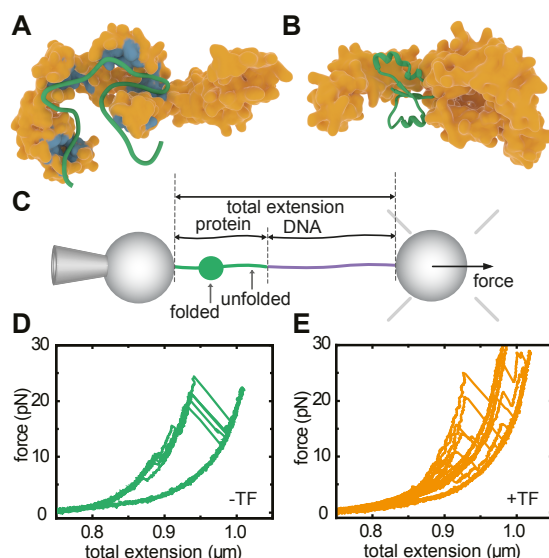
1

Figure 1.1: Interactions between Trigger factor and client proteins. **(A)** Interaction sites on TF for MBP as derived from NMR experiments [19]. **(B)** Interaction of TF with a partial fold of MBP, as determined by MD simulations [20], and observed by optical tweezers experiments (panels C-E). **(C)** Single-molecule optical tweezers experimental setup with MBP tethered between two polystyrene beads. One bead is held on a pipette, while the other is held by an optical trap that is also used to determine the applied force. Pulling experiments on MBP in isolation **(D)** and MBP with TF present **(E)** show an increased presence of partially folded states for the latter, during pulling and also during refolding at low force in between pulling cycles. Panel A is re-drawn from Saio et al. [19], panel B from Singhal et al. [20], panels C-E from Mashaghi et al. [21].

that have the lowest energy.

The dynamic nature of the TF-bound protein chain presses the question whether it can form tertiary structure. This issue requires the ability to trigger folding, and the ability to follow it in time, which can be achieved with optical tweezers. Using this single molecule technique, Mashaghi et al. tethered MBP to beads using DNA linkers (Fig. 1.1C), and unfolded and refolded them in repeated cycles of pulling and relaxation [21], showing a reproducible folding behavior of the protein (Fig. 1.1D). These force-extension curves changed dramatically upon addition of TF (Fig. 1.1E). First, unfolded proteins remained unfolded for longer, consistent with TF-chain binding, but then did form tertiary structure. More surprisingly, these partially folded states were stable for seconds and against applied force, and folding now proceeded via these intermediate states that were promoted by TF. Thus, this approach provided direct evidence for how folding guidance by chaperones is sometimes imagined: to continue interacting with a protein chain during the process of folding into tertiary structure. Local conformational plasticity of TF is relevant to this behavior; the flexibility of TF's two arms facilitates the binding of folded substrates of a variety of sizes

[22]. Moreover, TF was found to not only promote refolding of MBP monomers within a 4xMBP repeat-construct, but also to suppress misfolding interactions between them. These data suggested a generic mechanism to separate the good from the bad: multiple TF molecules bound to the different domains within multi-domain proteins that suppress erroneous interactions between domains while allowing native interactions within domains.

To gain further structural insight into substrate-TF complexes in these different stages of folding, MD simulations have been employed [20]. Simulations on unfolded MBP conformations revealed some of the same sites as detected by NMR. Intermediate folded states initially formed a 'touching complex' with the flexible tips of two TF domains. Interestingly, subsequent transfer to the TF cradle and embrace by its flexible arms provided a structural explanation of the experimentally observed stabilization of folded structures (Fig. 1.1B). Substrate-TF interactions became weaker for more fully folded states, which makes sense functionally.

A next step would be to assess the chaperoning action of TF on the ribosome. Biochemical assays on stalled ribosomes suggest that TF can then also delay folding of large multi-domain proteins [23] and even partially unfold some substrates [24]. Course grained simulations suggested that this folding delay is caused by kinetic trapping of unfolded ensembles, while smaller proteins could fold in between TF and the ribosome without delay. In this manner, the chaperone effectively lengthens the tunnel of the ribosome with its space-limited cradle [25]. Single-molecule experiments on stalled ribosomes have shown that formation of tertiary structure in nascent chains is suppressed due to confinement by proximity of the ribosome [26]. The chaperone SecB presents interesting similarities and differences with TF. SecB also interacts and stabilizes unfolded chains [18], which here facilitates their transport across the membrane by the SecA translocation machinery. A recent NMR study [27] revealed how an unfolded PhoA chain wraps around the chaperone SecB. Long, hydrophobic grooves on the chaperone tetramer facilitate binding of the substrate. The parts of PhoA in contact with SecB are fully unfolded with no secondary structure present, and the interaction surface – as deduced from the modeled structure of the complex - turns out to be much larger than that of PhoA with TF (250 vs 25 interacting residues). This larger interaction surface might explain the stronger anti-folding properties of SecB compared to TF. The latter was also consistent with single-molecule force spectroscopy, which indicated that SecB keeps MBP substrates in an unfolded state by preventing the formation of stable tertiary interactions [28].

A picture emerges of TF as a more versatile chaperone than commonly assumed. It forms the first line of defence against aggregation of the nascent chain, and protects freshly synthesized hydrophobic stretches from aggregation. TF also binds and transiently stabilizes partially folded structures, which protects them from long-range interactions at the cost of reduced folding rates. By holding unfolded as well as partially folded states in a transient and ATP-independent manner, TF can deepen corresponding energy valleys and guide folding trajectories along them. These initial insights into the structure and dynamics of TF-substrate complexes raise a host of novel questions. For instance, it remains unclear how the conformational dynamics of the substrate chain is affected when TF is bound, whether TF remains fully

bound during folding transitions, or rather leaves transiently, or how TF can partially compensate for deletions of ATP-dependent chaperones DnaK and GroEL [16, 17]. Another open question is how the role of TF differs at the ribosome [24]. A recent study suggests additional substrate binding sites on the tail of TF (ribosome binding domain) that only becomes available upon ribosome binding [29]. Single-molecule studies of ribosome-client-chaperone complexes are within reach [26], and could shed light on these important questions on co-translational chaperone action. A suggested hydrophilic binding mode of TF [15, 20] may also stimulate further structural and single-molecule investigation. Finally, it is of interest to determine how TF differs mechanistically from eukaryotic chaperone systems that fulfill similar functions (see [30] for an overview). These insights may also find practical use, for instance in helping to reduce the misfolding rate of bacterially produced eukaryotic proteins [23, 31].

1.3. HSP70, A CLAMP FOR UNFOLDED AND FOLDED PROTEIN STRUCTURES

The 70kDa heat shock proteins (Hsp70s) are one of the most ubiquitous families of chaperones, and are highly conserved across species. They are involved in a remarkably diverse range of cellular processes, well beyond assisting in de-novo protein folding. Other roles are for instance the disaggregation of already formed aggregates [3], assistance in protein trafficking across membranes, and regulating the activity of kinases and receptors [5]. Hsp70s are thought to interact with unfolded peptide chain segments extending from substrate proteins, which may be in (partially) unfolded of misfolded conformations. In addition, auxiliary co-chaperones interact with Hsp70s and regulate their activity [32].

Hsp70s consist of two distinct domains, a C-terminal substrate binding domain (SBD, 27 kDa) and an N-terminal nucleotide-binding domain (NBD, 43 kDa), connected through a highly conserved linker. An important feature of the SBD is its two subdomains, a twisted β-sandwich (SBDβ) and an α-helical (SBDα) subdomain ending in an unstructured stretch of about 30 residues, widely referred to as the chaperone lid [33, 34]. High-resolution crystal structure studies revealed two conformations of the chaperone that have been very instructive in understanding peptide binding, as discussed further below (Figure 2A). Hsp70 acts as a clamp: in the closed conformation, observed in the nucleotide-free and ADP-bound states, the lid (SBDα) is positioned closely against the peptide binding cleft on SBDβ and both subdomains are spatially separated from the NBD [33]. In the open conformation, the lid is detached from SBDβ, and both subdomains dock to different parts of the NBD [35].

The ATP cycle is important for its peptide-binding function: in the open ATP-bound state, association and dissociation rates are high, resulting in low substrate affinity. In the closed ADP-bound state, both rates are several orders of magnitude lower, leading to a higher affinity for polypeptides [36]. The chaperone-polypeptide interaction and the nucleotide state of the chaperone are strongly coupled and affect each other reciprocally. For instance, peptide binding catalyzes ATP hydrolysis, which is otherwise a rather slow process. ATP conversion to ADP, in return, stabilizes peptide binding

[37]. Certain co-chaperones, such as DnaJ in E. coli or Mdj1 in mitochondria, play important roles in these interactions [38].

Crystal structures provide a detailed yet static picture, without information on dynamics. For instance, they do not reveal whether the two Hsp70 conformations correspond strictly to a particular nucleotide state, or whether other intermediate conformations exist. This problem was addressed by Mapa et al., choosing the chaperone Ssc1, a mitochondrial member of the Hsp70 family, using ensemble and single-pair Förster Resonance Energy Transfer (FRET) as experimental techniques [39]. The authors engineered two FRET-constructs using strategically selected cysteine residues for labelling (Fig. 1.2A). For the first construct, two dyes were introduced in the NBD and the SBD in order to study the interaction between the two domains. The second contained cysteine residues in the lid and the base of the SBD, to study the dynamics of these subdomains. For the single-pair FRET experiments, a very low concentration of Ssc1 (20-40 pM) was used, ensuring that only individual proteins were probed with a confocal microscope and pulsed interleaved excitation. Both donor and acceptor emissions of at least 500 different molecules were independently recorded and combined in a FRET-efficiency histogram (Fig. 1.2). The results demonstrate that the conformation of Ssc1 in the ATP-bound state is well defined, with the lid detached from SBDβ and the NBD and SBD docked (Fig. 1.2B-C, left panels), in agreement with structural data. In contrast, the ADP state of Ssc1 is much more heterogeneous, both in the SBD-NBD interaction and in the lid conformation. Similar behavior was observed by the same authors for the bacterial Hsp70 (DnaK), with the NBD-SBD domains largely separated in the ADP state (Fig. 1.2B, right panel) as the only difference with its mitochondrial counterpart, showing similar heterogeneous lid dynamics (Fig. 1.2C, right panel) [39].
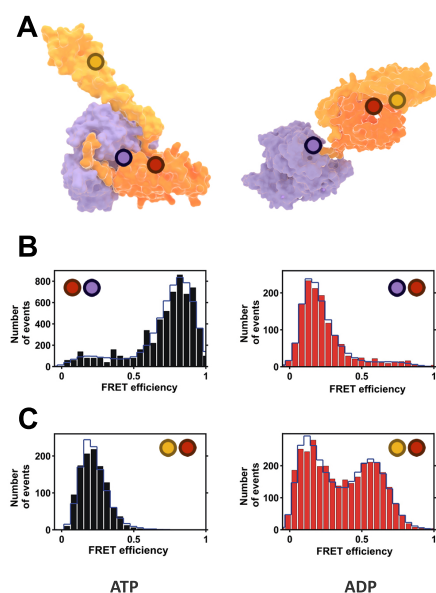

Figure 1.2: Single-molecule FRET experiments with Hsp70. (A) Crystal structures of Hsp70 open (left) and closed (right) conformations. Purple corresponds to the NBD subdomain and orange and yellow to the SBDβ and SBDα subdomains respectively. The circles denote the approximate location of the donor and acceptor labels described in [39]. **(B)** FRET histograms for the inter-domain dynamics under ATP (left panel, docked domains) and ADP (right panel, undocked domains) conditions. **(C)** FRET histograms for the lid dynamics under ATP (left panel, open lid) and ADP (right panel, heterogeneous state) conditions. (B) and (C) are adapted with permission from [39].

As mentioned, substrate-Hsp70 interaction and nucleotide cycle are tightly coupled and strongly dependent on co-chaperones. The details of this link were also explored [39]. The addition of a peptide substrate greatly accelerated ATP hydrolysis and resulting domain undocking and lid-closure; with evidence of these two processes happening virtually simultaneously. In presence of the Hsp40 Mdj1, a co-chaperone of Ssc1, both events were further accelerated. The lid was found to adopt a stable closed conformation over the binding cleft in presence of both the substrate and Mdj1. Interestingly, when the substrate is absent, Mdj1 is able to trigger domain undocking and lid closure, but after a short period of time the conformation changes back to that of the heterogeneous ADP-bound state. Another co-factor, nucleotide exchange factor (NEF), also plays a role in accelerating the ATP cycle by promoting the exchange of ADP by ATP in the NBD of Hsp70s.

These chaperone dynamics press an urgent question: how is the substrate affected? Single-molecule FRET was employed by Kellner et al. to study the conformation of different rhodanese-fluorophore constructs when they interact with the bacterial Hsp70 (DnaK) and its co-chaperone DnaJ [40]. Five different FRET pairs were engineered to monitor changes in different parts of the rhodanese polypeptide. A small probing volume was illuminated with pulsed excitation, and the emission of single molecules detected to obtain FRET histograms. Upon chemical denaturation, the FRET efficiency of all variants remained fairly high, suggesting a rather compact state. Importantly, refolding to the native state occurred spontaneously on a timescale of minutes without chaperones. Presence of DnaJ resulted in a broadening of the FRET distributions, indicating the blocking of refolding and the formation of heterogeneous non-folded conformations. Addition of DnaK to preformed DnaJ-rhodanese complexes led to a shift towards lower FRET values. This observation, supported by molecular dynamics simulations, suggested that several DnaK molecules bind to the peptide chain, resulting in its expansion by means of volume exclusion [40]. Interestingly, DnaK was not enough to drive substrate expansion, as the presence of DnaJ was essential for this process.

One of the most enigmatic aspects of Hsp70s is the role of the lid. While it is generally assumed to serve to stabilize peptide binding only [32], recent single-molecule studies have demonstrated a broader functional role. Mashaghi et al. recently employed optical tweezers [41] to mechanically control the folding state of MBP substrates and studied the response upon addition of DnaK. Surprisingly, these data showed that not only exposed peptide segments were stabilized, but also near-native folded protein structures: in presence of the bacterial Hsp70, the folded structures displayed high unfolding forces, or at times they could not be unfolded within the force limits of this method (up to 65 pN). Fully folded native structures were not stabilized – a minor unfolding transition that removes a number of external MBP alpha-helices was required to trigger stabilization by Hsp70. The authors further showed that both the lid and ADP are key to the stabilizing function. This mode of binding and stabilizing folded structures extends the longstanding canonical model of Hsp70, in which only extended peptides are bound and released. Notably, it has essentially the opposite effect to the known binding mode, as it stabilizes folded rather than unfolded states. Co-chaperones and nucleotide concentration may play an important

role in regulating the different modes of Hsp70s binding, as stabilization occurs only in the ADP state and is thus transient. This extended role of the lid is consistent with observations of its conformational dynamics. Marcinowski et al. used single-pair FRET, in a similar fashion as Mapa et al. [39], to resolve the conformational changes of the mammalian Hsp70 BiP (heavy chain-binding protein) [42]. The chaperone conformational dynamics during the nucleotide cycle was analogous to that of Ssc1 and DnaK, including lid closure upon peptide binding. When the peptide substrate was replaced by a larger unstructured client protein, however, the lid remained predominantly open, while cross-linking experiments revealed physical interactions between lid and bound substrate. Interestingly, it was also shown that the addition of ERdJ3, a co-chaperone of BiP, primes the conformation of the latter for protein substrate binding, while hindering peptide binding. Again, these results manifest the intricate interaction between chaperone, nucleotide, substrate and co-chaperones. Similar evidence for the lid versatility of Hsp70s was recently shown by Banerjee et al, which used smFRET to study the dynamics of the lid from DnaK [43]. Here, the lid remained mainly open in the presence of proteins in a molten globule state, in contrast to its closure upon peptide binding.

A number of intriguing questions arise from these findings, such as how Hsp70s can discriminate between partial folds that are native-like or misfolded, and more generally whether this direct binding of folded structures allows Hsp70 to actively fold proteins. The findings also suggest that in episodes of stress, when ATP levels are low and ADP levels are high, Hsp70-mediated stabilization may keep key proteins intact, as supported by increased unfolding temperatures of the protein RepE in the presence of Hsp70 and ADP [41]. It will also be interesting to determine how this novel binding mode affects the cooperation between Hsp70 and its co-chaperones, as well as other chaperones such as Hsp90.

1.4. HSP90, A VERSATILE REGULATORY CHAPERONE

The 90-kDa heat shock proteins (Hsp90s) constitute an essential chaperone family in bacteria and eukaryotic organisms. Like Hsp70, they participate in a broad spectrum of cellular processes, including heat stress protection, signal transduction and protein trafficking [44]. While the bacterial homologue HtpG is not essential in normal conditions, Hsp90s are critical in eukaryotic cells and make up to 1-2% of total soluble cell protein [45]. Hsp90 also plays an important role in certain disorders, including cancer, and has in the recent years emerged as a potential target for tumor treatment strategies [7]. Hsp90 forms a high affinity dimer (Fig. 1.3A), with each monomer consisting of three domains: a highly conserved N-terminal nucleotide binding domain (NTD), a middle domain (MD), and a C-terminal dimerization domain (CTD).

An interesting feature of Hsp90 is its unique ATPase activity. A series of early studies reported contradictory observations, suggesting both the existence and absence of the ATPase activity of Hsp90 [46, 47]. This controversy was resolved by a crystallographic study that revealed a nucleotide binding site on the N-terminal domain [48]. Subsequent work revealed that Hsp90 dimer adopts at least two conformations in a clip-like manner (Fig. 1.3A): an ATP-free open state in which only the C-terminal

domains interact, and an ATP-bound closed state, in which the other domains also interact (Fig. 1.3A) [49, 50]. In other studies, three to four states have been suggested [51, 52]. Biochemical assays have suggested that, in the particular case of bacterial HtpG, these structural changes are tightly coupled to the nucleotide cycle, contrary to eukaryotic Hsp90s, for which conformational dynamics are more subtle and yet not clearly related to the remarkably slow ATPase activity [53]. However, such heterogeneous ensembles of molecular states makes are difficult to characterize properly using bulk assays.
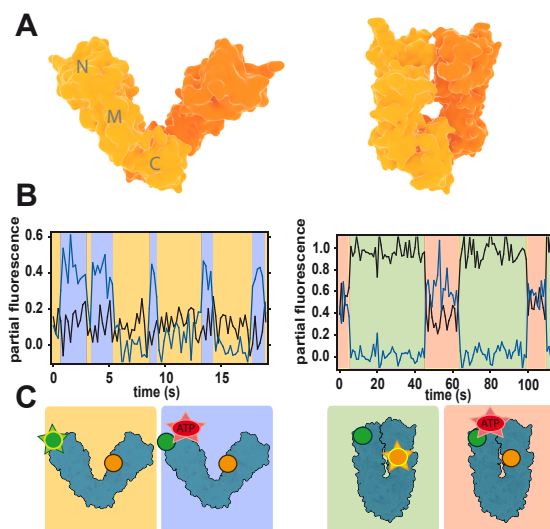

Figure 1.3: Conformational changes of Hsp90 studied with FRET. (A) Crystal structures of open (left) and closed (right) conformations of bacterial Hsp90 dimer (monomers are indicated by different color shades). **(B)** Partial fluorescence traces of two acceptors in 3-colour FRET experiments: black line corresponds to the NTD acceptor, blue line to nucleotide acceptor. The traces are calculated by dividing acceptor intensity by the total fluorescence signal [54]. Data shows that nucleotides can bind Hsp90 dimer in both open and closed conformations. **(C)** Scheme of the conformations and labeling of Hsp90. Green circle is donor, yellow is acceptor monitoring NTD dynamics and red is the nucleotide acceptor. Emission is represented by a star. Background colors link each conformation to the corresponding portion of the fluorescence traces in (B). Figures (B) and (C) are redrawn from [54].

This problem has been addressed in several single-molecule studies, the first of which used FRET to investigate the dynamics of the N-terminus dimerization [55]. Here, the authors created two different single-cysteine mutants of the yeast Hsp90, labeling each of them with an acceptor and a donor fluorophore, respectively. The formation of heterodimers produced an increase of the acceptor signal, further amplified when the chaperone adopted a closed state. Saturating ATP conditions led to a continuous switching between open and closed conformations on the seconds timescale, much faster than the expected 100 s ATPase cycle. Analysis of dwell times showed that the dynamics was best described by four states, two open and two closed,

revealing eight different kinetic rates between them, as well as a simplified energy landscape. In the presence of ATP, two of the kinetic barriers were lowered, but all states could be accessed spontaneously even in the absence of nucleotide. Together, all these results imply that the large conformational changes of the NTDs and the ATP-cycle are only weakly coupled for yeast Hsp90, and mostly driven by thermal fluctuations.

The same authors extended these observations in a second study using 3-color FRET [54]. Combining the labeled Hsp90 heterodimer with a second acceptor dye attached to the nucleotide (either ATP or ADP) allowed simultaneous detection of conformational changes and nucleotide binding events. Both ATP and ADP were found to bind open and closed states of Hsp90 with slightly different rates, strengthening the importance of thermal fluctuations (Fig. 1.3B-C). In addition, it was found that ATP binds on and off multiple times before it is hydrolyzed, contrasting with the long established idea that the chaperone remains in an ATP-bound "waiting state" until hydrolysis occurs.

In a parallel study, the differences between yeast and bacterial Hsp90 were investigated. For HtpG, it was found that the NTD conformational dynamics and the ATP cycle are strongly coupled by a mechanical ratchet mechanism [56]. These observations suggest that Hsp90s evolved from the rigid bacterial, nucleotide-regulated chaperone to its more flexible eukaryotic counterpart. This feature may have allowed the chaperone to adapt to a larger range of substrates without additional energy cost. It may also enable a more versatile and sophisticated regulation by co-chaperones, in addition to the ATPase activity. This might explain the extensive number of co-chaperones for eukaryotic Hsp90, while none has been found for the bacterial HtpG.

In order to investigate regulation of yeast Hsp90 by the co-chaperone p23, the FRET strategy was extended to four colors. The observations suggested that the ATP turnover regulates p23 binding, without a direct impact on the large NTD conformational changes, and is the interaction with the co-chaperone what provided the Hsp90 machinery its directionality [57]. Further evidence of co-chaperone regulation was found using single-molecule photoinduced electron transfer (PET) to study the intra-subunit domain interactions [58]. These local conformational changes in the NTD and MD, though more coupled to the ATPase activity than the inter-subunit NTD dimerization, are strongly catalyzed by Aha-1, another co-chaperone of yeast Hsp90.

Together, these results illustrate the novel insights that single-molecule experiments can provide. Key mechanisms have been revealed on the conformational dynamics of Hsp90 and its relation to the ATPase cycle and co-chaperones. An additional and critical element involved in the chaperone regulation is substrate binding, but its detailed impact on Hsp90 dynamics remains elusive. Conversely, how Hsp90 affects substrate conformations also remains largely unknown. A recent study showed that it is possible to monitor both monomer and dimer Hsp90 folding states using optical tweezers [59], while the same tool has been used to investigate how chaperones affect protein folding [41]. The extraordinary versatility of Hsp90 makes this an outstanding challenge, as diverse substrates may affect and be affected in different fashion.

1.5. GROEL, A CONFINING BARREL

The GroEL-GroES system is arguably the most studied molecular chaperone. The GroEL barrel-like structure is composed of 14 identical subunits of 57 kDa each that are stacked as two heptameric rings (see Fig. 1.4A) [60, 61]. Each monomer has three domains: an apical domain that binds with polypeptides and GroES, an intermediate domain, and the nucleotide binding equatorial domain [62]. GroEL functions with its co-chaperonin GroES that acts as a lid for encapsulating nonnative proteins (see Fig. 1.4A, left-top panel) and is composed of 7 subunits of 10 kDa each. ATP binding in the equatorial domain of GroEL results in large structural movements in the apical domain and exposure of hydrophobic residues that facilitate GroES binding, in turn doubling the size of the GroEL cavity, such that proteins of up to 60 kDa size can be encapsulated [63].

Early stopped flow fluorescence anisotropy and enzymatic activity studies have been instrumental in revealing many aspects of the ATP hydrolysis cycle [64–66], while cryo-EM and X-ray crystallography pushed understanding of the corresponding GroEL structural changes [67–69]. Non-native polypeptides are thought to first bind the GroEL apical domain, after which the binding of GroES drives them into the central GroEL cavity where folding takes place. Binding of a second substrate and GroES on the other ring of the double-barrel GroEL structure triggers release of the folded substrate protein. Despite the detailed biochemical and structural knowledge that has been amassed, the core folding mechanism remains contested [70]. GroEL-GroES could act as a passive folding cage by physically protecting substrates from aggregation, actively catalyze the folding of individual substrates, or exert pulling forces on misfolded conformers in order to unfold them and allow autonomous refolding to the native state. Detailed questions also remain unanswered on various other aspects, including the cooperation between the two rings, the precise role of GroES, and the substrate-accepting state of GroEL.

Compared to the other chaperones, the GroEL-GroES system has been extensively probed with single-molecule fluorescence approaches. One recent example is by Lin et al. [72], where using intra-molecular FRET the authors observed the binding of nonnative Rubisco to a nucleotide free ring of GroEL with ATP and GroES bound to the other ring (also termed as ATP bullet). The results suggested that the ATP bullet is the polypeptide accepting state, which is also consistent with the asymmetric functioning of GroEL. In this asymmetric model, release of ADP from the trans–ring is catalyzed by peptide binding, which in turn triggers ATP hydrolysis in the cis ring [73]. This notion was supported by recent single-molecule studies on symmetric or football-shaped GroEL-GroES complexes [74, 75]. Saturated substrate concentrations and slow ATP hydrolysis were found to promote symmetric complexes over asymmetric ones. Most studies on the symmetric complexes have been performed using the GroEL variant D398A that hydrolyzes ATP more slowly. Takei et al. [76] studied the football complex with a fluorescently labeled D398A variant and using GFP as the substrate protein. Single-molecule total internal reflection fluorescence (TIRF) microscopy was used to localize the positions of GFP molecules at the GroEL-GroES positions [77] (Fig. 1.4 B). The authors interestingly observed that two GFP molecules can fold simultaneously within the two GroEL cavities capped by GroES,
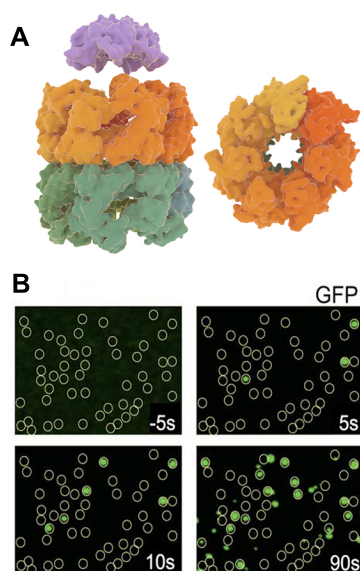
Figure 1.4: GroEL-GroES structure and folding of GFP by the complex. **(A)** GroEL side view (left – bottom) and top view (right) with its two heptameric rings and co-chaperonin GroES (left – top). **(B)** Fluorescence images acquired by total internal reflection fluorescence microscopy (TIRFM), showing GroEL positions as yellow circles and folded GFP molecules as green dots co-localized with GroEL [71]. Folding kinetics of individual GFP molecules was measured by acquiring the fluorescence images at different times. Panel B is adapted from [71].

with refolding kinetics similar to those observed in the asymmetric complex. The physiological significance of symmetric complexes has been a matter of debate, with a number of bulk studies considering only the asymmetric complexes as part of the functional cycle in-vivo [65, 78]. On the other hand, recent single-molecule studies have provided observations of symmetric complexes working as parallel folding machines, which may be a more productive protein folding state than the asymmetric complexes [76, 79, 80]. Moreover, during stress conditions when concentrations of nonnative proteins are high and negative cooperativity between the rings decreases [81], the formation of symmetric complexes should be favored.

Another conundrum about GroEL-GroES chaperonin is the process of protein folding itself. One debate is whether GroEL-GroES acts only as a passive or Anfinsen cage that simply prevents protein aggregation [82], with the encapsulated protein folding essentially autonomously [83], or whether the complex actively assist and accelerates protein folding, for instance by smoothening the folding landscape [84–86]. Another model, known as iterative annealing, proposes that GroEL-GroES functions by unfolding misfolded proteins, which subsequently fold autonomously inside or outside the GroEL cavity [87]. Ensemble measurements have limitations when aiming to eliminate the effect of protein aggregation on the overall folding rate [82]. A recent study [86] used FRET on a slow-folding Maltose Binding Protein variant (DM-MBP) to measure spontaneous and GroEL-GroES assisted folding rates [84]. The results suggested accelerated refolding rates by 8-fold. Using fluorescence correlation spectroscopy the authors estimated that at 100 pM DM-MBP, the observation volume contains only monomeric substrates. Consistently, a constant number of DM-MBP was detected in the observation volume, which indicated limited reversible aggregation that can confound the quantification of folding rates. We note that others have put forward arguments against active acceleration models [88]. Experiments on the

1

single ring GroEL (SR-1) variant that goes through just one round of ATP cycle without dissociating GroES [64] displayed similar folding environment for the substrate.

Theoretical models suggest that acceleration could be afforded by the confinement itself, as this can lower the entropic barrier that limits access to folded states, or an optimal range of hydrophobic interactions of the substrate GroEL cavity, which may restrict the formation of misfolded states [89–91]. The iterative annealing model has been considered for stringent substrates such as Rubisco, which are prone to form misfolded, kinetically trapped structures. GroEL-mediated unfolding could then provide another chance to refold from a high energy state [87]. The study by Lin et al. [71] found two different phases in the interaction between Rubisco monomers and ADP GroEL bullets. The first phase - before addition of ATP - displayed slowly decreasing FRET signals indicating passive unfolding by the trans ring, while ATP addition led to a rapid decrease in the FRET signal, consistent with unfolding of the monomer by GroEL. Using single-molecule FRET and rapid microfluidic mixing experiments, Hofmann et al. [92] studied the folding kinetics of the protein Rhodanese. Interestingly, the N and L regions of Rhodanese displayed similar refolding rates with or without chaperonin, while the C domain refolded two orders more slowly with the chaperonin. A possible reason for folding deceleration was postulated to be the lower diffusion constant of the polypeptide caused by interactions with the GroEL cavity surface.

Despite the large body of work on GroEL, many questions are still elusive. For instance, it is unclear whether GroEL-GroES functions typically as asymmetric complexes, with symmetric complexes favored under stress conditions. It also remains unknown how general the acceleration and stimulated annealing mechanisms are. The ability of GroEL to accelerate folding of proteins with diverse structure and folding kinetics would raise intriguing questions on how this is achieved. The specific ability of single-molecule methods to reveal individual conformational transitions will be central to resolving these important questions.

1.6. OUTLOOK

Polypeptide chains that emerge from the ribosomal tunnel are bound for a multifaceted journey guided by chaperones. The single-molecule approaches discussed here have begun to provide a glimpse of the intricate dynamics that these companions engage in. At the same time, these first results also underscore how much we do not yet know. Elementary questions are for instance whether and how chaperones such as Hsp70 directly promote folding, and how they switch to contrasting roles such as disaggregation and membrane translocation. The observed diverse modes of action also press questions on cooperation between chaperones and cochaperones. Existing models follow a rather hierachical view, with some chaperones acting upstream near the ribosome, and others downstream on mature or aggregated and damaged proteins. Observed action on near-mature proteins of supposedly upstream actors such as trigger factor challenges this logic. How, when and why different chaperones interact with a client is a crucial issue to begin addressing protein homeostasis at the cellular level. Another intriguing question is what happens early

on, at the ribosome itself. Nascent chains emerge vectorially, and hence can begin to fold before synthesis is complete. But the ribosome is also a busy platform that recruits a host of chaperones and other factors. The purpose of these actions and complex dynamics is filled with interesting open questions that are amenable to single-molecule approaches, as has been demonstrated [14, 26]. Small heat-shock proteins are a distinct class of chaperones not reviewed here but with important roles in preventing protein damage and aggregation, which can also be studied at the single-molecule level [93]. Another intriguing aspect of chaperones is their direct regulatory role. Specifically, Hsp70 and Hsp90 are involved in modulating the activity of a host of receptors and kinases, with key implications for medical conditions. Yet, how these roles are fulfilled is still obscure. Resolving these issues remains a key outstanding challenge, and single-molecule methods will be central in adressing them. At the same time, current methods are far from providing the full picture. Other rapidly advancing methods such as NMR, hydrogen exchange mass spectrometry, as well as combined fluorescence-manipulation techniques, will be crucial to arrive at the next level of models of these intriguing systems.

2

Simultaneous sensing and imaging of individual protein complexes

The vast majority of proteins form dynamic and transient complexes with DNA, RNA, small molecule ligands and other proteins. These interactions often induce protein conformational changes that are key to the function of the complex. Yet, methods to probe these dynamics are scarce. Here we advance an approach based on optical tweezers combined with fluorescence imaging to simultaneously monitor the conformational dynamics of individual proteins and the interactions with partners they form complexes with. Central to the method is a novel protein-DNA coupling strategy, which uses exonuclease digestion and partial re-synthesis to generate long DNA handles with overhangs of arbitrary length that can be efficiently ligated to oligo-labelled proteins. We demonstrate the fluorescence detection of a single yellow fluorescent protein while monitoring its folded state, and provide the first direct observation of the reciprocal dependence between the binding of a molecular chaperone (trigger factor) and the conformation of its client. We conjecture that this combined sensing and imaging approach will be an important tool to reveal the crucial conformational dynamics within larger biomolecular complexes.

2.1. INTRODUCTION

It is well known that conformational change is central to protein function and folding [94]. At the same time, ligands and protein binding partners that respond to, and affect these conformational changes, are crucial within the cellular context [94]. Indeed, in cells, proteins typically function transiently within functional complexes [95], respond to ligand binding in signaling pathways [96], and interact with the protein homeostasis machinery from synthesis to degradation [97, 98]. Yet, studying this interplay between protein interactions and conformational change is challenging. Advances in cryo-EM, NMR, and deuterium exchange mass spectrometry, are revealing protein complexes in increasing structural detail, but do not address the conformational and binding dynamics that play a central role in their function [99–101].

In the last decades, single-molecule force spectroscopy has provided key insights into diverse molecular systems and mechanisms [102]. In this approach, forces and displacements are measured on molecules tethered between trapped beads, Atomic Force Microscopy cantilevers, and surfaces. Recently, force spectroscopy has been combined with imaging techniques such as wide-field and confocal fluorescence microscopy [103], Förster Resonance Energy Transfer (FRET) [104, 105] or Stimulated Emission Depletion (STED) [106]. These approaches have so far mainly been applied to study protein-DNA interactions, with DNA strands being tethered to allow mechanical manipulation and force measurement, while interacting proteins and other molecules are detected using fluorescence imaging [107–111]. Combined sensing-imaging could be very useful when applied to protein-protein interactions as well.

However, a number of technical challenges have so far prevented these experiments on protein complexes. Specifically, it is difficult to couple proteins efficiently and strongly to long DNA handles. The latter are key to force sensing by optical tweezers, as they permit attachment to the trapped beads and limit interactions with the bead surface and laser damage [112, 113]. When combined with fluorescence imaging however, far longer tethers (over 4 kbp) are required because the parasitic fluorescence from the trapped beads readily obscures the emission signal of single molecules. Strong protein-DNA linkages are also important, as proteins are typically stabilized when complexed with other proteins or bound by ligands [41, 114]. Current coupling methods typically use cysteine chemistry to either directly attach DNA tethers [115], or first couple short DNA oligos and then hybridize longer DNA handles [59]. The former yields strong coupling but is practically limited to short tethers below 500 bp, mostly due to the electrostatic repulsion of large DNA molecules [115]. The latter has been used for tethers up to 3 kbp, but resulting constructs do not sustain large forces for long due to the susceptibility of short hybridized DNA segments to shearing forces [116, 117].

Here we present an approach for combined protein sensing and imaging using a new DNA-protein coupling method. In this modular DNA handle attachment by ligation (MODHAL) protocol, 20 nucleotide-long oligos, here called anchors, are coupled to proteins via cysteine chemistry or using an enzymatic reaction, and then linked to DNA tethers of over 5000 bp. In order to generate DNA handles with overhangs of any size, which we show is important to achieve efficient coupling, we used complete

digestion of one of the DNA strands, followed by partial re-synthesis. This strategy efficiently generates long constructs that sustain DNA overstretching forces (>60 pN) during long periods of time (>10 min). For simultaneous force sensing and imaging, we tethered the construct between beads trapped with optical tweezers, while scanning a confocal excitation beam and detecting the fluorescence emission. Using this approach, we visualize a fluorescent protein while monitoring its folding state, and demonstrate detection of chaperone binding to a tethered client.

2.2. RESULTS

First, we addressed the protein-anchor coupling, key to overall efficiency in hybridization approaches [118]. Specifically, we interrogated the effect of the anchor length. Maltose Binding Protein (MBP) with cysteines at both termini was incubated with a 4-fold excess of maleimide-modified anchors of 20, 34 and 40 nucleotides (nt) in length for one hour at 37 °C, and coupling results were analyzed by SDS-PAGE (Fig. 2.1A-B, S2.1A). For the longer 40 nt anchor, about 19% of the product corresponded to coupling of two oligos to the protein, and the rest either did not coupled or to one side only (Fig. 2.1D, S2.1A). Decreasing the length of the anchor resulted in a remarkable increase in efficiency however, with 36% dually-coupled anchor-protein-anchor for the 34 nt anchor, and 85% for the 20 nt anchor (Fig 2.1A-D, S2.1A).
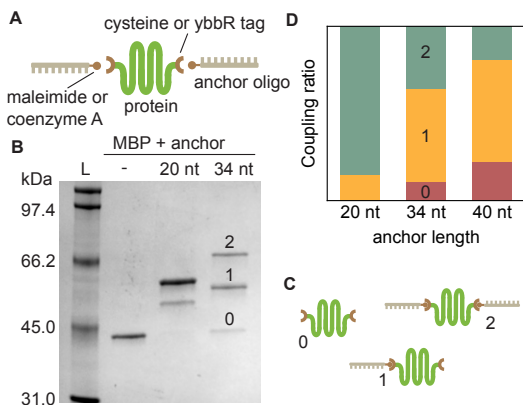

Figure 2.1: Protein-anchor coupling. **(A)**, scheme of the anchor oligo coupling to a modified protein. **(B)**, SDS-PAGE analysis of the coupling products. Lane L: protein ladder, next: cysteine-modified MBP (41.9 kDa); MBP reacted with 20 nt anchors; and 34 nt anchors. **(C)**, possible products of the coupling reaction. **(D)**, coupling ratios for different anchor lengths.

For proteins with essential cysteines, we modified the protocol. Specifically, we introduced a ybbR tag at each termini of the protein YPet. Anchors modified with coenzyme A (CoA) were then coupled to the ybbR tags using 4'-phosphopantetheinyl transferase (SFP synthase) [119]. Here we found that 27% of proteins coupled to two anchors (Fig. S2.1B). Next, we considered the anchor-handle linkage, which is central to the mechanical stability against applied forces. Previously, oligo anchors have been hybridized to a complementary overhang of the DNA handles, generated using abasic primers [59]. Hybridization yields non-covalent linkages that can limit mechanical stability against applied force, especially for shorter anchors [117]. This limitation can in principle be addressed by DNA ligation. However, the abasic site used in current methods hinders efficient ligation [120]. Restriction enzymes can be used

to generate ligation-compatible overhangs, but they are limited to 4-6 nucleotides and yield dual-handle coupling efficiencies lower than 5% even for DNA molecules that are too short for the present purpose (<400 bp) [121]. Therefore, we developed a strategy consisting of three consecutive rapid and specific enzymatic treatments to generate DNA overhangs unrestricted in length that can be covalently ligated to the coupled anchors (Fig 2.2).
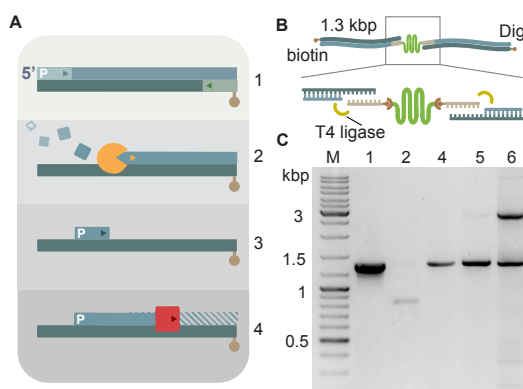

Figure 2.2: DNA handle generation and MODHAL coupling. (A) strategy for the overhang generation. 1) Initial PCR amplification of template DNA using phosphorylated and functionalized —brown circle— primers. 2) λ exonuclease (orange) digestion of the phosphorylated strand. 3) Tuning of the overhang length by selection of the appropriate primer. 4) Partial strand re-synthesis using Deep Vent (exo-) (magenta) that leaves the overhang intact (not abasic) for ligation. **(B)** MODHAL coupling scheme. **(C)** agarose gel electrophoresis analysis of MODHAL coupling. Lane M: DNA ladder, lane 1: initial 1333 bp template, lane 2: λ exonuclease digestion (dim signal because of ssDNA), lane 4: Deep Vent (exo-) re-synthesis, lane 5: ligation of overhang-DNA only (no anchor-MBP-anchor), lane 6: ligation of overhang-DNA with anchor-MBP-anchor hybrids.

First, a 1333 bp long DNA template was generated by PCR using a phosphorylated forward primer and a functionalized reverse primer for attachment to the bead or surface (Fig 2.2A,1 and Fig 2.2C, lane 1). Here we chose biotin or digoxigenin, but another link could be used [122, 123]. The product was digested with λ exonuclease (Fig 2.2A2) and the remaining functionalized ssDNA strand was then used for a partial re-synthesis, where the primer sequence is complementary to an inner segment of the strand, starting where the anchor-complementary sequence finishes (Fig 2.2A, 3). To preserve the overhang, we used Deep Vent (exo-) polymerase [124] (Fig 2.2C, lane 4). The overhang length can be easily tuned with this approach. More importantly, the generated overhang allows covalent DNA ligation and permits to use shorter anchors without limiting the resistance of the tethers.

The anchor-protein-anchor construct was ligated to the 1333 bp-long tethers (ratio 1:1:1) with T4 ligase (Fig 2.2B). Agarose gel electrophoresis showed that 45% of the handles were ligated together into a complex twice the size (Fig. 2.2C, lane 6). Consistently, in the absence of anchor-protein-anchor, almost none of the handles were

ligated (1%, Fig. 2.2C, lane 5).

To test mechanical stability, we linked the construct between functionalized polystyrene beads within the optical tweezers instrument (Fig 2.3A). Resulting force extension curves for MBP (Fig 2.3B) were fitted using a worm-like-chain (WLC) model, and showed the characteristic unfolding pattern in two steps [21]. We tested tether strength by recording the maximum tensions they reached without breaking when ramping up the applied force (Fig S2.2A). We considered traces that showed the characteristic unfolding behavior in the first pull.
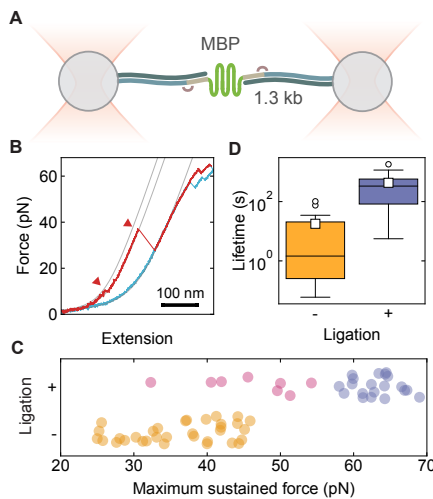

Figure 2.3: Mechanical tests with optical tweezers. (A) MBP tethered with DNA between two beads trapped with optical tweezers. **(B)** Force extension curve of MBP with 1300 kb DNA handles showing the characteristic two step unfolding pattern (red triangles) and the DNA overstretching regime above 60 pN (red curve: pulling, blue curve: relaxing, gray lines: WLC curves). **(C)** Distribution of maximum reached force for non-ligated (yellow) and ligated tethers. Points are scattered for visualization. Magenta: broken ligated tethers. Purple: unbroken ligated tethers. **(D)** Tether lifetime at 30 pN without and with ligation. Time scale is logarithmic.

Tethers that were generated by hybridization only, without ligation, were found to break below 47 pN (Fig 2.3C, $N = 33$), close to the predicted shearing force for our anchors (45 pN) [117]. In contrast, the majority of ligated tethers (71%, $N = 28$) could be pulled up to the DNA overstretching regime –above 60 pN– without rupturing for multiple cycles, thus demonstrating the improved mechanical stability (Fig 2.3C). We also measured tether lifetimes at 30 pN, well below the expected shearing force (Fig S2.2B). Ligation yielded a remarkable lifetime improvement of two orders of magnitude, to over 100 s ($N_+ = 21$, $N_- = 15$, Fig 2.3D). These data underscored the poor mechanical stability provided by short hybridized anchors, and the utility of the exonuclease approach to overcome these limitations and enable strong and efficient ligation.

Finally, we performed two proof of principle experiments. First, we related the activity of a single protein to its folding state, by monitoring the fluorescence emission of a single YPet that is mechanically manipulated (Fig. 2.4A). Confocal imaging showed significant parasitic autofluorescent signals emanating from the beads, several 100 nm beyond the bead surfaces, thus obscuring any relevant signal from YPet (Fig 2.4B). To overcome this issue, we generated even longer handles of 5 kbp each, and found that they could be ligated efficiently with our protocol (35%, Fig. S2.3). The key region between the beads now showed a minimal background photon count,
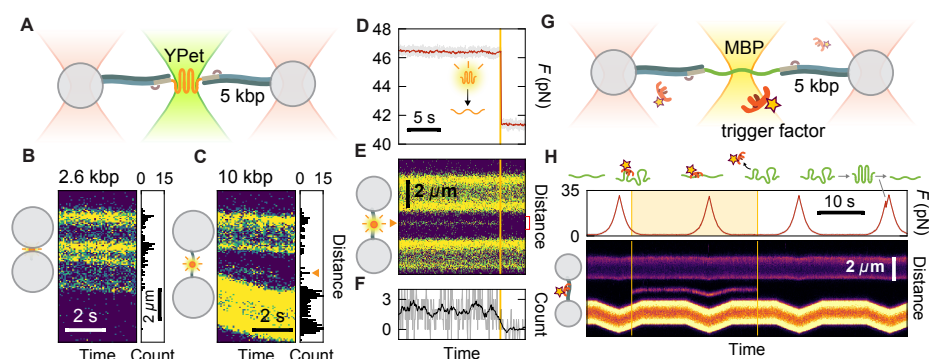
Figure 2.4: Dual monitoring of protein conformation and binding. (A) scheme of tethered YPet protein in optical tweezers with an additional 638 nm laser. **(B)** and **(C)**, confocal fluorescence kymographs of YPet using (b) 1.3 kbp and (c) 5 kbp handles, with a typical scanning line profile on the right. Parasitic fourescence of the beads prohibits detection of protein emission when using 1.3 kb handles (b), while 5kb tethers overcome this limitation. **(D)** force monitoring of YPet. Vertical yellow line indicates the unfolding event. **(E)** confocal scanning kymograph showing active Ypet between the beads. **(F)** photon count from YPet location (region indicated in red in g), after background subtraction. **(G)** scheme of unfolded MBP in optical tweezers with an additional 532 nm laser used for confocal scanning, and fluoresecently labeled Trigger Factor. **(H)** force monitoring and complex formation imaging for MBP-Trigger Factor. Trigger factor binds to MBP after unfolding, suppresses refolding, and remains bound during stretching to 35 pN.

indicating a lack of bead parasitic signals (Fig 4c). After establishing a single tether, a fluorescent spot was observed between the beads, indicating the presence of folded and active YPet, similar to previously reported for GFP [116] (Fig. 2.4E, F). The protein remained folded even in stretching and relaxation cycles up to 65 pN. However, our ligated tethers could be maintained at high forces (>45 pN) for tens of seconds, which resulted in YPet unfolding. Consistently, we observed a simultaneous drop in force and loss of fluorescence (Fig. 2.4D-F)

Next, we studied the dynamics of a protein-chaperone complex. We added fluorescently labelled trigger factor, a key *E. coli* chaperone, to tethered MBP (Fig. 2.4G). The parasitic signal from the beads was now even stronger due to unspecific trigger factor binding to the surfaces, but the 5 kbp handles alleviated the issue (Fig S4). The binding of single trigger factor chaperones could now be directly visualized (Fig 4h), occurred after MBP unfolding, and was here found to stabilize the unfolded state for over 10 s. This timescale is well below the lifetime of the fluorescent dye (Fig. S2.5). Thus, trigger factor binding depended on the client folded state and, conversely, substrate folding transitions depended on the binding of trigger factor. The nature of the trigger factor interaction with unfolded substrates remains incompletely understood [98], and has been suggested to involve multiple low-affinity contacts [19]. The present approach reveals that trigger factor can remain bound to both relaxed and stretched substrate chains, where the number of contacts is reduced. This direct ob-

servation of long-term binding also explains previously reported suppression of re-folding by trigger factor [21].

2.3. DISCUSSION

In summary, here we have presented a novel DNA-protein tethering strategy that enables combined protein sensing and imaging, and have demonstrated its potential by key proof-of-principle experiments. We envision that our strategy will provide important insights into the synergy between complex formation and protein conformation, as seen in molecular chaperones [98], intrinsically disordered protein networks [125], and DNA- and RNA- binding proteins including the CRISPR-Cas9 complex [126], tumor repressors [127] and steroid receptors [128], among many other systems.

2.4. MATERIALS AND METHODS

2.4.1. PROTEIN EXPRESSION AND PURIFICATION

MBP was modified with cysteine residues at both termini using the pET28 vector. The linker sequences are GRGS and RITK for N- and C-terminus respectively. YPet (a more stable and brighter variant of Yellow Fluorescent Protein) was fused to MBP to improve solubility and to enable affinity chromatography, and two YbbR tags (DSLEFI-ASKLA) were included at each termini. Proteins were purified from *E. coli* BL21(DE3) cells. For overexpression, overnight cultures were diluted 1:100 in fresh LB medium supplemented with 50 mg/l kanamycin, 0.2% glucose and incubated under vigorous shaking at 30 °C. Expression was induced at OD600=0.6 by addition of 1 mM IPTG and incubation overnight at room temperature. Cells were cooled, harvested by centrifugation at 5000g during 20 minutes, flash-frozen and stored at -80 °C. Cell pellets were resuspended in ice-cold buffer A (50 mM potassium phosphate pH 7.5, 0.15 M NaCl, 3 mM Chloramphenicol, 50 mM Glu-Arg, 10 mM Complete Protease Inhibitor Ultra from Roche, 10 mM EDTA) and lysed using a pressure homogenizer. The lysate was cleared from cell debris by centrifugation at 50000g for 60 min and incubated with Amylose resin (New England Biolabs) previously equilibrated in buffer A for 20 minutes at 4 °C. The resin was washed with buffer A three times by centrifugation and bound proteins were eluted in buffer A supplemented with 20 mM maltose. Purified proteins were aliquoted, flash frozen in liquid nitrogen and stored at -80 °C.

2.4.2. DETAILED PROTEIN-ANCHOR COUPLING PROTOCOL

Anchor oligos 5'-modified with maleimide or coenzyme A (biomers.net) were diluted in coupling buffer (Sodium Phosphate 100 mM pH 7.2, NaCl 150 mM, EDTA 10 mM) to a concentration of 300 μM or 500 μM, respectively. Purified proteins were thawed to room temperature and passed through a desalting column (PD-10, GE Healthcare) to get rid of reducing agents and elutants. If concentrations were below the 100 μM range, they were concentrated using an appropriate size Amicon centrifugal filter. Immediately after, they were set to the coupling reaction. For the cysteine chemistry coupling, the protein was mixed with the anchor oligos in a 1:4 ratio and incubated for 1 hour at RT or overnight at 4 °C. Addition of TCEP in the middle of the incu-

bation increased the coupling yield. For the SFP-mediated reaction, around 6 μM YbbR-modified YPet was incubated with 8 μM SFP and 25 CoA-modified oligos, 50 mM Hepes and 10 mM MgCl2, in a total volume of 20 μL. Excess anchor oligos were removed by affinity chromatography using amylose resin.

2.4.3. DNA OVERHANG GENERATION

Initial DNA templates were generated by PCR from 3 ng commercial pUC19 plasmid (ThermoFisher) or from pOSIP-TT (for 5kb tethers) using Phire Green Hot Start II polymerase Master Mix(ThermoFisher). The forward primer was phosphorylated at the 5' end and its sequence was
5'-CAGGGCTCTCTAGATTGACTTATGTATCCGCTCATGAGACAATAA-3', where underlined bases correspond to the annealing segment (and therefore to the internal primer for the subsequent fill-up PCR) and the rest constitutes the final overhang. The reverse primers were functionalized at the 5' end with biotin or digoxigenin in order to have asymmetric constructs. The amplification protocol is shown in Table 2.1.

Table 2.1: Polymerase chain reaction protocol for the initial amplification of DNA

Step	Temperature (°C)	Duration	
Denaturation	98	30 s	
Denaturation	98	5 s	35 cycles
Annealing and extension	72	18 s	
Extension	72	1 min	

Products were cleaned using a QIAquick PCR Purification Kit (Qiagen) and set to λ exonuclease (New England Biolabs) digestion for 2 hours at 37 °C, using 2 units of enzyme per μg of DNA. A heat treatment at 80 °C was then applied during 10 minutes to inactivate the exonuclease. The product was purified using 30 kDa Amicon centrifugal filters (Merck Millipore) and checked using agarose gel electrophoresis. If the digestion was successful, a linear PCR was performed on the ssDNA using 6 units of Deep Vent (exo-) from New England Labs per μg of ssDNA, and following the protocol showed in Table 2.2. The PCR product was purified and concentrated to approximately 500 nM using a 50 kDa Amicon filter.

Table 2.2: Polymerase chain reaction protocol for the partial fill up of the λ exonuclease treated DNA strand with Deep Vent (exo-) polymerase

Step	Temperature (°C)	Duration	
Denaturation	95	5 min	
Denaturation	95	30 s	
Annealing	53	20 s	10 cycles
Extension	74	2.5 min	
Extension	74	5 min	

2.4.4. OPTICAL TWEEZERS ASSAY

Carboxyl polystyrene beads (CP-20-10, diameter 2.1 μM, Spherotech) were covalently coated with sheep anti-digoxigenin antibody (Roche) via carbodiimide reaction (PolyLink Protein coupling kit, Polysciences Inc.). Around 50 ng of the generated construct were incubated with 2 μL beads in 10 μL HMK buffer (50 mM HEPES, ph 7.5, 5mM MgCl$_2$, 100 mM KCl) for 15 minutes in a rotary mixer at 4 $^\circ$C and rediluted in 350 μL HMK buffer. With our coupling strategy, approximately 50% of the constructs will be asymmetrically functionalized with digoxigenin and biotin in each side. In order to create the second connection, we employed Neutravidin coated polystyrene beads (NVP-20-5, diameter 2.1 μM, Spherotech). Once trapped, beads were brought into close proximity to allow binding, and tether formation was identified by an increase in force when the beads were moved apart. To mitigate photobleaching and tether damage we added an oxygen scavenging system (3 units/mL pyranose oxidase, 90 units/mL catalase and 50 mM glucose, all purchased from Sigma-Aldrich).

2.4.5. FORCE SPECTROSCOPY DATA ANALYSIS

Data was recorded at 500 Hz using a custom-built dual trap optical tweezers for the tether resistance assays and a C-Trap (Lumicks) for the dual monitoring experiments. Data was analyzed using custom scripts in Python. Optical traps were calibrated using the power spectrum of the Brownian motion of the trapped beads [129], obtaining average stiffness values of $\kappa = 0.39 \pm 0.04$ pN/nm. Force-extension curves were fitted to two worm-like-chain (WLC) models in series (Fig S1a), using the approximation of an extensible polymer reported by Petrosyan [130] for the DNA, and the Odijk inextensible approximation for the protein contribution [131].

$$x = L_c^* \left(\frac{4}{3}\left(1 - \frac{1}{\sqrt{\beta^*+1}}\right) - \left(\frac{10\exp\left(\sqrt[4]{\frac{900}{\beta^*}}\right)}{\sqrt{\beta^*}\left(\exp\left(\sqrt[4]{\frac{900}{\beta^*}}\right)-1\right)^2} + \frac{\beta^{*1.62}}{3.55+3.8\beta^{*2.2}}\right) + L_e\left(1 - \frac{1}{2}\sqrt{\beta}\right)\right)$$

Where $\beta^* = (FL_p^*)/(k_B T)$, being F the force, T the temperature, L_p^*, K and L_c^* the persistence length, stretch modulus and contour length of DNA, respectively. For the protein, $\beta = (FL_p)/(k_B T)$ with L_p, L_e being the persistence and extended length of the polypeptiden. L_c^* was 906 or 3500 nm for the two different DNA handles used (1.3 and 5 kb, respectively), L_c was 120 nm for the MBP and 105 nm for YPet, and L_p was fixed to 0.75 nm. L_p^* and K were fitted, yielding average values of 30 nm and 700 pN/nm respectively. Tether resistance was tested by slowly ramping up the tension on the tether and recording the rupture force. It is well known that the measured rupture force increases for higher pulling rates [132]. The pulling speed here used was 100 nm/s, too slow to bias the rupture force in any substantial way. If the DNA overstretching regime was reached, the tether was relaxed back. The rupturing force includes traces of the first pulls that showed proper MBP unfolding only (Figure S2.2A). For the lifetime experiments, the force was increased gradually to around 30 pN. The time between reaching this force and the rupture of the tether was recorded as the lifteme (Fig S2.2B).

2.4.6. Fluorescence imaging analysis

For dual monitoring experiments, an excitation laser beam (with wavelengths of either 532 nm or 647 nm) was scanned along the beads and tether at a line rate of 12 Hz. The excitation power was 1.3 mW for YPet experiments and 0.3 mW for trigger factor binding experiments. Force-spectroscopy and confocal microscopy data were synchronized based on the movement of the beads. The edge of the moving bead was tracked using a Gaussian fit and overlaid on top of the actual movement set in the mirror by minimizing the difference between the signals (Fig S2.6B). This same movement was used to trace a region of the scanning between the beads including the protein (Fig 2.4E, red lines). Signal was calculated by adding the intensity of all pixels in that region and subtracting the background, calculated similarly by summing the intensity in a region of the same size outside of the beads. We tested the emission lifetime of Atto647N under our experimental conditions using a labeled DNA construct (Fig. S2.5A). Photobleaching was not observed in any of the confocal scanning experiments, which terminated upon tether rupture ($t_r = 660 \pm 150$ s, $N = 6$, S2.5B), This timescale, which set a lower limit for the dye lifetime, is much longer than the tens of seconds observed for trigger factor binding.
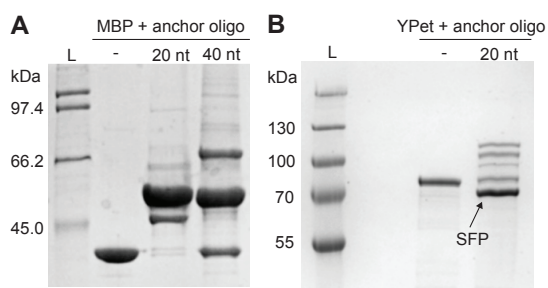
2.5. Supplementary figures

Figure S2.1: Additional anchor coupling to proteins. (A) Additional SDS-PAGE of MBP-anchor coupling for 20 and 40 nt oligos. Coupling yield quantification is shown in Figure 2.1D. (**B**) SDS-PAGE showing YPet (lane 1) coupling to a 20 nt anchor modified with coenzyme A (lane 2). SFP enzyme band is indicated with an arrow and lies below uncoupled YPet. Double anchor coupling yield is 27%. There are two bands corresponding to single-coupling due to the position of the tags, one at the C-terminus of YPet and the other one between MBP and YPet, giving rise to different migration behaviours.
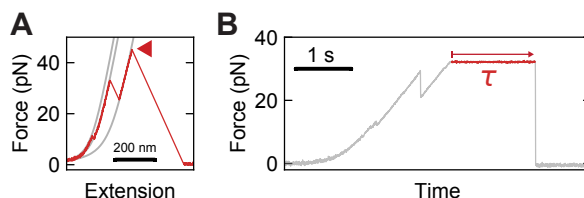
Figure S2.2: Tether resistance quantifications. **(A)** characterization of the maximum sustained force for a breaking tether (red triangle). Gray lines indicate WLC fits. **(B)** example of tether lifetime (τ) measurement at 30 pN. For both quantifications, only tethers that display the typical MBP unfolding pattern are considered.
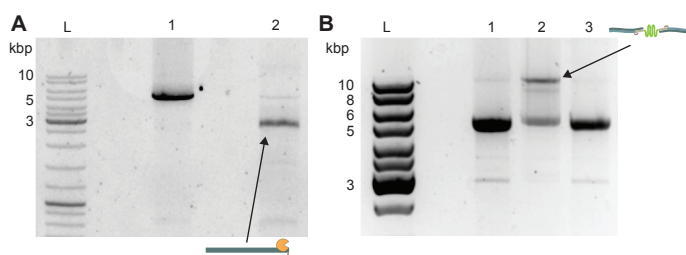

Figure S2.3: Overhang attachment with 5 kbp DNA handles. (A) First amplification of the 5000 bp template (lane 1) and exonuclease digestion (lane 2). **(B)** Lane 1 shows the final handle with the overhang after Deep Vent (exo-) PCR. In lane 2 the ligation of the handles with anchor-MBP-anchor shows a clear band at 10 kbp, indicating that the coupling is successful (35%). Lane 3 shows ligation in the absence of anchored protein, with no additional band at 10 kpb indicating that unspecific ligation is negligible.
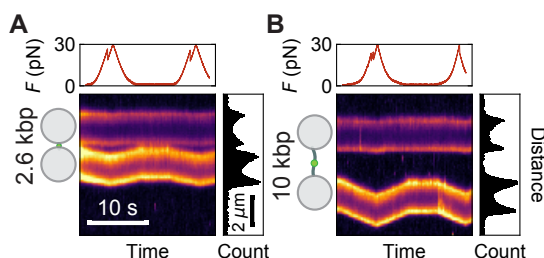

Figure S2.4: Confocal fluorescence kymographs of MBP in the presence of Trigger Factor using (A) 1.3 kbp and **(B)** 5 kbp handles, with the monitored force on top and a typical scanning line profile on the right. Noise from the beads hinders detection of fluorescence coming from single binding molecules when using 1.3 kb handles **(A)**, while 5kb tethers solve this limitation **(B)**.
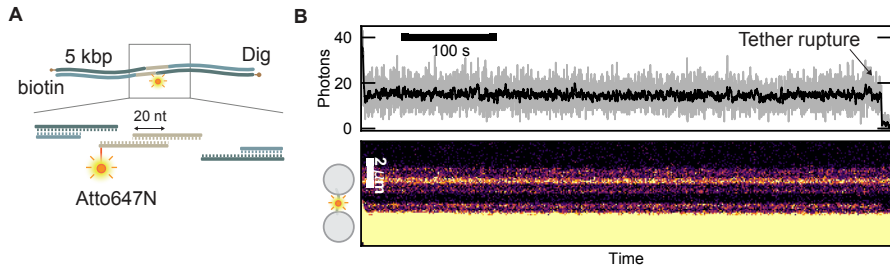
2

Figure S2.5: Photobleaching control for Atto647N (A) Scheme of the construct used for characterization. Two primers (biomers.net) present an overlapping sequence, which after annealing present identical overhangs that can be hybridized to the longer handles. One of the primers is modified with Atto647N. **(V)** Example kymograph of the fluorescence detection for the DNA construct. The bottom bead is coated with the DNA construct, and therefore its parasitic emission is higher. As a consequence, the position of the dye appears to be closer to the bead (red marker). In all experiments ($N = 6$) the (non-ligated) tether broke before the dye bleached, with an average tether lifetime of 660 ± 150 s.
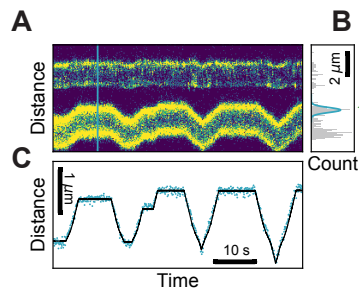

Figure S2.6: Force spectroscopy and fluorescence signals synchronization. (A) confocal scanning kymograph of the beads with YPet tethered in between. **(V)** profile of a scanning line (blue line in A), with a Gaussian fit of the edge of the moving bead in blue. **c** Synchronization of fluorescence (blue dots, position of bead calculated for each scanning line as shown in (V) and high-resolution position detector (black line, representing the actual movement of the bead) signals.

3

GroEL-mediated acceleration of protein folding by enhanced collapse

Most proteins appear to populate collapsed states during productive folding. However, misfolding and aggregation also proceed from collapsed conformations. How collapsed polypeptide states are impacted by molecular chaperones like GroEL-ES in order to promote folding is obscure. Here, we integrate single protein manipulation and imaging to probe substrate-chaperone complexes in real-time. GroEL is found to exert attractive forces that pull polypeptides into its cavity, and drive a collapse that triggers folding, even in slowly folding proteins. This collapse enhancement is strongest in the nucleotide-bound states of GroEL and is aided by its C-terminal tails and GroES. These results show that GroEL-ES employs an active folding mechanism, that collapse strength is a folding determinant directly targeted by GroEL-ES, and suggest that manipulation of protein collapse could be a general strategy used by other chaperones in cellular protein quality control.

3.1. INTRODUCTION

Even when not in their native conformation, most polypeptides spend the vast majority of their time in collapsed states, engaging in complex configurational dynamics that ultimately produce highly intricate protein structures [133]. However, this process can be subject to both perilous delays and toxic, aggregating side reactions [134]. It remains a matter of some controversy whether or not unfolded polypeptide conformations can be productively altered to accelerate folding and, if so, by what mechanism [70, 98, 135]. The archetypal chaperonin GroEL-ES has been proposed to enable such active folding by sterically confining polypeptides within its closed chamber, which can lower their entropy and thus effectively reduce folding barriers [62, 136], or by partially unfolding kinetically trapped, misfolded conformations [62, 86, 136]. However, diverse experiments suggest that GroEL-ES acts only as a passive Anfinsen cage that simply prevents protein aggregation [82, 137]. Testing these models is challenging, because the distinguishing driving forces and conformational effects exerted on substrate proteins by GroEL-ES are difficult to measure. Moreover, accelerated folding can be confounded by the impact of aggregation and population heterogeneity [84, 88, 138]. Elucidating these issues is key to developing a better understanding of the dynamics and stability of non-native protein states, how they govern folding delays and aggregation, whether and how these issues limit the scope of protein complexity, and to the mechanistic basis of stimulated protein folding by molecular chaperones.

3.2. RESULTS

We first aimed to test whether GroEL-ES mediates folding acceleration by following individual proteins in time (Fig. 3.1A-C). We tethered Maltose Binding Protein (MBP) between beads trapped with optical tweezers. After unfolding the protein, we relaxed it waiting at 0 pN for 5 s to allow refolding, and stretching to assess the result, first in the absence of chaperone (Fig. 3.1A-C). During relaxation, the measured forces and distances were consistent with the expected worm-like chain (WLC) curve for unfolded MBP, which has a contour length $L_c = 120$ nm (Fig. 3.1C, blue trace 1). As reported previously [21, 28, 41], subsequent stretching either showed the chain had remained unfolded (Fig. 3.1C, red trace 2b), or the data followed the WLC curve of the folded core MBP structure at $L_c = 30 \pm 10$ nm, with the detachment of C-terminal α-helices from the core sometimes visible at low force (Fig. 3.1C, red trace 2a). We quantified the fraction P_c of relax-stretch cycles that indicated such core refolding. P_c increased from 0.7 without chaperones, to 0.85 when GroEL, GroES and ATP were present (Fig. 3.1D, Extended Data Fig. 1, $p < 0.05$). The GroEL-ES-refolded cores displayed an unfolding force F_U that was similar to autonomously refolded cores (Fig. 3.1E)20. A reported slow-folding MBP mutant (dmMBP) [84] scarcely refolded when measured alone ($P_c = 0.06$). However, it refolded readily with GroEL-ES and ATP ($P_c = 0.9$, Fig. 3.1D, Fig. S3.1B). To investigate a different folding-rate limitation, we did not relax MBP chains fully, but kept them at 2 pN for 30 s before stretching again (Fig 3.1D and E). P_c was now 0.5, and increased to 0.9 with GroEL-ES and ATP (Fig. 3.1D, $p < 0.05$). These data demonstrated folding acceleration of single substrates in
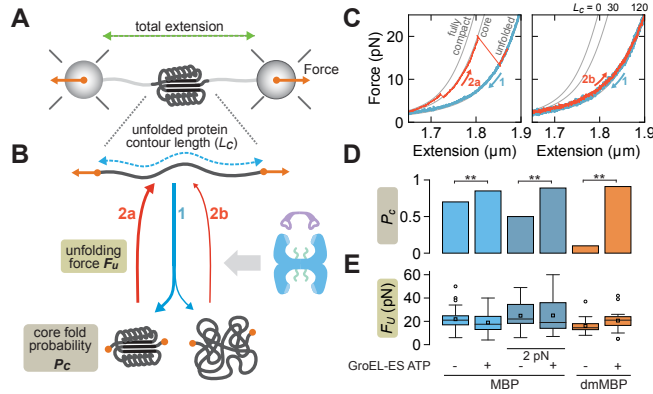
Figure 3.1: GroEL-ES accelerates folding in absence of aggregation. (A) Diagram of the optical tweezers experiments. **(B)** Diagram of relax-stretch cycles to quantify MBP core refolding. Unfolded chains (top) are relaxed, kept at 0 pN for 5 s (bottom, or alternatively at 2 pN for 30 s), and stretched to assess the new state. **(C)** Example force-extension data (without GroEL-ES). After relaxation of unfolded MBP (blue traces, 1), stretching data (left panel, red trace 2a) follows the WLC curve of the MBP core state (middle gray line). Up to about 5 pN, the data follows the WLC curve of fully folded MBP. Alternatively, the stretching follows the WLC curve of the unfolded state (right panel, red trace 2b). **(D)** Fraction of cycles showing core state refolding (P_c). Conditions: with and without 200 nM GroEL, 500 nM GroES, 1 mM ATP, for MBP at 0 pN ($N = 65$ and $N = 53$), for MBP at 2 pN ($N = 18$ and $N = 21$), and dmMBP at 0 pN ($N = 19$ and $N = 21$). Double stars: significant difference ($p < 0.05$). **(E)** Unfolding force F_U determined during stretching. For dmMBP alone, F_U of the first stretching curve is displayed because of the low refolding rate.

absence of aggregation, which also indicated that the DNA handles did not inhibit folding stimulation by GroEL-ES. The stimulation mechanism remained unresolved, however.

Because of the complex dynamics of the GroEL-ES cycle, we focused on GroEL in different nucleotide-bound states, without GroES. We now observed sudden switching to an unfolded state that persisted over multiple relax-stretch cycles until the tether broke (Fig. S3.2). This switching occurred frequently for the APO state, and less so for the ATP state, as mimicked by the slowly hydrolyzing GroEL398A [66], and the ADP state (50, 30, and 20% of tethers, respectively). These findings agree with the known stable binding of unfolded substrates to the apical domains at the rim of the GroEL cavity [77]. However, the polypeptides also displayed gradual decreases in contour length L_c during relaxation, down to dimensions of folded states (Fig. 3.2A and S3.3A). The chains were thus gradually compacted, as L_c measures the extended (non-compact) part of the polypeptide (Fig. 3.2B). Compaction also occurred for MBP alone, but was significantly stronger with GroEL, in particular in the ATP and ADP states - as quantified by the compaction energy E_c for the full relaxation process (Fig. 3.2B-C and S3.4). In principle, such a compaction could be the result of one protein substrate chain binding to many GroEL apical domain sites. However, a number
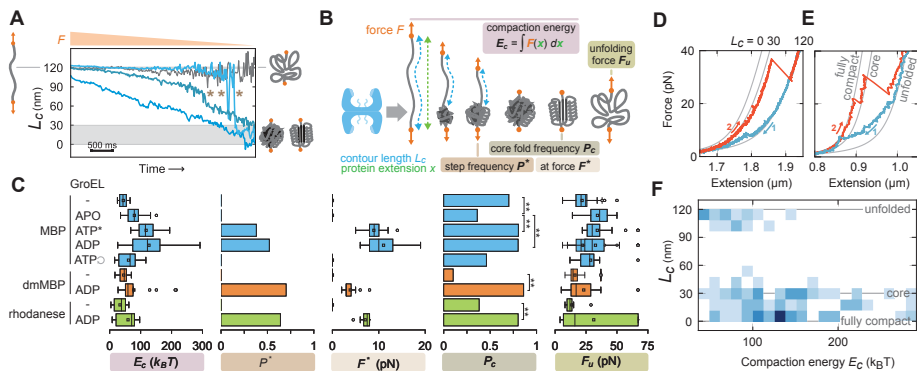
Figure 3.2: An open GroEL cavity can enhance protein chain collapse and folding. (A), Unfolded chain contour length (L_c) during relaxation, as determined from force-extension data (see blue curves panel D and E) and the WLC model. Measurement noise does not allow L_c determination below about 2 pN. Blue curves: dmMBP with 200 nM GroEL and 1mM ADP. Stars: sudden L_c steps indicating (partial) folding. Gray curve: data illustrating no detectable compaction in the accessible force range, with L_c remaining constant as the chain coils (data without GroEL). **(B)** Cartoon of a relax-stretch cycle and measured quantities. **(C)** Total compaction energy during relaxation (E_c), fraction of relaxation curves with steps (P^*), at force (F_*), fraction of cycles showing core refolding after relaxation and waiting at 0 pN for 5 s (P_c), unfolding force (or maximally sustained force when structure does not unfold) (F_U). Double stars: significant difference ($p < 0.05$). Conditions: MBP, no chaperones (-, $N = 65$), 200 nM GroEL (APO, $N = 33$), 200 nM GroEL398 and 1 mM ATP (ATP^*, $N = 43$), 200 nM GroEL and 1 mM ADP (ADP, $N = 96$), 200 nM GroEL and 1 mM ATP (ATP-cyclic arrow, $N = 21$), dmMBP no chaperone (-, $N = 19$), GroEL-ADP (ADP, $N = 21$), rhodanese no chaperones (-, $N = 16$), GroEL-ADP (ADP, N=20). **(D-E)** Example force-extension data of relax (blue) and stretch (red) curves for MBP with GroEL and ADP. **(F)** L_c of refolded states versus E_c of the previous relaxation (blue traces panels d and e). L_c is defined by the initial stretching data, and the WL_c curve is follows: $L_c = 30$ (panel D), $L_c = 0$ (panel E). Data of MBP with GroEL-ADP and with GroEL398-ATP was combined. The distribution indicated that strong compaction almost always results in a folded state between core and fully compacted. All curves in this latter category showed unfolding via the core state (as e.g. in panel E).

of features in the data indicated a different explanation.

First, the gradual compaction was accompanied by sudden step-wise compaction events (Fig. 3.2A, stars). These steps suggested folding transitions rather than stable binding: they were large in size (up to nearly the total chain length, Fig. S3.3B), occurred at high forces (up to 19 pN, Fig. 3.2A-C), and exhibited reversible "hopping" [139–141] (Fig. 3.2A and S3.3C). Second, for the ATP and ADP states, when the stabilization of unfolded states was weakest, compaction was strongest - as quantified by E_c and the fraction P^* of traces showing steps (Fig. 3.2B-C). Third, after relaxation and waiting, the polypeptides often had folded to the MBP core state, with subsequent stretching data following the core WLC curve (Fig. 3.2D to F), as seen before for spontaneous and GroEL-ES-assisted folding. Notably, the fraction of cycles showing refolded cores was particularly high for the ATP and ADP states ($P_c = 0.8$ for both,
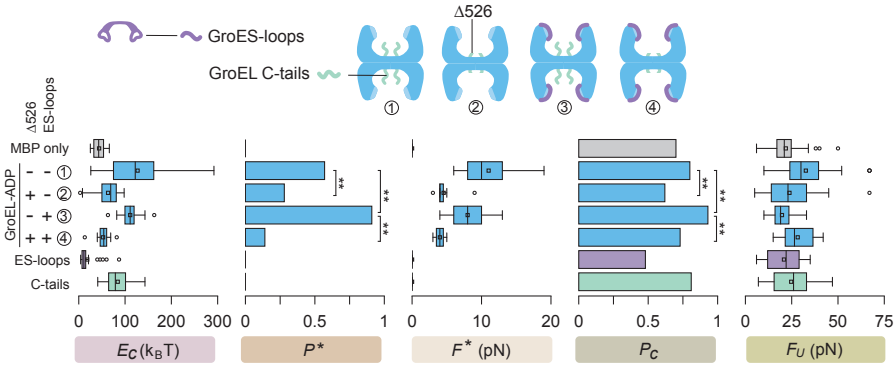
Figure 3.3: Roles of GroEL apical domains and cavity. The unstructured loops of GroES, added as separate polypeptides (purple), can bind the apical domains (light blue) of GroEL and reduce substrate affinity. The effect of the unstructured C-terminal tails (light green) of GroEL is tested by truncation (GroELΔ526). Quantification of (see Fig. 3.2B): total compaction energy during relaxation (E_c), fraction of relaxation curves with steps (P^*) at force (F_*), core refold probability after 5 s at 0 pN (P_c), unfolding force or maximally sustained force (F_U), as determined from MBP relax-stretch cycles. Double stars: significant difference ($p < 0.05$). Conditions are, from top to bottom: no chaperone ($N = 65$), 200 nM GroEL and 1 mM ADP (1, $N = 96$), 200 nM GroELΔ526 and 1 mM ADP (2, $N = 39$) 200 nM GroEL, 1 mM ADP and 1 µM GroES loops (3, $N = 55$), 200 nM GroELΔ526, 1 mM ADP and 1 µM loops (4, $N = 11$), 1 µM GroES-loops only ($N = 29$), and 1 µM C-tails only ($N = 32$).

Fig. 3.2C). Moreover, P_c even increased beyond 0.95 when the previous compaction was strong ($E_c > 100$ $k_B T$, Fig. 3.2F, S3.5). These findings showed that compaction was distinct from the stabilization of unfolded states, and played a role in stimulating folding instead.

GroEL thus displayed two different modes of interaction. In the first, unfolded substrates were bound, immobilized, and stabilized. In the second, they were compacted by attractive forces while preserving the necessary mobility to form tertiary structure. These compaction forces, which oppose and equal the applied forces (Fig. 3.2b, orange arrows), reached well above 10 pN (Fig. 3.2c, e). Such a collapse process, in which the protein chain becomes more compact and can form some secondary and tertiary structure, is considered key in autonomous folding [41, 142, 143]. Here, MBP alone also collapsed with low E_c and $P^* = 0$, and displayed increased E_c, P^*, and P_c when interacting with GroEL398-ATP and GroEL-ADP (Fig. 3.2C). GroEL mediated an enhanced collapse, with increased E_c, P^*, and P_c (Fig. 3.2C), as was also observed for two other substrates, namely dmMBP and rhodanese (Fig. 3.2C and S3.6).

Next, we tried to disentangle the contributions of the GroEL apical domains to unfolded chain stabilization and folding. We performed MBP relax-wait-stretch experiments in the presence of GroEL-ADP, which yielded the highest P^*, and small peptides equivalent to the unstructured GroES loops, which compete strongly for the

apical domain substrate binding sites [144]. The sudden stabilization of unfolded chains (Fig. S3.2) was no longer observed, consistent with reduced apical domain binding. Interestingly, P_c and P^* were now higher even than for GroEL-ADP only, both for MBP and dmMBP (Fig. 3.3 and S3.7). These increases were not caused by the GroES-loops directly, since they alone did not yield increases (Fig. 3.3). These data are consistent with the apical domains antagonizing folding.

The apical domains also exhibited another effect: the observed unfolding force F_U of MBP, which had increased with GroEL-ADP, now decreased back to MBP-only levels when the GroES loops were present (22, 33, 23 pN respectively, Fig. S3.2 and S3.8). dmMBP showed a similar trend (Fig. S3.7). GroEL thus can stabilize (partially) refolded states against forced unfolding, in addition to stabilizing unfolded states. Binding of the GroES-loops suppressed both effects, even as compaction and folding remained stimulated (Fig. 3.3). To test whether the GroEL interior played a role in the folding stimulation, we truncated the unstructured C-terminal tails at the cavity bottom (GroELΔ526) [145]. P_c and P^* were indeed more than two-fold lower for GroELΔ526-ADP than for GroEL-ADP, both with and without the GroES-loops present (Fig. 3.3 and S3.7). Interestingly, even alone the C-tails could promote some compaction (Fig. 3.3). Overall, the data showed that GroEL-mediated collapse and folding depended on the C-tails in the GroEL cavity. The enhancement of chain collapse while maintaining folding dynamics upon interaction with GroEL is notable, and suggests a balance between different cavity properties [84, 145].

Finally, we sought to verify two key interactions of the substrate-chaperone complex in these experiments, which required different approaches. To directly visualize GroEL binding, we scanned a fluorescence excitation beam along the tethered MBP during relax-wait-stretch cycles (Fig. 3.4A). ADP and Atto532-labeled GroEL were present, at reduced concentrations to limit background fluorescence. The appearance of a fluorescent spot between the beads indicated binding of a single GroEL tetradecamer (Fig. 3.4B). Consistently, during relaxation, such GroEL binding events always occurred first, and folding steps afterwards (Fig. 3.4C, D), thus confirming stimulated folding transitions in substrates complexed with GroEL.

Second, we used a buffer-exchange protocol to verify that ternary complexes (GroEL-GroES-MBP) can indeed be formed in the optical tweezers assay. Unfolded MBP was first complexed to a single-ring GroEL variant (SR1) with ATP [146]. SR1 binding was evidenced by unfolded-state stabilization of MBP (Fig. 3.4E to F). GroES is known to bind SR1 very tightly, thus trapping it in the ADP-bound state, and displacing substrates from the apical domains into the SR1-GroES cavity [70, 135]. The relaxed MBP-SR1 complex was therefore exposed to GroES and ATP, and unbound SR1 was washed away. This exposure triggered MBP refolding to the core state, which in subsequent stretching did not unfold fully (Fig. 3.4G). These data strongly suggested a SR1-ES-MBP complex: if not stabilized by SR1, MBP core states unfold below 40 pN (Fig. 3.1C), and if only SR1 and not GroES is bound, MBP remains unfolded (Fig. 3.4F). Because the MBP has DNA linkers attached, these data suggested the SR1-GroES cavity is not hermetically sealed [147]. It is thus possible that GroES does not require intimate contact with all seven SR1 monomers to form a stable complex capable of initiating folding.
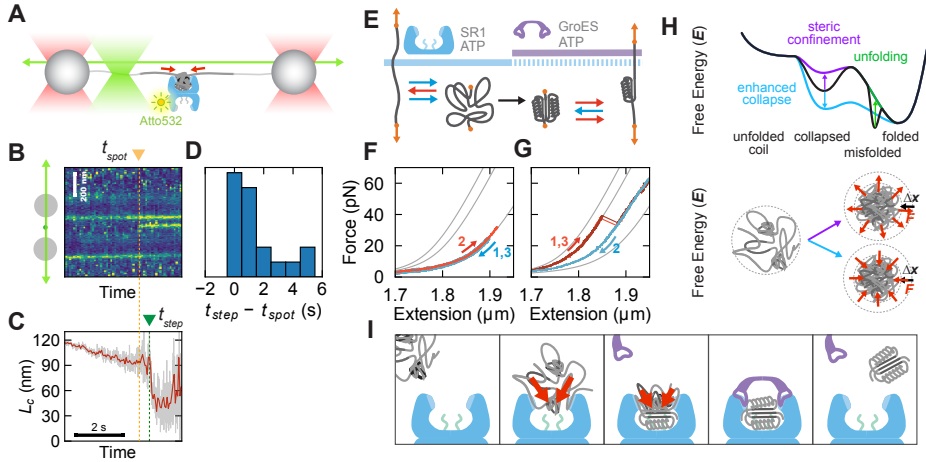
Figure 3.4: Compaction and folding in a single GroEL tetradecamer. **(A)** Experimental set-up: optical tweezers (red beams) and scanning fluorescence excitation (green beam), with GroEL (blue, 100 nM) labelled by Atto532 (yellow), and ADP (1 mM). **(B)** Resulting fluorescence emission kymograph. Fluorescent spot at time tspot shows single GroEL binding during relaxation. **(C)** Simultaneously measured contour length (L_c, red is filtered) shows step at time t_{step}. **(D)** Time between binding and L_c step ($N = 19$). **(E)** Flow-in protocol to test ternary complex (SR1-GroES-MBP) formation. **(F)** In presence of 200 nM SR1 and 1 mM ATP, MBP remains unfolded during relaxation (1), stretching (2), and relaxation (3), consistent with SR1 binding. **(G)** After buffer exchange to 500 nM GroES and 1 mM ATP, stretching (1) follows MBP core WLC curve (middle gray trace). This core state unfolds partially (right gray curve is WLC for fully unfolded state). After relaxation (2), stretching shows core state again (3). **(H)** Energy landscape cartoons for different acceleration models. Green: GroEL exerts pulling forces on misfolded states, raising their free energy to allow escape to other folded states. Purple: In steric confinement, chains are compacted (Δx) against counteracting entropic expansion forces, also raising their free energy, and effectively lowering the folding barrier. Blue: In enhanced collapse, chains are also compacted (Δx), but the mechanism differs and the associated attractive forces do not counteract, resulting in a lower free energy. The folding barrier is lowered, for instance by closer proximity of residues that must contact to fold, similar to autonomous chain collapse. **(I)** Cartoons of event sequence emerging from our data, without being complete. Substrate chain with limited autonomous collapse binds GroEL apical domains at different sites; GroEL-mediated forces draw unbound chain segments together; stimulation of partial folding; competitive GroES binding releases bound segments, which are thus free to collapse and fold in stimulated manner; GroES and substrate are released. Various factors including substrate sequence, initial conformation, and chance encounters with apical domains may affect this process.

3.3. DISCUSSION

The folding stimulation by enhanced collapse that our data suggests is distinct from current models [70, 135] (Fig. 3.4H, I). In confinement models, steric repulsion forces exerted by a closed cavity decrease the chain entropy and thus increase its free en-

3

ergy, which effectively lowers folding barriers. In unfolding models, pulling forces on misfolds also increase the chain free energy and allow escape to productive folding trajectories. Here, we measured attractive forces that enhanced protein chain compaction, mediated by open GroEL (not closed by GroES). These measurements rather suggested a decreased chain free energy, while the observed increase in folding probabilities indicated reduced folding barriers. Consistently, spontaneous collapse is believed to lower barriers [148]. Other aspects of our observations agreed with previous work as well. Though compaction mechanisms and their resulting states differ, both confinement and collapse bring chain segments together, making enhanced collapse consistent with diverse experimental findings [86, 135]. We also found unfolded chain segments to bind GroEL apical domains [135], though those that did not or were released by GroES appeared free to collapse, mediated by the attractive forces (Fig. 3.4I). The enhancement of collapse strength can also act in conjunction with other mechanisms. Notably, the attractive forces may help to unfold misfolded chains that are apically bound, while the steric constraints of open and closed cavities can help define compacted states.

GroEL-ES supports a broad range of proteins, suggesting that limited polypeptide collapse could be a more general folding impediment that these chaperonins help resolve. Collapse enhancement by GroEL-ES may be vital to expand the space of foldable proteins, limit the lifetime of aggregation-prone collapsed states, and to drive substrate transfer from Hsp70 to GroEL. Beyond GroEL-ES, one may speculate that the ability to manipulate collapse strength is exploited to influence conformations of intrinsically disordered proteins, to stimulate folding in type II chaperonins such as TRiC/CCT, and more generally in the protein quality control machinery.

3.4. MATERIALS AND METHODS

3.4.1. PROTEIN EXPRESSION AND PURIFICATION

Expression and purification of MBP, dmMBP and rhodanese. MBP and dmMBP were overexpressed in T7 competent cells (NEB laboratories) in LB medium supplemented with 0.2% glucose and 50 μg/ml kanamycin at 30 °C until OD600 0.6, induced with 0.4 μM IPTG (Sigma) and incubated at 18 °C overnight. The culture was harvested by centrifugation at 5000 g for 20 minutes at 4 °C. All following steps were carried out at 4 °C. The pellet was resuspended in ice-cold buffer A (50 mM phosphate buffer pH 7.5, 200 mM NaCl, 10 mM EDTA, 50 mM Glutamic Acid–Arginine (Sigma) and 3 mM β-mercaptoethanol (Sigma)) and lysed using an Emulsiflex homogenizer. The lysate was cleared from cell debris by centrifugation at 50000 g for 1 hour followed by incubation with Amylose resin (NEB) for 1 hour. After extensive washing with buffer A, the proteins were eluted using buffer A supplemented with 20 mM maltose. For rhodanese, the pellet was resuspended in buffer B (100 mM Tris-HCl pH 7.0, 5 mM EDTA, 20 mM $Na_2S_2O_3$, 2 mM β-mercaptoethanol) and lysed as described above. The lysate was mixed with Protino™ Ni-NTA Agarose (Macherey-Nagel) and incubated for 1 hour. After washing, the protein was eluted with buffer B supplemented with 250 mM imidazole.

GroEL was expressed from an inducible plasmid in E. coli BL21 in LB at 37 °C. After cell disruption, the crude lysate was clarified by ultracentrifugation (142000 rcf), followed by anion exchange chromatography (FastFlow Q, GE) equilibrated in buffer C (50 mM Tris pH 7.4, 0.5 mM EDTA, 2 mM DTT) and eluted by linear gradient from 7.5% to 35% with buffer D (50 mM Tris pH 7.4, 0.5 mM EDTA, 2 M NaCl, 2 mM DTT). GroEL fractions were concentrated by 70% (w/v) ammonium sulfate precipitation. This precipitate was solubilized and dialyzed against 50 mM Bis-Tris pH 6.0, 50 mM KCl, 0.5 mM EDTA, 2 mM DTT containing 25% (wild-type GroEL) or 12.5% (all GroEL mutants) methanol. A second round of strong anion exchange (FastFlow Q, GE), run in the same methanol-containing buffer at pH 6.0, was used to strip co-purifying small proteins and peptides from the GroEL oligomers. To further remove contaminating proteins and peptides that remain tightly associated through prior stages of purifiation, GroEL fractions were gently agitated in the same methanol-containing buffer and AffiBlue Gel (BioRad) resin overnight at 4 °C under an argon atmosphere. The final sample was dialyzed into storage buffer (25 mM Tris pH 7.4, 100 mM KCl, 0.5 mM EDTA, 2 mM DTT), supplemented with glycerol (15–20% v/v), concentrated, and snap frozen using liquid nitrogen.

GroES was expressed from an inducible plasmid in E. coli BL21(DE3) in LB at 37 °C. After cell disruption, the crude lysate was clarified by ultracentrifugation (142,000 rcf), followed by acidification with sodium acetate, and cation exchange chromatography (FastFlow S, GE) equilibrated in buffer E (50 mM NaOAc pH 4.6, 0.5 mM EDTA, 2 mM DTT) and eluted by linear gradient from 0% to 25% buffer F (50 mM NaOAc pH 4.6, 0.5 mM EDTA, 2 M NaCl, 2 mM DTT). The sample was dialyzed against 25 mM Tris pH 7.4, 0.5 mM EDTA, 50 mM KCl, 2 mM DTT and applied to a strong anion exchange column (Source Q, GE). GroES was eluted with NaCl and enriched fractions were pooled. The sample was dialyzed into storage buffer supplemented with glyc-

erol (15 – 20% v/v), concentrated, and snap frozen using liquid nitrogen.

For the expression of Single Ring GroEL (SR1), E. coli BL21 DE3 transformed with pSR1 was grown in LB-Ampicillin (100 μg/ml) at 30 °C to an OD600=0.5. Overexpression was induced by adding 1 mM IPTG and growth was continued for 3 hours. Cells were harvested by centrifugation and stored at -70 °C after flash freezing in liquid nitrogen. Frozen cells were resuspended in 20 mM Tris-HCl pH 7.4, 50 mM KCl, 1 mM EDTA, 1 mM DTT, 1 mM PMSF, lysed using a French Press and cell debris were removed by centrifugation. SR1 was enriched by fractionated $(NH_4)_2SO_4$ precipitation between 35% and 45% saturation. Following dialysis in 50 mM Tris-HCl pH 8, 1 mM EDTA at 4 °C, the protein solution was fractionated using a DEAE Sepharose Fast Flow anion exchange chromatography resin (GE healthcare) eluting with a gradient from 0 to 1 M NaCl and further fractionated by size-exclusion chromatography using a HiPrep 26/60 Sephacryl S-500 HR column. SR1 containing fractions were pooled, concentrated using Amicon Ultra centrifugal filters (Merck), frozen in liquid nitrogen and stored at -70 °C.

3.4.2. GroEL labeling

The GroEL variant (EL315C) [149] was labeled with Atto532 maleimide (Sigma). Reactive dyes were prepared fresh from dry powder in anhydrous dimethylformamide (DMF) immediately prior to use. All proteins were first buffer exchanged 300-400x the original volume by a Vivaspin Turbo 15 (Satorious) into 50 mM Tris buffer pH 7.4, 100 mM KCl, 0.5 mM EDTA, and 1 mM TCEP. The proteins were then run over gel filtration (PD-10 column; Pharmacia) equilibrated in reaction buffer (50 mM Tris pH 7.4, 100 mM KCl, 0.5 mM EDTA, 0.5 mM TCEP). EL315C was concentrated to a final concentration of 70 μM (monomer) in a volume of 5 mL. Protein samples were added to individual 5 mL conical Weaton reaction vials, followed by two sequential reactive dye additions. Freshly prepared Atto-532 maleimide in DMF was added at a molar ratio of 1:6.5 to EL315C monomer. Following each addition, the sample was incubated for 45 minutes in the dark at 23 °C. Following the full 1.5 hours reaction time, the sample was quenched by addition of 5 mM glutathione. The labeled EL315-Atto532 were separated from unreacted dye by four rounds of dilution and concentration in a Vivaspin Turbo 15 (Sartorious), followed by gel filtration (PD-10 column; Pharmacia). The labeled proteins were then supplemented with glycerol (15-20%) and snap frozen using liquid nitrogen. Protein concentration was determined using a calibrated Bradford assay, in which the protein standard was from a sample of wild-type GroEL whose concentration had been previously established. Conjugated dye concentrations are determined by absorption spectroscopy of the denatured proteins (in 6 M Gdm buffer) using the following corrected extinction coefficient: Atto532, 115000 M-1cm-1. GroEL-Atto532 activity was confirmed by MESG ATPase activity assay (EnzChek, Molecular Probes) and native gel filtration (Superdex 200, GE).

3.4.3. GroES mobile loops and GroEL C-tails

The GroES mobile loops [144] ETKSAGGIVLTGS and GroEL C-tails (GGM)$_4$M were ordered from Genscript. GroES mobile loops and C-tails were dissolved in MQ water

and snap frozen using liquid nitrogen. Prior to measurements the samples were dissolved in HMK buffer (50 mM HEPES, pH 7.5, 5 mM MgCl$_2$, 100 mM KCl). The GroES mobile loops were added in fivefold molar excess to GroEL during optical tweezers experiments (Fig. 3.3, Fig. S3.7).

3.4.4. PROTEIN-DNA COUPLING

The attachment of DNA handles (of 2600 or 5000 base pairs) was performed using the strategy described in Chapter 2.

3.4.5. OPTICAL TWEEZERS ASSAY

Carboxyl polystyrene beads (CP-20-10, diameter 2.1 μM, Spherotech) were covalently coated with sheep anti-digoxigenin antibody (Roche) via carbodiimide reaction (PolyLink Protein coupling kit, Polysciences Inc.). Around 50 ng of the generated construct were incubated with 2 μL beads in 10 μL HMK buffer (50 mM HEPES, ph 7.5, 5mM MgCl$_2$, 100 mM KCl) for 15 minutes in a rotary mixer at 4 °C and rediluted in 350 μL HMK buffer. With our coupling strategy, approximately 50% of the constructs will be asymmetrically functionalized with digoxigenin and biotin in each side. In order to create the second connection, we employed Neutravidin coated polystyrene beads (NVP-20-5, diameter 2.1 μM, Spherotech). Once trapped, beads were brought into close proximity to allow binding, and tether formation was identified by an increase in force when the beads were moved apart. The data was acquired at 500 kHz and averaged to 500 Hz. For constant force measurements, tension was held at 2 pN on average for 30 s using a proportional–integral–derivative (PID) feedback loop, before pulling again at constant velocity (Fig. 3.1E). In Fig. S3.3C, the distance between the traps is constant, while the extension of the protein is monitored as it changes conformation. Note that the beads can change position within the traps. For fluorescence measurements in combination with stretch-relax experiments (Fig. 3.4A to D), Atto-532 labelled GroEL proteins were visualized using a green excitation laser (532 nm), with 2 mM Trolox and 4 mM β-mercaptoethanol in the buffer. The excitation beam was used to scan along the tethered construct at 10 Hz during the force-spectroscopy measurements, generating fluorescence kymographs that were aligned to the force signal using a custom Python software.

3.4.6. FORCE SPECTROSCOPY DATA ANALYSIS

Several checks were performed to confirm that the data corresponded to a proper single tether, which include comparing the total measured unfolded length to the expected length, consistency with the WLC model (at higher forces), overstretching at 67 pN, and final tether breakage in one clean step. The unfolding forces (F_U), contour lengths (L_C), refolding forces (F^*) and compaction energies (E_C) were quantified from force extension data, using an open source MATLAB code after modifications. F_U was determined from stretching traces as the force required to fully unfold a protein (Fig. 3.1C left, E, 3.2C to E, 3.3 and S3.7). For stretching traces in which the protein did fully unfold below the maximum force that could be applied (67 pN, corresponding to the DNA overstretching plateau), F_U was determined as 67 pN, the maximally

sustained force (Fig. S3.8, Fig. 3.2C). The contour lengths (L_c) of refolded states were determined from the force-extension data of the stretching curve before the first unfolding transition, using the WLC model (Fig. S3.1A, B). The persistence lengths of the DNA (45 nm) and protein (1.5 nm), and the stretch modulus of DNA (1200 pN) were fixed parameters in the WLC model. In Fig. 3.2A and 3.4C, the instantaneous protein contour length was calculated using the same WLC model. Compaction energy (E_C) was calculated by quantifying the area under the relaxation curve and then subtracting the area under the WLC curve for fully unfolded protein (Fig. S3.4B, C, Fig. S3.7, and Fig. 3.2C). P^* was determined as the fraction of relaxation traces that show (one or more) steps in L_c of more than 15 nm (Fig. 3.2C, 3.3, Fig. S3.7). F^* was quantified as the measured force just before such a step in L_c (during relaxation, Fig. 3.2C, 3.3, and S3.7). The folding probability (P_c) was quantified as the fraction of relax-stretch cycles showing refolding to the core MBP state (Fig. 3.2C, 3.3, S3.1 and S3.7).

3.4.7. Additional statistical calculations

The statistical significance of differences in folding probability (P_c) and refolding at force probability (P^*) between experimental conditions was calculated using one tailed two proportion z-test. The statistical significance of differences in compaction energy (E_c) and maximally sustained forces (F_U) between experimental conditions was calculated using two sample assuming unequal variance t-Test. Test results are mentioned as p values in the main text. In box charts, whiskers indicate 90% and 10% extreme values, the inner line represents the median, the length of the box indicate interquartile range and the inner small square the mean of the population.
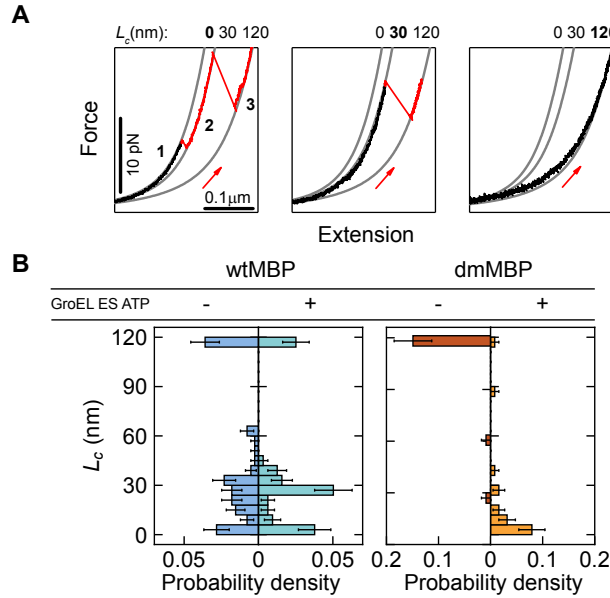
3.5. Supplementary figures

Figure S3.1: Contour lengths of refolded states. **(A)** Determination of contour lengths (L_c) from force-extension data taken during stretching. Displayed stretching curves are taken after relaxation of unfolded chains and waiting for 5 s at 0 pN to allow refolding. The protein states after this refolding window are characterized by their contour lengths, which represents the length of the non-compacted part of the protein chain. This L_c of refolded states is determined based on the first part of the stretching data, where no (detectable) unfolding has yet occurred (in black). This data is described by a single mean L_c value that is determined using the worm-like chain (WLC) model (gray curves). Gray curves: force-extension behavior of DNA tethers attached to an unfolded protein chain of three different contour lengths L_c. Indicated are WLC curves for fully compacted ($L_c = 0$ nm), MBP core state ($L_c = 30$ nm), and fully unfolded state ($L_c = 120$ nm). Panels indicate different example stretching curves that are observed for MBP with 200 nM GroEL, 500 nM GroES, and 1 mM ATP. Left panel: refolded state $L_c = 0$ nm (1), followed by unfolding to core state (2), and to fully unfolded state (3). Middle panel: refolded state $L_c = 30$ nm (core state), which then unfolds to fully unfolded state. Right panel: refolded state $L_c = 120$ nm (fully unfolded state). For traces that show no unfolding transitions, like the latter example, L_c is measured at 10 pN. **(B)** Probability density (P.D.) of contour lengths (L_c) of refolded states. Determination of L_c as described in panel A, for MBP (blue), and dmMBP (orange) in the presence and absence of GroEL-ES and ATP. This analysis is also performed to quantify the frequencies (P_c) of refolding to the core MBP structure. Also note that refolded states larger than the core structure always displays unfolding to (via) the core state, which involves detachment of MBP c-terminal helices from the core structure. Error bars are s.d.
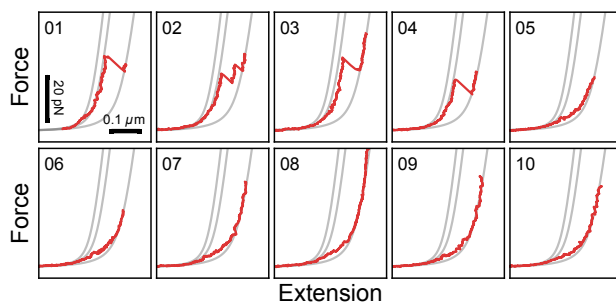
3

Figure S3.2: Irreversible switching to unfolded states in presence of GroEL. Force extension traces of MBP showing a sudden switch to a stable unfolded state. Successive stretching traces from relax-stretch cycles for MBP in the presence of 200 nM GroEL and 1 mM ADP, which initially (cycles 1 to 4) show the data following the worm-like chain (WLC) curve of the MBP core state (middle gray line) and thus indicating core refolding, followed by progressive unfolding to the fully unfolded state (right gray line). However, the data follows the WLC curve of the unfolded state and shows no unfolding transitions in the subsequent cycles (5 to 10), indicating the chain remained stabilized in the unfolded state. The data was close to the WLC curve corresponding to the unfolded state, though the deviation at lower forces suggested that a compacted yet non-folded state continued to be formed and disrupted.
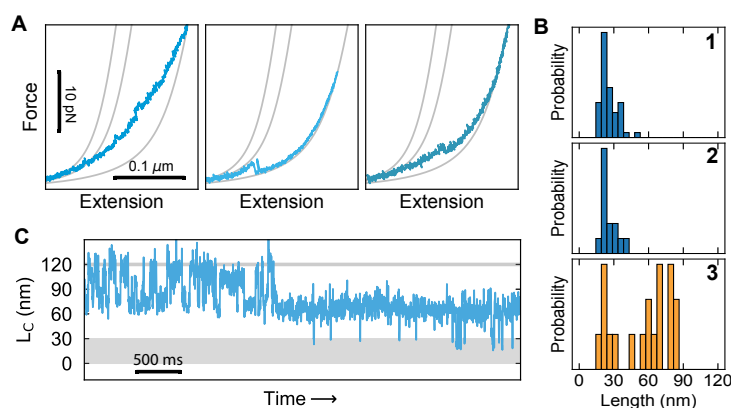
Figure S3.3: GroEL-mediated chain compaction and hopping transitions. (A) Three example relaxation curves, displayed in force-distance graph, showing gradual and step-wise compaction, for dmMBP in the presence of 200 nM GroEL and 1 mM ADP condition. **(B)** Histograms of contour length changes of step-wise compactions during relaxation (see panel A) for MBP, 200 nM GroEL and 1 mM ADP (1, $N = 52$), MBP, 200 nM GroEL398 and 1mM ATP (2, $N = 14$) and dmMBP, 200 nM GroEL and 1 mM ADP (3, $N = 19$) conditions. **(B)** Contour length vs time trace, showing repeated step-wise transitions (hopping) between states for dmMBP, mediated by GroEL. Data is taken with both traps at a constant position, in the presence of 200 nM GroEL and 1mM ADP. We stress that the details of these data, such as the folding step-sizes, are specific to this experiment and not generally observed. The latter may be expected. For hopping transitions of isolated proteins without chaperones, the energy landscape is defined only by the tethered protein that is in principle the same for different experiments. In contrast, here the energy landscape is also defined by GroEL, and how and where it is interacting with the substrate, which has a random aspect, and hence will produce differences between experiments.
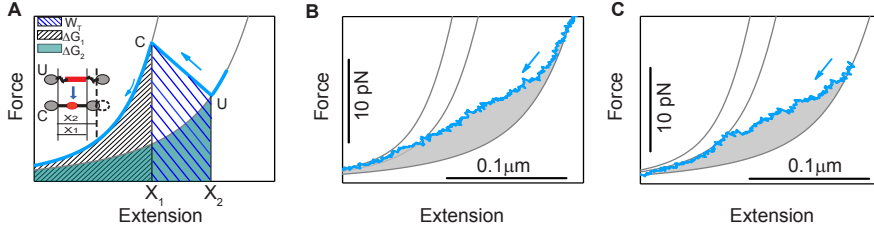
3

Figure S3.4: Determination of compaction energy. (A) Schematic diagram showing how the compaction energy (E_c) is determined from force extension cycles. We consider an unfolded protein chain during relaxation that suddenly compacts fully, but the derivation can also be used for gradual compaction composed of multiple smaller compaction events. At this single sudden event (U→C), the measured distance between beads (Extension, along the x-axis) suddenly decreases from x_2 to x_1, and the measured force $F(x)$ that acts on these beads and throughout the DNA-protein-DNA tether (Force, along the Y-axis) increases from $F(x_2)$ to $F(x_1)$, because the tether is now effectively shorter and hence its tension higher (see inset in panel A). In the case of slow relaxation where the system is in equilibrium and there is no heat dissipation, energy is conserved, and hence the corresponding increase in potential energy of the beads equals to the work done (WT) by DNA-protein-DNA tether. WT is then estimated as the area under the force-extension curve from x_2 to x_1 that quantifies the increase in bead potential energy. Note that displacing an object over distance dx against a force F costs an amount of energy F·dx. Thus, WT is quantified by (wide-hash region). WT can be decomposed into two contributions: the work done to compact the protein (E_c) and to increase the tension in the DNA-protein-DNA tether (W_{ext}), hence: $W_T = W_{ext} + E_c$. W_{ext} can be calculated using the WLC model. To use this, we write $W_{ext} = \Delta G_2 - \Delta G_1$, where ΔG_2 is the work done in extending the unfolded protein chain and DNA linkers from extension 0 to x_2 (green region), which is calculated with the WLC model. ΔG_1 is the work done in extending only DNA from 0 to x_1 (narrow-hash region), and is also calculated with the WLC model. ΔG_1 has no contribution from the protein chain because it is fully compacted in this state. The work done in compacting the protein (E_c) can then be calculated as $W_T + \Delta G_1 - \Delta G_2$. Thus, in graphical terms, E_c equals the size of the wide-hash (W_T) plus narrow-hash (ΔG_1) regions minus the size of the green region (ΔG_2). In more simple terms, this is thus the size of the area under the measured curve $F(x)$ minus the size of the area under the WLC curve for the unfolded protein (right gray curve), as illustrated for measured data in panels B and C. Note that in the latter one can integrate from $x = 0$ to any $x > x_2$, as beyond x_2 the chain is fully unfolded and hence there are no further area contributions. Perhaps counterintuitively, E_c is thus determined not only by $F(x)$ for x in between x_1 and x_2, but also by $F(x)$ for x in between 0 and x_1. Note that while the compacted chain (C) may in principle be deformed for $x < x_1$, the length changes as well as energies are negligible, owing to its high stiffness compared to the DNA, while the force is identical throughout the chain. Finally, we note that this estimate of E_c is a lower-bound, given that the system is not fully in equilibrium. (**B** to **C**) For gradual collapse and collapse with refolding jumps, the compaction energy E_c is thus determined by the size of the indicated gray area.
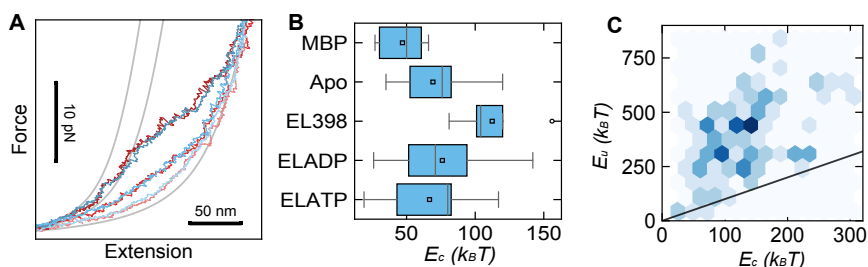
Figure S3.5: Collapse without folding, and stretching energy. (A) Force-Extension traces that show compaction during relaxation (blue), and after waiting at 0 pN for 5 s, are followed by stretching (red) that do not show unfolding features. These cycles indicate that significant compaction can occur without the formation of detectable folded states. Note that most cycles in these conditions rather show refolding instead (See Fig. 3.2). The displayed data is for MBP in presence of 200 nM GroEL and 1 mM ADP. **(B)** Corresponding compaction energy (E_c) of relaxation traces that do not produce detectable refolding (for examples see panel a) in different conditions. E_c is highest for EL398ATP and ELADP (as in Fig. 3.2C). **(C)** Area under the stretching curve (stretching energy Eu) against the compaction energy (E_c) from the prior relaxation trace for MBP with 200 nM GroEL and 1mM ADP condition. The data here includes both cycles that do show refolding and cycles that do not. Black line indicates $E_u = E_c$. Relax-stretch cycles producing no folding (panels A and B) are close to this line.
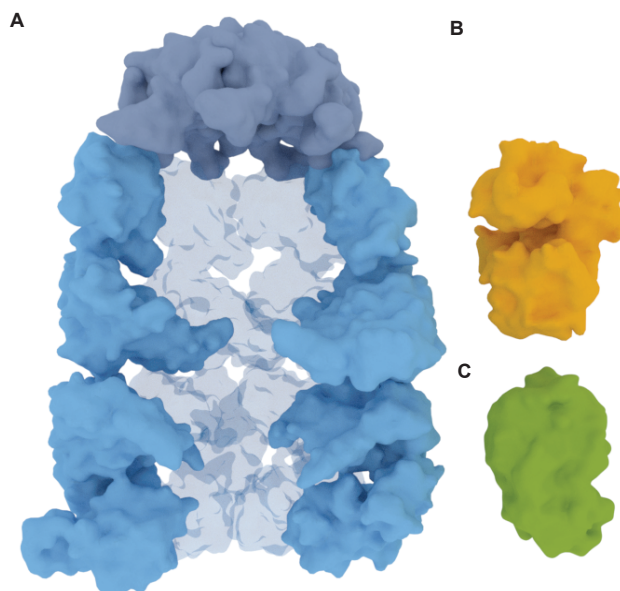
3

Figure S3.6: Protein structures. (A) A medial slice of GroEL-GroES ADP bullet, side view (PDB: 1PF9). **(B)** MBP (PDB ID: 2MV0) in orange. **(C)** Rhodanese (PDB ID: 1rhs) in green. Proteins are displayed in the same scale to compare their relative sizes.
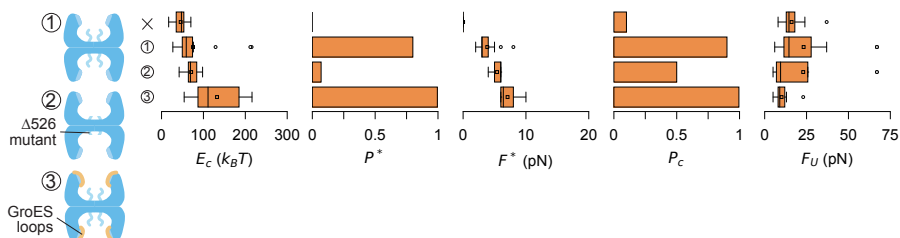
Figure S3.7: Roles of GroEL apical domains and cavity, dmMBP data. From dmMBP relax-stretch cycles, we quantified: total compaction energy during relaxation (E_c), probability (P^*) at force (F^*) of steps during relaxation, core refold probability after 5 s at 0 pN (P_c), unfolding force (F_U). Conditions are, from top to bottom: No Chaperone (x, $N = 19$), 200 nM GroEL and 1 mM ADP (1, $N = 21$), 200 nM ELΔ526 and 1mM ADP (2, $N = 11$) and 200nM GroEL, 1mM ADP and 1 μM loops (3, $N = 16$). For dmMBP without GroEL (x), F_U of first pulls is displayed because of the low refolding rate.
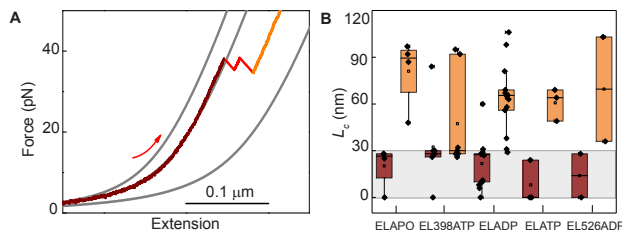

Figure S3.8: Lengths of stabilized partially folded structures. **(A)** Stretching curve showing stabilization of partially folded MBP states against forced unfolding, in the presence of GroEL-ADP. Displayed data initially follows worm-like chain (WLC) curve of MBP core state (dark red), then unfolds partially in two steps (red), to a partially folded state that is stable against high applied forces (orange). **(B)** Distributions of contour lengths Lc. Dark orange: Lc of initial MBP stretching data (panel a, dark red). Light orange: Lc of observed MBP structures that are stable against forces over 40 pN (panel A, orange). 40 pN is the maximum force in the absence of chaperonin; see Fig. 3.2C. The data indicates that the GroEL-stabilized MBP structures are typically smaller than the core state, in different nucleotide conditions.

4

Direct observation of processive peptide loop extrusion by a Hsp100 disaggregase

Hsp100 disaggregases perform the crucial task of dissociating proteins from aggregates and promoting their refolding. This process has not been observed directly, however, and its mechanistic basis is controversial. Here we demonstrate that the AAA+ chaperone ClpB processively extrudes polypeptide loops through its central pore in an ATP-driven process. By integrating optical tweezers with super-resolution imaging, we show that ClpB translocates both loop arms simultaneously, switching to single-arm translocation when encountering obstacles. Extrusion is strikingly powerful and fast, with exerted forces above 50 pN and speeds of over 450 residues per second, in periodic salvos of multiple ClpB monomers firing. Remarkably, extruded polypeptides refold upon exiting the translocation channel, similar to co-translational folding but in a looped topology. We conjecture that processive polypeptide loop extrusion is a general protein extraction principle underlying all Hsp100 disaggregases.

4.1. INTRODUCTION

How cells mitigate protein aggregation is a central problem in biology, and key to issues ranging from protein homeostasis to neurodegeneration and ageing [150–152]. The Hsp100 chaperone family of AAA+ disaggregases, which includes bacterial ClpB and yeast Hsp104, provide the critical capability to rescue aggregated proteins in bacteria, fungi and plant cells [153]. The mechanism of protein dissociation and refolding remains debated [154, 155]. It has been proposed that exposed polypeptides are translocated processively through the central Hsp100 pore, with multiple translocation steps in sequence before release [156]. On the other hand, Hsp100 disaggregases may instead act non-processively, as supported by recent work [157], which could involve single-step translocation, stabilization of spontaneously released polypeptide segments and entropic pulling, as proposed for Hsp70 disaggregation [158–160]. Unlike other peptide translocases, ClpB and Hsp104 hexamers rapidly exchange monomers, and display fast ring opening and closing dynamics [161, 162]. Initially, they may thus assemble around exposed polypeptides, or polypeptides may be inserted into closed rings at their termini or as a loop [163]. Hsp70 is believed to help targeting Hsp100 to substrates, activate ATP hydrolysis, and assist refolding of translocated polypeptides [153, 164]. Overall, the capacity to follow substrate-processing by Hsp100 chaperones in time is central to elucidate the principle of disaggregation in cells.

4.2. RESULTS

We studied the dynamics of Hsp100 disaggregases using single-molecule techniques. Maltose binding protein (MBP) was coupled to DNA handles at both termini and tethered between functionalized polystyrene beads that were trapped and manipulated with optical tweezers (Fig. 4.1A). After mechanical unfolding of the protein (Fig. 4.1A, fig. S4.1A), we reduced the tension and maintained it constant between 5 and 10 pN, high enough to prevent spontaneous refolding (Fig. 4.1A). Addition of ATP and ClpB-Y503D, a mutant altered in the regulatory M-domain that does not require Hsp70 (DnaK) binding for ATPase activation [165], resulted in isolated episodes of contractions in the bead-to-bead distance (Fig 4.1B). Zooming in showed continuous linear decreases in the effective polypeptide contour length Le, typically from 360 to 0 amino acids (aa), identical to the full MBP length (Fig. 4.1C, fig. S4.1B). Runs were followed by an arrest and/or an abrupt increase back to 360 aa (termed back-slips). These data indicated sequences of processive translocation runs of up-to the complete MBP chain, back-slips, and run restarts.

Consistent with a processive mode of action, longer polypeptide constructs of two and four MBP repeats in tandem arrangement showed longer runs (fig. S4.1C-F, S4.2). Furthermore, translocation was abolished in the presence of ADP, when mutating either of the two ClpB ATPase catalytic centers to prevent ATP hydrolysis (E279A, E678A) or the substrate-contacting pore loops (Y251A, Y653A), and by deleting the N-terminal domain that forms the pore entry (fig. S4.3A,B). Runs were less frequent but showed the same velocity for an M-domain mutant (K476C) that stimulates ATP hydrolysis less strongly [165], and for wild-type ClpB with the Hsp70 (DnaK) system

present [166] (Fig. 4.1D,E, fig. S4.3C,D, S4.4). These variants thus all differed in the time needed to start, but once initiated they translocated at the same speed. The latter appeared to peak at two values and extended beyond 500 aa/s (Fig. 4.1F), which is over 10-fold faster than other peptide translocases [167–169]. Strikingly, without slowing down, ClpB was able to exert high forces of over 50 pN, resulting in the melting of our DNA tethers (Fig. 4.1G). Overall, these findings further confirmed that the polypeptide chains were translocated processively through the central pore.
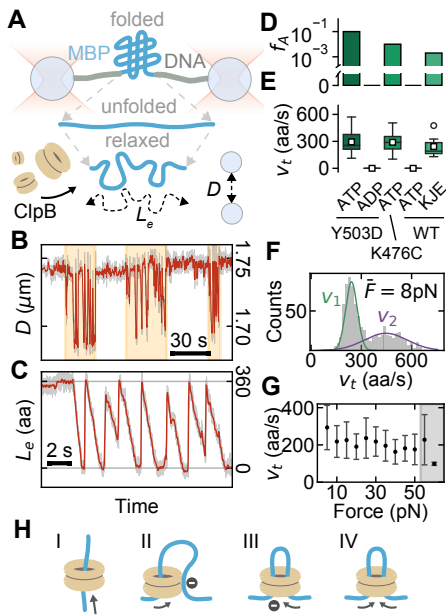

Figure 4.1: ClpB is a processive translocase. **(A)** Tethered MBP was unfolded with optical tweezers, relaxed to a low force that prohibits refolding, and moved to a region containing ClpB and ATP. **(B)** Tether contraction bursts (orange regions) in the presence of ClpB-Y503D and ATP. Raw signal in grey, filtered signal in red (2 Hz). **(C)** Protein contour length Le, computed using a WLC model, showing runs and back-slips. 360 aa is the contour length of MBP. Red: filtered signal (20 Hz). **(D)** Fraction of time showing ClpB translocation activity and **(E)** translocation speed for different chaperone compositions. KJE is the DnaK system (DnaJ and GrpE) with ATP. **(F)** Run speed distribution ($N = 800$) for ClpB-Y503D at 8 pN. Double Gaussian fit yields two velocities. **(G)** Translocation speed by ClpB-Y503D versus applied tension. Gray region indicates the start of DNA overstretching and upper force limit. **(H)** Possible ClpB translocation topologies.

Different models may explain how substrates are translocated through the ClpB channel (Fig. 4.1H). ClpB rings can dynamically assemble around polypeptides, which are then translocated (model I). However, this scenario would only result in the observed contractions if translocated chain segments also bind elsewhere on ClpB (model II), in a similar fashion to DNA extrusion by condensin [170]. Finally, the peptide could be inserted as a loop, with translocation of either one (model III) or both arms of the loop (model IV).

Testing these models with optical tweezers only is difficult. Hence, we developed a novel technique that integrates super-resolution imaging of ClpB with force sensing to determine which arm of the polypeptide loop is being translocated (Fig. 4.2A). We chose 2MBP as it yields longer runs. After unfolding, we exposed the construct to fluorescently labeled ClpB and ATP while scanning a confocal excitation beam along the tether and beads (Fig. 4.2B, fig. S4.5). To limit the parasitic signal emanating from the beads, caused by autofluorescence and bound ClpB, we used 5 kb-long DNA handles, thus providing additional spacing. Single ClpB binding events were identified by the appearance of a fluorescent spot between the beads (Fig. 4.2B), and translocation

was observed soon after (Fig. 4.2C). To reduce the background fluorescence of ClpB in solution and prevent additional ClpB binding, the tether was then moved to a solution containing only ATP (Fig. 4.2B, D). We determined the position of the fluorescent spot at a spatial resolution below the diffraction limit (Fig. 4.2E,F). Together with the tweezers measurements of the distance between MBP termini, this yielded two distances: between the MBP C-terminus and the bound ClpB, and between the MBP N-terminus and the bound ClpB, thus allowing us to distinguish between translocation of the two loop arms independently (Fig. 4.2A, fig. S4.6).
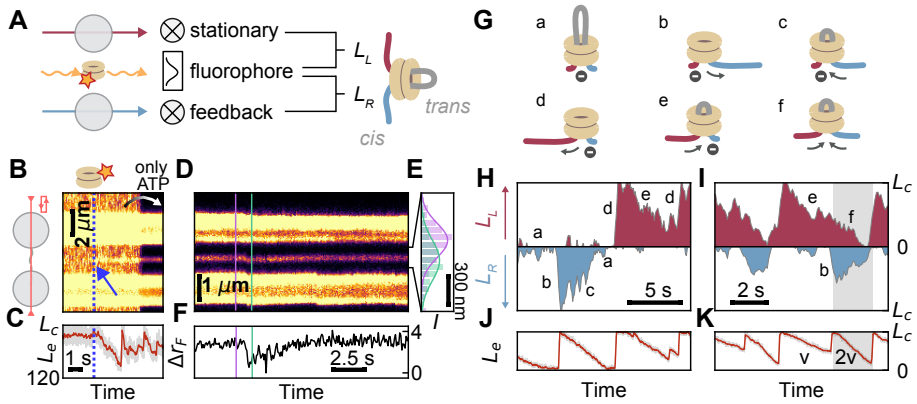

Figure 4.2: Optical tweezers combined with fluorescence reveals ClpB translocation of both loop arms. **(A)** Principle of approach: position of feedback-adjusted trap (blue) maintains force constant; the other trap is stationary (purple). Bead positions are detected at nm precision. Fluorescence detects ClpB-Atto633 position at sub-wavelength precision. Combining these signals allows quantification of the two non-translocated (*cis*) polypeptide segments: L_R and L_L. **(B)** Fluorescence kymograph during preparation, showing ClpB binding event (blue arrow), and movement to ClpB-free region. **(C)** Concurrent tweezers measurement of Le, showing translocation after binding. **(D)** Fluorescence kymograph during translocation. **(E)** Photon count of the ClpB spot along two scan lines (see panel D) and Gaussian fits. **(F)** Position of ClpB-Atto633, which here moves suddenly down and gradually up. **(G)** Cartoons of polypeptide processing events in panels H and I. **(H), (I)** *cis* polypeptide lengths of each arm (L_R and L_L, see panel A). Gray: double-speed translocation and, consistently, both arms shortening simultaneously. **(I), (J)** Total *cis* polypeptide length, $L_e = L_R + L_L$ as quantified by optical tweezers.

We found various sequences of events: after translocation of the entire chain, the right arm of the loop was released and slipped backwards until the chain was fully extended in *cis* again, and subsequently right-arm translocation re-started rapidly (Fig. 4.2G, event sequence a→b→c, Fig. 4.2H). A similar sequence on the left side occurred directly after (Fig. 4.2G, event sequence a→d→e, Fig. 4.2H). Strikingly, sometimes both arms were translocated simultaneously, each at approximately the same velocity (Fig. 4.2K, event f, Fig. 4.2I, gray region). Consistently, the total translocation speed, which reflects the speed by which the MBP termini approach each other, and is more accurate because it uses the optical tweezers signal only, was twice as high as in single-arm translocation runs (Fig. 4.2J,K). Models I and II do not yield two-arm
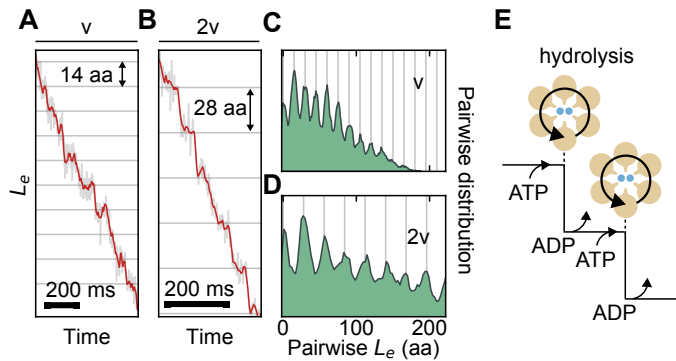
Figure 4.3: Translocation steps by ClpB. (A) Translocation runs with single-speed showing translocation steps. **(B)** idem for dual-speed runs. **(C)** Distribution of lengths between any two points in single-speed runs. **(D)** idem for dual-speed runs. **(E)** Translocation speed is doubled by doubling the step size (red) rather than the step frequency (blue), as evidenced by panels A-D.

translocation and hence can be ruled out, while models III and IV are both valid: polypeptide loop extrusion can occur by translocation of either of the two loop arms individually, or both loop arms simultaneously (Fig 4.1H). Switches between these two translocation modes took place after blockage of one arm, typically upon ClpB encountering the DNA tether. The data also confirmed directly that single ClpB rings remained intact and bound during runs, switches and back-slips (Fig. 4.2H,I), and explained the two peaks in the speed distribution (Fig. 4.2F).

This scenario was supported by multiple additional observations. First, 64% of the very first runs in a translocation burst initially showed the higher speed ($2v$) before switching to the lower speed (v), compared to 22% when considering all runs (Fig. S4.7A). This bias is logical, as when ClpB initially binds at a certain site on the MBP chain and starts translocating, both arms are most often still unobstructed. Consistently, the binding sites that we estimated from these data (Fig. S4.7B) agreed with high ClpB-NTD affinity regions on MBP, as determined using a peptide library binding assay (Fig. S4.7C-E). Second, lower-speed (v) runs displayed individual translocation steps of 14 aa, while higher-speed runs ($2v$) showed steps of 28 aa, which we could detect using shorter DNA handles to increase the resolution [171] (Fig. 4.3A-D, Fig. S4.8A, C). This data is consistent, since a decrease in the distance between termini (L_e) should be twice as large when both arms are translocated simultaneously. The data also showed that switching between translocation modes is achieved by changing the step size rather than the step frequency (Fig. 4.3E). The reported ATP-driven movement of single Hsp100 pore-loops corresponds to translocation steps of about 2 residues [172–174], which would not be detectable here, in part because of the high translocation speeds. The steps we observed are significantly larger, which implies the action of multiple Hsp100 monomers in rapid succession, in contrast to concerted models where they act simultaneously [175].

Finally, we wondered whether we could follow the fate of the extruded polypeptide loops, specifically if and when they could refold. Folding cannot be measured as usual, because translocated loops are not subjected to the applied forces that allow conformational sensing. Hence, we focused on the back-slip dynamics of the 2MBP construct (Fig. S4.9). We observed that substrate release to the *cis* side was often incomplete, with arrests at a measured length of about 310 aa (Fig. 4.4A, stars, and B). This length, which matches that of the MBP core (270 aa) plus two peptide segments inside the channel (~40 aa), suggested that one MBP core refolded in *trans* and subsequently arrested back-slipping by blocking the narrow ClpB pore (Fig. 4.4D and S4.9). Other observations also supported this hypothesis. First, the arrest probability increased as more of the polypeptide was translocated (Fig. 4.4C). Second, the back-slip arrests were not observed for a 2MBP mutant that cannot refold [84] (Extended Data Fig. S4.10). Overall, polypeptides thus started refolding when exiting the ClpB channel during translocation, analogous to co-translational folding of nascent chains. Notably, DnaK was not strictly required for refolding.
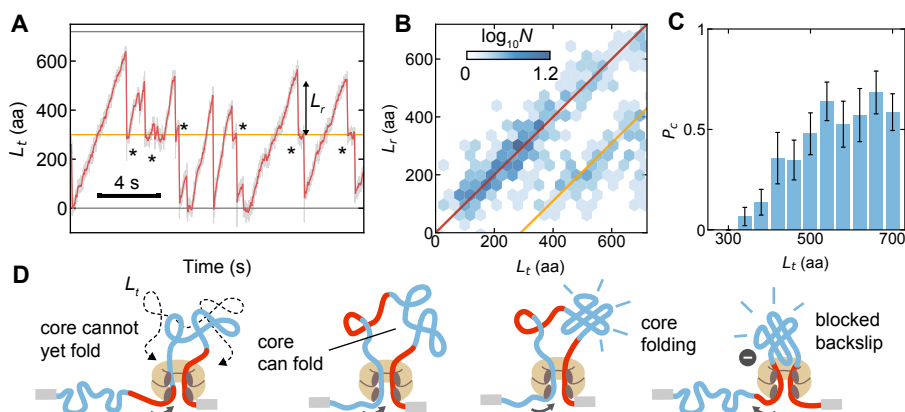

Figure 4.4: Substrate refolding during ClpB translocation. **(A)** Translocated polypeptide length during translocation for 2MBP. Gray lines indicate 0 and 720 aa, orange line 310 aa, the length of one MBP core plus the ≈40 aa inside the pore. Stars indicate back-slipping arrests at core length. **(B)** Red line: the released length (L_r) upon back-slipping equals all of the previously translocated length (L_t), indicating no arrest. Orange line: at higher L_t, L_r is reduced, leaving unreleased an amount of polypeptide equal to the length of one MBP core. **(C)** Probability of core folding (P_c) increases as more substrate has been translocated ($N = 427$). **(D)** Scheme of MBP refolding after ClpB translocation. Blue: core MBP. Red: non-core MBP.

4.3. DISCUSSION

In conclusion, our study on ClpB showed unambiguously that polypeptide loop extrusion is possible. Free substrate termini may also insert into the ClpB pore and be translocated in a non-loop topology, though we surmise that internal segments of aggregated proteins are targeted more readily and translocated as loops. ClpB is also

strikingly fast, processive, and powerful, and allows switching between single- and dual-arm translocation. ClpB thus appears to maintain a tight and long-term grip on both arms, with back-slips indicating a sporadic loss of contact, though it remains an intriguing open question how the independent handling of two arms is achieved at the structural level. These ClpB features are of direct relevance to efficient disaggregation (Fig. 4.5). Full dissolution of stable aggregates likely involves multiple ClpB rings and other chaperones such as Hsp70/DnaK, acting at different sites, at different moments in time, and involving many random dissociation and re-association events. Yet, ClpB translocation action itself is remarkably deterministic and processive once started. Overall, our findings define loop extrusion as the mechanistic basis of Hsp100 disaggregation, highlight the need for tight regulation of Hsp100 activity, and suggest that other polypeptide processing systems such as the Cdc48/p97 segregase, the ribosomal assembly factor Rix7/NVL, and the 26S proteasome may also exploit the capacity to handle multiple polypeptide strands in controlled manner.
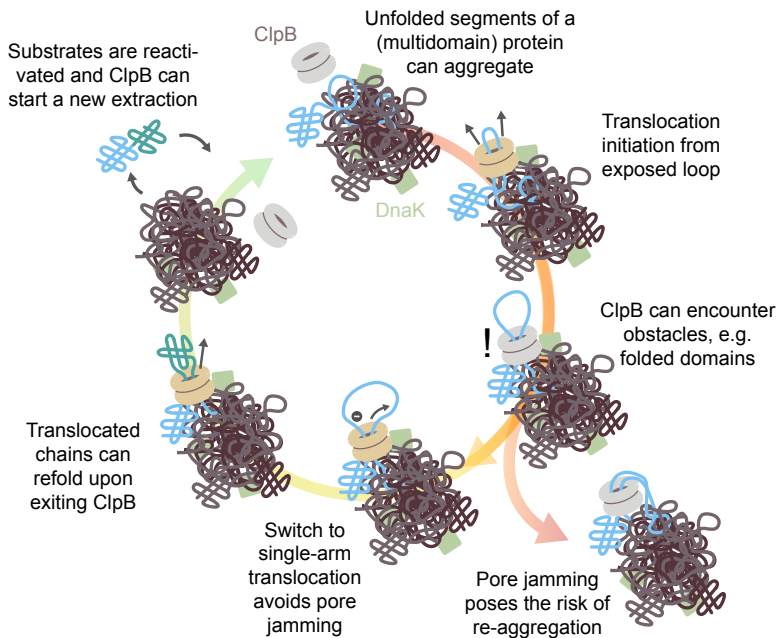
4

Figure 4.5: Loop extrusion and refolding as a disaggregation principle. Insertion and translocation of loops is key to efficient disaggregation of heterogeneous aggregates, which usually display few accessible polypeptide termini at the surface. Aggregated proteins often contain folded domains. The ability of Hsp100s to deal with obstacles is crucial to preserve those folds while preventing pore jamming. A highly processive, fast and powerful extraction process can effectively prevent re-aggregation of translocated polypeptides, which can instead refold while emerging from the central channel.

4.4. Materials and methods

4.4.1. Protein expression and purification

All MBP constructs were modified with cysteine residues using the pET28 vector. Linker sequences are GRGS and RITK for N- and C-terminus respectively. Proteins were purified from *E. coli* BL21(DE3) cells. For overexpression, overnight cultures were diluted 1:100 in fresh LB medium supplemented with 50 mg/l kanamycin, 0.2% glucose and incubated under vigorous shaking at 30 °C. Expression was induced at OD600=0.6 by addition of 1 mM IPTG and incubation overnight at room temperature. Cells were cooled, harvested by centrifugation at 5000g during 20 minutes, flash-frozen and stored at -80 °C. Cell pellets were resuspended in ice-cold buffer A (50 mM potassium phosphate pH 7.5, 150 mM NaCl, 3 mM Chloramphenicol, 50 mM Glu-Arg, 10 mM Complete Protease Inhibitor Ultra from Roche, 10 mM EDTA) and lysed using a pressure homogenizer. The lysate was cleared from cell debris by centrifugation at 50,000g for 60 min and incubated with Amylose resin (New England Biolabs) previously equilibrated in buffer A for 20 minutes at 4 °C. The resin was washed with buffer A three times by centrifugation and bound proteins were eluted in buffer A supplemented with 20 mM maltose. Purified proteins were aliquoted, flash frozen in liquid nitrogen and stored at -80 °C. ClpB and variants were overexpressed in *E. coli* Δ*clpB::kan* cells. Cell pellets were resuspended in LEW buffer (50 mM NaH$_2$PO$_4$ pH 8.0, 300 mM NaCl, 5 mM β-mercaptoethanol) and lysed by French press. Cleared supernatants were incubated with Protino Ni-IDA resin and bound proteins were eluted by LEW buffer containing 250 mM imidazole. ClpB containing fractions were pooled and subjected to Superdex S200 16/60 size exclusion chromatography in buffer BMDH buffer (50 mM Tris pH 7.5, 150 mM KCl, 20 mM MgCl$_2$, 2 mM DTT) containing 5% (v/v) glycerol.

4.4.2. ClpB-Atto633 labeling

Labeling of ClpB-E731C variants with Atto633-maleimide was performed in PBS buffer according to the instructions of the manufacturer (ATTO-TEC). Labeled ClpB-E731C was separated from non-reacted Atto633 by size exclusion chromatography using Superdex S200 HR10/30 in MDH buffer containing 5% (v/v) glycerol.

4.4.3. Protein-DNA coupling

The attachment of DNA handles (of 2600 or 5000 base pairs) was performed using the strategy described in Chapter 2.

4.4.4. Optical tweezers assay

Carboxyl polystyrene beads (CP-20-10, diameter 2.1 μM, Spherotech) were covalently coated with sheep anti-digoxigenin antibody (Roche) via carbodiimide reaction (PolyLink Protein coupling kit, Polysciences Inc.). Around 50 ng of the generated construct were incubated with 2 μL beads in 10 μL HMK buffer (50 mM HEPES, ph 7.5, 5mM MgCl$_2$, 100 mM KCl) for 15 minutes in a rotary mixer at 4 °C and rediluted in 350 μL HMK buffer. With our coupling strategy, approximately 50% of the constructs will

be asymmetrically functionalized with digoxigenin and biotin in each side. In order to create the second connection, we employed Neutravidin coated polystyrene beads (NVP-20-5, diameter 2.1 μM, Spherotech). Once trapped, beads were brought into close proximity to allow binding, and tether formation was identified by an increase in force when the beads were moved apart. ClpB was diluted in buffer C to a final concentration of 2 μM. For the ATP experiments, we used an ATP regeneration system (3 mM ATP, units/mL kinase, mM pep). To mitigate photobleaching and tether damage we added an oxygen scavenging system (3 units/mL pyranose oxidase, 90 units/mL catalase and 50 mM glucose, all purchased from Sigma-Aldrich).

4.4.5. FORCE SPECTROSCOPY DATA ANALYSIS

Data was recorded at 500 Hz using a custom-built dual trap optical tweezers for the tether resistance assays and a C-Trap (Lumicks) for the dual monitoring experiments. Data was analyzed using custom scripts in Python. Optical traps were calibrated using the power spectrum of the Brownian motion of the trapped beads [129], obtaining average stiffness values of $\kappa = 0.39 \pm 0.04$ pN/nm. Force-extension curves were fitted to two worm-like-chain (WLC) models in series (Fig S1a), using the approximation of an extensible polymer reported by Petrosyan [130] for the DNA, and the Odijk inextensible approximation for the protein contribution [131].

$$x = L_c^* \left(\frac{4}{3} \left(1 - \frac{1}{\sqrt{\beta^* + 1}} \right) - \left(\frac{10 \exp\left(\sqrt[4]{\frac{900}{\beta^*}} \right)}{\sqrt{\beta^*} \left(\exp\left(\sqrt[4]{\frac{900}{\beta^*}} \right) - 1 \right)^2} + \frac{\beta^{*1.62}}{3.55 + 3.8\beta^{*2.2}} \right) + L_e \left(1 - \frac{1}{2}\sqrt{\beta} \right) \right)$$

Where $\beta^* = (FL_p^*)/(k_B T)$, being F the force, T the temperature, L_p^*, K and L_c^* the persistence length, stretch modulus and contour length of DNA, respectively. For the protein, $\beta = (FL_p)/(k_B T)$ with L_p, L_e being the persistence and extended length of the polypeptiden. L_c^* was 906, 1750 or 3500 nm for the three different DNA handles used (1.3, 2.5 and 5 kb, respectively), L_c was 120 nm for the MBP and 105 nm for YPet, and L_p was fixed to 0.75 nm. L_p^* and K were fitted, yielding average values of 30 nm and 700 pN/nm respectively. These fitted parameters were then used to compute the instantaneous extended length of the protein (L_e) using the same WLC model (Fig. S4.1B). The translocated length (L_t) was computed by subtracting the extended length (L_e) to the total contour length of the protein (L_c). To classify translocation events, the translocated length signal was smoothed using a Savitzky-Golay filter [176] (Fig. S4.4A, black line), allowing to calculate its time derivative. Backslipping results in a large negative slope in the derivative, and was used as the criteria to separate individual translocation runs (Fig. S4.4B). Subsequent 1-dimensional dilation and erosion was performed to remove artifacts. Next, each individual run was similarly treated in order to find local changes in the slope (Fig. S4.4C), setting as threshold a value between the two known speeds (around 70 and 140 nm/s, Fig. S4.4D). Linear fits were performed in each identified region and reported as the local translocation speed (Fig. S4.4C). Only fits that yielded r values higher than 0.9 were considered.

4.4.6. FLUORESCENCE IMAGING ANALYSIS

For dual monitoring experiments, an excitation laser beam (with wavelengths of either 532 nm or 647 nm) was scanned along the beads and tether at a line rate of 12 Hz. The excitation power was 0.3 mW. To avoid parasitic noise from the beads, proteins were tethered using two 5 instead of 2.5 kbp DNA handles. Force-spectroscopy and confocal microscopy data were synchronized based on the movement of the beads (Fig. S4.5). The edge of the moving bead was tracked using a Gaussian fit (Fig S4.5B) and overlying it on top of the actual movement of the bead showed a clear time delay (Fig S4.5C). In order to find such delay, we computed the root-mean-square deviation (RMSD) between both signals for different time offsets (Fig S4.5D):

$$\text{RMSD}(r) = \frac{\sum_i^{N(r)} (X_i(r) - x_i(r))^2}{N(r)}$$

Where τ is the time offset applied to the tracked signal, $N(\tau)$ is the total number of points, $X(\tau)$ is the position of the bead according to the voltage of the mirror and $x(\tau)$ is the tracked position from the fluorescence kymograph. Minimization of RMSD(τ) provides an excellent estimate of the time delay between the signals (Fig S4.5D).

4.4.7. INTEGRATION OF OPTICAL TWEEZERS AND IMAGING SIGNALS TO COMPUTE THE LENGTH COMPONENTS

After ClpB binding and moving to a region containing only ATP, the fluorescent spot between the beads was fitted to a Gaussian distribution. To reduce the noise of the signal, we averaged the intensity profiles of three scanning lines before fitting. The resulting trajectory yielded the absolute position of ClpB with subpixel precision, which was then converted to nanometers using a factor of 80 nm/pixel. Next, we computed the position of each bead edge that is closest to ClpB (bottom edge for top bead and vice versa) using the trap position, bead displacement and bead radius. While it is possible to obtain these positions from the fluorescence kymograph as well, the optical tweezers data yields higher spatial resolution. We subtracted the ClpB tracked position from the position of bottom edge of the top bead, and the position of the top edge of the bottom bead from the ClpB position. The absolute value of these distances contains an arbitrary shift due to the mismatched reference system of the optical tweezers and confocal fluorescence images. In order to identify the offset, we used the fact that when the polypeptide is completely translocated (information present in the optical tweezers data, like Fig. 4.1 or fig. S4.2), both distances should be equal to each other and to half the distance between the edges of the beads D. After correcting for the shift, we obtained the absolute distance between ClpB and each of the beads (D_L and D_R). Since we use a force clamp, any change in distance is solely due to a change in the protein length ($x_L = D_L$ and $x_R = D_R$, fig. S4.6). Therefore, we removed the constant DNA contribution and computed the protein contour length from each distance (L_L and L_R) using the WLC model.

4.4.8. TRANSLOCATION STEPS CHARACTERIZATION

To increase the spatial resolution, we tethered a single MBP using 1300 bp DNA handles, 500 μM ATP and high tension (> 20 pN). Raw data was smoothed using a Savitzky-Golay filter of 5th order with a window of 21 data points. The difference between every distinct pair of data points was calculated and the sample binned to compute the pairwise length distribution. In order to calculate the periodicity more accurately, we computed the autocorrelation of the pairwise length distribution using the Pearson correlation coefficient for different lag lengths (Fig S4.8A and S4.8B) and its power spectrum employing the Welch method6 (Fig S4.8C and S4.8D). The autocorrelation distribution was fitted using the equation:

$$R(\tau) = \left(A\cos\left(\frac{2\pi}{T}t\right) + m\tau + n \right) Be^{-C\tau}$$

where $R(\tau)$ is the autocorrelation, τ is the time lag, T is the period of the steps and a linear and exponential function have been introduced to account for the decay in the signal (Fig S4.8A and S4.8B).

4.4.9. PEPTIDE LIBRARY DATA AND INITIAL ClpB BINDING LOCATION

The MBP peptide library was prepared by automated spot synthesis by JPT Peptide Technologies GmbH (PepSpotsTM). The library is composed of 13-mer peptides scanning the MBP primary sequence with an overlap of 10 residues. 1 μM ClpB-NTD (Met1-Ser148) was incubated for 30 min in buffer P (10 mM Tris pH 7.5, 150 mM KCl, 20 mM MgCl$_2$, 5% (w/v) sucrose, 0.005% (v/v) Tween 20) with the library. Afterwards buffer P was discarded and the membrane was washed with cold TPS. Fractionated Western blotting allowed for transfer of ClpB-NTD bound to peptide spots onto PVDF membranes and bound ClpB-NTD was detected by use of specific, polyclonal (rabbit) anti-ClpB-NTD serum.

4.4.10. ADDITIONAL STATISTICAL CALCULATIONS

Errorbars of proportion histograms (Fig 4.4C, S4.7A, and S4.10C) are calculated using the standard error of a binomial distribution:

$$\sigma = \sqrt{\frac{p(1-p)}{N}}$$

where p is the success proportion and N is the total number of observations.
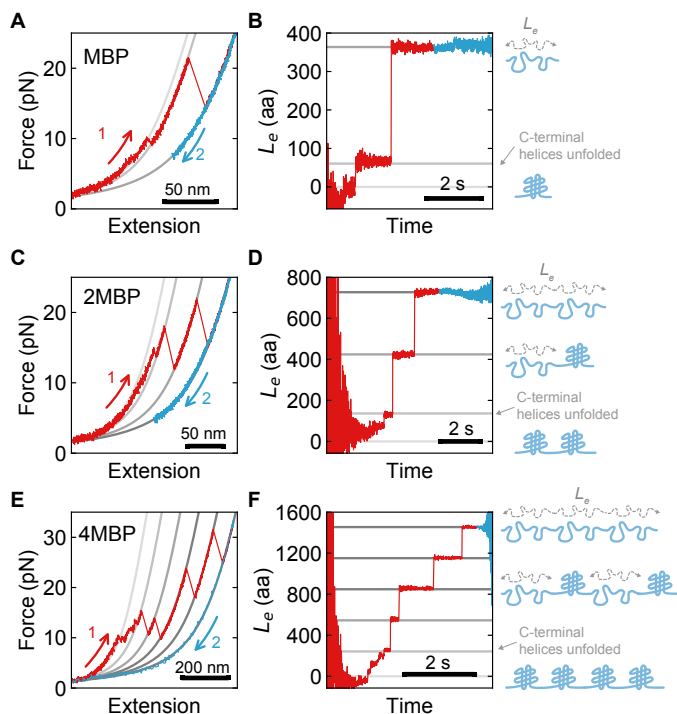
4.5. SUPPLEMENTARY FIGURES

Figure S4.1: Mechanical unfolding of substrates and extended length description. Force extension curves showing the characteristic unfolding pattern of **(A)** MBP, **(C)** the 2MBP and **(E)** the 4MBP construct, with first a gradual and discrete unfolding of C-terminal alpha-helices (fig. S9A) followed by a sharp unfolding of the core (two and four, for the 2MBP and 4MBP constructs, respectively). Gray lines show worm-like-chain fits to the data. Red indicates pulling while blue relaxing of the protein. **(B)**, **(D)** and **(F)** show the instantaneous extended length Le of the protein as it is unfolded (same traces as in the left panels). Gray lines correspond to the contour length values obtained from the fits. At low forces all WLCs of different contour lengths converge, yielding noisy data.
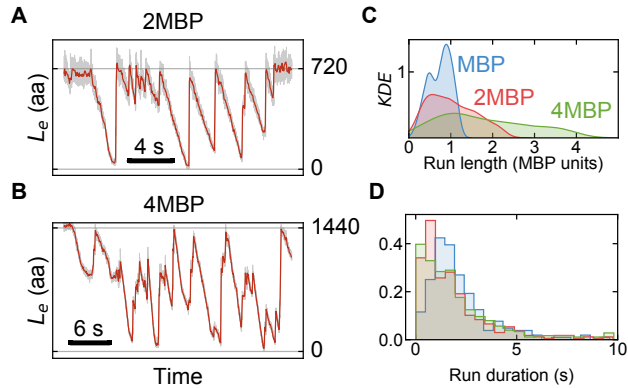
Figure S4.2: Larger constructs show longer translocation runs. Traces of protein extended length contractions in the presence of ClpB-Y503D for **(A)** 2MBP (L_c = 720 aa) and **(B)** 4MBP (L_c = 1440 aa). **(C)** Run length and **(D)** run duration (time between run start and back-slip) distribution for constructs of different lengths.
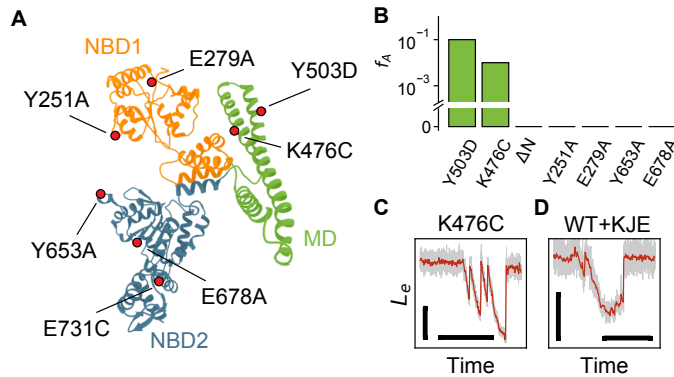

Figure S4.3: Translocation by ClpB variants. **(A)** Scheme of a ClpB monomer indicating all tested variants. These variants (except K467C) were generated in the constitutively active Y503D background. Variants E279A and E678A are Walker B mutants in the NBD1 and NBD2, respectively. These mutations abolish ATP hydrolysis at NBD1 or NBD2. Variants Y251A and Y653A are pore loops mutants in NBD1 and NBD2, respectively. These mutations affect substrate binding and translocation at either NBD1 or NBD2. The K476C variant undocks the MD, mimicking the effect of Hsp70 (DnaK) activation. MD undocking in the Y503D variant is more pronounced, and therefore activation more robust. An additional construct (ClpB-ΔN) lacked the NTD, hindering initial substrate binding. Finally, the variant E731C harbors a cysteine at the bottom of NBD2 for fluorophore labelling. **(B)** Fraction of time showing activity (fA) for different mutants (in Y503D background except K476C). **(C)** Translocation example for ClpB-K476C. Scale bars correspond to 200 aa and 10 s. **(D)** Translocation example for ClpB-WT with the DnaK system (DnaK, DnaJ, GrpE). Scale bars correspond to 200 aa and 5 s.
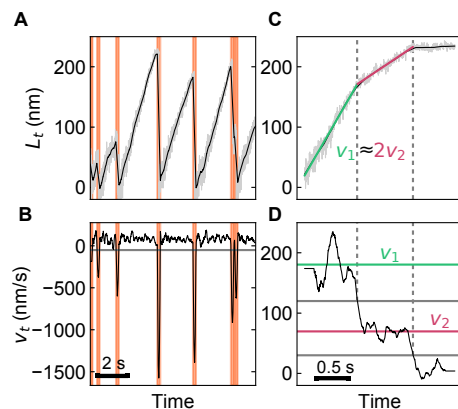
Figure S4.4: Speed characterization of translocation runs. (A) Translocated length (L_t) during threading of 2MBP by ClpB-Y503D. Raw data (light gray) is filtered using a Savitzky-Golay filter (black line). **(B)** Local translocation speed calculated as the time derivative of the translocated length after Savitzky-Golay filtering. Negative slopes below -50 nm/s (horizontal line) are considered backslipping events (green areas, also in e) and help determining isolated translocation runs. **(C)** Identification of different speeds within a single translocation run. Linear fits are used to calculate the speed of the run (green and magenta lines), most times revealing two main velocities, one double than the other. **(D)** Time derivative of the filtered translocated length for a single run, with solid black lines indicating the threshold speeds to distinguish no translocation from single and double translocation and green and magenta indicating the fitted velocity values (also shown in **C**).
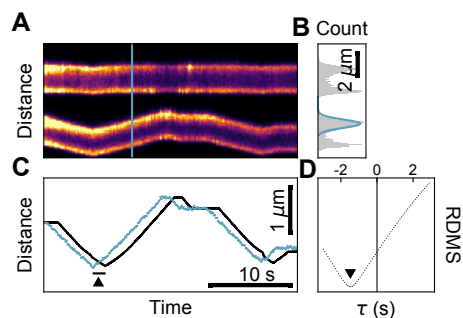

Figure S4.5: Synchronization of fluorescence and force spectroscopy signals. (A) Confocal scanning kymograph of two trapped beads. **(B)** Intensity profile of a scanning line (blue in **B**), with a Gaussian fit of the edge of the moving bead (bottom) in blue. **(C)** Offset between the fluorescence (blue dots, position of bead calculated for each scanning line as shown in B) and high-resolution position detector (black line, representing the actual movement of the bead) signals. **(D)** RDMS between both signals for different time shifts τ. The minimum is marked with a triangle and represents the best estimation of the offset between signals.
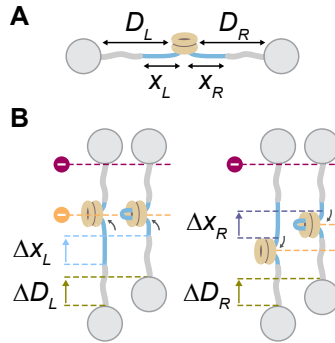
Figure S4.6: The force clamp and computation of the two length components. **(A)** Scheme of the lengths involved. **(B)** Response of the force clamp to translocation, and position changes of beads and fluorophore when the left or the right polypeptide arm is translocated.
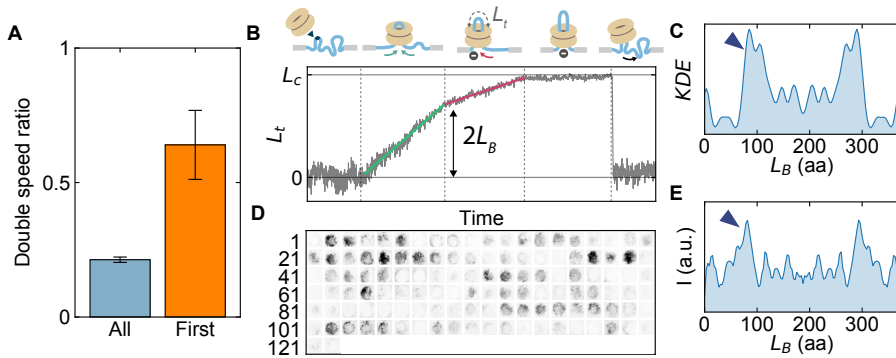

Figure S4.7: Initial binding of ClpB inferred from first runs confirms the model of a translocation mode switch. **(A)** Translocation speed for all runs (0.22, $N = 1704$) and for first runs only (0.64, $N = 30$). **(B)** Example of first translocation run. ClpB binds at a certain location within the polypeptide chain and starts translocating both strands (green) until it bumps into a DNA handle, switching to single strand translocation (red). If the second DNA handle is reached, translocation is stalled until backslipping occurs. The translocated length at which the first speed switch takes place is twice the distance between the firstly reached DNA handle and the initial binding position L_B. **(C)** KDE distribution of the inferred binding locations based on first runs ($N = 30$). The distribution is symmetric because N and C termini cannot be distinguished in our assay. **(D)** Peptide library data indicating regions of MBP which are bound by ClpB-NTD. ClpB-NTD spot intensities were quantified using a custom Python script **(D)**. For direct comparison with **(C)** the spot intensity distribution was mirrored.
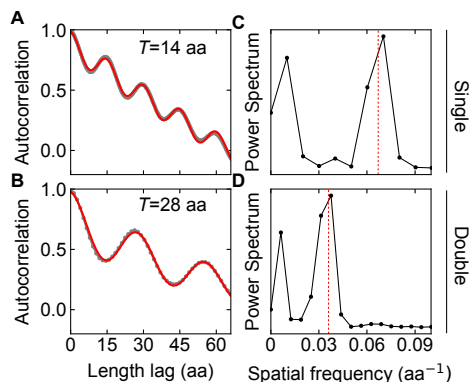
4

Figure S4.8: Step periodicity is confirmed by autocorrelation and power spectrum analysis.
(A), (B) Autocorrelation of the pairwise length distribution for single and double speed runs, respectively. The red line indicates a fit yielding period values of 2.3 and 4.4 nm. **(C), (D)** Power spectrum analysis of the pairwise length distribution, showing a peak at 0.42 and 0.22 nm^{-1}, in excellent agreement with the values obtained from the autocorrelation.
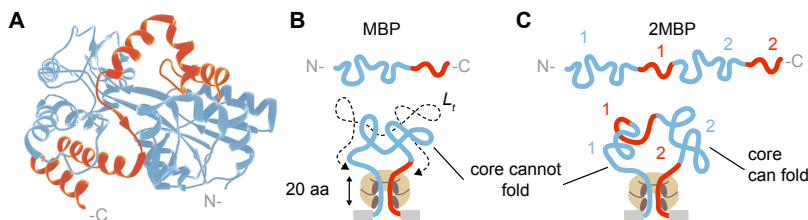
Figure S4.9: Substrate refolding after ClpB translocation. **(A)** Structure of MBP (PDB ID: 2MV0), showing in red the C-terminal helices (91 residues) not required for core folding [84] (blue). **(B)** Cartoon representation of the extended MBP chain showing the C-terminus domain in red. After complete translocation, a portion of the N- and C-termini will remain inside the ClpB pore, not available for tertiary structure. While the latter is not crucial for core formation, the former is. **(C)** Cartoon representation of the extended 2MBP. The C-terminal MBP core can form.
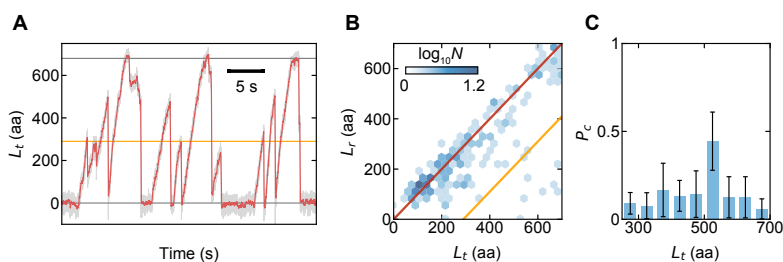
Figure S4.10: Introducing a double mutation in MBP inhibits its refolding after ClpB translocation. (A) Threading events for 2 double mutant MBP. Gray lines indicate 0 and 680 aa, while the orange line corresponds to 300 aa, the core length of MBP. **(B)** Most back-slipping events involve the release (Lr) of the whole polypeptide chain (red trend). **(C)** Probability of core folding (Pc) remains low regardless of how much chain is translocated ($N = 203$).

5

CATCH BONDS FORM DYNAMIC YET STRONG MATERIALS

A wide range of proteins form so called catch bonds, whose lifetime increases when subjected to moderate forces, in contrast to the monotonic lifetime decrease with force characterizing normal bonds. Cells are believed to exploit this remarkable property for different processes including migration and adhesion. Notably, many cytoskeletal cross-linkers are also catch bonds. However, their role in providing biological networks with their remarkable mechanical properties is not fully understood. In this chapter we use a novel single-molecule strategy to identify two variants of the actin cross-linker α-actinin 4 as an ideal slip-catch bond cross-linker pair. We next test their effect on the mechanical properties of an actin network using rheology. Surprisingly, we find that networks connected by catch bonds are stronger than slip bond networks even when the individual bonds are weaker. Our work provides insight into the widespread phenomenon of catch bonds in biology and offers a new design route towards life-like materials that combine adaptability with strength.

5.1. INTRODUCTION

Nature uses dynamic materials in order to combine mechanical resistance with the ability to adapt and self-heal. Biological soft matter is therefore made of transiently bonded interfaces and polymers [177–179]. Whereas synthetic transient networks readily fracture due to the inherent force sensitivity of dynamic bonds [180–182], biological networks are surprisingly strong [183–185]. How does biological matter achieve the ability to flow without risking mechanical failure [186, 187]? Remarkably, many biopolymer networks and virtually all cellular interfaces are crosslinked by catch bonds [188–190], proteins whose bond lifetime increases when moderate forces are applied via exposure of hidden binding sites [191]. After full exposure at high force, catch bonds behave like normal (slip) bonds whose lifetime monotonically decreases with force [191, 192]. Following this definition, catch bonds have a shorter or equal lifetime than the corresponding slip bonds, irrespective of the applied force (Fig. 5.1D).

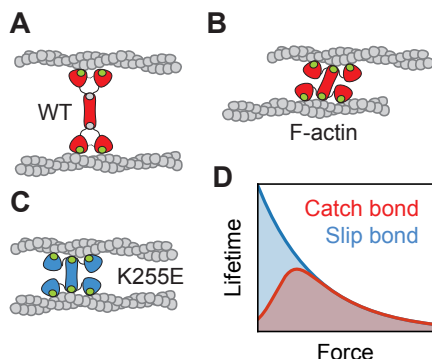
Figure 5.1: Catch-bond lifetime behaviour and ACTN4 scheme. **(A)** the actin-binding domains of α-actinin 4 present two exposed actin-binding sites, and one hidden that gets active under force **(B)**. **(C)** The mutant K255E has the three actin-binding sites exposed. **(D)** The lifetime of catch bonds presents a maximum at a non-zero force, in contrast to the exponential decrease of slip bonds.

Single-molecule experiments have demonstrated catch bond behaviour for a wide range of biological linkers [188–191]. Chief among them are actin-binding proteins, which crosslink actin filaments in the cytoskeleton of the cell [177], and transmembrane proteins known as cadherins [193] and integrins [194], which form cell-cell and cell-extracellular matrix adhesions. However, emergent phenomena in materials connected by catch bonds remain unexplored. Studying the impact of the catching nature on networks is not trivial, since many other bond features can be confounded. Therefore, it is key to choose a pair of crosslinkers whose only difference is their slip or catch nature, and thus can be directly compared. Prior structural evidence suggests that wild-type (WT) α-actinin 4 (ACTN4) and a mutant (K255E) may correspond to the catch/slip pair sought (Fig. 5.1B). WT ACTN4 has two weak actin-

binding sites that are always exposed, and one strong actin-binding site that is hidden [195]. The interaction between this third actin-binding site and actin is blocked by an intramolecular lysine-tryptophan interface, which is thought to be opened upon mechanical loading [196] (Fig. 5.1A,B). In contrast, it is constitutively active for the K255E point mutant as the intramolecular interface is destabilized (Fig. 5.1C) [197]. The wild type crosslinker and the K255E mutant could therefore serve as an ideal model system to experimentally test whether catch-bonding has a different impact on the network stability compared to crosslinking by a constitutively active linker (K255E mutant).
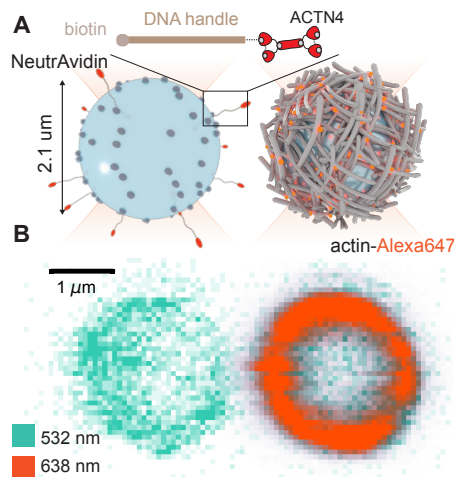

Figure 5.2: **Experimental scheme for the single-molecule ACTN4 characterization.** **(A)** Scheme of the experiments, with NeutrAvidin beads coated with actin filaments (right) or the ACTN4-DNA hybrid depicted on top (left). **(B)** Confocal scanning reveals the different bead identities. Green channel ($\lambda = 532$ nm) shows the autofluorescence of polystyrene beads, while the red channel ($\lambda = 638$ nm) indicates the presence of fluorescent actin on the surface of only one of the beads.

5.2. RESULTS

Studying the bond nature of ACTN4 with current single-molecule approaches is challenging, as they form very stable dimers and bind to actin filaments rather than single monomers. Therefore, we established a novel protocol to characterize the binding dynamics of the crosslinker. We engineered a construct consisting of ACTN4 covalently linked to a 2.5 kb DNA handle modified with biotin, via an SFP synthase mediated reaction as described in Chapter 2 (Fig. 5.2A). In order to ensure that only one of the monomers is coupled to DNA, we titrated the anchor concentration and analyzed the results with SDS-PAGE (Fig. S5.1). NeutrAvidin polystyrene beads were then prepared in two different ways: one sample was incubated with the ACTN4 construct and the other coated with biotinylated and Atto633-labeled actin filaments previously polymerized (Fig. 5.2A). We added 100 mM biotin to block any remaining free NeutrA-

vidin binding site. Beads were optically trapped in the optical tweezers and confocal fluorescence microscopy was used to verify the different identities of the beads. Polystyrene beads show intrinsic autofluorescence at 532 nm laser excitation, while the labeled actin filaments can be excited with 638 nm light (Fig. 5.2B).
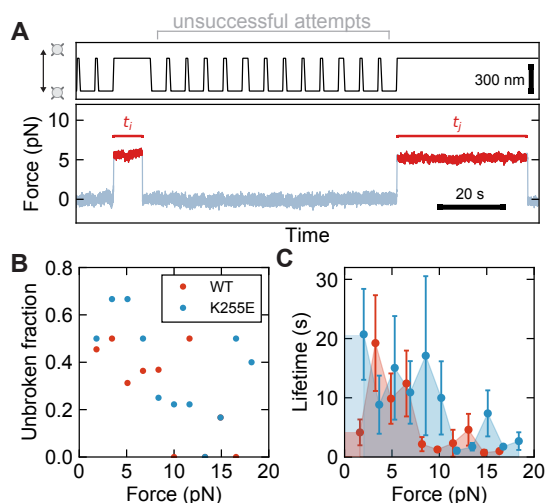

Figure 5.3: ACTN4 binding dynamics at the single-molecule level. (A) Approaching and retracting protocol to establish ACTN4-actin connections (top). An increase in the force while retracting indicates the presence of a tether, and the lifetime is measured until the instant the tether breaks (bottom). **(B)** Ratio of tethers that did not break after 1 minute. **(C)** Lifetime of breaking tethers at different forces.

Next, we started an approaching-retracting protocol in order to establish connections between ACTN4 and actin (Fig. 5.5A, top panel). In some occasions, retraction led to a marked increase in force (Fig. 5.5A), indicating the formation of a tether. The interpretation of such events is aided by the inclusion of DNA in the construct. First, it serves as a spacer, reducing surface effects and unspecific binding of actin to the other bead. Second, the DNA was specifically attached to ACTN4 close to the ABS3 domain, increasing the probability that the force exerted on ACTN4 by the optical tweezers is along the same direction as when ACTN4 is forced in stressed actin networks. Finally, its well-known mechanical properties can be used to discern between single and multiple tethers. In particular, individual DNA molecules show a characteristic overstretching plateau above 60 pN (Fig. S5.2). Once a tether was formed, it was classified as single or multiple by direct comparison with curves reaching DNA overstretching (Fig. S5.2), and its lifetime measured (Fig. 5.5A, bottom panel). Tether lifetimes ranged from subseconds to a few tens of seconds, and a portion of molecules did not break even after several minutes (Fig. 5.5B). We could not determined what exactly caused the later population, but unspecific interactions and cooperative binding of both monomers to actin are possible reasons. To estimate the lifetime, therefore, only tethers that broke before 1 minute were considered.

Wild-type ACTN4 showed a characteristic catch-bond lifetime distribution, peaking at around 6 pN. The distribution for the K255E mutant displayed more fluctuations, but the lifetime was overall higher than that of WT-ACTN4, with a clear clear increase at very low forces, consistent with a slip bond behaviour (Fig. 5.5C).

These results are in line with previous studies. In living cells, mechanical experiments have revealed that WT but not K255E ACTN4 strongly localizes to actin stress fibers [198] and externally stressed areas of the actin cortex [199, 200], and that its bound lifetime increases with the application of stress [199]. In vitro experiments have shown that this increase of lifetime is intrinsic to the ACTN4-actin interaction and not due to regulatory pathways, as a similar increase of bound lifetime upon application of mechanical stress was observed in a reconstituted network consisting of only actin filaments and ACTN4 cross linkers [183]. In contrast, the K255E mutant that has a constitutively exposed ABS1 domain was shown not to accumulate in stressed regions nor to increase its bound lifetime upon network stress [183, 198], suggesting that it is the exposure of the ABS1 domain upon force that is responsible for catch bonding. X-ray crystallography has identified ABS-1 as a hidden actin binding site [201] and MD-simulations have shown that this hidden binding site is exposed upon a K255E point mutation [195]. Together, these observations strongly support the notion that the WT ACTN4 requires force to be fully active whereas the K255E mutant is constitutively active.
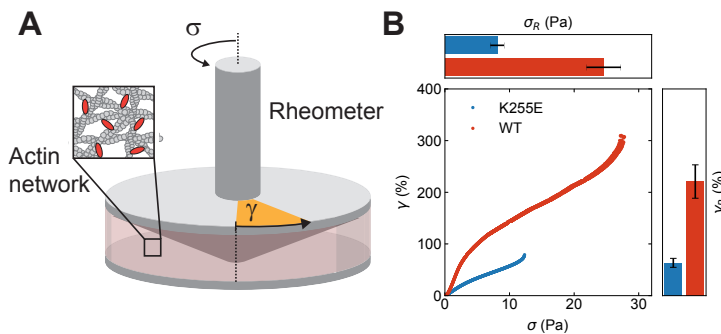
Figure 5.4: Trade-off between deformability and strength of transiently crosslinked actin networks. (A) Scheme of the rheometer used to characterize the crosslinked actin network. **(B)** The shear strain as a function of the shear stress. At the rupture stress, both quantities diverge. We define the rupture strain as the last data point where K' exceeds the linear storage modulus (Fig. S5.3. The top panel shows the average rupture stress and the right-most panel the rupture strain, with error bars representing the standard error on basis of 4 repeats per condition. All experiments are performed at a stress rate of 2.0 mPa/s.

After establishing ACTN4 WT and the K255E mutant as an ideal model system, we tested the impact of their different bond nature on actin networks using rheology (Fig. 5.4). In particular, we determined the rupture stress of crosslinked actin networks by linearly increasing the shear stress in time until the network fractured. During the stress ramp, we applied a superimposed small oscillatory stress to measure the

differential storage modulus, which represents the stiffness of the network. We find that the differential storage modulus is constant for small stresses, but increases for larger stresses as the thermal undulations of the filaments are pulled taut (Fig. S5.3). This network stiffening behavior is characteristic for semiflexible polymers such as actin [202]. At high stress, the differential storage modulus precipitously drops to zero (Fig. S5.3), indicating rupture. Strikingly, we find that WT crosslinked networks fracture at 27 ± 3 Pa, whereas K255E networks are significantly weaker with a rupture stress of 9 ± 1 Pa (Fig. 5.4B). In addition, we also found that WT connected networks are very ductile, with a rupture strain of $221 \pm 16\%$ as compared to the $63 \pm 4\%$ obtained for the K255E network (Fig. 5.4B).

5.3. DISCUSSION

How can catch bonds provide more long-lived networks than slip bonds, even though they are more short-lived at the single-molecule level? The reduced lifetime of ACTN4-WT at lower forces also implies a higher mobility throughout the network. Therefore, cross-linkers will more readily unbind and diffuse from areas where stress is low, towards highly stressed areas, providing a self-assembly mechanism against force inhomogeneity that can prevent rupture and increase the network lifetime. Future computational simulations can provide further support to this model.
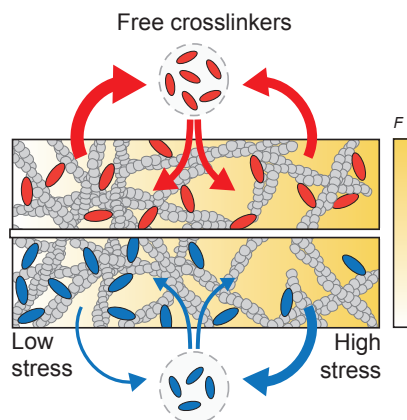
Figure 5.5: Mechanism of network strengthening by catch bonds. Catch bonding crosslinkers (red) can unbind more readily from low stress areas in the network and relocate randomly, effectively reinforcing high stress areas. In contrast, slip bonds lifetime is lower in high stress areas, and will therefore tend to accumulate at low stress areas.

Catch bonds are present in virtually any system of cellular adhesion, and in many biopolymer networks including the cytoskeleton. In all these cases, mechanical force is distributed over many bonds that can diffuse after unbinding. The functional relevance of catch bonds is often related only to mechanosensing, where the force-induced conformational transition could serve as a read-out for force [200, 203, 204]. Our results imply yet an additional possibility, namely that catch bonds evolved to

prevent fracturing. Indeed, the point mutation (K255E) in the actin crosslinker α-actinin-4 converts a catch bond into a constitutively active slip bond, which results in a disease known as focal segmental glomerulosclerosis type 1 that is associated with brittle cells and kidney failure [183, 196].

Outside biology, catch bonds provide a potential new biomimetic design principle towards smart materials. Typically, the addition of permanent or very stable bonds improves the strength of materials, but this limits the deformability [185, 205, 206]. However, for many applications such us regenerative medicine [207] both features are required. Our findings show that catch bonds uniquely combine the advantages of enhanced deformability and strength of materials.

5

5.4. MATERIAL AND METHODS

5.4.1. PROTEIN PURIFICATION

The actin crosslinker human α-actinin 4 (the wild type variant and its mutant counterpart K255E) was purified as described in reference [208]: *E. coli* cells were transformed to express recombinant crosslinkers with a 6xhis-tag. Induction was performed with 500 μM IDTA for eight hours at 25°C. After centrifugation at 6000 g for 15 minutes, cells were resuspended in 20 mM NaCl, 5 mg/ml lysozyme and 20 mM Hepes, pH 7.8. The cells were lysed by a freeze-thaw cycle, and centrifuged at 2,000 g for 30 min. The recombinant proteins were purified from the supernatant using a QIAGEN nickel column. Next, the column was washed with 20-bed columns of 500 mM NaCl, 25 mM imidazole, and 20 mM Hepes, pH 7.8. The recombinant proteins were eluted with 10-bed volumes of 500 mM NaCl, 500 mM imidazole, and 20 mM Hepes, pH 7.8. The proteins were concentrated using Centricon (Millipore) and purified by gel filtration in 150 mM NaCl, 20 mM Hepes pH 7.8, and 10 mM dithiothreitol (DTT). To ensure we use the same concentration of the WT and K255E ACTN4, we determined the ratio of stock concentrations by measuring the band intensity of both types of crosslinkers on a SDS-PAGE gel. We chose this method over determining the protein concentration by the absorbance of the stock solution at 280 nm, as with the SDS-PAGE gel we specifically measure the protein of interest and exclude the contribution of contaminants.

Actin was purified from rabbit psoas skeletal muscle as described in reference [209]. The concentration was determined by measuring the absorbance at 280 nm. Aliquots were stored at -80°C in G-buffer (2 mM tris-hydrochloride pH 8.0, 0.2 mM disodium adenosine triphosphate, 0.2 mM calcium chloride, 0.2 mM dithiothreitol) to prevent polymerization. After thawing, we store G-actin stock samples overnight at 4°C. The next day, we spin the sample at 120000 g to spin down any aggregates. We store the supernatant at 4°C for maximum of 7 days before use. Unless otherwise mentioned, we used a concentration of 48 μM, corresponding to 2 mg/ml, for all our experiments, and actin was polymerized in an F-buffer consisting of 50 mM KCl, 20 mM imidazole pH 7.4, 2 mM MgCl$_2$, 1 mM DTT and 0.5 mM MgATP. Unless otherwise mentioned, we used a crosslinker concentration of 0.48 μM to obtain a molar ratio of 1/100 crosslinker/actin and on average around 1 crosslinker per 0.5 μm of actin filament; under these conditions, networks are isotropic.

5.4.2. SINGLE-MOLECULE CONSTRUCTS GENERATION

Wild-type and K255E mutant were modified to include a ybbR tag (DSLEFIASKLA) right after the His-tag. Purified proteins were coupled to 20 nt CoA-modified DNA oligonucleotides using a SFP-mediated reaction [119]. A protein to DNA ratio of 10:1 ensured that coupling mostly occurred only on one monomer, as evidenced by SDS-PAGE analysis (Fig. S5.1). Next, 2.5 kbp DNA tethers were generated as described in Chapter 2 and ligated to the DNA-protein hybrid using T4 ligase (New England Biolabs) overnight at room temperature.

5.4.3. Single-molecule data analysis

Data was collected at 500 Hz using a custom-built dual trap optical tweezers and a commercial C-Trap (Lumicks). Tethers were classified as single if the extension-force characteristics were similar to molecules that showed DNA overstretching and only if they ruptured in a clean single step. The theoretical contour length of the construct is 850 nm, but the polydispersity of bead radii and the uneven distribution of the actin layer lead to uncertainties in the apparent extension. We shifted the extension so that overstretching curves matched a worm-like chain of 850 nm contour length. Then we considered an additional uncertainty margin of 50 nm and every point falling whithin this range was assigned as a single tether (Fig. S5.2).

5.4.4. Rheology

Rheology was performed using a stress-controlled Kinexus Malvern Pro rheometer with a stainless steel 20 mm radius cone plate geometry with a 1 degree angle. We loaded 40 μl of the ACTN4 cross linked actin networks on the rheometer plate at 10 °C directly after mixing the proteins into the polymerization buffer. A thin layer of Fluka mineral oil Type A was added around the geometry to prevent evaporation, and the sample was closed off with a hood to minimize effects of air flow. Polymerization of the network was followed by applying a small oscillatory shear with a strain amplitude of 0.5% and a frequency of 0.5 Hz. Then, we perform a fracturing experiment by linearly increasing the stress in time (2 mPa/s), until the network fractures. We define the rupture point as the last data point at which the differential storage modulus exceeds the linear modulus [Fig. 9a].

5

5.5. SUPPLEMENTARY FIGURES
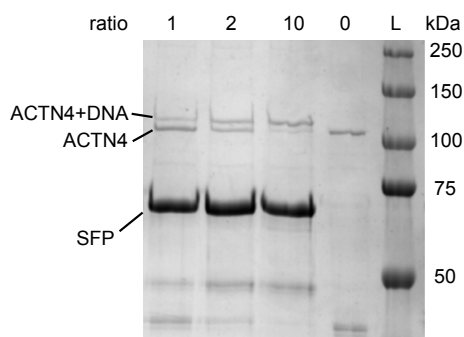
Figure S5.1: ACTN4-DNA coupling.   DNA was coupled to ACTN4 to prepare the single molecule constructs. We performed coupling reactions at several different molar ratios of DNA:ACTN4. These samples were run on a SDS-PAGE gel to measure the degree of coupling. The DNA:ACTN4 molar ratios is indicated on top of each band. We find that the fraction of coupled ACTN4 monomers increases with the amount of DNA added to the solution. At a molar ratio of 1:1, most of the ACTN4 is uncoupled.
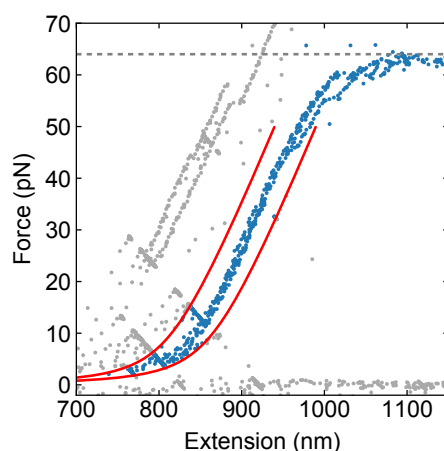
Figure S5.2: Classification criterion for single tethers. Force-extension curves revealed a force plateau between 63 and 65 pN that is characteristic of dsDNA [224] (dashed line). This was used as the criterion to classify tethers as single connections. Due to the uncertainty on bead radii and actin layer thickness, we generated two WLC models with contour lengths 25 nm shorter and longer than the theoretical 850 nm (red lines). Data lying within this range that broke in a clean step (and showed DNA overstretching when reaching 65 pN if pulled) was regarded as single tethers (blue dots), otherwise the data were discarded (grey dots).
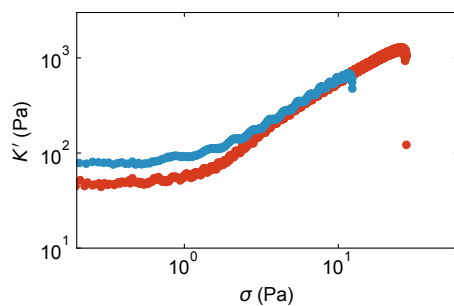
Figure S5.3: Network rupture. Representative example curves of the differential storage modulus measured at 0.5 Hz is plotted against the applied shear stress for actin networks crosslinked by ACTN4 WT (red) and K255E (blue). We define the rupture strain as the last data point where K' exceeds the linear storage modulus.

6

CONCLUSION AND OUTLOOK

In this final Chapter we briefly summarize the main discoveries included in this thesis. We also propose a series of potential applications using the approach described in Chapter 2. This includes technical innovations that can be employed to study other challenging aspects of protein folding and chaperone action. Moreover, we also present a series of preliminary experiments that complement the results of the thesis while setting the path to new studies that will provide additional important insights into the molecular aspects of chaperone-mediated protein rescue.

In this thesis we have presented a series of discoveries that provide important insight into the mechanistic basis of several chaperone systems. The studies are enabled by a new powerful strategy that combines single-molecule force spectroscopy with confocal imaging, and uses a novel protein tethering strategy. As a demonstration of the approach we first investigated how the binding dynamics of **Trigger Factor** depend on and affect substrate conformations. We directly show for the first time Trigger Factor binding to unfolded chains, for up to tens of seconds, and prevents refolding while bound. Next, we addressed the role of **GroEL-GroES** in accelerating protein folding. By following individual substrates in time, we confirmed that the chaperone can promote folding beyond mere aggregation prevention. Moreover, we observed that the collapse of the polypeptide chains was enhanced by the GroEL cavity, suggesting a new underlying mechanism that involves attractive forces. Finally, we studied how the **ClpB** disaggregase processes substrates. We observed that polypeptide *loops* are processively extruded through its central pore, at strikingly high speeds and exerting large pulling forces. Remarkably, ClpB can translocate both arms of the loop simultaneously, and switch to single-arm translocation when encountering obstacles, thus avoiding pore jamming. Moreover, substrates can refold while emerging from the ClpB channel, similarly to ribosomal co-translation folding, thus reducing the risk of re-aggregation.

These discoveries help improving our understanding of how chaperones achieve their critical tasks. Here we propose a series of potential questions that can be readily addressed with our approach.

Studying other chaperone systems.   In addition to the brief Trigger Factor example described in Chapter 2 and the GroEL and ClpB studies of Chapters 3 and 4 respectively, our robust tethering approach can help elucidate the mechanism underlying many other chaperone systems. This includes the more specific and technically challenging HdeA/B and Spy chaperones, for acid and solvent stress respectively. The extreme conditions necessary to study their physiological roles require resistant constructs, which our approach provides. Antigen-antibody linkages would not be suitable for bead attachment in this case, but a range of other more resistant coupling approaches are possible [122]. One solution consists on using very long (5 kbp) biotinylated tethers at both sides of the protein and adding them with fast flow to trapped beads. Upon binding to one bead, drag forces on the DNA exert a torque that rotates the bead and allows to create a connection with the other, as previously demonstrated for DNA alone [111]. Other exciting applications include the study of protein oligomerization, ribosome assembly or intrinsically disordered protein networks.

Monitoring folding when it is inaccessible to force sensing.   The assessment of substrate refolding in Chapters 3 and 4 relies, with a few exceptions (Fig. 3.1E), on the force sensing of states *posterior* to the actual folding event. For GroEL, the force-extension curve when pulling indicated the formation of tertiary structure during the waiting time at zero force. For ClpB, the dynamics of back-slipping contained information about the formation of folded domains too big to enter the central pore. However, several details remain unknown with these indirect approaches. When *ex-*
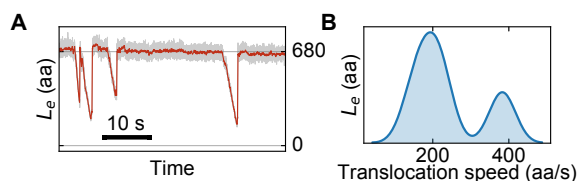
Figure 6.1: Translocation by the yeast Hsp104. **(A)** Translocation runs on 2MBP by the unrepressed Hsp104 variant Y508D. **(B)** Translocation speed distribution for the same variant ($N = 7$). The peaks are very close to those of ClpB.

actly is folding happening? Is the substrate not only forming tertiary structure, but gaining *activity* again? The dual monitoring strategy here described, combined with carefully selected substrates, can help elucidate these questions. Fluorescent proteins like YPet (used in Chapter 2 for this purpose) whose activity can be monitored by detecting their emission, or substrates that only bind their (fluorescently labelled) ligands when native, such as MBP binding of maltose, constitute suitable examples.

Folding-assistance mechanism of the human chaperonin TRiC. In Chapter 3 we have proposed an enhanced-collapse mechanism to accelerate folding that is exploited by the GroEL-GroES system. We have also opened the door to the possibility that such mechanism is more general and also employed by other chaperones. In particular, the human chaperonin TRiC is a strong candidate, as it is regarded as the human homologue of GroEL. The most striking difference in TRiC is the absence of a lid, replaced with retractile loops that can close the access to the cavity [210]. The chaperone is required for the proper folding of vital proteins, including actin and tubulin [211]. The approach presented here can readily be used to study the mechanism TRiC employs to accomplish its function.

Substrate processing by the Hsp104 disaggregase. We have proposed that processive translocation of polypeptide loops is a general mechanism exploited by all Hsp100 disaggregases. In fact, preliminary experiments on the yeast disaggregation chaperone Hsp104 show that it is able to translocate polypeptide loops in a very similar fashion to ClpB (Figure 6.1A), which is consistent with their high degree of structural homology. The translocation speed displayed by two different unrepressed mutants (K538C and Y508D) was almost identical to ClpB (Fig. 6.1B), and it also displayed switches from double-arm to single-arm translocation. The run initiation propensity, however, was lower for Hsp104, most likely due to differences in the non-conserved NTD, which is involved in substrate binding. Crystal structures show a number of dissimilarities between Hsp104 and ClpB. Studying the steps for Hsp104 could help elucidate whether the firing mechanism and subunit coordination is also conserved between all Hsp100 disaggregases.

Impact of obstacles on ClpB translocation. We also show in Chapter 4 that ClpB is able to switch from two- to single-arm translocation when encountering obstacles, namely the DNA tethers. But how does ClpB deal with other sorts of obstacles that

may be more relevant during disaggregation? If ClpB is not able to clear them rapidly enough, the consequent pauses in translocation would result in critical delays that can increase the risk of re-aggregation discussed in Chapter 4 (Fig. 4.5).

Here we propose an assay to address this question using labelled DnaJ as obstacles. This co-chaperone is part of the Hsp70 system in bacteria, and known to act as a holdase. Other chaperones are possible, including Hsp70 itself or SecB. First, the un-folded chain is immersed in a solution containing a high concentration of labelled DnaJ to ensure multiple binding. Next, the construct decorated with obstacles is moved to a region containing ClpB, while monitoring the total fluorescence on the chain (Fig. 6.2). Upon ClpB binding, it would be possible to correlate translocation pauses and restarts with decreases in the total fluorescence. The pause indicates that ClpB has encountered a roadblock, and the fluorescent drop and restart of the run shows that ClpB has been able to clear it (Fig. 6.2). Other obstacles that are present in aggregates include folded domains [167], and aggregated structures, which can be readily studied. Unfolding one of the subunits in a 2MBP construct allows initial ClpB engaging, which will translocate the unfolded polypeptide chain until it encounters the DNA tether and/or the folded MBP. Studying aggregates is technically more challenging, and explained in the next section.
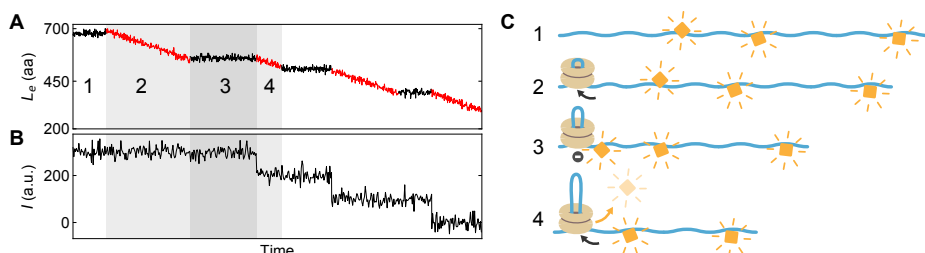
Figure 6.2: Translocation by ClpB could be slowed down by obstacles. (A) Simulated trace of translocation by ClpB in the presence of obstacles. Every time a roadblock is encountered, a pause in translocation occurs until the obstacle is cleared by ClpB. **(B)** Simulated fluorescence signal from the labelled obstacles. Discrete decreases in the total emission indicate the removal of roadblocks. **(C)** Scheme of the obstacle clearance.

Monitoring disaggregation at the single-molecule level. While seemingly paradox-ical, it is possible to form small protein aggregates using a 4MBP construct within the optical tweezers [93]. Unfolding a series of consecutive proteins and rapidly bring-ing the extended polypeptide chain together will result, with certain probability, in an aggregate-like structure. Such state is distinguished from native folding by its in-creased resistance to high forces and an erratic unfolding pattern. This mechanical characterization imposes a limitation on which kind of aggregates can be studied. On the one hand, very stable aggregates can be easily classified by force sampling, since they can handle high tensions. However, many chaperones are not capable of dealing with such tight structures. On the other hand, weak aggregates are more suitable for disaggregation assays, but will not resist the force test. Reducing the force threshold

for classification as an aggregate is not possible, since this would render them undistinguishable from native structures.

To overcome this limitation, one can introduce the two mutations into MBP that dramatically slow down folding [84] (Fig. 3.1E). In the optical tweezers assay, this variant presents a similar first unfolding pattern as the wild-type MBP, but in contrast rarely refolds. In fact, the refolding probability of a single dmMBP is below 10% (Fig. 3.1E), and therefore the formation of native conformations in several subunits simultaneously is extremely unlikely. Thus, one can safely assume that any compact structure will correspond to an aggregate-like state, without the need to test its resistance to high forces. Preliminary experiments show that such structures are formed upon unfolding and relaxation of a 2 dmMBP construct.

A similar approach might be employed to study the unique ability of yeast Hsp104 to dissolve amyloid aggregates [212]. Similar to the ones formed with MBP, it would be possible to generate such aggregates using prion-like proteins, such as α-synuclein [213].

6

BIBLIOGRAPHY

[1] J. C. Young, V. R. Agashe, K. Siegers, and F. U. Hartl, *Pathways of chaperone-mediated protein folding in the cytosol.* Nature reviews. Molecular cell biology **5**, 781 (2004).

[2] B. Bukau, J. Weissman, and A. Horwich, *Molecular Chaperones and Protein Quality Control,* Cell **125**, 443 (2006).

[3] K. Liberek, A. Lewandowska, and S. Zietkiewicz, *Chaperones in control of protein disaggregation.* The EMBO journal **27**, 328 (2008).

[4] L. Meunier, *A Subset of Chaperones and Folding Enzymes Form Multiprotein Complexes in Endoplasmic Reticulum to Bind Nascent Proteins,* Molecular Biology of the Cell **13**, 4456 (2002).

[5] W. B. Pratt and D. O. Toft, *Regulation of Signaling Protein Function and Trafficking by the hsp90/hsp70-Based Chaperone Machinery,* Experimental Biology and Medicine **228**, 111 (2003).

[6] P. J. Muchowski and J. L. Wacker, *Modulation of neurodegeneration by molecular chaperones,* Nature Reviews Neuroscience **6**, 11 (2005).

[7] L. Whitesell and S. L. Lindquist, *HSP90 and the chaperoning of cancer.* Nature reviews. Cancer **5**, 761 (2005).

[8] A.-P. Arrigo, *Hsp27: Novel Regulator of Intracellular Redox State,* IUBMB Life (International Union of Biochemistry and Molecular Biology: Life) **52**, 303 (2001).

[9] R. Kern, A. Malki, J. Abdallah, J. Tagourti, and G. Richarme, *Escherichia coli HdeB is an acid stress chaperone,* Journal of Bacteriology **189**, 603 (2007).

[10] A. Tissieres, H. K. Mitchell, and U. M. Tracy, *Protein synthesis in salivary glands of Drosophila melanogaster: Relation to chromosome puffs,* Journal of Molecular Biology **84**, 389 (1974).

[11] P. Clark, *Protein folding in the cell: reshaping the folding funnel,* Trends in Biochemical Sciences **29**, 527 (2004).

[12] E. Oh, A. H. Becker, A. Sandikci, D. Huber, R. Chaba, F. Gloge, R. J. Nichols, A. Typas, C. A. Gross, G. Kramer, J. S. Weissman, and B. Bukau, *Selective Ribosome Profiling Reveals the Cotranslational Chaperone Action of Trigger Factor In Vivo,* Cell **147**, 1295 (2011).

[13] L. Ferbitz, T. Maier, H. Patzelt, B. Bukau, E. Deuerling, and N. Ban, *Trigger factor in complex with the ribosome forms a molecular cradle for nascent proteins,* Nature **431**, 590 (2004).

[14] C. M. Kaiser, H.-C. Chang, V. R. Agashe, S. K. Lakshmipathy, S. A. Etchells, M. Hayer-Hartl, F. U. Hartl, and J. M. Barral, *Real-time observation of trigger factor function on translating ribosomes,* Nature **444**, 455 (2006).

[15] E. Martinez-Hackert and W. A. Hendrickson, *Promiscuous Substrate Recognition in Folding and Assembly Activities of the Trigger Factor Chaperone,* Cell **138**, 923 (2009).

[16] E. Deuerling, A. Schulze-Specking, T. Tomoyasu, A. Mogk, and B. Bukau, *Trigger factor and DnaK cooperate in folding of newly synthesized proteins,* Nature **400**, 693 (1999).

[17] S. Vorderwulbecke, G. Kramer, F. Merz, T. Kurz, T. Rauch, B. Zachmann-Brand, B. Bukau, and E. Deuerling, *Low temperature or GroEL/ES overproduction permits growth of Escherichia coli cells lacking trigger factor and DnaK,* FEBS Letters **559**, 181 (2004).

[18] R. S. Ullers, D. Ang, F. Schwager, C. Georgopoulos, and P. Genevaux, *Trigger Factor can antagonize both SecB and DnaK/DnaJ chaperone functions in Escherichia coli.* Proceedings of the National Academy of Sciences of the United States of America **104**, 3101 (2007).

[19] T. Saio, X. Guan, P. Rossi, A. Economou, and C. G. Kalodimos, *Structural Basis for Protein Antiaggregation Activity of the Trigger Factor Chaperone,* Science **344**, 1250494 (2014).

[20] K. Singhal, J. Vreede, A. Mashaghi, S. J. Tans, and P. G. Bolhuis, *The Trigger Factor Chaperone Encapsulates and Stabilizes Partial Folds of Substrate Proteins,* PLOS Computational Biology **11**, e1004444 (2015).

[21] 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).

[22] K. Singhal, J. Vreede, A. Mashaghi, S. J. Tans, and P. G. Bolhuis, *Hydrophobic Collapse of Trigger Factor Monomer in Solution,* PLoS ONE **8**, e59683 (2013).

[23] V. R. Agashe, S. Guha, H.-C. Chang, P. Genevaux, M. Hayer-Hartl, M. Stemp, C. Georgopoulos, F. Hartl, and J. M. Barral, *Function of Trigger Factor and DnaK in Multidomain Protein Folding,* Cell **117**, 199 (2004).

[24] A. Hoffmann, A. H. Becker, B. Zachmann-Brand, E. Deuerling, B. Bukau, and G. Kramer, *Concerted Action of the Ribosome and the Associated Chaperone Trigger Factor Confines Nascent Polypeptide Folding,* Molecular Cell **48**, 63 (2012).

[25] E. P. O'Brien, J. Christodoulou, M. Vendruscolo, and C. M. Dobson, *Trigger Factor Slows Co-translational Folding through Kinetic Trapping while Sterically Protecting the Nascent Chain from Aberrant Cytosolic Interactions,* Journal of the American Chemical Society **134**, 10920 (2012).

[26] C. M. Kaiser, D. H. Goldman, J. D. Chodera, I. Tinoco, and C. Bustamante, *The Ribosome Modulates Nascent Protein Folding,* Science **334**, 1723 (2011).

[27] C. Huang, P. Rossi, T. Saio, and C. G. Kalodimos, *Structural basis for the antifolding activity of a molecular chaperone,* Nature **537**, 202 (2016).

[28] 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 (New York, N.Y.) **318**, 1458 (2007).

[29] J. Deeng, K. Chan, E. van der Sluis, O. Berninghausen, W. Han, J. Gumbart, K. Schulten, B. Beatrix, and R. Beckmann, *Dynamic Behavior of Trigger Factor on the Ribosome,* Journal of Molecular Biology **428**, 3588 (2016).

[30] S. Preissler and E. Deuerling, *Ribosome-associated chaperones as key players in proteostasis,* Trends in Biochemical Sciences **37**, 274 (2012).

[31] F. Baneyx and M. Mujacic, *Recombinant protein folding and misfolding in Escherichia coli,* Nature Biotechnology **22**, 1399 (2004).

[32] M. P. Mayer and B. Bukau, *Hsp70 chaperones: cellular functions and molecular mechanism.* Cellular and molecular life sciences : CMLS **62**, 670 (2005).

[33] 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 (New York, N.Y.) **272**, 1606 (1996).

[34] R. C. Morshauser, W. Hu, H. Wang, Y. Pang, G. C. Flynn, and E. R. Zuiderweg, *High-resolution solution structure of the 18 kDa substrate-binding domain of the mammalian chaperone protein Hsc70.* Journal of molecular biology **289**, 1387 (1999).

[35] R. Sousa, *A Dancer Caught Midstep: The Structure of ATP-Bound Hsp70,* Molecular Cell **48**, 821 (2012).

[36] D. Schmid, a. Baici, H. Gehring, and P. Christen, *Kinetics of molecular chaperone action.* Science (New York, N.Y.) **263**, 971 (1994).

[37] J. S. McCarty, A. Buchberger, J. Reinstein, and B. Bukau, *The Role of ATP in the Functional Cycle of the DnaK Chaperone System,* Journal of Molecular Biology **249**, 126 (1995).

[38] T. Laufen, M. P. Mayer, C. Beisel, D. Klostermeier, A. Mogk, J. Reinstein, and B. Bukau, *Mechanism of regulation of Hsp70 chaperones by DnaJ cochaperones,* Proceedings of the National Academy of Sciences **96**, 5452 (1999).

[39] K. Mapa, M. Sikor, V. Kudryavtsev, K. Waegemann, S. Kalinin, C. a. M. Seidel, W. Neupert, D. C. Lamb, and D. Mokranjac, *The conformational dynamics of the mitochondrial Hsp70 chaperone.* Molecular cell **38**, 89 (2010).

[40] R. Kellner, H. Hofmann, A. Barducci, B. Wunderlich, D. Nettels, and B. Schuler, *Single-molecule spectroscopy reveals chaperone-mediated expansion of substrate protein.* Proceedings of the National Academy of Sciences of the United States of America **111**, 13355 (2014).

[41] A. Mashaghi, S. Bezrukavnikov, D. P. Minde, A. S. Wentink, R. Kityk, B. Zachmann-Brand, M. P. Mayer, G. Kramer, B. Bukau, and S. J. Tans, *Alternative modes of client binding enable functional plasticity of Hsp70,* Nature , 1 (2016).

[42] 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.* Nature structural & molecular biology **18**, 150 (2011).

[43] R. Banerjee, G. G. Jayaraj, J. J. Peter, V. Kumar, and K. Mapa, *Monitoring conformational heterogeneity of the lid of DnaK substrate binding domain during its chaperone cycle.* The FEBS journal **283**, 2853 (2016).

[44] J. C. Young, *Hsp90: a specialized but essential protein-folding tool,* The Journal of Cell Biology **154**, 267 (2001).

[45] B.-t. Lai, W. N. W. Chin, A. E. Stanek, W. Keh, and K. W. Lanks, *Quantitation and Intracellular Localization of the 85K Heat Shock Protein by Using Monoclonal and Polyclonal Antibodies,* **4**, 2802 (1984).

[46] K. Nadeau, M. a. Sullivan, M. Bradley, D. M. Engman, and C. T. Walsh, *83-kilodalton heat shock proteins of trypanosomes are potent peptide-stimulated ATPases.* Protein science : a publication of the Protein Society **1**, 970 (1992).

[47] T. Scheibel, S. Neuhofen, T. Weikl, C. Mayr, J. Reinstein, P. D. Vogel, and J. Buchner, *ATP-binding properties of human Hsp90,* Journal of Biological Chemistry **272**, 18608 (1997).

[48] C. Prodromou, S. M. Roe, R. O'Brien, J. E. Ladbury, P. W. Piper, and L. H. Pearl, *Identification and structural characterization of the ATP/ADP-binding site in the Hsp90 molecular chaperone,* Cell **90**, 65 (1997).

[49] M. M. U. Ali, S. M. Roe, C. K. Vaughan, P. Meyer, B. Panaretou, P. W. Piper, C. Prodromou, and L. H. Pearl, *Crystal structure of an Hsp90-nucleotide-p23/Sba1 closed chaperone complex.* Nature **440**, 1013 (2006).

[50] A. K. Shiau, S. F. Harris, D. R. Southworth, and D. A. Agard, *Structural Analysis of E. coli hsp90 Reveals Dramatic Nucleotide-Dependent Conformational Rearrangements,* Cell **127**, 329 (2006).

[51] T. Weikl, P. Muschler, K. Richter, T. Veit, J. Reinstein, and J. Buchner, *C-terminal regions of Hsp90 are important for trapping the nucleotide during the ATPase cycle,* Journal of molecular biology **303**, 583 (2000).

[52] P. Bron, E. Giudice, J.-P. Rolland, R. M. Buey, P. Barbier, J. F. Díaz, V. Peyrot, D. Thomas, and C. Garnier, *Apo-Hsp90 coexists in two open conformational states in solution.* Biology of the cell / under the auspices of the European Cell Biology Organization **100**, 413 (2008).

[53] L. H. Pearl and C. Prodromou, *Structure and mechanism of the Hsp90 molecular chaperone machinery.* Annual review of biochemistry **75**, 271 (2006).

[54] C. Ratzke, F. Berkemeier, and T. Hugel, *Heat shock protein 90's mechanochemical cycle is dominated by thermal fluctuations,* Proceedings of the National Academy of Sciences **109**, 161 (2012).

[55] M. Mickler, M. Hessling, C. Ratzke, J. Buchner, and T. Hugel, *The large conformational changes of Hsp90 are only weakly coupled to ATP hydrolysis.* Nature structural & molecular biology **16**, 281 (2009).

[56] C. Ratzke, M. N. T. Nguyen, M. P. Mayer, and T. Hugel, *From a ratchet mechanism to random fluctuations evolution of Hsp90's mechanochemical cycle,* Journal of Molecular Biology **423**, 462 (2012).

[57] C. Ratzke, B. Hellenkamp, and T. Hugel, *Four-colour FRET reveals directionality in the Hsp90 multicomponent machinery.* Nature communications **5**, 4192 (2014).

[58] A. Schulze, G. Beliu, D. A. Helmerich, J. Schubert, L. H. Pearl, C. Prodromou, and H. Neuweiler, *Cooperation of local motions in the Hsp90 molecular chaperone ATPase mechanism,* Nature Chemical Biology (2016).

[59] M. Jahn, J. Buchner, T. Hugel, and M. Rief, *Folding and assembly of the large molecular machine Hsp90 studied in single-molecule experiments,* Proceedings of the National Academy of Sciences **113**, 1232 (2016).

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

[61] D. C. Boisvert, J. Wang, Z. Otwinowski, A. L. Norwich, and P. B. Sigler, *The 2.4 Å crystal structure of the bacterial chaperonin GroEL complexed with ATPγS,* Nature Structural Biology **3**, 170 (1996).

[62] H. R. Saibil, W. A. Fenton, D. K. Clare, and A. L. Horwich, *Structure and Allostery of the Chaperonin GroEL,* Journal of Molecular Biology **425**, 1476 (2013).

[63] D. K. Clare, D. Vasishtan, S. Stagg, J. Quispe, G. W. Farr, M. Topf, A. L. Horwich, and H. R. Saibil, *ATP-Triggered Conformational Changes Delineate Substrate-Binding and -Folding Mechanics of the GroEL Chaperonin,* Cell **149**, 113 (2012).

[64] J. S. Weissman, H. S. Rye, W. A. Fenton, J. M. Beechem, and A. L. Horwich, *Characterization of the Active Intermediate of a GroEL–GroES-Mediated Protein Folding Reaction,* Cell **84**, 481 (1996).

[65] H. S. Rye, S. G. Burston, W. A. Fenton, J. M. Beechem, Z. Xu, P. B. Sigler, and A. L. Horwich, *Distinct actions of cis and trans ATP within the double ring of the chaperonin GroEL,* Nature **388**, 792 (1997).

[66] H. S. Rye, A. M. Roseman, S. Chen, K. Furtak, W. A. Fenton, H. R. Saibil, and A. L. Horwich, *GroEL-GroES Cycling: ATP and nonnative polypeptide direct alternation of folding-active rings,* Cell **97**, 325 (1999).

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

[68] N. A. Ranson, D. K. Clare, G. W. Farr, D. Houldershaw, A. L. Horwich, and H. R. Saibil, *Allosteric signaling of ATP hydrolysis in GroEL–GroES complexes,* Nature Structural & Molecular Biology **13**, 147 (2006).

[69] D. K. Clare, P. J. Bakkes, H. van Heerikhuizen, S. M. van der Vies, and H. R. Saibil, *Chaperonin complex with a newly folded protein encapsulated in the folding chamber,* Nature **457**, 107 (2009).

[70] M. Hayer-Hartl, A. Bracher, and F. U. Hartl, *The GroEL–GroES Chaperonin Machine: A Nano-Cage for Protein Folding,* Trends in Biochemical Sciences **41**, 62 (2016).

[71] Z. Lin, D. Madan, and H. S. Rye, *GroEL stimulates protein folding through forced unfolding,* Nature Structural & Molecular Biology **15**, 303 (2008).

[72] Z. Lin, J. Puchalla, D. Shoup, and H. S. Rye, *Repetitive protein unfolding by the trans ring of the GroEL-GroES chaperonin complex stimulates folding.* The Journal of biological chemistry **288**, 30944 (2013).

[73] D. Madan, Z. Lin, and H. S. Rye, *Triggering protein folding within the GroEL-GroES complex.* The Journal of biological chemistry **283**, 32003 (2008).

[74] T. Sameshima, R. Iizuka, T. Ueno, and T. Funatsu, *Denatured proteins facilitate the formation of the football-shaped GroEL–(GroES) 2 complex,* Biochemical Journal **427**, 247 (2010).

[75] X. Ye and G. H. Lorimer, *Substrate protein switches GroE chaperonins from asymmetric to symmetric cycling by catalyzing nucleotide exchange.* Proceedings of the National Academy of Sciences of the United States of America **110**, E4289 (2013).

[76] Y. Takei, R. Iizuka, T. Ueno, and T. Funatsu, *Single-molecule observation of protein folding in symmetric GroEL-(GroES)2 complexes.* The Journal of biological chemistry **287**, 41118 (2012).

[77] T. Ueno, H. Taguchi, H. Tadakuma, M. Yoshida, and T. Funatsu, *GroEL Mediates Protein Folding with a Two Successive Timer Mechanism,* Molecular Cell **14**, 423 (2004).

[78] M. K. Hayer-Hartl, K. L. Ewalt, and F. U. Hartl, *On the Role of Symmetrical and Asymmetrical Chaperonin Complexes in Assisted Protein Folding,* Biological Chemistry **380**, 531 (1999).

[79] D. Yang, X. Ye, and G. H. Lorimer, *Symmetric GroEL:GroES2 complexes are the protein-folding functional form of the chaperonin nanomachine.* Proceedings of the National Academy of Sciences of the United States of America **110**, E4298 (2013).

[80] X. Fei, X. Ye, N. A. LaRonde, and G. H. Lorimer, *Formation and structures of GroEL:GroES2 chaperonin footballs, the protein-folding functional form.* Proceedings of the National Academy of Sciences of the United States of America **111**, 12775 (2014).

[81] O. Llorca, A. Galán, J. L. Carrascosa, A. Muga, and J. M. Valpuesta, *GroEL under heat-shock. Switching from a folding to a storing function.* The Journal of biological chemistry **273**, 32587 (1998).

[82] A. C. Apetri and A. L. Horwich, *Chaperonin chamber accelerates protein folding through passive action of preventing aggregation.* Proceedings of the National Academy of Sciences of the United States of America **105**, 17351 (2008).

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

[84] K. Chakraborty, M. Chatila, J. Sinha, Q. Shi, B. C. Poschner, M. Sikor, G. Jiang, D. C. Lamb, F. U. Hartl, and M. Hayer-Hartl, *Chaperonin-Catalyzed Rescue of Kinetically Trapped States in Protein Folding,* Cell **142**, 112 (2010).

[85] F. Georgescauld, K. Popova, A. J. Gupta, A. Bracher, J. R. Engen, M. Hayer-Hartl, and F. U. Hartl, *GroEL/ES Chaperonin Modulates the Mechanism and Accelerates the Rate of TIM-Barrel Domain Folding,* Cell **157**, 922 (2014).

[86] A. J. Gupta, S. Haldar, G. Miličić, F. U. Hartl, and M. Hayer-Hartl, *Active Cage Mechanism of Chaperonin-Assisted Protein Folding Demonstrated at Single-Molecule Level,* Journal of Molecular Biology **426**, 2739 (2014).

[87] D. Thirumalai and G. H. Lorimer, *Chaperonin-Mediated Protein Folding,* Annual Review of Biophysics and Biomolecular Structure **30**, 245 (2001).

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

[89] M. R. Betancourt and D. Thirumalai, *Exploring the kinetic requirements for en-hancement of protein folding rates in the GroEL cavity,* Journal of Molecular Bi-ology **287**, 627 (1999).

[90] D. Thirumalai, D. K. Klimov, and G. H. Lorimer, *Caging helps proteins fold.* Pro-ceedings of the National Academy of Sciences of the United States of America **100**, 11195 (2003).

[91] A. I. Jewett, A. Baumketner, and J.-E. Shea, *Accelerated folding in the weak hy-drophobic environment of a chaperonin cavity: creation of an alternate fast fold-ing pathway.* Proceedings of the National Academy of Sciences of the United States of America **101**, 13192 (2004).

[92] H. Hofmann, F. Hillger, S. H. Pfeil, A. Hoffmann, D. Streich, D. Haenni, D. Net-tels, E. A. Lipman, and B. Schuler, *Single-molecule spectroscopy of protein fold-ing in a chaperonin cage.* Proceedings of the National Academy of Sciences of the United States of America **107**, 11793 (2010).

[93] S. Ungelenk, F. Moayed, C.-T. Ho, T. Grousl, A. Scharf, A. Mashaghi, S. Tans, M. P. Mayer, A. Mogk, and B. Bukau, *Small heat shock proteins sequester misfolding proteins in near-native conformation for cellular protection and efficient refold-ing,* Nature Communications **7**, 13673 (2016).

[94] C.-J. Tsai, S. Kumar, B. Ma, and R. Nussinov, *Folding funnels, binding funnels, and protein function,* Protein Science **8**, 1181 (1999).

[95] V. Spirin and L. A. Mirny, *Protein complexes and functional modules in molec-ular networks,* Proceedings of the National Academy of Sciences **100**, 12123 (2003).

[96] T. Pawson and J. D. Scott, *Signaling through scaffold, anchoring, and adaptor proteins,* Science **278**, 2075 (1997).

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

[98] M. J. Avellaneda, E. J. Koers, M. M. Naqvi, and S. J. Tans, *The chaperone toolbox at the single-molecule level : From clamping to confining,* Protein Science **26**, 1291 (2017).

[99] E. Nogales and S. H. Scheres, *Cryo-em: a unique tool for the visualization of macromolecular complexity,* Molecular cell **58**, 677 (2015).

[100] G. M. Clore and A. M. Gronenborn, *Determining the structures of large proteins and protein complexes by nmr,* Trends in biotechnology **16**, 22 (1998).

[101] R. A. Harrison and J. R. Engen, *Conformational insight into multi-protein sig-naling assemblies by hydrogen–deuterium exchange mass spectrometry,* Current opinion in structural biology **41**, 187 (2016).

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

[103] D. L. Vossen, A. van der Horst, M. Dogterom, and A. van Blaaderen, *Optical tweezers and confocal microscopy for simultaneous three-dimensional manipulation and imaging in concentrated colloidal dispersions,* Review of Scientific Instruments **75**, 2960 (2004).

[104] 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 (New York, N.Y.) **318**, 279 (2007).

[105] P. B. Tarsa, R. R. Brau, M. Barch, J. M. Ferrer, Y. Freyzon, P. Matsudaira, and M. J. Lang, *Detecting force-induced molecular transitions with fluorescence resonant energy transfer.* Angewandte Chemie (International ed. in English) **46**, 1999 (2007).

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

[107] J. van Mameren, P. Gross, G. Farge, P. Hooijman, M. Modesti, M. Falkenberg, G. J. Wuite, and E. J. Peterman, *Unraveling the structure of DNA during overstretching by using multicolor, single-molecule fluorescence imaging,* Proceedings of the National Academy of Sciences **106**, 18231 (2009).

[108] I. Heller, T. P. Hoekstra, G. A. King, E. J. Peterman, and G. J. Wuite, *Optical tweezers analysis of DNA–protein complexes,* Chemical reviews **114**, 3087 (2014).

[109] M. L. Bennink, O. D. Schärer, R. Kanaar, K. Sakata-Sogawa, J. M. Schins, J. S. Kanger, B. G. de Grooth, and J. Greve, *Single-molecule manipulation of double-stranded DNA using optical tweezers: Interaction studies of DNA with reca and yoyo-1,* Cytometry: The Journal of the International Society for Analytical Cytology **36**, 200 (1999).

[110] Y. Arai, R. Yasuda, K.-i. Akashi, Y. Harada, H. Miyata, K. Kinosita Jr, and H. Itoh, *Tying a molecular knot with optical tweezers,* Nature **399**, 446 (1999).

[111] P. Gross, G. Farge, E. J. Peterman, and G. J. Wuite, *Combining optical tweezers, single-molecule fluorescence microscopy, and microfluidics for studies of DNA–protein interactions,* in *Methods in enzymology,* Vol. 475 (Elsevier, 2010) pp. 427–453.

[112] P. Roach, D. Farrar, and C. C. Perry, *Interpretation of protein adsorption: surface-induced conformational changes,* Journal of the American Chemical Society **127**, 8168 (2005).

[113] E. J. G. Peterman, F. Gittes, and C. F. Schmidt, *Laser-induced heating in optical traps.* Biophysical journal **84**, 1308 (2003).

[114] S. R. K. Ainavarapu, L. Li, C. L. Badilla, and J. M. Fernandez, *Ligand binding modulates the mechanical stability of dihydrofolate reductase,* Biophysical journal **89**, 3337 (2005).

[115] C. Cecconi, E. A. Shank, F. W. Dahlquist, S. Marqusee, and C. Bustamante, *Protein-DNA chimeras for single molecule mechanical folding studies with the optical tweezers,* European Biophysics Journal **37**, 729 (2008).

[116] Z. Ganim and M. Rief, *Mechanically switching single-molecule fluorescence of gfp by unfolding and refolding,* Proceedings of the National Academy of Sciences , 201704937 (2017).

[117] K. Hatch, C. Danilowicz, V. Coljee, and M. Prentiss, *Demonstration that the shear force required to separate short double-stranded DNA does not increase significantly with sequence length for sequences longer than 25 base pairs,* Physical Review E **78**, 011920 (2008).

[118] A. Mukhortava and M. Schlierf, *Efficient Formation of Site-Specific Protein–DNA Hybrids Using Copper-Free Click Chemistry,* Bioconjugate Chemistry **27**, 1559 (2016).

[119] J. Yin, P. D. Straight, S. M. McLoughlin, Z. Zhou, A. J. Lin, D. E. Golan, N. L. Kelleher, R. Kolter, and C. T. Walsh, *Genetically encoded short peptide tag for versatile protein labeling by Sfp phosphopantetheinyl transferase,* Proceedings of the National Academy of Sciences **102**, 15815 (2005).

[120] C. Goffin, V. Bailly, and W. G. Verly, *Nicks 3' or 5' to ap sites or to mispaired bases, and one-nucleotide gaps can be sealed by t4 DNA ligase,* Nucleic acids research **15**, 8755 (1987).

[121] Y. Hao, C. Canavan, S. S. Taylor, and R. A. Maillard, *Integrated method to attach DNA handles and functionally select proteins to study folding and protein-ligand interactions with optical tweezers,* Scientific reports **7**, 10843 (2017).

[122] R. Janissen, B. a. Berghuis, D. Dulin, M. Wink, T. van Laar, and N. H. Dekker, *Invincible DNA tethers: covalent DNA anchoring for enhanced temporal and force stability in magnetic tweezers experiments.* Nucleic acids research **42**, gku677 (2014).

[123] D. P. Minde, E. F. Halff, and S. Tans, *Designing disorder: Tales of the unexpected tails,* Intrinsically disordered proteins **1**, e26790 (2013).

[124] H. W. Jannasch, C. O. Wirsen, S. J. Molyneaux, and T. A. Langworthy, *Comparative physiological studies on hyperthermophilic archaea isolated from deep-sea hot vents with emphasis on Pyrococcus strain GB-D,* Applied and Environmental Microbiology **58**, 3472 (1992).

[125] A. K. Dunker, M. S. Cortese, P. Romero, L. M. Iakoucheva, and V. N. Uversky, *Flexible nets: the roles of intrinsic disorder in protein interaction networks,* The FEBS journal **272**, 5129 (2005).

[126] S. H. Sternberg, B. LaFrance, M. Kaplan, and J. A. Doudna, *Conformational control of DNA target cleavage by CRISPR–Cas9,* Nature **527**, 110 (2015).

[127] M. Lambrughi, L. De Gioia, F. L. Gervasio, K. Lindorff-Larsen, R. Nussinov, C. Urani, M. Bruschi, and E. Papaleo, *DNA-binding protects p53 from interactions with cofactors involved in transcription-independent functions,* Nucleic acids research **44**, 9096 (2016).

[128] D. M. Presman, S. Ganguly, R. L. Schiltz, T. A. Johnson, T. S. Karpova, and G. L. Hager, *DNA binding triggers tetramerization of the glucocorticoid receptor in live cells,* Proceedings of the National Academy of Sciences **113**, 8236 (2016).

[129] K. Berg-Sørensen and H. Flyvbjerg, *Power spectrum analysis for optical tweezers,* Review of Scientific Instruments **75**, 594 (2004).

[130] R. Petrosyan, *Improved approximations for some polymer extension models,* Rheologica Acta **56**, 21 (2017).

[131] T. Odijk, *Stiff chains and filaments under tension,* Macromolecules **28**, 7016 (1995).

[132] O. K. Dudko, G. Hummer, and A. Szabo, *Theory, analysis, and interpretation of single-molecule force spectroscopy experiments.* Proceedings of the National Academy of Sciences of the United States of America **105**, 15755 (2008).

[133] N. Ferguson and A. R. Fersht, *Early events in protein folding,* Current opinion in structural biology **13**, 75 (2003).

[134] C. M. Dobson, *Protein folding and misfolding,* Nature **426**, 884 (2003).

[135] Z. Lin and H. S. Rye, *Groel-mediated protein folding: making the impossible, possible,* Crit Rev Biochem Mol Biol **41**, 211 (2006).

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

[137] R. J. Ellis, *Protein folding: Importance of the anfinsen cage,* Current Biology **13**, R881 (2003).

[138] F. Motojima, Y. Motojima-Miyazaki, and M. Yoshida, *Revisiting the contribution of negative charges on the chaperonin cage wall to the acceleration of protein folding,* Proceedings of the National Academy of Sciences of the United States of America **109**, 15740 (2012).

[139] C. Cecconi, E. A. Shank, C. Bustamante, and S. Marqusee, *Direct Observation of the Three-State Folding of a Single Protein Molecule,* Science **4174**, 2057 (2005).

[140] J. Stigler, F. Ziegler, A. Gieseke, J. C. M. Gebhardt, and M. Rief, *The Complex Folding Network of Single Calmodulin Molecules,* Science **334**, 512 (2011).

[141] M. M. Naqvi, P. O. Heidarsson, M. R. Otazo, A. Mossa, B. B. Kragelund, and C. Cecconi, *Single-Molecule Folding Mechanisms of the apo- and Mg2+-Bound States of Human Neuronal Calcium Sensor-1,* Biophysical Journal **109**, 113 (2015).

[142] J. A. Riback, M. A. Bowman, A. M. Zmyslowski, C. R. Knoverek, J. M. Jumper, J. R. Hinshaw, E. B. Kaye, K. F. Freed, P. L. Clark, and T. R. Sosnick, *Innovative scattering analysis shows that hydrophobic disordered proteins are expanded in water,* Science **358**, 238 (2017).

[143] A. Borgia, W. Zheng, K. Buholzer, M. B. Borgia, A. Schuler, H. Hofmann, A. Soranno, D. Nettels, K. Gast, A. Grishaev, *et al.*, *Consistent view of polypeptide chain expansion in chemical denaturants from multiple experimental methods,* Journal of the American Chemical Society **138**, 11714 (2016).

[144] S. Priya, S. K. Sharma, V. Sood, R. U. H. Mattoo, A. Finka, A. Azem, P. De Los Rios, and P. Goloubinoff, *Groel and cct are catalytic unfoldases mediating out-of-cage polypeptide refolding without atp,* Proceedings of the National Academy of Sciences of the United States of America **110**, 7199 (2013).

[145] J. Weaver and H. S. Rye, *The c-terminal tails of the bacterial chaperonin groel stimulate protein folding by directly altering the conformation of a substrate protein,* J Biol Chem **289**, 23219 (2014).

[146] J. S. Weissman, C. M. Hohl, O. Kovalenko, Y. Kashi, S. X. Chen, K. Braig, H. R. Saibil, W. A. Fenton, and A. L. Horwich, *Mechanism of groel action - productive release of polypeptide from a sequestered position under groes,* Cell **83**, 577 (1995).

[147] F. Motojima and M. Yoshida, *Polypeptide in the chaperonin cage partly protrudes out and then folds inside or escapes outside,* Embo Journal **29**, 4008 (2010).

[148] C. M. Dobson, A. Sali, and M. Karplus, *Protein folding: A perspective from theory and experiment,* Angew Chem Int Ed Engl **37**, 868 (1998).

[149] H. S. Rye, *Application of fluorescence resonance energy transfer to the groel-groes chaperonin reaction,* Methods **24**, 278 (2001).

[150] J. Tyedmers, A. Mogk, and B. Bukau, *Cellular strategies for controlling protein aggregation,* Nature reviews Molecular cell biology **11**, 777 (2010).

[151] P. Brundin, R. Melki, and R. Kopito, *Prion-like transmission of protein aggregates in neurodegenerative diseases,* Nature reviews Molecular cell biology **11**, 301 (2010).

[152] M. Jucker and L. C. Walker, *Self-propagation of pathogenic protein aggregates in neurodegenerative diseases,* Nature **501**, 45 (2013).

[153] A. Mogk, B. Bukau, and H. H. Kampinga, *Cellular handling of protein aggregates by disaggregation machines,* Molecular cell **69**, 214 (2018).

[154] E. C. Duran, C. L. Weaver, and A. L. Lucius, *Comparative Analysis of the Structure and Function of AAA+ Motors ClpA, ClpB, and Hsp104: Common Threads and Disparate Functions,* Frontiers in Molecular Biosciences **4**, 1 (2017).

[155] J. Shorter and D. R. Southworth, *Spiraling in control: Structures and mechanisms of the Hsp104 disaggregase,* Cold Spring Harbor perspectives in biology , a034033 (2019).

[156] J. Weibezahn, P. Tessarz, C. Schlieker, R. Zahn, Z. Maglica, S. Lee, H. Zentgraf, E. U. Weber-Ban, D. A. Dougan, and F. T. Tsai, *Thermotolerance requires refolding of aggregated proteins by substrate translocation through the central pore of clpb,* Cell **119**, 653 (2004).

[157] T. Li, C. L. Weaver, J. Lin, E. C. Duran, J. M. Miller, and A. L. Lucius, *Escherichia coli ClpB is a non-processive polypeptide translocase,* Biochemical Journal **470**, 39 (2015).

[158] P. De Los Rios and P. Goloubinoff, *Hsp70 chaperones use atp to remodel native protein oligomers and stable aggregates by entropic pulling,* Nature structural&molecular biology **23**, 766 (2016).

[159] V. Grimminger-Marquardt and H. A. Lashuel, *Structure and function of the molecular chaperone Hsp104 from yeast,* Biopolymers: Original Research on Biomolecules **93**, 252 (2010).

[160] J. A. Wagoner and K. A. Dill, *Molecular motors: Power strokes outperform brownian ratchets,* The Journal of Physical Chemistry B **120**, 6327 (2016).

[161] N. D. Werbeck, S. Schlee, and J. Reinstein, *Coupling and dynamics of subunits in the hexameric aaa+ chaperone clpb,* Journal of molecular biology **378**, 178 (2008).

[162] T. Uchihashi, Y.-h. Watanabe, Y. Nakazaki, T. Yamasaki, H. Watanabe, T. Maruno, K. Ishii, S. Uchiyama, C. Song, K. Murata, R. Iino, and T. Ando, *Dynamic structural states of clpb involved in its disaggregation function,* Nature communications **9**, 2147 (2018).

[163] T. Haslberger, A. Zdanowicz, I. Brand, J. Kirstein, K. Turgay, A. Mogk, and B. Bukau, *Protein disaggregation by the AAA+ chaperone ClpB involves partial threading of looped polypeptide segments.* Nature structural & molecular biology **15**, 641 (2008).

[164] E. Kummer, A. Szlachcic, K. B. Franke, S. Ungelenk, B. Bukau, and A. Mogk, *Bacterial and Yeast AAA + Disaggregases ClpB and Hsp104 Operate through Conserved Mechanism Involving Cooperation with Hsp70,* Journal of Molecular Biology **428**, 4378 (2016).

[165] Y. Oguchi, E. Kummer, F. Seyffer, M. Berynskyy, B. Anstett, R. Zahn, R. C. Wade, A. Mogk, and B. Bukau, *A tightly regulated molecular toggle controls aaa+ disaggregase,* Nature structural&molecular biology **19**, 1338 (2012).

[166] J. A. Fernández-Higuero, A. Aguado, J. Perales-Calvo, F. Moro, and A. Muga, *Activation of the dnak-clpb complex is regulated by the properties of the bound substrate,* Scientific reports **8**, 5796 (2018).

[167] 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).

[168] 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).

[169] A. O. Olivares, A. R. Nager, O. Iosefson, R. T. Sauer, and T. A. Baker, *Mechanochemical basis of protein degradation by a double-ring aaa+ machine,* Nature structural&molecular biology **21**, 871 (2014).

[170] M. Ganji, I. A. Shaltiel, S. Bisht, E. Kim, A. Kalichava, C. H. Haering, and C. Dekker, *Real-time imaging of DNA loop extrusion by condensin,* Science , eaar7831 (2018).

[171] N. Forns, S. de Lorenzo, M. Manosas, K. Hayashi, J. M. Huguet, and F. Ritort, *Improving signal/noise resolution in single-molecule experiments using molecular constructs with short handles,* Biophysical journal **100**, 1765 (2011).

[172] S. N. Gates, A. L. Yokom, J. Lin, M. E. Jackrel, A. N. Rizo, N. M. Kendsersky, C. E. Buell, E. A. Sweeny, K. L. Mack, E. Chuang, M. P. Torrente, M. Su, J. Shorter, and D. R. Southworth, *Ratchet-like polypeptide translocation mechanism of the AAA+ disaggregase Hsp104,* Science **357**, 273 (2017).

[173] C. Deville, M. Carroni, K. B. Franke, M. Topf, B. Bukau, A. Mogk, and H. R. Saibil, *Structural pathway of regulated substrate transfer and threading through an hsp100 disaggregase,* Science advances **3**, e1701726 (2017).

[174] H. Yu, T. J. Lupoli, A. Kovach, X. Meng, G. Zhao, C. F. Nathan, and H. Li, *Atp hydrolysis-coupled peptide translocation mechanism of mycobacterium tuberculosis clpb,* Proceedings of the National Academy of Sciences **115**, E9560 (2018).

[175] A. Y. Lyubimov, M. Strycharska, and J. M. Berger, *The nuts and bolts of ring-translocase structure and mechanism,* Current opinion in structural biology **21**, 240 (2011).

[176] A. Savitzky and M. J. Golay, *Smoothing and differentiation of data by simplified least squares procedures,* Analytical chemistry **36**, 1627 (1964).

[177] G. Salbreux, G. Charras, and E. Paluch, *Actin cortex mechanics and cellular morphogenesis.* Trends in cell biology **22**, 536 (2012).

[178] L. B. Case and C. M. Waterman, *Integration of actin dynamics and cell adhesion by a three-dimensional, mechanosensitive molecular clutch,* Nature Cell Biology **17**, 955 (2015).

[179] F. Meng and E. Terentjev, *Transient Network at Large Deformations: Elastic–Plastic Transition and Necking Instability,* Polymers **8**, 108 (2016).

[180] P. J. Skrzeszewska, J. Sprakel, F. A. de Wolf, R. Fokkink, M. A. Cohen Stuart, and J. van der Gucht, *Fracture and Self-Healing in a Well-Defined Self-Assembled Polymer Network,* Macromolecules **43**, 3542 (2010).

[181] Y. Mulla, A. Aufderhorst-Roberts, and G. H. Koenderink, *Shaping up synthetic cells,* Physical Biology **15**, 041001 (2018).

[182] S. Aime, L. Ramos, and L. Cipelletti, *Microscopic dynamics and failure precursors of a gel under mechanical load,* Proceedings of the National Academy of Sciences **115**, 3587 (2018).

[183] N. Y. Yao, C. P. Broedersz, M. Depken, D. J. Becker, M. R. Pollak, F. C. MacKintosh, and D. a. Weitz, *Stress-Enhanced Gelation: A Dynamic Nonlinearity of Elasticity,* Physical Review Letters **110**, 018103 (2013).

[184] A. P. Liu and D. a. Fletcher, *Actin polymerization serves as a membrane domain switch in model lipid bilayers.* Biophysical journal **91**, 4064 (2006).

[185] U. G. K. Wegst, H. Bai, E. Saiz, A. P. Tomsia, R. O. Ritchie, C. Ortiz, M. Boyce, U. G. K. Wegst, H. Bai, E. Saiz, A. P. Tomsia, and R. O. Ritchie, *Bioinspired structural materials,* Nature materials **14**, 23 (2014).

[186] L. Casares, R. Vincent, D. Zalvidea, N. Campillo, D. Navajas, M. Arroyo, and X. Trepat, *Hydraulic fracture during epithelial stretching,* Nature Materials **14**, 343 (2015).

[187] C. M. Denais, R. M. Gilbert, P. Isermann, A. L. McGregor, M. te Lindert, B. Weigelin, P. M. Davidson, P. Friedl, K. Wolf, and J. Lammerding, *Nuclear envelope rupture and repair during cancer cell migration,* Science **352**, 353 (2016).

[188] J. C. Friedland, M. H. Lee, and D. Boettiger, *Mechanically Activated Integrin Switch Controls 5 1 Function,* Science **323**, 642 (2009).

[189] S. Chakrabarti, M. Hinczewski, and D. Thirumalai, *Plasticity of hydrogen bond networks regulates mechanochemistry of cell adhesion complexes.* Proceedings of the National Academy of Sciences of the United States of America **111**, 9048 (2014).

[190] C. D. Buckley, J. Tan, K. L. Anderson, D. Hanein, N. Volkmann, W. I. Weis, W. J. Nelson, and A. R. Dunn, *Cell adhesion. The minimal cadherin-catenin complex binds to actin filaments under force.* Science **346**, 1254211 (2014), 15334406 .

[191] S. Chakrabarti, M. Hinczewski, and D. Thirumalai, *Phenomenological and microscopic theories for catch bonds,* Journal of Structural Biology **197**, 50 (2017), arXiv:1601.02050v1 .

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

[193] B.-A. Truong-Quang and P.-F. Lenne, *Membrane microdomains: from seeing to understanding.* Frontiers in plant science **5**, 18 (2014).

[194] T. A. Tsunoyama, Y. Watanabe, J. Goto, K. Naito, R. S. Kasai, K. G. N. Suzuki, T. K. Fujiwara, and A. Kusumi, *Super-long single-molecule tracking reveals dynamic-anchorage-induced integrin function,* Nature Chemical Biology **14**, 497 (2018).

[195] T. Travers, H. Shao, A. Wells, and C. J. Camacho, *Modeling the assembly of the multiple domains of α-actinin-4 and its role in actin cross-linking,* Biophysical Journal **104**, 705 (2013).

[196] D. Feng, J. Notbohm, A. Benjamin, S. He, M. Wang, L.-H. Ang, M. Bantawa, M. Bouzid, E. Del Gado, R. Krishnan, and M. R. Pollak, *Disease-causing mutation in α-actinin-4 promotes podocyte detachment through maladaptation to periodic stretch,* Proceedings of the National Academy of Sciences **115**, 1517 (2018).

[197] V. E. Galkin, A. Orlova, A. Salmazo, K. Djinovic-Carugo, and H. Edward Egelman, *Opening of tandem calponin homology domains regulates their affinity for F-actin,* nat struct mol biol. **17**, 614 (2010).

[198] A. J. Ehrlicher, R. Krishnan, M. Guo, C. M. Bidan, D. a. Weitz, and M. R. Pollak, *Alpha-actinin binding kinetics modulate cellular dynamics and force generation,* Proceedings of the National Academy of Sciences **112**, 201505652 (2015).

[199] E. S. Schiffhauer, T. Luo, K. Mohan, V. Srivastava, X. Qian, E. R. Griffis, P. A. Iglesias, and D. N. Robinson, *Mechanoaccumulative Elements of the Mammalian Actin Cytoskeleton,* Current Biology **26**, 1473 (2016).

[200] T. Luo, K. Mohan, P. a. Iglesias, and D. N. Robinson, *Molecular mechanisms of cellular mechanosensing.* Nature materials **12**, 1064 (2013).

[201] E. D. A. Ribeiro, N. Pinotsis, A. Ghisleni, A. Salmazo, P. V. Konarev, J. Kostan, B. Sjöblom, C. Schreiner, A. A. Polyansky, E. A. Gkougkoulia, M. R. Holt, F. L. Aachmann, B. Žagrović, E. Bordignon, K. F. Pirker, D. I. Svergun, M. Gautel, and K. Djinović-Carugo, *The structure and regulation of human muscle α-Actinin,* Cell **159**, 1447 (2014).

[202] C. Storm, J. J. Pastore, F. C. MacKintosh, T. C. Lubensky, and P. A. Janmey, *Nonlinear elasticity in biological gels,* Nature **435**, 191 (2005).

[203] E. A. Novikova and C. Storm, *Contractile Fibers and Catch-Bond Clusters: a Biological Force Sensor?* Biophysical Journal **105**, 1336 (2013).

[204] J. M. Laakso, J. H. Lewis, H. E. Shuman, and M. E. Ostap, *Myosin I Can Act As a Molecular Force Sensor,* Science **321**, 133 (2008).

[205] T. L. Sun, T. Kurokawa, S. Kuroda, A. B. Ihsan, T. Akasaki, K. Sato, M. A. Haque, T. Nakajima, and J. P. Gong, *Physical hydrogels composed of polyampholytes demonstrate high toughness and viscoelasticity,* Nature Materials **12**, 932 (2013).

[206] L. Sun, Q. H. Cheng, H. J. Gao, and Y. W. Zhang, *Effect of loading conditions on the dissociation behaviour of catch bond clusters,* Journal of The Royal Society Interface **9**, 928 (2012).

[207] A. M. Rosales and K. S. Anseth, *The design of reversible hydrogels to capture extracellular matrix dynamics,* Nature Reviews Materials **1**, 15012 (2016).

[208] V. W. Tang and W. M. Brieher, *alpha-Actinin-4/FSGS1 is required for Arp2/3-dependent actin assembly at the adherens junction,* Journal of Cell Biology **196**, 115 (2012).

[209] J. Alvarado, M. Sheinman, A. Sharma, F. C. MacKintosh, and G. H. Koenderink, *Molecular motors robustly drive active gels to a critically connected state,* Nature Physics **9**, 591 (2013).

[210] C. R. Booth, A. S. Meyer, Y. Cong, M. Topf, A. Sali, S. J. Ludtke, W. Chiu, and J. Frydman, *Mechanism of lid closure in the eukaryotic chaperonin tric/cct,* Nature structural & molecular biology **15**, 746 (2008).

[211] D. Balchin, G. Miličić, M. Strauss, M. Hayer-Hartl, and F. U. Hartl, *Pathway of actin folding directed by the eukaryotic chaperonin tric,* Cell **174**, 1507 (2018).

[212] M. E. DeSantis, E. H. Leung, E. A. Sweeny, M. E. Jackrel, M. Cushman-Nick, A. Neuhaus-Follini, S. Vashist, M. A. Sochor, M. N. Knight, and J. Shorter, *Operational plasticity enables hsp104 to disaggregate diverse amyloid and nonamyloid clients,* Cell **151**, 778 (2012).

[213] 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).

ABOUT THE AUTHOR

"Somewhere in La Mancha, in a place whose name I do not care to remember,[1] a gentleman was born not long ago,[2] one of those who...".[3]

Like the great Don Quixote manchego knight, Mario Avellaneda moved to the land of Valencia to pursue his dreams. There, he studied physics from 2008 until 2012 and learned how to distinguish giants from windmills in science.

The year 2012 turned out to be a special one. He received a Collaboration Scholarship from the Ministry of Education of Spain to study the multipactor effect under the supervision of Prof. Benito Gimeno, which resulted in his Bachelor Thesis. The same year, he read *What is life?* by Erwin Schrödinger, a brief but powerful book that changed his life in ways he could not have foreseen, and that made him decide to move towards the field of biophysics. He also received the prestigious scholarship from La Caixa and DAAD for postgraduate studies in Europe. With that support, in the fall of 2012, he moved to the beautiful city of Dresden, Germany, to join the Master in Nanobiophysics. In 2014, he wrote his Master thesis in Max Planck Institute of Molecular Cell Biology and Genetics under the supervision of Prof. Stephan Grill, and obtained his Master degree with highest honours.

That same year, he started his PhD in the Biophysics group of Prof. Sander Tans at AMOLF, Amsterdam, where he performed the research presented in this thesis.

---

[1] Albacete, Castilla-La Mancha, Spain

[2] 23 April 1990

[3] *"En un lugar de la Mancha, de cuyo nombre no quiero acordarme, no ha mucho tiempo vivía..."* is the famous beginning of the universal novel *Don Quijote de la Mancha*, by Miguel de Cervantes

LIST OF PUBLICATIONS

RELATED TO THIS THESIS

M. J. Avellaneda, E. J. Koers, M. M. Naqvi and S. J. Tans, *The chaperone toolbox at the single-molecule level: from clamping to confining*, Protein Science **26**, 7, 1291-1302 (2017). (Chapter 1)

F. Wruck, **M. J. Avellaneda**, E. J. Koers, D. P. Minde, M. Mayer, G. Kramer, A. Mashaghi and S. J. Tans, *Protein Folding Mediated by Trigger Factor and Hsp70: New Insights from Single-Molecule Approaches*, Journal of Molecular Biology **430**, 4, pages 438-449 (2018)

M. J. Avellaneda, E. J. Koers, D. P. Minde, V. Sunderlikova and S. J. Tans, *Simultaneous sensing and imaging of individual biomolecular complexes enabled by a modular DNA-protein coupling*, Commun. Chem., manuscript under review. (Chapter 2)

M. M. Naqvi, **M. J. Avellaneda**, A. Roth, E. J. Koers, V. Sunderlikova, G. Kramer, H. Rye and S. J. Tans, *GroEL-mediated acceleration of protein folding by enhanced collapse*, manuscript submitted. (Chapter 3)

M. J. Avellaneda, K. B. Franke, V. Sunderlikova, B. Bukau, A. Mogk and S. J. Tans, *Direct observation of processive peptide loop extrusion by a Hsp100 disaggregase*, Nature, manuscript under review. (Chapter 4)

Y. Mulla, **M. J. Avellaneda**, S. J. Tans, G.H. Koenderink, *Catch bonds form dynamic yet strong materials*, manuscript in preparation. (Chapter 5)

OTHER PUBLICATIONS

D. H. Murray, M. Jahnel, J. Lauer, **M. J. Avellaneda**, N. Brouilly, A. Cezanne, H. Morales-Navarrete, E. D. Perini, C. Ferguson, A. N. Lupas, Y. Kalaidzidis, R. G. Parton, S. W. Grill and M. Zerial, *An endosomal tether undergoes an entropic collapse to bring vesicles together*, Nature **537**, 107-11 (2016)

D. Choudhary, L. Mediani, **M. J. Avellaneda**, S. Alberti, E. Boczek, S. J Tans, S. Carra and C. Cecconi, *Single-molecule studies of the molecular mechanisms underlying the chaperone activity of HSPB8*, manuscript in preparation.

ACKNOWLEDGEMENTS

A PhD is much more than a handful of pages in a thesis. It is a unique experience mixing satisfaction and frustration, joy and bitterness, excitement and confusion. And as it is the case with most experiences, it is profoundly shaped by the people sharing the journey. I could not have wished for a better company in this adventure.

First, I would like to thank Sander. Working under your supervision was an enriching experience, and I am grateful for everything I have learned from you. I admire your unique ability to spot the important questions and design the best experiments to answer them. I really enjoyed all our discussions, not only scientific, but also those that started with a wide grin and the words *So, Mario, I read in the news that....* I knew a storm was coming, but always enjoyed the turbulences.

Being part of the Biophysics group was a lot of fun, and many people deserve my gratitude. Sergey and Fatemeh, you made the start of my PhD smooth and comfortable, and your initial help was a cornerstone for the rest of my journey. David, thanks for your invaluable help with the biochemistry. I managed to capture the sparks of genius that sometimes emerged among the chaos, and learned so much from you. Eline, you rescued the tethering efforts when things seemed almost lost, thank you for all your help and the many discussions. Mohsin, it was a pleasure to work with you in the GroEL project, we had to endure the long process of scientific publishing, and learned a lot along the way. Florian, the first person I would go to with questions about optical tweezers experiments. Your knowledge and scientific intuition were an inspiration and priceless help. So were the parties we enjoyed in conferences. Anne Bart, your political incorrectness and scientific correctness was a wonderful and explosive combination, and it was a great replacement for the departure of Sebastian. Antoine, we overlapped shortly, but I am sure your driving character will take you far, good luck with the p97 project! Martijn, it was sadly good to know that I was not the only one struggling through labwork and failed experiments. Thanks for all the enjoyable conversations! Xuan and Rutger, it was always exciting and fresh to hear about your research with organoids. There is one person here who deserves special mention. Vanda, you are simply the best technician ever. All the work of this thesis has been possible because of your incredible work in the lab. I feel fortunate and extremely grateful for your constant willingness to attempt all the crazy ideas we came up with (I am sorry for your nightmares with dmMBP-2MBP-SFP-linker-Luci-MBP-casein-cysteine and similar extravagant constructs). But more important than that, you have been a friend. I appreciate your patience while listening to those few explosions of frustration I experienced. I will miss you and all the hours spent chatting about all kinds of stuff.

My paranymphs deserve a big thank you. I feel fortunate to have you as friends, and I am extremely happy that you will accompany me so closely during my defence. We have shared many special moments during these years, and I am sure there are

many more to come. Marco, we have laughed to and with each other countless times, and yet actual punches were never involved. Thank you for your permanent good mood no matter what, and remember that your defence is a few days *after* mine. I will do my best to see a last hangover presentation from you. Fede, you are such a special and unique person. Thank you for all the comforting conversations, in both directions, and for all the fun together! Seeing you is always uplifting, and it always felt strange and sad when you were not in the office, I will miss you very much!

This work would not have been possible without the great help of collaborators. The most important of all, Axel, thanks for your immense support. Your knowledge and intuition has been a great inspiration and help. Also big thanks to Kamila, Bernd, Hays and Gunter. One not always needs to go far to find great collaborators. Yuval, or the walking irony, I hope you have the great scientific success your mind deserves, it was great working with you! Thanks for introducing me and my toenail to mountain biking in Texel! Lennard, I am still not convinced that all those rituals to catch beads and pull lipid tubes in the tweezers were just superstition. It was fun to rush and get some nice results in a few weeks! Gijsje, thanks for your trust, feedback and support with the catch bond and lipid tubes projects!

AMOLF has been a wonderful place to work, and Amsterdam a lovely city to live, thanks to the people whom I had the luck to meet. A special mention goes to Nicola. You soon became a true friend and the glue keeping many of us together. Thanks for so much fun, and good luck in my home country, I am sure you will enjoy! I hope our friendship stands the years and distance as it did with so many squash and Resistance clashes. Agata, you are one of the most determined and driving women I know! Our race to finish the PhD was as fierce as our tennis matches, but I am happy you won that one. Thank you for being there whenever I needed and for choosing me as your paranymph, I feel really honoured. Mathijs, a hurricane that managed to settle in the office. I really enjoyed our off-topic discussions, as well as playing sombrero-basketball, parabolic-ball-into-full-mug-over-the-screen and unexpected-frisbee. Thanks for walking to the coffee corner with me even when your mug was most often full. Other people made our office a lot of fun: Stijn, Adeline, Mick, Leonie, Simon, Rosalie and Faan. And of course Fede, the only person mentioned twice in these acknowledgements, but who certainly deserves it. Moritz and Laura, you are such a lovely couple. Time spent with you is always truly enjoyable, thank you for being there these years, I am happy you ended up living next door. Giulia and Judith, some of the best hosts in Amsterdam! I will always remember the cosy gatherings you mastered, full of laughs, warmth and delicious food! Thijs and Olesia, thanks for proving that neighbours can also be good friends. Celine, thank you for your unlimited kindness, and for so much fun during the many parties together all around the world! Ale, the failed tandem odyssey will remain a legend forever. Elena Tavkin, your visits were a nostalgic window to the great times in Dresden. Many other people deserve mention here, including Steffen, Joris, Johannes, Simone, the futsal crew, the Python team, Giorgos, Parisa, Jenny, Cristina, Guizela, Clyde, Sjoerd, Giorgio, Douwe, Kevin, Lukas and many more that I hope will forgive my forgetfulness. Support also came from Spain, thanks to Jorge and Pablo for continuously reminding me that I do not study real physics, to Víctor for remaining a close friend to always count on, to

Paco for always caring and specially to Fran, we are proving that 25 years of unique friendship weigh much more than a distance of thousands of kilometers.

Mamá, tu fortaleza y dedicación son una inspiración cada día. Papá, gracias por ser un ejemplo de cómo vivir con optimismo y sentido del humor. Gracias a los dos por hacerme saber que podía llegar lejos, y por apoyarme para conseguirlo. Ojalá sea tan buen padre como lo sois vosotros. Javi, gracias por los chistes con los que sólo papá se ríe, y por las dosis de realidad que son tan necesarias. Tío Jesús, tu saber estar es una inspiración, gracias por cuidar siempre de todos nosotros sin pedir nada a cambio.

While also part of my family, two people deserve a special paragraph. Olga, all these years have been like a dream. Coming to Amsterdam was a difficult decision, but here we have had the happiest and most special moments of our lives. Thank you for your unconditional support and care. Ariadna, you have been in our world for merely half a year, and have completely transformed it. Your smiles can lighten up the darkest days. This thesis is for you both.