STUDYING PROTEIN FOLDING AND PROTEIN-LIGAND INTERACTIONS AT SINGLE-MOLECULE LEVEL USING OPTICAL TWEEZERS

BY

SONAR PUNAM SURESH

THESIS IS SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR OBTAINING THE DEGREE OF

DOCTOR OF PHYLOSOPHY

IN

PHYSICS

THE SCHOOL OF GRADUATE STUDIES

CYCLE XXVII

(2012 – 2015)

UNIVERSITY OF MODENA AND REGGIO EMILIA

UNDER THE GUIDANCE OF

PROF. CIRO CECCONI (SUPERVISOR)

Università degli Studi di Modena e Reggio Emilia

Dipartimento di Scienze Fisiche Informatiche e Matematiche

Via Campi 213/a – 41100 Modena

Tel. +39 0592055243 - Fax +39 0592055235 - http: www.unimore.it

Dottorato di ricerca in ***Physics***

(nell'ambito della Scuola di dottorato in ***Physics and Nano Sciences***)

Ciclo **XXVII**

(titolo tesi)

"Studying protein folding and protein-ligand interactions at single-molecule level using optical tweezers"

Candidato : ***Ms. Sonar Punam Suresh***

Relatore (Tutor): : ***Prof. Andrea Alessandrini***

Correlatore (Tutor) : ***Prof. Ciro Cecconi***

Coordinatore del Dottorato: : ***Prof. Marco Affronte***

Direttore della Scuola di dottorato : ***Prof. Marco Affronte***

Prof. Marco Affronte

Prof. Ciro Cecconi

Dedicated to…

My lovely parents,

Former guide Dr. Vivek Jadhav,

Friends, Pradip Zamre & Sagar Varpe

Acknowledgements

This has been a wonderful journey. It would have not been possible without the guidance and motivations of individuals who helped me in one or another way for its completion. I would like to express my gratitude to all who have inspired, guided and extended their heart whelming full support.

First of all, I would like to express my sincere gratitude to my supervisor Prof. Ciro Cecconi for giving me Ph.D opportunity. His Joy, enthusiasm and dedication to research has always been motivational for me. Thank you very much for your guidance, encouragement, research facilities, support and friendly atmosphere in the lab. Thank you for all your teachings which surely will help me in the future research journey. I could not have imagined having better PI for my Ph.D. Thank you!!

My deepest thanks to my former guide and friend Dr. Vivek Jadhav. Thanks for your guidance from the starting phase of my career. Choosing research as a career, getting Ph.D and survival would have not be possible without your help. Words will be less to thank you for what I am today. Thank to Prof. Suvro Chatterjee, Vascular Biology Lab, MIT Campus of Anna University, India, for giving me an internship opportunity to start my career in this field.

I would like to thank Prof. Birthe Kragelund for giving me the opportunity to learn basic biochemistry in her lab at Department of Biology, University of Copenhagen, Denmark. I am thankful to Dr. Martina Caldarini and Dr. Pétur Heidarsson for providing us samples to carry out optical tweezers experiments.

I would like to thank my former and present colleagues Dr. Immanuel Valpapuram and Dr. Mohsin Naqvi, who willingly helped me in optical tweezers experiments and data analysis. My sincere thanks to Dr. Guido Tiana, Dr. Emanuele Paci, Dr. Luca Bellucci and Dr. Stefano Corni for their MD simulation work in my projects. Foremost, I would like to thank the Department of Physics and CNR Institute of Nanoscience for all conference funding during the tenure of Ph.D.

This Ph.D journey could have not possible without Pradip. Thanks for being motivation force to make this happen, thanks for all your guidance at each step that help me to understand things through a different angles. We have started this journey together and hope will finish it together.

I will like to extend my thanks to my friends Dr. Milind Jadhav, Dr. Manoj and Takshila Tripathi, Vikas Gunge and Wasif Khan for their moral support during my stay in Modena. Being away from family is always difficult, but I was lucky to have Sukhwinder Singh, Ashutosh and Veena Manohar around me, thank you very much for parental care, all festival celebrations and beautiful dinners. Dr. Deepak and Pradnya Banker thank you very much for your help during the early phase of research. Special thanks to Sagar Varpe for your unconditional support in this journey. Sagar thank you very much for always being there for all kinds of help.

My parents deserve special mention for their love and support. Being born in a traditional Indian family, it was not easy to come abroad for education. Mom and dad thank you so much to allow me to break old traditions and encourage me to perceive doctorate. It was your willingness and confidence in me that help me to live my dream. Thanks to my lovely grandma (Aai) and my brothers (Harshal and Ketan Sonar) for supporting me spiritually throughout my life.

Lastly, I am thankful to all of those who supported me in any respect during completion of research journey. My sincere apology to them whom I forget to mention.

Above all thanks to ALMIGHTY for his blessing.

Sonar Punam Suresh

Table of Contents

Abbreviations

OT	Optical tweezers
AFM	Atomic force microscopy
SMFS	Single molecule force spectroscopy
HIV-1-PR	HIV-1-protease
HIV	Human immunodeficiency virus
AIDS	Acquired immunodeficiency syndrome
N	Native state
I	Intermediate state
TS	Transition state
PF	Partially folded
NMR	Nuclear magnetic resonance
ACBP	Acyl-CoA binding protein
CoA	Coenzyme A
MD	Molecular dynamics
PDB	Protein data bank
RMD	Ratcheted molecular dynamics
WLC	Worm-like chain

Preface

Proteins are the fundamental units of life as they perform a wide variety of roles and affect virtually every property that characterizes a living organism. In order to function properly, protein has to be in their correct three-dimensional structure. If protein folding goes wrong then protein becomes useless and often harmful for cell, leading to various diseases such as Parkinson's diseases, Alzheimer's disease, Prion disorder and other neurodegenerative disorders. Thanks to rapid developments in genome sequencing, we can now identify gene responsible for the disease. However to understand the molecular basis of protein misfolding diseases a deep understanding of the basic rules governing the protein folding is required. Despite the theoretical and experimental efforts of many laboratories in the last 50 years, our comprehension of the mechanism by which protein attains their native structure is still incomplete. The advent of single-molecule manipulation techniques such as optical tweezers has allowed us to revisit protein folding with a completely new approach.

One approach to study protein folding is to use traditional ensemble techniques, such as nuclear magnetic resonance, circular dichroism, hydrogen/deuterium exchange mass spectroscopy. Although these techniques proven very informative, they are limited to the description of the overall properties of a large population of the molecules; the output of these measurements is the ensemble average generated from a large, and often dephased, population of molecules, in which time dependent dynamics of the individual molecules as well as rare but potentially important molecular events are hidden. These technical limitations have been overcome using single-molecule spectroscopy techniques which follow real time trajectories of single molecules, and study kinetics and thermodynamics of bimolecular reactions, also detect possible intermediates. Due to these advantages of single-molecule field, the main objective of my thesis was to go beyond traditional ensemble techniques and study the protein folding at single-molecule level. This thesis reports the application of dual-beam optical tweezers and molecular dynamics simulation to the study of protein folding and protein-ligand interactions. The focus of this dissertation was to understand the complex folding behavior of HIV-1-protease and to study the effect of ligand binding on mechanical properties of Acyl-CoA-binding protein. The thesis is organized as follows.

Chapter 1 explains the basic principles of force spectroscopy techniques, with a focus on optical tweezers, describe some of the theoretical models used to analyze and interpret single-molecule manipulation data. It also highlights on some recent and exciting results that have emerged from single-molecule research field on protein folding and protein-ligand interactions.

Chapter 2 gives the brief introduction of optical tweezers, its working principle and dual-beam optical tweezers setup at the University of Modena and Reggio Emilia. In the second part of this chapter, there is a brief description of model systems that I have used for my doctoral studies.

Retroviral protease from human immunodeficiency virus type 1 (HIV-1-PR) is essential for the maturation of the virus and therefore, it has been identified as potential targets for anti-AIDS therapies. Since it plays an essential role in the life cycle of the HIV virus, many drugs has been designed so far to inhibit the activity of HIV-1-PR, but all these attempts have failed. One possible reason for these failures could be a poor understanding of the molecular behavior of this protein. Folding mechanism of HIV-1-PR has already been studied by traditional bulk techniques. These studies had given the information of overall thermodynamics and kinetics of the process, but they failed to provide insight into the ensemble of folding routes that individual molecules follow to reach native state. To shed light on the molecular behavior of HIV-1-PR we decided to study unfolding and refolding processes of this protein using optical tweezers. With the help of molecular dynamics (MD) simulations, we gain the atomistic details of these processes. This study has been described in *chapter 3*.

Acyl-CoA-binding protein (ACBP) is involved in various biochemical processes such as signal transduction, regulation of lipid and energy metabolism, gene regulation and many more. It is well-known that ACBP binds to the acyl CoA esters with a broad range of acyl-chain length (C_4-C_{22}). This ligand binding may affect the molecular processes of ACBP. Using OT and MD simulations, we have investigated the effects of ligand binding on unfolding and refolding processes of ACBP. *Chapter 4* explains different methods such as force-ramp, force-jump we have used to extract kinetic parameters of ACBP. The results obtained from our studies were compared with the previous ACBP alone (without ligand) single-molecule study.

Finally, *chapter 5* summarizes the conclusions of this thesis.

Chapter: 1

Conformational dynamics of single protein molecules studied by direct mechanical manipulation

Chapter 1: Conformational dynamics of single protein molecules studied by direct mechanical manipulation

Abstract

Advances in single-molecule manipulation techniques have recently enabled researchers to study a growing array of biological processes in unprecedented detail. Individual molecules can now be manipulated with sub-nanometer precision along a simple and well-defined reaction coordinate, the molecular end-to-end distance, and their conformational changes can be monitored in real time with ever improving time resolution. The behavior of biomolecules under tension continues to unravel at an accelerated pace and often in combination with computational studies that reveal the atomistic details of the process under investigation. This chapter exclusively focuses on mechanical manipulation methods, specifically optical tweezers and atomic force microscopy (AFM), and illustrate the main experimental strategies used in these techniques. Briefly, it highlights on some pivotal single-molecule studies aimed at solving exciting biological problems, also describes some theoretical models used to analyze single-molecule force spectroscopic data. At the end of this chapter, the results that have emerged from single molecule force spectroscopy on protein folding and protein-ligand interactions has been reviewed.

Keywords: Single-molecule force spectroscopy, OT, AFM, Protein folding.

This work is published in Advances in protein Chemistry and Structural Biology, 92 (2013) 93-133.

Introduction

Sophisticated bulk methods have been developed to characterize the conformational changes of proteins as they carry out their biological functions. Traditional techniques such as circular dichroism, nuclear magnetic resonance (NMR) spectroscopy and hydrogen/deuterium exchange mass spectrometry [1-3] have been extensively used to study the structure and dynamics of proteins, both free in solution or bound to their molecular partners. These bulk studies have been very informative but limited to the description of the overall properties of a large population of proteins. This is because the output from these measurements is the ensemble average generated from a large, and often dephased, population of molecules, in which the time-dependent dynamics of the individual molecules, as well as rare but potentially important molecular events, are hidden. These technical limitations, which have restrained our ability to decipher the intricacies of many molecular processes, have recently been overcome with the advent of single-molecule methods. These novel experimental approaches enable us to follow the real-time trajectories of single molecules and describe the inherent heterogeneity of biological processes that are stochastic in nature. Briefly, these single-molecule experimental techniques are discussed below. The contents of this chapter are taken from [4] with permission.

Mechanical manipulation of single protein molecules

Just as chemicals and heat can be used to perturb the system under study, so can force be used for the same purpose. Force and the end-to-end distance of a protein molecule constitute the main variables during a mechanical manipulation experiment. Many methods have been developed to directly manipulate biomolecules, among them optical tweezers and AFM are the most popular experimental methods. Briefly, these techniques are explained below.

Optical tweezers

Ever since the pioneering work of Arthur Ashkin in the 1980´s [5], optical tweezers have continued to evolve and improve to tackle ever more complex systems. An optical trap can be formed by focusing a collimated beam of light through a microscope objective with a high numerical aperture. In this way, small objects of high refractive index can be optically trapped and manipulated. For a detailed description of the principles behind optical trapping, which is beyond the scope of this

chapter, the reader is referred to previous publications [5-7]. There are three main geometries that are nowadays used in optical tweezers setups *(Figure 1-1)*. In all cases the molecule under study is tethered between an optically trapped bead and: i) a substrate, ii) a second bead held at the end of micropipette by suction, or iii) another bead held in a second optical trap. In all cases, the force applied to the molecule is modulated by varying the distance between the two tethering points.

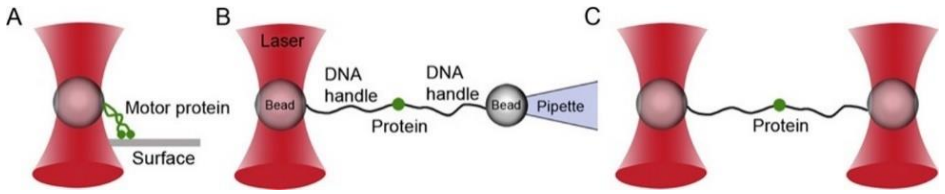


Figure 1-1 Optical tweezers experimental geometries. A) A single-beam optical tweezers setup where the molecule of interest is tethered between an optically trapped bead and a surface. The movements of the protein along the surface are revealed by the motions of the bead in the trap. B) A single-beam optical tweezers experimental setup where the protein is attached to two polystyrene beads through two double stranded DNA handles. One bead is held in the optical trap while the other is held at the end of a micropipette by suction. The micropipette can be mechanically moved relative to the trap, to induce the unfolding or refolding of a protein molecule. C) A double-beam setup where the protein molecule is tethered through DNA handles between two optically trapped beads. Adapted from [4] with permission.

Different approaches can be taken to manipulate a molecule. In what follows is a brief description of the experimental strategy used in: i) constant-velocity (also force-ramp), ii) constant-force (also force-clamp), iii) passive-mode, and iv) force-jump experiments. In constant-velocity experiments the force is increased and relaxed at constant speed (nm/s), to obtain force vs. extension cycles as shown in *figure 1-2A*. During stretching, the force is raised until the molecule is observed to unfold. This event is marked by a sudden increase in the extension of the molecule, as it goes from its compact native state (N) to a more extended unfolded state (U). During relaxation, the molecule is typically observed to refold around at ~ 5-10 pN, through a sharp transition that restores the original extension of the molecule. The changes in contour length of the protein associated with the unfolding and refolding events can be estimated by fitting the force-extension traces with the worm-like chain (WLC) model [8-10]. Constant-velocity experiments can provide information on both kinetics and thermodynamics of a protein folding

reaction. Kinetic parameters, such as rate constants and position of the transition state along the reaction coordinate, can for example be estimated by analyzing force distributions. Thermodynamics information (e.g. unfolding free energy) can instead be recovered by analyzing irreversible work distributions, *see section "Theoretical models of single-molecule force spectroscopy" for details*. In constant-force experiments the force applied to the molecule is kept constant through a feedback mechanism, while changes in the extension of the molecule are monitored over time *(Figure 1-2B)*.

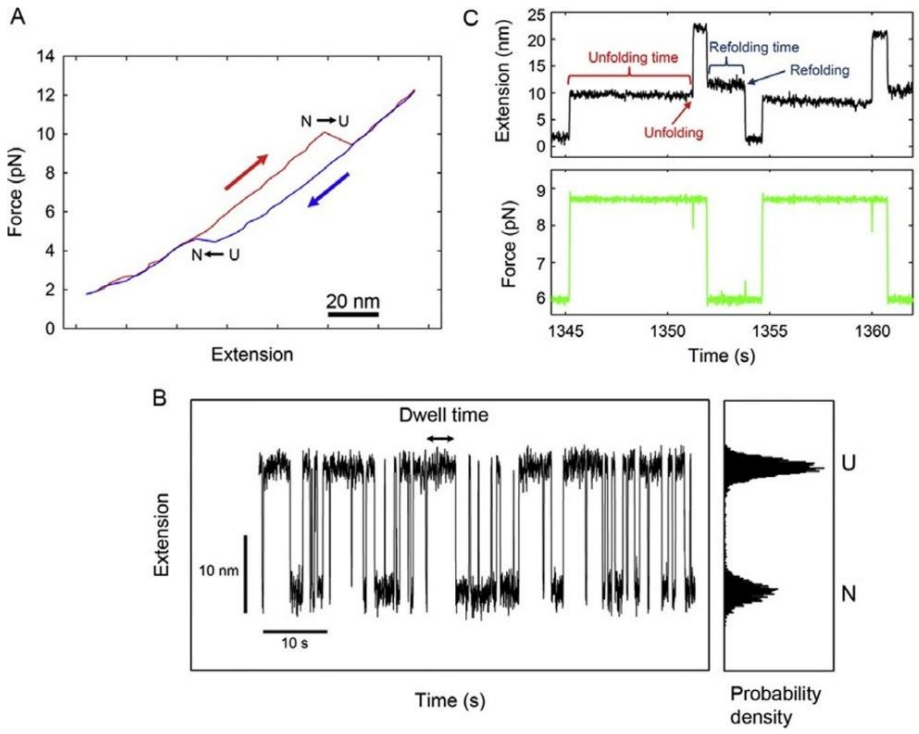


Figure 1-2 Basic mechanical manipulation experiments using optical tweezers. A) Constant-velocity experiment showing a two-state unfolding/refolding event. The arrows indicate the pulling direction. B) Constant-force experiment showing a molecule fluctuating between an unfolded *(U)* and folded *(N)* state. The dwell times of the unfolded and folded states contain both thermodynamic and kinetic information. A) and B) adapted from [4] with permission. C) Force-jump experiment. The force is rapidly jumped between two different set values, to increase the probability of observing unfolding/refolding events, see text for details. Adapted from [11] with permission.

In these measurements, rate coefficients can be obtained directly from the lifetimes of the folded and unfolded states, and free energies can be calculated from the ratio of the kinetics coefficients. As the molecule unfolds or refolds, the force is kept constant by reducing or increasing the distance between the tethering surfaces, respectively. These movements, which are controlled by the feedback mechanism, can however take place only at a certain rate and, because of this constant-force measurements can sometimes provide misleading results. This is the case for example when the response rate of the feedback mechanism is slower than rate at which the molecule can fluctuate. In these cases in fact, short-lived transitions are missed, leading to average dwell times biased towards larger values [12]. This potential instrumental artifact can be avoided, or at least reduced, by studying molecular fluctuations through passive-mode measurements. In this case, no feedback is employed. Instead, the distance between the tethering surface and the optical trap is kept constant, while force is allowed to increase or decrease as the molecule unfolds or refolds, respectively. In these experiments, the response time of the system is dictated by the corner frequency of the trapped bead, which is typically much higher than the response frequency of a feedback mechanism. Consequently, passive-mode measurements are better suited to study rapid conformational transitions. On the other hand, as no feedback mechanism is used, reliable passive-mode measurements require high mechanical and optical stability of the instrument itself.

Constant-force or passive-mode measurements can be effectively used to study molecular fluctuations at equilibrium only when, in a certain range of forces, the rate of both the forward and reverse reactions is high enough to allow the acquisition of a large number of events in a relative short amount of time. This however is not always the case, as for some molecules unfolding and refolding occurs at quite different forces. In these instances, at any given force, either the unfolding or refolding rate is so low that the acquisition of a significant number of events would require very long recordings. To overcome this problem, force-jump experiments can be performed. In a force-jump experiment the force is increased (jumped) or decreased (dropped) quickly to a preset force value and kept constant with a feedback mechanism until an unfolding or refolding event is observed *(Figure 1-2C).* These experiments allow the direct measurement of rate constants in force ranges where the probability of observing either an unfolding or refolding event is high [13].

Alternative, novel and hybrid optical tweezers instruments

Single-molecule manipulation instruments are advancing at a rapid pace. Considerable effort has been spent on designing optical tweezers that measure other single-molecule parameters such as

rotation or torque in combination with force and extension [14]. The ability to measure torque has significant importance because it addresses many aspects of cell and protein biology such as transcription, replication, recombination and protein folding [15]. The measurement of torque can be achieved with magnetic tweezers, a popular instrument due to its simple and cheap [16]. Spatial and temporal resolution of magnetic tweezers may in the near future advance to sub-nanometer precision, possibly with newly emerging camera technology that enhances the rate of data acquisition [16].

The integration of dual beam optical tweezers with magnetic tweezers allows manipulation of a single molecule with nanometer precision. This hybrid instrument has been used to investigate a wide range of complex system such as higher order chromatin interaction, DNA-DNA interaction mediated by proteins, measurement of intermolecular friction and localization, and binding strength analysis of DNA bound proteins [17]. The combination of magnetic or optical tweezers and fluorescence microscopy is a powerful tool for DNA manipulation and has been used to study DNA supercoiling, the dynamics of diffusion, hopping of plectonemics in DNA, and the torque and twist DNA-breathing dynamics [18, 19].

A recently designed Quad-trap optical tweezers instrument was used to study the condensation of bacterial chromosome DNA [20]. The instrument was able to independently trap four polystyrene beads at the same time, allowing the simultaneous mechanical manipulation of two independent DNA molecules. With this method, the authors were able to detect the complex and dynamic interactions of DNA and histone-like nucleoid structuring protein (H-NS), which is involved in mediating DNA-DNA contact. This remarkable technical feature helped explain the mechanism and role of H-NS in chromosomal DNA condensation.

Sample preparation

A major issue in optical tweezers experiments is to find an efficient method to manipulate the molecule of interest. Biomolecules such as proteins and RNAs are typically too small to be directly manipulated with micrometer-sized optical tweezers beads, as the tethering surfaces would come so close to each other that they would interact. To avoid these unspecific and unwanted interactions, a method has been developed that relies on the use of two DNA molecular handles *(Figure 1-1B and C)* [21]. One end of each handle (~ 500-1000 bp DNA molecule) is attached covalently to the side chain of a cysteine residue. The other end is attached to a polystyrene bead through either biotin-streptavidin interactions, or digoxigenin/antibody interactions. The handles

act as spacers between the protein and the beads to avoid unwanted interactions between the tethering surfaces that would compromise the experiment. This optical tweezers manipulation method was employed for the first time in 2005 [9], and it is now used by a growing number of laboratories around the world [22-26]. For details on how the DNA-protein coupling reaction is performed, reader can refer [27].

Atomic force microscopy

Atomic force microscopy (AFM) was initially developed for high-resolution imaging of surface contours of microscopic samples [28] and it still primarily serves this function. Later, AFM has also evolved into a versatile technique to manipulate single molecules and characterize their mechanical properties, in a mode of operation called "force spectroscopy" or "force measuring" mode. AFM force spectroscopy has been used to study a number of biological systems, including binding of antibodies to their antigens [29], ligand to receptors [30], binding forces of complementary DNA strands [31], conformational changes in biological polymers [32], and to study the mechanical properties of a large variety of protein molecules [33-36].

When AFM is used to study protein folding, the molecule of interest, typically a polyprotein made of a linear array of a globular domain, is tethered between a flat surface, usually made of gold, and a silicon nitrite AFM tip *(Figure 1-3A)*. One end of the polyprotein interacts with the gold surface through thiol groups of terminal cysteine residues, while the other end adheres to the tip unspecifically. The polyprotein is then stretched and relaxed by modulating the distance between the tethering surfaces, while the force applied on the molecule is determined by measuring the deflection of the cantilever. As the tip is pulled away from the surface the tension along the polymer increases until one globular domain stochastically unfolds. Upon unfolding, the contour length of the polymer suddenly increases, generating a sharp drop of the force on the cantilever. As the stretching of the polymer continues, all the other domains subsequently unfold generating a series of sudden force drops that give rise to a peculiar saw-tooth like pattern in the force-extension curve, where each peak corresponds to the unfolding of one domain *(Figure 1-3B)*. The rising phase of each peak reflects the elastic properties of the stretched polymer, while the distance between peaks is proportional to the number of amino acid constituting the protein domain. As in optical tweezers experiments molecular handles are necessary to keep the two polystyrene beads away from each other, in AFM studies the need for polyproteins comes from

the necessity of keeping a certain distance between the tip and the substrate. At short distances, in fact, tip-surface interactions could become the dominant features of a force trace. Different methods have been devised to generate a linear array of a globular domain, as described in [21, 37-40]. AFM force spectroscopy experiments on polyproteins can be performed at different pH and ionic strengths, as well as in the presence of other molecules to study the effect of binding partners on the energetics of the protein [41, 42].

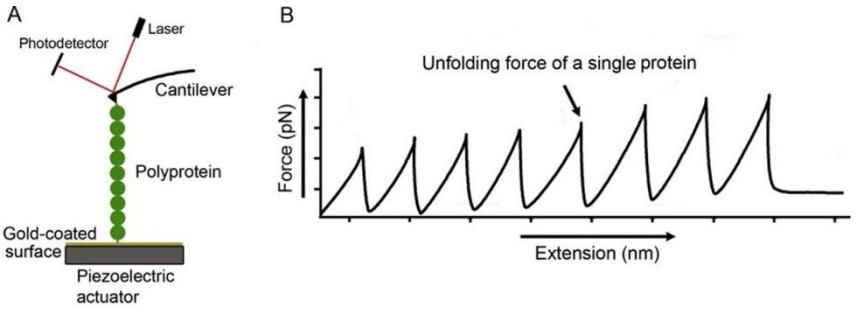


Figure 1-3 An AFM experiment. A) In a typical AFM experiment polyprotein constructs are picked up and unfolded by the cantilever tip. The position of the tip is determined by a photodectector from the deflection of a laser beam. B) The sequential unfolding of single proteins within the polyprotein construct results in a characteristic sawtooth pattern in the force vs. extension trace. Adapted from [4] with permission.

AFM and optical tweezers are complementary techniques for the study of the mechanical properties of biomolecules. An AFM cantilever has a higher root-mean square (RMS) force noise (~15 pN) than an optically trapped bead (~0.1 pN). As a consequence, optical tweezers is the technique of choice to study processes that take place at low forces, such as the refolding of a protein that typically occur in a 5-10 pN force range. In addition, the shallower harmonic potential of an optical trap compared to an AFM cantilever, makes laser tweezers a better technique to observe a molecule hopping between different molecular states. In fact, the steeper is the harmonic potential, the larger is the kinetic barrier that the system molecule-force transducer must overcome to hop between a folded and unfolded state of the molecule. On the other hand, optical tweezers are usually unable to measure forces that are larger than 100 pN, while AFM can measure forces even in the nanonewton range. It follows that AFM is the technique of choice to measure the rupture force of single covalent bonds (~2 nN) [43], the force at which polysaccharides switch to

different conformations (~270 pN) [44], or to characterize the mechanical properties of proteins that unfolds at high forces [45].

Theoretical models of single-molecule force spectroscopy

The mechanical manipulation of single molecules using optical tweezers and AFM can yield valuable information about the free energy surface of a single-molecule reaction. In this section, using basic thermodynamic and kinetic principles, the effect of mechanical force on the energy landscape of molecule is explained. Then, some theoretical models used to analyze and interpret single-molecule manipulation results are discussed.

Effect of force on the thermodynamics of a single-molecule reaction

The effect of force on a two-state reaction in which *A* is converted into *B* is depicted in *figure 1-4*. Along the mechanical reaction coordinate i.e extension of the molecule, states *A* and *B* occupy free energy minima separated by a distance Δx.

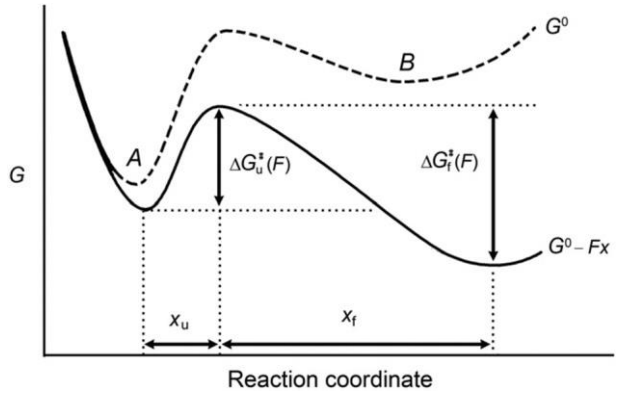


Figure 1-4 Energy landscape for a two-state unfolding reaction. The dotted-line curve is the free energy surface at force = 0 and the solid-line curve at force = *F*. The force *F* tilts the energy landscape by a factor *Fx* where *x* is the extension of the molecule. Adapted from [4] with permission.

The free energy difference between *A* and *B* at zero force is

$$\Delta G(F=0) = \Delta G^0 + k_B T ln\left(\frac{[B]}{[A]}\right) \qquad (1-1)$$

Where ΔG^0 is the standard state free energy, k_BT is the thermal energy, and [A] and [*B*] give the probabilities of populating states *A* and *B* in single-molecule experiments. To a first approximation, when a force *F* is applied to a molecule each point of its energy landscape is lowered by an amount equal to $F\Delta x$, where Δx is the distance between the point of interest and the native state. As a consequence, an applied force tilts the free energy surface along the mechanical reaction coordinate [46], such that

$$\Delta G(F) = \Delta G^0 - F(\Delta x) + k_B T ln\left(\frac{[B]}{[A]}\right) \qquad (1-2)$$

At equilibrium, $\Delta G = 0$ and

$$\Delta G^0 = -k_B T ln K_{eq}(F) + F\Delta x \qquad (1-3)$$

Equation (1-3) holds true only if the position of *A* and *B* along the reaction coordinate are unaffected by force. For most reactions, however, this is not true. Let's consider for example the case of a protein that under tension transits from its native state to its unfolded state. In this case the extension of the native state, and thus its position along the reaction coordinate, can to a good approximation be considered unchanged, but the extension of the unfolded state instead will significantly increase. This increment in the average end-to-end distance of the unfolded state corresponds to a change in the free energy of the system, as

$$\Delta G^0 = -K_B T\, ln\, K_{eq}\,(F) + F\Delta x - \Delta G_{stretch}(F) \qquad (1-4)$$

Where $\Delta G_{stretch}$ *(F)* is the free energy change due to stretching of the unfolded state at force *F*. When a molecule is manipulated at close to equilibrium conditions, it unfolds and refolds through transitions that take place around a force ($F_{1/2}$) at which the molecule has equal probability of being either in its folded or unfolded state. At $F_{1/2}$, $K_{eq} = 1$ and

$$\Delta G^0 = F_{1/2}\Delta x - \Delta G_{stretch}\left(F_{\frac{1}{2}}\right) \qquad (1-5)$$

Under these experimental conditions, $F_{1/2}\Delta x$ can be calculated as the area under the unfolding/refolding rips observed in force vs. extension curves. This free energy can then be compared with that measured in bulk measurements after subtraction of $\Delta G_{stretch}$ $(F_{1/2})$, which can

to a good approximation be calculated as the area under the WLC force-extension curve integrated from zero to the extension of the unfolded molecule at $F_{1/2}$ [9, 10].

Effect of force on the kinetics of a single-molecule reaction

As we know from transition state theory the rates of unfolding (k_u^0) and refolding (k_f^0) of a molecule at zero force are

$$k_u^0 = A\, exp\left(\frac{-\Delta G_{\,u}}{k_B T}\right) \qquad (1-6)$$

$$k_f^0 = A\, exp\left(\frac{-\Delta G_{\,f}}{k_B T}\right) \qquad (1-7)$$

Where A is the natural frequency of oscillation, ΔG_u and ΔG_f are the activation energies for unfolding and refolding, respectively. In the presence of force the unfolding (ΔG_u) and refolding (ΔG_f) activation energies will be lowered by an amount equal to $F\cdot\Delta x_u$ and $F\cdot\Delta x_f$, respectively, where Δx_u and Δx_f are the distances to the transition state from the unfolded and folded states [47]. It follows that the unfolding and refolding rates at force F are given by:

$$k_u(F) = k_u^0\, exp\left(\frac{F\,\Delta x_{\,u}}{k_B T}\right) \qquad (1-8)$$

$$k_f(F) = k_f^0\, exp\left(\frac{-F\,\Delta x_{\,f}}{k_B T}\right) \qquad (1-9)$$

From the ratio of $k_u(F)$ and $k_f(F)$, the equilibrium constant can be calculated as:

$$K_{eq}(F) = K_{eq}^0\, exp\left(\frac{F\,\Delta x}{k_B T}\right) \qquad (1-10)$$

Where K_{eq}^0 is the equilibrium constant at zero force and $\Delta x = \Delta x_u + \Delta x_f$. It is worth pointing out that in the above kinetics models the position of the transition state along the reaction coordinate is considered to be independent from the applied force. This approximation is usually correct when we consider very narrow range of forces. For more general cases, however, force-induced shifts in the position of the transition state must be accounted for. To this end, several improved models have been proposed and applied to different systems, as for example in [48-50].

Extracting kinetic parameters from force distributions

In constant-velocity optical tweezers experiments *(Figure 1-2A)*, above 3-4 pN the rate at which the force is applied on the molecule, loading rate *r* (dF/dt in units of pN/s), typically becomes constant [9, 10]. This allows us to analyze experimental data with analytical models, as explained below.

For a first order reaction with negligible refolding rate, the time dependence of the probability that the molecule has not unfolded is [51]:

$$\frac{dP_f(t)}{dt} = -k_u(t)P_f(t) \qquad (1-11)$$

If force varies linearly with time *t* as $F = r^* t$, where *r* is the loading rate, then variable *t* can be changed to *F* in the above equation as

$$\frac{dP_f(F)}{dF} = -\frac{k_u(F)}{r}P_f\,(\mathrm{F}) \qquad (1-12)$$

Integrating the above equation from 0 to *F* and using *equation (1-8)* we get

$$ln\{P_f(F)\} = \frac{k_u^0 k_B T}{r x_u}\left(1 - exp(F x_u/k_B T)\right) \qquad (1-13)$$

From *equation (1-13)*, the probability of unfolding as a function of force ($P_u(F)$) can be derived as

$$P_u(F) = 1 - P_f(F) = 1 - exp\left(-\frac{k_u^0 k_B T}{r x_u}(exp(F x_u/k_B T) - 1)\right) \qquad (1-14)$$

By differentiating *Pu(F)*, which is a sigmoidal function of force, we can calculate the probability density as

$$\frac{dP_u}{dF} = \frac{k_u^0}{r}exp(F x_u/k_B T) * exp\left(-\frac{k_u^0 k_B T}{r x_u}(exp(F x_u/k_B T) - 1)\right) \qquad (1-15)$$

Following similar steps the probability density function for the refolding force can be calculated as:

$$\frac{dP_f}{dF} = \frac{k_f^0}{r} exp\left(-Fx_f/k_BT\right) * \frac{exp\left(-\frac{k_f^0 k_B T}{r x_f}\left(\exp(-\frac{Fx_f}{k_B T})-1\right)\right)}{exp\left(\left(-\frac{k_f^0 k_B T}{r x_f}\right)-1)\right)} \qquad (1-16)$$

Normalized distributions of the unfolding and refolding forces of a molecule manipulated at constant loading rate can be fit to *equations (1-15) and (1-16)* in order to estimate rate constants at zero force (k_u^0, k_f^0) and distances to the transition state (x_u, x_f) [11, 52]. Often however, experimental force distributions are analyzed with a slightly different method. When $\exp(Fx_u/k_BT) > 10$, which is usually the case with biomolecules, *equation (1-13)* can be written as

$$ln\{P_f(F)\} = -\frac{k_u^0 k_B T}{r x_u}\left(exp(Fx_u/k_BT)\right) \qquad (1-17)$$

Which can then be linearized as

$$ln\{r\, ln(1/P_f(F))\} = ln\frac{k_u^0 k_B T}{x_u} + (x_u/k_BT)\, F \qquad (1-18)$$

Through similar considerations, for the refolding process, we have

$$ln\{-r\, ln(1/P_u(F))\} = ln\frac{k_f^0 k_B T}{x_f} + (x_f/k_BT)\, F \qquad (1-19)$$

Equations (1-18) and (1-19) are often used to fit *ln[rln[1/N]]* and *ln[-rln[1/U]]* vs. force graphs, where *N* and *U* are the folded and unfolded fractions, respectively, which are calculated by integrating the histograms of the force distributions over the corresponding range of forces [10, 13, 53].

Extracting thermodynamic parameters from non-equilibrium measurements

When the rate at which force is applied on a molecule is faster than its slowest relaxation rate, the unfolding and refolding processes occur out of equilibrium. Under these experimental conditions, the free energy change of the process is not equal to the work done on the molecule, and cannot be calculated as the area under the unfolding/refolding transitions. However, fluctuation theorems have been developed to extract equilibrium information from non-equilibrium measurements [54]. The first of these theorems to be successfully applied to single-molecule experimental data was

presented by *Jarzynski* [55]. He derived an equality that relates the free energy difference $\Delta G(z)$, separating states of a system at positions 0 and *z* along a reaction coordinate, to the work done to irreversibly switch the system between the two states:

$$\exp[-\beta \Delta G(z)] = Lim N_{N_{\rightarrow}\infty} < \exp[-\beta\, wi(z,r)] >_N \qquad (1-20)$$

Where < > denotes averaging over *N* work trajectories, $w_i(z,r)$ represents the work of the *i-th* of *N* trajectories, and *r* is the switching rate. The first application of this method in single-molecule force spectroscopy was reported by [56]. They manipulated a P5ab RNA hairpin out of equilibrium with optical tweezers and applied Jarzynski's equality to the irreversible work trajectories to extract the unfolding free energy of the molecule. Although effective for near-equilibrium processes like that observed for the P5ab RNA hairpin, in far from equilibrium systems Jarzynski's equality is hampered by large statistical uncertainties due to the exponential averaging of low work values [57]. These large uncertainties can be significantly reduced by using Crook's fluctuation theorem (CFT). CFT relates the amount of work done on a molecule that unfolds and refolds out of equilibrium, with the free energy change of the process [58]. Let $P_U(W)$ and $P_R(W)$ denote the probability distributions of the work performed on a molecule that is pulled (U) and relaxed (R) an infinite number times. The CFT then predicts that

$$\frac{P_U(W)}{P_R(W)} = \exp\left(\frac{W - \Delta G}{k_B T}\right) \qquad (1-21)$$

Where ΔG is the free-energy change between the final and the initial states of the molecule, and thus equal to the reversible work associated with this process. The value of ΔG can be determined as the point of intersection of the unfolding and refolding work distributions, where $W = \Delta G$ *(Figure 1-5)*. When the overlapping region of the two distributions is small because the unfolding/refolding process occurs very far from equilibrium, Bennett's acceptance ratio method [59] is often used to reduce the uncertainty in the estimation of ΔG. CFT has already been applied in several single molecule manipulation studies, as in [23, 60, 61].

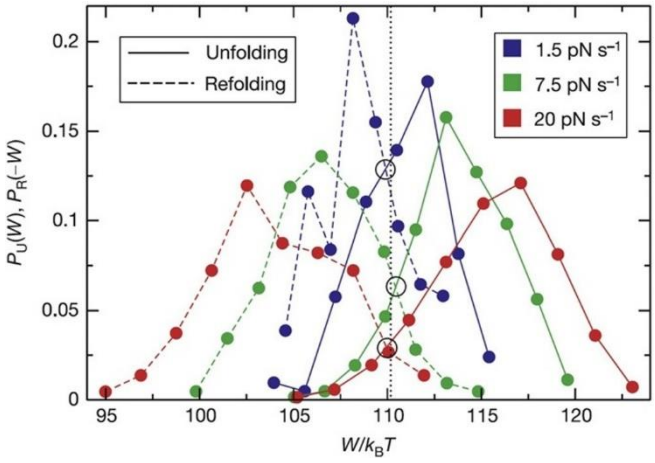


Figure 1-5 Work distributions $P_U(W)$ and $P_R(W)$ for RNA unfolding (solid curve) and refolding (dash curve) at different loading rates. The point of intersection of the two distributions gives $\Delta G = 110.3 k_B T$. Adapted from [60] with permission.

Extracting kinetics and thermodynamic parameters from equilibrium fluctuations

As discussed above, in constant-force and passive-mode measurements some molecules can be observed to fluctuate at equilibrium between different molecular conformations *(Figure 1-2B).* Although different theoretical approaches could be used, hidden Markov models (HMMs) are often the methods of choice to extract kinetics and thermodynamic information from these experimental data [12, 22, 62-64]. HMMs are powerful statistical tools introduced by Baum and colleagues in the late 1960s and early 1970s [65, 66], and were later implemented for speech processing applications by Baker [67]**.** *Figure 1-6* shows a general HMM, where the X_i represents the hidden state sequence, O_i the observation sequence, and A and B are the matrices of state transition and observation probabilities, respectively. HMM predicts the most likely state sequence X_i which has the maximum probability to give the observation sequence O_i by constructing a model using *A* and *B* matrices [68].

In the analysis of single-molecule force spectroscopy data, the HMM assumes that the observation sequence (force or extension trace) is generated by a Markov process in which the molecule makes history independent transitions governed by a transition matrix among its different conformational states. At a given force, each state of a molecule can be defined by the distribution of extension or force values. The HMM analysis provides the transition probability matrix *T* that can be used to calculate the matrix for rate constants *K* using the relation $T = exp(K\Delta t)$, where Δt

is the data acquisition time [69]. Using this method, rate constants can be calculated at different forces (*F*) and plots of *lnk* vs. *F* can be fit to linearized forms of *equations (1-8) and (1-9)* to estimate the distances to the transition state and the rate constants at zero force. The ratio of the rate constants (K_{eq}) will provide information about the free energy of the reaction according to *equation (1-4).*

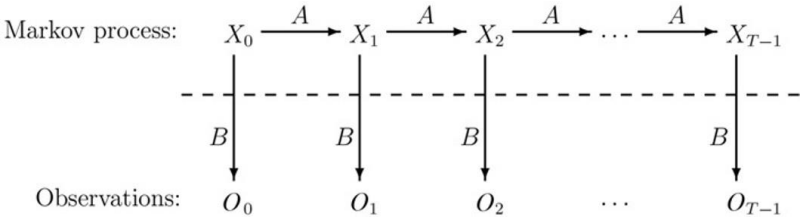


Figure 1-6 A general HMM where the hidden state sequence X_i is related to the observation sequence O_i by observation probability matrix *B* and transition probability matrix *A*. Adapted from [68] with permission.

With the advent of new experimental strategies, novel theoretical models have also been developed and applied to different systems. Reader can refer [48, 70, 71].

Biological Applications

Mechanical processes in the cell

A wide variety of cellular processes involves molecules that generate or are subjected to mechanical forces [46]. Molecular motors such as kinesin convert chemical energy into mechanical work and movement, usually through ATP hydrolysis [72]. Mechanosensors go through subtle conformational changes in response to mechanical stimuli to initiate signaling cascades [73]. Proteins can also be unfolded by force, for example by translocases, to target them for degradation or transport across membranes [74, 75]. Conversely, proteins perform work, and thus measurable force, through the compacting of their polypeptide chains during spontaneous folding. Following section briefly describes some recent studies on the conformational dynamics involved in protein folding and protein-ligand or protein-protein interactions.

Protein folding

Folding into a three-dimensional form is pivotal for the function and specificity of most proteins. The early classical experiments of *Christian B. Anfinsen* showed that a protein can spontaneously

fold into thermodynamically stable states *in vitro* and thus that all the necessary information for the process is encoded in the amino acid sequence [76]. Far from being a simple task, especially in the case of larger and of multi-domain proteins, the mechanism of protein folding has been studied extensively over many decades. Yet, our understanding of this complex process is still incomplete. Proteins fold within a biologically relevant timescale of microseconds to seconds despite the astronomical amount of conformations available to them, a paradox so famously stated by *Cyrus Levinthal* [77, 78]. The paradox, which highlights the fact that protein folding would be impossible with a random search through conformational space, ultimately through simplified models and pathways, led to the idea of folding funnels [79-83]. In this view, the folding of protein molecules is described as diffusion of a statistical ensemble over a funnel-shaped energy landscape with possibly numerous parallel pathways *(Figure 1-7A)* [80, 84].

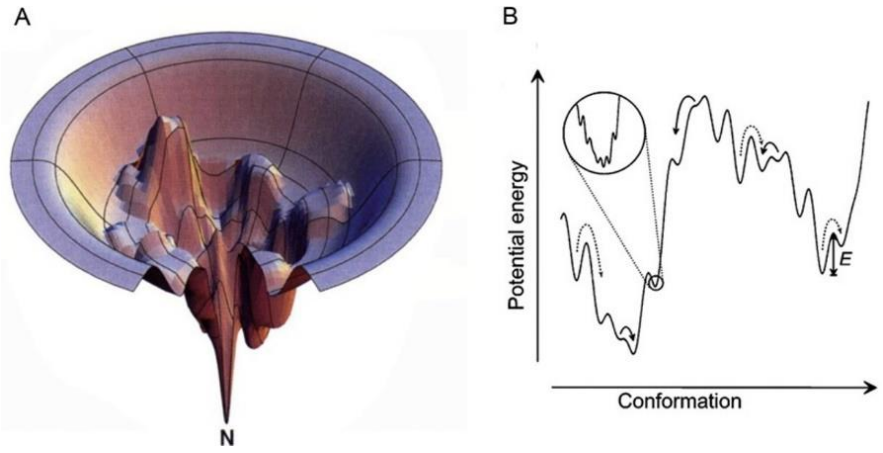


Figure 1-7 The funnel-shaped energy landscape for protein folding. A) Proteins are currently thought to fold to their native state by diffusion over a multi-dimensional funnel-shaped energy landscape. Adapted from [80] with permission. B) Two-dimensional scheme of the ruggedness of an energy landscape. Two main conformations are accessible with multiple smaller kinetic traps that are either short-lived (solid arrows) or long-lived (dotted arrows) conformations. E represents the height of the activation barrier for one such conformation. Adapted from [85] with permission.

The funnel shape provides an energetic bias towards folding into the native state and this can be described with two seemingly counteracting properties; a decrease in 1) configurational entropy (which is unfavorable) and 2) potential energy (which is favorable) [86]. The funneled energy landscape contains all the information to describe the path a protein can take on route to its native state. Theory predicts that a rugged energy landscape, with many small minima, will lead

to slower folding *(Figure 1-7B)* and this was recently demonstrated experimentally with spectrin domains [87, 88].

The ensemble view suggests that due to the ruggedness of the energy landscape a fraction of the molecules will fold slowly, due to trapping in energy minima, while other molecules will fold fast without populating intermediate states. This phenomenon is known as kinetic partitioning and explains the multi-phasic kinetics observed in many systems [89]. Kinetically trapped intermediate structures visited along a folding trajectory may be on-pathway, and thus productive towards the native state, or off-pathway misfolded states, which can form aggregates and lead to serious diseases *(see section "Intermediate states, misfolded states, and beyond").*

Traditional bulk techniques, such as NMR and protein engineering phi-value analysis [90], have unlocked a wealth of knowledge on the folding process but these methods are limited by an ensemble-averaged experimental output. Computational approaches to folding have also significantly enhanced our understanding but despite recent reports of millisecond-long simulations [91] they have mostly been limited to probing a relatively short timescale using small, fast folding proteins [92]. Single-molecule mechanical manipulation methods have added a new dimension to protein folding studies. The early mechanical manipulation instruments were only able to directly detect the unfolding process of single molecules, leaving the refolding process silent. The first of these experiments were performed simultaneously with AFM and optical tweezers on the giant muscle protein titin [93-95]. Quickly the field developed and the refolding of a single molecule into its native state was observed [9, 96]. In the following years came a surge of exciting results, establishing single-molecule force spectroscopy, often in combination with MD simulations, as a powerful tool to study protein folding.

Folding pathways and the energy landscape

Statistical ensemble views of protein folding predict that a rapid hydrophobic collapse into a subset of globular, unspecific structures, are the initial and necessary reduction in a polypeptide´s available conformational space. This collapse has been difficult to study from the bulk viewpoint due to the averaging of pathways and ensembles [97, 98]. AFM constant-force mode was used to study the early events of folding of individual ubiquitin molecules, induced by rapid force quenching [34]. Garcia-Maynes *et al.* were able to directly observe a collapse into an ensemble of lower energy structures for ubiquitin prior to folding into the native state. Polyubiquitin molecules were first unfolded followed by quenching to low force allowing the molecules to refold.

Subsequently, the molecules were unfolded again to monitor whether the proteins had folded into their native state. The natively folded ubiquitin molecules unfolded in discrete steps but structures apparently collapsed during refolding, unfolded in various size steps and had less mechanical stability. By varying the refolding time, they were able to extract the kinetics of the refolding process and showed that the collapsed states folded in a fast phase followed by a slower phase involving a barrier separated two-state mechanism.

The presence of multiple pathways is implied in the three-dimensional funnel-shaped energy landscape but experimental evidence for this phenomenon has remained limited [99, 100]. This is in part due to differential populations of pathways, where some pathways may be rarely visited by a protein. *He et al.* studied a slipknot protein by combining AFM and course-grained SMD simulations [101]. They found that the protein could untie the slipknot and unfold through either a two-state or a three-state mechanism, suggesting parallel pathways. SMD simulations revealed that the intermediate state in the three-state pathway was formed only if the unfolding initiated from the C-terminus and key structural elements were found that prevent the formation of a tightened knot structure in both pathways. The results support the idea of kinetic partitioning and thus add to a growing number of evidence suggesting it to be a general mechanism in protein folding.

Force acts on proteins locally in contrast to the global effect of temperature and chemicals. Furthermore, during mechanical manipulation the molecule is tethered, restraining its conformational flexibility. Certainly the unfolded state in these experiments is stretched as opposed to a random coil for freely diffusing molecules. A question therefore arises on whether the folding pathways are the same or not when determined with chemical/thermal denaturants or with force. Originally, both experimental and computational studies suggested that force-induced unfolding differs from that induced by chemical/thermal denaturants [102, 103]. Recently, however, evidence has emerged for similar pathways to exist. Simulations have suggested that at sufficiently low pulling speeds, where the molecule is able to sample a large part of its conformational space, mechanical unfolding pathways may start to resemble those observed in chemical denaturation studies [104]. A force-induced pathway-switch has indeed recently been observed experimentally for an SH3 domain in the low-force regime using optical tweezers [105]. Here, two different pulling axes were tested and one of them revealed two distinct and parallel unfolding pathways, giving evidence of a multi-dimensional energy landscape. Using both

constant-velocity experiments at different velocities and force-jump experiments, the authors determined the unfolding rates for both pulling axes. Although the absolute unfolding rates determined in bulk and using mechanical manipulation cannot generally be compared (due to contributions from the DNA handles and beads, however, a very recent study reports a data analysis method that accounts for these effects [106]), the effects of mutations on the rates could be compared in this study. The rates analysis indicated that the transition states in bulk and on the single-molecule level were populated to a similar extent, suggesting similar pathways.

Similar mechanical and chemical denaturation pathways have been suggested for acyl-CoA binding protein (ACBP) [53]. The structure and position along the reaction coordinate of the transition state of this small globular single-domain protein was investigated using a combination of optical tweezers and ratcheted MD simulations. Through both equilibrium and non-equilibrium optical tweezers experiments, Heidarsson *et al.* determined the "mechanical" Tanford β value ($m\beta_T$), which is analogous to a Tanford β-value in bulk studies [12, 107]. This value, which takes values between 0-1, can be regarded as a measure of the nativeness of the transition state and the determined $m\beta_T$ value for ACBP was very similar to the β-value in bulk [108]. An almost identical $m\beta_T$ value was also confirmed by ratcheted MD simulations, which were then used to estimate the structure of the transition state in atomic detail. The structure resembled the structure observed in bulk through phi-value analysis and NMR chemical shifts [109, 110]. It may therefore be protein-specific whether mechanical and chemical/thermal transition states and unfolding pathways are similar or not.

Related to transition states is the transition path, which is the seemingly instantaneous transition across the transition state barrier and includes all the mechanistic details of how a process happens. Transition path times, the actual time it takes to cross the barrier (not to be confused with transition rates which describe the frequency of barrier crossing), are impossible to measure by bulk methods as this is entirely a single molecule property. Yu *et al.* characterized the energy landscape of prion protein using optical tweezers [111]. From constant-force experiments they reconstructed a detailed free energy landscape and, using Kramers theory of barrier-limited diffusion, were able to directly determine the transition path times and folding rates. They found, in accordance with results from single-molecule fluorescence and MD results [91, 112], that transition path times are in the order of ~2 μs. Importantly, given the size of the prion protein this

study re-iterated that transition path times are relatively insensitive to protein size whereas folding rates can vary by many orders of magnitude.

Molecular response to force, secondary structure and topology

A remarkable feature of mechanical manipulation techniques is the ability to probe folding using various reaction coordinates. Almost any pulling geometry can be chosen by changing the position of the residues that define the load application. This offers the unique opportunity to directly explore the anisotropy of the energy landscape. Carrion-Vazquez *et al*. were the first to exploit this and using AFM they found that the direction of force application can dramatically change the mechanical stability of ubiquitin molecules [113]. Similar energy landscape anisotropy has been demonstrated with the green fluorescent protein (GFP) [45]. Dietz *et al*. performed AFM experiments on GFP using five different pulling geometries by varying the placement of cysteine residues which serve as attachment points in the polyprotein constructs *(Figure 1-8A)*. The mechanical stability of GFP displayed large variability depending on the pulling axis. From the resulting unfolding force distributions *(Figure 1-8B)* and using Monte Carlo simulations, the height of the transition state barriers and the potential energy well width of the native state could be determined. Remarkably, whereas the unfolding rates were not significantly affected by pulling direction, the width of the potential energy well showed significant variation, indicating either brittle or compliant behavior depending on the pulling geometry. Ultimately, the authors were able to describe the GFP structure in terms of different directional spring constants *(Figure 1-8C)*, highlighting the anisotropic nature of its energy landscape. These and similar studies demonstrate how directional control can allow exploration of regions of the energy landscape that are inaccessible to conventional methods.

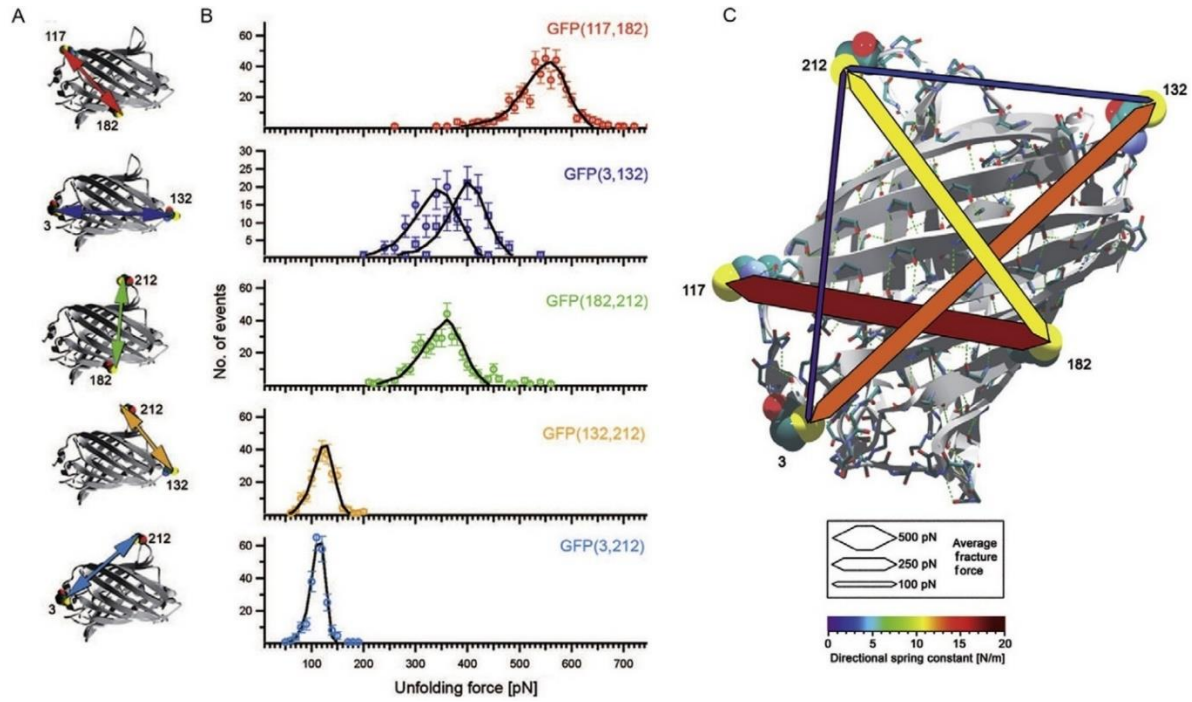


Figure 1-8 The anisotropy of GFP energy landscape probed with AFM. A) Five different pulling geometries of GFP were studied by generating polyproteins with various attachment points through cysteine mutations. The numbers correspond to the cysteine residues and the arrows show the pulling axis. B) Unfolding force distributions of the different variants. The solid lines represent results of a Monte Carlo simulation to reproduce the experimental data. C) Mechanical stability (indicated by the width of arrows) and directional spring constants of GFP (indicated by the color of arrows). Adapted from [45] with permission.

The determinants of anisotropic mechanical response are likely to be encoded in the structural architecture of proteins, involving features such as secondary structure, topology, and long-range contacts. It has been demonstrated that the mechanical stability of proteins is highly dependent on secondary structure. Proteins consisting mainly of β-sheets are better able to resist mechanical denaturation than α-helical proteins, and this has been attributed to the extensive network of hydrogen bonding in β-sheet proteins in contrast to mainly hydrophobic interactions between helices [114, 115]. Secondary structure also correlates with the distance from the native state to the transition state (x_u), where helical proteins have larger distances indicating softer, more compliant structures [116].

Topology, the arrangement of secondary elements along the sequence, has been shown to be important for cooperative folding during mechanical manipulation experiments. In T4 lysozyme (T4L), the A-helix is at the amino-terminal end but is part of the C-terminal structural domain. Shank *et al.* generated a circular permutant of T4L, effectively placing the A-helix at the C-terminal, and performed constant-velocity experiments using optical tweezers [61]. Even though the experiments provided only non-equilibrium information, equilibrium properties of the system could be revealed using CFT *(see section "Extracting thermodynamic parameters from non-equilibrium measurements").* In traditional ensemble studies, T4L unfolds in a single cooperative transition and this was also observed for the wild-type protein in constant-velocity experiments. However, the circular permutant displayed a three-state mechanism in which the N-domain now unfolded before the C-domain, indicating a decoupling of the two domains, which reduced cooperativity.

The mechanical properties of partially folded intermediate states are particularly interesting. Molten globule intermediate states have significant native-like secondary structure but lack the characteristic stable tertiary interactions of the native state. Elms *et al.* studied the molten globular state of apomyoglobin using optical tweezers [12]. They found that compared to native states, the molten globule is highly deformable (compliant), rendering its unfolding rate more sensitive to force, and speculated that this may be a general feature of molten globules. Surprisingly, the native state of ACBP has been shown to be even more deformable than the molten globule of apomyoglobin, a feature that may be important for its function as a lipid transporter [53]. The possibility that some intermediate states and even some proteins are more sensitive to force than others implies also that nature may have evolved this feature into proteins so that cells use less energy to unfold proteins that are targeted for degradation or translocation.

Intermediate states, misfolded states, and beyond

The role of intermediate states in protein folding has been debated and whether they are productive to the native state has remained an open question. *Cecconi et al.* used optical tweezers to characterize the refolding of ribonuclease H (RNase H) [9]. RNase H has been suggested to refold through a molten-globular intermediate state but it was unclear whether this state was on- or off-pathway [117]. By using constant-force experiments they observed that within a narrow range of forces the molecule "hopped" between the unfolded state and the intermediate state. In some cases, the hopping came to an abrupt stop followed by a further compaction into the native state. This

compaction occurred in the vast majority of cases directly from the intermediate state, providing direct evidence that it was on-pathway to the native state.

Predicted intermediate states have also been confirmed using force spectroscopy. Soluble N-ethylmaleimide–sensitive factor attachment protein receptor (SNARE) proteins mediate membrane fusion and are particularly important in vesicle fusion for neurotransmitter release [118]. Different SNARE proteins, attached to the vesicle membrane and the plasma membrane, are thought to assemble into a parallel four-helix bundle and it has been suggested that the bundle then zippers towards the membrane, producing sufficient force to enable fusion. The zippering action has, however, not been supported by direct evidence and the assembly intermediates have eluded detection. Using optical tweezers and some clever protein engineering, Gao *et al*. showed that a neuronal SNARE complex zippers in three distinct stages [119]. By applying forces in the same range as occurs during fusion, they were able to stabilize a half-zippered intermediate. The proposed zippering mechanism spawned from this study significantly impacts the study of neurotransmitter release.

Protein folding mechanisms have not evolved to perfection and can sometimes lead to incorrectly folded states. Misfolded proteins have attracted significant attention due to their well-established link to many severe diseases such as Alzheimer´s, Parkinson´s and Creutzfeldt-Jakobs' [120]. As the prime cause of the latter, the infectious form of the prion protein is one of the most widely studied misfolding systems. Yu *et al.* studied the prion protein using optical tweezers and found that folding occurred in a two-state mechanism, contrary to previous reports. They observed that besides the native state, three distinct short-lived misfolded states were also populated *(Figure 1- 9A)* [26] . These states were only accessible from the unfolded state, and could thus be classified as off-pathway, and remarkably, they were more frequently accessed than the native state under tension. A mutant that had higher aggregation propensity in bulk, populated two of the three misfolded states more frequently. The results of this study challenge the assumption that an on-pathway intermediate is responsible for aggregation and suggest that prion misfolding is mediated from the unfolded state, not the native state *(Figure 1-9B).*

Interestingly, proteins not directly linked to misfolding diseases have been shown to populate misfolded states. The ubiquitously expressed protein calmodulin (CaM) transduces changing levels in intracellular Ca^{2+} concentrations and is one of the most extensively studied calcium binding protein. Despite extensive information on CaM structural states and complex

formations from traditional methods, single-molecule methods have been able to impressively mine new features of its folding mechanism. Besides characterizing in detail the folding network of four distinct on-pathway states, Stigler *et al.* observed two additional misfolded states that were the result of incorrect pairing of EF-hands [64]. This misfolding slowed the overall folding kinetics of otherwise fast-folding individual domains.

Non-native interactions are not necessarily unwanted as elegantly demonstrated by [121]. The mechanosensory polycystin-1 PKD domain has been shown to be mechanically stronger than its native state predicts and MD simulations had suggested that this is due to a rearrangement leading to non-native hydrogen bonds that resist unfolding. Using genetic engineering, the authors produced mutant proteins that were designed to prevent formation of these non-native bonds. They subsequently showed through AFM experiments and MD simulations that despite having a negligible effect on the native state stability, the mutations caused dramatic mechanical destabilization. Thus, formation of non-native interactions during the mechanical unfolding of the PKD domain seems crucial for mechanical stability and, by inference, mechanosensory function.

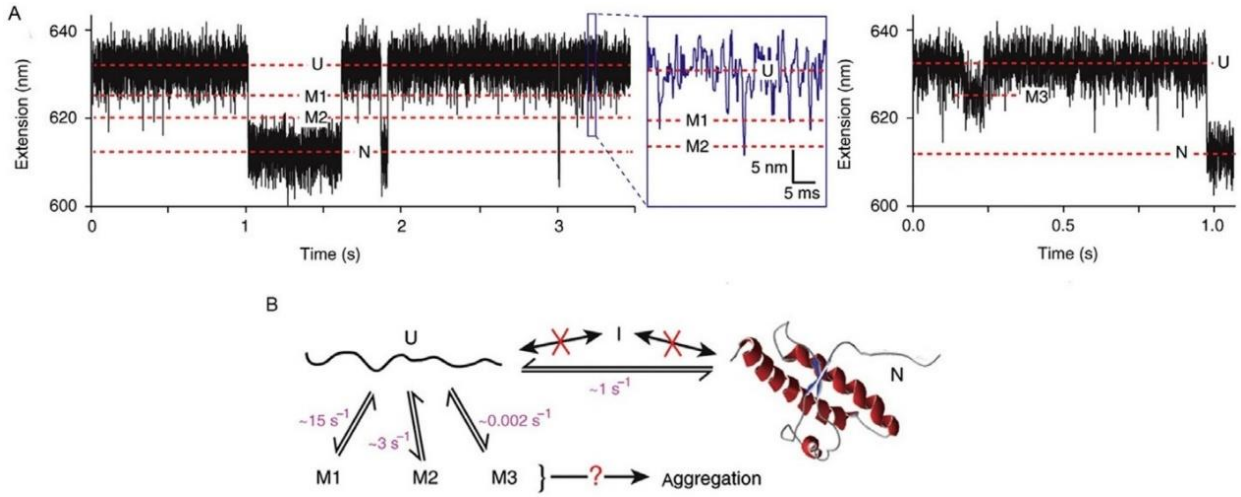


Figure 1-9 Misfolded states in single-molecule experiments. A) Extension vs. time traces of the prion protein. Left panel: The unfolded state frequently transitions into two distinct and short-lived misfolded states, M1 and M2. The inset shows a zoom into the region defined by the blue square. Right panel: Rarely, a third and longer-lived misfolded state was populated. B) An illustration of the folding/misfolding scheme of the prion protein along with the transition rates to the different misfolded states under tension. Adapted from [26] with permission.

Functional non-native interactions have also been suggested for coiled coils, molecules that have various roles in the cell often connected with sensing or generating force.

The coiled coils have characteristic *abcdefg* heptad repeats that engage in the dimer interface and native coiled coil dimer formation depends on the correct pairing of the *a* and *d* positions. Sometimes alternative pairings are formed when one helix is shifted, usually by one heptad repeat, and coiled coils have been suggested to undergo helix sliding to populate these so-called staggered helices. The non-native conformations may be important for biological function but have eluded detailed characterization due to their instability. Using high-resolution optical tweezers, *Xi et al.* studied two very stable coiled coils a variant of the GCN4 leucine zipper (pIL) and a heterodimeric coiled coil (pER)). The authors were able to directly observe partially folded states where the difference in the number of folded residues suggested these states to correspond to one helix shifted by one to three heptad repeats or in other words staggered helices. These states were populated less than 2% of the time with lifetimes decreasing inversely with the number of shifted heptad repeats. Usually, the partially folded states would only be accessed from the unfolded state (misfolding) but occasionally (<10%) the native state would transit directly into these states, which the authors interpreted as evidence for helix sliding.

Protein folding studies have greatly benefited from the advent of single-molecule manipulation studies from which new and exciting information continue to emerge. Kinetic partitioning seems to emerge as a general feature of protein folding kinetics and the frequent observation of misfolded states, even in small proteins thought to be efficient folders, challenge the idea of a highly evolved folding process. However, the traditional focus of protein folding studies needs to broaden from mostly single-domain proteins to more complex systems, as only a handful of larger proteins and double-domain systems has been explored on the single-molecule level. The current development of novel and hybrid instruments promises the ability to tackle ever more complex proteins and determine the details of multi-domain protein folding.

Protein-ligand and protein-protein interactions

Proteins interact with a multitude of binding partners *in vivo* and these interactions naturally affect the proteins structure and energetics. Molecular chaperones interact with proteins to improve the fidelity of their folding process, sometimes by rescuing misfolded states [122]. The binding of various ions can activate proteins by inducing conformational change, such as in the case of

calcium sensors, to initiate signaling pathways [123]. Conformational dynamics and flexibility are the hallmark of protein interactions [124] and single-molecule force spectroscopy is now being used to probe these events in significant detail.

Mechanical stability is largely governed by specific non-covalent interactions in the protein and ligand binding can induce conformational changes in the protein structure that lead to changes in mechanical stability. Alterations in mechanical stability upon ligand binding can therefore serve as an intrinsic reporter to identify the functional state of protein at the single-molecule level [125]. The effect of ligand binding on the mechanical stability of a protein is not trivial to predict. Using AFM, only moderate increases in mechanical stability were observed upon ligand binding to mouse dihydrofolate reductase (DHFR) [126] and to Im9 [127] while strong enhancement of mechanical stability occurs when protein G interacts with an Fc fragment from IgG. Stabilization of enzymatically inactive conformations by inhibitors might have significant importance in biomedical applications and drug design. For instance, the role of DHFR in cell proliferation and growth has rendered this enzyme a target for anticancer drug therapy [128]. Through AFM experiments Ainavarapu *et al*. showed that binding of the cancer chemotherapeutic agent methotrexate (MTX) to DHFR increases the mechanical stability of the enzyme [129]. The correlation between this result and the decrease in the degradation rate of DHFR inside the cell, supports the idea that force-induced denaturation is necessary for translocation and degradation of this protein. Also, through similar experiments, Junker *et al.*, have shown that simultaneous binding of MTX and the cofactor NADHP to DHFR increases the lifetime of one intermediate state of the protein, suggesting that MTX and NADHP could inhibit DHFR by trapping it in an enzymatically inactive intermediate structure [126].

Several proteins, like CaM, are activated by calcium ion binding, which allows them to interact with a vast array of binding partners. Using optical tweezers, Stigler and Rief studied CaM folding under conditions approaching physiological Ca^{2+} concentrations [24]. Variations in Ca^{2+} concentrations led to dramatic changes in the folding/unfolding kinetics of CaM. At high Ca^{2+} concentrations, CaM folding was characterized by a complex network of on- and off-pathway intermediate states *(Figure 1-10A)*. At low Ca^{2+} concentrations, the folding network remained the same but the mechanical stability of the protein drops and the unfolding and refolding rate constants changed significantly *(Figure 1-10B)*. In the absence of Ca^{2+}, the CaM folding network was essentially reduced to a two-state mechanism *(Figure 1-10C)*.

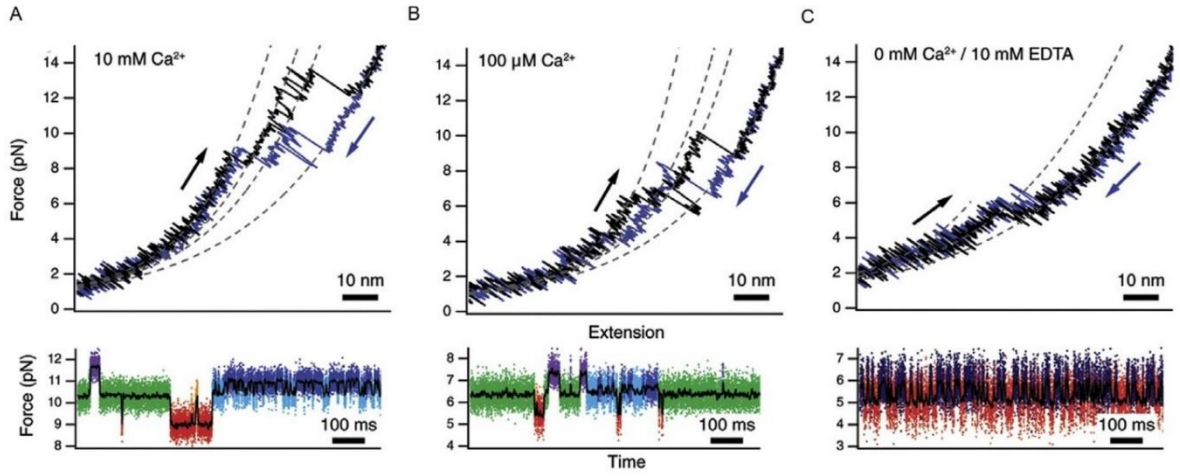


Figure 1-10 Calcium binding modulates the mechanical folding behavior of CaM. Force vs. extension cycles of CaM with varying Ca^{2+} concentration. A) 10 mM Ca^{2+}, B) 100 µm Ca^{2+}, and C) 0 mM Ca^{2+}. The lower panels show force vs. time traces, where the two beads tethering the molecule are kept at a constant distance. At high Ca^{2+} concentrations, CaM populates four intermediate structures between the unfolded and folded states, each color-coded. At lower Ca^{2+} concentrations, the same folding pattern is observed but at low forces. Under apo-conditions (0 mM Ca^{2+}) the kinetic pattern changes drastically and CaM fluctuates between only two states. Dotted lines in upper panels represent WLC fits to the data. Arrows indicate pulling direction. Adapted from [24] with permission.

Chaperones interact directly with polypeptide chains to increase the efficiency of folding. Some chaperones can prevent misfolding and aggregation, or facilitate protein translocation across membranes or for degradation [75]. However, the mechanism by which folding pathways are affected by chaperones is poorly understood. Bechtluft and co-workers used optical tweezers in combination with MD simulations to study the effect of the chaperone SecB on the folding pathway of maltose binding protein (MBP) [130]. They observed that, in the absence of SecB, MBP populated a molten-globule-like compacted state before folding into its native structure. Upon addition, SecB binds to the molten-globule state, stabilizing it and preventing the formation of stable native tertiary contacts of MBP. The effect of SecB on the structure of MBP might explain the mechanism by which this chaperone facilitates translocation of this protein through cellular membranes, as the absence of stable tertiary contacts likely facilitate the passage of the protein

through the translocation machinery. This work demonstrated how effective optical tweezers approaches can be to elucidate the effect of chaperones on protein folding landscapes.

Kim et al. developed an elegant optical tweezers method to perform repeated measurements of the binding/release kinetics between receptor and ligand [131]. Using this novel method, they were able to investigate the interaction between the A1 domain of von Willebrand factor (VWF) and the leucine-rich repeat (LRR) domain of glycoprotein Ib α subunit (GPIbα) by tethering the two molecules together via a flexible linker *(Figure 1-11A).*

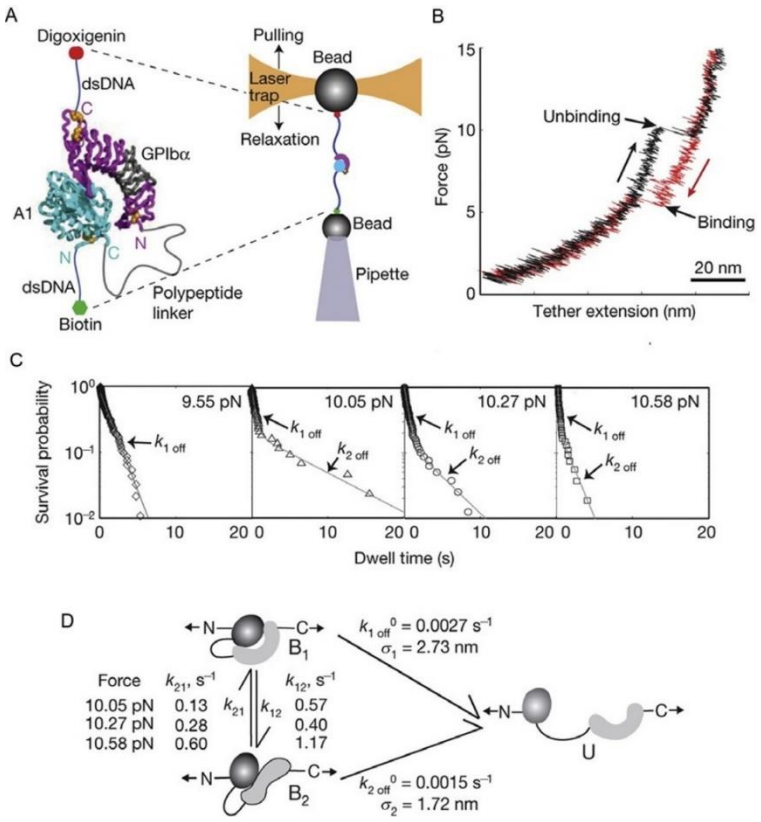


Figure 1-11 Novel approach to studying receptor-ligand binding/unbinding kinetics. A) The experimental setup. Left: The A1 domain-GPIbα LRR domain construct. Right: the optical tweezers setup. B) Force vs. extension trace for one cycle of stretching (black) and relaxing (red). C) Constant-force experiments were performed to determine the survival probability of the bound state as a function of time at different force values. D) A model of the A1- GPIbα flex-bond and the associated rate constants. Adapted from [131] with permission.

Force experiments at different pulling rates and constant-force showed that the A1- GPIbα interaction exists in two states *(Figure 1-11B and C)*, which the authors referred to as a flex-bond. One state was observed at low force whereas the other was mechanically more stable and had much longer lifetime. The kinetics of the bond formation *(Figure 1-11D)* help to explain how platelets bound to VWF are able to resist force to plug arterioles and how increased flow activates platelet plug formation.

Future perspectives

Single-molecule methods are making an impact in many aspects of cellular and molecular biology, and several new and exciting developments are on the horizon. Intrinsically disordered proteins (IDPs) are a recently recognized group of proteins where the central paradigm of a folded functional protein is challenged [132]. These proteins have seemingly no defined structure allowing significant functional plasticity but their properties are just beginning to be studied on the single-molecule level. So far, IDPs have exclusively been studied with single-molecule detection techniques but mechanical manipulation should offer an attractive approach to tackle IDPs as well. Optical tweezers have already successfully manipulated objects in living cells [133]. The dream is that near-future instrumental and methodological advances will allow us to directly observe life unfold by watching individual molecules *in vivo*, in real time.

References

[1] N.J. Greenfield, Using circular dichroism spectra to estimate protein secondary structure, Nat. Protocols, 1 (2007) 2876-2890.

[2] D. Kern, E.Z. Eisenmesser, M. Wolf-Watz, Enzyme Dynamics During Catalysis Measured by NMR Spectroscopy, in: L.J. Thomas (Ed.) Methods in Enzymology, Academic Press2005, pp. 507-524.

[3] H. Maity, M. Maity, M.M.G. Krishna, L. Mayne, S.W. Englander, Protein folding: The stepwise assembly of foldon units, Proceedings of the National Academy of Sciences of the United States of America, 102 (2005) 4741-4746.

[4] P.O. Heidarsson, M.M. Naqvi, P. Sonar, I. Valpapuram, C. Cecconi, Chapter Three - Conformational Dynamics of Single Protein Molecules Studied by Direct Mechanical Manipulation, in: K.-C. Tatyana (Ed.) Advances in Protein Chemistry and Structural Biology, Academic Press2013, pp. 93-133.

[5] A. Ashkin, J.M. Dziedzic, J.E. Bjorkholm, S. Chu, Observation of a single-beam gradient force optical trap for dielectric particles, Optics Letters, 11 (1986) 288-290.

[6] J.R. Moffitt, Y.R. Chemla, S.B. Smith, C. Bustamante, Recent Advances in Optical Tweezers, Annual Review of Biochemistry, 77 (2008) 205-228.

[7] S.B. Smith, Y. Cui, C. Bustamante, [7] Optical-trap force transducer that operates by direct measurement of light momentum, in: M. Gerard, P. Ian (Eds.) Methods in Enzymology, Academic Press2003, pp. 134-162.

[8] C. Bustamante, J. Marko, E. Siggia, S. Smith, Entropic elasticity of lambda-phage DNA, Science, 265 (1994) 1599-1600.

[9] C. Cecconi, E.A. Shank, C. Bustamante, S. Marqusee, Direct Observation of the Three-State Folding of a Single Protein Molecule, Science, 309 (2005) 2057-2060.

[10] J. Liphardt, B. Onoa, S.B. Smith, I. Tinoco, C. Bustamante, Reversible Unfolding of Single RNA Molecules by Mechanical Force, Science, 292 (2001) 733-737.

[11] P.O. Heidarsson, I. Valpapuram, C. Camilloni, A. Imparato, G. Tiana, F.M. Poulsen, B.B. Kragelund, C. Cecconi, A highly compliant protein native state with a spontaneous-like mechanical unfolding pathway, Journal of the American Chemical Society, 134 (2012) 17068-17075.

[12] P.J. Elms, J.D. Chodera, C. Bustamante, S. Marqusee, The molten globule state is unusually deformable under mechanical force, Proceedings of the National Academy of Sciences, 109 (2012a) 3796-3801.
[13] P.T.X. Li, D. Collin, S.B. Smith, C. Bustamante, I. Tinoco, Jr., Probing the Mechanical Folding Kinetics of TAR RNA by Hopping, Force-Jump, and Force-Ramp Methods, Biophysical journal, 90 (2006) 250-260.
[14] K.C. Neuman, A. Nagy, Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy, Nat Meth, 5 (2008) 491-505.
[15] I. De Vlaminck, I. Vidic, M.T.J. van Loenhout, R. Kanaar, J.H.G. Lebbink, C. Dekker, Torsional regulation of hRPA-induced unwinding of double-stranded DNA, Nucleic Acids Research, 38 (2010) 4133-4142.
[16] I. De Vlaminck, C. Dekker, Recent Advances in Magnetic Tweezers, Annual Review of Biophysics, 41 (2012) 453-472.
[17] M.C. Noom, B. van den Broek, J. van Mameren, G.J.L. Wuite, Visualizing single DNA-bound proteins using DNA as a scanning probe, Nat Meth, 4 (2007) 1031-1036.
[18] I. De Vlaminck, T. Henighan, M.T.J. van Loenhout, D.R. Burnham, C. Dekker, Magnetic Forces and DNA Mechanics in Multiplexed Magnetic Tweezers, PLoS ONE, 7 (2012) e41432.
[19] G. Sirinakis, Y. Ren, Y. Gao, Z. Xi, Y. Zhang, Combined versatile high-resolution optical tweezers and single-molecule fluorescence microscopy, Review of Scientific Instruments, 83 (2012) -.
[20] R.T. Dame, M.C. Noom, G.J.L. Wuite, Bacterial chromatin organization by H-NS protein unravelled using dual DNA manipulation, Nature, 444 (2006) 387-390.
[21] C. Cecconi, E. Shank, F. Dahlquist, S. Marqusee, C. Bustamante, Protein-DNA chimeras for single molecule mechanical folding studies with the optical tweezers, Eur Biophys J, 37 (2008) 729-738.
[22] Y. Gao, G. Sirinakis, Y. Zhang, Highly Anisotropic Stability and Folding Kinetics of a Single Coiled Coil Protein under Mechanical Tension, Journal of the American Chemical Society, 133 (2011) 12749-12757.
[23] J.C.M. Gebhardt, T. Bornschlögl, M. Rief, Full distance-resolved folding energy landscape of one single protein molecule, Proceedings of the National Academy of Sciences, 107 (2010) 2013-2018.

[24] J. Stigler, M. Rief, Calcium-dependent folding of single calmodulin molecules, Proceedings of the National Academy of Sciences, 109 (2012) 17814-17819.
[25] Z. Xi, Y. Gao, G. Sirinakis, H. Guo, Y. Zhang, Single-molecule observation of helix staggering, sliding, and coiled coil misfolding, Proceedings of the National Academy of Sciences, (2012).
[26] H. Yu, X. Liu, K. Neupane, A.N. Gupta, A.M. Brigley, A. Solanki, I. Sosova, M.T. Woodside, Direct observation of multiple misfolding pathways in a single prion protein molecule, Proceedings of the National Academy of Sciences, 109 (2012) 5283-5288.
[27] C. Cecconi, E.A. Shank, S. Marqusee, C. Bustamante, DNA Molecular Handles for Single-Molecule Protein-Folding Studies by Optical Tweezers, Methods Mol Biol, The Humana Press Inc (New York, NY, U.S.A.)2011, pp. 255-271.
[28] G. Binnig, C.F. Quate, C. Gerber, Atomic Force Microscope, Physical Review Letters, 56 (1986) 930-933.
[29] A. Raab, W. Han, D. Badt, S.J. Smith-Gill, S.M. Lindsay, H. Schindler, P. Hinterdorfer, Antibody recognition imaging by force microscopy, Nat Biotech, 17 (1999) 901-905.
[30] E. Florin, V. Moy, H. Gaub, Adhesion forces between individual ligand-receptor pairs, Science, 264 (1994) 415-417.
[31] G. Lee, L. Chrisey, R. Colton, Direct measurement of the forces between complementary strands of DNA, Science, 266 (1994) 771-773.
[32] M. Rief, F. Oesterhelt, B. Heymann, H.E. Gaub, Single Molecule Force Spectroscopy on Polysaccharides by Atomic Force Microscopy, Science, 275 (1997) 1295-1297.
[33] T. Bornschlögl, M. Rief, Single-Molecule Protein Unfolding and Refolding Using Atomic Force Microscopy, in: E.J.G. Peterman, G.J.L. Wuite (Eds.) Single Molecule Analysis, Humana Press2011, pp. 233-250.
[34] S. Garcia-Manyes, L. Dougan, C.L. Badilla, J. Brujić, J.M. Fernández, Direct observation of an ensemble of stable collapsed states in the mechanical folding of ubiquitin, Proceedings of the National Academy of Sciences, 106 (2009) 10534-10539.
[35] S. Ng, L. Randles, J. Clarke, Single Molecule Studies of Protein Folding Using Atomic Force Microscopy, in: Y. Bai, R. Nussinov (Eds.) Protein Folding Protocols, Humana Press2006, pp. 139-167.

[36] R. Rounsevell, J.R. Forman, J. Clarke, Atomic force microscopy: mechanical unfolding of proteins, Methods, 34 (2004) 100-111.
[37] M. Carrion-Vazquez, A.F. Oberhauser, S.B. Fowler, P.E. Marszalek, S.E. Broedel, J. Clarke, J.M. Fernandez, Mechanical and chemical unfolding of a single protein: A comparison, Proceedings of the National Academy of Sciences, 96 (1999) 3694-3699.
[38] H. Dietz, M. Bertz, M. Schlierf, F. Berkemeier, T. Bornschlogl, J.P. Junker, M. Rief, Cysteine engineering of polyproteins for single-molecule force spectroscopy, Nat. Protocols, 1 (2006) 80-84.
[39] A. Steward, J.L. Toca-Herrera, J. Clarke, Versatile cloning system for construction of multimeric proteins for use in atomic force microscopy, Protein Science, 11 (2002) 2179-2183.
[40] G. Yang, C. Cecconi, W.A. Baase, I.R. Vetter, W.A. Breyer, J.A. Haack, B.W. Matthews, F.W. Dahlquist, C. Bustamante, Solid-state synthesis and mechanical unfolding of polymers of T4 lysozyme, Proceedings of the National Academy of Sciences, 97 (2000) 139-144.
[41] J.P. Junker, M. Rief, Single-molecule force spectroscopy distinguishes target binding modes of calmodulin, Proceedings of the National Academy of Sciences, 106 (2009) 14361-14366.
[42] J.P. Junker, F. Ziegler, M. Rief, Ligand-Dependent Equilibrium Fluctuations of Single Calmodulin Molecules, Science, 323 (2009) 633-637.
[43] M. Grandbois, M. Beyer, M. Rief, H. Clausen-Schaumann, H.E. Gaub, How Strong Is a Covalent Bond?, Science, 283 (1999) 1727-1730.
[44] P.E. Marszalek, A.F. Oberhauser, Y.-P. Pang, J.M. Fernandez, Polysaccharide elasticity governed by chair-boat transitions of the glucopyranose ring, Nature, 396 (1998) 661-664.
[45] H. Dietz, F. Berkemeier, M. Bertz, M. Rief, Anisotropic deformation response of single protein molecules, Proceedings of the National Academy of Sciences, 103 (2006) 12724-12728.
[46] C. Bustamante, Y.R. Chemla, N.R. Forde, D. Izhaky, MECHANICAL PROCESSES IN BIOCHEMISTRY, Annual Review of Biochemistry, 73 (2004) 705-748.
[47] G. Bell, Models for the specific adhesion of cells to cells, Science, 200 (1978) 618-627.
[48] O.K. Dudko, T.G.W. Graham, R.B. Best, Locating the Barrier for Folding of Single Molecules under an External Force, Physical Review Letters, 107 (2011) 208301.
[49] M. Manosas, D. Collin, F. Ritort, Force-Dependent Fragility in RNA Hairpins, Physical Review Letters, 96 (2006) 218301.

[50] M. Schlierf, F. Berkemeier, M. Rief, Direct Observation of Active Protein Folding Using Lock-in Force Spectroscopy, Biophysical journal, 93 (2007) 3989-3998.
[51] I. Tinoco Jr, C. Bustamante, The effect of force on thermodynamics and kinetics of single molecule reactions, Biophysical chemistry, 101–102 (2002) 513-533.
[52] M. Schlierf, H. Li, J.M. Fernandez, The unfolding kinetics of ubiquitin captured with single-molecule force-clamp techniques, Proceedings of the National Academy of Sciences of the United States of America, 101 (2004) 7299-7304.
[53] P.O. Heidarsson, I. Valpapuram, C. Camilloni, A. Imparato, G. Tiana, F.M. Poulsen, B.B. Kragelund, C. Cecconi, A Highly Compliant Protein Native State with a Spontaneous-like Mechanical Unfolding Pathway, Journal of the American Chemical Society, 134 (2012) 17068-17075.
[54] C. Jarzynski, Equalities and Inequalities: Irreversibility and the Second Law of Thermodynamics at the Nanoscale, Annual Review of Condensed Matter Physics, 2 (2011) 329-351.
[55] C. Jarzynski, Nonequilibrium Equality for Free Energy Differences, Physical Review Letters, 78 (1997) 2690-2693.
[56] J. Liphardt, S. Dumont, S.B. Smith, I. Tinoco, C. Bustamante, Equilibrium Information from Nonequilibrium Measurements in an Experimental Test of Jarzynski's Equality, Science, 296 (2002) 1832-1835.
[57] J. Gore, F. Ritort, C. Bustamante, Bias and error in estimates of equilibrium free-energy differences from nonequilibrium measurements, Proceedings of the National Academy of Sciences, 100 (2003) 12564-12569.
[58] G.E. Crooks, Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences, Physical Review E, 60 (1999) 2721-2726.
[59] C.H. Bennett, Efficient estimation of free energy differences from Monte Carlo data, Journal of Computational Physics, 22 (1976) 245-268.
[60] D. Collin, F. Ritort, C. Jarzynski, S.B. Smith, I. Tinoco, C. Bustamante, Verification of the Crooks fluctuation theorem and recovery of RNA folding free energies, Nature, 437 (2005) 231-234.
[61] A.S. Elizabeth, C. Ciro, W.D. Jesse, M. Susan, B. Carlos, The folding cooperativity of a protein is controlled by its chain topology, Nature, 465 (2010) 637-640.

[62] A. Alemany, A. Mossa, I. Junier, F. Ritort, Experimental free-energy measurements of kinetic molecular states using fluctuation theorems, Nat Phys, 8 (2012) 688-694.
[63] C.M. Kaiser, D.H. Goldman, J.D. Chodera, I. Tinoco, C. Bustamante, The Ribosome Modulates Nascent Protein Folding, Science, 334 (2011) 1723-1727.
[64] J. Stigler, F. Ziegler, A. Gieseke, J.C.M. Gebhardt, M. Rief, The Complex Folding Network of Single Calmodulin Molecules, Science, 334 (2011) 512-516.
[65] L. Baum, An inequality and associated maximization technique in statistical estimation for probabilistic functions of Markov processes, Inequalities, 3 (1972) 1-8.
[66] L.E. Baum, T. Petrie, Statistical Inference for Probabilistic Functions of Finite State Markov Chains, (1966) 1554-1563.
[67] J. Baker, The DRAGON system--An overview, Acoustics, Speech and Signal Processing, IEEE Transactions on, 23 (1975) 24-29.
[68] M. Stamp, E. Le, Hamptonese and Hidden Markov Models, in: W. P. Dayawansa, A. Lindquist, Y. Zhou (Eds.) New Directions and Applications in Control Theory, Springer Berlin Heidelberg2005, pp. 367-378.
[69] J.D. Chodera, F. Noé, Probability distributions of molecular observables computed from Markov models. II. Uncertainties in observables and their time-evolution, The Journal of Chemical Physics, 133 (2010) -.
[70] O.K. Dudko, G. Hummer, A. Szabo, Theory, analysis, and interpretation of single-molecule force spectroscopy experiments, Proceedings of the National Academy of Sciences, 105 (2008) 15755-15760.
[71] G. Žoldák, M. Rief, Force as a single molecule probe of multidimensional protein energy landscapes, Current Opinion in Structural Biology, 23 (2013) 48-57.
[72] A.B. Kolomeisky, M.E. Fisher, Molecular Motors: A Theorist's Perspective, Annual Review of Physical Chemistry, 58 (2007) 675-695.
[73] D.E. Jaalouk, J. Lammerding, Mechanotransduction gone awry, Nat Rev Mol Cell Biol, 10 (2009) 63-73.
[74] R.W. King, R.J. Deshaies, J.-M. Peters, M.W. Kirschner, How Proteolysis Drives the Cell Cycle, Science, 274 (1996) 1652-1659.

[75] Rodrigo A. Maillard, G. Chistol, M. Sen, M. Righini, J. Tan, C.M. Kaiser, C. Hodges, A. Martin, C. Bustamante, ClpX(P) Generates Mechanical Force to Unfold and Translocate Its Protein Substrates, Cell, 145 (2011) 459-469.
[76] C.B. Anfinsen, E. Haber, M. Sela, F.H. White, THE KINETICS OF FORMATION OF NATIVE RIBONUCLEASE DURING OXIDATION OF THE REDUCED POLYPEPTIDE CHAIN, Proceedings of the National Academy of Sciences, 47 (1961) 1309-1314.
[77] C. Levinthal, Are there pathways for protein folding?, J. Chim. Phys., 65 (1968) 44-45.
[78] R. Zwanzig, A. Szabo, B. Bagchi, Levinthal's paradox, Proceedings of the National Academy of Sciences, 89 (1992) 20-22.
[79] J.D. Bryngelson, J.N. Onuchic, N.D. Socci, P.G. Wolynes, Funnels, pathways, and the energy landscape of protein folding: A synthesis, Proteins: Structure, Function, and Bioinformatics, 21 (1995) 167-195.
[80] K.A. Dill, H.S. Chan, From Levinthal to pathways to funnels, Nat Struct Biol, 4 (1997) 10-19.
[81] K.A. Dill, J.L. MacCallum, The Protein-Folding Problem, 50 Years On, Science, 338 (2012) 1042-1046.
[82] P.E. Leopold, M. Montal, J.N. Onuchic, Protein folding funnels: a kinetic approach to the sequence-structure relationship, Proceedings of the National Academy of Sciences, 89 (1992) 8721-8725.
[83] M. Oliveberg, P.G. Wolynes, The experimental survey of protein-folding energy landscapes, Quarterly Reviews of Biophysics, 38 (2005) 245-288.
[84] J.N. Onuchic, Z. Luthey-Schulten, P.G. Wolynes, THEORY OF PROTEIN FOLDING: The Energy Landscape Perspective, Annual Review of Physical Chemistry, 48 (1997) 545-600.
[85] L. Milanesi, J.P. Waltho, C.A. Hunter, D.J. Shaw, G.S. Beddard, G.D. Reid, S. Dev, M. Volk, Measurement of energy landscape roughness of folded and unfolded proteins, Proceedings of the National Academy of Sciences, 109 (2012) 19563-19568.
[86] M. Karplus, Behind the folding funnel diagram, Nat Chem Biol, 7 (2011) 401-404.
[87] J.D. Bryngelson, P.G. Wolynes, Intermediates and barrier crossing in a random energy model (with applications to protein folding), The Journal of Physical Chemistry, 93 (1989) 6902-6915.

[88] B.G. Wensley, S. Batey, F.A. Bone, Z.M. Chan, N.R. Tumelty, A. Steward, L.G. Kwa, A. Borgia, J. Clarke, Experimental evidence for a frustrated energy landscape in a three-helix-bundle protein family, Nature, 463 (2010) 685-688.
[89] D. Thirumalai, D.K. Klimov, S.A. Woodson, Kinetic partitioning mechanism as a unifying theme in the folding of biomolecules, Theor Chem Acc, 96 (1997) 14-22.
[90] A. Matouschek, J.T. Kellis, L. Serrano, A.R. Fersht, Mapping the transition state and pathway of protein folding by protein engineering, Nature, 340 (1989) 122-126.
[91] D.E. Shaw, P. Maragakis, K. Lindorff-Larsen, S. Piana, R.O. Dror, M.P. Eastwood, J.A. Bank, J.M. Jumper, J.K. Salmon, Y. Shan, W. Wriggers, Atomic-Level Characterization of the Structural Dynamics of Proteins, Science, 330 (2010) 341-346.
[92] H.A. Scheraga, M. Khalili, A. Liwo, Protein-Folding Dynamics: Overview of Molecular Simulation Techniques, Annual Review of Physical Chemistry, 58 (2007) 57-83.
[93] M.S.Z. Kellermayer, S.B. Smith, H.L. Granzier, C. Bustamante, Folding-Unfolding Transitions in Single Titin Molecules Characterized with Laser Tweezers, Science, 276 (1997) 1112-1116.
[94] M. Rief, M. Gautel, F. Oesterhelt, J.M. Fernandez, H.E. Gaub, Reversible Unfolding of Individual Titin Immunoglobulin Domains by AFM, Science, 276 (1997) 1109-1112.
[95] L. Tskhovrebova, J. Trinick, J.A. Sleep, R.M. Simmons, Elasticity and unfolding of single molecules of the giant muscle protein titin, Nature, 387 (1997) 308-312.
[96] J.M. Fernandez, H. Li, Force-Clamp Spectroscopy Monitors the Folding Trajectory of a Single Protein, Science, 303 (2004) 1674-1678.
[97] V.R. Agashe, M.C.R. Shastry, J.B. Udgaonkar, Initial hydrophobic collapse in the folding of barstar, Nature, 377 (1995) 754-757.
[98] A. Dasgupta, J.B. Udgaonkar, Evidence for Initial Non-specific Polypeptide Chain Collapse During the Refolding of the SH3 Domain of PI3 Kinase, Journal of Molecular Biology, 403 (2010) 430-445.
[99] S.E. Radford, C.M. Dobson, P.A. Evans, The folding of hen lysozyme involves partially structured intermediates and multiple pathways, Nature, 358 (1992) 302-307.
[100] C.F. Wright, K. Lindorff-Larsen, L.G. Randles, J. Clarke, Parallel protein-unfolding pathways revealed and mapped, Nat Struct Mol Biol, 10 (2003) 658-662.

[101] C. He, G.Z. Genchev, H. Lu, H. Li, Mechanically Untying a Protein Slipknot: Multiple Pathways Revealed by Force Spectroscopy and Steered Molecular Dynamics Simulations, Journal of the American Chemical Society, 134 (2012) 10428-10435.
[102] R.B. Best, S.B. Fowler, J.L. Toca Herrera, A. Steward, E. Paci, J. Clarke, Mechanical Unfolding of a Titin Ig Domain: Structure of Transition State Revealed by Combining Atomic Force Microscopy, Protein Engineering and Molecular Dynamics Simulations, Journal of Molecular Biology, 330 (2003) 867-877.
[103] S.P. Ng, R.W.S. Rounsevell, A. Steward, C.D. Geierhaas, P.M. Williams, E. Paci, J. Clarke, Mechanical Unfolding of TNfn3: The Unfolding Pathway of a fnIII Domain Probed by Protein Engineering, AFM and MD Simulation, Journal of Molecular Biology, 350 (2005) 776-789.
[104] D.K. West, P.D. Olmsted, E. Paci, Mechanical unfolding revisited through a simple but realistic model, The Journal of Chemical Physics, 124 (2006) -.
[105] B. Jagannathan, P.J. Elms, C. Bustamante, S. Marqusee, Direct observation of a force-induced switch in the anisotropic mechanical unfolding pathway of a protein, Proceedings of the National Academy of Sciences, 109 (2012) 17820-17825.
[106] M. Hinczewski, J.C.M. Gebhardt, M. Rief, D. Thirumalai, From mechanical folding trajectories to intrinsic energy landscapes of biopolymers, Proceedings of the National Academy of Sciences, 110 (2013) 4500-4505.
[107] C. Tanford, Protein Denaturation: Part C. Theoretical Models for The Mechanism of Denaturation, in: J.T.E. C.B. Anfinsen, M.R. Frederic (Eds.) Advances in Protein Chemistry, Academic Press1970, pp. 1-95.
[108] B.B. Kragelund, P. Højrup, M.S. Jensen, C.K. Schjerling, E. Juul, J. Knudsen, M. Poulsen, Fast and One-step Folding of Closely and Distantly Related Homologous Proteins of a Four-helix Bundle Family, Journal of Molecular Biology, 256 (1996) 187-200.
[109] S.W. Bruun, V. Iešmantavičius, J. Danielsson, F.M. Poulsen, Cooperative formation of native-like tertiary contacts in the ensemble of unfolded states of a four-helix protein, Proceedings of the National Academy of Sciences, 107 (2010) 13306-13311.
[110] B.B. Kragelund, P. Osmark, T.B. Neergaard, J. Schiodt, K. Kristiansen, J. Knudsen, F.M. Poulsen, The formation of a native-like structure containing eight conserved hydrophobic residues is rate limiting in two-state protein folding of ACBP, Nat Struct Mol Biol, 6 (1999) 594-601.

[111] H. Yu, A.N. Gupta, X. Liu, K. Neupane, A.M. Brigley, I. Sosova, M.T. Woodside, Energy landscape analysis of native folding of the prion protein yields the diffusion constant, transition path time, and rates, Proceedings of the National Academy of Sciences, 109 (2012) 14452-14457.
[112] H.S. Chung, K. McHale, J.M. Louis, W.A. Eaton, Single-Molecule Fluorescence Experiments Determine Protein Folding Transition Path Times, Science, 335 (2012) 981-984.
[113] M. Carrion-Vazquez, H. Li, H. Lu, P.E. Marszalek, A.F. Oberhauser, J.M. Fernandez, The mechanical stability of ubiquitin is linkage dependent, Nat Struct Mol Biol, 10 (2003) 738-743.
[114] D.J. Brockwell, G.S. Beddard, E. Paci, D.K. West, P.D. Olmsted, D.A. Smith, S.E. Radford, Mechanically Unfolding the Small, Topologically Simple Protein L, Biophysical journal, 89 (2005) 506-519.
[115] M. Schlierf, M. Rief, Temperature Softening of a Protein in Single-molecule Experiments, Journal of Molecular Biology, 354 (2005) 497-503.
[116] M.S. Li, Secondary Structure, Mechanical Stability, and Location of Transition State of Proteins, Biophysical journal, 93 (2007) 2644-2654.
[117] T.M. Raschke, J. Kho, S. Marqusee, Confirmation of the hierarchical folding of RNase H: a protein engineering study, Nat Struct Mol Biol, 6 (1999) 825-831.
[118] N.A. Ramakrishnan, M.J. Drescher, D.G. Drescher, The SNARE complex in neuronal and sensory cells, Molecular and cellular neurosciences, 50 (2012) 58-69.
[119] Y. Gao, S. Zorman, G. Gundersen, Z. Xi, L. Ma, G. Sirinakis, J.E. Rothman, Y. Zhang, Single Reconstituted Neuronal SNARE Complexes Zipper in Three Distinct Stages, Science, 337 (2012) 1340-1343.
[120] C.M. Dobson, Protein folding and misfolding, Nature, 426 (2003) 884-890.
[121] J.R. Forman, Z.T. Yew, S. Qamar, R.N. Sandford, E. Paci, J. Clarke, Non-Native Interactions Are Critical for Mechanical Strength in PKD Domains, Structure, 17 (2009) 1582-1590.
[122] F.U. Hartl, A. Bracher, M. Hayer-Hartl, Molecular chaperones in protein folding and proteostasis, Nature, 475 (2011) 324-332.
[123] W.J. Chazin, Relating Form and Function of EF-Hand Calcium Binding Proteins, Accounts of Chemical Research, 44 (2011) 171-179.
[124] K. Teilum, J.G. Olsen, B.B. Kragelund, Protein stability, flexibility and function, Biochimica et biophysica acta, 1814 (2011) 969-976.

[125] Y. Cao, M.M. Balamurali, D. Sharma, H. Li, A functional single-molecule binding assay via force spectroscopy, Proceedings of the National Academy of Sciences, 104 (2007) 15677-15681.
[126] J.P. Junker, K. Hell, M. Schlierf, W. Neupert, M. Rief, Influence of Substrate Binding on the Mechanical Stability of Mouse Dihydrofolate Reductase, Biophysical journal, 89 (2005) L46-L48.
[127] E. Hann, N. Kirkpatrick, C. Kleanthous, D.A. Smith, S.E. Radford, D.J. Brockwell, The Effect of Protein Complexation on the Mechanical Stability of Im9, Biophysical journal, 92 (2007) L79-L81.
[128] R. Li, R. Sirawaraporn, P. Chitnumsub, W. Sirawaraporn, J. Wooden, F. Athappilly, S. Turley, W.G.J. Hol, Three-dimensional structure of M. tuberculosis dihydrofolate reductase reveals opportunities for the design of novel tuberculosis drugs, Journal of Molecular Biology, 295 (2000) 307-323.
[129] S.R.K. Ainavarapu, L. Li, C.L. Badilla, J.M. Fernandez, Ligand Binding Modulates the Mechanical Stability of Dihydrofolate Reductase, Biophysical journal, 89 (2005) 3337-3344.
[130] P. Bechtluft, R.G.H. van Leeuwen, M. Tyreman, D. Tomkiewicz, N. Nouwen, H.L. Tepper, A.J.M. Driessen, S.J. Tans, Direct Observation of Chaperone-Induced Changes in a Protein Folding Pathway, Science, 318 (2007) 1458-1461.
[131] J. Kim, C.-Z. Zhang, X. Zhang, T.A. Springer, A mechanically stabilized receptor-ligand flex-bond important in the vasculature, Nature, 466 (2010) 992-995.
[132] P. Tompa, Unstructural biology coming of age, Current Opinion in Structural Biology, 21 (2011) 419-425.
[133] L.B. Oddershede, Force probing of individual molecules inside the living cell is now a reality, Nat Chem Biol, 8 (2012) 879-886.

Chapter: 2

Optical tweezers and Model systems

Chapter 2: Optical tweezers and Model systems

Abstract

Force spectroscopy has been emerging as very promising approach to study biological processes at single-molecule level. This chapter provides a brief overview on optical tweezers (OT), the physics behind optical trapping, the instrumental details of the OT setup at the University of Modena and Reggio Emilia and briefly describes the biological models used for these studies.

Part – A: Principles of Optical trapping and optical tweezers

Introduction

Optical tweezers (OT) invented by Arthur Ashkin and colleagues in 1986 [1], uses radiation pressure to accelerate, decelerate, deflect and even stably trap tiny objects. Ashkin and co-workers found that, by focusing a laser light beam through a microscopic objective, particles with high refractive index, such as glass, plastic, were attracted to intense regions in the beam and could be held permanently at a focal point. Later in 1987, Ashkin and co-workers made the discovery of damage free optical trapping and manipulation of biological samples using infrared lasers [2]. This discovery opened new doors for a revolutionary use of the OT in biology. Since then OT has developed very rapidly. OT has been modified several times by many groups to study different aspects and applications of biological objects [3, 4]. Coupling OT with other techniques such as, fluorescence spectroscopy, Raman spectroscopy significantly expands the capabilities of the traditional OT, allowing the manipulation of more complicated biological systems, providing additional variables such as torque or the simultaneous use of additional forms of single-molecule manipulation and detection [5]. After gaining tremendous success over last four decades, nowadays OT became an important tool to study different biological systems like nucleic acids, RNA, molecular motors, proteins.

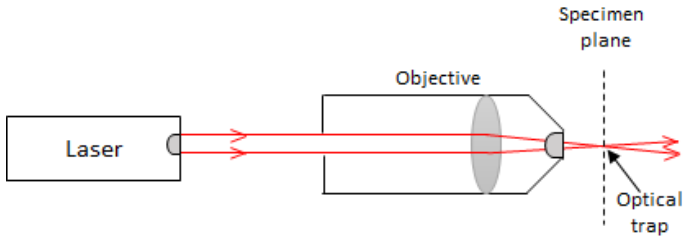
Physics behind OT

The basic principle behind OT is the momentum transfer associated with bending light. Light carries momentum that is proportional to its energy and in the direction of propagation. Light consists of photons, a photon of wavelength λ has a momentum $P = h/k$ (or $\hbar k$) where h is Plank's constant, $\hbar$ is $h/2\pi$, and k is a wave number. When light hits the dielectric object, it changes its direction due to reflection or refraction, which result in a change in the momentum of light. For the conservation of momentum, the affected object must undergoes an equal and opposite momentum change. This gives rise to a force acting on the object. The forces involved in the transfer of momentum from focused laser light to microscopic objects are in order of piconewton, but is sufficient to move micron size objects within a specific region.

In instances where the diameter of trapped particle is larger than the wavelength of light, the trapping phenomenon is explained by a simple ray optics (Mie regime). Consider a focused

beam of light with a Gaussian intensity profile hits the dielectric object with high refractive index than surrounding medium, situated off axis in the beam *(Figure 2-1A)*.

A.



B.

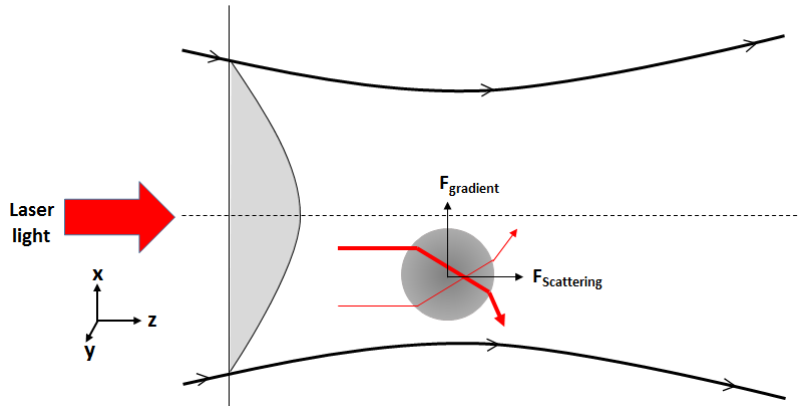


Figure 2-1 Schematic showing working principle of OT with the help of ray optics diagram. A) A laser beam is focused by a high quality microscope objective to spot in the specimen plane. This spot create an "optical trap" which is able to hold a small particle at its center. B) A focused beam of light with Gaussian intensity profile hits a high refractive index dielectric object. Due to this light-bead interactions, object feels scattering and gradient forces. The refraction of rays of different intensity (due to gradient) through the particle results in a change in total momentum of the exiting light beam, lead to reaction force on the object, which pushed it at highest light intensity of beam. Adapted from [6].

After interaction with object light rays bend according to the law of reflection and refraction. This changing direction of a beam changes the momentum of exiting light beam i.e. prior to entering into the object, rays travel parallel with zero vertical momentum, but after defection they pick up vertical momentum because any change in the direction of light by reflection or refraction, will results in a change in the momentum of light. For each ray in the beam the component of momentum flux parallel and perpendicular to the optic axis are given as:

$$(dp/dt)_{parallel} = n_B(W/c)\ cos\ \theta$$

and

$$(dp/dt)_{transverse} = n_B(W/c)\ sin\ \theta$$

Where, W is the power of laser light, c is speed of light, θ is the angle of ray to the optic axis, and n_B is the refractive index of the surrounding medium. As momentum is a conserve quantity, change of momentum of the light by its interactions with the object will exert an equal and opposite change momentum on the object, which gives rise to a force that pulls bead into the highest intensity region of light beam. The total force on an object due to refraction of light is:

$$F_{bead} = -\sum_{rays} (dp/dt)_{in} - (dp/dt)_{out}$$

Where, $\sum_{rays} =$ $Momentum\ flux\ over\ all\ rays\ passing\ through\ the\ bead.$

The sum of the forces from all rays get split into two components; F_s, the scattering force, pointing in the direction of incident light *(z, see axis in figure 2-1B),* and F_g, the gradient force, arising due to the gradient of the Gaussian intensity profile and pointing in the direction of the x-y plane towards the centre of the beam *(dotted line).* The gradient force is a restoring force that pull the bead at the centre of the beam. If the contribution to F_s of the refracted rays is larger than that of the reflected rays by focusing the light then a restoring force is also created along the z-axis, and stable trap will exist. Whenever the bead gets displaced from the trap center, the restoring force of optical trap works like an optical spring which pulls it back to the center. The foremost condition to make OT works is gradient force should be greater than scattering force ($F_g > F_s$). The equilibrium position is reached when the scattering force and gravity (which act to push bead downwards) is balanced by the axial gradient force (which pushes the bead upwards) [6].

To trap bead at a particular point in space, the beam of light must be focus *(Figure 2-1B).* However, in this situation scattering forces due to reflected rays pushes bead in forward direction. Therefore, for the effective trap in a single-beam trap, it is essential that the back aperture of the objective must be completely filled by the light beam so as to compensate the scattering force, the failure of this condition lead to the axial escape of the bead from the trap. Thus, it is necessary to have large focal cone angles of the beam, larger is the focal cone angle, stronger is the backward forces that oppose escape of the bead from the trap. For the stable trap and to avoid the use of large

focal cone angles of the beam, in 2003 *Smith and coworkers* developed new method called dual-beam trap which uses two opposite laser beams to trap an object at common foci of the beam *(Figure 2-2A).*

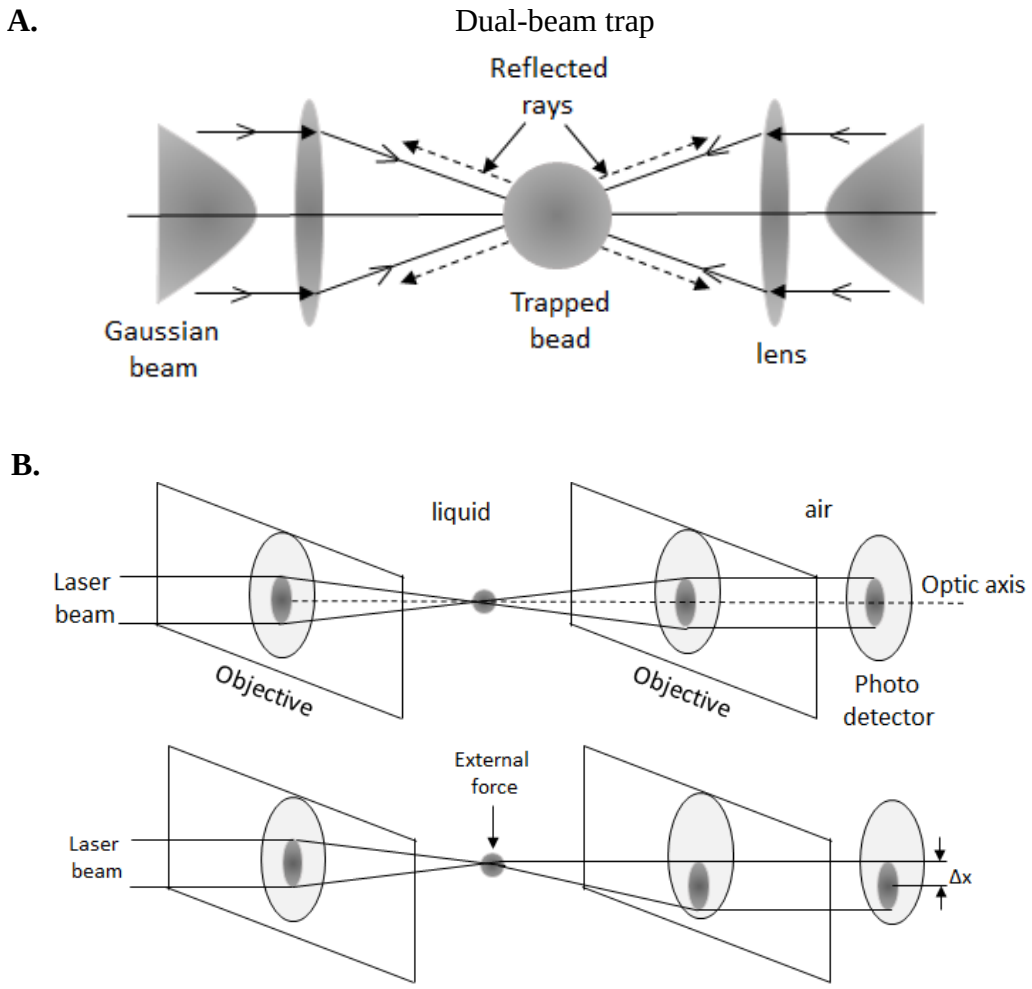


Figure 2-2 Dual-beam optical trap. A) Two low NA laser beams allowed to pass through back aperture of high NA ojectives and bead is trapped at their commom foci. In this method, scattering forces caused by reflected rays get cancelled, while gradient forces gets double up and stable trap is formed. B) Light momentum force sensor. In this set up, each objective used twice, to focus one beam and to collect other beam for analysis. Narrow cones used here allows the collection of nearly all the light even after singificant deflection of the beam. When the external force is applied on the bead, the bead moves slightly downwards and the light is refracted asymetrically. By collecting all the light momentum leaving the trap, it is possible to measure the force acting on

the bead as F= (W/c)* (x/RL), where W is the light intensity, c is the speed of light, x is the offset measured by photodetector and RL is the focal length of the lens. Adapted from [7].

In this method, two laser beams can be aligned to generate a dual-beam gradient trap. It consists of two microscopic objectives facing each other which focuses two separate laser beams at the common spot. Scattering forces generated by reflected rays (providing equal laser power) are canceled out, while transverse gradient forces get double up and stable trap is generated. This method uses low numerical aperture beams that are focused by high numerical aperture objectives. Each objective used here has dual functions; to focus one laser beam and collect the other for analysis. Directly measuring the change in momentum flux of trapping beam, the force applied to trapped object is determined. Developed by Smith and co-workers this method represents the best way to measure optical trap forces than previous methods [8-10], as it does not depend on particle's size and shape, buffer viscosity, refractive index, or laser power variation. One cannot apply this technique to measure forces in a single-beam trap as single beam of light with small cone angle cannot trap a bead effectively [7].

Dual-beam OT is used to characterize force induced unfolding and refolding of HIV-1-protease and ACBP.

OT setup at University of Modena and Reggio Emilia

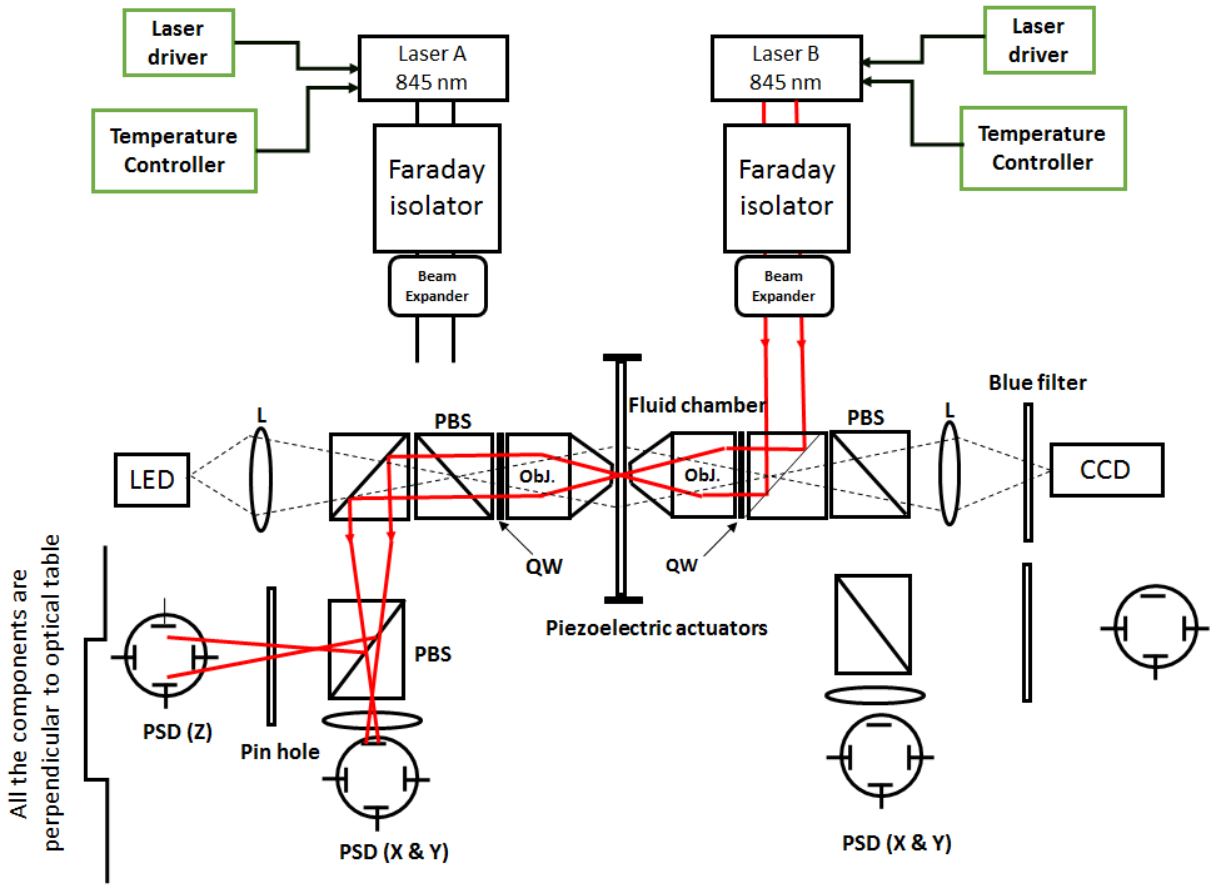


Figure 2-3 Schematic of dual beam OT setup at University of Modena and Reggio Emilia. The system is described in details elsewhere [11].

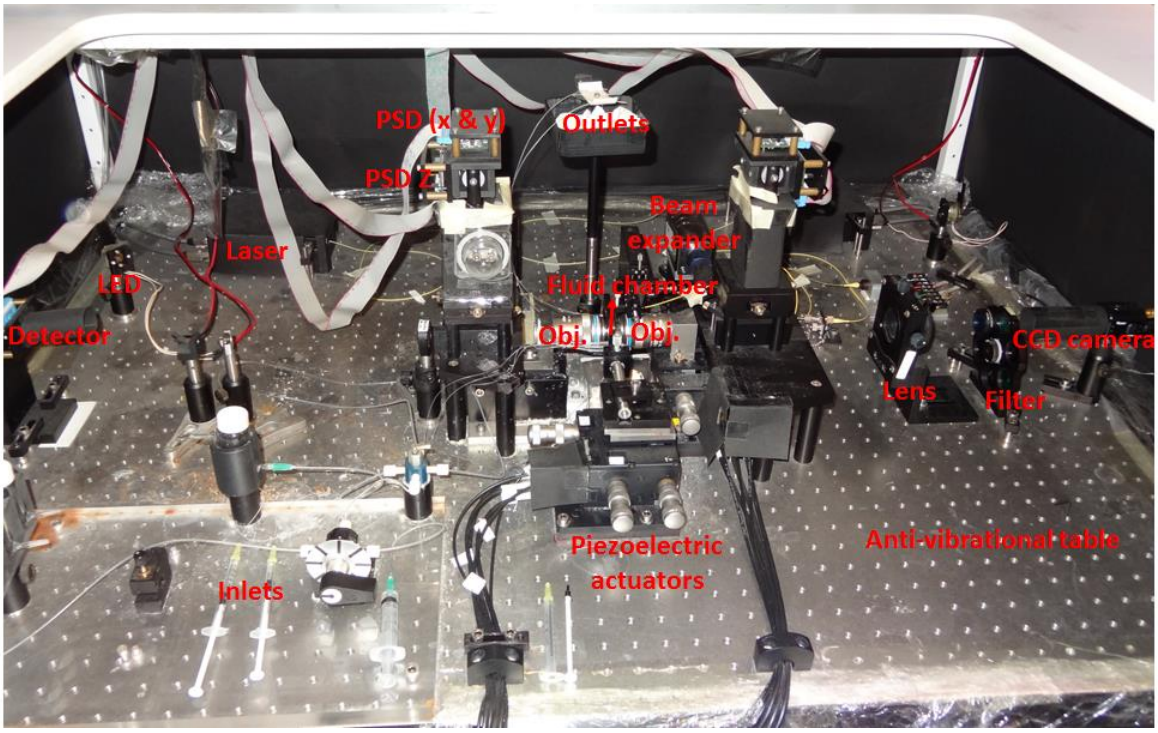


Figure 2-4 Picture of dual-beam OT setup at University of Modena and Reggio Emilia.

Fluid chamber

An experiment is carried out in fluid chamber. It contains two coverslips spaced 200 μM apart by parafilm layers and sealed by heat. The pipette is drawn from ~100 μm glass tubing down to point with an opening of ~0.5 μM. It consists of three channels: Upper, middle and lower *(Figure 2-5)*.

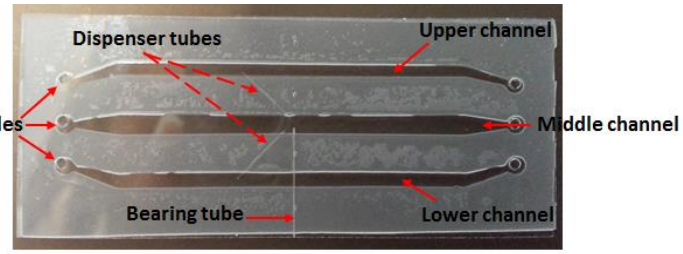


Figure 2-5 Picture of a fluid chamber (24mm x 60mm x 0.2 thick).

Each channel has one inlet and outlet for fluid. Streptavidin beads are flown through lower channel, while anti-dig beads with molecule are flown through upper channel. Two dispenser tubes present in the chamber; connecting lower and upper channel to middle channel that control the flows of beads. One bead is held at the end of glass micropipette by means of suction while another

one is trapped at the common foci of two objectives. Molecule of interest is stretched and relaxed multiple times by moving fluid chamber, and thus the micropipette to the optical trap. The movement of the fluid chamber is controlled by a piezoelectric actuator. The movement of two beads determines the extension of the molecule, while force applied on the protein is determined by measuring the change in the momentum flux of the light beams leaving the trap. The bead on micropipette is monitored by a "light lever" system *(Figure 2-6)* that track the position of the fluid chamber. During the experiment, beads are micropipette is visualized using LED (λ= 465 nm) and a CCD camera.

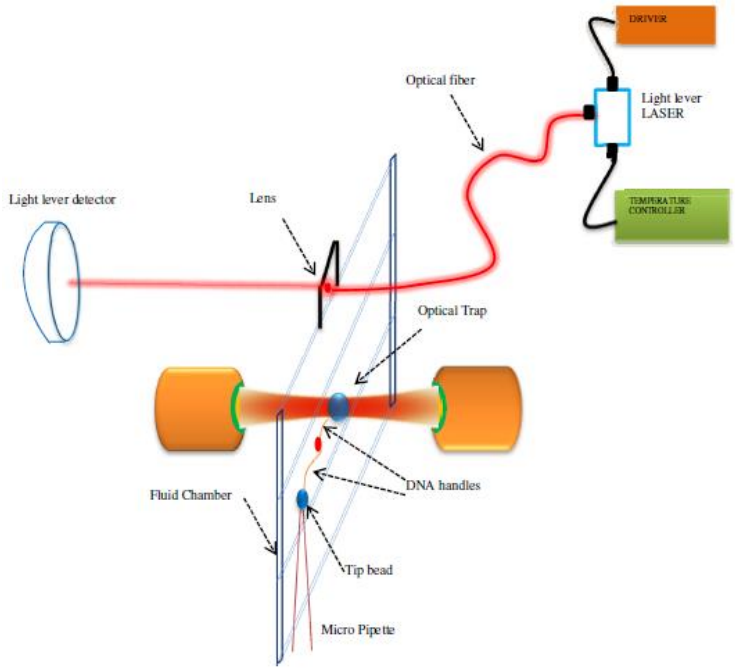


Figure 2-6 Light lever system. Adapted from [11].

Part – B: Model Systems

Introduction

The selection of a good model system is an essential step in research. About one-third genomic proteins are single-domain proteins. A single-domain protein is a relatively simple system that allows us to study basic principles of protein folding that would not be easily accessible using more complex multi-domain proteins. For the work presented in this thesis, we have chosen two different single-domain proteins namely, *Human Immunodeficiency Virus-1 Protease (HIV-1-PR)* and *Acyl-CoA-binding protein (ACBP).* The wealth of structural information already available for both proteins makes them particularly well-suited to study unfolding/folding reactions at single-molecule level. These proteins are described in detail below.

Human Immunodeficiency Virus type 1-Protease

Before going into the details of HIV-1-PR, first briefly I will explain the life cycle of HIV and importance of HIV-1-PR in the life cycle of the HIV virus.

Approximately three decades ago Acquired immunodeficiency syndrome (AIDS) was first time reported by the US Centre for Diseases control. Few years later it was found that Human immunodeficiency virus (HIV) is a causative agent for AIDS [12]. HIV was characterised into two main categories, HIV-1 and HIV-2. It was found that, compare to HIV-2, HIV-1 was more infectious and virulent. In a short time, it spread around the globe taking over millions of lives. After the invention of this dread disease, enormous efforts have been made to develop effective therapeutics for AIDS. Time to time different classes of anti-HIV drugs has been designed. Major classes were first HIV entry inhibitors (EIs), the nucleoside reverse transcriptase inhibitors (NRTIs), the non-nucleoside reverse transcriptase inhibitors (NNRTIs), maturation inhibitors (MIs) and Integrase inhibitors (IIs). Each class of these drug acts as a particular aspect of the viral life cycle, used to increase therapy efficacy therefore decrease the viral resistance. Despite of all efforts, unfortunately there is no cure for HIV and still it remains a serious public problem.

Life cycle of HIV and Role of HIV-1-PR

HIV belongs to retroviruses class. Life cycle of HIV consists of following steps. i) HIV virus enters in the body by various ways, then it finds T-cells and attaches to the receptors on the cell surface. ii) virus then enters into the cell's cytoplasm and RNA is released. iii) reverse transcription occurs

and cDNA is produced. iv) cDNA further moves from the cytoplasm to the nucleus and integrates into host DNA. v) DNA undergoes transcription and translation and newly synthesized viral protein chains are produced. vi) newly synthesized long precursor strand requires a specific enzyme for cutting activity. This task is performed by HIV-1-PR, it cuts long chains of proteins into more manageable lengths that will then produce new active copies of the virus. vii) protein gathers into vesicles called virions and are then released from cells. viii) virus moves on and spread infection to other cells.

HIV-1-PR: Target for anti-AIDS therapies

From the life cycle of HIV, it is very clear that HIV-1-PR plays an essential role [13], without the activity of HIV-1 protease HIV virions remain uninfecious. Any mutation in HIV protease or inhibition of its activity disrupts HIV's ability to replicate and further infection to other cells. Therefore, since 1990's HIV-1-PR became prime target for anti-AIDS drugs therapies and emerged as a major class of HIV protease inhibitors (PIs) today. Protease inhibitors help to stop the protease from cutting the proteins, so new copies of the virus do not contain any protein chains, and therefore it remains non-infectious. There are total eight approved HIV PIs; amprenavir, atazanavir, fosamprenavir, saquinavir, indinavir, lopinavir/ritonavir, nelfinavir, and ritonavir. Although effective during initial phase, these drugs were compromised by inconvenience, toxicity and less resistance to viral activity. One possible reason for ineffectiveness of these drugs could be a poor understanding of the molecular behavior of HIV-1-PR.

Structure

HIV-1-PR is a homodimer with 99 residues per subunit with only one binding site. Each subunit comprises of nine beta strands and one alpha helix. The X-ray crystallographic structure of the dimeric and monomeric form of HIV-1-PR is shown in *figure 2-7*. More than 140 structures of HIV-1-PR, its mutants, enzyme complex with inhibitors have been reported so far. Structural database of HIV-1-PR is available at the National Cancer Institute (http://www-fbsc.ncifcrf.gov/HIVdb/). Under biological conditions, HIV-1-PR takes about 50 seconds to reach its native conformation [14].

To study HIV-1-PR at single-molecule level following mutations has been made by standard genetic techniques; i) Arginine at position 87 was replaced by Lysine (R87K) to generate the monomeric form and ii) Leucine at position 5 and Asparagine at position 98 were replaced by

Cysteines (L5C, N98C) to attach DNA handles to perform optical tweezers experiments *(Figure 2-7).*

Figure 2-7 Structure of HIV-1-PR. A) Dimeric form of HIV-1-PR. B) A monomeric form of HIV-1-PR generated through site-directed mutagenesis, Arginine at position 87 replaced by Lysine (shown in green). Also, two amino acids Leucine and Asparagine at position 5 and 98 has been replaced by cysteine residues (shown in red).

Folding

Due to the importance of HIV-1-PR as a target for anti-AIDS therapies, enormous efforts have been taken in understanding the structures of the wild-type and variant HIV-1-PR forms. Likewise, the maturation process leading to the active dimer [15], compromising the folding of monomer, its cleavage from the polyprotein and the assembly of the monomers into the dimer, has been intensively investigated as each one of these steps could be potentially blocked to inhibit HIV replication [16, 17] . Simulations has suggested the existence of multiple pathways in unfolding and refolding processes of monomeric HIV-1-PR [18]. However, to date, no experimental evidence of this folding scenario has been provided. Folding mechanism of HIV-1-PR has been studied by a wide range of ensemble techniques such as differential scanning calorimetry, NMR [19], urea denaturation [20] and sedimentation equilibrium studies [21]. Although these studies have provided information on the overall thermodynamics and kinetics of the process, none has provided insight into the ensemble of folding routes that the individual molecules follow to reach their native state (N), as these experimental methods provide only average information. We address to this in *chapter 3*. Using optical tweezers and MD simulations, we characterised unfolding and refolding trajectories of monomeric HIV-1-PR at single-molecule level.

Acyl-CoA binding protein

Structure and functions

ACBP is a member of long chain fatty acid-CoA (LCFA-CoA) binding protein family. First time it was discovered in 1983 in brain and designated as diazepam binding inhibitor (DBI). ACBP found in a wide range of animals, including mammals, plants, yeasts, and insects. ACBP is a topologically simple and highly conserved 86 residue protein made of four alpha-helices (H_1-H_4). Four helices are organized in a unique up-down-down-up skewed bundle. Helix 2 (H_2) forms the central point and it has nonpolar contacts to other three helices. H_1 and H_4 run antiparallel to H_2 while H_3 runs almost parallel to H_2. Hydrophobic interactions between four helices form a stable three-dimensional structure which gives "shallow bowl" like appearance. The NMR solution structure of bovine ACBP is shown in *figure 2-8*.

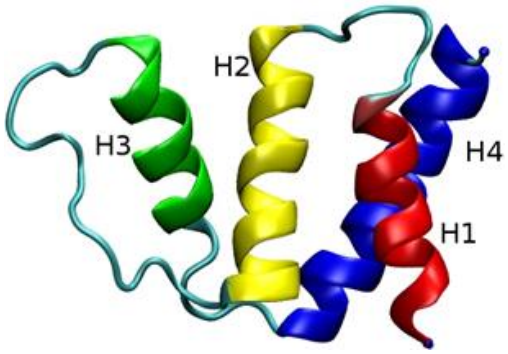


Figure 2-8 Acyl-Coenzyme A binding protein (ACBP).

The ACBP gene codes for an approximately 10 kDa protein that binds to acyl-CoA esters of chain length (C_4-C_{22}) with high affinity and specificity [22-24] . In vitro and in vivo experiments suggests that ACBP performs a wide range of functions, involved in multiple cellular tasks, for an instance ACBP has been recognized as being important for biosynthesis of fatty acids, membrane biogenesis, vesicular trafficking and donation of acyl-CoA esters for β oxidation [25-27]. It also involves in gene regulation, signal transduction, regulation of lipid and energy metabolism [28]. ACBP plays functional role in binding and transporting of acyl-CoA esters [25]. In vitro experiments showed that ACBP acts as an intracellular acyl-CoA transporter and pool former for acyl-CoA esters in order to function as intermediates in lipid metabolism [27, 29, 30]. Because of all these functions and knowing the fact that the genomic ACBP gene has all the characteristics of a housekeeping gene [31], indicate that ACBP is a housekeeping protein involved in acyl-CoA

metabolism, and it is not a protein with specific functions confined to specific cells [26, 32, 33]. Since its significant contribution in many biological processes, ACBP has been the subject of considerable interest for decades. In human ACBP gene is responsible for various diseases such as anxiety disorder with panic attack and type 2 diabetes [33, 34].

ACBP-Ligand complex

Present at several hundreds of μM in the cell, the majority of acyl-CoA esters are bound with ACBP. ACBP has specific binding sites for ligands located in the bottom inner surface. The arrangement of ligand is so that polar parts are exposed to the surface of the complex whereas the hydrophobic groups turn inward. A unique electrostatic network involving side chain of Try28, Lys32, Lys54 and 3'-phosphate plays key role in binding. More detailed information about ACBP-ligand interaction is given in [24]. ACBP-ligand complex has been studied by a variety of bulk methods [22, 35-37]. Using titration microcalorimetry, *Rasmussen et al.* showed that binding of acyl-CoA esters to ACBP is strongly chain length dependent. Rasmussen and coworkers were successful to determine the binding affinity of ACBP with acyl-CoA esters of different chain lengths [22]. Dissociation constant for ACBP-Octanoyl CoA (C_8), ACBP-dodecanoyl CoA (C_{10}) and ACBP-hexadecanoyl CoA (C_{16}) was found to be $0.24*10^{-6}$ M, $(0.65\pm 0.2)*10^{-8}$ M and $(0.45\pm 0.2)*10^{-13}$ M respectively. This study not only showed that binding affinity increases with increasing chain length but provided also evidence which clearly demonstrates that ACBP can mediate transport of long-chain acyl CoA esters from one membrane to another membrane. NMR studies by *Kragelund et al.* have also provided three-dimensional structure of recombinant bovine ACBP complexed with hexadecanoyl-CoA (C_{16}) [24]. Despite of all these information mechanical properties of ACBP-ligand complex remains unexplored. Our understanding of ACBP-ligand complex is still incomplete.

Folding

ACBP have a built-in pliability which is thought to be important both for binding to ligands of different size and length and also for ligand delivery. *Heidarsson et al.* were the first one to study the mechanical properties and the energy landscape of ACBP at single-molecule level using optical tweezers and ratcheted molecular dynamic simulation [38]. The results of their studies showed that, when manipulated from its N- and C-termini, ACBP unfolds and refolds in two-state manner and it is characterised by the more extended transition state (TS) to date known for a single globular

protein. The position of TS was observed about halfway between the native state and unfolded state, around 5.5 nm from native state. Ratcheted MD simulation has also provided a possible structure of the transition state. Though this study was successful, but it was restricted to the characterization of unfolding/refolding trajectories and the energy landscape of ACBP alone *that is* in the absence of its ligand. This study further creates many interesting questions for the ACBP-ligand complex. For an instance, how ligand binding will affect the mechanical properties of ACBP, i.e. unfolding and refolding of ACBP will still be a two-state process in the presence of ligand? Will there be any change in the mechanical stability upon ligand binding? We started to address this in *chapter 4*. Using optical tweezers and molecular dynamic simulations, we have investigated the effect of ligand binding on unfolding and refolding processes of ACBP.

Protein sequences

The amino acid sequences of HIV-1-PR and ACBP along with the sites of mutations are shown in *figure 2-9 and 2-10*. All mutations were generated by standard genetic engineering techniques.

HIV-1-PR

HIV-1-PR: Dimeric form

PQITLWKRPLVTIRIGGQLKEALLDTGADDTVIEEMNLPGKWKMIGGIGGFIKVRQYDQYDQIIIEICGHKAIGT
VLVGPTPVNIIGRNLLTQIGATLNF

HIV-1-PR: Monomeric form

5C
PQITLWKRPLVTIRIGGQLKEALLDTGADDTVIEEMNLPGKWKMIGGIGGFIKVRQYDQYDQIIIEICGHKAIGT
VLVGPTPVNIIGRNLLTQIGATLNF
K 98C

Figure 2-10 Protein sequence and sites of mutation of HIV-1-PR. The sites of Cysteine substitutions that define the pulling axes are indicated in red. To generate monomeric form of HIV-1-PR Arginine at position 87 was replaced by Lysine (R87K), is shown in blue.

ACBP

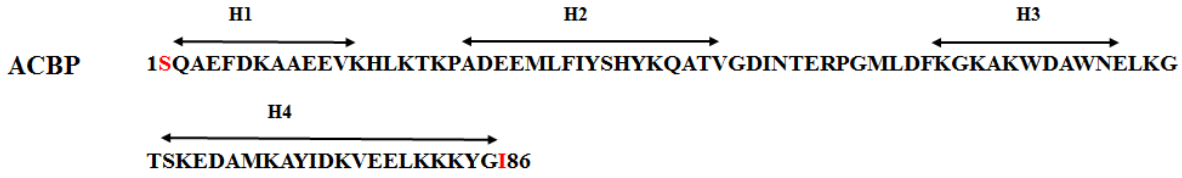


Figure 2-10 Protein sequence and sites of mutation of ACBP. The sites of Cysteine substitutions that define the pulling axes are indicated in red.

References

[1] A. Ashkin, J.M. Dziedzic, J.E. Bjorkholm, S. Chu, Observation of a single-beam gradient force optical trap for dielectric particles, Optics Letters, 11 (1986) 288-290.

[2] A. Ashkin, J. Dziedzic, Optical trapping and manipulation of viruses and bacteria, Science, 235 (1987) 1517-1520.

[3] C. Xie, Y.-q. Li, Confocal micro-Raman spectroscopy of single biological cells using optical trapping and shifted excitation difference techniques, Journal of Applied Physics, 93 (2003) 2982-2986.

[4] Remus T. Dame, Maarten C. Noom, G.J.L. Wuite, Bacterial chromatin organization by H-NS protein unravelled using dual DNA manipulation, Nature, 444 (2006) 387-390.

[5] J.R. Moffitt, Y.R. Chemla, S.B. Smith, C. Bustamante, Recent Advances in Optical Tweezers, Annual Review of Biochemistry, 77 (2008) 205-228.

[6] S. Husale, Single biomolcule studies using optical tweezers, Ph.D. thesis (2005), University of Basel, Switzerland.

[7] S.B. Smith, Y. Cui, C. Bustamante, Optical-trap force transducer that operates by direct measurement of light momentum., Biophotonics Pt B, 361 (2003) 134-162.

[8] S.B. Smith, L. Finzi, C. Bustamante, Direct mechanical measurements of the elasticity of single DNA molecules by using magnetic beads, Science, 258 (1992) 1122-1126.

[9] K. Svoboda, C.F. Schmidt, B.J. Schnapp, S.M. Block, Direct observation of kinesin stepping by optical trapping interferometry, Nature, 365 (1993) 721-727.

[10] A. Ashkin, K. Schutze, J.M. Dziedzic, U. Euteneuer, M. Schliwa, Force generation of organelle transport measured in vivo by an infrared laser trap, Nature, 348 (1990) 346-348.

[11] I. Valpapuram, Single molecule optical tweezers studies of the unfolding/refolding processes of Acyl-Coenzyme A binding Protein, Ph.D. thesis (2012), University of Modena and Reggio Emilia, Modena, Italy.

[12] R.C. Gallo, L. Montagnier, Sci. Am, 259 (1988).

[13] J.R. Huff, HIV protease: a novel chemotherapeutic target for AIDS, Journal of Medicinal Chemistry, 34 (1991) 2305-2314.

[14] A.F. Noel, O. Bilsel, A. Kundu, Y. Wu, J.A. Zitzewitz, C.R. Matthews, The Folding Free-Energy Surface of HIV-1 Protease: Insights into the Thermodynamic Basis for Resistance to Inhibitors, Journal of Molecular Biology, 387 (2009) 1002-1016.

[15] S. Kimura, M. Caldarini, R.A. Broglia, N.V. Dokholyan, G. Tiana, The maturation of HIV-1 protease precursor studied by discrete molecular dynamics, Proteins: Structure, Function, and Bioinformatics, 82 (2014) 633-639.
[16] R.A. Broglia, Y. Levy, G. Tiana, HIV-1 protease folding and the design of drugs which do not create resistance, Current Opinion in Structural Biology, 18 (2008) 60-66.
[17] R.A. Broglia, G. Tiana, L. Sutto, D. Provasi, F. Simona, Design of HIV-1-PR inhibitors that do not create resistance: Blocking the folding of single monomers, Protein Science, 14 (2005) 2668-2681.
[18] M. Bonomi, A. Barducci, F.L. Gervasio, M. Parrinello, Multiple Routes and Milestones in the Folding of HIV–1 Protease Monomer, PLoS ONE, 5 (2010) e13208.
[19] S.C. Panchal, N.S. Bhavesh, R.V. Hosur, Real time NMR monitoring of local unfolding of HIV-1 protease tethered dimer driven by autolysis, FEBS letters, 497 (2001) 59-64.
[20] S.K. Grant, I.C. Deckman, J.S. Culp, M.D. Minnich, I.S. Brooks, P. Hensley, C. Debouck, T.D. Meek, Use of protein unfolding studies to determine the conformational and dimeric stabilities of HIV-1 and SIV proteases, Biochemistry, 31 (1992) 9491-9501.
[21] S. Gulnik, J.W. Erickson, D. Xie, HIV protease: Enzyme function and drug resistance, Vitamins & Hormones, Academic Press2000, pp. 213-256.
[22] J.T. Rasmussen, N.J. Faergeman, K. Kristiansen, J. Knudsen, Acyl-CoA-binding protein (ACBP) can mediate intermembrane acyl-CoA transport and donate acyl-CoA for beta-oxidation and glycerolipid synthesis, The Biochemical journal, 299 (Pt 1) (1994) 165-170.
[23] J. Mikkelsen, J. Knudsen, Acyl-CoA-binding protein from cow. Binding characteristics and cellular and tissue distribution, Biochem. J., 248 (1987) 709-714.
[24] B.B. Kragelund, K.V. Andersen, J.C. Madsen, J. Knudsen, F.M. Poulsen, Three-dimensional Structure of the Complex between Acyl-Coenzyme A Binding Protein and Palmitoyl-Coenzyme A, Journal of Molecular Biology, 230 (1993) 1260-1277.
[25] M. Burton, T.M. Rose, N.J. Faergeman, J. Knudsen, Evolution of the acyl-CoA binding protein (ACBP), Biochem. J., 392 (2005) 299-307.
[26] R.E. Gossett, A.A. Frolov, J.B. Roths, W.D. Behnke, A.B. Kier, F. Schroeder, Acyl-CoA binding proteins: multiplicity and function, Lipids, 31 (1996) 895-918.
[27] N.J. Faergeman, S. Feddersen, J.K. Christiansen, M.K. Larsen, R. Schneiter, C. Ungermann, K. Mutenda, P. Roepstorff, J. Knudsen, Acyl-CoA-binding protein, Acb1p, is required for normal

vacuole function and ceramide synthesis in Saccharomyces cerevisiae, The Biochemical journal, 380 (2004) 907-918.
[28] J. Knudsen, M. Jensen, F\sgmaelig, N. rgeman, T.F. Neergaard, B. Gaigg, Role of acylCoA binding protein in acylCoA transport, metabolism and cell signaling, Mol Cell Biochem, 192 (1999) 95-103.
[29] J. Knudsen, J. Burton, N.J. Færgeman, Long chain acyl-CoA esters and acyl-CoA binding protein (ACBP) in cell function, Lipobiology (van der Vusse, G., ed.), Elsevier B.V., Amsterdam, 2004, pp. 123–153.
[30] S. Mandrup, N.J. Færgeman, J. Knudsen, Structure, function and phylogeny of acyl-CoA binding protein, Cellular Proteins and Their Fatty Acids in Health and Disease (Duttaroy, A. K. and Spener, F., eds.), Wiley-VCH gmbH & Co. Kaga, Weinheim2004, pp. 151–172.
[31] S. Mandrup, R. Hummel, S. Ravn, G. Jensen, P.H. Andreasen, N. Gregersen, J. Knudsen, K. Kristiansen, Acyl-CoA-binding protein/diazepam-binding inhibitor gene and pseudogenes: A typical housekeeping gene family, Journal of Molecular Biology, 228 (1992) 1011-1022.
[32] B.B. Kragelund, J. Knudsen, F.M. Poulsen, Acyl-coenzyme A binding protein (ACBP), Biochim Biophys Acta, 1441 (1999) 150-161.
[33] E. Fisher, I. Nitz, C. Gieger, H. Grallert, H. Gohlke, I. Lindner, S. Dahm, H. Boeing, B. Burwinkel, W. Rathmann, H.E. Wichmann, J. Schrezenmeir, T. Illig, F. Döring, Association of acyl-CoA-binding protein (ACBP) single nucleotide polymorphisms and type 2 diabetes in two German study populations, Molecular Nutrition & Food Research, 51 (2007) 178-184.
[34] J. Bhuiyan, P.H. Pritchard, S.V. Pande, D.W. Seccombe, Effects of high-fat diet and fasting on levels of acyl-coenzyme a binding protein in liver, kidney, and heart of rat, Metabolism, 44 (1995) 1185-1189.
[35] J.T. Rasmussen, T. Börchers, J. Knudsen, Comparison of the binding affinities of acyl-CoA-binding protein and fatty-acid-binding protein for long-chain acyl-CoA esters, Biochem. J., 265 (1990) 849-855.
[36] J. Rosendal, P. Ertbjerg, J. Knudsen, Characterization of ligand binding to acyl-CoA-binding protein, The Biochemical journal, 290 (Pt 2) (1993) 321-326.
[37] N.J. Færgeman, B.W. Sigurskjold, B.B. Kragelund, K.V. Andersen, J. Knudsen, Thermodynamics of Ligand Binding to Acyl-Coenzyme A Binding Protein Studied by Titration Calorimetry†, Biochemistry, 35 (1996) 14118-14126.

[38] P.O. Heidarsson, I. Valpapuram, C. Camilloni, A. Imparato, G. Tiana, F.M. Poulsen, B.B. Kragelund, C. Cecconi, A Highly Compliant Protein Native State with a Spontaneous-like Mechanical Unfolding Pathway, Journal of the American Chemical Society, 134 (2012) 17068-17075.

Chapter: 3

The complex folding behavior of HIV-1-protease monomer revealed by optical-tweezer single-molecule experiments and molecular-dynamic simulations

Chapter 3: The complex folding behavior of HIV-1-protease monomer revealed by optical-tweezer single-molecule experiments and molecular-dynamic simulations

Abstract

We have used optical tweezers and molecular dynamics simulations to investigate the unfolding and refolding process of a stable monomeric form of HIV-1-protease (PR). We have characterized the behavior under tension of the native state (*N*), and that of the ensemble of partially folded (*PF*) conformations the protein visits *en route* to *N*, which collectively act as a long-lived state controlling the slow kinetic phase of the folding process. Our results reveal a rich network of unfolding events, where the native state unfolds either in a two-state manner or by populating an intermediate state *I*, while the *PF* state unravels through a multitude of pathways, underscoring its structural heterogeneity. Refolding of mechanically-denatured HIV-1-PR monomers is also a multiple-pathway process. Molecular dynamics simulations allowed us to gain insight into possible conformations the protein adopts along the unfolding pathways, and provide information regarding possible structural features of the PF state.

Keywords: Single molecule studies, protein folding, HIV-1-protease, intermediate states, MD simulations

This work is published in Biophysical Chemistry journal, 195 (2014) 32-42.

Introduction

Human Immunodeficiency Virus type 1 protease (HIV-1-PR) is a dimeric aspartyl protease, made of two identical monomers of 99 residues each, which plays an essential role in the life cycle of HIV [1]. The virus expresses its protease embedded in a polyprotein, where it is largely monomeric [2] and the monomer adopts a fold very similar to that observed in the functional and mature dimer [3-5]. The structures of the wild-type and variant HIV-1-PR forms have been widely studied because of the importance of this enzyme as a target for antiviral therapy [6, 7]. Likewise, the maturation process leading to the active dimer [8], comprising the folding of the monomer, its cleavage from the polyprotein and the assembly of the monomers into the dimer, has been intensively investigated as each one of these steps could be potentially blocked to inhibit HIV replication [9-11].

In spite of its small size, the folding process of the HIV-1-PR monomer seems quite complex as compared to that of typical proteins of similar length. Under biological conditions, it takes about 50 s to reach its native conformation, not involving any proline isomerization, and events occurring at least on two different time scales were identified in stopped-flow fluorescence experiments [12]. The native conformation itself appears more complicated than that of proteins of similar size. In fact, if we quantify the complexity of the native topology with the Plaxco's contact order *CO [13]*, we find that compared to proteins of similar size, such as ACBP (*CO=11*), protein G (*CO=9*), CI2 (*CO=10*) and SH3 (*CO=10*), the monomer of the HIV-1-PR has a higher *CO* of 15.

Simulations have suggested the existence of multiple pathways in the unfolding and refolding processes of HIV-1-PR monomer [14]. However, to date no experimental evidence of this complex folding scenario has been provided. Several studies have examined the folding mechanism of HIV-1-PR using a variety of ensemble techniques, such as differential scanning calorimetry [15], NMR [16], urea denaturation [17] and sedimentation equilibrium studies [18]. Although these studies have given information on the overall thermodynamics and kinetics of the process, none have provided insight into the ensemble of folding routes that the individual molecules follow to reach their native state (*N*), as these experimental methods provide only averaged information. The advent of single molecule techniques, such as optical tweezers, has made it possible to revisit protein folding with a new approach, allowing us to go beyond the

ensemble average and to dissect folding mechanisms in unprecedented detail. Through mechanical manipulation, it is possible to follow in real time the trajectories of individual molecules and to describe the inherent heterogeneity of biological processes that are stochastic in nature, such as the folding of a polypeptide chain into a functional protein [19-24]. Moreover, these studies can be complemented by state-of-the-art molecular dynamics (MD) simulations, which, through a comparative analysis with the experimental data, can provide atomic details of the process under investigation [25, 26].

In this study, we used optical tweezers and MD simulations to shed light on the HIV-1-PR folding mechanism. We manipulated a stable monomeric form of HIV-1-PR to show that it unfolds mainly along two trajectories, sometimes populating a mechanically stable intermediate state. Taking advantage of the slow folding kinetics of HIV-1-PR observed by fluorescence and circular-dichroism experiments [12], we were also able to manipulate partially folded (*PF*) conformations that the molecule adopts along its journey to the native state. Our data indicate that the *PF* state is long lived, compact and structurally inhomogeneous, unfolding along a multitude of trajectories when mechanically stretched. State-of-the-art MD simulations revealed the sequence of events characterizing the unfolding and refolding trajectories, providing insight into possible structural features of the intermediate states populated by the protein and suggesting a role for non-native contacts in the unfolding process. The contents of this chapter are taken from [27] with permission.

Results

In order to study the folding process of HIV-1-protease at the single-molecule level while avoiding the complication of monomer/dimer equilibria [28], the R87K substitution was introduced [29] in the Q7K/C67A/C95A protease variant [30]. Cysteine residues were also introduced at the N- and C-terminal ends of the protein by engineering the L5C and N98C substitutions. Urea denaturation of the protease variant, as monitored in solution by tryptophan fluorescence, was reversible and well described by a two-step process *(Figure 3-1)*, as previously reported for both the well-characterized dimeric, but inactive, HIV-1 protease variant and the monomeric (inactive) protein form described by *Noel et al.* [12]. However, the mutant is marginally stable (ΔG^{o}=9.84±1.34 kJ/mol) and prone to aggregation. Thus, suitable preparation protocols had to be set up to generate protein-DNA chimeras for use in single-molecule experiments *(see section "Materials and methods" for details)*.

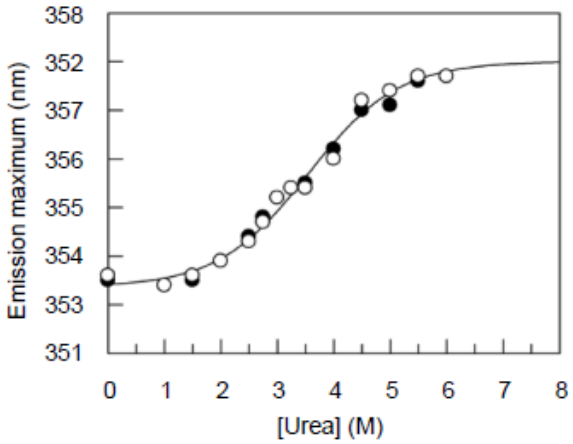


Figure 3-1 Equilibrium folding properties of the L5C/R87K/N98C HIV-1-PR variant. Fluorescence emission of protease samples (2 µM in 20 mM sodium phosphate buffer, pH 6.0, 1 mM DTT) containing increasing (empty symbols) or decreasing (closed symbols) urea concentrations was measured at 25 °C (excitation light, 290 nm) at equilibrium. The wavelength of the maximum of emission spectrum was plotted as a function of urea concentration. The curve shows the common fit to a two state model *(equation 3-1)* with a ΔG^{o} for folding of -9.84 ±1.34 kJ/mol and an *m* value of 2.85 ± 0.42 kJ/mol.

Single monomers of HIV-1-PR were manipulated as depicted in *figure 3-2*. The molecules were stretched and relaxed by moving the pipette relative to the optical trap, while the applied force and molecular extension were measured as previously described [19, 31]. As the protein unfolds and refolds under tension its extension suddenly changes, generating discontinuities (rips) in the stretching and relaxation traces that can be analyzed to gain insight into the kinetic and mechanical properties of the molecule [20]. We used this experimental approach to explore the unfolding and refolding processes of native and partially folded conformations of the monomeric HIV-1-PR.

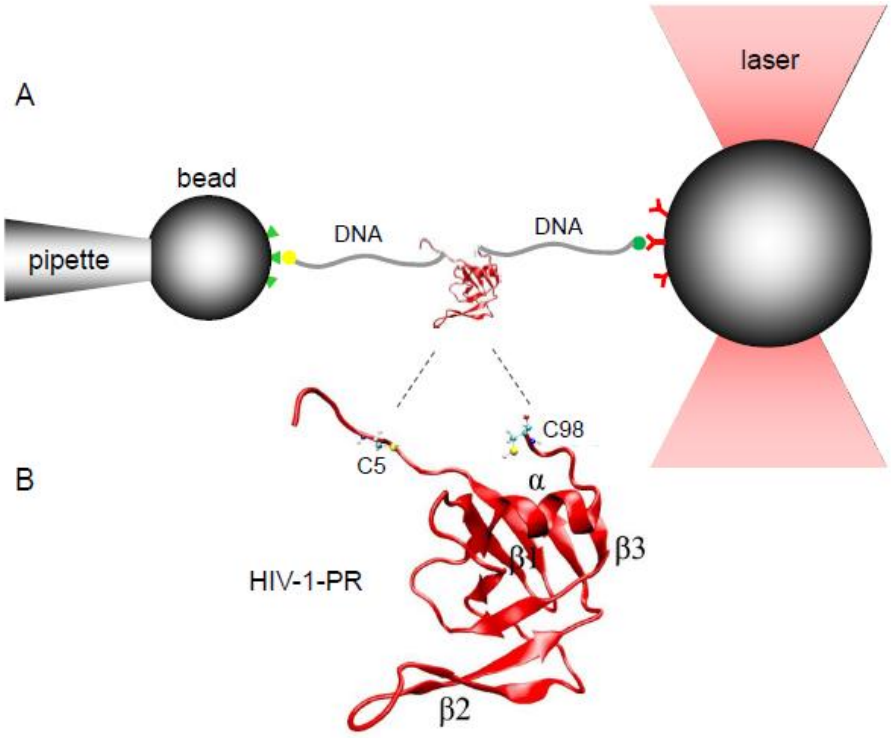


Figure 3-2 Experimental setup and structure of HIV-1-PR. A) Individual molecules were manipulated by means of DNA handles (~ 500 bp DNA molecules) covalently attached to cysteine residues engineered at positions 5 and 98. Each DNA handle is tethered to a functionalized bead, as previously described [19, 32]. B) NMR structure of the monomeric form of HIV-1-PR (Protein Data Bank code 1BVG).

The native state unfolds along two main pathways

To study the unfolding trajectories of the native state of monomeric HIV-1-PR, experiments were performed in which the force applied on the protein was first ramped (~ 5 pN/sec) to 55 pN to induce the unfolding of the molecule; then, it was decreased (~ -5 pN/sec) to 2 pN and kept constant at such low value for 90 s, before the next stretching-relaxation cycle started. In the absence of force the HIV-1-PR monomer takes ~ 50 s to refold into its native state from an urea-induced denatured state [12]. In the present experiments, the HIV-1-PR is given approximately twice as much time (90 s) to fold into *N* because under tension, although very small (2 pN), folding is slowed down [33, 34]. The force applied on the molecule is never lowered to zero, in order to avoid direct contact between the polystyrene beads, which would generate unwanted interactions between the tethering surfaces. The unfolding events observed during the first stretching of the

molecule, which corresponds to the unfolding of the native state, are indistinguishable from those observed during subsequent stretching-relaxation cycles, suggesting proper folding of HIV-1-PR during the 90 s wait period. It is worth mentioning that these experiments are technically very challenging for two main reasons. First, the interaction between digoxigenin and antibodies usually cannot stand forces larger than 25-30 pN. However, in spite of its poor thermodynamic stability, HIV-1-PR monomer is a mechanically resistant protein and force must be raised up to 55 pN to ensure complete unfolding. As a consequence, in each experiment many tethered molecules had to be tested to find one that could withstand such high tension for at least one cycle. Second, at 2 pN of force the two tethering surfaces are quite close to each other and in a 90 s period of time there is always a good chance that another molecule gets tethered between them, causing the end of the measurement.

Under these experimental conditions, the HIV-1-PR monomer was observed to unfold through two main mechanisms, as shown in *figure 3-3*. In 64% of the cases, it unfolded in an all-or-none fashion, transiting from *N* to the unfolded state (*U*) along a two-state unfolding pathway. Other times (31%) it unfolded in a three-state manner, populating an intermediate state (*I)*. Only rarely (5% of the cases) it unraveled through different mechanisms *(Figure 3-S1)*. During multiple stretching/relaxation cycles, the same molecule is observed to take different unfolding pathways *(Figure 3-3)*. When the HIV-1-PR unfolds in a two-state fashion, it generates a single discontinuity in the stretching trace at ~ 25 pN, which corresponds to the sudden increase in the extension of the molecule as it passes from its compact native state to its extended unfolded state (average rip width $d = 22 \pm 4$ nm), left trace in *figure 3-3B*. The increment in contour length (ΔL_c) upon unfolding of the molecule was estimated by fitting the force-extension curves to the worm-like chain model [35, 36]. This analysis yielded a ΔL_c of 30 ± 3 nm, which compares well with the value of 31.5 nm expected for the complete unfolding of the molecule *(see section "Materials and methods")*. Analysis of the force distribution of the two-state unfolding events *(Figure 3-3C)* provided information on the unfolding rate constant (k^0_u) and distance ($x^{\ddagger}_u$) between the native state and the transition state along the reaction coordinate [19, 37].

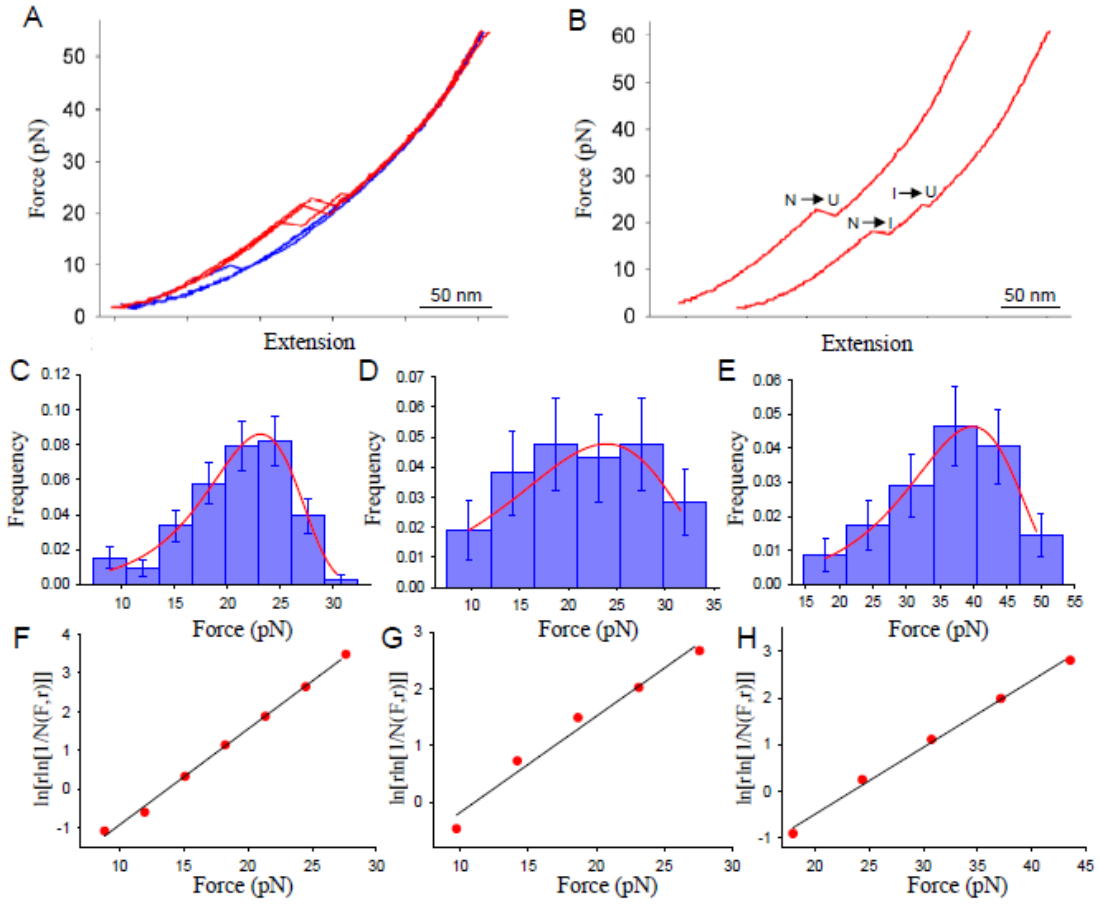


Figure 3-3 Forced unfolding of the monomeric HIV-1-PR native state. A) Three superimposed stretching (red) and relaxation (blue) cycles acquired by manipulating one HIV-1-PR molecule at 100 nm/s (~ 5 pN/sec). The molecule unfolded twice in a two-state manner, and once it populated an intermediate state *I*. B) Stretching traces from panel A) shifted laterally for clarity. The native state (*N*) transits into the unfolded state (*U*) through an all-or-none transition on the left trace, while it populates an intermediate state (*I*) in the other case. C) Force distribution of the two-state unfolding events (n = 105) fit to a probability distribution function [38], (red line, equation 2), to estimate the position of the transition state along the reaction coordinate and the zero force unfolding rate. The best fit (least squares) values gave $x^{\ddagger}_{u}$ = 1 ± 0.1 nm and k^{0}_{u} = 7 (+/-1) x 10^{-3} s^{-1}. D) and E) show the force distributions of the first and second unfolding events along the three-state unfolding pathway (n = 46) and the corresponding fit to the probability distribution function (red lines). The best fit values gave $x^{\ddagger}_{u}$ = 0.5 ± 0.1 nm and k^{0}_{u} = 4 (±1) x 10^{-2} s^{-1} for D), and $x^{\ddagger}_{u}$ = 0.5 ± 0.1 nm and k^{0}_{u} = 6 (±1) x 10^{-3} s^{-1} for E). Force distributions of C), D) and E) were also

linearized and analyzed as previously reported (equation 3) [25, 37] in F), G) and H) respectively. This analysis provided kinetic values quite similar to those reported above, that is $x^{\ddagger}_{u}$ = 1 ± 0.1 nm and k^{0}_{u} = 8 (±1) x 10^{-3} s^{-1} for F), $x^{\ddagger}_{u}$ = 0.7 ± 0.1 nm and k^{0}_{u} = 3 (±0.6) x 10^{-2} s^{-1} for G), and $x^{\ddagger}_{u}$ = 0.6 ± 0.1 nm and k^{0}_{u} = 5 (±0.7) x 10^{-3} s^{-1} for H).

However, it is important to notice that rate constants estimated through optical tweezers experiments contain contributions from experimental parameters, such as DNA handle length and bead size; thus, they cannot be compared directly to rate constants of the free protein in solution [37]. On the contrary, the estimated $x^{\ddagger}_{u}$ value reflects an intrinsic property of the molecule. The distance $x^{\ddagger}_{u}$ is a measure of the mechanical compliance of the native state, i.e.: how much it can be deformed along the pulling axis without crossing the transition state. The larger is $x^{\ddagger}_{u}$, the more pliant a protein structure is. Analysis of the force distribution of *figure 3-3C* yielded a $x^{\ddagger}_{u}$ of 1 ± 0.1 nm, indicating that the compliance of HIV-1-PR's native state is relative large compared to that of other all-β strand protein [39-41]. This unusual pliability of the three-dimensional structure, if maintained in the dimeric species, might play a role during the catalytic activity of the enzyme, allowing the protease to adopt the proper conformations to interact with its substrates.

When the HIV-1-PR unfolds in a three-state manner *(Figure 3-3B, right trace)*, it first populates an intermediate state *I* at ~ 25 pN (*N-I* transition; average rip width *d* = 17 ± 4 nm), and then it transits into its unfolded state at higher forces, (*I-U* transition; average rip width *d* = 7 ± 3 nm). The partial unfolding of the protein into *I* results in a change in contour length ΔL_c of 22 ± 4 nm, while the subsequent unfolding of the intermediate state results in a ΔL_c of 9 ± 3 nm. The sum of the two ΔL_c values (31 ± 5 nm) is consistent with the full unfolding of the protein. Assuming a value of 0.36 nm/residue, a ΔL_c of 9 ± 3 nm for the *I-U* transition indicates that ~ 25 residues are structured in the intermediate state [42]. Analysis of the unfolding force distributions provided kinetic information on the unfolding transitions *(Figure 3-3).*

Partially folded conformations unfold along multiple pathways

During the ~ 50 s needed to fold, the molecule likely explores different conformations before finding its native state. To gain insight into this ensemble of partially folded conformations, we performed measurements in which individual HIV-1-PR molecules were pulled before they could reach their native state. In these experiments, a tethered molecule was stretched and relaxed multiple times between an upper and lower force value (55 pN and 2 pN, respectively), with no

waiting time at low tension. Under these conditions, the whole stretching-relaxation cycle takes only 20 s to be completed. Thus, an unfolded HIV-1-PR has no time to fold back into its native state before it is pulled again, and the molecule is stretched while being in one of the many possible partially folded (*PF*) conformations it adopts *en route* to *N*, which collectively we call the *PF* state.

In these measurements, the HIV-1-PR monomer unfolds either through an all-or-none transition (23% of the cases) or, more often, in a three-state manner involving an intermediate state (71% of the cases), as shown in *figure 3-4*. Only rarely (6% of the cases), it unfolds through different mechanisms *(Figure 3-S1)*. The same molecule is observed to take different unfolding pathways when stretched and relaxed multiple times *(Figure 3-4)*.

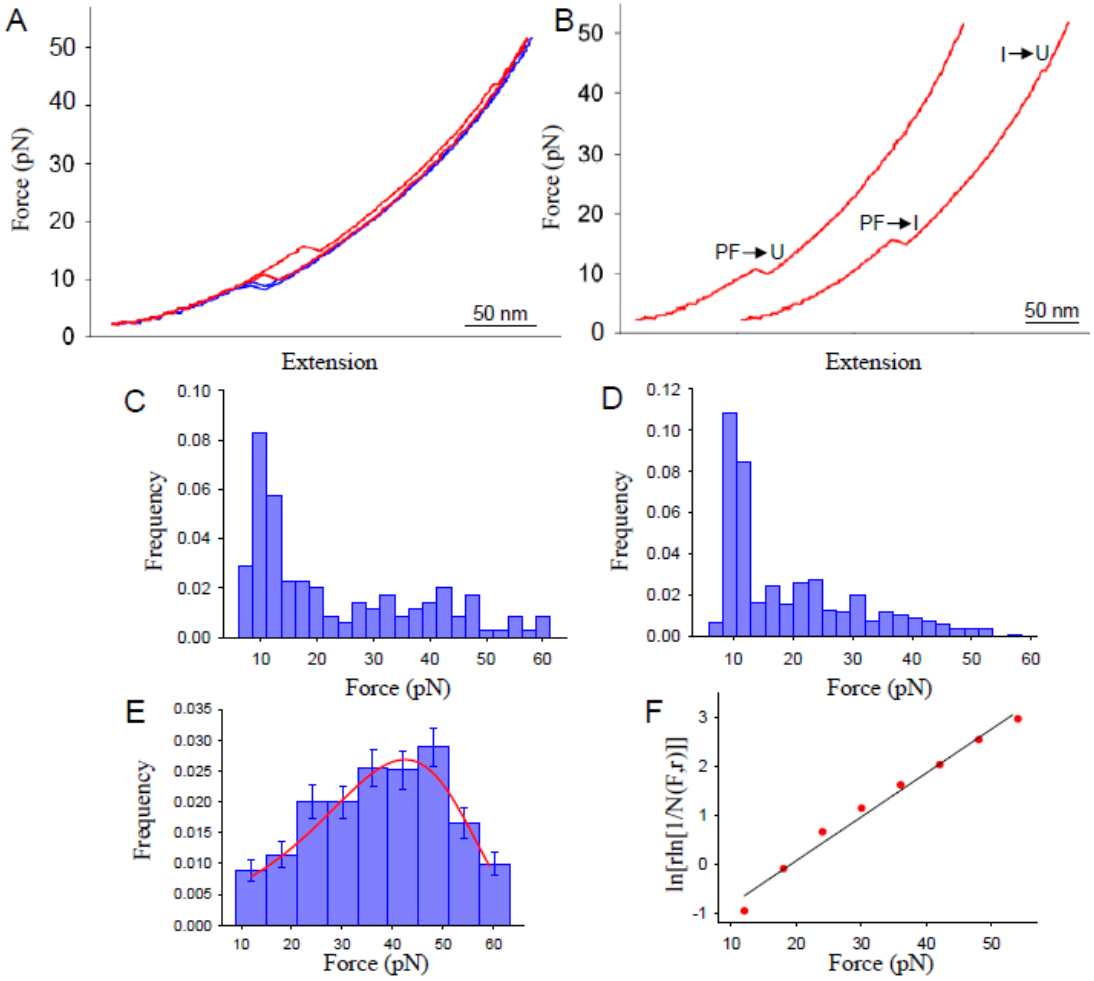


Figure 3-4 Forced unfolding of partially folded (*PF*) conformations of HIV-1-PR monomer. A) Two superimposed force vs extension cycles acquired by stretching and relaxing one HIV-1-PR molecule at 100 nm/s. B) Stretching traces of panel A), where the molecule transits into the

unfolded state either through an all-or-none transition (left trace), or by populating an intermediate state *I* (right trace). C) Force distribution of the two-state unfolding events (*n*=139). D) Force distribution of the first unfolding events along the three-state unfolding pathway (*n*=432). E) Force distribution of the second unfolding events along the three-state unfolding pathway (*n*=432). Fit of these data to a probability distribution function (red line, equation 2) [38] yielded $x^{\ddagger}_u$ = 0.3 ± 0.1 nm and k_0 = 3 (±0.5) x 10^{-2} s^{-1}. F) Plot of ln[*r* ln[1/*N(F,r)*]] *vs* force for the force distribution of E), and corresponding fit (black lines, equation 3) [37] yielding a $x^{\ddagger}_u$ = 0.4 ± 0.1 nm and k_0 = 2 (±0.2) x 10^{-2} s^{-1}.

The sharp transitions characterizing the stretching traces indicate that the *PF* state is a well-defined thermodynamic state, separated from other states of HIV-1-PR by free-energy barriers. The ΔL_c associated to the two-state unfolding events *(Figure 3-4B, left race)* was estimated to be 31 ± 3 nm, a value that is quite similar to that measured for the unfolding of the native state, revealing a native-like extension of the *PF* conformations. Similarly, the changes in contour length associated to the three-state unfolding pathway were estimated to be 22 ± 3 nm for the first transition, and 7 ± 2 nm for the second one, summing to a total ΔL_c of 29 ± 4 nm.

The ensemble of conformations associated with the *PF* state are, overall, less stable than the native state, unfolding mostly at 10 pN *(Figure 3-4C)*. However, unfolding events at higher forces are also observed, resulting in a broad force distribution. The shape and broadness of the force distribution suggests a continuum of two-state unfolding pathways, characterized by different kinetic barriers. Interestingly, the force distribution of the first unfolding event along the three-state unfolding pathway *(Figure 3-4D)* has a shape similar to that of *figure 3-4C*. These similarities suggest that the heterogeneity of the observed unfolding pathways most likely reflects a heterogeneity of the initial states of such pathways, indicating a high structural variability of the *PF* state. After crossing the first barrier for unfolding, the molecule transits directly into the unfolded state (23% of the cases), or is kinetically trapped into an intermediate state that then unfolds at higher forces (71% of the cases). However, the unfolding force distribution of the intermediate state *(Figure 3-4E)* is not scattered and can be well fitted by a two-state unfolding probability function, suggesting a well-defined molecular conformation for the intermediate.

HIV-1-PR monomer refolds through multiple pathways

Refolding of the monomeric HIV-1-PR is also a multi-pathway process, where the molecule regains its original extension either gradually, or through one or more transitions *(Figure 3-5)*. Notice that the refolding events displayed in this figure do not lead directly to the native state of the protein because, as mentioned above, folding into *N* is a very slow process. Mechanically denatured HIV-1-PR monomers compact into partially folded (*PF*) conformations that take at least 50 s to transit into the native state [12]. This *PF-N* transition is not detected in our measurements, implying that the change in extension involved in the associated folding process is too small to give rise to a noticeable signal in our recordings. These results indicate that the *PF* state is similar in compactness to the *N* state. The shape of the force distributions of *figure 3-5D, 5E and 5F* reveal the presence of multiple refolding pathways associated to the refolding events shown in *figure 3-5B and 5C*. No correlation has been noticed between type of refolding trajectory and type of unfolding events observed in the following stretching trace.

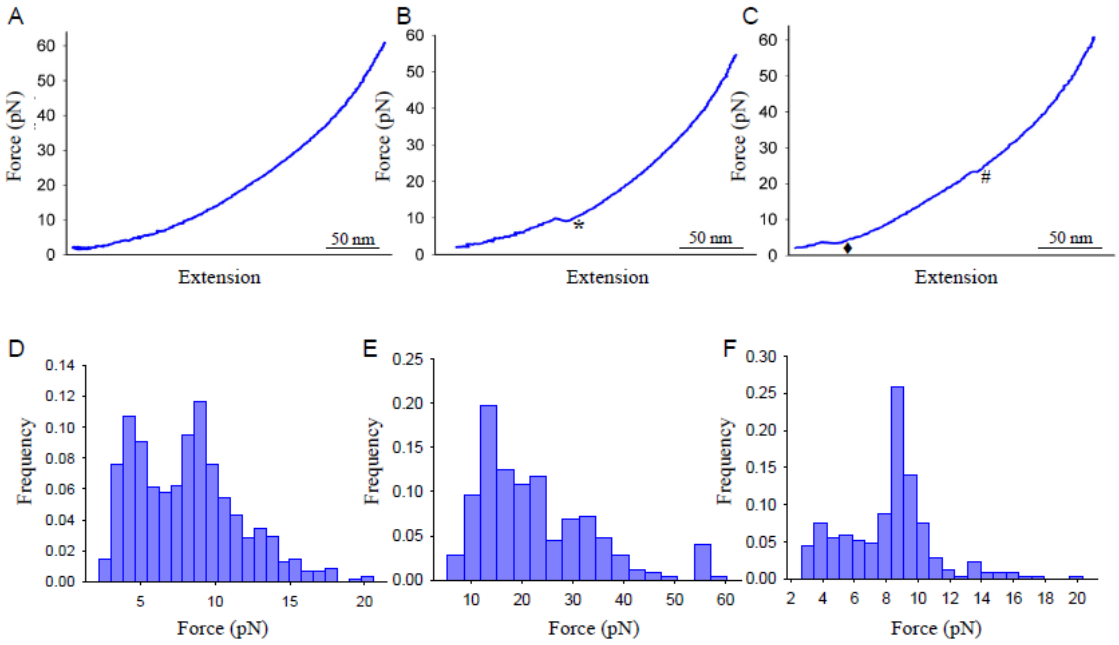


Figure 3-5 HIV-1-PR monomer refolds through multiple pathways. When the tension applied on a mechanically denatured HIV-1-PR monomer is relaxed, the molecule regains its original extension through either: a gradual compaction without any sharp transition (55% of the cases, panel A); a single sharp transition associated to a ΔL_c of 31 ± 4 nm (28% of the cases, panel B), or a two-step process with a first (small, ΔL_c = 8 ± 3 nm) and second (large, ΔL_c = 26 ± 5 nm) transition (12% of the cases, panel C)). Only rarely, it regains a compact conformation through different trajectories (5% of the cases; *figure 3-S1*). D), E) and F) show the force distributions of the refolding events indicated by *, # and ♦, respectively.

Molecular Dynamics Simulations

State-of-the-art ratcheted MD simulations of the mechanical denaturation of the HIV-1-PR monomer were carried out to gain insight into the molecular states visited by the protein during forced unfolding [43-45]. Starting from an equilibrated native-like conformation with a reference distance between Cys5 and Cys98 of 2.1 nm, and using a biasing algorithm to make the simulations feasible in a reasonable amount of time *(see section "Materials and methods" for details)* we calculated 7 unfolding trajectories in explicit solvent at a constant force of 20 pN. Snapshots of the molecular conformations visited by the protein during the forced denaturation are displayed in *figure 3-6*. The first elements to lose structure are hairpin β_1 (residues 10-23) and α helix (residues 87-93); then, β strand 83-87 is removed from the 23-34 loop, leaving hairpins β_2 (residues 43-58) and β_3 (residues 59-75) as the only native secondary structure *(see figure 3-2B)*. Only at the end of the simulation, these two β structures are disrupted, in an undefined order.

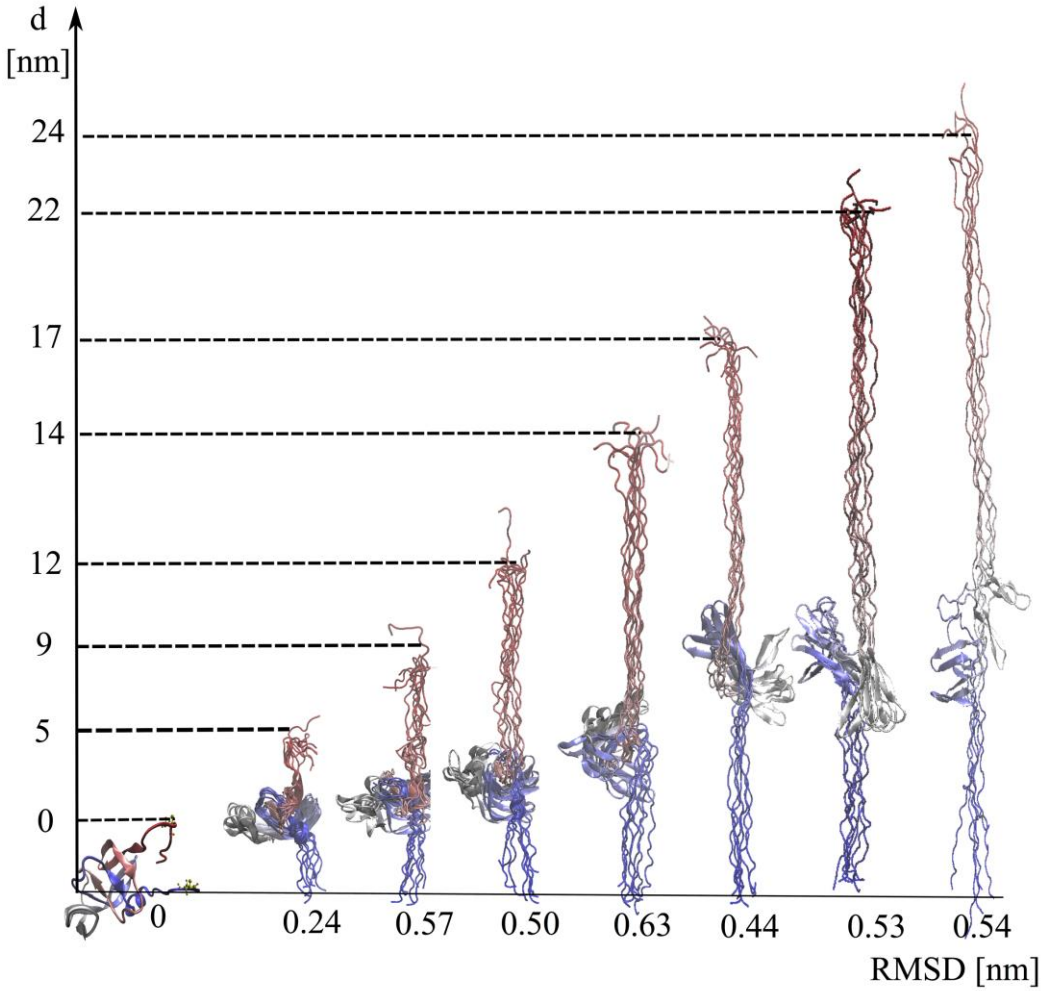


Figure 3-6 Molecular dynamics simulation of the forced unfolding of the HIV-1-PR monomer. Snapshots of the molecular structures visited by the protein along the unfolding trajectories at increasing distance *d* between Cys5 and Cys98. The RMSD values indicate the structural variability within each set.

The seven trajectories mainly differ from each other for number and type of non-native interactions. *Figure 3-7* shows the average number of non-native interactions as a function of the distance *d* between Cys5 and Cys98. Two trajectories display the formation of a limited number of non-native contacts (up to 5, points marked with #), which do not involve well-defined regions of the protein. In the other five trajectories, the number of non-native contacts can reach values between 8 and 10 (points marked with *). The regions of the chain involved in these contacts are more restrained, as shown in panel B of *figure 3-7 (see also Table 3-S1 in the Supplementary data).*

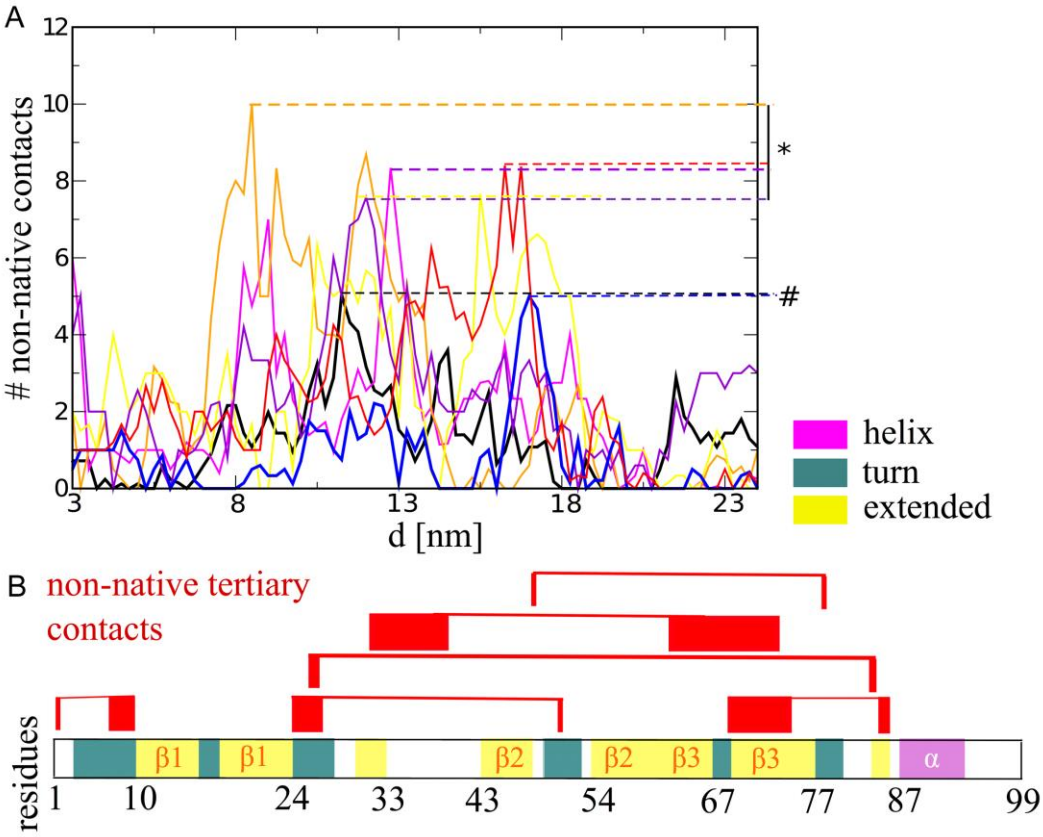


Figure 3-7 Non-native contacts forming during the forced unfolding of the HIV-1-PR monomer. A) Average number of non-native contacts as a function of the distance *d* between Cys5 and Cys98. Each color marks a different unfolding trajectory. The maximum number of non-native contacts is around 5 (#) in two of them, while in the other cases it ranges between 8 and 10 (*). B) The horizontal bar highlights the secondary structures of the protein, while the red connectors indicate non-native contacts forming in at least 2% of the time in a given trajectory.

If we assume that some non-native contacts can stabilize certain secondary and tertiary structures, then the observed heterogeneity along the experimental unfolding pathways *(Figure 3-3)*, in terms of presence or absence of an intermediate state, might reflect the variability in number and type of non-native contacts forming during the forced unfolding process. Moreover, by comparing the end-to-end distance *d* of the conformations displayed in *figure 3-6* with the unfolding transitions (rips) observed during the experiments *(Figure 3-3)*, we can propose that hairpins β_2 and β_3 are structured, at least to some extent, in the intermediate state observed by optical tweezers. In fact, at $d \approx 17$ *nm*, corresponding to the rip width of the *N-I* transition along

the three-state unfolding pathway *(Figure 3-3B)*, β_2 and β_3 are still partly formed, and surrounded by the structure-less part of the chain, with which they make a variable number of tertiary contacts.

Biased MD simulations were also carried out to gain insight into the mechanism by which denatured HIV-1-PR monomers travel toward the native state. Four refolding trajectories were simulated in explicit solvent at a constant force of 5 pN, starting from fully-elongated structures *(right-most conformation in figure 3-6)*. None of the simulations reached the native state of the protein. Rather, the molecules got trapped into partially folded conformations, characterized by native secondary and tertiary contacts in hairpins β_2 and β_3, and non-native contacts between segments 24-34 and 83-93 *(Figure 3-8)*. The overall structures of these conformations are rather heterogeneous, displaying a mutual RMSD of 0.47 nm. However, the distance between Cys5 and Cys98 is 1.3 ± 0.4 nm, which is essentially indistinguishable from that measured in the native state (0.6 ± 0.4 nm). Therefore, the trapped conformations displayed in *figure 3-8* appear to share common features with the *PF* conformations observed in single-molecule experiments, both for their native-like extension and for their structural heterogeneity.

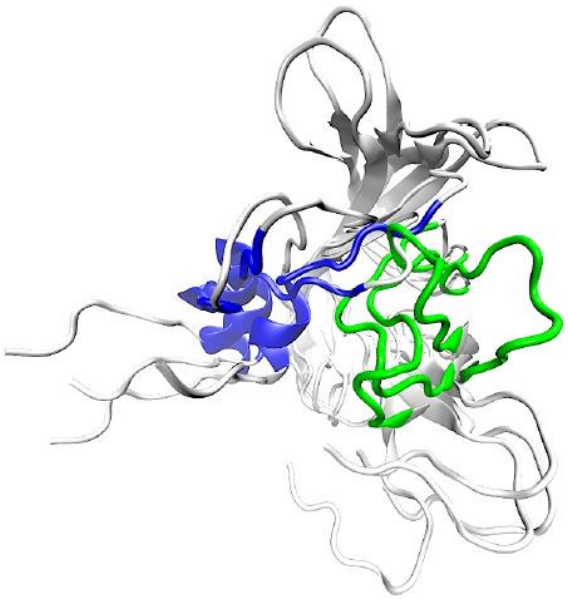


Figure 3-8 Molecular dynamics simulation of the refolding process of the HIV-1-PR monomer. Four superimposed structures adopted by the protein at the end of refolding simulations carried out at 5 pN. The trajectories ended with non-native conformations in which segments 24-34 (green) and 83-93 (blue) cannot build the native β sheet.

Discussion

The monomer of the HIV-1-PR is a complex protein, with a non-trivial β-rich native conformation and a particularly lengthy dynamics. This complexity is not always apparent in bulk experiments, in which the overall experimental signal is the average over a very large number of molecules. In fact, the fluorescence signal recorded in the equilibrium titration experiment shown in *figure 3-1* is compatible with a simple two-state model. Single-molecule force spectroscopy offers a powerful approach for deciphering the intricacies of complex folding scenarios. Through mechanical manipulation it is nowadays possible to follow the actual dynamics of single molecules as they undergo their transformations, avoiding the inherent averaging of ensemble methods. In this paper, we use optical tweezers and biased MD simulations to uncover information inaccessible to more traditional experimental techniques.

The molecular states and unfolding/refolding pathways of HIV-1-PR revealed by our studies are schematically depicted in *figure 3-9*. Starting from the native structure, the protein can either unfold directly, or through an intermediate state. Analysis of the force distributions of the unfolding events *(Figure 3-3)* suggests that the first kinetic barrier of the three-state pathway is probably different from that of the two-state pathway, both because of its position along the reaction coordinate ($x^{\ddagger}_u$=0.5 ± 0.1 nm *versus* $x^{\ddagger}_u$ =1 ± 0.1 nm, respectively) and its height (k^0_u = 4 (±1) x 10^{-2} s^{-1} *versus* k^0_u = 7 (±1) x 10^{-3} s^{-1}, respectively). These data support the idea that the bifurcation of the energy landscape of HIV-1-PR's native state does not take place after a major kinetic barrier, common to both pathways, as reported for other proteins [46]. Rather it takes place early on during the unfolding process [20]. The structure of the intermediate state populated by the protein is not known. However, the comparison of the rip width (~ 17 nm) of the *N-I* transition of the three-state unfolding process *(Figure 3-3B)* with the extension of the molecular conformations visited by the protein during the MD simulations *(Figure 3-6)*, seems to indicate that hairpins β_2 and β_3 are still at least partially structured in the intermediate state.

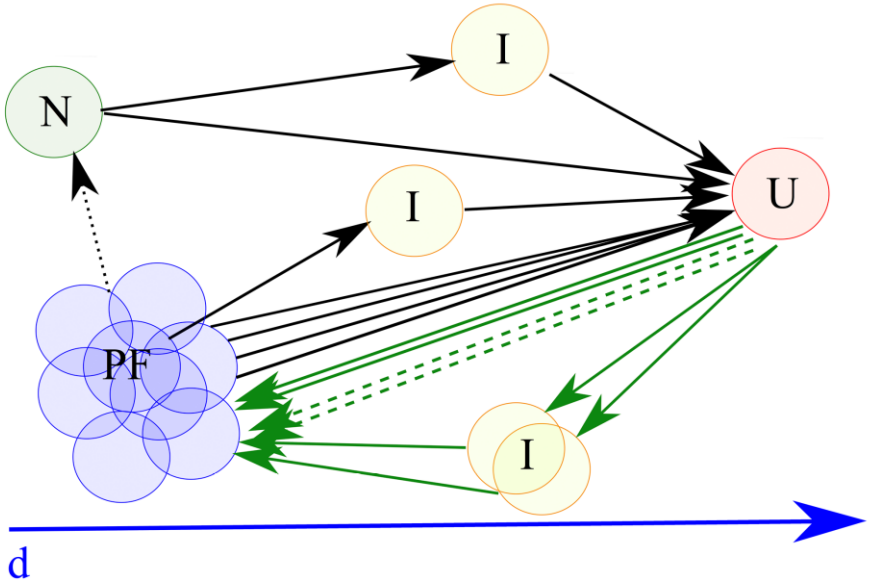


Figure 3-9 Sketch of the unfolding/refolding pathways and molecular states populated by the HIV-1-PR monomer. The native state unfolds along two main pathways, one involving an intermediate state. The unfolded state follows multiple trajectories to refold first into a structurally heterogeneous *PF* state, which then transits into the native state slowly. The *PF* state can unfold either in a two-state manner or, more often, by populating an intermediate state.

The heterogeneity of the unfolding pathways of the HIV-1-PR monomer has been previously proposed by Bonomi *et al.* [14], from calculations of the free-energy surface of the protein with a structure-based model. However, no specific intermediate state was observed. These results agree well with our data if we consider that, according to our MD simulations, the intermediate state could be stabilized by non-native contacts, and thus could not be detectable in structure-based models which, by definition, do not describe non-native interactions. It should also be noticed, however, that the pathways described in this paper apply only to the folding behavior of the molecule under tension. The folding pathways followed by HIV-1-PR free in solution might differ.

Refolding of the mechanically-denatured HIV-1-PR monomer into its native state is a multi-path process characterized by partially folded (*PF*) conformations that altogether act as a long-lived state controlling the slow kinetic phase of the process. In approximately half of the folding trajectories, the *PF* state is reached without encountering any major free-energy barrier *(Figure 3-5A)*. As a consequence, the overall folding kinetics is presumably solely controlled by

the large free-energy barrier between the *PF* and the native state. In the other half of the cases, the HIV-1-PR monomer folds towards the *PF* state by crossing different kinetic barriers, thus generating broad refolding force distributions *(Figure 3-5)*. The *PF* state is quite compact and its transition into the native state is undetectable in our measurements. Nonetheless, despite this experimental limitation, we were able to gain insight into this long-lived state by pulling on it and generating unfolding events *(Figure 3-4)*. The resulting broad and skewed distribution of unfolding forces suggests that the *PF* state comprises a structurally heterogeneous set of conformations that unravel through a multitude of unfolding pathways. All MD simulations of the refolding process lead to trapped non-native states that might share common features with the *PF* conformations. They are characterized by a native-like arrangement of hairpins β_2 and β_3, but a non-native positioning of the hydrophobic core of the protein. In the native conformation, the segment 83-86 is inserted in the looped segments 21-24 and 31-34 to form a β-sheet, which is further stabilized by the β-strand 74-78. In the conformations displayed in Figure 8 not only this β-sheet is not present, but also the relative positioning of segment 83-93 with respect to segment 24-34 is such that the molecule has to swell to reach a native-like conformation.

Despite the heterogeneity of the *PF* conformations, in ~ 70% of the unfolding trajectories the molecule populates an intermediate state that unravels producing a well-defined force distribution. This intermediate state appears to have similar properties to those of the intermediate state observed when pulling on the native state *(Figure 3-3)*. The height (k^0_u) and position ($x^{\ddagger}_u$) of the unfolding energy barriers are similar, and the changes in contour length associated with their unfolding are within the standard deviations. Likewise, their mechanical resistances are similar, as shown in *figure 3-3E and 4E*. These similarities might reflect common structural features of the two intermediate states, where hairpins β_2 and β_3 could be, at least partially, folded in both cases. However, our results do not allow for definitive conclusions on this matter and extracting detailed structural features of the two intermediate states will require further dedicated experiments.

Being a key enzyme in the maturation of the human immunodeficiency virus, the HIV-1-PR is the target of drugs used in anti-AIDS therapy. In essentially all these therapeutic approaches, molecules are designed to target the active site of the protease to inhibit its function. However, being associated with the error-prone replication mechanism of a retrovirus, the protease mutates under drug pressure, altering the binding site of the inhibitors and generating, in the process, pharmacological resistance. The intermediate states revealed in our study can be regarded as non-

conventional targets for molecules that, by stabilizing partially folded conformations, can block the maturation of HIV-1-PR, thereby inhibiting production of mature viral particles. Much work remains to be done to implement such a strategy. In particular, a systematic NMR study of the denatured state of the HIV-1-PR by itself, and in presence of selected exogenous molecules might be very helpful in this regard. Within this context, and in keeping with the fact that folding inhibition can be viewed as an extreme form for allosteric inhibition, one can refer to Kunze *et al.* [47] where it is reported that structure-based computational screening of dynamic pockets of the HIV-1-PR has allowed the identification of small molecules displaying robust (IC50 ~ 10-20 µM) inhibition.

Materials and methods

Protein Expression

Plasmid pJF19, a pXC35 derivative encoding the Q7K/C67A/C95A variant of HIV1-1-PR was a kind gift of Dr. Celia Schiffer (University of Massachusetts Medical School, Waltham, USA). Starting from this plasmid, a construct expressing the double cysteine, monomeric L5C/R87K/N98C protein variant used in single molecule experiments was generated with the QuikChange Multi Site-Directed Mutagenesis Kit (Stratagene) in conjunction with the following oligonucleotides:

oligo1: 5'-GCCGCAGATTACCTGCTGGAAACGCCCGC-3',

oligo2: 5'-CCGGTGAACATTATTGGCAAAAACCTGCTGACCCAG-3',

oligo3: 5'-GGCGCGACCCTGTGCTTTTAGGGATCC-3',

Where the codons introducing the L5C (oligo1), R87K (oligo2) and N98C (oligo3) amino acyl substitutions are underlined. The mutagenic reaction was performed according to manufacturer's instructions, while all other required DNA handling steps were carried out according to established protocols [48, 49]. Both DNA strands of the insert of the resulting plasmid were sequenced to confirm the presence of the desired mutations and to rule out the introduction of unwanted ones. The engineered protein was overproduced using the *Escherichia coli* TAP106 strain as the expression host, under the growth and induction conditions reported elsewhere [12, 28].

Protein purification

The recombinant protein was isolated from inclusion bodies under denaturing conditions, and stored frozen in dried form until needed. Cells obtained from a 5 L culture were resuspended in 20 mM Tris-HCl, pH 7.5, containing 1 mM EDTA, 1 mM 2-mercaptoethanol (buffer A) and 0.2 mg/ml mg/ml lysozyme, and disrupted by applying five cycles of 1 min sonication (Branson, Model 250 sonicator) on ice. After centrifugation of the homogenate at 20000xg for 20 min at 20 ºC, the recombinant protein was purified from inclusion bodies according to the protocol described in [50] Briefly, inclusion bodies were washed once with 2 M urea, 0.2% Triton X-100 in buffer A, once with 2 M urea in buffer A, and twice with buffer A. For each washing, 5 ml of buffer per gram of the initial cell pellet was used and the resuspended sample was centrifuged at 20000xg for 30 min at 20 ºC. The pellet was dissolved in glacial acetic acid (60 mg/ml) and then diluted with 50% acetic acid to a concentration of 12 mg/ml. After centrifugation at 45000xg for 30 min, the supernatant was passed through a 0.22 μm filter and chromatographed on a Sephadex G-75 Superfine column (2.6 cm x 85 cm, 450 ml) equilibrated and eluted with 50% acetic acid. Fractions containing homogenous protein, as judged by SDS-PAGE, were pooled, brought to dryness under vacuum and stored at -80 ºC. The resulting protein was monomeric as judged by dynamic light scattering measurements and analytical gel filtration.

Protein preparation for single molecule experiments

DNA handles were prepared essentially as described by [28, 32]. The following oligonucleotides were obtained from Primm srl (Milan, Italy) and they were used to amplify a 558 bp portion of the pGEMEX-1 vector.

oligo4: 5’Biotin-CAAAAAACCCCTCAAGACCC-3'

oligo5: 5’Digoxigenin-CAAAAAACCCCTCAAGACCC-3'

oligo6: 5’Thiol-GCTACCGTAATTGAGACCAC-3'

Handles (400–500 μg) were generated by PCR using ten 1.2 ml reaction mixtures. The two types of handles were generated using oligo6 with either oligo4 or oligo5. The reaction mixture contained 0.5 μM oligo4 or oligo5, 0.5 μM oligo6, 0.04 ng/μl p-GEMEX, 0.025 U/μl *Taq*-polymerase in PCR Buffer (Qiagen) 16.2 mM DTT, 1.5 mM $MgCl_2$. Amplification was performed

by applying 34 cycles consisting of denaturation at 94 ºC for 45 s, annealing at 67 ºC for 40 s for oligo4 or annealing at 63ºC for 40 s for oligo5, and extension at 72ºC for 90 s.

After purification with a HiSpeed Plasmid Maxi Kit (Qiagen), the DNA handles were attached to the protein by following the procedure described in [28, 32] with. In the light of the observed tendency of the protease to aggregate at concentrations above 1 mg/ml, activation of cysteine residues was carried out on the denatured protein, which was refolded before the attachment of the DNA handles.

The lyophilized purified protease (1 mg) was resuspended in 20 mM Na-phosphate, pH 6.5, containing 1 mM DTT and 7 M guanidine-HCl (200 µl). After 30 min incubation at 25 °C the sample was diluted to 3 mg/ml protein concentration (270 µM) with 20 mM Na-phosphate, pH 6.5, 1 mM DTT to lower the guanidine-HCl concentration to 4 M. DTT was removed by centrifugal chromatography on Bio Spin P6 columns (Bio-Rad) equilibrated and eluted with 20 mM Na-phosphate, pH 6.5, 4 M guanidine-HCl. The resulting protein (typically 80–100 µM) was reacted with 2- to 5-fold molar excess of dithiodipyridine (DTDP) added from a 10 mM stock solution in 20 mM Na-phosphate, pH 6.5, 4 M guanidine-HCl, 15% (v/v) acetonitrile. The reaction mix was incubated at room temperature for 30 min, which were sufficient to bring about the complete derivatization of the protease cysteine residues, as judged spectrophotometrically by monitoring absorbance changes at 343 nm [28, 32].

The sample was then diluted to a protein concentration of 0.33 mg/ml with 20 mM Na-phosphate, pH 6.5, 4 M guanidine-HCl. Excess DTDP was removed by dialysis at 4 °C against 0.1 M formic acid (two 1 l changes for 4 h each) and 30 mM formic acid, pH 2.8 (1 l, overnight). After centrifugation at 19000xg for 30 min at 4 °C to remove precipitated material, protein refolding was started by sample dilution in 5 volumes of 10 mM Na-acetate, pH 6.0, to shift the pH to 4.0, and dialysis for 4 h at 4 °C against 4 l of 20 mM Na-phosphate buffer, pH 6.0, which was made partially anaerobic by bubbling nitrogen gas. After concentration in centrifugal filter units (Millipore) followed by centrifugation at 16000xg for 30 min at 4 ºC, the protein (0.6-0.8 mg/ml) was stored in aliquots at -80ºC.

Attachment of the DNA handles to the protein was performed as described in [28] using an optimized DNA/protein ratio of 1:1 (mol/mol) instead of the 5:1 ratio suggested in Cecconi *et al.* [28]. Attachment of the DNA handles to the protein was verified by SDS-PAGE on 6% gels

prepared in TBE buffer. The coupling of protein−DNA chimera to polystyrene beads was performed as described in *Cecconi et al.* [28].

Denaturation/renaturation studies

Protein was refolded as described above by skipping the cysteine activation step. Protein folding and unfolding was studied by monitoring fluorescence emission changes of protease samples (2 µM) in 20 mM Na-phosphate buffer, pH 6.0, 1 mM DTT and decreasing or increasing urea concentration at 25°C with a Cary Eclipse fluorimeter (Varian; excitation light, 290 nm). At least 5 spectra were recorded after reaching equilibrium. The maximum wavelength of the emission spectrum was plotted as a function of urea concentration and the data were fitted to the two-state process equation:

$$\lambda_{\max obs} = \lambda_{\max F} + (\lambda_{\max U} - \lambda_{\max F}) * \left(\frac{\exp\left(\frac{-m * x + \Delta G^0}{RT}\right)}{1 + \left(\exp\left(\frac{-m * x + \Delta G^0}{RT}\right)\right)}\right) \qquad (3-1)$$

where ΔG^o is the unfolding free energy, x is the denaturant concentration, R is the gas constant (8.314 J/(K*mol)), T is the absolute temperature (283K), m is a measure of the dependence of ΔG on the denaturant concentration, $\lambda_{max\ F}$ is the maximum wavelength of the native emission spectrum and $\lambda_{max\ U}$ is the maximum wavelength of the unfolded emission spectrum. This analysis yielded a ΔG^o=9.84±1.34 kJ/mol.

Optical tweezers experiments

Single molecule manipulation experiments were performed using a custom-built optical tweezers instrument with a dual-beam laser trap of 840 nm wavelength [25]. The experiments were conducted at room temperatures in 10 mM Tris-HCl buffer, pH 7.0, 250 mM NaCl, 10 mM $CaCl_2$ or in 20 mM Na-phosphate buffer, pH 6.0, with similar results. DNA-protein constructs were manipulated between a 3.10 µm antidigoxigenin-coated bead (Spherotec) held in the optical trap, and a 2.2 µm streptavidin-coated bead (Spherotec) held at the end of a micropipette by suction. The force applied to the molecule was varied by moving the micropipette relative to the optical trap by means of a piezoelectric flexure stage (MAX311/M, Thorlabs, Newton, NJ). The applied force was determined by measuring the change in light momentum of the laser beams leaving the trap [31], while changes in the extension of the molecule were determined by video microscopy [31]. The molecules were stretched and relaxed at a constant speed of 100 nm s^{-1}, while force and

molecular extension were recorded at a rate of 40 Hz. Only those molecules that displayed the characteristic DNA overstretching transition at 67 pN were used in the analysis [19].

Changes in Contour Length of HIV-1 PR

Theoretical changes in contour lengths (ΔLc) upon unfolding/refolding of HIV-1-PR were calculated as described [19]. The distance between position 5 and 98 in the native structure was measured using the NMR structure of HIV-1-PR (Protein Data Bank code 1BVG).

Analysis of unfolding force distributions

Unfolding force distributions of two-state unfolding events were fit to the probability distribution function [38]:

$$P(F) = \left(\frac{{k^0}_u}{r}\right) \exp(F{x^\ddagger}_u/k_BT) \exp[-(({k^0}_u\, k_BT)/(r{x^\ddagger}_u))(\exp\left(\frac{F{x^\ddagger}_u}{k_BT}\right) - 1)] \qquad (3-2)$$

where ${k^0}_u$ is the unfolding rate at zero force, r is the loading rate (pN s^{-1}), ${x^\ddagger}_u$ is the distance from the native state to the transition state along the reaction coordinate, k_B is the Boltzmann's constant and T is the absolute temperature.

The same force distributions were also fit to the equation [25] [37]:

$$\ln\left[r ln\left[\frac{1}{N(F,r)}\right]\right] = \ln\left(\frac{{k^0}_u k_B T}{{x^\ddagger}_u}\right) + ({x^\ddagger}_u/k_BT)F \qquad (3-3)$$

where *N(F,r)* is the fraction of folded molecules at force F and loading rate r (pN s^{-1}). *N(F,r)* can be calculated by integrating the unfolding force distributions over the corresponding range of forces. *Equation 3-3* can be used to fit *ln[rln[1/N(F,r)]]* vs. F graphs.

Simulations

Unfolding simulations were carried out with the Amber03 force field [48], using a modified version of the GROMACS package [51, 52] . The system is solvated with 35000 water molecules in a fixed box of sides 6×6×30 nm. The temperature is maintained at 300 K by Nosè-Hoover thermostat [53].

First, we carried out a 100 ns MD simulation at constant forces of 8 pN, 16 pN, 22 pN and 35 pN. This time is much shorter than that needed for the system to undergo any consistent

conformational change. Thus, the simulations provided reference distances *d* between the Cys5 and Cys98 positions. According to these simulations *(Figure 3-S2)*, the application of a force on the two cysteine residues up to 22 pN sets the protein in a stationary initial conformation characterized by varying average values of *d*, ranging from 0.2 nm (calculated with respect to the native conformation) at 8 pN to 2.1 nm at 22 pN. Application of a force of 35 pN does not increase the average value of *d* beyond 2.1 nm. Such increase of *d* is due to the fact that the termini of the chain, which eventually, in presence of another monomer, build the dimeric interface of HIV-1-PR, are weakly stabilized in the monomer, and consequently can be unraveled even by the application of small forces *(Figure 3-S3)*. Since the transitions observed in the optical tweezer experiments occur typically at forces of ≈20 pN or higher, we used 2.1 nm as the reference distance between the cysteine residues in the native state.

The biasing scheme is that described in ref [25]. However, the ratchet is here applied both on the distance *d* between Cys5 and Cys98 and on the RMSD to a fully-unfolded, linear conformation of the protein. The biasing potential has thus the form

$$U_{rat}(t) = \frac{1}{2}k_d\left(\rho_d(t) - min_{t'<t}\,\rho_d(t')\right)^2 + \frac{1}{2}k_r\left(\rho_r(t) - min_{t'<t}\,\rho_r(t')\right)^2 \qquad (3-4)$$

where

$$\rho_d(t) = (d(t) - 30nm)^2$$

and

$$\rho_r(t) = [RMSD(t)]^2,$$

the harmonic constants being $k_d = k_r = 1$ kJ/mol/nm^2.

A contact between residues is defined if any pair of their atoms is closer than 0.35 nm. Non-native contacts are calculated as contacts between residues which are separated by at least 3 other residues along the chain, and which are at least 2 residues away from pairs of residues which form native contacts. They are considered relevant if they appear in at least 2% of the time in a given trajectory.

Conclusions

For the first time we study the folding mechanism of HIV-1-protease monomer at the single-molecule level, revealing information inaccessible to more traditional ensemble techniques. Our

results not only shed light on a complex protein folding network, improving our understanding of the basic rules governing the folding of a polypeptide chain into a functional three-dimensional structure. They also provide information which may be used at profit from an applicative point of view, as the intermediate states revealed by our measurements can be regarded as non-conventional targets for drugs that, by stabilizing partially folded conformations, may eventually block productive folding of HIV-1-PR.

References

[1] J.R. Huff, HIV protease: a novel chemotherapeutic target for AIDS, Journal of Medicinal Chemistry, 34 (1991) 2305-2314.

[2] J.M. Louis, A. Aniana, I.T. Weber, J.M. Sayer, Inhibition of autoprocessing of natural variants and multidrug resistant mutant precursors of HIV-1 protease by clinical inhibitors, Proceedings of the National Academy of Sciences, 108 (2011) 9072-9077.

[3] J.M. Louis, E.M. Wondrak, A.R. Kimmel, P.T. Wingfield, N.T. Nashed, Proteolytic Processing of HIV-1 Protease Precursor, Kinetics and Mechanism, Journal of Biological Chemistry, 274 (1999) 23437-23442.

[4] R. Ishima, D.A. Torchia, J.M. Louis, Mutational and Structural Studies Aimed at Characterizing the Monomer of HIV-1 Protease and Its Precursor, Journal of Biological Chemistry, 282 (2007) 17190-17199.

[5] Y. Levy, A. Caflisch, J.N. Onuchic, P.G. Wolynes, The folding and dimerization of HIV-1 protease: evidence for a stable monomer from simulations, J Mol Biol, 340 (2004) 67-79.

[6] T. Yamazaki, A.P. Hinck, Y.-X. Wang, L.K. Nicholson, D.A. Torchia, P. Wingfield, S.J. Stahl, J.D. Kaufman, C.-H. Chang, P.J. Domaille, P.Y.S. Lam, Three-dimensional solution structure of the HIV-1 protease complexed with DMP323, a novel cyclic urea-type inhibitor, determined by nuclear magnetic resonance spectroscopy, Protein Science, 5 (1996) 495-506.

[7] A.G. Tomasselli, R.L. Heinrikson, Targeting the HIV-protease in AIDS therapy: a current clinical perspective, Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology, 1477 (2000) 189-214.

[8] S. Kimura, M. Caldarini, R.A. Broglia, N.V. Dokholyan, G. Tiana, The maturation of HIV-1 protease precursor studied by discrete molecular dynamics, Proteins: Structure, Function, and Bioinformatics, 82 (2014) 633-639.

[9] R.A. Broglia, Y. Levy, G. Tiana, HIV-1 protease folding and the design of drugs which do not create resistance, Current Opinion in Structural Biology, 18 (2008) 60-66.

[10] R.A. Broglia, G. Tiana, L. Sutto, D. Provasi, F. Simona, Design of HIV-1-PR inhibitors that do not create resistance: Blocking the folding of single monomers, Protein Science, 14 (2005) 2668-2681.

[11] R. Zutshi, J. Franciskovich, M. Shultz, B. Schweitzer, P. Bishop, M. Wilson, J. Chmielewski, Targeting the Dimerization Interface of HIV-1 Protease: Inhibition with Cross-Linked Interfacial Peptides, Journal of the American Chemical Society, 119 (1997) 4841-4845.
[12] A.F. Noel, O. Bilsel, A. Kundu, Y. Wu, J.A. Zitzewitz, C.R. Matthews, The Folding Free-Energy Surface of HIV-1 Protease: Insights into the Thermodynamic Basis for Resistance to Inhibitors, Journal of Molecular Biology, 387 (2009) 1002-1016.
[13] K.W. Plaxco, K.T. Simons, D. Baker, Contact order, transition state placement and the refolding rates of single domain proteins, Journal of Molecular Biology, 277 (1998) 985-994.
[14] M. Bonomi, A. Barducci, F.L. Gervasio, M. Parrinello, Multiple Routes and Milestones in the Folding of HIV–1 Protease Monomer, PLoS ONE, 5 (2010) e13208.
[15] M.J. Todd, N. Semo, E. Freire, The structural stability of the HIV-1 protease, Journal of Molecular Biology, 283 (1998) 475-488.
[16] S.C. Panchal, N.S. Bhavesh, R.V. Hosur, Real time NMR monitoring of local unfolding of HIV-1 protease tethered dimer driven by autolysis, FEBS letters, 497 (2001) 59-64.
[17] S.K. Grant, I.C. Deckman, J.S. Culp, M.D. Minnich, I.S. Brooks, P. Hensley, C. Debouck, T.D. Meek, Use of protein unfolding studies to determine the conformational and dimeric stabilities of HIV-1 and SIV proteases, Biochemistry, 31 (1992) 9491-9501.
[18] D. Xie, S. Gulnik, L. Collins, E. Gustchina, L. Suvorov, J.W. Erickson, Dissection of the pH Dependence of Inhibitor Binding Energetics for an Aspartic Protease: Direct Measurement of the Protonation States of the Catalytic Aspartic Acid Residues†,‡, Biochemistry, 36 (1997) 16166-16172.
[19] C. Cecconi, E.A. Shank, C. Bustamante, S. Marqusee, Direct Observation of the Three-State Folding of a Single Protein Molecule, Science, 309 (2005) 2057-2060.
[20] E. A.Shank, C.Cecconi, J.W.Dill, S.Marqusee, C.Bustamante, The folding cooperativity of a protein is controlled by its chain topology, Nature, 465 (2010) 637-640.
[21] M. Shayegan, N. Rezaei, N.H. Lam, T. Altindal, A. Wieczorek, N.R. Forde, Probing multiscale mechanics of collagen with optical tweezers, Proc. SPIE, 8810 (2013) 88101P.
[22] N. Varongchayakul, S. Johnson, T. Quabili, J. Cappello, H. Ghandehari, S.D.J. Solares, W. Hwang, J. Seog, Direct Observation of Amyloid Nucleation under Nanomechanical Stretching, ACS Nano, 7 (2013) 7734-7743.

[23] A. Borgia, P.M. Williams, J. Clarke, Single-Molecule Studies of Protein Folding, Annual Review of Biochemistry, 77 (2008) 101-125.
[24] D. Aioanei, S. Lv, I. Tessari, A. Rampioni, L. Bubacco, H. Li, B. Samorì, M. Brucale, Single-Molecule-Level Evidence for the Osmophobic Effect, Angewandte Chemie International Edition, 50 (2011) 4394-4397.
[25] P.O. Heidarsson, I. Valpapuram, C. Camilloni, A. Imparato, G. Tiana, F.M. Poulsen, B.B. Kragelund, C. Cecconi, A highly compliant protein native state with a spontaneous-like mechanical unfolding pathway, Journal of the American Chemical Society, 134 (2012) 17068-17075.
[26] G. Arad-Haase, S.G. Chuartzman, S. Dagan, R. Nevo, M. Kouza, B.K. Mai, H.T. Nguyen, M.S. Li, Z. Reich, Mechanical Unfolding of Acylphosphatase Studied by Single-Molecule Force Spectroscopy and MD Simulations, Biophysical journal, 99 (2010) 238-247.
[27] M. Caldarini, P. Sonar, I. Valpapuram, D. Tavella, C. Volonté, V. Pandini, M.A. Vanoni, A. Aliverti, R.A. Broglia, G. Tiana, C. Cecconi, The complex folding behavior of HIV-1-protease monomer revealed by optical-tweezer single-molecule experiments and molecular dynamics simulations, Biophysical chemistry, 195 (2014) 32-42.
[28] C. Cecconi, E.A. Shank, S. Marqusee, C. Bustamante, DNA Molecular Handles for Single-Molecule Protein-Folding Studies by Optical Tweezers, Methods Mol Biol, The Humana Press Inc (New York, U.S.A.), 749 (2011) 255-271.
[29] R. Ishima, R. Ghirlando, J. Tozser, A.M. Gronenborn, D.A. Torchia, J.M. Louis, Folded monomer of HIV-1 protease, The Journal of biological chemistry, 276 (2001) 49110-49116.
[30] J.M. Louis, R. Ishima, I. Nesheiwat, L.K. Pannell, S.M. Lynch, D.A. Torchia, A.M. Gronenborn, Revisiting Monomeric HIV-1 Protease: Characterization and redesign for improved properties, Journal of Biological Chemistry, 278 (2003) 6085-6092.
[31] S.B. Smith, Y. Cui, C. Bustamante, [7] Optical-trap force transducer that operates by direct measurement of light momentum, in: M. Gerard, P. Ian (Eds.) Methods in Enzymology, Academic Press, 361 (2003) 134-162.
[32] C. Cecconi, E. Shank, F. Dahlquist, S. Marqusee, C. Bustamante, Protein-DNA chimeras for single molecule mechanical folding studies with the optical tweezers, Eur Biophys J, 37 (2008) 729-738.

[33] M. Carrion-Vazquez, A.F. Oberhauser, S.B. Fowler, P.E. Marszalek, S.E. Broedel, J. Clarke, J.M. Fernandez, Mechanical and chemical unfolding of a single protein: A comparison, Proceedings of the National Academy of Sciences, 96 (1999) 3694-3699.

[34] J. Stigler, F. Ziegler, A. Gieseke, J.C.M. Gebhardt, M. Rief, The Complex Folding Network of Single Calmodulin Molecules, Science, 334 (2011) 512-516.

[35] C. Bustamante, J. Marko, E. Siggia, S. Smith, Entropic elasticity of lambda-phage DNA, Science, 265 (1994) 1599-1600.

[36] J.F. Marko, E.D. Siggia, Stretching DNA, Macromolecules, 28 (1995) 8759-8770.

[37] J. Liphardt, B. Onoa, S.B. Smith, I. Tinoco, C. Bustamante, Reversible Unfolding of Single RNA Molecules by Mechanical Force, Science, 292 (2001) 733-737.

[38] E. Evans, K. Ritchie, Dynamic strength of molecular adhesion bonds, Biophysical journal, 72 (1997) 1541-1555.

[39] M. Schlierf, H. Li, J.M. Fernandez, The unfolding kinetics of ubiquitin captured with single-molecule force-clamp techniques, Proceedings of the National Academy of Sciences of the United States of America, 101 (2004) 7299-7304.

[40] H. Dietz, M. Rief, Exploring the energy landscape of GFP by single-molecule mechanical experiments, Proceedings of the National Academy of Sciences of the United States of America, 101 (2004) 16192-16197.

[41] H.C. Kotamarthi, R. Sharma, Sri R. Koti Ainavarapu, Single-Molecule Studies on PolySUMO Proteins Reveal Their Mechanical Flexibility, Biophysical journal, 104 (2013) 2273-2281.

[42] G. Yang, C. Cecconi, W.A. Baase, I.R. Vetter, W.A. Breyer, J.A. Haack, B.W. Matthews, F.W. Dahlquist, C. Bustamante, Solid-state synthesis and mechanical unfolding of polymers of T4 lysozyme, Proceedings of the National Academy of Sciences, 97 (2000) 139-144.

[43] M. Marchi, P. Ballone, Adiabatic bias molecular dynamics: A method to navigate the conformational space of complex molecular systems, The Journal of Chemical Physics, 110 (1999) 3697-3702.

[44] E. Paci, M. Karplus, Forced unfolding of fibronectin type 3 modules: an analysis by biased molecular dynamics simulations, Journal of Molecular Biology, 288 (1999) 441-459.

[45] C. Camilloni, R.A. Broglia, G. Tiana, Hierarchy of folding and unfolding events of protein G, CI2, and ACBP from explicit-solvent simulations, The Journal of Chemical Physics, 134 (2011) -.

[46] H. Dietz, F. Berkemeier, M. Bertz, M. Rief, Anisotropic deformation response of single protein molecules, Proceedings of the National Academy of Sciences, 103 (2006) 12724-12728.
[47] J. Kunze, N. Todoroff, P. Schneider, T. Rodrigues, T. Geppert, F. Reisen, H. Schreuder, J. Saas, G. Hessler, K.-H. Baringhaus, G. Schneider, Targeting Dynamic Pockets of HIV-1 Protease by Structure-Based Computational Screening for Allosteric Inhibitors, Journal of Chemical Information and Modeling, 54 (2014) 987-991.
[48] F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, K. Sthrul, Current Protocols in Molecular Biology, John Wiley & Sons. Inc. Hoboken NJ, (2005).
[49] T. Maniatis, E.F. Frisch, J. Sambrook, Molecular Cloning, a laboratory manual. Ed. , Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA., (1988).
[50] J. Hui, A. Tomasselli, I. Reardon, J. Lull, D. Brunner, C.-S. Tomich, R. Heinrikson, Large scale purification and refolding of HIV-1 protease fromEscherichia coli inclusion bodies, J Protein Chem, 12 (1993) 323-327.
[51] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, H.J.C. Berendsen, GROMACS: Fast, flexible, and free, Journal of Computational Chemistry, 26 (2005) 1701-1718.
[52] C. Camilloni, A. Guerini Rocco, I. Eberini, E. Gianazza, R.A. Broglia, G. Tiana, Urea and Guanidinium Chloride Denature Protein L in Different Ways in Molecular Dynamics Simulations, Biophysical journal, 94 (2008) 4654-4661.
[53] S. Nosé, A unified formulation of the constant temperature molecular dynamics methods, The Journal of Chemical Physics, 81 (1984) 511-519.

Supplementary Data

The complex folding behavior of HIV-1-protease monomer revealed by optical-tweezer single-molecule experiments and molecular-dynamics simulations

Supplementary data

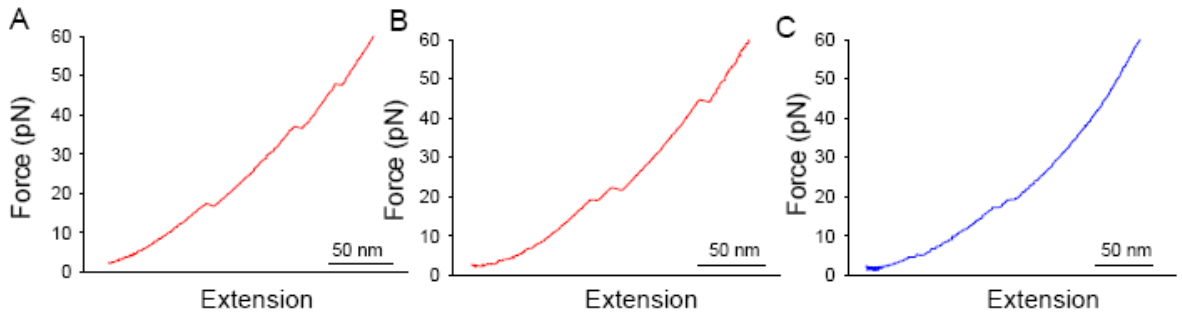


Figure 3-S1 Rare unfolding and refolding trajectories for HIV-1-PR. Occasionally (5% of the cases) the native state unfolds along trajectories that do not resemble those shown in *figure 3-3* (main text). Sometimes it populates multiple intermediate states, as shown in panel A). Sometimes it unfolds only partially through an event resembling the *N-I* transition of *figure 3-3B*. Sometimes it unfolds in a three-state manner with a small transition followed by a larger one. Similarly, the partially folded (*PF*) state occasionally (6% of the cases) follows rare unfolding trajectories, like the one shown in panel B). A rare relaxation trace characterized by multiple refolding events is shown in panel C).

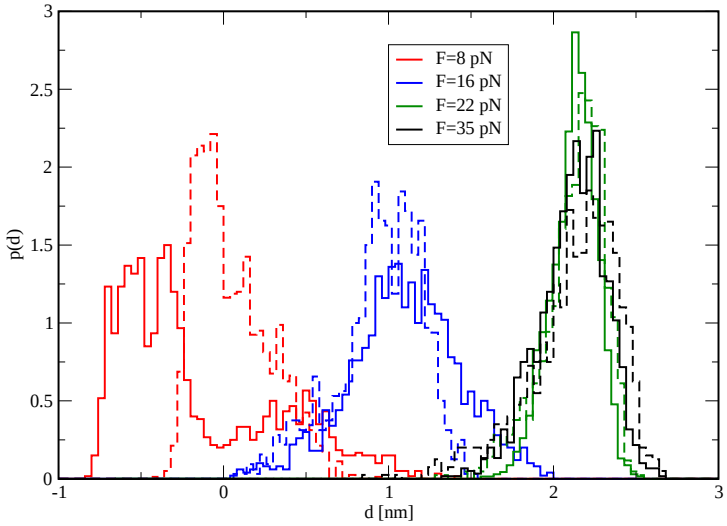


Figure 3-S2 Histograms of the distance *d* between Cys5 and Cys98 obtained from 100 ns plain MD simulations at 300 K at different forces. Dashed curves indicate histograms obtained from the second third of the simulation, while solid curves indicate histograms obtained from the last

third of the simulation. A similar position for a pair of corresponding histograms indicates that they represent a stationary state.

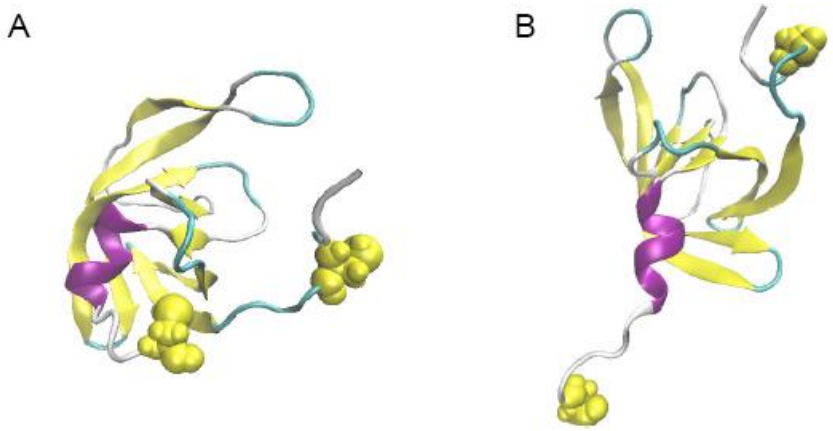


Figure 3-S3 Final conformations of HIV-1-PR after 100 ns MD simulation with an applied force of 8 pN A) and 35 pN B).

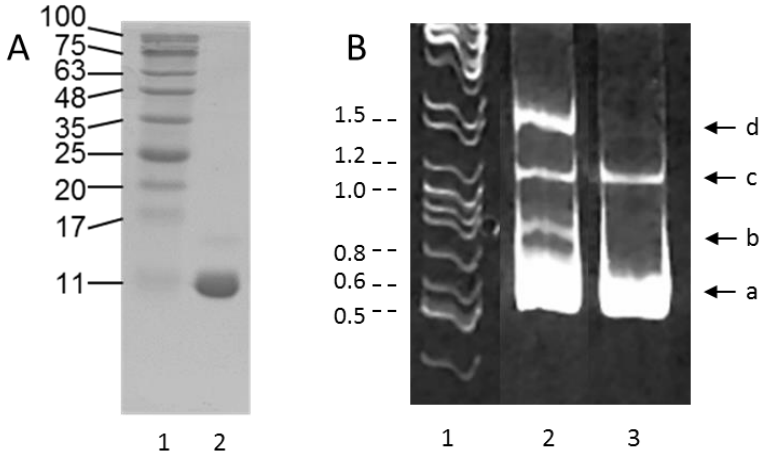


Figure 3-S4 DNA-protein chimera preparation. A) SDS-PAGE (16 % (w/v) acrylamide/bisacrylamide) of a sample of the purified L5C-R87K-N98C HIV1 PR (lane 2); the mass of protein standards (in kDa, lane 1) is indicated. The gel was stained with Coomassie blue. B) SDS-PAGE (6% (w/v) acrylamide/bisacrylamide), in TBE buffer [1, 2] of the product of the DNA-protein coupling reaction (lane 2). An aliquot of the control reaction from which the protein was omitted was loaded on lane 3. Single handles (558 bp DNA molecules) run at position *a*, while two handles attached to each other via a disulfide bond between the thiol groups

present at one end of each molecule run at position *c* [2]. After reaction of the DNA handles with activated HIV-1-PR, two new bands appear in the gel (lane 2). The band at position *b* corresponds to HIV-1-PR bound to one handle. The other band at position *d* corresponds to HIV-1-PR bound to two handles. No additional bands at higher mass were detected, indicating no reaction of the handles with covalent oligomeric forms of HIV-1-PR deriving from the reaction of cysteine residues of two or more proteins. Lane 1 show a 100-10000 bp DNA ladder (Fermentas), with the size of some of the fragments indicated in kbp. The samples were denatured in the absence of reducing agents to avoid cleavage of disulfide bonds, and the gel was stained with ethidium bromide.

Table 3-1 List of non-native contacts forming in at least 2% of the time in each of the five unfolding trajectories. Numbers indicate the amino acids involved in the non-native contacts.

Trajectory 1	Trajectory 2	Trajectory 3	Trajectory 4	Trajectory 5
23 51	1 7	1 7	1 6	1 7
24 70	1 8	1 9	1 7	1 9
25 51	1 9	1 10	26 51	28 51
31 65	8 83	24 51	26 82	29 51
33 65	20 67	24 70	27 51	30 51
33 67	21 67	25 51	27 82	31 71
34 67	21 68	26 51	28 51	31 72
36 66	24 70	27 51	28 80	32 67
36 67	26 82	31 65	28 81	32 70
47 80	27 82	33 63	28 82	32 72
66 76	28 81	33 65	31 65	33 65
69 86	28 82	33 67	33 63	33 67
69 87	31 65	33 70	33 65	33 70
70 87	31 70	36 67	33 67	33 72
71 86	31 72	36 70	33 70	36 67
71 87	33 63	36 71	33 72	36 70
75 86	33 65	36 72	34 67	36 72
	42 78	37 67	36 67	41 63
	45 62	38 67	36 70	50 82
	50 82	38 72	36 72	69 76
	69 86	39 67	37 67	69 86

	69 87 70 86 71 86 71 87	41 63 47 80 59 70 69 86 69 87 70 86 71 86 75 86 87 98	38 67 41 63 47 80 59 70 69 86 69 87 71 86 74 86 75 86	75 93

References

[1] T. Maniatis, E.F. Frisch, J. Sambrook, Book Chapter,Molecular Cloning, a laboratory manual. Ed, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA., (1988).

[2] C. Cecconi, E. Shank, F. Dahlquist, S. Marqusee, C. Bustamante, Protein-DNA chimeras for single molecule mechanical folding studies with the optical tweezers, Eur Biophys J, 37 (2008) 729-738.

Chapter: 4

Effect of ligand binding on the (un)refolding processes of Acyl-CoA-binding protein

Chapter 4: Effect of ligand binding on the (un)refolding processes of Acyl-CoA-binding protein

Abstract

Acyl-CoA-binding protein (ACBP) is a topologically simple and highly conserved 86 residue protein made of four α-helices (H_1-H_4) organized in a unique up-down-down-up skewed bundle. ACBP is involved in various cellular processes, including fatty acid elongation and membrane assembly, and it binds acyl-CoA esters with a broad range of acyl-chain length (C_4-C_{22}). We have previously described the behavior under tension of ACBP, revealing an unprecedented extended transition state (*TS*) located almost halfway between the unfolded and native states (Heidarsson et al., *JACS*, 2012). Here we use single-molecule optical tweezers and MD simulations to investigate the effect of Octanoyl-CoA (C_8) binding on unfolding and refolding processes of the ACBP. Our results show that Octanoyl-CoA binding enhances the activation free energies for the unfolding, without affecting the position of the transition state along the reaction coordinate. It follows that ligand binding increases the mechanical resistance of the protein, without changing its mechanical compliance. These results can be rationalized in terms of key interactions Octanoyl-CoA establishes with helices 2 and 3, comprising the structured part of the intermediate state. The effects of ligand binding on ACBP's mechanical properties are discussed in terms of protein function.

Keywords: Single-molecule studies, ACBP, Octanoyl-CoA, MD simulation.

Introduction

Proteins interacts with a multitude of binding partners in vivo and these interactions naturally affect protein structure and energetics. These binding partners play a crucial role when interacts with proteins as these interactions helps in many biological processes, such as signal transduction, molecular transport, cell regulation, gene expression control, enzyme inhibition, antibody–antigen recognition, and even the assembly of multi-domain proteins. Therefore, it is clear that the biological function of a protein depends on its interaction with ligand molecules.

Mechanical stability of the protein, is largely governed by specific noncovalent interactions and ligand binding can induce conformational changes in the protein structure that lead to changes in mechanical stability. Alternation in the mechanical stability upon ligand binding can therefore serve as an intrinsic reporter to identify the functional state of the protein. Knowing the resistance of the protein is important because many processes in living systems such as cell division, locomotion, enzyme activity depend critically on protein conformation changes and mechanical rigidity. Thus, understanding protein structure requires that the folding process must be studied in the presence of their binding partners. Understanding the principles of how a protein interacts with its natural ligands, will allow the development of other more potent substances that might serve as drug candidates. Therefore, since decades protein-ligand interactions have been studied both by traditional bulk methods [1-5] and single-molecule manipulation techniques [6-9]. As bulk methods are limited to the description of the overall properties of a large population of the molecules, it is advantageous to use single-molecule methods to study these interactions. ACBP is one such protein which binds with a medium and long chain of acyl-CoA esters. ACBP has been an excellent molecule to study as a wide range of kinetic and structural information has already known.

Long chain acyl-CoA esters function as a substrate and intermediate in lipid biosynthesis and catabolism and also play a role in regulating carbohydrate metabolism, protein sorting, gene expression, and signal transduction. Therefore, they are essential for numerous cellular function [10, 11]. ACBP is very simple and highly conserved 10 kDa protein binds to medium and long-chain acyl-coenzyme esters (C_{14}–C_{22}) with high specificity and affinity (K_d, 1–10 nM) [12]. First time it was discovered in 1983 in the brain and labelled as a diazepam-binding inhibitor (DBI) [13]. Now a days it is commonly referred as ACBP. It is a protein from long-chain fatty acyl-

Coenzyme A (LCFA-CoA) binding protein family found in all eukaryotes ranging from yeast to mammals [14]. Previous in vitro studies has indicated that ACBP is essential for various biochemical processes. For an instance, it plays an important role in the transport of acyl-CoA esters and it also deliver acyl-CoA esters to phospholipid [15, 16], glycerolipid [17], and cholesteryl ester synthesis [18]. ACBP is also involved in gene regulation, signal transduction, regulation of lipid and energy metabolism, vesicular trafficking, membrane biogenesis and many more functions [19]. The broad range of distribution throughout the animal and plant kingdom and its high degree of similarity among the different tissues and species suggest that ACBP is a housekeeping protein which was supported by *Mandrup et al*. whose study has proven that genomic gene of ACBP has all characteristics of a housekeeping gene [20].

ACBP is made up of four α-helices. The four α-helices are H_1 of residues 3-15, H_2 from 21-36, H_3; 51-62 and H_4; 65-84. ACBP is arranged in a very unique up-down-down-up manner. The compact form of the ligand organized so that polar parts are exposed to the surface of the complex whereas the hydrophobic groups turn inwards. The unique electrostatic network involving side chains of Tyr28, Lys32, Lys54 and the 3'-phosphate plays an important role in the binding. Ligand has both hydrophilic and hydrophobic determinants for making favorable interactions with ACBP. The molecular weight of the ligand is one-tenth of that of ACBP and is expected to have a significance influence on protein when binds to its ligand.

ACBP-ligand interactions has already been studied by bulk methods [17, 21, 22]. The results of these studies provide information not only on the ACBP binding affinity towards different acyl-chain length, but it also gives the structural information of ACBP with its ligand [23]. Despite of this fact, none of the above studies have provided information on mechanical properties or the energy landscape of ACBP. Previously, our group has characterized mechanical pathways of ACBP at single-molecule level using optical tweezers (OT) and Ratcheted molecular-dynamics simulations (RMSD). This study has shown that under tension ACBP reveals an unprecedented extended transition state (TS) that is located almost halfway between the unfolded and native states [24]. This study was restricted to ACBP alone (without its ligand), therefore we decided to go one step further and characterize mechanical pathways of ACBP in the presence of one of its ligand using OT and MD simulation. In the present study, using same earlier experimental conditions, we have characterized molecular processes of ACBP in the presence of Octanoyl CoA, C_8 and compare our results with ACBP alone.

Materials and Methods

Protein Expression, Purification, and sample preparation

The double cysteine variants $ACBP^{1-86}$ was expressed in Escherichia coli BL21(DE3)-pLysS cells transformed with a pET3a expression vector containing the mutated bovine ACBP gene [25]. Purification was performed as previously described in [26]. DNA handle preparation, their attachments to the protein and the coupling of protein-DNA chimera to polystyrene beads were performed as described in [27].

Optical tweezer experiments

We use DNA molecules as molecular handles to tether ACBP between two polystyrene beads, one 2.2 μm streptavidin-coated bead (Spherotec) is held at the end of a micropipette by suction, while the other 3.10 μm anti-digoxigenin-coated bead (Spherotec) is held in the optical trap. DNA handles allows us to manipulate protein while keeping tethering surfaces away from each other so that they does not interact and causes unnecessary non-specific interactions. For these experiments the DNA handles were attached to two cysteine residues present at position 1 and 86. The experiment is conducted at room temperature in 10 mM Tris-HCI buffer, pH 7.0, 250 mM NaCl, 10 mM CaCl2, and 44μM Octanoyl-CoA. The force applied to molecule is varied by moving the micropipette relative to the optical trap by means of a piezoelectric flexure stage (MAX3311/M, Thorlabs, Newton, NJ). The applied force was determined by measuring the change in the light momentum of the laser beams leaving the trap [28], while changes in the extension of the molecule were determined by video microscopy [28]. Molecules those showed the characteristic DNA overstretching transition at 67 pN were considered in the analysis [29].

Changes in contour length of ACBP

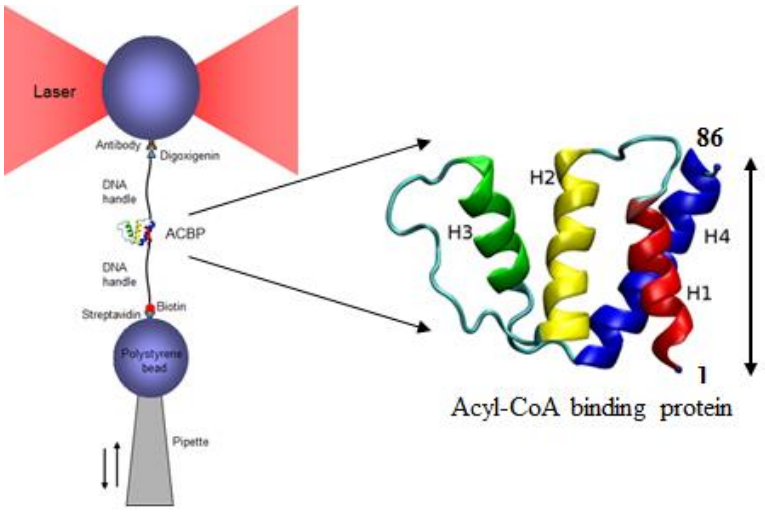
Theoretical changes in the contour length (ΔL_c) upon unfolding/refolding of ACBP is calculated as: (Number of residues between the attachment points* distance between two amino acids) - Distance between the attachment points in the folded structure. The number of amino acids: 85; Distance between two amino acids= 0.36; Distances between position 1 and 86 in the folded structure were measured using the solution structure of ACBP = 2.3nm (Protein Data Bank code 1NTI). We calculated a ΔL_c^{calc} of 28.3 nm for ACBP1-86. Experimentally, ΔL_c of ACBP associated with unfolding and refolding events was estimated by fitting force vs. extension cycles with WLC [30]. For ACBP, 1-86 was pulling axis of our experiment.

Results

Force-ramp experiment

Force-ramp is the first type experiment we have performed after tether is found between two beads. In this experiment, individual ACBP molecules were stretched and relax multiple times by moving the micropipette relative to trap. Piezoelectric stage moved with constant velocity (nm/sec) in one dimension to generate almost constant loading rates of pulling (pN/sec), likewise we have recorded data for various pulling speed (nm/s). The force applied on the protein can be determined by measuring the change in momentum flux of the light beams leaving the traps, while the extension of the molecule can be determined by measuring the bead separation distance [28]. Force is increased until a molecule unfolds. Unfolding event (rip) is marked by a sudden increase in the extension and decrease in the force of the molecule, as it goes from native compact state to extended unfolded state *(Figure 4-1B)*, while it refolds (zip), it generates transition in the opposite direction as it restores its original extension. In the presence of Octanoyl-CoA, ACBP unfolding occurs at ~12 pN, while refolding occurs at ~ 4 pN. From this data one can calculate the unfolding and refolding force distribution *(shown in figure 4-2D and E)*, which can be used to calculate the average unfolding and refolding forces at any given loading rate. Kinetic parameters such as rate constants and position of the transition state along the reaction coordinate can be calculated from force distributions. The force-ramp experiment is very simple and powerful way to profile the molecular transition.

A.



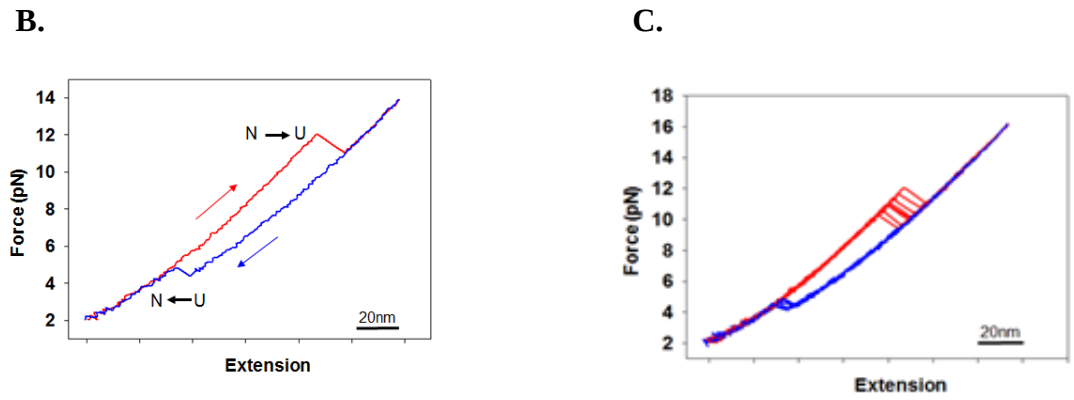


Figure 4-1 Mechanical manipulation of a single ACBP molecule by optical tweezers. A) Experimental strategy showing a single ACBP is mechanically manipulated between two polystyrenes beads by means of DNA molecules (~500 bp) that are covalently bound to two cysteine residues located at positions 1 and 86, therefore cysteine pair 1-86 is the pulling axis of experiment (shown by arrowhead line). During the experiment ACBP stretched and relaxed by moving the pipette relative to the optical trap, while monitoring force and extension. B) Force vs. extension cycle obtained by stretching (red) and relaxing (blue) of ACBP bound to Octanoyl-CoA complex with optical tweezers. Sudden discontinuity (rip) in the stretching trace ~12 pN corresponds to the mechanical denaturation of ACBP. During stretching the end-to-end distance of molecule suddenly increases as molecule goes from its compact native state (N) to extended unfolded state (U) (N to U transition). On the other hand, upon relaxation molecule regain its original extension, indicated by sudden decrease in the molecular extension (U to N transition) ~4 pN. C) Multiple force-extension traces superimposed, at constant speed of 50nm/sec.

Contour length measurement (ΔL_c)

In order to check whether molecule has completely unfolded or not we also did change in the contour length measurements. *Figure 4-2* shows fitting of force vs. extension cycles for unfolding and refolding with worm-like chain model (WLC). WLC fitting yielded a ΔL_c of 25.95 ± 2.10 (n=85) nm for unfolding, which compare well with the value of 28.3 nm (within a standard deviation) expected for the complete unfolding of the molecule *(see section "Materials and methods")*. ΔL_c value for refolding yielded, 26.52 ± 3.8 nm (n=85).

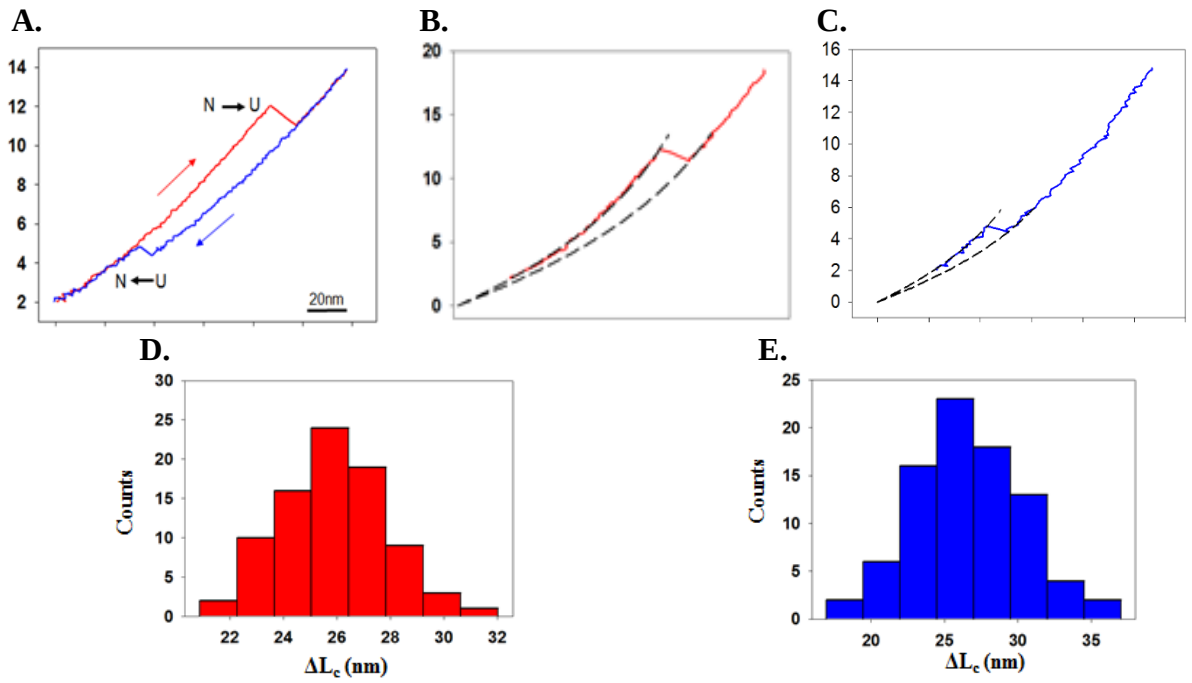


Figure 4-2 WLC fitting. A) Force vs. extension cycle of ACBP in the presence of Octanoyl CoA. B) and C) shows WLC fits (dashed line) to unfolding and refolding transitions, respectively. D) and E) are the distribution of the contour length changes for unfolding and refolding respectively. WLC fitting yielded a ΔL_c of 25.95 ± 2.10 (n=85) nm for unfolding and 26.52 ± 3.8 nm (n=85) for refolding. The data is collected at 40 Hz rate.

Figure 4-3 shows the average unfolding and refolding forces as a function of loading rates for ACBP and ACBP-Octanoyl CoA. From *figure 4-3A* it is clear that in the presence of Octanoyl-CoA, ACBP becomes more mechanical stable, while ligand binding does not affect the refolding process as there is no change in the refolding forces *(Figure 4-3B)*.

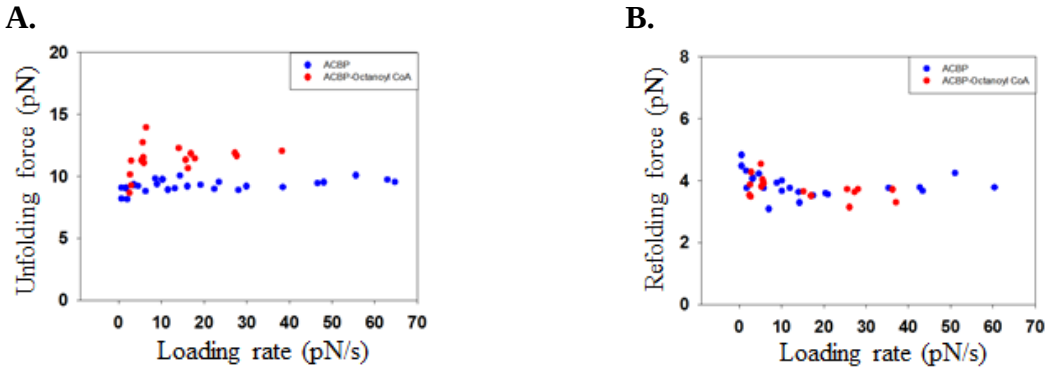


Figure 4-3 A) and B) are the average unfolding and refolding forces measured at different loading rates for ACBP (blue dots) and ACBP-Octanoyl CoA (red dots) respectively. In the presence of Octanoyl-CoA, ACBP becomes slightly more mechanical stable, while the refolding forces remain the same.

Determination of rate constants and position of TS

Unfolding and refolding force distributions can be used to calculate the position of the transition state along the pulling axis and the rate coefficient at zero force. The two state unfolding and refolding processes can be analysed with probability density function [31]:

For unfolding:

$$P(F) = ({k^0}_u/r)\exp(F\,{x^\ddagger}_u/k_B\,T)\exp\left[-\left(\frac{{k^0}_u k_B\,T}{r{x^\ddagger}_u}\right)exp\left(\frac{F{x^\ddagger}_u}{k_B\,T}\right)-1\right] \qquad (4-1)$$

For refolding:

$$P(F) = \left({k^0}_f/r\right)\exp\left(F\,{x^\ddagger}_f/k_B\,T\right)\exp\left[-\left(\frac{{k^0}_f k_B\,T}{r{x^\ddagger}_f}\right)exp\left(\frac{F{x^\ddagger}_f}{k_B\,T}\right)-1\right] \qquad (4-2)$$

Where ${k^0}_u$ and ${k^0}_f$ are the rate of unfolding and refolding at zero force respectively, r is the loading rate (pN/s), ${x^\ddagger}_u$ and ${x^\ddagger}_f$ are the distances to the transition state from native and unfolded state respectively. F is the applied force, k_B is Boltzmann constant and T is the absolute temperature. The unfolding and refolding force distributions can be fitted with probability density functions *equations (4-1) and (4-2)* to extract as the best (least square) parameters, the rate of unfolding and refolding at zero force (k_u^0 and k_f^0) from intercepts and distances to the transition state from native state (${x^\ddagger}_u$) and unfolded state (${x^\ddagger}_f$) via slopes *(Figure 4-4A and B) (for more details see section "Theoretical models of single-molecule force spectroscopy" of chapter 1)*. This analysis yielded: ${k^0}_u$= 4.99(± 4.21)E-7 sec^{-1}, ${x^\ddagger}_u$ =5.9 ± 0.35 nm, and ${k^0}_f$= 7.57(±3.02)E+3 sec^{-1}, ${x^\ddagger}_f$= 7.8 ± 0.36 nm.

Similarly, for a two state system, in the case of force increasing linearly with time, then position of TS and rate constants can be obtained from following equations [32, 33]:

$$ln\{r\,ln(1/N(F,r))\} = ln\frac{k_m k_u^0 k_B T}{{x^\ddagger}_u} + ({x^\ddagger}_u/k_B T)\,F \qquad (4-3)$$

And for refolding

$$ln\{-r\,ln(1/U(F,r))\} = ln\frac{k_m k_f^0 k_B T}{x^{\ddagger}_f} - (x^{\ddagger}{}_f/k_B T)\,F \qquad (4-4)$$

Where, $N(F,r)$ and $U(F,r)$ are the fractions of the folded and unfolded molecules at various forces F which can be calculated by integrating the histograms of the force distributions over the corresponding range of forces, r is the loading rate (pN/s). *Equations (4-3) and (4-4)* can be used to fit $ln\{r\,ln(1/N(F,r))\}$ vs. F, and $ln\{-r\,ln(1/U(F,r))\}$ vs. F to estimate distances to TS and rate constants from the slope and intercept of the graphs respectively *(Figure 4-4C)*. This analysis yielded: $k^0{}_u$= 3.48(±6.59)E-7 sec^{-1}, $x^{\ddagger}{}_u$ = 6.1 ± 0.73nm, and $k^0{}_f$= 6.75(± 4.93)E+3 sec^{-1}, $x^{\ddagger}{}_f$= 8.0 ± 0.63 nm.

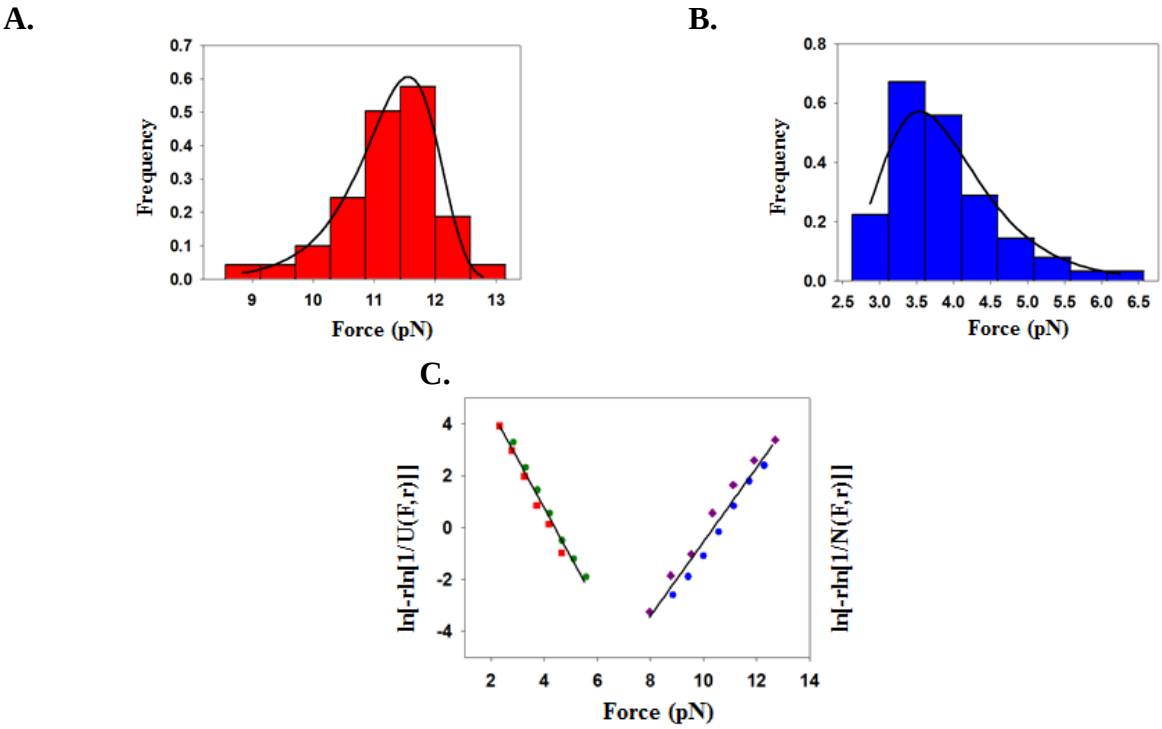


Figure 4-4 A) Unfolding force distribution (red) fitted to the probability density function $P(F)=(k_u^0/r)\exp\,(Fx^{\ddagger}_u/k_BT)\exp[-((k_u^0k_BT)/(rx^{\ddagger}_u))(\exp(Fx^{\ddagger}_u/k_BT)-1)]$, where k_u^0 is the unfolding rate at zero force, r is the loading rate, $x^{\ddagger}_u$ is the distance to the transition state from native state, k_B is the Boltzmann's constant and T is the absolute temperature. B) The refolding force distribution (blue) was also fitted with probability distribution function $P(F)=(k_f^0/r)\exp(Fx^{\ddagger}_f/k_BT)\exp[-((k_f^0k_BT)/(rx^{\ddagger}_f))(\exp(Fx^{\ddagger}_f/k_BT)-1)]$, where k_f^0 is the refolding rate at zero force, $x^{\ddagger}_f$ is the distance to the transition state from unfolded state. The unfolding and refolding force distribution at a loading rate of 2.9

pNs^{-1} and 26.1 pNs^{-1} respectively. Blank lines are the fitting lines produced by fitting distributions with these functions. The best fit (least square) values give $x^{\ddagger}_u$= 5.9 ± 0.35 nm and $x^{\ddagger}_f$ = 7.8 ± 0.36 nm and rate constants for unfolding and refolding are k_u^0: 4.99(±4.21)E-7 sec^{-1} and k_f^0: 7.57(±3.02)E+3 sec^{-1} respectively. C) Shows a plot of ln [rln(1/N)] and ln [-rln(1/U)] vs. Force. N and U are the folded and unfolded fractions respectively. Red solid circles and stars represent unfolding data collected at 2.9 and 5.9 pN/s. Blue solid circles and stars represent refolding data collected at 26.1 and 15.3 pN/s respectively. Fit of the unfolding data with equation yielded: k_u^0= 3.48(± 6.59)E-7 sec^{-1}, $x^{\ddagger}_u$ = 6.1 ± 0.73nm. Fit of the refolding data with equation yielded: k_f^0= 6.75(± 4.93)E+3 sec^{-1}, $x^{\ddagger}_f$ = 8.0 ± 0.63 nm.

Table 4-1 shows the kinetic parameters obtained by force-ramp method.

Variant	Experiment	x_u (nm)	x_f (nm)	k_u^0 (s^{-1})	k_f^0 (s^{-1})
ACBP-Octanoyl	Force-ramp	5.9 ± 0.35	7.8 ± 0.36	4.99(± 4.21)E-7	7.57(± 3.02)E+3

Table 4-1 Shows kinetic parameters obtained with force-ramp experiment.

Force-Jump experiment

If a molecule unfolds and refolds at equilibrium, then constant force measurements can be done, in which the force is kept constant to preset value with feedback mechanism and molecular extension is monitored over time. In these experiments, rate coefficients can be obtained directly from the lifetimes of the folded and unfolded states. To perform these experiments the only condition is that the rate of unfolding and refolding reaction should be high enough so that a large number of recordings is done in a shorter period of time.

In the presence of Octanoyl-CoA, ACBP unfolds ~ 12 pN and refolds ~ 4 pN *(see figure 4-1B)*. Because of the large hysteresis we cannot held force constant at equilibrium position, let's say around 8pN, as in this case molecule is far away from its unfolding or refolding forces and therefore to acquire a significant number of events, it requires long hours of recordings. Recording data for such a long time needs instrumental stability that our setup does not have. To overcome this problem, one can do another type of constant force experiments called force-jump experiments, in which the force is jumped between two constant force values. First it jumped to a force value which is close to unfolding force, and then force kept constant with feedback

mechanism until there is a change in the molecular extension then immediately force is jumped down to another force value which is close to refolding force and it kept constant until extension changes. By doing so we have characterized unfolding and refolding processes of ACBP-Octanoyl CoA.

A typical force-jump experiment with ACBP-Octanoyl CoA is shown in *figure 4-5A*. At the beginning molecule is folded at 4 pN, force is increased (jumped) from 4pN to 9.5 pN and then it held constant through the feedback mechanism until a change in the molecular extension, it marks unfolding event. At this point force is decreased (dropped) down to 4 pN and held constant until decrease in the molecular extension, this marks the refolding event. Then force is raised again and a new cycle can start. For each cycle, we can measure the time molecule takes to unfolds (τ_u) and to refolds (τ_f). We measured unfolding and refolding cycles for several cycles. Then we can plot dwell-time distribution of these measured times. Dwell-time distribution can be well fit with exponential function to estimate the rates of unfolding and refolding at these two specific forces *(Figure 4-5B and 5C)*. Then we can change force values, let's say 5 pN for refolding and 10 pN for unfolding and repeated same procedure and measured τu and τf values directly at that forces. Repeating same procedure for several force constants values, unfolding and refolding rates were measured through single exponential fit. Further, we can plot the estimated rates as a function of force which can be fitted to the Bell's model to calculate distances to the transition state and rate constants for unfolding and refolding.

Determination of rate constants and position of TS

The two states unfolding and refolding processes of ACBP-Octanoyl CoA can be analyzed with phenomenological Bell model, which postulates that the unfolding and refolding rate constants depend exponentially on force [32, 34].

For unfolding:

$$\mathrm{k_u}(F) = {k^0}_u \exp(F x_u / k_B T) \qquad (4-5)$$

For refolding:

$$\mathrm{k_f}(F) = {k^0}_f \exp\left(-F x_f / k_B T\right) \qquad (4-6)$$

In logarithmic form equation (4-5) and (4-6) can be written as:

$$\ln k_u(F) = ln{k^0}_u + (Fx_u/k_BT) \quad (4-7)$$

$$\ln k_f(F) = ln{k^0}_f - (Fx_f/k_BT) \quad (4-8)$$

Where, ${k^0}_u$ and ${k^0}_f$ are unfolding and refolding rate constants of molecule at zero force, x_u and x_f are the distances to the transition state from native state and unfolding state respectively. F is applied force, k_B is the Boltzmann constant and T is absolute temperature. Data obtained from force-jump experiments can be fitted with *equations (4-7) and (4-8)* to estimate rate constants and distances to the transition state from intercepts and slopes of the graph respectively *(Figure 4-6)*. This analysis yielded: ${k^0}_u$ = 2.91(±5.27)E-7 sec^{-1}, x_u= 6.0 ± 0.73 nm, and ${k^0}_f$= 5.44(± 4.92)E+3 sec^{-1}, x_f = 7.9 ± 0.73 nm.

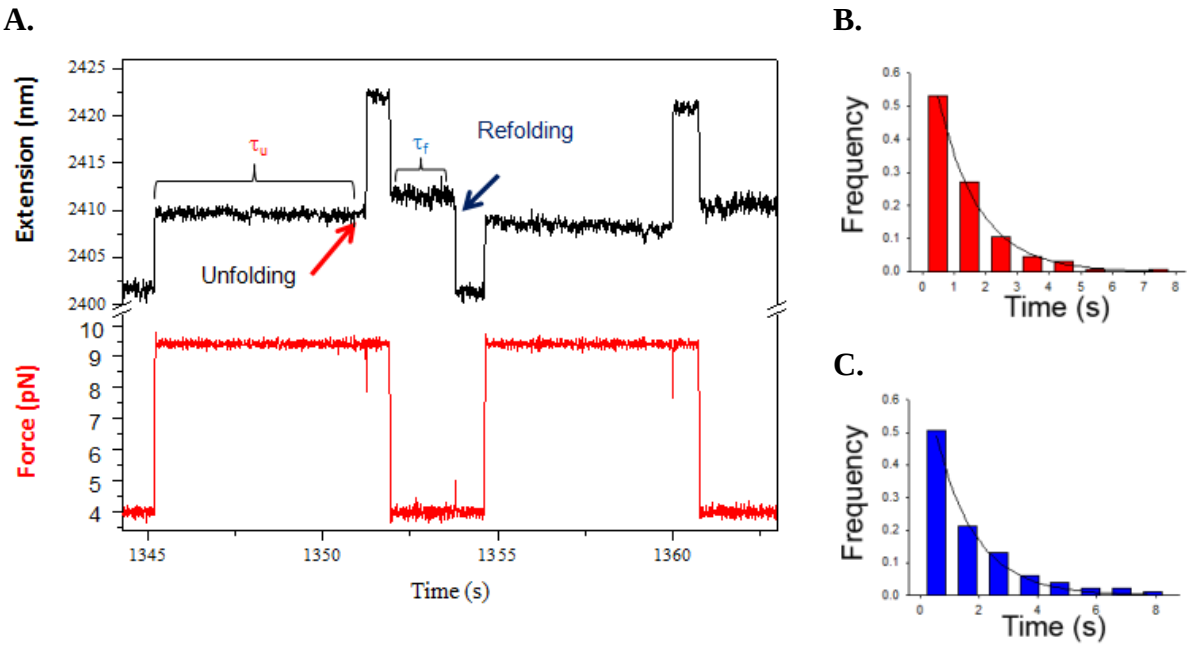


Figure 4-5 A) Two-force jump cycles. The black trace shows how extension of molecule suddenly increases or decreases respectively. The red trace shows the force applied on molecule. Panels B) and C) shows dwell-time histograms of the folded and unfolded states measured at 9.5 pN and 4.0 pN respectively. Solid lines (black) are the single exponential fits to the data. Force-extension data is recorded at the rate of 100 Hz.

Figure 4-6 shows the graph of ln k vs. F for ACBP-Octanoyl CoA obtained through force-jump experiment.

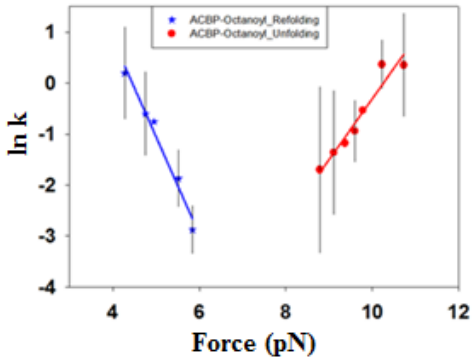


Figure 4-6 Plot of the logarithm of the unfolding (red dots) and refolding (blue stars) rate constants measured through force-jump experiments at different forces. Solid lines are fits with bell model. Fit with the unfolding data with *equation (4-7)* yielded: k_u^0= 2.91(±5.27)E-7 sec^{-1}, x_u = 6.0 ± 0.73 nm. Fit of the refolding data with *equation (4-8)* yielded k_f^0= 5.44(±4.92)E+3 sec^{-1}, x_f = 7.9 ± 0.73 nm.

Table 4-2 shows the kinetic parameters obtained by force-jump method.

Variant	Experiment	x_u (nm)	x_f (nm)	k_u^0 (s^{-1})	k_f^0 (s^{-1})
ACBP-Octanoyl	Force-jump	6.0 ± 0.73	7.9 ± 0.73	2.91(± 5.27)E-7	5.44(± 4.92)E+3

Table 4-2 Shows kinetic parameters obtained with force-jump experiments.

Comparison of force-ramp and force-jump results for ACBP-Octanoyl CoA and ACBP

Table 4-3 summarizes the kinetic values obtained with force-ramp and force-jump method for ACBP-Octanoyl complex and ACBP alone.

Variant	Experiment	x_u (nm)	x_f (nm)	k_u^0 (s^{-1})	k_f^0 (s^{-1})
ACBP-Octanoyl	Force-ramp	5.9 ± 0.35	7.8 ± 0.36	4.99(±4.21)E-7	7.57(±3.02)E+3
	Force-jump	6.0 ± 0.73	7.9 ± 0.73	2.91(±5.27)E-7	5.44(±4.92)E+3
ACBP	Force-ramp	5.4 ± 0.18	6.7 ± 0.16	2.84(±1.00)E -5	3.79(±0.80)E+3
	Force-jump	5.5 ± 0.60	6.5 ± 0.30	3.12(±2.50)E -5	4.25(±0.66)E+3

Table 4-3 Comparison of kinetic parameters obtained with force-ramp and force-jump experiments for ACBP-Octanoyl and ACBP. Upon ligand binding rate of unfolding decrease by magnitude of two, while the refolding rates remain unaffected which means that ligand bind to ACBP only after protein gets folded. The rate constants obtained through OT measurements contains contribution of DNA handles, beads therefore they cannot be compared with intrinsic rate constants of the protein.

If we compare results obtained from force-ramp and force-jump experiments for ACBP bound to Octanoyl-CoA with kinetic data obtained previously for ACBP alone [35], we can clearly observe that the distances to the transition state, both from the native state and the unfolded state are same (within a standard deviation) in both cases (ACBP-Octanoyl CoA and ACBP alone, as x_u and x_f remains about the same), while in the presence of ligand unfolding rate constants decreased by 2 order of magnitude (changes from 10E-5 to 10E-7, *Table 4-3*), because of this there is increase in the height of barrier that therefore leads to increase in the mechanical resistance of ACBP. The refolding process seems unaffected as k_f^0 remains same upon Octanoyl-CoA binding. The possible explanation for this can be, probably Octanoyl-CoA bind to ACBP only after the protein has folded. An unfolded ACBP first fold into its native state and then Octanoyl-CoA bind. In this way the ligand does not affect the refolding process.

Molecular Dynamics Simulations

To gain insights into the molecular mechanism by which Octanoyl-CoA affect the unfolding and refolding processes of ACBP, MD simulations were performed in collaboration with Dr. Luca Bellucci and Dr. Stefano Corni.

Figure 4-7 shows the force vs. distance (extension) traces generated by MD simulation. Black trace shows the mechanical denaturation of ACBP alone (APO) and in red is the ACBP bound to Octanoyl-CoA (HOLO). From the graph it is clear that when ACBP bound to Octanoyl-CoA, it became more mechanically resistant than ACBP alone.

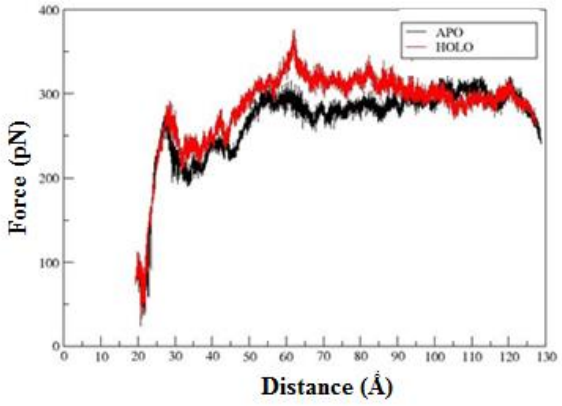
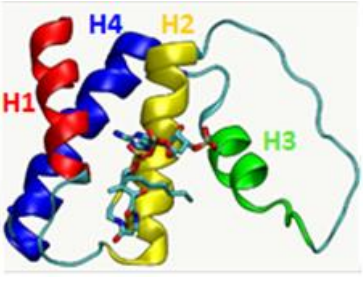


Figure 4-7 Force vs. distance trace of ACBP-Octanoyl CoA (red) and ACBP alone (black) respectively.

To understand the mechanism by which Octanoyl-CoA affects the unfolding process of ACBP we should look in detail at how Octanoyl-CoA interacts with the protein. The NMR structure of ACBP-Octanoyl CoA is shown in *figure 4-8A* which clearly shows that, Octanoyl-CoA interacts mostly with helices H_2 and H_3. The transition state of the mechanical denaturation of ACBP has been proposed by *Heidarsson et al.*[35] (*see figure 4-8B*), where helix H_1 is completely unfolded, H_4 is partially unfolded and H_2 and H_3 are still folded. According to this structure the extension of the transition state is determined primarily by the denaturation of H_1 and H_4, which are helices with which the ligand does not interact, therefore, there is no change in the position of the transition state, while the structured part of the transition state that probably is responsible for the height of the unfolding activation barrier is made of helices H_2 and H_3 with which the ligand interacts and probably stabilizes and therefore, we need additional force to cross the kinetic barrier, hence there is an increase in the mechanical stability of ACBP.

A.



B.

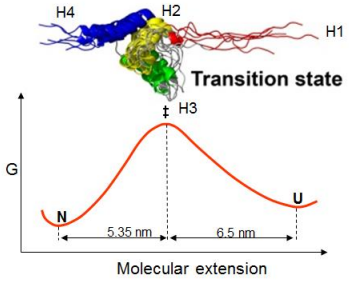


Figure 4-8 A) Three-dimensional structure of ACBP-Octanoyl CoA complex determined by NMR. B) Proposed structure of transition state,where helix H_1 is completely unfolded, H_4 is partially unfolded and H_2 and H_3 are still folded [35].

Figure 4-9 shows the snap shots of mechanical denaturation of ACBP and ACBP-Octanoyl CoA. We see that the unfolding of helices H_1 and H_4 take place more or less in the same manner in the two cases (ACBP and ACBP-Octanoyl CoA), and this is probably why the ligand does not affects the position of the transition state, while if we look at the structure having an extension of about 6 nm, which is more or less the extension of the transition state, in this case we see helices 2 and 3 slightly separated by force while in this case they are held together by the ligand, and this might be the mechanism by which the ligand increases the height of the unfolding activation barrier.

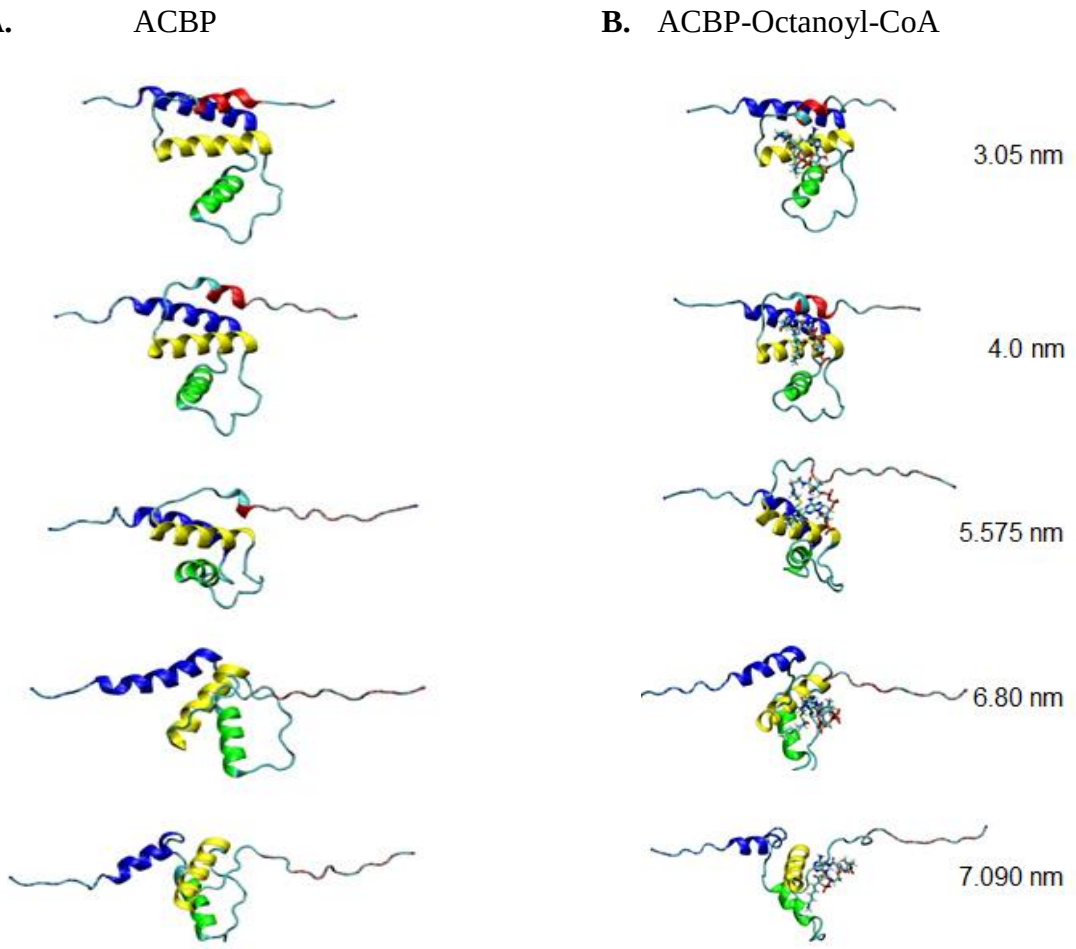


Figure 4-9 MD simulations of mechanical unfolding of ACBP A) and ACBP-Octanoyl-CoA B). Snapshots of molecular structures visited by the protein along the unfolding trajectories at incresing distace between Cys1 and Cys86 (end-to-end distance is shown in right side of *figure 4-8B)*.

From OT and MD simulations, we can say that Octanoyl-CoA binding changes kinetic parameters of ACBP. As protein-ligand interactions play a crucial role in almost all biological processes, studying these interactions will not only help us understand biological processes at the single molecular level but also help to design new drugs.

Conclusions

For the first time we have studied unfolding and refolding processes of ACBP-Octanoyl CoA complex at the single-molecule level using optical tweezers and MD simulations. In conclusion we can say that:

- Octanoyl-CoA binding decreases the unfolding rates of ACBP without changing the position of the transition state.
- Octanoyl-CoA binding increases the mechanical resistance of ACBP without changing its mechanical compliance.
- Octanoyl-CoA binding has no effect on the refolding rates of ACBP, suggesting that binding takes place after protein folding.
- MD simulations suggest that the effect of ligand binding on the unfolding and refolding processes of ACBP are due to interactions that Octanoyl-CoA undertakes with helices H_2 and H_3.

References

[1] J. Kool, N. Jonker, H. Irth, W.M. Niessen, Studying protein-protein affinity and immobilized ligand-protein affinity interactions using MS-based methods, Analytical and bioanalytical chemistry, 401 (2011) 1109-1125.

[2] M. Baldus, Molecular interactions investigated by multi-dimensional solid-state NMR, Curr Opin Struct Biol, 16 (2006) 618-623.

[3] S.W. Homans, Probing the binding entropy of ligand-protein interactions by NMR, Chembiochem, 6 (2005) 1585-1591.

[4] R. Perozzo, G. Folkers, L. Scapozza, Thermodynamics of protein-ligand interactions: history, presence, and future aspects, Journal of receptor and signal transduction research, 24 (2004) 1-52.

[5] B. Meyer, T. Peters, NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors, Angewandte Chemie (International ed. in English), 42 (2003) 864-890.

[6] P. Bechtluft, R.G.H. van Leeuwen, M. Tyreman, D. Tomkiewicz, N. Nouwen, H.L. Tepper, A.J.M. Driessen, S.J. Tans, Direct Observation of Chaperone-Induced Changes in a Protein Folding Pathway, Science, 318 (2007) 1458-1461.

[7] J. Stigler, F. Ziegler, A. Gieseke, J.C.M. Gebhardt, M. Rief, The Complex Folding Network of Single Calmodulin Molecules, Science, 334 (2011) 512-516.

[8] J.P. Junker, K. Hell, M. Schlierf, W. Neupert, M. Rief, Influence of Substrate Binding on the Mechanical Stability of Mouse Dihydrofolate Reductase, Biophysical journal, 89 (2005) L46-L48.

[9] E. Hann, N. Kirkpatrick, C. Kleanthous, D.A. Smith, S.E. Radford, D.J. Brockwell, The Effect of Protein Complexation on the Mechanical Stability of Im9, Biophysical journal, 92 (2007) L79-L81.

[10] N.J. Faergeman, J. Knudsen, Role of long-chain fatty acyl-CoA esters in the regulation of metabolism and in cell signalling, The Biochemical journal, 323 (Pt 1) (1997) 1-12.

[11] J. Knudsen, T.B. Neergaard, B. Gaigg, M.V. Jensen, J.K. Hansen, Role of acyl-CoA binding protein in acyl-CoA metabolism and acyl-CoA-mediated cell signaling, The Journal of nutrition, 130 (2000) 294s-298s.

[12] J. Knudsen, J. Burton, N.J. Færgeman, Long chain acyl-CoA esters and acyl-CoA binding protein (ACBP) in cell function, Lipobiology (van der Vusse, G., ed.), Elsevier B.V., Amsterdam, 2004, pp. 123–153.

[13] I. Mocchetti, M.R. Santi, Diazepam binding inhibitor peptide: cloning and gene expression, Neuropharmacology, 30 (1991) 1365-1371.
[14] M. Burton, T.M. Rose, N.J. Faergeman, J. Knudsen, Evolution of the acyl-CoA binding protein (ACBP), Biochem. J., 392 (2005) 299-307.
[15] H. Fyrst, J. Knudsen, M.A. Schott, B.H. Lubin, F.A. Kuypers, Detection of acyl-CoA-binding protein in human red blood cells and investigation of its role in membrane phospholipid renewal, The Biochemical journal, 306 (Pt 3) (1995) 793-799.
[16] L. Kannan, J. Knudsen, C.A. Jolly, Aging and acyl-CoA binding protein alter mitochondrial glycerol-3-phosphate acyltransferase activity, Biochim Biophys Acta, 1631 (2003) 12-16.
[17] J.T. Rasmussen, N.J. Faergeman, K. Kristiansen, J. Knudsen, Acyl-CoA-binding protein (ACBP) can mediate intermembrane acyl-CoA transport and donate acyl-CoA for beta-oxidation and glycerolipid synthesis, The Biochemical journal, 299 (Pt 1) (1994) 165-170.
[18] H. Chao, M. Zhou, A. McIntosh, F. Schroeder, A.B. Kier, ACBP and cholesterol differentially alter fatty acyl CoA utilization by microsomal ACAT, Journal of lipid research, 44 (2003) 72-83.
[19] J. Knudsen, M. Jensen, F\sgmaelig, N. rgeman, T.F. Neergaard, B. Gaigg, Role of acylCoA binding protein in acylCoA transport, metabolism and cell signaling, Mol Cell Biochem, 192 (1999) 95-103.
[20] S. Mandrup, R. Hummel, S. Ravn, G. Jensen, P.H. Andreasen, N. Gregersen, J. Knudsen, K. Kristiansen, Acyl-CoA-binding protein/diazepam-binding inhibitor gene and pseudogenes: A typical housekeeping gene family, Journal of Molecular Biology, 228 (1992) 1011-1022.
[21] J.T. Rasmussen, T. Börchers, J. Knudsen, Comparison of the binding affinities of acyl-CoA-binding protein and fatty-acid-binding protein for long-chain acyl-CoA esters, Biochem. J., 265 (1990) 849-855.
[22] J. Rosendal, P. Ertbjerg, J. Knudsen, Characterization of ligand binding to acyl-CoA-binding protein, The Biochemical journal, 290 (Pt 2) (1993) 321-326.
[23] B.B. Kragelund, K.V. Andersen, J.C. Madsen, J. Knudsen, F.M. Poulsen, Three-dimensional Structure of the Complex between Acyl-Coenzyme A Binding Protein and Palmitoyl-Coenzyme A, Journal of Molecular Biology, 230 (1993) 1260-1277.
[24] P.O. Heidarsson, I. Valpapuram, C. Camilloni, A. Imparato, G. Tiana, F.M. Poulsen, B.B. Kragelund, C. Cecconi, A Highly Compliant Protein Native State with a Spontaneous-like

Mechanical Unfolding Pathway, Journal of the American Chemical Society, 134 (2012) 17068-17075.
[25] K. Teilum, B.B. Kragelund, F.M. Poulsen, Transient Structure Formation in Unfolded Acyl-coenzyme A-binding Protein Observed by Site-directed Spin Labelling, Journal of Molecular Biology, 324 (2002) 349-357.
[26] S. Mandrup, P. Højrup, K. Kristiansen, J. Knudsen, Gene synthesis, expression in Escherichia coli, purification and characterization of the recombinant bovine acyl-CoA-binding protein., Biochem. J. , 276 (1991) 817-810.
[27] C. Cecconi, E. Shank, F. Dahlquist, S. Marqusee, C. Bustamante, Protein-DNA chimeras for single molecule mechanical folding studies with the optical tweezers, Eur Biophys J, 37 (2008) 729-738.
[28] S.B. Smith, Y. Cui, C. Bustamante, Optical-trap force transducer that operates by direct measurement of light momentum., Biophotonics Pt B, 361 (2003) 134-162.
[29] C. Cecconi, E.A. Shank, C. Bustamante, S. Marqusee, Direct Observation of the Three-State Folding of a Single Protein Molecule, Science, 309 (2005) 2057-2060.
[30] C. Bustamante, Z. Bryant, S.B. Smith, Ten years of tension: single-molecule DNA mechanics, Nature, 421 (2003) 423-427.
[31] I. Tinoco Jr, C. Bustamante, The effect of force on thermodynamics and kinetics of single molecule reactions, Biophysical chemistry, 101–102 (2002) 513-533.
[32] J. Liphardt, B. Onoa, S.B. Smith, I. Tinoco, C. Bustamante, Reversible Unfolding of Single RNA Molecules by Mechanical Force, Science, 292 (2001) 733-737.
[33] E. Evans, K. Ritchie, Dynamic strength of molecular adhesion bonds, Biophysical journal, 72 (1997) 1541-1555.
[34] G. Bell, Models for the specific adhesion of cells to cells, Science, 200 (1978) 618-627.
[35] P.O. Heidarsson, I. Valpapuram, C. Camilloni, A. Imparato, G. Tiana, F.M. Poulsen, B.B. Kragelund, C. Cecconi, A highly compliant protein native state with a spontaneous-like mechanical unfolding pathway, Journal of the American Chemical Society, 134 (2012) 17068-17075.

Chapter: 5

Conclusions

Chapter 5: Conclusions

Optical tweezers (OT) is an established tool in the field of life sciences and enable us to revisit protein folding with a completely new approach. Protein folding is one of the major unsolved challenges for modern biophysics. This thesis reports single-molecule studies on two proteins, HIV-1-protease and Acyl-CoA binding protein. Using custom-built dual beam optical tweezers and Molecular Dynamics (MD) simulation we have characterized unfolding and refolding pathways of these proteins. In both projects my role was i) to carry out optical tweezer experiments. ii) analysing experimental data with different theoretical models. iii) interpretation of the analysis in terms of structure and possible energy landscape and iv) writing manuscripts.

HIV-1-Protease (HIV-1-PR) is dimeric in nature. To study HIV-1-PR through OT experiments, a monomeric form of HIV-1-PR was generated by site-directed mutagenesis. Along with it some molecular biology such as, generation of DNA handles, and their covalent attachments to the protein has been done by our collaborator Martina Caldarini from Department of Physics, University of Milan, Italy. When manipulated from native state, we have observed that the molecule unfolds mainly along two pathways, either it unfolds in two state manner (64% cases) or it populated an intermediate state (30% cases), which might be characterized by the native-like arrangement of hairpins β_2 and β_3. Taking an advantage of its slow folding process, we have also characterized the unfolding behavior of the ensemble of partially folded (PF) conformations the molecule visits en route to the native state, which collectively we call the *PF* state. Our results reveal a variety of unfolding trajectories, reflecting the structural heterogeneity of this state. Remarkably, in 70% of the cases, partially folded conformations populate an intermediate state that might share common features with that observed during the unfolding of the native structure. Refolding of denatured HIV-1-PR molecules is also a multi-path process characterized by different intermediate states and energy barriers. In order to gain atomistic detail of the (un)folding processes of monomeric HIV-1-PR, MD simulation has been done in collaboration with Dr. Guido Tiana from Department of Physics, University of Milan, Italy.

Here for the first time, we have investigated unfolding and refolding pathways of monomeric HIV-1-PR at single-molecule level. Intermediate state (*I*) observed in our experiments can be regarded as the potential target for anti-AIDS therapies. A possible future study could be

to find out the exact structure of *I*, to then design a drug that can block *I*, thereby not allowing the molecule to go into its native state, which is the functional state of the protein. In this way one can block the activity of HIV-1-PR and therefore HIV.

Acyl-CoA binding protein (ACBP) functions as an intracellular carrier of acyl-CoA esters binding to medium and long chain acyl-CoA esters (C_4-C_{22}). Ligand binding may affect the molecular processes of ACBP. Therefore, the focus of my second project was to study the effect of ligand binding on unfolding and refolding processes of ACBP at single-molecule level.

To perform experiments using OT, we have received samples (ACBP-DNA chimeras) from Dr. Pétur Heidarsson from Department of Biology, University of Copenhagen, Denmark. The first type of experiment we have performed was force-ramp method in which we moves micropipette relative to optical trap, downwards and upwards movement of micropipette leads to stretching and relaxation of ACBP. We stretched/relaxed molecule at different constant pulling speeds in order to generate constant loading rates (pN/s). When manipulated ACBP unfolds ~ 12pN while refolds ~ 4 pN in the presence of ligand (Octanoyl-CoA). Unfolding and refolding force distributions generated from these experiments were then analysed with different theoretical models in order to calculate kinetic parameters. We have also performed force-jump experiments where the force was jumped between two preset values, one close to the average unfolding force, and one close to the average refolding force. By measuring the time that each molecule took to unfold and refolds, we have plotted dwell-time distributions. Fitting exponential function to these distributions provides rate constants for unfolding and refolding. By analysing this data with Bell's model, we extracted kinetic parameters. Kinetic parameters obtained from force-ramp and force-jump were in agreement.

Comparing our results with previous studies of ACBP alone, we observed that in the presence of Octanoyl-CoA, ligand binding increases the height of kinetic barrier, without changing the distances to the transition state (as x_u and x_f remains same in our experiments). Also in the presence of ligand, ACBP became more mechanically stable. In our experiments we have not observed any change in the refolding process of ACBP. MD simulation is a complementary technique to gain atomistic details of the (un)folding pathways. In collaboration with Dr. Luca Bellucci from Department of Physics, University of Modena and Reggio Emilia, Modena, Italy, we have performed MD simulation on ACBP-Octanoyl CoA and ACBP alone. When manipulated

from its native state in the presence of Octanoyl-CoA, we observed an increase in the mechanical stability of ACBP. From the snapshots of mechanical denaturation of ACBP-Octanoyl CoA complex, we can say that ligand binds to H_2 and H_3 helices of the ACBP, which is the structural part of an intermediate state. For complete denaturation of ACBP this part should break, requiring additional force that increases the kinetic barrier. Ligand does not binds to H_1 and H_4 which are responsible for the mechanical compliance of ACBP, therefore it does not change its compliance. Our MD results are in agreement to OT results.

Appendix

➢ Publications

Publications related to the Ph.D

- P.O. Heidarsson, M.M. Naqvi, **P. Sonar**, I. Valpapuram, C. Cecconi, Conformational dynamics of single protein molecules studied by direct mechanical manipulation, Advances in Protein Chemistry and Structural Biology, 92 (2013) 93-133.

- M. Caldarini*, **P. Sonar***, I. Valpapuram, D. Tavella, C. Volonté, V. Pandini, M.A. Vanoni, A. Aliverti, R.A. Broglia, G. Tiana, C. Cecconi, The complex folding behaviour of HIV-1-protease monomer revealed by optical-tweezer single-molecule experiments and molecular-dynamics simulations, Biophysical Chemistry, 195 (2014) 32-42. **[*First authorship shared]**.

- **P. Sonar,** P.O. Heidarsson, L. Bellucci, E. Paci, S. Corni, B.B. Kragelund ,C. Cecconi, Effect of ligand binding on the (un)refolding processes of Acyl-CoA-binding protein (***Manuscript under preparation***).

Other publication

- S. Nagarajan, S. Rajendran, U. Saran, M.K. Priya, A. Swaminathan, J.H. Siamwala, S. Sinha, V. Veeriah, **P. Sonar**, V. Jadhav, B.M. Jaffar Ali, S. Chatterjee, Nitric Oxide protects endothelium from cadmium mediated leakiness, Cell biology International, 37(2013) 495-506.

➢ Conferences

Oral presentations/Seminars related to the Ph.D

- *Protein folding studies at single-molecule level using optical tweezers and MD simulations*, 14 October, 2014, Soft Matter Lab, Bilkent University, Ankara, Turkey.
- *Effect of ligand binding on the energy landscape of Acyl-CoA-binding protein* at Gordon Research Seminar, Single molecule Approaches to Biology, Lucca (Barga), Italy. 12th -13th July, 2014.
- *Studying HIV-1-Protease (HIV-1-PR) folding pathways at single molecular level using Optical Tweezers* at Winter College on Optics: Treads in Laser Development and Multidisciplinary Applications to Science and Industry, ICTP, Trieste, Italy. 4th Feb-15th Feb, 2013.

Poster presentations related to the Ph.D

- Effect of ligand binding on the energy landscape of Acyl-CoA-binding protein" at Gordon Research Seminar, "Single molecule Approaches to Biology" Barga, Lucca, Italy. 12th - 13th July, 2014.
- Studying HIV1-Protease (HIV1-PR) folding pathways at single molecular level using Optical Tweezers " at "5th international symposium on Optical Tweezers in Life Sciences", Berlin, Germany, 18 June, 2013 (***First prize winner***).

Conferences attended during Ph.D tenure

- 2nd meeting of the Nanoscience Institute | CNRNANO at Modena, Italy, June 10-11, 2013.
- Winter College on Optics: Trends in laser development and multidisciplinary applications to science and industry at ICTP Italy, during 4th Feb – 15th Feb 2013.
- Preparatory School to winter College on Optics, Trieste, Italy, 28 Jan-1 Feb 2013.

Visiting student during Ph.D

Institute **:** Structural Biology and NMR laboratory, Department of Biology, University of Copenhagen, Denmark.

Supervisor **:** Prof. Birthe B. Kragelund

Duration **:** 14th Nov-11th Dec, 2014

*****