Single-molecule force spectroscopy of α-synuclein

by

Allison Kiran Solanki

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Physics
University of Alberta

Abstract

Intrinsically disordered proteins (IDPs) represent a large subpopulation of the proteome, and are characterized by high structural plasticity and a predisposition to aggregate. These aggregates can accumulate inside or outside the cell, often forming amyloid plaques that are implicated in a variety of maladies, including Alzheimer's disease, type II diabetes and Parkinson's disease. Despite many advances in elucidating aggregation mechanisms, much is still unknown about the early oligomeric states, which are the suspected toxic agents in disease, since their transient nature makes them hard to study with traditional experimental methods.

Here we have investigated the initial stages of aggregation of one such IDP, α-synuclein, a protein that is involved in Parkinson's disease and other related dementias. To do so, we developed assays for characterizing the transient structure, stability, and kinetics of monomers and oligomers using optical tweezers with high spatial and temporal resolution. Measurements of the molecular extension as the proteins unfolded under tension revealed that even small oligomers could form numerous metastable structures, with a surprisingly broad range of sizes. Moreover, our data also revealed rapid fluctuations at low force, arising from the folding of two different classes of structure that were only marginally stable. The energy landscape for these transitions was characterized via the force-dependent kinetics derived from correlation analysis of the extension trajectories. The barriers were small, only a few $k_{\mathrm{B}}T$, but the diffusion was slow,

revealing a landscape that is flat but rough. This thesis presents the first measurements of α-synuclein using optical tweezers, and also provides the first experimentally reconstructed energy landscape for an IDP.

Dedication

To my family,

for their selfless and unwavering support over the years.

Acknowledgements

To begin, I wish to thank my advisor, Prof. Michael Woodside, who has provided excellent guidance throughout my PhD. Also, I am greatly indebted to my committee members, Prof. Jack Tuszynski and Prof. Mark Freeman, who have offered helpful advice during the many supervisory meetings over the past six years. Additionally, it is a pleasure to have Prof. Steven Plotkin and Dr. Valerie Sim serve on my defense committee. Thank you for taking time out of your busy schedules to contribute insightful suggestions for my thesis.

Thank you to my α-synuclein partner-in-crime, Dr. Krishna Neupane, who has worked with me since the early days of this research. This project has been equal parts exciting and frustrating, and without your help, my PhD would have lasted at least twice as long. Thank you for training me on the optical tweezers and for your endless cooperation as we spent many an hour studying this stubborn little protein. It has been wonderful working with you.

Also, I want to acknowledge all of my fellow Woodside Lab colleagues, past and present. You are all dedicated scientists who have provided crucial suggestions and advice on my project. I would like to give special thanks to my fellow graduate student Daniel Foster, who has been a great friend and labmate since I arrived in Edmonton.

Thank you to our previous technicians, Dr. Iveta Sosova and Angela Brigley, who patiently taught me the way around a biochemistry lab. Dr. Sosova, in particular, designed the oligomer constructs used here and helped optimize

purification protocols. Also, thank you to Bethany Griffith and Meijing Wang for their help supplying me with reagents and performing a variety of crucial, often time-consuming tasks. It has been a pleasure spending time in the lab with everyone. Also, thank you to Dr. Miro Belov for building the optical trap I was fortunate enough to take my measurements on.

Last, but absolutely not least, I wish to thank my family, to whom I dedicate my thesis. To my parents, who instilled a love of learning into me at a very young age and have always been available to lend support and sage advice. Finally, I wish to thank my husband John, whose encouragement and patience has been invaluable during every step of my PhD experience. Thank you for being my biggest supporter and making our time in Edmonton a wonderful adventure.

Table of Contents

List of Tables

List of Figures

List of Symbols, Nomenclature, or Abbreviations

aa	Amino acid
AFM	Atomic force microscopy
AOD	Acousto-optic deflector
bp	Base pair
CD	Circular dichroism
DTDP	Dithiodipyridine
DFS	Dynamic force spectroscopy
dsDNA	Double stranded deoxyribonucleic acid
EM	Electron microscope
EOD	Electro-optic deflector
FCS	Fluorescence correlation spectroscopy
FEC	Force-extension curve
FPLC	Fast protein liquid chromatography
FRET	Förster resonance energy transfer
FTIR	Fourier transform infrared spectroscopy
IDP	Intrinsically-disordered protein
LB	Lewy body
MD	Molecular dynamics
MWCO	Molecular weight cut-off
NA	Numerical aperture
NAC	Non-Aβ component of Alzheimer's disease amyloid
NMR	Nuclear magnetic resonance
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PD	Parkinson's disease
PEG	Polyethylene glycol
pI	Isoelectric point
PSD	Position sensitive detector
RNA	Ribonucleic acid
SAXS	Small-angle x-ray scattering
SDS	Sodium dodecyl sulfate
SMFS	Single-molecule force spectroscopy
SNCA	α-synuclein gene
TEM	Transmission electron microscope
ThT	Thioflavin T
WLC	Worm-like chain
WT	Wild-type

F	force
p	photon momentum
t	time
n	refractive index
P	power of incident light
c	speed of light
R_F	Fresnel reflection coefficient
T_F	Fresnel transmission coefficient
θ	incidence angle
φ	refraction angle

$\hat{\imath}_{scat}$	scattering force unit vector
$\hat{\jmath}_{grad}$	gradient force unit vector
λ	wavelength
I	laser intensity
σ	particle scattering cross section
a	bead radius
E	electric field
α	polarizability
V	harmonic potential
κ_{trap}	trap stiffness
f	frequency of applied acoustic ultrasound wave
v	speed of acoustic wave
k_B	Boltzmann constant
T	temperature
Lc	contour length
Lp	persistence length
K	elastic modulus
ΔG	free energy difference
ΔG^0	free energy difference at zero force
$\Delta G^{\ddagger}$	height of the energy barriers
$\Delta G^{\ddagger,0}$	height of the energy barriers at zero force
$\Delta x^{\ddagger}$	distance to the transition state
k	Kramers rate
k_0	Kramers pre-factor
D	diffusion coefficient
κ_{well}	curvature of potential well
$\kappa_{barrier}$	curvature of energy barrier
τ_{unf}	reconfiguration time in the unfolded state
W	non-equilibrium work
$P_U(W)$	distribution of non-equilibrium work during unfolding
$P_R(W)$	distribution of non-equilibrium work during refolding
$k_{fold/unf}$	folding/unfolding rate at zero force
R	loading rate
$p(F)$	probability distribution of unfolding forces
$P(f)$	power spectrum
f_0	roll-off frequency
β	effective drag coefficient
η	viscosity
H	height of bead above surface
β'	effective drag coefficient with Faxen's Law correction
F_{unf}	unfolding force
$F_{1/2}$	equilibrium force in a two-state system
$x_H(F)$	extension of the handle
$x_{PU}(F)$	extension of fully-unfolded part of polypeptide chain
$g(\tau)$	autocorrelation of the molecular extension with delay time, τ
P_u	occupancy of the unfolded state
P_f	occupancy of the folded state
$\Delta G(x)_{stretch}$	energy required to stretch the unfolded protein at $F_{1/2}$

Bibliography of Published Papers from PhD Tenure

Solanki, A.*, Neupane, K.*, Woodside, M.T. (2014) Single-molecule force spectroscopy of rapidly-fluctuating, marginally-stable structures in the intrinsically-disordered protein α-synuclein. *Physical Review Letters* 112(15): 158103. * Co-first authors.

Neupane, K.*, **Solanki, A.***, Sosova, I., Belov, M., Woodside, M.T. (2014) Diverse metastable structures formed by small oligomers of α-synuclein probed by force spectroscopy. *PLoS One* 9(1): e86495. *Co-first authors.

Yu, H., Liu, X., Gupta, A.N., Brigley, A., **Solanki, A.**, Sosova, I., Woodside, M.T. (2012) Direct observation of misfolding transitions in single prion protein molecules. *Proceedings of the National Academy of Sciences USA* 109:5283-5288.

Chapter 1. Intrinsically Disordered Proteins and Amyloid Disease

1.1 Protein structure and function

Proteins are a ubiquitous component of every living organism's molecular machinery. They regulate, protect, and serve crucial roles in all biological processes, and their abundance in cells (~18% of total cellular weight in mammalian cells) indicates their utility (Alberts 2008). *In vivo*, proteins are encoded by mRNA and translated by the ribosome, where they start their life as a long polypeptide chain made of unique amino acid sequences. Upon exiting the ribosome, and often with the help of other chaperones or co-factors, most proteins cooperatively fold into a native, minimal-energy conformation. These structures are stabilized primarily by weak, non-covalent bonds (e.g., van der Waals interactions). Dating back to the 1950's, biochemists, most notably Christian Anfinsen and colleagues, performed experiments on proteins (e.g., ribonuclease A) that lead to the theory that a polypeptide's sequence and stable tertiary structure dictates its function (Anfinsen and Redfield 1956; Sela, Anfinsen et al. 1957). This prediction soon became adopted as a biological truth.

However, in the past couple of decades, it had become apparent that more and more proteins were actually 'intrinsically disordered'. These proteins were defined by their lack of a native tertiary structures, and instead were found to continuously sample a host of marginally-stable disordered confirmations (Dunker, Obradovic et al. 2000). Despite their structural plasticity, IDPs have been shown to serve imperative roles in key functions throughout our biological

machinery, from cell-cycle control to transcriptional and translational regulation (Tompa 2002; Dyson and Wright 2005). Further, their flexibility allows IDPs to associate with diverse chaperones, surfaces or other biological molecules, all of which may enable additional functionalities. This was a direct challenge to the principles laid out by Anfinsen and colleagues, and exposed a new chameleon-like protein-function phenomenon.

1.2 Intrinsically-disordered proteins

Unsurprisingly, IDPs have gained significant notoriety since they were first identified several decades ago, and their abundance has made them an especially compelling subgroup to study. In fact, a relatively recent 2007 survey of the Protein Data Bank found that ~10% of all entries contained long regions of disorder (>30 consecutive amino acids (aa)), while ~40% possessed shorter regions of disorder (10-30 aa) (Le Gall, Romero et al. 2007). The initial realization that proteins may be intrinsically disordered came more than two decades ago when scientists found that certain proteins were unable to form crystal structures necessary for X-ray crystallography analysis (Sedzik and Kirschner 1992). Under further investigation, circular dichroism (CD) spectra analysis found that these proteins also generally lacked a stable secondary structure (i.e., α-helices, β-sheets) under physiological conditions (Tompa 2010). Nuclear magnetic resonance (NMR), small-angle x-ray scattering (SAXS), and Fourier transform infrared spectroscopy (FTIR) later confirmed that a variety of proteins with clear functional roles (e.g., MAP2, tau, MBP) contained significant regions of disorder (Tompa 2010).

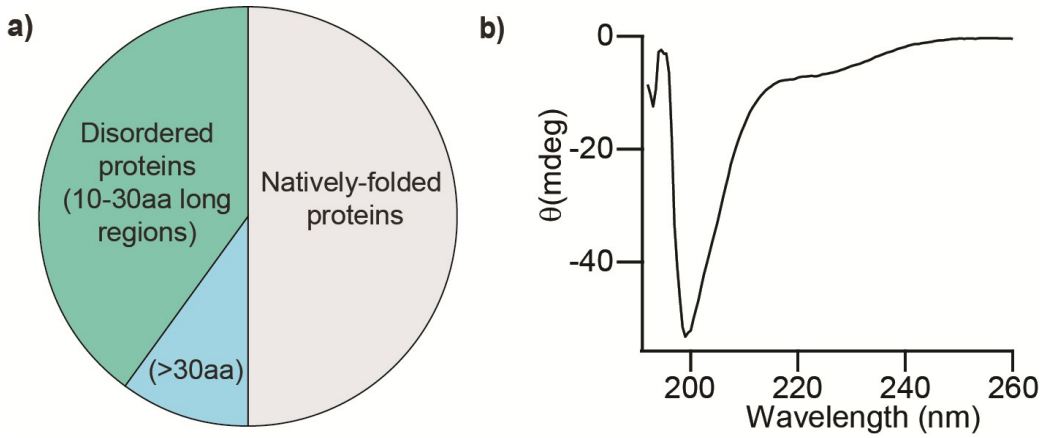

Figure 1.1 IDP pie chart and CD spectrum. a) Approximately half of all proteins registered in the Protein Data Bank contain disordered sequence regions (~40% contain short stretches (10-30 aa) and ~10% contain longer regions (>30 aa)). b) CD spectrum of an α-synuclein monomer sample, illustrating the typical 'random coil' motif that IDPs display in bulk secondary structure measurements.

Given the vast number of proteins that are identified as IDPs (or contain large regions of disorder), it was postulated that there may be sequence similarities that link together members of this proteome subgroup. However, this was not found to be the case. Despite their lack of sequence homology, most IDPs do share some similarities, including a general lack of cysteine and hydrophobic residues and a prevalence of polar amino acids (Uversky 2013). In fact, molecular dynamics simulations have investigated the polyampholytic nature of IDPs and correlated different distributions of oppositely charged residues to the varying levels of disorder in specific IDPs (Das and Pappu 2013). Regardless of these of these patterns, it is still extremely difficult to predict whether a protein is

intrinsically disordered based solely on sequence. Although there has been significant energy devoted to developing IDP prediction programs, they have ultimately proven to be unreliable thus far (Ferron, Longhi et al. 2006).

1.3 Amyloid diseases

As noted, the flexibility and general lack of stable structures in IDPs may be advantageous for their chameleon-like functionality, allowing them to serve multiple roles depending on specific co-factors (Uversky 2003; Dunker 2013). On the other hand, conformational disorder also seems to heighten a protein's susceptibility to misfold and aggregate. Indeed, many IDPs are involved in aggregation-related diseases, including Huntington's disease (huntingtin) and Parkinson's disease (α-synuclein) (Uversky, Oldfield et al. 2008). These are aptly-named 'amyloid diseases', due to the presence of amyloid plaques composed of highly-ordered protein assemblies in hallmark cross-β sheets found in many patients (Eisenberg and Jucker 2012). Cross-β formations, such that the β strands are oriented parallel to the trunk of the fibril, were first identified *in vitro* using various structural determination methods including x-ray diffraction and NMR (Eanes and Glenner 1968; Eisenberg and Jucker 2012).

The commonality of these β sheet fibril formations even among a diverse group of proteins is likely due to the involvement of the protein's main chain (Fandrich and Dobson 2002; Dobson 2003). In addition to the presence of similar proteinaceous, protease-resistant masses, there are additional commonalities that link these diseases together. For example, the majority of patients with amyloid

diseases begin to experience symptoms late in life, and the cases are generally sporadic in nature. And despite their prevalence, there are still many unknowns regarding the origins of these diseases, including the catalyst for amyloid formation and whether these large protein tangles are the toxic agents, all of which could prove crucial for identifying preventative treatment options.

1.4. Protein aggregation and amyloid formation

Numerous *in vitro* and *in vivo* assays have shed light on the disease-related protein aggregation mechanism that leads to characteristic amyloid formations (Soto 2003; Uversky and Eliezer 2009). Although the fine details remain elusive, it is evident that the process begins with a relatively slow lag stage. During this time, misfolded monomers (or fragments) interconvert with small, soluble oligomer intermediates (containing two or more monomer units) or other amorphous aggregates. A nucleation-type process is initiated at some point. Nucleation has been proposed to be the rate-limiting step, resulting in a relatively slow lag phase of days to weeks before the exponential growth phase commences (Eisenberg and Jucker 2012). As monomer species are depleted to form larger oligomeric and 'proto-fibrillar' species, a plateau is reached. By this time, highly ordered elongated fibrils have formed (Figure 1.2).

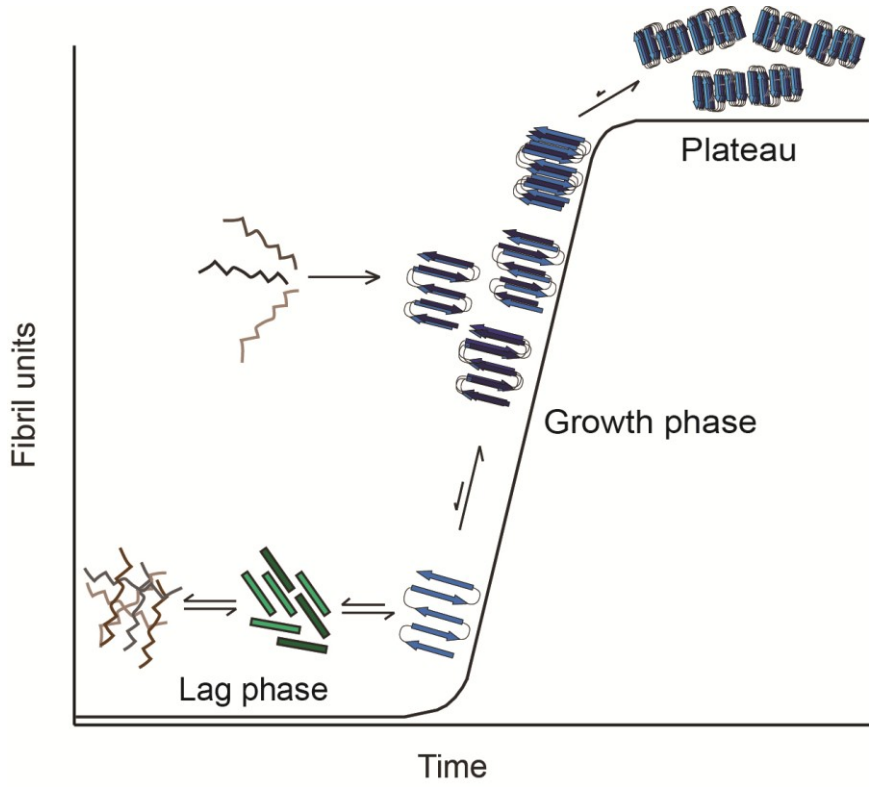

Figure 1.2 Model of characteristic protein aggregation kinetics. The proposed amyloid growth mechanism, which begins with a relatively slow lag phase. Misfolded monomer and oligomer species interconvert during this time. At some point, a nucleation event triggers an accelerated growth phase. Monomer units are depleted to form highly ordered fibril structures.

Specific dyes have been used extensively to track the formation of aggregates as they grow. For example, Thioflavin T fluoresces upon binding to the crossed β-sheets present in amyloids, and has proven particularly useful in determining the duration of lag phases (Ban, Hamada et al. 2003). In addition, the dye Congo Red, which exhibits green birefringence when associated with amyloids, provides an additional tracking tool for monitoring amyloid formation, although some studies indicate the dye itself may induce aggregation (Khurana,

Uversky et al. 2001). Techniques such as X-ray crystallography and solid-state NMR are also valuable for characterizing fibril structures *in vitro*, and more recently, imaging tools like transmission electron microscopy (TEM) and atomic force microscopy (AFM) have also been employed (Chiti and Dobson 2006). Together, they have led to the discovery that individual fibrils can possess significant heterogeneity (Eichner and Radford 2011), including 'straight' and 'twisting' motifs (Vilar, Chou et al. 2008; Breydo, Wu et al. 2012).

The original consensus in the field was that these fibrils and amyloid deposits were the toxic agents of the disease. However, years of studying the aggregated masses of various misfolded proteins have implicated oligomeric states instead, suggesting that amyloids may in fact serve a protective role by sequestering harmful intermediates (Conway, Lee et al. 2000; Dobson 2003). The toxicity of these protein intermediates (or oligomers) have been attributed to many different factors, including the exposure of normally-buried side chains in early stages of misfolding that can abnormally and adversely interact with other local cellular units leading to toxic losses-of-function (Bucciantini, Giannoni et al. 2002), or their ability to induce detrimental interactions with membranes (e.g., causing permeabilazation) (Lashuel, Petre et al. 2002; Lorenzen, Nielsen et al. 2014). Since protein fibrils and plaques are no longer considered to be the suspected culprit in amyloid diseases, there has been a shift in research toward investigating the initial stages of the aggregation. Despite the growing evidence of its importance, very little is actually known about the early aggregation process. This is primarily due to experimental limitations in studying small oligomers,

which are rapidly transient and diverse in nature (Goldberg and Lansbury 2000). Traditional bulk studies have little utility when studying these early aggregation states, since averaging over the ensemble can easily preclude identification of subtle conformational changes. To overcome these limitations, single-molecule tools, including Förster resonance energy transfer (FRET) and single molecule force spectroscopy (SMFS), are now emerging as powerful assets for studying individual protein molecules, providing clues to the highly complex and dynamic nature of potentially crucial aggregation subspecies. In this thesis, we use optical tweezers, an extremely powerful SMFS-based instrument, to study α-synuclein, the IDP linked to Parkinson's disease.

The outline of this thesis is as follows: In Chapter 2, an overview of the IDP α-synuclein is presented, as well as a general discussion of the current experimental techniques available to study the early stages of the aggregation process. The optical tweezers instrument is then introduced in Chapter 3, along with discussion of analysis techniques used in SMFS experiments. These are presented in the context of energy landscape theory. Chapter 4 presents the first optical trapping data on an IDP, comprised of measurements and analysis on the discrete unfolding events identified in monomer, dimer and tetramer α-synuclein constructs. The following chapter 5 explores the characterization of marginally stable states present in these same monomeric and early oligomeric systems. In turn, an IDP energy landscape is reconstructed by new analytical tools we developed and implemented. Lastly, I will summarize our results and present the future direction of this research in Chapter 6.

Chapter 2. α-Synuclein

2.1 Links to Parkinson's disease

PD is the second most common neurodegenerative disease worldwide. The majority (~ 80-90%) of cases are believed to be sporadic (i.e., not hereditary) and occur later in life; in fact, more than 1% of the population over 65 years old is afflicted (Ozansoy and Başak 2013). There are two hallmark events that occur in most PD patients: the loss of dopaminergic neurons and the formation of cytoplasmic amyloid inclusions, dubbed Lewy bodies (LBs). Clinical symptoms range from rigidity and tremors, to bradykinesia and instability, and were first described in 1817 by a surgeon, James Parkinson, for whom the disease is named (Parkinson 2002). Numerous environmental risks have been proposed to trigger PD (Foltynie, Michell et al. 2007), including pesticide (e.g., rotenone) and heavy metal (e.g., lead) exposure, while others appear to be protective (e.g., nicotine) (Sherer, Betarbet et al. 2003; Tsuboi 2012), However, the precise cause is still unknown, and as a result, there are currently no cures for PD. Available medications only suppress symptoms and lose their effectiveness over time.

In 1988, a new protein was identified by Maroteaux and colleagues using antiserum screening experiments (Maroteaux, Campanelli et al. 1988), and was named *syn-nuclein* based on its expression in close proximity to the nuclei in pre-synaptic nerve terminals. Interest in α-synuclein sky-rocketed once a mutation in the gene that codes for α-synuclein was linked to PD in 1997 (Polymeropoulos, Lavedan et al. 1997). Familial mutations may be implicated in as many as 20% of

all PD diagnoses, and are often linked to earlier symptom onset compared to sporadic cases (Ozansoy and Başak 2013). The first genetic mutations linked to the *SNCA* gene were A30P, E46K, and A53T, and this list has continued to expand in recent years (Polymeropoulos, Lavedan et al. 1997; Kruger, Kuhn et al. 1998; Zarranz, Javier Alegre et al. 2004; Alderson and Markley 2013; Appel-Cresswell, Vilarino-Guell et al. 2013). α-Synuclein was also soon identified as the major component in LBs (Spillantini, Schmidt et al. 1997). Since then, various loci in the *SNCA* gene were found to be duplicated and triplicated in certain PD cases, linking α-synuclein concentration to disease (Singleton, Farrer et al. 2003; Chartier-Harlin, Kachergus et al. 2004). As a result, there has been a significant effort to learn more about α-synuclein, particularly since it could play a key role in developing preventative PD treatment options.

α-Synuclein is a member of the IDP family, and is comprised of 140 aa (M.W. ~14.6kDa). Its sequence can be divided into three distinct units, starting with a membrane-binding N-terminal region (residues 1- 60), a relatively hydrophobic region called the non-Aβ component of Alzheimer's disease amyloid (or more commonly, the NAC region, residues 61-95), and a highly acidic C-terminal region (residues 96-140) (Figure 2.1). The overall pI of the α-synuclein monomer is 4.7, giving it a negative charge at buffers with a pH greater than the pI. Six imperfect KTKEGV repeats also span the N-terminal and NAC regions. α-Synuclein belongs to a family of three related proteins, including β-synuclein (134 aa) and γ-synuclein (127 aa), the latter of which has been implicated in breast and ovarian cancers (Lavedan 1998). Notably, β-synuclein lacks 11 amino

acids (residues 74-84) found in the NAC region of α-synuclein and fails to readily form fibrils, indicating a potential role of the NAC region in aggregation (Uversky, Li et al. 2002). The functionality of α-synuclein has long been contentious, but there have been many studies attempting to elucidate its role in the body, most often by utilizing knockout mice experiments. Possible functions range from playing a role in SNARE-mediated vesicle trafficking (Burré, Sharma et al. 2010) to protecting and modulating the release of neurotransmitters (Narkiewicz, Giachin et al. 2014).

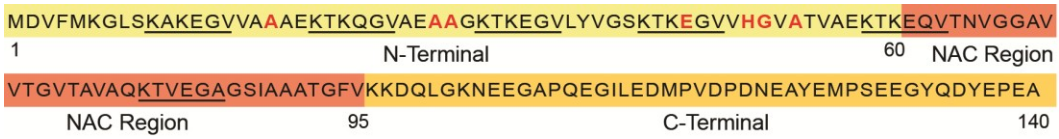

Figure 2.1 Human wild-type α-synuclein sequence. α-Synuclein is a 140 amino acid protein, whose sequence can be divided into three distinct regions: the membrane-binding N-terminal (yellow), the hydrophobic NAC region (red), and the highly acidic, disordered C-terminal (gold). Residues in red font indicate sites of known familial mutations: A18T, A29S, A30P, E46K, H50Q, G51D, A53T, and the six imperfect KTKEGV repeats are underlined.

2.2 IDP bulk structural studies

As mentioned in Chapter 1, traditional bulk techniques used to elucidate native protein conformational behavior are even less fruitful for IDPs due to the inherent diversity and transience of structures present in solution being averaged at any given time during data acquisition (Uversky 2013). CD spectra, which measures the asymmetric absorption of right or left circularly-polarized light and

is analyzed to extract secondary structure information, results in a classic 'random coil' signal for IDPs (Figure 1.1b). Extremely subtle shifts may be observed in the spectra if there are minority populations of secondary or partially folded structures present (Uversky, Li et al. 2001). NMR is also another popular tool to determine both the high-resolution structure and dynamics of proteins in solution. In typical experiments, an active nuclei absorbs electromagnetic radiation at a characteristic radiofrequency dictated by the isotope when placed in a magnetic field, and the resonant frequencies are analyzed to extract distance information. That said, the data is averaged and can mask any present subpopulations. In response to difficulties in studying IDPs *in vitro,* molecular dynamics simulations have also been employed to study proteins like α-synuclein. However, these can be computationally expensive simulations to run, and often theorists must resort to analyzing smaller sequence fragments instead of the entire protein (Sethi, Tian et al. 2012). Using experimentally derived distance or energetic constraints, such as pairwise distance values from smFRET assays, full-length protein computations have become more feasible (Nath, Sammalkorpi et al. 2012).

Despite the drawbacks of ensemble averaging, there has been numerous bulk experiments conducted on α-synuclein in solution, providing an overall macroscopic structural picture of a dynamic protein that is slightly more compact than a traditional random coil. For example, NMR, paramagnetic relaxation enhancement and molecular dynamics simulations were all used to detect the presence of weak long-range interactions in monomer α-synuclein (Bertoncini,

Jung et al. 2005; Dedmon, Lindorff-Larsen et al. 2005). Dedmon and co-workers reported that residues ~30-100 spanning the central protein region interact with the C-terminus (~residues 120-140). Meanwhile, Bertoncini and colleagues identified not only interactions between the C-terminal (residues 110-130) and the NAC region, but also longer range interactions with an N-terminal residue. Interestingly, C-terminal truncation studies have reported an increase in aggregation rates (Liu, Giasson et al. 2005), thus suggesting a role in protecting the hydrophobic NAC region.

Although many experiments (e.g., SAXS, FTIR, CD) indicate that α-synuclein is unstructured overall in solution, other experiments (namely fluorescence and NMR) have shown a propensity for the N-terminus to form α-helices *in vitro* when associated with membranes (e.g., micelles, vesicles). For example, upon binding to small sodium dodecyl sulphate (SDS) micelles, the N-terminus acquires an α-helical motif, consisting of two coils separated by a small spacer region (Chandra, Chen et al. 2003; Georgieva, Ramlall et al. 2010). There is additional evidence that it proceeds to form an extended helix upon binding to larger vesicles (Trexler and Rhoades 2009). The acidic C-terminal remains largely unstructured, regardless of the membrane surface. Despite existing primarily in a free and dynamic cytosolic state *in vivo*, α-synuclein has also been found to interact with presynaptic membranes, promoting similar helical structures (Davidson, Jonas et al. 1998; Jao, Der-Sarkissian et al. 2004; Ulmer, Bax et al. 2005; Pfefferkorn, Jiang et al. 2012). Recent work has proposed that α-synuclein actually exists as a stable helical tetramer *in-vivo* (Bartels, Choi et al. 2011;

Wang, Perovic et al. 2011), but this theory remains controversial in the field (Fauvet, Mbefo et al. 2012; Burré, Vivona et al. 2013).

2.3 α-Synuclein aggregation behavior

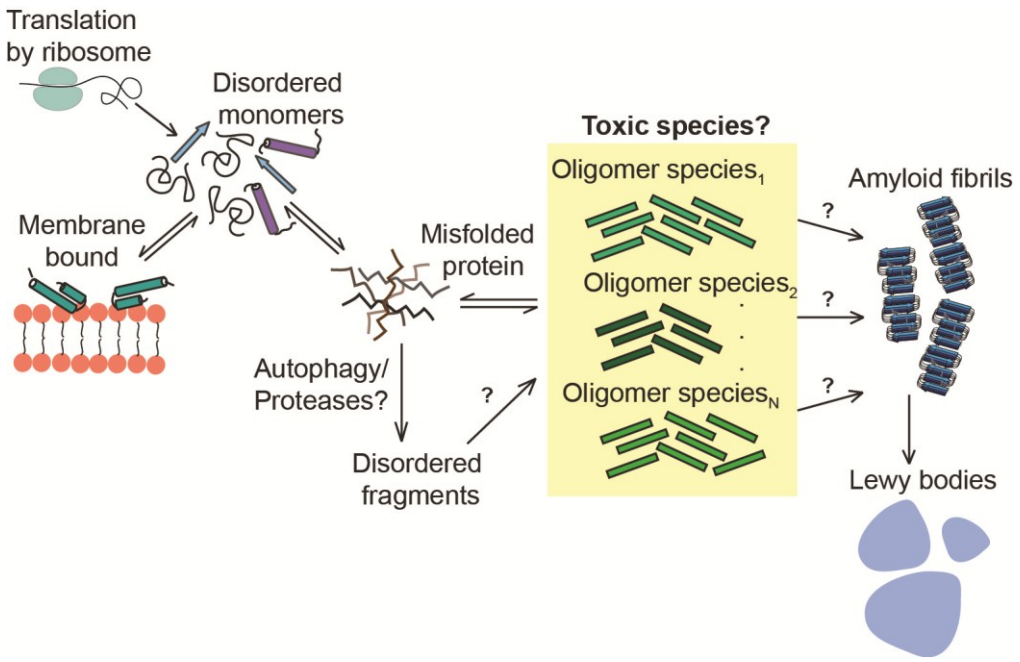

Figure 2.2 Proposed α-synuclein aggregation pathways. Upon exiting the ribosome, the IDP α-synuclein can traverse several different pathways, one or more of which may lead to the disease state. Although there are defence mechanisms in place in case protein misfolding occurs *in vivo*, these can sometimes fail and may lead to the formation of specific oligomeric species, one or more of which can be toxic. These small oligomers have been experimentally shown to coexist with other non-toxic oligomers and monomeric species. It is believed that the disease-related species go on to form hallmark amyloid fibrils and Lewy bodies found in many PD patients.

Understanding the transition from monomer α-synuclein units to ordered amyloid fibrils requires identifying the trigger that initiates aggregation, a process that is still mired in controversy (Kalia, Kalia et al. 2013) (Figure 2.2). Despite various theories concerning the properties of early oligomeric species and their role in the disease, one thing is clear: at some point following translation from the ribosome, α-synuclein can misfold and proceed through a classical aggregation growth mechanism with a lag phase, followed by an accelerated growth period (Figure 1.2). Once a plateau is reached and fibrils of varying sizes and morphologies coexist, a subpopulation(s) then proceeds to grow into larger protease-resistant LBs (Conway, Harper et al. 1998; Lundvig, Lindersson et al. 2005; Fink 2006). The lag time, during which nucleation begins, is on the order of weeks for α-synuclein samples comprised solely of monomer units (i.e., no higher order nucleation 'seeds' are present) in physiological buffers (Krishnan, Chi et al. 2002). This can be reduced to days, under shaking conditions at 37°C (Fink 2006). A minimum concentration of 0.2mg/ml is required to ensure fibril growth (van Raaij, van Gestel et al. 2008), with high starting monomer concentrations also reducing the lag phase significantly (Uversky, Li et al. 2001). This is another implication for over-expression as a factor in α-synuclein disease pathogenesis.

Numerous local environmental factors have been shown to accelerate α-synuclein aggregation *in vitro*, among them low pH, high ionic strength, high temperatures where denaturing typically occurs, and the presence of metals like Cu^{2+} (Lundvig, Lindersson et al. 2005; Uversky and Eliezer 2009; Giehm, Lorenzen et al. 2011). Interestingly, Mg^{2+} has shown to retard aggregation, as did

various other chemical compounds, including dopamine and curcumin (Mazzulli, Mishizen et al. 2006; Ahmad and Lapidus 2012). Environmental conditions have been also linked to different fibril morphologies, each with varying nucleation seeding ability, lag phases, and degrees of cytotoxicity. These differences have generally been attributed to shifts in electrostatic shielding abilities and hydrophobic effects involving the highly acidic C-terminal and hydrophobic central NAC region (Hoyer, Antony et al. 2002). For example, lowering a buffer's pH from neutral to 4.0 can catalyze a more rapid aggregation process, with ThT and turbidity assays recording lag phase reduction from days to minutes (Uversky, Li et al. 2002). Aggregation assays have thus shown the ease at which α-synuclein can aggregate, but identifying the crucial subpopulations involved in the disease-pathway nucleation process is still of supreme interest.

A partially-folded, β-sheeted intermediate was identified in recombinant α-synuclein using various bulk techniques (e.g., CD, FRET) (Uversky, Li et al. 2001), and proposed to seed fibril formation. Given how sensitive monomer α-synuclein is in the presence of different environmental conditions, it was unsurprising that the morphology of fibrils was also highly dependent on the environmental conditions during the growth phase. These studies found that high temperature or low pH conditions promoted shorter, denser fibrils, while neutral pH and physiological temperatures resulted in longer, less dense fibril formations. Bulk fluorescence measurements using tryptophan mutations also found that low-order and transient oligomer species co-exist with fibril populations (Kaylor, Bodner et al. 2005). The diversity of oligomeric species explains the presence of

heterogeneous proto-fibril species. In fact, EM was used to characterize a variety of proto-fibrils that formed larger fibril at markedly different rates (Lashuel, Petre et al. 2002).

Recently, experiments characterized two distinct oligomer populations that may coexist, formed from high concentrations of disordered monomers in physiological buffer (Lorenzen, Nielsen et al. 2014; Paslawski, Mysling et al. 2014). The latter study, using hydrogen-deuterium mass spectroscopy, found that one oligomer population is much more stable (on the order of minutes) than the other - a transient, low concentration oligomer that dissociates readily and exists in equilibrium with monomers in solution. These individual oligomer species were proposed to aggregate along distinct pathways, with the short lived oligomer able to form fibrils, while the more stable oligomer formed more amorphous clusters. These results complement older extensive studies on α-synuclein fibril morphologies, which identified two distinct subpopulations of similar length and widths (Fink 2006; Vilar, Chou et al. 2008).

Using classic biochemical and imaging techniques, a picture of the α-synuclein aggregation mechanism has emerged. The subtleties in the early oligomeric stages are still missing, due to the inherent difficulty of identifying conformational changes and transient oligomeric species in bulk. It has thus been essential to develop experimental assays that can probe the short lived and sparsely populated individual intermediate states in α-synuclein.

2.4 Single-molecule studies of α-synuclein

Throughout the past decade, increasingly sophisticated instrumentation has provided the impetus and technical feasibility for conducting single-molecule experiments. As the name suggests, single-molecule tools grant the ability to characterize individual molecules of interest, thus avoiding the drawbacks of ensemble averaging. Force and fluorescence experiments are most commonly employed to obtain information about structural dynamics, including distance information and folding kinetics (Figure 2.3). As with all experimental methods, each technique provides unique advantages and disadvantages. Yet, taken together they can provide useful information on disordered proteins like α-synuclein, characterizing the monomer state as well as the evolving aggregation process.

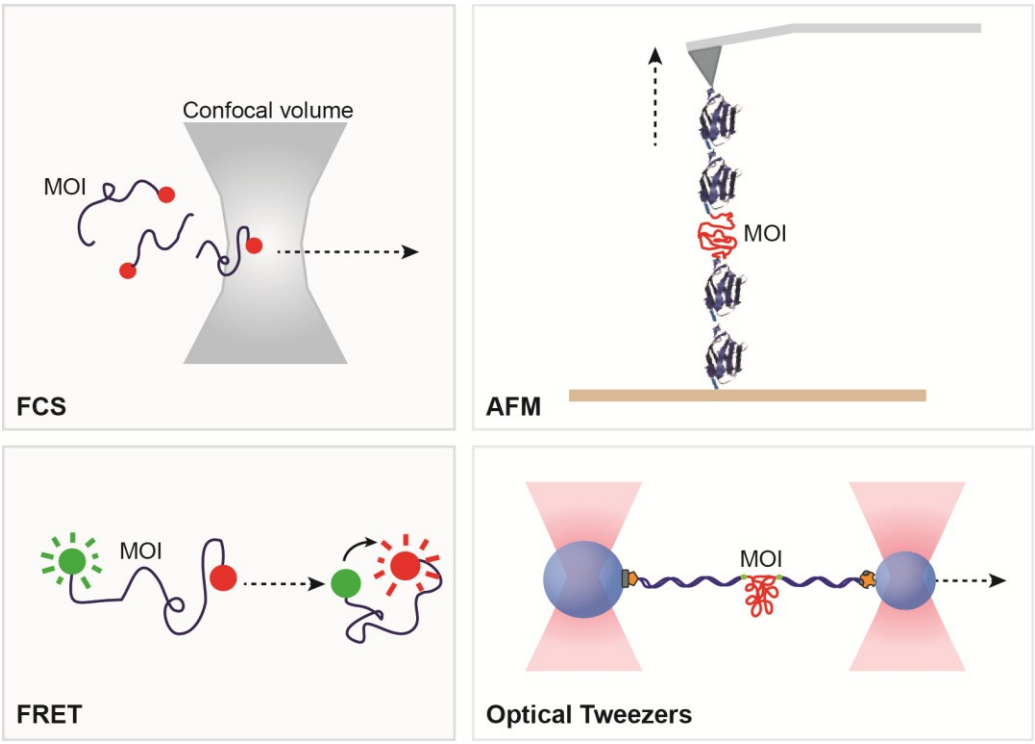

Figure 2.3 Summary of single-molecule techniques. Schematics of the most commonly- used single-molecule techniques to study α-synuclein, the molecule of interest (MOI): fluorescence based (Förster resonant energy transfer (FRET), fluorescence correlation spectroscopy (FCS)) and force based (AFM and optical tweezers). (Note: objects are not to scale).

2.4.1 Single-molecule fluorescence experiments

Two types of fluorescence experiments are commonly used to study single α-synuclein molecules and observe their aggregation process: Förster resonant energy transfer (FRET) and fluorescence correlation spectroscopy (FCS). Both involve attaching fluorescent dye molecules to specific regions on the protein of interest, then exciting them in different sample environments to extract dynamic structural information.

As the name suggests, FCS uses correlation analysis of the fluorescence fluctuations in low concentrations of dye-modified proteins as they are energetically excited and diffuse through a femtolitre confocal volume. The sample volume ensures that a minimum number of molecules are fluoresced at a time. Since it was first used over 40 years ago, FCS has proven to be a useful tool, providing kinetic details (e.g., diffusion rates) and structural information (e.g., identifying local aggregates) on numerous biomolecules (Rajagopalan, Huang et al. 2011; Paredes, Casares et al. 2012). More recently, FCS has been employed in numerous protein aggregation studies (Tjernberg, Pramanik et al. 1999; Nath, Meuvis et al. 2010). In these experiments, high concentrations of unlabeled proteins are mixed with low concentrations of dye-labeled protein, and over time, oligomer formation is observed. This is correlated with an increase in recorded diffusion times. For example, Nath and co-workers observed the formation of early oligomeric states in α-synuclein aggregation as monomer species disappeared, and concluded that the overall aggregation process was highly concentration dependent (Nath, Meuvis et al. 2010). It should be noted that while oligomer formation is observed, diffusion coefficients and oligomer sizes are averaged and produce a broad distribution due to FCS sampling rate limitations (Nath, Meuvis et al. 2010).

In contrast, FRET experiments involve labeling the protein of interest with two different fluorescent dyes (an acceptor and a donor), chosen to ensure their absorption and emission spectra overlap (Roy, Hohng et al. 2008). Interactions between two protein species can also be identified if one is labeled with a donor

dye and one with the acceptor dye. As the fluorophores are brought closer together or farther apart, the energy transfer efficiency between the two changes and is recorded as a ratio of acceptor to donor emission intensities. Depending on whether the two dyes are placed on the same molecule or on different molecules, intra- or inter-molecular conformational information and kinetics are obtained. Single-molecule FRET (smFRET) is achieved by lowering sample concentrations to the picomolar range. This technique has proven to be incredibly useful for studying protein folding and aggregation kinetics due to its ability to observe conformational changes on timescales as short as microseconds (Michalet, Weiss et al. 2006; Gambin and Deniz 2010).

smFRET studies on α-synuclein monomers and aggregates have both confirmed and built upon many prior bulk experiment results. For example, studies reestablished that membrane-protein interactions induced a N-terminal helical structural propensity, but also identified a rearrangement of the overall helical motif when SDS concentrations were increased (Ferreon, Gambin et al. 2009). As the detergent evolved from monomer to spherical to extended micelle formations, a transition from a broken helix to an extended helix formation was identified, each weakly bound to the membrane surfaces. smFRET studies of free α-synuclein monomers in different solution environments have also contributed to the notion that the IDP structural disorder can be tuned. In the presence of an osmolyte identified as an oligomerization promoter, the protein became more compact, while the commonly-used denaturant urea had the opposite effect (Ferreon, Moosa et al. 2012). Aggregation studies have also identified the

simultaneous presence of both soluble and compact, protease resistant oligomer populations (Cremades, Cohen et al. 2012). In fact, Cremades *et al.* found that insoluble oligomers possessed a longer lag time, which coincided with the generation of higher levels of reactive oxygen species and eventual fibril formation. Once these highly-ordered fibrils were present in the sample, the aggregation rate rapidly increased. These fluorescence based studies have provided a clear platform for investigating population shifts along the aggregation pathway. However, it should be kept in mind that, as with FCS experiments, there are some inherent limitations in smFRET data interpretations. These may include restrictions on spatial resolution, sample photobleaching concerns, as well as non-trivial dye attachments (Roy, Hohng et al. 2008).

2.4.2 SMFS experiments

In contrast to fluorescence-based single molecule assays, SMFS employs force to study the mechanical properties of the molecule of interest. This force is analogous to other *in vitro* experiments that use artificial denaturants (such as temperature, pH, pressure and chemicals such as urea), but here the denaturant is applied to the molecule of interest in a highly controlled geometry along the pulling axis. Examples of SMFS instruments include optical tweezers, magnetic tweezers and AFMs, which have probed everything from molecular motors to DNA and RNA polymerases (Neuman and Nagy 2008). In addition to studying individual protein molecules, SMFS is also an excellent tool for probing aggregation by studying tandem and/or oligomeric constructs. Even in multi-state systems, the conformations probed, as well as their rates and energies, can all be

described quantitatively (Borgia, Williams et al. 2008; Hoffmann and Woodside 2011; Stigler, Ziegler et al. 2011; Yu, Liu et al. 2012). Using this method, even structures that form very briefly or very rarely can be characterized (Yu, Liu et al. 2012). These qualities enable a very detailed picture of the events occurring during aggregation to be drawn.

AFM is the most commonly used SMFS instrument for studying α-synuclein. The typical experimental set-up is comprised of the molecule of interest, which is attached to a stiff cantilever probe (~10-100,000pN/nm), often via DNA or PEG tethers at one terminal, and to a functionalized (e.g., Au, Ni-NTA) sample surface on the opposite end. By moving the cantilever position, the protein can experience large denaturing forces. This may result in the presence of 'rips' on measured distance (of cantilever above the sample surface) versus force curves, and are attributed to the unfolding of stable proteins tructures. There have been only a handful of studies on both α-synuclein monomers and dimers, all of which have identified a relatively heterogeneous conformational landscape, with states possessing various stabilities and prevalence. In AFM SMFS experiments on monomeric α-synuclein, in which the protein of interest is placed in between tandem repeats of other well-characterized and stable natively folded protein (e.g., ubiquitin, titin), data was analyzed by fitting the unfolding of the stable protein and attributing the remaining unfolding events to the IDP behavior (Sandal, Valle et al. 2008; Hervas, Oroz et al. 2012).

Three distinct monomer structural subpopulations were identified in the work of Sandal *et al.*: mechanically weak interactions, random coil, and 'β-like',

assigned based on the distance of the unfolding events. The latter was attributed to aggregation-prone states, whose presence during unfolding events increased upon introducing a metal ion (Cu^{2+}), familial mutation (A30P), or increasing the ionic strength of the buffer (from 10mM to 500mM Tris/HCl, pH 7.5). These experiments complimented earlier bulk aggregation studies in similar conditions. In contrast, Hervas and colleagues found a much more diverse conformational landscape when they investigated both wild-type (WT) α-synuclein and two familial mutations (A30P, A53T). In their assay, a monomer α-synuclein protein was inserted into a specially designed plasmid, directly in the centre of a ubiquitin protein, and flanked by more ubiquitin proteins. Contour length changes ranging from a few nm up to ~60nm and unfolding forces measuring as much as 400pN were observed. Probing such high unfolding forces indicate the presence of extremely stable α-synuclein structures in solution, which is contradictory to the inherent instability of intrinsically disordered monomer units. It is likely that some of the observed interactions were due to α-synuclein interacting with nearby ubiquitin units or to the sample and/or cantilever surface, an assay-related issue also present the work of Sandel *et al*..

In AFM dimer studies, in which a single α-synuclein molecule is brought into contact with another monomer and then cantilever is retracted, two distinct states were observed at low pH (Yu, Malkova et al. 2008). Each possessed slightly different lifetimes on the order of seconds, but disappeared completely at neutral pH. The only unfolding events recorded at pH 7.0 were in the presence of metal ions (e.g., Al^{3+} and Zn^{2+}), dopamine, polyamines and following the

introduction of familial mutations (Yu, Warnke et al. 2011; Krasnoslobodtsev, Peng et al. 2012). By fitting plots of rupture force versus rupture length, population shifts in the presence of various chaperones and mutations were recorded, including increased intermolecular interactions between the two C-terminals. This was proposed to involve β-sheet formation based on the extracted contour length changes. Like the monomer studies, unfolding forces were quite high, with most events recording 50-100pN, but occasionally even as high as 250pN.

Conversely, we employed a high-resolution optical tweezers instrument to study both monomers and early oligomeric species of α-synuclein. Although there are numerous optical tweezers studies characterizing the behavior of natively – folded proteins (Gebhardt, Bornschlögl et al. 2010; Shank, Cecconi et al. 2010; Stigler, Ziegler et al. 2011; Yu, Liu et al. 2012), there have been no reported experiments on IDPs. One advantage optical tweezers hold over AFMs is their ability to probe a much lower force regime (down to 0.1 pN versus ~10 pN for an AFM). This is especially useful for studying the relatively unstable conformations present during the initial stages of oligomer formation. Additionally, optical tweezers can apply denaturing forces to proteins in a specific, well-defined geometry, void of potential surface and probe interactions that can complicate the interpretations of results. Optical tweezers, from their basic principles to data analysis, will be explored extensively in the following chapter.

Chapter 3. Optical Trapping

Ever since Arthur Ashkin and his coworkers at Bell Labs demonstrated the first optical trap over forty years ago (Ashkin 1970; Ashkin, Dziedzic et al. 1986), the technique has been exploited in the field of physics and biology. In brief, they found that a tightly focused laser beam forms a three-dimensional potential well that could trap a small dielectric object with a high refractive index (compared to the trapping medium). The object could then be moved around by manipulating the incoming beam. Over the years, numerous modifications and improvements have led to the sophisticated optical traps employed in biophysical experiments today. Optical trapping experiments are especially valuable in their ability to provide data that can be used to extract precise kinetic information and understand complex energy landscapes for a diverse set of biomolecules (Abbondanzieri, Greenleaf et al. 2005; Greenleaf, Frieda et al. 2008; Yu, Gupta et al. 2012). In this chapter, I will describe the basic physics behind the optical trap and detail the components that are essential for building an instrument with high spatial and temporal resolution. I will conclude by giving an overview of the basic experimental procedures we used to gather data on our α-synuclein protein constructs, and detail how various analytical methods may be employed to validate the model proposed for an IDP energy landscape.

3.1 Basics of optical trapping

The theory behind optical trapping can be first understood in terms of the Mie scattering regime, when the scattering object is much larger (at least 10x)

than the λ_{laser}. As the highly focused trapping beam interacts with a micron-sized dielectric bead and is deflected, it experiences a change in momentum. Recall that photon momentum (p) is related to the wavelength of light by $p = h/\lambda$, where h is Planck's constant. According to Newton's 2nd law, this necessitates that the bead will in turn experience an equal, yet opposite change in momentum, which ultimately pushes the bead back to the trap's centre, such that

$$\vec{F} \approx \frac{\Delta \vec{p}}{\Delta t}\,. \tag{3.1}$$

Mie scattering can also be described qualitatively by ray optics (Figure 3.1a). More specifically, each light ray imparts a force $\vec{F}$ on the bead, such that:

$$\vec{F} = \frac{n_m P}{c}\left(1 + R\cos 2\theta - \frac{{T_F}^2[\sin(2\theta - 2\varphi) + R_F \cos 2\theta]}{1 + {R_F}^2 + 2R_F \cos 2\varphi}\right)\hat{\imath}_{scat} + \frac{n_m P}{c}\left(R\sin 2\theta - \frac{{T_F}^2[\sin(2\theta - 2\varphi) + R_F \cos 2\theta]}{1 + {R_F}^2 + 2R_F \cos 2\varphi}\right)\hat{\jmath}_{grad.} \tag{3.2}$$

R and T are the Fresnel reflection and transmission coefficients, respectively, n_m is the index of refraction of the trapping medium, P is the power of the ray incident at angle θ, φ is the angle of refraction, and c is the speed of light (Svoboda and Block 1994). The unit vectors $\hat{\imath}_{scat}$ and $\hat{\jmath}_{grad}$ correspond to the scattering force and gradient force components, respectively.

Now let us consider another trapping regime, in which the scattering objects are much smaller than λ_{laser}. This is referred to as Rayleigh scattering. In this case, the dielectric bead that is held in the trap behaves like a point dipole in

an inhomogeneous electric field (Figure 3.1b). The scattering force can be described as:

$$F_{scat} = \frac{I\sigma n_m}{c}, \tag{3.3}$$

where I is the laser intensity and σ is the particle scattering cross section (Neuman and Block 2004). For a sphere of radius a,

$$\sigma = \frac{128}{3}\frac{\pi^5 a^6}{\lambda^4}\left(\frac{m^2-1}{m^2+2}\right), \text{where } m = \frac{n_p}{n_m}. \tag{3.4}$$

A Lorentz force on the dipole exerts a restoring gradient force on the dipole, which is represented by:

$$\vec{F}_{grad} = (\vec{p}\cdot\vec{\nabla})\vec{E}, \tag{3.5}$$

where $\vec{p} = \alpha\vec{E}$ and α is polarizability. For the time average of the force, the restoring force is proportional to the intensity gradient (Neuman and Block 2004; Woodside and Valentine 2009):

$$\langle\vec{F}_{grad}\rangle = \langle(\vec{p}\cdot\vec{\nabla})\vec{E}\rangle = \frac{2\pi n_m}{c}\left(\frac{m^2-1}{m^2+2}\right)a^3\vec{\nabla}I_0. \tag{3.6}$$

The dielectric beads are thus chosen such that a) they are small enough that the scattering force does not overcome the gradient force (since a scales to the power of 6 and 3 for scattering and gradient forces, respectively), and b) they are large enough to compensate for the motion of the bead due to the thermal energy of the system k_BT, ensuring that the bead does not hop out of the trap.

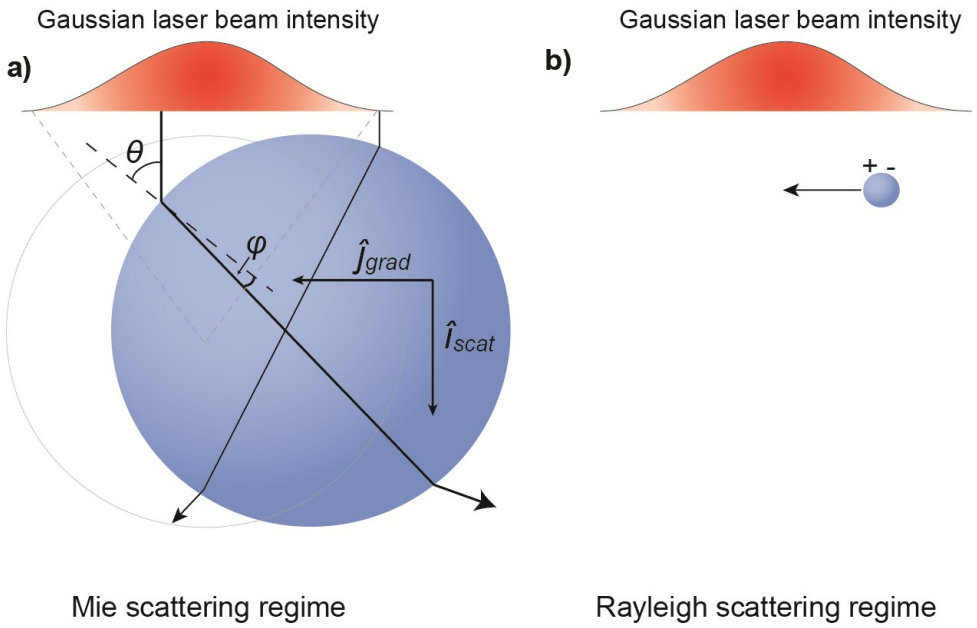

Figure 3.1 Basic principles of optical trapping. a) In the Mie scattering regime, a bead much larger than the wavelength of laser light scatters the incoming beam. This results in a restoring force, as dictated by conservation of momentum. θ is the incident angle, φ is the angle of refraction, and $\hat{\imath}_{scat}$ and $\hat{\jmath}_{grad}$ indicate the unit vectors for scattering force and gradient force components, respectively. b) The bead acts as a point dipole in the Rayleigh trapping regime, where it is subjected to a Lorentz force in the presence of the laser light. The balancing of scattering and restoring forces traps the bead. (Note: objects are not precisely to scale).

It should be noted that, in reality, the dielectric beads are similar in size to the 1064nm trapping laser ($a \sim \lambda$), such that system does not actually behave precisely in either the Mie or Rayleigh scattering regimes. That said, they adequately describe the behavior of the optical trap (Woodside and Valentine 2009) and render it unnecessary to solve the electromagnetic theory required to provide a completely accurate description of the system at hand (Almaas and Brevik 1995; Rohrbach and Stelzer 2002).

3.2 Optical tweezer components

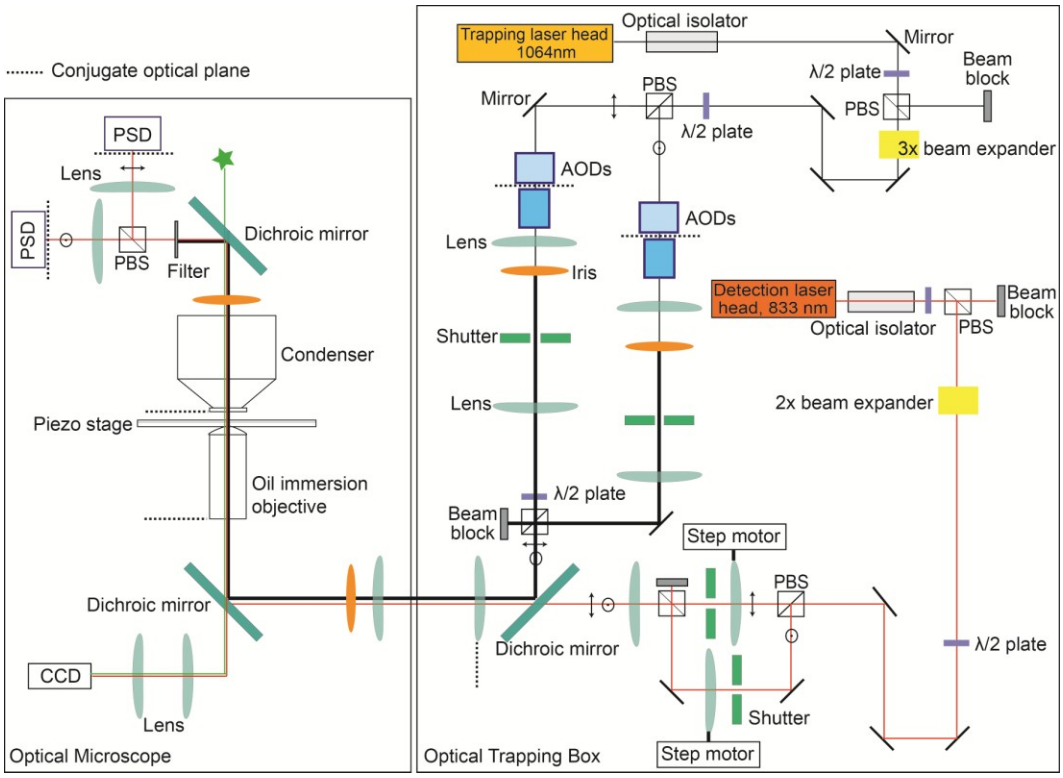

Figure 3.2 Schematic of optical tweezers system. Layout of the dual beam optical tweezers instrument, comprised of the following main components: two orthogonally-polarized trapping beams, originating from a 1064nm laser (black path), that form the two traps; two position sensitive detectors (PSDs) that determine bead positions by collecting light from orthogonally-polarized 833nm detection laser beams that scatter off of the trapped beads; and an inverted microscope where the sample slide is held. Labels in figure: PBS, polarizing beam splitter; AOD, acousto-optic deflector; CCD, charged-coupled device.

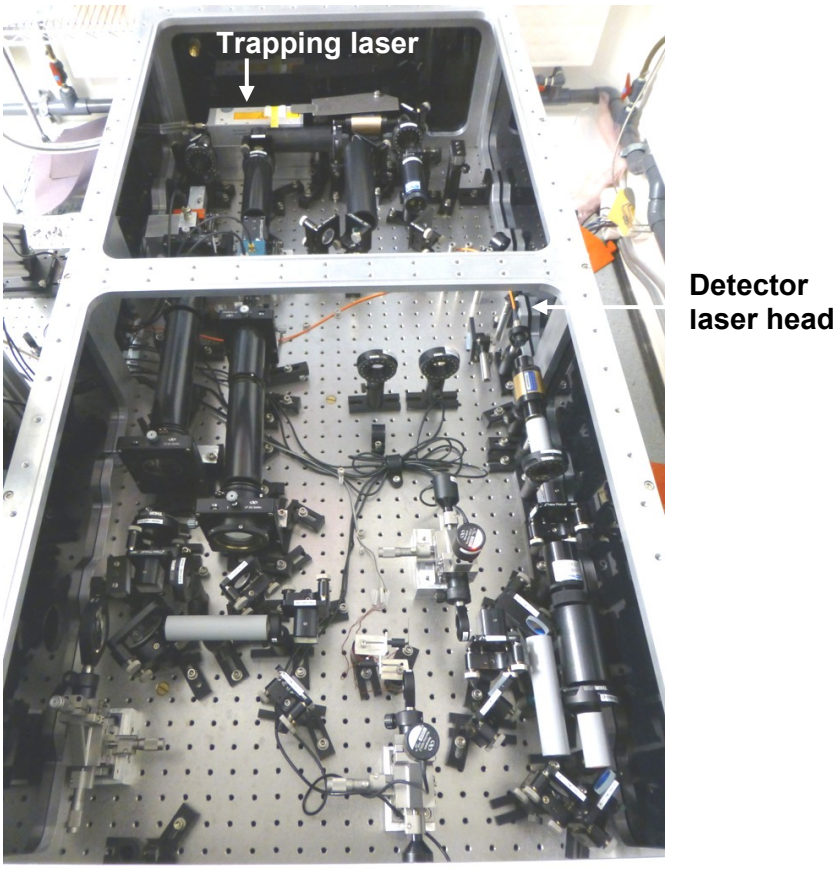

Figure 3.3 Picture of optical tweezers instrument. Photograph of the instrument components contained in the isolated optical box.

The dual-beam optical trap used in the experiments presented here (Figures 3.2 and 3.3) is created using two orthogonally-polarized laser beams from the same 5W, 1064-nm Nd:YVO_4 laser (Spectra Physics). An infrared laser is selected in part due to its relative biological transparency between ~750-1200nm, as to not damage the molecules of interest. Single TEM_{00} Gaussian beam profiles are used to create a harmonic trap. The intensity gradient is aligned with the center of the bead, such that the harmonic potential is $V = (1/2)\kappa_{trap}x^2$, where κ_{trap} is the trap stiffness and x is the bead displacement from trap centre. For a Gaussian beam, a linear region is observed near the trap center, where the bead's

position is linear to the applied force. In order to achieve the high gradient forces discussed earlier, the optical trap employs a collimated laser beam tightly focused to a diffraction limited spot by an objective with a high numerical aperture (NA ~ 1.35), which is housed in an inverted microscope (Nikon). The lens is immersed in transparent oil (n = 1.515), and was chosen as to minimize the refractive index mismatch with the sample solution while also maximizing the gradient near the focal surface (on the sample slide).

We use acousto-optic deflectors (AODs, Gooch-Housego) to steer the trapping beams. Made out of TeO_2 crystal, AODs modulate the diffracted light beam in response to a changing applied acoustic ultrasound wave frequency, f, such that the angle of the first order deflection $\theta = \lambda f / v$. Here, v is the speed of the acoustic wave and λ is the wavelength of the incoming laser beam. In contrast to electro-optic deflectors (EODs), which are also often used in optical traps, AODs exhibit fast response times and possess a shorter path length, allowing for easier alignment. However, they have a non-linear response, dissipate more heat and require careful alignment (Neuman and Block 2004). The sample slide is placed in the specimen plane, while the AODs are placed in planes conjugate to the back focal plane, so that deflections of the beam cause translations in the specimen plan (Woodside and Valentine 2009). To deflect both polarized beams (in the x- and y-axis), the AODs for each trap are placed orthogonally to each other.

Keplerian telescopes serve as beam expanders, ensuring that trap stiffness is maximized at the focal plane by slightly overfilling the objective lens (Neuman

and Block 2004). This stiffness, κ_{trap}, is directly related to recorded force by the expression $F=-\kappa_{trap}x$, where x is the displacement from the trapped position. Accurate determination of position is critical for achieving accurate displacement values, which are directly used to determine the applied force on the molecule. Bead positions within each trap are measured separately by collecting the light from two orthogonally-polarized 833-nm laser (Melles Griot) beams. The light scattered by the individual beads is split by polarization and then collected onto independent position sensitive detectors (PSDs, Pacific Silicon Sensors). The low operating power and different wavelength of the detection laser are chosen as to not interfere with the trapping potentials on the slide.

3.3 Force- ramp experiments

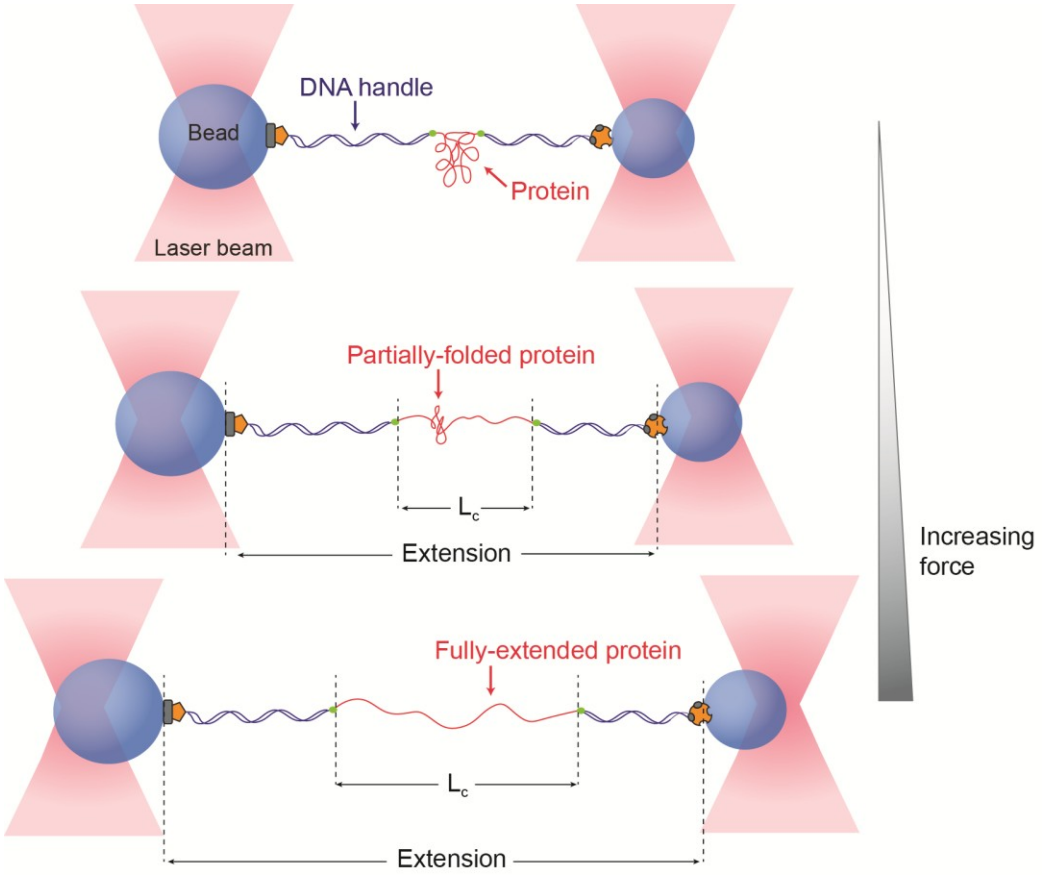

Figure 3.4 Schematic of dual-beam optical tweezers system and force-ramp experiments. The protein-dsDNA handle dumbbell construct is attached to functionalized polystyrene beads at each DNA terminal, which are held in the dual beam optical trap. As the traps are moved apart at a constant velocity, a denaturing force is applied to the protein, and the extension is recorded to generate force-extension curves (FECs).

Optical tweezers apply a well-defined force to the molecule of interest, effectively denaturing it in a controlled geometry. This is particularly suited to studying a protein like α-synuclein, since traditional denaturants (e.g., temperature, pH) have been shown to significantly perturb its conformational

landscape and aggregation properties (see Chapter 2), thus potentially masking physiological-like structural properties. In most SMFS experiments, force is applied to the molecule of interest via polystyrene beads, one held in each of the two traps. Two different bead sizes are selected for easy visual verification in the microscope. These beads are in turn bound to dsDNA 'handles', which serve as spacers between the biomolecule and bead. This eliminates both potential interactions between the molecule and beads, as well as prevents crosstalk between the two traps. dsDNA has been used extensively in SMFS in part because its behavior has been thoroughly characterized (Bustamante, Smith et al. 2000; Williams and Rouzina 2002). Both antibody-antigen and high affinity protein-protein partners are employed in the handle-bead attachments, while the protein-handle attachment utilizes disulfide bond chemistry. These attachment protocols will be explored in greater detail in Chapter 4.

The data presented in this thesis was taken using one particular type of SMFS measurement: the dynamic force experiment. In this procedure, force is applied to the protein by moving the traps apart at a constant speed (~100-1000 nm/s) and then returning them to their original position, allowing the biomolecule to relax (refold). These force ramps are repeated hundreds of times on a single molecule, while changing various measurement conditions, including the wait time at zero force (generally between 0-30s) and pulling speeds (~100 – 500 nm/s).The extension of the molecule, as well as the force, is determined by analyzing the displacement of the bead position in the traps as they move from their equilibrium positions (see Appendix: Basic experimental parameters). By

plotting force-extension curves (FECs), a wealth of information can be extracted from these experiments, including the presence of distinct folding pathways and intermediates, folding kinetics and diffusion times, and other key information about the energy landscape of the protein (Figure 3.5).

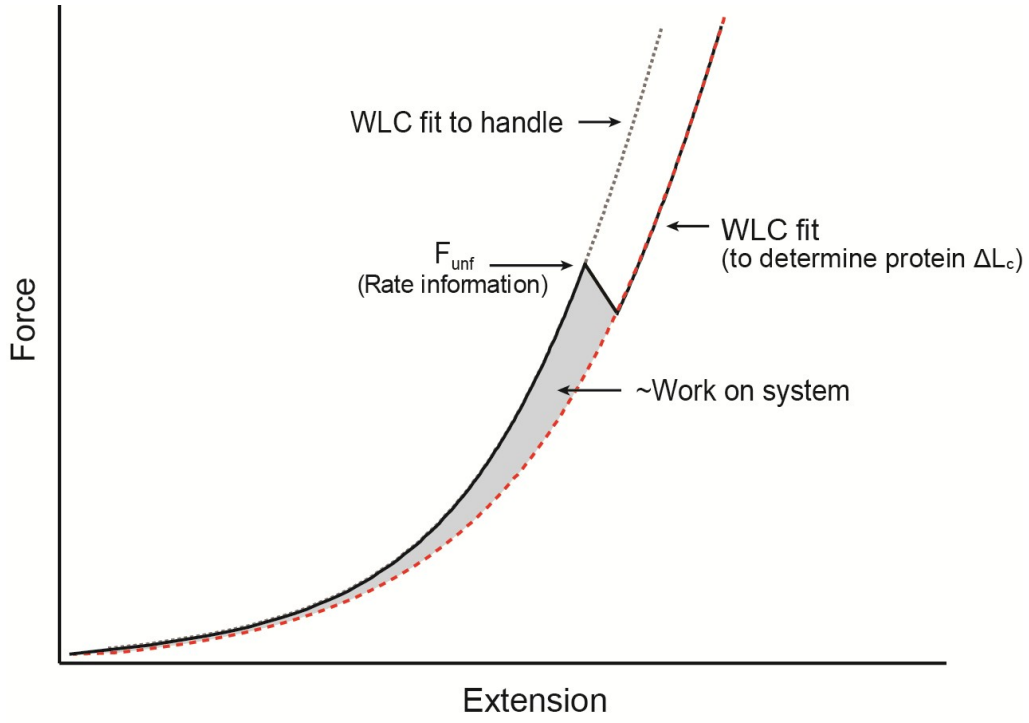

Figure 3.5 FEC measurement and analysis. Force extension curve (FEC, black) with a characteristic "rip", indicating two-state folding. The FEC is fit with two WLCs (black and red dashed lines). F_{unf}, ΔL_C and the area underneath the FEC can be analyzed to extract useful information on the kinetics, structural features, and work on the system, respectively.

In control constructs containing only dsDNA handles, without protein, a monotonic and non-linear rise in force is observed in FECs (Figure 3.6) and is well fit by a extensible version of the worm-like chain (WLC) model for entropic springs (Marko and Siggia 1995), which includes an enthalpic term (K):

$$F(x) = \frac{k_B T}{L_p}\left[\frac{1}{4}\left(1 - \frac{x}{L_c} + \frac{F}{K}\right)^{-2} - \frac{1}{4} + \frac{x}{L_c} - \frac{F}{K}\right], \tag{3.7}$$

where k_BT is the thermal energy, L_p is persistence length, x is the bead extension, L_C is the contour length, F is force and K is the elastic modulus (Wang, Yin et al. 1997). This WLC expression is generally considered to be a valid model until ~65pN, when the dsDNA handles begin to undergo an overstretching transition (Smith, Cui et al. 1996; van Mameren, Gross et al. 2009). All constructs presented here are pulled on well below this force threshold.

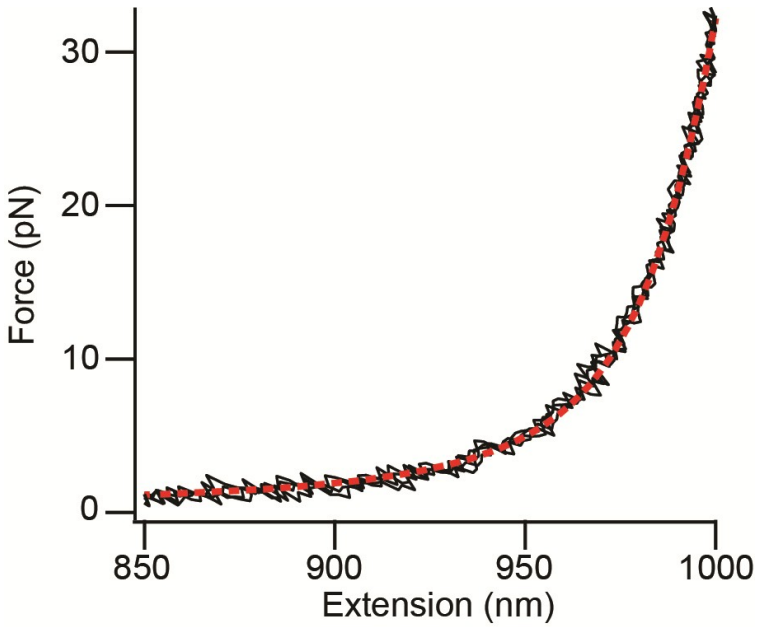

Figure 3.6 FEC of DNA handles only. A FEC measured on DNA handles only, without protein (black), is well-fit with a WLC model (red). Handle FECs never showed any discrete unfolding transitions like those that occurred when α-synuclein is present.

When probing dumbbell constructs containing a protein, any structures present may unfold depending on the magnitude of the applied tension. Cooperative two-state unfolding events are identified by rips on the force-extension curves (Figure 3.5, black), where an increase in molecular extension coincides with a sudden decrease in force. Intermediate states are identified by the presence of additional rips along a particular FEC. To determine the contour length change (ΔL_c) during the unfolding event(s), we begin by fitting the initial portion of the FEC (folded) to two WLCs in series: one for the handles and one for the protein (Figure 3.5, black dashed line). The handle parameters (L_p, L_C, and K) and protein parameters (L_p and K) are adjusted to obtain the best fit. Next, these parameters are held when fitting the unfolded portion of the FEC (Figure 3.5, red dashed line), such that the remaining free parameter returns ΔL_c of the protein.

The unfolding force (F_{unf}), or the force at which the first unfolding rip occurs, can also be recorded for each individual FEC. In turn, this quantity contains information concerning the height of the transition barrier between the unfolded and folded states. To reiterate earlier sentiments, the force ramp experiments can be used to extract key information about a single protein molecule's folding pathway, including the states that are probed by deducing structural information from ΔL_C values, as well as dynamic information (e.g., folding rates). A biophysical framework for interpreting observables from force ramp experiments can be understood in terms of energy landscapes, which is an elegant way of modeling and predicting a molecule's behavior. Original

landscape theory was applied only to natively-folded proteins, so I will present the theory in this framework before introducing predictions for IDP energy landscape properties.

3.4 Energy landscape overview

Energy landscapes provide a quantitative representation of a biomolecule's conformational landscape, such that every potential structure along a chosen conformational coordinate (e.g., protein backbone dihedral bond angles, native-like contacts) is correlated to a free energy value (Dill and Chan 1997; Onuchic, Luthey-Schulten et al. 1997; Plotkin, Onuchic et al. 2002). The shape of the landscape is thus unique to the system. In protein folding, the molecule begins in a highly disordered, high-energy random coil state, diffusing around the landscape due to the effects of Brownian motion, before it proceeds down folding pathways that are guided by the principle that free energy should be minimized (Bryngelson, Onuchic et al. 1995; Dill and Chan 1997). During this time, hydrophobic residues collapse to the interior and bonds form, which are concurrent with an entropy reduction as the number of available conformational states decreases during folding (and increase in solvent entropy) (Plotkin, Onuchic et al. 2002). The protein begins to form varying degrees of secondary structure, before it finally reaches its thermodynamically stable native state. In practice, this is conceptually represented by a 'funnel-like' energy landscape (Bryngelson, Onuchic et al. 1995; Dill and Chan 1997; Onuchic and Wolynes 2004). This funnel landscape can in turn explain the timescale of protein folding, which is often framed in terms of Levinthal's paradox. As proteins fold, they do

not experience a random search through all possible conformations before arriving at their native structure. If they did, Levinthal showed that it would take a medium-sized 150 residue protein $\sim 3^{150} \cdot 10^{-12}$s = $\sim 10^{52}$ years (assuming each amino acid could sample three conformations for $\sim 10^{-12}$ s each). In reality, an evolutionary process funnels proteins down biased folding pathways minimizing the free energy, ensures that most proteins fold on the order of microseconds to minutes (Kubelka, Hofrichter et al. 2004; Milanesi, Waltho et al. 2012).

Next, consider the addition of partially folded, on- or off-pathway intermediate states that may be essential for forming the native structure or may also represent misfolded states. These low energy states add 'ruggedness' to the funnel. An inherent randomness is thus added to the multi-dimensional energy landscape, since the molecule can now sample many different states and folding pathways (Figure 3.7a). The heterogeneity resulting from different folding pathways can in turn shift the magnitude of relative local free energy values, leading to preferential folding (Onuchic and Wolynes 2004). Ultimately, these unstable intermediate states directly influence the protein folding speed, as it transverses over barriers of varying roughness (Milanesi, Waltho et al. 2012).

Unsurprisingly, IDPs have been proposed to possess a unique energy landscape profile, which is directly linked to their dynamic and diverse conformation landscape that is easily perturbed by thermal fluctuations (Uversky 2013). IDPs are thus predicted to have a shallow, yet rough landscape, where states are stable on the order of a few $k_B T$s (Figure 3.7b), but one that can shift due to environmental changes and disruptive mutations (Uversky 2013; Brucale,

Schuler et al. 2014). These perturbations can alter the local energy landscape, leading to preferential folding with varying consequences. For example, a disease-related mutation can lead to a higher probability of the protein being trapped misfolded state. Due to the inherent complexity of measuring energy landscapes directly, calculations of IDP energy landscapes have been for the most part relegated to all-atom simulations (Higo and Umezawa 2014), until this year when we provided the first experimental evidence of an IDP's flat, yet rough landscape (see Chapter 5).

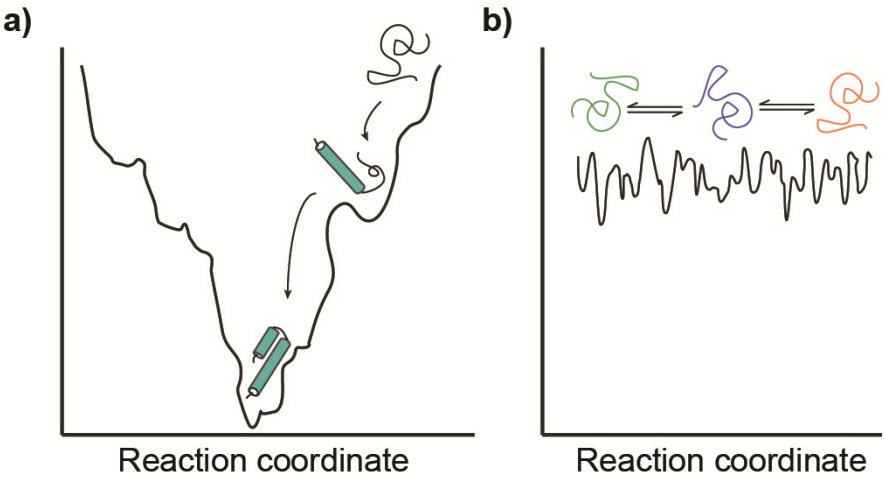

Figure 3.7 Natively-folded protein and IDP energy landscape comparison. Schematics of a) a typical rugged natively-folded protein energy landscape with intermediate states and b) the energy landscape for IDPs, which is predicted to be relatively flat and rough.

There are very few experimental tools available for reconstructing energy landscapes. Using an optical trap, we can effectively overcome the essentially impossible task of experimentally measuring all points of a multidimensional, highly complex energy landscape by projecting it onto a 1D landscape along an appropriate reaction coordinate. In SMFS experiments, this is the molecule's measured extension. Applying a force effectively tilts the energy landscape of the protein, increasing the population of conformational transitions by changing the thermodynamic stability of the molecule. The intermediate states present in FECs that occur between the folded native state and unfolded state are also identified if they are located along the reaction pathway.

3.5 Effect of force on the energy landscape

To fully understand the effect of force on a molecule's energy landscape, we can begin by thinking of the system in thermodynamic terms like entropy (S), temperature (T) and pressure (P). Considering that our experiments are conducted under constant T and P, the Gibbs free energy (G) is considered appropriate to describe the free energy change in an equilibrium system (Bryngelson, Onuchic et al. 1995; Tinoco Jr. and Bustamante 2002), where:

$$\mathrm{d}G = \mathrm{d}E - T\mathrm{d}S + P\mathrm{d}V = F \cdot dx \qquad (3.8)$$

In a two-state system framework, as force is applied to the molecule, the landscape changes along the reaction coordinate (i.e., extension x) by lowering the relative barrier height between folded and unfolded states (Figure 3.8), such that free energy difference at zero force can be written as:

$$\Delta G = \Delta G^{0} - F\Delta x. \quad (3.9)$$

ΔG^0 is the free energy difference at zero force. In two-state systems, a molecule's structural stability is directly related to the transition barrier height, and in this context, $F_{1/2}$ is the force where the molecule is equally probably to fold as it is to unfold. This means that the expression for ΔG^0 may be expressed by the simple relationship:

$$\Delta G^{0} = F_{1/2} \cdot dx, \quad (3.10)$$

Therefore, at any given force, the free energy difference is equal to:

$$\Delta G = (F_{1/2} - F)\Delta x. \quad (3.11)$$

In an equilibrium system such as this, the probability distribution upon applying force can be written as a Boltzmann distribution:

$$P_u(F) = [1 + \exp\{((F_{1/2} - F)\cdot\Delta x(F))/k_B T\}]^{-1}, \quad (3.12)$$

where the distance between adjacent states, $\Delta x(F)$, is also force-dependent (Woodside, Behnke-Parks et al. 2006). Thus, the observation of particular subpopulations (e.g., rare intermediate states) can be directly influenced by altering the applied force on the system.

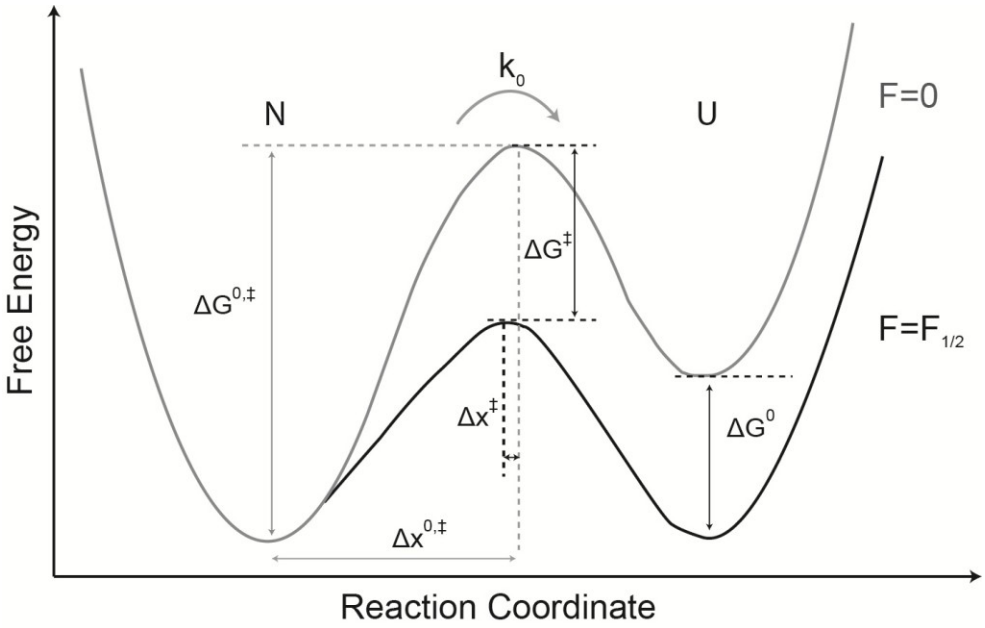

Figure 3.8 Effect of force on a two-state folder energy landscape. A protein that undergoes two-state folding must cross a transition barrier, of height $\Delta G^{0,\ddagger}$, to pass from the folded, native state (N) to the unfolded state (U). Folding rates are defined by the Kramers prefactor (k_0). The energy landscape tilts as force is applied to the system (black), leading to a shift in both the distance to transition state and the barrier height.

3.6 Extracting kinetic information from FECs

Using this framework, a molecule can be described by its diffusion around this 1D projected landscape, over the barrier of height $\Delta G^{\ddagger}$ and within the unfolded/folded state potential wells. Kinetic information can be described by two key parameters: the folding rate, k, which is directly related to the Kramers prefactor k_0 from reaction-rate theory and the diffusion coefficient, D. This prefactor approximates the molecule's speed limit, or the maximum folding rate in absence of a barrier (Kubelka, Hofrichter et al. 2004). Kramers' theory modifies the Bell model, originally used to describe cell-cell adhesion, to describe

1D movement in a double-well potential chemical reaction in a solution (Kramers 1940; Hanggi, Talkner et al. 1990). It has since been adopted to describe the diffusional process by which a protein travels over a 1D barrier, where motion is strongly coupled to the solvent conditions and friction is high. Using Kramers' theory, the folding rate is related to the barrier height (where $\Delta G^{0,\ddagger}$ is the folding free energy barrier at zero force) (Best and Hummer 2010) by:

$$k = k_0 \exp\left(-\frac{\Delta G^{0,\ddagger}}{k_B T}\right), \tag{3.13}$$

where

$$k_0 = \frac{\sqrt{\kappa_{well}\kappa_{barrier}}D}{2\pi k_B T}. \tag{3.14}$$

This folding rate, k, convolves the timescale for the molecule to diffuse within a kinetic trap, as well as the rate to cross over local barriers, the latter of which contributes most to D (Best and Hummer 2010). The diffusion coefficient is an inherent property of the molecule, and is in turn related to the roughness of the landscape. κ_{well} and $\kappa_{barrier}$ are the curvatures (stiffness) of the potential well or barrier, respectively. By inverting Equation 3.13 to obtain an expression for the reconfiguration time in the unfolded state (τ_{unf}), we can extract free energy parameters using experimentally observed folding times and the Kramers prefactor:

$$\tau_{unf} = \frac{2\pi k_B T}{\sqrt{\kappa_{well}\kappa_{barrier}}D} exp\left(\frac{\Delta G^{\ddagger}}{k_B T}\right). \tag{3.15}$$

In absence of a barrier, the folding time is simply equal to the Kramers prefactor.

3.7 Equilibrium free energy from non-equilibrium measurements

The majority of discrete unfolding events observed in FECs occur highly out of equilibrium, since the trap is extending the molecule fast enough to sample states with high energy barriers while work is irreversibly dissipated. Experimentally, these non-equilibrium events are identified by the presence of hysteresis in pulling and refolding curves. To interpret results from non-equilibrium pulling measurements, several widely used analytical tools have been developed, including theorems by Crooks (Crooks 1999) and Jarzynski (Jarzynski 1997). According to Crooks, the distribution of non-equilibrium work required to unfold a molecule, $P_U(W)$, is related to the distribution of work in the reverse direction (refolding), $P_R(W)$, by:

$$\frac{P_U(W)}{P_R(W)} = \left(\frac{W - \Delta G}{k_B T}\right) \tag{3.16}$$

ΔG is defined as the change in equilibrium free energy between the initial and final state for the system in question. Under equilibrium conditions, where no work is dissipated, the expression returns the relation $\Delta G = \langle W \rangle$. Similarly, the Jarzynski equality relates the change in free energy, ΔG, to the exponential average of the non-equilibrium work, W, by the expression

$$\Delta G_{eqbm} = -k_B T \ln \langle exp\left(-\frac{W}{k_B T}\right)\rangle. \tag{3.17}$$

In the FEC experiments described earlier, the work can effectively be estimated by integrating the area under each curve (Figure 3.5).

The unfolding force obtained from non-equilibrium FECs can also be used to extract kinetic information for the protein. Given that these unfolding states are stochastic and irreversible processes out of equilibrium, FECs collected over multiple pulls from a single molecule possess a distribution of unfolding forces and rupture times. From the model described in Dudko *et al*.'s 2006 paper, the probability distribution of unfolding forces, *p(F),* given a certain energy landscape shape:

$$p(F) \propto \frac{k(F)}{r} exp\left\{\frac{k_{unf}}{\Delta x^{\ddagger} r} - \frac{k(F)}{\Delta x^{\ddagger} r}\left(1 - \frac{\Delta x^{\ddagger} F}{\Delta G^{\ddagger}}\nu\right)^{1-\frac{1}{\nu}}\right\} \quad (3.18)$$

where

$$k(F) = k_{unf}\left(1 - \frac{\Delta x^{\ddagger} F}{\Delta G^{\ddagger}}\nu\right)^{\frac{1}{\nu}-1} exp\left\{\Delta G^{\ddagger}\left[1 - \left(1 - \frac{F\Delta x^{\ddagger}}{\Delta G^{\ddagger}}\nu\right)^{\frac{1}{\nu}}\right]\right\} \quad (3.19)$$

and k_{unf} is the unfolding rate at zero force, r is the loading rate, ν is the shape of the barrier (ν = 1/2 for a cusp potential, ν = 2/3 for a soft/cubic potential) (Dudko, Hummer et al. 2006). If ν = 1, then Bell's formula is recovered ($k = \exp(F\Delta x^{\ddagger}/k_B T)$. In the work presented here, we assumed ν = 2/3. The caveat to this discussion is that these types of analyses require statistics such that force distributions are well populated. However, in the case of IDPs like α-synuclein, where the folding landscape consists of a vast set of diverse structures, statistically significant analyses for a particular state are difficult to administer.

Taken as a whole, SMFS with optical tweezers is a powerful experimental technique which allows us to extract structural and kinetic information about a protein as it traverses its own unique energy landscape, including diffusion

coefficients and transition state information. In Chapters 4 and 5, for example, SMFS force ramp experiments help elucidate the complex behavior of early oligomeric states of α-synuclein by probing the transient and diverse conformational landscape.

Appendix:

Basic experimental parameters

Two orthogonally-polarized laser beams from the 1064 nm trapping laser were used to generate two traps, "T0" and "T1". The position and stiffness of each trap were controlled independently in each axis by two AODs. During measurements, only one bead (in trap T0) is tracked for position by the PSDs as it is displaced from its trap centre. The applied force is found by the expression $F=\kappa_{trap,0} \cdot dx_0$, where the stiffness and the bead displacement position are recorded. Trap stiffnesses were calibrated as described in the following section and values were set to 0.43 and 0.54 pN/nm for all experiments presented here. FEC data were acquired at 20 kHz and filtered online at 10 kHz with an 8-pole Bessel filter (Krohn-Hite). Traps were moved apart at a rate of 100, 250, or 400 nm/s.

Calibrating optical tweezers

Before measurements commence, the first order of business was to align the trapping beams on top of the detection beams, as well as calibrate the PSDs. The next task was to calibrate the trap stiffness. To ensure measurements were taken with the upmost accuracy, position and stiffness calibrations (both of which are directly linked to force via Hook's Law, $F=-\kappa_{trap}x$) were of the upmost priority. Three different types of calibrations were conducted: Stokes drag, power spectrum and variance. The last is perhaps the easiest, since the variance in the trap's stiffness is directly linked to the thermal fluctuations of a particle (Brownian motion) in a harmonic potential via the equipartition theorem:

$$\frac{1}{2}\kappa_{trap}\langle x^2\rangle = \frac{1}{2}k_B T. \tag{3.20}$$

The power spectrum calibration involved taking the Fourier transform of the bead position as it's driven by thermal energy, k_BT. The power spectrum was fit to the Lorentzian,

$$P(f) = \frac{k_B T}{\pi^2 \beta (f_0^2 + f^2)}, \tag{3.21}$$

where f_0 is the roll-off frequency ($f_0 = \kappa_{trap}/(2\pi\beta)$, which effectively gives the stiffness of the trap. β is the hydrodynamic drag coefficient given by the Stokes drag relation, and is related to viscosity (η) by

$$\beta = 6\pi\eta a. \tag{3.22}$$

Since the beads were trapped near the surface (~1μm away), the viscosity is more accurately described by Faxen's Law which provides a correction for the effective drag coefficient of a sphere with radius a above a surface H:

$$\beta' = \frac{6\pi\eta a}{1 - \frac{9}{16}\left(\frac{a}{H}\right) + \frac{1}{8}\left(\frac{a}{H}\right)^3 - \frac{45}{256}\left(\frac{a}{H}\right)^4 + \frac{1}{16}\left(\frac{a}{H}\right)^5}. \tag{3.23}$$

The final method, measuring the viscous drag of a stuck bead in the sample buffer conditions, involved moving the stage at different speeds and tracking the bead position. The stiffness was directly recovered using the following relation, where trap stiffness, κ_{trap}, is related to the bead position ($x(v)$ and stage speed (v) by:

$$\kappa_{trap} = \frac{F}{x} = \frac{\beta v}{x}. \quad (3.24)$$

Each calibration result was averaged over multiple beads. For additional background on optical trapping and its components, there are many excellent overviews on optical trapping (Svoboda and Block 1994; Neuman and Block 2004; Woodside and Valentine 2009). For more specific information on the calibration of optical tweezers, please refer to the master's thesis of Daniel A.N. Foster (Foster 2010) and PhD thesis of Hao Yu (Yu 2013).

Measurement resolution

There are several key factors that influence the resolution and experimental uncertainty in optical trapping measurements, arising both from the limitations of the instrument and the inherent properties of the molecule being measured. For example, instrumental drift may limit the ability to average multiple FECs, as done for fitting, when the curves are acquired over long time intervals. Drift may also affect the analysis of the correlations in the extension, thereby changing the apparent values for the microscopic rates. Another important factor limiting the position resolution is the force at which the structures unfold. At higher forces, the compliance of the tethered molecule is higher (owing to the non-linear stiffness of WLCs), hence Brownian motion is reduced and better position resolution can be achieved than at low forces. Noise can also be effectively coupled into the optical trap, thus greatly deteriorating both the stability and resolution of the measurements. Noise may arise from various sources, such as the laser itself (e.g., pointing stability), environmental vibrations

(e.g., air currents, stage movements) and mechanical noise (e.g., vacuum pumps nearby). There are several strategies to isolate the system. Firstly, the trap is housed in a soundproofed room, with temperature stability (~0.1°C deviations) atop a floating optics table. Beams are protected by plastic tubing, thus minimizing disruptions from air currents, while all of the optics components are enclosed in plastic boxes. Molecules themselves are isolated from the sample slide by conducting the measurements off the surface (such that the beads are roughly 600nm above the slide) in the dual-beam, dumbbell experimental set-up employed these experiments.

Chapter 4. Diverse Metastable Structures Formed by Small Oligomers of α-Synuclein Probed by Force Spectroscopy

In this chapter, the first results from optical tweezers studies of α-synuclein are presented. Monomers and tandem oligomer FECs from force ramp experiments were analyzed and information on the structural stability and kinetics of individual molecules uncovered a much more complex structural landscape than previously reported. A variation of this chapter was already published: Neupane, K.*, Solanki, A.*, Sosova, I., Belov, M., Woodside, M.T. "Diverse metastable structures formed by small oligomers of α-synuclein probed by force spectroscopy." *PLoS One* 9(1): e86495 (2014). *Co-first authors. In this work, AS contributed reagents and materials with IS, performed experiments with KN, and analyzed the data with KN and MTW. MB built the instrument.

4.1 Investigation of IDP aggregation behavior using optical tweezers

To probe the formation of aggregated structures in small α-synuclein oligomers systematically, we compared the behavior of monomeric human α-synuclein when unfolded by optical tweezers to that of engineered α-synuclein oligomers consisting of monomers linked in defined geometries by short peptides (Figure 4.1; also see Appendix: Protein purification and verification). Similar "tandem-repeat" oligomer constructs have been used previously to study aggregation in other peptides and proteins (Kellermayer, Bustamante et al. 2003; Bader, Bamford et al. 2006; Borgia, Borgia et al. 2011;

Laganowsky, Liu et al. 2012; Speretta, Jahn et al. 2012), and such oligomers made from disease-related proteins have shown toxicity in both neuronal cell cultures and animals (Simoneau, Rezaei et al. 2007; Speretta, Jahn et al. 2012). In the context of tandem-repeat oligomers, we define "aggregation" as the formation of a stable structure involving the association between two or more of the monomeric domains in the repeat. Beyond allowing the size of the aggregate to be controlled, the tandem-repeat oligomers enable a high local protein concentration to be maintained to encourage aggregation, while still keeping a low total concentration for work in the single-molecule regime. We focused on dimers and tetramers to study minimal aggregates and probe how the aggregation behavior changes with oligomer size.

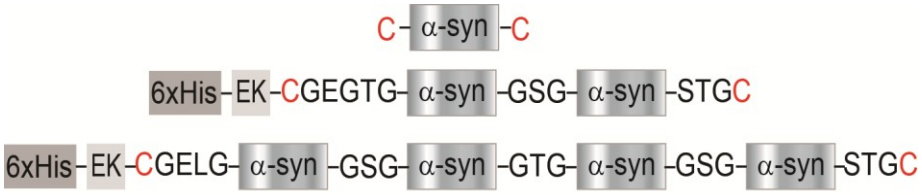

Figure 4.1 Engineered α-synuclein constructs. Schematic of the monomer and tandem-repeat protein constructs, the latter containing an N-terminal His-tag and enterokinase cleavage site (EK). The multimers were made from monomers linked at their termini via a short amino acid sequence. Constructs were attached to sulfhydryl-labeled DNA handles via disulfide bonds with cysteine residues (red) at each terminus of the protein.

4.2 Results from α-synuclein monomer and oligomer FECs

We first measured the unfolding of single α-synuclein monomers to provide a baseline for comparing the behavior of the oligomers in the SMFS experiments. The structure formed by the monomer was probed with FECs, measuring the end-to-end extension of the molecule while moving the traps apart at constant speed to ramp up the force until the protein was completely stretched out, as described in Chapter 3, Section 3. Repeated FEC measurements on the same molecule were typically separated by a 5 s delay at zero force to allow time for structures to form.

Although α-synuclein is nominally an intrinsically-disordered protein, a surprisingly wide diversity of structural behavior was observed in repeated pulling measurements. Most FECs (~ 85%) displayed a simple monotonic rise of force with extension (Figure 4.2a, cyan), well fit by two WLCs in series (see Chapter 3.3; Figure 4.2a, red). Similarly featureless WLC behavior was also seen in FECs of the DNA handles alone (Figure 3.6); the absence of discontinuities in these FECs indicates that α-synuclein did not contain any stable or metastable structures, as might be expected for an intrinsically-disordered protein. However, a significant minority of FECs (~ 15%) contained discrete rips (Figure 4.2a, black, orange, blue) consisting of an abrupt extension increase and concomitant force drop, which are the signature of cooperative unfolding of a well-defined structure (Borgia, Williams et al. 2008). The unfolding forces, F_{unf}, which are related to the structural stability and the height of the energy barrier to unfolding, were typically ~ 5–15 pN.

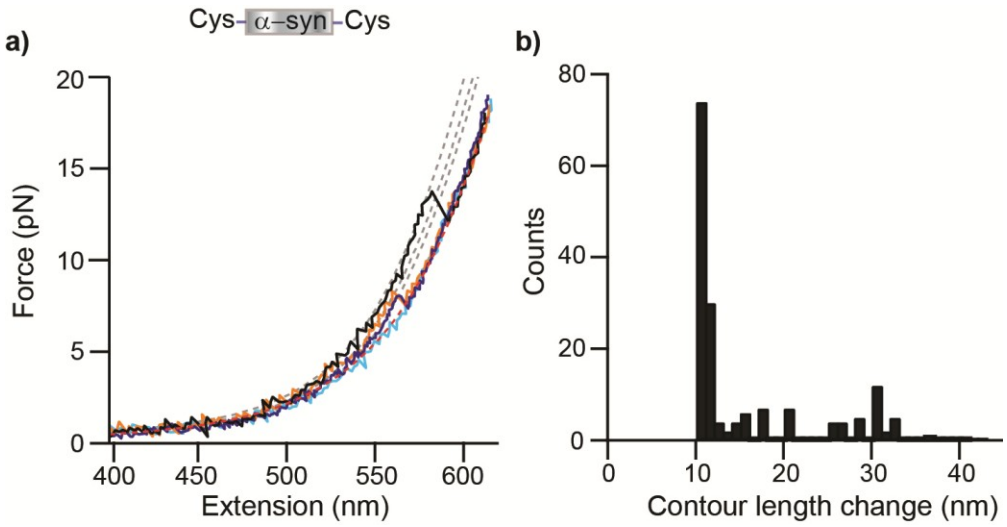

Figure 4.2 Force spectroscopy of α-synuclein monomers. a) Most FECs of a single monomer display no structure (cyan) and fit well to the WLC model expected for the unfolded-state (red). Some reveal discrete unfolding transitions (black, orange, blue) with different contour lengths, as found from WLC fits (grey). b) Histogram of ΔL_c for all identifiable transitions in FECs of the monomer.

The contour length change upon unfolding, ΔL_c, which reflects the number of amino acids unfolded during the structural transitions, was found by fitting the FECs to WLC models for the folded and unfolded states (Figure 4.2a, grey and red, respectively; see Appendix: Force ramp experiment details). ΔL_c values were most prominent (greater than two counts) between ~ 10 to 36 nm (Figure 4.2b), indicating the formation of several different structures. Given that the full contour length of α-synuclein is ~ 50 nm (140 amino acids with a contour length per amino acid of 0.36 nm (Pauling and Corey 1951), only part of the protein was folded in these structures. Surprisingly, many more structures were resolved than in previous single-molecule work using AFM (Sandal, Valle et al. 2008; Brucale,

Missimo et al. 2009) or fluorescence (Ferreon, Gambin et al. 2009; Ferreon, Moran et al. 2010; Trexler and Rhoades 2010) under similar conditions, most likely due to the high spatial resolution and force sensitivity of our trap, but there was no evidence of structures unfolding at very high F_{unf} (> 100 pN) as previously reported (Sandal, Valle et al. 2008; Hervas, Oroz et al. 2012).

We next probed the smallest possible oligomer—a dimer—by measuring FECs of the tandem dimer under the same conditions. Just as for the monomers, most dimer FECs (~ 85%) did not show discrete cooperative unfolding events. For those that did, however, the behavior was considerably more complex than in monomer FECs. Sometimes structures in the dimers unfolded completely in one step, but in many cases multiple unfolding transitions were observed within the same FEC, revealing the presence of several distinct structures or unfolding intermediates in the aggregate (Figure 4.3a). Most of the discrete unfolding transitions still had a relatively small ΔL_c of ~ 10–30 nm (Figure 4.3a), as determined by WLC fits (dashed lines), but sometimes much larger values were also observed (Figure 4.3b). Specific ΔL_c values were observed reproducibly in different pulls and with different molecules (Figure 4.3c). Most notably, in many of the FECs, ΔL_c during the unfolding was greater than the 50-nm contour length of monomeric α-synuclein. Such a result is not possible unless the structure being pulled apart involved interactions between amino acids from more than one monomer—*i.e.*, the two monomers in the tandem repeat were interacting to form an aggregated structure. The maximum observed ΔL_c was ~ 102 nm, equal to the full contour length of the dimeric construct, indicating the unfolding of a structure

that encompassed all the amino acids in both monomer units of the repeat. Although the structures observed in the dimer were larger than those observed in the monomer, F_{unf} was only slightly higher, typically ~ 5–20 pN.

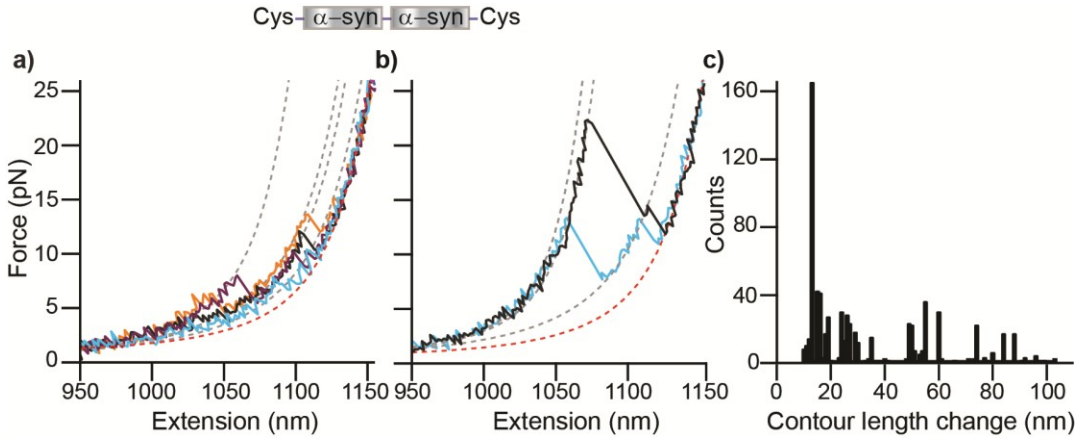

Figure 4.3 FECs of α-synuclein dimers. a,b) Representative FECs of a dimer show unfolding of stable structures with a wide range of sizes and unfolding forces. WLC fits to determine contour length changes are displayed as dashed lines (grey: folded states, red: unfolded state). Inset: the dimer contains two monomers connected by short, flexible peptides linkers. c) Histogram of ΔL_c for all identifiable transitions in dimer FECs.

Finally, we measured FECs of α-synuclein tetramers under the same conditions as the monomers and dimers. The behavior of the tetramers was qualitatively similar to that of the dimers, but even more complex, with cooperative unfolding observed in ~ 25% of the total pulls. Again, sometimes structures unfolded completely in one step (Figure 4.4a, black), but more often multiple intermediate steps were detected (Figure 4.4a, cyan, blue, purple; Figure 4.4b). ΔL_c varied over an even wider range than for the dimer, most commonly 10–50 nm but in some cases up to ~ 205 nm (equal to the full contour length of the tandem tetramer) (Figure 4.4c). Analogous to the dimers, any transitions with $\Delta L_c > 50$ nm must have involved the unfolding of structures formed by interactions between more than one monomeric unit of the tandem repeat. In the case of the largest transitions, those with $\Delta L_c > 150$ nm, the structure being unfolded must have involved interactions, at the very least, between the first and fourth α-synuclein monomer subunits, with the possibility that all four monomer units were sequestered to form a single aggregate (Figure 4.4a, black). Significantly higher forces were frequently found when unfolding the tetramer, with F_{unf} varying over a much wider range than for the dimer, from ~ 5–40 pN.

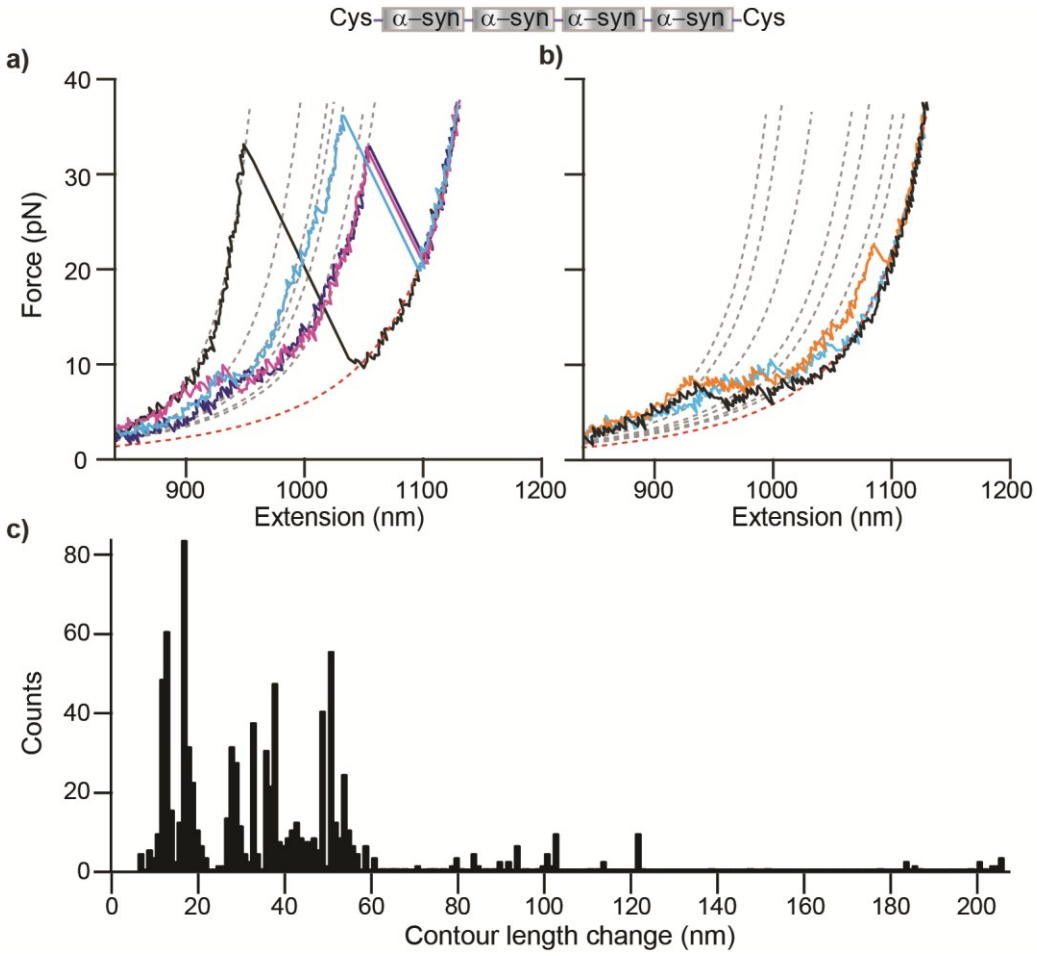

Figure 4.4 FECs of α-synuclein tetramers. a,b) Representative FECs of a tetramer reveal many structures with different sizes and unfolding forces. WLC fits are shown as dashed lines (grey: folded states, red: unfolded state). Inset: the tetramer contains four α-synuclein domains connected by short, flexible peptide linkers. c) Histogram of ΔL_c for all identifiable transitions in FECs of the tetramer.

One of the most notable features of these results is the diversity of structures that can form in monomers and oligomers of the intrinsically-disordered protein α-synuclein. The results for 2,498 FECs from 12 tetramer molecules, 1,769 FECs from 6 dimers, and 1,152 FECs from 4 monomers are

summarized in histograms of ΔL_c and F_{unf} (Figure 4.5), with the distributions of ΔL_c for all identifiable, discrete unfolding transitions shown in Figure 4.5a. The number of different states was estimated from the number of peaks in the distributions, given the resolution of 2 nm: approximately 5 states for the monomer (Figure 4.5a, black), ~ 15 for the dimer (Figure 4.5a, blue), and ~ 20–25 for the tetramer (Figure 4.5a, red). This behavior is much richer than previously reported for α-synuclein monomers (Sandal, Valle et al. 2008; Ferreon, Gambin et al. 2009) and oligomers (Yu, Malkova et al. 2008; Cremades, Cohen et al. 2012; Krasnoslobodtsev, Peng et al. 2012; Krasnoslobodtsev, Volkov et al. 2013), or indeed for other aggregation-prone intrinsically-disordered proteins such as Aβ (Kim, Palermo et al. 2011), indicating the extreme complexity of aggregation landscapes even for small oligomers of α-synuclein. The ability to distinguish such diverse sub-populations of structural transitions, some of which occurred very rarely (less than 0.1% of the time), is a direct result of the high resolution and sensitivity of our experimental approach.

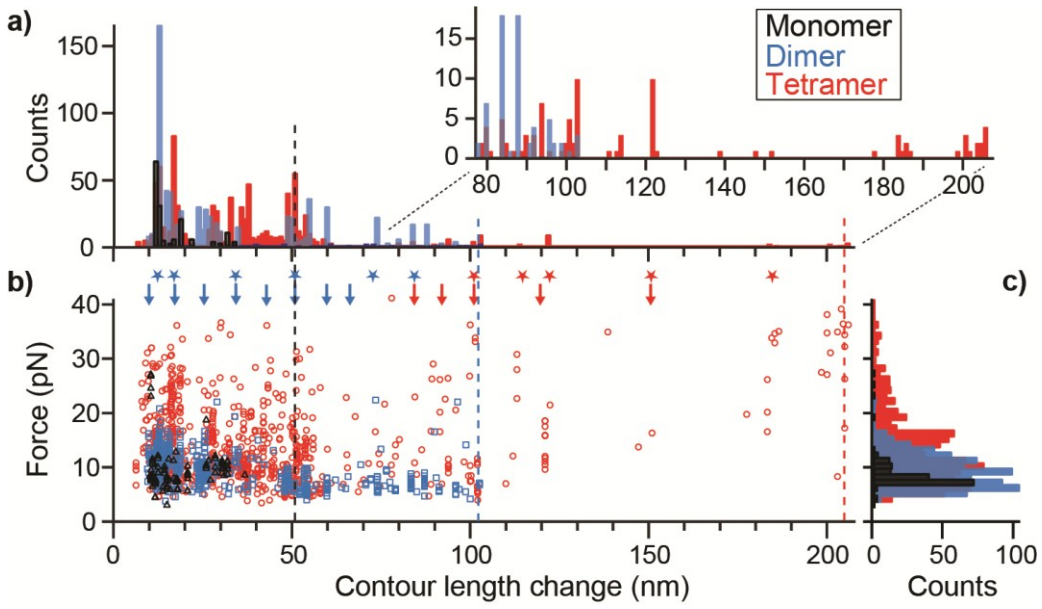

Figure 4.5 Contour length and unfolding force distributions. a) Histogram of ΔL_c for all identifiable transitions in FECs of the tetramer (red), dimer (blue), and monomer (black). b) Scatter plot of F_{unf} versus ΔL_c for tetramer (red), dimer (blue), and monomer (black). Arrows indicate ΔL_c values consistent with a β-sandwich structure, asterisks indicate ΔL_c values expected from a helical multimer structure (blue: dimer and tetramer, red: tetramer only). Dashed lines indicate the contour lengths of the entire monomer (black), dimer (blue), or tetramer (red). c) Histograms of F_{unf} for the tetramer (red), dimer (blue), and monomer (black) show an increase in F_{unf} with increasing oligomer size.

The existence of structures in α-synuclein that can withstand tens of pN of force might naively seem unlikely, given that α-synuclein is known to be intrinsically disordered under the conditions of these measurements. Circular dichroism spectra of the monomers, dimers, and tetramers used in these experiments confirm that the proteins remain largely unstructured (Figure 4.14; see Appendix: Protein biochemistry assays). Indeed most pulling measurements

did not show any evidence of cooperative unfolding transitions, suggesting that the disordered state is the minimum-energy configuration. The unfolding we observed is thus best understood in terms not of structures that are thermodynamically stable, but rather of kinetically-trapped conformations that are only metastable. Such states should form transiently as the protein undergoes thermally-driven conformational fluctuations, with a frequency and duration determined by the relative free energy of the state and the height of the energy barrier. When fluctuations into metastable states occur, they would be expected to generate discrete unfolding transitions, as in Figures 4.2-4.4, if the barriers for unfolding are sufficiently high so as to allow the structures to persist on the timescale of a pulling measurement (on the order of 0.5–1 s). Our results should thus be interpreted as reflecting the range of structures physically possible for α-synuclein to form via fluctuations into high-energy states, rather than only those structures that are stable under the specific conditions of the measurement.

4.3 Interpretation of FEC results

The length changes found in the unfolding transitions in Figure 4.5 can be compared to the values that would be expected from the different structures that α-synuclein is known or proposed to form under various conditions, to test if the kinetically-trapped metastable structures are consistent with any of these structural models. For example, monomeric α-synuclein is known to form helical structures under various conditions (Davidson, Jonas et al. 1998; Ulmer, Bax et al. 2005). These structures should produce $\Delta L_c \sim 7$, 10, 20, or 30 nm upon unfolding. Since these values are, in fact, seen in some of the monomer FECs, as well as in

dimer and tetramer FECs (Figure 4.5), it is possible that helical structures like those observed previously in the presence of micelles or membranes could be forming transiently in our measurements, despite the lack of the co-factors required to stabilize them. Computational simulations of α-synuclein have suggested another possibility, finding a metastable structure consisting of five β-strands arranged in a zig-zag pattern to form a "sandwich" (Jónsson, Mitternacht et al. 2013). Monte Carlo simulations of mechanical pulling on this structure indicated that it should produce unfolding ΔL_c distributions peaked at ~ 12, 18, 26, and 37 nm (Jónsson, Mitternacht et al. 2013). All of these values were observed in our measurements, on the monomer as well as the dimer and tetramer, indicating that the β-structured model is consistent with at least some of the results (See Appendix: Additional interpretation of unfolding distances for further L_c analysis).

Turning to structural models for oligomeric α-synuclein, two proposals based on experimental observations have sufficient detail to predict possible ΔL_c values in pulling experiments: a helical tetramer proposed as the native structure of endogenous α-synuclein (Wang, Perovic et al. 2011), and a 5-strand β-sandwich structural model proposed for α-synuclein amyloid (Vilar, Chou et al. 2008) (Figure 4.6). Many of the transitions in the tetramer FECs (Figure 4.5, stars) and even the dimer FECs (Figure 4.5, blue stars only) do indeed produce ΔL_c values consistent with those that might be expected from unfolding various structural elements of the helical model (Table 4.1). A significant number of the transitions also involve ΔL_c values that could be expected from the β-sandwich

model (Figure 4.5, arrows; blue: dimer and tetramer; red: tetramer only), suggesting that a β-sandwich structure might form in oligomers as small as dimers and tetramers. In total, ~ ½ of the dimer FECs contained ΔL_c values consistent with only the helical model, whereas ~ ⅓ contained ΔL_c values consistent with only the β-sandwich model (with some overlap between the populations due to degeneracies). For the tetramer, the ratios were reversed. However, roughly ⅓ of all FECs measured contained ΔL_c values that are inconsistent with any known model of α-synuclein, indicating that additional structural models remain to be developed.

Table 4.1 Potential unfolding distances estimated from structural models in literature.

Length changes for unfolding β-sandwich		**Length changes for unfolding helical tetramer**	
Part of the structure unfolded	**ΔL_c (nm)[a]**	**Part of the structure unfolded**	**ΔL_c (nm)[b]**
2 β-strands in a monomer (β1→ β2, β4→ β5)	8–10	N-terminal helix	11–12
3 β-strands in a monomer (β1→ β3, β3→ β5)	8–10	C-terminal helix	17–18
4 or 5 β-strands in a monomer (β1→ β4, β2→ β5, β1→ β5)	16–18	Helix-hairpin (both N- and C-terminal helices)	33
Unstacking of two β-sandwiches	33	Unstacking of two helix-hairpins	18
Complete unfolding monomer	51	Complete unfolding of a monomer	51
Complete unfolding of dimer	68	Complete unfolding of a dimer	85
Complete unfolding of trimer	128	Complete unfolding of a trimer	135
Complete unfolding of tetramer	168	Complete unfolding of a tetramer	185

[a]Contour length changes expected from unfolding different structural components of the stacked, 5-stranded β-sandwich structure in amyloid fibrils of α-synuclein, and the proposed α-helical tetramer of native α-synuclein. For the β-sandwich, ΔL_c values include unfolding of various β-strands, unstacking of β-sandwiches, and unfolding of complete monomers. [b]For the helical tetramer, ΔL_c values include unfolding individual helices, unfolding hairpins consisting of the N- and C-terminal helices, unstacking of neighboring helix-hairpins, and unfolding of complete monomers. Different permutations of these transitions are also possible.

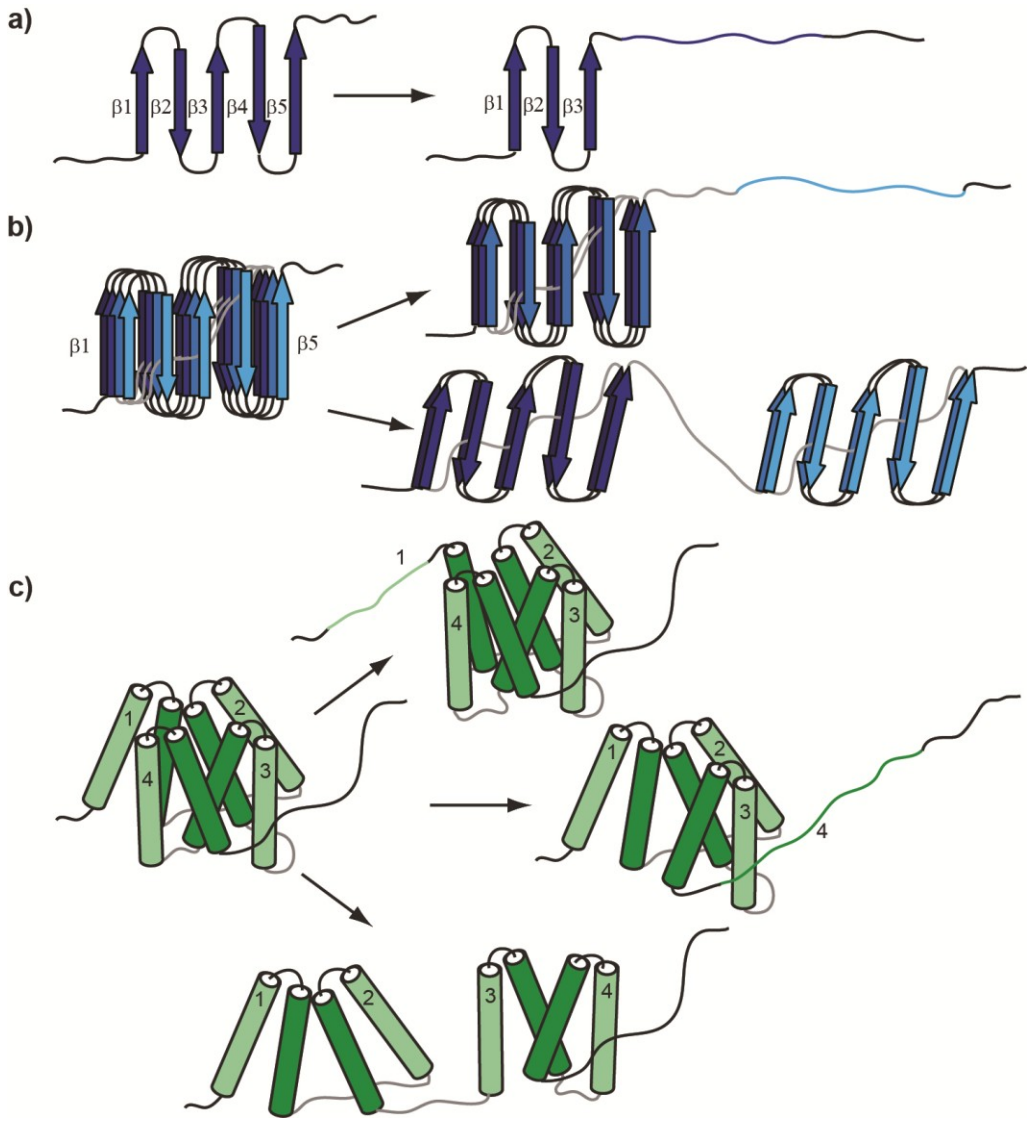

Figure 4.6 Examples of unfolding transitions from different structural models. a) Possible unfolding transitions in a monomer of the β-sandwich model. Unfolding of two β-strands (*e.g.* β4→ β5) produces ΔL_c ~ 8–10 nm. Unfolding all the β-strands in a monomer produces ΔL_c ~ 18 nm. b) Possible unfolding transitions in a stacked β-sandwich. Unfolding one β-sandwich completely from the tetramer produces ΔL_c ~ 50 nm (upper), while unstacking two β-sandwiches produces ΔL_c ~ 33 nm (lower). c) Possible unfolding transitions in the α-helical tetramer: unfolding of the N-terminal helix produces ΔL_c ~ 12 nm (top), unfolding the C-terminal monomer produces ΔL_c ~ 33 nm (middle), and unstacking neighboring helix-hairpins produces ΔL_c ~ 18 nm (bottom).

Yet another possibility is that the transient structures found in our measurements represent various "molten globule" conformations. Indeed, α-synuclein has been shown to form compact and partially folded conformations that are similar to the molten globule state of a natively-folded protein (a state with most of its native secondary structure but incomplete tertiary structure) (Uversky, Li et al. 2001; Bertoncini, Jung et al. 2005; Dedmon, Lindorff-Larsen et al. 2005). Molten globules have been shown to yield the same ΔL_c values upon unfolding as fully-folded structures but to be much more mechanically compliant, having a mechanical unfolding barrier that is much further from the folded state and hence more sensitive to the applied force (Elms, Chodera et al. 2012). To investigate this question, we used dynamic force spectroscopy (DFS) to compare the unfolding force for a particular transition at different loading rates (Evans and Ritchie 1997). The average unfolding force, $\langle F_{unf} \rangle$, varies with the loading rate (rate of change of force), r, as:

$$\langle F_{unf} \rangle \approx \frac{k_B T}{\Delta x^{\ddagger}} ln\left(\frac{r \Delta x^{\ddagger}}{k_{unf} k_B T}\right), \tag{4.1}$$

where $\Delta x^{\ddagger}$ is the distance to the unfolding barrier, k_{off} is the unfolding rate at zero force, and k_B is Boltzmann's constant (Evans and Ritchie 1997). Owing to the infrequent occurrence of any given transition (very few occur more than 1% of the time), only the most common ones could be analyzed by DFS: the 12-nm transition in dimers and the 17-nm transition in tetramers. Fitting the average unfolding force for these transitions as a function of loading rate to Equation 4.1 (Figure 4.7), we found that in each case the barrier was very close to the folded state (~ 1 nm), inconsistent with a molten globule.

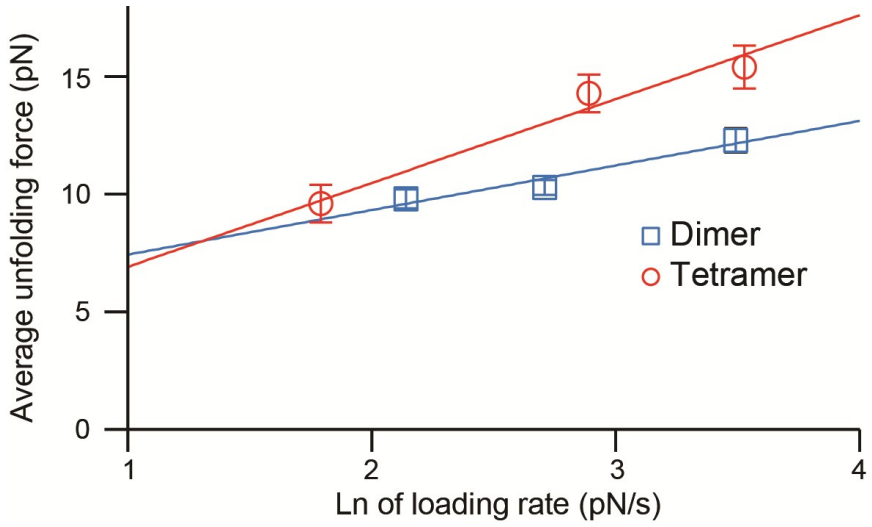

Figure 4.7 Dynamic force spectroscopy. Loading rate dependence of the average unfolding force for the two most frequent transitions: $\Delta L_c = 11–13$ nm for the dimer (blue) and $\Delta L_c = 16–18$ nm for the tetramer (red). Fits to Equation 4.1 yield the unfolding rates at zero force, ~ 0.1 s^{-1}, and the distance to the barrier for unfolding, ~ 1 nm.

Ramping the force down to induce refolding (at a speed identical to that used to unfold the molecule) (Figure 4.8) also did not show any evidence of refolding transitions or "hopping" between folded and unfolded states, as would be expected for a molten globule (Elms, Chodera et al. 2012). Interestingly, the unfolding rate at zero force from these fits was ~ $10^{-1\pm0.5}$ s^{-1}, faster than for stable natively-folded proteins (whether helical (Stigler, Ziegler et al. 2011; Yu, Gupta et al. 2012) or β-structured (Rief, Gautel et al. 1997; Schlierf, Li et al. 2004)) but slow enough that the structures would persist to be captured in the pulling measurements, consistent with the notion that we are probing transient metastable states.

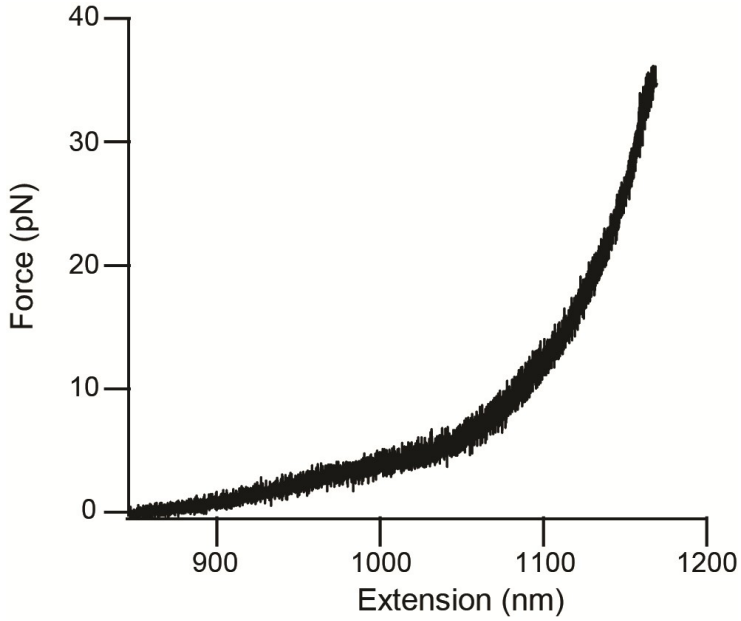

Figure 4.8 Refolding FECs. FECs measured while relaxing the force continuously from the fully-unfolded state did not show any discrete refolding transitions, in contrast to the behavior during unfolding curves, indicating that the structure formation occurs at or near zero force.

Considering the unfolding forces in more detail, many of the observed structures were mechanically quite stable, unfolding at forces similar to or higher than some thermodynamically-stable, natively-structured proteins (Stigler, Ziegler et al. 2011; Elms, Chodera et al. 2012; Jagannathan, Elms et al. 2012; Yu, Liu et al. 2012). The interpretation of the unfolding force is complicated, however, by the fact that F_{unf} depends on features of both the protein and the measurement. For example, faster pulling speeds lead to higher unfolding forces (Evans and Ritchie 1997), F_{unf} is generally higher for β-rich proteins (Rief, Gautel et al. 1997; Carrion-Vazquez, Li et al. 2003; Crampton and Brockwell 2010) than those that are α-rich (Cecconi, Shank et al. 2005; Gebhardt, Bornschlögl et al. 2010; Stigler, Ziegler et al. 2011; Yu, Liu et al. 2012), and tertiary interactions tend to increase

F_{unf} (Elms, Chodera et al. 2012). A key factor is the orientation of the applied force with respect to the structure (*i.e.*, geometry of pulling): a structure that is "sheared" by the applied force unfolds at much higher force than one that is "unzipped" (Jagannathan, Elms et al. 2012).

4.4 Comparison of results to previous SMFS experiments

Comparing our results to those of previous SMFS studies of α-synuclein, we find that the forces described here are similar in magnitude to the forces required to pull apart dimerized α-synuclein molecules with an AFM at similar pulling speeds (Yu, Malkova et al. 2008); the dimerized α-synuclein structure also had a brief lifetime at zero force, consistent with our DFS analysis. However, our results differ in important ways from two previous AFM studies of monomeric α-synuclein (Sandal, Valle et al. 2008; Hervas, Oroz et al. 2012), which found that the monomer was structured much more frequently (up to almost 50% of the time) and that it often required extremely high force to unfold—hundreds of pN, similar to what is required to unfold the most stable proteins. We attribute the differences to the design of the samples used in the AFM studies: the α-synuclein was incorporated into constructs containing other proteins in very close proximity, which could easily promote structures (such as co-aggregates) that would otherwise not form. This interpretation is bolstered not only by the increased tendency of α-synuclein to form mechanically-stable structures in the context of the AFM measurements, but also by the fact that the two studies—which used different sample designs—also found qualitatively different results (recall discussion in Chapter 2, Section 3.4). In one case, where the α-synuclein

was sandwiched between multiple copies of titin, the high-force unfolding events had a ΔL_c distribution with one main peak, which was attributed to a dominant structural subpopulation (Sandal, Valle et al. 2008); in the other case, where the α-synuclein was incorporated as a loop into a ubiquitin domain sandwiched between multiple copies of unmodified ubiquitin, a broad, essentially flat ΔL_c distribution for the high-force unfolding events indicated much higher structural heterogeneity (Hervas, Oroz et al. 2012). No such high forces were observed in over 5,400 pulls using our sample design; hence we believe that the high-force states are artifacts of interactions introduced by the extra proteins in the AFM samples. The avoidance of such extraneous protein-protein interactions, which can apparently affect the outcome for non-native folding strongly, is an important advantage of our assay design.

Several intriguing trends with oligomer size can be identified from the data summarized in Figure 4.5. Comparing the dimer to the monomer, we first note the surprisingly large increase in the number of different stable structures that can be formed by the dimer (~ 3 or more times as many as in the monomer). Interestingly, there are many transitions with $\Delta L_c < 50$ nm (the monomer length) that were observed for dimers but not for monomers. This suggests that such transitions in the dimer represent the unfolding of small structures formed by interactions between adjacent portions of different monomers in the dimeric construct, or else structures formed within a single monomer that are only stabilized in the context of a larger aggregate. A similar trend is seen when comparing the tetramer to the dimer, with many more transitions having

$\Delta L_c \sim 35–50$ nm observed for the tetramer than for the dimer. These trends correlate with an increased tendency to unfold via intermediates in the tetramer as compared to the dimer and monomer: ~ 40% of the tetramer FECs containing discrete transitions include unfolding intermediates, as opposed to only ~ 20% and 15% for dimers and monomers, respectively. Hence there is an increasing number of independently-stable structural elements as the number of monomers available to participate in structure formation increases. However, there does not seem to be any single, dominant structural intermediate mediating the aggregation process. The structural diversity observed in these small oligomers provides a possible origin for the wide diversity of structures observed in larger aggregates (Eichner and Radford 2011; Fauerbach, Yushchenko et al. 2012).

F_{unf} also varies with the oligomer size, trending consistently higher for larger oligomers (Figure 4.5c). The average F_{unf} for all discrete transitions increases from 9 pN for the monomer to 10 pN for the dimer and 14 pN for the tetramer. Most noticeably, there is a very prominent tail of high-F_{unf} events for the tetramer. Surprisingly, the high-F_{unf} tail in the force distribution for the tetramer is associated with transitions covering the entire range of ΔL_c values; even the smaller structural elements became more stable in a larger oligomer, indicating that the increased F_{unf} is not merely due to an increased chain length. Instead, the increase in F_{unf} with increasing oligomer size reveals that the aggregated structures are mechanically stabilized as the number of monomers involved grows. Although it is difficult to know the cause of the general trend of increasing F_{unf}, we speculate that it might indicate an increasing tendency to form β-rich

structures (which are typically more stable) in larger oligomers, consistent with the observation that the fraction of pulls having ΔL_c in agreement with the β-sandwich model increases from the dimer to the tetramer. Another possibility, especially for the high-force tail of the distribution in the tetramer, is that the greater structural complexity of the larger oligomers—and hence greater density of interactions between different parts of the structure—leads to a higher likelihood that at least some of the structures will experience a shearing force, rather than an unzipping force. If so, then the unfolding force would be expected to continue to rise as the size of the aggregate increases. Amyloid fibrils have been shown to be considerably more stable mechanically than the tetramers studied here, being able to withstand forces of up to 100–200 pN (Dong, Castro et al. 2010).

4.5 α-Synuclein folding rates

Finally, we consider the kinetics of structure formation in α-synuclein oligomers. The timescale for the formation of mechanically-stable, aggregated structures, as measured from the delay time between FECs, was relatively fast: on the order of tens of seconds. We used the frequency with which structures of different size were observed to estimate the apparent folding rate, taking the total contour length change (including all intermediates) in each pull, ΔL_c^{tot}, as the size of the structure. Given the known total amount of time spent waiting for refolding to occur in all FEC measurements, the number of occurrences divided by the total refolding time yielded an estimate of the apparent folding rate. Because some transitions were observed only very rarely, to improve the statistics for the rate

estimates when examining the length-dependence of the rates in Figure 4.9, the transitions were re-binned in 15-nm increments (roughly 1/3 the length of a monomer).

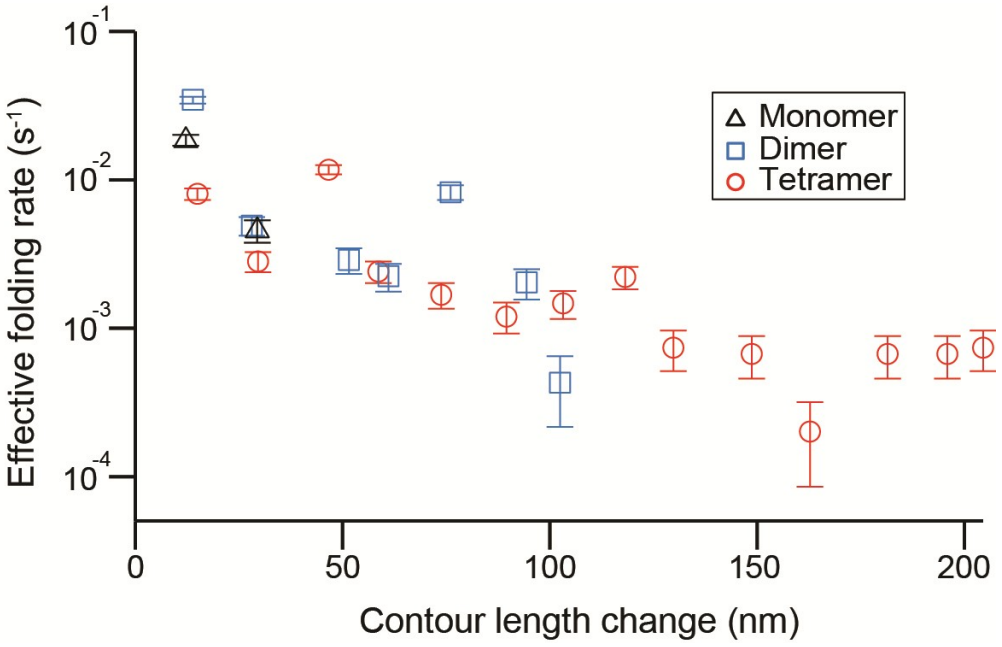

Figure 4.9 Size-dependent structure formation rates. The rate at which structures of a given total contour length change (including all intermediates) occur is similar for all constructs (tetramer: red, dimer: blue, monomer: black), but declines roughly exponentially with increasing length. Wait time at zero force was 5 sec. Rates were estimated from the occurrence frequency of specific ΔL_c values, binned in 15-nm increments to improve the statistics.

A structure larger than one monomer in length ($\Delta L_c^{tot} > 50$ nm) formed roughly once per minute (1.4×10^{-2} s^{-1}). This is faster than the oligomer-formation timescale found in fluorescence measurements (Orte, Birkett et al. 2008; Nath, Meuvis et al. 2010; Cremades, Cohen et al. 2012), likely due to a higher local protein concentration, and much faster than the typical lag phase of days during which aggregates nucleate before amyloid fibrils form (Wood, Wypych et al. 1999). We note that these apparent rates are not the true microscopic folding rates, since the apparent rates include the effects of other transitions (including dissociation). Strikingly, the apparent rate for stable structure formation decreased roughly exponentially with increasing ΔL_c^{tot}, but the rate for any given size was similar whether the structure formed in a tetramer, a dimer, or a monomer (Figure 4.9). The continuity of the apparent rates for the monomer, dimer, and tetramer suggests that the energy landscape for oligomerizaton is relatively flat, supporting the notion that the aggregation (at least in its early stages) is dominated by kinetics (Ahmad, Chen et al. 2012). Similar apparent rates were seen for different delay times between pulls, from 0.5–20 s (Figure 4.10), again underlining the picture of a random process that is not dominated by a single sequence of events during structure formation, at least at the time-scale probed.

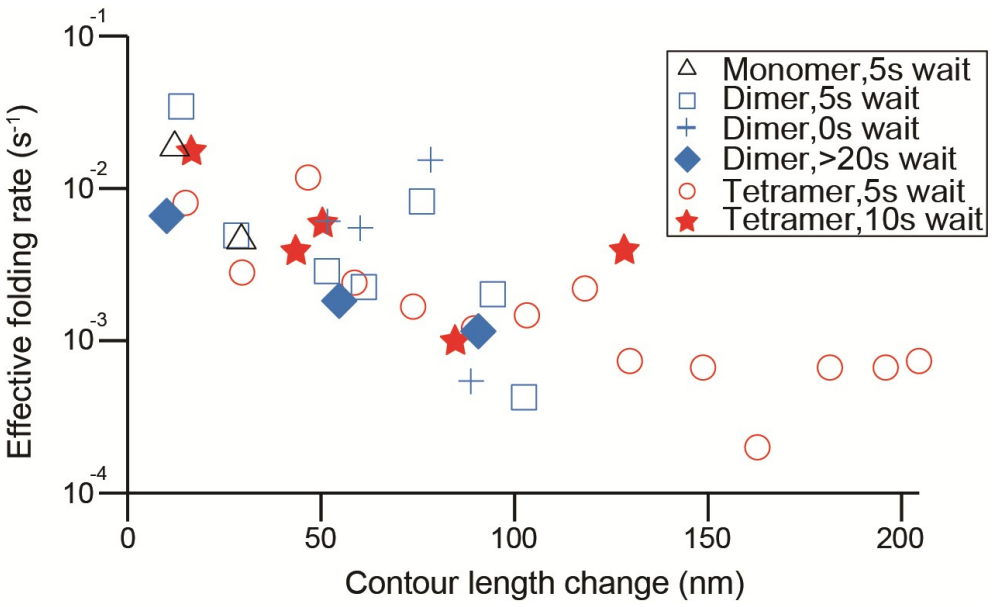

Figure 4.10 Effect of waiting time on size-dependent apparent folding rates. The frequency which structures of different size formed in the dimer and tetramer was not noticeably affected when the waiting time at zero force between each pull was changed. In addition to the data shown in Figure 4.9 are the results for dimers with 0 s waiting time (blue crosses) or over 20 s (blue diamonds) and tetramers with a 10 s waiting time (red stars). Apparent rates were estimated from the occurrence frequency of specific ΔL_c values, binned in 15-nm increments to improve the statistics. There are no data points at large ΔL_c values for the 10 s wait time due to the small data set acquired.

It is interesting to compare these rates to the folding rates for natively-structured proteins, many of which show a similar, roughly exponential decrease in rate with increasing size (Ivankov, Garbuzynskiy et al. 2003; Li, Klimov et al. 2004; Ivankov, Bogatyreva et al. 2009). Two key differences are observed: First, the apparent folding rates for small structures in α-synuclein are orders of magnitude lower than the rates for natively-folded proteins of similar size (Figure 4.11) and do not approach the folding "speed limit" as ΔL_c goes to zero (Kubelka, Hofrichter et al. 2004; Li, Klimov et al. 2004). Such low rates suggest a very rugged folding landscape for α-synuclein, qualitatively different from the funnel-shaped landscapes characteristic of natively-folded proteins (Thirumalai, O'Brien et al. 2010), as might be expected since α-synuclein is intrinsically-disordered. A second key difference is that the decrease of the rates with increasing size is much slower for the α-synuclein oligomers than for the natively-folded proteins, *i.e.* larger structures formed at a faster rate in the oligomers than would be expected based on the size-dependence of rates for natively-folded proteins. One possible explanation is that the smaller structures formed in the α-synuclein constructs are on-pathway to the larger aggregates, serving to nucleate additional structure formation within the aggregate and hence increasing the apparent formation rate of the larger structures.

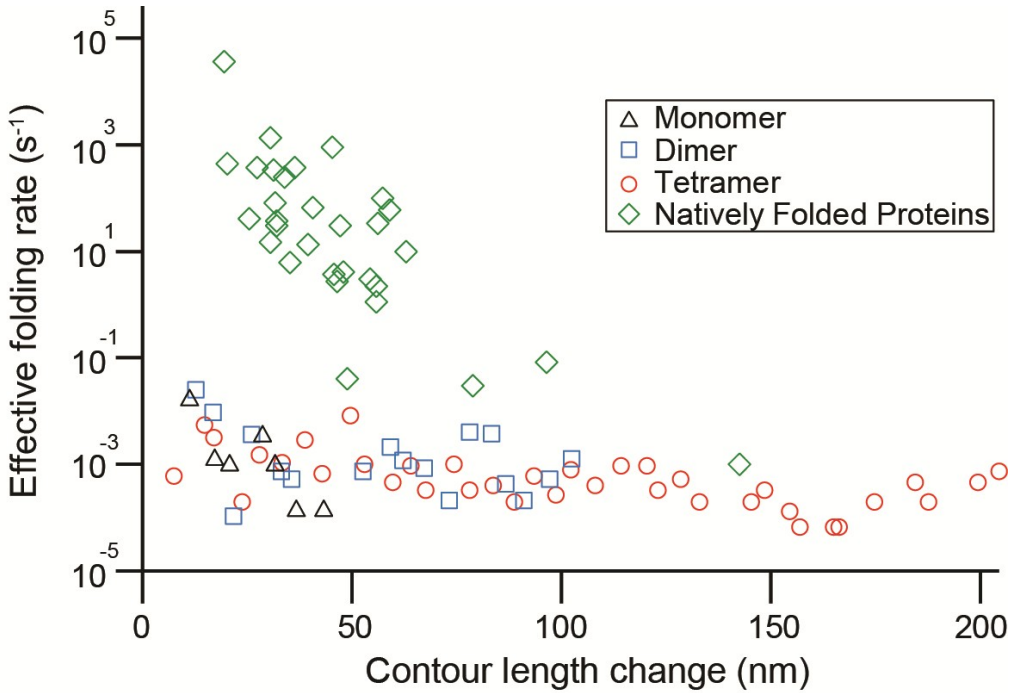

Figure 4.11 Size-dependent folding rate comparison. The apparent rate of formation of structures of different contour lengths in α-synuclein is compared to the folding rate of multi-state natively-folded proteins having different sizes taken from Ivankov et al., 2003. ΔL_c values were binned in 5-nm increments for comparison to the rates for natively-folded proteins.

To summarize, we have studied structure formation in monomers and small oligomers of the intrinsically-disordered protein α-synuclein using optical tweezers. We observed a very diverse set of mechanically-stable structures in α-synuclein, despite the fact that it is intrinsically-disordered, which were interpreted in terms of transient conformational fluctuations. The structural complexity increased as the size of the oligomers grew, both in terms of the number of different structural transitions observed and the number of folding intermediates that formed, revealing a more complex folding landscape for α-synuclein than previously reported. The unfolding force and apparent rate of structure formation both varied systematically with the size of the structural transitions (ΔL_c), pointing to changes in the structures that are able to form in the larger oligomeric constructs. These observations open a new window on the early events in the formation of aggregates of α-synuclein. More generally, they demonstrate the power of force spectroscopy as a tool for studying the diverse structures formed during protein aggregation, able to characterize the size, stability, and kinetics of species populated even at very low levels. The use of tandem-repeat protein constructs, while constraining the accessible aggregate structures due to the geometry of the construct, allows control over the size of oligomer being studied, thereby enabling a systematic exploration of how aggregates grow.

Appendix:

Protein purification and verification

The full length, human α-synuclein monomer protein from the *SNCA* gene was purchased as a GST fusion cloned into the pDEST15 plasmid. Site-directed mutagenesis (Stratagene) was conducted to make several plasmid modifications (e.g., inserting C- and N-terminal cysteine residues for handle attachments, introducing an N-terminal enterokinase cleavage site for GST-tag removal). The mutations were verified via DNA sequencing. The protein was transformed into and expressed in *E. coli* BL21-AI cells (Invitrogen), which utilizes an L-arabinose induction system. Expression cell pellets were resuspended in lysis buffer (PBS, pH 7.4 with 5 mM TCEP, 1 mM EDTA, 0.5 mM PMSF), sonicated for 60 s, centrifuged briefly before loading onto a dedicated 5ml- GST column (Qiagen), and purified using affinity chromatography by fast protein liquid chromatography (FPLC, GE Healthcare). The monomer was cleaved on-column using enterokinase protease (Novagen) overnight and eluted into PBS, pH 7.4. All purification steps were performed at 4°C.

Tetrameric α-synuclein protein was engineered to contain four copies of the 140-amino-acid sequence of human α-synuclein as a tandem repeat separated by 3-amino-acid peptide linkers: GSG, GTG, and GSG (Figure 4.1). The linkers were chosen to be short as to not introduce additional structural motifs or regions that could interfere with the protein of interest. Keeping in mind that the C-terminus of α-synuclein is naturally disordered, we assumed that the length of the linker was significant enough as to not constrain potential interactions between

protein units. Notably, the protein construct also contained a cleavable N-terminal His-tag for purification, as well as N- and C-terminal cysteines for attaching DNA handles. The ORF coding for the tetramer was designed with restriction sites for removing two of the copies of α-synuclein to create a dimer. The tetramer gene was synthesized and cloned into the bacterial expression vector pJexpress 406 (DNA2.0, Menlo Park CA), and expressed in *E. coli* C41(DE3) cells (Lucigen).

Cell pellets expressing tetrameric (or dimeric) α-synuclein were resuspended in lysis buffer A (50 mM NaH_2PO_4, 0.5 M NaCl, 6 M GdmCl, 0.5 mM PMSF, 20 mM imidazole, pH 7.4), 10 mM β-mercaptoethanol and 0.5% v/v Tween20, and sonicated for 40 s. The lysates were then centrifuged and filtered prior to loading onto a 5 mL Ni-NTA Superflow Cartridge (Qiagen). The column was previously equilibrated with 50mM NaH_2PO_4, 0.5 M NaCl, 20 mM imidazole, 6 M GdmCl, pH 7.4 using FPLC. All purification steps were again performed at 4°C. Unbound proteins were washed out with the equilibration buffer, and α-synuclein was then eluted with 50 mM NaH_2PO_4, 0.5 M NaCl, 6 M GdmCl, 500 mM imidazole pH 7.4. The 32-kDa dimer was successfully produced following restriction digests and was expressed and purified following the same protocol. Protein was finally dialyzed overnight at 4°C in an excess of PBS, pH 7.4 buffer using 3ml capacity dialysis cassettes, with a 10 kDa MWCO (Thermo Scientific). The tandem constructs presented here were designed by Iveta Sosova, as were the optimization of their expression and purification protocols. All protein was verified with SDS-PAGE and PageBlue protein staining solution (Thermo Scientific) (Figure 4.12b).

To further confirm that the proteins purified were indeed α-synuclein (or tandem repeats), Western blotting was conducted. Using either a 6xHis mAb/HRP conjugate (Clontech for the tetramer or dimer, and an α-synuclein mAb/HRP conjugate (Millipore), the proteins were verified to be the molecule of interest. Briefly, protein was run on a 12% SDS gel but not stained. The gel's protein contents were then transferred to a nitrocellulose membrane and soaked overnight in blocking buffer at 4°C. The following day, the membrane was washed and then incubated first with monoclonal antibody, washed, incubated with the conjugate antibody, and washed once more. Finally, the antibody binding was visualized using TMB stabilized substrate (Promega). The presence of α-synuclein was confirmed for all constructs.

Protein-dsDNA handle attachments

Two distinct dsDNA handles were amplified via PCR: one 2113-bp handle, labeled with digoxigenin, and the other measuring 798-bp, labeled with biotin (for the monomer, the digoxigenin-labeled handle was 1261 bp long). Each handle was functionalized with a sulfhydryl group on the opposite end. Every PCR reaction produced μg quantities of dsDNA, which was verified on 1% agarose gels. The DNA was EtOH precipitated overnight at -80C and resuspended in 15mM NaPi, pH 7.0 buffer. Following the removal of excess dNTPs, the DNA was precisely quantified using a NanoDrop UV spectrophotometer (Thermo Scientific).

α-Synuclein was then attached to the handles using previously described methods (Cecconi, Shank et al. 2008; Yu, Liu et al. 2012), employing 2,2'-

dithiodipyridine (Sigma-Aldrich) as the crosslinker between the protein and dsDNA. Maleimide-thiol reactions are commonly used due to their ability to preferentially form stable thioether bonds at neutral pH (between 6.5-7.5) (Hermanson 2008). Since DTDP utilizes thiol reactivity, the presence of reduced sulfhydryl groups was necessary to complete the reaction. We used the common reducing agent TCEP-HCl for several reasons, including its lack of thiol groups, which could potentially interfere with the reaction, and its stability even under oxidizing environments. That said, it is very acidic and also must be freshly prepared when solubilized in phosphate buffers.

Crosslinking consisted of three primary steps. First, 40uM of freshly purified protein was reduced at a 1:100 molar excess of TCEP for 1 hour, followed by three desalting rounds (7k MWCO Zeba desalting column, Thermo Scientific, equilibrated three times with 0.1M NaPi, pH 7.0) to remove surplus reducing agent. Next, an excess of DTDP was incubated with the reduced protein overnight. The following morning, the SH-modified dsDNA handles (~10μM) were reduced separately in a 200 molar excess of TCEP for 30 minutes at room temperature. Equal molar quantities of both DNA handles and the maleimide-protein sample were then mixed gently, before desalting together three times with equilibrated desalting columns. The sample was then incubated at 4°C for ~36 hours. The resulting construct was verified on a 1% native agarose gel by staining with EtBr for DNA visualization and identifying a band of the expected length (Figure 4.12c). The high handle and protein concentrations used in this reaction were critical for ensuring successful construct formation (see Figure 4.12c, lanes

3-4, where lower concentrations were used and no construct was observed). The resulting protein-DNA chimeras were incubated at ~ 100 pM with ~ 250 pM polystyrene beads (600-nm diameter labeled with avidin, 820-nm diameter labeled with anti-digoxigenin) to create the dumbbell constructs first described in Chapter 3.3 (Figure 4.12a). Dumbbells were then diluted to ~ 500 fM in 50 mM MOPS, pH 7.0, with 200 mM KCl and oxygen scavenging system (8 mU/μL glucose oxidase, 20 mU/μL catalase, 0.01% w/v D-glucose), before insertion into a sample flow cell for the optical tweezers experiments.

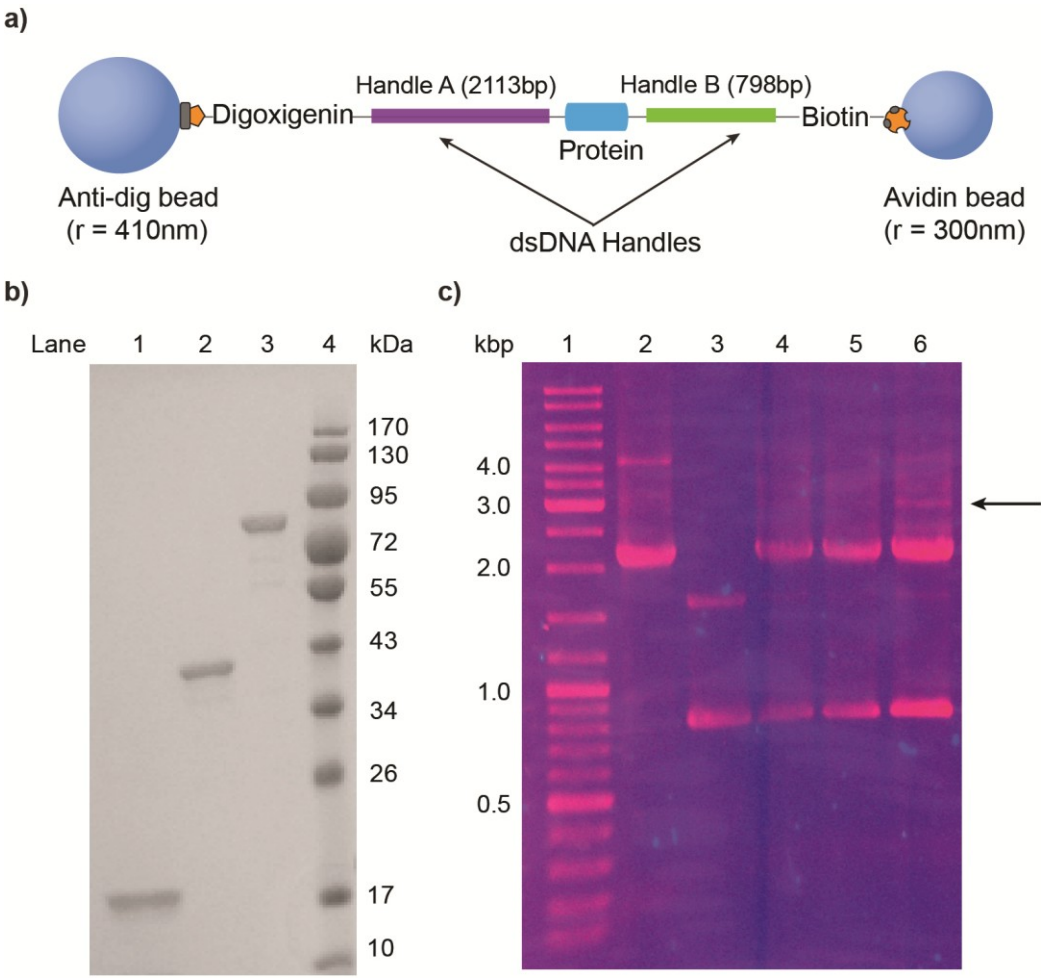

Figure 4.12 Protein-dsDNA handle attachments. a) Schematic of optical tweezers dumbbell construct, which is formed using maleimide-sulfhydryl chemistry. The terminal-functionalized handles are attached to beads via digoxigenin/anti-digoxigenin or biotin/avidin complexes. (Note: objects are not to scale). b) 12% SDS-PAGE gel of α-synuclein protein samples after affinity purification. Lane 1: monomer; lane 2: dimer; lane 3: tetramer; lane 4: protein ladder. c) 1% native agarose gel of dsDNA-protein handle attachments. Lane 1: 2113-bp Handle A monomer and dimer; Lane 2: 798-bp handle B monomer and dimer; Lane 3 – 5: handle-protein attachment reactions with different concentrations of protein and desalting steps, with arrow indicating construct in the final reaction.

Protein biochemistry assays

The ability of the tandem-repeat α-synuclein constructs to aggregate into amyloid fibers was tested by measuring ThT fluorescence similar to standard methods (Uversky, Li et al. 2001). Briefly, 50 μM tetrameric α-synuclein and 40 μM ThT dye were incubated at 37°C and shaken linearly at 20 and 30 Hz for 10 days. The ThT fluorescence, excited at 430 nm, was measured at 485 nm in a microplate reader (Gemini EM, Molecular Devices). As seen in Figure 4.13, the ThT fluorescence displayed the typical sigmoidal rise that occurs during amyloid formation after a lag time of approximately two days.

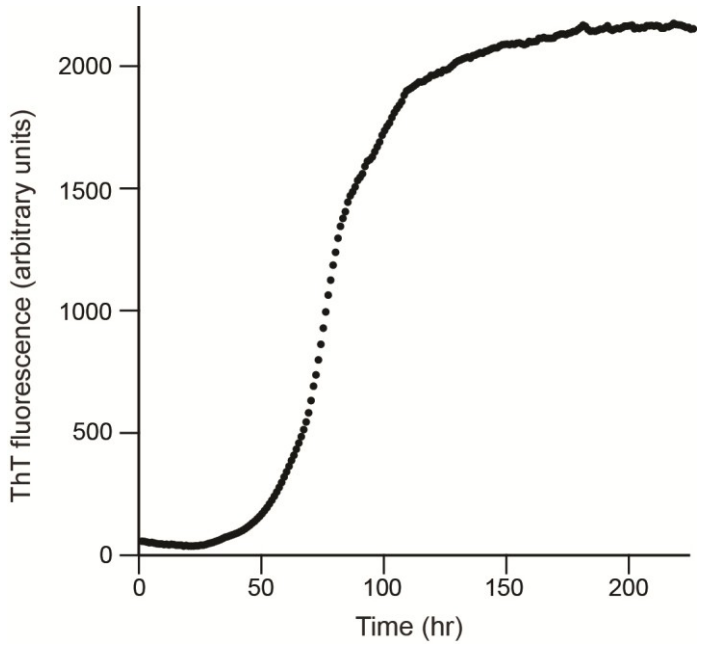

Figure 4.13 ThT monitored tetramer aggregation. α-Synuclein tetramers aggregated over the course of several days to form amyloid fibrils, as seen by the increase in ThT fluorescence.

Further, to assess the average structural content of the tandem-repeat oligomers, we measured the CD spectra of the three protein constructs (Jasco J–810 CD/ORD spectrometer) at a concentration of 10 μM in 10 mM phosphate buffer pH 7.0, using a 1-mm path length and subtracting the baseline spectrum of the buffer. All three spectra, shown in Figure 4.14, were characteristic of a largely disordered protein and similar to results found previously for α-synuclein (Uversky, Li et al. 2001).

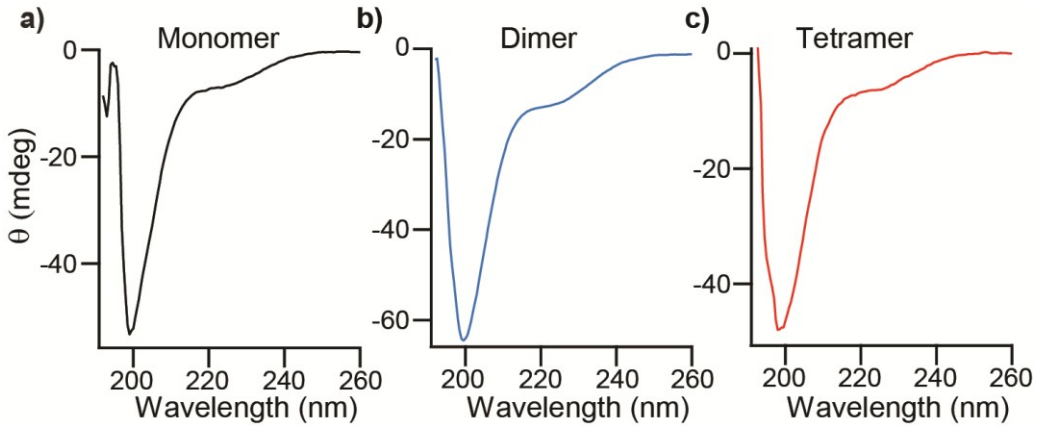

Figure 4.14 Comparison of CD spectra of monomer, dimer, and tetramer. CD measurements of the α-synuclein a) monomer, b) dimer, and c) tetramer all show spectra characteristic of unstructured proteins.

Force ramp experiment analysis details

Each single-molecule tether was pulled on under various conditions (including loading rate and wait time between pulls) until the construct broke, generally for a total of several hundred pulls. As discussed briefly in Chapter 3, each individual FEC was then analyzed by fitting the folded, unfolded and any present intermediate portions of the curve with a series of modified WLCs: one for the protein and one for the handle, such that

$$F(x) = \frac{k_B T}{L_p^{DNA}}\left[\frac{1}{4}\left(1 - \frac{x}{L_c^{DNA}} + \frac{F}{K^{DNA}}\right)^{-2} - \frac{1}{4} + \frac{x}{L_c^{DNA}} - \frac{F}{K^{DNA}}\right] + \frac{k_B T}{L_p^{protein}}\left[\frac{1}{4}\left(1 - \frac{x}{L_c^{protein}} + \frac{F}{K^{protein}}\right)^{-2} - \frac{1}{4} + \frac{x}{L_c^{protein}} - \frac{F}{K^{protein}}\right]. \quad (4.2)$$

Recall that L_p is the persistence length, K is the elastic modulus, L_c is contour length and k_BT is the thermal energy. For the FEC section corresponding purely to the stretching of the dsDNA (where any stable structures in the protein are presumed to be folded), the handle parameters were allowed to float and the protein values were fixed ($L_p \sim 0.8$ nm (Kim, Zhang et al. 2010), $K \sim 2000$ pN (Pauling and Corey 1951), and $L_c^{protein,f} = 0$, since the number of amino acids still unstructured was unknown). Typical fitting values for the dsDNA handles were: $L_p \sim 50$ nm, $K \sim 1500$ pN, and $L_c \sim 1000$ nm.

The next portion of the FEC was then fit, corresponding to the stretching of the handle in addition to the unfolded protein. By holding the handle and protein values from the portion of the FEC prior to the rip, only one variable was

free, $L_c^{protein,unf}$. The total protein ΔL_c was then calculated by taking the difference of L_c^{unf} - L_c^{f}. Since L_c = 0.36 nm per amino acid (Pauling and Corey 1951), the total number of residues involved in the structural unfolding event could be calculated. In these experiments, the resolution of the ΔL_c value obtained from fitting individual FECs was ~ 2 nm, as found from the standard deviation of the distribution of WLC fit results to multiple FEC measurements of a reference protein with a known structure, PrP (Yu, Liu et al. 2012).

Additional interpretation of unfolding distances

We have catalogued only some of the various different structures that have been observed or proposed for α-synuclein and listed the associated contour length changes expected for unfolding (Table 4.1). Contour length estimates were based on the number of amino acids involved in the structures and the distance between the points at which force is applied to the protein (as estimated from the structural models). Monomers of α-synuclein have been observed to form α-helical structures at the N-terminus under certain conditions (Davidson, Jonas et al. 1998; Jao, Der-Sarkissian et al. 2004; Ulmer, Bax et al. 2005): an extended α-helix of ~ 90 aa, or two broken α-helices – one ~ 35 aa long and the other ~ 48 aa, folded into a helix-turn-helix "hairpin". Since helices have a length of 0.15 nm/aa (Jao, Der-Sarkissian et al. 2004), unfolding these α-helices would give rise to $\Delta L_c \sim 20$ nm for the extended helix, ~ 7 and 10 nm for the short helices individually if not folded into a hairpin, and ~ 30 nm for the full hairpin. Suggestively, transitions with ΔL_c ~ 10, 20, and 30 nm are all seen for each of the constructs (monomer, dimer, and tetramer). Interestingly, the unfolding transitions

with $\Delta L_c \sim 30$ nm seen in FECs of the monomer match the results of a previous AFM study ($\Delta L_c \sim 28$ nm), which the authors suggested arose from β-sheet structures in the N-terminus (Sandal, Valle et al. 2008).

The structure of the amyloid fibril formed by α-synuclein (Vilar, Chou et al. 2008) involves each monomer forming a 5-strand β-sandwich, with individual sandwiches then aligned side by side. Many different ΔL_c values could be expected from unfolding different components of such β-sandwiches, whether the sandwiches are in isolation or stacked in multimers. For example, unfolding of different numbers of β-strands in one or more sandwiches would lead to $\Delta L_c \sim 8$–10 nm for 2–3 strands or ~ 16–19 nm for 4–5 strands from one or two sandwiches, separation of neighboring sandwiches would produce $\Delta L_c \sim 33$–34 nm, and unfolding of one complete monomer from a stack of sandwiches would produce $\Delta L_c \sim 50$ nm. Some of these unfolding transitions are illustrated in Figure 4.6. Several different transitions could in principle occur in combination, depending on the size of the construct (dimer or tetramer), giving rise to the large number of ΔL_c values listed in Table 4.1. Intriguingly, many of these values coincide with transitions observed in the FECs, suggesting that amyloid-like structural motifs may form even in the smallest oligomers.

A tetrameric native structure for α-synuclein was recently reported (Bartels, Choi et al. 2011; Wang, Perovic et al. 2011). The structural model involves monomers folded into helix-hairpins similar to the micelle-bound structure, stacked in parallel. As for the β-sandwich structure, multiple ΔL_c values could be expected from unfolding different combinations of structural

components. Many of these values would be degenerate with the ΔL_c values expected from other structures (*e.g.* 50 nm for removing a complete monomer, ~ 30 nm for unfolding a single helix-hairpin, ~ 18–19 nm for separating two neighboring helix-hairpins). ΔL_c for complete unfolding of a tetramer, trimer, or dimer are quite distinct, however, being ~ 185 nm, ~ 135 nm, and ~ 85 nm, respectively.

As a final note, α-synuclein has been shown to bind non-specifically to dsDNA (Hegde and Rao 2003; Cherny, Hoyer et al. 2004). Hence it is possible that the sawtooth patterns in the FECs reflect dissociation of α-synuclein bound to the DNA handles rather than cooperative unfolding transitions. If this were the case, however, we would expect to observe a smooth distribution of ΔL_c values, since non-specific interactions will not give rise repeatedly to the same discrete values. The fact that the actual distribution is highly peaked indicates that the transitions do indeed arise from protein structures, rather than protein-handle interactions.

Chapter 5. Single-Molecule Force Spectroscopy of Rapidly-Fluctuating, Marginally-Stable Structures in the Intrinsically Disordered Protein α-Synuclein

In the previous chapter, we only characterized the stable, yet transient structures that α-synuclein monomer, dimer and tetramer oligomers samples formed in approximately 25% of all measured FECs. In this chapter, we analyze the behavior observed in the majority of our force ramp experiments: rapid fluctuations at low force that arise from the folding of two different classes of structure that are only marginally stable. The first experimentally reconstructed energy landscape for an IDP is also characterized via the force-dependent kinetics derived from correlation analysis of the extension trajectories. A version of this chapter has been published: Solanki, A.*, Neupane, K.*, Woodside, M.T. "Single-molecule force spectroscopy of rapidly-fluctuating, marginally-stable structures in the intrinsically-disordered protein α-synuclein." *Physical Review Letters* 112(15): 158103 (2014). *Co-first authors. AS provided protein and constructs, some of which were designed by Dr. Iveta Sosova. AS also performed the experiments and analyzed data with KN.

5.1 Unique equilibrium folding behavior in α-synuclein FECs

As presented in Chapter 4, FECs of monomer, dimer and tetramer α-synuclein constructs occasionally displayed discrete rips, in which the extension increased abruptly concomitant with a sudden decrease in the force (Figure 5.1a-c, black). Such behavior, characteristic of the cooperative unfolding of

mechanically-stable structures, was qualitatively similar to that observed previously in AFM measurements (Sandal, Valle et al. 2008; Yu, Malkova et al. 2008; Hervas, Oroz et al. 2012). It reflects the formation, via transient conformational fluctuations, of metastable structures with sufficiently large energy barriers to withstand high unfolding forces during rapid force ramps, as shown previously (Neupane, Solanki et al. 2014). These FECs were well fit on each side of the rip by a polymer elasticity model consisting of two extensible worm-like chains (WLCs) in series (Equation 5.1), one for the DNA handles and the other for the unfolded protein (Yu, Liu et al. 2012). Here, we are interested instead in the majority of FECs that showed no discrete rips, but rather a monotonic rise in force with extension (Figure 5.1a-c, cyan).

Interestingly, the FECs without discrete unfolding rips did not fit well to a simple WLC model, as would be expected for a non-interacting (unfolded) polymer. Averaging the FECs for each set of curves measured on a single molecule to reduce the noise (Figure 5.1d-f, cyan), fits to the WLC expected for the fully-unfolded protein (Figure 5.1, dashed red line) yielded residuals with systematic deviations from zero (Figure 5.1d-f inset, red). The residuals were largest for the tetrameric construct (Figure 5.1f) and smallest for the monomer (Figure 5.1d), but in all cases non-random, indicating an incomplete fit. The deviation from pure WLC behavior produced a "shoulder" feature in the force range ~ 2-8 pN. Such deviations are not typically present in FECs of unfolded proteins (Schlierf, Berkemeier et al. 2007; Stigler, Ziegler et al. 2011; Yu, Liu et al. 2012), nor were they observed in control measurements of the DNA handles

alone, indicating that they are a property specifically of α-synuclein. Because the extension at a given force within the shoulder feature is lower than would be expected in a WLC, attractive interactions must be forming as we allow the molecules to refold, making the protein more compact than expected, even though no cooperative unfolding transitions could be discerned directly. The same characteristic shape was observed for refolding FECs as for unfolding FECs (Figure 5.2, black), indicating reversible (*i.e.* equilibrium) behavior.

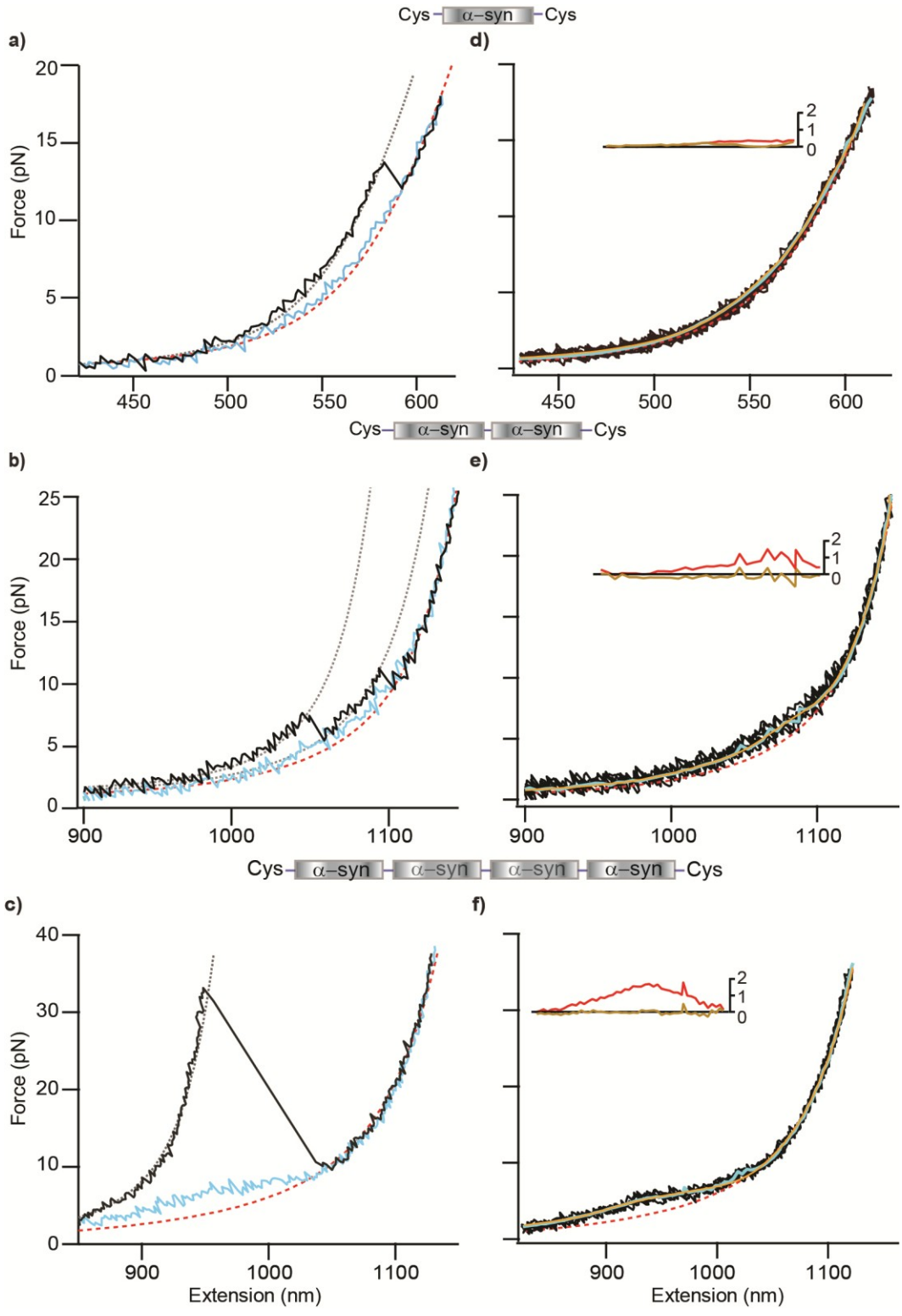
Cys α–syn Cys
a)
d)
Force (pN)
20
15
10
5
0
450
500
550
600
2
1
0
Cys α–syn α–syn Cys
b)
e)
25
20
15
10
5
0
900
1000
1100
Cys α–syn α–syn α–syn α–syn Cys
c)
f)
40
30
20
10
Extension (nm)

Figure 5.1 Force spectroscopy of α-synuclein. a) FECs of α-synuclein monomers sometimes revealed discrete unfolding rips (black), but usually showed a monotonic rise in force with extension without obvious rips (cyan). Similar behavior was seen for α-synuclein b) dimers and c) tetramers. FECs without rips nevertheless deviated from WLC behavior (red). d–f) FECs without discrete rips were averaged (cyan) and compared to polymer models. Data did not fit a non-interacting WLC model (red; residuals in inset), but did fit a model incorporating rapid structural fluctuations, Equation 5.1 (yellow; residuals in inset).

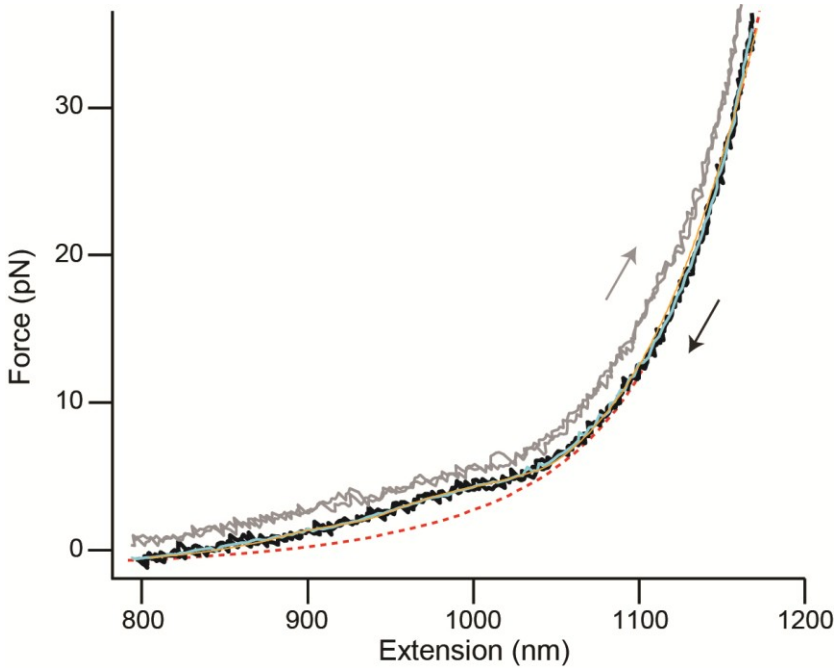

Figure 5.2 Refolding tetramer FECs. Refolding FECs measured for tetramers (black) retraced the same trajectory found in unfolding FECs for the same molecule (grey). Refolding FECs are offset by 1pN for clarity. Fitting the refolding FECs to Equation 5.1 yields the same results as found from fits of the unfolding FECs.

Motivated by qualitatively similar shoulder-like features seen in FECs of the ultra-fast folding protein villin (Žoldák, Stigler et al. 2013) and of mRNA transcripts that could form an ensemble of small, fast-folding hairpins (Dalal, Larson et al. 2006), we interpreted the shoulder features in terms of rapid, quasi-equilibrium unfolding/refolding of structures that are only marginally stable. In this picture, the fact that deviations from WLC behavior occur only at low force results from the marginal stability of the structures. In turn, the reversibility of the FECs and apparent lack of cooperativity both result from the fast kinetics. In particular, whereas the sharp, sawtooth-like rip patterns commonly seen in force spectroscopy measurements arise from folding and unfolding rates that are slow compared to the FEC step dwell time (1–5 ms), rates that are fast compared to the dwell times result in a quasi-equilibrium average over multiple transitions. Such averaging will produce a FEC that appears to move smoothly, as the force is increased, between the WLC curve expected for the folded state and that expected for the unfolded state, without any detectable rips (Žoldák, Stigler et al. 2013). In this case, the shoulder in the FEC will have a shape defined by the force-dependent probability that the structure is unfolded, $P_u(F)$.

5.2 Shoulder feature fits and analysis

To test this picture quantitatively, we constructed a minimal model for the protein, assuming that it can form some number of independent structures via two-state transitions in rapid equilibrium. The extension as a function of force was then taken as the sum of the extension of the handle, $x_H(F)$ (obtained by

inverting Equation 5.1 for the DNA), the extension of the fully-unfolded part of the polypeptide chain, $x_{PU}(F)$ (since the shoulder feature does not necessarily involve all residues), and the sum of the extensions of the various protein structures that are unfolding concurrently in rapid equilibrium:

$$x(F) = x_H(F) + x_{PU}(F) + \sum_{i=1}^{n} N_i\left[P_u^i(F)\Delta x_i(F)\right]. \qquad (5.1)$$

Here, n is the number of different types of transitions having distinct unfolding properties (each assumed to act as a two-state system) and $\Delta x_i(F)$ is the change in extension at a particular unfolding force. In these fits, the WLC parameters for the DNA handles and fully-unfolded portion of the protein were fixed at values found from fitting the discrete rips, as described previously in Chapter 4. The presence of a consistently unfolded region of the protein chain was attributed to the fact that the total extension in the shoulder region did not account for all available protein residues. To fit the shoulder feature, only the contour length change (ΔL_c) and equilibrium unfolding force ($F_{1/2}$) for the unfolding structures were allowed to vary. The model allows for the possibility that several structures might share similar unfolding properties, with the parameter N_i denoting the number of instances of each type of transition (note that this means the model cannot distinguish between distinct structures that have similar unfolding properties). If the structures involve only intra-monomer interactions, N_i should scale as the number of monomers in the protein construct. Recall that for two-state unfolding in equilibrium,

$$P_u(F) = [1 + \exp\{((F_{1/2} - F)\cdot\Delta x(F))/k_BT\}]^{-1}, \quad (3.12)$$

where $F_{1/2}$ is the force at which the structure has 50% probability of being unfolded and $\Delta x(F)$ is the extension change upon unfolding at force F. From this, the contour length change upon unfolding, ΔL_c, can be found by inverting the equation for an enthalpic WLC fit (Equation 3.7) (Woodside, Behnke-Parks et al. 2006). Each type of transition in the model is thus characterized by unique $F_{1/2}$ and ΔL_c values.

To determine the variant of the model that best explained the observed shoulder features, all individual FECs from a given molecule that did not contain discrete unfolding rips were aligned (to correct for any instrumental drift) and averaged. The averaged data were then fit to Equation 5.1, using the standard error on the mean as the uncertainty in the data. We applied three statistical tests to determine whether the fit of a particular model variant (n, N_i combination) was acceptable. First, a sum-of-squares lack-of-fit test was applied to each fit, to test if the model variant was able to fit the data. The reduced χ^2 of the fits was also used, as a second check to ensure that an acceptable fit could be achieved. If a particular model variant failed either of these tests, then that variant was judged inappropriate for describing the data, and the next variant with the least number of extra fitting parameters was tested. If a model variant passing these preliminary tests, we next examined the residuals to the fit, to determine if they were random or systematic, using the Wald-Wolfowitz runs test. The model variant with the fewest fitting parameters that passed all tests was chosen as the minimal model required to account for the observations.

Simple WLC fits as in Figure 5.1 (red) were tested first and they failed all three of the tests for goodness of fit. The simplest variant of Equation 5.1, with only a single transition per monomer ($n = 1$, $N_1 = 1$ per monomer), was tested next. This model failed the lack-of-fit test, indicating that it, too, was inadequate to explain the data (an example of an attempted fit is shown in Figure 5.3, red). Indeed, such a simplistic result would have been surprising for a nominally-disordered protein. The same failure of the lack-of-fit test was found for the variants of Equation 5.1 involving an ensemble of transitions having similar $F_{1/2}$ and ΔL_c (i.e., keeping $n = 1$ but allowing $N_1 > 1$ per monomer; illustrated for the same example in Figure 5.3, green), and the variant with two distinct but unique transitions per monomer ($n = 2$, $N_i = 1$/monomer for each transition; Figure 5.3, cyan). These variants of the model also failed the Wald-Wolfowitz runs test, showing systematic deviations from randomness in the residuals to the fits. The variant with three distinct but unique transitions per monomer ($n = 3$, $N_i = 1$/monomer for each transition: 6 free parameters) failed the runs test, too.

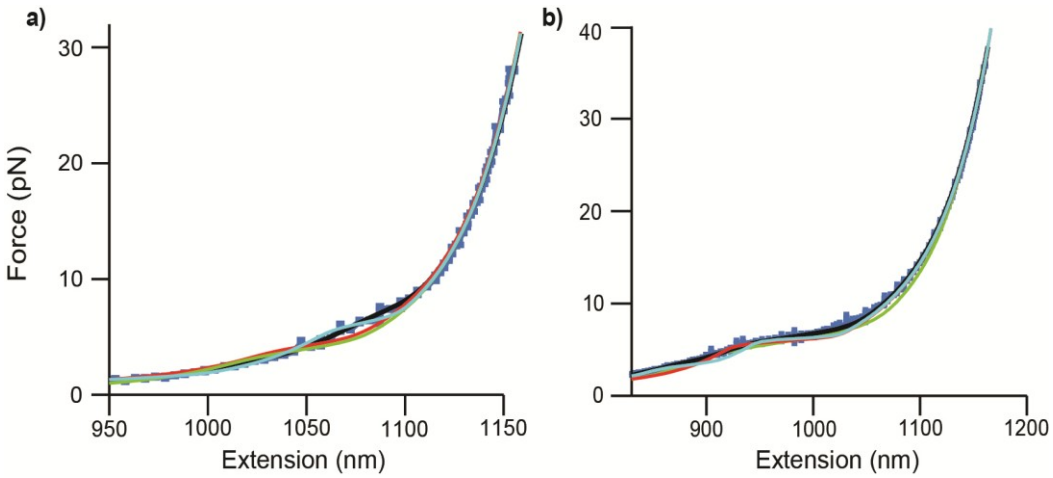

Figure 5.3 Testing models of the shoulder feature. The averages of the FECs without discrete unfolding transitions from a single a) dimer and b) tetramer molecule (blue) are not well fit by Equation 5.1 if the simplest variant, containing only a single type of transition per monomer (*i.e.*, $n = 1$ and $N_1 = 1$ per monomer), is assumed (red). The variant assuming two different structures per monomer ($n = 2$ and $N_1 = 1$ per monomer) also did not fit (cyan), nor did the variant in which there was a single ensemble of transitions having similar $F_{1/2}$ and $\Delta x(F)$ (green). The simplest model fitting all the data (black) involved two such ensembles, each occurring twice per monomer ($n = 2$ and $N_1,N_2 = 2$ per monomer).

The variant of Equation 5.1 requiring the fewest fitting parameters but still passing all three tests had $n = 2$ and $N_i > 1$ per monomer (again 6 parameters), specifically $N = 2$ per monomer for each type of transition (Figure 5.3, black). More complex variants of the model with additional fitting parameters, such as $n = 4$ and $N_i = 1$/monomer for each transition, also passed all three tests for goodness of fit. However, adding free parameters raises the possibility that the model might over-fit the data. Since over-fitting leads to more variable results for the fit parameters, we examined the standard deviations for the distributions of

fitting parameters, comparing those from the simplest model to pass the goodness-of-fit tests to those from models with additional parameters. The standard deviations were found to increase significantly in the models with more parameters, especially for the ΔL_c results, suggesting over-fitting and confirming the identification of the minimal model.

Therefore, we fit the shoulder features with $n = 2$ and $N_i > 1$, indicating the presence of two different types of transitions with distinct $F_{1/2}$ and ΔL_c values, each consisting of an ensemble of structures having similar unfolding properties (Figure 5.1d–f, yellow; Figure 5.3). Independent fits to the tetramer and dimer FECs yielded the same results for the properties of the two classes of transitions: one class had ΔL_c ~15 nm and $F_{1/2}$ ~ 3-4 pN (denoted type 1), the other, ΔL_c ~ 8 nm and $F_{1/2}$ ~ 7 pN (denoted type 2). N_1 and N_2 were twice as large (within error) for the tetramer as for the dimer, as would be expected from the scaling of the protein lengths, with the fit values indicating two transitions of each type per monomer. The shoulder feature for the monomer was often too small to provide reliable fits, however the data were fully consistent with the model obtained from the fits to the dimers and tetramers (Figure 5.4). The fit results are listed in Table 5.1. The linearity of the extension deviations from a simple WLC model with the number of tandem repeats, seen in the fitting results by the fact that the same number of transitions of each type were required per monomer to fit the data for both dimers and tetramers, implies that the transitions involved primarily structures arising from intra-domain interactions. Nevertheless, it is possible that

some of the transitions in the two ensembles (type 1 and/or type 2) might involve interactions between the domains in the tandem repeats.

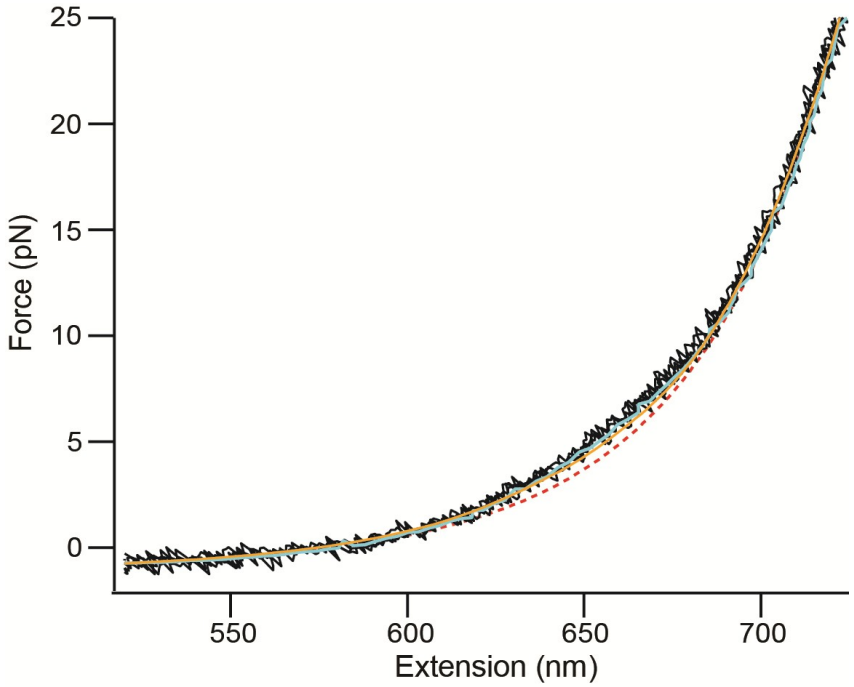

Figure 5.4 Monomer shoulder model fit. A monomer FEC is compared to the shape of the FEC predicted using the best-fit model parameters obtained from fitting the dimer and tetramer FECs. The measured (cyan) and predicted (yellow) FECs agree well, showing that the monomer FECs are consistent with the minimal model of the shoulder found from dimer and tetramer data.

Table 5.1 Fits of the FEC shoulder features to Equation 5.1. Fit parameters are the same for both dimer and tetramer, but there are twice as many transitions in the latter. Uncertainties represent standard errors.

Sample	N^1	ΔL_c^1 (nm)	$F_{½}^1$ (pN)	ΔG^1 (k_BT)	N^2	ΔL_c^2 (nm)	$F_{½}^2$ (pN)	ΔG^2 (k_BT)
Dimer	4 ± 1	14 ± 2	3.6 ± 0.4	1.5 ± 0.4	3 ± 1	9 ± 1	6.5 ± 0.7	2.4 ± 0.5
Tetramer	9 ± 1	15 ± 2	3.9 ± 0.4	1.8 ± 0.5	8 ± 1	8 ± 1	7.2 ± 0.5	2.7 ± 0.6

5.3 Kinetic analysis using autocorrelation analysis

To confirm that the shoulder features involved rapid structural transitions, as assumed in the model, we measured FECs for the tetramer without averaging the data at each step and investigated the kinetics of the extension fluctuations at different points in the curves. The autocorrelation of the molecular extension, g(τ), was computed at each force from the FEC data binned in increments of 0.2 pN using:

$$\mathrm{g}(\tau) = \frac{1}{N}\sum_{i=0}^{N-1} x(t_i)x(t_i+\tau), \qquad (5.2)$$

where N is the total number of data points in an extension trajectory, τ is the delay time; and $x(t_i)$ is the extension at time t_i. A 5 ms dwell time was used for these measurements, during which time we identified many unfolding/refolding traces in a short enough range of extensions where the force remained essentially constant. For reference, we repeated these calculations for FECs measured using DNA handles alone, without any protein present (Figure 5.5, circles). At all force values, the autocorrelation of the DNA handle extension showed a single-exponential decay, with a time constant τ near 50 μs as expected for the handle

and bead dynamics, which determine the time resolution of our trap (Neupane, Ritchie et al. 2012; Yu, Gupta et al. 2012). At forces well above the $F_{1/2}$ fitting values, where no structural transitions should occur, the autocorrelation for handles plus protein was indeed indistinguishable from that for handles alone (Figure 5.5a). In the range 2-8 pN, however, the autocorrelation for the extension of the handles plus protein revealed an additional component in the exponential decay, which had a force-dependent time constant (Figure 5.5b, triangles). This additional time constant was on the order of ~200–600 μs, slower than the time resolution of the instrument but faster than the step dwell time, consistent with the model described above.

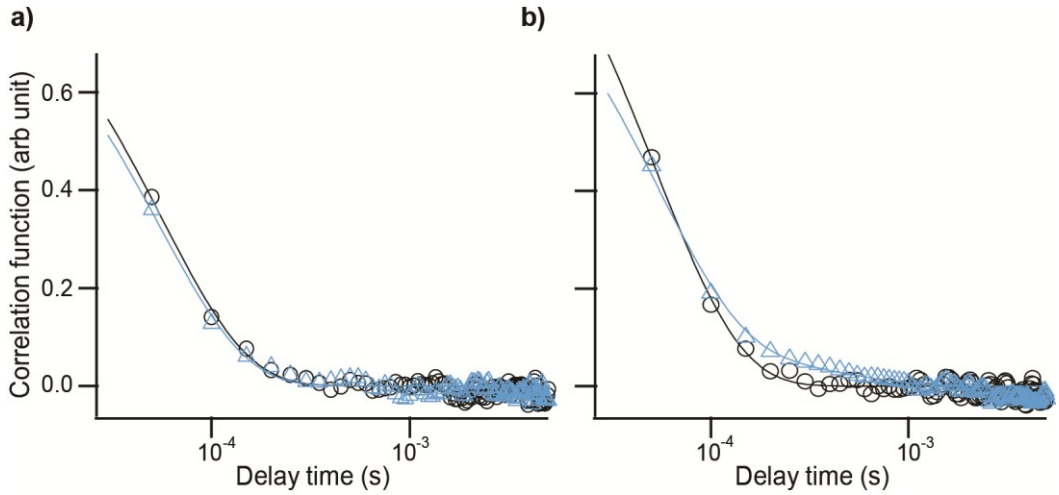

Figure 5.5 Autocorrelation analysis of α-synuclein tetramer. a) The extension autocorrelation of a tetramer construct in the high-force range (> 8pN, triangles) demonstrates a single-exponential decay, and is identical to the result for a construct containing DNA handles only, without protein (circles). b) In the mid-force range (2-8 pN), where the shoulder feature is present in FEC data, the extension autocorrelation exhibits a single-exponential decay for the handle only (circles), but a double-exponential decay for the protein construct (triangles) indicating the presence of an additional mode corresponding to structural transitions in the protein.

The correlation time constant arising from a structural transition, τ, can be related to the microscopic rates for folding (k_{fold}) and unfolding (k_{unf}) by $\tau(F) = [k_{fold}(F) + k_{unf}(F)]^{-1}$, where the microscopic rates are in turn related to the occupancies of the unfolded (P_u) and folded (P_f) states by $P_f/P_u = k_f/k_u$. We calculated P_u and P_f from the FEC fits (Table 5.1), finding that type 1 transitions dominated the occupancies in the range 1–4 pN (Figure 5.6, grey), whereas type 2 transitions dominated in the range 6–8 pN (Figure 5.6, black). The microscopic

rates for the transitions were then found from τ(*F*) as above, using the appropriate force ranges for type 1 (Figure 5.7a) and type 2 (Figure 5.7b) transitions. Notably, the force-dependence of the rates was well fit in each case by the landscape model of Dudko *et al.* (Dudko, Hummer et al. 2008), which describes transition kinetics in terms of the underlying energy landscape governing the conformational dynamics. Fitting the microscopic rates to Equation 3.19, with ν=2/3, (Figure 5.7a-b, solid lines)

$$k(F) = k_{unf}\left(1 - \frac{\Delta \mathrm{x}^{\ddagger} F}{\Delta \mathrm{G}^{\ddagger}}\nu\right)^{\frac{1}{\nu}-1} exp\left\{\Delta \mathrm{G}^{\ddagger}\left[1 - \left(1 - \frac{F\Delta \mathrm{x}^{\ddagger}}{\Delta \mathrm{G}^{\ddagger}}\nu\right)^{\frac{1}{\nu}}\right]\right\} \quad (3.19)$$

we found the distance to the energy barrier for the transition, $\Delta x^{\ddagger}$, the barrier height, $\Delta G^{\ddagger}$, and the transition rate at zero force, k_0 (Table 5.2).

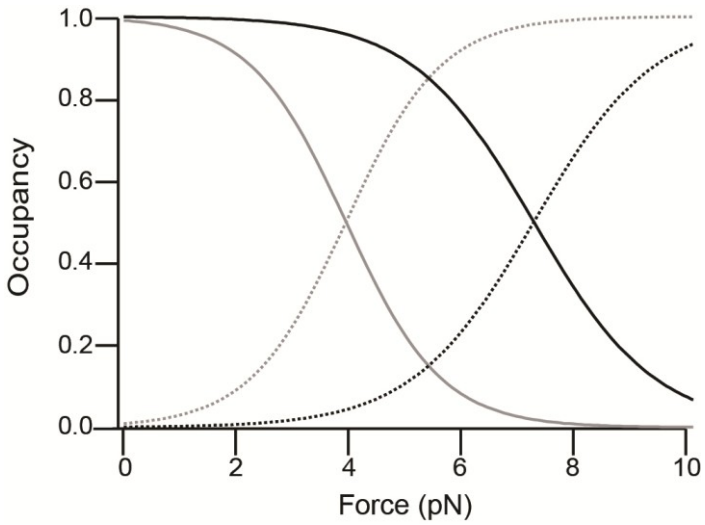

Figure 5.6 Occupancies of type 1 and type 2 structures. Folded (solid line) and unfolded (dashed line) occupancies for the two types of transitions (grey: type 1, black: type 2).

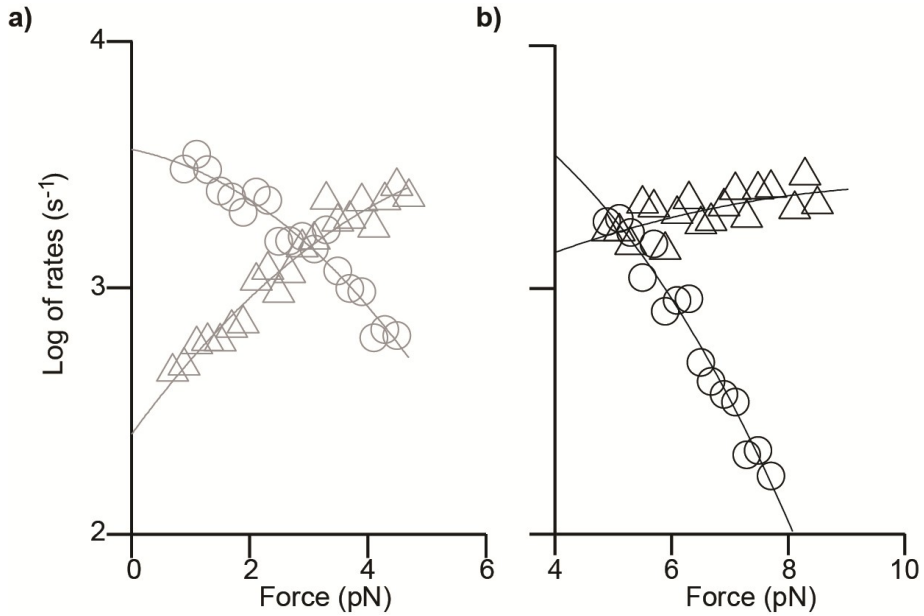

Figure 5.7 Microscopic rates for two α-synuclein transitions. a,b) The microscopic rates for tetramer unfolding (triangles) and refolding (circles) of type 1 (grey) and type 2 (black) transitions are well fit by Equation 3.19, yielding parameters describing the energy landscape.

Table 5.2 Energy landscape parameters obtained from fitting rates to Equation 3.19.

Parameters	k_{u0} (s^{-1})	k_{f0} (s^{-1})	$\Delta G_u^{\ddagger}$ (k_BT)	$\Delta G_f^{\ddagger}$ (k_BT)	$\Delta x_u^{\ddagger}$ (nm)	$\Delta x_f^{\ddagger}$ (nm)
Type 1	260 ± 50	3700 ± 300	4.4 ± 0.9	1.0 ± 0.5	3.6 ± 0.7	1.5 ± 0.6
Type 2	500 ± 100	8000 ± 1000	3.4 ± 0.5	0.5 ± 0.4	1.7 ± 0.9	1.0 ± 0.7

5.4 Free energy calculations

The equilibrium free energies for the structural transitions were calculated in three different ways. First, we estimated the free energies of the structures, ΔG, from the values for $F_{1/2}$ and $\Delta x(F_{1/2})$ obtained from the WLC fits of the FECs (Table 5.1) using

$$\Delta G = F_{1/2}\Delta x(F_{1/2}) - [\Delta G(x)]_{stretch}, \tag{5.3}$$

where $[\Delta G(x)]_{stretch}$ is the energy required to stretch the unfolded protein at $F_{1/2}$. Since the system is equilibrium, $[\Delta G(x)]_{stretch}$ is found by integrating the WLC curve for the protein from 0 to $F_{1/2}$ (Tinoco Jr. and Bustamante 2002), such that

$$\Delta G_{stretch}(x) = \frac{k_B T}{L_p}\frac{L_c}{4\left(1-\frac{x}{L_c}\right)}\left[3\left(\frac{x}{L_c}\right)^2 - 2\left(\frac{x}{L_c}\right)^3\right]. \tag{5.4}$$

Second, we calculated the free energies from the difference between the barrier heights for unfolding and refolding, $\Delta G = \Delta G^{\ddagger}_u - \Delta G^{\ddagger}_f$. Third, the free energies were calculated from the ratio of the folding and unfolding rates at zero force, $\Delta G = -k_B T \ln(k_f/k_u)$. The values found from the first method are listed in Table 5.1;

the values from the other methods are listed in Table 5.3. In each case, good agreement was found, indicating that all the kinetic and equilibrium fits are consistent with one another. We also validated the assumption that each of the transitions is two-state, by showing that the sum of the distances to the barrier from the folded and unfolded states is equal in each case to the total extension change upon unfolding found from the FEC fits (Table 5.3).

Table 5.3 Comparison of results from different methods. Free energy and extension changes for the two transition types obtained from different approaches.

Parameters	**Δx (nm)**	**Δx (nm)**	**ΔG ($k_B T$)**	**ΔG ($k_B T$)**
Method	**$\Delta x_u^{\ddagger} + \Delta x_f^{\ddagger}$**	**ΔL_c converted at $F_{½}$**	**$k_B T \cdot \ln(k_{f0}/k_{u0})$**	**$\Delta G_u^{\ddagger} - \Delta G_f^{\ddagger}$**
Type 1	5.1 ± 0.9	4.6 ± 0.6	3 ± 2	3 ± 1
Type 2	3 ± 1	3.8 ± 0.7	3 ± 2	2.9 ± 0.6

5.5 Structural interpretation of shoulder feature

What might these transitions represent structurally? Remember that although α-synuclein is largely disordered, structural and computational studies have found that it forms a condensed state, with long-range interactions between the negatively-charged C-terminal region and both the positively-charged N-terminal region and the hydrophobic central NAC region, which have been postulated to inhibit aggregation in the native state (Bertoncini, Jung et al. 2005; Dedmon, Lindorff-Larsen et al. 2005). Evidence was also found for interactions between the two halves of the N-terminal region, and a hydrophobic cluster at the C terminus (Bertoncini, Jung et al. 2005; Dedmon, Lindorff-Larsen et al. 2005).

Although we cannot directly identify the structures that form in our measurements, our observations are certainly consistent with the picture of a collapsed, molten-globule-like state held together by long-range contacts. From the ΔL_c values for the transitions (listed in Table 5.1), as many as 120 residues are sequestered within each monomer by the marginally-stable structures, similar to the distance between the longest-range contacts found previously (Bertoncini, Jung et al. 2005; Dedmon, Lindorff-Larsen et al. 2005), and the low unfolding forces are what would be expected from weak long-range interactions (Bertoncini, Jung et al. 2005; Dedmon, Lindorff-Larsen et al. 2005; McClendon, Rospigliosi et al. 2009). Further support for this picture arises from the landscape analysis, which shows that the structures are mechanically compliant, having barriers that are located closer to the unfolded state than the folded state (Table 5.2) and hence quite sensitive to force. Such compliant transitions are a hallmark of molten globule states, owing to the lack of tertiary contacts imparting mechanical rigidity (Elms, Chodera et al. 2012).

5.6 α-Synuclein energy landscape reconstruction

Considering the energy landscape for α-synuclein in more detail, we note that, to our knowledge, the landscape profile for conformational fluctuations in an IDP has never before been quantified. The structures formed here are only marginally more stable than the unfolded state, by 2–3 k_BT, in contrast to typical stabilities of ~ 10–20 k_BT for natively-structured proteins (Rief, Gautel et al. 1997; Yu, Gupta et al. 2012). The barriers are also very low, only ~ 3–5 k_BT for unfolding and ~ 0.5–1 k_BT for folding, accounting for the rapidity of the

fluctuations. By investigating the coefficient for conformational diffusion over the barrier, *D,* we also found that the landscape is quite rough.

D is the key parameter setting the time-scale for microscopic motions of the protein (via the prefactor in Kramers' theory (Hanggi, Talkner et al. 1990)). It can be deduced from the rates at zero force and the parameters describing the landscape profile (Neupane, Ritchie et al. 2012; Yu, Gupta et al. 2012):

$$D = \frac{\pi}{3}\left[\frac{k(\Delta x^{\ddagger})^2}{\Delta G^{\ddagger}/k_B T}\right] exp\left(\frac{\Delta G^{\ddagger}}{k_B T}\right). \quad (5.5)$$

D was calculated both for folding and unfolding, for each of the two transition types; in each case, the values found for folding and unfolding agreed. Similar values of *D* were found for both transition types, $D \sim 5\times10^{-14\pm0.7}$ m^2/s. This result is slower than the intrachain diffusion found for α-synuclein from the chain reconfiguration time in fluorescence-quenching measurements, $\sim10^{-11}$ m^2/s (Ahmad, Chen et al. 2012), as well as the diffusion coefficient found similarly in many other unfolded proteins and peptides (Nettels, Gopich et al. 2007; Singh and Lapidus 2008; Bouley Ford, Shin et al. 2013). It is also slower than *D* for crossing the native folding barrier in the protein PrP, $1\times10^{-12\pm0.4}$ m^2/s (Yu, Gupta et al. 2012), one of the few other measurements of *D* across a barrier. Assuming that the slower diffusion we observe arises primarily from roughness in the landscape at the barriers for formation of the marginally-stable structures, a random roughness distribution would imply fluctuations in the landscape of ~ 2–3 $k_B T$ (Zwanzig 1988). Such a roughness is quite large, comparable to the largest values reported for natively-structured proteins (Wensley, Kwa et al. 2012). It is also

similar to the energy differences between the states as well as the height of the barriers. The picture that emerges is thus one of a landscape that is quite flat but rugged (Figure 5.8), in contrast to the strongly funneled shape typical of landscapes for natively-structured proteins (Onuchic and Wolynes 2004; Dill and MacCallum 2012), directly confirming the qualitative landscape model that has been generally assumed for IDPs (Uversky 2013).

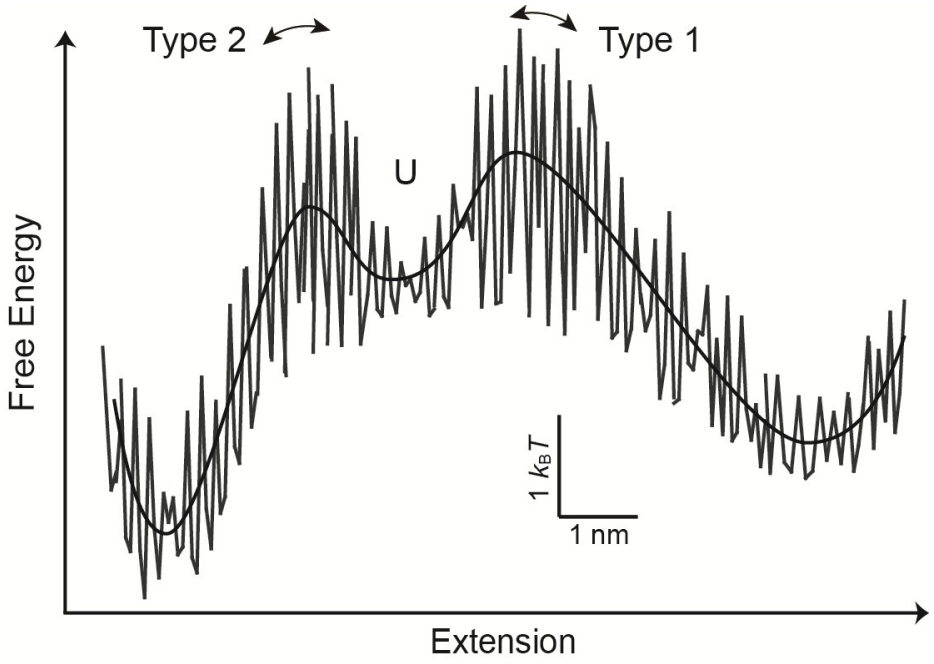

Figure 5.8 Schematic of the energy landscape at zero force, showing the barriers, roughness, and free energy changes from the unfolded state (U) to one of the two distinct states observed.

In summary, we have shown that the rapid but marginally-stable conformational fluctuations that are particularly relevant for IDPs can be observed and characterized using force spectroscopy. From the kinetics of these structural fluctuations, the energy landscape profile was quantified, revealing the flat but rugged landscape expected to be a hallmark of IDPs. Extending this approach to study α-synuclein under conditions in which it is more prone to aggregate, as well as to study other aggregation-prone IDPs, should provide quantitative insight into the features of the energy landscape that relate to the aggregation process.

Chapter 6. Conclusions and Future Work

6.1 Summary of results

Single-molecule approaches offer powerful advantages for characterizing the conformations of IDPs, since they can monitor the dynamic, transient structures formed in disordered proteins. However, it has proven difficult to obtain detailed, high-resolution information about the full range of structures formed in the aggregates—from the dominant populations to the rare, transient states—which is essential for building a microscopic picture of the aggregation process. Such information are ideally be complemented with simultaneous probes of the stability and formation rates of these structures, while also characterizing how all these properties change with the size of an oligomeric aggregate, in order to provide the most comprehensive analysis of the aggregation pathways. Thus far, only molecular simulations have elucidated IDP energy landscapes. Here, I have presented an alternate approach, using optical tweezers—in which the end-to-end molecular extension is measured as the conformation changes in response to an applied force —to characterize marginally stable structural fluctuations in single molecules of human α-synuclein. Given the low stiffness probe and ability to measure at the low forces that are necessary to probe these crucial unstable, and transient states, optical tweezers are ideal tools for identifying the conformational landscape that could prove critical in toxic oligomer formation.

Our measurements of single-molecules of α-synuclein monomers and short oligomers are the first to be reported from an optical trap for any IDP. The picture that emerged was that of a much more complex folding landscape than

previously reported from single molecule AFM and fluorescence experiments at neutral pH (Neupane, Solanki et al. 2014). Surprisingly, α-synuclein formed stable, discrete structures a small fraction of the time, as indicated by the presence of distinct rips in the recorded FECs. The remainder of the time, however, the protein sampled populations of only marginally stable states. We characterized their rates and free energies, allowing us to experimentally demonstrate for the first time that IDPs possess a rough, yet flat energy landscape as had been previously predicted (Solanki, Neupane et al. 2014). Finally, we were able to calculate unique folding rates and diffusion coefficients for these early aggregation states and paint a picture of the complex and dynamic early oligomeric states that may play a role in toxic α-synuclein related diseases Although experimentally rigorous, these results illustrate the utility for using optical tweezers to study the early oligomer states of α-synuclein, and also provide a new platform and analytical tools for studying other IDPs and aggregation prone proteins.

6.2 Future work

With the techniques and analytical tools presented here, the next objective in this project is to identify the specific toxic states present in the early aggregation stage. Identifying structures that are located on-pathway to the disease state requires significantly more work, but the techniques and analytical tools presented here provide the framework for doing so. The following sections detail future experiments that stem naturally from the results presented in this thesis.

6.2.1 SMFS of α-synuclein multimers

The motivation for studying monomers and small oligomers first was to develop an assay for studying oligomeric states, with the goal to eventually probe higher order aggregation species engineered in analogous ways to the dimers and tetramers (e.g., octamers, hexamers). This would help us to continue to track the structural and kinetic evolution that occurs during the early stages of the aggregation process. Given the interesting behavior of rates, stabilities and structural populations we observed in dimers and tetramers, studying larger oligomer constructs could provide more insight into states that stabilize and seed potentially toxic amyloid fibril growth.

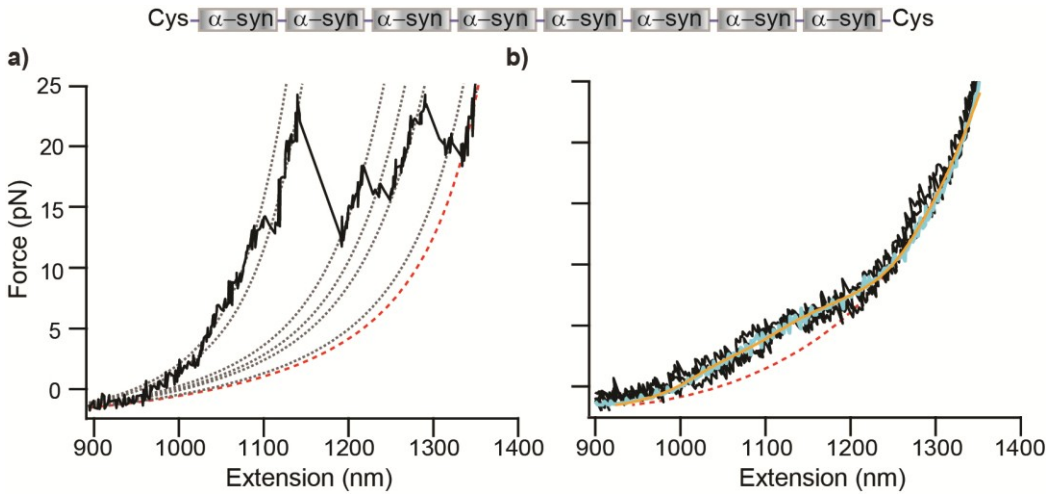

Figure 6.1 SMFS of α-synuclein octamer. a) Single unfolding FEC from an α-synuclein octamer construct, indicating the presence of numerous metastable intermediate structures. Fits to the WLC model are indicated by dashed red and black lines. b) Like the monomer, dimer and tetramer, FECs containing no discrete unfolding events (black) deviated from the WLC fit (red) at low force. The curves were averaged (cyan) and fit well to Equation 5.1 (yellow).

Preliminary data from an α-synuclein octamer has been obtained. A single FEC is shown in Figure 6.1a, indicating the presence of discrete unfolding events, including many intermediates with a relatively low F_{unf} distribution. Additionally, a significant shoulder feature at low force was identified in several pulls not containing stable conformations (Figure 6.2b), analogous to monomer, dimer and tetramer behavior. Due to the extremely low statistics from minimal FECs collected so far, little can be said for the evolution of structural and kinetic properties in higher order states. The intention is to gather enough data to apply the analysis methods presented in Chapters 4 and 5. Concurrent bulk measurements using traditional biochemistry assays will also be continued in parallel. This includes CD and ThT experiments, which could help identify any unique aggregation behavior or conformational changes, including further structural compaction and the formation of secondary structures.

6.2.2 SMFS of familial mutations

One method of determining potential toxic oligomer pathways is to investigate a disease-related mutation. As first introduced in Chapter 1, there are numerous familial mutations linked to PD. We have turned our attention to the mutant E46K for both experimental feasibility reasons and for its interesting structural and aggregation features. For example, E46K monomers haven been shown to aggregate faster than WT monomers in physiological buffers (Fredenburg, Rospigliosi et al. 2007). Further, recent fluorescence experiments studying regional dynamics of E46K fibrils found that the central NAC region is more dynamic and solvent exposed than in its WT counterpart (Sahay, Anoop et

al. 2014). Previous AFM force studies have also studied individual E46K monomers. Interestingly, E46K monomers, sandwiched between tandem titin subunits (Brucale, Missimo et al. 2009), displayed a slight increase in β-like conformations when compared to data from WT monomeric units. Additionally, slightly more long-range, mechanically weak structures were observed. This finding complemented NMR experiments of E46K monomer units that recorded additional contact regions between the protein's terminal regions, enhancing long-range interactions (Rospigliosi, McClendon et al. 2009). These results could perhaps debunk theories that long-range interactions protect against aggregation nucleation. AFM studies investigating the interactions in E46K dimer units reported signature 'rupture' events; however, these only occurred under low pH (3.7) conditions (Krasnoslobodtsev, Volkov et al. 2013). Further, the total number of reported peaks in the contour length distribution decreased in the E46K dimer studies in comparison to the analogous WT experiments.

Several data sets on the E46K tetramer mutant have been collected on our optical tweezers set-up. Most constructs were made using a slightly different handle-protein attachment design, where homobifunctionalized maleimide polyethylene glycol (PEG) units (Creative PEGWorks) replaced DTDP as the crosslinker (Figure 6.2a). The switch to a PEG crosslinker was implemented to help ensure that any potential electrostatic interactions between α-synuclein and the dsDNA handles (particularly between the C-terminal polar region and negatively-charged DNA) were suppressed. PEG is a widely available, soluble compound consisting of highly flexible repeating units of ethylene glycol. In

SMFS, it is often used to tether biomolecules to the sample surface (Baumgartner, Hinterdorfer et al. 2000; Yu, Malkova et al. 2008). PEG is available commercially with numerous sizes and terminal functionalization options. The length of each of our PEG units was ~10nm, and was chosen to provide a significant 'buffer region' between the dsDNA and protein, but not long enough to change the handle behavior which would complicate FEC analysis.

The constructs were made in an analogous way to the DTDP handle attachment protocol. Briefly, 40μM of freshly purified α-synuclein (in PBS, pH 7.4 buffer) was reduced with an approximately 50:1 molar excess of TCEP for 1 hour at 4°C. The excess TCEP was removed by passing the sample through desalting columns (Zeba, 7k MWCO, Thermo Scientific, equilibrated with 0.1M NaPi, pH 7.0) three times. The reduced protein was then incubated with a five molar excess of $PEG_{mal,1k}$ overnight at 4°C. The following morning, the excess PEG was removed with 0.5ml, 50k MWCO spin filters (Amicon). Meanwhile, 10μM dsDNA handles (798 bp-biotin labeled, 2113 bp-digoxigenin labeled) were reduced with a 100:1 molar excess of TCEP for 30 minutes at room temperature. Finally, the PEG reacted protein was mixed at a 3:1 molar excess to the handles, and the entire sample was desalted three times with desalting columns. The sample was then stored at 4°C for at least 24 hours to allow the reaction to go to completion, and then verified on a native 1% agarose gel (Figure 6.2b). These handle-attachments appeared to be slightly more efficient, and exhibited comparable FEC behavior from constructs containing DTDP crosslinkers (e.g., discrete unfolding events, similar F_{unf} and ΔL_c) (Figures 6.3a-c and 6.4).

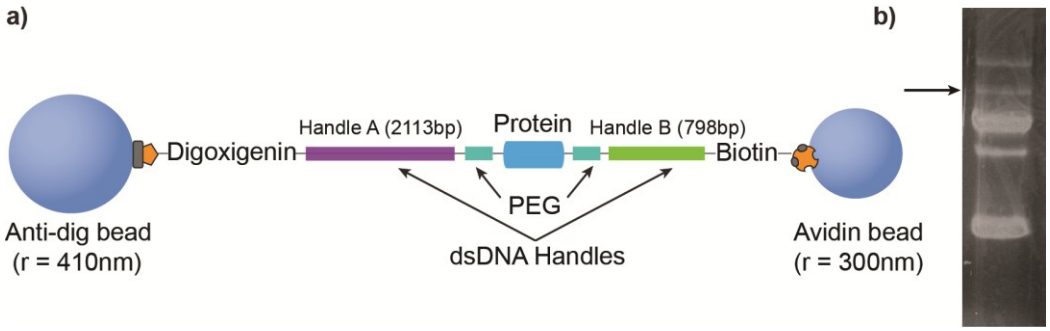

Figure 6.2 Protein-dsDNA construct with PEG crosslinker. a) A variation of the protein-dsDNA dumbbell construct, using a maleimide homobifunctionalized PEG_{1k} unit as the crosslinker instead of DTDP. For comparison to the previous construct design, please refer to Figure 4.12a. b) 1% agarose gel of E46K tetramer-dsDNA handle attachment using maleimide-PEG as the crosslinker. The construct is indicated by arrow, and other bands represent dsDNA handles and their respective dimers.

Preliminary FEC data from force-ramp measurements on 10 molecules (roughly 500 pulls) has been collected from the α-synuclein tetramer mutant E46K. The overall behavior of the E46K tetramer mutants was similar to WT tetramer FEC data: the majority of curves did not contain discrete unfolding events (~ 70%), but did generally display a shoulder feature at low force (Figure 6.3). Upon plotting ΔL_c and F_{unf} scatter plots for curves displaying characteristic rips (Figure 6.3a-c), we could identify correlations to the WT tetramer construct contour length changes. Despite the limited amount of data points, there already appeared to be roughly four clusters of E46K ΔL_c values at approximately 10, 18, 30 and 54 nm (Figure 6.4). No statistically significant events were recorded for much larger ΔL_c. Most interesting was the apparent increase in average unfolding forces for the mutation construct (~14pN for the WT tetramer and roughly

doubled for the E46K tetramer). Although we cannot assign these contour length changes to specific secondary structure, we can attribute this shift from less stable to structures with a higher transition barrier (e.g., from α-helix to β-sheet). Ultimately, significant more data must be collected to perform kinetic analysis and energy landscape reconstruction with any statistical significance. Despite having measured 10 different molecules, far fewer FECs were collected for each molecule in comparison to WT constructs. This may be due to the high forces we applied to the mutant in order to observe discrete unfolding events, which in turn destabilized the dumbbell construct over time.

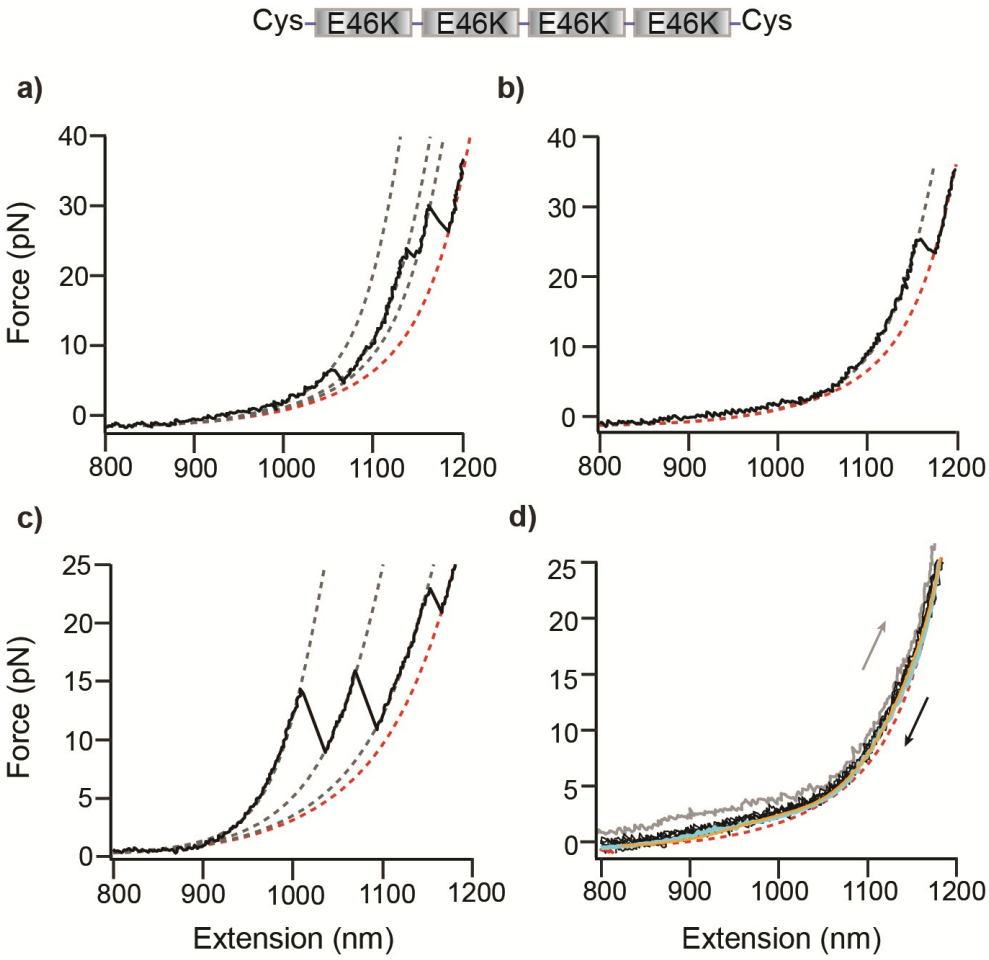

Figure 6.3 SMFS of α-synuclein E46K tetramer. a-c) Individual E46K tetramer FECs showing discrete unfolding events. d) Refolding FECs (black) looked identical to unfolding FECs (grey) from the same molecule when the shoulder feature was observed. The refolding curves were averaged (cyan) and well fit to Equation 5.1 (yellow). Unfolding curve is offset for clearness.

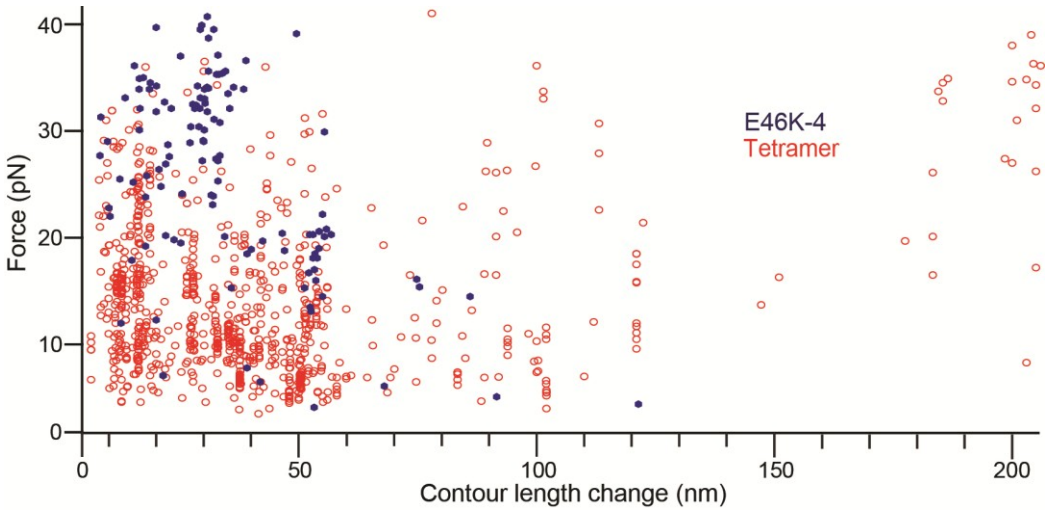

Figure 6.4 α-Synuclein E46K scatter plot comparison. a) Comparison of tetramer (red) and E46K tetramer (purple) F_{unf} versus ΔL_c values obtained from FEC analysis.

We also intend to perform parallel biochemistry assays on these familial mutants, similar to those presented for the WT constructs (see Chapter 4, Appendix: Protein biochemistry assays). Preliminary CD spectra from E46K tetramer samples produced spectra that displayed a classic random coil motif, identical to that of WT tetramers (Figure 6.5). Since CD assays are ensemble measurements, we may or may not be able to detect subtle changes in the conformational dynamics of individual tandem protein molecules, but is a quick and easy assay that may yield useful hints on secondary structure formations. It will also be particularly interesting to conduct ThT aggregation studies on familial mutants to identify potential changes in the lag phase. As a further comparison to the tetramer data, monomeric E46K mutant constructs have been expressed and will be available to measure. Together, they may shed light on the aggregation behavior in α-synuclein variants that are directly linked to PD.

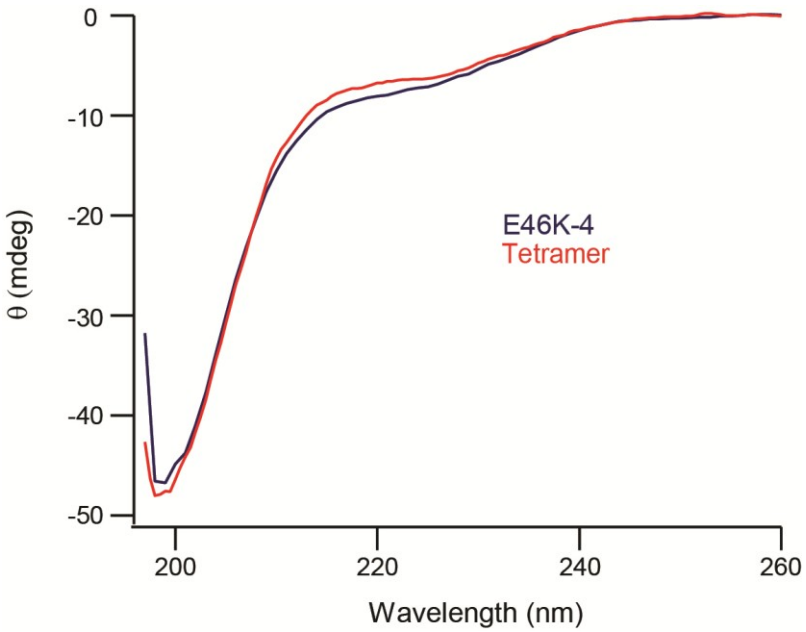

Figure 6.5 CD of α-synuclein E46K tetramer. Comparison of tetramer (red) and E46K tetramer (purple) CD spectrums. Both display the classic random coil motif expected for IDPs, with a very subtle shift observed between the two samples.

6.2.3 Network analysis of α-synuclein unfolding events

A series of FECs is collected for every molecule of interest, with some curves producing a diverse set of unfolding events (e.g., various contour lengths and stabilities) coupled with curves that display shoulder-like behavior. There is quite likely important sequential information embedded in each data set. For example, perhaps there is a particularly large unfolding event in monomer FECS that is always followed by another identifiable rip corresponding to a small ΔL_C at high force (Figure 6.6). If this sequence of events is more frequently populated in familial mutant and/or oligomeric FECs, a network analysis may contain useful information concerning potentially toxic folding pathways. For an IDP like α-synuclein, conducting such analysis is an extremely complex task given the vast

monomer and oligomer FEC data sets available. It is analysis worth pursuing, though, as it could prove especially useful when testing potential PD drug compounds. For example, a chemical that could bind to α-synuclein and eliminate and/or stabilize particular pathways could be identified by employing network analysis.

The ability to identify and test a drug compound that could suppress toxic oligomeric species formation would be of supreme interest. Currently, the most commonly prescribed medicine, L-Dopa, which supplements a PD patient's diminished dopamine supply due to a loss of dopaminergic neurons, loses its effectiveness over time and possesses undesired side effects. Thus, the quality of life for the patients decreases significantly over the course of treatment, especially since the majority of dopaminergic neurons are already dead by the time symptoms are evident (Fearnley and Lees 1991; Uversky and Eliezer 2009). Virtual drug screening will be especially useful for identifying promising drug candidates that we could then introduce into our optical trapping assays and characterize using network analysis.

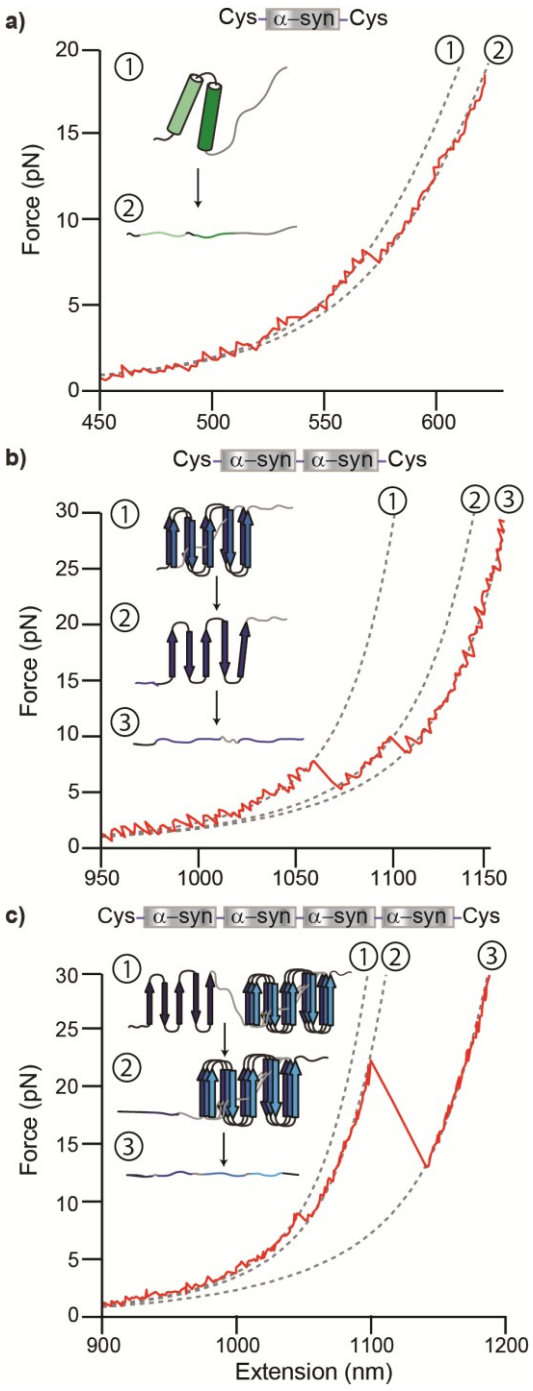

Figure 6.6 Network analysis schematic. Identifying patterns in FECs may be useful for determining folding pathways and testing potential PD drug compounds. Although we cannot assign specific conformations based solely on F_{unf} and ΔL_c data, we can make predictions based on structural models for a) monomers, b) dimers and c) tetramers. (Note: structures shown are hypothetical).

6.2.4 Microfluidic sample slides

Of additional interest is to implement a microfluidic sample slide arrangement for our optical tweezers experiments, similar to those used in single-molecule fluorescence assays (Gambin, VanDelinder et al. 2011; Wunderlich, Nettels et al. 2013). Currently, our experimental setup allows us to measure molecules under a single set of buffer conditions. Since there may be slight differences between individual proteins, it could be valuable to take measurements on a single protein, and then change the sample solution (e.g., salt concentration, introduce a PD drug compound) while continuing to collect data from the same molecule. This would be a particularly interesting experiment for an IDP like α-synuclein, whose conformational landscape can drastically change in a variety of different local environments (e.g., presence of micelles, low pH) (Uversky and Eliezer 2009). Network analysis of FECs could further yield valuable insight on folding pathways that are on-pathway to the nucleation state, including the formation of specific oligomeric intermediates, as well as identify compounds that suppress their presence by changing buffer conditions to induce aggregation. Certain considerations must be kept in mind when conducting measurements with microfluidic sample slides, though. Perhaps the most important is to ensure that mechanical noise from various buffer tubing is not coupled into the measurements. Ultimately, careful modifications would allow us to further enhance an already powerful SMFS instrument, thus opening up additional windows for studying the complex behavior of α-synuclein.

Bibliography

Abbondanzieri, E. A., W. J. Greenleaf, et al. (2005). Direct observation of base-pair stepping by RNA polymerase. *Nature* 438: 460-465.

Ahmad, B., Y. Chen, et al. (2012). Aggregation of α-synuclein is kinetically controlled by intramolecular diffusion. *Proceedings of the National Academy of Sciences USA* 109(7): 2336-2341.

Ahmad, B. and L. J. Lapidus (2012). Curcumin prevents aggregation in α-synuclein by increasing reconfiguration rate. *Journal of Biological Chemistry* 287(12): 9193-9199.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., Walter, P. (2008). Molecular Biology of the Cell. 5th Edition, Garland Science.

Alderson, T. R. and J. L. Markley (2013). Biophysical characterization of α-synuclein and its controversial structure. *Intrinsically Disordered Proteins* 1(1): 18-39.

Almaas, E. and I. Brevik (1995). Radiation forces on a micrometer-sized sphere in an evanescent field. *Journal of the Optical Society of America B* 12(12): 2429-2438.

Anfinsen, C. B. and R. R. Redfield (1956). Protein structure in relation to function and biosynthesis. Advances in Protein Chemistry. K. B. M.L. Anson and T. E. John, Academic Press. **11:** 1-100.

Appel-Cresswell, S., C. Vilarino-Guell, et al. (2013). Alpha-synuclein p.H50Q, a novel pathogenic mutation for Parkinson's disease. *Movement Disorders* 28(6): 811-813.

Ashkin, A. (1970). Acceleration and trapping of particles by radiation pressure. *Physical Review Letters* 24(4): 156-159.

Ashkin, A., J. M. Dziedzic, et al. (1986). Observation of a single-beam gradient force optical trap for dielectric particles. *Optics Letters* 11(5): 288-290.

Bader, R., R. Bamford, et al. (2006). Probing the mechanism of amyloidogenesis through a tandem repeat of the PI3-SH3 domain suggests a generic model for protein aggregation and fibril formation. *Journal of Molecular Biology* 356: 189-208.

Ban, T., D. Hamada, et al. (2003). Direct observation of amyloid fibril growth monitored by Thioflavin T fluorescence. *Journal of Biological Chemistry* 278(19): 16462-16465.

Bartels, T., J. G. Choi, et al. (2011). α-Synuclein occurs physiologically as a helically folded tetramer that resists aggregation. *Nature* 477: 107-110.

Baumgartner, W., P. Hinterdorfer, et al. (2000). Cadherin interaction probed by atomic force microscopy. *Proceedings of the National Academy of Sciences USA* 97(8): 4005-4010.

Bertoncini, C. W., Y.-S. Jung, et al. (2005). Release of long-range tertiary interactions potentiates aggregation of natively unstructured α-synuclein. *Proceedings of the National Academy of Sciences USA* 102(5): 1430-1435.

Best, R. B. and G. Hummer (2010). Coordinate-dependent diffusion in protein folding. *Proceedings of the National Academy of Sciences USA* 107(3): 1088-1093.

Borgia, A., P. M. Williams, et al. (2008). Single-molecule studies of protein folding. *Annual Review of Biochemistry* 77: 101-125.

Borgia, M. B., A. Borgia, et al. (2011). Single-molecule fluorescence reveals sequence-specific misfolding in multidomain proteins. *Nature* 474(7353): 662-665.

Bouley Ford, N. D., D.-W. Shin, et al. (2013). Intrachain contact dynamics in unfolded cytochrome cb562. *Journal of Physical Chemistry B* 117(42): 13206-13211.

Breydo, L., J. W. Wu, et al. (2012). α-Synuclein misfolding and Parkinson's disease. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* 1822(2): 261-285.

Brucale, M., S. Missimo, et al. (2009). Pathogenic mutations shift the equilibria of alpha-synuclein single molecules towards structured conformers. *ChemBioChem* 10: 176-183.

Brucale, M., B. Schuler, et al. (2014). Single-molecule studies of intrinsically disordered proteins. *Chemical Reviews* 114(6): 3281-3317.

Bryngelson, J. D., J. N. Onuchic, et al. (1995). Funnels, pathways, and the energy landscape of protein-folding - a synthesis. *Proteins-Structure Function and Genetics* 21(3): 167-195.

Bucciantini, M., E. Giannoni, et al. (2002). Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature* 416(6880): 507-511.

Burré, J., M. Sharma, et al. (2010). α-Synuclein promotes SNARE-complex assembly in vivo and in vitro. *Science* 329(5999): 1663-1667.

Burré, J., S. Vivona, et al. (2013). Properties of native brain α-synuclein. *Nature* 498(7453): E4-E6.

Bustamante, C., S. B. Smith, et al. (2000). Single-molecule studies of DNA mechanics. *Current Opinion in Structural Biology* 10(3): 279-285.

Carrion-Vazquez, M., H. B. Li, et al. (2003). The mechanical stability of ubiquitin is linkage dependent. *Nature Structural Biology* 10(9): 738-743.

Cecconi, C., E. Shank, et al. (2008). Protein-DNA chimeras for single molecule mechanical folding studies with the optical tweezers. *European Biophysics Journal* 37: 729-738.

Cecconi, C., E. A. Shank, et al. (2005). Direct observation of the three-state folding of a single protein molecule. *Science* 309: 2057-2060.

Chandra, S., X. Chen, et al. (2003). A broken α-helix in folded α-synuclein. *Journal of Biological Chemistry* 278(17): 15313-15318.

Chartier-Harlin, M.-C., J. Kachergus, et al. (2004). α-synuclein locus duplication as a cause of familial Parkinson's disease. *The Lancet* 364(9440): 1167-1169.

Cherny, D., W. Hoyer, et al. (2004). Double-stranded DNA stimulates the fibrillation of α-synuclein in vitro and is associated with the mature fibrils: an electron microscopy study. *Journal of Molecular Biology* 344(4): 929-938.

Chiti, F. and C. M. Dobson (2006). Protein misfolding, functional amyloid, and human disease. *Annual Review of Biochemistry* 75: 333-366.

Conway, K. A., J. D. Harper, et al. (1998). Accelerated in vitro fibril formation by a mutant α-synuclein linked to early-onset Parkinson disease. *Nature Medicine* 4(11): 1318-1320.

Conway, K. A., S.-J. Lee, et al. (2000). Acceleration of oligomerization, not fibrillization, is a shared property of both α-synuclein mutations linked to early-onset Parkinson's disease: Implications for pathogenesis and therapy. *Proceedings of the National Academy of Sciences USA* 97(2): 571-576.

Crampton, N. and D. J. Brockwell (2010). Unravelling the design principles for single protein mechanical strength. *Current Opinion of Structural Biology* 20(4): 508-517.

Cremades, N., Samuel I. A. Cohen, et al. (2012). Direct observation of the interconversion of normal and toxic forms of α-synuclein. *Cell* 149(5): 1048-1059.

Crooks, G. E. (1999). Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Physical Review E* 60(3): 2721-2726.

Dalal, R. V., M. H. Larson, et al. (2006). Pulling on the nascent RNA during transcription does not alter kinetics of elongation or ubiquitous pausing. *Molecular Cell* 23: 231-239.

Das, R. K. and R. V. Pappu (2013). Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. *Proceedings of the National Academy of Sciences USA* 110(33): 13392-13397.

Davidson, W. S., A. Jonas, et al. (1998). Stabilization of α-synuclein secondary structure upon binding to synthetic membranes. *Journal of Biological Chemistry* 273: 9443-9449.

Dedmon, M. M., K. Lindorff-Larsen, et al. (2005). Mapping long-range interactions in α-synuclein using spin-label NMR and ensemble molecular dynamics simulations. *Journal of the American Chemical Society* 127: 476-477.

Dill, K. A. and H. S. Chan (1997). From Levinthal to pathways to funnels. *Nature Structural and Molecular Biology* 4(1): 10-19.

Dill, K. A. and J. L. MacCallum (2012). The protein-folding problem, 50 years on. *Science* 338(6110): 1042-1046.

Dobson, C. M. (2003). Protein folding and misfolding. *Nature* 426(6968): 884-890.

Dong, J., C. E. Castro, et al. (2010). Optical trapping with high forces reveals unexpected behaviors of prion fibrils. *Nature Structural and Molecular Biology* 17: 1422-1430.

Dudko, O. K., G. Hummer, et al. (2006). Intrinsic rates and activation free energies from single-molecule pulling experiments. *Physical Review Letters* 96: 108101-108104.

Dudko, O. K., G. Hummer, et al. (2008). Theory, analysis, and interpretation of single-molecule force spectroscopy experiments. *Proceedings of the National Academy of Sciences USA* 105: 15755-15760.

Dunker, A. K. (2013). Another disordered chameleon: The micro-exon gene 14 protein from Schistosomiasis. *Biophysical Journal* 104(11): 2326-2328.

Dunker, A. K., Z. Obradovic, et al. (2000). Intrinsic protein disorder in complete genomes. *Genome Informatics* 11: 161-171.

Dyson, H. J. and P. E. Wright (2005). Intrinsically unstructured proteins and their functions. *Nature Reviews Molecular Cell Biology* 6(3): 197-208.

Eanes, E. D. and G. G. Glenner (1968). X-ray diffraction studies on amyloid filaments. *Journal of Histochemistry & Cytochemistry* 16(11): 673-&.

Eichner, T. and S. E. Radford (2011). A diversity of assembly mechanisms of a generic amyloid fold. *Molecular Cell* 43: 8-18.

Eisenberg, D. and M. Jucker (2012). The amyloid state of proteins in human diseases. *Cell* 148(6): 1188-1203.

Elms, P. J., J. D. Chodera, et al. (2012). The molten globule state is unusually deformable under mechanical force. *Proceedings of the National Academy of Sciences USA* 109(10): 3796-3801.

Evans, E. and K. Ritchie (1997). Dynamic strength of molecular adhesion bonds. *Biophysical Journal* 72(4): 1541-1555.

Fandrich, M. and C. M. Dobson (2002). The behaviour of polyamino acids reveals an inverse side chain effect in amyloid structure formation. *EMBO Journal* 21(21): 5682-5690.

Fauerbach, J. A., D. A. Yushchenko, et al. (2012). Supramolecular non-amyloid intermediates in the early stages of alpha-synuclein aggregation. *Biophysical Journal* 102(5): 1127-1136.

Fauvet, B., M. K. Mbefo, et al. (2012). Alpha-synuclein in central nervous system and from erythrocytes, mammalian cells, and escherichia coli exists predominantly as disordered monomer. *Journal of Biological Chemistry* 287(19): 15345-15364.

Fearnley, J. M. and A. J. Lees (1991). Ageing and Parkinson's Disease: substantia nigra regional selectivity. *Brain* 114(5): 2283-2301.

Ferreon, A. C., Y. Gambin, et al. (2009). Interplay of α-synuclein binding and conformational switching probed by single-molecule fluorescence. *Proceedings of the National Academy of Sciences USA* 106: 5645-5650.

Ferreon, A. C. M., M. M. Moosa, et al. (2012). Counteracting chemical chaperone effects on the single-molecule α-synuclein structural landscape. *Proceedings of the National Academy of Sciences USA* 109(44): 17826-17831.

Ferreon, A. C. M., C. R. Moran, et al. (2010). Alteration of the α-synuclein folding landscape by a mutation related to parkinson's disease. *Angewandte Chemie International Edition* 49: 3469-3472.

Ferron, F., S. Longhi, et al. (2006). A practical overview of protein disorder prediction methods. *Proteins-Structure Function and Bioinformatics* 65(1): 1-14.

Fink, A. L. (2006). The aggregation and fibrillation of alpha-synuclein. *Accounts of Chemical Research* 39(9): 628-634.

Foltynie, T., A. W. Michell, et al. (2007). Parkinson's Disease. Protein Misfolding in Neurodegenerative Diseases, CRC Press.

Foster, D. A. N. (2010). High resolution optical tweezers for single molecule studies of hierarchical folding in the *pbuE* riboswitch aptamer. Master's thesis, University of Alberta.

Fredenburg, R. A., C. Rospigliosi, et al. (2007). The impact of the E46K mutation on the properties of alpha-synuclein in its monomeric and oligomeric states. *Biochemistry* 46(24): 7107-7118.

Gambin, Y. and A. A. Deniz (2010). Multicolor single-molecule FRET to explore protein folding and binding. *Molecular BioSystems* 6(9): 1540-1547.

Gambin, Y., V. VanDelinder, et al. (2011). Visualizing a one-way protein encounter complex by ultrafast single-molecule mixing. *Nature Methods* 8(3): 239-241.

Gebhardt, J. C. M., T. Bornschlögl, et al. (2010). Full distance-resolved folding energy landscape of one single protein molecule. *Proceedings of the National Academy of Science USA* 107: 2013-2018.

Georgieva, E. R., T. F. Ramlall, et al. (2010). The lipid-binding domain of wild type and mutant alpha-synuclein: compactness and interconversion between the broken and extended helix forms. *Journal of Biological Chemistry* 285: 28261-28273.

Giehm, L., N. Lorenzen, et al. (2011). Assays for α-synuclein aggregation. *Methods* 53(3): 295-305.

Goldberg, M. S. and P. T. Lansbury (2000). Is there a cause-and-effect relationship between alpha-synuclein fibrillization and Parkinson's disease? *Nature Cell Biology* 2(7): E115-E119.

Greenleaf, W. J., K. L. Frieda, et al. (2008). Direct observation of hierarchical folding in single riboswitch aptamers. *Science* 319: 630-633.

Hanggi, P., P. Talkner, et al. (1990). Reaction-rate theory - 50 years after Kramers. *Reviews of Modern Physics* 62(2): 251-341.

Hegde, M. L. and K. S. J. Rao (2003). Challenges and complexities of alpha-synuclein toxicity: new postulates in unfolding the mystery associated with Parkinson's disease. *Archives of Biochemistry and Biophysics* 418(2): 169-178.

Hermanson, G. T. (2008). Bioconjugate Techniques. 2nd Edition, Academic Press.

Hervas, R., J. Oroz, et al. (2012). Common features at the start of the neurodegeneration cascade. *PLoS Biology* 10(5): e1001335.

Higo, J. and K. Umezawa (2014). Free-energy landscape of intrinsically disordered proteins investigated by all-atom multicanonical molecular dynamics. Protein

Conformational Dynamics. K.-l. Han, X. Zhang and M.-j. Yang, Springer International Publishing. **805:** 331-351.

Hoffmann, A. and M. T. Woodside (2011). Signal-pair correlation analysis of single-molecule trajectories. *Angewandte Chemie International Edition* 50: 12643-12646.

Hoyer, W., T. Antony, et al. (2002). Dependence of α-synuclein aggregate morphology on solution conditions. *Journal of Molecular Biology* 322(2): 383-393.

Ivankov, D. N., N. S. Bogatyreva, et al. (2009). Coupling between properties of the protein shape and the rate of protein folding. *PLoS ONE* 4: e6476.

Ivankov, D. N., S. O. Garbuzynskiy, et al. (2003). Contact order revisited: Influence of protein size on the folding rate. *Protein Science* 12: 2057-2062.

Jagannathan, B., P. J. Elms, et al. (2012). Direct observation of a force-induced switch in the anisotropic mechanical unfolding pathway of a protein. *Proceedings of the National Academy of Sciences USA* 109(44): 17820-17825.

Jao, C. C., A. Der-Sarkissian, et al. (2004). Structure of membrane-bound α-synuclein studied by site-directed spin labeling. *Proceedings of the National Academy of Sciences USA* 101(22): 8331-8336.

Jarzynski, C. (1997). Nonequilibrium equality for free energy differences. *Physical Review Letters* 78: 2690-2693.

Jónsson, Sigurður Æ., S. Mitternacht, et al. (2013). Mechanical resistance in unstructured proteins. *Biophysical Journal* 104(12): 2725-2732.

Kalia, L. V., S. K. Kalia, et al. (2013). α-Synuclein oligomers and clinical implications for Parkinson disease. *Annals of Neurology* 73(2): 155-169.

Kaylor, J., N. Bodner, et al. (2005). Characterization of oligomeric intermediates in α-synuclein fibrillation: FRET studies of Y125W/Y133F/Y136F α-synuclein. *Journal of Molecular Biology* 353(2): 357-372.

Kellermayer, M. S. Z., C. Bustamante, et al. (2003). Mechanics and structure of titin oligomers explored with atomic force microscopy. *Biochimica et Biophysica Acta - Bioenergetics* 1604: 105-114.

Khurana, R., V. N. Uversky, et al. (2001). Is Congo Red an amyloid-specific dye? *Journal of Biological Chemistry* 276(25): 22715-22721.

Kim, B.-H., N. Y. Palermo, et al. (2011). Single-molecule atomic force microscopy force spectroscopy study of Aβ-40 interactions. *Biochemistry* 50(23): 5154-5162.

Kim, J., C.-Z. Zhang, et al. (2010). A mechanically stabilized receptor-ligand flex-bond important in the vasculature. *Nature* 466: 992-995.

Kramers, H. A. (1940). Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* 7(4): 284-304.

Krasnoslobodtsev, A. V., J. Peng, et al. (2012). Effect of spermidine on misfolding and interactions of alpha-synuclein. *PLoS One* 7(5): e38099.

Krasnoslobodtsev, A. V., I. L. Volkov, et al. (2013). α-Synuclein misfolding assessed with single molecule AFM force spectroscopy: effect of pathogenic mutations. *Biochemistry* 52(42): 7377-7386.

Krishnan, S., E. Y. Chi, et al. (2002). Oxidative dimer formation is the critical rate-limiting step for Parkinson's disease α-synuclein fibrillogenesis. *Biochemistry* 42(3): 829-837.

Kruger, R., W. Kuhn, et al. (1998). Ala30Pro mutation in the gene encoding alpha-synuclein in Parkinson's disease. *Nature Genetics* 18(2): 106-108.

Kubelka, J., J. Hofrichter, et al. (2004). The protein folding 'speed limit'. *Current Opinion in Structural Biology* 14: 76-88.

Laganowsky, A., C. Liu, et al. (2012). Atomic view of a toxic amyloid small oligomer. *Science* 335(6073): 1228-1231.

Lashuel, H. A., B. M. Petre, et al. (2002). α-Synuclein, especially the Parkinson's disease-associated mutants, forms pore-like annular and tubular protofibrils. *Journal of Molecular Biology* 322(5): 1089-1102.

Lavedan, C. (1998). The Synuclein Family. *Genome Research* 8(9): 871-880.

Le Gall, T., P. R. Romero, et al. (2007). Intrinsic disorder in the protein data bank. *Journal of Biomolecular Structure & Dynamics* 24(4): 325-341.

Li, M. S., D. K. Klimov, et al. (2004). Thermal denaturation and folding rates of single domain proteins: size matters. *Polymer* 45: 573-579.

Liu, C.-W., B. I. Giasson, et al. (2005). A precipitating role for truncated α-synuclein and the proteasome in α-synuclein aggregation: implications for pathogenesis of Parkinson disease. *Journal of Biological Chemistry* 280(24): 22670-22678.

Lorenzen, N., S. B. Nielsen, et al. (2014). The role of stable α-synuclein oligomers in the molecular events underlying amyloid formation. *Journal of the American Chemical Society* 136(10): 3859-3868.

Lundvig, D., E. Lindersson, et al. (2005). Pathogenic effects of α-synuclein aggregation. *Molecular Brain Research* 134(1): 3-17.

Marko, J. and E. D. Siggia (1995). Stretching DNA. *Macromolecules* 28: 8759-8770.

Maroteaux, L., J. T. Campanelli, et al. (1988). Synuclein - a neuron-specific protein localized to the nucleus and presynaptic nerve-terminal. *The Journal of Neuroscience* 8(8): 2804-2815.

Mazzulli, J. R., A. J. Mishizen, et al. (2006). Cytosolic catechols inhibit α-synuclein aggregation and facilitate the formation of intracellular soluble oligomeric intermediates. *The Journal of Neuroscience* 26(39): 10068-10078.

McClendon, S., C. C. Rospigliosi, et al. (2009). Charge neutralization and collapse of the C-terminal tail of alpha-synuclein at low pH. *Protein Science* 18(7): 1531-1540.

Michalet, X., S. Weiss, et al. (2006). Single-molecule fluorescence studies of protein folding and conformational dynamics. *Chemical Reviews* 106(5): 1785-1813.

Milanesi, L., J. P. Waltho, et al. (2012). Measurement of energy landscape roughness of folded and unfolded proteins. *Proceedings of the National Academy of Sciences USA* 109(48): 19563-19568.

Narkiewicz, J., G. Giachin, et al. (2014). In vitro aggregation assays for the characterization of α-synuclein prion-like properties. *Prion* 8(1): 19-32.

Nath, A., M. Sammalkorpi, et al. (2012). The conformational ensembles of alpha-synuclein and tau: combining single-molecule FRET and simulations. *Biophysical Journal* 103(9): 1940-1949.

Nath, S., J. Meuvis, et al. (2010). Early aggregation steps in α-synuclein as measured by FCS and FRET: evidence for a contagious conformational change. *Biophysical Journal* 98: 1302-1311.

Nettels, D., I. V. Gopich, et al. (2007). Ultrafast dynamics of protein collapse from single-molecule photon statistics. *Proceedings of the National Academy of Sciences USA* 104(8): 2655-2660.

Neuman, K. C. and S. M. Block (2004). Optical trapping. *Review of Scientific Instruments* 75: 2787-2809.

Neuman, K. C. and A. Nagy (2008). Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. *Nature Methods* 5(6): 491-505.

Neupane, K., D. B. Ritchie, et al. (2012). Transition path times for nucleic acid folding determined from energy-landscape analysis of single-molecule trajectories. *Physical Review Letters* 109(6): 068102.

Neupane, K., A. Solanki, et al. (2014). Diverse metastable structures formed by small oligomers of α-synuclein probed by force spectroscopy. *PLoS One* 9(1): e86495.

Onuchic, J. N., Z. Luthey-Schulten, et al. (1997). Theory of protein folding: The energy landscape perspective. *Annual Review of Physical Chemistry* 48: 545-600.

Onuchic, J. N. and P. G. Wolynes (2004). Theory of protein folding. *Current Opinion in Structural Biology* 14: 70-75.

Orte, A., N. R. Birkett, et al. (2008). Direct characterization of amyloidogenic oligomers by single-molecule fluorescence. *Proceedings of the National Academy of Sciences USA* 105: 14424-14429.

Ozansoy, M. and A. N. Başak (2013). The central theme of Parkinson's disease: alpha-synuclein. *Molecular Neurobiology* 47(2): 460-465.

Paredes, J. M., S. Casares, et al. (2012). Early amyloidogenic oligomerization studied through fluorescence lifetime correlation spectroscopy. *International Journal of Molecular Sciences* 13(8): 9400-9418.

Parkinson, J. (2002). An essay on the shaking palsy. *The Journal of Neuropsychiatry and Clinical Neurosciences* 14(2): 223-236.

Paslawski, W., S. Mysling, et al. (2014). Co-existence of two different α-synuclein oligomers with different core structures determined by hydrogen/deuterium exchange mass spectrometry. *Angewandte Chemie International Edition*.

Pauling, L. and R. B. Corey (1951). The pleated sheet, a new layer configuration of polypeptide chains *Proceedings of the National Academy of Sciences USA* 37: 251.

Pfefferkorn, C. M., Z. Jiang, et al. (2012). Biophysics of α-synuclein membrane interactions. *Biochimica et Biophysica Acta - Biomembranes* 1818(2): 162-171.

Plotkin, S. S., J. Onuchic, et al. (2002). Understanding protein folding with energy landscape theory Part I: Basic concepts. *Quarterly Reviews of Biophysics* 35(02): 111-167.

Polymeropoulos, M. H., C. Lavedan, et al. (1997). Mutation in the a-synuclein gene identified in families with Parkinson's disease. *Science* 276: 2045-2047.

Rajagopalan, S., F. Huang, et al. (2011). Single-molecule characterization of oligomerization kinetics and equilibria of the tumor suppressor p53. *Nucleic Acids Research* 39(6): 2294-2303.

Rief, M., M. Gautel, et al. (1997). Reversible unfolding of individual titin immunoglobulin domains by AFM. *Science* 276: 1109-1112.

Rohrbach, A. and E. H. K. Stelzer (2002). Trapping forces, force constants, and potential depths for dielectric spheres in the presence of spherical aberrations. *Applied Optics* 41(13): 2494-2507.

Rospigliosi, C. C., S. McClendon, et al. (2009). E46K Parkinson's-linked mutation enhances C-Terminal-to-N-Terminal contacts in α-synuclein. *Journal of Molecular Biology* 388(5): 1022-1032.

Roy, R., S. Hohng, et al. (2008). A practical guide to single-molecule FRET. *Nature Methods* 5(6): 507-516.

Sahay, S., A. Anoop, et al. (2014). Site-specific fluorescence dynamics of α-synuclein fibrils using time-resolved fluorescence studies: effect of familial Parkinson's disease-associated mutations. *Biochemistry* 53(5): 807-809.

Sandal, M., F. Valle, et al. (2008). Conformational equilibria in monomeric α-synuclein at the single-molecule level. *PLoS Biology* 6: 99.

Schlierf, M., F. Berkemeier, et al. (2007). Direct observation of active protein folding using lock-in force spectroscopy. *Biophysical Journal* 93(11): 3989-3998.

Schlierf, M., H. Li, et al. (2004). The unfolding kinetics of ubiquitin captured with single-molecule force-clamp techniques. *Proceedings of the National Academy of Sciences USA* 101(19): 7299-7304.

Sedzik, J. and D. A. Kirschner (1992). Is myelin basic-protein crystallizable. *Neurochemical Research* 17(2): 157-166.

Sela, M., C. B. Anfinsen, et al. (1957). The correlation of ribonuclease activity with specific aspects of tertiary structure. *Biochimica et Biophysica Acta* 26(3): 502-512.

Sethi, A., J. Tian, et al. (2012). Identification of minimally interacting modules in an intrinsically disordered protein. *Biophysical Journal* 103(4): 748-757.

Shank, E. A., C. Cecconi, et al. (2010). The folding cooperativity of a protein is controlled by its chain topology. *Nature* 465: 637-640.

Sherer, T. B., R. Betarbet, et al. (2003). Mechanism of toxicity in rotenone models of Parkinson's disease. *The Journal of Neuroscience* 23(34): 10756-10764.

Simoneau, S., H. Rezaei, et al. (2007). In vitro and in vivo neurotoxicity of prion protein oligomers. *PLoS Pathogens* 3(8): 1175-1186.

Singh, V. R. and L. J. Lapidus (2008). The intrinsic stiffness of polyglutamine peptides. *Journal of Physical Chemistry B* 112(42): 13172-13176.

Singleton, A. B., M. Farrer, et al. (2003). α-Synuclein locus triplication causes Parkinson's disease. *Science* 302(5646): 841.

Smith, S. B., Y. Cui, et al. (1996). Overstretching B-DNA: The elastic response of individual double-stranded and single-stranded DNA molecules. *Science* 271: 795-799.

Solanki, A., K. Neupane, et al. (2014). Single-molecule force spectroscopy of rapidly fluctuating, marginally stable structures in the intrinsically disordered protein α-synuclein *Physical Review Letters* 112(15): 158103.

Soto, C. (2003). Unfolding the role of protein misfolding in neurodegenerative diseases. *Nature Reviews Neuroscience* 4(1): 49-60.

Speretta, E., T. R. Jahn, et al. (2012). Expression in drosophila of tandem amyloid beta peptides provides insights into links between aggregation and neurotoxicity. *Journal of Biological Chemistry* 287(24): 20748-20754.

Spillantini, M. G., M. L. Schmidt, et al. (1997). α-Synuclein in Lewy bodies. *Nature* 388(6645): 839-840.

Stigler, J., F. Ziegler, et al. (2011). The complex folding network of single calmodulin molecules. *Science* 334: 512.

Svoboda, K. and S. M. Block (1994). Biological applications of optical forces. *Annual Review of Biophysics and Biomolecular Structure* 23: 247-285.

Thirumalai, D., E. P. O'Brien, et al. (2010). Theoretical perspectives on protein folding. *Annual Review of Biophysics* 39: 159-183.

Tinoco Jr., I. and C. Bustamante (2002). The effect of force on thermodynamics and kinetics of single molecule reactions. *Biophysical Chemistry* 101-102: 513-533.

Tjernberg, L. O., A. Pramanik, et al. (1999). Amyloid β-peptide polymerization studied using fluorescence correlation spectroscopy. *Chemistry & Biology* 6(1): 53-62.

Tompa, P. (2002). Intrinsically unstructured proteins. *Trends in Biochemical Sciences* 27(10): 527-533.

Tompa, P. (2010). Structure and function of intrinsically disordered proteins, Chapman and Hall/CRC Press.

Trexler, A. J. and E. Rhoades (2009). α-Synuclein binds large unilamellar vesicles as an extended helix. *Biochemistry* 48: 2304-2306.

Trexler, A. J. and E. Rhoades (2010). Single molecule characterization of α-synuclein in aggregation-prone states. *Biophysical Journal* 99: 3048-3055.

Tsuboi, Y. (2012). Environmental-genetic interactions in the pathogenesis of Parkinson's disease. *Experimental Neurobiology* 21(3): 123-128.

Ulmer, T. S., A. Bax, et al. (2005). Structure and dynamics of micelle-bound human α-synuclein. *Journal of Biological Chemistry* 280: 9595-9603.

Uversky, V. N. (2003). A protein-chameleon: conformational plasticity of alpha-synuclein, a disordered protein involved in neurodegenerative disorders. *Journal of Biomolecular Structure and Dynamics* 21(2): 211-234.

Uversky, V. N. (2013). Unusual biophysics of intrinsically disordered proteins. *Biochimica et Biophysica Acta-Proteins and Proteomics* 1834(5): 932-951.

Uversky, V. N. and D. Eliezer (2009). Biophysics of Parkinson's disease: structure and aggregation of alpha-synuclein. *Current Protein and Peptide Science* 10(5): 483-499.

Uversky, V. N., J. Li, et al. (2001). Evidence for a partially folded intermediate in α-synuclein fibril formation. *Journal of Biological Chemistry* 276: 10737-10744.

Uversky, V. N., J. Li, et al. (2001). Metal-triggered transformations, aggregations, and fibrillation of human α-synuclein: a possible molecule NK between Parkinson's disease and heavy metal exposure. *Journal of Biological Chemistry* 276: 44284-44296.

Uversky, V. N., J. Li, et al. (2002). Biophysical properties of the synucleins and their propensities to fibrillate. *Journal of Biological Chemistry* 277: 11970-11978.

Uversky, V. N., C. J. Oldfield, et al. (2008). Intrinsically disordered proteins in human diseases: introducing the D2 concept. *Annual Review of Biophysics* 37(1): 215-246.

van Mameren, J., P. Gross, et al. (2009). Unraveling the structure of DNA during overstretching by using multicolor, single-molecule fluorescence imaging. *Proceedings of the National Academy of Sciences USA* 106(43): 18231-18236.

van Raaij, M. E., J. van Gestel, et al. (2008). Concentration dependence of alpha-synuclein fibril length assessed by quantitative atomic force microscopy and statistical-mechanical theory. *Biophysical Journal* 95(10): 4871-4878.

Vilar, M., H.-T. Chou, et al. (2008). The fold of α-synuclein fibrils. *Proceedings of the National Academy of Sciences USA* 105: 8637-8642.

Wang, M. D., H. Yin, et al. (1997). Stretching DNA with optical tweezers. *Biophysical Journal* 72: 1335-1346.

Wang, W., I. Perovic, et al. (2011). A soluble α-synuclein construct forms a dynamic tetramer. *Proceedings of the National Academy of Sciences USA* 108: 17797-17802.

Wensley, B. G., L. G. Kwa, et al. (2012). Separating the effects of internal friction and transition state energy to explain the slow, frustrated folding of spectrin domains. *Proceedings of the National Academy of Sciences USA* 109(44): 17795-17799.

Williams, M. C. and I. Rouzina (2002). Force spectroscopy of single DNA and RNA molecules. *Current Opinion in Structural Biology* 12(3): 330-336.

Wood, S. J., J. Wypych, et al. (1999). α-Synuclein fibrillogenesis is nucleation-dependent. *Journal of Biological Chemistry* 274: 19509-19512.

Woodside, M. T., W. M. Behnke-Parks, et al. (2006). Nanomechanical measurements of the sequence-dependent folding landscapes of single nucleic acid hairpins. *Proceedings of the National Academy of Sciences USA* 103: 6190-6195.

Woodside, M. T. and M. T. Valentine (2009). Single-molecule manipulation using optical traps. Handbook of Single-Molecule Biophysics. P. Hinterdorfer and A. Oijen, Springer US**:** 341-370.

Wunderlich, B., D. Nettels, et al. (2013). Microfluidic mixer designed for performing single-molecule kinetics with confocal detection on timescales from milliseconds to minutes. *Nature Protocols* 8(8): 1459-1474.

Yu, H. (2013). Single-molecule studies of prion protein folding and misfolding. PhD thesis, University of Alberta.

Yu, H., A. N. Gupta, et al. (2012). 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 USA* 109(36): 14452-14457.

Yu, H., X. Liu, et al. (2012). Direct observation of multiple misfolding pathways in a single prion protein molecule. *Proceedings of the National Academy of Sciences USA* 109(14): 5283-5288.

Yu, J., S. Malkova, et al. (2008). α-synuclein misfolding: single molecule AFM force spectroscopy study. *Journal of Molecular Biology* 384: 992-1001.

Yu, J., J. Warnke, et al. (2011). Nanoprobing of α-synuclein misfolding and aggregation with atomic force microscopy. *Nanomedicine: Nanotechnology, Biology and Medicine* 7(2): 146-152.

Zarranz, J. J., J. Javier Alegre, et al. (2004). The new mutation, E46K, of α-Synuclein causes Parkinson and Lewy Body Dementia. *Annals of Neurology* 55: 164-173.

Žoldák, G., J. Stigler, et al. (2013). Ultrafast folding kinetics and cooperativity of villin headpiece in single-molecule force spectroscopy. *Proceedings of the National Academy of Sciences USA* 110(45): 18156-18161.

Zwanzig, R. (1988). Diffusion in a rough potential. *Proceedings of the National Academy of Sciences USA* 85(7): 2029-2030.