NANOMECHANICAL PROPERTIES OF SINGLE PROTEIN MOLECULES AND PEPTIDES

by

NICOLETA T. PLOSCARIU

B.S. "Al. I. Cuza" University, 2009
M.Sc. "Al. I. Cuza" University, 2011

A THESIS

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Department of Physics
College of Arts and Science

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2014

Approved by:

Major Professor
Dr. Robert Szoszkiewicz

Copyright

Abstract

Proteins are involved in many of the essential cellular processes, such as cell adhesion, muscle function, enzymatic activity or signaling. It has been observed that the biological function of many proteins is critically connected to their folded conformation. Thus, the studies of the process of protein folding have become one of the central questions at the intersection of biophysics and biochemistry.

We propose to use the changes of the nanomechanical properties of these biomolecules as a proxy to study how the single proteins fold. In the first steps towards this goal, the work presented in this thesis is concentrated on studies of unfolding forces and pathways of one particular multidomain protein, as well as on development of the novel method to study elastic spring constant and mechanical energy dissipation factors of simple proteins and peptides.

In the first part of this thesis we present the results of the mean unfolding forces of the NRR region of the Notch1 protein. Those results are obtained using force spectroscopy techniques with the atomic force microscope (AFM) on a single molecule level. We study force-induced protein unfolding patterns and relate those to the conformational transitions within the protein using available crystal structure of the Notch protein and molecular dynamics simulations. Notch is an important protein, involved in triggering leukemia and breast cancers in metazoans, i.e., animals and humans.

In the second part of this thesis we develop a model to obtain quantitative measurements of the molecular stiffness and mechanical energy dissipation factors for selected simple proteins and polypeptides from the AFM force spectroscopy measurements. We have developed this model by measuring the shifts of several thermally excited resonance frequencies of atomic force microscopy cantilevers in contact with the biomolecules. Next, we provided partial experimental validation of this model using peptide films.

Ultimately, our results are expected to contribute in the future to the developments of medical sciences, which are advancing at a level, where human health and disease can be traced down to molecular scale.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would like to thank my advisor, Prof. Robert Szoszkiewicz for his guidance and support during the time I have worked in his group. I am also grateful to Prof. Larry Weaver for his helpful advices and insightful discussions. I would like to thank to my master committee members, Prof. Amit Chakrabarti and Prof. Bret Flanders.

I am also thankful to all former members of the group for helping me to learn some of the experimental skills and for their contribution to this work.

Finally, I would like to thank our numerous collaborators: Prof. Julio Fernandez lab, Columbia University, for the gift of pQE30-$I27_2$ plasmid; Prof. Tom Roche, KSU, for advising in molecular biology; Dr. Maureen Gorman, KSU and Magdalena Wawrzyniuk, for performing a Western Blott assay for Notch protein; Prof. Didem Vardar-Ulu from Wellesley College, MA, for the discussion about folding conditions for the NRR1 domain; Prof. Krzysztof Kuczera from the University of Kansas for providing the Steered Molecular Dynamics simulations of the Notch protein; Prof. John Tomich's lab, KSU, and in particular to Dr. Sushanth Gudlur for instrumental help with CD and fluorescence of Notch protein and to Dr. Pinakin Sukthankar for providing the peptide samples; Prof. Bret Flanders and Dr. Govind Paneru for instrumental help in optical measurements of the AFM cantilevers.

This work was supported by the KSU Physics Department, Prof. Szoszkiewicz Start-up Funds, the Thierry Johnson Cancer Research Center at KSU, and the NSF under Award No. EPS-0903806 and matching support from the State of Kansas through Kansas Technology Enterprise Corporation awarded to Prof. Szoszkiewicz.

"Biology is not simply writing information; it is doing something about it. A biological system can be exceedingly small. Many of the cells are very tiny, but they are very active; they manufacture various substances; they walk around; they wiggle; and they do all kinds of marvelous things – all on a very small scale. Also, they store information. Consider the possibility that we too can make a thing very small which does what we want – that we can manufacture an object that maneuvers at that level!"

Richard Feynman (December 29th 1959)

Overview

This thesis is structured in three major chapters, an introduction chapter presenting main concepts and methods used within this work and the two chapters (Part I and II) presenting the actual work done. In Part I we present the results of the force-induced unfolding patterns which can be related to conformational transitions within a NRR domain from a mammalian Notch protein. In Part II we derive a model to measure the stiffness and dissipation factors of polypeptides and simple proteins.

Introduction

Nowadays, the advancements in nanotechnology combined with the interdisciplinary collaborations allow scientists to work at the cellular and molecular levels and produce major developments in the life sciences and healthcare. We are now equipped with a set of tools which can be used for the detection of specific biomolecules and their structure within minutes and from small amounts of a given sample which have the potential to revolutionize medical healthcare. [1-3] The work presented in this thesis intendes to bring a contribution to the current developments of molecular biology and medical sciences which can be used to advance at a level in which human health and disease can be traced down to a molecular scale.

Molecular Structure of Proteins

Proteins are polymers chains for which the building blocks are smaller molecules called amino acids. Proteins participate in majority of processes needed to sustain life: all antibodies, enzymes and cell receptors are proteins and they have a variety of functions in the cell, both structural and enzymatic.[3,5]

Amino acids contain both amine and carbonyl functional groups, see Figure 2. The amino acids are linked together by covalent bonds to form a chain of amino acid residues, often abbreviated as *residues*. This sequence of residues is called the protein *primary structure.* The

primary structure is determined by the genetic makeup of the individual amino acids. It specifies the order of side-chain groups along the linear polypeptide "backbone".

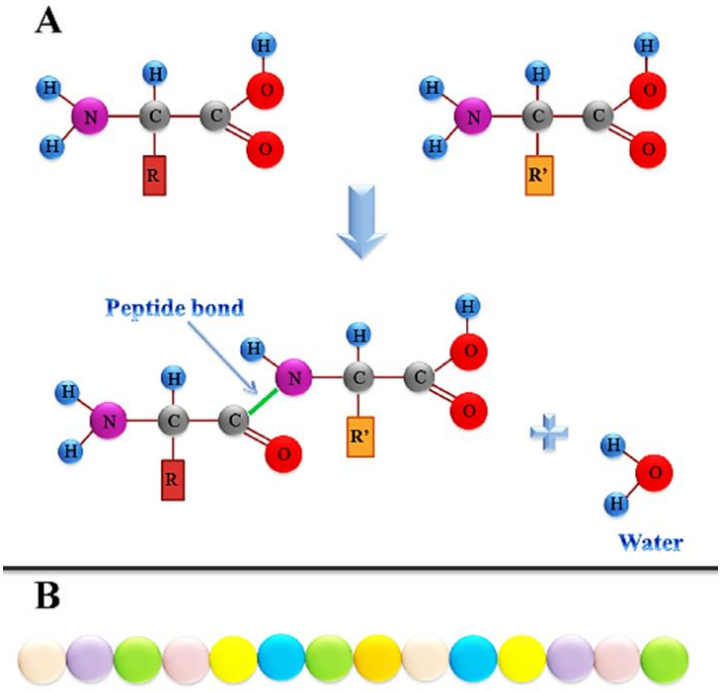

Figure 1. Schematic representation of amino acids, the structural units that make up proteins. A. The chemical reaction in which amino acids are bound together by peptide bond to form short polymer chains called peptides or longer chains called either polypeptides or proteins. B. Representation of the protein primary structure formed in a step-by-step addition of amino acids to a growing protein chain.

The three-dimensional structures of proteins show two major structural types: α- helices and β- sheets, which are called the *secondary structure* of the protein. The α-helices and β-sheets differ in the particular pattern of hydrogen bonds along the backbone, see figures 3 and 4.

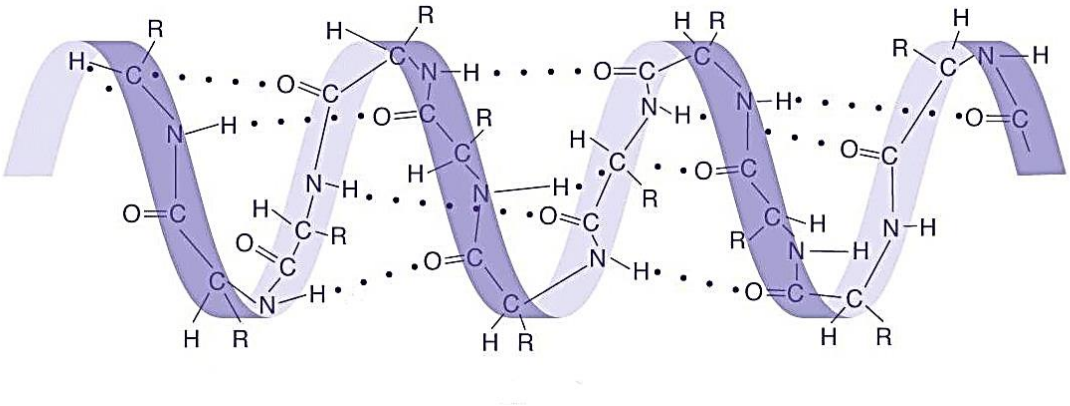

Figure 2. An example of protein secondary structure: α-helix structure showing the hydrogen bond formation between C=O and NH chemical groups.[6]

Typical α- helices are regular spirals stabilized by hydrogen bonds between a carbonyl group of one amino acid and a backbone amide group of the 4th residue down the protein chain. They form a spiral such that at every 3.6 residues, the spiral or helix makes one complete turn. The repeat distance of this helix is 5.4 Å. This structural motif is observed in most of the proteins.

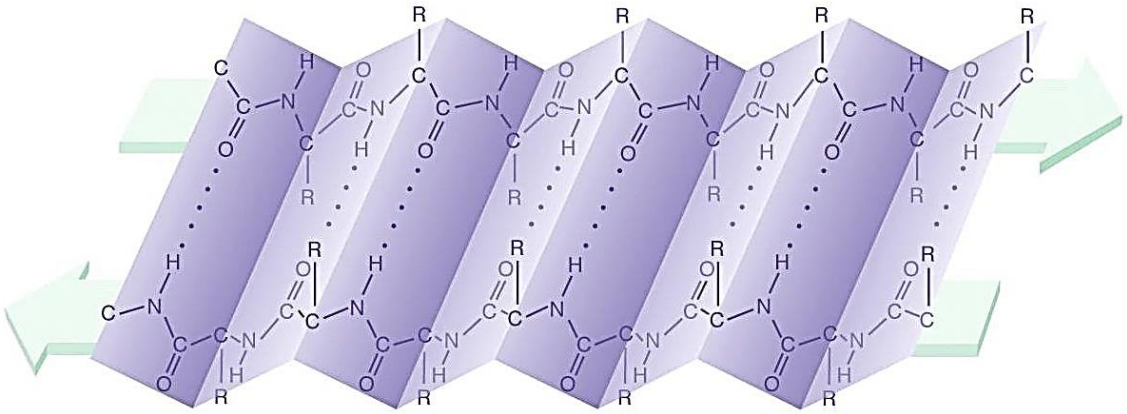

Figure 3. An example of protein secondary structure: β-sheet structure showing the hydrogen bond formation between C=O and NH chemical groups.[6]

The β- sheet conformation is formed by backbone hydrogen bonds between individual beta strands. These strands can be parallel or antiparallel to each other, with alternating side-chains above and below the sheet. The secondary-structure units can connect to one another by other parts of the protein chain, which are sometimes quite mobile or disordered but usually adopt a well-defined, stable arrangement.

The tertiary structure of a protein is a three-dimensional "fold" formed as a result of a final arrangement of all α-helices, β-sheets, and other regions of the proteins. Beyond hydrogen bonding other important interactions stabilizing a tertiary structure are disulfide bonds between selected amino acids containing sulphur, hydrophobic interactions between hydrophobic amino acids in their close vicinity, hydrophilic interactions, van der Waals forces and strong electrostatic interactions between charged amino acids.

Figure 4 shows that typical linear dimensions of proteins are within several nanometers. In addition, the range of the forces required to manipulate a single protein molecule within timescales of most of the biological processes are within sub-pN to pN range.[3,4] Consequently, mechanical properties of the single proteins and polypeptides need to be measured at the nano scale lengths and with appropriate tools sensible to pN forces. Such appropriate instrumental advancements opened an important avenue of research, by allowing us to study protein folding.

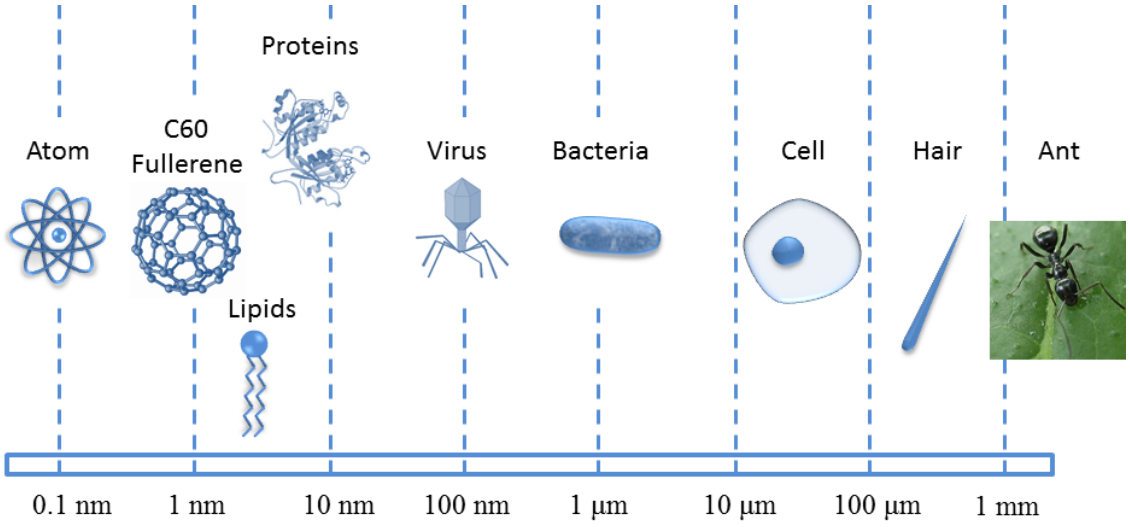

Figure 4. A relative scale of biological organisms. Proteins are in the nano scale level.

Proteins are involved in many of the essential cellular processes, such as cell adhesion, muscle function, enzymatic activity or signaling. There are several in ensemble techniques that use crystallography and spectroscopy to resolve the folded structure of many peptides and proteins in equilibrium states but this information is insufficient to understand their biological function. It has been observed that the biological function of many proteins is critically connected to their folded conformation. Thus, the studies of the process of protein folding have become one of the central questions at the intersection of biophysics and biochemistry.

In this thesis, we concentrate on the nanomechanical properties of single proteins, which will be measured as a proxy to study how the single proteins fold. By applying a mechanical force on a biomolecule it allows us to measure its mechanical responses (resistance) under stress. Nanomechanical properties represent the fundamental mechanical properties of a physical system at the nanoscale level, such as elastic, thermal, kinetic properties. Examples of elastic properties are Young modulus, sample spring constants and energy dissipation factors.

Techniques for the determination of protein structure

There are many currently used techniques to determine the structure of the proteins, as well as to determine variation of such structures with time, e.g., kinetics of the folding process. Generally, we would divide them into two classes: *in ensemble* techniques and single molecule techniques. Below we provide a brief review of the techniques used in this thesis.

Circular Dichroism

Circular dichroism (CD) is a valuable tool in biochemistry used to determine the structure of proteins *in ensemble*. This technique is based on the difference of absorption between left-circularly polarized light and right-circularly polarized light. The protein structures are asymmetric either because the secondary structures have a handedness or a twist-sense as in the case of the alpha helix, which is right-handed.[7,8]

Molecules with asymmetric structures absorb light asymmetrically, i.e., they preferentially absorb either the left- or right polarized light waves. The difference in absorption is described by Beer's law[8]:

$$\Delta\varepsilon = \varepsilon_L - \varepsilon_R = \frac{A_L - A_R}{cl} = \frac{\Delta A}{cl}$$

where $\Delta\varepsilon$ represents the differential molar extinction coefficient, ΔA is the difference in absorption between left-circularly polarized light (A_L) and right-circularly polarized light (A_R), c is the sample concentration, and l is a path length of the light within the measuring cuvette. In practice, *ellipticity* is reported instead of extinction coefficient and the link between them is defined by :

$$ellipticity = 3298 \cdot \Delta\varepsilon$$

Since the peptide bond has strong absorption in the far –UV wavelengths from 230 nm to about 180 nm, the CD spectra for each different type of the secondary structure is unique within this range of wavelengths, see Figure 5.[8]

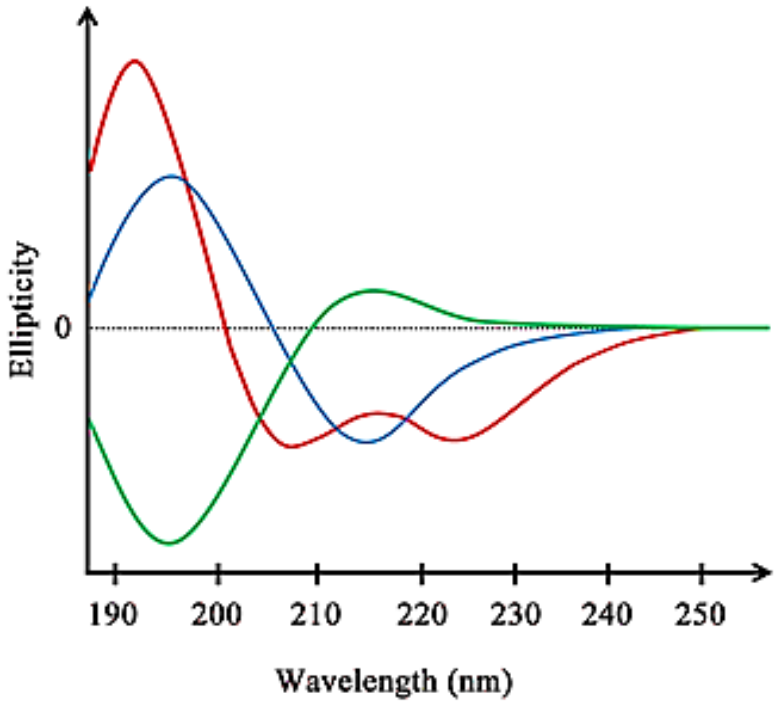

Figure 5. Representative far-UV CD spectra of protein secondary structure.[9] α-helix (red), β-sheet (blue), and unordered conformations (green)

CD can be used to determine the tertiary structure as well, using the signal in the near - UV wavelengths from 350 nm down to 250 nm[10]. The CD signal obtained in this region is due to the absorption, dipole orientation and the nature of the surrounding environment of the phenylalanine, tyrosine, cysteine (or S-S disulfide bridges) and tryptophan amino acids. The near-UV CD spectrum cannot be assigned to any particular three-dimensional structure, but it can provide structural information on the prosthetic groups in proteins, e.g., the heme groups in hemoglobin.

Fluorescence Spectroscopy

Similarly to CD, fluorescence spectroscopy is another technique to determine changes of the secondary structure of proteins in solution. Fluorescence is a multi-stage process that occurs in certain molecules called fluorophores or fluorescent dyes and leads to emitted light associated with direct de-excitation of an atom or a molecule to a ground state from the lowest excited state. The process is illustrated by the simple electronic-state diagram (Jablonski diagram), see Fig. 6.

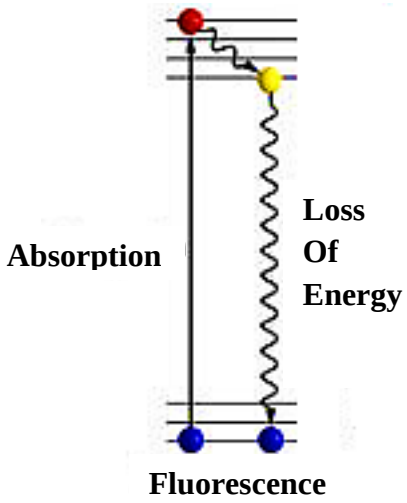

Figure 6. Jablonski diagram representing the fluorescence radiative process. A molecule sitting in the ground state absorbs energy and it is promoted to the excited state. On the way back to the ground state it will emit light.

In the case of proteins, there are three aromatic amino acids active in fluorescence: *phenylalanine*, *tyrosine*, and *tryptophan*. Out of these, only tyrosine and tryptophan are used experimentally, because their quantum yields (number of emitted photons out of excited photons) is high enough to give a good fluorescence signal. Therefore, this technique is limited to proteins containing either tryptophan or tyrosine or both. Tryptophan can be excited by light at 295 nm wavelength. For an excitation wavelength of 280 nm, both tryptophan and tyrosine will be excited. This method can be used to follow protein folding, because their fluorescence

properties are sensitive to their local environment which changes as the protein folds or unfolds, see Fig. 7.

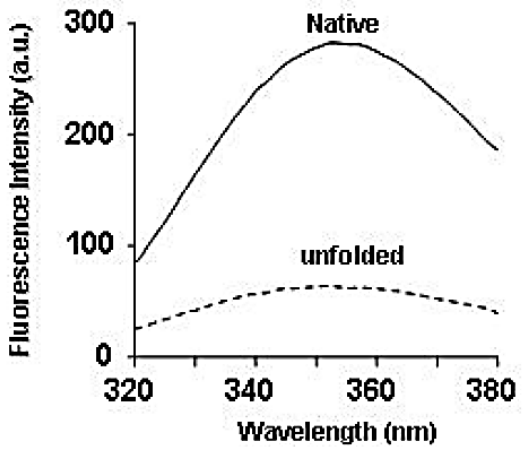

Figure 7. Schematic representation of the fluorescence spectra in the case of folded vs. unfolded protein. In the folded protein state amino acids are buried in the core of protein so which results in high fluorescence intensity while in the unfolded state amino acids are exposed to solvent, which decreases their fluorescence intensity.[11,12]

Single Molecule Force Spectroscopy

It has been shown that pulling biomolecules using single molecule force spectroscopy can be very useful in characterizing the mechanical properties and the unfolding pathways of proteins.[13] There are several promising tools and techniques to measure accurately the forces and displacement during both folding and unfolding processes of a single protein molecule. The most common force-sensitive techniques are: magnetic tweezers, optical traps, and force-extension and force-clamp modes of atomic force microscopy (AFM).

The magnetic tweezers technique uses a magnet to manipulate superparamagnetic nanoparticles. In experiments, the protein is attached between the nanoparticle and a substrate. The force exerted on the nanoparticles is controlled by the applied external magnetic field and the position of the nanoparticles is measured using optical microscopy.

In the case of optical traps, the concept is similar to magnetic tweezers but a focused laser beam is used to "trap" a dielectric bead which is attached to a protein.

In AFM force spectroscopy measurements, a single molecule is tethered between the tip of the AFM cantilever and a sample surface. A more detailed description is presented below.

Force Spectroscopy with Atomic Force Microscopy

Beyond the standard imaging, Atomic Force Microscope (AFM) can be operated to manipulate and stretch single protein molecules, see Figure 8. This is done using two main modes: force-extension and force-clamp mode. In the AFM force-extension experiments a single protein is stretched between the tip of a flexible cantilever and a flat substrate at a constant speed. The resulting observable is a saw-tooth pattern of forces vs. extension of the protein In the AFM force-clamp experiments, constant tensile force acting on a protein is maintained, and the protein's length is measured as a function of time. The substrate is typically a clean glass cover slide to which a layer of gold is adhered by evaporation. A small volume of a protein solution in the desired aqueous medium is added to the substrate and is either adsorbed or chemically cross-linked to the substrate.

By measuring lengths and forces acting on a molecule which is stretched and collapsed either in force-extension or in force-clamp modes, the unfolding and folding trajectories of individual proteins are generated. Those data is then used to gain insight into the physical mechanism of protein folding.[3,4]

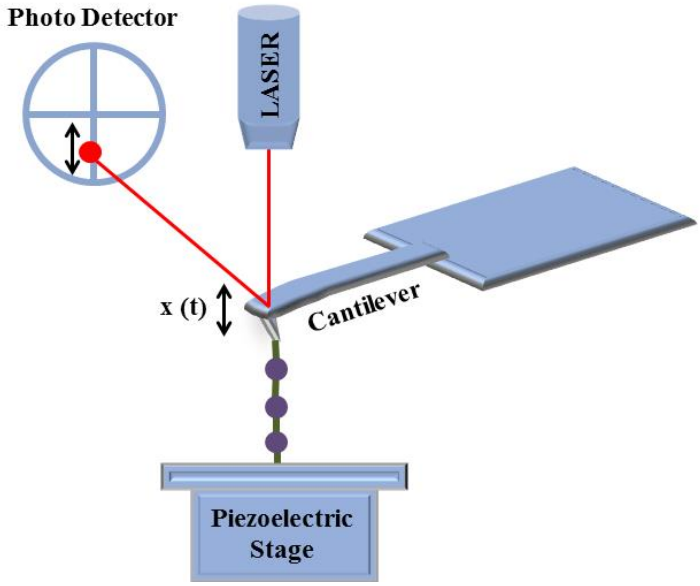

Figure 8. Schematic representation of the AFM setup in a configuration used to manipulate single protein molecules. A laser beam shines on the back side of an AFM cantilever, which is in contact with the studied sample. The sample is placed on a piezoelectric positioner that allows a precise control of the tip-substrate separation.[4] A very important part is the AFM cantilever. The tip of the cantilever is an extremely sharp spike mounted on the end of the cantilever and it is the only part that "touches" the sample. As the cantilever moves, the angle of the reflected laser beam changes, and this produces changes in intensity and position of the signal collected by a photo detector. The detector is a position sensitive photodiode (PSPD) detector which can measure both vertical and horizontal bending of the cantilever.

References

1. Kroll, Alexandra, *Cell. Mol. Life Sci* **2012**, *69*, 335-336.

2. Nussinov, R; Aleman, C., *Phys Biol* **2006**, *3(1)*

3. Toni Hoffmann and Lorna Dougan, *Chem. Soc. Rev.*, **2012**, *41*, 4781-4796

4. Dey, A.; Szoszkiewicz, R., *Nanotechnology* **2012,** *23*.

5. Sperelakis, N., Cell Physiology Sourcebook A Molecular Approach, Third Edition. *Academic Press*, **2001**; pp. 19-27

6. Source: http://www.mun.ca/biology/scarr/MGA2-03-25.jpg

7. van Holde, K.E., Physical Biochemistry. *Prentince Hall, Inc., Englewood Cliffs*, **1985**

8. Christopher Jones, Barbara Mulloy, Adrian H. Thomas, Microscopy, Optical Spectroscopy, and Macroscopic Techniques. *Springer-Verlag New York, LLC*, **1994**, Chapter 16

9. Reference: http://nptel.ac.in/courses/102103047/module2/lec10/1.html

10. D. T. Clarke, *Methods in Molecular Biology*, **2011**, *59-70*

11. Source from web site: http://www.physics.nus.edu.sg/~Biophysics/pc3267/Fluorescence%20Spectroscopy2007.pdf

12. A.G. Szabo, *Spectrophotometry & Spectrofluorimetry, New York: Oxford, University Press*, **2000**

13. M. Carrion-Vazquez, A. F. Oberhauser, T. E. Fisher, P. E. Marszalek, H. Li, J. M. Fernandez, *Progress in Biophysics & Molecular Biology*, **2000**, *74*, 63-91

Part I - Single molecule studies of force-induced S2 site exposure in the mammalian Notch Negative Regulatory Domain

I.1 Introduction

Highly conserved Notch cell-cell signaling pathway controls cell proliferation, cell death, specific cell fates and differentiation programs in all metazoans [1-3]. Aberrant Notch signaling causes developmental syndromes [4-5] and adult-onset diseases such as CADASIL [6]. In addition, Notch signaling emerged as a specific therapeutic target for T cell acute lymphoblastic leukemia[7], breast cancer [8-9], colon cancer [10] and squamous cell carcinomas [11]. Finally, manipulations of embryonic or adult stem cells also require development of receptor-specific antagonists and agonists of Notch signaling. Consequently, examining fine details of Notch activation is of growing translational value.

Canonical Notch signaling is activated when any of Delta-Serrate-Lag2 (DSL) family ligands binds to an ectodomain of a Notch receptor [12], see Figure 9.

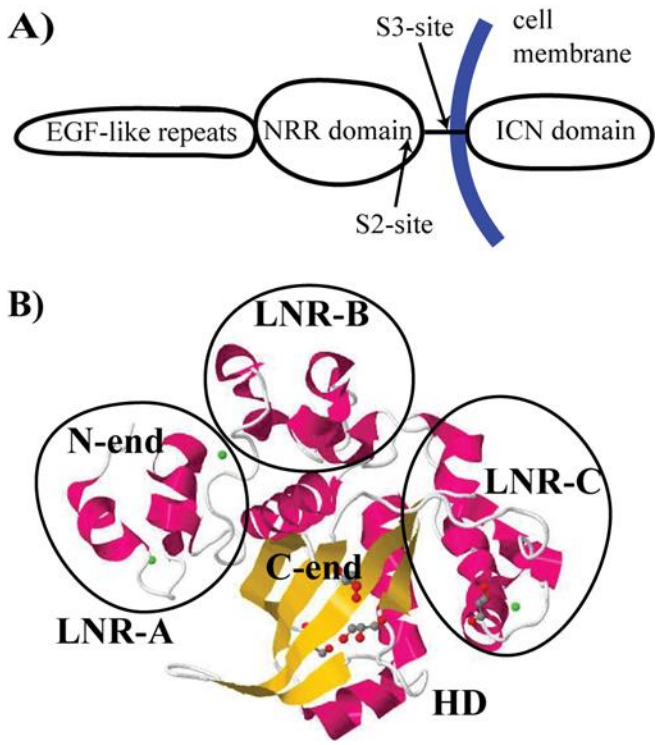

Figure 9 A. Canonical Notch signaling pathway. First, Notch ligand binds to the EGF repeats at the extracellular part of the signal receiving cell. Next steps comprise sequential proteolitic cleavage of the Notch receptor at the S2 site in the extracellular NRR domain, and later at the S3 site in the transmembrane domain. The S3 site cleavage releases an ICN domain, which translocates to the nucleus and activates target gene expression. B. Structure of the NRR1 domain with Ca^{2+} ions (green balls) within each LNR domain [17].

Ligand-receptor interactions enable cleavage of the S2 site in an extracellular negative regulatory region (NRR) by one of a disintegrin and metalloprotease (ADAM) proteins. The Notch extracellular truncation produced after the S2 site shedding is further cleaved at the S3 site in its transmembrane domain by γ-secretase. The S3 cleavage releases the intracellular Notch domain (ICN). The ICN translocates to the nucleus, where it binds to the CSL DNA-binding protein and activates target gene expression [13-14]. Vast biochemical and structural data describes Notch mediated transcriptional regulation of target genes [3, 13, 15]. However, an S2 site exposure and its cleavage are natural rate limiting steps in the Notch pathway [16], and their mechanistic details are still unsolved.

The high resolution crystal structure of NRR fragments from human Notch1 (hNRR1) [17] and Notch2 (hNRR2) [18] provided some clues for the S2 site exposure. First, NRR maintains a compact form, where three Lin12-Notch repeat (LNR) domains wrap around heterodimerization (HD) domain composed of HD-N and HD-C parts. Second, the S2 site, which is located only 12 residues away from the C-end of mammalian NRR domains, in a β5 strand, is quite buried. A linker between LNR-A and LNR-B domains interacts strongly with the HD domain by forming a hydrophobic plug, which occludes the S2 site. Additional S2 site protection is achieved by interactions between LNR-B and one of the alpha helices within the HD-C. When the LNRs are removed, intracellular domain cleavage occurs constitutively and in the absence of ligands [19]. Thus, large-scale conformational transitions of the LNR domains are necessary to expose the S2 site, and several models, including an allosteric model and a mechano-transduction model [3, 13] have been proposed.

The allosteric model proposes that ligand binding to the EGF-like repeats produces major rearrangements of the LNR modules with respect to the HD domain. However, the key region for ligand binding [13] is distal to the S2 site. Thus, allostery might exist in receptors with only very short extracellular domains, such as Notch receptors in *C. elegans*, but is less likely in receptors with a large number of EGF-like repeats separating the ligand binding site from NRR, such as in mammals and flies [3].

The mechano-transduction model proposes that Notch ligands exert mechanical force on the receptor to provoke the S2 site exposure. Indeed, productive interactions between Notch and its ligands occur only when these are present on neighboring cells, i.e., trans-endocytosis [20-21], and well immobilized [22]. Furthermore, clustering of Notch receptors at sites of contact with ligand-expressing cells [21] is reminiscent of clustering mechanically stimulated integrins [23-24]. It is

unclear, however, whether any additional reconfiguration of the HD domain beyond LNR unwrapping is required to expose the S2 site for ADAMs. Recent hydrogen exchange/chemical denaturation study on an NRR domain from human Notch1 have shown an increased accessibility of the S2 site already after partial destabilization of the LNR domains. However, chemical chelation of ions necessary for the LNR domain stability [25] might not exactly relate to the physiological S2 site exposure. In fact, structural results and computer simulations point out that a β5 strand must pop-out, at least partially, of the HD domain to expose the S2 site for cleavage [17-18, 26].

Force-induced exposure and cleavage of the S2 site in the Notch activation process can be addressed at a single molecule level. For example, in the force-extension (FX) experiments with atomic force microscopy (AFM) a single multi-domain protein is stretched between the tip of an AFM cantilever and a flat substrate (gold) that is mounted on a piezoelectric positioner [27-31]. As the distance between a cantilever and a substrate grows with time an extended protein generates a non-linear restoring force that is measured from the deflection of a pre-calibrated cantilever. The resulting force-distance sawtooth pattern informs directly about detectable conformational transitions during such a mechanical unfolding process.

Our preliminary FX AFM data on mechanical unfolding of the NRR domain from mouse Notch1 (mNRR1) pointed out that single molecule studies examining exposure of the S2 site induced by mechanical force are feasible [32]. We used a recombinant protein $I27_2$ – mNRR1 – $I27_2$, where mNRR1 was flanked by I27 proteins [27-31]. We are of the opinion that the use of well characterized unfolding benchmark, like I27, is necessary to obtain reliable AFM data [27-28, 30, 33]. A characteristic unfolding pattern from at least three native I27s in $I27_2$ – mNRR1 – $I27_2$ confirms proper pulling configuration and native, non-aggregated, protein structure. Lack of

internal signature was noticed in the other AFM study on extracellular Notch activation [34]. There, a recombinant protein with an NRR domain from human Notch 2 (hNRR2) was covalently immobilized on the surface via the Lys3 tag chemistry on its N terminus, and interacted with Ni-NTA functionalized AFM cantilevers via a His tag on its C terminus. The authors claimed to observe the S2 site cleavage of Lys3-hN2-His6 protein by ADAMs. However, due to lack of the internal signature, it is not clear whether they observed the S2 site cleavage or a drifting AFM cantilever, which remained in contact with the substrate due to strong interactions facilitated through a surface trapped protein molecule [27].

In this chapter we provide single molecule evidence of the S2 site exposure in the $I27_2$ – mNRR1 – $I27_2$ protein. Using single molecule FX AFM data we produce a histogram of the N to C termini lengths (N-to-C lengths) at which detectable force-induced conformational transitions occur within the mNRR1 domain. By contrasting the AFM data with the steered molecular dynamics (SMD) data obtained for unfolding of the $I27_2$ –NRR1 – $I27_2$ protein, we detect four classes of major conformational transitions within the mNRR1 domain. Our conditional probability analysis supports a sequential unfolding hypothesis for the mNRR1 domain, i.e., initial unwrapping and partial unfolding of the LNR domains, and then unfolding events in the HD domain. On the basis of the SMD results, the first three classes of the AFM detected mNRR1 unfolding events are attributed to the S2 site exposure. Mean forces associated with conformational events within those three classes are 69 ± 42 pN, 79 ± 45 pN, and 90 ± 50 pN respectively at 400 nm/s pulling speeds. Those substantial molecular forces constitute an effective barrier for the S2 site exposure and require continuous, not random, force application in at least several power strokes. In addition, our results agree with a recent physiological study on Notch activation. However, those forces would change depending on physiological pulling

speeds, which remain to be found out. Nevertheless, obtained here molecular fingerprint of the S2 site exposure can now be used in further FC-AFM single molecule studies on kinetics of the exposed S2 site cleavage by ADAMs.

I.2 Materials and Methods

I.2.1 Engineering, expression, and purification of the (I27)2-mNRR1-(I27)2 protein

The $I27_2$ – mNRR1 – $I27_2$ gene was obtained in two subcloning steps. First, cDNA fragments of mNRR1 from pCS2+mN1FL6MT plasmid were amplified by PCR using PfuTurbo DNA polymerase (Stratagen, USA), and subcloned into the pQE30-$I27_2$ vector. We used the *Bgl*II and *Sma*I restriction sites to digest the pQE30-$I27_2$ and the *Bam*HI and *Sma*I restriction sites to digest the mNRR1 insert [31]. The mNRR1 insert contained (from the 5' end): a *Bam*HI restriction site, followed by the mNRR1 coding 276 amino acids (numbered 1450–1725 in a pdb structure 3ETO for hNRR1, which is almost identical to the mNRR1) followed by the *Bgl*II restriction site, and finally by the *Sma*I restriction site. The pQE30-$I27_2$ was obtained by subcloning $I27_2$ into the pQE30 vector (Novagen, USA) as in Ref. 31. The pQE30-$I27_2$ vector contained (from the 5' end): a 6xHis tag, the *Bam*HI restriction site, I27, an inactive hybrid *Bam*HI/*Bgl*II site, I27, *Bgl*II restriction site, two Cys codons, two in-frame stop codons, and a *Sma*I restriction site. In a second step, the amplified cDNA of $I27_2$ obtained by *Bam*HI and *Sma*I digestion of the pQE30-$I27_2$ was subcloned into the pQE30-$I27_2$-mNRR1 digested with *Bgl*II and *Sma*I.

The recombinant $I27_2$ – mNRR1 – $I27_2$ protein has 657 residues: 89 residues for each I27 module, 277 residues for the mNRR1 module, and 24 extra residues for linkers. These extra residues are 12 residues for a His-tag at N-terminus (M-R-G-S-H-H-H-H-H-H-G-S), two residues (R-S) for each linker between each recombinant domain in a construct, and four extra residues (R-S-C-C) at the C-terminus of the construct.

Due to 10 disulfide bonds within a mNRR1 domain, i.e., three in each LNR domain and one in an HD domain the recombinant protein needs to be expressed in cells tolerating many disulfide bonds, kept in reducing environment to prevent aggregation, and equilibrated with calcium ions at redox conditions favoring formation of native disulfide bonds [17, 35-37]. Thus, a recombinant protein with a His tag was expressed in Rosetta(DE3)pLysS *E.Coli* (Novagen, USA) and lysed in the presence of 1 mM tris(2-carboxyethyl)phosphine (TCEP). We did not observe any inclusion bodies and obtained very small protein concentration in a pellet, which confirms non-appreciable aggregation. We suggest that non-appreciable aggregation is at least partially due to a protective role to an mNRR1 domain provided by the I27 modules. The protein was affinity-purified in 1 mM TCEP on Talon cobalt columns (Clontech, USA). Later, protein elution buffer (50mM sodium phosphate, 300mM sodium chloride, 250mM imidazole, pH 7.4) was exchanged into the equilibration buffer (50mM sodium phosphate, 300mM sodium chloride, pH 7.4) to remove imidazole, and the protein was dialyzed overnight into TBS + $CaCl_2$ buffer (50mM Tris, 300mM sodium chloride, 0.5mM calcium chloride, pH 8-8.5). In the last step, the so-called oxidative refolding [37], the protein was dialyzed into TBS + $CaCl_2$ + GSH/GSSG buffer (50mM Tris, 300mM sodium chloride, 0.5mM calcium chloride, 2mM GSH, 0.5mM GSSG, pH 8-8.5) for several days using MAXI Flex Tubes 25K/76bp MWCO (code: IB48250 from MidSci Scientific). The dialyzed protein was stored at 4^oC at typical concentrations of 0.3 to 1 mg/ml.

Size and purity of the obtained protein (73 kDa) were verified using SDS-PAGE, see Appendix A. Impact of Ca^{2+} ions and oxidative refolding of $I27_2$ – mNRR1 – $I27_2$ were investigated by circular dichroism (CD) and fluorescence,[38] as well as enzymatic cleavage with ADAMs detected by a Western Blot assay, see Figure 10.

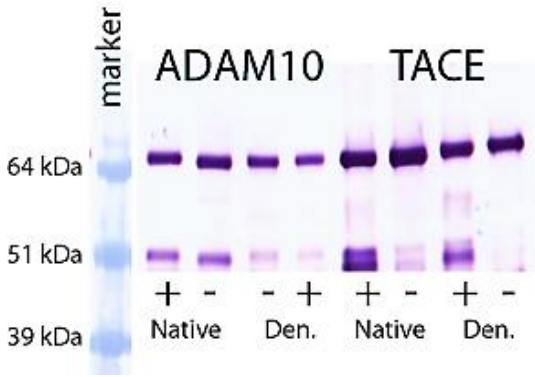

Figure 10. Western Blot gel showing His-tagged cleavage products of $I27_2$ – mNRR1 – $I27_2$ protein by ADAM10 and TACE proteases. "Native" refers to an oxidatively refolded protein; "denatured" is the "native" protein denatured overnight in 5M urea and 2 mM TCEP.

I.2.2 Western Blot assay

In order to identify a particular protein, we can use a "tagged" antibody to the protein and reveal its presence on the electrophosis gel. The antibody is "tagged" with an enzyme that produces a chromophoric reaction and therefore it can be identified.

Since the antibodies can't be added directly to the electrophoresis gel, the gel itself is blotted on a nitrocellulose membrane that contains all of the separated bands of proteins as they were on the original electrophoresis gel. Then, the nitrocellulose strip is incubated in a solution containing the antibody.

The Western Blot assay presented in this thesis was performed by Maureen Gorman and Magdalena Wawrzyniuk.

The $I27_2$ – mNRR1 – $I27_2$ protein after oxidative refolding (both: wild-type and denatured by overnight incubation at room temperature in 5M urea + 2mM TCEP) was diluted five times with an assay buffer (25 mM Tris-HCl in the case of ADAM10, 50mM Tris-HCl in the case of ADAM17 (TACE); pH 9.0 in both cases) and incubated at 37^{o}C for five hours with either 2 ng⁄µL ADAM10 or 2 ng⁄µL TACE. Cleavage products and negative controls (protein without enzyme) were separated on SDS-PAGE gradient gel (NuPAGE 4-12%, Invitrogen #NP0329BOX) and then transferred on a nitrocellulose membrane. The blot was blocked with 3% dry milk in 0.4 mM Tween20 + TBS (140mM NaCl + 2.7mM KCl + 25mM Tris) for 1 h, incubated with mouse anti-His6 antibody (BioRad, #620-0203) in blocking solution (overnight), washed, incubated for 1 h with AP-conjugated goat anti-mouse antibody (BioRad, #170-6520) and after another washing developed using the AP Conjugate Substrate Kit (BioRad #170-6432).

I.2.3 Single molecule AFM force spectroscopy

We used our high resolution FX, FC-AFM setup described in A. Dey et al. [32], see Fig. 11.

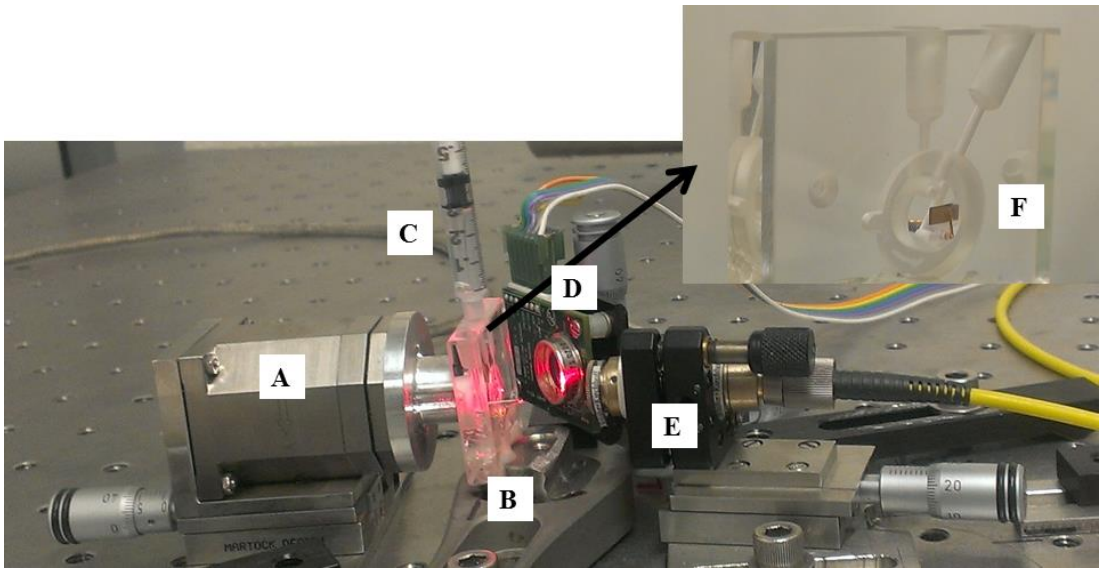

Figure 11. The AFM used for the experiments described in reference 32. The letters enumerate major parts: A—scanner, where the sample is placed; B and F—liquid cell, which holds the cantilever; C—syringe used to insert liquid in the system; E—laser beam focusing assembly; D—photodiode.

The FX AFM experiments were conducted at 23 ± 2 deg C in TBS + $CaCl_2$ buffer at a constant 400 nm/s cantilever-substrate approach/retraction speeds. Raw FX AFM data was filtered with a sub-kHz low pass filter to reduce noise. Protein samples were prepared by depositing between 5 to 30 μl of a protein solution onto freshly evaporated gold substrates, partially dried, and washed with a buffer [32,39]. Surface adsorbed protein molecules likely attached to gold through their C terminus cysteines [30]. A nonspecific binding strategy of N terminus His-tags to AFM cantilevers curbs protein accumulation on the cantilevers. We used MLCT types "C" and "D" cantilevers from Bruker and BioLevers from Olympus. The cantilevers were calibrated *in-situ* using an equipartition method[32]. Their force sensitivity was about 15 pN for the MLCT levers and about 5 pN for BioLevers.

There is no good reason to expect that the $(I27)_2$-mNRR1-$(I27)_2$ protein attaches to the AFM cantilever at its N-terminus. Thus, an unfolding pattern of $I27_2$ – mNRR1 – $I27_2$ with at least three native I27s was used to fingerprint conformational transitions within the mNRR1 domain. We excluded unevenly spaced I27 unfolding peaks, which are typically obtained when several proteins are pulled in series. To avoid analyzing protein agglomerates we selected FX AFM traces with a contact rupture force being significantly larger than 200 pN, which is obtained for the I27 unfolding events alone.

I.2.4 Histograms of the N-to-C distances

A contour length of ca. 55 nm for the NRR1 domain was obtained in the Ref. 26 assuming that none of the disulfide bonds within the mNRR1 domain breaks. This assumption is well justified for the force loading speeds in our FX AFM experiments [40]. Extended length of the whole 277 residue mNRR1 domain is expected to produce about 100 nm contour lengths, as calculated from the WLC chain model using a persistence length of 0.36 nm for each residue [27-31].

Using the AFM data we built a histogram of the N-to-C distances for major conformational transitions within the mNRR1 domain. First, we observed that the mNRR1 domain does not always unfold up to its contour length. Second, length of initially stretched and folded $(I27)_2$-mNRR1-$(I27)_2$ protein can vary between several nm up to 30 nm depending on the orientation of the protein modules on the surface and their binding to the AFM cantilever. A value of 30 nm is obtained as follows: the contour length of each I27 module is about 5 nm (pdb code: 1TIT), the resting N-to-C length of the mNRR1 domain is ca. 5 nm (pdb code: 3ETO), linkers and His-tag are 24 residues in total, which account a maximum additional length of ca. 9

nm. Thus, we have aligned the FX AFM traces to find the underlying unfolding pattern, and then built a histogram, similarly as described in [41]. Briefly, after establishing a zero stretched length for each FX AFM trace, each force peak within such a trace is fitted with the worm-like chain (WLC) model and its obtained contour length is taken for further analysis. Overall, we used 101 acceptable Notch unfolding events from 42 different FX AFM traces from about 100 attempted AFM experiments. We note that histogram results may depend on a bin size. A suggested bin size is about 1/3 of a standard deviation of the binned data [42], and a more advanced procedure for choosing a proper bin size has been described elsewhere [43]. However, due to a limited number of events in our experiments our bin size has been chosen to be 2.5 nm, which is a maximum estimated error of the N-to-C distance determination from the raw FX AFM data.

I.2.5 Steered molecular dynamics (SMD) simulations

The Steered Molecular Dynamics simulations were provided by Professor Krzysztof Kuczera from Department of Molecular Biosciences and Department of Chemistry, University of Kansas.

In order to interpret our AFM data we have performed SMD simulations of forced-induced unfolding of a NRR1 protein and a full length experimental construct $(I27)_2$-NRR1-$(I27)_2$. For the NRR1 domain we used the same linkers and terminating sequences as in our construct, and we utilized the structure and sequence of the NRR1 domain from the published crystal structure 3ETO, chain A. The simulated NRR1 domain has 240 residues, of which last 8 do not have coordinates in 3ETO. The simulated NRR1 domain differs from the experimental construct by the absence of the unstructured loop (amino acids 1623-1669 in the notation from

3ETO), but it is the same structure that was used in the previously described simulation of forced NRR1 unfolding [26, 34]. All ten disulfide bonds and three calcium ions were included. The simulations were performed with the CHARMM program, using the extended atom PARAM19 force field for the protein and the EEF1 implicit solvation model. The I27 structures were copied from the 1TIT structure. Linkers, His-tag and terminating sequence were built in their extended conformations. After a brief energy minimization a 100 ps equilibration was performed for the whole construct, which was followed by 10 ns of free molecular dynamics (MD). Starting with structures extracted every 1 ns from the free MD, we generated nine 20 ns SMD trajectories with a pulling rate of 0.1 A/ps. In all cases, a force constant of 10.0 kcal/(mol A^2) was employed. The resulting trajectories were used to describe the evolution of the lengths of the whole construct, and the NRR1 domain as a function of time and pulling force.

I.3 Results

I.3.1 Biochemical characterization

Figure 9 presents an overview of the Notch pathway and a molecular structure of the hNRR1 domain. In order to study force-induced unfolding of the NRR1 domain by single molecule AFM we constructed, expressed, and purified the $I27_2$ – mNRR – $I27_2$ protein, as described in the Appendix A-C. Using SDS-PAGE we verified its molecular weight (~ 73 kDa) and purity, see Fig. A1 in Appendix A. Next, we checked whether the mNRR1 domain folds correctly in solution. We used circular dichroism and fluorescence spectroscopy [36, 38,40]. However, due to lack of reference CD spectra for the $I27_2$ – mNRR1 – $I27_2$ protein we were unable to conclusively state that $I27_2$ – mNRR1 – $I27_2$ protein is folded correctly in the presence

of calcium ions. Thus, we performed an enzymatic cleavage with ADAM10 and ADAM17 (TACE) proteases known to cleave the S2 site [44-45]. This assay was followed by a Western blot detection of His-tagged proteolitic cleavage products.

The His-tag in our $I27_2$ – mNRR1 – $I27_2$ protein is at its N-terminus and the S2 cleavage site is about 12 residues from the C end of the mNRR1 module. Thus, His-tagged S2 site cleavage products are expected to produce about 50 kDa band in the Western blot gel. This band will comprise two I27 units and an almost full length mNRR1 unit. Figure 10 presents our Western blot results. First, we observed a small amount of intrinsic cleavage of the S2 site independent of the used ADAMs. The intrinsic cleavage is slightly higher in the native than in the denatured protein. Second, addition of ADAM10 did not produce any additional S2 site cleavage both in the native and in the denatured $I27_2$ – mNRR1 – $I27_2$ protein. Third, we observed a clear indication of the TACE induced cleavage in the native, i.e., oxidatively refolded, protein. Our first finding agrees with previous observations that the mNRR1 domain undergoes slow and spontaneous cleavage in solution [46]. Our second and third findings agree with previous observations that although ADAM10 is a predominant protease for the S2 cleavage *in vivo*, TACE is less specific, Notch ligand independent, and thus TACE induced S2 cleavage occurs more readily in bulk [45-46]. In addition, a notable decrease of the S2 site proteolitic cleavage in denatured conditions points out towards a protective role of the I27 domains to the S2 site, which mimics physiological conditions, where a C terminus of the mNRR1 domain is only 12 residues away from a stiff and buried transmembrane region [2-3]. Overall, the results of our enzymatic assay with ADAMs strongly suggest that our protein construct refolded in oxidative conditions is properly folded, i.e., the S2 site is protected against ADAM10 in its native state, while still partially accessible to TACE.

I.3.2 AFM experiments

Using a combination of AFM imaging with low resolution FX AFM data we observed that distinguishable unfolding events from native I27 molecules originate predominantly from single $I27_2$ – mNRR1 – $I27_2$ molecules [38]. Thus, we launched a thorough high-resolution FX AFM study to find a force-induced unfolding pattern of $I27_2$ – mNRR1 – $I27_2$ molecules.

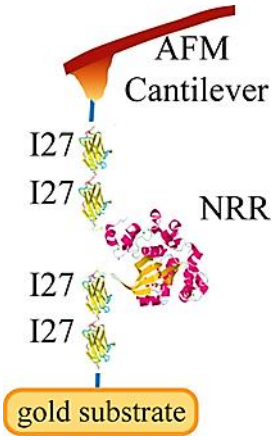

Figure 12 Schematics of the FX AFM experiment, where an AFM tip pulls and mechanically unfolds an $I27_2$ – mNRR1 – $I27_2$ protein absorbed on gold substrate.

Figure 12 presents the schematics of our FX AFM experiments and Figure 13A presents the typical data. Force peaks in the saw-tooth pattern of forces vs extension represent detectable protein unfolding events. I27 unfolds cooperatively at around 200 pN [28, 31, 47-49]. An I27 contour length of ~ 28 nm is recovered by fitting the WLC model to any two consecutive I27 unfolding events [32, 50].

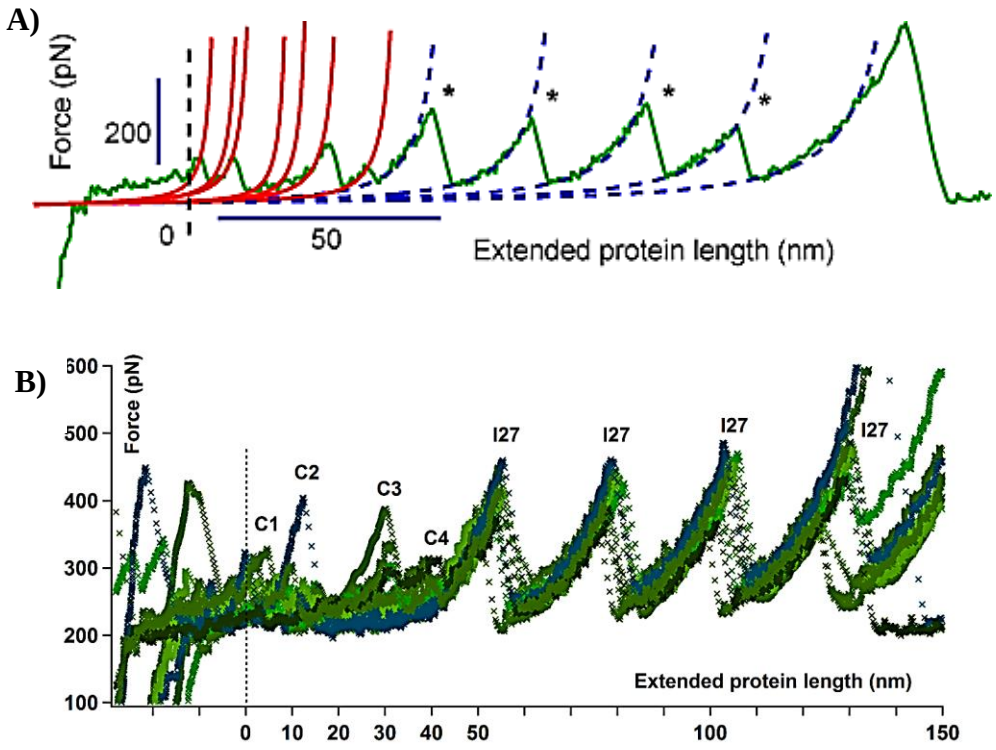

Figure 13. Sample of the FX AFM traces showing force-induced conformational events within the $I27_2$ – mNRR1 – $I27_2$ protein. A. An example of an FX AFM trace with WLC fits. The last force peak at highest extension and force marks contact rupture. Four peaks preceding the contact rupture and marked with a star belong to unfolding of each I27 modules. The peaks before zero extended protein length correspond to surface adhesions of the folded protein. Any other peaks relate to conformational transitions within the mNRR1 domain. B. Twelve FX AFM traces are superimposed to show an unfolding pattern of the mNRR1 domain, and attribution of major conformational events to the classes C1 to C4.

The I27 peaks follow some initial peaks at small extensions, which - within the limitations described in the Appendices A-H - are ascribed to conformational transitions within the mNRR1 domain. Thus, from the FX AFM data in Fig. 13A we obtain the N-to-C lengths corresponding to the force-induced conformational transitions within the mNRR1 domain. To do so, in Fig. 13B we align the FX AFM traces as described in the Appendices A-H. Next, Fig. 14A

plots a histogram of the N-to-C distances associated with detected conformational transitions within the mNRR1 domain. There, we identify four classes of events. Class 1 (C1) contains events occurring at up to 10 nm of the N-to-C distance; class C2 contains events between 10 to 20 nm; class C3: 20 – 35 nm, and class C4 encompasses events occurring between 35 to 45 nm.

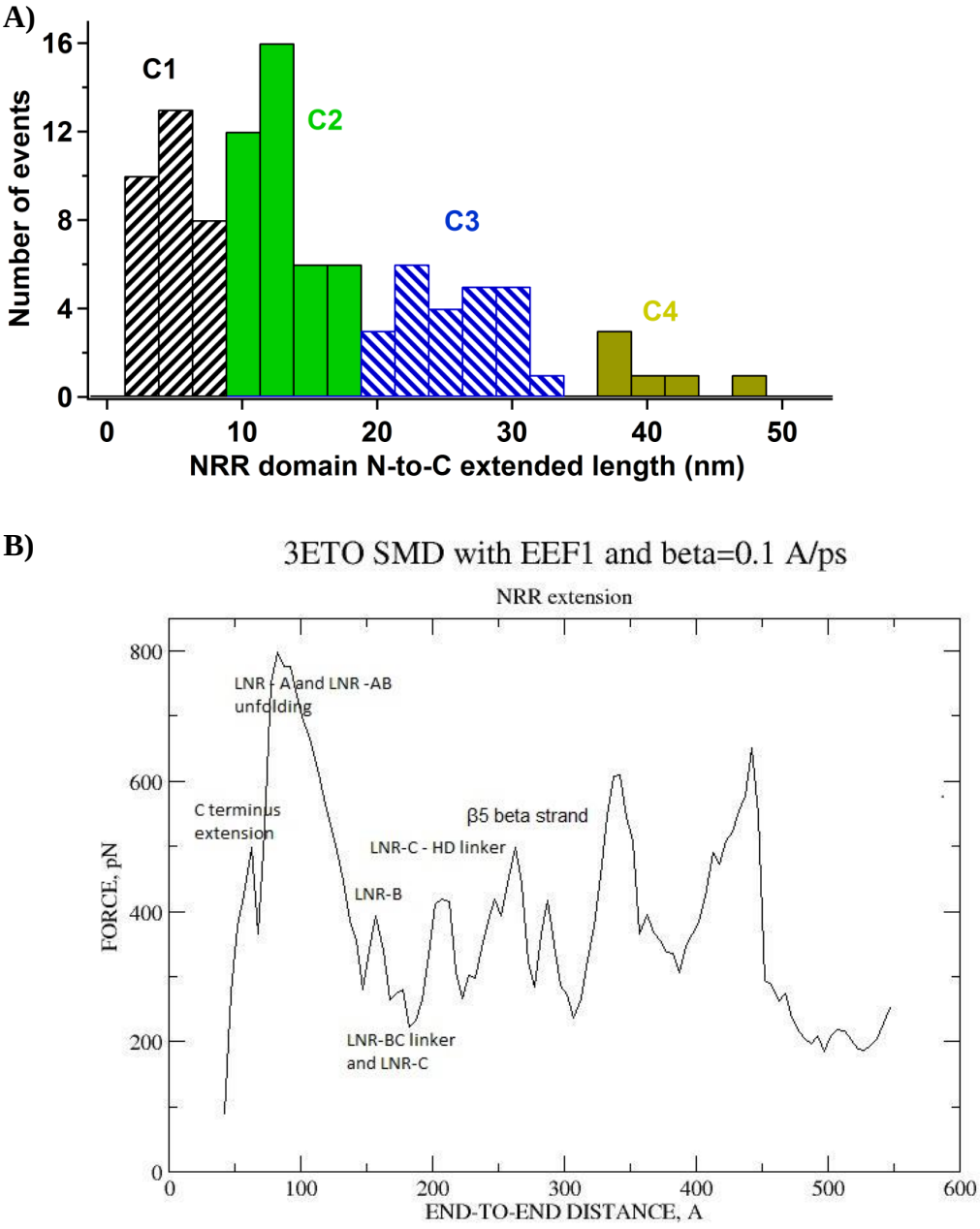

Figure 14 A. Histogram of the N-to-C distances within the mNRR1 domain at which major conformational transitions have been detected in FX AFM; B. Results of the SMD

simulations obtained to complement the AFM experiments. Crucial conformational transition events are labeled for further discussion in the text.

I.3.3 SMD simulations

To complement the FX AFM study we performed the SMD simulations. It is most convenient to analyze the SMD results only for the NRR domain, which we present in Fig. 14B and Table 1.

AFM	SMD	Relates to
Class C1: 4 – 10 nm	6 - 8 nm	Stretching of NRR C-terminus
	9 - 10 nm	Unfolding of LNR-A domain and LNR-AB linker
Class C2: 10 – 20 nm	13 - 15 nm	Unfolding of LNR-B domain
	18 - 20 nm	Unfolding of LNR-BC linker and LNR-C domain
Class C3: 20 – 35 nm	22 - 23 nm	Unfolding of LNR-C – HD linker
	25 - 28 nm	Unfolding of the β5 beta strand from within the HD domain; complete exposure of the S2 site
Class C4: 35 – 45 nm	Series of events	Other unfolding events within the HD domain

Table 1 Comparison of the N-to-C distances detected by AFM and SMD and corresponding to major conformational transitions within the NRR1 domain, see Fig. 14.

The data reported in Fig. 14B correspond to raw end-to-end distance of NRR (CA…CA distance between residues 195 and 434 of the construct), for which the resting value is 0.44 nm. Our SMD trajectories show the presence of several unfolding events. The initial events at distances in the 6-8 nm range correspond to stretching of the C terminus region of the NRR domain, i.e., the residues 427-434. Consecutive peaks in the force-extension plot reflect: i) unfolding of LNR-A and the LNR-AB linker at 9-10 nm, ii) unfolding of LNR-B at 13-15 nm, iii) unfolding of LNR-BC linker and LNR-C at 18-20 nm, iv) unfolding of the LNR-C and HD linker at 22-23 nm and v) unfolding of the β5 beta strand at 25-28 nm. In the individual trajectories we find significant cooperativity – i.e. unfolding of LNR-A, LNR-AB, LNR-B and LNR-BC tends to occur together. Furthermore, β5 beta strand unfolds either by itself as a last element or together with other elements e.g., LNR-C - HD linker, but not before any of the LNR domains and their linkers. Finally, since forces are applied to the ends of the construct, they rarely, induce unfolding of the I27 domains during the course of the NRR unfolding.

I.4 Discussion

I.4.1 Comparisons between FX AFM and SMD experiments

Peaks at similar locations appear in the force-extension curves from FX AFM experiments and in a histogram of distances from the SMD results. Thus, we assign the distinct four classes of unfolding events observed with AFM in Fig. 14A to the SMD results in Fig. 14B. The class C1 is ascribed to detachment and unfolding of the first LNR domain, i.e., the LNR-A domain, as well as its linker to the LNR-B domain. Based on previous crystallographic studies,

as well as previous coarse-grained SMD simulations [13, 26], those events require substantial force and are critical to the S2 site exposure, since they break the hydrophobic plug occluding the S2 site. Next, the class C2 is ascribed to complete unwrapping and maximum allowable unfolding of the remaining LNR domains and their linkers. Those events do not include breaking of the disulfide bonds within the LNR domains.[38] The class C3 relates to unfolding of the LNR-C linker to the HD domain, and most importantly to pulling out and unfolding the β5 beta strand from the HD domain. The β5 beta strand comprises the S2 site. Finally, the class C4 relates any further allowable unfolding events within the HD domain.

In addition, we find evidence for direct interactions between the I27 domains and the NRR1 domain. Namely, in the free MD simulations the five proteins of the construct can sometimes assume a helical arrangement, with contacts between NRR1 and its I27 neighbors, especially the third I27. Recent computer simulations [51] reported on a similar effect for I27 proteins flanking the C2A and C2B domains of Human Synaptotagmin 1. It was observed that the I27 interacts sporadically with a C2B sub-domain, and that such interactions lower the unfolding forces of the I27 modules.

Due to high pulling rates the SMD simulations generate much larger forces than those seen in the AFM experiments. We follow the typical assumption that similar paths are sampled at different pulling rates, only with fewer details visible at higher speeds. Then, we can calibrate the forces obtained by the SMD using the Evans-Ritchie model [52]. This widely accepted model for the ligand-receptor rupture, proposes a logarithmic dependence between the rupture force and the force loading rate. The SMD obtained unfolding events for I27 are observed at about 700 pN (Figure 14B). Same events are observed at about 200 pN in the FX AFM study (Figure 13). Applying such force calibration to the SMD predicated conformational transitions within the

NRR1 domain yields that such transition should occur at about 100 pN in the FX AFM. And indeed, typical forces detected by our FX AFM experiments for major mNRR1 conformational transitions fall within such a range of forces.

I.4.2 Sequential mNRR1 unfolding

Our analysis points out that the classes C1, C2, and C3 are necessary prerequisites of the S2 site exposure. Using the FX AFM data we test whether these classes of events appear in sequence or in random. To do so, we calculate probability of any particular unfolding sequence of events. For example we test a most intuitive sequence, *s1*, where C1 events precede C2, which precedes C3, and then C4 (C1 -> C2 -> C3 -> C4). Then we test permutations of *s1*. These tests, however, are hindered by three main limitations of the AFM experiments: 1) initial adhesion events masking some events, 2) tip-sample contact rupture before a complete mNRR1 unfolding, and 3) protein pickup not always on the end opposite to the adhering end. As a result, we measured only very few FX traces showing a complete mNRR1 unfolding sequence. Thus, we extend our calculations to include three-element and two-element chunks of each given sequence. For example, an extended sequence *s1* called "s1_ext" includes C1 -> C2-> C3, C2 -> C3 -> C4, C1 -> C2, C2 -> C3 and C3 -> C4. We obtain that the probability of the *s1* sequence $P(s1\_ext) = 0.16 \pm 0.02$ dwarfs any other probabilities for any other extended sequence, see Tables D1 and D2 in Appendix D. This finding is in strong support of the sequential mNRR1 unfolding model. The sequential model is justified as follows. An NRR domain, comprised of four sub-domains, maintains a compact form, where its three LNR sub-domains wrap around a hetero-dimerization (HD) sub-domain composed of HD-N and HD-C parts. The S2 site is quite

buried in a HD domain despite being located only 12 residues away from the NRR C-end. This is because a linker between LNR-A and LNR-B domains interacts strongly with the HD domain by forming a hydrophobic plug, which occludes the S2 site. Additional S2 site protection is achieved by interactions between LNR-B and one of the alpha helices within the HD-C. Thus, exposure of the S2 site is facilitated when the LNR domains are removed and those events occur within the classes C1 and C2.

I.4.3 Molecular forces in the S2 site exposure

Mechanical stability of proteins is governed by major conformational changes associated with rupture of hydrogen bonds (H-bonds) [29, 53-54]. H-bonds rupture forces are modulated by hydrophobic interactions [55-56] and tertiary structure interactions, like spatial orientations of the domains themselves [30, 57-58].

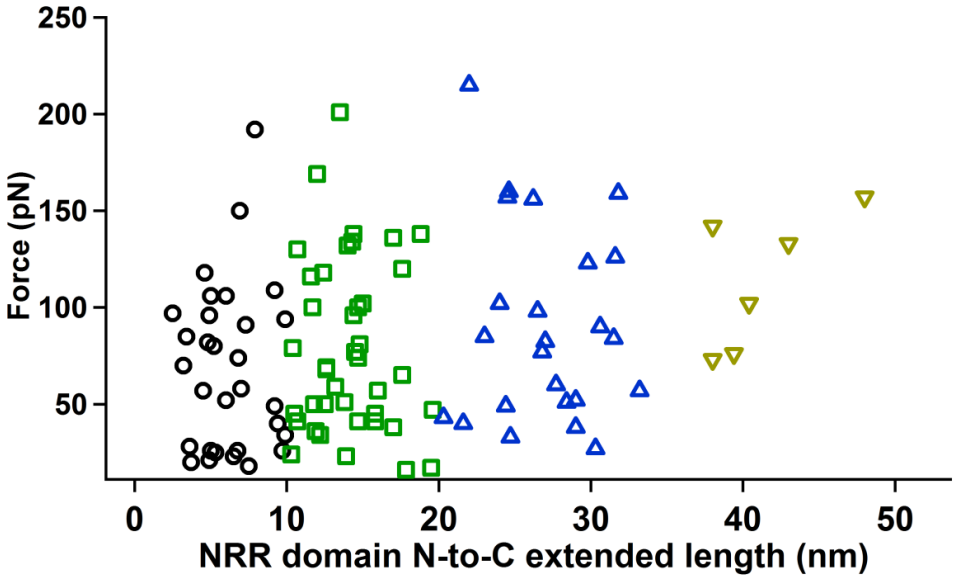

Figure 15 Scatter plot of forces associated with conformational transition events plotted in Fig. 14A. Each different class of events is distinguished using different symbols and colors.

Below, we discuss our experimentally obtained force values, see Figure 15, for the events related to the S2 site exposure and presented in Figure 14. We also note a substantial scatter of forces in Fig. 15. Based on a substantial scatter obtained for the I27 unfolding alone, see Ref. 38, we attribute the force scatter in Fig. 15 to at least two effects: rare interactions between I27 and NRR module (observed also in our SMD data), as well as systematic errors originating from the usage of three different types of AFM cantilevers (MLCT-C, MLCT-D, and BioLevers).

I.4.4 Classes C1-C2

Each LNR domain has some α helical structure and is stabilized by three disulfide bonds and calcium coordination. Unfolding forces of 69 ± 42 pN for the class C1 and 79 ± 45 pN for the class C2 correspond to unwrapping and partial unfolding of the LNR domains. This force range is much larger than the force of ca. 15 pN, which was obtained for purely α helical unfolding events like in calmodulins at unfolding rates similar to ours [30]. It relates better to the forces of 30 pN observed for disrupting the H-bonds between α-helix bundles in all-α proteins like spectrin [59]. However, even closer correspondence is found with 25 – 50 pN for disruption of α-helical bundles strengthened by strong hydrophobic interactions, like in all-α ankyrins [57, 60]. Such a scenario agrees quite well with the SMD predicted inter-helical H-bond breaking strengthened by substantial hydrophobic interactions between LNR-A domain, its LNR-AB linker, and HD domain.

I.4.5 Class C3

The HD domain, which immediately follows the LNR repeats, folds into an α/β sandwich with β1-α1-β2-β3-β4-α2-α3-β5 topology [17]. The forces related to the class C3 and a complete S2

site exposure are 90 ± 50 pN. We associated these events with unfolding and pulling out the β5 strand from the HD domain. Strand β5 H-bonds with a strand β1 and interacts hydrophobically with the α3 helix and the LNR-AB linker. Absence of the LNR-AB hydrophobic plug (class C1) decreases hydrophobic interactions between the β5 strand and the α3 helix, so that the β1- β5 H-bonding interactions are the last ones to be ruptured. Such a rupture is also expected to occur sequentially, i.e., in a zipper-like configuration. An example of such interactions is an all-β domain C2A, which unfolds at about 60 pN due to zipper-like rupture of the hydrogen bonds linking its two β strands.

I.4.6 Relation to molecular mechanisms of Notch activation

Unfolding forces detected here for each class C1 to C3 are around 100 pN and provide a substantial, but yet not insurmountable mechanical barrier. Such forces are much lower than 200 – 300 pN observed to unfold the Ig-like all β small proteins, which have a well-documented mechanical function [30]. Thus, the role of the mNRR1 domain is clearly not to withstand large forces. However, despite being only about 12 residues away from the C terminus of the mNRR1 domain, the S2 site is well mechanically protected against any accidental exposure. This is because, unfolding probability drops exponentially with required force and unfolding events within classes C1, C2, and C3 need to occur in sequence in order to expose the S2 site to ADAM10. Thus, strong Notch ligand-receptor interactions are necessary for the S2 site exposure, and consequently, Notch activation process. Measured here forces of ca. 100 pN and displacements of ca. 10 nm would expose the S2 site in a series of at least three steps, i.e., one step for each classes C1, C2, and C3. This would yield ca. 1000 pN*nm activation energy for each such step. Such energy is equivalent to 250 k_BT, which would need to be provided by

respective motor proteins. Motor proteins trap thermal fluctuations either by ratcheting diffusion of small, angstrom-sized, steps called "power strokes", or by rectifying nanometer-sized or larger, thermal displacements, i.e, "Brownian ratchet". However, an energy of 250 k_BT is too high to achieve in just one step by any known protein-based Brownian ratchet [61]. So what does exactly happen on a molecular level?

Two possibilities for the force-induced pulling of the NRR domains have been suggested: a "ligand recycling" model and a "pulling force" model [20-21, 62]. In the "ligand recycling" model, the ligands get more competitive through their posttranslational modifications, surface clustering, and/or cell trafficking [63] prior to the Notch binding. In the "pulling force" model, non-recycled ligands bind and pull the Notch receptor. The actual pulling has been suggested to occur through either clathrin-independent or clathrin-dependant trans-endocytosis [20-21, 61,62]. Clathrin dependent processes involve ligand ubiquitination and epsin adaptors binding to clathrin to produce clathrin-coated pits or clathrin-coat assembly around cargo molecules. Clathrin assembly is grabbed and pulled within the cell by a polymerizing network of actin filaments. Each actin polymerization event requires several k_BT of activation energy and generates up to several pN of force [60, 63]. Thus, very many of such steps would be necessary to overcome the 250 k_BT activation energy barrier.

Our results agree qualitatively with the results of Ahimou et al. obtained at similar pulling speeds [64]. Ahimou et al. used FX AFM to study adhesion forces between cells overexpressing the Delta ligands and cells overexpressing the Notch receptors. The Delta cells were fixed to tipless AFM cantilevers, and the Notch cells were fixed to a solid substrate. They used relatively stiff AFM cantilevers, and a low sensitivity AFM system, which limited their force sensitivity to sub-nanonewtons. Due to enormous cellular deformations at those conditions,

they probed large cell-cell contact areas and a large number of molecular adhesion events. Consequently, only cellular adhesion events of several nN were observed, and no force assessment for the rupture of a single Delta-Notch contact was provided. Noteworthy, Delta-Notch adhesion force is expected to be larger than the force necessary to expose the S2 site. Supposing that several tens of Notch receptors were affected at experimental conditions used by Ahimou et al., such an assumption holds true.

On the other hand, recent data by Wang and Ha [65] reported Notch activation forces of only about 12 pN, but due to the nature of their essay Notch activation reached an optimal level to be observed only after two days. Strikingly, the results of Wang and Ha agree very well with our findings. This is because, in the light of the previously mentioned Evans-Ritchie model [52], the unfolding forces are logarithmic function of the loading rates. In other words, even a very small force, of the order of several pN or less, is expected generate unfolding events within the NRR domain, but provided that it acts for a very long time. Thus, the decrease of the NRR unfolding forces is expected to be observed between our AFM study and the results of Wang and Ha. An amount of such a decrease is going to be larger than between the SMD simulations and AFM experiments. SMD simulations and AFM experiments differed by 5-6 orders of magnitude in the pulling speed and provided for 3-4 fold decrease in the observed unfolding forces. Our pulling speed of 400 nm/s requires about 20 – 30 ms to stretch the NRR domain by 10 nm steps required to expose the S2 site. This is 7-8 orders of magnitude faster than 2 days. Consequently, a decrease from about 100 pN forces to about 10 – 20 pN forces is expected between our AFM experiments and the results of Wang and Ha. In addition, explicit treatment of the potential energy barrier for the force-induced unfolding events, as well as accounting on the rebinding and refolding events is expected to produce lower forces at very small loading rates than predictions

of the Bell-Evans model.[66-68] Consequently, the mean S2 site exposure forces of the order of 10 pN as probed by Wang and Ha become plausible. However, it remains to be addressed how much time and force is needed for unbiased cells to activate Notch receptors.

Finally, the S2 site is not buried deeply within the mNRR1 domain, and some proteases like TACE are able to cut it in a native state. This has evolutionary benefits, since despite the fact that higher organisms like mammals have evolved to produce a more elaborate and more effective Notch activation mechanism than other species, they would still conserve a low output level S2 site cutting scheme by TACE. In fact, ligand independent activation of Notch supports recent suggestions by Kopan and Ilagan [3] that the NRR structures are dynamic and alternating between a "closed" and a hypothetical "open" state. Such a possibility would justify a low probability proteolitic access to the S2 site without ligand binding [46].

I.5 Conclusions and Future Work

We provided a molecular-level study for the force-induced conformational transitions in the NRR domain of mouse Notch 1. Using single molecule FX AFM experiments, we detected four classes of unfolding events. In the light of our SMD data, the first three of these classes, classes C1, C2, and C3 are necessary for major force-induced conformational transitions within the NRR domain. Mean unfolding forces associated with the S2 site exposure match well the typical forces associated with similar types of interactions in comparable systems. Through the conditional probability analysis of the most probable unfolding sequence we find that the S2 site is exposed only after C1, C2, and C3 events have occurred. Thus, the S2 site exposure requires passing several substantial mechanical barriers, which need to occur in a sequence. This mechanism provides an excellent control of the Notch activation processes and the need for

ligand binding in mammals. Finally, our results provide a clear fingerprint of the S2 site exposure, which is a stepping stone for further studies of the force-induced Notch activation by ADAM proteases. To fully unravel a potential of the S2 site as the checkpoint and target in Notch signaling, similar AFM studies should be realized on all four variants of Notch receptors and on a cellular level. The effects of most common mutations in the neighborhood of the S2 sites within the HD domain and within the LNR regions should be addressed as well.

In the future work an important step is to improve the experimental protocol to increase the frequency of single protein pick up. Some suggestions along these lines are to use better protein purification techniques, such as size exclusion FPLC purification or anion exchange. Another improvement can be achieved by chemically crosslinking the protein to a functionalized substrate.

The work done for understanding the unfolding pattern of the NRR domain of mouse Notch 1 can be extended to prove the S2 site exposure and enzymatic cutting using force-clamp AFM mode. Finally, using different concentrations of the enzymes we can study the kinetics of the S2 site cutting.

I.7 References

1. Artavanis-Tsakonas, S.; Muskavitch, M. A., *Curr Top Dev Biol* **2010,** *92*, 1-29.
2. Bray, S. J., *Nat Rev Mol Cell Bio* **2006,** *7*, 678-689.
3. Kopan, R.; Ilagan, M. X. G., *Cell* **2009,** *137*, 216-233.
4. Kokkoli, E.; Kasinskas, R. W.; Mardilovich, A.; Garg, A., *Biomacromolecules* **2005,** *6*, 1272-1279.
5. Gridley, T., *Hum Mol Genet* **2003,** *12*, R9-R13.
6. Louvi, A.; Arboleda-Velasquez, J. F.; Artavanis-Tsakonas, S., *Dev Neurosci-Basel* **2006,** *28*, 5-12.
7. Weng, A. P.; Ferrando, A. A.; Lee, W.; Morris, J. P.; Silverman, L. B.; Sanchez-Irizarry, C.; Blacklow, S. C.; Look, A. T.; Aster, J. C., *Science* **2004,** *306*, 269-271.
8. Politi, K.; Feirt, N.; Kitajewski, J., *Semin Cancer Biol* **2004,** *14*, 341 - 347.
9. Shi, W.; Harris, A. L., *J Mammary Gland Biol Neoplasia* **2006,** *11*, 41-52.
10. van Es, J. H.; van Gijn, M. E.; Riccio, O.; van den Born, M.; Vooijs, M.; Begthel, H.; Cozijnsen, M.; Robine, S.; Winton, D. J.; Radtke, F.; Clevers, H., *Nature* **2005,** *435*, 959-963.
11. Wang, X. Y.; Zhang, R.; Lian, S., *Clin Exp Dermatol* **2011,** *36*, 69-76.
12. Weinmaster, G., *Mol Cell Neurosci* **1997,** *9*, 91-102.
13. Kovall, R. A.; Blacklow, S. C., *Curr Top Dev Biol* **2010,** *92*, 31-71.
14. Kopan, R., Notch Signaling. *Csh Perspect Biol* **2012,** *4*.
15. Gordon, W. R.; Vardar-Ulu, D.; Histen, G.; Sanchez-Irizarry, C.; Aster, J. C.; Blacklow, S. C., *Nat Struct Mol Biol* **2007,** *14*, 295-300.
16. Mumm, J. S.; Schroeter, E. H.; Saxena, M. T.; Griesemer, A.; Tian, X. L.; Pan, D. J.; Ray, W. J.; Kopan, R., *Mol Cell* **2000,** *5*, 197-206.
17. Gordon, W. R.; Roy, M.; Vardar-Ulu, D.; Garfinkel, M.; Mansour, M. R.; Aster, J. C.; Blacklow, S. C., *Blood* **2009,** *113*, 4381-90.
18. Gordon, W. R.; Vardar-Ulu, D.; Histen, G.; Sanchez-Irizarry, C.; Aster, J. C.; Blacklow, S. C., S *Nature Structural & Molecular Biology* **2007,** *14*, 455-455.
19. Kopan, R.; Schroeter, E. H.; Weintraub, H.; Nye, J. S., *P Natl Acad Sci USA* **1996,** *93*, 1683-1688.

20. Parks, A. L.; Klueg, K. M.; Stout, J. R.; Muskavitch, M. A., *Development* **2000,** *127*, 1373-85.

21. Nichols, J. T.; Miyamoto, A.; Olsen, S. L.; D'Souza, B.; Yao, C.; Weinmaster, G., *J Cell Biol* **2007,** *176*, 445-458.

22. Varnum-Finney, B.; Wu, L. Z.; Yu, M.; Brashem-Stein, C.; Staats, S.; Flowers, D.; Griffin, J. D.; Bernstein, I. D., *J Cell Sci* **2000,** *113*, 4313-4318.

23. Miyamoto, H.; Nihonmatsu, I.; Kondo, S.; Ueda, R.; Togashi, S.; Hirata, K.; Ikegami, Y.; Yamamoto, D., *Gene Dev* **1995,** *9*, 612-625.

24. Beglova, N.; Blacklow, S. C.; Takagi, J.; Springer, T. A., *Nat Struct Biol* **2002,** *9*, 282-287.

25. Vardar, D.; North, C. L.; Sanchez-Irizarry, C.; Aster, J. C.; Blacklow, S. C., *Biochemistry-Us* **2003,** *42*, 7061-7067.

26. Chen, J.; Zolkiewska, A., *PLoS One* **2011,** *6*, e22837-1 - e22837-11.

27. Szoszkiewicz, R., Force-extension (FX) and force-clamp (FC) AFM spectroscopies in investigating mechanochemical reactions and mechanical properties of single biomolecules. In *Scanning Probe Microscopy in Nanoscience and Nanotechnology*, Bhushan, B., Ed. Springer-Verlag: Heidelberg, 2010; pp 395-424.

28. Rief, M.; Gautel, M.; Oesterhelt, F.; Fernandez, J. M.; Gaub, H. E., *Science* **1997,** *276*, 1109-1112.

29. Borgia, A.; Williams, P. M.; Clarke, J., *Annu Rev Biochem* **2008,** *77*, 101-125.

30. Carrion-Vazquez, M.; Oberhauser, A. F.; Fisher, T. E.; Marszalek, P. E.; Li, H. B.; Fernandez, J. M., *Prog Biophys Mol Bio* **2000,** *74*, 63-91.

31. Carrion-Vazquez, M.; Oberhauser, A. F.; Fowler, S. B.; Marszalek, P. E.; Broedel, S. E.; Clarke, J.; Fernandez, J. M., *Proc Natl Acad Sci U S A* **1999,** *96*, 3694-9.

32. Dey, A.; Szoszkiewicz, R., *Nanotechnology* **2012,** *23*.

33. Lu, H.; Isralewitz, B.; Krammer, A.; Vogel, V.; Schulten, K., *Biophys J* **1998,** *75*, 662-71.

34. Stephenson, N. L.; Avis, J. M., *P Natl Acad Sci USA* **2012,** *109*, E2757-E2765.

35. Tiyanont, K.; Wales, T. E.; Aste-Amezaga, M.; Aster, J. C.; Engen, J. R.; Blacklow, S. C., *Structure* **2011,** *19*, 546 - 554.

36. Gordon, W. R.; Vardar-Ulu, D.; L'Heureux, S.; Ashworth, T.; Malecki, M. J.; Sanchez-Irizarry, C.; McArthur, D. G.; Histen, G.; Mitchell, J. L.; Aster, J. C.; Blacklow, S. C., *PLoS One* **2009,** *4*, e6613.

37. Aster, J. C.; Simms, W. B.; Zavala-Ruiz, Z.; Patriub, V.; North, C. L.; Blacklow, S. C., *Biochemistry-Us* **1999,** *38*, 4736-4742.

38. Ploscariu, N.; Kuczera K.; Malek K. E.;Wawrzyniuk, M.; Dey, A.; Szoszkiewicz, R., *J. Phys Chem B.*, **2014**;*118(18)*, 4761-70

39. Malek, K.E.; Szoszkiewicz, R., *J. Biol. Phys.*, **2014**, 40, 15-23

40. Grandbois, M.; Beyer, M.; Rief, M.; Clausen-Schaumann, H.; Gaub, H. E., *Science* **1999,** *283*, 1727 - 1730.

41. Muller, D. J.; Kessler, M.; Oesterhelt, F.; Moller, C.; Oesterhelt, D.; Gaub, H., *Biophys J* **2002,** *83*, 3578-3588.

42. Taylor, J. R., *An Introduction to Error Analysis, Second Edition, p. 261-293*. University Science Books: Sausalito, CA, 1997; pp p. 261-293.

43. Szoszkiewicz, R.; Ainavarapu, S. R.; Wiita, A. P.; Perez-Jimenez, R.; Sanchez-Ruiz, J. M.; Fernandez, J. M., *Langmuir* **2008,** *24*, 1356 - 1364.

44. Zolkiewska, A., *Cell Mol Life Sci* **2008,** *65*, 2056-68.

45. Christian, L. M., *Fly (Austin)* **2012,** *6*, 30-4.

46. Delwig, A.; Rand, M. D., *Cell Mol Life Sci* **2008,** *65*, 2232-43.

47. Lu, H.; Isralewitz, B.; Krammer, A.; Vogel, V.; Schulten, K., *Biophys J* **1998,** *75*, 662-671.

48. Lu, H.; Schulten, K., *Biophys J* **2000,** *79*, 51-65.

49. Marszalek, P. E.; Lu, H.; Li, H.; Carrion-Vazquez, M.; Oberhauser, A. F.; Schulten, K.; Fernandez, J. M., *Biophys J* **2000,** *78*, 448a-448a.

50. Bustamante, C.; Marko, J. F.; Siggia, E. D.; Smith, S., *Science* **1994,** *265*, 1599-600.

51. Duan, L.; Zhmurov, A.; Barsegov, V.; Dima, R. I., *J Phys Chem B* **2011,** *115*, 10133-10146.

52. Evans, E.; Ritchie, K., *Biophysical Journal* **1997,** *72*, 1541-1555.

53. Kumar, S.; Li, M. S., *Phys Rep* **2010,** *486*, 1-74.

54. Isralewitz, B.; Gao, M.; Schulten, K., *Curr Opin Struc Biol* **2001,** *11*, 224-230.

55. Paci, E.; Karplus, M., *J Mol Biol* **1999,** *288*, 441-59.

56. Lee, G.; Abdi, K.; Jiang, Y.; Michaely, P.; Bennett, V.; Marszalek, P. E., *Nature* **2006,** *440*, 246-249.
57. I27 is the 27th module of a human cardiac titin.
58. Rief, M.; Pascual, J.; Saraste, M.; Gaub, H. E., *J Mol Biol* **1999,** *286*, 553-61.
59. Li, L.; Wetzel, S.; Pluckthun, A.; Fernandez, J. M., *Biophys J* **2006,** *90*, L30-2.
60. Mogilner, A.; Oster, G., *Curr Biol* **2003,** *13*, R721-33.
61. Musse, A. A.; Meloty-Kapella, L.; Weinmaster, G., *Semin Cell Dev Biol* **2012,** *23*, 429-36.
62. Le Borgne, R., *Curr Opin Cell Biol* **2006,** *18*, 213-22.
63. Kaksonen, M.; Toret, C. P.; Drubin, D. G., *Nat Rev Mol Cell Biol* **2006,** *7*, 404-14.
64. Ahimou, F.; Mok, L. P.; Bardot, B.; Wesley, C., *J Cell Biol* **2004,** *167*, 1217-29.
65. Wang, X. F.; Ha, T., *Science* **2013,** *340*, 991-994.
66. Hummer, G.; Szabo, A., *Biophys. J*, **2003**, *85*, 5-15
67. Dudko, O. K.; Hummer, G.; Szabo, A., *Phys. Rev. Lett.*, **2006**, *96*, 108101
68. Pierse, C. A.; Dudko, O. K., *Biophys. J.*, **2013**, *105*, L19-L22.

Part II - Study of Conformational Changes in Peptides and Proteins

In this part we develop a model to obtain quantitative measurements of the molecular stiffness and mechanical energy dissipation factors for selected simple proteins and polypeptides from the AFM force spectroscopy measurements. Using this model, we predict the shifts of several thermally excited resonance frequencies of atomic force microscopy cantilevers when unfolding a single molecule of a protein based on values of stiffness and dissipation factors found in literature. Next, we provide partial experimental validation of this model by measuring the shifts of excited resonance frequencies when AFM cantilever is in contact with protein and peptide samples.

II.1 A Method to Measure Nanomechanical Properties of Biological Objects

Key processes related to development and tissue homeostasis depend on mechanical properties of the involved proteins, cells, and other biological objects (BO).[1-6] It has become possible to interrogate such processes in situ and with a spatial resolution down to a single molecule.[1,7] Quantitative, fast, and non-destructive nanomechanical measurements of BOs are becoming possible too. For example, one can learn about forces associated with major conformational transitions during mechanical stretching of single proteins using optical and magnetic tweezers and AFM.[8-11]

Recent advances in high bandwidth AFM and compliant low-drift AFM cantilevers make it possible to visualize, manipulate, and indent single proteins, biological cells, and their films.[8,11,12] Calibrated AFM force—distance curves yield contact stiffness or elastic modulus of BOs.[13,14]

Techniques utilizing small-amplitude vibrations of the AFM cantilevers provide elastic moduli of agglomerated proteins and single cells non- destructively.[15,16] Use of ultrasonic techniques for nanomechanical measurements additionally eliminates mechanical hysteresis of the AFM cantilevers.[17]

Exploitation of a multi-frequency response of the AFM cantilever is expected to provide many topographical and nanomechanical parameters simultaneously and quickly.[18] Bimodal AFM methods have been already implemented.[18-20] These methods measure amplitudes and phases of the first two flexural resonance modes of the vibrating AFM cantilever in intermittent contact with the sample. The amplitudes and phases are manipulated to produce the maps of local stiffness, stiffness gradient, and the viscoelastic dissipation in contact with cells and protein films. Similar approaches have been also applied to torsional excitations of the AFM cantilevers.[21] While multifrequency AFM is highly accurate in theory, complicated and highly non-linear dependencies of the amplitudes and phases with measured tip-sample distance as well as their couplings can produce experimental artifacts.[18,20,22] Thus, complementary approaches to obtain quick and complete nanomechanical characterization of BOs are desirable.

We propose to measure stiffness and other nanomechanical properties of a BO from the shifts of the resonance frequencies for a thermally excited AFM cantilever in contact with such an object. The number of simultaneously elucidated nanomechanical parameters depends only on the number of the resonances measured, i.e., electronics AFM bandwidth.[23]

Using a similar approach, Dupas et al. elucidated local stiffness and internal friction of some engineering materials.[24] However, while measurements on engineering samples use stiff AFM cantilevers in air, the measurements on biological entities need to use compliant AFM cantilevers in biological media. For cantilevers with small aspect ratio, problems are exacerbated due to issues in providing analytical description of the hydrodynamic flow.[25] Currently, such cantilevers are among the most appropriate ones for probing compliant BOs. Thus, a comprehensive approach needs to be developed to accurately fit flexural resonances of compliant AFM cantilevers with a small aspect ratio[26] in contact with biological specimens in dissipative media.

In this subchapter, we develop a method to fit multiple resonance frequencies for compliant AFM cantilevers with a small aspect ratio in the biologically relevant phosphate buffered saline (PBS) buffer. The cantilevers are clamped on one end with the other end free. For each cantilever, we obtain geometrical and material properties. Properties with largest uncertainties, e.g., thickness, are determined from the fit of several consecutive resonance frequencies in air. Other geometrical and material parameters are measured or calculated.

To fit resonances in air, we use the model of Dupas et al. developed for a free cantilever in vacuum.[24] We obtain satisfactory agreement between fitted and measured resonances in air. Better agreement is obtained, when we correct the model of Dupas et al. for air damping using the results of Sader.[27] These developments are a starting point to fit the resonance frequencies of the cantilevers in the PBS buffer and introduce corrections to properly account for the hydrodynamic flow. We introduce a generalized hydrodynamic function, which we obtain from a set of several cantilevers.

We apply our model to obtain shifts in resonance frequencies expected in contact between a cantilever and a protein sample, and provide an error progression analysis.

We use Olympus AFM biolevers model BL-RC150VB, type "B," in air and in Dulbecco's PBS buffer (137mM NaCl, 3mM KCl, 2mM KH_2PO_4, and 8mM $Na_2HPO_4 \cdot 7H_2O$) from Midsci, USA. Thermal deflection signal of freely vibrating AFM cantilevers is fast Fourier transformed to produce amplitude spectra using our custom AFM setup as described in Ref. 28. Resonance frequencies from the cantilever's amplitude spectra are read using multipeak fit package with Voigt model in Igor Pro, Wavemetrics, USA. We fit the resonance frequencies using procedures written in Igor. Fit errors are the relative errors between fitted and measured resonance frequencies.[29] The electronics bandwidth is 250 kHz.[28] Figure 19 shows an AFM cantilever as a rectangular Euler- Bernoulli beam interacting with an arbitrary body.

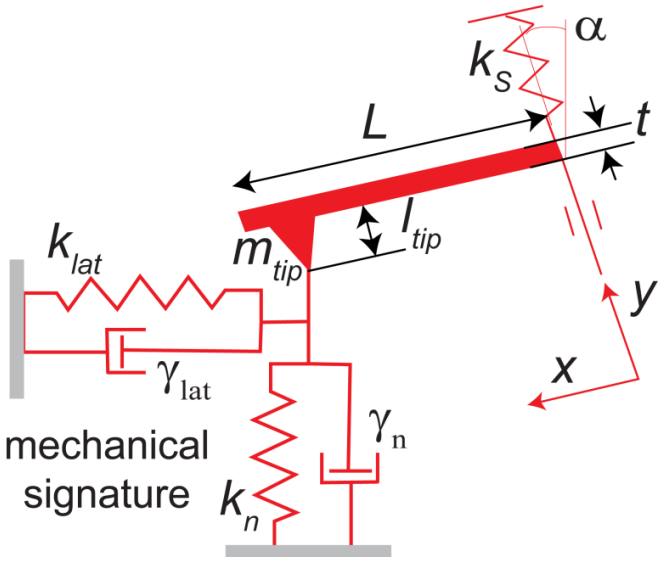

Figure 16. Adapted model to obtain mechanical signatures of a biological object in contact with an AFM cantilever.

The cantilever has length L, width b, thickness t, density ρ, Young's modulus E, tip length h_{tip}, and tip mass m_{tip} attached at a point βL along the beam. The cantilever is tilted at an angle α with respect to the normal to the substrate. One beam end is clamped by a support spring

with an elastic spring constant k_S. The other end is either left free or in contact—via its tip—with a body of interest. The body of interest is abstracted by an ensemble of dissipative springs providing its *mechanical signature*. We use the Kelvin-Voigt model, where *spring constants k* are in parallel with their corresponding *molecular damping factors γ*. BOs and proteins, in particular, exhibit distinctively different visco-elastic properties along each pulling/pushing direction.[30,31] Thus, in Figure 16, we adopt only a reduced mechanical signature with two dissipative and mutually perpendicular springs: one along a normal force-exerting direction with k_n and γ_n, and the other with k_{lat} and γ_{lat}.[32]

Dupas et al.[24] showed how to obtain the values of k and γ analytically for the cantilever in contact with a viscoelastic body as in Fig. 16 and obeying an equation of a moving Euler-Bernoulli beam:

$$EI\frac{\partial^4 y}{\partial x^4} + \mu\frac{\partial^2 y}{\partial t^2} = 0 \quad \textbf{(1)}$$

Here, I is the areal moment of inertia, y is the vertical deflection, and μ is the mass of the cantilever over its length. The solution of Eq. (1) is of the form

$$y(x,t) = y(x)e^{i\omega t} \quad \textbf{(2)}$$

with y(x) of the form

$$y(x) = A_1[\cos(\kappa x) + \cosh(\kappa x)] + A_2[\cos(\kappa x) - \cosh(\kappa x)] + A_3[\sin(\kappa x) + \sinh(\kappa x)] + A_4[\sin(\kappa x) - \sinh(\kappa x)] \quad \textbf{(3)}$$

Here: ω is an angular frequency, κ is a wave vector, and parameters A_1–A_4 are obtained from boundary conditions.

Equations (1)–(3) extend to any BO provided that an accurate model for the cantilever is developed in appropriate media.

Fig. 17 presents a typical amplitude vs. frequency spectrum for the BL-RC150VB cantilever obtained from its thermal excitations in air.[28] Three flexural resonances at frequencies of 11.6 kHz, 76.1 kHz, and 219 kHz are fitted using the model of Dupas et al., which depends on the following variables*: L, b, t, β, h_{tip}, m_{tip}, k_S, α, ε, E, a_0*. Here, ε is a mean position of the laser beam on the AFM cantilever, and a_0 is the cantilever's excitation amplitude. In order to get an accurate agreement between measured and modeled resonance frequencies, we fit only four variables, the values of *t, b,* h_{tip}, and *E,* which have the largest uncertainties, and treat the other variables as parameters.

We also constrain the four variables as follows. The values of β are estimated from optical images of the AFM cantilevers and constrained to 0.95 ± 0.05. Similarly, the values of h_{tip} are constrained to 7.5 ± 2.5 μm. Using the manufacturer's scanning electron microscopy (SEM) measurements, the cantilever's thickness is constrained to 200 ± 26 nm, and the value of *E* to 155 ± 10 GPa.[34-36] The cantilever's density ρ is related to thickness using a weighted average with the density of silicon nitride ρ_{SiNx} = 3100 kg/m^3, the density of the 10 nm chromium layer ρ_{Cr} =7140 kg/m^3, and the density of the 50 nm gold coating ρ_{Au} = 19320 kg/m^3.[33] The values of L and b are obtained within 1% and 2% relative errors, respectively, using optical microscopy.[35] The value of b is further constrained through measurements of the torsional resonance frequencies, when they are visible on the amplitude spectra.[37] Using the manufactures' SEM images of the tips, the value of m_{tip} is calculated supposing that a tip is half of a pyramidal

shell with thickness t.[38] The value of k_S = 600±30 N/m is measured using a dynamometer.[39] Since only resonance frequencies are fitted, and not their shape, the values of ε and a0 are arbitrary selected as 0.8 L, and 10^{-22} m, respectively.[24] The values of $k_n = k_{lat} = \gamma_n = \gamma_{lat} = 0$.

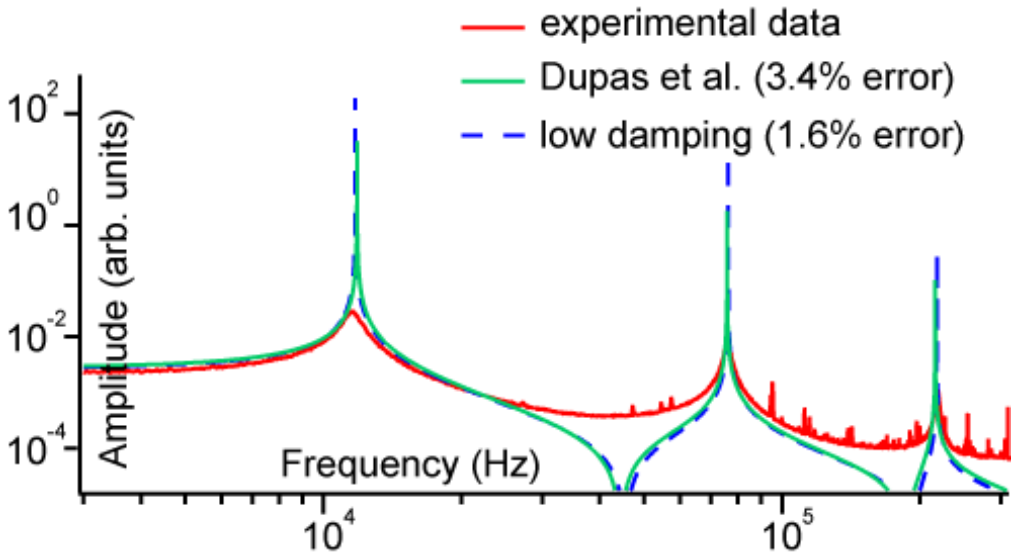

Figure 17. A typical thermal amplitude spectrum of flexural resonances of the cantilever c1, see Table I, in air. A solid line is a fit of the Dupas model.[24] A dashed line is a more accurate description obtained via Eq. (6).

The model of Dupas et al. produces a reasonable fit in Fig. 17 with an accumulated error of 3.4% over three flexural resonance frequencies. However, for cantilevers with high quality factors Q of 50, Dupas et al. obtained relative errors of less than 0.5% for each resonance frequency. This is because the model is essentially fitting the resonances in vacuum and, thus, with no damping. The cantilevers used here have modest quality factors of 10–15 in air, so air damping cannot be neglected.[40]

In the limit of Q >> 1, Sader et al. developed a correction to the resonance frequencies of the AFM cantilevers due to low damping by a hydrodynamic flow:[27]

$$\frac{\omega_n^{vac}}{\omega_n^{fluid}} = \left(1 + \frac{\pi \rho_{fluid} b}{4\rho t} \Gamma_r\right)^{1/2}$$

(4)

where $\rho_{fluid} = \rho_{air} = 1.18$ kg/m^3 is air density; ω_n^{fluid} and ω_n^{vac} are angular frequencies of the n–th resonance mode of the AFM cantilever in fluid (here: air) and vacuum, respectively; and Γ_r is the real part of the hydrodynamic function Γ_{rect} from the footnote (20) in Ref. 27.

The values of Γ_r apply to non-ideal rectangular cantilevers with an aspect ratio of 3.9 and more.[25] Thus, they are almost applicable to our cantilevers, which have an aspect ratio of 3.3 ± 0.1. Consequently, we upgrade the model of Dupas et al. by using the results of Sader et al.[27] To do so, we need to translate the corrections in resonance frequencies from Eq. (4) into a wave vector κ from Eq. (3). From Eqs. (1)–(3), we find

$$\kappa = (\omega_n^{vac})^{1/2} \left(\frac{\mu}{EI}\right)^{1/4}$$

(5)

Combining Eqs. (4) and (5), we calculate the wave vector κ_{fluid} in the arbitrary fluid

$$\kappa_{fluid} = \left(\frac{4\pi f}{t}\right)^{1/2} \left(\frac{3\rho}{E}\right)^{1/4} \left(1 + \frac{\pi \rho_{fluid} b}{4\rho t} \Gamma_r\right)^{1/4}$$

(6)

Here, $\omega_n^{fluid} = 2\pi f$ with f being frequency, and $I = t^3 b/12$. The dashed line in Fig. 17 plots the results of the model of Dupas et al. with low hydrodynamic damping, i.e., using κ_{fluid} from Eq. (6) in air. Excellent agreement with the experimental data is obtained and we extend

this analysis to three more BL-RC150VB cantilevers itemized as c2 to c4 in Table 2. Errors accumulated over their fitted resonance frequencies are 2% to 6%.[35]

We now want to fit the resonance frequencies in the PBS buffer. To start with, the dashed line in Fig. 18 plots the results of our upgraded model of Dupas et al. with Eq. (6), where $\kappa_{fluid} = \kappa_{PBS}$, and density of PBS ρ_{PBS} = 998 kg/m^3.[41,47]

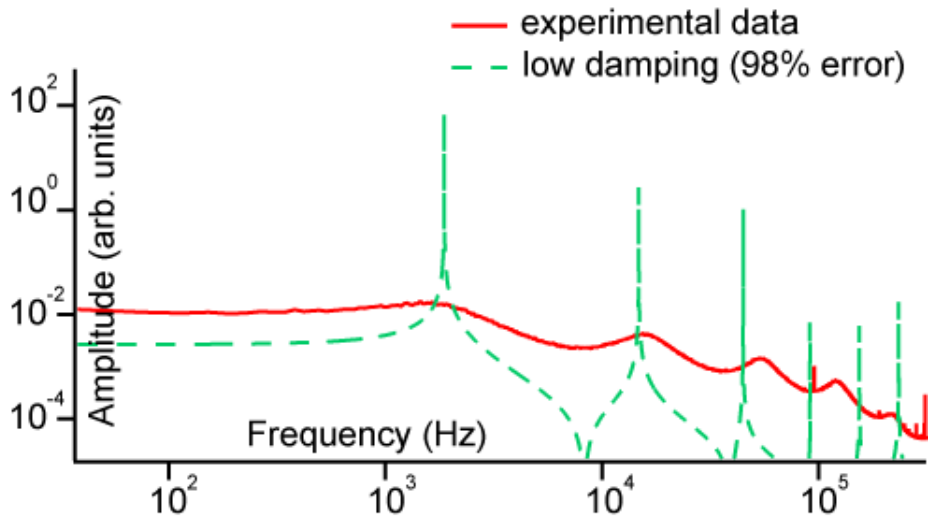

Figure 18. A typical thermal amplitude spectrum of flexural resonances of the cantilever c1 in the PBS buffer. A dashed line is a fit of the model via Eq. (6) applied to PBS.

Errors of 98% are obtained, so a more accurate description is needed.[35]

The quality factors of our cantilevers in PBS are about 1.5 for the first resonance at 1.50 kHz and about 2 for higher resonances. Those quality factors are larger than "1", but an actual hydrodynamic function is expected to differ from Γ_r .

Thus, we need to find the generalized hydrodynamic function $\mathcal{H}_r$ to substitute for Γ_r in Eq. (6). Sader et al.[25] suggested that for a rectangular cantilever with an arbitrary aspect ratio, an imaginary component of the generalized hydrodynamic function $\mathcal{H}_{im}$ can be approximated by a power law of the Reynolds number Re. The value of Re

$$Re = \frac{2\pi f \rho_{PBS} b^2}{4\eta_{PBS}} \tag{7}$$

where η_{PBS} is the viscosity of PBS. Thus, we suggest a complementary power law to describe the real component of the generalized hydrodynamic function $\mathcal{H}_r$. In order to find $\mathcal{H}_r$ in the limit of small damping, we manipulate Eq. (4) to yield

$$\mathcal{H}_r(Re) = \left[\left(\frac{\omega_n^{vac}}{\omega_n^{PBS}}\right)^2 - 1\right]\left(\frac{4\rho t}{\pi \rho_{PBS} b}\right) \tag{8}$$

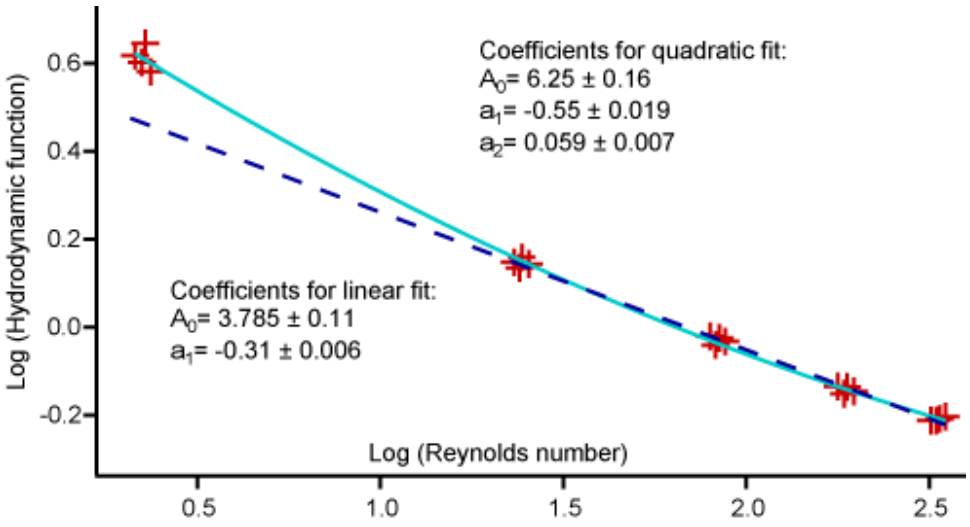

Figure 19. Log-log plot of $\mathcal{H}_r$ vs. Re for the cantilevers listed in Table 2.

The values of ω_n^{vac} are obtained from the model of Dupas et al. using the cantilever properties from Table 2.

Cantilever	L (μm)	b (μm)	t (nm)	h_{tip} (μm)	m_{tip} (pg)	ρ (kg/m^3)	β
c1	99.6	30.2	179	10.0	102	7856	0.91
c2	100.7	31.1	174	10.0	104	7993	0.97
c3	100.5	31.1	177	9.5	101	7910	0.95
c4	100.4	30.2	174	10.0	101	7993	0.91

Table 2. Properties of the cantilevers used in this study.

In addition, due to the lack of hydrodynamic damping in vacuum, we include two more resonances calculated using the model of Dupas et al. at 438 kHz and 727 kHz, respectively. These resonances become visible in PBS at 120 kHz and 215 kHz in Fig. 17, respectively.[35] To visualize the power law dependence of $\mathcal{H}_r$ with Re, Fig. 19 plots the decimal logarithm of $\mathcal{H}_r$ vs. the decimal logarithm of Re for the five resonances observed in PBS and for the cantilevers c1 to c4. Our data are best fitted with a quadratic relation, i.e.,

$$log\mathcal{H}_r = a_0 + a_1 logRe + a_2 (logRe)^2 \quad \textbf{(9)}$$

which yields:

$$\mathcal{H}_r = A_0 (Re)^{[a_1 logRe + a_2 (logRe)]} \quad \textbf{(10)}$$

with $A_0 = 10^{a}{}_0$.

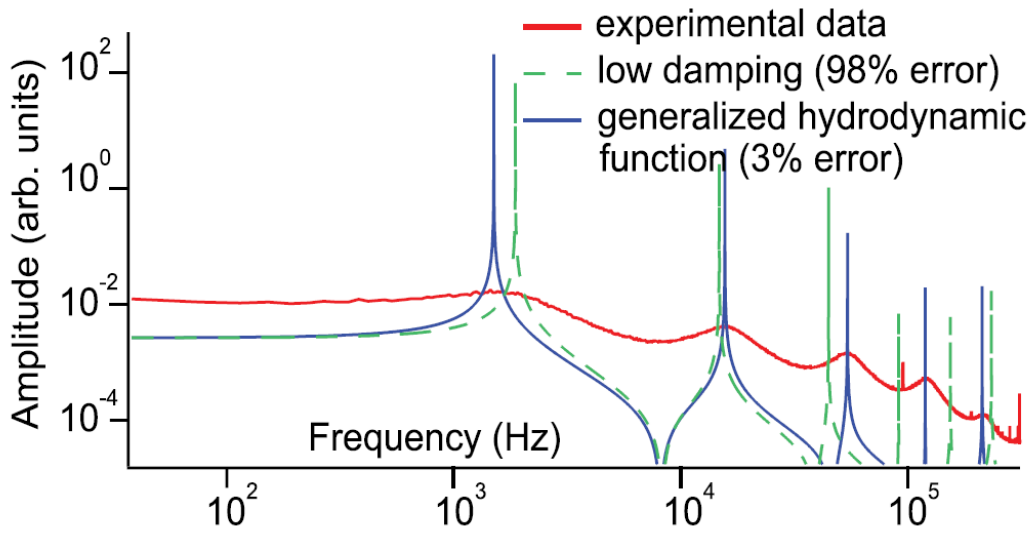

Figure 20. A typical thermal amplitude spectrum of flexural resonances of the cantilever c1 in the PBS buffer. A dashed line is a fit of the model via Eq. (6) applied to PBS. A solid line presents a more accurate description obtained via Eq. (11).

Numerical values of the fit coefficients are presented in Fig. 19. Fit quality is estimated by calculating the value of χ^2_{red}.[42] We obtain $\chi^2_{red} \simeq 2 \cdot 10^{-4}$ indicating a very good fit.

We begin our discussion of Fig. 19 by testing statistical significance of the quadratic term, i.e., a statistical hypothesis H_0: $a_2 = 0$. We find that with 99% confidence level H_0 is not true and a_2 is not zero.[43] However, the values of $log\mathcal{H}_r$ become linearly dependent on *log Re*, once we omit the data for the first resonance, i.e., with *log Re* < 0.5. A small value of $\chi^2_{red} \simeq 1 \cdot 10^{-4}$ suggests a cross-over between two regimes of the hydrodynamic function. Such a cross-over is expected, since the first resonance at 1.50 kHz is the most damped out of all the modes. Thus, hydrodynamic flow is expected to be described by a different functional dependence in the case of the first mode when compared to the other modes. However, with a quadratic fit, we capture a hydrodynamic correction to the wave vector κ_{PBS}, which applies to both hydrodynamic regimes

$$\kappa_{PBS} = \left(\frac{4\pi f_{PBS}}{t}\right)^{1/2} \left(\frac{3\rho}{E}\right)^{1/4} \left(1 + \frac{\pi \rho_{PBS} b}{4\rho t} A_0 (Re)^{[a_1 logRe + a_2 (logRe)]}\right)^{1/4} \quad \textbf{(11)}$$

Using the coefficients A_0, a_1, and a_2 within their errors, we refit the data in Fig. 20 and find only 3% agglomerated error over five resonances. Extending such analysis to the remaining cantilevers yields errors between 3% and 11%.[35] Propagating the errors from the material and geometrical parameters, we obtain errors between 10% and 40% with an average of 20%.[35] We apply our model to predict the shifts of five resonances in contact with a folded protein molecule in PBS.

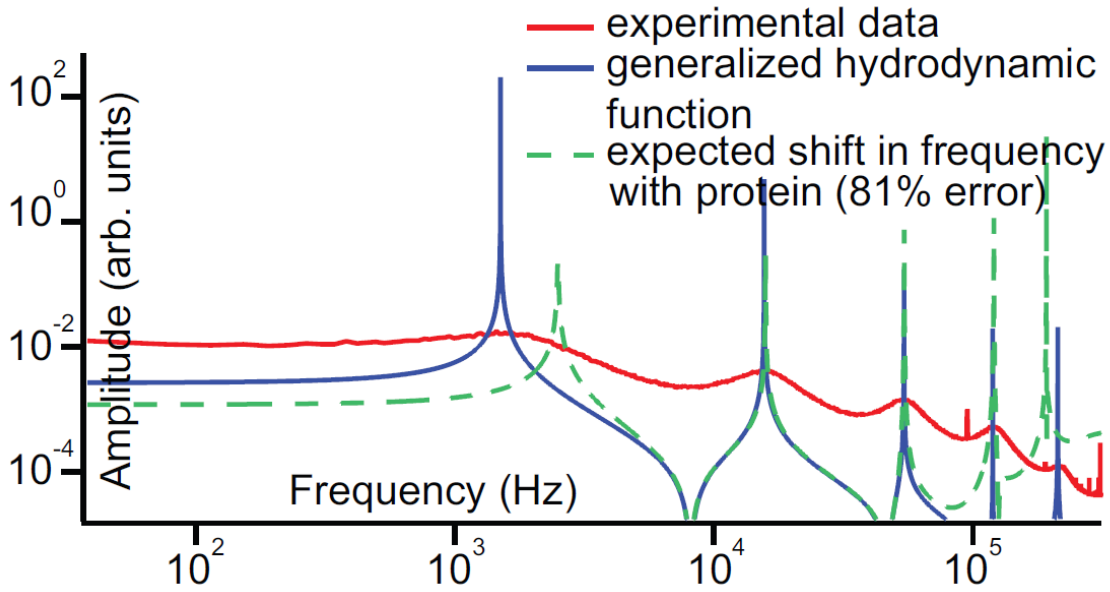

Figure 21. Application of our models to thermally excited flexural resonances of the cantilever described in the paper, while in contact with a folded protein molecule characterized by $k_n = k_{lat} = 10$ pN/nm and $\gamma_n = \gamma_{lat} = 10^{-8}$ kg/s in PBS.

Supposing that a protein has an expected normal elastic spring constant k_n of about 10 pN/nm[44-46] as well as $k_{lat} = k_n$ and $\gamma_n = \gamma_{lat} = 10^{-8}$ kg/s, see Ref. 45, one obtains well distinguishable 81% combined shift of the five resonance frequencies.[35] This is much larger than

our average propagated errors of 20%. However, 20% combined shifts in the five resonance frequencies would affect the values of k_n and γ_n as follows.

With $\gamma_n = 10^{-8}$ kg/s, k_n would need to change to either 5 or 15 pN/nm. With $k_n = 10$ pN/nm, γ_n would need to change to either $2 \cdot 10^{-8}$ or 10^{-9} kg/s. The values of $\gamma_n < 10^{-9}$ have no effect on error at $k_n = 10$ pN/nm. The variations in k_n and γ_n are substantial. However, the issue of elastic spring constants and dissipation factors for the proteins is still in its infancy, and it is not clear whether those changes are dramatic or not.[46] The results of our model can be improved by using SEM measurements of the geometrical properties of AFM cantilevers.[35] In addition, higher electronics bandwidth of the AFM setup will allow including more resonances of the cantilevers and obtaining lower uncertainties of k_n and γ_n.

We expect our results to be transferable to other cantilevers with similar aspect ratio and buffers with similar ionic strengths, e.g., tris-buffered saline. Further studies are needed to account for corrections coming from van der Waals and electrostatic forces in the proximity of BOs. However, once an AFM cantilever is in contact with an arbitrary body, the forces acting in the contact zone typically surpass any non-contact interactions.[13,14]

In conclusion, we have developed an accurate model and a method to fit thermal resonances for compliant AFM cantilevers in biological media like PBS. Greater numbers of observed resonances will provide more precise values of mechanical signatures. Other upgrades need to account for the non-contact corrections for the resonance frequencies of AFM cantilevers in proximity to BOs. Mechanical signatures of proteins and cells can now be obtained by fitting the frequency shifts of flexural resonances of AFM cantilevers in contact with BOs. Our model can also be used to describe changes in the mechanical signature with time, e.g., to describe single protein folding trajectories under force.

II.2 A Qualitative Study of Conformational Changes in Peptides and Proteins

Here we provide an application of the model presented in the previous subchapter to differentiate between different conformations of a selected simple protein and a particular polypeptide. This application is based on measuring and discussing several resonance frequencies of an AFM cantilever when the cantilever is a) in air, b) in liquid, c) in liquid and in contact with either a given protein or a peptide sample and a control gold sample.

Materials and Methods

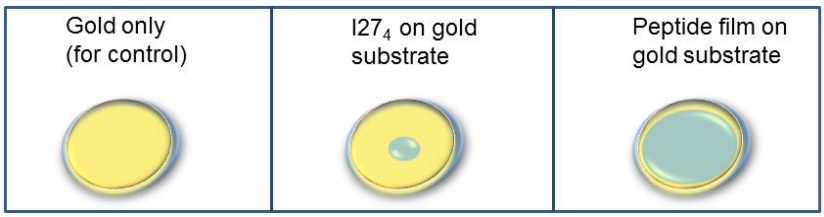

Figure 22. Types of samples used in experiments.

We use a simple protein, $I27_4$, as well as peptide samples in α- helix and β-sheet configuration deposited on a gold substrate, see Fig. 22. The gold substrates were prepared by evaporation on freshly cleaned glass slides. $I27_4$ protein was selected due to its known structure. Protein was expressed as described in Ref. 46. Each I27 module has 89 residues and there are 24 extra residues for linkers. These extra residues are 12 residues for a His-tag at N-terminus (M-R-G-S-H-H-H-H-H-H-G-S), two residues (R-S) for each linker between each recombinant domain in a construct, and four extra residues (R-S-C-C) at the C-terminus of the construct. The protein was affinity-purified on Talon cobalt columns (Clontech, USA) using a protein elution buffer (50mM sodium phosphate, 300mM sodium chloride, 250mM imidazole, pH 7.4) and then by FPLC size exclusion using S200 column (GE Healthcare).

Figure 23. The peptide amino acids sequence of the studied example of a peptide. [48]

The peptides used for experiments were characterized in Ref. 48. These peptides are in α-helical conformation in the case when the environment is 40% TFE and they would adopt a β-structure at pH 7.0 in 5 mM NaHCO3.[48] The peptide samples are prepared by dissolving lyophilized peptide in 100% ethanol. From this solution, 20μl are added to the gold coated sample and incubated for about 10 minutes. These molecules are absorbed on the gold substrate and can self-assemble in a α-helical conformation. To create a uniform film, the sample is washed with 100% ethanol to remove any excess. When the ethanol media is changed to DI water the peptides assemble in a β-sheet conformation.

We have used two types of rectangular AFM cantilevers: Olympus biolevers model BL-RC150VB, type "B," and Bruker MLCT type B. A set of experiments was performed in both air and in Dulbecco's PBS buffer (137mM NaCl, 3mM KCl, 2mM KH_2PO_4, and 8mM $Na_2HPO_4 \cdot 7H_2O$) from Midsci, USA. These measurements are used together with optical microscope measurements of the geometrical shape of the cantilever, see Appendix G.

Contact experiments were performed using force-clamp AFM force spectroscopy mode, for which we apply three different compressive forces: 30pN, 150pN and 300pN and then a tensile force of 130 pN, see Figure 24. In these experiments we recorded the fluctuations of the thermally – excited AFM cantilever at each point of the force ramp from Figure 24.

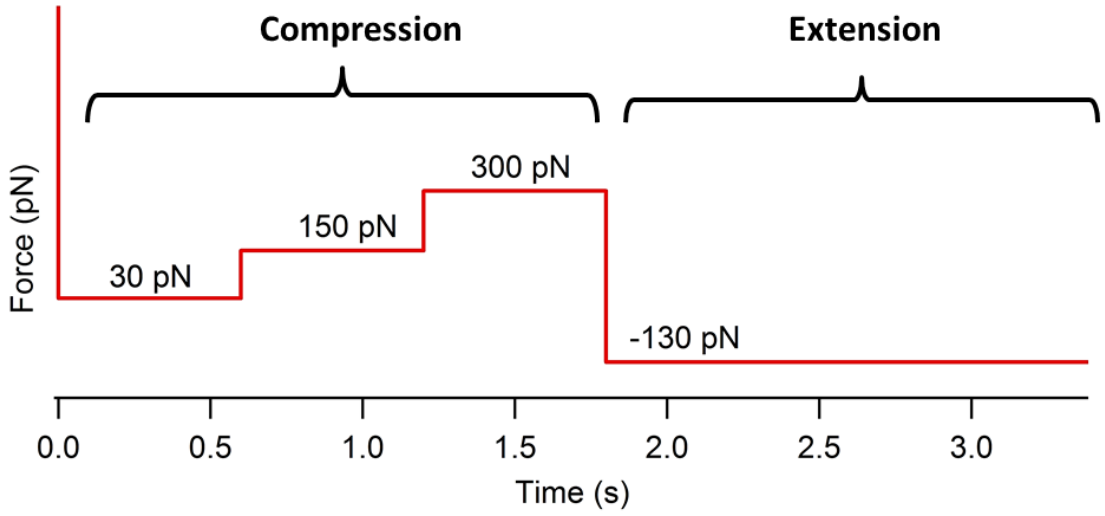

Figure 24. Force ramp used in the force-clamp experiments

Deflection signal of thermally vibrating AFM cantilevers is fast Fourier transformed to produce frequency spectra of the cantilevers, see Ref. 28. Resonance frequencies from the cantilever's frequency spectra are read using multipeak fit package with Voigt model in Igor Pro, Wavemetrics, USA. We fit the resonance frequencies using procedures written in Igor.

Results and Discussion

In order to assess the nanomechanical properties of the selected protein and polypeptide samples we have used the steps described in section II.1. First, we fit the resonance frequency spectra of the free cantilever in air to find the specific geometrical and mechanical parameters for a given cantilever, see Table 2 and Figure 17 in section II.1. Then, we fit the resonance frequency spectra of the free cantilever in PBS.

Below we plot the resonance frequency spectra of the free cantilever in PBS to find the hydrodynamic function away from the contact with the protein.

In Figure 25, we also provide the frequency spectra of the cantilever in contact or very close to the sample as obtained from the force-ramp in Figure 24.

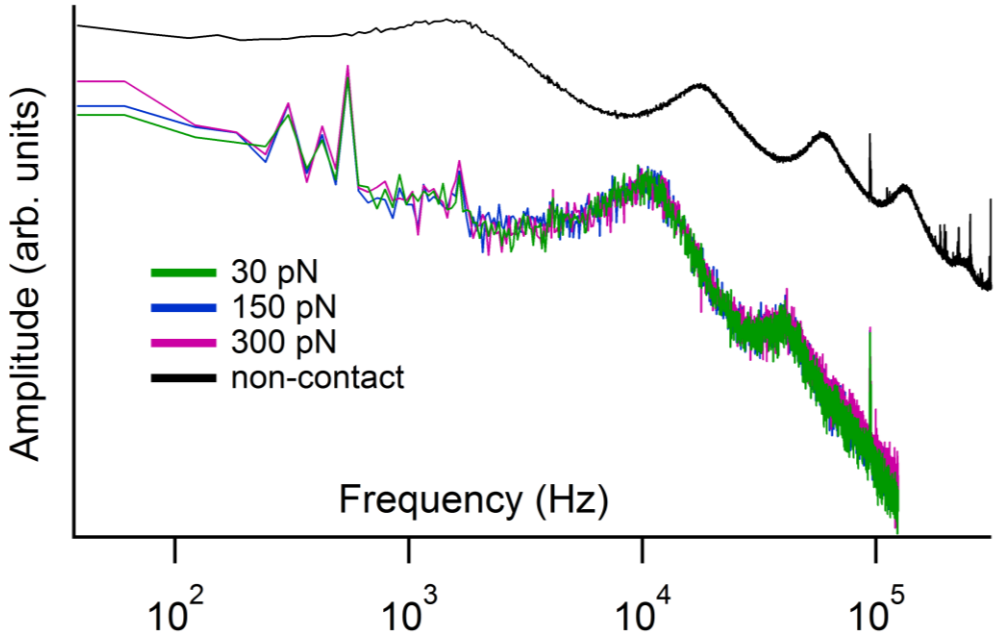

Figure 25. Plot of the frequency spectra at different applied forces on a gold sample in PBS buffer. The black curve represents the frequency spectra of the free cantilever and far away from the contact.

Using the steps described in II.1, we have fitted the free cantilever resonance frequency spectra in PBS using the generalized hydrodynamic function for a given cantilever, see Figure 20. Then, these parameters were used to fit the frequencies of the spectra in contact with the gold sample in order to determine the interaction parameters, k_n, k_{lat}, γ_n and γ_{lat} see section II.1 for definition of those parameters. We obtained that for a contact force of 300 pN, k_n = 0.6 N/m, k_{lat} = 12 N/m and $\gamma_n = \gamma_{lat} = 10^{-8}$ kg/s. These values are then cross-checked with the results of the contact mechanics describing the contact stiffness using specific models. For example, using Hertz model, we can estimate the values of k_n [24]

$$k_n = \sqrt[3]{6FRE^{*2}}$$

Where F is the applied force, R is the contact radius and E^* is the reduced Young modulus. Next, we can relate k_n to k_{lat} by[24]:

$$k_{lat} = 0.9k_n$$

Since the contact radius, R, is not known, we calculated the values of k_n and k_{lat} for several values of the contact radii expected for the used cantilevers, see Table 3.

F [N]	R [m]	E* [Pa]	k_n [N/m]	k_{lat} [N/m]
3.00E-10	2.00E-08	6.19E+10	51.66	46.49
3.00E-10	2.50E-08	6.19E+10	55.65	50.08
3.00E-10	3.00E-08	6.19E+10	59.13	53.22
3.00E-10	3.50E-08	6.19E+10	62.25	56.02

Table 3. Calculated values of k_n and k_{lat} using Hertz model.

Comparing the values of k_n and k_{lat} between our measurements and estimations, we can see that the values found from the fitted frequency spectra are very different than expected. A few possible explanation for this mismatch could be related to the bending of the cantilever tip, which is not accounted for within our model, or different than modeled behavior of the hydrodynamic function.

It has been observed that as the AFM cantilever approaches the surface of the sample the resonance frequencies shift and become more damped. [49] This is a result of adhesive interactions with the surface, but also a change in the hydrodynamic function of distance from the sample. In order to qualitatively estimate those effects, we present in Figure 26 thermally excited resonance frequencies of an Olympus cantilever at different distances from the surface.

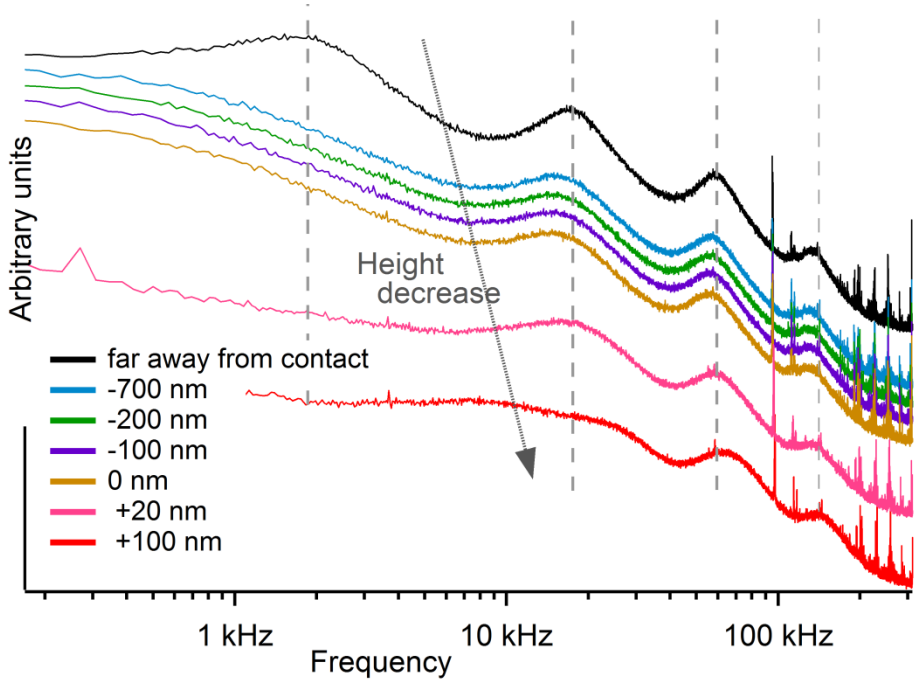

Figure 26. Plot of the frequency spectra at different heights from a gold sample in PBS buffer.

Due to rather marginal adhesive interactions between the AFM tip and the protein sample within PBS buffer, we tend to speculate that Figure 26 shows that when the cantilever is in the proximity of a sample, there are some changes in the hydrodynamic function; therefore, this should be dependent on the distance from the sample. Hydrodynamic forces depend on the hydrodynamic boundary condition therefore they depend on the mechanical properties of a liquid at a solid surface. In fluid mechanics it is generally assumed that the liquid molecules adjacent to a solid surface are stationary relative to the solid and that the viscosity is equal to the bulk viscosity, even when the liquid flows over the surface. This is a good assumption for macroscopic systems, but at nanometer scale it is not true and the viscosity of liquids can increase many orders of magnitude.

A possible solution would be to calculate the semi-empirical hydrodynamic function close to the sample in similar way as described in II.1. Unfortunately, when the AFM cantilever is close to the sample, the resonance frequencies shift towards lower frequencies and the first

resonance is significantly damped. This prevents us from accurately determining the first resonance frequency, see Figure 27. Without a good estimate of the first resonance frequency value, we are unable to calculate the hydrodynamic function.

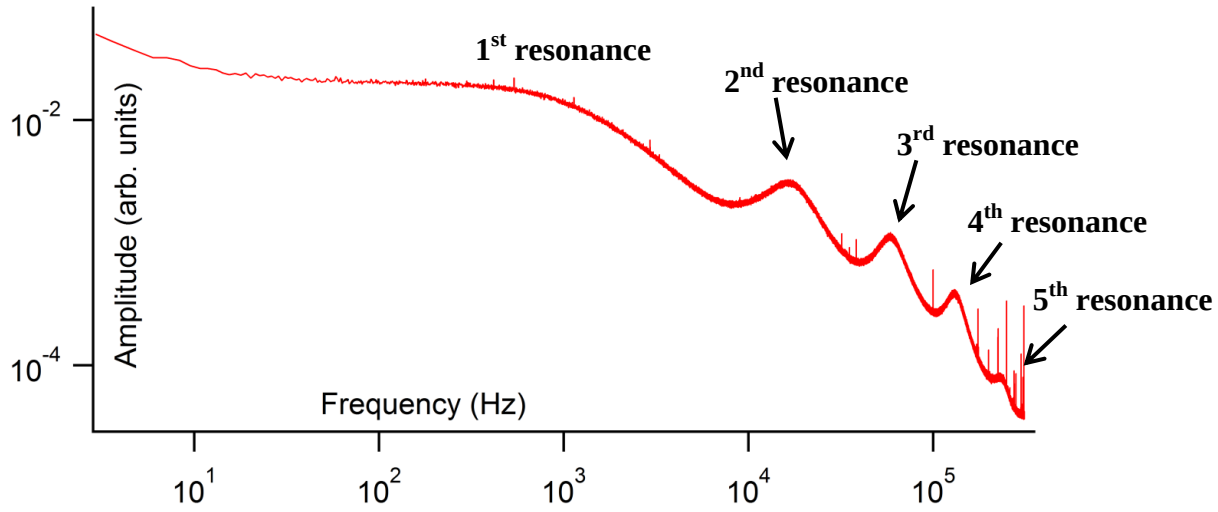

Figure 27. A typical thermal amplitude spectrum of flexural resonances of an Olympus cantilever in the PBS buffer close to contact.

Besides the hydrodynamic forces, there are other effects in the close proximity of the sample. Significant effects can be caused by the electrostatic double layer force, the hydration repulsion and the hydrophobic attraction.

Electrostatic double layer force it is caused by the surface charges at interfaces. Since we are performing experiments in a aquous environment, there are surface charges balanced by dissolved counterions, which are attracted back to the surface by the electric field, but spread away from the surface to increase the entropy. Together, the ions and charged surface are known as the electric double layer. When the cantilever approaches, there will be two surface approaching which perturbes the double layer resulting in a force known as the double-layer force. When the approaching surface charges have the same sign, the concentration of ions between the surfaces always increases. This results in a repulsive force. At large distances, this

electrostatic double-layer force decays roughly exponentially. Other effects might relate to hydration and hydrophobic forces. Hydration forces are a type of solvation forces which are acting between polar surfaces separated by a thin layer (<3 nm) of water (or some other polar solvents), which decays quasi-exponentially with the distance from the surface. Hydration forces are repulsive and, tend to be stronger and longer ranged with increasing salt concentration. We are using small concentration of salts, of the order of 0.1 M, so these effects are not very significant. Hydrophobic surfaces in water attract each other. It has been shown that the interaction between solid hydrophobic surfaces is stronger than the van der Waals attraction.

Also, we can use different models for the contact mechanics for better estimates of contact stifness. For soft samples, like the proteins, interacting with much more stiff AFM cantilever, the literature results would often apply either Hertz or Sneddon contact mechanics. The model derived by Sneddon assumes a rigid cone indenting a soft flat surface with a force:

$$F_{Sneddon} = \frac{2}{\pi}\frac{E}{1-\upsilon^2}\tan(\alpha)\delta^2$$

where E is the surface Young's modulus, υ is the surface Poisson's ratio, α is the indenter cone angle and δ is the indentation. Both Hertz and Sneddon models do not include adhesion and visco-elasticity. Hertz model is valid for small indentations, while in the case of Sneddon model the indentation is considered to be large such that the cone apex can be approximated as infinitely sharp.

For a moment, however, likely due to problems with hydrodynamic functions, we lack a quantitative method to translate the shifts in resonance frequencies into changes in molecular stiffness and mechanical energy dissipation factors. Thus, in the next part, we present a qualitative picture of these changes.

First, we look at variations of resonance frequency in different media for a given cantilever, see Table 4. In the case of liquid environment we present three repeats for the same conditions, PBS buffer and 70% ethanol solution.

Media	f1 [kHz]	f2 [kHz]	f3 [kHz]	f4 [kHz]	f5 [kHz]
AIR	13.6 ± 0.01	77.9 ± 0.04	84 ± 0.01	240 ± 0.01	-
PBS	1.42 ± 0.01	17.4 ± 0.02	58.9 ± 0.03	131.7 ± 0.1	233 ± 3.98
PBS	1.38 ± 0.01	17.4 ± 0.02	58.9 ± 0.04	131.9 ± 0.1	234.8 ± 0.87
PBS	1.43 ± 0.01	17.4 ± 0.02	59 ± 0.04	132.0 ± 0.1	233 ± 1.67
ETHANOL	1.41 ± 0.02	17.7 ± 0.03	58.3± 0.06	128.2 ± 0.1	219 ± 1.20
ETHANOL	1.34 ± 0.02	17.7 ± 0.03	58.2 ± 0.05	127.6 ± 0.1	218 ± 1.17
ETHANOL	1.44 ± 0.02	19.6 ± 0.03	57.3 ± 0.19	133.0 ± 0.1	-

Table 4. Values of resonance frequencies for a free Olympus cantilever in different media.

Observing the results within the Table 4, their errors, and variations of the results between the trials within the same media, one can notice very marginal frequency shifts between different media for the first four resonances (f1 to f4), but some changes of the fith resonance frequency (f5) between PBS and ethanol. Since the differences are not significant for the first four resonances, we can approximate the hydrodynamic function to be similar for the two liquid media. However, in the case that different liquid environment is used, the hydrodynamic function should be changed accordingly.

Next, we check the shift produced in the resonance frequencies when compared with the same cantilever in contact with a $I27_4$ molecule, see Figures 28 and 29.

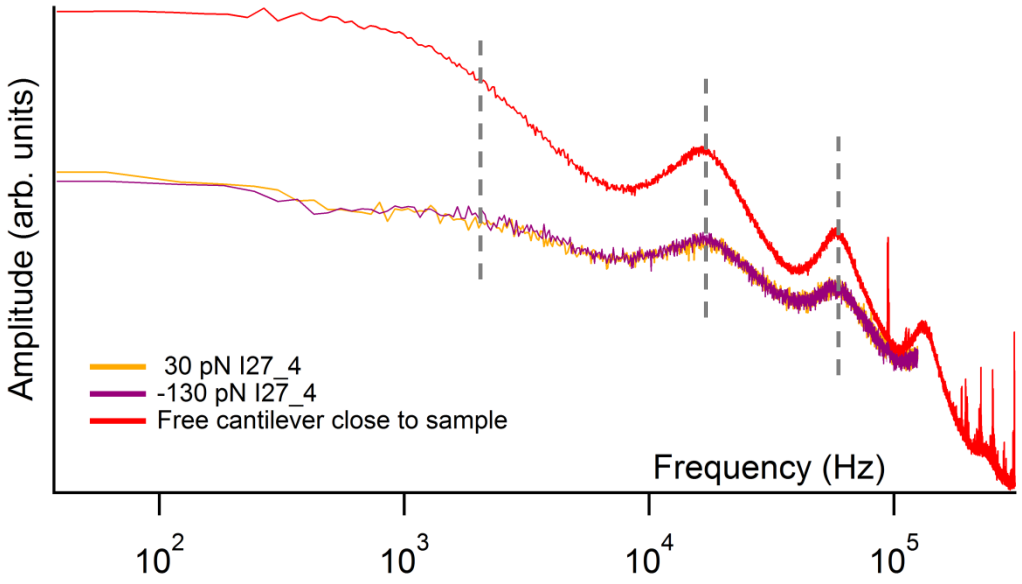

Figure 28. Plot of the frequency spectra at different applied forces on an $I27_4$ sample in PBS. The red curve represents the frequency spectra of the free cantilever close to contact in PBS used for comparison and a control.

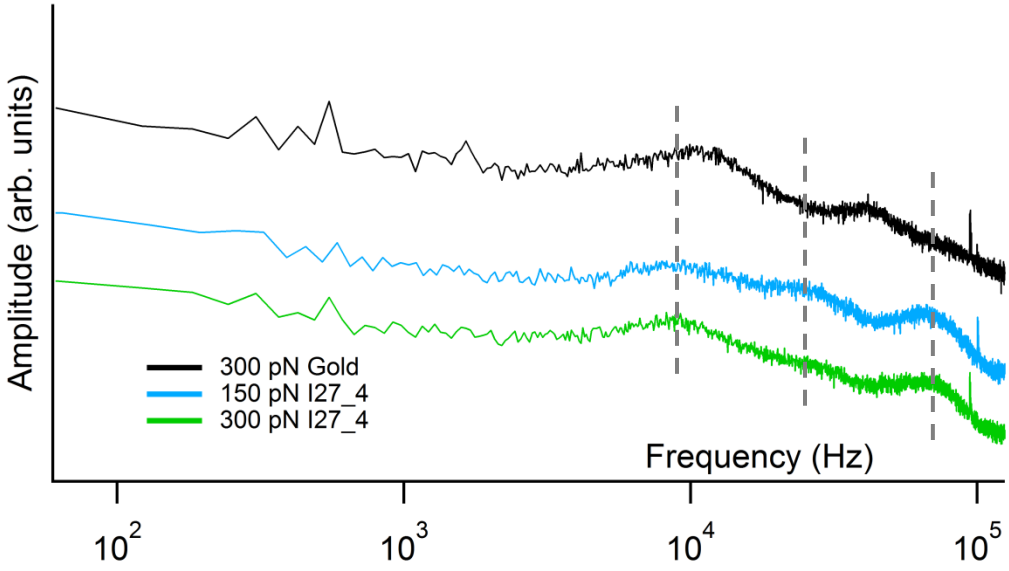

Figure 29. Plot of the frequency spectra at different compressive forces on an $I27_4$ sample in PBS. The black curve represents the frequency spectra at 300 pN contact force on gold substrate in PBS used for comparison as control.

There is a shift in the frequency spectra produced by the contact with protein compared to the frequency spectra in the case of contact with the gold sample.

Sample	f1 [kHz]	f2 [kHz]	f3 [kHz]
30 pN Gold	9.88	39.9	
150 pN Gold	10.1	39.5	
300 pN Gold	10.2	39.9	
-130 pN Gold	9.94	39.2	
30 pN $I27_4$ (unclear)	0.5?	16.4	56.9
150 pN $I27_4$	8.38	17.95	59.5
300 pN $I27_4$	8.52	?	62.4
-130 pN $I27_4$	0.92	16.4	56.7

Table 5. Comparison of resonance frequencies values of a cantilever in contact with gold versus in contact with an $I27_4$ sample in PBS buffer.

Observing the results in Table 5, we can see that resonance frequencies shift towards lower values in contact with protein+gold than in the case of contact with gold only. This is expected, since a protein makes the tip-surface contact more compliant.

We are unable to fit the model described in part II.1 because it is difficult to determine the first resonance frequency in some of the spectra but we can estimate the damping factors, from the width of the resonances. The resonance frequencies widths are similar, therefore, the differences in damping factors are not distinguishable between contact with gold+protein compared to contact with gold only.

In the next part, we repeated similar measurements as described for I27, but using peptide samples. First, we measured the control on gold sample in ethanol, see Figure 30.

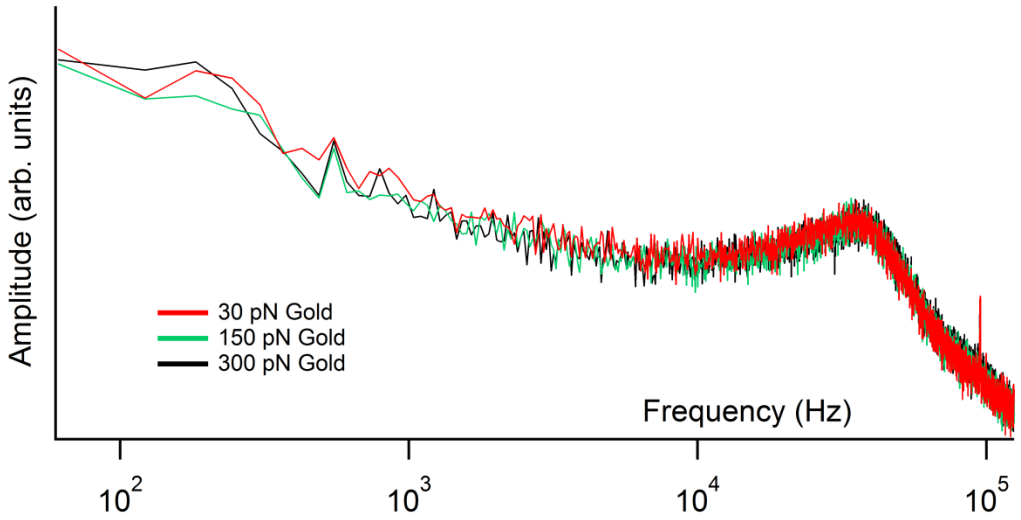

Figure 30. Plot of the frequency spectra at different compressive forces on a gold sample in ethanol.

There is a shift in the frequency spectra produced by the contact with peptide sample compared to the frequency spectra in the case of contact with the gold sample which can be observed from the graphs in figures 31 and 32.

We expect that the peptide in ethanol is in α-helix conformation.[48] Thus, as expected the resonances of the cantilever in contact with the peptide in α-helix conformation shift a bit towards lower values, which shows that the sample has become more compliant than the gold substrate only.

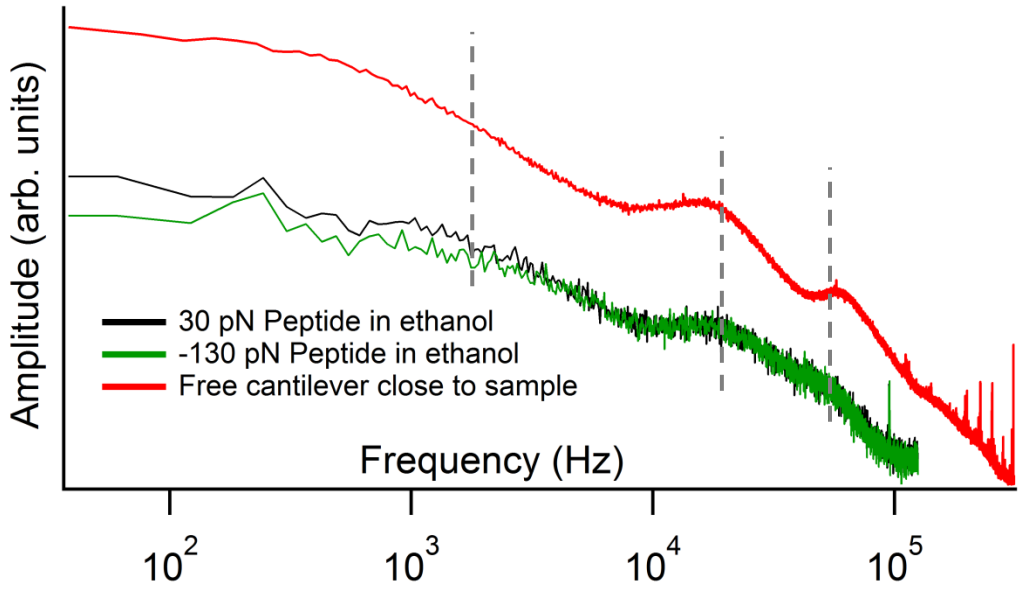

Figure 31. Plot of the frequency spectra at different applied forces on peptide sample in ethanol. The red curve represents the frequency spectra of the free cantilever close to contact in ethanol used for comparison as control.

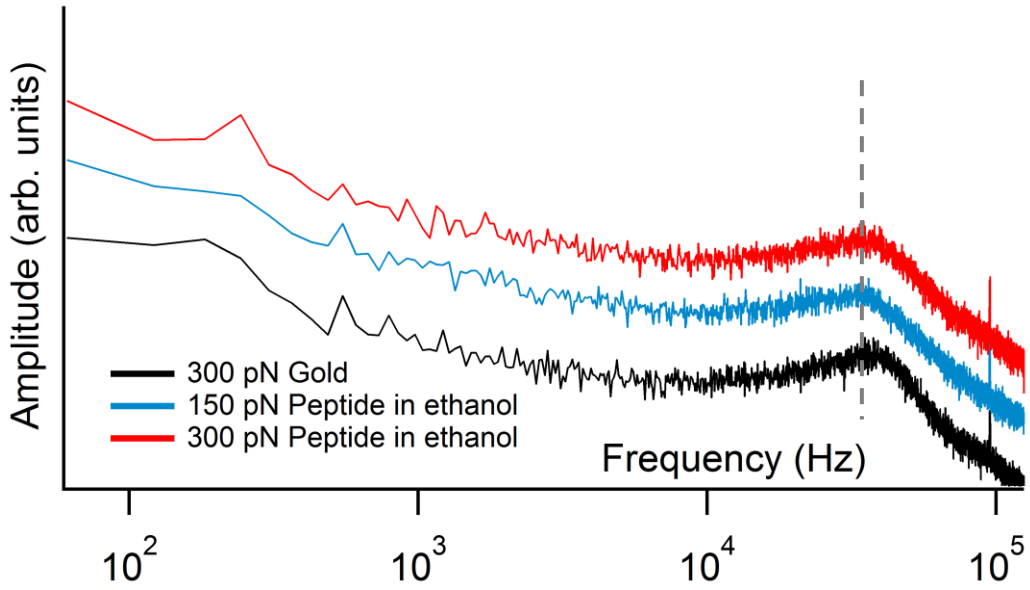

Figure 32. Plot of the frequency spectra at different compressive forces on peptide sample in ethanol. The black curve represents the frequency spectra at 300 pN contact force on gold substrate used for comparison as control in ethanol.

Next, we changed the media with water to change the conformation of the peptide sample into β-sheet and checked if this conformation would produce different results. Below are the results of peptide sample in β-sheet conformation.

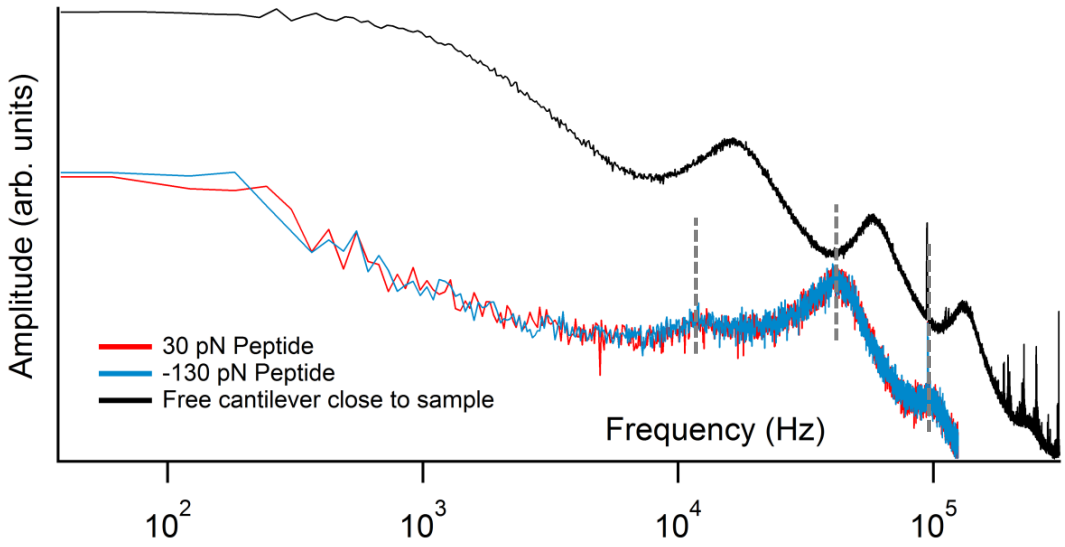

Figure 33 Plot of the frequency spectra at different applied forces on peptide sample in water. The black curve represents the frequency spectra of the free cantilever close to contact in water used for comparison as control.

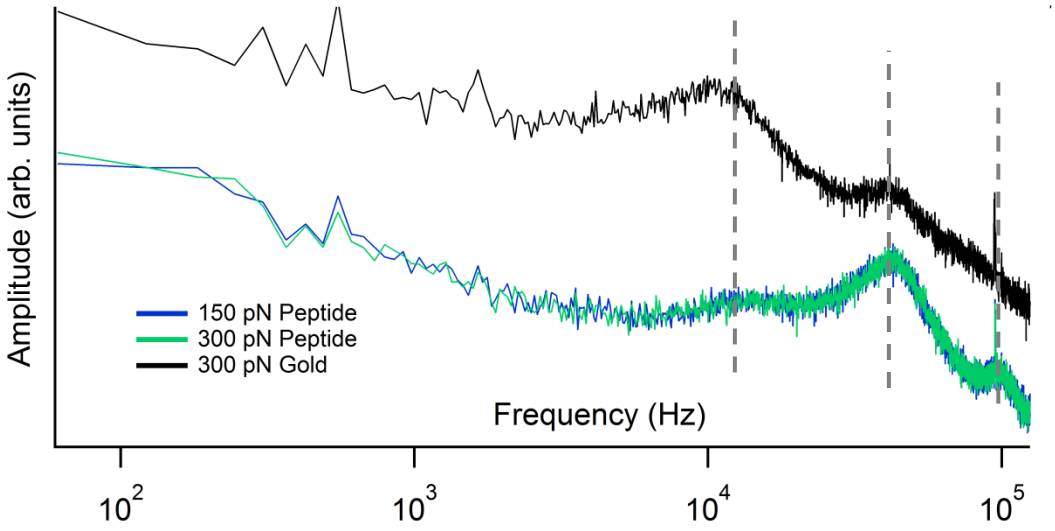

Figure 34. Plot of the frequency spectra at different compressive forces on peptide sample in water. The black curve represents the frequency spectra at 300 pN contact force on gold substrate used for comparison as control in water.

When comparing the frequency spectra for the two conformations of peptide sample, we can estimate that the α-helix conformation is more compliant than the β-sheet conformation.

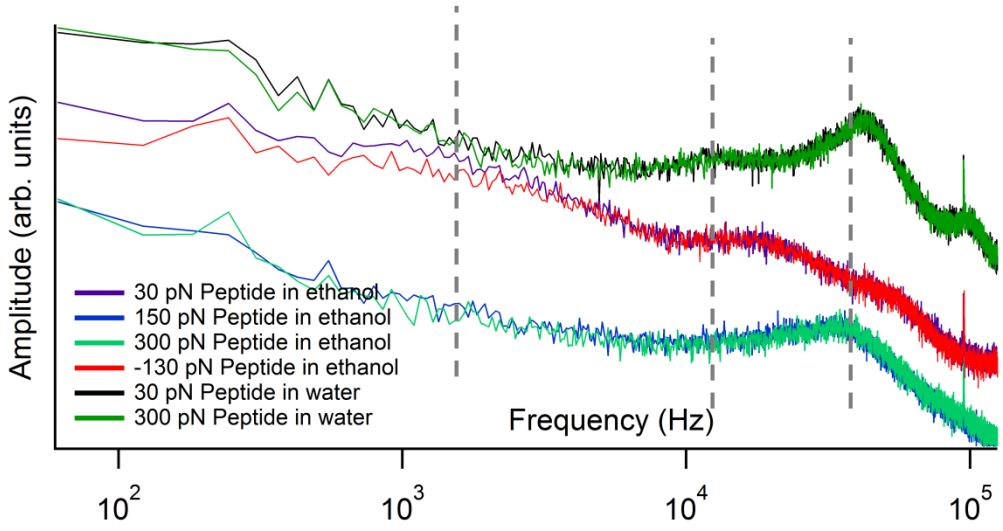

Figure 35. Plot of the comparison of frequency spectra at different applied forces on peptide sample in ethanol and in water.

These results were checked with another set of experiments using a different type of cantilever and we obtained similar results, indicating that we are able to detect different peptide conformations.

II.3 Future work for nanomechanical properties

In the future steps towards understanding the secondary protein structure folding we would need to combine the experimental data, computer simulations and visco-elastic models. First, we need to have an accurate model for the cantilever in contact with these samples to be able to de-convolve the nanomechanical properties of the samples from the substrate. We can then study a large variety of peptide secondary structures folding using mechanical signatures and create a data base of mechanical signatures presented by different peptides. Once we have a data basis of mechanical signatures we can conduct experimental measurements of mechanical

signatures at various rates of force quenching, pH, and ionic strengths. In the end, we need to perform extensive computer simulations of the experimentally studied systems and correlate those with experimentally measured mechanical signatures and use visco-elastic models to interpret the results.

II.4 References

1. C. Bustamante, Y. Chemla, N. Forde, and D. Izhaky, *Annu. Rev. Biochem.*, **2004**, 73, *705*

2. M. J. Buehler and T. Ackbarow, Comput. *Methods Biomech. Biomed. Eng.*, **2008**, 11, *595*

3. A. Sorkin and M. von Zastrow, *Nat. Rev. Mol. Cell Biol*., **2009**, 10, *609*

4. J. Arnadottir and M. Chalfie, *Annu. Rev. Biophys.*, **2010**, 39, *111*

5. L. Han, A. J. Grodzinsky, and C. Ortiz, *Annu. Rev. Mater. Res.*, **2011**, 41, *133–168*

6. N. B. Becker and R. Everaers, *J. Chem. Phys*., **2009**, 130, *135102-10*

7. P. V. Cornish and T. Ha, *ACS Chem. Biol*., **2007**, 2, *53–61*

8. J. M. Fernandez and H. B. Li, *Science*, **2004**, 303, *1674*

9. C. Cecconi, E. Shank, C. Bustamante, and S. Marqusee, *Science*, **2005**, 309, *2057*.

10. W. J. Greenleaf, M. T. Woodside, and S. M. Block, *Annu. Rev. Biophys. Biomol. Struct.*, **2007**, 36, *171*

11. G. Zoldak and M. Rief, *Curr. Opin. Struct. Biol.*, **2013**, 23, *48*

12. T. Ando, T. Uchihashi, and N. Kodera, *Annu. Rev. Biophys.*, **2013**, 42, *393*

13. B. Cappella and G. Dietler, *Surf. Sci. Rep.*, **1999**, 34, *1*

14. H.-J. Butt, B. Cappella, and M. Kappl, *Surf. Sci. Rep.*, **2005**, 59, *1*

15. R. Szoszkiewicz and E. Riedo, *in Applied Scanning Probe Methods V, edited by B. Bhushan, H. Fuchs, and S. Kawata* (Springer-Verlag, Heidelberg, **2007**), *pp. 269–286.*

16. J. Adamcik, C. Lara, I. Usov, J. Jeong, F. S. Ruggeri, G. Dietler, H. Lashuel, I. Hamley, and R. Mezzenga, *Nanoscale*, **2012**, 4, *4426*

17. R. Szoszkiewicz, A. Kulik, G. Gremaud, and M. Lekka, *Appl. Phys. Lett.*, **2005**, 86, *123901*

18. R. Garcia and E. Herruzo, *Nat. Nanotechnol*., **2012**, 7, *217*

19. J. Lozano and R. Garcia, *Phys. Rev. Lett.*, **2008**, 100, *076102*

20. A. Raman, S. Trigueros, A. Cartagena, A. Stevenson, M. Susilo, E. Nauman, and S. Antoranz-Contera, *Nat. Nanotechnol*., **2011**, 6, *809*

21. O. Sahin, C. Quate, O. Solgaard, and A. Atalar, *Nat. Nanotechnol.*, **2007**, 2, *507*

22. D. Kiracofe and A. Raman, *Phys. Rev. B*, **2012**, 86, *205405*

23. R. Szoszkiewicz, *Rev. Sci. Instrum*., **2012**, 83, *037101*

24. E. Dupas, G. Gremaud, A. Kulig, and J.-L. Loubet, *Rev. Sci. Instrum.*,**2001**, 72, *3891*

25. J. Sader, J. Sanelli, B. Adamson, J. Monty, X. Wei, S. Crawford, J. Friend, I. Marusic, P. Mulvaney, and E. Bieske, *Rev. Sci. Instrum.*, **2012**, 83, *103705*

26. J. Chon, P. Mulvaney, and J. Sader, *J. Appl. Phys.*, **2000**, 87, *3978*

27. J. Sader, *J. Appl. Phys.*, **1998**, 84, *64*

28. A. Dey and R. Szoszkiewicz, *Nanotechnology*, **2012**, 23, *175101*

29. The errors are calculated using a formula: $\sum_i \left| f^i_{measured} - f^i_{fitted} \right| / f^i_{measured}$ where $f^i_{measured}$ and f^i_{fitted} are the *i*-th measured and fitted frequencies, respectively.

30. M. Carrion-Vazquez, A. Oberhauser, T. Fisher, P. Marszalek, H. Li, and J. Fernandez, *Prog. Biophys. Mol. Biol.*, **2000**, *74*, 63

31. H. Dietz, F. Berkemeier, M. Bertz, and M. Rief, *Proc. Natl. Acad. Sci. U.S.A.*, **2006**, *103*, 12724

32. A torsional spring constant k_{tor} with its proper dissipation factor γ_{tor} can be used in addition or instead of k_{lat} and γ_{lat}.

33. A. Khan, J. Philip, and P. Hess, *J. Appl. Phys.*, **2004**, *95*, 1667

34. We calculate the Young's modulus using a formula for a two-layer composite beam—Ref. 27 in the paper of Gavan et al.[36]—comprised of 50 nm gold and silicon nitride. Thickness of the silicon nitride itself is estimated within 114 nm to 186 nm.[35] Error in the Young modulus is obtained using the results of of Gavan et al.,[36] who measured Young's moduli of thin silicon nitride films.

35. See supplementary material at http://dx.doi.org/10.1063/1.4858411 for optical and scanning electron microscopy images of AFM cantilevers, application of our model to proteins, and error propagation analysis.

36. K. Gavan, H. Westra, E. van der Drift, W. Venstra, and H. van der Zant, *Appl. Phys. Lett.*, **2009**, *94*, 233108

37. Frequency of the 1st torsional resonance is calculated from $f_{tor} = 0.5 \frac{t}{Lb} \left(\frac{E}{\rho(2+2\upsilon)} \right)^{1/2}$, where the Poisson $\upsilon = 0.2$ for a SiNx cantilever is obtained from Ref. 33.

38. Half a pyramid with a square base was used, with a side a = b/4 to yield the value of m_{tip}

$$m_{tip} = \rho \frac{t}{6} \left[(b + 4t)\left(h_{tip} + 2t\right) + \frac{b^2}{8} \right]$$

39. Displacements of the support spring of up to several millimeters were correlated with dynamometer's measurements of forces.
40. The quality factors are estimated from a ratio of the amplitudes on the resonance and at the arbitrarily chosen low frequency.
41. The PBS density of 998 kg/m3 was measured in Ref. 47 at temperature of about 22^{o}C.
42. J. Taylor, *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements*, 2nd ed. (University Science Books, **1996**).
43. t-Student test is used to test H_0: $a_2 = 0$. To do so, a calculated t-Student coefficient for a_2 is compared with its tabulated value for a given number of degrees of freedom and at a 99% confidence level. From the data in Figure 19, we get a value of $a_2 = 0.059$ and its standard deviation $s_{a2} = 0.011$.Thus, the calculated t-Student coefficient is $t(a_2) = a_2=s_{a2} = 5.4$. This value is larger than a tabulated value t(17;0.01) = 2.9 read from the statistical tables for 17^{o} of freedom and at 99% confidence level (P. Bevington, Data Reduction and Error Analysis for the Physical Sciences (McGraw-Hill Book Company, New York, 1969). Thus, H_0 is not accepted.
44. Y. Wang and G. Zocchi, *EPL*, **2011**, *96*, 18003
45. Y. Taniguchi, B. S. Khatri, D. J. Brockwell, E. Paci, and M. Kawakami, *Biophys. J.*, **2010**, 99, 257
46. K. E. Malek and R. Szoszkiewicz, "Changes of protein stiffness during folding detect protein folding intermediates," J. Biol. Phys. 2014.
47. J. Schiel and D. Hage, Talanta, **2005**, 65, 495
48. Sushanth Gudlur, Pinakin Sukthankar, Jian Gao, L Adriana Avila, Yasuaki Hiromasa, Jianhan Chen, Takeo Iwamoto, John M Tomich, *PLoS One*, **2012** *18*;7(9):e45374
49. Tung, R.C., Jana, A., and Raman, A., *Journal of Applied Physics*, **2008**, *104(11)*, 8

Appendix A - Complete SDS-PAGE results of the $I27_2$ – NRR – $I27_2$ protein

Different proteins have different net charges. Therefore, when placed in an electric field, positively charged proteins will migrate towards the negatively charged pole, while negatively charged proteins migrate to the positively charged pole and when the protein is present at a pH equal to its isoelectric point, it will generally not move.

Based on this principle, proteins are separated on gels made of polyacrylamide. One method is using Sodium Dodecylsulfate (SDS) electrophoresis.

SDS is a detergent that contains a long aliphatic chain and a sulfate group which interacts with denatured proteins to form a strongly negatively charged complex. The SDS-protein complexes all contain about the same negative charge and therefore the proteins all separate from each other strictly on the basis of their sizes.

SDS-PAGE: 1 mm thick denaturing SDS-PAGE gels were prepared and run according to a standard Laemli method with 9% separating gel solution and 3.9% stacking gel solution (1). We used about 20 μg of a protein per well, and we read off the bands using Perfect Protein Marker, 15-150 kDa, No.80030-960, from Novagen. We run the gels for 2 h at a constant voltage of 150 – 200 V in a standard SDS electrophoresis buffer. The gels were stained/destained using Coomassie Blue and 10 % acetic acid, similarly as in the Ref. (2). Destained gels were scanned using Epson Perfection V300 scanner.

Investigating the gel, and particularly intensity of the fractions corresponding to the $I27_2$ – NRR – $I27_2$ protein (lanes 7 and 8), we estimate the protein's purity at > 90%.

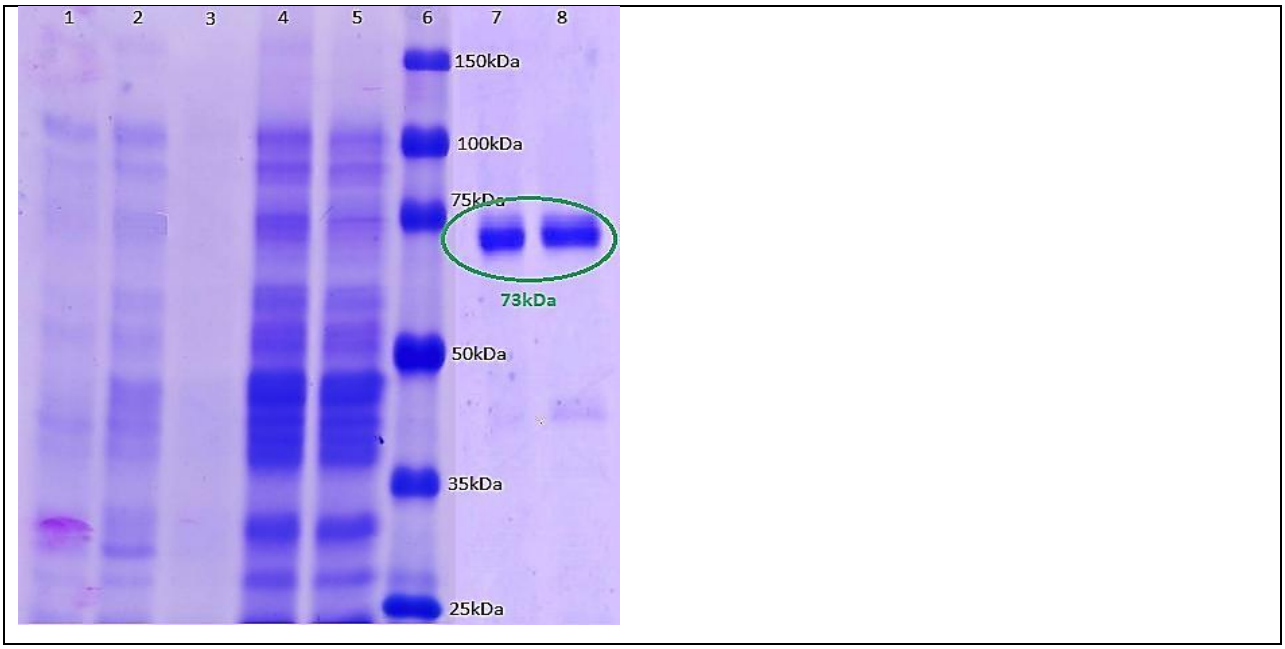

Fig. A1. *The SDS-PAGE gel for the I272 – NRR – $I27_2$. The lanes are: 1 – non induced (cells) control; 2 – induced (cells) control. IPTG was used for induction; 3 – supernatant after centrifuge (no protein, protein in pellet); 4 – supernatant after sonication and centrifuging (should contain protein); 5 – bacteria cells from the pellet after sonication and centrifuging (should not contain protein under optimized conditions, and it does not; furthermore, no inclusion bodies have been found in that pellet); 6 – protein marker (Perfect Protein Marker, 15-150 kDa, No.80030-960, from Novagen); 7 – fraction (elution "1") containing the $I27_2$ – NRR – $I27_2$ after the gravity column and without calcium ions; 8 – same as lane 7, but with 1 mM $CaCl_2$. The protein has been denatured, so addition of calcium (responsible for protein folding) should not influence the observed size of the $I27_2$ – NRR – $I27_2$, as we have found out to be the case.*

References:

1.Ausubel, F.M., Brent R., Kingston R.E., Moore D.D., Seidman J.G., Smith J.A., Struhl K. 2002. Short Protocols in Molecular Biology. 5th ed.Wiley,

2.Dong, W.H., Wang T.Y., Wang F., Zhang J.H. 2011. Simple, time-saving dye staining of proteins for sodium dodecyl sulfate-polyacrylamide gel electrophoresis using coomassie blue. PLoS One. 6(8):e22394. PMCID: 3151240.

Appendix B - Worm-Like Chain Model

A method to study the nanomechanical properties of a single biomolecule is by force-induced unfolding. For this mode we analyze the mean unfolding forces and the unfolding pattern for a given protein.

Typically, we explain the mechanical properties of elastic objects using Hooke's Law:

$$F = -kx$$

We expect that when we apply a force on some elastic object, it will increase in length linearly with the proportionality constant k. But in the case of polymer chains, the force is related to a fractional increase (x/L).

A model that can be used to estimate the mechanical properties of single polymer molecules under an applied force is the Worm-Like Chain model (WLC). This model can be applied with the assumption that the polymer is an elastic cylinder with a constant bending elasticity and of constant length.[1]

The force required to stretch a WLC-modeled polymer is given by formula in Figure B1.

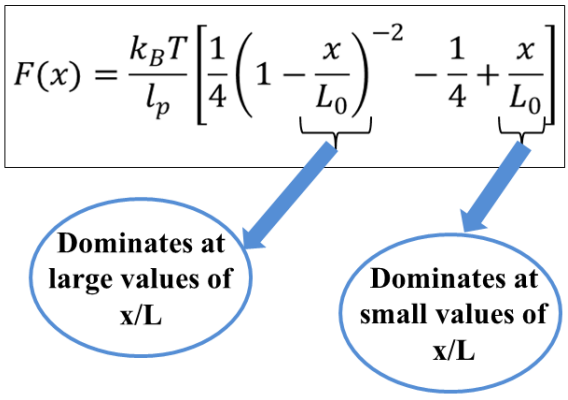

Figure B1. Force required to stretch a WLC-modeled polymer

Where the persistence length lp is the length of the unit block.

L0 is the full length of the unfolded polymer and x is the unfolded length of polymer.

The effect of lp is illustrated in figure B2.

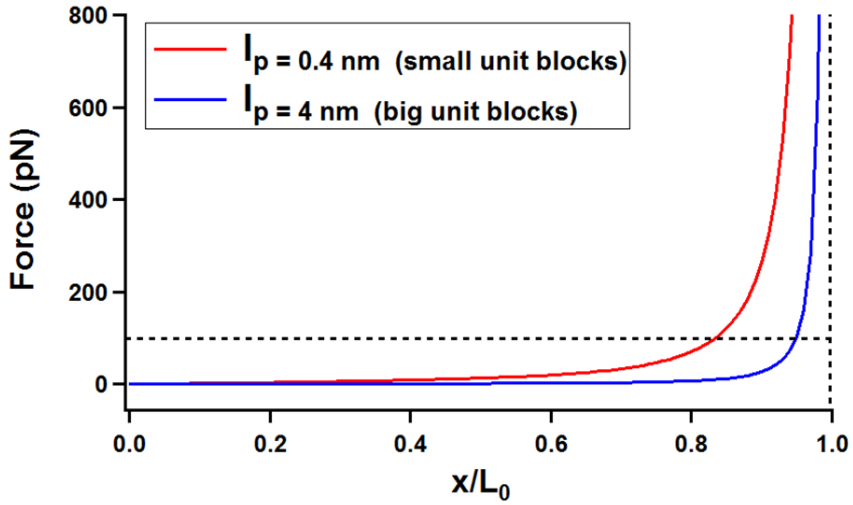

Figure B2. The effect of l_p

References:

1. Yuko Hori, Ashok Prasad, and Jané Kondev. *Phys. Rev. E*, **2007**, *75*, 041904

Appendix C - AFM imaging of the $I27_2$ – NRR – $I27_2$ protein adsorbed on gold substrate.

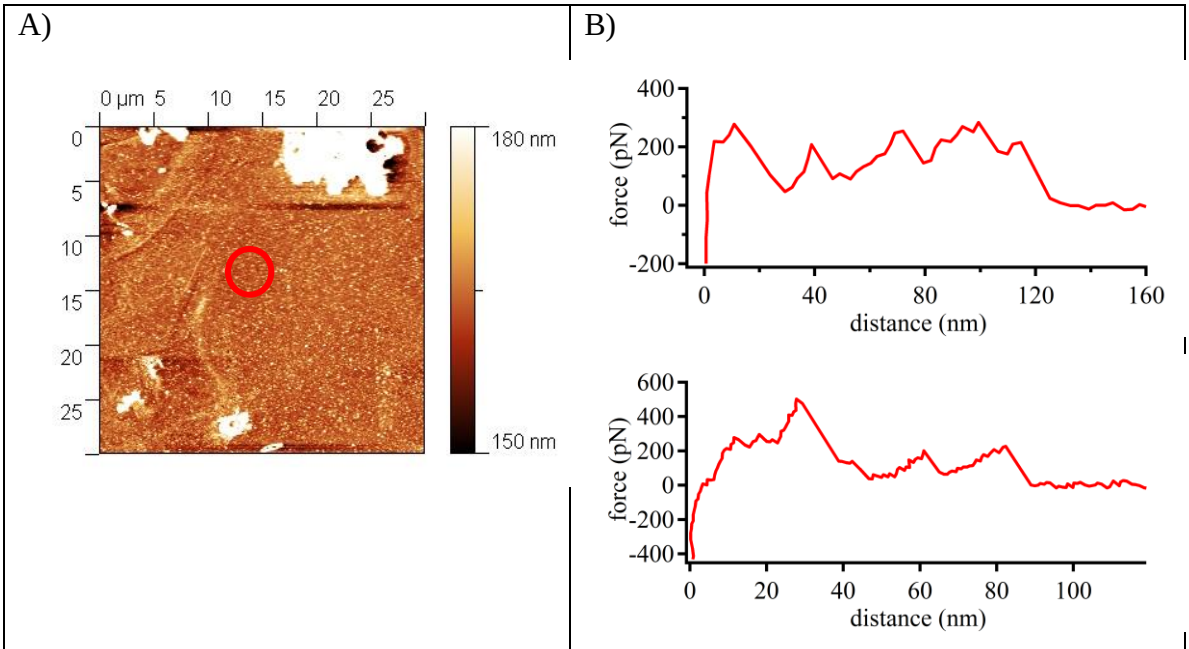

Fig. C1. (A) *30 µm x 30 µm topography scan (in water) of a gold surface with adsorbed proteins. There is potentially quite many single $I27_2$ – NRR – $I27_2$ proteins adsorbed there, as verified by a low quality f-d curves presented in* ***(B)*** *and collected within an area denoted by a red circle in (A).*

Appendix D - Unfolding probabilities

The unfolding probability for given extended sequences "S" of events. There are four classes of events, which can be arranged in 24 different ways (number of their permutations). However, in our FX traces we found very few complete sequences of any sort. Thus, we extend our calculations to include two-element and three-element chunks for each given sequence. For example, let us take one particular sequence (called "s1") and its extended components yielding the so-called "extended sequence", or "s1_ext". These results are presented in Table D1.

S1	C1->C2->C3->C4
S1_ext	C1->C2->C3
	C1->C2
	C2->C3->C4
	C2->C3
	C3->C4

***Table D1**. An example of an unfolding sequence (s1), and its extended sequence (s1_ext).*

We determine the probability of each extended sequence as: P(s"i"_ext) = (total # of traces where "s"i"_ext" is present) / (total number of traces considered), with i = 1, 2, ….24. We did not consider here the traces with only one NRR unfolding event present. The results are presented in Table D2. We obtain that the sequence "s1_ext" is the most probable, with P(s1_ext) = 0.23 ± 0.02.

S = sequence		No. of traces within S	P(S)
S1 extended	1->2->3->4	20	0.16
S2 extended	1->3->2->4	4	0.03
S3 extended	1->2->4->3	8	0.06
S4 extended	3->1->2->4	8	0.06
S5 extended	2->1->3->4	5	0.04
S6 extended	2->3->1->4	7	0.05
S7 extended	4->3->2->1	0	0.00
S8 extended	4->1->2->3	17	0.13
S9 extended	4->2->1->3	4	0.03

	S10 extended	4->3->1->2	8	0.06	
	S11 extended	4->2->3->1	6	0.05	
	S12 extended	4->1->3->2	4	0.03	
	S13 extended	3->4->1->2	8	0.06	
	S14 extended	3->2->1->4	1	0.01	
	S15 extended	3->4->2->1	0	0.00	
	S16 extended	3->1->4->2	1	0.01	
	S17 extended	3->2->4->1	0	0.00	
	S18 extended	2->4->3->1	0	0.00	
	S19 extended	2->3->4->1	6	0.05	
	S20 extended	2->1->3->4	5	0.04	
	S21 extended	2->4->1->3	4	0.03	
	S22 extended	1->4->3->2	1	0.01	
	S23 extended	1->4->2->3	7	0.05	
	S24 extended	1->3->4->2	5	0.04	

Table D2. *Probabilities of the extended unfolding sequences present in the traces.*

Appendix E - Igor macro for FFT transform

```
Function FFT_FC_SingleFragments()                    //FFT_FC_fragments(ctrlName) : ButtonControl
//String ctrlName
// The logic behind: goes on top graph gets the wave with a cursor "A", starts with the point at A and do "n" FFTs with "m" points each
// also adds communicate: Place a cursor on a RawFC_Forcexxx graph  -> get the info from
// ADD pop-up menu: with number of points for each FFT and number of averages
        Variable n, i, Pbegin, Pend
        String DestFolder = "root:FC_Analysis:FFT_Traces"
        String windowNameGraphRawForce, WaveToFFT, SpectrumWaveName, nameOfGraph, WaveNamePrefix      = "FFT_fragment_"
        if (DataFolderExists(DestFolder) == 0)
        NewDataFolder $DestFolder
        endif

        String waveCursorAIsOn = CsrWave(A)
        If(stringmatch(waveCursorAIsOn, "RawFC_Force*"))
        windowNameGraphRawForce = WinName(0,1)

        // make a folder with that given RawFC_Force wave
        String DestFolderRawForce = "root:FC_Analysis:FFT_Traces:"+waveCursorAIsOn
                                if (DataFolderExists(DestFolderRawForce) == 0)
        NewDataFolder $DestFolderRawForce
                                endif

        Wave W = CsrWaveRef(A)
        WaveToFFT = DestFolderRawForce+":"+waveCursorAIsOn+"_toFFT"
```

```
	Duplicate/O W, $WaveToFFT		///R=[pcsr(A),pcsr(B)]

	Variable numPointsInOneFFT=5000,numFFTs=10
	Prompt numPointsInOneFFT, "Enter number of Points in one FFT on a selected
RawForce trace: "		// Set prompt for x param
	Prompt numFFTs, "Enter number of FFTs to average: "	// Set prompt for y param
	DoPrompt "Enter numPointsInOneFFT and numFFTs", numPointsInOneFFT, numFFTs
		if (V_Flag)
			return -1
	// User canceled
		endif

	// do "numFFTs" and average
	SpectrumWaveName = (DestFolderRawForce+":"+waveCursorAIsOn+"_FFTSpectrum")
		Pbegin = pcsr(A)
		Pend = pcsr(A)+numPointsInOneFFT-1
FFT/MAGS/RP=[Pbegin,Pend]/DEST=$SpectrumWaveName	$WaveToFFT
//WINF=Hanning
// instead of puting it by hand, trust FFT to give the right  scale takes square by MAGS
		Duplicate/O $SpectrumWaveName, W_FFT
			n= numpnts(W_FFT)
		FastOp W_FFT = (0.5/(n^2))*W_FFT
//to go to RMS voltage (from ampl. voltage) and to make later summation of this RMS power
spectrum
	// "point" independant
nameOfGraph = "FFT_Transforms_"+waveCursorAIsOn
	DoWindow/F $nameOfGraph//display W_FFT and brings it to the front
				if(V_Flag==0)
					Display/K=1 W_FFT
					DoWindow/C/F $nameOfGraph
					Label bottom "Frequency (Hz)"
```

```
                ShowInfo
                //Cursor/P A Spectrum 1
                //Cursor/P B Spectrum 557
            DoUpdate
            endif
            ModifyGraph log=1
            i=1

            if(numFFTs ==1)

            else
                    do
// must bring back original graph with RawFC_Forcexxx
DoWindow/F $windowNameGraphRawForce
    Pbegin =pcsr(A)+i*numPointsInOneFFT
    Pend = pcsr(A)+(i+1)*numPointsInOneFFT-1

FFT/MAGS/RP=[Pbegin,Pend]/DEST=rtempWave $WaveToFFT ///WINF=Hanning
FastOp W_FFT =  W_FFT + (0.5/(n^2))*rtempWave
                                i += 1
                                PeriodicUpdate(0.5)
                            while (i<numFFTs)
            endif
//DeletePoints 0,1,W_FFT
//Delete zero frequency (DC) component
        W_FFT /= i
        Duplicate/O W_FFT, $SpectrumWaveName
//Power = sum(Spectrum,xcsr(A),xcsr(B))
//sum values between cursors A,B
WaveStats/Q Spectrum
Spectrum = 20*Log(Spectrum/power)
```

```
	//hange "spectrum" to display in dB
	//Spectrum = Log(Spectrum)

	ResumeUpdate
	KillWaves/Z W_FFT, rtempWave

	Else
	print "Make sure to be on the RawFC_Force wave on the top graph and call this function again :)"
		Beep
	EndIF
	//	SetDataFolder $SourceFolder


End



*************************************************************************
Function FFT_FC_MultipleFragments()		//FFT_FC_fragments(ctrlName) : ButtonControl
	//String ctrlName
	// The logic behind: goes on top graph gets the wave with a cursor "A", starts with the point at A and do "n" FFTs with "m" points each
	// also adds communicate: Place a cursor on a RawFC_Forcexxx graph  -> get the info from
	// ADD pop-up menu: with number of points for each FFT and number of averages
	Variable n, i, j, Pbegin, Pend
	String DestFolder = "root:FC_Analysis:FFT_Traces"
	String windowNameGraphRawForce, WaveToFFT, SpectrumWaveName, nameOfGraph, WaveNamePrefix	= "FFT_fragment_"
```

```
	if (DataFolderExists(DestFolder) == 0)
		NewDataFolder $DestFolder
	endif


	String waveCursorAIsOn = CsrWave(A)
	If(stringmatch(waveCursorAIsOn, "RawFC_Force*"))
	windowNameGraphRawForce = WinName(0,1)

	// make a folder with that given RawFC_Force wave
	String DestFolderRawForce = "root:FC_Analysis:FFT_Traces:"+waveCursorAIsOn
				if (DataFolderExists(DestFolderRawForce) == 0)
				NewDataFolder $DestFolderRawForce
				endif

	Wave W = CsrWaveRef(A)
	WaveToFFT = DestFolderRawForce+":"+waveCursorAIsOn+"_toFFT"
	Duplicate/O W, $WaveToFFT		//	/R=[pcsr(A),pcsr(B)]

	Variable numPointsInOneFFT=5000,numFFTs=10, numAvgFFTs=10
	Prompt numPointsInOneFFT, "Enter number of Points in one FFT on a selected RawForce trace: "		// Set prompt for x param
	Prompt numFFTs, "Enter number of FFTs to average in one group: "		// Set prompt for y param
	Prompt numAvgFFTs, "Enter number of FFT groups: "
	DoPrompt "Enter numPointsInOneFFT, numFFTs and numAvgFFTs", numPointsInOneFFT, numFFTs, numAvgFFTs
			if (V_Flag)
				return -1
// User canceled
			endif
```

```
// do "numFFTs" and averages

j=1                    // j will go till number of averaged FFTs

do

SpectrumWaveName = (DestFolderRawForce+":"+waveCursorAIsOn+"_FFTSpectrum"
+num2str(j) )
Pbegin =pcsr(A)+(j-1)*numPointsInOneFFT*numFFTs
//Pbegin = pcsr(A)
Pend = pcsr(A)+((j-1)*numFFTs*numPointsInOneFFT)+numPointsInOneFFT-1
 //Pend = pcsr(A)+numPointsInOneFFT-1

FFT/MAGS/RP=[Pbegin,Pend]/DEST=$SpectrumWaveName $WaveToFFT
//  /WINF=Hanning
// instead of puting it by hand, trust FFT to give the right scale takes square by MAGS
Duplicate/O $SpectrumWaveName, W_FFT
n= numpnts(W_FFT)
FastOp W_FFT = (0.5/(n^2))*W_FFT
//to go to RMS voltage (from ampl. voltage) and to make later summation of this RMS power
spectrum
// "point" independant
nameOfGraph = "FFT_Transforms_"+waveCursorAIsOn
DoWindow/F $nameOfGraph//display W_FFT and brings it to the front
if(V_Flag==0)
Display/K=1 W_FFT
DoWindow/C/F $nameOfGraph
Label bottom "Frequency (Hz)"
ShowInfo
//Cursor/P A Spectrum 1
```

```
                    //Cursor/P B Spectrum 557
                DoUpdate
                endif
                ModifyGraph log=1

                i=1             // will go till num of FFTs
                if(numFFTs ==1)

                else

                    do
// must bring back original graph with RawFC_Forcexxx
DoWindow/F $windowNameGraphRawForce
    Pbegin =pcsr(A)+ (j- )*(numFFTs*numPointsInOneFFT)+i*numPointsInOneFFT
    Pend = pcsr(A)+ (j- 1)*(numFFTs*numPointsInOneFFT) +((i+1)*numPointsInOneFFT-
1)
    FFT/MAGS/RP=[Pbegin,Pend]/DEST=rtempWave $WaveToFFT ///WINF=Hanning
    FastOp W_FFT =  W_FFT + (0.5/(n^2))*rtempWave
                        i += 1
                        PeriodicUpdate(0.5)
                    while (i<numFFTs)
                endif
                //DeletePoints 0,1,W_FFT
    //Delete zero frequency (DC) component

                W_FFT /= i
    Duplicate/O W_FFT, $SpectrumWaveName
                DoUpdate

            j+=1
```

```
            while (j<(numAvgFFTs+1))


        ResumeUpdate
        KillWaves/Z W_FFT, rtempWave

        Else
        print "Make sure to be on the RawFC_Force wave on the top graph and call this function again :)"
        Beep
        EndIF
//      SetDataFolder $SourceFolder

End
```

Appendix F - Igor macro for Frequency Spectra Fit

```
#pragma rtGlobals=1          // Use modern global access method.
#include <Readback ModifyStr>  //Contains a string utility function,
GetNumFromModifyStr( modstr, key, listChar, itemNo ), that helps with parsing the string
returned by TraceInfo or AxisInfo.
#include <Multi-peak fitting 2.0>

Function model_fit(w,f): FitFunc                                    //Model for AIR
    Wave w                                                          //model coefficients
    Variable f                                                      // frequency
    Variable E, alpha, I, l1, mt, k, rho, a0, b
    Variable  Ren, tau, omegare, omegaim, rhof, eta, RealLambda, ImLambda
                                                                    //auxiliar variables
    Variable/C omega, Lambda
    Variable alpha1r, alpha1i, alpha2r, alpha2i, alpha3r, alpha4r
    Variable beta1r, beta1i, beta2r, beta2i, beta3r, beta4r
    Variable gamma1r, gamma2r, gamma3r, gamma4r
    Variable delta1r, delta2r, delta3r, delta4r
    Variable epsilon1r, epsilon1i, epsilon2r, epsilon2i, epsilon3r, epsilon4r
    Variable Br, Bi, Cr, Ci, Dr, Di, Er, Eii, Fr, Fi, Gr, Gi, Hr, Hi
    Variable Jr, Ji, Kr, Ki, Lr, Li, Mr, Mi, Nr, Ni, Pr, Pim, Qr, Qi
    Variable A1r, A1i, A2r, A2i

    E=w[13]                     //Young's modulus of cantilever
    alpha=pi/180*w[6]           //angle in radians

    rhof = 1.18                 //density of fluid [ 1.18kg/m^3 for air ] at 25 degrees C
    eta = 1.86*10^(-5)          //viscosity of fluid [ 1.86*10^(-5) kg/(m*s) for air ]
    b=w[7]
```

```
//rho=(((w[9]-((w[14]+10)*10^(-9)))*3100+w[14]*10^(-9)*19320+10*10^(-9)*7140)/w[9])
//density of cantilever with parameter w[14] which is the thickness of Au
rho=(((w[9]-((w[14]+10)*10^(-9)))*3100+w[14]*10^(-9)*19320+10*10^(-9)*7140)/w[9])

l1=0.3  //0.33                                                    //center mass point

mt = (rho*((w[7]+4*w[9])*(w[3]+2*w[9])+((w[7])^2/8))*w[9])/6    // tip mass
a0=1e-22                                                        //amplitude of excitation
I=w[7]*w[9]^3/12                                                // inertia momentum


//*****************************************************************************
        //Real Gamma function from Sader 1998 - they also exist as functions but they cannot be introduced in FitFunct  (Igor doesn't like that)
        //Imaginary Gamma function from Sader 1998

Ren = 2*pi*f*rhof*b^2/(4*eta)                   //Reynolds number
tau = log(Ren)

omegare = (0.91324 - 0.48274*tau + 0.46842*(tau^2) - 0.12886*(tau^3) + 0.044055*(tau^4) - 0.0035117*(tau^5) + 0.00069085*(tau^6))/(1 - 0.56964*tau + 0.48690*(tau^2) - 0.13444*(tau^3) + 0.045155*(tau^4) - 0.0035862*(tau^5) + 0.00069085*(tau^6))

omegaim = (-0.024134 - 0.029256*tau + 0.016294*(tau^2) - 0.00010961*(tau^3) + 0.000064577*(tau^4) - 0.000044510*(tau^5))/(1 - 0.59702*tau + 0.55182*(tau^2) - 0.18357*(tau^3) + 0.079156*(tau^4) - 0.014369*(tau^5) + 0.0028361*(tau^6))

omega = cmplx(omegare, omegaim)
```

```
RealLambda = real(omega*( 1+(4*sqrt(-1)*besselk(1,-sqrt(-1)*sqrt((sqrt(-1))*Ren))/(sqrt((sqrt(-1))*Ren)*besselk(0,-(sqrt(-1))*sqrt((sqrt(-1))*Ren))))))

 ImLambda = imag(omega*( 1+(4*sqrt(-1)*besselk(1,-sqrt(-1)*sqrt((sqrt(-1))*Ren))/(sqrt((sqrt(-1))*Ren)*besselk(0,-(sqrt(-1))*sqrt((sqrt(-1))*Ren))))))

 Lambda = cmplx(RealLambda, ImLambda)


//*************************************************************************


//k=sqrt(sqrt(3*rho/E)*f*4*pi/w[9])        //Dupas model wave vector for air / vacuum

//upgrades for wavevector with hydrodynamic function from Sader JAP, vol.84, 1998
//Dupas model with Sader correction in the approximation of small dissipation:

 k=((sqrt(3*rho/E)*f*4*pi/w[9])^(1/2))*((1 +  (pi*rhof*w[7]/(4*rho*w[9]))*(RealLambda)
)^(1/4))

  //Dupas model with Sader correction with dissipation:

  //   k=real(((sqrt(3*rho/E)*f*4*pi/w[9])^(1/2))*((1 +  (pi*rhof*w[7]/(4*rho*w[9]))*(Lambda)
)^(1/4)) )

//*************************************************************************

alpha1r=k^2*E*I*(- cos(k*w[11]*w[8])+cosh(k*w[11]*w[8])
+w[12]/E/I/k^3*(sin(k*w[11]*w[8]) +sinh(k*w[11]*w[8]))) +k*w[3]^2*(w[0]*sin(alpha)^2+
w[10]*cos(alpha)^2+ mt*4*pi^2*f^2*l1)*(-sin(k*w[11]*w[8]) +sinh(k*w[11]*w[8])-
w[12]/E/I/k^3*(cos(k*w[11]*w[8])-
```

cosh(k*w[11]*w[8])))+w[3]*(cos(k*w[11]*w[8])+cosh(k*w[11]*w[8])-
w[12]/E/I/k^3*(sin(k*w[11]*w[8]) -sinh(k*w[11]*w[8])))*sin(alpha)*cos(alpha)*(w[10]-w[0])

alpha1i = (-sin(k*w[11]*w[8]) + sinh(k*w[11]*w[8]) - w[12]/E/I/k^3*(cos(k*w[11]*w[8]) -
cosh(k*w[11]*w[8])))*2*pi*f*k*(w[1]*sin(alpha)^2+w[2]*cos(alpha)^2)*w[3]^2 + w[3] *
(cos(k*w[11]*w[8]) + cosh(k*w[11]*w[8]) - w[12]/E/I/k^3*(sin(k*w[11]*w[8]) -
sinh(k*w[11]*w[8]))) * sin(alpha)*cos(alpha)*(w[2]-w[1]) * 2*pi*f

beta1r = k^2*E*I*(-cos(k*w[11]*w[8]) - cosh(k*w[11]*w[8])) + k*w[3]^2*(w[0]*sin(alpha)^2
+ w[10]*cos(alpha)^2 + mt*4*pi^2*f^2*l1) * (-sin(k*w[11]*w[8]) - sinh(k*w[11]*w[8])) +
w[3]*(cos(k*w[11]*w[8]) - cosh(k*w[11]*w[8]))*sin(alpha) * cos(alpha)*(w[10]-w[0])

beta1i = k*w[3]^2*2*pi*f* (w[1] * sin(alpha)^2 + w[2]*cos(alpha)^2) * (-sin(k*w[11]*w[8]) -
sinh(k*w[11]*w[8])) + w[3]*2*pi*f*(cos(k*w[11]*w[8]) - cosh(k*w[11]*w[8])) *
sin(alpha)*cos(alpha)*(w[2] - w[1])

gamma1r = k^2*E*I*(cos(k*(1-w[11])*w[8]) - cosh(k*(1-w[11])*w[8]))

delta1r = k^2*E*I*(sin(k*(1-w[11])*w[8]) - sinh(k*(1-w[11])*w[8]))

epsilon1r = -(-w[3]*mt*9.81*sin(alpha)*l1 - a0*w[12]/2/k*(sin(k*w[11]*w[8]) +
sinh(k*w[11]*w[8])) + k*w[3]^2*(w[0]*sin(alpha)^2 + w[10]*cos(alpha)^2 +
mt*4*pi^2*f^2*l1) * a0*w[12]/2/E/I/k^3* (cos(k*w[11]*w[8]) - cosh(k*w[11]*w[8])) +
w[3]*sin(alpha) * cos(alpha)*(w[10] - w[0])* (sin(k*w[11]*w[8]) -
sinh(k*w[11]*w[8]))*a0*w[12]/2/E/I/k^3)

epsilon1i = -(k*w[3]^2*2*pi*f*(w[1]*sin(alpha)^2 + w[2]*cos(alpha)^2) *
a0*w[12]/2/E/I/k^3*(cos(k*w[11]*w[8]) - cosh(k*w[11]*w[8])) + w[3]*sin(alpha) *
cos(alpha)*(w[2] - w[1])*(sin(k*w[11]*w[8]) - sinh(k*w[11]*w[8])) *a0*w[12]/2/E/I/k^3)

```
alpha2r=k^3*E*I*(sin(k*w[11]*w[8])+sinh(k*w[11]*w[8])+w[12]/E/I/k^3*(cos(k*w[11]*w[8])
+cosh(k*w[11]*w[8])))-k*w[3]*sin(alpha)*cos(alpha)*(w[10]-w[0])*(-
sin(k*w[11]*w[8])+sinh(k*w[11]*w[8])-w[12]/E/I/k^3*(cos(k*w[11]*w[8])-
cosh(k*w[11]*w[8])))-(w[0]*cos(alpha)^2+w[10]*sin(alpha)^2-
mt*4*pi^2*f^2)*(cos(k*w[11]*w[8])+cosh(k*w[11]*w[8])-w[12]/E/I/k^3*(sin(k*w[11]*w[8])-
sinh(k*w[11]*w[8])))

alpha2i=-2*pi*f*k*w[3]*sin(alpha)*cos(alpha)*(w[2]-w[1])*(-
sin(k*w[11]*w[8])+sinh(k*w[11]*w[8])-w[12]/E/I/k^3*(cos(k*w[11]*w[8])-
cosh(k*w[11]*w[8])))-
2*pi*f*(w[1]*cos(alpha)^2+w[2]*sin(alpha)^2)*(cos(k*w[11]*w[8])+cosh(k*w[11]*w[8])-
w[12]/E/I/k^3*(sin(k*w[11]*w[8])-sinh(k*w[11]*w[8])))

beta2r = k^3*E*I*(sin(k*w[11]*w[8]) - sinh(k*w[11]*w[8]) )+
        k*w[3]*sin(alpha)*cos(alpha) * (w[10] -w[0])*(sin(k*w[11]*w[8]) +
sinh(k*w[11]*w[8])) - (w[0]*cos(alpha)^2+w[10]*sin(alpha)^2 -
mt*4*pi^2*f^2)*(cos(k*w[11]*w[8]) - cosh(k*w[11]*w[8]))

beta2i = k*2*pi*f*k*w[3]*sin(alpha)*cos(alpha)*(w[2]-w[1]) * (sin(k*w[11]*w[8]) -
sinh(k*w[11]*w[8])) - 2*pi*f*(w[1]*cos(alpha)^2 + w[2]*sin(alpha)^2) * (cos(k*w[11]*w[8]) -
cosh(k*w[11]*w[8]))

gamma2r = k^3*E*I*(sin(k*(1-w[11])*w[8]) + sinh(k*(1-w[11])*w[8]))

delta2r = k^3*E*I*(-cos(k*(1-w[11])*w[8]) + cosh(k*(1-w[11])*w[8]))

epsilon2r = -(-mt*9.81*cos(alpha) - a0*w[12]/2*(cos(k*w[11]*w[8]) + cosh(k*w[11]*w[8])) -
k*w[3]*sin(alpha)*cos(alpha)* (w[10]-w[0])*w[12]*a0/2/E/I/k^3*(cos(k*w[11]*w[8]) -
cosh(k*w[11]*w[8])) - (w[0]*cos(alpha)^2 + w[10]*sin(alpha)^2 -
mt*4*pi^2*f^2)*w[12]*a0/2/E/I/k^3*(sin(k*w[11]*w[8]) - sin(k*w[11]*w[8])))
```

epsilon2i = -(-k*w[3]*sin(alpha)*cos(alpha)*2*pi*f* (w[2]-
w[1])*w[12]*a0/2/E/i/k^3*(cos(k*w[11]*w[8]) - cosh(k*w[11]*w[8])) -
2*pi*f*(w[1]*cos(alpha)^2 + w[2]*sin(alpha)^2)*(sin(k*w[11]*w[8]) -
sinh(k*w[11]*w[8]))*w[12]*a0/2/E/I/k^3)

alpha3r = cos(k*w[11]*w[8]) + cosh(k*w[11]*w[8]) -w[12]/E/I/K^3*(sin(k*w[11]*w[8]) -
sinh(k*w[11]*w[8]))

beta3r = cos(k*w[11]*w[8]) - cosh(k*w[11]*w[8])

gamma3r = -cos(k*(1-w[11])*w[8]) - cosh(k*(1-w[11])*w[8])

delta3r = -sin(k*(1-w[11])*w[8]) - sinh(k*(1-w[11])*w[8])

epsilon3r = -a0*w[12]/E/I/2/k^3* (sin(k*w[11]*w[8]) - sinh(k*w[11]*w[8]))

alpha4r = -sin(k*w[11]*w[8]) + sinh(k*w[11]*w[8]) -w[12]/E/I/k^3*(cos(k*w[11]*w[8]) -
cosh(k*w[11]*w[8]))

beta4r = -sin(k*w[11]*w[8]) - sinh(k*w[11]*w[8])

gamma4r = -sin(k*(1-w[11])*w[8]) +sinh(k*(1-w[11])*w[8])

delta4r = cos(k*(1-w[11])*w[8]) + cosh(k*(1-w[11])*w[8])

epsilon4r = -a0*w[12]/E/I/2/k^3*(cos(k*w[11]*w[8]) - cosh(k*w[11]*w[8]))

Mr=epsilon2r*delta1r - delta2r*epsilon1r
Mi = epsilon2i*delta1r - delta2r*epsilon1i
Nr = gamma2r*delta1r - gamma1r*delta2r
Pr = alpha2r*delta1r - delta2r*alpha1r

Pim = alpha2i*delta1r - delta2r*alpha1i

Qr = beta2r*delta1r - delta2r*beta1r

Qi = beta2i*delta1r - delta2r*beta1i

Br = Mr/Nr

Bi=Mi/Nr

Cr=Pr/Nr

Ci=Pim/Nr

Dr=Qr/Nr

Di=Qi/Nr

Jr = epsilon3r*delta1r - delta3r*epsilon1r + Br*(delta3r*gamma1r - gamma3r*delta1r)

Ji = -delta3r*epsilon1i + Bi*(delta3r*gamma1r - gamma3r*delta1r)

Kr = beta3r*delta1r - delta3r*beta1r + Dr*(delta3r*gamma1r - gamma3r*delta1r)

Ki = -delta3r*beta1i + Di*(delta3r*gamma1r - gamma3r*delta1r)

Lr = alpha3r*delta1r - delta3r*alpha1r + Cr*(delta3r*gamma1r - gamma3r*delta1r)

Li = -delta3r*alpha1i + Ci*(delta3r*gamma1r - gamma3r*delta1r)

Er = (Jr*Kr + Ji*Ki)/(Kr^2 + Ki^2)

Eii = (Ji*Kr - Jr*Ki)/(Kr^2 + Ki^2)

Fr = (Lr*Kr - Li*Ki)/(Kr^2 + Ki^2)

Fi = (Li*Kr - Lr*Ki)/(Kr^2 + Ki^2)

Gr = epsilon4r - beta4r*Er - gamma4r*Br + gamma4r*(Dr*Er - Di*Eii) - delta4r/delta1r*(epsilon1r - (beta1r*Er - beta1i*Eii) - gamma1r*Br + gamma1r*(Dr*Er - Di*Eii))

Gi = -beta4r*Eii - gamma4r*Bi + gamma4r*(Di*Er + Dr*Eii) - delta4r/delta1r*(epsilon1i - (beta1r*Eii + beta1i*Er) - gamma1r*Bi + gamma1r*(Di*Er + Dr*Eii))

Hr = alpha4r - beta4r*Fr - gamma4r*Cr + gamma4r*(Dr*Fr - Di*Fi) - delta4r/delta1r*(alpha1r - (beta1r*Fr - beta1i*Fi) - gamma1r*Cr + gamma1r*(Dr*Fr - Di*Fi))

```
Hi = -beta4r*Fi - gamma4r*Ci + gamma4r*(Di*Fr + Dr*Fi) - delta4r/delta1r*(alpha1i -
(beta1r*Fi + beta1i*Fr) - gamma1r*Ci + gamma1r*(Di*Fr + Dr*Fi))

                A1r = (Gr*Hr + Gi*Hi)/(Hr^2 + Hi^2)          //real part of A1
                A1i = (Gi*Hr - Gr*Hi)/(Hr^2 + Hi^2)          //imaginary  part of A1
                A2r = Er - (Fr*A1r - Fi*A1i)                 //real part of A2
                A2i = Eii - (Fr*A1i + Fi*A1r)                //imaginary part of A2

return k/a0*w[5]*sqrt((A1r*(-sin(w[4]*k*w[8]) + sinh(w[4]*k*w[8]) -
w[12]/E/I/k^3*(cos(w[4]*k*w[8]) - cosh(w[4]*k*w[8]))) + A2r*(-sin(w[4]*k*w[8]) -
sinh(w[4]*k*w[8])) + w[12]/2/E/I/k^3*(cos(w[4]*k*w[8]) - cosh(w[4]*k*w[8]))*a0)^2 +
(A1i*(-sin(w[4]*k*w[8]) + sinh(w[4]*k*w[8]) - w[12]/2/E/I/k^3*(cos(w[4]*k*w[8]) -
cosh(w[4]*k*w[8]))) + A2i*(-sin(w[4]*k*w[8]) - sinh(w[4]*k*w[8])))^2)

                End

//*****************************************************************************
        //Fitting model for liquid:

Function model_fit_water(w,f): FitFunc
        Wave w                                  //model coefficients
        Variable f                               // frequency
        Variable E, alpha, I, l1, mt, k, rho, a0, b
        Variable  Ren, tau, omegare, omegaim, rhof, eta, RealLambda, ImLambda
                                                //auxiliar variables
        Variable/C omega, Lambda
        Variable alpha1r, alpha1i, alpha2r, alpha2i, alpha3r, alpha4r
        Variable beta1r, beta1i, beta2r, beta2i, beta3r, beta4r
```

```
Variable gamma1r, gamma2r, gamma3r, gamma4r
Variable delta1r, delta2r, delta3r, delta4r
Variable epsilon1r, epsilon1i, epsilon2r, epsilon2i, epsilon3r, epsilon4r
Variable Br, Bi, Cr, Ci, Dr, Di, Er, Eii, Fr, Fi, Gr, Gi, Hr, Hi
Variable Jr, Ji, Kr, Ki, Lr, Li, Mr, Mi, Nr, Ni, Pr, Pim, Qr, Qi
Variable A1r, A1i, A2r, A2i


E=w[13]   //290e9                    //Young's modulus of cantilever
alpha=pi/180*w[6]                    //angle in radians

rhof = 998                           //density of fluid [ 998kg/m^3 for water ] at 25 degrees C
eta = 0.91*10^(-3)                   //viscosity of fluid [ 0.91*10^(-3) kg/(m*s) for water ]
b=w[7]

rho=(((w[9]-6*10^(-8))*3100+5*10^(-8)*19320+1*10^(-8)*7140)/w[9])*w[14]
                                      //density of cantilever
l1=0.3                                //center mass point

mt = (rho*((w[7]+4*w[9])*(w[3]+2*w[9])+((w[7])^2/8))*w[9])/6    // tip mass

a0=1e-22                              //amplitude of excitation
I=w[7]*w[9]^3/12                      // inertia momentum


//*****************************************************************************
      //Real Gamma function from Sader 1998 - they also exist as functions but they cannot be introduced in FitFunct  (Igor doesn't like that)
      //Imaginary Gamma function from Sader 1998
```

```
Ren = 2*pi*f*rhof*b^2/(4*eta)                    //Reynolds number
tau = log(Ren)

omegare = (0.91324 - 0.48274*tau + 0.46842*(tau^2) - 0.12886*(tau^3) + 0.044055*(tau^4) -
0.0035117*(tau^5) + 0.00069085*(tau^6))/(1 - 0.56964*tau + 0.48690*(tau^2) -
0.13444*(tau^3) + 0.045155*(tau^4) - 0.0035862*(tau^5) + 0.00069085*(tau^6))

omegaim = (-0.024134 - 0.029256*tau + 0.016294*(tau^2) - 0.00010961*(tau^3) +
0.000064577*(tau^4) - 0.000044510*(tau^5))/(1 - 0.59702*tau + 0.55182*(tau^2) -
0.18357*(tau^3) + 0.079156*(tau^4) - 0.014369*(tau^5) + 0.0028361*(tau^6))

omega = cmplx(omegare, omegaim)

RealLambda = real(omega*( 1+(4*sqrt(-1)*besselk(1,-sqrt(-1)*sqrt((sqrt(-1))*Ren))/(sqrt((sqrt(-
1))*Ren)*besselk(0,-(sqrt(-1))*sqrt((sqrt(-1))*Ren))))))

ImLambda = imag(omega*( 1+(4*sqrt(-1)*besselk(1,-sqrt(-1)*sqrt((sqrt(-1))*Ren))/(sqrt((sqrt(-
1))*Ren)*besselk(0,-(sqrt(-1))*sqrt((sqrt(-1))*Ren))))))

Lambda = cmplx(RealLambda, ImLambda)

//**************************************************************************

//upgrades for wavevector with hydrodynamic function from Sader JAP, vol.84, 1998
//Dupas model with Sader correction in the approximation of small dissipation:

 // k=((sqrt(3*rho/E)*f*4*pi/w[9])^(1/2))*((1 +  (pi*rhof*w[7]/(4*rho*w[9]))*(RealLambda)
)^(1/4))
//Dupas model with Sader correction with dissipation:
  // k=real(((sqrt(3*rho/E)*f*4*pi/w[9])^(1/2))*((1 +  (pi*rhof*w[7]/(4*rho*w[9]))*(Lambda)
)^(1/4)) )
```

//Dupas model with our fitted Hydrodynamic function:

// k=real(((sqrt(3*rho/E)*f*4*pi/w[9])^(1/2))*((1 +
(pi*rhof*w[7]/(4*rho*w[9]))*(w[15]*f^(w[16]+w[17]*log(f))))^(1/4)))

//**********

alpha1r=k^2*E*I*(-cos(k*w[11]*w[8]) +cosh(k*w[11]*w[8])
+w[12]/E/I/k^3*(sin(k*w[11]*w[8]) +sinh(k*w[11]*w[8]))) +k*w[3]^2*(w[0]*sin(alpha)^2
+w[10]*cos(alpha)^2+mt*4*pi^2*f^2*l1)*(-sin(k*w[11]*w[8])+sinh(k*w[11]*w[8])-
w[12]/E/I/k^3*(cos(k*w[11]*w[8])-cosh(k*w[11]*w[8]))) +w[3]*(cos(k*w[11]*w[8])
+cosh(k*w[11]*w[8])-w[12]/E/I/k^3*(sin(k*w[11]*w[8])-
sinh(k*w[11]*w[8])))*sin(alpha)*cos(alpha)*(w[10]-w[0])

alpha1i = (-sin(k*w[11]*w[8]) + sinh(k*w[11]*w[8]) - w[12]/E/I/k^3*(cos(k*w[11]*w[8]) -
cosh(k*w[11]*w[8])))*2*pi*f*k*(w[1]*sin(alpha)^2+w[2]*cos(alpha)^2)*w[3]^2 + w[3] *
(cos(k*w[11]*w[8]) + cosh(k*w[11]*w[8]) - w[12]/E/I/k^3*(sin(k*w[11]*w[8]) -
sinh(k*w[11]*w[8]))) * sin(alpha)*cos(alpha)*(w[2]-w[1]) * 2*pi*f

beta1r = k^2*E*I*(-cos(k*w[11]*w[8]) - cosh(k*w[11]*w[8])) + k*w[3]^2*(w[0]*sin(alpha)^2
+ w[10]*cos(alpha)^2 + mt*4*pi^2*f^2*l1) * (-sin(k*w[11]*w[8]) - sinh(k*w[11]*w[8])) +
w[3]*(cos(k*w[11]*w[8]) - cosh(k*w[11]*w[8]))*sin(alpha) * cos(alpha)*(w[10]-w[0])

beta1i = k*w[3]^2*2*pi*f* (w[1] * sin(alpha)^2 + w[2]*cos(alpha)^2) * (-sin(k*w[11]*w[8]) -
sinh(k*w[11]*w[8])) + w[3]*2*pi*f*(cos(k*w[11]*w[8]) - cosh(k*w[11]*w[8])) *
sin(alpha)*cos(alpha)*(w[2] - w[1])

gamma1r = k^2*E*I*(cos(k*(1-w[11])*w[8]) - cosh(k*(1-w[11])*w[8]))

delta1r = k^2*E*I*(sin(k*(1-w[11])*w[8]) - sinh(k*(1-w[11])*w[8]))

epsilon1r = -(-w[3]*mt*9.81*sin(alpha)*l1 - a0*w[12]/2/k*(sin(k*w[11]*w[8]) + sinh(k*w[11]*w[8])) + k*w[3]^2*(w[0]*sin(alpha)^2 + w[10]*cos(alpha)^2 + mt*4*pi^2*f^2*l1) * a0*w[12]/2/E/I/k^3* (cos(k*w[11]*w[8]) - cosh(k*w[11]*w[8])) + w[3]*sin(alpha) * cos(alpha)*(w[10] - w[0])* (sin(k*w[11]*w[8]) - sinh(k*w[11]*w[8]))*a0*w[12]/2/E/I/k^3)

epsilon1i = -(k*w[3]^2*2*pi*f*(w[1]*sin(alpha)^2 + w[2]*cos(alpha)^2) * a0*w[12]/2/E/I/k^3*(cos(k*w[11]*w[8]) - cosh(k*w[11]*w[8])) + w[3]*sin(alpha) * cos(alpha)*(w[2] - w[1])*(sin(k*w[11]*w[8]) - sinh(k*w[11]*w[8])) *a0*w[12]/2/E/I/k^3)

alpha2r=k^3*E*I*(sin(k*w[11]*w[8])+sinh(k*w[11]*w[8])+w[12]/E/I/k^3*(cos(k*w[11]*w[8]) +cosh(k*w[11]*w[8])))-k*w[3]*sin(alpha)*cos(alpha)*(w[10]-w[0])*(-sin(k*w[11]*w[8]) +sinh(k*w[11]*w[8])-w[12]/E/I/k^3*(cos(k*w[11]*w[8])-cosh(k*w[11]*w[8])))-(w[0]*cos(alpha)^2+w[10]*sin(alpha)^2-mt*4*pi^2*f^2)*(cos(k*w[11]*w[8]) +cosh(k*w[11]*w[8])-w[12]/E/I/k^3*(sin(k*w[11]*w[8])-sinh(k*w[11]*w[8])))

alpha2i=-2*pi*f*k*w[3]*sin(alpha)*cos(alpha)*(w[2]-w[1])*(-sin(k*w[11]*w[8]) +sinh(k*w[11]*w[8])- w[12]/E/I/k^3*(cos(k*w[11]*w[8])-cosh(k*w[11]*w[8])))-2*pi*f*(w[1]*cos(alpha)^2 +w[2]*sin(alpha)^2)*(cos(k*w[11]*w[8])+cosh(k*w[11]*w[8])-w[12]/E/I/k^3*(sin(k*w[11]*w[8])-sinh(k*w[11]*w[8])))

beta2r = k^3*E*I*(sin(k*w[11]*w[8]) - sinh(k*w[11]*w[8]))+ k*w[3]*sin(alpha)*cos(alpha) * (w[10] -w[0])*(sin(k*w[11]*w[8]) + sinh(k*w[11]*w[8])) - (w[0]*cos(alpha)^2+w[10]*sin(alpha)^2 -mt*4*pi^2*f^2)*(cos(k*w[11]*w[8]) - cosh(k*w[11]*w[8]))

beta2i = k*2*pi*f*k*w[3]*sin(alpha)*cos(alpha)*(w[2]-w[1]) * (sin(k*w[11]*w[8]) - sinh(k*w[11]*w[8])) - 2*pi*f*(w[1]*cos(alpha)^2 + w[2]*sin(alpha)^2) * (cos(k*w[11]*w[8]) - cosh(k*w[11]*w[8]))

gamma2r = k^3*E*I*(sin(k*(1-w[11])*w[8]) + sinh(k*(1-w[11])*w[8]))

delta2r = k^3*E*I*(-cos(k*(1-w[11])*w[8]) + cosh(k*(1-w[11])*w[8]))

epsilon2r = -(-mt*9.81*cos(alpha) - a0*w[12]/2*(cos(k*w[11]*w[8]) + cosh(k*w[11]*w[8])) - k*w[3]*sin(alpha)*cos(alpha)* (w[10]-w[0])*w[12]*a0/2/E/I/k^3*(cos(k*w[11]*w[8]) - cosh(k*w[11]*w[8])) - (w[0]*cos(alpha)^2 + w[10]*sin(alpha)^2 - mt*4*pi^2*f^2)*w[12]*a0/2/E/I/k^3*(sin(k*w[11]*w[8]) - sin(k*w[11]*w[8])))

epsilon2i = -(-k*w[3]*sin(alpha)*cos(alpha)*2*pi*f* (w[2]-w[1])*w[12]*a0/2/E/i/k^3*(cos(k*w[11]*w[8]) - cosh(k*w[11]*w[8])) - 2*pi*f*(w[1]*cos(alpha)^2 + w[2]*sin(alpha)^2)*(sin(k*w[11]*w[8]) - sinh(k*w[11]*w[8]))*w[12]*a0/2/E/I/k^3)

alpha3r = cos(k*w[11]*w[8]) + cosh(k*w[11]*w[8]) -w[12]/E/I/K^3*(sin(k*w[11]*w[8]) - sinh(k*w[11]*w[8]))

beta3r = cos(k*w[11]*w[8]) - cosh(k*w[11]*w[8])

gamma3r = -cos(k*(1-w[11])*w[8]) - cosh(k*(1-w[11])*w[8])

delta3r = -sin(k*(1-w[11])*w[8]) - sinh(k*(1-w[11])*w[8])

epsilon3r = -a0*w[12]/E/I/2/k^3* (sin(k*w[11]*w[8]) - sinh(k*w[11]*w[8]))

alpha4r = -sin(k*w[11]*w[8]) + sinh(k*w[11]*w[8]) -w[12]/E/I/k^3*(cos(k*w[11]*w[8]) - cosh(k*w[11]*w[8]))

beta4r = -sin(k*w[11]*w[8]) - sinh(k*w[11]*w[8])

gamma4r = -sin(k*(1-w[11])*w[8]) +sinh(k*(1-w[11])*w[8])

delta4r = cos(k*(1-w[11])*w[8]) + cosh(k*(1-w[11])*w[8])

epsilon4r = -a0*w[12]/E/I/2/k^3*(cos(k*w[11]*w[8]) - cosh(k*w[11]*w[8]))

Mr=epsilon2r*delta1r - delta2r*epsilon1r
Mi = epsilon2i*delta1r - delta2r*epsilon1i
Nr = gamma2r*delta1r - gamma1r*delta2r
Pr = alpha2r*delta1r - delta2r*alpha1r
Pim = alpha2i*delta1r - delta2r*alpha1i
Qr = beta2r*delta1r - delta2r*beta1r
Qi = beta2i*delta1r - delta2r*beta1i

Br = Mr/Nr
Bi=Mi/Nr
Cr=Pr/Nr
Ci=Pim/Nr
Dr=Qr/Nr
Di=Qi/Nr

Jr = epsilon3r*delta1r - delta3r*epsilon1r + Br*(delta3r*gamma1r - gamma3r*delta1r)
Ji = -delta3r*epsilon1i + Bi*(delta3r*gamma1r - gamma3r*delta1r)
Kr = beta3r*delta1r - delta3r*beta1r + Dr*(delta3r*gamma1r - gamma3r*delta1r)
Ki = -delta3r*beta1i + Di*(delta3r*gamma1r - gamma3r*delta1r)
Lr = alpha3r*delta1r - delta3r*alpha1r + Cr*(delta3r*gamma1r - gamma3r*delta1r)
Li = -delta3r*alpha1i + Ci*(delta3r*gamma1r - gamma3r*delta1r)
Er = (Jr*Kr + Ji*Ki)/(Kr^2 + Ki^2)
Eii = (Ji*Kr - Jr*Ki)/(Kr^2 + Ki^2)
Fr = (Lr*Kr - Li*Ki)/(Kr^2 + Ki^2)
Fi = (Li*Kr - Lr*Ki)/(Kr^2 + Ki^2)

```
Gr = epsilon4r - beta4r*Er - gamma4r*Br + gamma4r*(Dr*Er - Di*Eii) -
delta4r/delta1r*(epsilon1r - (beta1r*Er - beta1i*Eii) - gamma1r*Br + gamma1r*(Dr*Er -
Di*Eii))

Gi = -beta4r*Eii - gamma4r*Bi + gamma4r*(Di*Er + Dr*Eii) - delta4r/delta1r*(epsilon1i -
(beta1r*Eii + beta1i*Er) - gamma1r*Bi + gamma1r*(Di*Er + Dr*Eii))

Hr = alpha4r - beta4r*Fr - gamma4r*Cr + gamma4r*(Dr*Fr - Di*Fi) - delta4r/delta1r*(alpha1r -
(beta1r*Fr - beta1i*Fi) - gamma1r*Cr + gamma1r*(Dr*Fr - Di*Fi))

Hi = -beta4r*Fi - gamma4r*Ci + gamma4r*(Di*Fr + Dr*Fi) - delta4r/delta1r*(alpha1i -
(beta1r*Fi + beta1i*Fr) - gamma1r*Ci + gamma1r*(Di*Fr + Dr*Fi))

                A1r = (Gr*Hr + Gi*Hi)/(Hr^2 + Hi^2)          //real part of A1
                A1i = (Gi*Hr - Gr*Hi)/(Hr^2 + Hi^2)          //imaginary  part of A1
                A2r = Er - (Fr*A1r - Fi*A1i)                 //real part of A2
                A2i = Eii - (Fr*A1i + Fi*A1r)                //imaginary  part of A2


return k/a0*w[5]*sqrt((A1r*(-sin(w[4]*k*w[8]) + sinh(w[4]*k*w[8]) -
w[12]/E/I/k^3*(cos(w[4]*k*w[8]) - cosh(w[4]*k*w[8]))) + A2r*(-sin(w[4]*k*w[8]) -
sinh(w[4]*k*w[8])) + w[12]/2/E/I/k^3*(cos(w[4]*k*w[8]) - cosh(w[4]*k*w[8]))*a0)^2 +
(A1i*(-sin(w[4]*k*w[8]) + sinh(w[4]*k*w[8]) - w[12]/2/E/I/k^3*(cos(w[4]*k*w[8]) -
cosh(w[4]*k*w[8]))) + A2i*(-sin(w[4]*k*w[8]) - sinh(w[4]*k*w[8])))^2)

                End
```

Appendix G - Other cantilever fits

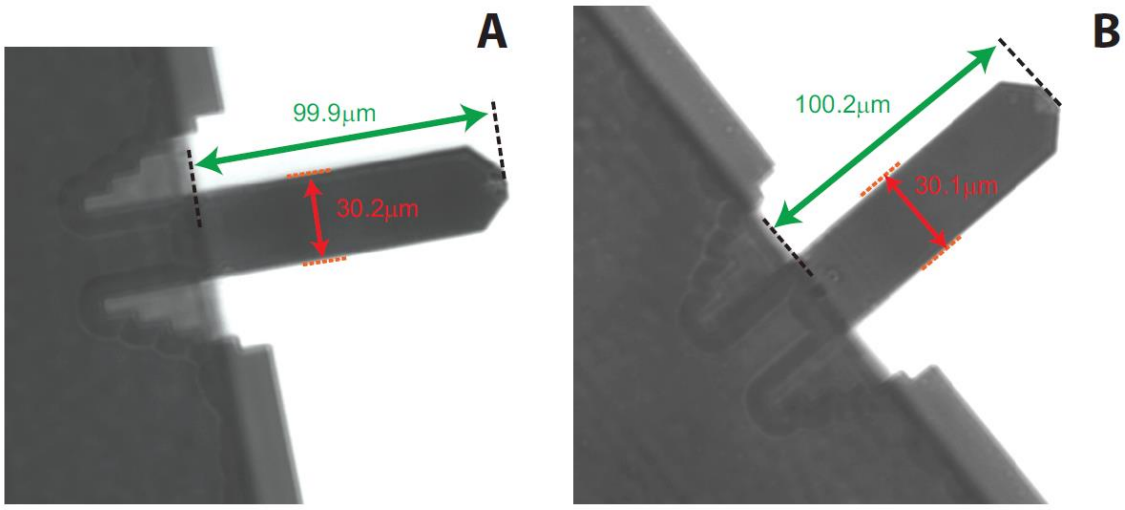

FIG. K1. Examples of optical microscope measurements of the aspect ratio for two AFM cantilevers used in our study.

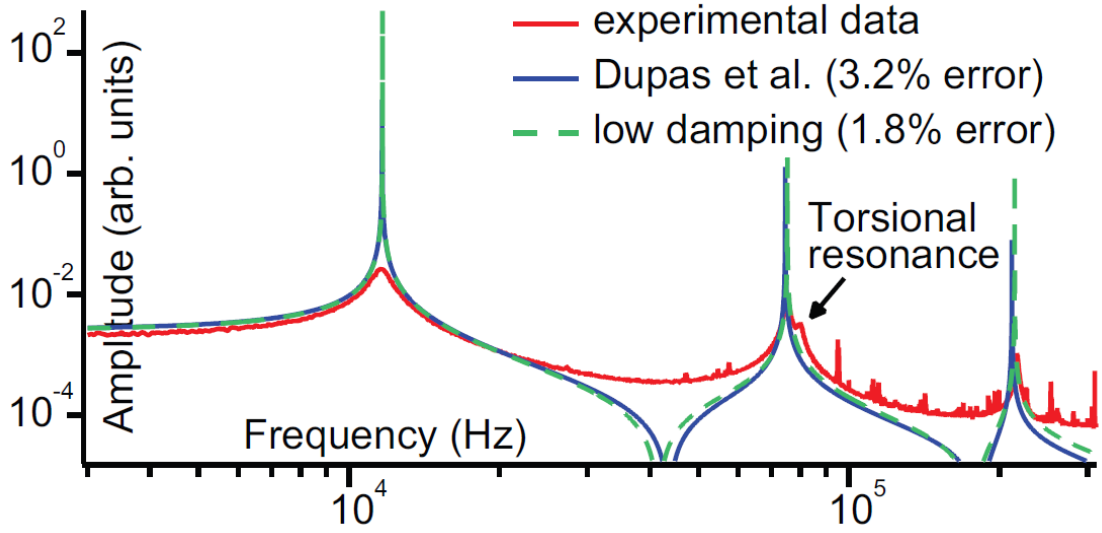

FIG. K2. Application of our models to another compliant AFM Biolevers model BL-RC150VB in air. First torsional resonance is marked by an arrow.

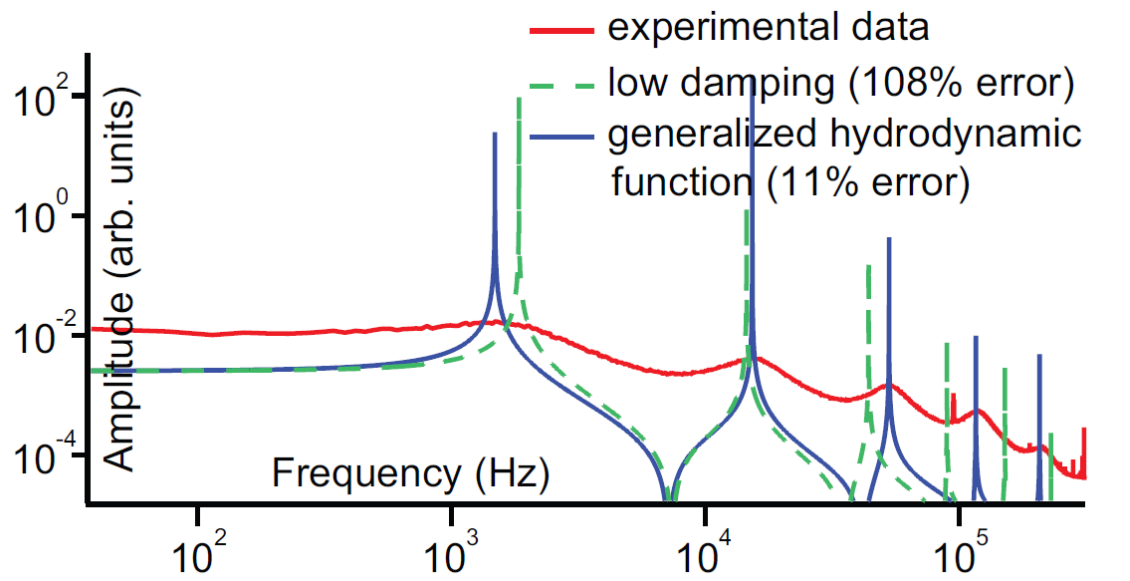

FIG. K3. Application of our models to another compliant AFM Biolevers model BL-RC150VB in water.

Appendix H - Analysis of Propagated Errors from Geometrical and Mechanical Parameters used in our Models

In our analysis of propagated errors we concentrate on the errors of our final model, i.e., the model used to fit the resonance frequency of a compliant cantilever in water. The errors arising from the manufacturers specifications for the cantilever c1 in the paper to the results in Fig. 18 in section II.1, would produce very large errors of about 290%. This error is calculated over five resonances in PBS and using a formula in Ref. 29 from Part II, and referenced here for completeness.1 The thickness values were provided by the manufacturer for a given batch of cantilevers. Tip height was varied in between the values provided by the manufacturer. See an overall error of 286% in Table I. The main culprit of such high errors are large errors coming from the manufacturer's specifications.

parameter	mean value	minimum value	maximum value
length (μm)	100	99	101
width (μm)	30.2	29.2	31.2
thickness (nm)	200	174	226
tip height (μm)	7.5	5.0	10.0
β	0.95	0.90	0.99
α (degrees)	20	20	20
ϵ	0.8	0.8	0.8
E (GPa)	155	145	165
density (kg/m^3)	7430	7993	6867
tip mass (pg)	90	62	118
relative error (%)		16	286

TABLE I. Properties of the cantilever c1 used in our error propagation study with values obtained from the manufacturer's specifications. Length and width were measured using optical microscopy.

However, the manufacturers specifications are only the starting point for our analysis. This is because we go substantially beyond fitting only one mechanical resonance for each cantilever. We fit simultaneously three flexular resonances and when visible - one torsional resonance in air, as well as up to five flexular resonances in water. We purposefully choose the critical parameters such as thickness, tip height, and Young's modulus to be fitted. Then, we calculate the fit error for each of these parameters keeping the other parameters constant and by varying the given parameter to match a maximum fit deviation from the experimental data. For example, for Figure 17 in section II.1, our model with low damping in air produced 1.6% error, meaning 1.6% combined frequency shifts for three resonance frequencies. Thus, to calculate an error of the thickness, we would keep all the other parameters fixed, and vary the value of thickness around its fitted value until a 1.6% departure from the fit is obtained.

We repeat such a procedure for the tip height, and the Young's modulus. Tables II – V gather the results of our analysis for the four cantilevers studied in the paper. These errors propagate to the fits in PBS, see Figure 18 in section II.1.

As seen in the Tables II - V we obtain the propagated errors in PBS between 10% to 40% with an average of 20%. Such errors are to be compared with 81% accumulated frequency shift expected for the cantilever c1 in contact with protein, see Fig. K4. Thus, contact with the protein is expected to produce shifts of the cantilevers resonance spectra, which are larger than the propagated errors from the material and geometrical properties of the used cantilevers.

parameter	fitted value	minimum value	maximum value
length (μm)	99.6	99.6	99.6
width (μm)	30.2	30.2	30.2
thickness (nm)	179.0	176.0	179.5
tip height (μm)	10.0	6.0	10.0
β	0.91	0.9	0.94
α (degrees)	20	20	20
ϵ	0.8	0.8	0.8
E (GPa)	155	145	147
density (kg/m^3)	7856	7938	7843
tip mass (pg)	102	72	102
relative error (%)	4	17	6

TABLE II. Properties of the cantilever c1 used in our error propagation study with values obtained from values fitted with low damping model.

parameter	fitted value	minimum value	maximum value
length (μm)	100.7	100.7	100.7
width (μm)	31.1	31.1	31.1
thickness (nm)	174.0	174.0	176.7
tip height (μm)	10.0	5.0	10.0
β	0.97	0.9	0.99
α (degrees)	20	20	20
ϵ	0.8	0.8	0.8
E (GPa)	145	145	150
density (kg/m^3)	7993	7993	7918
tip mass (pg)	104	67	105
relative error (%)	7	7	23

TABLE III. Properties of the cantilever c2 used in our error propagation study with values obtained from values fitted with low damping model

parameter	fitted value	minimum value	maximum value
length (μm)	100.5	100.5	100.5
width (μm)	31.1	31.1	31.1
thickness (nm)	177.0	176.3	178.0
tip height (μm)	9.5	5.5	10.0
β	0.95	0.92	0.97
α (degrees)	20	20	20
ϵ	0.8	0.8	0.8
E (GPa)	155	154	157
density (kg/m^3)	7910	7929	7883
tip mass (pg)	101	72	105
relative error (%)	11	40	17

TABLE IV. Properties of the cantilever c3 used in our error propagation study with values obtained from values fitted with low damping model fitted value minimum value maximum value

parameter	fitted value	minimum value	maximum value
length (μm)	100.4	100.4	100.4
width (μm)	30.2	30.2	30.2
thickness (nm)	174.0	174.0	175.5
tip height (μm)	10.0	5.0	10.0
β	0.91	0.9	0.97
α (degrees)	20	20	20
ϵ	0.8	0.8	0.8
E (GPa)	146	145	150
density (kg/m^3)	7993	7993	7951
tip mass (pg)	101	65	101
relative error (%)	10	8	16

TABLE V. Properties of the cantilever c4 used in our error propagation study with values obtained from values fitted with low damping model

Appendix I - SEM Analysis for one of the Cantilevers

Another way to obtain precise estimates of the critical geometrical parameters for AFM cantilevers is to use the scanning electron microscopy (SEM) to measure their length, width, tip height, and most importantly thickness. Figure I1 provides such SEM measurements, which we attempted a cantilever, which is similar to the cantilevers used in our paper. Table VI list the obtained parameters along with other geometrical and mechanical parameters, and their errors. Reproducing a similar error propagation analysis as for the cantilevers c1 to c4 studied in the paper produces the overall propagated error for the model in PBS to be up to 17 %. This is very similar as the results obtained for the other cantilevers.

Thus, we conclude that the SEM analysis provides similar upgrades on the manufacturer's specifications as our simultaneous fits in air.

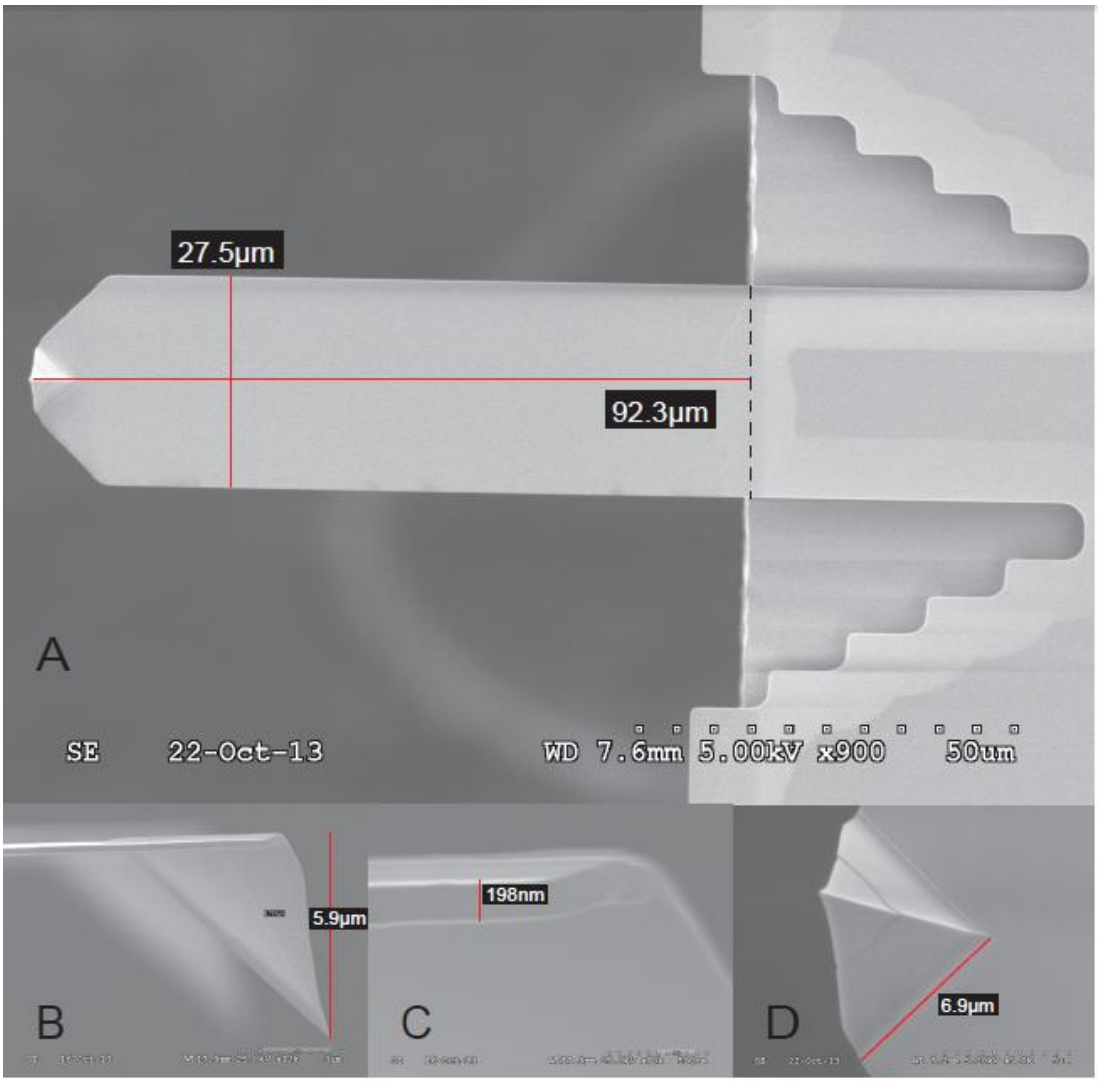

FIG. I1. SEM measurements of crucial geometrical parameters for one of the cantilevers used in our study.

parameter	fitted value	minimum value	maximum value
length (μm)	91.3	91.3	93.3
width (μm)	28	27	28
thickness (nm)	201	195	201
tip height (μm)	6.0	5.5	6.5
β	0.95	0.95	0.95
α (degrees)	20	20	20
ϵ	0.8	0.8	0.8
E (GPa)	150	150	150
density (kg/m^3)	7336	7466	7336
tip mass (pg)	69	62	73
relative error (%)	5	15	17

TABLE VI. Properties of the test cantilever used in our error propagation study with values obtained from values fitted with low damping model and parameters measured with SEM.

Appendix J - AFM cantilever models

In these experiments the AFM cantilevers were excited by thermal-noise and we recorded the fluctuations of the cantilever as a function of time as sketched in Figure J1. When a single molecule is being hold between the tip and substrate, the response is a convoluted fluctuation of both the cantilever and the molecule. Therefore, we can measure the nanomechanical properties of proteins and peptides single molecules by subtracting the response of the cantilever from the recorded fluctuations. In order to obtain accurate results for biomolecules, we need to start with an accurate model for the cantilever.

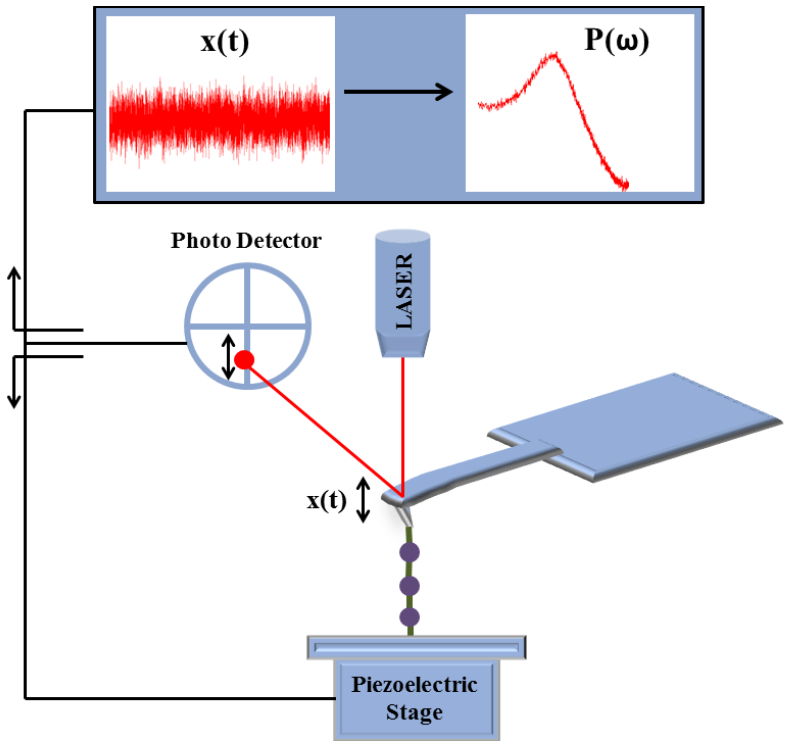

Figure J1. Schematic representation of the experimental setup from which we can extract the AFM cantilever fluctuation response as a function of time. The recorded signal is then Fourier transformed to obtain the frequency spectra of a cantilever.

The dynamic behavior of AFM cantilevers in liquid is completely different from the one in air due to the effect of the hydrodynamic force. There are several models used to theoretically describe the response of an AFM cantilever. Schaffer et al. proposed a simple model for the behavior of an oscillating cantilever in liquid media based on the assumption that the beam is driven by a uniform harmonic motion.[1] A more realistic model has been developed by Jai et al. modeling the cantilever as a point mass and spring.[2]

Mass on a damped spring model

This model is assuming that the cantilever can be considered a damped spring attached to a mass, m_c, the cantilever mass, see figure J2.

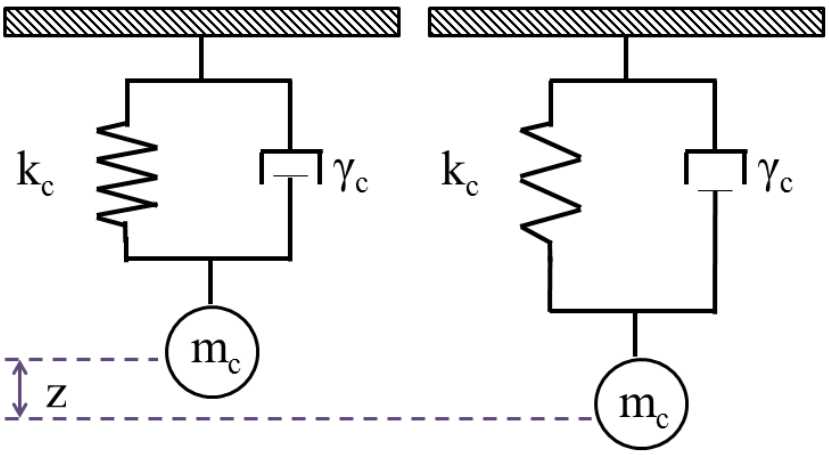

Figure J2. Mass on a damped spring model an AFM cantilever

The spring undergoes Hooke's Law:

$$F_s = -kz$$

and the damping can be described by:

$$F_d = \gamma_c \frac{dz}{dt}$$

Using Newton's second law, $\sum F = ma$, we obtain:

$$m\ddot{z} = -\gamma_c \dot{z} - kz$$

Rearranging, the equation becomes:

$$m\ddot{z} + \gamma_c \dot{z} + kz = 0$$

Dividing by m:

$$\ddot{z} + \frac{\gamma_c}{m}\dot{z} + \frac{k}{m}z = 0$$

And therefore, the natural frequency of the system is:

$$\omega_0 = 2\pi f_0 = \sqrt{\frac{k}{m}}$$

From this model we can only obtain a single resonance frequency, while a three-dimensional cantilever has an infinite number of resonance frequencies. Therefore, to describe the AFM cantilever, a better model is required.

Multi-resonance cantilever models

Several models were developed to describe the multi- resonance frequency spectra of a three-dimensional beam. In Sharabi et al., the hydrodynamic force has been modeled with string of spheres including the effect of the damping and the added mass of the cantilever.[3]

Sader, see Ref. 27 in Part II, proposed a general theoretical model with more rigorous assumptions. The cantilever is considered a continuous mass system excited by an arbitrary force and:

1. The cross section of the beam is uniform over its entire length;

2. The length of the beam greatly exceeds its nominal width;

3. Internal frictional effects are negligible;

4. Amplitude of vibration of the beam is far smaller than any length scale in the beam geometry.

Also, only the flexural modes are considered and the torsional effects are neglected.

Other models were derived similarly, using the Euler-Bernoulli beam theory, see Ref. 24, 27 in Part II.

References:

1. T. E. Schäffer, J. P. Cleveland, F. Ohnesorge, D. A. Walters and P. K. Hansma, *J. Appl. Phys*., **1996**, *80*, 3622

2. C. Jai, T. Cohen-Bouhacina, A. Maali, *App. Phys. Lett*, **2007**, *90 (11)*, 113512

3. Moharam Habibnejad Korayem and H. Jiryaei Sharahi, *Int J Adv Manuf Technol*, **2011**, *57*,477–489