

**The Structural Hierarchy and Mechanical Properties of Fibrin Described with a
Multi-scale Model**

Nathan E. Hudson

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill
in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the
Department of Physics and Astronomy

Chapel Hill
2011

Approved by:

Professor Richard Superfine

Professor Michael R. Falvo

Professor Susan T. Lord

Professor J. Christopher Clemens

Professor Nikolay V. Dokholyan

© 2011
Nathan E. Hudson
ALL RIGHTS RESERVED

Abstract

Nathan E. Hudson: The Structural Hierarchy and Mechanical Properties of Fibrin
Described with a Multi-scale Model
(Under the Direction of Professors Richard Superfine and Michael R. Falvo)

Fibrin networks form the structural scaffold of blood clots during hemostasis. To survive in the dynamic environment of the vasculature, these networks have a diverse set of mechanical and dynamical properties. In this work, a series of experiments and molecular dynamics simulations bridge the gap between the mechanical properties of fibrin networks and the origins of those properties at the fiber and molecular scales. Mechanical measurements on individual fibrin fibers indicated that fibrin fibers are soft in stretching, strain stiffen, and exhibit an increased stiffness when ligated by transglutaminase FXIIIa. We hypothesized that these properties derived from one particular part of the fibrin molecule, the α C region and developed a mechanical model for the fiber based on the extension of the α C region. Measurements on the recoil dynamics of fibrin fibers indicated that they recoil on μ s timescales and regain their full tension within a few ms of relaxation in agreement with the α C model. In spite of the success of the α C extension model, the fibrin molecule is complex and several other regions including the coiled coil region and the γ -nodule have previously been implicated as potential sources of extension. To test these hypotheses we ran constant force Steered Discrete Molecular Dynamics Simulations on each region of the fibrin molecule. The simulations confirmed that the α C region is the most likely to extend at forces as low as 10pN, but implicated other regions of the molecule as well. This led a more complete

model for the mechanical properties of individual fibrin fibers dubbed the SL α CK model. Finally, to probe how the mechanical properties of individual fibrin fibers affect network strength, 2-D fibrin networks suspended between channels were stretched to failure using and Atomic Force Microscope. The strain of individual fibers in the network was tracked, and it was shown that fibrin network strength is enhanced by the strain stiffening of individual fibrin fibers. This work provides a framework for a predictive model in which the affects from alterations at the molecular level could be observed in the mechanical properties of the higher levels of the hierarchy.

Acknowledgements

This work was made possible through the dedication of many individuals who poured parts of their lives into my own to make me the person and scientist that I am today. I am indebted to Professor Michael Falvo for his guidance and patience as I learned to think independently as a scientist. Mike always has his door open, and I used that access liberally. I am grateful to Professor Rich Superfine for pushing me to stretch myself and explore areas beyond my current level of comfort. Professor E. Timothy O'Brien provided my first training in biological sample preparation and has provided invaluable insight throughout this process. Professor Susan Lord has been a wonderful collaborator and has shown great patience in dealing with us physics types as we try to understand the complexity of biochemistry and fibrin in particular. I've had the pleasure of working with many others in the fibrinogen research group on campus including Barbara Cardinali, Lihong Huang, Lifang Ping, Kellie Beicker, and John Houser who have assisted with experiments and helped me in my understanding. A special thanks goes to Professor Oleg Gorkun who has provided many avenues of research to pursue and reviewed this document at the last second, and to Laurel Averett for allowing me to use her dissertation as a template.

I would be remiss if I did not thank the many undergraduate research students who have aided me along the way. Olamide Olusesi, Kyle Pridgen, Hans Peng, and Travis Hairfield all contributed to my research while performing their own projects. A special thanks to Igal Bucay who finished the recoil data analysis in record time.

Professor Nikolay Dokholyan, Professor Feng Ding and Pradeep Kota all provided their time and expertise in helping me perform the DMD simulations. Without their help chapter 4 would not have happened.

My colleagues in CISMM have provided a wonderful and thought provoking work environment. I am grateful for every time each of them has answered one of my ignorant questions with patience and clarity. The community in the group is rare, I believe, for a science lab; and much credit goes to Rich Superfine for the collaborative environment he created. I am also grateful to my computer science colleagues, especially Professor Russell Taylor for making our lives much easier with automated software.

I am grateful to the NIH and NSF for continued funding throughout my time at UNC. I also would like to thank the NIH/NCRR for funding the Biomedical Analysis and Simulation Supercomputer (BASS) which was used for all simulations described in the text. I also owe a debt of gratitude to Hilary Hicks and Photometrics for their generous loan of the Evolve 128 camera which enabled the analysis of the fiber recoil dynamics

Finally, I want to thank Professor Makis Petratos and Professor Mina Katramatou for encouraging me try experimental research. I want to thank my parents, Martin and Marilyn Hudson for their many prayers, and for their guidance as I grew and developed throughout life. Finally, I want to thank my wife Robin, who has been incredibly supportive throughout this process, never getting angry or frustrated in spite of my many frantic moments leading up to the completion of this document. You inspire me to work hard, love God and become a better man daily. I love you very much.

Preface

Many parts of this work were done in collaboration with other talented scientists. Chapter 2 represents a publication for which I was one of two co-first authors. My friend and collaborator John Houser developed the worm-like chain based model and performed the in-situ stiffening measurements described in the paper. The paper was published previous to writing this thesis with the following citation:

Houser, J.R.*, Hudson, N.E.*, Ping, L., O'Brien III, E.T., Superfine, R., Lord, S.T., Falvo, M.R. Evidence that the α C Region is Origin of Low Modulus, High Extensibility, and Strain Stiffening in Fibrin Fibers. *Biophys J.* 2010 November; 99(9): 3038–3047

Permission to include the article in its entirety in a PhD dissertation was retained from Elsevier Limited (publisher of BPJ).

Chapter 5 of this work represents a publication for which I was the primary first author. The paper was published previous to writing this thesis with the following citation:

Hudson, N.E., Houser, J.R., O'Brien III, E.T., Taylor II, R.M., Superfine, R., Lord, S.T., and Falvo, M.R. Stiffening of Individual Fibrin Fibers Equitably Distributes Strain and Strengthens Networks. *Biophys J.* 2010 April; 98(8): 1632–1640

Permission to include the article in its entirety in a PhD dissertation was retained from Elsevier Limited (publisher of BPJ).

All copyrighted material included in this dissertation is used with permission from the relevant copyright holders.

Table of Contents

List of Tables	xiv
List of Figures	xv
List of Abbreviations and Symbols.....	xviii
Chapter 1. Introduction	1
1.1 Fibrinogen and Fibrin: Structure and Polymerization.....	1
1.1.1 Fibrin(ogen) Molecular Structure	3
1.1.2 The Fibrin Fiber Structure	11
1.2 The Mechanical Properties of Fibrin.....	15
1.3 Mechanical Models for the Extension of Fibrin Fibers within a Network	17
1.4 Significance and Goals of this project	23
Chapter 2. Evidence that the α C Region is the Origin of the Low Modulus, High Extensibility, and Strain Stiffening in Fibrin fibers.....	27
2.1 An Introduction to Fibrin Fiber Mechanics	27
2.2 Methods	28
2.2.1 Structured Surface Preparation	28
2.2.2 Fibrinogen Preparation.....	29
2.2.3 Optical Microscopy.....	30
2.2.4 AFM.....	31
2.2.5 In-situ measurement of fiber Stiffening.....	32
2.3 Results & Discussion	34

2.3.1 Single Fiber Stress vs. Strain	34
2.3.2 In situ measurements.....	35
2.3.3 The α C Stretching Model.....	37
2.3.4 Enthalpic backbone stiffness of the fibrin fiber	44
2.3.5 Does the α C connector region mediate fibrin's initial extensibility?	46
2.3.6 Analysis of the Coiled-Coil (CC) unfolding model	48
2.3.7 Force per Monomer.....	55
2.3.8 Fibrin as Elastomeric Protein.....	58
2.4 Conclusions.....	59
Chapter 3. Fibrin Fiber Recoil Dynamics	61
3.1 Introduction.....	61
3.2 Materials and Methods.....	63
3.2.1 Optical microscopy	63
3.2.2 Discrete Molecular Dynamics.....	63
3.2.3 Fibrin Polymerization	64
3.3 Experimental Results	64
3.4 An α -helix to β -sheet transition	67
3.5 Conclusions.....	70
Chapter 4. Unfolding Energy Barriers of the Coiled coil, γ - γ and α C region of the Fibrin Molecule	73
4.1 Beyond the α C region	73
4.2 Developing a Structural Model for the α C Region of the Fibrin Molecule.....	74
4.2.1 Is the α C region truly unstructured?.....	74

4.2.2 A computationally Identified β helix structure of the fibrin α C region.....	76
4.2.3 Thermal unfolding of the α C region indicates the relative stability of the β helix.....	79
4.2.4 Comparing properties of the α C computational structure with known experimental data.....	80
4.2.5 Conclusions.....	87
4.3 The Unfolding Energy Barrier of the α C Domain.....	88
4.3.1 Methods.....	88
4.3.2 Patterns in Unfolding the α C domain.....	89
4.3.3 Calculating the Unfolding Energy Barrier of the α C Region.....	90
4.4 The Unfolding Energy Barrier of the Coiled Coil Region Probed by DMD.....	91
4.4.1 Methods.....	92
4.4.2 Results.....	93
4.5 Unfolding of the γ Regions of the Fibrin Molecule.....	97
4.5.1 Methods for SDMD simulations of the fibrin γ - γ interface.....	98
4.5.2 Methods for SDMD simulations of the fibrin γ -C terminus unfolding.....	100
4.5.3 γ -C and γ - γ unfolding energy barriers.....	101
4.6 The SL α CK Molecular Model for Fibrin Extension.....	107
Chapter 5. Stiffening of Individual Fibrin Fibers Equitably Distributes Strain and Strengthens Networks.....	113
5.1 The Mechanical Properties of Fibrin in the Context of Biopolymer Networks...	113
5.2 Materials and Methods.....	115
5.2.1 Fluorescent Microscopy Applied to Fibrin Fibers.....	115
5.2.2 Fibrinogen Preparation.....	116

5.2.3 Growing Fibrin Networks on Structured Surfaces.....	116
5.2.4 Stretching Networks with an AFM	117
5.2.5 Network Strain measurements	119
5.2.6 Using Webslinger to Simulate Network Stretching.....	120
5.3 Results and Discussion	121
5.3.1 Results from Stretching Fibrin Networks	121
5.3.2 Network Modeling.....	123
5.3.3 Network Strain Distribution Narrowing	128
5.3.4 Discussion	135
5.4 Conclusion	140
Chapter 6. Summary and Future Directions	142
6.1 Conclusions and Physiological Relevance	142
6.1.1 The Material Properties of Fibrin	142
6.1.2 Physiological Relevance	146
6.2 Future Directions	148
6.2.1 Measuring the mechanical properties of recombinant fibrinogen variants	148
6.2.2 Cryo-Electron Microscopy of Fibrin Fibers and Fibrin variants	151
6.2.3 FIONA/FRET Studies of Single Molecules within Fibrin Fibers	155
6.2.4 Single Molecule AFM Studies of Fibrin(ogen) Dynamics	158
6.2.5 Developing a Coarse-Grained Model for the Fibrin Fiber.....	160
6.2.6 Measuring the effects of other blood coagulation factors on Fibrin's mechanical properties	162
6.2.7 Testing for and α -helix to β -sheet Transition under stretching.....	163

Appendix A: An Overview of Molecular Dynamics Methods	165
A.1 Molecular Dynamics: A Primer	165
A.1.1 An Brief History to Molecular Dynamics	165
A.1.2 The Basic Molecular Dynamics Mechanism	166
A.1.3 Potential Energy or Force Field calculations in MD.....	167
A.2 The foundation of the MD method: statistical mechanics	171
A.2.1 A review of statistical ensembles	171
A.2.2 The ergodic hypothesis	174
A.2.3 Statistical ensembles in MD.....	176
A.3 Generating a Canonical Ensemble	177
A.3.1 Woodcock Thermostat	177
A.3.2 Anderson Thermostat.....	178
A.3.3 Berendsen Thermostat.....	178
A.3.4 Nose-Hoover Thermostat	179
A.3.5 Langevin dynamics	180
A.4 Molecular Dynamics of “Large” Systems	181
A.5 All Atom Discrete Molecular Dynamics (AADMD).....	183
A.5.1 DMD Equations of Motion	183
A.5.2 AADMD Potential Energy Terms.....	185
A.6 Calculating free energies from MD simulations	188
A.6.1 The Weighted Histogram Analysis Method (WHAM).....	188
A.6.2 Steered Molecular Dynamics (SMD) to calculate unfolding energy barriers.....	194

Appendix B: Optical Microscopy Techniques.....	197
Appendix C: Atomic Force Microscopy.....	206
Bibliography	212

List of Tables

Table 1.1 A comparison of the potential models for fibrin mechanical properties.....	25
Table 5.2 A comparison of network properties of 18 networks stretched in these experiments.....	144
Table 6.2 A comparison of the potential models for fibrin mechanical properties with the SL α CK.....	148

List of Figures

Figure 1.1 The Fibrin Hierarchy	3
Figure 1.2 The Fibrinogen molecule.....	7
Figure 1.3 Fibrin knob:hole interactions and Protofibril Formation.....	10
Figure 1.4 Fibrin Fiber Banding	14
Figure 1.5 Coiled-coil (CC) extension model.....	18
Figure 1.6 The Transverse Cross-linking (TCL) Extension Model.....	19
Figure 1.7 The α C extension model.....	21
Figure 2.1 Set up for single fiber and network stretching experiments	29
Figure 2.2 Stress vs. Strain plots of individual fibrin fibers.	35
Figure 2.3 Average tangent modulus at discrete strains for ligated (N=14) and un-ligated (N=14) fibers.	35
Figure 2.4 Left: Western blot of fibrin showing Right: In-situ fiber stiffness measurements. Fiber stiffness before (blue) and after (red) FXIII ligation.	37
Figure 2.5 Fibrin structure and corresponding mechanical model	39
Figure 2.6 Force vs. Strain data	43
Figure 2.7 Several representative experimentally obtained force curves	46
Figure 2.8 Single fiber force strain plot for coiled-coil model.	51
Figure 2.9 Fitting Single Fiber force vs. extension data with the Coiled-coil model.	54
Figure 2.10 Force per molecule (FPM) plots for one fiber with three different packing assumptions.....	57
Figure 2.11 Force per monomer as a function of fiber strain	58
Figure 3.1 Fibrin Fiber Recoil Models Based on the Mechanism of Extension.....	63
Figure 3.2 Experimental Setup.	65

Figure 3.3 Fibrin Recoil Dynamics Captured with High Speed Fluorescence Microscopy	66
Figure 3.4 β -sheet Relaxation	70
Figure 4.1 Structure of the fibrin α C region	78
Figure 4.2 Specific Heat of the α C Structure.....	80
Figure 4.3 β -hairpin vs. β -helix	82
Figure 4.4 Unfolding the α C region at 25pN.....	90
Figure 4.5 The Structure and Force Attachment Points for the DMD Simulations.....	93
Figure 4.6 Unfolding Energy Barrier.....	94
Figure 4.7 Coiled coil unfolded distance vs. Force	96
Figure 4.8 α -helix to β -sheet transition.....	97
Figure 4.9 Unhinging of th γ - γ connection under force.....	102
Figure 4.10 Step Heights between γ -C unfolding events.....	103
Figure 4.11 Forced Unfolding of the γ -C domain.....	106
Figure 4.12 The SL α CK model for Fibrin Fiber Extension.....	112
Figure 5.1 Experimental setup	119
Figure 5.2 Network Strain Measurements.	123
Figure 5.3 Experiment vs. Simulations.....	125
Figure 5.4 Single fibrin fiber force curve	127
Figure 5.5 Ligated network strain distributions	130
Figure 5.6 Linear springs with different stiffness.....	133
Figure 5.7 Unligated network strain distributions	135
Figure 5.8 The Network Strengthening Mechanism of Strain Stiffening.....	137
Figure 5.9 Force vs. Displacement of a fibrin network	138

Figure 6.1 Segmenting protofibrils within a Cryo-EM fibrin fiber section.....	154
Figure 6.2 Successive Levels of Coarse Graining	161
Figure 6.3 Fiber Tension Loss	163
Figure A.1 DMD Potential Energy Terms	188
Figure B.1 Epifluorescence vs. Transmission Illumination.....	199
Figure B.2 Snell’s Law	204
Figure C.1 Basic AFM Setup.....	207
Figure C.2 Normal and Lateral Cantilever deflections.....	209

List of Abbreviations and Symbols

~	approximately
>	greater than
<	less than
%	percentages
-	negative
+	positive
±	statistical uncertainty
2-D	two dimensions
α	fibrin alpha chain after fibrinopeptide A cleavage
α C	C-terminal region of the fibrino(ogen) α chain
α -helix	alpha-helix, a type of coiled protein structure
β	fibrin beta chain after fibrinopeptide B cleavage
β -sheet	beta-sheet, a type of folded protein structure
β -C or β -nodule	C-terminal region of the fibrino(ogen) β chain
γ	fibrin(ogen) gamma chain
γ -C or γ -nodule	C-terminal region of the fibrino(ogen) γ chain
γ - γ	interface between the γ -nodules of two adjacent fibrin molecules
Δ F	change in Force
Δ G _B	unfolding energy barrier
Δ l or Δ L	change in length
Δ t	change in time
Δ z	change in protein length

ε	strain
θ_C	Critical angle
θ_I	Incident angle
θ_T	Transmitted angle
κ	effective spring constant
λ	wavelength
τ	Relaxation time constant
τ_B	Unfolding Time
τ_o	escape attempt frequency
μm	micrometer
μs	microsecond
\AA	angstrom
A	cross-sectional area
AADMD	All Atom Discrete Molecular Dynamics
A α	fibrinogen α -chain before fibrinopeptide A cleavage
A α 251	recombinant fibrinogen truncated at α -chain residue 251
A:a	fibrin knob A bound to hole a
AFM	Atomic Force Microscope
AP	Attachment Point
ASA	Accessible Surface Area
B β	fibrinogen β -chain before fibrinopeptide B cleavage
B:b	fibrin knob B bound to hole b
BASS	Biomedical Analysis and Simulation Supercomputer

BPTI	Bovine Pancreatic Trypsin Inhibitor
Ca^{2+}	Calcium Ion
CaCl_2	Calcium Chloride
CC	Coiled-coil extension model
CCD	charge-coupled device
CD	Circular Dichroism
CHO	Chinese Hamster Ovary
Cryo-EM	Cryogenic Electron Microscopy
C_v	specific heat
Cys	amino acid Cystein
D_{fib}	fiber diameter
dE	incremental change in Energy
D_{mol}	fibrin molecule diameter
$D_{\text{mol}}^{\text{eff}}$	effective molecule diameter
DH	Donor Hydrogen
DMD	Discrete Molecular Dynamics
DNA	DeoxyriboNucleic Acid
d_o	distance from lens to sample
dR	incremental change in length
E	Young's Modulus or energy
EDTA	ethylenediaminetetraacetic acid
EM	Electron Microscope
exp	exponential function

Eqn	equation
F	Force
F_c	critical force
FEP	Fixed End Point
F_{fib}	Force on the fiber
FIONA	Fluorescence Imaging with One-Nanometer Accuracy
FpA	fibrinopeptide A
FpB	fibrinopeptide B
FPM	Force Per Molecule
fps	frames per second, a camera frame rate
f_o	focal length
FRET	Fluorescence Resonance Energy Transfer
FXIII	fibrin stabilizing factor thirteen
FXIIIa	activated fibrin stabilizing factor thirteen
G	Shear Modulus
GB	Generalized Born
Gln	amino acid Glutamine
Gly	amino acid Glycine
GPa	GigaPascal, A measure of Stress
GHRP	a ligand of the sequence Glycine, Histidine, Arginine and Proline
GPRP	a ligand of the sequence Glycine, Proline, Arginine and Proline
GPRV	a ligand of the sequence Glycine, Histidine, Arginine, and Valine
H	Helmholtz Free Energy
HB	Hydrogen Bond

HBS	HEPES Buffered Saline
HC	Human-Chicken hybrid fibrinogen
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His	amino acid Histidine
HPLC	High Performance Liquid Chromatography
Hz	Hertz, a measure of frequency
I-TASSER	Iterative Threading ASSEMBLY Refinement
K	Kelvin
k	spring constant
$k_B T$	Boltzmann's constant times Temperature, a measure of energy
kDa	kiloDalton, a measure of mass
K_o	protein backbone stiffness
L	Length
LJ	Leonard Jones
L_o	Initial Length
L_c or l_c	Contour Length
L_f	folded length
L_p	persistence Length
L_u	unfolded length
Lys	amino acid Lysine
MARTINI	a coarse grained MD force field
MD	Molecular Dynamics
mg	milligram
mL	milliLiter
mM	milliMolar

MPa	MegaPascal, a measure of stress
ms	millisecond
N	Number of trials or alternatively Number of Proteins
NA	Numerical Aperture
NaCl	Sodium Chloride
NALMS	NAnometer-Localized Multiple Single-molecule
NHT	Nose-Hoover Thermostat
NIH	National Institutes of Health
nm	nanometer, a measure of length
NMR	Nuclear Magnetic Resonance
N_{mol}	number of molecules
nN	nanoNewtons
ns	nanoseconds
OLS	Optical Lever Sensitivity
PDB	Protein Database
PDMS	Poly Di-Methyl Siloxane
pH	a measure of acidity
pM	picoMolar
PMF	Potential of Mean Force
pN	piconewtons, a measure of force
PP	Pull Point
PSF	Point Spread Function
Pro	amino acid Proline
QPD	Quadrant Photo Diode
R	position

REXMD	Replica Exchange Molecular Dynamics
s	seconds
s^{-1}	a shear rate or alternatively a strain rate
S	Entropy
SAXS	Small Angle X-ray Scattering
SDMD	Steered Discrete Molecular Dynamics
SDS	Sodium Dodecyl Sulfate
SDS-PAGE	Sodium Dodecyl Sulfate PolyAcrylamide Gel Electrophoresis
SEM	Scanning Electron Microscope
SL α CK	a model for fibrin fiber extension
SMD	Steered Molecular Dynamics
S_{L-R}	Left-Right Sensitivity
SS	Structured Surface
S_{T-B}	Top-Bottom Sensitivity
T	Temperature
TEM	Transmission Electron Microscopy
TCL	Transverse Cross-Linking extension model
TIRF	Total Internal Reflection
T_m	melting temperature
tPa	tissue Plasminogen Activator
U	Unit
UV	Ultra Violet
v	velocity
WAXS	Wide Angle X-ray Scattering

WHAM	Weighted Histogram Analysis Method
WLC	Worm-Like Chain, a polymer physics model
x_u	unfolding length
Z	Partition Function

Chapter 1. Introduction

1.1 Fibrinogen and Fibrin: Structure and Polymerization

The cardiovascular system consists of a network of arteries, capillaries and veins (the vasculature) designed to distribute oxygen and nutrients throughout the body. The vasculature can be a very dynamic environment with the heart pumping blood at 1Hz, and shear rates ranging from close to 0s^{-1} in near static conditions up to $100,000\text{s}^{-1}$ in areas of stenosis.(1) In spite of the wide range of environmental conditions, the body must have a functioning wound healing system that can withstand the high shear rates, without plugging the vasculature at low shear rates. The physiological ability to accommodate this dynamic range of behaviors is achieved in large part through the properties of the blood protein fibrinogen and its activated counterpart fibrin.

Fibrin and fibrinogen play a complex and important role in the process of hemostasis, the stopping of blood flow near a wound. The soluble fibrinogen molecule is the third most prevalent protein in the blood plasma, and acts as a molecular link to mediate platelet aggregation at the sight of a wound.(2) Subsequently, fibrinogen is converted into insoluble fibrin, which polymerizes sequentially, through specific interactions into larger linear fibers and eventually networks which form the structural backbone of blood clots. (3-7) Formation of these networks, which comprise only 0.25% of the volume of a physiological clot, is the last step of the blood coagulation cascade, the chemical process of blood clot formation.(8) After the wound healing process completes, an additional protein called plasmin can cleave the fibrin molecule at specific sites in a

process called fibrinolysis and dissolve the fibrin network, allowing blood to circulate freely again.(9, 10) Proper wound healing requires that the clot polymerize properly, withstand any shear stresses without failing, and dissolve rapidly after the healing is complete.

Clinically, patients with severe bleeding disorders can often have the roots of the diatheses traced back to an abnormality in fibrinogen (dysfibrinogenemias). Fibrinogen Gaurenas for example truncates 143 amino acids short in the C-terminus of its α -chain. Clots formed from this fibrinogen variant were three times stiffer than those of normal fibrinogen, while at the same time consisted of thinner and more highly branched fibers.(11) Thus, an abnormality at the molecular level led to morphology changes on both the fiber and network levels resulting in stiffer clots and severe bleeding. This example highlights the importance of having a model for fibrin polymerization, fiber and network structure, and the mechanical properties arising from these organizations. The goal of this dissertation is to generate a multi-scale model of the mechanical and dynamical properties of fibrin based on the structural hierarchy of the material. This model highlights the unique material properties of fibrin that allow it to function as a sealant within the vasculature while suggesting regions of the molecule that could be altered through protein design to tune the properties at higher levels in the hierarchy.

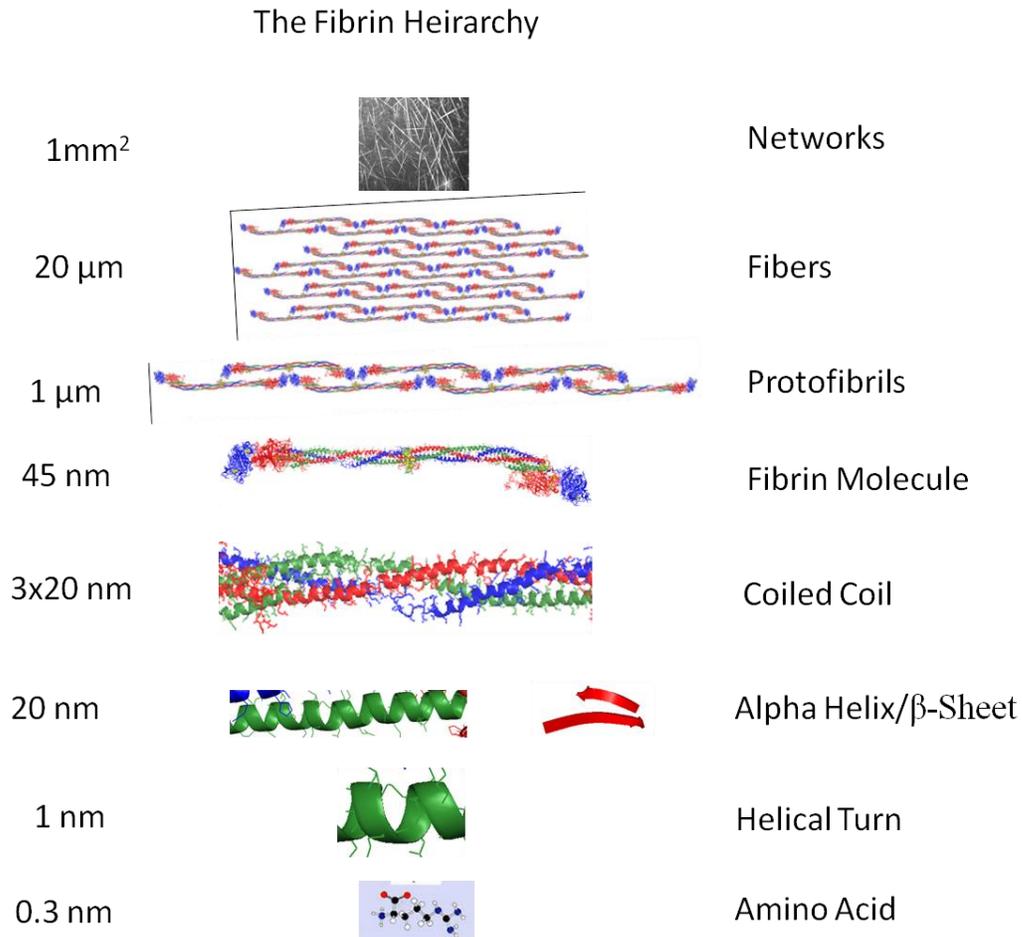


Figure 1.1 The Fibrin Hierarchy: Fibrin, as a protein based biopolymer, has a structural hierarchy spanning many orders of spatial magnitude. The lowest level is a chain of amino acids; these fold into secondary structures such as helical turns. These secondary structures form tertiary structures containing α -helices and β -sheets. Finally these structures combine to form the whole molecule. These molecules then polymerize into protofibrils, fibers and eventually networks. A complete model of the fibrin hierarchy, will be able to predict how a mutation in a particular amino acid will affect the mechanics of fibrin networks.

1.1.1 Fibrin(ogen) Molecular Structure

Human fibrinogen is a soluble 340 kDa glycoprotein found in the blood stream. The structure of the molecule can be observed in Figure 1.2. It is assembled as a dimer of two subunits of three amino acid chains (called $A\alpha$, $B\beta$, and γ) each.(6) The chains consist of 610, 461 and 411 amino acid residues respectively. Amino acid chains, or polypeptides, are structurally polar, with an amino end (N-terminus) and a carboxyl end

(C-terminus). Within the fibrinogen molecule, all six chains are oriented so that their N-termini are located in the central region (historically called the E region) of the molecule, and their C-termini are located in the periphery (See Figure 1.2). From both sides of the central nodule, three chains extend into coiled-coils that terminate in the two distal nodules. The C-terminal segments of the B β and γ chains fold independently to form the compact, globular β - and γ -nodules, respectively (the last part of the coiled-coil combined with the β -nodule and γ -nodule are referred to as the D region). The C-terminal segment of the A α chain goes through the D region, and folds back to form a brief fourth α -helix, but beyond this fourth coil the structure of the C-terminus of the A α chain (the α C region) is not visible in the X-ray structure indicating an inhering flexibility in this region.(12) The α C region is often grouped into two sub regions, the α C connector (221-391) and the α C domain (392-610). The α C connector region is thought to be unstructured while there exists some debate about the structure of the α C domain.(13, 14)

The molecule is stabilized by 29 di-sulfide bonds, pictured in yellow in Figure 1.2. Five of the di-sulfides link the two subunits of the molecule together near the N-termini of the A α , B β , and γ chains (A α ₁Cys28- A α ₂Cys28, γ ₁Cys8- γ ₂Cys9, γ ₁Cys9- γ ₂Cys8, A α ₁Cys36- B β ₂Cys65, and A α ₂Cys36- B β ₁Cys65). (15-17) The coiled coil region of the molecule is book-ended by a ring of interchain disulfides: α Cys45- γ Cys23, α Cys49- β Cys76, and β Cys80- γ Cys19 on the N-terminal side of the coiled coil and α Cys161- γ Cys135, α Cys165- β Cys193, and γ Cys139- β Cys197 in the C-terminal region.(18, 19) The rest of the disulfides are intrachain in the C-terminal portions of the molecule: the α C region is connected via A α Cys442-A α Cys472; the β C region via B β Cys201- B β Cys286, B β Cys211-B β Cys240, and B β Cys394- B β Cys407; and the γ -C

region via γ Cys153- γ Cys182 and γ Cys326- γ Cys339.(12, 19) Taken as a whole, the disulfide bonds make portions of the molecule much more mechanically stable to resist forced unfolding as will be discussed more in chapter 3. From one γ -nodule to the other, the structured part of the molecule measures 45nm, which gives it a half-distance of 22.5nm; and the coiled coil regions of the molecule measure 17nm each.

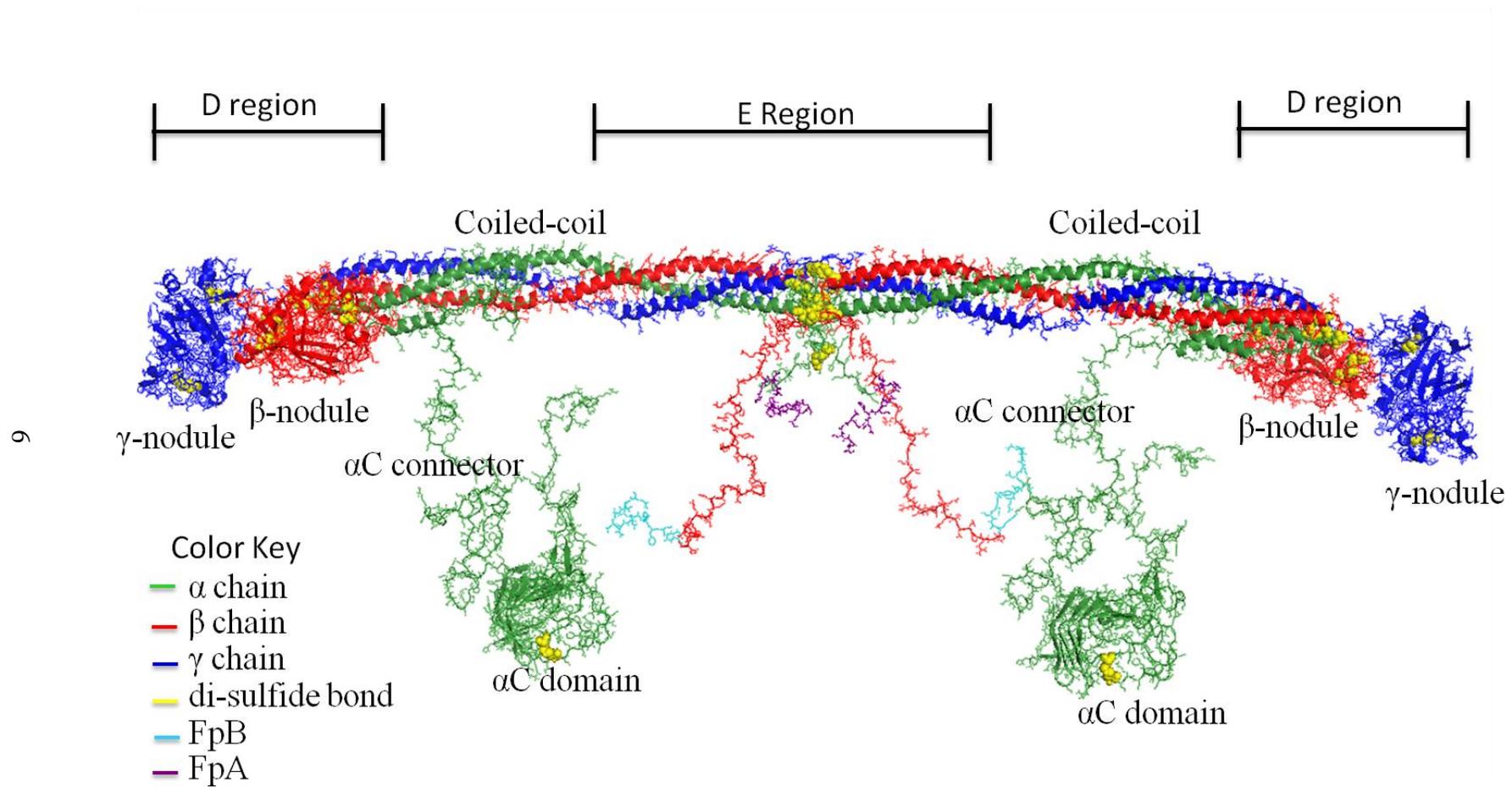


Figure 1.2 The Fibrinogen molecule: The fibrinogen molecule above was created using crystal structure 3GHG(12), combined with Discrete Molecular Dynamics Methods (DMD, see Appendix A) to fill in amino acids A α 1-26, A α 201-610 and B β 1-57. This image is intended to highlight various parts of the fibrinogen molecule, however it should not be treated as an exact representation of the structure of the molecule in solution or even within a fiber as it has inherent flexibility not represented in this view. A α 201-610 were modeled from homology modeling as described in section 3.3. The α C connector region is largely unstructured and links to the bulk of the model, while the α C domain is seen as a β -helix. The A α chain is colored in green, the B β chain is colored in red, and the γ chain is colored in blue. The 29 cysteins participating in di-sulfide bonds are represented as spheres and colored in yellow. FpA is colored in purple and FpB is colored in cyan, both are removed upon the conversion from fibrinogen to fibrin. It should be noted that the location of the α C regions is not known. Some data indicates that it is packed against the bulk of the molecule or attached to FpB. Here we have left it separate from the molecule as it would likely be in fibrin. The image was generated in PYMOL (The PyMOL Molecular Graphics System, Version 1.4, Schrödinger, LLC.).

Fibrinogen is converted into the insoluble fibrin when the enzyme thrombin cleaves fibrinogen at 4 sites: after residue 16 in each A α chain releasing fibrinopeptide A (FpA), and after residue 14 in the B β chain releasing fibrinopeptide B (FpB).(6) After FpA and FpB release, the fibrin chains are known simply as α , β , and γ , and exposed binding sites known respectively as knob A and knob B are active. Knob A consists of the α 17-19 Gly-Pro-Arg residues, while knob B consists of β 15-17 Gly-His-Arg residues.(20) The A and B knobs bind to complimentary “holes” (hole a and hole b respectively) in the peripheral D domains of nearby fibrin molecules; forming A:a and B:b interactions mediated primarily through electrostatic salt bridges and hydrogen bond interactions.(19, 21) These knob-hole interactions, allow for the formation of a half-staggered fibrin molecule arrangement in which the knobs in the E region of one molecule, bind to holes in two nearby molecules. (See Figure 1.3a) The half-staggered polymerization allows additional molecules to continuously bind to the unbound knobs at the end of the overlapping molecule. As this occurs, the chain extends longitudinally into a double-stranded fibrin polymer that has been termed a protofibril (Figure 1.3b).(22) Finally, the protofibrils bundle together laterally to form fibers. This process may be mediated by the B:b interactions, which are believed to form after the A:a interactions, or by the fibrin α C region.(23-26)

Simultaneously with FpA and FpB cleavage, thrombin also catalyzes the activation of FXIII to FXIIIa, which is bound to fibrinogen and fibrin. FXIIIa, also called fibrin stabilizing factor is a transglutaminase that catalyzes the formation of isopeptide bonds (intermittently known as ligation or cross-linking) between γ ₁Lys406- γ ₂Gln398 or γ ₂Gln399 at the C-terminal γ -nodules, forming longitudinal γ - γ dimers (See Figure 1.3a,

and note the nomenclature, the interface between two adjacent γ -nodules will be referred to as the γ - γ interface).(27) These interactions, serve to strengthen the protofibrils. It has been proposed that rather than γ - γ dimers, FXIIIa serves to cross-link latitudinally, (See Figure 1.6) but that is the minority view and seems to conflict with mechanical data.(28) Additionally, FXIIIa, can form bonds and between α C regions of several fibrin molecules, forming networks termed α polymers. (29, 30)

While this is the standard model of polymerization, it should be noted that there has been controversy even recently as to the exact molecular packing of protofibrils and the type of interactions and locations of those interactions within protofibrils.(7, 31-33) In developing our mechanical model, we have assumed the standard model of polymerization, but it is possible that future insights will require an update to the model including the alternative packing arrangements within the fiber.

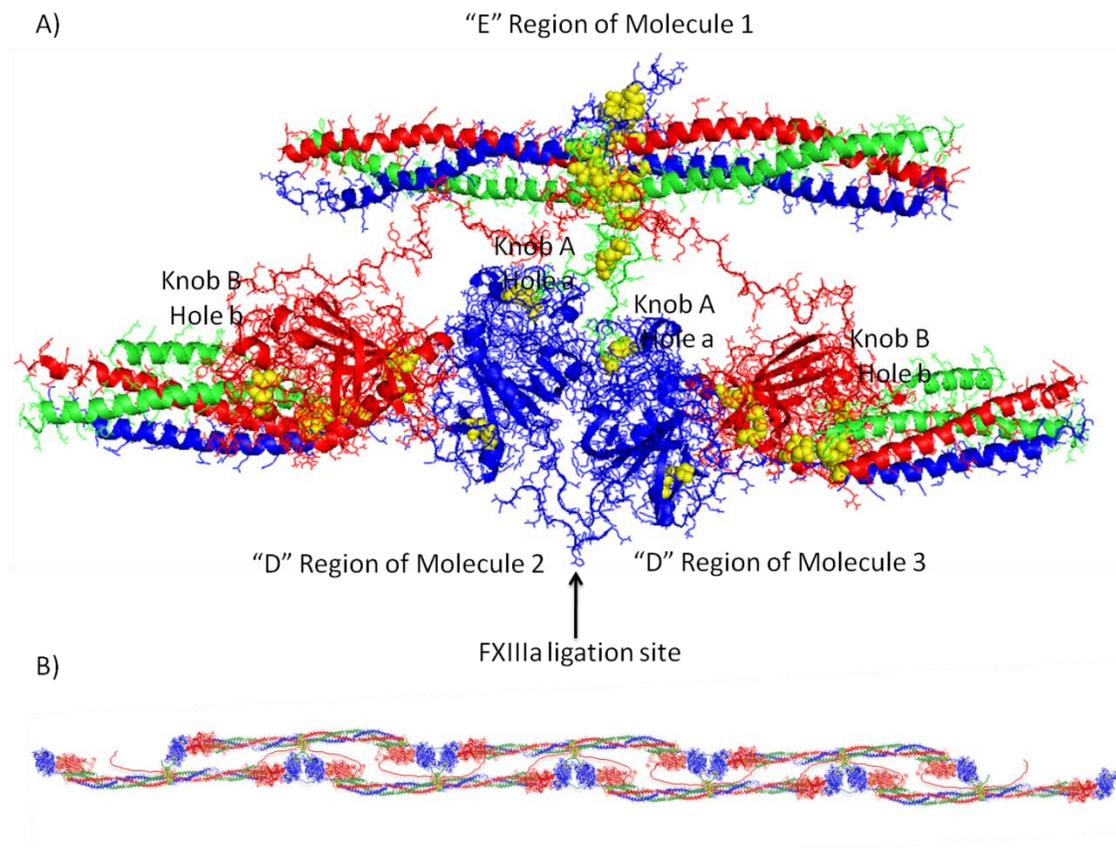


Figure 1.3 Fibrin knob:hole interactions and Protofibril Formation: Fibrin polymerization proceeds through a series of knob-hole interactions. Knob A is exposed by the thrombin cleavage of FpA and knob B is exposed by thrombin cleavage of FpB (See Figure 1.2 and text). Both A-knobs bind to corresponding "hole a's" located in the C-terminal region of the γ chain of adjacent molecules. In the figure, the molecule with the exposed knobs (green chains) is labeled as molecule 1, and the molecules with the holes are molecules 2 and 3. Similarly, the B-knobs (red chains) bind to "hole b's" in the C-terminal region of the β chain. The abutment of the D region of molecule with the D region of molecule 3 will be referred to in this document as the γ - γ interface. Image (A) was generated in PYMOL (The PyMOL Molecular Graphics System, Version 1.4, Schrödinger, LLC.). The original structure was taken from the D-D crystal structure published as 1FZC in the protein data base.⁽³⁴⁾ The FXIIIa ligation site located in amino acids γ 398-411 was not present in the crystal structure, and was built into the model using DMD. The A α chain is colored in green, the B β chain is colored in red, the γ chain is colored in blue, and the di-sulfide bonds are represented as yellow spheres. The knob:hole interactions allow the formation of a double-stranded protofibril as seen in image (B). The α C regions were omitted from the protofibril image for simplicity.

1.1.2 The Fibrin Fiber Structure

In order to understand the mechanical properties of fibrin fibers, one first must understand how the fibrin molecules are packed within the fibers. This is crucial because a fiber where protofibrils span the entire length of the fiber and are packed tightly in parallel across the width of the fiber has the potential to behave much differently than a fiber consisting of short protofibrils packed loosely within the fiber and held together by tethers such as the α C region.

Fibrin fibers were first imaged using electron microscopy by Hawn and Porter in 1947(35). They reported seeing fiber bundles, and a banding pattern of 25nm. They also reported that the banding pattern seemed to vary with fibers presumably under tension. C.E. Hall continued this work in 1949, where he measured a banding pattern on average of 22.7nm (although the actual banding ranged from 19-27nm) down the length of the fiber.(36, 37) This work was followed up by C. Cohen in 1966, who showed that fibrinogen, under certain ionic conditions, can also form structures with 22.6nm banding(38). The banding patterns indicate that there is a lateral registry between protofibrils in the fiber (See Figure 1.4). Based on these measurements, J Hermans in 1979 proposed that fibrin fibers have a crystalline structure, and proposed four possible packing orientations(39). In 1980, R Hantgan and J Hermans did a comparison of EM structures of fibrin at different stages of polymerization and compared those images to light scattering data(40). They saw that at many stages in polymerization, the fibers appear like loose bundles. In 1981, Weisel and Cohen used recent x-ray data to suggest that the fibrin monomer was more complicated than a simple dumb-bell shape, and had additional outer units, corresponding to the sub-banding pattern of 7-9nm often observed

between the 22.5nm bands in EM images(41). Voter and Erickson then used Transmission Electron Microscopy (TEM) sectioning and Small Angle X-ray Scattering (SASX) to look for crystallinity within the lateral cross-section of a fiber(42). They concluded that there did not seem to be any repeating pattern within the fiber cross-sections. In the same year, using refractive index matching, Voter and Erickson reported that fibers consist of 78% water(43). Also in 1986, Weisel attempted a variety of staining techniques with TEM imaging to look for lateral order. He noted that the banding in the fibers exhibited a saw-toothed pattern at the edges of the fibers. He reported seeing “diffuse equatorial reflections” in the lateral diffraction pattern across the edges of the fibers with a 4nm spacing using most stains, and in 6/300 fibers he reported a strong equatorial reflections with 19.5nm spacing(44). A year later using a “cryo-drying” technique with negative staining, Weisel reported seeing a twisting of the fibers with a pitch of 1.9 μ m, presumably coming from protofibrils wrapping around the fiber, limiting radial growth.(45) In 1990, negatively stained images from the early stages of fibrin polymerization by Medved and Weisel showed that protofibrils themselves are twisted.(46) These observations of twisting prompted the hypothesis that fibrin fibers form under tension because protofibrils must stretch to stay in registry while wrapping around the fiber.(45) In 2000, Yang and Doolittle, based on the first fibrinogen molecule crystal structure, proposed a packing arrangement within a fiber. The structure consisted of a way to arrange protofibrils within a 19x19x46nm unit cell within the fibers allowing lateral interactions between the β and γ domains of individual molecules within the protofibrils(47). In 2003, Caracciolo and Arcovito performed Energy Dispersive X-Ray Diffraction (EDXD) studies on hydrated fibrin fibers. These studies indicated two

diffraction peaks, one at 22.7nm and one at 18.4nm which they claimed agrees with the unit cell model of Yang(48). The Yang model is currently the generally accepted model for the packing of the fibrin fiber, however, it should be noted that the model does not account for the inherent flexibility of the fibrin molecule due to its crystalline basis.

In spite of this progress, there are still many unresolved questions about the fibrin structure including: 1) What is the length of the protofibrils in fibers? 2) Where does the water fit into the structure? 3) How structurally do larger molecules like FXIIIa fit into to fiber, while still allowing for lateral registry? 4) Where do the α C domains reside within the structure? 5) What is the native persistence length of a protofibril within the fiber 6) Where are the locations of the native contacts between protofibrils? 7) Some fibers do not readily exhibit banding in EM images, are these fibers structurally different? 8) If the fibers are crystalline, how is this broken at branch points?

Within this discussion of fibrin polymerization and structure, it may be important to note that in addition to fibers, fibrin has recently been reported to form molecularly thin 2-D sheets, and even bubbles (unpublished data from our lab).(49, 50) Sheets polymerize under static conditions on surfaces, and under flow conditions within a certain phase space of shear rate, and thrombin concentration.(50) Sheets appear continuous at magnifications of 300,000, act as a reversibly elastic membranes, and form under tension.(49) These results are difficult to reconcile with the standard model of fibrin polymerization and fiber packing as described above. It is possible that the sheets are composed of protofibrils laid flat in 2-D registry, but an alternative mechanism of fibrin polymerization could play a role in sheet formation as well. Other models of polymerization have included single stranded “protofibrils” arising from an interlacing

molecule assembly and “protofibrils” by end to end oligomerization.(33, 51) In this work, we have assumed the “standard” polymerization model with double stranded protofibrils, but given fibrin’s ability to form sheets, future models may need to include additional molecular interactions or polymerization morphologies not present in the standard model.

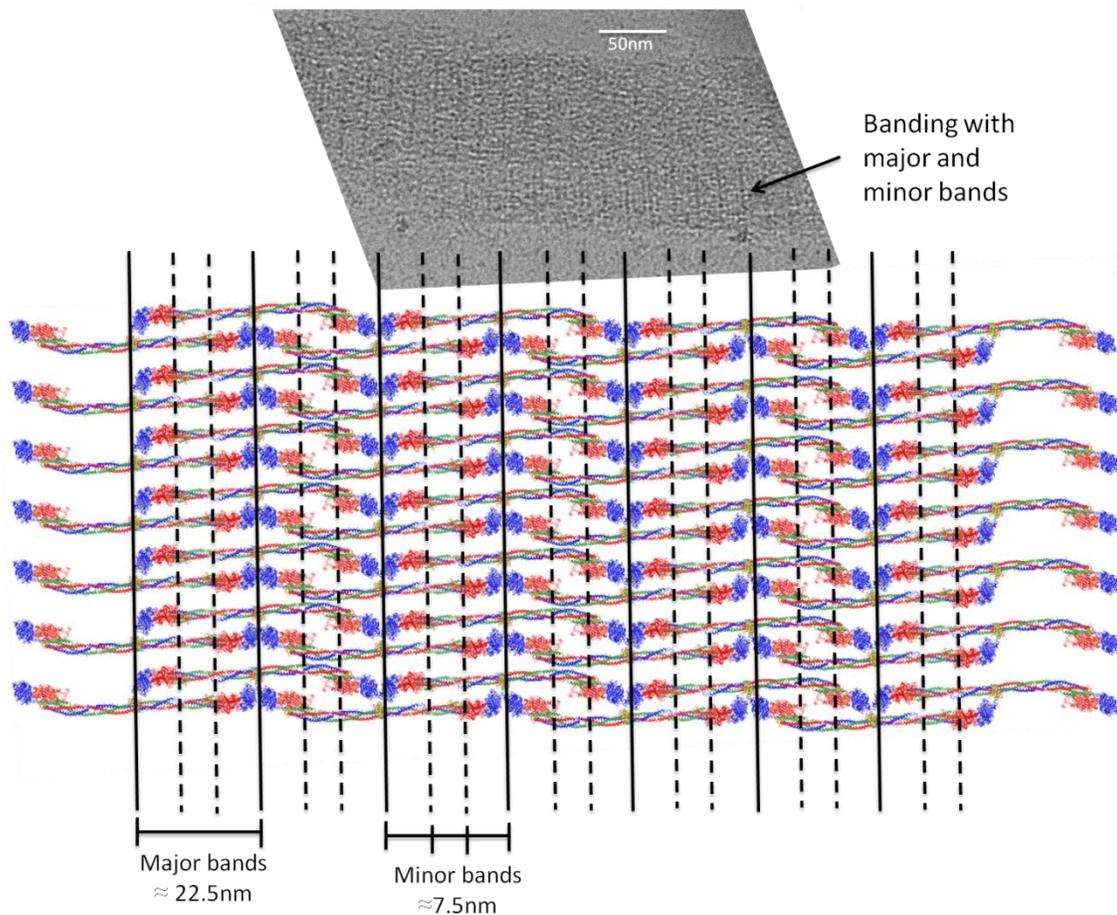


Figure 1.4 Fibrin Fiber Banding: Many EM images, including the Cryo-EM image above, indicate a longitudinal banding pattern within fibrin fibers. The banding pattern indicates a lateral registry between protofibrils. The primary banding distance is approximately 22.5nm, or half the length of the fibrin molecule. Minor bands exist between the major bands with a spacing of approximately 7.5nm. The above diagram gives the current view of how lateral registry between protofibrils gives rise to the observed banding. The mechanism of lateral registry between protofibrils is not well understood. The α C regions are not included in the figure.

1.2 The Mechanical Properties of Fibrin

Because fibrin networks provide the structural support for blood clots that must withstand the shear rates of up to $100,000\text{s}^{-1}$ in the vasculature, they perform an inherently mechanical and dynamical function.⁽¹⁾ The emergent mechanical properties of a fibrin network are a function of the mechanical properties of the components of each hierarchical level – protein monomer, protofibril, fiber, and network – and the architectural transitions between one level and the next – monomer-monomer interactions, lateral aggregation, and branching. In order to have an understanding of the emergent properties of the network, the mechanical properties and architecture of the lower levels of the hierarchy must be fully characterized.

The macroscopic mechanical properties of the highest rung in this structural hierarchy, fibrin gels, have been studied for decades going back to the pioneering work of John Ferry (52-54) and others (55-57). These studies showed that fibrin networks exhibit highly non-linear elasticity manifested in strain stiffening behavior (58-60) and negative normal stress (61). Macroscopic studies have also long established the importance of FXIIIa-induced ligation in stabilizing clots (5). Ligated clots show a much higher storage modulus, lower loss modulus and are more resistant to lysis by plasmin (5, 62). Additional rheological studies of whole clots and networks have focused on the correlation between clot structure (i.e., thickness of fibers, branch point density, and porosity) and viscoelastic dynamics and mechanics.^(53, 54, 63)

While the mechanical properties of fibrin gels have been studied since the 1940's, the mechanical properties of individual fibers have only received attention in the last ten years. Collet et al. used laser tweezers to measure the bending modulus of individual

fibrin fibers within a clot.(64) They reported that unligated fibrin fibers have an elastic modulus of 1-3MPa, while FXIIIa ligated fibers had an elastic modulus of 10-20MPa. AFM measurements on fibrin fibers suspended between microchannels indicated that fibrin fibers are among the most extensible in nature, with uncross-linked fibers being able to extend up to 3.3x their original length before failing, ligated fibers extending to 4.3x their original length before failure, and in both cases much of the extension was reversible.(65, 66) A recent paper by Liu et al. indicated that fully ligated (γ - γ and α C- α C) fibers are in fact less extensible and less elastic than unligated fibers, while partially ligated (only γ - γ cross-links) fibers are more extensible than unligated fibers.(67)

To understand the underlying mechanism behind these single fiber behaviors, several experiments have measured the mechanical properties of various parts of the fibrin molecule. Two groups have claimed to measure the stretching of the fibrinogen coiled-coil region. Brown et al. linked fibrinogen molecules end to end with FXIIIa to form fibrinogen oligomers. They reported a saw-tooth like unfolding pattern with an average force of 94pN and an average extension of 23nm. Lim et al. pulled on individual fibrinogen molecules, and measured a plateau in the force extension trace of 15nm at a force of 60pN and inferred that this indicated coiled coil unfolding. Additionally they pulled on fibrin protofibrils, showing a stepwise unfolding at forces around 1000pN, and a partial refolding of the protofibrils on the timescales of a few seconds.(68)

The strength of the knob:hole interactions has been probed by both optical tweezers spectroscopy and atomic force microscopy. In two seminal papers, Litvinov et al showed that the A:a interaction is critical for fibrin polymerization and can withstand forces of up to 120pN, (69) while the B:b interaction can withstand forces of 15-20pN

before failing.(70) This work was followed up in a beautiful study by Averett et al. showing that the A:a unbinding is actually a multi-step process, with a characteristic profile of four force peaks with two distinct populations of specific forces: $110 \pm 34\text{pN}$ and $224 \pm 31\text{pN}$. The four force peaks, seem to correspond to three unfolding events, corresponding to extension of the fibrin γ -C domain by extensions of 11nm, 9nm, and 18nm, before the final unbinding of the A-knob at the fourth peak.(71, 72)

Finally, the non-ligation-mediated interactions between the α C regions were studied using laser tweezers.(73) The authors used three α -chain fragments in the study: the “ α C region” corresponding to A α 221-610, the “ α C connector” corresponding to A α 221-391, and the “ α C domain” corresponding to A α 392-610. The authors found that α C region and α C region interact with each other with a higher frequency than other α C interaction and exhibit three distinct force thresholds of $19\pm 3\text{pN}$, $36\pm 2\text{pN}$, and $48\pm 2\text{pN}$. The α C region and α C-domain interact with three distinct force thresholds of $17\pm 5\text{pN}$, $36\pm 2\text{pN}$, and $49\pm 2\text{pN}$. The α C region and α C connector interact with each other less regularly and exhibit one force thresholds of $31\pm 6\text{pN}$. And finally the α C domain and α C connector exhibit one force threshold at $25\pm 3\text{pN}$. Combining the results indicates that the α C domain part of the α C region can bind to other α C domains to withstand forces of up to 50pN , while the α C domain can interact with the α C connector with much lower specificity and at lower forces $\sim 25\text{pN}$.(73)

1.3 Mechanical Models for the Extension of Fibrin Fibers within a Network

Given everything that is now known about the fibrin molecule and its interactions, three models have been proposed for the fibrin fiber mechanical properties, one based upon the unfolding of the coiled coil domains (See Figure 1.5), one based on an

alternative mechanism of cross-linking, and another which we proposed and is developed in more detail in Chapter 2, based upon the extension of the α C regions between sliding protofibrils. In this section, I will briefly describe each model, followed by a discussion of a few pros and cons of each.

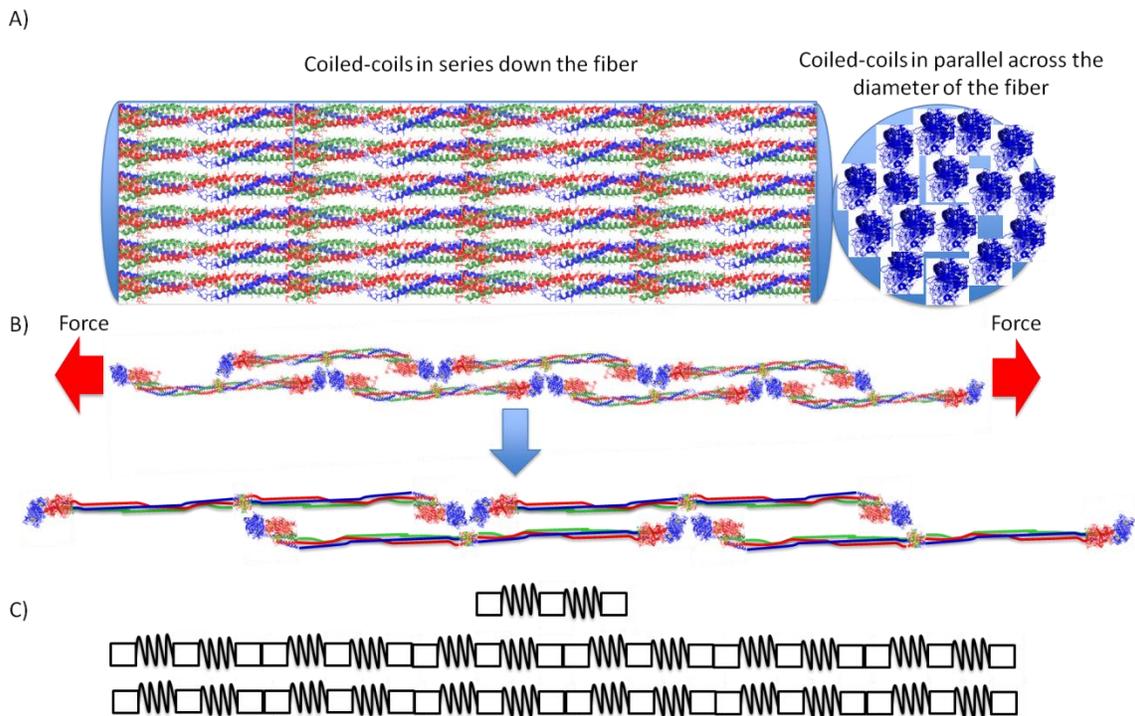


Figure 1.5 Coiled-coil (CC) extension model: In the coiled-coil extension model, the fibrin monomer can be thought of as two springs fixed between rigid beads (See panel “C” for the schematic version). The coiled coil in the fibrin molecule act as two state springs that can either be unstretched (folded) or stretched (unfolded). The fiber is composed of many springs in series and parallel down and across the fiber. When force is applied to the fiber, the coils unfold in a zipper-like manner. Protofibrils are assumed to span the entire length of the fiber, and little shear between protofibrils is assumed.

The coiled-coil unfolding (CC for short) model was proposed by Brown et. al after their fibrinogen oligomers experiment and was developed more fully in a recent paper showing that fibrin networks can stretch to strains of 250% before failing.(74, 75) The model views the fiber mechanically as a network of coiled-coil springs in series down the length of the fiber and in parallel across the diameter of the fiber (See Figure 1.5). The protofibrils within the fiber are assumed to be laterally attached in such a way

that they cannot slide. In the model, the fibrin extension comes from a zipper-like unfolding of the coils as the fiber is stretched.

This model is discussed in some detail in Chapter 2, briefly, the CC model is attractive in that it can account for fiber strains of 100%, and it explains the lack loss of 23nm signal in SAXS measurements of fibrin networks unfolding. The 23nm signal is attributed to the coiled coil structure of the molecules in the fiber.⁽⁷⁴⁾ However, there are several concerns with the coiled-coil unfolding model. The expected forces required to unfold the coiled-coil seem too high to explain the stretching of the fibrin fiber (this is discussed in great detail in section 2.3.6). In addition, while extensions of 100% strain are impressive, it is well established now that fibrin fibers can reversibly recover from strains of over 180%, and the coiled coil cannot account for all of this extension.⁽⁶⁶⁾

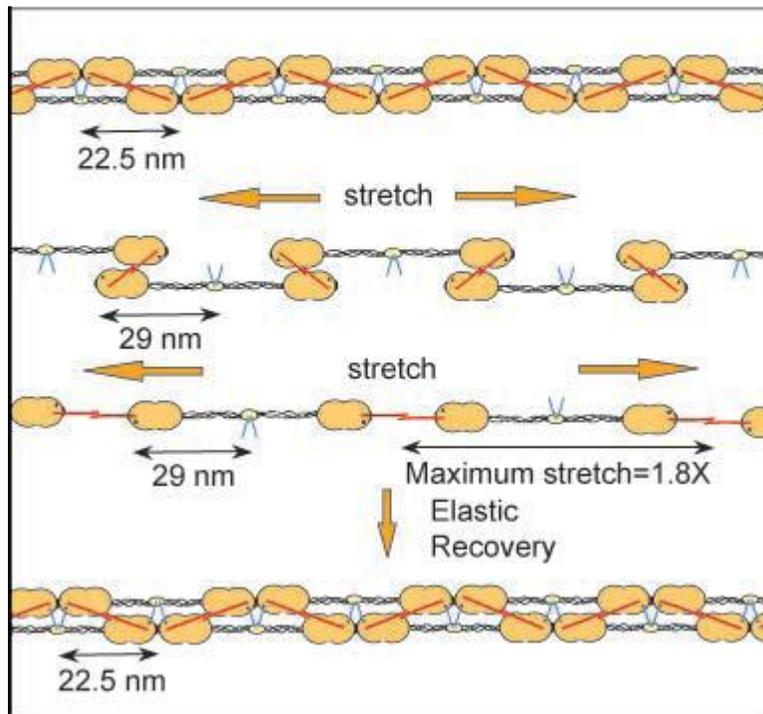


Figure 1.6: The Transverse Cross-linking (TCL) Extension Model: The model assumes that γ - γ cross-links in fibrin fibers occur transversely (red lines) between to γ - γ domains attached to half-staggered A-knobs, rather than longitudinally at the γ - γ interface as represented in Figure 1.3. Extension could come from breaking the A:a interactions

and intraprotofibril sliding of the fibrin molecules. Elastic recovery could occur for cross-linked fibers, but it is not clear how this model would account for un-cross-linked fibers. (Reprinted from Mosesson, M.W. "Cross-linked γ -Chains in fibrin fibrils bridge transversely between strands: yes" *Journal of Thrombosis and Haemostasis*, 2, 388-393. Used with Permission from John Wiley and Sons © 2004)

A second model for fibrin fiber extension was proposed by Mosesson in a discussion on the orientation of the γ cross-links between fibrin molecules within the fibrin.(32) In this proposal, the cross-links actually act transversely as seen in Figure 1.6. If the fiber was then stretched, and the A:a and B:b interactions were broken, the fiber would then be able to stretch up to 1.8 times its original length according to the authors.(32)

While this model would account for fiber strains of 180%, it presents other contradictions with existing data. First, the model is based on the assumption of transverse cross-linking, but evidence for transverse cross-linking has been disputed. (76) In addition, the model does not explain how un-ligated fibers can reversibly stretch to 1.8x their original length, which they are known to do.(67) Finally, it assumes that the A:a and B:b interactions will fail, but the fiber will retain its elasticity. While simulation results presented in Chapter 4 and AFM measurements indicate that A:a disruption can occur this only happens after unfolding the γ -C region which likely would cause plastic deformations in the fiber.(68, 70, 71, 75) Because of these issues, further discussion of it is not included in this document. However, if in the future it is shown conclusively that the main mode of cross-linking within fibers is transverse, then this model may need re-visiting.

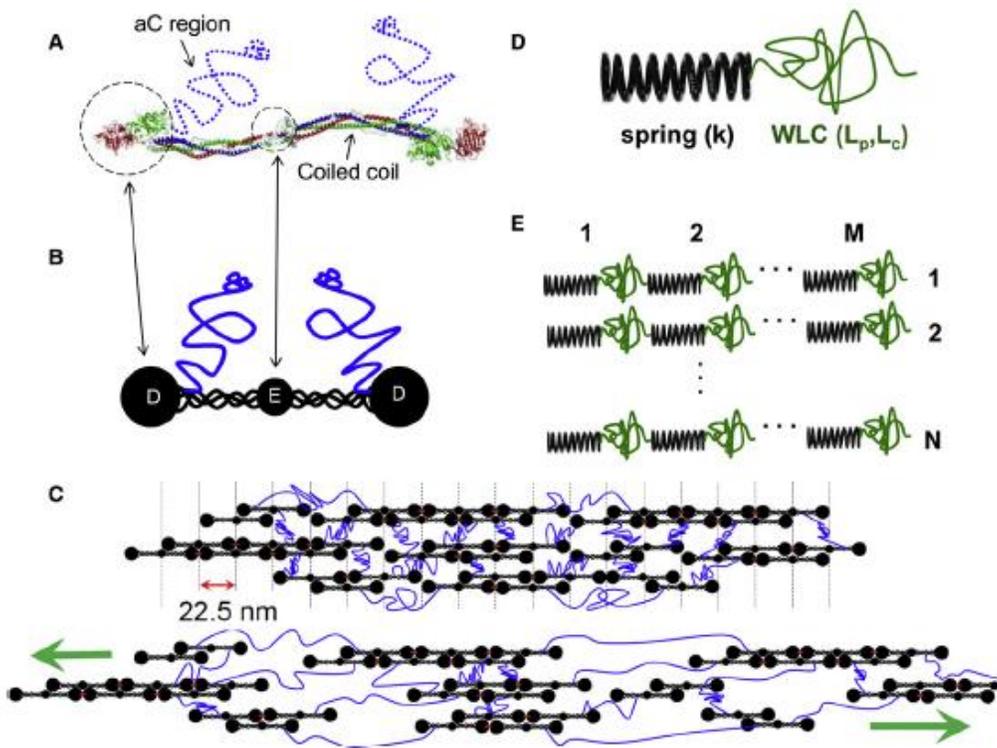


Figure 1.7: The α C extension model: This model is laid out in detail in Chapter 2. Briefly, the model considers the fibrin fiber as a series of stiff and soft springs in parallel and series down the fiber. The stiff springs represent the protofibrils consisting mainly of the coiled-coil, and the soft springs represent the α C regions linking the protofibrils. Stretching comes from sliding of the protofibrils.

We have recently proposed a third model (the α C model), based on evidence accumulated from single fiber stretching measurements, that the stretching of the fibrin fiber comes primarily from stretching of the α C region.(77, 78) The evidence and details for the model are laid on in Chapter 2. Briefly, the model treats the fiber as a series of soft springs in series with stiff strings (See Figure 1.7). The soft springs represent a natively unfolded region such as the α C region, while the stiff springs represent the structured protofibrils. While we believe that this model is a good starting point for fibrin mechanics, it needs to be developed to include all parts of the fibrin molecule and the higher order fibrin hierarchy.

To be complete, a hybrid model involving an α -helical to β -strand conversion of the coiled-coil region followed by an extension of the α C-region, followed by a partial unfolding of the γ -C domain was proposed by Guthold and Liu.(65, 67) While this model is more comprehensive than any of the other models as far as regions of the fibrin molecule involved, it is more abstract in nature, and not specific in details.

Thus, there are several models proposed to explain the mechanics of the fibrin fiber. Each model focuses on one part of the molecule (coiled-coil or α C region) or on a particular aspect of polymerization (transverse ligation mode) without accounting for other regions of the fibrin molecule. This, in part, is due to the fact that the structure of the α C domain, and the stability of the γ -C domains, and even the structure and topology of interactions within the fibrin fiber itself all have some level of uncertainty associated with them. In order to construct a full scale model of the fibrin fiber, these issues must be addressed. In chapter 4 of this paper we extend our original α C model into the fuller SL α CK model based on the results of our molecular dynamics simulations. We believe that the SL α CK model contains enough molecular detail to account for the full range of single fiber properties.

Ultimately, the properties of the single fibers will manifest themselves at the network level, so it is important to tie in the single fiber model with a network model. The attempts to do this thus far which can be broken down into two categories: shear models and stretching models. The shear models attempt to predict network behavior observed in cone and plate rheological studies originally pioneered by Ferry.(53, 54, 79) The original models considered the fiber as either a rigid, or semi-flexible rod and attempted to explain the network strain stiffening from that premise.(80, 81) A recent

model treated the fiber as a loosely bundled network of semi-flexible protofibrils linked by flexible cross-links (corresponding to the α C domains).(82) All of these models treated the fiber as either a rigid or semi-flexible polymer between network branch points, but were not able to originate the network behavior on the molecular level.

A molecular level model was employed by Brown et al. in their paper on fibrin networks under tensile stress.(83) The model implemented the CC stretching model within a model for biomacromolecular networks containing domains that unfold under force.(84) The model is elegant and generates fits that account for fibrin network behavior; however as discussed in 2.3.6 Analysis of the Coiled-Coil (CC) unfolding model, the molecular parameters needed to fit the network level behavior are non-physical. In chapter 5, we present a model for fibrin network stretching based on our α C stretching model for the single fiber. While the model is simple in form, it was able to predict the network response to the strain stiffening of individual fibers. At the network level, the CC model predicts the same strain stiffening behavior as the α C model; however the origins of the behavior are different, and the expected force levels differ by at least an order of magnitude.(85)

1.4 Significance and Goals of this project

There is increasing appreciation of the hierarchical design found in biomaterials like fibrin that give rise to their remarkable properties (86-88). At the same time, in order to build a mechanical model for these materials, not only the mechanical properties of the individual components, but also the connectivity of each component must be characterized. Fibrin has been studied for over seventy years, and the structure of the molecule, the interactions involved in protofibril assembly, and the overall role of

fibrin(ogen) in the blood coagulation cascade are now generally understood and accepted.(6) In addition, a series of measurements on the network, fiber, and molecular level have begun characterizing the mechanical properties of fibrin.(71, 83, 89) In spite of these advances, the complexity of the polymerization process and the hierarchical design of fibrin have proved to be difficult hurdles in the development of a complete multi-scale mechanical model of the fibrin network.

Fiber/Network Property	Coiled-coil model	α C Model	Transverse Ligation Model
Extensibility of 200% strain (Liu, 2006)	●	○	●
Stiffening after Ligation (Collet, 2006, Houser, 2010)	●	●	●
Fiber Strain Stiffening (Hudson, 2010, Liu, 2010)	●	●	●
Reversible extension of ligated fibers (Liu, 2006)	○	●	●
Reversible extension of un-ligated fibers (Lieu, 2006)	○	○	●
Sub Millisecond Recoil (this report)	○	●	○
Lateral Contacts between protofibrils (Weisel, 1986)	●	○	○
1MPa stiffness of fiber (Houser, 2010)	●	●	○
Force/monomer ~5-10pN at low strains (Houser, 2010)	●	●	○
A:a bonds withstanding >50pN forces (Averett, 2008)	●	●	●
Tension in Fibers	○	○	○
Negative clot compressibility(Brown, 2010)	●	○	○
SAXS 22nm banding pattern widening with stretching (Brown, 2010)	○	○	●
fibrin gamma domain stretching (Averett, 2008)	●	●	○
Fibrin Network Strain Stiffening (Brown, 2010, Ferry, 1974)	●	●	○

 Explained by model/included in model
  Data Inconsistent with model
  Model could be consistent, but more information needed
  No component of model to describe this data

Table 1.1 A comparison of the potential models for fibrin mechanical properties: A green circle represents a molecular, fiber level, or network level property that is explained by the model, a red filled circle represents a property that is not explained by the model, a red open circle represents a property that may be explained by the model but needs more investigation and a black open circle represents a property not addressed by the model.

In this work, I will present a series of experiments and discrete molecular dynamics (DMD) simulations designed to bridge the gap between the molecular mechanics and dynamics of fibrin, the manifestations of those molecular underpinnings in the mechanics and dynamics of individual fibrin fibers, and the higher order network

implications stemming from the mechanical properties of the molecule. The results of these projects will serve as a template to construct a model of fibrin mechanics and dynamics that bridges all levels of the fibrin hierarchy. In so doing, we will be able to provide a complete view of the mechanical role of fibrin in the process of Hemostasis.

This document is organized into measurements on three different length scales: the molecular length scale, the fiber length scale and the network length scale. The document begins by looking at the mechanical and dynamical properties at the intermediate length scale of the fibrin fiber, because the properties observed at the fiber level necessitate an investigation of their origins at the molecular level. The document then moves to the molecular length scale to investigate the origins of the fiber level properties and concludes at the network scale in an attempt to model the behavior of fibrin networks with a molecular model. The goal of Chapter 1 was to provide a brief overview of the fibrin(ogen) structure, function, polymerization and mechanics, while laying out the relevant questions remaining in understanding fibrin as a mechanical material. Chapter 2 details a series of AFM experiments indicating that fibrin fibers are elastomeric, meaning they have a low modulus, high extensibility, and strain stiffen. Evidence for these properties originating in the unstructured α C region of the molecule is presented, and a simple model for fibrin extension based on the Worm-like chain is presented. In chapter 3, fibrin recoil dynamics were probed using the atomic force microscope and high speed photography. This data was combined with DMD of the recoiling coiled-coil structure to understand the origins of fibrin's dynamic properties. Chapter 4 describes a series of steered discrete molecular dynamics simulations carried out to test the mechanical stability of various parts of the fibrinogen molecule. First a

structural model for the α C region was developed using homology modeling. Then simulations were run applying constant force to the α C region, the coiled-coil region, the γ - γ region, and the A:a interface to measure the relative stabilities of each part of the fibrinogen molecule. We conclude the chapter by presenting the SL α CK model for fibrin extension which extends our α C model to include stretching of the γ - γ interface and γ -C domains of the fibrin molecule. Chapter 5 focuses on the role of the individual fibrin fiber in fibrin network mechanics. We show using a simple model for network topology, that fibrin network strength is enhanced by the strain stiffening of individual fibrin fibers. Finally, a discussion of the conclusions and future directions of this work is presented in Chapter 6.

Chapter 2. Evidence that the α C Region is the Origin of the Low Modulus, High Extensibility, and Strain Stiffening in Fibrin fibers

2.1 An Introduction to Fibrin Fiber Mechanics

We recently reported that individual fibrin fibers exhibit extraordinary extensibility with an elastic regime in excess of a strain of 1.0, and strains at breaking of up to 3.0 (strain is defined here as fiber extension divided by original fiber length. A strain of 1.0 is a doubling of the fiber length) (65, 66). Several important and related questions emerge from that work: What are the full force extension characteristics of the fibrin fiber? Is fibrin an elastomeric fiber? What are the molecular origins of extensibility? How does FXIIIa ligation affect stiffness and extensibility? Along with very high extensibility, fibrin has a relatively low modulus (1-10 MPa range) and exhibits strain stiffening behavior above strains of 100% as we have recently reported (90) and as Liu et al. showed in their recent evaluation of the viscoelastic properties of fibrin fibers (91). These properties place fibrin squarely in the category of biomaterials such as elastin, resilin and spider silk (92, 93). This result is somewhat surprising given fibrin is predominantly a structured globular protein which polymerizes into ordered semi-crystalline arrangements in fibers (44). In contrast, elastomeric protein structures consist primarily of disordered networks of natively unfolded polypeptides with randomly distributed covalent crosslinks (94).

We present force vs. extension data on fibrin fibers with and without FXIIIa ligation, and show fibrin fibers have relatively low elastic moduli and exhibit pronounced strain stiffening. In situ measurements of fibrin fibers pre and post FXIIIa ligation

indicate a doubling in stiffness in the fiber after ligation. Based on the elastomeric nature of the fibrin fiber, along with previous measurements indicating that fibrin elasticity is tied to the length of the α C connector region, we present a mechanical model (henceforth called the α C model) of the fibrin fiber consisting of an ordered network of worm-like chain segments.⁽⁹⁵⁾ The success of this model in fitting our force vs. extension data suggests that fibrin's elasticity is entropic in origin, and that the observed stiffness and extensibility have their origins in unstructured regions of the protein.

2.2 Methods

For our measurements, we used an *Explorer* atomic force microscope (Explorer, Veeco Instruments, Woodbury, NY) resting on the manipulation stage of the optical epifluorescence (described in the next section) microscope enabling simultaneous AFM manipulation and optical data acquisition. For brief introductions to fluorescence microscopy and atomic force microscopy, the reader is referred to Appendices B and C respectively.

2.2.1 Structured Surface Preparation

Using PDMS (Sylgard 184; Dow Corning Co. Midland, MI) poured over an SU-8 master we created a stamp for the structured surfaces (SS). We then used the stamp to make channels on 24 x 50 mm 1.5 cover-glass (Corning) in Norland Optical #81 which we cured under UV light for 70 seconds. Channels were either 10 or 20 microns wide and ~10 microns deep. (See Figure 2.1)

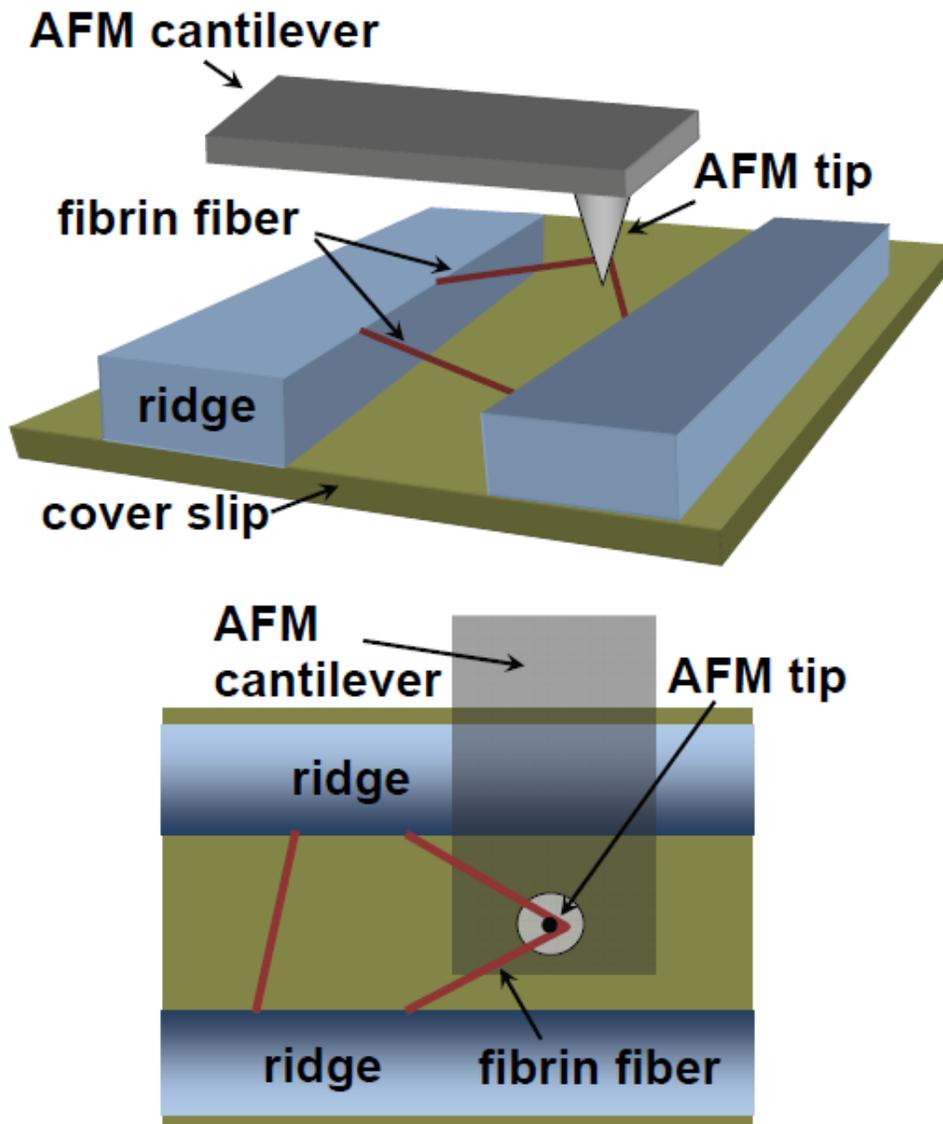


Figure 2.1 Set up for single fiber and network stretching experiments. Suspended fibrin fibers are labeled with fluorescent beads and then stretched with the AFM tip. Movies of the stretching are taken from below with epifluorescence imaging.

2.2.2 Fibrinogen Preparation

In all experiments we used recombinant human fibrinogen produced in Chinese Hamster Ovary (CHO) cells (96, 97). Recombinant human fibrinogen was grown in CHO cells without other blood coagulation factors and has been shown to be identical to

plasma fibrinogen in nearly all aspects. We assessed purity and homogeneity of this material through polyacrylamide gel electrophoresis (SDS-PAGE) and immunoblot analyses. Fibrinogen function was assessed through thrombin-catalyzed fibrinopeptide release monitored by High Performance Liquid Chromatography (HPLC), polymerization monitored by turbidity, and FXIII-catalyzed ligation monitored by SDS-PAGE, and blot analysis as described.

To form unligated fibrin fibers, fibrinogen was diluted in HBS (20 mM HEPES, 150 mM NaCl, pH 7.4) to 0.04 mg/ml and deposited onto the SS. Human thrombin (Enzyme Research Laboratories; final concentration 1 NIH U/ml in HBS with 10 mM CaCl₂) was added to the surface at the same volume as the fibrinogen (usually 8-10 μ L) and pipetted up and down to mix. The sample was incubated at 37°C for two hours in a water-saturated atmosphere, rinsed and stored in HBS. Fibers were fluorescently labeled by adding a 1/10,000 suspension of 24 nm volume-labeled red fluorescent carboxyl-coated microspheres (Invitrogen- Molecular Probes) in Ca-free HBS to the SS, incubated 5minutes and washed with buffer. To form fibers that were ligated during polymerization, we mixed FXIII (Enzyme Research Laboratories; final concentration: 0.5 mg/ml) with thrombin at 1 NIH U/mL in HBS with 10mM CaCl₂ to activate it into its active FXIIIa form and immediately added it to the SS containing 0.04 mg/mL fibrinogen. The samples were incubated for two hours at 37°C and then gently rinsed via buffer substitution. .

2.2.3 Optical Microscopy

For observation and manipulation of fibers, the coverslips were placed face up on an inverted Nikon Diaphot microscope with epifluorescence illumination (Nikon Diaphot

200, Southern Micro Instruments, Atlanta, GA), and imaged using a rhodamine HQ filter set (Chroma Technology, Rockingham, VT) and a 1.3 or 1.4 na 100X oil objective. Images were recorded using a high speed Cooke PCO 1600 camera with C-link and recorded with CamWare (Cooke Corp., Romulus, MI). All single fiber and network data was acquired as described previously (65, 66). For a full discussion of illumination techniques, see appendix B.

2.2.4 AFM

The atomic force microscope (Explorer, Veeco Instruments, Woodbury, NY) rests on the manipulation stage of the optical epifluorescence microscope enabling simultaneous AFM manipulation and optical data acquisition. Both OMCL-AC240TS-W2 (Olympus, Micro Cantilever) and RC150VB Biolever (Olympus, Asylum Research) AFM cantilevers (SiN) were used for manipulation. Force data was determined through calibration of the lateral deflection signal. The angular optical sensitivity of the twisting mode is the same as for the bending mode given we have a geometrically symmetric quadrant photodiode and similar gain settings for each quadrant. The twisting mode optical sensitivity and in deflection units and lateral spring constant are then determined using the specific geometry (length, tip length) of the cantilever using equations 2.1, 2.2 and C.6. The AFM tip was controlled using the Nanomanipulator software (3rdTech, Inc. Durham, NC). The tip was positioned within the channel next to a fiber and moved in one direction at $1\mu\text{m}/\text{sec}$ in 75nm increment steps to stretch the fiber. We note that this corresponds to a strain rate of roughly 0.1 s^{-1} ($\epsilon_{\text{max}}/\Delta t \sim (20\mu\text{m}/10\mu\text{m})/30\text{ s} \sim 2.0/30\text{ s} \sim 0.1\text{ s}^{-1}$) which is comparable to other force spectroscopy studies including those on myosin coiled coils (98). Care was taken to insure the fiber was contacted at the very end of the

AFM tip (< 1 micron from end)). For the stress vs. strain data depicted in Figure C.2 and Figure 2.2, force data was converted to stress by taking the calibrated force data and dividing by cross sectional area of the fibrin fiber as determined by AFM imaging (AFM measured fiber diameter which was converted to area assuming circular cross section). This was accomplished by imaging the fibrin fiber on the ridge immediately adjacent to the portion of the fiber suspended over the channel. Strain was determined by taking the extension (change in length, ΔL) of the fiber segments and dividing by the original length, L_0 (Strain : $\varepsilon = \Delta L / L_0$). . The twisting mode optical sensitivity in deflection units can be determined using the specific geometry (length, tip length) of the cantilever:

$$S_{L-R} = S_{T-B} \frac{2L}{3l_t} \quad (2.1)$$

Force can then be obtained using equations B.2 and B.4

$$F = k_L \frac{\Delta(L-R)}{S_{L-R}} \quad (2.2)$$

The lateral cantilever spring constant (k_L) was calculated from cantilever/geometry and SiN materials constants, as described in Equation C.7. Figure 2.1 depicts a diagram of the setup.

2.2.5 In-situ measurement of fiber Stiffening

To measure the change in stiffness due to ligation, of an individual fibrin fiber (In-situ), we prepared the fibrin on the SS without FXIII, and pulled the fiber to strains of up to 1.0 strain. Human fibrin fibers have been shown to be capable of strains up to 180% before sustaining plastic damage (66). To ligate the fibers we lifted the AFM from the sample, removed the buffer, and added HBS with Ca^{2+} (10 mM HEPES, 150 mM NaCl, 5mM $CaCl_2$, pH 7.4) with 5 mg/mL FXIII (which is activated to FXIIIa by thrombin) and 1 U/mL thrombin, and incubated for an hour at 37°C and then gently

rinsed via buffer substitution. After this, we replaced the AFM and pulled on the same fiber as we pulled on before ligation. As a control, we ran the whole experiment described above except doing a buffer transfer without adding FXIII and thrombin. All such controls showed no change in stiffness of the fibers. To confirm that ligation took place when we added FXIIIa to fibrin already deposited on the structured surfaces, we prepared additional fibrin samples on the structured surfaces without FXIIIa. We added FXIII and thrombin at 5mg/ML and 1 NIH U/mL respectively, incubated at 37°C for an hour, and then prepared the samples for Western blot analysis by replacing the fibrin buffer with 6X SDS, scraping the fibrin off the structured surfaces using a pipette tip (this was repeated three times) and combining all the samples in one tube, boiling the samples for 5 min, and then freezing them. A polyclonal antibody to fibrinogen (DAKO, Carpinteria, CA) was used for western blot analysis. Figure 2.4 shows that ligation of fibrin can be achieved after polymerization on the structured surface.

2.3 Results & Discussion

2.3.1 Single Fiber Stress vs. Strain

We prepared fibrin samples on micropatterned substrates with individual fibers suspended across channels (experimental geometry depicted in Figure 2.1). The sequence in Figure 2.2A-D depicts an AFM measurement of a single fibrin fiber stretched to the point of failure. The raw AFM force and position data were converted into fiber tensile strain and stress (see Methods). Typical stress vs. strain data for both FXIIIa ligated (red) and un-ligated fibers (blue) are depicted in Figure 2.2E, while Figure 2.2F shows their strain dependent tangent moduli (the slopes of the stress vs. strain curves in Figure 2.2E). At strain below 1.0, both fibers had a low and constant modulus. Above strain of 1.0 the fibers showed strain stiffening behavior. Both fibers exhibited roughly a ten-fold increase in stiffness between low and high strain. The bar plot in Figure 2.3 depicts the average differential modulus at discrete strains for fibers with and without FXIIIa ligation. The FXIIIa ligated fibers had an average elastic modulus of 2.1 ± 0.3 MPa at 0.25 strain rising to an average of 9.8 ± 1.2 MPa at failure. The un-ligated fibers showed 1.1 ± 0.2 MPa at 0.25 strain rising to 6.9 ± 1.3 MPa at failure. As a way of quantifying strain stiffening, we calculated the average ratio of maximum modulus to the modulus at 0.25 strain; the ligated fibers stiffened by a factor 6.2 ± 0.8 and un-ligated fibers stiffened by a factor 7.7 ± 1.5 .

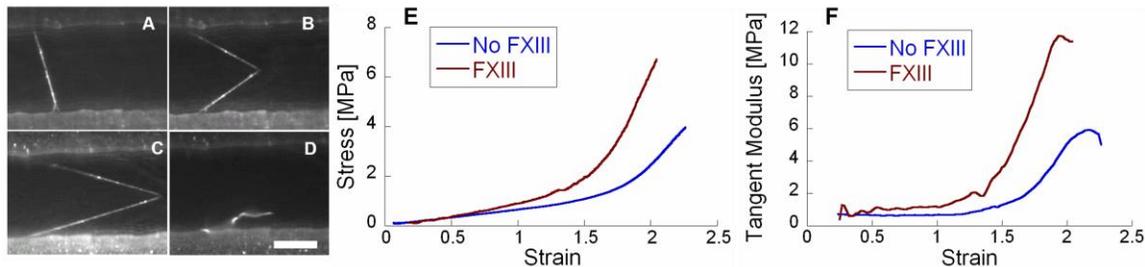


Figure 2.2 Stress vs. Strain plots of individual fibrin fibers. (A-D) AFM manipulation of a fibrin fiber suspended over micro patterned channel. The AFM tip (not visible) was brought in contact with a suspended fiber and stretched (B-C) to the point of failure (D). Scale bar = 10 μm . (E) Representative stress vs. strain data of individual fibers with and without FXIII ligation. Both fibers show relatively linear behavior up to strain just above 1.0 followed by significant stiffening. (F) Tangent modulus illustrating the strain dependent stiffness. These traces were found by numerically differentiating the traces in (E).

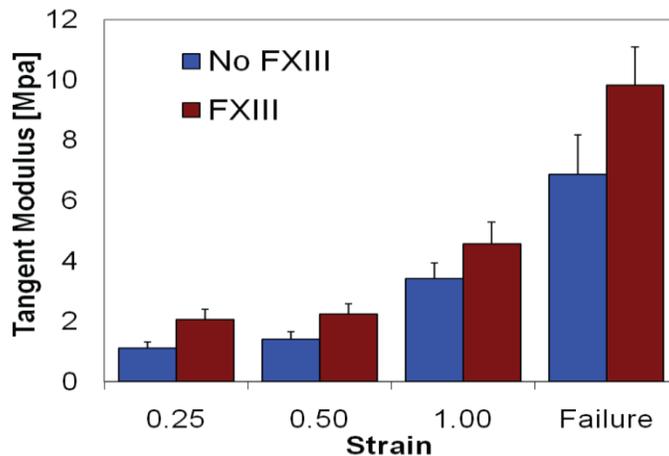


Figure 2.3 Average tangent modulus at discrete strains for ligated (N=14) and un-ligated (N=14) fibers. At 0.25 strain, the ligated fibers have an elastic modulus of 2.1 ± 0.3 MPa while the un-ligated have a modulus of 1.1 ± 0.2 MPa ($P < 0.003$). The average modulus rises to 9.8 ± 1.2 MPa for ligated and 6.9 ± 1.3 MPa for un-ligated fibers ($P < 0.05$).

2.3.2 In situ measurements

Determination of the stress vs. strain behavior of fibers yields intensive materials parameters, such as elastic modulus, by normalizing the raw force vs. elongation data by the length and diameter of the fibers. This allowed comparison of the stiffness of different fibers over many experiments. We estimated the fiber cross sectional area by using the AFM to measure the diameter of the fiber on the ridge. In calculating stress, we assumed that the fiber cross section is circular, that the fiber diameter on the ridge surface

is equivalent to the suspended fiber diameter, and that the diameter is constant over the suspended length. A comparison of the full stress vs. strain behavior from one experiment to another also requires comparing data from different AFM tips and force calibrations which adds additional uncertainties.

In order to address these uncertainties in comparing ligated to unligated fibers, and to corroborate the stress vs. strain data, we measured the relative change in stiffness of individual fibrin fibers before and after FXIIIa ligation in-situ (Figure 2.4). Results of a western blot indicating successful ligation is shown in Figure 2.4 (see Methods). These measurements provided direct comparisons of fiber stiffness with and without ligation in the same fibers, thus obviating the need for geometrical normalizations or force calibration. Fibrin was prepared on structured surfaces without ligation then stretched with the AFM. These initial stretches corresponded to strains of 0.40 or less in order to maintain mechanical reversibility. FXIIIa was then added to the sample and the same fiber was stretched again. Figure 2.4 shows representative data revealing increased stiffness with addition of FXIIIa. We observed an increase in stiffness after ligation by FXIIIa in all eight fibers we measured; the average increase was 80%, in agreement with an increase of $90 \pm 44\%$ obtained from the stress vs. strain measurements at strain of 0.25 (Figure 2.3). These results show that FXIIIa acts directly on the mechanical properties of the fiber itself; ligation more than doubles fiber stiffness. Our results are also consistent with low strain thermal noise spectrum measurements taken on fibers in clots (64).

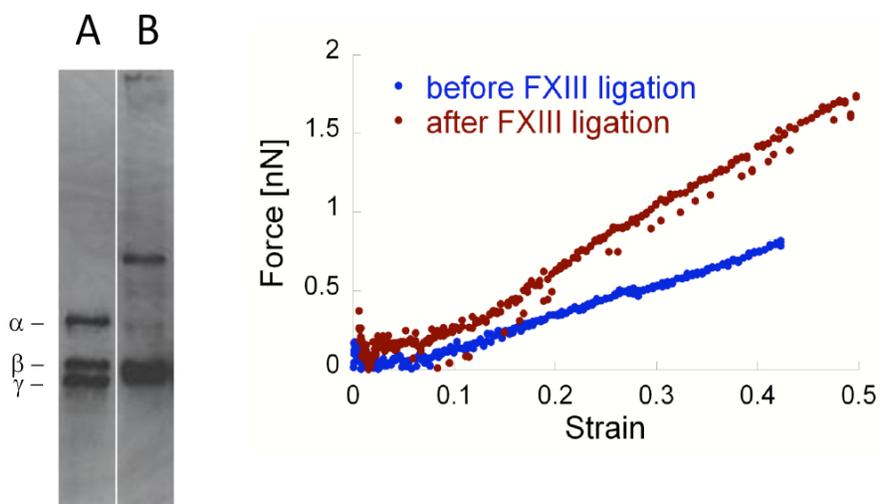


Figure 2.4 Left: Western blot of fibrin showing α , β and γ bands without FXIIIa (A) and α , β , and γ - γ dimer following ligation with FXIIIa (B). Right: In-situ fiber stiffness measurements. Fiber stiffness before (blue) and after (red) FXIII ligation.

2.3.3 The α C Stretching Model

To interpret the force vs extension data, we developed a model of the mechanics of fibrin fiber extension (the α C extension model). Figure 2.5 shows the structure of the fibrin monomer (Figure 2.5A&B), and our mechanical model of a fibrin fiber consisting of an ordered array of fundamental mechanical units (Figure 2.5D&E). Each fundamental unit has an “unstructured portion” modeled mechanically as a worm-like chain (WLC) and a structured portion modeled as a Hookean spring. Previous reports agree that the extreme extensibility of the fibrin fiber reflects properties of the fibrin monomer. These reports suggest that an unfolded segment in the molecule is critical: either a natively unfolded region such as the α C domain (95) or a force-unfolded region such as the coiled-coil (74, 75) or the gamma domain in the D region (71). In any case, a large portion of the monomer’s structured architecture will remain folded and is represented by the stiff spring. As this spring represents the structured globular regions of the protein, its stiffness is expected to be several orders of magnitude higher than the

unstructured region, and will not contribute significantly to the extension of the fiber [See section 2.3.4]. This force-extension model assumes that fiber extension is accommodated only by the unstructured region, which we model as a WLC. We have not explicitly included unfolding as a mechanism of the α C extension model. Unfolding may very well be happening, but our analysis shows it is unlikely until higher strain. Our force vs. extension data also shows no signatures or features indicative of abrupt structural changes within the fiber, and is suggestive of a single extension mechanism especially up to strains of 1.0. Our model is intended to present the simplest explanation of fibrin fiber extensibility that is consistent with what is known about fibrin fiber structure, the magnitudes of protein unfolding force thresholds, and the form of our force vs. extension data.

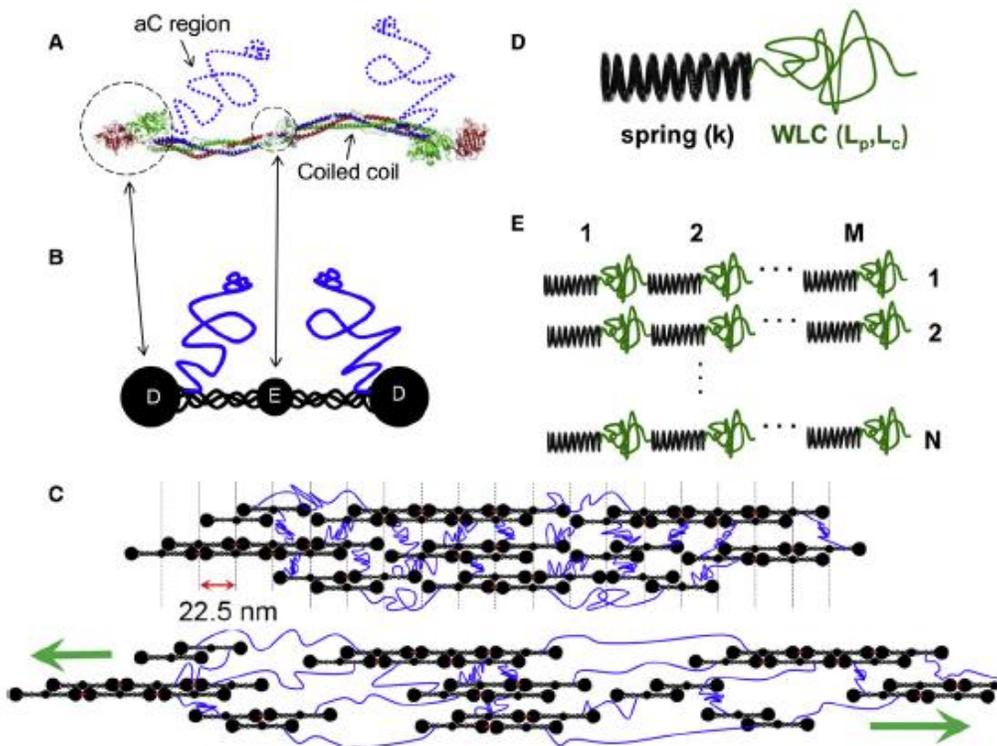


Figure 2.5 Fibrin structure and corresponding mechanical model. (A) Crystal structure of fibrinogen. α , β and γ polypeptides in blue, green and red respectively, with a cartoon of the C terminus region of the alpha chain (in dashed blue). (B) Cartoon depiction of the fibrinogen. (C) Upper model: Simplified Fibrin fiber structure. Intra-Protofibril FXIII induced covalent interactions are indicated by red dashes. Protofibrils are connected through α C region interactions (α C regions in blue). We note that this picture is simplified for clarity. In a real fiber, each monomer has two α C regions extending. Blue lines spanning protofibrils represent α C/ α C interactions. Though the figure suggests only pairwise α C interactions, it is known that α C regions typically form interactions with multiple other α C regions. Lower Model: Stretching of the fiber under stress. This cartoon depicts a model of one potential mechanism of extensibility. In this case the α C domains, though linking protofibrils laterally, accommodate the tensile strain induced by applied force (green arrows). The stiff protofibrils, within this model, are represented by the stiff springs and accommodate little of the strain. (D) Simple mechanical model of the fibrin monomer. The linear spring (black) represents the stiff structured portion of the monomer (spring constant k). The random coil (green) represents the unstructured portions of the protein. These could be either those regions natively unfolded (α C domain) or any mechanically unfolded region of the protein such as the coiled coil or portion of the D region. The WLC force vs. extension relation is parameterized with persistence length, L_p , and contour length, L_c . (E) Simplified mechanical model of the fibrin fiber. There are M monomers in series and N single monomer chains in parallel. The model includes no lateral interactions as they have no relevance to uniform tensile stretching.

Figure 2.5C depicts fibrin fiber structure in cartoon form, and a candidate model for fibrin extension. The model depicts protofibrils laterally aggregated through α C interactions. The α C domains (blue) accommodate the majority of the fiber strain and are represented in the mechanical model (Figure 2.5D&E) by the WLC. As the cartoon of the strain fiber indicates, the α C regions can accommodate tensile strain and act as spring in series within this proposed model despite acting as lateral tethers between protofibrils. As the fiber is strained, these tethers align along the fiber axis and act as series connections between the stiffer protofibrils. The distribution of protofibril lengths depicted within this cartoon is consistent with published studies of protofibril lengths.(40, 99)

A WLC model has successfully described the force-extension behavior of DNA, intrinsically unstructured polypeptides, and force-unfolded proteins (98, 100-104). The WLC is an idealized non-self-interacting flexible chain undergoing thermal fluctuations along its contour. The force of extension of the WLC is mediated entirely by entropic elasticity. Though this model has primarily been used to describe single molecule force spectroscopy data, we apply it here, in scaled form, to a full fibrin fiber. The form of the force vs. extension for a WLC in the Marko-Siggia form (104) is as follows.

$$F(\Delta l) = \frac{k_B T}{l_p} \left[\frac{1}{4} \left(1 - \frac{\Delta l}{l_c} \right)^{-2} - \frac{1}{4} + \frac{\Delta l}{l_c} \right] \quad (2.3)$$

Where, Δl is extension ($L - L_0$), k_B is the Boltzmann constant, T is temperature, l_p is persistence length, and l_c is the contour length. We note our use of Δl in Eqn. 2.3, rather than end-to-end length as is appropriate for a single random coil polymer chain. In our case we assume the relaxed length of the fiber segment, L_0 , is due entirely to the

structured portions of the monomer; the unstructured portion only contributes appreciable length to the fiber as it extends. In reality, the unstructured domains may contribute a small percentage to the unstretched fiber length, but we assume this contribution is negligible.

We model the force-extension behavior of the fundamental mechanical unit (Fig D) as a WLC (Eqn (1)). The whole fibrin fiber is modeled by M identical fundamental units in series and N of these chains in parallel (Figure 2.5E). A chain with M , identical, WLCs linked in series behaves as a WLC with a contour length equal to M times the contour length of the primitive unit:

$$L_c = M * l_c \quad (2.4)$$

Substituting Eqn. 2.4 into Eqn. 2.3, yields the force extension relation for a series of MWLC's

$$F_{M,1}(\Delta L) = \frac{k_B T}{l_p} \left[\frac{1}{4} \left(1 - \frac{\Delta L}{M * l_c} \right)^{-2} - \frac{1}{4} + \frac{\Delta L}{M * l_c} \right] \quad (2.5)$$

where ΔL is the extension of the entire fiber. The force of N WLC's in parallel is N times the force of the individual WLCs.

$$F_{M,N}(\varepsilon) = N * F_{M,1} = N * \frac{k_B T}{l_p} \left[\frac{1}{4} \left(1 - \frac{\varepsilon}{(M * l_c) / L_0} \right)^{-2} - \frac{1}{4} + \frac{\varepsilon}{(M * l_c) / L_0} \right] \quad (2.6)$$

where here we normalize the extension, ΔL , and contour length, $M * l_c$, to the original fiber segment length, L_0 , to produced force as a function of strain ($\varepsilon = \Delta L / L_0$). Because of the ordered arrangement of monomers within our fibrin fiber model, the force extension relation for the full fiber (Eqn. 2.6) is a scaled version of the single WLC relation with the persistence length l_p and contour length l_c replaced with l_p / N and $M l_c$ respectively.

As Fig. 4.8 shows, the scaled WLC equation (Eqn. 4) does a very good job of fitting force vs. strain data for both the FXIIIa ligated and unligated fibers. The fitting

parameters obtained are the scaled persistence length (l_p/N) and the scaled contour length (Ml_c/L_0). Both parameters provide insight into the molecular origins of fibrin's strain accommodation. To tease out the molecular parameters l_p and l_c , an estimate is made for M and N based on the geometry of the fiber determined by AFM of fibers on the ridges for diameter and fluorescence microscopy for the length. The number N , is found by taking the ratio of the cross sectional area of the fibrin fiber to the estimated cross sectional area of a monomer (or the ratio of the squares of the diameters). Using 3-7 nm as a rough estimated range for the effective individual monomer diameter –corresponding to 200 - 1100 monomers in parallel for a 100 nm fiber (see supplemental material)- we obtain an average persistence length value of 0.1 - 0.6 nm. This number is intended to be taken as an order of magnitude bench mark to compare with results from other single molecule force spectroscopy studies. Experimental persistence lengths for polypeptides fall in the range of 0.4-1.5nm (75, 101, 105, 106).

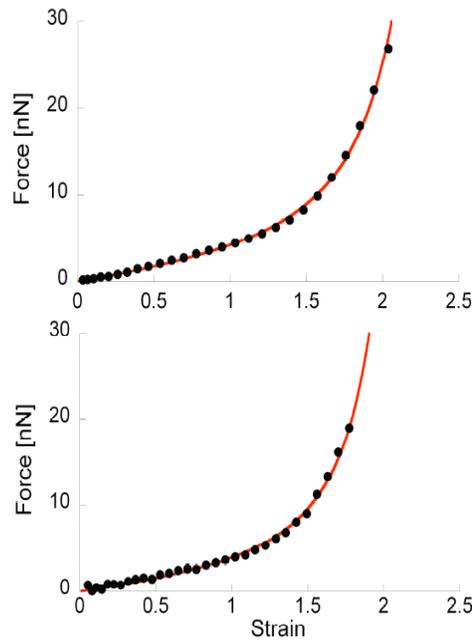


Figure 2.6 Force vs. Strain data (black circles) for un-ligated fiber (above) and ligated fiber below (5% of data points shown for clarity). The red curve is the fit of Eqn 2.6.

The second fitting parameter (MI_c / L_0) describes the effective contour length of the fiber as a strain. L_0 is simply M times the original length of the monomer ($L_0 = MI_0$) - for our fibers, L_0 is typically 10 microns - and l_0 is 45 nm, yielding $M = 220$ monomers in series. Substituting, the fitting parameter simplifies to (l_c / l_0) . This is a scaled contour length that expresses the contour length of the unstructured portion of the monomer as a fraction of the original monomer length, l_0 . We obtained a scaled contour length of 2.5 ± 0.4 ; at full extension of the unstructured portion of the monomer, the additional extension is 2.5 times the original length (and therefore the end to end length is 3.5 times the original end to end distance). For a monomer of 46 nm, this yields an additional length of ~ 115 nm. This length corresponds to 320 residues responsible for the WLC like extension of the monomer, which when split between the two identical sides of the fibrin monomer yields 160 residues. The tethered αC region contains nearly 400 amino acids, at least 200 of which are known to be part of an unstructured region and

could easily account for this additional length. An unfolded portion of the D region of the protein could also account for this additional length. Unfolding and completely extending both coiled coil regions leads to 46 nm extra length (75) accounting for less than half of the extension. We found no statistically significant difference between FXIIIa ligated and un-ligated fibers in either the persistence length ($p > 0.20$) or contour length ($p > 0.05$) parameters.

2.3.4 Enthalpic backbone stiffness of the fibrin fiber

The WLC model fitting to the force extension data described above (Eqn. 2.6) and Figure 2.6) takes into account entropic elasticity only, ignoring the effects arising from enthalpic effects (stretching of chemical or physical bonds). Following Wang et. al. (107) we can add enthalpic contributions to the extension of a WLC, which in the case of fibrin could result from stretching of the coiled coil or other structured regions. We modify the WLC by making the following substitution for Eqn 2.3:

$$\frac{\Delta l}{l_c} \rightarrow \frac{\Delta l}{l_c} - \frac{F}{K_0} \quad (2.7)$$

This substitution adds the contribution of the backbone strain to the overall extension; F/K_0 is equal to the strain under force F due to stretching of a Hookean spring element. The spring stiffness is represented by K_0 which has units of force per unit strain. Alternately, we can express the stiffness in terms of a Young's modulus $E=A*K_0$ where A is the cross sectional area of the spring element(s). Substituting Eqn. 2.7 into 2.3 yields

$$F = \frac{k_B T}{l_p} \left[\frac{1}{4} \left(1 - \frac{\Delta l}{l_c} + \frac{F}{K_0} \right)^{-2} - \frac{1}{4} + \frac{\Delta l}{l_c} - \frac{F}{K_0} \right] \quad (2.8)$$

The enthalpic correction, F/K_0 , is non-negligible in the case of relatively weak backbone stiffness or relatively large forces, i.e. $F \ll K_0$. For typical values for the radius of a fibrin fiber (50-100nm,) and we assume the differential modulus of the coiled coil is 1GPa - the same order of magnitude as other proteins consisting of coiled coils, such as alpha keratin - we get $K_0 \sim 7500 - 30,000$ nN. As the maximum force measured, before breaking, for fibrin fibers is in the 10's of nN range, it is clear that the condition $F \ll K_0$ is met. If the stiffness of the enthalpic components is ~ 10 MPa then $K_0 \sim 75 - 300$ nN and we would expect a noticeable change to the force extension curve. For extensions near and beyond the contour length the force curve will become linear, although, fibrin fibers are liable to break before reaching these extensions, even if K_0 is sufficiently small. The fact that the un-modified WLC model fits our data so closely seems to indicate that we are in the $F \ll K_0$ limit and the backbone stiffness can be ignored. Figure 2.7 shows several additional experimentally obtained force curves, fit using the WLC, entropic forces only, model.

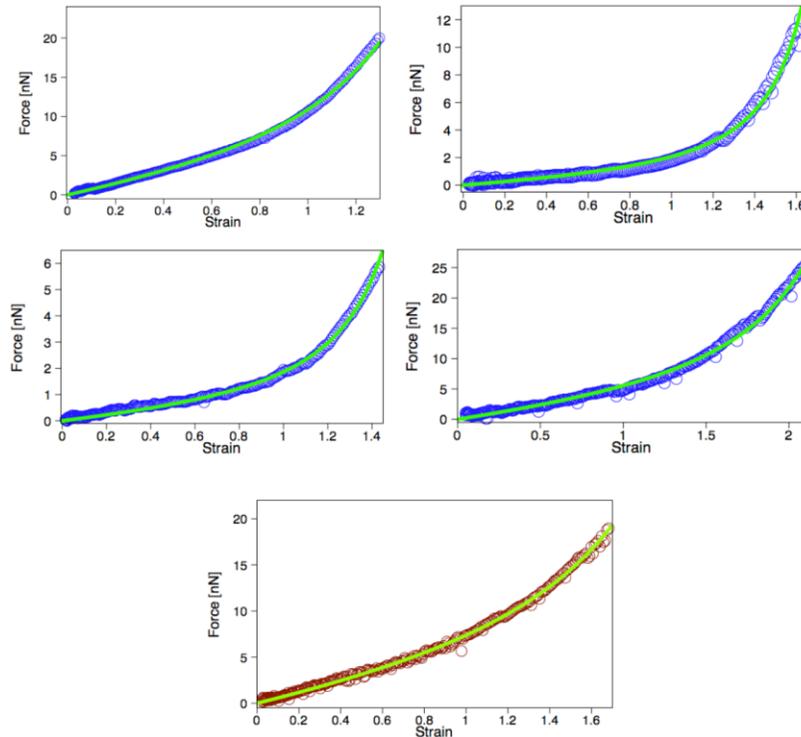


Figure 2.7 Several representative experimentally obtained force curves in addition to the two in the main text. These were fit using the WLC model. The green lines show the fitted values, the blue points are the experimentally observed data for un-ligated fibrin fibers, and the red points are the experimentally observed data for fibers ligated by FXIII.

2.3.5 Does the α C connector region mediate fibrin's initial extensibility?

In a recent study, we demonstrated that fibrin extensibility is correlated to the length of the α C connector region in the fibrin monomer(95). The longer the length of the unstructured region, the longer the fiber was able to extend. In order for the α C regions to play a key role in the elastic modulus and extensibility of fibrin, they must be arranged within the fiber such that they make series connections (longitudinal) between protofibrils in addition to parallel (lateral) connections. The conventional picture of fibrin structure is that the α C regions connect in a lateral arrangement. However, work by Hantgan and Ferry among others has shown that protofibrils have a distribution of lengths varying from a few monomers up to 20 or so in standard polymerization conditions (40, 99, 108). Given the protofibrils are relatively short - a few hundred nm at

most, as compared to the whole fiber which is tens of microns long - the inter-protofibril connections must be crucial to supporting tensile forces. The α C regions are therefore likely to play a role as “springs in series” with the stiffer protofibrils (See Figure 2.5E). This view is consistent with our WLC in series model for fibrin extension and is generally consistent with the observed low modulus and high extensibility. However, this model strongly argues that the series α C connections give rise to a significant fraction of the total extension of the fiber.

Another important piece of the experimental evidence that is relevant to this question is the stiffening of the fiber with FXIII ligation. The fibrin monomer has known ligation sites within its γ domains and its α C region (30, 109). Within a model in which coiled-coil domain unfolding leads to fibrin’s extensibility, and FXIII ligation only strengthens already existing gamma-gamma contacts (intra-protofibril interactions), it is hard to construct a scenario which FXIIIa ligation would result in fiber stiffening. An increase in fiber strength – a higher tensile stress before failure – would be expected, but not stiffening. Instead, stiffening suggests the mechanisms of extension are becoming more restricted through FXIIIa ligation. Addition of new constraints on conformational freedom are required to provide this, and we believe the most obvious places this could happen are in the inter-protofibril α C connections. The α C region is known to have multiple sites that can form dipeptide bonds through FXIIIa ligation. Like in a polymer gel, where increased crosslink density results in a stiffer material, additional covalent α C connections facilitated by FXIIIa would stiffen the inter-protofibril mechanical linkages. Within our model, this would effectively reduce the average contour length of the constituent WLC’s. Our fits to the force extension curves did not indicate clear change in

the contour length fitting parameter, though we did observe stiffening of fibrin fibers with FXIIIa ligation. This could be due to incomplete ligation of the α C chains by FXIIIa in our reaction. Recent work by Liu et. al. showed that full α C ligation reduces the extensibility of the fiber, reducing the contour length (91). Further refinement in the force vs. extension data as well as the WLC-in-series model may be required to tease out subtle changes in the effective contour length of the α C upon FXIIIa ligation.

2.3.6 Analysis of the Coiled-Coil (CC) unfolding model

A 2007 force spectroscopy study of oligomers of fibrinogen showed compelling evidence of coiled-coil unfolding. In these studies, the applied force was necessarily transferred through the D region, coiled-coil and E region, eliminating any contribution of the unstructured α C region. Thus, although these studies provide evidence that the coiled-coil is unfolding, it is not clear that the data are relevant to the stretching of physiological fibrin fibers, which are polymers with many parallel and series connections between monomers. More recently, a two-state model for fibrin extensibility has recently been proposed by Brown et al. (74) as part of an impressive multi-scale study of large fibrin networks (See Figure 1.5). This model assumes that unfolding of the coiled-coil region mediates fibrin's extensibility. In the model, the coiled-coils can be in two states: (a) the folded state where their stress-strain behavior is linear, and (b) the unfolded state where the coil is unfolded and behaves like a worm-like chain at high extensions (110). The distribution of folded and unfolded monomers is governed thermodynamically by the unfolding energy barrier and the distance between the energy wells of the folded and unfolded states. This model is appealing in that it meshes well with the conventional view of fibrin structure in which the tensile force is supported completely through

gamma-gamma contacts within the protofibrils with the coiled-coils acting as “springs in series”. However, the coiled-coil mediated extensibility model has difficulty accounting for the full stress-strain behavior of the fibrin fiber presented here and in other publications (90, 91).

We will now apply the two -state model recently proposed by Brown et al. to our single fiber data. In the model, the coiled-coils can be in two states: (a) the folded state where their stress-strain behavior is linear, and (b) the unfolded state where the coil is unfolded and behaves like a worm-like chain at high extensions (The Marko-Siggia worm-like chain also includes a linear region at low extensions which is not part of this two state model). (111) This model for individual fiber elasticity was then used in the 8-chain polymer physics model to represent the behavior of a fibrin network under tensile deformation. (84) The 8-chain model was fit to the network data using thermodynamic properties (the alpha-helix unfolding energy barrier and the alpha-helix unfolded extension length) as fitting parameters. While the authors were able to use the coiled-coil single fiber model to fit their network data, the single fiber force extension model embedded within the network model was not compared to single fiber force extension data.

The force curve given in the Brown paper is defined as the following (74):

$$\lambda = \frac{x}{L_f} = n_f(F) \left(1 + \frac{F}{EA} \right) + n_u(F) \frac{L_u}{L_f} \left(1 - \frac{\sqrt{k_B T N}}{2\sqrt{l_p F}} \right) \quad (2.9)$$

Where

$$n_f = \frac{1}{1 + \exp\left(-\frac{\Delta G_B - \frac{F\Delta z}{N}}{k_B T}\right)} \quad (2.10)$$

$$n_u = 1 - n_f \quad (2.11)$$

and x is the fiber extension, F is the force exerted on the fiber, N is the number of monomers/fiber cross-section, L_u is the unfolded coil contour length, L_f is the folded coil length, $\Delta\pi$ is the unfolding energy barrier, Δz is the distance between the folded and unfolded lengths and l_p is the WLC persistence length.

For the 8-chain model fit in the paper, L_u/L_f was assumed to be 2.1 based on the length of the unfolded coiled-coil, N was 1200, $EA=387.5\text{nN}$, $l_p=0.8\text{nm}$, and ΔG_B and Δz were kept as the fitting parameters. The fit yielded parameters of $\Delta G_B = 16.5k_B T$ and $\Delta z=19.2\text{nm}$. Equation 2.9 is plotted with these parameters in Figure 4.10. Our physical interpretation of this plot is that as the fiber is stretched, it will actually shrink from its folded length before beginning to stretch. Due to the large values for ΔG_B and Δz , this shrinking is likely due to the many coiled-coils unfolding within a relatively narrow force range, leaving the WLC's to entropically shrink before being stretched back out. This behavior is inconsistent with our force extension data which shows smooth, linear, low stiffness behavior up to strains of 100% or more. The plot is also hard to reconcile with single molecule force extension data for coiled-coils (103, 112-114).

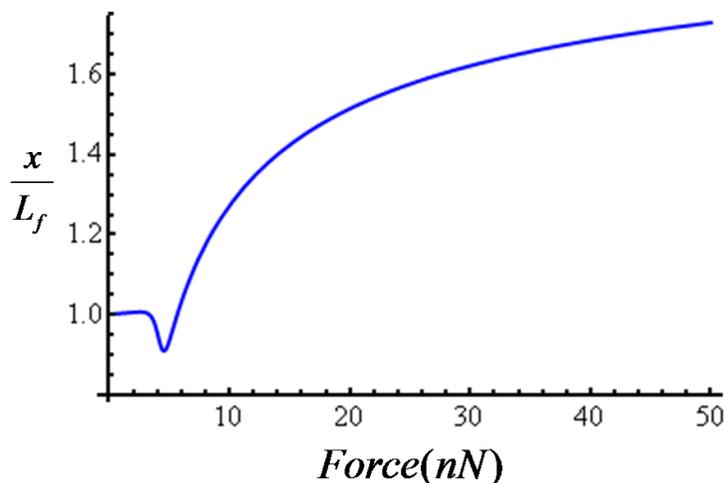


Figure 2.8 Single fiber force strain plot for coiled-coil model. The plot represents the extension-force curve arising from the assumptions and fitting parameters ($\Delta\pi = 16.5\text{kT}$ and $\Delta z=19.2\text{nm}$) from the coil-coil network fit .

In the model for single fiber elasticity presented in the Brown et al. study, a value of 14 MPa is used as the elastic modulus of the coiled-coil. This assumption is based on a previous experimental results indicating a 14.5 MPa elastic modulus for a fibrin fiber (64). However, the assumption that the elastic modulus of the fiber is similar to the modulus of the in-tact (unfolded) coiled-coil region is questionable given the known stiffness of coiled-coil structures. Biomaterials whose extensibility is known to be mediated by coiled-coil unfolding, including myosin, egg capsule(113), and intermediate filament based materials such as keratin(112), vimentin(114, 115), and others, all show a characteristic force vs. extension signature that includes three force regimes: an initial stiff regime ranging from 150 –1000 MPa corresponding the unfolding of the hydrogen bond mediated coiled coil, a softer regime corresponding to the extension of the uncoiled peptide chain, and a second stiffening regime where the uncoiled chain reaches its contour length. Gigapascal scale stiffness is also consistent with numerous theoretical treatments of the alpha helix stiffness (116-118). In some cases such as soft keratins and hagfish slime thread (119, 120), intermediate filament based materials show much lower

low strain modulus (a few MPa), but in these cases low stiffness is attributed to the unstructured elastomeric protein matrix connected in series with the stiffer coiled coils.

Thus, experimental, theoretical, and computational studies yield elastic modulus values for coiled-coil that are 3-4 orders of magnitude higher than that for the overall fiber measured in this study and others. The volume fraction of protein material within the fiber is one possible explanation for the discrepancy. However, estimates put the water fraction within the fiber at 70% which would provide a factor of 3 or 4 to work with, well short of the factor of 1000 to 10,000 needed to reconcile fibrin fiber stiffness and expected coil-coil stiffness.

We evaluated our single fiber force vs. extension data using the two state model assuming a three values of stiffness for the coiled-coil structure: 2 MPa, the measured modulus of the fiber measured using the AFM, 100 MPa (The low-end value for a material with stiffness known to be mediated by coiled-coil unfolding)(113), and 800 MPa, (The modulus of a 2 GPa stiff coiled coil structure surrounded by 80% water). While the fits using the low elastic modulus (2MPa) for the coils do match the data, they produce fitting parameters that are non-physical (e.g. picometer range persistence length). On the other hand, using a stiff spring model for the coiled coils requires an anomalously low unfolding energy barrier at low strain and the fit fails significantly at high strain. Three fits to one fiber force vs. extension curve are shown in Figure 2.9. The fiber had a radius of 45nm. The fits were performed using 4 fitting parameters, ΔG_B , Δz , l_p , and L_u/L_f . using Origin Pro (OriginLab Corp. Northampton, MA). The fits were dependent on the initial starting parameters, and most starting parameters did not lead to a convergent answer. In particular, fits attempted using a starting value of $\Delta G_B > 5k_B T$ did

not converge. Plots A and B in Figure 2.9 were performed using the starting parameters $G_B = 5kT$; $\Delta z = 3nm$; $L_c/L_o = 2.5$; $L_p = 0.8nm$, while Plot C was formed using : $\Delta G_B = 0.3kT$; $\Delta z = 0.3nm$; $L_c/L_o = 2.5$; $L_p = 0.2nm$.

Plot A in Figure 2.9 fits the data fairly well ($R^2 = 0.99$), but it assumes a coil-coiled modulus of 2 MPa, much lower than the measured modulus of a coiled coil. In addition, some of the fitting parameters do not make physical sense: $\Delta G_B = 4.1kT$; $\Delta z = 0.08nm$; $L_c/L_o = 3.66$; $L_p = 0.008nm$. While the ΔG_B value can be reconciled with theoretical coiled-coil unfolding barriers, $\Delta z = 0.08nm$ and $L_p = 0.008nm$ do not make sense physically. (121) The known L_p for unraveled or natively unstructured proteins is 0.2-0.9nm, (98, 100, 101) two orders of magnitude higher than the fits to our model, while the same experiments and molecular dynamics simulations show that coils unfold at 3-15% strain, (103, 112-114) the fitting parameter $\Delta z = 0.08$ corresponds to 0.1% strain, much lower than the known values.

On the other hand, plots B and C use a coiled-coil modulus of 0.8GPa, similar to the known values of the coiled-coil modulus. Plot B clearly does not fit the data at all, but does resemble other coiled-coil unfolding plots which have a very sharp peak in force before coil-unraveling begins to occur (103, 112-114). Plot C does fit the data at low extension before diverging in the strain stiffening regime, however once again some of the fitting parameters yielded do not make physical sense: $\Delta G_B = 1.99k_B T$; $\Delta z = 0.02nm$; $L_c/L_o = 2.827$; $L_p = 0.2nm$. These fitting parameters indicate that the coils would have to unfold at relatively low unfolding energies and at Δz values corresponding to 0.04% strain. In addition to the fits shown, other fits were done with a coil modulus of 100Mpa, yielding similar results to those done with the 0.8Gpa modulus.

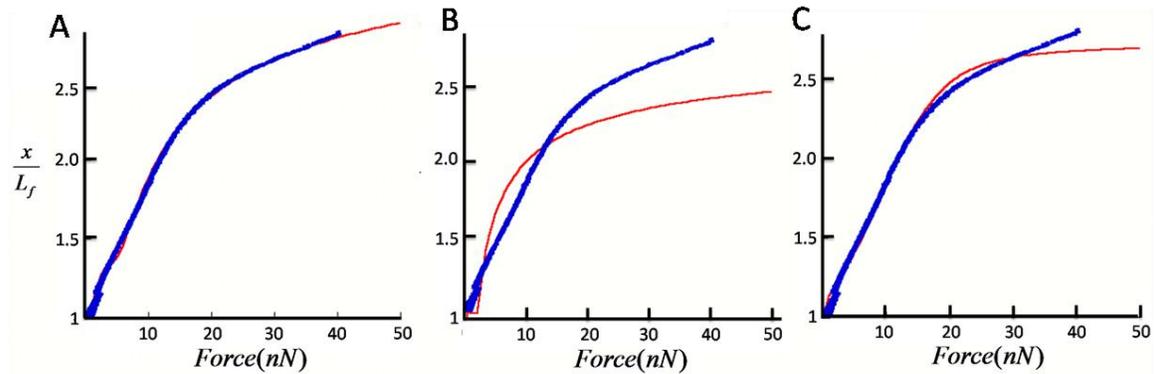


Figure 2.9 Fitting Single Fiber force vs. extension data with the Coiled-coil model. Plot A shows data assuming a coiled-coil modulus of 2MPa while plots B and C assume a coil modulus of 0.8GPa.

The SAXS data presented in the Brown et al. study are compelling, and do suggest coiled-coil unfolding may be occurring, however the data may also be explained by protofibril sliding within the fiber, causing the protofibrils to lose registry and thereby widen the 23nm signal. Interestingly the authors note that the fact the peak stays at 23nm indicates that a portion of the molecules stay folded giving rise to the 23nm signal, even though the width of the signal has widened.⁽⁷⁴⁾ It is not obvious how FXIII ligation would affect this signal. There are known ligation sites both in the α C domains and γ -domains of the fibrin monomer. (30, 109) γ - γ ligation has been shown longitudinally within protofibrils, but to our knowledge, not laterally between protofibrils. α C- α C and α C- γ interactions could be within protofibrils or between adjacent protofibrils, but it is not obvious that ligating these interactions would prevent protofibril shearing. Thus, with the known ligation sites within the fibrin fiber, it does not seem that FXIIIa crosslinking would necessarily prevent protofibrils from sliding and accounting for the decrease in the 23nm signal in the SAXS data.

2.3.7 Force per Monomer

AFM force spectroscopy studies over the past decade have established the range of forces required to unfold secondary and tertiary protein structure. (98, 101, 122). For protein unfolding to occur within strained fibrin fibers, the resolved forces (per monomer) would have to exceed these thresholds. To address this issue, we present estimates of force per monomer at all strains of fibrin fiber.

For our stress vs. strain curves, both the force vs. extension as well as the fiber diameter data were collected. With this information we made estimates of the average force per molecule (FPM) during the stretching process and compared to known protein unfolding forces determined in AFM force spectroscopy experiments. This estimate is subject to an assumption of the packing density of the monomers within the fiber. We calculated the FPM for three packing densities corresponding to a high “close packing”, a low “loose packing”, and an intermediate “intermediate packing” packing density. FPM calculations were made by assuming that the fibrin fiber consists of fibrin molecules acting mechanically as springs in series down the length of the fiber and in parallel across the cross-section of the fiber. FPM is then equal to total force on the fiber divided by N_{mol} , the number of fibrin molecules in parallel. For the first model, we assumed that fibrin monomers are close packed within the fibrin fiber, thus the number of monomers in parallel can be calculated by $N_{\text{mol}}=(D_{\text{fib}}/D_{\text{mol}})^2$, where D_{fib} is the diameter of the fiber and D_{mol} is the diameter of the monomer from x-ray crystallography, $D_{\text{mol}} = 4.5$ nm. It is widely assumed that fibrin fibers contain mostly water (70-80%), so close packing is likely not accurate for a full fiber (43, 123). The second “loose-packed model” followed that of Brown et. al (74), which assumed fibrin fibers contain 80% water, so the number

of monomers in parallel were calculated as $N_{\text{mol}}=0.2*(D_{\text{fib}}/D_{\text{mol}})^2$ (This makes the effective diameter per monomer, $D_{\text{mol}}^{\text{eff}}=10\text{nm}$). This second model assumes that the entire monomer as a space filling cylinder into which no water can penetrate. In actuality, the fibrin molecule is dumbbell shaped with 4.5nm diameter end domains, 2nm diameter coiled-coils, and a 4nm central domain (12). Based on these dimensions, we estimate that nearly 50% of the volume occupied by a fibrin protofibril consists of water. Thus, although most of the fiber is occupied by water, the majority of this water can fit within close-packed monomers. One can estimate an effective monomer diameter required to account for the other 20-30% of water occupying the fiber structure by assuming the cross-sectional area of a protofibril is roughly an ellipse, and estimating the effective area of the ellipse required to allow for 80% water volume. This ratio of areas can then be converted into an (upper end) effective monomer diameter by assuming that each monomer within the protofibril has the cross-sectional area of a circle. We estimate that an effective fibrin monomer diameter of 6.9 nm will account for the rest of the space needed for water within the fiber. Thus our third model is an “intermediate packing” model where the number of monomers in parallel is calculated by $N_{\text{mol}}=(D_{\text{fib}}/D_{\text{mol}}^{\text{eff}})^2$, where $D_{\text{mol}}^{\text{eff}}=6.9\text{nm}$. In each model, the force per monomer can then be calculated as $F_{\text{fib}}/N_{\text{mol}}$.

Figure 2.10 shows the FPM plots for one fiber, assuming the three different models. The different assumptions about packing lead to significant differences in the scale of the force on each monomer. The very lowest force we would expect a coiled-coil or globular protein domain to unfold would be in the 30-50 pN range (98, 103). Previous work has shown that the fibrin coiled-coil may unfold at even higher forces at around 100pN (74,

75). (See Figure 2.11) Note that even in the loose-packed case, these forces (the lower end of the range) are not reached until the fiber is strained beyond 0.50. For the intermediate and close packed models, this threshold is reached at strains of 1.0 and beyond. For the intermediate packing case, the fiber strain at which the 100pN coiled-coil unfolding threshold is reached is $113 \pm 23\%$ for ligated fibers and $108 \pm 15\%$ for unligated fibers. At 50% fiber strain, in the linear fiber regime, the average FPM value was 27 ± 8 pN for ligated fibers and 18 ± 4 pN for unligated fibers, suggesting that coiled-coil unfolding is not a prominent process in the linear strain regime. While coils have been seen to unfold at forces as low as 50pN, on average it seems that most fibers do not have much coil unfolding until after 100% strain.

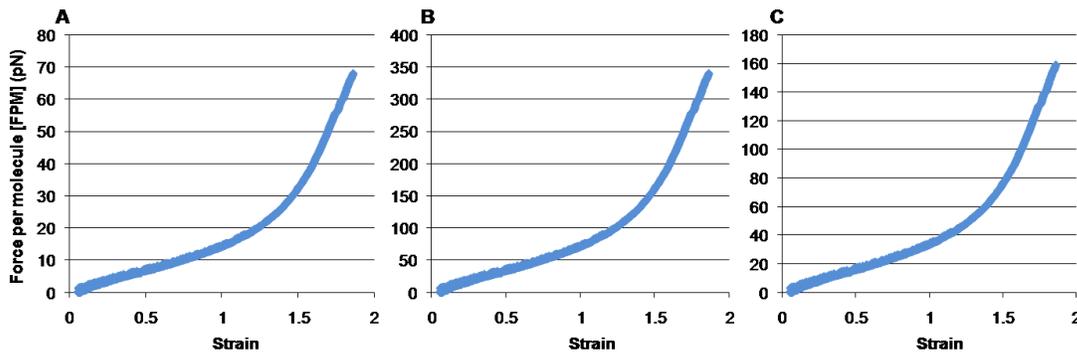


Figure 2.10 Force per molecule (FPM) plots for one fiber with three different packing assumptions. Frame [A] corresponds to a close-packed model fiber ($D_{\text{mol}} = 4.5$ nm), frame [B] corresponds to the loose-packed model ($D_{\text{mol}}^{\text{eff}} = 10\text{nm}$), and frame [C] corresponds to the effectively close-packed model ($D_{\text{mol}}^{\text{eff}} = 6.9\text{nm}$).

Our analysis indicates that though the FPM in our fibrin stretching experiments may reach the thresholds for unfolding the coiled-coil, these forces are not accessed until the fiber is strained beyond 0.5 for even the most conservative model of fibrin monomer packing (loose-packed). Most likely, this threshold is reached well beyond 0.5 strain. This suggests the unfolding mechanism is not a significant contributor to extension until very high strains are reached.

One counter-argument to this analysis that could be made is for concentrated forces within the fiber. Concentrated forces will make the FPM much higher locally than the average, inducing unfolding and then propagating across the fiber in a zipper like manner. While this may be occurring, there is no known structural data to support this model. We do point out that even if fibrin fiber diameter shrinking is occurring as indicated by other studies(74) , the number of monomers connected in parallel, down the fiber will remain constant and thus $FPM = F_{\text{fib}}/N_{\text{mol}}$ should remain the same at all points along the fiber. Additionally, some have claimed that the strain rates of the single molecule work were different than that in our whole fiber measurements, leading to inflated unfolding forces, however, we note that in our measurements the local strain rate – which is known to affect measured unfolding forces - is comparable to typical single molecule force spectroscopy measurements. (124)

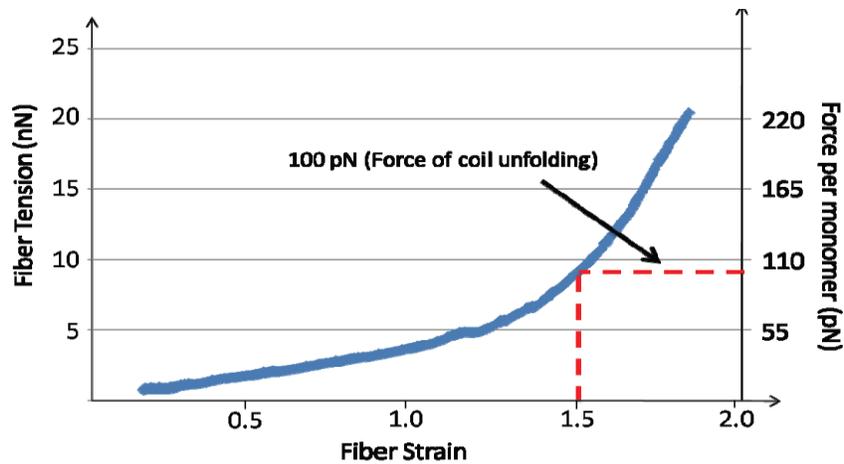


Figure 2.11 Force per monomer as a function of fiber strain: Assuming the loose-packed model, the force per monomer can be calculated over the entire fiber extension. The red line indicates the experimentally observed critical force for coiled-coil unfolding.(75)

2.3.8 Fibrin as Elastomeric Protein

The persistence length and contour length parameters that emerge from our model fitting together with the soft (MPa range) elastic modulus, strain stiffening, and

high extensibility of individual fibers presents a body of evidence that strongly suggests that the origin of fibrin's mechanical properties lie in the straightening of natively unstructured polypeptides. If this is the case, the tensile elasticity of the fibrin fiber itself is entropic in origin, mediated by straightening of a randomly coiled polypeptide, rather than enthalpic and mediated by straining of chemical bonds within the backbone. The entropic elastic behavior comes not from thermal fluctuations of the fiber segment's overall contour, but from internal degrees of freedom residing in the unstructured peptide sequences within the monomer. The argument has been made in the literature that the architectural parameters for fibrin gels (branch point density in particular) are not consistent with rubber like elasticity even though fibrin's extensibility and stiffness is comparable to rubber (63, 125). This is because unlike in rubbers where very stiff (GPa-scale tensile modulus), very thin, polymer molecules or polypeptides are polymerized into a complex highly crosslinked random coil networks, fibrin gels are made up of thick structurally complex fibers with very low branch point densities. Instead, we propose that for fibrin, the fibers themselves are rubber-like; the fiber is itself a flexible polymer network.

2.4 Conclusions

We have performed a full stress vs. strain evaluation of individual fibrin fibers, revealing elastomeric mechanical properties including low modulus (MPa), and high extensibility, and strain stiffening behavior. We found that FXIIIa ligation roughly doubles the stiffness of fibrin both in the low and high strain limit. This underlines a crucial point in fibrin mechanics: in addition to any architectural effects FXIIIa ligation confers to the overall fibrin network such as fiber diameter and branch point density, it

also directly affects individual fiber segment stiffness. We also point out that an increase in stiffness must come from the α C region of the molecule, not γ - γ ligation. We have proposed the first mechanical model of the fibrin fiber based on the stretching of the α C region of the molecule. The model depicts the fibrin fiber as a set of parallel chains of monomers linked in series. Each monomer consists of a soft WLC element representing an unstructured region of the fibrin protein monomer. Our force vs. strain data is fit well by the scaled WLC model and indicates that the unstructured portions of the monomer mediate the mechanical response of fibrin fibers.

Chapter 3. Fibrin Fiber Recoil Dynamics

3.1 Introduction

In the previous chapter we fully characterized the mechanical properties of fibrin fibers and developed a mechanical model for their elasticity. Furthermore, we compared our α C stretching model with the CC model, the other prominent hypothesis for the elastic behavior of fibrin fibers. We now turn our attention to a property perhaps even more pertinent to the survival of blood clots: fibrin fiber dynamics. The heart pumps blood at approximately 1Hz, so fibers must regain their full elasticity prior to the next pulse or risk plastic deformation and eventual failure over the course of several load cycles. It is crucial, then, to have an understanding of fibrin's dynamic response to force. In addition, measurements of the dynamic response of fibrin fibers can speak to the molecular underpinnings of the property. The two competing models for fibrin elasticity come from either protein unfolding or the entropic relaxation of a natively unfolded protein. In the CC model, the coiled-coil region must refold to re-gain the full elasticity of the fiber, while in the other model an entropic recoil without re-folding would be expected (See Figure 3.1).

We report measurements of fibrin fiber recoil dynamics, showing that fibrin fibers behave like rubber bands, recoiling on timescales of $\sim 500\mu\text{s}$, and tensing back into an elastically taut conformation on millisecond timescales. We probed the molecular origins of this behavior using discrete molecular dynamics simulations. We find that α -helical coiled-coil portions of the molecule unfold into an energetically stable β -sheet

conformation. If this does in fact occur upon fiber stretching, it would prevent a μs recoil response, however without experimental corroboration, we cannot rule out a model of coiled-coil unfolding without forming β -sheets, entropically relaxing and slowly re-forming coiled-coils on μs timescales. We conclude then, that the recoil dynamics are most likely governed by the unstructured αC regions of the molecule, however further experimental validation is required.

The dynamic behavior of fibrin fibers is quite surprising given the distinct 23nm banding pattern shown in SEM and TEM images indicating lateral contacts between protofibrils and a crystalline packing of the fiber (See Figure 3.1).(47) One could argue that this places fibers in a unique materials category...that of a crystalline rubber.

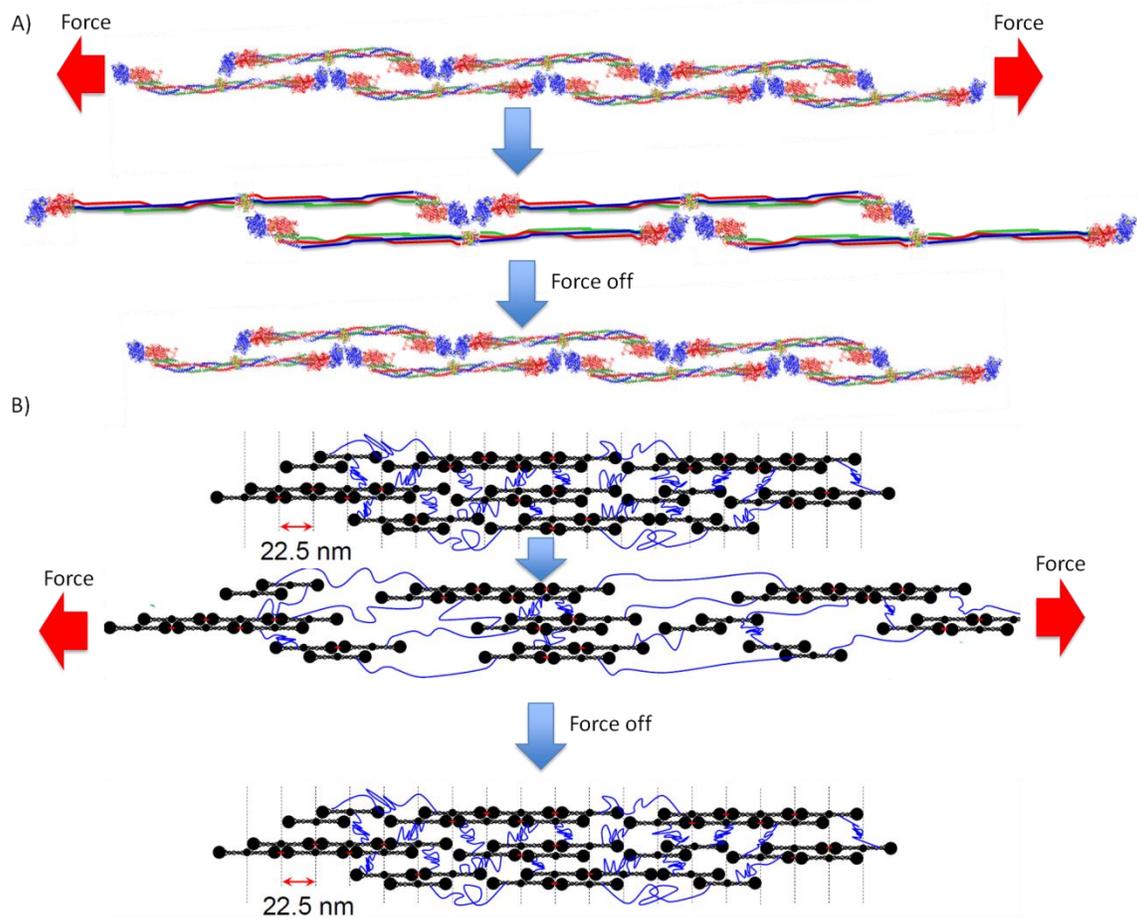


Figure 3.1 Fibrin Fiber Recoil Models Based on the Mechanism of Extension: (A) The Coiled-Coil (CC) model involves stretching of the coiled-coil region of the fibrin molecule. Upon the release of force, the model would predict that the coiled-coil region would refold. (B) The α C model involves stretching of the α C region of the molecule. Upon the removal of force the α C regions would be expected to entropically relax and allow the lateral connections to regain their periodicity.

3.2 Materials and Methods

3.2.1 Optical microscopy

Experiments were performed using a combined inverted optical microscope/Atomic force microscope as described in Ch. 2.2. (78) Fluorescently labeled fibers were suspended between ridges with a 20 μ m pitch and then stretched from the center with an AFM tip. Fibers were allowed to slip off the tip and fiber relaxation was measured optically. Because the timescales of relaxation were faster than the frame-rate of our current camera, an Evolve 128 camera (Photometrics, Tucson, AZ) was generously loaned to us by Photometrics. We measured the frame rate of the camera to be 120-150 frames per second (fps) in full frame mode, and up to 4,000 fps in region of interest (ROI) mode which only measures the photon counts in a certain region of the detector.

3.2.2 Discrete Molecular Dynamics

DMD simulations were carried out on the Biomedical Analysis and Simulation Supercomputer (BASS) at the University of North Carolina at Chapel Hill. The basics of DMD are discussed in Appendix A.5. Simulations first stretched the coiled-coil structure at constant force causing it to unfold into a stable β -sheet conformation which is described in more detail in chapter 4.4. To test the stability of this conformation, we ran additional simulations on this region after the force was turned off. The initial conformation used for these relaxation simulations was the final conformation of the

coiled-coil structure after 10ns of DMD simulation under 700pN of force (See Figure 3.4 top image for the initial conformation).

The coiled-coil structure was extracted from PDB structure 3GHG.⁽¹²⁾ Specifically amino acids α 27-200, β 58-198, and γ 14-130 from 3GHG were used to generate the initial structure for the simulation. Each chain of the coiled-coil structure terminates consists of 111 amino acids surrounded on each end by one (or more) Cystein residues. Di-sulfide bonds, between α Cys161- γ Cys135, α Cys165- β Cys193, and γ Cys139- β Cys197 on the C-terminal end and α Cys45- γ Cys23, α Cys49- β Cys76, and β Cys80- γ Cys19 on the N-terminal end were incorporated as described in chapter 4.4.1 Methods. The simulations were performed in a rectangular box of dimensions 100Åx100Åx2000Å with periodic boundary conditions. An Anderson thermostat (See Appendix A.3.2) was used to maintain room temperature (300K).⁽¹²⁶⁾

3.2.3 Fibrin Polymerization

Fibers were polymerized across channels at the same concentrations and in the same manner as described in Ch.2. Fibers were formed with and without FXIIIa cross-linking. Additional experiments were performed in the presence of ethylenediaminetetraacetic acid (EDTA) and after glutaraldehyde cross-linking. In this document only the data from non-ligated fibers is presented, but the data analysis from the additional experiments is ongoing.

3.3 Experimental Results

Fibrin fibers were non-specifically attached to an AFM tip by bringing the gold-coated tip in contact with the fiber surface. The tip was then moved in a direction away

from the fiber (See Figure 3.2) until the fiber detaches from the tip and recoils back into its taut conformation.

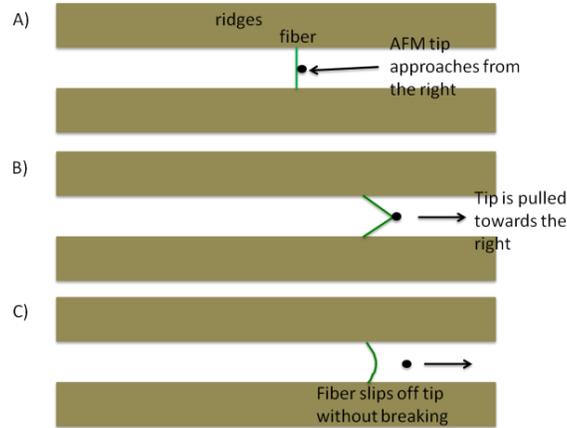
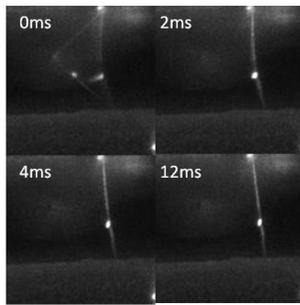


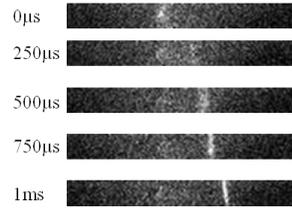
Figure 3.2 Experimental Setup: (A) The AFM tip is brought towards the fiber on the right side. Once in contact with the fiber, the tip is moved back to the right (B) before eventually slipping off and recoiling (C).

A) Full Frame Recoil

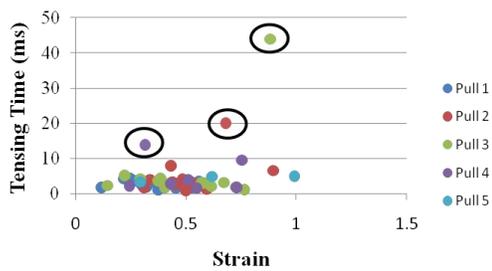


B)

Region of Interest (ROI) Recoil



C) Tensing Time vs. Strain (by pull)



D)

Tensing Time vs. Strain (by fiber)

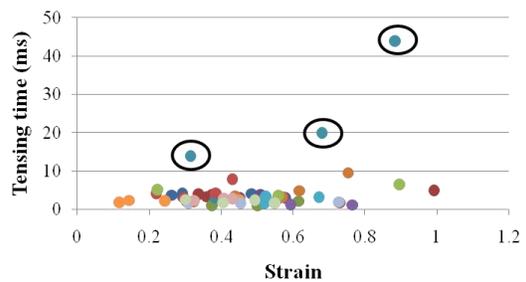


Figure 3.3 Fibrin Recoil Dynamics Captured with High Speed Fluorescence Microscopy: (A) A fiber suspended between two ridges and stretched by an AFM, slips off the AFM tip and recoils. The top left image shows a fast recoil which cannot be captured at the full frame camera frame rate, followed a recovery of tension within 10ms. B) Imaging only a region of interest allowed frame rates of 4000fps, enabling the initial recoil event to be captured. (C-D) The tensing time, defined as the decay constant of an exponential fit to the recoil data, was measured by plotting the average position of the center of the fiber vs. time. (C) Each fiber was pulled a minimum of three times, and the tensing time was measured for each recoil. Data is separated by color indicating whether the data point came from the first fiber pull, the second fiber pull...etc. Based on the data, there is not a direct correlation between tensing time and the number of times the fiber is pulled, indicating that fibers are not plastically deformed over a series of stretch and relax cycles. (D) Recoil was also analyzed on a per fiber basis to test whether certain fibers tensed on a characteristic timescale. Each color in plot “D” represents data taken from one fiber. The circled data points represent one fiber with a slower tensing time, but otherwise most fibers appeared to behave in a similar manner. Additionally, both (C) and (D) show no direct correlation between the tensing time and the fiber strain at the point when the fiber began relaxing.

In full frame mode, most fibers were observed to detach from the tip and regain their full tension within 5ms (Figure 3.3A). The initial time of the recoil could not be captured in full frame mode, so additional videos were taken using only the fiber center in a region of interest. ROI imaging revealed that the first part of the recoil occurs on timescales of approximately 500 μ s, which corresponds to a recoil strain rate of $\approx -4000\text{s}^{-1}$, while the average strain rate over the entire relaxation is 12s^{-1} . Fibers generally tensed back into their taut conformations within 5ms as seen in Figure 3.3.

. The tensing time (τ) was defined as the mean decay time of the fiber and was calculated for each fiber by fitting an exponential function to the fiber center position (R) as a function of time: $R=R_0\exp(-t/\tau)$. The most strained fibers observed in this set of experiments reached strains of 100%, and a full recovery of elasticity (based on a consistent tensing time) was observed upon repeated stretching cycles of the same fiber. There does not appear to be a correlation between τ and fiber strain at the point of detachment up to 100% strain. It should be noted that in previous experiments, plastic

deformation of fibers has been observed at strains above 120%, however due to the non-specific interactions between the fiber and tip in these experiments, fibers did not reach the point of plastic deformation before detachment. (66)

3.4 An α -helix to β -sheet transition

To probe the molecular origins of these properties, we used discrete molecular dynamics simulations to stretch coiled-coil region of the fibrin molecule under constant force. (See 3.2.2 Discrete Molecular Dynamics) We found that as the coiled-coils unfold, they form hydrogen bonds with adjacent chains, and form a stable β -sheet structure. (See Figure 3.4 and Figure 4.8). The unfolding force was then turned off, and the structure was allowed to relax. Instead of entropically collapsing into a molten globule state and re-folding into a coiled-coil structure, as would be expected from the coiled-coil stretching model, the β -sheets appeared to lock the chains into an energetically stable position that prevented immediate re-folding. The β -sheet rich structure was energetically stable for 100ns of simulation time. As a comparison, the single gamma chain of the coiled coil was stretched to its full contour length and then released. The chain immediately relaxed, collapsed, and began re-forming an α -helix. (See Figure 3.4)

This is not the first α -helix to β -sheet transition observed in nature. Astbury in his seminal work which actually defined the α -helix and β -sheet structures in X-ray signals, noticed that the keratin double helix coiled coil undergoes this transformation in wool and hair fibers after stretching.(127) This transition was re-characterized by Kreplak et al. who showed that in keratin, first the α -helices must unravel starting at 5% strain, and β -sheet formation does not occur until 20% strain, but between 20% strain and 40% strain both forms are present in the fibers.(128, 129) Interestingly, the same authors earlier

showed that keratins stretched in only 30% relative humidity did not form β -sheet structures after the disappearance of the α -helix signal indicating the presence of water is important in β -sheet formation.(129) An impressive study on hydrated hagfish slime, a keratin based material, showed that threads made of this material undergo an irreversible α -helix to β -sheet transition above strains of 30% leading to plastic deformation. (120) In conclusion, it seems that hydrated keratins undergo an irreversible α -helix to β -sheet transition at strains above 30%, while keratins not in the presence of water do not necessarily make the transition. Recent MD studies on keratins such as vimentin also indicate an irreversible α -helix to β -sheet transition.(114, 130) In one study, the authors predict that the α -helix to β -sheet transition is a universal feature of double helix coiled-coil structures greater than 40 amino acids.(130)

Single molecule protein refolding studies have offered a mixed bag of results. One impressive study showed that the 768 amino acid double helix of muscle motor protein myosin II can elastically stretch and refold when stretched on timescales of 2-7s. FRET and AFM studies on the coiled coil structure of the GCN4 leucine zipper show that it can refold on ms timescales after unzipping, however the direction of applied force in the AFM studies would seem to pull the coils apart preventing β -sheet formation.(131, 132) AFM measurements on 106 amino acid triple helix spectrin repeats indicated that in repeating load-unload cycles a fraction of the coiled-coil structures could re-fold on the order of 1s, but many did not re-fold(98). Finally, in perhaps the most relevant work to this paper, fibrin protofibrils were repeatedly stretched on timescales of a few seconds. It is not exactly clear what was unfolding in those measurements, but the authors interpreted it as coiled-coil unfolding and saw that several of the force peaks were present

in multiple extension-retraction curves indicating a partial refolding of some of the molecules.(68) Thus, the timescales of coiled-coil refolding from single molecule experiments do not necessarily rule out an α -helix to β -sheet transition, but they do not provide evidence that the a stretched coiled-coil would recoil on the order of μ s either. Conclusive single molecule work on fast-folding globular proteins such as the Ig domain have shown refolding times on the order of 40ms for repeated stretching cycles.(133) This likely rules out the re-folding of the fibrin γ -nodule or β -nodule as a mechanism of relaxation as they are rather large (~300 amino acids) globular regions and would not be expected to re-fold faster than the 100 amino acid Ig domain described above.

Our MD results combined with the experimental keratin data, MD simulations of other coiled-coil structures, and the single molecule measurements on coiled coil structures indicate that if the fibrin coiled coil does undergo an α -helix to β -sheet transition, the transformation is likely irreversible upon the release of force. Even in single molecule studies where an α -helix unfolding event did occur, the refolding was often only observed in a fraction of the molecules after several seconds of relaxation time.(68, 98) This would indicate that stretching of the coiled-coil cannot account for 500 μ s elastic response time of the fibrin fibers, and is likely not the predominant mechanism of extensibility within the fibrin fibers. This analysis is not complete however, and further studies should be carried out to measure whether fibrin does truly undergo an α -helix to β -sheet transition.

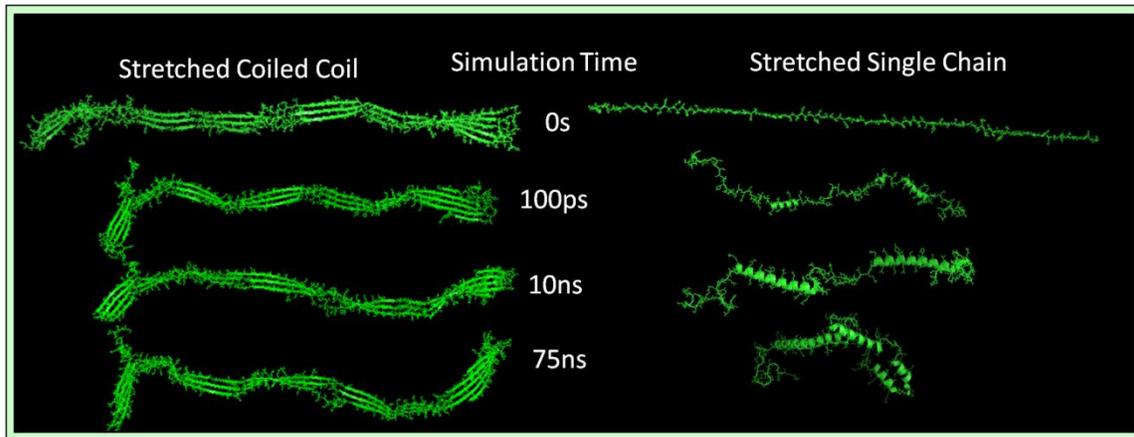


Figure 3.4 β -sheet Relaxation: Left: Simulations of the β -sheet form of the fibrin coiled-coil. The β -sheet form is obtained by stretching the coiled-coil at constant force. Once obtained, the β -sheet appears stable, and prevents, an immediate recoil. Right: Simulations of the γ chain of the coiled coil part of the molecule. Contrary, to the β -sheet structure, just the γ chain in isolation, immediately collapses and begins re-forming an α -helix structure. The collapse happens faster than it would physiologically due to the use of an implicit solvent in both simulations, and hence no viscosity. Nevertheless, the distinct difference in recoil behavior between the β -sheet structure and the γ chain in isolation indicate, that the β -sheet structure is not the source of the recoil dynamics properties of the fibrin fiber.

3.5 Conclusions

The cumulative evidence from experiments and simulations suggests that the most likely explanation for the initial $500\mu\text{s}$ relaxation followed by tensing on the order of 1-10ms comes from the entropic recoil of the αC region of the fibrin fiber in accordance with the αC stretching model (See Figure 3.1).(78) Mathematical modeling of the recoil behavior using polymer recoil dynamics (Rouse and Zimm dynamics) is the next step in the analysis, and preliminary work is ongoing. We will treat the fiber as a series of cylinders (representing the protofibrils) connected by polymer chains. The relaxation timescales will be calculated from the entropic restoring force of a stretched chain and the intrinsic viscosity of the polymer chain combined with the drag generated by the cylinder.

In previous work, the stress relaxation of fibrin fibers has been measured, where the fibers were stretched to a certain strain and then held constant.(67) In stress relaxation, the fibers also exhibited two relaxation times, however the times were on the order of 2s and 50s, indicating different mechanisms than those observed in our strain relaxation experiments.

The relaxation dynamics of fibrin fibers continue the mounting evidence that fibrin fibers are elastomeric in nature. Most other elastomeric materials are also comprised of networks of unstructured polymer chains. (92, 134) The most classic example is rubber, which consists of cross-linked hydrocarbon chains, and can recoil at average strain rates of -41s^{-1} in water.(135) Biopolymers such as elastin and resilin also consist of disordered networks of amino acid chains.(134, 136) The difference between fibrin, and these materials is that fibrin fibers may have a crystalline packing, while other elastomers start as a random array of unstructured chains. The trick with fibrin is that the part of the protein giving rise to the elastic properties, the αC regions, are independent from those forming the crystalline packing (the structured coiled-coil and D regions).

One type of material that also contains crystalline structure while also having large extensibilities is spider silk.(137-139) Spider silk is different than fibrin however in that it contains β -sheet nanocrystals embedded in a soft matrix of disordered material. Stretching of both viscid and dragline silks indicated that spider silk thread stretching was not reversible on ms timescales; rather silk threads after stretching resume a taut conformation if left slack for about 10 minutes.(92) Fibrin fibers are therefore unique elastomers. In an unstrained state, they have a regular packing giving rise to banding patterns across the fiber. At the same time, between the crystal-like structure, they

contain networks of disordered amino acid chains, allowing fibers to stretch to triple their length and resume taut conformations on ms timescales. The timescales observed in these recoil dynamics experience could be explained by: 1) an entropic recoil of the stretched α C regions, followed by a re-formation of the lateral contacts between protofibrils allowing the fiber to regain its crystalline packing. The plastic deformation observed in stretching fibers beyond their elastic limit may come from stretching the ordered parts of the fibrin molecule such as the γ -C domain as explained in Chapter 4.

. We conclude that fibrin fibers behave like rubber bands in their recoil behavior which is remarkable for a material with the internal order of a fibrin fiber. They contain unstructured polymer networks that can allow rapid stretching and recoiling on microsecond timescales. Unlike rubber bands however, they contain additional structured regions which must re-form contacts on ms timescales.

Chapter 4. Unfolding Energy Barriers of the Coiled coil, γ - γ and α C region of the Fibrin Molecule

4.1 Beyond the α C region

The previous two chapters describe a variety of fiber properties that ultimately must be explained on the molecular level. Fibrin fibers are highly extensible, exhibit strain stiffening behavior, stiffen when ligated by FXIIIa, recoil on μ s timescales, and regain tension on ms timescales. The evidence so far indicates that many of these mechanical properties arise from an unstructured part of the molecule, likely the α C region. However, the exact structure of the α C region is not known, making it difficult to devise the mechanism by which the extension can arise from this region. In addition, given that the α C regions allow extension between protofibrils (See Figure 2.5), the amount of fiber strain generated from this mechanism is related to the ratio of protofibril length to α C length. The distribution of protofibril lengths within a fiber has never been measured. Distribution of protofibril lengths during polymerization have ranged from 1 to 20 or more monomers which (assuming a half-staggered arrangement) corresponds to protofibril lengths ranging from 45 to 450nm or greater.⁽⁹⁹⁾ Given that the maximum length of the α C region is 410Amino Acids* 0.3nm/AA=123nm, it is likely, at least at higher strains, that additional regions of the fibrin molecule stretch in order to reach fiber strains of > 200%.

To predict potential other mechanisms of fibrin extension and to probe the relative stability of the various parts of the fibrin molecule, we have used Steered Discrete Molecular Dynamics (SDMD) simulations to assess the relative stability of the α C

region, the γ -C region, and the coiled-coil region. We started by generating a homology model for the α C domain, and then tested the stability of that region using a temperature scan and forced unfolding.⁽¹⁴⁰⁾ We proceeded to isolate the coiled coil region of the molecule, and the γ -nodule (sometimes referred to in this document as the γ -C domain), and the γ - γ interface between molecules for molecular stretching simulations using DMD. A series of simulations were performed stretching each region of the molecule at forces ranging from 5-1000pN. The critical force for unfolding was calculated for each region by assessing unfolding times at various forces. In so doing, we were able to measure the relative stability of each region of the fibrin molecule, and propose a complete model for fibrin fiber extension.

4.2 Developing a Structural Model for the α C Region of the Fibrin Molecule

4.2.1 Is the α C region truly unstructured?

The structure and function of the fibrin(ogen) α C region has been subject to much debate over the past 40 years. ^(13, 14, 26, 141-145) Originally, Doolittle proposed that these structures were largely unstructured and acted as “free swimming appendages.” ^(4, 142) This view was supported by earlier circular dichroism studies indicating that the α -C plasmin degradation product was a random coil and by amino acid sequence comparisons revealing a similarity between the α C region and other unstructured proteins. ^(146, 147) However, this work was contradicted by some calorimetric and EM studies which indicated that part of the α -C region contained compact structures. ^(143, 148-150) In 1983, Erickson and Fowler showed EM images indicating that the fibrinogen α -C region has some globular structure and is connected to the central region of the molecule.⁽¹⁴⁸⁾ This globular structure was dubbed the α C domain. This work has

been followed up with further EM studies of fibrinogen, fibrin, and various fragments, which show a structured portion of the α C domain in a high percentage of the molecules imaged. (150, 151) Recent NMR studies have indicated that recombinant truncated variants of both bovine and human fibrinogen α C-domains form beta sheet structure. (14, 152, 153) However, a debate still exists as to whether an ordered structure is present in the fibrin(ogen) α C region due to the lack of an electron density in crystal structures.(13)

EM images of fibrinogen, indicate that the α C domains interact with each other intramolecularly and are positioned near the central region of the molecule.(151) It has now been established that α C domains of fibrin molecules also interact intermolecularly through non-covalent interactions. (73, 152, 154) These interactions are concentration and pH dependent, and a recent laser tweezers study indicates some interactions between α C-domains that could withstand forces of up to 50pN at pH 7.4.(73, 152, 154)

Fibrin α C domains also interact via FXIIIa cross-linking to form networks termed α -polymers. FXIIIa is another blood protein that acts as a transglutaminase, catalyzing the formation of a covalent bond between the free amine group of a Lysine (Lys) and the gamma-carboxamid group of a glutamine (Gln). Each α C domain contains 23 potential Lys donor sites and 6 Gln sites, however not all are equally used in cross-linking. Sobel et al. showed that the primary Lys sites are Lys556 and Lys580, and Cottrell and co-workers indicated that the primary Gln sites are α 328 and α 366, indicating that some sites are more accessible to FXIIIa than other.(30, 155) The mechanism for this selection does not seem well understood.

To address these issues, we have undertaken a study of the α C region structure using discrete molecular dynamics modeling. We began by searching for a homology

model of the α C region using I-TASSER.(156, 157) I-TASSER is a web-based protein structure prediction tool, which is able to find homologs and even remote homologs (with relatively low sequence similarity). Using these identified homolog and remote homolog structures, a structural model of the query sequence was built. The model was then energy minimized and thermally unfolded to define the equilibrium structure at various temperatures. The resulting structure, which was not constrained by any of the previous, EM, calorimetry, or NMR data, is yet remarkably consistent with the aforementioned data. The model consists of an unstructured α C connector region, terminating in a β -helix structure which contains the di-sulfide bond. Beyond the β -helix is a second unstructured region, containing the most prevalent Lys donors for cross-linking. This structure is consistent with the binding and cross-linking requirements of the α C region and also suggests a possible mechanism by which FXIIIa can select residues for ligation.

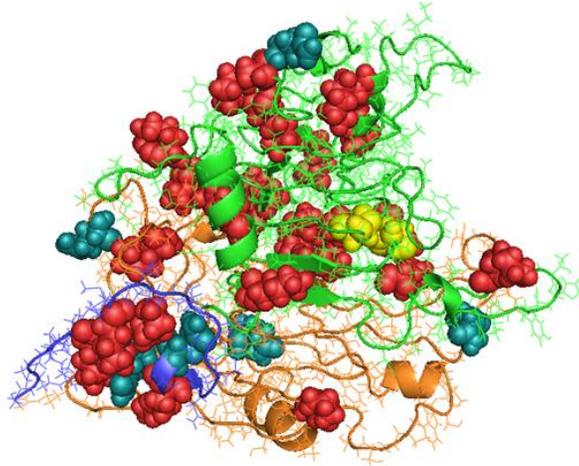
4.2.2 A computationally Identified β helix structure of the fibrin α C region

A series of homology models were generated for amino acids 196-610 of the fibrin α C domain using the online modeling resource I-TASSER.(140) One model of particular interest showed a β -helix structure for the α C region and within the structure Cys442 and Cys472 were within 1nm of each other indicating a potential for di-sulfide bond formation. Following homology modeling, the structure was relaxed using 400 steps. The di-sulfide bond was formed by adding constraint potentials between the Cys442 and Cys472,(158) and the system was allowed to relax for 10ns of DMD simulations at 276K. The system was then allowed to evolve for an additional 10ns of DMD simulations at 300K after di-sulfide formation.(159) Figure 4.1 shows the resulting

structure after equilibration at 300K along with the homology derived structure. The equilibrium resultant structure displays several remarkable structural features.

In Figure 4.1, the α C region has been separated by color to show several distinct regions. Amino acids 196-240 are seen folding back on each other, however in simulations of the whole fibrinogen molecule where α 196 was attached to the 4th coil, the bonding was not present. The repeat region, amino acids 264-391, displays a random coil arrangement, with no discernable fold, in agreement with previous studies.⁽¹⁴⁵⁾ A novel structure is seen between amino acids 392 and 520, where a partial β -helix structure is formed. The helix measures approximately 2nm in diameter, and appears destabilized on the side of the di-sulfide bond. The α C domain then terminates in a second random coil region containing the most active transglutaminase sites, Lys556 and Lys580, which account for 50% of all FXIIIa-catalyzed cross-links.

A)



B)

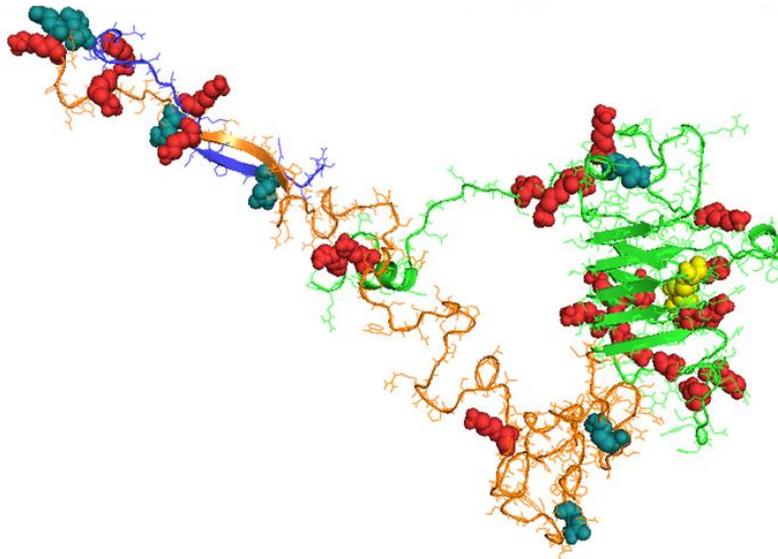


Figure 4.1 Structure of the fibrin α C region: A) The homology derived structure. B) The equilibrated structure at 300K. The blue region consists of amino acids 196-220, which are not in the human fibrinogen crystal structure implying an inherent flexibility (12), but have not historically been considered part of the α C region. (13) The orange region consists of amino acids 221-391, typically called the α C connector region. (160) Within this region, amino acids 264-391 contain the series of 10, non-identical (13 amino acid each), repeats. The green region, consisting of amino acids 392-610 is traditionally called the α C domain. The Cys442-Cys472 di-sulfide bond is highlighted in yellow, while all the potential cross-linking sights are highlighted in red (Lysines) and teal (Glutamine). All amino acids noted above are represented in atomic sphere mode, while the structure as a whole is represented in cartoon mode to highlight the beta sheet structures. Lys580 and Lys556 contribute to 50% of α C cross-linking.

4.2.3 Thermal unfolding of the α C region indicates the relative stability of the β helix

We performed equilibrium DMD simulations of the α C region for 10ns over a wide range of temperatures (276-390K). We used the weighted histogram analysis method (WHAM, see appendix A.6.1 The Weighted Histogram Analysis Method (WHAM))(161) to compute the specific heat as a function of temperature (Figure 4.2 top). The peaks in the specific heat vs. temperature plot correspond to the unfolding transitions. As can be seen, the α C region of fibrin displays three distinct unfolding transitions. Comparing the unfolding transition peaks to snapshots of the final structures of equilibrium simulations at 290K, 310K, and 325K, indicates the likely sources of each peak (Figure 4.2 bottom). The first peak comes from denaturing and α helix structure that forms at low temperatures near the C-terminal end of the α C domain. The second peak at 320K, comes from the thermal denaturing of the β -helix structure. The third peak centered at 345K is smaller and likely comes from the denaturing of residual contacts left after the β -helix unfolding.

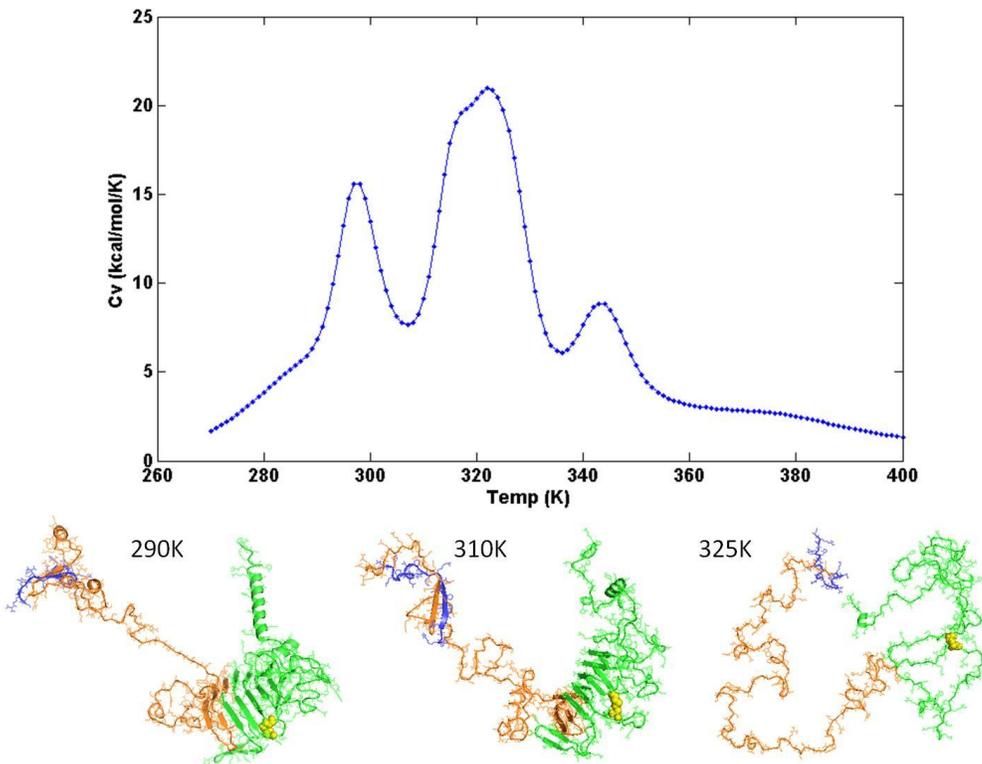


Figure 4.2 Specific Heat of the α C Structure: The specific heat of the fibrin α -C region β -helix structure was calculated by running a series of simulations at different temperatures. WHAM (See appendix A.6.1) was then used to calculate the specific heat using equation A.6.1. Peaks in the diagram correspond to unfolding transitions.

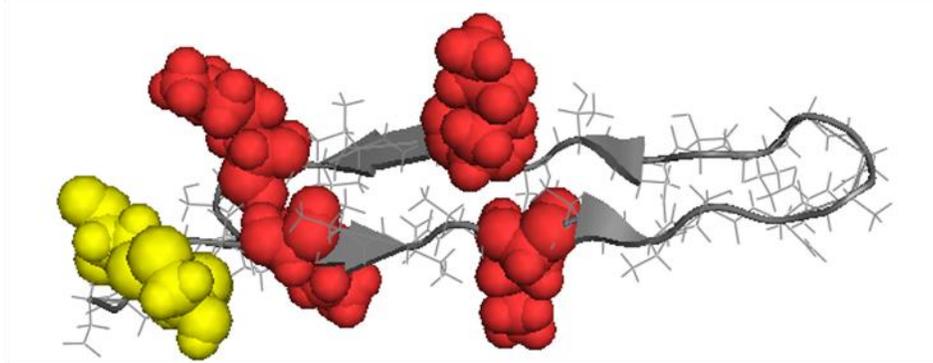
4.2.4 Comparing properties of the α C computational structure with known experimental data

Experimental evidence suggests that the α C region plays an important role in fibrin fiber polymerization and elasticity, yet the structure of the α C region has been difficult to pin down. (13, 145) In this work, we have presented a structure for the α C domain arrived at through homology modeling and molecular dynamics simulations that may solve several of the current structural conundrums. The entire α C region consists of an unstructured α C connector region (a.a. 221-391) followed by a loose β -helix structure (a.a. 392-522), and a second unstructured region (a.a. 522-610). The most interesting

part of the structure is the β -helix which makes sense of several other experimental clues, but has not been observed in crystallography studies. (12, 162)

Several groups have reported seeing “extra nodules” in EM studies of both human and bovine fibrin and fibrinogen and have attributed this structure to the α C domain. (150, 151) The location of the nodule(s) seems to vary based on sample type and pH, but they are typically seen either adjacent to the central region of the molecule or near the outer D-domains. They are not seen with every molecule imaged, but generally the nodules appear smaller than the C-termini of the β and γ chains which measure 5nm in diameter. Thus, the β -helix structure, which measures approximately 2nm in diameter is consistent with the size and existence of the extra nodules.

A)



B)

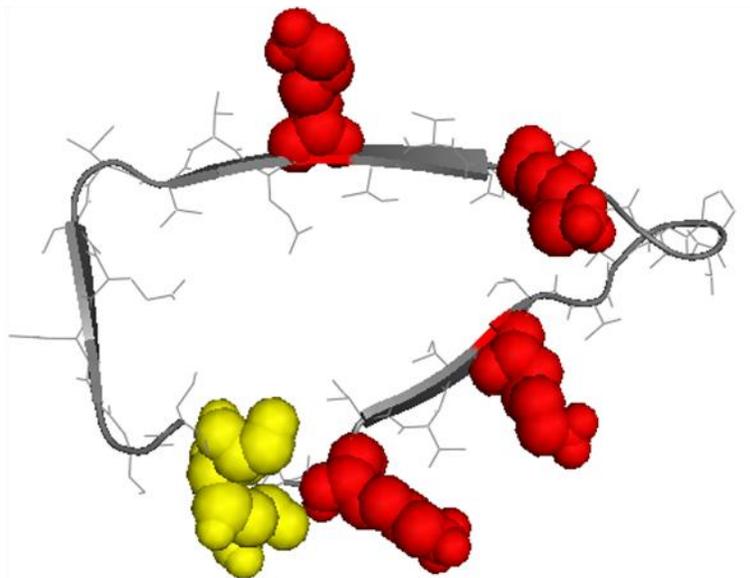


Figure 4.3 β -hairpin vs. β -helix: (A) shows the bovine fibrin β -hairpin structure found by Burton et al. in their NMR studies of truncated α C domains.(153) Between the two Cysteines(yellow), the structure folds back on itself forming hydrogen intra-chain β -sheets. (B) Identifies amino acids Cys442-Cys472 within our α C homology model. The open circle structure is stabilized by β -sheet formation with amino acids outside the immediate sequence (442-472) as can be seen in Figure 4.1.

Recent NMR data has shown that both bovine and human fibrinogen exhibit β -sheet structure in the α C domain. Work by Burton et. al in on recombinant α C fragments indicated that bovine fibrinogen exhibited a β -hairpin structure based around the disulfide bond (See Figure 4.3). (152, 153) Studies on a truncated bovine α C fragment indicated a second loose β hairpin located in the hydrophobic region beyond the di-

sulfide bond, in while a longer α C fragment indicated an undefined or flexible region in that space.(152, 153) Corresponding work on human α C fragments indicate similar β sheet structures, although with more structural flexibility. (14)

The structure for the human α C region derived using DMD shows intriguing similarity in β sheet structural features, indicating a high significance for our homology model. The structure consists of a loose β -helix structure, one side of which contains five parallel β sheets holding the helix together, while the opposite side of the helix consists of non-bonded, flexible regions, in the middle of which lies the di-sulfide bond (between β -sheets coils 2-3). In the NMR studies, hydrogen bond contacts were able to be assigned in the β hairpin structure of the bovine α C domain fragments. (152) Corresponding contacts and β sheets were not able to be identified in the human α C domain fragments using NMR due to larger flexibility of the human fragments, but the authors estimated β sheet regions using homology modeling to correspond to amino acids 444-452, 456-464, 475-480, 485-494.(14) In the β -helix structure on the other hand, β -sheets are observed in amino acids 393-397, 429-432, 434-438, 460-464, 466-470, 482-486, 488-492, 512-515, and 518-522.

The similarities and differences between the two structures (β hairpin vs. β -helix) can be examined in several ways (See Figure 4.3). First, there is a large amount of β -sheet forming potential in the α C domain (residues 392-522). Between the two structures over 50% of all amino acids participated in β sheet formation in one or the other structure. This also means that 50% of the amino acids did not participate in structural formation. This combination of β -sheets intertwined with flexible regions is possibly the

reason the structure of the α C domain has been difficult to pin down in crystallographic studies.

It is important to realize that the NMR derived β hairpin is not necessarily incompatible with the β -helix structure. In the β -hairpin, the amino acids between Cys442 and Cys 472 (the di-sulfide loop), fold back on each other and form hydrogen bonds with themselves.(152) If instead, this structure was opened up into a circular shape and allowed to interact with amino acids outside that immediate sequence, the circular structure could form into a three dimensional solenoid type shape by winding amino acids above and below the initial circle (Figure 4.3). It is possible that the α C region can form either a β hairpin or a β -helix depending upon whether the initial folding interactions stabilize it in a pinched hairpin or a circular shape. This may in part be driven by the length of the region on either side of the di-sulfide loop. The hairpin may have been more prevalent in the NMR studies due to the fact that truncated recombinant α C regions were used. (152) NMR studies on the full length α C region of both bovine and human fibrinogens were not able to deduce an exact structure. (14, 153) This also might explain why not every EM structure of the fibrinogen molecule displays an “extra nodule”, although it more likely has due do with protein-surface effects. Perhaps some form hairpins, which are less electron dense and may not be obvious in an electron micrographs, and some form β -barrels which show up as 2nm structures floating in space next to bulk molecule. β -hairpin vs. β -helix structure formation may also be pH dependent as work done on the α C region in the past has been performed over a wide variety of pH's. (145, 163)

The β -helix structure can also help explain the non-covalent oligomerization of fibrin α C domains. Previous reports have suggested that β -hairpin swapping is a relevant mechanism for α C multimerization.(152) A similar type of domain swapping mechanism could explain oligomerization with the β -helix structure. The helix consists of structured β -sheets on one side, and flexible non-bonded regions on the other side. Since these chains are not bound in the β -helix structure, but do have hydrogen bonding potential as explained above, these flexible chains would be able to form intermolecular hydrogen bonds with other α C domains, stabilizing the flexible regions of the β -helix and allowing for oligomerization. Previous circular dichroism (CD) measurements have indicated that upon oligomerization, the α C regions increase in regular structure, a result that would be expected in β -helix domain swapping.(14)

The thermal stability of the β -helix structure was tested by running a series of simulations at various temperatures. The specific heat and hence the unfolding transitions of the structure can then be deduced using the WHAM method. (161) The main peak in the C_v vs. Temperature plot for the β -helix centers around 45°C, indicating a melting temperature (T_m) between 40-50°C (See Figure 4.2). This is in good agreement with previous fluorescence and CD thermal denaturation measurements. A 2002 paper by Tsurupa et al. for the full length human α C fragment (at pH 8.0 in 150mM NaCl Tris buffer) indicated a T_m around 42°C. (145) A recent paper by the same author, indicated a melting temperature of 24°C for truncated α C fragment hA α 425-503 (conc. 6.3mg/ml, pH 7.4, 150mM, NaCl Tris buffer) and a T_m of 40°C for α C fragment hA α 392-610 (conc. 1.9mg/ml, pH 7.4, 150mM, NaCl Tris buffer).(14) Interestingly, in the CD studies, T_m of the longer fragment is 15°C higher than hA α 425-503. This number could be due to a

slightly higher number of oligomers in solution (6% vs. 15%), but also could be due to the presence of β -helix structures in hA α 392-610, not seen in hA α 425-503.

Finally it is important to understand how the β -helix model reconciles with the known binding and cross-linking sites in the fibrin(ogen) α C region. It has known binding sites for tPA, fibronectin and plasmin(ogen) as well as cross-linking sites for factor XIIIa, and α 2-antiplasmin. The tPA and plasminogen binding sites are known to reside on Lys residues within amino acids 392-610, and are only present in fibrin, not fibrinogen; however they are not competitive. (164) The β -helix structure provides a model for these binding sites as the nearly every Lys in the β -helix resides in the flexible region of the helix (Lys413, Lys 418, Lys421, Lys427, Lys444, and Lys457). α 2-antiplasmin cross-links to Lys303 in the α C connector region, and has recently been shown to non-covalently interact with the α C domain. It was additionally shown that α 2-antiplasmin was not competitive with tPA or plasminogen in binding.(165) The fibronectin binding site on the other hand is revealed in the α C connector region upon the conversion of fibrinogen to fibrin. (166) It is not clear that the β -helix structure would affect this site, unless it blocks binding in the fibrinogen state.

FXIIIa induced ligation serves to stabilize fibrin fibers and strengthen blood clots. It has previously been shown that three Gln residues (221, 237 and 328) (29) and at least 13 Lys donor residues can participate in FXIIIa reactions within the α C region, however not all Lys residues contribute in equal amounts to ligation. Lys556 and Lys580 account for 50% of Lys donors, while Lys539, Lys508, Lys418, and Lys448 contribute an additional 28%. The last 22% is split amongst Lys601, Lys606, Lys427, Lys429, Lys208, Lys224 and Lys219. (30) Because the α C region contains 23 potential Lys

donor sites, this means that an additional 10 Lysines are inert to FXIIIa. These data indicate some selection mechanism for FXIIIa, and the β -helix structure for the α C domain suggests a possible solution. Lys556 and Lys580 are the first two active Lys on the C-terminal side of the β -helix (Lys562 and Gln563 are adjacent and apparently inactive), while Lys539, Lys508, Lys418, and Lys448 are each the outermost Lys in their respective coil within the loose portion of the β -helix and line up vertically in the structure. Thus, a donor selection mechanism could be that FXIIIa first searches for the β -helix structure (in fact, a recent report has highlighted Glu396, a residue in the first coil of the β -helix, as being a key residue for the binding of FXIIIa to the α C region.(167) Upon finding the structure, the enzyme moves towards the C-terminal portion of the molecule. The first readily accessible donor sites are Lys 556 and Lys 580, which are used 50% of the time.(30) A smaller portion of the time, a donor Lys within the β -helix structure is available to be used for cross-linking. While only a hypothesis, this selection mechanism could be tested with site-directed mutagenesis.

The structure and interactions of the α C region is made more interesting by the recent proposal that the α C region, and the α C connector region in particular is responsible for the remarkable extensibility and elastic properties of fibrin fibers. (77, 78, 82, 85) If this is so, then the structure and intermolecular interactions of the α C regions will put a limit on the extension available to the fibrin fibers.

4.2.5 Conclusions

In conclusion, using homology modeling and discrete molecular dynamics, we have shown that the human fibrin(ogen) α C domain can fold into a compact β -helix structure that is thermodynamically stable at physiological temperatures. The structure

may help to solve many previously unresolved questions about the α C domain fold. The β -helix is of similar size to the “nodules” reported in EM images, contains both a structured and unstructured side which may explain the difficulty in crystallizing it, and thermally denatures at temperatures similar to those seen in previous calorimetry experiments.(14, 150, 151, 153)

4.3 The Unfolding Energy Barrier of the α C Domain

4.3.1 Methods

To further probe the stability of the β -helix structure of the α C region, we have used steered discrete molecular dynamics (SDMD) simulations to unfold the α C region of the molecule under force. Appendix A.6.2, gives an overview of SDMD techniques. To apply these techniques to the α C domain structure of amino acids 196-610 of the fibrin α -chain, the C_α carbon of Ser196 in the N-terminal region was fixed in space. The C_α atom of the C-terminal residue (henceforth referred to as PP for Pulled Point) was pulled at constant force towards a fixed end point (FEP) 1200Å away. The simulations were performed in a rectangular box of dimensions 200Åx200Åx3000Å with periodic boundary conditions. An Andersen thermostat (See Appendix A.3.2) was used to maintain room temperature (300K) for the simulation.(126) Constant-force pulling is achieved by applying a discretized step-function with a constant energy drop, dE , at the distance step of dR (0.1Å) between the PP and FEP. A step drop of -0.1kcal/0.1Å corresponds to a pulling force of 70pN. Simulations were run at constant forces of 5pN, 10pN, 15pN, 20pN, 25pN, 50pN, 70pN, 100pN, 150pN, and 300pN.

4.3.2 Patterns in Unfolding the α C domain

Calculating the critical force of unfolding for the α C region is slightly different than that of structured proteins such as discussed in subsequent sections. Unlike a folded protein where the force is distributed throughout the structure, when pulling on a natively unstructured protein, the natively unstructured part stretches first before unfolding will occur. Additionally, the direction of the applied force is important because the orientation of the structured region of the molecule with respect to the pulling direction affects how the molecule behaves. Direction, of course, is also important when pulling on a structured proteins, however in the case of structured proteins certain pulling directions are often more interesting due to their physiological relevance. In this study, the α C region was pulled from one direction, so we only probe one small region of the entire unfolding phase space. In future studies, it would be ideal to run a series of simulations pulling from different directions.

For an ideal (or freely jointed) polymer chain fixed at two ends, with end to end length \vec{R} the average restoring force $\langle \vec{F} \rangle$ to pull the chain into the state with maximum entropy (S) is given by(168):

$$\langle \vec{F} \rangle = T \frac{dS}{d\vec{R}} = \frac{k_B T}{P(\vec{R})} \frac{dP(\vec{R})}{d\vec{R}} \quad (4.1)$$

where $P(\vec{R})$ is the end to end probability distribution which follows a gaussian distribution. For a chain of N links of length l , at extensions not approaching the contour length of the fiber, this reduces to:

$$\langle \vec{F} \rangle = -k_B T \frac{3\vec{R}}{Nl^2} \quad (4.2)$$

Thus, throughout the pull, the unfolding force will constantly be working against a type of entropic force supplied by the velocity exchange algorithm in the Andersen thermostat.

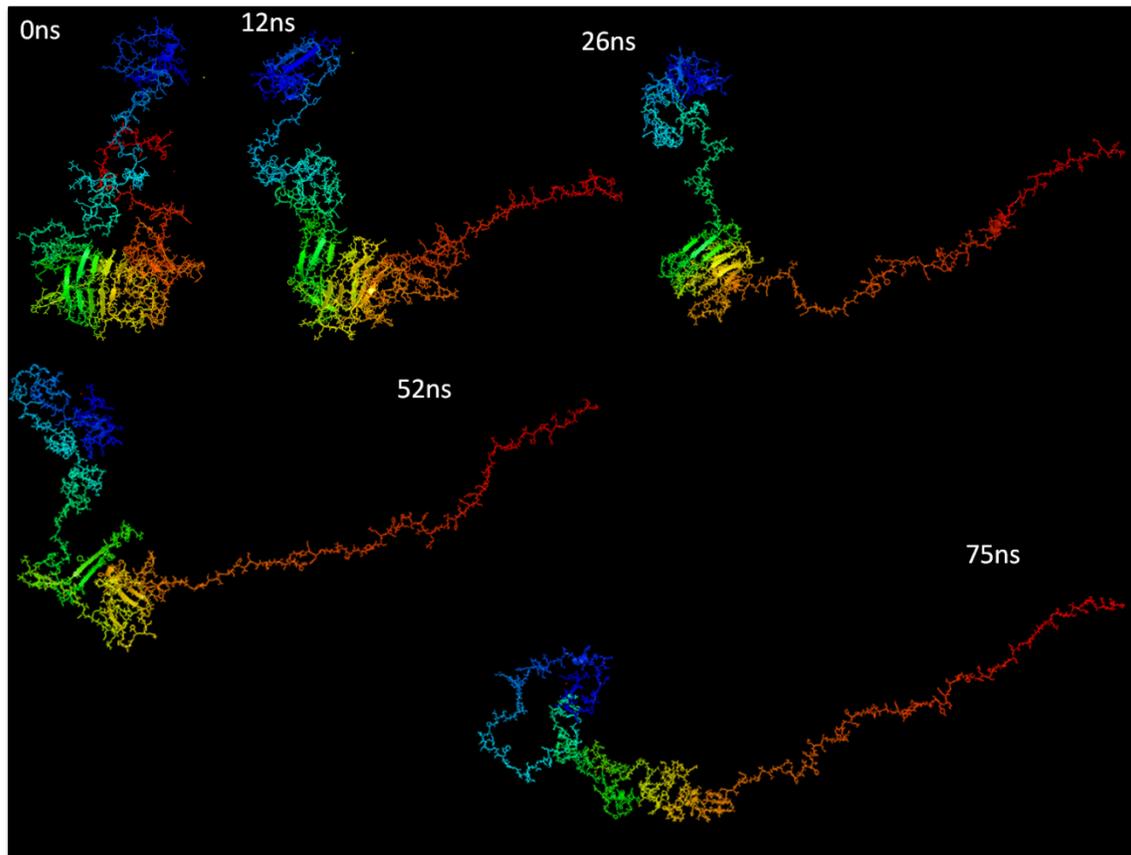


Figure 4.4 Unfolding the α C region at 25pN: Shown are five snapshots from a simulation of unfolding the α C region using 25pN of constant force. The C-terminus, where the force is applied is colored in red, while the N-terminus is colored in blue. At 12ns, the loose bundle on the C-terminus of the β -helix comes loose. At 26ns, the β -helix is reduced to β -sheet coils, and the repeat region is collapsing into a loose globule state. At 52ns, the N-terminal portion of the β -helix unfolds, and by 75ns, nearly all tertiary structure is gone.

4.3.3 Calculating the Unfolding Energy Barrier of the α C Region

As described above, unlike unfolding a mostly structured protein domain, unfolding the α C region likely has a significant entropic contribution to the energy barrier in addition to the enthalpic barrier due to two natively unstructured regions. The energy required to unfold the α C region can be calculated by integrating the stretching force over the stretching distance required to unfold:

$$\Delta G = \int F dx \quad (4.3)$$

For the C-terminal portion of the molecule, and a stretching length before unfolding of 17nm (observed in simulation), the integral becomes:

$$\Delta G = \int_{5nm}^{17nm} \left(\frac{3k_B T}{Nl^2} x \right) dx \quad (4.4)$$

Substituting $N=410$ and $l=0.3nm$ yields $U=43pN \cdot nm$.

In stretching simulations the the β -helix unfolds one coil at a time, with unfolding force of 10-20pN. Assuming an unfolding distance of 2nm, and a conservative critical force estimate of 25pN, unfolding the β -helix requires an additional 50pN·nm of energy. Stretching beyond this point will eventually lead to non-linear behavior force behavior of an ideal chain, typically modeled by the worm-like chain (see equation 2.3). Assuming a persistence length of 0.8nm (reasonable for an unstructured polymer), the αC can be stretched safely to 70nm before strong non-linear behavior occurs; again using eqn. 4.4 and integrating from 17nm to 70nm gives an additional energy barrier of 253pN·nm.(100) Thus, one can stretch αC region to 70nm of extension for a total energy output of 346pN·nm (~50kcal/mol), but only 50pN·nm of that entire energy barrier actually arises from unfolding the β -helix.

4.4 The Unfolding Energy Barrier of the Coiled Coil Region Probed by DMD

Previously the coiled-coil region of the fibrin molecule was probed using constant velocity steered molecular dynamics.(68) While the force-extension curves of these simulations provided an interesting comparison with AFM measurements, the simulations were run at velocities of 2.5-5m/s, much higher than experimental parameters. Thus, the forces were greatly inflated (1-3nN). To probe the stability of the coiled coil region of the structure at realistic forces and to compare its stability to that of the αC region, we

used steered discrete molecular dynamics (SDMD) simulations to unfold the isolated coiled-coil region of the molecule at constant force.

4.4.1 Methods

The coiled-coil structure was extracted from PDB structure 3GHG.⁽¹²⁾ Specifically amino acids α 27-200, β 58-198, and γ 14-130 from 3GHG were used to generate the initial structure for the simulation. Each chain of the coiled-coil structure consists of 111 amino acids surrounded on each end by one (or more) Cystein residues. At the C-terminal end of the coiled-coil, three di-sulfide bonds, between α Cys161- γ Cys135, α Cys165- β Cys193, and γ Cys139- β Cys197 link the three chains together. Correspondingly, at the N-termini of the molecule, di-sulfide bonds between α Cys45- γ Cys23, α Cys49- β Cys76, and β Cys80- γ Cys19 connect all three chains. Because of these di-sulfide bonds, all three chains of the coiled coil must unfold together. To pull on the coiled coil in the longitudinal direction, while generating minimal lateral forces, additional atoms (AP's for attachment points) were placed approximately 5nm beyond each of the Cystein clusters at the ends of the region.(See Figure 4.5) α Cys45, β Cys76, γ Cys19, were fixed at constant initial distance to one AP; α Cys165, β Cys197, γ Cys139 were fixed to the opposite AP and also allowed to oscillate. Di-sulfide bonds in the structure were parameterized with square well potentials in the same manner as Ding, Dokholyan 2008.⁽¹⁵⁸⁾ Similar to the α C region simulations, one final atom (the Fixed End Point FEP) was placed an additional 400Å beyond the C-terminal or N-terminal AP. One of the AP's was then held stationary while the other was pulled at constant force towards the FEP. The simulations were performed in a rectangular box of dimensions 100Åx100Åx2000Å with periodic boundary conditions. An Andersen thermostat (See Appendix A.3.2) was used to

maintain room temperature (300K).(126) Constant-force pulling is achieved by applying a discretized step-function with a constant energy drop, dE , at the distance step of dR (0.1\AA) between the PP and FEP. A step drop of $-0.1\text{kcal}/0.1\text{\AA}$ corresponds to a pulling force of 70pN. Simulations were run at constant forces of 25pN, 70pN, 100pN, 150pN, 225pN, 300pN, 400pN, 500pN, 700pN, and 1000pN.

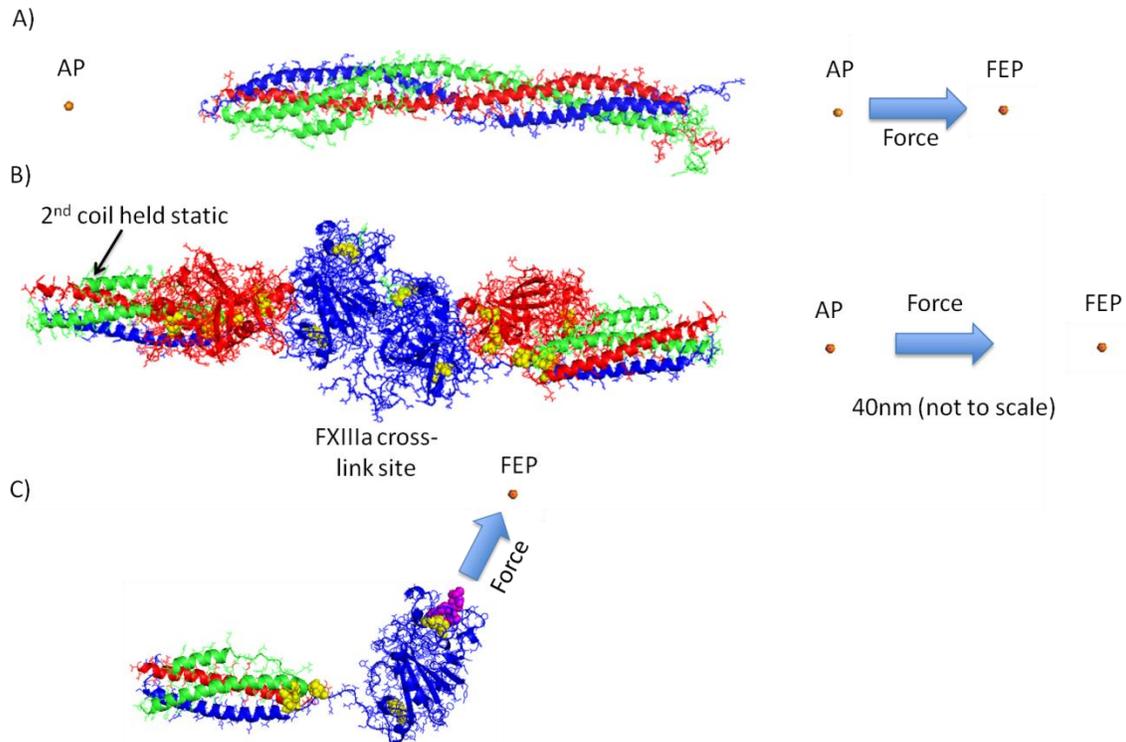


Figure 4.5 The Structure and Force Attachment Points for the DMD Simulations: In all images, di-sulfide bonds are colored in yellow, the α -chain is green, the β -chain in red, and the γ -chain in blue. A) The coiled-coil structure was generated from pdb 3GHG.(12) Attachment points (AP's) were located approximately 5nm from each end of the coiled coil, and were pulled towards the fixed end point. B) The γ - γ interface was generated from PDB structure 1FZC.(34) The triple- α -helix on the far left was held constant during the simulations. C) A:a unbinding interface was generated from PDB structure 2FFD.(169) The GPRV knob-A mimic is colored in purple. The β -C terminus was not included in the simulation.

4.4.2 Results

The Bell model can be applied to protein unfolding lifetimes under constant force.

In this model, the average protein lifetime τ_B is given by(170):

$$\tau_B = \tau_o \exp\left[\frac{[\Delta G_B - Fx_u]}{k_B T}\right] \quad (4.5)$$

where τ_B is the unfolding time for a protein under a given force (F), τ_o is the inverse of the energy well escape attempt frequency of the protein ($\sim 10^{-13}$ s), ΔG_B is the unfolding energy barrier in the absence of force, and x_u is distance between the unfolded energy state and the top of the unfolding energy barrier. (See Figure 4.6) Applying a constant force to a protein can be seen as lowering the energy barrier for unfolding. At the point where $F \cdot x_u \approx \Delta G_B$, the probability for the protein to be found in the folded and unfolded state are equal. The force at which $F \cdot x_u \approx \Delta G_B$ has been termed the critical force F_C .

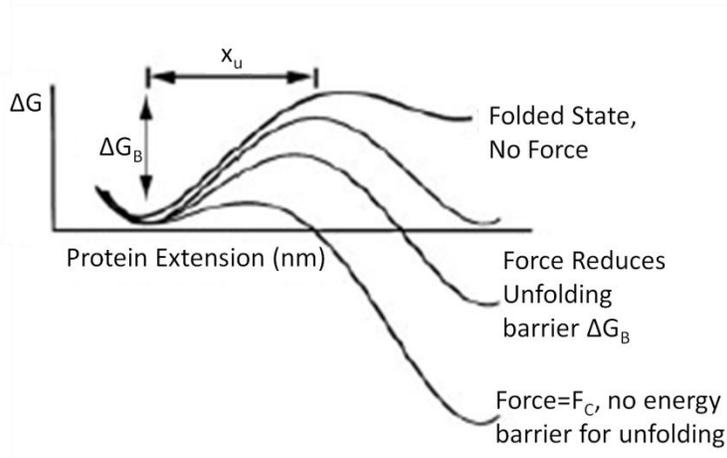


Figure 4.6 Unfolding Energy Barrier: Protein folding and unfolding can be thought of as an energy landscape. In its folded state, the protein resides at the lowest free energy. The energy required to kick it out of its folded state is the unfolding energy barrier (ΔG_B) and the extension in length to reach to height of the energy barrier is (x_u). By applying a constant force to a protein, one is essentially reducing the energy barrier by an amount $F \cdot x_u$. At the critical force (F_C), the protein is equally probable to be found in a folded or unfolded state. (Reprinted with modifications from Kesner, B.A., Ding, F., Temple, B.R., Dokholyan, N.V. “N-terminal Strands of Filamin IG domains act as a Conformational Switch Under Biological Forces” *Proteins* 2010, 78:12-24. Used with Permission from John Wiley and Sons. © 2009)

To estimate the critical force for coiled-coil unfolding, the simulations were allowed to run for 1,000,000 DMD time units (corresponding to 50ns), and then the protein length vs. force was plotted (See Figure 4.7). The plot indicates that the protein

remains folded at 100pN of force, but at 150pN the coiled coil region has unfolded a length of 12nm. This distance corresponds to the distance needed to unfold the triple helix portion of the coiled coil (Figure 4.8). This provides an estimate for the critical unfolding force (F_C) of the coiled coil of 125 ± 25 pN which is fairly consistent with the 94pN reported by Brown et. al in AFM experiments.(75) An estimate of x_u can be made by making a histogram of the time series of end-to-end distances of the coiled coil during extension at 100pN and 150pN of force (on either side of the critical force). Because the protein is unlikely to remain at the top of an unstable maxima in the energy landscape, x_u can be approximated as the extension at which the coiled-coil is least likely to reside (Figure 4.8). Using an x_u of 1.5nm gives a ΔG_B of $187.5 \text{ pN} \cdot \text{nm}$ (27 kcal/mol or $45 k_B T$), indicating a fairly high stability. To further refine this critical force value, a series of simulations should be run at a constant force of 125pN with different initial velocities. A series of unfolding trajectories will be generated and the critical residues in the unfolding process can be determined, allowing for further refinement of both F_C and x_u . The part of the coiled coil with a 4th coil was only observed unfolding in simulations at forces above 225pN, indicating a critical force between 150pN and 225pN. Extending the fourth coil can result in an additional 9nm of extension, giving a total extension for the coiled-coil region of 21nm.

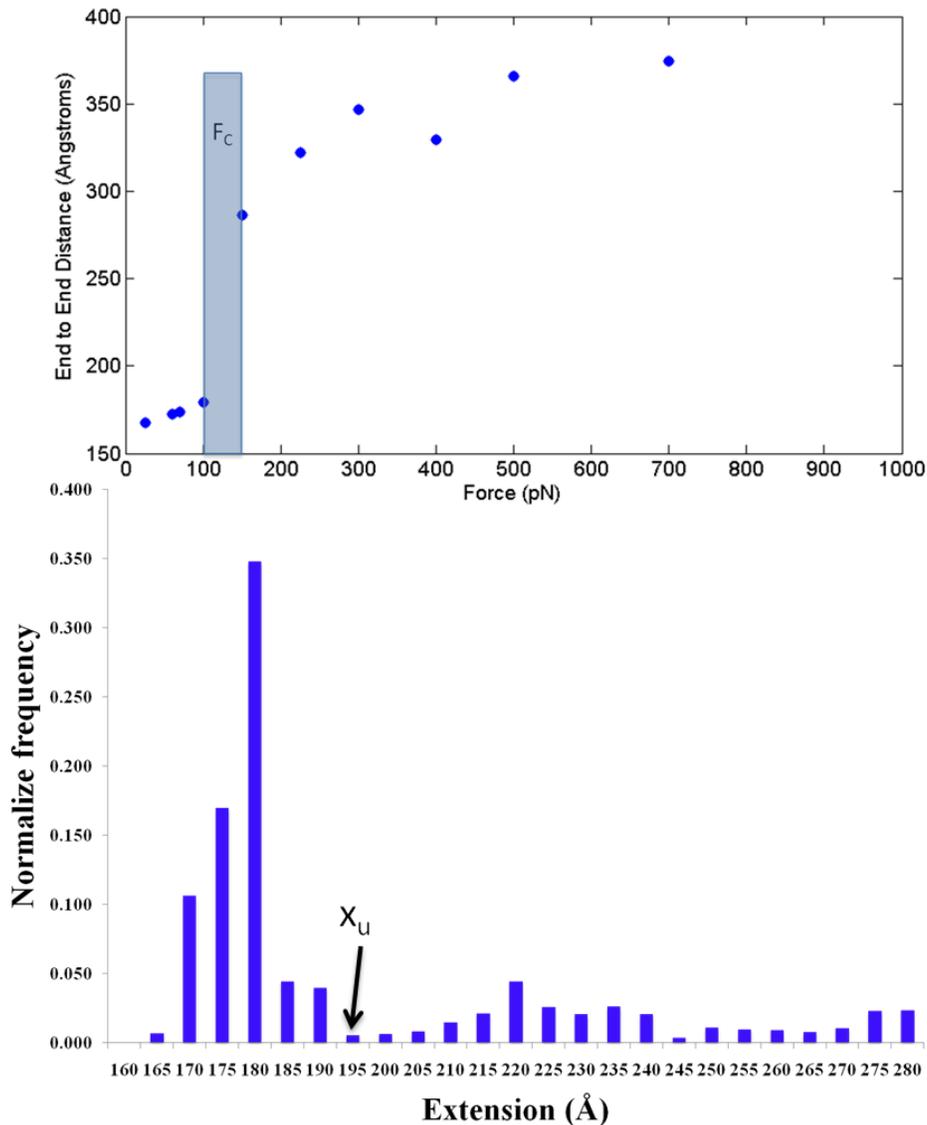


Figure 4.7: Coiled coil unfolded distance vs. Force: (Top) End to end distance (calculated as the distance between γ Cys19 and α Cys161) after ~ 50 ns was calculated for runs at each force. The critical force estimated to lie between 100pN and 150pN of force given the sudden jump in protein length at this juncture. At 150pN the triple helix of the coiled coil was unfolded, but the 4th coil remained folded. At higher forces, the 4th coil unraveled as well. (See Figure 4.8) (Bottom) Histogram of the end-to-end distances of the coiled-coil during the 100pN run (remained folded) and the 150pN run (unfolded). The critical unfolding length (x_u) can be estimated as the place of lowest frequency occurrence of end to end lengths because the protein resides at the top of an unstable energy landscape (Figure 4.6). Given the distributions, it appears that the critical distance occurs at 195Å extension. Assuming an unstretched coiled-coil length of 180Å gives $x_u \sim 15$ Å of extension.

These simulations also predict an intra-chain α -helix to inter-chain β -sheet hydrogen bond re-orientation upon unfolding of the coiled-coil region (See Figure 4.8). While this transformation has been speculated due to the similarity of the coiled coil structure to that of keratin, these simulations provide the first computationally based prediction of the effect.(65, 67) As discussed in Ch. 3, these results constrain the coiled coil region in its functionality as a reversible spring. Simulations are currently ongoing, but after 100ns of simulation time in the absence of force, the energy of the β -sheet conformation appears stable indicating that the region may not re-fold into the coiled coil structure after the removal of force.

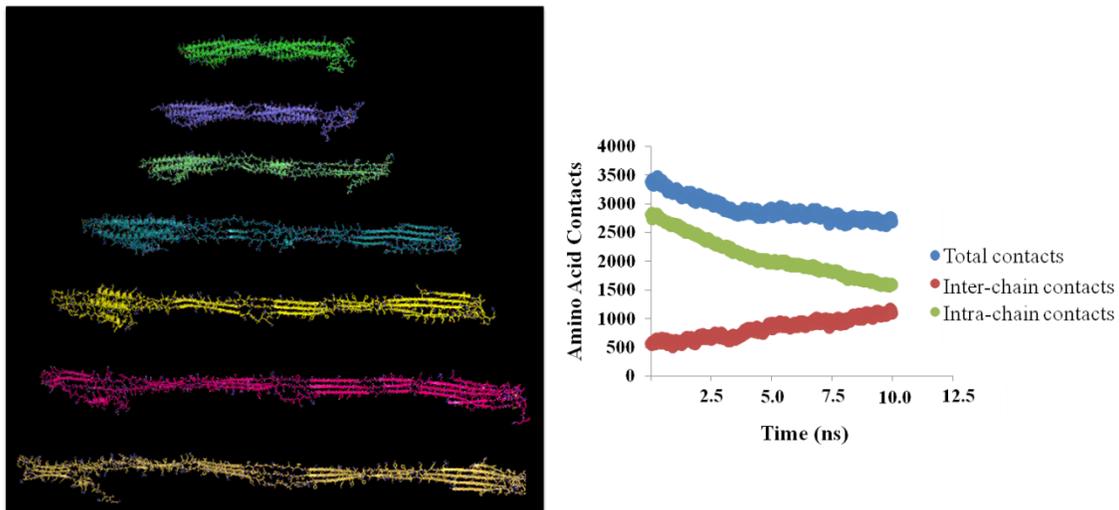


Figure 4.8 α -helix to β -sheet transition: As the α -helical coiled-coil region is stretched, the first part to open is the “kink” in the middle. Next, the N-terminal portion of the coiled-coil opens, and finally, the C-terminal portion with a 4th coil unravels. Interestingly, as the triple-helix unravels, inter chain β sheets begin to form. As can be seen in the plot on the right, The coiled-coil structure transitions from intra-chain contacts to inter-chain contacts indicating the α -helix to β -sheet transition. We define a contact when the distance between two C_{β} carbons is less than 7.5\AA .

4.5 Unfolding of the γ Regions of the Fibrin Molecule

Unfolding of the γ -C region (γ -nodule in other language) was proposed by Averett et al. as a potential mechanism for fibrin extension.(71, 72) In an elegant series of AFM experiments, they demonstrated that breaking the A:a interaction involved four separate

force rupture events, each with a characteristic unfolding length, critical force, and energy barrier. They proposed the pattern could be explained by the following series of events: the first force event involved the γ -C region of the protein detaching from the surface, this was followed by two unfolding events in the γ -C region of the molecule with characteristic lengths of $6\pm 2\text{nm}$ and $17\pm 6\text{nm}$ respectively, finally, the 4th event represented the rupture of the A:a interaction which occurred at lower forces due to the previous unfolding events. After finding the same pattern in several fibrin fragments, the authors concluded that they saw no evidence of coiled-coil unfolding in the A:a rupture pattern.(71) To investigate the relative stability and unfolding patterns of the γ -C regions we performed two sets of DMD simulations. In the first set of simulations, we investigated the γ - γ interface between two fibrinogen molecules. In the second set of simulations we investigated the unfolding pattern associated with pulling on the A:a interaction.

4.5.1 Methods for SDMD simulations of the fibrin γ - γ interface

The original structure was taken from the γ_1 - γ_2 crystal structure published as 1FZC in the protein data base, where γ_1 and γ_2 are the γ -C termini of two adjacent molecules (referred to in this section as M1 and M2).(34) The FXIIIa ligation site located in amino acids γ 398-411 was not present in the crystal structure. The residues γ 398- γ 411 were added to both γ_1 and γ_2 by creating a peptide bond between γ Gly397 and γ Gln398. An additional covalent-type bond potential was then added between γ_1 Gly399 and γ_2 Lys406 and γ_2 Gly399 and γ_1 Lys406 to represent the iso-peptide bond formed from FXIIIa ligation. The parameterization for both the peptide bond and iso-peptide bond are similar to those used in Ding, Dokholyan 2008.(159) Di-sulfide bonds in the structure

where parameterized with square well potentials in the same manner as Ding, Dokholyan 2008.(158) Ca^{2+} binding pockets were simulated as Zn^{2+} binding pockets because the DMD all atom force field currently does not include Ca^{2+} parameterization; the assumption being that having an ion bound in the pocket is better than none at all.(158) Zn^{2+} constraints were modeled by assigning distance constraints between each metal atom and the corresponding metal-coordinating atoms in the fibrin Ca^{2+} binding pockets observed in the crystal structure. Simulations were run with and without additional salt bridge constraints in the A:a and B:b binding pockets. In simulations performed without the additional salt bridge potential, the A knob mimic (GPRP) and the B knob mimic (GHRP) were observed to immediately leave the binding pocket. This indicates that hydrogen bonds play a minor role in the A:a and B:b interactions. To mimic the salt bridge and ensure that the peptides stay bound, the salt bridge was modeled as an infinite square well potential between the NH_2^+ arginine in GPRP and the delta oxygen of the γAsp234 . Similar constraints were applied for each binding pocket. The 1FZC crystal contains a partial coiled coil crystal structure starting at $\alpha 119$, $\beta 151$, and $\gamma 97$. To speed up the simulation, the atoms in the coiled coil region of M1 were fixed in place and not allowed to fluctuate. Force was applied to the coiled coil region of M2 in a similar manner as described before. The $\text{C}\alpha$ carbons of $\alpha 119$, $\beta 151$, and $\gamma 97$ were fixed at constant distance to an attachment point (AP) approximately 5nm beyond each of the $\text{C}\alpha$'s. (See Figure 4.5) One final atom (the Fixed End Point FEP) was placed an additional 400\AA beyond the AP. The simulations were performed in a rectangular box of dimensions $120\text{\AA} \times 120\text{\AA} \times 1800\text{\AA}$ with periodic boundary conditions. An Andersen thermostat was used to maintain room temperature (300K).(126) Constant-force pulling

is achieved by applying a discretized step-function with a constant energy drop, dE , at the distance step of dR (0.1\AA) between the PP and FEP. Simulations were run at constant forces of 25pN, 70pN, 100pN, 150pN, and 300pN.

4.5.2 Methods for SDMD simulations of the fibrin γ -C terminus unfolding

The original structure was taken from the crystal structure published as 2FFD in the protein data base.⁽¹⁶⁹⁾ Di-sulfide bonds in the structure and Zn^{2+} (substituted for Ca^{2+}) binding pockets were parameterized with square well potentials as described before.⁽¹⁵⁸⁾ The A-knob mimic GPRV was bound in the hole-A binding pocket using an infinite square well potential in a similar manner to section 4.5.1. The reasoning for this was based on results from AFM measurements indicating a sequence of 4 unfolding events before the A:a rupture.⁽⁷¹⁾ It was therefore assumed, that forcing the peptide to stay bound in the pocket should allow visualization of all unfolding events. To speed up the simulation, residues β 199-459 were omitted, based on the assumption that interactions with the β -C domain did not influence the A:a binding pocket. The 2FFD crystal contains a partial coiled coil crystal structure starting at α 126, β 157, and γ 96. Force was applied to the $C\alpha$ carbon of Val4 in the GPRP peptide in a similar manner as described before. One final atom (the Fixed End Point FEP) was placed an additional 400\AA beyond the AP. The simulations were performed in a rectangular box of dimensions $300\text{\AA}\times 300\text{\AA}\times 1200\text{\AA}$ with periodic boundary conditions. An Anderson thermostat was used to maintain room temperature (300K).⁽¹²⁶⁾ Constant-force pulling is achieved by applying a discretized step-function with a constant energy drop, dE , at the distance step of dR (0.1\AA) between the PP and FEP. Simulations were run at constant forces of 25pN, 50pN, 70pN, 100pN, 150pN, and 200pN.

4.5.3 γ -C and γ - γ unfolding energy barriers

The results presented in this section are preliminary as simulations of both the γ - γ interface and the γ -C region are ongoing; however patterns of unfolding that have already emerged are presented here. Simulations on the γ -C region predict a specific sequence of unfolding based on the relative energy barriers associated with the unfolding. First, simulation results from forced stretching of the γ - γ interface will be presented.

As can be seen in Figure 4.9, the weakest link in the γ - γ structure is actually the interface between the two molecules. At all force levels simulated, the γ - γ interface comes apart. Due to the direction of applied force combined with the γ - γ cross-linking, a torque is generated on the region after the interface separation, causing the γ -C and β -C regions in both molecules to propeller up into a position perpendicular to the coiled-coil. When the original γ - γ dimer crystal structure was presented, the authors noted that a crevice existed between the two molecules and the buried surface area came to only 750\AA^2 indicating only a small contribution to the polymerization free energy.⁽¹⁹⁾ An important work on protein oligomerization energies by Janin et al. indicated an average of 1 hydrogen bond per 200\AA^2 of binding surface area; so in retrospect, it is not surprising that the γ - γ interface is weakly bound.⁽¹⁷¹⁾ One particular residue, γ Arg275 has been implicated as being important in end-to end oligomerization, however it does not appear to play a role mechanically stabilizing the interface.⁽¹⁹⁾ While these simulations do not allow an exact calculation of the critical force, given that even at 25pN, the γ - γ interface comes apart within 20,000 DMD time units ($\sim 1\text{ns}$), the critical force for stretching this region may be as low as 5-10pN. It should be noted that within a fibrin protofibril, additional constrains from the A:a and B:b interactions would only

allow 4nm of separation between the interfaces (due to the lengths of the knob-A polypeptide chains) before larger forces would be required to stretch to molecule further.

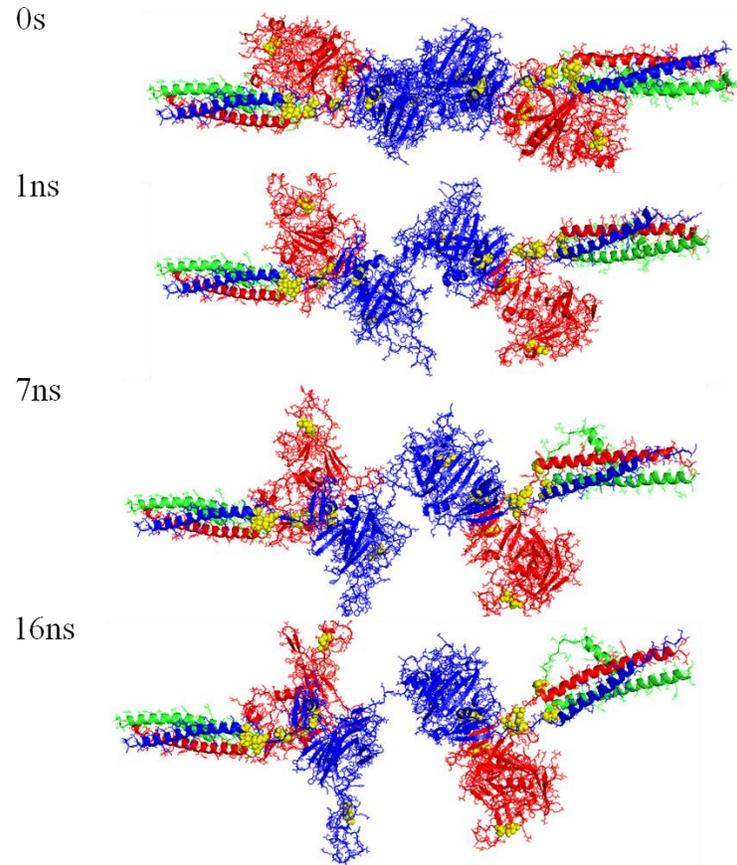


Figure 4.9 Unhinging of th γ - γ connection under force: While in the crystal structure, the γ - γ region appears stable, upon the application of force, even at 25pN as this series of simulations takes place, the first part of the structure to stretch is the γ - γ interface. After the initial separation, a torque generated by the FXIIIa ligation, rotates the β and γ -C domains perpendicular to the coiled-coil.

Additional simulations have been used to test the mechanical de-stabilization of the γ -C region resulting from forces applied to the A:a interaction. Because the DMD force field does not include long range electrostatic interactions, it did not properly account for the salt bridge binding that holds the A-knob peptide mimic in the binding pocket.(159) As described in the methods, this was overcome by using an infinite square well potential to keep the A-knob bound. The downside of this method is that knob unbinding will not be observed in the simulations; however, these simulations do provide

a vehicle to probe the sequential γ -C domain unfolding measured/predicted by recent AFM experiments. (71, 72)

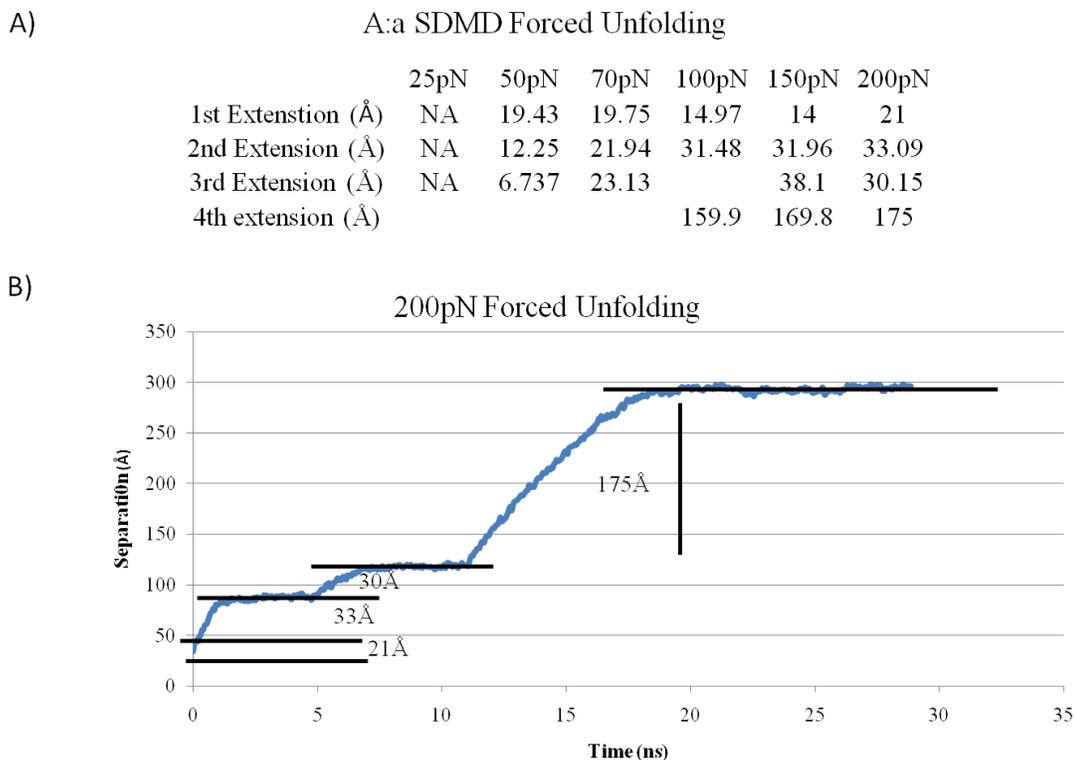


Figure 4.10: Step Heights between γ -C unfolding events: To quantify an unfolding event, we tracked the distance between γ Cys153 and the C α carbon of the Gly1 in the GPRV peptide. These residues were chosen based upon the observation that γ Cys153 forms a di-sulfide bond which stabilizes the “bottom” of the γ -C structure, and tracking the height of Gly1 would roughly correspond to the AFM measurements of Averett and coworkers.(72) (A) The top table shows the step heights between plateaus in the separation vs. time curve for each force. The γ -C domain did not unfold at 25pN of force and was not included. (B) Representative separation vs. time plot for the γ -C region under 200pN of applied force. Distances were measured from levels of steady plateaus in the curves as represented from the black lines. Separations were measured based on the distance between the average heights in the plateaus.

Intriguingly, simulation results after 25ns of DMD simulation indicate a sequential unfolding of the γ -C region under force similar to that measured by Averett et al.(71) A series of images depicting sequential unfolding are represented in Figure 4.11. Analysis of the step heights of the separation vs. time curves as seen in (Figure 4.10)

show a distribution of step heights at different forces. At each force, the first event of 15-20Å separation did not correspond to an unfolding event in the movies but rather an adjustment of the A-knob within the binding pocket. At each force above 50pN 2-3 distinct unfolding events were observed. The table has been adjusted in an attempt to line up unfolding events by separation distance. The second and third unfolding events ranged from 20-30Å of separation. If the two events were combined into one event (the reasoning will be motivated below), then the average separation for steps 2 and 3 would be one with an average separation of $59 \pm 13 \text{Å}$. Finally, at forces above 70pN, a 4th event was present with an average separation of $168 \pm 8 \text{Å}$. These results are remarkably similar to experimental results. In AFM experiments testing the stability of the A:a interaction, 4 distinct force peaks were observed. The first was attributed to the protein lifting off the surface, the second two were assigned as protein unfolding events prior to A:a rupture and weakening the A:a bond, and the 4th event was attributed to the unbinding of knob A from the binding pocket. The relative separations between these three events were $110 \pm 60 \text{Å}$, $60 \pm 20 \text{Å}$, and $170 \pm 60 \text{Å}$.⁽⁷²⁾ Comparing these results with our own indicate that our second and third separation events could be the third and fourth unfolding events captured in the AFM experiments. It's possible that the first event in the AFM experiments involved the β -nodule which was not included in our simulations for simplicity. It must be remembered that only one simulation was performed at each force, but the similarities are intriguing.

The critical force for each event can be determined by plotting the time needed to reach the unfolding event vs. force applied as seen in Figure 4.11. Linear Fits can be applied to determine the intersection between the barrierless region where the protein

unfolds without hindrance and the barrier region. This intersection provides the critical unfolding force for the event. The critical force for events #2-3 was 74 ± 7 pN, and for event #4 was 89 ± 21 pN. Uncertainties are reported as standard errors and were calculated by a simple monte carlo algorithm as described in Kesner et al, and should be reduced upon additional simulations at each force.(172) In terms of relative stability, the coiled-coil region was not observed unfolding at forces below 150pN, consistent with the coiled coil unfolding measurements and indicating a higher stability for the coiled coil region than the γ -C region.

While these simulations results are preliminary, they do indicate similarities with the AFM data. A second series of simulations performed near the critical force value for each event with a variety of randomized starting velocities will allow a characterization of the important amino acids in the unfolding events, and will give a better characterization of the critical force values. In both the experiments done by Averett et al. and these simulations, γ -C unfolding is observed before coiled-coil unfolding and the critical force values for unfolding are lower for the γ -C region, indicating that the γ -C region would unfold prior to coiled-coil stretching.(71, 72) It is not clear how to reconcile this with the AFM measurements by Lim and Brown indicating coiled-coil unfolding; neither paper described the AFM tip velocities, so it is not possible to do a direct comparison with the Averett force values.(68, 75) We are currently collaborating with Kellie Beicker to pursue follow up measurements of coiled-coil unfolding at constant force.

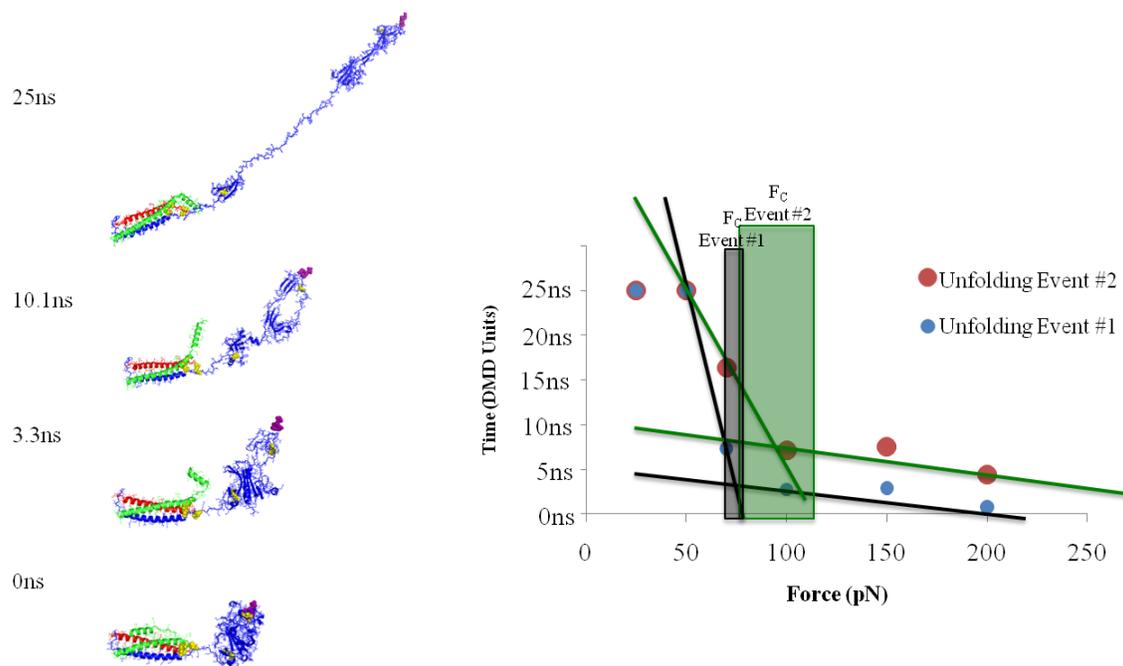


Figure 4.11 Forced Unfolding of the γ -C domain: Images on the right depict sequential unfolding of the γ -C domain through application of 100pN of constant force at the A:a interface. After 3.3ns, the first unfolding event is observed where one part of the region separates from the rest. After 10.1ns, a second unfolding event occurs. After 25ns the γ -C region contains three distinct domains with β -sheet structures separated by unstructured polypeptides. The distance between the first (containing the knob binding pocket) and second structures is ~ 6 nm. The distance between the second and third is ~ 16 nm. The plot on the right can be used to estimate the F_c for each event. Time before unfolding events #1 and #2 are plotted as a function of applied force. For each event, there is an approximately barrierless region where the unfolding occurs rapidly, and a region with a barrier, where the unfolding time for the event is observed to rise. For forces in which no unfolding events occurred, the unfolding time is registered as 25ns. The intersection of fits in the barrier and “barrierless” unfolding regions represents the critical force of unfolding. Event #1, represented in black has $F_c=74\pm 7$ pN, and for event #4 $F_c=89\pm 21$ pN.

4.6 The SL α CK Molecular Model for Fibrin Extension

The results of these simulations, combined with the known experimental mechanical data provide a more complete view of the molecular underpinnings of fibrin's mechanical properties. We now propose a new model for fibrin extension, which we refer to as the SL α CK model (**S**traightening of molecules, **γ -C L**igation site separation, **A**lpha (**α**)-C extension, **γ -C** unfolding / **C**oiled-coil extension, **K**nob unbinding); the main mechanism for extension in the SL α CK model comes from extending the loose α C domains (See Figure 4.12).

In order to develop a model for the mechanical properties of the fibrin fiber, one must first start with a model of the molecular connections within the fiber. Here I will briefly lay out our understanding of the molecular packing inside the fiber. The fiber consists of protofibrils in lateral registry in parallel across the diameter of the fiber.(44) The protofibrils vary in length from 1 to approximately 20 monomers within the fiber, and the distribution of lengths is bimodal (with peaks at 1 molecule and 20 molecules towards the end of polymerization) as described by Weisel et al.(99) Lateral contacts between protofibrils are strong enough to keep the fiber in registry without applied force, but weak enough that the fiber behaves as a loose bundle of protofibrils, rather than a close-packed bundle of protofibrils.(82) Protofibrils have an inherent twist to them and may wrap around the fibrin fiber.(45, 46) The α C domains connect the protofibrils both in series (we assume due to their long lengths) and in parallel down the length and across the diameter of the fiber.(144) The α C domains can be cross-linked with FXIIIa, or can associate with unbinding forces of up to 50pN in the absence of FXIIIa.(73, 154)

Protofibrils are assembled in a double-stranded, half-staggered arrangement due to A:a and B:b knob:hole interactions.⁽⁹⁹⁾ The γ -C termini within the protofibril are cross-linked by FXIIIa, but likely associate in the absence of cross-linking.^(34, 42) With the discovery of elastomeric fibrin sheets, it's possible that this model needs adjustment, but for the present report, we will assume the above polymerization model for fibrin fibers.⁽⁴⁹⁾

Given this structural model for the fibrin fiber, the SLACK extension model proceeds naturally from the results of the simulations combined with other experimental data. The first phase of SLACK is the straightening out of the fibrin molecules and protofibrils within the fiber. Others have estimated that this could generate up to 20% fiber strain or 9nm of extension.⁽⁸³⁾ Next, in the "L" phase the γ -C: γ -C interface between monomers in the protofibril stretches out. This was observed at all forces in the DMD simulations of the γ - γ interface discussed in section 4.5.3. Because the A:a and B:b associations constrain the γ - γ interface, this can only generate 4nm of extension (~8% strain) without unfolding the γ -C domain. However, this separation may be enough to disrupt lateral contacts between protofibrils, as they are believed to be weak anyways.⁽⁸²⁾ Both the "S" and "L" phases of SLACK should be reversible in the absence of force, as no protein unfolding occurs.

The third step is extension of the α C domains between protofibrils. Both our model and NMR fragment derived models for the α C domain indicate loose β -sheet structures.⁽¹⁴⁾ The results from our DMD stretching experiments indicate that the folded α C domain can unfold at forces as low as 10pN, while the natively unstructured portions of the α C region stretch as an entropic spring (spring constant of 0.3pN/nm). A few

experimental based assumptions can allow a rough estimate of the extension of the fibrin fiber allowed by α C mediated protofibril sliding.

The main α C ligation sites are Lys556 and Lys580 along with Gln328 and Gln366 (Gln221 and Gln237 have also been implicated), indicating that α -polymers are primarily ligated by interactions with Lysines on the C-terminal side of one α C domain and Gln on the N-terminal side of another α C domain.(29, 30, 155) Take for example a chain cross-linked between α_1 Lys580 and α_2 Gln326 where α_1 and α_2 represent two separate α C regions. The total extension length for α_2 would be 326-200 (the α -chain coiled coil ends at amino acid 200) =126 amino acids; the total extension for α_1 would be 580-200-30=350 amino acids (the extra 30 comes from amino acids Cys442-Cys472 which are linked by a di-sulfide bond and do not extend). Using an acid length of 0.3nm indicates that this cross-linking sequence could account for 140nm of extension. Using α_1 Lys556 and α_2 Gln221 (the shortest of the combinations) allows for 105nm extension, while the longest combination (α_1 Lys580 and α_2 Gln366) allows for 155nm of extension.

Assuming a simple model of protofibrils in series within the fiber, a distribution of protofibril lengths as measured at the 0.75 lag phase by Weisel et al.(99) and an α C length of 140nm, allows for fibrin strains of 75%. Assuming different distributions of protofibril lengths (such as that measured by Weisel et al. at the 0.5 lag phase), combined with different combinations of cross-linking can give fiber strains ranging from 50% up to 150%. Similar reasoning can be applied to un-crosslinked α C domains interacting non-covalently. As noted earlier, non-specific α C domains can interact with binding forces of up to 50pN, allowing moderate extensions (~120nm) of the α C region before irreversible α C domain unbinding would occur. It should be noted that within the fiber,

the α C domains are often found in large chains called α -polymers which have a more complex cross-linking pattern than the simple two chain model described above.(160) In spite of this, most of the potential Lys donors are located in α 411 and beyond indicating an extension of at least 100nm per α -polymer chain

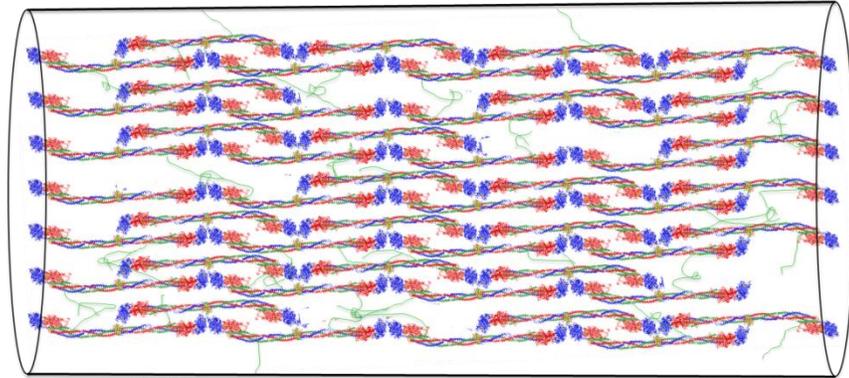
It seems then that the “ α ” phase of the SL α CK model can generate somewhere around 100% fiber strain in both ligated and unligated fibers while only requiring a force/protofibril of \sim 10pN at low extensions. This is in good agreement with the force/monomer calculations in Chapter2. In total, the S, L, and A regimes of the SL α CK model can account for \sim 130% reversible extension of the fibrin fiber. Fully cross-linking all α C regions at multiple cross-linking points would serve to reduce the available extension of the α C region, and generate a stiffer fiber as observed in equation 4.2, which may explain the results of Liu et al.(67) An intermediate level of cross-linking which provides stabilization between α C domain interactions without greatly decreasing the available α C contour length may provide an intermediate range where the fibers are more extensible than uncross-linked fibers while retaining the same stiffness as uncross-linked fibers, potentially explaining the apparent contradiction between (66) and (67).

As the α C regions stretch to higher extensions, WLC-like strain stiffening will increase the required pulling force to greater than 70pN to stretch the α C regions further. At this point the force required to stretch the α C region is higher than that needed to unfold the γ -C or coiled-coil regions leading to the “C” phase of the SLACK model. Results from the simulations in chapter 4, indicate that the coiled-coil region has a higher critical force than the γ -C region of the molecule (125pN vs. 74pN and 89pN). This is in agreement with the Averett et al. report finding that γ -C unfolding likely precedes coiled-

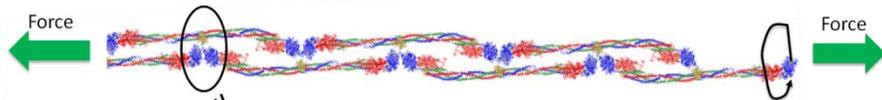
coil unfolding,(71, 72) but is contradicted by AFM measurements on the fibrin molecule and fibrinogen oligomers indicating that the coiled-coil region unfolds at 60pN and 95pN of force respectively, which were lower than the Averett values for the γ -C region (~100pN-200pN).(68, 75) Given our simulation results that the unfolding energy barriers are within 20-50pN, it seems possible that a portion of the molecules unfold via the γ -C region and a portion unfold the coiled-coil region. While it is possible that the “C” phase could allow for an additional 200nm of extension, it is likely that fiber failure as explained in the “K” phase will occur prior to this.

The “K” phase of the SLACK model refers to knob-hole unbinding which will occur after the sequential unfolding of the γ -C domain and lead to fiber failure.(71) Estimates of the force per fibrin molecule inside a fibrin fiber at the point of breaking (as explored in section 2.3.7 Force per Monomer) are consistent with A:a unbinding.(71) A:a unbinding and potentially coiled-coil and γ -C unfolding (the timescales of a few seconds may be too fast for a 400 amino acid protein segment to re-fold, see (173) and chapter 3.4 An α -helix to β -sheet transition for a discussion on re-folding timescales) would result in plastic deformation and irreversibility in fiber extension.

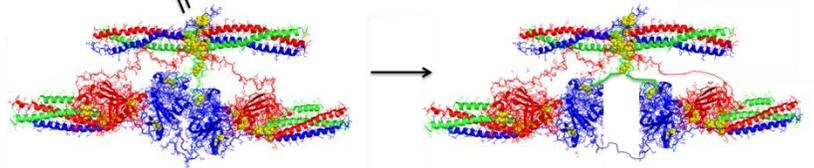
Initial Structure of a fibrin fiber: Protofibrils of varying lengths, laterally aggregated in registry, intertwined by interacting α C regions (green lines)



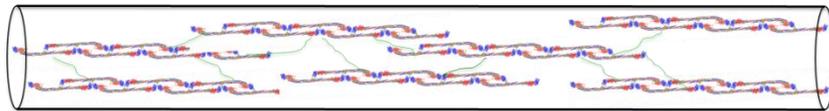
Step 1 (**S**): Monomer straightening, protofibril untwisting ~ 20% fiber strain



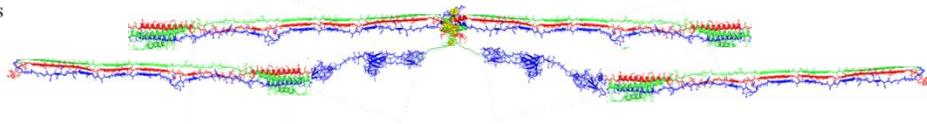
Step 2 (**L**): γ - γ interface separates =4nm ~additional 10% fiber strain; may also disrupt lateral contacts promoting protofibril sliding



Step 3 (**A**): α C regions between protofibrils stretch; additional ~50-150% strain; perhaps B:b interactions fail



Step 4 (**C**): γ C domains unravel and/or coiled coil region stretches ~could account for an additional 200% fiber strain



Step 5 (**K**): A:a knob: hole interactions fail, fiber breaks

Figure 4.12: The SL α CK model for Fibrin Fiber Extension: Our view of the packing in the fibrin molecule is described in the text. In the “S” phase the molecules straighten within the protofibrils. In the “L” phase, the γ - γ interface separates, while still being constrained by the A:a interaction. In the “A” phase the α C regions between the protofibrils stretch. The “C” phase comprises of some γ -C regions and some coiled-coils unfolding. And the “K” phase refers to knob unbinding resulting in fiber failure.

Chapter 5. Stiffening of Individual Fibrin Fibers Equitably Distributes Strain and Strengthens Networks

5.1 The Mechanical Properties of Fibrin in the Context of Biopolymer Networks

We now turn our attention to the highest rung of the mechanical hierarchy, the fibrin network. Fibrin networks belong to a class of biological materials that display a remarkable and diverse set of mechanical properties including high extensibility, non-linear elasticity (strain stiffening) and negative normal stress (59, 61, 63, 74, 79, 174-179). Though these behaviors have been well characterized in macroscopic studies, an understanding of their molecular and fiber-level origins has only recently begun to develop (74, 80, 178, 180). To construct a complete multi-scale picture of network behavior, a thorough characterization of microscopic properties must be accompanied by an understanding of how they conspire to produce the emergent bulk response under stress: How do the single fiber properties translate into the bulk properties of the network?

Materials strength - the maximum stress a material can withstand before failure - is particularly dependent on microscopic mechanical and geometrical details. Failure occurs at points of concentrated stress or points of mechanical weakness; it is determined by extreme values in the range of microscopic properties rather than in their average values. This is in contrast to properties, such as elastic moduli, which are amenable to a mean field analysis. A thorough understanding of how a fibrin network fails – a question of profound biomedical as well as materials interest – requires a detailed look at how individual fibrin fibers stretch, how they distribute strain, and ultimately how they fail.

Modeling studies focusing on the mechanics of biopolymer networks identify two phenomena responsible for a network's response to stress: stiffening behavior of the constituent fibers and geometrical rearrangement of the fibers (74, 80, 81, 181). The balance struck between these two mechanisms depends in large part on the flexibility of the fiber segments comprising the network. Models of semi-flexible fibers – fiber elements with persistence lengths comparable to their contour lengths - show that stiffening of a network can arise from the entropic stiffening of the constituent fibers comprising the network (80). On the other hand, models of stiff fibers show that stiffening develops from a geometrical reorientation and transition from bending at small strains to enthalpic stretching at larger strains (81). While these models have been successful at describing network behavior at low strains, (80, 180) the question of how the stiffening and re-arrangement of individual fibers affects stress and strain distribution and ultimately the overall network strength at high fiber strains, has received less attention.

Recent experimental work has begun to bridge the gap between bulk macroscopic properties and their origins at the individual fiber level. Advances in micro and nanoscale interrogative techniques have enabled mechanical studies of fibrin at the single fiber (65, 66, 77, 182) and molecular scale (68, 71, 75). These studies revealed that individual fibrin fibers possess remarkable elastic properties; they are capable of reversibly stretching as much as three times their original length, and can more than quadruple their length before breaking (65, 66, 77). We also report here direct measurements that show fibrin fibers themselves exhibit non-linear elasticity. The molecular origins of these properties are still an open question, but recent studies have

implicated several different regions of the fibrin molecule as possible sources for fibrin's extensibility including the coiled-coil region, (68, 74, 75) the globular γ domain (71), and the α C connector region (77). Further investigation is needed, but it is clear that the fibrin fibers themselves exhibit elastomeric behavior: low elastic modulus in stretching (MPa) and very high extensibility (>300%).

Within fibrin networks, microscopic measurements of filament re-orientation and alignment under stress have been made both at low strains in a shear cell (178, 180) and at high tensile strains (74). At high network strain, the fibers have aligned in the direction of the stress, and fiber stretching is the dominant deformation mechanism. Consequently, the elastomeric properties of the individual fiber dramatically affect network deformation in this regime, and play a direct role in determining network strength: failure originates ultimately at a single fiber (or branch point). Thus, network strength is determined by the maximum strain individual fibers can withstand. Determining how strain is shared among the constituent fiber segments in a network under imposed stress is therefore crucial to understanding failure modes of networks and their strength. We describe here that strain stiffening of the individual fibrin fibers produces a dramatic shift in strain distribution and effectively strengthens fibrin networks.

5.2 Materials and Methods

5.2.1 Fluorescent Microscopy Applied to Fibrin Fibers

For observation and manipulation of fibers, the coverslips were placed face up on an inverted Nikon Diaphot microscope with epifluorescence illumination (Nikon Diaphot 200, Southern Micro Instruments, Atlanta, GA), and imaged using a rhodamine HQ filter

set (Chroma Technology, Rockingham, VT) and a 1.3 or 1.4 NA 100X oil objective. Images were recorded using a high speed Cooke PCO 1600 camera with C-link and recorded with CamWare (Cooke Corp., Romulus, MI).

5.2.2 Fibrinogen Preparation

All experiments used recombinant human fibrinogen produced in Chinese Hamster Ovary (CHO) cells (97, 183). Fibrin clots formed with recombinant human fibrinogen are indistinguishable from clots formed with plasma fibrinogen. Recombinant fibrinogen was used in this study because it is free from other blood coagulation factors, in particular, FXIII (97). Fibrinogen purity and homogeneity was assessed by SDS polyacrylamide gel electrophoresis (SDS-PAGE) and immunoblot analyses. Fibrinogen function was assessed through thrombin-catalyzed polymerization monitored by turbidity, and FXIIIa-catalyzed crosslinking monitored by SDS-PAGE (184).

5.2.3 Growing Fibrin Networks on Structured Surfaces

Fibrin fibers and networks were assembled in situ onto Norland Optical cured structured surfaces as described in chapter 2. Briefly, a polydimethylsiloxane (PDMS) stamp with 20 μm wide and 10 μm deep channels was placed on a small drop of Norland Optical #81 (an ultraviolet-light-curable optical adhesive) in the middle of a 24 x 50 mm 1.5 coverglass (Corning, Lowell, MA). The adhesive was polymerized with long wavelength ultraviolet light for 2 minutes, and the stamp was peeled off, leaving a cured structured surface (SS).

Before use, the SS was cleaned for 2 additional minutes in a UV Cleaner, and a ring of silicon grease was applied with a cotton swab to confine fibrin assembly to the SS. Fibrinogen, stored at 0.6 mg/mL, thrombin (human α -thrombin, Enzyme Research Labs,

Indianapolis, IN), stored at 222 U/mL, and Factor XIII (human plasma FXIII, Enzyme Research Labs), stored at 68 $\mu\text{g/mL}$ at -80°C were thawed rapidly and placed on ice. Fibrinogen and FXIII were diluted to 0.04 mg/mL and 0.1 $\mu\text{g/mL}$ respectively in 20 mM HEPES, 150 mM NaCl, pH 7.4 (HBS), and one half the final volume needed to cover the SS (usually 8-10 μL) was pipetted onto the surface. Thrombin was diluted just before use to 2 U/mL in HBS with 10 mM CaCl_2 , and an equal volume was added to each SS and mixed by gently pipetting up and down several times. Final concentrations of reagents were: 0.02 mg/mL fibrinogen, 1.0 U/mL thrombin, 0.05 $\mu\text{g/mL}$ Factor XIII in HBS, 5 mM calcium.

Coverslips with fibrin reactions were placed in Petri dishes with a small square of wet paper towel to prevent drying, and incubated at 37° for 2 hours. Thrombin catalyzed the conversion of FXIII to FXIIIa and fibrinogen to fibrin forming crosslinked fibers as evidenced by gel electrophoresis. The reaction solution was removed and the samples washed twice with HBS. 24 nm volume-labeled red fluorescent carboxyl-coated microspheres (Stock comes at 2% solid in water; Invitrogen- Molecular Probes) were diluted 1/10,000 in HBS and added to each SS; samples were incubated for 5 minutes, washed with HBS, and used for experiments.

Samples were prepared with and without FXIII which is converted into FXIIIa in the presence of thrombin and calcium. Samples without FXIII were prepared in the same manner, without the addition of FXIII and incubating for only 15 minutes.

5.2.4 Stretching Networks with an AFM

The atomic force microscope (Explorer, Veeco Instruments, Woodbury, NY) rests on the manipulation stage of the optical epifluorescence microscope enabling simultaneous

AFM manipulation and optical data acquisition. Both OMCL-AC240TS-W2 (Olympus, Micro Cantilever) and RC150VB Biolever (Olympus, Asylum Research) AFM cantilevers (SiN) were used for manipulation. A detailed description of our setup and measurement can be found in prior publications (65, 66). Briefly, force data was determined through calibration of the lateral deflection signal. The angular optical sensitivity of the twisting mode is the same as for the bending mode given we have a geometrically symmetric quadrant photodiode and similar gain settings for each quadrant. The twisting mode optical sensitivity in deflection units is then determined using the specific geometry (length, tip length) of the cantilever. The lateral cantilever spring constant was calculated from cantilever/geometry and SiN materials constants. The AFM tip was controlled using the Nanomanipulator software (3rdTech, Inc. Durham, NC). The tip was set down inside the network and moved in one direction at 1 μ m/sec in 75nm increment steps to stretch the network. To facilitate the tracking of individual fibers using Video Spot Tracker software [Computer Integrated Systems for Microscopy and Manipulation at UNC-Chapel Hill (<http://www.cisimm.org/>)], networks containing 10-30 distinct fibers all around 1-10 microns in length and in the same plane of focus were selected for manipulation. Networks were stretched until failure of the pulled fiber; but when analyzing the data the network was considered to fail at the first rupture of any junction or fiber, not just the pulled fiber. Figure 5.1 depicts a diagram of the setup. Fiber diameters were determined with AFM topographical imaging. Fibers extending onto the ridge surface were located with fluorescence imaging. The AFM tip was engaged in proximity to the fiber and local imaging was performed.

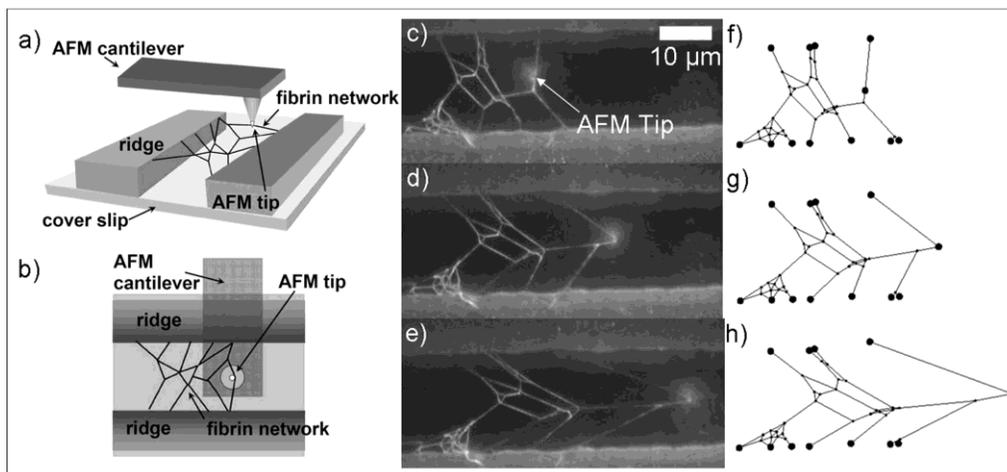


Figure 5.1 Experimental setup: (a-b) Side and bottom views depicting the AFM and a fibrin network suspended between ridges. (c-e) Three time point snapshots of a network undergoing deformation due to AFM manipulation. (f-h) Model network deformations at equivalent points to (c-e).

5.2.5 Network Strain measurements

Strain measurements on individual fibers within each network were calculated using video data recorded during the network manipulation. (See Figure 5.2) Strain was calculated for each fiber, measuring the change in the distances between junctions. In both FXIIIa ligated and unligated networks, fibers junctions were not observed sliding or slipping along the fibers during the pull; instead junctions appeared fixed with respect to the fiber. Data were only collected from samples for which fiber attachments to the ridge boundaries were fixed throughout the stretching. The fibers are adhered to the ridge through non-specific adhesion. Given the relatively low forces involved in the pulling experiments vs. the relatively high adhesive forces of the fibrin to the ridge, it is not surprising that they typically maintain fixed boundary attachments. Fibers pulled by the tip were assumed to converge to a point, for comparison with simulations. Manual distance measurements using the line segment measurement tool in the ImageJ (<http://rsbweb.nih.gov/ij/>) software package supplemented the automated software,

especially for junctions that Video Spot Tracker had difficulty tracking. Usually 4-8 fibers per network were tracked; the fibers tracked were selected based on original fiber length, focus, position relative to glue stamped surfaces and point of pull. We did not track fibers attached to the surfaces. Fiber strain was then plotted as a function of AFM tip movement. The derivative of the strain vs AFM step plots, dubbed the strain fraction, was calculated by taking the average slope between 10 points on strain vs AFM step data. (See Figure 5.2) The slope was then smoothed using the Matlab R2007b (The MathWorks, Natick, MA) *roess* function.

5.2.6 Using Webslinger to Simulate Network Stretching

Networks were simulated using Webslinger (<http://www.cisimm.org/>), a quasi-static mass-spring simulation which uses Euler integration with small time steps and damping to solve the system of partial differential equations governing a network of springs. After each step, the system is allowed to relax to equilibrium, and that solution is used for the initial conditions of the next step.

Arbitrary nonlinear spring behavior was specified by providing force-strain curves and rest lengths for each spring. For each experimental network geometry, a mesh model of the same geometry (node location accuracy within 150-200 nm) was created by hand from the video file image (Figure 5.1). Wall boundary nodes were simulated as immovable, but freely rotating pivot positions. The mesh node corresponding to location of the AFM tip was moved along the same trajectory taken by the tip in the experiment. (Figure 5.1) Patterns of fiber strain distribution within the simulated networks were then compared to those seen in the fibrin network experiments. Two force-strain curves were

used for the comparison with fibrin data, a linear model and a worm-like chain force model.

5.3 Results and Discussion

5.3.1 Results from Stretching Fibrin Networks

Using a combined fluorescence and atomic force microscope (AFM), we stretched fibrin networks to failure and measured tensile strain of individual constituent fibers. Focusing on small, (5-30 fibers) two dimensional fibrin networks suspended between micro-printed channels, we obtained a complete view of network geometry during deformation (Figure 5.1). For each network, the AFM tip was placed next to a fiber located at the periphery of the network, and that fiber was pulled in a direction parallel to the channel axis (Figure 5.2 a-d). Usually networks were pulled to the point of failure, but in some instances, networks slipped off the tip before failing and elastically recovered their original structure. Additionally, several experiments were performed where the network was stretched out twenty microns and then stepped back to its original position with no observable damage. These results indicate that network deformations were reversible for fiber strains at least equal to 1.0.

Five to ten fibers in close proximity to the point of pull were selected for strain measurements. The tensile strain of each tracked fiber in the network was plotted as a function of AFM tip movement (Figure 5.2e, Figure 5.3). At low strains, there was a noticeable transition regime in which all fibers began stretching more per incremental AFM tip movement. This is a geometrical effect; the fibers reorient, aligning in the direction of the applied force before entering a tensile stretching mode. When the most-strained fibers in the network reached strains above 1.0, there was a noticeable decrease

in strain per incremental AFM tip movement. That is, the change in strain with respect to tip movement decreased for the most-strained fibers (blue trace, Figure 5.2f). This derivative or “strain fraction”, reflects the share the overall network deformation taken on by each fiber per incremental AFM step. Accompanying the decrease in strain fraction of the most strained fibers, the strain fraction of lesser strained fibers commonly increased (green and cyan traces, Figure 5.2f). The two potential sources of this behavior are network geometry and individual fiber properties. In the former, the strain share (the distribution or “sharing” of the overall network deformation among the individual fibers) among fibers shifts as the network architecture reorients and gradually aligns in the direction of the applied force. Alternatively, the non-linear elasticity of the individual fibers redistributes the strain share as some fibers stiffen relative to the others. These two phenomena are not mutually exclusive and could both be contributing factors.

To assess the potential contributions of fiber non-linear elasticity to network deformation, we measured the force elongation behavior of individual fibrin fibers using a similar experimental approach (also see (65, 66)). Individual fibers were suspended between micro-printed channels and force data were obtained from calibrated lateral deflection of the AFM cantilever. The data showed a clear transition from low stiffness and linear elasticity at low strain to much higher stiffness at strains at or above 1.0 (Figure 5.4). This characteristic strain-stiffening force vs. elongation relation was used to quantitatively inform modeling of our network data.

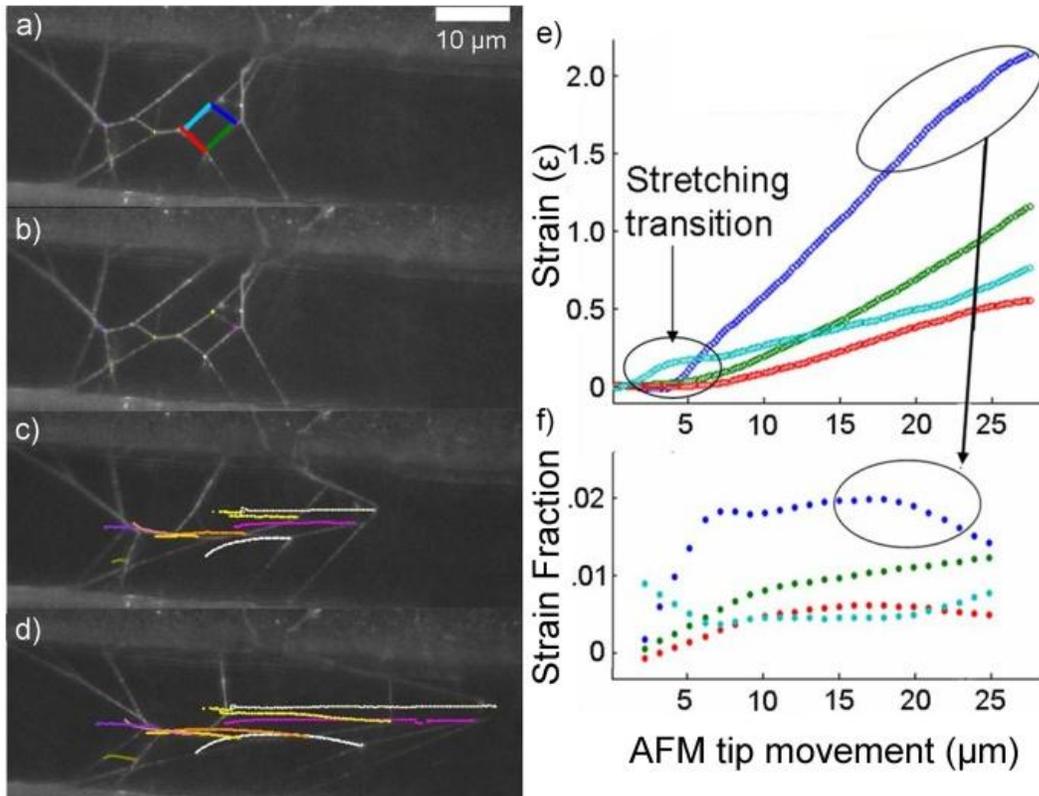


Figure 5.2 Network Strain Measurements: (a-d) Images from a video of network deformation. (a) The original network geometry displaying a color map of which fiber strains are plotted in frames e-f. (b-d) Fiber nodes in the network were selectively tracked, and fiber strains were calculated using the distances between the nodes. The colored lines in c-d show tracking data tracing the motion of network junctions throughout the deformation. (e) A plot of individual fiber strain vs. AFM tip movement. Each color represents the strain vs. AFM tip movement for one fiber corresponding to frame (a). (f) The strain fraction (derivative of strain with respect to tip movement, or slope, of plot in e). The circled data in the lower left of the plot in (e) indicates the stretching transition where fibers are reorienting and beginning to stretch. The other circles on the right side highlight a concave downward trend in the strain plot (e) and the corresponding decrease in the strain fraction (f) for the most strained fiber (blue trace). Note also that the strain fraction of the green and cyan traces trend upward in the strain fraction plot at high AFM tip movement. These trends indicate that as the most strained fiber stiffens, it transfers strain share to the less strained, softer fibers.

5.3.2 Network Modeling

To separate the effect of network geometry from the effect of the non-linear stiffness of the constituent fibers, experimental results were compared with quasi-static mass-spring simulations. Each experimental network was modeled by a simulated network of equivalent geometry. AFM stretching was replicated by incrementally moving the point

(a mass node in the model) corresponding to the AFM/fiber contact in the experiment. AFM imaging of fibers extending onto the structured surface ridges indicated an average diameter of 80 ± 20 nm (we found no statistical difference between ligated and unligated clots). However, due to the limited feasibility of accurately determining fiber diameters within the network before the pull (fibers suspended between the ridges could not be imaged with the AFM and the SEM preparation would alter network mechanical properties), we have assumed that all fibers have equal diameter. Though the equal diameter assumption is not ideal, it is supported by fluorescent confocal microscopy and SEM studies of in vitro clots that show a relatively narrow variation in fiber diameter (15-30% (63, 64)). As a consequence of this assumption, each fiber segment has a common force vs. strain relation. Two force extension models were implemented: a linear springs model and a non-linear strain stiffening model.

Within the linear springs model, each fiber segment is a linear spring with a spring constant inversely proportional to its original length (as is the case for an ideal spring), $k = \kappa/L_0$, where κ plays the role of spring constant in a force vs. strain relationship. This provides each fiber segment with the same force vs. strain relation, regardless of its length ($F = \kappa \epsilon$).

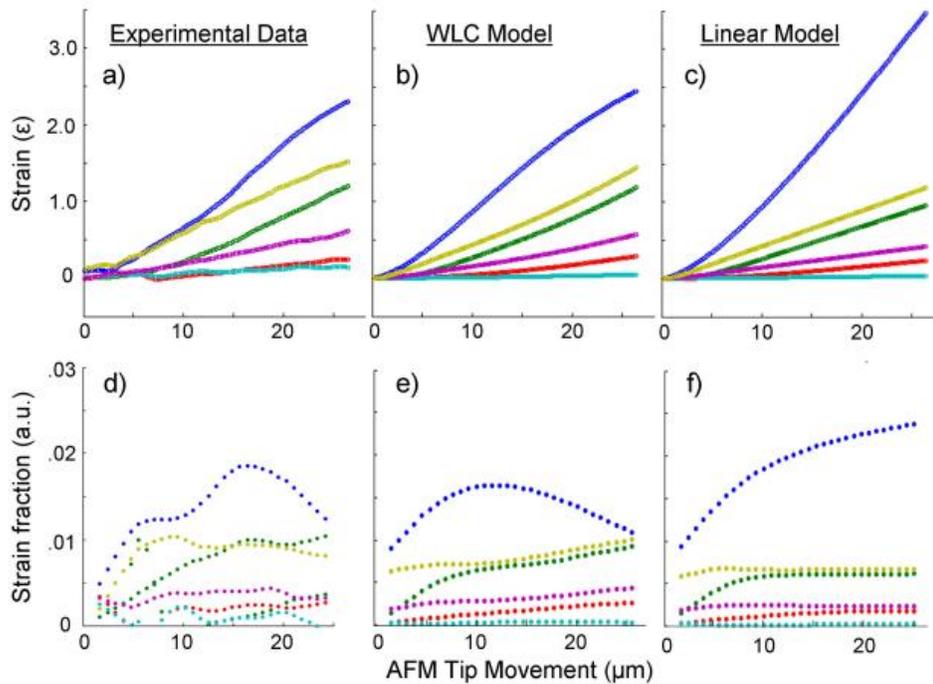


Figure 5.3 Experiment vs. Simulations: a-c) Fiber strain traces for a particular experimental network and for simulations of equivalent WLC and linear model networks. d-f) Plots of the strain fraction corresponding to plots a-c. The experimental strain and strain fraction data show a much closer correspondence to the WLC model than to the linear network particularly for the most strained fibers (blue, yellow, green) at high AFM tip movement. Two predominant features contrast the experiment and WLC model data from the linear model. First, the linear model shows much higher maximum strain (blue trace, c) than either experiment (a) or WLC model (b). Second, within the strain fraction plots, the most strained fiber (blue trace) shows a clear decrease above tip movement of 15 microns for both the experiment and WLC model, while the linear model shows no such increase. The strain fractions of the fibers in the linear model approach constant and highly dispersed values indicating each fiber takes on constant and inequitable strain share. The strain fraction of the experimental and WLC model fibers converge into a much narrower range at high AFM tip movement indicating that strain share is transferring from the most strained to the lesser strained fibers, more equitably distributing strain throughout the network.

We modeled the non-linear single fiber force vs. extension behavior (Figure 5.4), using an ideal chain model known as the worm-like chain (WLC), in accordance with our α C stretching model described in 2.3.3

$$F = \frac{k_b T}{P} \left[\frac{1}{4} \left(1 - \frac{\Delta L}{L_c} \right)^{-2} - \frac{1}{4} + \frac{\Delta L}{L_c} \right] \quad (5.1)$$

Where, ΔL is extension, T is temperature, P is persistence length, and L_c is the contour length, or the length of maximal fiber extension. The persistence length, P, is a measure of the bending rigidity of the chain (the fiber in our case). The contour length, L_c determines the shape of the force vs. extension curve, most importantly, the onset of strain stiffening. The red WLC fit in Figure 5.4, combined with the discussion in Chapter 2. Evidence that the α C Region is the Origin of the Low Modulus, High Extensibility, and Strain Stiffening in Fibrin fibers, show that the Marko-Siggia WLC is a more than adequate analytical expression for modeling the single fibrin fiber mechanics. Alternative force extension models that have been very successfully applied to semi-flexible biopolymer networks such as actin (185) don't fit our fibrin fiber data as well as the Marko-Siggia WLC.

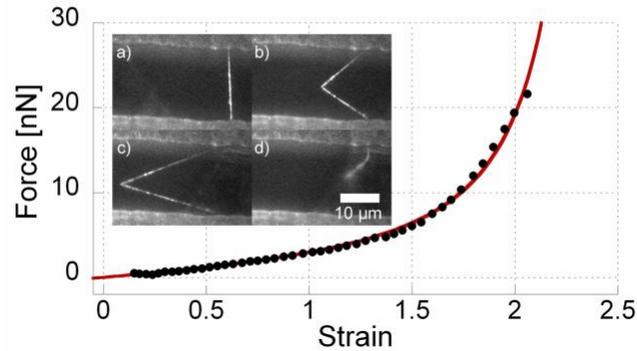


Figure 5.4 Single fibrin fiber force curve: Inset: a-d) depict a single suspended fiber stretched to breaking (d) by an AFM tip. Plot: Black point depict single fibrin fiber force-strain data determined by calibrated lateral AFM force measurement. The geometrical aspects of the measurement were taken into account to convert raw AFM force data to fiber tensile force. The strain of the fiber was determined from the calibrated video data. The red trace depicts the WLC fit using two fitting parameters: the persistence and contour lengths.

In the context of eqn. 5.1, each fiber having identical diameter corresponds to a common persistence length, P . Because P simply changes the force scale of the force vs. extension properties, its value will not affect the distribution of strains within a network (though it will affect the magnitude of the force required to strain it). On the other hand, the contour length parameter, L_c , does significantly affect the stretched network configuration and strain distributions, and is the parameter that was varied to best match our experimental observations.

To give each fiber a common force vs. strain relation, we modified eqn. 5.1 to provide force as a function of strain.

$$F = \frac{k_B T}{P} \left[\frac{1}{4} \left(1 - \frac{(\Delta L / L_0)}{(L_c / L_0)} \right)^{-2} - \frac{1}{4} + \frac{(\Delta L / L_0)}{(L_c / L_0)} \right] \quad (5.2)$$

$$F = \frac{k_B T}{P} \left[\frac{1}{4} \left(1 - \frac{\varepsilon}{(L_c / L_0)} \right)^{-2} - \frac{1}{4} + \frac{\varepsilon}{(L_c / L_0)} \right] \quad (5.3)$$

L_c/L_0 is the adjustable parameter that determines the onset of strain stiffening. Using $(L_c/L_0) = 5$ (meaning the fibers would reach their contour lengths around strains of 4.0, although they fail before this point) provided the best correspondence to the experimental network strain distributions (see Table 5.1). This contour length comes not from bending and straightening of the fiber itself, but from stretching out of internal molecular scale flexible segments within the sub-fiber structure.

5.3.3 Network Strain Distribution Narrowing

A set of fibrin networks were stretched to the point of breaking and the strains of the constituent fibers were analyzed. For each experimental network, the distributions of fiber strains were compared to those from equivalent linear and WLC model network simulations. Figure 5.3 compares the experimental and simulated strain data for a representative network. The experimental and WLC strain distributions differ markedly in several respects from the network of linear springs. First, the strain of the most strained fiber (blue trace) in the linear model is significantly higher than in the experimental data or in the WLC simulation at large AFM tip movement (Figure 5.3, a-c). Second, the strain fractions of the most strained fiber in both the experimental data and WLC simulation reach a maximum and then decreases above strains of ~ 1.0 , whereas no maximum is seen in the linear model (Figure 5.3, d-e). These differences were pronounced in all networks studied. The absence of decreasing strain fraction in the linear simulations indicates that this effect is due to the non-linear force vs. extension

properties of the constituent fibers rather than to the generic effect of reorientation and alignment of the fibers.

Strain distributions were compared at two points: “intermediate strains”, where the AFM had moved $15\mu\text{m}$ from its original location, and “high strains”, where the network failed (See inset of Figure 5.5). The intermediate strain point was chosen to fall just within the expected pre-stiffening linear elastic regime for all fibers. Figure 5.5 shows the strain distributions for nine different networks of FXIIIa-ligated fibers at both intermediate and high strains. Though the variations in results from network to network indicate that the distribution of strains depended on the specific network geometries, strong trends bridging all measured networks emerged from the analysis.

At intermediate strains (Figure 5.5 top) the model and experimental data were statistically similar. Within uncertainties, the maximum strain (strain of most strained fiber) and the standard deviation in strain agree for the experiment, WLC, and linear model (see Table 1). These results are consistent with the expectation for intermediate strain: all fibers were within the linear regime of elasticity (for experiment, WLC and linear model networks).

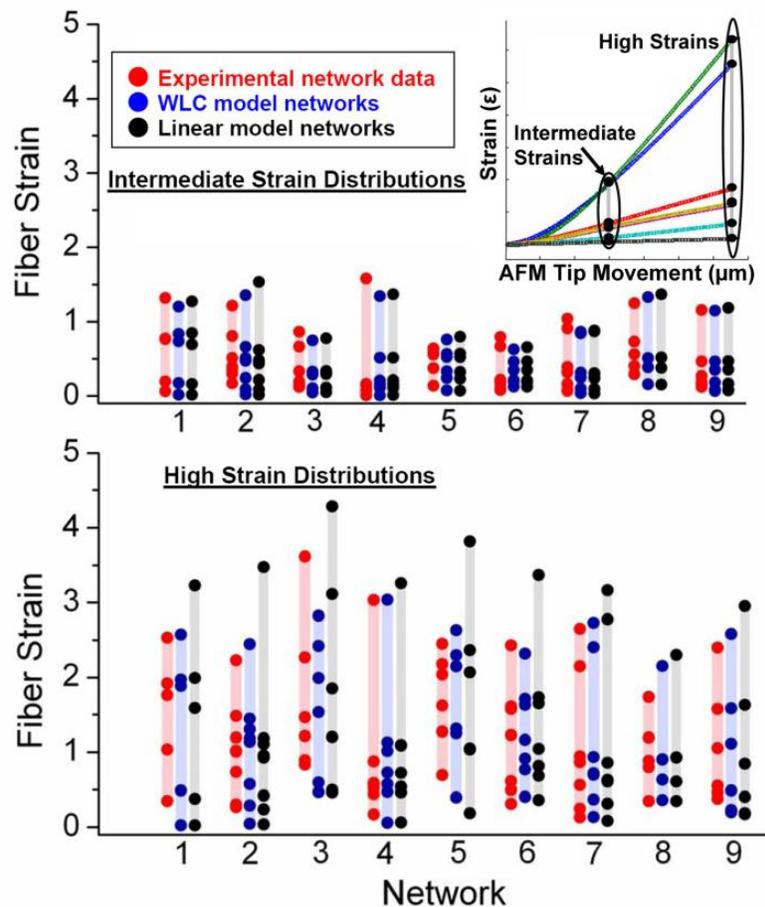


Figure 5.5 Ligated network strain distributions : A compilation of model vs. experimental strain distributions for nine different network geometries formed with FXIIIa ligation. Each dot in the plots indicates the strain of a fiber within a network. The shaded bars underneath the data points are simply guides to the eye to emphasize the strain ranges. Measurements were made after 15 μm of AFM movement (intermediate strain, top plot) and at the point of network failure (high strains, lower plot) (see inset). At intermediate strains there are clear variations, but no clear trends distinguish the experiment and models. At high strains, the linear model shows much higher maximum strain for all networks and lower minimum strain than for all but two of the experimental networks. There is much closer correspondence between the experimental and WLC distributions (see Table 5.1 for statistical analysis).

On the other hand, the data and models showed distinct differences in the high strain data (Figure 5.5 bottom). The linear distribution had higher maximum strain in all cases and the lowest minimum strain for most of the networks. Comparing the most-strained fibers of each network, there was on average $31\pm 6\%$ higher maximal strain in the linear model as compared to experiment, whereas there was only a $3\pm 4\%$ difference between the WLC model and experiment (Table 1). As a measure of the narrowness of the strain distributions, we calculated the standard deviation in fiber strain for each network. The average standard deviation in the experimental data was 30% less than that of the linear network (1.1 vs. 0.82), and was still 20% lower when removing the most and least strained fiber in each network from analysis. This finding indicates the narrowing of the strain distribution occurred network wide and was not due exclusively to the behavior of the most-strained fiber. The experimental networks behaved in a manner much more consistent with the WLC model at high strains, which indicates, not surprisingly given the single fiber data (Figure 5.4), that the most strained fiber in each network experienced significant strain stiffening prior to failure. As the network transitioned from the linear regime to the strain stiffening regime, the non-linear elasticity of the most-strained fibers reduced the maximum fiber strain as compared to a linear network, narrowing the strain distribution and preventing strain from being concentrated into only a few fibers. Though the gross features of the strain distributions showed good agreement between experiment and WLC, there are discrepancies in a fiber by fiber comparison. This discrepancy may be largely due to assumptions in the model, in particular, the use of a universal fiber diameter.

As a control, model networks composed of linear springs of varying spring constants were also tested vs. the experimental results (See Figure 5.6). While the distribution of strain is altered by varying fiber stiffness within the network, the decrease in strain fraction at high strain seen in the experimental networks does not occur. Instead, the strain fraction of fibers with a linear force-extension relation approaches a constant asymptotic value regardless of stiffness. Thus, for a network of fibers of linear but variant stiffness, once the fibers align in the direction of the pull, each fiber assumes a constant share of the strain.

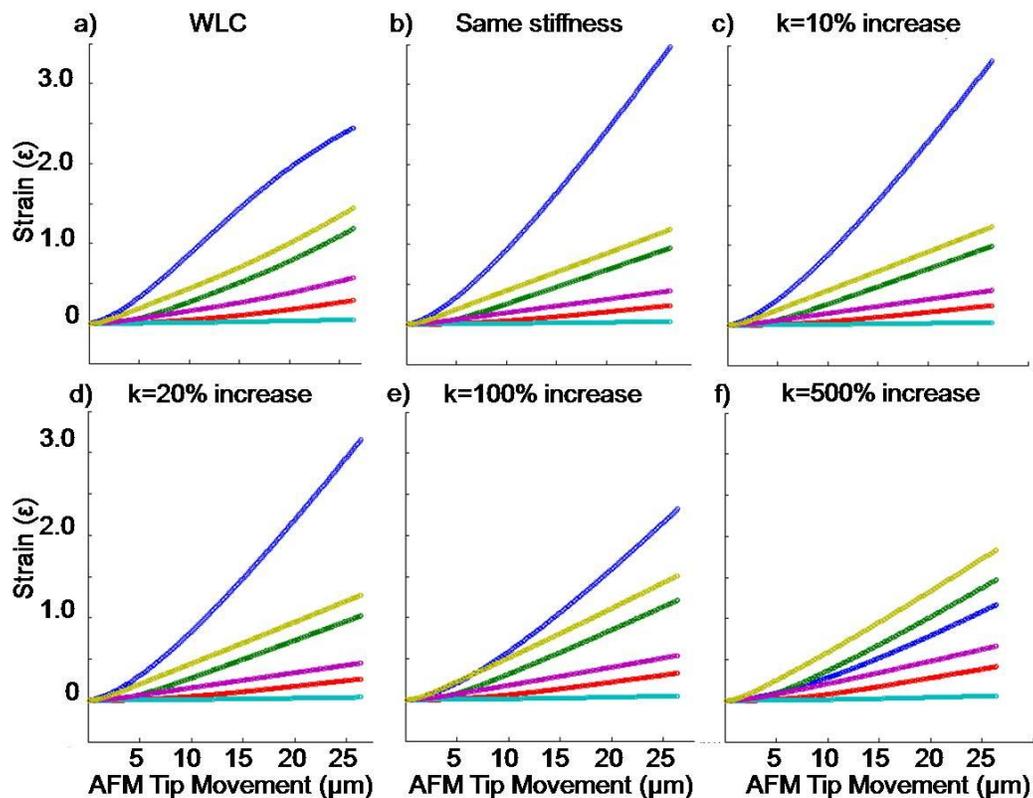


Figure 5.6 Linear springs with different stiffness: To determine whether strain distribution shifts require fibers with non-linear stiffness or can occur in linear networks of fibers of variable stiffness, we ran a series of simulations varying the stiffness of the most strained fiber while holding the stiffness of all other fibers constant. Panels (a-f) show the results of these simulations for one network geometry (corresponding to Figure 5.3). Panel (a) and (b) display the results from the worm-like chain modeling and from a simulation where all fibers were modeled as linear springs with the same stiffness respectively. In (c-f) the same simulation was run as in (b), except the stiffness of the most strained fiber (a-b, blue trace) was increased by: 10% in (c), 20% in (d), 100% in (e) and 500% in (f) while keeping all other fibers at the original stiffness. As the results show, the distribution of strains shifts: in figure (f) the blue fiber is no longer even the most strained fiber, but in no simulations is there any hint of the strain fraction decrease observed in the WLC model (panel (a)). Instead, for all fibers following a linear force extension behavior, the slope of the strain vs. tip movement plot asymptotically approaches a constant value. This fact held true for all network geometries tested. This leads us to conclude that the strain distribution narrowing effect comes from non-linear force extension behavior of the fibers, not differences in stiffness.

Physiological fibrin networks are formed in the presence of the transglutaminase FXIIIa, which incorporates covalent bonds between monomers - a process known as cross-linking or ligation. All of the experiments and data described thus far (Figure 5.5) reflect work on ligated fibers and networks. Studies investigating the effect of FXIIIa ligation on clot structure and mechanics reveal that ligated clots are stronger and more resistant to plasminogen mediated dissolution (lysis) (10) than their non-ligated counterparts (63). In addition, ligation also affects the mechanics of the individual fibrin fibers in both stiffness and extensibility (64, 66). As a comparison to the work with ligated networks, networks formed without FXIIIa were also stretched to gain insight into the mechanical influence of ligation. (Figure 5.7 and Table 1) These networks failed at significantly lower strain values than their ligated counterparts (max strain 1.82 ± 0.11 unligated, 2.56 ± 0.18 ligated). However, the same trends were observed at high strains: maximum fiber strains and strain standard deviations were lower for the experiment than in the linear models (16% and 17% lower respectively) and in agreement with the WLC model.

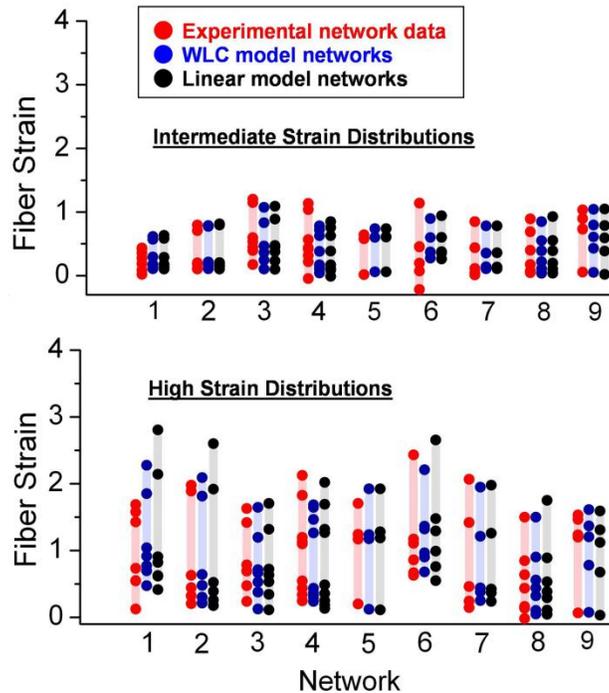


Figure 5.7 Unligated network strain distributions: A compilation of model vs. experimental strain distributions for nine network geometries formed without FXIIIa crosslinking.

5.3.4 Discussion

These studies indicate that strain stiffening of individual fibers plays a significant role in the larger scale response of fibrin networks. A closer look at the details in the plots of Figure 5.3 lends insight into the mechanisms responsible for the lower maximum strain and narrowing strain distribution of the fibrin networks as compared to the linear models. For the linear model, the fiber strain fraction vs. AFM tip movement curves asymptotically approach a constant value at high strain (Figure 5.5f), indicating each fiber supports relatively constant share of the strain. For networks of strain-stiffening fibers, the most strained fiber's plot trends downward indicating that the strain is shunted from areas of high strain (which have become relatively stiff) to areas of lower strain (which are relatively soft). This allows the maximum fiber strain for a given imposed network deformation to be significantly lower for the strain stiffening system than for the

linear network model. The failure of the most strain fiber is a crucial event as it relates to clot strength. This initial failure will further distribute stress that may lead to severe strain concentrations and a cascading catastrophic failure of the entire network. The stiffening of the individual fibers distributes strain throughout the network, preventing the strain from being concentrated in the few lead fibers and allowing larger imposed network deformations to be accommodated by the network prior to fiber failure; the network is therefore effectively strengthened (See Figure 5.8).

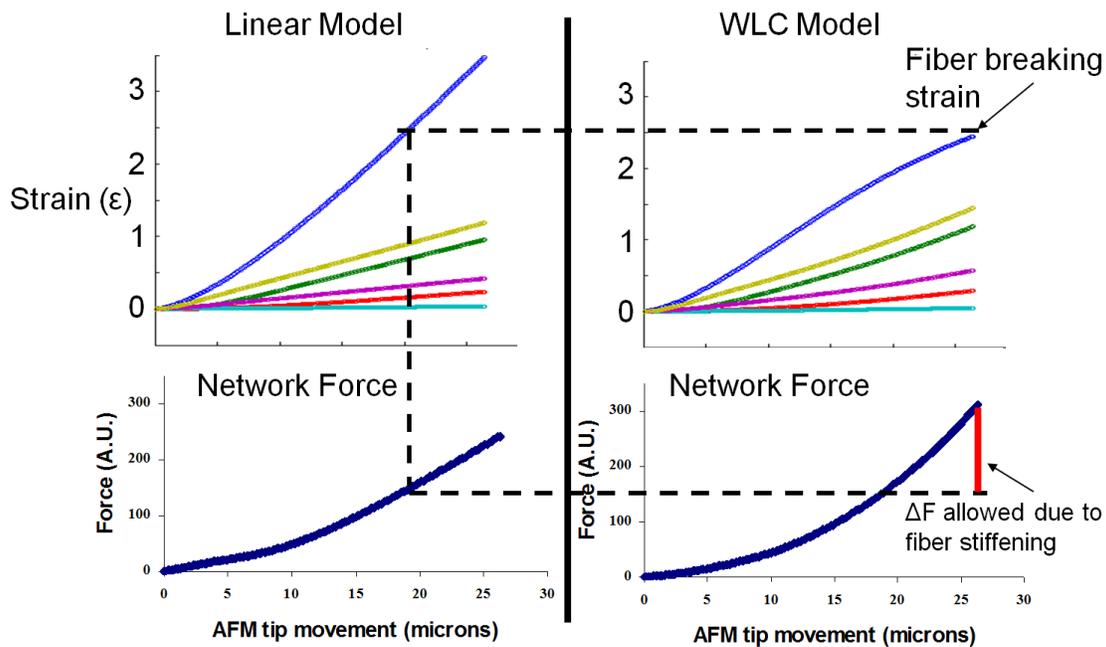


Figure 5.8 The Network Strengthening Mechanism of Strain Stiffening: Strength is defined as a material’s ability to withstand force before failing. To determine whether strain stiffening of individual fibers increases network strength, simulations of networks composed of linear fibers (left side of figure) and strain stiffening fibers (right side) were run, and the forces required to pull the networks was compared. As can be seen in the top plots, the distribution of strains in the network is compressed in the WLC networks. If one assumes that fibers fail around strains of 2.5, it becomes obvious that the WLC networks can be stretched an additional 7 microns before the first fibers cross the breaking strain threshold (dashed line). If one looks at the forces corresponding to the point where fibers begin failing, the WLC networks can withstand nearly double the force before fiber failure. This is clear evidence that strain stiffening does indeed increase network strength. Interestingly, looking at the two force curves, it can be seen that the linear model network exhibits two linear regimes of stiffness (likely corresponding to a transition from network re-orientation to stretching), while the WLC network is continually non-linear. This behavior is evident in experimental fibrin network behavior as observed in Figure 5.9.

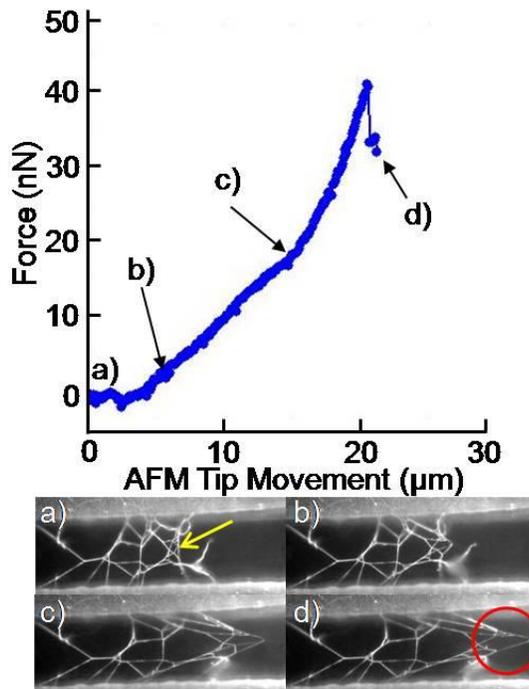


Figure 5.9 Force vs. Displacement of a fibrin network: Top: The force vs. AFM Tip Movement plot for a fibrin network manipulation. Four points are highlighted (a-d) corresponding to the four network displacement images shown at the bottom. The yellow arrow in figure (a) highlights the initial location of the tip. The abrupt force decrease at point (d) corresponds to the first fiber failure within the network. The red circle highlights the fibers that have snapped between frames c and d as evidenced at point (d) on the force plot. The force data shows the stiffness of the entire network during the deformation. The stiffness changes once the fibers align in the direction of the pull, and at high strains the network stiffens due to the strain stiffening of fibers within the network. In simulations of linear network deformations, the high strain network stiffening was not observed; instead, force-extension is linear (See Figure 5.8).

Our results underlined the importance of FXIIIa ligation on network strength (64). Fibers within ligated networks reached strains 30% higher than those in unligated networks; this result is consistent with previous measurements on single fibers (65, 66). Here however, we showed that FXIIIa ligation has a direct affect on fibrin network strength. Though biochemical factors affect many important aspects of fibrin clots (e.g. fiber diameter, cross link density etc.), attention must be paid to how these factors affect the mechanics of the constituent fibers. Subtle biochemically induced changes in the force vs. extension behavior of the individual fibers may have profound influence on overall clot mechanics.

This work also sheds light on the question of relative fiber strength vs. junction strength (66). Ligated networks failed first at junctions ~30% of the time, fibers ~40% with 30% of videos inconclusive (n=12). Unligated networks failed first at junctions ~45% of the time and fibers ~45% with 10% of videos inconclusive (n=9). In addition, the average value for the breaking strain of the most strained fibers in the network (2.56 ± 0.18 FXIII, 1.85 ± 0.11 non-FXIII) agreed within uncertainties with published values taken from single fiber measurements of fibrin extensibility (3.32 ± 0.71 ; 2.17 ± 0.47 FXII, 2.26 ± 0.52 non-FXIII) (66, 77). Although sample sizes are small, this data indicates that fibrin fibers and junctions have comparable strength. More comprehensive work is needed to quantify the strength of fiber junctions and validate this small data set.

5.4 Conclusion

To our knowledge, these are the first measurements attempting to quantify how individual fiber strains are distributed in a network. We have demonstrated that the fiber strain stiffening measured in

Chapter 2. Evidence that the α C Region is the Origin of the Low Modulus, High Extensibility, and Strain Stiffening in Fibrin fibers and characterized by the α C stretching model plays an observable role within a deforming network: strain stiffening equitably distributes strain in the network. As fibers are stretched, they become stiffer than any surrounding fibers at lower strains; this allows the more strained, stiffer fibers, to distribute the strain load to the less strained fibers and reduce strain concentrations. As network strength is directly related to the failure of individual fibers with the network, this reduction of strain concentration effectively strengthens the network. This non-linear regime of fiber elasticity is achieved prior to network failure and is relevant to understanding fibrin network strength. In addition, FXIIIa ligation directly affects network strength by increasing the extensibility of individual fibers within a network. The strain concentration reduction effect may be an important mechanism in vivo, enhancing blood clot strength under the high shear conditions of the blood stream.

	INTERMEDIATE NETWORK STRAIN				HIGH NETWORK STRAIN			
FXIII	Max fiber strain	% change in max fiber strain	Avg. fiber strain	Fiber strain Std. Dev.	Max fiber strain	% change in max fiber strain	Avg. fiber strain	Fiber strain Std. Dev.
Experiment	1.09±.10	-----	0.443	0.373	2.56±.18	-----	1.223	0.816
WLC	1.04±.10	-4±5%	0.414	0.355	2.58±.09	3 ± 4%	1.223	0.854
Linear	1.09±.11	1±6%	0.419	0.374	3.31±.17	31 ±6%	1.300	1.119
Non-FXIII								
Experiment	0.90±.08	-----	.428	.356	1.82±.11	-----	0.922	0.626
WLC	0.84±.05	2±7%	.400	.294	1.88±.09	4 ± 4%	0.920	0.623
Linear	0.87±.05	1±7%	.403	.304	2.11±.14	17 ± 7%	0.936	0.728

Table 5.1: A comparison of network properties of 18 networks stretched in these experiments: 9 networks were ligated by FXIII, and 9 were not. Intermediate network strain is defined as the point where the AFM had moved 15 μm ; High network strain is defined as the point of network failure. Qualitative differences were observed in the behaviors of the two types of networks, especially in the values of the most strained fibers of the network (bold) and the standard deviations. Uncertainty values are listed as standard errors. Max fiber strain: The average value of the most strained fiber in every network. % change in max fiber strain: On a fiber to fiber basis, the average percent change between the most strained fiber in the experimental network with the most strained fiber in the corresponding model network. Avg. strain: The average fiber strain over all of the networks. Fiber strain Std. Dev. : The standard deviation in fiber strains over all networks.

Chapter 6. Summary and Future Directions

6.1 Conclusions and Physiological Relevance

6.1.1 The Material Properties of Fibrin

In this project we have explored the mechanical properties of fibrin across its entire structural hierarchy in an attempt to create a bottom up understanding of fibrin mechanics. It is claimed that Isaac Newton once said “If I have seen further, it is only by standing on the shoulders of giants.” In the same manner, this work does not stand alone, but arises from the shoulders of the giants. The basis for this work comes from three previous observations about fibrin mechanics: 1) Fibrin fibers have remarkable elasticity and extensibility(66), 2) the extensibility of fibrin fibers appears directly related to the length of the tandem repeat region in the α C region of the fibrin molecule(95), and 3) fibrin networks strain stiffen.(53, 54, 79, 174) With this basis we’ve proceeded to produce a mechanical model for the fibrin fiber and fibrin networks based on the mechanical properties of the fibrin molecule, particularly the α C region.

In chapter 1, the groundwork was laid for the development of a molecular model for the fibrin fiber. The model must include as many as possible known details about fibrin structure, polymerization and the topology of interactions between molecules. In addition, the model must incorporate the known mechanical properties of the fibrin molecule. A comparison was made between the known mechanical properties of fibrin fibers and the simplistic models for fibrin extension. In each case, the models were incomplete in their description of the mechanical properties of the fibrin.

In chapter 2, the mechanical properties of fibrin fibers with and without FXIIIa cross-linking were explored. Fibrin fibers strain stiffen at strains above 100%, and have an average modulus of 1MPa for unligated fibers and 2MPa for ligated fibers at low extensions rising to 6MPa and 10MPa respectively at the point of failure. This stiffness is in line with other elastomeric materials such as rubber and elastin. (92, 125) A simple mechanical model was proposed based on modeling the α C region of the molecule as a worm-like chain, and thinking of the fiber as a series of stiff and soft spring in series and in parallel down the fiber. While the model fit the data well, it did not include other parts of the fibrin molecule and did not comment on modes of plastic deformation or fiber failure.

Besides the mechanical properties of the fibrin fiber, the dynamical properties of the fiber could play an important role in the fiber's response to force. In addition, measuring fibrin fiber dynamics could clarify potential molecular mechanisms of extension. The recoil dynamics and their potential molecular sources were described in chapter 3. Using an AFM and high frame rate fluorescence photography, the elastic recoil of fibrin fibers was measured. It was discovered that the fibers exhibited two regimes of elasticity: 1) An initial fast recoil which occurred on timescales of 500 μ s, and a slower tensing regime on the order 1-10 ms. The 500 μ s timescale eliminates protein re-folding as a mechanism for the fast regime, but the 1-10ms regime could involve protein folding. (133) The fast regime likely comes from an entropic recoil of an unstructured polypeptide such as the α C region. This could also come from a stretched coiled coil, but DMD simulations indicate the coiled coil stretches into a stable interchain β -sheet conformation which prevents the recoil; based on simulations and experiments of other

materials such as keratin, if this α -helix to β -sheet transition should be irreversible if it occurs in fibrin, eliminating the coiled-coil hypothesis altogether.(120, 121)

To probe the mechanical stability of additional parts of the fibrin molecule, we performed a series of steered discrete molecular dynamics simulations. First a homology model of for the α C region was generated. Then the α C model, the coiled-coil region, and the γ -C and γ - γ interface were stretched under force. The results of these simulations indicate a stability hierarchy for the fibrin molecule, with the γ - γ interface opening at the lowest forces, followed by α C unfolding, then γ -C stretching and finally coiled coil. These results prompted a new mechanical model for the fibrin fiber deemed the SL α CK model. As can be seen in Table 6.1, the SL α CK model is able to account for most known mechanical properties fibrin.

Finally, to test the role of individual fiber mechanics on network mechanics, a series of tests stretching fibrin networks were performed, and the individual fiber strain was tracked as a function of network strain. The fiber strain results were compared to model networks of linear and strain stiffening fibers. The results indicated that fiber strain stiffening plays the role of reducing stress concentrations and more equitable distributing force throughout the network. Thus, strain stiffening, which in the SL α CK model arises primarily from stretching out the fibrin α C regions, plays a role of strengthening fibrin networks. The mechanical properties of the individual monomer, therefore manifest themselves at the highest level of the fibrin hierarchy, the network level. This knowledge could provide insight in how to design stronger biopolymer materials in the future.

Fiber/Network Property	Coiled-coil model	α C Model	Transverse Ligation Model	SLaCK Model
Extensibility of 200% strain (Liu, 2006)	●	○	●	●
Stiffening after Ligation (Collet, 2006, Houser, 2010)	●	●	●	●
Fiber Strain Stiffening (Hudson, 2010, Liu, 2010)	●	●	●	●
Reversible extension of ligated fibers (Liu, 2006)	○	●	●	●
Reversible extension of un-ligated fibers (Lieu, 2006)	○	○	●	●
Sub Millisecond Recoil (this report)	○	●	○	●
Lateral Contacts between protofibrils (Weisel, 1986)	●	○	○	●
1MPa stiffness of fiber (Houser, 2010)	●	●	○	●
Force/monomer ~5-10pN at low strains (Houser, 2010)	●	●	○	●
A:a bonds withstanding >50pN forces (Averett, 2008)	●	●	●	●
Tension in Fibers	○	○	○	○
Negative clot compressibility (Brown, 2010)	●	○	○	○
SAXS 22nm banding pattern widening with stretching (Brown, 2010)	○	○	●	●
fibrin gamma domain stretching (Averett, 2008)	●	●	○	●
Fibrin Network Strain Stiffening (Brown, 2010, Ferry, 1974)	●	●	○	●

- Explained by model/included in model
- Data Inconsistent with model
- Model could be consistent, but more information needed
- No component of model to describe this data

Table 6.1: A comparison of the potential models for fibrin mechanical properties with the SLaCK model: A green circle represents a molecular, fiber level, or network level property that is explained by the model, a red filled circle represents a property that is not explained by the model, a red open circle represents a property that may be explained by the model but needs more investigation and a black open circle represents a property not addressed by the model.

The results from these investigations on all levels of the fibrin hierarchy lead to the conclusion that fibrin is a very remarkable material. Fibrin fibers are one of the most extensible materials in nature, and yet can recover their taut conformation on ms timescales after stretching. The molecule itself contains an unstructured region, a coiled coil region, and a globular region, each of which may play a different mechanical role in the fiber. Fibrin networks have an inherent strengthening mechanism built in by the strain stiffening properties of individual fibers. Taken as a whole, fibrin as a material appears designed to function in a dynamic environment in which continued resilience after multiple loading and unloading cycles is required.

6.1.2 Physiological Relevance

While this work has shown that fibrin has amazing materials properties that can be understood on all levels of the fibrin hierarchy, it is important to step back and analyze the physiological relevance of these properties. Stresses are applied in blood clots through the shearing of blood at the walls of blood vessels. The wall shear stress for a Newtonian fluid like blood can be obtained with the simple relation:

$$\tau_w = \dot{\gamma}\mu \quad (6.1)$$

Where $\dot{\gamma}$ is the shear rate parallel to the wall surface and μ is the fluid viscosity. The viscosity of blood is approximately 4×10^{-3} Pa·s, and the shear rate in arteries can range from 200 s^{-1} up to $100,000 \text{ s}^{-1}$ in severe areas of stenosis (constricting of the arteries due to disease such as atherosclerosis).⁽¹⁾ These correspond to wall shear stresses of 0.8 – 400 Pa. Fibrin fibers constitute only 0.25% of the total clot volume in physiological clots, and this value may range as low as 0.0025% based on fibrin's polymerization properties.⁽⁸⁾ Assuming a fiber modulus of 1 MPa, these shear stresses generate clot shear

strains of to 0.03% to 1600%, a huge range. These simple calculations indicate that fibrin stretching does not play a large role in hemodynamic response under normal shear rates, but plays an increasing role at higher shear rates in areas of stenosis. Interestingly, it's been noted that at low shear rates blood clots consist mostly of red blood cells and fibrin, while at high shear rates, platelet aggregation dominates thrombus formation.(186-188)

Nature is rarely design optimized for a set of properties such as high extension and strain stiffening that go unutilized. (86, 189) In the above calculation, an implicit assumption was a uniform distribution of stress throughout the network. If this is not the case, then locally within a network, stress concentrations could require larger fiber extensibilities. In this case, the mechanism of network strengthening arising from fiber strain stiffening described in Chapter 5 may play an important role in clot viability. Another important reason for fibrin elasticity could come from the polymerization of networks under flow conditions. It was recently shown using in vitro studies that networks formed under shear rates as low as 20s^{-1} show a nearly complete fiber alignment in the direction of flow.(50, 190, 191) The ability of αC domains to interact and stretch may play an important role in fibrin network formation under flow. Finally, after clot formation, platelets drive a retraction process in which the clot can shrink in volume by more than an order of magnitude.(192) Platelets induce retraction with a fibrin fiber loading rate of 10Pa/minute, which could lead to significant fiber deformations over the course of several hours.(193) We can conclude from this brief analysis that the extreme mechanical properties of fibrin may not play a role under normal shear rates, but could play a role in high shear rates, during clot formation, and during clot retraction.

6.2 Future Directions

6.2.1 Measuring the mechanical properties of recombinant fibrinogen variants

The work in this study has highlighted the potential mechanical role of various parts of the fibrin molecule. These studies must be followed up by work on both the single molecule level, fibrin level, and network level. As has been discussed earlier, fibrin is a complex hierarchical material, and the mechanical properties on the single molecule level could affect higher level properties in different ways. The effects of each part of the fibrin molecule can be tested by forming fibers and networks with mutated variants of the fibrin molecule. There are many naturally occurring mutants of fibrin, some of which have little effect on the efficacy of blood clotting, and others that are greatly deleterious to the normal functioning of the coagulation cascade. Two fibrinogen variants, γ' and α_E , actually compose 8% and 1% respectively of the total fraction of fibrinogen molecules in normally functioning adults. (194-196) While it is possible to find naturally occurring fibrinogen variants with specific mutations to the fibrin coiled-coil, γ -nodule, and α_C regions, these mutations are often heterozygous with the altered fibrinogen unable to polymerize or sometimes even be assembled.(11, 197, 198) A more specific method would be to make the fibrin proteins recombinantly, targeting specific amino acids for mutation.

Two main parts of the molecule that could be tested with recombinant protein technology are the α_C region and the coiled-coil parts of the molecule because of their relatively “simple” folds. Presumably the simplest part of the molecule to test would be the α_C region, as it can be shortened or lengthened in the repeat region without affecting any folding (although admittedly, with such a large protein it is difficult to predict the exact consequences of any alterations in amino acid sequenc). We have already begun

working with two recombinant proteins with altered α C regions, A α 251 (25) and a human-chicken hybrid (HC) protein. In A α 251, the amino acid sequence of the α chain was truncated after amino acid 251. In the human-chicken hybrid fibrinogen, the α chain contained amino acids 1-197 of the human fibrinogen molecule, and amino acids 198-490 of the chicken fibrinogen molecule. Surprisingly, A α 251 polymerized into fibers fairly normally, and fibers had extensibilities similar to fibers composed of normal human fibrin molecules. On the other hand, fibers composed of the HC protein polymerize abnormally as can be seen in both turbidity and electron microscopy measurements, and additionally the fibers were very inextensible.

This work needs to be followed up with several additional recombinant proteins targeting the α C region. I would suggest generating fibrinogen molecules with the following characteristics: 1) A α 199, where the α chain will consist of amino acids 1-199 of the human α chain sequence. This will eliminate any ambiguity arising from A α 251, because the known crystal structure of the fibrinogen molecule truncates at amino acid 200. Additionally, A α 251 contains potential cross-linking sites in Gln200, Gln221, Gln223, Gln237 and Lys206, Lys208, Lys219, Lys224, and Lys230 which would be eliminated in A α 199, isolating the effects of FXIIIa cross-linking to the γ chain. 2) A Chicken-human hybrid molecule consisting of amino acids ch1-196 of the chicken fibrinogen α chain, and hu196-610 of the human α chain. This will serve as a good comparison to the human-chicken hybrid data. In addition, we can compare it to the mechanical properties of purely chicken fibrin fibers. It should be noted that this variant may not polymerize correctly since the HC protein produced morphologically different fibers. 3) Hu-tdr, which will consist of amino acids hu1-220 juxtaposed with amino

acids hu392-610 of the human α -chain. This takes out the unstructured tandem repeats section of the human α -chain. It is expected that this molecule will lose the majority of its extensibility. 4) Hu+2tdr, which will consist of hu1-220, 20 repeats of the 13 amino acid sequence “RSGPGSTGNW,” terminated with amino acids hu392-610 of the human α -chain. In so doing, we will generate a fibrinogen molecule with double, the number of tandem repeats in the normal human molecule. If the extension of fibrin fibers comes from this region, Hu+20tdr fibrin fibers should be much more extensible than their counterparts. 5) Hu-Cys, which will consist of the full human α chain, except Cys442 and Cys472 will be replaced with Ser442 and Ser472. The main bond holding the α C domain structure will be removed; this may prevent the formation of secondary structure in the α C domain, and prevent specific intermolecular interactions between α C regions of neighboring molecules and have an influence on the extensibility of un-crosslinked fibrin fibers.

Creating recombinant fibrinogen with an altered coiled-coil region is tricky, but in principle is do-able as well. Typically, coiled coils form out of a specific sequence of seven amino acids, often called the heptad repeat $(abcdefg)_n$. In this sequence, residues in the *a* and *d* positions tend to be non-polar such as Leucine or Valine, and are positioned in the hydrophobic core between the α -helices. Residues in the *e* and *g* positions are usually charged, and are involved in inter-strand interactions, and amino acids in the *b*, *c*, and *f* positions are hydrophilic and exposed outward towards the solvent.(199-201) Fibrin’s coiled-coil structure is more loose than many other known coiled-coil domains, with numerous “stutters”,(200) nonetheless, great care would need to be taken in order to add or subtract amino acids from the α , β , and γ chains of the

coiled-coil structure and still retain its correct folding pattern. Each chain of the fibrin coiled-coil region (α , β , and γ) consists of bookend Cysteins with 111 amino acids in between the Cysteins. (202) One should be able to add one or more heptad repeats to each chain between the Cysteins and retain a coiled-coil structure while extending its length. A potential risk in this proposal is that adding an additional heptad repeat could twist the molecule in such a way the γ - γ interface or the A:a, B:b knob hole interactions are hindered altering polymerization. Thus, this is a high risk high reward protein.

Finally, our simulations indicate that removing the di-sulfide bond between γ Cys153 and γ Cys182 could result in an additional 87Å of extension in the γ -C region before A:a unbinding. It is not obvious that the γ -nodule would fold properly if these residues were mutated. In fact fibrinogen variant Matsumoto IV, has a γ Cys153 \rightarrow γ Arg153 heterozygous mutation which prevents fibrinogen assembly.(197) In spite of this, it may be possible to substitute a polar residue of similar size to Cysteine such as Serine and still achieve a similar fold. This would be another high risk-high reward fibrin variant, but would provide a model for future A:a knob-hole binding studies.

In these studies I have suggested a series of mutational studies which would provide good models for each phase of the SL α CK model of fibrin assembly. These experiments are risky, but if successful would test several of the hypothesis in our current model of fibrin mechanics.

6.2.2 Cryo-Electron Microscopy of Fibrin Fibers and Fibrin variants

As discussed in the introduction, the structure of fibrin fibers has been studied for nearly 60 years, but it is important to note that all of the data provided so far, be it by

EM or x-ray crystallography has been taken under dry conditions. As shown by Voter and Erickson, the fiber contains 80% water;(43) thus, the majority of the assumptions made about the crystalline structure of fibrin fibers, is based on dried out specimens. Since the formative work by Yang e. al. in 2000, little work on the fiber structure has been pursued.(47) In addition, the discovery of fibrin sheets under tension has called into question the protofibril twisting model for fiber tension.(49) These unresolved questions from the previous model combined with new questions arising from mechanical measurements demand a re-investigation of the fibrin fiber structure.

The structural model of fibrin fibers needs to be re-investigated at this point with an emphasis on understanding the structure of hydrated fibrin fibers, rather than dried out fibrin fibers. In 1994, Baur and Ogendal published a paper on light scattering of fibrin fibers during plasminogen dissolution. In the paper they show one panel of images taken on fibrin fibers using Cryo-Electron Microscopy (Cryo-EM)(9). The images show a banding pattern within the fibers, but no quantitative results were reported and no follow up work seems to have been pursued.

We have begun a full tomographic study of the fibrin fiber structure using Cryo-Electron Microscopy techniques. Cryo-EM is a type of transmission electron microscopy in which a frozen hydrated sample is studied. The advantage of Cryo-EM is that samples can be studied without being stained, fixed or dried. Cryo-EM conditions are obtained by flash-freezing an aqueous specimen in a cryogen like liquid ethane. If done quickly enough, water undergoes a phase transition in which the molecules are immobilized without crystallizing; this state is referred to as vitreous water. In Cryo-EM tomography, one takes a tilt series of low electron-dose images on the same part of the

sample. Using the Central Projection Theorem developed by Crowther in 1970, one can then take the Fourier Transform of each of the two dimensional tilt-series images, and develop a three-dimensional re-construction of the sample embedded in the vitrious water.(203)

Cryo-EM tomography has already solved several structural models of biopolymer-like materials. In one beautiful study by Nicastro et al., Cryo-EM was used to identify the packing of microtubules inside the axoneme, as well as identify Dyenine motors frozen in place between the microtubules.(204) In another, Schmid et al. investigate the actin packing within the acrosome, and show that it deviates significantly from standard, un-packed actin.(205) These studies have lead us to believe that the three dimensional structure of fibrin fibers could also be studied using Cryo-EM techniques.

In so doing, we hope to answer the following questions: 1) How are the protofibrils packed together and where does water fit into the structure? 2) What are the lengths of the protofibrils with fibers? 3) How do larger molecules like FXIIIa fit into to fiber, while still allowing for lateral registry? 4) Where do the α C domains reside within the structure? 5) What is the native persistence length of a protofibril within the fiber 6) Where are the locations of the native contacts between protofibrils? 7) Some fibers do not readily exhibit banding in EM images, are these fibers structurally different? 8) If the fibers are crystalline, how is this broken at branch points? 9) Is there a potential mechanism for tension generation in the arrangement of protofibrils within the fiber.

At this point, in collaboration with the National Center for Macromolecular Imaging (NCMI), we have successfully obtained a reconstruction of one segment of a fibrin fiber.

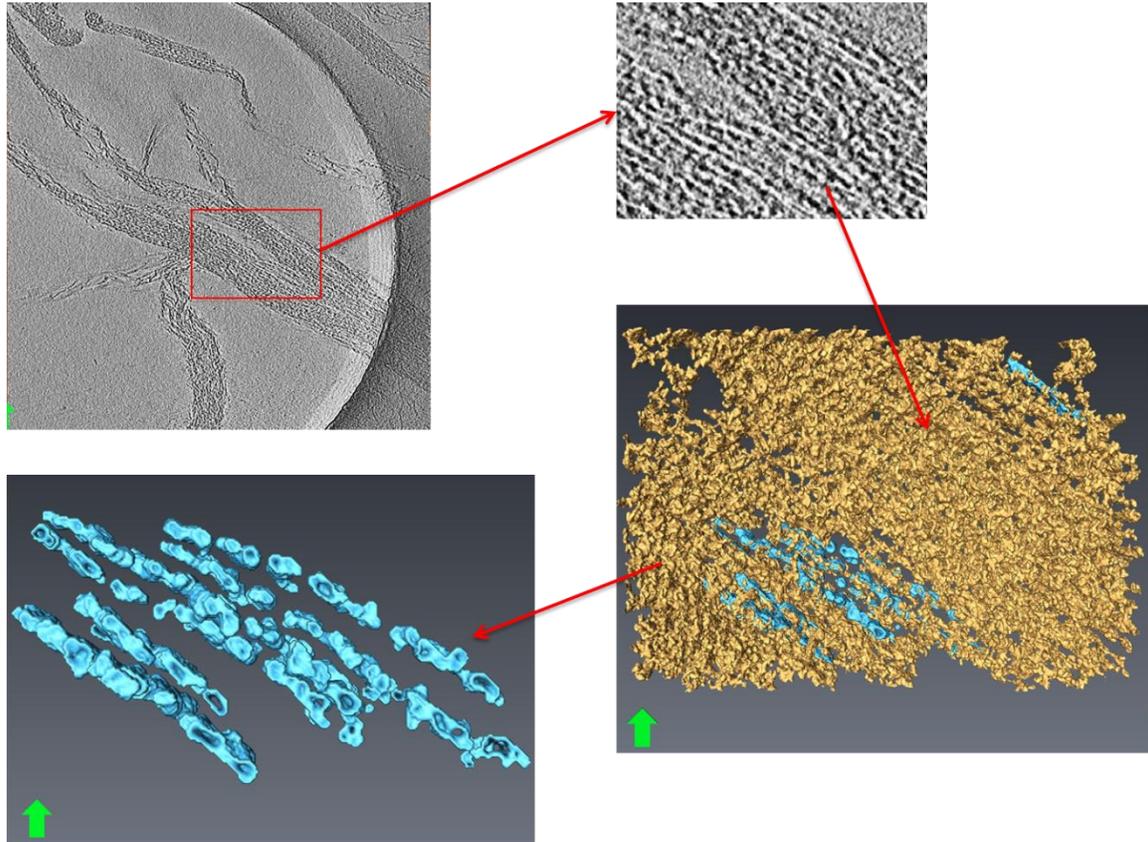


Figure 6.1: Segmenting protofibrils within a Cryo-EM fibrin fiber section: The above images represent the first preliminary results from fibrin fiber segmentation from Cryo-EM tomography. What appear to be protofibrils within the fiber can be identified with this method. Many more reconstructions must be made and images averaged before molecular resolution might be obtained.

While the results are exciting, we have been able to obtain isosurfaces of protofibrils within the fiber, this work must be repeated and refined over several fibers in order to obtain atomic resolution. Additionally, it would be useful to extend this work to look at recombinant fibrin fibers such as those formed by A α 251, or even better, the proposed variant A α 199 which would unambiguously show how protofibrils will pack in the absence of the α C region.

6.2.3 FIONA/FRET Studies of Single Molecules within Fibrin Fibers

A series of studies that were discussed during my time at UNC, but were not successfully pursued, involved using single molecule fluorescence techniques to study the motion of fibrin molecules within fibers during extension. Two of the most successful single molecule fluorescence techniques are Fluorescence Imaging with One-Nanometer Accuracy (FIONA) (206) or FIONA-like techniques and Fluorescence Resonance Energy Transfer (FRET).(207, 208) I will briefly describe the two techniques, some possible applications to the fibrin question, and some ways to overcome drawbacks.

FIONA is a technique in which a 2-D Gaussian profile is fit to the fluorescent point spread function (PSF) of an individual fluorophore in an image:

$$N(x, y) = B + N(x_o, y_o) \exp \left[-\frac{(x - x_o)^2}{2\sigma_x^2} - \frac{(y - y_o)^2}{2\sigma_y^2} \right] \quad (6.2)$$

where (x_o, y_o) is the centroid of the point source, $N(x,y)$ and $N(x_o,y_o)$ are the counts (or photons collected) at pixel (x, y) , and at the centroid position (x_o,y_o) on the CCD (Charge-Coupled Device) detector, σ_x and σ_y are the width of the point spread function in the x and y directions, and B is the baseline fluorescent intensity. The goal is to determine the mean value of the distribution ($N(x_o,y_o)$) and its uncertainty(σ_x and σ_y). FIONA has been used to localize individual fluorophores with 1nm accuracy.(206)

Due to the Raleigh criterion discussed in Appendix B: Optical Microscopy Techniques fluorophores need to be spatially separated by 200nm or more to use FIONA. Since fibrin fibers are 45nm long, true FIONA may not be applicable, however a technique that piggy-backs off FIONA, called NANometer-Localized Multiple Single-molecule (NALMS) fluorescence microscopy, allows spatial resolution of multiple fluorophores within 200nm. (209) With the NALMS technique, one applies a 2-D

Gaussian fit to a fluorescent region in an image. NALMS then uses the fact that all fluorophores eventually photobleach to resolve multiple fluorophores within the same region. When one fluorophore bleaches, the center of the Gaussian fit will adjust to incorporate only the remaining fluorophores. This will continue until all fluorophores have bleached; at this point one can use trigonometry to back out the location of each of the original fluorophores based on the movement of the center of the Gaussian fit after each photobleaching event.

NALMS could be used in fibrin fibers by placing fluorophores in specific locations within the fiber. For example, one could make recombinant fibrinogen molecules and add an additional Cysteine to the end of the β -chain of the molecule. This location is in the D-domain, but not near any Ca^{2+} or known ligand binding sites, so it should not affect polymerization. A fluorophore such as Cy3,(206) could then be specifically attached to the free Cysteine, using a second Cysteine as a linker. As a side note, free Cysteines are likely to lead to intermolecule di-sulfide bond formation, so care must be taken in the labeling and purification to eliminate any unwanted fibrin aggregates originating from the free Cysteines. The initial distance between fluorophores would measure about 35nm, sufficiently distant to localize with NALMS. Fibrin fibers could be created using a mixture of non-labeled fibrinogen molecules and labeled fibrinogen molecules so as to spot label the fiber. As the fiber is stretched, all of the molecules in the fiber would be stretching in some fashion. When one of the Cys-tagged fluorophores bleaches, the distance between the two fluorophores could be measured. One could then get a distribution of distances between fluorophores for all of the molecules making up the fiber at different time points.

The trick with FIONA-like techniques is to have a strong fluorophore signal above the background fluorescence in the sample and to have long-lasting fluorophores. Typically these techniques use TIRF illumination, which only travels 100nm into a sample and exponentially decays over the 100nm. TIRF provides a low background signal, but since fibrin fibers are only ~100nm thick themselves, parts of the fiber will not receive much illumination. Tests would have to be performed using multiple illumination techniques to see if one could generate a high enough signal to noise ratio to use FIONA-type techniques.(207)

FRET is a technique which exploits the fact that a donor fluorophore can transfer energy to an acceptor fluorophore non-radiatively if they are within a certain distance (called the Förster Radius), and their excitation and emissions spectra overlap. (210) The FRET technique can measure distances, by exciting a donor fluorophore and measuring whether or not fluorescence light is emitted in the acceptor fluorophore wavelength; if it is, then FRET is occurring, if it is emitted only in the donor fluorophore's wavelength then the molecules are too far apart and FRET cannot occur. There are many so-called FRET pairs, i.e. two fluorophores that emit at different wavelengths and can exchange energy via FRET. The FRET efficiency depends on the Förster Radius to the sixth power: (207)

$$E = \frac{1}{1 + (r/R_o)^6} \quad (6.2)$$

and thus dies off very rapidly beyond the Förster Radius. The Förster Radius is technically defined as the distance at which the FRET efficiency is 50%.

When FRET is used in single molecule experiments, TIRF illumination is used to maximize the signal to noise ratio; this would pose difficulties to fiber stretching

experiments due to the diameter of fibers. Recently however, FRET has been used to measure protein unfolding in fibronectin fibers using a confocal microscope.(211) This work indicates that a similar experimental setup could be used for fibrin fibers.

Recombinant fibrinogen molecules could be generated with specific FRET pairs at locations in the coiled-coil or α C region, designed such that the pairs can FRET in an un-strained state, but will not FRET in a stretched state. The labeling of large proteins for FRET is non-trivial, but several techniques such as the Split Intein technique proposed by Yang, et al. may be applicable to the fibrinogen case.(212) If these proteins can be generated, then pulling on fibers and measuring the FRET signal for various labeling schemes should give an unambiguous measurement of what part of the fibrin molecule is stretching.

6.2.4 Single Molecule AFM Studies of Fibrin(ogen) Dynamics

Previous studies have measured the unfolding of fibrinogen molecules, fibrinogen oligomers and fibrin protofibrils using the AFM pulling at constant velocity.(68, 75) However, in constant velocity mode, the force applied changes constantly over a wide range of forces, creating a complex time series of unfolding events that can sometimes be difficult to interpret. Mechanical unfolding, is a force dependent process, so a more detailed look at the unfolding pathway of a protein can be obtained by applying a constant force (Force Clamp mode).(170) We have recently begun force clamp measurements on the fibrin molecule, but at the time of writing, the data were not ready for publication.

In the results of the coiled-coil unfolding simulations (Chapter 4.), we showed that there is an α -helix to β -sheet transition undergone by the coiled coil and this β -sheet

transition prevents an immediate relaxation back into the coiled-coil conformation. These results tested experimentally by measuring the collapse dynamics of the coiled-coil after forced extension. This technique was recently used by Berkovich et. al to study the refolding of ubiquitin and I27 proteins. (213) In these studies the proteins were unfolded step-wise in an AFM force clamp mode, and then the force was reduced or quenched altogether, and the proteins were allowed to refold. The authors measured two distinct re-folding lengths, and characteristic times over which those lengths were reached.

For fibrin monomers, our simulations would indicate that re-folding does not immediately happen after the force quench, but rather the protein will stay in an elongated β -sheet conformation. This could be probed further by making measurements at various PH's or temperatures. Both the coiled-coil and β -sheet conformation should be less stable at PH's to either extreme of 7.4 or temperatures above 40 degrees.

Another property of interest is the intrinsic viscosity of the individual fibrin molecule. It has been shown that the whole fibrin fiber is viscoelastic,(67) meaning that it exhibits both elastic and viscous components when stretched. However, it has not been determined what gives rise to the viscoelastic behavior of the fibrin fiber. Recently, an AFM based technique has been developed to measure the viscoelastic behavior of proteins by stretching the proteins while undergoing driven oscillations.(214) This approach has been applied to unstructured polypeptide chains as well as the myosin coiled-coil. (215, 216) They've shown that the internal friction of the protein can be up to seven times greater than solvent-protein friction, but at the same time the internal friction of a folded protein is significantly higher than an unfolded poly-peptide. (214, 216, 217) This experiments could be repeated for the fibrin coiled coil-structure and the

α C region of the fibrin molecule. One would expect significantly different internal viscosities between the two regions, and this in turn could lead to a deeper understanding of the viscoelastic behavior of the entire fibrin fiber.

6.2.5 Developing a Coarse-Grained Model for the Fibrin Fiber

Discrete Molecular Dynamics has provided a powerful tool to probe the molecular origins of fibrin(ogen)'s mechanical properties. In so doing, we have been able to measure the relative stability of most parts of the fibrin(ogen) molecule providing a step by step description of what parts of the molecule should stretch under force. DMD falls under the methods category of a coarse-grained model as it uses simplified interaction potentials to simulate the reaction evolution of the system. In spite of these simplifications, the fibrin molecule is so large that simulating the entire fibrinogen molecule containing 28,400 atoms and 2923 amino acids for ~50ns of real time dynamics, will take approximately 1 year at our current computer speeds (estimate based on current speed of an ongoing simulation).

The next steps in generating a computational model of the fibrin fiber will involve additional coarse-graining of the fibrin molecule. The goal should be moving towards a hierarchical simulation of the entire fiber. The first level of coarse graining could involve simulating each amino acid as a virtual bead that interacts with a specific potential corresponding to its particular amino acid type. Simplified amino acid representations have been employed in DMD simulations in the past, and understandings of the interaction potentials for simplified amino acids have become more sophisticated in recent years.(218, 219) Recently, the coarse grained MARTINI force field was used to simulate the extension of vimentin trimer with results consistent with experiments.(114,

220) The same approach could be used in DMD to simulate a fibrin quatermer (~12,000 virtual amino acids). This would be a good building block for understanding the mechanics and inherent flexibility of a fibrin protofibril. (See Figure 6.2)

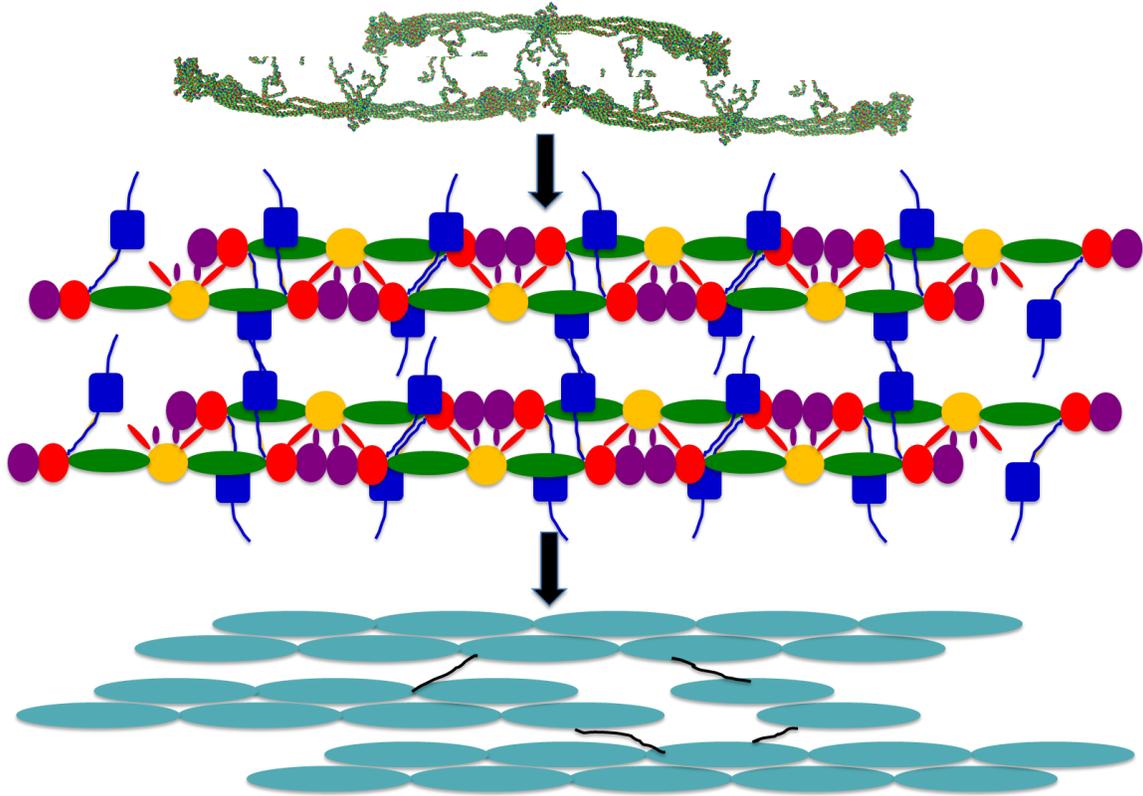


Figure 6.2 Successive Levels of Coarse Graining: An all molecule model of the fibrin fiber can be built starting with the simulation results from this work. The next step would be to do a protofibril level simulation with each amino acid being represented as a parameterized bead. The next level could parameterize the whole molecule as seven or eight different beads with different mechanical properties corresponding to results from the all atom simulations. Finally, a fiber level model could be generated by creating a virtual bead that represents the entire molecule and adding interaction terms that could be parameterized from the lower level simulations.

From the protofibril level simulations, the next step would involve simulating a series of protofibrils connected via α C domain chains. Additional coarse-graining must be employed to make these simulations computationally reachable. In this work, we have characterized the mechanical properties of most of the individual fibrin molecule, indicating that we could virtualize the molecule as a set of beads corresponding to the

mechanical properties of each of the individual part of the molecule. For example, a coiled coil could be represented as a two state (folded and unfolded) virtual bead with an energy barrier of $1.8nN \cdot nm$ separating the two states. While not mechanical models, similar protein coarse graining has been applied to membrane proteins in simulations of vesicle fusion.(221) This approach could be extended even further to simulations of whole fibers composed of virtual beads representing fibrin molecules linked into protofibrils within the fiber. Assuming a $10\mu m$ fiber with a diameter of $100nm$, approximately 16,000 individual fibrin molecules are packed into one fiber. At each level of simulation, the mechanical properties of that level could be carefully parameterized to represent the properties of a virtual bead at a higher level. By simulating a full fibrin fiber, a complete understanding the effects of the hierarchical topology of the fibrin fiber on the fibers mechanical properties can be achieved.

6.2.6 Measuring the effects of other blood coagulation factors on Fibrin's mechanical properties

Fibrin network formation is the final step of the blood coagulation cascade which itself exists within the larger framework of Hemostasis. Thus far, our mechanical measurements have been performed using only fibrin, thrombin and FXIII as variables. Future work should expand the single fiber measurements to include additional blood plasma components. We have recently begun investigating the role fiber diameter in fibrin lysis times. As an unexpected corollary, we've observed in repeated experiments, that as a fiber is lysed, it loses tension before fully breaking (See Figure 6.3). (experiments ongoing by I. Bucay). Interestingly, in these experiments we've also observed the lysing and loss of tension in fibrin sheets. Continuing these experiments

with recombinant fibrinogens such as A α 199 (proposed) or A α 251 may shed light on the mechanism of tension within fibrin fibers by eliminating the α C-region from consideration.

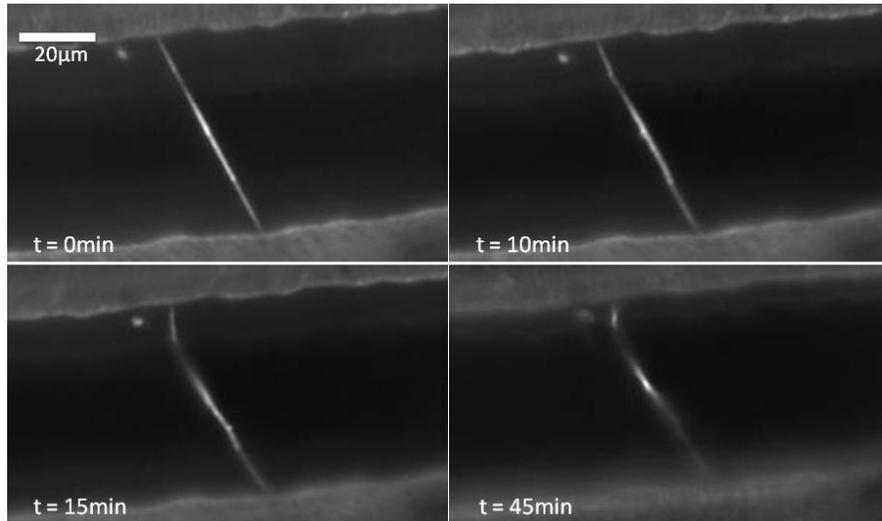


Figure 6.3: Fiber Tension Loss: The above time series of images was taken after addition of 20 μ M plasminogen and 100pM tPa. During the time series the fiber slowly loses tension as it is being lysed.

6.2.7 Testing for and α -helix to β -sheet Transition under stretching

The MD simulations described in Chapter 3. Fibrin Fiber Recoil Dynamics predict an α -helix to β -sheet transition in fibrin during stretching. This prediction should be tested experimentally. The model tests for this have been performed on keratin fibers.(120, 128) Kreplak et. al recently used Wide Angle X-ray Scattering (WAXS) to probe the transition in keratin fibers. To our knowledge, no one has attempted this with fibrin fibers, but based on Small Angle X-ray Scattering (SAXS) results from fibrin gels, WAXS should be applicable as well.

Another technique would be to use a dye such as congo red which creates a green birefringence when interacting with amyloid β -sheets.(119, 222) This was successfully used by Fudge et al. to measure the α -helix to β -sheet transition in hagfish slime

threads.(119) Measurements could be performed in a glass shear cell which would allow network visualization throughout the extension.(82, 178) If a coiled-coil to β -sheet transition does occur, then an increase in birefringence should be measurable. Since fibrin has native β -sheet structure, fibrin fibers will exhibit some inherent birefringence unlike keratin, but a difference in birefringence should be measurable if a large transition occurred.

Appendix A: An Overview of Molecular Dynamics Methods

A.1 Molecular Dynamics: A Primer

This entire appendix is meant to be used as a brief primer on MD methods. It is in no way meant to be complete, but it is presented here in the hopes that those following after me may have some basis for continuing my work. The first section discusses general MD methods, and the assumptions behind MD. Later, I describe in a bit more detail, the method and equations governing Discrete Molecular Dynamics (DMD) which was used to generate the results in Chapter 4. For additional resources on molecular dynamics the reader is directed to one of many good books on the subject.(223)

A.1.1 An Brief History to Molecular Dynamics

Molecular Dynamics (MD) is a form of computer simulation in which simplified models of atoms are allowed to interact for a period of time via various approximations of chemical and/or quantum mechanical forces, giving a view of the motion and evolution of the system. While initially applied to liquids, (224-226) MD simulations are now often applied to proteins and other biomolecules in an attempt to understand their motions in physiological settings.

Levitt and Warshel initiated molecular dynamics simulations on proteins in 1975 using Bovine Pancreatic Trypsin Inhibitor (BPTI) as a model system; they showed that unfolded chains of simplified amino acids would roughly collapse into the native BPTI conformation. (227) This work was followed up by Cooper (228) and McCammon et al. who simulated the dynamics of the atoms in BPTI, about the average structure, showing that proteins thermally fluctuate about an equilibrium structure and are not rigid. (229)

Since these seminal works, molecular dynamics simulations, have demonstrated many remarkable protein processes including protein folding pathways,(230-233) ligand binding,(234) normal mode oscillations,(235) forced protein unfolding, (68, 172, 236) and ion channel regulation. (237) In addition, MD simulations are currently being used for medical applications such as rational enzyme and drug design. (238)

A.1.2 The Basic Molecular Dynamics Mechanism

At its most basic level, the molecular dynamics method is simply an iterative application of Newtonian mechanics (or other types of mechanics such as Langevin (see below)) to a system of interacting particles. The force on each particle (i) in the system is given by:

$$F_i = -\nabla_i V = m_i \vec{a}_i = m_i \frac{d\vec{v}_i}{dt} = m_i \frac{d^2 \vec{x}_i}{dt^2} \quad (\text{A.1})$$

Where V is the potential energy of the system and ∇_i is the gradient of the potential energy applied to the i^{th} particle, m_i is the mass of the i^{th} particle, a_i is the acceleration of the i^{th} particle, v_i is the velocity of the i^{th} particle, x_i is the position of the i^{th} particle, t is time, and d/dt and d^2/dt^2 are the first and second derivatives with respect to time. The future position at time (τ) of every particle (i) in the system can then be calculated based on the starting positions (x_o) and velocities (v_o) of all the particles in the system using:

$$\vec{x}_i(\tau) = \vec{a}_i \tau^2 + \vec{v}_{oi} \tau + \vec{x}_{oi} \quad (\text{A.2})$$

Typically for MD simulations, the starting positions are based upon a known crystal structure, and the initial velocities are assigned randomly from a Maxwell-Boltzmann distribution at a given temperature.(223)

The potential energy of the system is sum of all the particle-particle interactions as well as solvent-particle interactions between the N particles. Because of this, it is

impossible to solve the Newtonian equations of motion analytically. Instead they must be solved iteratively, using numerical methods. Typically this is done by picking a time step, and then using an integration algorithm to solve for the position, velocity, and acceleration of each particle after that time step. The potential energy landscape must also be iteratively updated after each time step. There are many numerical integration algorithms each with benefits and disadvantages. An example of one is the Velocity Verlet method:(239, 240)

$$\begin{aligned}\vec{x}(t + d\tau) &= \vec{x}(t) + \vec{v}(t)d\tau + 0.5\vec{a}(t)d\tau^2 \\ \vec{v}(t + d\tau) &= \vec{v}(t) + 0.5[\vec{a}(t) + \vec{a}(t + d\tau)]d\tau\end{aligned}\quad (\text{A.3})$$

All integration methods lose accuracy over the course of many time-steps as they use cut-off Taylor series expansions to approximate the positions, velocities, and accelerations of all the particles in the simulation.(241)

MD simulations evolve for S steps of $d\tau$, to reach the chosen simulation time τ . Choosing the correct $d\tau$ is important, as longer time steps can speed up simulation time, but miss important phenomena that occur on faster timescales. Usually $d\tau$ is chosen to be smaller than the fastest timescales of interest. For protein simulations, $d\tau$ often chosen to be 1fs (10^{-15} s).(236, 242)

A.1.3 Potential Energy or Force Field calculations in MD

The most important part of the molecular dynamics method is calculating the potential energy (V) of the system, from which the acceleration can be derived and Newton's equations follow (Eq.A.1). The total potential energy of the system of N atoms is the sum of the potential energies of each individual atom.

$$V_{tot} = \sum_{i=1}^N V_i \quad (\text{A.4})$$

The potential energy of each individual atom is given by the sum of all of the different interactions that the atom will encounter.

$$V_i = V_{i,str} + V_{i,bend} + V_{i,tors} + V_{i,VDW} + V_{i,elec} + V_{solv}\dots \quad (\text{A.5})$$

Here, a brief description of the basic potential energy terms will be given as described by Höltje in one reference on MD techniques(243), and later a detailed description of the potential energy terms for a specific type of MD simulation, Discrete Molecular Dynamics will be provided.

$V_{i,str}$ is the bond stretching potential energy term. It describes the resistive force to an atom moving out of its lowest energy state or unstretched (resting) position. It is typically described as a harmonic potential with $V_{i,str} = 0.5k_{str}(\mathbf{r}-\mathbf{r}_0)^2$ where k_{str} is the stretching spring constant, \mathbf{r}_0 is the rest length, and \mathbf{r} is the atomic displacement from the rest length.

$V_{i,bend}$ is the angular potential energy term. It describes the resistive force to covalently bound atoms bending with respect to one another. It is typically described as a harmonic potential with $V_{i,bend} = 0.5k_{\theta}(\theta-\theta_0)^2$ where k_{θ} is the bending spring constant, θ_0 is the angular rest length, and θ is the angular displacement from the rest length.

$V_{i,tors}$ is the dihedral potential energy term. Dihedral angles are typically in found in conformations of 180° or 0° for $C_{\alpha}-C_{\alpha}$ interactions, or 180° , -60° or 60° for side chain dihedrals. The potential energy terms is often expressed as a cosine function: $V_{i,tors} = 0.5k_{\phi}[1+\cos(n\phi-\phi_0)]$ where $5k_{\phi}$ is the dihedral spring constant, n is the number of energy minima, ϕ is the torsion angle, and ϕ_0 is the rest or reference angle.

$V_{i,VDW}$ is the Van der Waals potential energy term and is typically described using the Leonard Jones potential. a square well version of the Lennard-Jones (LJ) potential[(244):

$$V_{i,VDW} = \sum_{i,j} 4\varepsilon_{ij} \left(\frac{b_{ij}}{\vec{r}_i - \vec{r}_j} \right)^{12} - \left(\frac{b_{ij}}{\vec{r}_i - \vec{r}_j} \right)^6 \quad (\text{A.6})$$

where $b_{ij} = b_i + b_j$ and is the Van der Waal radius between the i^{th} and j^{th} particle and ε_{ij} is the interaction strength, and \mathbf{r}_i and \mathbf{r}_j are the positions of the i^{th} and j^{th} atom.

$V_{i,elec}$ is the electrostatic potential term, governed by Coulombs law:

$$V_{i,elec} = \frac{1}{\varepsilon} \left(\frac{Q_i Q_j}{\vec{r}_i - \vec{r}_j} \right) \quad (\text{A.7})$$

where Q_i and Q_j are the charges of the i^{th} and j^{th} atom, and ε is the dielectric constant.

Finally, the V_{solv} term describes the potential energy for interactions with the solvent molecules. Solvent can either be included explicitly (including a representation of individual water molecules), or implicitly where the solvent is represented as a continuous medium. Including explicit water molecules requires much more computational time as all of the above potential terms must be calculated for each atom in the water molecules. In addition, simulations are limited to a small box of solvent molecules, and this can lead to boundary condition effects. Implicit solvation is less expensive computationally, do not suffer boundary condition effects, and usually give better sampling of conformational space. The downside is that they do not include viscosity, and come with the price of making a dramatic simplification to discrete solvent structures.(244) Different implicit solvation representations have been proposed, each

with trading off between computational time and exactness. Below I will briefly describe some of the prominently used implicit solvation models.

One of the first implicit solvation models was the accessible surface area (ASA) model, proposed by Eisenberg and McLachlan in 1986. (245) In this model, the free energy of solvation is calculated for a protein based on the accessible surface area of each of the atoms in the protein.

$$\Delta G_{solv} = \sum_i (\Delta G)_i ASA_i \quad (\text{A.8})$$

where $(\Delta G)_i$ is the free energy of solvation (atomic solvation parameter) of the i^{th} atom and ASA_i is the accessible surface area of the i^{th} atom.

A similar way to calculate the solvation potential was derived by Lazaridis and Karplus (LK or EEF1) in 1999.(246) In it, the free energy is calculated by using the reference free energy of a fully exposed small molecule, and calculating the free energy difference of all other atoms around the fully exposed molecule, assuming the solvation free energy density is given by a Gaussian function.

$$V_{solv} = \sum_{i,j} - \frac{2(\Delta G^{ref})_i V_i \exp[-((\vec{r}_i - \vec{r}_j) - 1.12a_o) / \lambda_i]^2}{4\pi^{1.5} \lambda_i (\vec{r}_i - \vec{r}_j)^2} - \frac{2(\Delta G^{ref})_j V_j \exp[-((\vec{r}_i - \vec{r}_j) - 1.12a_o) / \lambda_j]^2}{4\pi^{1.5} \lambda_j (\vec{r}_i - \vec{r}_j)^2} \quad (\text{A.9})$$

where the parameters are defined as: volume of atoms (V), correlation length (λ_i) and atomic radius (σ). In their paper Lazaridis and Karplus estimated values for each atom type.(246) Using the LK solvation model is only 50% slower than a vacuum simulation.

The other main type of solvation model comes from the Poisson-Boltzmann (PB) equation.(244) It is an electrostatics model that calculates the interactions of each of the atoms with a continuous solvent with a position dependent di-electric constant.

$$\vec{\nabla} \cdot [\epsilon(\vec{r}) \vec{\nabla} V_{solv}(\vec{r})] = -4\pi\rho(\vec{r}) - 4\pi \sum_i n_i^o Q_i \exp\left[\frac{-Q_i V_{solv}(\vec{r})}{k_B T}\right] \quad (\text{A.10})$$

Where n_i^o is the number density of the counterions of type i in the protein, Q_i is the charge of the counterions of type i , and $k_B T$ are Boltzmann's constant and temperature respectively. Solving the exact PB equation can be very computationally expensive for large proteins, thus in practice approximations to PB are typically used.

An approximation to the Poisson-Boltzman model is the Generalized Born (GB) approximation. (247) In the GB model, the protein is represented as a hard sphere with an internal dielectric constant.

$$V_{solv} = -\frac{1}{8\pi} \left(\frac{1}{\epsilon_o} - \frac{1}{\epsilon} \right) \sum_i \frac{Q_i^2}{b_i} - \frac{1}{8\pi} \left(\frac{1}{\epsilon_o} - \frac{1}{\epsilon} \right) \sum_{i,j} \frac{Q_i Q_j}{\vec{r}_i - \vec{r}_j} \quad (\text{A.11})$$

where ϵ is the dielectric constant of the solvent Q_i and Q_j are the charges of atoms i and j , and b_i is the effective Born radius of the i^{th} atom. (247) The Born radius is the effective radius of the atom in a given solute. It characterizes the amount of burial of the atom within the molecule, and must be known precisely to have a good approximation of V_{solv} .(248-250)

A.2 The foundation of the MD method: statistical mechanics

A.2.1 A review of statistical ensembles

Ideally during a simulation, one would like to measure macroscopic properties about the system as a whole. In order to do this one must apply methods learned in statistical mechanics. The molecular dynamics method is based in statistical mechanics, the branch of physics that applies probability theory to predict the macroscopic thermodynamics of systems based on the microscopic properties their constituent particles.(251, 252) In this section, a brief overview of statistical mechanics will be

presented, followed by a discussion on how this can be applied to molecular dynamics methods. **The basic idea behind statistical mechanics is to be able to extract macroscopic properties of a thermodynamic system, such as energy and entropy, from ensemble averages of the behaviors of all of the particles comprising the system.** The first concept in statistical mechanics is that of the microstate. A microstate (Ω) is a specific microscopic configuration of a thermodynamic system, such as one snapshot orientation of all of the atoms comprising a protein. All of the microstates taken as a whole consist of a thermodynamic ensemble of possible states; and within the ensemble, each microstate will be visited during the course of the systems' fluctuations with a certain probability. Using statistical mechanics, one can calculate macroscopic properties of the system based on the average fluctuations of the particles through each microstate.(252)

There are three types of ensembles typically described in statistical mechanics, and each ensemble carries certain assumptions and governing which thermodynamics properties can be calculated.(252) The microcanonical ensemble describes a completely isolated system having constant energy, number of particles, and volume. The macroscopic function that best describes this ensemble is entropy, as the energy is always constant. The equilibrium state of the system will be the one in which the entropy is maximized and the entropy (S) is defined as: $S=k_B \ln(\Omega_n(E))$. Where $\Omega(E)$ corresponds to the total number of microstates with the given ensemble energy.

The second type of ensemble is the canonical ensemble. In the canonical ensemble, the number of particles, the volume, and the temperature are fixed, meaning the system is in thermal equilibrium with its environment, but has local energy

fluctuations. Within the canonical ensemble, the probability $P(\Omega)$ of the system visiting any particular microstate (Ω) is governed by the Boltzman distribution: $e^{-E(\Omega)\beta}$

$$P(\Omega) = \frac{e^{-E(\Omega)\beta}}{\sum_{\Omega} e^{-E(\Omega)\beta}} \quad (\text{A.12})$$

$$\sum_{\Omega} P(\Omega) = 1 \quad (\text{A.13})$$

Where the denominator is a normalizing constant called the partition function (Z) defined such that the sum over all the probabilities for the system to be in each possible microstate is 1 and β is defined as $1/k_B T$ where k_B is the Boltzmann constant. Here the probabilities and partition functions are defined in discrete terms due to the correspondence in molecular dynamics simulations, but they can also be defined in continuous terms.

The partition function turns out to be the most important quantity in statistical mechanics. It contains information on how the probabilities are “partitioned” among the various energy microstates of the ensemble, and from the partition function most of the average macroscopic properties of the ensemble can be calculated. Any average property of interest ($X(\Omega)$) can be calculated via:

$$\langle X(\Omega) \rangle = \sum_{\Omega} X(\Omega) e^{-\beta E(\Omega)} \quad (\text{A.14})$$

Ones of interest for molecular dynamics simulations are the following(252):

$$H = \langle E \rangle - TS = -k_B T \ln(Z) \quad (\text{A.15})$$

$$\langle E \rangle = -k_B T^2 \frac{\partial \ln(Z)}{\partial T} \quad (\text{A.16})$$

$$\langle (\Delta E)^2 \rangle = \langle E^2 \rangle - \langle E \rangle^2 = \frac{\partial^2 \ln Z}{\partial \beta^2} \quad (\text{A.17})$$

$$C_v = \frac{\partial \langle E \rangle}{\partial T} = \frac{\langle (\Delta E)^2 \rangle}{k_B T^2} = \frac{1}{k_B T^2} \frac{\partial^2 \ln Z}{\partial \beta^2} \quad (\text{A.18})$$

Here H is the Helmholtz free energy, E is the internal energy, and C_v is the specific heat.

The third type of ensemble is not as useful for molecular dynamics, but it is called the grand canonical ensemble. In the grand canonical ensemble, the volume, temperature, and chemical potential are held fixed, where chemical potential is the change in free energy of the system with respect to particle species. Based on this definition, it is easy to see why protein molecular dynamics simulations often do not use the grand canonical ensemble, because typically the number of atoms in a protein is fixed.

One other property that can be derived from statistical mechanics is the system temperature. The equipartition theorem tells us that(251):

$$\langle E \rangle_{thermal} = \frac{f}{2} N k_B T \quad (\text{A.19})$$

Where N is the number of particles, and f is the degrees of freedom. This turns out to be particularly useful for molecular dynamics simulations where N is known.

A.2.2 The ergodic hypothesis

With this basic background in statistical mechanics, it is now time to see how statistical mechanics can be applied to molecular dynamics simulations, specifically molecular dynamics applied to a protein model. For more details, the reader is referred to an excellent online MD tutorial produced by users of CHARMM which was followed in these derivations.(241) Proteins consist of linear chains of amino acids in a specific

sequence that fold into a three-dimensional structure. In MD, a protein is reduced to the atoms that constitute the amino acids of the protein, with interaction potentials assigned between the atoms based on known electric and quantum mechanical interactions (a more in-depth look at these potentials will come later). The system is assigned an initial state of positions and momenta for each atom, and then allowed to evolve over time and sample various microstates of the system. Each microstate can be defined by the positions ($\mathbf{x}^N = \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$) and momenta ($\mathbf{p}^N = \mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N$) of the N atoms in the protein in a particular configuration. The ensemble average of a property of interest X (such as energy) is then defined as (using A.14):

$$\langle X \rangle_{ensemble} = \sum_{x^N=1}^{\infty} \sum_{p^N=1}^{\infty} X(x^N, p^N) P(x^N, p^N) \quad (\text{A.20})$$

Where the double sum (it's actually 3N sums, one for each dimension for N particles) is carried out over all atomic positions and momenta, $X(x^N, p^N)$ is the X microstate where all of the particles are in corresponding positions x^N and momenta p^N , and $P(x^N, p^N)$ is the probability to be in the microstate defined by (x^N, p^N) (equation A.12). This sum is nearly impossible to calculate for an MD simulation, since the simulation would have to pass through all possible states of the system, and would take an enormous amount of computational time. This leads us to the basic assumption of molecular dynamics: the ergodic hypothesis.(253) The ergodic hypothesis states that that over long time scales, all micro-states within the same energetic phase space will be equally sampled by the system.(254, 255) Thus, the statistical ensemble averages are equal to time averages of the molecular system or:

$$\langle X \rangle_{ensemble} = \langle X \rangle_{time} \quad (\text{A.21})$$

The time average of a property of interest can be defined as:

$$\langle X \rangle_{time} = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_{t=0}^{\tau} X(x^N(t), p^N(t)) dt \quad (\text{A.22})$$

However, simulations cannot be carried out to infinite time either, so in practice the time average for an MD simulation is calculated by:

$$\langle X \rangle_{time} = \frac{1}{S} \sum_{t=1}^S X(x^N, p^N) \quad (\text{A.23})$$

Where S is the number of time steps in the simulation. In order for molecular dynamics simulations to make experimentally relevant predictions, S must be large enough to approximately satisfy the ergodic hypothesis.

A.2.3 Statistical ensembles in MD

Before starting an MD simulation, one must first decide what properties about the system are of interest, this in turn will determine which type of statistical mechanics ensemble can be applied to the simulation system. One parameter of thermodynamic interest in an MD simulation is the temperature. Using the Equipartition theorem (Eqn. A.19), temperature in an MD simulation can be defined as:

$$\frac{1}{2} m \langle v^2 \rangle = \frac{3}{2} N k_B T \Rightarrow T = \frac{1}{3 N k_B} \sum_{i=1}^N m_i v_i^2 \quad (\text{A.24})$$

The equations of motion for molecular dynamics are typically formulated for a microcanonical ensemble (NVE) where N, V, and E are held constant while temperature and pressure are allowed to fluctuate.(223) This is due to energy conservation assumed in Newton's equations. Simulations run under these conditions are useful for investigating a constant-energy surface within the ensemble, but one must know a-priori what energy surface to investigate, as energy will be an input to the simulation and will fluctuate around an assigned value. The energy will drift due to the integration algorithms used, but in general will fluctuate around a constant value.

Often times however, in real systems, energy conservation is not nearly as important as keeping things at a constant temperature or pressure. Thus, different parameters must be added to Newton's equations to adjust the ensemble accordingly. In a canonical ensemble (NVT) MD simulation, temperature is held constant while the energy and pressure are allowed to fluctuate.(223)

A.3 Generating a Canonical Ensemble

For simulations requiring constant temperature or constant pressure, a “thermostat” or “barostat” must be added to the simulation to add or remove energy to keep these conditions constant. This write-up will briefly discuss some common thermostats used in MD simulations. In particular, the reader's attention is drawn to the Anderson Thermostat which is used in DMD simulations.

A.3.1 Woodcock Thermostat

The first attempt at isothermal molecular dynamics was done by a velocity re-scaling method proposed by Woodcock in 1971.(256) After every time-step he proposed that the velocities be re-scaled by a factor $f_{\alpha}(t)$ given by:

$$[f_{\alpha}(t)]^2 = \frac{\langle \dot{r}_i^2(t=0) \rangle_{\alpha}}{N_{\alpha}^{-1} \sum_{i=1}^{N_{\alpha}} \dot{r}_i^2(t)} \quad (\text{A.25})$$

Where N_{α} is the number of particles and \dot{r}_i is the velocity (time derivative of the position r) of each particle. The issue with this approach is that it does not allow fluctuations in temperature.

A.3.2 Anderson Thermostat

The concept was extended by Andersen in 1980 who proposed to include a stochastic element in the process. (126) At each step, a stochastically chosen set of particles would undergo an instantaneous “collision”, adjusting the momentum of those particles. This process is meant to replicate collisions with an external heat bath. No collision is actually calculated, but rather instead, after every time step a fraction of the particles exchange their velocities with a new set of velocities randomly chosen from a Gaussian distribution of velocities at the prescribed temperature. The strength of the thermostat can be adjusted by changing the mean rate at which each particle suffers a stochastic collision (ν). The probability of a collision is governed by a poison distribution:

$$P(t) = \nu \exp(-\nu t) \quad (\text{A.26})$$

Where $P(t)\tau$ is the probability that a collision will happen between t and τ . In this way, a canonical ensemble can be achieved. The downside is that the particle trajectories are not continuous. This process is generally referred to as the Andersen thermostat.

A.3.3 Berendsen Thermostat

A third mechanism of velocity re-scaling was proposed by Berendsen et al. in 1984. (257) The system is supposed to be weakly coupled to a heat bath at some constant temperature. At each step the velocities are scaled such that the rate of change of the temperature is proportional to the difference in temperatures between the system and the bath:

$$\frac{dT(t)}{dt} = \frac{1}{\tau}(T_o - T(t)) \quad (\text{A.27})$$

where τ is the coupling constant. Thus after each time step, the velocities can be rescaled by a factor of λ , $v \rightarrow \lambda v$ where λ is defined as:

$$\lambda = 1 + \frac{\Delta t}{\tau} \left(\frac{T_o}{T} - 1 \right) \quad (\text{A.28})$$

This is equivalent to adding a friction coefficient to the velocities of order $(1/\tau)(T_o/T-1)$. The advantages of the BT are that it is easy to implement, and can be adjusted to rapidly bring a system to a desired temperature using a strong coupling constant τ . The biggest disadvantage is that it does not correctly re-produce a canonical ensemble, so time-averaged properties cannot be calculated. Thus, it is often used to bring a system to a desired temperature, after which point a canonical thermostat is used. (258)

A.3.4 Nose-Hoover Thermostat

Because of the discontinuity of the Andersen thermostat, Nose (1984) and Hoover (1985) proposed a thermostat based on coupling the system to an external heat bath.(259, 260) The Nose-Hoover thermostat is typically derived in terms of Hamilton's equations of motion. (260) The coupling is typically achieved by adding an additional time scale degree of freedom (s), and a conjugate mass (Q) and momentum (p_s) to the Hamiltonian (Total System Energy). They also introduced virtual variables ($\mathbf{r}, \mathbf{p}, t$) which are related to the real variables ($\mathbf{r}', \mathbf{p}', t'$) by:

$$\begin{aligned} \vec{r}'_i &= \vec{r}_i \\ \vec{p}'_i &= \vec{p}_i / s \\ dt' &= dt / s \quad (\text{A.29}) \\ \frac{d\vec{r}'_i}{dt'} &= s \frac{d\vec{r}_i}{dt} \end{aligned}$$

The augmented Hamiltonian is given by:

$$H(\vec{p}, \vec{r}, p_s, s) = \sum_{i=1}^N \frac{\vec{p}_i^2}{2ms^2} + \frac{1}{2} \sum_{i \neq j} V(\vec{r}_i - \vec{r}_j) + \frac{p_s^2}{2Q} + gk_B T \ln(s) \quad (\text{A.30})$$

where g is the number of degrees of freedom. Thus, while the momenta are scaled by s , the particle coordinates are not. The resulting equations of motion are:

$$\begin{aligned} \ddot{r}_i &= -\frac{F}{m_i s^2} - \frac{p_s}{Q} \frac{2\dot{r}}{s} \\ \ddot{s} &= \frac{1}{Qs} \left(\sum_i m_i s^2 \dot{r}_i - gk_B T \right) \end{aligned} \quad (\text{A.31})$$

Thus the quantity p_s/Q acts as a friction coefficient for particle motion. It can also be seen that s is not constant, but rather fluctuates with particle velocity. Thus, the time scaling is not constant, so the trajectory is not sampled at even time intervals. The trick with the NHT is to use the appropriate friction constant Q . If this is not chosen to overlap with natural system frequencies, the NHT can act very slowly, or the temperatures can oscillate very rapidly, in either case the system energy will drift due to the accumulation of numerical errors. (258) The advantage of the NHT, is that Nose was able to show that the microcanonical distribution in the virtual variables is equivalent to a canonical distribution in real variables. The equations of motion can be re-formulated in of real system variables, and are time reversible, unlike the Andersen Thermostat.

A.3.5 Langevin dynamics

A third type of thermostat using a friction coefficient and a stochastic collision term deriving from the equations of Langevin Dynamics.(261) The basic idea of Langevin dynamics is that a particle is moving through a continuum of small “fictional” particles generating a drag force γv . The particles are also exhibit thermal fluctuations. The equations of motion of such particles would be given by:

$$m\ddot{r}_i = -\gamma\dot{r}_i + \eta_i(t) \quad (\text{A.32})$$

Where $\eta(t)$ is a “noise” force which produces a random kick in the particle position similar to a stochastic random walk. The function $\eta(t)$ is given by a Gaussian distribution at the desired temperature:

$$P(\eta(t)) = \frac{1}{\sqrt{8\pi(\gamma mk_B T)^2}} e^{\left(\frac{-\eta(t)^2}{8\pi(\gamma mk_B T)^2}\right)} \quad (\text{A.33})$$

Thus, with each step the particle position and velocity are scaled by both the friction and the random noise term. Given that the temperature is related to the particle velocities by Eq. A.24, the temperature of the system can be controlled by adjusting γ and $\eta(t)$. Langevin dynamics can reproduce a canonical ensemble. (258)

A.4 Molecular Dynamics of “Large” Systems

The calculation of the potential energy of the system at each step is typically the most computationally expensive task of the molecular dynamics method. Using the Verlet integration method and a stand potential energy function, the computational time A.24(real time on a computer) required to simulate a system scales as N^2 where N is the number of particles.(262) On the other hand, the simulation time required (the physical time need to be captured in the simulation) depends on the phenomena to be explored. Chemical reactions happened on the order of femto (10^{-15}) seconds, amino acid side chain motions happen on the order of nano (10^{-9}) seconds, while protein folding can occur on timescales of (10^{-6} to 10^1 seconds). (263) So, as N increases, simulating longer timescale phenomena, such as protein folding, becomes impractical in computational time.

In order to get around this problem, coarse-grained models have been developed to either simplify the potential energy terms or reduce the number of particles in the simulation and speed up the computational time. One common way for this to be done is

the united atom approximation, whereby several atoms or even several amino acids are grouped together and represented by an effective particle (bead).(264) The interaction potential energy between these effective particles is empirical and must be generated from protein structures, experimental data, all atom molecular dynamics simulations, or physics.(265-268) In spite of their limitations, effective bead approximations have given relatively good structural predictions of the folding of large proteins.(269)

Another useful technique for large molecules is the so-called Gō model named after Nobuhiro Gō.(270, 271) Gō models are typically applied to questions of protein folding. The basic assumption is that the only the interactions that matter for protein folding are those present in the native structure. In this case, one defines a set of contacts from the native structure, and then allows an unstructured chain of the same amino acid sequence re-fold into the native state. In doing so, Gō models are able to simplify the potential energy term significantly allowing for faster simulations. In spite of this Gō models have been able to give reasonable predictions for folding energy barriers, and can sample multiple folding pathways. (230-233)

The other useful coarse-grained approach is using simplified potential energy terms, the most successful of which is discrete molecular dynamics (DMD). The basic idea behind DMD is to extend accessible simulation time by using longer integration steps due to approximating interaction potentials as square wells. (272, 273) The DMD algorithm uses conservation of energy and momentum combined with the ballistic equations of motion (non-Newtonian), to speed of simulation time by a factor of 10^8 - 10^9 .(264) The results presented in this work, have all relied on the all-atom DMD

method, so the next section will describe the method, and energy potential assumptions in detail.

A.5 All Atom Discrete Molecular Dynamics (AADMD)

This section will provide a more detailed overview of the AADMD method because the method has been used to generate the results detailed in Chapter 4. Unfolding Energy Barriers of the Coiled coil, γ - γ and α C region of the Fibrin Molecule. It begins by describing the equations of motion in DMD and then describes the basic potential energy terms currently employed by the Dokholyan lab AADMD program.

A.5.1 DMD Equations of Motion

The DMD method was first used by Zhou and Karplus,(272) and was implemented by Dokholyan et al. in a Gō model study of a simple protein.(264) The idea is to model the atoms as hard spheres of unit mass. The interaction potentials between any given pair of atoms are approximated as finite square wells (See Eq. A.35 and Figure A.1) where the depth of the well is determined empirically. “Neighboring interactions” such as bonds, bond angles and dihedral angles are modeled as an infinitely high square wells. The potential energy of the whole system is given by:(273)

$$E = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N U_{i,j} \quad (\text{A.34})$$

$$U_{i,j} = \left\{ \begin{array}{ll} \infty & |\vec{r}_i - \vec{r}_j| \leq a_o \\ -\varepsilon & a_o < |\vec{r}_i - \vec{r}_j| \leq a_c \\ 0 & |\vec{r}_i - \vec{r}_j| > a_c \end{array} \right\} \quad (\text{A.35})$$

Where ε is the empirically calculated potential well depth, a_o is the radius of the atomic hard sphere, and a_c is the empirically defined potential cutoff length, and \vec{r}_i and \vec{r}_j are the

positions of the i^{th} and j^{th} atoms respectively. During the simulation the atoms evolve with constant velocity until encountering a potential. Here, Δ_{ij} is the native contact map, which equal 1 if the two atoms in contact in the native state, and 0 if the two atoms are not in contact in the native state. The cutoff distance a_c is used to define a contact. At that point the velocity changes instantaneously in accordance with the ballistic equations of motion (not Newton's equations of motion). The ballistic equations of motion are defined as follows: (274)

$$\vec{r}_i(t + \tau) = \vec{r}_i(t) + \vec{v}_i(t)\tau \quad (\text{A.36})$$

$$\vec{v}_i(t + \tau) = \vec{v}_i(t) + \Delta\vec{p}/m_i \quad (\text{A.37})$$

where τ is defined as the minimum time between collisions as described by equation A.43, and $\Delta\vec{p}/m_i$ is the change in velocity of the i^{th} particle during the collision. The new velocities can be calculated through the conservation of energy, and momentum:

$$m_i\vec{v}_{i,o} + m_j\vec{v}_{j,o} = m_i\vec{v}_{i,f} + m_j\vec{v}_{j,f} \quad (\text{A.38})$$

$$\frac{1}{2}m_iv_{i,o}^2 + \frac{1}{2}m_jv_{j,o}^2 = \frac{1}{2}m_iv_{i,f}^2 + \frac{1}{2}m_jv_{j,f}^2 + \varepsilon \quad (\text{A.39})$$

Leading to:

$$\Delta\vec{p} = \frac{m_im_j}{m_i + m_j} \left(\sqrt{(\vec{v}_{j,o} - \vec{v}_{i,o})^2 - 2\frac{m_i + m_j}{m_im_j}\varepsilon} - (\vec{v}_{j,o} - \vec{v}_{i,o}) \right) \quad (\text{A.40})$$

If the combined kinetic energy of the particles is high enough, i.e. if:

$$\varepsilon < \frac{m_im_j}{2(m_i + m_j)} (\vec{v}_{j,o} - \vec{v}_{i,o})^2 \quad (\text{A.41})$$

the particles can escape the potential well. If they remain in the well, then:

$$\Delta\vec{p} = \frac{2m_im_j}{m_i + m_j} (\vec{v}_{i,o} - \vec{v}_{j,o}) \quad (\text{A.42})$$

DMD simulations evolve as a series of collisions, where a sorting algorithm is used after each step to determine the next collision. The time between the next collision for any two particles can be calculated using:

$$t_{ij} = \frac{(\vec{r}_j - \vec{r}_i) \cdot (\vec{v}_j - \vec{v}_i) \pm \sqrt{[(\vec{r}_j - \vec{r}_i) \cdot (\vec{v}_j - \vec{v}_i)]^2 - (\vec{v}_j - \vec{v}_i)^2 [(\vec{r}_j - \vec{r}_i)^2 - a_c^2]}}{(\vec{v}_j - \vec{v}_i)^2} \quad (\text{A.43})$$

and τ , the time between collisions as defined in eq. A.36 is found using the minimum t_{ij} for each step. Because a constant integration time step is not used in DMD, τ is not exact from collision to collision. In spite of this, an approximate correlation to real time can be made using $KE=0.5mv^2$. Assuming a mass of a hydrogen atom (2×10^{-27} kg), a distance of 0.1nm between collisions, and an energy of 1 kcal/mol gives an average time step of ~ 50 femtoseconds.

A.5.2 AADMD Potential Energy Terms

All-Atom DMD simulations of proteins include all heavy atoms and polar hydrogen atoms of amino acids. The amino acids are connect via three different constraint potentials: covalent bonds between consecutive atoms (i, i+1), angular constraints between nearest neighbor atoms (i, i+2), and dihedral constraints between next nearest neighbors (i, i+3). (159) The covalent bands and bond angles are modeled as infinite square well potentials with effective bond lengths d_{ij} and θ_{ij} (where d_{ij} and θ_{ij} are the distance and angle between the i^{th} and j^{th} atom) respectively which are allowed vary by $\sigma_{d,ij}$ and $\sigma_{\theta,ij}$ (See Figure A.1). Dihedral angle potentials are modeled as multi-step well defined by a set of distance parameters (d_{\min} , d_o , d_1 , d_2 , d_{\max}). The dihedral distance parameters are determined empirically by a search of the all dihedral interactions in a database of protein structures and can be found in the supplemental material of Ding et al. 2005.(218)

The DMD non-bonded terms include Van der Waals interactions, implicit solvation, and hydrogen bond interactions. Each will be described in brief detail. Van

der Waals interactions are defined using a square well version of the Lennard-Jones (LJ) potential:

$$E_{LJ} = \sum_{i,j} 4\varepsilon_{ij} \left(\frac{b_{ij}}{\bar{r}_i - \bar{r}_j} \right)^{12} - \left(\frac{b_{ij}}{\bar{r}_i - \bar{r}_j} \right)^6 \quad (\text{A.44})$$

where $b_{ij} = b_i + b_j$ and is the Van der Waal radius between the i^{th} and j^{th} particle and ε_{ij} is the interaction strength. For the solvation term, DMD uses the Lazaridis-Karplus (LK or EEF1) solvation model.(246) Here the solvation energy is defined as:

$$E_{LK} = \sum_{i,j} - \frac{2\Delta G_i^{free} V_i \exp[-((\bar{r}_i - \bar{r}_j) - 1.12a_o) / \lambda_i]^2}{4\pi^{1.5} \lambda_i (\bar{r}_i - \bar{r}_j)^2} - \frac{2\Delta G_j^{free} V_j \exp[-((\bar{r}_i - \bar{r}_j) - 1.12a_o) / \lambda_j]^2}{4\pi^{1.5} \lambda_j (\bar{r}_i - \bar{r}_j)^2} \quad (\text{A.45})$$

where the parameters are defined as the reference solvation energy (ΔG^{free}), volume of atoms (V), correlation length (λ) and atomic radius (σ) and are all taken from Lazaridis and Karplus (Lazaridis et al., 1999).(246) The discretization of Equations A.44 and A.45 can be done by characterizing a hardcore cutoff distance, d_o , followed by a series of potential steps defined by d_i , and ε_i where d_i is the distance of the step and ε_i is the energy associated with the step (See Figure A.1). The cutoff distances for LJ and LK interactions are defined to be 0.65nm, and b_{ij} and ε_{ij} and all LK parameters are defined empirically; a table of the discretized parameters used in all atom DMD can be found in the supplemental material of Ding et al. 2005.(218)

Hydrogen bond (HB) interactions depend on distance and angle between atoms. In AADMD, hydrogen bond formation is calculated using the reaction algorithm proposed by Ding et al. (275) The method involves the formation of a HB between acceptor (A) and Donor Hydrogen (DH) atoms if they come within a certain distance of each other (0.42nm) and meet certain energetic requirements. If the HB reaction does

occur, then the atom type changes from A to A' and DH to DH'; A' and DH' cannot form hydrogen bonds with other atoms, and there is a potential energy gain ϵ_{HB} between the A' and DH' atoms. In addition, orientation (angle) potentials are included between A' and the donor atom D, as well as DH' and X (the heavy atom attached to A'). (See Figure A.1) In order for an HB reaction to occur, the total kinetic energy of A and DH must be greater than ϵ_{HB} , otherwise the atoms do not change their types and undergo a ballistic collision. If the reaction is successful, then A' and DH' interact with other atoms according to the interaction parameters related to their new types. The total potential energy change is given by:

$$\Delta E = \epsilon_{HB} + \sum_i \sum_j E(A', Y_k) - E(A, Y_k) + E(DH', Y_k) - E(DH, Y_k) \quad (A.46)$$

Where $E(A', Y_k)$ represents the energetic interactions between A' and all other (k) non HB-participating atoms (Y); $E(A, Y_k)$ represents the original energetic interactions between A and all other (k) atoms and so on. Because hydrogen bonding is affected by solvation, a correction term is included in the Lazaridis-Karplus (LK) solvation model, such that hydrogen bonded atoms have a weaker solvation energy ($\Delta G_{free}=3.85\text{kcal/mol}$) than their non hydrogen bonded counterparts ($\Delta G_{free}=5.85\text{kcal/mol}$), and hydrogen bonds buried in the core of the atom have different ΔG_{free} 's than hydrogen bonds exposed on the surface. All hydrogen bonding potential energies, for each type of atomic interaction are determined empirically; a table of the parameters can be found in the supplemental material of Ding et al. 2008. (159)

AADMD simulations performed at constant temperature are done so using an Andersen thermostat (See above). The strength of the thermostat can be controlled by adjusting the frequency and number of particles that have their velocities re-scaled.

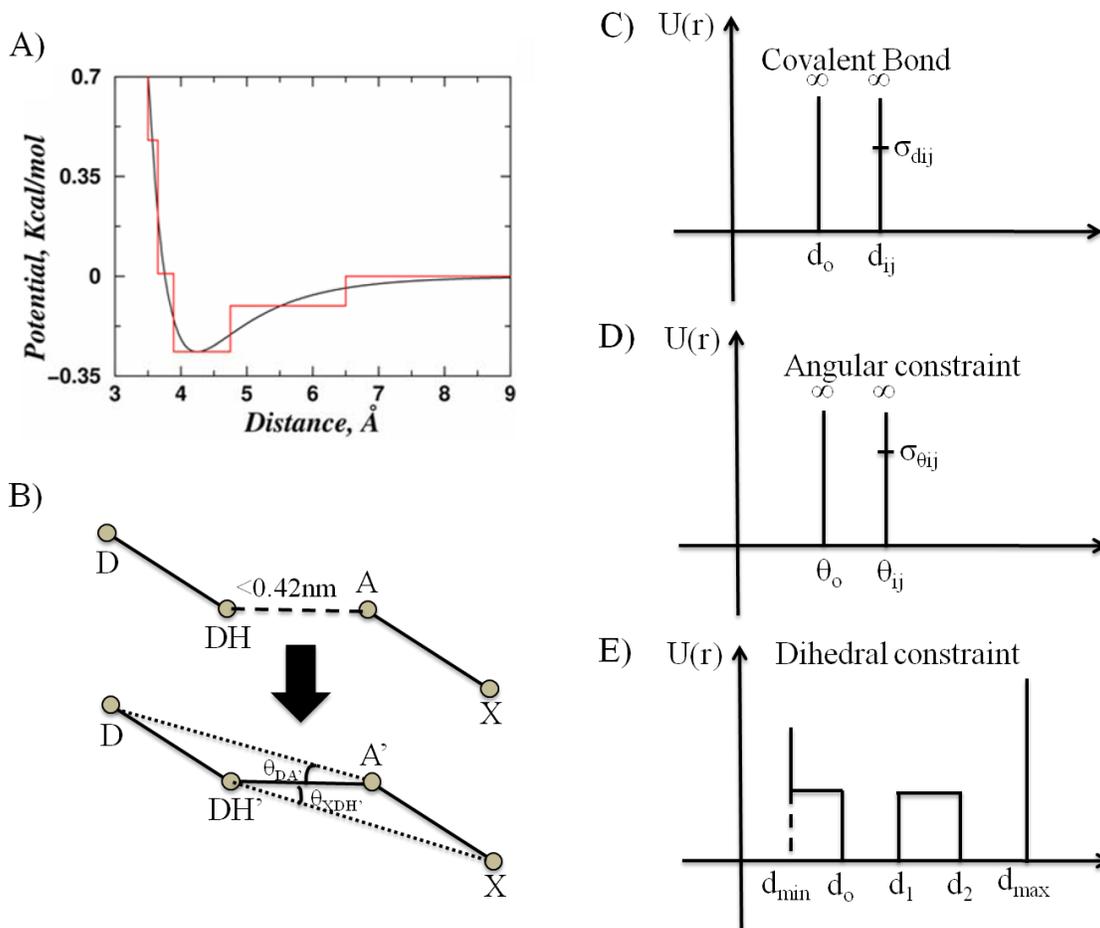


Figure A.1: DMD Potential Energy Terms: DMD employs discretized potential energy terms. As can be seen in (A) continuous potentials are approximated by a series of square well potentials (red). (B) Hydrogen bonds are simulated by changing the “atomic type” from DH to DH’ and A to A’. Hydrogen bonded vs. non-hydrogen bonded atoms have different solvation free energies as described in the text. (C-E) show potential energy terms for other types of potentials. (Reprinted with modifications from Ding, F., Tsao, D., Nie, H., Dokholyan, N.V. “Ab Initio Folding of Proteins with All-Atom Discrete Molecular Dynamics” *Structure* 2008, 16(7):1010-1018. Used with Permission from Elsevier Limited. © 2008)

A.6 Calculating free energies from MD simulations

A.6.1 The Weighted Histogram Analysis Method (WHAM)

In molecular dynamics simulations, a primary goal is to calculate the free energy landscape of the system. Free energies can be analyzed in two ways: 1) Calculating potential mean force (PMF), which is roughly equivalent to the free energy of the system

along a specific coordinate. [From the PMF most thermodynamic quantities can be derived](276) The other approach is to calculate the free energy differences between two states in the system. (161, 277, 278) While approach (2) is useful if a protein is known to reside in two primary conformations, it is often more interesting to study the PMF of a protein system as it evolves dynamically.

The challenge of MD is to sample enough of the protein conformational space to generate an accurate PMF, given the often high energy barriers separating parts of the space. Parallel Tempering molecular dynamics is one way of sampling more conformational space. This method involves running S copies (called replicas) of the system, randomly initialized, at different temperatures. The entire simulations can be carried out at these varying temperatures if the temperature differences are small and there is energy overlap between systems, or the configurations can randomly be chosen to be exchanged between temperatures. The second technique is called Replica Exchange Molecular Dynamics (REXMD). (279, 280) The challenge is then to generate a general PMF for the entire protein system from the various temperature runs, as information about the unfolding barriers is contained in the high temperature data and information about the low energy folded state is contained in the low temperature data.

The weighted histogram analysis method (WHAM) is a technique to combine data from multiple simulations at different temperatures (and/or biasing potentials) in order to obtain the PMF of the general system. The WHAM technique is carried out using a specific reaction coordinate of interest (such as the protein RMSD), and an unbiased PMF can be generated along that reaction coordinate.(281) The generalized WHAM method was first developed by Kumar et al., (161) and has subsequently been

modified to specific simulation cases by a variety of groups. (281-284) Here, I will follow the generalization of Gallicchio and Levy to lay out the basic WHAM theory. (284)

In statistical mechanics, the free energy can be found by taking the natural logarithm of the partition function: $F = -k_B T \ln(Z)$. Correspondingly, the PMF of a system is given by:

$$PMF(x) = -k_B T_o \ln(P^o(x, E)) \quad (A.47)$$

Where, $P^o(x, E)$ is defined as the probability density of the system along reaction coordinate x at temperature T_o . (161)

Suppose that we want to generate the PMF at room temperature (T_o). Assume that “**S**” number of simulations have been run on a given set of protein state replicas ($i=1 \dots S$). The replicas were run at S different temperatures ranging from $T_{i=1}$ to T_S . It is also possible to have additional biasing potentials $V_i(x)$ where x is the reaction coordinate of interest. For each replica simulation, the results of the unbiased potential energy E and the reaction coordinate x can be separated into **B** and **C** numbers of bins respectively as the simulations evolve. Choosing the limits of the **B** and **C** to be sufficiently large to contain the highest and lowest data points, one can generate a two dimensional histogram of the energy states ($E_k ; k=1 \dots C$) and reaction coordinate states ($x_j ; j=1 \dots B$). Thus for a given REXMD simulation, one can generate S such 2-D histograms. The indexing can get complicated, but I’ll represent the count in the j^{th} x bin and k^{th} E bin in the i^{th} replica simulation as $n_i(j, k)$. Thus $n_i(j, k)$ tells us how many times at a given temperature the simulation was in reaction coordinate state x and energy state k . The goal is to estimate the unbiased probability density of states of the ensemble at temperature T_o : $P^o(x, E)$. Based on our histograms $n_i(j, k)$, we can estimate the biased density of states for any given

temperature run: $p_i(j, k)$. This can be related to the unbiased density of states at reference temperature T_o , $p^o(j, k)$ by the following relation(284):

$$p_i(j, k) \propto p^o(j, k) \exp\left(\frac{E_k}{k_B T_o} - \frac{E_k}{k_B T_i}\right) \quad (\text{A.48})$$

or

$$p_i(j, k) = f_i c_i(j, k) p^o(j, k) \quad (\text{A.49})$$

Where f_i is a normalizing constant, i.e. $\frac{1}{f_i} = \sum_{j=1}^B \sum_{k=1}^C c_i(j, k) p^o(j, k)$

and $c_i(j, k)$ is the biasing factor which accounts for the effect of the temperature and any biasing potentials:

$$c_i(j, k) = \exp\left(\frac{E_k}{k_B T_o} - \frac{E_k}{k_B T_i}\right) \exp\left(\frac{-V_i(x)}{k_B T_i}\right) \quad (\text{A.50})$$

In simulations, there are also usually correlations between the counts in the $n_i(j, k)$ histograms, so in practice, each $n_i(j, k)$ is reduced by a factor $g_i(j, k) = 1 + 2\tau_i(j, k)$ where $\tau_i(j, k)$ is the correlation time of the simulation in the (j, k) bin of simulation i . In practice, the $g_i(j, k)$ factors are difficult to compute, but Kumar et al. showed that picking $g_i(j, k)$ values different by a factor of 10, only changed the respective relative PMF's by about 2%, so choosing a reasonable value for $g_i(j, k)$ such as 1.0 typically suffices, or even ignoring the $g_i(j, k)$'s altogether is typically ok. (161)

In order to calculate $p^o(j, k)$, Bayes theorem of conditional probabilities can be used and the maximum likelihood method can be applied. Bayes' theorem states that(282, 285):

$$\tilde{P}(A|B) = \frac{\tilde{P}(B|A)\tilde{P}(A)}{\tilde{P}(B)} \quad (\text{A.51})$$

Where $\tilde{P}(A|B)$ is the conditional probability of A , given B ; $\tilde{P}(A)$ is the unconditional probability of A , meaning it doesn't take into account any information about B ; $\tilde{P}(B)$ is the unconditional probability of B ; and $\tilde{P}(B|A)$ is the conditional probability of B , given A .

To find $p^o(j,k)$ we first to calculate $\tilde{P}(n_i(j,k) | p_i(j,k))$, the conditional probability of observing a particular set of counts $n_i(j,k)$ in the histogram, given a particular set of probabilities $p_i(j,k)$. Assuming a multinomial distribution for $n_i(j,k)$, allows the conditional probability to be defined as(284):

$$\tilde{P}(n_i(j,k) | p_i(j,k)) = \frac{\left(\sum_{i,j,k} n_i(j,k) \right)!}{\prod_{i,j,k} (n_i(j,k))!} \prod_{i,j,k} p_i(j,k) \propto \prod_{j=1}^B \prod_{k=1}^C \left(f_i c_i(j,k) p^o(j,k) \right)^{n_i(j,k)} \quad (\text{A.52})$$

Using this, estimates of $p^o(j,k)$ can then be determined by maximizing the likelihood function $L(p_i(j,k))$ with respect to the unbiased probabilities:

$$L(p_i(j,k)) = \sum_i \ln(\tilde{P}(n_i(j,k) | p_i(j,k))) \propto \prod_{i=1}^S \prod_{j=1}^B \prod_{k=1}^C \left(f_i c_i(j,k) p^o(j,k) \right)^{n_i(j,k)} \quad (\text{A.53})$$

$$\frac{dL(p_i(j,k))}{dp^o(j=1\dots B, k=1\dots C)} = 0 \quad (\text{A.0.54})$$

Doing this leads to two simultaneous nonlinear equations for $p^o(j,k)$ and f_i :

$$p^o(j,k) = \frac{\sum_i n_i(j,k)}{\sum_i N_i f_i c_i(j,k)} \quad (\text{solved for all } j,k) \quad (\text{A.55})$$

$$f_i^{-1} = \sum_j \sum_k c_i(j,k) p^o(j,k) \quad (\text{A.56})$$

Where N_i is the total number of counts in the i^{th} histogram. Equations (9) and (10) are together called the WHAM equations.(161, 284) They are typically solved iteratively, choosing reasonable an initial value for each of the f_i 's, substituting that into equation (9) and solving for the $p^o(j,k)$'s, and then using those values to re-solve for the f_i 's. The process is repeated until converging to a constant value. Finally the PMF can be solved for using equation (1):

$$PMF(x) = -k_B T_o \ln(P^o(x, E)) = -k_B T_o \ln \sum_k p^o(x, E_k) \quad (\text{A.57})$$

In addition to PMF(x), we can also generate a Specific Heat (C_v) profile for the simulation as a function of temperature. The partition function is given by:(161)

$$Z(T) = \sum_k \sum_j p^o(j,k) \exp\left(\frac{-E_k}{k_B T}\right) \quad (\text{A.58})$$

We can then use $Z(T)$ to solve for the average energy using: $\langle E \rangle = \sum_s E_s e^{-\beta E_s}$,

$$\langle E \rangle = \sum_k \sum_j E_k p^o(j,k) \exp\left(\frac{-E_k}{k_B T}\right) \quad (\text{A.59})$$

$$\langle E^2 \rangle = \sum_k \sum_j E_k^2 p^o(j,k) \exp\left(\frac{-E_k}{k_B T}\right) \quad (\text{A.60})$$

Finally using the definition of C_v (A.18) we arrive at:

$$C_v = \frac{\partial \langle E \rangle}{\partial T} = \frac{\langle E^2 \rangle - \langle E \rangle^2}{k_B T^2} = \frac{\sum_k \sum_j E_k^2 p^o(j, k) \exp\left(\frac{-E_k}{k_B T}\right) - \left(\sum_k \sum_j E_k p^o(j, k) \exp\left(\frac{-E_k}{k_B T}\right) \right)^2}{k_B T^2} \quad (\text{A.61})$$

A.6.2 Steered Molecular Dynamics (SMD) to calculate unfolding energy barriers

Many proteins such as actin, titin and fibrin are designed to respond to forces.(78, 100, 175) Because of this, there is an interest in the scientific community to understand the unfolding pathways of these proteins, to understand how natural materials are able to withstand high physiological stresses. Beautiful single molecule experiments using the atomic force microscope, laser tweezers, and fluorescence techniques such as FRET (Fluorescence Resonance Energy Transfer) have begun to explore the unfolding pathways of these proteins.(75, 100, 286) However, most of these techniques can only measure differences in distances between the folded and unfolded state under force, and the orientation of the unfolded state can only be speculated. Because of this, a new MD technique called steered molecular dynamics (SMD) has been developed to “watch” a protein unfold under force.

SMD was first developed by Izrailev et al. to study the biotin-avidin unbinding,(287) but has now been used on many force-related proteins such as fibrin, titin, and fibronectin.(68, 236, 288) SMD is typically performed in “constant velocity” mode, where one end (N or C terminus) of the protein is held fixed, and the other end is constrained by a harmonic spring to a pull-point, and the pull-point is moved with a constant velocity in a given direction. The force exerted on the protein is then $F = k_h(vt - x_o)$ where k_h is the harmonic spring constant, v is the velocity, and x_o is the original position of the pulled atom.(236) Experiments performed in this way are similar to AFM

experiments, where an AFM tip of given spring constant is pulled away from the surface at a given velocity. The downside of this technique, is that while AFM experiments are performed on the order of seconds, with pulling speeds in the 20nm/s-2000nm/s range,(103) MD simulations typically last 1-100ns. In order to unfold a protein on these timescales, pulling rates of m/s must be used, much faster than experimental rates, and likely faster than most physiological processes. Thus, while the shape of the force curve in an SMD simulation, will often match the shape of an experimentally unfolded protein, the magnitudes of the forces will be off by orders of magnitude.(68) In addition, the unfolding pathways may be altered under these extreme conditions.

Because of this, a new SMD technique of unfolding a protein at constant force has recently been used in combination with Discrete Molecular Dynamics (DMD).(172) In constant force mode, one end of a protein is held fixed, and the other end is constrained to a pull-point at a constant distance. A constant force is then applied to the pull point in a given direction by applying a discretized step-function with a constant energy jump, dE , for every distance step, dr , taken in that direction. The force can then be calculated as dE/dr ; a force of 70pN can be generated with $dE = 0.1\text{kcal/mol}$ and $dr=0.01\text{nm}$.

The average time that a protein can withstand a given force can be estimated by a Bell equation:(170)

$$\tau_B = \tau_o \exp\left[\frac{[\Delta G_B - Fx_u]}{k_B T}\right] \quad (\text{A.62})$$

where τ_B is the unfolding time for a protein under a given force (F), τ_o can be thought of as the time between energy well escape attempts, or the reciprocal of the oscillation frequency of atoms in a solid, ΔG_B is the unfolding energy barrier in the absence of force, and x_u is distance between the unfolded energy state and the top of the unfolding energy

barrier (See Figure 4.6).⁽¹⁷²⁾ One way to think of applying a constant force to a protein is that you are lowering the unfolding energy barrier ΔG , giving the protein a greater probability to escape the folded conformation. There should exist then, a critical force, F_c where the lowered unfolding energy barrier is comparable to the free energy of protein in its folded state, and it is equally probably to see the protein in a folded or unfolded state. By running a series of simulations at various constant forces, one can estimate F_c by looking for the τ_B for each simulation and finding the force where τ_B begins increasing exponentially (See Figure 4.11).

Appendix B: Optical Microscopy Techniques

Optical microscopy is a powerful and widely used technique to observe samples at higher magnifications. Optical microscopes date back to the early 1600's, perhaps being invented by Galileo Galilei; and at their simplest level can consist of visible light, a lens, and an image capturing device (ranging from the human eye to sophisticated cameras). Modern microscopes typically consist of a series of lenses designed to optimize illumination and resolution for a particular sample. In this paper it will be impossible to discuss the many techniques used in modern microscopy. The reader is referred one of many reviews for more details.(289) In this section, some basic concepts in optical microscopy will be presented, and a focused section on fluorescence microscopy will be presented. The goal is to provide an elementary background for the experimental techniques used in this research, as well as suggesting techniques which could be used for future studies on fibrin.

Figure 5.1 shows the basic components of a transmission light microscope. It consists of a light source, several sets of filters, a condenser lens to focus the light onto the sample, an objective lens to collect and magnify light coming off the sample, and a detector to capture the image. Light passing through the sample, either passes through the sample undisturbed, or diffracts off of the specimen. Abbe, who developed microscope theory in the late 19th century still used today, developed that idea that microscope specimens can be treated as complex diffraction gratings with openings of various sizes. The greater the number of diffracted orders collected by objective lens, the more accurately the image will represent the original object.(289) Thus, the objective lens is in some ways, the most important part of the microscope.

Important characteristics of a lens include its magnification, its numerical aperture (NA), and its immersion medium. The magnification, which is the enlargement factor of an image, is governed by the focal length (f_{ob}) of the objective lens (the distance from the lens to where the image is in focus), and the distance to the sample (d_{ob}) [$M=d_{ob}/f_{ob}$]. The numerical aperture defines the conical angle over which an objective can receive light. Typically the NA is defined as $n_1 \sin(\theta)$, where n_1 is the refractive index of the immersion medium and θ is the half-angle of the acceptance cone. Because $\sin(\theta)$ gets larger as θ approaches 90° , there is a trade-off between the working distance of the objective lens, and the NA of the lens. This equation also shows why the immersion medium is important. Having a larger n_1 , allows for higher NA irrespective of θ . This happens because (as described in detail below in the discussion of Snell's Law), when light passes from one medium into another of a different index of refraction, the light path is altered.

In any type of microscopy, the natural question that arises is: "What is the smallest object resolvable with this technique?" For optical microscopy, two objects can be resolved if the principal diffraction maximum of one object coincides with the first diffraction minimum of the other. For a circular objective lens, the angular distance from the diffraction maximum (also known as the Airy Disk after physicist Gerorge Airy) to the first diffraction minimum is governed by the Raleigh criterion(289):

$$\sin(\theta) \approx 1.22 \frac{\lambda}{d} \text{ (B.1)}$$

where λ is the wavelength of light (inversely proportional to the frequency ν) and d is the diameter of the lens aperture. This can be converted into a resolution length by multiplying to the focal length of the lens.(290)

$$\Delta l \approx 1.22 \frac{f_{ob}\lambda}{d} \approx 1.22 \frac{\lambda}{2xNA} \quad (\text{B.2})$$

Where Δl is the minimum distance over which two objects would appear distinct, and f_{ob} is the focal length of the objective, and NA is the numerical aperture of the lens system. For visible, light and an N.A. of 1.4, this puts the spatial resolution at about 200nm. So, fluorophores closer than 200nm would appear to be the same thing.

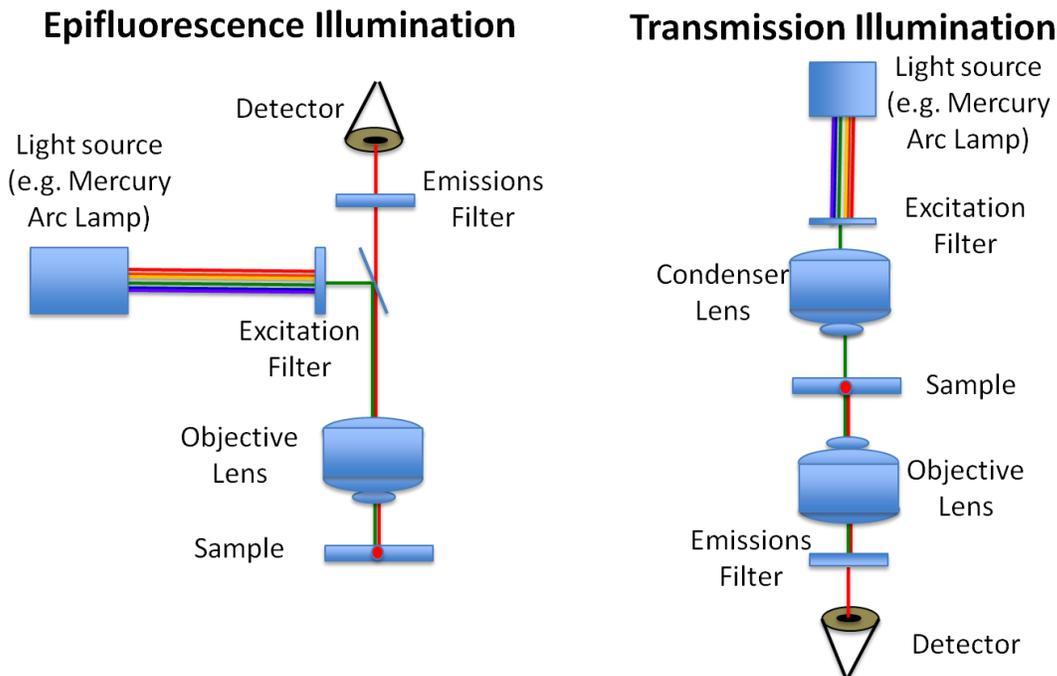
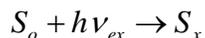


Figure B.1 Epifluorescence vs. Transmission Illumination: In Transmission illumination, the condenser lens focuses light through the sample, into the objective lens which magnifies the image of the sample into the detector. In contrast, in Epifluorescence illumination, the objective lens acts to both focus light into the sample, and collect light reflected (or emitted) from the sample. The acts to reduce ambient illumination light from reaching your detector.

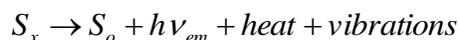
With this background, we now turn to a particular type of light microscopy, fluorescence microscopy. Fluorescence microscopy allows the study of a wide variety of

biological processes from cell motility down to the motion individual protein and DNA molecules, by labeling parts of the sample with a light emitting fluorophores.(206, 292-294) There are a wide variety of fluorescence microscopy methods, but they all follow the same general principle. The sample is embedded with (or contains natural i.e. autofluorescence) fluorescent molecules (fluorophores). Illumination of the sample with light of a particular wavelength then causes the fluorophore to fluoresce. The light emitted by fluorescence is then detected through the microscope. For a full review of fluorescence techniques the reader is directed to many excellent books and reviews (295-298), here a brief and incomplete overview will be presented.

Fluorescence can occur when light excites the electrons in a chemical compound into a higher energy state (for a deeper understanding of singlet and triplet energy states and their role in the fluorescence process the reader is directed to the aforementioned reviews) via the reaction:



Where S_o is the ground state energy, h is Planck's constant and ν_{ex} is the frequency of the excitation light, and S_x is the energy of the excited state. Nearly all chemical compounds can absorb light, however only those whose $\Delta E = S_x - S_o = h\nu_{ex}$ can fluoresce, because fluorescence only occurs when the incoming energy matches a specific energy gap (known as the Stokes shift) between two electron energy states (a direct result of quantum mechanics). When an electron does jump into the S_x state, it will then relax back down to the ground state, by emitting another electron:



where ν_{em} is the frequency of the emitted electron. ν_{em} is typically smaller than ν_{ex} , meaning that the emitted photon has a lower energy than the excitation light.

Because of the quantized nature of fluorescence, new research is constantly looking for better fluorophores. Fluorophores can be characterized by several properties, including their excitation and emissions spectra, their quantum efficiency, their fluorescence lifetime, and their photobleaching lifetime. The quantum efficiency is generally the ratio of photons emitted by the fluorophore to photons absorbed(299), while the fluorescence life time is the time in which an electron will stay in its excited state before fluorescing. The photobleaching lifetime, on the other hand, is the length of time over which a fluorophore can continue absorbing and emitting photons before it losing its ability to fluoresce typically due to light induced chemical modification of the fluorophore. “Good” fluorophores can undergo 10,000-40,000 cycles and last ~1s before bleaching. (295)

Fluorescence microscopy can utilized several different types of illumination. The canonical method of fluorescence illumination is epifluorescence. In this method of illumination, the excitation light passes through an objective lens, directly into the sample, the emitted light then passes back through the same objective lens and into a viewing area.(295) Thus, the objective lens acts both as the illumination condenser and the fluorescent light collector. Epifluorescence illumination is typically done using a series of filters and a dichroic mirror which reflects light of one wavelength and is transparent to light of another wavelength. The advantage of this type of illumination is that only a small percentage of the exciting light that is reflected off the sample back towards the detector, as opposed to a transmission microscope where the illumination

light is transmitted through a sample and all of the non-absorbed light passes to the other side. (See Figure B.1) Each dichroic mirror must specifically be designed for an excitation wavelength and an emission wavelength.

In spite of its advantages over transmission microscopy, epifluorescence illumination can often suffer due to large background fluorescence from the fact that the entire sample is illuminated during excitation. To overcome these deficiencies several methods have been developed. One common technique is the use of a confocal microscope.^(300, 301) In a confocal microscope, a pinhole filter is placed in front of the detector so that only focused light passes to the detector while all out of focus light misses the pinhole. In so doing, the user is able to focus on one focal plane at a time, while light from other focal planes is eliminated, removing much of the background light. To get a full image in confocal microscopy, the focus must be scanned across a plane of the sample. The confocal technique can additionally be used to image at various sample depths by adjusting the focal plane of the imaging.

Another set of low-background illumination techniques are the so-called multiphoton techniques.^(302, 303) These techniques involve illuminating the sample with a wavelength around double the actual emission wavelength (since energy is inversely related to wavelength, $E=h/\lambda$ where h is Planck's constant and λ is the wavelength, this corresponds to roughly $\frac{1}{2}$ the needed excitation energy). However, if a pulsed laser is used, two photons (or more) could reach the point of focus nearly simultaneously, providing enough energy for electron excitation.⁽³⁰²⁾ In this way, fluorophore excitation will only occur at the point of focus, thereby eliminating excitation of out-of-focus fluorophores. Multiphoton techniques allow for deeper sample imaging

than confocal techniques, due to the fact that the longer wavelengths suffer less scattering in the sample, and the that illumination light is focused on one point in the sample.

One final low background illumination technique that should be mentioned is Total Internal Reflection Illumination (TIRF).(304, 305) TIRF takes advantage of the fact that an evanescent wave is developed when light is totally internally reflected at an interface of two media with different refractive indices. As a brief refresher, Snell's Law states that when light (polarized in the \hat{y} direction in this example), is incident on an interface, the angle of transmission, the transmitted plane wave vector, and the electric field can be given by (See Figure B.2):(306)

$$\sin(\theta_I)n_1 = \sin(\theta_T)n_2 \quad (\text{B.3})$$

$$\vec{k}_T = k_T \sin(\theta_T)\hat{x} + k_T \cos(\theta_T)\hat{z} \quad (\text{B.4})$$

$$\vec{E}_T(x, z, t) = \vec{E}_o \exp(2\pi i [xk_T \sin(\theta_T) + zk_T \cos(\theta_T) - \nu t]) \quad (\text{B.5})$$

Where θ_I is the angle of incidence, n_1 is the index of refraction on the incident side, θ_T is the exit angle on the transmission side of the interface, and n_2 is the corresponding index of refraction, and ν is the wave frequency.

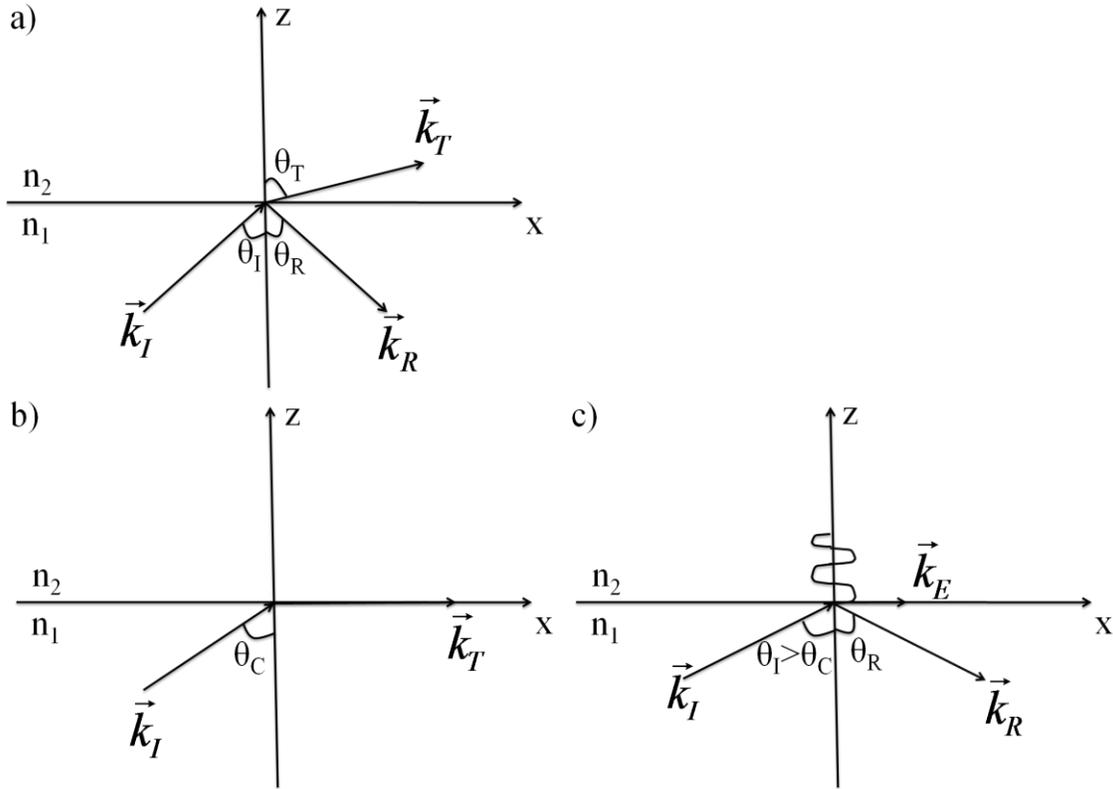


Figure B.2 Snell's Law when $n_2 < n_1$: a) $\theta_I < \theta_C$, the light passes into the sample; b) $\theta_I = \theta_C$, the wave is reflected along the interface c) $\theta_I > \theta_C$ generates an evanescent into the sample.

If $n_2 < n_1$, there is an angle θ_C at which and beyond which all of the light will be reflected.

$$\sin(\theta_C)n_1 = \sin(90^\circ)n_2 \text{ (B.6)}$$

At angles beyond θ_C :

$$\sin(\theta_T) = \sin(\theta_I > \theta_C) \frac{n_1}{n_2} > 1 \text{ (B.7)}$$

making the \hat{z} component of the \vec{k}_T vector imaginary. It can then be shown that the electromagnetic wave propagates as an evanescent wave:

$$\vec{E}_E(x, t) = \vec{E}_o \exp(-\kappa z) \exp(2\pi i[k_T x - \nu t]) \text{ (B.8)}$$

where the attenuation constant, κ is a function of the angle of incidence, and the refractive indices. The evanescent wave propagates in the \hat{x} direction, but exponentially dies off in the n_2 medium. The intensity of the evanescent field can be written as:

$$I(x,t) = \vec{E}(x,t)^2 \approx I(t=0) \exp(-z/\delta) \text{ (B.9)}$$

$$\delta \approx \frac{1}{v_o n_2 \sqrt{[\sin(\theta_1)/\sin(\theta_c)]^2 - 1}} \text{ (B.10)}$$

Thus, using TIRF illumination, one can generate an evanescent wave that's intensity exponentially dies in the transmission medium. Experimental δ values range around 100nm, meaning that only the first 100nm of a sample will be excited fluorescently, yielding very low fluorescence background. The downside is that, in practice, the sample must also be $< 100\text{nm}$ thick, so TIRF illumination can only be used primarily in single molecule studies.

Appendix C: Atomic Force Microscopy

The primary tool for measuring forces in this work was the Atomic Force Microscope (AFM). The AFM was invented by Binnig, Quate and Gerber in 1986, and has quickly risen to prominence as a tool for atomic resolution imaging and force sensing, among other things. (307) Figure C.1 shows a schematic design of a basic AFM. It consists of cantilever (a thin beam supported on one end), with a sharp tip hanging off (called a probe) at the unsuspended end. Probes are often conical in shape, and can taper down to radii of 1-5nm. The cantilevers are typically rectangular or triangular in shape and composed of silicon or silicon nitride. A laser is focused on the back of the cantilever and deflected into a quadrant photodiode detector (QPD) to sense horizontal or torsional deflections in the cantilever.(308) The cantilever is typically mounted on a piezo-electric (piezo) material that allows it to adjust to changes in the laser signal through a feedback loop.

For details on AFM imaging techniques, the reader is directed to a few of many reviews available.(308-312) Basically, imaging with the AFM is done by raster-scanning the tip back and forth over a surface. Imaging can either be performed in contact mode, where a feedback loop adjusts the piezo height to keep the deflection of the cantilever constant; or in dynamic mode where the cantilever is driven near its resonant frequency and the feedback loop adjusts the piezo to keep the amplitude of oscillations constant.

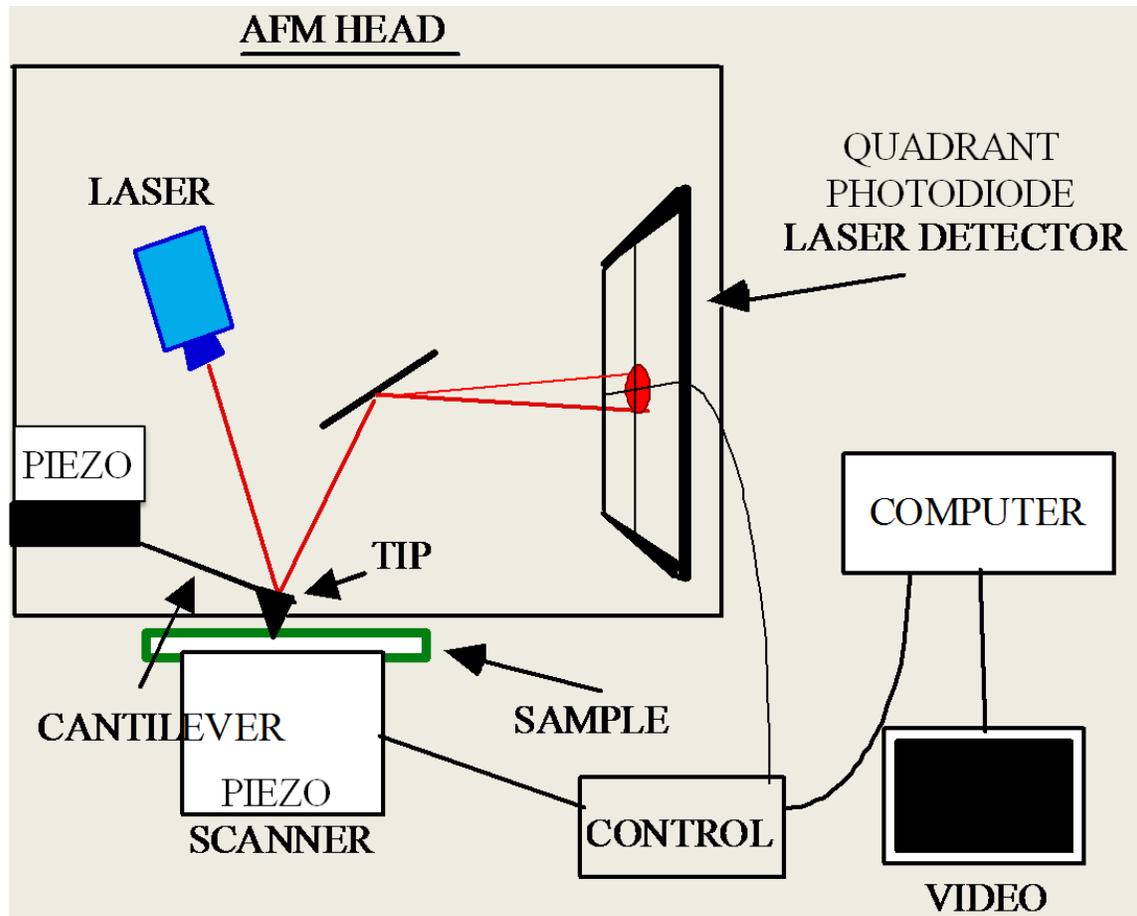


Figure C.1 Basic AFM Setup: One can think of an AFM as a tip mounted on a cantilever which can be used to scan a surface. The cantilever and/or scanner are mounted on a piezo electric material allowing constant adjustments through a feed-back loop. A laser reflects off the tip and into a quadrant photo diode to measure the tip-surface interactions.

There are two main types of dynamic mode operation: non-contact and tapping mode. The difference between the two modes is governed by whether the AFM tip actually touches the surface. In non-contact mode, the AFM tip floats 5-15nm above the surface, and interacts with the attractive Van der Waals forces of the surface. (311) The tip is driven slightly below its resonance frequency at small oscillations. The topography of the sample can then be re-constructed by measuring changes in the oscillation amplitude from Van der Waals forces. In tapping mode, the cantilever is driven at much larger amplitudes, typically between 20-100nm.(311) The tip is scanned back and forth

at a height where it intermittently touches the surface. Changes in the height of the surface cause the vibration amplitude to increase or decrease, and a feedback loop adjusts the tip height accordingly. The difference between the instantaneous tip oscillation amplitude, and the reference amplitude (above the surface), read out as an error signal, can be used to plot the topography of the sample.

In contact mode, the AFM is kept in close contact with the surface as it scans. The tip can either be held at a constant height, or at constant force during scanning. In both modes, the topography of the surface can be measured from the deflection of the signal off the cantilever. Constant height mode imaging is not typically used, as the tip will quickly be damaged if there are high places on the sample surface. In constant force mode, the deflection of the cantilever is kept constant through a feedback loop, where the piezo adjusts the height of the scanner in response to changes in sample height. Because of the deflection can directly be related to a force through calibration (see below), this mode is called constant force mode. The sample topography can be calculated from the error between the instantaneous deflection signal and the reference (set point) signal ($\Delta T-B$ in Figure C.2). The main drawback of contact mode imaging is that soft samples can be damaged from tip-surface interactions.

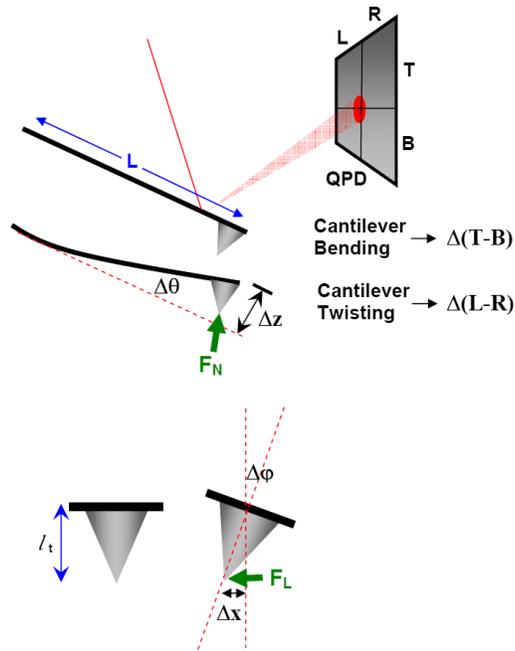


Figure C.2 Normal and Lateral Cantilever deflections: The AFM cantilever can either deflect normally, shifting the top-bottom ($\Delta T-B$) signal of the QPD, or can twist laterally which will show up in the left-right ($\Delta L-R$) QPD signal. After calibration, the $\Delta T-B$ and $\Delta L-R$ signals can be converted into an length or angular deflection of the cantilever which can then be converted into a force through Equations C.1 and C.2.

In contact mode, the AFM can also be used to directly measure forces. This has become a powerful technique in protein unfolding studies.^(98, 102, 103) The AFM cantilever can bend and twist when interacting with the surface (or a protein) as seen in Figure C.2. Due to its inherent length, width, thickness, and composition material each cantilever has a natural spring constant in both the bending (k_N) and twisting or lateral (k_L) modes. Using Hooke's law, the force on the cantilever can then be calculated for each mode as:

$$F_N = k_N \Delta z \text{ (C.1)}$$

$$F_L = k_L \Delta x \text{ (C.2)}$$

according to Figure C.2. The changes in bending (Δz) and lateral (Δx) deflections can be measured from the change in the top-bottom ($\Delta T-B$), and left-right ($\Delta L-R$) laser signals in the QPD according to:

$$\Delta z = \frac{\Delta(T - B)}{S_{T-B}} \text{ (C.3)}$$

$$\Delta x = \frac{\Delta(L - R)}{S_{L-R}} \text{ (C.4)}$$

where S_{T-B} and S_{L-R} are calibration sensitivities (often called the optical lever sensitivity or OLS). S_{T-B} , the bending sensitivity, can be measured by pushing the cantilever into an “infinitely hard” surface. In principle, this is done by pushing the cantilever onto a surface whose stiffness is much higher than that of the cantilever. The sensitivity is measured in units of current or voltage/nanometer and is a measure of how much the T-B signal changes per nanometer or indentation into the surface.

Calculating the S_{L-R} is not as straightforward, but techniques have recently been improving. Early efforts to calculate S_{L-R} often involved some combination of beam bending mechanics, and assumptions about the photodetector.(313-315) Other techniques including the glass rod techniques, and trapezoid techniques involved running the tip into a known surface and inducing a twisting of the cantilever.(316-318) Others have used piezoelectric force sensors as part of their calibration techniques.(319) A recent and promising approach has shown that the torsional OLS can be calculated directly from the thermal noise spectrum of the cantilever, if the spring constant is known a priori. (320) For our measurements, we typically use the geometrical calibration because of its simplicity, and the observation that using other methods (the wedge method for example) did not provide drastically more accurate results.

The spring constants k_N and k_L also can be solved through a variety of methods. One of the simplest and most common methods is using the equipartition function to calculate the spring constant based on the cantilevers thermal fluctuations.(321, 322)

$$k = k_B T / \langle z^2 \rangle \quad (C.5)$$

Where k is the spring constant and $\langle z^2 \rangle$ is the mean squared cantilever fluctuations, which can be calculated for the bending or torsional modes. Updates to the thermal calibration procedures have been produced to account for the non-idealness of the cantilever as a spring.(323) There have also been several geometrical calibration procedures produced derived from Euler-Bernoulli beam bending equations, the most common of which is the Sader method. (314, 315, 324) For rectangular cantilevers the constants are:(319)

$$k_L = \frac{Gwt^3}{3Ll_t^2} \quad (C.6)$$

$$k_N = \frac{Ewt^3}{4L^3} \quad (C.7)$$

where G and E are the shear modulus and Young's modulus respectively of the cantilever, and the geometrical parameters are the width of the cantilever, w , the thickness of the cantilever, t , the length of the cantilever, L , and the length of the tip, l_t . In addition, there have been more involved methods proposed such as adding known masses to the cantilever and measuring the decrease in resonance frequency.(325)

Bibliography

1. D. M. Wootton, D. N. Ku, *Annual Review of Biomedical Engineering* **1**, 299 (1999).
2. S. T. Lord, *Curr Opin Hematol* **14**, 236 (May, 2007).
3. S. T. Lord, *Current Opinion in Hematology* **14**, 236 (May, 2007).
4. R. F. Doolittle, *Annu Rev Biochem* **53**, 195 (1984).
5. J. W. Weisel, *Biophys Chem* **112**, 267 (Dec 20, 2004).
6. J. W. Weisel, *Adv Protein Chem* **70**, 247 (2005).
7. M. W. Mosesson, *Journal of Thrombosis and Haemostasis* **3**, 1894 (Aug, 2005).
8. J. W. Weisel, *Science* **320**, 456 (Apr 25, 2008).
9. R. Bauer *et al.*, *Eur Biophys J* **23**, 239 (1994).
10. C. W. Francis, V. J. Marder, *Blood* **71**, 1361 (May, 1988).
11. R. Marchi *et al.*, *Thromb Res* **118**, 637 (2006).
12. J. M. Kollman, L. Pandi, M. R. Sawaya, M. Riley, R. F. Doolittle, *Biochemistry* **48**, 3877 (May 12, 2009).
13. R. F. Doolittle, J. M. Kollman, *Proteins* **63**, 391 (May 1, 2006).
14. G. Tsurupa *et al.*, *Biochem J* **48**, 12191 (Nov 24, 2009).
15. J. Z. Zhang, C. M. Redman, *Journal of Biological Chemistry* **269**, 652 (Jan 7, 1994).
16. E. V. Lugovskoi, P. G. Gritsenko, S. V. Komisarenko, *Russian Journal of Bioorganic Chemistry* **35**, 393 (Jul, 2009).
17. B. Blomback, B. Hessel, D. Hogg, *Thrombosis Research* **8**, 639 (1976).
18. H. Bouma, T. Takagi, R. F. Doolittle, *Thrombosis Research* **13**, 557 (1978).
19. G. Spraggon, S. J. Everse, R. F. Doolittle, *Nature* **389**, 455 (Oct 2, 1997).
20. A. P. Laudano, R. F. Doolittle, *Proceedings of the National Academy of Sciences of the United States of America* **75**, 3085 (1978).

21. S. J. Everse, G. Spraggon, L. Veerapandian, M. Riley, R. F. Doolittle, *Blood* **90**, 1128 (Nov 15, 1997).
22. J. D. Ferry, *Proc Natl Acad Sci U S A* **38**, 566 (Jul, 1952).
23. P. Wiltzius, G. Dietler, W. Kanzig, A. Haberli, P. W. Straub, *Biopolymers* **21**, 2205 (1982).
24. E. Mihalyi, *Biochemistry* **27**, 967 (Feb 9, 1988).
25. O. V. Gorkun, A. H. Henschen-Edman, L. F. Ping, S. T. Lord, *Biochemistry* **37**, 15434 (Nov 3, 1998).
26. O. V. Gorkun, Y. I. Veklich, L. V. Medved, A. H. Henschen, J. W. Weisel, *Biochemistry* **33**, 6986 (Jun 7, 1994).
27. L. Purves, M. Purves, W. Brandt, *Biochemistry* **26**, 4640 (Jul 28, 1987).
28. M. Guthold, C. Carlisle, *Journal of Thrombosis and Haemostasis* **8**, 2090 (Sep, 2010).
29. Y. V. Matsuka, L. V. Medved, M. M. Migliorini, K. C. Ingham, *Biochemistry* **35**, 5810 (May 7, 1996).
30. J. H. Sobel, M. A. Gawinowicz, *The Journal of biological chemistry* **271**, 19288 (Aug 9, 1996).
31. J. W. Weisel, *J Thromb Haemost* **5**, 2340 (Dec, 2007).
32. M. W. Mosesson, *Journal of Thrombosis and Haemostasis* **2**, 388 (Mar, 2004).
33. E. B. Hunziker, P. W. Straub, A. Haeberli, *The Journal of biological chemistry* **265**, 7455 (May 5, 1990).
34. S. J. Everse, G. Spraggon, L. Veerapandian, M. Riley, R. F. Doolittle, *Biochemistry* **37**, 8637 (Jun 16, 1998).
35. K. R. Porter, C. V. Hawn, *Proc Soc Exp Biol Med* **65**, 309 (Jun, 1947).
36. C. E. Hall, *J Biol Chem* **179**, 857 (Jun, 1949).
37. C. E. Hall, *J Am Chem Soc* **71**, 1138 (Mar, 1949).
38. C. Cohen, H. Slayter, L. Goldstein, J. Kucera, C. Hall, *J Mol Biol* **22**, 385 (Dec 28, 1966).

39. J. Hermans, *Proc Natl Acad Sci U S A* **76**, 1189 (Mar, 1979).
40. R. R. Hantgan, J. Hermans, *J Biol Chem* **254**, 11272 (Nov 25, 1979).
41. J. W. Weisel, G. N. Phillips, Jr., C. Cohen, *Nature* **289**, 263 (Jan 22, 1981).
42. W. A. Voter, C. Lucaveche, A. E. Blaurock, H. P. Erickson, *Biopolymers* **25**, 2359 (Dec, 1986).
43. W. A. Voter, C. Lucaveche, H. P. Erickson, *Biopolymers* **25**, 2375 (Dec, 1986).
44. J. W. Weisel, *J Ultrastruct Mol Struct Res* **96**, 176 (Jul-Sep, 1986).
45. J. W. Weisel, C. Nagaswami, L. Makowski, *Proc Natl Acad Sci U S A* **84**, 8991 (Dec, 1987).
46. L. Medved, T. Ugarova, Y. Veklich, N. Lukinova, J. Weisel, *J Mol Biol* **216**, 503 (Dec 5, 1990).
47. Z. Yang, I. Mochalkin, R. F. Doolittle, *Proc Natl Acad Sci U S A* **97**, 14156 (Dec 19, 2000).
48. G. Caracciolo *et al.*, *Thromb Haemost* **89**, 632 (Apr, 2003).
49. E. T. O'Brien, 3rd *et al.*, *Proc Natl Acad Sci U S A* **105**, 19438 (Dec 9, 2008).
50. K. B. Neeves, D. A. Illing, S. L. Diamond, *Biophys J* **98**, 1344 (Apr 7, 2010).
51. E. G. Zavyalova, A. D. Protopopova, A. M. Kopylov, I. V. Yaminsky, *Langmuir* **27**, 4922 (Apr 19, 2011).
52. J. D. Ferry, P. R. Morrison, *J Clin Invest* **23**, 566 (Jul, 1944).
53. C. Gerth, W. W. Roberts, J. D. Ferry, *Biophys Chem* **2**, 208 (Oct, 1974).
54. W. W. Roberts, O. Kramer, R. W. Rosser, F. H. Nestler, J. D. Ferry, *Biophys Chem* **1**, 152 (Feb, 1974).
55. L. Lorand, *Nature* **166**, 694 (Oct 21, 1950).
56. W. W. Roberts, L. Lorand, L. F. Mockros, *Biorheology* **10**, 29 (Mar, 1973).
57. L. F. Mockros, W. W. Roberts, L. Lorand, *Biophys Chem* **2**, 164 (Aug, 1974).
58. H. Kang *et al.*, *J Phys Chem B* **113**, 3799 (Mar 26, 2009).

59. M. D. Bale, J. D. Ferry, *Thromb Res* **52**, 565 (Dec 15, 1988).
60. J. V. Shah, P. A. Janmey, *Rheologica Acta* **36**, 262 (May-Jun, 1997).
61. P. A. Janmey *et al.*, *Nat Mater* **6**, 48 (Jan, 2007).
62. E. A. Ryan, L. F. Mockros, A. M. Stern, L. Lorand, *Biophys J* **77**, 2827 (Nov, 1999).
63. E. A. Ryan, L. F. Mockros, J. W. Weisel, L. Lorand, *Biophys J* **77**, 2813 (Nov, 1999).
64. J. P. Collet, H. Shuman, R. E. Ledger, S. Lee, J. W. Weisel, *Proc Natl Acad Sci U S A* **102**, 9133 (Jun 28, 2005).
65. M. Guthold *et al.*, *Cell Biochem Biophys* **49**, 165 (2007).
66. W. Liu *et al.*, *Science* **313**, 634 (Aug 4, 2006).
67. W. Liu, C. R. Carlisle, E. A. Sparks, M. Guthold, *J Thromb Haemost* **8**, 1030 (May, 2010).
68. B. B. Lim, E. H. Lee, M. Sotomayor, K. Schulten, *Structure* **16**, 449 (Mar, 2008).
69. R. I. Litvinov, O. V. Gorkun, S. F. Owen, H. Shuman, J. W. Weisel, *Blood* **106**, 2944 (Nov 1, 2005).
70. R. I. Litvinov *et al.*, *Blood* **109**, 130 (Jan 1, 2007).
71. L. E. Averett *et al.*, *Langmuir* **24**, 4979 (May 6, 2008).
72. L. E. Averett, M. H. Schoenfisch, B. B. Akhremitchev, O. V. Gorkun, *Biophysical Journal* **97**, 2820 (Nov 18, 2009).
73. R. I. Litvinov *et al.*, *Biochemistry* **46**, 9133 (Aug 7, 2007).
74. A. E. Brown, R. I. Litvinov, D. E. Discher, P. K. Purohit, J. W. Weisel, *Science* **325**, 741 (Aug 7, 2009).
75. A. E. Brown, R. I. Litvinov, D. E. Discher, J. W. Weisel, *Biophys J* **92**, L39 (Mar 1, 2007).
76. J. W. Weisel, *Journal of Thrombosis and Haemostasis* **2**, 394 (Mar, 2004).
77. M. R. Falvo, D. Millard, E. T. O'Brien III, R. Superfine, S. T. Lord, *J Thromb Haemost* **6**, 1991 (2008).

78. J. R. Houser *et al.*, *Biophys J* **99**, 3038 (Nov 3, 2010).
79. J. D. Ferry, P. R. Morrison, *Journal of the American Chemical Society* **69**, 400 (1947).
80. C. Storm, J. J. Pastore, F. C. MacKintosh, T. C. Lubensky, P. A. Janmey, *Nature* **435**, 191 (May 12, 2005).
81. P. R. Onck, T. Koeman, T. van Dillen, E. van der Giessen, *Phys Rev Lett* **95**, 178102 (Oct 21, 2005).
82. I. K. Piechocka, R. G. Bacabac, M. Potters, F. C. Mackintosh, G. H. Koenderink, *Biophys J* **98**, 2281 (May 19, 2010).
83. A. E. X. Brown, R. I. Litvinov, D. E. Discher, P. K. Purohit, J. W. Weisel, *Science* **325**, 741 (Aug 7, 2009).
84. H. J. Qi, C. Ortiz, M. C. Boyce, *Journal of Engineering Materials and Technology-Transactions of the Asme* **128**, 509 (Oct, 2006).
85. N. E. Hudson *et al.*, *Biophys J* **98**, 1632 (Apr 21, 2010).
86. M. J. Buehler, *Proc Natl Acad Sci U S A* **103**, 12285 (Aug 15, 2006).
87. M. J. Buehler, Y. C. Yung, *Nat Mater* **8**, 175 (Mar, 2009).
88. J. Aizenberg *et al.*, *Science* **309**, 275 (Jul 8, 2005).
89. P. A. Janmey, J. P. Winer, J. W. Weisel, *J R Soc Interface* **6**, 1 (Jan 6, 2009).
90. N. E. Hudson *et al.*, *Biophysical Journal* **98**, 1632 (2010).
91. W. Liu, C. R. Carlisle, E. A. Sparks, M. Guthold, *J Thromb Haemost*, (2010).
92. J. Gosline *et al.*, *Philos Trans R Soc Lond B Biol Sci* **357**, 121 (Feb 28, 2002).
93. J. M. Gosline, P. A. Guerette, C. S. Ortlepp, K. N. Savage, *J Exp Biol* **202**, 3295 (Dec, 1999).
94. A. S. Tatham, P. R. Shewry, *Philos Trans R Soc Lond B Biol Sci* **357**, 229 (Feb 28, 2002).
95. M. R. Falvo, D. Millard, E. T. O'Brien, 3rd, R. Superfine, S. T. Lord, *J Thromb Haemost* **6**, 1991 (Nov, 2008).
96. C. G. Binnie, S. T. Lord, *Blood* **81**, 3186 (Jun 15, 1993).

97. O. V. Gorkun, Y. I. Veklich, J. W. Weisel, S. T. Lord, *Blood* **89**, 4407 (Jun 15, 1997).
98. M. Rief, J. Pascual, M. Saraste, H. E. Gaub, *J Mol Biol* **286**, 553 (Feb 19, 1999).
99. J. W. Weisel, Y. Veklich, O. Gorkun, *J Mol Biol* **232**, 285 (Jul 5, 1993).
100. M. S. Kellermayer, S. B. Smith, H. L. Granzier, C. Bustamante, *Science* **276**, 1112 (May 16, 1997).
101. A. F. Oberhauser, P. E. Marszalek, H. P. Erickson, J. M. Fernandez, *Nature* **393**, 181 (May 14, 1998).
102. M. Rief, M. Gautel, A. Schemmel, H. E. Gaub, *Biophys J* **75**, 3008 (Dec, 1998).
103. I. Schwaiger, C. Sattler, D. R. Hostetter, M. Rief, *Nat Mater* **1**, 232 (Dec, 2002).
104. C. Bustamante, J. F. Marko, E. D. Siggia, S. Smith, *Science* **265**, 1599 (Sep 9, 1994).
105. M. Rief, M. Gautel, F. Oesterhelt, J. M. Fernandez, H. E. Gaub, *Science* **276**, 1109 (May 16, 1997).
106. M. S. Kellermayer, C. Bustamante, H. L. Granzier, *Biochim Biophys Acta* **1604**, 105 (Jun 5, 2003).
107. M. D. Wang, H. Yin, R. Landick, J. Gelles, S. M. Block, *Biophysical Journal* **72**, 1335 (Mar, 1997).
108. G. W. Nelb, G. W. Kamykowski, J. D. Ferry, *The Journal of biological chemistry* **255**, 6398 (Jul 10, 1980).
109. K. F. Standeven *et al.*, *Blood* **110**, 902 (Aug 1, 2007).
110. J. F. Marko, E. D. Siggia, *Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics* **52**, 2912 (Sep, 1995).
111. J. Marko, S. E., *Macromolecules* **28**, 8759 (1995).
112. J. W. Hearle, *Int J Biol Macromol* **27**, 123 (Apr 12, 2000).
113. A. Miserez, S. S. Wasko, C. F. Carpenter, J. H. Waite, *Nat Mater* **8**, 910 (Nov, 2009).
114. Z. Qin, L. Kreplak, M. J. Buehler, *PloS one* **4**, e7294 (2009).

115. L. Kreplak, D. Fudge, *Bioessays* **29**, 26 (Jan, 2007).
116. S. Choe, S. X. Sun, *The Journal of chemical physics* **122**, 244912 (Jun 22, 2005).
117. S. Choe, S. X. Sun, *Biophys J* **92**, 1204 (Feb 15, 2007).
118. C. W. Wolgemuth, S. X. Sun, *Phys Rev Lett* **97**, 248101 (Dec 15, 2006).
119. D. S. Fudge, J. M. Gosline, *Proc Biol Sci* **271**, 291 (Feb 7, 2004).
120. D. S. Fudge, K. H. Gardner, V. T. Forsyth, C. Riekkel, J. M. Gosline, *Biophys J* **85**, 2015 (Sep, 2003).
121. J. Bertaud, J. Hester, D. D. Jimenez, M. J. Buehler, *Journal of Physics-Condensed Matter* **22**, (Jan 27, 2010).
122. M. Rief, M. Gautel, H. E. Gaub, *Adv Exp Med Biol* **481**, 129 (2000).
123. M. E. Carr, Jr., J. Hermans, *Macromolecules* **11**, 46 (Jan-Feb, 1978).
124. P. K. Purohit, R. I. Litvinov, A. E. Brown, D. E. Discher, J. W. Weisel, *Acta Biomater* **7**, 2374 (2011).
125. L. R. G. Treloar, *Transactions of the Faraday Society* **40**, 59 (1944).
126. H. C. Andersen, *Journal of Chemical Physics* **72**, 2384 (1980).
127. W. T. Astbury, H. J. Woods, *Philos Trans R Soc Lond A Math or Phys Sci* **232**, 333 (1934).
128. L. Kreplak, J. Doucet, P. Dumas, F. Briki, *Biophys J* **87**, 640 (Jul, 2004).
129. L. Kreplak, J. Doucet, F. Briki, *Biopolymers* **58**, 526 (Apr 15, 2001).
130. Z. Qin, M. J. Buehler, *Physical review letters* **104**, 198304 (May 14, 2010).
131. D. S. Talaga *et al.*, *Proc Natl Acad Sci U S A* **97**, 13021 (Nov 21, 2000).
132. T. Bornschlogl, M. Rief, *Physical review letters* **96**, 118102 (Mar 24, 2006).
133. I. Schwaiger, M. Schleicher, A. A. Noegel, M. Rief, *EMBO Rep* **6**, 46 (Jan, 2005).
134. B. B. Aaron, J. M. Gosline, *Biopolymers* **20**, 1247 (1981).
135. R. Vermorel, N. Vandenberghe, E. Villermaux, *Proceedings of the Royal Society A-Mathematical Physical and Engineering Sciences* **463**, 641 (2007).

136. C. M. Elvin *et al.*, *Nature* **437**, 999 (Oct 13, 2005).
137. N. Becker *et al.*, *Nat Mater* **2**, 278 (Apr, 2003).
138. S. Keten, M. J. Buehler, *J R Soc Interface* **7**, 1709 (Dec 6, 2010).
139. S. Keten, Z. Xu, B. Ihle, M. J. Buehler, *Nat Mater* **9**, 359 (Apr, 2010).
140. A. Roy, A. Kucukural, Y. Zhang, *Nat Protoc* **5**, 725 (2010).
141. J. P. Collet *et al.*, *Blood* **106**, 3824 (Dec 1, 2005).
142. R. F. Doolittle, *Adv Protein Chem* **27**, 1 (1973).
143. P. L. Privalov, L. V. Medved, *J Mol Biol* **159**, 665 (Aug 25, 1982).
144. J. W. Weisel, L. Medved, *Ann N Y Acad Sci* **936**, 312 (2001).
145. G. Tsurupa, L. Tsonev, L. Medved, *Biochemistry* **41**, 6449 (May 21, 2002).
146. Budzynsk.Az, *Biochimica Et Biophysica Acta* **229**, 663 (1971).
147. M. Murakawa *et al.*, *Thrombosis and Haemostasis* **69**, 351 (Apr 1, 1993).
148. H. P. Erickson, W. E. Fowler, *Annals of the New York Academy of Sciences* **408**, 146 (1983).
149. L. V. Medved, O. V. Gorkun, P. L. Privalov, *Febs Letters* **160**, 291 (1983).
150. J. W. Weisel, C. V. Stauffacher, E. Bullitt, C. Cohen, *Science* **230**, 1388 (Dec 20, 1985).
151. Y. I. Veklich, O. V. Gorkun, L. V. Medved, W. Nieuwenhuizen, J. W. Weisel, *J Biol Chem* **268**, 13577 (Jun 25, 1993).
152. R. A. Burton, G. Tsurupa, R. R. Hantgan, N. Tjandra, L. Medved, *Biochemistry* **46**, 8550 (Jul 24, 2007).
153. R. A. Burton, G. Tsurupa, L. Medved, N. Tjandra, *Biochemistry* **46**, 6674 (Jun 5, 2007).
154. G. Tsurupa *et al.*, *Biophysical Chemistry* **112**, 257 (Dec 20, 2004).
155. B. A. Cottrell, D. D. Strong, K. W. Watt, R. F. Doolittle, *Biochemistry* **18**, 5405 (Nov 27, 1979).

156. Y. Zhang, *Proteins-Structure Function and Bioinformatics* **69**, 108 (2007).
157. Y. Zhang, *Bmc Bioinformatics* **9**, (Jan 23, 2008).
158. F. Ding, N. V. Dokholyan, *Proc Natl Acad Sci U S A* **105**, 19696 (Dec 16, 2008).
159. F. Ding, D. Tsao, H. F. Nie, N. V. Dokholyan, *Structure* **16**, 1010 (Jul, 2008).
160. L. Medved, J. W. Weisel, *J Thromb Haemost* **7**, 355 (Feb, 2009).
161. S. Kumar, D. Bouzida, R. H. Swendsen, P. A. Kollman, J. M. Rosenberg, *Journal of Computational Chemistry* **13**, 1011 (Oct, 1992).
162. Z. Yang, J. M. Kollman, L. Pandi, R. F. Doolittle, *Biochemistry* **40**, 12515 (Oct 23, 2001).
163. L. V. Medved, O. V. Gorkun, P. L. Privalov, *FEBS Lett* **160**, 291 (Aug 22, 1983).
164. L. Medved, W. Nieuwenhuizen, *Thromb Haemost* **89**, 409 (Mar, 2003).
165. G. Tsurupa, S. Yakovlev, P. McKee, L. Medved, *Biochemistry* **49**, 7643 (Sep 7, 2010).
166. E. Makogonenko, G. Tsurupa, K. Ingham, L. Medved, *Biochemistry* **41**, 7907 (Jun 25, 2002).
167. K. A. Smith *et al.*, *Blood*, (Jan 11, 2011).
168. M. Rubinstein, R. H. Colby, *Polymer physics*. (Oxford University Press, Oxford ; New York, 2003), pp. xi, 440 p.
169. L. Betts, B. K. Merenbloom, S. T. Lord, *J Thromb Haemost* **4**, 1139 (May, 2006).
170. G. I. Bell, *Science* **200**, 618 (1978).
171. J. Janin, S. Miller, C. Chothia, *J Mol Biol* **204**, 155 (Nov 5, 1988).
172. B. A. Kesner, F. Ding, B. R. Temple, N. V. Dokholyan, *Proteins-Structure Function and Bioinformatics* **78**, 12 (Jan, 2010).
173. N. D. Socci, J. N. Onuchic, P. G. Wolynes, *Journal of Chemical Physics* **104**, 5860 (Apr 15, 1996).
174. P. A. Janmey, E. J. Amis, J. D. Ferry, *J Rheol* **27**, 135 (1983).
175. M. L. Gardel *et al.*, *Science* **304**, 1301 (May 28, 2004).

176. P. A. Janmey, U. Euteneuer, P. Traub, M. Schliwa, *J Cell Biol* **113**, 155 (Apr, 1991).
177. J. V. Shah, P. A. Janmey, *Rheol Acta* **36**, 262 (1997).
178. Z. Wen, A. Basu, J. Winer, A. Yodh, P. A. Janmey, *New Journal of Physics*, (2007).
179. J. Xu, Y. Tseng, D. Wirtz, *J Biol Chem* **275**, 35886 (Nov 17, 2000).
180. H. Kang *et al.*, *J Phys Chem B* **113**, 3799 (Mar 26, 2009).
181. C. Heussinger, E. Frey, *Phys Rev Lett* **97**, 105501 (Sep 8, 2006).
182. M. Guthold *et al.*, *Biophys J* **87**, 4226 (Dec, 2004).
183. C. G. Binnie, J. M. Hettasch, E. Strickland, S. T. Lord, *Biochemistry* **32**, 107 (Jan 12, 1993).
184. R. F. Doolittle, *Biochem J* **94**, 735 (Mar, 1965).
185. J. S. Palmer, M. C. Boyce, *Acta Biomater* **4**, 597 (May, 2008).
186. Y. Cadroy, T. A. Horbett, S. R. Hanson, *Journal of Laboratory and Clinical Medicine* **113**, 436 (Apr, 1989).
187. K. S. Sakariassen *et al.*, *Arteriosclerosis* **10**, 276 (Mar-Apr, 1990).
188. J. Strony, A. Beaudoin, D. Brands, B. Adelman, *Am J Physiol* **265**, H1787 (Nov, 1993).
189. D. Weaire, R. Phelan, *Nature* **367**, 123 (Jan 13, 1994).
190. K. C. Gersh, K. E. Edmondson, J. W. Weisel, *Journal of Thrombosis and Haemostasis* **8**, 2826 (Dec, 2010).
191. R. A. Campbell, M. M. Aleman, L. D. Gray, M. R. Falvo, A. S. Wolberg, *Thrombosis and Haemostasis* **104**, 1281 (Dec 10, 2010).
192. M. M. Rooney, L. V. Parise, S. T. Lord, *The Journal of biological chemistry* **271**, 8553 (Apr 12, 1996).
193. I. Cohen, J. M. Gerrard, J. G. White, *J Cell Biol* **93**, 775 (Jun, 1982).
194. A. V. Cooper, K. F. Standeven, R. A. S. Ariens, *Blood* **102**, 535 (Jul 15, 2003).

195. Y. P. Fu, G. Grieninger, *Proceedings of the National Academy of Sciences of the United States of America* **91**, 2625 (Mar 29, 1994).
196. Y. P. Fu, J. Z. Zhang, C. M. Redman, G. Grieninger, *Blood* **92**, 3302 (Nov 1, 1998).
197. F. Terasawa *et al.*, *Blood* **94**, 4122 (Dec 15, 1999).
198. S. O. Brennan, A. P. Fellowes, J. M. Faed, P. M. George, *Blood* **95**, 1709 (Mar 1, 2000).
199. P. Jing, J. S. Rudra, A. B. Herr, J. H. Collier, *Biomacromolecules* **9**, 2438 (Sep, 2008).
200. C. Cohen, D. A. Parry, *Proteins* **7**, 1 (1990).
201. A. Lupas, M. Van Dyke, J. Stock, *Science* **252**, 1162 (May 24, 1991).
202. R. F. Doolittle, D. M. Goldbaum, L. R. Doolittle, *Journal of Molecular Biology* **120**, 311 (1978).
203. R. A. Crowther, L. A. Amos, J. T. Finch, D. J. Derosier, A. Klug, *Nature* **226**, 421 (1970).
204. D. Nicastro *et al.*, *Science* **313**, 944 (Aug 18, 2006).
205. M. F. Schmid, M. B. Sherman, P. Matsudaira, W. Chiu, *Nature* **431**, 104 (Sep 2, 2004).
206. A. Yildiz *et al.*, *Science* **300**, 2061 (Jun 27, 2003).
207. T. Ha *et al.*, *Proceedings of the National Academy of Sciences of the United States of America* **93**, 6264 (Jun 25, 1996).
208. P. R. Selvin *et al.*, *Biophysical Journal* **70**, Wp302 (Feb, 1996).
209. X. H. Qu, D. Wu, L. Mets, N. F. Scherer, *Proceedings of the National Academy of Sciences of the United States of America* **101**, 11298 (Aug 3, 2004).
210. T. Forster, in *Modern Quantum Chemistry, Istanbul Lectures*, O. Sinanoglu, Ed. (Academic Press, New York and London, 1965), pp. 93-137.
211. E. Klotzsch *et al.*, *Proceedings of the National Academy of Sciences of the United States of America* **106**, 18267 (Oct 27, 2009).

212. J. Y. Yang, W. Y. Yang, *Journal of the American Chemical Society* **131**, 11644 (Aug 26, 2009).
213. R. Berkovich, S. Garcia-Manyes, M. Urbakh, J. Klafter, J. M. Fernandez, *Biophysical Journal* **98**, 2692 (Jun 2, 2010).
214. Y. Taniguchi, B. S. Khatri, D. J. Brockwell, E. Paci, M. Kawakami, *Biophysical Journal* **99**, 257 (Jul 7, 2010).
215. M. Kawakami, K. Byrne, B. S. Khatri, T. C. B. McLeish, D. A. Smith, *Chemphyschem* **7**, 1710 (Aug 11, 2006).
216. M. Kawakami, K. Byrne, D. J. Brockwell, S. E. Radford, D. A. Smith, *Biophysical Journal* **91**, L16 (Jul 15, 2006).
217. B. S. Khatri *et al.*, *Faraday Discussions* **139**, 35 (2008).
218. F. Ding, S. V. Buldyrev, N. V. Dokholyan, *Biophysical Journal* **88**, 147 (Jan, 2005).
219. S. Peng *et al.*, *Phys Rev E Stat Nonlin Soft Matter Phys* **69**, 041908 (Apr, 2004).
220. L. Monticelli *et al.*, *J. Chem. Theory and Comput.* **4**, 819 (2008).
221. A. F. Smeijers, K. Pieterse, A. J. Markvoort, P. A. Hilbers, *J Phys Chem B* **110**, 13614 (Jul 13, 2006).
222. H. Puchtler, F. S. Waldrop, S. N. Meloan, *Appl Pathol* **3**, 5 (1985).
223. D. C. Rapaport, *The art of molecular dynamics simulation.* (Cambridge University Press, Cambridge, UK ; New York, NY, ed. 2nd, 2004), pp. xiii, 549 p.
224. B. J. Alder, T. E. Wainwright, *Journal of Chemical Physics* **27**, 1208 (1957).
225. B. J. Alder, T. E. Wainwright, *Journal of Chemical Physics* **31**, 459 (1959).
226. Stilling.Fh, A. Rahman, *Journal of Chemical Physics* **60**, 1545 (1974).
227. M. Levitt, A. Warshel, *Nature* **253**, 694 (1975).
228. A. Cooper, *Proceedings of the National Academy of Sciences of the United States of America* **73**, 2740 (1976).
229. J. A. Mccammon, B. R. Gelin, M. Karplus, *Nature* **267**, 585 (1977).

230. O. V. Galzitskaya, A. V. Finkelstein, *Proceedings of the National Academy of Sciences of the United States of America* **96**, 11299 (Sep 28, 1999).
231. E. Alm, D. Baker, *Proceedings of the National Academy of Sciences of the United States of America* **96**, 11305 (Sep 28, 1999).
232. B. A. Shoemaker, J. Wang, P. G. Wolynes, *Proceedings of the National Academy of Sciences of the United States of America* **94**, 777 (Feb 4, 1997).
233. J. J. Portman, S. Takada, P. G. Wolynes, *Phys Rev Lett* **81**, 5237 (Dec 7, 1998).
234. D. A. Case, M. Karplus, *Journal of Molecular Biology* **132**, 343 (1979).
235. B. Brooks, M. Karplus, *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences* **80**, 6571 (1983).
236. H. Lu, B. Isralewitz, A. Krammer, V. Vogel, K. Schulten, *Biophysical Journal* **75**, 662 (Aug, 1998).
237. S. Berneche, B. Roux, *Nature* **414**, 73 (Nov 01, 2001).
238. D. Rothlisberger *et al.*, *Nature* **453**, 190 (May 8, 2008).
239. W. C. Swope, H. C. Andersen, P. H. Berens, K. R. Wilson, *Journal of Chemical Physics* **76**, 637 (1982).
240. L. Verlet, *Physical Review* **159**, 98 (1967).
241. R. Stone, A. Dejaegere, D. Kuznetsov, L. Falquet, R. Stone, Ed. http://www.ch.embnet.org/MD_tutorial/ (1999).
242. E. H. Lee, J. Hsin, O. Mayans, K. Schulten, *Biophys J* **93**, 1719 (Sep 1, 2007).
243. H.-D. Höltje, *Molecular modeling : basic principles and applications*. (Wiley-VCH, Weinheim, ed. 3rd, rev. and expanded, 2008), pp. x, 310 p.
244. R. Luo, J. Wang, C. H. Tan, Y. H. Tan, Q. Lu, *Communications in Computational Physics* **3**, 1010 (May, 2008).
245. D. Eisenberg, A. D. McLachlan, *Nature* **319**, 199 (Jan 16, 1986).
246. T. Lazaridis, M. Karplus, *Proteins-Structure Function and Genetics* **35**, 133 (May 1, 1999).
247. T. Kozaki, K. Morihashi, O. Kikuchi, *Journal of the American Chemical Society* **111**, 1547 (Mar 1, 1989).

248. A. Onufriev, D. A. Case, D. Bashford, *Journal of Computational Chemistry* **23**, 1297 (Nov 15, 2002).
249. A. Onufriev, J. Mongan, W. A. Svrcek-Seiler, *Journal of Chemical Physics* **127**, (Nov 14, 2007).
250. W. C. Still, A. Tempczyk, R. C. Hawley, T. Hendrickson, *Journal of the American Chemical Society* **112**, 6127 (Aug 1, 1990).
251. D. V. Schroeder, *An introduction to thermal physics*. (Addison Wesley, San Francisco, CA, 2000), pp. x, 422 p.
252. K. Huang, *Statistical mechanics*. (Wiley, New York, ed. 2nd, 1987), pp. xiv, 493 p.
253. W. W. Wood, J. J. Erpenbeck, G. A. Baker, J. D. Johnson, *Physical Review E* **6301**, (Jan, 2001).
254. C. R. Oliveira, T. Werlang, *Revista Brasileira De Ensino De Fisica* **29**, 189 (Apr-Jun, 2007).
255. A. Patrascioiu, *Los Alamos Science* **15**, 17 (1987).
256. L. V. Woodcock, *Chemical Physics Letters* **10**, 257 (1971).
257. H. J. C. Berendsen, J. P. M. Postma, W. F. Vangunsteren, A. Dinola, J. R. Haak, *Journal of Chemical Physics* **81**, 3684 (1984).
258. P. Hunenberger, *Advanced Computer Simulation Approaches for Soft Matter Sciences I* **173**, 105 (2005).
259. W. G. Hoover, *Physical Review A* **31**, 1695 (1985).
260. S. Nose, *Journal of Chemical Physics* **81**, 511 (1984).
261. T. Schneider, E. Stoll, *Physical Review B* **17**, 1302 (1978).
262. D. D. Humphreys, R. A. Friesner, B. J. Berne, *Journal of Physical Chemistry* **98**, 6885 (Jul 7, 1994).
263. F. Ding, N. V. Dokholyan, *Trends in Biotechnology* **23**, 450 (Sep, 2005).
264. N. V. Dokholyan, *Current Opinion in Structural Biology* **16**, 79 (Feb, 2006).
265. N. Go, H. Abe, *Biopolymers* **20**, 991 (1981).

266. R. D. S. Dixon *et al.*, *Structure* **12**, 2161 (Dec, 2004).
267. R. Miller *et al.*, *Journal of Chemical Physics* **96**, 768 (Jan 1, 1992).
268. R. Srinivasan, G. D. Rose, *Proceedings of the National Academy of Sciences of the United States of America* **96**, 14258 (Dec 7, 1999).
269. P. Bradley, K. M. S. Misura, D. Baker, *Science* **309**, 1868 (Sep 16, 2005).
270. N. Go, *Annual Review of Biophysics and Bioengineering* **12**, 183 (1983).
271. S. Takada, *Proceedings of the National Academy of Sciences of the United States of America* **96**, 11698 (Oct 12, 1999).
272. Y. Zhou, M. Karplus, *Proc Natl Acad Sci U S A* **94**, 14429 (Dec 23, 1997).
273. N. V. Dokholyan, S. V. Buldyrev, H. E. Stanley, E. I. Shakhnovich, *Fold Des* **3**, 577 (1998).
274. A. Emperador, T. Meyer, M. Orozco, *Proteins-Structure Function and Bioinformatics* **78**, 83 (Jan, 2010).
275. F. Ding, J. M. Borreguero, S. V. Buldyrey, H. E. Stanley, N. V. Dokholyan, *Proteins-Structure Function and Genetics* **53**, 220 (Nov 1, 2003).
276. J. G. Kirkwood, *The Journal of chemical physics* **3**, 14 (1935).
277. P. Kollman, *Chemical Reviews* **93**, 2395 (Nov, 1993).
278. M. Souaille, B. Roux, *Computer Physics Communications* **135**, 40 (Mar 15, 2001).
279. Y. Sugita, Y. Okamoto, *Chemical Physics Letters* **314**, 141 (Nov 26, 1999).
280. R. H. Swendsen, J. S. Wang, *Phys Rev Lett* **57**, 2607 (Nov 24, 1986).
281. Y. Sugita, A. Kitao, Y. Okamoto, *Journal of Chemical Physics* **113**, 6042 (Oct 15, 2000).
282. C. Bartels, M. Karplus, *Journal of Computational Chemistry* **18**, 1450 (Sep, 1997).
283. E. M. Boczko, C. L. Brooks, *Journal of Physical Chemistry* **97**, 4509 (Apr 29, 1993).

284. E. Gallicchio, M. Andrec, A. K. Felts, R. M. Levy, *Journal of Physical Chemistry B* **109**, 6722 (Apr 14, 2005).
285. E. T. Jaynes, G. L. Bretthorst, *Probability Theory: the Logic of Science*. (Cambridge University Press, Cambridge, UK, 2003), vol. 1.
286. M. S. Z. Kellermayer, L. Grama, *Journal of Muscle Research and Cell Motility* **23**, 499 (2002).
287. S. Izrailev, S. Stepaniants, M. Balsera, Y. Oono, K. Schulten, *Biophysical Journal* **72**, 1568 (Apr, 1997).
288. M. Chabria, S. Hertig, M. L. Smith, V. Vogel, *Nat Commun* **1**, 135 (Dec 7, 2010).
289. M. W. Davidson, M. Abramowitz, in *Encyclopedia of Imagin Science and technology* J. P. Hornak, Ed. (J. Wiley, New York, 2002).
290. H. G. Kapitzka, *Carl Zeiss*, 40 (1994).
291. M. Abramowitz, K. R. Spring, H. E. Keller, M. W. Davidson, *Biotechniques* **33**, 772 (Oct, 2002).
292. C. G. Galbraith, R. Skalak, S. Chien, *Cell Motil Cytoskeleton* **40**, 317 (1998).
293. S. R. Quake, H. Babcock, S. Chu, *Nature* **388**, 151 (Jul 10, 1997).
294. I. Spector, N. R. Shochet, D. Blasberger, Y. Kashman, *Cell Motil Cytoskeleton* **13**, 127 (1989).
295. J. W. Lichtman, J. A. Conchello, *Nat Methods* **2**, 910 (Dec, 2005).
296. B. Valeur, *Molecular Fluorescence: Principles and Applications*. (Wiley-VCH, Weinheim, 2002).
297. K. Suhling, P. M. French, D. Phillips, *Photochem Photobiol Sci* **4**, 13 (Jan, 2005).
298. K. Suhling *et al.*, *Opt Lett* **29**, 584 (Mar 15, 2004).
299. C. A. Parker, W. T. Rees, *Analyst* **85**, 587 (1960).
300. C. J. R. Sheppard, T. Wilson, *Journal of Microscopy-Oxford* **124**, 107 (1981).
301. C. J. R. Sheppard, T. Wilson, *Optik* **55**, 331 (1980).
302. W. Denk, J. H. Strickler, W. W. Webb, *Science* **248**, 73 (Apr 6, 1990).

303. J. H. Strickler, W. Denk, W. W. Webb, *Biophysical Journal* **57**, A374 (Feb, 1990).
304. D. Axelrod, *Traffic* **2**, 764 (Nov, 2001).
305. X. W. Zhuang *et al.*, *Science* **288**, 2048 (Jun 16, 2000).
306. M. Kramer, C. Zeiss, *Photonik*, (2004).
307. G. Binnig, C. F. Quate, C. Gerber, *Phys Rev Lett* **56**, 930 (Mar 3, 1986).
308. R. J. Colton, D. R. Baselt, Y. F. Dufrene, J. B. Green, G. U. Lee, *Curr Opin Chem Biol* **1**, 370 (Oct, 1997).
309. J. Preiner, J. L. Tang, V. Pastushenko, P. Hinterdorfer, *Phys Rev Lett* **99**, (Jul 27, 2007).
310. R. Garcia, R. Perez, *Surface Science Reports* **47**, 197 (2002).
311. N. Jalili, K. Laxminarayana, *Mechatronics* **14**, 907 (Oct, 2004).
312. P. L. T. M. Frederix, P. D. Bosshart, A. Engel, *Biophysical Journal* **96**, 329 (Jan 21, 2009).
313. J. E. Sader, *Review of Scientific Instruments* **66**, 4583 (Sep, 1995).
314. J. E. Sader, I. Larson, P. Mulvaney, L. R. White, *Review of Scientific Instruments* **66**, 3789 (Jul, 1995).
315. J. E. Sader, J. W. M. Chon, P. Mulvaney, *Review of Scientific Instruments* **70**, 3967 (Oct, 1999).
316. M. Varenberg, I. Etsion, G. Halperin, *Review of Scientific Instruments* **74**, 3362 (Jul, 2003).
317. W. H. Liu, K. Bonin, M. Guthold, *Review of Scientific Instruments* **78**, (Jun, 2007).
318. D. F. Ogletree, R. W. Carpick, M. Salmeron, *Review of Scientific Instruments* **67**, 3298 (Sep, 1996).
319. H. Xie, J. Vitard, S. Haliyo, S. Regnier, M. Boukallel, *Review of Scientific Instruments* **79**, (Mar, 2008).
320. K. Wagner, P. Cheng, D. Vezenov, *Langmuir* **27**, 4635 (Apr 19, 2011).

321. J. L. Hutter, J. Bechhoefer, *Review of Scientific Instruments* **64**, 1868 (Jul, 1993).
322. D. A. Walters *et al.*, *Review of Scientific Instruments* **67**, 3583 (Oct, 1996).
323. J. L. Hutter, *Langmuir* **21**, 2630 (Mar 15, 2005).
324. C. P. Green *et al.*, *Review of Scientific Instruments* **75**, 1988 (Jun, 2004).
325. J. P. Cleveland, S. Manne, D. Bocek, P. K. Hansma, *Review of Scientific Instruments* **64**, 403 (Feb, 1993).