POLITECNICO DI MILANO

SCUOLA DI INGEGNERIA INDUSTRIALE E DELL'INFORMAZIONE

Corso di Laurea in Ingegneria Biomedica

Study of the AGEs effect on the mechanical properties of collagen fibrils through molecular models and experimental validation

Relatore: Prof. SIMONE VESENTINI

Correlatori: Dott. ALFONSO GAUTIERI

Prof. JESS GERRIT SNEDEKER

Tesi di laurea di:

LAURA BERNARDI

Matr. 781281

FEDERICA CRIPPA

Matr. 782047

Anno Accademico 2012 - 2013

I computer sono incredibilmente veloci, accurati e stupidi.
Gli uomini sono incredibilmente lenti, inaccurati e intelligenti.
L'insieme dei due costituisce una forza incalcolabile.
Albert Einstein

Contents

Abstract **xv**

Sommario **xxi**

1 INTRODUCTION **1**

1.1 Aim of the work 1

1.2 Collagen 2

1.2.1 Collagen structure 2

1.2.2 Enzymatic cross-links 3

1.2.3 Non-enzymatic cross-links 5

1.3 Fibrillogenesis 7

1.4 Mechanical tests on single fibrils 10

1.5 Computational models of collagen fibrils 12

2 MATERIAL AND METHODS **15**

2.1 Molecular Dynamics 15

2.1.1 Introduction 15

2.1.2 Basics of atomistic modeling: structure of MD Program . 16

2.1.3 Reactive Force Field and Non Reactive Force Field 19

2.1.4 Fibril implementation in LAMMPS 21

2.1.5 Mechanical characterization of enzymatic cross-links . . . 23

2.1.6 Mechanical characterization of non-enzymatic cross-links 26

2.1.7 Mechanical characterization of a tropocollagen molecule . 27

2.1.8 Lennard Jones parameters 30

2.1.9 Physiological fibril 31

2.1.10 Aged fibril 35
2.2 Experimental Validation 39
2.2.1 Fibrillation 39
2.2.2 Fibrils isolation from rat tail tendon 41
2.2.3 Sample preparation for SEM analysis 42
2.2.4 Fibrils visualization under light microscope 44
2.2.5 Set-up development for the pulling tests 45
2.2.6 Mechanical tests 48

3 RESULTS 51
3.1 Molecular Dynamics 51
3.1.1 Mechanical characterization of enzymatic cross-links . . . 51
3.1.2 Mechanical characterization of non-enzymatic cross-links 54
3.1.3 Mechanical characterization of a tropocollagen molecule . 57
3.1.4 Lennard Jones parameters 61
3.1.5 Physiological fibril 62
3.1.6 Aged fibril 66
3.2 Experimental Validation 70
3.2.1 Fibrillation 70
3.2.2 Fibrils isolation from rat tail tendon 70
3.2.3 Fibrils observation under light microscope 72
3.2.4 Set-up development for the pulling tests 75
3.2.5 Mechanical tests 76

4 DISCUSSIONS 83
4.1 Computational Model 83
4.2 Experimental Validation 87

5 CONCLUSIONS 91

A Characterization of cross-links 93

B C++ scripts 107

C LAMMPS input files 115

Bibliography **125**

List of Figures

1.1 Hierarchical structure of collagen, Buehler [6]. 2
1.2 Chemical pathway of crosslinking interactions, Eyre et al. [8]. . . 4
1.3 Sites of enzyymatic cross-links formation, Eyre et al. [8]. 5
1.4 Effect of non-enymatic cross-links on the physical properties of collagen, Avery et al. [9]. 6
1.5 Glucosepane. 6
1.6 Synthesis of fibrillar collagens, Kadler et al. [12]. 7
1.7 Typical turbidity curve, Williams et al. [13]. 9
1.8 Schematic view of the AFM tests, Yang et al. [20]. 11
1.9 SEM image of the fibril breakage with MEMS, Shen et Al. [21]. . 12
1.10 Schematic representation of fibril (a) and the model where cross-links are represented increasing the adhesion on the tropocollagen edges (b). Adapted from Buehler Markus J. [23]. 13
1.11 Schematic representation of collagen fibril (a) and the viscoelastic model based on Kelvin-Voigt units (b). Adapted from Gautieri et al. [24]. 14

2.1 Schematic structure of MD program. 16
2.2 Lennard-Jones potential. 18
2.3 Buehler [6], Non Reactive Force Field versus Reactive Force Field. 20
2.4 Buehler [6], modeling fracture of a single tropocollagen. 20
2.5 Force - Length chart for alanine pulling at 1 m/s and 100 m/s. . . . 23
2.6 Hydroxylysyl pyridinoline, pulling directions: 2-12, 2-7, 7-12. . . 24
2.7 Input script. 25

2.8 Glucosepane, pulling direction 1-22. 26
2.9 Coarse-grained model of a peptide. 27
2.10 Full atomistic simulation of a peptide. 28
2.11 Script for the peptide simulation. 28
2.12 Script for the tropocollagen simulation. 30
2.13 Fibril model for the Lennard Jones parameters calculation: a) longitudinal view b) cross section. 31
2.14 Pseudo physiological fibril: one edge of itl is pulled, the other one is mantained fixed. 32
2.15 Input file for the pseudo physiological fibril, the first group of atoms (my pull) represents one edge of the fibril and it is pulled at 100 m/s, the other edge (myfix) is mantained fixed. 32
2.16 Atomistic vs coarse-grained model. 33
2.17 C terminus configuration, Uzel et al. [33]. 33
2.18 Input script for enzymatic bonds: the first bond coeff represents bond interactions in the tropocollagen chains, bond coeff 2 describes enzymatic cross-link bonds. 34
2.19 Schematic representation of force versus length behavior of enzym cross-link: The force is mantained null untill the molecule reaches a length of 4.1 nm, then it increases (spring constant = 620.7 kcal/molA^2). When F = Fmax the bond breaks and the force is null again. 34
2.20 Non enzymatic crosslinks present in the whole section of the fibril. 36
2.21 Input script for bonds type *3*. 37
2.22 Determination of the centre coordinates of the fibril using VMD. . 38
2.23 Non-enzymatic cross-links in the external shell of the fibril. 39
2.24 Fibrillation set-up. 40
2.25 Incubation overnight in Eppendorf tubes. 40
2.26 Fascicles extracted from rat tail tendon and stored in PBS. 41
2.27 Procedure used to extract fibrils from fascicles. 41
2.28 Sonication procedure applied to cleaning glasses. 42
2.29 Plasma treatment. 43
2.30 Coating of samples in the sputtering machine. 43

2.31 A,B: Two different views of FemtoTool system equipped with the needle-sphere. C: The FemtoTool is integrated with the inverted microscope. 45
2.32 A: the needle (black shadow) is drived above the spheres placed on the sample; B: the sphere is attached to the needle. 46
2.33 Both edges of the fibril are glued to the substate. 47
2.34 Both edges of the fibril are glued to the substate and the needle-sphere component (black shadow) is glued to one edge. 47
2.35 A: the real probe, B: new arm configuration. 48
2.36 Fibrils: A-B physiological fibrils, C-D aged fibrils. 49

3.1 Force - Length chart of hydroxylysyl pyridinoline streched in 2-12 direction at 1 m/s. 52
3.2 Force - Length chart of hydroxylysyl pyridinoline streched in 2-12 direction at 1 m/s, the linear interpolation provides the spring constant. 53
3.3 Force - Length chart of glucosepane (structure 1) streched at 1 m/s. 54
3.4 Force-length chart of glucosepane (structure 1) streched at 1 m/s, the spring constant is evaluated using the linear interpolation tool. 56
3.5 Adapted from Grandbois et al.[36], stretching of a single polysaccharide chain that is attached to the tip and the substrate. 57
3.6 Force - Length curve for the coarse-grained model. 58
3.7 Comparison between the two models. 58
3.8 Input script used for the persistence length calculation. 59
3.9 Curve obtained elaborating the values from matlab. 59
3.10 Comparison between the coarse-grained and the full atomistic models for the tropocollagen. 60
3.11 Calculation of the persistence length. 61
3.12 Input script for the Lennard Jones simulation 61
3.13 Screenshot of the simulation. Note that the distance among the molecules is the expected one. 62
3.14 Force - Length chart for the pseudo physiological fibril stretched at 100 m/s. 63

3.15 VMD image of the rupture of pseudo physiological fibril. 63

3.16 Stress - Strain chart for pseudo physiological fibril stretched at 100 m/s. The elastic modulus is evaluated using the linear interpolation tool. 64

3.17 Force - Length chart for physiological fibril stretched at 100 m/s. . 65

3.18 Stress - Strain chart for physiological fibril stretched at 100 m/s. . 66

3.19 Force - Length chart for aged fibril with cross-links through the whole section, stretched at 100 m/s. 67

3.20 Stress - Strain chart for aged fibril with cross-links through the whole section, stretched at 100 m/s. 68

3.21 Force - Length chart for aged fibril with cross-links in the external shell, stretched at 100 m/s. 69

3.22 Stress - Strain chart for aged fibril with cross-links in the external shell, stretched at 100 m/s. 69

3.23 SEM images of self-assembled fibrils. A: on the right edge a microfibril that does not complete the fibrillation process can be visualized. B: diameter evaluation. 70

3.24 Fibrils extracted from rat tail tendon. A: the fibril has a diameter of $\sim$192 nm and a D-banding of $\sim$67 nm. B: The fibril has a length of $\sim$16μm. C: The fibril has a diameter of $\sim$212 nm and a D-banding of $\sim$68 nm. 71

3.25 Fluorescent fibrils extracted from rat tail tendon. 72

3.26 Dark field microscope used to visualize the fibrils from rat tail tendon. 73

3.27 Fibrils visualized using the inverted microscope. 74

3.28 A fibril aggregate visualized using (A) dark field, (B) bright field in an inverted microscope, (C) SEM. Higer magnification of (C) where sigle fibrils are shown, (c1) and (c2). 74

3.29 Fibril detachment from the substrate. In all the images the focus is mantained on the substrate to better visualize the detachment of the fibril. A: the tip is still in contact with the substrate and the whole fibril lays on it, B-E: the tip is raised and the fibril gradually loses contact with the silicon substrate, F: the fibril is attached to the substrate only by the drop of glue. 75
3.30 Fibril after failure visualized under inverted microscope. 76
3.31 Force - Length chart for fibril A. 77
3.32 Stress - Strain chart for fibril A. 78
3.33 Force - Length chart for fibril B. 78
3.34 Stress - Strain chart for fibril B. 79
3.35 Force - Length chart for fibril C. 80
3.36 Stress - Strain chart for fibril C. 80
3.37 Force - Length chart for fibril D. 81
3.38 Stress - Strain chart for fibril D. 81

4.1 Stress - Strain curves comparison. 84
4.2 Force - Length curves comparison. 88
4.3 Stress - Strain curves comparison. 89
4.4 SEM images of fibrils after pulling tests. A: fibril A lays onto the drop of glue. B: fibril B has been pulled out of the glue, note the cast impressed on the PDMS substrate. C and D: aged fibrils broken and laying on the substrate. 89

List of Tables

2.1 Fibrils features. 48

3.1 Values of force and length at failure for 2-12 pulling direction at 100, 10, 1, 0.5, 0.1, 0.05 m/s. 52
3.2 Values of force and length at failure for 7-12 pulling direction at 100, 10, 1, 0.5, 0.1, 0.05 m/s. 52
3.3 Values of force and length at failure for 2-7 pulling direction at 100, 10, 1, 0.5, 0.1, 0.05 m/s. 53
3.4 Values of spring constant for direction 2-12, 7-12 and 2-7 at pulling rate of 100, 10, 1, 0.5, 0.1, 0.05 m/s. 54
3.5 Values of force and length at failure for structure 1 at 100, 10, 1, 0.5, 0.1, 0.05 m/s. 55
3.6 Values of force and length at failure for structure 2 at 100, 10, 1, 0.5, 0.1, 0.05 m/s. 55
3.7 Values of force and length at failure for structure 3 at 100, 10, 1, 0.5, 0.1, 0.05 m/s. 55
3.8 Values of spring constant for structure 1, 2 and 3 at pulling rate of 100, 10, 1, 0.5, 0.1, 0.05 m/s. 56

4.1 Values of fibril elastic modulus. 87
4.2 Mechanical tests. 90

Abstract

Introduction

Collagen is the most abundant protein in the human body. It confers mechanical properties to many tissues, such as tendon, skin, bone and cartilage. In particular, it is responsable for mechanical stability, strenght and toughness. Due to natural ageing or disease, such as diabetes, the structural function and integrity of tissues that contain collagen can be compromised. During tissue ageing, non enzymatic cross-links between tropocollagen molecules are generated due to the reaction of glucose and amino-acids residues in the collagen molecule. A large amount of non enzymatic cross-links is not physiological and it leads to an increase in collagen stiffness, breaking load and denaturation temperature.

Collagen has a highly organised hierarchical structure, and although the mechanical properties on the macroscale have been diffusely investigated, little is known about properties of individual fibrils at the nanoscale. Collagen fibrils have a diameter in the range of 50 - 500 nm and a variable length in the order of micrometers, therefore their mechanical characterization presents many challanges. In order to support the *in vitro* characterization, computational models have been developed. They use the molecular dynamics to create fibrils models and test them under different conditions.

The present work aims to mechanically characterize both physiological and aged fibrils, in order to study the difference between them in terms of mechanical properties. Specifically, the influence of non enzymatic cross-links, or AGEs (Advanced Glycation End products) on the mechanical properties of single fibrils is analysed. In order to reach this goal, the work is divided into two parts: a

computational modeling of collagen fibrils, and an experimental validation of the results obtained in the computational part. In particular, both computational and real fibrils are stretched until failure and values of force, length, stress and strain at failure, and elastic modulus are calculated.

Material, Methods and Results

Molecular Dynamics

The first activity of this work concerns the generation of molecular models of collagen fibrils and their pulling at a defined rate. Different models have been investigated: a physiological fibril, in which tropocollagen molecules are linked by enzymatic cross-links at the N and C termini; and an aged fibril, where non enzymatic cross-links are randomly added to the structure. The non-enzymatic cross-links density is known from literature, but their distribution is not clear. For this reason, two aged fibrils have been modeled: the first one contains cross-links homogeneously distribuited through the cross-section, the second one contains cross-links only in the external shell. The main goal of this part of the work is to analyse how the mechanical properties change among the three models. In order to create them, several steps have been performed.

First of all, mechanical characterization of an enzymatic cross-link and a non-enzymatic one has been done. From literature it is known that the most abundant enzymatic cross-link is hydroxylysyl pyridinoline and the major non-enzymatic one is glucosepane. A full atomistic model of the two cross-links has been created. The cross-links have been pulled at different strain rate: 100 m/s, 10 m/s, 1 m/s, 0.5 m/s, 0.1 m/s and 0.05 m/s. During this first step, a reactive force field (ReaxFF) has been applied. Differently from conventional force fields, ReaxFF allows the bond breakege because it calculates the bond length and the bond order at each step, but simulations using RRF take more time and more computational cost than simulations using non reactive force field, that are quick but do not describe the bond failure. Since the main goal of this part is to analyse the cross-link force and length at failure, the RFF has been used. The results show that after a first period of unrolling, the force starts increasing untill it reaches the breakage

point. Elaborationg the force-length curve, the spring constant has been calculated and the value obtained will be used in the fibril model. There are no significant differences among the different strain rates.

As second step in the fibril building, pulling tests on tropocollagen molecules have been performed in order to calculate bond and angle constants. A coarse-grained approach has been used. This approach, contrary to the full atomistic one, allows to reduce the computational cost of a simulation describing a group of atoms (amino-acids in this case) with a single structure, called bead. A bead expresses the mechanical properties of the group of atoms it describes, but it can not illustrate the chemistry of the model. A validation of the coarse-grained tropocollagen model, in which each bead represents 6 amino-acids, has been performed comparing the pulling test results with the results of the same tropocollagen molecule describes with a full atomistic model. Bond and angle constant have been calculated.

Once all parameters have been calculated, the physiological fibril model has been created. The fibril has a length of 960 nm and a radius of 10 nm and it contains 332 tropocollagen molecules. It is thiner than a real one (average radius among 50 and 500 nm) but due to the computational cost only small fibrils have been tested in this work. Firstly a fibril without cross-links has been stretched in order to compare the behavior without cross-links and the one with them. Then, enzymatic cross-links are added between the C and N termini and the correct bead of a neighboring molecule. The value for the cross-link stiffness has been selected from the previous analysis on the cross-link stretching. Finally, the fibril has been stretched. The data have been elaborated in order to obtain Force - Length and Stress - Strain curves. Mechanical parameters have been extrapolated from the curves. Maximum stress and strain are due to the unravel of the fibril: σ_{max}=0.445 GPa, ε_{max}=17.72%. The elastic modulus has been calculated as linear interpolation of the first part of the curve: E=2.06 GPa.

Finally, aged fibrils have been created. Two different models have been developed. The first one contains non enzymatic cross-links (glucosepane) through the whole section of the fibrils; the second one contains cross-links only in the external shell. From literature it is known that aged fibrils contain 1 cross-link every 5 tropocollagen molecules, since the fibril model contains 332 molecules,

66 glucosepane bonds have been created. In order to generate the cross-links, a C++ script has been written. It contains a function that randomly creates 66 bonds between two beads of two neighboring fibrils. The bond rigidity value has been selected from the analysis on glucosepane. Pulling tests on both the models have been performed. Concerning the first model, the Force - Length curve has the same shape of the physiological one. The maximum values for stress and strain, and the elastic modulus are: σ_{max}=0.512 GPa, ε_{max}=18.82%, E=2.39 GPa. Concerning the second model, maximum values are: σ_{max}=0.539 GPa, ε_{max}=19.96% and the elastic modulus is 2.47 GPa. As expected aged models are stiffer than the physiological one and elastic modulus are in the literature range. Since the pulling rate is high and the fibril dimensions are not the natural ones, the peak does not represent the failure of the fibril but it is due to the unreavel of the structure. Further analysis must focus on the development of a more realistic model in terms of dimensions and pulling rates.

Experimental Validation

In order to validate the computational model, the experimental work aimed to develop an experimental set-up and perform preliminary mechanical tensile tests on single physiological and aged fibrils. Using a micromechanical testing instrument both kind of fibrils have been stretched until failure and values of force, length, stress, strain and elastic modulus have been collected.

The first purpose of the work was to obtain single fibrils, and two different approaches have been followed: *in vitro* fibrillation and isolation from rat tail tendon. The fibrillation technique has been investigated as reconstucted fibrils do not contain proteoglicans or biological residues that could be found in natural fibrils and may affect mechanical properties. Furthermore they do not contain cross-links so it is possible to introduce a known amount of cross-links. In order to obtain self-assembled collagen type I fibrils a fibrillation protocol has been developed. The fibrils have been reconstructed from a commercial solution of collagen, acetic acid and a buffer at a specific collagen concentration, pH and temperature. SEM images have been taken to verify the protocol efficacy and spaced fibrils with considerable diameters and D-period have been found. The main limitation of the

fibrillation process is its highly sensitivity to enviromental changes that often influences the results. For this reason self-assembled fibrils have not been used for mechanical testing. The second approach was the isolation of fibrils from rat tail tendon: fascicles have been extracted from the tail tendon, deposited on a glass surface with milli-Q water or PBS (phospate buffer saline) and gently scratched with tweezers. Fibrils have been removed from the fascicle and deposited in the liquid drop, then the drop has been collected and stored in an Eppendorf. SEM images have been taken to validate the procedure of extraction and deposition: the samples presented short segments of isolated fibrils with D-banding and variable diameters. Overall the isolation technique has been considered more suitable than the fibrillation because of its highly reproducibility and simplicity, therefore mechanical test have been performed on fibrils from rat tail tendon. Meanwhile, in order to create aged fibrils with a consistent amount of non enzymatic cross-links, fibrils from rat tail tendons have been aged using Methylglyoxal (MGO) following a standard procedure.

To perform mechanical tests, fibrils have to be visualized under microscope. In order to find the best procedure for visualization of a single fibril, three different microscopy techniques have been tested. Firstly fibrils have been labelled with primary and secondary antibodies in order to be fluorescent and have been visualized under fluorescence microscope. From fluorescence images fibrils with different diameters were clearly visible but the fluorophore quikly decayed and it was not possible to extract a single fibril. After, samples containing fibrils have been visualized using an optical microscope in dark field, an optical technique where the direct illumination is excluded and the light comes laterally. Fibrils were still distinctly visible but the working distance between the lens and the sample was low so the insertion and movimentation of tensile machine tips has not been feasible. Finally fibrils have been visualized under inverted optical microscope, only fibrils with high diameters (500 - 1000 nm) have been displayed due to the magnification which was not optimal. Even if the resolution did not allow the visualization of small fibrils the inverted microscope will be used to extract fibrils as the movimentation of the FemtoTool was easier.

After the visualization of a speciment that contains fibrils a specific one has be choosen, the micromechanical testing instrument has be drived near the fibril and,

using an epoxy glue, the edges of the fibril has been glued to substrate. Finally the tensile probe has been glued to one edge of the fibril and it has been driven above the other edge in order to perform pulling tests.

The experimental set-up developed in the present work has been validated through preliminary tests on single fibrils. Four fibrils have been tested, two physiological and two aged. Values of force (11 - 16 μN), length (55 - 64 μm), stress (0.3 - 0.5 GPa), strain (160-350 %) at failure and elastic modulus (0.1 - 0.4 GPa) obtained from the tests are in the physiological range but data analysis must be improved increasing the number of samples.

Conclusions

The molecular model and the experimental one have been developed in order to evaluate mechanical properties of single collagen fibrils and they both improve the state of art. Computational studies present in literature simplify the fibril structure while the model developed in this work allows a more complete and detailed description of single fibril biomechanics. Concerning the experimental approach, the method proposed here permits to mechanically test single fibrils with the use of an easy and reproducible set-up. The limitations of the computational model are mainly the fibrils dimensions that are too small and the pulling rate that is not physiological so the model must be improved. Concerning the experimental set-up, a larger number of samples must be tested in order to increase the data. To conclude, further improvements are necessary in order to better undestand the mechanical behavior of collagen fibrils, and the AGEs effects on them but the methods proposed in this work are good starting point to reach this goal.

Sommario

Introduzione

Il collagene è la proteina più abbondante nel corpo umano. Conferisce proprietà meccaniche a molti tessuti, tra i quali tendine, osso e cartilagine. In particolare, il collagene è responsabile della stabilità meccanica e della resistenza dei tessuti. A causa del naturale invecchiamento o di alcune patologie, come il diabete, la funzionalità strutturale e l'integrità dei tessuti che contengono collagene può essere compromessa. Durante l'invecchiamento del tessuto, la reazione tra glucosio e residui amminoacidici presenti nelle molecole di collagene adiacenti produce cross-links non enzimatici. Qualora il numero di cross-link non enzimatici aumenti considerevolmente, le proprietà meccaniche e biologiche del collagene, come il carico a rottura o la temperatura di denaturazione, vengono compromesse, e il tessuto risulta più rigido.

Il collagene presenta una struttura gerarchica. Nella macroscala le proprietà del collagene sono state ampiamente studiate, mentre poco si conosce ancora circa il comportamento delle singole fibrille nella nanoscala. Le fibrille di collagene hanno un diametro di 50 - 500 nm e una lunghezza variabile nell'ordine dei micrometri, di conseguenza la loro caratterizzazione meccanica presenta molte difficoltà. Modelli molecolari sono stati sviluppati al fine di supportare la caratterizzazione *in vitro*: essi permettono lo sviluppo di modelli di fibrilla che possono essere testati in varie condizioni.

Lo scopo del presente lavoro è la caratterizzazione meccanica di fibrille fisiologiche e invecchiate, al fine di studiare come le proprietà meccaniche variano in presenza di cross-links non enzimatici, o AGEs (Advanced Glycation End prod-

ucts). Il lavoro è suddiviso in due parti: nella prima parte viene sviluppato un modello computazione di una singola fibrilla, nella seconda parte vengono effettuati test sperimentali per la validazione del modello molecolare. Entrambi i modelli sono sottoposti a test di trazione fino a rottura al fine di valutare le proprietà meccaniche in termini di forza, lunghezza, sforzo e deformazione a rottura, e modulo elastico.

Materiali, Metodi e Risultati

Dinamica Molecolare

Come prima attività si sono sviluppati modelli molecolari che descrivono singole molecole di fibrille di collagene fisiologiche ed invecchiate e si sono effettuate simulazioni di trazione a rottura. Tre diversi modelli sono stati creati, il primo descrive una fibrilla di collagene fisiologica che presenta i cross-links enzimatici ai terminali N e C delle molecole di tropocollagene. Il secondo e il terzo modello rappresentano fibrille invecchiate e includono, oltre ai cross-links fisiologici, anche quelli non enzimatici. Siccome la densità dei cross-links non enzimatici è nota in letteratura ma la loro distribuzione non è specificata due diverse fibrille sono state modellizzate: la prima contiene i cross-links non enzimatici distribuiti omogeneamente lungo tutta la sua sezione, la seconda fibrilla li presenta solo sulla superficie esterna. Il principale obiettivo di questa parte del lavoro è l'analisi della variazione delle proprietà meccaniche fra i tre modelli ma per il loro sviluppo sono stati svolti diversi studi preliminari.

In primo luogo è stata svolta la caratterizzazione meccanica dei cross-links enzimatici e non enzimatici. Da letteratura è noto che il cross-link enzymatico preponderante all'interno delle fibrille di collagene di tipo I è l'hydroxylysyl pyridinoline e che quello non enzymatico è il glucosepane. Di conseguenza si è creato un modello full atomistic dei due cross-links e si sono svolte diverse simulazioni di trazione a rottura a con velocità (100 m/s, 10 m/s, 1 m/s, 0.5 m/s, 0.1 m/s e 0.05 m/s) utilizzando il Reactive Force Field (RFF) implementato in LAMMPS. Questo tipo di force field premette la rottura dei legami poichè, a ogni step della simulazione, ricalcola la distanza fra gli atomi e ridefinisce le interazioni fra di

essi ma richiede una capacit di calcolo superiore e lunghi tempi computazionali. I risultati ottenuti mostrano come, dopo un periodo di srotolamento della struttura molecolare, le forze sostenute dal cross-link aumentano linearmente fino a raggiungere un valore massimo. Successivamente i legami si rompono e le forze si annullano. Dall'analisi dei grafici di Forza - Lunghezza si è ricavata la costante elastica che è stata successivamente implementata all'interno del modelli di fibrille. Si è notato che la velocità di trazione non influenza significativamente il valore della costante elastica.

Successivamente si è modellizzata e caratterizzata una singola molecola di tropocollagene utilizzando un approccio coarse-grained e il Force Field nonReattivo implementato in LAMMPS. Questo approccio, al contrario della modellizzazione full atomistic, riduce fortemente il costo computazionale poichè modellizza un gruppo di atomi come una singola struttura, chiamata bead, dotata delle medesime proprietà meccaniche del gruppo che rappresenta ma ne tralascia la struttura chimica. Il modello coarse-grained proposto prevede l'aggregazione di 6 amino acidi all'interno di un singolo bead, la struttura intera è caratterizzata dalle costanti di legame e di angolo che sono state validate confrontando simulazioni di trazione su un modello full atomistic dello stessa molecola di tropocollagene con quelle svolte nel presente lavoro.

Infine si è creato il modello fisiologico della fibrilla. La fibrilla contiene 332 molecole di tropocollagene, è lunga 960 nm ed ha un raggio di 10 nm. Le sue dimensioni risultano quindi essere molto inferiori ai valori reali (raggio medio 50 - 500 nm) ma, per ragioni computazional, in questo lavoro sono state simulate solo strutture molto piccole. Per investigare l'effetto dei soli cross-links enzimatici si è sviluppata inizialmente una fibrilla di sole molecole di troocollagene e la si è trazionata a 100 m/s. Si è notato che per per valori di deformazione del $\sim$20% gli sforzi raggiungono valori massimi e poi diminuiscono poichè un gruppo compatto di molecole di tropocollagene viene sfilato dalla struttura. In seguito si sono aggiunti i cross-links enzymatici ai terminali C ed N legandoli ai corretti beads delle molecole vicine utilizzando uno specifico script in C++. Il valore di rigidezza del cross-link è stato ottenuto dalle analisi fatte precendentemente sull' hydroxylysyl pyridinoline e i cross-links sono stati modellizzati come legami covalenti fra dei bead specifici. Infine la fibrilla è stata trazionata: i dati ottenuti sono stati

elaborati in grafici di Forza - Lunghezza e Stress - Strain. I grafici presentano un andamento lineare di incremento della forza e un successivo decremento, anche in questo caso il picco massimo viene associato allo sfilamento della fibrilla e non è da riferirsi alla rottura dei cross-links. I valori massimi di sforzo e deformazione sono: σ_{max}=0.445 GPa, ε_{max}=17.72%. Il modulo elastico, calcolato come interpolazione lineare del grafico Stress - Strain è: E=2.06 GPa.

I modelli delle fibrille invecchiate sono stati creati. Nel primo i cross-links non enzimatici sono stati distribuiti omogeneamente lungo tutta la sezione della fibrilla. Nel secondo modello essi sono stati distribuiti solo sulla superficie esterna. Poichè la loro densità è nota, 1 cross-link ogni 5 molecole di tropocollagene, e il modello contiene 332 molecole, 66 cross-links non enzimatici sono stati modellizzati. Per la loro generazione è stato scritto uno script in C++ contenente una funzione che randomicamente crea legami fra i bead di molecole vicine. La rigidezza del legame è stata selezionata dall'analisi precendente svolta per la caratterizzazione meccanica del glucosepane. A causa dell'elevato costo computazionale, le simulazioni di trazione sono state svolte su entrambi i modelli a 100 m/s. Per quanto riguarda il primo modello, le curve di Forza - Lunghezza e Stress - Strain presentano lo stesso andamento riscontrato per il modello fisiologico di fibrilla. Di conseguenza i valori massimi di sforzo e deformazione (σ_{max}=0.512 GPa, ε_{max}=18.82%), sono da attribuirsi allo sfilamento di un gruppo compatto di molecole di tropocollagene. Il modulo elastico, come aspettato, è incrementato rispetto al modello fisiologico a causa della presenza dei cross-links non enzimatici e vale 2.39 GPa. Il secondo modello, che presenta i cross-links non enzimatici sono sulla superficie esterna, presenta i seguenti valori di sforzo e deformazione massimi σ_{max}=0.539 GPa, ε_{max}=19.96%. Il modulo elastico, anche in questo caso, è incrementato rispetto al modello fisiologico e vale: 2.47 GPa.

Validazione Sperimentale

Lo scopo del lavoro sperimentale è la validazione del modello computazionale attraverso lo sviluppo di un set-up sperimentale e l'esecuzione di test meccanici su singole fibrille, sia fisiologiche che invecchiate. I test meccanici di trazione sono stati effettuati fino a rottura.

Come prima attività sono state ottenute singole fibrille di collagene attraverso due differenti approcci: fibrillogenesi *in vitro* e isolamento di fibrille a partire da tendine di code di ratto. Il primo approccio, la fibrillogenesi, è stata studiata in quanto le fibrille ricostituite da soluzione commerciale non contengono proteoglicani o residui biologici che, al contrario, si trovano nelle fibrille naturali e che possono influenzare le proprietà meccaniche del collagene. Inoltre, le fibrille così formate non presentano cross-links che possono quindi essere aggiunti in quantità note. Fibrille di tipo I sono state ottenute facendo reagire collagene commerciale in soluzione con acido acetico e un adeguato tampone, a determinati pH e temperatura. Le fibrille sono state analizzate attraverso l'utilizzo del SEM: il protocollo seguito ha permesso di ottenere singole fibrille con D-banding e diametri nel range fisiologico ma la sensibilità del processo alle condizioni ambientali influenza considerevolmente la riuscita dello stesso e ne limita l'ultizzo. Per questa ragione le fibrille ottenute tramite fibrillogenesi *in vitro* non sono state testate. Il secondo approccio riguarda l'isolamento di fibrille da tendine di coda di ratto: i fascicoli sono stati estratti dal tendine, depositati su una superficie di vetro contenente acqua milliQ o soluzione tampone (PBS, phosphate saline buffer) e grattate contro di essa tramite l'uso di pinzette. I campioni così ottenuti sono stati analizzati al SEM, le immagini mostrano fibrille isolate che presentano D-banding e diametri variabili. I test meccanici sono stati eseguiti su fibrille ottenute con questa tecnica in quanto si è dimostrata altamente riproducibile. Infine, alcuni fascicoli sono stati incubati in MGO (Methylglyoxal) per permettere la creazione di cross-links non enzimatici e ottenere fibrille invecchiate.

Al fine di testare le fibrille, esse devono essere visualizzate con l'utilizzo della microscopia. Sono stati testati tre differenti metodi. Primo, le fibrille sono state marcate con anticorpi primari e secondari, rese fluorescenti e visualizzate attraverso il microscopio a fluorescenza. Le immagini così ottenute hanno mostrato che la visualizzazione con questo metodo è possibile, ma la rapida decadenza del fluoroforo ne limita il tempo di utilizzo. Come secondo metodo è stato testato un microscopio in dark field. La tecnica di dark field esclude la diretta illuminazione del campione ma permette alla luce di incidere su di esso solo lateralmente; le fibrille sono visibili chiaramente anche con questa tecnica ma la distanza di lavoro tra la lente e il campione risulta insufficiente per l'inserimento della macchina di

test e pertanto non è utilizzata. La terza tecnica di visualizzazione utilizza un microscopio a inversione. Esso permette la visualizzazione delle fibrille anche se i diametri calcolati non risultano reali ma sono affetti da distorsione ottica. Nonostante questo limite, il microscopio a inversione è utilizzato per la visualizzazione delle fibrille durante i test meccanici in quanto la movimentazione dello strumento di test risulta migliore e più facile. Le fibrille scelte per i test meccanici sono state incollate al substrato attraverso l'utilizzo di un ago montato sullo strumento di prova. Infine, l'ago è stato sostituito con un sensore che presenta una punta, essa è stata incollata ad una estremità della fibrilla precedentemente incollata al substrato e trazionata fino a rottura del campione. Il set-up sviluppato è stato utilizzato per testare quattro fibrille, due fisiologiche e due invecchiate. Valori di forza (11 - 16 μm), lunghezza (55 - 64 μm), sforzo (0.3 - 0.5 GPa) e deformazione (160 - 350 %) a rottura, e i moduli elastici (0.1 - 0.4 GPa) ottenuti dai test sono nel range aspettato, ma l'analisi sperimentale deve essere migliorata testando un numero maggiore di campioni.

Conclusioni

Il modello molecolare e quello sperimentale descritti nel presente lavoro sono stati sviluppati per investigare le proprietà meccaniche delle singole fibrille di collagene. Entrambi arricchiscono lo stato dell'arte dal momento che gli studi computazionali presenti in letteratura semplificano molto la struttura della fibrilla e i metodi sperimentali risultano poco riproducibili o non analizzano le proprietà meccaniche a rottura. Il modello computazionale proposto descrive dettagliatamente la struttura meccanica delle singole fibrille e permette la modellizzazione di diversi tipi di cross-links mentre il set-up sperimentale utilizzato è facilmente riproducibile e permette la caratterizzazione meccanica completa delle fibrille. I modelli, però, presentano diversi limiti: in quello computazionale la fibrilla è molto piccola e viene trazionata a velocità non fisiologiche. Di conseguenza si rende necessario incrementare le sue dimensioni e ridurre drasticamente la velocità attraverso l'uso di una maggiore capacità computazionale. Mentre, sebbene il set-up sperimentale sia stato validato, è necessario testare un numero elevato di fibrille, fisiologiche e invecchiate, per poter ottenere un'analisi significativa. In-

oltre il modello presentato analizza le proprietà meccaniche delle fibrille in condizione non idratata e quindi non realistiche.

In conclusione, si rende necessario sviluppare maggiormente entrambi i modelli per comprendere il comportamento meccanico delle fibrille e l'effetto degli AGEs. Nel complesso i modellii proposti si possono considerare dei validi puntii di partenza per il raggiungimento di questi obiettivi.

Chapter 1

INTRODUCTION

1.1 Aim of the work

Structural functions and integrity of connective tissues can be partially compromised by pathologies like diabetes and the natural tissue ageing. At the molecular level, connective tissues observe an increase of Advanced Glycation End-products (AGEs) [1]. AGEs form cross-links inside collagen fibrils which may cause alterations in their mechanical behavior and in the interactions among collagen and cells or other molecules in the extra cellular matrix [2]. The density of AGEs, that describes the quantity of cross-links present in the tissue (pmol/mg), has been already evaluated [3] but their position in the fibril is still unknown. This work aims to develop a computational model of a collagen fibril in presence of different distributions of AGEs in order to analyse the mechanical properties and to validate the model through experimental tests.

The first activity consists in the development of a computational model of phisiological and aged collagen fibrils, followed by pulling simulation, conceived to obtain the mechanical behavior in the two states. This activity is performed using the Molecular Dynamics technique described in Sec.2.1.

The experimental work includes the production of fibrils through the fibrillogenesis technique and the isolation of fibrils from rat tail tendon. Some of the fibrils are incubated with MGO in order to induce the formation of AGEs. A bench for the mechanical tests is developed, then stretching tests are performed

for the validation of the computational model, Sec.2.2.

1.2 Collagen

1.2.1 Collagen structure

Collagen is the most abundant protein in the human body. It is a structural protein that provides mechanical properties to many tissues, like tendon, bone, teeth and cartilage. More than 19 types of collagen have been described [4], but the most abundant are type I, type II and type III [5]. Type I is present in skin, tendon and bone; type II forms cartilage and type III is present in the cardiovascular system tissues and in the granulation tissue.

Collagen has a highly organised hierarchical structure (Fig.1.1).

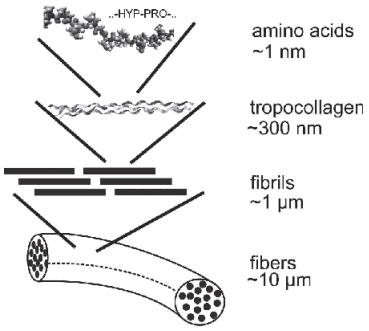


Figure 1.1: Hierarchical structure of collagen, Buehler [6].

The structural block is a collagen fibril, made of several building blocks: the tropocollagen molecules have a lenght of 300 nm and a diameter of approximately 1.5 nm. They consist of three polipeptide chains (called α chain) rolled into a right handed triple helical structure. Each α chain is composed of a big central region with the repetition of the sequence (Gly-X-Y)$_n$; where n=337-349, Gly is Glycine and X, Y can be any amino acid, but Proline and Hydroxyproline are the most present. Type II and type III collagen are homotrimeric molecules, they contain 3 identical α chains; type I collagen is a heterotrimeric molecule, formed by 2 α_1 chains and 1 α_2 chain. Collagen type I, II and III have two non helical regions at the N and C termini called telopeptides of about 20 residues length. Telopeptides

are very important for the mechanical properties of the collagen because they allow the formation of cross-links between different tropocollage molecules.

The collagen fibril is composed of several tropocollagen molecules packaged in an axial structure. It is possible to observe a repetition of 67 nm made of gap and overlap regions, called D-period. Due to eletrostatic and hydrophobic interactions the tropocollagen molecules align beside each others, each molecule is shifted relative to the previous one and the packaging of shifted molecules creates the D-period. Fibrils associate to form a larger structure, the fiber.

Mechanical properties of fibrils are guaranteed by intermolecular cross-links, formed between two different tropocollagen molecules. Inibition of cross-link formation leads to loss of strenght in the tissue. Two types of cross-links are known: the enzymatic and the non-enzymatic ones. The enzymatic cross-links are responsable for mechanical properties, tissue toughness and strenght. They take place between two specific sites, and the two molecules are linked head to tail. The non-enzymatic cross-links are due to the tissue aging or to specific diseases, such as diabetes. They are due to non-enzymatic reaction of glucose or its products, with lysine and arginine residues present in the helical region of the collagen molecule. This process is called glycation and the non-enzymatic cross-links are often called AGEs (advanced glycation end products). The increase of AGEs involves an increase in the tissue stiffness and brittleness. Collagen loses its compliance and the fibers functionality is compromised. The location of glycation cross-links is not clearly determined.

1.2.2 Enzymatic cross-links

The mechanical properties of collagen, such as strenght and toughness, are due to enzymatic cross-links. The enzymatic cross-links formation is initiated by enzyme lysyl oxidase, that acts on specific lysines in telopeptides [7]. Depending on the precursors (telopeptide lysine, helix lysine, telopeptide hydroxylysine, helix hydroxylysine), 4 different mature cross-links can be formed (lysyl and hydroxylysil pyrrole, lysyl and hydroxylysil pyridinoline), and all of them are trivalent (Fig.1.2).

Eyre et al., [8] report that pyrroles and pyrydinolines in the same amount are

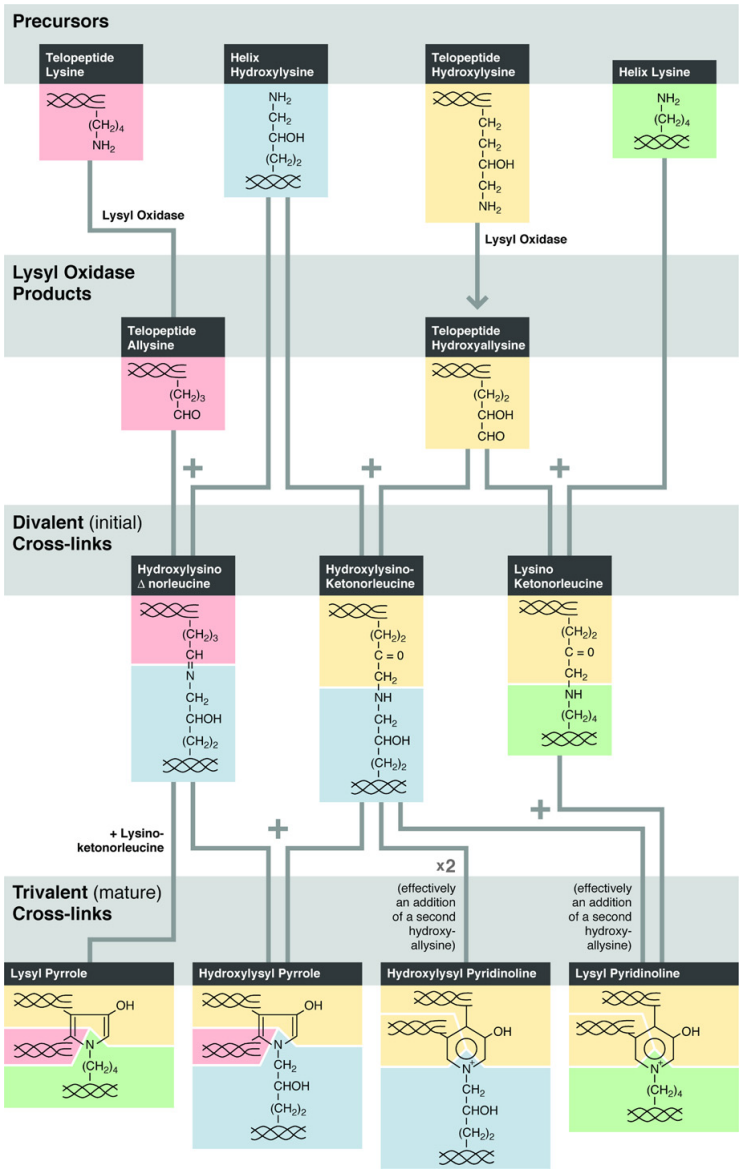


Figure 1.2: Chemical pathway of crosslinking interactions, Eyre et al. [8].

characteristic cross-links of bone collagen, and hydroxylysil pyridinoline is predominant in cartilages. They also show that the telopeptide at the N termini links to the aminoacid 930 and the telopeptide at the C termini links to the amino acid 87 (Fig.1.3).

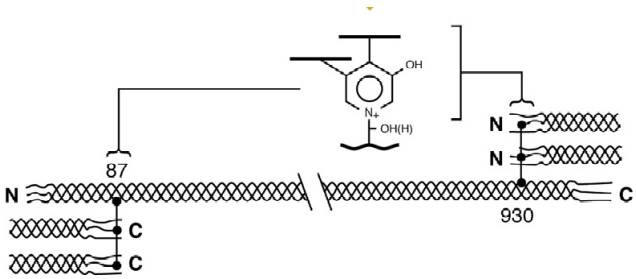


Figure 1.3: Sites of enzyymatic cross-links formation, Eyre et al. [8].

1.2.3 Non-enzymatic cross-links

Non-enzymatic cross-links are created in presence of glucose or its products. The glucose reacts with lysine or arginine residues in the helical region of the molecule and it forms AGEs. AGEs increase with tissue maturation or in presence of particular diseases, for example diabetes. AGEs influence the mechanical properties of the tissue because they increase the collagen stiffness, the breaking load, the denaturation temperature, the solubility and they decrease the susceptibility of the collagen to degradatives enzymes [9] (Fig.1.4). Due to these modifications, the tissue becomes brittler and it loses its elasticity. The formation process is not clearly known yet and the sites of formation of these type of cross-links have to be discovered. Some studies have shown that mature tissues have 1 cross-link per 5 collagen molecules, while diabetic tissues have 1 cross-link per 2 molecules [3]. A very important non-enzymatic cross-link is glucosepane, created from the reaction of a glucose molecule with lysine and arginine (Fig1.5). Verzijl et al. [10], studied the effect of AGEs on mechanical properties of human cartilage. They incubated a sample of cartilage with thereose and they found out that an increase in AGEs causes an increase of tissue stiffness that may contribute to failure of collagen network to resist damage. In turn, Sell and collegues, [3], determined the level of AGEs in human skin in both diabetic and non diabetic subjects as a func-

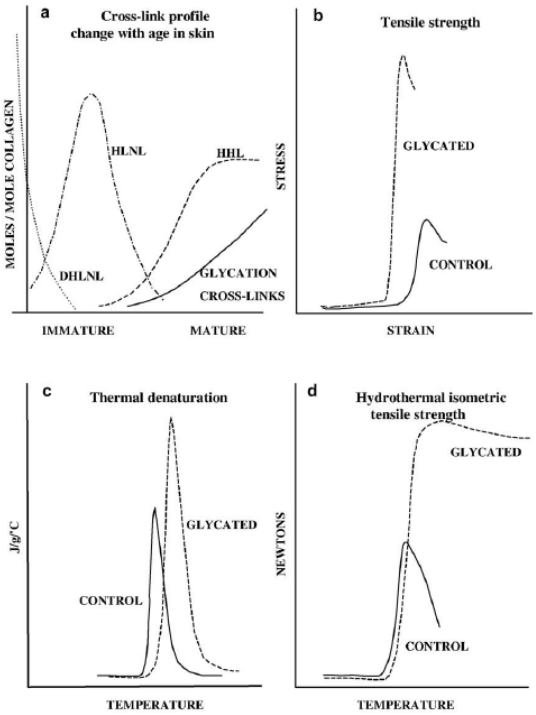


Figure 1.4: Effect of non-enymatic cross-links on the physical properties of collagen, Avery et al. [9].

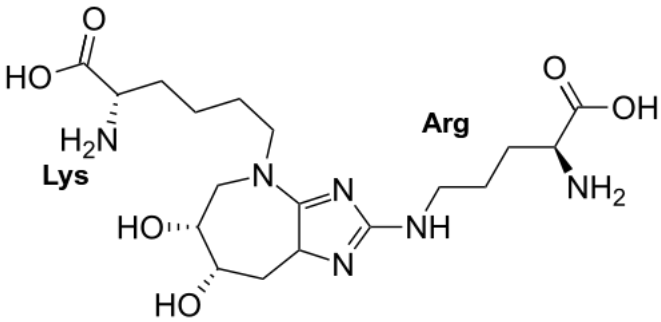


Figure 1.5: Glucosepane.

tion of age. They found out that in non diabetic subjects glucosepane significantly increased with age ($p<0.0001$) and it reached a value of 2000 pmol/mg of collagen in 90 year old subjects; they also found out that diabets significantly increases glucosepane level up to 5000 pmol/mg of collagen ($p<0.0001$). Normal level of glucosepane in lens protein have been studied by Biemel et al., and in [11] they declear that glucosepane level is 132-240 pmol/mg of protein. All these studies agree that glucosepane is the major and dominant non-enzymatic cross-link and that high quantities of it contribute to collagen disfunctions and properties. They also analyse the density of AGEs in different conditions but none of them evaluate their distribution in the fibril.

1.3 Fibrillogenesis

The process of *in vivo* fibrillogenesis is very complex and much remains to be understood. The process involves a collagen precursors molecule (or procollagen) which is converted into collagen by the removal of its N- and C-terminal sequences by procollagen N-proteinase and procollagen C-proteinase. Then the collagen spontaneously self-assembles into cross-strainded fibrils that are stabilized by intermolecular cross-links [12] (Fig.1.6).

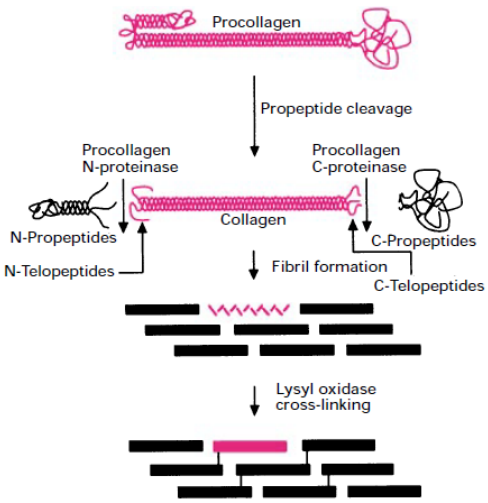


Figure 1.6: Synthesis of fibrillar collagens, Kadler et al. [12].

In order to better comprehend this process, *in vitro* fibrillation has been widely studied. The initial studies tried to reconstruct the physiological conditions (pH and temperature) in order to obtain spontaneous fibril formation with D-banding. The fibrils were reconstructed from a solution of collagen in acid, for example acetic acid, to obtain a low pH.

Further studies demontrated that the fibrillation process is influenced by many factors [5]:

- buffer composition
- temperature
- intactness of N and C telopeptides
- presence of other types of collagen
- presence of other macromolecules

The buffer composition has been studied by Williams et al. [13] and Liu et al. [14]. Their results show that two different buffer solutions (for example phosphate buffered saline, PBS, and sodium phospate, SP, in [14]) significantly influence the fibrils diameter.

Intactness of N and C telopeptide is an important parameter in fibrils formation because telopeptide are involved in cross-links formation. There are two types of commercial collagens avaiable [12]:

- acid soluble; in this type of collagen telopeptides are intact and the fibrils result long and cylindrical.
- pepsin-soluble; the collagen is extracted with pepsin that digests some protein structures (except for the helical region of collagen molecules), the absence of telopeptides causes short and cigar shaped fibrils.

The temperature influences the fibrils diameter. Low temperature (for example 20°C) causes large diameters, up to 200 nm; high temperature (34°C) causes smaller diameters [12].

The fibril formation is an entropy driven process. It can be monitored by turbidimetry at $\lambda \approx 300$ nm. The amount of fibrillar material formed can be described by light scattering, that is proportional to the collagen concentration and to the mass of assembled structures. The turbidity curve tipically has 3 phases (Fig.1.7):

1. a lag phase: no change in turbidity is observed, and only few molecules self associate
2. a rapid growth phase: big amount of molecule self associate
3. a plateau region

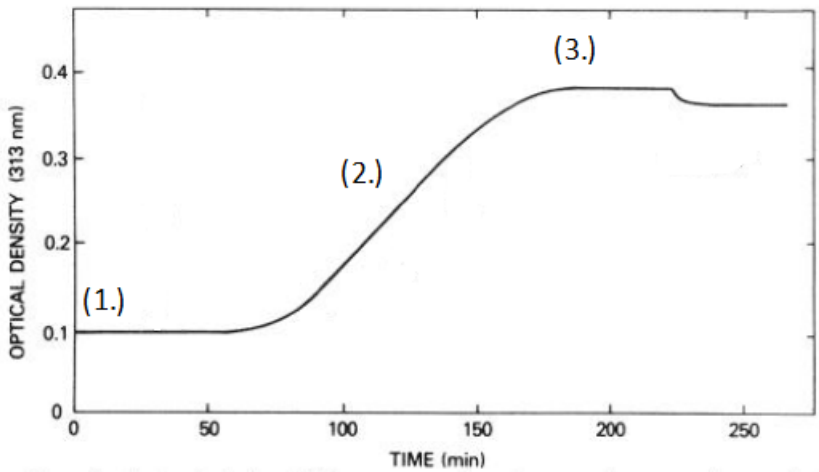


Figure 1.7: Typical turbidity curve, Williams et al. [13].

Williams et al. [13], evaluated the best conditions for *in vitro* fibrillation. They used rat tail tendon collagen with intact non helical ends. Polymerization was initiated by raising the pH to a value of 7.3 and warming a cold solution of collagen. They followed the process by turbidity. Their resuts show that the optimal conditions are: 30 mM phosphate, 30 mM N-[tris(hydroxymethyl)methyl-2-amino]ethanesulfonic acid, NaCl, pH=7.3, T=20°C - 30°C and collagen concentration between 0.02 and 0.05 mg/ml. They observed that in case of a small amount of phosphate concentration the banding is still evident but the aligment

and lateral packaging are poorer. They also studied the effect of pH on fibril formation and concluded that in the pH range of 7.0-7.5 there is not difference in the fibrils morphology but above a value of 8.0 the fibrils surface is distorted and there is not lateral packing. They varied the temperature between 12°C and 37°C: between 20°C and 30°C no change in fibrils structure has been observed; below this range protein appears a mixture of filaments agregates with occasional thin fibrils; above the range banding was not observed.

In turn, Liu et al. [14] studied the effect of different temperatures and different buffer solutions on single fibril formation. They used two different buffers (PBS and SP) and three temperatures (37°C, 33°C, 29°C) and incubated the collagen at pH 7.4 for 18 to 24 hours. Their results show significantly different average lengths for single fibrils incubated with the same buffer at different temperatures and between the to buffers.

Raspanti et al. [15][16], studied the influence of collagen initial concentration and the effect of decorin or glycosaminoglycans. Concerning the initial collagen concentration, they discovered that it influences the fibril diameter. High concentrations (1mg/ml) cause partially lateral fused small fibrils, while small concentrations (0.1mg/ml) result in huge superfibrils. They also found out that decorin prevents fibril fusion and it restores the normal fibrils appearance, and that in presence of GAGs the fibrils have a smaller diameter compared to fibrils polymerized without GAGs.

1.4 Mechanical tests on single fibrils

Collagen type I fibrils have a wide distribution of diameters [5], from 10 nm to 500 nm [17]. Due to the fibrils dimensions, mechanical tests on collagen single fibrils is challenging. Furthermore, isolation and visualisation of fibrils present many challanges and finally, collagen fibrils have a viscoelastic behaviour that make the strain calculation difficult. When fibrils are stretched at small strain rates (≤ 6 %) they have a reversible behaviour and they return to the initial configuration, so the stress-strain curve is difficult to obtain.

Two approches have been proposed in order to mechanically test fibrils. The

first one uses the AFM force spectroscopy [18], [19], [20]; the second one involves the use of microelectromechanical systems (MEMS) [21], [22].

The AFM method consists in attaching a single fibril between an AFM cantilever and a glass surface with the use of epoxy glue. When the attachment is completed, the cantilever is raised from the surface and the test begins (Fig.1.8). The main limitations of this method are the small-strain regime of investigation

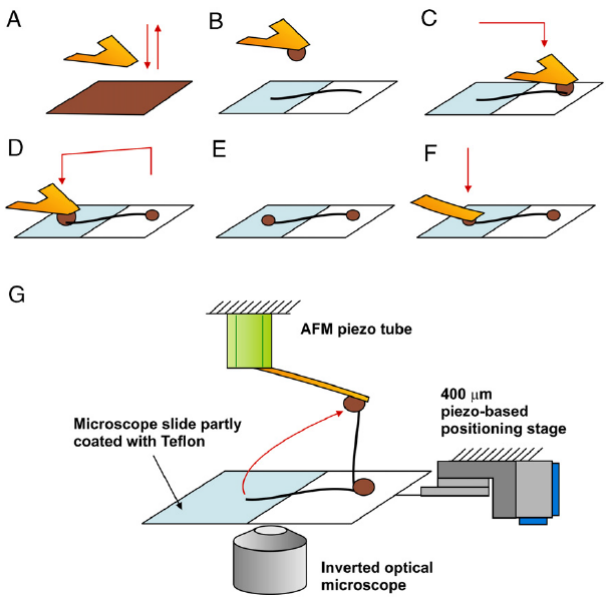


Figure 1.8: Schematic view of the AFM tests, Yang et al. [20].

(<5%), that does not allow failure tests; and the impossibility of *in situ* imaging, that makes some interesting informations unavaiable. Van der Rijt et al., [19] stretched collagen fibrils in both dried and hydrated conditions. Concerning the test in ambient conditions, they stretched the fibrils up to 90 MPa in stress, that means about 3% strain. They did not reach the fibril breakage, they found a linear strain-stress relation, with an elastic modulus of 5±2 GPa. Concerning the tests in hydrated conditions, they use a PBS solution, and they found an increase in diameter of 73 ± 15%, a different shape of the stress-strain curve, and different values for the elastic modulus, that varies between 250 and 450 MPa. Their results show that test conditions have a great influence on the results and that fibrils behave different in dried or hydrated environments. Yang and collegues,

[20] found values for the elastic modulus similar to the ones obtain by Van der Rijt. They obtained an elastic modulus of 0.6 ± 0.2 GPa. They also study the effect of cross-links on mechanical properties incubating the fibrils with EDC/NHS (water-soluble carbodimide 1-ethyl-3-(3-dimethyl-aminopropyl)carbodiimide hydrochloride in presence of N-hydrpxysuccinimide) and performing the tests with the AFM. They found an elastic modulus $\sim$ 3 times higher than the one of native fibrils.

The MEMS approach allows to stretch single fibrils until the yield point and, in some cases, until failure [21] (Fig.1.9). This method allows also *in situ* imaging, so the breakage point can be analysed and stress-strain curves can be created. The main limitation of this method is the small mechanical resolution.

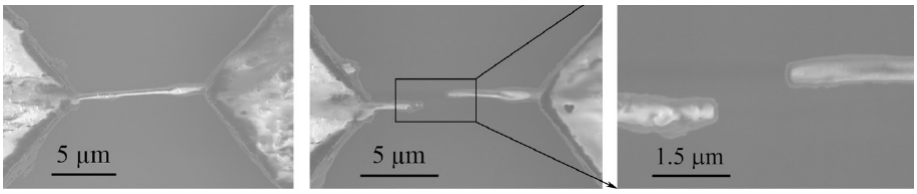


Figure 1.9: SEM image of the fibril breakage with MEMS, Shen et Al. [21].

Shen et al., [21] tested 13 specimen using a MEMS device. They obtained an elastic modulus of 0.86 ± 0.45 GPa, they conclude that fibrils are capable of sustaining stresses higher than 0.60 GPa without fracturing. They finally expose fibrils to a damaging scenario, that allows them to break the fibril.

To conclude, research in this field is open and in future mechanical tests on single fibrils have to be improved.

1.5 Computational models of collagen fibrils

Since mechanical tests on single collagen fibril require complex test benches and present several limitations, the use of computational models can be a valid option in order to investigate the mechanical behavior of these structures. In literature only two works focused on this purpuse and are both based on the Molecular Dynamics (Sec.2.1).

Buehler [23] proposes a 2D periodic coarse-grained fibril model based on an array of 2 x 5 tropocollagen molecules as shown in Fig.1.10. The parameters for a single tropocollagen molecule are obtained from full atomistic simulations [6] and the molecule is modelled as a collection of beads that describes a group of atoms. Cross-links are modelled as incrementation of non bond interactions at the ends of each tropocollagen molecule. In order to evaluate the influence of the cross-links and the elastic properties at failure, pulling test simulations are performed in presence of different cross-links densities with a pulling rate of 0.4 m/s.

Figure 1.10: Schematic representation of fibril (a) and the model where cross-links are represented increasing the adhesion on the tropocollagen edges (b). Adapted from Buehler Markus J. [23].

The results show that higher cross-link densities lead to the incrementation of yield stress and rupture stress (from 0.2 GPa to 6.2 GPa) and to a brittler behavior. The main limitations of this model are that it is not 3D and the spatial cross-link distribution is homogeneous. Furthermore, cross-links inserted in the model have the same characteristics and it is impossible to differentiate enzymatic and non enzymatic ones.

Gautieri et al. [24] have developed a viscoelastic model of collagen fibril based on Kelvin-Voigt elements (consisting in a spring and a dashpot in parallel) that describe tropocollagen molecules. Performing full atomistic simulations of creep on a short chain of solvated and dry collagen peptide, Young's modulus and viscosity have been calculated for the tropocollagen molecule. These values have been included in the Kelvin-Voigt element, then the units have been connected in series and in parallel in order to create the fibril (Fig.1.11).

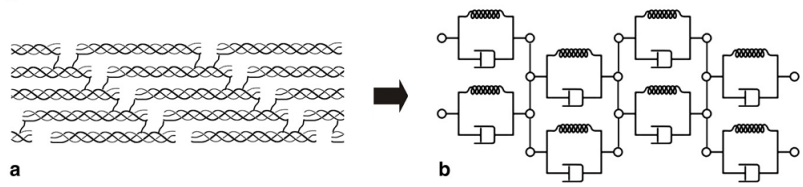


Figure 1.11: Schematic representation of collagen fibril (a) and the viscoelastic model based on Kelvin-Voigt units (b). Adapted from Gautieri et al. [24].

Young's modulus values, 2.5 GPa for wet collagen fibril and 9.0 GPa for the dry one, are a good estimation of the elastic properties while viscous ones are not well described by the model because it does not consider the sliding of tropocollagen molecules on each other. Since the fibril model is theoretical and it is based on an assembly of Kelvin-Voigt elements connected in series and in parallel, the presence of different cross-links distribution and their effect can not be analysed and it represents the main limitation as it restricts the flexibility of the model.

As it shown in literature, few investigations have been performed in this field and the developmet of a more realistic and adaptable fibril computational model is desirable to better understand mechanical properties.

Chapter 2

MATERIAL AND METHODS

2.1 Molecular Dynamics

2.1.1 Introduction

Molecular Dynamics (MD) is one of the principal computational techniques that allows to simulate the time dependent behavior of atoms and molecular systems. This method is widely used to investigate thermodynamic and mechanic properties at microscopic level of several molecular and atomic structures as it can easily study complex systems. Each system of particles could be completely described by a quantistic approach based on Schrodinger's equation

$$\hat{H}\Psi(r,t) = -i\frac{h}{2\pi}\frac{\partial}{\partial t}\Psi(r,t) \tag{2.1}$$

where the Hamilton's operator $\hat{H}$ describes the kinetic energy and the interactions among particles. The behaviour of the system could be obtained from the wave function Ψ(r,t) but for big and complex structures this approach is not appropriate because it is hard to find Ψ(r,t) that satisfies the Schrodinger's equation. In order to obtain a solution for these kind of systems it is important to analyse the dynamic behaviour of molecular systems that is composed by the motion of nuclei and the motion of electrons. The Born-Oppenheimer approssimation permits to separate them and also allows to describe the molecular system as a system of point particles. Their motion can be described by classical mechanics and the effect of

electrons is included in the potential force field. The MD approach uses the Born-Oppenheimer approssimations so it strongly reduces the calculation complexity and it lets the simulation of complex systems for nanoseconds [25].

2.1.2 Basics of atomistic modeling: structure of MD Program

In this paragraph a schematic overview of the structure of MD program is reported. The input parameters are the number of particles and their mass, the volume of the

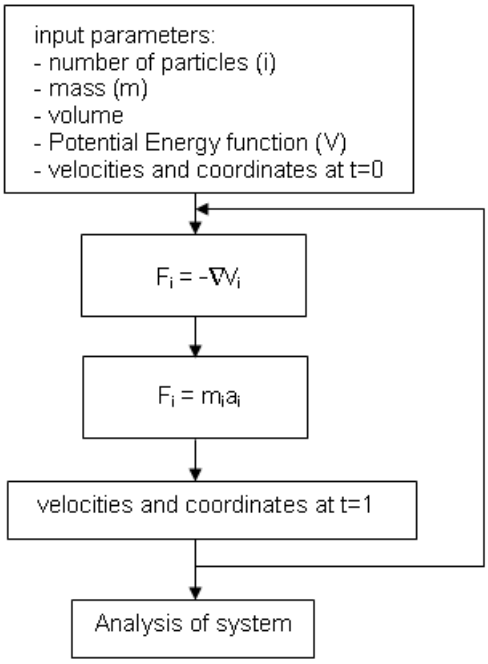


Figure 2.1: Schematic structure of MD program.

system, the potential energy function, the coordinates and velocities for each particle at t=0. Forces are calculated using the potential energy derivative, Newton's laws are integrated and new coordinates and velocities are found. These steps are repeted several times in order to propagate the system through time, at each step coordinates and velocities are evaluated and memorized. Then dynamic properties can be calculated from the MD trajectory obtained by the evolution in time of coordinates and velocities.

The *Potential Energy Function* describes the energy of the system and it is the

sum of bond terms and non-bond terms.

$$V(r) = \sum V_{bonded} + \sum V_{non-bonded} \tag{2.2}$$

The V_{bonded} term represents the interaction among bonded atoms and it is composed by bonds, angles and bond rotations terms

$$V_{bonded} = V_{b-stretch} + V_{angle-bend} + V_{diedral-tor} \tag{2.3}$$

The first term is the potential of bond stretching and it describes the interaction between atomic pairs, in the following equation it is modeled as an harmonic spring:

$$V_{b-stretch} = \frac{1}{2} k_b (r - r_0)^2 \tag{2.4}$$

k_b is the stiffness constant of the bond which depends on chemical type of atoms, r the distance between atoms and r_0 the equilibrium distance.

The second term is the potential of angle bending. It is represented again by an harmonic potential, k_a is the angle stiffness, θ the current angle and θ_0 the optimized one.

$$V_{angle-bend} = \frac{1}{2} k_a (\theta - \theta_0)^2 \tag{2.5}$$

The third term is the torsion angle potential that describes the interaction among four atoms separated by three covalent bonds. This potential evaluates rotations by the dihedral angle ϕ, the coefficient of symmetry n and it is assumed to be periodic.

$$V_{diedral-tor} = k_\phi (1 - \cos(n\phi)) \tag{2.6}$$

All the parameters described previously (k_b, r_0, k_a, θ_0, k_ϕ) can be obtained from quantistic and spectroscopic analysis.

The $V_{non-bonded}$ term represents the contribution of non-bonded interactions, the Van der Waals interaction energy and the Coulomb energy.

$$V_{non-bonded} = V_{VdW} + V_{Coulomb} \tag{2.7}$$

The Van der Waals potential is the balance between repulsive and attractive forces

for a pair of non charged atoms. At short distance electronic clouds are overlapped and repulsive force prevales, when atoms are far the attractive one becomes dominant. When the distance between atoms is optimized the energy reachs a minimum that depends on chemical types of atoms. This potential is often modelled using the Lennard-Jones potential

$$V_{VdW} = 4\varepsilon[\underbrace{(\frac{\sigma}{r})^{12}}_{repulsive} - \underbrace{(\frac{\sigma}{r})^{6}}_{attractive}] \tag{2.8}$$

at the distance σ the potential is null and for r_{min}, the potential reachs the minimum ε (Fig.2.2).

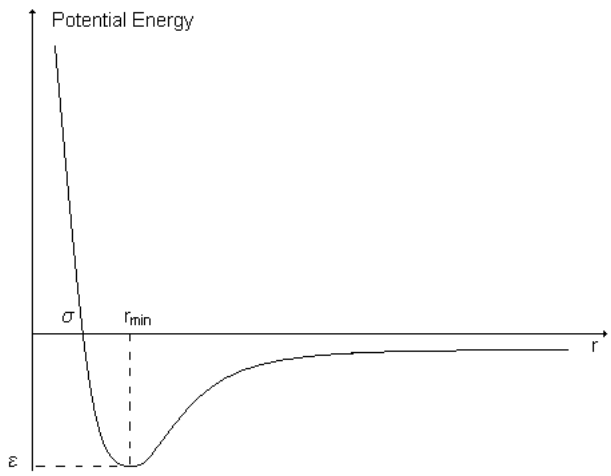


Figure 2.2: Lennard-Jones potential.

The Coulomb potential represents the electrostatic interaction between two non-bonded atoms

$$V_{Coulomb} = \frac{1}{4\pi\varepsilon_0\varepsilon_r}\frac{q_1 q_2}{r_{1-2}} \tag{2.9}$$

ε_0 is the dielectric constant in vacuum, ε_r is the relative dielectric constant, q_1 and q_2 are the charges and r_{1-2} is the distance between the pair of atoms.

The force that acts on particle i, mass m_i and acceleration a_i is obtained from the *Newton's motion equation* (2.10) but also it is given by the gradient of the Potential Energy (2.11).

$$F_i = m_i a_i \tag{2.10}$$

$$F_i = -\nabla V \tag{2.11}$$

Combining (2.10) and (2.11) yields

$$-\frac{\partial V}{\partial r_i} = m_i \frac{\partial^2 r_i}{\partial t^2} \tag{2.12}$$

From equation (2.12) it is not possible to obtain motion evolution (trajectory and velocity) of each particle because this equation has not analytical solution then it has to be solved numerically. In order to find the trajectory of particles several integration algorithms, like Verlet algorithm or Leap-frog algorithm, have been developed. All of them are generated from Taylor series expansion of position (2.13) and velocity (2.14). They all respect the conservation of energy and momentum and they must be computationally efficient.

$$r(t+\partial t) = r(t) + v(t)\partial t + \frac{1}{2}a(t)\partial t^2 + .. \tag{2.13}$$

$$v(t+\partial t) = v(t) + a(t)\partial t + \frac{1}{2}b(t)\partial t^2 + .. \tag{2.14}$$

This specific example represents the Leap-frog algorithm:

$$r(t+\partial t) = r(t) + v(t+\frac{1}{2}\partial t)\partial t \tag{2.15}$$

$$v(t+\frac{1}{2}\partial t) = v(t-\frac{1}{2}\partial t) + a(t)\partial t \tag{2.16}$$

Velocities are explicitly calculated at time t+$\frac{1}{2}\partial t$ and they are used to obtain positions at time t+∂t, so positions and velocity are not calculated at the same time. In order to evaluate velocities at time t+∂t the average between v(t-$\frac{1}{2}\partial t$) and v(t+$\frac{1}{2}\partial t$) is calculated [26].

2.1.3 Reactive Force Field and Non Reactive Force Field

The force field describes the atomic interactions via the potential energy fuction. It is the most important part of MD methods as it characterizes the molecular system behavior [6]. Force fields can be divided into two groups: Reactive Force

Fields (RFFs) and Non Reactive Force Fields (NRFFs).

NRFFs simulate bond interaction as an harmonic spring and bonds can not be created or broken because when the distance between two atoms increases the energy increases too. RFFs allow bonds breakage, the potential energy function is not harmonic and when bond separation increases the energy becomes null and the bond is broken (Fig.2.3). Reactive Force Fields riformulate the bond length and the bond order at each step so simulations take more time and computational capacity than NRFFs where bonds are fixed.

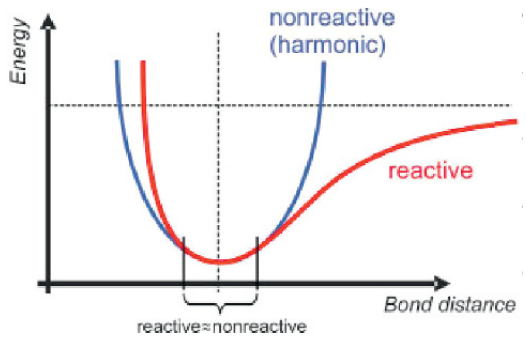


Figure 2.3: Buehler [6], Non Reactive Force Field versus Reactive Force Field.

As it is shown in the work of Buehler [6], for a stretching simulation of a single tropocollagen molecule, (Fig.2.4), NRFF (a) and RFFs (b-c) give same results of forces for small strains but NRFF (a) can not allow bond breakage so for higher strains, values of force are not realistic and do not match with RFFs results.

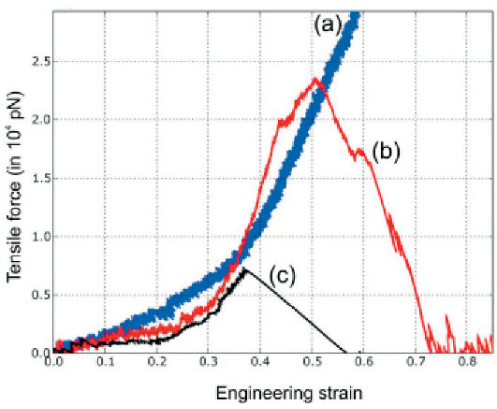


Figure 2.4: Buehler [6], modeling fracture of a single tropocollagen.

Then NRFFs are fine for simulations near the equilibrium point and for small strain while RFFs must be used for tests at failure.

There are several MD programs that implement NRFFs. NAMD [27], for example, uses the non reactive force field CHARMM [26] but it does not support RFFs. The MD program that supports both force fields and that is used in this work is LAMMPS [28].

2.1.4 Fibril implementation in LAMMPS

Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) is a classical molecular dynamics code developed by Steve Plimpton, Aidan Thompson, and Paul Crozier and it is designed to efficiently compute Newton's motion equation for a system of particles. It is generally used to model atoms, coarse-grained particles, polymers, organic molecules, proteins, DNA, metals and granular materials. It can run on single processors or in parallel architecture, it supports several types of force fields and it is useful to obtain thermodynamic and kinetic informations.

A LAMMPS input script typically has 4 parts:

1. initialization
2. atom definition
3. setting
4. run

In the first section, units, dimensions, force field and style of atoms and bonds are set. Then the system of atoms must be defined reading one of the following files:

- file.pdb.: It contains topology informations (3D coordinates) of each atom, the atom type that is related to the chemical structure, bonds and angles among atoms. This file can be taken from the RCSB Protein Data Bank and from cristallography analysis.

- data.file: It contains the same informations of file.pdb but it is created using VMD[29], TopoTools, a plugin for manipulating topology informations, and a pdb file.

The third part is the setting, here force field coefficients and the simulation type (e.g. energy minimization, pulling..) have to be defined, boundary conditions, time integration and time step are set. Then output options like thermodynamic informations, force, displacement and trajectory are specified. Finally the simulation is run and output files are generated. When the simulation ends it is possible to analyze the outcomes.

LAMMPS is the MD program used for the development of the fibril model as it allows to use both Reactive and Non Reactive Force Field. In order to characterize the behavior at failure of enzymatic and non-enzymatic cross-links, respectivly hydroxylysil pyridinoline and glucosepane, the Reactive Force Field implemented in LAMMPS is required. Several stretching simulation are performed, values of force and displacement at failure for different rates of pulling are collected.

To develop a tropocollagen molecule with low computational weight, a coarse-grained approach, already presented in literature [30] [6], is used. Groups of atoms are collected in pseudoatoms called beads so the number of degrees of freedom rapidly decreases and more complex structures can be simulated. The coarse-grained structure has to be completely characterize to best fit full atomistic analysis [31]. For this activity the use of Non-Reactive FF implemented in LAMMPS is appropriate. Tropocollagen molecules and cross-links are assembled together. Inside the fibril, cross-links are modelled as bonds between beads of different tropocollagens. These bonds contain breakage informations that are obtained from the values at failure previously calculated for the cross-links. Then a stretching simulation is performed on the entire structure. Crosslinks failure parameters are already inside bonds. Values of force and displacement are collected and analyzed using a spreadsheet. After several tests on different structure of fibrils both physiologic and not, it is possible to compare results of force, displacement, stress and strain. More over elastic modulus can be evaluated to verify possible changes in the stiffness of fibril in presence of glucosepane.

Before starting the mechanical characterization of cross-links the validation of

LAMMPS RFF is required since pulling tests on those cross-links are not reported in literature. Therefore deca-alanine, a small amino acid, is stretched using a NRFF in NAMD and using the RFF in LAMMPS at different pulling rates (1m/s and 100 m/s). The results (Fig.2.5) agree with Buehler's work[6]: for small strains forces match whereas for major strains they are different.

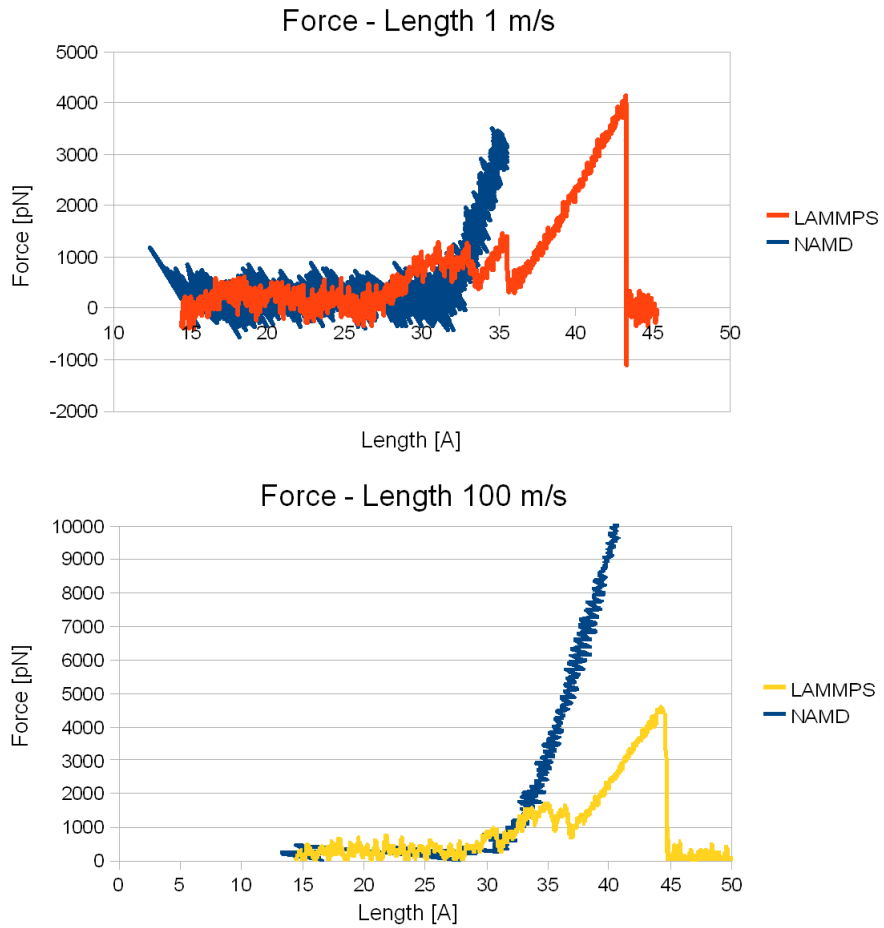


Figure 2.5: Force - Length chart for alanine pulling at 1 m/s and 100 m/s.

2.1.5 Mechanical characterization of enzymatic cross-links

Hydroxylysyl pyridinoline prevails in cross-linked peptides from adult cartilage [8], and bond and tendon collagen present a large amount of it. Other trivalent cross-links such as lysil pyrrole, hydroxylysys pyrrole an lysyl pyridinoline can

be found in mature tissues but none of them is predominant. For this reason hydroxylysyl pyridinoline is mechanically evaluated and it is implemented as the enzymatic cross-link in the model of fibril. In order to characterize the behavior at failure of hydroxylysyl pyridinoline, the RFF of LAMMPS is used and several pulling tests are performed. The trivalent cross-link is stretched in three different pulling directions (2-12, 2-7, 7-12) where 2, 7, 12 indicate carbon atoms (Fig.2.6). Tensile tests are performed at different rates: 100 m/s, 10 m/s, 1 m/s, 0.5 m/s, 0.1 m/s, 0.05 m/s.

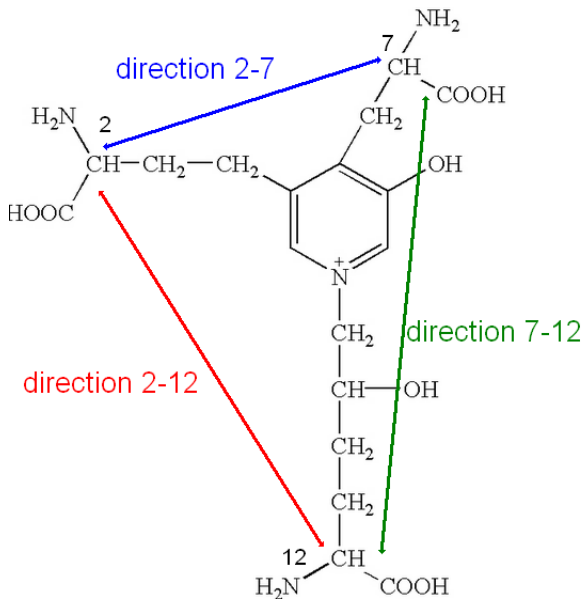


Figure 2.6: Hydroxylysyl pyridinoline, pulling directions: 2-12, 2-7, 7-12.

The tests are developed to understand if different pulling directions lead to different mechanical behavior and if pulling velocity may influence values of force and length at failure.

An input script for direction 2-12 at 1m/s is reported in Fig.2.7. In the first section the force field is defined, in this case the reactive one. In the *groupinfo* section atoms that are stretched are defined (2 and 12), in the equilibrium section the pulling rate is in Å/fs where

$$1m/s = 10^{-5}\ \mathring{A}\ /fs \tag{2.17}$$

```
# REAX potential for CHO system
# .....

units      real

atom_style charge
read_data  data.enzym            # molecular structure

pair_style reax/c lmp control
pair_coeff * * ffield.reax 1 2 4 3

neighbor   2 bin
neigh_modify    every 10 delay 0 check no

########  groupinfo
group           mypull id 2                # first group for
stretching
group           myfix id 12                # second group for
stretching

#######  Equlibration
unfix           1
fix        1 all nve
fix             2 all qeq/reax 1 0.0 10.0 1e-6 param.qeq
fix        3 all langevin 300.0 300.0 1000.0 48279

fix             pull mypull smd cvel 10.0 0.00001 couple myfix
auto auto auto 0      # Stretchig Info
fix        myforce all ave/time 1 100 10000 f_pull[1] f_pull[2]
f_pull[3] f_pull[4] f_pull[5] f_pull[6] f_pull[7] file
smd.force  # Stretching output

timestep        0.25
run             6000000
```

Figure 2.7: Input script.

Dynamic values of force and length are obtained from the file smd.force. Length is evaluated in Å, force in kcal/mol/Å, and it has to be converted into piconewtons using the following equation:

$$F_{pN} = -69F_{kcal/mol/\mathring{A}} \tag{2.18}$$

Then charts of force versus length are obtained using a spreadsheet and parameters at failure are collected for each pulling direction and each velocity. Finally the spring constant is evauated using the linear interpolation tool.

2.1.6 Mechanical characterization of non-enzymatic cross-links

Glucosepane is the most abundant non-enzymatic cross-link [3] [11], therefore in this work it is mechanically characterized and used in the model of aged fibril. As shown in the previous section, pulling tests are performed using the RFF of LAMMPS. Three structures of glucosepane, characterized by different initial configurations are obtained from three pdb files. Then pulling tests are performed on each structure at different rates (100 m/s, 10 m/s, 1 m/s, 0.5 m/s, 0.1 m/s, 0.05 m/s) in direction 1-22 where 1 and 22 indicate the carbon atoms (Fig. 2.8).

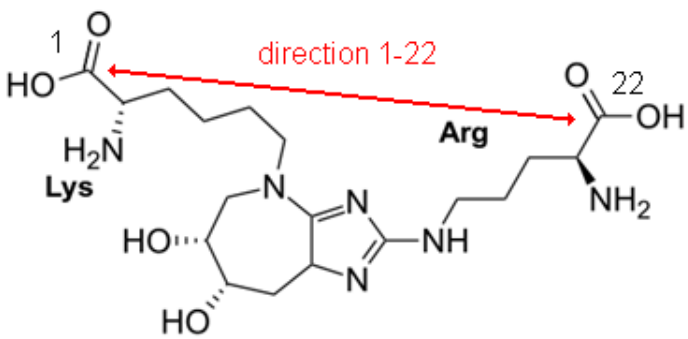


Figure 2.8: Glucosepane, pulling direction 1-22.

The smd.force file is analyzed, charts of force versus length are elaborated, values at failure are collected and the spring constant is calculated.

2.1.7 Mechanical characterization of a tropocollagen molecule

In order to investigate the mechanical properties of fibrils, a system with a high amount of atoms has to be created. This kind of system requires a huge computational cost. It is possible to reduce the computational cost using a corse-grained model. In this model a group of atoms is expressed by a single pseudoatom, called bead, that describes the mechanical properties of the group of atoms. A coarse-grained model expresses the mechanics of a system but it is not able to define the chemical properties of it.

As a first step in the model development, a validation of the coarse-grained approch is performed. A small peptide chain is created and tested in order to compare the results with a full-atomistic model (Fig.2.9).

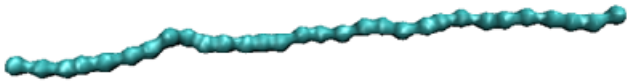


Figure 2.9: Coarse-grained model of a peptide.

The peptide chain contains 30 beads that simulate a triple helix of 84 nm. The mass of a single bead is 1776 kDalton. A previous work on a single peptide chain simulated via a full atomistic model [31] shows that the force-length behaviour is not linear (Fig.2.10).

In particular, two different stiffness values can be seen: in the first part of the curve, the stiffness is very low; in the second part of the curve the stiffness increases. The coarse-grained model developed in this work must be validated matching the behavior of the full atomistic model. First, the two stiffness values are calculated from the full atomistic curve using the linear interpolation tool. Second, the two values are used in the coarse-grained input file, the simulation is performed and the force-length curve is evaluated. Finally, the two curves, one

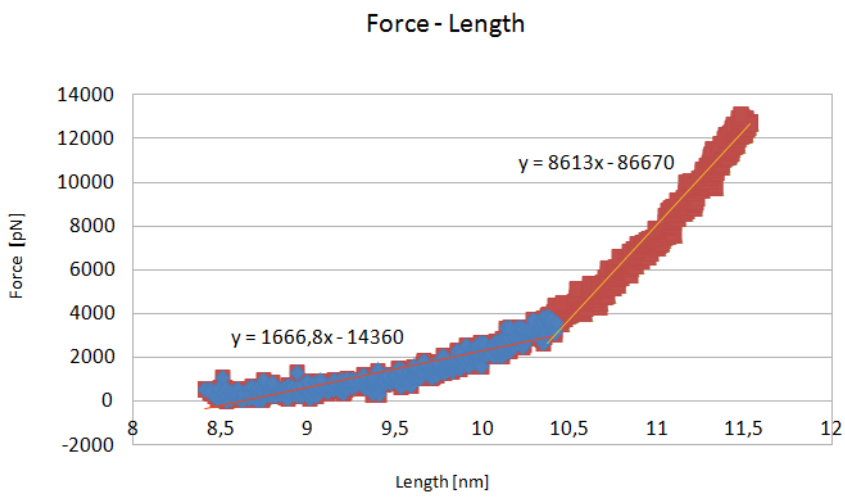


Figure 2.10: Full atomistic simulation of a peptide.

from the full atomistic model and the other from the coarse-grained approach, are compared: if they match, the coarse-grained model is validated. In order to simulate the two-phase behaviour, it is necessary to use two different spring constants. An updated version of LAMMPS allows to use a bond style that can simulate more than one spring constant in the same bond: the *piecewise* bond style. In Fig.2.11 the script used for the simulation is shown.

```
units           real
boundary        f f f

atom_style      bond
read_data       data.test

bond_style      piecewise
bond_coeff      * 3.04 70 0.5 362

pair_style      lj/cut 0
pair_coeff      * * 1.0 1.0 2.5

neighbor        100 bin
neigh_modify    every 10 delay 0 check no

########  groupinfo
group           mypull id 30
group           myfix id 1
```

Figure 2.11: Script for the peptide simulation.

The first bond coefficient ($*$) indicates that the parameters must be applied to all bonds. The second coefficient (3.04 Å) indicates the distance between two

beads after which the first spring constant must be applied. The third coefficient is the first spring constant (70 kcal/mol/A^2); the fourth coefficient (0.5 Å) shows the increase in length after which the second spring constant must be applied and, finally, the last parameter is the second spring constant (362 kcal/mol/A^2). The peptide has been stretched at a pull rate of 1m/s. The results are compared with the full atomistic model.

The peptide chain is also used to evaluate the angle constant of the bond. In order to find the correct value for the constant, the persistence length is evaluated. The persistance length measures the chain rigidity; it is the length over which there is no correlation among atoms. It is described as follow:

$$log(cos(\theta)) = -\frac{D}{L_p} \tag{2.19}$$

Where θ is the angle between a vector that is tangent to the chain at position 0 and a tangent vector at a distance D away from 0; and L_p is the persistence length. Atomistic simulations showed that the collagen has a persistence length of $\sim$ 23.4 nm [6]. In order to calculate the persistence length, simulations with different angle constants are performed. Using a matlab script [32] the results from each simulations are analysed. The script gives as output the persistence length for each angle constant used. The constant that gives the value closest to 24.3 nm is choosen.

A single molecule of collagen, called tropocollagen, is made of 1014 amino acids and it is 300 nm long. Using the coarse-grained approach, a tropocollagen model is created. In this model, a single bead represents 6 amino acids. The model contains 169 beads; the mass of a single bead is the sum of the masses of 6 amino acids: 1776 KDalton. The bond constants and the angle constant are evaluated and a pulling simuation is performed. The results are compared with the full atomistic model in order to validate the model. Fig.2.12 shows the input script used for the simulation. The constants used in the simulation are:

$$K_1 = 1208 \frac{kcal}{mol \cdot nm^2} \tag{2.20}$$

$$K_2 = 6240 \frac{kcal}{mol \cdot nm^2} \tag{2.21}$$

$$K_a = 4.0 \frac{kcal}{mol \cdot rad^2} \tag{2.22}$$

```
atom_style      angle
read_data       data.tropo

bond_style      piecewise
bond_coeff      * 1.756 1208 0.4 6240

angle_style     harmonic
angle_coeff     * 4.0 180

pair_style      lj/cut 0
pair_coeff      * * 1.0 1.0 2.5

neighbor        100 bin
neigh_modify    every 10 delay 0 check no
```

Figure 2.12: Script for the tropocollagen simulation.

2.1.8 Lennard Jones parameters

The Lennard Jones potential describes the non-bond interaction between a pair of atoms or molecules. It is described as follow:

$$V_{LJ} = 4\varepsilon[(\frac{\sigma}{r})^{12} - (\frac{\sigma}{r})^{6}] \tag{2.23}$$

where σ is the distance at which the potential is zero, ε is the depth of the potential well and r is distance between the particles. The calculation of the Lennard Jones parameters is very important in order to have an equilibrated and stable system. The model of a single fibril is developed to calculate and validate the Lennard Jones parameters. The model contains 32 molecules of tropocollagen, each molecule has 169 beads for a total of 5408 beads. The fibril has a 10 nm diameter and it is $\sim$ 290 nm long. The distance among the tropocollagen molecules is set up to be 4 nm (Fig.2.13).

The goal of this part of the work is to calculate the correct Lennard Jones parameters in order to reduce the distance among the molecules until the physiological value: 1.6 nm. This goal is achieved with a simulation that begins with the non physiological distance (4.0 nm) and, if the parameters are correct, must end

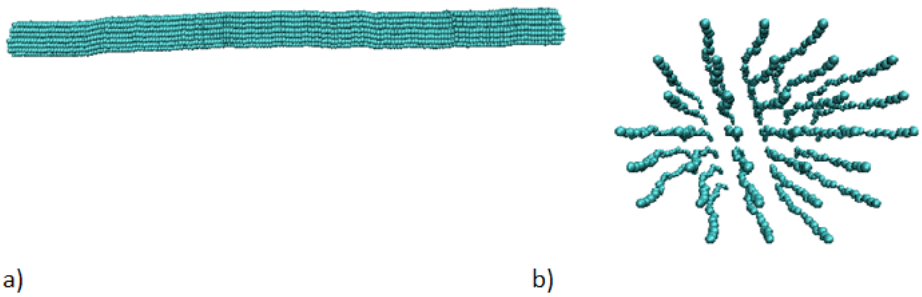


Figure 2.13: Fibril model for the Lennard Jones parameters calculation: a) longitudinal view b) cross section.

with the physiological one (1.6 nm) and must be stable. The parameters used in the simulation are σ, r_{cut} and ε. The cut radius, r_{cut} [nm], is ten times higher than the physiological distance among the tropocollagen molecules: 16 nm. σ [nm] is calculated as follow:

$$\sigma = \frac{r_{cut}}{2^{\frac{1}{6}}} \tag{2.24}$$

This leads to a value of 1.425 nm. Finally, ε is choosen from a literature work [6] and it is 7.0 kcal/mol.

2.1.9 Physiological fibril

A first model of a pseudo-physiological fibril, that does not present cross-links, is developed. A pdb file containing spatial coordinates of a short and thin fibril (length 960 nm, radius 10 nm, atoms number 65108) is converted into a data file. In the data file each atom is transformed in a bead (mass=1776 KDalton) and 169 beads are linked together in order to create a tropocollagen molecule. Values of bond and angle stiffness are taken from previous analysis. The fibril obtained is composed by 332 molecules of tropocollagen and the non-bond interactions are regulated by Lennard Jones parameters. Due to computational cost the fibril is stretched at 100 m/s in implicit solvent, beads at one edge of the fibril are pulled, beads at the other edge are mantained fixed (Fig.2.14).

The identification numbers of pulled and fixed atoms are found using VMD and are implemented in the input file (Fig.2.15).

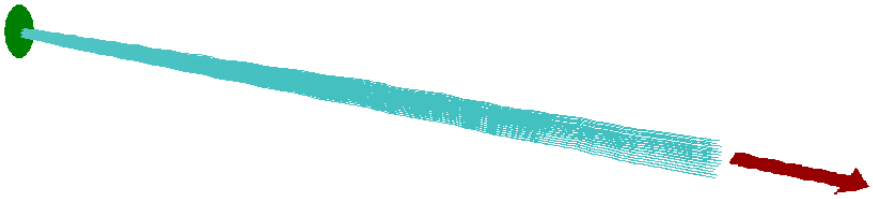


Figure 2.14: Pseudo physiological fibril: one edge of itl is pulled, the other one is mantained fixed.

```
########  groupinfo
group           mypull id 50869 51038 51207 51376 51545 51714
51883 52052 52221 52390 52559 52728 52897 53066 53235 53404
53573 53742 53911 54080 54249 54418 54587 54756 54925 55094
55263 55432 55601 55770 55939 56108              # first group
for stretching

group           myfix id 1 170 339 508 677 846 1015 1184 1353
1522 1691 1860 2029 2198 2367 2536 2705 2874 3043 3212 3381
3550 3719 3888 4057 4226 4395 4564 4733 4902 5071 5240
     # second group for stretching
```

Figure 2.15: Input file for the pseudo physiological fibril, the first group of atoms (my pull) represents one edge of the fibril and it is pulled at 100 m/s, the other edge (myfix) is mantained fixed.

Values of force and length during the pulling process are collected in the smd.force file. The chart of force versus length is obtained using a spreadsheet. Using the following equations, stress and strain are calculated and the chart is plotted, analyzed and compared with literature and experimental data.

$$\sigma = \frac{F}{A_0} \tag{2.25}$$

$$\varepsilon = \frac{l - l_0}{l_0} \tag{2.26}$$

In order to obtain a physiological model of the fibril, the data file used previously has been modify. Cross-links normaly bind the N-termini telopeptide to the aminoacid number 930, that corresponds to bead number 155, and the C termini telopetide to the amino acid 87 that is bead number 15. Enzymatic crosslinks are obtained using a C++ script and modelled as bonds in the coarse-grain model. The data file is scanned, the beads number 1 and 155 of the two closest different tropocollagen molecules are linked together in order to create N termini

crosslinks, and new bonds are added in the file. The same is done with beads 169 (C termini) and 15 (Fig.2.16). The model proposed simplifies the trivalent structure of the hydroxylysyl pyridinoline into a divalent cross-link structure.

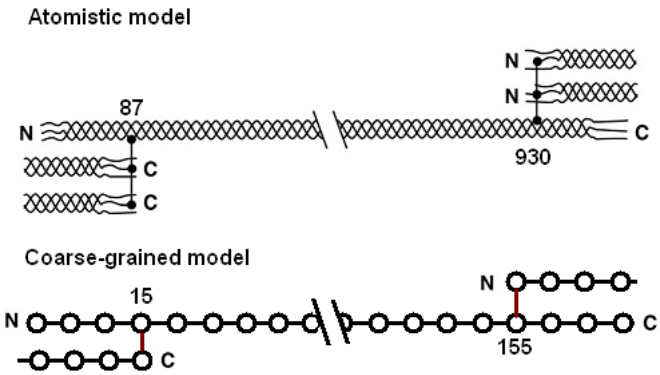


Figure 2.16: Atomistic vs coarse-grained model.

Crosslinks are now represented as new type of bonds (type 2), and must be characterized. In order to define the bond stiffness for enzymatic cross-links, values of rupture force, spring constants found in Sec 3.1 and the folded configuration of C termini and N termini are considered. In the work of Uzel and Al. [33] it has been shown that C terminus mantains the configuration shown in Fig.2.17 and binds the cross-link in the folded domain.

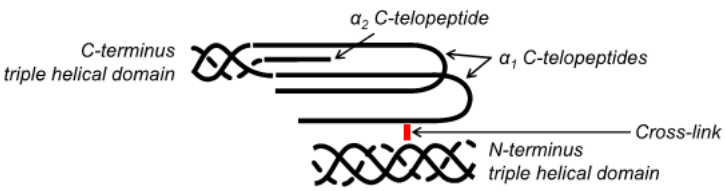


Figure 2.17: C terminus configuration, Uzel et al. [33].

During stretching the telopeptide and the cross-link firstly unroll themselves untill they are linear (about 4 nm), then they are both stretched and the cross-link withstands high forces before failure. In order to consider this behavior the cross-link bond is modelled by the *piecewise* bond style that allows to use different bond constants. The script for enzymatic bonds is reported in Fig.2.18.

```
bond_style      piecewise
bond coeff      1 1.756 1208 0.4 6240
bond coeff      2 1.6 0 2.5 610.7 3.1
```

Figure 2.18: Input script for enzymatic bonds: the first bond coeff represents bond interactions in the tropocollagen chains, bond coeff 2 describes enzymatic cross-link bonds.

The first coefficient (2) describes the bond type of enzymatic cross-link, the second value (1.6 nm) is the equilibrium distance between the pair of atoms, the third one is the first spring constant and, during the unrolling phase, it is mantained null; 2.5 nm is the increases in distance from the equilibrium point, after which the spring constant becomes 610.7 kcal/mol/nm^2. The last constant is the increase in distance from the equilibrium point at which the bond breaks (Fig.2.19).

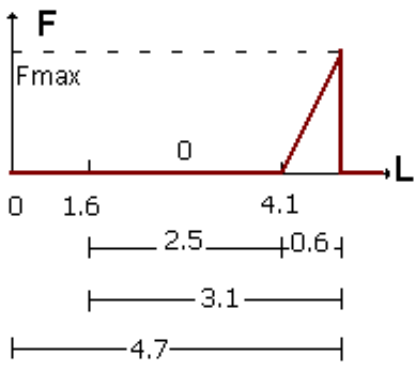


Figure 2.19: Schematic representation of force versus length behavior of enzym cross-link: The force is mantained null untill the molecule reaches a length of 4.1 nm, then it increases (spring constant = 620.7 kcal/molA^2). When F = Fmax the bond breaks and the force is null again.

Since the simulation of the fibril requires high computational cost and it must be performed at 100 m/s, the rupture force is calculated from Sec 3.1, Table 3.1 making an average between values at 100 m/s for direction 2-12 and 7-12. The direction 2-7 describes the interaction between two telopeptides and it is not considered.

$$F_{max} = \frac{2562.2 + 2862.9}{2} = 2712.6pN \tag{2.27}$$

For the same reason the bond stiffness is calculated from Sec 3.1 Table 3.4 making an average between values at 100 m/s for direction 2-12 and 7-12.

$$K_{enzymatic-bond} = \frac{369.2+473.6}{2} = 421.4\frac{pN}{A} = 610.7\frac{kcal}{molnm^2} \tag{2.28}$$

The rupture length is calculated with the following equation:

$$\varepsilon = \frac{F_{max}}{K_{enzymatic-bond}} = 0.643nm \tag{2.29}$$

2.1.10 Aged fibril

The aim of this work is to study how the mechanical properties change between a physiological collagen fibril and an aged one. Due to aging and some diseases, such as mellitus diabets, the non-enzymatic cross-links, called AGEs, increase, fibrils become stiffer and they lose their elasticity. In this work the non-enzymatic cross-link used is glucosepane. Sell et al. [3], demonstate that the number of non-enzymatic cross-links in old but non diabetic individuals reaches a level of $\sim$ 2 nmol/mg and in old and diabetic individuals is $\sim$ 4.3 nmol/mg. This leads to one cross-link every five collagen molecules in old individuals and one cross-link every two in diabetic ones.

A model of a aged fibril has been created. This model describes collagen of old, non-diabetic individuals. It contains 332 tropocollagen molecules, so the number of cross-links in the whole fibril is $\sim$ 66. The distribution of cross-links in the fibril is unknown. In this work two different distributions are analysed: in the first one, cross-links are distribuited through the whole section of the fibril, in the second model cross-links are distribuited only in the external shell of the fibril. The hypotesis of the second model is that the cross-links are not able to penetrate inside the fibril so they can only be distribuited in the external shell. The second problem in the model creation is the random distribution of cross-links. In order to simulate this kind of distribution, a script in C++ has been created. The script contains a function, called RANDOM, that gets as input a list of elements to be analysed (molecules, beads...) and gives as output a defined number (66 in this case) of elements choosen in a random way.

Concerning the first distribution, the one in which the cross-links are distributed through the whole section of the fibril, the C++ script contains 3 steps:

1. RANDOM analyses all the 332 tropocollagen molecules and it randomly chooses 66 molecules

2. RANDOM analyses all the beads of the 66 selected molecules and it randomly chooses 1 bead for each molecule, for a total of 66 beads

3. A bond between each selected bead and the closest one in a neighboring molecule is created. During this step, for each selected bead all the beads of all the molecules are analysed one by one; the distance between the selected bead and the analysed one is registered. When all the beads are analysed, the bond is created between the selected bead and the one with the smallest distance from it.

The C++ script gives as output the list of the 66 pair of beads that must be bonded. Fig.2.20 shows the non enzymatic cross-links created.

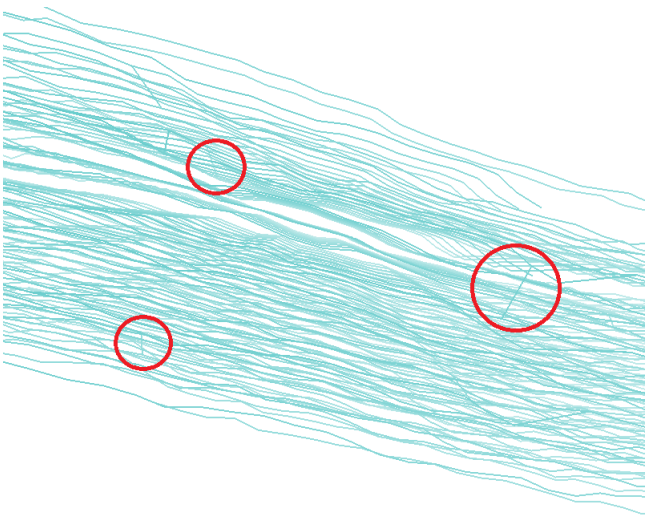


Figure 2.20: Non enzymatic crosslinks present in the whole section of the fibril.

In the LAMMPS input script (Fig.2.21), this bonds are called *3*. The values of the constant coefficients are calculated from the previous analysis on glucosepane.

The avarage bond constant for simulations at 100 m/s, calculated in Sec.3.1.2, Table 3.8, is $\sim$ 528.5 pN/Å= $\sim$ 766 kcal/mol/nm^2. Since the computational model is a coarse-grained representation of the fibril and not a full atomistic one, the cross-link length is not the real one and the simulation parameters have to be adjusted. The average glucosepane initial length is calculated from an analysis on the data.file: it is 2.7 nm, whereas the physiological glucosepane length is $\sim$ 1.6 nm. The average initial length is the second parameter in the input script, that specifies the length at which the constant bond has to be applied (third value in the input script). The last value indicates the length at which the bond breaks. This value, 1 nm, represents the glucosepane increase in length until failure, and it is calculated from the previous analysis in Sec. 3.1.2.

```
atom_style      angle
read_data       data.fibrillacompleta

bond_style      piecewise
bond_coeff      1 1.756 1208 0.4 6240
bond_coeff          2 1.6 0 2.5 610.7 3.1
bond_coeff          3 2.7 766.0 1
```

Figure 2.21: Input script for bonds type *3*.

For computational reasons, the simulation is performed at a pulling rate of 100 m/s.

Concerning the second distribution, in which the cross-links are distribuited only in the external shell, the first step is the identification of the external molecules in the fibril. The coordinates of the fibril section centre are (Fig.2.22):

$$x_c = 9.5nm \tag{2.30}$$

$$y_c = 45nm \tag{2.31}$$

The radius of the fibril is 10 nm and it has been decided that the external molecules are the molecules that have a radius greater than 6.5 nm. The C++ script analyses all the molecules and it calculates the distance of each one from the centre. If the distance is greater than 6.5 nm the molecule is stored in a list, if the distance is smaller than 6.5 nm the molecule is not stored in the list. The

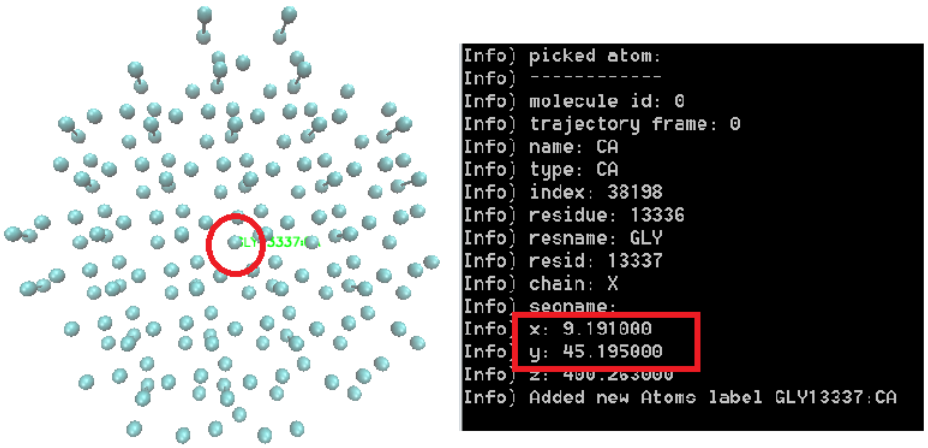


Figure 2.22: Determination of the centre coordinates of the fibril using VMD.

calculation of the distance of each molecule is performed as follow:

$$d_m = \sqrt{(x - x_c)^2 + (y - y_c)^2} \tag{2.32}$$

Where x and y are the coordinates of the analysed fibril while x_c and y_c are the coordinates of the centre. The second step is the choice of 66 fibrils among the external fibrils of the list, using the RANDOM function. The third step is the ramdom choice of 66 beads, 1 bead each fibril. Finally, the bond between each selected bead and the closest one in a neighboring fibril is generated, as described previously. Fig.2.23 shows a screenshot of the fibril with cross-links in the external shell.

The simulation is performed at a pulling rate of 100 m/s due to computational limitations.

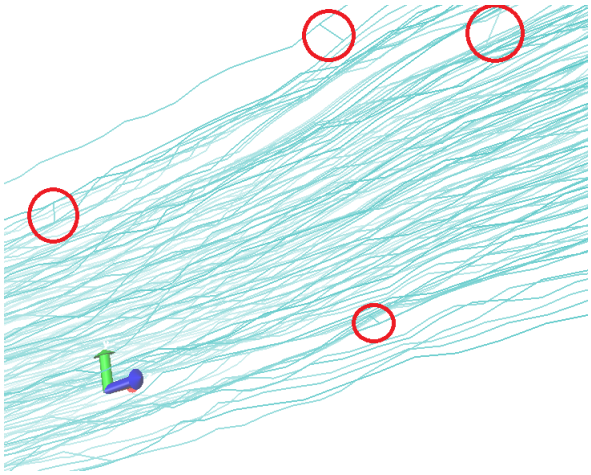


Figure 2.23: Non-enzymatic cross-links in the external shell of the fibril.

2.2 Experimental Validation

2.2.1 Fibrillation

The first aim of the experimental work is the development of a protocol for reconstitution of collagen fibrils. In oder to obtain self-assembled fibrils with consistent diameters (50-400 nm) and visible D-banding several tests have been done. The protocol that yieldest best results is briefly summerized here. Solubilyzed type I collagen derived from rat tail tendon (SERVA, 2 mg/ml, collagen R solution 0.2%), is filtrated sterile and aliquoted at 4°C. The buffer used ((Na_2HPO_4 30mM, KH_2PO_4 10 mM, KCl 200 mM), from[34]) is filtrated sterile and aliquoted at -20°C. A fibrillation solution with a collagen concentration of 30 μg/ml is prepared mixing the same amount of buffer and collagen solution in acetic acid 0.1 N at room temperature (Fig.2.24). To induce fibrillation, samples were incubated in an Eppendorf at 24°C over night (Fig.2.25). Then samples were treated as described in Sec.2.2.2 and fibrils were visualized using the scanning electron microscope (SEM).

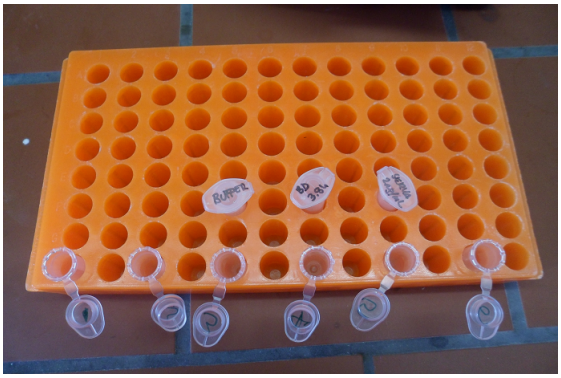


Figure 2.24: Fibrillation set-up.

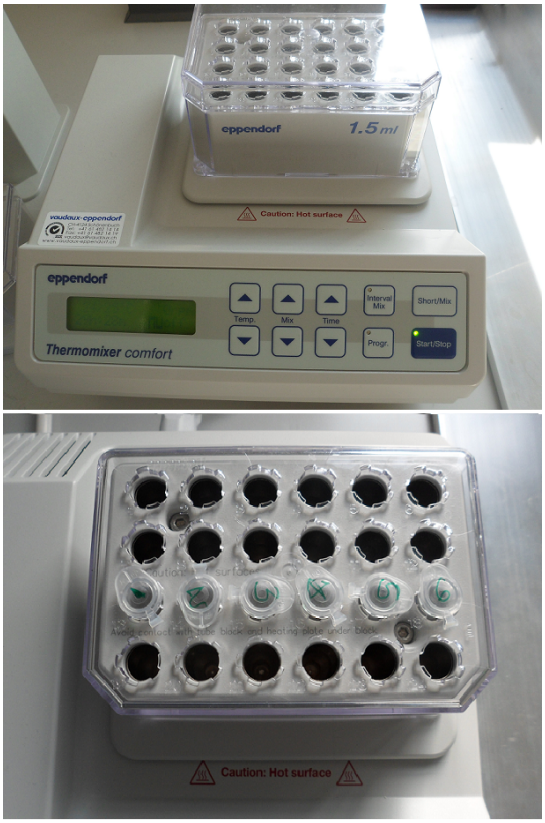


Figure 2.25: Incubation overnight in Eppendorf tubes.

2.2.2 Fibrils isolation from rat tail tendon

In another approach collagen fibrils were isolated from rat tail tendons. Tendons were stored at -20°C. In order to use them, they were first immersed in PBS until defrosted, then they are cut into small pieces of $\sim$ 2 - 3 cm. Using tweezers, the fascicles were extracted from the tendon (Fig.2.26).

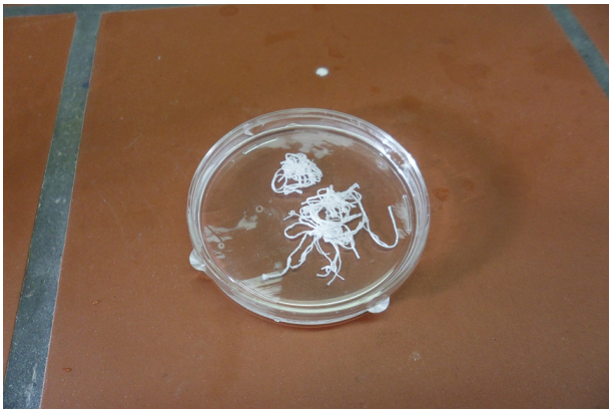


Figure 2.26: Fascicles extracted from rat tail tendon and stored in PBS.

In order to extract fibrils from fascicles, a single fascicle is placed onto a glass with a drop of PBS or milliQ water, so the fascicle does not dehydrate. After this, with tweezers one edge of the fascicle is keep fixed and with a different tweezers the fascicle is gently scratched against the glass (Fig.2.27).

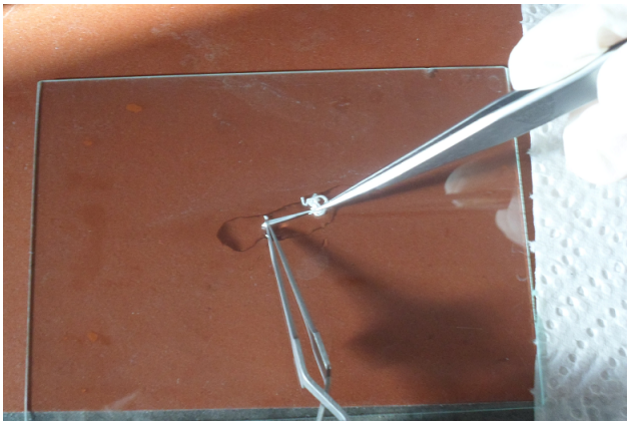


Figure 2.27: Procedure used to extract fibrils from fascicles.

Finally, the drop of PBS or milliQ water that contains fibrils is collected with

a pipette and placed on an Eppendorf tube.

In order to age the fibrils the protocol developed by Li et al.[1] has been adapted for the purpuse of this work. Some of the fascicles were incubated in 20 mM methylglyoxal (MGO, Sigma-Aldrich) in 100 mM EPPS (4-(2-Hydroxyethyl)-1-Piperazinepropanesulfonic acid, Sigma-Aldrich) at 26°C for 6 hours. Then fibrils were extracted as previously described in Sec.2.2.2.

2.2.3 Sample preparation for SEM analysis

The fibril solution previously prepared (from rat tail tendon or from commercial collagen) was used to make samples for SEM analysis. First of all, glasses for SEM were cleaned and coated. In order to clean the glasses, a three step procedure was applied. First, coverslips were immersed in milliQ water and placed in a sonication bath (Emmi 20HC, EMAG Technologies) for 5 minutes. Then, milliQ water was replaced with pure ethanol and glasses were let into the ultrasound chamber for 5 additional minutes. Finally, a last washing with isopropanol was performed (Fig.2.28).

Figure 2.28: Sonication procedure applied to cleaning glasses.

Glasses were let to dry. The cleaning was improved using a plasma machine (Electronic Diener, Fig.2.29).

Finally, they were coated with HMDS (Hexamethyldisilazane, SIGMA) in order to make them hydrophobic: they were placed inside a vacuum chamber with a small amount of HMDS overnight.

In order to prepare the samples for SEM analysis, a drop of 100 μl of solution was placed onto a glass for about 1 hour in order to dry it. After this, the glass was cleaned with milliQ water and placed into a small beaker filled with ethanol. Five

Figure 2.29: Plasma treatment.

steps of 5 minutes each were applied, in each step the concentration of ethanol was varied: first ethanol 10%, then 25%, 50%, 80%, 100% (for 20 minutes instead of 5). Afterwards, ethanol was replaced with HMDS for 30 minutes in order to fix the sample on the glass. Finally, the glass was fixed onto the SEM stub. In order to image the sample with the SEM, it has to be conductive. Samples were coated with platinum using the sputtering machine (Sputter Coater SCD500, BAL-TEC). A coating of 3 nm was performed (Fig.2.30).

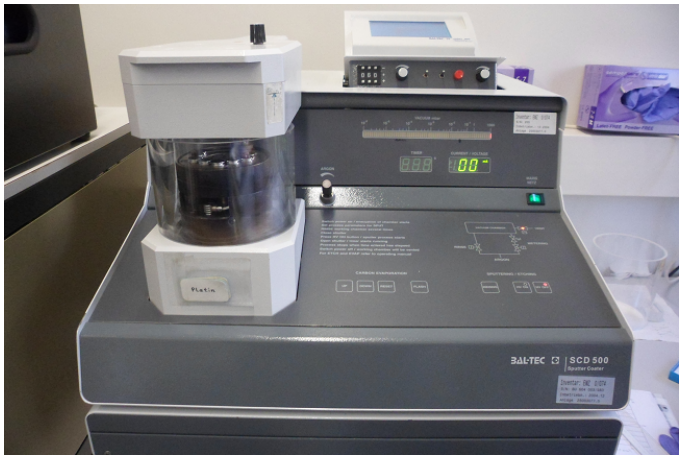


Figure 2.30: Coating of samples in the sputtering machine.

2.2.4 Fibrils visualization under light microscope

In order to isolate single fibrils and perform tensile tests on them using the mechanical device, collagen fibrils must be visualized under optical microscope. Due to the small dimensions of fibrils, an inverted microscope, a fluorescence microscope and a dark field microscope were tested.

Fluorescence microscopy is a technique perfomed with an optical microscope that uses fluorescence instead of transmitted light to generate an image. It must be carry out in a dark chamber and samples have to be fluorescent. To prepare samples with fluorescent fibrils the following protocol was followed. A drop of 20 μl containing fibrils isolated from rat tail tendon and PBS was deposited onto a previously cleaned glass. Then it was washed three times for 5 minutes with PBS and incubated with the primary antibodies (monoclonal anti-collagen type I, mouse IgG1 isotype, SIGMA) for one hour at a 1:1000 or 1:2000 dilution. To prevent evaporation incubations were performed in a humidity chamber. The sample was carefully washed with PBS (3x5 minutes) in order to remove residual antibodies. Then it was incubated for one hour with secondary antibodies (1:50 dilution) (Fluorescein (FITC)-conjugated AffiniPure Donkey Anti-Mouse igG (H+L), JacksonImmunoReserch). They bind primary antibodies and contain a green fluorophore. Finally the sample was washed with PBS (3x5 minutes), mounted with Mowiol mounting medium and visualized using the fluorescence microscope. Several samples were prepared using this protocol and varing the dilution of primary antibodies. To validate the secondary antibodies selective adhesion to primaries, a negative control was prepared without the use of primary antibodies.

The second tested method is the visualization of fibrils under a dark field microscope. Dark field microscopy is a method that excludes the direct illumination of the sample with the use of a sized disc, and allows only the scattered light to hit the sample and produce the image, resulting in a specimen surrounded by a dark field. The Nikon microscope (Nikon Eclipse Ni) was set up in order to obtain dark field illumination. A 20x objective lens was used, and a 6 mm diameter disc was obtained cutting an adhesive black paper and it was placed on the light source in order to exclude the direct illumination of the sample. The sample, containing

fibrils from rat tail tendon, was placed under the lens. The focus was adjusted and the fibrils were visualized.

The last visualization technique involves the use of the inverted microscope: light source and condenser are on the top and the lens are below the stage. The same samples visualized in the dark field were placed on the stage above the lens and are displayed.

2.2.5 Set-up development for the pulling tests

Pulling tests on single fibrils can be performed using AFM method or MEMS devices [18], [19], [22]. In the present work a new and different set-up has been developed. The set-up is primarily made of two components: an inverted microscope (Nikon Eclipse TS100), used to visualize the fibrils on the sample, and a FemtoTool (FemtoTools, FT-FS10000). The FemtoTool is a high precision micromechanical testing instrument that can provide valuable insights into mechanical properties, structure and functionality of small samples. It is equipped with sensible probes capable of measuring forces from single nanoNewton to milliNetwons. The FemtoTools is placed on a platform integrated to the microscope (Fig.2.31). In a first approach the set-up was validated using a needle instead of a sensor.

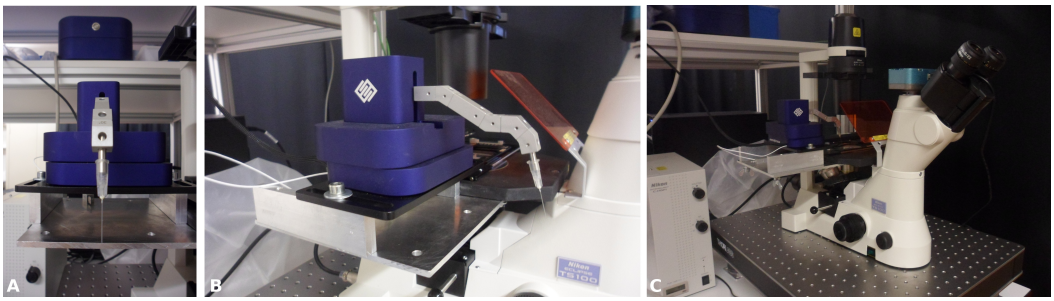


Figure 2.31: A,B: Two different views of FemtoTool system equipped with the needle-sphere. C: The FemtoTool is integrated with the inverted microscope.

As first step in the set-up development, a sphere (Duke Scientific Corporation, 100 μm, Dry Soda Lime Glass) was attached to the needle using a two component epoxy glue (Araldite®Rapid 2 x 15 ml tube). Spheres were placed on a silicon substrate and visualized under the inverted microscope using the 10x objective lens. The glue was prepared and placed near the sample; the FemtoTools was

firstly driven onto the glue and a small drop was picked up, and then it was drived on the sphere. It was carefully set down untill contact and let cure for 15 minutes (Fig.2.32).

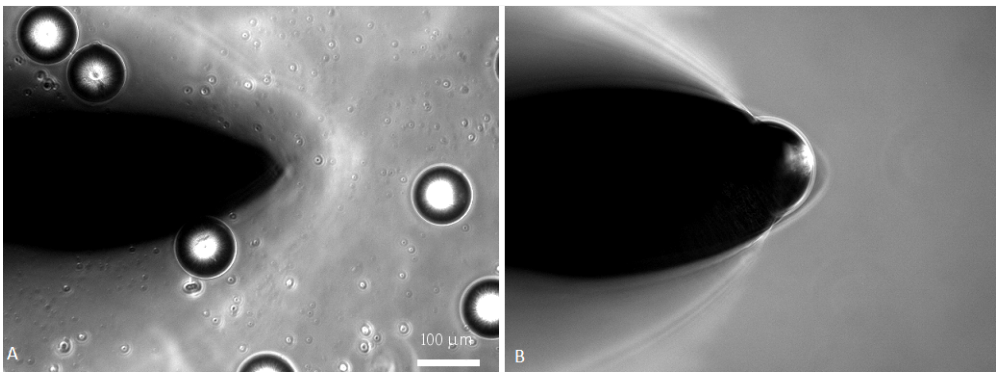


Figure 2.32: A: the needle (black shadow) is drived above the spheres placed on the sample; B: the sphere is attached to the needle.

As second step, a fibril was selected and it was pulled using the needle-sphere component. Fibrils were extracted from rat tail fascicles with the procedure previously described and they were placed into a PDMS sustrate. This subtrate has been proved to do not stick the fibrils that can consequently be pulled out from it. Fibrils were visualized under the inverted microscope using a 20x objective lens and a long one was selected and centered. After this, the FemtoTool was driven under the microscope light and the needle-sphere component was visualized using first a 2x objective lens, then a 10x and finally the 20x one. The needle-sphere component was driven above an edge of the selected fibril and the x-y-z coordinates were registered in order to later find the correct position. A little quantity of epoxy glue was prepared mixing the two components together for about 30 seconds and it was placed near the sample. The FemtoTool tip was driven above and then immersed into the drop in order to catch a small amount of glue with the sphere. Then, the sphere with the glue was driven above the sample using the coordinates previously registered and the glue was attatched to the fibril. This procedure was repeted for both the edges of the fibril. The glue was allowed to cure for about 15 minutes (Fig.2.33).

Fresh glue was then prepared and placed near the sample; the needle-sphere component was driven above it and a small quantity of glue was picked up. The

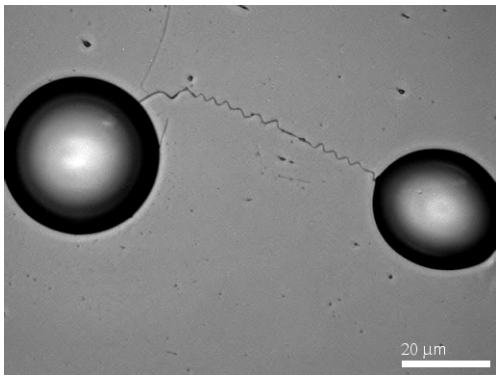


Figure 2.33: Both edges of the fibril are glued to the substate.

tip was driven above the drop of cured glue previously placed on an edge of the fibril, it was set down until contact and let cured for about 20 minutes. At the end of this procedure the selected fibril had an edge glued to the silicon substrate and the other one glued to the substrate and to the needle (Fig.2.34).

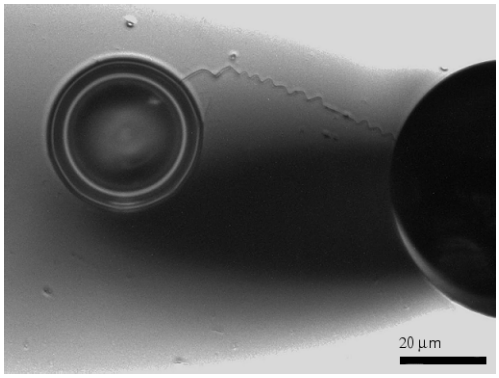


Figure 2.34: Both edges of the fibril are glued to the substate and the needle-sphere component (black shadow) is glued to one edge.

The tip was carefully raised and the edge of the fibril was gently detached from the silicon substrate. Then it was gently moved above the glue of drop which was still stick to the substrate in order to position the fibril perpendicular to the substrate. This procedure must be performed at low velocity since the fibril has not to be damaged. Finally the tip probe was lifted vertically and the pulling test was performed.

2.2.6 Mechanical tests

Since the set-up has been validated, mechanical tests were performed on single fibrils using a force probe (FT-S100, FemtoTools) and a different arm configuration, Fig.2.35.

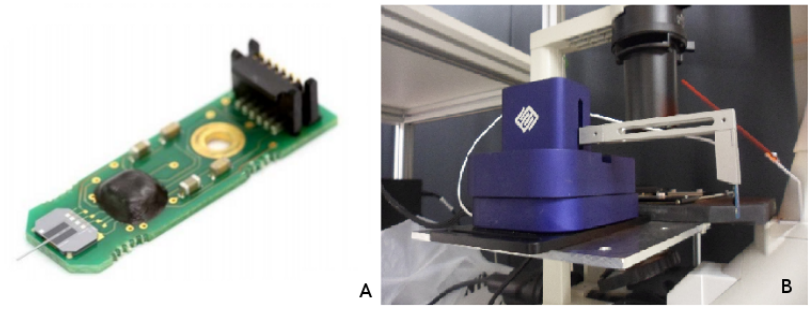


Figure 2.35: A: the real probe, B: new arm configuration.

Four different fibrils, two physiological and two aged, were choosen and glued onto the substrate as described in the previous section (Fig.2.36). In Tab.2.1 their features are reported.

Name	Type	Length [μm]
Fibril A	physiological	14.6
Fibril B	physiological	16.1
Fibril C	aged	12.3
Fibril D	aged	22.0

Table 2.1: Fibrils features.

Pulling tests were performed at 0.5 μm/s and data were recorded using the FemtoTools FT-WF502-CT software, version 1.0.3. Data were elaborated and Force - Length, Stress - Strain charts and elastic moduli are reported in Sec.3.2.5.

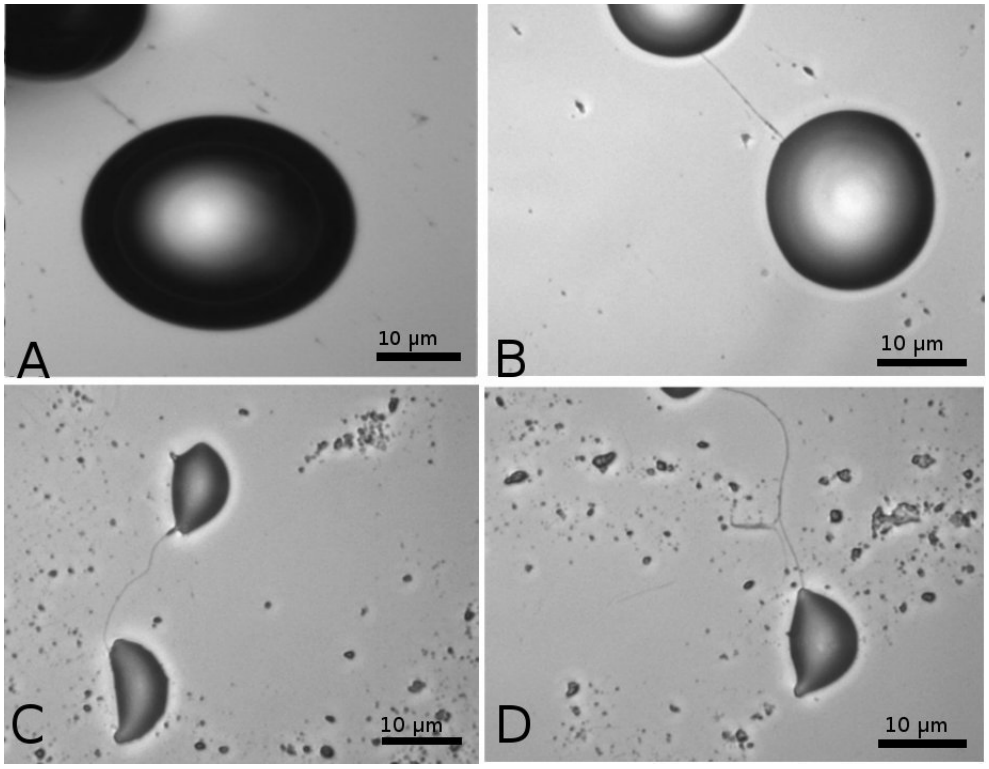


Figure 2.36: Fibrils: A-B physiological fibrils, C-D aged fibrils.

Chapter 3

RESULTS

3.1 Molecular Dynamics

3.1.1 Mechanical characterization of enzymatic cross-links

Pulling tests on hydroxylysyl pyridonoline are performed using the Reactive Force Field of LAMMPS. The enzymatic cross-link is stretched at different pulling rates (100, 10, 1, 0.5, 0.1, 0.05 m/s). The smd.force file provides the evolution in time of the molecule motion and here the Force - Length chart in direction 2-12 with a pulling rate of 1 m/s is reported (Fig.3.1). In Appendix A all the charts are reported. In the chart it is possible to identify three different parts. In the first one the cross-link rearranges its molecular structure and the force is null, when hydroxylysyl pyridinoline is linear force starts increasing untill it reachs the failure point, the molecule chain breaks and the force becomes zero again. This three-stage behavior is present in all the tests performed. Values of force and length at failure for each pulling direction and each velocity are collected and reported in the following tables (Tab.3.1, Tab.3.2, Tab.3.3)

Direction 2-7 better resists streaching as the chain of bond pulled is shorter than the one in 2-12 or in 7-12 and, since the bond is represented as series of springs, a small number of springs leads to a major stiffness.

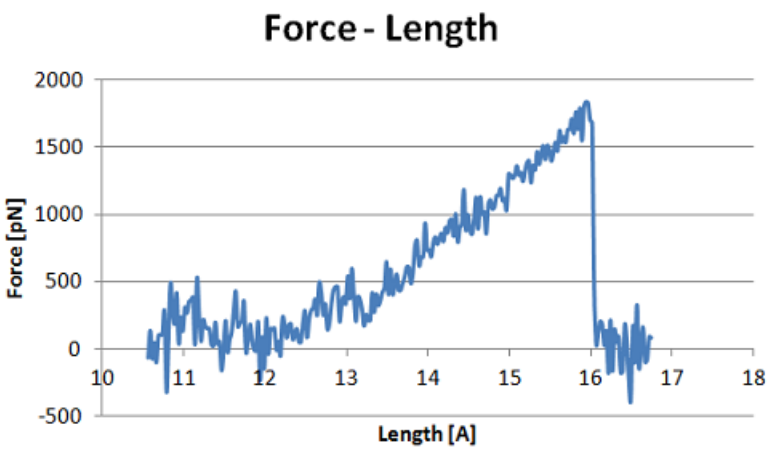


Figure 3.1: Force - Length chart of hydroxylysyl pyridinoline streched in 2-12 direction at 1 m/s.

direction 2-12		
rate [m/s]	Force at failure [pN]	Length at failure [Å]
100	2562.2	17.2
10	2128.5	16.6
1	1801.5	15.9
0.5	1553.2	15.4
0.1	1489.2	15.2
0.05	5700.8	22.4

Table 3.1: Values of force and length at failure for 2-12 pulling direction at 100, 10, 1, 0.5, 0.1, 0.05 m/s.

direction 7-12		
rate [m/s]	Force at failure[pN]	Length at failure[Å]
100	2862.9	17.1
10	2422.5	16.5
1	2172.8	15.7
0.5	1895.9	15.4
0.1	1458.9	14.4
0.05	2138.4	15.6

Table 3.2: Values of force and length at failure for 7-12 pulling direction at 100, 10, 1, 0.5, 0.1, 0.05 m/s.

direction 2-7		
rate [m/s]	Force at failure [pN]	Length at failure [Å]
100	7288.1	19.7
10	7223.5	19.9
1	6601.1	18.4
0.5	6252.9	17.7
0.1	6174.4	17.7
0.05	6070.4	17.6

Table 3.3: Values of force and length at failure for 2-7 pulling direction at 100, 10, 1, 0.5, 0.1, 0.05 m/s.

The spring constant, defined as:

$$K = \frac{\Delta F}{\Delta L} \tag{3.1}$$

is calculated using the linear interpolation as shown in Fig.3.2.

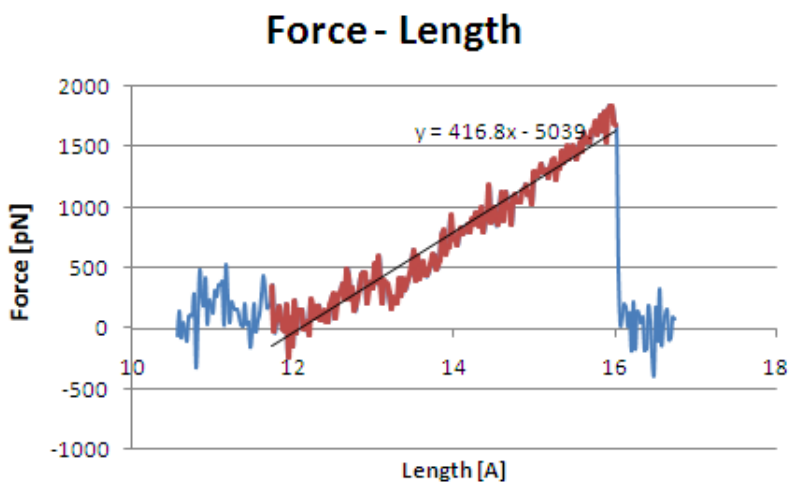


Figure 3.2: Force - Length chart of hydroxylysyl pyridinoline streched in 2-12 direction at 1 m/s, the linear interpolation provides the spring constant.

In Tab.3.4 the spring constants are reported for all the tests.

spring constant [pN/Å]			
rate [m/s]	direction 2-12	direction 7-12	direction 2-7
100	369.2	473.6	543.8
10	337.7	455.6	519.1
1	416.8	446.4179	541.8
0.5	287.4	426.8	571.2
0.1	227.6	372.6	538.6
0.05	471.3	438.5	511.3

Table 3.4: Values of spring constant for direction 2-12, 7-12 and 2-7 at pulling rate of 100, 10, 1, 0.5, 0.1, 0.05 m/s.

3.1.2 Mechanical characterization of non-enzymatic cross-links

Same analysis are repeated for the glucosepane. Pulling tests are performed using the RFF of LAMMPS on three different structures of glucosepane at different pulling rate. After the stretching simulation, charts of force versus length of glucosepane are elabotated from the smd.force file (Appendix A). Here the chart for structure 1 at 1 m/s is reported, it presents the same three-stages behavior observed before for the hydroxylysyl pyridinoline: the molecule first unrolls then it is streched and it breaks (Fig.3.3).

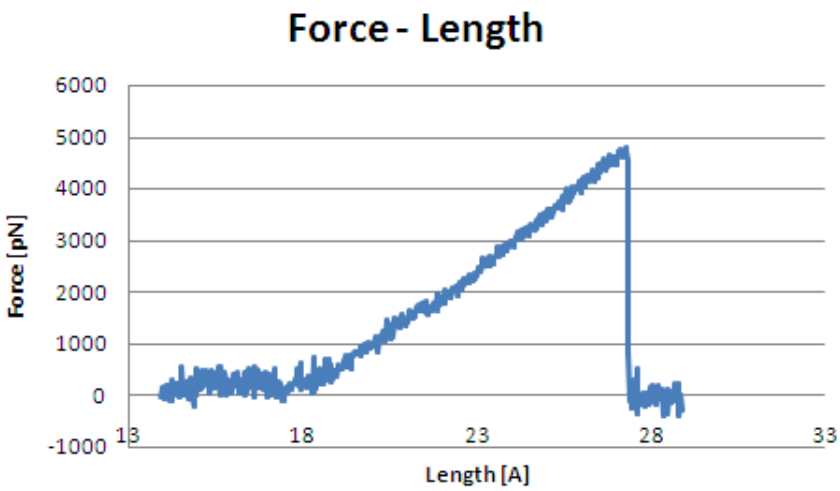


Figure 3.3: Force - Length chart of glucosepane (structure 1) streched at 1 m/s.

structure 1		
rate [m/s]	Force at failure [pN]	Length at failure [Å]
100	4267.4	26.3
10	4198.0	26.1
1	4802.3	27.3
0.5	4828.1	27.3
0.1	3985.1	23.0
0.05	3627.3	22.7

Table 3.5: Values of force and length at failure for structure 1 at 100, 10, 1, 0.5, 0.1, 0.05 m/s.

structure 2		
rate [m/s]	Force at failure [pN]	Length at failure [Å]
100	4132.1	25.9
10	4977.1	27.5
1	4878.5	27.4
0.5	4769.7	27.3
0.1	3971.6	22.9
0.05	4618.3	26.8

Table 3.6: Values of force and length at failure for structure 2 at 100, 10, 1, 0.5, 0.1, 0.05 m/s.

structure 3		
rate [m/s]	Force at failure [pN]	Length at failure [Å]
100	4638.2	22.8
10	5030.2	27.7
1	4651.5	27.1
0.5	3889.2	22.9
0.1	3892.1	22.9
0.05	4558.7	26.9

Table 3.7: Values of force and length at failure for structure 3 at 100, 10, 1, 0.5, 0.1, 0.05 m/s.

Values of force and length at failure are collected and summarized in Tables 3.5, 3.6, 3.7. The spring constant is calculated using the linear interpolation tool (Fig.3.4). In Tab.3.8 values of spring constant are reported for all the tests.

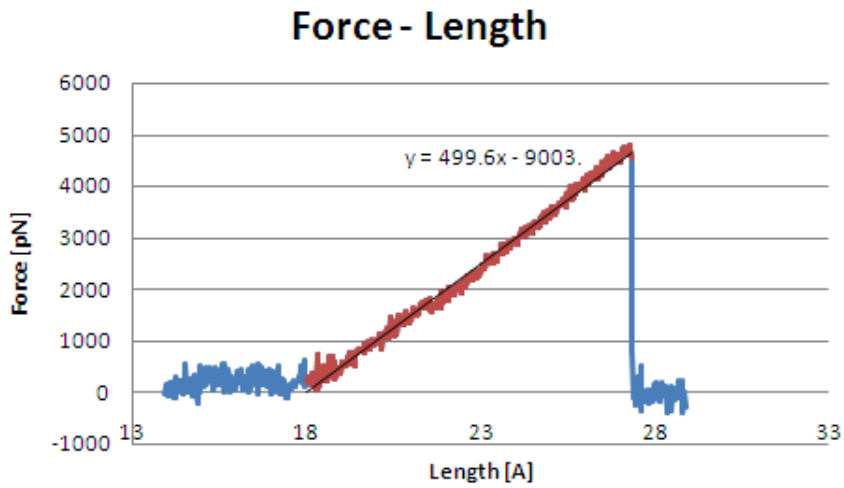


Figure 3.4: Force-length chart of glucosepane (structure 1) streched at 1 m/s, the spring constant is evaluated using the linear interpolation tool.

spring constant [pN/Å]				
rate [m/s]	structure 1	structure 2	structure 3	average
100	543.4	525.1	517.1	528.5
10	538.7	550.7	545.6	545.0
1	499.6	502.2	554.40	518.9
0.5	507.9	531.4	477.5	505.6
0.1	524.4	498.8	502.2	508.5
0.05	507.1	503.9	516.1	509.0

Table 3.8: Values of spring constant for structure 1, 2 and 3 at pulling rate of 100, 10, 1, 0.5, 0.1, 0.05 m/s.

In order to validate the results obtained for glucosepane and hydroxylysyl pyridonoline a literature research is done. Several techniques as scanning force microscope, microneedles and AFM cantilevers, can be used to evaluate single molecule mechanical properties [35]. Grandbois et al. [36] evaluate the rupture

force of single covalent bond with the atomic force microscope (AFM). They mechanically stretched a polysaccaride molecule attached between the surface and the AFM tip as shown in Fig.3.5, they found out that the bond rupture occoured at 1000-2000 pN. Values of force at failure calculated in this work for enzymatic and non-enzymatic crosslinks are in the range of 1000-5000 pN and they agree with Grandbois's results. No values of rupture length and constant spring have been found in literature as they are specific for each kind of molecule and glucosepane and hydroxylysyl pyridonoline have never been tested mechanically in laboratory.

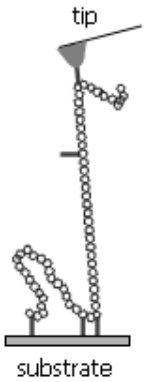


Figure 3.5: Adapted from Grandbois et al.[36], stretching of a single polysaccharide chain that is attached to the tip and the substrate.

3.1.3 Mechanical characterization of a tropocollagen molecule

The peptide chain has been stretched at a velocity of 1 m/s. The values of stiffness obtained from the stretching curve of the full atomistic models, used also in the coarse-grained model, are:

$$K_1 = 70\frac{kcal}{molA^2} \tag{3.2}$$

$$K_2 = 362\frac{kcal}{molA^2} \tag{3.3}$$

Fig.3.6 shows the results obtained for the pulling of the peptide at a rate of 1 m/s. The curve shows a non linear behaviour, as expected from the use of the *piecewise* bond style. In particular, two stiffness values can be individuated: the peptide

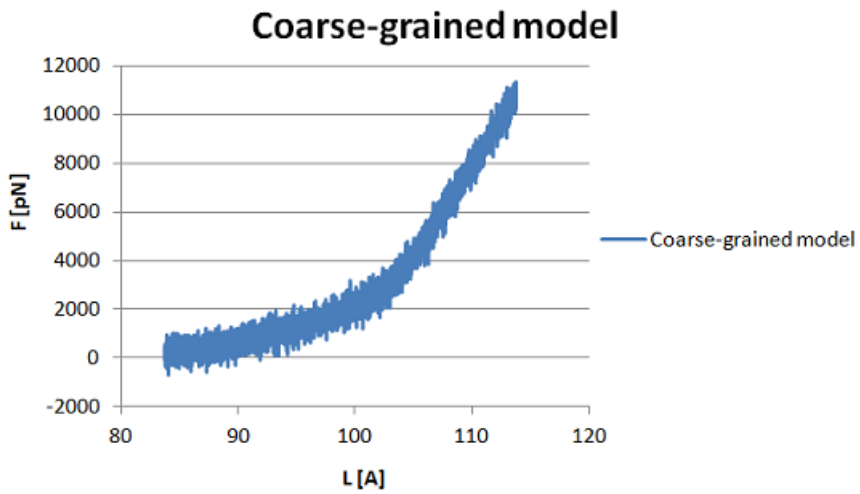


Figure 3.6: Force - Length curve for the coarse-grained model.

presents a low rigidity during the first part of the pulling, while in the second part the peptide rigidity increases. The curve obtained is compared with the curve for the full atomistic model. The comparison is shown in Fig.3.7. The two curves

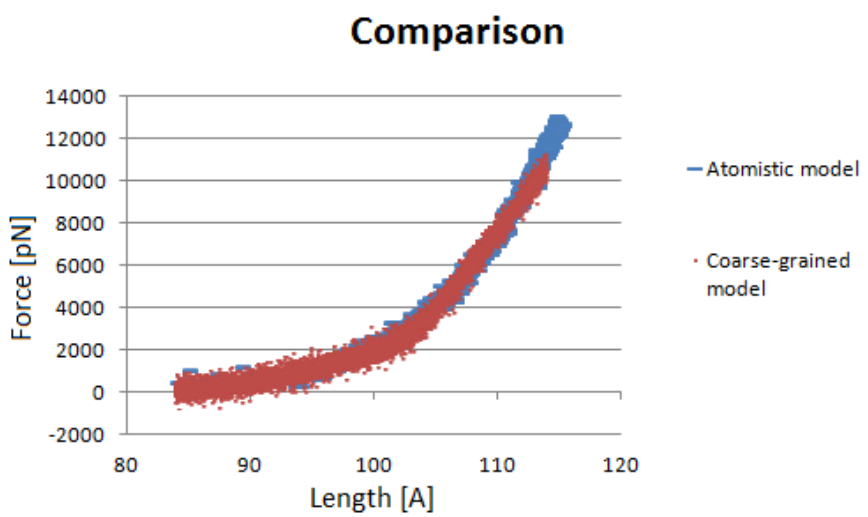


Figure 3.7: Comparison between the two models.

match. It is possible to conclude that the coarse-grained model is validated and it can be used for the purpose of this work. It allows to decrease the computational cost of the system and it well describes the mechanical properties of the model.

The persistence length has been used in order to evaluate the angle constant of the model. Different simulations are performed using different angle constants and the simulation that allows to obtain a persistence length close to $\sim$ 23.4 nm is choosen. In Fig.3.8 the input script used for the simulation is reported. The

```
units           real
boundary        f f f

atom_style      angle
read_data       data.angle

bond_style      piecewise
bond_coeff      * 3.04 70 0.5 362

angle_style     harmonic
angle_coeff     * 17 180

pair_style      lj/cut 0
pair_coeff      * * 1.0 1.0 2.5
```

Figure 3.8: Input script used for the persistence length calculation.

matlab script used to evaluate the persistence length gives as output values that are elaborated in a graph. The inverse of the curve slope is the persistence length evaluated in Å (Fig.3.9).

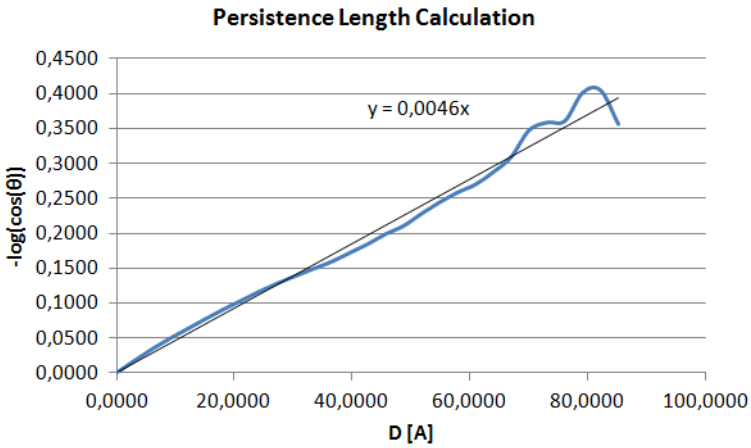


Figure 3.9: Curve obtained elaborating the values from matlab.

The slope of the curve is:

$$y = 0.0046 \ \mathring{A} \tag{3.4}$$

The persistence length is evaluated from the slope in the following method:

$$L_p = \frac{1}{0.0046} \text{Å} = 217 \text{Å} = 21.7nm \tag{3.5}$$

The value obtained is close to the goal. The angle constant is:

$$K_a = 17 \frac{kcal}{mol \cdot rad^2} \tag{3.6}$$

Buehler in [6] found a angle constant of 14.98 kcal/mol/rad^2.

Concerning the tropocollagen molecule, a pulling simulation with a rate of 1 m/s is performed. The results obtained from the simulation are compared with the full atomistic model. The comparison is shown in Fig.3.10. The two curves

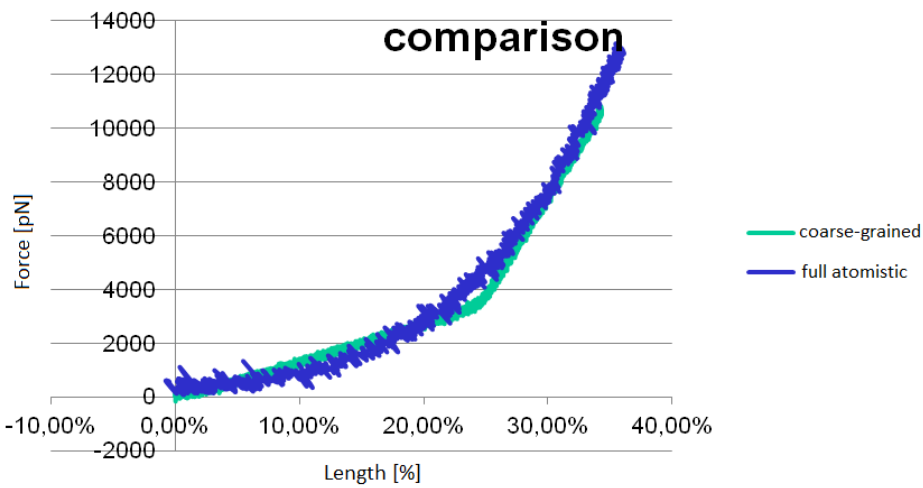


Figure 3.10: Comparison between the coarse-grained and the full atomistic models for the tropocollagen.

substantially fitted so the bond constants are correct. Fig.3.11 shows the curve obtained using the matlab data elaboration. The persistence length is calculated as follow:

$$L_p = \frac{1}{0.0414} nm = 24.2nm \tag{3.7}$$

This value is very close to the goal.

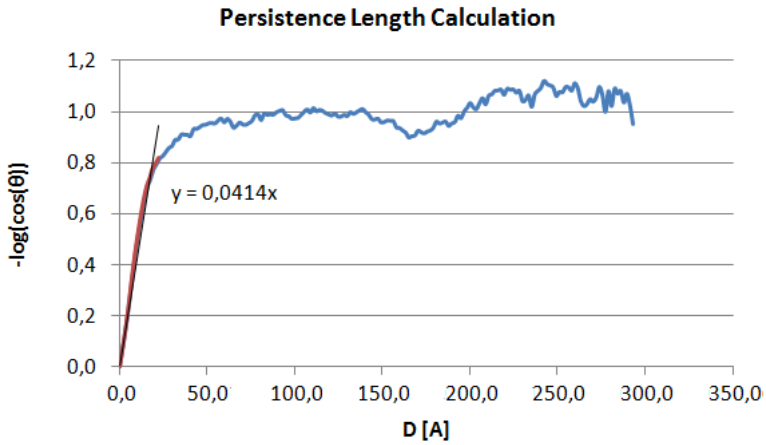


Figure 3.11: Calculation of the persistence length.

3.1.4 Lennard Jones parameters

A simulation until equilibrium has been perfomed in order to verify if the calculated parameters for the Lennard Jones potential are correct. Fig.3.12 shows the input script for the simulation, where:

$$r = 16nm \tag{3.8}$$

$$\sigma = 1.425nm \tag{3.9}$$

$$\varepsilon = 7.0\frac{kcal}{mol} \tag{3.10}$$

```
atom_style full
read_data  data.fibrilla

bond_style      piecewise
bond_coeff      * 1.756 1208 0.4 6240

angle_style     harmonic
angle_coeff     * 4.0 180

pair_style lj/cut 16
pair_coeff * * 7 1.425
```

Figure 3.12: Input script for the Lennard Jones simulation

At the end of the equilibration process the distance among the tropocollagen

molecules reaches a value close to the physiological one, as shown in Fig.3.13.

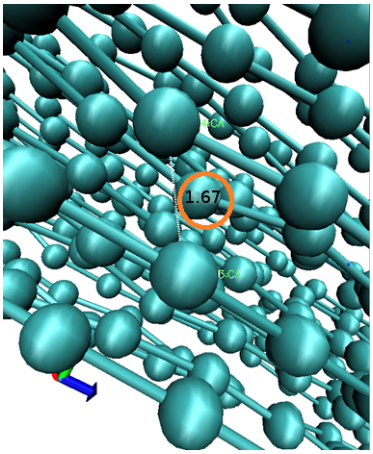


Figure 3.13: Screenshot of the simulation. Note that the distance among the molecules is the expected one.

3.1.5 Physiological fibril

The pseudo physiological fibril, that does not present cross-links, is stretched at 100 m/s then the chart of force versus length is obtained from the smd.force file and displayed in Fig.3.14. In order to describe the mechanical responce of the fibril, the chart is analyzed in relation to the dynamic trajectory using VMD. During the first part of the pulling test (length<1141 nm) shear forces, defined by Lennard Jones potential, and the covalent bond interactions inside tropocollagen molecules withstand the stretching and mutually contribute to the increase of force. For length>1141 nm a compact bunch of tropocollagen molecules is gradually pulled out, the force begins to decrease and the fibril is damaged. Peak values are led by the unravel of the structure and are:

$$F_{max} = 139nN \tag{3.11}$$

$$L_{max} = 1141nm \tag{3.12}$$

Rising and descending slopes are approximately linear and similar as shear forces and bond interactions are still strong. For length=1267 nm the bunch of tropocol-

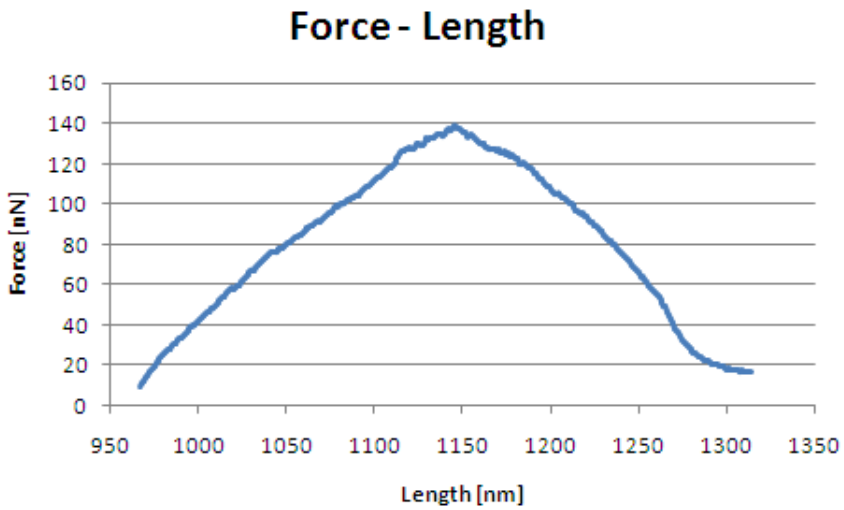


Figure 3.14: Force - Length chart for the pseudo physiological fibril stretched at 100 m/s.

lagen molecules is completely pulled out as shown in Fig.3.15 and the force rapidly decrease until it reachs negligible values. The use of NRFF does not allow the bond rupture inside tropocollagen molecules and it induces some alteration during the rupture phase: even when the fibril fails completely the force remains at a constant value around 16.2 nN since the extracted molecules still have inertia forces.

Figure 3.15: VMD image of the rupture of pseudo physiological fibril.

Then values of stress and strain are calculated using the following formulas:

$$\sigma = \frac{F}{A} = \frac{F[nN]}{10^2\pi[nm^2]} = \frac{F}{\pi}10^{-2}[GPa] \tag{3.13}$$

$$\varepsilon = \frac{l - l_0}{l_0} \tag{3.14}$$

where A is the cross-section area and l_0 is the initial length of the fibril. In Fig.3.16 is reported the Stress - Strain chart. Maximum values of stress and strain are:

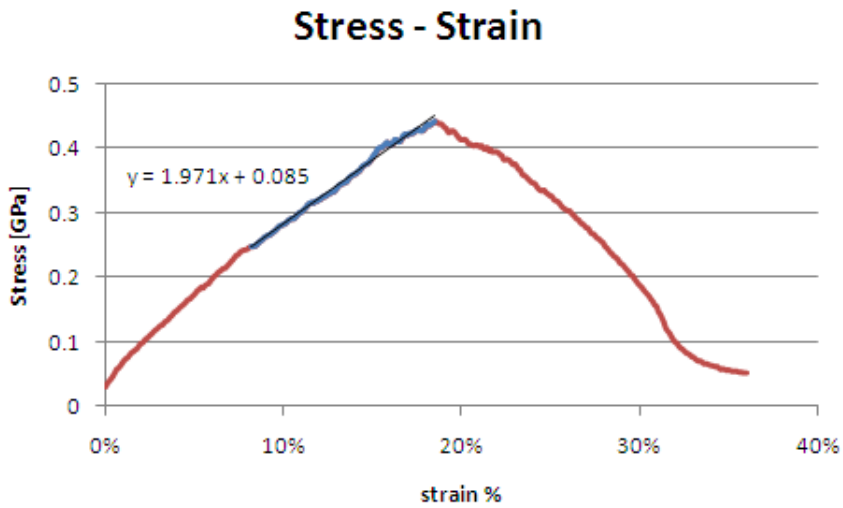


Figure 3.16: Stress - Strain chart for pseudo physiological fibril stretched at 100 m/s. The elastic modulus is evaluated using the linear interpolation tool.

$$\sigma_{max} = 0.442GPa \tag{3.15}$$

$$\varepsilon_{max} = 18.51\% \tag{3.16}$$

The elastic modulus is evaluated for values of strain between 8% and 19% as shown in Fig.3.16 using the linear interpolation tool and it is about 1.971 GPa.

In order to compare the effect of enzymatic cross-links to the one determined only by shear forces and covalent bond interactions inside tropocollagen molecules, a physiological fibril that contains the natural amount of enzymatic cross-links is stretched at 100m/s. The Force - Length chart is reported in Fig.3.17. The analysis of trajectory is performed using the output dcd file and VMD and Force - Length chart is described in relation to it. During the first part of the pulling test (length$<$1043 nm) enzymatic cross-links are not stretched and only tropocollagen bonds and non bond interactions among beads induce the increase in force. Then cross-links are gradually stretched and start to bear the load. As for the pseudo physiological fibril, the peak of force is determined by the unravel of the fibril: from the analysis of trajectory it is found that force begins to decrease when a compact bundle of tropocollagens starts to be pulled out. Peak values of force and

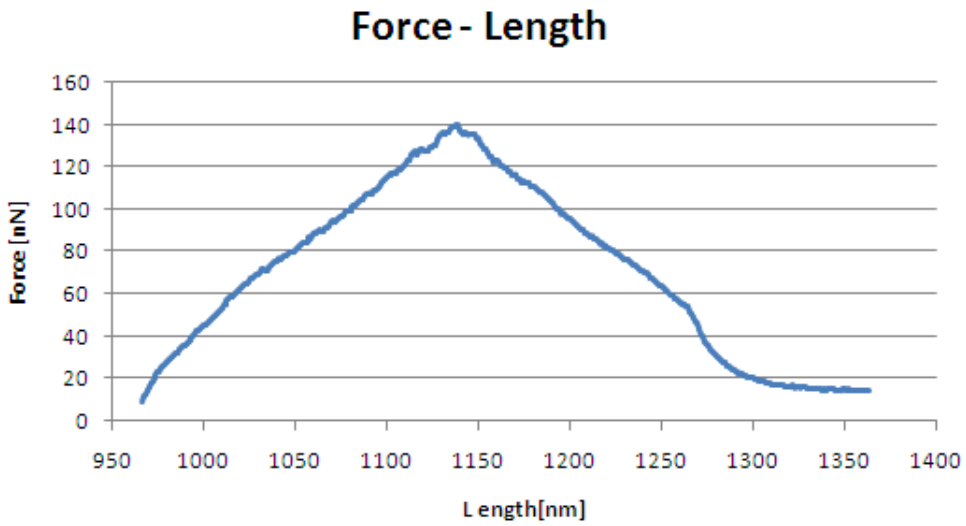


Figure 3.17: Force - Length chart for physiological fibril stretched at 100 m/s.

length are:

$$F_{max} = 140nN \tag{3.17}$$

$$L_{max} = 1138nm \tag{3.18}$$

For length>1138 nm cross-links are still stressed untill they reach rupture, meanwhile tropocollagen molecules shear reciprocally, non bond interactions diminish their effect then, for length=1276 nm, the fibril completely fails. From the dcd analysis the clear rupture can not be visualized as cross-links are still represented in VMD frames even if they do not contribute anymore. After rupture the force remains at a constant value around 14.5 nN as tropocollagen molecules are still stretched.

Values of stress and strain are calculated using the formulas (3.13) and (3.14) and the Stress - Strain chart is elaborated and displayed in Fig.3.18.

Maximum values of stress and strain are:

$$\sigma_{max} = 0.445GPa \tag{3.19}$$

$$\varepsilon_{max} = 17.72\% \tag{3.20}$$

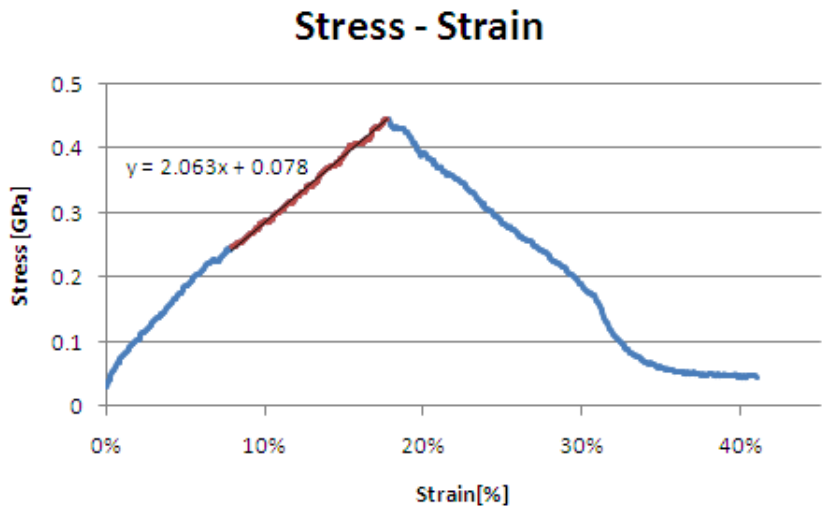


Figure 3.18: Stress - Strain chart for physiological fibril stretched at 100 m/s.

The elastic modulus is evaluated as previously described and is about 2.063 GPa.

3.1.6 Aged fibril

In order to study the influence of cross-links distribution on the fibril section two different models are created. In the first one, cross-links are distribuited through the whole section of the fibril. A pulling test is performed at a rate of 100 m/s. The smd.force file contains the output data that are elaborated with the use of a sheet-spread in order to obtain the Force - Length and Stress - Strain curves. Fig.3.19 reports the Force - Length curve. Trajectory analysis is performed as described before. For length$<$1028 nm tropocollagen bonds and non bond interactions simultaneously contribute to the increase of the force while cross-links do not participate. For length$>$1028 nm both enzymatic and non enzymatic cross-links are stretched. As for the models analyzed before, the peak of force is determined by the unravel of the fibril and maximum values are:

$$F_{max} = 161nN \tag{3.21}$$

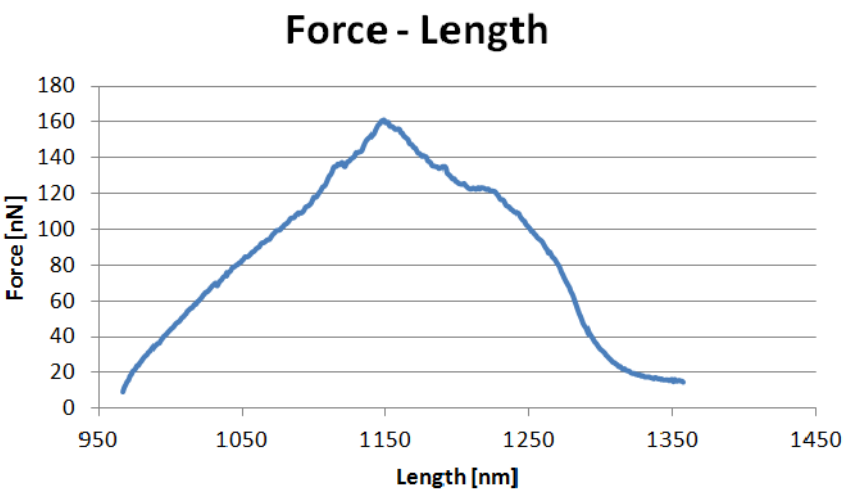


Figure 3.19: Force - Length chart for aged fibril with cross-links through the whole section, stretched at 100 m/s.

$$L_{max} = 1149nm \tag{3.22}$$

In the second part of the chart the force presents some oscillations since non enzymatic cross-links progressively fail. The final rupture of the fibril occours for length = 1291 nm and the residual force is 16.6 nN.

Then the data are elaborated in order to obtain a Stress - Strain curve, as described in the previous section. Fig.3.20 shows the obtained results. Maximum values of stress and strain are:

$$\sigma_{max} = 0.512GPa \tag{3.23}$$

$$\varepsilon_{max} = 18.82\% \tag{3.24}$$

The elastic modulus is calculated as linear interpolation of the first part of the curve, it has a value of 2.391 GPa.

The second fibril model contains cross-links only in the external shell. The pulling test at 100 m/s is performed, data are analysed and the curves are reported in Fig.3.21 and Fig.3.22. The simulation did not finish, from the output file analisys it is inferred that this computational error is to due the expulsion of a bead from the structure during the simulation. This error could be overcome decreas-

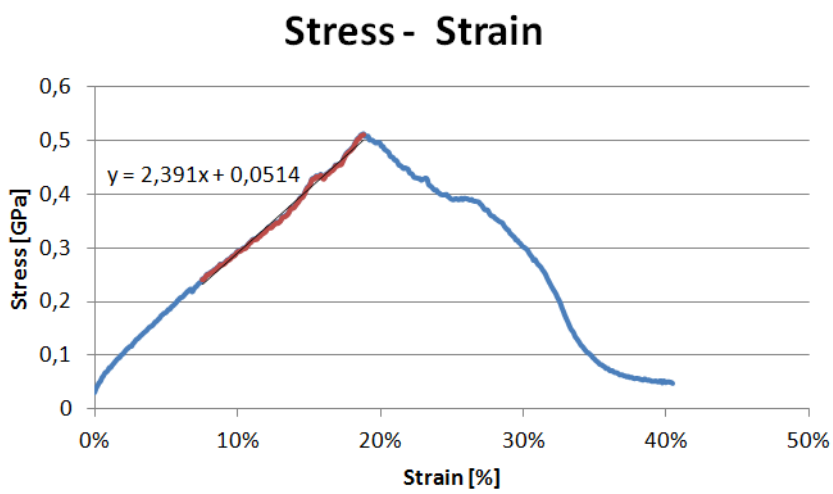


Figure 3.20: Stress - Strain chart for aged fibril with cross-links through the whole section, stretched at 100 m/s.

ing the time step from 10 fs to lower values but it would lead to an increase of computatial time that is prohibitive for the current work. The trajectory analysis is performed as described before, for length<1042 nm cross-links do not bear the load but they start the stretching only for superior length. The peak is due to the unravel of the fibril and the values of force and length are:

$$F_{max} = 168nN \tag{3.25}$$

$$L_{max} = 1157nm \tag{3.26}$$

Concernig the Stress - Strain chart, the maximum values are:

$$\sigma_{max} = 0.539GPa \tag{3.27}$$

$$\varepsilon_{max} = 19.96\% \tag{3.28}$$

The elastic modulus has been calculated as linear interpolation of the first part of the Stress - Strain curve as described before. The value is 2.467 GPa.

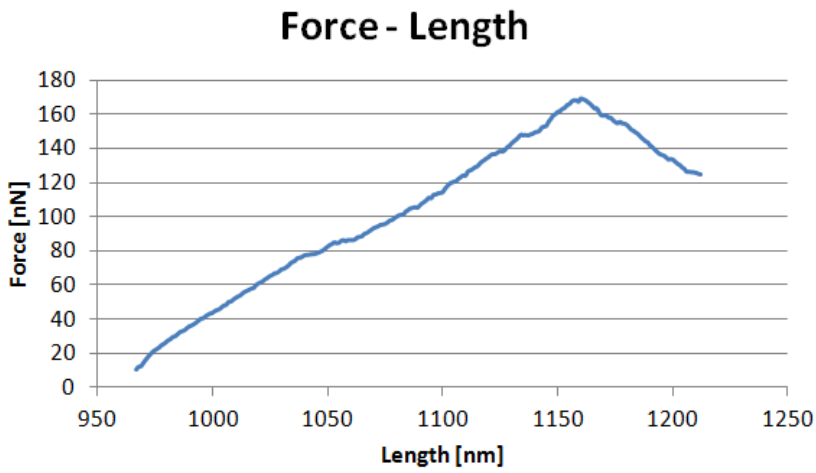


Figure 3.21: Force - Length chart for aged fibril with cross-links in the external shell, stretched at 100 m/s.

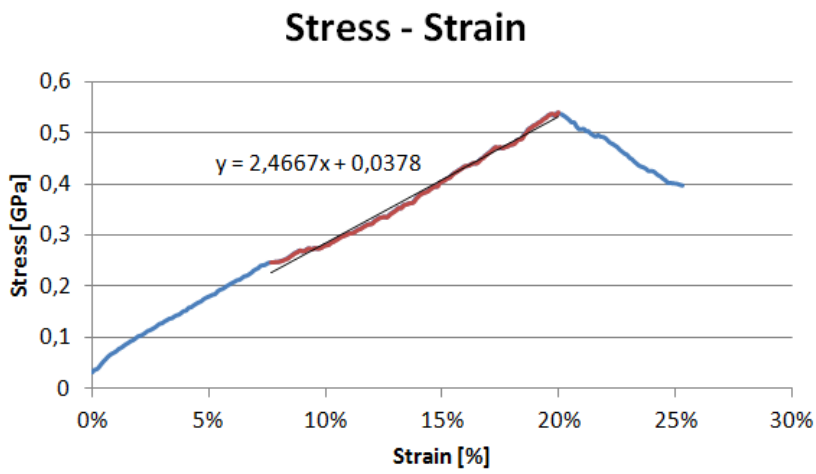


Figure 3.22: Stress - Strain chart for aged fibril with cross-links in the external shell, stretched at 100 m/s.

3.2 Experimental Validation

3.2.1 Fibrillation

Type I collagen fibrils were self-assembled in Eppendorf tubes from a solution containing 30 μg/ml of solubilyzed type I collagen (SERVA). After the reconstitution process SEM images of fibrils were acquired using Zeiss SUPRA 50 VP with InLens detector. Sporadical spaced fibrils with considerable diameter and D-period are found in the samples.

In Fig.3.23.A a portion of a fibril with a diameter of about 425.5 nm and D-banding clearly visible is shown. A D-banding evaluation is shown in Fig.3.23.B, it results lower than 67 nm probably due to the fact that the measure is done manually and could not be precise.

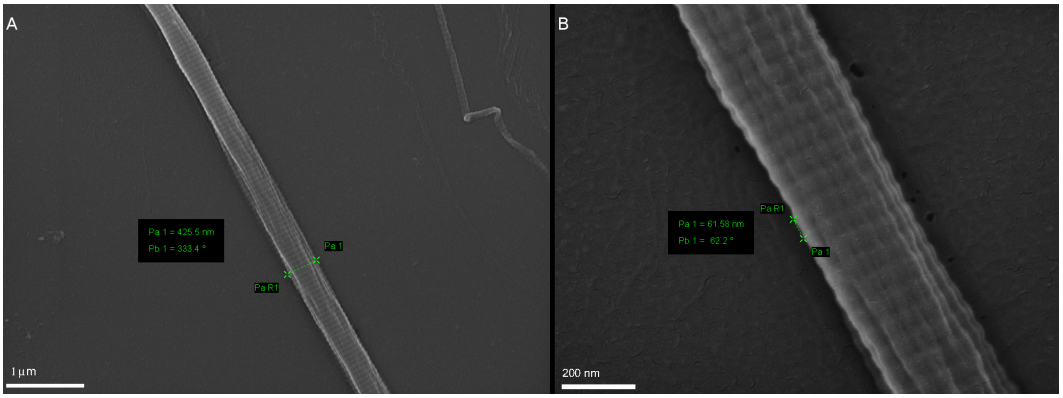


Figure 3.23: SEM images of self-assembled fibrils. A: on the right edge a microfibril that does not complete the fibrillation process can be visualized. B: diameter evaluation.

In order to validate the recostitution protocol shown in Sec.2.2.1 several tests were performed. As the reproducibility is strongly determined by temperature, pH and collagen concentration and not every test was successful, the mechanical tests were carried out on fibrils extracted of rat tail-tendons.

3.2.2 Fibrils isolation from rat tail tendon

Fibrils were isolated from rat tail tendon as previously described (Sec.2.2.2). In order to evaluate if the isolation procedure was successful, the samples were anal-

ysed under the SEM microscope. The sample contained single fibrils on the substrate. The fibrils presented D-banding, it was calculated using a SEM tool and the values obtained are in the range of 64-68 nm. The diameters of the fibrils were between 100 and 250 nm. The fibrils length was very variable, among 15 and 210 μm. The huge variability among the fibrils is probably due to the technique of isolation from the rat fascicle: the fibrils were pulled out from the fascicle using tweezers and this methods could cause the isolation of fibril fragments. Images from SEM analysis are shown in Fig.3.24.

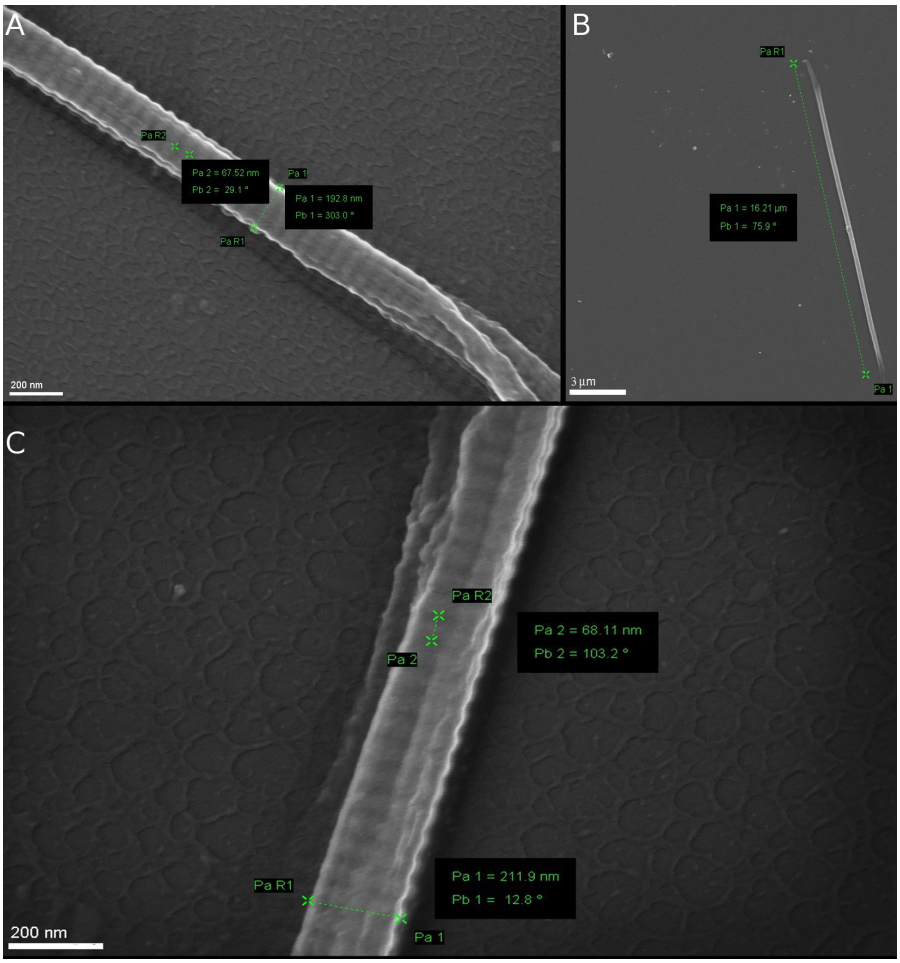


Figure 3.24: Fibrils extracted from rat tail tendon. A: the fibril has a diameter of $\sim$192 nm and a D-banding of $\sim$67 nm. B: The fibril has a length of $\sim$16μm. C: The fibril has a diameter of $\sim$212 nm and a D-banding of $\sim$68 nm.

3.2.3 Fibrils observation under light microscope

Fluorescent samples of fibrils isolated from rat tail tendon were obtained using the protocol described in Sec.2.2 and visualized using the fluorescence microscope (Leica DM5500B). As the negative control image results black, the secondary antibodies bind specifically primaries without labelling the sample. Therefore the fluorescence image of fibrils can be considered true and not the results of unreal interferences between the fluorophore and the substrate. Fig.3.25 shows fluorescent fibrils, pointed out with white arrows, and several residues of biological material: small microfibrils or filaments that are deposited during the isolation process. The

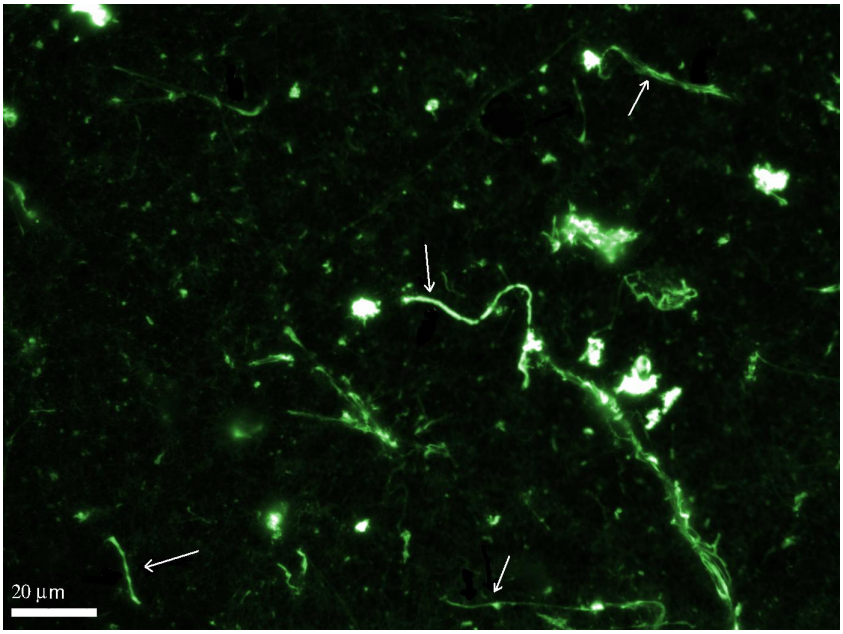


Figure 3.25: Fluorescent fibrils extracted from rat tail tendon.

main issue of fluorescence technique that strongly limited its use in this work is the time of decay of the fluorophore. After 30 seconds of exposition at light the fluorophore deteriorates and fibrils were not longer visible so their extraction from the sample was not easy.

Concerning the dark field, the results show that the fibrils visualization using this method is possible. An example of image obtained with dark field is shown in Fig. 3.26. It shows a dark field in which it is possible to recognize some salt particles, due to the PBS in which fascicles are stored, and some isolated fibrils

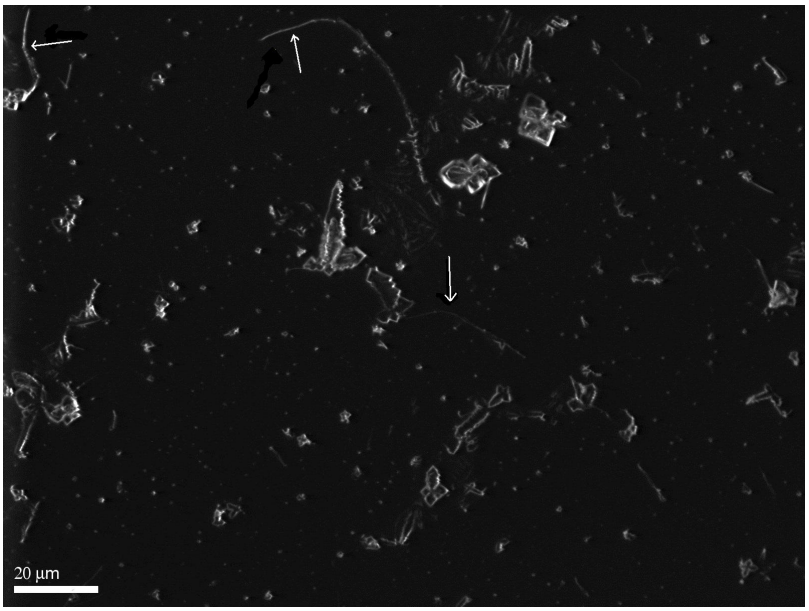


Figure 3.26: Dark field microscope used to visualize the fibrils from rat tail tendon.

(marked by white arrows). In order to measure the length of the fibrils, an optical ruler was used: it was placed under the microscope for estimating the length. The fibrils analyzed have a length in the range of 20-30 μm. These values confirm that the specimens are fibrils.

Finally samples containing fibrils from rat tail tendons were visualized using the inverted microscope (Nikon Eclipse TS100). In order to evaluate if the structures displayed in Fig.3.27 are fibrils or larger structures like fibers or fibrils aggregates, diameters were firstly calculated using a micrometer and a image analysis program (Fiji Is Just ImageJ). Values about 500-800 nm were obtained but the truthfulness of these data is deeply influenced by optical focus, distortion and measurement errors. Therefore further analysis were done: a single sample is visualized under inverted microscope, dark field and SEM. Fig.3.28 shows the same fibrillar structure observed using dark field (A), inverted microscope (B) and SEM (C). Using the SEM two magnifications of the same structure were acquired: c1 and c2 show single fibrils with diameters of 262 nm and 259 nm. Those structures are visible in A and B, hence the structures visualized under the inverted microscope are single fibrils, or at least bundles of fibrils.

To conclude the dark field method is easier and quicker than the immunofluo-

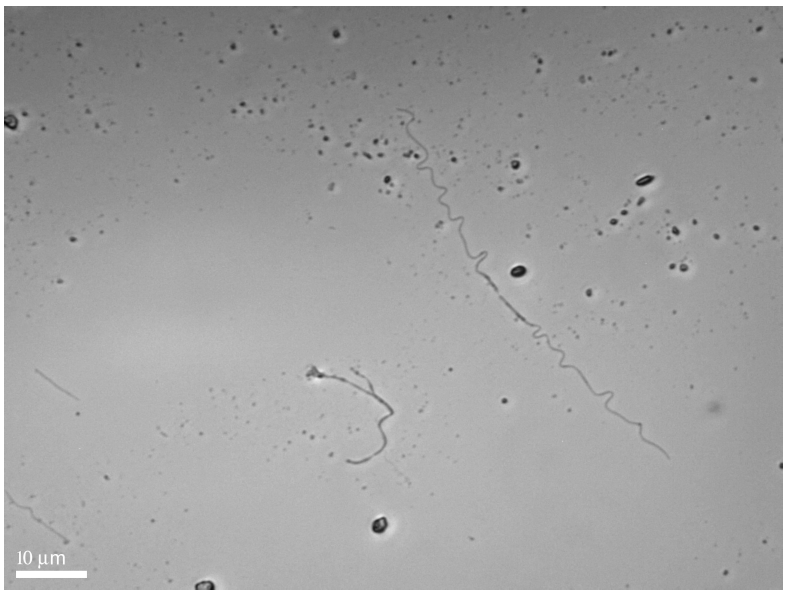


Figure 3.27: Fibrils visualized using the inverted microscope.

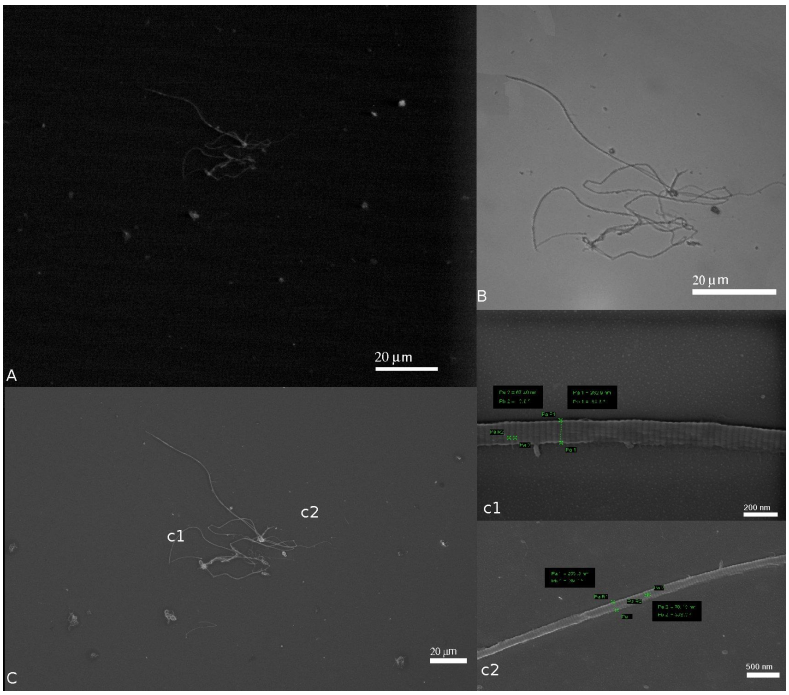


Figure 3.28: A fibril aggregate visualized using (A) dark field, (B) bright field in an inverted microscope, (C) SEM. Higer magnification of (C) where sigle fibrils are shown, (c1) and (c2).

rescence but the working distance between the lens and the sample is low and the insertion and movimentation of test machine tips is not feasible. Then, even if the resolution is not optimal, the technique based on the inverted microscope allows the fibril visualization and it will be applied in order to perform mechanical tests.

3.2.4 Set-up development for the pulling tests

An experimental set-up for pulling tests of single fibrils was created using the FemtoTools, the inverted microscope and a tip probe. In order to verify if the fibril can be detached from the silicon substrate and stretched, a needle sphere was used as a tip probe and the procedure described in Sec.2.2.5 was performed. Results shown in Fig.3.29 validate the technique, the fibril can be lifted from the egde stick to the tip and moved while the other edge remains fix on the substrate. After

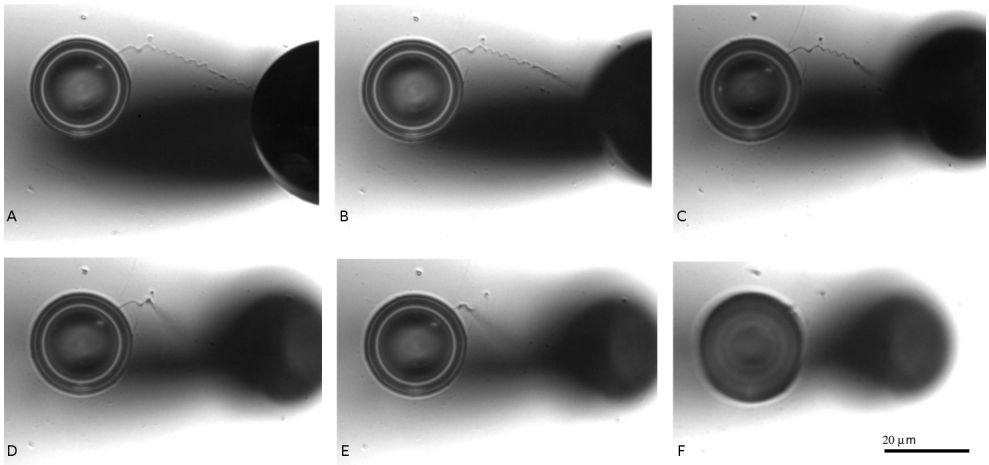


Figure 3.29: Fibril detachment from the substrate. In all the images the focus is mantained on the substrate to better visualize the detachment of the fibril. A: the tip is still in contact with the substrate and the whole fibril lays on it, B-E: the tip is raised and the fibril gradually loses contact with the silicon substrate, F: the fibril is attached to the substrate only by the drop of glue.

the pulling test the fibril is analysed, Fig.3.30 shows the fibril after failure: the left edge is still attached to the drop and the substrate while the other is disconnected from the tip and damaged. Since the technique here described is appropriate to test a single fibril, the experiment were repeated using a real tip sensor able to record values of force and displacement.

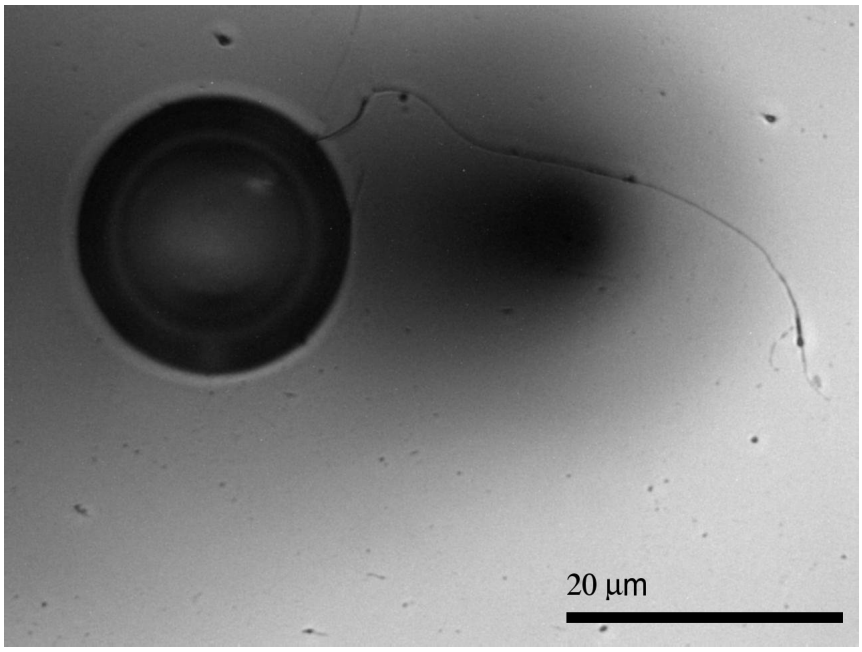


Figure 3.30: Fibril after failure visualized under inverted microscope.

3.2.5 Mechanical tests

Mechanical test on single fibrils have been performed following the procedure described in Sec.2.2.6. Data were elaborated in order to obtain Force - Length and Stress - Strain charts. Force and length values were recorded by the FemtoTools FT-WF502-CT software while stress and strain values were calculated using the following equations:

$$\sigma = \frac{F}{\pi R^2} \tag{3.29}$$

$$\varepsilon = \frac{l - l_0}{l_0}\% \tag{3.30}$$

After the test the broken fibrils were image using the SEM and diameters were evaluated. Since the SEM analysis requires vacuum application on the sample, the diameters calculation is not realistic because the PDMS substrate is altered by the procedure. Consequently the diameter is calculated as average of 15 fibrils diameters from the same rat tail tendon deposited on a glass substrate instead of PDMS one. The average leads to a diameter of 198 nm. Finally elastic modulus for each test was extrapolated from the Stress - Strain chart using the interpolation

tool.

Results for the physiological fibril A are reported in Fig.3.31, Fig.3.32. Values

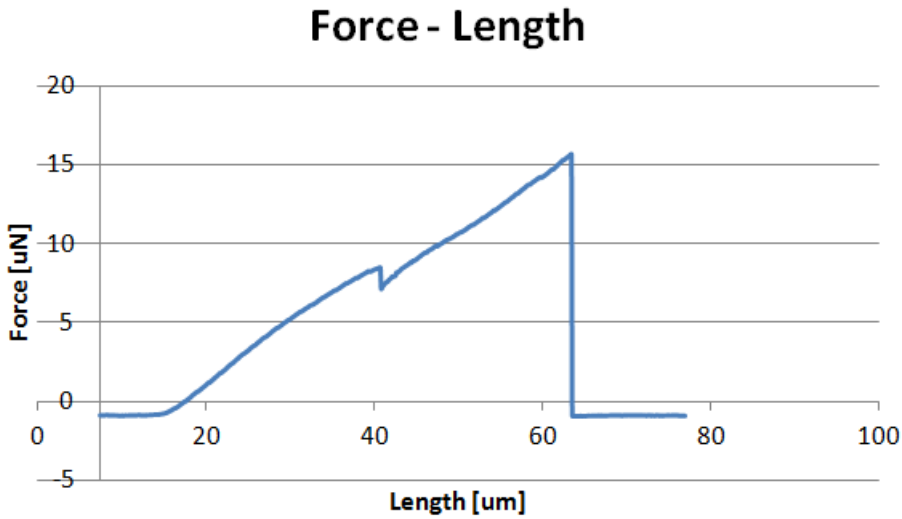


Figure 3.31: Force - Length chart for fibril A.

of force and length at failure are:

$$F_{max} = 15.7\mu N \tag{3.31}$$

$$L_{max} = 63.3\mu m \tag{3.32}$$

Values of stress and strain at failure are:

$$\sigma_{max} = 0.51GPa \tag{3.33}$$

$$\varepsilon_{max} = 311\% \tag{3.34}$$

Elastic modulus is 0.16 GPa.

Results for the physiological fibril B are reported in Fig.3.33, Fig.3.34. Values of force and length at failure are:

$$F_{max} = 3.1\mu N \tag{3.35}$$

$$L_{max} = 42.6\mu m \tag{3.36}$$

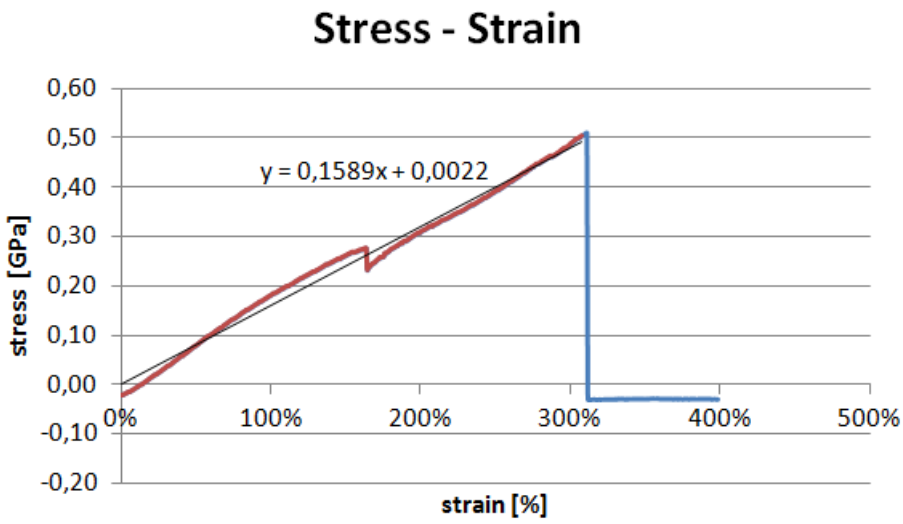


Figure 3.32: Stress - Strain chart for fibril A.

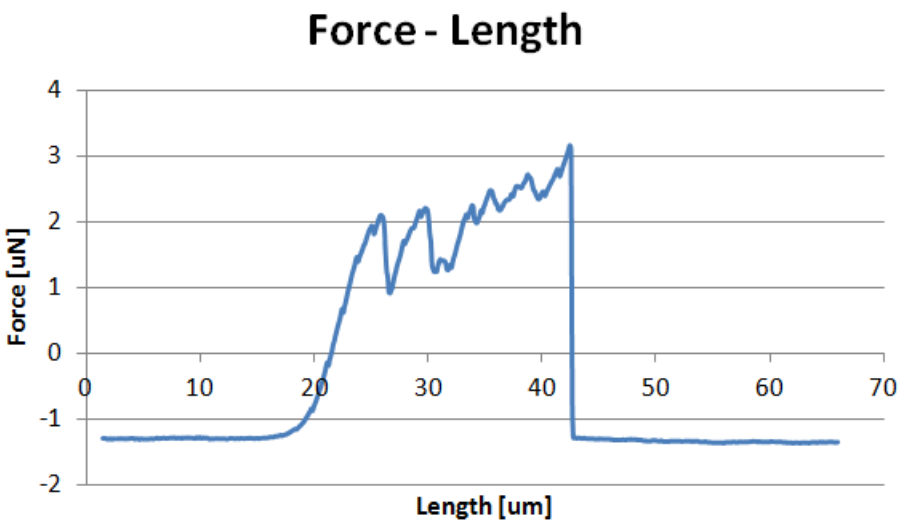


Figure 3.33: Force - Length chart for fibril B.

Values of stress and strain at failure are:

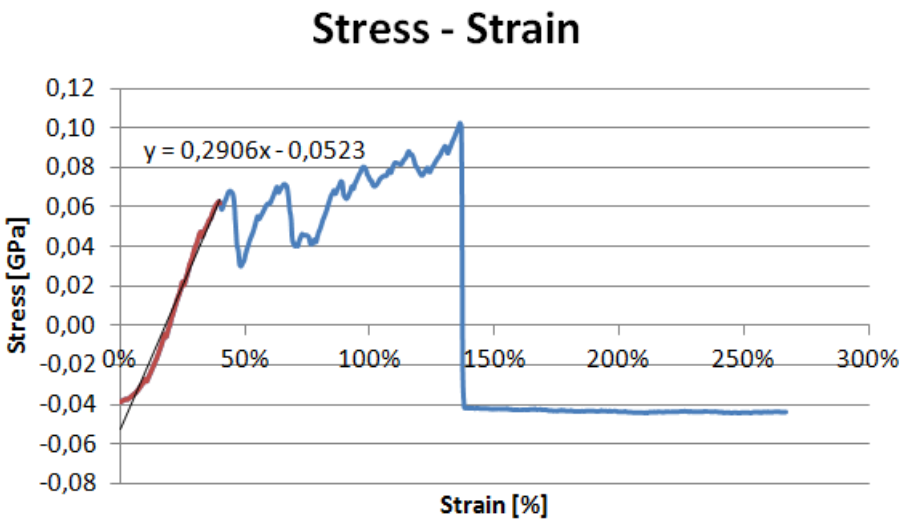


Figure 3.34: Stress - Strain chart for fibril B.

$$\sigma_{max} = 0.1GPa \tag{3.37}$$

$$\varepsilon_{max} = 137\% \tag{3.38}$$

Elastic modulus is 0.29 GPa.

Results for the aged fibril C are reported in Fig.3.35, Fig.3.36. Values of force and length at failure are:

$$F_{max} = 11.7\mu N \tag{3.39}$$

$$L_{max} = 55.3\mu m \tag{3.40}$$

Values of stress and strain at failure are:

$$\sigma_{max} = 0.38GPa \tag{3.41}$$

$$\varepsilon_{max} = 350\% \tag{3.42}$$

Elastic modulus is 0.11 GPa.

Results for the aged fibril D are reported in Fig.3.37, Fig.3.38.

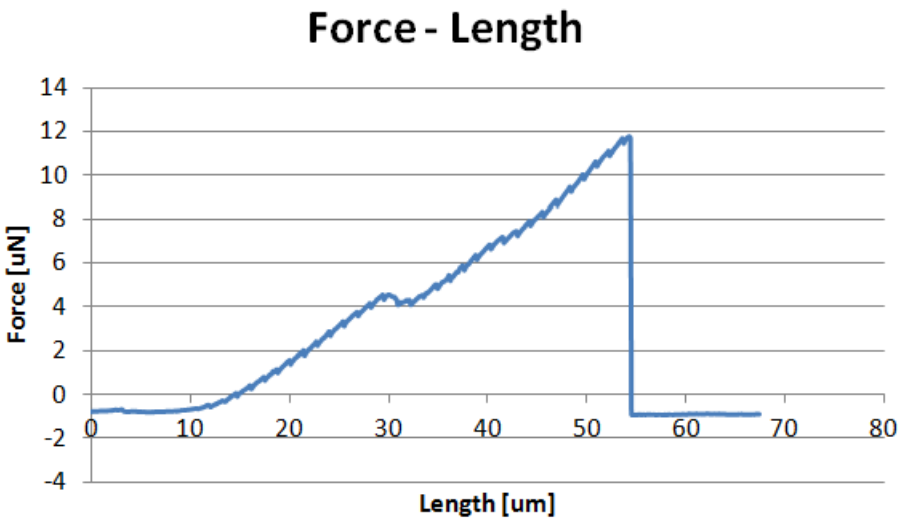


Figure 3.35: Force - Length chart for fibril C.

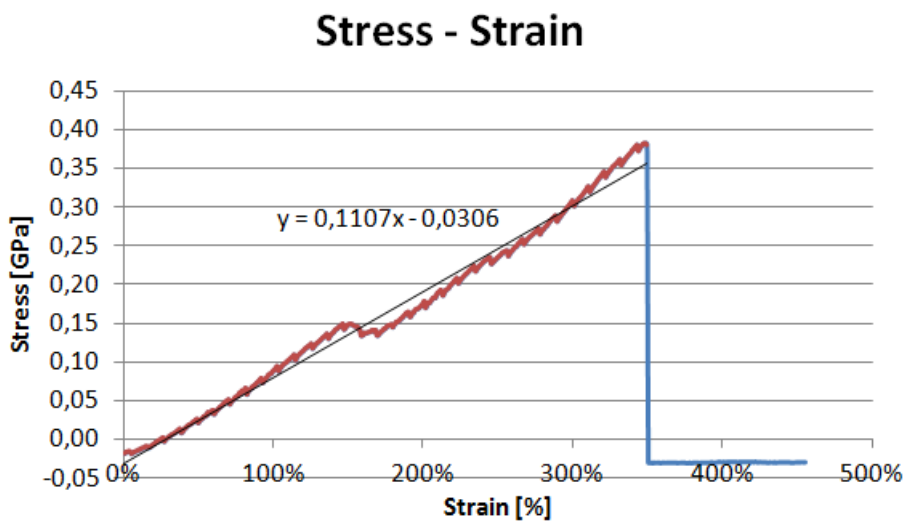


Figure 3.36: Stress - Strain chart for fibril C.

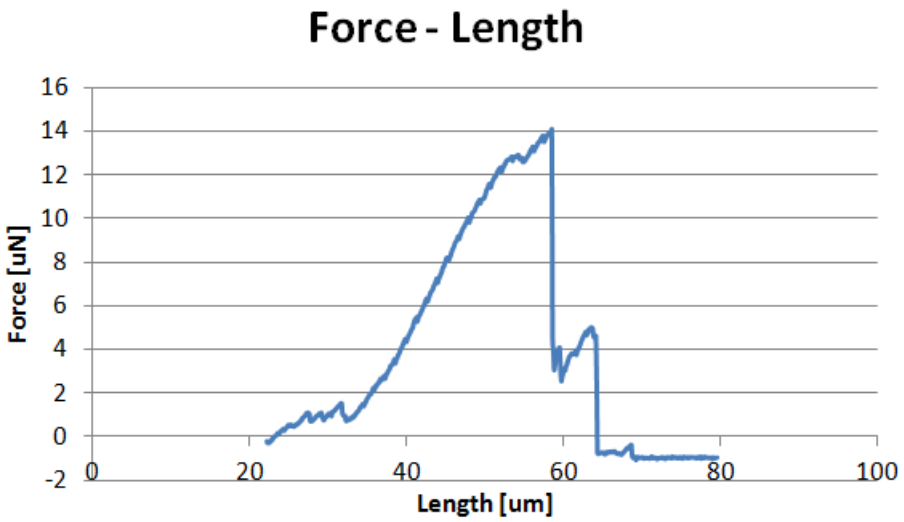


Figure 3.37: Force - Length chart for fibril D.

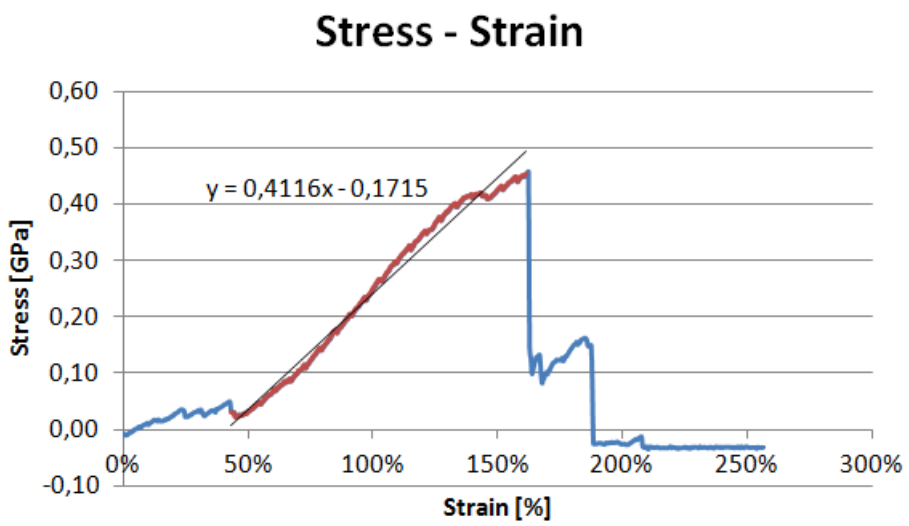


Figure 3.38: Stress - Strain chart for fibril D.

Values of force and length at failure are:

$$F_{max} = 14.1\mu N \tag{3.43}$$

$$L_{max} = 58.5\mu m \tag{3.44}$$

Values of stress and strain at failure are:

$$\sigma_{max} = 0.46GPa \tag{3.45}$$

$$\varepsilon_{max} = 162\% \tag{3.46}$$

Elastic modulus is 0.41 GPa.

Chapter 4

DISCUSSIONS

4.1 Computational Model

The aim of this work is to develop a collagen fibril model using the MD technique. The molecular model describes the mechanical properties of a single fibril and allows to analyse the influence of cross-links (enzymatic and non-enzymatic ones) during a pulling test. The model uses a coarse-grained approach in which a group of amino-acids is described by a single structure, called bead. Four different models have been created:

- pseudo-physiological model: it contains tropocollagen molecules joined by non bond interactions, enzymatic cross-links are not included.
- physiological model: tropocollagen molecules are linked together by enzymatic cross-links. The enzymatic cross-links describe the mechanical properties of hydroxylysyl pyridinoline and are inserted between the N or C termini and the corresponding bead.
- non-enzymatic total model: non enzymatic cross-links (that describe the mechanical properties of glucosepane) are ramdomly added at the structure, they are placed through the whole section.
- non-enzymatic shell model: non enzymatic cross-links are placed only in the external shell of the fibril.

Simulations have been performed using LAMMPS and pulling tests have been carried out at a rate of 100 m/s in vacuum. Stress - Strain curves are shown in Fig.4.1. In the first part (strain <8%, point A) the four curves are overlapped because the

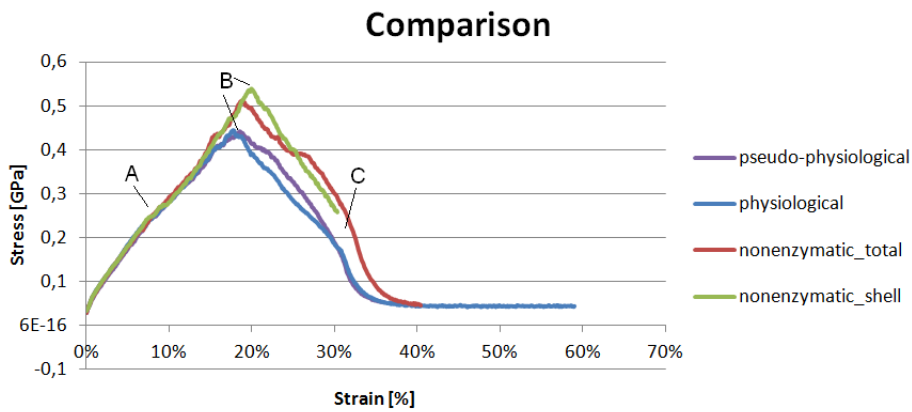


Figure 4.1: Stress - Strain curves comparison.

fibrils behave in the same way since cross-links are not already stretched. The stress increase is due to shear forces and covalent bond interaction inside tropocollagen molecules. In the second part of the curves (A - B) cross-links start to influence the behavior. They are gradually stretched and bear the load. The curves slope is different and their elastic modulus changes in respect to cross-links. The peak point (point B) describes the unravel of the fibril: from the trajectory analysis is evident that a compact group of tropocollagen molecules starts to be pulled out from the fibril. This leads to a decrease in shear forces and a consequent decrease in the stress supported by the fibril. After the peak (B - C), the stress decreases in different ways among the curves because the presence of cross-links, still stretched, is clearly evident. In this part of the curve, three different events occour: cross-links progressively break, tropocollagen molecules separate from each other therefore shear forces decreases, covalent bonds inside tropocollagen molecules are stil stretched because the use of Non Reactive FF does not allow the bond breakage. All these events determine the differences among the curves in slopes and shapes during the B-C section. Point C corresponds to the complete fibril failure, stress rapidly decreases because, as shown in Fig.3.15, a bundle of

tropocollagen molecules is pulled out from the fibril.

The effect of enzymatic cross-links has been evaluated comparing the pseudo-physiological Stress - Strain curve and the physiological one. The enzymatic cross-links begin to influence mechanical behavior in the last part of the A - B section: curves start to separate and the elastic modulus are slightly different. As expected the physiological fibril is stiffer than the pseudo-physiological one since its elastic modulus is 2.06 GPa and the other is 1.97 GPa. The major influence of enzymatic cross-links is evident in the B-C section where the curves are clearly different. Pulled cross-links determine an increase in fibril stiffness indeed the curve slope is higher than the one without cross-links. When both fibrils fail (from point C to the end) their curves are overlapped as the cross-links do not influence the behavior anymore and the two models correspond.

Non enzymatic cross-links have been included in non-enzymatic total model and non-enzymatic shell model. Their influence is clearly evident comparing the non-enzymatic curves to the enzymatic one, in the A - B section their elastic modulus are discrepant: since enzymatic models contain both kind of cross-links their stiffness is higher. Elastic modulus values are: 2.06 GPa for the enzymatic model, 2.39 GPa for non-enzymatic total model and 2.47 GPa for non-enzymatic shell model. Peak values of stress are higher for non-enzymatic models than for the enzymatic one while peak values of strain are similar. Simulation results show that the peaks are due to the unravel of a compact group of tropocollagen molecules. Since this behavior depends on the overlapped region length among tropocollagen molecules, the peak strains are the same in all models but the non-enzymatic models are more resistent and need more load to be unravel, then maximum values of stress are greater. The influence of non-enzymatic cross-links can be evaluated as an increase in the elastic modulus:

$$(1 - \frac{E_{enzymatic}}{E_{non-enzymatic-total}}) \cdot 100 = 13.8\% \tag{4.1}$$

$$(1 - \frac{E_{enzymatic}}{E_{non-enzymatic-shell}}) \cdot 100 = 16.6\% \tag{4.2}$$

$$(1 - \frac{E_{non-enzymatic-total}}{E_{non-enzymatic-shell}}) \cdot 100 = 3.3\% \tag{4.3}$$

Results show that the presence of non-enzymatic cross-links leads to an increase in the fibril stiffness. Between the two non-enzymatic models, a little difference can be seen: the shell model shows a stiffer behavior as the elastic modulus increases of 17% while in the total model the elastic modulus increases only of 14%. In order to define which AGEs distribution is the realistic one, further studies using a more complete model and lower pulling rates are required.

The models presented in this study contain some limitations that compromise the results truthfulness. Due to computational limitation fibril models are only partly realistic since the dimensions are reduced and the velocity is extremely high. Natural fibrils diameter is among 50 - 500 nm and length can vary from 10 to 200 μm [5] while the one presented here is 960 nm long and has a diameter of 20 nm. *In vitro* pulling tests are normally performed at slow velocities in the range of μm /s [17] [37] [38], the model proposed is pulled at 100 m/s. Reduced dimensions and high velocity are required as each simulation takes about six days to complete on eight processors therefore a more realistic model would lead to extremely high computational times. Furthermore the simulations are carried out in implicit solvent as the creation of a water box would have strongly increased the computational times. Works that analyse the influence of environmental conditions are presented in literature [21] [19], they show that dried conditions lead to a decrease of diameters of about 30%. Consequently elastic modulus are higher in dried conditions than in the physiological ones. Finally the enzymatic cross-link is modelled as divalent bond even if hydroxylysyl pyridinoline is a trivalent cross-link. Among all the limitations, the last one less influences the results. The limitations determine non realistic Stress - Strain curves in fact their shapes and stress peaks shown in Fig.4.1 can not be compared to the experimental works presented in literature. Consequently values of stress and strain at failure are not clearly determinable. Elastic modulus values are in the range of literature ones (Tab.4.1) and they correctly describe the cross-links influence.

Overview of fibril elastic modulus			
Technique	Specimens	E [GPa]	Reference No.
Molecular model	None	2.06 - 2.47	Present study
AMF tensile	Human patellar tendon	2.8 $\pm$ 0.34	[39]
AFM tensile	Bovine achilles	0.2 - 0.8	[19]
MEMS tensile	sea cucumber dermis	0.86 $\pm$ 0.45	[21]
Molecular model	None	4.36	[23]
Molecular model	None	0.3 - 1.2	[40]

Table 4.1: Values of fibril elastic modulus.

Despite computational limitations, the models proposed are solid start points to the development of more complex and realistic ones, since in literature single collagen fibrils models are not fully developed yet.

The only two models presented in literature contain substantial semplifications in the whole fibril description. The one proposed by Buehler [23] is a 2D periodic coarse-grained fibril model based on an array of 2 x 5 tropocollagen molecules. Cross-links are not differentiated between enzymatic and non enzymatic ones and they are modelled as incrementation of non bond interaction at the ends of each tropocollagen molecules. Gautieri et al. [24] have developed a viscoelastic fibril model based on Kelvin-Voigt elements that describe tropocollagen molecules. The model proposed in this study is therefore the first one that try to describe a 3D single fibril that can contain different cross-links distributions: cross-links density and their mechanical properties are realistic and their position can be varied. To conclude the fibril model is very interesting for future developments since the only limitation is related to computational time and can be avoided with the use of a higher calculating capacity.

4.2 Experimental Validation

Mechanical tests on single fibrils have been performed using different tecniques (MEMS [21] and AFM [7]) that contain some limitations and do not allow fibril

breakage. In order to overcome this, in the present work a new set-up was developed and validated with preliminary tests. The experimental set-up is composed by relatively standard components (micromanipulator and inverted microscope) therefore it can be easily replicated. Preliminary tests have been performed on four fibrils, two physiological and two aged as described previously. In order to test them a new protocol has been developed: fibrils are deposited onto a PDMS substrate, visualized under inverted microscope and pulled using a micromanipulator. Obtained curves of Force - Length and Stress - Strain are compared in Fig.4.2 and Fig.4.3.

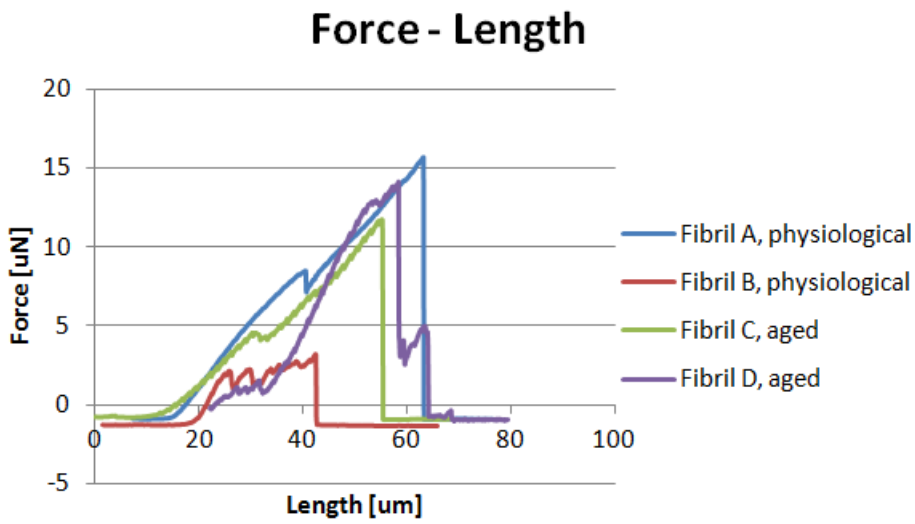


Figure 4.2: Force - Length curves comparison.

Since SEM analysis shows that the failure of fibril B is due to its pulling out from the drop of glue while others break along their structures (Fig.4.4), fibril B results have no significancy and are not analized.

Values of force, length, stress, strain at failure and elastic modulus are reported in Tab.4.2.

This work shows for the first time Stress - Strain curves of fibrils until failure. Works present in literature do not reach the natural breakage point so the comparison is not possible. Shen in [21] demonstrates that fibrils withstand stresses up to 0.6 GPa and strain of 100% without fracturing. The obtained values of

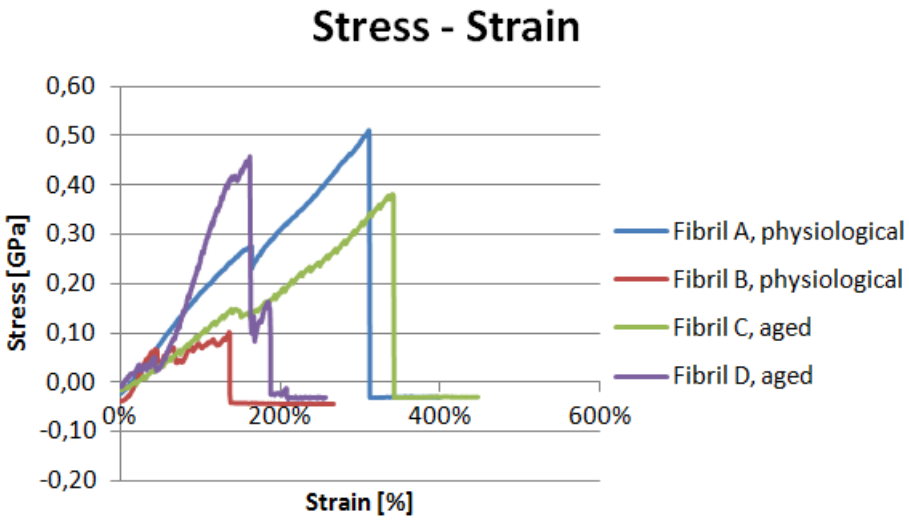


Figure 4.3: Stress - Strain curves comparison.

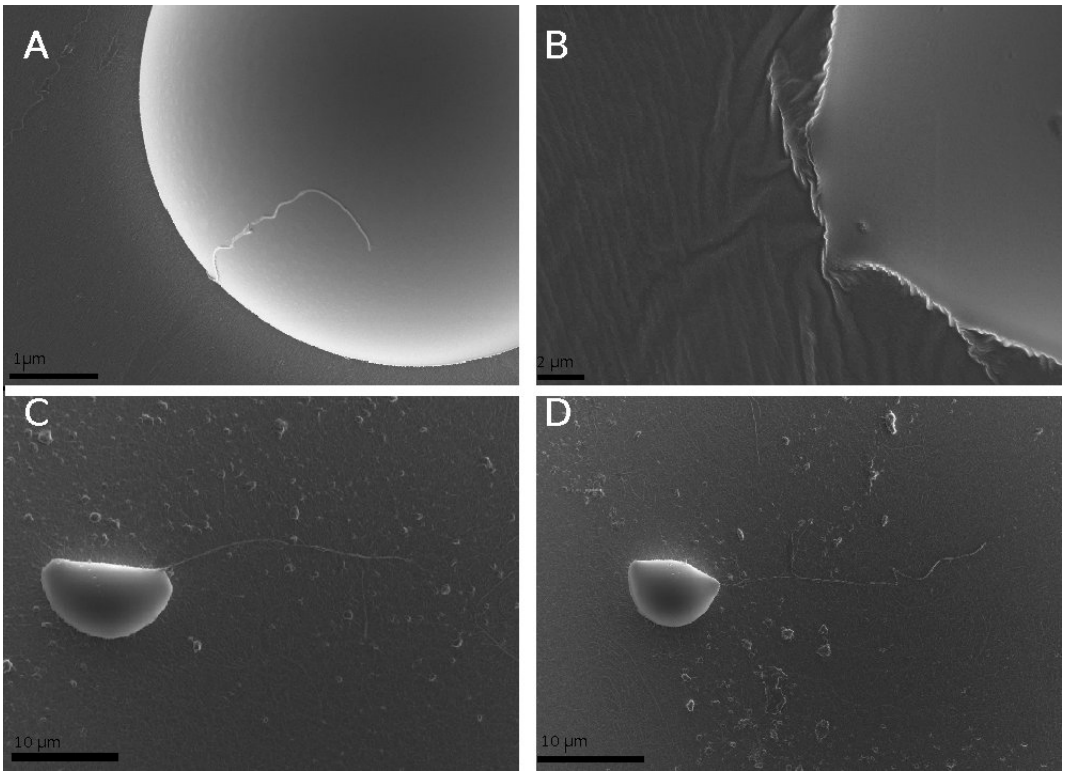


Figure 4.4: SEM images of fibrils after pulling tests. A: fibril A lays onto the drop of glue. B: fibril B has been pulled out of the glue, note the cast impressed on the PDMS substrate. C and D: aged fibrils broken and laying on the substrate.

Fibril	Force [μN]	Length [μm]	Stress [GPa]	Strain [%]	E [GPa]
A	15.7	63.3	0.51	331	0.16
C	11.7	55.3	0.38	350	0.11
D	14.1	58.5	0.46	162	0.41

Table 4.2: Mechanical tests.

elastic modulus agree with literature range [19]. Therefore the values presented here can be considered realistic, even if their precision is influenced by the use of an average diameter instead of a real one, and the testing protocol is suitable for the purpose of the work. Due to the small number of samples tested, a comparison among physiological and aged fibrils is not possible. Moreover, a cross-links quantification would be necessary in order to validate the protocol for single fibrils and to know the exact density of non-enzymatic cross-links. Finally tests can be inproved using hydrated conditions that better mimic the physiological environment.

In order to obtain a statistically significant validation, a larger number of samples must be tested and the aging protocol must be improved; moreover a method to evaluate the exact diameter of pulled fibrils should be developed. Nevertheless these limitations, the preliminary tests presented here are a good start point for further investigations, especially concerning the breakage analysis of single fibrils.

Chapter 5

CONCLUSIONS

Collagen is the most abundant protein in the human body. Its hierarchical structure determines mechanical properties of connective tissues as tendon, bone and cartilage. Mechanical behavior at macroscopic level has been deeply studied, while properties of single collagen fibrils have not been completely evaluated yet. Moreover aging and diabetes, due to the formation of AGEs, could make a change in the biomechanics of fibrils. Consequently a good knowledge of the fibril mechanics is desiderable. Due to the small dimensions of collagen fibrils, mechanical tests are challanging and they present several limitations. Meanwhile, molecular models can be used in order to support experimental analysis. The present work aimed to develop a molecular model of single fibrils in order to analyse their mechanical properties and the influence of AGEs. At the same time a new experimental set-up has been developed and preliminary tests have been performed on physiological and aged fibrils.

The two models have been separately developed and they both improve the current state of art: concerning the computational approach, few and simplified studies are present in literature but they do not fully describe the fibril behavior, while the model presented here is more complete; concerning the experimental approach, the experimental method proposed is the first one that allows the failure test of single fibrils. Despite the models are good start points for further investigations, both have some limitations and can not be compared. Further studies must focus on the improvement of the models proposed here. Regarding the computa-

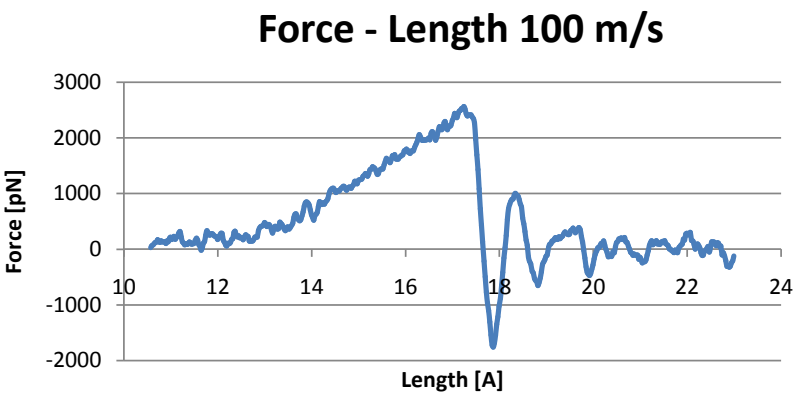
tional model, fibrils dimensions and pulling rate must be optimised through the use of high calculating capacity. The experimental set-up has been validated but more samples have to be tested in order to obtain a statistical analysis and the aging protocol must be validated. Through these improvements both approaches could be combined in order to understand the mechanical behavior at the fibril level, the effects of AGEs and their real distribution through the fibril cross-section.

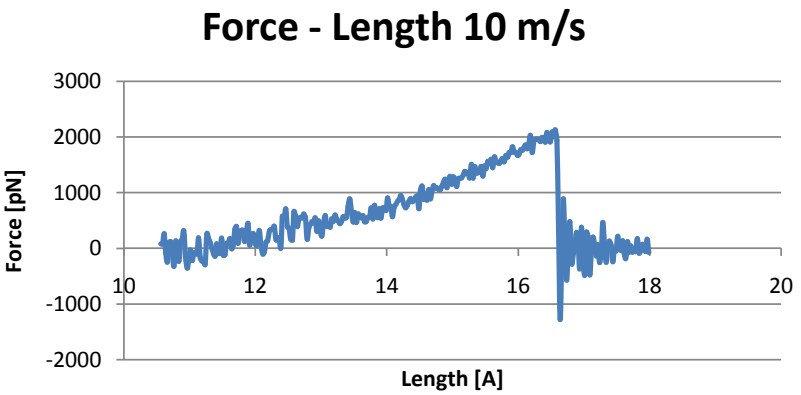
Appendix A

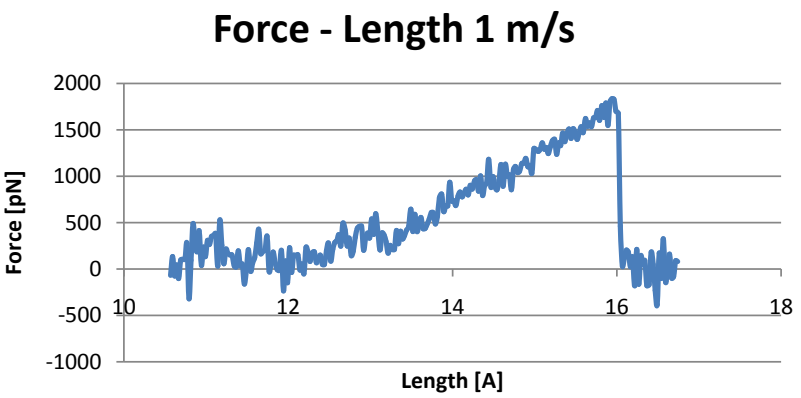
Characterization of cross-links

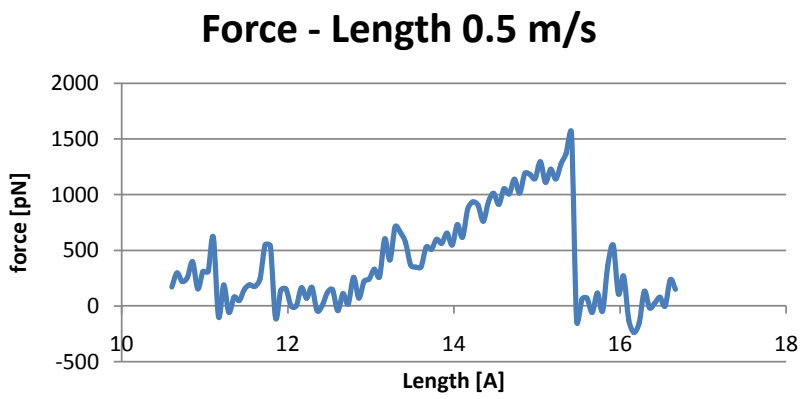
In the following pages Force - Length charts for hydroxylysyl pyridinoline and glucosepane are reported.

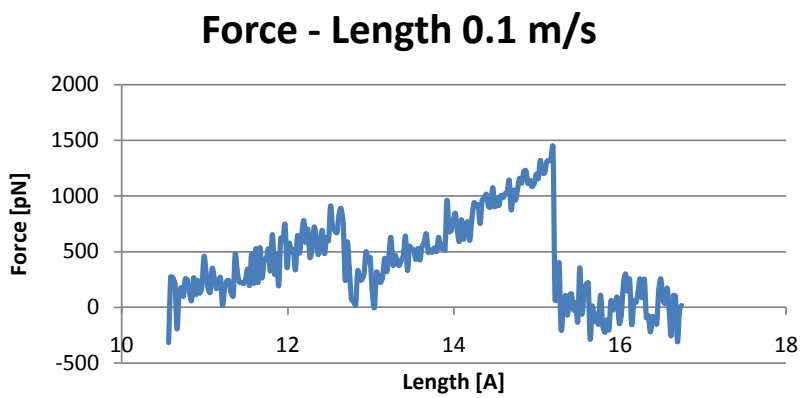
Force – Length charts for hydroxylysyl pyridinoline, direction 2-12

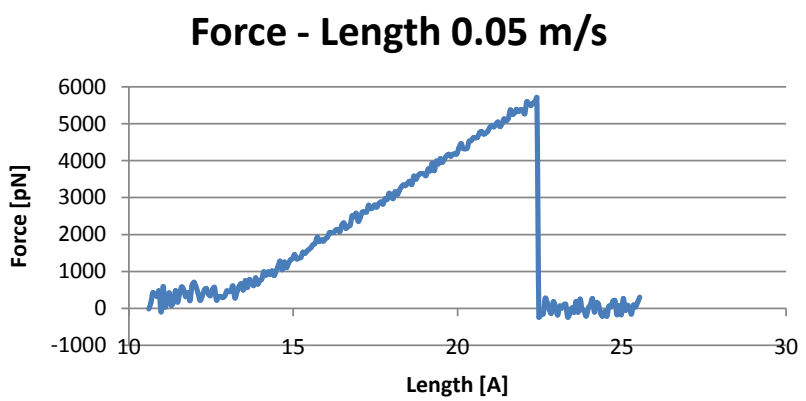




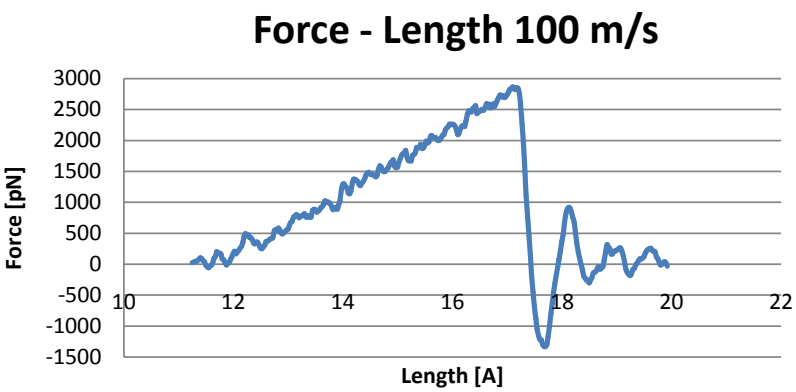


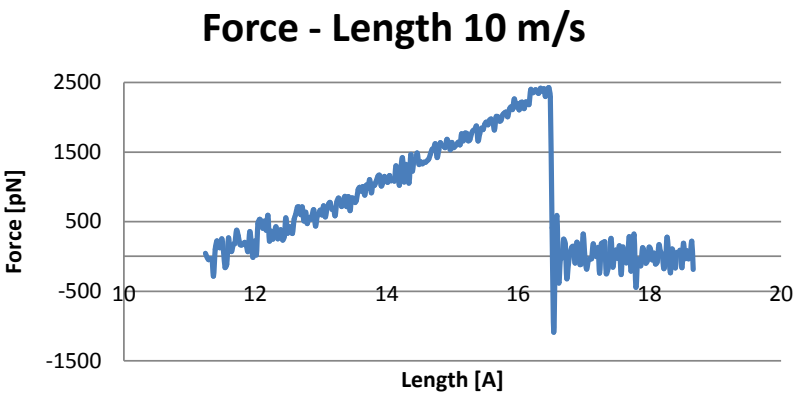

Force - Length 0.5 m/s
force [pN]
2000
1500
1000
500
0
-500
10
12
14
16
18
Length [A]

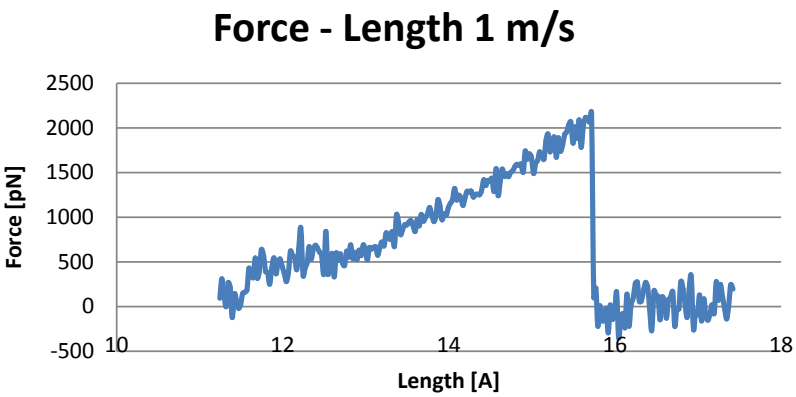

Force - Length 0.1 m/s
Force [pN]
2000
1500
1000
500
0
-500
10
12
14
16
18
Length [A]

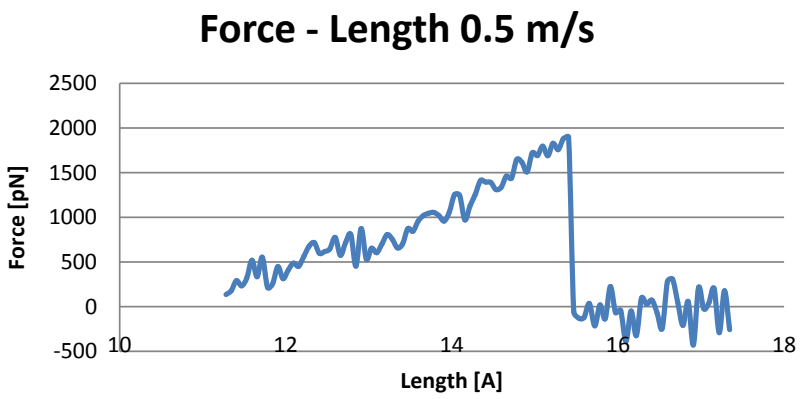

Force - Length 0.05 m/s
Force [pN]
6000
5000
4000
3000
2000
1000
0
-1000
10
15
20
25
30
Length [A]

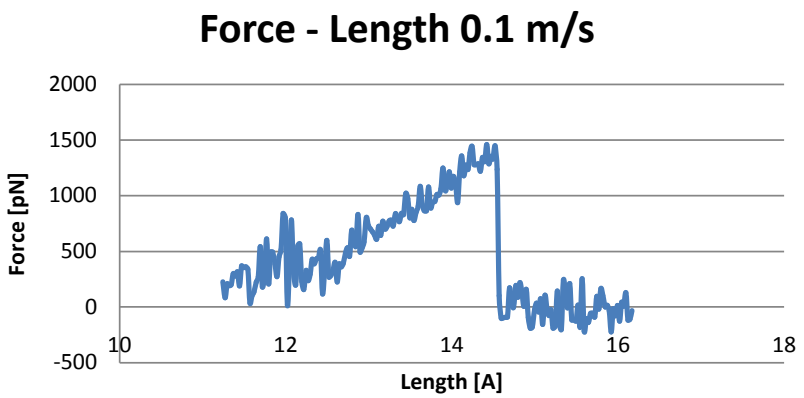
Force – Length charts for hydroxylysyl pyridinoline, direction 7-12

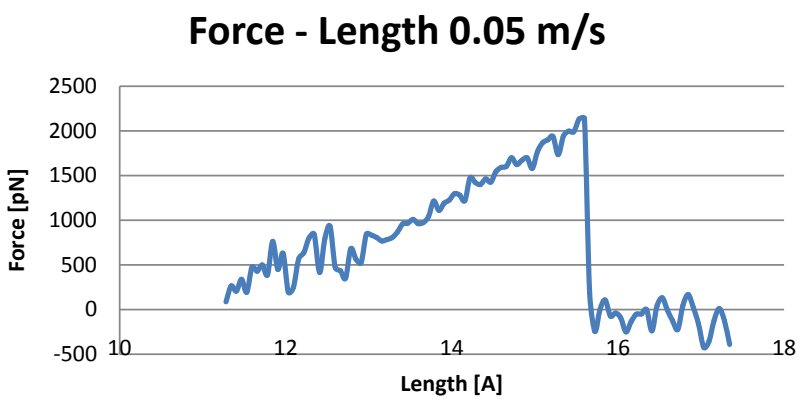




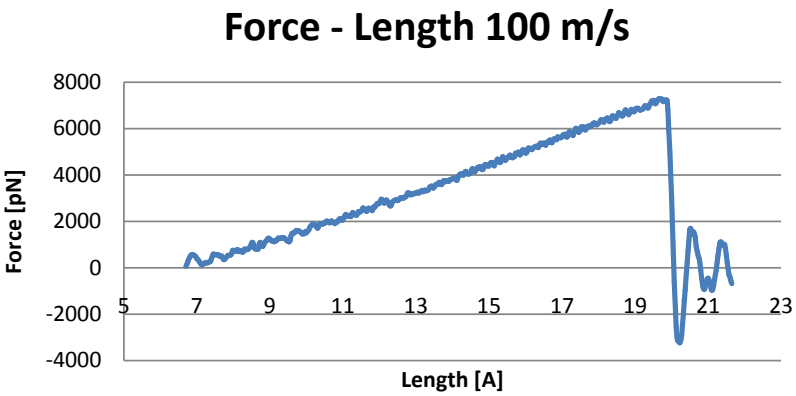


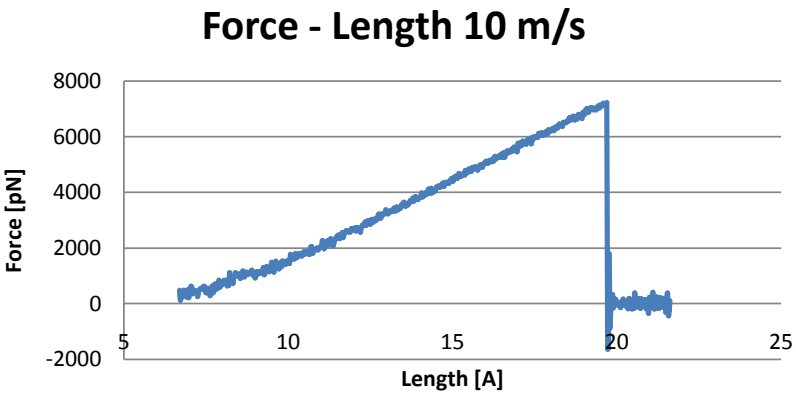

Force - Length 0.5 m/s
2500
2000
1500
1000
500
0
-500
Force [pN]
10
12
14
16
18
Length [A]

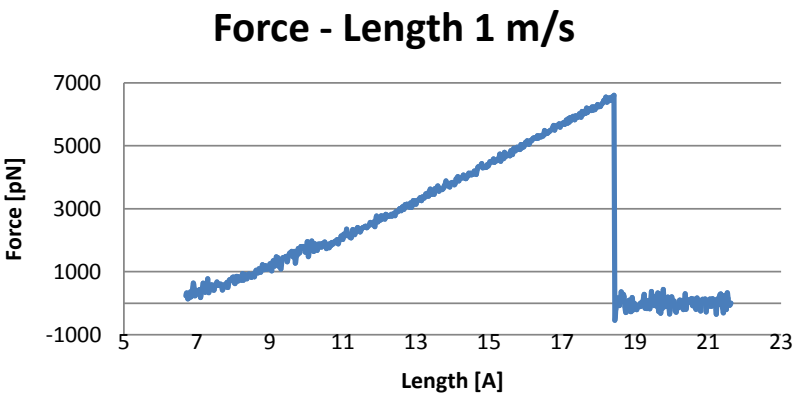

Force - Length 0.1 m/s
2000
1500
1000
500
0
-500
Force [pN]
10
12
14
16
18
Length [A]

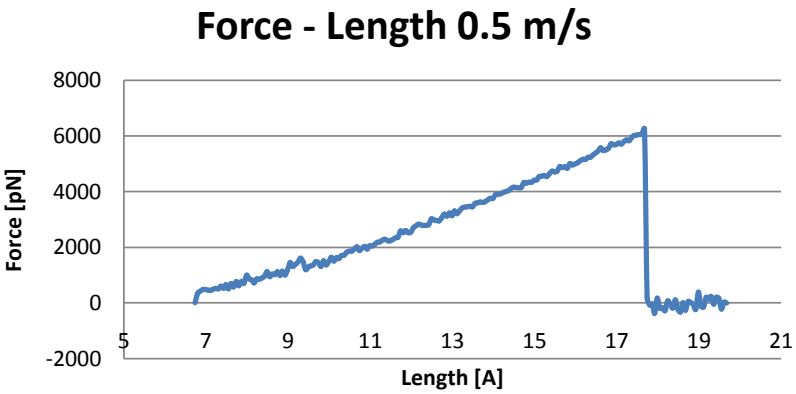

Force - Length 0.05 m/s
2500
2000
1500
1000
500
0
-500
Force [pN]
10
12
14
16
18
Length [A]

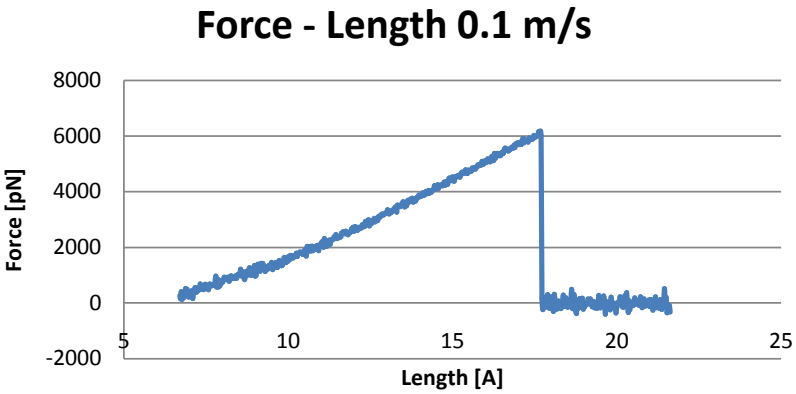
Force – Length charts for hydroxylysyl pyridinoline, direction 2-7

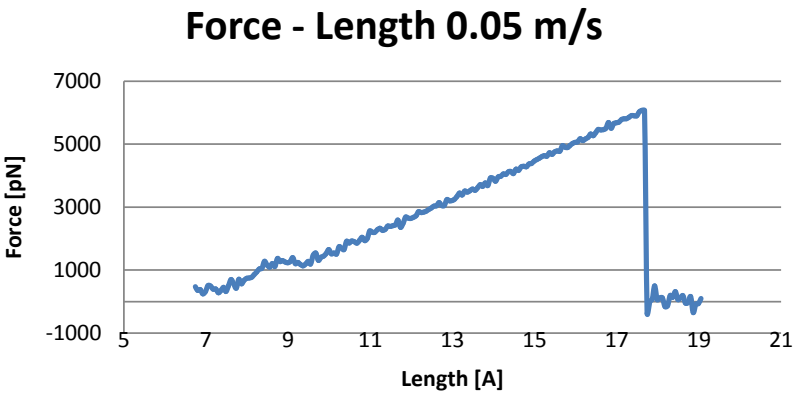




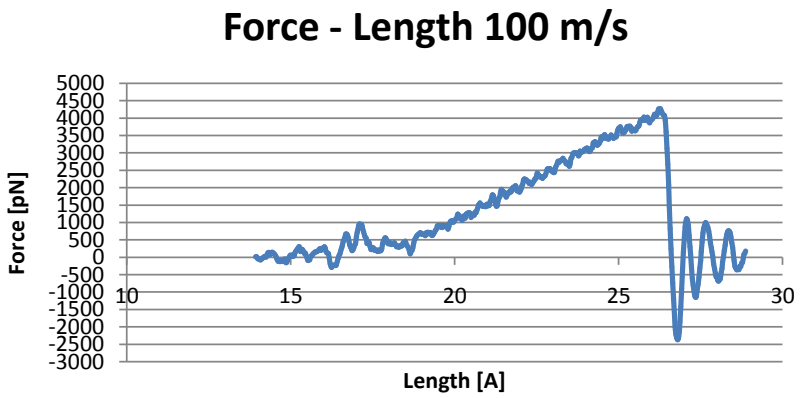


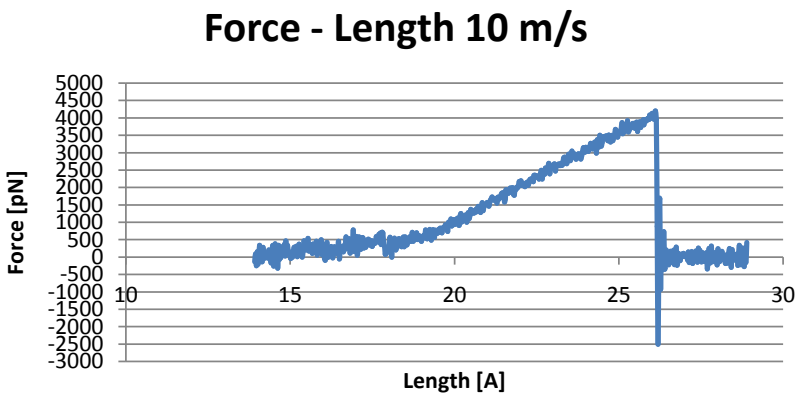

Force - Length 0.5 m/s
8000
6000
4000
2000
0
-2000
Force [pN]
5
7
9
11
13
15
17
19
21
Length [A]

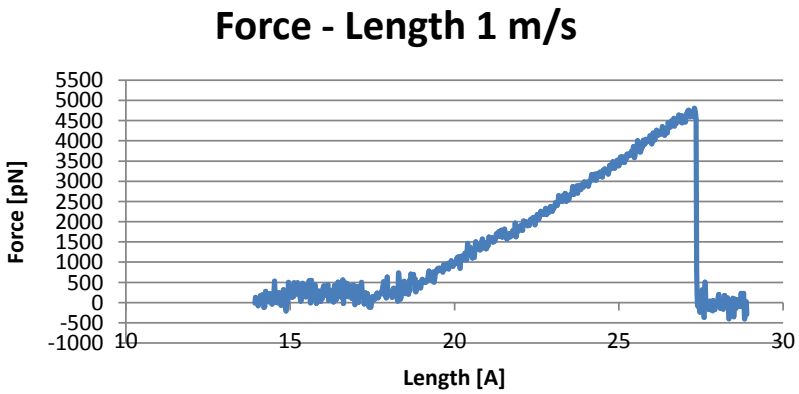

Force - Length 0.1 m/s
8000
6000
4000
2000
0
-2000
Force [pN]
5
10
15
20
25
Length [A]

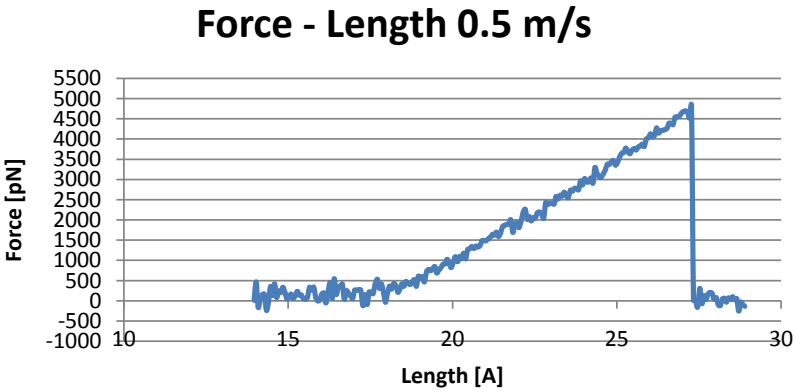

Force - Length 0.05 m/s
7000
5000
3000
1000
-1000
Force [pN]
5
7
9
11
13
15
17
19
21
Length [A]

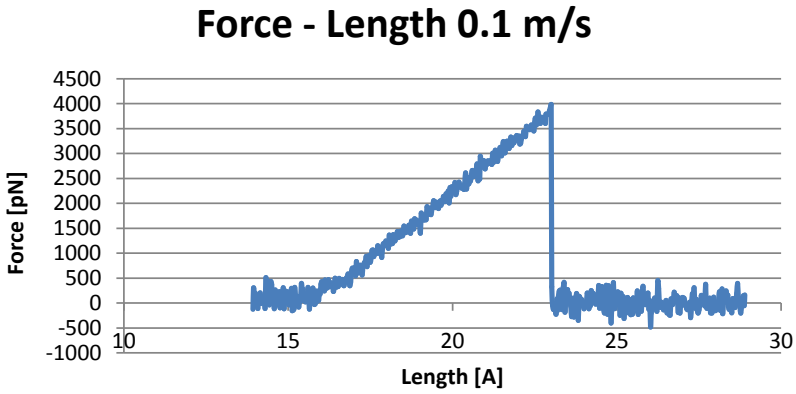
Force – Length charts for Glucosepane, structure 1

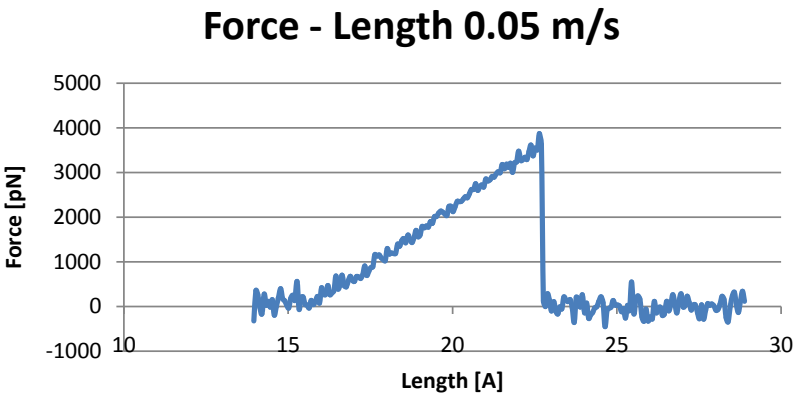




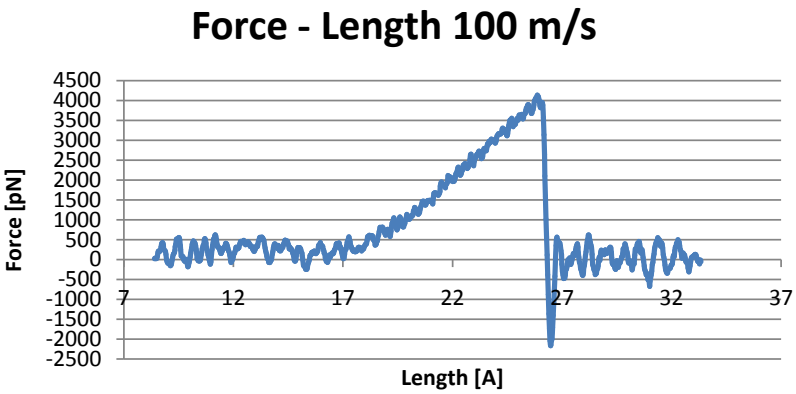


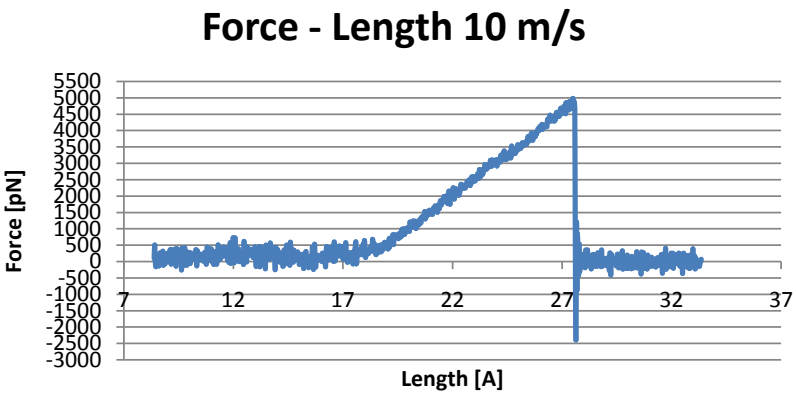

Force - Length 0.5 m/s
Force [pN]
5500
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
-500
-1000
10
15
20
25
30
Length [A]

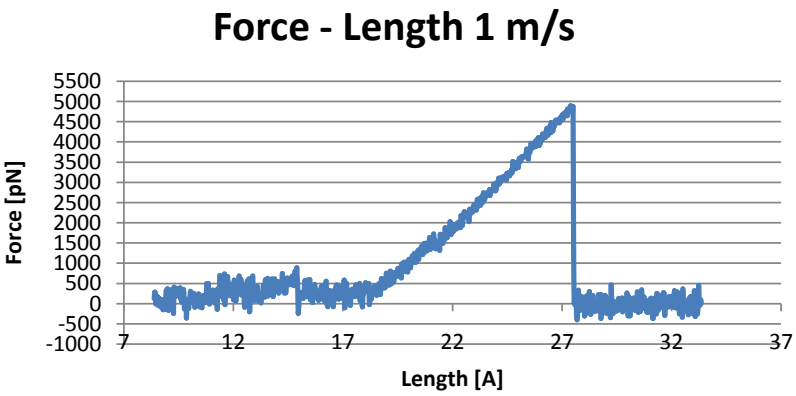

Force - Length 0.1 m/s
Force [pN]
4500
4000
3500
3000
2500
2000
1500
1000
500
0
-500
-1000
10
15
20
25
30
Length [A]

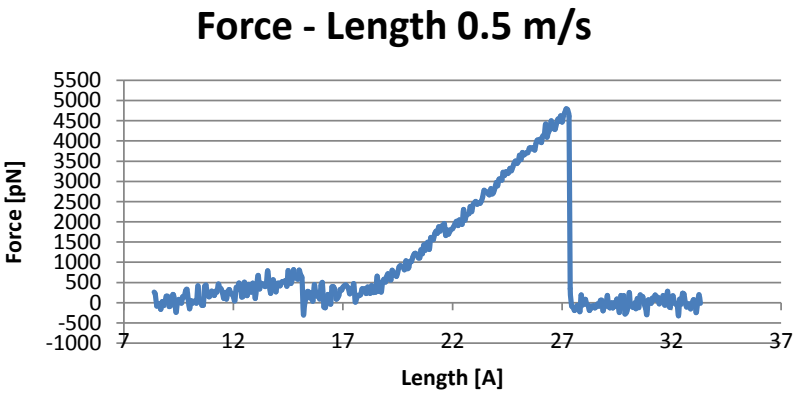

Force - Length 0.05 m/s
Force [pN]
5000
4000
3000
2000
1000
0
-1000
10
15
20
25
30
Length [A]

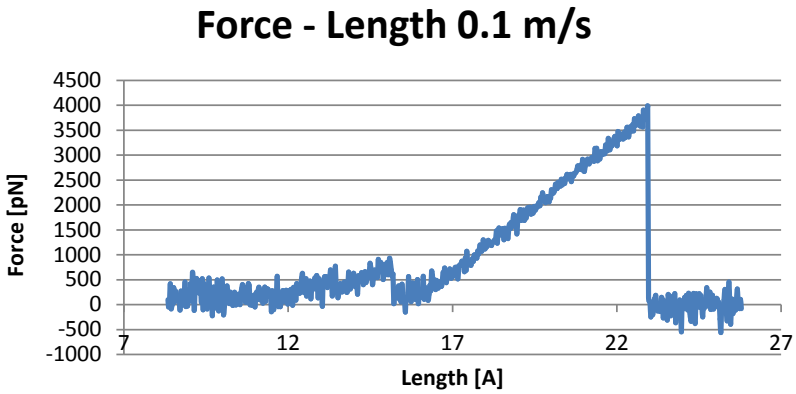
Force – Length charts for Glucosepane, structure 2

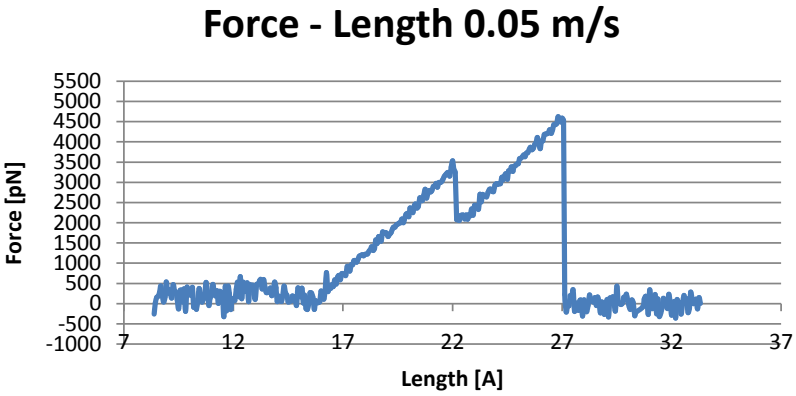




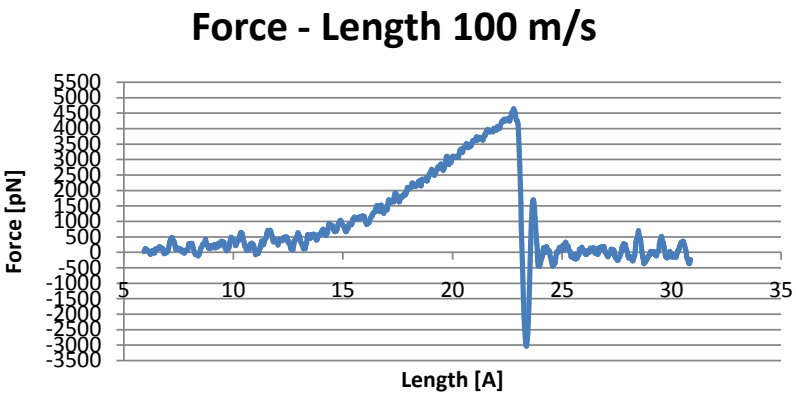


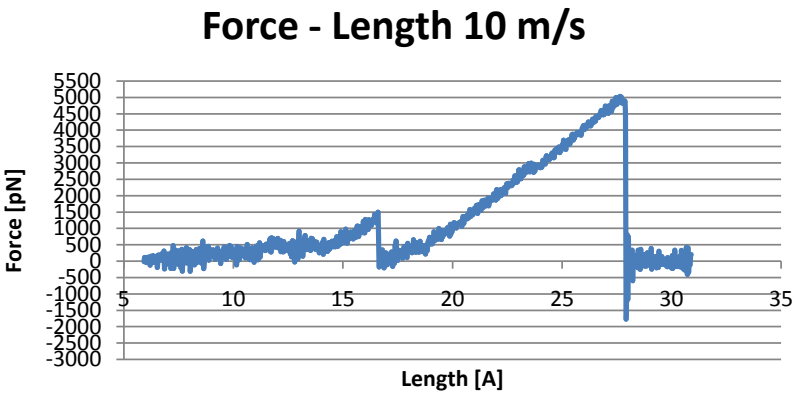

Force - Length 0.5 m/s
Force [pN]
5500
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
-500
-1000
7
12
17
22
27
32
37
Length [A]

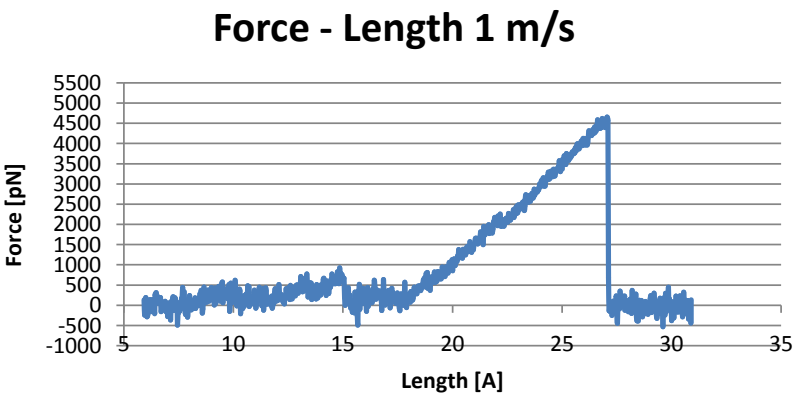

Force - Length 0.1 m/s
Force [pN]
4500
4000
3500
3000
2500
2000
1500
1000
500
0
-500
-1000
7
12
17
22
27
Length [A]

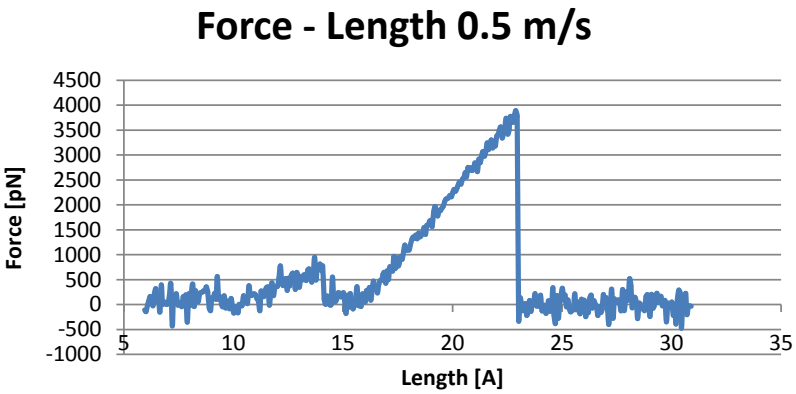

Force - Length 0.05 m/s
Force [pN]
5500
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
-500
-1000
7
12
17
22
27
32
37
Length [A]

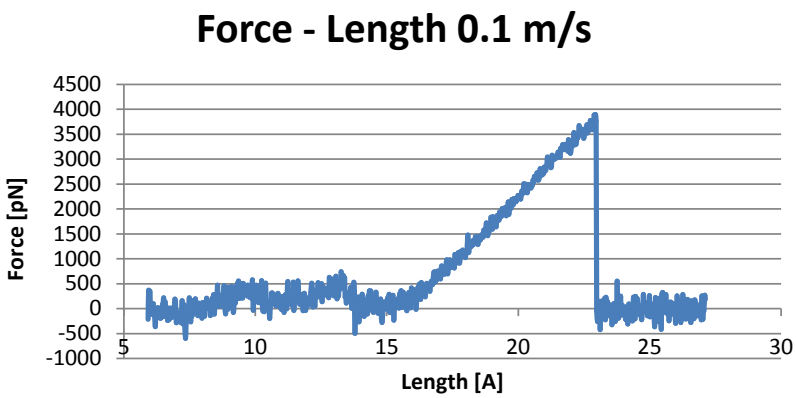
Force – Length charts for Glucosepane, structure 3

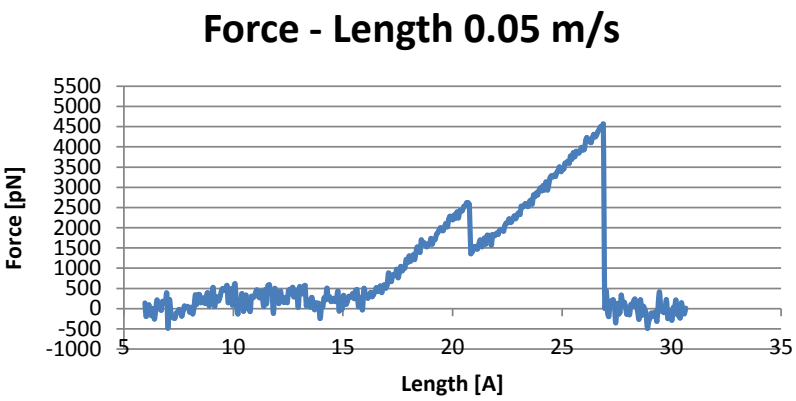







Force - Length 0.5 m/s
Force [pN]
4500
4000
3500
3000
2500
2000
1500
1000
500
0
-500
-1000
5
10
15
20
25
30
35
Length [A]


Force - Length 0.1 m/s
Force [pN]
4500
4000
3500
3000
2500
2000
1500
1000
500
0
-500
-1000
5
10
15
20
25
30
Length [A]


Force - Length 0.05 m/s
Force [pN]
5500
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
-500
-1000
5
10
15
20
25
30
35
Length [A]

Appendix B

C++ scripts

In the following pages C++ scripts used in the work are reported.

Scripts recive and input file that contains ID, spatial coordinates of beads and bonds of tropocollagen molecules. Then they manipulate the input file in order to create new different bonds among beads. Output file containing new bonds can be paste in the bond section of data file. Threee different scripts have been developed:

- ENZYMATIC BONDS generates enzymatic cross-links among the tropocollagen molecules. The bonds are implemented in the models of physiological and aged fibril.
- NON-ENZYMATIC BONDS THROUGH THE WHOLE SECTION allows to obtain non-enzymatic bonds between tropocollagen molecules along the whole section of the fibril.
- NON-ENZYMATIC BONDS IN THE EXTERNAL SHELL is used to generate non-enzymatic bonds among the tropocollagen molecules. Those bonds are used to implement the model of partially crosslinked fibril.

```
//ENZYMATIC BONDS

#include <stdlib.h> //libraries
#include <stdio.h>
#include <malloc.h>

int
main (){
  float dx,dy,dz,distmin,dist;
  float xi,yi,zi;
  float eps=10.; //cut radius
  float *x, *y, *z;
  int NT,NTL;
  int *idb, *ida1, *ida2;
  int *ida;
  int i, j, k, l;
  int kk,ll;
  int iflag;
  int natom, nbond, nbb;
  int *idn, *idnl, *idc, *idcl;
  int *NTL_VET;
  nbb = 333;

  scanf ("%d", &natom);
  scanf ("%d", &nbond);
  natom++;
  nbond++;
  x = (float *) malloc (sizeof (float) * natom);
  y = (float *) malloc (sizeof (float) * natom);
  z = (float *) malloc (sizeof (float) * natom);
  ida = (int *) malloc (sizeof (int) * natom);

  idb = (int *) malloc (sizeof (int) * nbond);
  ida1 = (int *) malloc (sizeof (int) * nbond);
  ida2 = (int *) malloc (sizeof (int) * nbond);

  idn = (int *) malloc (sizeof (int) * nbb);
  idnl = (int *) malloc (sizeof (int) * nbb);
  idc = (int *) malloc (sizeof (int) * nbb);
  idcl = (int *) malloc (sizeof (int) * nbb);
  NTL_VET = (int *) malloc (sizeof (int) * nbb);

//lecture of input file with atoms ID and bonds
  for (i = 1; i < natom; i++)
    {scanf ("%d%*d%*d%f%f%f%*s%*s%*s", &ida[i], &x[i], &y[i], &z[i]);}
  for (i = 1; i < nbond; i++)
    {scanf ("%d%*d%d%d", &idb[i], &ida1[i], &ida2[i]);}
//definition of atoms that will create bonds
for (i =0 ; i < nbb; i++) {
      idn[i] = i * 169 + 1; //Ntermini beads
      idnl[i] = i * 169 + 155; //beads n°155
      idc[i] = i * 169 + 169; //Ctermini beads
      idcl[i] = i * 169 + 15; //beads n°15
}
//ENZYMATIC BONDS
// bonds Ntermini-bead n°155
l=0;
ll=0;
```

```
  for (k = 1; k < nbb; k++){
      distmin=100.;
      i = idn[k];
      xi = x[i];
      yi = y[i];
      zi = z[i];
      for (j = 1; j < nbb; j++){
        dx = abs (xi - x[idnl[j]]);
        dy = abs (yi - y[idnl[j]]);
        dz = abs (zi - z[idnl[j]]);
        if (dx < eps && dy < eps && dz < eps){
            dist = sqrt (dx * dx + dy * dy + dz * dz);
            if (dist < distmin){
              distmin = dist;
              NT = i;
              NTL = idnl[j];}}}

        if(distmin <100){
         NTL_VET[ll]=NTL;
         ll++;
         iflag=0;
         for(kk=0;kk<ll;kk++)
           if(NTL_VET[kk]==NTL) iflag++;
           if(iflag==1 ) {printf("%d 2 %d %d\n",nbond+l,NT,NTL); l++;}}}
// bonds Ctermini-bead n°15
ll=0;
  for (k = 1; k < nbb; k++){
      distmin=100.;
      i = idc[k];
      xi = x[i];
      yi = y[i];
      zi = z[i];
      for (j = 1; j < nbb; j++){
        dx = abs (xi - x[idcl[j]]);
        dy = abs (yi - y[idcl[j]]);
        dz = abs (zi - z[idcl[j]]);
         if (dx < eps && dy < eps && dz < eps){
            dist = sqrt (dx * dx + dy * dy + dz * dz);
            if (dist < distmin){
              distmin = dist;
              NT = i;
              NTL = idcl[j];}}}

      if(distmin <100) {
       NTL_VET[ll]=NTL;
         ll++;
         iflag=0;
         for(kk=0;kk<ll;kk++)
           if(NTL_VET[kk]==NTL) iflag++;
           if(iflag==1 ) {printf("%d 2 %d %d\n",nbond+l,NT,NTL); l++;}}}}
```

```
//NON ENZYMATIC BONDS THROUGH THE WHOLE SECTION

#include <stdlib.h>           //libraries
#include <stdio.h>
#include <malloc.h>
#include <time.h>

//function declatation
void RANDOM(int vett[100], int num);

int main(){
float dx,dy,dz,distmin,dist;
  float xi,yi,zi;
  float eps=2.; //cut radius
  float *x, *y, *z;
  int *idb, *ida1, *ida2;
  int *ida;
  int i, j, k, l, ll, kk, iflag;
  int natom, nbond, nbb, nenz, idl, b2;
  int *VET;
  int num, x1, y1, z1;
  int vet_mol[100], vet_bead[100], vet_id[100], vet_b1[100], vet_b2[66];
  nbb = 333;

  scanf ("%d", &natom);
  scanf ("%d", &nbond);
  natom++;
  nbond++;
  x = (float *) malloc (sizeof (float) * natom);
  y = (float *) malloc (sizeof (float) * natom);
  z = (float *) malloc (sizeof (float) * natom);
  ida = (int *) malloc (sizeof (int) * natom);

  idb = (int *) malloc (sizeof (int) * nbond);
  ida1 = (int *) malloc (sizeof (int) * nbond);
  ida2 = (int *) malloc (sizeof (int) * nbond);

  VET = (int *) malloc (sizeof (int) * nbb);

  for (i = 1; i < natom; i++) //lecture of input file with atoms ID and bonds
    { scanf ("%d%*d%*d%f%f%f%*s%*s%*s", &ida[i], &x[i], &y[i], &z[i]);}

  for (i = 1; i < nbond; i++)
    { scanf ("%d%*d%d%d", &idb[i], &ida1[i], &ida2[i]);}

//RANDOM function is used to find 66 molecules for cross-links
RANDOM(vet_mol, 332);
printf("molecules:\n");
for (i=0;i<66;i++)
printf("%d\n", vet_mol[i]);

//RANDOM function is used to find 66 beads for cross-links
RANDOM(vet_bead, 169);
printf("beads:\n");
for (i=0;i<66;i++)
printf("%d\n", vet_bead[i]);

//vector of ID-beads for cross-links
```

```
for (i=0;i<66;i++)
{ vet_id[i]=169*vet_mol[i]+vet_bead[i];}
printf("ID non enzym\n");
for (i=0;i<66;i++)
printf("%d\n", vet_id[i]);

//NON ENZYMATIC BONDS
distmin=100.;
ll=0;
kk=0;
l=0;
for (i=0; i<66; i++)
    {nenz=vet_id[i];
     xi = x[nenz];
     yi = y[nenz];
     zi = z[nenz];
      for (j = 1; j < natom; j++){
            idl=ida[j];
            x1=x[idl];
            y1=y[idl];
            z1=z[idl];
              dx = abs (xi - x1);
              dy = abs (yi - y1);
            dz = abs (zi - z1);

  if (dx < eps && dy < eps && dz < eps  && idl!=nenz && idl!=nenz+1 && idl!=nenz-1 &&
idl!=nenz-2 && idl!=nenz+2)
          {  dist = sqrt(dx * dx + dy * dy + dz * dz);
             distmin = dist;
             b2=idl;}}
      if(distmin <100) {
         VET[ll]=b2;
         ll++;
         iflag=0;
         for(kk=0;kk<ll;kk++)
           if(VET[kk]==b2) iflag++;
           if(iflag==1 ) {printf("%d 3 %d %d %d\n",nbond+l,nenz,b2,distmin); l++;}}}
}

//function definition
void RANDOM(int vett[100], int num) {
int i;
srand(time(NULL));
for(i=0;i<66;i++)
vett[i] = (rand()%num+1);}


//NON ENZYMATIC BONDS IN THE EXTERNAL SHELL

#include <stdlib.h>           //libraries
#include <stdio.h>
#include <malloc.h>
#include <time.h>

//function declatation
void RANDOM(int vett[100], int num);

int main(){
```

```
float dx,dy,dz,distmin,dist;
float xi,yi,zi,xn,yn,x1,y1,z1,xc,yc,rc,xd2,yd2,r;
  float eps=2.; //cut radius
  float *x, *y, *z;
  int *idb, *ida1, *ida2;
  int *ida;
  int i, j, k, l, ll, kk, ii, n, iflag;
  int natom, nbond, nbb, nenz, idl, b2;
  int *VET;
  int num;
  int vet_moltot[200], vet_mol[100], vet_bead[100], vet_id[100], vet_b1[100],
vet_b2[66];
  nbb = 333;

  scanf ("%d", &natom);
  scanf ("%d", &nbond);
  natom++;
  nbond++;
  x = (float *) malloc (sizeof (float) * natom);
  y = (float *) malloc (sizeof (float) * natom);
  z = (float *) malloc (sizeof (float) * natom);
  ida = (int *) malloc (sizeof (int) * natom);

  idb = (int *) malloc (sizeof (int) * nbond);
  ida1 = (int *) malloc (sizeof (int) * nbond);
  ida2 = (int *) malloc (sizeof (int) * nbond);

  VET = (int *) malloc (sizeof (int) * nbb);

  for (i = 1; i < natom; i++) //lecture of input file with atoms ID and bonds
    { scanf ("%d%*d%*d%f%f%f%*s%*s%*s", &ida[i], &x[i], &y[i], &z[i]);}
  for (i = 1; i < nbond; i++)
    { scanf ("%d%*d%d%d", &idb[i], &ida1[i], &ida2[i]);}

//66 molecules in the external shell for cross-links
xc=9.5;
yc=45;
rc=6.5;
ii=0;
srand(time(NULL));
for(i=0;i<200;i++)
vet_moltot[i] = (rand()%332+1);
for (i=1;i<200;i++)
{printf("%d\n", vet_moltot[i]);
 n=vet_moltot[i]*169;
 xn=x[n];
 yn=y[n];
 xd2=(xn-xc)*(xn-xc);
 yd2=(yn-yc)*(yn-yc);
 r=sqrt(xd2+yd2);
 if(r>rc) {vet_mol[ii]=n/169;ii++;}}
printf("molecules\n");
for(i=0; i<65; i++)
printf("%d\n",vet_mol[i]);

//RANDOM function is used to find 66 beads for cross-links
RANDOM(vet_bead, 169);
 printf("beads:\n");
```

```
for (i=0;i<66;i++)
printf("%d\n", vet_bead[i]);
//vector of ID-beads for cross-links
for (i=0;i<66;i++)
{ vet_id[i]=169*vet_mol[i]+vet_bead[i];}
printf("ID non enzym\n");
for (i=0;i<66;i++)
printf("%d\n", vet_id[i]);

//NON ENZYMATIC BONDS
distmin=100.;
ll=0;
kk=0;
l=0;
for (i=0; i<66; i++)
    {nenz=vet_id[i];
     xi = x[nenz];
     yi = y[nenz];
     zi = z[nenz];
       for (j = 1; j < natom; j++){
            idl=ida[j];
            x1=x[idl];
            y1=y[idl];
            z1=z[idl];
              dx = abs (xi - x1);
              dy = abs (yi - y1);
            dz = abs (zi - z1);
        if (dx < eps && dy < eps && dz < eps  && idl!=nenz && idl!=nenz+1 && idl!=nenz-1
&& idl!=nenz-2 && idl!=nenz+2){
            dist = sqrt(dx * dx + dy * dy + dz * dz);
            distmin = dist;
            b2=idl;}}

      if(distmin <100) {
       VET[ll]=b2;
         ll++;
         iflag=0;
         for(kk=0;kk<ll;kk++)
           if(VET[kk]==b2) iflag++;
           if(iflag==1 ) {printf("%d 3 %d %d\n",nbond+l,nenz,b2); l++;}}}
}

//function definition
void RANDOM(int vett[100], int num) {
int i;
srand(time(NULL));
for(i=0;i<66;i++)
vett[i] = (rand()%num+1);}
```

Appendix C

LAMMPS input files

In the following pages LAMMPS input files used in the work are reported.

- INPUT FILE FOR GLUCOSEPANE 1m/s
- INPUT FILE FOR HYDROXYLYSYL PYRIDINOLINE DIRECTION 2-12 1m/s

All the simulations on glucosepane and hydroxylysyl pyridinoline are performed using the same input script changing the pulling rate following the conversion table here reported:

pulling rate conversion table	
rate [m/s]	rate [Å/fs]
100	10^{-3}
10	10^{-4}
1	10^{-5}
0.5	$5{\cdot}10^{-6}$
0.1	10^{-6}
0.05	$5{\cdot}10^{-7}$

- INPUT FILE FOR TROPOCOLLAGEN MOLECULE 1 BEAD = 6 AA
- INPUT FILE FOR THE CALCULATION OF LENNARD JONES PARAMETERS

- INPUT FILE FOR THE PSEUDO PHYSIOLOGICAL FIBRIL 100 m/s
- INPUT FILE FOR THE PHYSIOLOGICAL FIBRIL 100 m/s
- INPUT FILE FOR AGED FIBRILS 100 m/s: this script can be used for each model of aged fibril but the data.fibrilla file must be changed in order to simulate the whole section cross-linked or the partially cross-linked fibril.

```
# INPUT FILE FOR GLUCOSEPANE 1 m/s
# REAX potential

units       real

atom_style  charge
read_data   data.gcp                      # molecular structure

pair_style  reax/c lmp_control
pair_coeff  * * ffield.reax 1 2 4 3

neighbor    2 bin
neigh_modify      every 10 delay 0 check no

########  groupinfo
group             mypull id 22            # first group for stretching
group             myfix id 1              # second group for stretching

########  minimize

thermo          5000                      # frequency of energy output
thermo_style    custom step ke pe etotal temp
fix             1 all qeq/reax 1 0.0 10.0 1e-6 param.qeq
dump            1 all dcd 5000 glucosepane.dcd  # frequency of trajectory output

min_style       sd
min_modify      dmax 0.5
minimize        0.0 1.0e-8 100 1000

min_style       cg
min_modify      dmax 0.2
minimize        0.0 1.0e-8 100 10000

#######  Equilibration
unfix             1
fix         1 all nve
fix             2 all qeq/reax 1 0.0 10.0 1e-6 param.qeq
fix         3 all langevin 300.0 300.0 1000.0 48279

fix             pull mypull smd cvel 10.0 0.00001 couple myfix auto auto auto 0
               # Stretchig Info
fix         myforce all ave/time 1 100 10000 f_pull[1] f_pull[2] f_pull[3] f_pull[4]
f_pull[5] f_pull[6] f_pull[7] file smd.force    # Stretching output

timestep        0.25
run             6000000
```

```
# INPUT FILE FOR HYDROXYLYSYL PYRIDINOLINE DIRECTION 2-12 1 m/s
# REAX potential
units       real

atom_style  charge
read_data   data.enzym                    # molecular structure

pair_style  reax/c lmp_control
pair_coeff  * * ffield.reax 1 2 4 3

neighbor    2 bin
neigh_modify      every 10 delay 0 check no

########  groupinfo
group             mypull id 2             # first group for stretching
group             myfix id 12             # second group for stretching

########  minimize

thermo          5000                      # frequency of energy output
thermo_style    custom step ke pe etotal temp
fix             1 all qeq/reax 1 0.0 10.0 1e-6 param.qeq
dump            1 all dcd 5000 enzym.dcd  # frequency of trajectory output

min_style       sd
min_modify      dmax 0.5
minimize        0.0 1.0e-8 100 1000

min_style       cg
min_modify      dmax 0.2
minimize        0.0 1.0e-8 100 10000

#######  Equilibration
unfix             1
fix         1 all nve
fix             2 all qeq/reax 1 0.0 10.0 1e-6 param.qeq
fix         3 all langevin 300.0 300.0 1000.0 48279

fix             pull mypull smd cvel 10.0 0.00001 couple myfix auto auto auto 0
               # Stretchig Info
fix         myforce all ave/time 1 100 10000 f_pull[1] f_pull[2] f_pull[3] f_pull[4]
f_pull[5] f_pull[6] f_pull[7] file smd.force    # Stretching output

timestep        0.25
run             6000000
```

```
# INPUT FILE FOR TROPOCOLLAGEN MOLECULE 1 BEAD = 6 AA
# NONREAX potential

units       real
boundary    f f f

atom_style  angle
read_data   data.tropo              # molecular structure

bond_style  piecewise
bond_coeff  * 1.756 1208 0.4 6240

angle_style harmonic
angle_coeff * 4.0 180

pair_style  lj/cut 0
pair_coeff  * * 1.0 1.0 2.5

neighbor    100 bin
neigh_modify      every 10 delay 0 check no

########  minimize

thermo          100                       # frequency of energy output
thermo_style    custom step ke pe etotal temp
dump            1 all dcd 100 tropo.dcd   # frequency of trajectory output

min_style       sd
min_modify      dmax 0.5
minimize        0.0 1.0e-8 100 1000

min_style       cg
min_modify      dmax 0.2
minimize        0.0 1.0e-8 100 10000

#######  Equilibration
fix         1 all nve
fix         2 all langevin 300.0 300.0 1000.0 12345


timestep        10
run             16000000
```

```
# INPUT FILE FOR THE CALCULATION OF LENNARD JONES PARAMETERS
# NONREAX potential

units       real
boundary    f f f

atom_style  full
read_data   data.fibrilla                 # molecular structure

bond_style  piecewise
bond_coeff  * 1.756 1208 0.4 6240

angle_style harmonic
angle_coeff * 4.0 180

pair_style  lj/cut 16
pair_coeff  * * 7 1.425

neighbor    1 bin
neigh_modify      every 100 delay 0 check no


########  minimize

thermo          100                       # frequency of energy output
thermo_style    custom step ke pe etotal temp
dump            1 all dcd 100 fibrilla.dcd      # frequency of trajectory output

min_style       sd
min_modify      dmax 0.5
minimize        0.0 1.0e-8 100 1000

min_style       cg
min_modify      dmax 0.2
minimize        0.0 1.0e-8 100 10000

#######  Equlibration
fix         1 all nve
fix         2 all langevin 300.0 300.0 1000.0 12345


timestep        10
run             1000000
```

 # INPUT FILE FOR THE PSEUDO PHYSIOLOGICAL FIBRIL 100 m/s

```
# NON REAX potential

units       real
boundary    f f f

atom_style  angle
read_data   data.fibrilla                 # molecular structure

bond_style   piecewise
bond_coeff   1 1.756 1208 0.4 6240

angle_style harmonic
angle_coeff 1 4.0 180

pair_style  lj/cut 16
pair_coeff  * * 7 1.425

neighbor    1 bin
neigh_modify      every 100 delay 0 check no

########  groupinfo
group             mypull id 50869 51038 51207 51376 51545 51714 51883 52052 52221 52390
52559 52728 52897 53066 53235 53404 53573 53742 53911 54080 54249 54418 54587 54756 54925
55094 55263 55432 55601 55770 55939 56108             # first group for stretching

group             myfix id 1 170 339 508 677 846 1015 1184 1353 1522 1691 1860 2029 2198
2367 2536 2705 2874 3043 3212 3381 3550 3719 3888 4057 4226 4395 4564 4733 4902 5071
5240              # second group for stretching


########  minimize

thermo          100                       # frequency of energy output
thermo_style    custom step ke pe etotal temp
dump            1 all dcd 1000 fibrilla.dcd     # frequency of trajectory output

min_style       sd
min_modify      dmax 0.5
minimize        0.0 1.0e-8 100 1000

min_style       cg
min_modify      dmax 0.2
minimize        0.0 1.0e-8 100 10000

#######  Equlibration
fix         1 all nve
fix         2 all langevin 300.0 300.0 1000.0 12345

fix             pull mypull smd cvel 4000.0 0.0001 couple myfix auto auto auto 0
      # Stretchig Info
fix         myforce all ave/time 1 50 1000 f_pull[1] f_pull[2] f_pull[3] f_pull[4]
f_pull[5] f_pull[6] f_pull[7] file smd.force    # Stretching output

timestep        10
run             100000000
```

```
# INPUT FILE FOR THE PHYSIOLOGICAL FIBRIL 100 m/s
# NON REAX potential

units       real
boundary    f f f

atom_style  angle
read_data   data.fibrilla                 # molecular structure

bond_style   piecewise
bond_coeff   1 1.756 1208 0.4 6240
bond_coeff   2 1.6 0 2.5 610.7 3.1

angle_style harmonic
angle_coeff 1 4.0 180

pair_style  lj/cut 16
pair_coeff  * * 7 1.425

neighbor    1 bin
neigh_modify      every 100 delay 0 check no

########  groupinfo
group             mypull id 50869 51038 51207 51376 51545 51714 51883 52052 52221 52390
52559 52728 52897 53066 53235 53404 53573 53742 53911 54080 54249 54418 54587 54756 54925
55094 55263 55432 55601 55770 55939 56108             # first group for stretching

group             myfix id 1 170 339 508 677 846 1015 1184 1353 1522 1691 1860 2029 2198
2367 2536 2705 2874 3043 3212 3381 3550 3719 3888 4057 4226 4395 4564 4733 4902 5071
5240              # second group for stretching


########  minimize

thermo          100                       # frequency of energy output
thermo_style    custom step ke pe etotal temp
dump            1 all dcd 1000 fibrilla.dcd     # frequency of trajectory output

min_style       sd
min_modify      dmax 0.5
minimize        0.0 1.0e-8 100 1000

min_style       cg
min_modify      dmax 0.2
minimize        0.0 1.0e-8 100 10000

#######  Equlibration
fix         1 all nve
fix         2 all langevin 300.0 300.0 1000.0 12345

fix             pull mypull smd cvel 4000.0 0.0001 couple myfix auto auto auto 0
      # Stretchig Info
fix         myforce all ave/time 1 50 1000 f_pull[1] f_pull[2] f_pull[3] f_pull[4]
f_pull[5] f_pull[6] f_pull[7] file smd.force    # Stretching output

timestep        10
run             100000000
```

```
# INPUT FILE FOR AGED FIBRILS 100 m/s
# NONREAX potential

units       real
boundary    f f f

atom_style  angle
read_data   data.fibrilla                 # molecular structure

bond_style      piecewise
bond_coeff      1 1.756 1208 0.4 6240
bond_coeff      2 1.6 0 2.5 610.7 3.1
bond_coeff      3 2.7 766.0 1

angle_style harmonic
angle_coeff 1 4.0 180

pair_style  lj/cut 16
pair_coeff  * * 7 1.425

neighbor    1 bin
neigh_modify      every 100 delay 0 check no

########  groupinfo
group             mypull id 50869 51038 51207 51376 51545 51714 51883 52052 52221 52390
52559 52728 52897 53066 53235 53404 53573 53742 53911 54080 54249 54418 54587 54756 54925
55094 55263 55432 55601 55770 55939 56108             # first group for stretching

group             myfix id 1 170 339 508 677 846 1015 1184 1353 1522 1691 1860 2029 2198
2367 2536 2705 2874 3043 3212 3381 3550 3719 3888 4057 4226 4395 4564 4733 4902 5071
5240              # second group for stretching


########  minimize

thermo          100                       # frequency of energy output
thermo_style    custom step ke pe etotal temp
dump            1 all dcd 1000 fibrilla.dcd     # frequency of trajectory output

min_style       sd
min_modify      dmax 0.5
minimize        0.0 1.0e-8 100 1000

min_style       cg
min_modify      dmax 0.2
minimize        0.0 1.0e-8 100 10000

#######  Equlibration
fix         1 all nve
fix         2 all langevin 300.0 300.0 1000.0 12345

fix             pull mypull smd cvel 4000.0 0.0001 couple myfix auto auto auto 0
      # Stretchig Info
fix         myforce all ave/time 1 50 1000 f_pull[1] f_pull[2] f_pull[3] f_pull[4]
f_pull[5] f_pull[6] f_pull[7] file smd.force    # Stretching output

timestep        10
run             100000000
```

Bibliography

[1] Li et Al. Advanced glycation end-products diminish tendon collagen fiber sliding. Department of Orthopaedics, Balgrist Hospital, University of Zurich, Switzerland. Department of Health Sciences and Technology, ETH Zurich, Switzerland.

[2] Monnier et al. Nonenzymatic browning in vivo: possible process for aging of long-lived proteins. *Science*, 211:492–493, 1981.

[3] Sell et Al. Glucosepane is the major protein cross-link of the senescent human extracellular matrix. *The Journal of Biological Chemistry*, 280:12310–12315, 2005.

[4] Montevecchi Redaelli. *Biomeccanica. Analisi multiscala di tessuti biologici*. Ptron editore, 2007.

[5] Peter Frantzl, editor. *Collagen. Structure and Mechanics*. Springer, 2006.

[6] Markus J. Buehler. Atomistic and continuum modeling of mechanical properties of collagen: elasticity, fracture and self-assembly. *Journal of Material Research*, 21:1947–1961, 2006.

[7] Svensson et Al. Fracture mechanics of collagen fibrils: influence of natural cross-links. *Biophysical Journal*, 104:2476–2484, 2013.

[8] Eyre et Al. Advanced in collagen cross-link analysis. *Methods*, 45:65–74, 2008.

[9] Bailey Avery. The effects of the maillard reaction on the physical properties and cell interactions of collagen. *Pathologie Biologie*, 54:387–395, 2006.

[10] Verzijl et Al. Crosslinking by advanced glycation end products increases the stiffness of the collagen network in human articular cartilage. *Arthritis & Rheumatism*, 46:114–123, 2002.

[11] Biemel et Al. Identification and quantification of major maillard cross-links in human serum albumin and lens protein. *The Journal of Biological Chemistry*, 277:24907–24915, 2002.

[12] Kadler et Al. Collagen fibril formation. *Biochemistry*, 316:1–11, 1996.

[13] Williams et Al. Collagen fibril formation. optimal in vitro conditions and preliminary kinetic results. *The Journal of Biological Chemistry*, 253:6578–6585, 1978.

[14] Liu et Al. In vitro regulation of single collagen fibril ength by buffer compositions and temperature. *Bio-Medical Materials and Engineering*, 15:413–420, 2005.

[15] Raspanti et Al. Collagen fibril ctructure is affected by collagen concentration and decorin. *Biomacromolecules*, 8:2087–2091, 2007.

[16] Raspanti et Al. Glycosaminoglycans show a specific periodic interaction with type i collagen fibrils. *Journal of Structural Biology*, 164:134–139, 2008.

[17] F. Barthelat J. Poissant. In situ mechanical testing of hydrated biological nanofibers using a nanoindenter transducer. *Experimental Mechanics*, 52:1287–1295, 2012.

[18] Yang et Al. Micromechanical bending of single fibrils using tomic force microscopy. *Journal of Biomedical Research A*, 82A:160–168, 2006.

[19] van der Rijt et Al. Micromechanical testing of individual collagen fibrils. *Macromolecular Bioscience*, 6:697–702, 2006.

[20] Yang et Al. Micromechanical analysis of native and cross-linked collagen type i fibrils support the existence of microfibrils. *Journal of the mechanical behavior of biomedical materials*, 6:148–158, 2012.

[21] Shen et Al. Stress-strain experiments on individual collagen fibrils. *Biophysical Journal*, 95:3956–3963, 2008.

[22] Eppel et Al. Nano measurements with micro-devices: mechanical properties of hydrated collagen fibrils. *Journal of the Royal Society Interface*, 3:117–121, 2006.

[23] Buehler Markus J. Nanomechanics of collagen fibrils under varying cross-link densities: atomistic and continuum studies. *Journal of the mechanical behavior of biomedical materials*, 1:59–67, 2008.

[24] Gautieri et al. Viscoelastic properties of model segments of collagen molecules. *Matrix Biology*, 31:141–149, 2012.

[25] Danilo Roccatano. Introduzione alla dinamica molecolare. Dipartimento di Chimica, Ingegneria Chimica e Materiali. Universit de L'Aquila. roccata@caspur.it.

[26] Brooks et Al. Charmm: The biomolecular simulation program. *Journal of Computational Chemistry*, 30:1545–1615, 2009.

[27] James C. Phillips et Al. Scalable molecular dynamics with namd. *Journal of Computational Chemistry*, 26:1781–1802, 2005.

[28] Simon Plimpton. Fast parallel algorithms for short-range molecular dynamics. *Journal of Computational Phys.*, 117:1–19, 1995.

[29] Humphrey et Al. Vmd - visual molecular dynamics. *J. Mol. Graphics*, 14:33–38, 1996.

[30] Gautieri et Al. Coarse-grained model of collagen molecules using an extended martini force field. *J. Chem. Theory Comput.*, 6:1210–1218, 2010.

[31] Gautieri et Al. Deformation rate controls elasticity and unfolding pathway of single tropocollagen molecules. *Journal of the Mechanical Behavior of Biomedical Materials*, 2:130–137, 2009.

[32] Andrea Mezzanzanica. Studio della risposta meccanica dei glicosamminoglicani della cartilagine mediante modelli molecolari. 2013.

[33] Uzel et Al. Molecular structure, mechanical behavior and failure mechanism of the c-terminal cross-link domain in type i collagen,. *Journal of the mechanical behavior of biomedical materials*, 4:153–161, 2011.

[34] Fang et Al. Type l collagen self-assembly: The roles of substrate and concentration. *Langmuir*, 29:2330–2338, 2013.

[35] Bustamante et al. Grabbing the cat by the tail: manipulating molecules one by one. *Nature reviews*, 1:130–136, 2000.

[36] Grandbois et Al. How strong is a covalent bond? *Science*, 283:1727–1730, 1999.

[37] Gutsmann et al. Force spectroscopy of collagen fibers to investigate their mechanical properties and structural organization. *Biophysical journal*, 86:3186–3193, 2004.

[38] Sun et al. Stretching type ll collagen with optical tweezers. *Journal of biomechanics*, 37:1665–1669, 2004.

[39] Svensson et Al. Mechanical properties of human patellar tendon at the hierarchical levels of tendon and fibril. *J. Appl. Physiol.*, 112:419–426, 2012.

[40] Gautieri et al. Hierarchical structure and nanomechanics of collagen microfibrils from the atomistic scale up. *Nano Lett.*, 11:757–766, 2011.

Ringraziamenti

Ringraziamo le nostre famiglie, per tutto.
Ringraziamo Simone, Alfonso, Mario, Manuela, Jess, Unai e Nils per tutti i consigli e per averci supportato e sopportato in questo lungo periodo.
Ringraziamo gli amici di sempre che hanno condiviso con noi questa avventura.
Ringraziamo Stiffler ed Ares per aver lavorato anche in agosto, 24 ore su 24.
Infine, ringraziamo Lauderica.