Computational Approaches to the Calculation of Spectroscopic, Structural and Mechanical Properties of Polysaccharide Chains

by

Padmesh Anjukandi

A thesis submitted in partial fulfillment for the
degree of Doctor of Philosophy
at Massey University

September 2010

Declaration of Authorship

I, Padmesh Anjukandi, declare that this thesis titled, 'Computational Approaches to the Calculation of Spectroscopic, Structural and Mechanical Properties of Polysaccharide Chains' and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: Padmesh Anjukandi

Date: 08/09/2010

"God Gave Me Nothing I Wanted, He Gave Me Everything I Needed."

Swami Vivekananda...

Abstract

In this thesis atomistic, statistical mechanical and coarse grained simulation techniques are used to study the properties of biopolymers and in particular the plant polysaccharide pectin. Spectroscopic aspects, structural and conformational behavior, and mechanical properties of the molecule in different physical states are addressed.

After an introduction to the area and the theoretical techniques utilised herein (chapter 1), chapter 2 deals with the spectroscopic characterisation of pectin. Spectra were obtained theoretically by undertaking complete energy minimisation and Hessien calculations using DFT techniques implemented in Gamess (PC & US) packages. The calculated IR absorptions of different pectinic species and oligomers coupled on different surfaces were compared with experimental results. Herein, it is confirmed that experimental FTIR studies coupled with DFT calculations can be used as an effective tool for the characterisation of pectin, and studying chemical coupling of the biopolymer to surfaces.

In chapter 3, the properties of single chain polymer systems in controlled solvent conditions were studied using Brownian dynamics simulations, motivated by the formation of secondary structure architectures in biopolymer systems. We focus on the conformational properties of the chain in the presence of an additional torsional potential. New, interesting, and biologically relevant structures were found at the single molecule scale when a torsional potential was considered in the calculations.

In chapter 4, results from DFT calculations carried out on single pectin sugar molecules (lengths and the free energies) are incorporated into a statistical mechanical model of polymer stretching, in order to obtain the force-extension behaviour of a single molecule pectin. This captures a good deal of the phenomenology of the single molecule stretching behavior of pectin.

Chapter 5 summarises the conclusions of the work and finally chapter 6 suggests direction for further work.

Acknowledgements

There are many people whom I would like to thank for making my stay at IFS, Massey enjoyable. First of all, I would like to thank my supervisor, Bill, for guiding me through my research and for helping me overcome the usual hurdles of my Doctoral research. Over the years he motivated me and helped a lot shaping my research career (but as far as supporting Arsenal, still no compromise). I was also lucky to have Gerald as my co-supervisor and Mark Waterland for his help with Quantum Chemistry works. Bill and Gerald always found time to talk with me if I needed their help or advice. A special thanks to all the administrative staff at IFS, Massey, who have been helpful beyond words through the three years of my PhD. I also acknowledge Marsden and the MacDiarmid Institute for my Ph.D funding and for providing access to the computational facilities (Bluefern, UCSC).

I would like to mention my boundless thanks to Prof. K. L. Sebastian for being the first person to expose me to the world of scientific research. During the six months stay there at the IPC, IISc, Bangalore, I had benefited a lot by working with an exceptionally experienced scientist like him. I am deeply indebted to my teachers Prof. M. Padmanabhan, Dr. A. S. Padmanabhan, Dr. Pius Kuruvilla, Mr. Vijayan, Miss. Sujaya, and Mr. Krishnakumar for their valuable guidance throughout my career and helping me mold my career.

I would also like to thank all the past and present Biomaterials group members. I will always be grateful to Jonny for helping me adjust with the new atmosphere and getting me started up with my work. Romaric, Aurelie, Abdanour, Erich, Rob, Van, Davide, Ive really enjoyed working with you guys and thanks for sharing the office. Special thanks to Naser, Steve and Aaron for all those wonderful discussions in and out of the office, especially you Naser, the discussions in the coffee room were lovely. I wish you guys all the best in your current and future endeavours.

My stay at Massey would not be nearly as memorable had it not been for some wonderful friends that I made along: Bo Zhao, Naser, Roony, Rajesh, Manoj, Shijoy and Antoney. I am really glad that I got to know them, and I truly hope that our friendships will continue after we all leave (some of them left already) Massey. It was my friends at the flatting place that kept me happy and sane even when I

had hit dead-ends in my research and nothing seemed to work. Hari, Sachin bhai, Pranav, Litty, Poonacha and Robin, you people were always there for me. Thank you very much for all your help and friendship.

Thanks to all who have been there with me at different stages of my PhD, to share all my joys and sorrow. Friends and seniors from Mahatma Gandhi University; Vinod, Vivek, Renji, Rani, Shyla, Resmi, Jousheed, Sreeprasad, Habeeb, Anumon, Sreedevi, Jisha, Babby and Jessy for their constant support and friendship during my doctoral years. Also many thanks to those lovely people at Sir Syed college, Jithesh, Shabeer, Viju, Rejesh, Meby, Nikesh, Devidas, Sheik. It was wonderful to be with you guys.

None of what I have achieved, in sport or academia, would have been possible without the support of my family. Even since I can remember myself, my parents have been there for me, carefully guiding me and supporting my decisions without exerting pressure to behave a certain way or to choose a certain path. I would like to thank my father (late: A Sreenivasan) my mother (Padmini), my brothers (Sunil and Sujith) and my sis (Retty); they have always supported my dreams and aspirations. Throughout my endeavours, their love, support, guidance, and endless patience have been truly inspirational - "thanks" will never suffice.

Finally, I would like to thank everybody who was important to the successful realization of this thesis, as well as expressing my sincere apology that I could not mention personally one by one.

Padmesh Anjukandi

...

Contents

List of Figures

List of Tables

Abbreviations

FJC	**F**reely **J**ointed **C**hain
FRC	**F**reely **R**otating **C**hain
WLC	**W**orm **L**ike **C**hain
AFM	**A**tomic **F**orce **M**icroscopy
e-WLC	**e**xtensible **W**orm **L**ike **C**hain
DFT	**D**ensity **F**unctional **T**heory
HF	**H**artree **F**ock
MP	**M**Ø**ller** **P**lesset
CC	**C**oupled **C**luster
GVB	**G**eneralised **V**alence **B**ond
MCSCF	**M**ulti **C**onfigurations **S**elf **C**onsistant **F**ield
KS	**K**ohn **S**ham
MD	**M**olecular **D**ynamics
BD	**B**rownian **D**ynamics
LJ	**L**ennard **J**ones
FENE	**F**inite **E**xtensible **N**onlinear **E**longation
DP	**D**egree of **P**olymerisation
DM	**D**egree of **M**ethylesterification

Publications

M.A.K. Williams, A.T. Marshall, P. Anjukandi and R. G. Haverkamp; Investigation of the effects of fine structure on the nanomechanical properties of pectin, Phys. Rev. E, 76, 021927, 2007.

Abdenor Fellah, Padmesh Anjukandi, Mark Waterland and Martin A.K. Williams; Determining the degree of methylesterification of pectin by ATR/FT-IR : methodology optimisation and comparison with theoretical calculations, Carbohydrate polymers, 78(4), 847-853, 2009.

Padmesh Anjukandi, Gerald G. Pereira and Martin A.K. Williams; Langevin Dynamics Simulations Reveal Biologically Relevant Folds Arising from the Incorporation of a Torsional Potential, Journal of Theoretical Biology, 265(3), 245-249, 2010.

Anjukandi, P., Fellah, A., Patty, P., and Williams, M. A. K; Immobilisation of Pectin on Solid Surfaces, its Characterisation and Theoretical Calculations, Manuscript in preperation, 2010.

Dedicated to my beloved father . . .

Chapter 1

Introduction.

Polymers are substances whose molecules have high molar masses and are composed of a large number of repeating units. Such molecules are composed of long sequences of one or more types of atoms or groups of atoms linked to each other by primary, usually covalent, bonds [1]. The long chain nature of the macromolecules is responsible for many characteristic properties of the polymer. There are both naturally occurring and synthetic polymers. Among naturally occurring polymers are proteins, starches, cellulose, and latex. Synthetic polymers are produced commercially on a very large scale and have a wide range of properties and uses. The materials commonly called plastics are all synthetic polymers. Polymers are formed by chemical reactions in which a large number of molecules called monomers (the repeating units) are joined sequentially, forming a chain.

Biopolymers are a group of polymers synthesized by living organisms. There are three main sets of biopolymers which sustain all life, namely the polysaccharides, polypeptides, and polynucleotides, whose monomer units are respectively constituted by sugars, amino acids and nucleic acids. Many studies of these polymers have been carried out in order to get a deeper understanding of their structure-function relations. Much of this work has focused on structural polymers such as the proteins and carbohydrates, the building blocks of all organisms. Carbohydrates are the most abundant class of organic compounds found in living

organisms. They have many roles such as being used in the storage and transport of energy and as structural components in plants and animals. They also serve as lubricants [2] and provide support to the fibrous and cellular elements of the tissue. Most of the carbohydrates found in nature occur in the form of high molecular weight polymers called polysaccharides. Polysaccharides composed of a single type of monosaccharide unit are known as the homopolysaccharides and those composed of more than one type are called the heteropolysaccharides.

1.1 Pectins: Introduction to Pectin and its Properties.

Pectins are a complex set of plant polysaccharides that are present in the primary cell walls (the non-woody part) of terrestrial plants [3]. They are also seen in the middle lamella between plant cells, where they help to bind cells together. This heterogeneous polysaccharide has a complex structure in vivo. Nevertheless the pectin chain backbone is mainly (1-4)-linked α-D-galacturonic acid with substitution by some amount of its methyl ester (fig 1.1) [4]. Around 80% of the carboxyl groups in naturally occurring galacturonic acid are esterified with methanol. The ratio of the methyl esterified galacturonic acid monomer to that of the total number of monomers present in the polymer chain is known as the degree of methyl esterification (DM) of the polymer chain. The properties of pectins depend on this DM [5, 6] and this is how pectins are classified as high and low esterified pectins, with more or less than half of all the galacturonic acid esterified respectively. In certain cases, some ethyl esterified [7] and amidated [8] analogues have been synthesized, but they do not occur naturally. Some plants like sugar beet, potatoes and pears contain pectins with acetylated galacturonic acid [9] in addition to methyl esters. In between the regular array of the polygalacturonate, the pectin structure in plants is interrupted by (1-2)-linked rhamnose residues. They also contain branches of side chains of up to 20 residues, of mainly L-arabinose and D-galactose, giving pectin a complex structure. But generally, almost all of the

mechanical properties of pectin are governed by the dominant polygalacturonate back bone and during the extraction of the pectin, the side chains are usually removed and DM of the final pectin sample will depend not only on its in vivo composition but also on how it is extracted and processed. Due to their high viscosity and gelling properties, extracted pectins are mainly used as gelling agents, and as thickeners and stabilizers in the food industry.

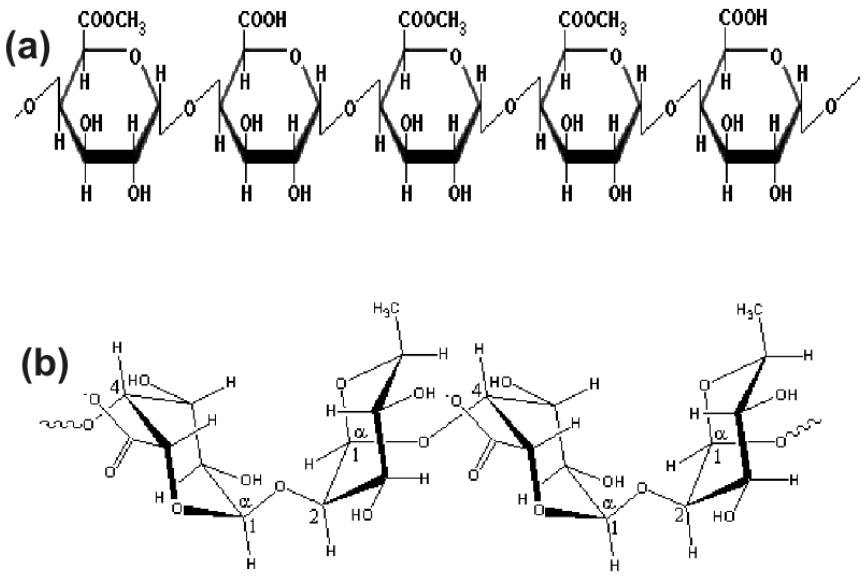

FIGURE 1.1: (a) The chemical structure of pectin. (b) The 1-4 linkage in pectin with the initial ($^4C_1$) pyranose conformation.

1.2 Polymer Chain Models

Polymers are composed of sequentially bonded monomeric units. As the bonds joining the monomers are mostly free to rotate (flexible), they can adopt a very

large number of accessible microscopic states. At a given temperature T, the probability of ith configuration P$_i$ is given by

$$P_i = \frac{\exp(-\beta E_i)}{Q}. \tag{1.1}$$

where $\beta = \frac{1}{k_B T}$, with k_B, the Boltzmann's constant, E_i is the energy of the i^{th} configuration and Q is the canonical partition function. The knowledge of Q provides complete thermodynamic information about the system. Thus, it is often a key quantity to be calculated in statistical mechanical studies, but it can be seldom known exactly. However a great deal can be learnt about the average properties of a polymer using a very simple model in which, all conformations, obtained by free rotations about the bonds are given equal importance. Experimental measurements of polymer properties typically involve averages over a large ensemble of identical molecules or are carried out on time scales over which many possible configurations are sampled.

Chain models can be sub-divided into those dealing with *Ideal Chains* or *Real Chains* [10]. An *Ideal Chain* represents the polymer as a random walk and neglects any kind of interactions among monomers. Thus it is a simple model for a polymer chain, but nevertheless in general is a good model to understand certain polymer physics. *Real Chains* on the other hand, take into account much more indepth knowledge of the chain and the interactions between chain monomers can be modeled here by incorporating the excluded volume interactions. A few important polymer chain models are the 'Random walk model', the 'Gaussian model or the Freely Jointed Chain model', the 'Freely Rotating Chain model', the 'Worm Like Chain (WLC) model' etc [11] and these are breifly described below.

1.2.1 The Lattice Model.

The lattice model or the random walk model can be used to study the statistical properties of long molecules as it provides a simple and convenient description of chain configurations in solutions. It was introduced by Einstein in 1905 [12] and Pearson in 1905-06 [13, 14], to find the probability that a walker covers a distance R in N steps assuming that each step is taken at random.

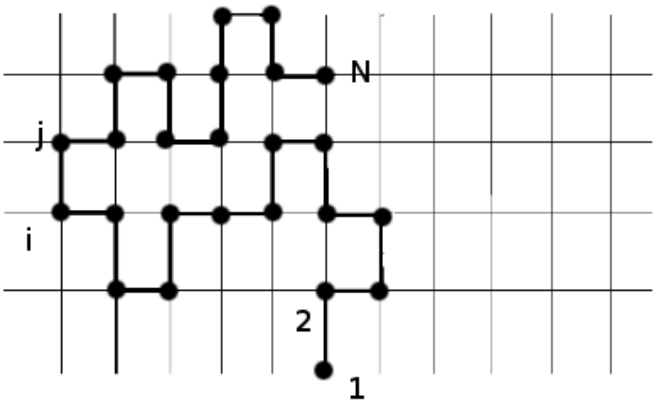

FIGURE 1.2: Random walk model on a periodic lattice (two dimensional).

The random walk model views the polymer as a sequence of uncorrelated bond vectors of fixed length and all other microscopic details are ignored. The probability 'P' that the chain has a particular configuration can be written in terms of the random distribution of a bond vector of constant length l, ie $P(R_i)$, that a given bond i has a certain orientation

$$\begin{aligned} \mathbf{R_i} &= \mathbf{r_i} - \mathbf{r_{i-1}} \\ P &= \prod_{i=1}^{N} P(\mathbf{R_i}). \end{aligned} \tag{1.2}$$

'$\mathbf{r}_i$' defines the position vector for the i^{th} segment. The model accurately describes the bulk behavior of dilute polymer solutions. The physics holds adequately when the chains are comparatively large as only the back bone connectivity is accounted for.

1.2.2 Freely Jointed Chain (FJC) and Freely Rotating Chain (FRC) Model.

Among the non-lattice models of polymer chains, the FJC is the simplest and is most convenient for the study of many properties of polymer systems [15]. Being a Gaussian model, the chain conformation can be represented by the set position vectors $S = \{r_0, r_1.....,r_n\}$ of N beads or segments, which can be thought of as N repeating monomer units of the chain. The conformation of an ideal macromolecule coincides with the random walk path of a Brownian particle. Let $P(\mathbf{r_n}, \mathbf{r_{n-1}})$ represent the linear memory that describes the bond between a pair of linked neighbors.

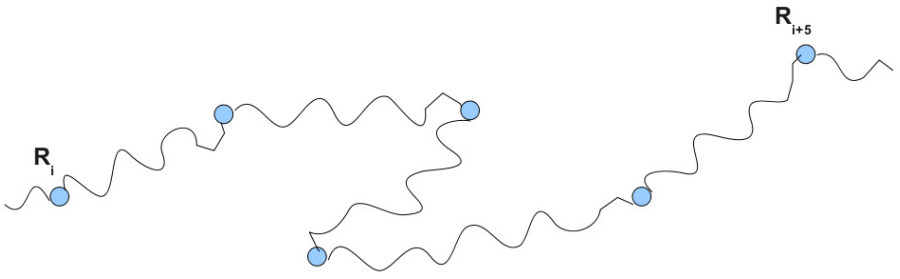

FIGURE 1.3: A Gaussian chain (or the bead spring) model.

The memory of the chain direction is lost over a distance comparable to its persistence length, persistence length$\approx l$. The probability of a given polymer configuration S is thus given by $P[S] = \prod_{n=2}^{N} P(\mathbf{r_n}, \mathbf{r_{n-1}})$ and represents the connectivity of the ideal macromolecule. For the bead-spring model, the bond lengths have the Gaussian distribution, such that $P(\mathbf{r_n}, \mathbf{r_{n-1}})$ is defined as follows;

$$P(\mathbf{r_n}, \mathbf{r_{n-1}}) = \left(\frac{3}{2\pi l^2}\right)^{\frac{3}{2}} \exp\left(-\frac{3(\mathbf{r}_n - \mathbf{r}_{n-1})^2}{2l^2}\right). \tag{1.3}$$

Like the FJC model, the freely rotating chain model (FRC) also consists of n bonds, but unlike the FJC, they are at a fixed angle. Hence the correlations between bond directions are specified by the bond angle θ.

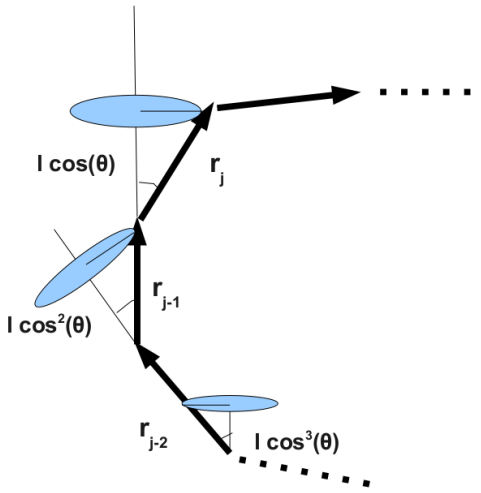

FIGURE 1.4: A freely rotating chain model.

The mean square end-to-end distance in a FRC is given by

$$\left\langle \mathbf{R}^2 \right\rangle = \langle \mathbf{R}_i . \mathbf{R}_j \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} \langle \mathbf{r}_i . \mathbf{r}_j \rangle \,, \tag{1.4}$$

and the correlation between the bond vectors $\langle \mathbf{r}_i . \mathbf{r}_j \rangle$ is given by

$$\langle \mathbf{r}_n . \mathbf{r}_n \rangle = l^2 (\cos\theta)^{|j-i|}. \tag{1.5}$$

Where, 'l' is the length of the individual bond vector. If there is no correlation between the adjecent bonds, then $< \cos\theta_{ij} >= 0$ for all $i \neq j$. For all $i = j$,

$< \cos\theta_{ij} > = 1$ and thus the mean square end-to-end distance of a FJC with 'n' uncorrelated bonds then converges to

$$\left\langle \mathbf{R}^2 \right\rangle = nl^2. \tag{1.6}$$

In general however, polymer chains do have correlations among the adjecent bond vectors and this needs to be incorporated into the mean square displacement. Now, $< \cos\theta_{ij} > \neq 0$, but still there is no interaction between the distant bond vectors and thus we get $\lim_{|i-j|\to\infty} < \cos\theta_{ij} >= 0$. The correlation C_i' for a bond vector 'i' over all the 'j' bond vectors is given by $\sum_{j=1}^{n} < \cos\theta_{ij} >$. Thus, the mean square end-to-end displacement in this case is given by

$$\left\langle \mathbf{R}^2 \right\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} < \cos\theta_{ij} >= l^2 \sum_{i=1}^{n} C_i' = C_n n l^2 \tag{1.7}$$

C_n is called the Flory's characteristic ratio. A freely jointed chain has got 'N' free segments of length 'l' jointed together and these segments are known as the 'Kuhn segment' or the 'Kuhn length'.

Using the probability distribution for the end-to-end length, we can evaluate the partition function and from the partition function, we can derive all other thermodynamic quantities. Thus, from this distribution, an expression for extended length of the chain due to an end force on the polymer chain can be calculated and for a FJC and FRC, [16, 17]

$$x[F] = [coth(\frac{Fl}{k_B T}) - (\frac{k_B T}{Fl})]L. \tag{1.8}$$

where l is the length of the Kuhn segment, L, the contour length of the chain and F is the force with which the chain is being pulled.

1.2.3 Wormlike (WLC) or Kratky-Porod Chain Model.

The wormlike chain can be considered as a limiting case of the FRC, where the rotating bond angle takes up very small values [$\cos(\theta) \approxeq 1 - \frac{\theta^2}{2}$]. It can be pictured as a thin, elastic tube which obeys Hooke's law at low deformations. The WLC model well describes stiff polymer chains [18]. The mean square end-to-end distance in the case of a WLC is given by

$$\left\langle \mathbf{R}^2 \right\rangle = l^2 \sum_{i=1}^{n} \sum_{j=1}^{n} (\cos\theta)^{|j-i|} = l^2 \sum_{i=1}^{n} \sum_{j=1}^{n} \exp(-\frac{|j-i|}{l_p} l), \tag{1.9}$$

$$\left\langle \mathbf{R}^2 \right\rangle = 2 l_p \mathbf{R}_{max} - 2 l_p^2 [1 - exp(-\frac{\mathbf{R}_{max}}{l_p})]. \tag{1.10}$$

Here, $\mathbf{R}_{max}$ is the constant chain contour length and l_p is the persistence length of the chain. An expression for the extension ($\frac{l}{l_c}$) due to an end force on the polymer chain in the case of a WLC can be obtained as

$$F = \frac{k_B T}{l_p} [\frac{1}{4(1 - \frac{l}{l_c})^2} + \frac{l}{l_c} - \frac{1}{4}], \tag{1.11}$$

where l_c is the contour length. Unlike the other chain models, the WLC also takes in to account the stiffness of the chain through the persistence length (which can be considered as the ratio of the bending stiffness to the thermal energy). Thus they more closely mimic biopolymers that are often semi-flexible in nature.

1.3 Mechanical properties of Polymers: Pulling Single Chain Polymers.

Pulling single polymer chains in order to study their physical properties has remained a hot area of studies, ever since the theoretical consideration of the problem

by *P J Flory* [19] and *P G deGennes* [20]. Standard elastic materials give standard *'Hooke's law'* behavior upon deformation, ie a linear stress-strain relation or a linear "force curve". Different polymer physics models also show similar force-extension behavior at low forces (or at very high forces when the chain is stretched beyond its contour length). The force curve initially has a linear, Hookean region, where the entropy of the chain generates the restoring force. The FJC and WLC chain both predict that the force then increases more rapidly, and eventually diverges at the contour length. Figure 1.5 [21] gives us an idea of the force-extension behavior of different polymer models.

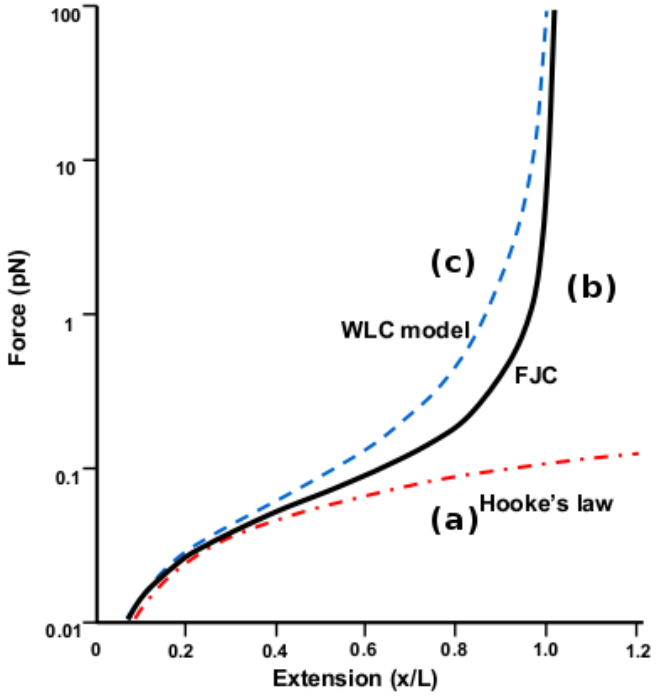

FIGURE 1.5: (a) A typical force-extension curve of a elastic hookean spring. (b) Force curve observed for FJC model. (c) Force curve for a WLC chain. After Bustamante et al.

The development of experimental techniques such as the Atomic Force Microscopy (AFM) [22] and Optical Tweezers (OT) [23] have added a great deal to the study

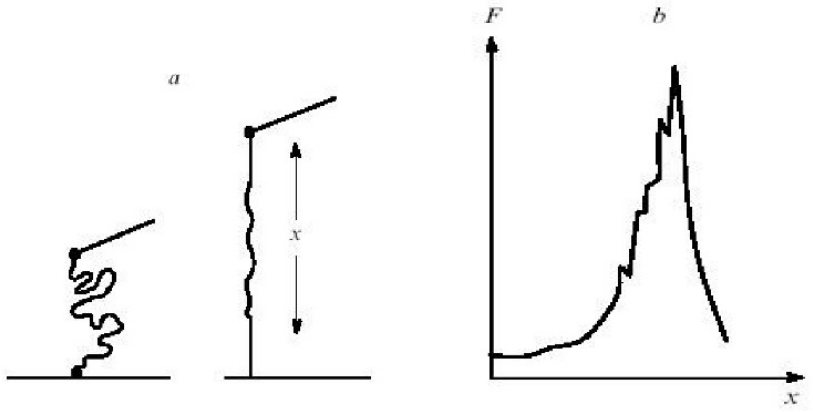

FIGURE 1.6: (a)Schematic diagram for the stretching of a polymer chain by AFM. (b) The nature of the force extension profile of a polymer chain as obtained from the AFM experiment.

of the mechanical properties of single chain polymers and biopolymers, allowing the experimental determination of the elastic properties of these molecules. In AFM force spectroscopy measurements, the experimental restoring force (measured from the deflection of the tip of a cantilever) is studied as a function of the distance between the tip and the surface, between which the polymer is attached. The pulling of the polymer chain is attained by the movement of the surface up and down, mounted on a piezoelectric material. In most stretching experiments, a single molecule absorbed between a substrate and the cantilever of the AFM, is extended vertically (fig 1.6(a)) at a constant rate while measuring the resultant force from the deflection of the cantilever. Figure 1.6(b) is the schematic of a force-extension measured by an AFM experiment. The irregularities in the curve show noise and the sharp drop of the force at the end of the profile shows the detachment of the chain from the AFM tip or the chain snapping. Like the AFM instrument, Optical Tweezers are also used for manipulating nanometer and micrometer sized objects by giving them a small force by the means of highly focused laser beam. Of late, these instruments have proved be of great use in nanomechanical studies of biopolymers and especially in the single molecule stretching at low forces. Generally such pulling studies have been helpful in the analysis of the conformational chemistry of proteins and polysaccharides and in providing insights into how such

processes are important in a range of living systems.

1.3.1 Pulling of Proteins and Polysaccharides.

The introduction of AFM and OT techniques have made a revolutionary change in the studies of the nano-mechanical properties of biopolymers like the proteins [24] and polysaccharides [25], which form the basic structural components of all forms of life. Bustamante et al [26] in their defining work demonstrated how the optical tweezers could be utilized for force-extension studies. Here they attached titin filaments to latex beads and used the tweezers to study the stretching behavior. When we interpret the relation between the force $F(x)$ and the resultant change in length (x), we see that the force-extension profile has two distinct regimes, an initial Hookean regime and a non-Hookean regime. As mentioned above, the Hookean regime is attributed to entropic extension or the unraveling of the folded nature of the chain and the non-Hookean regime takes care of the stretching of the chain. The freely jointed chain model (FJC), developed by Kuhn, as described in equation 1.8 successfully describes the behavior of flexible chains force regimes of the order of 10-500 pN force. Stretched further, all standard polymer chain models exhibit diverging forces near the contour length. But in experiments, its not the case as the individual bonds are stretchable and hence they deviate from the predictions of standard polymer chain models. Hence to account for the force-extension regimes at higher forces, the force-extension relations discussed earlier have been modified taking into consideration the elasticity or the extensible nature of the Kuhn segments [27] as

$$x[F] = [coth(\frac{Fl}{k_BT}) - (\frac{k_BT}{Fl})](L + \frac{nF}{K}). \tag{1.12}$$

Here, n is the number of Kuhn segments $(n = L/l)$ and 'K' is the elastic modulus of the Kuhn segment. An alternative model that fits well with the experimental force-extension data of stiffer chains is the extended form of the wormlike chain model (e-WLC) [28]. The standard WLC model is modified by a force dependent

extension, $\frac{F}{\phi}$, where ϕ is the elastic modulus of the chain. Thus the modified extensible-WLC model is given by

$$F = \frac{k_B T}{l_p}[\frac{1}{4}(1 - \frac{l}{l_c} + \frac{F}{\phi})^{-2} + \frac{l}{l_c} - \frac{F}{\phi} - \frac{1}{4}]. \tag{1.13}$$

Here, l_p is the persistence length, l_c, the contour length and ϕ is the elastic modulus. Essentially these modifications ensure Hookean behavior is recovered at high forces (now enthalpic and not entropic in origin) and ensures that divergence is not a problem.

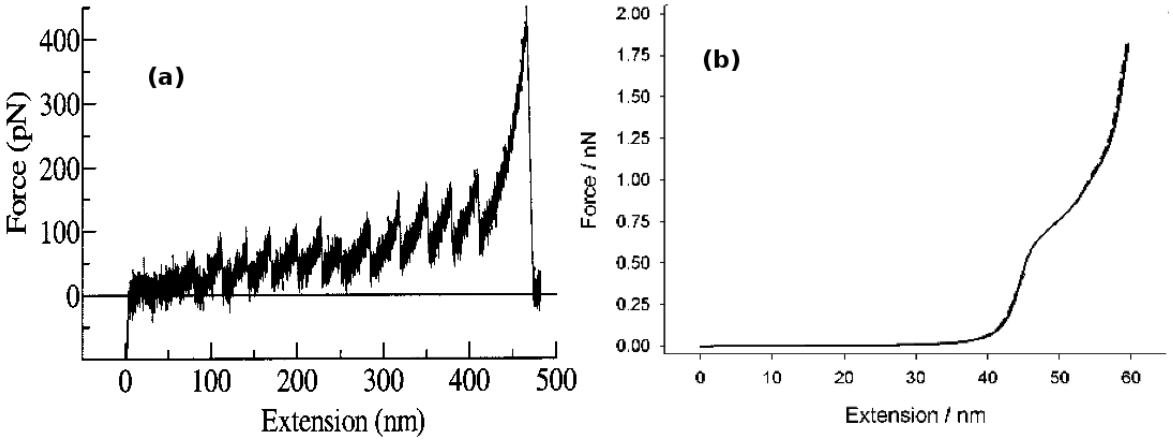

FIGURE 1.7: The nature of the force extension profile of (a) single chain protein, (b) for polysaccharides, as obtained from the AFM experiment.

Often interesting features have been experimentally observed when stretching proteins or polysaccharides. For example, the force-extension curve for intra molecularly bundled proteins exhibits a saw tooth pattern [29] while in certain other cases, found in particular polysaccharides, it exhibits some changes in the slope (fig 1.7). In the case of proteins, this is attributed to the folding-unfolding nature or unbundling of the different protein conformations and in the case of the polysaccharides, it is proposed that this arises from the conformational transitions taking place in the polysaccharide monomers. The sudden decay of the curve at high forces in either case is due to the snapping of the polymer chain or detatchment

from the tip of the AFM cantilever, while being stretched. It has been shown that experimental data of the form shown in figure 1.7(b) fits well with a modified extensible worm like chain (e-WLC) model, which takes into account a force dependent extension of the monomers comprising the polymer chain [30]. (See Appendix A for more details):

$$F = \frac{k_B T}{l_p}[\frac{1}{4}(1 - \frac{l}{l_e} + \frac{F}{\phi})^{-2} + \frac{l}{l_e} - \frac{F}{\phi} - \frac{1}{4}]. \tag{1.14}$$

Here, l_e is the force dependent contour length. Now, an important question is that how do these conformational transitions arise in the polymer chain? We know that many carbohydrates are composed of 6-membered pyranose rings as the monomer units, which can be linked to each other, typically by 1-3, 1-4 and 1-6 glycosidic linkages. The conformational transitions arise from the structure of these pyranose rings, which can exist in different stable conformers of varying lengths [31], while being stretched. Normally, pyranoses exists in their most stable $^4C_1$ chair conformation of length around 0.45nm. When being stretched, they can transform to different boats (B) of length range 0.49-0.52nm (of which the skew boat conformation $^3S_5$ of length 0.52 nm is the most stable one) and finally to the inverted chair conformation $^1C_4$, at a length of 0.55 nm. Corresponding to these different stable pyranose conformations attained during stretching, we see a region of different slope or inflection [32] in the force-extension profile of the macromolecule which shows that during this forced conformational change, little force is required to make an effective elongation to the contour length of the macromolecule. These inflections in the force-extension profile are also known by "clicks". Marszalek et al [33] showed that the number of "clicks" in the force-extension profile reflects the nature of the glycosidic linkages present in the carbohydrate chain and their spatial alignment. Thus the plateaus in the force-extension curve arise from different conformational transitions of the sugar rings in the polymer chain such as chair to boat ($^4C_1 \longrightarrow$B) or, when there are two clicks in the force-extension curve, ie chair to boat to inverted chair ($^4C_1 \longrightarrow$B$\longrightarrow$ $^1C_4$).

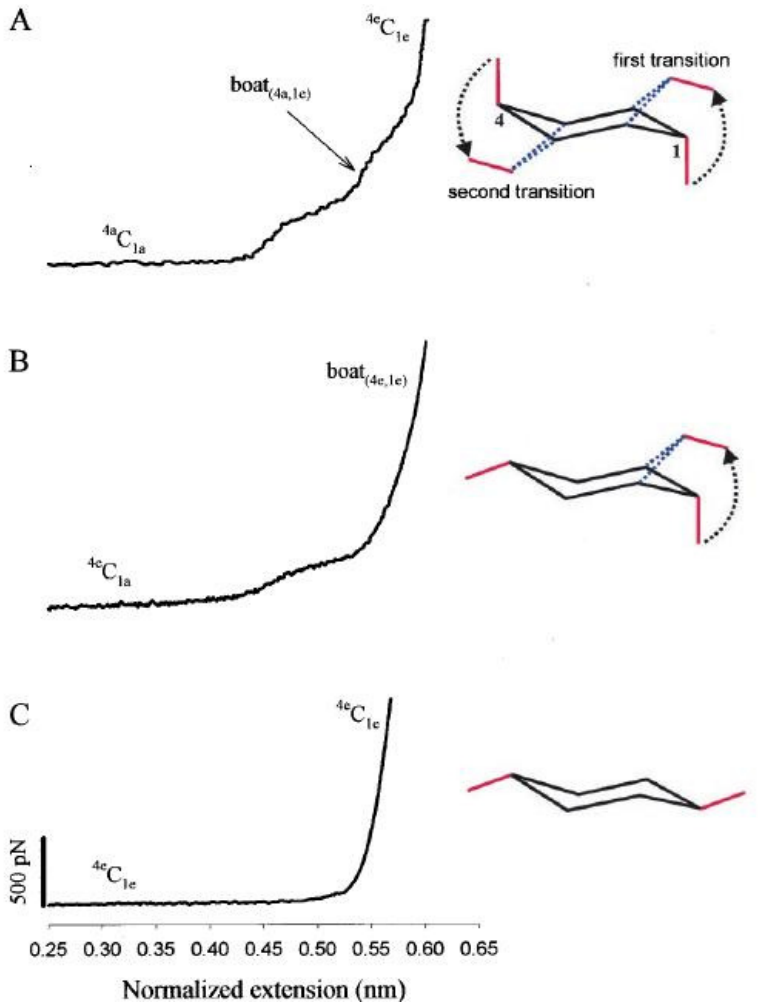

FIGURE 1.8: (A) Force-extension profile for a 1a-4a linked polysaccharide (like pectin), which exhibits two clicks; (B) This shows a single click nature for a molecule like amylose or β-D-galactose, which has an axial and an equatorial (1a-4e and 1e-4a respectively) linkage; (C) Force-extension profile for cellulose, which has a 1e-4e linkage and doesn't have any clicks or deviations.

Generally, 6 membered monomer units in the macromolecule possesing 1-4 glycosidic linkages being equatorially linked (1e-4e), such as cellulose, do not give any conformational transition. Here the sugar rings are already in their most elongated form. Those having 1-4 glycosidic linkages with one equatorial and one axial, such as amylose(1a-4e), β-D-galactose(1e-4a) or α-D-glucose(1a-4e) etc, can undergo one main force-induced conformational transition. The monomer units with (1a-4a) geometry such as pectin, can undergo two conformational transitions (fig 1.8). It has been shown in experiments that indeed cellulose doesn't show any "click",

when being stretched using the AFM instrument. In contrast β-D-galactose and amylose show "clicks" in the force-extension curves at around 640 pN and 280 pN respectively. If we look at the pectin molecule, the glycosidic linkages in pectin are of 1a-4a nature. Pectin when stretched, shows two elongations or "clicks" in its force extension curve, at forces around 350 pN and 900 pN (fig 1.9). It has been suggested that the plateau corresponding to the 350 pN transition in the force-extension curve is due the chair-to-boat ($^4C_1 \longrightarrow B$) transition in the pectin monosaccharide and that at 900 pN, the transition arises from the [34] boat-inverted chair ($B \longrightarrow {}^1C_4$) conformational transition.

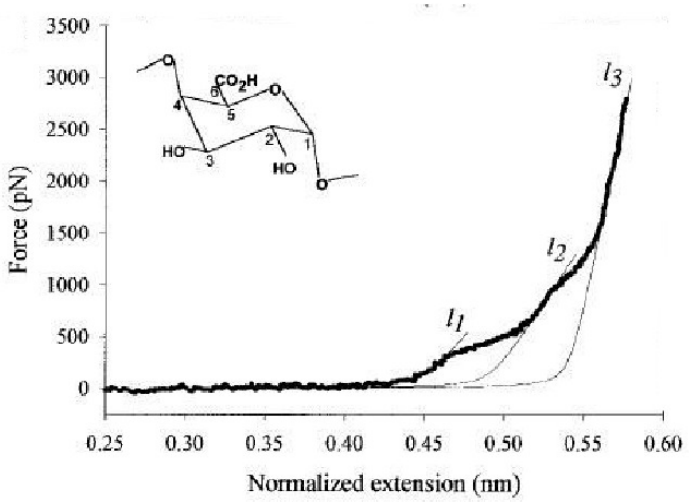

FIGURE 1.9: Force-Extension relation for single molecule pectin by *Marzalek et al.*

With pectins exhibiting this interesting force-extension behavior, its worth asking the question whether the DM has any influence on the stretching behavior of pectins? This question gains its importance owing to fact that pectins form an important entity in the structural composition of plant cell wall and its DM is known to be remodelled by enzymes in vivo, where it starts at around 80% methylesterified. Nevertheless the only experimental work to date has been carried out on a very low DM sample.

1.4 Theoretical Methods and Techniques.

In the recent past theoretical methods such as the *ab initio*/Density Functional Theory [35] and classical dynamics [36] techniques have been extensively used to study different molecular systems qualitatively and quantitatively (both living and non-living). They remain a good tool to help to understand experimental observations, even in those systems where experimental procedures are still a matter of great difficulty. Using these mathematical tools the dependence of specific properties on molecular parameters can be studied. One of the big disadvantages of these studies in the past has been with the scaling of these programs and the need for large times to run the computer codes. With the advancement in recent computational technologies and algorithms, now comparatively large systems (200 atoms) can be realistically examined.

With many well established experimental techniques available to characterize pectins, theoretical methods play an important role of verification of the interpretation of results. The *ab initio*/DFT studies can be utilized to look at the molecular details of the polymers at the finest level for comparison with spectroscopic analysis of these fine structures. With classical dynamics studies, properties of the chain such as its folding pattern and its mechanical stretching behavior can be examined.

1.4.1 *ab initio*/DFT Techniques.

The term *ab initio* comes from *latin* which means the beginning. These are basically calculations which are performed from first principles, solving the Schrödinger equation [37] for the system and no experimental data is used as such.

$$\widehat{H}\Psi = E\Psi \tag{1.15}$$

Here, Ψ is the wave function associated with the system. The probability (P) of finding the electron in an interval (a,b) is given by

$$P = \int_a^b |\Psi^2| dx. \tag{1.16}$$

When these methods are used, the quantum chemistry wave functions are defined using multi-electron basis functions for the system and are solved essentially by using a single electron Slater determinant method [38]. One of the most basic *ab initio* calculation methods is the Hartree Fock (HF) method [39]. HF dosen't take into consideration the electron-electron coloumb repulsion or the electronic correlation in the total energy and hence the energy values are approximate. Rather it takes into consideration an overall approximate value for the electron correlation. MØller-Plesset perturbation theory (MPn) [40], coupled cluster theory (CC) [41], Generalized Valence Bond (GVB) [42], Multi-Configurations Self Consistent Field (MCSCF) [43] etc, are all examples of post HF calculation methods where the correlation energy is explicitly incorporated.

Density functional theory (DFT) is yet another quantum mechanical approach adopted to study the ground states of multi-particle systems. This method was put forward by Hohenberg and Kohn in 1964 [44] and then further developed by Kohn and Sham (KS Scheme) [45] in the year 1965. Here, rather than analyzing the wave function of the system, importance is given to the electron densities $\rho(r)$ of the many-body system. The probability of finding an electron is given by

$$\rho(r) = N \int ... \int |\Psi(x_1, x_2, ..., ..., x_N)|^2 dx_1 dx_2 ..., ..., dx_N. \tag{1.17}$$

The pair density or the probability of finding a pair of electrons differentiated by their spins is given by

$$\rho(x_1, x_2) = N(N-1) \int ... \int |\Psi(x_1, x_2, ..., ..., x_N)|^2 dx_3, ...dx_N. \tag{1.18}$$

Two of the Hohenberg-Kohn theorems made way for the new method of DFT. These theorems states that

- The properties of a molecule in its electronic ground state are the direct consequence of its ground state electron density function $\rho(r)$.

- Any trial electron density function will give an energy greater than or equal to the true ground state energy. This is an analogue to the wavefunction variation theorem adopted in HF method for *ab initio* calculations.

In 1965, Kohn & Sham put forward a method to calculate the molecular properties from the electron densities, which is the present day DFT technique.

- The molecular energy is expressed as a sum of the electron kinetic energy(T), the nucleus-electron attraction potential energy(V_{Ne}) and electron-electron repulsion potential energy(V_{ee}). K-S energy for the ground state electronic configuration is given by

$$E = \langle T[\rho] \rangle + \langle V_{Ne}[\rho] \rangle + \langle V_{ee}[\rho] \rangle \,. \tag{1.19}$$

- An initial guess is used in order to define the electron density (ρ) in the KS equations and hence to get the KS orbitals. This in turn is used to refine the orbitals and the final KS orbitals are then used to calculate electron density and hence the molecular energy.

Thus, similar to the variation principle for HF method, Kohn and Sham formulated the KS equations and solving this set of equations results in the K-S energy and all other molecular properties.
To define the wave functions and electron densities in the case of these calculations, a set of mathematical functions are designed, which are a combination of atomic orbitals (the molecular orbitals) known as the basis sets. In a physical sense, basis sets define the electron distribution around an atom. The larger the basis set, the more accurately it approximates the orbitals by imposing fewer restrictions on the locations of electrons in the space. Different types of basis sets utilized in *ab initio*/DFT calculations are the Minimal basis set, Split-Valence basis set, Polarized basis set, Diffuse functions, etc.

The major applications of *ab initio*/DFT calculations are in assigning geometries or conformations, energies, thermodynamics (calculating heats of formation, reaction rates), spectroscopic frequencies etc. The study of molecular geometry is of the upmost importance as it plays a huge role in designing molecular materials, obtaining information on molecules which are short lived in the course of a reaction, modeling the docking of drugs to the active sites of enzymes and so on. Next to the molecular structure or geometry, another important aspect that can be addressed in *ab initio*/DFT studies is the calculation of the ground state energies of different species. The output of these simulations provides an energy, which describes the stability of the molecule related to the corresponding positions of electrons and nuclei, when bonded together. Thus these calculations furnish energies which are negative to the energy needed to dissociate the molecule to separate electrons and nuclei. Another important aspect of *ab initio*/DFT calculations is in the study of thermodynamics (heat or enthalpy of formation) of a system. This is an important quantification as the enthalpy of formation helps to calculate the heats of formation of different independent process and to determine whether a reaction is endothermic or exothermic. This also helps in the determination of the standard reaction enthalpy for the formation of a system. Reaction enthalpy helps in the study of the kinetics of a reaction and the calculation of reaction rates.
Yet another important application of theoretical simulations that has been used extensively in this work is the calculation of the normal mode frequencies of a molecular system so that the vibrational spectra of the molecule might be calculated. For a diatomic molecule, the normal-mode of vibrational frequency ($\widetilde{\nu}$) is given by

$$\widetilde{\nu} = \frac{1}{2\pi c}(\frac{k}{m_{eff}})^{1/2}, \tag{1.20}$$

where 'k' is the force constant of a vibrational mode and 'm_{eff}' is the reduced mass of the molecule. In the case of a diatomic molecule AB, its given by (m_A m_B)/(m_A + m_B). In the case of a system with a quadratic dependence of the energy on the

geometry of the molecule, the force constant is defined by the second derivative of energy (Hessien) to the molecular geometry. ie

$$E = \frac{k}{2}(q - q_{eq})^2. \tag{1.21}$$

$$\frac{\partial^2 E}{\partial q^2} = k. \tag{1.22}$$

Here, 'q' defines the geometry (any geometry changes such as bond lengths, angles etc) of the molecule and $'q'_{eq}$ is the equilibrium geometry. This is in the case of a diatomic molecule and in the case of a considerably larger molecule, higher powers of 'q' can be incorporated in the above energy equation and their coefficients, the anharmonicity corrections are added. Apart from the wave numbers or the positions in the IR spectrum, intensities of the IR bands can be estimated as well. The intensity of an unscaled IR band is obtained from the change in dipole moment (μ) of the system with respect to the geometry (q) of the molecule. Thus, the intensity I is given by

$$I = K(\frac{d\mu}{dq})^2. \tag{1.23}$$

Frequency calculations are important in:

- Calculating the zero point energy of a molecule from the frequencies and hence obtaining the accurate energies.
- Defining the stationary state of a system. ie the determining whether the optimised geometry is a minima, a transition state or a saddle point.
- Analyzing the IR properties of modeled molecules and in characterizing experimentally unidentified bands.

1.4.2 Classical Dynamics.

In classical dynamics techniques, the classical equations of motion for a system of 'N' molecules interacting via a potential 'U' are solved or in other words, the Newtonian mechanics on a system of 'N' atoms under a particular interaction potential 'U' [46] are calculated. The fundamental process in all classical dynamics methods like Molecular dynamics simulations and Brownian dynamics simulations is solving Newton's laws [47] for the system. In a conventional MD simulation,

$$F_i = -\nabla U_i = m_i a_i; (i = 1, 2, ...N), \tag{1.24}$$

where, m_i is the mass, a_i is the acceleration and F_i is the force acting on the i^{th} particle in the system. Knowing the position of the i^{th} particle r_i at time 't', the velocity and acceleration can be obtained by $v_i = \frac{dr_i}{dt}$ and $a_i = \frac{d^2r_i}{dt^2}$. Now, the time evolution of these interacting 'N' atoms can be easily studied by defining the potential 'U_i' for the i^{th} particle in the system and then integrating Newton's equations of motion.

In Langevin dynamics simulations, the basic approach is to consider the molecular motions as diffusion controlled. So Newton's equations become Langevin's equation [48]:

$$\frac{d^2r_i}{dt^2} = -\nabla U_i - \Gamma\frac{dr_i}{dt} + W_i. \tag{1.25}$$

Here, 'Γ' is the frictional constant that takes into account the viscous drag force in a medium and W_i is a stochastic force which accounts for the random collision of the solvent molecules that are not explicitly included. For purely viscous media Γ can be obtained from the Stokes equation, ie $\Gamma = 6\pi a_i^{stokes}\eta$, where a_i^{stokes} is the Stoke's radius of the diffusing particle 'i' and η is the viscosity of the solvent. The noise is related to the viscous drag by the fluctuation dissipation theorem [49]:

$$\langle W_i(t)\rangle = 0, \tag{1.26}$$

$$\langle W_i(t)W_j(t+\tau)\rangle = 6\Gamma k_B T\delta_{ij}\delta(\tau). \tag{1.27}$$

where $'\delta'_{ij}$ is the Kronecker delta function. Brownian dynamics simulation is a simplified version of Langevin dynamics and assumes an over damped state of the system. Hence there is no average acceleration during the simulation run. Thus the inertia vanishes from the above Langevin equation and it is modified as

$$0 = -\nabla U_i - \Gamma\frac{dr_i}{dt} + W_i. \tag{1.28}$$

Now the next important step is defining the potential U. This is crucial as the entire system of study depends how we define the interaction potentials.
We use the Lennard Jones potential to account for the excluded volume and solvent interactions [50, 51] given by:

$$U_{LJ} = 4\epsilon[(\frac{\sigma}{r_{ij}})^{12} - s(\frac{\sigma}{r_{ij}})^{6}], r_{ij} < r_{cutoff} \tag{1.29}$$

Here, $r_{ij} = |r_i - r_j|$ is the distance between two interacting particles i and j, ϵ is the minimum energy and σ is the length at which $U_{LJ} = 0$. This is known as the 12-6 LJ potential where $\frac{1}{(r_{ij})^{12}}$ is a strong repulsive term and $\frac{1}{(r_{ij})^{6}}$ is the attractive part. A coefficient s is introduced in the LJ potential to fine tune the attractive interactions. U_{LJ} is truncated at a certain value of r_{ij} given by r_{cutoff} ($r_{cutoff} = 2^{\frac{1}{6}}\sigma$). For a polymer chain, the stiffness can be accounted by including in this potential a bending contribution $U_{bending}$ given by

$$U_{bending} = \frac{1}{2}k_\theta(\theta_i - \theta_i^0)^2, \tag{1.30}$$

where k_θ defines the bending stiffness constant of the polymer chain and θ_i is the angle between the adjecent bonds. θ_i^0 gives the equilibrium bond angle. A torsional potential ($U_{torsion}$) can also be defined as:

$$U_{torsion} = \sum_{k} C_k cos^k(\phi_i). \tag{1.31}$$

where the coefficient C_k defines the potential and ϕ_i is the torsional angle between three adjacent bonds or four adjacent atoms. Normally the summation is truncated at a value of k=6.

The potential associated with the chemical bonding (U_{bond}) between the adjacent particles can be accounted by a Gaussian potential. ie

$$U_{bond} = \frac{k}{2} \sum_{i=1}^{N} (|\mathbf{r}_i - \mathbf{r}_{i-1}| - l)^2. \tag{1.32}$$

Where 'k' is the spring constant associated with spring joining the bonds in the bead spring model for the polymer chain and 'l' is the equilibrium bond length. The main drawback of this model is that this force law is linearly elastic, which is not the case in the real system. Nevertheless, this remains a good model to study the bond fluctuations in polymer simulations. To investigate its limitations, the FENE model (finitely extensible nonlinear elastic model) [52] for the polymer bonds has also been used. The FENE potential is given by:

$$U_{FENE} = \begin{cases} -0.5kR_0^2 \ln[1 - (\frac{r_{ij}}{R_0})^2] & , \ r_{ij} \leq R_0 \\ \infty & , \ r_{ij} > R_0 \end{cases} \tag{1.33}$$

This is an empirical relation, which at lower extensions ($\frac{r_{ij}}{R_0} < 0.2$) has the form of a simple harmonic potential but deviates at a higher chain extensions. It limits the spring extensibility to R_0.

We make use of velocity Verlet scheme [53] for the integration of Newton's equation of motion, where the position and velocity of each of the particles are updated by the equations

$$r(t + \Delta t) = r(t) + v(t)\Delta t + \frac{1}{2} a(t) \Delta t^2. \tag{1.34}$$

and

$$v(t + \frac{1}{2}\Delta t) = v(t) + \frac{1}{2}a(t)\Delta t. \tag{1.35}$$

Using these updated positions and velocities, the new updated force can be calculated.

Important applications of classical dynamics simulations like MDS and BDS are:

- In conformational analysis or geometry searching. Since thermal energy is introduced into the system, the system navigates through different degrees of freedom, thus exploring the potential energy surface in the immediate vicinity of its local minima.

- In studies of simulated annealing. Here in, the system is deliberately heated to a high temperature (more than 1000 K) to allow it to sample different alternate conformations. Then the system is cooled slowly, or annealed, by taking off the kinetic energy part from the simulation, to get back a stable conformation, as its internal strains are now relaxed.

- In determining the force-extension behavior of single long chain molecules and enabling the studies of different chain conformations.

Chapter 2

Characterization of Pectins.

In this chapter computational tools are used to aid the development of experimental methods for the characterization of pectin samples(in both powder form and when attached to different surfaces). In the first chapter, we have seen that pectin exists in different esterified forms, of which the highly methylesterified pectins are perhaps the most biologically relevant ones. The degree of methylesterification (DM) plays a relevant part in all the structural and functional roles of pectin and as such it is the first priority to distinguish samples on the basis of their DM when studying their functions. Several methods have been developed over the years for the DM determination of pectins. Acid-base titration [54] was the first method used while later alternative methods using alkaline hydrolysis to split the ester linkage and estimate the quantity of methanol released using HPLC analysis, headspace GC, or NMR have also been demonstrated [55–60]. These methods present certain disadvantages, mainly due to the complexity of the procedures involved and the sample destruction. Unlike many other natural polysaccharides, pectin has both charge and a UV chromophore (the carboxyl group) which makes Capillary Electrophoresis (CE) a practical analytical tool for its study (Appendix B) [61–64]. This technique provides a simple, rapid and non-destructive method for the quantitative detection and separation of pectins with different degree of

methyl esterification in aqueous solution, with the advantage of gaining information regarding the distribution of the substituents among chains. Another important method for the characterization of these species is the infrared spectroscopy technique. FT-IR (Fourier transform infra red spectroscopy) has great potential for the characterization of these molecules due to its non-destructive nature and more importantly owing to the spectroscopic aspect of the technique. This means that new groups and bonds formed through the interaction of pectin chains with different substrates might be detected.

2.1 Analysis of DM in Pectins: IR Simulation Studies and Experiment.

One objective of this study was to develop a more robust method for the DM determination in pectin using Attenuated Total Reflectance Fourier Transform Infrared Spectroscopy (ATR / FT-IR), which, unlike FT-IR does not require the sample to be heated, pressed or ground in order to collect the spectrum; (which takes time and can cause structural changes). Further more, it was hoped that developing such robust spectroscopy methods might prove of value in validating specific bond formation in proposed coupling reactions involving pectin macromolecules and substrates, which have been carried out in order to facilitate various biophysical experiments.

The application of FT-IR to the study of the structure and interactions of the major plant-cell-wall polysaccharide pectin has been reported for many decades [65–67]. Nevertheless, here we show that the generally reported methodology for one of its most commonly utilized applications, the measurement of the determination of the degree of methylesterification, requires careful interpretation and sample handling; including consideration of the moisture content and ionization state. We propose instead a different methodology based on the assessment of the magnitude of C-H stretches in the methyl groups relative to those in the backbone and demonstrate experimentally the advantage of this method. In addition we add

a theoretical dimension to our work performing full quantum-chemical calculations of monomeric-, dimeric-, and trimeric-pectic compounds, in various states of partial methylesterification. These extensive calculations confirm the identity of the proposed methyl-band and illustrate its scaling with DM. Furthermore, by carrying out calculations on unmethylesterified galacturonic oligomers of increasing DP and comparing these with experiment, insight into the origins of the broadness seen in the results of typical experiments on polymeric pectins compared with the calculations can be gleaned.

We calculated IR spectra using density functional theory (DFT) in order to gain confidence in our understanding of the experiments, validate previous proposed assignments, confirm the frequencies associated with our new proposed methodology, and assign uncharacterized frequencies. The question of the calculation of saccharide IR spectra has been considered previously for *single* sugar rings of differing substitutions [68, 69]. However of late, there have been extensive studies of the IR and Raman properties of molecular systems using theoretical simulations [70, 71] and the revolutionary development in recent years in available computational facilities and large clusters have now made the novel simulations of comparatively large systems (such as the three sugar rings examined herein) possible.

2.1.1 DFT Calculations.

Complete Hessien calculations were performed on a monomer, dimer and trimer of α-D-galacturonic acid and their methylesterified analogues using DFT calculations implemented in the Gamess U S package [72]. The computational facilities used were an IBM- Bluegene cluster at the University of Canterbury, NZ.

Bluegene / L nodes have two processors and two modes of execution are supported: co-processor (CO) and virtual node (VN) modes. In CO mode the first processor runs the program and the second processor handles I/O and communication on behalf of the first processor. In virtual node mode, both processors run the program. In CO mode all node memory is allocated to the first processor, while in the VN mode, the node memory is shared between the two processors. The single rack

system we utilized provides a maximum of 1024 processors with 512 MB of RAM each, or 2048 processors with 256MB of RAM each; with the calculations reported here exploiting the first option. Initially, the respective systems were geometry optimised using the B3LYP/6-31G* basis and subsequently a complete optimization and Hessien calculation was undertaken using the B3LYP/6-311++G** basis set [73–75], incorporating the diffuse functionals. All simulations were performed under conditions corresponding to 0 K and in vacuum condition. The convergence criteria in the energy minimization for energy differences between cycles of optimization were less than 1×10^{-6} Hartree with the gradient set to be less than 1×10^{-4}a.u [76]. The scaling and assignment of the vibrational modes were carried out using the Chemcraft programe [77].

2.1.2 Experiment.

The pectin samples used in the analysis included pectins of DM 0%, 30%, 60% and 90% purchased from Sigma (polygalacturonic acid salt, P9311, P9436, P9561); samples of DM 0%, 60% and 78% purchased from Fluka (polygalacturonic acid, apple and citrus pectins); pectins of DM 35%, 71% and 74% kindly provided by HenkSchols of the University of Wageningen (F, C and X6904); samples of DM 35%, 37%, 48% and 65% kindly provided provided by CP Kelco (randomly distributed methyl groups, LM 12, and blocky samples 0001-8-F, 0001-8-D and 0001-8-A respectively)and samples of DM 58%, 62%, 65% and 70% with randomly distributed methyl groups, alkali deesterified in previous studies [78, 79]. All the DMs were provided by the suppliers, with the exception of the homemade deesterified samples, which had their values determined by Capillary Electrophoresis, as described previously [80, 81].

Selected samples, P9311, P9436, P9561, Apple, 0001-8-D, 0001-8-F, were dissolved (0.5% w/w in MilliQ water) and the resultant solutions were acidified by addition of 0.2M HCl solution until a low pH (≈1-2) was reached. Then, they were dialyzed, freeze-dried and finally dried under reduced pressure, in a vacuum oven,

at T=30-40°C. Spectral acquisition was performed on solid samples using a Nicolet 5700 FT-IR spectrometer equipped with Omnic software (version 7.1) and a Smart Omni-Sampler (ATR cell with single reflectance germanium crystal). Each recorded spectrum is the average of 32 scans with a spectral resolution of 4 cm^{-1} from 400 to 4000 cm^{-1} on a dried sample, with a background spectrum recorded before each analysis. Three spectra were measured and each one was analyzed and fitted using Origin software (version 7.5) equipped with Peak-Fitting Module (PFM). Fitting involved first ensuring that the baseline was averaged to zero in the appropriate spectral range where no bands were detectable (1800-1900 cm^{-1}) followed by a normalization by the intensity of the largest peak. The spectral region for fitting was then selected (typically from 950-1900 cm^{-1}; or 950-1500 cm^{-1}). Typically an automated peak-picking algorithm found around twelve main peaks in such a region (peaks could be added or subtracted manually if required). Subsequently, Lorentzian peak shapes were selected and placed at the peak-picked positions whereupon their relative intensities and widths were modified by a fitting algorithm until a best fit was found, as determined by a minimized reduced chi squared. The peak areas of the Lorentzian components, centred at the peak-picked frequencies, are then output. The uncertainty in the repeat measurements performed on the same sample was less than 2%.

2.1.3 Result and Discussion.

Figure 2.1(a) shows the ATR / FT-IR spectra obtained from a number of representative pectin samples run simply as powders untreated, with the inset revealing that indeed their appears to be a promising relationship between the DM and the relative contribution of the 2 peaks centred at about 1750 cm^{-1} and 1630 cm^{-1}, due respectively to the infrared absorption of the carboxylic ester and to any protonated carboxylic acid groups, and to the carboxylic anion [82, 83] as reported previously. However, although a linear relationship between the DM and the ratio of the area underneath these peaks has been proposed [84] it was found here that

when the results from a larger set of samples was studied (with no specific sample treatment protocol) that the relationship did not seem to be strongly adhered to

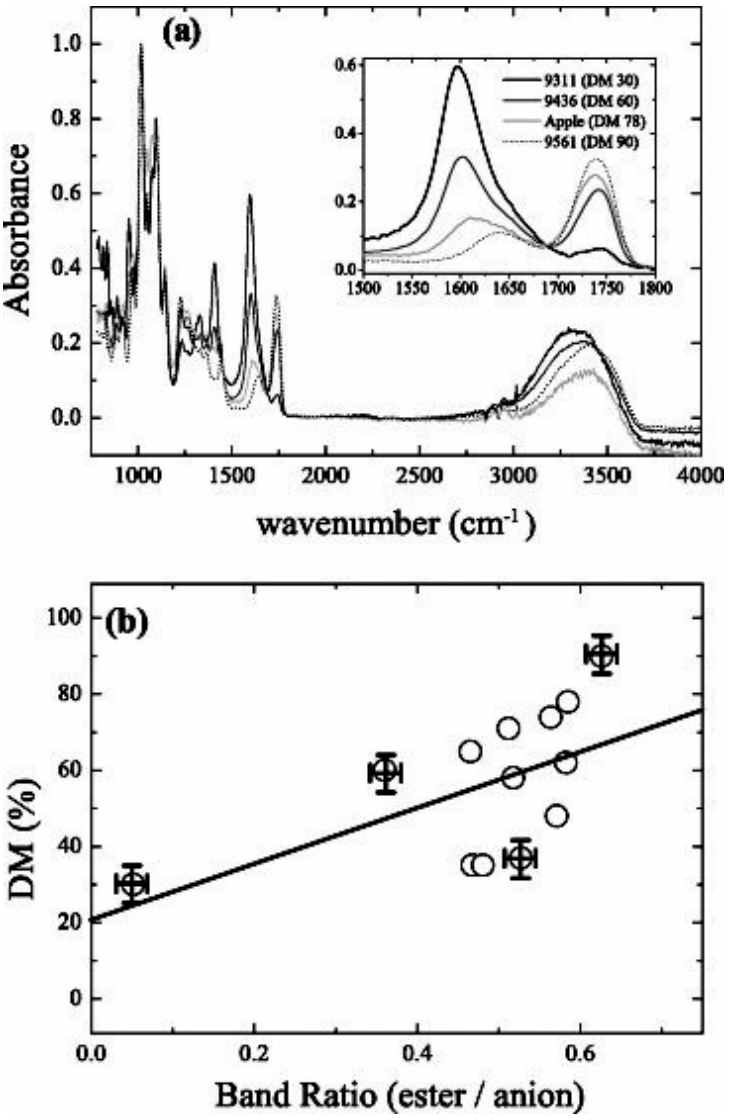

FIGURE 2.1: (a) ATR/FT-IR spectra of untreated pectins with different DM. Insert - the area of these 2 peaks conventionally used to determine the DM. (b) Regression analysis for crude samples using these carboxyl bands at 1630 cm^{-1} and 1750 cm^{-1}.

(figure 2.1(b)). One explanation could possibly be interference from other carboxylates and carbonyl ester groups, such as those from cell wall phenolics, or more concerningly from water (the in-plane deformation band of water δ (H_2O) occurs at 1645 cm^{-1}). As such, samples were extensively dried and while this

improved the correlation between the ratio of the suggested band intensities and DM somewhat, there were still evident problems with obtaining accurate DM determinations. Further rationalization of the results lead us to consider that the intensity of the symmetric (1450 cm^{-1}) and asymmetric (1630 cm^{-1}) carboxylate stretch modes are in fact be very sensitive to the ratio of protonated to ionic carboxylate groups.

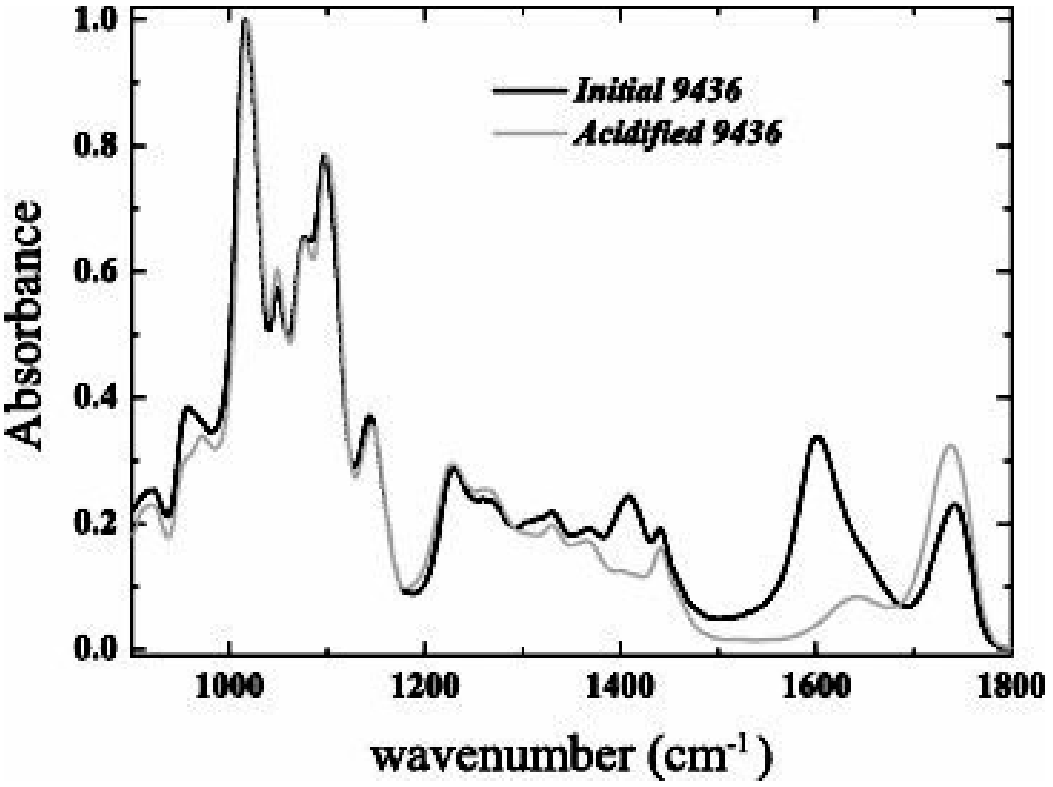

FIGURE 2.2: Effect of drying of pectins from acidified solutions on the measured IR spectrum.

By carefully precipitating and drying the pectins from acidified solutions, the carboxylic anion band largely disappears, and the band at 1750 cm^{-1} can be seen to increase owing to the presence of the newly formed carboxylic acid groups (figure 2.2). This may seem an unusual procedure in deference to attempting to manipulate the sample conditions in the opposite direction in order to obtain all charged

and no protonated groups, as now the protonated and methylesterifed species essentially absorb at the same frequency and the originally proposed method is no longer applicable. However, another alternative now becomes possible, which

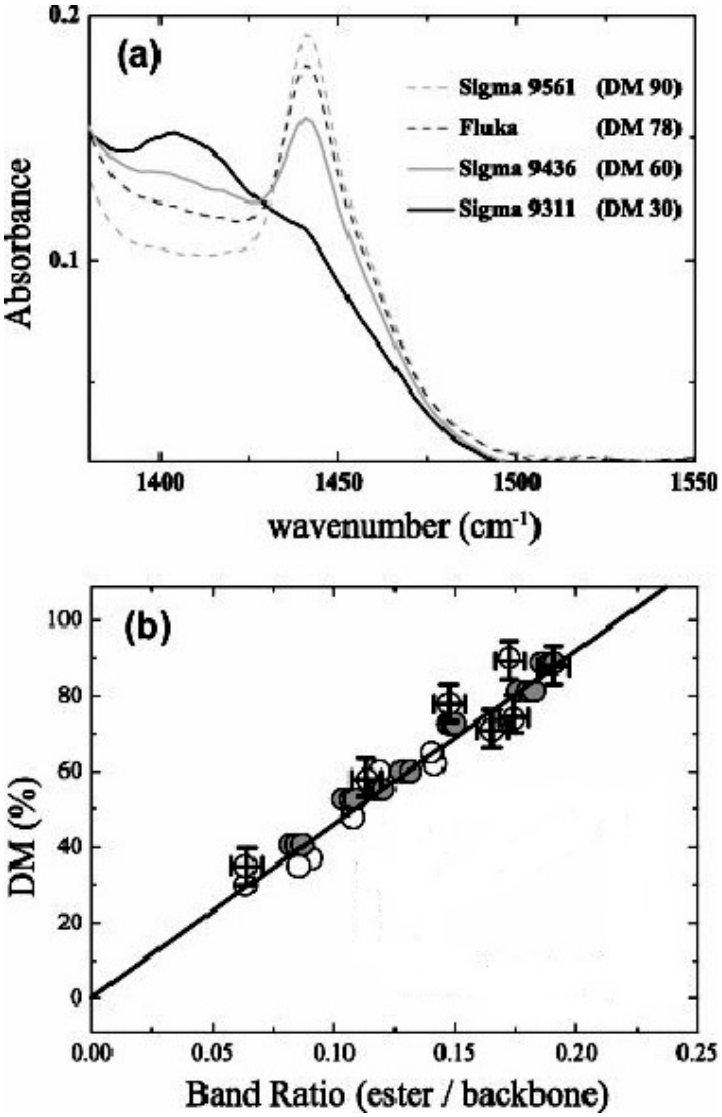

FIGURE 2.3: (a) ATR/FT-IR spectra of pectins with different DM in the region of 1380-1500 cm^{-1} (b) Regression analysis for acidified samples using CH_3 stretching and backbone vibration bands. The uncertainties in the DM are reasonable estimates based on the techniques used and those in the ratio are from 3 repeat experiments. The results in grey are obtained by reanalysing the results of a previous study.

completely eliminates problems introduced by the presence of the signals from hydration water.

Previous IR studies on pectins have suggested that the esterified CH_3 group presents bands in the 1350-1450 cm^{-1} range, one at 1380 cm^{-1} corresponding to the symmetric stretching of CH_3 and one at 1440 cm^{-1} corresponding to the asymmetric stretching of CH_3.

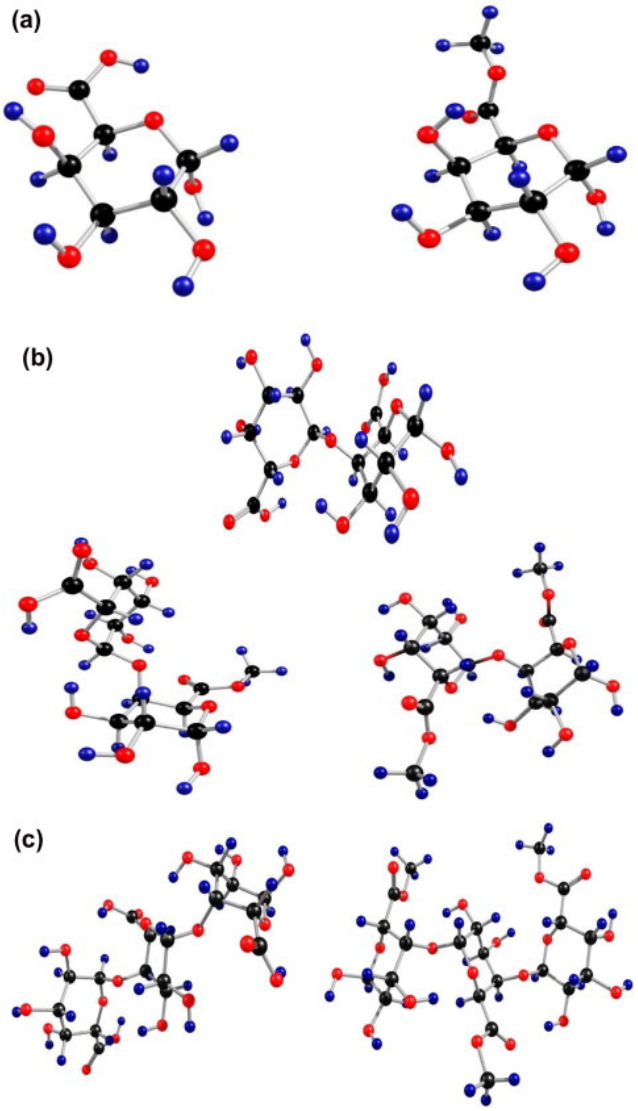

FIGURE 2.4: (a) Pure monomer of α-D-galacturonic acid and its methyl ester analogue. (b) Galacturonic acid dimers of 0%, 50% and 100% DM depending on the number of the esterified groups in the molecule. (c) Trimers of α-D-galacturonic acid and its completely methyl-esteried analogue.

Therefore, rather than attempting to quantify the bands corresponding to (1) the absorption of the carboxylic ester and the carboxylic acid groups; and (2) to

the absorption of the carboxylic anion, we investigated whether the DM could be calculated simply by using the intensity of the asymmetric stretching of CH_3 [85] at 1440 cm^{-1} relative to a backbone vibration at 1010 cm^{-1}.

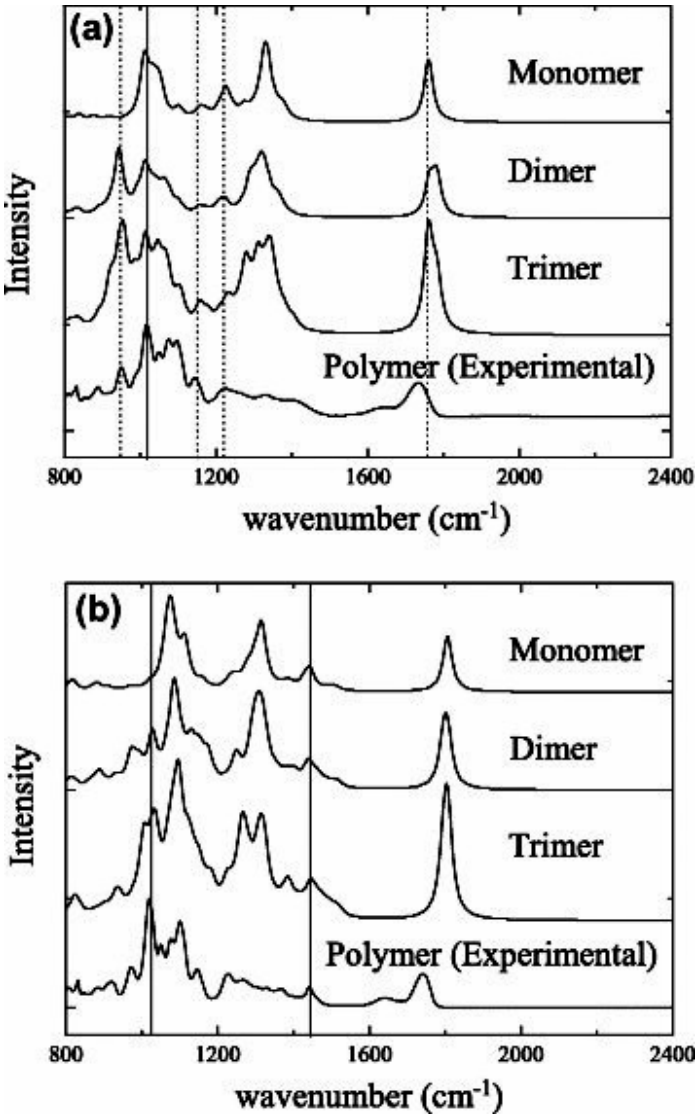

FIGURE 2.5: (a) Comparison of experimental IR spectra of a 0% DM pectin with calculations of monomer, dimer and trimer of α-D-galacturonic acid. (b) Comparison of experimental IR spectra of a 90% DM pectin with calculations of 100% methylesterified monomer, dimer and trimer of α-D-galacturonic acid. The dotted line shows the position of the CH_3 stretch.

While pectin samples can contain small amounts of neutral sugars (for example in rhamnogalacturonan), the backbone vibration band is largely due to the homogalacturonan and a reasonably good correlation with the classically measured

DM could still be expected. In fact to perform the quantification of the proposed band, acidification of the sample (to limit the interference of the

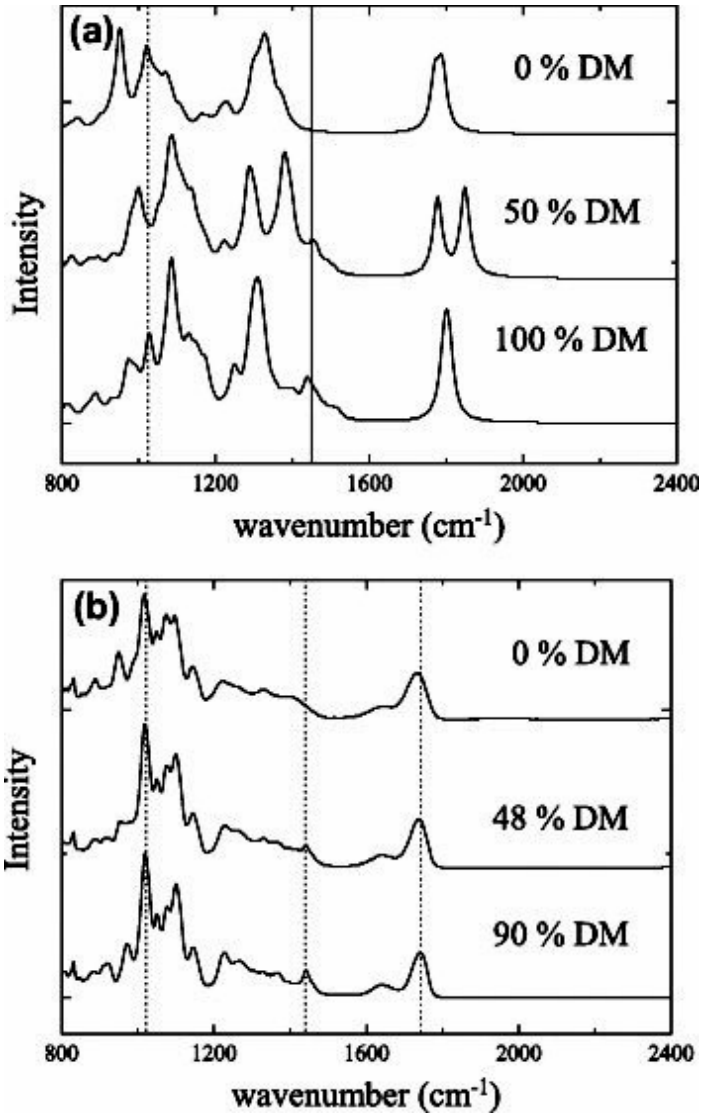

FIGURE 2.6: a) Simulated IR spectra for galacturonic acid dimmers of 0%, 50% and 100% DM; and (b) Experimental IR spectra of pectins of different (comparable) DMs. The central dotted line shows the position of the CH_3 stretch (compared with the solid line in (a)).

symmetrical stretching of the carboxylate group (figure2.2)) and robust peak fitting were both required. Under these conditions a relationship was established between the degree of methylesterification of the pectin samples and the ratio of

the peak areas from bands at 1440 cm^{-1} and at 1010 cm^{-1}, which yielded a considerably improved correlation coefficient (figure 2.3).

For further confidence in the proposed methodology, we carried out full Density Functional Theory (DFT) quantum chemical calculations and obtained the spectra of the monomer, dimer and trimer of α-D-galacturonic acid and their methyl-esterified analogues (structures shown in figure 2.4). First we compared the experimental IR spectra of 0% and 90% DM pectin samples with simulated spectra of the monomer, dimer and the trimer of α-D-galacturonic acid and their fully methylesterified counterparts (Fig 2.5) (A few of the important bands are specifically highlighted in the spectra and they are characterized and assigned in Table 2.1). Fig. 2.5(a) shows the comparison of the 0% pectin with simulated spectra of oligomers of 0% DM and different DPs up to the point at which the calculation become uncomfortably expensive. There is in general good accordance, in particular to the carbonyl group vibration frequency (1600 - 1800 cm^{-1}) and the backbone vibration frequencies (900 - 1200 cm^{-1}) respectively. Since the 0% DM pectin samples dont have any methylester groups, we don't find the corresponding vibrations, occurring at around 1430 - 1490 cm^{-1} as expected; and clearly the water interference in the region 1600 - 1650 cm^{-1} in the experimental result is not expected to be present in the calculated spectra. Five bands are specifically highlighted and there assignments are shown in Table 2.1. It should be emphasized that the calculations performed on systems of this size are highly complex, and the fact that the raw calculated frequencies only deviate by some 40 wave numbers or less from those found experimentally constitutes good agreement. The main reasons for the observed differences result from approximations inherent in the calculations including: they are essentially carried out in vacuum at 0 K, and assume a strictly harmonic form for the relevant potentials. In light of this it is routine to scale the frequency axis by a small amount so as to align major bands with those experimentally observed. The co-alignment of multiple calculated and experimental bands via the same scaling, as seen here, constitutes good agreement. Fig 2.5(b) shows the comparison between the experimental 90% DM data and simulations again of varying DP, but here with all residues methylesterified.

TABLE 2.1: The assignment of relevant IR peaks and the comparison of the experimentally measured frequencies with those found directly from DFT calculation expounded herein (PGA, polygalacturonic acid).

Raw Calculated (cm^{-1})	Experimental (cm^{-1})	Assignment
1786	1750	$\nu(C=O), PGA-ester$
1798	1750	$\nu(C=O), PGA$
–	1630	$\nu(C=O), PGA-ion$
–	1640	$\nu(HOH)$
1497	1440	$\nu_{as}(CH_3), ester(CH_3)$
1442	1380	$\nu_{s}(CH_3), ester(CH_3)$
1215	1275	$\nu(CO), \delta(OCH)$
1155	1107	$\nu(CO), \nu(CC), Ring$
1051-1062(acid & ester)	1010	$\nu(CO), \nu(CC), \delta(OCH), Ring$
983-1013(acid & ester)	934	$\nu(CO)glycosidic$

Now, in addition to observing all the pectin peaks in the IR spectra as described for the 0% DM oligomers, we do find the -CH_3 group vibrations at around 1450 cm^{-1}.

In order to look at the difference in the -CH_3 group vibrations of different DM pectins, we took a galacturonic acid dimer for simulation, as this is the simplest molecule in which it is possible to study three different DMs by simulation. By varying the number of methyl groups along the molecule, we can easily generate a 0%, 50% and 100% sample for simulation. Comparing the calculated spectra of these 0%, 50% and 100% dimers (Fig 2.6 (a)), we could clearly see an increase in the peak areas of the -CH_3 group vibrations, as the DM increases, validating the philosophy behind our proposed methodology. The simulation results are also compared with the experimentally determined spectra of different DM pectins (0%, 48% and 90%), which are comparable with the ones obtained through the DFT simulations (Fig 2.6(b)). Again we found that the simulations and the experimental data agree well. In all the raw simulated spectra, when compared with those obtained from experiments on the polymeric samples, we could see that the bands owing to the backbone -CH vibrations (1300 - 1350 cm^{-1}) are much sharper in simulation than in experiment. The spectra shown in 2.5 and 2.6 have had Lorentzian broadening applied to them in order to aid the comparison with experiment. To investigate the origins of the broadness, which we hypothesize is due

to the increased complexity in the polymeric backbone, we took the experimental spectrum of a pure pectin monomer and dimer (0% DE) and compared them with the polymeric result (Fig 2.7).It was indeed observed that the experimental spectra of a sole monomer showed sharp regions in 1300 - 1350 cm^{-1} accounting for the -CH vibrations and the experimental and the theoretical spectra (shown in figure 2.5a) were in reasonable accordance with each other. Already some broadening of the peaks is seen

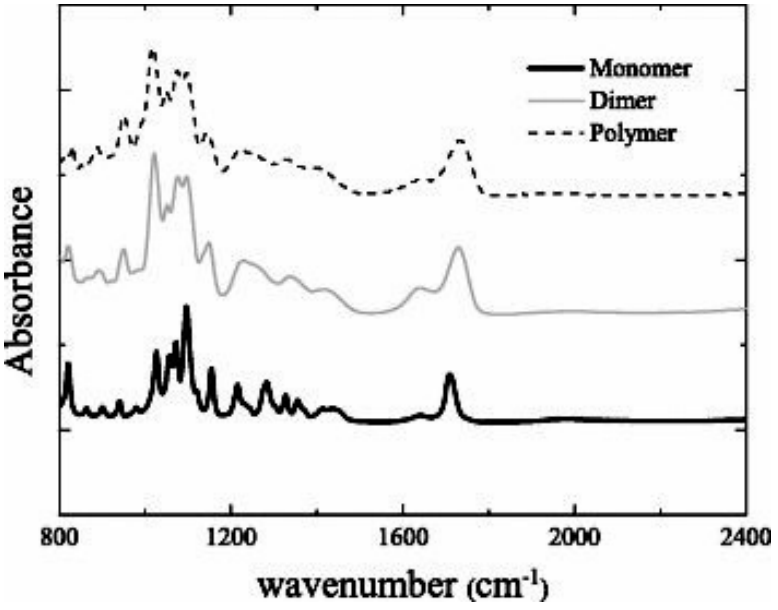

FIGURE 2.7: Comparison of IR spectra recorded for α-D-galacturonic acid monomer, dimmer and polymer, illustrating the increasing broadening owing to degree of polymerization.

experimentally as the dimer spectra is measured, suggesting that the broadening of the experimental pectic spectra in this regime (1300 - 1350 cm^{-1}) is a facet of the increasing degree of polymerization.

2.2 Pectins on Surfaces.

In the earlier section, we have seen how pectin molecules could be characterized using an ATR-FTIR technique [86]. One of the important uses of this method

of characterization could be in determining the chemical attachment of pectins on different surfaces. While many indirect methods can asses the "sticking" of polymers to substrates, we aimed to investigate whether the spectroscopy aspect of FTIR could allow the detection of newly formed bonds with surfaces. Functionalization of different surfaces with pectin molecules could be utilized as the starting step for AFM and OT studies, which are a few of the modern day techniques available to obtain high resolution structural and mechanical properties of these functionally important molecules [87–89].

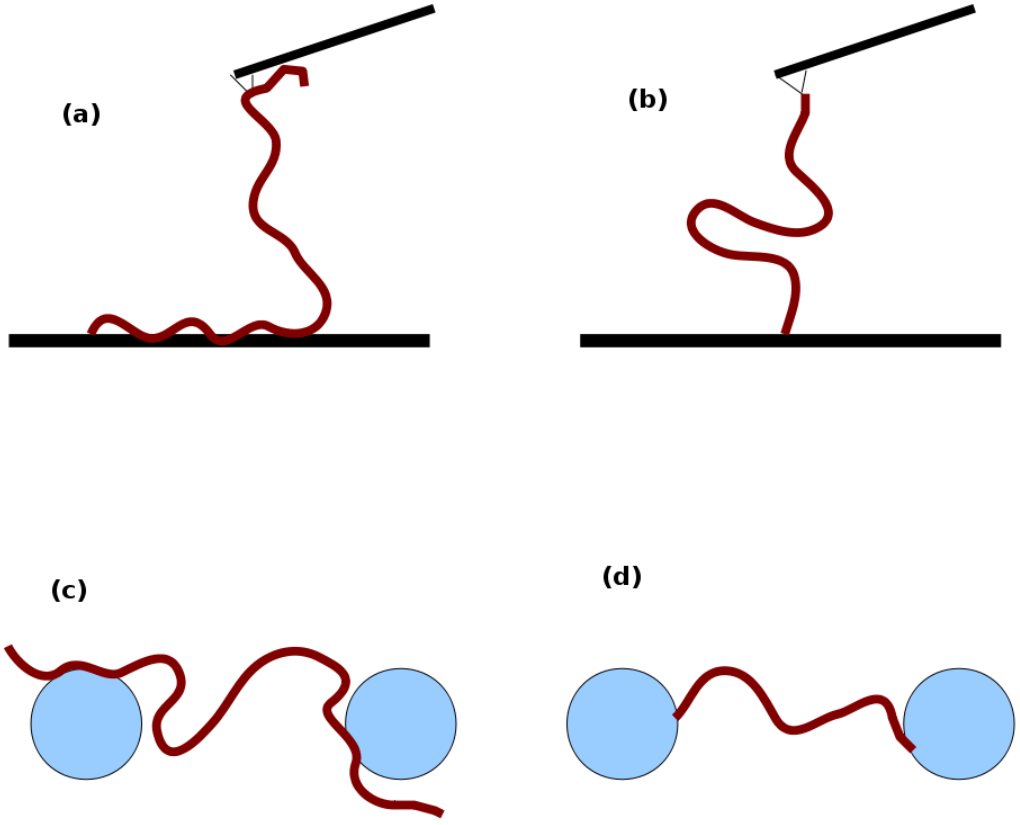

FIGURE 2.8: (a) Polymers physisorbed on a plain slide being pulled. (b) polymer chain functionalised on an AFM tip and tethered on a plain slide, being pulled. (c) Polymer chain physisorbed on spherical surfaces and (d) Polymer chain being tethered on two beads as in Optical Tweezer experiments.

Conventional AFM stretching experiments are hugely dependent on picking up physisorbed polymers by chance with the AFM cantilever, from different surfaces and pulling them. Physisorption of these substances on surfaces tends to be weak,

owing to the intrinsic lack of strength of physical forces such as the van der Waals forces [90]. A few of the drawbacks in carrying out these experiments are

- The number of scans needed to perform to detect an effective stretch.
- The snapping or detachment of the polymer chain off the AFM cantilever.
- Difficulty in performing multiple scanning and stretch reversal studies.
- Unknown length scales of these polymer chains and lack of understanding and control over these experiments.

Now, if we could attach known points of these polymer molecules onto surfaces like glass, gold or mica and if the AFM tip could similarly be functionalised, this would reduce some of the difficulties during our experiment, which are highlighted above. Schematic diagram 2.8 gives us an idea of pulling the physically attached (a & b) polymer chains from different surfaces and pulling of polymer molecules chemically attached (c & d) on these kinds of surfaces. Previous studies have attempted such specific methods for nucleic acids and proteins like titin, where an abundance of the functional groups with well differentiated chemistry makes such an endeavor comparatively easy. However in sugars and carbohydrates, owing to their smaller length scales and the number of similar active groups, it becomes hard to carry out these experiments. Having many functional -OH groups in the system, it becomes a complex procedure to perform chemistry at preferred positions. Herein, we have made an effort to attach one end of a pectin molecule to different beads and to characterize specific bonds formed using spectroscopy supported by quantum chemical calculations as a starting step in understanding the interactions of these class of polymers on different surfaces and AFM tips. Characterisation is carried out on the basis of their IR properties.

2.2.1 Experiments and Techniques

Apple pectin, DM 78, was purchased from Sigma, LM 12 and 0001-8-D pectins were provided by CP Kelco. All chemicals were purchased from Acros Organics

except $SnCl_2$ from Scharlau, Boc-Cys(trt)-OH from Bachem and Hydrazine monohydrate from Alfa Aesar. Polystyrene beads (100 and 500 nm, 2.5% w/v solids aqueous suspension, Polysciences, USA) and amino-terminated beads (100, 200 and 500 nm, 2.5% w/v solids aqueous suspension, Polysciences, Polysciences, USA) used in experiments were removed from their respective solutions by centrifugation and subsequently dried. Immobilization of pectins on to these aminated beads was carried out via Reductive Amination and Thiazolidine Formation method. These methods have been successfully carried out for pectin oligomers[91],

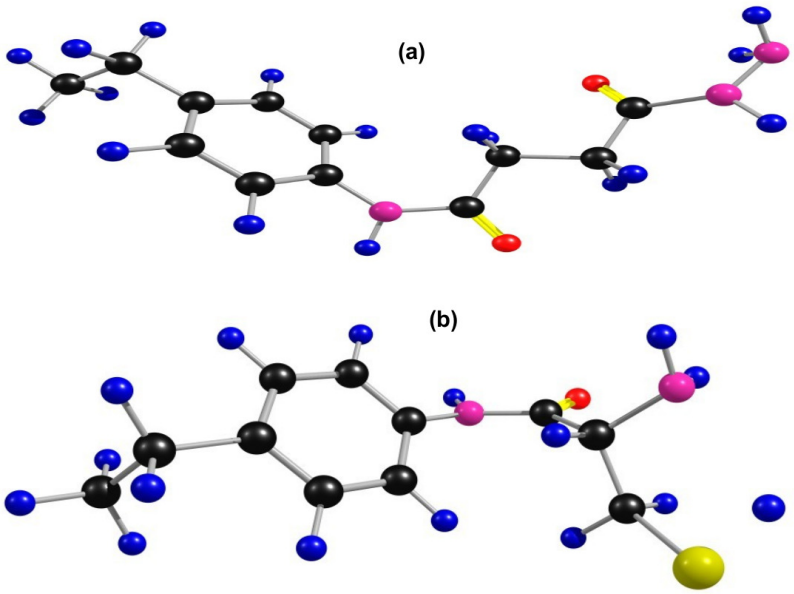

FIGURE 2.9: (a) Minimal model for the intermediate formed during Reductive amination method. (b) Minimal model for the intermediate formed during Thizolidine amination method.

but to our knowledge, this work with polymers is novel (see Appendix C) [92]. Spectral acquisition was performed on solid samples using a Nicolet 5700 FT-IR spectrometer equipped with Omnic software (version 7.1) and a Smart Omni-Sampler (ATR cell with single reflectance germanium crystal). Each recorded spectrum is the average of 32 scans with a spectral resolution of 4 cm^{-1} from 400 to 4000 cm^{-1}, with a background spectrum recorded before each analysis.

DFT calculations were again investigated for their usefulness in the prediction of

changes in the IR spectra resulting from the coupling of pectins to the bead. Molecular models for the pectin moiety immobilized on polystyrene beads (by Reductive Amination as well as Thizolidine formation) were built using the Ghemical [93] software and they were subsequently minimized using the semi empirical AM1 basis set. To make the system computationally less expensive, a pectinic acid dimer

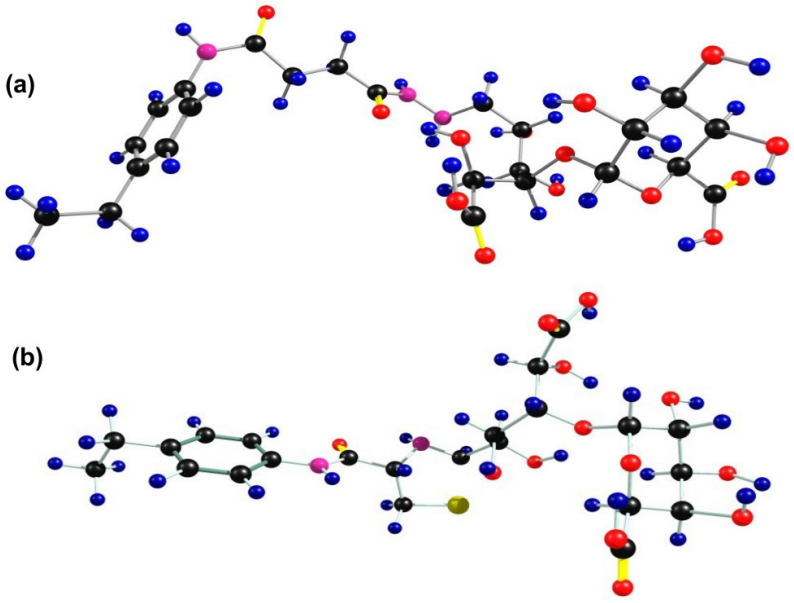

FIGURE 2.10: (a) Pectin dimer immobilized on 'polystyrene beads' by Reductive Amination method. (b) Pectin dimer immobilized on 'polystyrene beads' by Thizolidine formation method.

molecule attached to only one unit of polystyrene is considered. The free valencies of the single monomer polystyrene was satisfied by H atoms. Figures 2.9 and 2.10 are the minimal models used for intermediate state formation and pectin attachment to the polystyrene beads. From then on, complete Hessien calculations were performed on these molecules using DFT calculations as described in more detail in the introduction. Calculations of the proposed reaction intermediate were also carried out (Figure 2.9) in order to monitor the reaction. These structures were first geometry optimised using the B3LYP/6-31G* basis and subsequently a complete optimization and Hessien calculation was undertaken using

the B3LYP/6-311++G** basis set, incorporating the diffuse functionals. All simulations were performed at 0 K and in vacuum condition. The convergence criteria in the energy minimization for energy differences between cycles of optimization were less than 1 $\times 10^{-6}$ Hartree with the gradient set to be less than 1 $\times 10^{-4}$a.u. The computational facilities used were an IBM-Bluegene cluster at the University of Canterbury, NZ. The scaling and assignment of the vibrational modes were carried out using the Chemcraft programe.

2.2.2 Results and Discussions

The previous section suggests that IR spectroscopy coupled with DFT Hessien studies should be a relaible tool in biopolymer characterization studies. Figures 2.11 and 2.12 show the comparison between the IR spectra of the aminated beads after intermediate steps directly prior to pectin attatchment and the beads with putative pectin coupling via the (a) Reductive amination method and (b) Thizolidine formation method. The most interesting feature in either of the reactions is whether the formation of a new -C-N bond after the pectin immobilization could be monitored in the IR spectrum. Green dotted lines indicate spectral features clearly associated with pectin while the blue ones originate from the bead intermediate. While the detailed calculated spectra are shown later, the black lines in the figure show two regions where the formation of the new bond is predicted to be manifest in the IR spectra. In the case of the reductive amination technique, this new pectin-substrate bond is unique as it is the only -C-N bond in the system. Though there is a pre-existing -C-N bond in the thizolidine coupled system prior to pectin attachment, we may still be able to monitor the formation of the new bond after the immobilization, due to the broadening of the pre-existing -C-N band. As can be seen in figures 2.11 and 2.12 there does appear to be extra intensity in the spectra of the coupled system at the positions indicated by the calculations. While the bands at 1120-1150 cm^{-1} show the typical -C-N vibrations, the ones around 1580-1670 cm^{-1} reflect the delocalised character of the -C$\simeq$N; being coupled with -C=O in the reductive amination case and strained in the five membered ring

formed in the thizolidine case. Thus this work provides evidence that not only is pectin attached to these beads but it is via the formation of a specific recognizable bond and that in the course of its chemistry of formation pectin is coupled at its reducing end with the intended -CN bond.

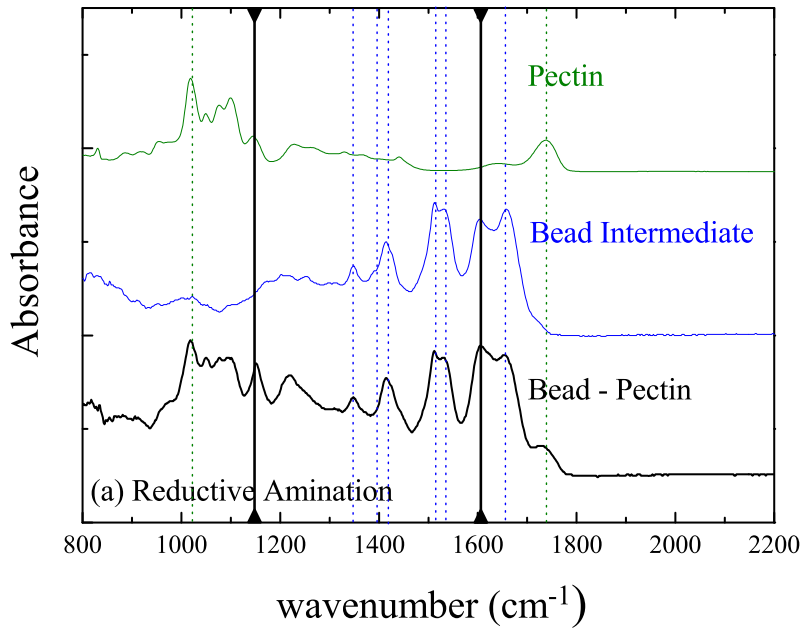

FIGURE 2.11: Experimental IR spectra for pectins immobilized on polystyrene beads by Reductive amination method and its comparison with aminated intermediate polystyrene beads.

Other prominent bond vibrations in the IR spectra are those corresponding to a set of carbonyl (-C=O) groups in the system at around 1650-1750 cm^{-1} and the benzene ring vibrations at around 1550 cm^{-1}. In pectin immobilized beads, the pectin backbone vibrations can be seen at 1000-1100 cm^{-1}. The addition of a new shoulder peak in the IR spectra, owing to the -C=O stretching of the pectin entity at around 1715 cm^{-1} also adds up as a support to the attachment of the pectin molecules on the bead.

Similar behavior is shown by the simulated IR frequencies. Figures 2.13 and 2.14

depict the comparison of the simulated IR spectra of chemically attached pectins on the polystyrene beads via the two methods discussed above with aminated beads calculated as described earlier. Again green dotted lines indicate spectral features clearly associated with pectin while the blue ones originate from the bead intermediate. Furthermore, similar to the experiment, the simulated spectra show distinct peaks associated with the newly formed -C-N bond after pectin immobilization as described above.

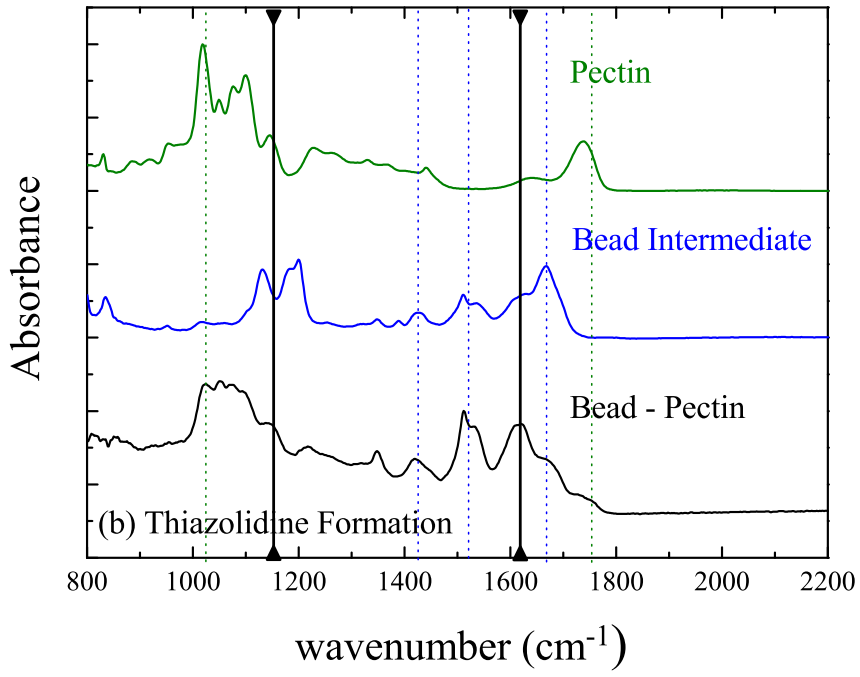

FIGURE 2.12: Experimental IR spectra for pectins immobolised on polystyrene beads by Thizolidine formation method and its comparison with aminated intermediate polystyrene beads.

Thus it can be seen that the simulation compliments the experimental findings well. Tables 2.2 and 2.3 characterise and compare the different coupled states on the basis of the peak assignment. On a closer look of this table, we can very well judge the aminated bead from the pectin attached bead by the distinction of the pectin backbone peak regions and the newly formed C-N bond region, which stands out uniquely.

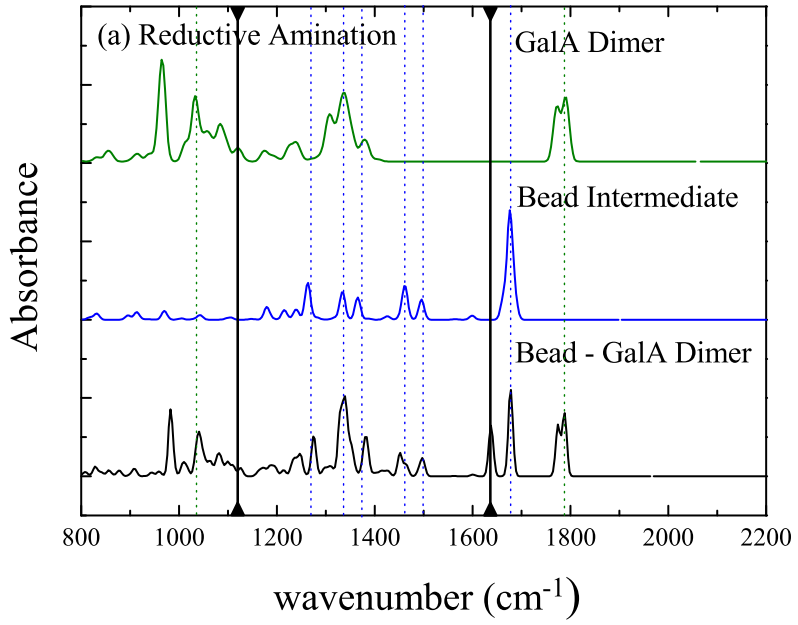

FIGURE 2.13: Simulated IR spectra for pectins immobilized on polystyrene beads by Reductive amination method and its comparison with aminated polystyrene beads.

Unlike the experimental bond vibrations, some peaks in the simulated spectra are quite sharp. This is due to the crystalline nature or the single unit nature of the substrate in the simulations. Here also, the calculations are complex and the raw calculated frequencies are bound to deviate the experimental ones by 40-50 wave numbers and any value in this range can be considered as good. So it is routine to scale the frequency axis by a small amount so as to align major bands with those experimentally observed. As in the previous section, co-alignment of multiple calculated and experimental bands via the same scaling, as seen here, constitutes good agreement. In the real experiments, bigger volume of the samples are analyzed and in turn, the vibration frequencies among single units could overlap making broader peaks in the spectrum. It is also to be considered that the simulation are carried out at vaccum and 0 K unlike the experiments that are performed in the room temperature.

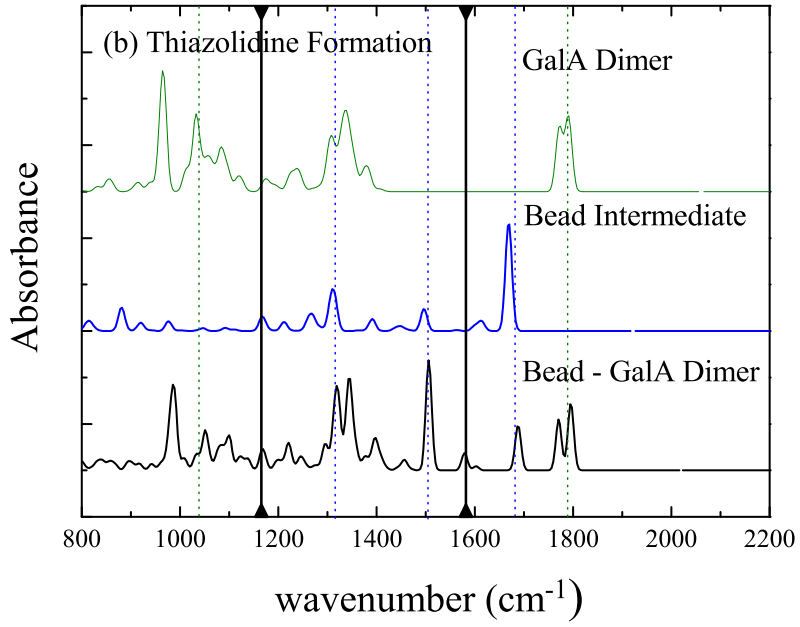

FIGURE 2.14: Simulated IR spectra for pectins immobilized on polystyrene beads by Thizolidine formation method and its comparison with aminated polystyrene beads.

2.3 Summary.

In first part of the chapter (section 2.1), we see that the results of DFT calculations compliment the main features of the experimental pectin spectra well, with a clear evolution towards the polymer result as the degree of polymerization of the saccharide molecule in the calculation was increased from 1 to 2, and ultimately 3. The FT-IR method described in this study for the determination of the degree of methylesterification of pectins is non-invasive and is significantly more accurate than other reported IR methods; although the expected uncertainty in the measured DM value still stands at around 7%. In order to get the best results from this method the samples should be treated (precipitated from acidic solution and dried) as described. The ratio of the intensities of the asymmetric vibrations of the -CH_3 group and backbone bands can then be used to determine the DM

TABLE 2.2: The assignment of relevant Reductive Amination IR peaks and the comparison of the experimentally measured frequencies with those found directly from DFT calculation expounded herein (PIB :- Pectin immobilized on beads, AB :- Aminated intermediate beads, PGA :- Polygalacturonic acid).

PIB Cal: (cm^{-1})	PIB Exp: (cm^{-1})	AB Cal: (cm^{-1})	AB Exp: (cm^{-1})	Assignment
1795	1720	-	-	$\nu(C=O), PGA$
1720	1660	1735	1660	$\nu(C=O), bead$
1650	1520	1635	1520	ν Benzene ring
1230	1210	-	-	$\nu(CO), \nu(CC)$, Pectin Ring
1125	1145	-	-	$\nu(CN)$ new covalent coupling
1610	1660	-	-	$\nu(C \simeq N)$ new covalent coupling
1040	1010	-	-	$\nu(CO), \nu(CC), \delta(OCH), Ring$

TABLE 2.3: The assignment of relevant Thizolidine immobilization IR peaks and the comparison of the experimentally measured frequencies with those found directly from DFT calculation expounded herein (PIB :- Pectin immobilized on beads, AB :- Aminated intermediate beads, PGA :- Polygalacturonic acid).

PIB Cal: (cm^{-1})	PIB Exp: (cm^{-1})	AB Cal: (cm^{-1})	AB Exp: (cm^{-1})	Assignment
1815	1760	-	-	$\nu(C=O), PGA$
1720	1680	1720	1660	$\nu(C=O), bead$
1620	1550	1620	1550	ν Benzene ring
1267	1210	-	-	$\nu(CO), \nu(CC), PectinRing$
1202	1210	1205	1215	$\nu(CN)bead$
1125	1145	-	-	$\nu(CN)$ new covalent coupling
1590	1610	-	-	$\nu(C \simeq N)$ new covalent coupling
1040	1010	-	-	$\nu(CO), \nu(CC), \delta(OCH), Ring$

(rather than the commonly used carboxylic bands), which permits the elimination of the interferences from other cell wall components such as water and proteins. The spectroscopic aspect of characterization makes FT-IR coupled with DFT an analysis method that has the potential to investigate pectin structure in more complex systems, such as the post-coupling reactions.

In section 2.2, we could successfully justify our hypothesis put forward in the first section and the IR characterization technique was successfully utilised on bulkier and coupled systems. Here in, it was confirmed that IR can be used as an effective tool in studying these coupled systems. Tables 2.2 and 2.3 summarise our effort

to characterize the coupling reactions using IR spectroscopy. A good agreement between the experimental and simulated spectra confirm the reliability of this technique. This work provides evidence that not only is pectin attached to these beads but it is via the formation of a specific recognizable bond and that pectin is coupled at its reducing end with the intended -CN bond.

Chapter 3

Compact Packing of Polymers.

In chapter 2, we examined the spectral characterization of biopolymer molecules, in particular pectin. Using oligomers allowed full DFT calculations to be carried out in a useful way. However here we are motivated by the interesting question of how these polymers arrange themselves in solutions. In such cases of systems with many atoms, DFT calculations becomes unrealistic and here we examine the use of coarse grained techniques to address questions regarding the conformations and the packing of long chain polymers. Since the surroundings play an important role in governing the structural properties of these chains in all living beings and chemical systems, it is important that we study these polymers taking the solvent properties into account. These studies are crucial in understanding signaling properties, structure-function relations and equilibrium structure states of different biological systems and materials.

Flexible and semi-flexible chain models for polymers have played an important role in studying different classes of polymers in the recent past [94, 95]. They are good tools in the study of many biologically important polymers such as the DNA, proteins, carbohydrates etc [96, 97]. Some of the thrust areas of research using these different polymer models have been the analysis of nanomechanical properties, the consideration of conformational analysis or folding patterns, and the deformation behavior of polymer networks comprising these single chains. Here, we use a semi-flexible model for a polymer chain that incorporates chain stiffness

as it is known that most biopolymers are stiff due to intra and inter molecular bonding. In the past the folding pattern and conformations of flexible and semi flexible chains in different solvent conditions and kinetic pathways have been studied by Brownian dynamics simulations (BDS) [98]. Here we use BDS to perform similar calculations but additionally focus on the implications of the introduction of a torsional potential (described in the introduction).

3.1 Relevant Conformations of a Polymer chain in Different Environments.

The solvent system or the surroundings that a polymer chain encounters can be crudely classified as either (a) a good solvent, (b) a theta (θ) solvent or (c) a poor solvent.

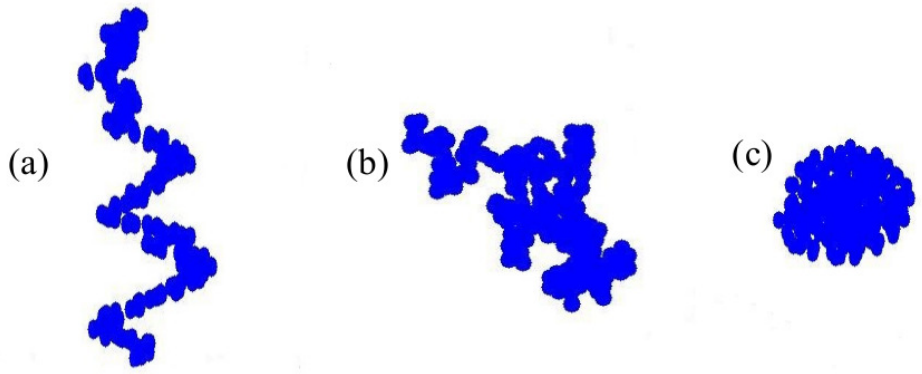

FIGURE 3.1: Flexible chain in a (a) Good solvent, (b) Theta solvent and (c) Poor solvent.

Good solvents are those which favor the expansion of the polymer molecules and show a favorable interaction with the polymer chains that helps the chain to extend. Poor solvents on the other hand, tend to separate the polymer chains from themselves and the chain prefers itself, being compact rather than exploring its

surrounding exposing surface area to the solvent. Theta solvents behave as an intermediate system, where the polymer chain exhibits properties intermediate to good and poor solvents (where the enthalpic interactions with the solvent balance the excluded volume interactions that tend to expand the coil). A flexible polymer chain in good solvent conditions takes up an extended or random coil state and spreads itself into the surroundings while in poor solvent conditions, the polymer chain confines amongst itself and scrunches up like a sphere or a ball, since the sphere like structure possesses the least surface area and minimum energy. Figure 3.1 shows a schematic representation of a flexible polymer chain in different solvents.

However, unlike a flexible chain, a semi-flexible chain has a significant chain bending energy that resists deformation. Since most biopolymers are stiff,

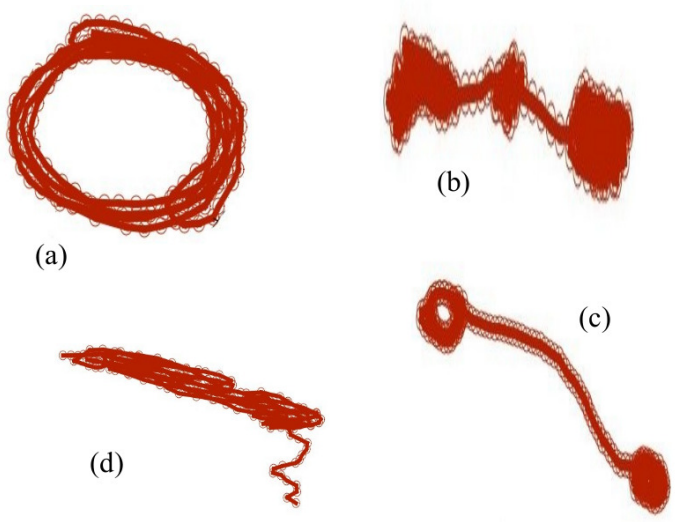

FIGURE 3.2: Semi-flexible chain structures (a) toroid formation, (b) pearl-necklace model in a poor solvent (c) tadpole model (d) bundling of a less stiff semi flexible chain in a poor solvent.

investigations of semi-flexible chains in different solvent conditions are generally good models for the studies of this class of polymers. In the case of a stiff semi-flexible chain, in good or theta solvent the observed behaviour is similar to that seen in the case of a flexible chain. But in poor solvent conditions, they collapse

into toroidal structures rather than more "spherical" ones [99]. This is due to the excessive bending in the collapsed polymer conformation which penalises the spherical state. Depending on the relative importance of the bending rigidity in the chain, the polymer chain can explore different conformations in the poor solvent state such as polymer bundles, pearl-necklace models [100], tadpole loops and so on. Figure 3.2 shows a schematic representation of different conformations exhibited by a semi-flexible polymer chain with different chain stiffnesses in different solvents that we have calculated.

3.2 Secondary and Tertiary Conformations.

While rarely mentioned explicitly in generic coarse grained polymer simulations such as those mentioned above an interesting aspect of experimentally observed biopolymer conformations is the formation of 'secondary' and 'tertiary' structures. It is well-known that proteins [101], nucleotides [102] and indeed certain carbohydrates like amylose [103], curdlan [104] and hyaluronan [105] possess distinct secondary structures referred to as helices, β-sheets, hairpin loops, etc.

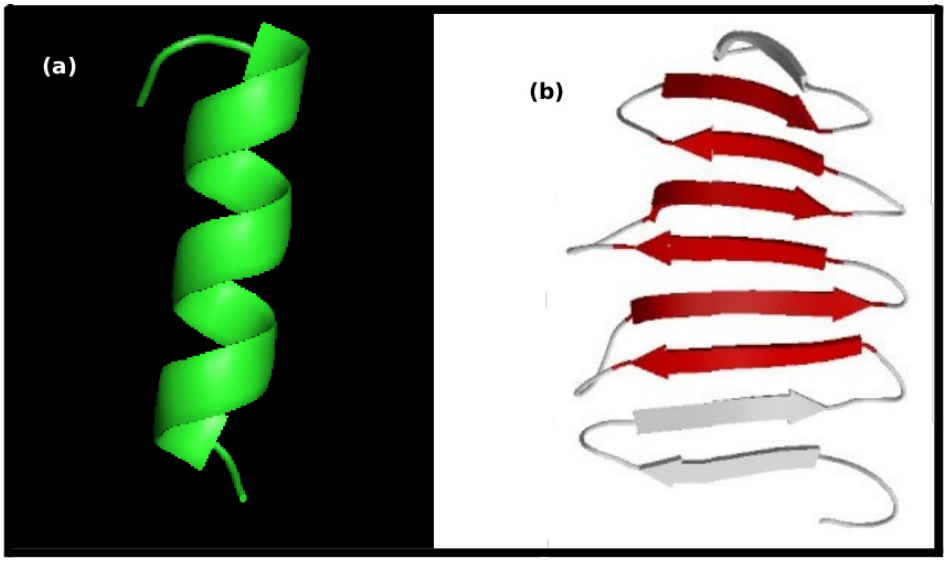

FIGURE 3.3: Schematic of the most common secondary structures found in proteins (a) α-helix. and (b) β-sheet.

One of the most regularly found secondary structures, especially in polypeptides, are the so called α-helices. These are conventionally described as being stabilized by hydrogen bonds between the adjacent atoms in the peptide groups along the main chain. A helix is seen to establish a hydrogen bond between the n^{th} and $(n-4)^{th}$ residue. Its not necessary that all stretches in a protein should form an α-helix. Another relevant secondary motif is the β-sheet. Contrary to the α-helices making intra molecular hydrogen bonds locally for stabilization, the β-sheets form hydrogen bonds with adjacent polymer strands running parallel or anti-parallel to each other. The number of polypeptide sections in a protein forming a β-sheet structure can also vary. The major difference between the

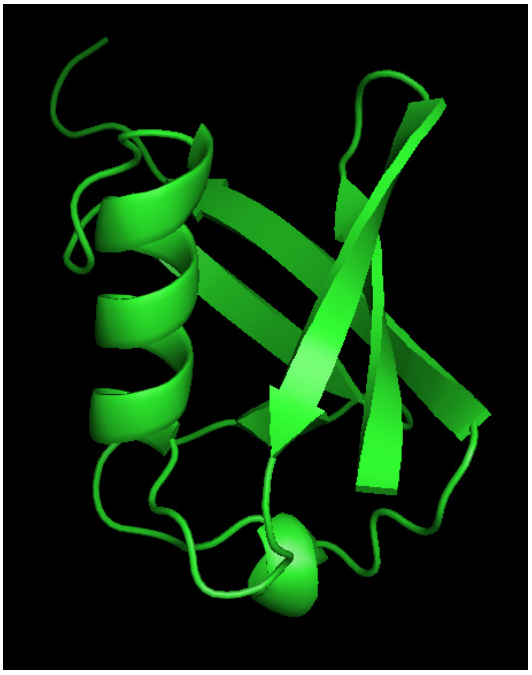

FIGURE 3.4: A global protein structure (ubiquitin molecule obtained from PDB, showing both α-helical and β-sheet regions within it).

α-helix and the β-sheet is then, that in former, the interactions take place amongst residues which are in close proximity in the primary atomic structure, whereas in the latter case, interactions take place between the residues of different parts of the protein that can be far apart in the primary structure. Figure 3.3 shows a schematic representation of these most relevant secondary structures. Apart from

these, other secondary forms encountered are the 3_{10}-helix and the π-helix [101]. They show some prominent differences in there spatial arrangement compared to the α-helix.

The 'tertiary' structure for these classes of polymers describes the global three dimensional arrangement of all its comprising atoms. This generally includes several commonly found 'secondary architectures' such as the α-helix and β-sheet, as discussed above, which can vary in size, in a single structure as shown in the figure 3.4. Other than these, it should also be noted that a number of protein molecules are found to be intrinsically disordered in their native state [106].

3.3 Relevance of Thickness in Polymer Models.

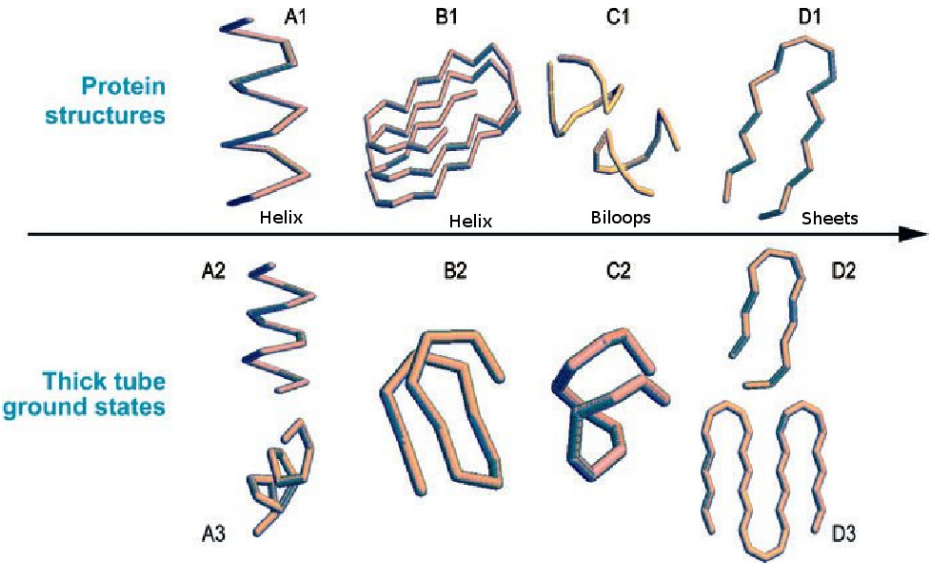

FIGURE 3.5: The experimentally observed polymer secondary structures compared with the simulated geometries using the 'tube model' - *Banavaer et al*

In the past, people have used many different models for simulating the physical properties of polymers and biopolymers. These models could successfully demonstrate the emergence of relevant global polymer architectures such as the swollen,

collapsed, toroidal, bundled etc, as discussed. However, until recently the emergence of distinct secondary structures of the type discussed here seemed to be absent from the predictions of these models. Instead the secondary structures have largely been considered to be the outcome of the fine structure of the polymer chains with volume interactions between different residues during assembly stabilising the structures [107–109]. However a recently proposed 'tube model' [110] suggests that polymers could exhibit different secondary structures in simulations parameterized simply by the total length of the polymer chain, the lengthscale of the short range interactions in the chain and its thickness. The thickness essentially prevents the toroidal collapse of semi-flexible polymers and marginally compact structures like helices and sheets now emerge.

While investigating the result of the incorporation of a realistic torsional potential into the BDS scheme seems an interesting line of inquiry in its own right, here we are specifically motivated by the abstracted physics of the "tube model": a competition between solvent quality and bending with thickness included. In particular it was investigated whether a torsional term might have the same effect as an apparent thickness and allow the formation of biologically relevant secondary structures to be investigated in a BDS framework.

3.3.1 Polymer Model.

A Brownian dynamics simulation technique was used to study polymeric systems, described in general in Chapter 1. The model comprises of N spherical beads connected by (N-1) springs in a 3D-space. According to the Langevin equation, the dynamics of a single bead is described as:

$$m\mathbf{r_i''} = -\nabla U_i - \zeta \mathbf{r_i'} - \eta_i(t), \tag{3.1}$$

where $\mathbf{r_i}$ is the position vector for the monomer i, U_i is the total potential energy of the monomer, $\zeta \mathbf{r_i'}$ is the viscous drag or the frictional force and $\eta_i(t)$ accounts for the noise attributed to random collisions with the solvent molecules. The frictional

coefficient ζ is set to unity for all simulation purposes and accounts for a moderate damping. The noise term $\eta_i(t)$ is assumed to be Gaussian with zero mean and a variance of $6k_BT\zeta\delta_{ij}\delta(t-t')$. The total molecular interaction in the polymer chain U_i is given by (3.2)

$$U_i = U_{bond} + U_{bending} + U_{interactions}, \tag{3.2}$$

$$U_i = \frac{K}{2}(|\mathbf{r_i} - \mathbf{r_{i-1}}| - \mathbf{l})^{\mathbf{2}} + k_BT\frac{\kappa}{2l}{\theta_i}^2 + \frac{\epsilon}{2}\sum_{i,j=1,i\neq j}^{N}(\frac{\sigma}{|\mathbf{r_{ij}}|})^{12} - s(\frac{\sigma}{|\mathbf{r_{ij}}|})^6, \tag{3.3}$$

where ϵ is the magnitude of the Lennard-Jones interaction between the beads and σ is the effective range of its interaction. *'s'* is incorporated as a scaling factor in order to model very poor solvent conditions [111] (3.3). It should be noted that using a FENE potential to describe the bond stretching rather than the Gaussian force as shown here does not change the conclusions of the work.

The tube model discussed in the introduction shows that by introducing a "thickness" to the chain the modeled polymers could take up some relevant, biopolymeric conformations. Motivated by this the effect of the incorporation of a thickness in our bead-spring model was investigated in order to observe if this feature makes any difference to the accessible conformational space. In order to mimic the thickness, the addition of a torsional term in our force law was used with the torsional angle ϕ defined by three successive bond vectors [46], penalizing cis co-planar arrangements and in that aspect exhibiting a "thickness". Figure 3.6(a) and 3.6(b) gives the pictorial representation of the torsional angle. Thus our modified potential is given by:

$$U_i = U_{bond} + U_{bending} + U_{interactions} + U_{torsional}, \tag{3.4}$$

where $U_{torsional}$ is given by

$$U_{torsional} = \sum_k C_k \cos^k \phi_a \tag{3.5}$$

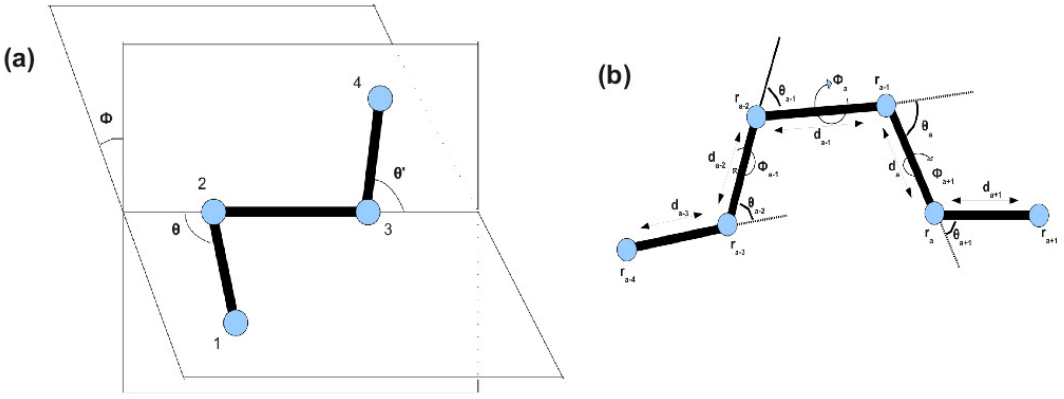

FIGURE 3.6: (a) Schematic representation of the torsional angle. (b) Calculation of the torsional potential.

$$cos\phi_a = \frac{(\mathbf{d}_a \times \mathbf{d}_{a-1}) \cdot (\mathbf{d}_{a-1} \times \mathbf{d}_{a-2})}{|\mathbf{d}_a \times \mathbf{d}_{a-1}||\mathbf{d}_{a-1} \times \mathbf{d}_{a-2}|}, \tag{3.6}$$

and the coefficients C_k define the relative importance of the torsional potential term.

In all our simulations, time is measured as $\tau = (\frac{\epsilon}{ml^2})^{\frac{1}{2}}t$ and the equations of motion are integrated with a time step $\Delta t = 10^{-4}(\frac{\epsilon}{ml^2})^{\frac{-1}{2}}$. All the lengths are measured with respect to l and the energies are measured in terms of ϵ, i.e $\rho_i = \frac{r_i}{l}$ and $U^* = \frac{U}{\epsilon}$. The dimensionless temperature is given by $T^* = \frac{k_B T}{\epsilon}$. Initially, the ground states of these polymers were mapped without the addition of $U_{torsional}$ in the force law, under different solvent conditions and for chains with different chain stiffness, as previously carried out in such a framework. Having gleaned confidence from the good agreement of these initial results with previous work, similar simulations were subsequently carried out with the incorporation of $U_{torsional}$ in the total potential and the ground state conformations examined in these cases. Simulations were carried out multiple times yielding consistent results. Polymer chains were initialized in a random configuration and allowed to equilibrate to the

ground state. Simulations were carried out for polymer chains with N=50, 100, and 200 beads; with no convincing differences observed in phenomenology. Unless stated all data shown are for 200 beads. The values of ϵ and s were tuned in order to simulate different solvent states; $K = \frac{200k_BT}{l^2}$ was used with dimensionless temperature $T^* = 1$, the bending stiffness constant κ was varied between $3l - 15l$ and the values of C_k were initially set to unity, in order to study chains of different stiffnesses. To quantatively describe the simulation ground states and the growth of different conformations the radius of gyration(R_g) and an orientational order parameter(S) of both global and local order [112, 113] in the chain were monitored. These are given by:

$$\langle R_g \rangle^2 = \frac{1}{2N^2} \sum_{i,j} (\mathbf{r}_i - \mathbf{r}_j)^2, \tag{3.7}$$

$$S = \frac{\langle 3cos^2(\theta) - 1 \rangle}{2}. \tag{3.8}$$

S is defined by 'θ', which is the angle between the orientation bond vectors for all beads i and j. The orientation bond vector for a particular bead i is given by $(\mathbf{r}_{i+1} - \mathbf{r}_{i-1})$. The local order (LO) parameter refers to the order between any two adjacent bonds and the global order (GO) parameter is the one between any two bonds. Thus, we can see that S defines a four bond parameter and this value can be used to define the helicity of helicoid's [114]. R_g and S were monitored over a period of every 10^4 iterations.

3.3.2 Results and Discussion

Initially standard semi-flexible chains in different solvents were equilibrated and conformations were examined as a function of varying stiffness. Figure 3.7 summarizes this simulation set and shows the ground states of such semi-flexible polymer chains, varying with respect to the solvent conditions and to different chain stiffnesses, using the potential described by equations 3.2 and 3.3. In this schematic

it is indeed clear that in the case of a less stiff polymer chain in a poor solvent regime, the polymer just collapses on itself minimising its surface area and hence the energy. On increasing the stiffness of the chain, intermediate structures like two-headed tadpoles(not shown here), and ultimately toroids are formed.

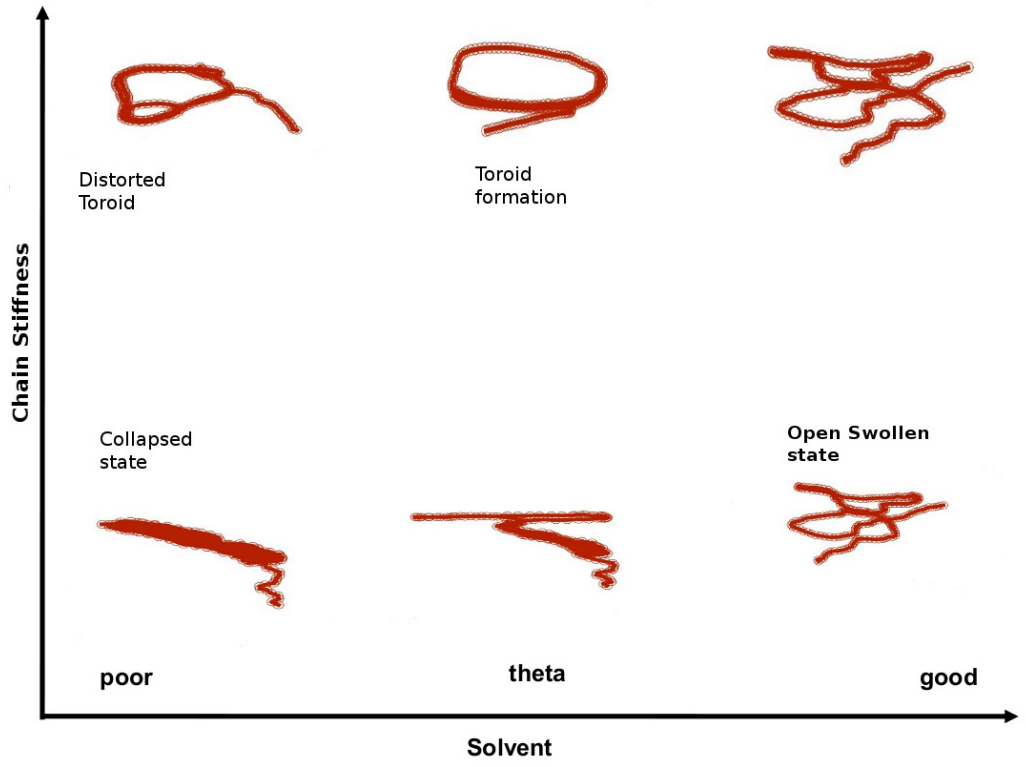

FIGURE 3.7: Typical ground states of semi-flexible polymer chains in different solvent conditions and at different chain stiffnesses, **without** a torsional potential.

These structure are in good accordance with the equilibrium structures proposed by Williams *et al* [98, 115] in their earlier work. Furthermore, electron microscopy experiments on stiff DNA duplexes and triplexes of the polysaccharide curdlan and xanthan [116] confirm some of these toroidal structures can be manifest by real biopolymers.

Figure 3.8 summarizes the results of the same type of simulation set, but where we have endeavored to introduce a thickness like effect, in the spirit of the physics of the tube-models discussed, with the introduction of a torsional term in the

potential, as described by equation 3.4. Incorporating a torsion in the chain limits the excessive bending of the chain by penalising cis co-planar arrangements and thereby results in a chain conformation which resists collapsing.

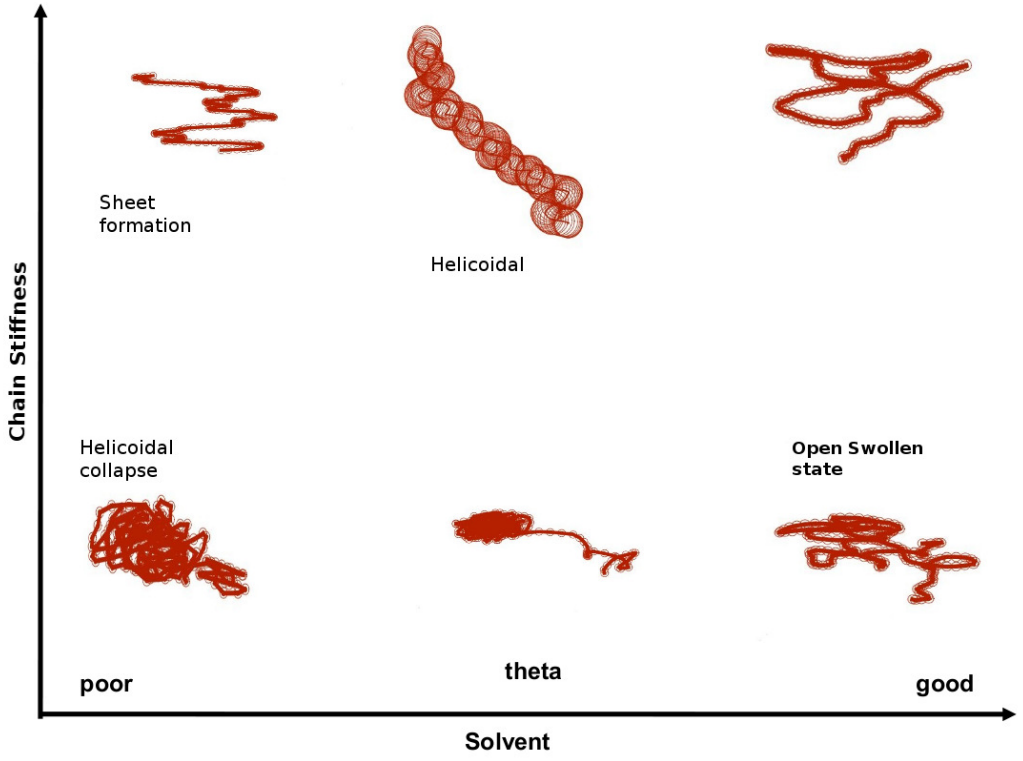

FIGURE 3.8: Typical ground states of semi-flexible polymer chains in different solvent conditions and at different chain stiffness, **with** a torsional potential.

This figure shows that in good solvents less stiff, random chains behave similarly to those in Figure 3.7, ie they have an open, swollen state to begin with and then as the stiffness increases, a more open chain conformation results. However, in the poor solvent regime, the behavior appears to diverge somewhat from that found without the torsional term. For a less stiff chain, the polymer attempts to minimize its surface contact with the solvent but with the torsional term included, the chain collapses to novel marginally compact states rather than collapsing to a toroid. More specifically when the chain stiffness is increased further, the collapse of the polymer chain has a large bending energy penalty and preferred compact shapes such as toroids would be expected, as observed in the earlier case. However

the contribution of the torsional term prevents the chain from simply winding up by collapsing rings on top of one another; and a planar sheet like structure is manifest.

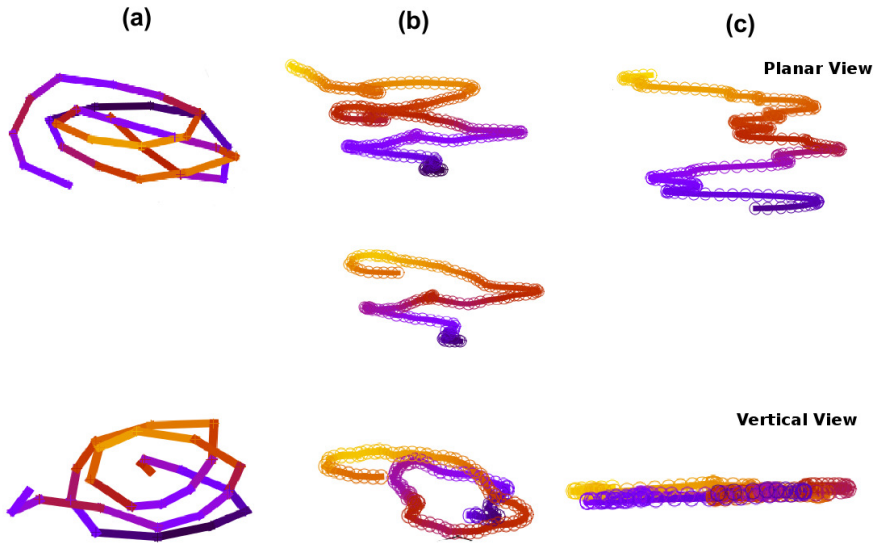

FIGURE 3.9: Different views of helicoid and sheet structures from our simulations: (a) Helicoidal collapse of a chain obtained for a less stiff chain in poor solvent, (b) Helicoid formation of a stiff polymer chain in theta solvent and (c) Sheet like collapse of a stiff chain in poor solvent.

The theta solvent also reveals interesting behavior and shows considerable deviation from Figure 3.7 with the chain driven towards a more helicoidal [Figure 3.9(b)] phase, opening up the toroid. Different 3-D views of these structures are shown in Figure 3.9. These ground-state structures found in Figures 3.8 and 3.9 are not only in accordance with the results of the tube-model [117–119], but also resemble standard motifs commonly found in biopolymers.

To quantify these observations and allow a comparison with not only with previous work but also real biological motifs, the physical chain parameters described by the radius of gyration (R_g) and an orientational order parameter (S) were analyzed. In order to differentiate these states, the R_g and S of the extended and

the collapsed conformations of a polymer chain having different stiffnesses, when applying the different potentials are calculated. Figure 3.10 shows the value of R_g with respect to

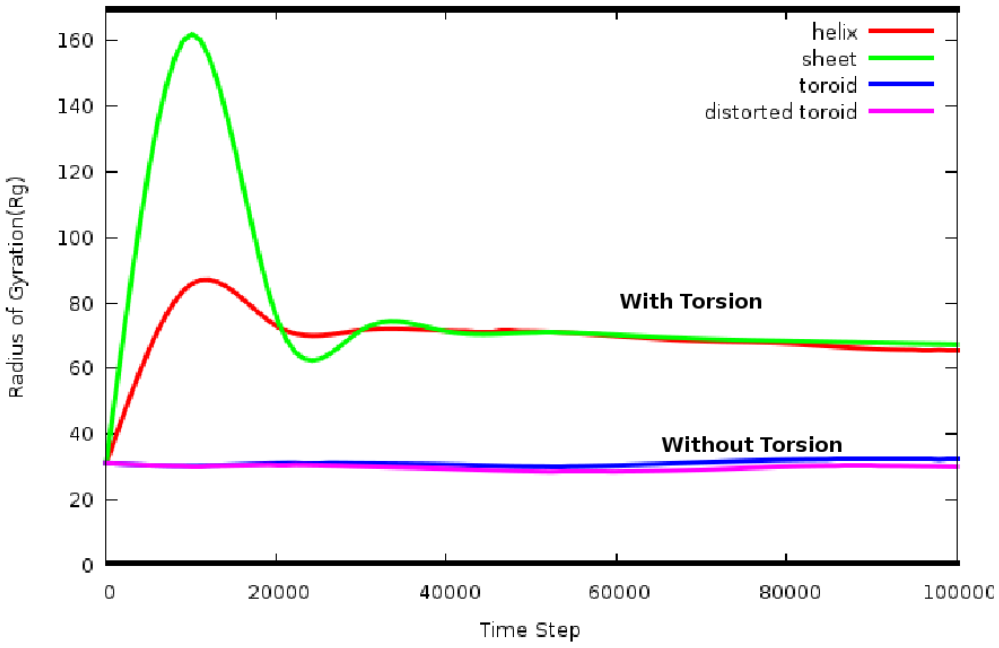

FIGURE 3.10: Propagation of the radius of gyration of the system of stiff and less stiff chains with and without a torsional potential.

time, reporting on the collapse trajectory of a chain of different stiffness with and without a torsional potential. In the figure, there is a convincing difference between the different conformations observed with and without the torsional part of the potential. The former results in structures like sheets and helicoids, where as the latter ends in a toroidal phase. The higher value of R_g obtained by incorporating the torsional term into the generating potential does indeed reflect the characteristics of a helicoid and sheet like structures, with the ordered helical structure tending to stay extended along its helical axis. As expected, the radius of gyration for toroids is always found to be comparatively small.

Furthermore, we examined the orientational order parameter (S) for these different states (Figure 3.11(a) and 3.11(b)). Figure 3.11(a) shows a value of $S(GO) \approx 0.4$ for the helicoid phase formed with the incorporation of the torsional term, which

TABLE 3.1: The global order parameters for realistic sheets and helices from specific proteins residues. "GO" refers to the global order parameter.

Protein	Residue (atom's)	N	Motif	GO(S)
Ubqn(1UBQ)	1-16	48	Sheet	0.27
SynpVII(2D8K)	12-36	76	Sheet	0.27
Lzme(1AKI)	41-61	67	Sheet	0.29
β-lgn(2AKQ)	51-71	67	Sheet	0.27
Ovlbn(1JTI)	217-231	75	Sheet	0.25
Lzme(1AKI)	90-100	32	Helix	0.012
Ovlbn(1JTI)	2-21	61	Helix	0.0064
Ubqn(1UBQ)	22-34	39	Helix	0.0019
β-lgn(2AKQ)	130-140	33	Helix	0.014
Memb: Protein(2KE4)	390-419	90	Helix	0.0078

TABLE 3.2: The global order parameters for simulated polymer architectures with a torsional potential incorporated. "GO" is the global order parameter.

Simulation	C_K	K_s	K_b	ϵ	N	Motif	GO(S)
1	1.0	200	15.0	2.0	200	Helicoid	0.43
2	0.7	200	15.0	2.0	200	Helicoid	0.27
3	0.1	200	15.0	2.0	200	Helicoid	0.027
4	1.0	200	3.0	2.0	200	Sheet	0.301
5	0.7	200	3.0	2.0	200	Sheet	0.27
6	0.3	200	3.0	2.0	200	Sheet	0.25

is comparable to that obtained for a helical phase by Kemp *et al* [114]. Toroids, on the other hand, have a small global order.

To look at the chain order on a smaller length scale, we monitored the local order as described in 3.8. Both helical and toroidal structures show similar behavior as they both possess well ordered local regions in them. Similar relations could be seen in the case of the order parameter for the sheet-like versus distorted-toroid conformations generated from chains of different stiffness (Figure 3.11(b)). While this clearly distinguishes the conformations with and without a torsional potential, to investigate further the degree to which these simulations capture the

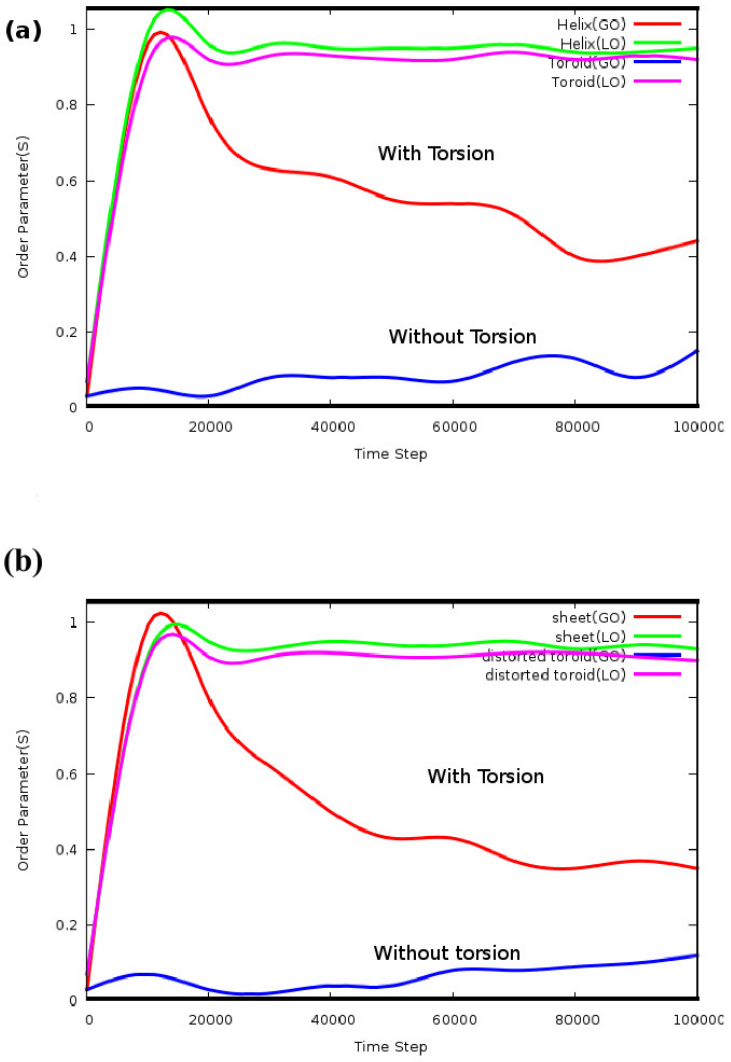

FIGURE 3.11: (a)Global and Local order parameters of the system of stiff and less stiff chains with and without a torsional potential, forming a helicoid. (b) Order parameters of the system of stiff and less stiff chains with and without a torsional potential, forming a sheet. "LO" corresponds to the local order and "GO", the global order parameter.

features of actual biologically relevant structures, we selected five proteins from Protein Data Bank (PDB), an online database that contains crystallographically determined coordinates of the bonds in many proteins. Protein backbone coordinates of the individual helix and sheet regions from these global protein structures were isolated and the global order parameter of these regions was determined. The global order seems to be the most relevant since in real structures, the coordinates reflect the atomistic details and thereby local order parameters simply the presence of specific bond angles. The bead spring model on the other hand is coarse

grained and lacks any atomistic details in the molecule. Tables 3.1 and 3.2 summarize the global order parameters for realistic protein α-helix and β-sheet residues and those from our simulations. The values from the table suggest that indeed by tuning the relative importance of the torsional term (C_k value), global order comparable with real protein secondary structures can be achieved. The fact that such marginally compact structures can be reproduced without explicitly considering volume interactions or indeed charge interactions is surprising considering the polyelectrolytic nature of most biopolymers and perhaps suggests that such interactions are significantly screened in-vivo, or alternatively that they can be reasonably accounted for by the equivalent interaction strength.

3.4 Summary.

In this work, the folding pattern of a standard bead-spring polymer chain is investigated by Langevin dynamics simulations. In all cases, a bending potential takes care of the chain stiffness and a L-J term accounts for the excluded volume and the solvent effects. Our focus has been on the differences in the conformational phase space resulting from the inclusion or not of a torsional potential. Without such a torsional term the chains exhibit conformations previously observed by such coarse-grained models. However, the addition of a torsional term into the energy expands the breadth of the conformational space; and the polymer now exhibits previously un-manifest marginally-compact structures that resemble well-known biological forms; secondary structures such as helices and sheets. The radius of gyration (R_g) and order parameter (S) shows that these novel structures are not only distinct from the previously found structures, but also that the relative importance of the torsional term can be modified in order to achieve structures that are comparable with real α-helices and β-sheets.

With structures of Sodium and calcium pectates, pectic acid and pectinic acid all reported as right-handed (3_1) helices [120] and other polysaccharides such as xanthan and amylose exhibiting helical motifs, this work on standard polymer chain

conformations in different solvent conditions and chain stiffness can also be seen to be relevant to polysaccharide (pectin) chain structures in different environments.

Chapter 4

Mechanical Properties of Single Chain Pectins.

Single molecule force spectroscopy using the atomic force microscope has given access to an unprecedented level of information regarding the stress response of a host of biopolymers [121–124]. Studies of polysaccharide stretching have yielded particularly interesting data, the interpretation of which requires the marriage of statistical mechanical theories of polymer physics to the complexities afforded by possible force induced rearrangements or even conformational transitions of the constituent sugar rings. Such monomer transitions during stretching, from classical chair forms of the pyranose ring to more elongated arrangements, function as molecular sacrificial bonds [125], increasing the polymers contour length and thus producing characteristic deviations, or clicks, in the slope of the force-extension curve. Indeed, many curves exhibiting features of this type have been measured [25, 32, 33, 103, 126–128], and while accounting for their detailed origin is still a work in progress [129, 130], the fact that they are observed in polysaccharides in which the rings are axially linked and are completely absent in equatorially linked analogues, gives great credence to the idea that axial bonds facilitate these transitions by acting as molecular levers. Thus, it is central to the interpretation of polysaccharide force-extension curves to understand that those polymers possessing only equatorial glycosidic linkages such as cellulose (1e-4e) do not have the

possibility of undergoing stretching driven conformational transitions, while those with one axial and one equatorial bond, such as amylose (1a-4e) or dextran (1a-6e) can undergo one conformational transformation, and those with both bonds axial, such as pectin, two [131]. This is explained in detail further in Chapter 1.
Previously, force-extension studies performed on pectin have discussed the possible biological significance of force induced conformational transitions in processes such as signaling by for example modifying its binding to lectins. The magnitude of force that triggers these conformational changes in the sugar rings of the pectin moiety has been measured and indeed those thought to exist in biological systems approach this value. As described in Chapter 1 pectins extracted from different sources and/or by different methods are typically obtained with different amounts of methyl esterification in the chain backbone (DM). In this regard, an important question is to address is: how does the different degree of esterification in pectin chains influence its stretching pattern? While the initial pectin stretching study was carried out on a sample of limited DM (9%), such molecules that might be stretched in the cell wall are likely to be closer to 80% esterified. Also we know that degree of esterification plays a crucial role in calcium binding properties and the gelation of pectin. So it becomes important to ask, how the click behavior of single pectin chains might impact on the nature or the structure of pectin gels formed and its rheological behaviour?

4.1 Stretching Single Chain Pectins: AFM and DFT Studies.

Pectin is one of the few polysaccharides in which the chain linkages at both sides of the constituent pyranose rings are axial (1a-4a), but nevertheless it has received comparatively little attention since an initial study of its single chain stretching behavior was reported almost a decade ago [33], despite the fact that it is an important structural polysaccharide. It is found in the cell walls of all land plants

and although this is a complex biological matrix in which hemicelluloses, pectins, cellulose, proteins and lignin all play a role in determining structure and properties, it is known that the pectin component has considerable mechanical utility within the cell wall [132, 133]. Pectin is extracted commercially from lemon peel and apple pomace and consists mainly of a 1,4 linked α-D-galacturonate residues (typically $\approx$ 90%). While this dominant part of the pectin, referred to as homogalacturonan, is a linear and relatively stiff anionic polymer [134], the sugar residue also naturally occurs in a substituted, methylesterified form as described in chapter 1.

It has been speculated that the force-induced conformational transitions in pectin molecules, proposed to explain the origin of two clicks observed in the single molecule force-extension curves, may have biological significance. This could involve acting as a sensor with a signaling role [33], or as a compliant element possessing extra elastic extensibility, directly contributing to the control and maintenance of the mechanical properties of the cell wall [128]. In this context it should be noted that the fine structure of many polysaccharides, particularly those fulfilling a structural role, can be modified spatially and temporally by the orchestrated action of specific enzymes, and that the nature, amount and pattern of substituents thus generated is known to significantly modify the intermolecular interactions of the polymer. It particular, it is clear from in vitro studies that the degree and pattern of methylesterification modulates calcium binding and the propensity for intermolecular hydrogen bonding in pectin systems [135, 136]. It is therefore interesting to consider the effect that such modifications might have on the nanomechanical behaviour of the single chains and thereby their implications for the proposed mechanisms by which force-induced conformational transitions might be functional in muro.

In fact it has been predicted in a molecular dynamics study that sugar ring substituents could significantly alter the force-extension behavior of single chains [137] and experimentally observed that methylcarboxylation has a large effect on the force extension curve of dextran [25]. In addition, it would be expected that substituent groups could extensively modify interring hydrogen bonding and in that

way may also have a significant effect on the nanomechanics of single chains. Indeed, recent studies have shown that the importance, and even the strength, of interring hydrogen bonds might be assessed by carrying out single chain stretching studies in solvents of different dielectric constant [138, 139].

Here, we specifically investigate the consequences of the degree of methylesterification for the single molecule stretching behavior of pectin. After first observing the behavior of polygalacturonic acid, a completely unsubstituted chain, we use DFT calculations to investigate the effects of methylesterification upon the intrinsic stabilities of different ring conformers. Full force-extension curves are subsequently predicted using the results of these calculations by employing an extensible wormlike chain (e-WLC) model, incorporating the concept of an equilibrium of states to account for the conformational transformations [30, 140, 141], and these are compared with the experimental results for pectin samples of known methylester contents. Experiments were also performed in solvents of low dielectric constant in order to address the significance of interring hydrogen bonding in highly methylesterified samples.

4.2 Experiment.

Pectin samples derived from lemon peel were kindly supplied by CP Kelco ApS, DK 4623 Lille Skensved, Denmark. These samples were manufactured from the same highly methylesterified mother pectin extraction. Their DM (31 and 78 % respectively) was controlled via treatment with a pectin methyl esterase (PME) of fungal origin (CP Kelco, private communication). This enzyme is believed to generate intrachain distributions of methylesterification that are close to random. Polygalcturonic acid (DM 0) and highly methylesterified (90 %) samples were purchased form SigmaAldrich. The sample average degree of methylesterification and also the width of the intermolecular DM distribution of each sample were experimentally determined by capillary electrophoresis as previously described [61, 63, 80]. The CE method is explained in Appendix 2. An estimate of the width of this distribution of DM among polymers is particularly important when it is

being sampled one chain at a time [142]. It should be noted that often methods give a sample average DM value so that the DM of each chain picked up could be different.

In the AFM experiments, a single molecule was physisorbed to a substrate and an AFM tip. As the tip and substrate separation is increased by moving the piezo on which the substrate sits, the molecule straightens and stretches, with the force applied to the molecule determined from the deflection of the flexible AFM cantilever. The samples were prepared by applying 20 μ l of 0.01 $w/w\%$ solutions in deionized H_2O to clean glass discs, which were then dried at 11.3% relative humidity overnight. This was then extensively rinsed with deionized H_2O leaving only the tightly bound molecules on the surface. After drying the sample was mounted in the AFM and the liquid cell filled with the appropriate solvent just prior to the force curve measurements. Force-distance curves were recorded by pulling the molecules at 0.54 μ m s^{-1} using a scanning probe microscope (Veeco Nanoscope E) with a Si AFM tip that had been calibrated using the thermal excitation method. We performed our AFM experiments in deionized water, 0.1 M sodium phosphate buffer, or in hexadecane.

4.3 DFT Calculations.

The B3LYP functionals with the basis sets denoted by 631G* and 6311++G** were used as previously described in the introduction and in detail if Chapter 2 [75], to study α-D-galacturonic acid, methyl esterified α-D-galacturonic acid and α-D-galacturonic acid ion. The results described in the next section are based on the galacturonic acid ion and its methyl ester, unless otherwise specified, in order to best simulate the charged sugar species relevant to the pH of the experiment. B3LYP/6311++G** density functional studies on glucopyranoses and disaccharides like maltose and cellobiose shows that this level of theory is reliable for the prediction of geometries and energies for saccharides [143] when compared with the experimental ones. Here, the density functionals and the basis sets, B3LYP/631G*

and B3LYP/6-311++G**, were used in the *pcgamess* software. The starting structure for the galactouronan was generated with the help of SWEET II [144, 145], a database for carbohydrate building. The bond length of the sugar molecule was fixed by the distance between the two consecutive glycosidic oxygen atoms ($O1 - O4$). Increasing the constrained $O1 - O4$ distances for consecutive calculations, we performed full geometry optimization of our molecule at each step during stretching, to mimic our AFM experiment. The residue vector ($O1 - O4$) was changed by $0.1A^{\circ}$ between calculations. In the region of interest (transition to different conformations), we narrowed the range going to 0.05 and 0.025 A°, to get a better understanding of the path followed in the conformational process. The preliminary geometry optimization was carried out at B3LYP/631G* level and the optimization continued from that result using DFT and a higher order of 6-311++G** basis. The convergence criteria energy differences between cycles of optimization were less than 1×10^6 Hartree and that for the gradient was set to be less than 1×10^4a.u [146]. These calculation procedures were repeated with the ionised and methyl esterified anomers of α-D-galacturonan. Due care was taken while working with the pectin ion, so that there was no interference of the interaction of the O^- of the acid group and the -OH proton at the 3rd and 4th position of the sugar ring. In real systems, this never happens as the pectin molecules and highly solvated and thus preventing the O^- ion forming hydrogen bondings with these -OH groups. At the extended state of 5.376 and 5,476 A° respectively, the position of the -OH proton on the 3rd position of the ring C is restricted away from the O^- of acid group of the ion. The free energies and the length scales of the relevant conformations are thus used in the e-WLC model to determine the population of different stable conformers and hence the force-extension profile is estimated.

4.4 e-WLC Model for Chain Extension.

To develop the model we consider a polymer consisting of glycan rings, which are able to exist in different (two or three) conformers that have different lengths.

These conformers may have different energies and may have an energy barrier for the transition between the states.

4.4.1 Single-click model

It has previously been shown, by using a Monte Carlo simulation, that at extension rates used in AFM stretching experiments (typically around 500 nms^{-1}) polysaccharide molecules maintain an equilibrium between the "clicked" and "unclicked" states throughout the stretch. This has also been assumed on the basis of the analogy with the conformational kinetics of cyclohexane. In other words, the height of the energy barrier between the clicked and unclicked states, $\Delta G^{\ddagger}$, does not influence the shape of the force curve. It was calculated that it is not until extension rates approach 1 cms^{-1} that a significant departure from equilibrium conditions might be expected to be observed. We could also describe this by saying that the shape of the force curve is not time dependent under experimental conditions.
We therefore use this notion of an equilibrium state being maintained during the extension process as the basis for our thinking. The number of clicked or unclicked sugar rings in the molecule is given by the Gibbs equation

$$\Delta G_0 = -k_B T \ln K \tag{4.1}$$

where ΔG_0 is the energy difference between the "clicked" and "unclicked" states for one sugar ring (notice we say nothing about the transition between these two states), k_B is the Boltzmann constant, and T is the temperature. The equilibrium constant $'K'$ is the ratio of the number of clicked to unclicked sugar rings.
If a force is applied to the ends of the molecule, this equilibrium is perturbed in favor of the clicked form. This force gradually increases from zero to a large force (nN) over the entire range of the stretching; however, the region where the clicking takes place is only over small portion of this range. The effect of the force can be considered to lower the Gibbs energy difference between the unclicked and clicked states. This lowering of energy will be quantized such that for each click

that takes place the applied force will contribute an amount of energy $F\Delta x$ where F is the applied stretching force and Δx is the increase in length of one unit of the molecule (one sugar ring) along the axis of stretching as a result of the conformational transformation. This Δx is the difference between the length of an unclicked and a clicked X_c sugar ring.

The effective Gibbs energy, or Gibbs energy under an applied force, ΔG_F, for the transformation between the clicked and unclicked polymers is therefore modified by this energy,

$$\Delta G_F = \Delta G_0 - F\Delta x \tag{4.2}$$

We do not have to use Δx as a fitting parameter as this length can be obtained by other means [76].

Now we can calculate the equilibrium constant (the ratio of clicked to unclicked states in the molecule) under a stretching force K_F to be

$$K_F = exp(\frac{-\Delta G_F}{k_B T}) = exp(-\frac{\Delta G_0 - F\Delta x}{k_B T}) \tag{4.3}$$

We then consider the fraction of the sugar rings that have clicked, f_c, under any applied force. This can be related to the equilibrium constant thus

$$f_c = \frac{K_F}{1 + K_F} \tag{4.4}$$

Combining Equations 4.3 and 4.4 and rearranging gives

$$f_c = [1 + exp(\frac{\Delta G_0 - F\Delta x}{k_B T})] \tag{4.5}$$

It is now possible to incorporate this fraction into any of the existing polymer models that incorporate the concept of a contour length l_c - for example, the freely jointed chain (FJC) model or the e-WLC model. For each ring that clicks,

the contour length can be considered to increase. We can therefore adapt the e-WLC model by substituting an expression for l_c. We will use the term l_e, the effective contour length, for this substitution.

The effective contour length depends on the total number of clickable rings, N, the fraction of clicked rings (at the applied force), and the difference in the lengths, Δx, of the unclicked, X_u, and clicked polysaccharide units in the form

$$l_e = N[X_u + f_c \Delta x]. \tag{4.6}$$

If we want to relate this to the original contour length, we have $l_c = NX_u$. Substituting the expression for f_c Equation 4.5 we obtain

$$l_e = NX_u + \Delta x[1 + exp(\frac{\Delta G_0 - F\Delta x}{k_B T})]. \tag{4.7}$$

This term l_e is then incorporated into the EWLC model in place of l_c, which was previously a constant, to give

$$F = \frac{k_B T}{l_p}[\frac{1}{4}(1 - \frac{l}{l_e} + \frac{F}{\Phi})^2 + \frac{l}{l_e} - \frac{F}{\Phi} - \frac{1}{4}]. \tag{4.8}$$

This expression may be readily used to simulate force extension curves for clicking polysaccharides or fit experimental data and extract N (or l_c), ΔG_0, the persistence length l_p, and the modulus of elasticity, Φ.

So far we have assumed every sugar ring within the molecule is able to undertake an extending conformational change. However, many polysaccharide molecules are made up of a mixture of monomers configured such that some may click and others may not. For example, dermochondan sulfate can contain a mixture of D-glucuronate and L-iduronate, of which only the L-iduronate can undergo force induce conformational transformations. Therefore f_c can be modified by multiplying by the proportion of sugar rings that may undergo this conformational change in the molecule and different lengths can be used for Δx.

4.4.2 Double-click model

With previous attempts at modeling the force-extension curves for single-molecule stretching it has been very difficult to extend the models to more than one type of conformational transformation within a molecule. However, it is a relatively simple exercise to extend the model described in this work to two or more clicks. It is also possible to take into account various arrangements that may exist in a heteropolymer but for brevity we do not include all of these permutations here. We imagine three states: (i) where the sugar ring is in it lowest energy form (e.g., starting at $^4C_1$), (ii) where either one end of the ring or the other has clicked (e.g., to $^{1,4}B$), and (iii) where both ends of the ring have clicked (e.g., to $^1C_4$).
The total number of sugar rings in the molecule, N, is therefore made up of the number of unclicked, N_u, the number with one clicked, N_{c1}, and the number with both ends clicked, N_{c2}:

$$N = N_u + N_{c1} + N_{c2} \tag{4.9}$$

These states are in equilibrium with each other with equilibrium constants K_{1F} and K_{2F} for the first and second clicks under an applied force:

$$N_u = N_{c1} = N_{c2} \tag{4.10}$$

The effective contour length can then be represented as a function of the fractions of rings in each of these states if we take Δx_1 and to be the extension of each unit of the molecule to the intermediate state, Δx_2 the extension from the intermediate state to the final doubly clicked state, and X as the length of the unclicked sugar ring:

$$l_e = NX + (N_{c1} + N_{c2})\Delta X_1 + N_{c2}\Delta X_2 \tag{4.11}$$

or, in terms of fractions clicked,

$$l_e = N[X + (f_{c1} + f_{c2})\Delta X_1 + f_{c2}\Delta X_2] \tag{4.12}$$

where the fraction in each state (f_{c1} clicked through to the intermediate form, f_{c2} to the fully extended form) is given by

$$f_{c1} = \frac{K_{1F}K_{2F}}{1 + K_{1F} + K_{1F}K_{2F}} \tag{4.13}$$

$$f_{c2} = \frac{K_{2F}}{1 + K_{1F} + K_{1F}K_{2F}} \tag{4.14}$$

This results in the expression for the effective contour length:

$$l_e = N[X + K_{1F}[\frac{(1 + K_{2F})\Delta x_1 + K_{2F}\Delta x2}{1 + K_{1F} + K_{1F}K_{2F}}] \tag{4.15}$$

where

$$K_{1F} = exp(-\frac{\Delta G_{0_1} - F\Delta X_1}{k_B T}) \tag{4.16}$$

$$K_{2F} = exp(-\frac{\Delta G_{0_2} - F\Delta X_2}{k_B T}) \tag{4.17}$$

Here, ΔG_{0_1} is the Gibbs energy differences for the first click and ΔG_{0_2} is for the second click. Extending the model to include three or more clicks is possible with similar reasoning to that used to develop the two-click model.

4.5 Results and Discussion

The starting geometry of all pyranose rings was considered to be that of lowest energy: the $^4C_1$ chair form. Naturally occurring pectin has some amount of the

methyl esterified residue present and it was confirmed from our calculations that the methyl esterification doesnt make much difference in the molecular length of the single sugar ring in pectin (pectin:- 4.592$A°$ and e-pectin:- 4.560$A°$). We found that, in the course of an increasing applied force, when the stress exceeds some critical amount, the stretched α-D-galacturonan flips into a skew boat ($^{3}S_{5}$ or $^{0}S_{2}$) [31], with flagpoles (refers to the two 1^{st} and 4^{th} 'C' atom in the boat conformation of a sugar ring) occupied by -OH and H groups. When stretched further, the molecule goes to an inverted skew boat ($^{5}S_{3}$ or $^{2}S_{0}$), where the flagpoles are then occupied by the bulky acid -COOH and the H groups. Finally on the application of further force, the molecule transforms to an inverted chair $^{1}C_{4}$ conformation (4.1), where all the axial groups are transformed to be equatorial. We also performed these calculations with the methyl esterified residue of pectin which also showed a similar pathway.

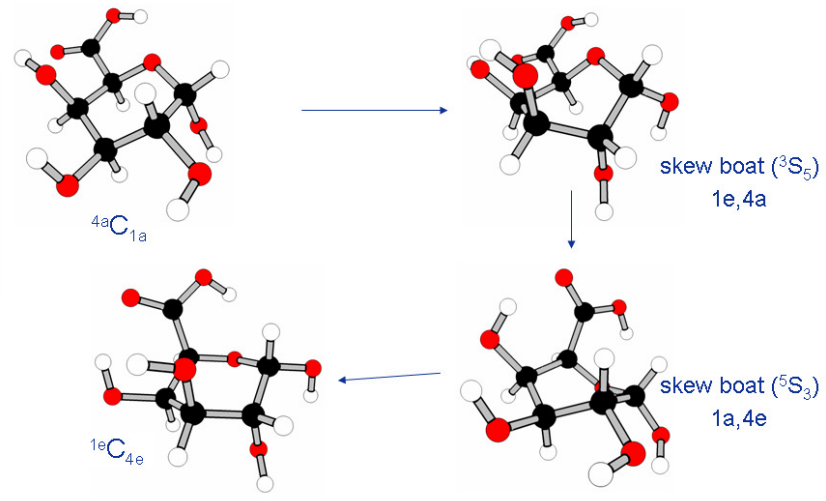

FIGURE 4.1: The geometry of a monomer, in the course of conformational transition in pectin during the forced stretching of a single chain.

Looking at the calculated energy profiles figure (4.2) and table 4.1, it can be seen that the different conformational states reached in the pulling of α-D-galacturonic acid and methyl esterified α-D-galacturonic acid can be well identified. The

TABLE 4.1: Length of α-D-galacturonic acid anion and methoxylgalactose during elongation of the O1-O4 distance using the B3LYP/6311++G** basis set.

Conformation	Galacturonic acid / nm	Galacturonic acid ester / nm
Chair ($^4C_1$)	0.4592	0.4560
Skew Boat ($^3S_5$)	0.5176	0.5147
Inverted Chair ($^1C_4$)	0.5576	0.5547

stretched state conformers (boat conformers, at around $5.15A^\circ$, differ by an energy of 1.3 kcal/mol with the methyl esterified sugar rings slightly lower in energy compared to α-D-galacturonic acid ion. We will show that this energy difference would correspond to a difference in the force required to trigger the first "click" in AFM experiments in the range of 100 pN for between a completly unesterified chain and a completly methyl esterified one.

In conclusion, we could infer from this part of the theoretical study that, a small difference in the energy of the preferred conformation in the stretched state (skew boat structures) could alter the force-extension curves of α-D-galacturonan compared to methyl esterified α-D-galacturonan. That is; an effective substitution could alter the behavior of the substance by a small but measurable amount.

Also, it is seen that lowest energy inverted chair in acid pectin (not shown in the energy profile 4.2) is at around 5 kcal/mole higher in energy to the ground state, where as that in the case of the methyl esterified residue and the pectinic ion is around 1.5 Kcal/moles. This suggests that stretching single molecule pectins could be highly pH sensitive as the ion and the acid have different energy curves. The solvent conditions (pH and ionic strength) could play an important role in the stretching.

In our AFM experiments, pectin chains from samples with average DM's of 31, 78 and 90% were stretched in water at a pH in agreement with our calculations where the unmethylated sugar rings are considered to be charged. It should be noted again that the different ways by which these pectins have been characterized record a sample average property, ie not every chain in the 31 or 90% DM pectin are methylesterified to its total percentage, it is an average property of the total

number of chains in the sample. So, depending on the width of the DM distribution, deviations in the DM of the actual stretched chain might be of the order of 10%.

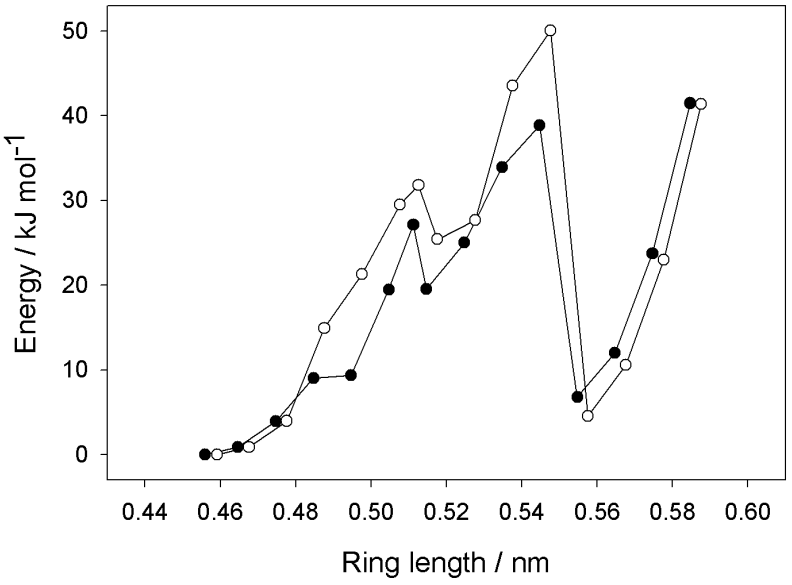

FIGURE 4.2: Comparison of the energy curves of pectinic ion (hollow dots) and methyl esterified pectin (solid dots).

During our stretching experiments, it was difficult for us to get good data in the region of second click owing to the high forces required compared to the typical physisorption adhesion strength. The chain would typically detach before good data could be acquired. Improving this forms a part of further work. The experimental data figure (4.3) shows the concatination of many normalised force-extension curves for different DM pectins in water and shows that in water the methylesterification indeed does not influence the force-extension behavior of pectin by a huge amount. However on a closer inspection, by interpolating the concatinated datasets, it can be seen that the stretches of 31% and 90% DM pectin do exhibit a measurable difference in the click forces of the range 50-100pN (fig 4.4), which is in accordance to that predicted based on the energy calculation predictions from

our DFT simulations performed on single monomer sugars. The predicted force-extension relations obtained from the energy calculations and by our e-WLC model (a two state equilibrium model as explained in Appendix A, which is essentially an e-WLC, but now it takes into consideration the extended contour length due to the conformational transition) are shown in figures 4.5 and 4.6.

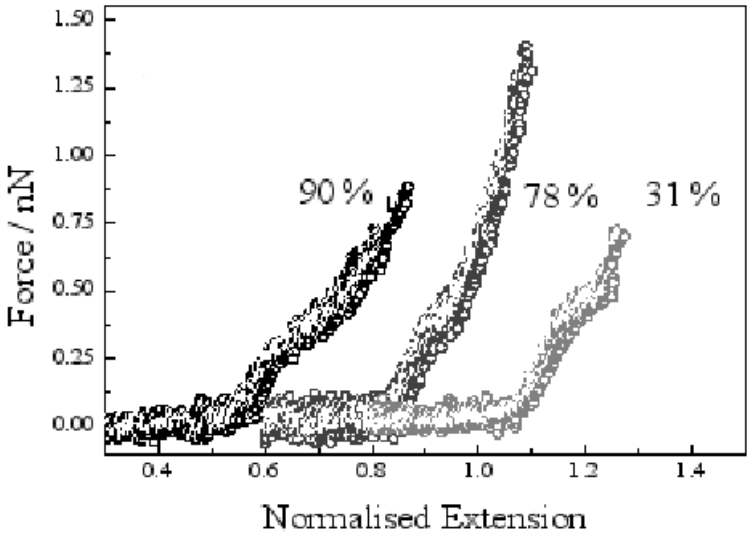

FIGURE 4.3: The experimental, normalized, concatinated force-extension curves for different DM pectins in water.

Despite the success of the model in accounting for the first "click", how to explain the force-extension curve at higher forces where a second click is observed experimentally is unclear in light of our energy calculations. While in fact there is significant general precedent for the relative stability of inverted chairs compared to boat structures, it is also significantly more stable in our calculations than the skew boat structure was at the smaller extension (Figure 4.2). This raises a difficult question of how it can account for the second transition observed in the AFM stretching experiments given that the predicted force required to extend the ring to the skew boat would be greater than that required to obtain a greater extension

ending in an inverted chair. The consequence of this would be that as soon as a ring is stretched enough to populate the skew boat state it would immediately reduce the force acting on it by elongating further and dropping into the inverted chair, which is significantly less strained regardless of its more extended conformation. This is clearly demonstrated by taking the three state equilibrium model [30] at face value and using the calculated DFT conformer lengths and energy differences to generate a force-extension curve (Figure 4.5).

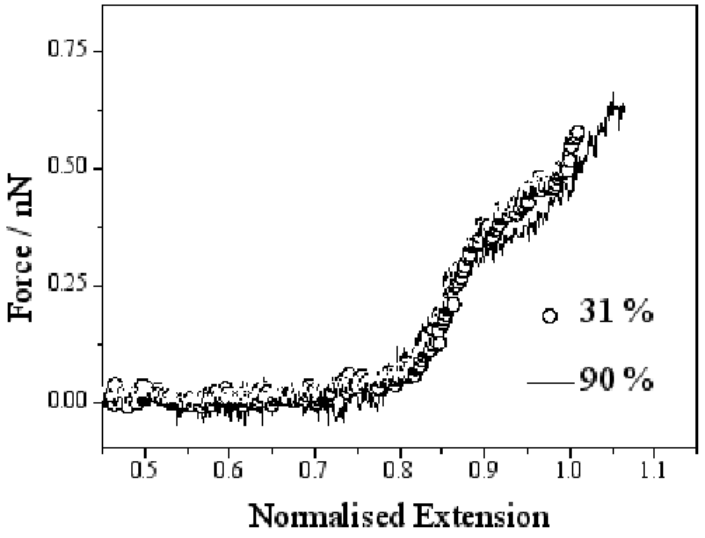

FIGURE 4.4: Interpolations of the concatenated normalized experimental force-extension datasets for 31 and 90 % DM sample pectins, showing a small (50-100 pN)force difference in the force activating the first conformational transition of the pyranose ring.

Evidently this results in an undistinguishable single click (centred at 77pN), very different from the experimental result, where only the chair and inverted chairs are significantly populated during the extension process. This suggests that either i) the inverted chair is significantly destabilised in the actual polymeric system as opposed to the monomer, owing to interaction effects neglected in the DFT, ii) the conformational states involved in the second transition are not in equilibrium

so that the transition barriers have a significant role to play or iii) that the second transition originates predominantly from a change in some other molecular feature or secondary structure. Improving our high force, large extension, measurements to clearly observe the second transition remains a primary part of ongoing work, so that its dependence on rate might be examined.

For the purposes of calculating what the effects of the differences obtained in the DFT calculations would be on the force-extension curves of completely methylesterified or unesterified pectin, the energy differences of the skew boats to inverted chair conformations has been assumed to be 25 kJ mol^{-1}, in line with the fitting of experimental data. The curves have then been calculated using the eWLC modified to take into account a threestate equilibrium, as described previously. A small but clear difference is observed in the force required for the chair to boat transformation depending whether the polymer is methylesterified or not (Figure 4.6). The model we have employed here is able to predict many features of the force-extension curves using values calculated from DFT rather well compared with previously reported attempts, including the values of the forces of the first click (500-700 pN simulated vs 400-500 pN experimental). At this stage we conclude that DFT calculations indicate that methylesterification does influence the conformational energy landscape of the monomers, albeit by a reasonably small amount, and therefore may result in a small change in the force-extension behaviour of polymers built from these monomers, with the conformational transitions occurring at slightly lower forces in the methylesterified form.

It shows that while the DM is critically important in the calcium binding properties and gelling of pectin and is known to be remodeled in-vivo, only subtle changes occur in the force-extension behavior of pectin as the DM is altered. It has been confirmed in previous works that the force-extension behavior of amylose and pectin depend on the nature of the solvent used as the dielectric constant of the solvent modulates the intramolecular H-bondings formed.

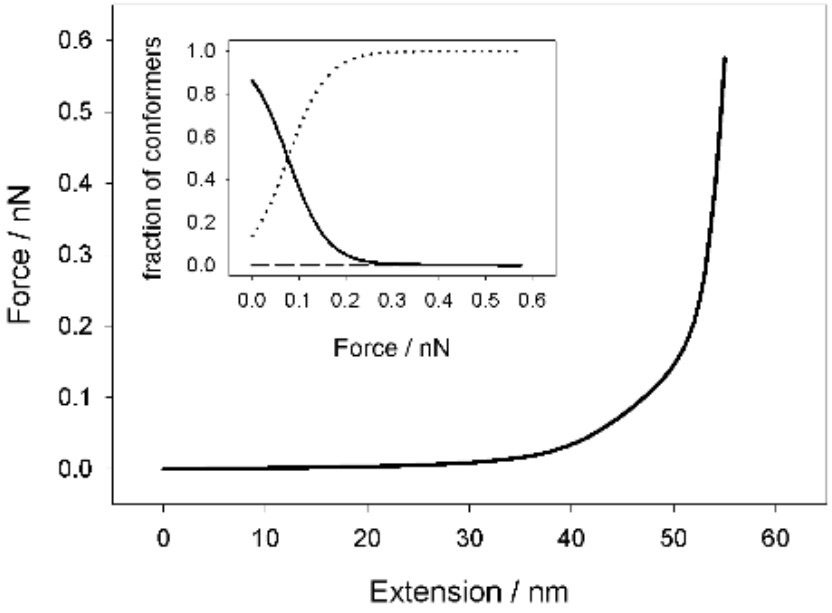

FIGURE 4.5: Simulated force-extension curve for polygalacturonic acid using parameters obtained by DFT calculations of the α-D-galacturonic acid anion using the B3LYP/6-311++G** basis set. Insert - The fraction of conformers as a function of applied force, solid line - chair, dashed line - boat, dotted line - inverted chair. Simulation parameters: l_p = 1nm, N_{total} = 100rings, chair length = 0.4592 nm, boat length = 0.5176 nm, inverted chair length = 0.5576 nm, Φ = 20 nN, ΔG_{01} = 25.4 kJmol^{-1}, ΔG_{02} = 20.85 kJmol^{-1}.

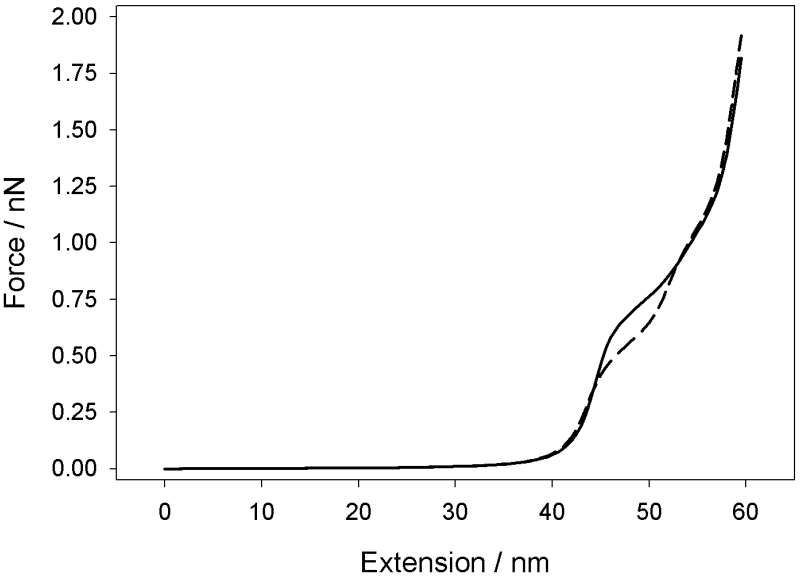

FIGURE 4.6: Simulated force-extension curves for polygalacturonic acid (solid line) and its methylesterified analogue (dashed line)

The ring hydroxyl group of amylose, when substituted by acetyl groups, results in a solvent independent force extension behavior [139]. This shows that H-bonding in the system can be modified (by substitution or the solvent quality) and its affect on the force-extension curve might be detected. Now, if the carboxyl side groups in pectin act as hydrogen bond donors, forming bonds with neighbouring ring oxygen atoms, then we might expect that by studying a highly methylesterified (DM=90%) sample, we would remove the ability for such bonds to exist and thereby be able to clearly differentiate substituted from unsubstituted samples by their single molecule stretching behaviour in low dielectric strength solvents,

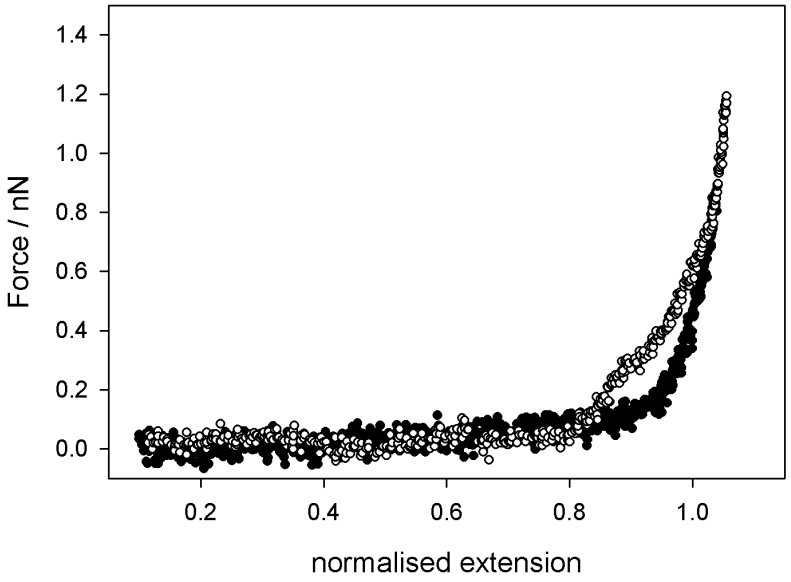

FIGURE 4.7: Force-extension curves of highly methylesterified pectin (DM=90 %) in water and hexadecane. The extension is normalized to the length at 1nN.

where inter ring hydrogen bonding is maximised. However, stretching highly methylesterified (DM=90%) pectin in water and hexadecane, essentially reproduced the result obtained from experiments that used a sample with a low degree of methylesterification, reported previously (Figure 4.7); that is, a significant strengthening of hydrogen bonds was found in hexadecane. As the largely methylesterified pectin sample still exhibited solvent dependent force-extension

behaviour, it is unlikely that differences in interring hydrogen bonding would be sufficient to modify the single molecule stretching behaviour between samples with differing degrees of methylesterification. It should be noted that the result obtained at low dielectric constant shows that there are strong hydrogen bonds in both samples, and by implication that, as the bonding is not changed by methylesterification, the carboxyl group acts as a H bond acceptor, not donor.

4.6 Summary

Using the results of DFT calculations (energies and sugar ring lengths of the conformers as a function of extension) allowed us to generate an extension-dependent contour length that when incorporated into a standard statistical mechanical model of chain behavior captured a good deal of the phenomenology of the single molecule stretching behavior of pectin. On extending galacturonic acid monomers the initial ($^4C_1$) chair structure is transformed to a ($^3S_5$) skew boat and upon further elongation, via an intermediate inverted skew boat ($^5S_3$), the inverted chair ($^1C_4$) is populated. The glycosidic oxygen distances of all forms, the energy difference of the first step and furthermore the prediction of the effects of methylesterification are all in reasonable agreement with experiment. However, the predicted stability of the inverted chair is problematic in terms of explaining the observed stretching behavior using an equilibrium model.

The force-extension curve of highly methylesterified pectin is solvent dependent in a similar vein as the unesterified sample, with interring hydrogen bonding a significant determinant of stretch behavior in solvents with low dielectric constants. In such hydrogen bonding the carboxyl oxygen acts as a proton acceptor, not donor as has previously been thought. As only subtle changes in the force-extension behavior of pectin result as the degree of methylesterification is modified, previous speculations about the role of force induced transformations in low DM polymers hold for more methylesterified analogues which are more likely to be the predominant fine structure in vivo.

Chapter 5

Conclusions

This work focused on computational approaches whose results were used to aid the understanding of (a) the characterisation of biopolymers, (b) the behavior of a polymer chains in different environments and in the presence of stiffness and torsional elasticity and (c) the mechanical properties of single chains of pectin.
The first intention was to validate an alternative means of characterisation of pectin by FT-IR techniques, using quantum chemical calculations. Secondly the folding pattern of polymer chains and how this is related to its environment, the stiffness and tension of the chain has been examined. Thirdly, the result of quantum chemical calculations on the structure and energetics of sugar ring conformations have been combined with a statistical mechanical model of chain stretching to determine the force-extension pattern exhibited by single pectin chains of differing DM. These theoretical studies (very well backed up by the experiments) succeeded in generating fruitful results in characterising spectroscopic properties of pectin, the pulling of single chain pectins of different DM's, and pH and also the relevance of stiffness, solvent and torsion in modelling secondary structures in stiff polymers.

5.1 The characterisation of biopolymers

The FT-IR method described in this study for the determination of the degree of methylesterification of pectins is non-invasive and is significantly more accurate than other reported IR methods; although the expected uncertainty in the measured DM value still stands at around 7%. In order to get the best results from this method the samples should be treated (precipitated from acidic solution and dried) as described. The ratio of the intensities of the asymmetric vibrations of the CH3 group and backbone bands can then be used to determine the DM (rather than the commonly used carboxylic bands), which permits the elimination of the interferences from other cell wall components such as water and proteins. In addition, DFT calculations have been shown to reproduce the main features of the experimental spectra, with a clear evolution towards the polymer result as the DP of the saccharide molecule in the calculation was increased from 1 to 2, and ultimately 3. This spectral elucidation makes FT-IR coupled with DFT an analysis method that has the potential to investigate pectin structure in more complex systems, such as post-coupling reactions. We persued this line and our work provides evidence that not only is pectin attached to these beads but it is via the formation of a specific recognizable bond and that pectin is coupled at its reducing end with the intended -CN bond.

5.2 The polymer in different environments

The compact packing of polymers has been a significant area of studies for past 3-4 decades. The folding pattern of a standard bead-spring model for a polymer chain has been investigated by Langevin dynamics simulations. In all cases, a bending potential takes care of the chain stiffness and a L-J term accounts for the excluded volume and the solvent effects; while our focus has been on the differences in the conformational phase space resulting from the inclusion or not of a torsional potential. Without such a torsional term the chains exhibit conformations previously observed by such coarse-grained models. However, the addition of a torsional term

into the energy expands the breadth of the conformational space; and the polymer now exhibits previously un-manifest marginally-compact structures that resemble well- known biological forms; secondary structures such as helices and sheets.

5.3 The mechanical properties of single pectins

The DFT calculations regarding the conformational space of individual sugar rings have been used in order to generate an extension-dependent contour length and incorporating this into standard statistical mechanical models of chain behavior captures a good deal of the phenomenology of the single molecule stretching behavior of pectin. On extending galacturonic acid monomers the initial ($^4C_1$) chair structure is transformed to a ($^3S_5$) skew boat and upon further elongation, via an intermediate inverted skew boat ($^5S_3$), the inverted chair ($^1C_4$) is populated. The glycosidic oxygen distances of all forms, the energy difference of the first step and furthermore the prediction of the effects of methylesterification are all in reasonable agreement with experiment. However, the predicted stability of the inverted chair is problematic in terms of explaining the observed stretching behavior using an equilibrium model. The force-extension curve of highly methylesterified pectin is solvent dependent in a similar vein as the unesterified sample, with interring hydrogen bonding a significant determinant of stretch behavior in solvents with low dielectric constants. In such hydrogen bonding the carboxyl oxygen acts as a proton acceptor, not donor, as was previously thought. As only subtle changes in the force-extension behavior of pectin result as the degree of methylesterification is modified, previous speculations about the role of force induced transformations in low DM polymers hold for more methylesterified analogues which are more likely to behave in this way invivo.

To conclude, our calculations were in good agreement with the experimental findings for IR characterisation of pectins, pectins coupled on surfaces and also in looking at the mechanical properties of pectin chain. Also we could successfully monitor the emergence of biologically relevant structures when a polymer chain is

allowed to minimise in different situations. Finally, this work could be extended in looking at the stretching properties of long chain molecules, which shows a "click" or deviation from the normal force-extension profile and also looking the properties of these polymers, when they form a network structure.

Chapter 6

Future Work.

6.1 Pulling Polymers Incorporating the Biopolymer Stretch Behavior - BDS Approach

In this thesis, we have successfully coupled the results of DFT techniques pertaining to the energy of different sugar ring conformations with a statistical mechanical model to describe the mechanical stretching of a pectin chain. It is well in accordance with the AFM stretching studies, but here we should be aware of the fact the chain parameters used in the e-WLC model such as the pectin monomer length scales and the free energies of different extended states come from single monomer calculations. Carrying out DFT simulations of larger carbohydrate molecules to study bigger pectin chains is simply computationally too expensive. People in the recent past have tried using alternative simulation techniques such as coarse grained simulations [95, 147, 148] and steered molecular dynamics (SMD) studies [130, 149, 150], in order to study the mechanical behaviour of these bulky systems and to reduce the expense of the calculations. Herein, we initially try simulating the pulling of a polymer chain in poor solvent and good solvent conditions and look at the force-extension behaviour of these with a view to investigating the possibility of incorporating the "clicking" .

6.1.1 Polymer Model & Results

The details of BDS simulations are as given in the Introduction and further elaborated in Chapter 3. A Lennard-Jones potential was used to account for the solvent-monomer and monomer-monomer interactions. The Fluctuation-dissipation theorem takes care of the random forces inside the system. Gaussian potentials are used to model the connection of monomers in the polymer chain. A time step of $\Delta t = 0.01$ was used for the simulation and the monomer-monomer bond length was fixed to be unity. The system was equilibriated for 10^6 time steps.

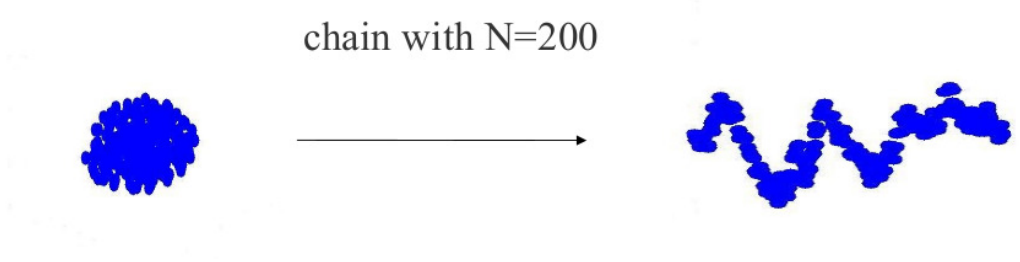

FIGURE 6.1: BDS of the Pulling of a polymer chain in poor solvent condition.

Firstly, a random chain configuration was allowed to collapse to a globule in poor solvent conditions. Pulling of the chain was carried out by fixing the Z-component of end-to-end distance of the chain at each interval and allowing the chain to equilibriate, as demonstrated by Cooke & Williams [148]. This is similar to what we do in real AFM experiments. Figure 6.1 represents the two states of a polymer chain, before stretching and after stretching, in poor solvent conditions. Three sets of simulations were carried out and figure 6.2 represents the averaged force (for the last 1000 iterations) profiles for these different runs. Force obtained from the derivative of energy also gives similar force profile (data not shown). It is clearly visible from the force-extension profile that

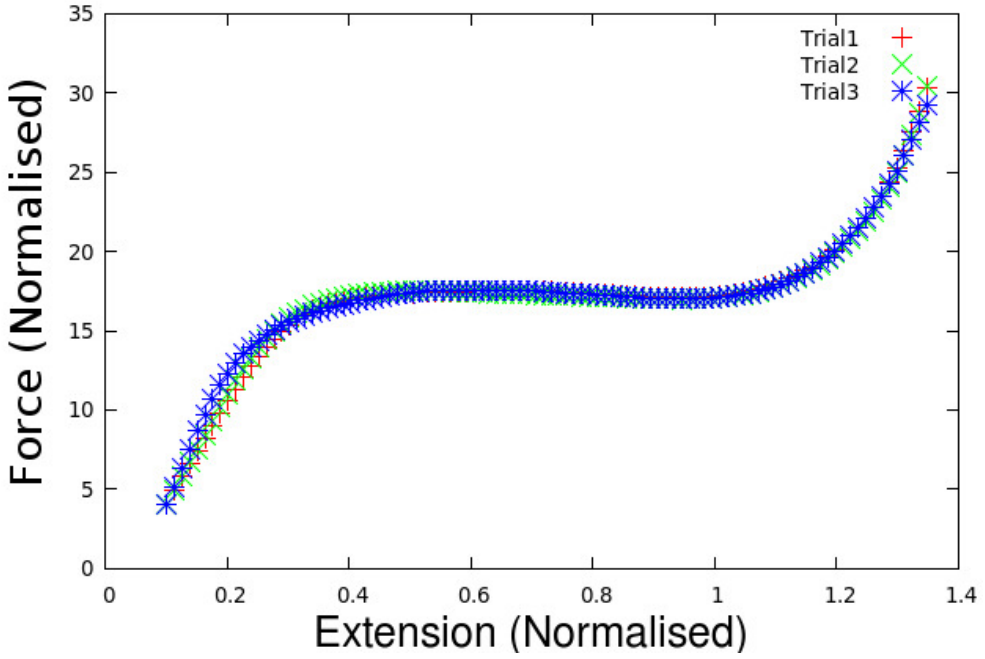

FIGURE 6.2: Force profile for a polymer chain pulled in a poor solvent condition.

that there exist a "phase separation" between the folded and the extended configurations, when being pulled in a poor environment. There is minimal extension until the applied force reaches some critical value, whereupon there is a sudden power-law increase in the extension. This is found to be in accordance with earlier proposed results by Williams et al [148] and Matthai et al [95] in their previous works.

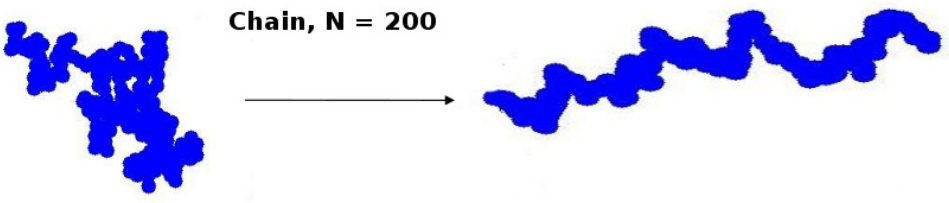

FIGURE 6.3: Pulling of a polymer chain in good solvent condition.

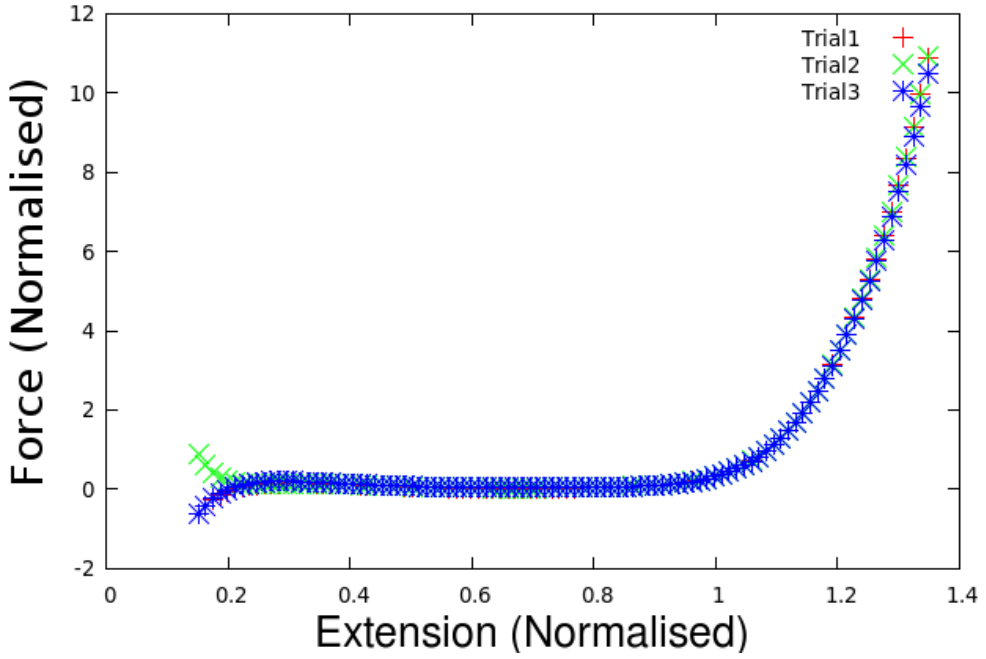

FIGURE 6.4: Force profile for a polymer chain pulled in a good solvent condition.

Since most biological processes take place at a variety of pHs and temperatures, studying a good solvent environment is also important. Also during the AFM stretching, most of the systems are studied in a convenient "solubilising" environment. Figure 6.3 shows the stretching of a polymer chain in a good solvent. In the case of a good solvent, the force-extension curve exhibits Hookes law type linear behavior for small applied forces and from the contour length, they exhibit a power law relation. Figure 6.4 shows the results of a BDS simulation of the force-extension profile for the stretching of a polymer chain in good solvent.

Gaining confidence from our code and the simulations successfully reproducing the previous works in the case of a standard polymer chain, lastly we try and simulate a biopolymer stretch model that incorporates a "clicking" type feature using by Langevin dynamics techniques. As in the previous case, the model incorporates a Lennard-Jones potential for internal interactions and the fluctuation-dissipation theorem takes care of the random forces in the system. Deviating from the above, we incorporate the analytical e-WLC model [151] itself to calculate the

bond stretching and bond bending forces in BD simulations. Thus, the modified bond force law for the springs in the bead-spring model is given by

$$F_{WLC} = \frac{k_B T}{l_p}[\frac{1}{4}(1 - \frac{l}{l_e} + \frac{F}{\phi})^{-2} + \frac{l}{l_e} - \frac{F}{\phi} - \frac{1}{4}]. \tag{6.1}$$

Where l_e is the extensible contour length as mentioned in Appendix 1. The corresponding bond potential is obtained by integrating the above expression and is given by

$$U_{WLC} = k_B T(\frac{l_e}{l_p})[\frac{1}{2}(\frac{l}{l_e})^2 - \frac{l}{l_e}(\frac{1}{4} + \frac{F}{\phi}) + \frac{1}{4}(\frac{1}{1 - \frac{l}{l_e} + \frac{F}{\phi}})]. \tag{6.2}$$

The polymer model simulation results are in good accordance with the statistical chain model. In poor solvent system, the chain shows an unravelling transition where as in a good solvent, the force-extension profile is comparable with FJC model. This has encouraged to extend this model to biopolymer stretching simulations.

Appendix A

CE Characterization of Pectin

Capillary electrophoresis (CE) is a separation technique used in the analysis of

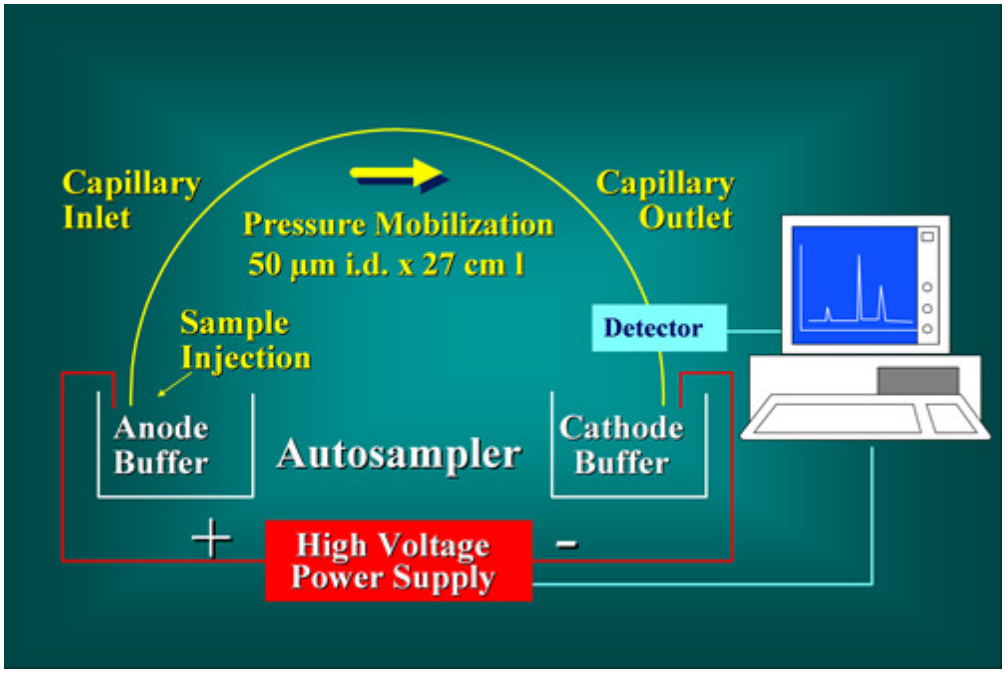

FIGURE A.1: Capillary Electrophoresis system.

charged molecules which are separated by the virtue of their charge and viscous forces. The separation is attained by the presence of a strong electric field. In the presence of an electric field, electroosmotic flow of the background electrolyte is

generated, and in addition charged analytes are also mobilised by the field. Figure A.1 (http://www.chnola-research.org/node/218) demonstrates the schematic working of a CE instrument. As shown in the figure, a CE instrument consist of a high voltage source, electrodes, vials(source, sample and destination), a detector and a data processor. Most CE instruments use UV/Visible spectroscopy for the detection purpose. The mobility $'\mu'$ of a molecule in a CE is given by

$$\mu = \mu_{obs} - \mu_{eo} = (\frac{IL}{V})(\frac{1}{t} - \frac{1}{t_0}). \tag{A.1}$$

and specifically for pectin the average charge per residue (z) and hence the DM can be calculated as

$$z = \frac{-(100 - DM)}{100} \tag{A.2}$$

Here in, 't' is the migration time of the molecule compared to the migration time of a neutral marker 't'_0. 'I' and 'L' are the distance from the inlet to the detector and total length of the capillary. $'\mu'_{obs}$ and $'\mu'_{eo}$ are respectively the observed mobility and electroosmotic mobility. 'DM' refers to the degree of methyl esterification of the sample used.

A.1 Pectin in CE

Unlike many other natural polysaccharides, pectin has both charge and a UV chromophore (the carboxyl group) which makes Capillary Electrophoresis (CE) a practical analytical tool for its study. The charge density of pectin depends on its degree of esterification.

The migration time of pectin in capillary electrophoresis (CE) is a function of the DM, and separation of pectins with different DM is thus obtained. Plotting mobility against the charge density, a linear relationship is found between pectin charge and its mobility in the range 30-80 % DM. Since the left hand sides of

the equations A.1 and A.2 are now in proportion, it is infered that the DM and the reduced mobility $(\frac{1}{t} - \frac{1}{t_0})$ will also possess a linear relationship. Hence on comparison with a standard sample of known DM, degree of methyl esterification of a new samples could be obtained.

One of the main advantages of the CE over other conventional methods lies in the fact that it is possible to access the intermolecular DM distribution by direct measurement. Also it is seen that the mobility of the sample is not effected by the intramolecular distribution of the methyl ester, thus giving only little difference between the random and blocky methyl esterified pectins. Also in the case of pectin with DM smaller than 30%, counter ion condensation is observed and hence it is a complicated procedure to characterise such molecules in CE.

Appendix B

Coupling Reactions on Pectin

The simple mechanism for the immobilisation of pectins on spherical surfaces such as the Polystyrene or Silica beads can be schematically represented as in figure B.1.

B.1 Bead amination

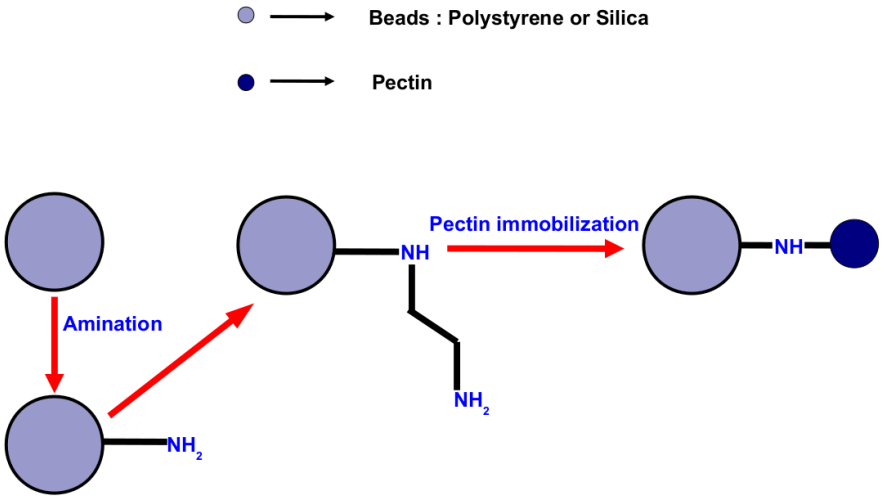

FIGURE B.1: Immobilisation of pectins on spherical surfaces.

266 mg of polystyrene beads, 500 nm, in a nitrating mixture of 0.532 mL of HNO_3 and 1.33 mL of H_2SO_4 on an oil bath, was stirred at T = 60°C for 30 min. The mixture was then poured over cold water in a refrigerated bath and washed several times with distilled water to remove excess acid. The polystyrene beads were then transferred to a mixture of 2.8 g of $SnCl_2$ to which 2.34 mL of HCl and 2.6 mL of ethanol were added, and refluxed for 10 h at T = 90 °C. Upon completion of the reaction, the aminated polystyrene beads were washed with distilled water and 2 M NaOH, removed from solution by centrifugation and dried overnight under reduced pressure.

B.2 Immobilization techniques

B.2.1 Immobilization via Reductive Amination

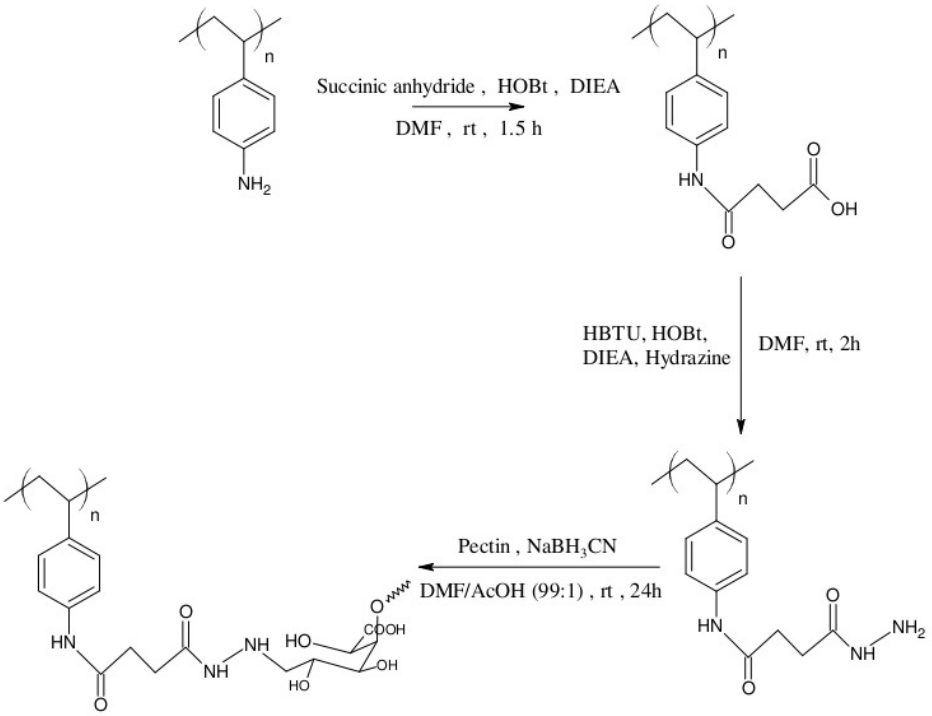

FIGURE B.2: Mechanism for Reductive Amination.

Dried amino-terminated beads (50mg) were washed with DMF. Succinic anhydride (320 mg, 3.2 mmol) and HOBt (320 mg, 2.4 mmol) were dissolved in DMF (15 mL). DIEA (600 μL, 3.6mmol) was added and the solution was added to the beads. The mixture was then stirred at room temperature for 1.5 h. Excess reagents and solvents were filtered off and the beads were washed with DMF, centrifuged and dried overnight under reduced pressure.

The dried carboxylic acid-terminated beads were transferred and washed with DMF. HBTU (200 mg, 0.5 mmol) and HOBt (360 mg, 2.7 mmol) were dissolved in DMF (15 mL). DIEA (540 μL, 3.24 mmol) was then added, followed by hydrazine monohydrate (180 μL, 3.6 mmol) and the mixture was added to the beads. After 2 hour coupling at room temperature, solvents and unreacted compounds were removed by filtration and the beads were washed with DMF, centrifuged and dried overnight under reduced pressure.

Dried hydrazide-terminated beads were transferred and washed with DMF. Pectin was dissolved in DMF-AcOH (99:1), 0.5% w/w. The solution (3mL) was then added to the beads and following addition of $NaBH_3CN$ (30 mg, 0.48 mmol), the mixture was then stirred at room temperature. After 24 hour, the solvent and excess reagents were filtered off and the beads were washed with DMF, centrifuged and dried overnight under reduced pressure.

B.2.2 Immobilization via Thiazolidine Formation

Dried amino-terminated beads (15 mg) were washed with DMF. Boc-Cys(trt)-OH (100 mg, 0.22 mmol) was coupled in presence of PyBOP (80 mg, 0.15 mmol), HOBt (15 mg, 0.12 mmol) and DIEA (30 μL, 0.18 mmol) in DMF (5 mL) at room temperature. After 1 hour, the beads were washed with DMF, centrifuged and dried overnight under reduced pressure. Deprotection of Boc and trt groups was accomplished with TFA-CH_2Cl_2 (1:1) (5 mL) in the presence of Et_3SiH (50 μL, 0.3 mmol) for 1 hour at room temperature. The beads were finally washed with CH_3CN, centrifuged and dried overnight under reduced pressure.

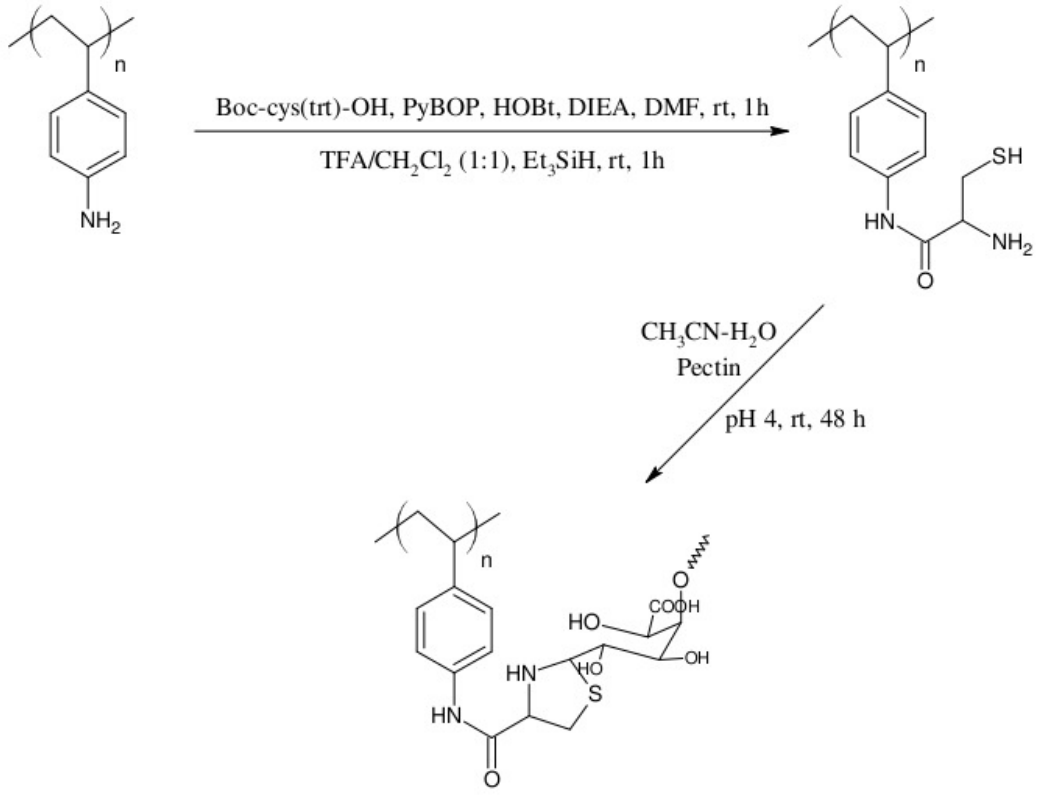

FIGURE B.3: Mechanism for Thizolidine Formation.

Dried cys-terminated beads were transferred and washed with CH_3CN. Pectin was dissolved into CH_3CN-H_2O (1:2), 0.5% w/w, and the solution, whose pH was decreased to 4 with 0.2 M HCl, was added to the beads. After 48 hours at room temperature, the beads were washed with CH_3CN, centrifuged and dried overnight under reduced pressure.

Bibliography

[1] J.M.G. Cowie. *Polymers: Chemistry and Physics of Modern Materials.* Blackie, 2nd edition, 1991.

[2] R.W. Forsey, J. Fisher, J. Thompson, M.H. Stone, C. Bell, and E. Ingham. The effect of hyaluronic acid and phospholipid based lubricants on friction within a human cartilage damage model. *Biomaterials*, 27(26):4581–4590, 2006.

[3] M.L. Fishmann and J.J. Jen. *Chemistry and Function of Pectins.* ACS, Washington, 1986.

[4] P.A. Levene and L.C. Kreider. On the structure of pectin polygalacturonic acid. *Science*, 85:610, 1937.

[5] F. Goubet, A. Strom, P. Dupree, and M.A.K. Williams. An investigation of pectin methylesterification patterns by two independent methods: capillary electrophoresis and polysaccharide analysis using carbohydrate gel electrophoresis. *Carbohydrate Research*, 340(6):1193–1199, 2005.

[6] E.L. Pippen, T.H. Schultz, and H.S. Owens. Effect of degree of esterification on viscosity and gelation behaviour of pectin. *Journal Of Colloid Science*, 8:97–104, 1953.

[7] G.J.W.M. van Alebeek, K. van Scherpenzeel, G. Beldman, H.A. Schols, and A.G.J. Voragen. Partially esterified oligogalacturonides are the preferred substrates for pectin methylesterase of aspergillus niger. *Biochemical Journal*, 372:211–218, 2003.

[8] R.K. Mishra, P. Singhal, M. Datt, and A.K. Banthia. Amidated pectin based hydrogels: synthesis, characterization and cytocompatibility study. *Journal of Applied Biomaterials & Biomechanics*, 5:88–94, 2007.

[9] E.A. Mccomb and R.M. Mccready. Determination of acetyl in pectin and in acetylated carbohydrate polymers: - hydroxamic acid reaction. *Analytical Chemistry*, 29:819–821, 1957.

[10] M. Doi and H. See. *Introduction to Polymer Physics*. Clarendon Press, Oxford., 1996.

[11] M. Rubinstein and R.H. Colby. *Polymer Physics*. Oxford University Press, 2003.

[12] A. Einstein. On the movement of small particles suspended in stationary liquids required by the molecular-kinetic theory of heat. *Annalen der Physik*, 17:549560, 1905.

[13] K. Pearson. The problem of the random walk. *Nature*, 72:294, 1905.

[14] K. Pearson and J. Blakeman. *A mathematical theory of random migration*. Dulau and co, 1906.

[15] M. Doi and S.F. Edwards. *The Theory of Polymer Dynamics*. Oxford University Press, New York., 1986.

[16] W. Kuhn. Uber die kinetik des abbaues hochmolekularer ketten. *Berchite*, 63:1503, 1930.

[17] W. Kuhn. Uber die gestalt fadenfrmiger molekle in lsungen. *Kolloid, Z.*, 68: 2–15, 1934.

[18] O. Kratky and G. Porod. Rntgenuntersuchung gelster fadenmolekle. *Rec. Trav. Chim. Pays-Bas*, 68:1106, 1949.

[19] P.J. Flory. *Statistical Mechanics of Chain Molecules*. Interscience Publishers, 1969.

[20] P.G. deGennes. *Scaling Concepts in Polymer Physics.* Cornell University Press, Ithaca., 1979.

[21] C. Bustamante, S.B. Smith, J. Liphardt, and D. Smith. Single-molecule studies of dna mechanics. *Current Opinion in Structural Biology*, 10:279–285, 2000.

[22] D. Sarid. *Scanning Force Microscopy.* Oxford University Press, New York., 1991.

[23] A. Ashkin, J.M. Dziedzic, J.E. Bjorkholm, and S. Chu. Observation of a single-beam gradient force optical trap for dielectric particles. *Optics Letters*, 11:288, 1986.

[24] C. Bustamante, J.F. Marko, E.D. Siggia, and S. Smith. Entropic elasticity of a-phage dna. *Science*, 265:1599, 1994.

[25] M. Reif, F. Oesterhelt, B. Heymann, and H.E. Gaub. Single molecule force spectroscopy on polysaccharides by atomic force microscopy. *Science*, 275: 1295, 1997.

[26] M.S.Z. Kellermayer, S.B. Smith, H.L. Granzier, and C. Bustamante. Folding-unfording transitions in single titin molecules characterized with laser tweezers. *Science*, 276:1112, 1997.

[27] A.L. Buchachenko. New horizons of chemistry : single molecules. *Russian Chemical Reviews*, 75(1):122, 2006.

[28] M.D. Wang, H. Yin, R. Landick, J. Gelles, and S.M. Block. Stretching dna with optical tweezers. *Biophysical Journal*, 72:1335–1346, 1997.

[29] M. Rief, M. Gautel, F. Oesterhelt, J.M. Fernandez, and H.E. Gaub. Reversible unfolding of individual titin immunoglobulin domains by afm. *Science*, 276:1109, 1997.

[30] R.G. Haverkamp, A.T. Marshall, and M.A.K. Williams. A model for stretching elastic biopolymers which exhibit conformational transformations. *Physical Review E*, 75:021907, 2007.

[31] V.S.R. Rao, P.K. Qasba, P.V. Balaji, and R. Chandrasekaran. *Conformation of Carbohydrates*. Harwood Academic Publishers, 1998.

[32] P.E. Marszalek, H. Li, F.A. Oberhauser, and J.M. Fernandez. Chair-boat transitions in single polysaccharide molecules observed with force-ramp afm. *Proceedings of National Academy of Sciences*, 99:4278–4283, 2002.

[33] P.E. Marszalek, Y.P. Pang, H. Li, F.A. Oberhauser, J.E. Yazal, and J.M. Fernandez. Atomic levers control pyranose ring conformations. *Proceedings of National Academy of Sciences*, 96:7894–7898, 1999.

[34] Q. Zhang, G. Lee, and P.E. Marszalek. Atomic cranks and levers control sugar ring conformations. *Journal of Physics: Condensed Matter*, 17:S1427, 2005.

[35] E. Lewars. *Computational Chemistry - Introduction to the Theory and Application of Molecular and Quantum Mechanics*. Kluwer Academic Publications, 2003.

[36] J.M. Haile. *Molecular Dynamics Simulations - Elementary Methods*. John Wiley & Sons, INC, 1992.

[37] E. Schrdinger. An undulatory theory of the mechanics of atoms and molecules. *Physical Review*, 28 (6):10491070, 1926.

[38] I.N. Levine. *Quantum Chemistry*. Pearson Prentice Hall, N.J., 2009.

[39] I.N. Levine. *Quantum Chemistry*. Prentice Hall, New jersey., 1991.

[40] C. M∅ller and M.S. Plesset. Note on an approximation treatment for many-electron systems. *Physical Review*, 46:618622, 1934.

[41] D.T. Crawford and H.F. Schaefer. *An Introduction to Coupled Cluster Theory for Computational Chemists, Reviews in Computational Chemistry*. Wiley-VCH, 1998, Vol. 14.

[42] M.M. Goodgame and W.A. Goddard. Modified generalized valence-bond method: A simple correction for the electron correlation missing in generalized valence-bond wave functions; prediction of double-well states for cr2 and mo2. *Physical Review Letters*, 54 (7):661664, 1985.

[43] R. McWeeny. *Coulson's Valence*. Oxford University Press, 1979.

[44] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Physical Review B*, 136:864871, 1964.

[45] W. Kohn and L.J. Sham. Self-consistent equations including exchange and correlation effects. *Physical Review A*, 140:1133–1138, 1965.

[46] M.P. Allen and D.J. Tildesley. *Computer simulation of liquids*. Clarendon Press, Oxford, 1987.

[47] I. Newton. *Philosophi Naturalis Principia Mathematica*. 1687.

[48] F. Reif. *Fundamentals of Statistical and Thermal Physics*. McGraw Hill, New York, 1965.

[49] J. Weber. Fluctuation dissipation theorem. *Physical Review*, 101:16201626, 1956.

[50] J.E. Jones. On the determination of molecular fields. i. from the variation of the viscosity of a gas with temperature. *Proceedings of Royal Society*, 106: 441–462, 1924.

[51] J.E. Jones. On the determination of molecular fields. ii. from the equation of state of a gas. *Proceedings of Royal Society*, 106:463–477, 1924.

[52] H.R. Warner. Kinetic theory and rheology of dilute suspensions of finitely extendible dumbbells. *Industrial and Engineering Chemistry Fundamentals*, 11:379387, 1972.

[53] W.C. Swope, H.C. Andersen, P.H. Berens, and K.R. Wilson. Computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *Journal of Chemical Physics*, 76(1):637649, 1982.

[54] T.H. Schulz. *Methods in Carbohydrate Chemistry.* Academic Press, 1965.

[55] L. Bdouet, B. Courtois, and J. Courtois. Rapid quantification of o-acetyl and o-methyl residiues in pectin extracts. *Carbohydrate Research*, 338 (4): 379–383, 2003.

[56] M.M.H. Huisman, A Oosterveld, and H.A. Schols. Fast determination of the degree of methylesterification of pectins by head-space gc. *Food Hydrocolloids,*, 18 (4):665–668, 2004.

[57] N.O. Maness, J.D. Ryan, and A.J. Mort. Determination of the degree of methylesterification of pectins in small samples by selective reduction of esterified galacturonic acid to galactose. *Analytical Biochemistry,*, 185 (2): 346–352, 1990.

[58] P. Massiot, V. Perron, A. Baron, and J.F. Drilleau. Release of methanol and depolymerisation of highly methylesterified apple pectin with an endopolygalacturonase from aspergillus niger and pectin methylesterases from a niger or from orange. *Food Science and Technology-Lebensmittel-Wissenschaft & Technologie*, 30 (7):697–702, 1997.

[59] C. Rosenbohm, I. Lundt, T.M.I.E. Christensen, and N.W.G. Young. Chemically methylated and reduced pectins: preparation, characterization by 1h nmr spectroscopy, enzymatic degradation and gelling properties. *Carbohydrate Research*, 338:637–649, 2003.

[60] A.G.J. Voragen, H.A. Schols, and W. Pilnok. Determination of the degree of methylation and acetylation of pectins by hplc. *Food Hydrocolloid*, 1:65–70, 1986.

[61] A. Strom, M.C. Ralet, J.F. Thibault, and M.A.K. Williams. Capillary electrophoresis of homogeneous pectin fractions. *Carbohydrate Polymers*, 60: 467–473, 2005.

[62] A. Strom and M.A.K. Williams. On the separation, detection and quantification of pectin derived oligosaccharides by capillary electrophoresis. *Carbohydrate Research*, 339:1711–1716, 2004.

[63] H.J. Zhong, M.A.K. Williams, D.M. Goodall, and Hansen M.E. Capillary electrophoresis studies of pectins. *Carbohydrate Research*, 308:1–8, 1998.

[64] H.J. Zhong, M.A.K. Williams, R.D. Keenan, D.M. Goodall, and C. Rolin. Separation and quantification of pectins using capillary electrophoresis : a preliminary study. *Carbohydrate Polymers*, 32:27–32, 1996.

[65] A.S. Barros, I. Mafra, D. Ferreira, S. Cardoso, A. Reis, J.A. Lopes da Silva, I. Delgadillo, D.N. Rutledge, and Coimbra M.A. Determination of the degree of methylesterification of pectic polysaccharides by ft-ir using an outer product pls1 regression. *Carbohydrate Polymers*, 50:85–94, 2002.

[66] U. Haas and M. Jager. Degree of esterification of pectins determined by photoacoustic near infrared spectroscopy. *Journal of Food Science*, 51:1087–1088, 1986.

[67] M.A. Monsoor, U. Kalapathy, and A. Proctor. Improved method for determination of pectin degree of esterification by diffuse reflectance fourier transform infrared spectroscopy. *Journal of Agricultural and Food chemistry*, 49:2756–2760, 2001.

[68] M.V Korolevich and R.G. Zhbankov. Theoretical study of the influence of structure on the vibrational spectra of crystalline nitrates of substituted glucopyranosides. *Journal of Molecular Structure*, 474:187–196, 1999.

[69] M.V Korolevich and R.G. Zhbankov. A theoretical study of vibrational spectra of tetranitrate-methyl--dglucopyranoside and 2,3-di-o-nitro-methyl--d-glucopyranoside. *Journal of Molecular Structure*, 440:247–263, 1998.

[70] V. Krishnakumar and N. Prabhavathi. Simulation of ir and raman spectral based on scaled dft force fields: A case study of 2-amino 4-hydroxy 6-trifluoromethylpyrimidine, with emphesis on band assignment. *Spectrochemica Acta Part A,*, 71:449–457, 2008.

[71] V. Krishnakumar, G. Keresztury, T. Sundius, and R. Ramasamy. Simulation of ir and raman spectral based on scaled dft force fields: A case study of 2-(methylthio)benzonitrile, with emphesis on band assignment. *Spectrochemica Acta Part A*, 71:449–457, 2004.

[72] J.A. Pople and M.S. Gordon. Molecular orbital theory of the electronic structure of organic compounds. i. substituent effects and dipole moments. *Journal of the American Chemical Society*, 89:4253–4261, 1967.

[73] M. Appell, G. Strati, J.L. Willett, and F.A. Momany. B3lyp/6-311++g**study of α- and β-d-glucopyranose and 1,5-anhydro-d-glucitol: $^{4}c_{1}$ and $^{1}c_{4}$ chairs, $^{3,0}b$ and $b_{3,0}$ boats, and skew boat conformations. *Carbohydrate Research*, 339:537–551, 2004.

[74] F.A. Momany and J.L. Willett. Computational studies on carbohydrates: I. density functional ab-initio geometry optimisation on maltose conformations. *Journal of Computational Chemistry*, 21:1204–1219, 2000.

[75] F.A. Momany, M. Appell, G. Strati, and J.L. Willett. B3-lyp/6-311++g** study of α- and β-d-glucopyranose hydrogen bonding, stress energies, and effect of hydration on internal coordinates. *Carbohydrate Research*, 339: 553–567, 2004.

[76] M.A.K. Williams, A.T. Marshall, P. Anjukandi, and R.G. Haverkamp. Investigation of the effects of fine structure on the nanomechanical properties of pectin. *Physical Review E*, 76:021927, 2007.

[77] G.A. Zhurko. Chemcraft. Technical report, http://www.chemcraftprog.com.

[78] R.R. Vincent, A. Cucheval, Y. Hemar, and M.A.K. Williams. Bio-inspired network optimization in soft materials- insights from the plant cell wall. *The European Physical Journal E*, 28:79–87, 2009.

[79] R.R. Vincent and M.A.K. Williams. Microrheological investigations give insights into the structure and functionality of pectin gels. *Carbohydrate Research*, 344:1863–1871, 2009.

[80] M.A.K. Williams, T.J. Foster, and H.A. Schols. Elucidation of pectin methylester distributions by capillary electrophoresis. *Journal of Agricultural and Food Chemistry*, 51(7):1777–1782, 2003.

[81] A Strom. *Characterization of Pectin Fine Structure and its Effect on Supramolecular Properties*. PhD Thesis, University of Cork., 2006.

[82] S.M. Bociek and D. Welti. The quantitative analysis of uronic acid polymers by infrared spectroscopy. *Carbohydrate Research*, 42:217–226, 1975.

[83] S.Y. Venyaminov and N.N. Kalnin. Quantitative ir spectroscopy of peptide compounds in water solutions. 1. spectral parameters of amino acid absorption bands. *Biopolymers*, 30:1243–1257, 1990.

[84] A.K. Chatjigakis, C. Pappas, N. Proxenia, O. Kalantzi, P. Rodis, and M. Polissiou. Ft-ir spectroscopic determination of the degree of esterification of cell wall pectins from stored peaches and correlation to textural changes. *Carbohydrate Polymers*, 37:395–408, 1998.

[85] A. Synytsya, J. opkov, P. Matjka, and V. Machovi. Fourier transform raman and infrared spectroscopy of pectins. *Carbohydrate Polymers*, 54:97–106, 2003.

[86] A. Fellah, P. Anjukandi, M.R. Waterland, and M.A.K Williams. Determining the degree of methylesterification of pectin by atr/ft-ir: Methodology optimisation and comparison with theoretical calculations. *Carbohydrate Polymers*, 78:847–853, 2009.

[87] Y.L. Sun, Z.P. Luo, and K.N. An. Stretching short biopolymers using optical tweezers. *Biochemical and Biophysical Research Communications*, 286:826–830, 2001.

[88] B.P.B. Downing, A.V.D. Horst, M. Miao, F.W. Keeley, and N.R. Forde. Probing the elasticity of short proteins with optical tweezers. *Biophotonics Applications*, OTuA3, 2009.

[89] K. A. Walther, J. Bruji, H. Li, and J.M. Fernndez. Sub-angstrom conformational changes of a single molecule captured by afm variance analysis. *Biophysical Journal*, 90:3806–3812, 2006.

[90] M.-C. Desjonqueres and D. Spanjaard. *Concepts in Surface Physics*. Springer-Verlag, 1993.

[91] F. Guillaumie, O.R.T. Thomas, and K.J. Jensen. Immobilization of pectin fragments on solid supports: Novel coupling by thiazolidine formation. *Bioconjugate Chemistry*, 13:285294, 2002.

[92] P. Anjukandi, A. Fellah, P. Patty, and M.A.K Williams. Immobilisation of pectin on solid surfaces, its characterisation and theoretical calculations. Manuscript in preperation.

[93] T. Hassinen and M. Peräkylä. New energy terms for reduced protein models implemented in an off-lattice force field. *Journal of Computational Chemistry*, 22:1229–1242, 2001.

[94] J.G.H. Cifre, R. Pamies, A.-L. Kjniksen, K.D. Knudsen, B. Nystrm, and J.G. de la Torre. Brownian dynamics simulation of reversible polymer networks using a non-interacting bead-and-spring chain model. *Journal of Non-Newtonian Fluid Mechanics*, 146:3–10, 2007.

[95] R.G. Maurice and C.C. Matthai. Force-extension curves for a single polymer chain under varying solvent conditions. *Physical Review E*, 60:3165–3169, 1999.

[96] J.R. Banavar and A. Maritan. Computational approach to protein folding problem. *PROTEINS: Structure, Function and Genetics*, 42:433–435, 2007.

[97] I.C.B. Miller, M. Keentok, G.G. Pereira, and D.R.M. Williams. Semiflexible polymer condensates in poor solvents: Toroid versus spherical geometries. *Physical Review E*, 71:031802–031810, 2005.

[98] G.G. Pereira and D.R.M. Williams. Kinetic pathways for the collapse of semiflexible polymers. *ANZIAM Journal*, 45(E):c163–c173, 2004.

[99] B. Schnurr, F.C. MacKintosh, and D.R.M. Williams. Dynamical intermediates in the collapse of semiflexible polymers in poor solvents. *Europhysics Letters*, 51(3):279–285, 2000.

[100] L. Degrve and A. Cal iri. Geometric constraints in polymer chains: analysis on the pearl-necklace model by monte carlo simulation. *Journal of Molecular Structure: THEOCHEM*, 335:123–127, 1995.

[101] T.E. Creighton. *PROTEINS - Structure and Molecular properties*. 2nd edn. Freeman. New York., 1999.

[102] J.D. Watson and F.H.C. Crick. A structure for deoxyribose nuclic acid. *Nature*, 3:737–738, 1953.

[103] Q. Zhang, Z. Lu, L. Hu, W. Yang, and P.E. Marszalek. Direct detection of the formation of v-amylose helix by single molecule force spectroscopy. *Journal Of American Chemical Society*, 128(29):9387–9393, 2006.

[104] L. Zhang, C. Wang, S. Cui, Z. Wang, and X. Zhang. Single molecule force spectroscopy on curdlan: unwinding helical structures and random coils. *Nano Letters*, 3:1119–1124, 2003.

[105] I. Hargittai and M. Hargittai. Molecular structure of hyaluronan: an introduction. *Structural Chemistry*, 19:697–717, 2008.

[106] E. Wright, P and J. Dyson, H. Intrinsically unstructured proteins: reassessing the protein structure-function paradigm. *Journal of molecular biology*, 293(2):321–31, 1999.

[107] S. Kumar, I. Jenson, J.L. Jacobsen, and A.J. Guttmann. Role of conformational entropy in force-induced biopolymer unfolding. *Physical Review Letters*, 98:128101(1)–128101(4), 2007.

[108] V.S. Pande and D.S. Rokhsar. Folding pathway of a lattice model for proteins. *Proceedings of National Academy of Sciences*, 96:1273–1278, 1999.

[109] J.D. Bryngelson and P.G. Wolynes. Spin glasses and the statistical mechanics of protein folding. *Proceedings of National Academy of Sciences*, 84::7524–7528., 1987.

[110] J.R. Banavar and A. Maritan. Physics of proteins. *Annual Review on Biophysics and Biomolecular Structure*, pages 261–280., 2007.

[111] S.A. Edwards and D.R.M. Williams. Stretching a single diblock copolymer in a selective solvent: Langevin dynamics simulations. *Macromolecules*, 38: 10590–10595, 2005.

[112] C. Liu and M. Muthukumar. Langevin dynamics simulations of early-stage polymer nucleation. *Journal of Chemical Physics*, 109:2536–2542, 1998.

[113] P. Welch and M. Muthukumar. Molecular mechanisms of polymer crystallization from solution. *Physical Review Letters*, 87:218302(1)–218302(4), 2001.

[114] J.P. Kemp and Z.Y. Chen. Formation of helical states in wormlike polymer chains. *Physical Review Letters*, 81:3880–3883, 1998.

[115] G.G. Pereira and D.R.M. Williams. Crystalline toroidal globules of dna and other semi-flexible polymers: Jumps in radius caused by hexagonal packing. *Europhysics Letters*, 50(4):559–564, 2000.

[116] M. Sletmoen, G. Maurstad, and B.T. Stokke. Potentials of bionanotechnology in the study and manufacturing of self-assembled biopolymer complexes and gels. *Food Hydrocolloids*, 22:2–11, 2008.

[117] J.R. Banavar and A. Maritan. Colloquium: Geometrical approach to protein folding: a tube picture. *Reviews of Mordern Physics*, 75:23–34, 2003.

[118] A. Maritan, C. Micheletti, A. Trovato, and J.R. Banavar. Optimal shapes of compact strings. *Nature*, 406:287–290, 2000.

[119] T. Vogel, T. Neuhaus, M. Bachmann, and W. Janke. Thickness-dependent secondary structure formation of tubelike polymers. *Europhysics Letters*, 85: 10003(p1)–10003(p5), 2009.

[120] J N. BeMiller. *An Introduction to Pectins: Structure and Properties.* ACS Symposium Series, American Chemical Society, 1986.

[121] H.J. Butt, B. Cappella, and M. Kappl. Force measurements with the atomic force microscope: Technique, interpretation and applications. *Surface Science Reports*, 59:1–152, 2005.

[122] S.A. Harris. The physics of dna stretching. *Contemporary Physics*, 45:11–30, 2004.

[123] A. Janshoff, M. Neitzert, Y. Oberdrfer, and Fuchs. H. Force spectroscopy of molecular systems - single molecule spectroscopy of polymers and biomolecules. *Angewandte Chemie*, 39:3212 – 3237, 2000.

[124] T.E. Fisher, P.E Marszalek, A.F. Oberhauser, M. CarrionVazquez, and J.M. Fernandez. The micro-mechanics of single molecules studied with atomic force microscopy. *Journal of Physiology*, 50:5, 1999.

[125] B.L. Smith, T.E. Schffer, M. Viani, J. B. Thompson, N.A. Fredrick, J. Kindt, A. Belcher, G.D. Stucky, D.E. Morse, and P.K. Hansma. Molecular mechanistic origin of the toughness of natural adhesives, fibres and composites. *Nature*, 399:761, 1999.

[126] P.E. Marszalek, A.F. Oberhauser, Y.P. Pang, and J.M. Fernandez. Polysaccharide elasticity governed by chair-boat transitions of the glucopyranose ring. *Nature*, 396:661664, 1998.

[127] P.E. Marszalek, A.F. Oberhauser, H. Li, and J.M. Fernandez. The force-driven conformations of heparin studied with single molecule force spectroscopy. *Biophysical Journal*, 85(4):2696–2704, 2003.

[128] R.G. Haverkamp, M.A.K. Williams, and J.E. Scott. Stretching single molecules of connective tissue glycans to characterise their shape-maintaining elasticity. *Biomacromolecules*, 6:1816, 2005.

[129] M. Kuttel and K.J. Naidoo. Glycosidic linkage rotations determine amylose stretching mechanism. *Journal Of American Chemical Society*, 127:12, 2005.

[130] I. Neelov, D. Adolf, M. Ratner, O. Zhicol, and T. McLeish. Molecular dynamics simulation of dextran extension at constant pulling speed. *Macromolecular Symposia*, 237:81, 2006.

[131] P.E. Marszalek, H. Li, and J.M. Fernandez. Fingerprinting polysaccharides with single molecule atomic force microscopy. *Nature Biotechnology*, 19:258, 2001.

[132] S.C. Fry. *The Growing Plant Cell Wall: Chemical and Metabolic Analysis.* The Blackburn: Caldwell, 2000.

[133] W.G.T. Willats, P. Knox, and J.D. Mikkelson. Pectin: new insights into an old polymer are starting to gel. *Trends in Food Science and Technology*, 17: 97104, 2006.

[134] E.R. Morris, D.A. Powell, M.J. Gidley, and D.A. Rees. Conformations and interactions of pectins. i. polymorphism between gel and solid states of calcium polygalacturonate. *Journal of Molecular Biology*, 155:507, 1982.

[135] D.A. Powell, E.R. Morris, M.J. Gidley, and D.A. Rees. Conformations and interactions of pectins : Ii. influence of residue sequence on chain association in calcium pectate gels. *Journal of Molecular Biology*, 155:517, 1982.

[136] A. Strom, P. Ribelles, L. Lundin, I.T. Norton, E.R. Morris, and M.A.K. Williams. Influence of pectin fine structure on the mechanical properties of calciumpectin and acidpectin gels. *Biomacromolecules*, 8:2668–2674, 2007.

[137] B. Heymann and H. Grubmller. chairboat' transitions and side groups affect the stiffness of polysaccharides. *Chemical Physics Letters*, 305:202, 1999.

[138] Q. Zhang, J. Jaroniec, G. Lee, and P.E. Marszalek. Direct detection of inter-residue hydrogen bonds in polysaccharides by single-molecule force spectroscopy. *Angewandte Chemie*, 44:2723, 2005.

[139] Q. Zhang and P.E. Marszalek. Solvent effects on the elasticity of polysaccharide molecules in disordered and ordered states by single-molecule force spectroscopy. *Polymer*, 47:2526, 2006.

[140] F. Khner, M. Erdmann, and H. Gaub. Scaling exponent and kuhn length of pinned polymers by single molecule force spectroscopy. *Physical Review Letters*, 97:218301, 2006.

[141] B.S. Khatri, M. Kawakami, K. Byrne, D.A. Smith, and T.C.B. McLeish. Entropy and barrier-controlled fluctuations determine conformational viscoelasticity of single biomolecules. *Biophysical Journal*, 92:1825, 2007.

[142] M.A.K. Williams, A.T. Marshall, R.G. Haverkamp, and K.I. Draget. Stretching single polysaccharide molecules using afm: A potential method for the investigation of the intermolecular uronate distribution of alginate? *Food Hydrocolloids*, 22:18–23, 2008.

[143] F.A. Momany and J.L. Willett. Computational studies on carbohydrates: in vacuo studies using a revised amber force field, amb99c, designed for α-(1$\rightarrow$ 4) linkages. *Carbohydrate Research*, 326:194–209, 2000.

[144] A. Bohne, E. Lang, and C.W. von der Lieth. W3-sweet: Carbohydrate modeling by internet. *Journal of Molecular Modelling*, 4(1):33–43, 1998.

[145] A. Bohne, E. Lang, and C.W. von der Lieth. Sweet - www-based rapid 3d construction of oligo- and polysaccharides. *Bioinformatics*, 15:767–768, 1999.

[146] M. FernandezAlonso, J. Asensio, F. Canada, J. JimenezBarbero, and G. Cuevas. G2 and dft rigorous description of the inversion process of oxane and thiane used as simple ring systems to model sugar components. *ChemPhysChem*, 4:754–757, 2003.

[147] N.K. Lee and D. Thirumalai. Pulling speed dependent force-extension profiles for semiflexible chains. *Biophysical Journal*, 86(5):2641–2649, 2004.

[148] I.R. Cooke and D.R.M. Williams. Stretching polymers in poor solvents: pullout peaks and an unraveling transition. *Europhysics Letters*, 64:267273, 2003.

[149] Z. Lu, H. Hu, W. Yang, and P.E. Marszalek. Simulating force-induced conformational transitions in polysaccharides with the smd replica exchange method. *Biophysical Journal*, 91:L57L59, 2006.

[150] I. Neelov, D.B. Adolf, and T.C.B. McLeish. Molecular dynamics of pectin extension. *Macromolecular Symposia*, 252:140148, 2007.

[151] W. Wang, K.A. Kistler, K. Sadeghipour, and G. Baran. Molecular dynamics simulation off afm studiesof a single polymer chain. *Physics Letters A*, 372: 7007–7010, 2008.