

Molecular Dynamics Investigation of the Arabinan-Cellulose Interface for Cellulose Nanocomposite Applications

Luke Thornley

Cal Poly San Luis Obispo

June 2015

Advisor: Trevor Harding

Abstract:

Atom level computer simulations of the arabinan and cellulose interface were performed to better understand the mechanisms that give arabinan-cellulose composites (ArCCs) their strength with the goal to improve man-made ArCCs. The molecular dynamics (MD) software LAMMPS was used in conjunction with the ReaxFF/c force field to model the bond between cellulose and arabinan. A cellulose nanocrystal with dimensions 51 x 32 x 8 Å was minimized with various weight percent of water, 0%, 3%, 5%, 8%, 10%, and 12%. After the system was equilibrated for at least 100,000 femtoseconds, an arabinan molecule composed of 8 arabinose rings was added close to the cellulose's surface and equilibrated until fully adsorbed. A straightening force was applied to the arabinan during adsorption so the arabinan would lay flat on the {200} plane of the cellulose. A simulated AFM pull was performed to measure the force needed to desorb the arabinan from the cellulose. Due to computational resource limits, the pull speed was much faster than physical experiments, 500 m/s and 50 m/s. In general, the force needed for desorption increased with increasing water content with the force plateauing at 8wt% water. This increase in strength is probably due to water forming bridging hydrogen bonds between the relatively flat cellulose and crimped arabinan. Without water, fewer hydrogen bonds would form between cellulose and arabinan. This is an effect that will probably only be seen at high strain rates. Pull speeds of slower than 0.5 m/s must be performed to get accurate results.

Table of Contents

1	Background.....	1
1.1	Plastics.....	1
1.2	Arabinan Cellulose Composite (ArCC).....	1
1.2.1	A Composite Inspired By Cactus Spines.....	1
1.2.2	Arabinan.....	2
1.2.3	Cellulose.....	3
1.3	Molecular Dynamics (MD).....	5
1.3.1	History.....	5
1.3.2	Basics of MD.....	5
1.3.3	MD Software.....	6
1.3.4	LAMMPS.....	6
1.3.5	Experiments using MD.....	7
2	MD and Biopolymers.....	8
2.1	MD Studies with Cellulose.....	8
2.2	MD Studies with Arabinan.....	9
3	Research Question.....	9
4	Experimental Procedure.....	9
4.1	Produce an MD model of Arabinan, Crystalline Cellulose, and Water.....	9
4.2	Adsorption of Arabinan onto Cellulose.....	12
4.3	Measuring Arabinan-Cellulose Adhesion Strength.....	12
5	Results.....	13
5.1	Organizing Data.....	13
5.2	Effect of adding water to ArCC and Strain Rate Effects.....	14
5.3	Future Research.....	15
6	Conclusion.....	16
7	References.....	17
8	Appendix.....	20
8.1	Example Code.....	20
8.2	Using OVITO.....	23

1 Background

1.1 Plastics

Plastic production is a \$2.5 trillion per year industry¹ and bioplastics are a growing part of the industry. Bioplastics are important because they are not reliant on petroleum supplies, are environmentally neutral over their lifecycle, and most are biodegradable.

Traditional plastics in the U.S. are made from liquid petroleum gases (LPG) and natural gas liquids (NGL), which are byproducts of petroleum and natural gas refining. Plastic production takes a significant amount of resources on a national scale. In 2010, 2.7% of petroleum, 1.7% of natural gas, and 1.7% of electricity consumption in the U.S. was used to make plastics². Cost of production for traditional plastics relies on the cost and availability of petroleum. The world has a finite oil supply. Humanity will burn through the 1.3 trillion barrels left in the major oil fields in about 40-50 years³. Even with the discovery of new oil fields, a new source of plastics must be found to maintain the world's current standard of living.

Bioplastics are derived from agricultural products. In 2012, the global production of bioplastics was 1.4 million tonnes while global production of traditional plastics was around 300 million tonnes⁴. Despite bioplastics being almost 1% of the global plastic industry, the land used to make that plastic only accounted for 0.01% of global agricultural area. The bioplastics industry is growing at an incredible rate and is becoming a larger part of the market share, while being low impact on global food production. Additionally, because bioplastics are made from biological material, most can be composted or will decompose in landfill faster than petroleum-based plastics.

1.2 Arabinan Cellulose Composite (ArCC)

1.2.1 A Composite Inspired By Cactus Spines

Cactus spines are a strong natural composite that have attracted the interest of researchers in recent years. There has been specific interest in the spines of the *Opuntia ficus-indica* (OFI) cactus. OFI spines are a fiber-matrix composite made of a 50/50 blend

of arabinan and cellulose where the arabinan acts as a binder for aligned cellulose fibers. Previous research at Cal Poly San Luis Obispo has focused on creating a composite made from arabinan and cellulose⁵ with the hopes of duplicating OFI's fiber-matrix structure and achieving high strengths and stiffness's⁶. The ArCC made at Cal Poly San Luis Obispo had a stiffness of 6 GPA, which is higher than PET, a commonly used plastic in packaging. There is hope that with further research, commercially viable biocomposites can be manufactured. In order to achieve a better product, more must be learned of the nature of the bond between the components in ArCC.

1.2.2 Arabinan

Pectins are one of the components in cell walls that hold cellulose fibers together. Arabinan is a polysaccharide that exists as a side chain on pectins. There is evidence that arabinan side chains have the strongest interaction with cellulose out of all polysaccharides in pectin⁷. The adsorbing ends of arabinan are made of arabinose, which is a 5 sided sugar ring. Arabinan is found in large quantities in agricultural products like apples and sugar beets. Chemical arabinan derived from sugar beets is about \$20 per gram⁸, making it too expensive for making plastics outside of a research setting. However, because arabinan is common in agricultural products, large-scale production could drive down cost if there is demand. The actual structure that chemical arabinan takes when purified from sugar beets varies. The arabinose forms straight or branched chains involving 3 to 8 arabinose units (Figure 1)⁹.

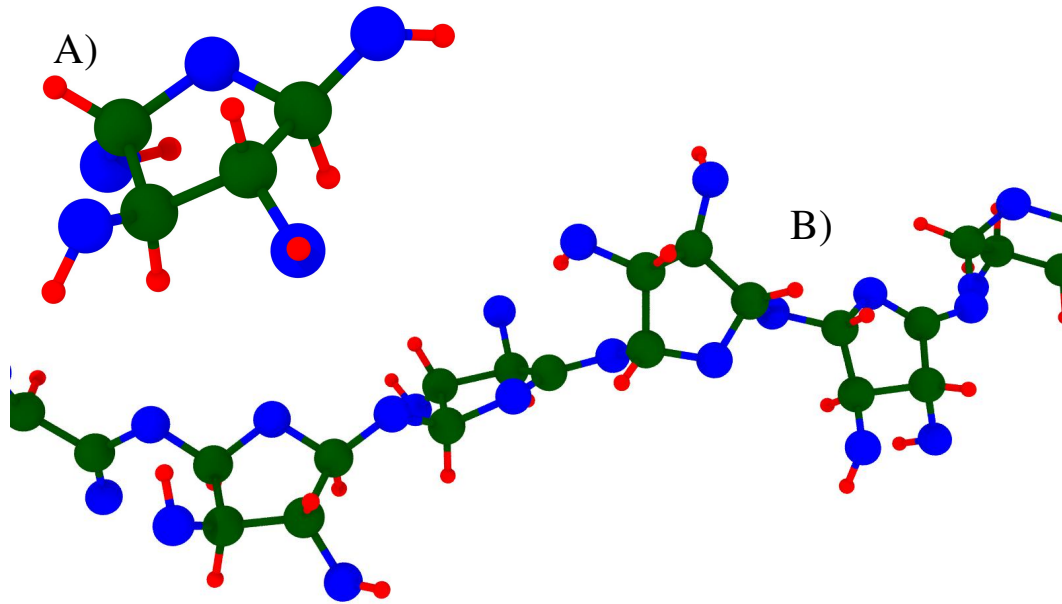


Figure 1- Arabinan is composed of branching chains of the arabinose 5-sided sugar ring. A) Rendering of the arabinose unit cell. B) The arabinan chain used in this study is 8 arabinose monomers long.

1.2.3 Cellulose

Cellulose is the most common natural material in the world with the global biosphere producing $10^{10} - 10^{12}$ tons per year¹⁰. Cellulose is a polymer of glucose where each glucose unit is rotated 180° with respect to its neighbor. The cellulose unit is called cellobiose, which is a dimer of the glucose unit (Figure 2)¹¹. Cellulose can have a degree of polymerization (DP) of anywhere from a handful of cellobiose units to 20,000.

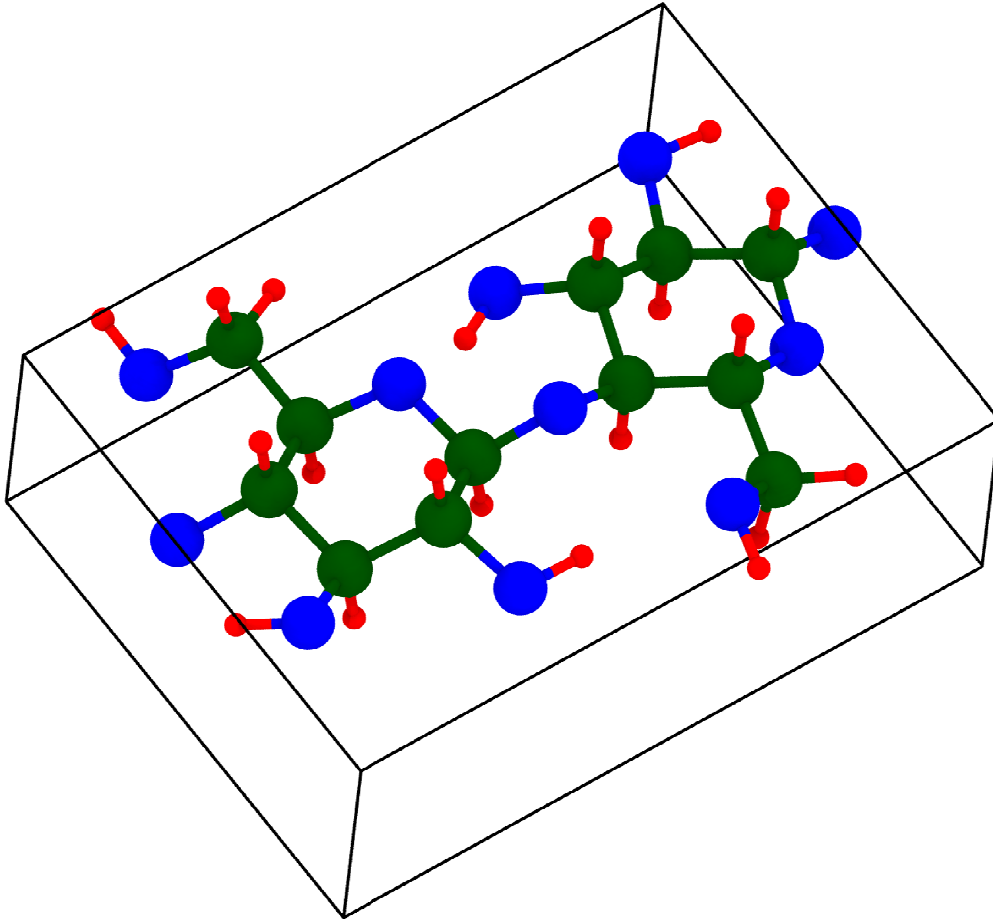


Figure 2- Cellobiose is well understood. The angle between the two glucose molecules is preferentially lies at one of three orientations called gt, tg, or gg. Cellobiose is the repeating unit that forms cellulose.

Cellulose is of interest because of its high specific strength and renewable nature¹². While amorphous cellulose has been studied because it is easier to process¹³, crystalline cellulose is much stronger. Natural cellulose itself is a composite consisting of crystalline nanocellulose (CN) in an amorphous cellulose matrix. It should be noted that CN are also called cellulose whiskers, cellulose nanofiber (CNF), microfibrillated cellulose (MFC), nanocrystalline cellulose (NCC), and nanocellulose. CN is typically 5-20 nm in length depending on how it is are processed¹⁴. Cellulose naturally exists as crystalline regions separated by amorphous cellulose. The amorphous cellulose can be etched away, leaving behind CN.

1.3 Molecular Dynamics (MD)

1.3.1 History

Alder and Wainwright conceived molecular Dynamics (MD) in the 1950s by calculating the interactions between hard spheres, which represented atoms¹⁵. The first MD simulations were in liquid systems. MD has come a long way since then, and has branched out into several similar fields due to larger scale simulations becoming possible through increased computational power. There are modeling methods that are more precise and others that are less precise than MD. The most accurate method is called Quantum Mechanics (QM), but is much more computationally intensive. There are many types of modeling for larger scales, but they cannot capture the phenomena that researchers are looking for in MD. MD is at the right level of accuracy and computational efficiency to help solve problems in fracture mechanics, shock physics, dislocations, and environments that are hard to study, such as high temperature systems.

1.3.2 Basics of MD

Molecular dynamics is a way to use computers to simulate material systems. MD is an all-atoms (AA) simulation method where the atoms are explicitly defined and there is a set of potentials, also called force fields, which describe the interactions between the atoms. Just to clarify, force fields and potentials are the same thing. Potentials are a set of equations that determine the forces that atoms exert on each other. Forces are calculated at each time step, which can be as small as tenths of femtoseconds. Once potentials determine the forces, the atoms are subjected to the Newtonian laws of motions based on those forces. Over the course of the simulation, the atoms form bonds, break bonds, and move around. Depending on the potential used, Van der Waals bonds can be explicitly defined and tracked or are *emergent behavior*. Explicitly defined secondary bonds are less accurate but take less computing power.

MD simulations can have a few dozen to a few million atoms, and run anywhere from femtoseconds to a microsecond. The main limitation for MD is that simulations longer than a microsecond are too computationally intensive for our current technology¹⁶. While computing power is always increasing, it will be a long time before large time scale simulation of macro-sized systems is possible. There are coarse-grained (CG)

modeling systems that approximate groups of atoms, which are able to more economically model large system¹⁷. Also, there is software like finite element analysis (FEA) that only looks at the macro scale. However, this paper will only be looking at AA MD simulations.

1.3.3 MD Software

To run an MD simulation, MD software, a force field, initial atom coordinates, and initial environmental conditions are needed. First, there is the MD software, like LAMMPS, AMBER, GROMACS, GROMOS, and CHARMM. The MD software is what the force fields, initial conditions, and actual atoms coordinates are inputted into. Each MD package is different, but many of them can do the same things, and some are better at some things than others.

Next, there are the potentials. Potentials are made specifically for the materials and environmental conditions that are going to be modeled. For example, a model for water at room temperature may be unable to accurately reproduce water at temperatures close to its boiling point. Potentials are calibrated using empirical data^{18,19,20} ensuring that the models closely represent reality. Creating a force field is outside of the scope of this paper, so a force field will be selected from the existing library of MD software. The ReaxFF potential is ideal for this study because the secondary bonds are emergent behavior of the potential and therefore more accurate, and ReaxFF is calibrated for just the atoms in arabinan, cellulose, and water.

The last two elements needed to run a simulation are the initial conditions of the system and the starting locations of all of the atoms. A bulk material is generated by populating the materials repeating unit cell with atoms. Many MD software packages have crystal unit cells built in, like FCC, BCC, HCP, etc. Giving LAMMPS the x-y-z coordinates of each atom can generate a non-uniform polymer. Variables like temperature, boundary conditions, and time can be controlled through the software. The temperature is controlled by an algorithm, like the Nose-Hoover thermostat^{21,22,23,24}.

1.3.4 LAMMPS

For this study the MD software Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) was selected because it is free, constantly updated, compatible

with organic polymers, and can easily take advantage of multiple computer processors. LAMMPS was developed by Sandia National Lab²⁵. LAMMPS is most commonly used to model metal systems, but is just as well equipped to model organic polymers²⁶. LAMMPS is compatible with common force fields designed for organic systems like ReaxFF and GLYCAM. LAMMPS makes use of multiple processors by breaking up the simulation into 3D sections for each processor. LAMMPS can also output the data needed to study adhesion strength, such as pulling forces on groups of atoms.

The three main bonds present are hydrogen bonds, Columbic interactions, and Van der Waals interactions. These forces are described by the potential and manifest as emergent behavior so they are all considered in the simulation. In MD, a hydrogen bond can be geometrically defined with a cutoff distance of 2 - 3 Å and a bond angle of 90° - 120°. The value of this cutoff distance and angle can greatly affect the number of hydrogen bonds recorded.²⁷⁻³². However, since hydrogen bonds are not explicitly described in ReaxFF, LAMMPS cannot output the number of hydrogen bonds. Other data is needed to quantify adhesion strength in systems.

A visualization software called OVITO is used to generate images and movies of the simulations.³³

1.3.5 Experiments using MD

There are three steps for mechanically testing a material using MD: minimization, equilibration, and testing. Energy minimization is a process where the MD software moves atoms around to find to lowest local energy state for the atoms. Minimization ensures that atoms are not too close or far from each other at the start of the simulation. Equilibration is the process of running a simulation with no external forces until a stable configuration is reached. Equilibration ensures that the system is in a *natural* state before mechanical testing. For this study, a simulated Atomic Force Microscope (AFM) pull will be used for mechanical testing. This involves attaching a virtual spring to a single atom or a group of atoms and moving the other end of the spring away from the attached atoms.

2 MD and Biopolymers

2.1 MD Studies with Cellulose

Cellulose is a well-studied molecule in the MD field. These studies range from confirming the crystal structure of cellulose³⁴, finding the strength of adsorption of a cellulose strand on crystalline cellulose³⁵, to the structural changes due to temperature in certain cellulose systems³⁶.

All of these studies agree that cellulose forms a crystal bundle (Figure 3)³⁷. Natural cellulose forms into I α or I β configuration, with I β being more common for higher order plants. Processing the cellulose will change the crystalline cellulose into a different crystal structure. The most stable cellulose is cellulose II³⁸, but this study is only going to focus on cellulose I β because that is what is found in nature. If a cellulose bundle is large enough, a slight twist is seen along the length of the CN³⁹.

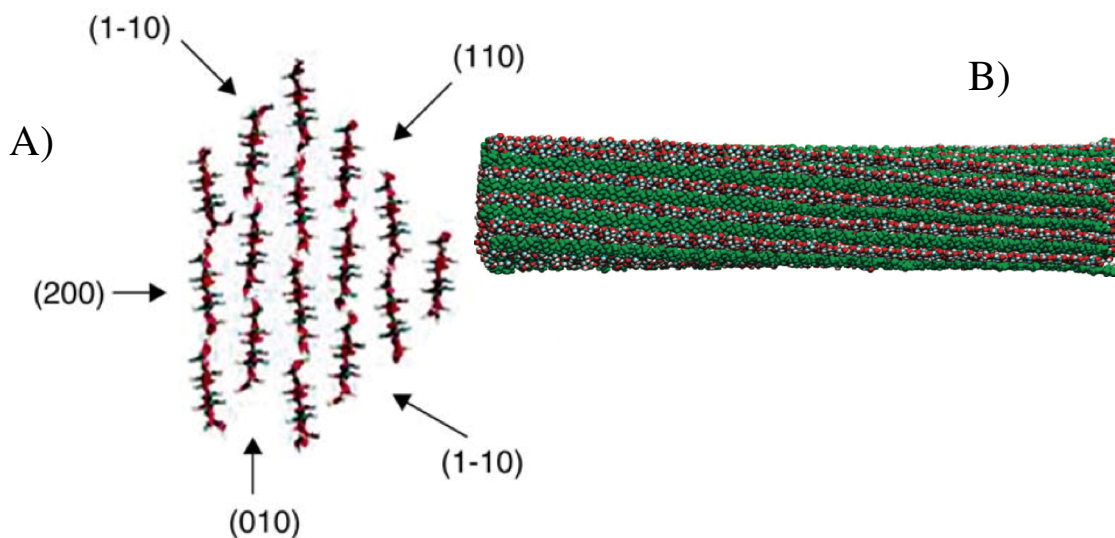


Figure 3- Cellulose I α is similar, but the similar crystal faces are labeled differently. (110) in I α = (200) in I β . A) Cellulose I β crystal cross-section with labeled crystal faces³⁰. B) Side view of cellulose I β with a visible twist. This twist depends on the temperature of the system³².

While most of these articles have no direct application, a few articles modeled real-world composites. The first was a study on the cellulose lignin system⁴⁰. This study concluded that the lignin closest to the cellulose would align well, but adding more lignin would result in a disordered structure. Also, some of the crystal faces did not promote a good bond, resulting in more disorder. The mix of a disorder and order in the simulations might explain confusing experimental results where some characterization methods, like

X-ray diffraction, show disorder, and others, like photoconductimetry, suggest order. This is an excellent example of how MD can be used to inform and explain physical experiments.

Another article models a polyethylene-vinyl acetate (EVA) – CN composite⁴¹. The article states that composites are not fully understood because of lack of knowledge of the atomic interactions between the matrix and filler. The estimated strength of the EVA-CN composite was stronger than the measured strength, probably due to the actual materials having rougher surfaces than the idealized particles in the model. Because it is difficult to model natural defects, most strengths in modeling are overestimated.

2.2 MD Studies with Arabinan

Arabinan has been modeled since at least the 1990s⁴², but it is not nearly as studied as cellulose. The literature on modeling of arabinan only goes as far as measuring bond angle in energy minimized arabinan. The type of arabinan that is most commonly modeled is α -L-arabinofuranose, which is the same arabinan found in sugar beets. Most of the literature on arabinan is biological studies. In my research, I did not find any articles that did a MD study of the arabinan-cellulose system.

3 Research Question

The objective of this study is to analyze the cellulose arabinan interface in a man-made composite. The simulations will reveal the nature of the bond between cellulose and arabinan. Another area of interest is the effect of different amounts of water on the strength of bonding. Hydrogen bonds are sensitive to residual water in the system, and quantifying the effect water has on strength will inform the experimentalists who will be processing and testing ArCC.

4 Experimental Procedure

4.1 Produce an MD model of Arabinan, Crystalline Cellulose, and Water

The monomer units of arabinan and cellulose were generated in an online application that can be exported into a .xyz data format that can be used in OVITO and LAMMPS⁴³. These monomers were manipulated in OVITO to generate the polymers by

repeating the monomers along the x, y, and z axis. The cellulose was placed into the cellulose I β configuration because that is the most common natural cellulose crystal.

OVITO exported the polymer configurations into a .xyz file, which is then manually converted into a data.filename.txt file according to the *charge* LAMMPS data type. Once the data.filename.txt file had been created, the polymers were simulated in LAMMPS. Each polymer was minimized and equilibrated for about 15,000 femtoseconds to ensure that LAMMPS with ReaxFF can accurately model these materials.

To ensure that ReaxFF could accurately model cellulose, 2 layers of cellulose chains 3 wide and 5 cellobiose units long were modeled with periodic boundaries, simulating an infinite sheet 2 layers thick. After the cellulose was minimized and equilibrated, the lattice parameters of the cellulose matched the expected values for cellulose I β , 3.9 Å between sheets, 8.2 Å between chains in a sheet, and with the cellobiose units being 10.3 Å long (Figure 4). This process was repeated with water present near the cellulose so the water would be equilibrated on the cellulose surface. Water was added in increments of 35 water molecules up to 175 water molecules and a percent weight of the whole system was calculated.

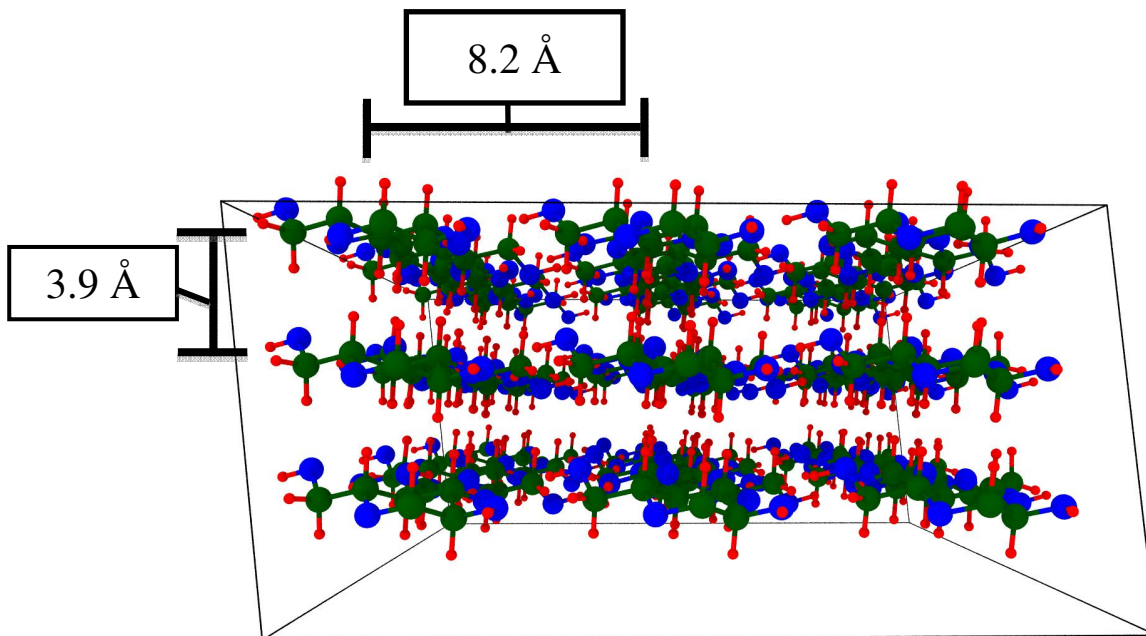


Figure 4- The crystal structure of cellulose is well understood. However, since cellulose is organic, the crystal structure is not as neat as it is in metals. The lattice parameters are averages, but can change slightly from time step to time step. This figure represents pre-equilibrated cellulose.

While arabinan has not been well studied with ReaxFF, because cellulose and arabinan have the same atoms: oxygen, hydrogen, and carbon, ReaxFF should work just as well for arabinan. Potentials are designed to correctly model a narrow set of interactions, like the pairwise interactions between the atoms. Since there are no new atoms or interactions in arabinan, the same potential should work for cellulose and arabinan. To verify that arabinan was being correctly modeled a simulation of just arabinan with a small straightening force was performed and the structure of the atoms was monitored throughout. The 5-sided ring structure of the arabinan persisted as expected, and the bond lengths between atoms remained consistent.

To verify water was correctly modeled, 27 waters were simulated until they formed a sphere and the approximate density of the waters was determined to be what the density of water should be (Figure 5).

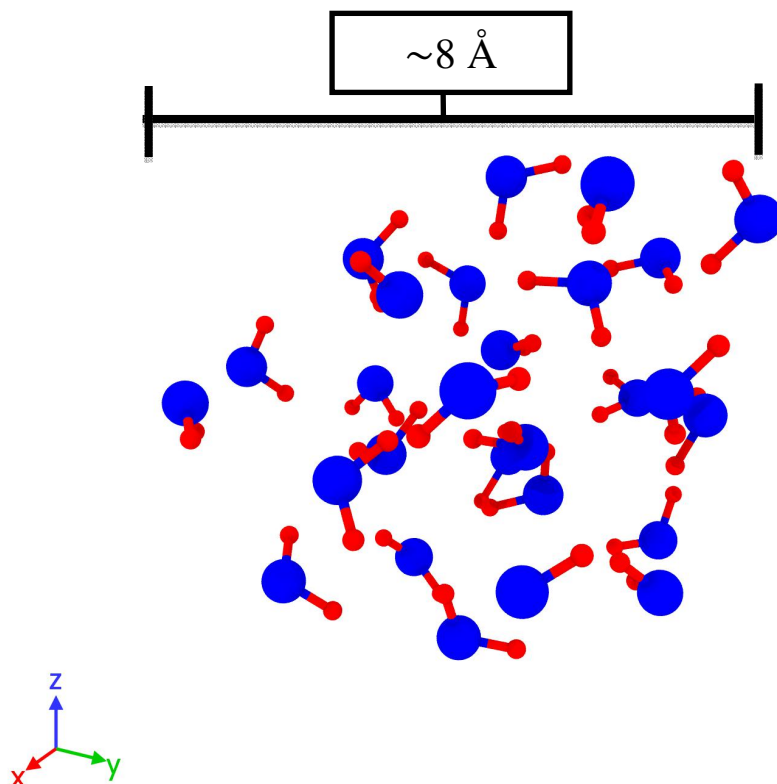


Figure 5- 27 water molecules forming a rough sphere after being minimized and equilibrated.

4.2 Adsorption of Arabinan onto Cellulose

After arabinan and cellulose were independently verified to be physically accurate, pre-minimized and pre-equilibrated arabinan and cellulose were placed in the same simulation. The arabinan was placed 5 – 8 Å above the cellulose, aligned in the same direction as the cellulose chains. The system was equilibrated for 10,000 – 17,500 femtoseconds. The system was equilibrated for different times so there would be different results during the pull phase. A small straightening force was applied along the axis of the arabinan so it would not bunch up on itself, but rather adsorb onto the cellulose, parallel to the cellulose. The process was repeated for each cellulose system with water present.

4.3 Measuring Arabinan-Cellulose Adhesion Strength

To measure the strength of the arabinan-cellulose bond, an AFM pull was performed by attaching a virtual spring to the arabinan on one side and a fixed-point hydrogen atom on the other side (Figure 6). This hydrogen's position was controlled in such a way that it moved 0.005 and 0.0005 Å per femtosecond, which is equivalent to 500 and 50 m/s. The simulations were run until the arabinan had fully desorbed from the cellulose. The pull speeds were selected due to computational resource limits for this project. At these speeds, there is a clear effect due to strain rate, which can be eliminated if the pulls are completed slower than 0.5 m/s⁴⁴.

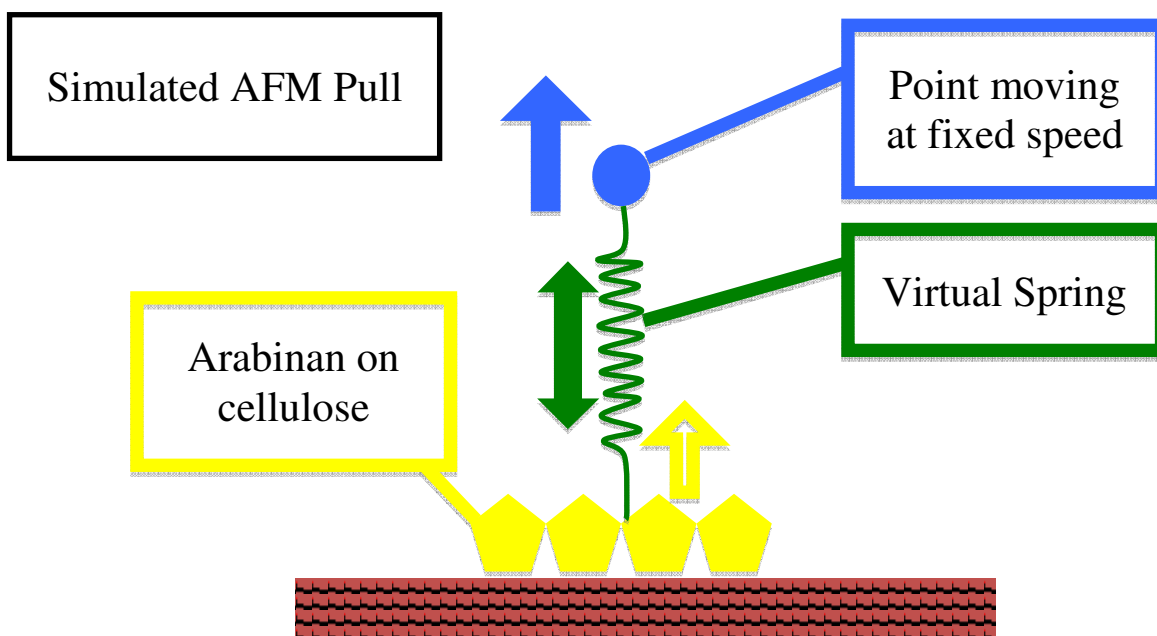


Figure 6- In an AFM pull, the probe tip is used to exert a force on the target atoms. In an MD simulated pull, the probe tip is replaced with a spring fixed to a point. Using this method the pull rate, spring constant, and spring equilibrium distance can be controlled.

LAMMPS exported a dump.filename.xyz file, which contains the xyz data for each atom per time step, and an output.filename.txt file, which contains the various data as defined by the code. For this study, the important output information is the total force the spring exerts on the arabinan at each time step. This simulation was repeated for each variation of water content with 4 – 6 runs at 500 m/s and 50m/s.

5 Results

5.1 Organizing Data

The maximum force during desorption was used to quantify the adhesion strength between the cellulose and arabinan. As the fixed point on the other side of the spring is moved, the force the spring exerts on the arabinan increases, until there is a rapid decrease in force. This corresponds to the breaking of several hydrogen bonds at once. The force was converted into an approximate stress and plotted as a function of time (Figure 7). The max pull stresses ranged from 1 to 3 GPa, which is an acceptable range. However, since the simulations pulled so fast, actual assumptions of the actual strength of the arabinan cellulose interface cannot be made.

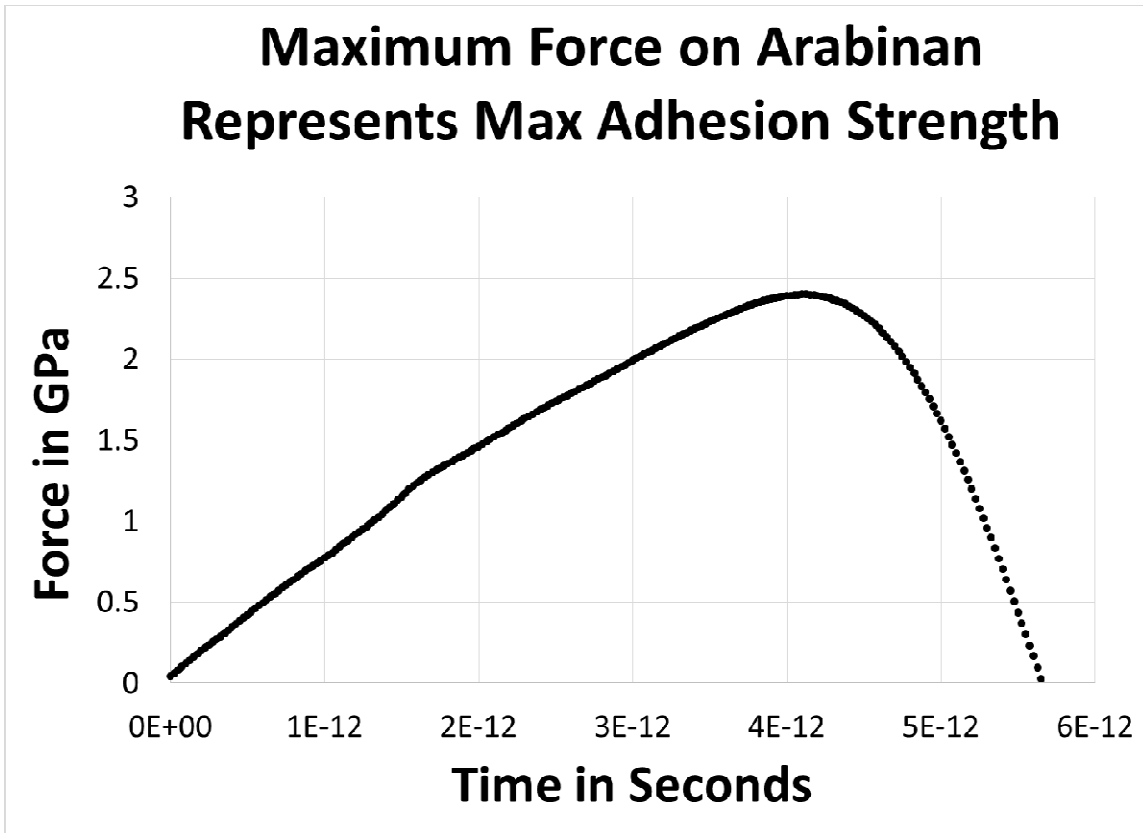


Figure 7- The force exerted on the arabinan increases until the hydrogen bonds break and the force decreases. The maximum force before the bonds break gives an idea of the adhesion strength.

5.2 Effect of adding water to ArCC and Strain Rate Effects

Adding water to the simulation increased the stress needed to desorb the arabinan. This is probably a strain rate effect. Water forms bridging hydrogen bonds between the cellulose and arabinan. Since the arabinan is being pulled so fast, the hydrogen bonds do not have time to equilibrate and break once they are past their stretching limit (Figure 8).

The effects of stain rate are apparent in the numerical data (Figure 9). Lowering the strain rate lowered the force needed for desorption.

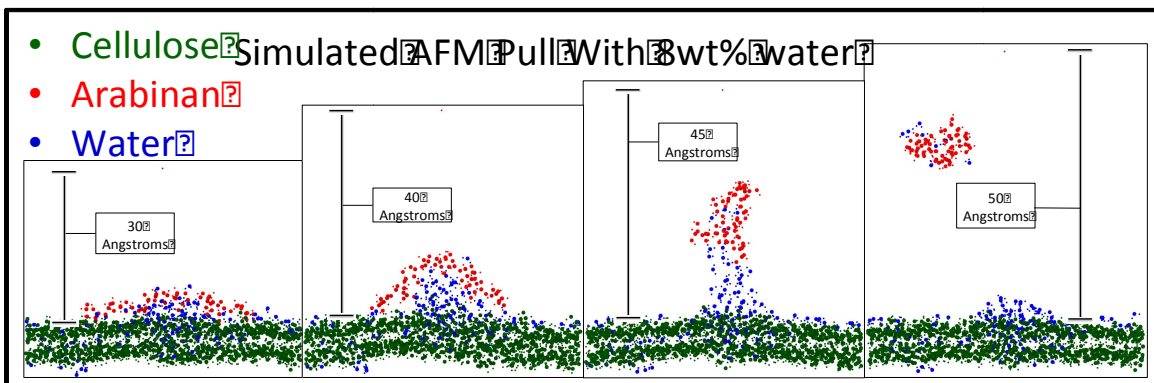


Figure 8- Renderings of one of the simulated AFM pulls. The water is clearly creating bridging bonds between the cellulose and arabinan. At the third step, water is the only thing connecting the arabinan and cellulose. This strange behavior of water is probably a high strain rate effect.

5.3 Future Research

These simulations should be run at pulls of 0.5 m/s and slower, to determine the effect of water on adhesion without strain rate effects. Other things that could be changed in the simulations are the geometry of adsorbed arabinan on cellulose, the difference in adhesion on different cellulose crystal faces, the effect of arabinan chain length, branched arabinan, and different simulation sizes. The best way to test the effect of water would be to run the simulation with a non-polar solvent instead of water. This would help determine the effect the hydrogen's bond with water has on the strength of the system.

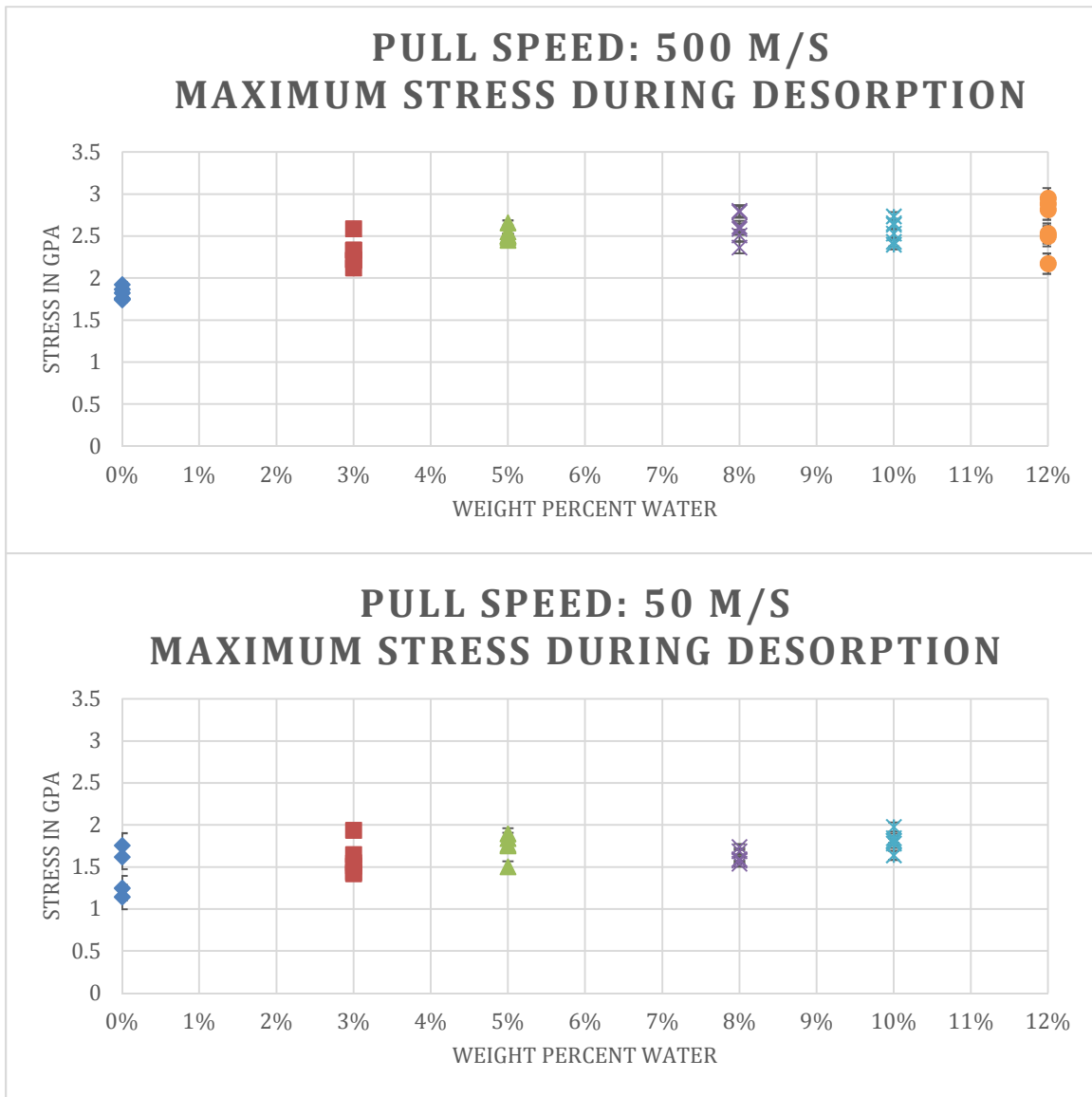


Figure 9- Graphs showing the maximum stress needed to desorb arabinan from cellulose. At 50 m/s 12wt% water data was not obtained due to time constraints.

6 Conclusion

Cellulose, arabinan, and water were successfully modeled using the ReaxFF potential with LAMMPS. Arabinan was adsorbed onto the cellulose in the absence and presence of water. A simulated AFM pull determined that water increases the adhesion strength between cellulose and arabinan at high strain rates. However, there was also more scatter at high strain rate with high amounts of water. More testing needs to be done to determine the effect of water on the cellulose arabinan interface at realistic pull speeds.

7 References

- ¹ Huff, E. "Plant-Based Plastics to Replace Petroleum Plastics?" *NaturalNews*. Natural News Network, 30 Jan. 2010. Web. 28 Nov. 2014.
- ² US EIA. "How Much Oil Is Used to Make Plastic?" *EIA FAQs*. US Energy Information Administration, 18 June 2014. Web. 28 Nov. 2014.
- ³ "When Will Oil Run Out?" *The Institution of Mechanical Engineers IMechE*. Institution of Mechanical Engineers, n.d. Web. 29 Nov. 2014.
- ⁴ *Bioplastics*. N.p.: Agrotechnology & Food Sciences Group, n.d. *Corbion*. European Bioplastics, Dec. 2013. Web. 30 Nov. 2014.
- ⁵ Smithfield, Marianne, Carl Petterson, and Robbie Miller. "An Investigation of Arabinan and Cellulose Based Nano-composite Fabrication Methods." Thesis. Cal Poly San Luis Obispo, 2013. *Digital Commons*. California Polytechnic State University, June 2013. Web. 1 Dec. 2014.
- ⁶ Johnson, Ross, Austin Rosso, and Nick Semansky. *Fabrication and Testing of Arabinan Cellulose Nanocomposites*. *Cal Poly Digital Commons*. California Polytechnic State University, June 2014. Web. 1 Dec. 2014.
- ⁷ Zykwiniska, A. W. "Evidence for In Vitro Binding of Pectin Side Chains to Cellulose." *Plant Physiology* 139.1 (2005): 397-407. *JSTOR*. Web.
- ⁸ Megazyme International Ireland. *Arabinan Sugar Beet Polysaccharides*. N.p., n.d. Web. 01 Dec. 2014.
- ⁹ Westphal, Yvonne, Stefan Kühnel, Henk A. Schols, Alphons G.j. Voragen, and Harry Gruppen. "LC/CE-MS Tools for the Analysis of Complex Arabino-oligosaccharides." *Carbohydrate Research* 345.15 (2010): 2239-251. *Science Direct*. Elsevier Ltd., 21 Aug. 2010. Web. 3 Dec. 2014.
- ¹⁰ Habibi, Youssef, Lucian A. Lucia, and Orlando J. Rojas. "Cellulose Nanocrystals: Chemistry, Self-Assembly, and Applications." *Chemical Reviews* 110.6 (2010): 3479-500. Web.
- ¹¹ Yui, Toshofumi, and Sachio Hayashi. "Structural Stability of the Solvated Cellulose III Crystal Models: A Molecular Dynamics Study." *Cellulose* 16.2 (2009): 151-65. *Web of Science*. Web. 25 Oct. 2014.
- ¹² Zhao, H., J. Kwak, Z. Conradzhang, H. Brown, B. Arey, and J. Holladay. "Studying Cellulose Fiber Structure by SEM, XRD, NMR and Acid Hydrolysis." *Carbohydrate Polymers* 68.2 (2007): 235-41. *Science Direct*. Elsevier Ltd. Web.
- ¹³ Chen, Wei, Gary C. Lickfield, and Charles Q. Yang. "Molecular Modeling of Cellulose in Amorphous State. Part I: Model Building and Plastic Deformation Study." *Polymer* 45.3 (2004): 1063-071. *Science Direct*. Elsevier Ltd., 13 Nov. 2003. Web. 25 Oct. 2014.
- ¹⁴ Zhao
- ¹⁵ Buehler, Markus. "Introduction to Atomistic Modeling Techniques." Lecture Series. MIT. Spring 2006. *MIT*. Web. 2 Dec. 2014.
- ¹⁶ Hahn, Eric. "Introduction to Molecular Dynamics Simulation and Modeling." MESS Meeting. San Luis Obispo. Fall 2014. Speech.
- ¹⁷ Wohlert, Jakob, and Lars A. Berglund. "A Coarse-Grained Model for Molecular Dynamics Simulations of Native Cellulose." *Journal of Chemical Theory and Computation* 7.3 (2011): 753-60. *ACS*. Web. 25 Oct. 2014.
- ¹⁸ Hanus, Jaroslav, and Karim Mazeau. "The Xyloglucan-cellulose Assembly at the Atomic Scale." *Biopolymers* 82.1 (2006): 59-73. *Web of Science*. Web. 25 Oct. 2014.
- ¹⁹ Mackerell, A. D., D. Bashford, Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-Mccarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C.

-
- Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, and M. Karplus. "All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins." *The Journal of Physical Chemistry B* 102.18 (1998): 3586-616. ACS. Web. 25 Oct. 2014.
- ²⁰ Mazeau, K., and L. Heux. "Molecular Dynamics Simulations of Bulk Native Crystalline and Amorphous Structures of Cellulose." *The Journal of Physical Chemistry B* 107.10 (2003): 2394-403. ACS. Web. 25 Oct. 2014.
- ²¹ Liu, Hanbin, Kenneth L. Sale, Bradley M. Holmes, Blake A. Simmons, and Seema Singh. "Understanding the Interactions of Cellulose with Ionic Liquids: A Molecular Dynamics Study." *The Journal of Physical Chemistry B* 114.12 (2010): 4293-301. ACS. Web. 25 Oct. 2014.
- ²² Yang, Mingjun, Vasileios Koutsos, and Michael Zaiser. "Interactions between Polymers and Carbon Nanotubes: A Molecular Dynamics Study." *The Journal of Physical Chemistry B* 109.20 (2005): 10009-0014. ACS. Web. 25 Oct. 2014.
- ²³ Hanus
- ²⁴ Buehler
- ²⁵ Sandia National Laboratories. "LAMMPS Molecular Dynamics Simulator." *LAMMPS Molecular Dynamics Simulator*. Sandia National Lab, n.d. Web. 02 Dec. 2014.
- ²⁶ Gómez Zuluaga, Mateo; Fernando Luis Dri; Robert J. Moon; Pablo Daniel Zavattieri (2014), "Crystalline Cellulose – Atomistic Modeling Toolkit," <https://nanohub.org/resources/20703>.
- ²⁷ Chen,
- ²⁸ Chenoweth
- ²⁹ Mackerell
- ³⁰ Miyamoto
- ³¹ Heiner
- ³² Paavilainen
- ³³ A. Stukowski, *Modelling Simul. Mater. Sci. Eng.* 18, 015012 (2010)
- ³⁴ Baker, A.a., W. Helbert, J. Sugiyama, and M.j. Miles. "New Insight into Cellulose Structure by Atomic Force Microscopy Shows the I α Crystal Phase at Near-Atomic Resolution." *Biophysical Journal* 79.2 (2000): 1139-145. *Biophysical Society*. Web. 25 Oct. 2014.
- ³⁵ Bergenstråhle, Malin, Esben Thormann, Niklas Nordgren, and Lars A. Berglund. "Force Pulling of Single Cellulose Chains at the Crystalline Cellulose–Liquid Interface: A Molecular Dynamics Study." *Langmuir* 25.8 (2009): 4635-642. ACS. Web. 25 Oct. 2014.
- ³⁶ Bergenstråhle, Malin, Lars A. Berglund, and Karim Mazeau. "Thermal Response in Crystalline I β Cellulose: A Molecular Dynamics Study." *The Journal of Physical Chemistry B* 111.30 (2007): 9138-145. ACS. Web. 25 Oct. 2014.
- ³⁷ Besombes, Stéphane, and Karim Mazeau. "The Cellulose/lignin Assembly Assessed by Molecular Modeling. Part 1: Adsorption of a Threo Guaiacyl β -O-4 Dimer onto a I β Cellulose Whisker." *Plant Physiology and Biochemistry* 43.3 (2005): 299-308. *Science Direct*. Elsevier Ltd., 16 Mar. 2005. Web. 25 Oct. 2014.
- ³⁸ Miyamoto, Hitomi, Myco Umemura, Takeshi Aoyagi, Chihiro Yamane, Kazuyoshi Ueda, and Kazuhiro Takahashi. "Structural Reorganization of Molecular Sheets Derived from Cellulose II by Molecular Dynamics Simulations." *Carbohydrate Research* 344.9 (2009): 1085-094. *Science Direct*. Elsevier Ltd., 20 Mar. 2009. Web. 25 Oct. 2014.
- ³⁹ Paavilainen, Sami, Tomasz Róg, and Ilpo Vattulainen. "Analysis of Twisting of Cellulose Nanofibrils in Atomistic Molecular Dynamics Simulations." *The Journal of Physical Chemistry B* 115.14 (2011): 3747-755. ACS. Web. 25 Oct. 2014.
- ⁴⁰ Besombes, Stéphane, and Karim Mazeau. "The Cellulose/lignin Assembly Assessed by Molecular Modeling. Part 2: Seeking for Evidence of Organization of Lignin Molecules at the Interface with

-
- Cellulose." *Plant Physiology and Biochemistry* 43.3 (2005): 277-86. *Science Direct*. Elsevier Ltd., 16 Mar. 2005. Web. 25 Oct. 2014.
- ⁴¹ Chauve, Gregory, Laurent Heux, Rachid Arouini, and Karim Mazeau. "Cellulose Poly(Ethylene--vinyl Acetate) Nanocomposites Studied by Molecular Modeling and Mechanical Spectroscopy." *Biomacromolecules* 6.4 (2005): 2025-031. *ACS*. Web. 25 Oct. 2014.
- ⁴² Cros, Soizic, Anne Imberty, Nadia Bouchemal, Catherine Herv Du Penhoat, and Serge Perez. "Modeling of Arabinofuranose and Arabinan, II. Nmr and Conformational Analysis of Arabinobiose and Arabinan." *Biopolymers* 34.10 (1994): 1433-447. *Wiley Online Library*. John Wiley & Sons Inc., 1 Feb. 2004. Web. 3 Dec. 2014.
- ⁴³ Ertl, Peter, Bruno Bienfait, and CORNIA. *JSME Molecular Editor*. Computer software. *Molecular Networks*. Molecular Networks GmbH, 2014. Web. 2 Dec. 2014.
- ⁴⁴ Gautieri, Alfonso, Markus J. Buehler, and Alberto Redaelli. "Deformation Rate Controls Elasticity and Unfolding Pathway of Single Tropocollagen Molecules." *Journal of the Mechanical Behavior of Biomedical Materials* 2.2 (2009): 130-37. *Science Direct*. Elsevier, 14 Mar. 2008. Web. 29 May 2015.

8 Appendix

8.1 Example Code

The code in grey boxes comes from the file “in.5%.run1.slowpull.mac.txt” that was used to run one of the pull tests. This lines with “#” at the front are comments.

```
##### Equilibrating and pulling arabinan from cellulose with 5% water#####  
#### By Luke Thornley ####  
#### Purpose: adsorb arabinan and pull arabinan to get data.
```

Title at the top of code is always a good idea for later use.

```
## Run this command in lammgs:  
# mpiexec.exe -np 6 lmp_mpi.exe -in in.5%.run1.slowpull.mac.txt #
```

Comment to remind user what to run in the command terminal to run this code in lammgs

```
## Name of Data File saved to /Simulations  
variable frame index data.5%.run2.txt  
## Name associated with the dump and output files  
variable simname index 5wt.slowpull.run1
```

Variable is a command to store a string of text to be called later in the program. These lines save the name of the *data* file to be used for the initial atom geometry and what to call the output files. The rest of the grey boxes should be well explained by their in-line comments.

```
### Setup ###  
# This section contains the setup needed for any lammgs reax/c simulation.  
clear  
#      makes run more efficient  
units      real  
processors  3 2 1  
           # tells lammgs to divide the simulation into 6 parts, one for each process, 3 parts /  
           #in the x-direction, 2 parts in the y-direction, and none in the z direction. This /  
           #allocation minimized atom bonds across processor boundaries and makes sure /  
           #there are no blank processors.  
boundary p p p  
           # periodic boundaries in all direcitons  
atom_style  charge  
log         log.${simname}.txt  
           # created log file with name defiend earlier  
read_data  ${fname}  
           # read initial geometry file  
pair_style reax/c lmp_control  
           # sets reaxff/c potential and calls control file in current directory
```

```

pair_coeff * *ffield.reax.cho H C O
      # calls the ReaxFF file and sets atoms types 1 2 3 to their respective atoms
neighbor 2 bin
neigh_modify      every 100 delay 0 check no
timestep 0.25
      # 1 timestep = 0.25 femtoseconds

### Defining Groups ###
# Define the groups used later in simulation
region base block -25.76 26.14 -12.9885 20.303 -5 0
group base region base
      # regions to later be used to hold down cellulose during pull
group hydrogen type 1
group carbon type 2
group oxygen type 3
group A id 1890:2000
      #group defining arabinan
group C id 1:1680
      #group defining cellulose
group F id 2001
      #group for fix point
group H2O id 1681:1889
group pullA id 1895
      #pull point to flatten arabinan on negative x side
group pullB id 1989
      # pull point to flatten arabinan on positive x side

### Computes ###
## For dumping info later
compute Acom A com
      # computes center of mass for arabinan and gives output by c_Acom[i] where i=1,2,3 /
      gives x,y,z coords in angstroms
compute Ccom C com
      # center of mass for cellulose
compute peratom all pe/atom
      # potential energy per atom

### Set up the Pull ###
## The tether point is 0 angstroms away from the pull atom, the springs equilibrium length is 0
## MUST COME BEFORE TEMP FIXES
fix fixpoint F move linear 0.0 0.0 0.0
      #last number is the pull rate in angstroms per femtosecond
fix pullA pullA spring tether 1 -17 5 8 0
      #springs equilibrium is 5 short of fully equilibrated system
fix pullB pullB spring tether 1 19 5 8 0
      # both these fixes are sued to straighten the arabinan during equilibration

```

Specific Run Conditions

```
fix      1 all nve
fix      2 all qeq/reax 1 0.0 10.0 1e-6 param.qeq
fix      3 all temp/berendsen 300.0 300.0 100.0
        # these are temperature fixes. Fix 3 sets the temp to 300 K
thermo   500
        # dumps info about the system to the screen every 500 steps.
```

Data Dump

```
dump     1 all custom 100 dump.${simname}.ovi id type q x y z
        # Dumps the xyz file for ovito
variable p1 equal "step"
variable p2 equal "f_pull"
        # energy as  $1/2 * K * r^2$  where r is the difference between original COM and current position.
variable p3 equal "f_pull[1]" # x force magnitude
variable p4 equal "f_pull[2]" # y force magnitude
variable p5 equal "f_pull[3]" # z force magnitude
variable p6 equal "f_pull[4]" # total force magnitude
variable p10 equal "pe" # does not include spring energy
variable p11 equal "c_Ccom[3]" # center of mas of cellulose
variable p12 equal "c_Acom[3]" # center of mass of arabinan
        # all of these variables will be called later and can be used for post-simulation analysis
```

Run the simulation

```
run 60000
        # adsorbs the arabinan and equilibrates they system
unfix 3
        # gets rid of berendsen algorithm so system is not limited to a set atomic motion
unfix pullA
unfix pullB
        #gets rid of straightening factor on arabinan
fix fixpoint F move linear 0 0 0.00005
        # last number is the pull rate in angstroms per femtosecond
        # sets rate at which arabinan is pulled
fix base base recenter INIT INIT INIT
        #keeps cellulose from being pulled up
fix pull A spring couple F 4.0 0.0 0.0 0.0 22
        # creates virtual spring
fix printme all print 1000 "${p1} ${p2} ${p3} ${p4} ${p5} ${p6} ${p10} ${p11} ${p12}" /
        file output.${simname}.txt screen no
        # dumps output data for force vs distance graph
run 3000000
        # runs with settings to pull arabinan
```


8.2 Using OVITO

OVITO is a free atom visualization software useful for visualizing atom position data. OVITO is updated constantly so things may change after this short guide is written. OVITO can read many data file types such as .xyz, .cfg, and LAMMPS dump files. OVITO should automatically load up the dump file after it is opened from the "Load File" option. Select the checkbox to note that the file contains multiple time steps and then use the slide bar to look at the data.