

THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

**Molecular-level Simulations of Cellulose Dissolution by Steam and SC-CO₂
Explosion**

FARANAK BAZOORYAR



CHALMERS

Department of Chemical and Biological Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2014



UNIVERSITY OF BORÅS
SCIENCE FOR THE PROFESSIONS

Swedish Centre for Resource Recovery
UNIVERSITY OF BORÅS
Borås, Sweden, 2014

Molecular-level Simulations of Cellulose Dissolution by Steam and SC-CO₂ Explosion

FARANAK BAZOONYAR
ISBN 978-91-7597-058-5

© FARANAK BAZOONYAR, 2014.

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie nr 3739
ISSN 0346-718X

Department of Chemical and Biological Engineering
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden
Telephone + 46 (0)31-772 1000

Skrifter från Högskolan i Borås, nr 51
ISSN 0280-381X

University of Borås
SE-501 90 Borås
Sweden
Telephone + 46 (0)33-435 4000

Cover: Changes in the cellulose crystal structure during steam explosion at 250 °C and 39.7 bar.

Printed by Chalmers Reproservice
Gothenburg, Sweden 2014

Molecular-level Simulations of Cellulose Dissolution by Steam and SC-CO₂ Explosion

Faranak Bazooyar

Department of Chemical Engineering
Chalmers University of Technology

Swedish Centre for Resource Recovery
University of Borås

ABSTRACT

Dissolution of cellulose is an important but difficult step in biofuel production from lignocellulosic materials. Steam and supercritical carbon dioxide (SC-CO₂) explosion are two effective methods for dissolution of some lignocellulosic materials. Loading and explosion are the major processes of these methods. Studies of these processes were performed using grand canonical Monte Carlo and molecular dynamics simulations at different pressure/temperature conditions on the crystalline structure of cellulose. The COMPASS force field was used for both methods.

The validity of the COMPASS force field for these calculations was confirmed by comparing the energies and structures obtained from this force field with first principles calculations. The structures that were studied are cellobiose (the repeat unit of cellulose), water-cellobiose, water-cellobiose pair and CO₂-cellobiose pair systems. The first principles methods were preliminary based on B3LYP density functional theory with and without dispersion correction.

A larger disruption of the cellulose crystal structure was seen during loading than that during the explosion process. This was seen by an increased separation of the cellulose chains from the centre of mass of the crystal during the initial stages of the loading, especially for chains in the outer shell of the crystalline structure. The ends of the cellulose crystal showed larger disruption than the central core; this leads to increasing susceptibility to enzymatic attack in these end regions. There was also change from the *syn* to the *anti* torsion angle conformations during steam explosion, especially for chains in the outer cellulose shell. Increasing the temperature increased the disruption of the crystalline structure during loading and explosion.

Keywords: Molecular modelling, Cellulose, Steam explosion, SC-CO₂ explosion

LIST OF PUBLICATIONS

This thesis is based on the following papers which are referred by roman numerals in the text:

- I Bazooyar, F., Momany, F. A., & Bolton, K. (2012). Validating empirical force fields for molecular-level simulation of cellulose dissolution. *Computational and Theoretical Chemistry*, 984, 119-127.
- II Bazooyar, F., Taherzadeh, M., Niklasson, C., & Bolton, K. (2013). Molecular Modelling of Cellulose Dissolution. *Journal of Computational and Theoretical Nanoscience*, 10(11), 2639-2646.
- III Bazooyar F. and Bolton K. (2014). "Molecular-level simulations of cellulose steam explosion." *Quantum Matter*. Accepted.
- IV Bazooyar F., Bohlén M. and Bolton K. (2014). "Computational studies of water and carbon dioxide interactions with cellobiose." *Journal of Molecular Modelling*. Submitted.
- V Bazooyar F., Bohlén M, Taherzadeh M., Niklasson C. and Bolton K. (2014). "Molecular-level calculations of cellulose explosion using supercritical CO₂." Manuscript.

Publication by the author that is not included in this thesis:

- 1 Samadikhah, K., Larsson, R., Bazooyar, F., & Bolton, K. (2012). Continuum-molecular modelling of graphene. *Computational Materials Science*, 53(1), 37-43.

Contribution to the Publications

Faranak Bazooyar's contributions to the appended papers:

Paper I: FB performed all calculations and wrote the first draft of the paper.

Paper II: FB performed all calculations and wrote the first draft of the paper.

Paper III: FB performed all calculations and wrote the first draft of the paper.

Paper IV: FB performed the molecular mechanics calculations and wrote the first draft of the paper. The first principles calculations were performed by Dr. Martin Bohlén.

Paper V: FB performed all calculations and wrote the first draft of the paper. First principles calculations were performed by Dr. Martin Bohlén.

Faranak Bazooyar's contributions to the out of scope papers:

1 FB performed first principle calculations.

Table of Contents

1 Introduction.....	1
1.1 Biofuels from lignocellulosic materials.....	1
1.2 Lignocellulosic materials.....	2
1.2.1 Lignin.....	2
1.2.2 Hemicellulose.....	3
1.2.3 Cellulose.....	3
1.3 Pretreatment methods of lignocellulosic materials.....	6
1.3.1 Steam explosion.....	7
1.3.2 SC-CO ₂ explosion.....	8
2 Computational Methods.....	10
2.1 Quantum Mechanics.....	10
2.1.1 Born-Oppenheimer approximation.....	11
2.1.2 Hartree-Fock approximation.....	11
2.1.3 Post- Hartree-Fock methods.....	14
2.1.4 Density Functional Theory (DFT).....	14
2.2 Molecular Mechanics.....	17
2.2.1 COMPASS force field.....	18
2.2.2 Dreiding force field.....	19
2.2.3 Universal force field.....	20
2.3 Molecular Dynamics.....	20
2.4 Monte Carlo.....	21
3 Summary of Papers I-V.....	24
3.1 Papers I & II.....	24
3.2 Paper III.....	30
3.3 Paper IV.....	35
3.3.1 H ₂ O-cellobiose pair.....	35
3.3.2 CO ₂ -cellobiose pair.....	38
3.4 Paper V.....	40
3.4.1 Effect of temperature.....	43
3.4.2 Effect of pressure.....	44
4 Conclusions and Outlook.....	47
Future work.....	49
Acknowledgments.....	50
References.....	51

1 Introduction

One of the main concerns in the last few decades is substitution of fossil fuels by an appropriate energy supply. More than 80% of the world's energy demand is afforded by fossil fuels like oil, natural gas and coal [1]. Rapid growth of global population, limitation, depletion and high price of fossil fuels as well as climate changes due to emission of greenhouse gases, mainly due to using these fuels, have promoted the motivation of finding new sources. Several techniques have been developed to utilize lignocellulosic feedstock in biofuel production. Dissolution of cellulose is an important but difficult step in biofuel production from lignocellulosic materials (biomass). Steam and supercritical carbon dioxide (SC-CO₂) explosion are two effective pretreatment methods for this purpose. Loading and explosion are the major processes of these methods.

In this thesis, a molecular-level simulation study of these processes was performed using grand canonical Monte Carlo and molecular dynamics simulations at different temperature/pressure conditions on the crystalline structure of cellulose. The COMPASS force field was used in both methods.

1.1 Biofuels from lignocellulosic materials

Lignocellulosic biomass is a potential feedstock to substitute fossil fuels and is known as a good source for biofuels like biogas (bio-methane) and bioethanol (cellulosic ethanol). It is known as the most abundant organic material in the biosphere [2]. Besides the economic benefits of converting lignocellulosic biomass to biofuels, its sustainability and lower environmental impacts [3] have made it as a favoured feedstock. These materials can be provided in large-scale from inexpensive natural resources such as agricultural plant wastes, non-edible plant materials, paper pulp and industrial and municipal waste materials. Depending on the availability of feedstock, different materials are supplied in different areas [4]. An initial LCA analysis shows that, compared to gasoline, using sugar-fermented ethanol and bioethanol can reduce about 18-25% and 89% emission of greenhouse gases, respectively [5].

Production of bioethanol from lignocellulosic materials covers two main steps: hydrolysis and fermentation. During the hydrolysis, cellulose and hemicellulose of lignocellulosic biomass is decomposed by means of enzymes or chemicals to fermentable reducing sugars by

cutting the glycosidic linkages between glucose units. During the fermentation step microorganisms like yeasts or bacteria reduce the sugars to ethanol [4].

However, the bottleneck of the process is recalcitrance of lignocellulosic materials that hinders enzymatic hydrolysis; first due to presence of lignin that covers cellulose and hemicellulose and second because of high crystallinity of cellulose structure. These problems can be solved by adding a pretreatment step. Distillation and dehydration processes help to purify the produced bioethanol [6]. Figure 1 illustrates a very simple view of these processes.

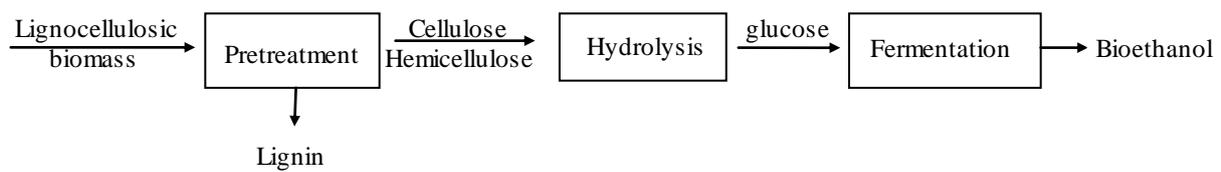


Figure 1- Simple view of different steps in bioethanol production from lignocellulosic materials.

1.2 Lignocellulosic materials

The main components of lignocellulosic biomass are lignin, hemicellulose and cellulose. Lignin and hemicellulose are in non-crystalline phase, where microfibrils of cellulose are ordered in crystalline phase. Inter-linkages (via glycosidic, esteric or etheric linkages) between lignin and hemicellulose as well as cellulose, give stiffness to the lignocellulosic structure [7]. Proteins, coumaric acid, ferulic acid and other polysaccharides such as pectin also can be found in the non-crystalline phase [8]. The relative quantities of these components varies in different feedstock [9].

1.2.1 Lignin

Lignin is an amorphous, three-dimensional branched polymer complex. It is an aromatic-containing hydrocarbon polymer mainly consisting of phenyl-propanes that gives stiffness to the structure of lignocellulosic materials, holds polysaccharides together and supports the structure against swelling [10]. Lignin is covalently linked to cellulose, directly or through a bridging molecule like hydroxycinnamate. Most covalent bonding between lignin and cellulose are ester-ether cross links [11].

1.2.2 Hemicellulose

Hemicellulose, which fills the empty spaces between cellulose microfibrils, has a random, amorphous and branched structure. It is not rigid and can be hydrolyzed easily [12]. Hemicellulose is a polymer containing five and six-carbon sugars (mostly substitute with acetic acid) and uronic acid. Common five-carbon sugars in hemicellulose are D-xylose and L-arabinose, and the six-carbon sugars are D-galactose, D-glucose, and D-mannose. About 25-30% of total dry wood weight is hemicellulose [13].

1.2.3 Cellulose

Cellulose, the main structural part of plant cells and biomass, is a linear polymer of β -1,4 D-glucose repeat units. Cellulose is known as the most abundant organic material worldwide that can be found not only in all plants, primitive and unicellular creatures such as bacteria, algae, etc., but in some parts of animal world like horse-tail and tunicin. Table 1 shows the different amounts of cellulose in some living cells [14].

Table 1. Cellulose content in different living cells

Living cells	Cellulose content %
Bark	20-30
Wood	40-50
Bamboo	40-50
Ramie	80-90
Cotton	95-99
Bacteria	20-30
Horse-tail	20-25

The properties of cellulose motivate its use in a variety of applications. Due to the excellent strength of cellulose, its applications in synthesized composites have been increasing; because of its flexibility, it is the main material in paper manufacturing; and its good tensile properties have increased its usage in textile fibres [15].

Cellulose is synthesized in the cell's plasma membrane [16]. Native cellulose is structured in fibrils with a high degree of polymerization. In general, these fibrils are known as microfibrils and each microfibrils consists of numbers of cellulose chain or elementary fibril and a mixture of hemicelluloses which covered its surface [16]. Cellulose chains are stabilized via van der Waals and hydrogen bonds that give strength and crystalline structure to the elementary fibrils. Depending on the source of cellulose, the number and dimension of

the microfibrils varies. The diameter of the microfibrils in plants cell walls is about 3-10 nm. However, in *Valonia* (an alga), *Acetobacter xylinum* (bacteria) and Ramie it is 18-20, 2 and 10-20 nm, respectively [17].

Four types of cellulose allomorphs have been identified: cellulose *I*, *II*, *III* and *IV*. Cellulose *I* is the main type of cellulose in nature which can be found in two forms, I_α and I_β . Cellulose *II* is the result of mercerization of cellulose *I*. Cellulose *III* is produced by treating cellulose *I* and *II* with amines such as liquid ammonia. Cellulose *IV* is the product of treating cellulose *I*, *II*, *III* with glycerol at high temperature. Cellulose I_α and I_β have different crystalline forms. I_α contains a single-chain triclinic unit cell, while I_β is in the form of two-chain monoclinic unit cell [18].

There are several reports of experimental, modelling and biological studies that identified cellulose microfibril structures [16, 19-21]. These studies show that different plant cell walls have different crystalline structures. In most of the proposed models, cellulose is assumed to organize the crystalline core structure that interacts with hemicellulose which forms the noncrystalline sheath.

Figure 2 shows four proposed crystal structures for plant cell walls of cabbage and onion (a), pineapple (b), apple cell walls (c) and Italian raygrass (d).

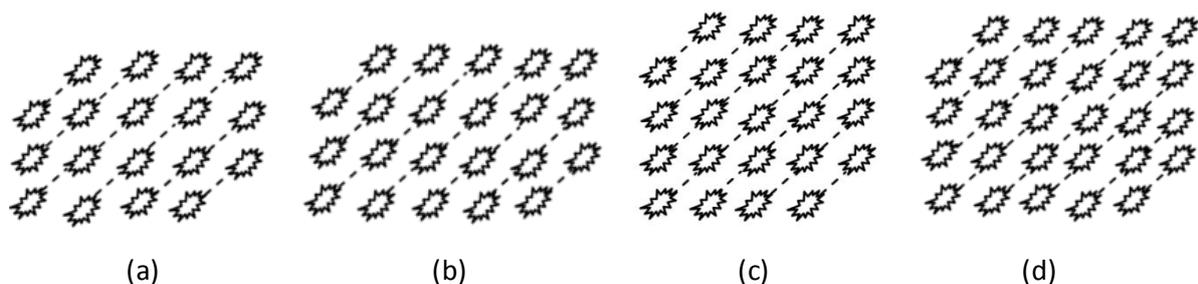


Figure 2. Order of cellulose chains in cabbage and onion (a), pineapple (b), apple cell wall (c) and Italian raygrass (d) cell walls.

Model (a) shows the order of 18 cellulose chains in cabbage and onion, containing 33% crystalline core chains (i.e., the chains that are not on the surface) while model (b) belongs to pineapple cell wall with 22 cellulose chains, containing 36% crystalline core chains [21]. The models for apple cell walls (c) and Italian raygrass (d) contain 23 and 28 cellulose chains with about 39 and 43% of core crystallinity, respectively. It is believed that the crystalline part of cellulose may be affected by the non-cellulosic materials present in the cell. Due to

the wide variety of non-cellulosic materials in different plants, the crystallite cellulose can have different dimensions [16, 19, 21].

The studies presented in this thesis are based on a new model proposed for cellulose I_β by Ding and Himmel in 2006 [16, 22]. In their model 36-cellulose chains are arranged in three layers as shown in Figure 3(a). Six crystalline core chains are surrounded by 12 sub-crystalline chains and 18 non-crystalline chains support them in the outer layer, in which the chains are fixed by hydrogen bonds (O3'- O5 and O2- O6').

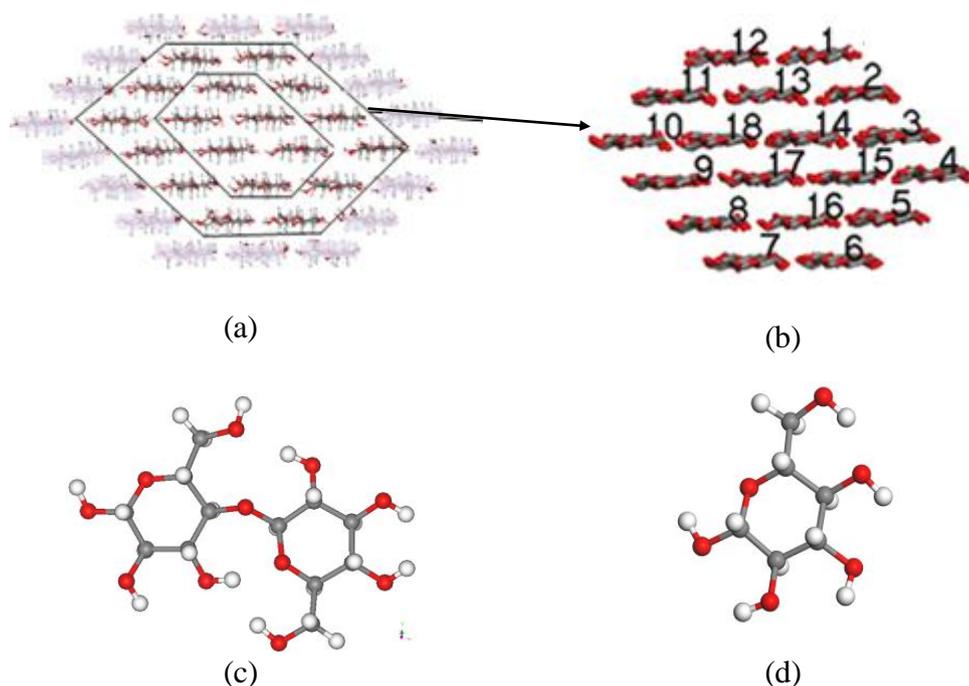


Figure 3. Illustration of a cellulose microfibril containing 36 cellulose chains (a), numbering of the cellulose chains in the crystalline structure of cellulose (b), cellobiose (*syn*) (c) and glucose (d) molecules.

Each cellulose chain is a linear polymer of β -1,4 D-glucose repeat units. Cellobiose is the shortest cellulose chain with two glucose repeat units. These structures are shown in Figures 3 (c) and (d).

Two conformers are known for cellobiose: *syn* and *anti*. Figures 3(c) and 4 show the *syn* and *anti* conformer of cellobiose. Numbering of atoms of cellobiose *anti* is also showed in Figure 4; φ_H is torsion angle between atoms H1-C1-O-C4'. In the *anti* conformer, φ_H either ψ_H (C1-O-C4'-C5') lies near -180° [23].

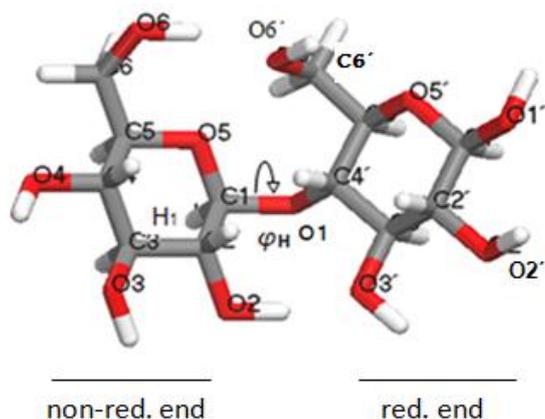


Figure 4. Numbers of atoms in the *anti* conformation of cellobiose, showing the non-reducing and reducing ends.

The *anti* conformer can be found in vacuum where the *syn* is generally found in hydrated environment and in crystalline structure. However, due to entropic effects, at elevated temperatures the *syn* conformer is preferred in both vacuum and hydrated environments [23-25]. This is discussed in Paper I.

1.3 Pretreatment methods of lignocellulosic materials

As mentioned, the most important part of lignocellulosic materials for biofuel production is its hemicellulose and cellulose content. However, in the plant cell wall, lignin has bolstered cellulose and hemicellulose in a way that access of enzymes to these molecules is tough. The pretreatment step is an essential issue in biofuel production process to increase the accessibility of enzymes to cellulose chains by increasing the accessible surface area by removing the lignin from the surface of microfibrils and decreasing the crystallinity of cellulose crystals or dissolving the cellulose. Several methods have been developed for this purpose. Selection of the best pretreatment method should be based on several features such as effectiveness of the method for the chosen feedstock to increase the digestibility of the components, avoid to produce inhibitors, process economic issues and the environmental impact [26].

Four main pretreatment methods for lignocellulosic materials have been identified. These methods are classified as biological, physical, chemical and physico-chemical pretreatment.

The main goal of biological pretreatment is degrading of lignin. Microorganisms such as some bacteria, and fungi such as white-, brown- and soft-rot fungi can degrade lignin and make hemicellulose more soluble. However, the rate of cellulose dissolution is slow [27].

Physical pretreatment includes different methods such as milling (ball milling, hammer milling, vibro energy milling, colloid milling and two-roll milling), ultrasound and irradiation methods (gamma-ray, electron-beam and microwave irradiation), hydrothermal methods, expansion, extrusion and pyrolysis to reduce the particle size and crystallinity [26, 28].

During chemical pretreatment, several chemicals such as sodium hydroxide, ammonia, sulphuric acid, phosphoric acid, sulphur dioxide, hydrogen peroxide and ozone are used to disrupt biomass structure through chemical reactions [26, 29].

The most important processes of physico-chemical pretreatment are steam explosion, ammonia fiber explosion (AFEX), *N*-Methylmorpholine *N*-oxide (NMMO), supercritical carbon dioxide (SC-CO₂) explosion, SO₂ explosion and liquid hot-water pretreatment [26-29].

A brief description of steam and SC-CO₂ explosion methods that are studied in this thesis can be found in the following sections.

1.3.1 Steam explosion

Steam explosion was developed by Mason in 1925 [30], and Babcock used it as a pretreatment method for bioethanol production in 1932 [31]. This method is a successful and economical method with low environmental impact that can be applied for pretreatment of several types of lignocellulosic biomass. Steam explosion is the most effective pretreatment method in commercial production of bioethanol from feedstock such as wheat straw [32].

Steam explosion pretreatment includes two main steps, steaming (steam loading) and explosion. During steaming step the lignocellulosic material is subjected to high pressure saturated steam. It is followed by a pressure drop to atmospheric pressure called the explosion step. The process helps to remove, depolymerise and dissolve lignin and hemicellulose into lower molecular-weight products as well as reduces the size and crystallinity of the cellulose structure.

Several experimental studies have investigated the effect of temperature- pressure and retention time during steam explosion of different feedstock such as aspen wood, sweet sorghum, wheat straw and hardwood chips. It is believed that during pretreatment the crystalline structure of cellulose becomes disordered by disruption of the ordered cellulose chains in the crystal. This increases the surface area and enhances the accessibility to enzymes that leads to more effective hydrolysis. Typically, an increase in temperature- pressure and/or residence time increases the disruption of the cellulose structure. Optimum conditions of temperature- pressure and retention time of the steam explosion procedure differ for different types of feedstock [32-40].

1.3.2 SC-CO₂ explosion

Among the supercritical fluids [41], supercritical carbon dioxide (SC-CO₂) is known as a green solvent for dissolving lignocellulosic biomass during biofuel production that was proposed by Zheng in 1995 [42]. CO₂ has a low critical temperature and pressure (31.1 °C and 1067 psi) and higher temperature and pressure give the CO₂ gas like mass transfer, liquid like solvating power and low viscosity characteristics [36, 43]. SC-CO₂ is economical, non-flammable, non-toxic, environmental friendly and easy to recycle [44]. In pulping production process, SC-CO₂ explosion enhances the penetration of chemicals. SC-CO₂ explosion has been widely used for treating different materials like corn stover, switchgrass, aspen, rice straw, southern yellow pine (SYP), cellulose-containing waste from cotton production, cotton fibre, Avicel and wheat straw [36, 45-53].

Experiments show that lignocellulosic materials answer to SC-CO₂ pretreatment differently. For example, the method is effective for Avicel and increases the glucose yield by 50% while pine wood does not show significant changes in its microstructure arrangement during SC-CO₂ pretreatment [42, 53].

Several experimental studies have worked on SC-CO₂ pretreatment with and without explosion for different feedstock at different conditions (temperature, pressure and residence time). For instance, temperature/ pressure/ residence time combinations of 40-110 °C/ 1450-4350 psi/ 15-45 min, 25-80 °C/ 1100-4000 psi/ 60 min, 160-210 °C/ 2900 psi/ 60 min, 80-160 °C/ 2900 and 3500 psi/ 10-60 min and 112-165 °C/ 3100 psi/ 10-60 min have been used for treatment of rice straw, bagasse, switch grass, corn stover and aspen and SYP, respectively [36, 46, 49, 52, 53]. Similar to steam explosion, SC-CO₂ explosion consists of two steps. In

the first step, the material is subjected to high temperature and pressure CO₂ (above its critical temperature and pressure) and in the second step, the pressure drops to atmospheric pressure rapidly. SC-CO₂ explosion helps in removing, dissolving and depolymerisation of lignin and hemicellulose into lower molecular-weight products as well as reduction of the crystallinity of the cellulose structure.

The effect of increasing temperature and pressure has been studied for some feedstock. While increasing temperature has been generally known as an effective parameter during this pretreatment [47, 52], pressure effect has been reported differently for different materials. It is believed that high pressure facilitates penetration of the fluid into the pores of the biomass. Studies of the effect of increasing temperature by Zheng *et al.* [52] showed that pretreatment of Avicel with subcritical CO₂ at 25 °C gave small yields of glucose whereas raising the temperature to 35°C increased it significantly. According to Narayanaswamy *et al.* [49], increasing pressure from 2500 to 3500 psi doubles the yield from corn stover compare to non-treated material, while increasing pressure from 3100 to 4000 psi shows negative effects for aspen [53].

Most of the experimental SC-CO₂ pretreatments use specific amount of water. It is believed that the presence of moisture during SC-CO₂ increases the enzymatic hydrolysis [52]. For example, dry lignocellulosic materials like aspen and SYP show no significant hydrolysis yield; but the presence of water between 40- 73% increases the sugar yield [53]. There are some reasons for the positive effect of presence of water, like formation of weak carbonic acid due to reaction of water and CO₂ which consequently hydrolyses some parts of hemicelluloses surrounding cellulose, breaking up cellulose-hemicellulose hydrogen bonds and the hydrogen bonds between cellulose microfibrils. Water also causes swelling of the biomass and prepares the material for more penetration of CO₂ into the pores of biomass. Enzymes are also more active in microaqueous environments [49, 51, 52].

2 Computational Methods

Computational studies complement experimental studies. Computational chemistry or molecular modelling uses a set of theories and techniques to solve chemical problems like molecular energies and geometries, transition states, chemical reactions, spectroscopy (IR, UV and NMR), electrostatic potentials and charges on a computer. Connection between theory and experiment helps to have a better understanding of vague and inconsistent results, optimization of design or progress of chemical processes and prediction of the results of difficult or dangerous experiments. However computational chemistry techniques are expensive and models cannot be computed accurately and need some approximations.

Computational chemistry is based on classical and quantum mechanics and covers a wide range of areas like statistical mechanics, cheminformatics, semi-empirical methods, molecular mechanics and quantum chemistry. Various methods have been developed to study the structures and the energies of molecules, either using quantum mechanics or molecular mechanics. Due to its cost, quantum mechanics can be used for small molecules or systems containing with a good accuracy while molecular mechanics can be applied to larger systems containing thousands of atoms [54].

In the following paragraphs a brief description of the methods that have been used in this thesis is given.

2.1 Quantum mechanics

Quantum mechanics describes the behaviour of the electrons mathematically and describes electron density using a wave function, $\Psi(r)$. Quantum mechanics is based on the time-independent Schrödinger equation (Eq.1):

$$H(r) \Psi(r) = E(r) \Psi(r) \quad \text{Eq. 1}$$

where $H(r)$ is the Hamiltonian operator, $\Psi(r)$ is the wave function and $E(r)$ is the total energy (kinetic and potential) of the system and r denotes the nuclear and electronic positions [55, 56]. When no external field is present, the Hamiltonian operator is given by factors related to interaction of electrons, interaction of nuclei and interaction between electrons and nuclei according to Eq. 2:

$$H(r) = -\frac{\hbar^2}{2m_{(el)}} \sum_i \nabla_i^2 + \frac{1}{2} \sum_i \sum_{j \neq i} \frac{e^2}{|r_i - r_j|} - \sum_{i,l} \frac{e^2 Z_l}{|r_i - R_l|} - \frac{\hbar^2}{2} \sum_l \frac{\nabla_l^2}{M_l} + \sum_{l,k} \frac{e^2 Z_l Z_k}{|R_k - R_l|} \quad \text{Eq. 2}$$

where the terms describe kinetic energy of electrons, the electrostatic potential between electrons i and j , the electrostatic potential between electron i and nucleus l , the kinetic energy of nucleus l , and the electrostatic potential between nuclei l and k , respectively [47].

Quantum mechanics has very accurate prediction of a single atom or molecule, but practically can solve equations for systems containing one electron like hydrogen atom. Systems with M atoms and N number of electrons have $3 \times (M + N)$ variables, and solving the Schrödinger equation for such systems needs some approximations [57].

2.1.1 Born-Oppenheimer approximation

The Born-Oppenheimer approximation [58] is one of the most fundamental approximations in chemistry that decouples the motions of electrons and the nuclei. Compared to electrons, nuclei are very heavy and the Born-Oppenheimer approximation considers the nuclei as fixed particles and implies that the electronic wave function is dependent on the nuclei position but independent of nuclei momenta. In this case the Schrödinger equation will be for electrons [47, 54]:

$$H_{(el)} \Psi_{(el)} = E_{(el)} \Psi_{(el)} \quad \text{Eq. 3}$$

where the Hamiltonian operator is:

$$H(r) = -\frac{\hbar^2}{2m_{(el)}} \sum_i \nabla_i^2 + \frac{1}{2} \sum_i \sum_{j \neq i} \frac{e^2}{|r_i - r_j|} - \sum_{i,l} \frac{e^2 Z_l}{|r_i - R_l|} \quad \text{Eq. 4}$$

and the total potential energy of the molecule is calculated according to Eq. 5:

$$E_{\text{total}} = E_{(el)} + \sum_{l,k} \frac{e^2 Z_l Z_k}{|R_k - R_l|} \quad \text{Eq. 5}$$

2.1.2 Hartree-Fock approximation

The Hartree-Fock approximation [47, 54-56] is a useful approximation for the many-electron Schrödinger equation that gives a correct picture of electron motions by considering the electrons as independent particles.

The Hartree-Fock approximation describes the electrons as orbitals, limited to molecular orbitals (MO), Ψ .

$$\Psi = \psi_1 \psi_2 \psi_3 \dots \psi_N \quad \text{Eq. 6}$$

where ψ_i is single-electron orbitals.

It is assumed that electrons move within an average field of all the other electrons and that the total wave function can be written in the form of a single determinant called the Slater-determinant (SD).

Considering the antisymmetry principle (Pauli Exclusion Principle) [59, 60], the N-electron wave function is defined as a product of N one-electron wave functions, $\psi_i(i)$. The Slater-determinant creates molecular orbitals (MO) as a linear combination of atomic orbitals (LCAO).

$$\Psi = \Phi_{SD} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(1) & \psi_2(1) & \dots & \psi_N(1) \\ \psi_1(2) & \psi_2(2) & \dots & \psi_N(2) \\ \vdots & \vdots & & \vdots \\ \psi_1(N) & \psi_2(N) & \dots & \psi_N(N) \end{vmatrix} \quad \text{Eq.7}$$

Atomic orbitals are linear combinations of a set of basis functions (ϕ) known as basis sets:

$$\psi(r) = \sum_k C_k \phi_k(r) \quad \text{Eq. 8}$$

where C_k is the wave function's coefficient.

In principle, an exact molecular orbital can be achieved by choosing a complete basis set and if the basis set is large enough, this could be a fairly accurate approximation [61]. Two main basis sets that are developed for calculating the molecular orbitals are Slater type orbitals (STO) and Gaussian type orbitals (GTO) [47, 56].

The mathematical description of a Slater type orbital (STO) is given in Eq. 9:

$$\eta^{STO} = N r^{n-1} e^{-\zeta r} Y_{lm}(\theta, \phi) \quad \text{Eq. 9}$$

where N is a normalization factor, n is the quantum number, ζ corresponds to the orbital exponent, r is the radius and Y_{lm} describes the angular part of the function. However, Gaussian type orbitals (GTO) is given as the mathematical form in Eq. 10:

$$\eta^{GTO} = Nx^l y^m z^n e^{-\alpha r^2} \quad \text{Eq.10}$$

where N is a normalization factor, α corresponds to the orbital exponent, r is the radius and l , m , n are quantum numbers such that $L = l + m + n$ gives the angular momentum of η . A linear combination of Gaussian functions or “Contracted Gaussians” (CGs) in the form of STO-MG are widely used that approximate Slater-type orbitals (STOs) by M primitive Gaussians (GTOs). STO-3G is called a “minimal basis set”, that is simplest possible atomic orbital that has the lowest basis functions.

Extending the basis sets is possible by adding the *double zeta*, *triple* or *quadruple zeta* to the basis sets, so that the set of functions are doubled, tripled or quartet. *Split-valence* basis sets apply two or three more basis functions to each valence orbital. Addition of *polarization* (*) and *diffusion* (+) functions to the basis sets can extend the basis sets even more. *Polarization* functions add orbitals higher in energy than the valence orbitals of each atom, e. g., adding the *p-functions* for hydrogen or *d-functions* for the first-row elements of the periodic table. *Diffusion* functions allow the electrons to be distributed far from the ionic positions. This function is useful for description of systems where the electrons need to move far from nuclei, like anions [62].

Hartree-Fock is a molecular orbital approximation that gives a set of coupled differential equations but cannot explain the correlation between electrons. It gives good description for many equilibrium geometries in the ground state but cannot describe thermochemistry where bonds are broken or formed. Using adequate basis sets, Hartree-Fock wave function can predict 99% of the total energy where the 1% remaining energy belongs to correlation interactions between electrons. The post-Hartree-Fock methods and Density Functional Theory (DFT) are useful methods that give more flexibility to Hartree-Fock methods. The so-called second-order Møller-Plesset model (MP2) is a commonly used method that describes thermochemistry where bonds are broken or formed [56].

2.1.3 Post-Hartree-Fock methods

Configuration interaction (CI) [63] and Møller-Plesset (MP) [64] are two of the useful methods that improve the flexibility of the Hartree-Fock through mixing the ground-state wave functions with excited-state wave functions. They also give a good description of electron correlations. However they are more expensive than Hartree-Fock methods. The correlation energy (E_C) is the difference between the real energy of the molecule and the energy calculated by Hartree-Fock methods.

$$E_C = E_{real} - E_{HF} \quad \text{Eq. 11}$$

MP methods are based on perturbation theory. Simply, the Møller-Plesset model mixes ground-state and excited-state wave functions together, i.e. when MP2 is applied, one or two electrons from occupied orbitals in the Hartree-Fock configuration will move to the unoccupied orbitals (excited state) to calculate the contribution to the correlation energy. Different orders of MP methods give different description of electronic structures. If MP0 considers electron repulsion in one molecular orbital, MP1 can be regarded as the Hartree-Fock wave function, considering an average of inter-electronic repulsions [64]. Higher orders of MP by addition of more functions can improve the calculations and give more accurate correlation energies but require large computational resources. MP2 methods account for ~ 80-90% of the correlation energy, while higher orders of MP like MP3 and MP4 account for ~ 90-95% and ~ 95-98%, respectively [56, 65].

2.1.4 Density Functional Theory (DFT)

DFT [47, 66] is first principles method based on the electron density $\rho(r)$, Eq. 12,

$$\rho(r) = N_{el} \int \dots \int |\Psi_0(r_1 r_2, \dots, r_{N_{el}})|^2 dr_2 \dots dr_{N_{el}} \quad \text{Eq. 12}$$

where N_{el} denotes the total number of electrons.

The idea of DFT theory was born in the late 1920s by Thomas-Fermi model, but the density functional theory as we know it today was introduced in the contributions by Hohenberg-Kohn (1964) and Kohn-Sham (1965). DFT methods can be applied to larger systems than the post-Hartree-Fock methods. Hohenberg and Kohn showed that properties and the ground

state energy of a system can be defined solely by the electron density [67]. In DFT method, energy functional can be calculated as Eq. 13:

$$E[\rho(r)] = \int V_{ext}(r)\rho(r)dr + F[\rho(r)] \quad \text{Eq. 13}$$

where $V_{ext}(r)$ is the external potential due to the Coulomb interaction between electrons and nuclei, and $F[\rho(r)]$ is kinetic energy of the electrons and the energy obtained from interaction of electrons. The problem with this definition was that the function $F[\rho(r)]$ was not clear; one year later Kohn and Sham extended the equation to Eq. 14 :

$$F[\rho(r)] = E_{KE}[\rho(r)] + E_H[\rho(r)] + E_{XC}[\rho(r)] \quad \text{Eq. 14}$$

in which $E_{KE}[\rho(r)]$ is kinetic energy of non-interacting electrons , $E_H[\rho(r)]$ is Coulombic energy between electrons and $E_{XC}[\rho(r)]$ is the energy due to exchange and correlation. The kinetic energy of non-interacting electrons, $E_{KE}[\rho(r)]$, can be obtained according to Eq. 15:

$$E_{KE}[\rho(r)] = \sum_{i=1}^N \int \psi_i(r) \left(-\frac{\nabla^2}{2} \right) \psi_i(r) dr \quad \text{Eq. 15}$$

$E_H[\rho(r)]$ or Hartree electrostatic energy, which is the electrostatic energy due to interaction between charge densities, is given in Eq. 16 :

$$E_H[\rho(r)] = \frac{1}{2} \iint \frac{\rho(r_1)\rho(r_2)}{|r_1-r_2|} dr_1 dr_2 \quad \text{Eq. 16}$$

Considering the Coulomb interaction between electrons and nuclei, $V_{ext}(r)$, Eq. 13 can be written as:

$$E[\rho(r)] = \sum_{i=1}^N \int \psi_i(r) \left(-\frac{\nabla^2}{2} \right) \psi_i(r) dr + \frac{1}{2} \iint \frac{\rho(r_1)\rho(r_2)}{|r_1-r_2|} dr_1 dr_2 + E_{XC}[\rho(r)] - \sum_{A=1}^M \int \frac{Z_A}{|r-R_A|} \rho(r) dr \quad \text{Eq. 17}$$

where M is the number of ions in the system.

The exchange-correlation functional, $E_{XC}[\rho(r)]$, is developed by several approaches. The simplest one is the Local Density Approximation (LDA) [47, 68, 69] which states that exchange-correlational energy is only affected by the local electron density and is given by Eq. 18:

$$E_{XC}^{LDA}[\rho(r)] = \int \rho(r) \epsilon_{XC}(\rho(r)) dr \quad \text{Eq. 18}$$

$\epsilon_{XC}(\rho)$ can be obtained from simulations of a homogeneous electron gas.

Adding electron spins α and β to Eq. 18 yields a modified version of LDA called Local Spin Density approximation (LSD):

$$E_{XC}^{LSD}[\rho(r)] = \int \rho(r) \epsilon_{XC}(\rho_{\alpha}(r), \rho_{\beta}(r)) dr \quad \text{Eq. 19}$$

An improved method beyond LDA is Generalized Gradient Approximation (GGA) [47, 70] which includes both local electron density and the gradient of the charge density. The gradient term shows the rate of density changes and is known as non-local functional.

$$E_{XC}^{GGA}[\rho(r)] = \int f(\rho(r), |\nabla\rho(r)|) dr \quad \text{Eq. 20}$$

PW91 [71] and PBE [70] are two general GGA functionals. Further improvement to the GGA is possible by including a certain amount of Hartree-Fock (HF) exchange. These functionals are known as hybrid functionals [72]. One of the most popular functionals, that has been used in the calculations presented here, is B3LYP (Becke three-parameter exchange and the Lee–Yang–Parr correlation functionals) [73-76] that includes LSD, Hartree-Fock and Becke (B) exchange functionals and LSD and Lee-Yang-Parr (LYP) correlation functionals [47]:

$$E_{XC}^{B3LYP} = (1 - a)E_X^{LSD} + aE_X^{HF} + bE_X^B + dE_C^{LYP} + (1 - d)E_C^{LSD} \quad \text{Eq. 21}$$

a , b and d are equal to 0.2, 0.72 and 0.81, respectively; the values are fitted to the empirical data like atomization energies, ionization potentials and proton affinities.

Nowadays, Kohn-Sham-DFT is one of the most common methods for calculating electronic structures of molecules in quantum chemistry, but one of the main challenges for DFT is its deficiency to find a correct description of dispersion interactions for long-range van der Waals forces. Many approaches have been proposed for the inclusion of dispersion interactions [77-79] [67]. One of the useful methods is DFT-D with B3LYP and PBE functionals that was developed by Grimme:

$$E_{\text{DFT-D}} = E_{\text{DFT}} + E_{\text{disp}} \quad \text{Eq. 22}$$

where E_{disp} is an empirical dispersion correction including an energy term of the $\frac{1}{R_{ij}^6}$:

$$E_{\text{disp}} = -s_6 \sum_{i=1}^{N-1} \sum_{j=i+1}^N \frac{C_6^{ij}}{R_{ij}^6} f_{\text{damp}}(R_{ij}) \quad \text{Eq. 23}$$

N is the number of atoms, C_6^{ij} is the dispersion coefficient for ij atom pair, s_6 is the global scaling factor that depends on the DFT method and R_{ij} is the distance between atoms i and j .

When the sum of van der Waals radii is R_r , the damping function is given by:

$$f_{\text{damp}}(R_{ij}) = \frac{1}{1 + e^{-d(R_{ij}/R_r - 1)}} \quad \text{Eq. 24}$$

In this thesis, the DFT-D method has been used in the Papers IV and V.

2.2 Molecular Mechanics

Molecular mechanics (MM) is widely used for conformational analysis. It is less accurate but more economical than quantum mechanics methods and can be applied to large systems like organic materials (oligonucleotides, hydrocarbons and peptides) and in some cases to metallo-organics and inorganics. In molecular mechanics, there is no reference to electrons and molecules are considered as a collection of balls joined by springs; the energy of a system (Eq. 25) caused by the geometry of the molecules in terms of a sum of contributions of bonding (stretching, bending, torsion and inversion) and non-bonding (electrostatic and van der Waals) energies between atoms [80, 81].

$$E_{\text{total}} = E_{\text{str.}} + E_{\text{bend.}} + E_{\text{tors.}} + E_{\text{inv.}} + E_{\text{el.}} + E_{\text{vdW}} \quad \text{Eq. 25}$$

These energy terms are illustrated in Figure. 5:

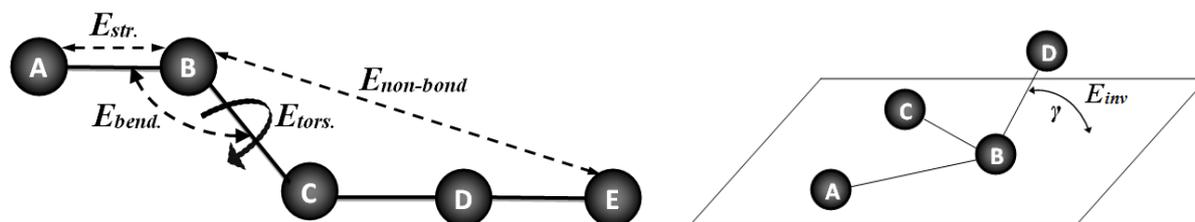


Figure 5- Schematic view of contributions of stretching, bending, torsion, inversion and non-bonding energies in a molecule

Molecular mechanics uses a set of mathematical functions that are so-called force fields to describe the potential energy of molecular systems, thus generally molecular mechanics methods denoted as force field methods [56, 65]. The parameters of force field functions are fitted to both quantum mechanics and empirical data to describe entire types of atoms in the molecules; hence the choice of the molecular model and the force field is an essential step in prediction of the geometry and conformation of the molecules. Several groups of force fields have been developed for this purpose; classical force fields, second-generation force fields, special-purpose force fields and rule-based force fields are main groups of force fields [82].

Molecular mechanics employs one or more minimization method to find the local minima on the potential energy surface (PES). Steepest descent [83], conjugate gradient [84] and Newton-Raphson [85] are a number of well-known algorithms that are mostly used by molecular mechanics to find the geometry of the structure related to the local minima [65].

Three force fields, COMPASS, Dreiding and Universal that are applicable for polymers have been used in this thesis and will be discussed briefly. The COMPASS force field belongs to second-generation force fields while Dreiding and Universal (UFF) fit in the Rule-based force fields.

2.2.1 COMPASS force field

COMPASS (Condensed-phase Optimized Molecular Potentials for Atomistic Simulation Studies) [86] is an *ab initio* force field that is based on the PCFF (Polymer Consistent Force Field) [87]. The COMPASS force field gives good prediction of geometry, conformational, vibrational, and thermophysical properties of a broad range of molecules, especially polymers in both isolation and condensed phases [88].

The COMPASS force field potential energy expression is given in Eq. 26. The first four terms account for bond stretch, angle bend, out-of-plane torsion and out-of-plane wag energies, terms five to ten are for cross-coupling and the last two terms show electrostatic interactions and van der Waals energies, respectively.

$$\begin{aligned}
E_{total} = & \sum_b [K_2(b - b_0)^2 + K_3(b - b_0)^3 + K_4(b - b_0)^4] + \sum_{\theta} [K_2(\theta - \theta_0)^2 \\
& + K_3(\theta - \theta_0)^3 + K_4(\theta - \theta_0)^4] + \sum_{\varphi} K_1(1 - \cos \varphi) \\
& + K_2(1 - \cos 2\varphi) + K_3(1 - \cos 3\varphi)] + \sum_x K_2 \gamma^2 \\
& + \sum_{b,b'} K(b - b_0)(b' - b'_0) + \sum_{b,\theta} K(b - b_0)(\theta - \theta_0) \\
& + \sum_{\theta,\theta',\varphi} K(\theta - \theta_0)(\theta' - \theta'_0) \cos \varphi \\
& + \sum_{b,\varphi} K(b - b_0)[K_1 \cos \varphi + K_2 \cos 2\varphi + K_3 \cos 3\varphi] + \sum_{b,\varphi} K_1(\theta \\
& - \theta_0)[K_1 \cos \varphi + K_2 \cos 2\varphi + K_3 \cos 3\varphi] + \sum_{\theta,\theta'} K(\theta - \theta_0)(\theta' - \theta'_0) \\
& + \sum_{i,j} Q_i Q_j / R_{ij} + \sum_{i,j} \epsilon_{ij} [2(R_e/R_{ij})^9 - 3(R_e/R_{ij})^6]
\end{aligned}$$

Eq. 26

b , θ , φ and γ are bond length, angle, torsion angle and out-of-plane wag or inversion angle, respectively, Q_i and Q_j are atomic charges and R_{ij} is the interatomic separation. Parameters of b_0 and θ_0 are equilibrium values and K and K_1 - K_3 are constants. However parameters for the intramolecular terms as well as the atomic charges have been fit to *ab initio* data, and those for the intermolecular terms are fit to empirical data [86, 89].

2.2.2 Dreiding force field

The Dreiding force field [90] has been fitted to biological, organic and some inorganic molecules where atomic hybridization has been considered when fitting the parameters and force constants. The Dreiding force field calculates the potential energy by considering bonded energy i.e., bond stretch, angle bend, out-of-plane torsion and inversion (out-of-plane wag angle), and non-bonded energy i.e. electrostatic interactions, the Van der Waals interactions and hydrogen bond energies. For calculation of bond stretch energy, in some applications harmonic functions can be replaced by Morse functions that are more accurate [91, 92]. The van der Waals interactions are based on the Lennard-Jones potential. Hydrogen bond energy is calculated according to Eq. 27:

$$E_{Hb} = D_e [5(R_e/R_{DA})^{12} - 6(R_e/R_{DA})^{10}] \cos^4(\theta_{DHA}) \quad \text{Eq. 27}$$

D_e is the energy for bond dissociation; R_e is equilibrium distance and R_{DA} is the length between electron donor and acceptor atoms. Θ_{DHA} is the bond angle between atoms A, D and H (hydrogen) [90].

2.2.3 Universal force field

The Universal force field (UFF) [91] is a biological force field which can cover the entire elements of the periodic table. This force field is reasonably precise for geometry estimation and energy calculation of organic conformers, metal complexes and organo-metallic molecules. The Universal energy expressions are the sum of bonding and non-bonding energies. Like Dreiding, both harmonic oscillator and Morse functions can be used for calculation of bond stretch energy. Angular bend is given by General Fourier extension and inversion term is according to *Cosine Fourier* expansion. Non-bond interactions are introduced as van der Waals (Lennard-Jones potential) and electrostatic interaction [91].

In brief, molecular mechanics is a rapid and simple force field based method that can be applied for systems comprising several thousand atoms but is limited to finding particular conformation in equilibrium state. However, it is not able to explain the transition state and time evolution of the system.

2.3 Molecular Dynamics

Molecular dynamics focuses on molecules in motion through the study of nuclear motions by step-by-step solving the Newton's equation of motion (Eq. 28) to calculate the trajectory of all atoms [55, 65, 93]:

$$F_i = m_i \frac{d^2 r_i(t)}{dt^2} \quad \text{Eq. 28}$$

where F_i is the force acting on particle i at time t , m_i is the mass of the particle and r_i is the position vector of i th particle. The trajectories show the position, velocity and acceleration of the particles in the system under a period of time. Choosing an adequate time step is needed to have an acceptable time evolution. A too short time step is very expensive and covers a limited phase space, but a big time step makes the system unstable and leads to errors. In molecular dynamics, parameters like positions, velocities and also accelerations are often approximated as Taylor expansion (Eq. 29) and (Eq. 30) using time step δt :

$$r_i(t + \delta t) = r_i(t) + \delta t v_i(t) + \frac{\delta t^2}{2} a_i(t) + \dots \quad \text{Eq. 29}$$

$$v_i(t + \delta t) = v_i(t) + \frac{\delta t}{2} [a_i(t) + a_i(t + \delta t)] + \dots \quad \text{Eq. 30}$$

r_i shows the position, v_i stand for velocity and a_i is acceleration of particle i . An integrator is needed to project the trajectory over a small time step δt [93]. There are several integrators for this purpose like the Verlet, Verlet leap frog, Gear fixed time step, Gear variable time step, Runge-Kutta, and Gauss-Radau algorithms [94]. The Verlet algorithm (Eq. 31) is a *time-reversible* algorithm [95] that, because of its simplicity and stability, is a widely used integration algorithm:

$$r_i(t + \delta t) = 2r_i(t) - r_i(t - \delta t) + \frac{F(t)}{m} \delta t^2 \quad \text{Eq. 31}$$

and the velocity will be:

$$v_i(t) = [r_i(t + \delta t) - r_i(t - \delta t)]/2\delta t \quad \text{Eq. 32}$$

In brief, in MD simulation, calculating time evolution is performed by numerically integrating Newton's equation of motion for interacting atoms.

Molecular dynamics describes the potential energy of molecules by using an appropriate molecular mechanics force field. Choosing an ensemble such as NVT, NpT or NVE will identify which parameters are constant during the simulation. NVT keeps the number of particles, volume and temperature constant, while in NpT the number of particles, pressure and temperature remain unchanged and in NVE as well as the number of particles and volume, energy is kept constant too. NVT, NpT and NVE ensembles have been used in this thesis for study of steam and SC-CO₂ explosion.

2.4 Monte Carlo methods

Monte Carlo (MC) is a stochastic method based on probabilities, where random numbers are used to create a sequence of possible configurations.

In the MC methods, the particles can be positioned, for example, by transition between points, or random insertion-deletion of particles. The new conformations are then accepted or

rejected according to some filter. Many states are generated, and the energy of each conformation is calculated, often using a molecular mechanics force field.

However random numbers decide how atoms or molecules move to generate new conformations or geometric arrangements. In other words, configurations are chosen randomly and then their impact is weighted with $\exp(-\Delta E / k_B T)$, where ΔE is the energy difference between two configuration, k_B is Boltzman constant and T is temperature.

Metropolis method is a development of Monte Carlo method [96]. According to this method, configurations are chosen with a probability distribution of $\exp(-\Delta E / k_B T)$ and then weighted equally.

Simply, if i is the configuration of a system of particles, the Metropolis Monte Carlo algorithm generates a new configuration j with a transition probability of $P(i \rightarrow j)$:

$$P(i \rightarrow j) = \exp(-\Delta E_{ij} / k_B T) \quad \text{Eq. 33}$$

where ΔE_{ij} is the energy difference between configuration i and j . If the energy of the new configuration j is lower than the old one (i), i.e., $\Delta E_{ij} \leq 0$, the new configuration j is accepted for the new positioning; but if j has a higher energy than i or $\Delta E_{ij} > 0$, $P(i \rightarrow j)$ is compared to a random number ζ where $0 < \zeta < 1$; if $P(i \rightarrow j) > \zeta$, the new configuration is accepted, otherwise j is rejected and a new configuration is generated [55, 56, 65, 97].

Metropolis Monte Carlo is a faster method with high quality of the statistics that ensures that accepted structures have a Boltzmann distribution. However, the magnitude of the particle displacements should be selected carefully, since a small change in displacement leads to high acceptance, but is a slow procedure. However big change is faster but the probability of acceptance is lower and the number of sampled configurations is few [56].

Grand canonical Monte Carlo (GCMC) simulation is an appropriate tool to, for example, study physical interactions of fluids with solid systems. An example is when one simulates a solid sorbent phase and a liquid or gas phase at equilibrium with a specified chemical potential [98].

In grand canonical Monte Carlo, which has a partition function denoted by $\Xi(\mu, V, T)$, the volume, temperature and chemical potential are conserved. The system is open and the

numbers of particles are allowed to fluctuate by discontinuously creating new particles and destroying them during the simulation. This helps to minimize ergodic difficulties of the system.

In the grand canonical ensemble, the probability of a configuration m , is given by Eq. 34:

$$\rho_m = CF(\{N\}_m) \exp[-\beta E_m] \quad \text{Eq. 34}$$

where C is an arbitrary normalization constant, $\beta = \frac{1}{T k_B}$, E_m is the total energy of configuration m , and the function $F(N)$ is calculated by Eq. 35:

$$F(N) = \left(\frac{(\beta f V)^N}{N!} \right) \cdot e^{-\beta N \mu} \quad \text{Eq. 35}$$

where, f is the fugacity, μ is the intramolecular chemical potential and N is the loading of the component. Probability of accepting the proposed configuration n is then calculated according to Eq. 36:

$$P_{mn} = \min \left[1, \frac{F(\{N\}_n)}{F(\{N\}_m)} \cdot e^{-\beta(E_n - E_m)} \right] \quad \text{Eq. 36}$$

3 Summary of Papers I-V

Molecular-level studies of dissolution of crystalline structure of cellulose during steam and supercritical carbon dioxide (SC-CO₂) were performed using grand canonical Monte Carlo and molecular dynamics. For both simulations, COMPASS force field was used. The validity of this force field for these systems was tested by comparing the energy and structures obtained from quantum and molecular mechanics. These studies are presented in Papers I to V.

Quantum mechanics calculations were performed in the GAUSSIAN 09 program package at Neolith, AKKA and C3SE, and GAMESS-US program at the high performance computer cluster Kalkyl at UPPMAX. Molecular mechanics, Monte Carlo and molecular dynamics calculations were performed using the Materials Studio package version 6.0 (Accelrys Software Inc).

3.1 Papers I & II

Paper I presents the results from the COMPASS, Dreiding and Universal force fields for studies of cellulose systems. These force fields are widely used for studies of polymeric systems. The validity of the force field is tested by comparing structures and energies obtained by the force fields with data obtained from first principles calculations. The use of first principles methods requires that the comparison is limited to small systems of importance to cellulose, and we therefore focus on glucose and cellobiose molecules as well as their interaction with water molecules. The results indicate that the COMPASS force field is preferred over the Dreiding and Universal force fields for studying dissolution of large cellulose structures.

Figure 6 illustrates the annealed structure of cellobiose obtained from each of the three force fields, as well as the corresponding structures obtained after B3LYP/6-311++G** minimization. Similar structures are obtained after geometry optimization with the other DFT and MP2 calculations. It is clear that the annealed (and first principles optimized) cellobiose structure depends on the force field used for the annealing. The annealed structures obtained from COMPASS did not show significant change during the subsequent optimization with the first principles methods. For example, when performing geometry optimization with B3LYP/6-311++G** the bond lengths changed by less than 0.02 Å and the change in bond

angle was less than 2 degrees. The cellobiose structure obtained from Dreiding shows a larger change during the subsequent optimization with DFT (where an OH group rotates). Cellobiose structures obtained from Universal also show large changes during DFT geometry optimization. Together with the relative energies of the first principles methods discussed below with respect to Table 2, this indicates that the COMPASS force field yields the preferred cellobiose structures.

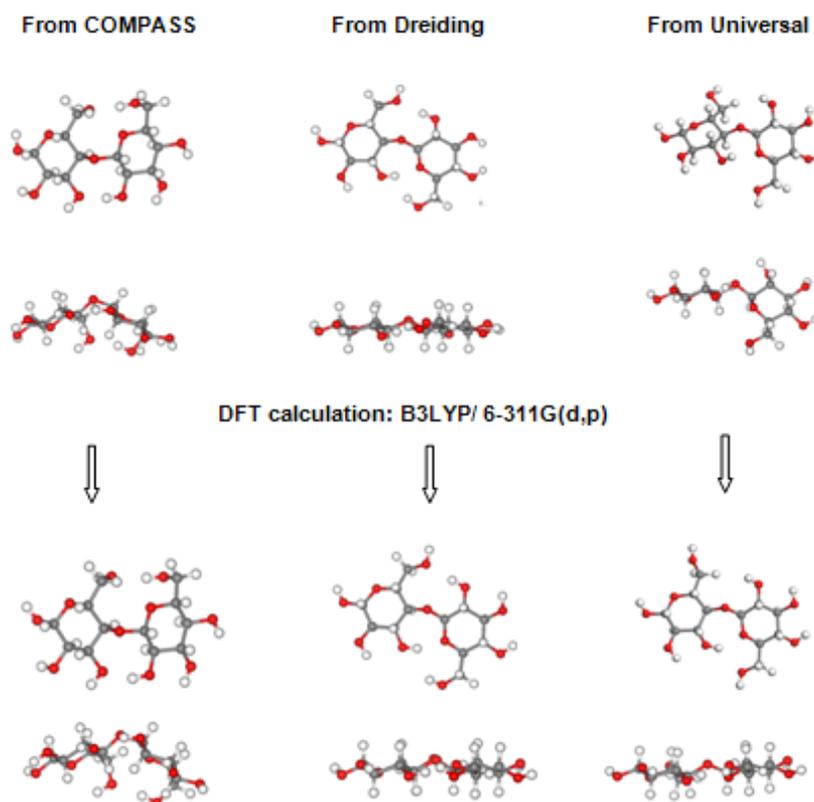


Figure 6- Cellobiose molecular structures (top and side view) obtained after annealing with the COMPASS, Dreiding and Universal force fields (upper three figures) and after further geometry optimization with B3LYP/6-311++G**.

The three force fields yield different structures for the cellobiose molecule. Figure 6 also shows that there is a large difference in the cellobiose structures obtained from the different force fields. Dreiding and Universal force fields yield *syn* structures whereas COMPASS yields an *anti* structure. The torsions are $\varphi_H = 30.1^\circ$ (φ_H is defined in Figure.4) for the Dreiding force field, $\varphi_H = 51.4^\circ$ for the Universal force field and $\varphi_H = -179^\circ$ for the COMPASS force field. More structural details like torsion angles that exemplify differences in the structures can be found in appended Papers I and II.

Relative energies of the cellulose molecules that were optimized using the different quantum mechanics methods and basis sets are listed in Table 2. The energies in columns 3, 4 and 5 are obtained when the initial glucose structure is from the COMPASS, Dreiding and Universal force fields, respectively.

Table 2- Relative energies (kcal/mol) for the cellobiose molecule obtained after geometry optimization with the first principles methods. Energies are given relative to the results obtained when the initial structure is from the COMPASS force field. ‡ These results are from MP2/6-311++G**//B3LYP/6-311++G** calculations.

Method	Basis-set	Initial structure from COMPASS	Initial structure from Dreiding	Initial structure from Universal
MP2	6-311G	0.0	10.3	11.7
	6-311G**	0.0	9.3	8.5
	‡ 6-311++G**	0.0	7.7	8.2
B3LYP	6-311G	0.0	8.6	11.0
	6-311G**	0.0	6.9	11.6
	6-311++G**	0.0	5.1	5.0
PBE	6-311G	0.0	9.3	11.5
	6-311G**	0.0	7.3	12.3
	6-311++G**	0.0	5.4	6.2
B3PW91	6-311G	0.0	8.7	10.7
	6-311G**	0.0	6.8	11.1
	6-311++G**	0.0	5.2	5.7

Table 2 also shows that all methods and basis sets yield the lowest energy for the cellobiose structure that was obtained from annealing using the COMPASS force field. There is a rather large energy difference between the structures obtained from the various force fields. For example, the B3LYP/6-311++G** energy from the COMPASS structure is 5.1 and 5.0 kcal/mol lower than the structures obtained from the Dreiding and Universal force fields, respectively.

This conclusion is also supported by comparing the energies of the structures optimised by B3LYP. The energies of the cellobiose structures that are geometry optimised using B3LYP and when starting with structures obtained after annealing with the COMPASS, Dreiding and Universal force fields are -814729.84, -814724.72 and -814724.83 kcal mol⁻¹, respectively. Hence, the cellobiose structure obtained from the COMPASS force field not only has a structure that is in good agreement with B3LYP, but it is also the energetically preferred structure.

Similar results were obtained for glucose structures. All first principles methods and basis sets yield the COMPASS structure as the lowest energy structure that can be found in the original appended Paper. This indicates that COMPASS is the preferred force field when studying the glucose and cellobiose molecule.

Since the COMPASS force field is preferred, it was used in a more detailed comparison for the cellobiose molecule hydrated with between 0 and 4 water molecules. The comparison was made with relative energies and structures calculated using the B3LYP/6-311++G** method that were initially geometry optimized using AMB02C empirical force field. The COMPASS force field yields results for cellobiose that are in excellent agreement with these DFT results. The quantitative agreement seen for cellobiose deteriorates as more water molecules are added to the system.

According to previous studies, the *syn* conformers have a larger entropy contribution than the *anti* conformers, and may therefore be thermodynamically stable at higher temperatures. MD simulations at various temperatures under NpT conditions at 1 bar showed that at very low temperatures (e.g., 100 K) the conformer that is observed in the simulations depends on the initial cellobiose structure; i.e., the *syn* (*anti*) conformer is seen when starting with the *syn* (*anti*) structure since the energy barrier for isomerization has not been passed within the simulation time. However, this is not the case for higher temperatures.

Figure 7 shows the distribution of ϕ_H when the cellobiose (in vacuum) initially had a *syn* conformation and the temperature is 298 K. The data shown in the figure were obtained from the last part of the simulation, once the *syn* had isomerized to *anti*. Note that the same results were obtained at 325 K, and in neither case did the *anti* revert back to the *syn*. Hence, at these temperatures the *anti* conformer is thermodynamically stable.

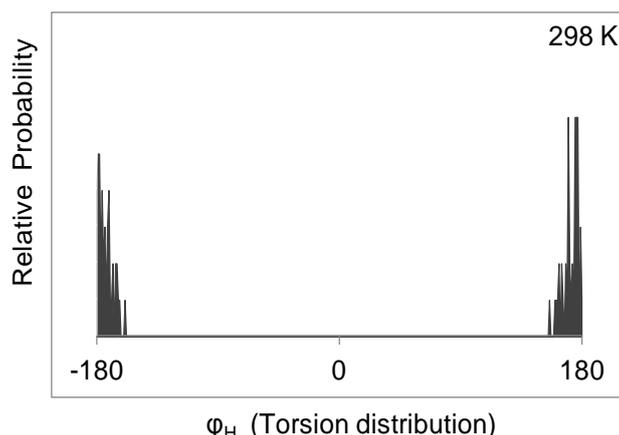


Figure 7- Torsion distribution, ϕ_H , of cellobiose in vacuum at 298 K when the initial structure is *syn*.

Figure 8 shows that the *anti* conformation remains in this conformation at 298 K, which is expected since this is the thermodynamically stable structure. However, at 375 K there are peaks in the distribution that belong to both *anti* and *syn* conformations. It is important to note that multiple barrier crossings occur, i.e., many *anti* \leftrightarrow *syn* isomerisation occur in the simulation time) showing that both parts of coordinate space are sampled at this temperature. As expected, the same behaviour is found at even higher temperatures, with the fraction of time spent in the *syn* conformation increasing with temperature. This trend does not depend on which initial conformer is used.

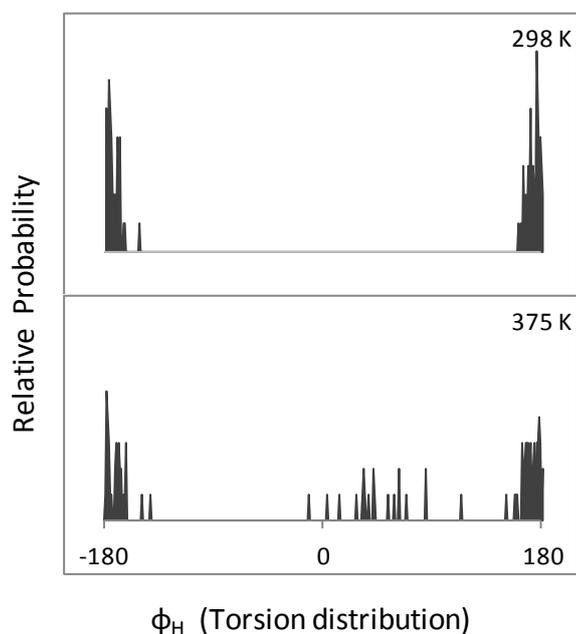


Figure 8- Torsion distribution, ϕ_H , of cellobiose in vacuum at 298K and 375 K when the initial structure is *anti*.

The distributions of φ_H for cellobiose in bulk water at 1 bar and 100, 298, 350 and 475 K are shown in Fig. 9.

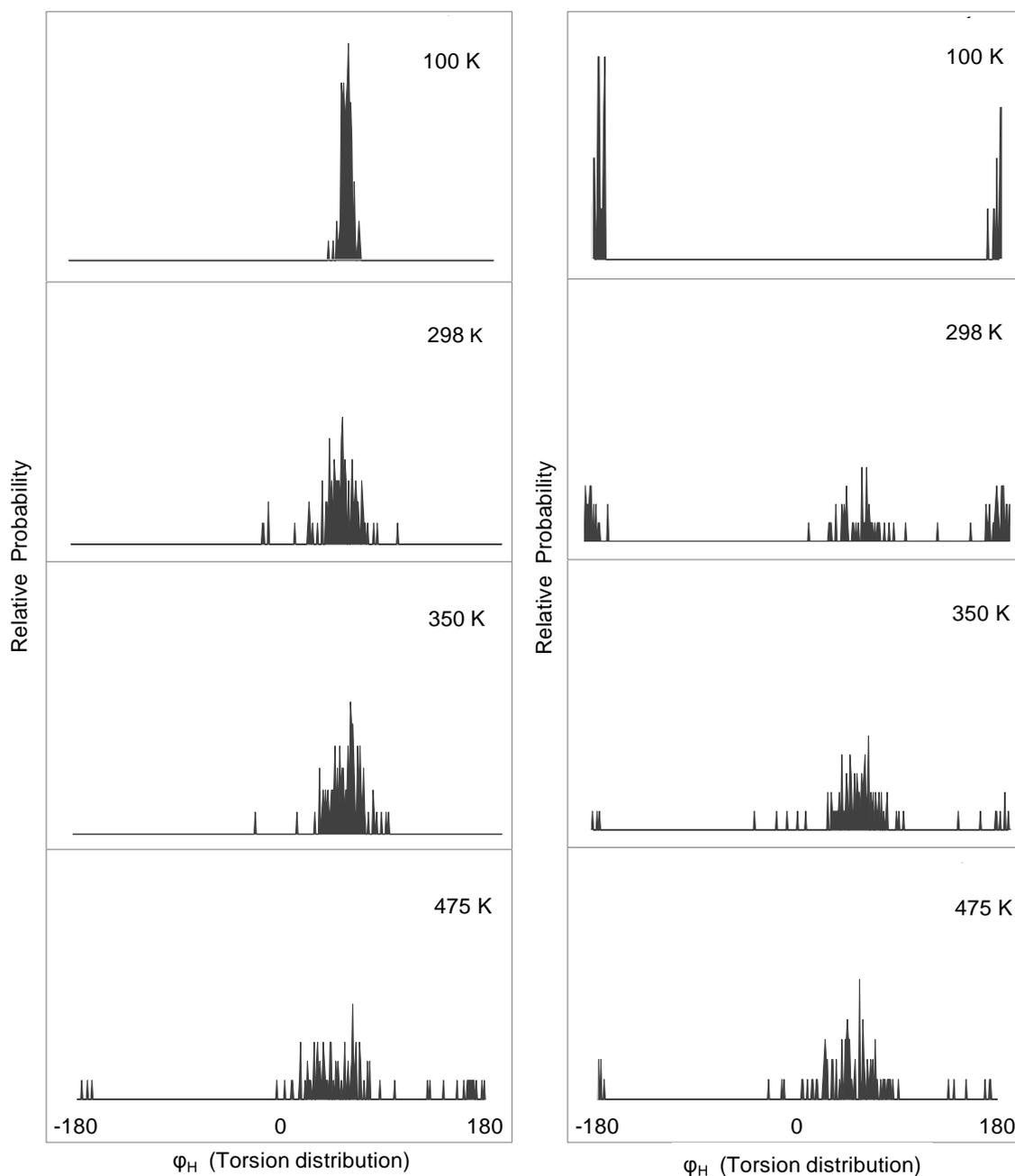


Figure 9- Torsion distribution, φ_H , of cellobiose in bulk water at 100, 298, 350 and 475 K, when the initial structures are *syn* (left) and *anti* (right).

The COMPASS force field was also used to study the binding strength between two parallel glucose and two parallel cellobiose molecules. This binding strength is expected to be important when dissolving cellulose in a pretreatment process. The glucose-glucose and cellobiose-cellobiose structures obtained from annealing with the COMPASS force field are

shown in the left column of Figure 10, and the structures after further geometry optimisation with B3LYP/6-311++G** are shown in the right column. There is very little change in the structure after geometry optimisation. For example, the glucose-glucose and cellobiose-cellobiose centre of mass distances are 5.16 and 4.52 Å after annealing, and they change to 5.36 and 4.31 Å after geometry optimisation. The glucose-glucose and cellobiose-cellobiose binding energies are 13 and 29 kcal mol⁻¹ according to the COMPASS force field and 14 and 41.6 kcal mol⁻¹ according to B3LYP/6-311++G**. Hence, both methods yield strong binding between the molecules, indicating the COMPASS force field will produce valid mechanisms and trends when studying the formation and breaking of glucose-glucose and cellobiose-cellobiose intermolecular bonds.

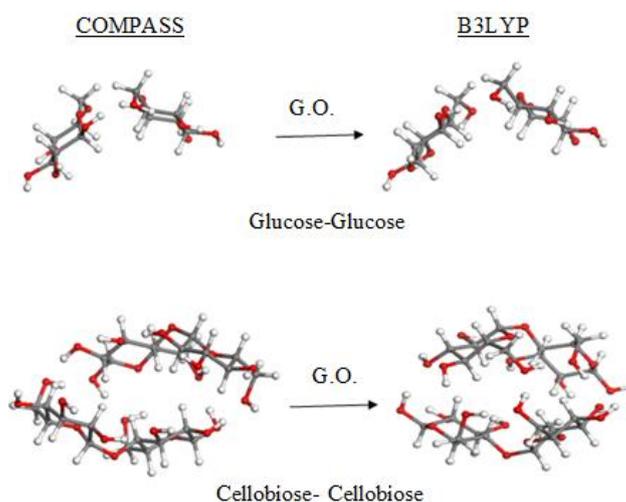


Figure 10- Glucose-glucose and cellobiose-cellobiose structures obtained from the COMPASS force field (left) and after subsequent geometry optimisation with B3LYP/6-311++G** (right).

The COMPASS force field was also used to study the interaction of glucose-glucose and cellobiose-cellobiose pairs with a water molecule (which is important for cellulose dissolution in water and steam explosion). More details of these calculations will be given in the summary of results of Paper IV.

3.2 Paper III

Molecular-level studies of dissolution of the crystalline structure of cellulose during steam explosion were performed at (100 °C, 1.0 bar), (160 °C, 6.2 bar), (210 °C, 19.0 bar) and (250 °C, 39.7 bar). These studies were based on the grand canonical Monte Carlo and molecular dynamics methods.

Figure 11 illustrates the change in the crystal structure after steam explosion at 250 °C and 39.7 bar. These changes are quantified below.

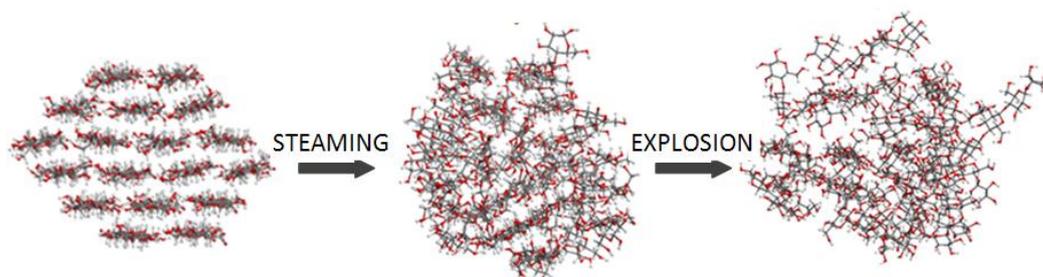


Figure 11- Illustration of the changes in the cellulose crystal structure during steam explosion at 250 °C and 39.7 bar.

Separation of the center of mass of each chain from the center of mass of the crystal for the initial structure and after the steaming simulations is presented in Figure 12. The figure reveals that there is substantial disruption of the crystal structure during the steaming stage at all (temperature, pressure) pairs, and an increase in temperature and pressure leads to a larger distortion.

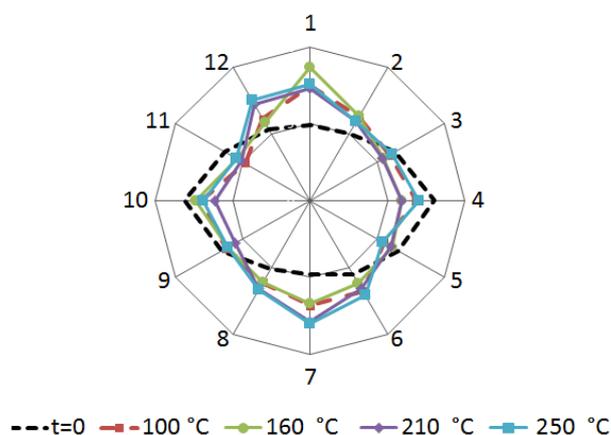


Figure 12- Separation of the center of mass of each chain in the outer shell from the center of mass of the crystal for the initial structure ($t=0$) and after the steaming simulations at (100 °C, 1.0 bar), (160 °C, 6.2 bar), (210 °C, 19.0 bar) and (250 °C, 39.7 bar). The chain numbers are according to Figure 3.

Figure 13 shows the distance of the centers of mass of the chains from the center of mass of the crystal after steaming and after explosion (NpT) at 1 bar and constant temperature. As expected, there is no change in the center of mass separations at the (100 °C, 1 bar) combination. This is because these conditions were used for both the steaming and explosion

simulations. The results are included here since they confirm that the system has equilibrated during steaming and there are therefore no further changes during the subsequent simulation.

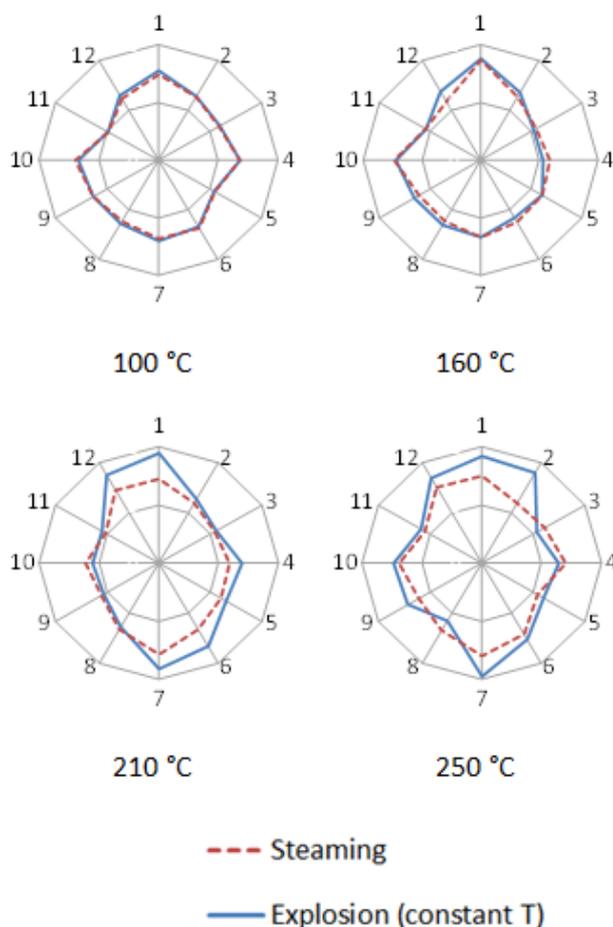


Figure 13- Separation of the centres of mass of outer shell chains from the centre of mass of the crystal after steaming (red dashed line) and after explosion at 1 bar and constant temperature (solid blue line). The four panels show results at different temperatures.

Similar to the change in the crystal structure during steaming, an increase in temperature and pressure leads to a larger disruption of the cellulose crystal structure during explosion. Also, although not shown in the figure (for the sake of clarity), the change in center of mass separation is largest for the chains in the outer shell compared to those in the core region. That is, the central cellulose chains (13-18) show very small changes compared to many of the chains in the outer shell (1-12). For example, the absolute value of the change in centers of mass of chains 1-12 during the explosion stage at (250 °C, 39.7 bar) is, on average, 1.9 Å, whereas for the chains in the core it is 0.8 Å.

Figure 14 shows the same results as those discussed with reference to Figures 13, but where the explosion simulations, starting with one of the structures obtained from steaming at (250 °C, 39.7 bar), are performed using NVE molecular dynamics with a volume 26 times larger than that used for steaming.

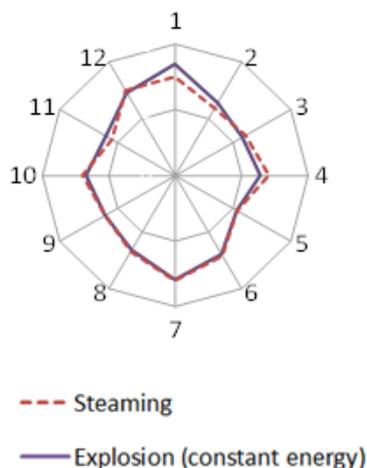


Figure 14- Same as Figure 13 but using NVE with the large box to simulate the explosion stage.

The figure shows that there is very little change in the centers of mass of the chains when performing the explosion simulations under constant energy conditions. Hence, it is the constant temperature, and not the drop in pressure, that caused the change in centers of mass during the NpT simulations. Since experimental explosion is performed under NVE conditions (but where the final pressure is 1 bar), the results obtained here indicate that most of the disruption of the crystal structure occurs during the steaming stage.

The change in radius of gyration for each chain as a function of its change in center of mass during explosion at 250 °C and 1 bar is shown in Figure 15. The results are typical for all initial structures and (temperature, pressure) pairs.

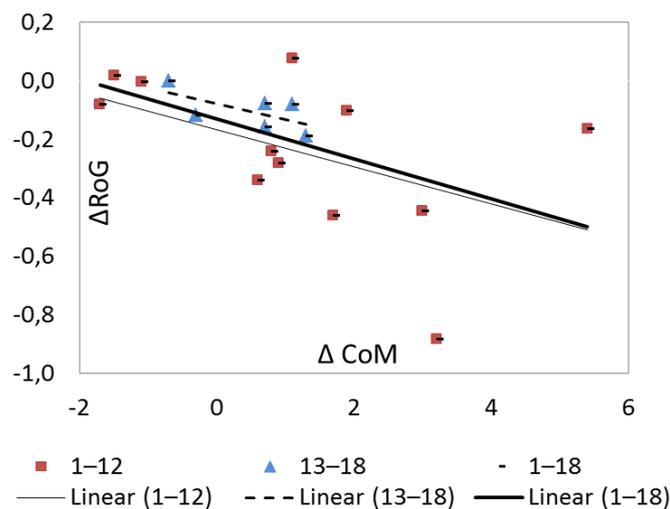


Figure 15- Change in radius of gyration for each chain as a function of its change in centre of mass during explosion at 250 °C and 1 bar. The lines are best-fit straight lines to the core chains (dashed line), the chains in the outer shell (thin solid line) and all eighteen chains (thick solid line).

Although all three curves show a trend of decreasing change in radius of gyration with increasing change in centre of mass, there is a large scattering (the R-squared values are 0.42, 0.24 and 0.27 for each of the fits, respectively). This means that there is no statistically relevant correlation between the change in radius of gyration of each chain and its change in centre of mass.

The change in separation between the non-reducing end of a chain in the outer shell and the same end of the neighbouring core chain is shown in Figure 16.

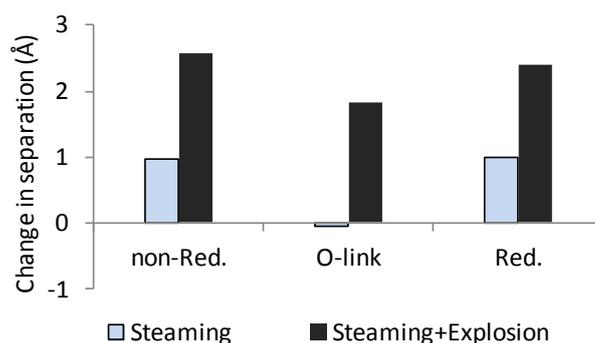


Figure 16- Change in the separation of centre of mass of the non-reducing end (non-Red.), the O-link in the centre of each chain and reducing end (Red.), between the outer chains and their neighbouring inner chains. The results are for steaming (red) and NpT explosion (blue) at 250 °C.

The figure reveals that there is larger disruption at the ends of the chains than at the middle during both steaming and explosion. It should also be noted that the values in the figure are

averages over all outer chains, and that some chain ends are separated by as much as 16.1 Å from the neighboring inner chain ends. Hence, after steaming and explosion the ends of the crystalline elementary fibrils are more accessible to enzymatic attack than other regions of the fibril.

At the beginning of the simulations all torsion angles in the chains in the crystal structure are *syn*. There was significant change from *syn* to *anti* conformation during the steaming and NpT explosion simulations. For example, after steaming at (250 °C, 39.7 bar) ~15.7% and ~9.9% of the torsions in the chains in the outer shell and core chains were *anti*, respectively. The corresponding numbers at (160 °C, 6.2 bar) were ~17.1% and ~7.4%. Explosion at 250 °C resulted in a further increase to ~21.4% for chains in the outer shell and ~14.4% in the core chains, whereas NpT explosion at 160 °C did not lead to a large increase in the percent of *anti* torsions (~17.9% for chains in the outer shell and ~9.1% in the core chains). These trends are typical for all structures and temperatures, and show that the chains in the outer shell, which also showed the largest change in center of mass motion, change more readily to *anti* torsions than the confined chains in the core region. In addition, a larger percentage of torsions change to the *anti* conformer at the higher temperatures and pressures and there is no significant correlation between the changes in percent *anti* conformer in a chain with its change in center of mass.

3.3 Paper IV

Computational studies of water and carbon dioxide interactions with a pair of cellobiose molecules was performed using the B3LYP/6-311++G** and Grimme's dispersion correction. This study yields information on cellobiose-cellobiose bonding mechanisms and interactions between an H₂O or CO₂ with the cellobiose pair that can give a deeper understanding of the steam and SC-CO₂ explosion mechanisms.

The goal of this study is to determine if the CO₂ molecule yields significantly different low energy structures compared to when the H₂O interacts with the cellobiose pair and to investigate the relative importance of the inter-cellobiose hydrogen and van der Waals bonding and how this may differ between the H₂O and CO₂ complexes.

3.3.1 H₂O-cellobiose pair

Figure 17 shows the relative energies, ΔE , of the 90 unique H₂O-cellobiose pair local minimum energy structures, ordered according to relative energies obtained from the DFT with dispersion correction (DFT-D) calculations. These structures were initially geometry optimized with the COMPASS force field. The energies are relative to the DFT-D energy of the lowest energy structure. The lowest energy structure obtained from DFT-D also has the lowest DFT energy (i.e., when dispersion corrections are not included).

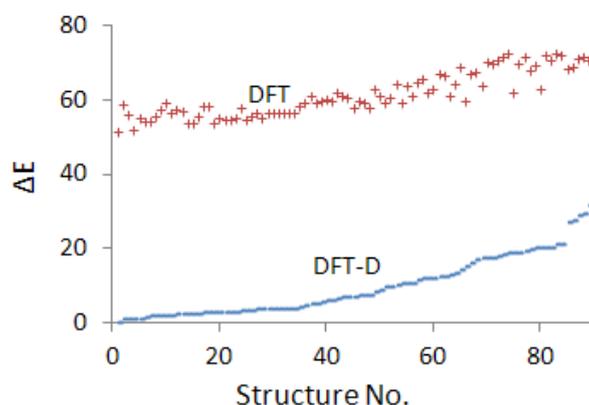


Figure 17- Relative energies (in kcal/mol) of local minimum H₂O-cellobiose pair structures obtained from DFT-D and DFT.

The Figure 17 shows that the energy difference between the high and low energy structures is ~25 kcal/mol within the DFT-D series and ~15 kcal/mol within the DFT series. This difference between the change in DFT-D and DFT energies is due to the extra stability that the dispersion contributes to the low energy structures. The dispersion correction yields DFT energies that are ~50 kcal/mol lower in energy than the DFT results.

Structures that have low energy consist of cellobiose molecules that are parallel to each other and where the glucose units on one of the molecules lie directly above the glucose units on the second molecule. This maximises the number of hydrogen bonds (which is between 5 and 7 in the low energy structures) and the van der Waals energy. Structures with intermediate energies consist of cellobiose molecules that lie parallel to each other, but the cellobiose molecules are shifted relative to each other that result in fewer hydrogen bonds (3-5) and reduced van der Waals attraction. In the structures with the highest energies, the cellobiose molecules have almost no overlap of the glucose units. There are only 1 or 2 hydrogen bonds and the van der Waals interactions are far weaker.

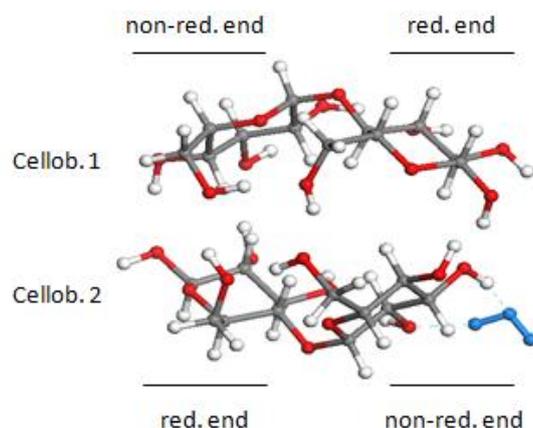


Figure 18- Lowest energy H₂O-cellobiose pair structure obtained from DFT-D. The H₂O molecule is shown in blue.

Figure 18 shows the lowest energy structure obtained from DFT-D. There are six hydrogen bonds in this structure that are between O3--H-O2', O6--H-O6', H-O4--O1', O4--H-O1', O1'-H--O3 and O6'-H--O6, where the first number in each bond is for Cellob.1 and the second for Cellob.2. The H₂O molecule is attached to the cellobiose pair by two hydrogen bonds in the minimum energy structure. The distance between centre of masses of the H₂O and Cellob.2 is 5.63 Å. The intermolecular energy between the H₂O and the cellobiose pair is -15.5 kcal/mol.

Table 3- Cellobiose-cellobiose intermolecular energies ($E_{\text{inter-pair}}$) and the dispersion correction energies (E_{disp}) in kcal/mol for some structures shown in Figure 17.

Structure No.	$E_{\text{inter-pair}}$	E_{disp} (%)
1	-51.5	-17.8 (35)
6	-46.8	-18.5 (40)
14	-51.2	-17.6 (34)
37	-42.2	-20.8 (49)
40	-27.4	-11.2 (41)
46	-38.2	-14.6 (38)
48	-33.0	-20.3 (62)
78	-34.6	-16.7 (48)
82	-29.4	-18.6 (63)
86	-12.6	-9.7 (77)
88	-13.8	-4.4 (32)
90	-17.3	-6.3 (36)

Table 3 represents DFT-D intermolecular energies ($E_{\text{inter-pair}}$) and the dispersion correction energies (E_{disp}) of cellobiose-cellobiose pairs in kcal/mol for some structures of Figure 17. It is evident that the intermolecular energy decreases with increasing structure number i.e. when relative energy increases. The dispersion contributions to the inter-cellobiose energies and the

non-dispersion contribution typically decrease or get weaker as the structures become less stable or relative energies increase. There is no clear trend of the percentage contribution increasing or decreasing with increasing the relative energies, which indicates that the dispersion and non-dispersion contributions decrease equally rapidly as the energy of the structure increases.

3.3.2 CO₂-cellobiose pair

Figure 19 shows the relative energies of the 80 unique CO₂-cellobiose pair local minimum energy structures. These structures were initially geometry optimized with the COMPASS force field.

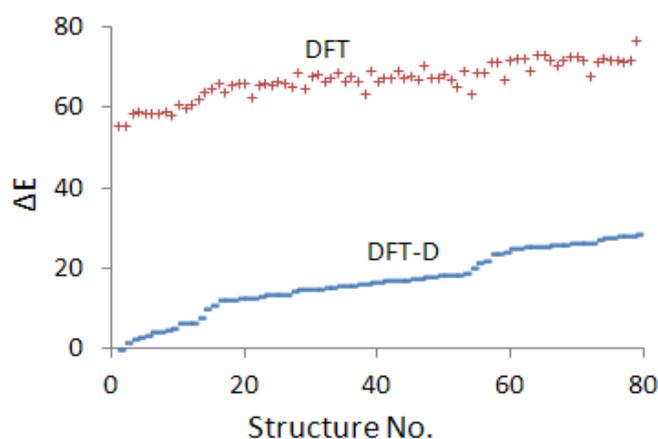


Figure 19- Same as Fig 2 but for the CO₂-cellobiose pair structures.

The minimum energy structure obtained from the DFT-D calculations is also the DFT minimum energy structure, and the trend of increasing relative energies from Structures 1 through 80 is the same for DFT-D and DFT calculations. The difference in DFT-D energies between the highest and lowest energy structures is ~25 kcal/mol (which was the same for the H₂O-cellobiose pair systems) and this difference in DFT energies is ~15 kcal/mol (which was also the same for the H₂O systems). Since DFT does not include the dispersion contribution to the stabilisation of the low energy structures, the DFT energies are ~60 kcal/mol higher than the DFT-D energies.

The lowest energy CO₂-cellobiose pair structures are parallel such that the inter-cellobiose attraction is maximised. There are 5-7 H-bonds in the minimum energy CO₂-cellobiose pair structure (~Structures 1-15), while cellobiose molecules that have intermediate energies (~Structures 16-60) results in fewer (~3-5) H-bonds and weaker dispersion attractions. The

structures with the highest relative energies (~Structures 61-80) have even fewer H-bonds and weaker van der Waals attraction.

Figure 20 is the minimum energy CO₂-cellobiose pair structure. The separation between the cellobiose centres of mass is 4.10 Å. Both cellobiose molecules have the *anti* conformation, with $\varphi_H = 177.5$ and 177.0° for Cellob.1 and Cellob.2, respectively. There are seven H-bonds, which are located between O6--H-O3', O3'--H-O6, O3-H--O3, O2'-H--O6, O3'-H--O6', O4'-H--O5' and O6'-H--O4', where the first number in each bond refers to the Cellob.1 molecule and the second number to Cellob.2.

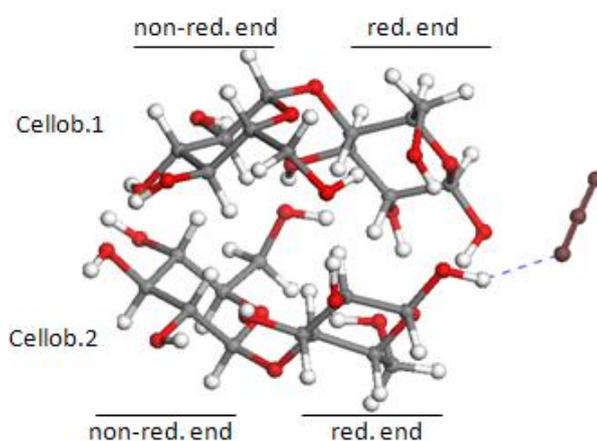


Figure 20- CO₂-cellobiose pair minimum energy structure. The CO₂ molecule is shown in brown.

The distance between the centre of mass of the CO₂ molecule and the cellobiose pair in the minimum energy structure shown in Figure 20 is 7.51 Å. Neither this distance, nor the CO₂-cellobiose intermolecular energy, shows a systematic change with increasing structure number. For example, this energy is -8.2, -9.4 and -5.6 kcal/mol for Structures 1, 21 and 72, respectively.

Table 4 shows that the inter-cellobiose energy decreases by increasing relative energy. The dispersion energy and the non-dispersion contribution to the intermolecular energy also decrease. Similarly to the H₂O-cellobiose pair systems, there is no clear evidence that the relative contribution of the dispersion energy (shown as percent in parenthesis in the table) either increases or decreases with increasing structure number.

Table 4- same as Table 3, but for CO₂-cellobiose pair systems

Structure No.	E _{inter-pair}	E _{disp} (%)
1	-56.6	-23.3 (41)
3	-54.9	-22.4 (41)
11	-46.7	-19.1 (41)
19	-38.0	-20.7 (54)
21	-34.0	-13.4 (39)
22	-25.6	-17.8 (70)
28	-30.2	-20.0 (66)
34	-29.4	-19.2 (65)
46	-23.6	-12.2 (52)
54	-18.9	-6.0 (32)
65	-12.9	-11.2 (87)
72	-17.4	-6.1 (35)

3.4 Paper V

The relative energies obtained from DFT-D, DFT and the COMPASS force field for CO₂- cellobiose pair systems are shown in Figure 21. The numbering of the structures is according to the lowest energy structure obtained from DFT-D method. DFT energies are obtained from the DFT-D optimisations and where the dispersion corrections have not been added to the total energy.

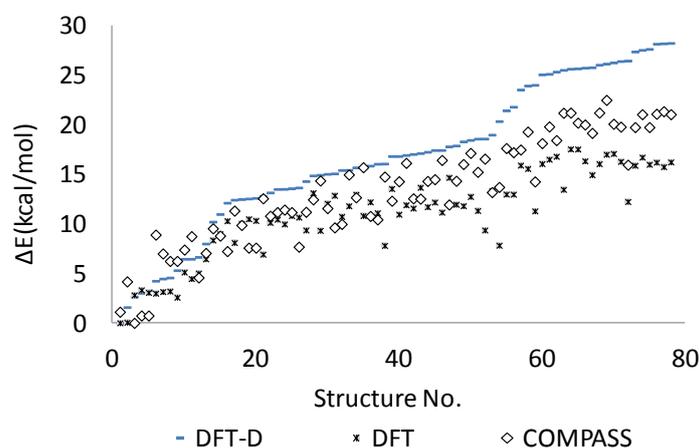


Figure 21- Relative energies of local minimum CO₂-cellobiose pair structures obtained from DFT-D, DFT and the COMPASS force field.

Figure 21 shows that the total energies obtained from DFT-D, DFT and COMPASS follow the same trends, such that the low and high energy structures obtained from DFT and the COMPASS force field are typically the same as the structures obtained from DFT-D. The

lowest energy structure that is obtained from DFT-D (first structure) has a COMPASS relative energy of 1.1 kcal/mol. All of these low energy structures are similar in geometry and are more compact compared to the higher energy structures. Hence, the COMPASS force field correctly describes the trends of compact structures having lower relative energies than structures where the cellobiose units have shifted relative to each other.

The lowest energy CO₂-cellobiose pair structure obtained from DFT-D (right) and the initial structure obtained from geometry optimisation using the COMPASS force field (left) are shown in Figure 22.

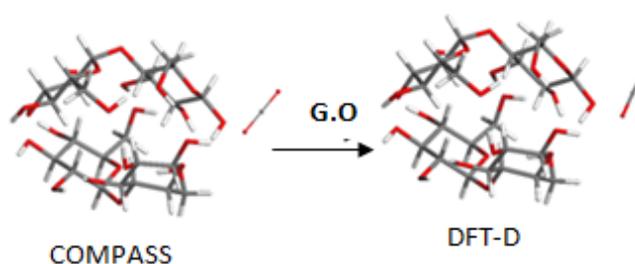


Figure 22- Lowest energy CO₂-cellobiose pair structure obtained from DFT-D. The initial structure obtained from geometry optimisation using the COMPASS force field is shown on the left and that obtained from the DFT-D optimization is shown on the right.

It is evident that these structures are similar; hence, geometry optimisation with DFT-D only leads to small changes in the structure obtained from the COMPASS force field. For instance, the separation between the cellobiose centres of mass is 3.85 Å in the structure obtained from the COMPASS force field and it is 4.10 Å according to DFT-D (more details can be found in Paper V). In both structures, seven H-bonds are formed by the same atoms. Both methods yield cellobiose molecules with the *anti* conformation. These similarities indicate that the COMPASS force field provides a valid description for the trends of CO₂ interacting with cellulose and for the disruption of the cellulose crystal structure. It was therefore used to study these trends during SC-CO₂ explosion of cellulose.

Similarly to the steam explosion, the molecular-level studies of dissolution of the crystalline structure of cellulose were simulated using grand canonical Monte Carlo and molecular dynamics. The SC-CO₂ explosion was at different temperature/ pressure combinations of 110 °C/ 2500 psi, 110 °C/ 3000 psi, 110 °C/3500, 110 °C/ 4000 psi, 135 °C/ 3500 psi, 165 °C/ 3500 psi and 200 °C/ 3500 psi.

Figure 23 illustrates the change in the crystal structure after steam explosion at 250 °C and 39.7 bar. These changes are quantified below.

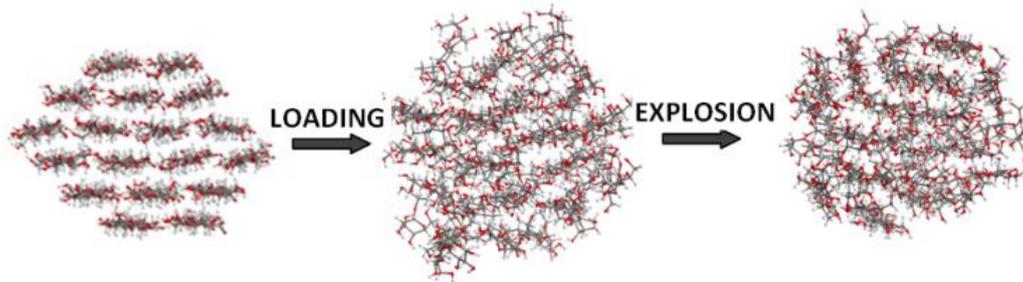


Figure 23- Illustration of the changes in the cellulose crystal structure during SC-CO₂ explosion at 200 °C and 3500 psi.

The changes in the centres of mass for each chain relative to the centre of mass of the crystal are shown in Figure 24.

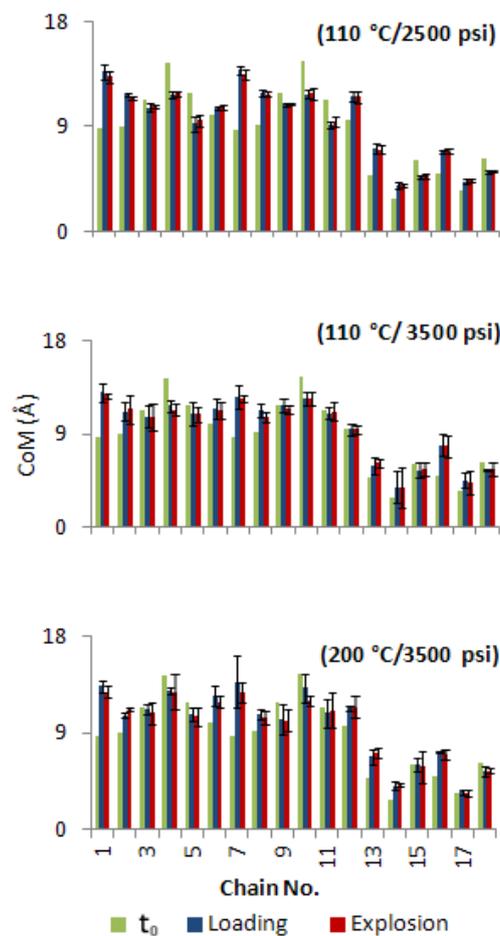


Figure 24- Centre of mass of each chain 1-18 relative to the centre of mass of the crystal for the initial structure (t_0) and after loading and explosion. Results are for Systems (110 °C/ 2500 psi), (110 °C/ 3500 psi) and (200 °C/ 3500) and the standard deviations are from the three trajectories

propagated at each temperature/ pressure combination. The standard deviations are shown as error bars.

Figure 24 shows that during SC-CO₂ loading and explosion the centres of mass of chains 3-5, 9-11, 15 and 18 decrease while other chains show an increase in centres of mass relative to the centre of mass of the crystal. This occurs since the pressure of the SC-CO₂ makes the cross section of the crystal structure more circular. The figure also shows that displacement of the chains in the outer layer (chains 1-12) is larger than chains in the core and changes in the chains in the outer shell affect the displacement of their neighbouring chains in the core. For example, during loading, chains 1 and 7 move away from the centre of the crystal by about 4.3 and 3.9 Å, while chains 4 and 10 move closer to the centre about 2.7 and 2.1 Å. Similarly, chains 13 and 16, which neighbour chains 1 and 7, respectively, move out by about 1.1 and 2.9 Å, respectively, while chains 15 and 18 move inwards by about 0.7 Å. As shown in Figure 22, most of the changes occur during loading of the SC-CO₂, and explosion has a far smaller effect on the disruption of the structure.

The effect of temperature and pressure on SC-CO₂ loading and explosion was studied using two sets of chains. The first set was chains 1, 6, 7 and 12 and the second set was chains 4, 5, 10 and 11. These sets were chosen since the chains are all in the outer shell and the chains in the first set increase their centre of mass during loading and explosion and the chains in the second set decrease their centre of mass.

3.4.1 Effect of temperature

The left panel in Figure 25 shows the effect of increasing temperature on the average change in centre of mass of chains 1, 6, 7 and 12 after loading and the combined loading and explosion. The pressure is 3500 psi and the temperature increases from 110 to 200 °C. The right panel is the same but for chains 4, 5, 10 and 11.

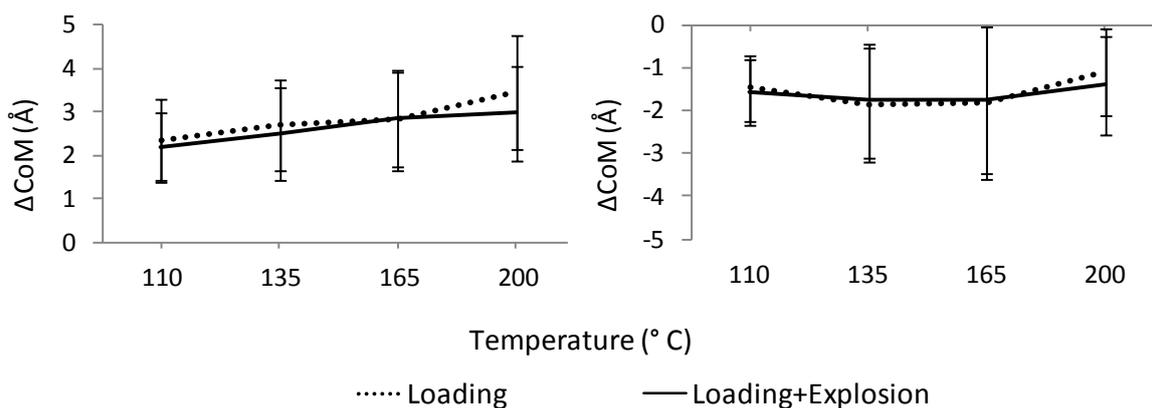


Figure 25- Effect of temperature on the average of separation of centre of mass of chains 1, 6, 7 and 12 (left) and chains 4, 5, 10 and 11 (right) from the centre of mass of the crystal after loading and explosion at 3500 psi, for systems at 110, 135, 165 and 200 °C. The error bars are standard deviations obtained from the three trajectories propagated at each temperature/ pressure combination.

The results presented in the figure reveal that, although there is no significant effect of increasing temperature (within the statistical uncertainty of the error bars); there is a trend of increasing disruption of the crystal structure with increasing temperature. The average change in the centres of mass of chains 1, 6, 7 and 12 when increasing from 110 to 135 °C is 0.35 Å, and this increases to 0.48 and 1.11 Å with a further increase in temperature to 165 and 200 °C. Increasing the temperature has a smaller effect on the average centre of mass of chains 4, 5, 10 and 11, which is about 0.3 Å when the temperature increases from 110 to 200 °C.

Similarly to the discussion with reference to Figure 24, Figure 25 shows that explosion does not lead to further disruption of the crystal structure. For example, explosion at 110, 135, 165 and 200 °C changes the average centres of mass of chains 1, 6, 7 and 12 by only about 0.14, 0.19, 0.05 and 0.48 Å (compared to the centres of mass after loading).

3.4.2 Effect of pressure

Figure 26 is the same as Figure 25 but when the pressure is increased from 2500 to 4000 psi at a constant temperature of 110 °C.

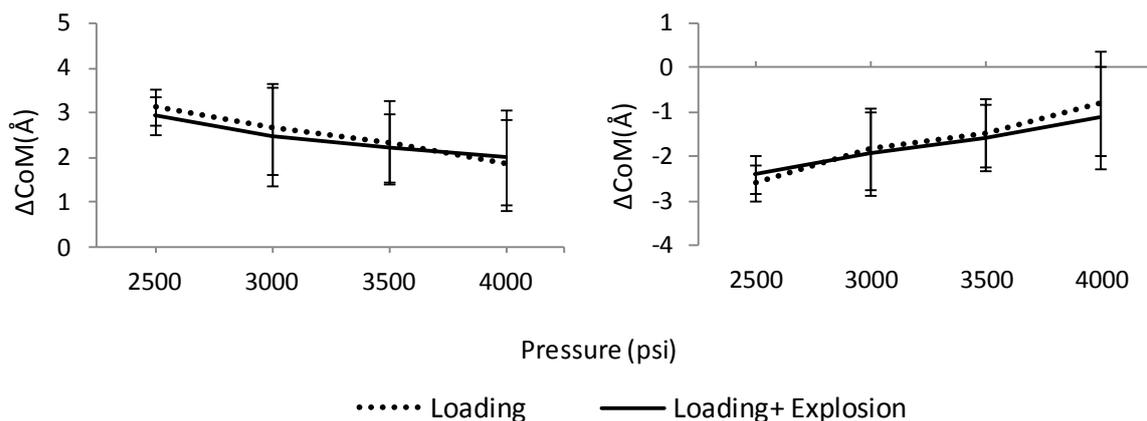


Figure 26- Same as for Figure 25 but for systems 2500, 3000, 3500 and 4000 psi. The temperature is 110 °C for all systems.

There is a larger change in separation of the centre of mass of the chains relative to the centre of the crystal at lower pressures. This is probably due to the fact that lower pressures allow chains 1, 6, 7 and 12 to move further away from the centre of the crystal than at higher pressures. This, in turn, means that chains 4, 5, 7 and 10 can move towards the centre of the crystal at the lower pressures. For example, increasing the pressure from 2500 to 3000 psi decreases the average change in centre of mass separation for chains 1, 6, 7 and 12 by 0.48 Å during loading. Further increases in pressure to 3500 and 4000 psi leads to larger decreases of 0.78 and 1.27 Å compared to 2500 psi. Similarly, the average change in centre of mass separation for chains 4, 5, 7 and 10 differs by 0.8 Å between pressures of 2500 and 3000 psi, 1.1 Å between 2500 and 3500 psi and 1.8 Å between 2500 and 4000 psi.

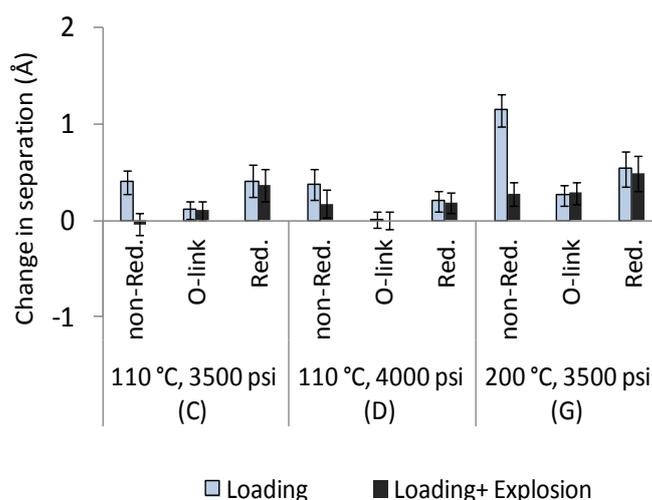


Figure 27- Change in the separation of centre of mass of the non-reducing end (non-Red.), the O-link in the centre of each chain and reducing end (Red.) between the outer chains and their neighbouring

inner chains during loading and explosion. The results are shown for loading and explosion at 110 °C and 3500 psi, 110 °C and 4000 psi and 200 °C and 3500 psi.

Figure 27 reveals that there is a larger disruption at the ends of the chains than in the middle (O-link) during both loading and explosion. Hence, there is larger disruption at the ends of the crystal structure (i.e., at the chain ends) than in the middle of the crystal. This would lead to enhanced accessibility for enzymatic hydrolysis at the crystal ends rather than in the middle of the crystal.

No significant correlation was observed between the magnitude of the change in centres of mass of chains during loading and explosion with changes in their radii of gyration.

4 Conclusions and Outlook

The COMPASS force field was preferred over the Dreiding and Universal force fields for study of glucose and cellobiose molecules as well as cellulose structural changes during dissolution with steam and supercritical carbon dioxide (SC-CO₂).

The validity of the COMPASS force field was checked by comparing the resulted structures and energies from first principles calculations and force field methods for glucose and cellobiose. These molecular systems were selected since they are sufficiently small to perform the first principles calculations in a tractable time and also being relevant to the dissolution of cellulose in water. The comparison was also made with relative energies and structures calculated by others using the B3LYP/6-311++G** method. The COMPASS force field was the only force field that yields the correct *anti* conformer of cellobiose and gave the minimum energy structure in vacuum and *syn* conformer in vacuum at high temperature and in aqueous environment at 1 bar and 298 K. The COMPASS force field also indicates that the crystal structure of cellulose has lower energy than the separated cellulose chains.

According to previous studies, the B3LYP/6-311++G** density functional method yields valid energies and structures for cellobiose and H₂O-cellobiose systems. This method, including Grimme's dispersion correction, was used to study the interactions between several structures containing H₂O or CO₂ and two cellobiose molecules that were initially optimized using the COMPASS force field.

Comparison of the intermolecular energies of cellobiose-cellobiose obtained from DFT and DFT-D showed that both the non-dispersion and dispersion terms have large contributions (between 30 and 70 %.) to the intermolecular energies.

Geometry optimisation with the DFT-D method showed that the H₂O and CO₂ molecules prefer to bond to the surface of the cellobiose pair and opposed to being located between the cellobiose molecules.

The intermolecular energies obtained from DFT-D calculations also showed that when a H₂O molecule interacts with two cellobiose molecules, the intermolecular interactions between the two cellobiose molecules are weaker than when a CO₂ molecule interacts with the cellobiose

molecules. This is due to larger electron density that is located between the cellobiose pair and the water molecule.

Studies of steam and supercritical carbon dioxide (SC-CO₂) explosion were performed using grand canonical Monte Carlo and molecular dynamics based on the COMPASS force field. The calculations are based on the crystalline structure of cellulose that was proposed by Ding and Himmel. In these studies restructuring of the cellulose crystal were investigated by changes in the centres of mass of cellulose chains during loading of saturated steam (steaming) and SC-CO₂ and explosion steps. For both solvents disruption was larger for the chains in the outer shell compared to the core chains and comparison between chains within either the outer shell or core region showed that there was no significant correlation between changes in the radius of gyration and the change in centre of mass of the chains. Also, for both methods increasing the temperature leads to more disruption in the cellulose crystal. During SC-CO₂ explosion, increasing the pressure decreases the disruption of the crystal.

Comparing the centre of mass of neighbour chains, larger distortion of the crystal structure was seen at the ends of the crystal. Increase in disruption of the cellulose crystals due to increasing temperature, enhances accessibility of enzymes to the cellulose chains which particularly happens at the ends of the elementary fibrils and is important for the cellulose pretreatment in production of biofuel.

Future work

Several pretreatment processes have been developed for decreasing the recalcitrance and increasing dissolution of cellulose in several solvents, but only a few of them seem to be promising. N-methyl morpholine-N-oxide (NMMO) hydrate is an effective and direct solvent for dissolution of cellulose. Different proportion of water has different effects on the dissolution of cellulose. For example NMMO–water mixture containing 83% and 87% (w/w) NMMO (monohydrate) is known as dissolution mode and the cellulose fibres can be dissolve completely, while in a mixture with 76–82% NMMO, cellulose dissolves partially in the produced fibre balloons. Lower NMMO contents of 70–75% cause partial ballooning and swelling. Less NMMO content shows no effect on dissolution of cellulose fibres [99, 100]. The studies presented here, can be applied for molecular-level investigation of the dissolution of cellulose with different NMMO-water percentage.

The studies presented in this thesis shows the effect of steam and SC-CO₂ explosion on an 18 chains elementary fibril, the study can be extended to the 36 chains model and to microfibrils containing several elementary fibrils.

There is also no molecular-level study of the pretreatment for other proposed models of cellulose structures; these studies can also be executed on the different models of cellulose elementary and microfibrils.

Acknowledgements

First, I would like to express my deep gratitude to my supervisor Prof. Kim Bolton for accepting me as a student and for his patient guidance, enthusiastic encouragement and useful critiques of this research work.

My grateful thanks are also extended to my examiner Prof. Claes Niklasson and my co-supervisor Assoc. Prof. Peter Ahlström.

I would also like to thank Prof. Frank Momany for his great and helpful contribution in my first paper and Prof. Mohammad Taherzadeh for his assistance.

I am very thankful to my colleagues in modelling group: Anders, Edvin, Kavitha, Martin, Shayesteh and Abas for their good collaboration.

I would like to express my appreciation to all staffs at the School of Engineering, University of Borås for providing such a friendly environment. Peter, Ilona, Thomas S., Solveg, Louise and Sari are gratefully acknowledged for their helps.

My sincere thank also goes to Martin Bohlén for being such a good colleague.

I wish also to thanks my friends and all former and present PhD students at the University of Borås especially Behnaz, Azadeh, Shayesteh, Supansa, Päivi, Farzad, Abas, Tariq, Kamran and Mofoluwake.

I am grateful to Stiftelsen Foreningssparbanken Sjuharad and the Carl Trygger Foundation for Scientific Research for the financial support. The Quantum Mechanics calculations were performed on resources provided by the Swedish National Infrastructure for Computing (SNIC) at NSC, HPC2N and C3SE. Molecular Mechanics results were obtained using program from Accelrys Software Inc.

My special thanks to my parents, Zahra and Heidar for their endless love, supports and encouragements. I wish also to thanks my lovely sisters, Parisa and Bitá.

Thanks with love to my kind husband, Mehdi for his support, love, patience and encouragements.

And finally my greatest thanks to my wonderful children, Sahand and Soren for their patience and understanding.

References

1. Goldemberg, J. (2007). Ethanol for a sustainable energy future. *science*, 315(5813), 808-810.
2. Orfao, J. J. M., Antunes, F. J. A., & Figueiredo, J. L. (1999). Pyrolysis kinetics of lignocellulosic materials—three independent reactions model. *Fuel*, 78(3), 349-358.
3. Lucia, L. A. (2008). Lignocellulosic biomass: A potential feedstock to replace petroleum. *BioResources*, 3(4), 981-982.
4. Limayem, A., & Ricke, S. C. (2012). Lignocellulosic biomass for bioethanol production: current perspectives, potential issues and future prospects. *Progress in Energy and Combustion Science*, 38(4), 449-467.
5. Ho, M. W. (2006). Ethanol from cellulose biomass not sustainable nor environmentally benign. *Science in Society*, (30), 32-35.
6. Taherzadeh, M. J., & Karimi, K. (2007). Enzymatic-based hydrolysis processes for Ethanol. *BioResources*, 2(4), 707-738.
7. Jin, Z., Katsumata, K. S., Lam, T. B. T., & Iiyama, K. (2006). Covalent linkages between cellulose and lignin in cell walls of coniferous and nonconiferous woods. *Biopolymers*, 83(2), 103-110.
8. Festucci-Buselli, R. A., Otoni, W. C., & Joshi, C. P. (2007). Structure, organization, and functions of cellulose synthase complexes in higher plants. *Brazilian Journal of Plant Physiology*, 19(1), 1-13.
9. Esteghlalian, A., Hashimoto, A. G., Fenske, J. J., & Penner, M. H. (1997). Modeling and optimization of the dilute-sulfuric-acid pretreatment of corn stover, poplar and switchgrass. *Bioresource Technology*, 59(2), 129-136.
10. Davin, L. B., Patten, A. M., Jourdes, M., & Lewis, N. G. (2008). Lignins: a twenty-first century challenge. *Biomass Recalcitrance: Deconstructing the Plant Cell Wall for Bioenergy*, 213-305.
11. Mudhoo, A. (Ed.). (2012). *Biogas production: pretreatment methods in anaerobic digestion*. John Wiley & Sons.
12. Pethrick, R. A. (Ed.). (2007). *Polymer structure characterization: from nano to macro organization*. Royal Society of Chemistry.
13. Pérez, J., Muñoz-Dorado, J., de la Rubia, T. D. L. R., & Martínez, J. (2002). Biodegradation and biological treatments of cellulose, hemicellulose and lignin: an overview. *International Microbiology*, 5(2), 53-63.
14. Fengel, D., & Wegener, G. (Eds.). (1983). *Wood: chemistry, ultrastructure, reactions*. Walter de Gruyter.
15. Zugenmaier, P. (2008). *Cellulose* (pp. 101-174). Springer Berlin Heidelberg.

16. Ding, S. Y., & Himmel, M. E. (2006). The maize primary cell wall microfibril: a new model derived from direct visualization. *Journal of Agricultural and Food Chemistry*, 54(3), 597-606.
17. Stone, B. (2005). Cellulose: Structure and distribution. *eLS*.
18. O'SULLIVAN, A. C. (1997). Cellulose: the structure slowly unravels. *Cellulose*, 4(3), 173-207.
19. Newman, R. H., Ha, M. A., & Melton, L. D. (1994). Solid-state ¹³C NMR investigation of molecular ordering in the cellulose of apple cell walls. *Journal of agricultural and food chemistry*, 42(7), 1402-1406.
20. Newman, R. H., Davies, L. M., & Harris, P. J. (1996). Solid-state ¹³C nuclear magnetic resonance characterization of cellulose in the cell walls of Arabidopsis thaliana leaves. *Plant physiology*, 111(2), 475-485.
21. Smith, B. G., Harris, P. J., Melton, L. D., & Newman, R. H. (1998). Crystalline cellulose in hydrated primary cell walls of three monocotyledons and one dicotyledon. *Plant and cell physiology*, 39(7), 711-720.
22. Himmel, M. E. (2009). *Biomass recalcitrance: deconstructing the plant cell wall for bioenergy*. Wiley-Blackwell.
23. Strati, G. L., Willett, J. L., & Momany, F. A. (2002). Ab initio computational study of β -cellobiose conformers using B3LYP/6-311++ G**. *Carbohydrate research*, 337(20), 1833-1849.
24. Bazooyar F., B.M., Bolton K., *Computational Studies of Water and Carbon Dioxide Interactions with Cellobiose*. Journal of Molecular Modeling, 2014. Submitted.
25. Bosma, W. B., Appell, M., Willett, J. L., & Momany, F. A. (2006). Stepwise hydration of cellobiose by DFT methods: 1. Conformational and structural changes brought about by the addition of one to four water molecules. *Journal of Molecular Structure: THEOCHEM*, 776(1), 1-19.
26. Taherzadeh, M. J., & Karimi, K. (2008). Pretreatment of lignocellulosic wastes to improve ethanol and biogas production: a review. *International journal of molecular sciences*, 9(9), 1621-1651.
27. Silverstein, R. A. (2005). A comparison of chemical pretreatment methods for converting cotton stalks to ethanol.
28. Johnson, D. K., & Elander, R. T. (2008). Pretreatments for enhanced digestibility of feedstocks. *Biomass recalcitrance: Deconstructing the plant cell wall for bioenergy*, 436-453.
29. Kumar, P., Barrett, D. M., Delwiche, M. J., & Stroeve, P. (2009). Methods for pretreatment of lignocellulosic biomass for efficient hydrolysis and biofuel production. *Industrial & Engineering Chemistry Research*, 48(8), 3713-3729.

30. Mason, W. H. (1926). *U.S. Patent No. 1,578,609*. Washington, DC: U.S. Patent and Trademark Office.
31. BABCOCK, L. W. (1932). *U.S. Patent No. 1,855,464*. Washington, DC: U.S. Patent and Trademark Office.
32. Volynets, B., & Dahman, Y. (2011). Assessment of pretreatments and enzymatic hydrolysis of wheat straw as a sugar source for bioprocess industry. *Int J Energy Environ*, 2(3), 427-446.
33. Josefsson, T., Lennholm, H., & Gellerstedt, G. (2002). Steam explosion of aspen wood. Characterisation of reaction products. *Holzforschung*, 56(3), 289-297.
34. Ballesteros, I., Negro, M. J., Oliva, J. M., Cabañas, A., Manzanares, P., & Ballesteros, M. (2006, January). Ethanol production from steam-explosion pretreated wheat straw. In *Twenty-Seventh Symposium on Biotechnology for Fuels and Chemicals* (pp. 496-508). Humana Press.
35. Schultz, T. P., Biermann, C. J., & McGinnis, G. D. (1983). Steam explosion of mixed hardwood chips as a biomass pretreatment. *Industrial & Engineering Chemistry Product Research and Development*, 22(2), 344-348.
36. Gao, M., Xu, F., Li, S., Ji, X., Chen, S., & Zhang, D. (2010). Effect of SC-CO₂ pretreatment in increasing rice straw biomass conversion. *Biosystems engineering*, 106(4), 470-475.
37. Donaldson, L. A., Wong, K. K. Y., & Mackie, K. L. (1988). Ultrastructure of steam-exploded wood. *Wood science and technology*, 22(2), 103-114.
38. Martin-Sampedro, R., Capanema, E. A., Hoeger, I., Villar, J. C., & Rojas, O. J. (2011). Lignin changes after steam explosion and laccase-mediator treatment of eucalyptus wood chips. *Journal of agricultural and food chemistry*, 59(16), 8761-8769.
39. Jedvert, K., Saltberg, A., Lindström, M., & Theliander, H. (2011). Extraction of hemicelluloses after chemical pretreatment combined with mild steam explosion. In *16th International Symposium on Wood, Fiber and Pulping Chemistry, ISWFPC, Tianjin, 8 June through 10 June 2011* (Vol. 2, pp. 867-871).
40. Chacha, N., Toven, K., Mtui, G., Katima, J., & Mrema, G. (2011). Steam pretreatment of pine (*Pinus patula*) wood residue for the production of reducing sugars. *Cellulose Chemistry and Technology*, 45(7), 495.
41. Eckert, C. A. (1996). Supercritical fluids as solvents for chemical and materials processing. *Nature*, 383, 313-318.
42. Zheng, Y., Lin, H. M., Wen, J., Cao, N., Yu, X., & Tsao, G. T. (1995). Supercritical carbon dioxide explosion as a pretreatment for cellulose hydrolysis. *Biotechnology Letters*, 17(8), 845-850.

43. Paljevac, M., Primožič, M., Habulin, M., Novak, Z., & Knez, Ž. (2007). Hydrolysis of carboxymethyl cellulose catalyzed by cellulase immobilized on silica gels at low and high pressures. *The Journal of Supercritical Fluids*, 43(1), 74-80.
44. Canelas, D. A., & DeSimone, J. M. (1997). Polymerizations in liquid and supercritical carbon dioxide. In *Metal Complex Catalysts Supercritical Fluid Polymerization Supramolecular Architecture* (pp. 103-140). Springer Berlin Heidelberg.
45. Alinia, R., Zabihi, S., Esmaeilzadeh, F., & Kalajahi, J. F. (2010). Pretreatment of wheat straw by supercritical CO₂ and its enzymatic hydrolysis for sugar production. *Biosystems engineering*, 107(1), 61-66.
46. Alvira, P., Tomás-Pejó, E., Ballesteros, M., & Negro, M. J. (2010). Pretreatment technologies for an efficient bioethanol production process based on enzymatic hydrolysis: a review. *Bioresource technology*, 101(13), 4851-4861.
47. Koch, W., Holthausen, M. C., & Holthausen, M. C. (2001). *A chemist's guide to density functional theory* (Vol. 2). Weinheim: Wiley-Vch.
48. Luterbacher, J. S., Tester, J. W., & Walker, L. P. (2010). High-solids biphasic CO₂-H₂O pretreatment of lignocellulosic biomass. *Biotechnology and bioengineering*, 107(3), 451-460.
49. Narayanaswamy, N., Faik, A., Goetz, D. J., & Gu, T. (2011). Supercritical carbon dioxide pretreatment of corn stover and switchgrass for lignocellulosic ethanol production. *Bioresource technology*, 102(13), 6995-7000.
50. Ritter, D. C., & Campbell, A. G. (1991). Supercritical carbon dioxide extraction of southern pine and ponderosa pine. *Wood and Fiber Science*, 23(1), 98-113.
51. Zheng, Y., & Tsao, G. T. (1996). Avicel hydrolysis by cellulase enzyme in supercritical CO₂. *Biotechnology letters*, 18(4), 451-454.
52. Zheng, Y., Lin, H. M., & Tsao, G. T. (1998). Pretreatment for cellulose hydrolysis by carbon dioxide explosion. *Biotechnology progress*, 14(6), 890-896.
53. Kim, K. H., & Hong, J. (2001). Supercritical CO₂ pretreatment of lignocellulose enhances enzymatic cellulose hydrolysis. *Bioresource Technology*, 77(2), 139-144.
54. Hehre, W. J. (2003). *A guide to molecular mechanics and quantum chemical calculations* (p. p19). Irvine, CA: Wavefunction.
55. Cramer, C. J., & Bickelhaupt, F. M. (2003). Essentials of computational chemistry. *ANGEWANDTE CHEMIE-INTERNATIONAL EDITION IN ENGLISH*, 42(4), 381-381.
56. Jensen, F. (2007). *Introduction to computational chemistry*. John Wiley & Sons.
57. Börjesson, A. (2010). *In Silico Studies of Carbon Nanotubes and Metal Clusters*. Department of Physics; Institutionen för fysik.
58. Born, M., & Oppenheimer, R. (1927). Zur quantentheorie der molekeln. *Annalen der Physik*, 389(20), 457-484.

59. Schweber, S. S. (2011). *An introduction to relativistic quantum field theory*. Courier Dover Publications.
60. Heisenberg, W. (1974). *The uncertainty principle*. Educational Materials and Equipment.
61. Sadlej, A. J. (1991). Medium-size polarized basis sets for high-level-correlated calculations of molecular electric properties. *Theoretica chimica acta*, 81(1-2), 45-63.
62. Young, D. (2004). *Computational chemistry: a practical guide for applying techniques to real world problems*. John Wiley & Sons.
63. Nesbet, R. K. (1955). Configuration interaction in orbital theories. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 230(1182), 312-321.
64. Møller, C., & Plesset, M. S. (1934). Note on an approximation treatment for many-electron systems. *Physical Review*, 46(7), 618.
65. Hinchliffe, A. (2005). *Molecular modelling for beginners*. John Wiley & Sons.
66. Andrew R. Leach. (2001). *Molecular modelling: principles and applications*. Pearson Education.
67. Steinmann, S. N. (2012). *Understanding and Minimizing Density Functional Failures Using Dispersion Corrections* (Doctoral dissertation, ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE).
68. Dirac, P. A. (1930, July). Note on exchange phenomena in the Thomas atom. In *Mathematical Proceedings of the Cambridge Philosophical Society* (Vol. 26, No. 03, pp. 376-385). Cambridge University Press.
69. Slater, J. C. (1951). A simplification of the Hartree-Fock method. *Physical Review*, 81(3), 385.
70. Perdew, J. P., Burke, K., & Wang, Y. (1996). Generalized gradient approximation for the exchange-correlation hole of a many-electron system. *Physical Review B*, 54(23), 16533.
71. Perdew, J. P., Chevary, J. A., Vosko, S. H., Jackson, K. A., Pederson, M. R., Singh, D. J., & Fiolhais, C. (1992). Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation. *Physical Review B*, 46(11), 6671.
72. Heyd, J., Scuseria, G. E., & Ernzerhof, M. (2003). Hybrid functionals based on a screened Coulomb potential. *The Journal of Chemical Physics*, 118(18), 8207-8215.
73. Becke, A. D. (1993). Density-functional thermochemistry. III. The role of exact exchange. *The Journal of Chemical Physics*, 98(7), 5648-5652.
74. Lee, C., Yang, W., & Parr, R. G. (1988). Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Physical Review B*, 37(2), 785.

75. Miehlich, B., Savin, A., Stoll, H., & Preuss, H. (1989). Results obtained with the correlation energy density functionals of Becke and Lee, Yang and Parr. *Chemical Physics Letters*, 157(3), 200-206.
76. Stephens, P. J., Devlin, F. J., Chabalowski, C. F., & Frisch, M. J. (1994). Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *The Journal of Physical Chemistry*, 98(45), 11623-11627.
77. Cohen, A. J., Mori-Sánchez, P., & Yang, W. (2011). Challenges for density functional theory. *Chemical reviews*, 112(1), 289-320.
78. Heyd, J., & Scuseria, G. E. (2004). Efficient hybrid density functional calculations in solids: assessment of the Heyd–Scuseria–Ernzerhof screened Coulomb hybrid functional. *The Journal of chemical physics*, 121(3), 1187-1192.
79. Grimme, S. (2006). Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *Journal of computational chemistry*, 27(15), 1787-1799.
80. Shattuck, T. W. (2003). Colby College Molecular Mechanics Tutorial QUANTA Version.
81. Lewars, E. (2010). *Computational chemistry: introduction to the theory and applications of molecular and quantum mechanics*. Springer.
82. Moloy, E. C., Davila, L. P., Shackelford, J. F., & Navrotsky, A. (2002). High-silica zeolites: a relationship between energetics and internal surface areas. *Microporous and mesoporous materials*, 54(1), 1-13.
83. Levitt, M., & Lifson, S. (1969). Refinement of protein conformations using a macromolecular energy minimization procedure. *Journal of molecular biology*, 46(2), 269-279.
84. Fletcher, R., & Reeves, C. M. (1964). Function minimization by conjugate gradients. *The computer journal*, 7(2), 149-154.
85. Ermer, O. (1976). Calculation of molecular properties using force fields. Applications in organic chemistry. In *Bonding forces* (pp. 161-211). Springer Berlin Heidelberg.
86. Sun, H. (1998). COMPASS: an ab initio force-field optimized for condensed-phase applications overview with details on alkane and benzene compounds. *The Journal of Physical Chemistry B*, 102(38), 7338-7364.
87. Jorgensen, W. L., Maxwell, D. S., & Tirado-Rives, J. (1996). Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society*, 118(45), 11225-11236.
88. Bunte, S. W., & Sun, H. (2000). Molecular modeling of energetic materials: the parameterization and validation of nitrate esters in the COMPASS force field. *The Journal of Physical Chemistry B*, 104(11), 2477-2489.

89. Mark, J. E. (2007). *Physical properties of polymers handbook* (Vol. 1076). New York: Springer.
90. Mayo, S. L., Olafson, B. D., & Goddard, W. A. (1990). DREIDING: a generic force field for molecular simulations. *Journal of Physical Chemistry*, *94*(26), 8897-8909.
91. Rappé, A. K., Casewit, C. J., Colwell, K. S., Goddard Iii, W. A., & Skiff, W. M. (1992). UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *Journal of the American Chemical Society*, *114*(25), 10024-10035.
92. Conery, J. S., Peticolas, W. L., Rush III, T., Shanmugam, K., & Dominguez, J. (1995, January). A parallel algorithm for calculating the potential energy in DNA. In *System Sciences, 1995. Proceedings of the Twenty-Eighth Hawaii International Conference on* (Vol. 5, pp. 123-131). IEEE.
93. Frenkel, D., & Smit, B. (2001). *Understanding molecular simulation: from algorithms to applications* (Vol. 1). Academic press.
94. Keil, F. J. (1996). Application of numerical methods in chemical process engineering. In *Scientific Computing in Chemical Engineering* (pp. 92-110). Springer Berlin Heidelberg.
95. Bolton, K., & Nordholm, S. (1994). An evaluation of the Gauss-Radau algorithm for the simulation of chemical dynamics. *Journal of Computational Physics*, *113*(2), 320-335.
96. Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of state calculations by fast computing machines. *The journal of chemical physics*, *21*(6), 1087-1092.
97. Frenkel, D. (2004). Introduction to monte carlo methods. *Computational Soft Matter: From Synthetic Polymers to Proteins*, *23*, 29-59.
98. Allen, M. P., & Tildesley, D. J. (1987). Computer simulation of liquids.
99. Jeihanipour, A., Karimi, K., & Taherzadeh, M. J. (2010). Enhancement of ethanol and biogas production from high-crystalline cellulose by different modes of NMO pretreatment. *Biotechnology and bioengineering*, *105*(3), 469-476.
100. Cuissinat, C., Navard, P., & Heinze, T. (2008). Swelling and dissolution of cellulose. Part IV: Free floating cotton and wood fibres in ionic liquids. *Carbohydrate polymers*, *72*(4), 590-596.