J Mol Model (2015) 21: 16
DOI 10.1007/s00894-014-2553-5

ORIGINAL PAPER

Computational studies of water and carbon dioxide interactions with cellobiose

Faranak Bazooyar · Martin Bohlén · Kim Bolton

Received: 22 June 2014 / Accepted: 30 November 2014 / Published online: 24 January 2015

Abstract B3LYP/6–311++G** with dispersion correction (DFT-D) was used to study local and global minimum energy structures of water (H_2O) or carbon dioxide (CO_2) bonding with a pair of cellobiose molecules. The calculations showed that neither the H_2O nor the CO_2 prefer to be between the cellobiose molecules, and that the minimum energy structures occur when these molecules bond to the outer surface of the cellobiose pair. The calculations also showed that the low energy structures have a larger number of inter-cellobiose hydrogen bonds than the high energy structures. These results indicate that penetration of H_2O or CO_2 between adjacent cellobiose pairs, which would assist steam or supercritical CO_2 (SC-CO_2) explosion of cellulose, is not energetically favored. Comparison of the energies obtained with DFT-D and DFT (the same method but without dispersion correction) show that both hydrogen bonds and van der Waals interactions play an important role in cellobiose-cellobiose interactions.

Keywords Cellobiose · CO_2 · DFT · Dispersion correction · H_2O

Introduction

Fossil fuel and natural gas reserves are limited and the use of these energy sources has a large environmental impact. Hence, alternative and preferably renewable sources need to be identified to support social and technological development. One such source is lignocellulosic biomass which can, for example, be converted to biofuel.

Lignocellulosic biomass, which stems mainly from forestry waste, agricultural residue, and some municipal waste, is currently the largest source of biofuel [1, 2]. Due to this, there has been increasing focus on the conversion of this biomass to fuel, including improving the conversion efficiency. Different methods and materials have been used to break down the lignocellulosic structure, which is required for its conversion into the smaller biofuel molecules (such as ethanol and methane) [3]. Although the presence of lignin and hemicelluloses increases the recalcitrance of lignocellulosic material to hydrolysis, it is believed that it is the intermolecular bonding between cellulose chains as well as its crystalline structure that is the bottleneck for efficient conversion into biofuel [4, 5].

In order to improve the conversion efficiency of the cellulosic material into biofuels, a pretreatment step is usually included in the production [3]. The aim of this step is to remove lignin and hemicellulose, dissolve cellulose microfibrils and disrupt their crystallinity so that they are more susceptible to, for example, biological attack. Numerous solvents have been examined for the pretreatment, and several physical, chemical, and physico-chemical methods have been studied [3]. Steam explosion and supercritical CO_2 (SC-CO_2) explosion are two physico-chemical methods that are commonly used. In these pretreatment processes the biomass is exposed to H_2O or CO_2 at high temperatures and pressures, before there is a sudden drop in pressure [6–19].

Computational studies complement experimental research by offering easy control, manipulation and analysis at the molecular level. It is expected that insights obtained at this level can assist in understanding experimental results and identifying improved or new experimental methods. Molecular-level studies can be performed using accurate first principles techniques or methods based on analytic force fields. Although first principles techniques may yield reliable

F. Bazooyar (✉) · M. Bohlén · K. Bolton
School of Engineering, University of Borås, 501 90 Borås, Sweden
e-mail: faranak.bazooyar@hb.se

F. Bazooyar
Department of Chemical Engineering, Chalmers University of Technology, Gothenburg 412 96, Sweden

results, their computational expense limits them to studies of small model systems. The methods based on force fields can be used to study larger systems for longer times, but the chemical relevance of the results depends, among other things, on the validity of the force field [20, 21].

First principles studies of cellulose often use cellobiose as the model system since it is the smallest repeat unit of cellulose [22–32]. For example, Stortz et al. have shown that the B3LYP functional with a basis set that includes diffuse terms yields accurate structures and relative energies for cellobiose [20]. Calculations performed with B3LYP/6–311++G** yielded the correct *anti* conformer as the lowest energy structure for cellobiose in vacuum, and showed that the addition of at least two H_2O molecules that surround the conformer change this lowest energy structure to the *syn* conformer. Similar studies showed that the COMPASS force field capture these properties, although the *syn* conformer is only obtained in bulk water at temperatures above 298 K (when the pressure is 1 bar). Since the COMPASS force field also predicts the correct crystalline geometry [33], it was suggested that this force field may be used in simulations of larger models of steam explosion [34].

In a previous study [26] we used DFT with the B3LYP/6-311G basis set to study the interaction of glucose and cellobiose pairs with one water molecule. This method has the limitation that mainly van der Waals interactions are not included. Since intermolecular interactions such as hydrogen bonds and van der Waals interactions are expected to hinder dissolution of the cellulose crystal, the present contribution extends this work by using the B3LYP/6–311++G** basis set with Grimme's dispersion correction (DFT-D) [35–37].

The present study, which yields information on cellobiose-cellobiose bonding mechanisms and interactions between an H_2O or CO_2 with the cellobiose pair, is a first step toward using computational methods to gain a deeper understanding of the far more complex steam and SC-CO_2 explosion mechanisms. A more complete investigation of explosion may well require molecular dynamics or Monte Carlo simulations at high pressures and temperatures. To perform these calculations in a tractable computational time one needs a valid force field. This is beyond the scope of the present contribution. The first goal of this study is therefore to determine if the CO_2 molecule yields significantly different low energy structures compared to when the H_2O interacts with the cellobiose pair. If this is the case then different mechanisms may be expected for steam and SC-CO_2 explosion. Although the crystalline structure present in the steam and SC-CO_2 explosion is very different from the cellobiose studied here, the types of intermolecular interactions — hydrogen and van der Waals bonding — will be qualitatively the same. The second goal therefore is to investigate the relative importance of the inter-cellobiose hydrogen and van der Waals bonding and how this may differ between the H_2O and CO_2 complexes. This is achieved by comparing the B3LYP/6–311++G** with dispersion correction and B3LYP/6–311++G** results.

Computational methods

First principles methods

First principles methods are expected to give reasonably accurate data for disaccharides, including the glycosidic bond strength, which may be affected by the electron pairs on the oxygen atom that is involved in the bond as well as those of nearby oxygen atoms [32]. Previous studies have shown that, among different density functional theory (DFT) methods, the B3LYP [38–42] functional combined with a basis set that includes diffuse and polarization terms yields accurate relative energies and structures of hydroxyl-containing compounds like cellobiose [20, 31]. In addition, results that are obtained from these large basis sets do not need to be corrected for basis set superposition (BSSE) errors for the systems studied here [43]. Hence, similar to those studies, the B3LYP/6–311++G** method is used here.

Since DFT methods underestimate van der Waals energies, which may be important between the cellobiose molecules and between the H_2O or CO_2 and the cellobiose pair, we consider the effect of including dispersion corrections to the B3LYP/6–311++G** results. This was done by including Grimme's dispersion corrections to the B3LYP/6–311++G** results. For the sake of brevity this method is called DFT-D (DFT with dispersion corrections). To quantify the contribution of the dispersion to the total intermolecular energy we also calculate the DFT energy (without dispersion) for the DFT-D optimized geometries, i.e., we calculate the DFT//DFT-D energies. The difference between the DFT-D and DFT//DFT-D energies is the contribution from the dispersion, which is a measure of the van der Waals interactions. These can be compared to the non-dispersion contribution (mainly hydrogen bonds) obtained by subtracting the dispersion contribution from the total (DFT-D) interaction energy.

The first principles calculations, which are described below, were done using the general atomic and molecular electronic structure system (GAMESS) program [44].

Molecular mechanics force field

The condensed-phase optimized molecular potentials for atomistic simulation studies (COMPASS) force field has been discussed in detail by Sun [45] and is only briefly described here for the sake of completeness. The intramolecular terms are bond stretching, angle bending, cross-terms, and out-of-plane torsions and wags, while intermolecular interactions include electrostatic and van der Waals terms. The 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 experimental data. The fitting was done for a variety of materials including metals, metal oxides, some metal ions, inorganic small molecules, most common organics, and polymers [46]. This force field is also suitable for studies of cellulose and cellobiose [26, 27, 47–52]. Calculations done with this force field were performed using the Materials Studio Software (Accelrys Software Inc.).

Simulation methods

As described below, the initial structures for most of the DFT-D geometry optimizations were obtained from annealing simulations using the COMPASS force field. Since none of the structures had the H_2O or CO_2 molecule between the cellobiose molecules (since this was not a preferred structure according to the COMPASS force field), six DFT-D geometry optimizations were initialized with the H_2O or CO_2 between the cellobiose molecules. These geometries were therefore constructed by hand. The cellobiose molecules, as well as the H_2O or CO_2, were initially in their minimum energy structures and the cellobiose molecules were parallel to each other. The H_2O or CO_2 was placed between the center of masses of the cellobiose molecules or between neighboring glucose units, with the separation between the nearest atoms on the H_2O/CO_2 and the nearest atom on cellobiose molecules ranging from 1.6 to 4.7 Å (to avoid starting with a structure that was too high in energy). All of these geometry optimizations resulted in the H_2O or CO_2 moving from being between the molecules to the outside of the cellobiose pair. That is, in the geometry optimized structures the H_2O or CO_2 bonded to the outer surface of the cellobiose pair (similar structures are discussed below with reference to Figs. 3 and 5). The same trends were observed when using the COMPASS force field. Hence, the DFT-D and COMPASS methods predict that the H_2O or CO_2 prefers to bond to the outer surface of the cellobiose pair, and the COMPASS force field was used to identify many H_2O-cellobiose pair and CO_2-cellobiose pair local minimum energy structures, which were used as input for the DFT-D geometry optimizations.

These structures were obtained using simulated annealing. Since the goal was to obtain different high and low energy local minimum energy structures, 50–100 cycles with 4–8 million simulation steps per cycle were simulated. The Verlet integration algorithm, which has the advantage of being formally time-reversible [53], was used with a step size of 1 fs. The mid-cycle temperature for the H_2O systems was between 300 and 340 K, and for the CO_2 systems it was between 170 and 275 K. These temperatures were sufficiently high to allow for sampling of large regions of configuration space while still preventing excessive evaporation of the H_2O or CO_2 molecule from the cellobiose pair.

Ten geometries were used as input for the annealing to further increase the configuration space that was sampled. These geometries had different orientations of the cellobiose molecules relative to each other (parallel, anti-parallel, perpendicular, and when one of the cellobiose was rotated so that there was a 90° angle between the molecular planes of the cellobiose molecules) and where the H_2O or CO_2 molecule was placed between the cellobiose molecules or at different sites on the surface of the cellobiose pair. These structures were geometry optimized before being used as input for the annealing simulations. The choice of the annealing parameters enabled identification of local minimum energy structures from all of these regions of configuration space, and many of the annealed structures obtained from the different initial structures were very similar (i.e., the annealing linked the regions of configuration space spanned by the initial structures).

Ten geometries were typically chosen from each of the ten annealing simulations for further analysis. The selection was done so that both high and low energy (including the lowest energy) structures were included. These structures were geometry optimized with the COMPASS force field using a combination of conjugate gradient [54], Newton [55], and steepest descent [56] methods. The structures were considered to be minimized once the change in energy between subsequent steps was less than 2.0×10^{-5} kcal mol^{-1}. Since some of these structures were the same, this procedure resulted in 90 unique structures for the H_2O-cellobiose pair and 80 unique structures for the CO_2-cellobiose pair. These structures were used as input for the DFT-D geometry optimizations. These geometry optimizations were performed using a gradient convergence tolerance of 6.28×10^{-3} kcal $(mol\times bohr)^{-1}$ and a RMS gradient tolerance of 2.09×10^{-3} kcal $(mol\times bohr)^{-1}$.

Analysis

Several parameters were analyzed to ascertain whether H_2O and CO_2 induced significantly different local minimum energy structures. These included the relative energies of the cellobiose pair and the size of the dispersion correction for the different structures. Several geometrical parameters were analyzed (e.g., the relative orientation of the cellobiose molecules, the relative positioning of the reducing and non-reducing ends, the orientation of the carbonyl groups, the positions of the H_2O and CO_2 molecules), but the clearest difference between the high and low energy geometries was the number of hydrogen-bonds (H-bonds) that linked the two cellobiose molecules. Hence, this is discussed in detail below, where the H-bond is defined by a maximum separation of 2.5 Å between the H and O atoms on the different molecules and a minimum angle of 90° formed by the H–O bond on one molecule and the H atom on the second molecule. The trends

discussed below are not expected to be sensitive to this definition.

The strength of the H_2O/CO_2–cellobiose pair interaction is:

$$E_{(X-pair)} = E_{(X+pair)} - E_{pair} - E_X, \quad (1)$$

where $E_{(X+pair)}$ is the energy of the geometry optimized structure, E_{pair} is the energy of the cellobiose pair and E_X is the energy of the H_2O or CO_2. The cellobiose pair structure (used to obtain E_{pair}) was subsequently used to obtain the intermolecular energy between the cellobiose molecules, which is:

$$E_{(inter-cellob)} = E_{pair} - E_{Cellob.1} - E_{Cellob.2}, \quad (2)$$

where $E_{Cellob.1}$ and $E_{Cellob.2}$ are the energies of the separated cellobiose molecules. Note that the structures used to obtain E_{pair}, $E_{Cellob.1}$, and $E_{Cellob.2}$ were the same as those obtained from geometry optimization of H_2O/CO_2–cellobiose pair system (i.e., there was no further optimization of the individual cellobiose molecules or the cellobiose pair). This was done since the aim was to analyze the strength of the intermolecular interactions in the H_2O/CO_2–cellobiose pair complex, and further geometry optimization would also have included the contribution of the intramolecular energies. As discussed below, comparison between the DFT-D and DFT//DFT-D results reveals the relative importance of the non-dispersion and dispersion interactions.

The conformation (*anti* or *syn*) of the cellobiose was also analyzed since it is known that it is *syn* in the cellulose crystal structure and *anti* in the cellobiose structure in vacuum. Possible changes from *syn* to *anti* will be important during steam or SC-CO_2 explosion since this would distort the crystal structure and make the crystal more susceptible to biological attack. The conformation is given by φ_H, which is the dihedral angle defined by H1–C1–O1–C4′ as shown in Fig. 1. A φ_H near 180° (or −180°) reveals the *anti* (flipped) conformer and a φ_H near 60° (or −60°) reveals the *syn* (normal) conformer [27, 40].

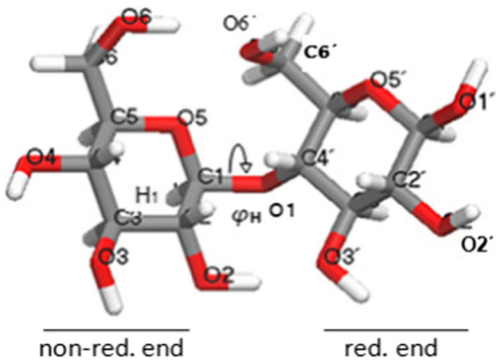

Fig. 1 Structure of the *anti* (*flipped*) conformer of β-cellobiose. The atom numbering is used in the discussion of the H-bonding below, and the dihedral angle φ_H is defined by H1–C1–O1–C4′. The non-reducing and reducing ends are also shown

Results and discussion

H_2O-cellobiose pair

Figure 2 shows the relative energies, ΔE, of the 90 unique H_2O-cellobiose pair local minimum energy structures, ordered according to energies obtained from the DFT-D calculations. The energies are relative to the DFT-D energy of the lowest energy structure. Several aspects are revealed from the figure. First the lowest energy structure obtained from DFT-D also has the lowest DFT//DFT-D energy. This lowest energy structure is discussed in more detail below. Second, the energy difference between the high and low energy structures is ~25 kcal mol^{-1} within the DFT-D series and ~15 kcal mol^{-1} within the DFT//DFT-D series. As discussed below, the difference between the change in DFT-D and DFT//DFT-D energies is due to the extra stability that the dispersion contributes to the low energy structure. Third, the dispersion correction yields DFT-D energies that are ~50 kcal mol^{-1} lower in energy than the DFT//DFT-D results.

The local minimum structures were analyzed to ascertain if there are any general differences and similarities between the high and low energy structures. There are no significant trends regarding differences in the binding position of the H_2O molecule on the surface of the cellobiose pair, the relative positions of the reducing and non-reducing ends of the cellobiose molecules, the conformations of the cellobiose molecules (which are typically *anti*) or the relative orientations of the carbonyl groups on the cellobiose molecules. However, as exemplified in Fig. 3, the low energy structures 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 maximizes the

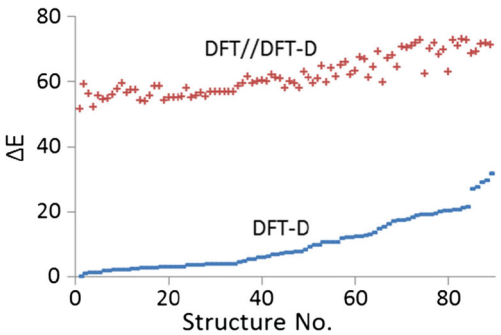

Fig. 2 Relative energies (in kcal mol^{-1}) of local minimum H_2O-cellobiose pair structures obtained from DFT-D and DFT//DFT-D (i.e., DFT energies of the DFT-D optimized structures)

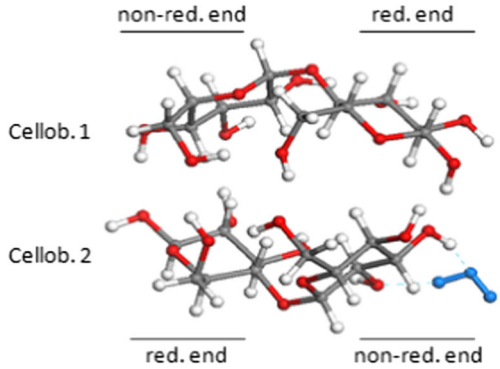

Fig. 3 Lowest energy H_2O-cellobiose pair structure obtained from DFT-D. The H_2O molecule is shown in blue. The reducing and non-reducing ends of each cellobiose molecule are shown

number of hydrogen bonds (which is between 5 and 7 in the low energy structures) and the van der Waals energy.

The structures with intermediate energies (~structures 40–85 in Fig. 2) also consist of cellobiose molecules that lie parallel to each other, but the cellobiose molecules are shifted relative to each other such that the two glucose units of the first molecule are not directly above those on the second molecule. This results in fewer hydrogen bonds (3–5) and reduced van der Waals attraction. The cellobiose molecules in the structures with the highest energies have almost no overlap of the glucose units. There are only one or two hydrogen bonds and the van der Waals interactions are far weaker (as quantified below).

Table 1 lists bond lengths, angles and torsions of the two cellobiose molecules in the minimum energy structure shown in Fig. 3. The data in the table were chosen since they represent different types of bonds, angles and torsions. It is evident that the bond lengths and angles are similar for the two cellobiose molecules. The torsion angles that are formed by carbon and oxygen atoms are also similar, whereas torsion angles that end with a hydrogen atom can show a significant difference between the two molecules (e.g., C5′-C6′-O6′-H). This is expected since the hydrogen atoms readily rotate between local minimum structures that have similar energies so that the intermolecular interactions — including H-bonding — are strengthened. The separation between the cellobiose centers of mass is 4.59 Å. Both cellobiose molecules have the *anti* conformation, with φ_H=−173.7 and 179.2° for Cellob.1 and Cellob.2, respectively.

There are six hydrogen bonds in the lowest energy structure (shown in Fig. 3) and these 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 atom numbers are given in Fig. 1. The second column in Table 2 shows the DFT-D intermolecular energy between the cellobiose molecules for this structure and 11 other local minimum energy structures. The structures were chosen to represent low, intermediate, and high energy

Table 1 Representative bond lengths (Å), bond angles (°), and torsions (°) in the cellobiose molecules (Cellob.1 and Cellob.2 in Fig. 3) of the DFT-D minimum energy structure

		Cellob.1	Cellob.2
Bond lengths	O1-C4′	1.436	1.435
	C4′-C5′	1.543	1.546
	C5′- O5′	1.432	1.436
	O6′-C6′	1.423	1.417
	O1′-C1′	1.389	1.401
Angles	C1-O1-C4′	120	118
	O1-C4′-C5′	110	108
	C4′-C5′-O5′	110	112
	C4′-C5′-C6′	114	114
	O5′-C1′-O1′	109	108
Torsions	C1-O1-C4′-C5′	118	123
	O1-C4′-C5′-O5′	173	169
	C4′-C5′-O5′-C1′	63	54
	C4′-C5′-C6′-O6′	57	56
	C5′-O5′-C1′-O1′	176	179
	O5′-C1′-C2′-O2′	174	176
	C5′-C6′-O6′-H	102	60
	O5′-C1′-O1′-H	75	64
	C1′-C2′-O2′-H	61	93

structures (the numbers in the table are the same as those in Fig. 2). It is evident that the intermolecular energy decreases with increasing structure number (increasing relative energy). The difference between DFT-D and DFT//DFT-D energies for each structure is shown in the third column in Table 2. These are the dispersion contributions to the inter-cellobiose energies, and they also decrease with increasing structure number. The difference between the values in the second and third columns is the non-dispersion contribution, which also decreases with increasing structure number. Hence, both types of inter-cellobiose interactions get weaker as the structures

Table 2 Cellobiose-cellobiose intermolecular energies ($E_{inter\text{-}pair}$) and the dispersion correction energies (E_{disp}) in kcal mol^{-1} for some structures shown in Fig. 2

Structure no.	$E_{inter\text{-}pair}$	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)

become less stable (as their energy increases). The numbers in parentheses in the third column are the percentage contribution of the dispersion interactions to the inter-cellobiose energies. There is no clear trend of this percentage contribution increasing or decreasing with increasing structure number, which indicates that the dispersion and non-dispersion contributions decrease equally rapidly as the energy of the structure increases.

The H_2O molecule is attached to the cellobiose pair by two hydrogen bonds in the minimum energy structure shown in Fig. 1. The H-bonds are with the H-O3 and the O2 atoms of Cellob.2 shown in Fig. 3, and the distance between the center of masses of H_2O and Cellob.2 is 5.63 Å. The intermolecular energy between the H_2O and the cellulose pair is −15.5 kcal mol^{-1}. As mentioned above, there are no clear trends of changes in this energy as the total energy of the H_2O-cellobiose pair structure increases. For example, the energy between the H_2O and the cellulose pair is −14.1 kcal mol^{-1} for structure 46, and it is −18.2 kcal mol^{-1} for structure 88.

CO_2-cellobiose pair

The trends observed for the H_2O-cellobiose pair systems are also seen for the CO_2-cellobiose pair systems. Figure 4 shows the relative energies of the 80 unique CO_2-cellobiose pair local minimum energy structures. The minimum energy structure obtained from the DFT-D calculations also has the lowest DFT//DFT-D energy, and the trend of increasing relative energies from structures 1 through 80 is the same for DFT-D and DFT//DFT-D calculations. The difference in DFT-D energies between the highest and lowest energies structures is ~25 kcal mol^{-1} (which was the same for the H_2O-cellobiose pair systems) and this difference in DFT//DFT-D energies is ~15 kcal mol^{-1} (which was also the same for the H_2O systems). Since DFT//DFT-D does not include the dispersion contribution to the stabilization of the low energy structures, the DFT//DFT-D energies are ~60 kcal mol^{-1} higher than the DFT-D energies.

Similar to the H_2O-cellobiose pair systems, the lowest energy CO_2-cellobiose pair structures are parallel such that the inter-cellobiose attraction is maximized (the minimum energy CO_2-cellobiose pair structure is discussed below with reference to Fig. 5). There are 5–7 H-bonds in the structures with the lowest energies (~structures 1–15 in Fig. 4) and, as discussed below, there are strong dispersion attractions. The cellobiose molecules are shifted relative to each other in the structures that have intermediate energies (~structures 16–60) which 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.

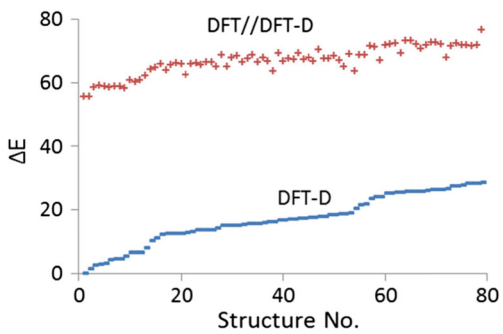

Fig. 4 The same as Fig. 2 but for the CO_2-cellobiose pair structures

The minimum energy structure for the CO_2-cellobiose pair system is shown in Fig. 5 and data for this structure is shown in Table 3. The separation between the cellobiose centers of mass is 4.10 Å. Both cellobiose molecules have the *anti* conformation, with φ_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. The atom numbers are those given in Fig. 1.

It is evident from Table 3 that the structures of Cellob.1 and Cellob.2 (Fig. 5) are very similar. The only large differences are in some of the torsion angles that have hydrogen as an end atom. As discussed above with reference to the H_2O-cellobiose pair system, this is because there is a small torsion barrier between these local minima, and the energy difference between the minima is also small.

The second column in Table 4 shows that the inter-cellobiose energy decreases with increasing structure number (i.e., with increasing relative energy). The dispersion energy and the non-dispersion contribution to the intermolecular energy (difference between columns two and three) also decrease. Similar to the H_2O-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.

The distance between the center of mass of the CO_2 molecule and the cellobiose pair in the minimum energy structure shown in Fig. 5 is 7.51 Å. Neither this distance, nor the CO_2-

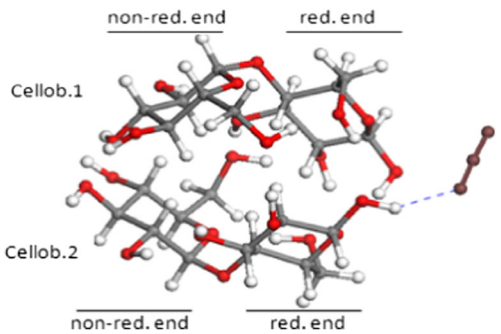

Fig. 5 The same as Fig. 3 but for the CO_2-cellobiose pair minimum energy structure. The CO_2 molecule is shown in brown

Table 3 The same as Table 1 but for the CO_2-cellobiose pair minimum energy structure

		Cellob.1	Cellob.2
Bond lengths	O1-C4′	1.429	1.432
	C4′-C5′	1.540	1.545
	C5′- O5′	1.432	1.438
	O6′-C6′	1.426	1.419
	O1′-C1′	1.394	1.388
Angles	C1-O1-C4′	117	118
	O1-C4′-C5′	109	109
	C4′-C5′-O5′	110	108
	C4′-C5′-C6′	115	116
	O5′-C1′-O1′	107	108
Torsions	C1-O1-C4′-C5′	−123	−132
	O1-C4′-C5′-O5′	−179	174
	C4′-C5′-O5′-C1′	57	67
	C4′-C5′-C6′-O6′	66	49
	C5′-O5′-C1′-O1′	−173	−177
	O5′-C1′-C2′-O2′	174	171
	C5′-C6′-O6′-H	46	−70
	O5′-C1′-O1′-H	68	−64
	C1′-C2′-O2′-H	80	106

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^{-1} for structures 1, 21, and 72, respectively. In the minimum energy structure (structure 1) the dispersion between CO_2 and cellobiose pair contributes −3.7 kcal mol^{-1} to the total intermolecular energy (−8.2 kcal mol^{-1}). This can be compared to the DFT-D intermolecular energy of −5.49 kcal mol^{-1} between CO_2 and a single molecule of cellobiose, of which −3.22 kcal mol^{-1} is due to dispersion forces. The large interaction energy is due to the fact that the CO_2 molecule interacts with two –OH groups of the cellobiose molecule. To put this into perspective, and to compare with previous calculations, we note that the DFT-D interaction energy (using the B3LYP/6–311++G**) between CO_2 and ethanol is −4.28 kcal mol^{-1} (of which −1.97 kcal mol^{-1} is due to dispersion). This can be compared to an MP2/6–311++G** intermolecular energy of −4.22 kcal mol^{-1}. This is slightly larger than the value of −2.96 kcal mol^{-1} that was obtained previously using MP2 with a smaller (Dunning's aug-cc-pVDZ) basis set [57].

Table 4 The same as Table 2 but for CO_2-cellobiose pair systems

Structure no.	$E_{inter\text{-}pair}$	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)

Comparison between the H_2O-cellobiose pair and CO_2-cellobiose pair results

The results obtained from the H_2O and CO_2 systems are similar, indicating that the intermolecular bonding between the H_2O or CO_2 molecule and the cellobiose pair does not significantly influence the minimum energy structures, or the trends observed between structures with increasing relative energies. For example, the structures with the lowest energies have the largest number of H-bonds and strongest van der Waals attractions, and are approximately 25 kcal mol^{-1} (DFT-D) or 15 kcal mol^{-1} (DFT//DFT-D) lower in energy than the high energy structures. Both non-dispersion and dispersion interactions have large relative contributions to the cellobiose-cellobiose intermolecular interactions for all of the structures. Also, the H_2O and CO_2 molecules prefer to bind to the outer surface of the cellobiose pair instead of being located between the cellobiose molecules.

However, there are some differences. For example, the distance between the centers of mass between the H_2O and the cellobiose pair (5.63 Å) in the lowest energy structure is smaller than between the CO_2 and the cellobiose pair (7.51 Å). This is expected since the water is more strongly bound to the cellobiose pair. It is also of interest that the distance between the centers of mass of the two cellobiose molecules is larger (4.59 Å) when they interact with the H_2O molecule than when they interact with the CO_2 molecule (4.10 Å). This is consistent with the weaker cellobiose-cellobiose intermolecular energy for the pair that interacts with the H_2O molecule (−51.1 compared to −56.6 kcal mol^{-1} for the CO_2 complex). Hence, the increase in intermolecular attraction to the water molecule — which results in increased electron density being located between the cellobiose pair and the water molecule — decreases the interaction strength (electron density) between the cellobiose molecules.

Conclusions

Previous studies [26, 27] have shown that the B3LYP/6–311++G** density functional method yields valid energies and structures for cellobiose and H_2O-cellobiose systems. This method, including Grimme's dispersion correction, was

used to study the interactions between H_2O and two cellobiose molecules, as well as CO_2 and two cellobiose molecules. The results obtained with and without the dispersion corrections are presented, since they enable an estimation of the relative contributions of non-dispersion (mainly H-bonding) and dispersion (van der Waals bonding) terms to the intermolecular energies.

Geometry optimization with the DFT-D method showed that the H_2O and CO_2 molecules prefer to bond to the surface of the cellobiose pair as opposed to being located between the cellobiose molecules. Also, comparison of 90 unique H_2O-cellobiose pair and 80 unique CO_2-cellobiose pair local minimum energy structures showed that the trends in relative energies between the low and high energy structures were the same with and without dispersion correction. The structures with lower energies typically have a larger number of H-bonds and stronger van der Waals interactions.

Comparison of the DFT//DFT-D and DFT-D cellobiose-cellobiose intermolecular energies showed that both the non-dispersion and dispersion terms have large contributions to the intermolecular energies. The contribution of the dispersion (and non-dispersion) energies was typically between 30 and 70 %, and there was no clear trend in increasing or decreasing this contribution with weaker intermolecular bonds.

The distance between the centers of mass of the two cellobiose molecules is larger (4.59 Å) when they interact with the H_2O molecule than when they interact with the CO_2 molecule (4.10 Å). This is consistent with the weaker cellobiose-cellobiose intermolecular energy for the pair that interacts with the H_2O molecule. Hence, the increase in intermolecular attraction to the water molecule — which results in increased electron density being located between the cellobiose pair and the water molecule — decreases the interaction strength between the cellobiose molecules.

Acknowledgments Financial support from Stiftelsen Föreningssparbanken Sjuhärad and The Carl Trygger Foundation for Scientific Research is gratefully acknowledged. GAMESS-US, version 1 OCT 2010 (R1) was used at the high performance computer cluster Kalkyl at UPPMAX (Uppsala Multidisciplinary Centre for Advanced Computational Science), Uppsala Sweden. Calculations using the COMPASS force field were done using the program from Accelrys Software Inc.

References

1. Zhang Q, Bulone V, Ågren H, Tu Y (2011) A molecular dynamics study of the thermal response of crystalline cellulose Iβ. Cellulose 18:207–221
2. Wyman CE (2007) What is (and is not) vital to advancing cellulosic ethanol. Trends Biotechnol 25:153–157
3. Taherzadeh MJ, Karimi K (2008) Pretreatment of lignocellulosic wastes to improve ethanol and biogas production: a review. Int J Mol Sci 9:1621–1651
4. Fan LT, Lee Y, Beardmore DH (1980) Mechanism of the enzymatic hydrolysis of cellulose: effects of major structural features of cellulose on enzymatic hydrolysis. Biotechnol Bioeng 22:177–199
5. Wyman CE (1996) 1: Ethanol production from lignocellulosic biomass: overview. In: Wyman CE (ed) Handbook on bioethanol: production and utilization. Taylor and Francis, Washington, pp 1–18
6. Mason W H (1926) Process and apparatus for disintegration of wood and the like. US Pat. 1578609
7. Babcock L W (1932) Method of producing fermentable sugars and alcohol from wood. US Pat. 1855464
8. Dahman Y, Volynets B (2011) Assessment of pretreatments and enzymatic hydrolysis of wheat straw as a sugar source for bioprocess industry. Int J Energy Environ (IJEE) 2:427–446
9. Martin-Sampedro R, Rojas OJ, Capanema EA, Hoeger I, Villar JC (2011) Lignin changes after steam explosion and laccase-mediator treatment of eucalyptus wood chips. J Agric Food Chem 59:8761–8769
10. Jedvert K, Theliander H, Saltberg A, Lindström ME (2012) Mild steam explosion and chemical pre-treatment of Norway spruce. Bio Resources 7:2051–2074
11. 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. Cellul Chem Technol 45:495–501
12. Kumar P, Barrett DM, Delwiche MJ, Stroeve P (2009) Methods for pretreatment of lignocellulosic biomass for efficient hydrolysis and biofuel production. Ind Eng Chem Res 48:3713–3729
13. Zheng Y, Lin HM, Wen J, Cao N, Yu X, Tsao GT (1995) Supercritical carbon dioxide explosion as a pretreatment for cellulose hydrolysis. Biotechnol Lett 17:845–850
14. Zheng Y, Tsao GT (1996) Avisel hydrolysis by cellulose enzyme in supercritical CO2. Biotechnol Lett 18:451–454
15. Zheng Y, Lin H, Tsao GT (1998) Pretreatment for cellulose hydrolysis by carbon dioxide explosion. Biotechnol Prog 14:890–896
16. Narayanaswamya N, Faik A, Goetz DJ, Gu T (2011) Supercritical carbon dioxide pretreatment of corn stover and switchgrass for lignocellulosic ethanol production. Bioresour Technol 102:6995–7000
17. Alinia R, Zabihi S, Esmaeilzadeh F, Kalajahi F (2010) Pretreatment of wheat straw by supercritical CO_2 and its enzymatic hydrolysis for sugar production. J Biosystems Eng 107:61–66
18. Muratov G, Kim C (2002) Enzymatic hydrolysis of cotton fibers in supercritical CO2. Biotechnol Bioprocess Eng 7:85–88
19. Gua T, Held MA, Faikc A (2013) Supercritical CO2 and ionic liquids for the pretreatment of lignocellulosic biomass in bioethanol production Environ. Technol 34:1735–1749
20. Stortz CA, Johnson GP, French AD, Csonka GI (2009) Comparison of different force fields for the study of disaccharides. J Carbohydr Res 344:2217–2228
21. Bergenstråhle M, Matthews J, Crowley M, Brady J (2010) Cellulose crystal structure and force fields. International Conference on Nanotechnology for the forest products industry. Otaniemi, Espoo, Finland
22. Payal RS, Bharath R, Periyasamy G, Balasubramanian S (2012) Density functional theory investigations on the structure and dissolution mechanisms for cellobiose and xylan in an ionic liquid: gas phase and cluster calculations. J Phys Chem B 116:833–840
23. Momany FA, Schnupf U (2011) DFTMD studies of β-cellobiose: conformational preference using implicit solvent. Carbohydr Res 346:619–630
24. French AD, Johnson GP (2006) Quantum mechanics studies of cellobiose conformations. Can J Chem 84:603–612
25. French AD, Johnson GP (2008) Roles of starting geometries in quantum mechanics studies of cellobiose. Mol Simul 34:365–372
26. Bazooyar F, Taherzadeh M, Niklasson C, Bolton K (2013) Molecular modelling of cellulose dissolution. J Comput Theor Nanosc 10:2639–2646

27. Bazooyar F, Momany FA, Bolton K (2012) Validating empirical force fields for molecular-level simulation of cellulose dissolution. Comput Theor Chem 984:119–127
28. Strati GL, Willett JL, Momany FA (2002) Ab initio computational study of β-cellobiose conformers using B3LYP/6-311++ G**. Carbohydr Res 337:1833–1849
29. Strati GL, Willett JL, Momany FA (2002) A DFT/ab initio study of hydrogen bonding and conformational preference in model cellobiose analogs using B3LYP/6-311++ G**. Carbohydr Res 337:1851–1859
30. Bosma WB, Appell M, Willett JL, Momany FA (2006) Stepwise hydration of cellobiose by DFT methods: 2. Energy contributions to relative stabilities of cellobiose $(H_2O)_{1-4}$ complexes. THEOCHEM 776:1–19
31. Momany FA, Willett JL (2000) Computational studies on carbohydrates: I. Density functional ab initio geometry optimization on maltose conformations. J Comput Chem 21:1204–1219
32. French AD, Johnson GP (2004) Advanced conformational energy surfaces for cellobiose. Cellulose 11:449–462
33. Aldred P http://www.accelrys.com/resource-center/case-studies / archive /studies /cellulose.html. Accessed 12 April 2014
34. Bazooyar F, Bolton K (2014) Molecular-level simulations of cellulose steam explosion. Quantum Matter. doi:10.1166/qm.2015.1178
35. Grimme S (2004) Accurate description of van der Waals complexes by density functional theory including empirical corrections. J Comput Chem 25:1463–1473
36. Piacenza M, Grimme S (2005) Van der Waals interactions in aromatic systems: structure and energetics of dimers and trimers of pyridine. Chem Phys Chem 6:1554–1558
37. Grimme S, Antony J, Ehrlich S, Krieg H (2010) A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. J Chem Phys 132:154104
38. Grimme S (2006) Semiempirical GGA-type density functional constructed with a long-range dispersion correction. J Comput Chem 27: 1787–1799
39. Becke AD (1993) Density-functional thermochemistry. III. The role of exact exchange. J Chem Phys 98:5648
40. Lee C, Yang W, Parr RG (1998) Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. Phys Rev B 37:785–789
41. Miehlich B, Savin A, Stoll H, Preuss H (1989) Results obtained with the correlation energy density functionals of becke and Lee, Yang and Parr. Chem Phys Lett 157:200–206
42. Stephens PJ, Devlin FJ, Chabalowski CF, Frisch MJ (1994) Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. J Phys Chem 98:11623–11627
43. Lii JH, Ma B, Allinger NL (1999) Importance of selecting proper basis set in quantum mechanical studies of potential energy surfaces of carbohydrates. J Comput Chem 20:1593–1603
44. Schmidt MW, Baldridge KK, Boatz JA, Elbert ST, Gordon MS, Jensen JH, Matsunaga N, Koseki S, Nguyen KA, Su SJ, Windus TL, Dupuis M, Montgomery JA (1993) General atomic and molecular electronic structure system. J Comput Chem 14: 1347–1363
45. Sun H (1998) COMPASS: an ab initio force-field optimized for condensed-phase applications overview with details on alkane and benzene compounds. J Phys Chem B 102:7338–7364
46. Bunte SW, Sun H (2000) Molecular modeling of energetic materials: the parameterization and validation of nitrate esters in the COMPASS force field. J Phys Chem B 104:2477–2489
47. Sarko A, Muggli R (1974) Packing analysis of carbohydrates and polysaccharides. III. Valonia cellulose and cellulose II. Macromolecules 7:486–494
48. Wu X, Wagner R, Raman A, Moon R, Martini A (2010) Elastic deformation mechanics of cellulose nanocrystals. Miner, Met Mater Soc (TMS) 2:689–695
49. Wu X, Moon R, Martini A (2011) Calculation of single chain cellulose elasticity using fully atomistic modelling. Tech Assoc Pulp Paper Indust J (TAPPI J) 10:37–43
50. Miyamoto H, Umemura M, Aoyagi T, Yamane C, Ueda K, Takahashi K (2009) Structural reorganization of molecular sheets derived from cellulose II by molecular dynamics simulations. Carbohydr Res 344: 1085–1094
51. Miyamoto H, Yamane C, Ueda K (2013) Structural changes in the molecular sheets along (hk0) planes derived from cellulose Iβ by molecular dynamics simulations. Cellulose 20:1089–1098
52. Eichhorn SJ, Davies GR (2006) Modelling the crystalline deformation of native and regenerated cellulose. Cellulose 13:291–307
53. Bolton K, Nordholm S (1994) An evaluation of the Gauss-Radau algorithm for the simulation of chemical dynamics. J Comput Phys 113:320–335
54. Fletcher R, Reeves CM (1964) Function minimization by conjugate gradients. Comput J 7:149–154
55. Ermer O (1976) Calculation of molecular properties using force fields. Applications in organic chemistry. Struct Bond 27:161–211
56. Levitt M, Lifson S (1969) Refinement of protein conformations using a macromolecular. J Mol Biol 46:269–279
57. Lalanne P, Tassaing T, Danten Y, Cansell F, Tucker SC, Besnard M (2004) CO2-ethanol interaction studied by vibrational spectroscopy in supercritical CO2. J Phys Chem A 108:2617–2624