Bull. Chem. Soc. Ethiop. **2025**, 39(10), 2041-2064.
DOI: https://dx.doi.org/10.4314/bcse.v39i10.12

ISSN 1011-3924
Printed in Ethiopia
Online ISSN 1726-801X

AN INVESTIGATION ON THE VIBRATIONAL ANALYSIS, MOLECULAR STRUCTURE AND INTERACTION WITH THE CELLULASE ENZYMES OF THE CELLULOSE I ALPHA AND CELLULOSE I BETA: DFT CALCULATIONS, DOCKING SIMULATIONS

Aliye Demet Demirag[1], Sefa Celik[2*], Aysen E. Ozel[2], Sevim Akyuz[3]

[1]Yeditepe University, Faculty of Engineering, Department of Genetics and Bioengineering, 34755, Istanbul/Türkiye
[2]Istanbul University, Faculty of Science, Department of Physics, 34134, Istanbul/Türkiye
[3]Istanbul Kultur University, Faculty of Science and Letters, Department of Physics, 34156, Istanbul/Türkiye

(Received February 4, 2025; Revised June 12, 2025; Accepted June 17, 2025)

ABSTRACT. Molecular docking is one of the most widely used techniques for simulating molecular interactions between molecules and forecasting the mode of binding and affinities between them. Due to the presence of structure-function relationship, in this study firstly, the molecular structures of the cellulose I(alpha) and I(beta) molecules were optimized and their most stable structures were determined by density functional theory (DFT) using B3LYP method with 6-31G(d,p) basis set. The vibrational wavenumbers of 1-ring, 2-ring, 3-ring, and 4-ring structures of cellulose I(alpha) and I(beta) were calculated using the same level of theory. Reliable vibrational assignments were made based on potential energy distribution (PED %) of the vibrational modes. The energy gap ($E_g = E_{LUMO} - E_{HOMO}$) of the cellulose I_α and cellulose I_β was found to be 8.286 eV and 7.965 eV, respectively. To identify the molecular interactions between cellulose I_α and I_β ligands and the cellulase enzymes, molecular docking studies were performed. The molecular docking results revealed the strong interaction of the cellulose I_α and I_β with Endoglucanase enzyme (-6.4 and -6.3 kcal/mol, respectively), enzyme β-glucosidase (-5.3 and -5.2 kcal/mol, respectively), and Exoglucanase enzyme (-6.1 and -6.2 kcal/mol, respectively).

KEY WORDS: Cellulose I alpha, Cellulose I beta, Density functional theory, Molecular docking

INTRODUCTION

The most prevalent polymer that is produced through biosynthesis is cellulose, which is a chain polymer of the glucose [1]. It functions as a structural element, providing rigidity and durability, along with lignin and hemicelluloses, in the cell walls of annual plants and trees [2]. Cellulose is produced by specific bacteria (such as Acetobacter xylinus) and occurs in specific algae (such as Valonia ventricosa), especially in pure form and high crystallinity [3]. Hemicellulose is a term used to describe polymeric amorphous hydrocarbons found in the cell wall other than cellulose. Depending on the species, different amounts of cellulose and other compounds are present in plants [4]. Its heterogeneous structure is made up of elements from plants, including resin, cellulose, and hemicellulose as well as lignin and mineral substances.

The primary and secondary hydroxyl functional groups are the most common reactive groups in cellulose. A primary and two secondary hydroxyl functional groups can participate in the typical chemical reactions of hydroxyl groups and are present in each repeating anhydroglucose unit. Compared to secondary hydroxyls, primary hydroxyls are easier to access and more receptive. However, both kinds of hydroxyls go through a lot of the cellulose-specific chemical processes [5]. Plants do contain cellulose in its pure form, but this is less common than hemicelluloses, lignins, and other extractive materials. It is made directly from a fibrous web

*Corresponding author. E-mail: scelik@istanbul.edu.tr

and is polycrystalline with a high degree of polymerization (DP); it does not contain bacterial cellulose, lignin, pectin, hemicellulose, or any other biogenic products [6, 7]. Different three-dimensional configurations are produced by hydrogen bonds. Intramolecular hydrogen bonds and the -glycosidic covalent bond both contribute to the stiffness of the cellulose polymer [8]. The cellulose chains interact strongly with one another thanks to the intermolecular hydrogen bond. The oxygen atom at C3 and the OH atom at C6 interact to form bonds between the cellulose macromolecules of cellulose I (natural cellulose) [9]. Synchrotron X-ray and neutron diffraction experiments have provided an explanation for the layering of cellulose, namely hydrogen bonding, weak C-H-O bonds, and hydrophobic interactions [9]. Natural cellulose (cellulose I) forms of two crystal phases; Cellulose I_α and I_β, Both forms can be found within the same cellulose sample, and even along a given microfibril [10-12]. However, the relative amounts of the I_α and I_β varies in the samples from different origins [12]. In contrast to land plants, which typically produce the majority of cellulose I_β, algae and bacteria primarily produce cellulose I_α [13, 14]. By applying heat, cellulose I_α can be permanently changed back into cellulose I_β [15].

The enzymes that hydrolyze the cellulose chain into glucose monomers are known as cellulases, an enzymatic system typically made up of endoglucanase, exoglucanase, and -glucosidase [16]. The third most frequently produced industrial enzyme, cellulases account for about 20% of the global enzyme market [17]. The cost of industrial cellulose degradation is jeopardized by the high cost of cellulases and the substantial amount of these enzymes required. The effectiveness of existing cellulases has been improved, and new cellulase-producing microorganisms have been sought after [18, 19]. Depending on which component of cellulose they degrade, cellulases are categorized. Exoglucanases break down the exposed fiber ends into smaller pieces while endoglucanases randomly sever cellulose fibers. Endoglucanase, which is typically regarded as a member of the cellulase family, has a stronger affinity for cellulose and also reacts with xylan and mixed -(1-3,1-4)-glucan. Similar to endoxylanase, -glucosidase can produce a reducing and a non-reducing end by rupturing internal -(1-4) or -(1-3) bonds within glucose chains [20].

Any interaction between or within biomolecules is referred to as a "biomolecular interaction" (intermolecular interactions). The interactions can be classified as bound or unbound depending on their chemical structure. Unbonded interactions include the formation of hydrogen bonds, vdW interactions, salt bridges, and electrostatic interactions. Bonded interactions primarily refer to covalent bonds and disulfide bonds. The primary forces behind biomolecular stability, functionality, and signal transmission are these interactions. In areas like drug design, nanotechnology, the development of biosensors, and the design of biomarkers, it is crucial to investigate the nature and modeling of these interactions [21].

In order to identify biologically active sites of a molecule and to investigate its interactions with enzymes, the precise determination of its optimized structure is crucial [22]. In this study the most stable structures of cellulose I_α and I_β molecules were obtained by DFT/B3LYP/6-31G(d,p) level of theory. IR and Raman vibrational frequencies were calculated, and frontier orbital (HOMO, LUMO) and molecular electrostatic potential (MEP) analyses were performed for the optimized structures. In order to shed light on the degradation of cellulose, molecular docking studies of cellulose I_α and I_β into the cellulase enzymes (endoglucanase, exogluconase, and β-glucosidase) were performed. In our previous studies, molecular structures of cellulose II [20] and cellulose triacetate II [23] molecules were investigated. In this study, 1-4 ring structures of cellulose I_α and cellulose I_β monomers, dimers, trimers and tetramers were investigated.

EXPERIMENTAL

The accurate determination of the lowest energy structure of any molecule is necessary to determine its active sites that will interact with an enzyme. For this purpose, the most stable structures of cellulose Iα and Iβ units were determined. The Gaussian 09 package and the Gauss View 05 graphical user interface programs were used and all quantum chemical calculations were performed by using the DFT/B3LYP/6-31G(d,p) level of theory.

The initial structures of cellulose I_α and cellulose I_β units were taken straight from the crystal structures [12, 24]. In order to create the units of these molecules, H atoms were added to the oxygen atoms forming the crystallization, and the scale factors of the force constants of the internal coordinates were assumed to be close to zero. Afterwards, the created units were optimized using DFT/B3LYP/6-31G(d,p) level of theory. As a result of calculations, optimized structures of cellulose I_α and cellulose I_β monomers were found to be in good agreement with the corresponding crystal results. The two-, three- and four-ring structures of cellulose I_α and cellulose I_β were constructed by combining optimized each cellulose unit by glycosidic bonds and then the obtained structures were optimized by the DFT/B3LYP/6-31G(d,p) level of theory.

The vibrational frequencies of the created molecules were calculated using the same level of theory. The harmonic force fields of cellulose I_α and cellulose I_β were assessed using the scaled quantum mechanical force field procedure proposed by Pulay, Fogarasi, *et al.* [25]. By converting force fields in Cartesian coordinates to natural internal coordinates, the MOLVIB program was used to calculate the IR intensities and potential energy distributions (PEDs) of the vibrational modes [26]. The calculations with MOLVIB program are performed in mass-weighted Cartesian coordinates, which eliminates most problems with redundant coordinates. The classical molecular force fields are calculated with this program. Thus, the harmonic force field can be refined by scaling, and the scaling factors are calculated from a least-squares fit of the calculated and observed frequencies [26]. In this study the following scale factors were used:

O-H stretch	0.87
C-H deformation	0.92
C-H stretch	0.91
All others	0.98

The *in silico* molecular docking studies were carried out with the AutoDock-Vina program [27]. Three-dimensional structures of Endoglucanase enzyme (PDB ID: 1EG1), β-glucosidase enzyme (PDB ID: 3AHY), and exoglucanase enzyme (PDB ID: 5YJ6) were taken from the protein data bank [28-30], The initial preparations of target proteins and compounds (ligands) were performed in Auto Dock Tool to dock in the compatible format. The target proteins (Endoglucanase, β-glucosidase, and exoglucanase enzymes) were prepared for molecular docking by removing water molecules and adding polar hydrogens. The Kollman charges of enzymes were also determined and dispersed. The active sites of the enzymes were identified with a grid size of 40 Å x 40 Å x 40 Å. The optimized structures of ligands (cellulose I_α and cellulose I_β) were adapted for docking analysis. The partial charges of cellulose I_α and cellulose I_β molecules were calculated with the Geistenger method.

RESULTS AND DISCUSSION

Geometry optimization

The optimized single-ring structures of cellulose I_α and cellulose I_β are shown in Figure 1 and the optimized geometrical parameters of single ring cellulose- I_α and I_β are given in Table 1. The

differences in geometrical parameters of I_α and cellulose I_β are due to the difference in the symmetrical structures of the OH groups in the ring structure. The optimized 4-ring I_α and cellulose I_β structures are found to have nearly the same energy. However, when detailed calculation results were taken into account, optimized cellulose 4-ring I_β beta structure was found to be 0.000053 kJ/mol lower energy than 4-ring I_α. The slight variation in dihedral angles causes this small energy difference. The conversion of cellulose I_α to I_β on heating was previously reported [31]. Nishiyama *et al.* (2003) investigated the crystal structure and H-bonding interactions of cellulose I_α and discussed discuss possible routes for the solid-state conversion of cellulose I_α to I_β [12].

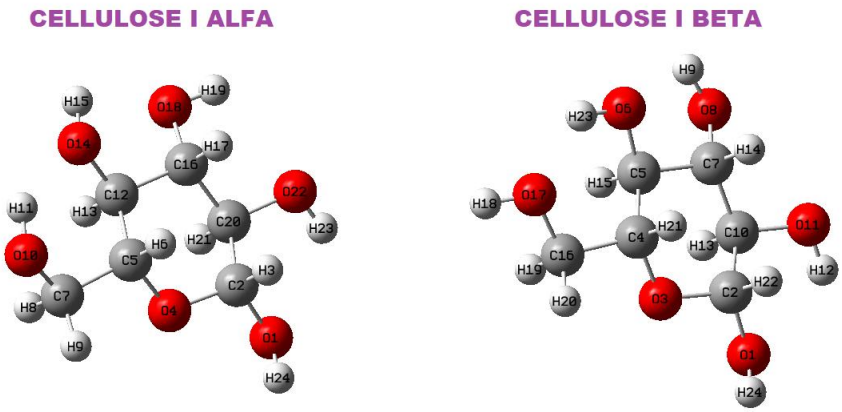

Figure 1. The optimized single-ring (glucose) structures of cellulose Iα and cellulose Iβ molecules (24 atoms) obtained by DFT/B3LYP/6-31G(d,p) level of theory.

The chemical structure of cellulose molecules is governed by primary and secondary alcohol groups and the glycosidic bonds that connect these parts. Their highly resistant glycosidic bonds make it difficult to break them down. Alcohollysis and hydrolysis reactions can dissolve these bonds [32].

Table 1. A single ring of cellulose I_α and I_β structural parameters[a] as determined by DFT/B3LYP/6-31G(d,p) level of theory.

Atoms	I_α	I_β	Atoms	I_α	I_β	Atoms	I_α	I_β
R(1,2)	1.395	1.395	R(16,17)	1.102	1.103	D(24,1,2,3)	57.6	53.9
R(1,24)	0.967	0.967	R(16,18)	1.421	1.415	D(24,1,2,4)	-62.9	-65.9
R(2,3)	1.106	1.106	R(16,20)	1.523	1.523	D(24,1,2,20)	177.5	173.1
R(2,4)	1.421	1.423	R(18,19)	0.968	0.969	D(1,2,4,5)	179.6	-180.0
R(2,20)	1.529	1.533	R(20,21)	1.100	1.101	D(3,2,4,5)	57.6	58.6
R(4,5)	1.432	1.427	R(22,23)	0.967	0.968	D(1,2,20,22)	-65.6	-65.0
R(5,6)	1.102	1.103	A(2,1,24)	108.2	108.3	D(3,2,20,16)	-63.4	-64.1
R(5,7)	1.532	1.527	A(1,2,3)	111.5	111.4	D(4,2,20,21)	-61.6	-61.9
R(5,12)	1.537	1.539	A(1,2,4)	109.0	108.5	D(2,4,5,6)	-57.2	-60.4
R(7,8)	1.103	1.099	A(1,2,20)	107.2	107.1	D(6,5,7,10)	49.4	62.5
R(7,9)	1.094	1.098	A(3,2,20)	109.6	109.1	D(12,5,7,9)	170.3	-180.0
R(7,10)	1.413	1.431	A(4,2,20)	110.4	111.9	D(9,7,10,11)	169.8	-67.7
R(10,11)	0.970	0.965	A(2,4,5)	113.0	112.8	D(13,12,14,15)	-68.9	71.5
R(12,13)	1.102	1.104	A(4,5,6)	110.5	110.9	D(12,16,20,22)	-175.4	-171.4
R(12,14)	1.425	1.420	A(4,5,7)	107.7	104.9	D(17,16,20,2)	64.9	66.5

R(12,16)	1.524	1.533	A(4,5,12)	108.6	110.0	D(18,16,20,21)	-55.9	-55.5
R(14,15)	0.968	0.971	A(6,5,7)	108.2	109.4	D(21,20,22,23)	-60.9	-61.5

[a]R and A are abbreviations for bond (Å), angle (°), respectively.

Frontier molecular orbital analysis

The frontier molecular orbitals (HOMO, LUMO) allow us to predict the reactivity of the molecule. They also give information about the electric and optical properties, as well as in UV–Vis spectrum of the molecule [33, 34]. The energies of the Highest Occupied Molecular Orbital, HOMO and the Lowest Unoccupied Molecular Orbital, LUMO, of the single ring structures of cellulose Iα and cellulose Iβ molecules were calculated using DFT/B3LYP/6-31G(d,p) level of theory. The E_{LUMO}-E_{HOMO} energy difference is most crucial for the prediction of the chemical stability of the molecule [35]. Ionization potential and electron affinity are both directly correlated with HOMO and LUMO energy, respectively. A wide HOMO-LUMO gap denotes a molecule's high stability and low reactivity in chemical reactions [34, 36]. The DFT/B3LYP/6-31G(d,p) theory level was used to calculate the HOMO and LUMO energies of the single-ring structures of the cellulose I alpha and cellulose I beta molecules. The frontier molecular orbitals are displayed in Figure 2.

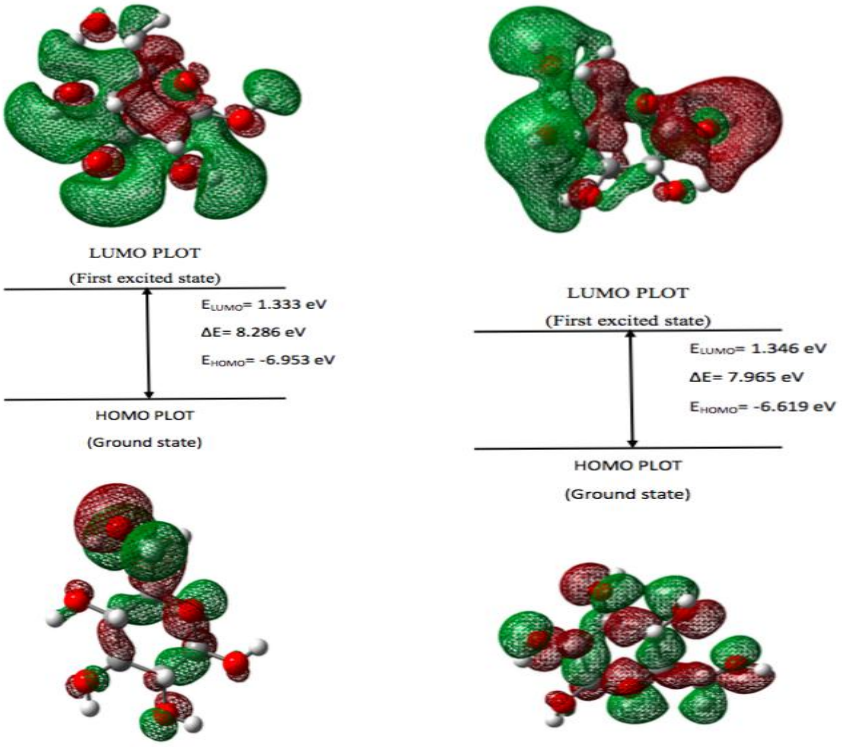

Figure 2. The frontier molecular orbital for cellulose I alfa and cellulose I beta are composed of atomic orbitals with the HOMO-LUMO configuration.

The calculation results of cellulose I_α and cellulose I_β revealed that HOMO energies -6.953 eV and -6.619 eV and LUMO energies 1.333 eV and 1.346 eV, respectively. The energy gap

($E_g = E_{LUMO} - E_{HOMO}$) of the cellulose I_α and cellulose I_β was found to be 8.286 eV and 7.965 eV, respectively. The investigated molecules have a relatively high $E_{HOMO-LUMO}$ value, which suggests that they are highly chemically stable and have a low reactivity [34]. Moreover, the results show that cellulose Iα structure is more stable then cellulose I_β. By using Koopman theory [37] the electron affinity ($A = E_{LUMO}$) the ionization potential ($I = E_{HOMO}$), the electronegativity ($\chi = (I + A) / 2$), hardness ($\eta = (I - A) / 2$), and softness ($\sigma = 1/ \eta$) parameters of the investigated molecules were calculated [38] and the results are shown in Table 2.

The Koopman theory [38] states that the electron affinity (A) is equal to E_{LUMO} and the ionization potential (I) is equal to E_{HOMO}. From the HOMO-LUMO energies of neutral molecules, the values of the molecule's electronegativity ($\chi = (I + A) / 2$), hardness ($\eta = (I - A) / 2$), and softness ($\sigma = 1/ \eta$) were calculated [37]; the results are shown in Table 2.

Table 2. Calculated electrochemical parameters of cellulose I_α and I_β by DFT/B3LYP/6-31G(d,p) level of theory.

	Ionization potential (I)	Electron affinity (A)	Electronegativity (χ)	Hardness (η)	Softness (σ)
Cellulose I_α	6.953	-1.333	2.81	4.143	0.24
Cellulose I_β	6.619	-1.346	2.6365	3.9825	0.251

Molecular electrostatic potential

Molecular electrostatic potential (MEP) is a visual method used to determine the active sites and the relative polarity of the molecule. The MEP map shows the potential in red and blue, with red representing the most negative potential. The most electrophilic attack is most likely to occur in a region where MEP of a molecule is most negative, whereas the most nucleophilic attack is most likely to occur in a region where it is most positive [39-43]. The molecular electrostatic potential (MEP) plots of cellulose I_α and I_β molecules are shown in Figures 3a and 3e. In the MEP plot the different values of the electrostatic potential on the surface are represented by different colors and the potential increases in the order of red < orange < yellow < green < blue. The color code of cellulose I_α is in the range between − 0.06125 a.u. (deepest red) and 0.06125 a.u. (deepest blue), while that of cellulose I_β is − 0.07502 and 0.07502 a.u. range. As seen in Figures 3a, and 3e the regions with negative potential value (red regions) are localized on the oxygen atoms, that are the possible electrophilic attack regions, and the positive potential regions (blue regions) are found to close to the hydrogen atoms of the OH groups, that are the possible nucleophilic attack sites.

Molecular electrostatic potential (ESP) is a valuable concept in molecular modeling calculations as it offers information on the active sites of molecules [44]. The ESP explains the distribution of charge in molecule with the visualization of variable charged regions, so that charge distribution gives wide information that how a molecule interacts with another molecule. Furthermore, it is crucial in assessing electrophilic and nucleophilic regions. In Figure 3, electrostatic potential (ESP) array plot, ESP iso-surface contour map and total density array plot of cellulose I_α, and cellulose I_β molecules, respectively, were given.

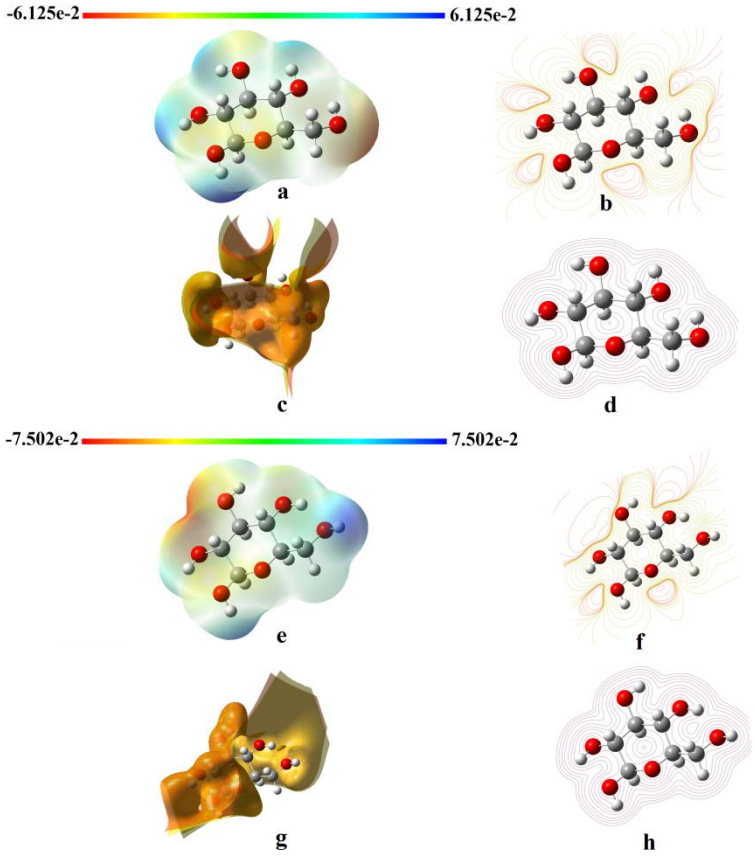

Figure 3. Molecular electrostatic potential (MEP) surface map plot (a,e) electrostatic potential (ESP) array plot (b,f), ESP iso-surface contour map (c,g) and total density array plot (d,h) of cellulose I_α and I_β, calculated using DFT/B3LYP/6-31G(d,p) method, respectively.

Vibrational spectral analysis

The calculated scaled wavenumbers of the 1-ring, 2-ring, 3-ring, and 4-ring structures of cellulose I_α and I_β are given in Tables 3 and 4, respectively, as compared with the ATR-FTIR wavenumbers of cellulose I [45] and Raman wavenumbers of cotton cellulose [46]. The calculated potential energy distribution of the vibrational modes of the 1-ring structure (PED %) of the cellulose was also included in the tables.

The OH stretching vibrations are generally observed in the region 3200-3700 cm^{-1}. The O-H stretching vibrations of the 1-ring cellulose Iα and Iβ molecules were calculated at 3553 cm^{-1}, 2548 $cm^{-1}$and 3496 cm^{-1}, and 3581 cm^{-1}, 3547 cm^{-1} and 3533 cm^{-1}, respectively as pure modes (PED %, 99-100). The OH stretching wavenumber was observed at 3352 cm^{-1} in the FTIR-ATR spectrum of cellulose-I [45]. On going from a one-ring structure to a 2, 3 or 4-ring structure, the frequency alterations occurred, due to increasing or decreasing intra molecular interactions by increasing the ring units.

The CH stretching vibrations of one ring-Cellulose I alpha and I beta were calculated in the regions 2954-2810 cm^{-1} and 2920-2812 cm^{-1}, respectively. Their PED contributions were 97-92 % and 99-89 %. The corresponding vibrations were observed at 2901 cm^{-1} in the FTIR-ATR spectrum of cellulose I, and observed at 2948, 2894, and 2870 cm^{-1} in the Raman spectrum of cotton cellulose [46].

The HCH bending (CH_2 scissoring) vibration was observed at 1478 cm^{-1} in Raman spectrum of cotton cellulose [46]. The HCH bending modes of the cellulose I alpha molecule for 1-4 ring structures were calculated in the 1499-1486 cm^{-1} region, and for 1-4 ring structures of I beta molecule in the 1500-1488 cm^{-1} region.

The COH bending modes were observed at 1431 cm^{-1}, 1373, 1337, and 1319 cm^{-1} in the FTIR-ATR spectrum of cellulose I [45] and 1462, 1410, 1380, 1338 cm^{-1}, in the Raman spectrum of cotton cellulose [46]. These vibrations were calculated in the 1423-1413 cm^{-1}, 1337-1319 cm^{-1}, and 1317-1222 cm^{-1} wavenumber ranges for cellulose I alpha, with PED contributions of 29-22%, 28-27%, and 51-49%, respectively and in the 1476-1411 cm^{-1}, 1389-1236 cm^{-1} and 1249-1203 cm^{-1} wavenumber ranges for cellulose I beta, with PED contributions of 38-27%, 50-16%, and 28%, respectively.

The CO stretching mode, mixing with other vibrations, calculated at 1199 cm^{-1} (36% PED) and 1204 cm^{-1} (17% PED) for 1-ring cellulose Iα, and Iβ, respectively, was observed at 1202 cm^{-1} in the FTIR-ATR spectrum of cellulose 1. The other computed CO stretching vibrations for 1-ring cellulose I_α were found at 1145 (59% PED), 1134 (45% PED), 1111 (53% PED), 1100 (62% PED), 1085 (55% PED), 1077 (43% PED) and 1072 cm^{-1} (55% PED), as mixed modes. The corresponding vibrations for 1-ring cellulose I_β were computed at 1168 (74% PED), 1144 (70% PED), 1137 (39% PED), 1112 (50% PED), 1102 (51% PED), 1099 (46% PED), 1091 (53% PED), 1071 (55% PED) and 1059 cm^{-1} (56% PED).

The 897 cm^{-1} and 900 cm^{-1} bands observed in the FTIR-ATR spectrum of cellulose I, and the Raman spectrum of cotton cellulose assigned as CCH bending mode [45, 46]. This mode was calculated as 942-690 cm^{-1} for the cellulose I_α molecule and 999-962 cm^{-1} for the cellulose I_β molecule.

Table 3. Scaled wavenumbers (cm-1) and the 1-ring PED (%) potential energy distribution are used to compare the 1-ring, 2-ring, 3-ring, and 4-ring structures of the cellulose I alpha molecule.

Assign.	FTIR	Raman	$\nu_{cal}^{(S)}$ Cellulose I alpha 6- 31G(d,p)				
	[45]	[46]	$\nu_{cal}^{(S)}$ 1 ring	$\nu_{cal}^{(S)}$ 2 ring	$\nu_{cal}^{(S)}$ 3 ring	$\nu_{cal}^{(S)}$ 4 ring	PED (≥ 5%)
ν_{OH}			3553	3548: 3363	3546: 3544: 3396	3546: 3396: 3393: 3360	ν_{OH}(99)
ν_{OH}			3548	3545: 3443	3444: 3442: 3361	3496: 3464: 3461: 3460	ν_{OH}(99)
ν_{OH}			3496	3568: 3497	3570: 3565: 3496	3570: 3567: 3565: 3543	ν_{OH}(100)
ν_{CH}		2948	2954	2950: 2933	2950: 2934: 2929	2950: 2935: 2929: 2929	ν_{CH}(96)
ν_{CH}	2901	2894	2886	2874: 2873	2893: 2893:	2874: 2874:	ν_{CH}(96)

					2883	2871: 2869	
ν_{CH}		2870	2870	2959: 2890	2882: 2874: 2871	2894: 2893: 2893: 2883	$\nu_{CH}(95)$
ν_{CH}			2857	2879: 2861	2872: 2873: 2861	2882: 2879: 2873: 2872	$\nu_{CH}(94)$
ν_{CH}			2852	2860: 2848	2847: 2846: 2860	2860: 2849: 2847: 2846	$\nu_{CH}(92)$
ν_{CH}			2843	2879: 2844	2878: 2849: 2842	2879: 2861: 2843: 2846	$\nu_{CH}(93)$
ν_{CH}			2810	2842: 2815	2841: 2840: 2813	2842: 2840: 2838: 2814	$\nu_{CH}(97)$
δ_{HCH}		1478	1492	1499: 1491	1493: 1489: 1487	1495: 1490: 1489: 1486	$\delta_{HCH}(33) + \Gamma_{HCOH}(13) + \Gamma_{HCCO}(12) + \Gamma_{HCCC}(6)$
δ_{COH}	1431	1462	1423	1479: 1429	1474: 1468: 1426	1475: 1472: 1466: 1429	$\delta_{COH}(29) + \nu_{CC}(14)$
δ_{COH}		1410	1407	1456: 1435	1455: 1423: 1413	1455: 1424: 1423: 1413	$\delta_{COH}(22) + \nu_{CC}(16) + \delta_{HCO}(6)\ ) + \nu_{CO}(6)\ ) + \Gamma_{HCCO}(5)$
ν_{CO}, δ_{HCO}			1397	1415: 1395	1418: 1389: 1370	1418: 1408: 1388: 1370	$\nu_{CO}(13) + \delta_{HCO}(7) + \Gamma_{HCCO}(8) + \Gamma_{HCCC}(7) + \Gamma_{HCOC}(7) + \delta_{COH}(5)$
δ_{COH}	1373	1380	1374	1429: 1422	1410: 1389: 1365	1410: 1387: 1365: 1356	$\delta_{COH}(7)$
ν_{CC}			1363	1381: 1368	1385: 1383: 1376	1383: 1377: 1374: 1371	$\nu_{CC}(24) + \delta_{HCC}(19) + \Gamma_{COCH}(7) + \delta_{HCO}(7)$
δ_{COH}	1337	1338	1339	1364: 1361	1360: 1357: 1336	1369: 1361: 1339: 1334	$\delta_{COH}(28) + \delta_{HCC}(7) + \Gamma_{HCCO}(6) + \Gamma_{HCCC}(6)$
δ_{COH}	1319		1326	1340: 1333	1335: 1331: 1323	1352: 1327: 1320: 1319	$\delta_{COH}(27) + \delta_{HCO}(6) + \delta_{CCH}(5) + \Gamma_{HCCH}(6) + \Gamma_{HCCO}(5)$
δ_{HCO}		1294	1299	1331: 1311	1318: 1278:	1332: 1293:	$\delta_{HCO}(18) + \Gamma_{HCCH}(17) + \Gamma_{HCCC}(9) + \Gamma_{HCCO}(8)$

					1273	1287: 1271	
δ_{COH}	1282		1284	1317: 1296	1311: 1304: 1300	1309: 1307: 1302: 1299	$\delta_{COH}(49) + \nu_{CC}(15)$
δ_{COH}	1236		1222	1248: 1238	1292: 1289: 1269	1297: 1287: 1280: 1272	$\delta_{CCH}(10) + \delta_{COH}(51) + \Gamma_{HCCH}(6)$
ν_{CO}	1202		1199	1222: 1210	1236: 1224: 1209	1268: 1266: 1265: 1249	$\nu_{CO}(36) + \nu_{CC}(9) + \delta_{COH}(7) + \delta_{CCH}(7)$
ν_{CO}			1189	1201: 1198	1199: 1196: 1193	1199: 1197: 1194: 1191	$\nu_{CO}(34) + \nu_{CC}(5) + \delta_{COH}(6) + \delta_{CCH}(6) + \delta_{HCO}(11)$
ν_{CO}	1165	1153	1145	1151: 1148	1149: 1147: 1143	1151: 1147: 1144: 1142	$\nu_{CO}(59) + \nu_{CC}(16)$
ν_{CO}			1134	1139: 1131	1136: 1135: 1132	1135: 1134: 1132: 1132	$\nu_{CO}(45) + \nu_{CC}(18)$
ν_{CO}		1122	1111	1127: 1123	1125: 1122: 1119	1126: 1123: 1120: 1117	$\nu_{CO}(53) + \nu_{CC}(20)$
ν_{CO}	1114	1097	1100	1112: 1106	1114: 1111: 1105	1116: 1111: 1108: 1105	$\nu_{CO}(62) + \nu_{CC}(5)$
ν_{CO}			1085	1094: 1093	1095: 1093: 1090	1095: 1094: 1093: 1089	$\nu_{CO}(55) + \delta_{COH}(11)$
ν_{CO}			1077	1092: 1083	1085: 1083: 1081	1084: 1082: 1080: 1077	$\nu_{CO}(43) + \nu_{CC}(8)$
ν_{CO}		1072	1072	1073: 1065	1072: 1066: 1063	1070: 1066: 1065: 1062	$\nu_{CO}(55) + \nu_{CC}(5)$
ν_{CC}	1058	1059	1046	1056: 1053	1055: 1050: 1048	1055: 1050: 1049: 1048	$\nu_{CC}(23) + \nu_{CO}(22) + \delta_{COH}(10)$
ν_{CO}	1032	1038	1040	1043: 1020	1045: 1032: 1016	1043: 1032: 1026: 1016	$\nu_{CO}(32) + \delta_{CCH}(11)$
ν_{CC}			1010	1014: 1006	1015: 1007:	1017: 1008:	$\nu_{CO}(19) + \nu_{CC}(32)$

					1002	1003: 1000	
ν_{CO}	983	999	995	994: 972	999: 987: 980	999: 992: 982: 979	$\nu_{CO}(27) + \nu_{CC}(19)$
ν_{CO}		972	962	965: 940	969: 962: 905	972: 961: 906: 904	$\nu_{CO}(15) + \nu_{CC}(8) + \delta_{CCH}(13) + \delta_{COH}(6)$
δ_{CCH}	897	900	942	894: 868	903: 867: 692	903: 869: 698: 690	$\delta_{CCH}(19) + \delta_{HCO}(12) + \Gamma_{HCOC}(10) + \nu_{CO}(9) + \Gamma_{HOCH}(7)$
ν_{CO}	713		673	643: 630	711: 626	714: 657: 619	$\nu_{CO}(21) + \delta_{OCO}(8) + \delta_{CCO}(8)$
δ_{CCO}			585	605: 589	607: 502: 467	602: 590: 544: 469	$\delta_{CCO}(20) + \delta_{CCC}(6)$
ν_{CO}			577	577: 555	641: 597: 587	642: 607: 511: 492	$\nu_{CO}(9) + \nu_{CC}(6) + \delta_{COC}(7) + \delta_{OCC}(6)$
Γ_{HCOH}		542	540	539: 526	573: 571: 564	587: 571: 568: 565	$\Gamma_{HCOH}(33) + \Gamma_{CCOH}(11)$
Γ_{HCOH}		520	524	498: 475	561: 548: 527	563: 559: 555: 546	$\Gamma_{HCOH}(64) + \Gamma_{CCOH}(29)$
ν_{CC}		460	466	450: 435	454: 451: 417	453: 418: 412:	$\nu_{CO}(5) + \nu_{CC}(19)$
δ_{CCO}		437	435	421: 410	435: 424: 411	435: 426: 422:	$\nu_{CC}(8) + \delta_{CCO}(18)$) + $\delta_{CCC}(5)$
Γ_{CCOH}			405	405: 393	406: 395: 394	447: 405: 394: 393	$\Gamma_{CCOH}(45) + \Gamma_{HCOH}(14)$
Γ_{CCOH}			400	388: 375	387: 374: 359	389: 383: 370: 359	$\Gamma_{CCOH}(38) + \Gamma_{HCOH}(14)$
Γ_{CCOH}		380	379	349: 321	345: 341: 334	355: 343: 341: 336	$\Gamma_{CCOH}(54) + \Gamma_{HCOH}(16)$
ν_{CC} , δ_{CCO}		333	329	340: 308	334: 324: 316	333: 326: 317: 311	$\nu_{CC}(15) + \delta_{CCO}(15)$
δ_{CCO}			311	307:	304:	302:	$\delta_{CCO}(27) + \nu_{CO}(7) +$

				287	300: 296	298: 293: 291	ν_{CC}(10)
Γ_{CCOH}			287	282: 261	287: 281: 270	285: 280: 278: 259	Γ_{CCOH}(24)
Γ_{CCOH}		252	272	252: 232	258: 248: 229	254: 242: 230: 227	Γ_{CCOH}(51) + Γ_{HCOH}(6)
Γ_{CCOH}			237	216: 198	223: 214: 205	217: 210:	Γ_{CCOH}(61) + Γ_{HCOH}(10)
Γ_{CCOH}			210	187: 165	190: 173: 160	202: 193: 185: 168	Γ_{CCOH}(36) + Γ_{HCOH}(6)
Γ_{CCOH}			200	151: 149	152: 150: 143	158: 153: 153: 148	Γ_{CCOH}(53) + Γ_{HCOH}(9)
$\Gamma_{CCCO,}$ Γ_{OCOC}		172	171	122: 115	135: 122: 116	133: 122: 99: 76	Γ_{CCCO}(6) + Γ_{OCOC}(6) + Γ_{OCCO}(5) + Γ_{CCCH}(5)
Γ_{OCCO}			122	95: 78	98: 95: 82	144: 136: 114: 91	Γ_{OCCO}(4)
Γ_{OCOC}			111	71: 53	76: 71: 58	57: 32	Γ_{OCOC}(6)
Γ_{COCC}			71	33: 21	40: 31: 28	83: 75: 24	Γ_{COCC}(5)
ν_{OH}			12	12	12	19: 12: 12: 12	ν_{OH}(28) + Γ_{HOCO}(20) + δ_{COH}(18) + Γ_{HOCC}(16) + Γ_{HOCH}(14)
Γ_{HOCO}			11	11	11	11	ν_{OH}(5) + Γ_{HOCO}(33) + δ_{COH}(9) + Γ_{HOCC}(25) + Γ_{HOCH}(22)
δ_{COH}			11	11	11	11	ν_{OH}(28) + Γ_{CCOH}(18) + δ_{COH}(34) + Γ_{HCOH}(6)

Table 4. Scaled wave numbers (cm^{-1}) and the 1-ring PED (%) potential energy distribution are used to compare the 1-ring, 2-ring, 3-ring, and 4-ring structures of the cellulose I beta molecule.

Assign.	FTIR	Raman	$\nu_{cal}^{(S)}$ Cellulose I beta 6- 31G(d,p)				
	[45]	[46]	$\nu_{cal}^{(S)}$ 1 ring	$\nu_{cal}^{(S)}$ 2 ring	$\nu_{cal}^{(S)}$ 3 ring	$\nu_{cal}^{(S)}$ 4 ring	PED (≥ 5%)
ν_{OH}			3581	3581- 3564	3581- 3566- 3565	3570- 3567- 3565-	ν_{OH}(100)

						3546	
ν_{OH}			3547	3546-3387	3546-3398-3396	3543-3395-3392-3359	$\nu_{OH}(100)$
ν_{OH}	3352		3533	3532-3434	3533-3441-3435	3496-3446-3443-3441	$\nu_{OH}(100)$
ν_{CH}		2948	2920	2927-2920	2929-2928-2920	2950-2935-2929-2929	$\nu_{CH}(98)$
ν_{CH}	2901	2894	2878	2877-2877	2894-2879-2878	2894-2893-2882-2879	$\nu_{CH}(89)$
ν_{CH}		2870	2876	2874-2873	2877-2874-2874	2883-2874-2874-2873	$\nu_{CH}(99)$
ν_{CH}			2857	2868-2862	2872-2870-2868	2879-2872-2871-2870	$\nu_{CH}(97)$
ν_{CH}			2843	2846-2834	2861-2843-2835	2861-2860-2849-2843	$\nu_{CH}(95)$
ν_{CH}			2835	2891-2843	2892-2846-2845	2893-2847-2846-2846	$\nu_{CH}(97)$
ν_{CH}			2813	2846-2812	2845-2839-2813	2842-2840-2838-2813	$\nu_{CH}(96)$
δ_{HCH}		1478	1500	1498-1491	1499-1493-1491	1492-1489-1488-1488	$\delta_{HCH}(32) + \delta_{OCH}(14) + \Gamma_{HCCO}(12) + \Gamma_{HCOH}(6)$
δ_{COH}		1462	1456	1471-1468	1475-1469-1467	1476-1473-1465-1453	$\delta_{COH}(38) + \nu_{CC}(22)$
ν_{CC}	1431		1432	1435-1435	1441-1437-1434	1442-1440-1437-1436	$\nu_{CC}(23) + \delta_{COH}(20) + \delta_{CCH}(18) + \delta_{HCO}(12)$
δ_{COH}			1403	1419-1411	1430-1418-1413	1429-1425-1424-1413	$\delta_{COH}(27) + \nu_{CO}(11) + \delta_{HCO}(9) + \Gamma_{HCCO}(6)$
ν_{CO}			1389	1406-1389	1410-1405-1389	1418-1411-1409-	$\nu_{CO}(9) + \nu_{CC}(6) + \delta_{COH}(6)$

						1385	
δ_{COH}	1373		1374	1387-1333	1389-1358-1335	1388-1372-1361-1359	$\delta_{COH}(45) + \Gamma_{HCCH}(6)$
ν_{CC}			1346	1375-1368	1379-1374-1372	1383-1379-1375-1372	$\nu_{CC}(15) + \delta_{COH}(11)$
δ_{COH}	1337	1338	1337	1343-1318	1343-1335-1324	1369-1356-1355-1337	$\delta_{COH}(16) + \delta_{CCH}(11) + \Gamma_{HCCO}(7) + \Gamma_{HCCC}(6) + \Gamma_{COCH}(6)$
δ_{COH}	1319		1308	1313-1310	1321-1315-1310	1336-1335-1328-1321	$\delta_{COH}(19) + \delta_{HCO}(7) + \Gamma_{HCCH}(18)$
δ_{COH}		1294	1298	1301-1294	1306-1304-1297	1320-1310-1307-1303	$\delta_{COH}(29) + \Gamma_{HCCH}(7) + \delta_{HCO}(5)$
δ_{COH}	1282		1279	1288-1269	1293-1289-1274	1300-1297-1293-1286	$\delta_{COH}(42) + \delta_{CCH}(5)$
δ_{COH}	1236		1236	1264-1237	1270-1267-1266	1280-1275-1272-1269	$\delta_{COH}(50) + \delta_{CCH}(9) + \nu_{CC}(8)$
δ_{HCO}			1209	1235-1221	1241-1236-1224	1265-1224-1217-1206	$\delta_{HCO}(36) + \delta_{CCH}(19) + \delta_{COH}(6)$
δ_{COH}	1202		1204	1212-1203	1219-1212-1204	1249-1242-1235-1220	$\delta_{COH}(28) + \nu_{CO}(17)$
ν_{CO}	1165		1168	1199-1184	1199-1195-1187	1198-1196-1194-1192	$\nu_{CO}(74)$
ν_{CO}		1153	1144	1167-1149	1179-1167-1148	1187-1184-1178-1167	$\nu_{CO}(70) + \nu_{CC}(6)$
ν_{CO}			1137	1142-1135	1143-1142-1134	1150-1146-1144-1142	$\nu_{CO}(39) + \nu_{CC}(20)$
ν_{CO}		1122	1112	1131-1122	1133-1131-1120	1136-1134-1133-1132	$\nu_{CO}(50) + \nu_{CC}(11) + \delta_{COH}(6)$
ν_{CO}	1114		1102	1113-1111	1119-1114-1113	1126-1123-1120-	$\nu_{CO}(51) + \nu_{CC}(22) + \delta_{COH}(6)$

						1117	
ν_{CO}		1097	1099	1106-1098	1111-1105-1098	1116-1111-1108-1105	$\nu_{CO}(46) + \nu_{CC}(9)$
ν_{CO}		1072	1091	1094-1081	1095-1088-1084	1095-1094-1091-1090	$\nu_{CO}(53) + \nu_{CC}(5)$
ν_{CO}			1071	1074-1070	1079-1071-1069	1084-1082-1079-1077	$\nu_{CO}(55)$
ν_{CO}	1058	1059	1059	1067-1061	1066-1066-1062	1069-1066-1065-1062	$\nu_{CO}(56) + \nu_{CC}(23)$
ν_{CC}	1032	1038	1042	1049-1042	1050-1047-1044	1054-1051-1050-1047	$\nu_{CC}(26) + \nu_{CO}(16) + \delta_{COH}(11)$
ν_{CO}		999	1015	1017-1011	1029-1015-1013	1043-1032-1027-1017	$\nu_{CO}(34) + \nu_{CC}(31)$
ν_{CC}			992	1005-998	1005-1001-988	1016-1008-1005-1003	$\nu_{CC}(31) + \nu_{CO}(18) + \delta_{COH}(6)$
δ_{CCH}	983	972	962	981-971	980-970-962	999-994-983-977	$\delta_{CCH}(24) + \delta_{HCO}(15) + \Gamma_{HCOC}(11) + \Gamma_{HOCH}(9) + \nu_{CO}(13)$
ν_{CO}	897	900	959	962-904	909-905	962-929-906-905	$\nu_{CO}(17) + \nu_{CC}(10) + \delta_{COH}(11)$
ν_{CC}			872	879-701	879-701-672	903-866-714-698	$\nu_{CC}(31) + \Gamma_{HCOH}(11) + \delta_{CCH}(8)$
ν_{CO}	668		675	666-608	675-670-608	691-622	$\nu_{CO}(22) + \delta_{CCO}(7) + \delta_{OCO}(8)$
δ_{CCO}			589	576-543	544-472	607-543-466-453	$\delta_{CCO}(27) + \delta_{CCC}(9)$
ν_{CO} , δ_{COC}			582	644-627	640-628-624	657-642-630-626	$\nu_{CO}(8) + \delta_{COC}(8) + \nu_{CC}(7) + \delta_{CCO}(6)$
Γ_{CCOC}		542	539	591-584			$\Gamma_{CCOC}(5)$
ν_{CC}			473	500-	451	527-	$\nu_{CC}(22) + \delta_{CCO}(6)$

				479		514-510-492	
Γ_{CCOH}		460	439	453-436	436-430-423	679-670-603-591	$\Gamma_{CCOH}(48) + \Gamma_{HCOH}(16)$
Γ_{CCOH}		437	421	430-416	416	587-571-569-565	$\Gamma_{CCOH}(64) + \Gamma_{HCOH}(18)$
Γ_{CCOH}			411	409-397		563-559-555-545	$\Gamma_{CCOH}(35) + \Gamma_{HCOH}(12) + \delta_{CCO}(7)$
Γ_{CCOH}			393	392-380	393-387-376	394-393-389-384	$\Gamma_{CCOH}(59) + \Gamma_{HCOH}(20)$
Γ_{CCOH}		380	384	359-343	359-357-328	370-359-356-343	$\Gamma_{CCOH}(9) + \Gamma_{HCOH}(6)$
Γ_{HCOH}		333	337	326-315	324-312-302	341-337-333-326	$\Gamma_{HCOH}(29) + \Gamma_{CCOH}(10) + \nu_{CC}(8) + \delta_{CCO}(6)$
δ_{CCO}			321	313-293		291-259	$\delta_{CCO}(25) + \nu_{CC}(17) + \nu_{CO}(7)$
Γ_{OCCO}			286	289			$\Gamma_{OCCO}(6)$
Γ_{HCOH}			278	272-271	287-277-271	295-288-280-278	$\Gamma_{HCOH}(73) + \Gamma_{CCOH}(24)$
δ_{CCO}			265	258-239	265-257-232		$\delta_{CCO}(8) + \Gamma_{CCOH}(5) + \Gamma_{OCOC}(5)$
Γ_{CCOH}		252	252	206-152	224-194		$\Gamma_{CCOH}(15) + \delta_{CCO}(6) + \Gamma_{OCOC}(6)$
δ_{CCC}			205	190	207		$\delta_{CCC}(11) + \Gamma_{COCC}(9) + \delta_{CCO}(5)$
Γ_{OCCO}, Γ_{COCC}		172	178	163-109			$\Gamma_{OCCO}(7) + \Gamma_{COCC}(7) + \Gamma_{CCCO}(6) + \Gamma_{OCOC}(6) + \Gamma_{CCCH}(6) + \Gamma_{OCCH}(6)$
Γ_{OCCO}			120	143-118		153-144-137-122	$\Gamma_{OCCO}(6)$
Γ_{CCOH}			102	97-52			$\Gamma_{CCOH}(5)$
Γ_{CCOC}			77	80-21	83-79	80-46-19	$\Gamma_{CCOC}(5)$
Γ_{CCOH}			13	71-	31-	71-	$\Gamma_{CCOH}(40) + \delta_{COH}(34) +$

				13	13-13	57-25-24	$\Gamma_{HCOH}(13) + \nu_{OH}(6)$
ν_{OH}			12	32-12	58-40	12-12-12	$\nu_{OH}(35) + \delta_{COH}(21) + \Gamma_{HOCO}(17) + \Gamma_{HOCC}(14) + \Gamma_{HOCH}(12)$
Γ_{HOC0}			11	11	11	11	$\Gamma_{HOCO}(32) + \Gamma_{HOCC}(25) + \Gamma_{HOCH}(22) + \delta_{COH}(10) + \nu_{OH}(6)$
ν_{OH}			11	11	11	11	$\nu_{OH}(67) + \delta_{COH}(19)$
Γ_{CCOH}			4	4	4	4	$\Gamma_{CCOH}(76) + \Gamma_{HCOH}(24)$
Γ_{HOCO}			3	3	3	3	$\Gamma_{HOCO}(41) + \Gamma_{HOCC}(31) + \Gamma_{HOCH}(28)$

Molecular docking

Molecular docking is a common method used to study the interaction of a biomolecule with an enzyme or DNA, thus has importance for drug discovery. Since its first appearance in the mid-1970s, docking has become a heavily used method in drug discovery and understanding how chemical compounds interact with their biomolecular targets [47-50]. Cellulose is degraded by cellulose enzymes called "cellulases" that hydrolyze β-1,4-glycoside bonds. The breakdown of cellulose is important in making an important component of plants to ready for consumption and use in chemical reactions. Cellulose enzymes are endoglucanase, exoglucanase, and β-glucosidase enzymes that work together for degradation of cellulose. Therefore, we aimed to reveal the interactions of I_α and cellulose I_β molecules with cellulose enzymes through molecular docking studies. Molecular docking studies of I_α and cellulose I_β were performed against Endoglucanase (PDB ID: 1EG1), β-glucosidase (PDB ID: 3AHY), and exoglucanase (PDB ID: 5YJ6) enzymes.

Figures 4-5 show 2D and 3D depictions of the interactions of the cellulose I_α and cellulose I_β ligands in the target receptors. Table 5 lists the binding affinities of the cellulose I_α and cellulose I_β ligands docked in the active side of cellulose enzymes, as well as the various residues implicated in the binding. The interactions between cellulose I_α and the Endoglucanase enzyme are shown in Figure 4a-b. Seven conventional hydrogen bonding interactions and an unfavorable donor-donor interaction between cellulose I_α and Endoglucanase enzyme residues were revealed (see Figure 4a-b and Table 5). Figure 4c-d shows the docking results of cellulose I_α into the β-glucosidase enzyme. Cellulose I_α forms four conventional hydrogen bonding interactions and a carbon-hydrogen bonding interaction with β-glucosidase enzyme. The interactions between Cellulose I_α and the exoglucanase enzyme are shown in Figure 4e-f. As seen in Figure 4e-f, cellulose I_α forms four conventional hydrogen bonding interactions and an unfavorable acceptor-acceptor interaction with exoglucanase enzyme.

Figure 5 shows the docking results of cellulose I_β into Endoglucanase, enzyme β-glucosidase and exoglucanase enzymes, respectively. Cellulose I_β forms six convertional hydrogen bonding interactions and unfavorable donor-donor and pi-sigma interactions with Endoglucanase residues (see Figure 5a-b and Table 5). Figure 5c-d shows the interactions between Cellulose I_β and the enzyme β-glucosidase. Cellulose I_β forms three conventional hydrogen bonding interaction with ALA344, PRO342. The 2D and 3D depictions of the interactions of the I_β with the exoglucanase enzyme are shown in Figure e-f. Cellulose I_β forms four conventional hydrogen bonding interactions and an unfavorable donor-donor interaction with exoglucanase enzyme (see Table 5).

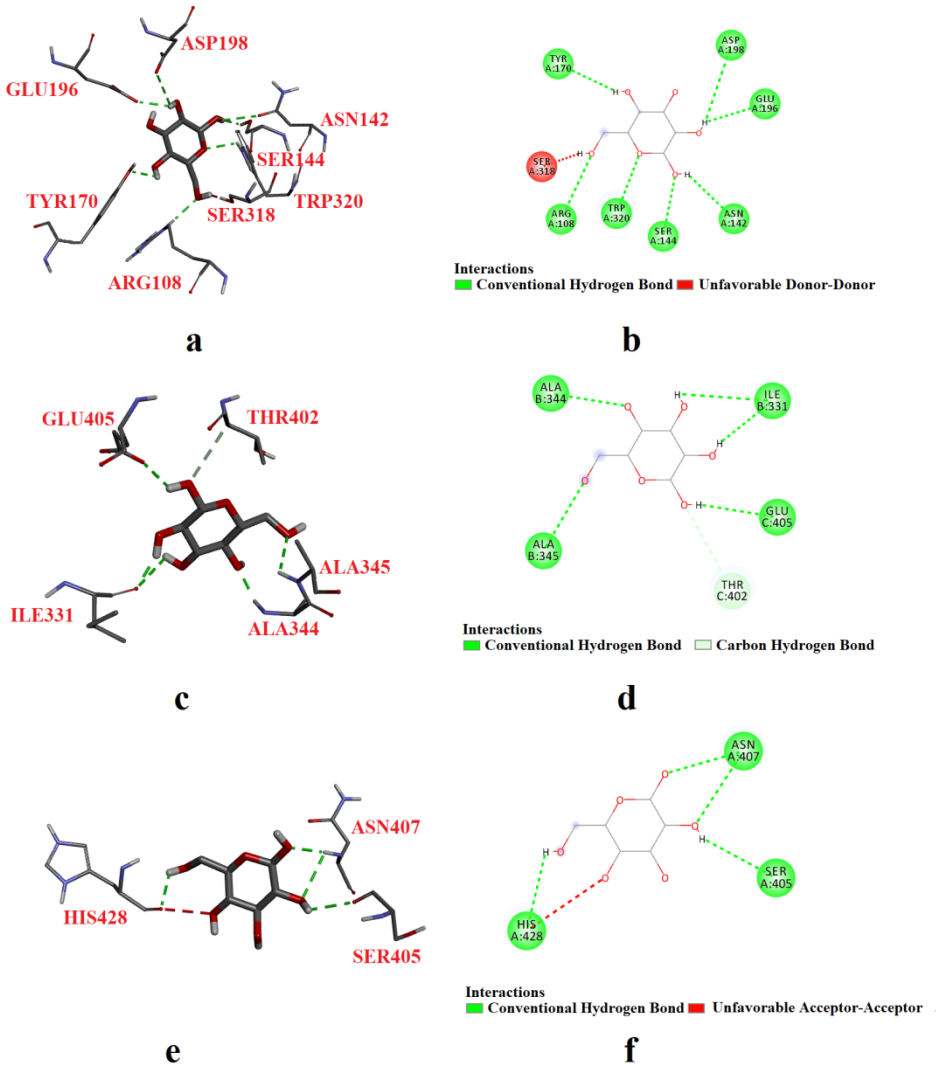

Figure 4. 3D (a,c,e) and 2D (b,d,f) visualizations of the interactions of cellulose I_α within the active side of Endoglucanase enzyme (binding affinity -6.4 kcal/mol), enzyme β-glucosidase (binding affinity -5.3 kcal/mol) and Exoglucanase enzyme (binding affinity -6.1 kcal/mol), respectively.

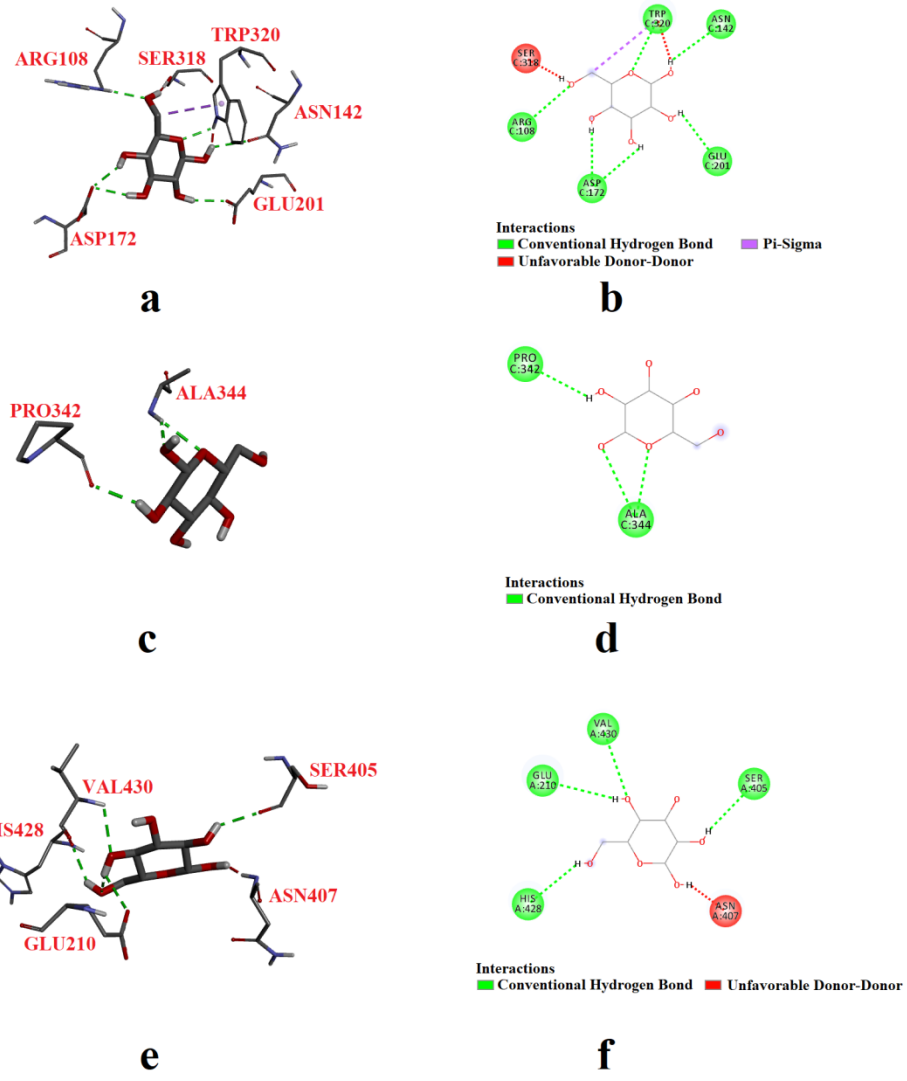

Figure 5. 3D (a,c,e) and 2D (b,d,f) visualizations of the interactions of cellulose I_β within the active side of Endoglucanase enzyme (binding affinity -6.3 kcal/mol), β-glucosidase enzyme (binding affinity -5.2 kcal/mol) and Exoglucanase enzyme (binding affinity -6.2 kcal/mol), respectively.

Table 5. The binding affinities (kcal/mol), bonded residues and bond lengths (Å) of cellulose I_α and cellulose I_β, docked with target proteins*.

		Cellulose I_α		Cellulose I_β
Targets (PDB ID)	Binding affinity	Interactions (Å)	Binding affinity	Interactions (Å)
1EG1	-6.4	H-bonding: ARG 108 (2.27)	-6.3	H-bonding: ARG108 (2.3)

		ASN142 (2.54) SER144 (2.2) TYR170 (2.39) GLU196 (2.25) ASP198 (3.01) TRP320 (2.26) Unfavorable donor-donor: SER318 (1.21)		ASN142 (2.3) ASP172 (2.14; 2.29) GLU201 (2.2); TRP320 (2.26) Unfavorable donor-donor: SER318 (1.25) TRP320 (1.32) Pi-Sigma: TRP320 (3.93)
3AHY	-5.3	H-bonding: ILE331 (1.89; 2.91) GLU405 (2.39) ALA345 (2.67) ALA344 (1.94) Carbon-H bond: THR402 (3.45).	-5.2	H-bonding: ALA344 (2.26; 2.29) PRO342 (2.12)
5YJ6	-6.1	H-bonding: ASN407 (2.12; 2.79) SER405 (2.39) HIS428 (2.26) Unfavorable acceptor-acceptor: HIS428 (2.9).	-6.2	H-bonding: GLU210 (2.75) SER405 (2.21) HIS428 (2.22) VAL430 (2.58) Unfavorable donor-donor: ASN407 (1.2)

*Bond lengths are given in patanthesis in Å. The interactions of both ligands with the same target residues are shown in bold.

CONCLUSIONS

Cellulose is degraded by cellulase enzymes called "cellulase" that hydrolyze β-1,4-glycosidic bonds, but the exact mechanism is not wholly understood. The current study performed to reveal the mechanism of cellulose I_α and cellulose I_β binding to the the cellulase enzymes endoglucanase, exoglucanase, and β-glucosidase. The most stable structures of cellulose I_α and I_β molecules were obtained by DFT/B3LYP/6-31G(d,p) level of theory. A complete vibrational analyses of 1-,2-,3-, and 4-ring structures of cellulose I_α and I_β molecules were performed at DFT/B3LYP/6-31G(d,p) level of theory. In addition HOMO and LUMO energies and the molecular electrostatic potential contour plots (MEP and ESP) of one ring structures of cellulose I_α and I_β molecules have been cellulose I_α calculated, to predict the reactive behaviours. The HOMO and LUMO energy values were calculated and the band gap energy of the cellulose I_α and I_β molecules were determined to be 8.286 eV and 7.965 eV in gas phase, respectively. The docking simulations of the cellulose I_α and I_β molecules with the cellulase enzymes were performed for comparatively evaluation of the degrade activity of the and I_β molecules. The molecular docking of the cellulose I_α and I_β with Endoglucanase enzyme, enzyme β-glucosidase, and exoglucanase enzyme revealed the strong interactions between ligand and

targets (-6.4 and -6.3; -5.3 and -5.2; -6.1 and -6.2 kcal/mol, respectively). Docking results indicate that although the binding affinity of cellulose I_α for Endoglucanase and β-glucosidase enzymes are found to be slightly better than cellulose I_β molecule, but for exoglucanase enzyme cellulose I_β is better than I_α.

ACKNOWLEDGEMENTS

This work was supported by the Scientific Research Projects Coordination Unit of Istanbul University [ÖNAP-2423], [FDK-2019-34884].

Disclosure statement

The authors declare that they have no competing interests.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

1. Brown Jr, R.M.; Saxena, I.M. Cellulose biosynthesis: A model for understanding the assembly of biopolymers. *Plant Physiol. Biochem.* **2000**, 38, 57-67.
2. Sorieul, M.; Dickson, A.; Hill, S.J.; Pearson, H. Plant fibre: Molecular structure and biomechanical properties, of a complex living material, influencing its deconstruction towards a biobased composite. *Mater.* **2016**, 9, 618.
3. Ganster, J.; Fink, H.P. *Cellulose and Cellulose Acetate. Bio-Based Plastics: Materials and Applications*, 1st ed., John Wiley & Sons, Ltd., England; **2013**; pp. 35-62.
4. Pauly, M.; Keegstra, K. Plant cell wall polymers as precursors for biofuels. *Curr. Opin. Plant Biol.* **2010**, 13, 304-311.
5. Heinze, T.; Liebert, T. Unconventional methods in cellulose functionalization. *Prog. Polym. Sci.* **2001**, 26, 1689-1762.
6. Pettersen, R.C. The chemical composition of wood. *ACS* **1984**, 207, 57-126.
7. Klemm, D.; Schumann, D.; Udhardt, U.; Marsch, S. Bacterial synthesized cellulose-artificial blood vessels for microsurgery. *Prog. Polym. Sci.* **2001**, 26, 1561-1603.
8. Rao, V.S.R.; Sundararajan, P.R.; Ramakrishnan, C.; Ramachandran, G.N. *Conformational studies of amylose*. In: Ramachandran GN (ed) Conformation of biopolymers, Vol 2. Academic, London; **1967**; p. 721-737.
9. Gardner, K. H.; Blackwell, J. The structure of native cellulose. *Biopolymers* **1974**, 13, 1975-2001.
10. Sugiyama, J.; Okano, T.; Yamamoto, H.; Horii, F. Transformatıon of valonia cellulose crystals by an alkaline hydrothermal treatment macromolecules. *Macromol.* **1990**, 23, 3196-3198.
11. Sugiyama, J.; Vuong, R.; Chanzy, H. Electron-dıffractıon study on the 2 crystalline phases occurrıng in natıve cellulose from an algal cell-wall. *Macromol.* **1991**, 24, 4168-4175.
12. Nishiyama,Y.; Sugiyama, J.; Chanzy, H.; Langan, P. Crystal structure and hydrogen bonding system in cellulose Iα from synchrotron X-ray and neutron fiber diffraction. *J. Am. Chem. Soc.* **2003**, 125, 14300-14306.
13. Atalla, R.H.; Vanderhart, D.L. Native cellulose – a composite of 2 distinct crystalline forms. *Sci.* **1984**, 223, 283-285.

14. Jarvis, M. Cellulose stacks up. *Nature* **2003**, 426, 611-612.
15. Horii, F.; Yamamoto, H.; Kitamaru, R.; Tanahashi, M.; Higuchi, T. Transformation of native cellulose crystals induced by saturated steam at high-temperatures. *Macromol.* **1987**, 20, 2946-2949.
16. Henrissat, B.; Bairoch, A. New families in the classifi- cation of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* **1993**, 293, 781-788.
17. Yoon, L.W.; Ang, T.N.; Ngoh, G.C.; Chua, A.S.M. Fungal solid-state fermentation and various methods of enhancement in cellulase production. *Biomass Bioenerg.* **2014**, 67, 319-338.
18. Dhillon, G.S.; Kaur, S.; Brar, S.K.; Verma, M. Potential of apple pomace as a solid substrate for fungal cellulase and hemicellulase bioproduction through solid-state fermenta- tion. *Ind. Crops Prod.* **2012**, 38, 6-13.
19. Nava-Cruz, N.Y.; Contreras-Esquivel, J.C.; Aguilar-Gonzalez, M.A.; Nuncio, A.; Rodriguez-Herrera, R.; Aguilar, C.N. Agave atrovirens fibers as substrate and support for solid-state fermentation for cellulase production by Trichoderma asperellum. *3 Biotech*. **2016**, 6, 115.
20. Celik, S.; Demirag, A.D.; Ozel, A.E.; Akyuz, S. Molecular structure, molecular docking and absorption, distribution, metabolism, excretion and toxicity study of cellulose II. *J. Chin. Chem. Soc.* **2021**, 68, 1250-1262.
21. Arnold, K.; Bordoli, L.; Kopp, J.; Schwede, T. The SWISS-MODEL workspace: A web-based environment for protein structure homology modelling. *Bioinform.* **2006**, 22, 195-201.
22. Tomberg, A. Gaussian 09w Tutorial an Introduction to Computational Chemistry Using G09w and Avogadro Software, **2020**.
23. Celik, S.; Demirag, A. D.; Ozel, A. E.; Akyuz, S. Molecular structure, vibrational spectra, molecular docking, and ADMET study of cellulose triacetate II. *Opt. Spectrosc.* **2020**, 128, 1138-1150.
24. Nishiyama, Y.; Langan, P.; Chanzy, H. Crystal structure and hydrogenbonding system in cellulose I_β from synchrotron X-ray and neutron fiber diffraction. *J. Am. Chem. Soc.* **2002**, 124, 9074-9082.
25. Pulay, P.; Fogarasi, G.; Pongor, G.; Boggs, J.E.; Vargha, A. Combination of theoretical ab initio and experimental information to obtain reliable harmonic force constants. Scaled quantum mechanical (QM) force fields for glyoxal, acrolein, butadiene, formaldehyde, and ethylene. *J. Am. Chem. Soc.* **1983**, 105, 7037-7047.
26. Sundius, T. Molvib-A flexible program for force field calculations. *J. Mol. Struct.* **1990**, 218, 321-326.
27. Trott, O.; Olson, A.J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2010**, 31, 455-461.
28. Kleywegt, G.J.; Zou, J.Y.; Divne, C.; Davies, G.J.; Sinning, I.; Stahlberg, J.; Reinikainen, T.; Srisodsuk, M.; Teeri, T.T.; Jones, T.A. The crystal structure of the catalytic core domain of endoglucanase I from Trichoderma reesei at 3.6 Å resolution, and a comparison with related enzymes. *J. Mol. Biol.* **1997**, 272, 383-397.
29. Jeng, W.Y.; Wang, N.C.; Lin, M.H.; Lin, C.T.; Liaw, Y.C.; Chang, W.J.; Liu, C.I.; Liang, P. H.; Wang, A.H.J. Structural and functional analysis of three beta-glucosidases from bacterium clostridium cellulovorans, fungus *Trichoderma reesei* and termite *Neotermes koshunensis. J. Struct. Biol.* **2011**, 173, 46-56.
30. Liu, Y.J.; Liu, S.; Dong, S.; Li, R.; Feng, Y.; Cui, Q. Determination of the native features of the exoglucanase Cel48S from *Clostridium thermocellum*. *Biotechnol. Biofuels* **2018**, 11, 6.
31. Wada, M.; Kondo, T.; Okano, T. Thermally induced crystal transformation from cellulose Iα to Iβ. *Polym. J.* **2003**, 35, 155-159.

32. McNamara, G.; Gupta, A.; Reynaert, J.; Coates, T.D.; Boswell, C. Spectral imaging microscopy web sites and data. *Cytometry Part A* **2006**, 69, 863-871.
33. Rocha, M.; Santo, A.D.; Arias, J.M.; Gil, D.M.; Altabef, A.B. Ab-initio and DFT calculations on molecular structure, NBO, HOMO–LUMO study and a new vibrational analysis of 4-(Dimethylamino) Benzaldehyde. *Spectrochim. Acta Part A. Mol. Biomol. Spectrosc*. **2015**, 136, 635-643.
34. Bahgat, K.; Fraihat, S. Normal coordinate analysis, molecular structure, vibrational, electronicspectraand NMR investigation of 4-amino-3-phenyl-1H-1,2,4-triazole-5(4H)-thionebyab initio HF and DFT method. *Spectrochim. Acta Part A. Mol. Biomol. Spectrosc.* **2015**, 135, 1145-1155.
35. Nemetova, U.; Önem, A.N.; Er, A.; Celik, S.; Ozel, A.E.; Akyuz, S.; Ozyurek, M.; Ayla, S.S. A fast and responsive turn-on fluorescent probe based on a quinone conjugated alkoxy derivative for biothiols and a cellular imaging study. *Turk. J. Chem.* **2024**, 48, 830-842.
36. Politzer, P.; Laurence, P.R.; Jayasuriya, K. Molecular electrostatic potentials: An effective tool for the elucidation of biochemical phenomena. *Environ. Health. Persp*. **1985**, 61, 191-202.
37. Parr, R.J.; Szentpaly L.V.; Liu, S. Electrophilicity index. *J. Am. Chem. Soc.* **1999**, 121, 1922-1924.
38. Koopmans, T.A. Ordering of wave functions and eigenenergies to the individual electrons of an atom. *Physica*. **1993**, 1, 104-113.
39. Oktemer, T.S.; Celik, S.; Ozel, A.E.; Akyuz, S.; Er, A. Vibrational spectroscopic characterization, quantum chemical, molecular docking and molecular dynamics investigations of cyclo (L-phenylalanyl-L-proline), an anticancer agent. *Spectrosc. Lett.* **2024**, 58, 1-19.
40. Celik, S.; Ozel, A.E.; Durak, V.; Akyuz, S. Vibrational spectroscopic characterization, quantum chemical and molecular docking studies of Valyl-Methionine dipeptide. *Spectrosc. Lett.* **2020**, 53, 648-663.
41. Kinaytürk, N.Y.; Kalaycı, T.; Tunalı, B. Experimental and computational investigations on the molecular structure, vibrational spectra, electronic properties, and molecular electrostatic potential analysis of phenylenediamine isomers. *Spectrosc. Lett.* **2021**, 54, 693-706.
42. Kinaytürk, N.K.; Önem, E.; Oturak, H. Benzoic acid derivatives: Anti-biofilm activity in Pseudomonas aeruginosa PAO1, quantum chemical calculations by DFT and molecular docking study. *Bull. Chem. Soc. Ethiop*. **2023**, 37, 171-181.
43. Çatal, G.S.; Kaya Kinaytürk, N. Comprehensive analysis of indacaterol and theophylline molecules: Insights from DFT, FTIR, Raman, UV, molecular docking, and ADMET studies. *Spectrosc. Lett.* **2024**, 57, 499-526.
44. Bayoumy, A.M.; Ibrahim, M.; Omar, A. Mapping molecular electrostatic potential (MESP) for fulleropyrrolidine and its derivatives. *Opt. Quantum Electron.* **2020**, 52, 1-13.
45. Oh, S.Y.; Yoo, D.I.; Shin, Y.; Kim, H.C.; Kim, H.Y.; Chung, Y.S.; Youk, J.H. Crystalline structure analysis of cellulose treated with sodium hydroxide and carbon dioxide by means of X-ray diffraction and FTIR spectroscopy. *Carbohydr. Res.* **2005**, 340, 2376-2391.
46. Edwards, H.G.; Nikhassan, N.F.; Farwell, D.W.; Garside, P.; Wyeth, P. Raman spectroscopic analysis of a unique linen artefact: The HMS Victory Trafalgar sail. *J. Raman Spectrosc.* **2006**, 37, 1193-1200.
47. Derafa, W.; Elkanzi, N.A.; Ali, A.M.; Abdou, A. Three Co(II), Ni(II) and Cu(II) Schiff base complexes incorporating 2-[(4-{[(4-methylphenyl) sulfonothioyl] oxy} phenyl) methylene] amino} benzoic acid: Synthesis, structural, DFT, biological and molecular docking investigation. *Bull. Chem. Soc. Ethiop.* **2024**, 38, 325-346.
48. Eğlence-Bakır, S.; Celik, S.; Şahin, M.; Ozel, A. E.; Akyuz, S.; Ülküseven, B. Synthesis, molecular modelling, FT-IR, Raman and NMR characterization, molecular docking and

ADMET study of new nickel(II) complex with an N4-tetradentate thiosemicarbazone. *J. Biomol. Struct. Dyn.* **2021**, 39, 4212-4224.
49. Er, A.; Celik, S.; Ozel, A. E.; Akyuz, S. *Molecular Docking Studies of Cephalosporin and Oxytetracycline with Escherichia Coli DNA Gyrase, In New Approaches in Science and Engineering. Artikel Akademi* **2024**, pp. 95-103.
50. Onem, Z.C.; Er, A.; Celik, S.; Ozel, A. E.; Akyuz, S. The significance of doxylamine, an antihistamine compound, in the context of its potential efficacy against Covid-19: Conformational analysis, molecular docking, and molecular dynamics studies. *Russ. J. Phys. Chem.* **2024**, 99, 1-22.