PCCP

PAPER

View Article Online
View Journal | View Issue

Cite this: *Phys. Chem. Chem. Phys.*, 2017, **19**, 28195

Elastic moduli of biological fibers in a coarse-grained model: crystalline cellulose and β-amyloids†

Adolfo B. Poma, * Mateusz Chwastyk and Marek Cieplak*

We study the mechanical response of cellulose and β-amyloid microfibrils to three types of deformation: tensile, indentational, and shear. The cellulose microfibrils correspond to the allomorphs Iα or Iβ whereas the β-amyloid microfibrils correspond to the polymorphs of either two- or three-fold symmetry. This response can be characterized by three elastic moduli, namely, Y_L, Y_T, and S. We use a structure-based coarse-grained model to analyze the deformations in a unified manner. We find that each of the moduli is almost the same for the two allomorphs of cellulose but Y_L is about 20 times larger than Y_T (140 GPa *vs.* 7 GPa), indicating the existence of significant anisotropy. For cellulose we note that the anisotropy results from the involvement of covalent bonds in stretching. For β-amyloid, the sense of anisotropy is opposite to that of cellulose. In the three-fold symmetry case, Y_L is about half of Y_T (3 *vs.* 7) whereas for two-fold symmetry the anisotropy is much larger (1.6 *vs.* 21 GPa). The S modulus is derived to be 1.2 GPa for three-fold symmetry and one half of it for the other symmetry and 3.0 GPa for cellulose. The values of the moduli reflect deformations in the hydrogen-bond network. Unlike in our theoretical approach, no experiment can measure all three elastic moduli with the same apparatus. However, our theoretical results are consistent with various measured values: typical Y_L for cellulose Iβ ranges from 133 to 155 GPa, Y_T from 2 to 25 GPa, and S from 1.8 to 3.8 GPa. For β-amyloid, the experimental values of S and Y_T are about 0.3 GPa and 3.3 GPa respectively, while the value of Y_L has not been reported.

Received 3rd August 2017,
Accepted 25th September 2017

DOI: 10.1039/c7cp05269c

rsc.li/pccp

I. Introduction

There are many types of biological fibers, such as those made of proteins or those made of carbohydrates. Examples of protein fibers are collagen, actin, spider silk, fibrin, microtubules, elastin, and amyloid systems. Examples of carbohydrate fibers are cellulose, chitin, dextran, and alginate. All of these fibers are made of smaller units, fibrils or microfibrils, and they come with a wide range of elastic properties that relate to their function. Some fibers, like cellulose, are tough and extend little. Others, like fibronectin, are endowed with large extensibility. Fibronectin fibrils can be extended more than 8-fold before they break.[1] Similar extensibility has been reported for spider silk[2,3] and fibrin.[4] The design of novel materials with biological components often requires the knowledge of the response to deformations of various kinds.

The fibers/fibrils are systems that are highly anisotropic and the elastic properties along the fiber axis are expected to be different than along the transverse directions. Experimentally, there are various ways to assess these properties. One may determine the tensile Young's modulus, Y_L (the subscript L stands for longitudinal or axial), by stretching the fibrils. One may also place the fibril on a support plane and press on it by a rounded tip toward the plane to determine the transverse Young's modulus, Y_T. This procedure is called nanoindentation and is typically implemented through atomic force microscopy. Then there is shear modulus, S, that can be determined by anchoring one side of the fibril and moving the opposite side, for instance, through exposure to a fluid flow. One can also perform a variant of nanoindentation in which the fibril is suspended on two beams and then indented from a side at a place between the two beams. The resulting bending modulus, Y_b, depends on both Y_T and S.[5] All of these parameters are useful for characterization and understanding of the fibrils. However, quite often, only some of these techniques can actually be applied to a particular biomolecular system, so the information about the elastic properties is fragmentary.

It is thus desirable to develop computational methods that would allow for an estimation of various elastic constants within one model and provide insights into the deformation

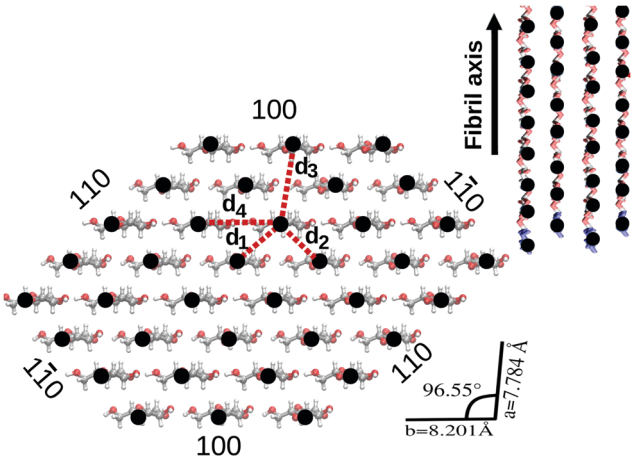
Institute of Physics, Polish Academy of Sciences, Aleja Lotników 32/46, PL-02668 Warsaw, Poland. E-mail: poma@ifpan.edu.pl, cieplak@ifpan.edu.pl

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c7cp05269c

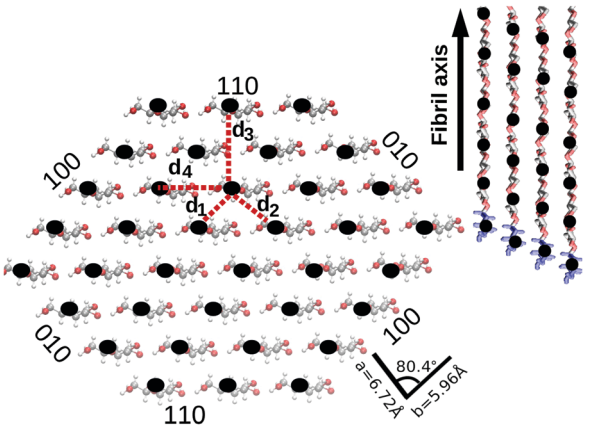
process. Here, we present a structure-based coarse-grained (CG) model which yields results which are consistent with disparate experimental data. Considering CG models seems to be particularly appropriate in this context because of the large sizes of the fibrils and because of the molecularly long timescales involved in the deformation processes. The natural CG degrees of freedom are associated with the effective particles placed at the locations of the C4 atoms in many carbohydrates and at the α-C atoms in proteins. We have developed a structure based CG model, together with the relevant parameters, for polysaccharide–protein complexes.[6] Here, we use it to determine Y_L, Y_T, and S for crystalline cellulose microfibrils and for two polymorphic forms of Aβ(1–40) amyloid. It should be noted that our CG theory has been originally developed[6] for a sugar–protein complex by using the common methodology. Thus the two different types of fibers considered here are studied within the same scheme.

Cellulose is one of the most abundant biopolymers on earth and it may constitute up to 50% of plant cell walls, as measured under dry conditions. Cellulose microfibrils wrap around plant cells. They provide resistance to osmotic pressure and generate turgor that allows plants to grow in an erect manner.[7] In crystallographic studies,[8] the microfibrils correspond to two possible crystalline allomorphs, denoted as Iα[9] and Iβ,[10] the predominance of which depends on the source organism (in nature, microfibrils can also be amorphous). A microfibril of cellulose consists of, typically, 36 parallel chains of cellobiose[11] as illustrated in Fig. 1. Each cellobiose unit consists of two D-glucose molecules (denoted here by D-GLC) which are connected by a β-1,4-glycosidic covalent bond (C–O–C). The number of repeat units is between 250 and 7000. Cellulose has attracted scientific interest not only because of its remarkable mechanical properties, but also because of its surface chemistry resulting from hydration.[12,13] For cellulose I allomorphs, the tensile and shear deformations are implemented along the *c*-axis direction, whereas indentation takes place perpendicular to the 100 plane for Iβ and $1\bar{1}0$ plane for Iα. The measured typical Y_L for cellulose Iβ ranges from 133 to 155 GPa,[14] Y_T from 2 to 25 GPa,[15] and S from 1.8 to 3.8 GPa.[16] The theoretical estimate based on density functional theory (DFT) is 206 GPa for Iβ.[17]

Aβ(1–40) and the closely related Aβ(1–42) (with an extra C-terminal ionic bridge) are involved in Alzheimer's disease.[18–20] The chains of these proteins may aggregate in a variety of ways. Two of them, corresponding to two-fold[21] and three-fold[18] symmetry axes, are shown in Fig. 3. It appears that, *in vitro*, the two-fold symmetry has lower energy than the three-fold symmetry. In our calculations, we focus on both polymorphic forms because they have not been studied theoretically yet (in the context of deformations) and moreover because the three-fold system has been found in the brain tissues of patients with Alzheimer's disease.[18] Moreover, it has been suggested that the three-fold amyloid is considered as one of the most neurotoxic forms of low-molecular weight β-amyloid oligomers. They damage membranes and alter calcium ion homeostasis.[22,23] The propensity to find aggregates *in vivo* with a three-fold symmetry could be due to some selection mechanism in the brain environment or due to an unknown



(a) Cellulose Iβ



(b) Cellulose Iα

Fig. 1 Panels (a) and (b) show a view of the 36-chain microfibril of cellulose Iβ and Iα in the plane perpendicular to the fibril axis and along the axis. The crystallographic directions assigned for each side are displayed. Axes *a*, *b* and *c* define the monoclinic unit cell for Iβ formed by two chains. The elementary unit cell of Iα is triclinic and it contains one chain. The position of the CG beads located at the C4 atoms is represented by black circles. Typical non-bonded interactions between CG beads are shown by colored dashed lines.

mechanism of fibril clearance.[18,24] The structure of the β-amyloid fibril forms through the self-assembly of peptide chains, either 40[25] or 42[26] amino acid residues in length, within layers which is followed by stacking of the layers. Amyloid is an example of a fibril with a very different architecture than cellulose. The juxtaposition of two kinds of fibrils provides a good test of the proposed CG theory. We focus on the more abundant and faster aggregating Aβ(1–40) instead of Aβ(1–42) but we expect that the two systems should be similar mechanically. For β-amyloid, the experimental values of S and Y_T are about 0.3 GPa and 3.3 GPa respectively.[27] They were obtained for Aβ(1–42) *in vitro*, but the experimental conditions did not allow for identification of the type of symmetry. The experimental samples are expected to be composed of fibrils of various symmetries and also to contain amorphous aggregates. The value of Y_L has not been reported.

Our structure based CG model involving the C4 atoms[6] is different from several other CG descriptions proposed for cellulose. In our CG model, the strength of the bonded and non-bonded potentials was obtained by matching their respective energy scales from all-atom simulations. The model by Srinivas *et al.*[28] was designed to reach long time scales which are necessary to study the crystalline–amorphous transition but it was not applied to the determination of the elastic properties of the fiber. Their model introduces two types of effective atoms. The crystalline unit cell of cellulose Iβ is monoclinic and it contains two cellobiose chains, denoted as origin (OR) and center (CE). In this way, an iterative version of the Boltzmann method which is applied between CE–CE, OR–OR and CE–OR all-atom radial distribution (RDF) functions is used to derive tabulated potentials for the effective non-bonded interactions. Fan *et al.*[29] designed a model to characterize the local effects of the axial deformation of longer fibrils (>100 nm) without, however, determining Y_L. Their method uses one CG bead per D-GLC monomer positioned at the center of mass of the ring. The non-bonded interactions are mapped to the Morse potential with the depth set to 5 kcal mol^{-1}, which is representative of the moderate O–H···O hydrogen bond (HB) energy strength in solids. By considering the RDF for the rings, they have associated each RDF peak with one type of HB and by Boltzmann Inversion of the Gaussian fitting of the widths of the peaks they derived the effective non-bonded potentials.

Glass *et al.*[30] have proposed a model, known as REACH (for Realistic Extension Algorithm *via* Covariance Hessian), that makes use of the elastic network system in which the contact interactions are represented by unbreakable harmonic bonds instead of the Lennard-Jones interactions. The REACH model makes cellulose to be necessarily much stiffer than in reality and yields Y_L of 162 GPa and Y_T of 25 GPa. On the other hand, an approach based on the MARTINI force field[31] (in which groups of several atoms, including the molecules of water, are represented by single effective atoms) yields[32] Y_T of 10 GPa – closer to our estimate. For amyloid fibrils, there is a CG approach which involves the normal-mode analysis of the α-C-atom-based elastic network model,[33,34] combined with the continuum deformation theory. For Aβ(1–40), it leads to Y_L of 20 GPa and S of 5.6 GPa.[35] However, this approach relies on the arbitrary selection of normal modes to analyze. As a result, the elastic constants obtained appear to be way too high. Our model allows for calculation of all elastic constants in a unified manner and yields results that are consistent with the experimental data.

In the next section, we describe the systems we study and outline our methodology for the simulations. We then present results pertaining to the tensile, shear and indentational deformations for the systems studied.

II. Methods

A. Determination of the elastic moduli

The unit cell for cellulose Iα is triclinic[9] and it contains one chain. For Iβ it is monoclinic[10] and contains two chains. The structural arrangements corresponding to the Iβ and Iα allomorphs are shown in Fig. 1, where d_i indicates the distance to atoms which are the nearest neighbors to the central atom indicated. Our CG model for cellulose I allomorphs was tested against all-atom simulations.[36] In particular, the CG model was able to capture the radial distribution function for these two systems. Then, the free energy difference, $F_\alpha - F_\beta$, was found to be about ~1.3 kcal mol^{-1}, agreeing with the expectation that Iβ is more stable than Iα.[37,38] One application of our model has been the determination of the transition pathways between Iα and Iβ at room temperature. The transition takes place through an amorphous phase in which about 10% of non-bonded contacts are missing compared to the crystalline phases. For β-amyloid, the CG model preserves the all-atom derived fibril structure within the α-C RMSD (root-mean-square-deviation) of 3 Å.

Fig. 3 shows the schematic setups that are used to induce tensile, shear and indentational deformations in the biological fibers. It also shows the indentational direction and its expected outcome. We induce the deformations in the constant speed mode. In the tensile case, we attach elastic springs to the ends of individual chains of initial length L_0. On one side, the springs have the elastic constants of k_1 (100 kcal mol^{-1} Å^{-1}) and are attached to an unmovable support. On the other side, the elastic constants are k_2 (0.1 kcal mol^{-1} Å^{-1}) and the outer ends of the springs are moving with a speed of v_{pull}. The initial length of these springs is considered to be numerically negligible. As a result of stretching, the fibril extends to length L and the strain, ϕ, is defined as $(L - L_0)/L_0$. The stress is determined as the net force exerted by springs k_2 divided by the cross-sectional area, A, of the fibril. Y_L is obtained in the regime in which the dependence of F/A on ϕ is linear. For cellulose Iα and Iβ, the cross section is hexagonal in shape and A can be determined from the locations of the C4 atoms on the edges. A hexagon can be considered as the sum of a number of triangles with corners defined by three nearby C4 particles. The sides of each triangle are along the 110, 1$\bar{1}$0 and 100 crystallographic directions for Iβ and along 110, 100 and 010 for Iα defined in Fig. 1. We project the triangles into the single plane of the hexagon. The area of each triangle can be estimated to be $A = \sqrt{s(s-b)(s-d_1)(s-d_2)}$ with $s = \dfrac{b + d_1 + d_2}{2}$. Here d_1, d_2 and b are distances defined in Fig. 1. In the case of amyloid fibrils, the cross-sectional area depends on the symmetry. For the three-fold symmetry, we calculate the triangular area defined by the locations of the α-C atoms on the first residues in each peptide chain denoted as 1, 1′ and 1″ (see Fig. 2(a)). For the two-fold symmetry, we calculate the rectangular area defined by amino acid residues 14 and 14′ (see Fig. 2(b)). Each elementary area was monitored during the simulation.

The shear deformation is also implemented at a constant speed. In this case, we attach an elastic spring to each effective atom on the top (k_2) and the bottom (k_1) planes. The top layer atoms are made to move along the fiber and the bottom ones are anchored. If the top plane moves by X, then the shear-strain is defined as $\phi = X/y$ (see Fig. 3). The shear-stress is determined by the total force acting on the top plane divided by the area of the plane. The top plane is rectangular in shape, so its area is

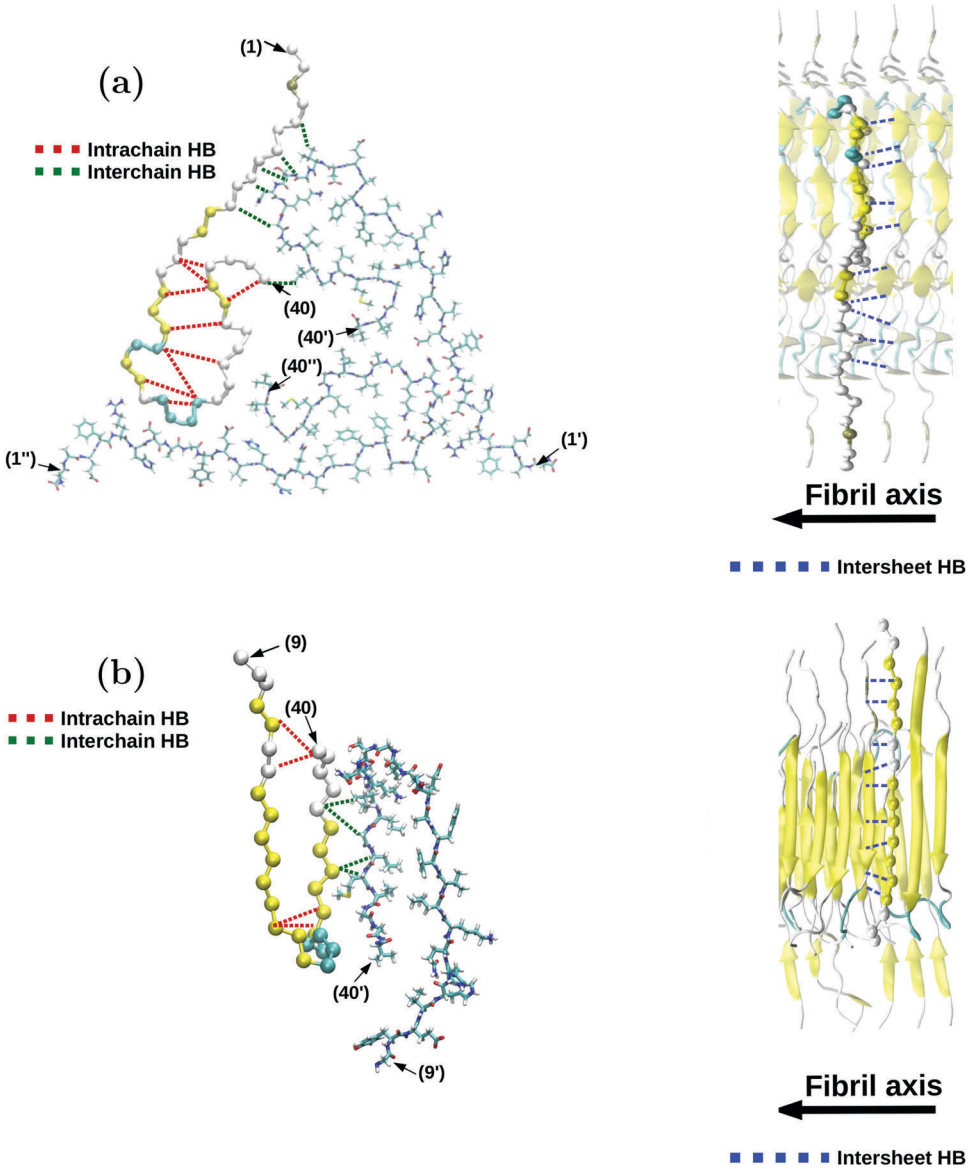


Fig. 2 Top panels show similar projections as in Fig. 1, but for Aβ(1–40)-amyloid with three-fold symmetry. The bottom panels are for β-amyloid with two-fold symmetry. The position of the α-C atoms for one peptide chain is represented by circles colored according to the type of the secondary structure. Panel (a) shows the first and last residues in a peptide. They are denoted by 1, 1′ and 1″ and 40, 40′ and 40″ and indicated by arrows. Panel (b) shows the same for two-fold symmetry. Typical non-bonded interactions between CG beads are shown by colored dashed lines.

equal to the product of the length of the fibril and its width. For the three-fold symmetric amyloid, the width is given by the distance between α-C atoms denoted as 1′ and 1″ in Fig. 2(a) and the anchored bottom part of the fibril consists of a line of α-C atoms denoted as 1. For the case of the two-fold symmetry, the shear plane is defined by amino acids in the range of residues 14–20 and the same set are anchored in the other chain.

In the case of nanoindentation one considers a rigid object with the radius of curvature R_{ind} that is approaching the fibril from the side with a speed of v_{ind}. When the forward tip of the object travels inside the fibril a distance of h away from the undeformed situation, then the force of reaction exerted by the fibril is F and the object ceases to touch the deformed fibril at the distance of a_{c} away from the vertical central axis. The relationship between F and h is given by

$$F = h^{3/2}\frac{4{R_{\mathrm{ind}}}^{1/2}Y_{\mathrm{T}}}{3(1-\nu^2)} \quad (1)$$

where ν denotes the Poisson ratio of the microfibril. It is defined as the ratio of the transverse contraction strain to the axial strain in the direction of the stretching deformation. We have computed this coefficient (Fig. S1 in ESI†) from our simulation obtaining 0.33 for cellulose allomorphs and 0.40 for both of the amyloid polymorphs. Our estimate for cellulose agrees with the experimental value in the literature.[39] For amyloids the value has not been reported. The relationship follows from the Hertz theory[40] of mechanical interactions between two spheres as extended by Yoffe.[41] It is assumed that

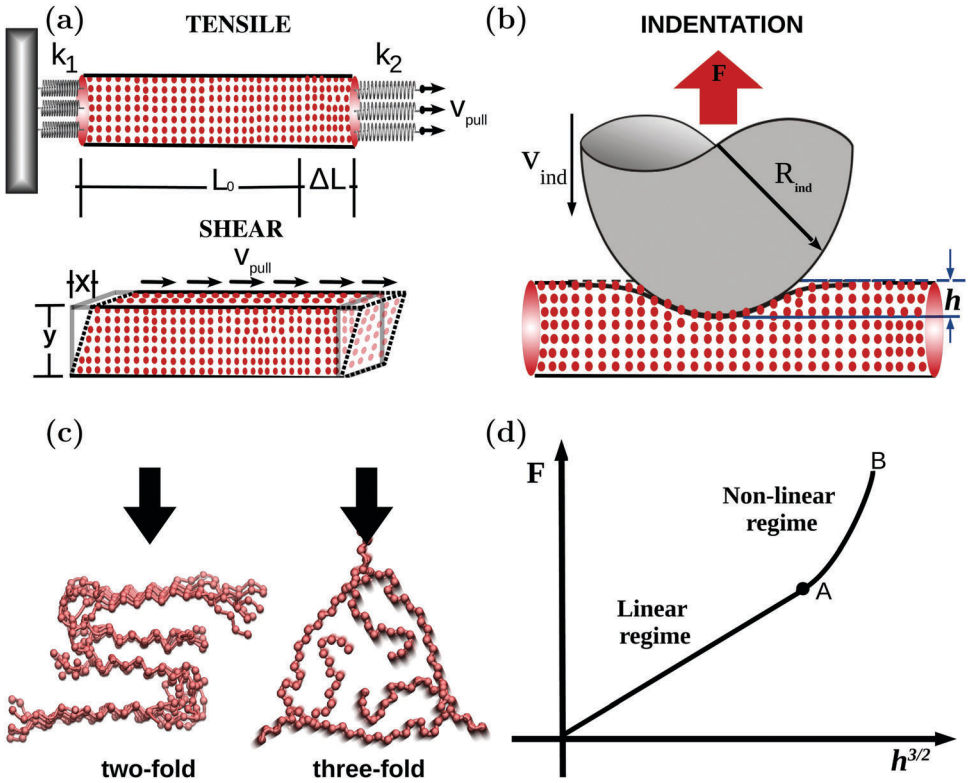


Fig. 3 Schematic representation of the three types of deformations considered and the characteristic parameters used in this paper. Panels (a) and (b) correspond to the tensile and shear, and indentational deformations respectively. Panel (c) shows the indentational direction for the Aβ(1–40) system. Panel (d) shows the typical shape of the experimental indentational curve: a linear response to deformation is followed by a non-linear, or plastic, deformation. The values of the elastic moduli are estimated in the linear regime for the process. The transition from the linear to non-linear regime corresponds to point A.

the initially flat microfibril corresponds to a sphere with an infinite radius of curvature and that the indenter is incompressible (infinite Young's modulus). Thus plotting F *vs.* $h^{3/2}$ allows one to deduce Y_T from the linear part of the plot.

In real experiments, the indenter has a finite stiffness and the resulting compliance has to be subtracted from the full indentation depth.[42,43] In simulations, we can eliminate the compliance of the indenter by making it sufficiently stiff. It should be noted that a reversal in the direction of motion of the indenter (known as unloading) may introduce an effective irreversibility which shows as a hysteresis. This happens if one enters the nonlinear part of the plot and is due to the longer timescale of the fibril recovery compared to the motion of the indenter, which, in turn, is due to heat dissipation. This point is discussed, for instance, in the context of modelling of nanoindentation of virus capsids.[44,45]

B. The structure-based CG model

The CG model employed here uses the positions of atoms in a well-relaxed atomistic structure and then it sets the bonded and non-bonded interactions between the effective atoms representing the structural units. This approach would also be valid for systems with mutations that may alter the HB network and lead to structural rearrangements. In this case, the mutated (average) structure has to be first derived, for instance, through all-atom simulations. Our CG model for allomorphs of cellulose[6,36] is based on representing D-GLC monomers by beads located at the C4 atoms (Fig. S2 in ESI†). For amyloid, it is based on representing amino acids by beads located at the α-C atoms. In both cases, the potential energy is given by

$$V^{\mathrm{CG}} = \sum_{\text{bonds}} K_{\mathrm{r}}(r-r_0)^2 + \sum_{\text{angles}} K_\theta(\theta-\theta_0)^2 + \sum_{\text{dihedrals}} K_\phi(\phi-\phi_0)^2 + \sum_{i<j}^{\mathrm{CON}} 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \sum_{i<j}^{\mathrm{NO\text{-}CON}} 4\varepsilon'\left(\frac{r_{\mathrm{cut}}}{r_{ij}}\right)^{12}. \quad (2)$$

The first three terms on the right hand side of eqn (2) correspond to the harmonic pseudo-bond, bond angle and dihedral potentials. Table 1 shows the values of elastic constants, K_{r}, K_θ and K_ϕ, for each of the systems studied here as derived in ref. 6 and 36 through all-atom simulations. The values for r_0, θ_0 and ϕ_0 correspond to the equilibrium distance between two, three and four beads respectively. They are listed in Table 1. In the case of amyloid, we first determine an effective native conformation (NC) through all-atom simulations as described later, and the angles and distances correspond to the derived structure. Notice that the harmonic description of the covalent bonds between the consecutive effective C4 atoms precludes

Table 1 top table shows bonded parameters used to perform our CG simulation of the native cellulose I allomorphs and the Aβ(1–40) microfibrils. The equilibrium values of bending (θ_0) and dihedral (ϕ_0) angles for Aβ(1–40) are taken from the native conformation (NC) and used in eqn (2) to define the potential energy. The bottom table shows non-bonded potential parameters for three systems. The unit of ε is given in kcal mol^{-1}. The subscript d_i for the cellulose parameters corresponds to the four nearest neighbors depicted in Fig. 1(a)

	k_r [kcal mol^{-1} Å^{-2}]	r_0 [Å]	k_θ [kcal mol^{-1} rad^{-2}]	θ_0 [°]	k_ϕ [kcal mol^{-1} rad^{-2}]	ϕ_0 [°]
Cellulose Iβ	51.6 ± 7.5	5.27	160.1 ± 42.6	169	2.1 ± 0.3	183.6
Cellulose Iα	51.7 ± 7.1	5.28	182.0 ± 46.0	170	2.1 ± 0.2	181.0
Aβ(1–40)	100.0 ± 11.0	3.80	45.0 ± 11.1	NC	5.0 ± 2.2	NC

	ε_{d_1}	σ_{d_1} [Å]	ε_{d_2}	σ_{d_2} [Å]	ε_{d_3}	σ_{d_3} [Å]	ε_{d_4}	σ_{d_4} [Å]	ε	σ [Å]
Cellulose Iβ	2.3	5.60	2.3	5.60	3.0	7.10	7.4	7.43	—	
Cellulose Iα	1.9	5.70	1.9	5.70	2.5	6.90	7.3	7.50	—	
Aβ(1–40)	—		—		—		—		1.5	NC

their rupture. Thus the simulations cannot lead to the stress–strain relationship illustrated in Fig. 3. Instead, in the non-linear regime, the stress must grow indefinitely whereas breakage of the bonds would lead to leveling off followed by fragmentation.

The fourth term on the right-hand side of eqn (2) takes into account the non-bonded contact interactions, as described by the Lennard-Jones potentials. For cellulose, the contact interactions correspond to four different distances d_1, d_2, d_3, and d_4 between the C4 atoms (see Fig. 1). These interactions are due to the HBs between specific atoms (like O and C) which are in distances shorter than those between the C4 atoms. The corresponding values of the depth, ε_{ij}, of the potential well and the length parameters σ_{ij} are listed in Table 1. For Aβ(1–40), we take ε_{ij} to be uniform and equal to ε, also listed in the table. This approach has been shown to work well when comparing the theoretical results on protein stretching to the experimental data[46–48] and on nanoindentation of virus capsids.[44,45] The value of 1.5 kcal mol^{-1} for ε is close to the one obtained by comparison to stretching. In all systems, the strength, ε', of the repulsive term is set as equal to ε.

For proteins, the presence of a native contact is determined by the overlap criterion.[46,49] This criterion works as follows: each heavy atom is assigned a van der Waals radius as determined in ref. 50. A sphere with the radius enlarged by 1.24 is built around the atom. If two amino acids have heavy atoms with overlapping spheres then the contact is declared to exist. These contacts can be due to HBs, hydrophobicity and ionic bridges. Only contacts with $|i-j| > 4$ are included in the contact map. The parameters σ_{ij} are given by $r_{ij0}/2^{1/6}$, where r_{ij0} is the distance between two α-C atoms that form the native contact. The last term in eqn (2) describes repulsion in non-native contacts. In the protein case, we take r_{cut} = 4 Å, and in the case of cellulose, r_{cut} = 5 Å. For proteins, we also include one more term: the local backbone stiffness. We take it as described by the bond-angle and dihedral potentials that favor the native geometry.[46]

The CG simulations were carried out with an implicit solvent at 300 K which corresponds to k_BT = 0.59 kcal mol^{-1} (k_B is the Boltzmann constant, T the temperature). The Langevin velocity-dependent (over) damping and random forces were used to represent effects due to the solvent, including thermostating. The characteristic time scale, τ, in the problem is of order 1 ns. In our model, this time-scale corresponds to the diffusional time to move by 5 Å a CG bead (*e.g.* α-C atom). A similar time scale for an α-C protein model, $\tau \sim 3$ ns, was derived by Veitshans *et al.*[51] It was also obtained through consideration of stretching of a protein by a fluid flow at the same Peclet number as in experiments.[52] The equations of motion for the beads are solved by a fifth order predictor-corrector method. The integration step is $\tau/200$. The systems were studied typically for a duration of 1000τ after equilibrating them for 100τ. To capture characteristic thermal fluctuations for each system, our results were averaged over 50 trajectories. The forces are monitored at time intervals of τ. We have considered averaging the forces over an interval of 10τ. This was sufficient to capture the characteristic information for stress–strain and F–$h^{3/2}$ curves.

It should be noted that the uniform-ε model used here has been selected out of 62 models studied in ref. 47, including those with specificity. When compared to the experimental data on protein stretching, it was matching or outperforming descriptions based on non-uniform values of ε. The sequence specificity is contained in the native structure that is used to construct the contact map which is the most important feature of the model.

In nanoindentation studies, the microfibril or the amyloid fibril, of length listed in Table 2, is placed above a repulsive plate and the deformation is induced towards the plate. The interaction with a flat plate is described by a repulsive potential[53] that scales as z_0^{-10}, where z_0 is the distance between the plate and the CG bead. The dynamical role of the indenter is described by a sphere that generates a purely repulsive

Table 2 List of geometric parameters of the fibril structures used in our studies of the deformations

	Cellulose Iβ	Cellulose Iα	Aβ three-fold	Aβ two-fold
Initial length, L_0 [nm]	39.10 ± 0.12	39.24 ± 0.23	42.21 ± 0.34	41.10 ± 0.33
Cross-sectional area, A [nm^2]	10.02 ± 0.14	10.30 ± 0.20	21.11 ± 0.33	16.02 ± 0.41
Shear plane area, A [nm^2]	56.03 ± 0.22	52.10 ± 0.30	170.20 ± 0.41	160.01 ± 0.11

potential: $V_{ind} = 4\varepsilon_{tip}\left[\left(\frac{\sigma_0}{r}\right)^{12}\right]$ for $r < R_{ind}$ and with $\sigma_0 = R_{ind}/2^{1/6}$. R_{ind} defines the curvature of the indenter. The value of ε_{tip} has to be sufficiently large so that the effects of the compliance of the indenter are negligible. We find that this happens for ε_{tip} larger than 10 kcal mol^{-1}. In practice, we run the simulations with ε_{tip} = 20 kcal mol^{-1}. Most of the indentations were implemented at a speed of v_{ind} = 0.005 Å/τ($\sim$0.5 $\times$ 10^{-3} m s^{-1}). This value was found to be sufficiently slow to generate results that are close to the experimental nanoindentation profiles obtained for cellulose I[54–56] at a speed of $\sim$0.8 $\times$ 10^{-5} m s^{-1}. It is also five orders of magnitude slower than the corresponding all-atom simulations.[57]

C. The structural relaxation through all-atom simulations

The molecular dynamics (MD) simulations were conducted with version 2.9 of the NAMD package.[58] The cellulose[59] and β-amyloid[60] fibrils were parametrized using the CHARMM all-atom force field. A solvation box with TIP3P water molecules[61] was used to allow for structural relaxation of the initial structures. The number of water molecules was 120 000 and 165 000 for cellulose and β-amyloid with three-fold symmetry. The crystalline initial states of cellulose were prepared by using the cellulose-builder toolkit.[62] The cellulose fibrils were made of 40 cellobiose units with each extending for 1.04 nm along the *c*-axis.

The amyloid with three-fold symmetry was neutralized by 675 sodium counterions in the solvation box. We assemble the three-fold system fibril of length 35 nm by starting from a nine-peptide and three-layer structure with the PDB code 2M4J[18] (see Fig. 2). The other amyloid fibril with two-fold symmetry and a length of 40 nm has been constructed from a 12-peptide and six-layer structure with the PDB code 2LMO.[63] After equilibration and relaxation, the three-fold amyloid fibril acquires a length of about 42 nm. The three-fold system was replicated 25 times along the fibril axis using the CreateFibril 2.5 toolkit,[64] whereas the two-fold symmetry case was replicated 12 times. Afterwords, all-atom simulation of the whole system was performed to minimize the energy and to obtain a conformation in the solution. This process was carried out for both cellulose allomorphs and the amyloid with three-fold symmetry. The two-fold symmetry case was studied without the need of an exhaustive all-atom relaxation. As we discuss later, our CG methodology was found to be not very sensitive to the initial structure (*i.e.* solvated or in vacuum conditions) when determining the elastic moduli. The simulation box reached the following dimensions: L_x = 13.2 nm, L_y = 12.7 nm and L_z = 39 nm (z is the direction along the axis of the fibril) for the amyloid (three-fold) fibril and L_x = 6.0 nm, L_y = 6.2 nm and L_z = 46 nm for cellulose allomorphs. Periodic boundary conditions were used. Numerical integration of Newton's equations of motion involved a time step of 1 fs and the atomic coordinates were saved every 1 ps for analysis. The system equilibration was carried out in the following way: the first 1000 steps of energy minimization were followed by a short 0.5 ns run in the *NPT* ensemble to achieve a pressure of 1 bar. The production runs were carried out in the *NVT* ensemble for 40 ns at T = 300 K. The temperature was controlled by the standard Langevin algorithm and the pressure by the Langevin piston pressure control algorithm.[65] The structure obtained at the end of these simulations was used as an input for deformation studies.

III. Results and discussion

A. Tensile deformation

The stress–strain curves for the four systems studied are shown in Fig. 4. The result for cellulose Iα is similar to that for Iβ (Fig. S3 in ESI†). The initial lengths and cross-sectional areas, A, are determined based on the equilibrated structures. They are summarized in Table 2. The insets in the left-hand panels of Fig. 4 show A as a function of strain. The deviations are seen to be small compared to the mean, especially in the amyloid

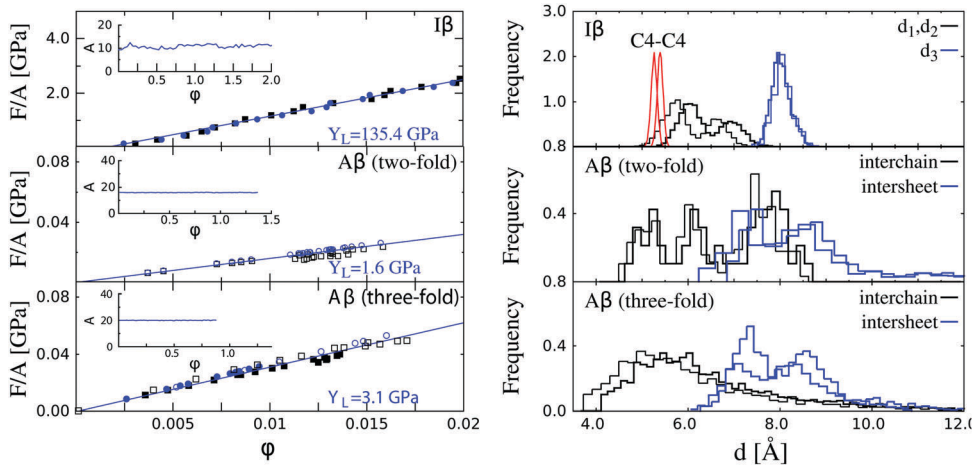


Fig. 4 Tensile deformation at different pulling speeds. The left panels show stress–strain curves of cellulose Iβ and β-amyloids with two-fold and three-fold symmetry. The square and circle data points correspond to v_{pull} of 0.005 and 0.0005 Å/τ. The error bars are the same as the symbol size and they are based on 50 independent simulations for each structure. The insets show the corresponding cross-sectional areas in nm^2 for the slowest pulling speed. The right panel shows distributions of HB lengths for ϕ = 0 (solid lines) and for a finite ϕ corresponding to the end of the linear regime (dashed lines): for cellulose allomorphs this ϕ is equal to 0.02, whereas for Aβ(1–40) it is 0.013. The types of HB contacts are denoted as in Fig. 1 and 3.

cases, so the stress can be calculated using the initial value of A. The values of Y_L obtained for ν_{pull} = 0.005 Å/τ are listed in Table 3. For cellulose, they are seen not to depend on the type of the allomorph, and are essentially the same for ν_{pull} = 0.05 Å/τ. They are also comparable to the values obtained experimentally and through all-atom simulations.[14] It should be noted however that the experimental value of Y_L for Iα is visibly smaller than for Iβ. The differences may have to do with the actual degree of crystallinity of the sample. In our studies, the deformation is implemented precisely along the c-axis for cellulose and the main fibril axis for β-amyloid. However, the experimental determination of Y_L involves microscopic thin strips of compressed cellulose so that an accurate alignment along the c-axis is difficult to obtain. The dependence on ν_{pull} is much stronger for Aβ(1–40) amyloid. For ν_{pull} = 0.05 Å/τ we report a Y_L of 13.0 GPa and 7.20 GPa for three-fold and two-fold symmetries respectively. We have found that, on lowering ν_{pull}, the stress–strain curve saturates at 0.005 Å/τ and for 0.0005 Å/τ we get essentially the same results: Y_L of 3.1 GPa and 1.6 GPa for three-fold and two-fold symmetries respectively. All-atom simulations for Aβ(1–40) (at constant force) result in Y_L in the range 2.3–12.4 GPa[66] and those for the related elastic network of human islet amyloid polypeptides (hIAPP) −13 GPa.[67] The precise symmetry of the fibril is important for the mechanical properties. The all-atom simulation study[66] was done for the two-fold symmetry *in vitro* model and it has shown a large variation of Y_L compared to our CG study of the two-fold symmetry model. Such variability in the Y_L value can be explained in terms of the large forces applied to initiate the deformation process in the all-atom simulations, which results in a strong response of the system. Our result for the three-fold-symmetry system shows that it can withstand larger stresses than the two-fold symmetry one. Experimentally, the two-fold system is thermodynamically more stable in terms of the size of aggregates than the three-fold symmetry system.[24] This result may imply that a two-fold symmetry is more relevant in the progression of Alzheimer's disease.

Table 3 The elastic moduli for cellulose I allomorphs and the Aβ(1–40) amyloid as obtained by experimental and theoretical methods. The listed CG results are obtained in this paper. The all-atom results for cellulose I allomorphs are cited after ref. 14 and 57 and after ref. 66 for β-amyloid. The tensile modulus computed *via* DFT calculations along the c-lattice is cited after ref. 17. All-atom tensile deformation was implemented by changing the c-lattice length parameter and then minimizing the energy. The experimental entries are taken from ref. 68 and 56 respectively. The nanoindentation velocities of cellulose Iβ *via* all-atom, CG and experiment are 50 m s^{-1}, 0.5 × 10^{-3} m s^{-1} and 0.8 × 10^{-5} m s^{-1} respectively. The structural symmetry of β-amyloid (if specified in the literature) is given next to calculated elastic modulus. Experimental nanoindentation and shear data corresponding to Aβ(1–42) are taken from ref. 27

	Cellulose Iβ	Cellulose Iα	Aβ(1–40) (symmetry)
Y_L (GPa)			
All-atom[14]	150	152	—
All-atom[57]	139.5	—	—
All-atom[66]	—	—	(2.3–12.4) (two-fold)
DFT	206	—	—
Experiment	(133–155)	114	—
CG	135.4	138.3	3.1 (three-fold) 1.6 (two-fold)
S (GPa)			
All-atom[16]	1.6	—	—
Experiment	(1.8–3.8)	—	0.3 (not known)
CG	3.2	3.1	1.2 (three-fold) 0.6 (two-fold)
Y_T (GPa)			
All-atom[57]	5.1	—	—
Experiment	(2–25)	—	3.3 (not known)
CG	7.2	7.1	7.0 (three-fold) 21.3 (two-fold)

The right panels in Fig. 4 show the distributions of the lengths, d, in the contacts as determined first in the starting state (the solid lines) and then at the point of the largest strain considered (the dashed lines) at ν_{pull} = 0.005 Å/τ. One would expect elongation of the distances on stretching. For cellulose, this is indeed observed for the C4–C4 direction along which the interactions are mainly covalent. Fig. 6 shows a substantial shift to larger values for the C4–C4 bond distribution. However, the intersheet HB contacts d_1 and d_2 (see Fig. 1) get compressed and there is very little change in d_3. We determine the distributions of d_1 and d_2 and observe that they are spread due to thermal fluctuations. Regardless of the type of the symmetry for Aβ(1–40) amyloid, the intersheet contacts get stretched but the interchain contacts within each sheet get compressed. In order to quantify these effects, we adopt a criterion that an at least 25% change in the contact length, relative to its native value, counts either as an extension or a compression. By this criterion, we find that for the three-fold symmetry case, about 2% of the interchain contacts are compressed and 1% of the intersheet contacts are stretched. For two-fold symmetry, about the same number for interchain contacts and only 0.5% of intersheet contacts are similarly affected.

B. Shear deformation

Our results for all the systems are shown in Fig. 5. The result for cellulose Iα is similar to that for Iβ (Fig. S4 in ESI†). The shear deformation for cellulose I allomorphs and β-amyloid takes place along the same directions as in the tensile deformation. The initial values of the top-plane areas are listed in Table 2. The insets in the left panels of Fig. 5 demonstrate that A does not change when the strain is increasing. The values of S obtained for ν_{pull} = 0.005 Å/τ are listed in Table 3. They do not depend on the type of the cellulose allomorph and are in agreement with the experimental results[15] and all-atom simulations.[16] S for Aβ(1–40) is 2.7 times smaller than for cellulose allomorphs. No experimental data on S for Aβ(1–40) has been reported, but it is close to 1.4 GPa measured for microtubules[5] which are assemblies of proteins known as α- and β-tubulins.

The right-hand panels in Fig. 5 show the distributions of the characteristic distances. The intersheet HB contacts d_1 and d_2 in cellulose (see Fig. 1) get compressed while there is essentially no change in d_3 and no change in the lengths of the covalent bonds C4–C4. For both symmetries of βA(1–40) considered here, the intersheet HB contacts get slightly stretched but the

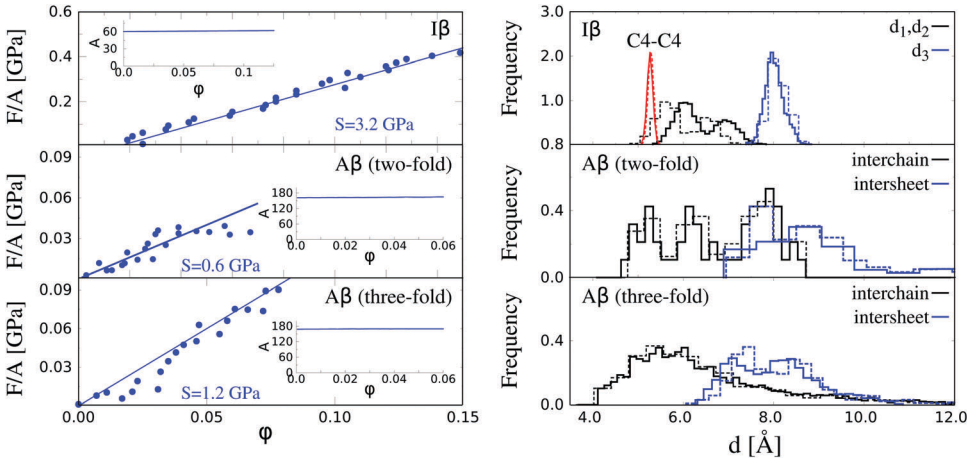


Fig. 5 Shear deformation at a pulling speed of 0.005 Å/τ. The left panels show stress–strain curves of cellulose Iβ and β-amyloids with two-fold and three-fold symmetry. The error bars are the same as the symbol size and they are based on 50 independent simulations for each structure. The inset shows the corresponding cross-sectional area in nm^2. The right panels show the distributions of the HB lengths for ϕ = 0 (solid lines) and for a finite ϕ corresponding to the end of the linear regime (dashed lines): for cellulose allomorphs this ϕ is equal to 0.15, whereas for Aβ(1–40) with two-fold symmetry it is 0.04 and 0.075 for the three-fold case. The types of HB contacts are denoted as in Fig. 1 and 3.

distances in the interchain contacts within each sheet are not affected.

C. Nanoindentation

Motivated by the experimental procedures[55] used for cellulose, we consider the tip with R_{ind} of 100 Å and expect F to be between 10 and 30 nN when h is about 10 Å. This is indeed observed for ε_{tip} of 10 and 20 kcal mol^{-1} but not for 1 kcal mol^{-1}. For the indentational calculations we consider the latter value for ε_{tip}. We take v_{ind} of 0.005 Å/τ to match the tensile situation. Fig. 6 shows our results for both allomorphs of cellulose. The initial plot of F *vs.* $h^{3/2}$ is indeed linear which allows us to determine Y_T. We find that the values of Y_T for Iα and Iβ are nearly the same (7.1 and 7.2 GPa respectively). The REACH-based approach[30] yields 25 GPa for Y_T. In this case, the covalent couplings influence Y_T only in a minor way and yet they are effectively overemphasized in REACH by relying on the elastic network approach and hence the factor of 4 larger value of Y_T. Still, the CG and REACH models yield elastic moduli which are strongly anisotropic. We observe that the intersheet HB contacts of types d_1 and d_2 get compressed on indentation while those of the type d_3 are affected little (Fig. S5 in ESI†). This is because the d_1 and d_2 contacts tend to align parallel to the indentation direction. No deformation is observed for covalent C4–C4 bonds as they align perpendicular to the direction of indentation.

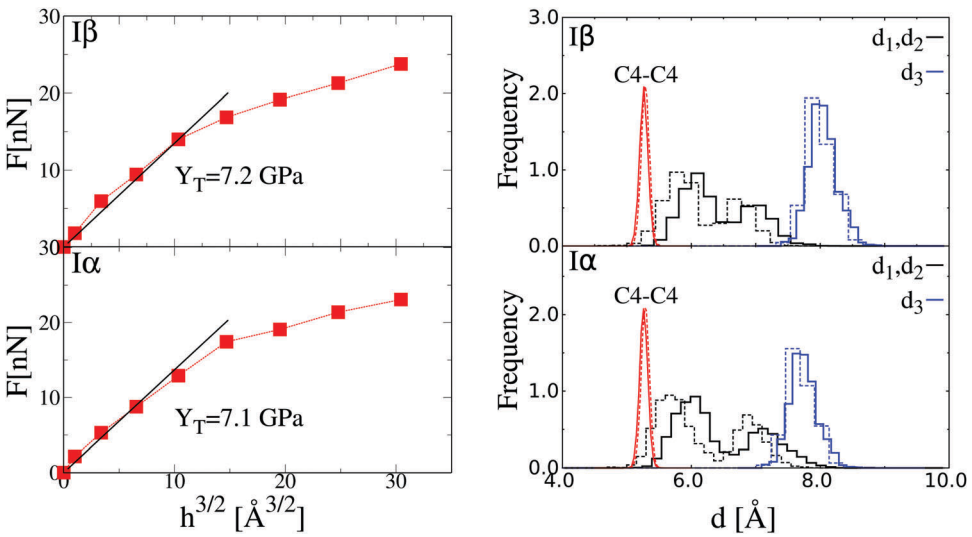


Fig. 6 The left panel shows the F–$h^{3/2}$ curve for cellulose I allomorphs at v_{ind} = 0.005 Å/τ and R_{ind} = 100 Å. The error bars are the same as the symbol size and they are based on 50 independent simulations for each structure. The corresponding Y_T is shown next to the linear fit (black solid line). The right panel shows the distribution of lengths of HB contact involved during the process for $h^{3/2}$ = 0 (solid line) and $h^{3/2}$ = 10 (dashed line). The types of HB contacts are denoted as in Fig. 1 and 3.

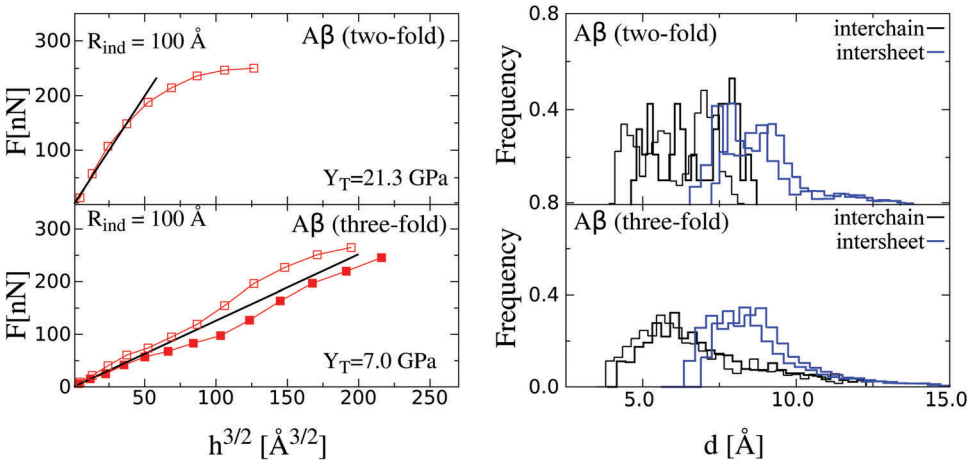


Fig. 7 Same plots as in Fig. 6, but for β-amyloid with indenter radius (R_{ind} = 100 Å). The open square symbols correspond to the structure derived from the PDB files without any equilibration in water. The filled symbols correspond to the relaxed structure. The distributions are calculated for h = 0 (solid line) and for h = 28 (three-fold symmetry) and h = 11 (two-fold symmetry) (dashed line).

Fig. 7 shows our nanoindentation results for both symmetries in Aβ(1–40). We consider two indenter radii: 50 Å and 100 Å. The results depend on R_{ind} rather weakly and the initial slopes of the plot of F against $h^{3/2}$ are only about 2 GPa smaller than for cellulose. However, the extension of the linear regime based on indentation depth (h) is at least 7 times larger when compared with three-fold symmetry and only 3 times larger for the other symmetry. We also observe that the interchain contacts become shorter while the intersheet contacts get longer.

A summary of our CG results is presented in Table 3. For cellulose Iβ, we find that Y_T is 7.2 GPa whereas Y_L is 135.4 GPa, *i.e.* about 20 times larger – both results are consistent with the experimental results[56,68] (see also ref. 69 and 70). The S modulus for cellulose allomorphs is calculated to be 3.2 GPa whereas experimental results on S fall in the range of 1.8–3.8 GPa, depending on the source of cellulose.[15] Interestingly, the values of Y_L and Y_T are larger than Y_L of 1.2 GPa reported for spider silk,[71,72] still larger than of order of 0.1 GPa for microtubules,[5] and much larger than up to 0.015 GPa reported for fibrin.[4] The large value of Y_L in cellulose correlates well with the fact that cellulose is the most recalcitrant carbohydrate polymer to catalytic degradation compared to other plant cell wall polysaccharides.[73,74] For Aβ(1–40), we find Y_L to vary between 1.6 and 3.1 GPa and Y_T between 7.0 and 21.3 GPa according to the type of amyloid symmetry. The anisotropy between the longitudinal and transverse directions is significantly reduced compared to the cellulosic case and, more interestingly, its sense is reversed. The experimental value of Y_T for the related Aβ(1–42) is 3.3 GPa[27] which differs from our value by a factor of 2 but not by an order of magnitude. Our values correspond to a highly ordered amyloid system, whereas in experiment it is difficult to achieve such well defined conditions. The existence of amorphous aggregates/fibrils across the sample may in fact decrease the real value of Y_T. Moreover, the experimental value of Y_L appears to be unavailable. Adamcik *et al.*[75] find that Y_T for many amyloid fibrils (assembled from α-synuclein, heptapeptides, insulin, β-lactoglobulin, tau protein, lysozyme, ovalbumin, and bovine serum albuminum) all fall in the range of 2–4 GPa (see also ref. 76 and 77). The same observation applies to collagen.[78] The theoretical prediction of S is 1.2 GPa and it is four times larger than the experimentally obtained value for the insulin amyloid fibril[76] and about the same as for microtubules.[5] The data can be explained in terms of the deformation of HBs between the β-sheets. Notice that Y_T for cellulose Iβ is comparable to that of the amyloids but Y_L is much larger. This is because stretching in cellulose must affect covalent bonds[17,30,79] but in the amyloids, only the HB contacts.

Aβ(1–40) is found to be much more isotropic elastically than the cellulose microfibrils since Y_L and Y_T differ by a factor of 0.5 instead of 20. The cellulose microfibrils studied here are softer along the transverse direction compared to the longitudinal direction because tensile deformation involves covalent bonds. The opposite is seen for Aβ(1–40). In this case, all deformations affect primarily the network of the hydrogen bonds.

IV. Conclusions

We have considered the tensile, shear, and indentational deformations in crystalline cellulose I allomorphs and two polymorphs of the Aβ(1–40) system within one structure-based CG model. The resulting elastic moduli are found to be consistent with the experimental data and all-atom simulations whenever they are available. The nanomechanical characterization of the four systems based on the elastic anisotropy shows that $Y_L > Y_T$ for cellulose allomorphs and for the amyloid systems this inequality is inverted and the result is independent of the axial symmetry considered. The larger anisotropy for crystalline cellulose is associated with stretching of covalent bonds, while in the amyloids, stretching mainly involves hydrogen bonds.

The fibril axis in cellulose is along the covalently connected chains but this is not so in the fibrils made of amyloids. The tensile and indentational deformations are defined with respect to the fibril axis and hence larger anisotropy in cellulose

and smaller anisotropy in β-amyloids. The differences in the degree of anisotropy are enhanced by the features of the HBs as characterized by the depths of the corresponding potential wells. In cellulose, the values of ε vary between 1.9 and 7.4 kcal mol^{-1}. The largest values are in the d_4 direction – between two chains belonging to one plane. In β-amyloid, ε is smaller, 1.5 kcal mol^{-1}, and (approximately) the same in all directions. This is because the protein structure is much less regular than cellulose and directional features are smeared out by the presence of 20 types of amino acids. The anisotropy between the longitudinal and transverse directions in the β-amyloid fibers results from the fact that the interchain hydrogen bonds get compressed during nanoindentation in a more pronounced way than during stretching, as illustrated in Fig. 4 and 7.

Our CG model can be used to study the structural transitions between the cellulose allomorphs, in a way that is similar to the one based on the Martini CG force field for cellulose as developed by Lopez *et al.*[32] However, our method is computationally more efficient since it employs an implicit solvent. Our model would be convenient to use in studies of the behavior at various temperatures at which cellulose starts to ''melt'' or at which the β-amyloid fibril gets disordered. The elastic moduli are expected to become reduced on increasing the temperature. It would be interesting to determine the functional form of this reduction. The parameters of the CG models are determined at room temperature but they can be used at other temperatures as an approximation.

Another area of interest would be to study the elastic properties of the amorphous cellulose. As a first step in this direction, we have considered a model in which 10% of the hydrogen bonds in the crystalline Iβ are removed in places that miss the bonds in the transition state between Iα and Iβ.[36] In this approach, we avoid the kinetic effects of the crystallization in the amorphous state that take place on the time scale of nanoindentation. As expected, the bond removal makes the system softer. The interesting result is that Y_T becomes 5.1 GPa (Fig. S6 in ESI†) which is only about 70% of the crystalline value, indicating that the softening is not a linear function of the fraction of the bonds removed.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

A. Poma acknowledges helpful discussions with R. Szoszkiewicz regarding the theory of AFM-nanoindentation. This research has been supported by the European Framework Programme VII NMP grant 604530-2 (CellulosomePlus) and by the EU Joint Programme in Neurodegenerative Diseases project (JPND CD FP-688-059) through the National Science Centre (2014/15/Z/NZ1/00037) in Poland. It was also co-financed by the Polish Ministry of Science and Higher Education from the resources granted for the years 2014–2017 in support of international scientific projects. This research was supported in part by PLGrid Infrastructure.

References

1 E. Klotzsch, M. L. Smith, K. E. Kubow, S. Muntwyler, W. C. Little, F. Beyeler, D. Gourdon, B. J. Nelson and V. Vogel, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 18267–18272.
2 B. D. Oppel and J. E. Bond, *Biol. J. Linn. Soc.*, 2000, **70**, 107–120.
3 N. Becker, E. Oroudjev, S. Mutz, J. P. Cleveland, P. K. Hansma, C. Y. Hayashi, D. E. Makarov and H. G. Hansma, *Nat. Mater.*, 2003, **2**, 278–283.
4 W. Liu, L. M. Jawerth, E. A. Sparks, M. R. Falvo, R. R. Hantgan, R. Superfine, S. T. Lord and M. Guthold, *Science*, 2006, **313**, 634.
5 A. Kis, S. Kasas, B. Babic, A. J. Kulik, W. Benoit, G. A. D. Briggs, C. Schonenberger, S. Catsikas and L. Forro, *Phys. Rev. Lett.*, 2002, **89**, 248101.
6 A. B. Poma, M. Chwastyk and M. Cieplak, *J. Phys. Chem. B*, 2015, **119**, 12028–12041.
7 D. J. Cosgrove, *Nature*, 2000, **407**, 321–326.
8 R. H. Atalla and D. L. Vanderhart, *Science*, 1984, **223**, 283–285.
9 Y. Nishiyama, J. Sugiyama, H. Chanzy and P. Langan, *J. Am. Chem. Soc.*, 2003, **125**, 14300–14306.
10 Y. Nishiyama, G. P. Johnson, A. D. French, V. T. Forsyth and P. Langan, *Biomacromolecules*, 2008, **9**, 3133–3140.
11 Y. Nishiyama, *J. Wood Sci.*, 2009, **55**, 241–249.
12 K. Kulasinski, S. Keten, S. V. Churakov, D. Derome and J. Carmeliet, *Cellulose*, 2014, **21**, 1103–1116.
13 G. Nawrocki, P. A. Cazade, D. Thompson and M. Cieplak, *J. Phys. Chem. C*, 2015, **119**, 24404–24416.
14 S. Eichhorn and G. Davies, *Cellulose*, 2006, **13**, 291–307.
15 M. G. Northolt, H. Boertoel, H. Maatman, R. Huisman, J. Verunik and H. Elzerman, *Polymer*, 2011, **42**, 8249–8264.
16 Z. Zhao, O. E. Shklyaev, A. Nili, M. N. A. Mohamed, J. D. Kubicki, V. H. Crespi and L. Zhong, *J. Phys. Chem. A*, 2013, **117**, 2580–2589.
17 F. L. Dri, L. G. Hector Jr., R. J. Moon and P. D. Zavattieri, *Cellulose*, 2013, **20**, 2703–2718.
18 J.-X. Lu, W. Qiang, W.-M. Yau, C. D. Schwieters, S. C. Meredith and R. Tycko, *Cell*, 2013, **154**, 1257–1268.
19 M. Goedert, *Science*, 2015, **349**, 1255555.
20 Y. Xiao, B. Ma, D. McElheny, S. Parthasarathy, F. Long, M. Hoshi, R. Nussinov and Y. Ishii, *Nat. Struct. Mol. Biol.*, 2015, **22**, 499–505.
21 C. Sachse, M. Fändrick and N. Grigorieff, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 7462–7466.
22 S. J. C. Lee, E. Nam, H. J. Lee, M. G. Savelieff and M. H. Lim, *Chem. Soc. Rev.*, 2017, **46**, 310.
23 M. K. Jana, R. Cappai, C. L. L. Pham and G. D. Ciccotosto, *J. Neurochem.*, 2016, **136**, 594–608.
24 E. J. Alred, M. Phillips, W. M. Berhanu and U. H. E. Hansmann, *Protein Sci.*, 2015, **24**, 923–935.

25 C. Sachse, C. Xu, K. Wieligmann, S. Diekmann, N. Grigorieff and M. Fändrick, *J. Mol. Biol.*, 2006, **362**, 347–354.
26 M. Schmidt, A. Rohou, K. Lasker, J. K. Yadav, C. Schiene-Fisher, M. Fändrick and N. Grigorieff, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, 11858–11863.
27 F. S. Ruggeri, J. Adamcik, J. S. Jeong, H. A. Lashuel, R. Mezzenga and G. Dietler, *Angew. Chem., Int. Ed.*, 2007, **54**, 2462–2466.
28 G. Srinivas, X. Cheng and J. C. Smith, *J. Chem. Theory Comput.*, 2011, **7**, 2539–2548.
29 B. Fan and J. K. Maranas, *Cellulose*, 2015, **22**, 31–44.
30 D. C. Glass, K. Moritsugu, X. Cheng and J. C. Smith, *Biomacromolecules*, 2012, **13**, 2634–2644.
31 J. Wohlert and L. A. Berglund, *J. Chem. Theory Comput.*, 2011, **7**, 753–760.
32 C. A. Lopez, G. Bellesia, A. Redondo, P. Langan, S. P. S. Chundawat, B. E. Dale, S. J. Marrink and S. Gnanakaran, *J. Phys. Chem. B*, 2015, **119**, 465–473.
33 M. M. Tirion, *Phys. Rev. Lett.*, 1996, **77**, 1905.
34 K. Hinsen, A. Thomas and M. J. Field, *Proteins*, 1999, **34**, 369–382.
35 Z. Xu, R. Paparcone and M. J. Buehler, *Biophys. J.*, 2010, **98**, 2053–2063.
36 A. B. Poma, M. Chwastyk and M. Cieplak, *Cellulose*, 2016, **23**, 1573–1591.
37 E. M. Debzi, H. Chanzy, J. Sugiyama, P. Tekely and G. Excoffier, *Macromolecules*, 1991, **24**, 6816–6822.
38 J. F. Matthews, M. E. Himmel and F. Crowley, *Cellulose*, 2012, **19**, 297–306.
39 R. J. Roberts, R. C. Rowe and P. York, *Int. J. Pharm.*, 1994, **105**, 177–180.
40 H. Hertz, *J. Reine Angew. Math.*, 1882, **92**, 156–171.
41 E. Yoffe, *Philos. Mag. A*, 1984, **50**, 813–828.
42 D. Maugis, *Contact, adhesion and rupture of elastic solids*, Springer Science & Business Media, Berlin, 2013.
43 R. Szoszkiewicz, B. Bhushan, B. Huey, A. Kulik and G. Gremaud, *J. Chem. Phys.*, 2005, **122**, 144708.
44 M. Cieplak and M. O. Robbins, *J. Chem. Phys.*, 2010, **132**, 015101.
45 M. Cieplak and M. O. Robbins, *PLoS One*, 2013, **8**, e63640.
46 J. I. Sułkowska and M. Cieplak, *J. Phys.: Condens. Matter*, 2007, **19**, 283201.
47 J. I. Sułkowska and M. Cieplak, *Biophys. J.*, 2008, **95**, 3174–3191.
48 M. Sikora, J. I. Sułkowska and M. Cieplak, *PLoS Comput. Biol.*, 2009, **5**, e1000547.
49 K. Wołek, A. Gomez-Sicilia and M. Cieplak, *J. Chem. Phys.*, 2015, **143**, 243105.
50 J. Tsai, R. Taylor, C. Chothia and M. Gerstein, *J. Mol. Biol.*, 1999, **290**, 253–266.
51 T. Veitshans, D. Klimov and D. Thirumalai, *Folding Des.*, 1997, **2**, 1–22.
52 P. Szymczak and M. Cieplak, *J. Chem. Phys.*, 2006, **125**, 164903.
53 W. A. Steele, *Surf. Sci.*, 1973, **36**, 317–352.
54 R. Wagner, A. Raman and R. Moon, *Cellulose*, 2010, **7**, 27.
55 R. Wagner, R. Moon, J. Pratt, D. Shaw and A. Raman, *Nanotechnology*, 2011, **22**, 455703.
56 Y.-C. Hsieh, H. Yano, M. Nogi and S. Eichhorn, *Cellulose*, 2008, **15**, 507–513.
57 X. Wu, R. J. Moon and A. Martini, *Cellulose*, 2013, **20**, 43–55.
58 J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale and K. Schulten, *J. Comput. Chem.*, 2005, **26**, 1781–1802.
59 O. Guvench, S. S. Mallajosyula, E. P. Raman, E. Hatcher, K. Vanommeslaeghe, T. J. Foster, F. W. JamisonII and A. D. MacKerell Jr, *J. Chem. Theory Comput.*, 2011, **7**, 3162.
60 A. D. MacKerell Jr, D. Bashford, M. Bellott, R. L. Dunbrack Jr, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo and S. Ha, *et al.*, *J. Phys. Chem. B*, 1998, **102**, 3586–3616.
61 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.
62 T. C. F. Gomes and M. S. Skaf, *J. Comput. Chem.*, 2012, **33**, 1338–1346.
63 A. K. Paravastu, R. D. Leapman, W. M. Yau and R. Tycko, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 18349–18354.
64 M. R. Smaoui, F. Poitevin, M. Delarue, P. Koehl, H. Orland and J. Waldisuhl, *Biophys. J.*, 2013, **104**, 683–693.
65 S. E. Feller, Y. Zhang, R. W. Pastor and B. R. Brooks, *J. Chem. Phys.*, 1995, **103**, 4613–4621.
66 R. Paparcone, S. Keten and M. J. Buehler, *J. Biomech.*, 2010, **43**, 1196–1201.
67 G. Yoon, J. Kwak, J. I. Kim, S. Na and K. Eom, *Adv. Funct. Mater.*, 2011, **21**, 3454–3463.
68 S. Iwamoto, W. Kai, A. Isogai and T. Iwata, *Biomacromolecules*, 2009, **10**, 2571–2576.
69 K.-Y. Lee, Y. Aitomki, L. A. Berglund, K. Oksman and A. Bismarck, *Compos. Sci. Technol.*, 2014, **105**, 15–27.
70 K. M. O. Hakansson, A. B. Fall, F. Lundell, S. Yu, C. Krywka, S. V. Roth, G. Santor, M. Kvick, L. P. Wittberg, L. Wagberg and L. D. Soderberg, *Nat. Commun.*, 2014, **5**, 4018.
71 C. P. Brown, C. Harnagea, H. S. Gill, A. J. Price, E. Traversa, S. Licoccia and F. Rosei, *ACS Nano*, 2012, **6**, 1961–1969.
72 M. B. Hinman, J. A. Jones and R. Lewis, *Trends Biotechnol.*, 2000, **18**, 374–379.
73 A. J. Ragauskas, C. K. Williams, B. H. Davison, G. Britovsek, J. Cairney, C. A. Eckert, W. J. Frederick Jr., J. P. Hallet, D. J. Leak, C. L. Liotta, J. R. Mielenz, R. Murphy, R. Templer and T. Tschaplinski, *Science*, 2006, **311**, 484–489.
74 Q. Xu, A. Singh and M. E. Himmel, *Curr. Opin. Biotechnol.*, 2009, **20**, 364–371.
75 J. Adamcik, C. Lara, I. Usov, J. S. Jeong, F. S. Ruggeri, G. Dietler, H. A. Lashuel and R. Mezzenga, *Nanoscale*, 2012, **4**, 4426–4429.
76 J. F. Smith, T. P. Knowles, C. M. Dobson, C. E. MacPhee and M. E. Welland, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 15806–15811.
77 T. P. Knowles and M. J. Buehler, *Nat. Nanotechnol.*, 2011, **6**, 469–479.
78 M. P. Wenger, L. Bozec, M. A. Horton and P. Mesquida, *Biophys. J.*, 2007, **93**, 1255–1263.
79 X. M. Zhang, S. Q. Shi and J. Cao, *Appl. Mech. Mater.*, 2013, **416**, 1726–1730.

