Electronic Supplementary Material (ESI) for RSC Advances.

Structure and Dynamics of Chitin Nanofibrils in Aqueous Environment Revealed by Molecular Dynamics Simulations

Zora Střelcová,[ab] Petr Kulhánek,[ab] Martin Friák,[cd] Helge-Otto Fabritius,[c] Michal Petrov,[c] Jörg Neugebauer,[c] and Jaroslav Koča*[ab]

a) CEITEC - Central European Institute of Technology, Masaryk University, Kamenice 5, 625 00 Brno, Czech Republic
b) National Centre for Biomolecular Research, Faculty of Science, Masaryk University, Kamenice 5, 625 00 Brno, Czech Republic
c) Max-Planck-Institut für Eisenforschung GmbH, Max-Planck-Straße 1, 402 37 Düsseldorf, Germany
d) Institute of Physics of Materials, Academy of Sciences of the Czech Republic, Žižkova 22, 616 00 Brno, Czech Republic

Electronic Supplementary Information (ESI)

Corresponding author: jaroslav.koca@ceitec.muni.cz

Table S1 Composition of simulated systems.

System	Number of Chitin Chains	Number of Chitin Atoms	Number of All Atoms	R_{in} [Å]
1000	1	543	123696	48.0
1110	3	1629	135609	49.5
2200	4	2172	136137	49.5
2220	6	3258	135687	49.5
3300	6	3258	136200	49.5
3330	9	4887	140496	50.0
4400	8	4344	136362	49.5
6000	6	3258	162870	53.0
6600	12	6516	145599	51.0
6760	19	10317	153705	51.9
8998	34	18462	163032	52.7

R_{in} – radius of the largest inscribed sphere into the simulation box

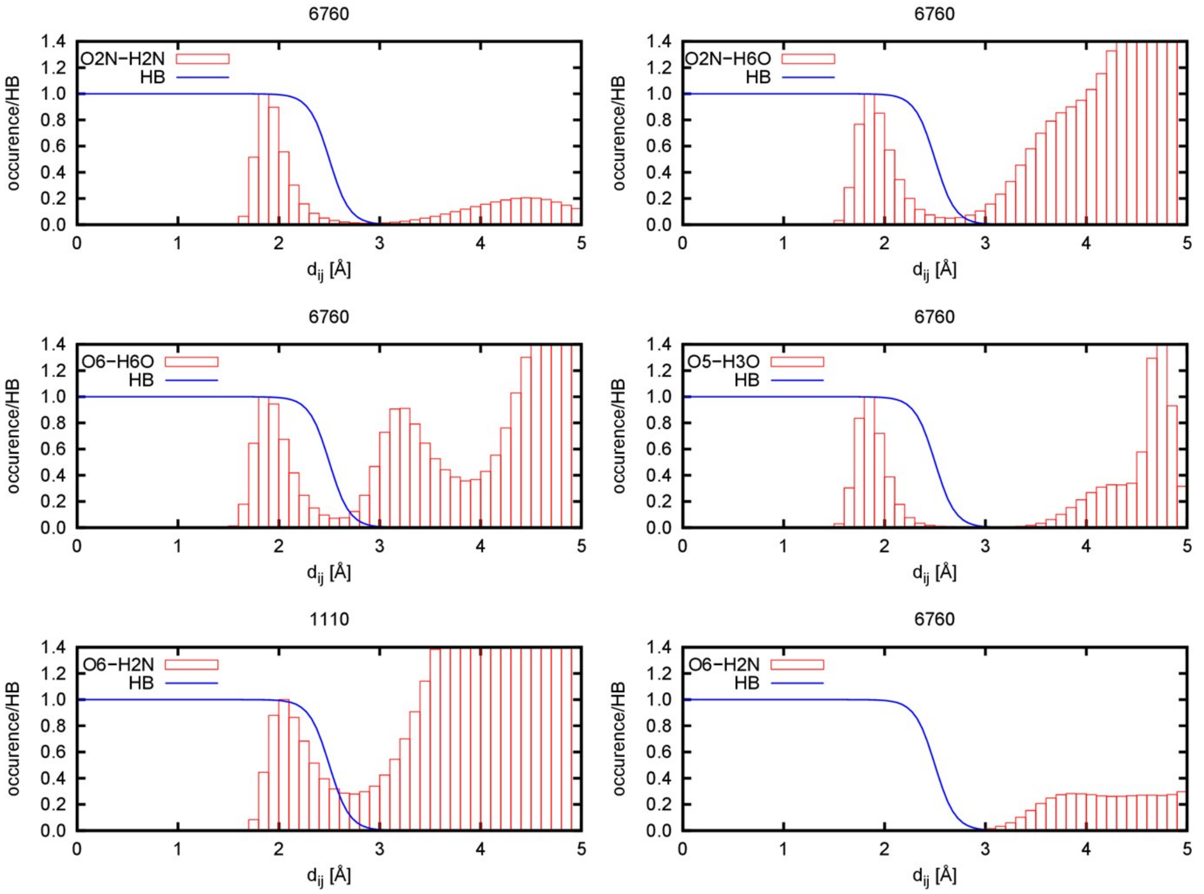

Figure S1 Comparison of the switch function (blue) and histograms (red) of distances of hydrogen bond donor and acceptor atoms. The histograms are arbitrarily scaled and the switch function is shown for one hydrogen bond only. The two top rows represent specific hydrogen bonds, the bottom row represent a selected non-specific hydrogen bond.

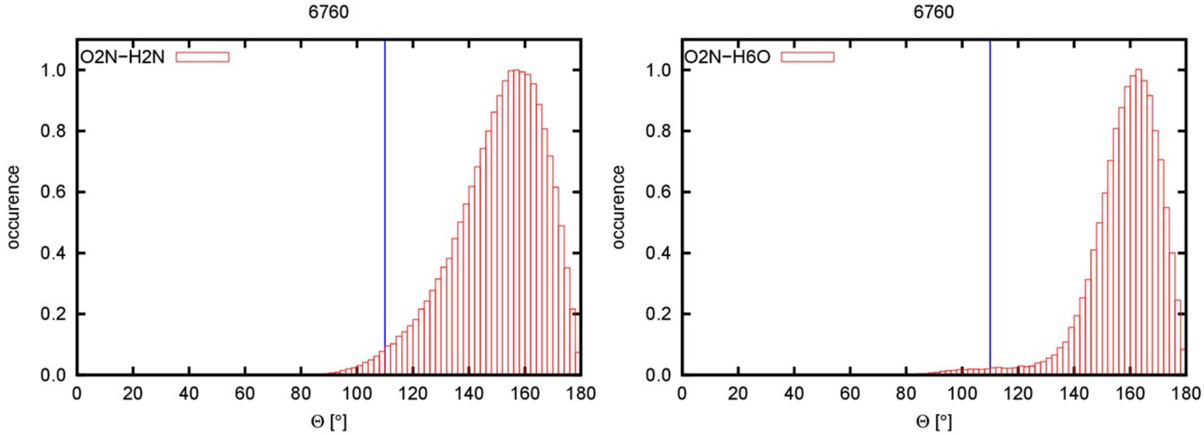

Figure S2 Distribution of the Θ(D-H···A) angle of selected hydrogen bonds for values of HB higher than 0.01 per a single hydrogen bond. The blue vertical line (110˚) is a minimal value recommended by IUPAC.

Table S2 Performance of our <HB> and <HB$_c$> conventional approaches for determination of number of hydrogen bonds. Criteria are the distance between the donor D and acceptor A atoms (please note that in our definition we use distance between the hydrogen H and the acceptor A atoms) and the angle between the donor, hydrogen, and acceptor atoms. RD is relative difference towards <HB>.

	3330					**8998**				
		3.0 Å/135˚		3.5 Å/120˚			3.0 Å/135˚		3.5 Å/120˚	
	<HB>	**<HB$_c$>**	**RD [%]**	**<HB$_c$>**	**RD [%]**	**<HB>**	**<HB$_c$>**	**RD [%]**	**<HB$_c$>**	**RD [%]**
O5-H3O	166.4	150.8	9.4	165.1	0.8	634.0	579.8	8.5	631.2	0.4
O2N-H2N	113.4	86.3	23.9	114.2	-0.7	567.3	431.1	24.0	565.2	0.4
O2N-H6O	28.6	24.2	15.5	29.3	-2.4	192.7	166.5	13.6	196.6	-2.0
O6-H6O	59.0	45.4	23.1	61.0	-3.3	266.3	191.2	28.2	279.4	-4.7

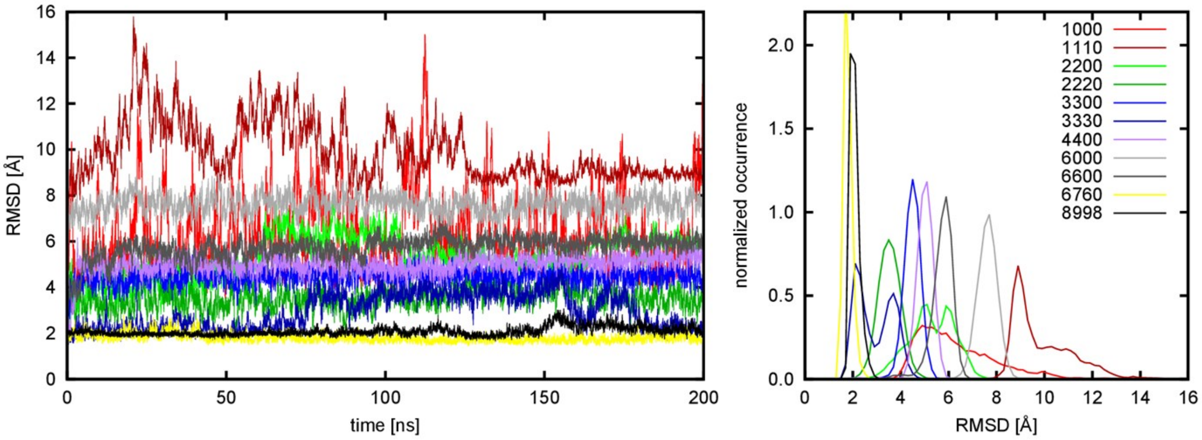

Figure S3 Time evolution of RMSD (left); normalized histograms of RMSD (right).

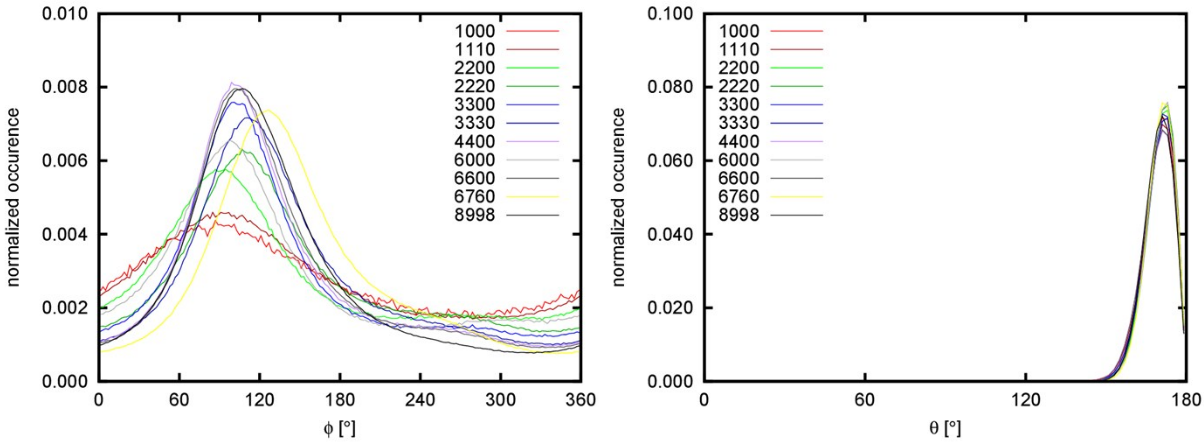

Figure S4 Distributions of the ϕ and θ ring-puckering coordinates on the sugar rings (C1, C2, C3, C4, C5, and O5).

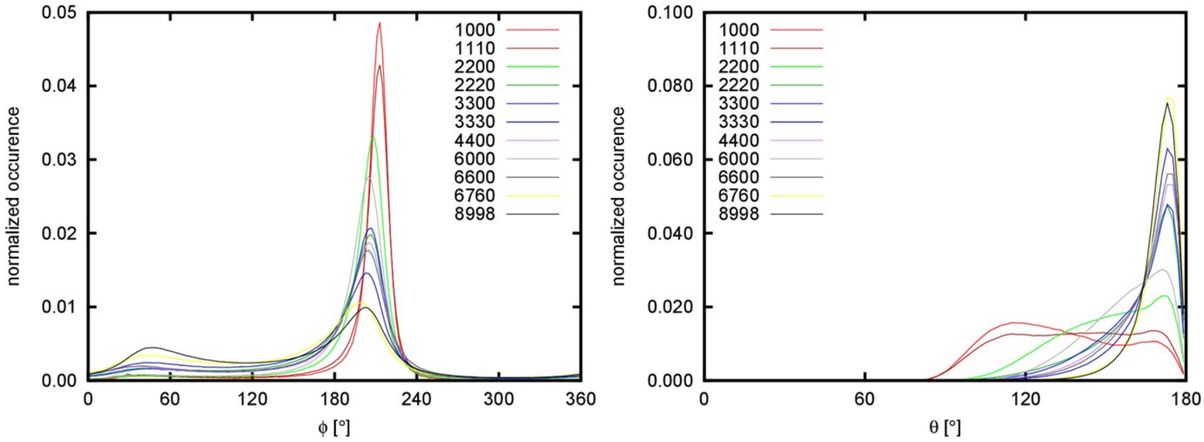

Figure S5 Distributions of the ϕ and θ ring-puckering coordinates on pseudo-rings formed by the $O4^n$, $C4^n$, $C3^n$, $O3^n$, $O5^{n+1}$ and $C1^{n+1}$ atoms. These rings are formed by the H3O-O5 hydrogen bonds. Due to technical restrictions related to ring-puckering coordinates, the H3O atom is excluded from the calculations of the coordinates.

Table S3 Relationship between number of hydrogen bonds (*<HB>*) and number of stacked layers exposed to the solvent (l_w) or the internal area of the fibril (l_i). $<HB>_p$ is predicted number of hydrogen bonds by the model and *RMSE* is the root-mean-square error of the prediction.

			O2N-H6O		**O6-H60**	
system	**l_w**	**l_i**	**<HB>**	**$<HB>_p$**	**<HB>**	**$<HB>_p$**
2200	2	2	9.126	9.91	10.89	14.42
2220	2	4	13.48	17.75	13.48	25.57
3300	4	4	18.66	19.83	31.79	28.85
3330	4	8	28.63	35.50	59.01	51.14
4400	6	6	26.09	29.74	44.27	43.27
6000	10	0	16.19	10.39	12.53	16.37
6600	10	10	45.63	49.57	67.58	72.11
6760	12	20	89.73	90.83	143.50	131.13
8998	16	44	192.7	189.03	266.30	271.47
RMSE				4.01		7.03

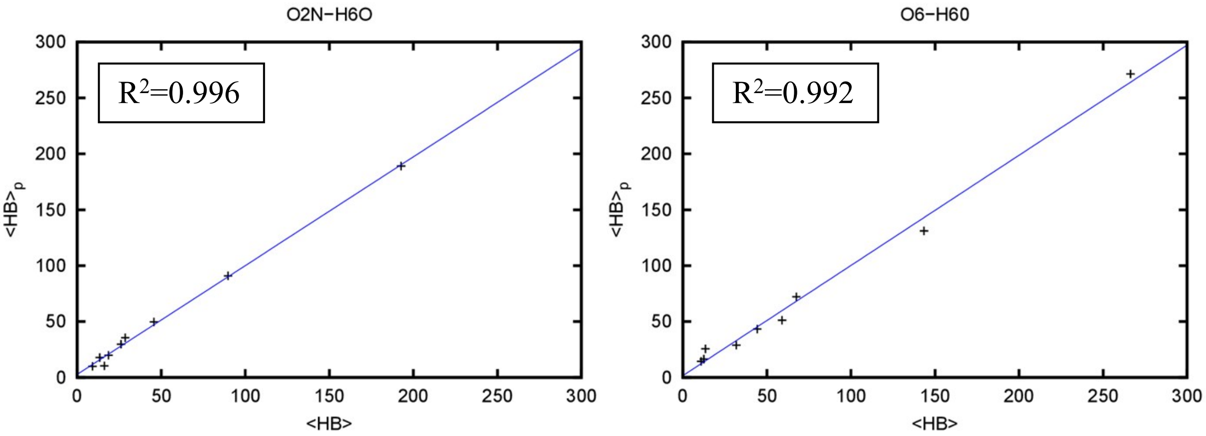

Figure S6 Correlation of the source *<HB>* and predicted $<HB>_p$ number of hydrogen bonds.

Table S4 Statistics of non-specific hydrogen bonds.

		O6-H3O			O6-H2N		
system	**n**	**<HB>**	**σ(HB)**	**<HB>/n**	**<HB>**	**σ(HB)**	**<HB>/n**
1000	1	0.23	0.31	0.23	0.00	0.05	0.00
1110	3	1.93	1.04	0.64	6.55	3.14	**2.18**
2200	4	3.77	1.47	0.94	0.13	0.28	0.03
2220	6	3.73	1.29	0.62	0.09	0.20	0.01
3300	6	4.44	1.39	0.74	0.06	0.16	0.01
3330	9	5.60	1.69	0.62	0.07	0.15	0.01
4400	8	7.85	1.96	0.98	0.12	0.23	0.01
6000	6	4.96	1.65	0.83	0.14	0.25	0.02
6600	12	10.45	2.50	0.87	0.19	0.30	0.02
6760	19	9.89	2.08	0.52	0.08	0.14	0.00
8998	34	13.45	3.64	0.40	0.10	0.15	0.00

		O3-H6O			O3-H2N		
system	**n**	**<HB>**	**σ(HB)**	**<HB>/n**	**<HB>**	**σ(HB)**	**<HB>/n**
1000	1	0.08	0.21	0.08	0.11	0.13	0.11
1110	3	2.74	1.19	0.91	1.09	0.87	0.36
2200	4	0.88	0.76	0.22	0.59	0.68	0.15
2220	6	0.92	0.76	0.15	0.34	0.35	0.06
3300	6	2.09	1.16	0.35	0.73	0.53	0.12
3330	9	1.87	1.16	0.21	0.41	0.32	0.05
4400	8	3.61	1.71	0.45	1.82	1.35	0.23
6000	6	1.60	1.11	0.27	2.04	1.33	0.34
6600	12	4.82	2.01	0.40	2.47	1.13	0.21
6760	19	2.97	1.42	0.16	0.78	0.25	0.04
8998	34	3.95	2.25	0.12	1.07	0.29	0.03

		O2N-H3O		
system	**n**	**<HB>**	**σ(HB)**	**<HB>/n**
1000	1	0.04	0.16	0.04
1110	3	6.08	3.55	**2.03**
2200	4	0.17	0.36	0.04
2220	6	0.30	0.52	0.05
3300	6	0.21	0.39	0.03
3330	9	0.31	0.48	0.03
4400	8	0.26	0.43	0.03
6000	6	0.26	0.44	0.04
6600	12	0.41	0.55	0.03
6760	19	0.46	0.58	0.02
8998	34	0.70	0.73	0.02

n – number of chitin chains; <HB> – average number of hydrogen bonds; σ(HB) – variance (fluctuation) of number of hydrogen bonds; <HB>/n – average value of number of hydrogen bonds related to a single chitin chain

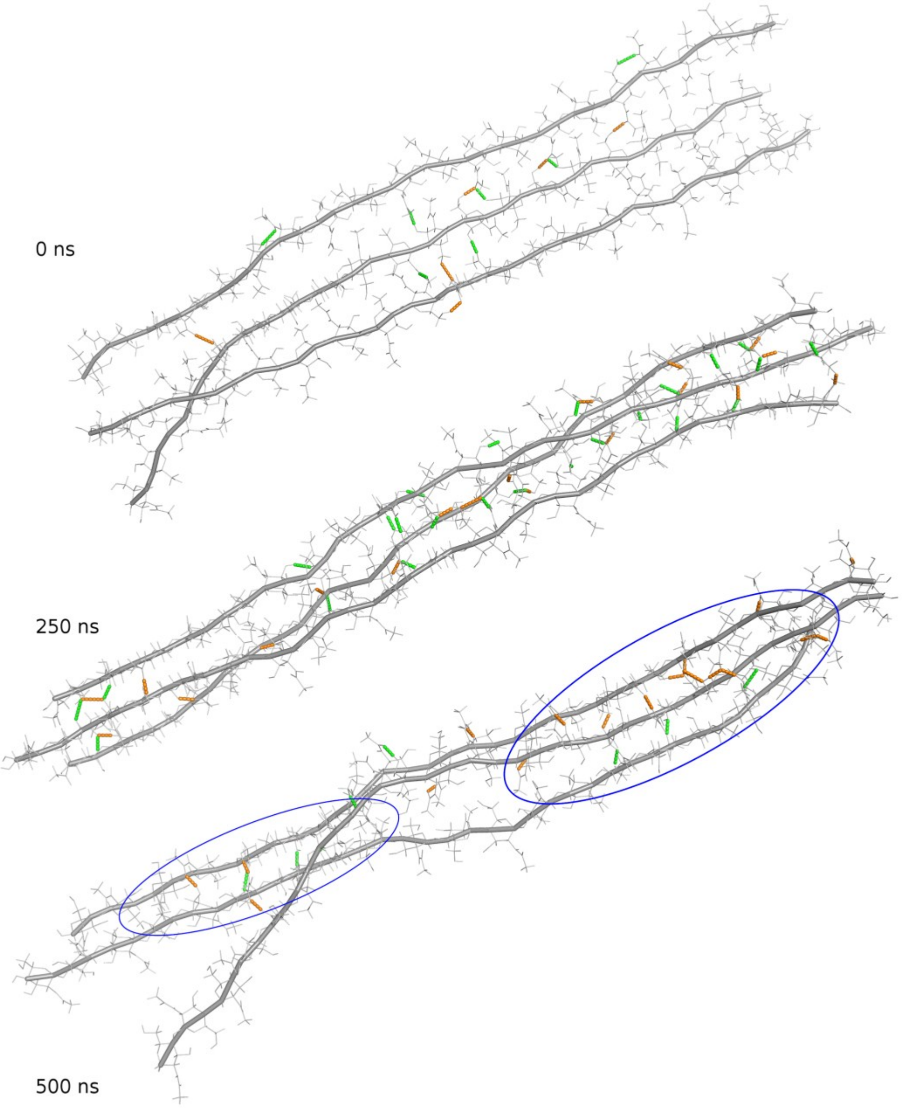

Figure S7 Selected structures of the 1110 system observed at different times during the MD simulations with highlighted hydrogen bonds: specific hydrogen bonds (orange) and non-specific hydrogen bonds (green). Emerging stacked regions are highlighted in ovals. The chain backbones are shown as gray tubes.

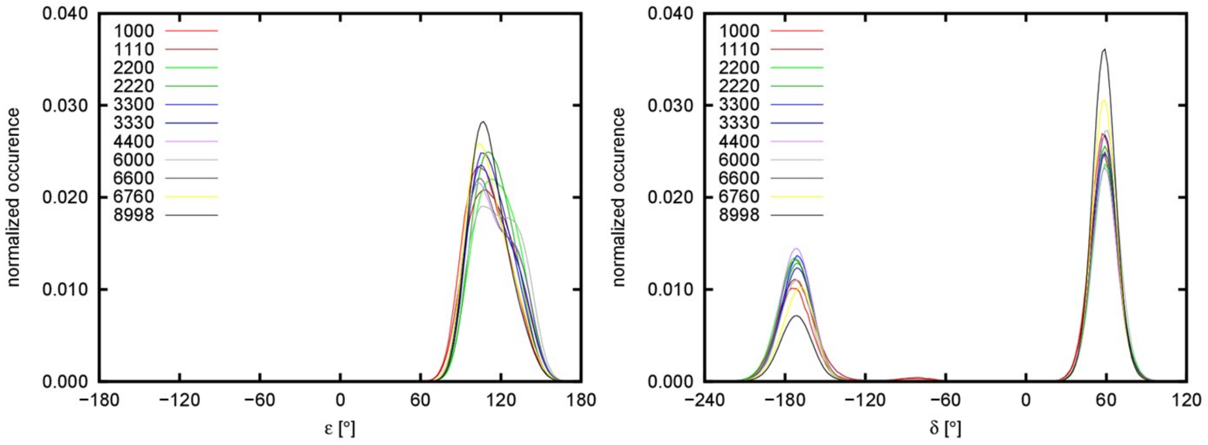

Figure S8 Distributions of dihedral angle ε showing orientation of the N-acetylamine group and dihedral angle δ showing population of rotamers on the C5-C6 bonds.

Table S5 Population of rotamers on the C5-C6 bond.

	rotamer			
system	**-171˚**	**-83˚**	**60˚**	**ratio(-171˚/60˚)**
1000	0.33	0.01	0.66	0.49
1110	0.38	0.02	0.61	0.63
2200	0.42	0.00	0.57	0.74
2220	0.39	0.00	0.61	0.64
3300	0.41	0.00	0.59	0.69
3330	0.37	0.00	0.63	0.60
4400	0.44	0.01	0.56	0.79
6000	0.34	0.00	0.66	0.51
6600	0.40	0.01	0.59	0.68
6760	0.31	0.00	0.69	0.46
8998	0.21	0.00	0.78	0.27

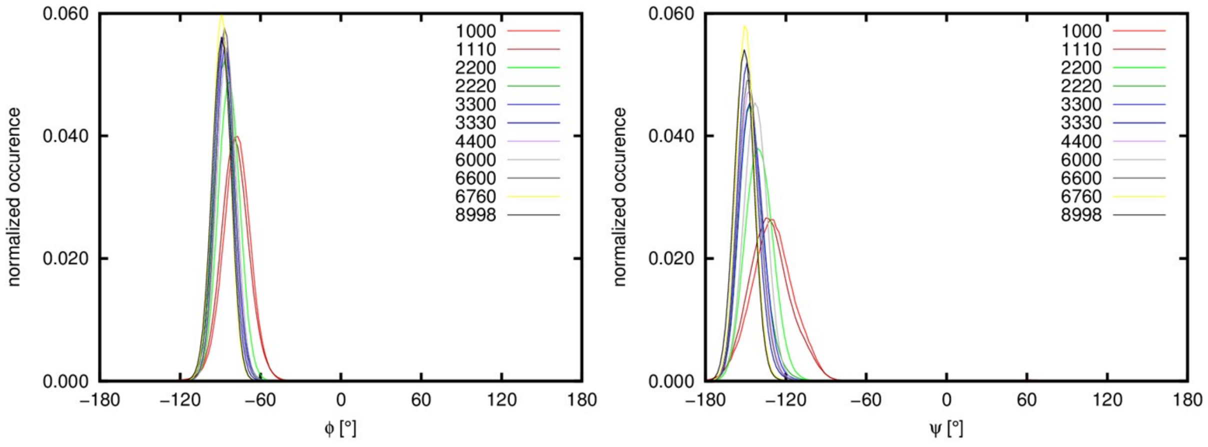

Figure S9 Distributions of glycosidic bond dihedral angles ϕ and ψ.

Parameters that do not correlate with the pseudo-dihedral angle α

The distribution of dihedral angle ε exhibits a Gaussian distribution for the majority of the systems, but for some of them it is split into two strongly overlapping distributions (this is most noticeable for the 6000 system) (Figure S7 left). This splitting indicates that there are at least two very close rotamer states. One of them appears to be the result of a lack of stabilization between adjacent stacked layers (in the *b* direction), which is most obvious with the 6000 system. The average values range from 111° to 117° and they only poorly correlate with the pseudo-dihedral angle α, with an R^2 of 0.85. Since the acetamide group is involved in the formation of hydrogen bonds formed by the O2N-H2N and O2N-H6O atom pairs, we also tested the correlation of their average values with the pseudo-dihedral angle. Their correlation was even worse, with R^2 of 0.7 and 0.6, respectively.

The distribution of the dihedral angle δ shows two distinct rotamer states (Figure S7 right). The most populated state has its maximum at 60°, while the second state has its maximum at about -171°. In a few systems, a third rotamer state appears at about -83°, but with a very low population. With varying the size of the nanofibrils, the ratio between both states varies (Table S4) from nearly 1:1 (2200 and 4400) to 1:3 (8998). The ratio does not correlate well with the pseudo-dihedral angle (R^2 is 0.4). Unexpectedly, this is also the case with the amount of hydrogen bonds between the O2N-H6O and O6-H60 atom pairs. In the latter case, one could expect a correlation between the orientation of the CH$_2$-OH and the number of hydrogen bonds that include acceptor and donor atoms from this group. The most likely reason for this observation is that these hydrogen bonds are weaker and strongly influenced by the presence of the solvent.

Table S6 Correlation between pseudo-dihedral angle α and selected structural descriptors.

			<HB>/n or <HB>/l			**C5-C6**			**pseudo-ring**
system	**<α>**	**<ε>**	**O2N-H2N**	**O2N-H6O**	**O5-H3O**	**rot. ratio**	**<ψ>**	**<ϕ>**	**<θ>**
2200	236.2	116.6	18.1	0.5	17.9	0.741	-139.7	-83.4	150.7
2220	100.3	114.8	18.7	1.1	18.3	0.637	-145.0	-86.8	160.9
3300	117.7	113.9	18.2	1.0	18.5	0.685	-145.5	-86.2	161.7
3330	59.2	113.2	18.9	1.9	18.5	0.595	-148.0	-88.0	166.1
4400	112.4	114.3	18.1	1.0	18.4	0.795	-146.4	-86.2	163.8
6000	169.4	117.4	18.0	0.7	18.2	0.505	-142.7	-85.3	157.0
6600	99.9	114.5	18.2	1.1	18.4	0.681	-146.9	-86.6	164.8
6760	15.4	111.3	19.3	7.2	18.6	0.455	-149.9	-89.1	169.3
8998	11.3	111.5	18.9	9.8	18.6	0.273	-150.9	-89.2	169.1
R^2		0.854	0.700	0.588	**0.940**	0.398	**0.970**	**0.989**	**0.952**

<α> – average value of pseudo-dihedral angle showing axial twist distortion; <ε> – average value of dihedral angle showing orientation of the N-acetylamine group; <HB> – average number of hydrogen bonds; n – number of chitin chains; l – number of stacked layers; rot. ratio – ratio between -171°/60° rotamers on C5-C6 bond; <ψ> – average value of glycosidic bond dihedral angle; <ϕ> – average value of glycosidic bond dihedral angle; <θ> – average value of ring-puckering coordinate on pseudo-rings formed by the O4^n, C4^n, C3^n, O3^n, O5^{n+1} and C1^{n+1} atoms; R^2 – Pearson correlation coefficient between given property and <α>.

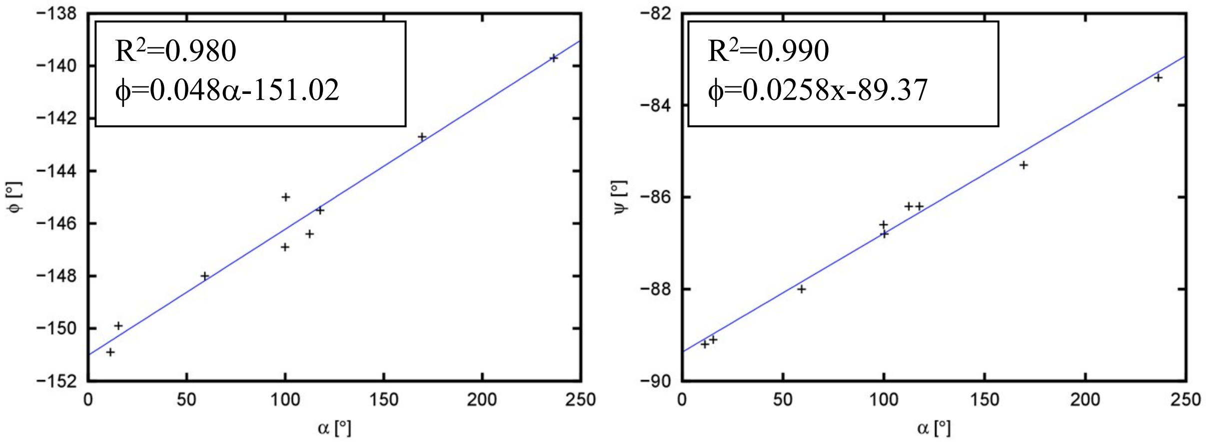

Figure S10 Correlation between glycosidic bond dihedral angles ϕ and ψ and the axial chirality dihedral angle α.

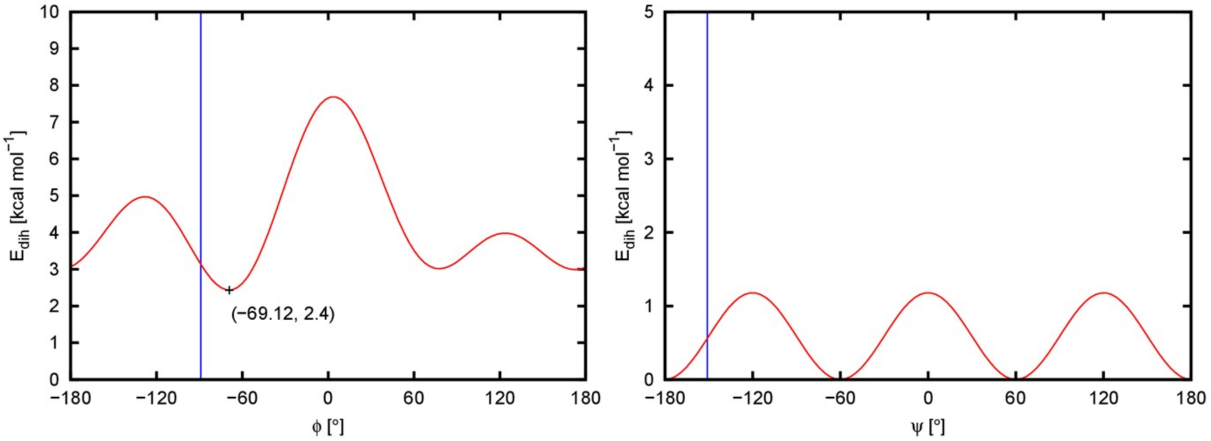

Figure S11 Contributions to the potential energy caused by rotation on glycosidic bonds described by dihedral angles ϕ and ψ. Each potential is composed from three contributions: ϕ [C4^n, O4^n, C1^{n+1}, O5^{n+1} (+0.0); C4^n, O4^n, C1^{n+1}, H1^{n+1} (+120.0); C4^n, O4^n, C1^{n+1}, C2^{n+1} (-120.0)] and ψ [C5^n, C4^n, O4^n, C1^{n+1} (+0.0); H4^n, C4^n, O4^n, C1^{n+1} (+120.0); C3^n, C4^n, O4^n, C1^{n+1} (-120.0)]. Values in brackets are dihedral angle offsets as results of sp^3 configuration of C1 and C4 carbon atoms. The vertical lines represent extrapolated values of ϕ and ψ for zero value of axial chirality dihedral angle α.

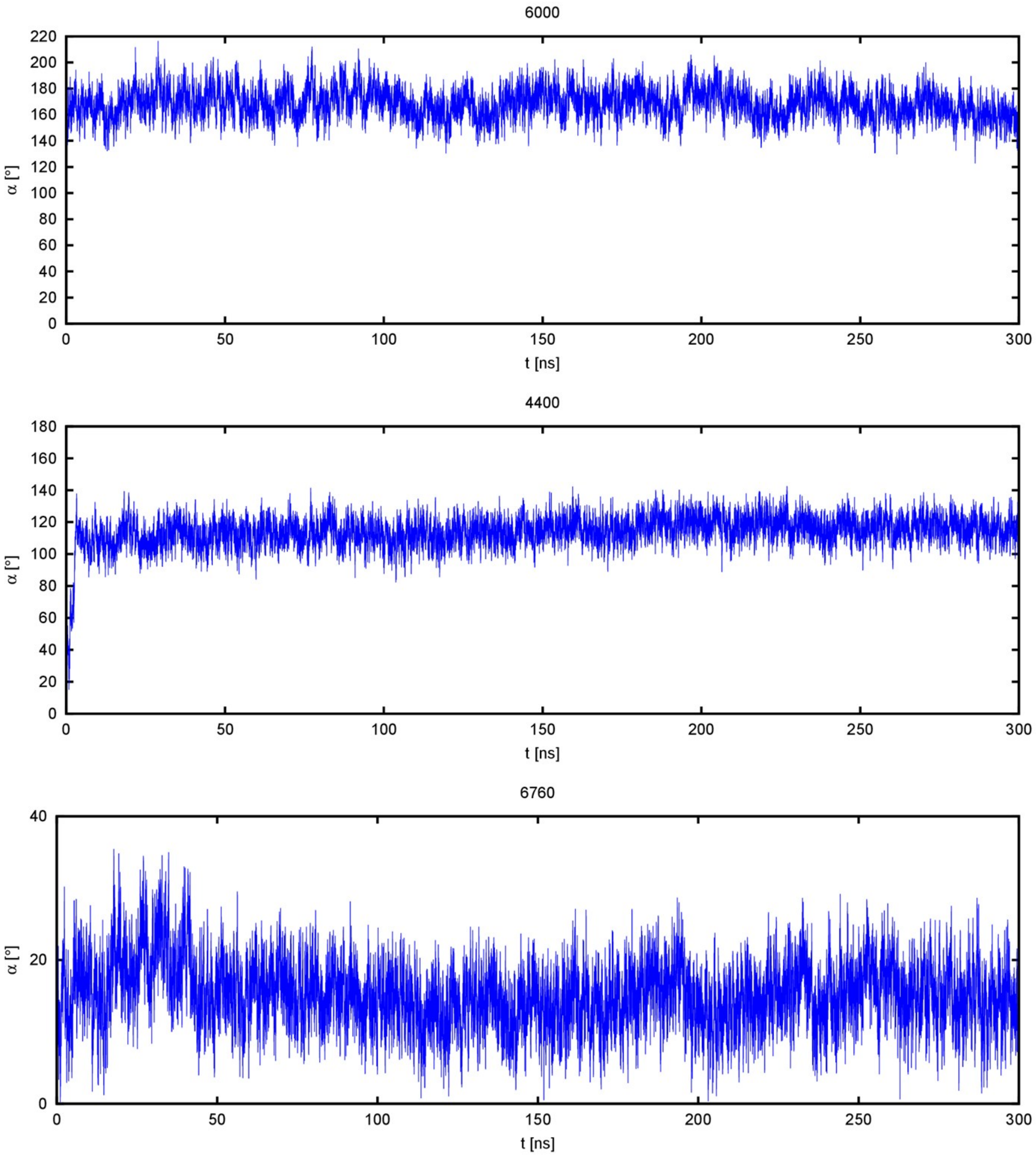

Figure S12 Time evolution of the twist angle α for the three selected systems.

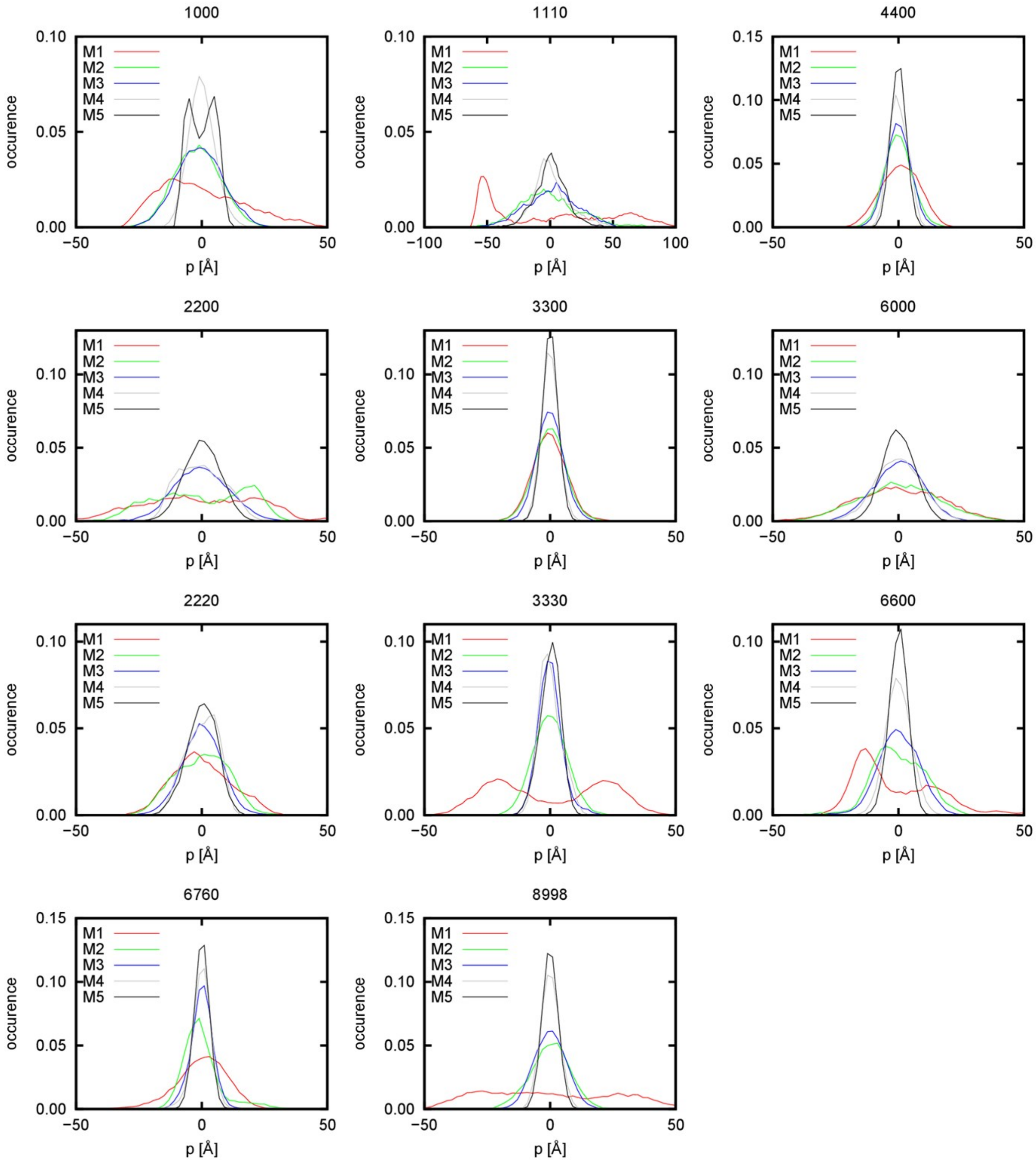

Figure S13 Distribution of trajectory snapshot projections (p) on the five largest essential modes (M1-M5).

Table S7 Energy contributions to the reaction enthalpy related per a single chitin chain.

system	**$\langle \Delta E_{int,r} \rangle /n$**	**$\langle \Delta E_{ele,r} \rangle /n$**	**$\langle \Delta E_{vdw,r} \rangle /n$**	**$\Delta H_r/n$**
1110	2.4	-77.6	-80.7	-155.8
2200	-4.1	-120.2	-116.8	-241.1
2220	-5.1	-147.9	-139.4	-292.4
3300	-4.0	-177.0	-172.3	-353.2
3330	-5.0	-187.4	-187.5	-380.0
4400	-2.9	-194.9	-193.4	-391.2
6000	-2.1	-182.5	-158.2	-342.8
6600	-1.0	-215.2	-213.7	-429.9
6760	-5.3	-238.5	-234.7	-478.6
8998	-4.7	-256.6	-257.2	-518.4

$\Delta E_{int,r}$ -.change of internal energy (bonds, angles, dihedral angles); $\Delta E_{ele,r}$ – change of electrostatic energy; $\Delta E_{vdw,r}$ – change of van der Waals energy, ΔH_r – total reaction enthalpy; n – number of chitin chains in the nanofibril; all energies in kcal mol^{-1}

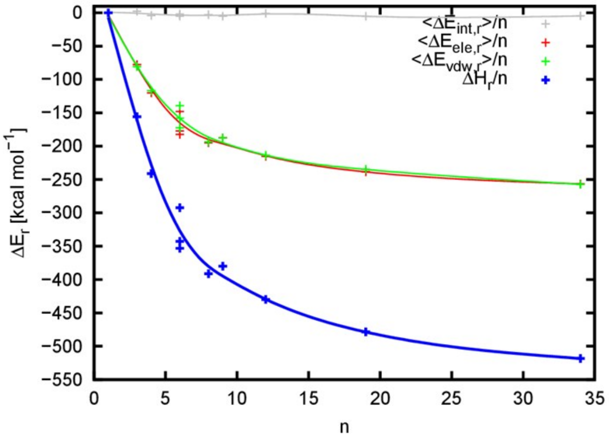

Figure S14 Dependence of reaction enthalpy related per a single chitin chain and its contributions on number of chitin chains *n* in the nanofibril. Lines are meant to only highlight progress of individual energy contributions.

Table S8 Change in reaction enthalpy ΔH_r, electrostatic $\Delta G_{el,r}$ and non-electrostatic $\Delta G_{ne,r}$ contributions of solvation free reaction energies and total free reaction energy ΔG_r related to a single chitin chain without the entropy contribution.

			MM/PBSA		
system	**n**	**ΔH_r/n**	**<$\Delta G_{el,r}$>/n**	**<$\Delta G_{ne,r}$>/n**	**<ΔG_r>/n**
1110	3	-155.8	110.2	-9.3	-54.9
2200	4	-241.1	156.6	-12.6	-97.1
2220	6	-292.4	190.1	-15.1	-117.4
3300	6	-353.2	228.3	-16.6	-141.5
3330	9	-380.0	247.2	-18.2	-150.9
4400	8	-391.2	256.5	-18.1	-152.7
6000	6	-342.8	225.7	-15.3	-132.3
6600	12	-429.9	284.0	-19.4	-165.4
6760	19	-478.6	312.4	-21.2	-187.4
8998	34	-518.4	336.3	-22.5	-204.7

n – number of chains in the nanofibril; all energies in kcal mol^{-1}

Table S9 Change of reaction translational and rotational entropy $\Delta S_{tr,r}/n$, vibrational entropy $\Delta S_{v,r}/n$, total reaction entropy $\Delta S_r/n$, and associated energy $-T\Delta S_r/n$ related per a single chitin chain.

system	n	$\Delta S_{tr,r}/n$	$\Delta S_{v,r}/n$	$\Delta S_r/n$	$-T\Delta S_r/n$
1000	1	0	0	0.0	0.0
1110	3	-66.1	-14.4	-80.5	24.2
2200	4	-74.9	-77.2	-152.1	45.6
2220	6	-83.8	-116.8	-200.6	60.2
3300	6	-83.9	-144.1	-227.9	68.4
3330	9	-89.9	-145.5	-235.4	70.6
4400	8	-88.4	-147.2	-235.6	70.7
6000	6	-83.8	-122.0	-205.8	61.7
6600	12	-93.0	-148.6	-241.7	72.5
6760	19	-96.5	-159.2	-255.8	76.7
8998	34	-99.3	-166.5	-265.8	79.7

n – number of chitin chains in the nanofibril; all entropy in cal mol^{-1} K^{-1}; all energy in kcal mol^{-1}.

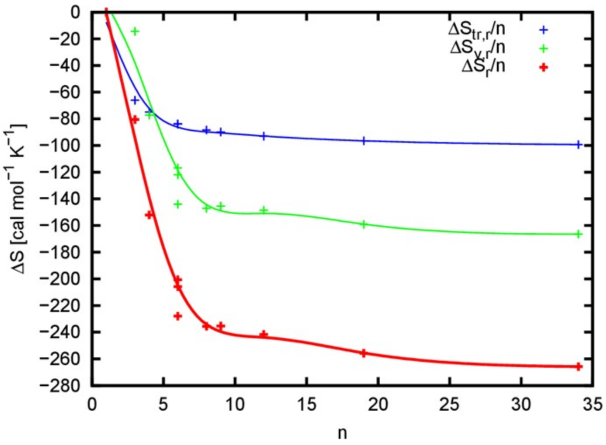

Figure S15 Dependence of change of reaction translational and rotational entropy $\Delta S_{tr,r}/n$, vibrational entropy $\Delta S_{v,r}/n$, total reaction entropy $\Delta S_r/n$ related per a single chitin chain on number of chitin chains *n* in the nanofibril. Lines are meant to only highlight progress of individual entropy contributions.