Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Supporting Information

Mineralization of phosphorylated cellulose: Crucial role of surface structure and monovalent ions for optimizing calcium content

N. V. Lukasheva[a], D. A. Tolmachev[a], and Mikko Karttunen[a,b,c]

[a]Institute of Macromolecular Compounds, Russian Academy of Sciences, Bolshoj pr. V.O., 31, 199004 St. Petersburg, Russia

[b] Department of Chemistry, The University of Western Ontario, 1151 Richmond St, London, Ontario, Canada, N6A 5B7

[c] Department of Applied Mathematics, The University of Western Ontario, 1151 Richmond St, London, Ontario, Canada, N6A 5B7

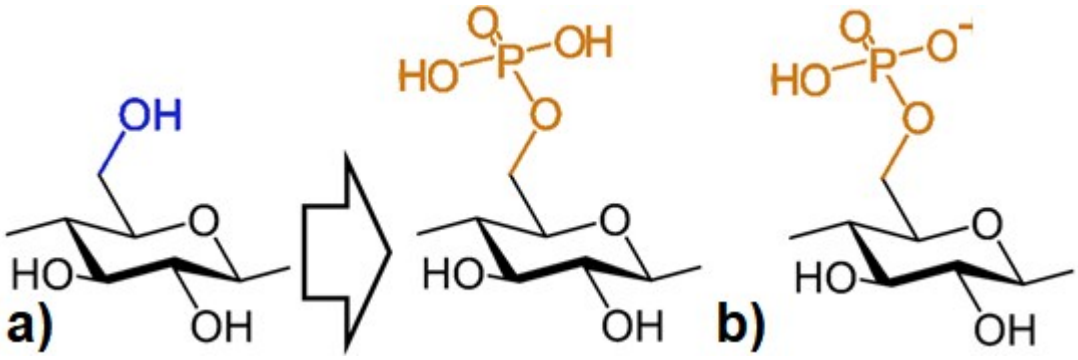

Figure S1. Illustration of cellulose phosphorylation (a). Primary hydroxyl groups (blue) on cellulose fibril surface are substituted by phosphates groups (orange). Deprotonated state of phosphate group in water environment (b).

Computational details.

For a correct description of ionized molecules in quantum-chemical simulation, the basic set should include a diffuse function[1]. In the CSFF force field[2] used in the simulations in this paper, the partial charges

were obtained from quantum chemical calculations using the basis set HF/6-31G (d) with the Mulliken charge distribution. A series of quantum chemical calculations using the Hartree-Fock method with a variety of basis sets (6-31G (d, p), 6-31G (d), 6 -31 + G (d), 6-31 + G (d, p)) was performed to determine the best basis set for calculating the partial charges of the phosphate ions and phosphorylated cellulose molecules. To test consistency, we also used Löwdin population analysis[3] as discussed below.

Calculations were carried out for the deprotonated phosphate ion and deprotonated cellulose oligomers consisting of two glucopyranose residues. Calculations were made for two cellulose oligomers differing in their arrangements of glucopyranose residues with substituted and unsubstituted primary hydroxyl groups. The values of the charges on the atoms of the "central regions" (Figure S2) of the oligomers were used to model polymer molecule.

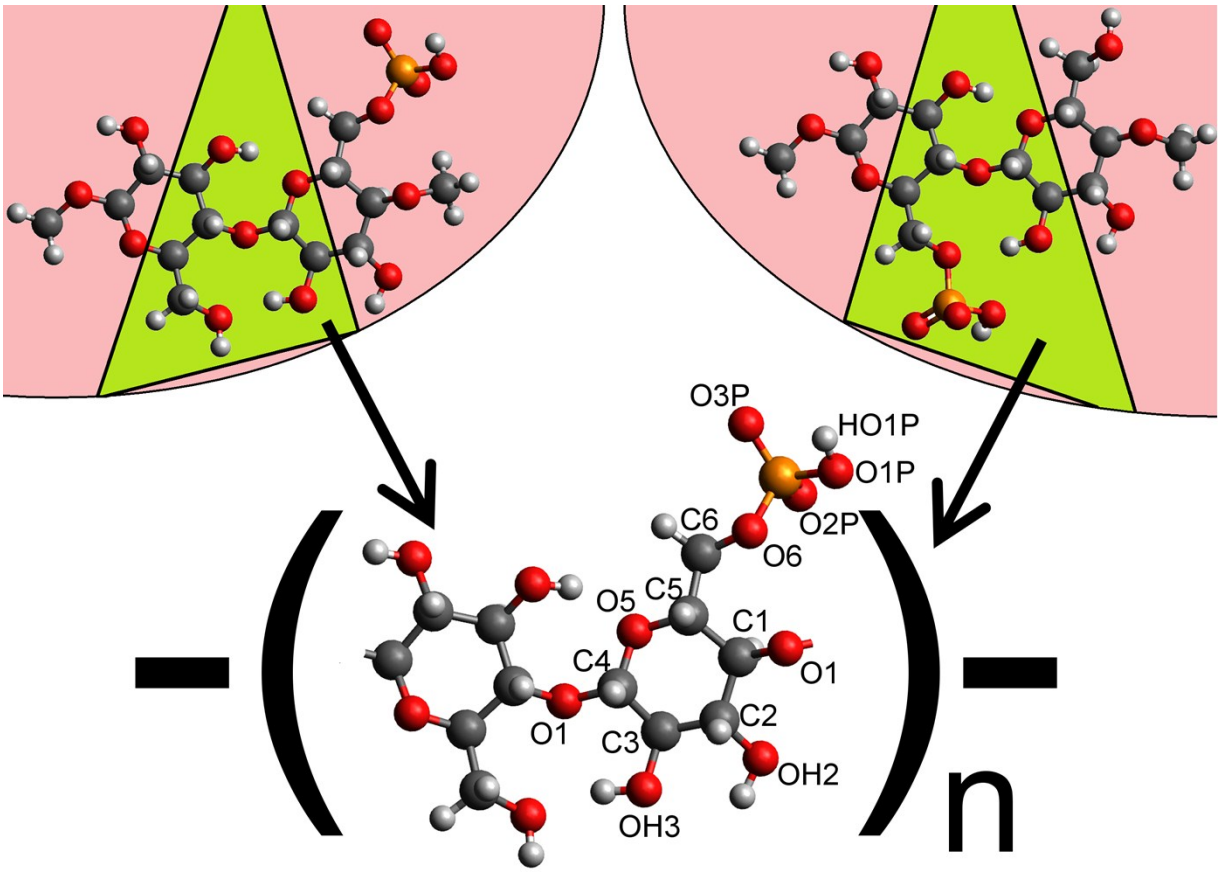

Figure S2. Scheme for the obtaining the distribution of charges on atoms on cellulose monomer.

To investigate the effect of the distribution of the partial charges on the phosphate ion upon its interaction with water molecules, a series of MD simulations of the phosphate ion in an aqueous environment with charge distributions obtained from quantum chemical calculations with different basis sets was carried

out. The simulation cell contained one H_2PO_4 ion, Na^+ ion and 4,136 water molecules. The simulations were carried out in the NPT ensemble at a temperature of 300 K and a pressure of 1 bar for 10 ns in steps of 0.001 ps. To maintain constant temperature, the Nosé-Hoover thermostat[4,5] with a time constant $\tau_T = 0.2$ ps was used. For constant pressure, the Parrinello-Rahman barostat[5] with a time constant of $\tau_R = 0.5$ ps was used. The long-range Coulomb interactions were calculated using the particle-mesh Ewald summation procedure[7]. To describe water molecules, the TIPs3P[8] water model was used. The water molecules were kept rigid using the SETTLE procedure[9]. The van der Waals interactions, as well as the valence interactions of the phosphate ion, were described by using the CHARMM27[10] parameters. The *ab initio* (Hartree-Fock) calculations were performed using the Firefly software package[11,12]. MD simulations were carried out with the Gromacs 4.5.6[13] software package. Two criteria were used to guide the selection of the partial charges: 1) experimental data for solvation of phosphate ions[14] (12 water molecules in the first coordination shell) and 2) the fact that phosphorylation does not affect the crystal structure of cellulose fibril[15].

Calculations of the partial charges

The results from the quantum chemical calculations are presented in Table SI.

Table SI. Partial charges obtained by using different basic sets

	6-31G(d)	6-31+G(d)	6-31G(d,p)	6-31+G(d,p)
P	1.469	2.336	1.477	2.293
O(H)	-0.846	-1.090	-0.742	-0.937
O(H)	-0.846	-1.090	-0.742	-0.937
O	-0.828	-1.091	-0.829	-1.079
O	-0.828	-1.091	-0.828	-1.079
H	0.440	0.513	0.332	0.370
H	0.440	0.513	0.332	0.370

The addition of the diffuse function to the basis set leads to an increase in the values of the partial charges on all atoms of the phosphate ion (in particular more than by 50% for the phosphorus atom). The addition of the polarization function for hydrogen atoms leads to a decrease in the partial charge on the hydrogen atoms and on the nearest oxygen atoms.

The results of the series of the MD simulations of the ion in water are presented in Figure S3. Figure S3a shows the degree of hydration of the phosphate ion and Figure S3b shows the potential of mean force for the interactions between water molecules and the phosphate ion.

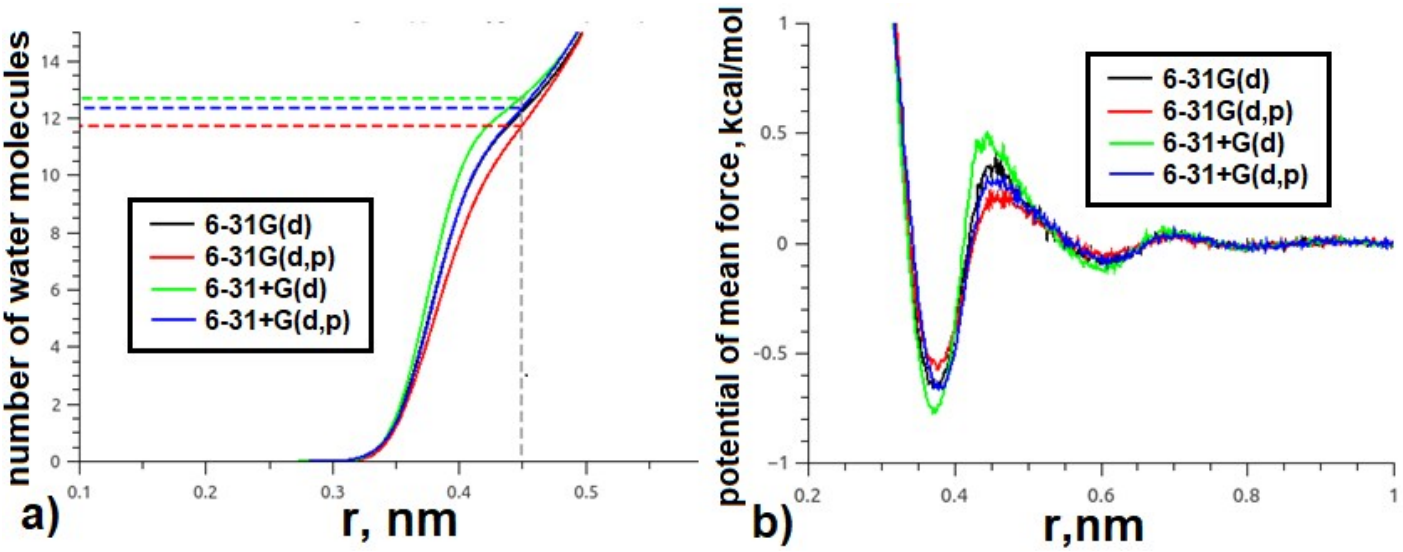

Figure S3. (a) The number of water molecules near the phosphate ion and (b) the potential of mean force between the phosphate ion and water, for simulations with the partial charges on the atoms of the phosphate ion obtained with different basis sets.

As Figure S3a shows, the difference in the average number of water molecules in the first coordination shell of the phosphate ion does not exceed one molecule. On average, there are 12 water molecules. This result agrees with the experimental data from neutron diffraction[10]. On the other hand, the differences in the ion/water contact energy as well as in the energy barriers that need to be overcome in order to break this contact are less than the thermal energy (k_BT). Since the obtained data are not sensitive to the choice of the basis set, these characteristics cannot be used as criteria for the correct choice of the partial charge distribution.

The results of the calculations of the partial atomic charges on the phosphorylated cellulose molecule are presented in Table IIS.

Table SII. The values of the partial charges on atoms of the hydroxyl groups involved in the formation of intrafibrillar hydrogen bonds and for phosphate groups of the phosphorylated and native cellulose obtained using different basic sets.

	6-31G(d)	6-31+G(d)	6-31G(d,p)	6-31+G(d,p)	CSFF
Native cellulose					
OH2	-0.757	-0.755	-0.661	-0.608	-0.600
OH3	-0.795	-0.811	-0.688	-0.644	-0.600
OH6	-0.739	-0.813	-0.643	-0.665	-0.660
Phosphorylated cellulose					
OH2	-0.762	-0.766	-0.664	-0.616	-0.600
OH3	-0.800	-0.805	-0.693	-0.626	-0.600
OH6	-0.756	-0.811	-0.656	-0.686	-0.660
O6	-0.794	-0.851	-0.805	-0.985	
P	1.535	2.483	1.543	2.467	
O1P	-0.853	-1.187	-0.746	-1.005	
O2P	-0.805	-1.070	-0.807	-1.061	
O3P	-0.801	-1.021	-0.801	-1.026	
HO1P	0.454	0.554	0.346	0.395	
O1			-0.6	-0.42	-0.31
O5			-0.6	-0.46	-0.36

For further use in MD modeling of phosphorylated cellulose fibrils, the partial atomic charges obtained with the basis set 6-31+G (d, p) are selected, because the charges on the oxygen atoms of hydroxyl groups participating in intrafibrillar intermolecular hydrogen bonds are closest to the charges used in the CSFF force field.

System equilibration

We have judged equilibration by monitoring the time-evolution of the RDFs and the values of their maxima. Example in the case of 12% phosphorylation is shown in Figure S4. The interval for data collection and averaging was then chosen to be 100−200 ns.

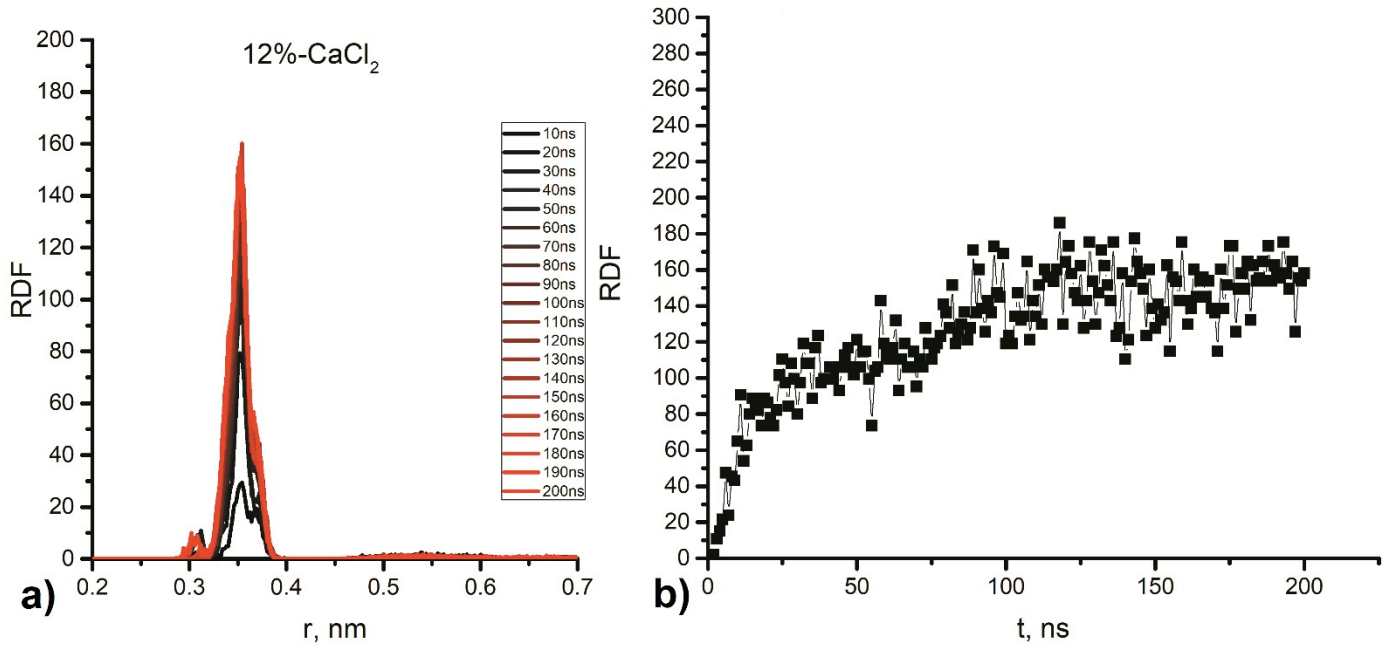

Figure S4. The time-evolution of the radial distribution functions for Ca^{2+} ions at phosphate groups (a) and the evolution of the values of their maxima in time (b).

The length scales characterizing electrostatic interactions in the simulated systems.

The lateral separation for counterions is given by $a\perp \sim\sqrt{q/\sigma s}$, where q is charge and σ_s is the surface charge density. Using σ_s =0.2 e/nm^2 and 0.4 e/nm^2 for 12% and 25%, respectively, the a⊥ values are 2.24 nm and 1.83 nm.

Gouy-Chapman length $\mu=\frac{1}{2\pi q l_B \sigma s}$, where l_B is the so-called Bjerrum length, which measures the distance at which two elementary charges interact with thermal energy k_BT (in water and at room temperature $l_b \approx 0.71$nm). The Gouy-Chapman length values are 1.14 and 0.57 nm for 12% and 25% cases respectively.

Coupling parameter: $\Xi=l_b/\mu$, Ξ = 0.63 (12%) and 1.24 (25%).

Comparison of the distance between the two charged surfaces in the simulated systems (the thickness of the water layer d=7.7 nm) shows that **μ<a⊥<d**.

This allows[15] to consider the simulated system decoupled into two subsystems each consisting of a charged surface and corresponding counterions. Thus, ion adsorption at the two surfaces can be considered independently and each adsorption layer can be described separately.

Conformations of cellulose molecules

To study the effects of phosphorylation on cellulose conformations, the distributions of the dihedral angles φ, ψ, χ, τ_6 and ω (Figure S5) in the surface layer were measured.

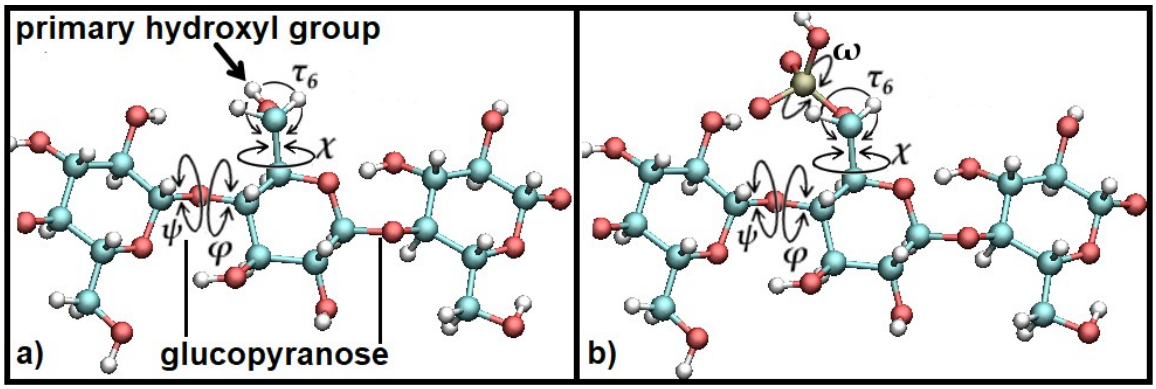

Figure S5. Fragments of native (a) and phosphorylated (b) cellulose molecules. Color code: cyan: carbons, red: oxygens, white: hydrogens, yellow: phosphates. Dihedral angles φ, ψ, χ, τ_6 and ω are indicated.

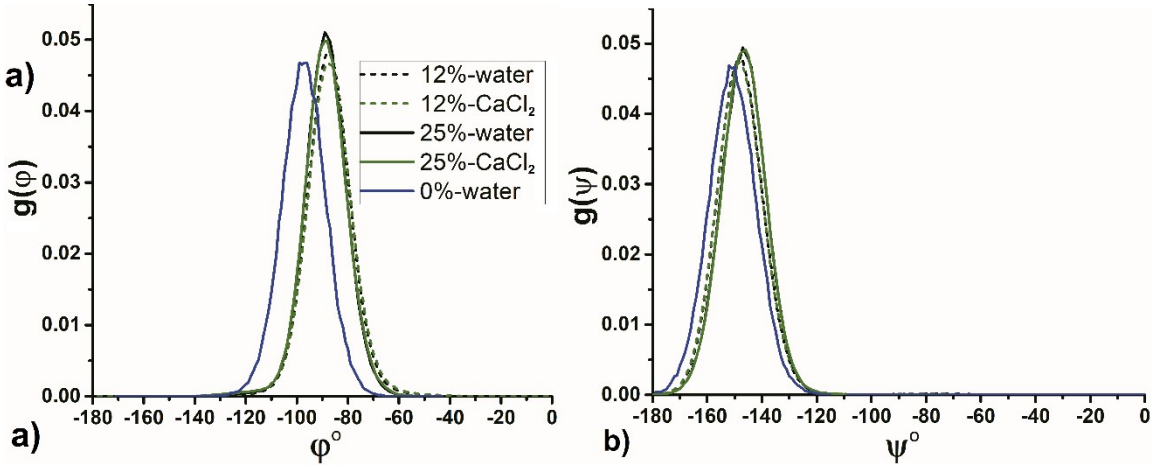

Figure S7. Dihedral angle distributions, for φ (a) and ψ (b). See Figure 2 for definitions. Native[24] (0%, blue solid line) and modified fibrils at 12% and 25% phosphorylation in water (black solid and dashed lines, respectively), and in $CaCl_2$ salt solution (green solid and dashed lines, respectively).

The distributions of the dihedral angles φ and ψ of the main chain (see Figure S6) are shown in Figure S6. The figure shows that upon phosphorylation, the maxima for both φ and ψ shift toward less negative angles and that the shifts for φ are larger; φ defines the orientation of the glucopyranose cycle with the phosphate group. The differences between 12% and 25% phosphorylation are not significant and virtually independent of $CaCl_2$ concentration.

The distributions for the dihedral angle of the internal rotation χ that determine the conformation of the primary hydroxyl (see the definition in Figure S3) and the $-HPO_3^-$ groups are presented in Figure S7.

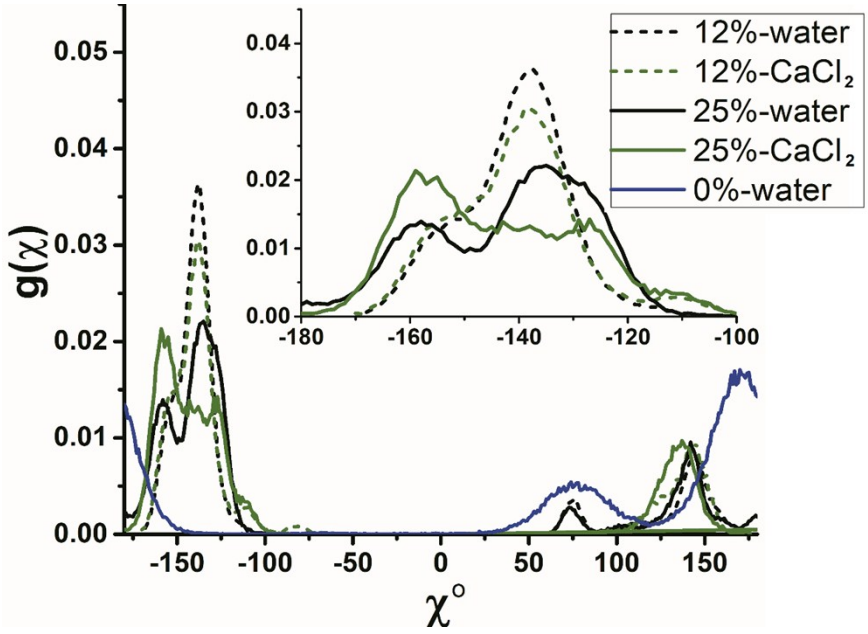

Figure S7. Distributions for the dihedral angle of the internal rotation χ for native[24] (solid blue line) and phosphorylated fibrils in water (12%: black solid and 25%: dashed lines) and in salt solution (green solid and dashed lines, respectively). The inset shows the region -180 – -100°.

Figure S7 shows that the distribution of the angle χ (see Figure S5) for the phosphorylated fibril differs significantly from that of the native one. Not only the positions and the number of the maxima are different, but also the widths. There are two domains of allowed values. The position of the more intense maximum corresponds to the most frequently visited orientation of the C6-O6 bond vector parallel to the long axis of the

molecule (which means the formation of intramolecular H-bonds). In water at both 12% and 25% phosphorylation and along with the two weak maxima (one coinciding with the peak for native cellulose and the other being slightly shifted), a new pronounced peak at -140º appears. Importantly, this corresponds to the orientation of the phosphate group being able to form an H-bond with the hydroxyl group of a neighboring molecule. In $CaCl_2$ solution, there are changes in the intensity of the peak at -140º and its shoulder at -160º. This is a result of the surface/ion interactions and will be discussed in detail in the next section.

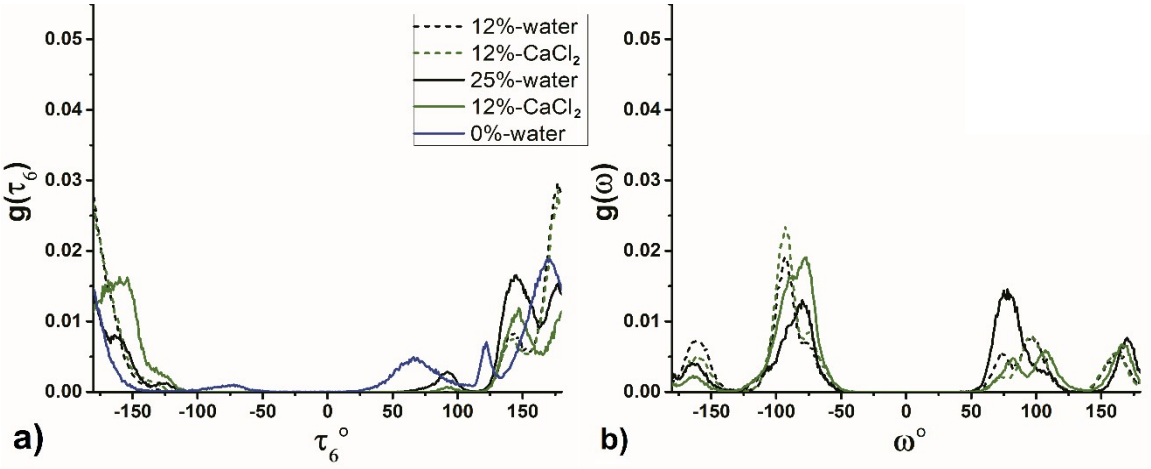

Figure S8. Distributions of the dihedral angles of the internal rotation τ_6 (a) and ω (b) for the native[24] (solid blue line) and the phosphorylated fibrils in water (12%: black solid and 25%: dashed lines), and in salt solution (green solid and dashed lines, respectively). For the definitions of the angles, see Figure S5.

The angle τ_6 determines the orientation of the hydrogen atom of the primary hydroxyl group (the O6-H vector, see Figure S5) in native cellulose and the oxygen atoms of the phosphate group (the O6-P vector) in phosphorylated cellulose. Similar behavior as above was observed for the distributions (Figure S8a), that is, upon phosphorylation, new peaks appear. Moreover, there are some changes in the intensities of the peaks of the distribution of the dihedral angle ω (determines the orientation of the P-O2 vector in phosphorylated cellulose molecules, Figure S8b) depending on the amount of phosphorylation and salt.

Mobility of the cations attracted by phosphate group.

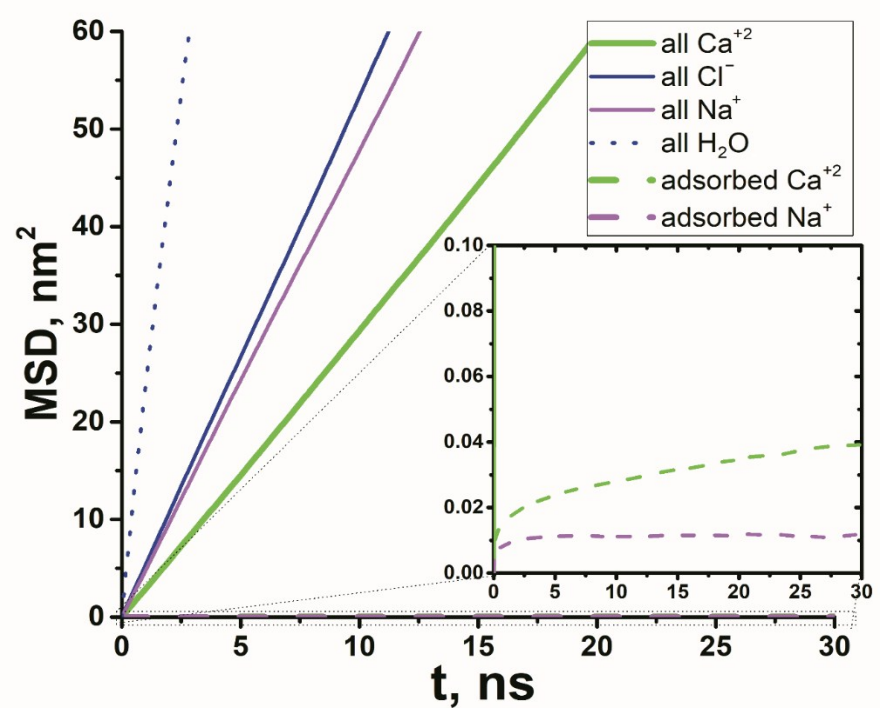

Figure S9. Mean square displacement of the cations attracted by phosphate group in comparison with mean square displacement of all (Ca^{2+}, Na^{+}, Cl^{-}) ions and water molecules in the system (for the system with 25% phosphorylated cellulose in $CaCl_2$ solution).

Numbers of Na^+ and Ca^{2+} ions adsorbed by oxygen atoms.

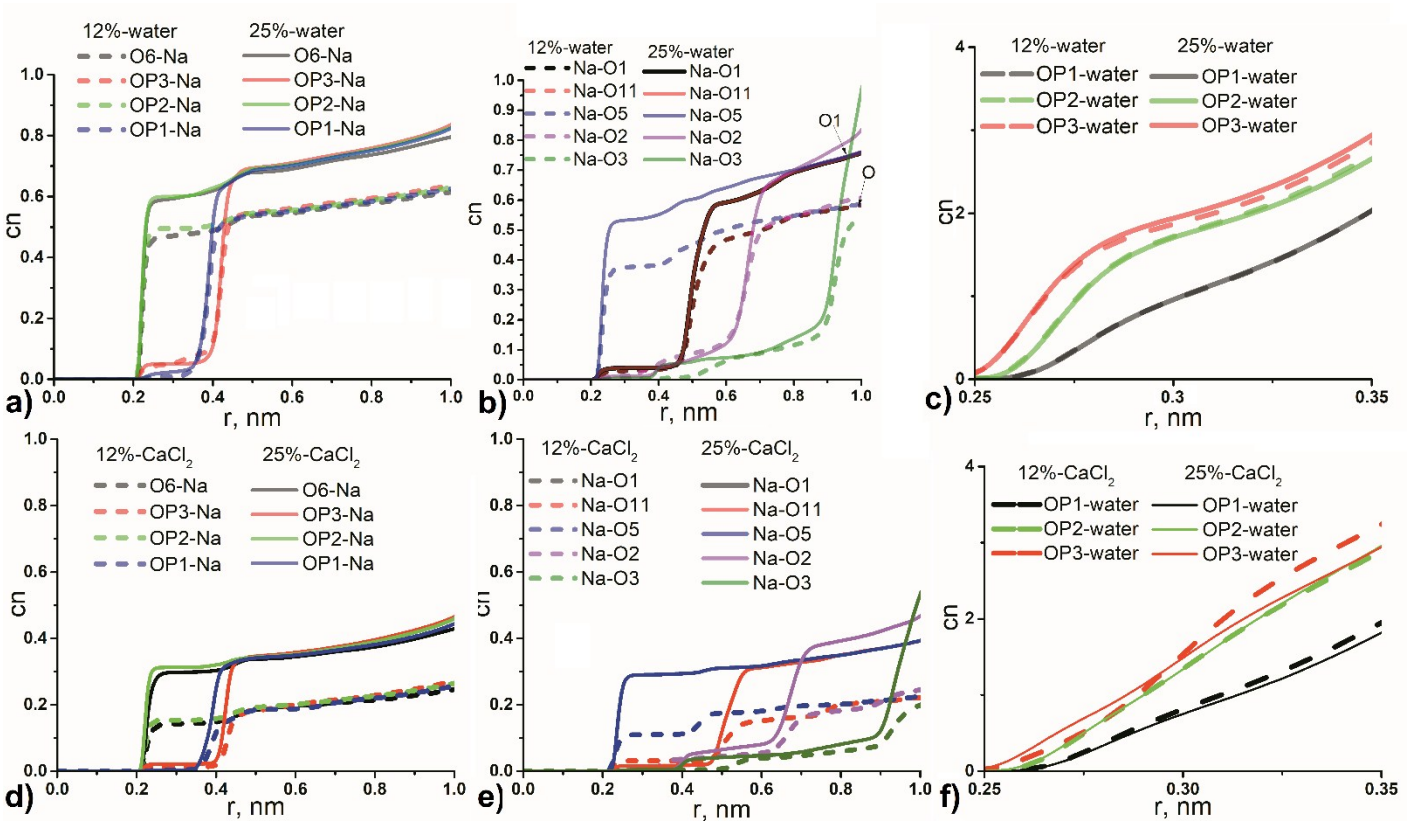

Figure S10. The numbers of the Na^+ ions (a,b,d,e) and water molecules (c,f) near the oxygen atoms of the phosphate groups (a,c,d,f) and glucopyranoze cycles (b.e) for the phosphorylated surfaces with 12% (dashed lines) and 25% (solid lines) of the substitution in water (a,b,c) and in $CaCl_2$ (d,e,f).

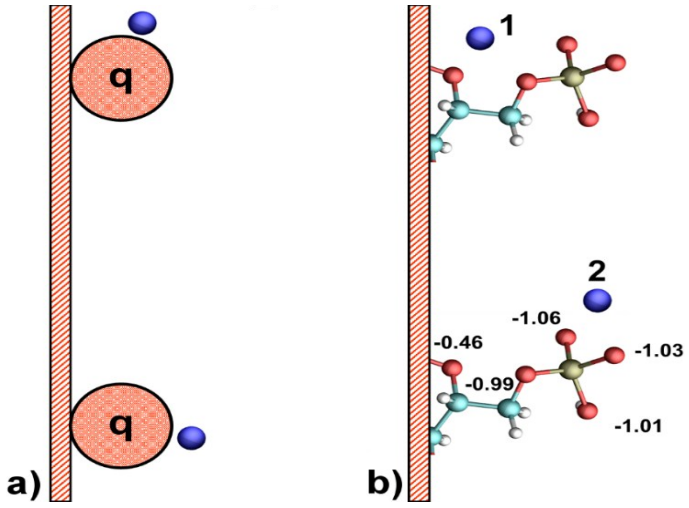

Figure S11. Illustration of the cation's positions on the surface for the model of a decorated surface (a) and in the current simulation (b).

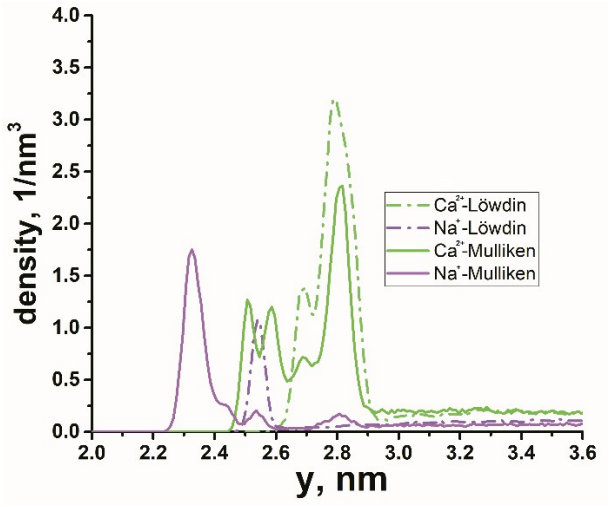

Figure S12. The number density profiles for Na^+ (violet color) and Ca^{2+} (green color) ions obtained with the Mulliken (solid lines) and Löwdin (dash dot lines) charges for the surface with 25% phosphorylation for the system in $CaCl_2$ solution.

Because of the lower electronegativity of the oxygen atoms in the Löwdin population analysis (Table SIII) the density profiles for both cations are shifted by about 2 Å further from the surface in comparison with those obtained by using Mulliken charges (Figure S12).

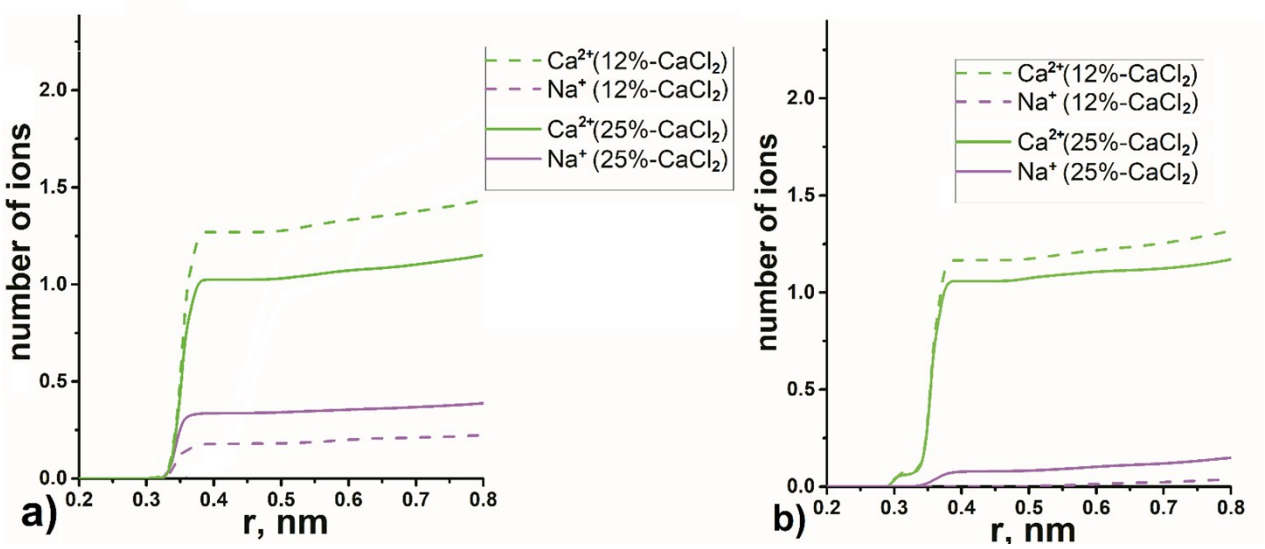

Figure S13. The numbers of Na^+ and Ca^{2+} ions in the first coordination shell of P atoms in $CaCl_2$ solution: for Mulliken (a) and Löwdin (b) charges.

Table SIII. The values of partial charges on atoms of the hydroxyl groups involved in the formation of intrafibrillar hydrogen bonds and for phosphate groups of phosphorylated cellulose obtained using Mulliken and Löwdin population analyses.

	Mulliken/6-31+G(d,p)	Löwdin/6-31+G(d)	CSFF
OH2	-0.616	-0.624	-0.600
OH3	-0.626	-0.624	-0.600
OH6	-0.686	-0.662	-0.660
O6	-0.985	-0.693	
P	2.467	1.538	
O1P	-1.005	-0.787	
O2P	-1.061	-0.891	
O3P	-1.026	-0.861	
HO1P	0.395	0.439	
O1	-0.42	-0.444	-0.31
O5	-0.46	-0.506	-0.36

REFERENCES

1 N. Treitel, R. Shenhar, I. Aprahamian, T. Sheradsky, M. Rabinovitz, Calculations of PAH Anions: When Are Diffuse Functions Necessary? *Phys. Chem. Chem. Phys.* 2004, **6**, 1113–1121.

2 M. Kuttel, J.W. Brady, K.J. Naidoo, Carbohydrate Solution Simulations: Producing a Force Field with Experimentally Consistent Primary Alcohol Rotational Frequencies and Populations. *J. Comput. Chem.* 2002, **23**, 1236–1243.

3 P.-O. Löwdin, On the Non-Orthogonality Problem Connected with the Use of Atomic Wave Functions in the Theory of Molecules and Crystals, *J. Chem. Phys.,* 1950, **18**, 365–375.

4 S. Nosé, A Molecular Dynamics Method for Simulations in the Canonical Ensemble. *Mol. Phys.* 1984, **52**, 255–268.

5 W.G. Hoover, Canonical Dynamics: Equilibrium Phase-Space Distributions. *Phys. Rev. A* 1985, **31**, 1695–1697.

6 M. Parrinello, A. Rahman, Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J. Appl. Phys.* 1981, **52**, 7182–7190.

7 U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, L.G.Pedersen, A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* 1995, **103**, 8577–8593.

8 W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, Comparison of simple potential functions for simulating liquid water, *J. Chem. Phys.*, 1983, **79**, 926-935.

9 S. Miyamoto, P.A. Kollman, Settle: An Analytical Version of the SHAKE and RATTLE Algorithm for Rigid Water Models. *J. Comput. Chem.* 1992, **13**, 952–962.

10 A.D. MacKerell, N. Banavali, N. Foloppe, Development and Current Status of the CHARMM Force Field for Nucleic Acids. *Biopolymers.* 2001, **56**, 257–265.

11 A.A. Granovsky, Firefly version 8, www http://classic.chem.msu.su/gran/firefly/index.html

12 M.W. Schmidt, K.K. Baldridge, J.A. Boatz, S.T. Elbert, M.S. Gordon, J.H. Jensen, S. Koseki, N. Matsunaga, K.A. Nguyen, S. Su, T.L. Windus, M. Dupuis, J.A. Montgomery, General atomic and molecular electronic structure system, *J.Comput.Chem.* 1993, **14**, 1347-1363.

13 M.J. Abraham, T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, E. Lindah, Gromacs: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* 2015, **1–2**, 19–25.

14 P.E. Mason, J.M. Cruickshank, G.W. Neilson, P. Buchanan, Neutron Scattering Studies on the Hydration of Phosphate Ions in Aqueous Solutions of K3PO4, K2HPO4 and KH2PO4. *Phys. Chem. Chem. Phys.* 2003, **5**, 4686.

15 C. Gao, G.Y. Xiong, H.L. Luo, K.J. Ren, Y. Huang, Y.Z. Wan, Dynamic Interaction between the Growing Ca-P Minerals and Bacterial Cellulose Nanofibers during Early Biomineralization Process. *Cellulose* 2010, **17**, 365–373.

16 A. Naji, S. Jungblut, A. G. Moreira and R. R. Netz, Electrostatic interactions in strongly coupled soft matter, Phys. A Stat. Mech. its Appl., 2005, **352**, 131–170.