Supporting Information

pH-responsive self-assembly of polysaccharide through a rugged energy landscape

Brian H. Morrow[†], Gregory F. Payne[‡§], and Jana Shen[†*]

[†]Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, Baltimore, MD 21201

[‡]Fischell Department of Bioengineering, University of Maryland, College Park, MD 20742
[§]Institute for Biosystems and Biotechnology Research, University of Maryland, College Park, MD 20742

[*]Corresponding author. Phone: (410) 706-4187; Fax: (410) 706-5017; E-mail: jshen@rx.umaryland.edu.

1 Methods and Protocol

Self-assembly simulations. The conventional all-atom molecular dynamics simulations were performed using GROMACS version 4.5.5[1]. The CHARMM36 forcefield for carbohydrates was used to represent chitosan[2]. The initial configurations for the self-assembly simulations were generated using the *genbox* and *genion* utilities in GROMACS. 24 chitosan chains, each consisting of 10 glucosamine units, were randomly placed in an 83.5 Å cubic box. The systems were solvated with CHARMM-style TIP3P waters. In the systems with 0.5 M NaCl, 175 NaCl pairs were added via random replacement of solvent molecules. The systems with and without salt contained $\sim$17,250 and $\sim$17,600 water molecules, respectively. The simulations were performed under NPT conditions. Temperature was maintained at 300 K with a modified Berendsen thermostat using velocity re-scaling with a stochastic term[3] with a time constant of 0.1 ps, while pressure was maintained with the isotropic Parrinello-Rahman pressure coupling method[4] using a time constant of 2.0 ps and a compressibility of 4.5 10^{-5} bar^{-1}. The van der Waals interactions were smoothly switched to zero between 10 and 12 Å. The particle mesh Ewald method was used to calculate long-range electrostatics, with a real-space cutoff of 12 Å and a fourth-order interpolation with 1.6 Å^{-1} grid spacing. Bonds involving hydrogen were constrained using the LINCS algorithm,[5] enabling a 2-fs time step.

Structure preparation and equilibration for the model crystallite. Starting from the asymmetric unit of the hydrated crystal structure of chitosan[6], crystallites consisting of various numbers of chains and sheets were built. These crystallites were solvated in a cubic water box with a minimum distance of 15 Å between the solute and the edge of the box. The solvated systems were subjected to conventional molecular dynamics simulations with the amine groups fixed at the unprotonated (neutral) states. The same settings as in the self-assembly simulations were used. The smallest stable crystallite was found to contain 16 chains of 10-unit chitosan. The chains were initially arranged in sheets consisting of 4 hydrogen-bonded chains, with 4 sheets stacked vertically (Figure S3a,b). This crystallite was stable over the course of a 50 ns NPT molecular dynamics. The final configuration is shown in Figure S3c.

All-atom continuous constant pH molecular dynamics. In continuous constant-pH molecular dynamics (CpHMD) method, a titration coordinate λ_i is introduced for each group i that can change protonation state, with λ_i values bound between 0 and 1, corresponding to the protonated and unprotonated states, respectively[7,8]. The titration coordinates are coupled to and propagated simultaneously with the spatial coordinates using an extended Hamiltonian.[7,8]. We employed the all-atom version of CpHMD[9], which unlike the previous CpHMD methodologies[7,8], does not rely on the generalized Born implicit-solvent model. To accelerate sampling convergence, the pH-based replica-exchange protocol was applied[10].

Multiple replicas subjected to different pH conditions are run simultaneously, with an exchange of adjacent pH conditions attempted every 500 MD steps (1 ps) according to the Metropolis criterion. All 160 amine groups are allowed to simultaneously titrate. The newly developed titratable water model was used to maintain charge neutrality for the simulation system[11]. Each glucosasmine was coupled to a titrating hydronium ion such that the pair has a constant +1 net charge and can be neutralized with a chloride ion. Simulations were performed at constant temperature of 300 K and pressure of 1 atm using Hoover thermostat[12] and Langevin piston pressure coupling algorithm[13]. A switching function starting at 10 Å reduces the van der Waals interactions to zero at 12 Å. The generalized reaction field method[14] was used to compute long-range electrostatic interactions, with a cutoff of 14 Å. The SHAKE algorithm was used to constrain bonds involving hydrogen. The equations of motion were integrated with a 2 fs timestep using an in-house modified version of CHARMM c37b[15].

Potential of mean force of the model compound titration. CpHMD simulations require the potential of mean force (PMF) of the protonation/deprotonation reaction of a model compound, in this case a single glucosamine. The PMF along λ for titrating a glucosamine unit in water was determined via thermodynamic integration, $\Delta G = \int \langle dU/d\theta \rangle_\theta d\theta$, where θ is related to λ via $\lambda = sin^2(\theta)$[7]. The average force, $\langle dU/d\theta \rangle_\theta$, was calculated from 1 ns NPT simulations at θ values of 0.2, 0.4, 0.6, 0.8, 1.0, 1.2, and 1.4. Assuming the PMF is a quadratic function, we fit the average force to a linear function of lambda, $\langle dU/d\lambda \rangle = 2A(\lambda - B)$, which resulted in $A = -63.1846$ and $B = 0.887983$.

REX-CpHMD simulation of a single chain. The 5-unit chain was solvated in an octahedral water box, with at least 15 Å between the chain and box edge. One hydronium co-ion and one chloride counter-ion were added for each glucosamine unit. 8 replicas evenly spaced between pH 6.0 and 9.5 were used. The simulation was 15 ns per replica, for an aggregate sampling time of 120 ns.

REX-CpHMD simulation of the model crystallite. The CpHMD simulation of the chitosan crystallite used 16 chains each consisting of 10 glucosamines. The final configuration from the 50-ns simulation of the neutral crystallite was placed in a 90.45 Å cubic box with 160 hydroniums and 160 chlorides and solvated with 21,979 TIP3P water molecules. 19 replicas with 0.25-unit spacing covered the pH range 4.0–8.5. The simulation was run for 17 ns per replica, for a an aggregate sampling time of 323 ns.

Calculation of pK_a's To calculate unprotonated fractions, we considered $\lambda < 0.1$ and $\lambda > 0.9$ to correspond to the protonated and unprotonated states, respectively[7,8]. Intermediate values of λ were discarded as unphysical[7,8]. On average, the fraction of these intermediate values was $< 10\%$. The unprotonated fraction S is computed as $S = N_u/(N_u + N_p)$, where N_u and N_p are the number of unprotonated and protonated states, respectively. To obtain the pK_a we fitted the unprotonated fraction as a function of pH to the Hill equation, $S = 1/(1+10^{n(pK_a - pH)})$, where n is the Hill coefficient.

Calculation of pH-dependent free energy of dissociation. The free energy of dissociation of the model crystallite (also referred to as stability in the main text) was calculated using the Wyman-Tanford linkage equation[16,17]:

$$\partial \Delta G / \partial \mathrm{pH} = 2.303 RT (Q^{\mathrm{D}} - Q^{\mathrm{A}}) \tag{1}$$

where Q^{A} and Q^{D} refer to the charge of the fully-associated crystallite and fully-dissociated chains, respectively. The value of Q^{D} at each pH condition was obtained using the amine pK_{a}'s from the single-chain simulation. Since the crystallite dissociates rapidly at low pH, we were unable to obtain precise Q^{A} values below the transition pH. To estimate Q^{A} we used the chitosan charge from the final 3 ns of the dissociation simulation. This gives an upper bound for ΔG for pH $<$ 6.5. To verify this estimate, we also calculated ΔG assuming a two-state equilibrium between associated and dissociated states[18] based on the number of inter-chain hydrogen bonds. This method gave a relative stability of $\sim$40 kcal/mol at pH 4, in qualitative agreement with the result calculated using the Wyman-Tanford linkage equation.

2 Supplementary Figures

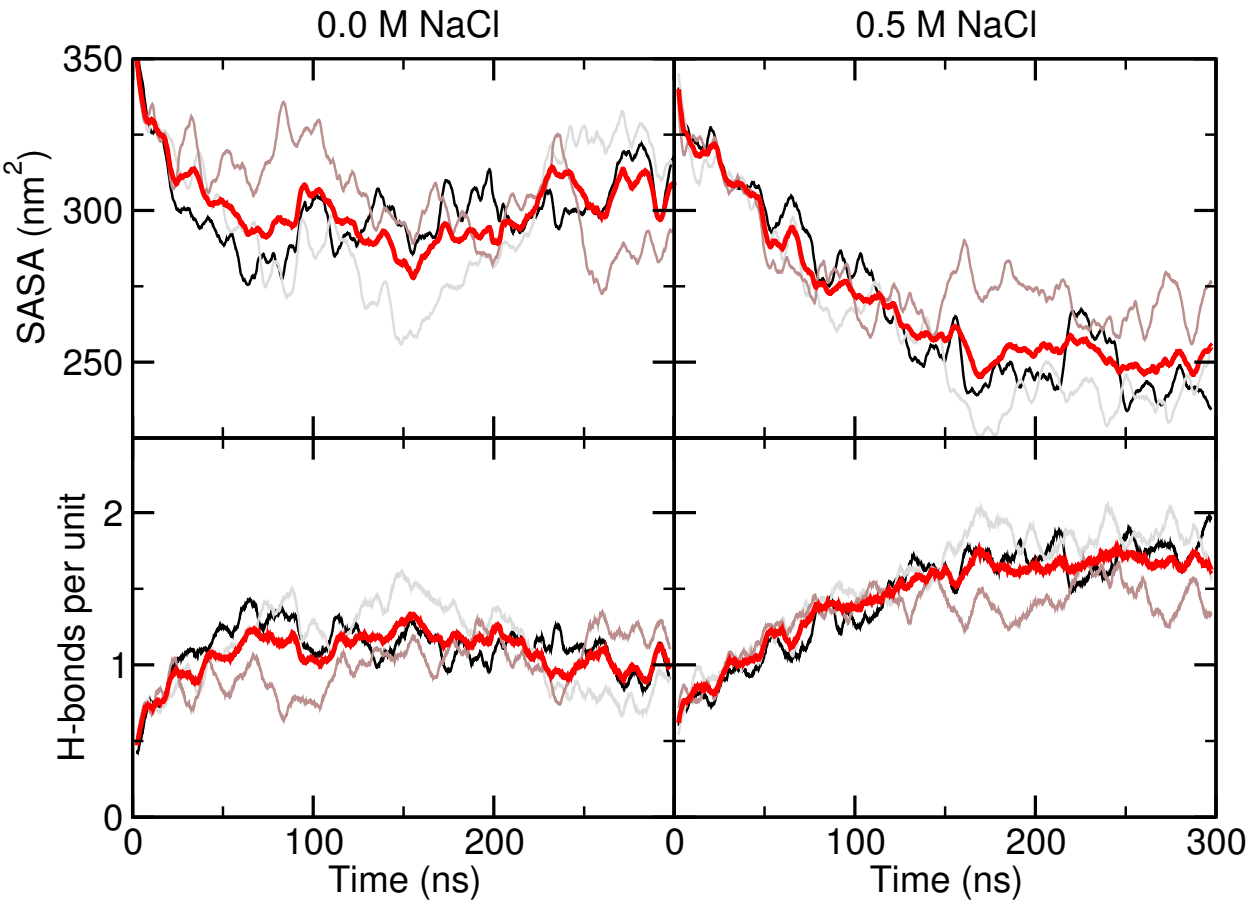

Figure S1: **Progression of the chitosan self-assembly**. **Top**: Time series of the total solvent-accessible surface area (SASA) of chitosan in the 3 independent sets of self-assembly simulations with and without 0.5 M NaCl. The SASA was calculated using the *g_sas* utility in GROMACS with a probe radius of 1.4 Å. **Bottom**: Time series of the number of inter-chain hydrogen bonds per glucosamine unit. Values for the three independent runs are shown in black, grey, and brown, while the average is shown in red.

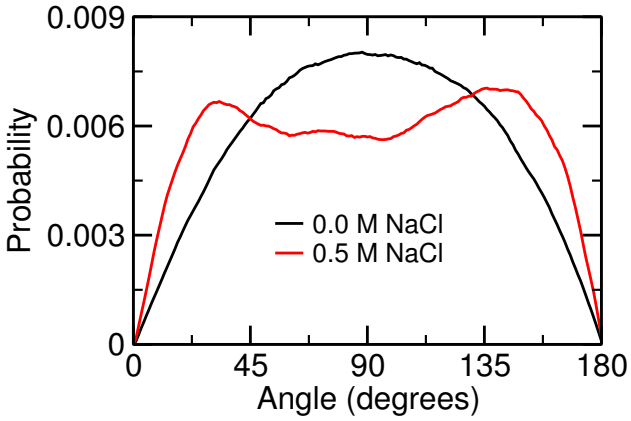

Figure S2: **Difference between the self-assembled aggregate with salt and without salt**. Distribution of angles between chitosan chains, averaged over the final 100 ns of each self-assembly simulation. Angles are between the vectors joining the first and last glucosamine unit of each chain.

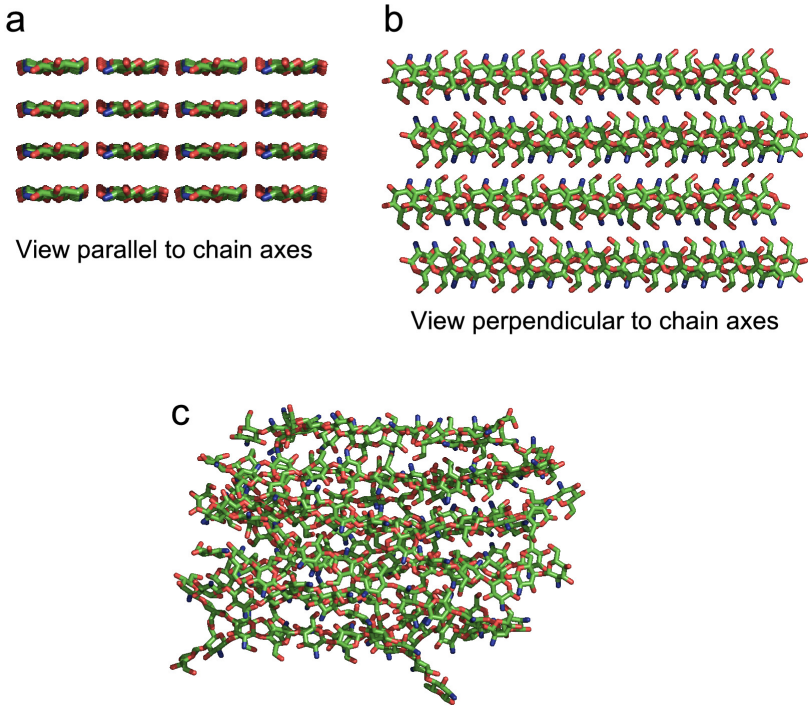

Figure S3: **Snapshots of the initial and final structure of the chitosan crystallite**. The structure consists of 16 chains of 10-unit chitosan, arranged in a 4×4 pattern, viewed along (**a**) and perpendicular to (**b**) the chain axes. **c**, Final structure after 50-ns simulation with all amines fixed in the unprotonated states.

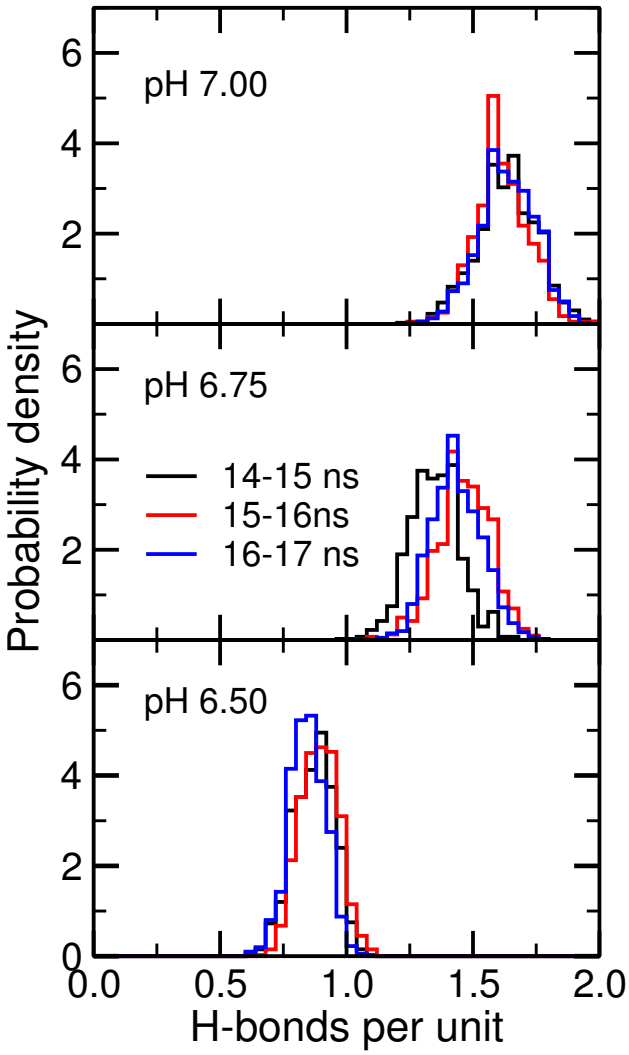

Figure S4: **Convergence of the dissociation simulation**. Distribution of the number of inter-chain hydrogen bonds for the last three time windows. Only replicas near the transition pH are shown because they are the ones with the slowest convergence.

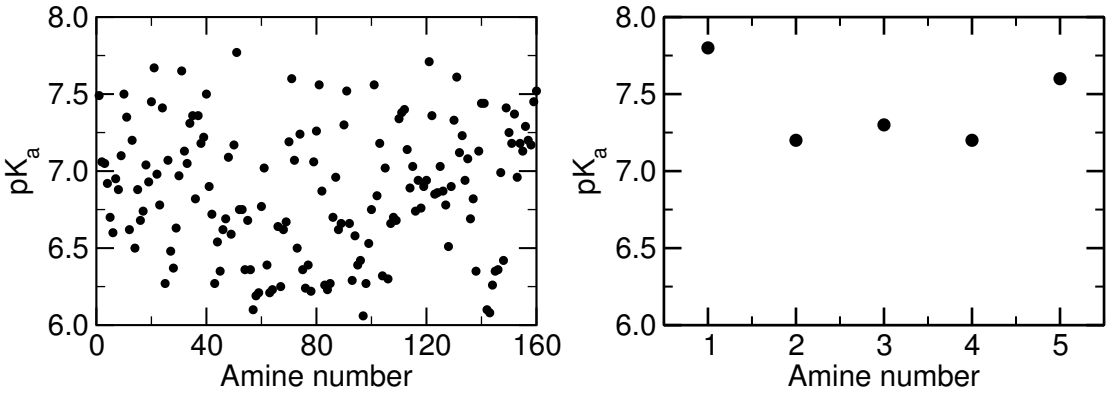

Figure S5: **Amine pK_{a} values**. Microscope pK_{a}'s of amines in the model crystallite (left) and 5-unit single chain (right).

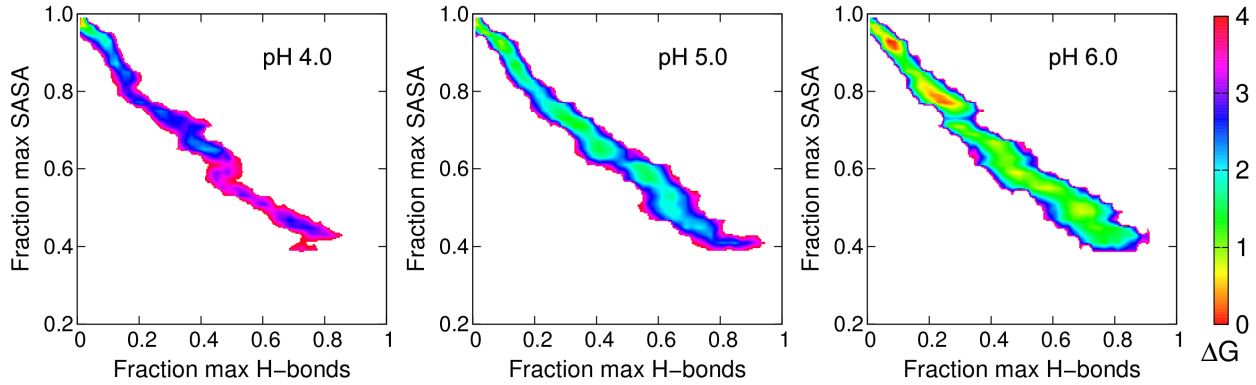

Figure S6: **Mechanism of the pH-dependent dissociation of chitosan crystallite**. Free energy surface as a function of the fraction of the maximum solvent-accessible surface area (SASA) and the fraction of the maximum number of inter-chain hydrogen bonds at three pH conditions below the transition pH. Relative free energy is defined as -RTln(P_i/P_0), where P_i is the probability of state i and P_0 is the probability of the most probable state. Calculation is based on the entire 15 ns (per replica) of the CpHMD simulations.

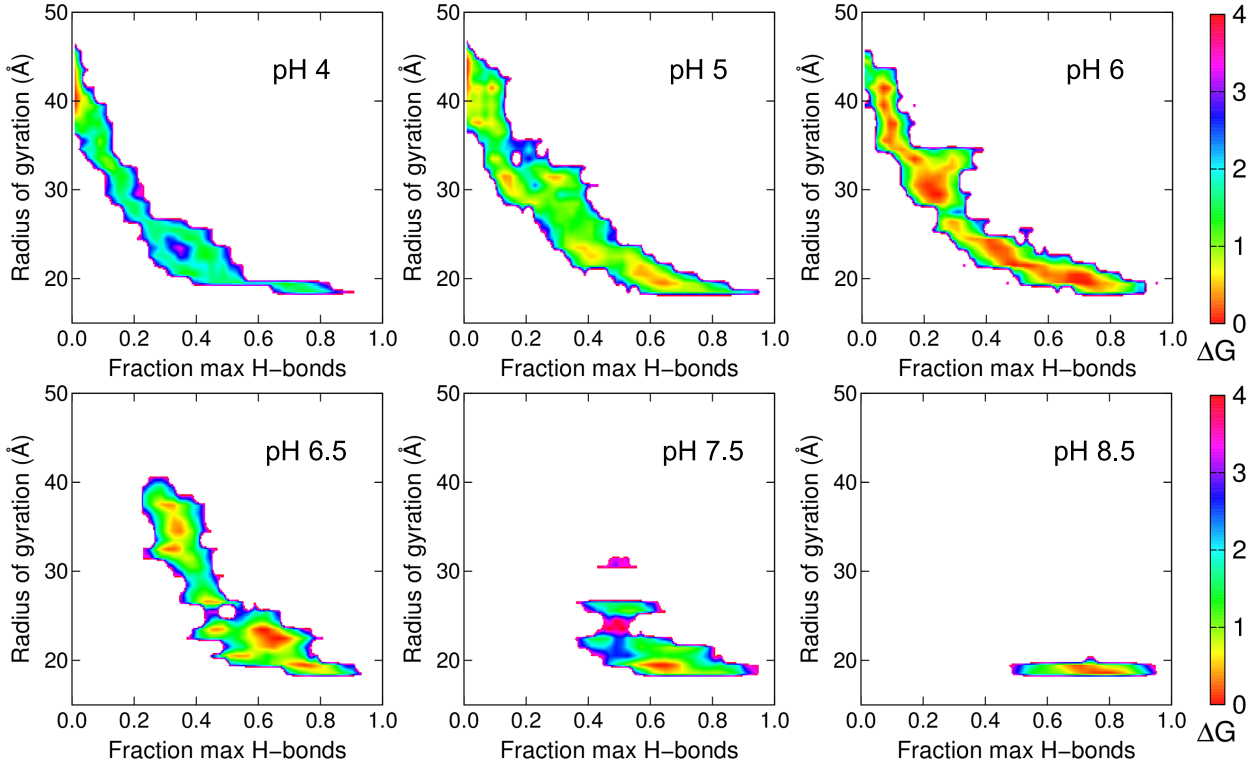

Figure S7: **Mechanism of the pH-dependent dissociation of chitosan crystallite**. Free energy surface as a function of the radius of gyration of the model crystallite and the fraction of the maximum number of inter-chain hydrogen bonds at three pH conditions below the transition pH. Relative free energy is defined as -RTln(P_i/P_0), where P_i is the probability of state i and P_0 is the probability of the most probable state. Calculation is based on the entire 15 ns (per replica) of the CpHMD simulations.

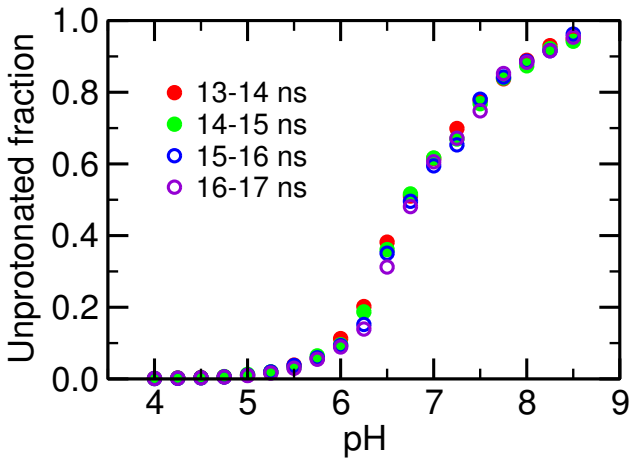

Figure S8: **Convergence of the unprotonated fractions**. Unprotonated fraction of the chitosan crystallite calculated over 1-ns time windows.

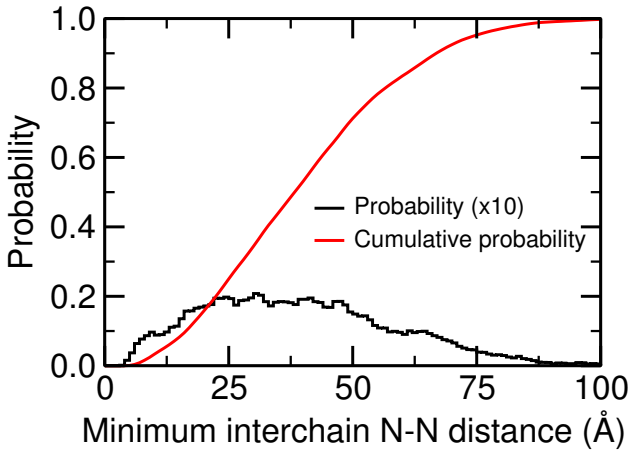

Figure S9: **Potential interaction of protonated chains in the CpHMD simulation**. Probability distribution of the minimum interchain N-N distance for all pairs of chitosan chains at pH 4.0, averaged over the final 3 ns.

References

1. Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.
2. Guvench, O.; Mallajosyula, S. S.; Raman, E. P.; Hatcher, E.; Vanommeslaeghe, K.; Foster, T. J.; Jamison, II, F. W.; MacKerell, Jr., A. D. *J. Chem. Theory Comput.* **2011**, *7*, 3162–3180.
3. Bussi, G.; Donadio, D.; Parrinello, M. *J. Chem. Phys.* **2007**, *127*, 14101–14107.
4. Parrinello, M.; Rahman, A. *J. Appl. Phys.* **1981**, *52*, 7182–7190.
5. Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. *J. Comput. Chem.* **1997**, *18*, 1463–1472.
6. Okuyama, K.; Noguchi, K.; Miyazawa, T.; Yui, T.; Ogawa, K. *Macromolecules* **1997**, *30*, 5849–5855.
7. Lee, M. S.; Salsbury, Jr., F. R.; Brooks III, C. L. *Proteins* **2004**, *56*, 738–752.
8. Khandogin, J.; Brooks III, C. L. *Biophys. J.* **2005**, *89*, 141–157.
9. Wallace, J. A.; Shen, J. K. *J. Chem. Phys.* **2012**, *137*, 184105.
10. Wallace, J. A.; Shen, J. K. *J. Chem. Theory Comput.* **2011**, *7*, 2617–2629.
11. Chen, W.; Wallace, J.; Yue, Z.; Shen, J. *Biophys. J.* **2013**, *105*, L15–L17.
12. Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695–1697.
13. Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. *J. Chem. Phys.* **1995**, *103*, 4613–4621.
14. Tironi, I. G.; Sperb, R.; Smith, P. E.; van Gunsteren, W. F. *J. Chem. Phys.* **1995**, *102*, 5451–5459.
15. Brooks, B. R.; Brooks III, C. L.; Mackerell Jr., A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartles, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Lazaridis, K. K. T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. *J. Comput. Chem.* **2009**, *30*, 1545–1614.
16. Wyman, Jr., J. *Adv. Protein Chem.* **1964**, *19*, 223–286.
17. Tanford, C. *Adv. Protein Chem.* **1970**, *24*, 1–95.
18. Park, C.; Marqusee, S. *Protein Sci.* **2004**, *13*, 2553–2558.