An implicit coarse-grained model for alginate hydrogels crosslinked with chitosan oligomers

Lukas Baldauf

June 10, 2022

Abstract

A coarse-grained approach is adopted in this master's project to investigate hydrogels at the microscopic scale using molecular dynamics simulations. The hydrogels studied are composed of alginates crosslinked with chitosan oligomers of lengths 4 to 8 and are particularly relevant for biomedical applications. Obtaining insights into how the mechanical properties of these hydrogels are affected by certain factors at the microscopic scale would allow the design of novel materials tailored toward specific applications. Therefore, this project seeks knowledge of how the underlying microscopic phenomena affect the macroscopic behavior of these hydrogels. The main objective was to develop an implicit solvent coarse-grained model to study the role of alginate composition, counterions, chitosan oligomers, and polymeric effects on gel strength. The model was developed based on atomistic reference simulations with the GLYCAM06 and CHARMM36 force fields. The conformational properties of the coarse-grained model were compared to experimental data as a validation step. A satisfactory agreement was found between simulations and experiments. The model was further applied to large-scale systems representative of a hydrogel. However, a thorough analysis of the gel strength based on previously mentioned factors was not performed and is left for future studies. Still, the model reproduced experimentally observed trends and shed light on some of the underlying dynamics that may affect the macroscopic behavior of these gels.

Acknowledgements

I am grateful my supervisor, Anders Lervik, for guiding me throughout this long project. I would also like to thank my family, friends, and better half for always offering their support.

Contents

Symbols and abbreviations

Table 1: List of key variables.

Symbols	
b	Kuhn length
$\mathbf{C}$	cross-link matrix
$g(r)$	radial distribution function
$G(t)$	shear stress relaxation modulus
η	viscosity
k_b	Boltzmann constant
l	distance between monomers in a polymer
M	molecular mass
N	number of repeating units in a polymer
p	p'th Rouse mode
q	persistence length
R_{G}	radius of gyration
ρ	mass concentration (mass per volume)
τ_R	Rouse time, the longest relaxation time
τ_p	relaxation time of the p'th Rouse mode
T	temperature
V	volume or a potential function
ξ	friction constant

Table 2: List of abbreviations.

Abbreviations	
AA	all-atom
CSVR	canonical sampling through velocity rescaling
CG	coarse-grained
DP	degree of polymerization
ESP	electrostatic potential
G	α-L-guluronate
HF	Hartree Fock
M	β-D-mannuronate
NB1	non-bonded parameter set 1
NB2	non-bonded parameter set 2
NPT	isothermal-isobaric
PEC	polyelectrolyte complex
NVT	canonical
MD	molecular dynamics
PME	particle mesh Ewald
PMF	potential of mean force
RDF	radial distribution function
RESP	restrained electrostatic potential
WHAM	weighted histogram analysis method
Y	coarse-grained chitosan residue

1 Introduction

Alginates are linear polymers comprised of two different monomers, abbreviated G and M, and are important in the Norwegian industry. They are found in brown algae and some bacteria. In Norway, they are harvested mainly around the west coast, totaling an annual production of around 6'000 tons, with a global production of around 30'000 tons. Most of their applications are based on their ability to form gels when calcium salts are added [1].

A recent article studied gelling systems composed of alginates crosslinked with chitosan oligomers [2]. One of the studied gelling systems in which the alginate composition strictly alternates in its G- and M-residue composition (poly-MG) revealed surprisingly high gel strengths when crosslinked with chitosan oligomers of lengths 7 to 9. The poly-M analog displayed much lower gel strengths, while poly-G did not form gels but rather a turbid viscous solution. The authors suggested that hydrophobic interactions, phase separations, or both play a vital role in the poly-MG systems. The high gel strengths and turbidity of the poly-MG gelling systems also indicated a fundamentally different gelling mechanism compared to its poly-M counterparts. Additionally, the poly-MG gelling systems revealed high tolerances to a broad range of salt concentrations, making them relevant for biomedical applications [2].

Obtaining insights into how the mechanical properties of these hydrogels are affected by certain factors at the microscopic scale would allow the design of novel materials tailored toward specific applications. However, obtaining microscopic insights from an atomistic modeling approach is challenging due to the large spatial and temporal scales of the underlying polymer dynamics. For example, the chain lengths of the alginates studied in [2] were around 800 - 1200, but treating alginate chain lengths greater than 50 at atomic resolution becomes impractical due to a large number of solvent molecules. In addition, the forming and breaking of bonds between alginate chains and chitosan oligomers further slows down the chain dynamics [3].

A coarse-grained approach in which several atoms are combined to form larger units and the solvent is treated implicitly may be worth exploring. Such a model leads to a drastic reduction in the computational demands and allows exploring the effects of specific parameters on the behavior of these hydrogels.

All these considerations lead to the main objective of this master's project:

Developing an implicit solvent coarse-grained model describing the above-mentioned gelling systems to study the role of alginate composition, counterions, chitosan oligomers and polymeric effects on gel strength.

This thesis is outlined as follows. First, the theory part describes the necessary background for understanding this text, covering the structure of alginates and chitosan, the basics of molecular dynamics simulations, and coarse-graining. The first section concludes with an introduction to the theory of viscoelasticity, which is essential for describing the properties of hydrogels.

The methods section describes the development and subsequent application of a coarse-grained model. However, adequate reference simulations must be performed before such a model can be developed, and these are described first.

The results section covers the microscopic insights obtained from the reference simulations. The results of the coarse-grained model development are then presented, and the applications of this model to large scales systems are explored.

Finally, a summary of the conclusions drawn from the large-scale applications and suggestions for future studies are given in the last section.

2 Theory

An introduction to the relevant theory is given in the following section. First, the structure of alginates and chitosan, the basics of molecular dynamics simulations, and coarse-graining are covered. Then, a simple model describing viscoelasticity is developed as an introduction to the subject. This section concludes with an expression relating the viscoelastic properties of a material to the microscopic origins of the phenomenon.

2.1 Alginates, chitosans and hydrogels

Alginates are linear polymers of (1→4)-linked α-L-guluronate residues (G) and β-D-mannuronate residues (M), as illustrated in figure 1. The carboxylic acid groups on alginate have a pKa of around 3.5, meaning they are polyanions in neutral to mildly acidic solutions [4].

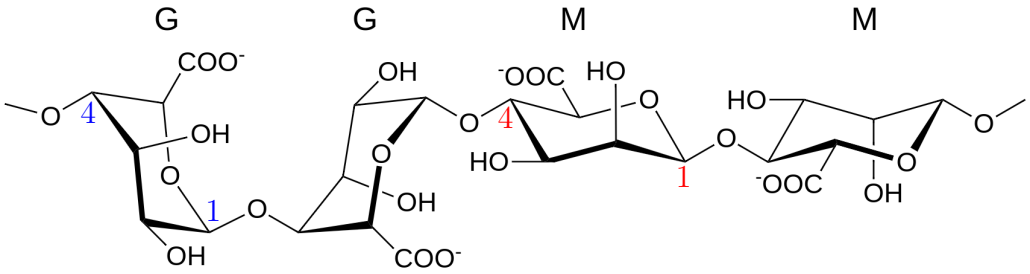


Figure 1: The structure of alginate. (1→4)-linked α-L-guluronate (G) in the $^1C_4$ chair conformation and β-D-mannuronate (M) in the $^4C_1$ conformation. The residue composition of the alginate in this figure can be specified as GGMM. The chair conformation is given by the position of the atoms labeled with colored numbers relative to the sugar ring.

Chitosans are linear polysaccharides with varying amounts of β-(1→4)-linked D-glucosamine and N-acetyl-D-glucosamine residues. Figure 5 illustrates the repeating structure of chitosan containing only D-glucosamine, which is most relevant in this project. Each residue has a pKa of around 6.5, making them polycations at low pH [4].

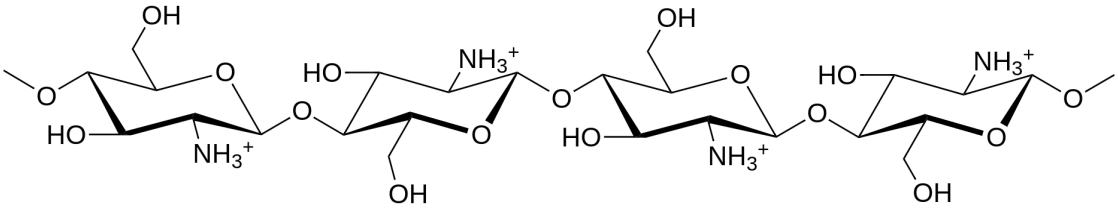


Figure 2: The structure of chitosan. Each residue consists of β-(1→4)-linked D-glucosamine in the $^4C_1$ chair conformation.

At pH values around 4.5, a large fraction of the carboxyl groups of alginate and the amino groups of chitosan are ionized, and the two polymers readily form complexes in solution. These so-called polyelectrolyte complexes (PEC) give rise to many interesting solution properties, such as the formation of hydrogels.

Hydrogels are 3-dimensional networks of polymers that can retain their structure and hold large amounts of water due to chemical or physical crosslinking. Physical hydrogels,

such as those formed by alginate-chitosan complexes, can transition from a liquid phase to a gel in response to factors such as ionic strength, pH, or changes in the concentration of its components. In addition, alginate and chitosan are naturally derived polymers and often interact favorably in vivo. These properties make their hydrogels excellent candidates for biomedical applications. Examples include drug delivery and tissue engineering [5]. The poly-MG gelling systems mentioned in the introduction are especially relevant. A deeper understanding of these gels at the microscopic level is therefore desired.

2.2 Molecular Dynamics

The movement of atoms and molecules can explain many macroscopic phenomena around us. Molecular dynamics (MD) is a computational tool for exploring the microscopic origins that give rise to these phenomena. Given some system of particles, MD evolves the atomic positions $\mathbf{r}_i$ and velocities $\mathbf{v}_i$ of these particles by numerically integrating Newton's equation of motion

$$\frac{d^2\mathbf{r}_i}{dt^2} = -\frac{1}{m_i}\frac{\partial V}{\partial \mathbf{r}_i} \tag{2.1}$$

Here, $V(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)$ is a potential function specific to the system at hand, m_i is the mass of particle i, and $\mathbf{r}_i$ is the position of particle i. In many cases, a simple form of V is often a reasonable approximation for the movement of atoms:

$$V = V_{\text{LJ}} + V_{\text{coul}} + V_{\text{bond}} + V_{\text{angle}} + V_{\text{dihedral}} \tag{2.2}$$

Within this assumption, a Lennard-Jones (LJ) potential V_{LJ} and a Coulomb potential V_{coul} describe the so-called non-bonded interactions between particles. The bond, angle, and dihedral potentials (V_{bond}, V_{angle} and V_{dihedral}, respectively) are assumed to represent the conformational properties of molecules. The potentials in equation 2.2 are, in many cases, derived from quantum mechanical calculations of small reference molecules performed in vacuum. The derivative $\partial V/\partial \mathbf{r}_i$ of the total potential with respect to all atoms in the system then defines the so-called force field [6]. Examples include the GLYCAM06 force field [7] and the carbohydrate extension of the CHARMM36 force field [8], which are both used in this project. Both these force fields have been used in numerous applications involving alginates and chitosan [9–11]. However, the atomic charges defining V_{Coul} are not available in the literature and are therefore developed in this project.

The bond potential between two atoms i and j within the GLYCAM06 and CHARMM36 force field is represented by harmonic function of the distance between the atoms r_{ij}

$$V_{\text{bond}}(r_{ij}) = \frac{1}{2}k_{ij}(r_{ij} - b_{ij})^2 \tag{2.3}$$

where k_{ij} is the force constant and b_{ij} the equilibrium bond length.

The potential of an angle θ formed by a triplet of atoms i, j and k is also represented by a harmonic function within the GLYCAM06 force field

$$V_{\text{angle}}(\theta) = \frac{1}{2}k_\theta(\theta - b_\theta)^2 \tag{2.4}$$

where k_θ is the force constant and b_θ the equilibrium angle. Within the CHARMM36 force field, the angle potential contains an additional correction term between the two outer atoms i and k

$$V_{\text{angle}}(\theta, r_{ik}) = \frac{1}{2}k_\theta(\theta - b_\theta)^2 + \frac{1}{2}k_{ik}(r_{ik} - b_{ik})^2 \tag{2.5}$$

where the first term is identical to equation 2.4 and the last term is identical to equation 2.3.

The dihedral angles within the GLYCAM06 force field are given by a periodic function

$$V_{\text{dihedral}}(\phi) = k_\phi(1 + \cos(n\phi - b_\phi)) \tag{2.6}$$

where n is the periodicity, ϕ is the angle between the planes formed by the ijk and jkl atoms, b_ϕ is the equilibrium angle and k_ϕ the force constant. The same function describes most of the dihedral angles within the CHARMM36 force field. The so-called improper dihedral angles in the CHARMM36 force field are described by a harmonic potential

$$V_{\text{dihedral}}(\xi) = \frac{1}{2}k_\xi(\xi - k_\xi)^2 \tag{2.7}$$

where ξ is the angle between the planes formed by the ijk and jkl atoms, k_ξ is the force constant and b_ξ the equilibrium angle. Improper dihedral angles may be used to keep certain groups planar, which in this case are the C-CO_2^- groups of alginates (see figure 1).

The LJ potential between two atoms i and j is given by

$$V_{\text{LJ}}(r_{ij}) = 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \tag{2.8}$$

where r_{ij} is the distance between atoms i and j, -ε_{ij} is the minimum energy of the LJ potential, and σ_{ij} is related to distance at the energy minimum by $2^{1/6}\sigma_{ij}$.

The coulomb potential between two particles with charge q_i and q_j is given by

$$V_{\text{coul}}(r_{ij}) = f\frac{q_i q_j}{\varepsilon_r r_{ij}} \tag{2.9}$$

where f is a constant and ε_r is the relative dielectric constant, which is taken to be 1 for both the GLYCAM06 and CHARMM36 force fields [12]. Restrained electrostatic potential (RESP) charges are used to describe the charges on atoms within the GLYCAM06 force field. These charges are obtained by a fitting procedure that reproduces electrostatic potential (ESP) calculated from quantum mechanical calculations at many grid points around the molecule. The ESP charges of buried carbon atoms are often dependent on the molecular conformation. Introducing restraints in the ESP fitting procedure considerably reduces this problem [13]. These charges are then denoted RESP charges.

The non-bonded interactions (LJ and Coulomb interactions) between 1-2 and 1-3 atom pairs are excluded in the CHARMM36 and GLYCAM06 force field. The 1-2 atom pairs include all atom pairs directly bonded to each-other, while 1-3 atom pairs include all atom pairs separated by exactly one bond. In some force fields, the 1-4 atom pair interactions are additionally reduced or excluded because they are implicitly incorporated in the dihedral potentials. Such an approach is used for the coarse-grained model developed later, in which also 1-4 interactions are excluded. However, the CHARMM36 and GLYCAM06 force fields treat these interactions at full strength.

Radial distribution functions are useful for describing the average structure of a dynamic system during a simulation. Therefore, they are extensively used as target properties in this project, which the coarse-grained model should reproduce. The radial distribution function $g(r)$ gives the probability of finding a particle at a distance r from another particle compared to the ideal gas distribution [6, p. 310].

If some molecules in the system interact strongly, the radial distribution function obtained from a finite simulation may be dominated by the more probable regions of the configuration space. Therefore, the simulation can be biased to force molecules or particles into lower probability regions separated by large energy barriers, which are not explored efficiently during conventional simulations. Forcing molecules along a given coordinate, which may be the distance between the centers of two molecules, can be used to obtain the free energy as a function of the coordinate. This free energy along the given coordinate is then referred to as the potential of mean force (PMF), which also incorporates solvent effects if performed in solution [6, pp. 580–581].

2.3 Coarse-graining

The molecular weight of the alginates studied in [2] ranged from 158 - 232 kDa, corresponding to a chain length of around 800 - 1200 monomers. Conventional atomistic simulations of such large systems are expensive because of the slow time scales governing the chain motion and a large number of solvent molecules in the simulation box. The alginate concentration of the relevant gels in [2] was 1% (10 g L^{-1}), meaning that most computations would be spent on water-water interactions. A coarse-grained (CG) approach in which multiple atoms are grouped into larger units and the solvent is treated implicitly enables the exploration of system sizes unreachable by atomistic simulations.

Coarse-graining is a promising approach to understanding the properties of polymeric systems too large for atomic resolution simulations. In a CG procedure, the number of degrees of freedom of the atomistic model is reduced by mapping atoms into larger units, called beads. The mapping is achieved by applying an operator $\mathbf{M}$ to the atomistic coordinates $\mathbf{r}$ to produce a set of CG coordinates $\mathbf{R}$. In this project, $\mathbf{M}$ is defined by a center-of-geometry grouping of specific atoms from the AA model. This grouping is denoted as the mapping scheme.

The reduced degrees of freedom allow for a larger integration time step and a more efficient phase-space exploration by decreasing energy barriers. Adopting a CG model also requires less time on force calculations and less memory to store trajectories. However, this smoother energy landscapes leads to timescales that are not directly relatable to the timescales of the AA model [14, 15]. The timescale of the CG model may therefore be scaled to match either the atomistic model or experimental data.

An implicit solvent treatment is necessary for the polymer lengths and concentrations considered in this project. Replacing the solvent with an implicit solvent yields a significant gain in computational efficiency because the expensive calculation of water-water interactions is avoided. The use of an implicit solvent model may be justified by the fact that, in dilute solutions, the viscoelastic properties polymeric system can be split into a sum of two terms. One term represents the property of a pure solvent, and the other term represents the effect of the polymer [16, pp. 108–109]. Several implicit solvent models have been developed for carbohydrates [17, 18].

The potential functions defining the CG model in this project are identical to the

potentials used in the carbohydrate extension of the well-known martini CG force field [19]. The bond potential between two atoms is given by equation 2.3. The potential of an angle θ formed by three consecutively connected atoms is represented by a cosine-harmonic potential

$$V_{\text{angle}}(\theta) = \frac{1}{2} k_\theta (\cos(\theta) - \cos(b_\theta))^2 \tag{2.10}$$

where k_θ is the force constant and b_θ the equilibrium angle. The dihedral potential between four consecutively connected atoms are represented by equation 2.6 with a multiplicity of $n = 1$. A LJ potential (equation 2.8) represents the non-bonded interactions between different CG particle types.

Two different potentials are used for the electrostatic treatment in the CG model. These two potentials give rise to two different CG models. The first is the regular Coulomb potential from equation 2.9, and the second is the reaction-field potential. The reaction-field potential is a modification of the Coulomb potential in which a constant dielectric environment is assumed beyond a given cutoff. The dielectric constant beyond the cutoff is infinity for the CG model in this project. The force and the electrostatic potential go to zero at the cutoff, which eliminates the need for expensive long-range electrostatic treatments and may provide a more realistic description of an implicit solvent. The reaction-field potential V_{rf} between two particles with charge q_i and q_j is given by

$$V_{\text{rf}} = f \frac{q_i q_j}{\varepsilon_r} \left[\frac{1}{r_{ij}} + k_{rf} r_{ij}^2 - c_{rf} \right] \tag{2.11}$$

where f, k_{rf} and c_{rf} are constants, ε_r is the relative dielectric constant and r_{ij} is the distance between the two particles [12]. The relative dielectric constant of the CG model is set to 80 to mimic the solvent.

A plot of the non-bonded potentials used in the CG model of this project is given in figure 3.

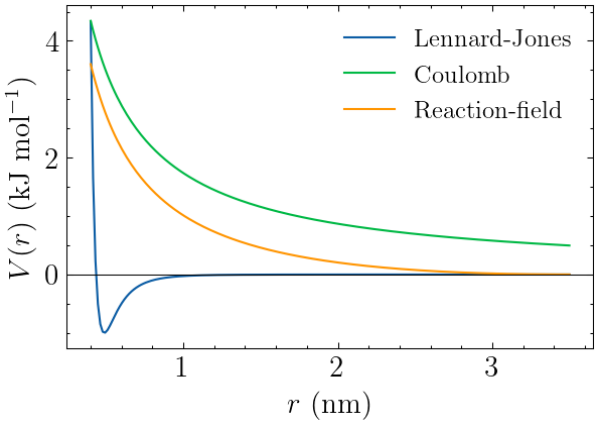


Figure 3: A plot of the Lennard-Jones potential with $\sigma = 0.435$ nm, $\varepsilon = 1.0$ kJ mol^{-1}, and the Coulomb and reaction-field potentials between two particles with equal signed charges. The relative dielectric constant is $\varepsilon_r = 80$. The dielectric constant beyond the cutoff at 3.5 nm is infinity for the reaction-field potential.

2.4 Viscoelasticity

A simple model system is considered in detail in this section to provide an introduction to the subject of viscoelasticity. A differential equation is derived that governs the viscoelastic response of the model to external stimuli. In the end, a route for obtaining the viscoelastic properties from molecular simulations is given.

Viscoelastic materials display both viscous and elastic properties. The Maxwell material is a convenient starting point for a discussion on viscoelasticity. It can be visualized as a spring connected in series with a dashpot and is illustrated in figure 4.

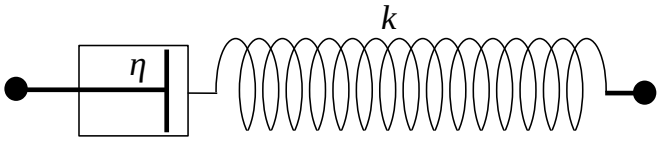


Figure 4: The Maxwell material is an idealized model for a viscoelastic fluid. It is represented by a Hookean spring with spring constant k connected in series to a Newtonian dashpot with viscosity η.

When the right-hand side of the spring is strained (i.e., displaced) by an amount ε, a certain stress σ (i.e. internal force) arises. The spring is assumed to follow Hooke's law, yielding an expression for the stress present in the spring at time t

$$\sigma_k(t) = k\varepsilon_k(t). \tag{2.12}$$

As time progresses, the spring begins to pull the dashpot towards it, represented by a Newtonian fluid. In a Newtonian fluid, the viscous stresses are linearly related to the *strain rate* $\dot{\varepsilon}$ by the viscosity η

$$\sigma_\eta(t) = \eta\dot{\varepsilon}_\eta(t) \tag{2.13}$$

The series configuration of the dashpot and the spring within the Maxwell material demands that the stress in the dashpot and the stress in the spring are at any time equal

$$\sigma = \sigma_k = \sigma_\eta, \tag{2.14}$$

while the total strain is the sum of the strain in each element [20]

$$\varepsilon = \varepsilon_k + \varepsilon_\eta. \tag{2.15}$$

Taking the time derivative of equation 2.15 and plugging in equation 2.12 and 2.13 yields

$$\dot{\varepsilon} = \dot{\varepsilon}_k + \dot{\varepsilon}_\eta = \frac{\dot{\sigma}_k}{k} + \frac{\sigma_\eta}{\eta} = \frac{\dot{\sigma}}{k} + \frac{\sigma}{\eta} \tag{2.16}$$

Assuming that an instantaneous deformation ε_0 takes place at time $t = 0$ such that $\dot{\varepsilon} = 0$, one obtains

$$\dot{\sigma} = -\frac{k}{\eta}\sigma. \tag{2.17}$$

This differential equation has the solution

$$\sigma(t) = \sigma_0 \exp(-kt/\eta) \tag{2.18}$$

where σ_0 is a constant of integration determined by the initial conditions. Normalizing the stress $\sigma(t)$ with respect to the strain ε_0 yields an expression for the stress relaxation modulus $G(t)$, corresponding to how much stress remains in the material at some time t after the sudden deformation. Assuming that the initial response of the Maxwell material to the sudden strain is purely elastic ($\sigma_0 = k\epsilon_0$), the stress relaxation modulus can be expressed as

$$G(t) = k \exp(-kt/\eta). \tag{2.19}$$

$G(t)$ is of central importance in the field of viscoelasticity because it allows predicting all linear viscoelastic properties. For example, the viscosity η is given by

$$\eta = \int_0^\infty G(t)dt. \tag{2.20}$$

The storage modulus, representing the elastic portion of the stored energy in a system, is given by

$$G'(\omega) = \int_0^\infty G(t) \sin(\omega t)dt \tag{2.21}$$

which is directly relatable to gel strength.

The viscoelastic properties of polymer solutions originate at the microscopic level from the dynamics of the polymer chains. Therefore, the so-called Rouse model provides a more realistic description and gives more insight than the simple Maxwell material. The Rouse model is a chain of N beads connected by springs with spring constant

$$k = \frac{3k_BT}{b^2} \tag{2.22}$$

which are moving according to a Langevin equation of motion with friction coefficient ξ [21]. Here, b is the statistical segment length, or Kuhn length, related to the stiffness of a continuous chain and can be interpreted as follows. Several consecutively connected monomers of a real polymer chain may be grouped into segments. If enough monomers along the polymer backbone are grouped into segments, the orientation of these segments will be independent of neighboring segments. The length of the real polymer chain in these segments is then defined by the Kuhn length b [22]. Dividing the polymer into independent Kuhn segments results in an equivalent chain that is much simpler to treat.

The connectivity of the Rouse chain beads gives rise to a set of coupled linear differential equations, which can be uncoupled and solved by matrix diagonalization. The resulting expression for the relaxation modulus has contributions from several different modes p

$$G(t) = \frac{\rho RT}{M} \sum_{p=1}^{N-1} \exp(-2tp^2/\tau_R) \tag{2.23}$$

where ρ is the mass density of the polymer, R is the gas constant, M is the molecular weight and τ_R is the first Rouse mode, i.e. the longest relaxation time. The p'th Rouse mode τ_p is given by

$$\tau_p = \frac{\xi b^2}{3k_B T}\left[4\sin^2\left(\frac{p\pi}{2N}\right)\right]^{-1} \approx \frac{\xi N^2 b^2}{3\pi^2 k_B T}\frac{1}{p^2} \tag{2.24}$$

where the approximation is valid for small p/N [16, 21].

$G(t)$ can be estimated in an MD simulation from the autocorrelation function of the off-diagonal elements σ_{ij} of the (symmetric) stress tensor $\mathbf{P}$

$$G(t) = \frac{V}{3k_B T}\big(\langle\sigma_{xy}(t)\sigma_{xy}(0)\rangle + \langle\sigma_{yz}(t)\sigma_{yz}(0)\rangle + \langle\sigma_{xz}(t)\sigma_{xz}(0)\rangle\big). \tag{2.25}$$

The elements of the stress tensor $\mathbf{P}$ are defined as

$$\sigma_{ij} = \frac{1}{V}\sum_{k=1}^{N} m_k v_{ki} v_{kj} + \frac{1}{V}\sum_{k=1}^{N} r_{ki} f_{kj} \tag{2.26}$$

where m_k is the mass of particle k, and $v_{k\alpha}$, $r_{k\alpha}$ and $f_{k\alpha}$ are the α-components of the velocity, position and total force acting on particle k, respectively [23]. While on the subject, the diagonal elements of $\mathbf{P}$ are used to compute the pressure during an MD simulation.

The time correlation function $\langle A(t)A(0)\rangle$ of a physical quantity A can be defined as the equilibrium ensemble average of $A(t)A(0)$. That is, take the value of $A(0)$ at some time origin $t = 0$ and multiply by $A(t)$ at some later time t, and take the average of all possible time origins during an equilibrium simulation.

Significant memory and computer time requirements complicate the calculation of time correlation functions for systems whose dynamics are characterized by long time scales. Polymer solutions are prime examples of such systems. For example, the longest simulations performed in this project are on the order of 10^8 integration steps, meaning that 10^8 values must be stored for calculating a correlation function in the usual way. Instead, the multiple-tau correlator method [24] allows for efficient and accurate calculations of time correlation functions on the fly during a simulation. Instead of storing all values of some property during a simulation, the multiple-tau algorithm uses an averaging procedure over some interval without significant loss of information. This averaging is performed with varying averaging times and drastically reduces the memory requirements. It has been used to estimate $G(t)$ in systems containing entangled polymers [25] and associating polymers [26].

As an illustration, the autocorrelation function of the CG Martini water model (a single LJ particle) above and below the freezing point is given in figure 5. The properties of the two phases are apparent in the left figure, where $G(t)$ approaches a constant value for the solid and goes to zero for the liquid.

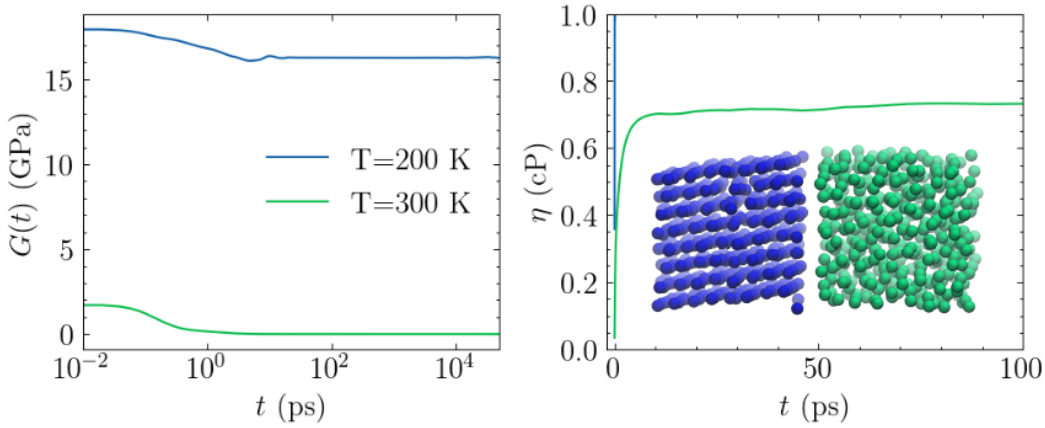


Figure 5: Left) Example of the stress relaxation modulus $G(t)$ estimated from MD simulations for the Martini water model above and below the freezing point. **Right)** The shear viscosity η calculated from equation 2.20. The inset shows a snapshot of the two systems.

3 Methods

Before a CG model can be developed, a suitable atomistic model must be established to form the reference simulations. Two atomistic force fields are used: the CHARMM36-TIP3Pm force field and the GLYCAM06 force field. All parameters, except for the atomic charges of alginates within the GLYCAM06 force field, are available in the literature. Therefore, the first part of this project concerns the development of these missing charges for alginate. Then, the atomistic reference simulations that form the basis of the CG model are described. Following that, the CG parameter development is considered. Two sets of LJ parameters are tested for the CG model: non-bonded parameters set 1 (NB1) and non-bonded parameter set 2 (NB2). NB1 is based on the CHARMM36-TIP3Pm force field and a Coulomb potential, while NB2 is based on the GLYCAM06-TIP3P force field and a reaction-field potential. The two different potentials were plotted in figure 3. Finally, the two parameter sets NB1 and NB2 are applied to large-scale alginate solutions and "hydrogels".

3.1 GLYCAM06 atomic charge calculations for alginates

The electrostatic interactions within the GLYCAM06 force field are described by RESP charges obtained from quantum mechanical calculations. These charges are not available in the literature for alginates and had to be developed. This section describes the charge calculation procedure, which is identical to the general procedure of the GLYCAM06 force field [7] In short, charges for similar sugars are taken as an initial charge set. Then, MD simulations are run with these charge sets, and coordinates are saved at regular intervals. After the simulation, the coordinates of the sugar are fed into quantum mechanical software. Partial charges are calculated for each simulation snapshot, and the ensemble-averaged charges can be calculated by taking the mean. This ensemble-averaged charge set is then used to start another simulation. The partial charges for each snapshot of this new simulation are calculated to yield a final set of ensemble-averaged charges.

3.1.1 Initial parameters

Protein Data Bank files containing the atom types of methylated β-D-glucuronate and methylated α-L-galacturonate were prepared using the online carbohydrate builder [27]. These sugars differ in the position of a single hydrogen atom and a hydroxyl group at a specific carbon atom from the desired monomers of alginate. The epimer of β-D-glucuronate at carbon 2 is β-D-mannuronate (M residue of alginate), while the epimer of α-L-galacturonate at carbon 3 is α-L-guluronate (G residue of alginate). These epimers were the closest resembling structures available, considering their chair conformations. The hydroxyl group and the hydrogen atom attached to the epimeric centers were further interchanged manually to produce the desired methyl glycosides. Their structures are given in figure 6.

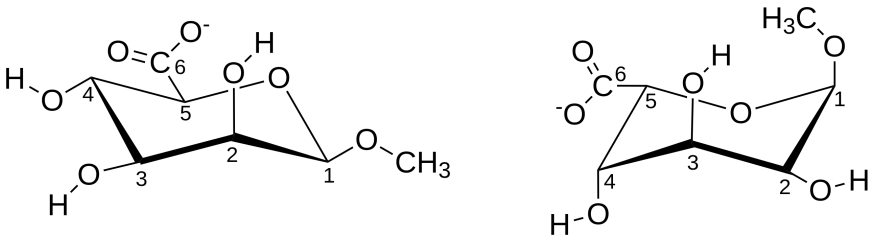


Figure 6: The structures from which partial atomic charges for alginate are developed. **Left)** Methyl β-D-mannuronate. **Right)** Methyl α-L-guluronate.

The online produced Protein Data Bank files were processed with *tleap* from AmberTools [28] to generate the GLYCAM06 parameters. They were further converted to GROMACS format using ACPYPE [29], and the resulting parameters served as the initial charge sets. As mentioned above, these charge sets correspond to the sugars galacturonate and glucuronate instead of guluronate and mannuronate. However, using the epimer charge sets as initial guesses is justified because ensemble-averaging can absorb some inaccuracies in the initial charge sets [30].

3.1.2 Equilibrium simulations

Each methyl glycoside was solvated in a periodic box using the TIP3P water model [31], and a single sodium ion was added to neutralize the charge. A straight non-bonded cutoff of 1.2 nm was used to truncate the short-ranged LJ interactions. Long-range electrostatics were treated with the Particle-Mesh Ewald (PME) method [32]. Dispersion corrections were applied to the energy and pressure. The equations of motion were integrated with a leap-frog algorithm using a timestep of 2 fs, and all bonds to hydrogen were constrained using LINCS [33]. The system's energy was first minimized using the steepest descent method, after which an equilibration of 100 ps was performed in the NVT ensemble. The temperature was maintained at 298 K using the canonical sampling through velocity rescaling (CSVR) thermostat [34] with a time constant of 0.1 ps. The system was further simulated for 210 ns in the NPT ensemble using Parrinello-Rahman pressure coupling [35] with a time constant of 2 ps to maintain the pressure at 1 bar. The first 10 ns served as NPT equilibration, and the last 200 ns were used to extract 200 evenly spaced snapshots of the methyl glycosides.

3.1.3 RESP charges

Quantum mechanical calculations were performed with NWChem [36] (version 7.0.0). The 200 extracted equilibrium coordinates of the methyl gylcosides were used as initial structures for geometry optimizations. All dihedral angles were constrained to their MD snapshot values during optimization. To be compatible with GYLCAM06, the geometry optimizations are performed at the HF/6-31G* level of theory. The molecular electrostatic potential calculations are then performed at the HF/6-31G* level of theory on the optimized geometries [7]. This geometry optimization and subsequent electrostatic potential calculation with the given basis sets is referred to as HF/6-31G*//HF/6-31G*. The 6-31G* basis set is generally considered an excellent choice for reproducing the properties of biomolecular systems in aqueous solutions due to fortunate error cancelling [13]. RESP charges were fitted using a grid spacing of 0.02 nm and a hyperbolic restraint value of 0.01. Charges on hydrogen atoms connected directly to carbon were constrained to

zero, and the charges of the two carboxylate oxygens were constrained to be equal. The oxygen and carbon atom of the methoxy group were constrained to -0.458 and 0.264, respectively. These are default values within the GLYCAM06 force field for the sugar linkages in alginates. The python script used for generating the NWChem input files is available at `https://github.com/lukasbaldauf/ensemble-average-q`. An intermediate ensemble averaged charge set was derived from the mean of these 200 charge sets. The whole simulation procedure (equilibrium simulation, geometry optimizations and RESP charge calculations) was repeated using the newly derived charges to obtain a final set of partial atomic charges.

3.2 All-atom reference simulations

This section describes the atomistic reference simulation setups that form the basis of the coarse-grained model.

Six systems were prepared to develop all bonded parameter combinations. Each system contained an oligomer with degree of polymerization DP=7, neutralized with counterions (sodium or chloride). To capture all possible MG-block parameter combinations, five out of the six systems were alginates with the composition given in table 3.

Table 3: The oligomeric systems for which bonded parameters are developed, and their abbreviations. Y refers to a chitosan unit, G a guluronate unit, and M a mannuronate unit.

7CH: `Y-Y-Y-Y-Y-Y-Y`	7MM: `M-M-M-M-M-M-M`	7GG: `G-G-G-G-G-G-G`
7MG: `M-G-M-G-M-G-M`	`M-M-G-G-G-M-M`	`G-G-M-M-M-G-G`

The systems in table 3 were simulated with the GLYCAM06 forcefield, such that the GLYCAM06 force field formed the basis of the bonded CG parameters. The GLYCAM06 force field is specifically targeted toward the modeling of carbohydrates, and the parameters have been extensively validated against experimental NMR data of charged sugars [11, 37].

The choice of DP=7 was based on a compromise between computational efficiency and avoiding end effects. The computational efficiency depends on heavily the chain length when a minimum image convention is used, as the number of solvent molecules increases with the cube of the polymer length. However, the chain length should be long enough to obtain conformational distributions and inter-molecular interactions that are representative of the target molecules.

Regarding the non-bonded parameter development, seven different systems, each containing two oligomers with DP=7, formed the AA reference simulations:

Table 4: The oligomeric systems for which non-bonded parameters are developed, and their abbreviations. Y refers to a chitosan unit, G a guluronate unit, and M a mannuronate unit.

2-7CH: Y-Y-Y-Y-Y-Y-Y Y-Y-Y-Y-Y-Y-Y	2-7MM: M-M-M-M-M-M-M M-M-M-M-M-M-M	2-7GG: G-G-G-G-G-G-G G-G-G-G-G-G-G	2-7MG: M-G-M-G-M-G-M M-G-M-G-M-G-M
7MM-7CH: M-M-M-M-M-M-M Y-Y-Y-Y-Y-Y-Y	7GG-7CH: G-G-G-G-G-G-G Y-Y-Y-Y-Y-Y-Y	7MG-7CH: M-G-M-G-M-G-M Y-Y-Y-Y-Y-Y-Y	

Each of the 7 systems in table 4 was simulated with both the CHARMM36 and GLYCAM06 force fields, such that two sets of different non-bonded parameters were developed for the CG model. These two sets of non-bonded parameters are denoted NB1 for the CHARMM36 force field and NB2 for the GLYCAM06 force field.

The GLYCAM06 force field has been found to overestimate solute-solute interactions in some applications, while the CHARMM36 force field may underestimate them [38–40]. Reassignment of the LJ potential-depth has been found to improve agreement with experiments [41, 42]. However, the literature is scarce on relevant experimental data for alginate and chitosan oligomers, which prevents validating such a rescaling procedure. It has also been suggested to use TIP5P water in combination with the GLYCAM06 forcefield [38]. However, the resulting RDFs resembled the CHARMM36 RDFs but at an increased computational cost. The coarse-grained non-bonded parameters were therefore based on two force fields; GLYCAM06/TIP3P, which may overestimate the solute-solute interactions, and CHARMM36/TIP3Pm, which may underestimate them. The real situation may then lie somewhere between these two.

3.2.1 CHARMM36 setup

Simulations with the CHARMM36 force field formed the basis of the first set of non-bonded parameters for the CG model, denoted NB1. The systems considered were given in table 4.

The CHARMM36 parameters were generated using the glycan modeler [43] provided by CHARMM-GUI [44] by building the 1→4)-linked epimers β-D-glucuronate (epimer of β-D-mannuronate at C2) and α-L-iduronate (epimer of α-L-guluronate at C2). The parameters for the uronates are identical within the CHARMM36 force field. Therefore, interchanging the positions of the hydroxyl group and the hydrogen atom attached to the epimeric centers produced the desired input file. The chitosan were directly generated using the CHARMM-GUI. The CHARMM36 parameters were further converted to GROMACS format using the force field converter tool [45].

The simulation setup was as follows. A time step of 2 fs was used to integrate the equations of motion. A cut-off of 1.2 nm was used for non-bonded interactions. Long-range electrostatics were treated with PME. A force-switched smoothing function was applied to Lennard-Jones (LJ) interactions between 1.0 and 1.2 nm. All systems were subject to 1000 steps of steepest-descent energy minimization, followed by 250 ps NVT-equilibration using the CSVR thermostat [34] with a time constant of 0.1 ps. A time

constant of 0.2 ps for the CSVR thermostat was used for simulations in the NPT ensemble, additionally using a Parrinello-Rahman barostat [35] with a time constant of 2 ps for a total of 405 ns. The first 5 ns served as equilibration. Bonds to hydrogen were constrained using LINCS [33]. All systems were solvated in the modified version of the original TIP3P water model in which LJ interactions are added to the hydrogens [46]. The oligomers present in all systems were neutralized individually by adding the corresponding number of sodium or chloride ions. The temperature was set to 300K and the pressure to 1 bar.

Production simulations were run for 400 ns, and coordinates were extracted every 20 ps. The RDFs were calculated and used as target properties for the non-bonded parameterization of the CG model.

3.2.2 GLYCAM06 setup

The GLYCAM06 simulations formed the basis of the CG bonded parameterization in addition to the second set of non-bonded parameters, denoted NB2.

The GLYCAM06 force field parameters for alginates were generated as described in section 3.1, while the parameters for chitosan were taken from [37]. The oligomeric systems simulated are listed in tables 3 and 4. Each system was solvated in TIP3P water and neutralized with counterions (sodium or chloride). Production simulations were run for 200 ns and 400 ns for bonded and non-bonded parameters, respectively, with the setup from section 3.1.2. Coordinates were saved every 20 ps and served as target properties for the CG model bonded parameter development.

PMFs calculations between chitosan and alginate oligomers of different compositions were performed with GPUs on the Idun cluster [47]. The three systems considered were 4GG-4CH, 4MG-4CH, and 4MM-4CH. These three systems were solvated (without counterions) in a rectangular box that extended in the x-direction. Each alginate oligomer was placed at a fixed position and held in place during all simulations by harmonic position restraints on the ring atoms. The chitosan oligomer was placed at increasing distances from the alginate. 24 windows with a minimum COM distance of 0.5 nm and a maximum COM distance of 3.2 between alginate and chitosan were simulated for 10 ns each. Each window was equilibrated during 50 ps in both the NVT and NPT ensemble with position restraints on alginate and chitosan. A harmonic potential kept the alginate and chitosan oligomers separated at the pre-set distance during all simulations. The harmonic potential was described by a force constant of 500 kJ mol^{-1} nm^{-2} acting along the x-direction. Each atom experienced the force weighted by its mass fraction in the molecule.

PMFs were then calculated using the weighted histogram analysis method (WHAM) implemented in GROMACS. Errors were estimated using Bayesian bootstrapping [48].

3.3 Coarse-grained parameter optimization

This subsection describes the development of an implicit solvent coarse-grained (CG) model based on the atomistic reference simulations of oligomers consisting with DP=7. It was heavily inspired by the the martini force field [49, 50], its extension to carbohydrates [19], and the dry-martini force field [51]. Some effort is put into the transferability of the model to the Martini 3 force field with an explicit solvent.

The development of bonded parameters is similar to the Martini procedure [50] in which bonded parameters are optimized to reproduce atomistic distributions mapped to CG coordinates.

The non-bonded parameters were optimized to best reproduce atomistic radial distributions functions (RDF) mapped to CG coordinates, PMFs between alginate and chitosan, and polymer solution properties. Two sets of non-bonded parameters were developed, one based on the CHARMM36 force field and one based on the GLYCAM06 force field. This procedure differs from the MARTINI forcefield in which CG beads are described by specific building blocks with fixed interactions [50].

3.3.1 Mapping scheme and bonded parameters

The mapping scheme, i.e., the transformation between the AA and CG coordinates, is based on a compromise between specificity and computational performance. The mapping scheme used is identical to that used for carbohydrates in the Martini 3 forcefield. On average, the center-of-geometry of 4 heavy atoms and their connected hydrogens represent a single CG bead.

Bond lengths, angles and dihedral parameters of the CG model are updated iteratively to match as closely as possible the bond length, angle and dihedral distributions of the mapped atomistic trajectories. These parameters were based on the GLYCAM06 AA simulations. The CG setup was as follows. Force field parameter for the oligomers from table 3 were generated. The oligomers were further centered in a dodecahedron box and neutralized with 7 sodium or chloride ions. The rest of the simulation setup was identical to that used in section 3.3.2 below. The bonded distributions were rather insensitive to the electrostatic treatment and the non-bonded parameters used. Therefore, almost indistinguishable distributions were obtained with NB1, NB2, and the crude inital non-bonded parameters. The inital simulations were relatively short, and the duration was increased as better agreement between CG and AA distributions were obtained. The final CG simulations were run for 500 ns.

3.3.2 Non-bonded parameter set 1

Once the bonded parameters adequately reproduced the AA bonded distributions, a set of non-bonded parameters based on the CHARMM36 force field was developed. These non-bonded parameters were determined by directly fitting the CG RDFs to the mapped AA RDFs.

The equations of motion were integrated with a time step of 20 fs using the stochastic dynamics integrator [52] at a temperature of 300K. The stochastic dynamics integrator adds a friction term and a stochastic noise term to the regular equations of motion (eq. 2.1). The friction constant was set to 1/4 ps^{-1}, which was also chosen in the dry-martini forcefield based on a compromise between not imposing too much friction and maintaining effective temperature control [51]. No pressure coupling is applied for the CG systems. All CG simulations are therefore performed in the NVT ensemble. Equilibration simulations were performed by first relaxing the systems during 10'000 steps with a timestep of 1 fs and a friction constant of 1/0.1 ps^{-1}, followed by 10'000 steps with the regular timestep of 20 fs and a friction constant of 1/4 ps^{-1}. The dielectric constant was set to 80, which mimics the solvent (water). Non-bonded interactions were excluded between neighbouring atoms that were no further than 3 bonds away, i.e. all atom pair interactions up to and including 1-4 atom pairs are excluded. This is common practice for sugars within the Martini mapping scheme, as the 1-4 beads often overlap for polysaccharides. Lennard-Jones interactions were cut off at 1.2 nm. Long range electrostatic interactions were treated with the PME method. The real space coulomb cutoff

(*rcoulomb* in GROMACS) was set to half the shortest box length for small systems. For larger systems *rcoulomb* was set such that the computational load of the PME mesh part was around 0.5 (around 7 nm for the largest systems). The fourier grid size was set by *fourierspacing*, which was taken to be 1/10 of *rcoulomb*. The neighbour list was set to *rcoulomb* + 0.2 nm and updated every 10 steps.

Lennard-Jones potentials were initially assigned by brute force optimization of an objective function defining the goodness of a parameter set. The Nelder-Mead method was used to iteratively perform CG simulations with a given set of non-bonded parameters **p**. The accuracy or goodness of the parameter set $\boldsymbol{p}$ was defined by the error integral [53]

$$f(\boldsymbol{p}) = \int_0^2 \left[g_{\mathrm{CG}}(r, \boldsymbol{p}) - g_{\mathrm{AA}}(r)\right]^2 dr \tag{3.1}$$

where g_{CG} and g_{AA} are the RDFs calculated from the CG trajectory and the atomistic trajectory mapped to CG coordinates, respectively, and r is is given in units of nm. One such integral can be calculated for all bead type combinations, and the total accuracy was taken as the sum of each contribution.

However, this procedure was rather slow for the large number of particle type combinations, and a trial and error procure was adopted. Simulation durations were increased as the match betewen CG and AA RDFs improved. The final CG RDFs were calculated from simulations lasting 400 ns for each of the systems in table 4.

The development of the first non-bonded parameter set (which also formed the initial parameter set of NB2) was one of the most challenging and time-consuming parts of the parameterization procedure and the whole project. Moreover, this assignment process is very prone to error, even for the parameterization of the limited number of molecules considered in this project.

First, the non-bonded interaction treatment must be chosen. Many CG models in the literature use tabulated potentials [10, 54]. These tabulated potentials are flexible and can accurately reproduce atomistic RDF by using iterative Boltzmann inversion or inverse Monte Carlo [14, 55], which iteratively update the pair-interactions during successive CG simulations based on specific criteria. However, tabulated potentials are not implemented in the GROMACS versions used in this project. Older versions support tabulated potentials, but the computational performance was then drastically reduced due to the long-ranged cut-off necessary to model long-ranged electrostatic interactions accurately. Therefore, a regular Lennard-Jones interaction potential was chosen. It could be argued that a more flexibe MD software such as LAMMPS would be productive at this point, but the GROMACS engine was significantly faster for these simulations.

The selection of target properties is also non-trivial. Target features in the literature often include radial distribution functions obtained from simulations or reproducing thermodynamic properties from experiments. However, thermodynamic data for the polyelectrolytes considered in this project is scarce and usually limited to large polymeric systems that are challenging calculate using molecular simulations. Therefore, although validated against conformational properties from NMR and quantum mechanical calculations [8, 11, 37], the underlying atomistic force fields remain largely unvalidated against condensed phase properties for the systems relevant to this project.

After selecting a non-bonded interaction treatment, a particle type assignment must be chosen for all CG beads in the system. A large number of particle types quickly results in huge interaction tables with many parameters. Using many particle types

reproduced the atomistic properties well, but different particle type assignments with very different non-bonded parameters resulted in equally good agreements. Therefore, identical particle types were assigned to the polymer backbone beads, and all interactions were discarded initially by starting with a "non-interaction" system with approximately repulsive potentials between all beads. Then, an attractive interaction was only included if it significantly improved the fit between the CG and AA RDFs.

3.3.3 Non-bonded parameter set 2

A more specific parameterization procedure is used to develop NB2, which is based on the GLYCAM06 force field. Firstly, the particle types SQ5n and SP4 within an alginate dimer (GG, MM, MG, or GM) are now described by bead types specific to that residue in that dimer. For example, the carboxylate groups on G residues connected to M residues are assigned bead type SQ5n(GM). Secondly, in addition to atomistic RDFs, two other properties served as targets in the non-bonded parameterization; PMF between alginates and chitosan, and the aggregation behavior of large alginate chains.

The CG potentials of mean force were calculated as described in section 3.2.2, with position restraints placed on the B2 atoms of alginate. Each window was equilibrated during 100 ps and thereafter simulated for 200 ns.

The aggregation behavior was screened for 50 pre-equilibrated 300GG chains neutralized with sodium ions (10 g L^{-1} and 300 K) over the course of 50 ns. This aggregation behavior is qualitative in nature and is determined by visual inspection of the simulation snapshots. The poly-G systems in [2] were soluble at a concentration of 1%, and the interactions between the 300GG alginate and sodium ions are therefore tuned such that no chain aggregation is observed during the simulation. The pre-equilibrated structures were those described in section 3.4.1 with NB1 after 1.5 μs of simulations.

The simulation setup for NB2 differed in the treatment of electrostatic interactions, which were modeled by a reaction-field potential instead of a Coulomb potential. The reaction-field potential is a modification of the coulomb potential in which a constant dielectric environment is assumed beyond a given cutoff. The dielectric constant beyond the cutoff distance at 3.8 nm was set to infinity, such that the force and the electrostatic potential goes to zero at the cutoff. Treating electrostatic interactions with the reaction-field potential instead of a coulomb potential avoids the long-range behavior of CG PMFs between oppositely charged molecules (alginate and chitosan), which was not observed in AA PMFs. The neighbour list was extended to 4.2 nm and updated every 20 steps. The remaining setup was identical to the setup of NB1 in section 3.3.2.

3.4 Application polymer solutions

This section describes the application of the developed CG model to polymer solutions consisting of 50 alginate chains with DP=300 in cubic boxes with dimensions 79.1 nm $\times$ 79.1 nm $\times$ 79.1 nm. This yields a polymer concentration of 1% (10 g L^{-1} with monomer mass 199 g mol^{-1}). An identical setup was used to simulate chitosan oligomer solutions at a concentration of 0.3%. Chitosan systems with DP=4 contained 1128 oligomers and systems with DP=8 contained 564 oligomers. A cubic box was chosen such that the Z1-code [56–59] can be used for chain entanglement calculations. Polymer entanglements originate from the uncrossability of the alginate chains. A chain length of 300 was chosen for alginate to avoid excessive aggregation after adding chitosan, which was observed

for chains of length 150 and 200. Three systems are considered; 300GG, 300MM, and 300MG, where 300MG strictly alternates between G- and M-block residues. All alginate and chitosan residues are assumed to be ionized and are neutralized with sodium ions or chloride, respectively. The systems are then simulated with the setups of NB1 and NB2 described previously. Static and dynamics properties are calculated from the simulations. The static properties can be directly compared to experiments, which serves as a validation step.

The static properties calculated include the persistence length q and the mean squared radius of gyration $\langle R_{\mathrm{G}}^2 \rangle$. The persistence lenght $q = b/2$, where b is the Kuhn length, may be determined from an equilibrium simulation of these large systems by calculating the number of bonds at which the average cosine between bond vectors reaches a value of 1/e, multiplied by the average bond lengths l for the respective alginate residues. Table 6 gives l for the different alginate blocks. The mean squared radius of gyration $\langle R_{\mathrm{G}}^2 \rangle$ for polymer with any molecular weight may then be estimated using the "Worm-like" chain model

$$\langle R_{\mathrm{G}}^2 \rangle = \frac{qM}{3M_L} - q^2 + \frac{2q^3 M_L}{M}\left[1 - \frac{qM_L}{M}\left(1 - \exp\left(-\frac{M}{qM_L}\right)\right)\right] \tag{3.2}$$

where

$$M_L = \frac{M}{Nl} \tag{3.3}$$

with N being the number of monomers in a chain and l the bond length between monomers [1, 60, 61].

The multiple-tau correlator was used to estimate $G(t)$ and was called sequentially every 1'000'000 steps. The pressure tensor elements were saved to disk every timestep during each of these 1'000'000 steps. When reaching 1'000'000 steps, the elements were pushed to the correlator and discarded thereafter to save disk space. After the simulation finished, the shear stress relaxation modulus $G(t)$ was calculated (equation 2.25) from the values in the correlator. The Python implementation of the correlator is available at `https://github.com/lukasbaldauf/multiple-tau`, which is a Python translation of the code provided in [62]. The Z1-code was used to calculate the average number of entanglements from the simulation snapshots every 200 ns.

The calculated moduli $G(t)$ can be compared to those obtained from the Rouse model (equation 2.23 without the approximation). The Rouse time τ_R (equation 2.24 with $p = 1$) was calculated as follows. b was taken to be $2\langle q \rangle$, where $\langle q \rangle$ is the average persistence length obtained during the simulation. It was again taken to be number of bonds at which the average cosine between bond vectors reaches a value of 1/e (denoted N_q), multiplied by the average bond lengths l for the respective alginate residues. The values for l were listed in table 6. N in equation 2.24 was then taken to be $300/N_q$, because the alginate chains have length 300. The friction constant ξ was set to 1/8 ps^{-1} for NB1 (to speed up the dynamics) and 1/4 ps^{-1} for NB2. The prefactor in equation 2.23 was calculated using $M = 176$ g mol^{-1}, which is the monomer mass of alginate less the sodium ion. This yields a prefactor of 418.45 Pa for 50 alginate chains with length 300. The relaxation modulus $G(t)$ can then finally be obtained from equation 2.23 by summing over $p = 1$ up to some large value (e.g. $p = 1$ to $p = 10000$).

The simulations in this section were performed on the Idun cluster [47] using GPUs. The computational performance averaged around 550 ns and 650 ns for NB1 and NB2,

respectively. This was reduced to around 450 ns per day after chitosan was added (section 3.5). The large-scale simulations described in this and the next section were carried out towards the end of this project, and time became a limiting factor for investigations of other systems.

3.4.1 NB1

Energy minimized straight alginate chains were introduced at random positions in the simulation box. An equilibration run of 1.5 μs was then performed using the simulation parameters described in section 3.3.2. Production runs lasting 1.2 μs were then performed from which the static properties were calculated. After these 2.7 ns, simulations involving $G(t)$ were initialized. The setup from section 3.3.2 was used, except that the friction coefficient was reduced from 1/4 ps^{-1} to 1/8 ps^{-1} for the simulation involving $G(t)$ to speed up the dynamics.

3.4.2 NB2

The structures obtained after 1.5 μs of equilibration and 1.2 μs of productions runs with NB1 were used as initial structures in this step. First, around 1.3 μs of equilibration runs were performed, after which around 1.8 μs of production runs were performed to estimate the static properties and $G(t)$. The rest of the simulation parameters were given in section 3.3.3.

3.5 Application to hydrogels

3.5.1 NB1

The final snapshots after 2.7 μs from section 3.4.1 were used as initial coordinates. Chitosan oligomers with DP=8 or DP=4 were then introduced in random non-overlapping positions at a concentration of 3 g L^{-1}, which was the experimental concentration of chitosan in [2].

The chitosan cluster size and the alginate network size was quantified using the *AgglomerativeClustering* method provided by Scikit-learn [63]. The clustering was performed every 10 ns over the last 600 ns of 1.6 μs trajectories with the setup described in section 3.3.2. For systems with length 8 chitosan oligomers, two additional simulations were performed lasting 800 ns with chitosans inserted at independent positions, and clustering is performed on the last 200 ns every 10 ns.

The chitosan cluster size was determined as follows. The distance matrix between chitosan B2 beads (see figure 10) was calculated using MDAnalysis [64, 65] and used for agglomerative clustering with a single linkage and distance threshold of 2.0 nm. Thus, two chitosan oligomers belong to the same cluster if any of their B2 beads are within 2.0 nm of each other.

The alginate network size was calculated as follows. First, for each chitosan B2 bead, the indices of the alginate chains within 1.2 nm of the B2 chitosan bead are obtained. If a chitosan B2 bead is within 1.2 nm of alginate chains i and j, the value of a (symmetric) cross-link matrix $\mathbf{C}$ at element ij is updated by 1. The elements c_{ij} of $\mathbf{C}$ quantify the number of crosslinks between alginate chains i and j, and a chitosan oligomer with DP=4 can contribute with a maximum of 4 crosslinks. The diagonal elements of the matrix are set to zero. The networks size is then estimated from -$\mathbf{C}$, again using agglomerative

clustering. This is because the negative crosslink matrix -**C** in a sense quantifies how tightly two chains i and j are bound, with large negative elements implying that elements i and j are heavily crosslinked. Therefore, the clustering method is called iteratively, starting with a single linkage distance threshold of -1 and stopping at the most negative element of -**C**, i.e. the negative of the largest number of crosslinks between two chains. This first iteration thus gives the number of clusters in the system that are crosslinked by at least 1 chitosan monomer. For example, 11 out of the 12 alginate chains may at some time t be crosslinked together in a network, while the last of the 12 chains is not surrounded by any chitosans and drifts freely in the solution. This results in 2 alginate clusters, one cluster with 11 crosslinked chains, and one cluster for the single chain drifting freely within the network formed by the 11 other chains, as illustrated in figure 7.

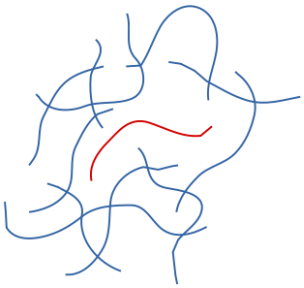


Figure 7: An example of 12 alginate chains, 11 of which are crosslinked in a network. The last chain is not crosslinked and is drifting freely within the cavity of the network. The chitosan oligomers between crosslinks are not shown.

The final gel-network size N_{network} is calulated from a weighted average

$$N_{\text{network}} = \sum_{k=1}^{n} N_k \frac{N_k}{N_t} \tag{3.4}$$

where N_k is the number of alginate chains in cluster k, N_t is the total number of chains in the system, and the sum runs over all n clusters of the system. This average was inspired by the weight average molecular weight, which is often used to characterized polymer masses. The network size from the above example would then be $1\frac{1}{12} + 11\frac{11}{12} \approx 10.2$. It should be noted that crosslinks between the same alginate chains through the periodic boundaries are not counted.

Averaged binding energies $\Delta\langle E\rangle$ between chitosan and alginates were estimated using

$$\Delta\langle E\rangle = \langle E\rangle_{\text{Alg-Chit}} - \left(\langle E\rangle_{\text{Alg}} + \langle E\rangle_{\text{Chit}}\right) \tag{3.5}$$

where $\langle E\rangle_{\text{Alg}}$ and $\langle E\rangle_{\text{Chit}}$ are the averaged total energies of the pure alginate and chitosan solutions calcualted from the last 20 ns, respectively. $\langle E\rangle_{\text{Alg-Chit}}$ is the averaged total energy of the alginate-chitosan mixture during the last 20 ns.

3.5.2 NB2

The final snapshots after 2.7 μs from section 3.4.1 were also used as initial coordinates in this step. Chitosan oligomers with DP=8 were then introduced in random non-overlapping positions at a concentration of 3 g L^{-1}. Around 800 μs of simulations were performed with the setup described in section 3.3.3. No further analysis was carried out, except for visual inspection.

4 Results and discussion

This section first discusses the microscopic insights obtained from the partial charge calculations for alginate and the atomistic reference simulations. Next, the results of the coarse-grained parameterization are presented and discussed. Using two force fields for the atomistic reference simulations allowed the development of two coarse-grained models, NB1 and NB2, which differ in the LJ potentials between particles and the treatment of non-bonded interactions. A python script that generates the correct GROMACS force-field parameters for any desired alginate or chitosan polymer with NB1 is available at `https://github.com/lukasbaldauf/topogen`. These two coarse-grained models were further used to investigate large-scale alginate solutions and hydrogels, which would be impractical to study at full atomistic resolution. The results from the alginate solution simulations are compared to experiments as a validation step. Finally, the results from the hydrogel simulations are presented and discussed in light of the experimentally observed behavior of the corresponding systems in [2].

4.1 GLYCAM06 atomic charge calculations for alginates

Partial charges for the mannuronate and guluronate residues found in alginates are calculated according to the method outlined in the standard GLYCAM06 paper [7]. A Python script for generating the NWChem input files is available at `https://github.com/lukasbaldauf/ensemble-average-q`. The final charges are visualized in figure 8 and listed in table 5. Similar charge values for the hydroxyl hydrogen atoms (H2O, H3O, H4O*) and the carboxylate groups (C6, O6A, O6B) are evident. However, the hydroxyl oxygen atoms (O2, O3) on guluronate carry a greater charge than the corresponding oxygen atoms on mannuronate. The greater charge on these two oxygen atoms may contribute to a more favorable counterion interaction.

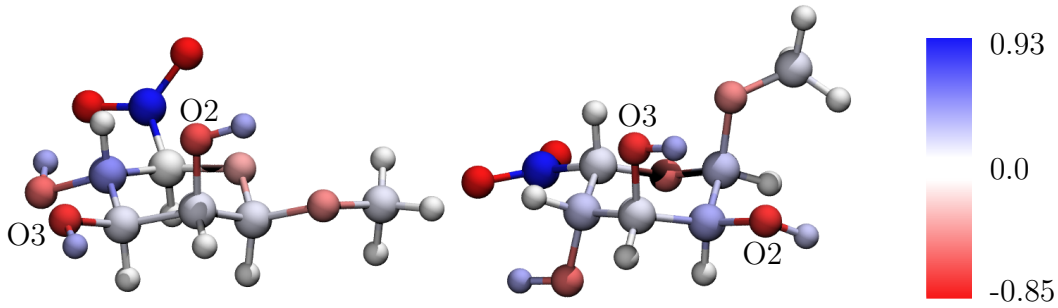


Figure 8: Second iteration ensemble-averaged RESP charges computed at the HF/6-31G*//HF/6-31G* level of theory. Atoms are colored according to charge value. The charges were obtained by averaging 200 charge sets obtained from a 200 ns equilibrium simulation. **Left)** methyl β-D-mannuronate. **Right)** methyl α-L-guluronate.

Table 5: Ensemble averaged RESP charges for methyl β-D-mannuronate (MAN) and methyl α-L-guluronate (GUL) in TIP3P water. Hydrogens connceted directly to carbon have zero charge. The values of methyl α-L-galacturonate (GAL) were taken directly from the GLYCAM06 parameter set and are given for comparability. See figure 9 below for naming conventions.

Atom	MAN	GUL	comment	GAL
CMe	0.264	0.264	constrained	0.264
OMe	-0.458	-0.458	constrained	-0.458
C1	0.264	0.356		0.305
O5	-0.374	-0.514		-0.424
C5	0.016	0.216		0.054
C6	0.931	0.906		0.896
O6B	-0.844	-0.850		-0.822
O6A	-0.844	-0.850		-0.822
C4	0.513	0.325		0.267
O4*	-0.767	-0.753	terminal oxygen	-0.704
H4O*	0.409	0.401	terminal hydrogen	0.429
C3	0.210	0.250		0.251
O3	-0.705	-0.750		-0.725
H3O	0.402	0.417		0.416
C2	0.232	0.410		0.413
O2	-0.653	-0.768		-0.771
H2O	0.405	0.399		0.431
O4	-0.552	-0.546	= O4* + H4O* + OMe + CMe	-0.469
H1O	0.445	0.445	glycam default	0.445
O1	-0.639	-0.639	glycam default	-0.639

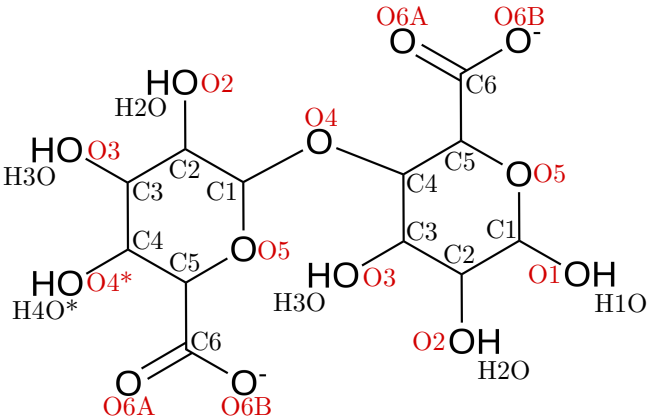


Figure 9: The atom naming convention within the GLYCAM06 force field. Hydrogens directly attached to carbon carry no charge and are not shown.

4.2 Coarse-grained parameters

This section lists the optimized coarse-grained bonded parameters and the non-bonded parameters for NB1 and NB2.

4.2.1 Mapping scheme and bonded parameters

The mapping scheme is illustrated in figure 10. The atom type naming conventions are chosen such that the force field parameters listed may be directly transferred to the Martini 3 force field simply by exchanging the non-bonded parameter file.

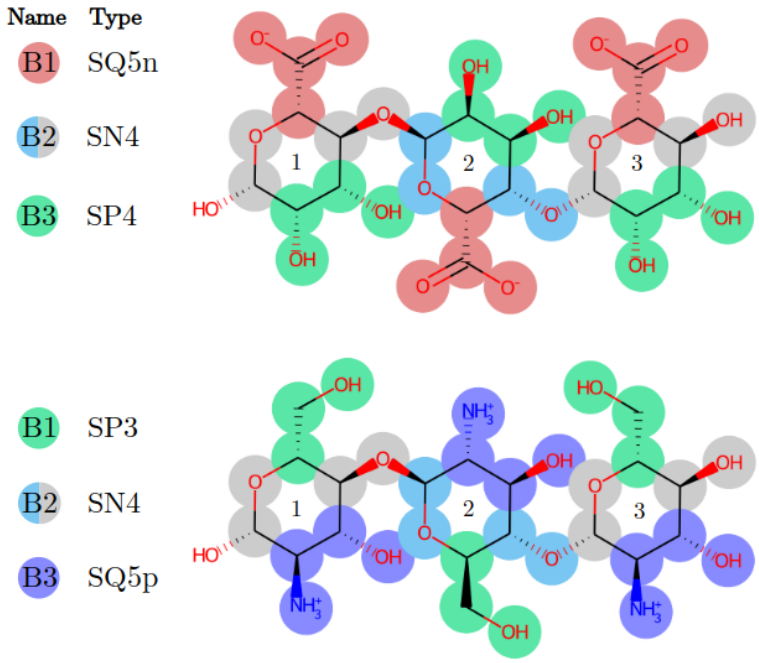


Figure 10: **Top)** The mapping scheme for alginate. **Bottom)** The mapping scheme for chitosan. The CG beads are position on the center of geometry of the underlying AA beads. Hydrogen atoms attached to carbon are not shown, but their contribution to the center of geometry of the respective coarse-grained beads is considered.

As an example, the atomistic structure of 7MG and the mapped CG structure is given in figure 11.

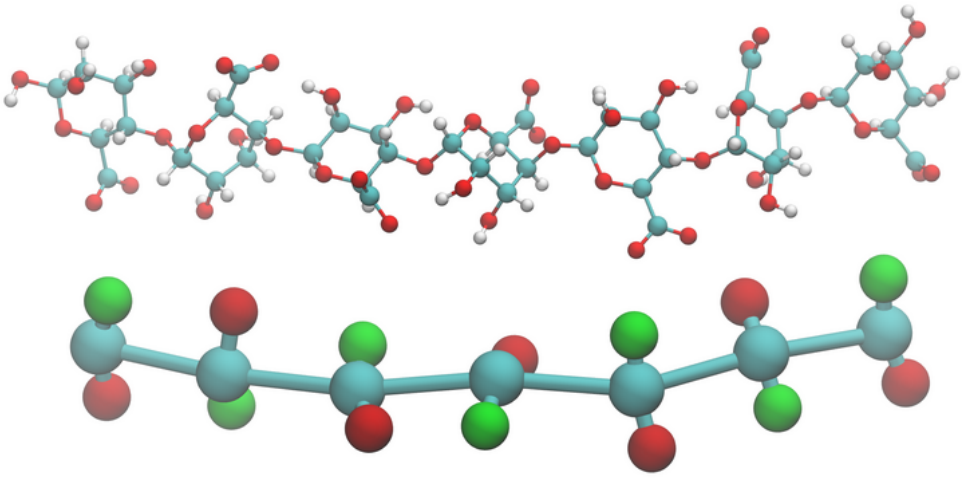


Figure 11: The final snapshot of 7MG after 200 ns simulated with the GLYCAM06 force field and the corresponding center-of-geometry mapped CG structure.

The mapping scheme is fine enough to reproduce essential polymer properties for the present systems, such as the orientation of functional groups along the backbone. This orientation is vital for the interaction between alginate and chitosan and the chelation of sodium ions with specific alginate residues (GG and alternating MG blocks). Ideally, back mapping from CG to AA coordinates should also be possible without too much loss of information. At the same time, the mapping scheme is coarse enough for a significant increase in computational performance. The greatest increase in efficiency is achieved by replacing the solvent molecules with an implicit solvent and merging atoms into larger beads. Additionally, grouping multiple atoms into a single CG bead allows for a larger integration time step by reducing the oscillation frequencies of stiff bonds.

The final bonded parameters that most accurately match the mapped GLYCAM06 bonded distributions are given in table 6. A comparison between the mapped AA and the optimized CG bonded distributions is shown in figure 22-25 in the appendix.

Table 6: Optimized bonded parameters for the implicit solvent coarse-grained model. The corresponding GROMACS function types are; constraints: 1, bonds: 1, angles: 2, dihedrals: 9 with multiplicity 1. The mapping direction is defined from left to right when the reducing end of the carbohydrate is on the left, as illustrated in figure 10. GROMACS parameters can be generated for any alginate or chitosan at `https://github.com/lukasbaldauf/topogen`.

bond	b_0 (nm)	k_b (kJ mol^{-1}nm^{-2})
G1-G2	0.240	constraint
G2-G3	0.216	constraint
M1-M2	0.221	constraint
M2-M3	0.203	constraint
Y1-Y2	0.248	constraint
Y2-Y3	0.243	constraint
G2-G2	0.453	12000
M2-M2	0.537	24000
M2-G2	0.504	23000
G2-M2	0.512	13000
Y2-Y2	0.532	28000

dihedral	ϕ_s (deg)	k_ϕ (kJ mol^{-1})
G1-G2-G2-G3	-120	20
G3-G2-G2-G1	162	20
M1-M2-M2-M3	-107	20
M3-M2-M2-M1	-171	20
M1-M2-G2-G3	-140	20
M3-M2-G2-G1	145	20
G1-G2-M2-M3	-91	20
G3-G2-M2-M1	152	20
Y1-Y2-Y2-Y3	-113	20
Y3-Y2-Y2-Y1	-143	20

angle	θ_0 (deg)	k_θ (kJ mol^{-1})
G1-G2-G3	142	675
G1-G2-G2	98	150
G2-G2-G2	180	700
G2-G2-G1	93	200
G2-G2-G3	76	220
G3-G2-G2	83	380
M1-M2-M3	160	1800
M1-M2-M2	69	130
M2-M2-M2	180	400
M2-M2-M1	104	340
M2-M2-M3	95	200
M3-M2-M2	77	120
M1-M2-G2	81	200
M2-G2-M2	180	2000
M2-G2-G1	95	340
M2-G2-G3	83	250
M3-M2-G2	79	320
G2-M2-M1	110	200
G2-M2-M3	87	200
G2-M2-G2	180	900
G1-G2-M2	89	40
G3-G2-M2	77	60
M2-M2-G2	180	350
M2-G2-G2	180	300
G2-G2-M2	160	1200
G2-M2-M2	155	600
Y1-Y2-Y3	180	5500
Y1-Y2-Y2	63	110
Y2-Y2-Y2	167	1000
Y2-Y2-Y1	96	350
Y2-Y2-Y3	86	800
Y3-Y2-Y2	88	180

The bonds B1-B2 and B3-B2 are rather stiff because of the sugar rings. These bonds are therefore replaced by constraints with bond lengths equal to the average distances of the corresponding mapped AA distributions.

All possible angles formed by three consecutive bonded CG beads are parameterized, defined over a maximum of 3 residues (B2-B2-B2, see figure 10). This avoids non-physical conformational changes upon binding with oppositely charged polymers

Regarding the dihedral angles, 2 out of the 4 possible dihedrals ranging over two residues are parameterized (B1-B2-B2-B3 and B3-B2-B2-B1). These 2 dihedral angles were adequate to reproduce the AA dihedral distributions ranging over 2 residues. However, dihedrals ranging over 3 residues (e.g. B2-B2-B2-B3) and 4 residues (B2-B2-B2-B2) could not be parameterized because of the instability when the angle B2-B2-B2 visits values close to 180°. The restricted bending potentials [66] and a shorter timestep solved the issue. However, the resulting B2-B2-B2 angle distribution matched the AA model inadequately. As the B2-B2-B2 angle is assumed important for the polymer backbone stiffness, no restricted bending potentials were used. The dihedrals ranging over 3 to 4 residues were therefore not parameterized, and some of their distributions are bimodal.

4.2.2 NB1

The first coarse-grained non-bonded parameter set was based on the CHARMM36 RDFs. Interestingly, it turned out that a very simple set of parameters captured most of the atomistic RDF structures with equal or better accuracy than any other set of parameters. The simplicity of the parameter set is apparent in table 7. Most of the interactions between atom types are (approximately) purely repulsive, as illustrated by the SN4-SN4 and SP4-SP4 potentials in the left of figure 12. These interactions capture the excluded volume effects between beads.

Table 7: The Lennard-Jones interaction parameters between beads B_i and B_j that define NB1. Monovalent ions are described by the TQ5 bead types. The bead type naming convention is given in the right of figure 12 below.

B_i	B_j	σ_{ij} (nm)	ϵ_{ij} (kJ mol^{-1})	B_i	B_j	σ_{ij} (nm)	ϵ_{ij} (kJ mol^{-1})
SQ5n	SQ5n	0.900	$1 \cdot 10^{-5}$	SP4	SP4	0.900	$1 \cdot 10^{-5}$
SQ5p	SQ5n	0.350	$8 \cdot 10^{0}$	SP4	SP3	0.900	$1 \cdot 10^{-5}$
SP4	SQ5n	0.900	$1 \cdot 10^{-5}$	SP4	SN4	0.900	$1 \cdot 10^{-5}$
SP3	SQ5n	0.350	$2 \cdot 10^{0}$	SP4	TQ5	0.900	$1 \cdot 10^{-5}$
SN4	SQ5n	1.100	$1 \cdot 10^{-5}$	SP3	SP3	0.900	$1 \cdot 10^{-5}$
TQ5	SQ5n	0.900	$1 \cdot 10^{-5}$	SP3	SN4	0.900	$1 \cdot 10^{-5}$
SQ5p	SQ5p	0.900	$1 \cdot 10^{-5}$	SP3	TQ5	0.900	$1 \cdot 10^{-5}$
SP4	SQ5p	0.900	$1 \cdot 10^{-5}$	SN4	SN4	1.100	$1 \cdot 10^{-5}$
SP3	SQ5p	0.900	$1 \cdot 10^{-5}$	SN4	TQ5	0.900	$1 \cdot 10^{-5}$
SN4	SQ5p	1.100	$1 \cdot 10^{-5}$	TQ5	TQ5	0.462	$1 \cdot 10^{-1}$
TQ5	SQ5p	0.900	$1 \cdot 10^{-5}$				

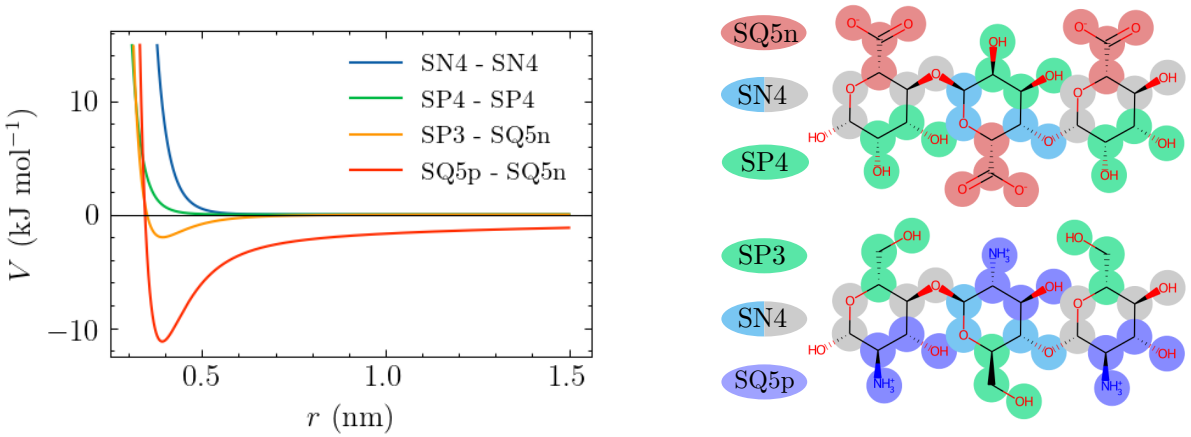


Figure 12: **Left)** Potential energy, V, as a function of the distance, r, between bead types. The SQ5n - SQ5p potential includes an attractive coulomb interactions with $\varepsilon = 80$. The SQ5n and SQ5p beads have a charge of -1 and +1, respectively. **Right)** Bead type definitions for alginate (top) and chitosan (bottom).

The strongest interaction is between the carboxylate group on alginate (SQ5n) and the ammonium group on chitosan (SQ5p). The charged groups interact through a Coulomb potential with a relative permittivity of 80. Treating the solvent as a dielectric continuum is unrealistic at short distances. Therefore, the short-ranged contribution is factored into the LJ parameters, which gives rise to a rather large well depth for the SQ5p-SQ5n interaction. Sodium and chloride ions are represented by TQ5 beads. Their interactions were parameterized by matching the RDFs of these univalent ions in Martini 3 water.

The second strongest interaction is between the -CH2-OH group on chitosan (SP3) and SQ5n. Most of the mapped AA RDFs were reproduced by *not* including the SP3-SQ5n attractive interaction. However, this interaction describes the hydrogen bonding between chitosan and alginate. It was therefore included, which also improved the fit of the SP3-SQ5n RDFs (see figure 28 in the appendix). However, adding more interactions between beads with underlying hydrogen bond donating or accepting capabilities did not drastically improve the fit between AA and CG RDFs. Therefore, no other attractive potentials are assigned.

It is apparent from the RDFs in figure 28 in the appendix that the 7MG oligomer can bind tighter to the 7CH oligomer, which is not captured well by the current CG parameterization. This binding mode may be essential for gel formation. An approach to replicating such specific binding modes in the CG model is by differentiating between atom types based on the alginate residue composition. Such an approach is used to develop NB2. However, avoiding these specific interactions makes the CG model more general, and the comparison of properties between different alginate systems is straightforward. For example, differences in complexation energies between alginate and chitosan can be attributed to the conformational differences arising from the specific alginate composition.

The alginate-alginate and chitosan-chitosan RDF structures (or their absence, see figure 29 in the appendix) are reproduced well by a purely repulsive potential (again, approximately). A repulsive LJ potential also describes the interactions between oligomers and ions.

The final model may be transferred to the martini force field, as the bonded distributions are rather insensitive to the non-bonded treatment and even the presence of an

explicit solvent. However, simulations with the default martini 3 parameters suggest an overestimation of electrostatic interactions when treating long-range interactions explicitly, as ions condensated completely along the alginate chains. Direct transfer of the model to the current martini 3 force field should therefore be done with caution. Instead, using the martini 2 force field with the polarizable water model [67] yields better properties. A future extension of the martini 3 force field with a polarizable water model may solve this issue.

4.2.3 NB2

The second non-bonded parameter set is given in table 8, and a comparison between the AA and CG RDFs is given in figure 30 and 31 the appendix.

Table 8: The Lennard-Jones interaction parameters between beads B_i and B_j that define NB2. Monovalent ions are described by the TQ5 bead types. The bead type naming convention was given in figure 12. The identifier J(KL) refers to a J bead on a K residue within a JK-block.

B_i	B_j	σ_{ij} (nm)	ϵ_{ij} (kJ mol^{-1})	B_i	B_j	σ_{ij} (nm)	ϵ_{ij} (kJ mol^{-1})
SQ5p	SQ5n(GM)	0.350	5.100	TQ5	SP4(MG)	0.300	0.500
SQ5p	SQ5n(MG)	0.350	5.100	TQ5	SP4(MM)	0.300	0.500
SQ5p	SQ5n(MM)	0.350	2.500	TQ5	SP4(GG)	0.300	4.000
SQ5p	SQ5n(GG)	0.350	3.750	TQ5	SQ5n(GM)	0.300	2.500
SP3	SQ5n(GM)	0.350	5.250	TQ5	SQ5n(MG)	0.300	5.000
SP3	SQ5n(MG)	0.350	5.250	TQ5	SQ5n(MM)	0.300	2.500
SP3	SQ5n(MM)	0.350	2.500	TQ5	SQ5n(GG)	0.300	5.000
SP3	SQ5n(GG)	0.350	3.750	TQ5	TQ5	0.462	0.100
TQ5	SP4(GM)	0.300	4.000	rest	rest	0.900	$1 \cdot 10^{-5}$

A significant coordination between sodium ions and the oligomers 7GG and 7MG is apparent from simulations with the GLYCAM06 force field (see figure 31). This was due to GG-blocks on 7GG and GM-blocks on 7MG. This binding mode was not observed for MM-blocks on 7MM or MG-blocks on 7MG. The involved functional groups that bind to sodium are the carboxylate groups (SQ5n) on an M or G residue, and the two hydroxyl groups (SP4) on the G residues, with additional contribution from the oxygen in the glycosidic bond. The atomic charges on the two oxygen atoms on the hydroxyl groups (O2 and O3) were given in table 5 above. It was noted that these charges are greater for G residues than M residues.

This binding mode is often called the "egg-box" model because ions bound to two GG-blocks of different chains resemble an egg in a box. This binding mode is well known for alginates with high G-block contents [11, 68]. However, a similar interaction for GM-blocks was unexpected, but it has been observed experimentally with calcium ions [69–72]. A representative snapshot is given in figure 13.

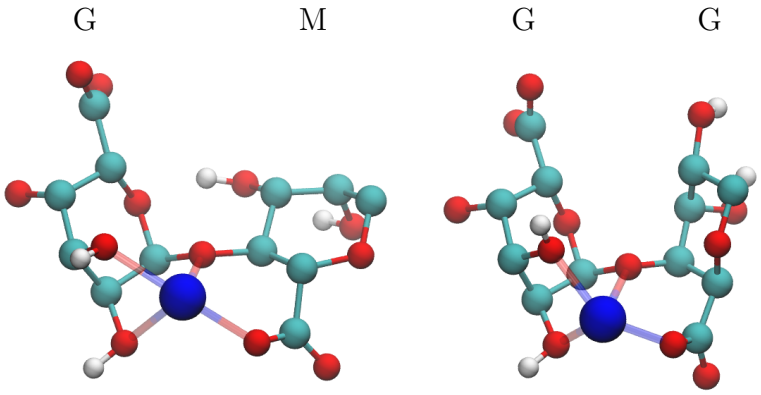


Figure 13: The "egg-box" model with sodium ions (colored blue), as observed in the equilibrium simulations of oligomers 7MG and 7GG with the GLYCAM06 force field. The bonds between sodium and oxygen atoms are based on a distance criterion, in which a bond is drawn if the distance between sodium and oxygen is less than 2.4 nm.

The LJ parameters were initially optimized to reproduce the GLYCAM06 RDFs, and the new bead types allowed reproduction of this "egg-box" binding mode for GG- and GM-blocks. However, directly fitting the CG model to the mapped GLYCAM06 RDFs resulted in simulation instabilities and excessive chain aggregation for larger systems. The instabilities resulted from the small radius and great well depth of the SP4-TQ5 interaction, which left the integration time step too large. The aggregation was due to the great well depth required to reproduce the SQ5n-TQ5 and SQ5n-SP4 peaks in the GLYCAM06 RDFs. Therefore, a qualitative procedure was used to optimize the interactions between sodium and alginate.

It was assumed that the "egg-box" interaction is characterized by a coordination of sodium with GG- and GM- blocks, but not with MG- and MM-blocks (as shown in figure 13). The interaction strength between sodium (TQ5) and the two hydroxyl groups on G residues of a GG- or GM-block (SP4(GG), SP4(GM)) are given a default value of 4 kJ mol^{-1}. The SP4 beads on G- and M-residues of a MG- or MM-block (SP4(MG), SP4(MM)) are given by a lower default value of 0.5 kJ mol^{-1} because they do not participate in such a binding mode with sodium. Similarly, the carboxylate groups on M-residues of MM-blocks and on G-residues of GM-blocks (SQ5n(MM), SQ5n(GM)) do not participate in egg-box binding, and these interactions are given a default value of 2.5 kJ mol^{-1}. These default values reproduce well the binding of sodium to 7GG and 7MM. Finally, the interaction strength between sodium and the carboxylate groups on G-residues of GG-blocks and on M-residues of GM-blocks (SQ5n(GG), SQ5n(MG)) are given a value ε. This value is then adjusted in simulations with large alginate chains such that no chain aggregation occurs (the poly-G system was soluble at 10 g L^{-1} in [2]) . This is illustrated in figure 14.

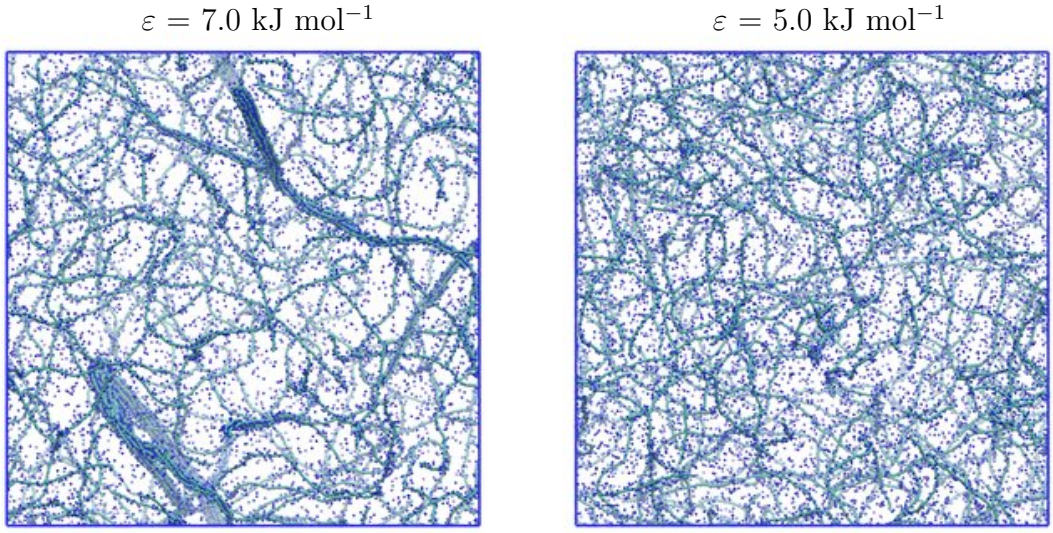


Figure 14: Simulation snapshots at $t = 50$ ns for a solution of 50 pre-equilibrated poly-G chains neutralized with sodium ions (DP=300 at 10 g L^{-1}). The two systems differ in their LJ interactions strengths ε between sodium and the carboxylate group on alginate (bead types TQ5-SQ5n(GG)). Alginate chains are colored cyan and soidum ions are colored dark blue.

The new bead type assignment also gives room for residue-dependent interaction strengths between alginate and chitosan, as observed in the CHARMM36 and GLYCAM06 simulations. The PMFs between alginates of different compositions and chitosan are used to establish these interactions. In addition, the PMF incorporates the effects of the solvent. Using the PMF obtained from biased simulations allows sampling configurations of low probability that do not occur during equilibrium simulations at practical time scales. A comparison between AA PMFs and those obtained with NB2 are given in figure 15 and table 9.

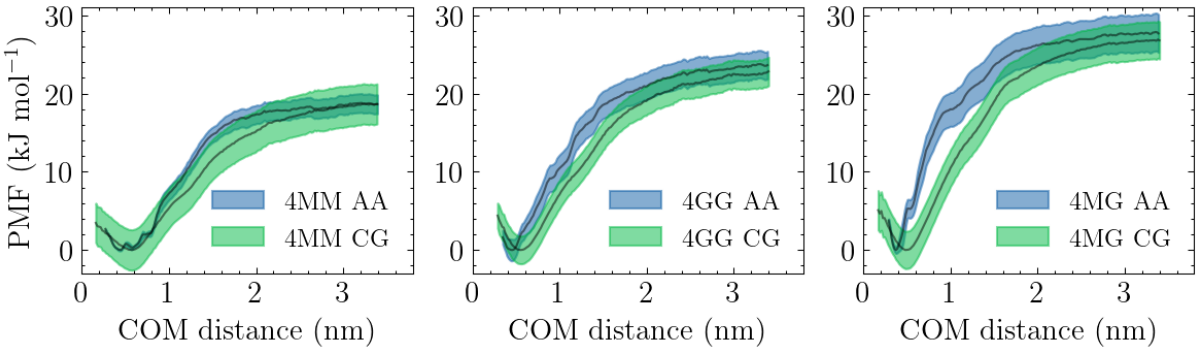


Figure 15: Potentials of mean force (PMF) as a function of the center of mass (COM) distance between alginate and chitosan oligomers of length 4. The shaded areas are the standard deviations obtained from 100 iterations of Bayesian bootstrapping [48].

Table 9: Average potentials of mean force $\langle \text{PMF} \rangle$ between chitosan and different alginates of length 4 at 3.4 nm for the AA and CG systems. The corresponding standard deviations σ were obtained from 100 iterations of Bayesian bootstrapping.

System	$\langle \text{PMF} \rangle_{\text{AA}}$ (kJ mol^{-1})	σ_{AA} (kJ mol^{-1})	$\langle \text{PMF} \rangle_{\text{CG}}$ (kJ mol^{-1})	σ_{CG} (kJ mol^{-1})
4MM-4CH	18.6	1.2	18.7	2.6
4GG-4CH	23.6	1.8	22.8	1.8
4MG-4CH	27.7	2.4	26.8	2.4

The atomistic binding free energies are in the order MM<GG<MG and are adequately reproduced with this CG parameter set. The CG PMFs are lower at 3.4 nm compared to the AA PMFs. However, the CG PMFs would increase at larger COM distances until all CG beads are outside the electrostatic cut-off at 3.8 nm. Therefore, more accurate AA PMFs should be calculated in future studies using larger alginate chains and ensuring that the profiles level off at large COM distances.

The use of relatively short oligomers for binding free energy estimates in this project was based on computational efficiency and obtaining adequate sampling. Binding free energies calculated for the systems 12MG-4CH and 12MG-8CH were approximately equal to and twice as large as for 4MG-4CH, respectively (see figure 32 in the appendix). However, these larger systems were expensive to simulate, and their PMF profiles converged slowly.

4.3 Polymers solutions: NB1 and NB2

This subsection describes the results of the large-scale polymer solution simulations, which are compare to experimental values as a validation step.

The equilibrium values for the persistence length and radius of gyration of the alginate 300-mers are given in table 10 together with experimentally determined values.

Table 10: Comparison of calculated static polymer properties with experiments [73, 74] and other simulations [75]. The average persistence length $\langle q \rangle$ and standard deviation during the simulation σ_q for 300GG, 300MG and 300MM using NB1 and NB2. The radius of gyration $\sqrt{\langle R_G^2 \rangle}$ was calculated directly from the simulations, while $\sqrt{\langle R_G^2 \rangle}_{955}$ was based on equation 3.2 with N=955 and M=190 kDa.

System	$\langle q \rangle$ (nm)	σ_q (nm)	$\sqrt{\langle R_G^2 \rangle}$ (nm)	$\sqrt{\langle R_G^2 \rangle}_{955}$ (nm)
300GG A	21.3	1.1	25.2	51.5
300GG B	19.5	0.8	24.1	49.7
300MM A	14.0	0.6	24.5	47.0
300MM B	13.4	0.5	23.7	46.0
300MG A	22.7	1.2	27.8	56.6
300MG B	22.6	1.4	27.8	56.5
M/G ratio≈1.6 [73]	10 ± 6			50 ± 7
poly-M [74]	14.5 ± 1.8		25.8*	
poly-MG [74]	14.9 ± 1.2		25.8*	
poly-G [74]	16.5 ± 0.5		25.8*	
poly-M [75]	11.9			
poly-G [75]	21.0			

* Calculated from eq. 2 in [74]: $R_{\mathrm{G}} = 0.0352 M^{0.60}$.

The persistence lengths and radii of gyration calculated for the same systems with NB1 and NB2 are in agreement. The persistence lengths for 300MM also agree well with experiments and other simulations of comparable systems (poly-M). The values for 300GG are somewhat larger than the experimental values (poly-G), but agree well with other simulations. Experimentally, stiffness parameters are estimated using experimental properties that are challenging to determine accurately. Therefore, the experimentally determined persistence lengths vary greatly in the literature [74]. It is noted that the persistence lengths are greater for 300MG than for 300GG, which deviates from the usual stiffness trend MM < MG < GG observed for alginates, and the 300MG systems have a significantly larger persistence length as compared to experiments (poly-MG) There is also a large jump between the persistence lengths of 300MM and those of 300MG and 300GG. This may be due to an inadequate parameterization of dihedral angles ranging over 4 residues mentioned in section 4.2.1, which is apparent from the bimodal dihedral angle distribution for 7MM (B2-B2-B2-B2) in figure 22. This problem remained unexplored, but it should be reviewed by future investigators aiming to apply the model. However, the calculated radii of gyration $\sqrt{\langle R_G^2 \rangle}$ in table 10 are in better agreement with experiments. The radius of gyration of a large chain with M/G ratio≈1.6 (around 61.5% M residues) is lower than for 300MG, but the values are within the experimental error estimate.

The stress relaxation moduli $G(t)$ of the alginate solutions were calculated with NB2. This non-bonded parameter set incorporates the specific "egg-box" interaction for the 300MG and 300GG systems. The calculated moduli are shown in figure 16 and compared to the moduli predicted by the Rouse model.

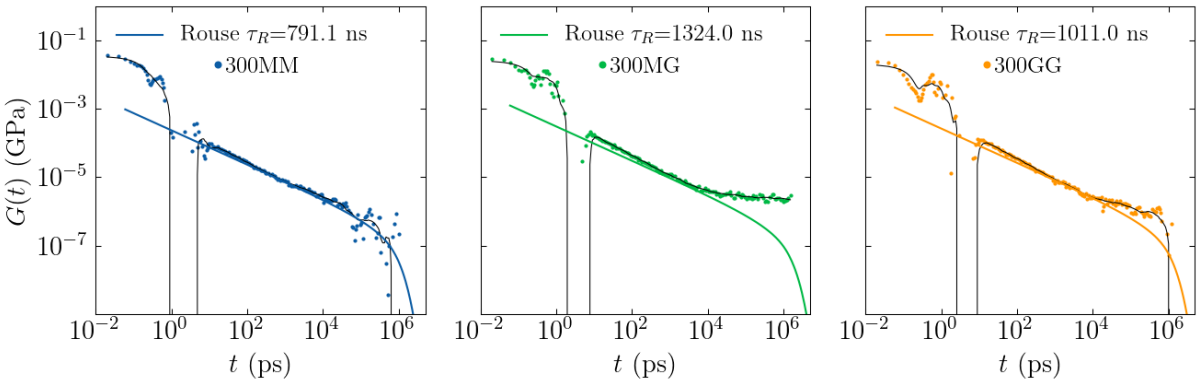


Figure 16: Shear stress relaxation moduli $G(t)$ calculated for 50 alginate chains (DP=300) with NB2 at a concentration of at 10 g L^{-1}. The corresponding moduli predicted from the Rouse model (equation 2.23) are also given. The black lines are smoothed signals.

The 300MM system agrees with the predictions from the Rouse model, except at very short timescales. The disagreement at short time scales is due to short-lived relaxation processes such as interactions with counterions and bond lenght relaxations [25]. The 300MG and 300GG systems deviate from the Rouse behavior at long timescales, and it is apparent that the simulations are too short of obtaining the complete relaxation spectrum. Deviations at long timescales occur because the Rouse model assumes that the polymer beads interact only with neighboring beads. Other intra- and inter-chain interactions are neglected.

The deviations of Rouse behavior at long time scales are often attributed to chain entanglements, which originate from the uncrossability of the chains. The chains follow Rouse behavior up to an entanglement time, after which the well-known tube model better explains the effect of entanglements [16]. A key parameter of the tube model is the average number of entanglements per chain, which were calculated using the Z1-code [56–59]. The values were 2.65, 3.31, and 2.58 for 300MM, 300MG, and 300GG, respectively. The number of entanglements is approximately equal for 300MM and 300GG. Thus, the deviation from Rouse behavior may be attributed to inter-chain correlations due to the "egg-box" interaction modes between 300GG chains. This interaction is weaker for the 300MG chains, but the greater number of entanglements may explain the more pronounced deviation. Therefore, specific interactions between sodium and certain alginate blocks may significantly affect the viscoelastic properties. Ongoing simulations with NB1, which do not capture the specific interactions, confirm this (see figure 33 in the appendix).

The intrinsic viscosity, which is the polymer contribution to the viscosity, may be estimated by integrating $G(t)$ (equation 2.20) in figure 16 because the solvent is neglected . It can be shown that the intrinsic viscosity of the Rouse model is proportional to M^1 [16, p. 114], which is characteristic for non-entangled solutions. The M^1 scaling behavior of the intrinsic viscosity has been found experimentally for alginates with chain lengths of around 100 - 500 with compositions similar to those used in this project (see the correction notice in [74]). Relating the intrinsic viscosities of the CG model to experiments allows scaling the CG timescale such that the two viscosities match. Rescaling the CG timescale to match the atomistic dynamics is often performed in CG applications [76]. However, it may be worthwhile to establish a connection between the time scales of the CG model and real systems such that the absolute values of other calculated properties may be more accurately related to experiments. This opens the way for further validation steps or

refinements of the model or even accurately predicting the absolute values of properties that are not easily accessible from experiments.

It is noted that no replicates were performed for the simulations in this section, and the results are expected to depend on the initial conditions.

4.4 Hydrogels

In this section, the "hydrogel" systems are investigated, in which chitosan oligomers were inserted into the equilibrated structures from the previous section. The results are discussed in light of the experimentally observed behavior of the corresponding systems in [2].

4.4.1 NB1

The final snapshots after around 1.5 us of simulations using NB1 are given in the appendix (figures 34 and 35). The three systems resemble each other upon visual inspection. Some larger structures are evident for 300MM and 300GG.

A clustering procedure is adopted to quantify the number of chitosan oligomers bundling together every 10 ns during the final 600 ns of the simulations. As mentioned in the methods section, two or more chitosan oligomers form a cluster if one can draw lines less than 2 nm connecting the oligomers B2 beads. The results for chitosan 8-mers and 4-mers are given in figures 17 and 18, respectively. Results for the shorter replicate runs are shown in figures 37 and 38 in the appendix.

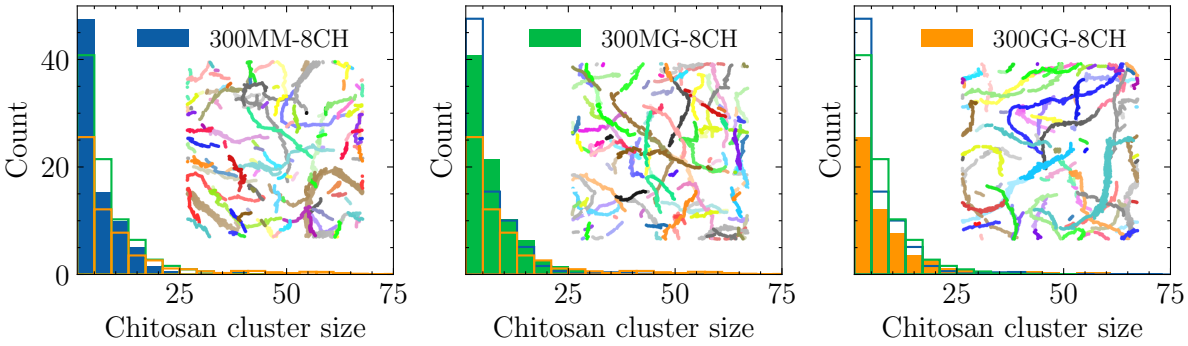


Figure 17: The distribution of chitosan cluster sizes during the last 600 ns of the simulations calculated with NB1. The chitosan cluster size quantifies the number of clusters formed by chitosan molecules that are within 2.0 nm of each other. The total number of chitosan 8-mers is 564, and the histogram bin width is 4. The insets illustrate all chitosan clusters colored according to their cluster indices in the final simulation snapshot.

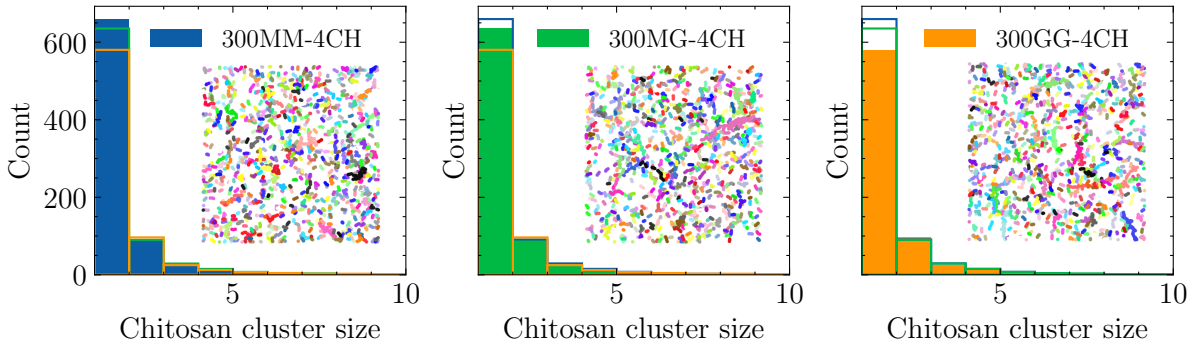


Figure 18: The distribution of chitosan cluster sizes during the last 600 ns of the simulations calculated with NB1. The chitosan cluster size quantifies the number of clusters formed by chitosan molecules that are within 2.0 nm of each other. The total number of chitosan 4-mers is 1128, and the histogram bin width is 1. The insets illustrate all chitosan clusters colored according to their cluster indices in the final simulation snapshot.

Few of the 8-mers detach from the alginate chains, and the majority of the oligomers form crosslinks between alginate chains. The situation differs for systems involving the chitosan 4-mers, in which a considerable number of oligomers are attached to single chains without forming crosslinks. A clear trend is evident in figures 17 and 18 (and in figures 38 and 37 in the appendix for the replicate runs). The chitosan cluster size parallels MM $<$ MG $<$ GG, meaning that chitosan oligomers form fewer, but larger, bundles between 300GG chains (see the orange histogram lines within the plots of 300MM-8CH in figure 17 at large cluster sizes) The 300MG systems form more numerous medium-sized clusters, while the 300MM system forms a large number of small sized clusters. This trend is most apparent for the 8-mer systems in figure 17.

Furthermore, it turns out that the chitosan oligomers are significantly more mobile along the 300GG chains, which may explain this trend. The chitosan oligomers tend to "slide" along the 300GG chain but are more rigidly attached along the 300MM and 300MG chains (see the left of figure 19). The RDF between the chitosan backbone beads and the alginate backbone beads indicates the existence of energy barriers along the chain of 300MM and 300MG, which are absent for 300GG (right and bottom of figure 19). This property is attributed to the mismatch of the distances between functional groups (the B1 and B3 beads) of chitosan and 300GG. A chitosan oligomer attaching to the 300GG chain can therefore slide efficiently along the backbone to a low energy position, often where other oligomers are crosslinked to adjacent chains (see e.g. 300GG-8CH in figure 34). This resembles the successive addition of beads to a pendant necklace.

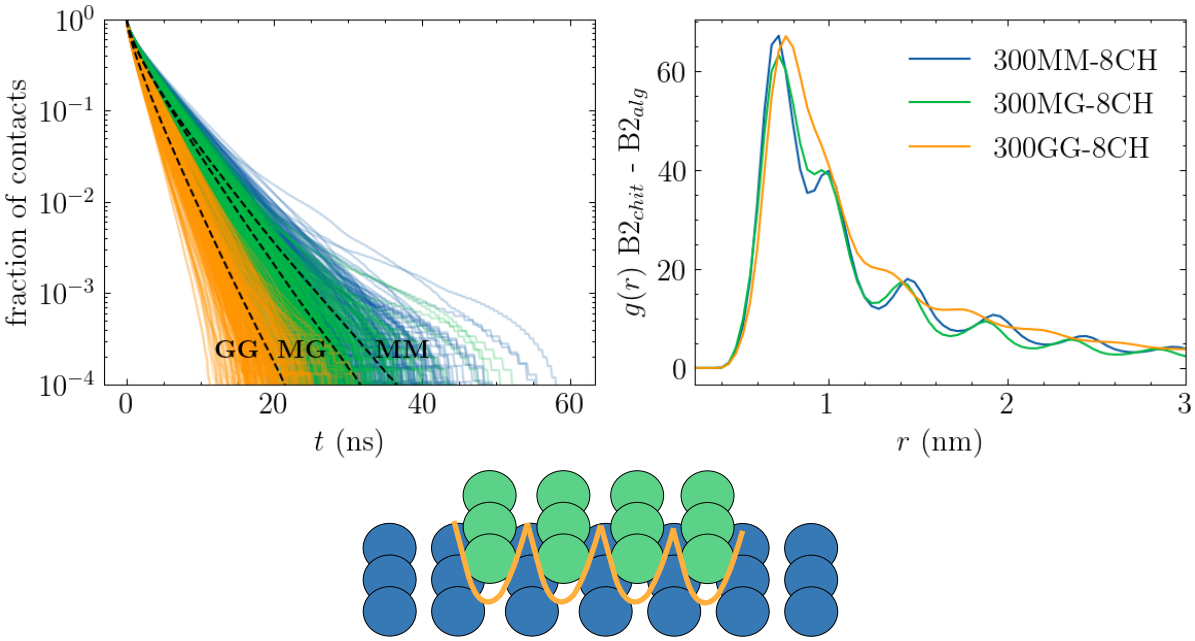


Figure 19: **Left)** The fraction of contacts within 1.2 nm between 8-mer chitosan backbone beads and alginate backbone beads remaining at time t, averaged over the last 1.2 μs of the simulation. The colored lines corresponds to 1 of the 50 alginate molecules. The dotted lines are the averages of all 50 alginate molecules. **Right)** The radial distribution function $g(r)$ between 8-mer chitosan backbone beads and alginate backbone beads. **Bottom)** An illustration of the energy barriers (orange lines) that avoid the sliding of a chitosan oligomer (green) along an alginate chain (blue), which is observed for 300MM and 300MG, but not 300GG. Bonds between monomers are not shown.

The lifetime of crosslinks between polymer chains is an essential factor and influences the viscoelastic properties of associating polymers. Therefore, the Rouse model has been extended to include the effect of associative groups, or "stickers", that can form transient bonds between polymer chains [77]. This Sticky Rouse model may also be applied when phase separations occur by treating them as fixed crosslinks within the network. The theoretical treatment shows that a greater sticker lifetime, related to the crosslink energy through an Arrhenius relationship, increased the solid elastic characteristics [78] (the storage modulus in equation 2.21). It is, therefore, interesting to note that even though the binding energies are similar for the alginates with NB1 (see table 11), the greater mobility of the chitosan oligomers along the 300GG chain may give rise to different viscoelastic behaviors. This phenomenon would not be easily detected during atomistic simulations of smaller systems and illustrates the capability of the current CG approach. The differences in chitosan mobility along the various alginates should be investigated in future atomistic simulations.

The binding energies between alginates and chitosan oligomers of lengths 8 and 4 are given in table 11.

Table 11: Averaged binding energies $\Delta\langle E\rangle$ and the corresponding standard deviations $\sigma_{\Delta\langle E\rangle}$ between alginate 300-mers and chitosan oligmers calculated with NB1. Units are in kJ per mol of chitosan oligomer (DP=8 or DP=4)

System	$\Delta\langle E\rangle$ (kJ mol^{-1})	$\sigma_{\Delta\langle E\rangle}$ (kJ mol^{-1})
300MM-8CH	-103.8	1.9
300MG-8CH	-104.6	1.8
300GG-8CH	-104.6	1.8
300MM-4CH	-34.2	0.9
300MG-4CH	-35.9	1.0
300GG-4CH	-36.4	0.8

The binding energies of the chitosan 8-mers approach the same values in the final states of the simulation. For the chitosan 4-mers, the 300MM binding energies are slightly lower than for 300MG and 300GG.

The alginate network size formed by chitosan crosslinks is investigated next. As mentioned in the methods section, a slightly different clustering procedure quantifies the network size. Two alginate chains belong to the same network if they are crosslinked by a given number of chitosan monomers. This number is referred to as the crosslink number threshold, or simply the number of crosslinks between chains. The results for the two longest simulations involving chitosan 8-mers and 4-mers are shown in figures 20 and 21, respectively.

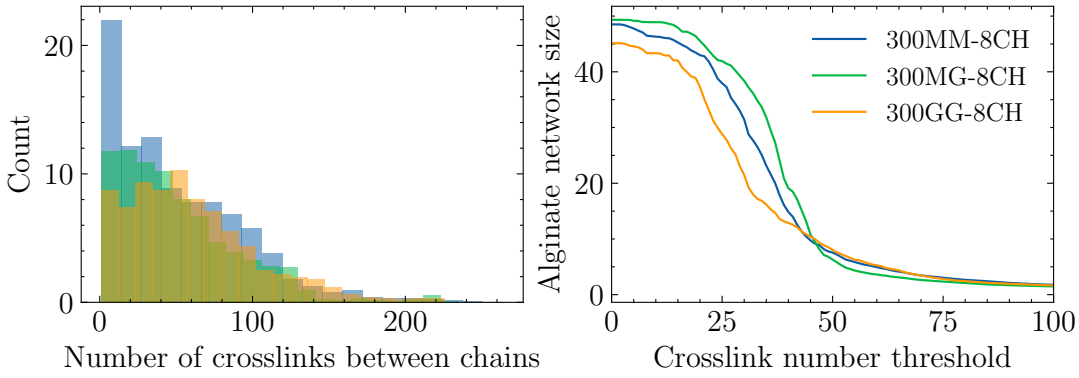


Figure 20: **Left)** The number of chitosan crosslinks between alginate chains. One chitosan 8-mer can contribute with a maximum of 8 crosslinks. **Right)** The alginate network size as a function of the number of crosslinks between alginate chains. The largest possible network size is 50. The definition of the network size was given in equation 3.4.

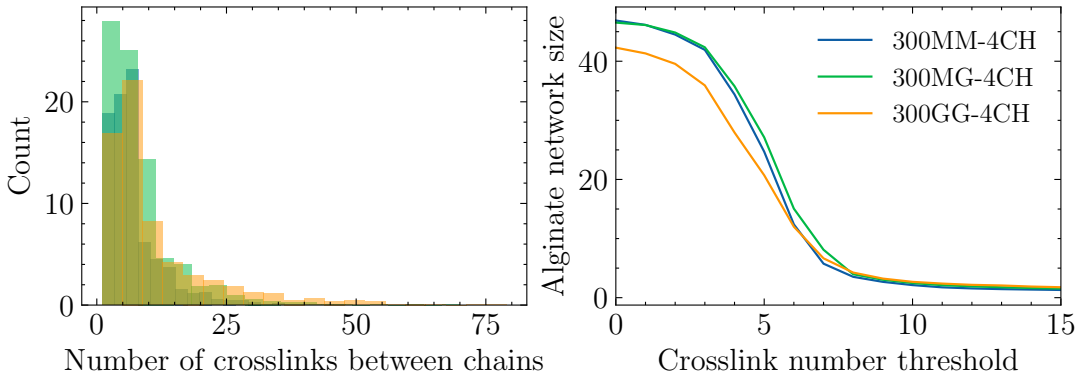


Figure 21: **Left)** The number of chitosan crosslinks between alginate chains. One chitosan 4-mer can contribute with a maximum of 4 crosslinks. **Right)** The alginate network size as a function of the number of crosslinks between alginate chains. The largest possible network size is 50. The definition of the network size was given in equation 3.4.

A trend similar to the chitosan cluster size is evident on the left of figures 20 and 21; the number of crosslinks between chains parallels the chitosan cluster size. The same is evident in figures 39 and 40 for the shorter replicate runs given in the appendix.

Regarding the network sizes on the right of figures 20 and 21, it is apparent that the 300GG systems form considerably smaller polymer networks. Even at a crosslink number threshold of 1, the network size does not reach the maximum values, meaning that the system does not constitute a single network. In contrast, almost all 50 300MM and 300MG chains are crosslinked with many oligomers. The less extensive networks for 300GG are even more evident in the two replicate runs (figures 39 and 40 in the appendix).

An accurate estimation of the relaxation moduli $G(t)$ for the systems crosslinked with chitosan oligomers of lengths 4 and 8 was challenging due to the increased relaxation times after introducing chitosan. The simulation duration was too short for obtaining relaxation moduli without significant noise at large time scales. At the same time, smaller systems with alginate lengths 200 and 150 displayed excessive aggregation in combination with chitosan oligomers of length 8, and eventually formed a single cluster. Systems with length 4 oligomers were not investigated.

4.4.2 NB2

The final snapshots of simulations carried out for alginate-chitosan hydrogels with NB2 are given in figure 36 in the appendix. A visual inspection reveals remarkably different behaviors for the three systems.

Compared with the experiments mentioned in the introduction [2], the gelling systems involving poly-MG and poly-G with chitosan 8-mers were highly turbid. The turbidity suggests the existence of phase separations in these systems. This hypothesis is strengthened based on a visual inspection of figure 36. In the simulation snapshots, the 300MG and 300GG systems indicate a phase separation, which is more pronounced for the 300GG system.

NB2 was based on reproducing free energies of binding between chitosan and alginate, and the binding strengths with chitosan were in the order MM < GG < MG. In addition, the "egg-box" interaction was incorporated for GG and GM residues. The 300GG can therefore form "egg-box" interaction with sodium between each GG residue at each side of

the chain. For example, the oligomer GGGG can bind six sodium ions tightly. In reality, such interactions may only be possible at alternating sides of each GG residue. However, the current CG model can not capture this, and the alginate-alginate complexation with sodium ions in-between may be overestimated.

The hypothesis of phase separations may be augmented with the polymer network analysis for the chitosan 8-mers with the NB1 model (figure 20). Regarding the 300MM system, the crosslinks are evenly distributed and form a continuous network, and the snapshots (300MM-8CH NB2 in figure 36) do not suggest a phase separation. This may give rise to the transparent gels of moderate strengths observed experimentally in the poly-M gelling systems [2]. In contrast, the possible phase separation (300GG-8CH NB2 in figure 36) and the less evenly crosslinked 300GG system (figure 20) may result in the viscous and turbid polymer solutions that were observed experimentally for the poly-G gelling systems. On the other hand, the poly-MG systems displayed much larger gel strengths. Due to a more extensively crosslinked network in the 300MG system (figure 17) and the possible phase separation (300MG-8CH B in figure 36), the large gel strengths observed experimentally may therefore arise due to the formation evenly crosslinked fibers of alginate and chitosan chains.

Furthermore, it was experimentally observed that increasing the salt concentration of the poly-MG gelling system increased the gel strength at moderate salt concentrations [2]. During the simulations, considerable fractions of ions are situated between the crosslinks of 300MG and 300GG chains (this is not visible in figure 36). Therefore, increasing the salt concentration may further stabilize the junction zones along the 300GG and 300MG chains and therefore increase the elastic characteristics, which was observed for the poly-MG gelling system when increasing the salt concentration [2]. It may be constructive at this point to revisit figure 5 in the theory section, in which the different behavior of the solid and liquid phases are apparent from the stress relaxation function $G(t)$.

The conclusions that are drawn in this section parallel the conclusions from [2] and illustrate the potential of the current coarse-grained modeling approach. However, these results may seem speculative due to the limited number of systems studied and the limited simulation durations.

5 Conclusion and perspectives

It was found in a recent study that the hydrogels formed by alginates with strictly alternating MG residues (poly-MG) crosslinked with chitosan oligomers displayed surprisingly high gel strengths, which was attributed to a phase separation. Moreover, the gel strength increased at moderate salt concentration, making these gels relevant to biomedical applications. The poly-M gelling system displayed weaker gel strengths and was less tolerant to increasing salt concentrations. The poly-G system did not form gels but rather a turbid viscous solution.

A coarse-grained approach was adopted in this project to study these gelling systems at the microscopic level. The coarse-grained model was based on the CHARMM36 and GLYCAM06 force fields. Using two force fields gave rise to two flavors of the CG model, NB1 and NB2. NB1 was based on a general treatment of the nonbonded interactions, while NB2 included composition-dependent alginate-chitosan and alginate-sodium interactions.

Large-scale alginate solutions with DP=300 were studied with the two CG models, and the static properties derived from the simulations were compared to experimental data to validate the conformational properties of the CG model. Persistence lengths and radii of gyration obtained from the simulations agreed reasonably well with the experimental values of comparable systems, considering that no replicate simulations were performed. Replicate simulations were not performed because of the long runs necessary for obtaining independent equilibrated initial structures.

The main objective of this project was to study the role of certain factors on the gel strengths of the systems mentioned above. These factors included the alginate composition, the role of counterions, chitosan oligomers, and polymeric effects. The gel strength was challenging to calculate using the direct approach adopted in this project due to the slow dynamics (even at the CG scale) and the onset of aggregations for smaller systems. Therefore, qualitative arguments were given instead of calculating the actual gel strength.

It was found with NB1 that chitosan oligomers are significantly more mobile along the 300GG chains compared to the 300MG and 300MM chains. The lower mobility of chitosan oligomers along the 300MG and 300MM chains was attributed to energy barriers along the 300MM and 300MG chains. These energy barriers were absent for the 300GG system. It was suggested that these energy barriers originated from a closer match between distances of the chitosan functional groups and the functional groups along the 300MM and 300MG chains. It was further suggested that the greater mobility of the chitosan oligomers along the 300GG chains resulted in less evenly crosslinked networks. A chitosan oligomer attaching to a 300GG chain could efficiently "slide" towards low energy positions where other oligomers attached to the same chain are crosslinked to an adjacent chain.

With NB2, a drastic difference between the three alginate systems was observed. The visual appearance of the 300MG and 300GG systems resulted in greater phase separations, in which sodium further stabilized the junction zones. This effect was most pronounced for the 300GG system. Therefore, the large gel strengths observed experimentally for the poly-MG systems may arise from both alginate-chitosan and alginate-sodium interactions, which explains the increased gel strengths with increasing salt concentrations of the poly-MG system.

In conclusion, the coarse-grained model developed in this project reproduced experimentally observed trends. The final application of the model sheds light on some of

the effects of alginate composition, the complexation with chitosan oligomers, and polymeric effects. The simulations performed with the model also highlight the importance of specific interactions with counterions, which are believed to be essential for the gelling systems investigated. However, a thorough investigation of the factors mentioned in the main objective of this project was not performed and is left for future studies. Still, the model provides insights into the large-scale structures of the hydrogels and a starting point for future investigators. Some ideas for future studies are listed below.

- The phase separations are challenging to treat with the current model and may require a different approach. A finer model could provide a deeper understanding of the phase separations. The model developed in [17] may be an interesting starting point in which the orientations of the hydroxyl groups of the sugars are modeled explicitly.

- The simulations performed in this project were expensive, even with a coarse-grained representation, due to the large number of simulation steps required for equilibration. Therefore, obtaining equilibrated structures from the direct approach adopted in this project is problematic. Therefore, "smart" Monte Carlo methods that produce less correlated configurations with fewer computational resources may be beneficial for generating initial configurations [59]. In addition, more replicates and longer simulations should be carried out, especially for the polyelectrolyte complex systems, due to the slower dynamics when introducing chitosan oligomers.

- The modified sticky Rouse model [79] could be used to treat the systems in this project, in which phase separations are treated as fixed crosslinks in space. The parameters entering the sticky Rouse model may be estimated from atomistic molecular dynamics simulations. The sound theoretical foundation provides a much cheaper route for obtaining the stress relaxation modulus $G(t)$, from which the other lienar viscoelastic properties can be derived.

- The solvent was neglected in the current coarse-grained model, and a realistic description of the phase separations may require the inclusion of an explicit solvent. However, a dilemma emerges at some model resolution. High-resolution modeling methods are restricted by the achievable length and time scales, while low-resolution methods suffer from the loss of accuracy. The atomistic force fields are, in turn, approximations of quantum mechanical calculations. Therefore, bypassing the atomistic force field and directly parameterizing a coarse-grained model based on quantum mechanical calculations may provide a more accurate description [80]. A new set of nonbonded parameters would be straightforward to incorporate into the model developed in this project.

References

[1] Bjørn E. Christensen. *TBT4135 Compendium: Biopolymers*. Department of Biotechnology and Food Science, NTNU, 2018.

[2] Georg Kopplin, Anders Lervik, Kurt I Draget, and Finn L Aachmann. Alginate gels crosslinked with chitosan oligomers–a systematic investigation into alginate block structure and chitosan oligomer interaction. *RSC Advances*, 11(23):13780–13798, 2021.

[3] Zhijie Zhang, Quan Chen, and Ralph H Colby. Dynamics of associative polymers. *Soft Matter*, 14(16):2961–2977, 2018.

[4] Alistair M Stephen and Glyn O Phillips. *Food polysaccharides and their applications*. CRC press, 2006.

[5] Sutapa Biswas Majee. *Emerging concepts in analysis and applications of hydrogels*. BoD–Books on Demand, 2016.

[6] Andrew R Leach and Andrew R Leach. *Molecular modelling: principles and applications*. Pearson education, 2001.

[7] Karl N Kirschner, Austin B Yongye, Sarah M Tschampel, Jorge González-Outeiriño, Charlisa R Daniels, B Lachele Foley, and Robert J Woods. Glycam06: a generalizable biomolecular force field. carbohydrates. *Journal of computational chemistry*, 29(4):622–655, 2008.

[8] Olgun Guvench, Sairam S Mallajosyula, E Prabhu Raman, Elizabeth Hatcher, Kenno Vanommeslaeghe, Theresa J Foster, Francis W Jamison, and Alexander D MacKerell Jr. Charmm additive all-atom force field for carbohydrate derivatives and its utility in polysaccharide and carbohydrate–protein modeling. *Journal of chemical theory and computation*, 7(10):3162–3180, 2011.

[9] Brian H Morrow, Gregory F Payne, and Jana Shen. ph-responsive self-assembly of polysaccharide through a rugged energy landscape. *Journal of the American Chemical Society*, 137(40):13024–13030, 2015.

[10] Steven W Benner and Carol K Hall. Development of a coarse-grained model of chitosan for predicting solution behavior. *The Journal of Physical Chemistry B*, 120(29):7253–7264, 2016.

[11] Wojciech Plazinski and Mateusz Drach. Calcium-α-l-guluronate complexes: Ca2+ binding modes from dft-md simulations. *The Journal of Physical Chemistry B*, 117(40):12105–12112, 2013.

[12] Paul Bauer, Berk Hess, and Erik Lindahl. Gromacs 2022.1 manual, April 2022.

[13] Christopher I Bayly, Piotr Cieplak, Wendy Cornell, and Peter A Kollman. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the resp model. *The Journal of Physical Chemistry*, 97(40):10269–10280, 1993.

[14] Satyen Dhamankar and Michael A Webb. Chemically specific coarse-graining of polymers: Methods and prospects. *Journal of Polymer Science*, 59(22):2613–2643, 2021.

[15] William George Noid. Perspective: Coarse-grained models for biomolecular systems. *The Journal of chemical physics*, 139(9):09B201_1, 2013.

[16] Masao Doi, Samuel Frederick Edwards, and Samuel Frederick Edwards. *The theory of polymer dynamics*, volume 73. oxford university press, 1988.

[17] Daniel J Beltran-Villegas, Daniel Intriago, Kyle HC Kim, Natnael Behabtu, J David Londono, and Arthi Jayaraman. Coarse-grained molecular dynamics simulations of α-1, 3-glucan. *Soft matter*, 15(23):4669–4681, 2019.

[18] Levan Tsereteli and Andrea Grafmüller. An accurate coarse-grained model for chitosan polysaccharides in aqueous solution. *PLoS One*, 12(7):e0180938, 2017.

[19] Cesar A López, Andrzej J Rzepiela, Alex H de Vries, Lubbert Dijkhuizen, Philippe H Hünenberger, and Siewert J Marrink. Martini coarse-grained force field: extension to carbohydrates. *Journal of Chemical Theory and Computation*, 5(12):3195–3210, 2009.

[20] David Roylance. Engineering viscoelasticity. *Department of Materials Science and Engineering–Massachusetts Institute of Technology, Cambridge MA*, 2139:1–37, 2001.

[21] JT Padding and Willem J Briels. Zero-shear stress relaxation and long time dynamics of a linear polyethylene melt: A test of rouse theory. *The Journal of Chemical Physics*, 114(19):8685–8693, 2001.

[22] Johan T Padding. Theory of polymer dynamics. *Advanced Courses in Macroscopic Physical Chemistry.(Han-sur-lesse winterschool 2005)*, 21, 2005.

[23] David Brown and Sylvie Neyertz. A general pressure tensor calculation for molecular dynamics simulations. *Molecular Physics*, 84(3):577–595, 1995.

[24] Jorge Ramírez, Sathish K Sukumaran, Bart Vorselaars, and Alexei E Likhtman. Efficient on the fly calculation of time correlation functions in computer simulations. *The Journal of chemical physics*, 133(15):154103, 2010.

[25] Alexei E Likhtman, Sathish K Sukumaran, and Jorge Ramirez. Linear viscoelasticity from molecular dynamics simulation of entangled polymers. *Macromolecules*, 40(18):6748–6757, 2007.

[26] Nuofei Jiang, Hongdong Zhang, Yuliang Yang, and Ping Tang. Molecular dynamics simulation of associative polymers: Understanding linear viscoelasticity from the sticky rouse model. *Journal of Rheology*, 65(4):527–547, 2021.

[27] Complex Carbohydrate Research Center, University of Georgia, Athens, GA. Glycam web. `http://legacy.glycam.org`, 2005-2021.

[28] D.A. Case, H.M. Aktulga, K. Belfon, I.Y. Ben-Shalom, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, G.A. Cisneros, V.W.D. Cruzeiro, T.A. Darden, R.E. Duke, G. Giambasu, M.K. Gilson, H. Gohlke, A.W. Goetz, R. Harris, S. Izadi, S.A. Izmailov, C. Jin, K. Kasavajhala, M.C. Kaymak, E. King, A. Kovalenko, T. Kurtzman, T.S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, M. Machado, V. Man, M. Manathunga, K.M. Merz, Y. Miao, O. Mikhailovskii, G. Monard, H. Nguyen, K.A. O'Hearn, A. Onufriev, F. Pan, S. Pantano, R. Qi, A. Rahnamoun, D.R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C.L. Simmerling, N.R. Skrynnikov, J. Smith, J. Swails, R.C. Walker, J. Wang, H. Wei, R.M. Wolf, X. Wu, Y. Xue, D.M. York, S. Zhao, and P.A. Kollman . *Amber 2021*. University of California Press, 2021.

[29] Austen Bernardi, Roland Faller, Dirk Reith, and Karl N Kirschner. Acpype update for nonuniform 1–4 scale factors: Conversion of the glycam06 force field from amber to gromacs. *SoftwareX*, 10:100241, 2019.

[30] Maral Basma, S Sundara, Dilek Çalgan, Tereza Vernali, and Robert J Woods. Solvated ensemble averaging in the calculation of partial atomic charges. *Journal of computational chemistry*, 22(11):1125–1137, 2001.

[31] William L Jorgensen, Jayaraman Chandrasekhar, Jeffry D Madura, Roger W Impey, and Michael L Klein. Comparison of simple potential functions for simulating liquid water. *The Journal of chemical physics*, 79(2):926–935, 1983.

[32] Ulrich Essmann, Lalith Perera, Max L Berkowitz, Tom Darden, Hsing Lee, and Lee G Pedersen. A smooth particle mesh ewald method. *The Journal of chemical physics*, 103(19):8577–8593, 1995.

[33] Berk Hess, Henk Bekker, Herman JC Berendsen, and Johannes GEM Fraaije. Lincs: a linear constraint solver for molecular simulations. *Journal of computational chemistry*, 18(12):1463–1472, 1997.

[34] Giovanni Bussi, Davide Donadio, and Michele Parrinello. Canonical sampling through velocity rescaling. *The Journal of chemical physics*, 126(1):014101, 2007.

[35] Michele Parrinello and Aneesur Rahman. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied physics*, 52(12):7182–7190, 1981.

[36] Marat Valiev, Eric J Bylaska, Niranjan Govind, Karol Kowalski, Tjerk P Straatsma, Hubertus JJ Van Dam, Dunyou Wang, Jarek Nieplocha, Edoardo Apra, Theresa L Windus, et al. Nwchem: A comprehensive and scalable open-source solution for large scale molecular simulations. *Computer Physics Communications*, 181(9):1477–1489, 2010.

[37] Arunima Singh, Matthew B Tessier, Kari Pederson, Xiaocong Wang, Andre P Venot, Geert-Jan Boons, James H Prestegard, and Robert J Woods. Extension and validation of the glycam force field parameters for modeling glycosaminoglycans. *Canadian journal of chemistry*, 94(11):927–935, 2016.

[38] Jörg Sauter and Andrea Grafmüller. Solution properties of hemicellulose polysaccharides with four common carbohydrate force fields. *Journal of Chemical Theory and Computation*, 11(4):1765–1774, 2015.

[39] Anita Plazinska and Wojciech Plazinski. Comparison of carbohydrate force fields in molecular dynamics simulations of protein–carbohydrate complexes. *Journal of chemical theory and computation*, 17(4):2575–2585, 2021.

[40] Aishwary T Shivgan, Jan K Marzinek, Roland G Huber, Alexander Krah, Richard H Henchman, Paul Matsudaira, Chandra S Verma, and Peter J Bond. Extending the martini coarse-grained force field to n-glycans. *Journal of Chemical Information and Modeling*, 60(8):3864–3883, 2020.

[41] Wesley K Lay, Mark S Miller, and Adrian H Elcock. Optimizing solute–solute interactions in the glycam06 and charmm36 carbohydrate force fields using osmotic pressure measurements. *Journal of chemical theory and computation*, 12(4):1401–1407, 2016.

[42] Jörg Sauter and Andrea Grafmüller. Predicting the chemical potential and osmotic pressure of polysaccharide solutions by molecular simulations. *Journal of chemical theory and computation*, 12(9):4375–4384, 2016.

[43] Sang-Jun Park, Jumin Lee, Yifei Qi, Nathan R Kern, Hui Sun Lee, Sunhwan Jo, InSuk Joung, Keehyung Joo, Jooyoung Lee, and Wonpil Im. Charmm-gui glycan modeler for modeling and simulation of carbohydrates and glycoconjugates. *Glycobiology*, 29(4):320–331, 2019.

[44] Sunhwan Jo, Taehoon Kim, Vidyashankara G Iyer, and Wonpil Im. Charmm-gui: a web-based graphical user interface for charmm. *Journal of computational chemistry*, 29(11):1859–1865, 2008.

[45] Jumin Lee, Xi Cheng, Jason M Swails, Min Sun Yeom, Peter K Eastman, Justin A Lemkul, Shuai Wei, Joshua Buckner, Jong Cheol Jeong, Yifei Qi, et al. Charmm-gui input generator for namd, gromacs, amber, openmm, and charmm/openmm simulations using the charmm36 additive force field. *Journal of chemical theory and computation*, 12(1):405–413, 2016.

[46] Stewart R Durell, Bernard R Brooks, and Arieh Ben-Naim. Solvent-induced forces between two hydrophilic groups. *The Journal of Physical Chemistry*, 98(8):2198–2202, 1994.

[47] Magnus Själander, Magnus Jahre, Gunnar Tufte, and Nico Reissmann. Epic: An energy-efficient, high-performance gpgpu computing research infrastructure. *arXiv preprint arXiv:1912.05848*, 2019.

[48] Jochen S Hub, Bert L De Groot, and David Van Der Spoel. g_wham - a free weighted histogram analysis implementation including robust error and autocorrelation estimates. *Journal of chemical theory and computation*, 6(12):3713–3720, 2010.

[49] Siewert J Marrink, H Jelger Risselada, Serge Yefimov, D Peter Tieleman, and Alex H De Vries. The martini force field: coarse grained model for biomolecular simulations. *The journal of physical chemistry B*, 111(27):7812–7824, 2007.

[50] Paulo CT Souza, Riccardo Alessandri, Jonathan Barnoud, Sebastian Thallmair, Ignacio Faustino, Fabian Grünewald, Ilias Patmanidis, Haleh Abdizadeh, Bart MH Bruininks, Tsjerk A Wassenaar, et al. Martini 3: a general purpose force field for coarse-grained molecular dynamics. *Nature methods*, 18(4):382–388, 2021.

[51] Clément Arnarez, Jaakko J Uusitalo, Marcelo F Masman, Helgi I Ingólfsson, Djurre H De Jong, Manuel N Melo, Xavier Periole, Alex H De Vries, and Siewert J Marrink. Dry martini, a coarse-grained force field for lipid membrane simulations with implicit solvent. *Journal of chemical theory and computation*, 11(1):260–275, 2015.

[52] N Goga, AJ Rzepiela, AH De Vries, SJ Marrink, and HJC Berendsen. Efficient algorithms for langevin and dpd dynamics. *Journal of chemical theory and computation*, 8(10):3637–3649, 2012.

[53] Florian Müller-Plathe. Coarse-graining in polymer simulation: from the atomistic to the mesoscopic scale and back. *ChemPhysChem*, 3(9):754–769, 2002.

[54] Hossein Ali Karimi-Varzaneh, Nico FA Van Der Vegt, Florian Müller-Plathe, and Paola Carbone. How good are coarse-grained polymer models? a comparison for atactic polystyrene. *ChemPhysChem*, 13(15):3428–3439, 2012.

[55] Alexander P Lyubartsev and Aatto Laaksonen. Calculation of effective interaction potentials from radial distribution functions: A reverse monte carlo approach. *Physical Review E*, 52(4):3730, 1995.

[56] Martin Kröger. Shortest multiple disconnected path for the analysis of entanglements in two-and three-dimensional polymeric systems. *Computer physics communications*, 168(3):209–232, 2005.

[57] Sachin Shanbhag and Martin Kröger. Primitive path networks generated by annealing and geometrical methods: Insights into differences. *Macromolecules*, 40(8):2897–2903, 2007.

[58] Robert S Hoy, Katerina Foteinopoulou, and Martin Kröger. Topological analysis of polymeric melts: Chain-length effects and fast-converging estimators for entanglement length. *Physical Review E*, 80(3):031803, 2009.

[59] Nikos Ch Karayiannis and Martin Kröger. Combined molecular algorithms for the generation, equilibration and topological analysis of entangled polymers: Methodology and performance. *International journal of molecular sciences*, 10(11):5054–5089, 2009.

[60] H_ Benoit and P_ Doty. Light scattering from non-gaussian chains. *The Journal of Physical Chemistry*, 57(9):958–963, 1953.

[61] Takahiro Sato, Takashi Norisuye, and Hiroshi Fujita. Double-stranded helix of xanthan: dimensional and hydrodynamic properties in 0.1 m aqueous sodium chloride. *Macromolecules*, 17(12):2696–2700, 1984.

[62] Jorge Ramírez. Computational Soft Matter Lab: Multiple-tau Correlator, 2010. `https://blogs.upm.es/compsoftmatter/software/multiple-tau-correlator/` [accessed 09.06.2022].

[63] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[64] Naveen Michaud-Agrawal, Elizabeth J Denning, Thomas B Woolf, and Oliver Beckstein. Mdanalysis: a toolkit for the analysis of molecular dynamics simulations. *Journal of computational chemistry*, 32(10):2319–2327, 2011.

[65] Richard J Gowers, Max Linke, Jonathan Barnoud, Tyler John Edward Reddy, Manuel N Melo, Sean L Seyler, Jan Domanski, David L Dotson, Sébastien Buchoux, Ian M Kenney, et al. Mdanalysis: a python package for the rapid analysis of molecular dynamics simulations. Technical report, Los Alamos National Lab.(LANL), Los Alamos, NM (United States), 2019.

[66] Monica Bulacu, Nicolae Goga, Wei Zhao, Giulia Rossi, Luca Monticelli, Xavier Periole, D Peter Tieleman, and Siewert J Marrink. Improved angle potentials for coarse-grained molecular dynamics simulations. *Journal of chemical theory and computation*, 9(8):3282–3292, 2013.

[67] Semen O Yesylevskyy, Lars V Schäfer, Durba Sengupta, and Siewert J Marrink. Polarizable water model for the coarse-grained martini force field. *PLoS computational biology*, 6(6):e1000810, 2010.

[68] Robert Seale, Edwin R Morris, and David A Rees. Interactions of alginates with univalent cations. *Carbohydrate Research*, 110(1):101–112, 1982.

[69] Kurt Ingar Draget, Olav Gåserød, Ingrid Aune, Peder O Andersen, Bente Storbakken, Bjørn Torger Stokke, and Olav Smidsrød. Effects of molecular weight and elastic segment flexibility on syneresis in ca-alginate gels. *Food Hydrocolloids*, 15(4-6):485–490, 2001.

[70] Ivan Donati, Synnøve Holtan, Yrr A Mørch, Massimiliano Borgogna, Mariella Dentini, and Gudmund Skjåk-Bræk. New hypothesis on the role of alternating sequences in calcium- alginate gels. *Biomacromolecules*, 6(2):1031–1040, 2005.

[71] Ida Tungesvik Leirvåg. Strategies for stabilising calcium alginate gel beads: Studies of chitosan oligomers, alginate molecular weight and concentration. Master's thesis, NTNU, 2017.

[72] Lianqi Cao, Wei Lu, Analucia Mata, Katsuyoshi Nishinari, and Yapeng Fang. Egg-box model-based gelation of alginate and pectin: A review. *Carbohydrate polymers*, 242:116389, 2020.

[73] Knut Arne Strand, A Bøe, PS Dalberg, T Sikkeland, and Olav Smidsroed. Dynamic and static light scattering on aqueous solutions of sodium alginate. *Macromolecules*, 15(2):570–579, 1982.

[74] Inger Mari Nygård Vold, Kåre A Kristiansen, and Bjørn E Christensen. A study of the chain stiffness and extension of alginates, in vitro epimerized alginates, and periodate-oxidized alginates using size-exclusion chromatography combined with light scattering and viscosity detectors. *Biomacromolecules*, 7(7):2136–2146, 2006.

[75] Isabelle Braccini, Robert P Grasso, and Serge Pérez. Conformational and configurational features of acidic polysaccharides and their interactions with calcium ions: a molecular modeling investigation. *Carbohydrate Research*, 317(1-4):119–130, 1999.

[76] Ying Li, Brendan C Abberton, Martin Kröger, and Wing Kam Liu. Challenges in multiscale modeling of polymer dynamics. *Polymers*, 5(2):751–832, 2013.

[77] David R Hansen and Mitchel Shen. Viscoelastic retardation time computations for homogeneous block copolymers. *Macromolecules*, 8(3):343–348, 1975.

[78] Jingyu Shao, Nuofei Jiang, Hongdong Zhang, Yuliang Yang, and Ping Tang. Sticky rouse model and molecular dynamics simulation for dual polymer networks. *Macromolecules*, 2022.

[79] Nuofei Jiang, Hongdong Zhang, Ping Tang, and Yuliang Yang. Linear viscoelasticity of associative polymers: Sticky rouse model and the role of bridges. *Macromolecules*, 53(9):3438–3451, 2020.

[80] Mei Zheng, Andres Jaramillo-Botero, Xue-hai Ju, and William A Goddard. Coarse-grained force-field for large scale molecular dynamics simulations of polyacrylamide and polyacrylamide-gels based on quantum mechanics. *Physical Chemistry Chemical Physics*, 23(18):10909–10918, 2021.

6 Appendix

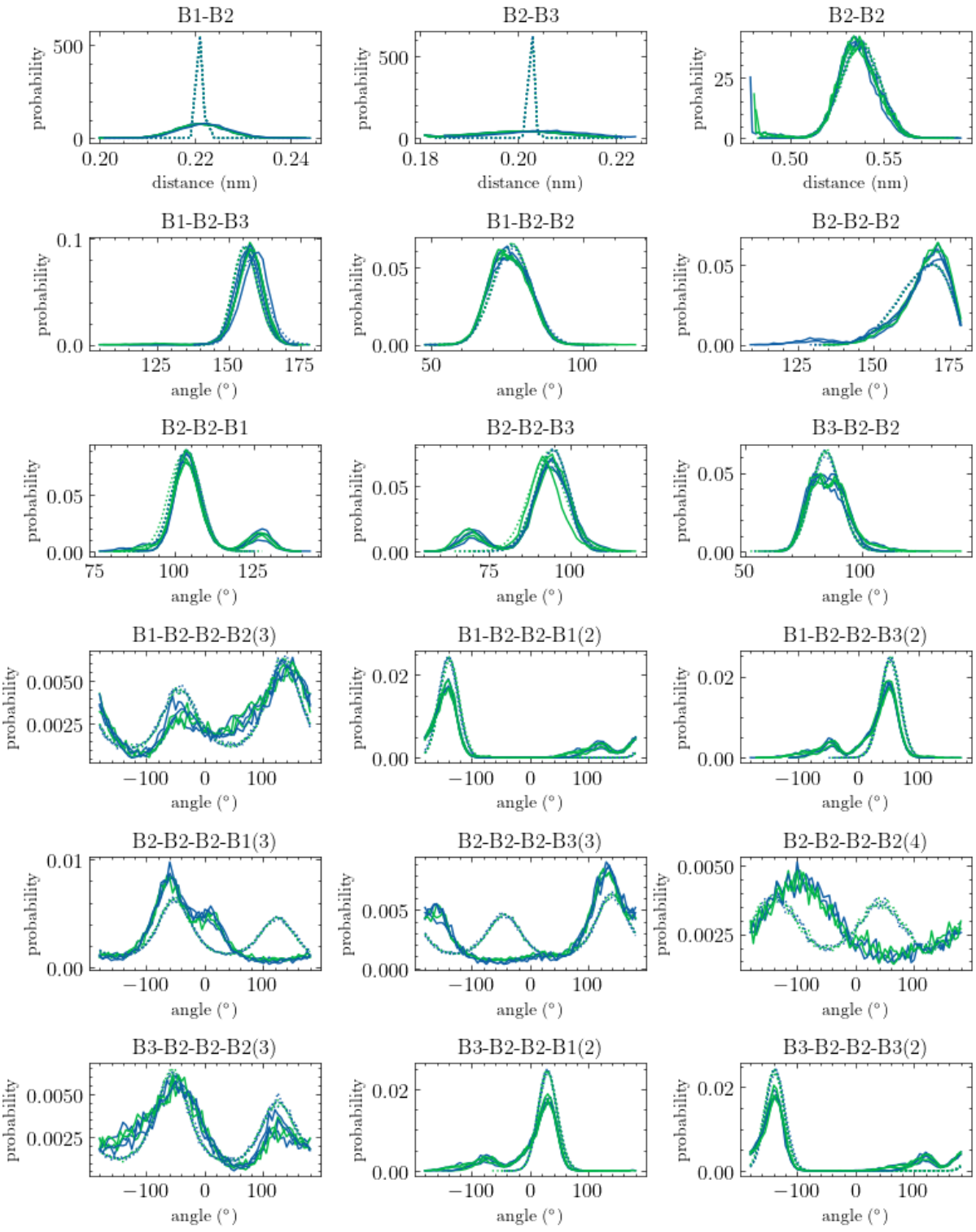


Figure 22: Mapped GLYCAM06 and CG bonded distributions for 7MM. Whole lines represent GLYCAM06 and dotted lines the CG model. The number in parenthesis following the dihedral identifiers refers to the number of consecutive residues over which the dihedrals runs.

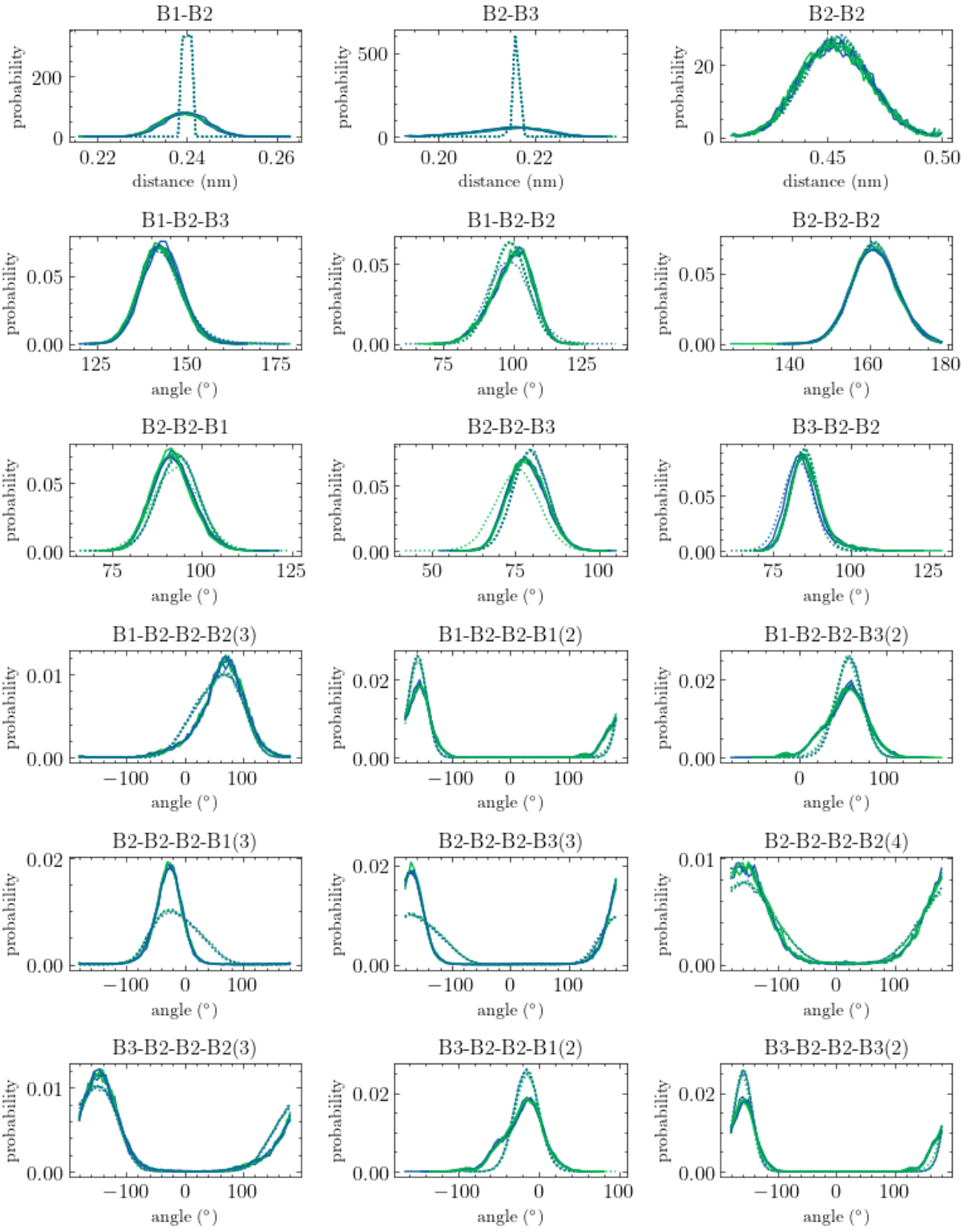


Figure 23: Mapped GLYCAM06 and CG bonded distributions for 7GG. Whole lines represent GLYCAM06 and dotted lines the CG model. The number in parenthesis following the dihedral identifiers refers to the number of consecutive residues over which the dihedrals runs.

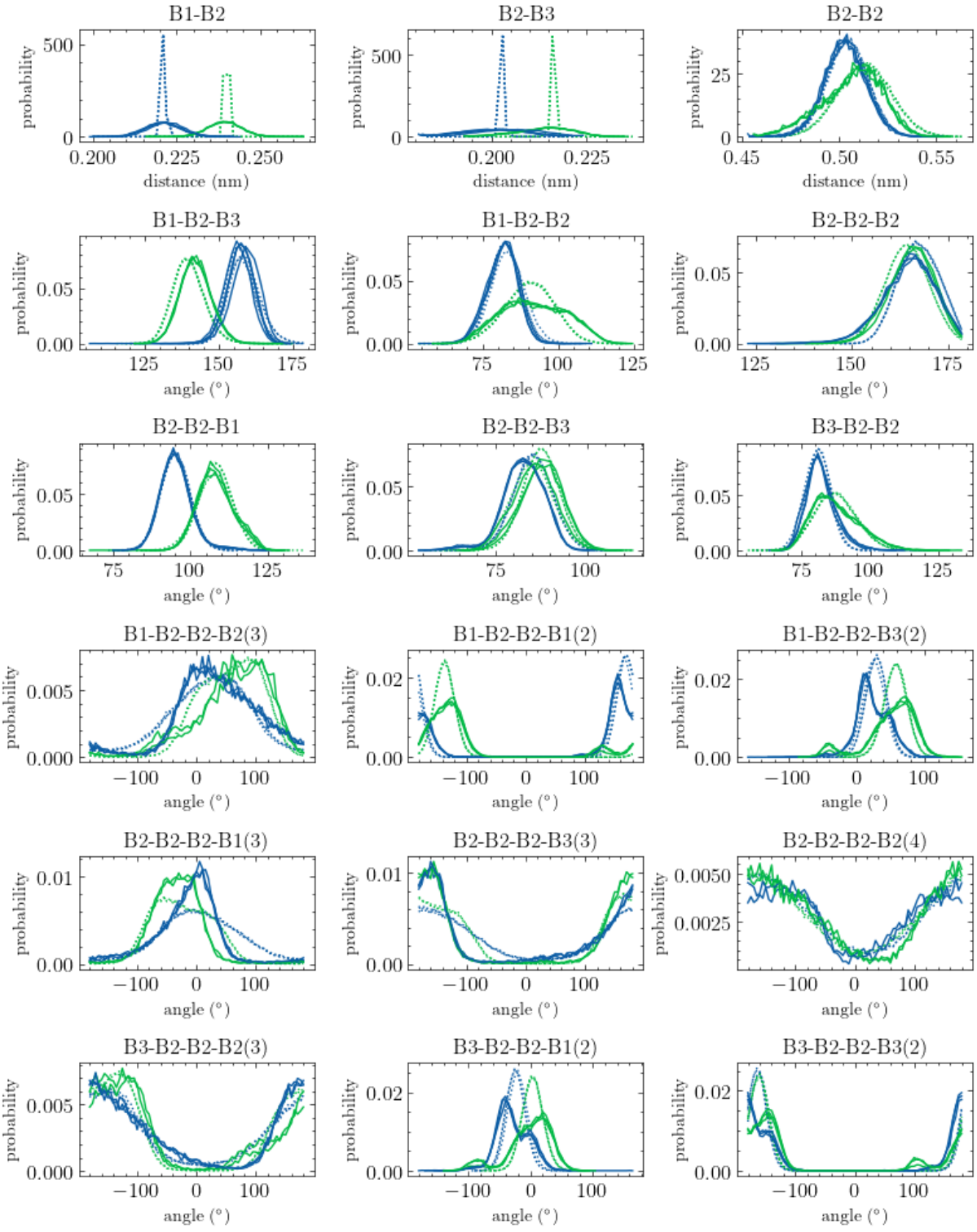


Figure 24: Mapped GLYCAM06 and CG bonded distributions for 7MG. Whole lines represent GLYCAM06 and dotted lines the CG model. The number in parenthesis following the dihedral identifiers refers to the number of consecutive residues over which the dihedrals runs.

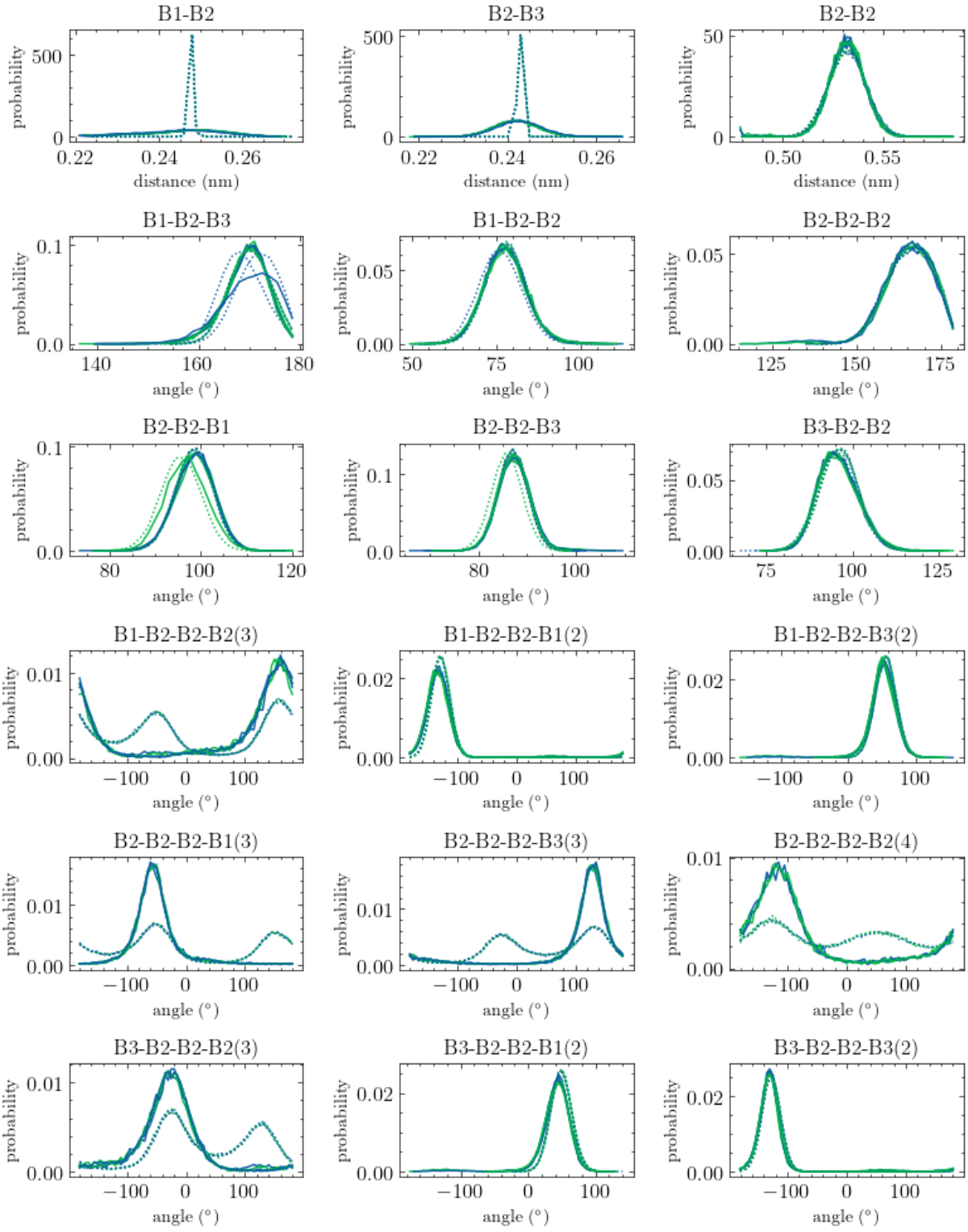


Figure 25: Mapped GLYCAM06 and CG bonded distributions for 7CH. Whole lines represent GLYCAM06 and dotted lines the CG model. The number in parenthesis following the dihedral identifiers refers to the number of consecutive residues over which the dihedrals runs.

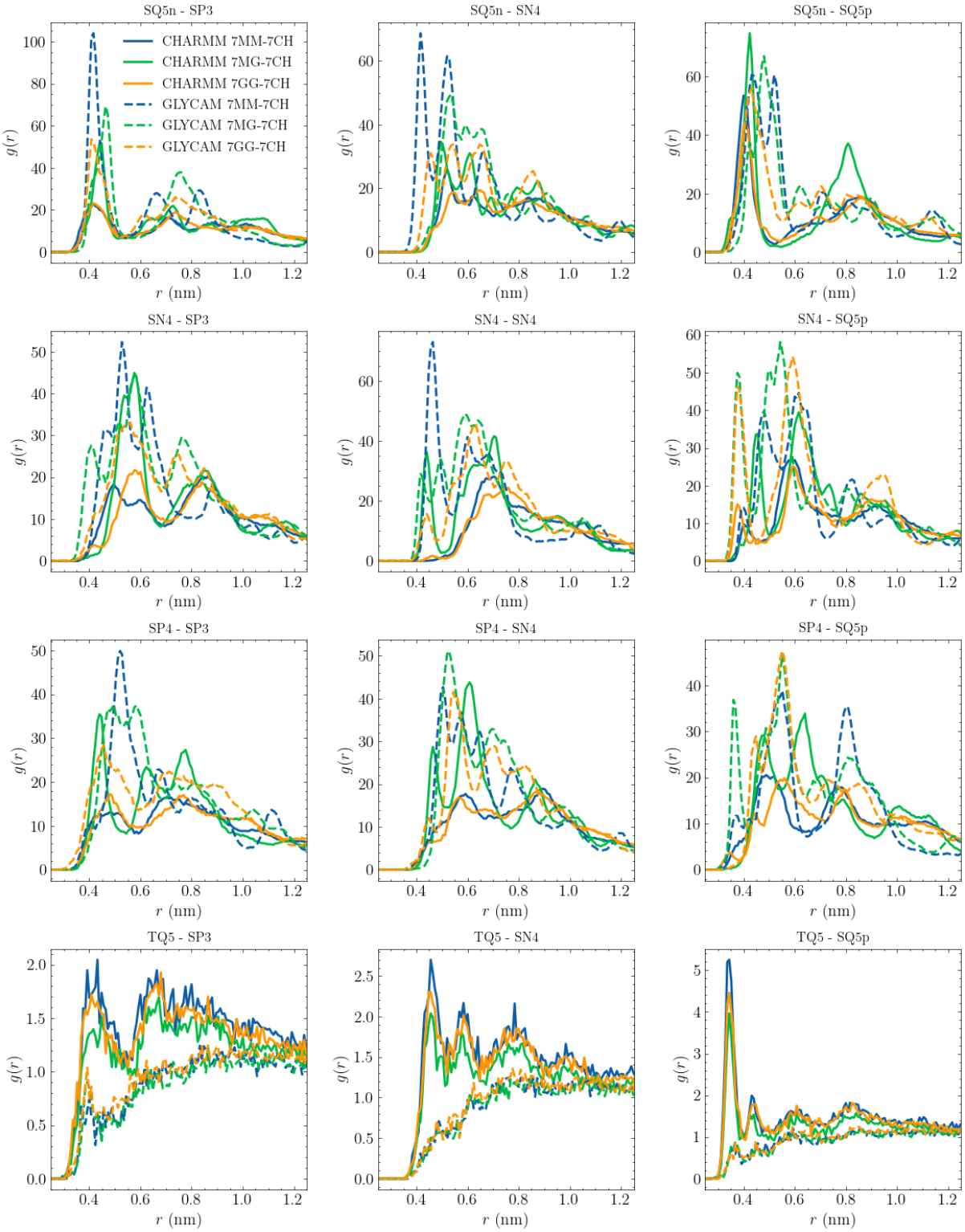


Figure 26: Radial distribution functions for the mapped CHARMM36 trajectory and the mapped GLYCAM06 trajectory for systems 7MM-7CH, 7MG-7CH and 7GG-7CH.

Figure 27: Radial distribution functions for the mapped CHARMM36 trajectory and the mapped GLYCAM06 trajectory for systems 2-7MM, 2-7MG and 2-7GG.

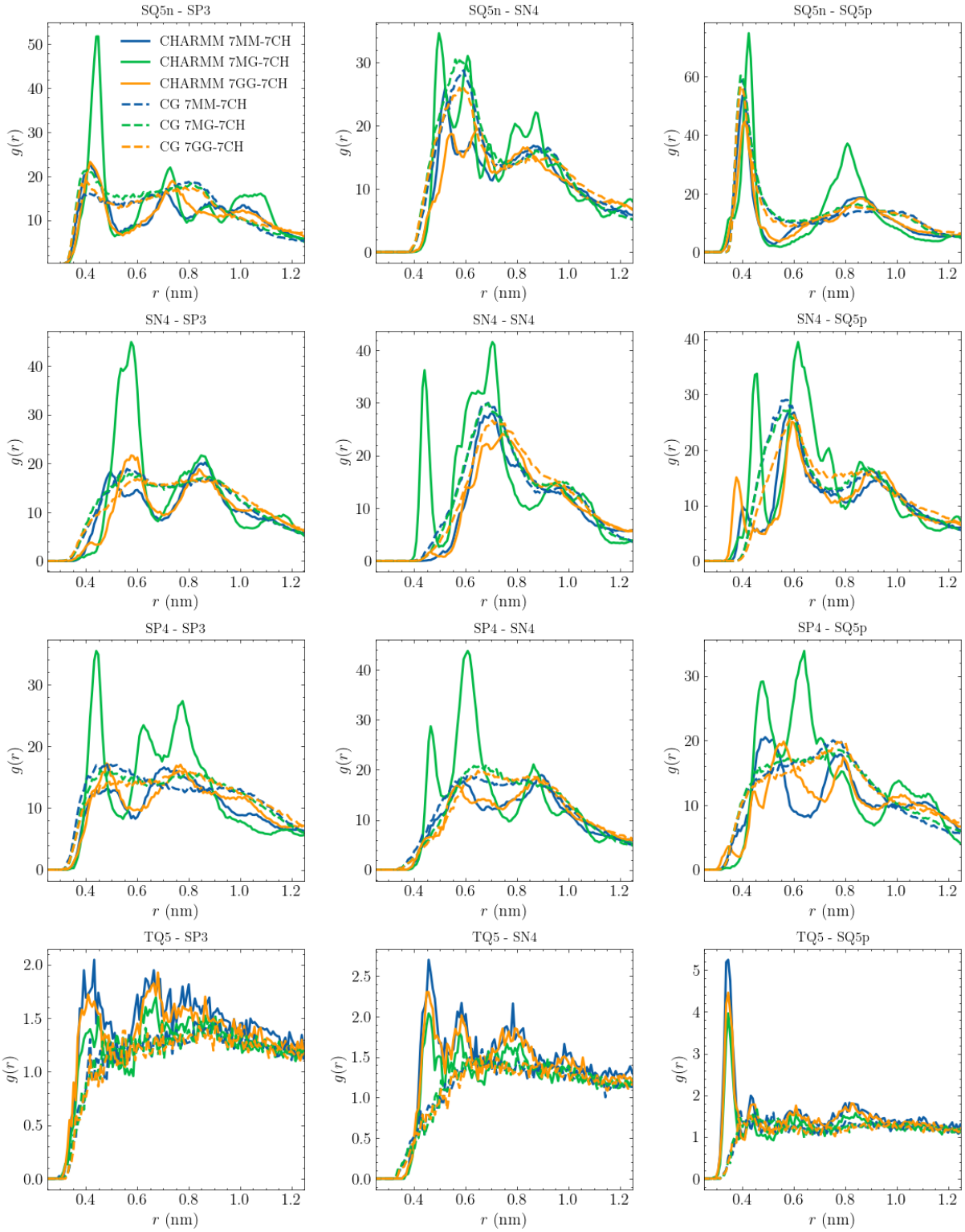


Figure 28: Radial distribution functions for the mapped CHARMM36 trajectory and the CG model with NB1 for systems 7MM-7CH, 7MG-7CH, and 7GG-7CH.

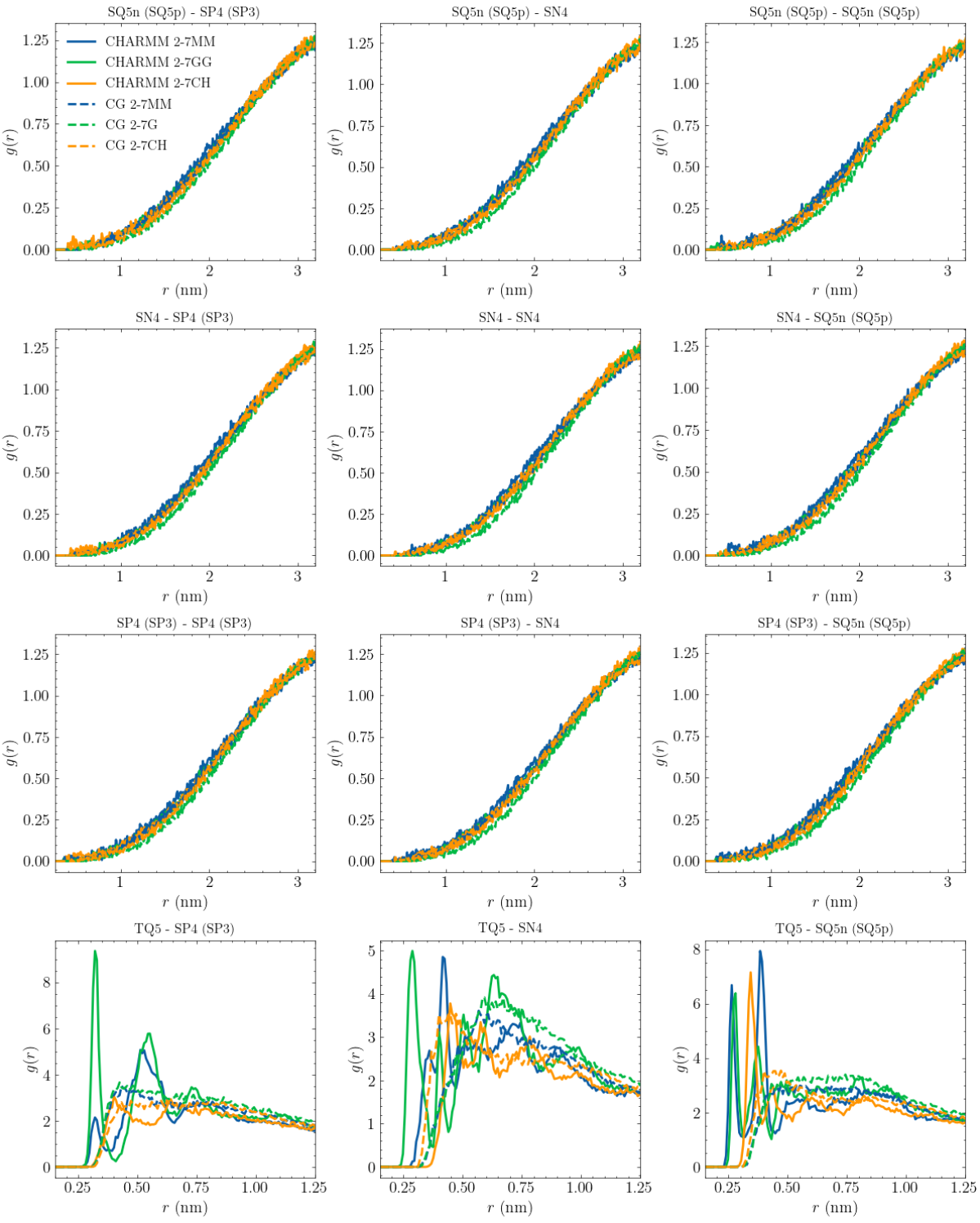


Figure 29: Radial distribution functions for the mapped CHARMM36 trajectory and the CG model with NB1 for systems 2-7MM, 2-7GG, and 2-7CH. Note that the 2-7MG system is exchanged with the 2-7CH system. Also, note the different x-axis limits in the uppermost plots.

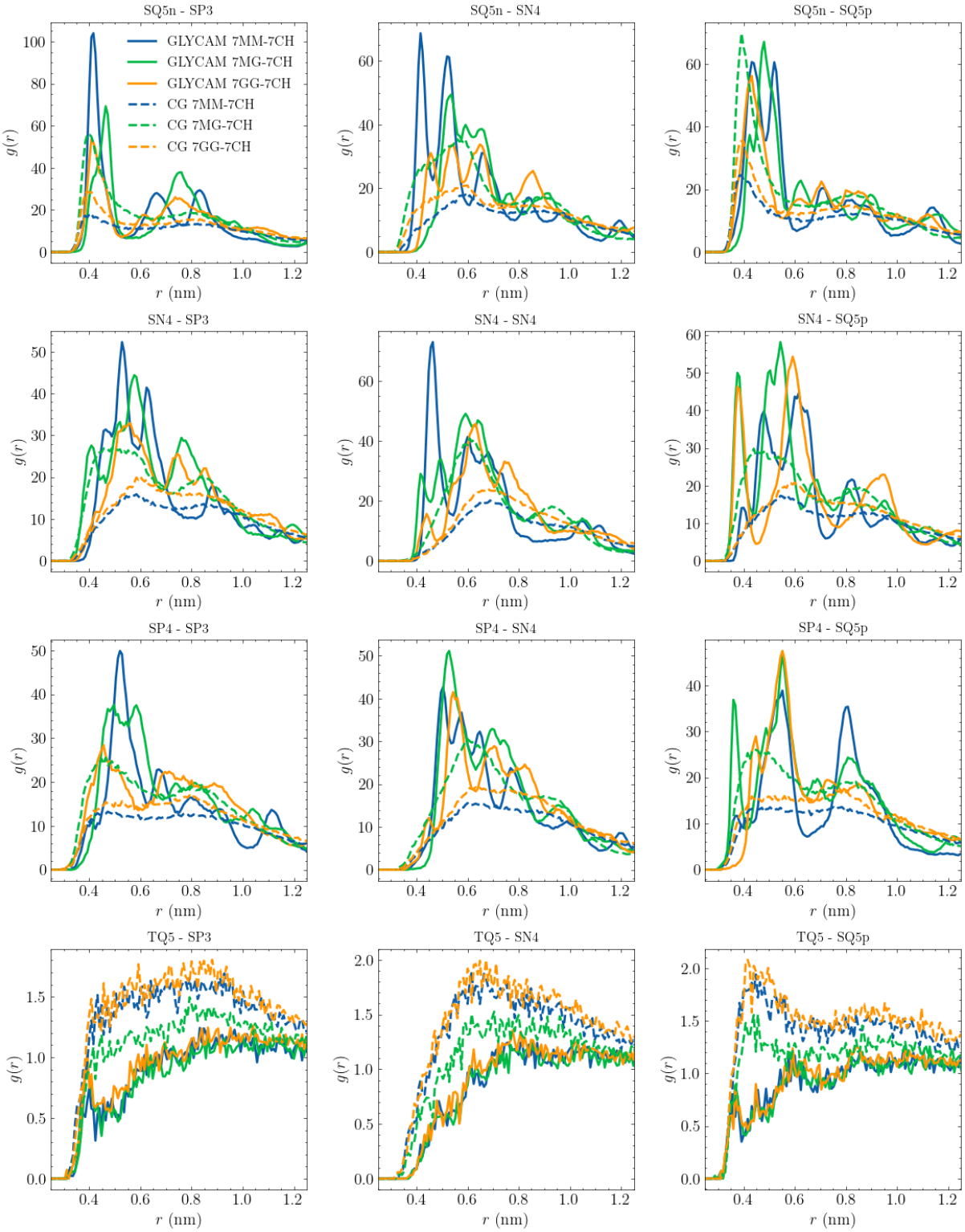


Figure 30: Radial distribution functions for the mapped GLYCAM06 trajectory and the CG model with NB2 for systems 7MM-7CH, 7MG-7CH, and 7GG-7CH.

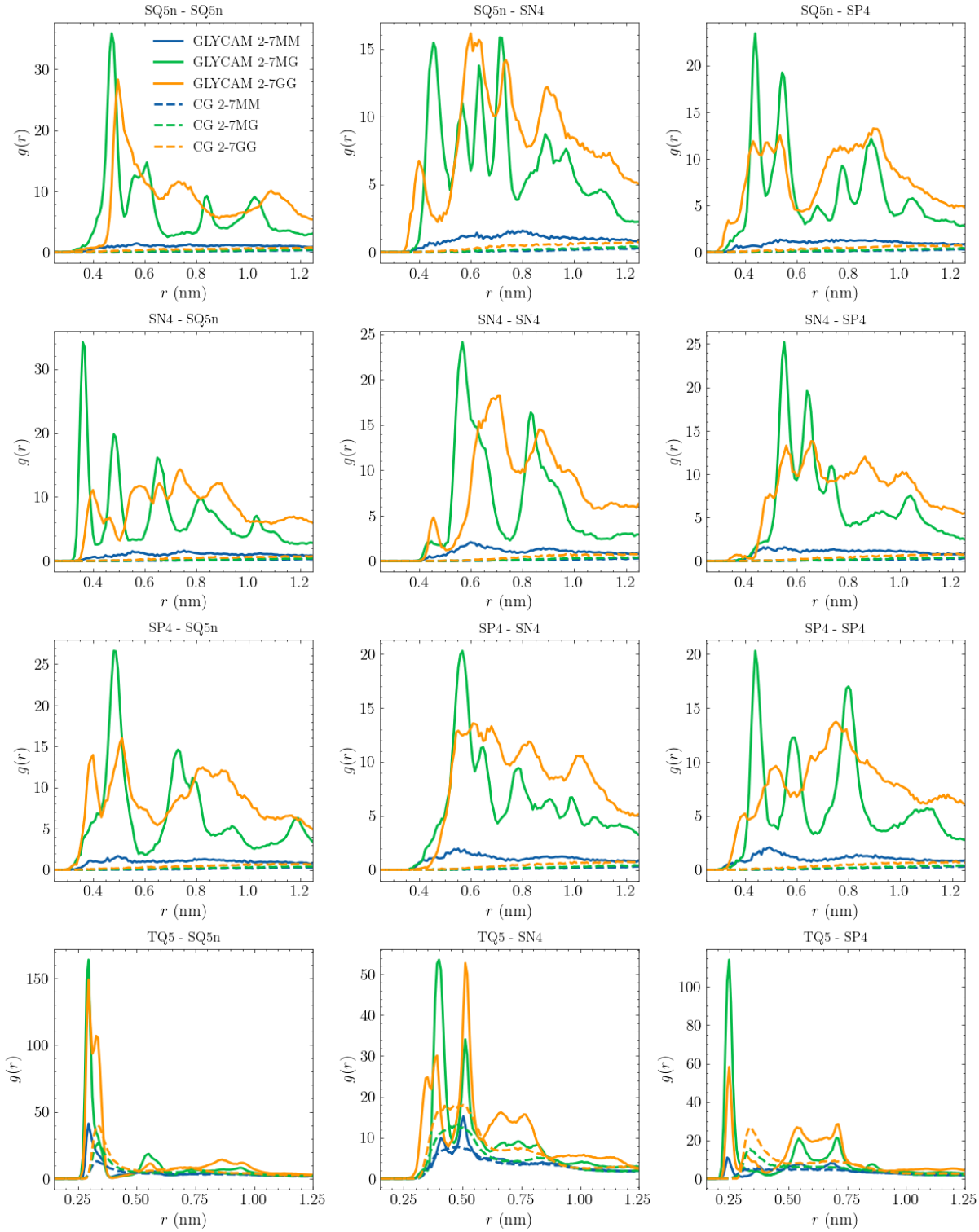


Figure 31: Radial distribution functions for the mapped GLYCAM06 trajectory and the CG model with NB2 for systems 2-7MM, 2-7MG, and 2-7GG.

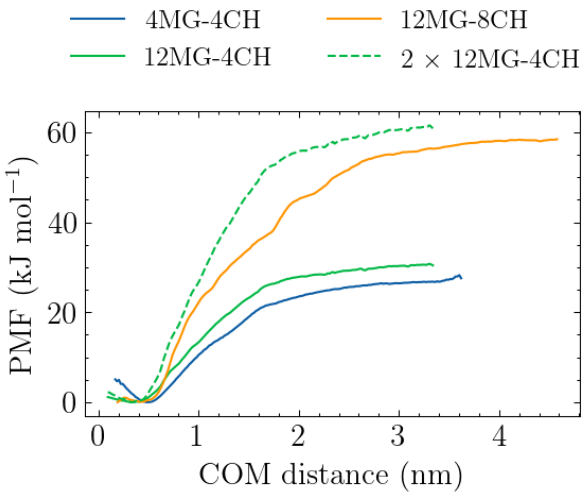


Figure 32: PMFs as a function of alginate - chitosan COM distance obtained from atomistic simulations. The dashed line "2×12MG-4CH" refers to the profile of 12MG-4CH multiplied by two.

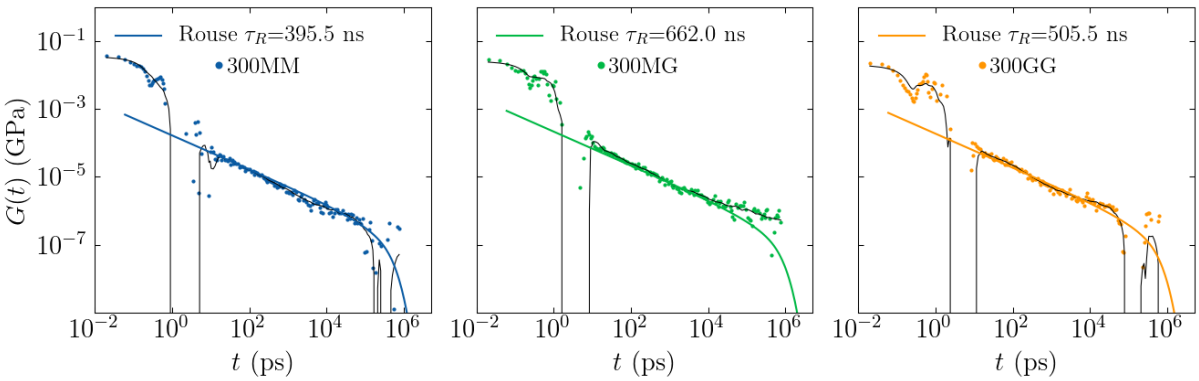
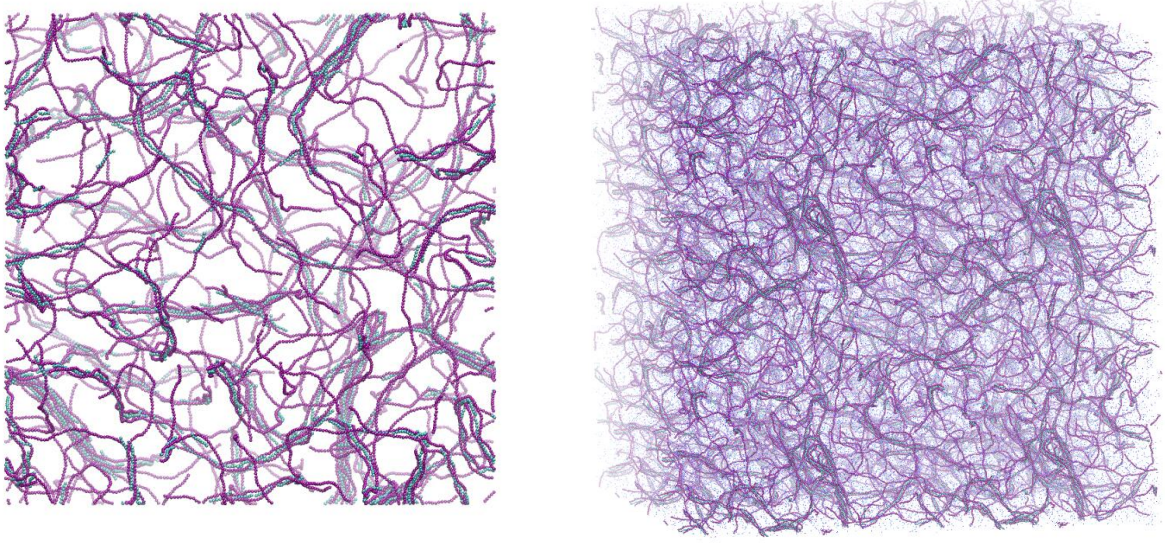
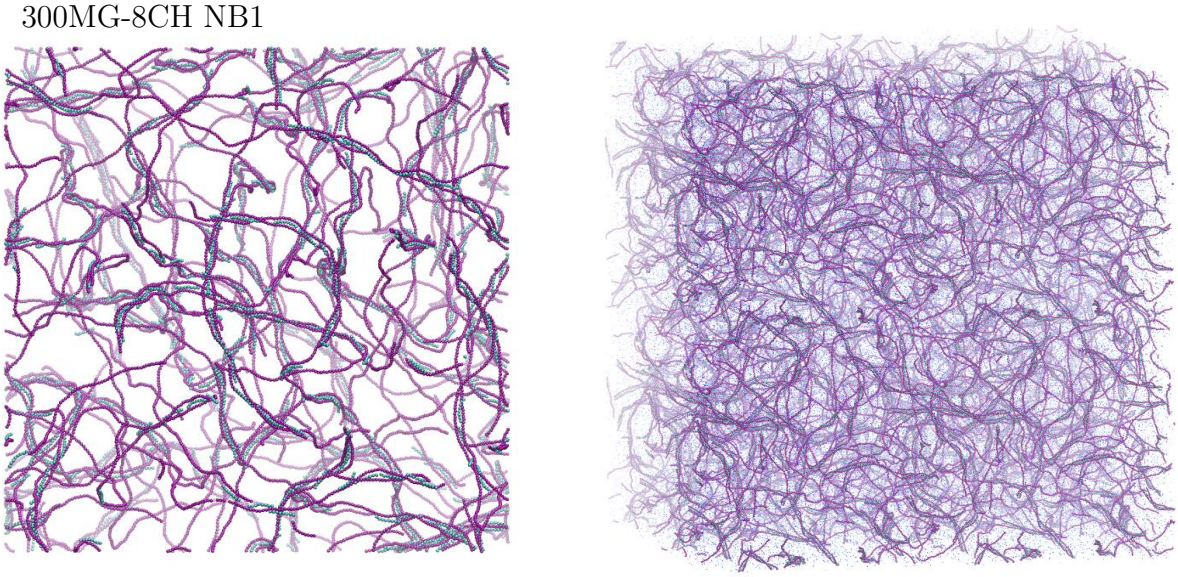


Figure 33: SShear stress relaxation moduli $G(t)$ calculated for 50 alginate chains (DP=300) with NB1 at a concentration of 10 g L^{-1}. The corresponding moduli predicted from the Rouse model (equation 2.23) are also given. The black lines are smoothed signals. The friction coefficient ξ is 1/8 ps^{-1}, as compare to 1/4 ps^{-1} in figre 16.

300MM-8CH NB1



300MG-8CH NB1



300GG-8CH NB1

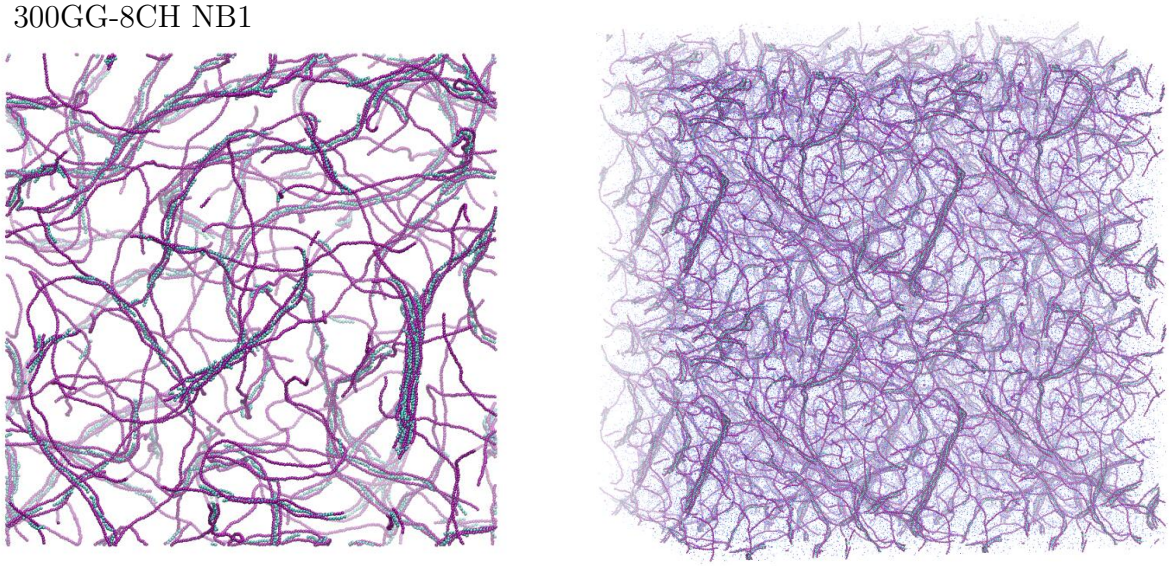
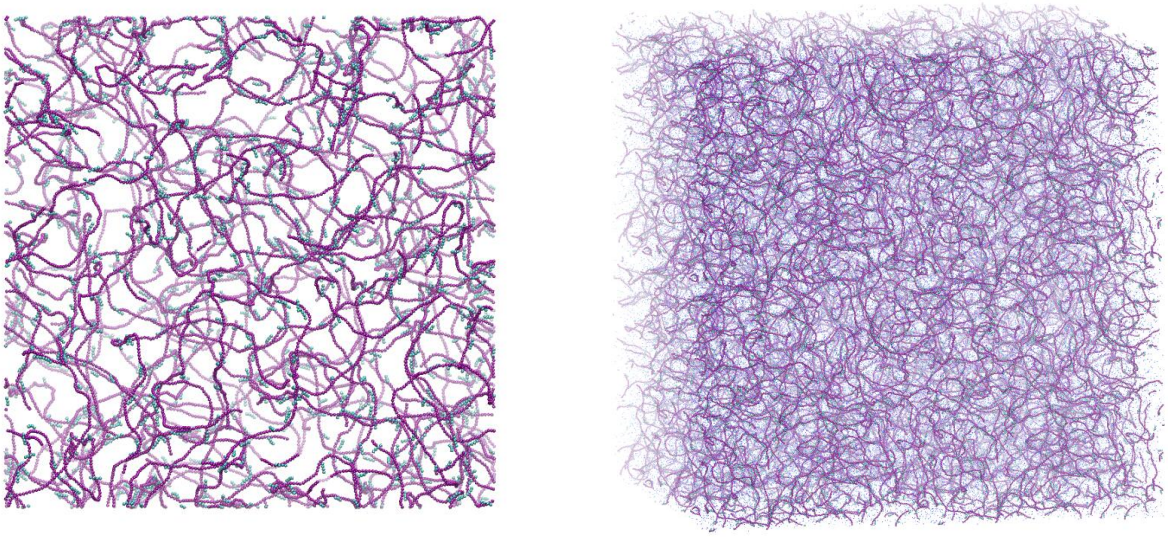
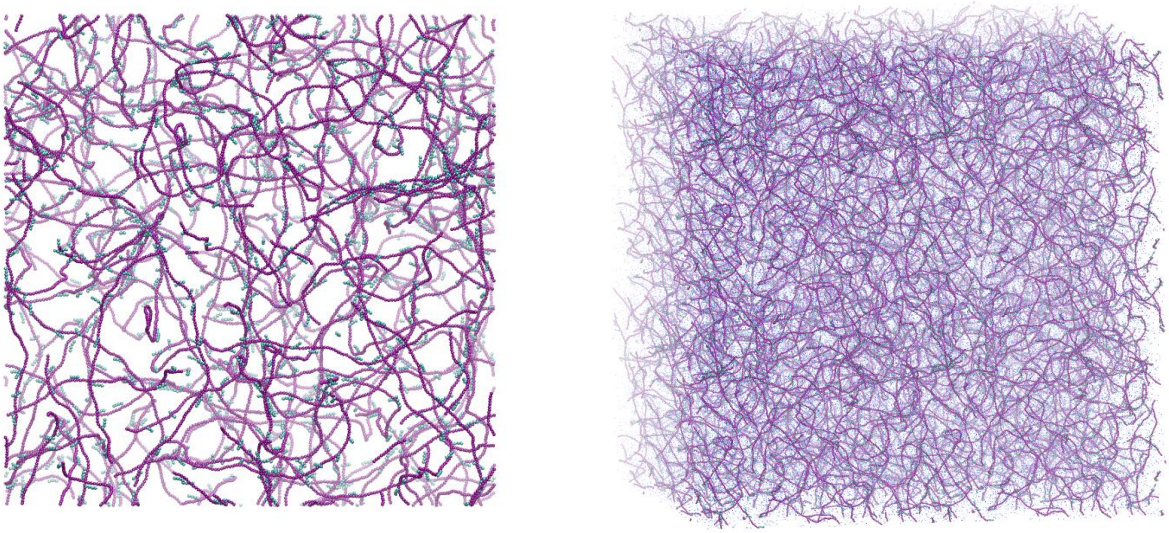


Figure 34: Left) Final snapshots at 1.6 μs of 300-mers with 8CH using NB1. Right) Same as left but rotated, with counterions visible (barely) and one periodic image in each direction. Alginate chains are colored purple and chitosan chains are colored cyan.

300MM-4CH NB1



300MG-4CH NB1



300GG-4CH NB1

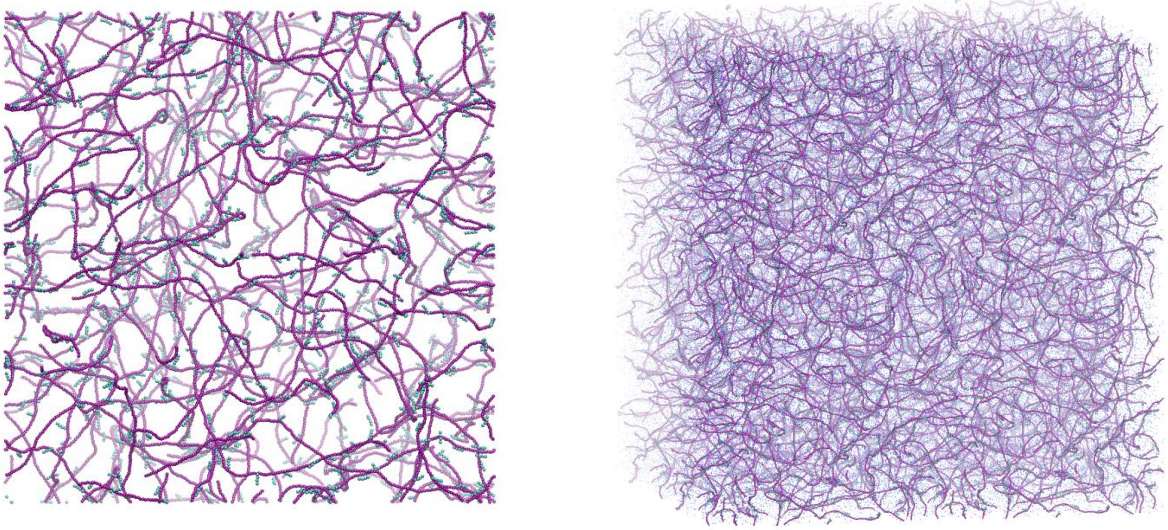
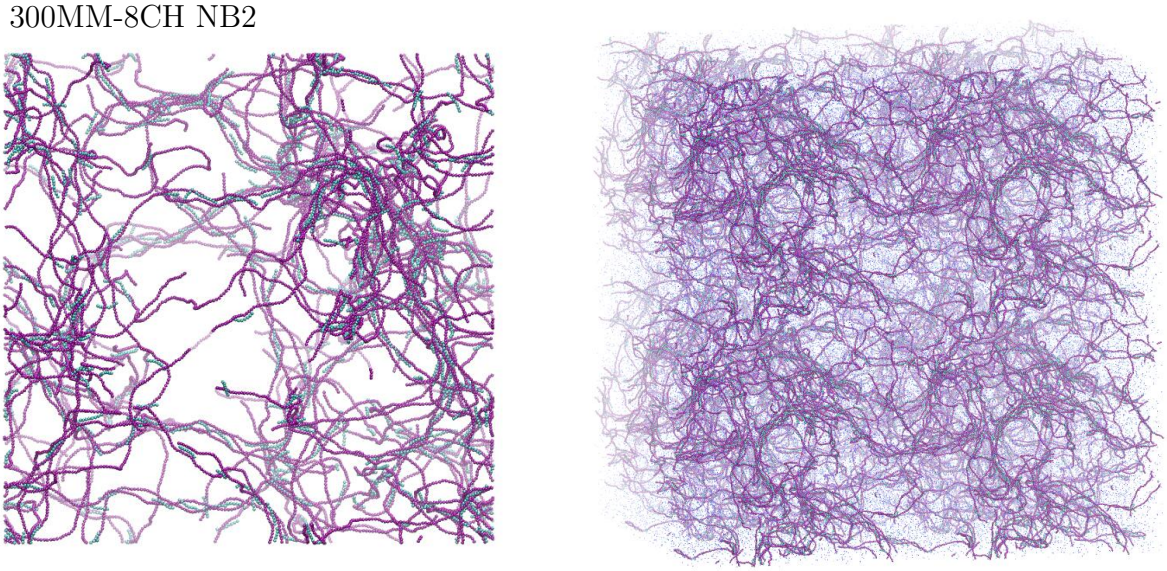
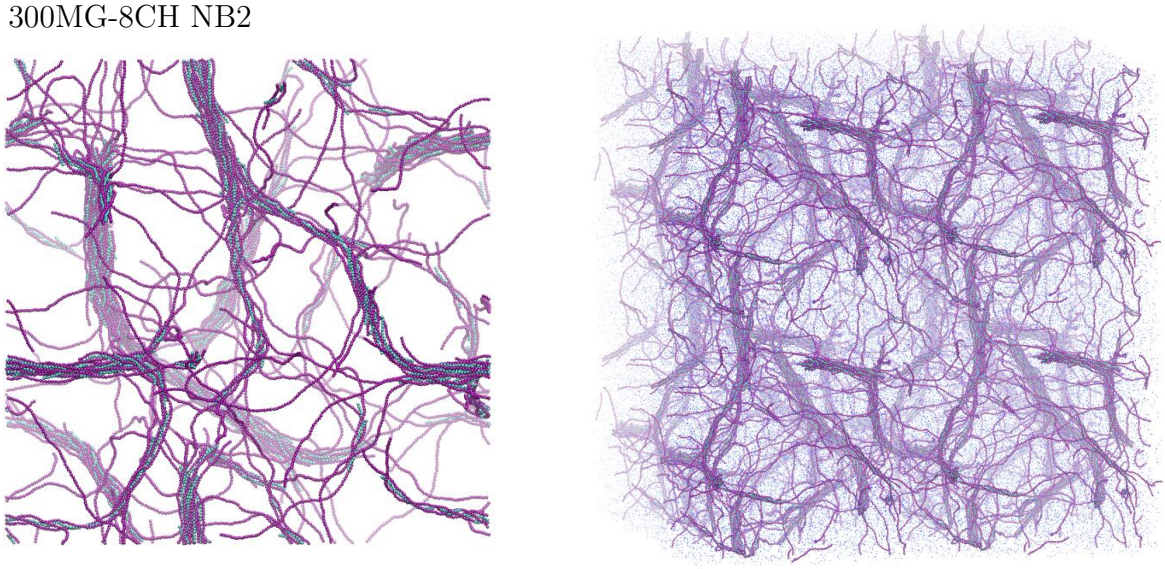


Figure 35: Left) Final snapshots at 1.3 μs of 300-mers with 4CH using NB1. Right) Same as left but rotated, with counterions visible (barely) and one periodic image in each direction. Alginate chains are colored purple and chitosan chains are colored cyan.

300MM-8CH NB2



300MG-8CH NB2



300GG-8CH NB2

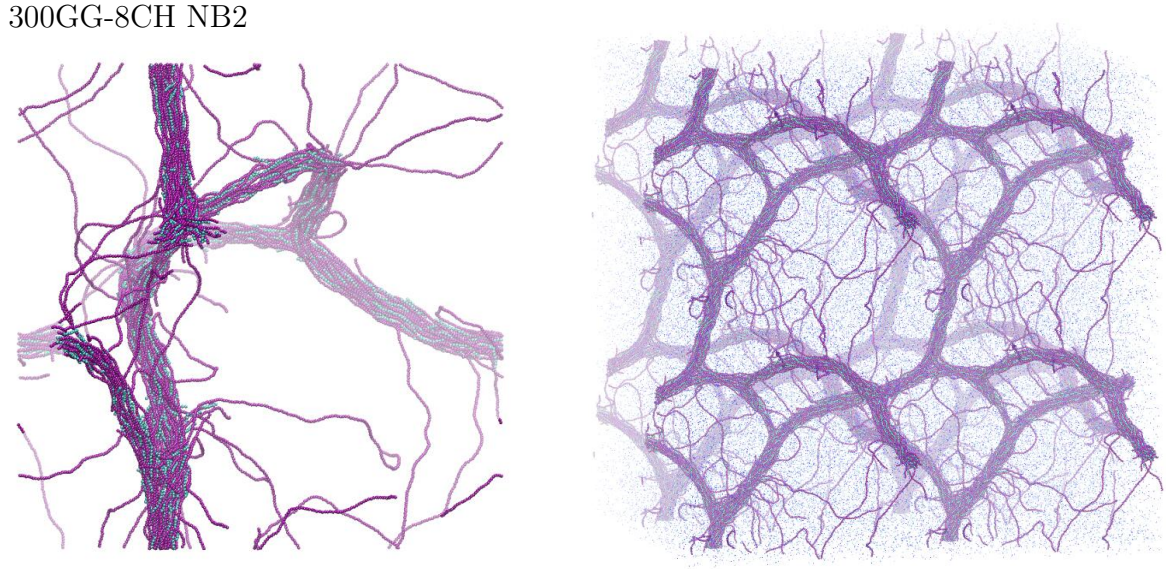


Figure 36: Left) Final snapshots at 0.8 μs of 300-mers with 8CH using NB2. Right) Same as left but rotated, with counterions visible (barely) and one periodic image in each direction. Alginate chains are colored purple and chitosan chains are colored cyan.

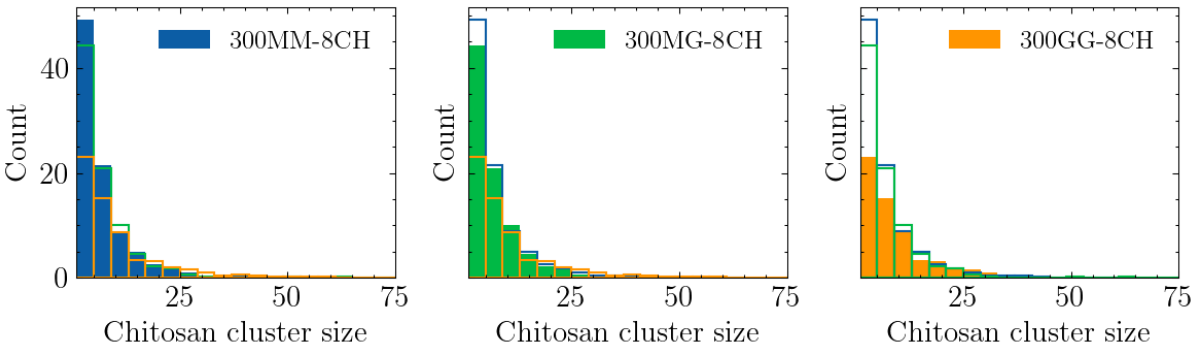


Figure 37: The distribution of chitosan cluster sizes during the last 200 ns of the second replicate simulations calculated with NB1. The chitosan cluster size quantifies the number of clusters formed by chitosan molecules that are within 2.0 nm of each other. The total number of chitosan 8-mers is 564, and the histogram bin width is 4. The insets illustrate all chitosan clusters colored according to their cluster indices in the final simulation snapshot.

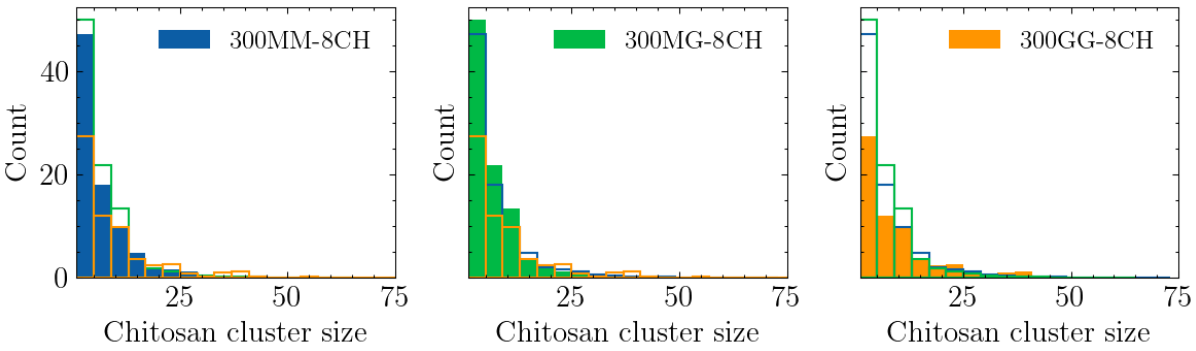


Figure 38: The distribution of chitosan cluster sizes during the last 200 ns of the third replicate simulations calculated with NB1. The chitosan cluster size quantifies the number of clusters formed by chitosan molecules that are within 2.0 nm of each other. The total number of chitosan 8-mers is 564, and the histogram bin width is 4. The insets illustrate all chitosan clusters colored according to their cluster indices in the final simulation snapshot.

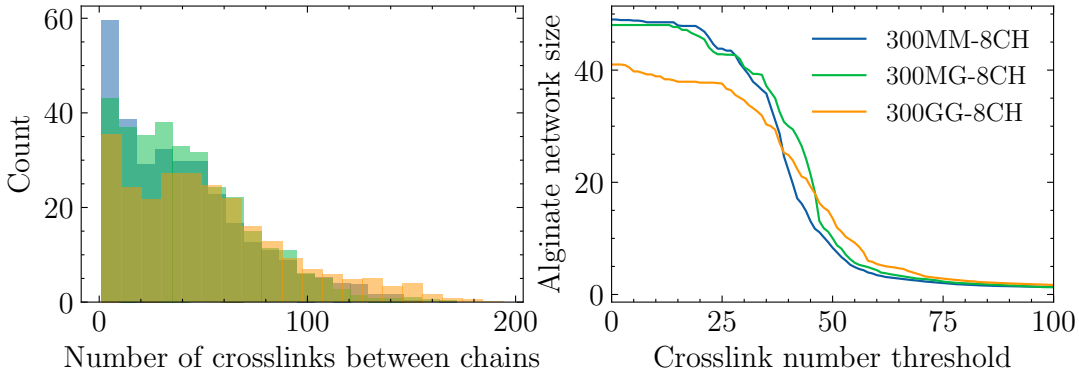


Figure 39: Figures for the third replicate run during the last 200 ns of the simulation with NB1. **Left)** The number of chitosan crosslinks between alginate chains. One chitosan 8-mer can contribute with a maximum of 8 crosslinks. **Right)** The alginate network size as a function of the number of crosslinks between alginate chains. The largest possible network size is 50. The definition of the network size was given in equation 3.4.

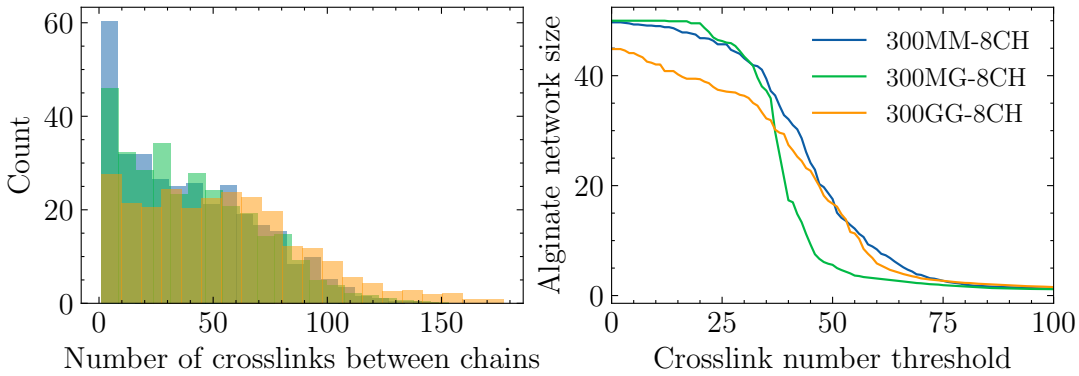


Figure 40: Figures for the third replicate run during the last 200 ns of the simulation with NB1. **Left)** The number of chitosan crosslinks between alginate chains. One chitosan 8-mer can contribute with a maximum of 8 crosslinks. **Right)** The alginate network size as a function of the number of crosslinks between alginate chains. The largest possible network size is 50. The definition of the network size was given in equation 3.4.