Soft Matter

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/softmatter

Journal Name

ARTICLE TYPE

Cite this: DOI: 10.1039/xxxxxxxxxx

Trigger sequence can influence final morphology in the self-assembly of asymmetric telechelic polymers

Aatish Kumar,[a] Christopher P. Lowe,[a] Martien A. Cohen Stuart,[b] and Peter G. Bolhuis*[a]

Received Date
Accepted Date

DOI: 10.1039/xxxxxxxxxx

www.rsc.org/journalname

We report on a numerical study of polymer network formation of asymmetric biomimetic telechelic polymers with two reactive ends based on a self-assembling collagen, elastin or silk-like polypeptide sequence. The two reactive ends of the polymer can be triggered independently by e.g. temperature and pH. We show, using a simple coarse grained model that the order in which this triggering occurs influences the final morphology. For a collagen-silk and the elastin-silk topology we find that for relatively short connector chains the morphology of the assembly is greatly influenced by the order of the trigger, whereas for longer chains the equilibrium situation is more easily achieved. Moreover, self-assembly is greatly enhanced at moderate collagen interaction strength, due to facilitated binding and unbinding of the peptides. This finding indicates that both the trigger sequence and strength can be used to steer self-assembly in these biomimetic polymer systems.

1 Introduction

Telechelic polymers (TP), chain molecules with two reactive end groups, are used as precursors and building blocks for macromolecular structures such as dendrimers.[1] TPs belong to the class of self-associating polymers.[2,3] Such polymers have one or more strongly associating groups (referred to as "stickers") and an inert flexible-chain part (referred to as "connector"). Through the association of stickers, self-associating polymers can form aggregates, which in turn can produce characteristic superstructures with specific structural, mechanical and rheological properties. Telechelic polymers have stickers at both ends of the chain, which can be identical or different, resulting in symmetric or asymmetric TPs, respectively (see Fig. 1).

Recently, biomimetic telechelic polymers with amino acid sequences as the connector and sticker units have attracted much interest.[4] Such polymers are biostable and biodegradable, and are useful for a variety of medical applications[5] such as tissue scaffolding.[6] The connector is a water soluble sequence of several hundreds of residues that does not respond to physicochemical triggers like temperature, pH or additives, and assumes a random coil structure in water under all conditions. The sticker units self-assemble upon appropriate changes in ambient conditions. This self-assembly leads to different morphologies which can be the basis of new materials comprising of nanofibers, nanoparticles and hydrogels.[4,7–10]

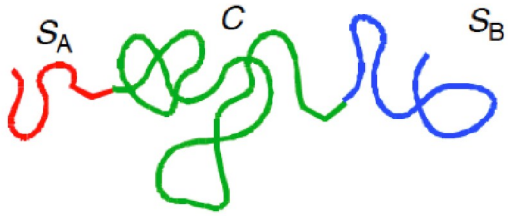


Fig. 1 Cartoon of asymmetric telechelic polymer (S_A, S_B are stickers, C is the connector).

One instance of such a biomimetic polymer is a triblock polymer consisting of silk-like and elastin-like polypeptide blocks separated by a random coil polypeptide block connecting the two end groups.[11,12] The silk-inspired self-assembling element contains several repeats of a glycine-alanine rich octapeptide (GAGAGAGE), in which the eighth amino acid (glutamic acid) is chosen to disrupt the tight, hydrogen-bonded fibrin-like packing promoted by the glycine-alanine repeats. Upon lowering the pH, such silk-inspired polypeptides form highly defined elongated fibers, which are a few nanometer wide and hundreds of nanometers to micrometers long.[13,14] The elastin-like block is composed of a pentapeptide repeat motif (VPGXG), where the X position is occupied by either valine, alanine, or glycine in a 5:3:2 ratio, respectively. Elastins with this kind of sequence are known to undergo a temperature-induced phase transition: above the transition temperature (lower critical solution temperature, LCST),[15] the random-coiled structure changes into a β-spiral, thereby forming aggregates.

Another possibility is a triblock polymer consisting of silk-like

[a] *Amsterdam Center for Multiscale Modeling, van 't Hoff Institute for Molecular Sciences, University of Amsterdam, PO Box 94157 1090 GD Amsterdam, The Netherlands*
[b] *Laboratory of Physical Chemistry and Colloid Science, Wageningen University, Dreijenplein 6, 6703 HB Wageningen, The Netherlands*
* E-mail:p.g.bolhuis@uva.nl

and collagen-like polypeptide blocks separated by a random coil polypeptide block connecting the two end groups. The silk blocks of such TPs can again form the above-mentioned fibril structures. The collagen part consists of several repeats of the tripeptide motif PGP.[16] While being flexible and solvated at high temperature, collagen blocks of three different polymers assemble into a triple helix when the temperature is lowered below the helix melting temperature.

In both triblock TP systems, there are two triggers, pH and temperature, which can be controlled independently. Upon triggering collagen in the collagen-silk TP, triple helix formation results in a solution of star polymers with three arms. The subsequent triggering of silk self-assembly results in a polymer network. Depending on the initial density this will be either a system spanning cluster, or a dispersion of finite clusters. On the other hand, triggering the silk first will result in long silk fibers, which can subsequently be connected by collagen bridges. In the elastin-silk TP, a similar mechanism takes place with the elastin-like block self-assembling in micelles of around twelve monomers instead of star polymers with three arms as is the case for the collagen-like block.

While in principle both trigger sequences should end in the same equilibrium situation, kinetic trapping can leave the system in a deep meta-stable minimum which is different for each route. Indeed, this behavior is often found in polymer systems.[1] However, the molecular origin of the self-assembly process, the final topology, and the structures of the trapped intermediates, are not well understood.

The aim of this work is to provide a physical basis for understanding the non-equilibrium behavior of self-assembling telechelic polymer solutions using molecular simulations,[17] and to gain insight into whether and under what circumstances, the sticker triggering sequence (silk followed by elastin or vice-versa) influences the final network structure. Molecular simulation can predict behavior in atomic detail based on classical force fields. While in principle one could perform an all-atom simulation of these systems, the required time and length-scales are prohibitively large. Therefore we opt for a highly coarse grained (CG) description.[18,19] Such a model cannot account for atomic detail, but can predict large-scale topology, morphology and self-assembly properties at a fraction of the computational cost.

Hydrophilic polymers in a good solvent in the semi-dilute regime can be described by the so-called blob model,[20] in which polymer chains are replaced by a chain of soft repulsive blobs connected by harmonic springs. The softness of the potential arises from the integrating out of the monomer and solvent degrees of freedom. A polymer of thousands of monomers can in this way be modeled by just a few blobs.[21] On top of this soft repulsion a sticker interaction will force the polymers to connect into networks. In this work we apply a generic attractive interaction that essentially distinguishes between bound and unbound states.

Since we are interested in a natural time evolution of the system, we perform regular molecular dynamics (MD) to sample the system. Of course, this dynamics is artificial, since we make use of highly coarse grained systems. However, the use of coarse grained potentials would only enhance equilibration, not prevent it. Here, we show that even the coarse grained systems can already be kinetically trapped. The natural system is constrained by the molecular details, e.g. collagen will only form trimers, and silk will form linear fibrils. To mimic this in the coarse grained system, we restrict the topology of the bound stickers, and impose this constraint using Monte Carlo (MC).[2,22] This MC procedure should mimic the stochastic nature of the binding and unbinding, and assumes that (un)binding, when possible, occurs on a faster timescale than the polymer diffusion timescale.

In this work we focus on the collagen-silk and the elastin-silk TPs. These systems have been recently studied experimentally by Cohen Stuart and co-workers.[4,11–13] The important control variables are the sticker attraction strength, representing temperature, pH and salt concentration. In addition, we vary the polymer length, and the concentration. For different conditions, we compute the long term assembly behavior using our hybrid MD/MC scheme. The morphology of the resulting (finite) clusters is characterized using graph topology and cluster analysis, as a function of polymer density and strength of sticker attraction. Due to our use of a highly coarse grained force field we cannot expect the molecular structure to be correctly reproduced. Hence, direct comparison between experiments and simulations in terms of molecular structure and geometry is at this stage premature. Instead we aim to answer the question whether and how the triggering sequence can influence the assembly morphology and topology.

The remainder of the paper is organized as follows. In the methods section we introduce the model and explain the simulation methodology. In section 3, we present the results for the collagen-silk and the elastin-silk systems. We discuss these results in section 4 and end with concluding remarks in section 5.

2 Methods

We use a highly coarse grained description for the telechelic polymer.[2,21,23] Each telechelic polymer is considered to be a linear chain consisting of N_b units (blobs), with each blob representing several hundreds of amino acids. These blobs are modeled as very soft spheres characterized by the center of mass of the polymer segment, and its radius of gyration R_g. In this simple model both the solvent and the monomer degrees of freedom are integrated out.[21] For a solution of telechelic polymers, we consider a system of N_c identical polymers. The two end units of each polymer are designated as "stickers". Thus, the total number of sticker sites in the system is equal to $N_s = 2N_c$, and total number of blobs is equal to $N_{tot} = N_c N_b$.

The formation of finite clusters is studied using hybrid molecular dynamics (MD) simulations.[2,22] Sticker pair interactions are handled by interspersing MD simulations with Monte Carlo (MC) moves to perform binding of stickers or unbinding of already bound stickers. This hybrid simulation approach[2,22] enables better control of sticker dynamics.

2.1 Interaction potentials

All blob pairs interact via a soft repulsive Gaussian potential U_A[21]

$$\frac{U_A(r_{ij})}{k_B T} = 1.8 \exp\left(\frac{-3r_{ij}^2}{4\bar{R}_g^2}\right) \quad (1)$$

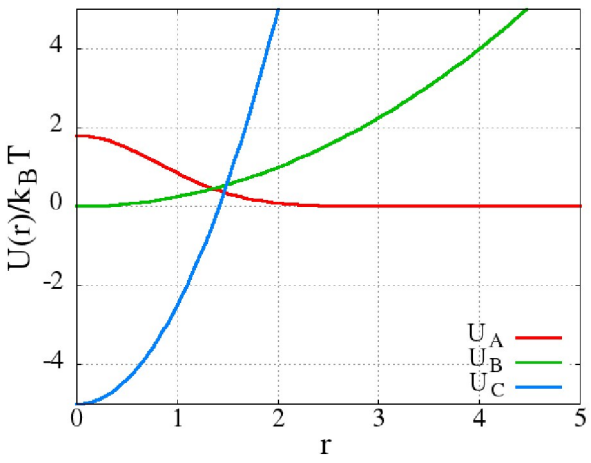


Fig. 2 Interaction potentials (U_A between all blob pairs, U_B between neighboring blobs, and U_C between sticker pairs).

where $r_{ij} = |\vec{r}_i - \vec{r}_j|$ is the separation between blobs i and j, $\bar{R}_g$ is the average radius of gyration of the blobs, k_B is Boltzmann's constant, and T is the temperature. In this work we use a reduced length unit, by setting $\bar{R}_g = 1$. The potential is shifted and truncated sightly so that it vanishes at the cut-off radius (R_{cut}), which is set to 2.

Neighboring blobs in the same polymer chain interact via a harmonic potential U_B,[23,24]

$$\frac{U_B(r_{ij})}{k_B T} = \frac{3}{2}\left(\frac{r_{ij}}{\sqrt{6}\bar{R}_g}\right)^2 \tag{2}$$

where r_{ij} and $\bar{R}_g$ are defined as above.

The sticker blobs interact via a shifted harmonic potential U_C,

$$\frac{U_C(r_{ij})}{k_B T} = \frac{1}{2}k_{attr} r_{ij}^2 - \varepsilon_{attr} \tag{3}$$

where k_{attr} and ε_{attr} determine the range and strength of attraction between the sticker units, and r_{ij} is the separation between stickers i and j. Thus, the binding energy of a sticker pair is given by ε_{attr}. All the above interactions are summarized in Fig. 2. k_{attr} and ε_{attr} are simulation variables (set to $k_{attr} = \varepsilon_{attr} = 6$ in Fig. 2). In all the simulations, k_{attr} is set equal to ε_{attr}, which leads to $r_{ij} = \sqrt{2}$ when $U_C(r_{ij}) = 0$. Therefore, the range of sticker attraction (R_{attr}) remains constant for all simulations. Note the attractive potential is minimal at zero sticker distance, which is allowed because the sticker blobs can overlap.

The total potential energy of the system is given by

$$U = \sum_{i,j=0,i<j}^{N_{tot}-1} U_A(r_{ij}) + \sum_{i=0}^{N_c-1}\sum_{k=0}^{N_b-2} U_B(r_{iN_b+k,iN_b+k+1}) + \sum_{i,j=0,i<j}^{N_c-1}\left[U_C^h(r_{iN_b,jN_b})C_{ij}^h + U_C^t(r_{iN_b+N_b-1,jN_b+N_b-1})C_{ij}^t\right], \tag{4}$$

where the superscripts h and t in the last sum refer to the first and final sticker blob in the TP, respectively. C_{ij} denotes the bond matrix for the stickers, in which an element is unity when a bond is formed between stickers of chain i and j, and zero otherwise. Note that the blob indices run between zero and $N_{tot} - 1$. To avoid complete collapse of all stickers we restrict the number of stickers to which a particular sticker can bind to, and/or the maximum number of bound stickers in a cluster. Note that the bond matrix entries are not directly implied by the positions of particles, but constitute additional degrees of freedom that are turned on or off based on the available energy.

2.2 Simulation system

We investigate two different telechelic polymer systems. One is the collagen-silk system, in which one of the stickers is inspired by a collagen motif while the other sticker is inspired by silk motif. The second system is the elastin-silk system in which the collagen is replaced by an elastin inspired motif. The collagen motif forms a triple helix when triggered, while the elastin motif forms micellar structures when triggered (see Fig. 3). The silk motif has the tendency to form long filaments when triggered (see Fig. 3). As mentioned above, for collagen and elastin the cluster size is restricted to a maximum of three and twelve respectively, whereas the silk can grow indefinitely, with the condition that each sticker can only bind to two other silk stickers and no rings are formed.

We simulate two telechelic polymer chain lengths, a short one ($N_b = 3$) and moderately long one ($N_b = 9$). The number of chains is set to 200 ($N_c = 200$) for all the simulations. To study the effect of concentration (or blob volume fraction) we vary the (cubic) simulation box length. Table 1 lists the used concentrations for the two polymer chain lengths systems. We can define two different volume fractions: the blob volume fraction (ϕ_b) determined by all the blobs together (sticker and connector), and the sticker volume fraction (ϕ_s) determined by sticker blobs only. The blob volume fraction is set to 0.01, 0.015, and 0.5 for the short chain systems, and 0.01, 0.05 and 0.15 for moderately long chain systems. As depicted in table 1, this enables us to compare blob volume fractions and sticker volume fractions directly between the two chain lengths, thus allowing to see how the behavior for polymer concentration and chain length is related.

We note that the our bead volume fraction never exceeds the overlap concentration. Indeed, the potential U_A was developed for dilute polymer solution.[21] Because of the chain of beads model, an effective semidilute concentration for the whole chain is possible.[23,24]

Table 1 Simulated Volume Fractions

ϕ_b	ϕ_s	
	$N_b = 3$	$N_b = 9$
0.01	-	-
0.015	0.010	-
0.05	0.033	0.011
0.15	-	0.033

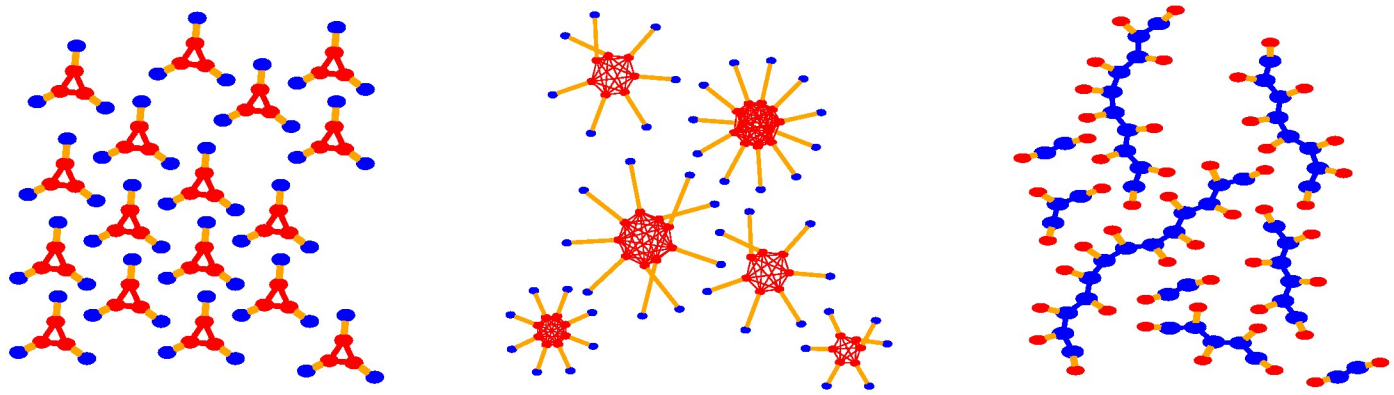
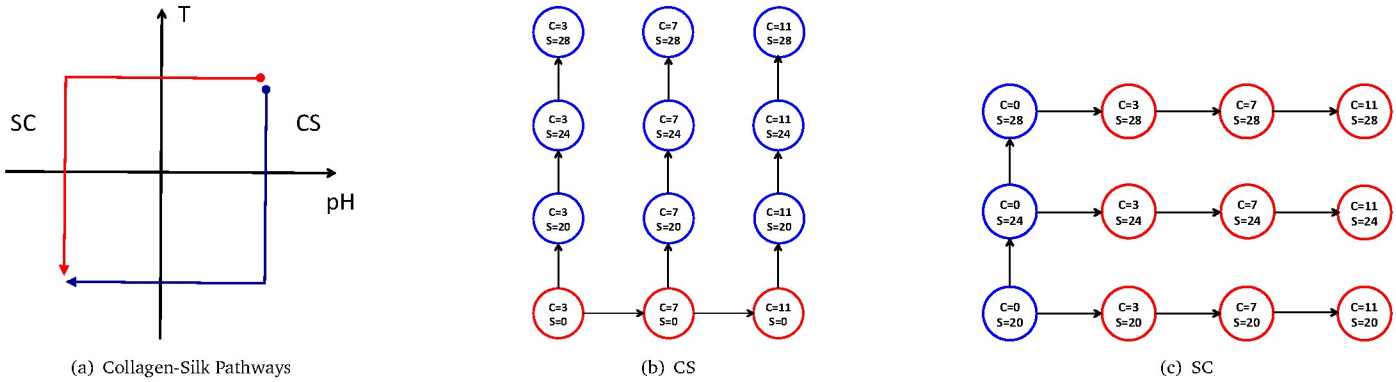


Fig. 3 Topology of resulting network structures from different triggering sequences. Left: collagen (red circles). Middle: elastin (red circles) Right: silk (blue circles). The orange lines between blue and red indicate connectors, while the lines between circles of same color indicate bonds.



(a) Collagen-Silk Pathways (b) CS (c) SC

Fig. 4 Left: the two major triggering pathways for Collagen-Silk TP: *SC*: starting at high pH and T, pH reduction is followed by temperature decrease. *CS*: temperature decrease is followed by pH reduction. Middle: Simulation protocol for the *CS* pathway, Right: simulation protocol for the *SC* route.

2.3 Simulation methodology

The simulations are performed in two stages. In first stage, either the collagen/elastin or silk sticker is activated (triggered), and in the second stage both the collagen/elastin and silk stickers are activated. In each stage, the system evolves according to a hybrid MD/MC scheme.[2,22] This hybrid MD/MC scheme performs regular molecular dynamics in the NVT ensemble using the velocity-Verlet integration scheme[25] and the Lowe-Andersen thermostat.[26] The MD run is interrupted periodically, to perform a set of MC moves that attempt to bind stickers or unbind already bound stickers, by switching the bond matrix entry C_{ij} from zero to 1 or vice versa.[2] These MC moves conserve momentum as well as total energy. After the set of MC moves are completed the MD run continues. The stickers which are connected during the MC phase start interacting with U_C during the subsequent MD runs. They continue to interact until the connection is broken in an subsequent MC phase. A detailed description can be found in Appendix A.

The strength of attraction (ε_{attr}, see Eq. 3) between the sticker units is varied. For collagen and elastin stickers ε_{attr} is varied from 3 to 11 k_BT and for silk stickers ε_{attr} is varied from 20 to 28 k_BT. These values are chosen such that the assembly of structure occurs in this range. Note that the jump in silk interaction from zero (before triggering) to 20 k_BT is a larger quench than for collagen or elastin. This is in line with the fact that in experiments the collagen/elastin effective interaction strength is influenced by temperature (gradually), whereas for silk it is primarily determined by changing the pH (instantaneously). The sticker attraction strengths will be denoted by C, E and S for collagen, elastin and silk stickers respectively in the next sections.

An MD simulation time step of $\delta t = 0.1\tau$ is used for both the equilibration and production runs. Here the MD time unit $\tau = (mR_g^2/k_BT)^{1/2}$ is determined by the mass and the size of the polymer beads. Assuming a single polymer bead to contain 100 residues, the R_g is around 3 nm[27] and it weighs around 11 kDa ($\sim 1.8 \times 10^{-20}$grams), we obtain $\tau \sim 0.2$ ns at room temperature. We note that this time scale is not representing the experimental timescale due to the high coarse-graining, and the lack of solvent in the model. The diffusion constant for a polymer in solution is much smaller than the effective diffusion we simulate with the hybrid MD/MC simulation. Therefore, one MD time step in our simple model corresponds to much longer times in a corresponding experimental set up.

The equilibration run is performed for 3 million steps and pro-

duction run for 1 million steps. The production run is considered for computing the averages. The thermostat is activated every 10^{th} time-step with the collision frequency set to 2.5 τ^{-1}. The MD run is paused every 100^{th} time-step and 20 thousand MC moves are attempted for binding or unbinding of stickers. The acceptance ratio of these moves is usually very small, because most of the sticker pairs are more distant than the cutoff range, thus rejected. The acceptance varies from 0 to 0.1%. Of course, the acceptance would be higher when choosing only neighboring pairs, but then obeying detailed balance is more involved.

For each blob volume fraction, a separate set of simulations is performed. In all the simulations, we start with the 200 chains distributed randomly (Additionally, we perform also a few simulations of larger systems, $N_c = 400$ and $N_c = 600$, to investigate the effect of the system size). First, an initial set of simulations is performed with only either the collagen/elastin or the silk stickers activated. The resulting final conformation is then used as the starting configuration for the series of simulations with increasing sticker attraction strength. The increasing of the sticker attraction strength mimics lowering of pH (silk) or temperature (collagen), or raising the temperature (elastin). Subsequently, a set of simulations is performed with both the collagen/elastin and silk stickers activated, with the final structures resulting from the first set of simulations as the starting configuration. Thus, we perform two major sets of simulations resembling two different pathways indicated in fig. 4(a). In the *CS* pathway (fig. 4(b)) the collagen sticker is triggered first, followed by triggering of the silk sticker. In the *SC* (fig. 4(c)) pathway, the collagen is triggered after the silk stickers. A similar methodology is followed for the elastin-silk system leading to *ES* and *SE* pathways. Each triggering sequence was repeated multiple (15-25) times to obtain statistics.

As pointed out above, triggering only one of the stickers leads to small clusters for collagen/elastin, or filaments for silk. On triggering both the silk and collagen stickers, these small clusters can join to form polymer networks in the shape of larger clusters and possibly even a system spanning polymer network. These polymer networks can be characterized by the size of clusters and nodes. Cluster size refers to the number of polymer chains in a joint polymer network and node size refers to number of stickers connected together in each collagen/elastin or silk node. The cluster sizes are determined using a clustering algorithm.[28]

3 Results

3.1 Self-assembly process

We performed extensive simulations of the two pathways indicated in Fig. 4(a) for both the silk-collagen and the silk-elastin systems. In all cases we found that triggering the stickers leads to self-assembly and cluster formation. Figure 5 shows a typical evolution of an MD equilibration and production run. These particular figures are for *CS* with strength $C = 5$ and $S = 20$. At the start the collagen interaction is already triggered, and has formed triple helices, but not completely, as the average node size is about 2. After triggering, the equilibration run shows a quick increase of the mean silk node size as well as the mean polymer network cluster size, until it settles at around $N_{silk} \approx 10$ and $N_{clust} \approx 20$.

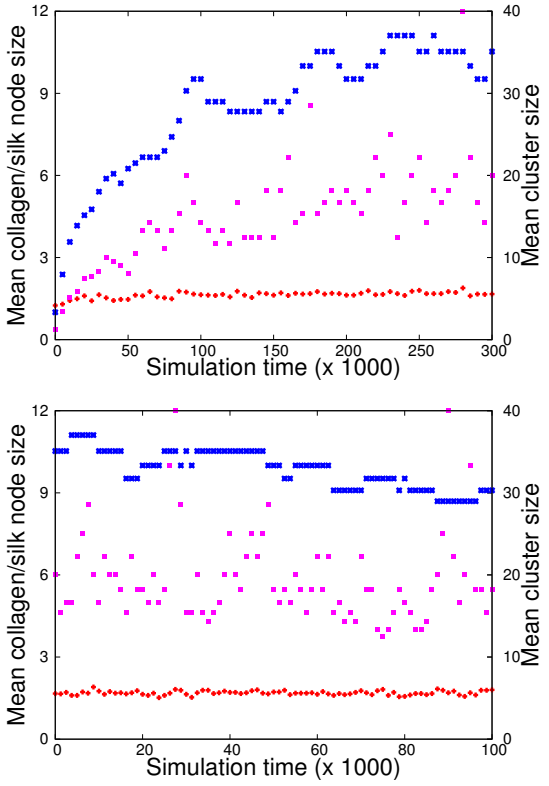


Fig. 5 A typical evolution of an equilibration run (top) and production run (bottom) [red/blue: mean collagen/silk node size, pink: mean cluster size]. These runs were taken from a CS pathway for $N_b = 3$ at $\phi_b = 0.05$ with $C = 5, S = 20$.

The production run shows fluctuations in these cluster sizes, indicating that there are still clusters being formed and broken.

The final morphology depends on the initial overall polymer density. At low density, we find a number of separate finite-sized clusters floating around in the (implicit) solvent, whereas at high density we observe a single large polymer network. A typical graph representation of both situations is shown in fig. 6. In the next subsections, we present results for the two systems, collagen-silk and elastin-silk.

3.2 Collagen-Silk

One of the most direct ways to characterize the morphology of the system is the average polymer network cluster size and its distribution. For the two triggering sequences, Fig. 7 lists the average cluster sizes for a blob volume fraction $\phi_b = 0.05$ for short ($N_b = 3$) and long ($N_b = 9$) chain systems respectively. These averages were computed over the last 20% of the production trajectories. The standard deviation of all averages over the production trajectories (error in the mean) is given as a subscript. This standard deviation can be quite large, especially at large cluster sizes, due to the small number of clusters involved. Clearly, triggering along the pathways increases the cluster size, as expected. However, surprisingly for the smaller chain length ($N_b = 3$), the different pathways result in different system configurations. In Fig. 8 snapshots of the identical end situation $C = 11, S = 28$ are

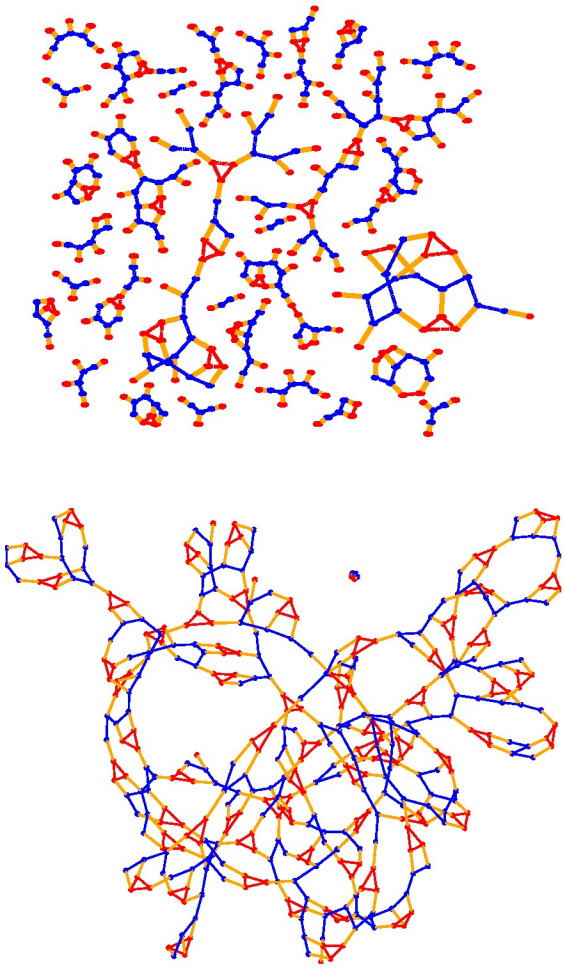


Fig. 6 Top: a graph representation of a solution of finite sized clusters for $N_b = 3$ at $\phi_b = 0.01$. Collagen (red circles), silk (blue circles), connector (orange solid lines), intermolecular silk-silk connection (blue dotted lines), collagen-collagen interaction (red dotted lines). Bottom: a graph of one large cluster obtained for $N_b = 3$ at $\phi_b = 0.05$.

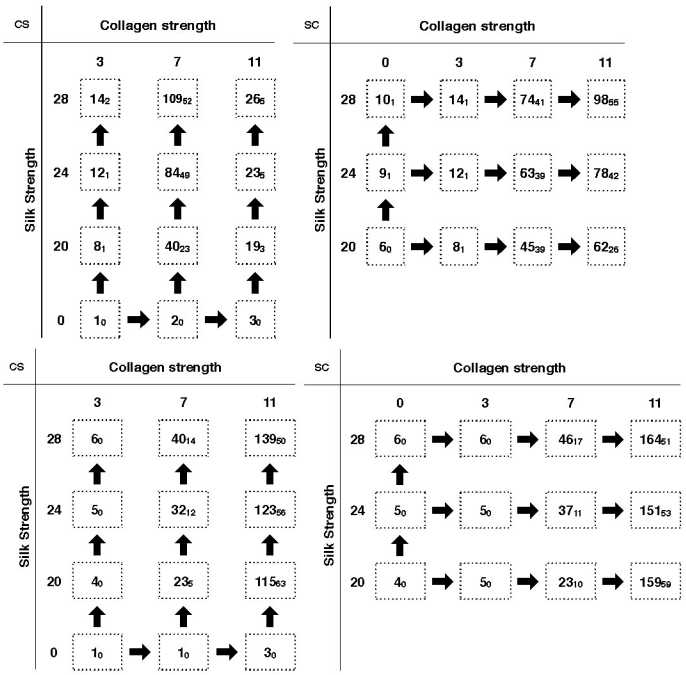


Fig. 7 Average cluster sizes for the collagen-silk TP systems, for $N_b = 3$ (top) and $N_b = 9$ (bottom) for $\phi_b = 0.05$. The left and right panels show the average cluster size for CS and SC trigger sequences respectively. The arrows denote the order in which the simulations have been performed, using the output of one simulation as input for the next. The subscript numbers indicates the error in the mean of the last digit(s).

compared for the two pathways *CS* and *SC*. Clearly, there are large differences. For the *CS* pathway (top row), one ends with a distribution of smaller finite size clusters, whereas for the *SC* route (bottom row) we observe a single cluster. This behavior is also visible in the cluster size table of Fig. 7. For the long chain this non-equilibrium behavior does not seem to occur, and cluster sizes are relatively more independent of the pathway (see Fig. 7).

Another peculiar observation for the short connector system ($N_b = 3$) is that in the pathway in which collagen is triggered first (*CS*), applying intermediate collagen attraction strength ($7k_BT$) shows more aggregation than higher attraction strength ($11k_BT$). Intermediate collagen strength also leads to higher errors in the mean, and also larger fluctuations. This behavior is replicated across all the blob volume fractions (figure S1 in the Supplmentary Information, SI). Again, the longer chain system $N_b = 9$ does not show such behavior. A possible explanation for this finding will be discussed in section 4. To check whether these observations are dependent on the used network metric, i.e. the cluster sizes, we present a different way of analyzing the network in Fig. S9 in the SI. Here we counted the number of bridging collagens between two silk nodes. The results are qualitatively the same, showing more bridging collagens in the SC sequence for $N_b = 3$, while for $N_b = 9$ this difference disappears.

Fig. 9 summarizes the average cluster size and the average node size for collagen and silk for the $N_b = 3$ and $N_b = 9$ system at a blob volume fraction $\phi_b = 0.05$. The x and y axes of each plot refer to collagen and silk sticker attraction strengths, respectively. The color indicates the value for each metric and can be read from the color bar next to each plot. In the lower panels, the cluster size distribution for the highest silk sticker attraction strengths are depicted (labeled by letters in the top row size plots). These cluster size distributions show that for increasing collagen interaction strength, in both trigger sequences, the cluster size distribution can jump from small clusters at lower interaction strength to large ($N_C = 200$) aggregates at high interaction. This resembles a (phase) transition in which the enthalpy of aggregation becomes more dominant than the entropy of the distributed clusters. To check whether this behavior depends on system size, we performed runs for N_c = 400 and 600. The results, shown in Fig. S2 of the SI, are very similar to the N_c = 200 case. Note that the cluster size distribution can become bimodal (as seen in the lower panel of figures 9 and S2), contributing to the large standard deviations in the average cluster size reported in fig. 7.

From Fig. 9 it is also clear that for the pathway in which collagen is triggered first (*CS*), intermediate collagen attraction strength ($7k_BT$) causes significantly larger silk node sizes than

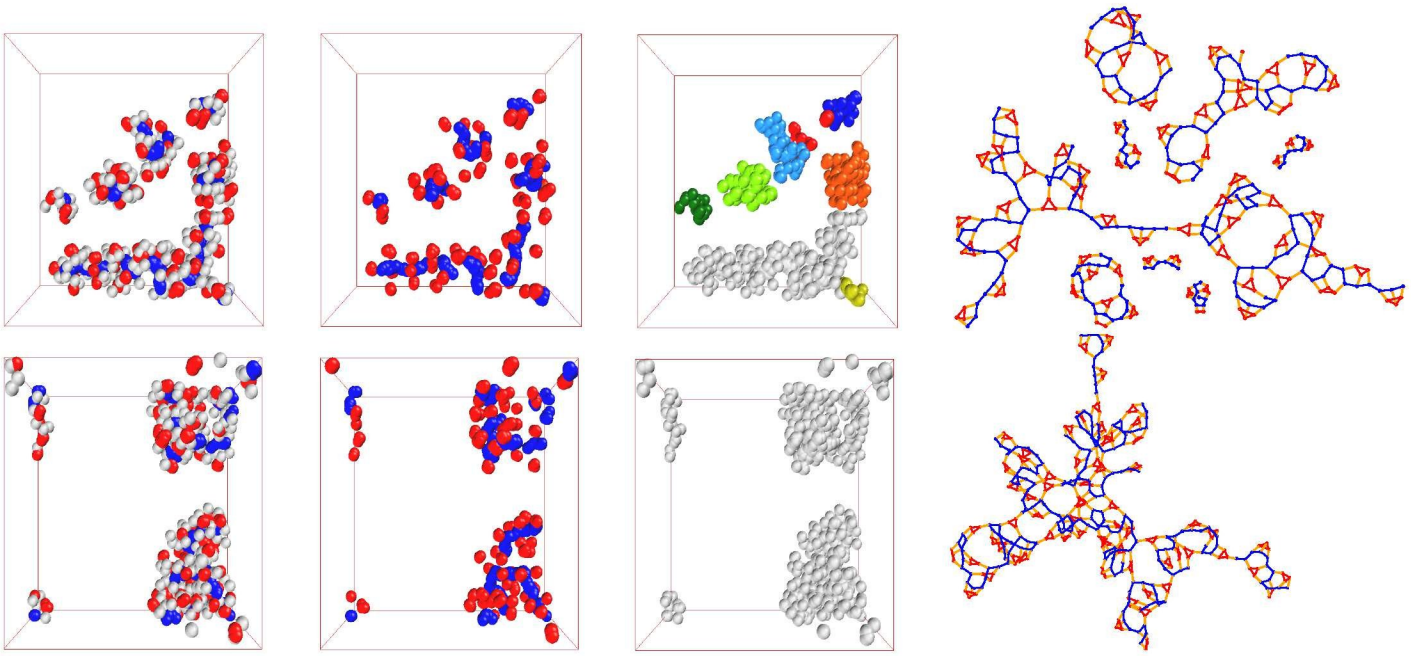


Fig. 8 Snapshots of the collagen-silk system,$N_b = 3$, $\phi_b = 0.05$, at the strongest interaction strengths $C = 11, S = 28$. Top row: the collagen-first trigger sequence (CS). From left to right: the entire system with collagen (red beads), silk (blue beads) and connector (grey beads), the system showing the stickers only, showing cluster identity by different colors, and the connectivity graph (colors as in figure 6). Bottom row: Snapshots of full system, stickers only, cluster identity, and connectivity graph for the silk-first trigger sequence SC).

the higher attraction strength ($11k_BT$). Indeed, for this high collagen attraction strength, the cluster distribution is not bimodal, indicating that larger (system wide) aggregates are not formed.

In Fig. S3 of the SI, we plot the average cluster size (distributions), the average node size for collagen and silk for the $N_b = 3$ system for several blob volume fractions. From these figures it follows that for low density there are only small clusters of finite size. When increasing the blob volume fraction, the cluster size distribution can suddenly jump from small clusters, towards large aggregates of TPs as in Fig. 9. This is akin to a (percolation) transition often observed in aggregating systems.[29] The observation that at intermediate collagen attraction strength ($7k_BT$) significantly larger silk node sizes form than for higher attraction strength ($11k_BT$), is replicated across all the blob volume fractions.

In Fig. S4 of the SI, the cluster and node averages, and the cluster size distribution is shown for the TP with the long connector ($N_b = 9$) for several blob volume fractions. For low blob volume fractions only small clusters form, even when fully turning on the attraction strength. When increasing the blob volume fraction, the cluster size distribution can suddenly jump from small clusters, towards large aggregates of TPs. The cluster size distribution is more markedly bimodal. Unlike the short chain polymer case (Fig. S3 of SI), the different pathways do not result in significantly different system configurations for long chain length polymer system. A single large cluster is formed for the higher collagen attraction strengths for the blob volume fraction ϕ_b of 0.05 and 0.15.

3.3 Elastin-Silk

Figures 10 and S5 of the SI summarize the simulation results for the elastin-silk system with short chain length ($N_b = 3$) and long chain length ($N_b = 9$) for blob volume fraction $\phi_b = 0.01$ and $\phi_b = 0.05$, respectively. At the low blob volume fraction of 0.01, neither pathway leads to aggregated structures. Like the collagen-silk system, the two pathways result in different final structures only for short chain length system. The pathway in which silk is triggered first (SE) leads to larger aggregates at high elastin and silk attraction strengths, as can be seen from both the average node and cluster sizes. This behavior is illustrated in the snapshots of Fig. 11. In contrast, while we observe an enhancement of the cluster size at low elastin attraction, the ES path does not show the strong non-monotonic behavior that was found for the CS path.

For all the trigger sequence simulations, the cluster size distributions moves towards larger clusters, except when elastin is triggered first (ES) in the short chain system. The transition is much more stark for the long chain length system. Similar to the collagen-silk system, a single large cluster is formed for the higher elastin attraction strengths for the long chain length system.

4 Discussion

The hybrid MD simulations using the CG model leads to either small aggregates or large system-wide clusters. Aggregation takes place after triggering a single sticker, which leads to formation of star polymers with three arms (collagen), micellar structure (elastin) and fibril-like polymers (silk). The equilibrium distribution of these clusters depends, besides on the molecular model, on the overall concentration of the system and the temperature. As-

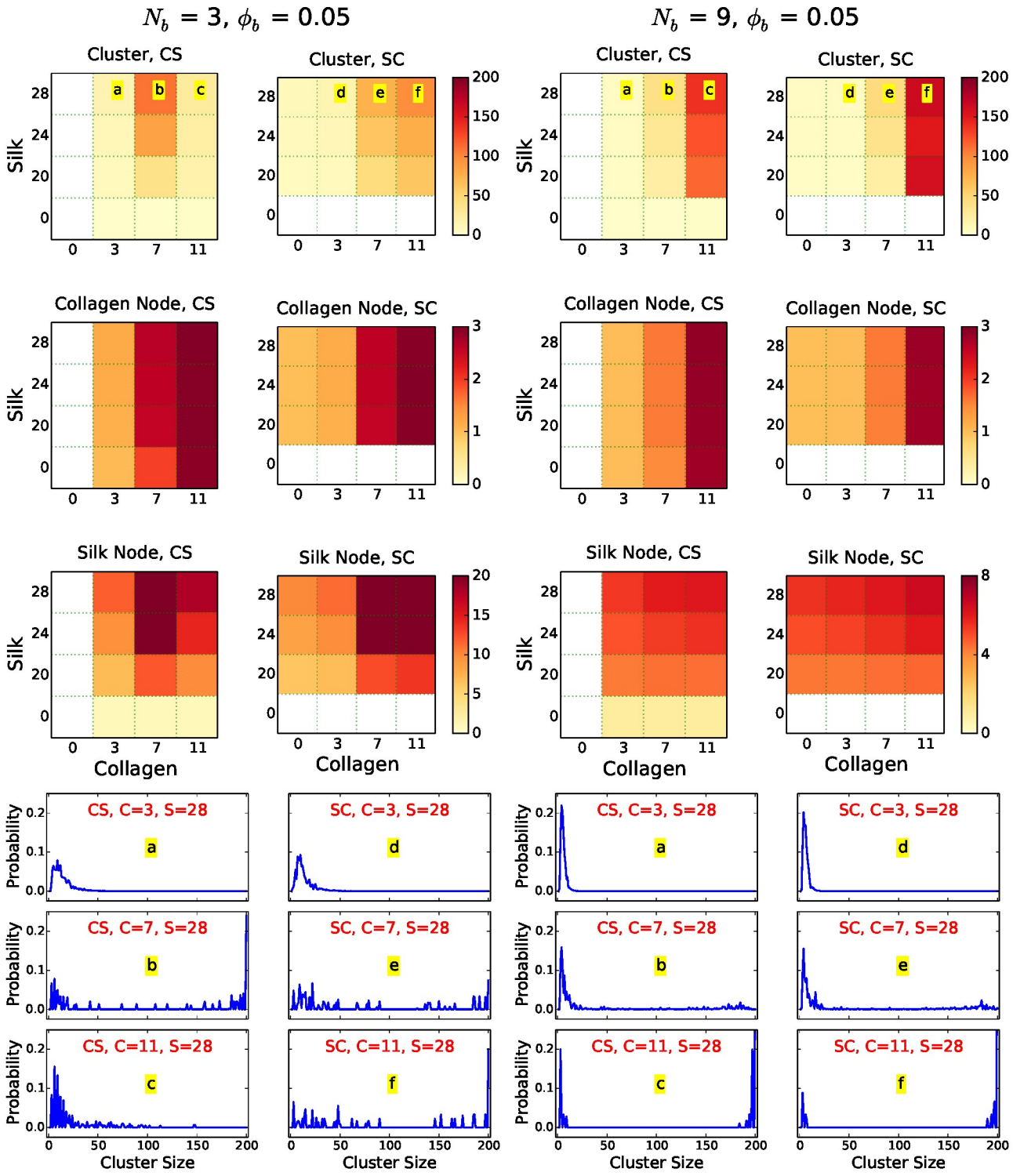


Fig. 9 Top: Average cluster and node sizes for CS and SC routes for collagen-silk TP for $N_b = 3$ (left) and $N_b = 9$ (right) for a blob volume fraction $\phi_b = 0.05$. Each color-map indicates the average size as a function of collagen and silk strength. Bottom: cluster size distributions corresponding to settings indicated by labels a-f in top row.

suming a simple aggregation/polymerisation/micellisation model the cluster size distribution is either exponential (for linear polymers) or bimodal (for micelles).[30] The bimodal distribution for the latter case arises due to the presence of monomers along with the aggregated micelles.

A stronger sticker attraction and a higher density lead to larger

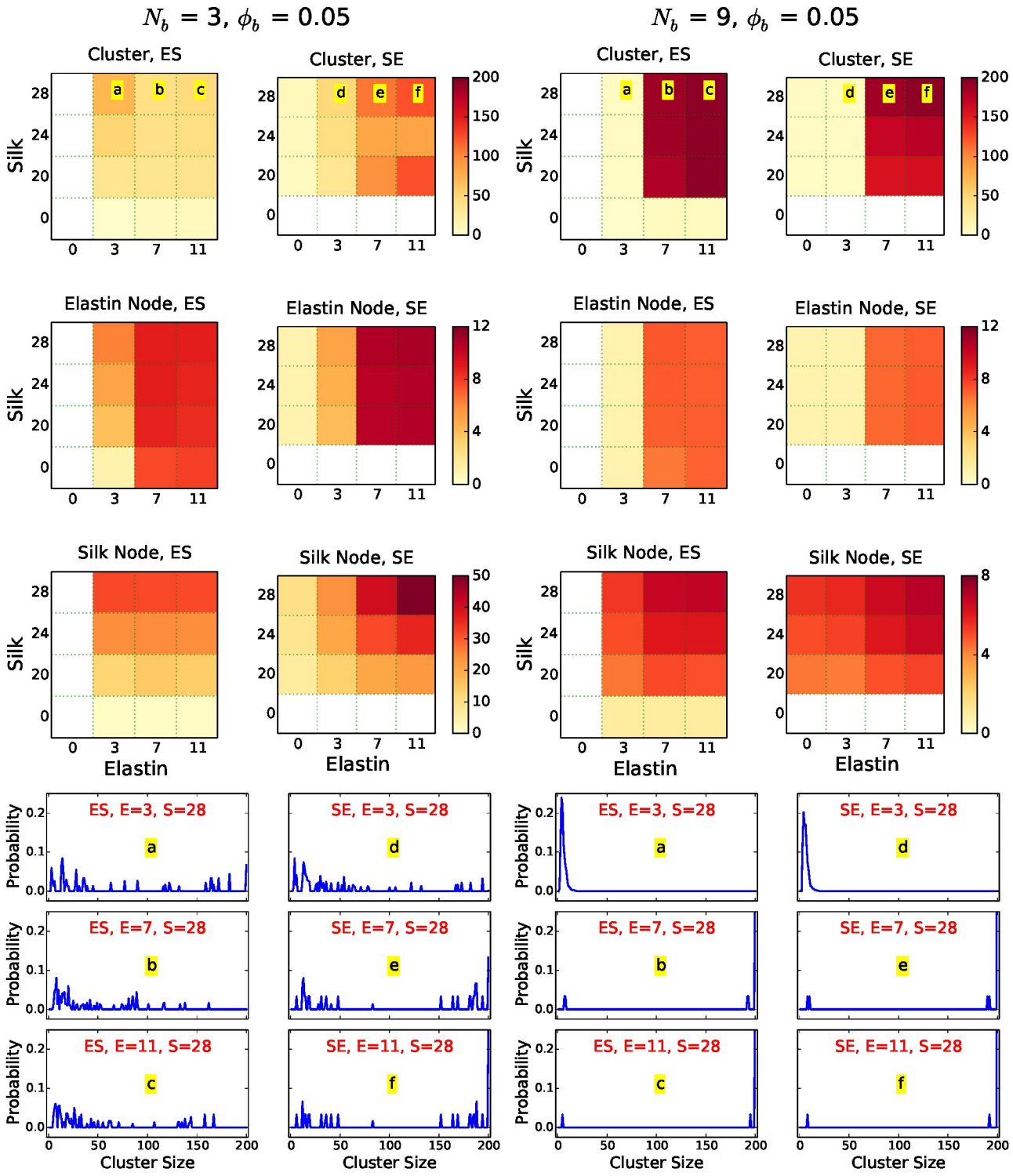


Fig. 10 Top: Average cluster and node sizes for ES and SE routes for elastin-silk TP for $N_b = 3$ (left) and $N_b = 9$ (right) for a blob volume fraction $\phi_b = 0.05$. Each color-map indicates the average size as a function of elastin and silk strength. Bottom: cluster size distributions corresponding to settings indicated by labels a-f in top row.

average cluster sizes. For collagen and elastin the cluster size (node size) is limited by the molecular details. The aggregation fraction as a function of attraction strength for this type of aggregation usually shows a sigmoidal behavior as is clearly demonstrated in Fig. 12. Size distributions for silk-only aggregation, shown in Fig. S6 of the SI, exhibit a maximum beyond which they

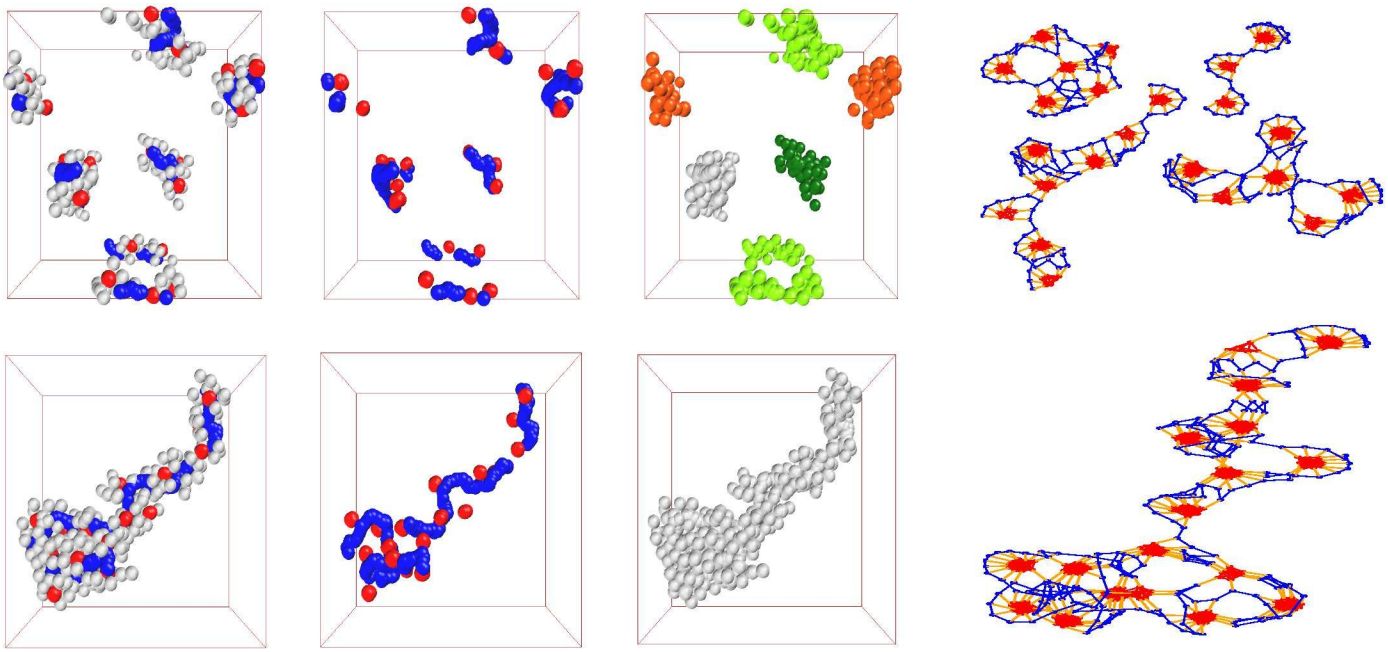


Fig. 11 Snapshots of the elastin-silk system, $N_b = 3$, $\phi_b = 0.05$, at the strongest interaction strengths $E = 11, S = 28$. Top row, left to right: Snapshots of full system, stickers only, cluster identity, and connectivity graph for the ES trigger sequence . Bottom row: Snapshots of full system, stickers only, cluster identity, and connectivity graph for the SE trigger sequence. Colors as in figure 8.

decay roughly exponentially, as would be expected for an association reaction. The initial increase in these distributions and the presence of a maximum suggests micelle-like behavior in which a critical number of molecules is required to make a stable cluster. Note that the applied density is far above the critical aggregation concentration, and we hardly observe monomers.

When the other sticker is also triggered, the already formed aggregates fuse further to form an interconnected network. Since the previously triggered sticker remains activated, dissociation of earlier formed aggregates is not likely to occur. Hence, the initial configuration at the moment of triggering of the second sticker depends on the triggering sequence. For example, for high collagen and elastin attraction energies, the system comprises of trimers or micelles with twelve stickers, making the system roughly monodisperse. On the other hand, the silk stickers aggregate into filaments of different sizes which leads to a polydisperse starting configuration (see Fig. S8 of the SI). Also, the silk stickers in the trimer (collagen) and 12-mer (elastin) are slightly closer to each other with respect to the collagen/elastin stickers in silk filaments (see Fig. S7 of the SI). For instance, for $N_b = 3$, the average SS distance in a collagen trimer is 5.3 while the average CC distance in a silk filament is 6.7. These (small) differences in the starting configurations can lead to different final network structures depending on the triggering sequence.

From an equilibrium viewpoint, when the density is high one obtains a single large cluster, while for low density, a solution of individual clusters is stabilized by entropy. Here we observe non-equilibrium behavior because even if the equilibrium state would favor one large cluster, individual clusters with chains looping back to themselves cannot easily fuse with each other due to the fulfillment of the bonds internally in the cluster. Once a small cluster is formed which has all the collagen/elastin bonds fulfilled, the formation of large clusters via merging of such smaller clusters is hampered due to the exchange of stickers among individual clusters being a low probability process. Only the silk fibrils are able to fuse indefinitely for high enough interaction strengths. However, the reactive ends of these fibrils can be hard to reach because they are surrounded by a buffer of polymer chains consisting of connector and collagen/elastin blocks. These factors render merging of the clusters into a single large cluster difficult, even when equilibrium behavior would dictate so.

The connector length also plays a role here as we find that strong non-equilibration behavior is observed in the chains with short connectors, leading to different final network structures for the different triggering sequences. We speculate that this behavior is caused by the fact that short connectors are more restricted than the longer connectors. A collagen/silk sticker that accidentally breaks loose is then also more likely to rebind. For longer connectors, the free sticker ends can explore more space, which in turn leads to a faster equilibration. On the other hand, for the longer connectors we observe smaller silk node sizes (see Fig.9 and Fig. S8 of the SI), supporting the presence of the above-mentioned buffer effect.

For the collagen-silk system, the chains with shorter connectors are able to form a network much more easily for partially aggregated collagen stickers (intermediate attraction strength) than with the highly aggregated collagen stickers (large attraction strength). This effect relies on the ability of stickers to bind and unbind depending on the extent of aggregation. Since for low attraction strength binding hardly takes place, and for high attraction strength unbinding is unlikely, only at intermediate strength (C=7) sufficient unbinding and binding occurs to allow cluster fu-

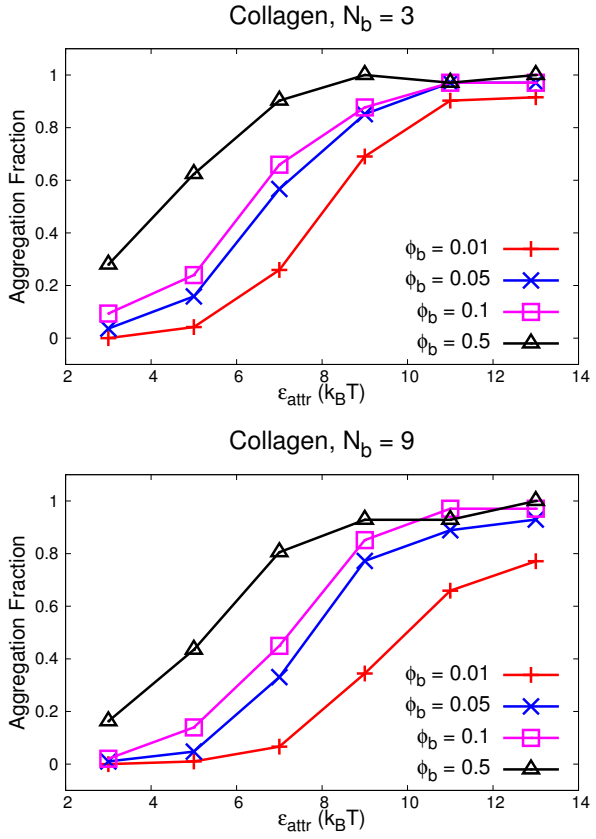


Fig. 12 Collagen aggregation curves for $N_b = 3$ and $N_b = 9$ for different blob volume fractions ϕ_b.

sion. This larger fluctuation in bonding energy follows from the aggregation curves in Fig.12, where the derivative of the aggregation fraction is proportional to the bond energy fluctuations, and hence the number of bonds being formed or broken. Such strong fluctuations enable the system to respond faster to changing circumstances. When the conditions would ~~have~~ favor extensive aggregation, the large fluctuations allow the stickers to switch their initial binding partner for another one, and form a network. These fluctuations correspond to a faster timescale (see the Appendix B for an alternative kinetic argument). It is thus for the intermediate attraction strength of the collagen sticker that we find large network sizes. Another way of looking at this effect is to ask why at high collagen strength (C=11) in the CS trigger sequence cluster growth is limited. Our results suggest that strong collagen binding does not allow rearrangement of collagen bonds, and the subsequent silk trigger quenches the system in a set of smaller clusters, which are largely internally bound. This effect is also visible in the bridging collagen data (see Fig. S9 in the SI). The physical reason for this quench is that while collagen is triggered in experiments by temperature, which usually is changed gradually, silk is triggered by pH, which changes almost instantaneously, and immediately alters the attraction of the molecules. The silk binding is thus not showing large fluctuations. Instead, further growth occurs via fusion of silk ends. Coincidentally, for the longer connectors we observe smaller silk node sizes at intermediate collagen strength (see Fig.9), providing further evidence to the above-mentioned buffer effect. The fluctuations also explain the hysteresis in the triggering sequence for the collagen-silk system, since for the SC trigger sequence, the intermediate collagen regime is reached only after the silk has already formed long clusters, leading to additional fusion of silk clusters by collagen.

A comparison of the behavior for the two chain lengths should be performed at the same densities. However, as shown in Table 1 one can compare either the polymer (bead) density or the sticker density. From Figs. 9, S3 and S4 we conclude that the same blob volume fraction, e.g. of 0.05 lead to roughly similar equilibrium cluster size distribution after the trigger sequence (although of course the $N_b = 3$ case shows non-equilibrium behavior). For both cases we observe a bimodal distribution. In contrast, when we compare equal sticker density e.g. $\phi_b = 0.05$, for $N_b = 9$ and $\phi_b = 0.015$, for $N_b = 3$, we observed very different cluster distributions. Even the average node sizes are different despite the sticker densities being identical. This shows that the topology of entire polymer chain is relevant for the cluster aggregation instead of only the stickers.

Comparing the two TP systems considered in this work, namely collagen-silk and elastin-silk, we notice that unlike the $N_b = 3$ collagen-silk system for blob volume fraction $\phi_b = 0.05$, no enhanced aggregation into large clusters is observed for intermediate elastin sticker attraction strengths. This can be explained by the fact that elastin aggregation shows a much sharper transition, at lower interaction strength, leaving less room for fluctuations. Nevertheless, the elastin is only able to aggregate into the maximum node size of twelve for the case when silk is triggered first, indicating that the silk filament formation assists in further growth of elastin micelle. The elastin-silk system behaves similar to the collagen-silk system for the chain with longer connector length. The silk node size is considerably larger for the elastin-silk system for high attraction strengths (∼50) as compared to collagen-silk systems (∼20).

Experiments by Pham *et al.*[12] were performed at roughly the same density as the simulations (a density of 5 grams/liter corresponds roughly to $\phi_b \approx 0.03^*$). Our results agree qualitatively with the experimental finding that the trigger sequence leads to hysteresis, and loss of reversibility. However, our coarse grained simulations are not accurate enough to reproduce structural and molecular geometry observed. For instance, in Ref.[11] the distinction is made between networks that are aggregates of micelles versus networks that are bundles of fibers. For the SE trigger, one might expect bundles of fibres, and for the ES triggering route one might expect aggregates of micelles. While our model does produce elastin micelles, the silk filaments are much shorter, more floppy than is experimentally observed. Indeed, there are several shortcomings of the model: no folding transition, the interaction is isotropic, there is no stiffing transition upon aggregation of silk

∗ This estimate is based on an experimental telechelic polymer length of around 700 residues, which translates to a weight of 77 kDa, or ∼ 1.3e-19 grams per molecule with one CG bead comprising of roughly 100 residues, leading to an $R_g = 3nm$.[27]

and so on. Better models might be employed to alleviate this problem in future work.

5 Conclusion

We have investigated the self-assembling behavior of biomimetic telechelic (sticker-connector-sticker) polymers using a simple CG model. We found that different sticker activation pathways lead to different aggregation behavior only for relatively short connectors. Upon unbinding of a sticker, long connectors can explore larger space so that the free sticker can more easily establish a bond with a new sticker; this allows a system with long connectors to relax more extensively on the time scale of the simulation than systems with short connectors. As a result, the systems assembled from long chains can equilibrate better and hence do not show any significant influence of sticker activation pathways. Longer chains are able to form large system-spanning clusters when the sticker activation energies are high. Shorter chains show bimodal cluster size distributions and are not able to form a system-spanning structure.

In both collagen-silk and elastin-silk systems with short connectors, activating the silk sticker first (*SC* and *SE*) leads to more aggregated structures. However, for the collagen-silk system, activating collagen first (*CS*) also leads to increased aggregation, surprisingly, at moderate sticker attraction strengths. Fluctuations allow, partially aggregated collagen stickers to form the network more easily than the highly aggregated collagen stickers.

Our results are in qualitative agreement with the experimental finding that when two blocks are sequentially triggered, reversibility is lost, and the final morphology and mechanical properties of the aggregates depend on the pathway chosen.[12] However, as we use a extremely course grained model, a direct comparison with experiment is at this stage premature.

These findings indicate that morphologies of asymmetric telechelic polymers can be influenced and steered experimentally by making use of the non-equilibrium effects occurring in the different self-assembly pathways.

Acknowledgments

This work is funded by the Industrial Partnership Program (IPP) Bio-related materials (8BRM01-2) of the Foundation for Fundamental Research on Matter (FOM), which is part of the Netherlands Organization for Scientific Research (NWO).

A MC scheme

This appendix details the hybrid MD/Monte Carlo scheme used. The MD integration is paused at regular intervals and several Monte Carlo steps are executed to the existing sticker connectivity network. This mimics the spontaneous binding and unbinding of the stickers. These MC steps are required to obtain control over the sticker functionality (e.g. the maximum number of stickers attached to one node), which is not easily achieved in MD. The MC scheme is as follows:

1. Select at random two stickers i and j of the same type.

2. Calculate the distance r_{ij} between the stickers. Reject the move if the distance is larger than the cut-off distance R_{cut} (see section 2.1 for a explanation of R_{cut}).

3. If the stickers are already bonded ($C_{ij} = 1$ in Eq. 4), an attempt is made to break the bond by setting $C_{ij} = 0$. The MC move is accepted or rejected depending on the kinetic energy available to the stickers.

4. If the stickers are not bonded ($C_{ij} = 0$ in Eq. 4), an attempt is made to create the bond by setting $C_{ij} = 1$. The MC move is accepted or rejected depending on the node constraints (collagen implies a maximum of three stickers in a node, elastin has a maximum of twelve, silk has maximum of two neighbors only, and no loops are allowed). The energy that is released at bond formation is translated into kinetic energy of the stickers.

The MD stage in the hybrid MD/MC scheme occurs in the canonical NVT ensemble. During the MC stage, bond creation or breaking takes place, in which (large) amounts of energy is released or absorbed. To ensure that the MC algorithm obeys detailed balance, it is necessary that we use a scheme that conserves energy and momentum. To account for these conservation laws we slow down or speed up the stickers.

For the bond creation move between stickers labeled 1 and 2, the initial kinetic energy has to be equal to the final kinetic energy plus the bond energy, $K_1 + K_2 = E_B + K_1^{'} + K_2^{'}$. In contrast, for the bond break move, we have $K_1 + K_2 + E_B = K_1^{'} + K_2^{'}$. Here K_1 and K_2 denote the kinetic energy of the two chosen stickers prior to MC move, $K_1^{'}$ and $K_2^{'}$ denote the kinetic energy of the two chosen stickers after the MC move and $E_B = U_C(r_{12})$ denotes the bond energy absorbed during breaking of the bond or released during the creation of a bond. The kinetic energy is defined in the usual way as $K = \frac{1}{2}mv^2$.

To conserve the total momentum, the velocities of the stickers are perturbed by equal amounts in opposite direction along the displacement vector, i.e. for sticker 1: $\vec{v}_1^{'} = \vec{v}_1 + \alpha\vec{r}_{12}$ and sticker 2: $\vec{v}_2^{'} = \vec{v}_2 - \alpha\vec{r}_{12}$, where $\vec{r}_{12} = \vec{r}_1 - \vec{r}_2$, $\vec{r}_1$ and $\vec{r}_2$ being the positions of stickers 1 and 2 respectively. $\vec{v}_1$ and $\vec{v}_2$ are the initial velocities of stickers 1 and 2 respectively, whereas $\vec{v}_1^{'}$ and $\vec{v}_2^{'}$ are the final velocities of stickers 1 and 2 respectively and α denotes the perturbation strength.

We first check whether breaking or creation of the bonds will lead to the sum of kinetic energies of the stickers to be negative. If this occurs, the move is rejected. Otherwise substituting the velocities, we obtain a quadratic equation in α. For the bond creation this is

$$\vec{v}_1^2 + \vec{v}_2^2 = (\vec{v}_1 + \alpha\vec{r}_{12})^2 + (\vec{v}_2 - \alpha\vec{r}_{12})^2 + E_B$$

If the equation has real roots, the MC move to create or break a bond is accepted, or rejected otherwise. The α with lower magnitude is considered.

Once two unbound stickers bind in a given MC phase, they remain bound until they become separated in another MC phase of the simulation. During the subsequent MD phases, $C_{ij} = 1$, and the bound stickers interact under an additional potential U_C. Thus, the stickers experience U_A and U_B during all the MD phases

and U_C when they are bound, as given by Eq. 4. Since U_C becomes repulsive at a distance larger than R_{attr}, we reject any move to bind stickers at a distance more than R_{cut} (note, $R_{cut} > R_{attr}$). To ensure detailed balance, we also do not break any bound stickers at a distance more than R_{cut}. It is however possible to form and break bonds with a positive U_C value, although it is less likely.

This hybrid MC/MD scheme conserves energy and total momentum, obeys detailed balance and is completely reversible. Further coupling to a stochastic heat bath using the Lowe-Andersen thermostat ensures canonical behavior.

B Kinetic argument for effect of intermediate binding strength

An alternative way to look at the effect of the intermediate binding strength is to consider the probability for an escape from a bound state followed by another binding event that then survives for a certain (sufficiently long) time. The first part of this conditional probability, the probability to react in a time t_0 is proportional to $1-\exp(-t_0/\tau)$, while the second part is the survival probability, the complement of the reaction probability $\exp(-t_0/\tau)$. Clearly, this conditional probability will be largest for intermediate relaxation time τ. The relaxation time is given by $\tau = 1/(V+\exp[-\Delta G/k_BT]$, and thus depends on the binding strength ΔG, or equivalently the temperature. The resulting probability $p(\Delta G) = (1-\exp(-t_0/\tau))\exp(-t_0/\tau)$ will thus show a peak as a function of the binding strength ΔG. This rather crude argument can be made more precise, but this is beyond the scope of this work.

References

1 F. L. Verso and C. N. Likos, *Polymer*, 2008, **49**, 1425 – 1434.

2 R. S. Hoy and G. H. Fredrickson, *The Journal of Chemical Physics*, 2009, **131**, 224902.

3 P. G. Khalatur, A. R. Khokhlov, J. N. Kovalenko and D. A. Mologin, *The Journal of Chemical Physics*, 1999, **110**, 6039–6049.

4 T. Pham, *PhD thesis*, Wageningen University, 2013.

5 M. Sutter, J. Siepmann, W. E. Hennink and W. Jiskoot, *Journal of Controlled Release*, 2007, **119**, 301 – 312.

6 J. G. Hardy, J. Y. Lee and C. E. Schmidt, *Current Opinion in Biotechnology*, 2013, **24**, 847 – 854.

7 W. Qiu, W. Teng, J. Cappello and X. Wu, *Biomacromolecules*, 2009, **10**, 602–608.

8 X.-X. Xia, Q. Xu, X. Hu, G. Qin and D. L. Kaplan, *Biomacromolecules*, 2011, **12**, 3844–3850.

9 W. Hwang, B.-H. Kim, R. Dandu, J. Cappello, H. Ghandehari and J. Seog, *Langmuir*, 2009, **25**, 12682–12686.

10 A. A. Dinerman, J. Cappello, H. Ghandehari and S. W. Hoag, *Biomaterials*, 2002, **23**, 4203 – 4210.

11 M. D. Golinska, T. T. H. Pham, M. W. T. Werten, F. A. de Wolf, M. A. Cohen Stuart and J. van der Gucht, *Biomacromolecules*, 2013, **14**, 48–55.

12 T. T. H. Pham, F. A. de Wolf, M. A. Cohen Stuart and J. van der Gucht, *Soft Matter*, 2013, **9**, 8737–8744.

13 A. A. Martens, J. van der Gucht, G. Eggink, F. A. de Wolf and M. A. Cohen Stuart, *Soft Matter*, 2009, **5**, 4191–4197.

14 M. Schor, A. A. Martens, F. A. deWolf, M. A. Cohen Stuart and P. G. Bolhuis, *Soft Matter*, 2009, **5**, 2658–2665.

15 N. K. Li, F. G. Quiroz, C. K. Hall, A. Chilkoti and Y. G. Yingling, *Biomacromolecules*, 2014, **15**, 3522–3530.

16 M. W. T. Werten, H. Teles, A. P. H. A. Moers, E. J. H. Wolbert, J. Sprakel, G. Eggink and F. A. de Wolf, *Biomacromolecules*, 2009, **10**, 1106–1113.

17 A. C. Balazs, C. Anderson and M. Muthukumar, *Macromolecules*, 1987, **20**, 1999–2003.

18 J. Baschnagel, J. P. Wittmer and H. Meyer, in *Computational Soft Matter: From Synthetic Polymers to Proteins*, ed. N. Attig, K. Binder, H. Grubmüller and K. Kremer, John von Neumann Institute for Computing, Jülich, 2004, ch. Monte Carlo Simulation of Polymers: Coarse Grained Models, pp. 83–140.

19 F. Müller-Plathe, *ChemPhysChem*, 2002, **3**, 754–769.

20 P. T. Underhill and P. S. Doyle, *Journal of Non-Newtonian Fluid Mechanics*, 2004, **122**, 3 – 31.

21 P. G. Bolhuis, A. A. Louis, J. P. Hansen and E. J. Meijer, *The Journal of Chemical Physics*, 2001, **114**, 4296–4311.

22 M. Wilson, A. Rabinovitch and A. R. C. Baljon, *Phys. Rev. E*, 2011, **84**, 061801.

23 A. Berkenbos, *PhD thesis*, Universiteit van Amsterdam, 2008.

24 C. Pierleoni, B. Capone and J.-P. Hansen, *The Journal of Chemical Physics*, 2007, **127**, 171102.

25 W. C. Swope, H. C. Andersen, P. H. Berens and K. R. Wilson, *The Journal of Chemical Physics*, 1982, **76**, 637–649.

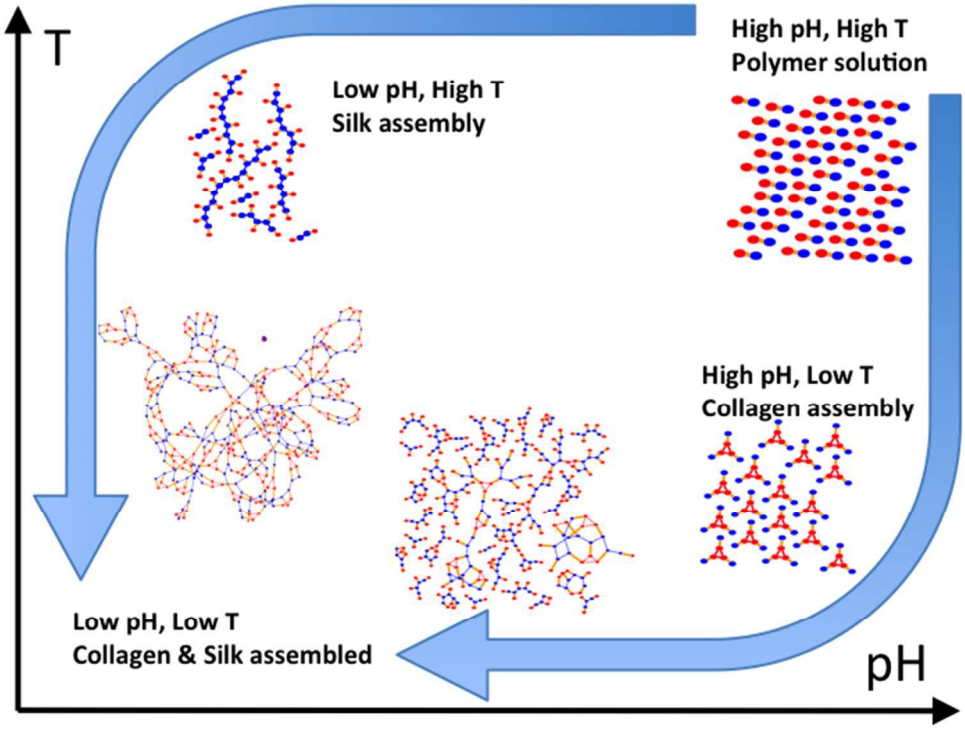
26 E. A. Koopman and C. P. Lowe, *The Journal of Chemical Physics*, 2006, **124**, 204103.

27 M. Schor, B. Ensing and P. G. Bolhuis, *Faraday Discuss.*, 2010, **144**, 127–141.

28 R. Pool and P. G. Bolhuis, *The Journal of Physical Chemistry B*, 2005, **109**, 6650–6657.

29 M. Aizenman and D. J. Barsky, *Communications in Mathematical Physics*, 1987, **108**, 489–526.

30 J. N. Israelachvili, *Intermolecular and surface forces : with applications to colloidal and biological systems*, Academic Press London, 1985.



254x190mm (72 x 72 DPI)