The role of hydrodynamic flow in the self-assembly of dragline spider silk proteins

Ana M. Herrera-Rodríguez[*,1], Anil Kumar Dasanna[*,2,3], Csaba Daday[1], Eduardo R. Cruz-Chú[4], Camilo Aponte-Santamaría[1], Ulrich S. Schwarz[2,3,+], and Frauke Gräter[1,5,+]

[1]Heidelberg Institute for Theoretical Studies, Schloß-Wolfsbrunnenweg 35, 69118 Heidelberg, Germany
[2]BioQuant, Heidelberg University, Im Neuenheimer Feld 267, 69120 Heidelberg, Germany
[3]Institute for Theoretical Physics, Heidelberg University, Philosophenweg 19, 69120 Heidelberg, Germany
[4]Laboratorios de Investigación y Desarrollo, Facultad de Ciencias y Filosofía, Universidad Peruana Cayetano Heredia, Lima, Perú
[5]Interdisciplinary Center for Scientific Computing, Heidelberg University, Im Neuenheimer Feld 205, 69120 Heidelberg, Germany
[*]These authors contributed equally.
[+]frauke.graeter@h-its.org; schwarz@thphys.uni-heidelberg.de

October 24, 2022

Abstract

Hydrodynamic flow in the spider duct induces conformational changes in dragline spider silk proteins (spidroins) and drives their assembly, but the underlying physical mechanisms are still elusive. Here we address this challenging multiscale problem with a complementary strategy of atomistic and coarse-grained Molecular Dynamics (MD) simulations with uniform flow. The conformational changes at the molecular level were analyzed for single tethered spider silk peptides. Uniform flow leads to coiled-to-stretch transitions and pushes alanine residues into β-sheet and Poly-Proline II (PPII) conformations. Coarse-grained simulations of the assembly process of multiple semi-flexible block copolymers using multi-particle collision dynamics reveal that the spidroins aggregate faster but into low-order assemblies when they are less extended. At medium-to-large peptide extensions (50%-80%), assembly slows down and becomes reversible with frequent association and dissociation events, while spidroin alignment increases and alanine repeats form ordered regions. Our work highlights the role of flow in guiding silk self-assembly into tough fibers by enhancing alignment and kinetic reversibility, a mechanism likely relevant for other proteins whose function depends on hydrodynamic flow.

1   Introduction

Water plays a major role for the stabilization, structure and dynamics of proteins [1, 2], mainly via hydrogen bond networking [3] and screening of electrostatic interactions [4]. Although most proteins function in the context of a quiescent fluid, there are important situation in which hydrodynamic flow impact protein structure by inducing non-uniform drag force along the protein. Hydrodynamic flow can mediate a large variety of processes such as material synthesis, blood coagulation and protein misfolding [5, 6, 7]. For example, the flow-induced activation of von Willebrand factor (vWf) protein is crucial in hemostasis [8, 9, 10, 11]. Shear and elongational flows can lead to protein unfolding, misfolding and aggregation, and can accelerate fibrillation of amyloid-β peptide [12, 13, 14, 15]. A complete molecular understanding of how proteins respond to flow, however, remains elusive.

Here we study an important example from material synthesis, namely assembly of dragline spider silk proteins (spidroins) into a fiber, which occurs in the *Major Ampullate* (MA) spinning gland and outperforms mechanically any biomaterial made by nature [7, 16, 17, 18, 19, 20, 21, 22]. Major ampullate spidroins consists of pH-dependent folded N-terminal and C-terminal domains, which experience conformational changes from pH 7.0 down to pH 5.0 in the spider spinning gland, thereby initiating assembly [20]. Inbetween these terminal domains, spidroins feature a long repetitive region, which makes up roughly 90% of their total sequence (Figure 1a) [23]. This repetitive region of silk proteins comprises alternating blocks of hydrophobic alanines, which form β-sheet stacks in the silk fiber, and hydrophilic glycine-rich repeats, which form the amorphous and extensible matrix of silk [16, 22, 24].

A large body of work explored the dependence of silk fiber assembly under elongational and shear flows using microfluidics [6, 18, 25, 19, 20, 26]. In single molecule force spectroscopy experiments, single spidroins molecules largely followed a worm-like chain behavior, but also exhibited unique unfolding steps [24, 27, 28, 29]. However, flow-dependent dynamics of single silk proteins and flow-induced spidroins self-assembly remain largely unknown.

Unraveling the physical effects of elongational flow during formation of spider silk would fill an important gap in our understanding of the assembly pathway of the spidroins. In particular, it could help to solve the open question whether the spinnability of silk requires intermediate formation of micelles or not [30, 31]. Additionally, it could help to better describe liquid crystalline flows inside the s-duct [17, 32, 33]. The lack of a more detailed understanding of silk self-assembly is, among others, due to the challenge to experimentally work with the highly aggregation-prone silk proteins. In principle, the proteins should be soluble to prevent fatal aggregation. It is difficult to experimentally manipulate the silk dopes due to their high concentrations $\sim 50\%$ wt/v and their sensitivity to changes in the micro-environment (pH, ion conditions and shear forces). Moreover, the lack of structuring of the silk monomers in most of its domains limits the use of structural studies.

Theory and simulations have shed further light into the stretching and assembly behavior of biopolymers under flow. For the simplest case of biopolymers under uniform flow, Monte Carlo simulations of tethered DNA chains revealed that the nonlinear elastic dumbbell model predicts their deformation [34]. The simulations included Brownian motion, entropic elasticity, and variations in the drag coefficient while the chain is deformed. Brownian dynamics simulations of DNA molecules in steady flow showed interesting differences between WLC and Rouse models in a strong steady flow, with convective as opposed to diffusive propagation of tension along the chain, respectively, leading to a distinct time evolution of chain extension [35]. In addition, bead-spring models have proven useful for describing (bio)polymers, including proteins under flow conditions. A prominent example is the coarse-grained (CG) modelling of vWf multimers in shear flow, with one bead representing roughly one monomer and a Lennard-Jones potential to account for average attractive interactions [36, 37]. A variety of methods have been used in such mesoscopic simulations for the hydrodynamic interactions of the polymer with the solvent, based on Brownian Dynamics (BD) [9, 11, 36, 38], multi-particle collision dynamics (MPCD) [39, 40, 41, 42], dissipative particle dynamics (DPD) [19, 37, 43], or lattice-Boltzmann [44]. In the context of spider silk SA by flow, a mesoscopic modelling using DPD served to find that intermediate ratios of hydrophobic and hydrophilic blocks observed in spider silks lead to exceptional silk fiber formation [19]. Shear flow was shown to lead to enhancement of β-sheet crystals, confirming a previous experimental research[6, 7, 18]. However, the effect of hydrodynamic flow on the structure of silk fibers remains elusive.

In contrast to these mesoscale simulations, atomistic MD simulations can give detailed insight into the conformational changes of proteins induced by flow, such as ubiquitin unfolding [45] or secondary structure transitions in the platelet receptor glycoprotein Ib [46]. The high degree of detail offered by atomistic MD in explicit solvent comes at a high computational cost, resulting in high flow rates used in these studies. Additionally, the high computational costs also limits the possibility to study protein assembly processes like spider silk protein self-assembly, thus most of the studies are associated with single protein dynamics under flow.

In this study, we attack these open questions in a two-pronged approach of simultaneously performing atomistic and coarse-grained simulations at physiologically relevant flow velocities. As representative example, we chose an ~ 80 amino acid-long sequence from the repetitive part of MA *Euphrosthenops Australis* dragline spidroin, which consists of three poly-ALA repeats, and two amorphous regions (Figure 1a). We modeled these proteins at the mesoscopic scale as a simple block-copolymer using a Go-like potential [47] combined with Multiparticle-Collision Dynamics (MPCD) [39] to model a uniform flow. For comparison, we performed atomistic MD simulations of the same spidroin fragments in uniform flow of explicit water [48]. We find the proteins to largely follow the expected freely-jointed chain behavior as known from single molecule stretching experiments and simulations, though with interesting differences. Flow unravels long-range poly-alanine interactions and promotes β-sheet and poly-proline II conformations, both known to be present in silk fibers, in particular within the poly-alanine repeats. Regarding self-assembly, we found an enhancement of oligomerization at low flow rates. The alanines are more prone to form interchain contacts than the residues in the amorphous region and this favors formation of β-sheet crystals during fibrillation. Our results give molecular insight into the drag force and stretching dynamics of disordered and unfolded proteins in flow as well as into flow-induced assembly of spider silk and related protein-based materials.

2   Methods

In all simulations, we model a representative fragment of silk spidroin, consisting of three polyalanine repeats separated by two glycine rich regions, with a total length of ~ 80 amino acids. Figure 1b shows all-atom and CG representations of the repetitive fragment we simulated. The simulations comprise both single silk peptides, at both all-atom and coarse-grained level, and multiple silk peptides, only at coarse-grained level, under uniform flow.
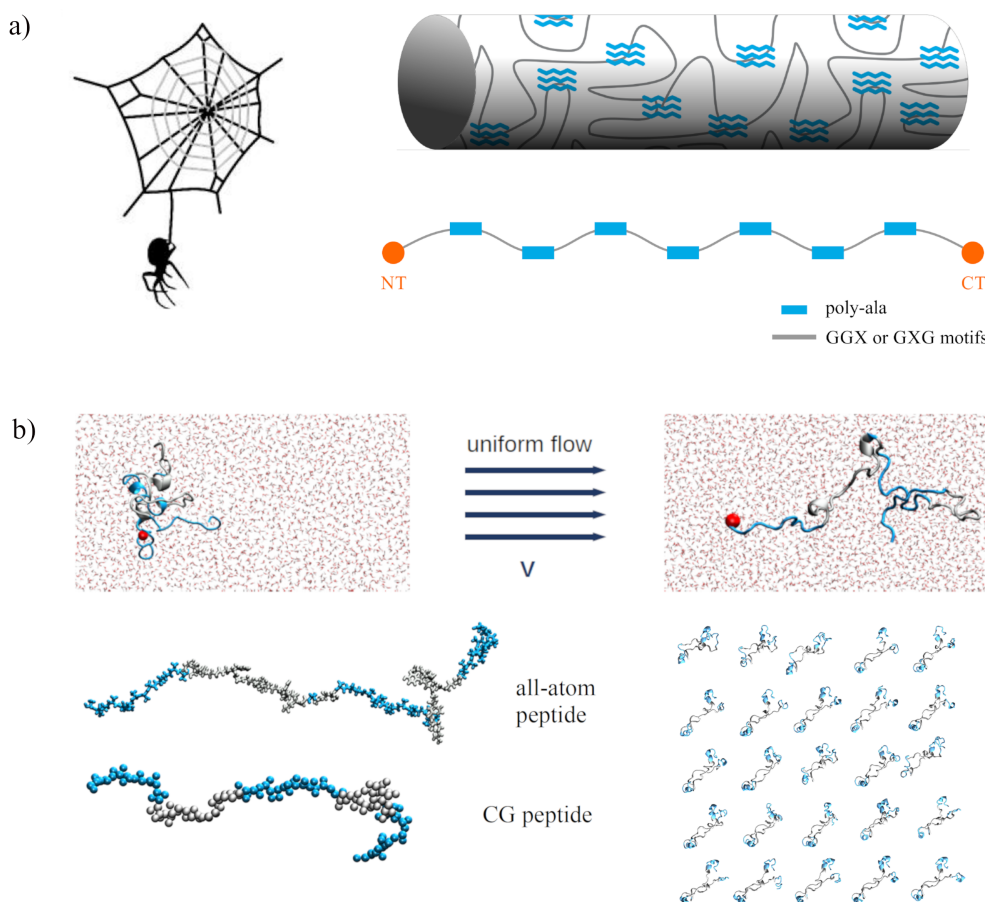
Figure 1: **Dragline silk and computational model for spidroins under uniform flow**. (a) Dragline silk corresponds to the structural part of spider's webframe (top left), which fiber microstructure consists of embedded β-sheet crystals into an amorphous viscoelastic matrix. The fiber is formed upon directed SA induced by flow of the spidroins, an scheme of the *MA Euphrosthenops Australis dragline spidroin* sequence is shown. Its molecular structure comprises three regions: pH dependent N- and C-terminal domains (NT and CT, red), as well as a long repetitive region with alternating poly-alanine (Poly-Ala, blue) repeats and glycine-rich motifs (GGX or GXG, with X being another amino acid, grey). (b) The MD model of silk under flow considers atomistic (all-atom) and coarse-grained (coarse-grained) models of a fragment of the repetitive region, containing three poly-Ala and two amorphous regions, resulting in $\sim$ 80 amino acids in total. Spheres represent atoms (for all-atom) or backbone and sidechain beads for each amino acid (coarse-grained). The spidroin fragment is tethered at the N-terminus (red sphere), and a uniform flow is produced to stretch the peptide. To model SA under uniform flow, sets of 25 of peptides (bottom right) are tethered along the flow direction and able to move transversely. Color coding as in the spidroin scheme in (a).

All-atom MD simulations of single peptides

We used the Gromacs 2018 version [49]. The sequence of the peptide originates from the repetitive part of the *Euprosthenops Australis* spidroin at the MA spinning gland and is: $[A]_{13}$GQGGQGGYGGLGQGGYGQGAGSS$[A]_{14}$GRGQGGYGQGSGGN$[A]_{12}$. It is organized in alternating Poly-Ala repeats and glycine-rich disordered fragments. We used the Amber force field Amberff99sb-star-ILDN in combination with the Tip4pD water model, which previously proved appropriate to model intrinsically disordered proteins such as the one investigated here [50, 51, 52]. Periodic boundary conditions were used in all three dimensions for all simulations. The concatenated simulation time including system preparation and production runs with flow is 18.5 μs. Before introducing uniform flow, we prepared and equilibrated the system as follows: an energy minimization with 100000 steps starting from a randomly chosen conformation of the protein was followed by 5 ns of MD simulations in the NVT ensemble and 10 ns of MD in the NpT ensemble with the protein with position restraint. Temperature coupling was done by a V-rescale thermostat during equilibration and production runs [53]. The Parrinello-Raman barostat [54] was used for pressure coupling in the NpT runs. LINCS constraints were used for all bonds to allow a 2 fs integration step. After preparation, we ran 300 ns in the NpT ensemble to equilibrate the protein without restraining the protein. After protein equilibration, we took 10 different conformations and prepared larger simulation systems along the flow direction. We simulated

10 flows with mean velocities from 0.0 m/s up to 0.48 m/s, each for 600 ns in an NVT ensemble. During the simulation we pulled waters inside a slice of 4 nm size with a constant force, using a modified Gromacs version to introduce uniform flow [48]. The V-rescale thermostat was used, which relaxed the system to achieve a constant velocity flow given by the externally applied velocity. The peptide was placed 4 nm away from the pulling region to avoid undesired stretching not associated with the flow, and artifacts due to the pressure drop inside the slice and at its interface. For the production run, the first alpha carbon of the peptide was position restraint to prevent translation. A timestep of 2 fs and LINCS constraint on h-bonds were used. We simulated for three different replicates for every flow velocity, each starting from a different velocity seed.

Coarse-grained spidroin model

We simulated the same system described in the previous section at coarse-grained level, i.e. at the resolution of two beads per amino acid (backbone bead and sidechain bead) rather than atoms. The two beads are located on the C_α and C_β atoms of each amino acid. The peptide is simplified as a block copolymer composed of two types of amino acids: alanine and a non-specific amino acid type for the disordered region. The amount of Ala residues in every region is the same as in all-atom simulations. Since the primary intermolecular interaction within silk fibers are β-sheets formed from the Poly-Ala regions, we considered attractive (native) interactions only between Ala residues in our model, more specifically between pairs of alanine backbone beads (representing the backbone hydrogen bonding within β-sheets) and between pairs of alanine sidechain beads (representing hydrophobic packing between methyl groups across β-sheets [55]. These 'native' interactions were modelled as attractive Lennard Jones interactions and all other bead-bead interactions as purely repulsive terms. We now describe all interactions of our spidroin mesoscopic SOP model including sidechains (SOP-SC) in detail.

The total potential energy U_{total} to describe the interaction between these monomers is based on the SOP-SC polymer model [47], is defined as

$$U_{\text{total}} = U_{\text{FENE}} + U_{\text{nat}} + U_{\text{exc}} . \tag{1}$$

The first term U_{FENE} is a modified form of Finitely Extensible Nonlinear Elastic (FENE) potential for the bonded interactions, i.e. beads are connected by nonlinear springs. The summation runs over all bonded backbone-backbone (BB) and backbone-sidechain (BS) beads. The potential is:

$$U_{\text{FENE}} = -\sum_{i}^{N} \frac{k}{2} R_0^2 \log \left(1 - \frac{(r_i - r_0)^2}{R_0^2} \right) \tag{2}$$

where r_0 is equilibrium length of the bond, R_0 is the maximum extent of the bond and k is the spring constant. The second term in equation (1) represents native interactions between the alanine pairs i,j such that $|i - j| > 2$. It is modeled with Lennard Jones (LJ) potential with equilibrium distance σ and energy ϵ:

$$U_{\text{nat}} = \sum_{i,j}^{\text{Ala-pairs}} U_{\text{LJ}}^{i_{\text{ala}},j_{\text{ala}}} \tag{3}$$

with

$$U_{\text{LJ}} = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right] . \tag{4}$$

The distance between BB and sidechain-sidechain (SS) beads is 0.5 nm, corresponding to the distance when there is β-sheet secondary structure formation in silk. Therefore, $\sigma = 0.5$ nm. BS interactions within alanines are not treated as attractive interactions, but are included in the repulsive term (see below).

The last term represents exclusion interactions:

$$U_{\text{exc}} = \sum_{i,j \notin \text{Ala-pairs}} U_{\text{LJ,rep}}^{i,j}, \tag{5}$$

and

$$U_{\text{LJ,rep}} = \begin{cases} U_{\text{LJ}}, & r < r_{\text{cut}} \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

where $r_{\text{cut}} = r_m$ is the distance at which the potential has a minimum value. The summation in equation 5 runs over all possible pairs of beads, except over native interaction pairs.

The parameters associated with the potential energy are for:

- U_{FENE} potential: bond length $r_0 = 1.0$ [a], $k = 500$ [$k_B T/a^2$] and $R_0 = 0.5$ [a].

- U_{nat} potential: $\epsilon_{nat} = 0.25$ [$k_B T$] and $\sigma_{nat} = 2\,r_0$ [a] of the LJ potential for both BB and SS interactions.

- U_{exc} potential: the parameters have the same value for BB, SS and BS interactions, which are $\epsilon_{exc} = 1.0$ [$k_B T$] and $\sigma_{exc} = r_0$ [a].

4

Multi-particle Collision Dynamics simulations

The CG model of the peptide was combined with hydrodynamic flow using Multiparticle Collision Dynamics (MPCD), which is a particle-based mesoscopic simulation technique for the Navier-Stokes equation. It incorporates hydrodynamic interactions and thermal fluctuations [56, 39]. We employ a local cell-level scaling thermostat [57] to maintain the temperature of the system constant. The fluid in MPCD consists of point particles m, position $\vec{r}(t)$ and velocity $\vec{v}(t)$. Particles are sorted into a cubic lattice with a lattice constant a and subjected to the dynamics of two steps:

- a streaming step where the particles move ballistically during a collision time step Δt_c. Particle positions at every timestep are described by

$$\vec{r}_i(t + \Delta t_c) = \vec{r}_i(t) + \Delta t_c \vec{v}_i(t), \tag{7}$$

  where $\vec{r}_i$ and $\vec{v}_i$ are the positions and velocities of every particle. Position update follows thus a standard integration of Newton's equation of motion

- The collision step where particle velocities are altered. The center-of-mass velocity of the cells $\vec{u}_{\text{com}}$ is computed, and subsequently the relative velocities of the particles with respect to $\vec{u}_{\text{com}}$ are rotated around a random axis by an angle α, which ensures linear momentum conservation and thermal fluctuations. The velocity evolution for every particle in time is given by

$$\vec{v}_i(t + \Delta t_c) = \vec{u}_{\text{com},j}(t) + \vec{\Omega}_j(\alpha)[\vec{v}_i(t) - \vec{u}_{\text{com},j}(t)]. \tag{8}$$

$\vec{\Omega}$ and j correspond to the rotation matrix in spherical coordinates and the cell index respectively. The beads of the peptide are included in the MPC collision step, in order to take into account solvent-peptide interactions. The integration time step Δt was chosen to be smaller than the collision time step Δt_c. To induce flow, a constant force F was applied to every solvent particle.

The fluid viscosity η in MPCD depends on the collision time step Δt_c, particle density ρ and mass m. It has two contributions: collision viscosity η_{col} which results of momentum transfer between the particles during the collision step and kinetic viscosity η_{kin} which results from the stage of particle streaming, and are given by

$$\eta_{kin} = \frac{\rho k_B T \Delta t_c}{a^3} \left[\frac{\rho}{\rho - 5/4} - \frac{1}{2} \right], \tag{9}$$

$$\eta_{col} = \frac{m(\rho - 7/5)}{24a\Delta t_c}. \tag{10}$$

For low Reynolds numbers η_{kin} can be neglected and the main contribution to the total viscosity η is provided by the collision viscosity $\eta \simeq \eta_{col}$

The simulation parameters are expressed in basic units $k_B T = 1$, $a = 1$ and $m = 1$, where k_B is the Boltzmann constant and T is the temperature. The box volume is $V = 100.0 \times 50.0 \times 50.0 \ a^3$ and the particle density is set to $\rho = 10$. The mass of every monomer is $M = \rho m$, the integration time step Δt is $0.00625\sqrt{ma^2/k_B T}$ and the collision time step is $\Delta t_c = 0.1 \ \sqrt{ma^2/k_B T}$.

The simulations consisted of an equilibration of the system for 12000 $\sqrt{ma^2/k_B T}$ starting from a fully extended chain, from which its final conformation was taken to simulate under 10 different flows. The force applied to the solvent particles to simulate uniform flow at a range of velocities was in the interval of $(5 \cdot 10^{-07} - 2.7 \cdot 10^{-4})[(k_B T/a)]$. We carried out three replicas with different velocity seeds for each flow velocity in both single peptide and multiple peptide simulations.

Coarse-grained simulations of silk assembly

To monitor the self-assembly of spider silk proteins, we performed simulations of peptides sets under uniform flow in coarse-grained simulations. Every spidroin corresponds to a random conformation taken from the single peptide simulations in a given flow regime. Like in the single silk spidroins simulations, the silk peptides are tethered as well along the flow direction. Initially, the peptides are placed equi-distant from each other in a matrix of 5x5, where the positions (y, z) of the tether atoms correspond to the nodes of the matrix. Figure 1b (bottom right) shows the initial configuration of the peptides for one replicate of the all-atom simulations. The peptides are allowed to move transversely (in the yz plane) and are tethered at the same position in the flow component x.

To monitor the self-assembly of spider silk proteins, we performed simulations of peptides sets under uniform flow in coarse-grained simulations. Every spidroin corresponds to a random conformation taken from the single peptide simulations in a given flow regime. Like in the single silk spidroins simulations, the silk peptides are tethered as well along the flow direction. Initially, the peptides are placed equi-distant from each other in a

Mean flow velocity mpcd units	Mean fractional extension of peptides L	Simulation time mpcd units	Number of replicates	Number of peptides
0.1	0.4	800	3	25
0.2	0.7	700	3	25
0.3	0.8	600	3	25
0.4	0.9	600	3	25
0.5	0.4	500	3	25
0.7	0.7	300	3	25
0.8	0.8	200	3	25
1.0	0.9	200	3	25

matrix of 5x5, where the positions (y, z) of the tether atoms correspond to the nodes of the matrix. Figure 1b (bottom right) shows the initial configuration of the peptides. The peptides are allowed to move transversely (in the yz plane) and are tethered at the same position in the flow component x.

We used the same protocol and simulation parameters as in the single peptide case (Subsection 2) except the bond length was decreased to $r_0 = 0.7$ from $r_0 = 1.0$ to reduce the overall system size and increase computational speed. This only alters the effective friction of a single bead. The initial peptide separation is 7,6 nm and the system concentration is $\sim 4\%$ wt/v for the CG system. In total, we performed 6 simulations with 3 replicates each. A description of each simulation is contained in Table 2.

Persistence and Kuhn lengths

To obtain the Kuhn length of the polymer at all-atom and coarse-grained scale, we fitted an exponential function to the auto-correlation function $C(s)$ of the silk peptide backbone segments s, using:

$$C(s) = exp(-2s/l_k). \tag{11}$$

The Kuhn length for all-atom and coarse-grained simulations was 0.2 nm and 1.6 a respectively. With $l_k = 2l_p$ we obtained the Kuhn length for both models.

Drag force calculation

We calculated the drag force acting on the protein in all-atom MD simulations from inter-atomic forces F_{ij} between protein atom i and protein atom j, summed the x-component of all these forces up (with x being the flow direction), and averaged over time. We used the same definition of the drag force for the coarse-grained simulations, just that in this case, the forces only included the momenta transferred in the collision steps as only local protein-solvent forces are considered in MPCD.

Secondary structure analysis

For the secondary structure analysis, we defined residues to be in a β-sheet conformation when the dihedrals fall into the intervals of $90 \leq \psi \leq 180$ and $-180 \leq \phi \leq -90$ (circles), and in a PPII conformation if $128 \leq \psi \leq 180$ and $-90 \leq \phi \leq -58$ (squares). In order to observe the relevance of alanines in the formation of β-sheets we show the percentage of Ala and non-Ala residues like: $\%Ala_{\beta/PPII} = N_{Ala_{\beta/PPII}} * 100/N_{Ala}$ (blue symbols) and $\%nonAla_{\beta/PPII} = N_{nonAla_{\beta/PPII}} * 100/N_{nonAla}$ (gray symbols). $N_{Ala_{\beta/PPII}}$ and N_{Ala} correspond to the number of Ala residues in β-sheet or PPII helix and the total number of alanines in the peptide sequence, while $N_{nonAla_{\beta/PPII}}$ and N_{nonAla} are the number of residues of the amorphous region in β-sheet or PPII and the total number of residues different from alanines.

Friend of friends algorithm

To monitor whether a silk peptide forms part of an oligomer of certain size or not, we adapted a group selection algorithm called Friend of Friends (FoF) to our case [58]. First, we chose a peptide, which has not been previously assigned to an oligomer group i. Next, we search for partners of this peptide based on the inter-chain C_α carbons contact formation, where the condition to have a contact of two atoms between peptide 1 and peptide 2 is defined as $d_{C_\alpha}^{12} \leq 0.5$ nm (see the spidroin inter-contact sketch at Figure S1). In case, the peptide does not have a partner, we assign it to an "isolated" peptide list. When partners are found, all of them are added to the list

of i oligomers. We search repeatedly until no further members can be found. This process runs over all frames of the simulation. Therefore, we can monitor not only the time evolution of individual chains that form oligomers, but also their contact formation with other chains.

3   Results

3.1   Single silk spidroins under flow

We first asked how the flow velocities are related to the drag force acting on the peptide. As the drag force changes dynamically with the peptide conformation, a simple analytical model of the drag as a function of velocity might not be appropriate. We instead calculated the drag force directly from the forces between the solvent particles and the protein particles (atoms or beads respectively). In the all-atom case, we computed the vector-sum of the non-bonded forces between all atoms of the water molecules and of the peptide over time. In the coarse-grained case, we obtained the drag force from averaging the collisions of the solvent particles with the beads, as the solvent-protein interactions solely occur through local momentum exchange. In both all-atom and coarse-grained simulations, the force components along the flow yields a non-zero average force, the drag force. We averaged the drag force over the last half of each trajectory and over all three replicates. Interestingly, independent on the resolution of the simulations, i.e. for both all-atom and coarse-grained, we obtain an overall linear increase in drag force with flow velocity (Figure 2a and Figure 2b). Thus, the drag force effectively follows Stokes' law, with an effective hydrodynamic radius of 2.52 nm (obtained from the linear fit to the AA data of Figure 2a). We conclude that our spidroin peptide does not exhibit a pronounced globule stretch transition at a critical flow speed, which would lead to a deviation from a linear drag-flow relationship, but instead steadily extends with increasing flow velocity. This is in sharp contrast to the behavior seen previously for proteins in shear flow, which showed a rather abrupt globule-stretch transition (in case of VWF) [9] or a one-step unfolding (in case of ubiquitin) [40, 41], respectively. However, a gradual stretching through multiple unfolding intermediates were observed for ubiquitin in the case of elongational flow, [40, 41] directly in line with the gradual extension of the spidroin with increasing flow velocities.
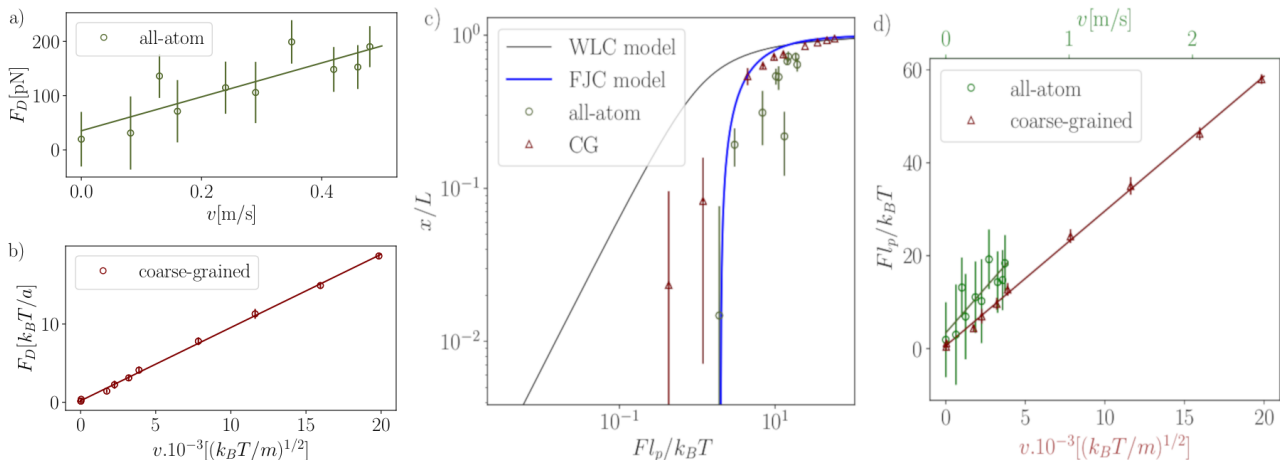

Figure 2: **Stretching silk peptides under uniform flow**. (a) Averaged peptide drag force over mean flow velocity for all-atom simulations. The drag force in the peptide originates from the non-bonded interactions between the water and protein atoms. The green solid curve is a linear fit. (b) Averaged peptide drag force for coarse-grained simulations, the drag force was computed from the collisions between the solvent particles and the amino acids of the peptide. The maroon solid curve is a linear fit. (c) Dimensionless peptide drag force as a function of the normalized peptide extension, L is the contour length. Circles and triangles correspond to all-atom and coarse-grained simulations, respectively. The persistence length was extracted from fitting to equation (11). Black and blue solid curves: normalized theoretical prediction from the WLC and FJC models, $f_{FJC}(x) = 1/(1 - x)$ and $f_{WLC}(x) = 1/4(1 - x/L)^{-2} - 1/4 + x/L$. Error bars in all the plots are standard errors of the mean for nine data points, obtained from time-averages over three 100 ns windows of the second half of each of the three trajectories per flow speed. (d) Dimensionless drag force over flow velocity for all-atom and coarse-grained simulations, l_p is the persistence length. The highest velocity of the atomistic simulations is about one fifth the highest velocity of the coarse-grained simulations. The maroon and green solid curves are linear fits.

To quantitatively compare the all-atom and coarse-grained scales, Figure 2c shows for both sets of simulations the dimensionless drag force in the peptide as a function of the normalized extension along the flow. L and l_p correspond to the contour length and persistence length, which are 24.3 nm and 0.4 nm for all-atom simulations

as well as 80 a and 3.1 a for coarse-grained simulations. We recover a non-linear force-extension curve typical for polymers when stretched by a pulling force. The all-atom model yields overall lower extensions for a given drag force, but given that they are fully independently parameterized, the agreement is satisfying. The coarse-grained model extends more readily is likely due to our approximation that only attractive interactions are present within alanine residues, neglecting unspecific favorable interactions involving non-alanine residues of the disordered region. The force-extension behavior is reminiscent of the behavior predicted by the freely-jointed chain (FJC) [59] or worm-like chain (WLC) model [60] of polymers. Indeed, we find a good agreement of our data with the FJC model (solid blue curve Figure 2c). We find slightly less agreement between the WLC model and our particle-based simulations (Figure 2c, solid black line), which we explain by the nature of the protein backbone the (albeit only partially) rotatable bonds of which are better depicted by a FJC model than by the rod-like WLC with uniform bending stiffness. Stretching the silk peptide to ~80 % of its contour length ($x/L = 0.8$) requires a drag force from uniform flow in the range of 100-200 pN, according to both all-atom and coarse-grained simulations. Interestingly, this range covers the rupture forces of ~176±73 pN required to stretch single dragline silk molecules by an atomic force microscope [27]. To estimate the velocities achieved in the coarse-grained model, the Figure 2d shows the normalized drag force as a function of velocity for all-atom and coarse-grained models, the velocities in the all-atom model are re-scaled by a factor α =7.75. The highest flow velocity in the coarse-grained simulations corresponds to ~2.6 m/s.

Alanine residues are expected to play an important role during shear-induced unfolding of silk peptides and their assembly into a fiber. To analyze the relevance of Ala residues in the silk protein conformational ensembles in flow, we quantified the number of Ala contacts along flow velocities in all-atom (Figure 3a) and coarse-grained simulations (Figure 3b). In the all-atom case, we observe a drop of Ala-Ala contacts between 0.16 and 0.24 m/s of flow rate. Also the fractional extension of the all-atom peptide as a function of the flow rate confirms this behavior (Supplementary Figure S2a): initially the peptide is extended up to about 30% of the contour length from 0.08 m/s up to 0.16 m/s. After this threshold, the peptide is stretched 20% further in the mean flow velocity interval of (0.24 - 0.29) m/s up to around 80% of its contour length at 0.48 m/s.
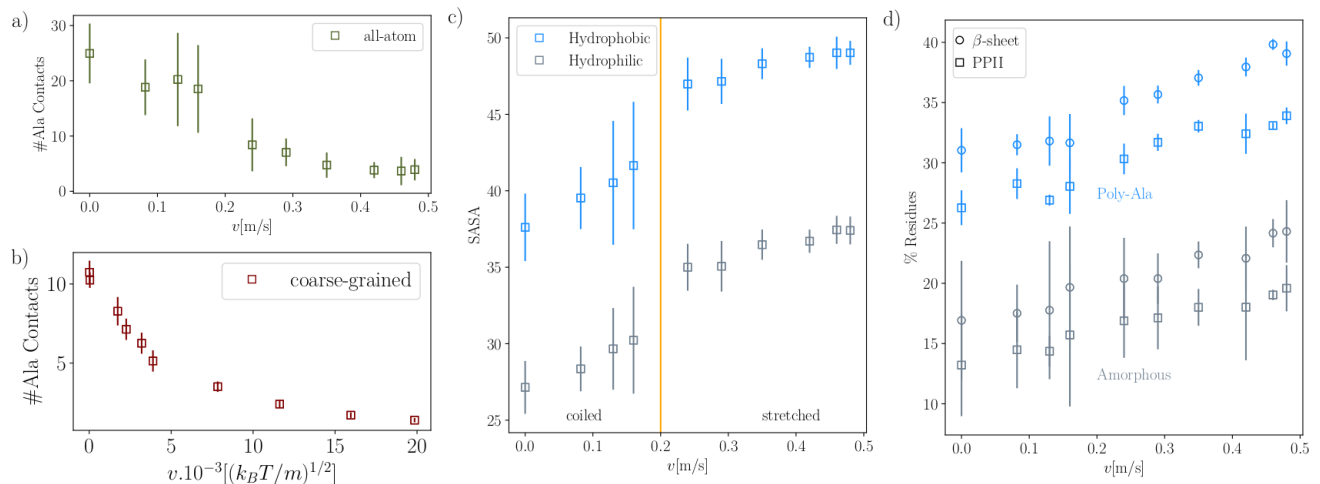

Figure 3: **Molecular details during silk elongation**. a) Number of contacts of Ala residues as a function of time in all-atom MD simulations. The cutoff radius within we examine a contact formation is 0.5 nm. b) Number of contacts of Ala residues as a function of time in coarse-grained MD simulations. The cutoff radius within we examine a contact formation is 0.5 nm. c) Hydrophobic and hydrophilic solvent accessible surface areas over flow rate for all-atom simulations. At about 0.2 m/s a transition of the peptide to a extended state is observed (solid vertical orange line). The highest velocity of the atomistic simulations is about one fifth the highest velocity of the coarse-grained simulations. d) Flow populates β-sheet and PPII conformations. Percentage of dihedral angles of amino acids in β-sheet (circles) and PPII conformation (squares) as a function of mean flow velocity, as observed in all-atom simulations. The data is separated into alanine (blue symbols) and amorphous regions (gray symbols), and normalized by the total number of residues in these two regions. Error bars show the standard error of the mean. See Methods for the definition of the dihedral intervals for both secondary structures.

In the case of the coarse-grained peptide we observe a less abrupt transition to a fully extended state without Ala-Ala contacts. The contacts steadily decrease from 10 contacts to nearly 0 contacts between 0.02 and $20 \times 10^{-3} (k_B T/m)^{1/2}$ (Figure 3b). The number of total and amorphous contacts over mean flow rate in Figure 3b shows the same tendency. With regard to the coarse-grained fractional extension (Supplementary Figure S2b) over the mean velocity, Ala contacts are maximal only at 0-10% fractional extensions, and are largely lost only beyond 90% extensions. Thus, the CG model does not reproduce the comparably sudden transition observed at all-atom. In addition, the coarse-grained model shows less contacts initially, i.e. at low flow velocities (~12

instead of $\sim$24 for all-atom). We conclude that the CG model reproduces the overall loss of contacts in flow, though it cannot resolve the collapse-stretch transition of all-atom simulations.

The observed abrupt transition of the peptide from coiled to stretched conformation in the all-atom is associated with a decrease of the hydrophobic effect. It can be observed at the Figure 3c, which shows the solvent accessible surface area (SASA) over flow rate. Upon force, the hydrophobic residues are not buried and the silk peptide suddenly transits from a coiled to a extended state. This transition occurs about at $0.2\,\mathrm{m/s}$ of flow rate (see the orange vertical line at Figure 3c).

The formation of β-sheets is a key factor in the self-assembly of spider silk and requires flow [18]. However, how flow induces β-sheet formation is poorly understood. In recent studies, a significant population, namely $\sim$ 24%, of poly-proline II (PPII) helix in the repetitive region of MA spidroin dopes, i.e. prior to flow-induced elongation during spinning, from the *Euphrostenops Australis* spider has been observed [61]. It was suggested that PPII helices can form a rigid structure that can be quickly transformed into β-sheets. We here asked to what extent our spidroin fragment samples β-sheet and PPII conformations and how these backbone propensities are influenced by flow. We defined for each amino acid the respective backbone configurations based on its two backbone dihedral angles (see Methods for details). We averaged over all alanines and all amorphous-phase residues (non-alanines) and normalized by the total number of alanines and non-alanines, respectively. We note that we analyzed the secondary structure propensities only for the AA simulations, as the CG lacks the required resolution.

Figure Figure 3d shows the percentage of amino acids with dihedral angles either in β-sheet or PPII conformation as a function of mean flow velocity. The percentage of alanines in both, β-sheet and PPII conformations is increasing with the mean flow velocity for both the poly-Ala and amorphous regions. Both, β-sheet and PPII conformations populate the upper left corner of the Ramachandran plot and are close to the most extended backbone configuration with $\psi = 180$ and $\phi = -180$. For this reason, flow is expected to drive the amino acids into this region of the dihedral space, as previously also observed for disordered proteins under a stretching force [62]. The content of the backbone torsion angles in β-sheet and PPII in the poly-Ala regions remains around 10% higher than in the amorphous regions, largely independent of the flow velocity. This higher ratios are in line with the view that the poly-alanine repeats preferentially drive β-sheet formation, i.e. crystallization, during fiber formation.

In both the alanine and amorphous phase, β-sheet conformations are slightly preferred over the PPII conformation, and they together represent the majority of secondary structure conformations. More specifically, in the poly-alanine regions, at the absence of flow, β-sheet and PPII dihedral conformations are on average sampled by more than 50% of the residues and reach $\sim$70% at the largest flow rates we have used. Taking into account that the transition from PPII-helix to *beta*-sheet is very likely to happen due to their close proximity in terms of dihedrals, we can conclude that poly-alanine regions in our spidroin are primed for β-sheet formation by uniform flow. The predominantly proportion of alanines of $\sim$70% in β-sheet or close to β-sheet (that is, PPII) conformations is already very close to the 90% percent β-sheet content by alanines predicted in the model of Alexandra H. Simmons et al. [63], which is based in NMR experimental studies. The occurence of torsional angles in PPII-helix conformation might result in a prefibrillar form of the repetitive region of the spidroins before forming stronger β-sheet interactions, as proposed by Nur Alia Oktaviani et al. [61].

3.2 Silk spidroin assembly under flow

Having established a coarse-grained model of the spidroin protein, MD simulations of multiple proteins offer the possibility of analyzing in detail how flow drives self-assembly. We set up systems of 25 spidroin peptides in the coarse-grained model and monitored their dynamics, interactions and oligomerization at a range of flow velocities. The simulation setup is explained in the Methods section 2.2. All-atom MD simulations were not performed in this stage due to their significantly higher computational expense and the formation of voids observed at higher velocities which limits range of accessible flow velocities [48].

As in the simulations of single chains, the silk fragment comprises $\sim$80 residues, that is, three poly-Ala repeats and two intrinsically disorder regions. An initial pre-alignment of the peptides was performed, mimicking the multimerization at the N-terminal domains (Figure 1). The pre-alignment is maintained by tethering the peptides along the flow direction at the same position, while they can move transversely. Imposing this pre-alignment will enhance the packing of the chains (via formation of contacts mostly within the poly-Ala regions). In this way, we monitored assembly in an idealized setting with all chains in phase and subjected to uniform flow. The mean peptide fractional extension as a function of the mean flow velocity, averaged over the replicates simulated at the same flow, is shown in the Supplementary Figure S3). We covered mean fractional extensions in the range of $[0.0 - 0.76]L$. In fact, the lower four out of the total 8 velocities applied all resulted in fractional extensions close to zero. This is in sharp contrast to the single-chain simulations, in which the same range of flow velocities has lead to measurable expansions of the chain (compare Figure 2c). Here, with a relatively high density of spidroins, the lower range of flows was not sufficient to compete with the high tendency of the chains to interact and could not expand the chains prior to interchain aggregation, which further reduced their extensions drastically.
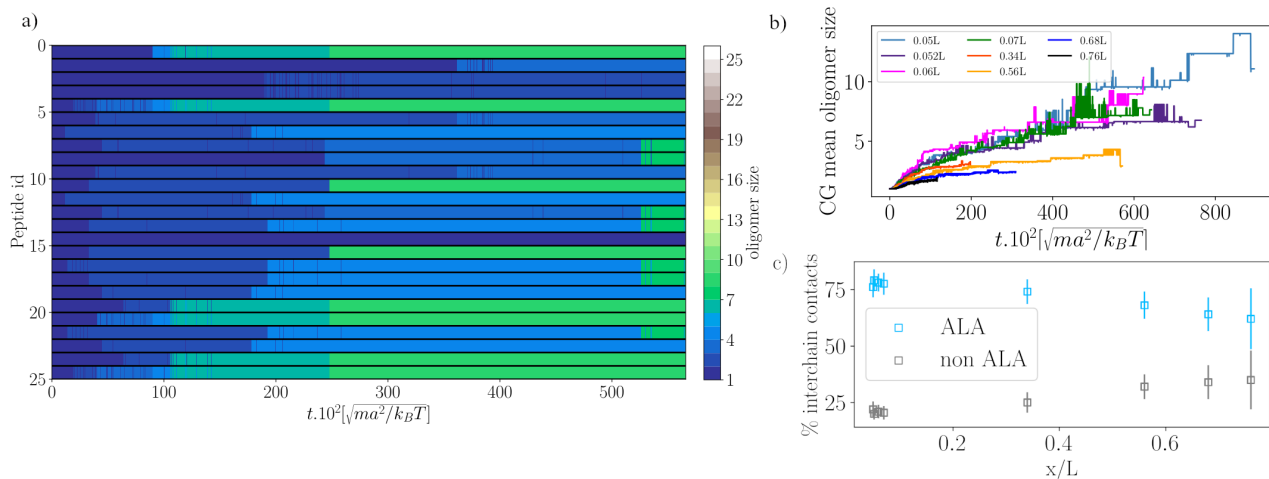
Figure 4: Silk oligomerization. a) Oligomer formation over time for every peptide. The color code indicates the size of the oligomer they form. Shown is the oligomerization for one replicate at a mean fractional extension of 0.56L. b) Mean oligomer size as a function of time. c) Average percentage of interchain contacts for the poly-Alanine and amorphous regions over the mean CG peptide fractional extension x/L.

In order to analyze the self-assembly of the spidroins over time, we monitored when a certain peptide takes part of an oligomer. This was done through a clustering algorithm called Friends of Friends (FoF) (see Methods section 2.2). The criterion to form an oligomer of certain size is based on the formation of a single interchain $C_\alpha - C_\alpha$ interaction. In Figure 4a, we visualize the oligomerization process. Data is shown for one replicate at an intermediate flow velocity, where the mean fractional extension of the peptides corresponds to $0.56L$. We observe successive growth of oligomers, and an overall increase in the number of peptides involved in dimers or larger assemblies. The largest oligomer formed within the simulated time is a 10-mer in this specific case. A reduction in oligomer size is observed rarely, as fluctuations at short time scales. At a lower flow rate, with mean peptide fractional extensions close to $0.0L$, oligomers form at shorter time scales and reach larger sizes (Supplementary Figure S4). Also, the partial reversibility observed at $0.56L$ is not observed at lower flow rates, and assembly becomes fully irreversible in this case.

To further quantify the effect of flow velocities on the spidroin assembly, we computed the average oligomer size as a function of time over all replicates (Figure 4b). Average oligmer sizes for individual replicates, exemplarily for two different flows, are also shown in Supplementary Figure S5). Overall, self-assembly of silk peptides slows down with increasing flow velocity (simulation snapshots for two different flows are shown in supplementary Figure S6). Slower flow causes a higher mean oligomerization at a given time (light blue, purple, green and magenta curves in Figure 4b) compared to faster flow regimes (red, orange, dark blue, and black curves in Figure 4b). This suggests that flow not only extends the spidroin chains, but in this way also reduces oligomerization. We attribute this slowdown in oligomerization to fluctuations of the protein chains orthogonal to the flow direction, which reduces the likelihood of diffusional encounters. Thus, random associations are reduced, and controlled assembly into β-sheet is enhanced. Indeed, the inter-protein contacts between the spidroins are strongly affected by flow. By definition, oligomerization involves the formation of inter-chain contacts. In addition, the type of contacts changes with flow. Namely, higher flow velocities and extensions result in more non-alanine contacts, due to the fact that chains are aligned in the oligomers (Figure 4c).

To investigate the oligomerization kinetics, we computed association and dissociation time intervals at every flow regime. Figures 5a and 5b show log-normal cumulative distributions of the time intervals at which association and dissociation events, respectively, occur. The time interval for a certain event (association or dissociation) of an oligomer of size N_o corresponds to $\delta t^{io}_{olig} = t^{io}_f - t^{io}_0$. The io index is the oligomer identifier, t^{io}_0 is the simulation time where an oligomer io exists and t^{io}_f is the simulation time where either the io oligomer increased its size (association) or decreased its size (dissociation). As expected, for all flow velocities, association dominates over dissociation, reflecting the overall growth in assemblies over time (Fig. 3b). Interestingly, both the numbers of association and dissociation events increase with flow, with differences particularly large at short times. Thus, higher flow overall slows down assembly, but it does so by primarily accelerating dissociation, i.e. disassembly. Overall, flow prevents rapid aggregation into large oligomers by enhancing reversibility.

4 Discussion and Conclusion

We aimed to understand the influence of uniform flow on the dynamics of silk self-assembly. To this end, we focused on one-end tethered single and multiple dragline spidroins, and performed MD simulations under uniform
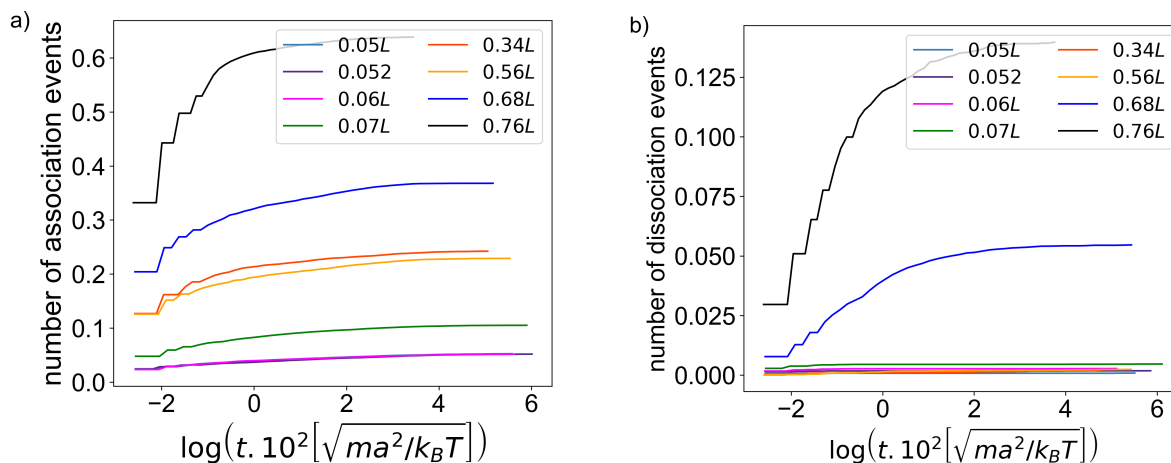
Figure 5: **Oligomerization kinetics.** Log-normal cumulative distributions of the time intervals of (a) association and (b) dissociation events, for every flow velocity (or every average protein extension). The distributions are based on data from three replicates at every flow velocity, and are normalized by the total number of time steps simulated at every flow rate. Higher flow, or larger average protein extensions, give rise to both faster association and dissociation events.

flow on two scales: atomistic (AA) and amino-acid (CG) resolution. In this way, we could consider atomistic details from AA simulations of single chains and on the other hand covered larger time and length scales using CG simulations for the assembly. In this combination, AA simulations served to analyse in detail the structural changes of spidroins with flow (Figure 3c and Figure 3b), while the CG simulations helped us to simulate larger systems with an adequate description of the hydrodynamic interactions at play.

In the single peptide simulations, the explicit computation of the drag force from the interactions between the solvent particles and every amino acid of the spidroins allowed us to reproduce the force-extension behavior for silk proteins observed in atomic force microscopy experiments [27]. Thus, the sum of drag forces acting along a peptide by colliding water molecules can be equated with the pulling forces acting on a peptide on its termini, when estimating the resulting end-to-end distance. For the dynamic response of the spidroin protein to flow, we find a good agreement between the AA and CG model, with a loss of intrachain interactions and an increase in extension largely following the WLC model. However, the CG model is not able to reproduce the step-wise transition from a collapsed to an extended state that is observed in the AA model. In contrast, it shows a steady decrease of contacts with the mean flow velocity. We attribute this to the approximation of the spidroin as a block copolymer that lacks favorable interactions and specificity for the amorphous amino acids. We also note that the mesoscopic solvent of the CG model captures hydrodynamic interactions but the molecular nature of these interactions via hydrogen bonds and the hydrophobic effect are not explicitly included, another potential reason for the lack of a well-defined collapse-stretch transition in the CG simulations.

One interesting observation of the AA simulations is the increase of alanines in beta-sheet conformations with flow, increasing their propensity for fibrillation and the formation of beta-sheet crystals as found in silk fibers [55].

Due to the high computational cost and the limited flow velocity that can be achieved in AA MD simulations [48], we only simulated CG multiple spidroins to monitor their assembly. Silk fiber formation is a highly complex and multiscale process occuring in the gland, and capturing the full complexity, from the gradual increase in the flow rate, the transition from a micellar to an extended state of the proteins, to the change in pH and salt concentration, to only name a few, is prohibitive. Our highly simplified setup was motivated by the following scenario: The N and C-terminal domains (NT and CT) experience conformational changes upon assembly inside the S-gland. The CT is dimerized from the tail of the gland [21], and the NT domain dimerizes inside the first part of the s-duct after a drop of pH [64]. Therefore, both terminal domains promote the initial interconnections of the proteins, and thereby their pre-alignment. To mimic this role of the terminal domains of the spidroins for multimerization (Figure 1), we pre-aligned the repetitive silk peptides to maintain them in phase. This setup facilitated the assembly via crystal formation in the poly-Ala regions, and allowed to focus mostly in the flow effects involved in the spidroins' packing.

Regarding the assembly by the formation of crystals, several studies have pointed out the high tendency of β−sheet structures between the poly-Ala regions of the spidroins [63, 65, 66]. CG simulations can extend the limited time scales of the AA systems, but at the expense of an accurate structural model. We here chose a CG model that still captures the sidechain packing of alanines across beta-sheets in crystals by using two beads per amino acid. Our CG model was able to identify two key tendencies of flow-induced silk fiber assembly. Flow extends the chains and overall slows down the oligomerization. The underlying kinetics shows that this is a

consequence primarily of enhancing and accelerating also dissociation. As a result, flow introduces reversibility and removes kinetic traps. This situation is similar to assembly of protein complexes under quiescent conditions, like virus capsids, which is stalled in kinetic traps if the free energy gain is too high [67, 68, 69]. For the case of silk discussed here, moderate amounts of shear, but not too high values increase the quality of the assembly, allowing proper beta-sheet configurations to form not only because of the mere extension and alignment of the poly-alanine repeats, but also by speeding up their reversible association and dissociation, as a mechanism of quality control. We note that our observations on silk assembly at the CG level, by comparison to the all-atom simulations, correspond to flows in the range of 0.5-3 m/s (Figure 2c). These speeds are relevant speeds, albeit at the upper boundary, of those used in a laboratory setting or observed within in the spider gland [70]

Overall, we have investigated by computer simulations the effect of flows on spider silk peptides. Our study gives insights into their flow-induced non-equilibrium conformational dynamics and how flow influences their self-assembly. Microfluidic experiments have shown how elongational flow enhances $\beta-$sheet formation [18], which we here can confirm. However, due to the high hydrophobicity of the silk dope, a careful control of the solvent during the assembly process is required. Microscopy experiments under controlled conditions of pH and flow to prevent fatal aggregation, and where the assembly of the spidroins can be monitored in real-time are much desired. Such experiments can help to put the mechanisms of flow-induced $\beta-$sheet formation we here put forward to test.

5   Acknowledgments

References

[1] W. Huang, S. Krishnaji, O. R. Tokareva, D. Kaplan, and P. Cebe, "Influence of water on protein transitions: Thermal analysis," *Macromolecules*, vol. 47, no. 22, pp. 8098–8106, 2014.

[2] W. Huang, S. Krishnaji, O. R. Tokareva, D. Kaplan, and P. Cebe, "Influence of water on protein transitions: Morphology and secondary structure," *Macromolecules*, vol. 47, no. 22, pp. 8107–8114, 2014.

[3] Y. Kim, H. Choi, I. Baek, and S. Na, "Spider silk with weaker bonding resulting in higher strength and toughness through progressive unfolding and load transfer," *Journal of the Mechanical Behavior of Biomedical Materials*, vol. 108, p. 103773, 2020.

[4] H. Boroudjerdi, Y.-W. Kim, A. Naji, R. R. Netz, X. Schlagberger, and A. Serr, "Statics and dynamics of strongly charged soft matter," *Physics reports*, vol. 416, no. 3-4, pp. 129–199, 2005.

[5] P. Szymczak and M. Cieplak, "Hydrodynamics effects in proteins," *Journal of physics: Condensed Matter*, vol. 23, no. 3, p. 033102, 2010.

[6] L. Lu, S. Fan, L. Geng, J. Lin, X. Yao, and Y. Zhang, "Flow analysis of regenerated silk fibroin/cellulose nanofiber suspensions via a bioinspired microfluidic chip," *Advanced Materials Technologies*, vol. 6, no. 10, p. 2100124, 2021.

[7] Y. Han, M. Langer, M. Medved', M. Otyepka, and P. Král, "Biomimetic materials: Stretch-healable molecular nanofibers," *Advanced Theory and Simulations*, vol. 3, no. 10, p. 2070023, 2020.

[8] C. Baldauf, R. Schneppenheim, W. Stacklies, T. Obser, A. Pieconka, S. Schneppenheim, U. Budde, J. Zhou, and F. Gräter, "Shear-induced unfolding activates von willebrand factor a2 domain for proteolysis," *Journal of Thrombosis and Haemostasis*, vol. 7, no. 12, pp. 2096–2105, 2009.

[9] S. Schneider, S. Nuschele, A. Wixforth, C. Gorzelanny, A. Alexander-Katz, R. R. Netz, and M. Schneider, "Shear-induced unfolding triggers adhesion of von willebrand factor fibers," *Proceedings of the National Academy of Sciences*, vol. 104, no. 19, pp. 7899–7903, 2007.

[10] H. Chen, M. A. Fallah, V. Huck, J. I. Angerer, A. J. Reininger, S. W. Schneider, M. F. Schneider, and A. Alexander-Katz, "Blood-clotting-inspired reversible polymer-colloid composite assembly in flow," *Nature Communications*, vol. 4, no. 1333, 2013.

[11] H. Chen, M. A. Fallah, V. Huck, J. I. Angerer, A. J. Reininger, S. W. Shneider, M. F. Schneider, and A. Alexander-Katz, "Blood-clotting-inspired reversible polymer-colloid composite assembly in flow," *Nature Communications*, vol. 4, no. 1333, 2013.

[12] D. Dunstan, P. Hamilton-Brown, P. Asimakis, and W. Ducker, "Shear flow promotes amyloid- fibrilization," *Protein engineering, design selection : PEDS*, vol. 22, pp. 741–6, 10 2009.

[13] J. Dobson, A. Kumar, L. F. Willis, R. Toma, D. R. Higazi, R. Turner, D. C. Lowe, A. E. Ashcroft, S. E. Radford, N. Kapur, and D. J. Brockwell, "Inducing protein aggregation by extensional flow," *Proceedings of the National Academy of Sciences*, vol. 114, no. 18, pp. 4673–4678, 2017.

[14] A. Kamada, N. Mittal, L. D. Söderberg, T. Ingverud, W. Ohm, S. V. Roth, F. Lundell, and L. Christofer, "Flow assisted assembly of nanostructured protein microfibers," *Proceedings of the National Academy of Sciences*, vol. 114, no. 6, pp. 1232–1237, 2017.

[15] W. Gospodarczyk and M. Kozak, "Microchip circulation drastically accelerates amyloid aggregation of 1–42 β-amyloid peptide from felis catus," *ACS Chemical Neuroscience*, vol. 8, no. 11, pp. 2558–2567, 2017.

[16] E. Lukas, S. Smith, and T. Scheibel, "Decoding the secrets of spider silk," *materials today*, vol. 14, no. 3, pp. 80–86, 2011.

[17] N. Kojic, J. Bico, C. Clasen, and G. H. McKinley, "Ex vivo rheology of spider silk," *Journal of Experimental Biology*, vol. 209, no. 21, pp. 4355–4362, 2006.

[18] S. Rammensee, U. Slotta, T. Scheibel, and A. Bausch, "Assembly mechanism of recombinant spider silk proteins," *Proceedings of the National Academy of Sciences*, vol. 105, no. 18, pp. 6590–6595, 2008.

[19] S. Lin, S. Ryu, O. Tokareva, G. Gronau, M. M. Jacobsen, W. Huang, D. J. Rizzo, D. Li, C. Staii, N. M. Pugno, J. Y. Wong, D. L. Kaplan, and M. J. Buehler, "Predictive modelling-based design and experiments for synthesis and spinning of bioinspired silk fibres," *Nature communications*, vol. 6, no. 6892, 2015.

[20] M. Andersson, Q. Jia, A. Abella, X. Y. Lee, M. Landreh, P. Purhonen, H. Hebert, M. Tenje, C. Robinson V, Q. Meng, G. Plaza R, J. Johansson, and A. Rising, "Biomimetic spinning of artificial spider silk from a chimeric minispidroin," *Nature Chemical Biology*, vol. 13, pp. 262–264, 2017.

[21] C. Rat, J. C. Heiby, J. P. Bunz, and H. Neuweiler, "Two-step self-assembly of a spider silk molecular clamp," *Nature Communications*, vol. 9, no. 4779, 2018.

[22] D. Onofrei, D. Stengel, D. Jia, H. R. Johnson, S. Trescott, A. Soni, B. Addison, M. Muthukumar, and G. P. Holland, "Investigating the atomic and mesoscale interactions that facilitate spider silk protein pre-assembly," *Biomacromolecules*, vol. 22, no. 8, pp. 3377–3385, 2021.

[23] K. N. Ali D. Malay, Kazuharu Arakawa, "Analysis of repetitive amino acid motifs reveals the essential features of spider dragline silk proteins," *PLoS ONE*, vol. 12, no. 8, p. 025025, 2017.

[24] T. Asakura, "Structure and dynamics of spider silk studied with solid-state nuclear magnetic resonance and molecular dynamics simulation," *Molecules*, vol. 25, no. 11, 2020.

[25] D. Li, M. M. Jacobsen, N. G. Rim, D. Backman, D. L. Kaplan, and J. Y. Wong, "Introducing biomimetic shear and ion gradients to microfluidic spinning improves silk fiber strength," *Biofabrication*, vol. 9, no. 2, p. 025025, 2017.

[26] C. G. Copeland, B. E. Bell, C. D. Christensen, and R. V. Lewis, "Development of a process for the spinning of synthetic spider silk," *ACS Biomaterials Science & Engineering*, vol. 1, no. 7, pp. 577–584, 2015.

[27] E. Oroudjev, J. Soares, S. Arcidiacono, J. B. Thompson, S. A. Fossey, and H. G. Hansma, "Segmented nanofibers of spider dragline silk: Atomic force microscopy and single-molecule force spectroscopy," *Proceedings of the National Academy of Sciences*, vol. 99, no. suppl 2, pp. 6460–6465, 2002.

[28] M. Geisler, S. Xiao, E. M. Puchner, F. Gräter, and T. Hugel, "Controlling the structure of proteins at surfaces," *Journal of the American Chemical Society*, vol. 132, no. 48, pp. 17277–17281, 2010.

[29] W. Hu, "The physics of polymer chain-folding," *Physics Reports*, vol. 747, pp. 1–50, 2018. The physics of polymer chain-folding.

[30] H.-J. Jin and D. L. Kaplan, "Mechanism of silk processing in insects and spiders," *Nature*, vol. 424, pp. 1057–1061, 2003.

[31] L. R. Parent, D. Onofrei, D. Xu, D. Stengel, J. D. Roehling, J. B. Addison, C. Forman, S. A. Amin, B. R. Cherry, J. L. Yarger, N. C. Gianneschi, and G. P. Holland, "Hierarchical spidroin micellar nanoparticles as the fundamental precursors of spider silks," *Proceedings of the National Academy of Sciences*, vol. 115, no. 45, pp. 11507–11512, 2018.

[32] F. Vollrath and D. Knight, "Liquid crystalline spinning of spider silk," *Nature*, vol. 410, pp. 541–8, 04 2001.

[33] J. Sparkes and C. Holland, "Analysis of the pressure requirements for silk spinning reveals a pultrusion dominated process," *Nature Communications*, vol. 8, no. 594, 2017.

[34] R. G. Larson, T. T. Perkins, D. E. Smith, and S. Chu, "Hydrodynamics of a dna molecule in a flow field," *Physical review E*, vol. 55, pp. 1794–1797, 1997.

[35] A. Mohan and P. S. Doyle, "Unraveling of a tethered polymer chain in uniform solvent flow," *Macromolecules*, vol. 40, pp. 4301–4312, 2007.

[36] A. Alexander-Katz, M. Schneider, S. Schneider, A. Wixforth, and R. Netz, "Shear-flow-induced unfolding of polymeric globules," *Physical review letters*, vol. 97, no. 13, p. 138101, 2006.

[37] K. Rack, V. Huck, H. Masoud, F. D. A, S. W. Schneider, and G. Gompper, "Margination and stretching of von willebrand factor in the blood stream enable adhesion," *Scientific Reports*, vol. 7, no. 14278, 2017.

[38] C. E. Sing and A. Alexander-Katz, "Globule-stretch transitions of collapsed polymers in elongational flow," *Macromolecules*, vol. 43, pp. 3532–3541, 2010.

[39] G. Gompper, T. Ihle, D. Kroll, and R. Winkler, "Multi-particle collision dynamics: a particle-based mesoscale simulation approach to the hydrodynamics of complex fluids," in *Advanced computer simulation approaches for soft matter sciences III*, pp. 1–87, Springer, 2009.

[40] A. S. Lemak, J. R. Lepock, and J. Z. Y. Chen, "Molecular dynamics simulations of a protein model in uniform and elongational flows," *Proteins: Structure, Function, and Bioinformatics*, vol. 51, no. 2, pp. 224–235, 2003.

[41] A. Lemak, J. Lepock, and J. Chen, "Unfolding proteins in an external field: Can we always observe the intermediate states?," *Physical review. E*, vol. 67, p. 031910, 2003.

[42] N. Kikuchi, F. J Ryder, C. M. Pooley, and J. M. Yeomans, "Kinetics of the polymer collapse transition: The role of hydrodynamics," *Phys. Rev. E*, vol. 71, no. 061804, 2005.

[43] P. Hoogerbrugge and J. Koelman, "Simulating microscopic hydrodynamic phenomena with dissipative particle dynamics," *EPL (Europhysics Letters)*, vol. 19, no. 3, p. 155, 1992.

[44] T. Krueger, H. Kusumaatmaja, A. Kuzmin, O. Shardt, G. Silva, and E. Viggen, "The lattice boltzmann method principles and practice," 2017.

[45] G. M. Wang and W. C. Sandberg, "Complete all-atom hydrodynamics of protein unfolding in uniform flow," *Nanotechnology*, vol. 21, no. 235101, 2010.

[46] Z. Chen, J. Lou, C. Zhu, and K. Schulten, "Flow-induced structural transition in the β-switch region of glycoprotein ib," *Biophysical journal*, vol. 95, no. 3, pp. 1303–1313, 2008.

[47] G. Reddy and D. Thirumalai, "Dissecting ubiquitin folding using the self-organized polymer model," *The Journal of Physical Chemistry B*, vol. 119, no. 34, pp. 11358–11370, 2015.

[48] H.-R. A. M, M. Vedran, A.-S. Camilo, and F. Gräter, "Molecular dynamics simulations of molecules in uniform flow," *Biophysical Journal*, vol. 116, pp. 1579–1585, 2019.

[49] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindahl, "Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers," *SoftwareX*, vol. 1, pp. 19–25, 2015.

[50] S. Piana, A. G. Donchev, P. Robustelli, and D. E. Shaw, "Water dispersion interactions strongly influence simulated structural properties of disordered protein states," *The Journal of Physical Chemistry*, vol. 119, pp. 5113–5123, 2015.

[51] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror, and D. E. Shaw, "Improved side-chain torsion potentials for the amber ff99sb protein force field," *Proteins: Structure, Function, and Bioinformatics*, vol. 78, no. 8, pp. 1950–1958, 2010.

[52] J. a. Henriques and M. Skepö, "Molecular dynamics simulations of intrinsically disordered proteins: On the accuracy of the tip4p-d water model and the representativeness of protein disorder models," *Journal of Chemical Theory and Computation*, vol. 12, no. 7, p. 3407–3415, 2016.

[53] G. Bussi, D. Donadio, and M. Parrinello, "Canonical sampling through velocity rescaling," *The Journal of Physical Chemistry*, vol. 126, no. 014101, 2007.

[54] M. Parrinello and A. Rahman, "Polymorphic transitions in single crystals: A new molecular dynamics method," *Journal of Applied Physics*, vol. 52, pp. 7182–7190, 1981.

[55] S. Xiao, M. Cetinkaya, B. Markert, and F. Gräter, "Mechanical response of silk crystalline units from force-distribution analysis," *Biophysical Journal*, vol. 96, pp. 3997–4005, 2009.

[56] A. Malevanets and R. Kapral, "Mesoscopic model for solvent dynamics," *The Journal of chemical physics*, vol. 110, no. 17, pp. 8605–8613, 1999.

[57] C.-C. Huang, A. Varghese, G. Gompper, and R. G. Winkler, "Thermostat for nonequilibrium multiparticle-collision-dynamics simulations," *Physical Review E*, vol. 91, no. 1, p. 013310, 2015.

[58] J. P. Huchra and M. J. Geller, "Groups of galaxies. i - nearby groups," *Astrophysical Journal*, vol. 257, no. 1, pp. 224–235, 1982.

[59] P. J. Flory and M. Volkenstein, "Statistical mechanics of chain molecules," *Biopolymers*, vol. 8, no. 5, pp. 699–700, 1969.

[60] O. Kratky and G. Porod, "Röntgenuntersuchung gelöster fadenmoleküle," *Recueil des Travaux Chimiques des Pays-Bas*, vol. 68, no. 12, pp. 1106–1122, 1949.

[61] N. A. Oktaviani, A. Matsugami, A. D. Malay, F. Hayashi, D. L. Kaplan, and K. Numata, "Conformation and dynamics of soluble repetitive domain elucidates the initial beta-sheet conformation of spider silk," *Nature communications*, vol. 9, no. 2121, 2018.

[62] S. Cheng, M. Cetinkaya, and F. Gräter, "How sequence determines elasticity of disordered proteins," *Biophysical journal*, vol. 99, no. 12, pp. 3863–3869, 2010.

[63] A. H. Simmons, C. A. Michal, and L. W. Jelinski, "Molecular orientation and two-component nature of the crystalline fraction of spider dragline silk," *Science*, vol. 271, no. 5245, pp. 84–87, 1996.

[64] A. Rising and J. Johansson, "Toward spinning artificial spider silk," *Nature chemical biology*, vol. 11, pp. 309–315, 05 2015.

[65] S. Keten, Z. Xu, B. Ihle, and M. J. Buehler, "Nanoconfinement controls stiffness, strength and mechanical toughness of β-sheet crystals in silk," *Nature Materials*, vol. 9, pp. 359–367, 2010.

[66] S. Keten and M. J. Buehler, "Nanostructure and molecular mechanics of spider dragline silk protein assemblies," *J.R. Soc. Interface*, vol. 7, no. 53, 2010.

[67] M. F. Hagan and D. Chandler, "Dynamic pathways for viral capsid assembly," *Biophysical journal*, vol. 91, no. 1, pp. 42–54, 2006.

[68] J. E. Baschek, H. C. R Klein, and U. S. Schwarz, "Stochastic dynamics of virus capsid formation: direct versus hierarchical self-assembly," *BMC biophysics*, vol. 5, no. 1, pp. 1–18, 2012.

[69] J. D. Perlmutter and M. F. Hagan, "Mechanisms of virus assembly," *Annual review of physical chemistry*, vol. 66, p. 217, 2015.

[70] D. N. Breslauer, L. P. Lee, and S. J. Muller, "Simulation of flow in the silk gland," *Biomacromolecules*, vol. 10, no. 1, pp. 49–57, 2009.