Biophysical Journal
Article

Biophysical Society

The role of 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, Heidelberg, Germany; [2]BioQuant, Heidelberg University, Heidelberg, Germany; [3]Institute for Theoretical Physics, Heidelberg University, Heidelberg, Germany; [4]Laboratorios de Investigación y Desarrollo, Facultad de Ciencias y Filosofía, Universidad Peruana Cayetano Heredia, Lima, Peru; and [5]Interdisciplinary Center for Scientific Computing, Heidelberg University, Heidelberg, Germany

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 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 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, whereas 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 also for other proteins whose function depends on hydrodynamic flow.

SIGNIFICANCE Spider silk has an extraordinary toughness because it is made from fibers with highly ordered crystalline regions embedded into a soft and disordered matrix. We here used computer simulations with molecular details at two different scales, at the atomistic scale and at a coarser scale, to analyze how silk proteins assemble under flow. A key result is that flow elongates the proteins, which enhances the reversibility of the assembly process, allowing oligomers to grow more slowly and form more crystalline regions compared to no-flow conditions or very low flow rates. Our results can aid the optimization of silk spinning processes and are likely relevant also for other proteins whose biological function relies on hydrodynamic flow.

INTRODUCTION

Water plays a major role in 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 situations in which hydrodynamic flow affects protein structure by inducing nonuniform drag force along the protein. Hydrodynamic flow can mediate a large variety of processes such as material synthesis, blood coagulation, and protein misfolding (5–7). For example, the flow-induced activation of von Willebrand factor (vWf) protein is crucial in hemostasis (8–10). Shear and elongational flows can lead to protein unfolding, misfolding, and aggregation, and can accelerate fibrillation of amyloid-β peptide (11–14). 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 (7,15–21). MA spidroins consist of pH-dependent folded N-terminal and C-terminal

Submitted October 25, 2022, and accepted for publication September 29, 2023.

*Correspondence: schwarz@thphys.uni-heidelberg.de or frauke.graeter@h-its.org

Ana M. Herrera-Rodríguez and Anil Kumar Dasanna contributed equally to this work.

Editor: Jeetain Mittal.

https://doi.org/10.1016/j.bpj.2023.09.020

domains, which experience conformational changes from pH 7.0 down to pH 5.0 in the spider spinning gland, thereby initiating assembly (19). In between these terminal domains, spidroins feature a long repetitive region, which makes up roughly 90% of their total sequence (Fig. 1 *a*) (22). 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 (15,21,23). A large body of work explored the dependence of silk fiber assembly under elongational and shear flows using microfluidics (6,17–19,24,25). In single-molecule force spectroscopy experiments, single spidroin molecules largely followed a worm-like chain behavior but also exhibited unique unfolding steps (23,26–28). However, flow-dependent dynamics of single silk proteins and flow-induced spidroin 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 of whether the spinnability of silk requires the intermediate formation of micelles or not (29,30). Additionally, it could help to better describe liquid crystalline flows inside the s-duct (16,31,32). The lack of a more detailed understanding

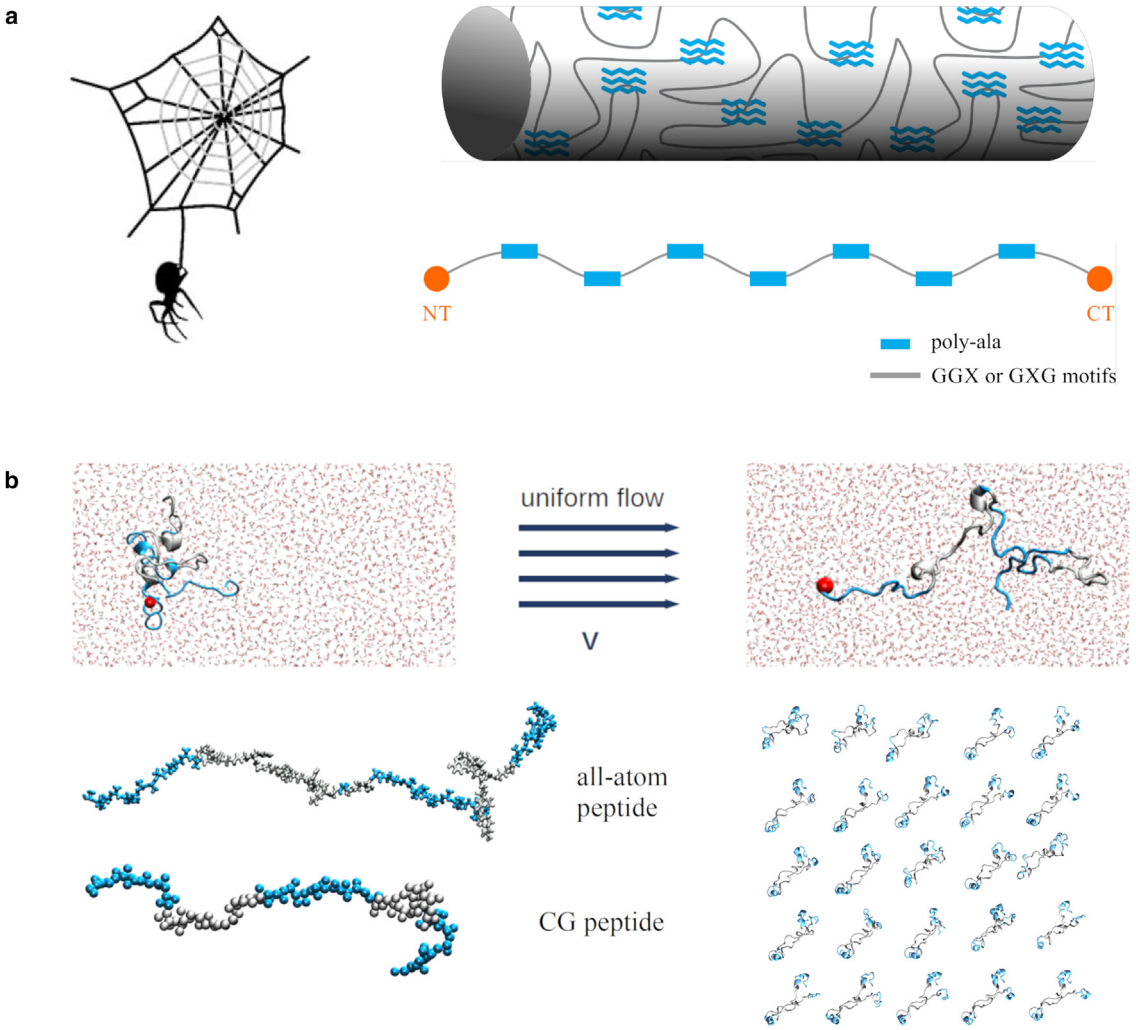

FIGURE 1 Dragline silk and computational model for spidroins under uniform flow. (*a*) Dragline silk corresponds to the structural part of the spider's web frame (top left), fiber microstructure of which consists of embedded β sheet crystals into an amorphous viscoelastic matrix. The fiber is formed upon directed self-assembly induced by flow of the spidroins; a 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, gray). (*b*) The MD model of silk under flow considers all-atom (AA) and coarse-grained (CG) models of a fragment of the repetitive region, containing three poly-Ala and two amorphous regions, resulting in ∼ 80 amino acids in total. Spheres represent atoms (for AA) or backbone and sidechain beads for each amino acid (CG). The spidroin fragment is tethered at the N terminus (red sphere), and a uniform flow is produced to stretch the peptide. To model self-assembly under uniform flow, sets of 25 peptides (bottom right) are tethered along the flow direction and able to move transversely. Color coding as in the spidroin scheme in (*a*). To see this figure in color, go online.

of silk self-assembly is, among other things, due to the challenge of experimentally working with 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 of $\sim 50\%$ w/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 on 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 (33). 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 worm-like chain (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 (34). In addition, bead-spring models have proved useful for describing (bio)polymers, including proteins under flow conditions. A prominent example is the coarse-grained (CG) modeling of vWf multimers in shear flow, with one bead representing roughly one monomer and a Lennard-Jones (LJ) potential to account for average attractive interactions (35,36). 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 (9,10,35,37), multi-particle collision dynamics (MPCD) (38–41), dissipative particle dynamics (18,36,42), or lattice Boltzmann (43). Using MPCD, various studies have explored dynamical and rheological properties of polymer solutions in a shear flow (44,45). In the context of spider silk self-assembly by flow, a mesoscopic modeling using dissipative particle dynamics found that intermediate ratios of hydrophobic and hydrophilic blocks observed in spider silks lead to exceptional silk fiber formation (18). Shear flow was shown to lead to enhancement of β sheet crystals, confirming previous experimental research (6,7,17). 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 (46) or secondary structure transitions in the platelet receptor glycoprotein Ib (47). 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 cost also limits the possibility to study protein assembly processes such as spider silk protein self-assembly; thus, most of the studies are associated with single protein dynamics under flow. The evolution of the secondary structure transition of *Nephila clavipes* MaSp1 dragline silk has been previously studied using steered MD simulations (48), where the single pre-assembled protein chains were subjected to terminal pulling forces and shear stresses were indirectly inferred from these boundary conditions.

In this study, we attack these open questions in a two-pronged approach of simultaneously performing all-atom (AA) and CG simulations at physiologically relevant flow velocities. As a 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-alanine repeats and two amorphous regions (Fig. 1 *a*). We modeled these proteins at the mesoscopic scale as a simple block copolymer using a Go-like potential (49) combined with MPCD (38) to model a uniform flow. For comparison, we performed atomistic MD simulations of the same spidroin fragments in uniform flow of explicit water (50). We find the proteins to largely follow the expected freely jointed chain behavior as known from single-molecule stretching experiments and simulations, although with interesting differences. Flow unravels long-range poly-alanine interactions and promotes β sheet and poly-proline II (PPII) 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 inter-chain contacts than the residues in the amorphous region and this favors the 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.

The reported flow velocities of the simulations (up to $\sim 1\,\mathrm{m/s}$ for AA and up to $\sim 2.6\,\mathrm{m/s}$ for CG) are in the range of both the physiological spinning speeds of dragline threads and the flow rates of dragline spidroins in microfluidic experiments. Dragline threads from *Nephila edulis* spider spun at high reeling speeds of about 0.4 m/s showed less extensibility than threads spun more slowly (51). Jumping spiders (Salticidae) attach silk to a surface and jump to spin high-performance dragline threads at speeds between 0.5 and 0.7 m/s (52). Enhanced β sheet formation was observed using microfluidics and elongational flow (17), with an elongational rate of $1000\,\mathrm{s}^{-1}$ in a microfluidic chamber of a longitudinal size of 1000 μs, which corresponds to a speed of about 1 m/s.

METHODS

In all simulations, we model a representative fragment of silk spidroin, consisting of three poly-alanine repeats separated by two glycine-rich regions,

with a total length of ~80 amino acids. Fig. 1 *b* shows AA and CG representations of the repetitive fragment we simulated. The simulations comprise both single silk peptides, at both AA and CG level, and multiple silk peptides, only at CG level, under uniform flow.

AA MD simulations of single peptides

We used the Gromacs 2018 version (53). The sequence of the peptide originates from the repetitive part of the *E. australis* spidroin at the MA spinning gland: $[A]_{13}$GQGGQGGYGGLGQGGYGQGAGSS$[A]_{14}$GRGQGGYGQG SGGN$[A]_{12}$. 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 (54–56). 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 100,000 steps starting from a randomly chosen conformation of the protein was followed by 5 ns of MD simulations in the NVT (constant number of particles, volume and temperature) ensemble and 10 ns of MD in the NpT (constant number of particles, pressure and temperature) ensemble with the protein with position restraint. Temperature coupling was done by a V-rescale thermostat during equilibration and production runs (57). The Parrinello-Raman barostat (58) 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 (50). 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 hydrogen bonds was used. We simulated three different replicates for every flow velocity, each starting from a different velocity seed.

CG spidroin model

We simulated the same system described in the previous section at CG 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 nonspecific amino acid type for the disordered region. The amount of Ala residues in every region is the same as in AA 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) (59). These "native" interactions were modeled as attractive LJ interactions and all other bead-bead interactions as purely repulsive terms. We now describe all interactions of our spidroin mesoscopic self-organized polymer (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 (49), and is defined as

$$U_{\text{total}} = U_{\text{FENE}} + U_{\text{nat}} + U_{\text{exc}} . \quad (1)$$

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

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

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

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

with

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

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:

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

and

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

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

The parameters associated with the potential energy are for The following:

- U_{FENE} potential: bond length $r_0 = 1.0$ $[a]$, $k = 500$ $[k_BT/a^2]$ and $R_0 = 0.5$ $[a]$.
- U_{nat} potential: $\epsilon_{nat} = 0.25$ $[k_BT]$ 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 $\epsilon_{exc} = 1.0$ $[k_BT]$ and $\sigma_{exc} = r_0$ $[a]$.

MPCD simulations

The CG model of the peptide was combined with hydrodynamic flow using MPCD, which is a particle-based mesoscopic simulation technique for the Navier-Stokes equation. It incorporates hydrodynamic interactions and thermal fluctuations (38,60). We employ a local cell-level scaling thermostat (61) to maintain the temperature of the system constant. A comparison of this thermostat and a Maxwell-Boltzmann thermostat is shown in the Fig. S1, where there is no significant difference in the velocities in all directions. 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 thus follows 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

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

$\vec{\Omega}$ and j correspond to the rotation matrix around a fixed angle $\alpha = \frac{\pi}{2}$ and the cell index respectively. The beads of the peptide are included in the MPCD collision step, 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 from momentum transfer between the particles during the collision step, and kinetic viscosity η_{kin}, which results from the stage of particle streaming. They are given by (60,62)

$$\eta_{kin} \;=\; \frac{k_B T \Delta t_c \rho}{a^3} \times \left(\frac{5\rho}{(4 \;-\; 2\cos(\alpha) \;-\; 2\cos(2\alpha))(\rho \;-\; 1)} - \frac{1}{2}\right), \tag{9}$$

$$\eta_{col} \;=\; \frac{(1 \;-\; \cos(\alpha))(\rho \;-\; 1)}{18 a \Delta t_c}. \tag{10}$$

The total viscosity is $\eta \;=\; \eta_{kin} + \eta_{col}$.

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 $\rho \;=\; 10\, a^{-3}$. 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 $\Delta t_c \;=\; 0.1\sqrt{ma^2/k_B T}$. For the chosen model parameters, the viscosity of the solvent becomes $\eta \;=\; 5.42\sqrt{k_B T a^2/m}$.

The simulations consisted of an equilibration of the system for 12,000 $\sqrt{ma^2/k_B T}$ 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. Furthermore, the first beads of the backbones are fixed in x-direction, allowing them only to diffuse in yz-directions.

CG simulations of silk assembly

To monitor the self-assembly of spider silk proteins, we performed simulations of peptides sets under uniform flow in CG simulations. Every spidroin corresponds to a random conformation taken from the single-peptide simulations in a given flow regime. As in the single silk spidroins simulations, the first bead of the silk peptides is tethered longitudinally. Initially, the peptides are placed equidistant from each other in a matrix of 5×5, where the positions (y, z) of the tether atoms correspond to the nodes of the matrix. Fig. 1 *b* (bottom right) shows the initial configuration of the peptides. The peptides are allowed to move transversely (in the yz plane) and are tethered at their starting position by a harmonic restraint along the flow direction x. We note that this simplified setup is tailored toward monitoring assembly at the nanoscale in which physiological shear rates are negligible, and in this regard different from microfluidic experiments in which assembly was monitored of nontethered chains in elongational flow.

We used the same protocol and simulation parameters as in the single peptide case (All-atom MD simulations of single peptides) 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 $\sim 4\%$ w/v for the CG system. In total, we performed six simulations with three replicates each. A description of each simulation is contained in Table 1. The initial mean fractional extension in Table 1 corresponds to the mean of the extensions of the initial conformation of the peptides. The spidroins are initially subjected to position restraint while the flow is being equilibrated during an MPCD time of 625 $\sqrt{(ma^2/k_B T)}$.

Persistence and Kuhn lengths

To obtain the Kuhn length of the polymer at AA and CG 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 AA and CG simulations was 0.2 nm and 1.6 a respectively. With $l_k \;=\; 2l_p$ we obtained the Kuhn length for both models.

TABLE 1 Description of CG Simulations

Mean flow velocity ($\sqrt{k_B T/m}$)	Initial mean fractional extension of peptides ($t \;=\; 0$) L	Mean fractional extension of peptides L	Simulation time ($\sqrt{ma^2/k_B T}$)	Number of replicates	Number of peptides
0.0005	0.4	0.05	800	3	25
0.0009	0.7	0.052	700	3	25
0.002	0.8	0.06	600	3	25
0.004	0.9	0.07	600	3	25
0.007	0.4	0.34	500	3	25
0.011	0.7	0.56	300	3	25
0.012	0.8	0.68	200	3	25
0.014	0.9	0.76	200	3	25

Drag force calculation

We calculated the drag force acting on the protein in AA 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 CG 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 $-180 \leq \varphi \leq -90$ (circles), and in a PPII conformation if $128 \leq \psi \leq 180$ and $-90 \leq \varphi \leq -58$ (squares). To observe the relevance of alanines in the formation of β sheets, we show the percentage of Ala and non-Ala residues as $\%Ala_{\beta/PPII} = N_{Ala_{\beta/PPII}} * 100/N_{Ala}$ (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, whereas $N_{nonAla_{\beta/PPII}}$ 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 a certain size or not, we adapted a group selection algorithm called friend of friends to our case (63). 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 $d^{12}_{C_\alpha} \leq 0.5$ nm (see the spidroin inter-contact sketch at Fig. S2). 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.

RESULTS

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 AA case, we computed the vector sum of the nonbonded forces between all atoms of the water molecules and of the peptide over time. In the CG 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 AA and CG simulations, the force components along the flow yield a nonzero average force, the drag force. We averaged the drag force over the last half of each trajectory and over all three replicates. Interestingly, regardless of the resolution of the simulations, i.e., for both AA and CG, we obtain an overall linear increase in drag force with flow velocity (Fig. 2 *a* and *b*). Thus, the drag force effectively follows Stokes' law, with

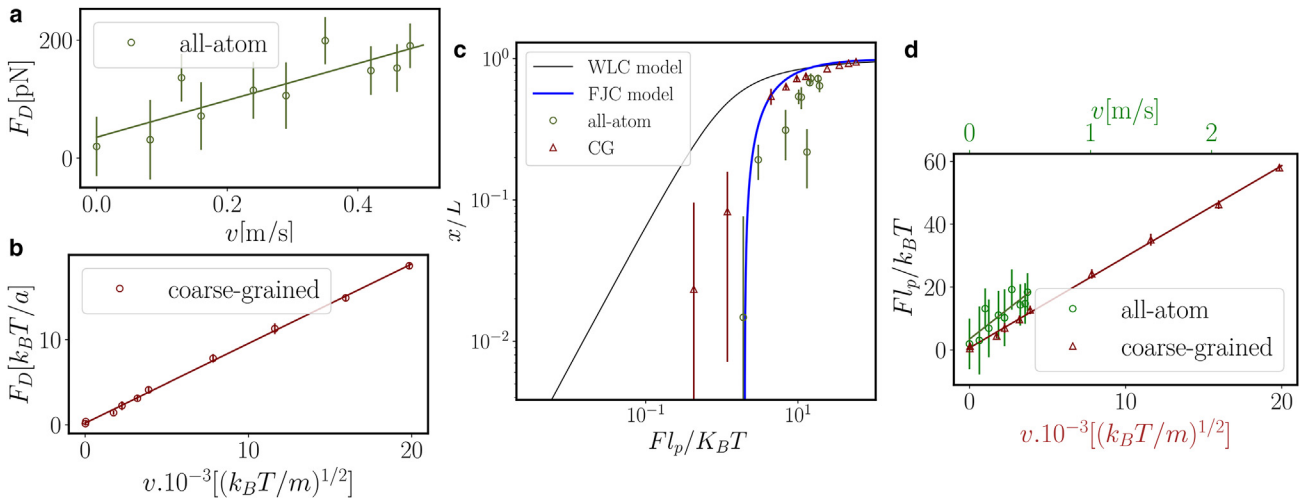

FIGURE 2 Stretching silk peptides under uniform flow. (*a*) Averaged peptide drag force over mean flow velocity for AA simulations. The drag force in the peptide originates from the nonbonded interactions between the water and protein atoms. The green solid curve is a linear fit. (*b*) Averaged peptide drag force for CG 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 AA and CG simulations, respectively. The persistence length was extracted from fitting to Eq. 10. 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 SEs 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 AA and CG 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 see this figure in color, go online.

an effective hydrodynamic radius of 2.52 nm (obtained from the linear fit to the AA data of Fig. 2 *a*). 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) (39,40), respectively. However, a gradual stretching through multiple unfolding intermediates was observed for ubiquitin in the case of elongational flow (39,40) and proposed recently for flowing vWF (64), directly in line with the gradual extension of the spidroin with increasing flow velocities.

To quantitatively compare the AA and CG scales, Fig. 2 *c* 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 and 0.4 nm for AA simulations as well as 80 *a* and 3.1 *a* for CG simulations. We recover a nonlinear force-extension curve typical for polymers when stretched by a pulling force. The AA model yields overall lower extensions for a given drag force, but, given that they are fully independently parameterized, the agreement is satisfying. The CG model extends more readily, which is likely due to our approximation that only attractive interactions are present within alanine residues, neglecting unspecific favorable interactions involving nonalanine residues of the disordered region. The force-extension behavior is reminiscent of the behavior predicted by the freely jointed chain (FJC) (65) or WLC model (66) of polymers. Indeed, we find a good agreement of our data with the FJC model (solid blue curve Fig. 2 *c*). We find slightly less agreement between the WLC model and our particle-based simulations (Fig. 2 *c*, solid black line), which we explain by the nature of the protein backbone and the (albeit only partially) rotatable bonds, which are better depicted by an 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 AA and CG simulations. Interestingly, this range covers the rupture forces of $\sim 176 \pm 73$ pN required to stretch single-dragline silk molecules by an atomic force microscope (26). To estimate the velocities achieved in the CG model, Fig. 2 *d* shows the normalized drag force as a function of velocity for AA and CG models; the velocities in the AA model are re-scaled by a factor $\alpha = 7.75$. The highest flow velocity in the CG 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 AA (Fig. 3 *a*) and CG simulations (Fig. 3 *b*). In the AA case, we observe a drop of Ala-Ala contacts between

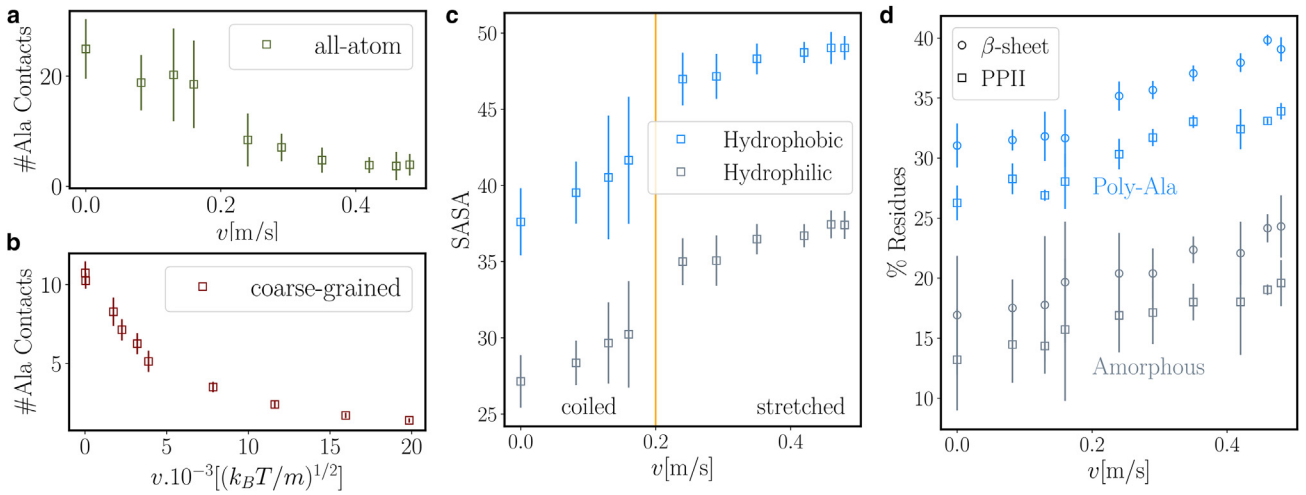

FIGURE 3 Molecular details during silk elongation. (*a*) Number of contacts of Ala residues as a function of time in AA MD simulations. The cutoff radius within which we examine a contact formation is 0.5 nm between $C_\alpha - C_\alpha$ atoms of the peptide. (*b*) Number of contacts of Ala residues as a function of time in CG MD simulations. The cutoff radius within which we examine a contact formation is 0.5 nm between $C_\alpha - C_\alpha$ atoms of the peptide. (*c*) Hydrophobic and hydrophilic solvent-accessible surface areas over flow rate for AA simulations. At about 0.2 m/s, a transition of the peptide to an extended state is observed (solid vertical orange line). The highest velocity of the atomistic simulations is about one-fifth the highest velocity of the CG 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 AA simulations. The data are 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 SE of the mean. See section "methods" for the definition of the dihedral intervals for both secondary structures. To see this figure in color, go online.

0.16 and 0.24 m/s of flow rate. Also, the fractional extension of the AA peptide as a function of the flow rate confirms this behavior (Fig. S3 *A*): 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.

In the case of the CG 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 $20 \times 10^{-3}(k_BT/m)^{1/2}$ (Fig. 3 *b*). The number of total and amorphous contacts over mean flow rate in Fig. 3 *b* shows the same tendency. With regard to the CG fractional extension (Fig. S3 *B*) 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 AA. In addition, the CG model shows fewer contacts initially, i.e., at low flow velocities ($\sim$12 instead of $\sim$ 24 for AA). We conclude that the CG model reproduces the overall loss of contacts in flow, although it cannot resolve the collapse-stretch transition of AA simulations.

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

The formation of β sheets is a key factor in the self-assembly of spider silk and requires flow (17). However, how flow induces β sheet formation is poorly understood. In recent studies, a significant population, namely $\sim$ 24%, of PPII helix in the repetitive region of MA spidroin dopes, i.e., before flow-induced elongation during spinning, has been observed (67). It has been suggested that PPII helices in *N. clavipes* can form a rigid structure that can be quickly transformed into β sheets and having higher flexibility in the glycine-rich region of *E. australis* makes this domain prone to aggregation. 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 section "methods" for details). We averaged over all alanines and all amorphous-phase residues (nonalanines) and normalized by the total number of alanines and nonalanines, respectively. We analyzed the secondary structure propensities only for the AA simulations, as the CG lacks the required resolution.

Fig. 3 *d* 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 $\varphi = -180$. 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 (68). 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. These higher ratios are in line with the view that the poly-alanine repeats preferentially drive β sheet formation, i.e., crystallization, during fiber formation. Furthermore, the poly-Ala regions show rather extended backbone conformations at a higher flow (> 70% PPII and β sheet), whereas the amorphous fragments only exhibit a minor fraction of residues in these conformations (< 50% PPII and β sheet). We have measured explicitly the rate of change by fitting the data to a straight line. We obtained a rate $rate_{ALA} = 36$ [% s/m] for the poly-Ala region and $rate_{nonALA} = 28$ [% s/m] for the amorphous one.

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. Looking into more detail at the PPII content, in the absence of flow, there are on average 10 ALA residues and six amorphous residues in the PPII conformation. This suggests that, before the initiation of flow, the PPII content is approximately 22% (see Fig. 3 *d*), a value comparable with the experimental estimate (67). This comparison suggests that the used force field predicts to some extent the proper amount of poly-proline content. Considerations of other force fields, for instance, the recent ones that reproduce well both disordered and folded regions (69,70), is certainly very important. However, it falls outside the scope of the present study. Taking into account that the transition from PPII helix to β 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 predominant 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% β sheet content by alanines predicted in the model of Simmons et al. (71), which is based in nuclear magnetic resonance experimental studies. The occurrence 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 Oktaviani et al. (67).

Silk spidroin assembly under flow

Having established a CG 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 CG model and monitored their dynamics, interactions, and oligomerization at a range of flow velocities. The simulation setup is explained in the section "methods." AA 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. Voids can appear at the boundary of the slice within which solvent particles are subjected to additional forces (50). For all simulations included in this study, both AA simulations of single-protein chains as well as MPCD simulations of single and multiple protein chains, density remained constant throughout the systems.

As in the simulations of single chains, the silk fragment comprises ∼80 residues; that is, three poly-Ala repeats and two intrinsically disordered regions. An initial pre-alignment of the peptides was performed, mimicking the multimerization at the N-terminal domains (Fig. 1). The pre-alignment is maintained by tethering the peptides along the flow direction at the same position, whereas they can move transversely. Imposing this pre-alignment will enhance the packing of the chains (via the 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 Fig. S3. We covered mean fractional extensions in the range of $[0.0 - 0.76]L$. In fact, the lower four out of the total eight 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 led to measurable expansions of the chain (compare Fig. 2 *c*). 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 before inter-chain aggregation, which further reduced their extensions drastically.

To analyze the self-assembly of the spidroins over time, we monitored when a certain peptide takes part in an oligomer. This was done through a clustering algorithm called friends of friends (see section "methods"). The criterion to form an oligomer of a certain size is based on the formation of a single inter-chain $C_\alpha - C_\alpha$ interaction. In Fig. 4 *a*, we visualize the oligomerization process. Data are 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 timescales. At a lower flow rate, with mean peptide fractional extensions close to $0.0L$, oligomers form at shorter timescales and reach larger sizes (Fig. S7). 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 (Fig. 4 *b*). Average oligomer sizes for individual replicates, exemplarily for two different flows, are also shown in Fig. S7). Overall, self-assembly of silk peptides slows down with increasing flow velocity (simulation snapshots for two different flows are shown in Fig. S7). Slower flow causes a higher mean oligomerization at a given time (light blue, purple, green, and

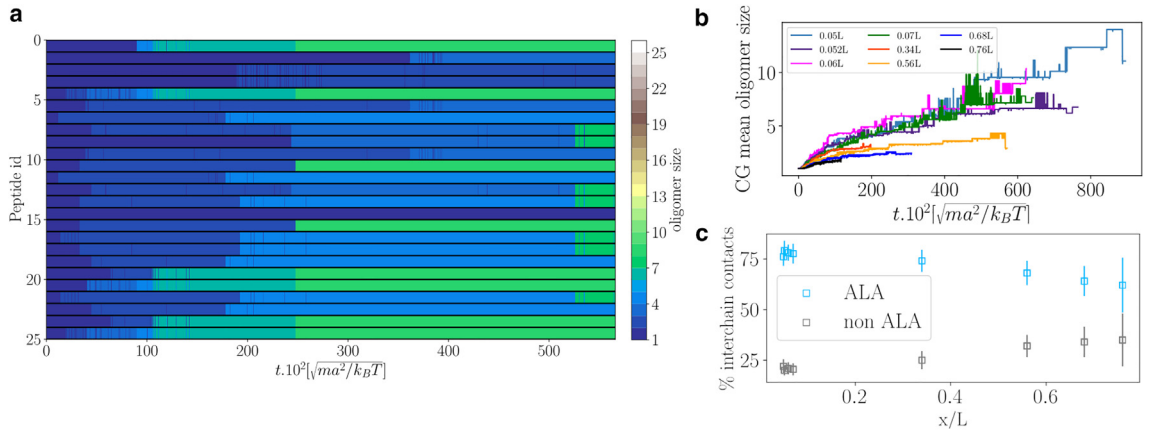

FIGURE 4 Silk oligomerization. (*a*) Oligomer formation over time for every peptide. The color code indicates the size of the oligomer they form. Shown is the oligomerization for one replicate at a mean fractional extension of 0.56L. (*b*) Mean oligomer size as a function of time. (*c*) Average percentage of interchain contacts for the poly-alanine and amorphous regions over the mean CG peptide fractional extension x/L. To see this figure in color, go online.

magenta curves in Fig. 4 *b*) compared to faster flow regimes (red, orange, dark blue, and black curves in Fig. 4 *b*). We attribute this slowdown in oligomerization to a reduction of the fluctuations of the protein chains orthogonal to the flow direction at high flow rates (flow rates larger than 1 m/s or polymer extensions larger than 0.4*L*), which reduces the lateral diffusion of the peptides and in fact the likelihood of lateral diffusional encounters. However, the inter-chain contact formation and controlled assembly into β sheet is enhanced due to the flattening and alignment of the chains along the flow (see Fig. S8). This enhancement can be observed by comparing Fig. 4 *b* and *c*: the mean oligomerization size is lower at higher rates but the non-ALA inter-chain contact formation increases with the flow, and the number of ALA contacts is not drastically reduced despite a lower oligomerization because of a more favorable interaction between the inter-chain residues. A similar behavior has been observed for one-end tethered polymers on surfaces under uniform flow, where the adsorption of the polymers on the surface was enhanced due to the flattening of the polymer chain along the flow (72). 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 nonalanine contacts due to the fact that chains are aligned in the oligomers (Fig. 4 *c*).

To investigate the oligomerization kinetics, we computed association and dissociation time intervals at every flow regime. Fig. 5 *a* and *b* 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. 3 *b*). Interestingly, both the numbers of association and dissociation events increase with the 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. Fig. 5 *c* shows transversal (*yz*) and longitudinal (*yx*) views of multimers at two flow rates corresponding to mean fractional extensions of 0.05*L* and 0.76*L*. One can observe a higher chain alignment at high flow rates, whereas the aggregates are micelle-like at low flow rates.

DISCUSSION

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 flow on two scales: atomistic (AA) and amino acid (CG) resolution. In this way, we could consider atomistic details from AA

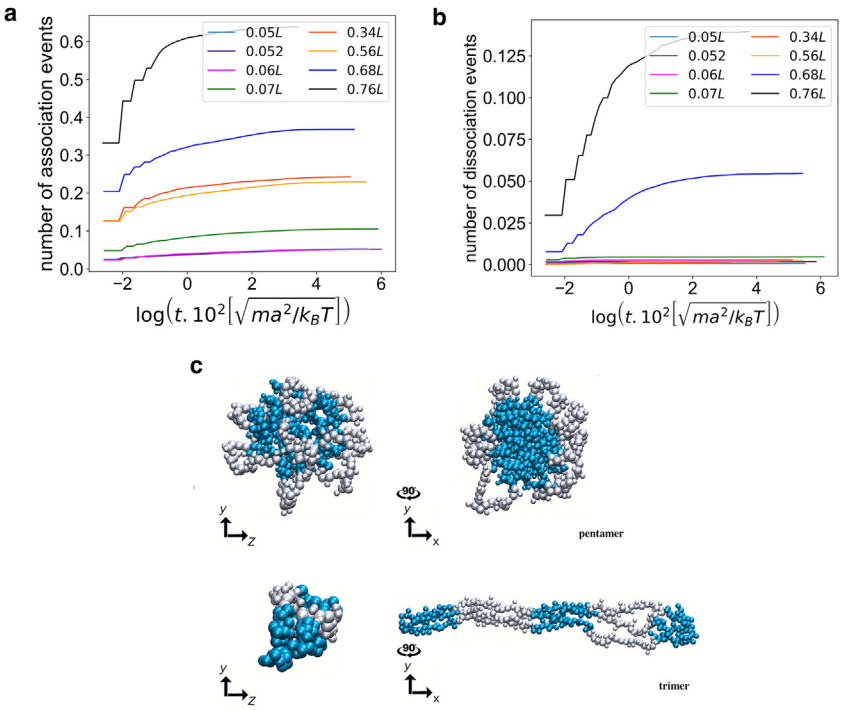

FIGURE 5 Oligomerization kinetics. Log-normal cumulative distributions of the time intervals of (*a*) association and (*b*) dissociation events, for every flow velocity (or every average protein extension). The distributions are based on data from three replicates at every flow velocity and are normalized 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. (*c*) Assembly snapshots at mean fractional extensions 0.05*L* (upper part) and 0.76*L* (lower part). They correspond to the last time step of one of the replicates simulated at the respective flow. Transversal (*yz*) and longitudinal (*yx*) views are being shown for a pentamer at 0.05*L* (low flow velocity) and a trimer at 0.76*L* (high flow velocity). Alanine regions are in light blue and amorphous regions are in gray. To see this figure in color, go online.

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 analyze in detail the structural changes of spidroins with flow (Fig. 3 *c* and *b*), whereas 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 (26). 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 in 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 β sheet conformations with the flow, increasing their propensity for fibrillation and the formation of β sheet crystals as found in silk fibers (59).

Due to the high computational cost and the limited flow velocity that can be achieved in AA MD simulations (50), we only simulated CG multiple spidroins to monitor their assembly. Silk fiber formation is a highly complex and multiscale process occurring 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 (20), and the NT domain dimerizes inside the first part of the s-duct after a drop of pH (73). 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 (Fig. 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 on 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 (71,74,75). CG simulations can extend the limited timescales 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 β 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, such as virus capsids, which are stalled in kinetic traps if the free energy gain is too high (76–78). For the case of silk discussed here, moderate amounts of shear, but not too-high values, increase the quality of the assembly, allowing proper β 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 AA simulations, correspond to flows in the range of 0.5–3 m/s (Fig. 2 *c*). These speeds are relevant, albeit at the upper boundary, to those used in a laboratory setting or observed within the spider gland (79)

CONCLUSION

Overall, we have investigated by computer simulations the effect of flows on spider silk peptides. Our study gives insights into their flow-induced nonequilibrium conformational dynamics and how flow influences their self-assembly. Microfluidic experiments have shown how elongational flow enhances β sheet formation (17), which we here can confirm. However, due to the high hydrophobicity of the silk dope, 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 test the mechanisms of flow-induced β sheet formation that are suggested by our computational results.

SUPPORTING MATERIAL

Supporting material can be found online at https://doi.org/10.1016/j.bpj.2023.09.020.

AUTHOR CONTRIBUTIONS

F.G. and U.S.S. designed the project. A.M.H.R., A.K.D., C.D., E.R.C.C., and C.A.S. developed the methods and set up the simulations. A.M.H.R. and A.K.D. carried out the simulations. A.M.H.R., A.K.D., and C.A.S analyzed the data. A.M.H.R., A.K.D., C.A.S., U.S.S., and F.G. wrote the paper.

ACKNOWLEDGMENTS

This work was supported by the Klaus Tschira Foundation. We also acknowledge funding through the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy (2082/1 – 390761711) (cluster of excellence 3DMM2O) and through research grant GR 3494/7-3. This research was conducted within the Max Planck School Matter to Life supported by the German Federal Ministry of Education and Research (BMBF) in collaboration with the Max Planck Society. Computing time was allocated by the state of Baden-Württemberg through bwHPC and DFG through grant INST 35/1134-1 FUGG.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Huang, W., S. Krishnaji, …, P. Cebe. 2014. Influence of Water on Protein Transitions: Thermal Analysis. *Macromolecules.* 47:8098–8106.
2. Huang, W., S. Krishnaji, …, P. Cebe. 2014. Influence of Water on Protein Transitions: Morphology and Secondary Structure. *Macromolecules.* 47:8107–8114.
3. Kim, Y., H. Choi, …, S. Na. 2020. Spider silk with weaker bonding resulting in higher strength and toughness through progressive unfolding and load transfer. *J. Mech. Behav. Biomed. Mater.* 108, 103773.
4. Boroudjerdi, H., Y. Kim, …, A. Serr. 2005. Statics and dynamics of strongly charged soft matter. *Phys. Rep.* 416:129–199.
5. Szymczak, P., and M. Cieplak. 2011. Hydrodynamics effects in proteins. *J. Phys. Condens. Matter.* 23, 033102.
6. Lu, L., S. Fan, …, Y. Zhang. 2021. Flow Analysis of Regenerated Silk Fibroin/Cellulose Nanofiber Suspensions via a Bioinspired Microfluidic Chip. *Adv. Mater. Technol.* 6, 2100124.
7. Han, Y., M. Langer, …, P. Král. 2020. Biomimetic Materials: Stretch-Healable Molecular Nanofibers. *Advanced Theory and Simulations.* 3, 2070023.
8. Baldauf, C., R. Schneppenheim, …, F. Gräter. 2009. Shear-induced unfolding activates von Willebrand factor A2 domain for proteolysis. *J. Thromb. Haemostasis.* 7:2096–2105.
9. Schneider, S. W., S. Nuschele, …, M. F. Schneider. 2007. Shear-induced unfolding triggers adhesion of von Willebrand factor fibers. *Proc. Natl. Acad. Sci. USA.* 104:7899–7903.
10. Chen, H., M. A. Fallah, …, A. Alexander-Katz. 2013. Blood-clotting-inspired reversible polymer-colloid composite assembly in flow. *Nat. Commun.* 4:1333.
11. Dunstan, D. E., P. Hamilton-Brown, …, J. Bertolini. 2009. Shear flow promotes amyloid- fibrilization. *Protein Eng. Des. Sel.* 22:741–746.
12. Dobson, J., A. Kumar, …, D. J. Brockwell. 2017. Inducing protein aggregation by extensional flow. *Proc. Natl. Acad. Sci. USA.* 114:4673–4678.
13. Kamada, A., N. Mittal, …, C. Lendel. 2017. Flow assisted assembly of nanostructured protein microfibers. *Proc. Natl. Acad. Sci. USA.* 114:1232–1237.
14. Gospodarczyk, W., and M. Kozak. 2017. Microchip Circulation Drastically Accelerates Amyloid Aggregation of 1–42 β-amyloid Peptide from Felis catus. *ACS Chem. Neurosci.* 8:2558–2567.
15. Eisoldt, L., A. Smith, and T. Scheibel. 2011. Decoding the secrets of spider silk. *Mater. Today.* 14:80–86.
16. Kojić, N., J. Bico, …, G. H. McKinley. 2006. Ex vivo rheology of spider silk. *J. Exp. Biol.* 209:4355–4362.
17. Rammensee, S., U. Slotta, …, A. R. Bausch. 2008. Assembly mechanism of recombinant spider silk proteins. *Proc. Natl. Acad. Sci. USA.* 105:6590–6595.
18. Lin, S., S. Ryu, …, M. J. Buehler. 2015. Predictive modelling-based design and experiments for synthesis and spinning of bioinspired silk fibres. *Nat. Commun.* 6:6892.
19. Andersson, M., Q. Jia, …, A. Rising. 2017. Biomimetic spinning of artificial spider silk from a chimeric minispidroin. *Nat. Chem. Biol.* 13:262–264.
20. Rat, C., J. C. Heiby, …, H. Neuweiler. 2018. Two-step self-assembly of a spider silk molecular clamp. *Nat. Commun.* 9:4779.
21. Onofrei, D., D. Stengel, …, G. P. Holland. 2021. Investigating the Atomic and Mesoscale Interactions that Facilitate Spider Silk Protein Pre-Assembly. *Biomacromolecules.* 22:3377–3385.
22. Malay, A. D., K. Arakawa, and K. Numata. 2017. Analysis of repetitive amino acid motifs reveals the essential features of spider dragline silk proteins. *PLoS One.* 12, e0183397.
23. Asakura, T. 2020. Structure and Dynamics of Spider Silk Studied with Solid-State Nuclear Magnetic Resonance and Molecular Dynamics Simulation. *Molecules.* 25.
24. Li, D., M. M. Jacobsen, …, J. Y. Wong. 2017. Introducing biomimetic shear and ion gradients to microfluidic spinning improves silk fiber strength. *Biofabrication.* 9, 025025.
25. Copeland, C. G., B. E. Bell, …, R. V. Lewis. 2015. Development of a Process for the Spinning of Synthetic Spider Silk. *ACS Biomater. Sci. Eng.* 1:577–584.
26. Oroudjev, E., J. Soares, …, H. G. Hansma. 2002. Segmented nanofibers of spider dragline silk: Atomic force microscopy and single-molecule force spectroscopy. *Proc. Natl. Acad. Sci. USA.* 99:6460–6465.
27. Geisler, M., S. Xiao, …, T. Hugel. 2010. Controlling the Structure of Proteins at Surfaces. *J. Am. Chem. Soc.* 132:17277–17281.
28. Hu, W. 2018. The physics of polymer chain-folding. *Phys. Rep.* 747:1–50.
29. Jin, H.-J., and D. L. Kaplan. 2003. Mechanism of silk processing in insects and spiders. *Nature.* 424:1057–1061.
30. Parent, L. R., D. Onofrei, …, G. P. Holland. 2018. Hierarchical spidroin micellar nanoparticles as the fundamental precursors of spider silks. *Proc. Natl. Acad. Sci. USA.* 115:11507–11512.
31. Vollrath, F., and D. P. Knight. 2001. Liquid crystalline spinning of spider silk. *Nature.* 410:541–548.
32. Sparkes, J., and C. Holland. 2017. Analysis of the pressure requirements for silk spinning reveals a pultrusion dominated process. *Nat. Commun.* 8:594.
33. Larson, R. G., T. T. Perkins, …, S. Chu. 1997. Hydrodynamics of a DNA molecule in a flow field. *Phys. Rev. E.* 55:1794–1797.
34. Mohan, A., and P. S. Doyle. 2007. Unraveling of a tethered polymer chain in uniform solvent flow. *Macromolecules.* 40:4301–4312.
35. Alexander-Katz, A., M. F. Schneider, …, R. R. Netz. 2006. Shear-flow-induced unfolding of polymeric globules. *Phys. Rev. Lett.* 97, 138101.
36. Rack, K., V. Huck, …, G. Gompper. 2017. Margination and stretching of von Willebrand factor in the blood stream enable adhesion. *Sci. Rep.* 7:14278.
37. Sing, C. E., and A. Alexander-Katz. 2010. Globule-stretch transitions of collapsed polymers in elongational flow. *Macromolecules.* 43:3532–3541.
38. Gompper, G., T. Ihle, …, R. Winkler. 2009. 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 Springer, pp. 1–87.

39. Lemak, A. S., J. R. Lepock, and J. Z. Y. Chen. 2003. Molecular dynamics simulations of a protein model in uniform and elongational flows. *Proteins.* 51:224–235.
40. Lemak, A. S., J. R. Lepock, and J. Z. Y. Chen. 2003. Unfolding proteins in an external field: Can we always observe the intermediate states? *Phys. Rev.* 67, 031910.
41. Kikuchi, N., J. F. Ryder, …, J. M. Yeomans. 2005. Kinetics of the polymer collapse transition: The role of hydrodynamics. *Phys. Rev. E.* 71:061804.
42. Hoogerbrugge, P. J., and J. M. V. A. Koelman. 1992. Simulating Microscopic Hydrodynamic Phenomena with Dissipative Particle Dynamics. *EPL.* 19:155–160.
43. Krueger, T., H. Kusumaatmaja, …, E. Viggen. 2017. The Lattice Boltzmann Method Principles and Practice.
44. Winkler, R. G. 2010. Conformational and rheological properties of semiflexible polymers in shear flow. *J. Chem. Phys.* 133, 164905.
45. Huang, C.-C., R. G. Winkler, …, G. Gompper. 2010. Semidilute polymer solutions at equilibrium and under shear flow. *Macromolecules.* 43:10107–10116.
46. Wang, G. M., and W. C. Sandberg. 2010. Complete all-atom hydrodynamics of protein unfolding in uniform flow. *Nanotechnology.* 21:235101.
47. Chen, Z., J. Lou, …, K. Schulten. 2008. Flow-induced structural transition in the β-switch region of glycoprotein Ib. *Biophys. J.* 95:1303–1313.
48. Giesa, T., C. C. Perry, and M. J. Buehler. 2016. Secondary Structure Transition and Critical Stress for a Model of Spider Silk Assembly. *Biomacromolecules.* 17:427–436.
49. Reddy, G., and D. Thirumalai. 2015. Dissecting Ubiquitin Folding Using the Self-Organized Polymer Model. *J. Phys. Chem. B.* 119:11358–11370.
50. Herrera-Rodríguez, A. M., V. Miletić, …, F. Gräter. 2019. Molecular Dynamics Simulations of Molecules in Uniform Flow. *Biophys. J.* 116:1579–1585.
51. Vollrath, F., B. Madsen, and Z. Shao. 2001. The effect of spinning conditions on the mechanics of a spider's dragline silk. *Proc. Biol. Sci.* 268:2339–2346.
52. Chen, A., K. Kim, and P. S. Shamble. 2021. Rapid mid-jump production of high-performance silk by jumping spiders. *Curr. Biol.* 31:R1422–R1423. https://www.sciencedirect.com/science/article/pii/S0960982221012938.
53. Abraham, M. J., T. Murtola, …, E. Lindahl. 2015. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX.* 1–2:19–25.
54. Piana, S., A. G. Donchev, …, D. E. Shaw. 2015. Water Dispersion Interactions Strongly Influence Simulated Structural Properties of Disordered Protein States. *J. Phys. Chem. B.* 119:5113–5123.
55. Lindorff-Larsen, K., S. Piana, …, D. E. Shaw. 2010. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins.* 78:1950–1958.
56. Henriques, J., and M. Skepö. 2016. Molecular Dynamics Simulations of Intrinsically Disordered Proteins: On the Accuracy of the TIP4P-D Water Model and the Representativeness of Protein Disorder Models. *J. Chem. Theor. Comput.* 12:3407–3415.
57. Bussi, G., D. Donadio, and M. Parrinello. 2007. Canonical sampling through velocity rescaling. *J. Chem. Phys.* 126:014101.
58. Parrinello, M., and A. Rahman. 1981. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* 52:7182–7190.
59. Xiao, S., W. Stacklies, …, F. Gräter. 2009. Mechanical response of silk crystalline units from force-distribution analysis. *Biophys. J.* 96:3997–4005.
60. Malevanets, A., and R. Kapral. 1999. Mesoscopic model for solvent dynamics. *J. Chem. Phys.* 110:8605–8613.
61. Huang, C.-C., A. Varghese, …, R. G. Winkler. 2015. Thermostat for nonequilibrium multiparticle-collision-dynamics simulations. *Phys. Rev.* 91, 013310.
62. Kikuchi, N., C. M. Pooley, …, J. M. Yeomans. 2003. Transport coefficients of a mesoscopic fluid dynamics model. *J. Chem. Phys.* 119:6388–6395.
63. Huchra, J. P., and M. J. Geller. 1982. Groups of galaxies. I - Nearby groups. *Astrophys. J.* 257:423.
64. Bergal, H. T., Y. Jiang, …, W. P. Wong. 2022. Conformation of von Willebrand factor in shear flow revealed with stroboscopic single-molecule imaging. *Blood.* 140:2490–2499.
65. Flory, P. J., and M. Volkenstein. 1969. Statistical mechanics of chain molecules. *Biopolymers.* 8:699–700.
66. Kratky, O., and G. Porod. 1949. Röntgenuntersuchung gelöster Fadenmoleküle. *Recl. Trav. Chim. Pays. Bas.* 68:1106–1122.
67. Oktaviani, N. A., A. Matsugami, …, K. Numata. 2018. Conformation and dynamics of soluble repetitive domain elucidates the initial beta-sheet conformation of spider silk. *Nat. Commun.* 9:2121.
68. Cheng, S., M. Cetinkaya, and F. Gräter. 2010. How sequence determines elasticity of disordered proteins. *Biophys. J.* 99:3863–3869.
69. Robustelli, P., S. Piana, and D. E. Shaw. 2018. Developing a molecular dynamics force field for both folded and disordered protein states. *Proc. Natl. Acad. Sci. USA.* 115:E4758–E4766.
70. Piana, S., P. Robustelli, …, D. E. Shaw. 2020. Development of a Force Field for the Simulation of Single-Chain Proteins and Protein–Protein Complexes. *J. Chem. Theor. Comput.* 16:2494–2507.
71. Simmons, A. H., C. A. Michal, and L. W. Jelinski. 1996. Molecular Orientation and Two-Component Nature of the Crystalline Fraction of Spider Dragline Silk. *Science.* 271:84–87.
72. Serr, A., and R. R. Netz. 2007. Enhancing polymer adsorption by lateral pulling. *Europhys. Lett.* 78, 68006.
73. Rising, A., and J. Johansson. 2015. Toward spinning artificial spider silk. *Nat. Chem. Biol.* 11:309–315.
74. Keten, S., Z. Xu, …, M. J. Buehler. 2010. Nanoconfinement controls stiffness, strength and mechanical toughness of β-sheet crystals in silk. *Nat. Mater.* 9:359–367.
75. Keten, S., and M. J. Buehler. 2010. Nanostructure and molecular mechanics of spider dragline silk protein assemblies. *J. R. Soc. Interface.* 7:1709–1721.
76. Hagan, M. F., and D. Chandler. 2006. Dynamic pathways for viral capsid assembly. *Biophys. J.* 91:42–54.
77. Baschek, J. E., H. C. R Klein, and U. S. Schwarz. 2012. Stochastic dynamics of virus capsid formation: direct versus hierarchical self-assembly. *BMC Biophys.* 5:22.
78. Perlmutter, J. D., and M. F. Hagan. 2015. Mechanisms of virus assembly. *Annu. Rev. Phys. Chem.* 66:217–239.
79. Breslauer, D. N., L. P. Lee, and S. J. Muller. 2009. Simulation of Flow in the Silk Gland. *Biomacromolecules.* 10:49–57.