

# Bi-terminal fusion of intrinsically-disordered mussel foot protein fragments boosts mechanical strength for protein fibers

Received: 27 September 2022

Accepted: 22 March 2023

Published online: 14 April 2023

 Check for updatesJingyao Li<sup>1</sup>, Bojing Jiang<sup>1</sup>, Xinyuan Chang<sup>1</sup>, Han Yu<sup>1</sup>, Yichao Han<sup>1</sup> & Fuzhong Zhang<sup>1,2,3</sup>✉

Microbially-synthesized protein-based materials are attractive replacements for petroleum-derived synthetic polymers. However, the high molecular weight, high repetitiveness, and highly-biased amino acid composition of high-performance protein-based materials have restricted their production and widespread use. Here we present a general strategy for enhancing both strength and toughness of low-molecular-weight protein-based materials by fusing intrinsically-disordered mussel foot protein fragments to their termini, thereby promoting end-to-end protein-protein interactions. We demonstrate that fibers of a ~60 kDa bi-terminally fused amyloid-silk protein exhibit ultimate tensile strength up to  $481 \pm 31$  MPa and toughness of  $179 \pm 39$  MJ\*m<sup>-3</sup>, while achieving a high titer of  $8.0 \pm 0.70$  g/L by bioreactor production. We show that bi-terminal fusion of Mfp5 fragments significantly enhances the alignment of  $\beta$ -nanocrystals, and intermolecular interactions are promoted by cation- $\pi$  and  $\pi$ - $\pi$  interactions between terminal fragments. Our approach highlights the advantage of self-interacting intrinsically-disordered proteins in enhancing material mechanical properties and can be applied to a wide range of protein-based materials.

Microbially-synthesized materials are attractive replacements for petroleum-derived synthetic polymers as they are renewable and biodegradable. Recent advances in synthetic biology have offered the possibility to produce protein-based materials (PBMs) with precisely controlled protein sequences, functional domains, and uniform molecular weights, leading to high yet tunable mechanical performances beyond the reach of natural protein materials and petroleum-derived synthetic polymers<sup>1-3</sup>. One example is spider silk, best known for its unique combination of high tensile strength and high toughness<sup>4,5</sup>. This attractive mechanical property has motivated decades of research to produce silk and silk-mimetic materials from fast-growing microbial hosts<sup>6-8</sup>. Most high mechanical-performing protein-based materials have high molecular weight (MW), repetitive protein sequences, and highly biased amino acid compositions<sup>7,9-12</sup>. These

sequence features are essential to the mechanical properties of PBMs but have presented major technical obstacles for their biosynthesis in microbial hosts, such as limited choices of amino acids and genetic instability. Current approaches to solve these problems include post-translational protein polymerization<sup>12,13</sup>, host engineering<sup>9</sup>, and engineering biomimetic fiber spinning<sup>14</sup>. Through these efforts, recombinant spider silk protein and engineered silk-amyloid hybrid protein containing 192 and 128 repeating units with MW of 556 kDa and 378 kDa were synthesized from microbial hosts, respectively<sup>10,12</sup>. Fibers made of these high MW PBMs displayed high mechanical properties (tensile strength up to ~1 GPa, toughness up to ~160 MJ\*m<sup>-3</sup>), similar to or higher than those of natural silk fibers<sup>10</sup>. Despite this progress, producing high MW PBMs with high mechanical performances in microbial hosts often leads to low titers and yields. Metabolic and

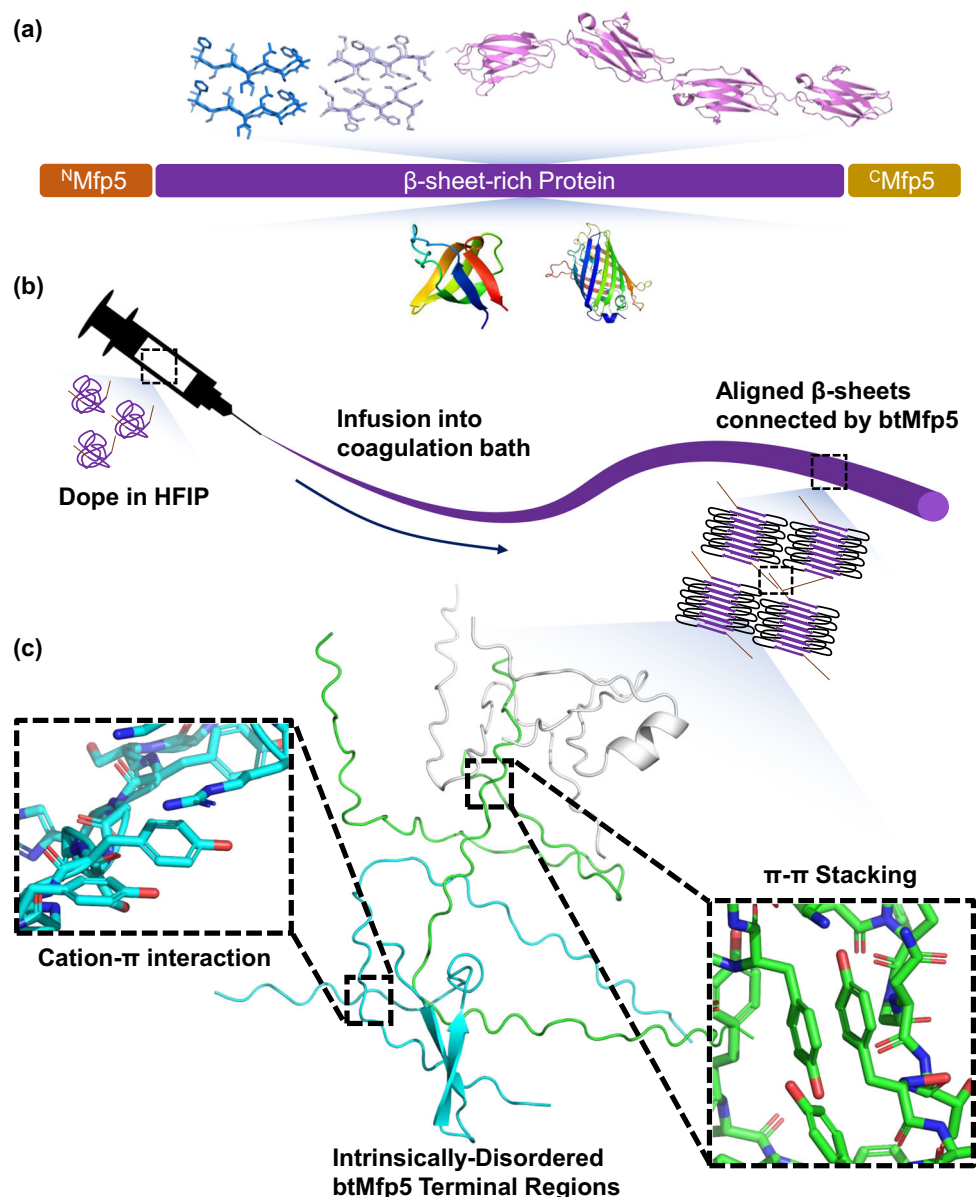
<sup>1</sup>Department of Energy, Environmental and Chemical Engineering, Washington University in St. Louis, One Brookings Drive, Saint Louis, MO 63130, USA.<sup>2</sup>Division of Biological & Biomedical Sciences, Washington University in St. Louis, One Brookings Drive, Saint Louis, MO 63130, USA. <sup>3</sup>Institute of Materials Science & Engineering, Washington University in St. Louis, One Brookings Drive, Saint Louis, MO 63130, USA. ✉e-mail: [fzhang@seas.wustl.edu](mailto:fzhang@seas.wustl.edu)

genetic engineering methods are useful in enhancing protein titers and yields<sup>15</sup>; however, these strategies have limited effects on high MW PBMs<sup>9,12</sup>. Meanwhile, a less explored strategy is to use protein engineering to achieve high mechanical performance from low MW proteins that can be produced in high titers and yields from microbial hosts.

Mussel foot protein 5 (Mfp5) is an intrinsically disordered protein secreted at the tip of mussel byssus and is used for surface adhesion underwater<sup>16,17</sup>. Mfp5 contains multiple tyrosine residues that can be enzymatically oxidized to L-3,4-dihydroxyphenylalanine (DOPA) and play important roles in surface interacting. Beside surface adhesion, the disordered and flexible Mfp5 also self-interacts under neutral and basic conditions through  $\pi$ - $\pi$  interaction, cation- $\pi$  interaction, and hydrophobic effects, thus playing important cohesive roles<sup>18</sup>. Moreover, recent experimental and modeling results have shown that mussel foot proteins remain adhesive and cohesive even without post-translational modifications<sup>19,20</sup>.

Inspired by the molecular glue Mfp5, here we demonstrate a general approach to promote both the strength and the toughness

of low MW protein fibers by genetically fusing Mfp5 fragments to their termini, therefore promoting end-to-end protein-protein interactions (Fig. 1). We applied this strategy to multiple proteins, including recombinant silk repeats, amyloid-silk hybrid proteins, a titin fragment, Fyn tyrosine kinase SH3 domain (SH3), and green fluorescent protein (GFP). Our results showed that in all these cases, bi-terminal Mfp5 (btMfp5) fusion promoted fiber ultimate tensile strength (by up to 345%) and toughness (by up to 1,970%). Using site-directed mutagenesis, we further showed that the significantly enhanced fiber mechanical properties are mostly contributed by interactions of N-half tyrosine with C-half tyrosine and/or cationic residues. One of the btMfp5-fused amyloid-silk polymers, <sup>N</sup>M-16xFGA-<sup>C</sup>M(YtoS), has an MW of 57.3 kDa, but its fiber displayed an ultimate tensile strength of  $481 \pm 31$  MPa and a toughness of  $179 \pm 39$  MJ\*m<sup>-3</sup>. These properties are comparable to recombinant spider silk of 285 kDa. This low MW protein has enabled high titer/yield protein production, resulting in a titer of  $8.0 \pm 0.70$  g/L and a yield of  $13.5 \pm 1.0\%$  of the total proteome. Our protein engineering strategy together with proteins designed in this project will enable a broad



**Fig. 1 | Schematics of the btMfp5-fused protein fiber spinning.** **a** Schematics of recombinant proteins fused with split Mfp5 fragments at their termini. **b** Schematics of protein structural change during fiber spinning. **c** Schematics of interactions between split Mfp5 fragments. Only Mfp5 fragments are shown.

range of material applications in the fields of protection, sustainability, and biomedical technology<sup>21,22</sup>.

## Results

### Bi-terminal fusion of Mfp5 fragments to amyloid-silk protein 16xFGA promotes its fiber mechanical properties

Fiber wet-spinning requires the preparation of highly concentrated protein dope and the protrusion of the dope into a coagulation bath that promotes protein interaction. To avoid loss-of-function during fiber spinning, we focused on IDPs that can strongly interact in disordered forms while remain soluble for dope preparation. One such candidate is Mfp5, whose self-interaction can be controlled by pH and solvents. At acidic pH or the presence of disrupting solvents (e.g. acetic acid or 1,1,1,3,3,3-hexafluoroisopropanol, HFIP), Mfp5 is soluble due to its high positive charge and favorable protein-solvent interactions. When placed in neutral or basic aqueous solutions, Mfp5 self-interacts by forming multiple cation- $\pi$  and  $\pi$ - $\pi$  interactions between positively charged residues and tyrosine/DOPA sidechain<sup>20,23</sup>. Such self-interaction can form even in the absence of DOPA residues, thus does not require additional enzymatic oxidation in protein preparation. The full-length Mfp5 has a MW of 8.9 kDa and contains 74 amino acids, including 20 tyrosine and 16 lysine residues. We split Mfp5 in the middle into two halves with similar number of residues and MW on each segment (Fig. 1a). Upon splitting, the N-half (<sup>N</sup>Mfp5, <sup>N</sup>M) has 10 Ys but only 3 Ks and has a net charge of +1 at pH=7, while the C-half (<sup>C</sup>Mfp5, <sup>C</sup>M) contains 10 Ys, 13 Ks, and 2 Rs with a net charge of +15 at pH=7 in aqueous solution (Supplementary Table 1, Supplementary Dataset 1).

We first fused the split Mfp5 fragments to the termini of a 16-repeat of artificially-designed amyloid-silk protein 16xFGA (Fig. 2a)<sup>10</sup>. The resulting fusion protein, namely <sup>N</sup>M-16xFGA-<sup>C</sup>M, was expressed and purified as a soluble protein in denaturing or acidic aqueous solutions (Supplementary Figure 1). <sup>N</sup>M-16xFGA-<sup>C</sup>M is also highly soluble in HFIP (up to 30% w/v) as HFIP acts as a good H-bond donor, thus a concentrated protein dope can be readily prepared and remains stable (without precipitation or aggregation). Protein dope was protruded into an 95% methanol bath, which effectively removed HFIP and promoted non-covalent interactions between Mfp5 fragments (Fig. 1)<sup>24,25</sup>. Fibers produced from this procedure were compared to fibers made of the 16xFGA protein<sup>10</sup>. Optical microscopy and scanning electron microscopy (SEM) revealed that the <sup>N</sup>M-16xFGA-<sup>C</sup>M fibers had regular cylindrical shapes and consistent diameters of  $19 \pm 0.70 \mu\text{m}$  (Fig. 2b, c). Standard tensile tests showed that btMfp5 fusion has enhanced both fiber strength and toughness. The ultimate tensile strength increased by 77% from  $230 \pm 34 \text{ MPa}$  of 16xFGA fibers<sup>10</sup> to  $406 \pm 36 \text{ MPa}$  of <sup>N</sup>M-16xFGA-<sup>C</sup>M fibers, and toughness increased by 100% from  $59 \pm 17 \text{ MJ}\cdot\text{m}^{-3}$  of 16xFGA fibers to  $118 \pm 21 \text{ MJ}\cdot\text{m}^{-3}$  of <sup>N</sup>M-16xFGA-<sup>C</sup>M fibers (Fig. 2d–g, Supplementary Figure 2, Supplementary Table 2)<sup>10</sup>. Mechanical enhancement on recombinant silk fibers was observed when globular terminal domains of spidroins were fused to a minimal repetitive domain (2 repeats)<sup>14,26</sup>. Here we have shown that bi-terminal fusion of the intrinsically disordered self-interacting Mfp5 fragments, which have significantly different structure and sequence than silk terminal domains, can also enhance fiber mechanical properties. Our result thus highlighted the potential of engineering strong fiber materials from the vast protein sequence space.

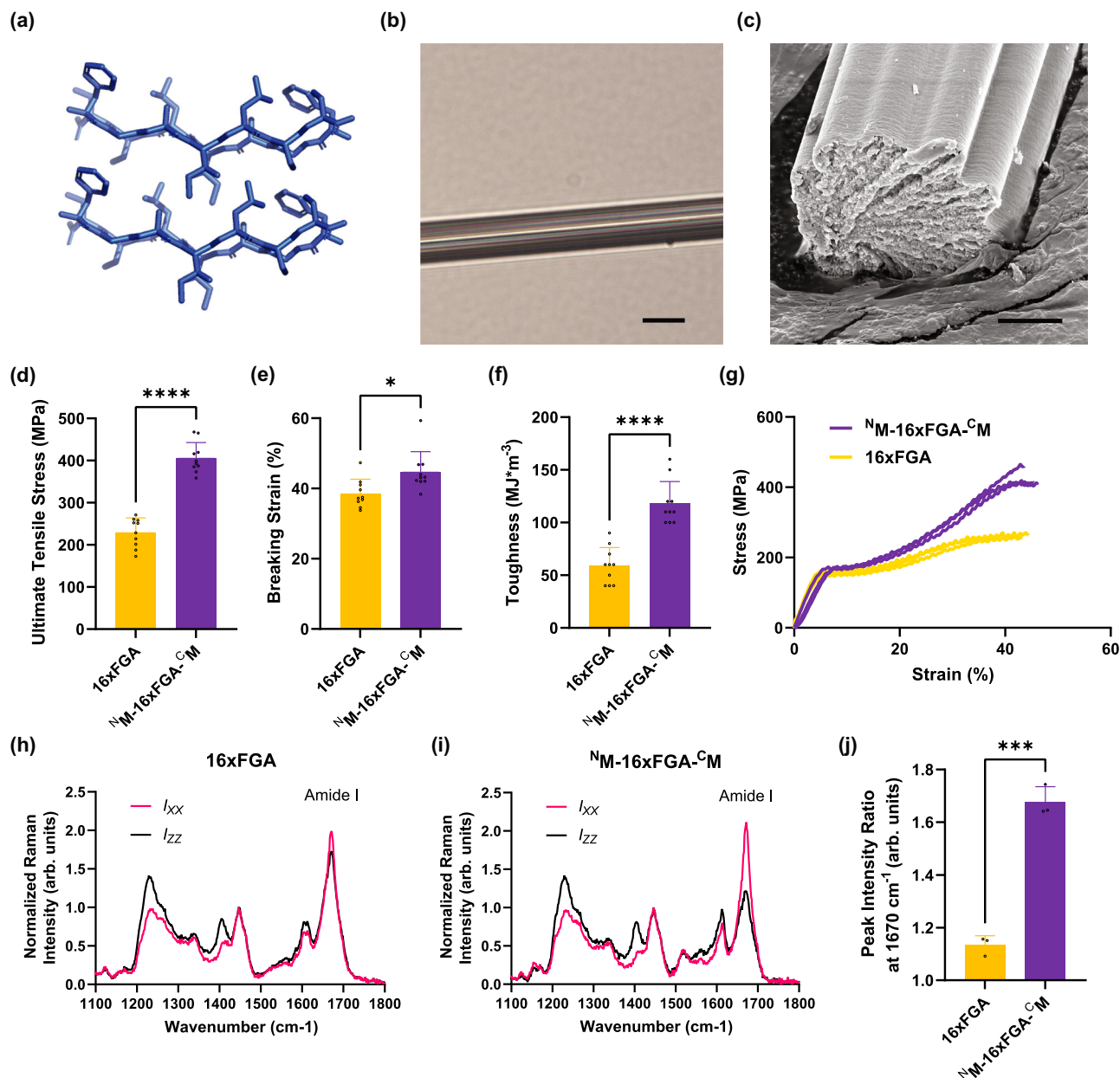
### btMfp5 fusion promotes the alignment of $\beta$ -sheets along fiber axis

After observing the enhanced mechanical properties from btMfp5 fusion, we aimed to investigate the underlying mechanism and structural influence on fibers. Except for oligomerization to reduce breaks at chain ends, terminal domains of natural spidroins also involve in fiber assembly by triggering  $\beta$ -sheet formation and their alignment along the fiber axis<sup>27–30</sup>. In natural silk fibers, the alignment of  $\beta$ -sheets to the fiber axis is an important factor contributing to the exceptional

mechanical properties<sup>7,31,32</sup>. Our previous work has shown that the 16xFGA amyloid-silk repeats can trigger  $\beta$ -nanocrystal formation<sup>10</sup>, and here we hypothesized that inter-molecular interactions between bi-terminal fused Mfp5 fragments would facilitate the alignment of  $\beta$ -sheet along fiber axis. To test this hypothesis, we used polarized Raman spectrometry to examine the amide I  $\beta$ -sheet peak ( $1670 \text{ cm}^{-1}$ ) and compared the Raman spectra when fibers was oriented either parallel (ZZ) and perpendicular (XX) to the direction of laser polarization<sup>12,28,33,34</sup>. For 16xFGA fibers, their amide I  $\beta$ -sheet peak measured at perpendicular position ( $I_{XX}$ ) is only slightly stronger than that at parallel position ( $I_{ZZ}$ ). The peak intensity ratio between perpendicular and parallel positions  $I_{YX}$  was  $1.13 \pm 0.04$ , suggesting that  $\beta$ -sheets in 16xFGA fibers are only weakly aligned to the fiber axes. In contrast, amide I  $\beta$ -sheet peak of <sup>N</sup>M-16xFGA-<sup>C</sup>M fibers are dramatically higher at the perpendicular position ( $I_{XX}$ ) than at the parallel position ( $I_{ZZ}$ ), indicating a substantially enhanced  $\beta$ -sheet alignment compared to 16xFGA fibers (Fig. 2j). The peak intensity ratio reached  $1.68 \pm 0.06$ , significantly higher than that of previously reported high MW (556 kDa) recombinant spider silk<sup>12</sup>.

### Tyrosine and positively charged residues form end-to-end intermolecular interactions between Mfp5 fragments

Upon observing the enhanced mechanical properties from btMfp5 fusion, we aimed to understand the interaction mechanism. We first fused the full-length Mfp5 to the C-terminus of 16xFGA. The resulting fusion protein 16xFGA-M has the same amino acid composition and MW with <sup>N</sup>M-16xFGA-<sup>C</sup>M, but its Mfp5 can only interact with other protein molecules from one end (Supplementary Figure 3). We found that 16xFGA-M fibers failed to exhibit comparable ultimate tensile strength with <sup>N</sup>M-16xFGA-<sup>C</sup>M fibers (Supplementary Table 3), indicating that it is necessary for the Mfp5 fragments to be fused to both protein termini to promote end-to-end intermolecular interactions. Next, we examined which residues are involved in the end-to-end interactions. The dominant form of interaction between Mfp5 are cation- $\pi$  interaction (Y-K/R)<sup>20,35,36</sup> and  $\pi$ - $\pi$  stacking (Y-Y)<sup>37</sup>. Previous research mostly emphasized the former and estimated a binding energy of  $-20.8 \text{ kcal/mol}$  per Y-K pair<sup>38</sup>. More recently, ab initio quantum mechanical simulation suggested that total binding energy from  $\pi$ - $\pi$  stacking can add up to 33–50% of the cation- $\pi$  interactions<sup>36</sup>. To differentiate these interactions, we designed, biosynthesized, and purified a series of <sup>N</sup>M-16xFGA-<sup>C</sup>M mutant proteins (Fig. 3a–c). The N-half Mfp5 (<sup>N</sup>M) is most enriched with Y with only 3 positively charged residues at neutral pH. We thus mutated all Ys to S, resulting in <sup>N</sup>M(YtoS)-16xFGA-<sup>C</sup>M (Fig. 3a, Supplementary Figure 4–6). The ultimate tensile strength of <sup>N</sup>M(YtoS)-16xFGA-<sup>C</sup>M fibers was close to those of 16xFGA fibers but was significantly lower than those of <sup>N</sup>M-16xFGA-<sup>C</sup>M fibers (Fig. 3d–f, Supplementary Figure 4–6, Supplementary Table 4), indicating that tyrosine residues in <sup>N</sup>M play vital roles in the end-to-end Mfp5 interaction. Interestingly, <sup>N</sup>M(YtoS)-16xFGA-<sup>C</sup>M fibers increased fiber-breaking strain compared to 16xFGA fibers, leading to an increase in fiber toughness. This suggests that although YtoS mutant cannot hold tensile stress beyond 260 MPa, the flexible <sup>N</sup>M(YtoS) mutant allows protein chains to be straightened and weakly-interact under extension. Additionally, the C-half Mfp5 (<sup>C</sup>M) has 10 tyrosine and 15 positively charged residues (13 Ks and 2 Rs). We separately mutated all Ys and positively charged residues to S, resulting in two <sup>C</sup>M mutants, <sup>N</sup>M-16xFGA-<sup>C</sup>M(YtoS) and <sup>N</sup>M-16xFGA-<sup>C</sup>M(KRtoS), respectively (Fig. 3b, Supplementary Figure 4–6). Fibers of both <sup>C</sup>M mutants exhibited similar mechanical properties with <sup>N</sup>M-16xFGA-<sup>C</sup>M fibers (Fig. 3d–f, Supplementary Figure 4–6, Supplementary Table 5–6), suggesting that either cation- $\pi$  interaction between C-half positively-charged residues and N-half Ys or  $\pi$ - $\pi$  stacking between Ys in both halves are sufficient to enhance fiber mechanical properties. The ultimate tensile strength and toughness of <sup>N</sup>M-16xFGA-<sup>C</sup>M(KRtoS) fibers reached  $424 \pm 58 \text{ MPa}$  and  $130 \pm 14 \text{ MJ}\cdot\text{m}^{-3}$ .



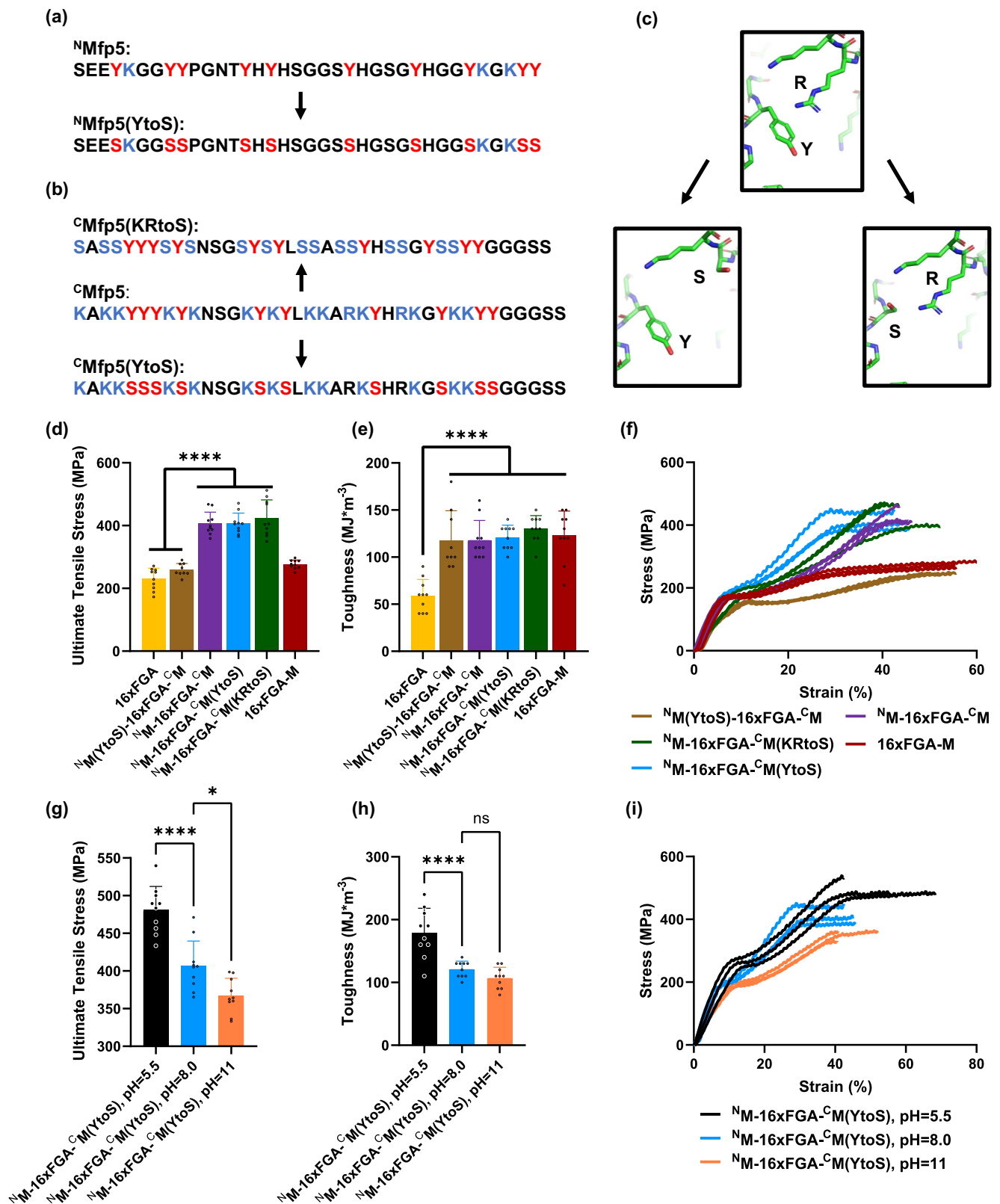
**Fig. 2 | Mechanical properties and polarized Raman spectra of 16xFGA and <sup>N</sup>M-16xFGA-<sup>C</sup>M fibers.** **a** Crystal structures of self-assembled FGAILSS (PDB ID: 5E61) in the cross-β form. **b** Optical microscopy image of <sup>N</sup>M-16xFGA-<sup>C</sup>M fiber. The scale bar in the image was 20 μm. **c** SEM image of <sup>N</sup>M-16xFGA-<sup>C</sup>M fiber. The scale bar in the image was 5 μm. **d-f** Ultimate tensile strength (**d**, \*\*\*\* $p=1.6 \times 10^{-9}$ ), breaking strain (**e**, \* $p=1.1 \times 10^{-2}$ ) and toughness (**f**, \*\*\*\* $p=2.0 \times 10^{-6}$ ) of 16xFGA and <sup>N</sup>M-16xFGA-<sup>C</sup>M fibers, two-tailed unpaired *t*-test. **g** Representative stress-strain curves of 16xFGA and <sup>N</sup>M-16xFGA-<sup>C</sup>M fibers. **h-j** Polarized Raman spectra of 16xFGA (**h**) and <sup>N</sup>M-16xFGA-<sup>C</sup>M fibers (**i**), and peak intensity ratio at 1670 cm<sup>-1</sup> of the two fibers

(**j**, \*\*\*\* $p=1.6 \times 10^{-4}$ ), two-tailed unpaired *t*-test. Data are presented as mean values ± standard deviations. Error bars represent standard deviation. 16xFGA mechanical data reproduced from previous publication (reproduced with permission<sup>10</sup>, Copyright 2021 American Chemical Society) serve as a comparison. For all mechanical data,  $n=10$  independent tensile test measurements; for peak intensity ratio,  $n=3$  ratios of a pair of peak intensity at 1670 cm<sup>-1</sup> from independently acquired Raman spectra on the *I*<sub>xx</sub> and the *I*<sub>zz</sub> directions. Details of source data and statistical analyses are provided in the Source Data file.

To further confirm the roles of cation-π interaction, we examined fiber properties under different pH. As pH decreases, Mfp5 fragments are expected to be more strongly positively charged, thus promoting more cation-π interactions and further enhancing fiber mechanical properties. To test this hypothesis, we spun the <sup>N</sup>M-16xFGA-<sup>C</sup>M(YtoS) protein at acidic and basic pH from 5.5 to 11 (see Methods). We chose the <sup>N</sup>M-16xFGA-<sup>C</sup>M(YtoS) protein because its <sup>C</sup>M lacks tyrosine, thus reducing the influence from π-π interaction. Consistent with our hypothesis, ultimate tensile strength of fibers spun at pH 5.5 was 18% and 31% higher than those spun at pH 8.0 and pH 11, while toughness was 48% and 67% higher than those spun at pH 8.0 and pH 11,

respectively (Fig. 3g-i, Supplementary Fig. 7, Supplementary Tables 8–9). <sup>N</sup>M-16xFGA-<sup>C</sup>M(YtoS) fibers spun at pH 5.5 exhibited an impressive ultimate tensile strength of  $481 \pm 31$  MPa and toughness of  $179 \pm 39$  MJ·m<sup>-3</sup>. These results not only support the role of cation-π interactions between Mfp5 fragments as an important mechanism, but also suggest strategies in optimizing process conditions (e.g. spin in acidic coagulation bath) to promote fiber mechanical properties.

Additionally, because Mfp5 is intrinsically disordered, Mfp5 fragments in these fibers most likely do not fold into a specific structure when they interact. The end-to-end intermolecular interactions may occur between any positively charged residue



and tyrosine from each bi-terminally-fused Mfp5 fragment. And such end-to-end intermolecular interactions are not limited to the current split Mfp5 fragments but can be potentially achieved by other IDPs that are rich in tyrosine and positively charged residues. To test so, we created another mutant <sup>C</sup>M-16xFGA-<sup>C</sup>M that has <sup>C</sup>M on both termini (Supplementary Fig. 8). Mechanical

properties of <sup>C</sup>M-16xFGA-<sup>C</sup>M fibers are similar to those of <sup>N</sup>M-16xFGA-<sup>C</sup>M fibers, and significantly higher than those of 16xFGA fibers (Supplementary Fig. 9, Supplementary Table 7). These results indicate that interactions between disordered <sup>C</sup>M and <sup>C</sup>M can also promote fiber mechanical properties when fused to both protein termini.

**Fig. 3 | Design and mechanical properties of btMfp5-fused 16xFGA variant fibers.** **a** Design of  $^N$ Mfp5(YtoS) sequence from  $^N$ Mfp5 with all tyrosine residues mutated to serine. Y and S mutated from Y were labeled in red, while K/R were labeled in blue. **b** Design of  $^C$ Mfp5(YtoS) and  $^C$ Mfp5(KRtoS) from  $^C$ Mfp5 with all tyrosine or lysine/arginine residues mutated to serine. Y and S mutated from Y were labeled in red, while K/R and S mutated from K/R were labeled in blue. **c** Schematics of mutations introduced to the split Mfp5 fragments. **d** Ultimate tensile strength of 16xFGA, 16xFGA-M, four btMfp5-fused 16xFGA variant fibers. **e** Toughness of 16xFGA, 16xFGA-M, four btMfp5-fused 16xFGA variant fibers. **f** Representative stress-strain curves of four btMfp5-fused 16xFGA variant fibers. **g** Summary of ultimate tensile stress for  $^N$ M-16xFGA- $^C$ M(YtoS) spun in different pH in 95%

methanol; **h** summary of toughness for  $^N$ M-16xFGA- $^C$ M(YtoS) spun in different pH in 95% methanol; **i** representative stress-strain curves of  $^N$ M-16xFGA- $^C$ M(YtoS) spun in different pH in 95% methanol. Data are presented as mean values  $\pm$  standard deviations. Error bars represent standard deviation. \* $p < 0.05$ , \*\*\*\* $p < 0.0001$ , ordinary one-way ANOVA multiple comparisons. 16xFGA mechanical data reproduced from previous publication (reproduced with permission<sup>10</sup>, Copyright 2021 American Chemical Society) serve as a comparison. For all mechanical data other than  $^N$ M(YtoS)-16xFGA- $^C$ M fibers,  $n = 10$  independent tensile test measurements; for  $^N$ M(YtoS)-16xFGA- $^C$ M fibers,  $n = 9$  independent tensile test measurements. Details of source data and statistical analyses are provided in the Source Data file.

### btMfp5 fusion promotes mechanical properties for a wide range of protein fibers

Upon demonstrating the enhanced mechanical property for 16xFGA fibers, we examined whether btMfp5 fusion can be a general method for promoting the mechanical properties of other protein fibers. We fused split Mfp5 fragments to three types of fiber proteins, including a recombinant spider silk 16xAAA (polyalanine)<sup>9</sup>, another amyloid-silk hybrid protein 16xKLV<sup>10</sup>, and a titin fragment Ig67-70<sup>11</sup>. These fiber materials use different molecular interaction mechanisms and form different sizes and portion of  $\beta$ -nanocrystals<sup>10,11</sup>. Bi-terminal fusion of split Mfp5 fragments enhanced fiber ultimate tensile strength by 65.5–139% for all three proteins (Fig. 4a–i, Supplementary Figs. 10–12, Supplementary Tables 10–13). Interestingly, we observed decreased breaking strain for btMfp5-fused 16xKLV and titin fibers, which may be a consequence of the interacting Mfp5 fragments hindering the free movement of peptide chains in the amorphous region, therefore limiting their capacity to extend before the load is transferred to the crystalline region. Polarized Raman spectrometry showed increased  $I_{VX}$  peak intensity ratios for all three btMfp5-fused fibers, confirming enhanced  $\beta$ -sheet alignment along the fiber axis for all three types of protein materials (Supplementary Fig. 13). In the case of 16xAAA fibers, the degree of crystallinity was much lower as indicated by previous WAXS results<sup>10</sup>, and the interaction among Mfp5 fragments was likely to have enhanced the chain entanglement in the amorphous region, which could explain the elevated breaking strain with btMfp5 fusion.

We further tested whether btMfp5 fusion could convert globular proteins that have never been used as fiber materials. We chose GFP which might endow fibers with fluorescence properties, and the SH3 domain, which is commonly used as a model globular protein for protein folding (Fig. 4m, q). Both proteins were not capable of forming strong intermolecular interaction. Therefore, they were not expected to yield strong fibers. Using the same wet-spinning technique as described above, we obtained GFP and SH3 fibers, and both fibers were indeed very weak and brittle (Fig. 4n, r). However, btMfp5 fusion ( $^N$ M-GFP- $^C$ M and  $^N$ M-SH3- $^C$ M) substantially enhanced both fibers' major mechanical properties. The strengths of GFP and SH3 fibers are enhanced by 345% and 54.9% to 249 MPa and 127 MPa, respectively. The breaking strains were increased by 11.6- and 3.43-folds, respectively. Most impressively, the toughnesses were enhanced by 19.7- and 19.2-fold to 59 and 100 MJ m<sup>-3</sup>, respectively (Fig. 4m–t, Supplementary Figure 14–16, Supplementary Table 14–17). The  $^N$ M-GFP- $^C$ M fibers remained fluorescent under blue light excitation (Fig. 5a, Supplementary Figure 17). Overall, these results showed that btMfp5 fusion is a general and highly effective strategy to boost mechanical properties of protein fibers.

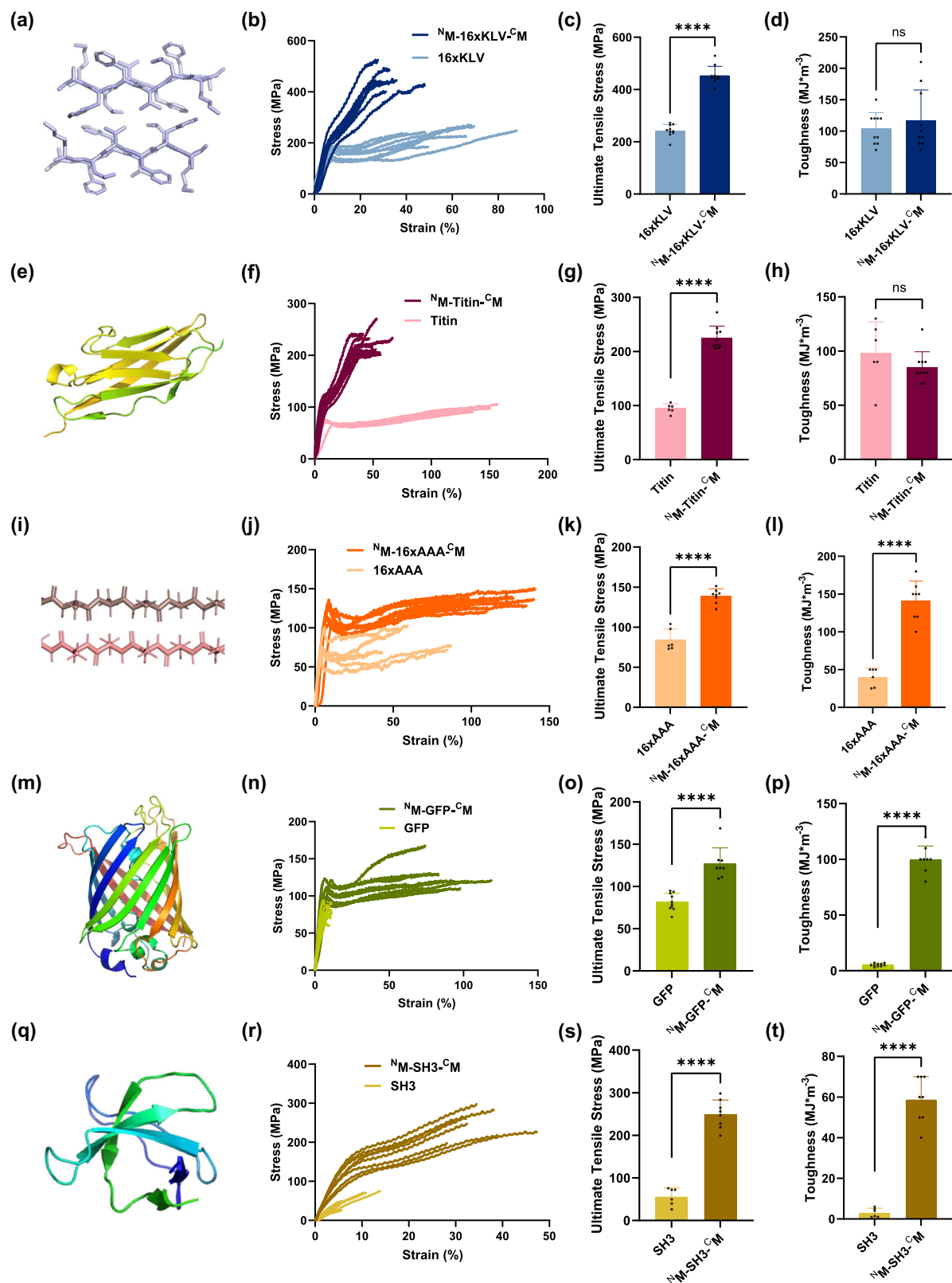
At last, to demonstrate the potential of btMfp5 fusion in facilitating scaled-up production, we expressed  $^N$ M-16xFGA- $^C$ M in a 2 L fed-batch bioreactor (Supplementary Fig. 18). The final OD<sub>600</sub> of the culture reached 155. A high protein expression level of  $13.5 \pm 1.0\%$  was obtained, and the titer of the protein was determined as  $8.0 \pm 0.70$  g/L with the standard curve, which was generated with bands of purified btMfp5-fused 16xFGA proteins with known concentrations (Supplementary Figs. 19–21, Supplementary Note). This represents a five-to-ten-fold enhancement on the protein titer compared to that past

studies, such as a 285 kDa recombinant silk protein (containing 96 repeats)<sup>39</sup>, while their fibers display comparable strength and toughness, falling well within the range of conventional natural and synthetic fibers (Fig. 5b, Supplementary Fig. 22 and Supplementary Table 18)<sup>12</sup>.

### Discussions

Our results demonstrate a general approach to improve mechanical properties of protein fibers using btMfp5 fusion, allowing fibers made of relatively low MW proteins of 20–60 kDa to display high tensile strength and toughness (up to 481 MPa and 179 MJ/m<sup>3</sup>), which has only been observed previously on high MW recombinant silk proteins with MW >285 kDa. While these large proteins can eventually generate strong fibers, their yields were significantly lower than low MW proteins as demonstrated in this study<sup>12,39</sup>. Furthermore, high MW recombinant silk proteins often require complicated purification processes such as ammonia sulfate precipitation and multi-step chromatography (e.g. size exclusion chromatography)<sup>9,10,12,40</sup>, while our btMfp5-fused low MW proteins can be purified using one-step affinity chromatography. Additionally, we show that the enhanced strength and toughness was caused by end-to-end intermolecular interactions between tyrosine in  $^N$ Mfp5 and tyrosine or positively charged residues in  $^C$ Mfp5. This approach is effective on a wide range of proteins, including traditional fiber materials (i.e. silk and amyloid-silk) as well as globular proteins (i.e. GFP and SH3) that are not traditionally used as bulk materials, thereby creating fibers with interesting properties, such as fluorescence.

This approach is widely applicable because the fused Mfp5 fragments are disordered and the fiber spinning process is not affected by the unfolding of proteins. Previously, water-soluble terminal domains of natural spidroins were fused to two repeating units of recombinant silk protein<sup>14,26</sup>. During fiber spinning, these terminal domains undergo NTD-NTD and CTD-CTD dimerization in response to pH change and shearing force, thus allowing multiple molecules to form non-covalently-associates and increasing *de facto* MW. This strategy elegantly enabled aqueous spinning similar to that used by spiders<sup>14,26,41–44</sup>. Nevertheless, fusing the terminal domains to a large number of repeats or other proteins often leads to misfolding of the chimeric proteins, thus limiting its use<sup>27,45</sup>. Our method exploited the design principle of silk terminal domains, but used the disordered Mfp5, whose self-interactions were not affected by misfolding and could be controlled during fiber spinning. Additionally, disordered Mfp5 fragments have an estimated radius of gyration of 12–16 Å<sup>46,47</sup>, allowing them to sample a large space for protein-protein interaction. Interaction between Mfp5 fragments can also adopt numerous configurations (e.g. between different residues) and orientations (Fig. 1c). Both advantageous features of IDP have greatly enhanced the chance of end-to-end protein interaction, thereby effectively reducing defects caused by chain-ends and promoting fiber mechanical properties. Although in solution, Mfp5 fragments within a chimeric protein can undergo intramolecular interactions if the N- and C-termini of the fused protein are in close proximity, the shear force during fiber spinning is expected to stretch protein chains along the fiber axis and effectively separate them, thus promoting the oligomerization of fused PBMs.

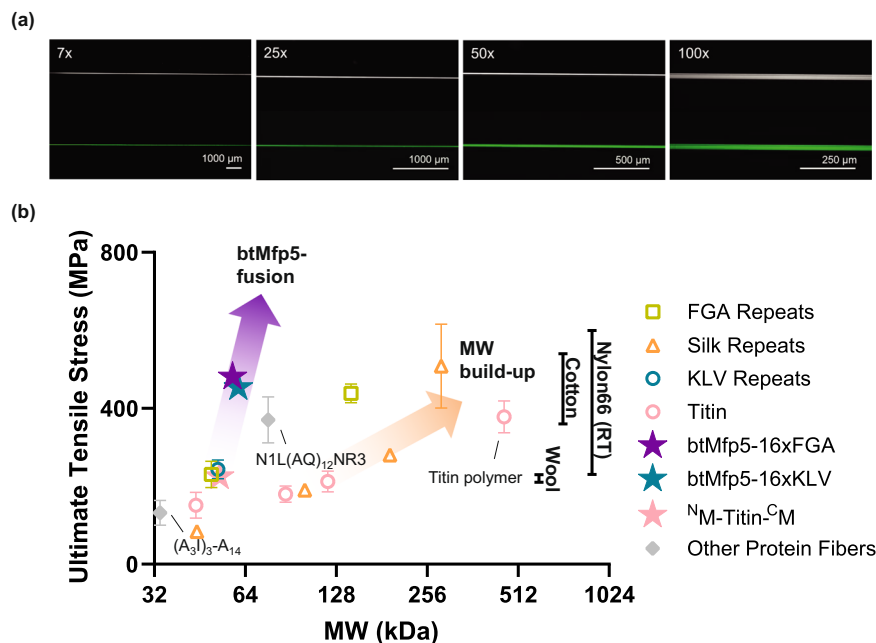


Tyrosine residues of natural Mfp5 are enzymatically oxidized to DOPA<sup>48</sup>, which play key roles in protein-protein cohesive interaction and surface adhesion<sup>23</sup>. Adhesions to metal, metal oxide and silica surfaces are particularly strong due to the formation of high energy chelation and bi-dentate hydrogen bonds<sup>23,49</sup>. However, recent studies demonstrated that Mfp and Mfp-derived peptides without tyrosine-

DOPA conversion showed a comparable level of cohesive-interaction and adhesion to surfaces such as mica or hydrophilic glass<sup>19,20,37,50</sup>. Experimental and molecular dynamic (MD) studies further revealed the strong interaction between tyrosine sidechains and positively charged residues in Mfp-Mfp interaction<sup>19,20</sup>. Tyrosine in Mfp5 fragments used in this study remained unmodified, thus largely

**Fig. 4 | Bi-terminal Mfp5 fusion boosted mechanical properties of fibers spun from other silk-forming/ $\beta$ -sheet-rich non-silk-forming proteins.** **a–d** Amyloid steric zipper structure (PDB ID: 3OW9), stress-strain curves, ultimate tensile stress and toughness of 16xKLV and  $^N$ M-16xKLV- $^C$ M fibers; **e–h** representative Ig structure (PDB ID: 3B43), stress-strain curves, ultimate tensile stress and toughness of methanol-spun titin and  $^N$ M-titin- $^C$ M fibers; **i–l** representative secondary structure element in  $\beta$ -nanocrystals in polyalanine spider silk, stress-strain curves, ultimate tensile stress and toughness of 16xAAA and  $^N$ M-16xAAA- $^C$ M fibers; **m–p** structure (PDB ID: 1EMB), stress-strain curves, ultimate tensile stress and toughness of GFP and  $^N$ M-GFP- $^C$ M fibers; **q–t** Structure (PDB ID: 3UA6), stress-strain curves, ultimate

tensile stress and toughness of SH3 and  $^N$ M-SH3- $^C$ M fibers. Data are presented as mean values  $\pm$  standard deviations. Error bars represent standard deviation. ns  $p \geq 0.05$ , no significant difference; \*\*\*\* $p < 0.0001$ , two-tailed unpaired  $t$ -test. 16xKLV and 16xAAA mechanical data reproduced from previous publication (reproduced with permission<sup>10</sup>, copyright 2021 American Chemical Society) serve as comparisons. For 16xKLV,  $^N$ M-16xKLV- $^C$ M,  $^N$ M-titin- $^C$ M and GFP fibers,  $n = 10$  independent tensile test measurements; for  $^N$ M-16xAAA- $^C$ M,  $^N$ M-GFP- $^C$ M fibers and  $^N$ M-SH3- $^C$ M fibers,  $n = 8$  independent tensile test measurements; for methanol-spun titin, 16xAAA and SH3 fibers,  $n = 6$  independent tensile test measurements. Details of source data and statistical analyses are provided in the Source Data file.



**Fig. 5 | btMfp5-fusion as an alternative strategy other than MW build-up to boost mechanical properties of protein-based fiber materials.** **a** Phase-contrast and fluorescent images of  $^N$ M-GFP- $^C$ M fiber at different magnifications; **b** strength-MW plot of various protein-based fiber materials from this work and from previous works<sup>9–11</sup>. Data are presented as mean values  $\pm$  standard deviations. Error bars represent standard deviation. Fiber average ultimate tensile strength, standard deviation and sample size data from previous works include FGA Repeats: 16xFGA,

48xFGA (left to right)<sup>10</sup>; Silk Repeats: 16-mer, 32-mer, 64-mer and 96-mer (left to right)<sup>9</sup>; KLV Repeats: 16xKLV<sup>10</sup>; Titin (water): Titin 4Ig, Titin 8Ig, Titin 12Ig and Titin polymer (left to right)<sup>11</sup>; Other Protein Fibers: (A3I)3-A14<sup>44</sup>, NIL(AQ)12NR3<sup>26</sup>. Black lines on the right indicate typical values of strength of different conventional fibers (RT: regular tenacity) and are independent from the  $x$ -axis. The two arrows indicate the two strategies in designing mechanically-strong protein-based fiber materials. Source data are provided in the Source Data file.

simplifying protein synthesis, resulting in high protein titer ( $8.0 \pm 0.70$  g/L). Our observation on the role of cation- $\pi$  and  $\pi$ - $\pi$  interaction are also consistent with these previous studies.

In summary, this work demonstrated that bi-terminal fusion of Mfp5 fragments boosted the mechanical properties of multiple recombinant protein fibers. Low MW proteins designed with this strategy ( $\sim 60$  kDa, e.g.  $^N$ M-16FGA- $^C$ M) can be produced at high titers, and their fibers displayed impressively high strength and toughness. The combination of high protein yield and strong mechanical performance makes this engineered protein a promising target for industrial-scale manufacturing to enable practical applications. Our strategy highlights the advantage of self-interacting IDPs in promoting fiber mechanical properties. While btMfp5 fusion can be potentially used on a wide range of protein-based materials, this design principle opens opportunities to engineer other IDPs that better promote material properties, even beyond fiber materials.

## Methods

### Chemicals and reagents

All chemicals and reagents were procured from MilliporeSigma (Burlington, MA), unless specified otherwise. Plasmid cloning experiments were completed using gel extraction kits and plasmid miniprep kits

(iNtRON Biotechnology, Seoul, Republic of Korea) following the manufacturer's instructions. Phusion DNA polymerase, FastDigest restriction enzymes and T4 DNA ligase (Thermo Fisher Scientific, MA) were employed for all the PCR, digestions and ligations according to the manufacturer's recommended protocols. The Ni-NTA columns used in this study were purchased from (GE Healthcare, IL).

### Plasmid assembly

DNA sequences of 16x polymeric amyloid-silk and 16x recombinant spider silk repeats were obtained from plasmids pB6c-A-16xKLVFFAE-H10 and pE8a-A-16xFGAILSS-H10, respectively<sup>10,51</sup>. The titin fragment was originated from rabbit soleus titin domains Ig67-70 and was amplified from plasmid p-4XT<sup>11,52</sup>. The human Fyn tyrosine kinase SH3 domain was prepared using synthetic DNA (IDT Inc, Coralville, IA)<sup>53</sup>. The GFP sequence was obtained from a plasmid as used in a previous study<sup>54</sup>. The *Mytilus galloprovincialis* Mfp5 sequence was obtained from plasmid pE7a-mfp5<sup>(1)</sup> as described in recent studies<sup>18,49</sup>. Mfp5 mutants were prepared from synthetic DNAs (IDT Inc, Coralville, IA) and codon-optimized for *E. coli* expression. The N- and C-half Mfp5 coding sequences were incorporated between the Acc65I/NheI and BclI/Kpn2I restriction sites of the pE8a plasmid. Primers used in this study were chemically synthesized (IDT Inc, Coralville, IA), and their



sequence information are provided in the associated supplementary data file. *E. coli* NEB10 $\beta$  strain (New England Biolabs, MA) was used for all plasmid construction and protein expression. *E. coli* cells were cultured in Luria Broth (LB) media (5 g/L yeast extract, 10 g/L tryptone and 10 g/L NaCl) with appropriate antibiotics of 50  $\mu$ g/mL kanamycin or 100  $\mu$ g/mL ampicillin on orbital shakers at 37 °C for plasmid cloning purposes.

### Protein expression in shake flasks

Protein expressions were performed in shake flasks using Terrific Broth (TB) media (24 g/L yeast extract, 20 g/L tryptone, 0.4% v/v glycerol, 17 mM KH<sub>2</sub>PO<sub>4</sub> and 72 mM K<sub>2</sub>HPO<sub>4</sub>, pH = 7.2) containing 100  $\mu$ g/mL ampicillin. Protein expression in shake flasks began with transformation of corresponding plasmids into *E. coli* NEB10 $\beta$  strain. Once colonies grew on LB agar (5 g/L yeast extract, 10 g/L tryptone, 10 g/L NaCl and 15 g/L agar) plates, a single colony was selected to inoculate a seed culture of 50 mL LB media and cultivated at 37 °C with orbital shaking until OD<sub>600</sub> reached 0.8. This seed culture was then used to inoculate TB media in 2 L shake flasks at 37 °C with orbital shaking until OD<sub>600</sub> reached 3–5. All cultures were induced by addition of 0.4% arabinose and were allowed to grow at 30 °C with orbital shaking for additional 24 h. Cells pellets were collected by centrifugation and were stored at –80 °C before proceeding to purification.

### Protein purification

Protein purification protocol was modified from previous methods<sup>10,12,55</sup>. For insoluble proteins, including all FGA and KLV proteins, with or without btMfp5 fusion, cell pellets were lysed in buffer A (6 M guanidine hydrochloride, 50 mM K<sub>2</sub>HPO<sub>4</sub> and 300 mM NaCl, pH = 7.0) at a ratio of 40 mL buffer A per liter of cell culture, followed by centrifugation. The supernatant was subjected to Ni-NTA column chromatography and washed sequentially with buffer B (8 M urea, 50 mM K<sub>2</sub>HPO<sub>4</sub> and 300 mM NaCl, pH = 7.0) containing 0 mM, 20 mM, and 50 mM imidazole. The target proteins were eluted using buffer B containing 300 mM imidazole. For water-soluble proteins, such as titin, GFP, and SH3, cell pellets were lysed in PBS buffer (137 mM NaCl, 2.7 mM KCl, 1.8 mM KH<sub>2</sub>PO<sub>4</sub> and 10 mM Na<sub>2</sub>HPO<sub>4</sub>, pH = 7.4) at a ratio of 40 mL PBS buffer per liter of cell culture, using sonication followed by centrifugation. The supernatant was loaded onto a Ni-NTA column, washed with PBS buffer containing 0 mM, 20 mM, and 50 mM imidazole, and eluted with PBS buffer containing 300 mM imidazole. The purified proteins were dialyzed against 5% acetic acid, lyophilized, and stored at –80 °C before dissolved in HFIP to make spinning dopes.

### SDS-PAGE

All SDS-PAGE gels were 1 mm thick and discontinuous with 5% stacking gel on the top and indicated percentages separation gels on the bottom. Samples were resuspended in 1x Laemmli sample buffer (10% v/v glycerol, 2% w/v SDS, 60 mM Tris-HCl, 0.01% w/v bromophenol blue and 100  $\mu$ M DTT, pH = 6.8) and incubated at 95–100 °C for 10 min. Electrophoresis were performed on Mini-PROTEAN Tetra Cells (Bio-Rad, CA) in 1x Tris-glycine SDS buffer (25 mM Tris-HCl, 250 mM glycine and 0.1% w/v SDS, pH = 8.8). For staining, all SDS-PAGE gels were incubated in a staining buffer (40% v/v methanol, 10% v/v acetic acid, 0.1% w/v Coomassie Brilliant Blue G-250 in water) for at least 1 h, before transferred to a destaining buffer (40% v/v methanol, 7% v/v acetic acid in water) and destained on orbital shakers at room temperature. Gel images were acquired using Azure Biosystems cSeries (Azure Biosystems Inc., CA) and were analyzed by ImageJ 1.53 software<sup>56</sup>.

### Bioproduction in fed-batch bioreactors

Fed-batch fermentation of <sup>15</sup>N-M-16xFGA-CM from pET28a vector in *E. coli* BL21(DE3) strain (New England Biolabs, MA) was performed following published protocols<sup>41,57</sup>. Specifically, cells were cultivated in 600 mL

TB media with 50  $\mu$ g/mL kanamycin until OD<sub>600</sub> reached ~5. The culture was then centrifuged, and the pellet was resuspended in 1 L batch media (10 g/L glucose, 60 g/L glycerol, 20 g/L tryptone, 24 g/L yeast extract, 0.5 g/L magnesium sulfate heptahydrate, 3.4 g/L potassium phosphate monobasic, 3.6 g/L sodium phosphate dibasic, 2.7 g/L ammonium chloride, 0.7 g/L sodium sulfate, 100.8 mg/L ferric citrate, 2.5 mg/L cobalt chloride hexahydrate, 15 mg/L manganese chloride tetrahydrate, 1.5 mg/L copper chloride dihydrate, 3 mg/L boric acid, 2.1 mg/L sodium molybdate dihydrate, 33.8 mg/L zinc acetate hydrate, 14.1 mg/L EDTA, 50 mg/L kanamycin, 0.01% v/v antifoam 204), and transferred to a 2 L Bioflo120 fed-batch bioreactor (Eppendorf, Hamburg, Germany). The culturing was controlled at 29 °C, and the pH was maintained at 7.1 with either 3 M phosphoric acid or 25% ammonia. Autoclaved water was supplemented to the system in compensation for water losses during fermentation. The DO level was initially set at 70% and maintained at no less than 30% by adjusting the stirring speed (200–1200 rpm) and the airflow rate. Cells were induced at OD<sub>600</sub>  $\approx$  68 with a final IPTG concentration of 0.25 mM. Feed media (400 g/L glycerol, 20 g/L tryptone, 24 g/L yeast extract, 40 g/L magnesium sulfate heptahydrate, 3.4 g/L potassium phosphate monobasic, 3.6 g/L sodium phosphate dibasic, 2.7 g/L ammonium chloride, 0.7 g/L sodium sulfate, 40 mg/L ferric citrate, 4 mg/L cobalt chloride hexahydrate, 23.5 mg/L manganese chloride tetrahydrate, 2.3 mg/L copper chloride dihydrate, 4.7 mg/L boric acid, 4 mg/L sodium molybdate dihydrate, 16 mg/L zinc acetate hydrate, 13 mg/L EDTA, 50 mg/L kanamycin, 0.01% v/v antifoam 204) was added at a rate of 0.2 mL/min. Pellets were collected 20 h after induction. Estimation of protein expression level and titer was completed following the protocol in Supplementary Notes.

### Fiber spinning

Fiber spinning protocol was adopted from previous publications with some modifications<sup>9,10,12</sup>. Spinning dopes were prepared by the dissolution of lyophilized protein powders in HFIP. The resulting solution was loaded to a 100  $\mu$ L Hamilton syringe (Hamilton Robotics, NV), and extruded into a 95% v/v methanol bath (pH = 8.0) using a Harvard Apparatus Pump II Elite syringe pump (Harvard Apparatus, MA) at a controlled protrusion rate of 10  $\mu$ L/min. To achieve pH-controlled spinning, either 5 mM CH<sub>3</sub>COONa or 5 mM Na<sub>2</sub>CO<sub>3</sub> was added to the 95% v/v methanol bath to adjust the pH to 5.5 or 11, respectively. The extruded fibers were then transferred to a 75–80% v/v methanol bath and extended to five times their original length. Following extension, the fibers were removed from the methanol bath and allowed to dry in a ventilated, dark, and dry room at room temperature until further testing.

### Mechanical testing

Tensile tests were conducted on an MTS Criterion Model 41 universal test frame equipped with a 1 N load cell (MTS Systems Corporation, MN). The tests were carried out under controlled environmental conditions of 20% relative humidity (RH) and 25 °C temperature, and with a constant pulling speed of 10 mm/min. Stress-strain curves were obtained at a 50 Hz sampling rate. To calculate the mechanical properties, engineer stress and strain were employed. Specifically, the ultimate tensile strength was calculated as the maximum load divided by the initial fiber cross-sectional area, which was determined by light microscopy before testing, assuming a circular cross-section. The modulus was obtained by fitting a linear least-square regression line to the initial elastic region of the stress-strain curve. The ultimate breaking strain was calculated as the percentage of fiber elongation relative to the initial gauge length of 5 mm before fracture. The toughness was determined as the area under the total stress-strain curve. The MTS TestSuite TW Elite software (MTS Systems Corporation, MN) was used for data collection and mechanical properties calculation.

### Light microscopy

Fiber diameters were measured using a Zeiss Axio Observer ZI inverted microscope (Zeiss, Stuttgart, Germany) with a phase-contrast 20x objective lens. The quantification was completed using the Axiovision LE software (Zeiss, Stuttgart, Germany).

### SEM

Fibers samples were acquired after tensile tests and were mounted onto a sample holder using conductive tapes. The sample holder was sputtered with 10 nm gold using a Leica EM ACE600 high-vacuum sputter coater (Leica Microsystems, Wetzlar, Germany). Imaging was completed with a ThermoFisher Quattro S Environmental Scanning Electron Microscope (ThermoFisher Scientific, MA) at an accelerating voltage of 10 kV and a spot size of 3.0, and images were acquired using AutoScript 4 software (ThermoFisher Scientific, MA).

### Fluorescence microscopy

Phase-contrast images of the <sup>N</sup>M-GFP-<sup>C</sup>M fiber were observed by the fluorescence stereomicroscope (ZEISS Axio Zoom V16) equipped with a 7x, 25x, 50, and 100x lens. For the fluorescence images, the 7x lens required 550 msec exposure time, others were set as 280 msec. Trichannel (EGFP, DAPI, DsRed) fluorescent images were captured at the same magnification (80x), with exposure time of 90 ms, using Hamamatsu ORCA-Flash4.0 V3 sCMOS camera equipped in Zeiss Axio Zoom V16 in 16 bit and processed using Zen Software (Zeiss, Stuttgart, Germany).

### Polarized Raman spectroscopy analysis

Polarized Raman Spectroscopy were recorded following published methods<sup>12,58–60</sup>. Specifically fiber samples were fixed by tape on a glass slide. Raman spectra were acquired with a Renishaw RM1000 InVia confocal Raman spectrometer (Renishaw, Wotton-under-Edge, United Kingdom) coupled to a Leica DM LM microscope with rotating stage (Leica Microsystems, Wetzlar, Germany). Fibers were initially oriented along the *x*-axis and were irradiated using a 514 nm line argon laser with polarization fixed along the *x*-axis and focused through a 50x objective (NA = 0.75). Fibers were then aligned to the *y*-axis, followed by the same data acquisition process. Spectra were recorded from 1150 to 1750 cm<sup>-1</sup> with an 1800 lines/mm grating. For each acquisition, a total of 16 spectra were accumulated, each for 10 s. All fibers remained intact after acquisition with no visual sign of degradation under the incident laser. Spectra collected were analyzed with Fityk 1.3.1<sup>61</sup>. Baseline subtraction is accomplished using a built-in Fityk automatic convex hull algorithm. All spectra were normalized to the intensity of the 1447 cm<sup>-1</sup> peaks, which arises from CH<sub>3</sub> asymmetric stretching and CH<sub>2</sub> bending<sup>59</sup> and is insensitive to protein conformation.

### Statistics and reproducibility

Statistical analyses, including student *t*-test and ANOVA, were conducted with Microsoft Excel 365 (Microsoft Corporation, WA) and GraphPad Prism 9 (GraphPad Software, CA). All data are presented means ± SD as depicted in the figure caption. Significant statistical difference is considered at *p* < 0.05 and is denoted with asterisks (\*), while no significant statistical difference is considered at *p* ≥ 0.05 and is denoted with 'ns' in figures. Statistical tests for significance were described in individual figure legends. Details about statistical analyses can also be found in the Source Data file. Details about statistical analyses in Fig. 5b can be found in the corresponding references if those results were not generated in this study. No data were excluded from the analyses. For representative experiments (imaging), each experiment has been repeated for at least three times independently with similar results.

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

All experimental data from mechanical testing, Raman spectra, uncropped raw gel images, protein sequences and oligonucleotide sequences generated in this study underlying all main figures and supplementary figures are provided as a Source Data file. All protein sequences in this study have been provided in Supplementary Dataset 1. Protein structures with accession numbers “SE61”, “3OW9”, “3B43”, “1EMB” and “3UA6” were used in this work and obtained from the RCSB Protein Data Bank. All remaining data generated, processed, or analyzed in this study are available within the article and Supplementary Information or from the corresponding author upon reasonable request. Source data are provided with this paper.

### References

- Hu, X., Cebe, P., Weiss, A. S., Omenetto, F. & Kaplan, D. L. Protein-based composite materials. *Mater. Today* **15**, 208–215 (2012).
- Sun, J. et al. Fabrication and mechanical properties of engineered protein-based adhesives and fibers. *Adv. Mater.* **32**, e1906360 (2020).
- Partlow, B. P. et al. Highly tunable elastomeric silk biomaterials. *Adv. Funct. Mater.* **24**, 4615–4624 (2014).
- Lewis, R. V. Spider silk—the unraveling of a mystery. *Acc. Chem. Res.* **25**, 392–398 (1992).
- Work, R. W. Mechanisms of major ampullate silk fiber formation by orb-web-spinning spiders. *J. Microsc. Soc.* **96**, 170–189 (1977).
- Chung, H., Kim, T. Y. & Lee, S. Y. Recent advances in production of recombinant spider silk proteins. *Curr. Opin. Biotechnol.* **23**, 957–964 (2012).
- Yarger, J. L., Cherry, B. R. & van der Vaart, A. Uncovering the structure-function relationship in spider silk. *Nat. Rev. Mater.* **3**, 18008 (2018).
- Omenetto, F. G. & Kaplan, D. L. New opportunities for an ancient material. *Science* **329**, 528–531 (2010).
- Xia, X. X. et al. Native-sized recombinant spider silk protein produced in metabolically engineered *Escherichia coli* results in a strong fiber. *Proc. Natl Acad. Sci. USA* **107**, 14059–14063 (2010).
- Li, J. et al. Microbially Synthesized polymeric amyloid fiber promotes beta-nanocrystal formation and displays gigapascal tensile strength. *ACS Nano* **15**, 11843–11853 (2021).
- Bowen, C. H. et al. Microbial production of megadalton titin yields fibers with advantageous mechanical properties. *Nat. Commun.* **12**, 5182 (2021).
- Bowen, C. H. et al. Recombinant spidroins fully replicate primary mechanical properties of natural spider silk. *Biomacromolecules* **19**, 3853–3860 (2018).
- Bowen, C. H. et al. Seeded chain-growth polymerization of proteins in living bacterial cells. *ACS Synth. Biol.* **8**, 2651–2658 (2019).
- Andersson, M. et al. Biomimetic spinning of artificial spider silk from a chimeric minispidroin. *Nat. Chem. Biol.* **13**, 262–264 (2017).
- Silva, F., Queiroz, J. A. & Domingues, F. C. Evaluating metabolic stress and plasmid stability in plasmid DNA production by *Escherichia coli*. *Biotechnol. Adv.* **30**, 691–708 (2012).
- Brennan, M. J., Kilbride, B. F., Wilker, J. J. & Liu, J. C. A bioinspired elastin-based protein for a cytocompatible underwater adhesive. *Biomaterials* **124**, 116–125 (2017).
- Wheat, J. C. & Wolf, J. S. Jr. Advances in bioadhesives, tissue sealants, and hemostatic agents. *Urol. Clin. North Am.* **36**, 265–275 (2009).
- Kim, E. et al. Microbially synthesized repeats of mussel foot protein display enhanced underwater adhesion. *ACS Appl Mater. Interfaces* **10**, 43003–43012 (2018).
- Ou, X. et al. Structure and sequence features of mussel adhesive protein lead to its salt-tolerant adhesion ability. *Sci. Adv.* **6**, 39 (2020).
- Gebbie, M. A. et al. Tuning underwater adhesion with cation- $\pi$  interactions. *Nat. Chem.* **9**, 473–479 (2017).

21. Aizenberg, J. New nanofabrication strategies: inspired by biomineralization. *MRS Bull.* **35**, 323–330 (2011).
22. Kiseleva, A. P., Krivoschapkin, P. V. & Krivoschapkina, E. F. Recent advances in development of functional spider silk-based hybrid materials. *Front Chem.* **8**, 554 (2020).
23. Lee, H., Scherer, N. F. & Messersmith, P. B. Single-molecule mechanics of mussel adhesion. *Proc. Natl Acad. Sci. USA* **103**, 12999–13003 (2006).
24. Zhu, Y. et al. Water and the Cation- $\pi$  Interaction. *J. Am. Chem. Soc.* **143**, 12397–12403 (2021).
25. Bania, K. K., Guha, A. K., Bhattacharyya, P. K. & Sinha, S. Effect of substituent and solvent on cation- $\pi$  interactions in benzene and borazine: a computational study. *Dalton Trans.* **43**, 1769–1784 (2014).
26. Heidebrecht, A. et al. Biomimetic fibers made of recombinant spidroins with the same toughness as natural spider silk. *Adv. Mater.* **27**, 2189–2194 (2015).
27. Askarieh, G. et al. Self-assembly of spider silk proteins is controlled by a pH-sensitive relay. *Nature* **465**, 236–238 (2010).
28. Xu, L. et al. Structural and mechanical roles for the C-terminal nonrepetitive domain become apparent in recombinant spider aciniform silk. *Biomacromolecules* **18**, 3678–3686 (2017).
29. Hagn, F. et al. A conserved spider silk domain acts as a molecular switch that controls fibre assembly. *Nature* **465**, 239–242 (2010).
30. Hagn, F. A structural view on spider silk proteins and their role in fiber assembly. *J. Pept. Sci.* **18**, 357–365 (2012).
31. Rising, A., Hjalm, G., Engstrom, W. & Johansson, J. N-terminal nonrepetitive domain common to dragline, flagelliform, and cylindrical spider silk proteins. *Biomacromolecules* **7**, 3120–3124 (2006).
32. Motriuk-Smith, D., Smith, A., Hayashi, C. Y. & Lewis, R. V. Analysis of the conserved N-terminal domains in major ampullate spider silk proteins. *Biomacromolecules* **6**, 3152–3159 (2005).
33. Lefevre, T., Boudreault, S., Cloutier, C. & Pezolet, M. Conformational and orientational transformation of silk proteins in the major ampullate gland of *Nephila clavipes* spiders. *Biomacromolecules* **9**, 2399–2407 (2008).
34. Liu, Y., Shao, Z. & Vollrath, F. Extended wet-spinning can modify spider silk properties. *Chem. Commun. (Camb.)* **19**, 2489–2491 (2005).
35. Park, S., Kim, S., Jho, Y. & Hwang, D. S. Cation- $\pi$  Interactions and their contribution to mussel underwater adhesion studied using a surface forces apparatus: a mini-review. *Langmuir* **35**, 16002–16012 (2019).
36. Kim, S. et al. Cation- $\pi$  interaction in DOPA-deficient mussel adhesive protein mfp-1. *J. Mater. Chem. B* **3**, 738–743 (2015).
37. Deepankumar, K. et al. Liquid-liquid phase separation of the green mussel adhesive protein Pvf-5 is regulated by the post-translated dopa amino acid. *Adv. Mater.* <https://doi.org/10.1002/adma.202103828> (2021).
38. Yildiz, I. A computational insight into the interaction of methylated lysines with aromatic amino acid cages. *J. Phys. Org. Chem.* <https://doi.org/10.1002/poc.3660> (2017).
39. Whittall, D. R., Baker, K. V., Breiting, R. & Takano, E. Host systems for the production of recombinant spider silk. *Trends Biotechnol.* **39**, 560–573 (2021).
40. Li, J. & Zhang, F. Amyloids as building blocks for macroscopic functional materials: designs, applications and challenges. *Int. J. Mol. Sci.* **22**, 10698. (2021).
41. Schmuck, B. et al. High-yield production of a super-soluble miniaturized spidroin for biomimetic high-performance materials. *Mater. Today* **50**, 16–23 (2021).
42. Saric, M., Eisoldt, L., Doring, V. & Scheibel, T. Interplay of different major ampullate spidroins during assembly and implications for fiber mechanics. *Adv. Mater.* **33**, e2006499 (2021).
43. Huemmerich, D. et al. Primary structure elements of spider dragline silks and their contribution to protein solubility. *Biochemistry* **43**, 13604–13612 (2004).
44. Arndt, T. et al. Engineered spider silk proteins for biomimetic spinning of fibers with toughness equal to dragline silks. *Adv. Funct. Mater.* **32**, 2200986 (2022).
45. Andersson, M. et al. Carbonic anhydrase generates CO<sub>2</sub> and H<sup>+</sup> that drive spider silk formation via opposite effects on the terminal domains. *PLoS Biol.* **12**, e1001921 (2014).
46. Kohn, J. E. et al. Random-coil behavior and the dimensions of chemically unfolded proteins. *Proc. Natl Acad. Sci. USA* **101**, 12491–12496 (2004).
47. Sneha, P. & Doss, C. G. Molecular dynamics: new frontier in personalized medicine. *Adv. Protein Chem. Struct. Biol.* **102**, 181–224 (2016).
48. Waite, J. H. Mussel adhesion—essential footwork. *J. Exp. Biol.* **220**, 517–530 (2017).
49. Kim, E. et al. A biosynthetic hybrid spidroin-amyloid-mussel foot protein for underwater adhesion on diverse surfaces. *ACS Appl. Mater. Interfaces* **13**, 48457–48468 (2021).
50. Bilotto, P. et al. Adhesive properties of adsorbed layers of two recombinant mussel foot proteins with different levels of DOPA and tyrosine. *Langmuir* **35**, 15481–15490 (2019).
51. Dai, B., Sargent, C. J., Gui, X., Liu, C. & Zhang, F. Fibril self-assembly of amyloid-spider silk block polypeptides. *Biomacromolecules* **20**, 2015–2023 (2019).
52. von Castelmur, E. et al. A regular pattern of Ig super-motifs defines segmental flexibility as the elastic mechanism of the titin chain. *Proc. Natl Acad. Sci. USA* **105**, 1186–1191 (2008).
53. Zhang, F. et al. Structure-based approach to the photocontrol of protein folding. *J. Am. Chem. Soc.* **131**, 2283–2289 (2009).
54. Han, Y. & Zhang, F. Heterogeneity coordinates bacterial multi-gene expression in single cells. *PLoS Comput. Biol.* **16**, e1007643 (2020).
55. Bai, W., Sargent, C. J., Choi, J. M., Pappu, R. V. & Zhang, F. Covalently-assembled single-chain protein nanostructures with ultra-high stability. *Nat. Commun.* **10**, 3317 (2019).
56. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675 (2012).
57. da Silva, A. J. et al. Non-conventional induction strategies for production of subunit swine erysipelas vaccine antigen in *rE. coli* fed-batch cultures. *Springerplus* **2**, 322 (2013).
58. Elices, M. et al. Bioinspired fibers follow the track of natural spider silk. *Macromolecules* **44**, 1166–1176 (2011).
59. Lefevre, T., Paquet-Mercier, F., Rioux-Dube, J. F. & Pezolet, M. Review structure of silk by raman spectromicroscopy: from the spinning glands to the fibers. *Biopolymers* **97**, 322–336 (2012).
60. Shao, Z., Vollrath, F., Sirichaisit, J. & Young, R. J. Analysis of spider silk in native and supercontracted states using Raman spectroscopy. *Polymer* **40**, 2493–2500 (1999).
61. Wojdyr, M. Fityk: a general-purpose peak fitting program. *J. Appl. Crystallogr.* **43**, 1126–1128 (2010).

## Acknowledgements

The authors would like to thank Dr. Srikanth Singamaneni and Yixuan Wang for their assistance with the Raman spectrometry. SEM images were acquired within the facilities of IMSE at Washington University. This research was funded by United States Department of Agriculture (grant number 20196702129943 to F.Z.) and National Science Foundation (award numbers DMR-2207879 and OIA-2219142 to F.Z.).

## Author contributions

F.Z. and J.L. conceived the project. J.L. performed plasmid construction, strain engineering, cell culture, protein expression, protein purification, SDS-PAGE, fiber spinning, optical microscopy, SEM imaging, fiber

mechanical testing, Raman analyses and associated analyses. B.J. performed fluorescence imaging. X.C. and H.Y. performed protein expression and protein purification. Y.H. helped with the plasmid design. J.L. and F.Z. prepared the manuscript with comments from all authors.

### Competing interests

The authors declare the following competing financial interest(s): J.L. and F.Z. have filed a provisional patent application based on the protein sequences and methods of fabrication presented in this work (US provisional application No. 63/488,637, filed Mar. 6th, 2023). The remaining authors declare no competing interests.

### Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-023-37563-0>.

**Correspondence** and requests for materials should be addressed to Fuzhong Zhang.

**Peer review information** *Nature Communications* thanks the other anonymous reviewer(s) for their contribution to the peer review of this work. Peer review reports are available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023