Soft Matter

In silico Stress-Strain Measurements on Self-assembled Protein Lattices

SCHOLARONE™
Manuscripts

Soft Matter

ARTICLE TYPE

In silico Stress-Strain Measurements on Self-assembled Protein Lattices

Rachel A. Baarda[a], Tegan L. Marianchuk[b], Michael D. Toney[c], and Daniel L. Cox[a]

Due to their large mechanical strength and potential for functionalization, beta solenoid proteins show promise as building blocks in biomaterials applications such as two- and three- dimensional scaffolds. We have designed simulation models of two-dimensional square and honeycomb protein lattices by covalently linking a beta solenoid protein, the spruce budworm antifreeze protein (SBAFP), to symmetric protein multimers. Periodic boundary conditions applied to the simulation cell allow for the simulation of an infinite lattice. We use molecular dynamics to strain the lattice by deforming the simulation cell and measuring the resulting stress tensor. We evaluate the linear portion of stress-strain curves to extract the corresponding bulk and shear elastic moduli. When strained at a rate of 0.3 nm/ps, the lattices yield a bulk modulus of approximately 3 GPa. This large elastic modulus demonstrates that 2-dimensional structures designed from beta solenoid proteins can be expected to retain the exceptional material strength of their building blocks.

1 Introduction

The development of materials and devices whose structures are precisely controlled down to nanoscale dimensions is a key research area in nanotechnology. In particular, many research groups are currently working to develop so-called decorated scaffolds (molecular lattices bonded to functional structures) with a variety of applications in mind. These applications include sacrificial templates, on which metals are deposited and then the underlying template removed via degradation or dissolution[1]; biological sensors and probes[2]; and passive molecular sieves (see Figure 1). Although scaffold-like nanomaterials can be made synthetically through processes such as electrospinning, many groups choose to utilize existing biological nano-structure and self-assembly processes by constructing lattices from biomolecules, especially DNA and proteins. Self-assembled DNA arrays have been engineered to grow two-dimensional protein and nanoparticle lattices[3,4], but these are not as environmentally robust or industrially scalable as protein based arrays[5]. Several groups have made progress in designing 2-dimensional protein scaffolds[6–10].

We focus here on lattices constructed primarily of proteins belonging to a family known as beta-solenoids. The beta-solenoid protein structure consists of a solenoid-like coil of parallel beta sheets (see Figure 2). Because of their fibrillar shape and the dominance of these beta sheets in their structures, these proteins can be described as amyloid-like, and bear similitude in structure and strength to amyloid proteins such as Curli[11]. This structure is reinforced by hydrogen bonding between beta-sheets and hydrophobic packing in the interior. Previous molecular dynamics studies have shown that members of this family exhibit similar behavior under mechanical stress[12], and are able to sustain exceptionally large axial load[13]. These protein structures also quite robust to a variety of perturbations that often cause less stable proteins to unfold, including high temperatures, extremes of basic or acidic pH, high concentrations of protein denaturants, and additive modifications to the original sequence[14]. This last point is particularly crucial to their implementation as the basis for decorated scaffolds, since many protein-based scaffolds are prone to losing their well-folded structure upon the addition of functional structures[1]. In addition, the hydrogen bonding network allows individual proteins to self-assemble end-to-end into long fibrils. Self-assembly is necessary for realistic production of supramolecular structures from individual molecules.

We anticipate that 2-D lattices constructed from beta-solenoid proteins would be both tunable (in length and functionality) and robust to environmental perturbations. Because the beta-solenoid sequence-to-structure map is regular and well-understood, the length of each protein and therefore the scaffold dimensions can be easily and reliably modified. In addition, the stability of the beta-solenoid structure enhances the decorability of the scaffold through the ability to mutate in functional molecules. Finally, when complemented with other particularly stable proteins as the multimeric junctions, such as those found in thermophilic archea[15], we expect that beta-solenoid-based materials will ex-

[a] Department of Physics, University of California, Davis, California, USA. Fax: 1(530)752-4717; Tel: 1(530)752-1500; E-mail: rabaarda@ucdavis.edu
[b] Department of Physics, Arizona State University, Tempe, Arizona, USA.
[c] Department of Chemistry, University of California, Davis, California, USA.
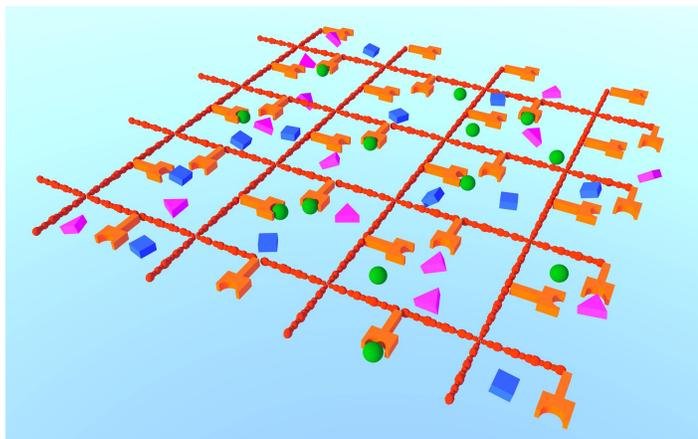
Fig. 1 Cartoon illustration of a potential application of a nano-scale decorated lattice as a passive molecular sieve. A protein lattice (red) is decorated with a functional molecule (orange) that selectively binds to a particular target (green spheres) leaving other molecules (pink triangles, blue squares) free to diffuse through the lattice. As a practical example, the orange molecules could be antibodies and the green spheres antigens, creating a medical assay.

Fig. 2 Cartoon representation of the Spruce Budworm Antifreeze Protein (SBAFP)[18], the beta-solenoid protein used throughout this study. Beta-solenoid protein structure is characterized by parallel beta-sheets coiled around a central axis. Left: side view, illustrating structural motif. Right: front view, illustrating regular cross-section.

hibit exceptionally high material strength.

Where the present research differs from earlier design strategy for two-dimensional protein arrays (c.f., Refs.[7,9,10,16,17]) is in the following regards: (i) the use of modified beta-solenoid proteins fused to archaeal extremophile multimers allows for the design of arrays that can in principle be used in extreme environments. Notably, modified beta-solenoid protein fibrils have been shown to persist in a wide variety of solvent and temperature environments[14]. This could be of relevance to applications in nanoscale filtration in, say, reverse osmosis. It is not clear whether this is possible with other designs. (ii) Because it is easy to augment the length of the modified beta-solenoid fibrils by adding additional repeats, the present approach allows for the design of porous arrays with readily tunable pore sizes, and the open faces of the beta-solenoid proteins can be modified for binding organic or inorganic ligands. (iii) To our knowledge, for no other designed lattices have the two-dimensional elastic properties been simulated and shown to be of a large magnitude.

As the field of engineered biomaterials continues to develop, it is valuable to be able to predict mechanical properties of these proposed novel materials. Although no model can perfectly re-

flect the real-world system, computer simulations are a useful complement to wet bench work in developing these biomaterials. All-atom simulations such as those performed here allow for more detailed knowledge of the system than is typically possible in a wet bench experiment, since every atom is tracked for each ~femtosecond timestep of the simulation. The speed, ease of access, and wealth of quantitative data provided by simulations can be used to advise experimentalists with suggestions such as which proteins to use, how to modify them for optimal performance, and the approximate values to expect from quantitative measurements. This paper describes a technique for measuring stress-strain curves using molecular dynamics simulations with the goal of characterizing the material properties that a nano-scaffold constructed from beta-solenoid proteins would be expected to exhibit.

2 Methods

2.1 Construction of Protein Lattices

The beta-solenoid protein used in this work is the Spruce Budworm Antifreeze Protein (SBAFP), which is available from the Protein Data Bank under the code 1M8N[18]; see Figure 2. The wild-type SBAFP contains a capping region of anti-parallel beta sheets which prohibits fibrillization due to its disruption of the otherwise continuous hydrogen bonding network between parallel beta sheets. When this capping region is removed, it has been shown that fibrillization can take place. Thus, all simulations used the protein structure without this capping region.

We have designed a number of symmetric structures that can form the basis of a protein lattice, including a simple fiber (1-D lattice) and 2-D square and honeycomb lattices which are measured in this work. In the 2-D lattices, beta-solenoid proteins comprise the lattice edges, which are linked together by symmetric multimeric proteins that form the lattice vertices. The beta-solenoid proteins and the symmetric multimers can be covalently bonded together at the time of gene expression[10].

The hydrogen bonding network at the face of each beta-solenoid protein allows assembly between the N-terminus of one protein and the C-terminus of another, and prohibits other configurations (N-to-N or C-to-C). This constraint necessitates a binary alphabet of beta-solenoid building blocks in order to construct a complete lattice: the N-termini of the beta-solenoid proteins in one building block must be exposed with their C-termini covalently bonded to the symmetric multimer, and the C-termini of the other building block type must be exposed with their N-termini covalently bonded to the symmetric multimer. This concept is illustrated in Figure 3. For this reason, it is desirable to identify multimers with both N-terminus and C-terminus close to the exterior of the protein, facilitating covalent linkage (see Figure 4). Complete lattice structures consisting of beta-solenoids bonded to the symmetric multimers are shown in Figure 8.

The symmetric multimers (the "linker" proteins) were identified using the symmetry browser tool on the protein data bank. A trimer with C3 symmetry from the foldon domain of a bacteriophage, pdb code 4NCV[19], was used in the honeycomb lattice. An archaeal proteasome activator tetramer from the extremophile
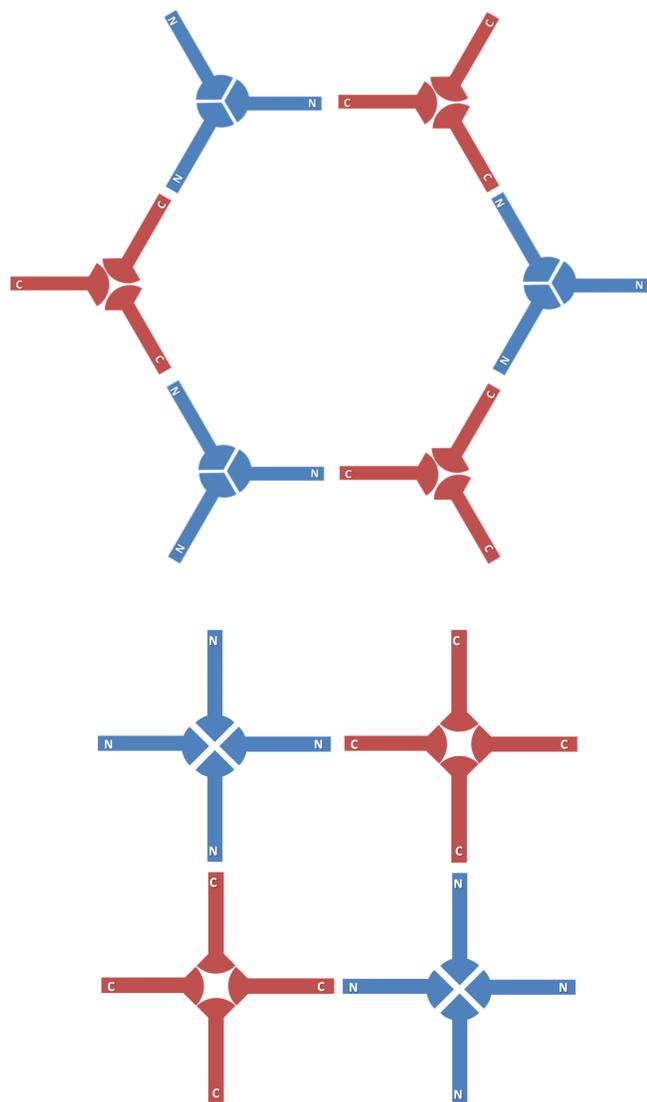
Fig. 3 Binary alphabet of lattice building blocks, illustrated here for the honeycomb lattice (top) and the square lattice (bottom). One block type exposes the C terminal ends of its beta-solenoid proteins, while the other exposes the N terminal ends, allowing for hydrogen bond formation between building blocks.
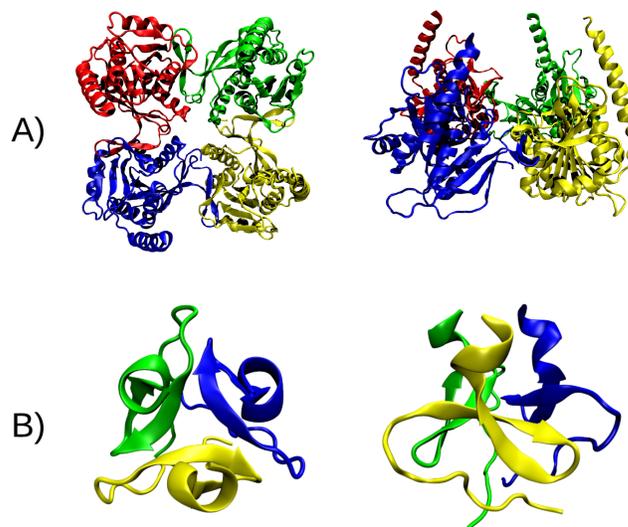
Fig. 4 Symmetric multimers used to join beta-solenoid proteins at lattice vertices. A): tetramer used to form square lattice: archaeal proteasome activator from Pyrococcus furiosus. B): trimer used to form honeycomb lattice: foldon domain of Escheria coli virus T4. Left column: top view, illustrating symmetry. Right column: front view, illustrating exposed terminal ends.

Pyrococcus furiosus with C4 symmetry, pdb code 3VR0[15], was used in the square lattice. These proteins are shown in Figure 4. As discussed above, these proteins were chosen both for their symmetry and the ease of access to both terminal ends of each protein monomer. Note that the tetramer used in the square lattice belongs to a thermophile and is stable at 100° C.

2.2 Elastic Response to Deformation

As a gauge of the material strength of these protein lattices, stress-strain simulations were performed under bulk and shear deformation to determine the corresponding elastic moduli. This section reviews the definition and properties of the stress, strain, and elasticity tensors.

2.2.1 Strain

Strain describes the amount of displacement endured under a given deformation with respect to the original configuration. For example, a 1-D tensile deformation of a rod of length L that results in a displacement of ΔL corresponds to a strain of $\Delta L/L$. Since strain is unitless, it is often expressed in terms such as percent and millimeters per meter.

A given deformation can be described using the deformation map D, which describes the difference between the original configuration S and the deformed configuration S':

$$D = S'(\{x',y',z'\}) - S(\{x,y,z\}) \tag{1}$$

The strain tensor ε can then be defined in terms of the deforma-

tion map as the symmetric sum of derivatives

$$\varepsilon_{ij} = \frac{1}{2}\left[\frac{\partial D_i}{\partial r_j} + \frac{\partial D_j}{\partial r_i}\right] \tag{2}$$

In the case of two-dimensional bulk (uniform) deformation, the relationship between the deformed and original coordinates is given by $x' = x(1+\Delta)$ and $y' = y(1+\Delta)$ where Δ scales the deformation. The deformation map is then given by $D = \Delta(x\hat{x} + y\hat{y})$, yielding the strain tensor $\varepsilon_{ij}^{bulk} = \Delta\delta_{ij}$:

$$\varepsilon^{bulk} = \begin{pmatrix} \Delta & 0 \\ 0 & \Delta \end{pmatrix} \tag{3}$$

Similarly, two-dimensional shear (volume-conserving) deformation can be described by the relationship $x' = x - \Delta \cdot y$ and $y' = y$, yielding $D = -\Delta \cdot y\hat{x}$. Thus the shear strain tensor will be given by Equation 4:

$$\varepsilon^{shear} = \begin{pmatrix} 0 & -\Delta/2 \\ -\Delta/2 & 0 \end{pmatrix} \tag{4}$$

These modes of deformation are illustrated in Figure 5.

2.2.2 Stress

Stress describes a material's response to applied strain. Stress is defined as applied force per cross-sectional area, or equivalently as energy per unit volume, and hence has units of pressure. The energy density interpretation is applied when determining the stress tensor for a microscopic ensemble by finding the derivative with respect to strain of the total kinetic and potential energies per unit volume, considering contributions from each individual atom. This results in a kinetic energy term and a virial term. Section 3.3 discusses the relative contributions of each of these terms.

$$\sigma_{ij}^{\alpha} = -\frac{1}{V}\left(M^{\alpha}v_i^{\alpha}v_j^{\alpha} + \Sigma_{\beta}F_i^{\alpha\beta}r_j^{\alpha\beta}\right) \tag{5}$$

Here V is the volume of the simulation box. Other parameters in this expression are illustrated in Figure 6.

2.2.3 Elasticity

For small deformations, stress is linear in strain: this linear region defines the material's elastic response and is described by the material's elasticity tensor, C. Since both stress and strain are rank-2 tensors, the elasticity tensor must be rank-4 to connect them:

$$\sigma_{ij} = C_{ijkl}\varepsilon_{kl} \tag{6}$$

where i, j, k, l all index $x, y,$ and z. Consequently, the elasticity tensor contains 81 components; however, because all three tensors are symmetric, only 21 of these components are independent. Any symmetry or dimensionality reduction in the lattice comprising the material further reduces the number of independent components in the elasticity tensor: a 2-D square lattice has only 3 independent components, while a 2-D honeycomb lattice has only 2 (see Figure 7). The form of the elasticity tensors for each of these lattices is given by the following (terms coupling to the third dimension have been removed):
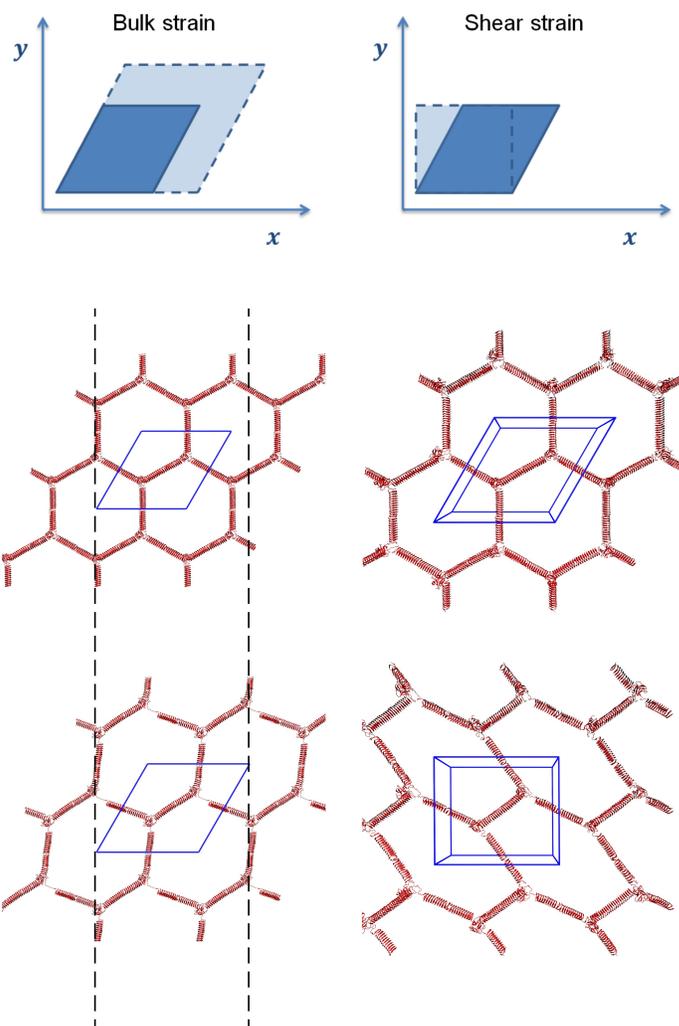
Fig. 5 Applied strain configurations. Left column: bulk (isotropic) strain. Right column: shear (volume-conserving) strain. Honeycomb lattice deformation illustrated below diagrams. Dashed vertical lines in left column illustrate expansion of the unit cell.
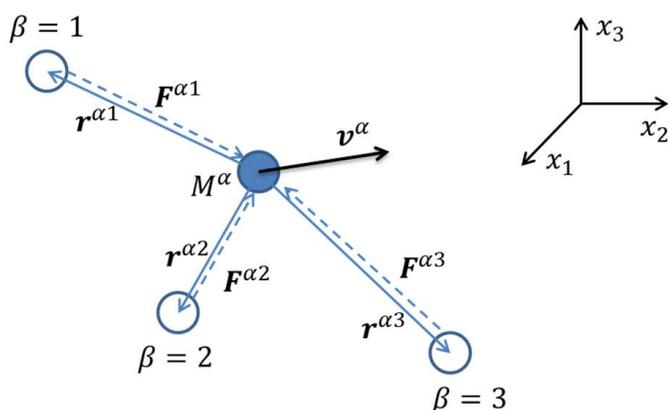
Fig. 6 Illustration of contributions to the stress tensor from a given atom α as expressed in equation 5. M^α: mass of atom α. v^α: velocity of atom α. β indexes over neighboring atoms; $F^{\alpha\beta}$ describes the inter-atomic forces acting at a distance $r^{\alpha\beta}$. Figure modified from reference [20]

Fig. 7 Symmetries of elasticity tensors of honeycomb and square lattices. Red bars indicate couplings to the third dimension which are ignored for 2-D lattices. The 2-D honeycomb lattice has two independent components (C_{11} and C_{12}) and the 2-D square lattice has three ($C_{11}, C_{12},$ and C_{66}). Tensor notation following reference [21].

$$C_{honeycomb} = \begin{pmatrix} C_{11} & C_{12} & 0 \\ C_{12} & C_{11} & 0 \\ 0 & 0 & \frac{1}{2}(C_{11}-C_{12}) \end{pmatrix} \tag{7}$$

$$C_{square} = \begin{pmatrix} C_{11} & C_{12} & 0 \\ C_{12} & C_{11} & 0 \\ 0 & 0 & C_{66} \end{pmatrix} \tag{8}$$

Note that the tensors are expressed using Voigt notation, which contracts the four indices of the elasticity tensor into two indices according to the convention in equation 9:

$$\begin{aligned} 11 &\to 1 \\ 22 &\to 2 \\ 33 &\to 3 \\ 23, 32 &\to 4 \\ 31, 13 &\to 5 \\ 12, 21 &\to 6 \end{aligned} \tag{9}$$

Knowing the form of both the elasticity tensor and the strain tensor corresponding to a particular deformation allows for the determination of the stress tensor that describes the response to the deformation. Expanding Equation 6 in matrix multiplication form (again, removing terms that couple to the third dimension) yields the following:

$$\begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_6 \end{pmatrix} = \begin{pmatrix} C_{11} & C_{12} & C_{16} \\ C_{21} & C_{22} & C_{26} \\ C_{61} & C_{62} & C_{66} \end{pmatrix} \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ 2\varepsilon_6 \end{pmatrix} \tag{10}$$

Then for a deformation of the form

$$\begin{pmatrix} \varepsilon_{11} \\ \varepsilon_{22} \\ 2\varepsilon_{12} \end{pmatrix} = \begin{pmatrix} \alpha \\ \alpha \\ 2\beta \end{pmatrix} \tag{11}$$

where the diagonal elements α describe the amount of bulk strain and the off-diagonal elements β describe the shear strain, the stress response exhibited by each of the lattices can be found by referencing Equations 7 and 8:

$$\sigma_{honeycomb} = \begin{pmatrix} C_{11} & C_{12} & 0 \\ C_{12} & C_{11} & 0 \\ 0 & 0 & \frac{1}{2}(C_{11}-C_{12}) \end{pmatrix} \begin{pmatrix} \alpha \\ \alpha \\ 2\beta \end{pmatrix} \tag{12}$$

$$\sigma_{square} = \begin{pmatrix} C_{11} & C_{12} & 0 \\ C_{12} & C_{11} & 0 \\ 0 & 0 & C_{66} \end{pmatrix} \begin{pmatrix} \alpha \\ \alpha \\ 2\beta \end{pmatrix} \tag{13}$$

The bulk modulus, K, and shear modulus, G, for a two-dimensional isotropic material are defined as follows:

$$\begin{aligned} 2K &\equiv C_{11} + C_{12} \\ 2G &\equiv C_{11} - C_{12} \end{aligned} \tag{14}$$

This allows the stress tensors to be re-written as

$$\sigma_{honeycomb} = \begin{pmatrix} 2K\alpha & 2G\beta \\ 2G\beta & 2K\alpha \end{pmatrix} \tag{15}$$

$$\sigma_{square} = \begin{pmatrix} 2K\alpha & 2C_{66}\beta \\ 2C_{66}\beta & 2K\alpha \end{pmatrix} \tag{16}$$

Hence the diagonal elements of the stress tensors are equal to

twice the bulk modulus K times the amount of bulk strain α that has been applied. Similarly, the off-diagonal elements of the honeycomb lattice stress tensor are equal to twice the shear modulus G of the lattice times the amount of shear strain β that has been applied. Due to the lesser symmetry of the square lattice compared to the honeycomb lattice (see Figure 7), the off-diagonal elements of the square lattice do not contain the shear modulus G. Nevertheless, since the C_{66} element describes coupling to the shear deformation, it can be loosely interpreted as a shear modulus. It is in this sense that we use it in this work.

2.3 Simulations

The molecular building and modeling suite YASARA[22] was used to remove the anti-fibrillizing cap, add covalent bonds between the beta-solenoid proteins and the symmetric multimers, and align the structures to correctly tile the lattice in two dimensions. The Gromacs[23–25] molecular dynamics software package was used for all simulations.

Pressure coupling was performed with the anisotropic Parrinello-Rahman barostat[26]. The isothermal compressibility in the directions along which strain was applied (i.e., in the plane of the lattice) was set to zero so that this direction would not be coupled to the pressure bath. The compressibility in the direction transverse to the strain were set to that of water (4.5×10^{-5} bar^{-1}). The reference pressure was set to 1 bar. The Nose-Hoover thermostat algorithm[27,28] was used to maintain a constant temperature of 300 Kelvin.

The Charmm27 force field[29] was used for all simulations. Protein images were produced using the visualization software VMD[30]. The protein was solvated in explicit water molecules using the TIP3P (transferable intermolecular potential with 3 points) potential.

The protein was placed in a simulation box with a sidelengths and angles calculated to tile the lattice in the x- and y-directions. Periodic boundary conditions were applied to all boundaries of the box, making the system effectively infinite in the plane of the lattice (see Figure 8). The box height was made to be sufficiently large as to avoid interactions between the protein's periodic images in the out-of-plane directions (i.e., larger than the long-range interaction cut-off radii).

Prior to the production runs, steepest descent and L-BFGS minimization were performed to relax the system into a low potential energy configuration. Temperatures were instantiated by randomly prescribing values from a Boltzmann distribution to each atom. Temperature was then allowed to equilibrate under the velocity-rescale thermostat. Finally, pressure was allowed to equilibrate under the Parrinello-Rahman barostat.

Strain was applied to the system by deforming the simulation box at a constant rate throughout the simulation using Gromacs' built-in molecular dynamics parameter option "deform". In each integration timestep, the atomic coordinates were first updated according to the leapfrog force integration algorithm. Velocities and atomic coordinates were then updated according to temperature and pressure coupling, respectively. The atomic coordinates and box lengths were then rescaled according to the specified
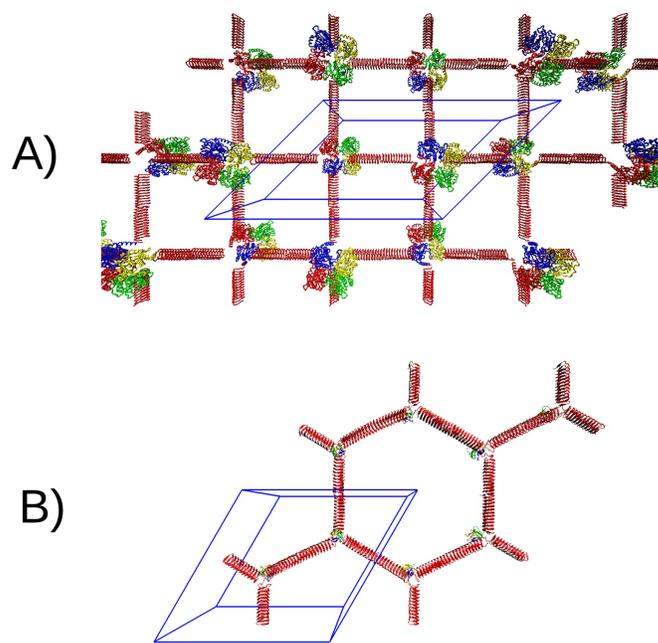
Fig. 8 Simulation box geometries used to periodically tile unit cell onto lattice. A): square lattice. B): honeycomb lattice. Note that in each case the unit cell consists of one representative of each of the two building blocks discussed in Figure 3.

deformation (either bulk or shear). These simulations were performed at several deformation speeds for each lattice.

After each production run, the relevant elements of the stress tensor were calculated using Gromacs' "gmx energy" program. Stress-strain curves were plotted, and a linear regression with a fitting constraint of zero intercept was applied to points corresponding to less than one percent strain (see Figure 9). The slope of this curve was used to find the lattice's elastic modulus: K for bulk strain and G for shear strain, according to equations 15 and 16. These elastic modulus values are plotted against the logarithm of the strain rate in Figure 11.

3 Results and Discussion

3.1 Small-Strain Regime

As expected, the lattice's initial stress response is linear in strain, as emphasized in the right column of Figure 9. We examine the mechanisms that produce the elasticity values exhibited by this material, and find that predominant contribution during small-strain bulk deformation arises from contraction and lengthening of the solenoids that compose the lattice. This section illustrates the dominance of this mechanism through the Cauchy-Born approximation and a comparison of bulk moduli across the two lattices, makes general comments about trends in the elastic moduli, then briefly comments on shear strength. Finally, a comparison to some other materials is made.

In order to apply the Cauchy-Born approximation to this system, we formulate it as an idealized lattice composed of compressible rods with spring constant k, allowing the bulk modulus to be found using purely geometric means. In the case of the hon-

eycomb lattice whose rods each have equilibrium length s and are stretched to a length $s' = s(1 + \alpha)$, the internal energy U in each unit cell will be given by

$$U = 6 \cdot \frac{1}{2} k (s\alpha)^2 \tag{17}$$

where the factor of 6 refers to the number of rods per unit cell (see Figure 8 B). The unit cell has sidelengths $L = 2\sqrt{3}s$ and inscribed angle $60°$, yielding an area $A = L^2 \sin(60°) = 6\sqrt{3}s^2$. Hence the areal energy density will be

$$u = \frac{U}{A} = \frac{k\alpha^2}{2\sqrt{3}} \tag{18}$$

Finally, differentiating the areal energy density with respect to the applied strain yields the resultant stress:

$$\sigma = \frac{\partial u}{\partial \varepsilon} = \frac{\partial u}{\partial \alpha} = \frac{1}{\sqrt{3}} k\alpha \tag{19}$$

Comparing this result to equation 15, we see that in terms of the bulk modulus the equivalent deformation would yield

$$\sigma = 2K\alpha \tag{20}$$

Hence, in this idealized lattice, the relationship between the spring constant of the individual rods and the elastic modulus of the entire lattice will be

$$K = \frac{1}{2\sqrt{3}} k \tag{21}$$

In order to make this comparison, a single beta-solenoid was subjected to tensile extension and analyzed in the same manner as the lattice* in order to extract its effective spring constant. The results are plotted in Figure 10 A). Circular data points indicate the bulk modulus as derived from this spring constant (i.e., multiplied by the geometric factor $\frac{1}{2\sqrt{3}}$). Triangular data points indicate the bulk modulus as derived from expansion simulations of the entire lattice; both of these moduli are plotted against the logarithm of the speed at which the structure was deformed. Although the single beta-solenoid produced a more erratic curve, its similitude to the lattice's curve indicates that the idealized rod-lattice is a reasonable model for the honeycomb lattice under small-strain bulk deformation, and therefore that the predominant elasticity mechanism of the lattice in this case is determined by the spring-like behavior of the constituent beta-solenoids.

The dominance of the mechanism of individual beta-solenoids is further indicated in Figure 11, which shows that the honeycomb lattice and square lattice exhibit nearly identical bulk moduli in the range of strain rates tested. This would not be expected if a mechanism other than beta-solenoid stretching dominated at low
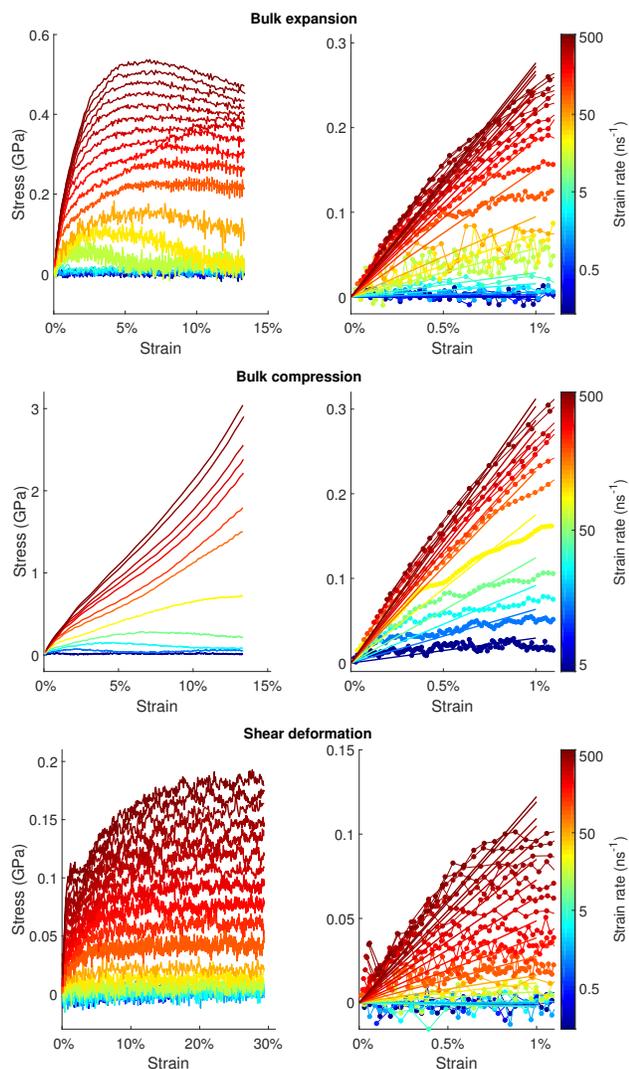
Fig. 9 Stress-strain curves measured under each deformation mode for a range of strain rates as indicated by colorbar. Left column: stress-strain behavior over the duration of each simulation. Right column: detail of curve during initial linear response region, showing simulation data (points) and linear fits (solid lines). Plots shown measured on honeycomb lattice.

---

* The stress-strain curves produced from the single beta-solenoid simulations were found to have a slight non-zero offset, seemingly indicating that the starting structure was somewhat out of equilibrium. For this reason, these linear fits were not constrained to yield zero intercept. Correspondingly, the lattice stress-strain curves were re-fit without a zero-intercept constraint in order to yield a more consistent comparison. This results in some slight differences in the shape of the bulk modulus-strain rate curve compared to Figure 11 .

A)

B)
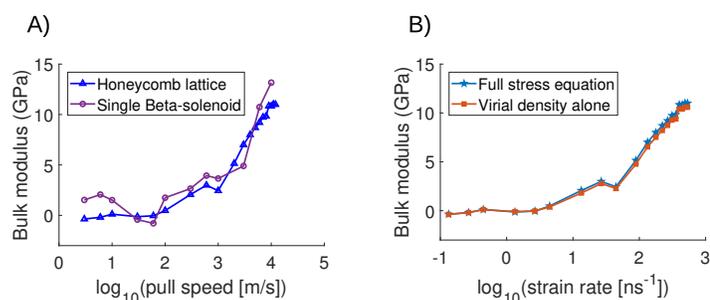
Fig. 10 Contributions to the honeycomb lattice's elastic modulus value. A): comparison between the honeycomb lattice bulk modulus as estimated by single beta-solenoid tensile extension simulations (circular data points) and as measured by whole-lattice bulk expansion simulations (triangular data points). B): Comparison of the elastic modulus of the honeycomb lattice under bulk expansion as inferred from two measures of stress. The curves shown represent the diagonal component of the virial density tensor alone, and the full stress tensor of equation 5 which also includes a kinetic term. The data points for each curve represent twice the slope of that measure when plotted against strain, as a function of the logarithm of the strain rate.

Fig. 11 Elastic modulus for bulk and shear deformation plotted against the logarithm of the strain rate, demonstrating increasing stiffness as strain rate increases.

Material	$K^{\star}$ or $E^{\dagger}$ (GPa)	Pulling speed (m/s)
Honeycomb lattice*	2.9	300
Square lattice*	3.2	300
Suffolk sheep wool† [34,35]	0.5 - 1	5×10^{-3}
Collagen fibers† [36]	0.3 - 0.5	1.7×10^{-4}
Silkworm silk† [37]	8 - 18	8.5×10^{-5}
Glass-epoxy composite† [38]	6.2	4.2
Carbon-epoxy composite† [39]	7.3 - 8.2	2.8
Graphene† [40]	1,000.0	2.3×10^{-7}

Table 1 Elastic modulus values of protein lattices compared with select other materials for context. Star ($\star$) superscripts indicate the bulk modulus, and dagger ($\dagger$) superscripts indicate the material's Young's modulus; Young's modulus data are more readily available in the literature.

lattice (see supplemental materials). Another option for strengthening these lattices is choosing symmetric multimers with large inherent strength. Comparing the trimeric protein used in the honeycomb lattice to the tetrameric protein used in the square lattice (Figure 4), it is clear that the tetramer is much larger and has more internal hydrogen bonds reinforcing its structure. Correspondingly, the shear modulus of the square lattice is noticeably larger than that of the honeycomb lattice, even though the bulk modulus of the square lattice is roughly equal to that of the honeycomb lattice (see Figure 11). This appears to indicate the efficacy of strengthening the overall lattice by choosing inherently stronger constituent proteins.

For context, the elastic moduli of the simulated protein lattices tested here are compared against that of other materials, both natural and man-made, in Table 1. Due to the difficulty in finding measurements of the bulk modulus (K) of materials in which the strain rate was stated, the values listed in the table indicate materials' Young's modulus (E) except in the case of the two simulated lattices.

In the interest of providing a more direct comparison, strain rates that were listed as 1/s were multiplied by the sample gauge length to yield a result in m/s. In the case of the sheep wool measurement, the elastic modulus was measured in grams per grex (g/gx); grex is a unit of linear density used in textile science. To convert g/gx to Pa, one must multiply by the mass density of the sample and the gravitational acceleration $g = 9.8\text{m/s}^2$ (see e.g.[33]). A wool density of 1 g/cm^3 was assumed for this conversion.

3.2 Large-Strain Regime

When the amount of applied strain is large enough, the structure of the protein lattices begins to visibly deform, as shown in Figure 12. Under bulk compression, the beta-solenoids are forced to bend out of the plane of the lattice. This transverse degree of freedom allows compression to continue while keeping the overall folded structure of the lattice intact. In contrast, under bulk expansion and shear deformation, large strains cause the lattice to unravel.

As these unfolding events occur, stabilizing forces such as hy-

strain values, such as linker protein deformation, since the linker proteins differ substantially between the two lattices.

At small deformation speeds, stress-strain curves exhibited large noise (see Figure 9), especially in the case of the square lattice. Hence, these curves were measured in a smaller range of strain rates in the neighborhood of 50 per nanosecond. Note that in all cases, the material becomes less compliant (stiffer) at faster deformation speeds, consistent with the results of many other material studies[31,32] (see Figure 11). Note also that expansion and compression of the honeycomb lattice yield quite similar bulk modulus values, indicating local symmetry in the potential around equilibrium.

Although the lattices' bulk moduli benefit from the large intrinsic strength of the beta-solenoid proteins, the shear moduli depend to a large degree on the linker proteins' resistance to torsion. Figure 11 shows that the bulk modulus is significantly greater than the shear modulus, especially in the case of the honeycomb lattice. In order to further optimize these structures, it is desirable to reinforce these weak links. One approach considered in this work was inserting cysteine residues to the symmetric multimers in order to form disulfide bonds near the lattice vertices. Curiously, this did very little to impact the elastic modulus of the

drogen bonds between beta-sheets are removed. This results in a plateau in the strength of the inter-atomic restoring forces that contribute to the stress tensor, as seen in the left column of Figure 9. Conversely, the bulk compression stress-strain curve lacks such a plateau by virtue of the continuous presence of inter-atomic repulsion forces.

3.3 Contributions to Endured Stress

Figure 10 B) illustrates the contributions of each term in equation 5, namely the virial density and the kinetic energy density, in the case of bulk expansion of the honeycomb lattice. The *xx* element of both the virial tensor and the complete stress tensor were used in this comparison. Note that both the virial and the kinetic energy are multiplied by a factor of two when they appear in equation 5; this factor was included in this comparison. Also note that the shape of the "full stress equation" curve differs from Figure 11 because the linear fit in this case was not constrained to have zero offset to be more consistent with the other ensemble variables (similar reasoning was applied to the analysis of the single beta-solenoid simulations - see footnote in section 3.1).

The virial density term nearly exactly reproduces the curve formed from the full stress equation, indicating that this term is by far the majority contributor to the elastic modulus, producing only a slight under-estimate. Hence, kinetic energy is a negligible contribution to the stress endured by the lattice during these simulations.

4 Conclusion

We have demonstrated a method of using the molecular dynamics package GROMACS to perform stress-strain simulations on atomic systems to find values for the system's bulk and shear elastic moduli. Through this method, we have shown that these beta-solenoid-based protein lattices exhibit large mechanical strength, supporting the premise that these are promising candidates for use in robust biomaterials. Without optimizing the strength of the multimeric linkers, we already achieve respectable values for the bulk and shear modulus. Thus, the combination of tunable, highly functionalizable beta solenoid proteins possessing known structure/sequence maps with robust physical and mechanical properties with natural symmetric oligomers evolved to survive severe conditions in extremophiles opens a new avenue for designed biomaterial self-assembly.

5 Conflicts of Interest

There are no conflicts of interest to declare.

6 Acknowledgements

References

1 D. N. Woolfson and Z. N. Mahmoud, *Chemical Society Reviews*, 2010, **39**, 3464.
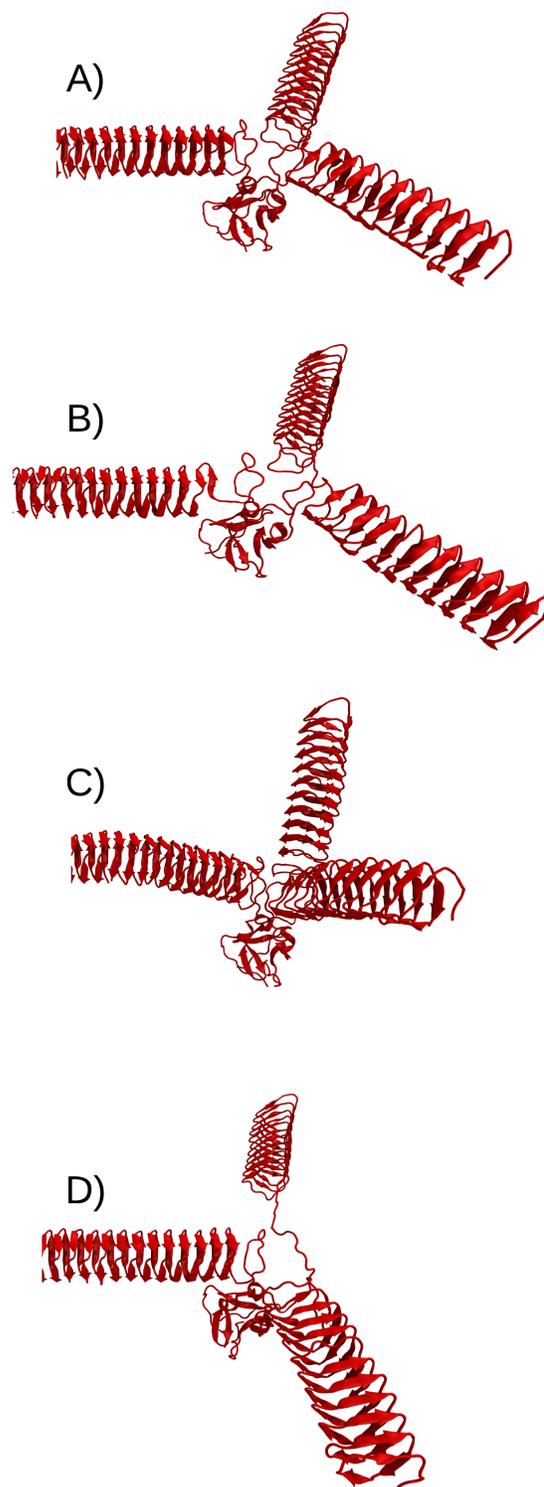
Fig. 12 Snapshots of a single vertex of the honeycomb lattice near the end of typical simulations illustrating modes of failure of the lattice during each type of deformation. A): equilibrium configuration. B): bulk expansion. C): bulk compression. D): shear deformation. Note that both bulk expansion and shear deformation cause structural unraveling near the beta-solenoid-linker junction, whereas bulk compression causes the beta-solenoids to bend out-of-plane without unraveling.

2   G. Colombo, P. Soto and E. Gazit, *Trends in Biotechnology*, 2007, **25**, 211–218.

3   H. Y. Li, S. H. Park, J. H. Reif, T. H. LaBean and H. Yan, *Journal of the American Chemical Society*, 2004, **126**, 418–419.

4   H. Yan, P. Yin, S. H. Park, H. Y. Li, L. P. Feng, X. J. Guan, D. G. Liu, J. H. Reif and T. H. LaBean, in *Self-assembled DNA structures for nanoconstruction*, ed. W. Fritzsche, 2004, vol. 725, pp. 43–52.

5   C. Castro and et al., *Nature Methods*, 2011, **8**, 221–229.

6   J. T. MacDonald, B. V. Kabasakal, D. Godding, S. Kraatz, L. Henderson, J. Barber, P. S. Freemont and J. W. Murray, *Proceedings of the National Academy of Sciences*, 2016, **113**, 10346–10351.

7   F. Baneyx and J. F. Matthaei, *Current Opinion in Biotechnology*, 2014, **28**, 39–45.

8   J. D. Brodin, X. I. Ambroggio, C. Tang, K. N. Parent, T. S. Baker and F. A. Tezcan, *Nature Chemistry*, 2012, **4**, 375–382.

9   J. F. Matthaei, F. DiMaio, J. J. Richards, L. D. Pozzo, D. Baker and F. Baneyx, *Nano Letters*, 2015, **15**, 5235–5239.

10 J. C. Sinclair, K. M. Davies, C. Vénien-Bryan and M. E. M. Noble, *Nature Nanotechnology*, 2011, **6**, 558–562.

11 E. P. Debenedictis, D. Ma and S. Keten, *RSC Advances*, 2017, **7**, 48102–48112.

12 M. Solar and M. J. Buehler, *Nanotechnology*, 2014, **25**, 1–8.

13 S. Keten and M. J. Buehler, *Computer Methods in Applied Mechanics and Engineering*, 2008, **197**, 3203–3214.

14 Z. Peng, M. Peralta and M. Toney, *Biochemistry*, 2017, **56**, 6041–6050.

15 K. Kumoi, T. Satoh, K. Murata, T. Hiromoto, T. Mizushima, Y. Kamiya, M. Noda, S. Uchiyama, H. Yagi and K. Kato, *PLoS ONE*, 2013, **8**, 1–9.

16 S. Gonen, F. DiMaio, T. Gonen and D. Baker, *Science*, 2015, **348**, 1365–1368.

17 Y. Suzuki, G. Cardone, D. Restrepo, P. D. Zavattieri, T. S. Baker and F. A. Tezcan, *Nature*, 2016, **533**, 369–373.

18 E. Leinala, P. Davies, D. Doucet, M. Tychenko, V. Walker and Z. Jia, *J. Biol. Chem.*, 2002, **277**, 33349–33352.

19 A. Berthelmann, J. Lach, M. A. Gräwert, M. Groll and J. Eichler, *Organic & biomolecular chemistry*, 2014, **12**, 2606–14.

20 S. J. Frankland, V. M. Harik, G. M. Odegard, D. W. Brenner and T. S. Gates, *Composites Science and Technology*, 2003, **63**, 1655–1661.

21 J. Nye, *Press (clarendon) London*, 1957.

22 E. Krieger and G. Vriend, *Bioinformatics (Oxford, England)*, 2014, **30**, 2981–2982.

23 B. Hess, C. Kutzner, D. van der Spoel and E. Lindahl, *J. Chem. Theory comput.*, 2008, 435–447.

24 D. van der Spoel, E. Lindahl, B. Hess and the GROMACS development team, *www.gromacs.org*, 2013.

25 *SoftwareX*, 2015, **1-2**, 19–25.

26 M. Parrinello and A. Rahman, *Journal of Applied physics*, 1981, **52**, 7182–7190.

27 S. Nosé, *Molecular physics*, 1984, **52**, 255–268.

28 W. G. Hoover, *Physical review A*, 1985, **31**, 1695.

29 B. Brooks, C. Brooks, A. Mackerell, L. Nilsson, R. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch and et al., *J. Comput. Chem.*, 2009, **30**, 1545–1614.

30 W. Humphrey, A. Dalke and K. Schulten, *Journal of Molecular Graphics*, 1996, **14**, 33–38.

31 A. Gautieri, M. Buehler and A. Redaelli, *Journal of the Mechanical Behavior of Biomedical Materials*, 2008, 130–137.

32 G. C. Jacob, J. M. Starbuck, J. F. Fellers, S. Simunovic and R. G. Boeman, *Journal of Applied Polymer Science*, 2004, **94**, 296–301.

33 G. Susich, *Mechanical properties of decrystallized cotton*, Quartermaster research and development center natick ma technical report, 1953.

34 R. O'Connell and A. Yeiser, *Textile Research Journal*, 1954, **24**, 629–632.

35 H. Preusser, R. O'Connell, A. Yeiser and H. Lundgren, *Textile Research Journal*, 1954, **24**, 118–122.

36 G. D. Pins, D. L. Christiansen, R. Patel and F. H. Silver, *Biophysical journal*, 1997, **73**, 2164.

37 P.-R. J., V. C., L. J. and E. M., *Journal of Applied Polymer Science*, 1998, **70**, 2439–2447.

38 J. M. Lifshitz, *Journal of Composite Materials*, 1976, **10**, 92–101.

39 L. G. Melin and L. E. Asp, *Composites Part A: Applied Science and Manufacturing*, 1999, **30**, 305 – 316.

40 C. Lee, X. Wei, J. W. Kysar and J. Hone, *science*, 2008, **321**, 385–388.

We build a model 2-D nano-scaffold from beta-solenoid proteins fused with symmetric microbial multimers and characterize it using stress-strain simulations