In Silico Stress-Strain Measurements of a β-Solenoid Protein Lattice

Tegan L. Marianchuk[1], Rachel A. Baarda[2], and Daniel L. Cox[2]
[1]*Department of Physics, Arizona State University*
[2]*Department of Physics, University of California, Davis*

Abstract
β-solenoid proteins are exceptionally strong biopolymers, so measuring their bulk mechanical properties is useful for ascertaining their viability for biomaterials application. I have computationally engineered a protein lattice by covalently binding the β-solenoid spruce budworm antifreeze protein, SBAFP, (1M8N) and the symmetric archaeal protein (3VR0) into a 2-dimensional square geometry. Periodic boundary conditions were applied to the unit cell to form an infinitely repeating lattice and implicit waters were simulated to eliminate the solvent strain from the results. Utilizing molecular dynamics (MD) software, stress was applied to the system and computed as twice the difference between the virial and kinetic energy over the volume of the simulation cell. The elastic region of stress-strain curves was evaluated to obtain a bulk modulus of 6.44-10.71 GPa and shear modulus of 2.41-5.14 GPa over a range of pulling velocities, 500-1500 m/s. The calculated material strength of the engineered 2-D protein lattice is comparable to that of spider silk and much greater than that of bacterial S-layers.

Introduction
Naturally occurring protein lattice structures called bacterial S-layers have been discovered on the surface of bacteria and modified for application as nanoscale biomaterials and devices[6]. Once removed from the surface of bacteria they will reassemble in solution and adhere to various surfaces. The first application of these S-layers was in the production of ultrafiltration membranes with molecular sieving properties[1]. By fusing deliberate functional domains, the S-layers have proven useful in a multitude of biotech applications from drug delivery systems to biochip development. Motivated by the promise that these naturally occurring S-layers have shown, we have developed a method for the design of protein lattice geometries that can be willfully and precisely manufactured with nanoscale precision. The advantage of this bottom up engineering approach allows for tunable mechanical and functional application.

The development of highly customizable materials and devices using molecular building blocks is a prominent research area in nanotechnology, but remains a somewhat elusory objective. Biological lattices that act as platforms ornamented with functional domains—called decorated bioscaffolds—are of interest for biomedical and environmental applications like biodegradable templates for synthesizing inorganic material or carrying biomolecular cargo; biosensors and probes; and molecular sieves[10]. While synthetic materials can be utilized, we choose to exploit the self-assembly and existing nano-structures of proteins to construct tailored lattices.

In particular, we investigate a class of mechanically and environmentally robust protein aggregates called β-solenoids. Their tertiary structure forms a rigid, rod-like structure reinforced by interior hydrophobic packing and a secondary structure composed of hydrogen bonds connecting its coil-like layers[3]. This robustness allows them to remain folded against

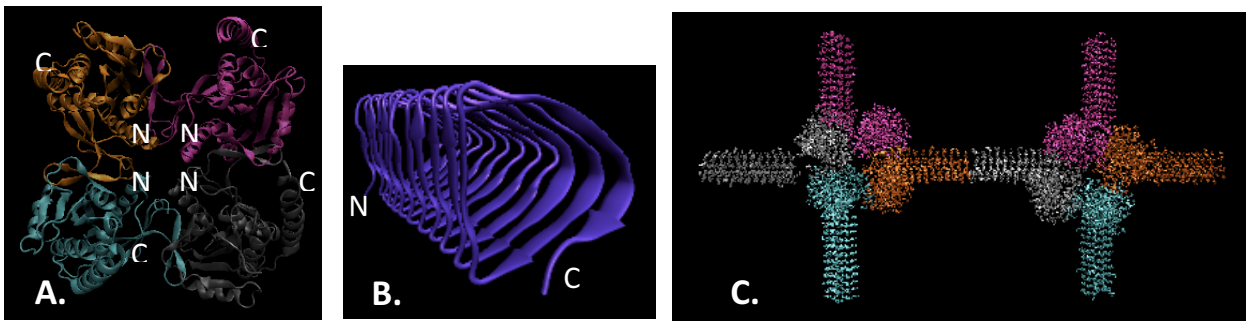
perturbations like the addition of protein denaturant, uric acid, extreme temperatures, and sequence modifications[4]. The ability to create decorated scaffolds is dependent on the maintenance of the β-solenoid's stable, folded structure with the implementation of accessory functional groups to the original amino acid sequence. This property, in addition to the hydrogen bonding network that allows for the self-assembly of individual monomers into long protein fibrils, make β-solenoids ideal candidates for the molecular scaffold structures used in bionanotechnology.

As computational modeling methods for large and complex systems have improved and technologies in engineered biomaterials grows, the accuracy and value of these simulation models have increased to better complement, predict, and demonstrate experimental work in developing these biomaterials. Experimental techniques including protein expression, purification, and characterization take significant time and resources that can be reduced with the employment of computational methods. Using downloadable, virtual proteins has the advantage of variation and accessibility that is limited by experiments performed in a lab. While simulation procedures simplify realistic conditions, they still provide predictive results that can inform experimentalists on how to modify their protein system for optimal performance and produce approximate values for quantitative analysis. Each of the system's atoms in the all-atom simulations used is tracked at every femtosecond time step which allows for a detailed assessment of the system's properties. In this investigation, the mechanical properties that a bioscaffold with square geometry composed primarily of β-solenoids is expected to exhibit are extrapolated from simulated stress-strain measurements.

Methods

SYSTEM

The protein used as the linker for the protein lattice unit is an archaeal proteasome activator from *Pyrococcus furiosus,* (Protein Data Bank code 3VR0). It is a hyperthermophile; therefore, it can withstand extremely high temperatures upwards of 100 degrees Celsius. It has a four-fold rotation axis—C4 symmetry—which is ideal for tiling a square lattice (Figure A). The protein used as the arms of the unit is a beta-solenoid, SBAFP (PDB code 1M8N), which has a triangular cross section and a rod-like tertiary structure resulting from the hydrogen binding of its secondary structure (Figure B). Wild-type SBAFP contain a capping region that prohibits fibrilization which has been removed so that fibrilization can occur. The resulting covalently bound unit is composed of two linkers and eight SBAFPs as depicted in Figure C.



*The N- and C-termini of the proteins have been labeled accordingly in Figures A and B

Due to the hydrogen bond network of the β-solenoids, there is a preferential N-terminus to C-terminus binding directionality for assembly that prohibits alternative binding configurations. Therefore, the N-terminus of each β-solenoid must be covalently bound to the C-terminus of one protein linker. The C-terminus of the β-solenoid is then free to covalently link to the nearest exposed N-terminus of another β-solenoid and the C-terminus of that β-solenoid is covalently bound to the N-terminus of another linker protein. Thus, in addition to its symmetry, the linker protein was chosen because of its terminal accessibility as labeled in Figure A.

MEASUREMENTS/CALCULATIONS

To quantify the material strength of the square protein lattice, it was subjected to stress-strain simulations in which it underwent bulk and shear deformation to determine its corresponding elastic moduli. This section reviews the definitions of stress and strain and details bulk and shear deformation.

Stress describes a material's response to an applied strain. Stress has the units of pressure as it is defined by an applied force over a change in cross-sectional area or equivalently, energy per unit volume. To determine the stress tensor of a microscopic ensemble, the energy density interpretation is used in which the derivative with respect to strain is found for the combined kinetic and potential energies per unit volume, accounting for the contribution of each individual atom in the ensemble. The stress tensor, σ, is pictorially described by Figure 1 and is represented as follows:

$$\sigma = \frac{Energy}{Vol_{protein}} = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{pmatrix}$$

- Bulk stress
- Shear stress

*The black components of the stress tensor are zero as the protein lattice is not deformed in those dimensions

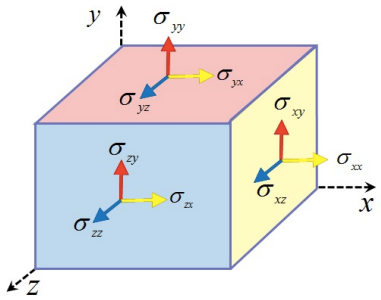


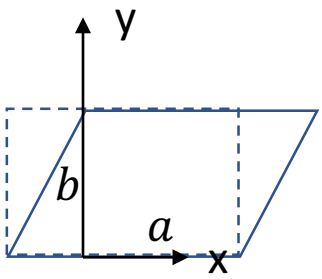
Figure 1: Stress tensor in 3-D

Strain describes the magnitude of deformation or a ratio of displacement in the dimension of deformation to the original dimension. For instance, the amount of strain on a 1-D rod of length L under deformation conditions that result in a displacement of ΔL would be expressed as ΔL/L. Strain is typically dimensionless and is expressed as a percent. The strain tensor is mathematically defined as the symmetric sum of derivatives:

$$\varepsilon_{ij} = \frac{1}{2}\left(\frac{\partial \mu_i}{\partial r_i} + \frac{\partial \mu_j}{\partial r_j}\right) \quad (1)$$

Shear is a type of strain described as the lateral deformation of a material. The simulated shearing (volume-conserving) can be illustrated by the following deformation map in which the solid line represents the original configuration of the material, the dotted outline representing its resulting

configuration, and sides a and b representing the material's x and y positions, respectively:



The coordinates are transformed as follows, in which the primed values represent the coordinates resulting from shear deformation:

$$x' = x - \mu y \; ; \; \mu = \frac{v_x t}{b}$$

$$y' = y$$

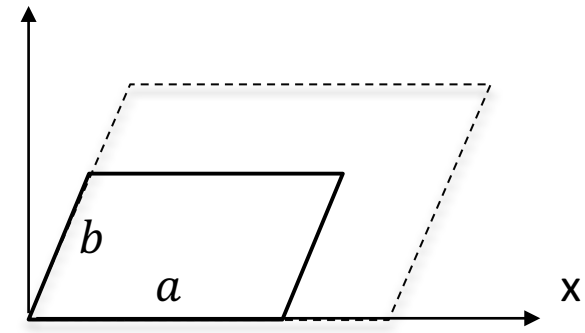
Figure 2: Shear deformation map

The strain tensor mathematically is represented by:

$$\epsilon_{shear} = \begin{pmatrix} 0 & -\frac{v_x t}{2b} & 0 \\ -\frac{v_x t}{2b} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (2)$$

Bulk is a type of strain described by the uniform deformation of a material. The simulated bulk deformation is described by Figure 3 in which the solid line represents the original configuration of the material, the dotted outline representing its resulting configuration, and sides a and b representing the material's x and y positions, respectively:

y



The coordinates are transformed as follows, in which the primed values represent the coordinates resulting from bulk deformation:

$$x' = x + \mu x \; ; \; \mu = \frac{v_x t}{a}$$

$$y' = y + \mu y \; ; \; \mu = \frac{v_y t}{b}$$

Figure 3: Bulk deformation map

The bulk strain tensor is mathematically represented by:

$$\epsilon_{bulk} = \begin{pmatrix} \frac{v_x t}{2a} & 0 & 0 \\ 0 & \frac{v_y t}{2b} & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (3)$$

For small deformations, stress is linear in strain and this linear region defines the material's elastic response, which is described by its elasticity tensor, C, where *i,j,k,l* index *x,y,z*:

$$\sigma_{ij} = C_{ijkl}\epsilon_{kl} \tag{4}$$

For the 2-dimensional square lattice, the elasticity tensor is mathematically represented by:

$$C_{tetragonal} = \begin{pmatrix} C_{1111} & C_{1122} & 0 \\ C_{1122} & C_{1111} & 0 \\ 0 & 0 & C_{1212} \end{pmatrix} \tag{5}$$

SOFTWARE/SIMULATION

The visual molecular dynamics graphics software, VMD, was utilized to identify the N and C termini of the linker and β-solenoid proteins. Visual representations of the protein lattice as well as the molecular dynamics simulations produced in GROMACS were created with VMD.

YASARA, a molecular graphics, modeling and simulation software program, was used to remove the caps from the β-solenoids and covalently bind the unit illustrated in Figure C. The PDB protein structure coordinates were uploaded to this software and aligned with a C++ scripted template to create a uniform tiling pattern for the 2-dimensional square lattice.

Non-equilibrium pulling simulations were used to probe the mechanical properties of the protein lattice. GROMACS 4.6.7 is the molecular dynamics software package used to conduct all the simulations. The Parinello-Rahman barostat was used for anisotropic pressure coupling whose matrix equation of motion is[8]:

$$\frac{d\boldsymbol{b}^2}{dt^2} = V\boldsymbol{W}^{-1}\boldsymbol{b}^{-1}(\boldsymbol{P} - \boldsymbol{P}_{ref}) \tag{6}$$

Where **b** is the simulation box matrix, V is the simulation box volume, **W** is a parameter matrix that controls the coupling strength and box deformation restrictions, and **P** and $\mathbf{P_{ref}}$ are the current and reference pressure matrices, respectively. The isothermal compressibility in the directions transverse to the strain were set to that of water, 4.5 x 10^{-5} bar^{-1} and along the directions of the applied strain (i.e. the plane of the lattice) were set to zero to prevent coupling to the pressure bath. The reference pressure was set to 1 bar. The Nose-Hoover thermostat was used for temperature coupling to maintain a constant temperature of 300 Kelvin.

The protein unit (Fig C) is placed in a triclinic, simulation box with side lengths and angles measured to tile the lattice in the x and y directions. The box dimensions include a 1 nm spacing (i.e. larger than the long-range interaction radii) between its edges and the protein unit to ensure no interaction between periodic images. Then, periodic boundary conditions are applied to all boundaries of the box to form an effectively infinite lattice. The system is then solvated in three-point simple point charge water, annealed, equilibrated, and put through production.

Production occurs as strain is applied to the system by increasing the box length at a constant rate. The inter-atomic forces account for the initial atomic coordinates and then over each integration time step in the simulation the velocities and atomic coordinates are updated according to temperature and pressure coupling, respectively. The simulation software solves Newton's equations of motion for the large-scale system and using velocity pulling within a force field—in this case the OPLS-AA (all-atom) force field was used for all simulations—produces trajectories from which the stress-strain relationship is obtained. GROMACS provides values for the pressure at each time step of production which is used to determine the stress and the percent strain is set pre-production. Stress-strain curves are generated to quantify elasticity or material strength.

Results and Discussion

The elastic modulus is extrapolated from the slope of a linear regression of stress plotted versus strain. A model stress-strain curve is depicted in Figure 4.

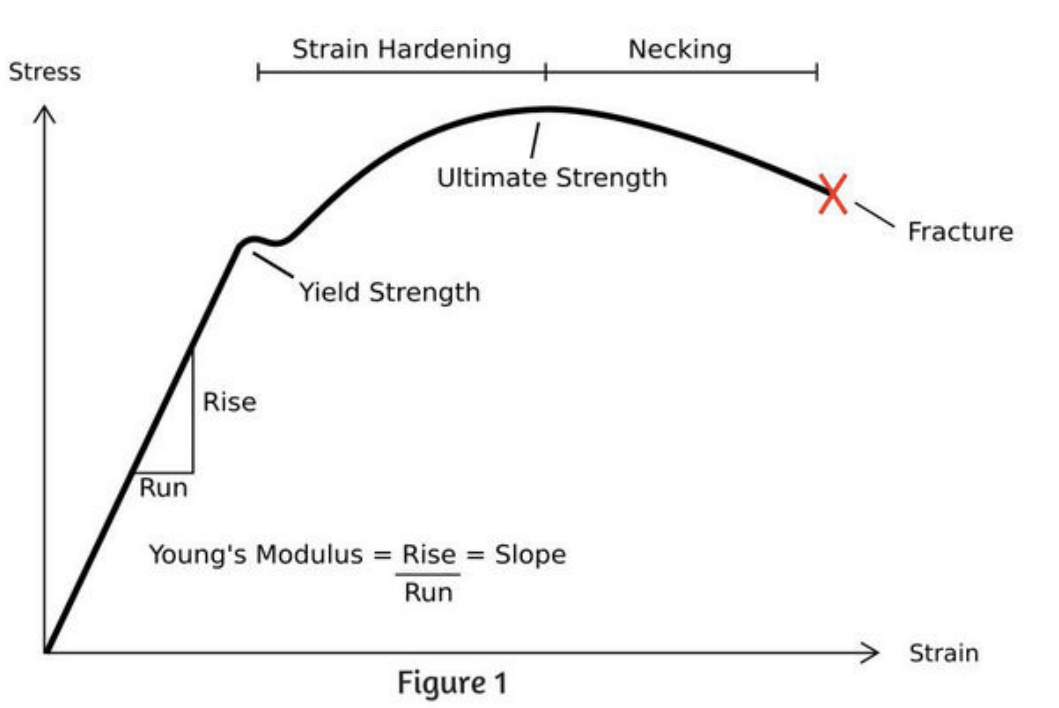


Figure 4: Typical stress-strain curve

A stress-strain curve resulting from a typical simulation is illustrated in Figure 5. A linear regression was applied to points corresponding to less than 1% bulk strain and less than 0.5% shear strain. All elastic moduli measurements for the 2-D square lattice are shown in Figures 6 and 7, which span over speeds ranging from 500 to 1500 m/s and are values averaged from measurements obtained from four different equilibrium starting points. The square lattice exhibits greater resistance to deformation at higher pulling velocities, which corresponds to the results of numerous materials studies. It is noted that the relationship between elastic modulus

and velocity is linear over the range of pulling velocities used. Also, that the bulk moduli are around twice that of the shear moduli, indicating that the lattice is more capable of withstanding isotropic changes in pressure than it is to parallel, anisotropic deformation.

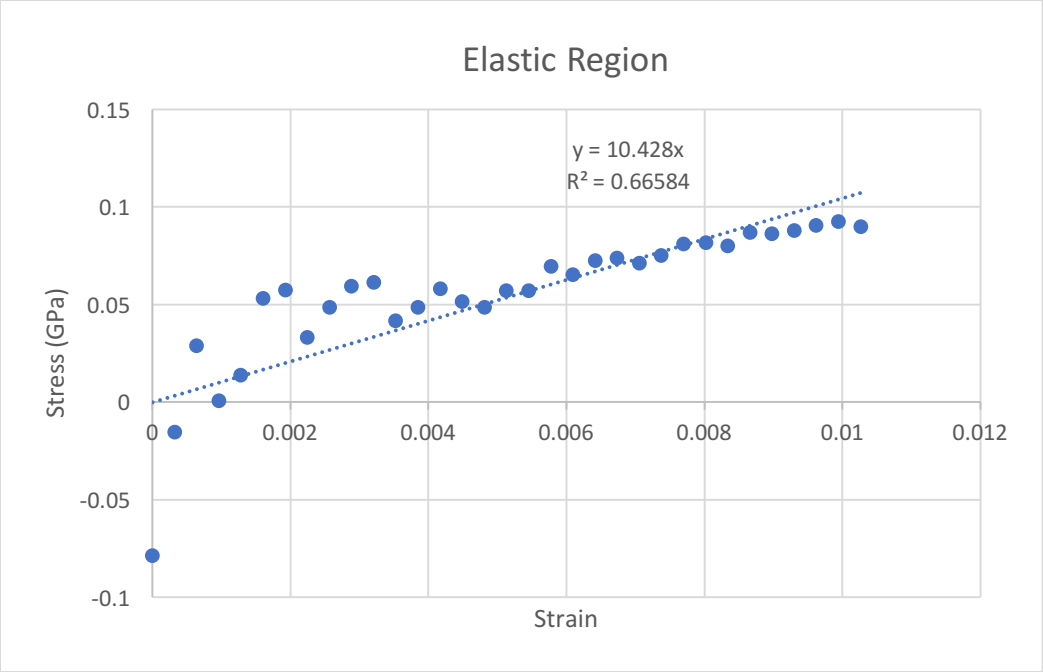


Figure 5: Linear regression for the bulk modulus at a 1% strain and pulling velocity of 1100 nm/ns

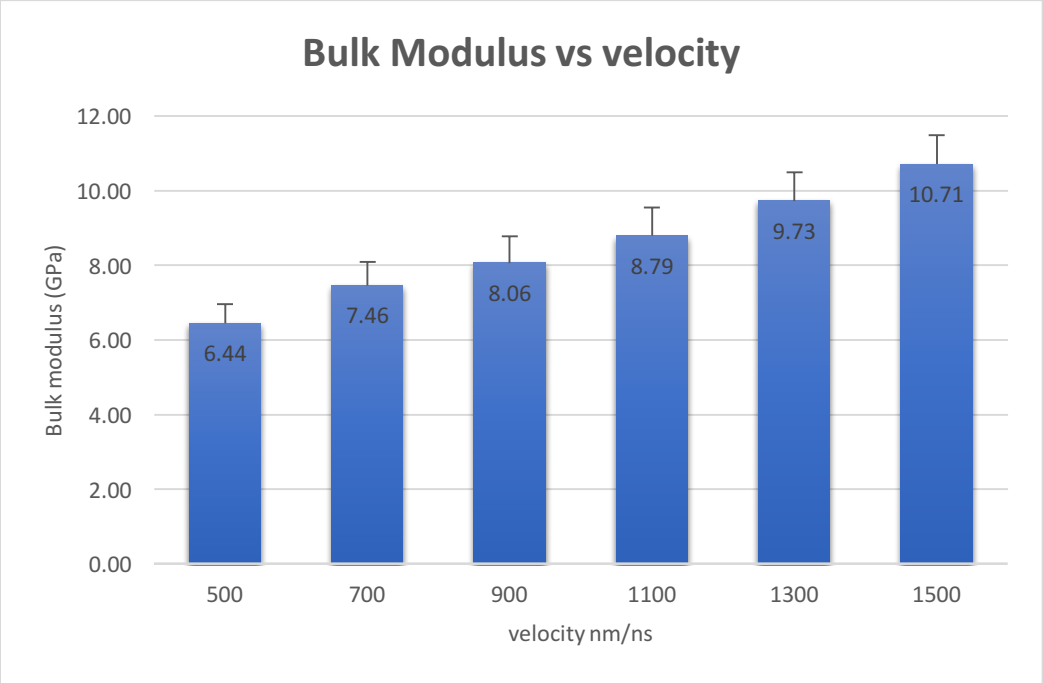


Figure 6: Bulk modulus vs pulling velocity of the square protein lattice at a 1% strain

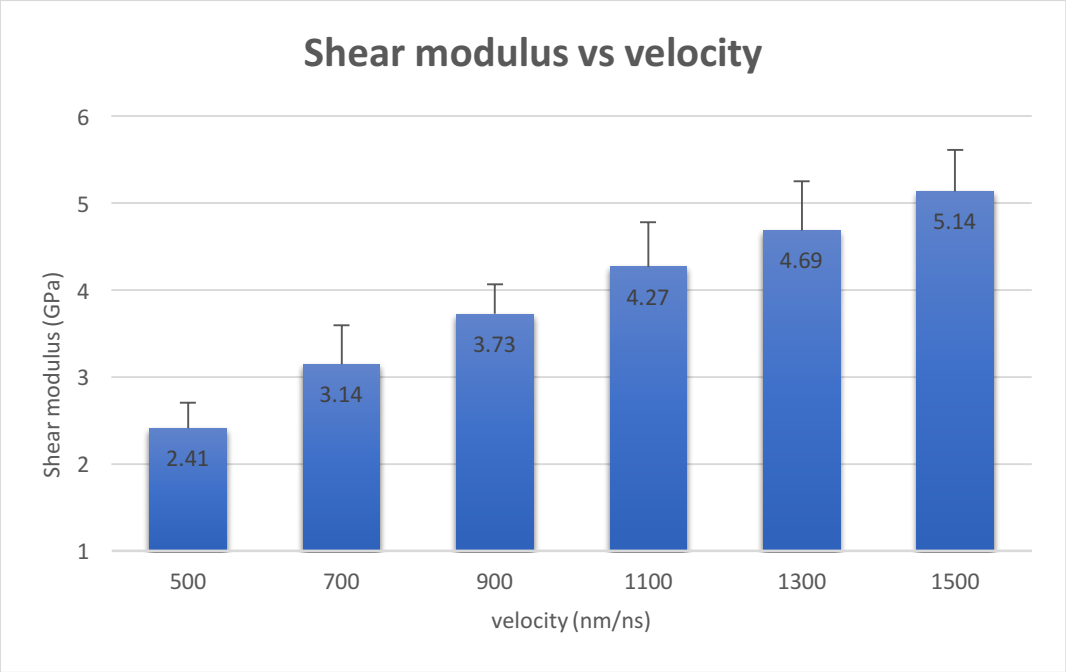


Figure 7: Shear modulus vs pulling velocity of the square protein lattice at a 0.5% strain

To assess the material strength of the square protein lattice, Table 1 provides some context comparing the elastic moduli to that of various elastomers. Although the mechanical strength of these proteins does not rival materials like Kevlar and fiberglass, it does exceed that of materials like rubber, chitin, and the bacterial S-layers found in nature. Its mechanical strength is comparable to that of spider silk which is composed of mostly β-solenoids and β-sheets.

Table 1: Mechanical moduli of conventional materials for context

Material	Bulk Modulus (GPa)	Shear Modulus (GPa)
Kevlar[5]	71-112	2.8-4.1
Fiber Glass[5]	43-50	30-36
N. Clavpipe spider silk[9]	12.71	2.38
Square Protein Lattice	6.44-10.71	2.41-5.14
Vulcanized Rubber[5]	2.7	0.005
Chitin/Chitosan[7]	0.350-0.421	0.198-0.217
Bacterial S-layers[2]	0.006	0.002

Conclusion

Simulated stress-strain curves have been determined for a 2-D square protein lattice system using the molecular dynamics software program GROMACS. With this method, the values for the bulk and shear tensors were evaluated and substantiate the β-solenoid based protein lattice's appreciable mechanical strength given that stresses encountered by biological systems occur between the 0.1 and 1 GPa range. These results support the premise that when assembled into two or three-dimensional lattice structures, these proteins make promising candidates as a viable, robust foundation for creating novel biomaterials.

Future work

To determine if this β-solenoid protein lattice would be a viable biosensor it is necessary to calculate the piezoelectric tensor. This would require velocity pulling while measuring the change in the electric dipole moment of the lattice. Pulling will simulate the conformational change induced by a binding event of an antigen to its corresponding antibody which would be attached via functionalized loops bound to the scaffold-like protein lattice. The magnitude of the piezoelectric tensor would dictate whether binding to the lattice would deform it enough to generate a significant—measurable and distinguishable—voltage. This would have potential application for cheap and compact disease detection devices. Furthermore, we will investigate the role of geometric shape in mechanical performance by performing stress-strain measurements on other lattice geometries such as oblique and hexagonal, which are found in nature.

References

1. Sleytr, Uwe B. et al. S-layers: Principles and Applications (2014) *FEMS Microbiology Review*
2. Lopez AE, et al. Influence of Surface Chemistry and Protein Concentration on the Adsorption Rate and S-layer Crystal Formation (2011) *Phys. Chem.*
3. Heinz, Leonard P. et al. In Silico Measurements of Twist and Bend Moduli for -solenoid Protein Self-Assembly Units (2015) *ACS NanoLetters*
4. Peralta, Maria D.R. et al. Engineering Amyloid Fibrils from Beta-solenoid Proteins for Biomaterials Applications (2015) *ACS Nano*
5. AZO Materials: https://www.azom.com/properties.aspx
6. Pum, Dietmar et al. S-Layer Protein Self-Assembly (2013) *International Journal of Molecular Science*
7. Nurul, Hazwani M.H. et al. Effect of Pressure on Elastic Properties of Chitosan (2011) *Solid State Science and Technology*
8. Abraham, Mark et al. GROMACS Reference Manual (2016) *Department of Physical Chemistry,* Netherlands *& Royal Institute of Technology*, *Sweden*
9. Romer, Lin and Scheibel, Thomas. The Elaborate Structure of Spider Silk (2008) *Prion*
10. Hong, Seungpyo and Kim, Dongsup Library of Binding Protein Scaffolds (2016) *Bioinformatics*

Acknowledgement

National Science Foundation Research Experience for Undergraduates Program (NSF grant PHY-1560482)