Perspective

Engineering Materials from Proteins

Bradley D. Olsen
Dept. of Chemical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139

DOI 10.1002/aic.14223
Published online September 6, 2013 in Wiley Online Library (wileyonlinelibrary.com)

Keywords: polymer, protein, biomaterials

Introduction

Nature provides great inspiration for engineering efforts, having developed complex materials that can sustain life in a variety of environments. A common category of macromolecules underlies many of the engineering accomplishments exhibited in nature: strong fibers,[1] underwater adhesives,[2] tough composites,[3] fatigue-resistant hinges,[4] selective filters,[5] solar energy converters,[6] and highly active catalysts[7] are all made wholly or in part from proteins. Emerging technology needs, particularly in medicine but also in fields such as security, energy, and sustainability, have encouraged engineers to look to proteins as a source of building blocks and design inspiration for the next generation of polymeric materials. Even though the study of polymer science and protein science developed as intellectually separate disciplines, a growing community of researchers is designing materials at the interface of these two fields. A few examples are bioinspired polymers and synthetic enzymes that use traditional polymer materials to mimic the performance of proteins,[8] protein-polymer hybrids that incorporate both types of molecules into a single material,[9] and artificially engineered protein polymers that apply the design principles of polymer science to the engineering of new proteins.[10]

This convergence of protein engineering and polymer science and the rapid progress at the interface of the two fields is indicative of a larger trend: several previously distinct fields of soft materials are being drawn together into an intellectually cohesive subject, unified by common considerations of time, length, and energy scales governing the underlying physics, a common set of building blocks for constructing molecules, and a common set of characterization tools. Riding both this fundamental shift in materials engineering and the continued rapid advances in molecular biotechnology, protein materials engineering has become well-established as a discipline with all the tools to make a significant impact on modern technology.

Correspondence concerning this article should be addressed to B. D. Olsen at bdolsen@mit.edu.

While our ability to construct a myriad of protein materials is already well-developed and continues to grow, our understanding of the materials engineering and physics of these systems lags significantly. At the level of covalent bonds, proteins are poly(amino acids), polymers synthesized from a set of 20 naturally occurring monomers and a number of chemical modifications thereof. However, proteins demonstrate capabilities far beyond what is achievable with synthetic polymers due to their sequence-specific chemistry. The specific monomer sequences in proteins enable them to fold into the complex hierarchical structures which give rise to their function (Figure 1) and differentiate them from their synthetic polymer cousins.

The goal of this perspective is to explore the similarities and differences between proteins and polymers in a materials context and to highlight several important questions relevant to the physics and engineering of these materials. First, approaches to the design of protein materials and protein-polymer hybrids are reviewed, highlighting large advances in chemistry that enable the design of new materials. Several key examples are given that illustrate how the non-Gaussian chain conformations and specific sequences present in proteins have a powerful impact on the physical properties of the resulting polymers, pushing the limits of polymer physics. Soft matter physics challenges are then highlighted in the self-assembly of proteins into functional materials, leveraging proteins as templates for the construction of synthetic materials, preserving the folded structure of proteins within materials, and understanding structure-mechanical property relationships in structural proteins that lead to their high performance.

While many of these problems are old, recent developments in chemistry, gene synthesis, and protein engineering make protein materials increasingly accessible to the physicist and materials engineer, positioning the field for renewed progress in understanding the physics of these sequence-specific polymers. In each case, critical applications that will benefit from a deep understanding of the relevant physical principles are discussed. It is hoped that these challenges will stir further collaborative efforts between polymer physicists and protein engineers to develop a deep understanding of the principles guiding protein material design, enabling the next generation of unexpected technologies.

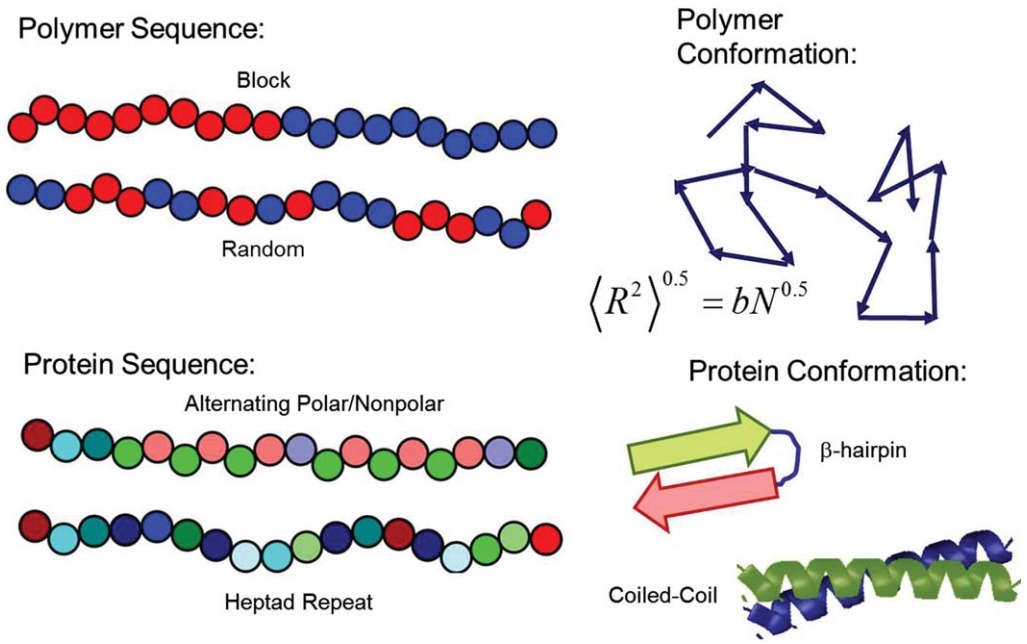

Figure 1. Difference between polymer and protein structure.

Sequence specificity in proteins, which also implies perfectly monodisperse chains, enables specific interactions within a polymer molecule that determine the ensemble of chain configurations adopted by the polymer and the interactions between different polymer chains. These effects are being exploited by the materials community to significantly broaden the scope of polymer science and polymer engineering.

Chemistry of Protein-Based Materials

The ability to consider problems of design and engineering in protein-based materials for science and technology applications is driven by stunning advances in both polymer chemistry and molecular biology over the last 50 years, providing the ability to synthesize such molecules in a relatively pure state in large quantities. However, the design of sequence-specific polymers continues to be a grand challenge in polymer synthesis, and three complementary approaches have been developed to address this problem as applied to the synthesis of protein materials: solid-state synthesis, N-carboxy anhydride (NCA) polymerization, and biological synthesis. The solid-phase synthesis of proteins, originally developed by Merrifield,[11] uses alternating coupling and deprotection reactions to grow a protein chain immobilized on a solid support. The protein chain is cleaved from the support at completion of the synthesis to yield the target molecule. The step-by-step chemistry enables sequence-specific molecular designs to be prepared, and the solid-phase immobilization enables easy purification.[12] The resulting popularity of the technique has led to commercialization of the protected amino acids required for the synthesis such that materials can be routinely prepared on automated peptide synthesis machines. However, limits on conversion at each synthetic step translate into limits on the number of amino acids that can be polymerized (currently degrees of polymerization less than ~100 are readily achievable).

If sequence-specific polymerization is not required, high-molecular-weight polypeptides can be prepared using traditional chain growth polymerization. NCA polymerization[13,14] uses an activated amino acid monomer that undergoes ring-opening polymerization initiated by a primary amine or transition metal catalyst, allowing the synthesis of polypeptides with sequences similar to traditional homopolymers, random copolymers, or block copolymers. While the synthetic methods are easily scaled to meet materials challenges, sequence specificity is sacrificed. Therefore, NCA polymerization and solid phase peptide synthesis produce significantly different materials with complementary roles.

Currently, the only way to achieve high-molecular-weight materials with sequence specificity is to use the natural machinery of biology to synthesize proteins. The biological method for protein synthesis is attractive for several reasons. First, a protein can be designed using the principles of polymer engineering to form materials such as gels[15] or elastomers,[16] and biological functionality can be simply incorporated into the material through the use of functional protein sequences such as RGD.[17] Second, once a gene is synthesized, it can be multiplied indefinitely to allow a protein to be expressed at an arbitrarily large scale. Third, biological synthesis allows the facile design of fusion proteins which can behave as block copolymers,[18] often removing the need for bioconjugation as a part of molecular design. As DNA synthesis technology has advanced, genes encoding for many such proteins can now be purchased, significantly reducing the up-front burden for cloning required in biological synthesis.

After synthesis of a gene, expression and purification of the single encoded protein species from the complex cellular milieu presents additional challenges. For enzymes, the protein's tertiary or quaternary structure is difficult to recreate

and must often be preserved throughout expression and purification; however, for many artificially engineered protein polymers the folded structures are simple and may be refolded after denaturing purification. This flexibility of purification strategy makes many artificially engineered protein polymers easier to isolate in high yield than globular proteins. In addition, advances in tag-based purification[19] that enable single-step, high-purity protein to be produced for a wide variety of chemistries have significantly advanced the ability to isolate research quantities of these materials with minimal effort. Strategies for the thermal or salt-based purification of proteins have also been developed, including tag strategies, which avoid the cost of chromatographic resins and enable highly scalable purification methods.[20] These advances have made a great many proteins available to the research community in quantities suitable for material preparation.

A drawback of biological protein synthesis is that natural systems are limited to a set of 20 canonical amino acids. The incorporation of noncanonical amino acids provides an extremely valuable technology for the synthesis of new protein materials, enabling new bioconjugation chemistries[9,21] and tuning of protein interactions.[22] While the chemistries used for solid-phase synthesis and NCA polymerization are amenable to the incorporation of many such noncanonical amino acids, their incorporation into proteins through biological synthesis is more challenging. Among a variety of different approaches, residue-specific incororpation[23] and amber codon suppression[24] have emerged as two leading technologies to incorporate noncanonical amino acids. In the residue-specific method developed in the Tirrell laboratory,[23] one of the 20 natural amino acids is replaced with a synthetic analogue either by charging a natural tRNA with a noncanonical amino acid or by replacing one of the natural tRNAs with a synthetic analogue that binds to a noncanonical amino acid. This method offers the advantage of simultaneously replacing all instances of an amino acid within a protein, allowing a high level of noncanonical functionality to be readily achieved. In contrast, the amber stop codon method popularized by the Schultz laboratory[24] reassigns the amber stop codon to encode for a 21st amino acid. This allows a noncanonical amino acid to be incorporated without replacing one of the natural amino acids and has proven particularly useful for producing proteins with a single noncanonical monomer.

Coarse-Grained Physics in Protein Polymers

How much does the physics of protein polymers differ significantly from that of traditional synthetic polymers? This question is not just a matter of whether existing models can be applied to the characterization of polymers composed of amino acids, but touches on the deeper concept of what level of chemical detail must be considered to describe sequence-specific macromolecules. Polymer physics takes a coarse-grained approach to formulating descriptions of chain conformation, thermodynamics, and dynamics, suggesting that all linear, flexible polymer chains should have broadly similar properties.[25,26] While this coarse graining is broadly effective for polymers with a low degree of chemical complexity, the chain conformations and thermodynamics of proteins originate from their monodisperse, sequence-specific primary structures. Given the profound impacts of sequence control on molecular structure and function, it is natural to expect that protein polymers and protein materials will exhibit different physics than their polymer cousins; the question as to how far coarse grained descriptions can go toward treating systems with the level of chemical control present in proteins naturally arises. Several key examples demonstrate that coarse-grained polymer models continue to have a role to play in describing protein polymers, but that material design with proteins can dramatically expand the scope of physics observed in polymeric systems, providing rich new fields for exploration.

An area attracting recent interest that highlights both similarities and differences between proteins and traditional polymer physics is the study of the force-extension curves of single proteins. Single molecule force puling experiments have been used to understand the force-induced unfolding of proteins, and subsequently this understanding has been applied to engineer protein hydrogels with improved mechanical properties. In these experiments, a single protein domain is cloned into a fusion polymer containing several such domains arranged in a linear sequence (Figure 2).[27] When the protein is extended, the force extension curve follows that of a short Gaussian chain until a point at which one of the protein domains unfolds, causing relaxation of the stress as the effective length of the Gaussian polymer chain increases. Subsequent unfolding of the additional protein domains at increased strain creates a sawtooth pattern in the force-extension curve for the polymer chain due to the unfolding of quanta of stored length when each individual protein domain unfolds. The energetics of the unfolding process affect the way that energy is recovered upon relaxation, enabling energy dissipation mechanisms to be engineered. This example shows that many protein force-extension curves can be well modeled by traditional coarse-grained polymer physics, but the specific effects of protein unfolding/refolding dynamics must be incorporated. This concept is useful in hydrogel engineering: it provides a molecular mechanism for stress relaxation in polymer gels due to the release of stored length, an effect that is challenging to capture in synthetic materials. Stored length significantly increases the extensibility of protein gels[28] and is thought to be a mechanism to prevent tearing in muscle tissue.[29] Clearly, both protein science and polymer physics have a role to play in applying this concept to material design.

Another example of how protein polymers move beyond the scope of traditional coarse grained polymer theories is engineering elastin-like proteins (ELPs).[31]These proteins, with sequences derived from natural elastin, consist of pentapeptide repeat sequences VPGXG or close relatives where the guest residue X can be changed to affect the physical properties of the resulting ELP. One of the most remarkable properties of these proteins is that they demonstrate lower critical solution behavior, similar to synthetic thermoresponsive polymers such as poly(N-isopropylacrylamide) (PNIPAM). However, the transition is different in that the chains do not transition from a random coil to a globule upon chain collapse, but rather have been shown to undergo a random coil to β-turn transition.[32] Urry and coworkers demonstrated relationships between the guest residue and the transition temperature of the polymer;[33] the group contribution

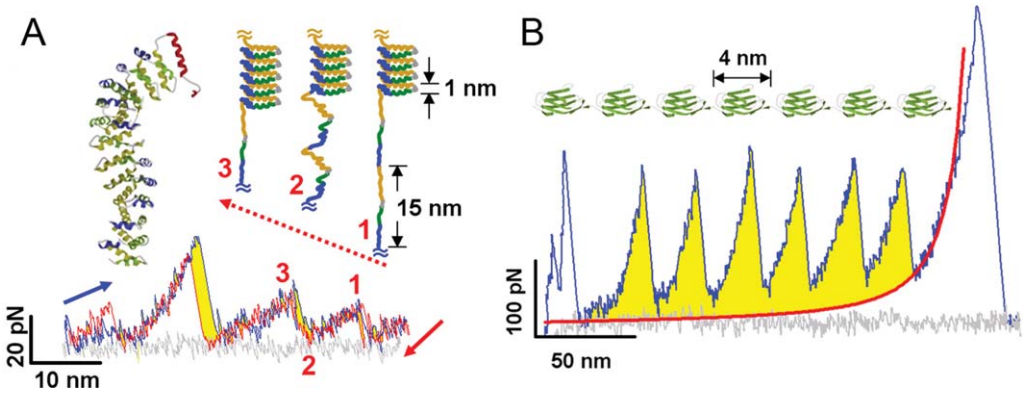

Figure 2. Chain extension of protein polymers.

Force-extension curves during extension (blue) and relaxation (red) for (a) 12 α-helical ARM repeats of the protein b-catenin, and (b) seven repeats of the Ig 27 domain. The yellow areas show energy dissipated during a single-loading cycle. α-helical repeat proteins dissipate significantly less energy than the Ig 27 domain upon unfolding. Reproduced with permission.[30] Copyright 2010, Elsevier, Ltd.

predictions for transition temperatures derived from Urry's data have been used to guide the application and design of ELPs for a wide variety of biomedical applications. Polymer physics suggests that molecular weight and concentration should also have a large impact on the thermal transition temperature according to the Flory–Huggins Theory,[25] and thermodynamic models are being developed that predict solution properties of the polymers as a function of these parameters as well. The most recent models indicate that group contribution theories for prediction transition temperature as a function of guest residue cannot capture observed sequence-dependent effects,[34] suggesting complex solution thermodynamics and an opportunity for further advances in polymer solution theory.

Self-Assembly of Folded Protein Domains

As the aforementioned examples illustrate, protein polymers present complex new behaviors that expand the scope of polymer physics. One of the most interesting aspects of polymers and soft materials is their ability to self-assemble, forming nanostructured materials with a great diversity of structures that can be tuned through molecular design. Although the self-assembly of systems such as block copolymers[35,36] and surfactants[37] is well-known, the folds of proteins provide access to new molecular designs. The challenge of self-assembling these molecules represents an emerging frontier in polymer science. Studies of self-assembly of folded proteins are motivated by the challenge of forming structured materials out of molecules that typically exist only in solution. For example, a wide variety of industrially useful enzymes are soluble, and many other useful proteins operate within lipid bilayers. Incorporating these folded and functional enzymes and other proteins into new materials represents a large opportunity for the development of new products such as sensors or heterogenous catalysts.[38,39] For proteins which do not naturally form mechanically robust materials, either immobilization/encapsulation in a polymer matrix or self-assembly provide routes to capturing the biological functionality in an engineered system. Self-assembly of the materials through construction of protein-polymer conjugate or fusion protein block copolymers provides a valuable technique for the control of structure on length scales ranging from the individual protein through the nanoscale all the way up to the micro- and macroscale morphology.

The folded shape of the protein is important both for the function of a protein material (i.e., catalytic activity) and for the physics of block copolymer self-assembly. Proteins have a diverse variety of shapes due to their secondary, tertiary, and quaternary structures, with typical dimensions of 3–8 nm for globular proteins. While comparatively simple structures such as α-helices,[40–42] β-strands,[43] ring peptides,[44] and coiled-coils[45] were first characterized, more complex structures such as β-barrels and membrane proteins represent specific examples of the broad category "globular proteins" that contains an almost unlimited diversity of folded shapes. For the few protein folds that have been studied, the effect of these shapes on self-assembly has been shown to be profound. α-helical peptides provide the most thoroughly investigated example. Gallot and coworkers were the earliest to synthesize a-helical polypeptide-polymer conjugates, demonstrating a propensity to self-assemble into lamellar phases (Figure 3).[40] Further studies of similar protein-polymer conjugates and emerging knowledge of self-assembly in the broader category of rod-coil polymers (the α-helical block is rigid while the coil block is flexible)[46] have demonstrated that the phase diagram is significantly different than that of traditional block copolymers due to the effects of chain topology and anisotropic molecular interactions. Studies on ring peptides, coiled-coils, β-sheets, and β-barrel proteins[47,48] have also shown significant differences from traditional block copolymer self-assembly for the formation of solid materials.

In addition to the effect of protein shape, the specific interactions between proteins are anticipated to have a large impact on protein self-assembly. In the same way that the Flory–Huggins theory of polymer solutions serves as a basis for theories of block copolymer self-assembly, theories of protein solutions will be critical to formulating as yet undeveloped theories of protein self-assembly. Despite the

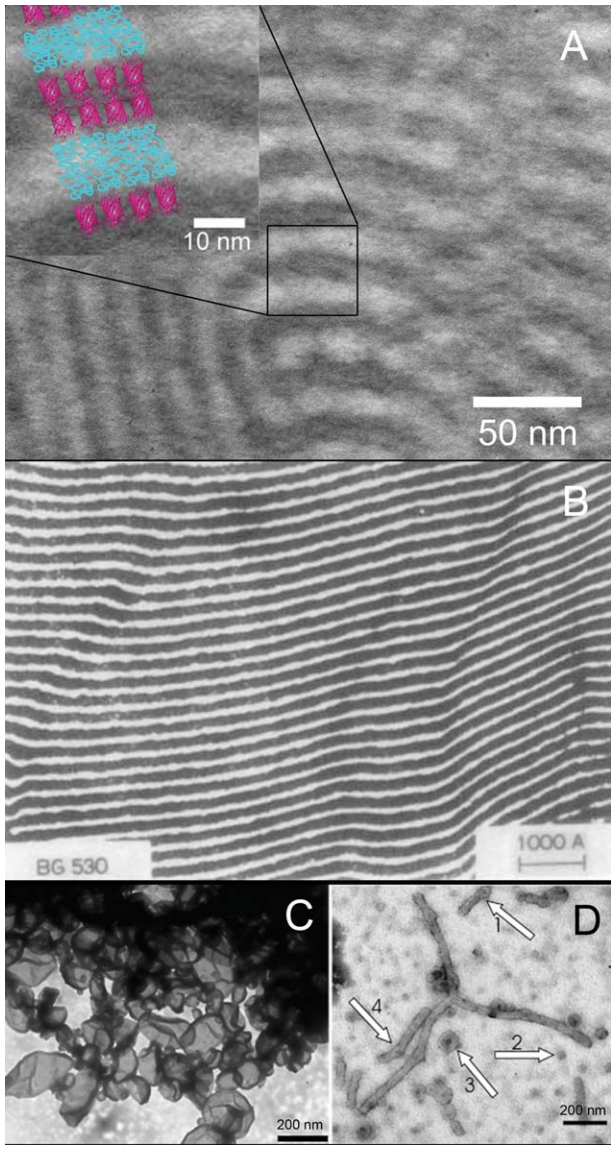

Figure 3. Protein-polymer conjugate sefe-assembly.

Self-assembled nanostructures formed from protein-containing block copolymers include (a) solid-state lamellar structures formed from globular proteins such as mCherry-*b*-poly(N-isopropylacrylamide), (b) lamellae formed from polybutadiene-*b*-poly(benzyl-L-glutamate) swollen with polymerized 2,3-dichloro-1-propene, and (c, d) horseradish peroxidase-*b*-polystyrene-*b*-poly(ethylene oxide) forming vesicular aggregates and a mix of rod, toroid, and junction structures. (a) reproduced with permission.[48] Copyright 2011, American Chemical Society. (b) reproduced with permission.[40] Copyright 1976, John Wiley and Sons, Inc. (c, d) reproduced with permission.[49] Copyright 2007, Elsevier, Ltd.

complexity of protein interactions, a variety of theories based on simplified protein-protein interaction models have been developed and shown success in predicting key aspects of protein phase behavior in solution. These models use averaged parameters such as the second virial coefficient, the protein solubility, or the surface excess interaction with a solute to characterize the protein solution thermodynamics. They have been demonstrated to predict both conditions for protein crystallization[50] and the qualitative shape of the protein phase diagram in solution.[51,52] Protein-polymer interactions have also been explored in solution, illustrating the importance of excluded volume (depletion interactions) in governing phase behavior.[53] The translation of these elements to the self-assembly of protein materials has the potential to fill a gap in our quantitative understanding of bioconjugate and fusion protein self-assembly by producing theories for the thermodynamics that governs structure formation.

The self-assembly of protein-polymer conjugates or fusion proteins to form nanomaterials in solution (i.e., micelles, vesicles) has attracted a great deal of interest for applications in drug delivery, yielding a fascinating observation: not all proteins form stable micelles. For example, EGFP-PNIPAM conjugates form stable micelles,[54] but structurally similar mCherry-PNIPAM micelles do not.[55] Stable micelles can also be formed from globular protein-ELP fusions above the ELP thermoresponsive transition; however but the stability of the micelle depends on the identity of the globular protein.[56] Clearly, the steric effects in a micelle with a folded protein corona are significantly different than in a micelle with a Gaussian coil polymer corona, and these comparisons suggest that a detailed understanding of the effects of electrostatics, hydrophobicity, and sterics on protein micelles is required to enable rational design of stable self-assembling systems in solution. Advances in our understanding of self-assembly in this area will provide important input to material design for medical applications.

Proteins as Templates for Self-Assembly

While the construction of copolymers enables the self-assembly of soluble proteins and membrane proteins, a number of proteins form robust self-assembled structures that may be used to guide the self-assembly of other molecules. Materials such as viral capsids[57] and amyloid fibers[58] present examples of such natural assemblies that can be exploited as templates. Engineered repeat proteins can be used as modular platforms for controlling the placement of functional groups,[59] and attachment to a rigid α-helix backbone can control the spacing between organic chromophores.[60]

Protein nanostructures offer two important advantages as templates for nanomaterial self-assembly: first, the protein template self-assembly is often orthogonal to the behavior of attached functional groups, enabling the protein assembly to be designed separately from the functional group and to govern the placement of the functional group within the final nanostructure. Second, changes in the sequence of the self-assembling protein provide a valuable method to control the functionalization location and the structure of the self-assembly, allowing the chemistry and physics of the material to be tuned. The *de novo* design of protein templates which can guide assembly into a predetermined structure remains a challenge; therefore, efforts focus on developing platform materials adapted from specific natural systems.

Viral capsids provide an example of protein-based nanoparticles that may be readily exploited as templates for self-assembly or as nanocapsules for the delivery of cargo. Spherical viral capsids have attracted interest as models for

templating on nanoparticles. Many of these capsids have clearly defined inner and outer surfaces that can be separately functionalized, allowing particles to be prepared with controlled properties on both their outer surface and in an inner compartment that may be used to contain cargo.[61] In some cases it has been observed that functionalization of the template protein causes self-assembly into a different structure, suggesting that in some cases the free energy of template assembly may not dominate system behavior.[57] Rodlike viral capsids, including the M13 bacteriophage and the tobacco mosaic virus (TMV), have also inspired the design of a large number of materials, starting with early investigations of these capsids as model rodlike particles capable of self-assembly into nematic and smectic liquid-crystalline phases.[62] These viruses have since been developed into versatile platforms for the templated growth and assembly of a wide variety of inorganic and organic molecules.[63]

Protein nanofibers (Figure 4) are another large category of materials that have attracted interest in the templating of materials. Inspired by the pathogenic assembly of proteins into amyloid fibrils, researchers have designed short oligopeptide sequences[58] that assemble into β-sheets and β-turns, which aggregate into long fibers or induced globular proteins to refold into β-sheet rich fibers.[64] When prepared in solution, these fibers physically entangle to form shear-thinning hydrogels. Responsive changes in protein configuration may be used to trigger the sol–gel transition, different gel designs have been shown to be highly effective as tissue engineering matrices.[65] Covalent attachment of functional groups to these fibers enables the preparation of bioactive hydrogels[66] or the control over the spatial positioning of organic chromophores.[67]

Enhancing the Activity and Stability of Proteins

Because the chain configuration of a protein is so important to its functional role, engineering materials from proteins raises the question of how to stabilize a macromolecule in a specific chain configuration. There are two broad approaches that have been taken to address this challenge (a) engineering proteins to improve the stability of the desired folded structure, or (b) using a material environment to stabilize the protein. Directed evolution and protein engineering approaches use changes in the amino acid sequence of a protein to increase its thermodynamic or kinetic stability.[69] These techniques have been shown to be highly successful in improving protein thermostability and chemostability. The idea that engineering a protein sequence can improve its stability has also motivated extensive research on extremophiles, organisms that have evolved proteins to withstand extremes of pH, temperature, or ionic strength.[70] Typically, protein engineering approaches focus on a single protein molecule, but advances in high-throughput characterization suggest that these methods may be extended further. Perhaps, one day it will be common to apply directed evolution using screens of proteins in materials to evolve the optimal protein performance in a material context.

A complementary method to protein engineering is to encapsulate/immobilize the protein in a matrix or additive that modifies its surrounding environment to improve stability.[39] The challenge is to develop material chemistries that allow preparation of the encapsulating layer without denaturation of the protein due to heat, exposure to organic chemicals, or changes in protein shape due to the formation of chemical bonds. Encapsulation methods enable the formation of solid layers with proteins embedded, and commonly employed methods include incorporation into layer-by-layer assemblies,[71] immobilization in polyurethanes,[72] or the incorporation into self-assembled block copolymer templates.[73] The formation of chemical and physical bonds at the surface of the protein can be used to provide rigid cross-links that stabilize the protein's tertiary or quaternary structure,[74] and the matrix may additionally contribute hydrogen-bonding or osmolytic properties that help to stabilize the protein.[39] These solid encapsulants are particularly useful for the preparpation of heterogeneous catalysts and sensors, where they can be prepared in the form of microparticles or films, respectively. In addition to controlling enzyme stability, it is important for catalyst development to engineer diffusive transport and electrical transport through the encapsulating materials so that mass transfer limitations or a poor conductivity do not limit the enzyme's performance.[38]

Motivated largely by challenges in stabilizing pharmaceutical proteins in the dehydrated state, approaches to improving

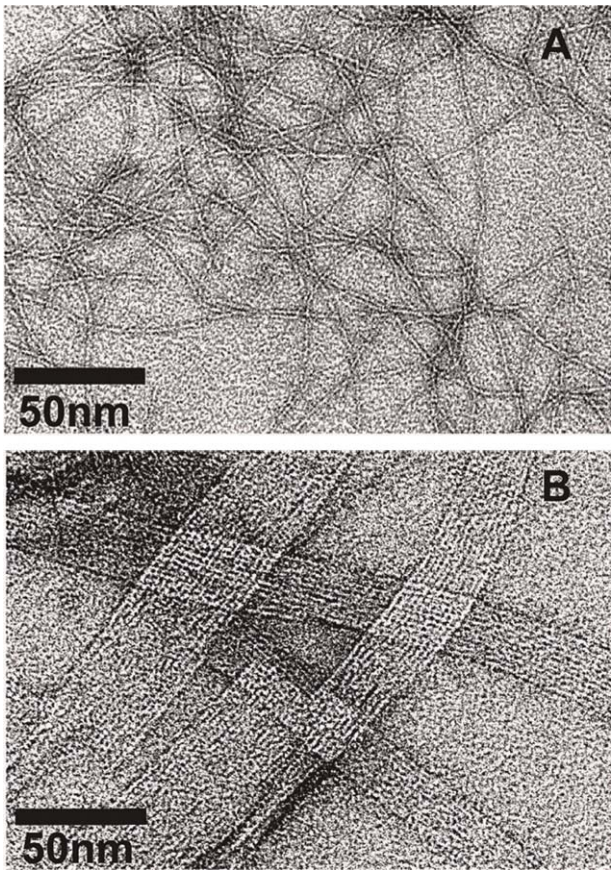

Figure 4. Self assembly of protein fibrils.

Transmission electron micrograph (TEM) illustrating (a) entangled fibril, and (b) laminated ribbonlike morphologies formed by self-assembly of β-hairpin and extended β-strand self-assembling peptides, respectively. Reproduced with permission.[68] Copyright 2005, American Chemical Society.

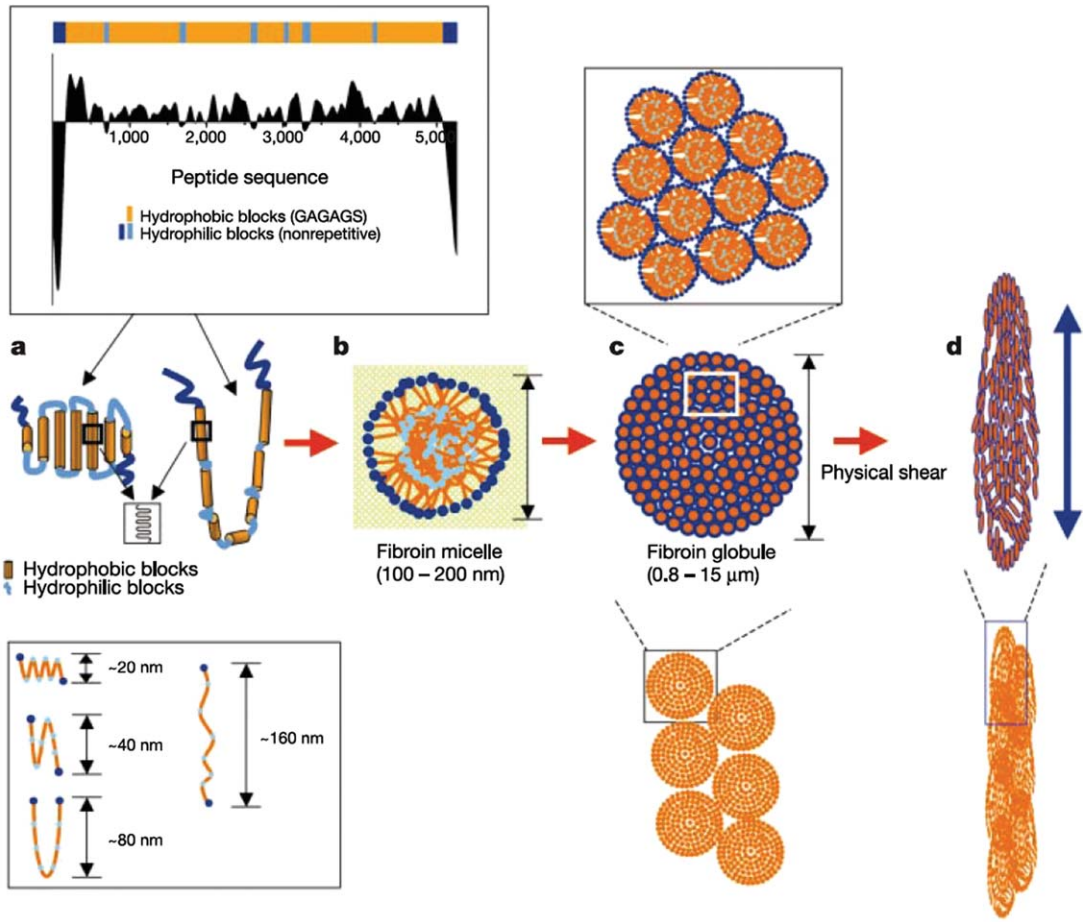

Figure 5. Processing of reconstituted silkworm silk fibroin to mimic the structure of the natural fibers.

The hydrophobicity pattern in silk (a) leads to initial chain folding into micellar structures, (b) which then aggregate into larger globules, (c) as concentration is increased and through the addition of crowding agents. Elongation of globules in physical shear, and (d) leads to the formation of fibrillar structures. Reproduced with permission.[86] Copyright 2003, Nature publishing group.

protein stability based on the addition of small molecule additives have also been developed that are particularly useful for formulating materials in the dry state. These additives include osmolytes (i.e., sugars), polymers,[75] chelating agents,[76] salts,[77] and surfactants. They are thought to operate by two mechanisms. First, some additives may form rigid, glassy structures around proteins similar to encapsulants, providing a method to increase stability by reducing thermal motion of the protein. Trehalose, a high-glass transition sugar, is a very common osmolyte that is capable of forming such a rigid encapsulating matrix. Second, additives may replace hydrogen bonds, serving as water replacements in the dehydrated state. Molecules such as polyols are thought to form these bonds using hydroxyl functionalities.

A major hurdle for studies of protein stability in materials is the limited set of analytical tools available to the researcher for characterizing protein secondary and tertiary structure within materials. This limit is a large contributor to the lack of fundamental knowledge about protein denaturation pathways in materials and quantitative structure–property relationships that are critical to maximizing protein activity. Consequently, it is difficult to understand how stabilization strategies developed for one protein will apply to another, and each protein must currently be engineered on a case-by-case basis. Spectroscopic tools, such as circular dichroism (CD) and Fourier-transform infrared (FTIR) spectroscopy, provide information about the secondary structure of the protein, but often the secondary structure shows relatively minor changes even while a protein undergoes major rearrangements in tertiary or quaternary structure. Scattering methods such as small-angle neutron scattering can measure tertiary structure directly, but resolution is limited and scattering contrast must be carefully controlled. All of these methods measure an ensemble average of the structure of many proteins which are potentially in many different unfolded states. Therefore, techniques that are able to probe single molecules to gain specific information about their unfolded structure would be extremely helpful in identifying mechanisms for the loss of protein activity and would enable rapid advances in the design of materials to stabilize proteins. Advances in protein modeling or in the development of protein-like polymers[78] as model synthetic systems provide further routes for developing general principles to understand the stabilization of a desired macromolecular configuration.

Structure–Property Relationships in Structural Proteins

Structural proteins are the most abundant proteins in nature. Collagen and elastin form the primary components of many tissues, keratin is a protein found in hair, spider silk is used to form strong webs, and resilin is found in the flight and jumping

systems of insects. Understanding the way nature has designed structural proteins can have significant application to structural polymers or for biomaterials, where control over mechanical properties is critical to controlling cellular response and engineering properties such as injectability. Often, protein structural materials exhibit mechanical properties (for example, strain hardening[79]) that are difficult to replicate with current synthetic polymers; therefore, understanding their design promises to enable materials with improved mechanical properties and new responsive transitions.

Resilin is a fascinating example of a structural protein where the properties of the natural system can be engineered into recombinant protein materials. Resilin exhibits an exceptionally high resilience, storing greater than 97% of the work of deformation as elastic energy and undergoing a large number of cyclical motions with minimal fatiguing. It is naturally found in insect wing hinges and composites with cuticle used to store energy for jumping. Native resilin is composed of three exons, each with its own distinct structure. Until recently, it was thought that resilin did not exhibit a well-formed secondary structure, suggesting that it behaves as an ideal polymeric elastomer.[80] However, it has recently been proposed that a reversible β-turn transition in exon III is responsible for the superelastic properties of the material.[81] Because of its properties, resilin has attracted significant interest for engineering artificial protein analogues, and several minimal sequences have been developed. These are actively being investigated for the development of new biomaterials,[82] where the minimal sequences are able to produce hydrogels with a relatively high resilience. The relationship between design of the minimal sequence, material processing and crosslinking, and the resilience of the final hydrogel continues to be an area of ongoing research as emerging knowledge from studies of native resilins is transferred into the design of artificially engineered protein polymer analogues.

An example of the impact of processing on material performance is engineered silk proteins. Spider dragline silk has been shown to have a tensile strength that exceeds that of steel per on a mass basis,[83] making it one of the strongest fibers known. Silk's morphology has a great deal of similarity to tough synthetic polymers such as polyurethanes and polyolefins: highly crystalline domains are dispersed within an amorphous matrix, with individual chains alternately through crystalline and amorophous domains to chemically link them together. A great deal of effort has gone into developing expression systems that can produce high-molecular-weight variants of the slik-forming proteins in high yield with notable success.[84,85] However, the simple formation of the protein is not enough. The translation of the primary chemical structure of silk into a tough material requires careful control over the processing of the polymer to induce the appropriate morphology within the final material, a classical problem in polymer science. Techniques have been developed that mimic the formation of native silk fibers (Figure 5);[86] while these are able to replicate many of the properties of natural materials, work remains to be done to develop optimized processing strategies for a wide range of silks. Much of the knowledge gained from engineering fibers has also been applied to develop silk-based biomaterials and in silk-based optical materials fabricated from silkworm silk.[87]

These examples illustrate that protein chain configuration, material morphology, and processing conditions have all been shown to play an important role in the properties of materials engineered from structural proteins. However, it is unclear that common design principles apply to all structural proteins, and for many proteins the specific structure–property relationships are in the early stages of investigation. Nonetheless, several unifying themes are emerging. Many of the protein materials have significant changes in their properties depending on the hydration of the polymers; therefore, understanding the impact of hydration on the molecular-level chain dynamics and mechanical response is a critical aspect of design. Also, control over protein conformation through substrate binding or force-induced unfolding enables engineering of stored length in protein materials, enhancing toughness or triggering responsive swelling and deswelling.[88] Finally, supramolecular interactions such as the hierarchical assembly of collagen fibers or the crystal structures in silks are critical to the performance of materials, and understanding how primary sequence and the resulting secondary structures are translated by processing into the appropriate supramolecular strucutures is an active area. Engineering of minimal protein sequences, as has been demonstrated for silks and elastins, provides a very powerful tool to gain insight into the structure–property relationships through controlled modifications in the protein sequence. Adaptation of the molecular designs discovered in proteins into synthetic or biohybrid polymer systems will provide an important test of the fundamental design principles.

Outlook and Challenges

Despite the development of an active research community in protein materials, commercial products remain relatively few. Commercially viable commodity products such as food additives can be derived from sustainable protein sources, including gelatin and soy protein. Industrial biocatalysis has also found widespread application, for example in the synthesis of penicillin, use of lipases for esterification, or enzymes in laundry detergent formulations. Finally, proteins have found application in glucose sensors and home pregnancy tests. In all cases, cost is one of the primary considerations for the adoption of protein-based materials. For bulk materials (fibers, gels), the cost of the protein must be extremely low to compete with synthetic materials and the material must be readily available in large quantities through scalable manufacturing. Industrial biocatalysis faces similar challenges: the enzyme must be able to be produced in high volume at low cost, and often purity is of secondary concern. As a result of these considerations, continuing developments that lower the cost of engineering protein-based materials should be a high-priority goal for the scientific community.

The most promising markets for future breakthrough protein products are those where the product adds high value. Not surprisingly, there are active research or commercialization efforts in applications such as enzyme technology for the detection of chemical and biological warfare agents,[72] biomedical applications such as targeted drug delivery,[89] and the preparation of tissue engineering matrices.[90] An increased recent focus on biomass utilization and sustainable materials also poses interesting

challenges for researchers and engineers interested in protein materials. Proteins are nature's principle source of nitrogen containing polymers, making them natural materials for exploration of sustainable alternatives to common synthetic polymers such as polyurethanes, polyamides, or polyesteramides. The development of new sustainable resins as replacement polymers or to fill currently undefined market needs will push researchers to develop lower cost production strategies, to understand and control feedstock diversity while producing consistent products, and to leverage the most advanced physics emerging from studies of model systems in ways that apply to polydisperse, impure mixtures of polymers.

As highlighted in this perspective, the development of new concepts in polymer science to understand the complex physics of sequence-specific macromolecules will play an important role in the engineering of these materials. While the chain conformations and thermodynamics of sequence-specific protein polymers provide access to a variety of properties difficult to replicate in synthetic materials, the need to understand their self-assembly, mechanical response, solution properties, and thermodynamic stability will require new scientific advances. As scientists and engineers tackle these challenges, the future of protein materials holds great promise. Advances in molecular biology and polymer chemistry have opened the door to materials that can be used to advance our understanding of soft matter physics and to engineer new products at the biological/synthetic polymer interface. By combining their understanding of basic chemistry and biology, thermodynamics, mechanics, and process systems engineering, chemical engineers have a significant role to play in both answering fundamental scientific questions and defining the future targets for engineering efforts.

Acknowledgments

This work was supported by the National Science Foundation under award DMR-1253306. The author thanks R.E. Cohen and G. Stephanopoulos for helpful comments.

Literature Cited

1. Lewis RV. Spider silk: Ancient ideas for new biomaterials. *Chem Rev*. 2006;106:3762.
2. Lee BP, Messersmith PB, Israelachvili JN, Waite JH. Mussel-inspired adhesives and coatings. In: Clarke DR, Fratzl P, eds. *Ann Rev Mater Res*. 2011;41:99.
3. Miserez A, Weaver JC, Thurner PJ, Aizenberg J, Dauphin Y, Fratzl P, Morse DE, Zok FW. Effects of laminate architecture on fracture resistance of sponge biosilica: Lessons from nature. *Adv Funct Mater*. 2008; 18:1241.
4. Elvin CM, Carr AG, Huson MG, Maxwell JM, Pearson RD, Vuocolo T, Liyou NE, Wong DCC, Merritt DJ, Dixon NE. Synthesis and properties of crosslinked recombinant pro-resilin. *Nature*. 2005;437:999.
5. Jovanovic-Talisman T, Tetenbaum-Novatt J, McKenney AS, Zilman A, Peters R, Rout MP, Chait BT. Artificial nanopores that mimic the transport selectivity of the nuclear pore complex. *Nature*. 2009;457:1023.
6. Jordan P, Fromme P, Witt HT, Klukas O, Saenger W, Krauss N. Three-dimensional structure of cyanobacterial photosystem I at 2.5 angstrom resolution. *Nature*. 2001; 411:909.
7. Schoemaker HE, Mink D, Wubbolts MG. Dispelling the myths – Biocatalysis in industrial synthesis. *Science*. 2003;299:1694.
8. Wang QG, Yang ZM, Zhang XQ, Xiao XD, Chang CK, Xu B. A supramolecular-hydrogel-encapsulated hemin as an artificial enzyme to mimic peroxidase. *Angew Chem Int*. 2007;6:4285.
9. Canalle LA, Lowik D, van Hest JCM. Polypeptide-polymer bioconjugates. *Chem Soc Rev*. 2010;39:329.
10. Maskarinec SA, Tirrell DA. Protein engineering approaches to biomaterials design. *Curr Opin Biotechnol*. 2005;16:422.
11. Merrifie Rb. Automated synthesis of peptides. *Science*. 1965;150:178.
12. Guillier F, Orain D, Bradley M. Linkers and cleavage strategies in solid-phase organic synthesis and combinatorial chemistry. *Chem Rev*. 2000;100;2091.
13. Deming TJ. Polypeptide materials: New synthetic methods and applications. *Adv Mater*. 1997;9:299.
14. Engler AC, Lee HI, Hammond PT. Highly efficient "grafting onto" a polypeptide backbone using click chemistry. *Angew Chem Int*. 2009;48:9334.
15. Fong E,Tzlil S, Tirrell DA. Boundary crossing in epithelial wound healing. *Proc Nat Acad Sci USA*. 2010;107:19302.
16. Wright ER, Conticello V P. Self-assembly of block copolymers derived from elastin-mimetic polypeptide sequences. *Adv Drug Delivery Rev*. 2002;54:1057.
17. Liu JC, Heilshorn SC, Tirrell DA. Comparative cell response to artincial extracellular matrix proteins containing the RGD and CSS cell-binding domains. *Biomacromol*. 2004;5:497.
18. Rabotyagova OS, Cebe P, Kaplan DL. Protein-based block copolymers. *Biomacromol*. 2011;12:269.
19. Terpe K. Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. *Appl Microbiol Biotechnol*. 2003;60:523.
20. Meyer DE, Chilkoti A. Purification of recombinant proteins by fusion with thermally-responsive polypeptides. *Nature Biotechnol*. 1999;17:1112.
21. Kalia J, Raines RT. Advances in Bioconjugation. *Curr Org Chem*. 2010;14:138.
22. Tang Y, Ghirlanda G, Petka W, Nakajima T, DeGrado WF, Tirrell DA. Fluorinated coiled-coil proteins prepared in vivo display enhanced thermal and chemical stability. *Angew Chem Int*. 2001;40:1494.
23. Connor RE, Tirrell DA. Non-canonical amino acids in protein polymer design. *Polym Rev*. 2007;47:9.
24. Xie JM, Schultz PG. Innovation: A chemical toolkit for proteins - an expanded genetic code. *Nat Rev Mol Cell Biol*. 2006;7:775.
25. Flory PJ. *Principles of Polymer Chemistry*. Ithaca, NY: Cornell University Press; 1953.
26. Rubinstein M, Colby RH. *Polymer Physics*. New York: Oxford University Press; 2003.
27. Marszalek PE, Dufrene YF. Stretching single polysaccharides and proteins using atomic force microscopy. *Chem Soc Rev*. 2012;41:3523.

28. Lv S, Dudek DM, Cao Y, Balamurali MM, Gosline J, Li HB. Designed biomaterials to mimic the mechanical properties of muscles. *Nature*. 2010;465:69.
29. Tskhovrebova L, Trinick J. Titin: Properties and family relationships. *Nature Rev Mol Cell Biol*. 2003;4:679.
30. Kim M, Abdi K, Lee G, Rabbi M, Lee W, Yang M, Schofield CJ, Bennett V, Marszalek PE. Fast and forceful refolding of stretched a-helical solenoid proteins. *Biophys J*. 2010;98:3086.
31. Nettles DL, Chilkoti A, Setton LA. Applications of elastin-like polypeptides in tissue engineering. *Adv Drug Delivery Rev*. 2010;62:1479.
32. Kim W, McMillan RA, Snyder JP, Conticello VP. A stereoelectronic effect on turn formation due to proline substitution in elastin-mimetic polypeptides. *J Am Chem Soc*. 2005;12:18121.
33. Urry DW, Luan CH, Parker TM, Gowda DC, Prasad KU, Reid MC, Safavy A. Temperature of polypeptide inverse temperature transition depends on mean residue hydrophobicity. *J Am Chem Soc*. 1991; 113:4346.
34. McDaniel JR, Radford CD, Chilkoti A. A unified model for de novo design of elastin-like polypeptides with inverse transition tmeperatures. *Biomacromol*. 2013. In Press.
35. Bates FS, Fredrickson GH. Block copolymers designer soft materials. *Phys Today*. 1999;52:32.
36. Albert JNL, Epps TH. Self-assembly of block copolymer thin films. *Mater Today*. 2010;13:24.
37. Svenson S. Controlling surfactant self-assembly. *Curr Opin Colloid Interface Sci*. 2004;9:201.
38. Cracknell JA, Vincent KA, Armstrong FA. Enzymes as working or inspirational electrocatalysts for fuel cells and electrolysis. *Chem Rev*. 2008;108;2439.
39. Minteer SD, ed. *Enzyme Stabilization and Immobilization Methods and Protocols*. New York: Springer; 2010;679.
40. Perly B, Douy A, Gallot B. Block copolymers polybutadiene-poly(benzyl-l-glutamate) and polybutadiene-poly(n5-hydroxypropylglutamine) preparation and structural study by x-ray and electron-microscopy. *Macromol Chem Phys*. 1976;177:2569.
41. Lecommandoux S, Achard MF, Langenwalter JF, Klok HA. Self-assembly of rod-coil diblock oligomers based on alpha-helical peptides. *Macromol*. 2001;34:9100.
42. Checot F, Rodriguez-Hernandez J, Gnanou Y, Lecommandoux S. pH-responsive micelles and vesicles nanocapsules based on polypeptide diblock copolymers. *Biomol Eng*. 2007;24:81.
43. Smeenk JM, Otten MBJ, Thies J, Tirrell DA, Stunnenberg HG, van Hest JCM. Controlled assembly of macromolecular beta-sheet fibrils. *Angew Chem Int*. 2005;44:1968.
44. Xu T, Zhao NNk, Ren F, Hourani R, Lee MT, Shu JY, Mao S, Helms BA. Subnanometer Porous Thin Films by the Co-assembly of Nanotube Subunits and Block Copolymers. *Acs Nano*. 2011;5:1376.
45. Shu JY, Tan C, DeGrado W F, Xu T. New design of helix bundle peptide-polymer conjugates. *Biomacromol*. 2008;9:2111.
46. Olsen BD, Segalman RA. Self-assembly of rod-coil block copolymers. *Mater Sci Eng Reports*. 2008;62:37.
47. Lam CN, Olsen BD. Phase transitions in concentrated solution self-assembly of globular protein-polymer block copolymers. *Soft Matter*. 2013;9:2393.
48. Thomas CS, Glassman MJ, Olsen BD. Solid-state nanostructured materials from self-assembly of a globular protein-polymer diblock copolymer. *Acs Nano*. 2011;5: 5697.
49. Reynhout IC, Cornelissen J, Nolte RJM. Self-assembled architectures from biohybrid triblock copolymers. *J Am Chem Soc*. 2007;129:2327.
50. George A, Wilson WW. Predicting protein crystallization from a dilute-solution property. *Acta Crystallogr Sect D Biol Crystallogr*. 1994;50:361.
51. Liu H, Kumar SK, Sciortino F. Vapor-liquid coexistence of patchy models: Relevance to protein phase behavior. *J Chem Phys*. 2007;127:084902.
52. Dumetz AC, Chockla AM, Kaler EW, Lenhoff AM. Protein phase behavior in aqueous solutions: Crystallization, liquid-liquid phase separation, gels, and aggregates. *Biophys J*. 2008;94:570.
53. Odijk T. Protein-macromolecule interactions. *Macromolecules*. 1996;29:1842.
54. Xia Y, Tang SC, Olsen BD. Site-specific conjugation of RAFT polymers to proteins via expressed protein ligation. *Chemical Commun*. 2013;49:2566.
55. Thomas CS, Xu LZ, Olsen BD. Kinetically controlled nanostructure formation in self-assembled globular protein-polymer diblock copolymers. *Biomacromol*. 2012: 13;2781.
56. Hassouneh W, Fischer K, MacEwan SR, Branscheid R, Fu CL, Liu RH, Schmidt M, Chilkoti A. Unexpected multivalent display of proteins by temperature triggered self-assembly of elastin-like polypeptide block copolymers. *Biomacromol*. 2012;13:1598.
57. van Eldijk, MB,Wang JCY, Minten IJ, Li CL, Zlotnick A, Nolte RJM, Cornelissen J, van Hest JCM. Designing two self-assembly mechanisms into one viral capsid protein. *J Am Chem Soc*. 2012;134:18506.
58. Zhang SG. Fabrication of novel biomaterials through molecular self-assembly. *Nature Biotechnol*. 2003:21; 1171.
59. Grove TZ, Forster J, Pimienta G, Dufresne E, Regan L. A modular approach to the design of protein-based smart gels. *Biopolymers*. 2012;97:508.
60. Kas OY, Charati MB, Rothberg LJ, Galvin ME, Kiick KL. Regulation of electronic behavior via confinement of PPV-based oligomers on peptide scaffolds. *J Mater Chem*. 2008;18:3847.
61. Kovacs EW, Hooker JM, Romanini DW, Holder PG, Berry KE, Francis MB. Dual-surface-modified bacteriophage MS2 as an ideal scaffold for a viral capsid-based drug delivery system. *Bioconjugate Chem*. 2007;18:1140.
62. Dogic Z, Fraden S. Smectic phase in a colloidal suspension of semiflexible virus particles. *Phys Rev Lett*. 1997; 78:2417.
63. Lee SW, Mao CB, Flynn CE, Belcher AM. Ordering of quantum dots using genetically engineered viruses. *Science*. 2002;296:892.

64. Yan H, Saiani A, Gough JE, Miller AF. Thermoreversible protein hydrogel as cell scaffold. *Biomacromol*. 2006;7:2776.
65. Yan CQ, Altunbas A, Yucel T, Nagarkar RP, Schneider JP, Pochan D J. Injectable solid hydrogel: mechanism of shear-thinning and immediate recovery of injectable beta-hairpin pq)tide hydrogels. *Soft Matter*. 2010;6: 5143.
66. Storrie H, Guler MO, Abu-Amara SN, Volberg T, Rao M, Geiger B, Stupp SI. Supramolecular crafting of cell adhesion. *Biomater*. 2007:28:4608.
67. Wall BD, Diegelmann SR, Zhang SM, Dawidczyk TJ, Wilson WL, Katz HE, Mao HQ, Tovar JD. Aligned Macroscopic Domains of Optoelectronic Nanostructures Prepared via Shear-Flow Assembly of Peptide Hydrogels. *Adv Mater*. 2011;23:5009.
68. Lamm MS, Rajagopal K, Schneider JP, Pochan DJ. Laminated morphology of nontwisting beta-sheet fibrils constructed via peptide self-assembly. *J Am Chem Soc*. 2005;127:16692.
69. Tracewell CA, Arnold FH. Directed enzyme evolution: climbing fitness peaks one amino acid at a time. *Curr Opin Chem Biol*. 2009;13:3.
70. Schiraldi C, De Rosa M. The production of biocatalysts and biomolecules from extremophiles. *Trends Biotechnol*. 2002;20:515.
71. Ferreira M, Fiorito PA, Oliveira ON, de Torresi SIC. Enzyme-mediated amperometric biosensors prepared with the Layer-by-Layer (LbL) adsorption technique. *Biosens Bioelectron*. 2004;19:1611.
72. Drevon GF, Russell AJ. Irreversible immobilization of diisopropylfluorophosphatase in polyurethane polymers. *Biomacromol*. 2000;1:571.
73. Cardoso MB, Smolensky D, Heller WT, Hong KL, O'Neill H. Supramolecular assembly of biohybrid photoconversion systems. *Energy Environ Sci*. 2011;4:181.
74. Govardhan CP. Crosslinking of enzymes for improved stability and performance. *Curr Opin Biotechnol*. 1999; 10;331.
75. Mohidem NA, Bin Mat H. Catalytic activity and stability of laccase entrapped in sol-gel silica with additives. *J Sol-Gel Sci Technol*. 2012;61:96.
76. Andersson MM, Breccia JD, Hatti-Kaul R. Stabilizing effect of chemical additives against oxidation of lactate dehydrogenase. *Biotechnol Appl Biochem*. 2000;32:145.
77. Triantafyllou AO, Wehtje E, Adlercreutz P, Mattiasson B. How do additives affect enzyme activity and stability in nonaqueous media? *Biotechnol Bioeng*. 1997;5:67.
78. Malik R, Hall CK, Genzer J. Protein-like copolymers (plcs) as compatibilizers for homopolymer blends. *Macromolecules*. 2010;43:5149.
79. Storm C, Pastore JJ, MacKintosh FC, Lubensky TC, Janmey PA. Nonlinear elasticity in biological gels. *Nature*. 2005;435:191.
80. Nairn KM, Lyons RE, Mulder RJ, Mudie ST, Cookson DJ, Lesieur E, Kim M, Lau D, Scholes FH, Elvin CM. A synthetic resilin is largely unstructured. *Biophys J*. 2008;95:3358.
81. Qin GK, Hu X, Cebe P, Kaplan DL. Mechanism of resilin elasticity. *Nature Commun*. 2012:3.
82. Charati MB, Ifkovits J L, Burdick JA, Linhardt JG, Kiick KL. Hydrophilic elastomeric biomaterials based on resilin-like polypeptides. *Soft Matter*. 2009;5: 3412.
83. Teule F, Miao YG, Sohn BH, Kim YS, Hull JJ, Fraser MJ, Lewis RV, Jarvis DL. Silkworms transformed with chimeric silkworm/spider silk genes spin composite silk fibers with improved mechanical properties. *Proc Nat Acad Sci USA*. 2012;109:923.
84. Hinman MB, Jones JA, Lewis RV. Synthetic spider silk: a modular fiber. *Trends Biotechnol*. 2000:18;374.
85. Vepari C, Kaplan DL. Silk as a biomaterial. *Progr Polym Sci*. 2007;32:991.
86. Jin HJ, Kaplan DL. Mechanism of silk processing in insects and spiders. *Nature*. 2003;424:1057.
87. Parker ST, Domachuk P, Amsden J, Bressner J, Lewis JA, Kaplan DL, Omenetto FG. Biocompatible silk printed optical waveguides. *Adv Mater*. 2009;21:2411.
88. Sui ZJ, King WJ, Murphy WL. Dynamic materials based on a protein conformational change. *Adv Mater*. 2007; 19:3377.
89. Callahan DJ, Liu WE, Li XH, Dreher MR, Hassouneh W, Kim M, Marszalek P, Chilkoti A. Triple stimulus-responsive polypeptide nanoparticles that enhance intratumoral spatial distribution. *Nano Lett*. 2012;12:2165.
90. Foo C, Lee JS, Mulyasasmita W, Parisi-Amon A, Heilshorn SC. Two-component protein-engineered physical hydrogels for cell encapsulation. *Proc Nat Acad Sci USA*. 2009;106:22067.