FULL PAPER

The Effect of Protein Fusions on the Production and Mechanical Properties of Protein-Based Materials

*Shang-Pu Tsai, David W. Howell, Zhao Huang, Hao-Ching Hsiao, Yang Lu, Kathleen S. Matthews, Jun Lou, and Sarah E. Bondos**

Proteins implement most of the vital molecular functions of living organisms, including structural support, energy generation, biomolecule sensing, and chemical catalysis, storage, and degradation. While capturing proteins in materials could create devices that mimic these functions, this process is challenging due to the sensitivity of protein structure to the chemical environment. Using recombinant DNA methods, specific functions can be incorporated by fusing the gene encoding a self-assembling protein and the desired functional protein, to produce a single polypeptide that self-assembles into functionalized materials. However, the functional protein has the potential to disrupt protein production, protein assembly, and/or the structure and mechanical properties of the resulting materials. 24 fusion proteins are created based on Ultrabithorax, a *Drosophila* transcription factor that self-assembles into materials in vitro. The appended proteins dictate the solubility and purification yield of the corresponding protein fusions. Any loss of solubility and yield can be mitigated by fusing a third protein that is highly soluble. All protein fusions self-assemble equally well to produce materials with similar morphologies. Fusing enhanced green fluorescent protein to Ultrabithorax influences mechanical properties of the resulting fibers. It is concluded that a far wider range of proteins can be successfully incorporated into elastomeric protein-based materials than originally anticipated.

1. Introduction

Proteins implement most of the vital molecular functions of living organisms, including providing structural support, generating energy, sensing biomolecules, and catalyzing, storing, and degrading important biomolecules. Proteins are particularly adept at highly specific molecular recognition, which can be adapted to bind virtually any ligand, ranging from small chemicals to specific types of cells. Such interactions are often regulated by the chemical environment, by binding additional ligands (allostery), or by post-translational modification of the protein.

Devices that capture proteins in materials have the potential to mimic these functions and be regulated by eternally applied factors. However, this approach is technically challenging. The three-dimensional structures of proteins are maintained by noncovalent bonds, which are sensitive to the surrounding medium. Perturbations of this chemical environment can easily result in loss of protein function. Proteins are commonly incorporated into materials either by physically trapping them within the matrix of the materials during assembly or by covalently crosslinking them to the surface of the materials post-assembly. Both of these approaches can result in the loss of functional proteins. Physically trapped proteins can be inactivated by the harsh chemical environment often used to instigate the assembly of the proteins into materials. Furthermore, the functional proteins are free to diffuse out of the materials.[1] Although covalent crosslinking tethers functional proteins to the materials, the crosslinking agent can also inactivate the appended protein or remain embedded in the materials, rendering them toxic to cells.[1,2] In addition, materials held together by noncovalent bonds may be too fragile for chemical modifications after assembly.[1] Finally, depending on the specificity of the cross-linking strategy, some portion of the functional protein may be oriented such that the materials structure blocks ligand binding and hence protein function.[3]

For materials composed of recombinant proteins, the use of protein fusions provides an attractive alternative. By fusing a gene encoding a functional protein to a gene encoding a self-assembling protein, a single polypeptide can be produced that contains the sequences of both proteins. This fusion protein should retain both the functional and self-assembly properties of the parent proteins. This method offers several advantages for incorporating functional proteins: (i) materials assembly and functionalization can be combined into a single step,

S.-P. Tsai, Dr. D. W. Howell, H.-C. Hsiao, Prof. S. E. Bondos
Department of Molecular and Cellular Medicine
Texas A&M Health Science Center
College Station, TX 77843, USA
E-mail: sebondos@tamhsc.edu
Dr. Z. Huang, Prof. K. S. Matthews, Prof. S. E. Bondos
Department of Biochemistry and Cell Biology
Rice University
Houston, TX 77005, USA
Prof. Y. Lu
Department of Mechanical and Biomedical Engineering
City University of Hong Kong
Kowloon, Hong Kong
Prof. J. Lou
Department of Materials Science and Nanoengineering
Rice University
Houston, TX 77005, USA

DOI: 10.1002/adfm.201402997

(ii) stoichiometric levels of functionalization can be achieved, (iii) covalent attachment prevents loss of the functional protein due to diffusion, (iv) toxic by-products (remnants of chemical cross-linking) are not created, and (v) the functional proteins can be patterned within the materials.[1,4] In addition, all of the appended proteins have a uniform orientation, although there are only two possible points of attachment – the N- and C-termini of the self-assembling protein. The gene fusion approach requires mild conditions for materials assembly that will not perturb the structure of the appended functional protein. Even so, full-length protein fusions have been successfully used to functionalize materials composed of many proteins, including elastin, silk, amyloid-forming proteins, and the *Drosophila melanogaster* transcription factor Ultrabithorax (Ubx).[4–6]

Not every functional protein is likely to be a good candidate for incorporation into materials via gene fusion. For instance, unstable or insoluble functional proteins could hamper expression of the fusion protein, or large/multimeric functional proteins may misposition the self-assembling protein, thus altering the mechanical properties of the materials or even preventing materials assembly. We have fused 24 proteins to Ubx to examine their impact on Ubx protein production and assembly. Ubx self-assembles rapidly in gentle buffers to form films and fibers.[7] Ubx materials are biocompatible, strong, and remarkably extensible. Furthermore, proteins fused to Ubx retain their activity once incorporated into materials.[4,8,9] The proteins fused to Ubx were selected to test a range of sizes, structures, stabilities, solubilities, and charges. We found that the appended protein had a large effect on protein production, with the solubility and quaternary structure of the appended protein being the best predictors for success. In contrast, the ability to self-assemble into materials was dominated by Ubx, and the presence or identity of a fused protein had little impact on materials assembly. Although the appended protein can alter the mechanical properties of the materials as demonstrated by the enhanced green fluorescent protein (EGFP)–Ubx fusion, these properties, like those of Ubx fibers, (i) can be adjusted by varying fiber diameter and (ii) remain similar to the extracellular matrix protein elastin.[7,8]

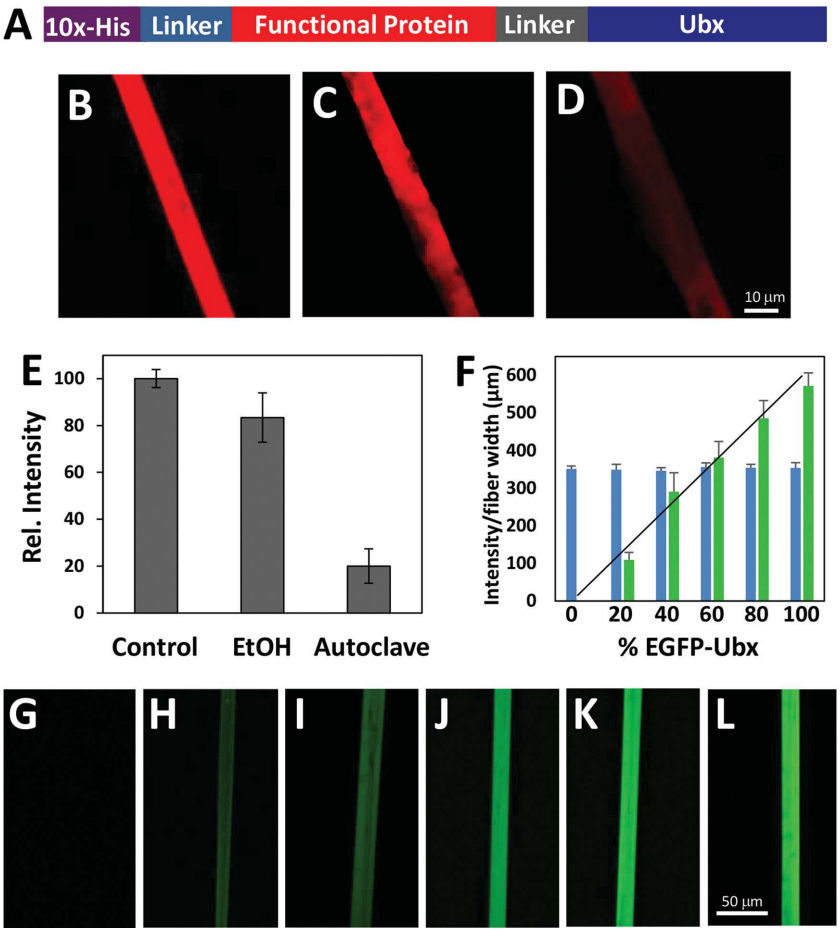

Figure 1. Incorporating proteins into Ubx materials by gene fusion. A) Schematic depicting a Ubx fusion protein, in which an N-terminal histidine tag, linker, and a functional protein are connected to Ubx via a flexible linker in a single polypeptide chain. B–E) mCherry protein, assembled into Ubx materials, retains some activity under harsh sterilization methods. B) Photomicrograph of mCherry–Ubx fibers. C) An mCherry–Ubx fiber that has been incubated for 30 min in 100% EtOH. D) An mCherry–Ubx fiber that has been autoclaved for 30 min. E) Relative fluorescence of treated and untreated mCherry–Ubx fibers, $N = 3$. F) Fluorescence from fibers produced from mixtures of EGFP–Ubx and Ubx correlates with the percentage of EGFP–Ubx in the mixture, $N = 3$. G–L) Fluorescent micrographs of fibers composed of 0%, 20%, 40%, 60%, 80%, and 100% EGFP–Ubx, respectively.

2. Results and Discussion

2.1. Generating Ubx Fusion Proteins

Previously, we demonstrated that four proteins could be fused to Ubx and retain their function once incorporated into materials.[4] However, all of these proteins were stable monomers—traits that may have facilitated their successful use. Conversely, less stable functional proteins could potentially hamper expression of the fusion protein, highly charged proteins could inhibit assembly by charge–charge repulsion, or very large/multimeric functional proteins may misposition the self-assembling region of Ubx, creating weaker materials or possibly even preventing assembly. Our goal was to test the impact of fusing single proteins to Ubx on protein production, materials assembly, and the mechanical properties of materials, and identify properties of the appended proteins that predict whether a specific fusion can be successfully produced and assembled into useful materials. To accomplish this, we fused 24 peptides or proteins to the N-terminus of Ubx (**Figure 1**A and **Table 1**). We also tested several fusions to the C-terminus of Ubx, but these proteins did not express well in *E. coli*, and further C-terminal fusions were not pursued. A similar effect of

Table 1. Properties of proteins fused to Ubx.

Fusion category	Fusion	Size [kDa]	Charge at pH 8	SCOP classification	Quaternary structure	Linker
Peptides[a)]	GSGSGS	0.45	–0.5	Peptide	Monomer	H
	RGD	0.78	–0.5	Peptide	Monomer	GSGSH
	WRW	0.98	0.5	Peptide	Monomer	GSGSH
	YKLKYY	1.3	1.5	Peptide	Monomer	GSGSH
Cytokines	SDF-1α[b)]	8.2	7.5	α+β	Dimer	GH
	bFGF	17.4	8.5	β	Dimer	GH
	VEGF	19.6	–3.2	β	Dimer	GH
	Osteopontin	35.6	–46.9	N/A[c)]	Monomer	GH
Fluorescent protein	AmCyan	25.5	–2.6	α+β	Tetramer	GH
	mCherry	26.9	–6.6	α+β	Monomer	GH
	EBFP	27.1	–8.9	α+β	Monomer	GH
	EGFP	27.1	–9.0	α+β	Monomer	GH
Fusion tags	SUMO	12.2	–6.5	α+β	Monomer	GSGSH
	Thioredoxin	12.8	–6.0	α/β	Dimer	GSGSH
	GST	26.5	–4.6	α	Dimer	GSGSH
	MBP	41.3	–10.9	α/β	Monomer	GSGSH
	NusA	55.8	–41.4	α	Monomer	GSGSH
Enzymes	PFK	34.6	–4.1	α/β	Tetramer	GSGSH
	L-PYK	58.9	–4.1	β	Tetramer	GGSGSH
	Luciferase	60.9	–5.6	α+β, α/β, β	Monomer	GH
Ligand binding protein	TneSSB	16.7	–7.5	β	Tetramer	GSGSH
	TmaSSB	16.7	–7.5	β	Tetramer	GSGSH
	Myoglobin	17.6	2.7	α	Monomer	GH
	FN	30.0	–6.4	β	Monomer	GH

[a)]The sequences of appended peptides are listed; [b)]Abbreviations: SDF-1α, stromal cell-derived factor-1; bFGF, basic fibroblast growth factor; VEGF, vascular endothelial growth factor; EBFP, enhanced blue fluorescent protein; EGFP, enhanced green fluorescent protein; SUMO, small ubiquitin-like modifier protein; GST, glutathione S-transferase; MBP, maltose binding protein; PFK, phosphofructokinase; L-PYK, liver pyruvate kinase; TneSSB, *Thermotoga neapolitana* single-stranded DNA binding protein; TmaSSB, *Thermatoga maritime* single-stranded DNA binding protein; FN, type III domain 8–10 of Fibronectin; [c)]Not applicable.

fusion order on protein expression and activity has been previously reported for other protein systems.[5,10]

By carefully selecting peptides and proteins, we were able to test a wide range of physical properties: size (0.45–60.9 kDa), predicted charge (–47 to +8.5 at pH = 8), stability (intrinsically disordered as well as thermostable proteins), secondary structure (α, β, α+β, α/β), quaternary structure (monomer, dimer, and tetramer), and solubility. When possible, groups of related proteins were selected so that as few of these properties as possible were simultaneously varied. Consequently, we focused on 6 categories of fusions: peptides, cytokines, fluorescent proteins, solubility tags, enzymes, and ligand binding proteins (Table 1). A series of charged peptides were generated to test the impact of charge density on the appended moiety as well as provide additional positively charged fusions. In particular, monomeric and tetrameric fluorescent proteins, which have similar charges, monomer sizes, and tertiary structures, provided a direct test of the role of quaternary structure in materials production. The series of solubility tags are collectively far more soluble than the other proteins. To examine stability, we compared Ubx fused to osteopontin (Opn), an intrinsically disordered protein, with mesostable human proteins, and proteins derived from thermophilic bacteria.

Peptide linking sequences were included between this his-tag and the appended protein, and between the appended proteins and Ubx to ensure both proteins had sufficient space to fold and function. These sequences are rich in glycine (to provide flexibility) and serine or histidine (to provide solubility) (Table 1). For fusions of Ubx to peptides, we used a Gly–Ser–Gly–Ser–His linking sequence. For full-length proteins fused to Ubx, the sequence linking the two proteins was Gly–His. The one exception was pyruvate kinase–Ubx, in which the C-terminus of this tetrameric protein has the potential to interfere with subunit contacts. Therefore, in pyruvate kinase–Ubx we used a longer and more flexible Gly–Gly–Ser–Gly–Ser–His linker.

2.2. Ubx Protein Fusions Are More Stable in Materials Than as Free Monomers

Purified recombinant proteins are both expensive and labile.[11] Single-pot synthesis of functionalized materials would

Table 2. Definition of terms.

Term	Definition
Expression	The amount of soluble and insoluble Ubx or fusion protein generated in *E. coli*
Yield	The amount of soluble protein purified from a 1 L culture of *E. coli* after 16 h induction with IPTG at 25 °C
Solubility	The maximum amount of protein that can be dissolved in 1.1 M ammonium sulfate
Assembly	The ability of a protein to self-assemble into materials

significantly lower production costs. The structure and function of protein monomers are notoriously sensitive to environmental conditions. Since the material is in an altered environment relative to bulk solvent, it is possible that the material itself could stabilize or destabilize a fused protein. Conversely, confining the functional protein within the materials could stabilize the functional protein by preventing the motions required for denaturation. To determine whether incorporation of protein fusions alters the stability of the appended protein, we examined the structure and fluorescence of mCherry–Ubx fibers under the denaturing conditions imposed by sterilization procedures: autoclaving and ethanol washing. Ubx materials are remarkably robust and can even survive boiling.[7] Interestingly, mCherry–Ubx fibers remained intact and retained almost 80% of fluorescence intensity after incubation in 100% ethanol for 30 min (Figure 1B–E). After autoclaving mCherry–Ubx fibers in steam sterilizer for 20 min (121 °C, 1.27 kg cm^{-2}), 20% of the fluorescence intensity remained. In both cases, the fiber structure appeared undamaged by treatment. For comparison, treatment with alcohol and autoclaving caused monomeric mCherry–Ubx and mCherry to unfold, release their chromophores, and precipitate (data not shown). These results imply that Ubx fibers remarkably stabilize proteins incorporated by gene fusion. These results are consistent with prior studies by other laboratories in which immobilization improved enzyme efficiency or stability.[12]

2.3. The Appended Proteins Determine Ubx Fusion Protein Expression Levels

In addition to the fibers altering the fused protein, the fused protein could impact various aspects of materials production, including expression, solubility, purification yield, and assembly (**Table 2**). Materials assembly is strongly dependent on protein concentration (Figure S1, Supporting Information), and therefore any factors that alter protein production or assembly will also impact the amount of materials that can be produced. In general, high levels of protein expression are desirable because they increase the yield of purified protein and ultimately the amount of materials produced. However, robust expression of a polymer-forming protein, such as Ubx, in bacteria may lead to aggregation/inclusion body formation and thus lower the total yield. It is well established that protein fusions can dramatically alter the expression of proteins.[13] To determine how fusions alter Ubx expression in *E. coli*, we used Western blots to compare the expression of Ubx and its variants 16 h after induction with Isopropyl β-D-1-thiogalactopyranoside (IPTG). This method measures total (soluble and insoluble) protein produced by the bacteria (**Figure 2**A and Figure S2, Supporting Information). We found the results varied significantly among the Ubx fusion proteins, ranging from more than tripling protein expression to reducing expression to nearly undetectable levels.

We also measured the expression of a subset of the functional proteins not fused to Ubx. Proteins were selected to test

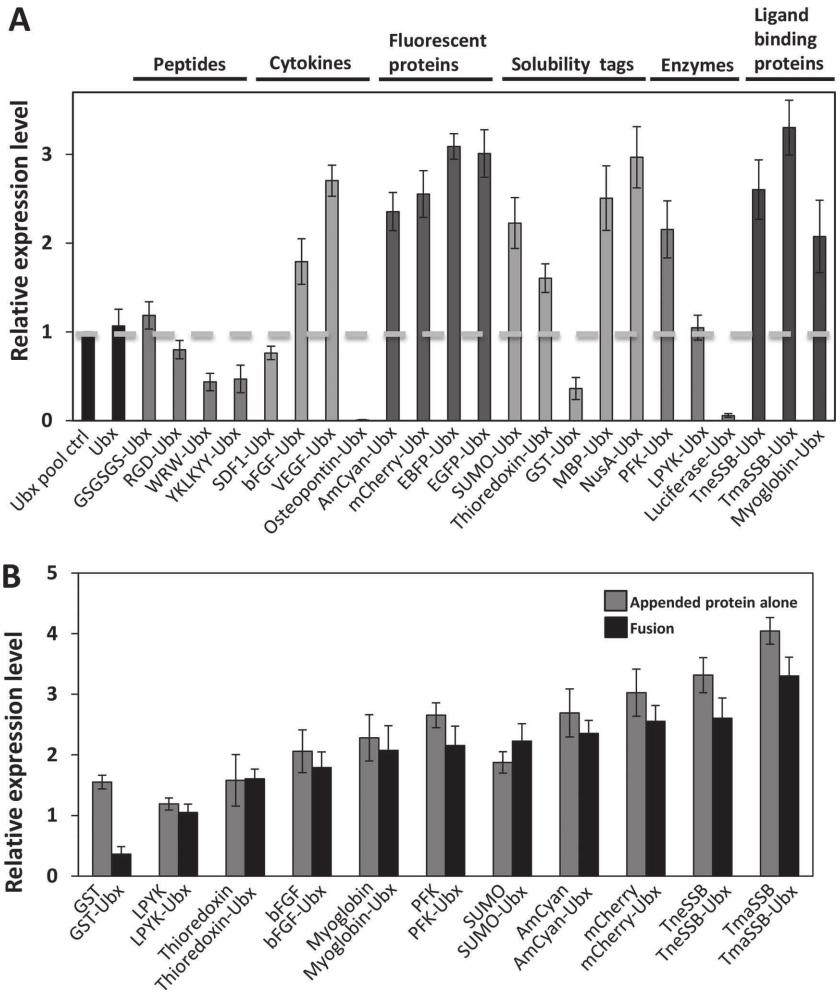

Figure 2. Expression of Ubx fusion proteins varies significantly with the identity of the appended protein. A) Expression levels for all tested Ubx fusion proteins. B) Comparison of the expression of a subset of Ubx fusions proteins (black bars) with the isolated functional proteins. Proteins were selected to sample the range of observed expression levels. Expression was measured by Western blot of whole-cell lysates. $N = 3$ for experiments in both panels.

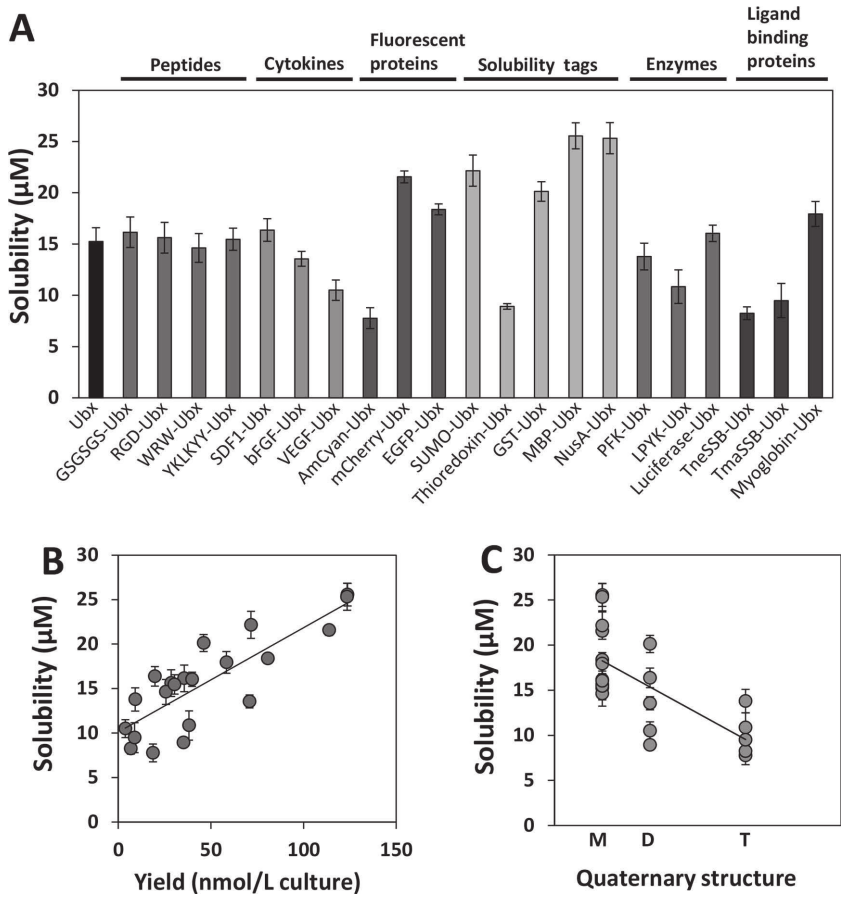

Figure 3. Assessing the solubility of Ubx and Ubx fusion proteins. A) Relative solubility of Ubx-fusions measured by ammonium sulfate precipitation, measured as the concentration of protein remaining in a solution containing 1.1 M ammonium sulfate (Figure S4, Supporting Information), $N = 3$. B) Solubility correlates with protein yield ($r = 0.83$), $N = 2$ for measurements of yield. C) The quaternary structure of the appended protein correlates with reduced protein solubility ($r = 0.69$). M, monomer; D, dimer; T, tetramer. $N = 3$.

the range of productive expression levels observed. The expression of the isolated functional protein and the corresponding Ubx fusion protein correlate very well, suggesting that the protein appended to Ubx determines the expression level (Figure 2B). We examined whether properties of the fused protein could predict expression levels. Expression does not appear to correlate with the molecular weight, the predicted charge density, or the quaternary structure of the appended proteins (Figure S3, Supporting Information).

Protein solubility is another factor that influences the production of recombinant proteins.[13] Protein solubility was determined by measuring the resistance of each protein to ammonium sulfate precipitation. In the presence of ammonium sulfate, solubility determines the concentration of protein that remains in solution.[14] Our measurements were performed at 1.1 M ammonium sulfate, at which concentration differences in the solubility of Ubx fusions are most pronounced (Figure S4, Supporting Information). This observation is consistent with data from other proteins.[14] The solubility varies >3-fold among Ubx fusion proteins (**Figure 3**), and fusions can either increase or decrease solubility relative to Ubx. The measured solubilities roughly correlate with the yield from the protein purification ($r = 0.83$). Monomeric proteins were more soluble, on average, when fused with Ubx than dimeric or tetrameric proteins. The reduced solubility of dimeric and tetrameric Ubx fusions is not an artifact of protein selection; the corresponding isolated proteins have solubilities similar to the monomeric proteins tested (Figure S5, Supporting Information). Because fusion solubility is related to the solubility of the appended protein,[13] the solubility and quaternary structure of the appended protein can be used to estimate the relative purification yield of the Ubx fusion. Neither the size nor the charge density of the appended proteins correlated with the solubility (Figure S6, Supporting Information). Furthermore, the expression level of the fusion protein also did not correlate with the solubility of the appended protein (Figure S6, Supporting Information).

2.4. The Appended Proteins Have No Significant Effect on Materials Assembly

Solubility and stability tags are widely used to prevent aggregation, assist folding, and improve protein yield.[13] Thus, they have the potential to positively contribute to the production of protein-based materials. On the other hand, very soluble proteins may also impede protein polymerization and thus reduce materials production. For example, fusing glutathione S-transferase (GST) to a Huntington protein fragment can prevent aggregation to amyloid.[15] In contrast, GST improves Ubx solubility without impacting self-assembly (Figure 3 and **Figure 4**). Although the reason for this difference is not known, one reasonable hypothesis is that the GST protein (26.5 kDa) is more easily accommodated in the extensible Ubx materials than in rigid amyloid fibrils. Furthermore, the density of GST would be lower in Ubx materials because Ubx is a much larger protein (13.7 kDa Huntington fragment vs 40 kDa Ubx). Nevertheless, proteins appended to Ubx can potentially impede self-assembly by many additional mechanisms. If the appended protein carries a large charge, then charge–charge repulsion could inhibit the intermolecular interactions required for self-assembly. The appended protein could bind Ubx and block the assembly interface. This mechanism is especially concerning for Ubx, since, like many transcription factors, it has a net positive charge, whereas many cytosolic or secreted proteins have a negative charge. Finally, fusing large or multimeric proteins could misposition the self-assembling protein such that it cannot form the necessary intermolecular interactions.

To determine whether the presence of the functional proteins compromises material assembly, we compared the ability

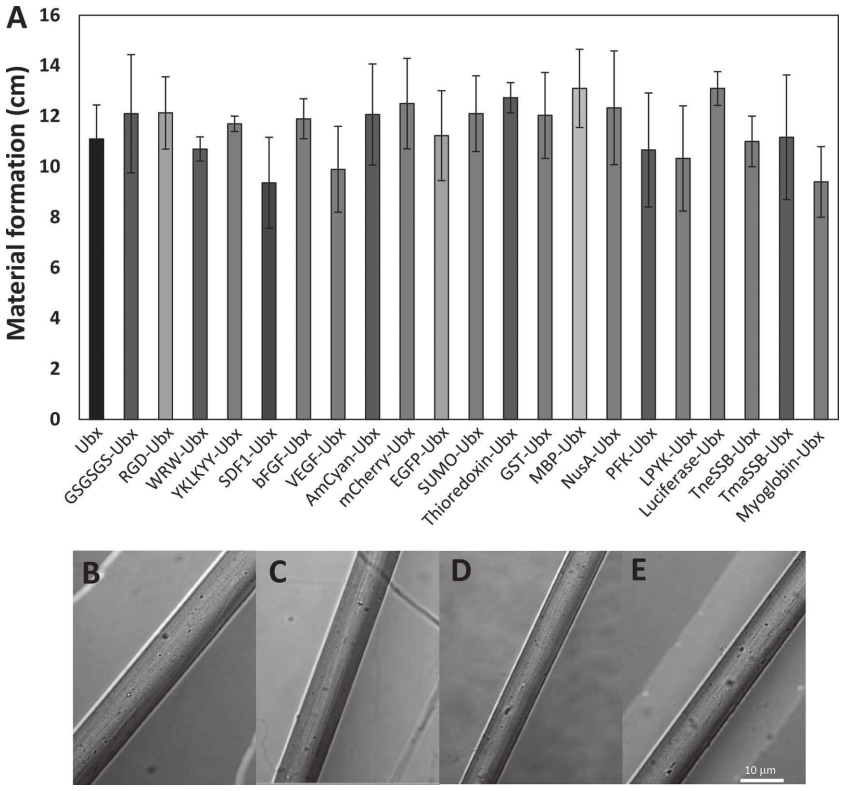

Figure 4. Materials assembly is not significantly impacted by protein fusion. For all experiments, 50×10^{-9} M of the Ubx variants were incubated at room temperature for 16 h. Materials formation was measured as length of fiber that can be drawn from the surface of the assembly solution, $N = 3$. B–E) Differential interference contrast microscopy of fibers formed by Ubx, VEGF–Ubx, MBP–Ubx, and phosphofructokinase–Ubx, respectively. Images of the remaining Ubx fusions are in Figure S8, Supporting Information.

of Ubx fusions to assemble under constant conditions. Ubx fusions were incubated at a fixed protein concentration (50×10^{-9} M), temperature (25 °C), and humidity (40%–60%) to form a film at the air–water interface. This film was subsequently drawn into fibers. The length of fiber produced depends, in part, on the ability of Ubx variants to assemble; proteins that assemble poorly produce less film and thus shorter fibers.[7] Surprisingly, despite the variations in size, charge, stability, and quaternary structure, all Ubx-fusions formed materials equally well (Figure 4A). This similarity was not an artifact of the selected conditions, since the similarities persist at different protein concentrations (Figure S1, Supporting Information). Given we observed no significant variation, there was of course no correlation with molecular weight, charge density, or quaternary structure (Figure S7, Supporting Information). Confocal microscopy confirms that fibers produced by all Ubx-fusions have nearly identical morphology (Figure 4B–E and Figure S8, Supporting Information).

If Ubx and Ubx fusion proteins do indeed form materials equally well, then a mixture of the two proteins should produce fibers containing the same ratio of the two proteins. However, if one protein assembles more efficiently than the other, then the protein content in the fiber should be biased toward the protein that assembles more efficiently. We generated a series of assembly trays containing the same total concentration of protein, but different ratios of Ubx to EGFP–Ubx. The EGFP content of fibers drawn from these trays was measured by fluorescence microscopy (Figure 1F–L). The intensity of the natural blue fluorescence of Ubx fibers was used to gauge the reproducibility of these measurements. We found the EGFP–Ubx content in fibers linearly correlates with percentage of EGFP–Ubx in the original protein mixture, indicating that these two proteins are equally able to self-assemble into materials (Figure 1F–L). Together, these results indicate that the strong tendency of Ubx protein to self-assemble is not altered by the use of fusion proteins.

2.5. Double Ubx Fusion Proteins Self-Assemble

Despite the differences in the properties of the appended proteins, the possibility exists that all Ubx variants only appear to self-assemble equally well because we did not select proteins that produce a difference. In particular, the size and charge of the appended proteins were expected to have a large impact on Ubx self-assembly. To push both of these boundaries, we fused two NusA proteins to Ubx. Thus, the appended portion adds 112 kDa, which is approximately 3 times the size of Ubx. Furthermore, the double fusion has a theoretical charge of –83 at pH = 8.0. To our surprise, the NusA–NusA–Ubx protein still self-assembled into materials (**Figure 5**A).

The fact that double fusion proteins can be successfully produced and incorporated into materials creates two additional advantages. First, by using double fusions, two proteins can be simultaneously incorporated into Ubx at stoichiometric ratios,

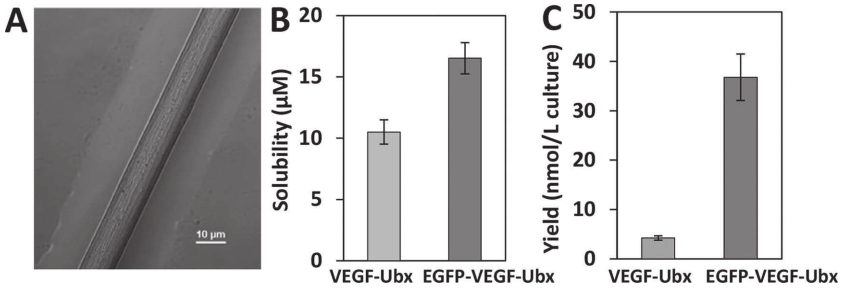

Figure 5. Double fusions solve low solubility/yield problems. A) Ubx can assemble large, highly charged triple fusion proteins, i.e., NusA–NusA–Ubx (112 kDa, –82.8 charge at pH = 8), into materials. B,C) Fusing VEGF–Ubx with EGFP, a soluble protein, increases both the solubility ($N = 3$) and the purification yield ($N = 2$).

increasing the functional capacity of the materials. Second, the solubility or purification yield of difficult fusions can be potentially enhanced by adding a well-behaved protein. To test this possibility, we created EGFP–vascular endothelial growth factor (VEGF)–Ubx. Both the solubility and purification yield of the double fusion was improved relative to VEGF–Ubx (Figure 5). In a second test, we fused NusA to the N-terminus of Opn–Ubx, which does not express in *E. coli*. In contrast, the NusA–Opn–Ubx double fusion does express (Figure S9A,B, Supporting Information). Finally, a fusion of domains 8–10 of fibronectin type III to Ubx could not be reliably expressed in *E. coli* due to severe proteolysis. Additional fusion of GST to the N-terminus reduces this proteolysis, allowing production of the full-length fusion protein (Figure S9C, Supporting Information). Thus, double fusions can facilitate the production of many problem proteins. Notably, this solution requires a self-assembling protein, like Ubx, that can accommodate large fusions.

2.6. A Comparison of the Mechanical Properties of Ubx and EGFP–Ubx

In addition to altering monomer production and materials assembly, fusing proteins to Ubx may also impact the mechanical properties of the resulting materials. Fiber diameter dictates the mechanical properties of unmodified Ubx materials: small diameter (5–10 μm) Ubx fibers have high breaking stress and low breaking strain and undergo elastic deformation, whereas larger Ubx fibers exhibit a combination of elastic and plastic deformation, are highly extensible, and rupture at lower stress (8). To test whether protein fusions can impact the mechanical properties of Ubx fibers, we examined EGFP–Ubx fibers. Similar to Ubx fibers, the mechanical properties of EGFP–Ubx fibers were strongly dependent on fiber diameter (**Figure 6**). In addition, EGFP–Ubx fibers wrinkle, similar to Ubx fibers, upon unloading, to an extent that depends on both strain and fiber diameter (Figure S10, Supporting Information).[8] Compared to Ubx fibers of the same diameter, the breaking strain is reduced in the EGFP–Ubx fusion fibers. Indeed, a 20 μm Ubx fiber is ≈120% extensible, whereas a 20 μm EGFP–Ubx fiber is only 40% extensible. Since the mechanical properties of all fibers are diameter-dependent, a specific breaking strain can still be achieved by generating fibers with the appropriate diameter. Even with these alterations, it is important to note that the mechanical properties of EGFP–Ubx fibers are still similar to natural elastin, and thus lie in a biologically relevant range.

At any given fiber diameter, the breaking stress of the EGFP–Ubx chimera fibers is slightly higher than that for the Ubx fibers, suggesting that the presence of EGFP may contribute to the tensile strength of the fibers. This increased strength may be due to intermolecular interactions in the materials between the negatively charged EGFP and the positively charged DNA binding domain in the Ubx protein.[16] From these data, we conclude that the chimeric partner for Ubx is capable of influencing the mechanical properties of the fiber. Consequently, mechanical properties should be considered when engineering a novel chimeric monomer for materials assembly.

3. Conclusions

Protein fusions provide a facile mechanism to incorporate a broad array of chemical functions into protein-based materials. Not only can proteins be safely incorporated by this method, incorporation into materials can also enhance their stability. However, the presence of the appended protein has the potential to compromise production of the fusion protein monomers, assembly of the monomers into materials, and the mechanical properties of the functionalized materials. We have fused 24 proteins to the *D. melanogaster* transcription factor Ubx to examine their impact on Ubx assembly and properties. In this study, we find that inclusion in Ubx materials stabilizes the appended proteins. Because the fusions do not impact Ubx assembly, the concentration of functional protein in the materials can be predictably controlled by co-assembling the Ubx fusion protein with unmodified Ubx. The proteins appended to Ubx by gene fusion were selected to vary, as systematically as possible, the size, fold, quaternary structure, stability, solubility, and charge density. We find that the appended protein had a large effect on the production of the fusion protein, with the solubility and quaternary structure of the appended protein best predicting successful expression and purification. In contrast, the ability to self-assemble into materials was dominated by Ubx, and the presence or identity of a fused protein had little impact on materials assembly. Although an appended protein can alter the mechanical properties of the materials, these properties still lie in a biologically relevant range. Finally, two proteins can be simultaneously fused to Ubx to create double fusions that still assemble into functionally active materials. These double fusions can be used to incorporate multiple activities at stoichiometric levels into the protein materials, or to improve the solubility or expression of a poorly performing single Ubx fusion.

Generating materials from fusion proteins and from a mixture of protein monomers allows easy control of the concentration of the functional protein and enables incorporation of multiple functionalities. Because all Ubx-based materials self-assemble equally well, mixing Ubx fusions with each other or with plain Ubx precisely and predictably determines the concentration of

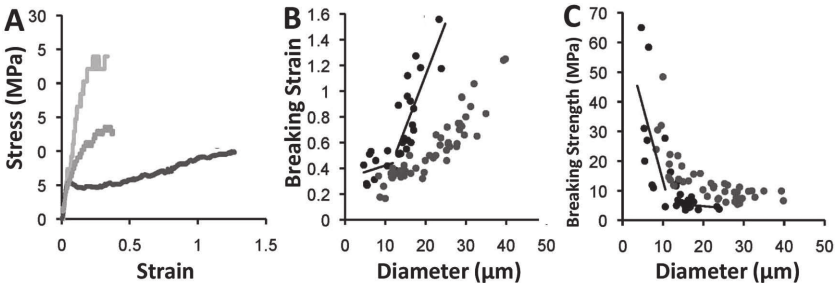

Figure 6. Although fusions do not impact fiber morphology, the mechanical properties of fibers are altered. A) Stress–strain curves for small (<10 μm, light grey), medium (10–15 μm, medium grey), and wide (>15 μm, dark grey) EGFP–Ubx fibers. B,C) Comparing engineering stress and engineering strain for Ubx (black) and EGFP–Ubx (grey) fibers as a function of fiber diameter.

the functional proteins in the resulting materials. This ability allows specific mixtures of functionalities to be incorporated into materials. The potential to functionalize Ubx materials is further enhanced by creating double fusion proteins, in which two functional proteins are present at stoichiometric levels. In summary, the range of functional proteins that can be incorporated into elastomeric protein-based materials is far greater than previously established.

4. Experimental Section

Construction of Plasmids: pET19b–Ubx plasmid was used as the parent vector for all fusion proteins. In this plasmid, the gene encoding Ubx mRNA splicing isoform Ia (GenBank AAN13718.1) was inserted between the NdeI and BamHI sites of pET19b vector (Novagen), thus adding a 10×His-tag and a hydrophilic linker (SSGGHDDDDK) to the amino-terminus of Ubx.[7] DNA sequences encoding all protein fusion partners were inserted into the NdeI site of pET19b–Ubx, between the N-terminal His-tag and Ubx. DNAs encoding the genes for SDF-1a and bFGF were cloned from a cDNA library derived from human cells and provided by Dr. Bayless (Texas A&M Health Science Center). The plasmids expressing the isolated functional proteins were generated from plasmids that encode corresponding Ubx fusions by introducing a stop codon before the *ubx* gene using the following primers: Forward: 5′-CATATGAACTCGTAGTTTGAACAGGCCTCC-3′ and Reverse: 5′-GGAGGCCTGTTCAAACTACGAGTTCATATG-3′. Table 1 lists the sequence of Gly- and Ser-rich linkers which separate the fused protein from Ubx and provide sufficient flexibility for Ubx assembly and appended protein function. A general schematic of Ubx and Ubx fusions is shown in Figure 1A.

Predicting the Charge of Ubx Fusion Proteins: Protein charge was predicted using the online calculator http://protcalc.sourceforge.net from the Scripps Research Institute. The pH of the buffer in the materials assembly trays (pH = 8.0) was used for these calculations.

Protein Expression and Purification: All Ubx-fusion expressing plasmids were transformed individually into Rosetta(DE3)pLysS competent cells. For each growth, a single colony was inoculated into l00 mL of Luria broth plus 100 mg L^{-1} carbenicillin and 34 mg L^{-1} chloroamphenicol and incubated with shaking overnight. The overnight culture (5 mL) was then subcultivated into 1 L Luria broth plus the same antibiotics at 37 °C. Once the absorbance at 600 nm reached 0.6–0.65, 1×10^{-3} M IPTG was added to induce protein expression. For protein expression, all cultures were grown for 16 h at 25 °C. Cells were harvested by centrifugation at 3500 relative centrifugal force (RCF) for 30 min at 4 °C and stored as frozen pellet at –20 °C. For purification, all procedures were performed carefully on ice to prevent protease degradation. In order to compare the final protein yield of all fusions, a more complete lysis protocol was performed. Each aliquot was thawed and lysed in 10 mL of lysis buffer (50×10^{-3} M sodium phosphate buffer, pH 8.0, 5% glucose w/v, 500×10^{-3} M NaCl, 1 protease inhibitor tablet (Roche), 0.8 mg L^{-1} DNase I, 5 mg Lysozyme, and 2×10^{-3} M dithiothreitol (DTT)). Thawed lysate was frozen again at –80 °C. The freeze and thaw procedure was performed twice more, followed by ultrasound sonication (Branson digital sonicator, Model S250) at 25% intensity in 10 s intervals for 5 min. Cell lysates were centrifuged at 37 000 RCF for 30 min. The supernatant was loaded on a gravity column with 3–5 mL nickel-nitrilotriacetic acid (Ni-NTA) agarose resin (Qiagen), which was pre-equilibrated with equilibration buffer (5% glucose w/v, 500×10^{-3} M NaCl, 50×10^{-3} M sodium phosphate buffer, 1×10^{-3} M DTT, pH 8.0). The column was then washed by 10 column volumes of W1 buffer, 10 column volumes of W2 buffer, and 5 column volumes of W3 buffer (W1, W2, W3 buffers are equilibration buffer containing 20×10^{-3}, 40×10^{-3}, and 80×10^{-3} M imidazole, respectively). Protein was eluted with 10 mL of elution buffer (equilibration buffer plus 200×10^{-3} M imidazole). Protein purity was estimated based on gel electrophoresis results, and was similar for purification of Ubx and all Ubx fusion proteins. Concentrations of the purified Ubx samples were determined using the BioRad protein assay (BioRad).

Quantification of Total (Soluble + Insoluble) Protein: For all fusion proteins, the relative levels of expressed protein were measured by western blot. A single colony containing transformed cells was inoculated into a 250 mL flask containing 50 mL Luria broth with 100 mg L^{-1} carbenicillin and 34 mg L^{-1} chloroamphenicol and grown at 37 °C at 250 rpm. When the absorbance at 600 nm reached 0.6 to 0.65, 1×10^{-3} M IPTG was added to induce protein expression and the temperature was lowered to 25 °C for an additional 16 h. The final optical density at 600 nm for fermentations to produce all Ubx fusion proteins ranged from 1.2 to 1.3. An aliquot (0.5 mL) of cell culture was collected from each flask and spun down by centrifugation at 16 000 RCF for 3 min. Pelleted cells were lysed by disolving them in 0.5 mL sample buffer for sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS PAGE) followed by heating at 95 °C for 10 min. Protein sample and cell debris were then separated by centrifugation at 16 000 RCF for 15 min. The protein sample was then resolved using 10% SDS PAGE followed by western blotting (BioRad). Low molecular weight proteins (39.0–61.4 kD) were transferred to nitrocellulose membrane at 50 V, 150 mA for 2 h. High molecular weight proteins (67.2–102.7 kD) were transferred at 50 V, 200 mA for 3 h. A monoclonal mouse anti His-tag antibody (Qiagen, Cat# 34670) was used as primary antibody at a 1:5000 dilution. An antibody conjugated to a near infrared fluorescence dye, IRDye 800CW Goat anti-Mouse IgG (Li-Cor), was used as secondary antibody. Signal generated by the secondary antibody was detected by an Odyssey Imaging System. All fusions were tested 3 times from the plasmid transformation step to final scanning step.

Solubility Assays: To test solubility of all fusions, an ammonium sulfate tolerance assay was performed.[14] Purified Ubx-fusions were concentrated and dialyzed into assay buffer (50×10^{-3} M sodium phosphate buffer, pH 8.0, 5% glucose w/v, 500×10^{-3} M NaCl, 100×10^{-3} M imidazole). All measurements were performed at room temperature (25 °C). The solubility of fusion proteins was measured as the quantity of 32×10^{-6} M fusion protein that remains in solution in 1.1 M ammonium sulfate dissolved in assay buffer. The mixture was allowed to equilibrate for 3 min at room temperature. Samples were then centrifuged for 3 min at 18 000 RCF. The protein concentrations in the supernatants were quantified by BioRad protein assay (BioRad). Assays of all fusions were replicated 3 times, and reported as an average with standard deviation.

Assembly of Ubx Fibers: All assays were performed with the same protein concentration, fixed incubation time, and controlled environmental parameters. Ubx fusion protein (30 nmol) was added into a teflon coated tray filled with Material Forming Buffer (50×10^{-3} M sodium phosphate buffer, pH 8.0, 5% glucose w/v, 500×10^{-3} M NaCl). The tray was then incubated for 16 h at room temperature (approximately 25 °C) and 30%–40% humidity. Fibers were pulled by hand using a 4 mm-diameter inoculating loop from the air–water interface. In prior tests, Ubx fibers mechanically drawn at different rates had similar lengths and mechanical properties. Therefore pulling fibers by hand should not impact fiber length. Only the length of the first fiber drawn from each tray was measured. At least 5 fibers, each pulled from a different tray, were measured for each Ubx fusion, and the results were evaluated by statistical analysis.

Ethanol Washing and Autoclaving of Ubx Materials: mCherry–Ubx fibers were pulled using a 4 mm diameter plastic inoculating loop (VWR International) from teflon tray and dried in air for 30 min as described above. Loops with mCherry fiber were either dipped into 100% ethanol solution for 30 min or autoclaved, the loops were sterilized in steam sterilizer for 20 min (121 °C, 1.27 kg cm^{-1}). mCherry–Ubx fibers were then observed, and the fluorescence intensity was measured by confocal microscopy.

Mixing Ubx and EGFP–Ubx: Protein mixtures to assess the efficiency of fiber assembly were made by first purifying and quantifying pET19b–Ubx and EGFP–Ubx as previously described. The concentrations of each protein were used to create mixtures of 0%, 20%, 40%, 60%, 80%, and 100% EGFP–Ubx by pipetting the appropriate amount of each protein

into 6 micro centrifuge tubes and allowing them to incubate at room temperature for 15 min. Each mixture was added into a separate teflon coated tray filled with material forming and incubated for 20 h at room temperature (approximately 25 °C) and 30%–40% humidity. Fibers ($N = 5$) were pulled by a 4 mm-diameter inoculating loop from the air–water interface, covered, and allowed to dry at room temperature for 30 min. The loops were then placed on a 22 × 55 mm coverslip, and imaged immediately using a Nikon Eclipse Ti A1R inverted confocal microscope equipped with NIS Elements AR 4.10.01 software to analyze fluorescent intensity of both the 4',6-diamidino-2-phenylindole (DAPI) and fluorescein isothiocynate (FITC) channels.

Analysis of EGFP–Ubx Mechanical Properties: All fibers used to measure the mechanical properties of EGFP–Ubx were drawn using the drop method as defined by Greer et al.[7] All experiments were performed as previously described.[8] In brief, a Gatan microtest tensile tester was used with a 2 N load capacity, a sensitivity of 0.0001 N, and a displacement resolution of 0.001 mm. A loading speed of 0.1 mm min^{-1} was used to mitigate any pulling rate effects. Samples were attached using double-sided carbon tape and Loctite 495 adhesive. Initial and final sample diameters were measured by scanning electron microscopy.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

This work was supported by grants from the Robert A. Welch Foundation to J.L. (C-1716) and to K.S.M. (C-576), and the Ted Nash Long Life Foundation (M1500779) and the Texas A&M Health Science Center Research Development and Enhancement Awards Program and NSF CAREER Program (1151394) to S.E.B. The authors would like to thank Marcin Olszewski (Gdańsk University of Technology) for DNA encoding TmaSSB and TneSSB, Aron Fenton (University of Kansas Medical Center) for DNA encoding liver pyruvate kinase, Greg Reinhart (Texas A&M University) for DNA encoding phosphofructokinase, and Kayla Bayless (Texas A&M Health Science Center) for DNA encoding osteopontin and the human umbilical vein endothelial cell cDNA library, and members of the Bondos lab for assistance and helpful discussions.

Received: August 29, 2014
Revised: November 13, 2014
Published online: January 27, 2015

[1] D. N. Woolfson, Z. N. Mahmoud, *Chem. Soc. Rev.* **2010**, *39*, 3464.
[2] U. Hershel, C. Dahmen, H. Kessler, *Biomaterials* **2003**, *24*, 4385.
[3] A. K. Adak, B. Y. Li, L. D. Huang, T. W. Lin, T. C. Chang, K. C. Hwang, C. C. Lin, *ACS Appl. Mater. Interfaces* **2014**, *6*, 10452.
[4] Z. Huang, T. Salim, A. Brawley, J. Patterson, K. S. Matthews, S. E. Bondos, J. Lou, *Adv. Funct. Mater.* **2011**, *21*, 2633.
[5] R. Jansson, N. Thatikonda, D. Lindberg, A. Rising, J. Johansson, P. A. Nygren, M. Hedhammar, *Biomacromolecules* **2014**, *15*, 1696.
[6] a) J. Huang, C. Wong, A. Geroge, D. L. Kaplan, *Biomaterials* **2007**, *28*, 2358; b) M. Kumar, K. J. Sanford, W. A. Cuevas, M. Du, K. D. Collier, N. Chow, *Biomacromolecules* **2006**, *7*, 2543; c) M. Nagaoka, H. L. Jiang, T. Hoshiba, T. Akaike, C. S. Cho, *Ann. Biomed. Eng.* **2010**, *38*, 683; d) M. Sackewitz, S. von Einem, G. Hause, M. Wunderlich, F. X. Schmid, E. Schwartz, *Protein Sci.* **2008**, *17*, 1044.
[7] A. M. Greer, Z. Huang, A. Oriakhi, Y. Lu, J. Lou, K. S. Matthews, S. E. Bondos, *Biomacromolecules* **2009**, *10*, 829.
[8] Z. Huang, Y. Lu, R. Majithia, J. Shah, K. Meissner, K. S. Matthews, S. E. Bondos, J. Lou, *Biomacromolecules* **2010**, *11*, 3644.
[9] a) J. L. Patterson, C. A. Abbey, K. J. Bayless, S. E. Bondos, *J. Biomed. Mater. Res. A* **2014**, *102*, 97; b) J. Patterson, A. Arenas-Gamboa, T. Y. Wang, H. C. Hsiao, J. P. Pellois, A. Rice-Ficht, S. E. Bondos, *J. Biomed. Mater. Res. A.* DOI: 10.1002/jbm.a.35295.
[10] T. Christensen, M. Amiram, S. Dagher, K. Trabbic-Carlson, M. Shamji, L. A. Setton, A. Chikloti, *Protein Sci.* **2009**, *18*, 1377.
[11] K. Chatterjee, S. Lin-Gibson, W. E. Wallace, S. H. Parekh, Y. J. Lee, M. T. Cicerone, M. F. Young, C. G. Simon Jr., *Biomaterials* **2010**, *31*, 5051.
[12] a) C. Cummings, H. Murata, R. Koepsel, A. J. Russell, *Biomaterials* **2013**, *34*, 7437; b) J. H. Forstater, A. Kleinhammes, Y. Wu, *Langmuir* **2013**, *29*, 15013; c) J. Ma, L. Zhang, Z. Liang, W. Zhang, Y. Zhang, *J. Sep. Sci.* **2007**, *30*, 3050.
[13] a) S. E. Bondos, *Curr. Anal. Chem.* **2006**, *2*, 157; b) G. D. Davis, C. Elisee, D. M. Newham, R. G. Harrison, *Biotechnol. Bioeng.* **2000**, *65*, 382; c) P. Forrer, R. Jaussi, *Gene* **1998**, *224*, 45; d) Y. B. Zhang, J. Howitt, S. McCorkle, P. Lawrence, K. Springer, P. Freimuth, *Protein Expr. Purif.* **2004**, *36*, 207.
[14] a) S. Trevino, J. M. Schotlz, C. N. Pace, *J. Mol. Biol.* **2007**, *16*, 449; b) S. R. Trevino, J. M. Scholtz, C. N. Pace, *J. Pharm. Sci.* **2008**, *97*, 4155.
[15] E. Scherzinger, R. Lurz, M. Turmaine, L. Mangiarini, B. Hollenbach, R. Hasenbank, G. P. Bates, S. W. Davies, H. Lehrach, E. E. Wanker, *Cell* **1997**, *90*, 549.
[16] R. Majithia, J. Patterson, S. E. Bondos, K. E. Meissner, *Biomacromolecules* **2011**, *12*, 3629.