ACCOUNTS
of chemical research

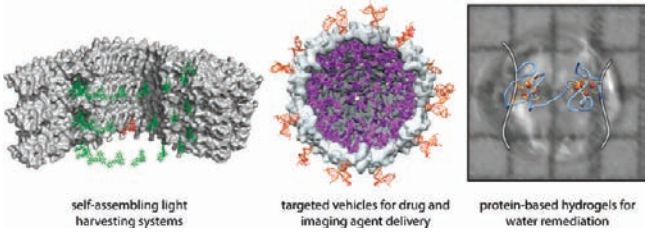
Using Synthetically Modified Proteins to Make New Materials

LEAH S. WITUS AND MATTHEW B. FRANCIS*

Department of Chemistry, University of California, Berkeley, California 94720-1460, United States, and Materials Sciences Division, Lawrence Berkeley National Laboratories, Berkeley, California 94720-1460, United States

RECEIVED ON MAY 6, 2011

CONSPECTUS



The uniquely diverse structures and functions of proteins offer many exciting opportunities for creating new materials with advanced properties. Exploiting these capabilities requires a set of versatile chemical reactions that can attach nonnatural groups to specific locations on protein surfaces. Over the years, we and others have developed a series of new techniques for protein bioconjugation, with a particular emphasis on achieving high site selectivity and yield. Using these reactions, we have been able to prepare a number of new materials with functions that depend on both the natural and the synthetic components. In this Account, we discuss our progress in protein bioconjugation over the past decade, focusing on three distinct projects.

We first consider our work to harness sunlight artificially by mimicking features of the photosynthetic apparatus, with its beautifully integrated system of chromophores, electron transfer groups, and catalytic centers. Central to these photosystems are light-harvesting antennae having hundreds of precisely aligned chromophores with positions that are dictated by the proteins within the arrays. Our approach to generating similar arrangements involves the self-assembly of tobacco mosaic virus coat proteins bearing synthetic chromophore groups. These systems offer efficient light collection, are easy to prepare, and can be used to build complex photocatalytic systems through the modification of multiple sites on the protein surfaces.

We then discuss protein-based carriers that can deliver drugs and imaging agents to diseased tissues. The nanoscale agents we have built for this purpose are based on the hollow protein shell of bacteriophage MS2. These 27 nm capsids have 32 pores, which allow the entry of relatively large organic molecules into the protein shell without requiring disassembly. Our group has developed a series of chemical strategies that can install dyes, radiolabels, MRI contrast agents, and anticancer drugs on the inside surface of these capsids. We have also developed methods to decorate the external surfaces with binders for specific proteins on cancer cells.

As a third research area, our group has developed protein–polymer hybrid materials for water remediation. To reduce the toxicity of heavy metals in living cells, Nature has evolved metallothioneins, which are sulfur-rich polypeptides that bind mercury, cadmium, and other toxic ions at sub-parts-per-billion concentrations. Unfortunately, these proteins are very difficult to incorporate into polymers, largely because typical protein modification reactions target the very cysteine, lysine, and carboxylate-containing residues that are required for their proper function. To address this challenge, we developed a new way to attach these (and many other) proteins to polymer chains by expressing them as part of an N- and C-terminal modification "cassette". The resulting materials retain their selectivity and can remove trace amounts of toxic metal ions from ocean water.

Each of these examples has presented a new set of protein bioconjugation challenges that have been met through the development of new reaction methodology. Future progress in the generation of protein-based materials will require scalable synthetic techniques with improved yields and selectivities, inexpensive purification methods for bioconjugates, and theoretical and dynamical treatments for designing new materials through protein self-assembly.

Published on the Web 08/03/2011 www.pubs.acs.org/accounts
10.1021/ar2001292

Introduction

Proteins exhibit unrivaled diversity in terms of their structure and function. Using only a limited set of side chains and auxiliary groups, they have evolved unparalleled abilities to accelerate chemical transformations, facilitate the delivery of genetic cargo to targeted cells, bind specific analytes in complex mixtures, transduce energy, and generate elaborate three-dimensional structures through self-assembly. It would be highly desirable to import these capabilities into the context of new materials, and as a result we and other laboratories have worked to augment the native functions of proteins through the covalent attachment of synthetic groups, polymer chains, inorganic particles, and surfaces.

Initially, our primary goal was to use self-assembling proteins (such as actin and the protein capsids of bacteriophage MS2 and the tobacco mosaic virus) as rigid scaffolds for the creation of well-defined nanoscale materials. To do so required the attachment of synthetic functional groups to specific sites on the protein surfaces either before or after the self-assembly process. We quickly found, however, that the available bioconjugation reactions were insufficient to access many of our targets. In particular, many of our designs required two or more different chemical groups to be installed in distinct locations on the same protein. At the time there were very few combinations of chemically orthogonal modification techniques that could be used to generate bioconjugates with this level of complexity.

As a result, our pursuit of protein based materials has been accompanied by significant efforts to expand the options for bioconjugate preparation.[1,2] To do so, we have reported a series of new reactions that can modify underutilized amino acids, such as tyrosines and tryptophans, and extend the pool of organic substrates that can be considered. We have also approached site-specific modification through new reactions that target unique positions in the protein sequence (such as the N-terminus) and unnatural amino acids that can be introduced genetically. The development of all of these reactions required that we learn to optimize chemical reactivity in aqueous solution at near-neutral pH, ambient temperatures, and low micromolar substrate concentrations, all without the use of protecting groups.

By applying various combinations of these strategies, we have been able to generate a number of well-defined protein bioconjugates for use in materials applications. Each has presented a unique set of chemical and biological challenges, which have in turn guided our efforts in reaction discovery. This Account describes three examples of the protein-based materials we have prepared, with a particular emphasis on the chemical strategies that have been developed to access them.

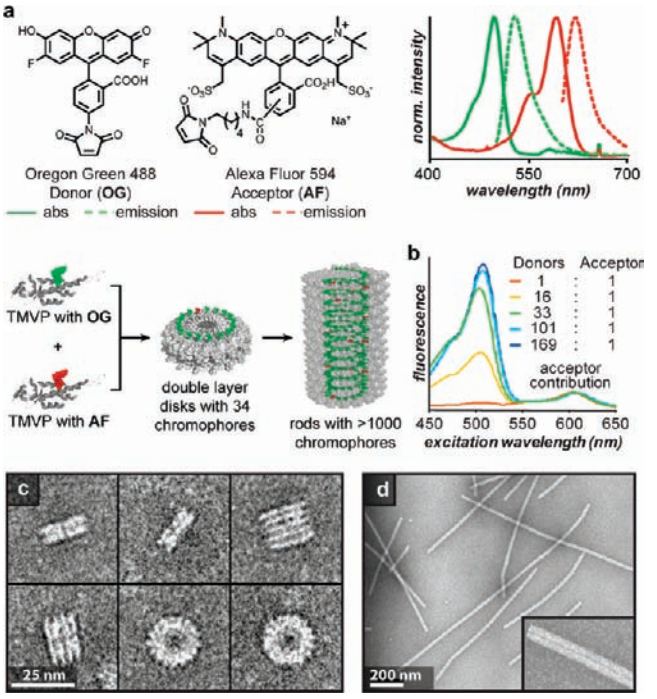


FIGURE 1. Self-assembling light harvesting systems based on the coat protein of the tobacco mosaic virus (TMVP). (a) S123C TMVP monomers covalently labeled with either OG or AF were combined and self-assembled under disk and rod-forming conditions. (b) Excitation spectra for rod assemblies indicated light harvesting behavior through the increase in sensitivity at 500 nm with increasing donor fraction. TEM images of chromophore-labeled TMVP disks (c) and rods (d) are shown. Reprinted in part with permission from ref 6. Copyright 2007 American Chemical Society.

Synthetic Light Harvesting Systems Based on the Tobacco Mosaic Virus

Photosynthetic organisms collect and store the energy of sunlight through a beautifully integrated system of chromophores, electron transfer groups, and catalytic centers. Key to the function of these photosystems are complex membrane-bound light harvesting antennae that provide essential increases in the number of photons that can be collected.[3a] These systems also allow the full solar emission spectrum to be utilized, rather than just the portion that chlorophyll absorbs directly. The hallmark of light harvesting systems is the presence of hundreds of precisely aligned chromophores with positions that are dictated by the proteins within the arrays.[3b] To mimic these systems for applications in optical sensing and solar cell light collection, many research groups have devised elegant strategies to create similarly integrated arrangements of synthetic chromophores and

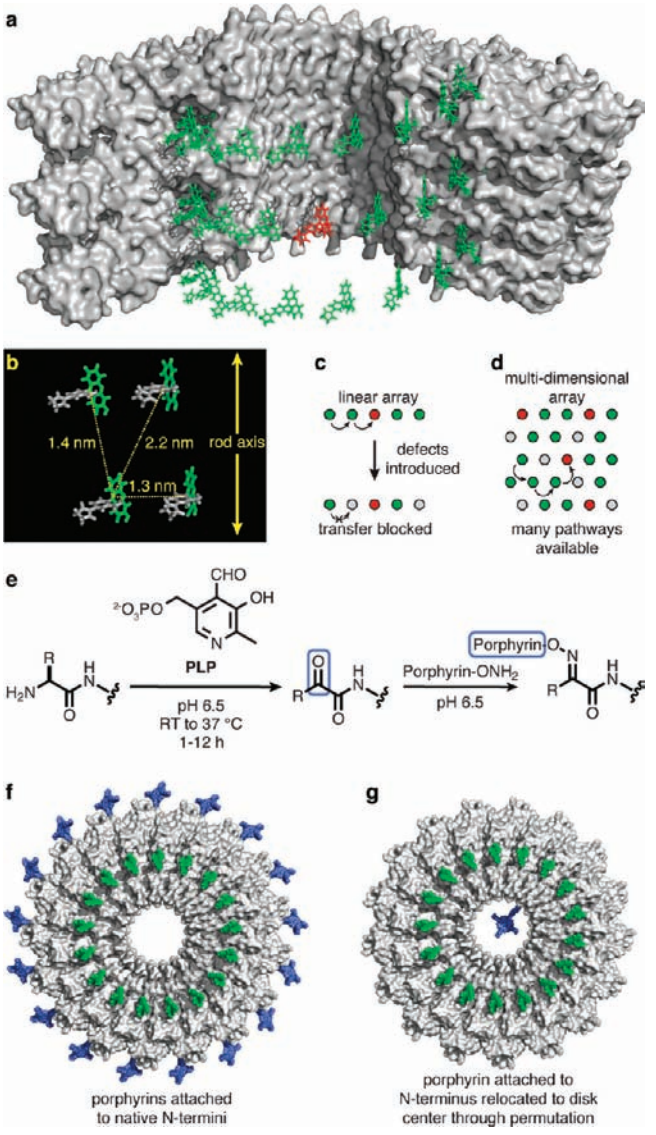


FIGURE 2. Energy transfer in multidimensional light harvesting arrays. (a) A model of the chromophore distribution in TMVP is provided. The calculated positions and orientations of the chromophores are shown in (b), indicating that the transition dipoles (green) align along the long axis of the rod. (c) In a 1D array or a ring, defects can block transfer pathways. (d) In multidimensional systems, redundant energy transfer pathways can circumvent these sites. (e) Electron transfer groups (blue) can be attached to the N-termini through transamination followed by oxime formation.[11,12] These positions are shown for the native sequence in (f) and a cyclic permuatant[14] in (g). Reprinted in part with permission from ref 7. Copyright 2010 American Chemical Society.

electron transfer groups.[4] It has remained difficult, however, to establish and maintain the long-range order that is observed in the natural systems.

Our approach for developing synthetic light harvesting systems has involved the positioning of synthetic chromophores into rigid periodic arrays through the self-assembly of the tobacco mosaic virus coat protein (TMVP) (Figure 1). This protein was already known to assemble into well-defined disk and rod aggregates under differing pH and ionic strength conditions (see Figure 1c and d).[5] Both structure types have a diameter of 18 nm and an interior pore diameter of 2 nm. In order to install commercially available dye molecules, we introduced a highly reactive cysteine residue in position 123 to allow modification with maleimides.[6] Despite the large size of the chromophores, the labeled TMVP monomers retained their ability to self-assemble into both disk and rod structures, positioning the dye molecules with 1.4 nm spacings. In the rod assemblies, the chromophores are believed to occupy the spiral channel normally filled with the genomic RNA strand, as shown in Figure 2a. These rods can reach many micrometers in length, generating light collection systems containing many thousands of individual chromophores. Particular advantages of this protein-based approach include the high synthetic efficiency of the self-assembly process, the regularity of the chromophore attachment sites, and the potential to alter the distances between the groups by changing the individual bioconjugation locations. These latter two points are difficult to achieve using noncovalent chromophore assembly methods.

Using suitable donor and acceptor dyes, we were able to assemble both disk and rod structures capable of collecting a broad range of wavelengths within the solar spectrum.[6] The overlap between the emission and absorption bands of these dyes allowed Förster resonance energy transfer (FRET) to occur with efficiencies ranging from 33 to 55% (Figure 1a,b). We were also able to vary the ratio of the labeled monomers, resulting in rod assemblies with different compositions of donor and acceptor chromophores. Greater overall performance was obtained by constructing three-chromophore systems, which could achieve over 90% transfer efficiency across a broad wavelength range.

An interesting finding in these studies was that the rods could achieve more efficient energy transfer than disks with the same chromophore composition. Our current explanation for this behavior is that the rods orient the principal transition dipoles of the chromophores along the long axis (see Figure 2b) to increase the FRET efficiency in a way that the disks cannot access. Subsequent experimental and theoretical studies in collaboration with Prof. Phillip Geissler have indicated that the rods can also transfer energy through multiple redundant pathways, giving them a greater ability to circumvent assembly defects or photobleached sites (Figure 2c,d).[7]

These studies have yielded light harvesting systems that can be tuned to match any input wavelength through the selection of the appropriate dye molecules. To make use of the collected energy, however, additional modifications are

needed to allow the attachment of photocatalytic and electron transfer groups. We chose to introduce these groups either in the inner pore or on the external surface of the TMVP structures to minimize their interference with the self-assembly process. As one option, we had shown in previous work that tyrosine 149 could be modified to display new functionality through a diazonium coupling reaction.[8] This resulted in the introduction of thousands of uniquely reactive ketone groups on the exterior surface of the rods. Through subsequent oxime bond formation with alkoxyamines, we have shown that polymer chains[8,9] and carbon nanotubes[10] can be introduced in these locations. Unfortunately, this modification technique introduces an azo bond that could quench fluorescence, and the tyrosine selectivity could only be achieved for RNA-containing TMV rods obtained from plants.

As a more versatile alternative strategy, we used a reaction developed in our lab for the transamination[11] of the N-terminus, resulting in the introduction of a ketone or aldehyde group for further elaboration as described above. Our lab had previously shown that pyridoxal 5′-phosphate (PLP) can achieve the conversion of N-terminal amino groups to carbonyls (Figure 2e).[11] PLP-mediated transamination is site specific for the N-terminus of a protein, proceeds under mild reaction conditions, is compatible with the cysteine residues that are required for chromophore attachment, and is readily scalable. Research in our group has shown that this reaction is applicable to many N-terminal residues but leads to highest conversion when the N-terminal side chains are Ala, Gly, Asp, Glu, and Asn.[12] In more recent work, we have enhanced the reliability of this reaction by identifying a highly reactive short sequence, Ala-Lys-Thr, which results in particularly high yields (up to 90% conversion for GFP, compared to negligible wild-type reactivity under the same conditions) when installed at the N-terminus of a protein.[13] Following these reactivity guidelines, the wild-type sequence of TMVP was extended at the N-terminus through the addition of an Ala-Gly group. This allowed the installation of dyes and polymers on the disk peripheries with much higher levels of conversion than could be obtained for the wild-type sequence (Figure 2f).[12] To relocate the N-terminal modification sites to the center of the disks, we have also reported a circular permutant of TMVP that can still assemble into disk and short rod structures.[14] This strategy allowed access to concentric arrangements more reminiscent of natural light harvesting complexes (Figure 2g).

Moving forward, our current efforts involve the use of these compatible modification strategies to attach porphyrins, phthalocyanines, and other electron transfer groups to different sites on the TMVP assemblies in order to convert the collected light into useful chemical and electrical energy. Parallel studies by our lab have already shown that a spherical MS2 viral capsid (see below) can be used to integrate these types of light collecting and electron transfer groups.[15] We are also introducing new modification sites to allow the attachment of TMVP assemblies to electroactive surfaces for photocurrent collection. In collaboration with other groups at UC Berkeley, we are currently characterizing the systems using time-resolved spectroscopy.[16]

Targeted Imaging and Drug Delivery Agents Based on Bacteriophage MS2

Tissue-specific delivery strategies offer many promising avenues for the detection and treatment of disease. The overall approach involves coupling a targeting group, which specifically binds to receptors on the surface of certain cell types, to an imaging or therapeutic agent. Using a nanoscale carrier to link these components allows multiple targeting groups to be attached (potentially increasing binding avidity through multivalency effects) and offers the ability to deliver multiple copies of an imaging or drug cargo with each binding event. Hollow or permeable carriers with multiple attachment sites in both interior and exterior positions are thus ideal in many respects, and additional desirable properties include monodispersity and biodegradability. While many different scaffolds are being explored for delivery purposes,[17] we have found viral coat proteins to be particularly amenable to meeting these design criteria.[18]

Our research in this area has centered around the chemical modification of the bacteriophage MS2 viral capsid. The protein shell of this virus consists of 180 sequence-identical subunits that can be expressed recombinantly in *Escherichia coli* hosts. The 27 nm capsids are isolated in fully assembled form and lack the native genomic RNA strand. An unusual feature of this capsid, and the reason we originally selected it, is the presence of 32 pores (each 1.8 nm in diameter) that allow the access of relatively large organic molecules to the interior of the protein shell without requiring disassembly. Protein modification chemistry can then be used to attach molecules to the inside surface. In the case of drug molecules, the attachment linkages can be engineered to release the cargo upon endocytosis. Through the use of a second chemical strategy, we envisioned the installation of tissue-specific targeting groups on the external surface. An unforeseen practical advantage of MS2 capsids has been their stability from pH 3 to 10, up to 60 °C,[19] and in the presence of modest amounts of organic cosolvents.

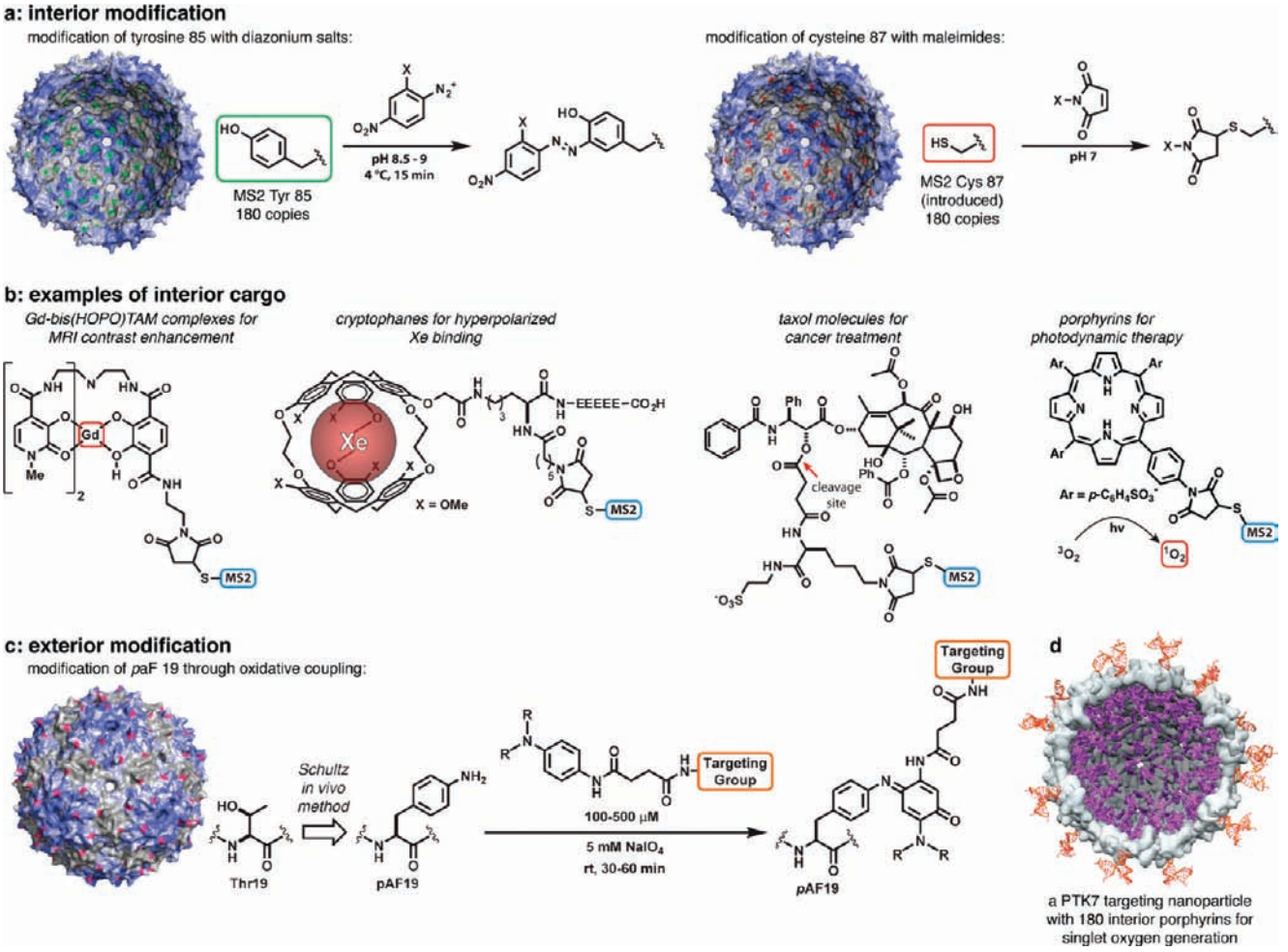


FIGURE 3. Synthesis of targeted delivery agents from bacteriophage MS2 capsids. (a) The interior surface can be modified by targeting a native tyrosine[19] or an introduced cysteine residue.[20] (b) Several types of cargo molecules (90−180 copies of each) have been attached to these sites.[19−24] (c) Targeting groups can be installed on the external surface by modifying an artificial amino acid using a new oxidative coupling reaction.[27,31] These strategies can be combined to yield targeted particles for therapeutic applications, such as the structure in (d), which combines protein tyrosine kinase 7 (PTK7) binding aptamers with porphyrins for singlet oxygen generation.[33]

We have developed two different strategies for the attachment of synthetic cargo to the interior of the capsids. For the first, we have found that electron-deficient diazonium salts react selectively with tyrosine 85, 180 copies of which are displayed on the internal surface of the capsids (Figure 3a).[19] By adding aldehydes and other substituents to the diazonium salts, this strategy has been used to install chromophores and other groups for use in imaging applications (see below). As an alternative method, we have also installed a uniquely reactive cysteine residue in position 87 to provide 180 sulfahydryl groups for modification using maleimides (Figure 3a).[20]

The ability of both of these strategies to introduce 100 or more copies of a given functional group on the inside each capsid provides an important increase in the specific activity of the carriers. As one example of this concept, we have collaborated with Prof. Ken Raymond to attach 90 copies of a Gd-bis(HOPO)TAM complex to the inside of each MS2 capsid (Figure 3b).[21] The assemblies remained fully water-soluble and could achieve very high relaxivities per particle (as compared to currently used small molecule contrast agents) due to the large number of complexes and the enhancement in relaxivity values afforded by the slow rotation of the 2.5 MDa protein shells. As another example, we collaborated with Profs. Alex Pines and David Wemmer to install 120 copies of hydrophobic cryptophane cages inside the MS2 capsids by targeting the cysteine residues.[22] This gave the protein shells the ability to bind hyperpolarized xenon atoms in solution, providing a distinct chemical shift that could be detected at sub-pM concentrations using ^{129}Xe NMR. As a third imaging example, the diazonium coupling method was used to install functionality for the

installation of ^{18}F-labels inside the capsid shells, in collaboration with Dr. Jim O'Neil and Dr. Scott Taylor at LBNL.[23] Subsequent live animal PET experiments showed that the circulation and excretion of the capsids could be followed in vivo, providing the basis for future, targeted imaging applications.

To explore the ability of MS2 to house chemotherapuetic agents, we have also attached taxol to the interior of the capsid using the cysteine modification strategy (Figure 3b).[24] Fully water-soluble capsids were obtained with 120 taxol molecules inside each, as determined using RP-HPLC and mass spectrometry analysis. Even in the absence of targeting groups, these capsids were found to be effective for the treatment of MDA-MB-231 cancer cells after endocytosis and linker cleavage.

With the interior modification strategies in place, we set out to develop a flexible way to attach multiple copies of receptor-binding molecules on the exterior surface. Although this could be achieved to a limited extent by inserting peptide-coding sequences into the capsid protein gene before expression,[25] we chose not to pursue this route because these sequence additions can change the capsid self-assembly behavior in unpredictable ways. We also wanted to take advantage of the full range of cancer targeting molecules that were becoming available in the literature, including antibody fragments, designed proteins, nucleic acid aptamers, and peptoids, none of which can be introduced into the monomers on the genetic level.

Identifying a compatible reaction for the attachment of targeting moieties to the exterior proved to be very challenging, and ultimately required the development of a new bioorthogonal "click"-type bioconjugation reaction.[26] We had previously found that anilines reacted rapidly and cleanly with *N*-acyl phenylene diamines in the presence of sodium periodate.[27] The resulting "Bandrowski bases"[28] were stable with respect to hydrolysis across a wide pH range and could withstand strong reductants, oxidants, nucleophiles, and high temperatures. Although the original reaction yielded a trimeric species, we found that the product symmetry could be broken by adding two alkyl groups to the phenylene diamine group.[27] This prevented it from reacting with itself, but still allowed an additional aniline in solution to couple in an "A+B" fashion with very high efficiency, Figure 3c. As a key feature for bioconjugation, this oxidative coupling strategy could be used in the presence of the natural amino acids and nucleic acids without interference. Most importantly, it took place in aqueous solution at low micromolar concentrations, thus offering great potential for the attachment of complex targeting groups to our MS2 capsids.

To use this coupling reaction, we required the site-specific introduction of the aniline group into the capsid proteins. This was achieved using the powerful method of amber codon suppression developed by the Schultz lab.[29] Specifically, this technique was used to introduce a *p*-aminophenylalanine (*p*aF) residue[30] at position 19 in each capsid monomer. This provided 180 copies of the aniline group that were evenly distributed on the exterior surface of the assemblies.[31] In our original study of this approach, we found that the oxidative coupling method could be used to attach up to 135 copies of a breast cancer targeting peptide to MS2 capsids in under 1 h (Figure 3c).

We have also used this artificial amino acid modification strategy to install DNA aptamers reported by the Tan lab[32] to bind to tyrosine kinase 7 (PTK7), a receptor found on Jurkat T-lymphocytes and many other cancer cell lines. This 40-base sequence was obtained with an amino group at the 5′-end and coupled to an *N*,*N*-diethylphenylene diamine for protein attachment.[20] In the presence of sodium periodate, 20−60 copies of the DNA strands could be coupled to each capsid surface. The interior surface of these capsids had previously been modified with a fluorophore by targeting the cysteine residues, providing our first prototype of cancer targeting imaging agent. Flow cytometry data confirmed that the aptamer-modified capsids bound to Jurkat cells at concentrations as low as 11 nM, while capsids bearing a randomized DNA sequence did not.

Following the success of these experiments, we next synthesized a system for targeted therapy that included a cytotoxic group in addition to the cell-binding moieties. In this case, the capsids displayed ∼20 copies of the Jurkat-specific DNA aptamer outside and housed 180 photodynamic therapy agents introduced as maleimide-functionalized porphyrins (Figure 3d).[33] Upon irradiation with 420 nm light, these capsids were able to produce 300 000 singlet oxygen molecules in 20 min with minimal disassembly. Using the targeting capabilities of the DNA aptamer, we were able to show the selective killing of a PTK7-positive cell line in the presence of other cells.

As a result of these studies, the chemistry that has been developed for the dual-surface modification of MS2 capsids should now be sufficiently flexible to wed the majority of available targeting groups with large numbers of a desired drug or imaging payload. In current experiments, we are adapting this system to develop optical, PET, and ^{129}Xe-MRI based agents for the sensitive and specific detection of breast cancer markers in vivo.

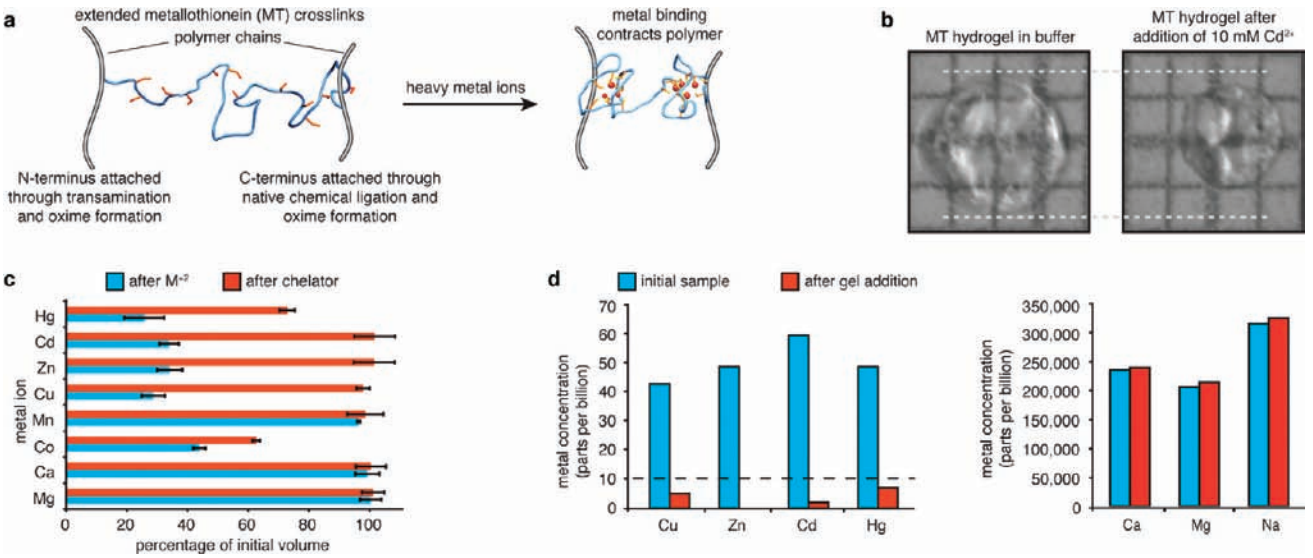


FIGURE 4. Modification of metallothioneins (MTs) for the creation of metal-binding hydrogels. These small proteins have 12 to 20 cysteine residues (orange in (a)). A dual modification strategy was used to cross-link poly(2-hydroxypropyl methacrylamide) (HPMA) polymers with the MT sequences. (b) The initial gels (left) shrink considerably in the presence of contaminant ions (right). (c) The gels respond to a number of toxic ions. The blue bars indicate the percentage of initial gel volume after the binding and contraction have occurred. In most cases, the gels can be recycled using inexpensive chelators, such as EDTA (red bars). (d) Water from a saline marsh was spiked with four contaminant ions at ~50 ppb each. Upon addition of the MT hydrogel, the concentrations of all four toxic species were simultaneously reduced to below 10 ppb (left graph). In contrast, the concentrations of much more abundant Na^+, K^+, Ca^{2+}, and Mg^{2+} ions were unchanged (right graph). Reprinted in part with permission from ref 38. Copyright 2008 American Chemical Society.

Protein–Polymer Hybrid Materials for Water Remediation

In addition to using proteins as new structural components, our group is also interested in harnessing their unique *functions* in a materials context. Proteins have evolved to perform the majority of the tasks in living cells, providing numerous useful sequences with capabilities that are unmatched by fully synthetic materials. However, proteins usually suffer from poor handling and processing characteristics and are difficult to immobilize for the purposes of removal and recycling. By attaching proteins with desirable binding, sensing, or enzymatic capabilities to inexpensive polymer matrices, many groups have prepared new materials with advanced functions.[34] Inspired by these examples, we have applied our experience in bioconjugation chemistry to develop highly general and scalable methods that can couple proteins to synthetic materials through well-defined linkages.

One example of a challenge that can be addressed using protein–polymer hybrid materials is water remediation. Binding selectivity is of paramount importance for these applications because heavy metals, such as cadmium, mercury, and arsenic, are toxic in low ppb quantities. However, drinking water supplies and industrial waste streams also contain relatively high concentrations of innocuous metal ions, such as sodium, potassum, calcium, and magnesium that can compete for binding sites on a given treatment agent.

The challenge of selective heavy metal ion sequestration has already been addressed by Nature. Metallothioneins (MTs) are short protein sequences (~70–100 amino acids) that contain up to 20 cysteine residues (Figure 4a).[35] These polypeptides are produced by organisms from every biological kingdom to capture heavy metal ions and thus reduce their toxicity. While one would expect them to exhibit some selectivity for heavy metal ions due to their thiol-rich nature alone, the intervening portions of the sequences serve to orient the cysteines into well-defined binding pockets that can capture multiple metal ions with picomolar binding constants. Perhaps the most impressive (and useful) aspect of these proteins is their ability to bind sub-ppb concentrations of toxic metal atoms while in the presence of very high concentrations of harmless ions.

Despite their potential for use in synthetic remediation materials, these proteins are very difficult to incorporate into polymers and other processable supports. This is largely because typical protein modification reactions target the very cysteine, lysine, and carboxylate-containing residues that are required for their proper function. To address this challenge, we developed a new way to attach these (and

many other) proteins to polymer chains by expressing them as part of an N- and C-terminal modification "cassette".[36] This strategy involves the use of an intein domain that is fused to the C-terminus of a protein sequence to allow for a native chemical ligation[37] with a ketone labeled cysteine residue. The N-terminal ketone is next installed using the biomimetic transamination reaction described above.[11] The two ketone groups then serve as chemospecific attachment points for polymers and other materials that possess pendant alkoxyamine groups.

As an example of this strategy, we have reported the attachment of doubly labeled MTs to alkoxyamine groups extending from *N*-hydroxypropyl methacrylamide (HPMA) polymers.[38] Because they had been labeled at both termini, the MTs served to cross-link the polymer strands, resulting in the formation of hydrogels with swelling states that were intimately correlated to the folded structure of the proteins (Figure 4a,b). As a result of this configuration, the binding of metal ions could be determined visually by a contraction of the hydrogel as the binding pockets were formed. Much like freely soluble MTs, the gels responded to the presence of many toxic metal ions, including Hg^{2+}, Cd^{2+}, Zn^{2+}, Cu^{2+}, and Co^{2+} (Figure 4c). If desired, the MT gels can be recovered and regenerated through the addition of an EDTA solution, which produces small volumes of highly concentrated metal ions that are easy to dispose in a safe and inexpensive manner. Further experiments showed that these gels were capable of removing ppb quantities of heavy metals from authentic samples taken from a saline marsh (Figure 4d).

In current efforts, we are coupling the functionalized metallothioneins to low cost polymeric materials that are already used for water purification, and we are applying this approach to the development of new hybrid materials that are capable of binding to organic pollutants and actinides. We are also expanding the scope of this readily generalizable synthetic strategy for the construction of a large number of new protein-based materials[39] with properties that would be very difficult to achieve using more conventional components.

Conclusion and Outlook

Considering their uniquely broad range of functions and properties in the natural world, proteins clearly have much to offer in terms of creating new materials. Our experiences in this area over the past decade have bolstered our confidence in the power of combining biomolecular structures with non-natural components. It is also abundantly clear that new reaction development must accompany these efforts, as each new reactive method increases the range of building blocks that can be incorporated in addition to solving the challenges at hand. In particular, our work in this area has brought to light a number of future challenges that are serving as our current research emphases.

First, even the most thoroughly tested bioconjugation reactions occasionally fail or produce unexpected results. This stems from our poor understanding of the ways in which the complex and dynamic environments presented by protein surfaces influence chemical reactivity. This is especially true when multiple modifications are sought, as there are still few examples of using tandem reaction strategies to modify a given protein target in more than one location. This is a critically important capability for the creation of complex materials that achieve emergent behavior through the interaction of multiple components.

Despite one's best efforts, however, protein modification reactions often do not reach full conversion or suffer from erosions in site selectivity. Unfortunately, there are no general methods that can be used to isolate modified proteins from unmodified species or separate regioisomers. The development of inexpensive and scalable chromatography methods that can achieve this would therefore be of tremendous value to the field.

As another consideration, very few protein modification reactions have been demonstrated on a large scale. Rather, many require hundreds or even thousands of equivalents of a given reactant. This adds significant cost to the reactions and requires subsequent purification steps that are often difficult to scale. As protein-based materials move toward widespread use, the yield, expense, and purification requirements of each method will weigh heavily on its level of use. Protein production can already be achieved on a significant scale through bacterial fermentation and plant-based expression, and further increases are anticipated as protein-based therapeutics are pursued by the pharmaceutical industry. These opportunities will require protein bioconjugation researchers to address issues of scale and efficiency in ways that are seldom done today.

From the protein context, we are only beginning to understand the subtle factors that influence protein self-assembly. As a result, the creation of new three-dimensional structures with built-in asymmetry will require many new design principles that are not yet in place. Theoretical treatments and dynamical simulations are likely to play a key role in the development of these concepts, ushering the

development of complex multicomponent materials that currently do not exist outside of the natural context.

First and foremost, M.B.F would like to thank the inspiring, creative, and adventurous group of students and post docs he has worked with over the last 10 years. Our studies regarding the development of protein modification reactions have been supported by the NIH (GM 072700). The DOD (BCRP Grant BC016995) has funded the development of MS2-based breast cancer imaging agents, and the NSF (0449772) has supported our work on metallothionein−polymer conjugates. The TMV-based light harvesting systems were supported by the Director, Office of Science, Materials Sciences and Engineering Division, of the DOE (DE-AC02-05CH11231). The UC Berkeley Chemical Biology Graduate Program (Training Grant 1 T32 GMO66698) has provided student support for many of these projects.

BIOGRAPHICAL INFORMATION

Leah S. Witus grew up in Michigan. She obtained her B.S. in Chemistry in 2007 from Rice University. At Rice she worked with Prof. Jeffrey Hartgerink on the noncovalent functionalization of carbon nanotubes using peptides. She is currently pursuing her Ph.D. in chemistry at U.C. Berkeley in the laboratory of Prof. Matthew Francis. Her research involves the optimization of site-specific protein bioconjugation reactions using combinatorial peptide libraries.

Matthew B. Francis received his undergraduate degree in Chemistry from Miami University in Oxford, OH in 1994. From 1994 to 1999 he attended graduate school at Harvard University, working with Prof. Eric Jacobsen. His Ph.D. research involved the development of combinatorial strategies for the discovery and optimization of transition metal catalysts. He then moved to UC Berkeley, where he was a Postdoctoral Fellow working with Prof. Jean Fréchet. Matt started his independent career in the UC Berkeley Chemistry Department in 2001, and he is also a faculty member at the Lawrence Berkeley National Laboratory.

FOOTNOTES

*To whom correspondence should be addressed. E-mail: mfrancis@berkeley.edu.

REFERENCES

1 Hermanson, G. T. *Bioconjugate Techniques*, 2nd ed.; Academic Press: San Diego, 2008.

2 Tilley, S. D.; Joshi, N. S.; Francis, M. B. The Chemistry And Chemical Reactivity of Proteins. In *The Wiley Encyclopedia of Chemical Biology*; Begley, T., Ed.; Wiley-VCH: Weinheim, 2008.

3 (a) Freer, A.; Prince, S.; Sauer, K.; Papiz, M.; Hawthornthwaite-Lawless, A.; McDermott, G.; Cogdell, R.; Isaacs, N. W. Pigment-Pigment Interactions and Energy Transfer in the Antenna Complex of the Photosynthetic Bacterium Rhodopseudomonas acidophila. *Structure* **1996**, *4*, 449–462. (b) Jordan, P.; Fromme, p.; Witt, H. T.; Klukas, O.; Saenger, W.; Kraus, N. Three-dimensional structure of cyanobacterial photosystem I at 2.5 Å resolution. *Nature* **2001**, *411*, 909–917.

4 Balzani, V; Credi, A; Venturi, M. Photochemical Conversion of Solar Energy. *ChemSusChem.* **2008**, *1*, 26–58.

5 Klug, A. The Tobacco Mosaic Virus Particle: Structure and Assembly. *Philos. Trans. R. Soc. London, Ser. B* **1999**, *354*, 531–535.

6 Miller, R. A.; Presley, A. D.; Francis, M. B. Self-Assembling Light Harvesting Systems From Tobacco Mosaic Virus Capsid Proteins. *J. Am. Chem. Soc.* **2007**, *129*, 3104–3109.

7 Miller, R. A.; Stephanopoulos, N.; McFarland, J. M.; Rosko, A. S.; Geissler, P. L.; Francis, M. B. The Impact of Assembly State on the Defect Tolerance of TMV-based Light Harvesting Arrays. *J. Am. Chem. Soc.* **2010**, *132*, 6068–6074.

8 Schlick, T. L.; Ding, Z.; Francis, M. B. Dual Surface Modification of the Tobacco Mosaic Virus. *J. Am. Chem. Soc.* **2005**, *127*, 3718–3723.

9 Holder, P. G.; Finley, D. T.; Clark, D. S.; Francis, M. B. Dramatic Thermal Stability of Virus-Polymer Conjugates in Hydrophobic Solvents. *Langmuir* **2010**, *26*, 17383–17388.

10 Holder, P. G.; Francis, M. B. Integration of a Self-Assembling Protein Scaffold with Water-Soluble Single-Walled Carbon Nanotubes. *Angew. Chem., Int. Ed.* **2007**, *23*, 4370–4373.

11 (a) Gilmore, J. M.; Scheck, R. A.; Esser-Kahn, A. P.; Joshi, N. S.; Francis, M. B. N-Terminal Protein Modification Using a Biomimetic Transamination Reaction. *Angew. Chem., Int. Ed.* **2006**, *45*, 5307–5311. (b) For previous studies on PLP-mediated transamination, see: Snell, E. E. The vitamin B6 group. V. The reversible interconversion of pyridoxal and pyridoxamine by transamination reactions. *J. Am. Chem. Soc.* **1945**, *67*, 194–197. Dixon, H. B. F. N-Terminal Modification Of Proteins−A Review. *J. Protein Chem.* **1984**, *3*, 99–108.

12 Scheck, R. A.; Dedeo, M. T.; Iavarone, A.; Francis, M. B. A Solid-Phase Approach for the Optimization of a Biomimetic Transamination Reaction. *J. Am. Chem. Soc.* **2008**, *130*, 11762–11770.

13 Witus, L.; Moore, T. A.; Thuronyi, B.; Esser-Kahn, A. P.; Scheck, R. A.; Iavarone, A.; Francis, M. B. Identification of Highly Reactive Sequences For PLP-Mediated Bioconjugation Using a Combinatorial Peptide Library. *J. Am. Chem. Soc.* **2010**, *132*, 16812–16817.

14 Dedeo, M. T.; Duderstadt, K. E.; Berger, J. T.; Francis, M. B. Nanoscale Protein Assemblies from a Circular Permutant of the Tobacco Mosaic Virus. *Nano Lett.* **2010**, *10*, 181–186.

15 Stephanopoulos, N.; Carrico, Z. M.; Francis, M. B. Nanoscale Integration of Sensitizing Chromophores and Photocatalytic Groups Using Bacteriophage MS2. *Angew. Chem., Int. Ed.* **2009**, *48*, 9498–9502.

16 Ma, Y.; Miller, R. A.; Fleming, G. R; Francis, M. B. Energy Transfer Dynamics in Light-Harvesting Assemblies Templated by the Tobacco Mosaic Virus Coat Protein. *J. Phys. Chem. B.* **2008**, *112*, 6887–6892.

17 Alexis, F.; Pridgen, E. M.; Langer, R.; Farokhzad, O. C. Nanoparticle Technologies for Cancer Therapy. *Handb. Exp. Pharmacol.* **2010**, *197*, 55–86.

18 Koudelka, K. J.; Manchester, M. Chemically modified viruses: principles and applications. *Curr. Opin. Chem. Biol.* **2010**, *14*, 810–817.

19 Hooker, J. M.; Kovacs, E. W.; Francis, M. B. Interior Modification of Bacteriophage MS2. *J. Am. Chem. Soc.* **2004**, *126*, 3718–3719.

20 Tong, G. J.; Hsiao, S. C.; Carrico, Z. M. Viral Capsid DNA Aptamer Conjugates as Multivalent Cell Targeting Vehicles. *J. Am. Chem. Soc.* **2009**, *131*, 11174–11178.

21 Hooker, J. M.; Datta, A.; Botta, M.; Raymond, K. N.; Francis, M. B. MR Contrast Agents from Viral Capsid Shells: A Comparison of Exterior and Interior Cargo Strategies. *Nano Lett.* **2007**, *7*, 2207–2210.

22 Meldrum, T.; Seim, K. L.; Bajaj, V. S.; Palaniappan, K.; Wu, W.; Francis, M. B.; Wemmer, D. E.; Pines, A. A Xenon-based Molecular Sensor Assembled on an MS2 Viral Capsid Scaffold. *J. Am. Chem. Soc.* **2010**, *132*, 5936–5937.

23 Hooker, J. M.; O'Neil, J. P.; Romanini, D. W.; Taylor, S. E.; Francis, M. B. Genome-free Viral Capsids as Carriers for Positron Emission Tomography Radiolabels. *Mol. Imaging Biol.* **2008**, *10*, 1536–1632.

24 Wu, W.; Hsiao, S. C.; Carrico, Z. M.; Francis, M. B. Genome-Free Viral Capsids as Multivalent Carriers for Taxol Delivery. *Angew. Chem., Int. Ed.* **2009**, *48*, 9493–9497.

25 Mastico, R. A.; Talbot, S. J.; Stockley, P. G. Multiple Presentation Of Foreign Peptides On The Surface Of An RNA-Free Spherical Bacteriophage Capsid. *J. Gen. Virol.* **1993**, *74*, 541–548.

26 Sletten, E. M.; Bertozzi, C. R. Bioorthogonal Chemistry: Fishing for Selectivity in a Sea of Functionality. *Angew. Chem., Int. Ed.* **2009**, *48*, 6974–6998.

27 Hooker, J. M.; Esser-Kahn, A. P.; Francis, M. B. Modification of Aniline-Containing Proteins Using a Chemoselective Oxidative Coupling Reaction. *J. Am. Chem. Soc.* **2006**, *128*, 15558–15559.

28 Bandrowski, E. Über die Oxidation des Paraphenylendiamins. *Chem. Ber.* **1894**, *27*, 480.

29 Wang, L.; Schultz, P. G. Expanding The Genetic Code. *Chem. Commun.* **2002**, *1*, 1–11.

30 Mehl, R. A.; Anderson, J. C.; Santoro, S. W.; Wang, L.; Martin, A. B.; King, D. S.; Horn, D. M.; Schultz, P. G. Generation Of A Bacterium With A 21 Amino Acid Genetic Code. *J. Am. Chem. Soc.* **2003**, *125*, 935–939.

31 Carrico, Z. M.; Romanini, D. W.; Mehl, R. A.; Francis, M. B. Oxidative Coupling of Peptides to a Virus Capsid Containing Unnatural Amino Acids. *Chem. Commun.* **2008**, *10*, 1205–1207.

32 Shangguan, D.; Li, Y.; Tang, Z.; Cao, Z. C.; Chen, H. W.; Mallikaratchy, P.; Sefah, K.; Yang, C. J.; Tan, W. Aptamers Evolved From Live Cells As Effective Molecular Probes For Cancer Study. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 11838–11843.

33 Stephanopoulos, N.; Tong, G. J.; Hsiao, S. C.; Francis, M. B. Dual-Surface Modified Virus Capsids for Targeted Delivery of Photodynamic Agents to Cancer Cells. *ACS Nano* **2010**, *4*, 6014–6020.

34 van Hest, J. C.; Tirrell, D. A. Protein-Based Materials: Toward A New Level Of Structural Control. *Chem. Commun.* **2001**, *19*, 1897–1904.

35 Coyle, P.; Philcox, J. C.; Carey, L. C.; Rofe, A. M. Metallothionein: the multipurpose protein. *Cell. Mol. Life. Sci.* **2002**, *59*, 627–647.

36 Esser-Kahn, A. P.; Francis, M. B. Protein-Crosslinked Polymeric Materials Through Site-Selective Bioconjugation. *Angew. Chem., Int. Ed.* **2008**, *47*, 3751–3754.

37 Hofmann, R. M.; Muir, T. W. Recent Advances In The Application Of Expressed Protein Ligation To Protein Engineering. *Curr. Opin. Biotechnol.* **2002**, *13*, 297–303.

38 Esser-Kahn, A. P.; Iavarone, A. I.; Francis, M. B. Metallothionein-Crosslinked Hydrogels for the Selective Removal of Heavy Metals from Water. *J. Am. Chem. Soc.* **2008**, *130*, 15820–15822.

39 Esser-Kahn, A. P.; Trang, V.; Francis, M. B. Incorporation of Anti-Freeze Proteins into Polymer Coatings Using Site-Selective Bioconjugation. *J. Am. Chem. Soc.* **2010**, *132*, 13264–13269.