Exploiting Chemoenzymatic and Photochemical Reactions to Control 4D Protein Presentation in Cell-Laden Hydrogels

Jared Shadish

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2019

Reading Committee:

Cole DeForest, Chair

François Baneyx

Elizabeth Nance

Program Authorized to Offer Degree:

Department of Chemical Engineering

University of Washington

Abstract

Exploiting Chemoenzymatic and Photochemical Reactions to Control 4D Protein Presentation in Cell-Laden Hydrogels

Jared Shadish

Chair of the Supervisory Committee:
Assistant Professor Cole DeForest
Chemical Engineering

As the primary workhorses of biology, proteins exhibit a wide range of functions including the ability to regulate cell fate with exquisite specificity. Though proteins represent powerful tools for tissue engineering and drug delivery applications, methods that permit their dynamic and heterogeneous presentation from and within materials remain limited. This thesis exploits recently developed chemoenzymatic strategies to tether biomacromolecules reversibly and site-specifically to hydrogel biomaterials, an approach that uniquely preserves their full bioactivity. First, the versatile *Staphylococcus aureus* enzyme Sortase A was employed to generate a diverse library of homogenous, singly functionalized fluorescent proteins, enzymes, and growth factors with bioorthogonal reactive handles for biomaterial modification. Additionally, N-myristoyltranferase was used in combination with an engineered photocleavable protein (PhoCl) to create

photoreleasable protein chimeras. Photolithographic patterning of these protein constructs was used to pattern dynamic biological response within 3D materials for the first time with subcellular resolution. The techniques presented here enable the creation of next-generation biomaterials that fully utilize the power and versatility of full-length proteins to probe and direct cell function in 4D.

Table of Contents

ACKNOWLEDGEMENTS

First of all, thank you to my committee for guiding me through the academic process. Particularly to Cole DeForest, without whose mentorship I would not have been able to accomplish the things that I have. I also would like to thank Elizabeth Nance, who lent me her ear and her time whenever I needed advice.

This work would not have been possible without the help of several extremely talented undergraduate students: Gabrielle Benuska, Mira Liu, Austin Im, and Alder Strange. This work would not have been possible without their commitment of time and mental energy to moving the projects forward. Their dedication to the lab while also studying for class, taking exams, and trying to have a normal undergraduate life is extraordinary.

I would also like to thank Brittney Hellner, Alexander Thomas, and Jessica Soto-Rodriguez from the Baneyx Lab for useful discussions on *E. coli* expression. I would never have been able to stick through the years of research without the support of my friends in the department. Thank you to Yi-Ting Lee and Brian Swift for getting me out of the basement of Benson Hall and into Seattle to see some sunlight. Special thanks also go to my lab mates Barry Badeau and Emily Ruskowitz for all of the help along the way, scientifically and otherwise. I would also like to thank Karina Miller, my biology consultant and girlfriend extraordinaire, whose love and patience got me through all of the ups and downs of graduate school. Thank you to my family for all of the support that they have given me despite the long distances involved.

Finally, thank you to the readers of this document. Your readership prevents this work from becoming just another dusty tome that is lost to history.

DEDICATION

This work is dedicated to all of my friends and family, without whose support none of the science would have been possible.

Chapter 1

Introduction and Background

Cells *in vivo* are constantly being exposed to a myriad of dynamically presented biochemical cues that govern critical aspects of cellular behavior, such as proliferation, migration, apoptosis, and a host of other phenotypic responses[1]. As an example of the power of these signals can have on cell fate, the developing *Drosophila* embryo anterior-posterior polarity is established through a gradient of the protein Bicoid[2]. Changes in Bicoid concentration during early embryogenesis can lead to dramatic downstream phenotypic changes such as mutant or repeated anterior regions in the developed organism[3]. In addition to spatial organization, the timing of signal presentation is also crucial. During embryogenesis, the protein Wnt shows biphasic activity in that it can drive cardiac cell differentiation, or it can act as a repressor for this very same process depending on when the signal is presented. This change in function occurs on the order of only a few hours[4]. Clearly, if there is ever to be significant improvement in the tissue engineering of whole organs or probing complicated biological questions, it is crucial to be able to recapitulate this dynamic signal presentation. As a first step, transitioning from two-dimensional polystyrene substrates to three-dimensional cell cultures platforms is crucial for capturing the intricate spatiotemporal signal presentation that exists *in vivo*. Hydrogels, water-swollen polymeric networks, have emerged as a tool for recapitulating some of this complexity[5,6].

Some of the earliest hydrogels in use were formed by extracting major components of the extracellular matrix, such as collagen, from natural sources and reconstituting them as a cell scaffold. These naturally derived materials benefit from high biocompatibility and degradability[7]. However, a disadvantage of these systems is that they suffer from high batch-to-batch variability

that is dependent upon the source of these naturally derived polymers, and experimental replication is complicated by such poorly defined systems[8]. There has been a push recently to move towards fully synthetic hydrogel systems to minimize the number of experimental degrees of freedom and to allow for greater control over the precise biochemical and biomechanical properties of the hydrogel[9].

Polymer chemists have developed an array of synthetic, biocompatible polymers that can be used as three-dimensional cell scaffolds with fully defined porosity, chemical composition, and mechanical stiffness[10]. Poly(ethylene glycol) (PEG) is the gold standard polymer used in tissue engineering applications as it is widely considered bioinert, optically clear, and can be engineered to have very similar water content and porosity to natural tissue[11,12]. Using PEG as a platform, hydrogels can be modified with biochemical moieties to promote vital cellular functions such as matrix remodeling and adhesion[13,14]. Incorporation of these chemical functionalities relies upon the ability to perform bio-orthogonal cytocompatible chemistries.

A variety of mild, catalyst-free chemistries have been developed in the search for new, facile methods for engineering materials in the presence of cells[15]. One particularly powerful chemistry utilizes the cycloaddition of an azide and an alkyne, two chemical functionalities not found in biology[16]. The kinetics of this reaction can be dramatically improved by adding ring strain to the alkyne, known as strain-promoted azide-alkyne cycloaddition (SPAAC)[17]. This chemistry can be used to form hydrogels between azide-functionalized and bicyclononyne (BCN) containing PEG monomers on the order of minutes, readily encapsulating cells with high viability. Additionally RGDS peptides, commonly used to promote cell adhesion, and degradable crosslinks can be included stochastically using this method[18].

While SPAAC allows the facile encapsulation of living cells and biochemical moieties, it lacks the ability to generate gradients or arbitrarily defined regions of these cues. Light-based chemistries enable dynamic 3D control over biochemical immobilization because exposure to light can be easily controlled in space and time. As techniques to expose samples to controlled wavelengths of light have been refined for use in microscopy, these techniques can be co-opted to initiate photochemical reactions with extremely high resolution, and with the advent of multiphoton lithography, photochemical reactions can be fully localized to 3D subvolumes in a material.

Initially, photochemical patterning techniques were limited to peptides, which are easily synthesized with the precise functional handles needed, avoiding the vast heterogeneity of chemistries found in protein samples. The Stoichet and West labs are early pioneers of this work, publishing work on photomediated biochemical coupling in 2008 and 2010 respectively[19,20]. The West lab used the photoinitiated crosslinking of acrylate containing molecules within a PEG-acrylate hydrogel. The Stoichet lab used a photoprotected thiol and patterned in maleimide-functionalized biomolecules. While these techniques were successful in creating arbitrarily defined 3D regions of relatively simple biomolecules, they lack the broad applicability to pattern full-length proteins. Acrylate crosslinking involves a free-radical mechanism which has broad cross reactivity with other functional groups and may be cytotoxic. Maleimide-thiol reactions on the other hand lack specificity in complex biological systems which naturally contain thiols, such as live cells or proteins with exposed cysteine residues.

One technique recently used by DeForest et al. is the combination of a 2-(2-nitrophenyl)propoxycarbonyl (NPPOC) photocaged alkoxyamine and oxime ligation with an aldehyde labeled protein[21]. NPPOC undergoes cleavage upon exposure to cytocompatible 365 nm

light, liberating an alkoxyamine[22]. The alkoxyamine reacts with an aldehyde to form a covalent oxime linkage that is stable at neutral pH[23]. By combining this technique with an *ortho*-nitrobenzyl ester (*o*NB), a photocleavable linker which also cleaves upon exposure to 365 nm light, proteins were reversibly patterned with full 3D control within PEG based hydrogels. As aldehydes and other clickable handles are not found in any of the canonical amino acids, a technique to incorporate this functionality onto proteins is needed. The following section will investigate the strategies currently available for protein modification.

1.1 Site-Selective Protein Modification: From Functionalized Protein to Functional Biomaterial

Abstract:

Proteins are the primary working unit of biology. Their diverse range of activities and exquisite specificity have made them attractive candidates for providing solutions to complicated chemical and biological problems that have proven intractable to small molecule techniques. However, selectively modifying such complex biomolecules that contain a wide array of functional groups in virtually limitless sequence and conformational configurations for immobilization in biomaterials is an enormous challenge. In this section, we focus on methods for the site-specific modification of proteins, with an emphasis on techniques that enable the creation of functional biomaterials. We highlight promising strategies that use cutting-edge small molecule techniques, genetic engineering approaches, and chemoenzymatic strategies that offer unprecedented control over the extent and location of protein modification. Finally, we discuss future applications of these techniques to create the next generation of biomaterials.

1.2 Introduction:

Proteins are the main workhorses of biology with nearly infinite possibilities for sequence, structure, and function. Given their potential for exquisite control over biological functions, they are of prime interest in controlling delicate cellular functions for the creation of therapeutics and biomaterials[24,25]. It is becoming increasingly apparent that control over protein orientation and spatiotemporal presentation within a material are important for fully harnessing the power of full-length proteins. Typically some modification must be done to the protein to label it at a specific site for selective incorporation into a material among complex biological mixtures. However, modifying such complicated biopolymers for use in these contexts without perturbing their

function remains an open challenge[26]. Random modifications have the potential to block enzymatic active sites, binding pockets, or disrupt the structure of the protein. Despite the difficulties in achieving site-specific protein modification, early pioneering work using stochastic incorporation of functional groups onto antibodies has revolutionized biology, enabling the now widely used techniques of immunofluorescence and ELISA[27,28]. With these early successes, techniques to further refine specificity for modification of more delicate proteins have been of prime interest for utilizing these fragile but powerful tools in novel applications.

No matter the technique, careful consideration must be paid to the reaction conditions for protein modification. The reactions must occur under mild, aqueous conditions and proceed relatively rapidly on substrates at very low molar concentrations[29]. The protein targets themselves are composed of a wide array of functional groups with varying nucleophilicity, pKa, and redox potential. Targeting a specific functional group from the many that may be available on the protein's surface, let alone a single functional group among nearly identical copies, is a tall order. Several strategies have emerged to increase the specificity of reactions that can be broadly broken into two categories: endogenous protein modification using small molecules, and genetic or protein engineering approaches that append sequences for subsequent selective modification.

Typically, chemoselectivity is achieved by utilizing the differences in nucleophilicity, pKa, and redox potential of a given side chain. The most widely targeted side chains have historically been lysine and cysteine due to their high nucleophilicity[30]. N-hydroxysuccinimide and maleimides act as electrophiles for conjugation to amines and thiols, respectively. While these techniques enabled the creation of functional bioconjugates, they often have deleterious effects on protein activity, and further refinement over selectivity is required.

Genetic engineering approaches have proven quite successful in refining the selectivity of bioconjugation. Genetic code expansion allows for the incorporation of amino acids containing non-natural functional side chains into a specific location within the primary sequence of a protein. However, this technique often suffers significant decreases in protein expression yields with each non-natural amino acid incorporated and requires specific evolved tRNA / tRNA synthetase pairs for each unique non-natural amino acid to be loaded. Encoding longer sequences of natural amino acids can increase specificity while still utilizing the cell's native translational machinery.

Protein engineering strategies to append short peptide sequences to full-length proteins has proven very successful in increasing selectivity. Refinement of the microenvironment around a particular amino acid allows for increased specificity by subtly changing the reactivity of a specific residue. In addition, chemoenzymatic solutions have become increasingly popular. Combining the selectivity of an engineered peptide tag with the exquisite selectivity of enzymes has allowed for unparalleled control over the site of modification while retaining high conversions.

Ultimately, proteins need to be modified mildly, selectively, and with high conversion to retain their activity for use in biomaterials. There is increasing demand for techniques to control orientation of proteins on surfaces, create crosslinks within hydrogels using proteins with novel functions, and for spatiotemporal control over biochemical presentation. Each technique for protein conjugation comes with its own set of reaction conditions and amenable chemistries that require careful consideration for the protein and application of interest.

1.3 Chemoselective Protein Modification

Most historically prevalent examples of protein modification have relied on inherent differences in reactivity among the side chains to dictate conjugation. While these techniques have

traditionally had relatively good chemoselectivity, they have typically had poor regioselectivity. Careful engineering of the reagents for protein modification have led to increased specificity. An attractive aspect of this technique is that off-the-shelf proteins may be used as there is no need for genetic engineering. However, there is still much room for improvement by increasing conversion while maintaining specificity.

1.4 Amine Modification

Lysine has proven to be a popular target for modification due to its ubiquity (~6% of residues in the human proteome) and the nucleophilicity of the ε-amine located on the side chain. However, this abundance and inherent reactivity make site-specific incorporation challenging. The most widely used electrophiles for conjugation with primary amines on the protein surface are N-hydroxysuccinimide esters (NHS), which have a long history for use in peptide synthesis[31]. Despite the stochastic incorporation of functionalization for protein modification, the ease and simplicity of NHS modification of proteins have led to their use in the creation of elegantly patterned biomaterials. The Stoichet lab NHS-functionalized epidermal growth factor (EGF) with biotin for immobilization in a streptavidin patterned hyaluronic acid (HA) hydrogel[32]. They found that immobilized EGF was still able to affect invasion of MDA-MB-231 and MDA-MB-468 breast cancer cells. In another example, DeForest and Tirrell used NHS chemistry to modify a variety of proteins with aldehydes and photoreactive ortho-nitrobenzyl esters for photopatterning of polyethylene glycol (PEG) hydrogels[21]. By reversibly immobilizing vitronectin during human mesenchymal stem cell culture, they found that osteogenesis could be spatiotemporally defined. NHS chemistry offers a quick off the shelf method for modifying proteins that has proven successful in many applications.

Despite these successes, stochastic conjugation to amines can have a deleterious effect on many proteins. As a good nucleophile and with the ability to act as a base in the proper pH, lysine is often part of the catalytic triad of enzymes thus necessary to maintain protein activity. Beta-lactamases in particular are known to utilize lysine in their active pocket[33]. In one interesting example, this enhanced reactivity of a lysine in the catalytic domain of a beta-lactamase fused to a monoclonal antibody was utilized to site-specifically incorporate a beta-lactam containing chemotherapeutic[34].

Computational approaches to electrophile design have also been exploited to take advantages of minute differences in the pKa between lysine residues to site-specifically modify proteins. Sulfonyl acrylate reagents were designed to exploit subtle differences in reactivity on native lysine residues to give site-specific incorporation on a variety of proteins, including installation of a fluorophore onto Trastuzumab without destroying its activity[35]. While an impressive demonstration of the power of computational design of small molecule reagents, there remains significant room for improvement. Conversion occurred at the lysine with the lowest pKa, which may be a critical or inaccessible residue on some proteins, and high excesses of sulfonyl acrylate reagents (1000- to 10000-fold excess) were required for quantitative conversion.

Multistep reactions have also been used in an attempt to control the extent and location of lysine modification. Using a multi-component reaction using an aldehyde, an acetylene, and a Cu-ligand complex, propargylamine was appended site-specifically onto nine different proteins[36]. While the modified proteins retained their activity, this technique required long reaction times (up to 72 hours) with relatively low conversion (generally 40-60%). Lysine modifications remain one of the most common techniques for protein modification, but achieving high conversion while maintaining specificity remains a challenge.

1.5 Thiol Modification

Cysteine residues are also a popular choice for protein modification due to the high nucleophilicity of the thiol on the side chain, and the relative lack of free sulfhydryl groups on native proteins. Incorporation of a single cysteine residue via genetic engineering is commonly employed to give a functional handle for conjugation. Traditionally, maleimides have been used for thiol alkylation and protein conjugation. Maleimides have long been used for the conjugation of Fab and Horseradish Peroxidase for immobilization onto surfaces for the creation of ELISA assays[37]. However, it was recently discovered that the thioether bond may undergo thiol exchange or ring opening and decompose *in vivo*[38]. In addition, many proteins do not contain native free thiols that can be targeted, and reduction of disulfide bonds can have deleterious effects on activity by disrupting protein structure. Despite this limitation, there have been several FDA approved maleimide conjugates, such as brentuximab vedotin, trastuzumab emtansine, and certolizumab pegol. To further increase clinical translation of these conjugates, more stable and specific chemistries are required.

To overcome some of the limitations of maleimide compounds, thiol targeting reagents have been developed for increased aqueous solubility and metabolic stability. Oxetanes have found use for thiol modification, and were shown to be stable in human plasma and in the presence of thiol-containing nucleophiles such as glutathione and β-mercaptoethanol[39]. Amine to thiol crosslinking reagents, such as succinimidyl 4-(N-maleimidomethyl)cyclohexane-1-carboxylate (SMCC), have also proven popular. These compounds contain maleimide and NHS functionalities permitting conjugation to lysines on a protein of interest that does not contain free thiols through amine reactivity. Then, the sulfhydryl groups of a second protein are conjugated using the maleimide group to create protein-protein heterodimers. Sodium 4-((4-

(cyanoethynyl)benzoyl)oxy)-2,3,5,6-tetrafluorobenzenesulfonate (CBTF) is a recently developed amine to thiol coupling reagent that was used to produce antibody drug conjugates with improved plasma stability as compared to maleimide conjugation[40].

Where free sulfhydryl groups are not present, disulfide bonds can be targeted using disulfide rebridging. In this technique, native disulfide bonds are rebridged via bisalkylation using a linker that includes exogenous functionality. The Brocchini group demonstrated that interferon alpha-2a and antibody fragments could be site-specifically PEGylated and retain their activity. By increasing the stoichiometric ratio of the PEG monosulfone, alkylation occurs at each of the thiol groups leading to double labeling of the protein[41]. Further development of alkylation reagents has enabled the creation of compounds capable of reducing and rebridging disulfides with one reagent[42], photo-mediated disulfide rebridging using ultraviolet (UV) light[43], and allyl sulfones with improved water solubility for the functionalization of peptides and proteins[44]. Recently, disulfide rebridging has been used to conjugate toxic payloads to antibodies for targeted therapeutic delivery[45]. Disulfide rebridging is a powerful technique for modifying proteins with accessible disulfide bonds, especially in cases where maintaining the disulfide linkage is critical to function.

A promising class of thiol-based reactions due to its amenability to photomediated conjugation is thiol-ene click chemistry. These reactions proceed through either a Micheal-type addition or radical mediated conjugation that leads to anti-Markovnikov addition of a thiol to an alkene with high selectivity and yield, enabling the creation of protein functionalized biomaterials. Photochemical thiol-ene coupling typically requires the addition of a photoinitiator to generate radicals for radical-mediated addition. A recently developed allyl sulfide chain-transfer agent

enables the reversible immobilization of proteins into hydrogels using UV light. Thiolated transferrin and ovalbumin were photoimmobilized in the presence of lithium phenyl-2,4,6-trimethylbenzoylphosphinate (LAP) and released with 3D control using multiphoton lithography[46]. The Weber group used a thiolated phytochrome-based protein (Cph1) to create a hydrogel with cyclic stiffening properties[47]. Cph1 engineered to contain a free sulfhydryl was conjugated to an 8-arm PEG-vinyl sulfone, and the Cph1 was allowed to dimerize to form gels. Upon exposure to 660 nm light, the proteins became monomeric, leading to gel softening. The process could be reversed upon exposure to 740 nm light, which favored protein dimerization. While UV mediated thiol-ene conjugations proceed rapidly, there are generally concerns over the cytocompatibility of the photoinitatiors and radical mediated reactions[48].

1.6 Aromatic Residue Modification

Occasionally the aromaticity of tyrosine, tryptophan, and phenylalanine residues has been a target due to the relatively low abundance of aromatic residues on solvent accessible regions of protein surfaces, making site-specific incorporation more tractable. However, there is difficulty in maintaining selectivity for one particular aromatic residue over another due to the small difference in reactivity between the side chains.

Sodium trifluoromethanesulfinate has been used to selectively install CF_3 groups onto tryptophan residues in full-length proteins for use in F-NMR studies[49]. This technique enables the curious ability to create "Teflon proteins", which are highly hydrophobic (the CF_3 group is twice as hydrophobic as CH_3), potentially increasing the bioavailability and reducing the metabolic detoxification of pharmaceutically active proteins. A photocatalytic method for modification of the beta-position of tryptophan has also been developed[50]. An iridium photocatalyst, which is active under blue light, was used to incorporate a methyl acrylate onto the tryptophan residue on

model peptides with good selectivity over other aromatic amino acids. While relatively unexplored, the unique physiochemical effects of modifying aromatic residues enable novel functionalities that are otherwise inaccessible when modifying nucleophilic residues.

1.7 N/C-Terminal Modification

Much effort has been expended to find reagents suitable for modifying the termini of protein substrates[51]. As every polypeptide chain contains a unique N- and C-terminus, and these ends are often solvent accessible, modification of these groups facilitates a site-specific modification without the need for extensive protein engineering. Typically, slight differences in reactivity, pKa, and redox potential between the termini and amines or carboxylic acids off of side chains are used to localize modification to the termini. Generally, it has proven difficult to obtain quantitative conversions of the termini without leading to off-target modification of the side chains. However, the benefits of performing site-specific modification on native proteins proves alluring, and many chemistries continue to be developed with great promise in achieving chemo- and regioselectivity.

Several small molecule approaches to selectively modify the N-terminus have emerged. O-aminophenols react readily with the N-terminus of proteins using a potassium ferricyanide catalyst; although, the reaction tends to prefer N-terminal proline residues as a substrate[52]. However, free cysteines are also modified by the reaction, and the conversion is relatively low (20-60%). Aldehyde containing reagents such as 2-pyridinecarbaldehyde (2-PCA) have proven popular for selective reductive alkylation of the N-terminus[53]. One advantage of this technique is that reductive alkylation preserves the amine on the N-terminus, theoretically leading to more native charge distribution on the protein. Antibodies have been site-selectively modified with an azide-functionalized PCA for further click chemistry conjugation[54]. However, achieving complete

functionalization of both chains of the antibody proved difficult, even at a very high excess of the coupling reagent. In a biomaterials application, polyacrylamide (PAAm) hydrogels were functionalized with 2-PCA and used to immobilize ECM proteins by their N-termini[55]. These ECM functionalized substrates promoted cell adhesion and spreading, and supported the creation of large collagen fibers, an effect that could not be achieved using succinimide chemistry for lysine modification.

Another common chemistry is for N-terminal transamination through the use of pyridoxal 5'-phophate (PLP). The reaction oxidizes the N-terminus to a ketone or aldehyde, which can be subsequently used for oxime ligation. Originally used to modify peptides, PLP gave 65% conversion of the N-terminus into a ketone in 2 h at 37 °C[56]. Where modification of native N-termini on full-length proteins proves challenging, such as in the case of monoclonal antibodies, site-directed mutagenesis can be used to achieve high levels of conversion by increasing the reactivity of the terminus[57]. Streptavidin was immobilized through an N-terminal oxime ligation onto micropatterned aminooxy surfaces[58]. This selective control over the location of protein immobilization allowed for controlling the protein orientation on surfaces, which has important implications for making enzymatically reactive surfaces for industrial or biosensing applications.

N-terminal cysteine residues provide a unique handle for subsequent modification. Native chemical ligation utilizes the free thiol of the cysteine which is reacted with a C-terminal thioester. The bond then undergoes an S-N acyl shift to generate an amide bond. However, engineering proteins to have an N-terminal cysteine residue requires careful consideration, as the methionine from the start codon must often be excised, and it can prove difficult to create C-terminal thioesters during solid phase peptide synthesis for protein modification[59]. Expressed protein ligation, using recombinantly expressed polypeptides which self-cleave to contain the required functionalities,

has also seen some development. In one example, expressed protein ligation was used to create site-specifically SUMOylated or methylated histones to investigate the impact of post-translational modifications on histone regulation[60].

N-terminal cysteines can also react with aldehydes to produce thiazolidines, however the reactions tend to be inefficient. Ortho-boronic acid substituted benzaldehydes have been used to give site-specific conjugation of N-terminal cysteines with a unique thiazolidino boronate structure. However, this bond was found to degrade under mildly acidic conditions and gave modest conversions (~50%).

While not as thoroughly explored as the N-terminal modifications, selective C-terminal modification has also been performed. Mild differences in the oxidation potential of the C-terminus versus side chains allowed for decarboxylative alkylation to be performed exclusively at the C-terminus[61]. Using a visible-light photoredox catalyst, insulin was modified to add a C-terminal alkyne with 40-50% conversion. Investigation into post-translational modifications in the proteome reveal proteins with C-terminal methylation and α-amidation, suggesting the possibility of future chemistries to be developed to target this functionality[62].

1.8 Metabolic Labeling

While small molecules have proven successful at site-specific protein modification and biomaterial functionalization, they require the exogenous treatment of the protein of interest with small molecule reagents. There is growing interest in repurposing protein translational tools for incorporation of non-natural functionality during polypeptide elongation to enable co-translational protein modification. An early example of this technique is the metabolic incorporation of azidohomoalanine (AHA), an azide containing methionine analog that is incorporated by the native

tRNA synthetase into the growing polypeptide chain stochastically when added during protein expression[63]. Similarly, an alkyne containing methionine analog homopropargylglycine (HPG) may also be incorporated[64]. Both of these tags proved useful for subsequent purification and fluorophore conjugation onto newly synthesized proteins using click chemistry.

To increase the specificity of metabolic engineering, genetic engineering may be performed to enrich a given domain with the residue to be incorporated. In one example, GFP was engineered to have a C-terminal -HMHMHMHM tag[65]. After metabolic incorporation of AHA, the methionine residues were replaced with azide-containing side chains. This allowed for purification via immobilized metal affinity chromatography using the histidine residues, and subsequent immobilization onto an alkyne containing agarose resin.

1.9 Amber Codon Suppression and Genetic Code Expansion

Genetic code expansion may be used to localize the site of modification to a single genetically encoded residue. Typically, the amber stop codon (UAG) is overridden to instead incorporate non-natural functionality using mutated tRNA/tRNA synthetase pairs. The Schultz group has done pioneering work in developing orthogonal tRNA/tRNA synthetase pairs which are specific to the amber stop codon without cross recognition with endogenous synthetases, typically by utilizing synthetases from archaea and other distantly related organisms. Subsequently, this technique has been further developed to create site-specific post-translational modifications onto proteins, as well as to install light activatable functionality.

Post-translational modifications that govern protein-protein interactions are highly sensitive to the site of incorporation. To investigate this, many natural post-translational modifications have been engineered for incorporation using genetic code expansion. An ε-N-2-

Hydroxyisobutyryl-lysine (HibK) was site-specifically incorporated into histones to probe its function[66]. In addition, the effects of protein phosphorylation has been investigated through the site-specific incorporation of genetically encoded phosphotyrosine residue[67]. A dimethylated lysine residue was site-specifically incorporated into histones and p53 in order to probe the function of epigenetic regulators such as histone demethylase LSD1 and histone acetyltransferase Tip60[68]. Beyond natural post-translational modifications, there is significant interest in introducing novel functionalities into proteins using genetic engineering as well.

Light activatable amino acids have proven popular to give spatiotemporal control over protein activation. By protecting a side chain critical for protein activity with a photoprotecting group, protein activity can be turned on upon exposure to light. A photocaged lysine was incorporated into live zebrafish embryos to control the activity of luciferase and various kinases using light[69]. Further, a photocaged cysteine was used to control TEV protease activity in mammalian cells[70]. While these techniques enable control over protein activity in time and space, photochemistries which can localize protein immobilization are also useful in developing materials with engineered functionalities.

Several non-canonical amino acids have been developed which are able to undergo photomediated ligation. A genetically encoded photo-lysine has an alkyne tag for labeling as well as the ability to crosslink nearby proteins upon exposure to UV light[71]. The Bertozzi group incorporated a quadricyclane that could be used for quadricyclane ligation where the ligation product would undergo photocleavage upon exposure to light[72]. Reactive aryl carbamate side chains have been incorporated that react with adjacent nucleophiles to create a crosslink[73]. A dual functional phenylazide / phenylamine side chain (AmAzZLys) was developed with a photoreactive azide for UV induced crosslinking and a phenylamine for bioconjugation[74]. This was used to site-

selectively modify an antibody with a fluorophore and subsequently crosslink it to an antigen. A photoswitchable click amino acid which contains azo-benzene photoswitches was incorporated into calmodulin and shown to modify protein structure in response to light[75]. Finally, genetic code expansion was used for the conversion of an O-phosphoserine (Sep) residue into a ligated product using a two-step reaction of dephosphorylating the Sep residue to a dehydroalanine followed by zinc and copper mediated addition of alkyl iodides to the dehydroalanine[76]. While genetic code expansion has enabled many novel functionalities to be incorporated site-specifically into proteins, it remains limited in the substrates which can realistically be incorporated into the tRNA synthetase while maintaining orthogonality, and in maintaining high protein expression yields through multiple non-canonical amino acid insertions.

1.10 Motif Insertion

Going beyond genetic code expansion, installation of motifs, short peptide domains, onto proteins enables unique selectivity over protein modification. These domains function by enabling careful control over the microenvironment of specific residues to increase the efficiency and specificity of reactions[77]. When used in combination with small molecule chemistries for protein modification, an impressive degree of control over protein functionalization can be achieved. However, these still tend to be fairly limited in scope with non-quantitative conversions, though computational modeling of microenvironments may change this.

Perhaps the best known example of a peptide epitope leading to enhanced binding is the "His-Tag". Polyhistidine sequences (usually either His6 or His10 tags) on proteins allow for non-covalent binding with metal ions (Ni^{2+}, Co^{2+}, Cu^{2+}) which can be used for protein purification in immobilized metal affinity chromatography[78]. It has been found that His-tagged proteins may undergo N-terminal acylation with D-gluconic acid δ-lactone. The GDL tag ($GHHH_n$) has been

developed to enable N-terminal acylation of proteins. Gluconolactone and 4-methoxyphenyl esters were used as acylating agents to PEGylate the N-terminus of multiple full-length proteins[79]. This modification was found to be reversible after extended incubation in PBS.

Phage display typically has been used to determine epitopes which bind specific substrates. Fluorettes are one such short peptide sequence that have been found to bind common flurophores such as Rhodamine Red, Oregon Green 514, and fluorescein. When a maleimide functionalized Texas Red probe is used, site-specific conjugation at a cysteine residue in the fluorette can be achieved[80].

When combined with chemoselective modification techniques, motif insertion gives impressive regioselectivity. One peptide sequence identified through library screenings is the dibenzocyclooctyne (DBCO) tag (LCYPWVY) which when reacted with a DBCO for thiol-yne click conjugation forms a stable covalent bond on the free sulfhydryl group of the encoded cysteine[81]. This engineered microenvironment was found to lead to a 220 fold increase in reaction rate as compared to a cysteine peptide control. Trastuzumab and EGFP with terminal DBCO tags were modified site-selectively with 80-90% conversion.

Increased reactivity of specific cysteine residues has been engineered through the addition of a pi clamp (FCPF)[82]. Reactions with perfluoroaryl compounds and cysteines are typically very slow, except in the presence of a catalytic enzyme. Thus the microenvironment around specific cysteine residues was explored to find sequences specific to increasing the reaction rate. The FCPF sequence was found to allow for site-specific conjugation in the presence of endogenous thiols, and used to conjugate affinity tags, fluorophores, alkynes, and PEG onto protein substrates at either the N- or C-terminus, or in an internal loop. This technique was found to lead to >95% conversion of the target thiol in 30 minutes.

Creating sites for efficient disulfide rebridging has also been explored through the addition of a Dis-Tag (CISTCC), which contains three cysteines as potential sites of modification[83]. The Dis-Tag allows for stepwise conjugation of maleimide containing compounds for conjugation of the free thiol followed by the addition of allyl sulfones for disulfide rebridging. EYFP was engineered to include the Dis-tag and subsequently dual functionalized with two different fluorophores. Interleukin-2, which contains a native disulfide was also functionalized through rebridging and shown to have no loss in bioactivity. This is particularly noteworthy as reduction of the native disulfide bond of IL-2 leads to a 90% loss in activity.

Histidine bis-alkylation has been used to increase the selectivity of histidine modification. Histidine is a relatively rare residue, and by constraining the conjugation chemistry to require two nearby histidine residues, the specificity is increased. The HGH tag was incorporated into IFN away from the active site and used for site-specific PEGylation[84]. Conversion was found to be modest (~20-40%), though the activity of modified IFN (74% compared to wild type) was the highest reported for any PEGylated IFN. Generally, increasing the length of the motif leads to increased specificity. This leads to the idea of fusing entire enzymes, which have exquisite selectivity, to the protein of interest to facilitate modification.

1.11 Direct Ligand Immobilization

Proteins have evolved to have strong binding affinity selectively for their binding partners in the extremely crowded and complex biological environment within living cells. The ability of proteins to form strong noncovalent interactions or even covalent bonds with their substrates allows for protein modification at the active site of the enzyme. When expressed as a fusion protein with a protein of interest, this allows for the site of substrate conjugation to be localized away from the protein of interest.

Perhaps the most commonly used example of direct ligand immobilization are the SNAP tag and Halo tag. The SNAP tag is a modified form of O6-alkylguanine DNA Alkyltransferase which enables conjugation of O6-benzylguanine analogs to the SNAP domain[85]. This technique has been used to functionalize proteins with ferrocene for immobilization onto cyclodextrin micropatterned surfaces and vesicles[86]. The Halo tag is an engineered haloalkane dehalogenase that irreversibly binds to chloroalkane substrates. Fusions with the Halo tag have been used for conjugating full-length proteins to small molecule fluorophores and agarose resins[87]. Another example is the enzyme Cutinase, which forms a covalent adduct with phosphonate ligands. Fusions between Cutinase and antibodies have allowed for the creation of immobilized antibody arrays with controlled orientation[88].

One of the most exciting recent developments in site-specific protein modification has been the SpyTag-SpyCatcher system. The Howarth group has engineered fibronectin-binding protein FbaB into two fragments which are capable of forming an isopeptide bond between a lysine and aspartic acid residue. Renamed SpyTag and SpyCatcher for the two fragments, isopeptide bond formation was found to form rapidly and near quantitatively in diverse reaction conditions[89]. The domains have been further split to create a SpyLigase, which is an enzyme that covalently conjugates SpyTag (AHIVMVDAYKPTK) with the KTag (ATHIKFSKRD)[90].

Protein hydrogels have been formed by creating protein fusions with multiple SpyTag or SpyCatcher domains. Upon mixing the two components together, rapid gelation is observed as the SpyCatcher forms a covalent bond with SpyTag. In one example, Elastin Like Proteins (ELPs) are fused to either 3 SpyTag domains or 2 SpyCatcher domains[91]. After mixing, hydrogels form within 5 minutes. Live cells can be encapsulated within the network with high viability, and proteins such as SpyTag functionalized mCherry and leukemia inhibitory factor were immobilized into the

network during gelation. Similar protein hydrogels based on globular domains (GB1 and FNIII) as fusions with SpyTag and SpyCatcher were also formed and shown to be able to encapsulate viable cells[92]. The Clark group created PEG hydrogels with SpyTag crosslinks for subsequent immobilization of SpyCatcher labeled proteins[93]. This allowed for immobilization of proteins at very high ligand density. When co-immobilizing mCerulean and mVenus, FRET efficiencies of up to 45% were observed, pointing to the close proximity of tethered proteins. While enabling site-specific conjugation, fusions with the bulky enzymatic groups may have deleterious effects on protein function[94]. The FbaB domains have been further split to create a SpyLigase, which hacts as an enzyme that covalently conjugates SpyTag (AHIVMVDAYKPTK) with the KTag (ATHIKFSKRD)[90]. This technique harnesses the benefits of protein specificity without necessitating the addition of the bulky SpyCatcher group; an advantage that other groups have sought to exploit through engineering other enzymes for protein modification.

1.12 Chemoenzymatic Protein Modification

Enzymes are biocatalysts that are able to perform exquisitely selective reactions by utilizing the sterics and unique microenvironment present in large biomacromolecules to change the energy landscape of a reaction. Taking advantage of these enzymes for their selectivity, and engineering them for expanded function has enabled unparalleled site-selective modification of full-length proteins[95]. Many enzymes have been discovered with unique functionalities for the creation of modified bioactive proteins for use in materials.

1.12.1 Sortase A

Sortase A (SrtA) is a transpeptidase that canonically recognizes the LPXTG motif on proteins and conjugates an aminoglycine nucleophile in the presence of calcium. Originally found in *Staphylococcus aureus* as the enzyme responsible for tethering proteins to the cell wall[96], SrtA's

unique functionality has gained significant popularity for the "sortagging" reaction. In brief, the cysteine residue in the active pocket of SrtA forms a thioester intermediate by cleaving between the threonine and glycine in the LPXTG motif. The thioester is then displaced by an aminoglycine nucleophile, typically an N-terminal polyglycine sequence on a peptide or a protein, forming a stable amide bond. This reaction reforms the LPXTG motif, making the process reversible. SrtA has been evolved to have a higher catalytic activity, be calcium independent, and to recognize different substrates[97].

Sortase has been used not only for PEGylation and in the creation of protein-protein conjugates, but also for immobilizing proteins into materials. By immobilizing a pentaglycine peptide onto surfaces, LPETG-containing proteins could be reversibly conjugated[98]. The Griffith group created synthetic PEG hydrogels that contained an LPRTG sortaggable substrate[99]. GGG-EGF was subsequently immobilized within the gels using sortase, and the tethered EGF was found to promote DNA synthesis in primary human hepatocytes and endometrial epithelial cells. In another example, EGFP with a C-terminal LPETG tag was immobilized onto polystyrene beads containing GGG[100]. Interestingly, it was also found that alkylamines were also a suitable substrate for protein immobilization in lieu of the canonical polyglycine domain.

Protein engineering to create protein-sortase fusions for intramolecular sortagging has also been performed to create single step purification / conjugation strategies[101]. Proteins of interest are expressed as an N-terminal fusion to a 6xHis containing SrtA enzyme with an LPETG spacer. Upon addition of calcium and a GGG-tagged probe of interest, intramolecular transpeptidation by the SrtA enzyme conjugates the GGG-tagged probe to the C-terminus of the protein of interest while displacing the 6xHis SrtA domain during immobilized metal ion chromatography. This technique has been used to conjugate photoactive and click moieties onto various proteins for the

creation of elegantly protein-patterned hydrogels[102]. Sortagged and tethered proteins were found to be bioactive and able to spatiotemporally govern cellular response.

While traditionally SrtA has been used to modify the N/C terminus of proteins, recent work has enabled sortagging of internal lysine residues through isopeptide bond formation. Isopeptide bond formation is achieved by encoding the pilin domain (WX3VWVYPKH) to make the ε-amino group of lysine a more reactive nucleophile. This reaction was found to be selective and achieved high conversion (>95%) on peptide substrates and on antibodies[103]. This technique enables the facile incorporation of multiple sortagged modifications per protein, and could be used in cases where the N/C terminus of a protein is inaccessible or would have deleterious effects on protein activity.

1.12.2 Subtiligase

Subtiligase is an engineered ligase which conjugates C-terminal activated esters, such as a thioester or a glycolate phenylalanyl amide activated ester, with the N-terminus of peptides and proteins[104]. Although similar to native chemical ligation, subtiligase was found to enhance the reaction rate by approximately three-fold. Subtiligase has been used in the synthesis of full length ribonuclease A from six peptide fragments, as well as in the conjugation of phosphopeptides to phosphatase and tensin homolog PTEN in the investigation of post-translational phosphorylation on protein activity[105]. Conversion is relatively good (~70%) after 60 min reaction. Subtiligase has found utility in the proteomics community, where labeling free N-termini is useful in identifying proteolytic cleavage sites[106,107].

1.12.3 Microbial Transglutaminase

Microbial transglutaminase facilitates the transamidation of glutamine side chains to lysine residues to create an ε-(γ-glutamyl)lysine isopeptide bond[108]. In its original incarnation it was not specific, and has long been used to improve the texture in foods. Recent developments in using transglutaminase for protein modification have enabled site-specific protein modification and biomaterial formation[109].

Site-specificity has typically been achieved through the engineering of glutamyl or amine donor peptides with specific transglutaminases for example, the Q-tag, YAHQAHY, or K-tag, KKKKKK, for *Bacillus subtilis*[110]. These substrates increase the reactivity of the specific residue, allowing for site-specific translglutaminase conjugation. In one example, biotin ligase with an N-terminal Q tag was immobilized onto magnetic microspheres and was shown to retain >95% activity[111]. Newly discovered microbial transglutaminases, such as that from *Kutzneria albida* show high specificity for their substrates (YRYRQ and RYESK) and have been used for the site-specific labeling of antibodies with biotin[112].

Hydrogels have also been formed through transglutaminase crosslinking. Hyaluronic acid was functionalized with glutamine, and an 8-arm PEG was end functionalized with lysine residues for crosslinking through transglutaminase[113]. Hydrogels formed in this manner showed increased adhesion and spreading of MCF7 and C2C12 cells. To facilitate spatial protein immobilization using microbial transglutaminase, the active lysine in the K domain of the substrate of an engineered transglutaminase (FXIIIa) was photocaged[114,115]. This photocaged substrate could be stochastically immobilized into PEG hydrogels and spatiotemporally activated upon exposure to UV light. Q peptides were engineered into proteins, such as VEGF and PDGF, and upon treatment with FXIIIa would be crosslinked into the activated regions of the hydrogel. Immobilized growth factors were found to drive human mesenchymal stem cell migration.

1.12.4 Farnesyltransferase

Farnesyltransferase recognizes the CA_1A_2X motif (where A_1 and A_2 are any aliphatic amino acid, and X is any amino acid) on the C-terminus of proteins and installs a farnesyl pyrophosphate group at the cysteine residue to form a thioether bond, a process called prenylation. Farnesyl analogs have been synthesized with a variety of functional handles, including aldehydes, azides, and alkynes[116,117]. Azide and alkyne containing analogs showed the best incorporation using farnesyltransferase, with kinetic constants comparable to incorporation of the native substrate[118]. Azide functionalized GFP was immobilized onto alkyne functionalized agarose beads and shown to retain its fluorescence. Further investigation revealed a benzaldehyde containing substrate that could be efficiently incorporated onto EGFP[119]. This benzaldehyde could then be used for hydrazone or oxime ligation to create a FRET pair between GFP and Texas Red or GFP and TAMRA, respectively. Aldehyde functionalized proteins could also be immobilized onto hydrazide beads and released through the addition of aminooxy-PEG, leading to a single step purification / PEGylation strategy.

Other prenyltransferases, such as geranylgeranyltransferase (GGTase), have also been used to enable orthogonal labeling strategies for dual site-specific protein labeling. The Distefano group found that farnesyltransferase best recognized the CVIA sequence, and GGTase recognized CVLL. GGTase was also found to incorporate bulkier and longer isoprenoids, enabling one pot labeling of GFP-CVLL and RFP-CVIA with an azide containing farnesyltransferase substrate and an alkyne containing GGTase substrate[120]. The two functionalized proteins could then be conjugated together through copper-catalyzed alkyne-azide cycloaddition reaction.

1.12.5 N-myristoyltransferase

N-myristoyltransferase (NMT) is a eukaryotic enzyme which installs a fatty acid group to the N-terminus of proteins for anchoring to the cell membrane. NMT has been shown to be promiscuous in substrate affinity, and azide-functionalized myristic acid analogs can be incorporated site-specifically when an N-terminal NMT recognition motif is appended to the protein of interest. Expression systems in *E. coli* have been developed which co-express the protein of interest with the NMT enzyme and a methionine aminopeptidase for excision of the N-terminal methionine and subsequent near quantitative functionalization[121]. As *E. coli* does not naturally utilize the NMT enzyme, NMT functionalization was found to be orthogonal to the *E. coli* proteome. Myristic acid has been installed onto an elastin-like polypeptide to generate temperature responsive liposomes for doxorubicin and paclitaxel encapsulation[122].

NMT has also been used to great success to make modified proteins for incorporation into biomaterials. One example used both NMT and the SrtA to functionalize the N- and C-terminus of the light responsive protein LOV2 with azides[123]. The dual-functionalized LOV2 was then used to crosslink a PEG hydrogel, such that the light-driven conformational change of LOV2 could be used to change the bulk mechanical properties of the gel. In another example, an azide was installed onto the N-terminus of another light responsive protein, PhoCl, which was expressed as a fusion with other proteins. PhoCl is an engineered fluorescent protein from *Clavularia sp.* that undergoes intramolecular bond cleavage upon exposure to violet light[124]. After incorporation into PEG hydrogels, exposure to violet light causes cleavage in the PhoCl domain, liberating the protein of interest. This allowed for the facile creation of protein patterned biomaterials.

1.12.6 Phosphopantetheinyl Transferase

Coenzyme A (CoA) analogs can be efficiently labeled onto the serine residue of the DSLEFIASKLA (ybbR) tag via Phosphopantetheinyl transferase (Sfp). This tag can be inserted

into internal loops of the protein, as well as the N- or C-terminus, and allows for efficient incorporation of diverse CoA analogs containing biotin, fluorescein, TAMRA, Texas Red and other compounds[125]. By attaching CoA functionalized PEG to a PEG acrylamide resin, luciferase and glutathione-S-transferase were covalently immobilized directly from cell lysate and retained high activity[126]. Sfb has also been used to modify proteins quantitatively with biotin for immobilization on streptavidin surfaces for phage display binding experiments[127]. However, the Gaub group found that Sfb was not able to immobilize TagGFP2 onto a CoA modified biolayer interferometer surface where SrtA was able to functionalize polyglycine modified surfaces[128]. Thus careful consideration of sterics should be taken into account when choosing a chemoenzymatic method for protein modification.

1.12.7 Tubulin Tyrosine Ligase

Tubulin tyrosine ligase (TTL) is an enzyme which is naturally involved in microtubial maintenance[129,130]. It recognizes the "Tub-Tag" (VDSVEGEGEEEGEE) and appends tyrosine analogs including azide and aldehyde functional handles to the C-terminus[131]. Nanobodies, GFP, and ubiquitin were all modified with high conversion (>99%). Azide-containing proteins could be further functionalized using click chemistry for super-resolution microscopy and immunoprecipitation experiments. Further investigation of the substrate tolerance of TTL has revealed further promiscuity in substrates, including biotin and fluorophore containing tyrosine analogs[132]. This allowed for single step protein functionalization with coumarin, and TTL coumarin labeled Annexin V was found to be as active as a commercial Alexa 350 conjugated Annexin V probe.

1.12.8 Lipoic Acid Ligase

Lipoic acid ligase (LplA) recognizes the lipoic acceptor peptide (LAP) sequence (GFEIDKVWYDLDA) and conjugates a lipoic acid group to the side chain of the lysine residue. Pioneered by the Ting group for fluorophore conjugation, lipoic acid ligase has been shown to accept other substrates[133]. The LAP tag can be added to an internal loop or the N- or C-terminus of the protein of interest, and treatment with LplA proceeds rapidly and near quantitatively. Further investigation into substrate specificity revealed substrates with clickable handles for subsequent conjugation.

Azide groups were installed to the N-terminus and two internal sites of GFP using an azide containing lipoic acid analog, 10-azidodecanoic acid[134]. Near complete conversion of LAP-GFP was achieved after an hour incubation with LplA and 60 fold molar excess of substrate. These sites were then PEGylated with an alkyne PEG to give a multiple site-specifically PEGylated protein. Use of a mutant LplA allowed for the incorporation of a stable norbornene derivative for subsequent reaction with tetrazine containing compounds[135]. The technique has been used for cell surface labeling of a LAP-containing transmembrane protein[136].

1.12.9 Biotin Ligase

Biotin ligase (BirA) is an *E. coli* derived enzyme which appends a biotin group to the ε-amine of lysine in an ATP dependent reaction. Biotin ligase's natural substrate is the biotin carboxyl carrier protein (BCCP), a relatively large domain of 75 amino acids. The binding of BirA to BCCP has been used to immobilize immunoglobulin G (IGG) to gold surfaces, however the long length of BCCP may be problematic for some proteins[137]. A more promiscuous biotin ligase has also been developed, BioD2, which when fused to a bait protein allows for biotinylation of proximate proteins for investigation into protein-protein interactions[138]. However for site-specific labeling, phage display has revealed a short peptide substrate for localized biotinylation[139]. Biotin

labeling of full-length proteins has been achieved by adding the AviTag (GLNDIFEAQKIEWHE) to a termini or internal loop[140]. The biotin-streptavidin bond is one of the most stable non-covalent bonds, making it a promising strategy for immobilization within streptavidin containing materials.

In one exquisite example exploiting the specificity and orthogonality of this non-covalent interaction, the Stoichet lab labeled ciliary neutrophic factor with the AviTag for biotinylation, and sonic hedgehog was labeled with barnstar, which strongly interacts with its binding partner barnase[141]. Streptavidin and barnase were subsequently immobilized into agarose gels using two-photon lithography. Upon swelling in of biotinylated and barnstar labeled proteins, exquisite 3D control over protein immobilization was achieved. These immobilized proteins were found to be able to guide migration of adult neural precursor cells within the hydrogels. The Leipzig lab added an N-terminal AviTag to interferon-γ, platelet derived growth factor-AA, and BMP-2 for site-specific biotinylation of growth factors and subsequent immobilization into streptavidin functionalized methacrylamide chitosan gels[142]. They found that the immobilized growth factors led to more cells differentiating into the expected phenotype as compared to adsorbed growth factor treatments.

1.12.10 Formylglycine Generating Enzyme

Unlike many chemical and enzymatic modifications that append bulky substituents to a protein, Formylglycine generating enzyme (FGE) enables the direct oxidation of a cysteine residue to an aldehyde containing formylglycine[143]. Aldehydes are convenient tags for bioconjugation via oxime or hydrazine ligation[144]. FGE was found to selectively modify the cysteine in the LCTPSR motif which can be incorporated in any accessible region of the protein[145]. FGE was used to modify a haloacid dehalogenase enzyme ST2570 for immobilization into amine-functionalized flow reactors for conversion of L-2-haloalkanoic[146]. The immobilized enzyme was found to have a

higher catalytic activity and thermal stability than the free enzyme. FGE is a promising, though relatively unexplored, enzyme for materials applications. This may be in part explained by the limited functionality of the enzyme. While other conversion techniques have broad substrate specificity and are able to append many functional groups, FGE is limited to relatively simple aldehyde modification.

1.12.11 Butelase 1 Ligase

Butelase 1 is a recently discovered Asx peptide ligase which is able to append C-terminal NHV or DHV sequences to nucleophilic N-terminal domains[147]. Butelase has very high activity, roughly 20,000 times faster than SrtA, and has been used for peptide ligation and protein cyclization[148]. One potential downside of this technique is that the enzyme can also act as a peptidase, although careful control of the reaction conditions can limit hydrolyzed product to <5%. While a rapid and efficient enzyme, the promiscuity towards the N-terminus means that the N-terminus of the NHV or DHV containing domain must be blocked or otherwise hindered to prevent cyclization when performing peptide ligation. One possible solution is to engineer an N-terminal proline residue, which is unable to act as a nucleophile for ligation. While butelase 1 has not been used in a materials setting as of this publication, there is significant interest in creating cyclic peptides and proteins, which often have unique therapeutic effects[149].

1.13 Conclusion and Future Outlook

While this review has highlighted many newly developed chemistries which offer unparalleled control over protein functionalization, there is no one-size-fits-all solution for a given application. For the modification of wild-type proteins, labeling of endogenous residues or a termini may prove amenable if sufficient conversion and selectivity can be achieved. Genetic

engineering approaches, either by genetic code expansion or chemoenzymatic strategies, enable exquisite selectivity on otherwise intractable protein domains. Additional consideration must be paid when attempting multiple modifications per protein, whether it be to create protein crosslinks or to add even more functionalities to an enzyme. Ultimately, combinations of several modification strategies may be necessary for engineering advanced materials.

While peptide screening and phage display can often prove useful screens for identifying new site-selective modification chemistries, prediction of the effects of the complex tertiary structure of full-length proteins remains challenging. Increasing computational abilities will accelerate the development of new techniques and allow for more accurate predictions on their efficacy in modifying a given protein[150,151].

Even with the wide range of strategies for protein modification, applying these conjugation schemes to make the next generation of protein-functionalized materials is still largely unexplored. Incorporation of functional stimuli-responsive proteins into hydrogels allows for the creation of "smart" biomaterials, materials which can be engineered to undergo physiochemical changes upon exposure to a stimuli[152]. While there exist thousands of stimuli-responsive proteins, only a few have been used in materials. Calmodulin, a protein which undergoes a conformational change to bind calcium ions, has been used to create hydrogels whose mechanical properties are a function of calcium concentration[153]. The field of optogenetics has diligently been identifying, engineering, and utilizing a host of light responsive proteins to spatiotemporally control cell fate[154,155,156]. Use of these light responsive proteins in a material complex would enable unprecedented fully reversible spatiotemporal control over chemical and physical properties.

In summary, there exist a wide range of techniques for the modification of proteins. There is a trade-off between "off the shelf" labeling of endogenous proteins and the specificity of

reaction. While chemoselective methods like amine and thiol modification have historically been used with great success in creating functionalized materials, they have typically only been applied to robust proteins or in applications where retaining even a fraction of activity is tolerable. Protein engineering and chemoenzymatic approaches by contrast tend to require in house knowledge of protein expression techniques, as commercial sources for proteins encoding the correct tags are rarely available. However, these approaches enable unprecedented control over the site of modification, enabling the creation of protein based materials with novel functionalities that would be impossible to replicate with small molecules. The protein modification field has seen exponential growth in the number of techniques available, and the time is ripe for applying these methodologies to create the next generation of biomaterials.

1.14 References

1. DeForest, C. A. & Anseth, K. S. Advances in Bioactive Hydrogels to Probe and Direct Cell Fate. *Annu. Rev. Chem. Biomol. Eng.* **3,** 421–444 (2012).

2. Gilbert, S. F. The Generation of Dorsal-Ventral Polarity. in *Developmental Biology. 6th edition* (Sinauer Associates, 2000).

3. The bicoid Protein Determines Position in the Drosophila Embryo in a Concentration-Dependent Manner. *Cell* **54,** 95–104 (1988).

4. Ueno, S. *et al.* Biphasic role for Wnt/beta-catenin signaling in cardiac specification in zebrafish and embryonic stem cells. *Proc. Natl. Acad. Sci. U. S. A.* **104,** 9685–90 (2007).

5. Hydrogel: Preparation, characterization, and applications: A review. *J. Adv. Res.* **6,** 105–121 (2015).

6. Ratner, B. D. & Bryant, S. J. Biomaterials: Where We Have Been and Where We Are Going. *Annu. Rev. Biomed. Eng.* **6,** 41–75 (2004).

7. Dhandayuthapani, B., Yoshida, Y., Maekawa, T. & Kumar, D. S. Polymeric Scaffolds in Tissue Engineering Application: A Review. *Int. J. Polym. Sci.* **2011,** 1–19 (2011).

8. Antoine, E. E., Vlachos, P. P. & Rylander, M. N. Review of collagen I hydrogels for bioengineered tissue microenvironments: characterization of mechanics, structure, and transport. *Tissue Eng. Part B. Rev.* **20,** 683–96 (2014).

9. Caliari, S. R. & Burdick, J. A. A practical guide to hydrogels for cell culture. *Nat. Methods* **13,** 405–414 (2016).

10. Langer, R. & Tirrell, D. A. Designing materials for biology and medicine. *Nature* **428,** 487–492 (2004).

11. Zhu, J. Bioactive modification of poly(ethylene glycol) hydrogels for tissue engineering. *Biomaterials* **31,** 4639–56 (2010).

12. Lee, K. Y. & Mooney*, D. J. Hydrogels for Tissue Engineering. (2001). doi:10.1021/CR000108X

13. Hern, D. L. & Hubbell, J. A. Incorporation of adhesion peptides into nonadhesive hydrogels useful for tissue resurfacing. *J. Biomed. Mater. Res.* **39,** 266–276 (1998).

14. Lutolf, M. P. *et al.* Synthetic matrix metalloproteinase-sensitive hydrogels for the conduction of tissue regeneration: Engineering cell-invasion characteristics.

15. Kolb, H. C., Finn, M. G. & Sharpless, K. B. Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angew. Chemie Int. Ed.* **40,** 2004–2021 (2001).

16. Sletten, E. M. & Bertozzi, C. R. From mechanism to mouse: a tale of two bioorthogonal reactions. *Acc. Chem. Res.* **44,** 666–76 (2011).

17. Dommerholt, J. *et al.* Readily Accessible Bicyclononynes for Bioorthogonal Labeling and Three-Dimensional Imaging of Living Cells. *Angew. Chemie Int. Ed.* **49,** 9422–9425 (2010).

18. Kloxin, A. M., Tibbitt, M. W. & Anseth, K. S. Synthesis of photodegradable hydrogels as dynamically tunable cell culture platforms. *Nat. Protoc.* **5,** 1867 (2010).

19. Hoffmann, J. C. & West, J. L. Three-dimensional photolithographic patterning of multiple bioactive ligands in poly(ethylene glycol) hydrogels. doi:10.1039/c0sm00140f

20. Wosnick, J. H. & Shoichet, M. S. Three-dimensional Chemical Patterning of Transparent Hydrogels. doi:10.1021/cm071158m

21. DeForest, C. A. & Tirrell, D. A. A photoreversible protein-patterning approach for guiding stem cell fate in three-dimensional gels. *Nat. Mater.* **14,** 523–531 (2015).

22. Wöll, D. *et al.* Intramolecular Sensitization of Photocleavage of the Photolabile 2-(2-Nitrophenyl)propoxycarbonyl (NPPOC) Protecting Group: Photoproducts and Photokinetics of the Release of Nucleosides. *Chem. - A Eur. J.* **14,** 6490–6497 (2008).

23. Ulrich, S., Boturyn, D., Marra, A., Renaudet, O. & Dumy, P. Oxime Ligation: A Chemoselective Click-Type Reaction for Accessing Multifunctional Biomolecular Constructs. *Chem. - A Eur. J.* **20,** 34–41 (2014).

24. Lagassé, H. A. D. *et al.* Recent advances in (therapeutic protein) drug development. *F1000Research* **6,** 113 (2017).

25. Leader, B., Baca, Q. J. & Golan, D. E. Protein therapeutics: a summary and pharmacological classification. *Nat Rev Drug Discov* **7,** 21–39 (2008).

26. Fisher, S. A., Baker, A. E. G. & Shoichet, M. S. Designing Peptide and Protein Modified Hydrogels: Selecting the Optimal Conjugation Strategy. *J. Am. Chem. Soc.* **139,** 7416–7427 (2017).

27. Coons, A. H., Creech, J., Norman, R. & Berliner, E. The Demonstration of Pneumococcal Antigen in Tissues by the Use of Fluorescent Antibody 1. *J. Immunol.* **45,** (1942).

28. Engvall, E. & Perlmann, P. Enzyme-linked immunosorbent assay (ELISA) quantitative assay of immunoglobulin G. *Immunochemistry* **8,** 871–874 (1971).

29. Tamura, T. & Hamachi, I. Chemistry for Covalent Modification of Endogenous/Native Proteins: From Test Tubes to Complex Biological Systems. *J. Am. Chem. Soc.* **141,** 2782–2799 (2019).

30. Spicer, C. D. & Davis, B. G. Selective chemical protein modification. *Nat. Commun.* **5,** 4740 (2014).

31. Anderson, G. W., Zimmerman, J. E. & Callahan, F. M. The Use of Esters of N-Hydroxysuccinimide in Peptide Synthesis. *J. Am. Chem. Soc.* **86,** 1839–1842 (1964).

32. Fisher, S. A. *et al.* Photo-immobilized EGF chemical gradients differentially impact breast cancer cell invasion and drug response in defined 3D hydrogels. *Biomaterials* **178,** 751–766 (2018).

33. Dodson, G. & Wlodawer, A. Catalytic triads and their relatives. *Trends Biochem. Sci.* **23,** 347–52 (1998).

34. Nanna, A. R. *et al.* Harnessing a catalytic lysine residue for the one-step preparation of homogeneous antibody-drug conjugates. *Nat. Commun.* **8,** 1112 (2017).

35. Matos, M. J. *et al.* Chemo- and Regioselective Lysine Modification on Native Proteins. *J. Am. Chem. Soc.* **140,** 4004–4017 (2018).

36. Chilamari, M., Purushottam, L. & Rai, V. Site-Selective Labeling of Native Proteins by a Multicomponent Approach. *Chem. - A Eur. J.* **23,** 3819–3823 (2017).

37. Yoshitake, S. *et al.* Mild and Efficient Conjugation of Rabbit Fab' and Horseradish Peroxidase Using a Maleimide Compound and Its Use for Enzyme Immunoassay. *J. Biochem.* **92,** 1413–1424 (1982).

38. Fontaine, S. D., Reid, R., Robinson, L., Ashley, G. W. & Santi, D. V. Long-Term Stabilization of Maleimide–Thiol Conjugates. *Bioconjug. Chem.* **26,** 145–152 (2015).

39. Boutureira, O. *et al.* Site-Selective Modification of Proteins with Oxetanes. *Chemistry* **23,** 6483–6489 (2017).

40. Kolodych, S. *et al.* CBTF: New Amine-to-Thiol Coupling Reagent for Preparation of Antibody Conjugates with Increased Plasma Stability. *Bioconjug. Chem.* **26,** 197–200 (2015).

41. Shaunak, S. *et al.* Site-specific PEGylation of native disulfide bonds in therapeutic proteins. *Nat. Chem. Biol.* **2,** 312–313 (2006).

42. Lee, M. T. W., Maruani, A., Baker, J. R., Caddick, S. & Chudasama, V. Next-generation disulfide stapling: reduction and functional re-bridging all in one. *Chem. Sci.* **7,** 799–802 (2016).

43. Griebenow, N., Dilmaç, A. M., Greven, S. & Bräse, S. Site-Specific Conjugation of Peptides and Proteins via Rebridging of Disulfide Bonds Using the Thiol–Yne Coupling Reaction. *Bioconjug. Chem.* **27,** 911–917 (2016).

44. Wang, T. *et al.* Water-soluble allyl sulfones for dual site-specific labelling of proteins and cyclic peptides. *Chem. Sci.* **7,** 3234 (2016).

45. Ruddle, B., Fleming, R., Wu, H., Gao, C. & Dimasi, N. Characterization of Disulfide Bond Rebridged Fab-Drug Conjugates Prepared Using a Dual Maleimide Pyrrolobenzodiazepine Cytotoxic Payload. *ChemMedChem* (2019). doi:10.1002/CMDC.201900077

46. Grim, J. C. *et al.* A Reversible and Repeatable Thiol–Ene Bioconjugationfor Dynamic Patterning of Signaling Proteins in Hydrogels. *ACS Cent. Sci.* **4,** 909 (2018).

47. Hörner, M. *et al.* Phytochrome-Based Extracellular Matrix with Reversibly Tunable Mechanical Properties. *Adv. Mater.* **31,** 1806727 (2019).

48. Williams, C. G., Malik, A. N., Kim, T. K., Manson, P. N. & Elisseeff, J. H. Variable cytocompatibility of six cell lines with photoinitiators used for polymerizing hydrogels and cell encapsulation. *Biomaterials* **26,** 1211–1218 (2005).

49. Imiołek, M. *et al.* Selective Radical Trifluoromethylation of Native Residues in Proteins. *J. Am. Chem. Soc.* **140,** 1568–1571 (2018).

50. Yu, Y. *et al.* Chemoselective Peptide Modification via Photocatalytic Tryptophan β-Position Conjugation. *J. Am. Chem. Soc.* **140,** 6797–6800 (2018).

51. Rosen, C. B. & Francis, M. B. Targeting the N terminus for site-selective protein modification. *Nat. Chem. Biol.* **13,** 697–705 (2017).

52. Obermeyer, A. C., Jarman, J. B. & Francis, M. B. N-Terminal Modification of Proteins with *o* -Aminophenols. *J. Am. Chem. Soc.* **136,** 9572–9579 (2014).

53. Chen, D., Disotuar, M. M., Xiong, X., Wang, Y. & Chou, D. H.-C. Selective N-terminal functionalization of native peptides and proteins. *Chem. Sci.* **8,** 2717–2722 (2017).

54. Li, D. *et al.* N-terminal α-amino group modification of antibodies using a site-selective click chemistry method. *MAbs* **10,** 712–719 (2018).

55. Lee, J. P., Kassianidou, E., MacDonald, J. I., Francis, M. B. & Kumar, S. N-terminal specific conjugation of extracellular matrix proteins to 2-pyridinecarboxaldehyde functionalized polyacrylamide hydrogels. *Biomaterials* **102,** 268–76 (2016).

56. Gilmore, J. M., Scheck, R. A., Esser-Kahn, A. P., Joshi, N. S. & Francis, M. B. N-Terminal Protein Modification through a Biomimetic Transamination Reaction. *Angew. Chemie Int. Ed.* **45,** 5307–5311 (2006).

57. Witus, L. S. & Francis, M. Site-Specific Protein Bioconjugation via a Pyridoxal 5′-Phosphate-Mediated N-Terminal Transamination Reaction. *Curr. Protoc. Chem. Biol.* **2,** 125–134 (2010).

58. Christman, K. L., Broyer, R. M., Tolstyka, Z. P. & Maynard, H. D. Site-specific protein immobilization through N-terminal oxime linkages. *J. Mater. Chem.* **17,** 2021 (2007).

59. Conibear, A. C., Watson, E. E., Payne, R. J. & Becker, C. F. W. Native chemical ligation

in protein synthesis and semi-synthesis. *Chem. Soc. Rev.* **47,** 9046–9068 (2018).

60. Dhall, A., Weller, C. E., Chu, A., Shelton, P. M. M. & Chatterjee, C. Chemically Sumoylated Histone H4 Stimulates Intranucleosomal Demethylation by the LSD1–CoREST Complex. *ACS Chem. Biol.* **12,** 2275–2280 (2017).

61. Bandyopadhyay, A., Cambray, S. & Gao, J. Fast and selective labeling of N-terminal cysteines at neutral pH via thiazolidino boronate formation. *Chem. Sci.* **7,** 4589–4593 (2016).

62. Marino, G., Eckhard, U. & Overall, C. M. Protein Termini and Their Modifications Revealed by Positional Proteomics. (2015). doi:10.1021/acschembio.5b00189

63. Dieterich, D. C. *et al.* Labeling, detection and identification of newly synthesized proteomes with bioorthogonal non-canonical amino-acid tagging. *Nat. Protoc.* **2,** 532–540 (2007).

64. tom Dieck, S. *et al.* Metabolic Labeling with Noncanonical Amino Acids and Visualization by Chemoselective Fluorescent Tagging. in *Current Protocols in Cell Biology* **Chapter 7,** Unit7.11 (John Wiley & Sons, Inc., 2012).

65. Zhang, X., Lu, W., Kwan, K., Bhattacharyya, D. & Wei, Y. Dual-Functional-Tag-Facilitated Protein Labeling and Immobilization. (2017). doi:10.1021/acsomega.6b00512

66. Xiao, H., Xuan, W., Shao, S., Liu, T. & Schultz, P. G. Genetic Incorporation of ε-N-2-Hydroxyisobutyryl-lysine into Recombinant Histones. (2015). doi:10.1021/cb501055h

67. Luo, X. *et al.* Genetically encoding phosphotyrosine and its nonhydrolyzable analog in bacteria. *Nat. Chem. Biol.* **13,** 845–849 (2017).

68. Wang, Z. A. *et al.* A Genetically Encoded Allysine for the Synthesis of Proteins with Site-Specific Lysine Dimethylation. *Angew. Chem. Int. Ed. Engl.* **56,** 212–216 (2017).

69. Liu, J., Hemphill, J., Samanta, S., Tsang, M. & Deiters, A. Genetic Code Expansion in Zebrafish Embryos and Its Application to Optical Control of Cell Signaling. *J. Am. Chem. Soc.* **139,** 9100–9103 (2017).

70. Nguyen, D. P. *et al.* Genetic Encoding of Photocaged Cysteine Allows Photoactivation of TEV Protease in Live Mammalian Cells. *J. Am. Chem. Soc.* **136,** 2240–2243 (2014).

71. Yang, T., Li, X.-M., Bao, X., Fung, Y. M. E. & Li, X. D. Photo-lysine captures proteins that bind lysine post-translational modifications. *Nat. Chem. Biol.* **12,** 70–72 (2016).

72. Tomlin, F. M. *et al.* Site-specific incorporation of quadricyclane into a protein and photocleavage of the quadricyclane ligation adduct. *Bioorg. Med. Chem.* **26,** 5280–5290 (2018).

73. Xuan, W., Shao, S. & Schultz, P. G. Protein Crosslinking by Genetically Encoded Noncanonical Amino Acids with Reactive Aryl Carbamate Side Chains. *Angew. Chemie Int. Ed.* **56,** 5096–5100 (2017).

74. Yamaguchi, A. *et al.* Incorporation of a Doubly Functionalized Synthetic Amino Acid into Proteins for Creating Chemical and Light-Induced Conjugates. *Bioconjug. Chem.* **27,** 198–206 (2016).

75. Hoppmann, C. *et al.* Genetically Encoding Photoswitchable Click Amino Acids in *Escherichia coli* and Mammalian Cells. *Angew. Chemie Int. Ed.* **53,** 3932–3936 (2014).

76. Yang, A. *et al.* A chemical biology route to site-specific authentic protein modifications. *Science* **354,** 623–626 (2016).

77. Lotze, J., Reinhardt, U., Seitz, O. & Beck-Sickinger, A. G. Peptide-tags for site-specific protein labelling in vitro and in vivo. *Mol. Biosyst.* **12,** 1731–1745 (2016).

78. Block, H. *et al.* Immobilized-Metal Affinity Chromatography (IMAC). in *Methods in enzymology* **463,** 439–473 (2009).

79. Martos-Maldonado, M. C. *et al.* Selective N-terminal acylation of peptides and proteins with a Gly-His tag sequence. *Nat. Commun.* **9,** 3307 (2018).

80. Sunbul, M., Nacheva, L. & Jäschke, A. Proximity-Induced Covalent Labeling of Proteins with a Reactive Fluorophore-Binding Peptide Tag. *Bioconjug. Chem.* **26,** 1466–1469 (2015).

81. Zhang, C., Dai, P., Vinogradov, A. A., Gates, Z. P. & Pentelute, B. L. Site-Selective Cysteine-Cyclooctyne Conjugation. *Angew. Chemie Int. Ed.* **57,** 6459–6463 (2018).

82. Zhang, C. *et al.* π-Clamp-mediated cysteine conjugation. *Nat. Chem.* **8,** 120–128 (2016).

83. Agrawalla, B. K. *et al.* Chemoselective Dual Labeling of Native and Recombinant Proteins. *Bioconjug. Chem.* **29,** 29–34 (2018).

84. Peciak, K., Laurine, E., Tommasi, R., Choi, J. & Brocchini, S. Site-selective protein conjugation at histidine. *Chem. Sci.* **10,** 427–439 (2019).

85. Cole, N. B. Site-specific protein labeling with SNAP-tags. *Curr. Protoc. protein Sci.* **73,** Unit 30.1 (2013).

86. Wasserberg, D. *et al.* Immobilization of Ferrocene-Modified SNAP-Fusion Proteins. *Int. J. Mol. Sci.* **14,** 4066–80 (2013).

87. Los, G. V. *et al.* HaloTag: A Novel Protein Labeling Technology for Cell Imaging and Protein Analysis. *ACS Chem. Biol.* **3,** 373–382 (2008).

88. Kwon, Y., Han, Z., Karatan, E., Mrksich, M. & Kay, B. K. Antibody Arrays Prepared by Cutinase-Mediated Immobilization on Self-Assembled Monolayers. *Anal. Chem.* **76,** 5713–5720 (2004).

89. Zakeri, B. *et al.* Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. *Proc. Natl. Acad. Sci. U. S. A.* **109,** E690-7 (2012).

90. Fierer, J. O., Veggiani, G. & Howarth, M. SpyLigase peptide–peptide ligation polymerizes affibodies to enhance magnetic cancer cell capture. *Proc. Natl. Acad. Sci.* **111,** E1176–E1181 (2014).

91. Sun, F., Zhang, W.-B., Mahdavi, A., Arnold, F. H. & Tirrell, D. A. Synthesis of bioactive protein hydrogels by genetically encoded SpyTag-SpyCatcher chemistry. *Proc. Natl. Acad. Sci. U. S. A.* **111,** 11269–74 (2014).

92. Gao, X., Fang, J., Xue, B., Fu, L. & Li, H. Engineering Protein Hydrogels Using SpyCatcher-SpyTag Chemistry. *Biomacromolecules* **17,** 39 (2016).

93. Lim, S., Jung, G. A., Muckom, R. J., Glover, D. J. & Clark, D. S. Engineering bioorthogonal protein–polymer hybrid hydrogel as a functional protein immobilization platform. *Chem. Commun.* **55,** 806–809 (2019).

94. Terpe, K. Overview of tag protein fusions: from molecular and biochemical fundamentals

to commercial systems. *Appl. Microbiol. Biotechnol.* **60,** 523–533 (2003).

95. Zhang, Y., Park, K.-Y., Suazo, K. F. & Distefano, M. D. Recent progress in enzymatic protein labelling techniques and their applications. *Chem. Soc. Rev.* **47,** 9106–9136 (2018).

96. Mazmanian, S. K., Liu, G., Ton-That, H. & Schneewind, O. Staphylococcus aureus sortase, an enzyme that anchors surface proteins to the cell wall. *Science* **285,** 760–3 (1999).

97. Antos, J. M., Truttmann, M. C. & Ploegh, H. L. Recent advances in sortase-catalyzed ligation methodology. *Curr. Opin. Struct. Biol.* **38,** 111–118 (2016).

98. Ham, H. O. *et al.* In situ regeneration of bioactive coatings enabled by an evolved Staphylococcus aureus sortase A. *Nat. Commun.* **7,** 11140 (2016).

99. Cambria, E. *et al.* Covalent Modification of Synthetic Hydrogels with Bioactive Proteins via Sortase-Mediated Ligation. *Biomacromolecules* **16,** 2316–2326 (2015).

100. Parthasarathy, R., Subramanian, S. & Boder, E. T. Sortase A as a Novel Molecular 'Stapler' for Sequence-Specific Protein Conjugation. *Bioconjug. Chem.* **18,** 469–476 (2007).

101. Warden-Rothman, R., Caturegli, I., Popik, V. & Tsourkas, A. Sortase-tag expressed protein ligation: combining protein purification and site-specific bioconjugation into a single step. *Anal. Chem.* **85,** 11090–7 (2013).

102. Shadish, J. A., Benuska, G. M. & DeForest, C. A. Bioactive site-specifically modified proteins for 4D patterning of gel biomaterials. *Nat. Mater.* 1 (2019). doi:10.1038/s41563-019-0367-7

103. McConnell, S. A. *et al.* Protein Labeling via a Specific Lysine-Isopeptide Bond Using the Pilin Polymerizing Sortase from *Corynebacterium diphtheriae*. *J. Am. Chem. Soc.* **140,** 8420–8423 (2018).

104. Chang, T. K., Jackson, D. Y., Burnier, J. P. & Wells, J. A. Subtiligase: a tool for semisynthesis of proteins. *Proc. Natl. Acad. Sci. U. S. A.* **91,** 12544–8 (1994).

105. Henager, S. H. *et al.* Enzyme-catalyzed expressed protein ligation. *Nat. Methods* **13,** 925–927 (2016).

106. Yoshihara, H. A. I., Mahrus, S. & Wells, J. A. Tags for labeling protein N-termini with

subtiligase for proteomics. *Bioorg. Med. Chem. Lett.* **18,** 6000–6003 (2008).

107. Wiita, A. P., Seaman, J. E. & Wells, J. A. Global Analysis of Cellular Proteolysis by Selective Enzymatic Labeling of Protein N-Termini. in *Methods in enzymology* **544,** 327–358 (2014).

108. Ando, H. *et al.* Purification and Characteristics of a Novel Transglutaminase Derived from Microorganisms. *Agric. Biol. Chem.* **53,** 2613–2617 (1989).

109. Strop, P. Versatility of Microbial Transglutaminase. *Bioconjug. Chem.* **25,** 855–862 (2014).

110. Oteng-Pabi, S. K., Clouthier, C. M. & Keillor, J. W. Design of a glutamine substrate tag enabling protein labelling mediated by Bacillus subtilis transglutaminase. *PLoS One* **13,** e0197956 (2018).

111. Yu, C.-M., Zhou, H., Zhang, W.-F., Yang, H.-M. & Tang, J.-B. Site-specific, covalent immobilization of BirA by microbial transglutaminase: A reusable biocatalyst for in vitro biotinylation. *Anal. Biochem.* **511,** 10–12 (2016).

112. Steffen, W. *et al.* Discovery of a microbial transglutaminase enabling highly site-specific labeling of proteins. *J. Biol. Chem.* **292,** 15622–15635 (2017).

113. Ranga, A., Lutolf, M. P., Hilborn, J. & Ossipov, D. A. Hyaluronic Acid Hydrogels Formed in Situ by Transglutaminase-Catalyzed Reaction. *Biomacromolecules* **17,** 1553–1560 (2016).

114. Mosiewicz, K. A. *et al.* In situ cell manipulation through enzymatic hydrogel photopatterning. *Nat. Mater.* **12,** 1072–1078 (2013).

115. Griffin, D. R. *et al.* Hybrid Photopatterned Enzymatic Reaction (HyPER) for in Situ Cell Manipulation. *ChemBioChem* **15,** 233–242 (2014).

116. Labadie, G.R., Viswanathan, R. & Poulter, C. D. Farnesyl Diphosphate Analogues with ω-Bioorthogonal Azide and Alkyne Functional Groups for Protein Farnesyl Transferase-Catalyzed Ligation Reactions. (2007). doi:10.1021/JO7017747

117. Jennings, B. C. *et al.* Analogs of farnesyl diphosphate alter CaaX substrate specificity and reactions rates of protein farnesyltransferase. *Bioorg. Med. Chem. Lett.* **26,** 1333–1336

(2016).

118. Duckworth, B.P., Xu, J., Taton, A., Guo, A., and & Distefano, M.D. Site-Specific, Covalent Attachment of Proteins to a Solid Surface. (2006). doi:10.1021/BC060125E

119. Rashidian, M., Song, J. M., Pricer, R. E. & Distefano, M. D. Chemoenzymatic Reversible Immobilization and Labeling of Proteins without Prior Purification. *J. Am. Chem. Soc.* **134,** 8455–8467 (2012).

120. Zhang, Y. *et al.* Simultaneous Site-Specific Dual Protein Labeling Using Protein Prenyltransferases. *Bioconjug. Chem.* **26,** 2542–2553 (2015).

121. Kulkarni, C., Kinzer-Ursem, T. L. & Tirrell, D. A. Selective Functionalization of the Protein N Terminus with N-Myristoyl Transferase for Bioconjugation in Cell Lysate. *ChemBioChem* **14,** 1958–1962 (2013).

122. Luginbuhl, K. M. *et al.* Recombinant Synthesis of Hybrid Lipid-Peptide Polymer Fusions that Self-Assemble and Encapsulate Hydrophobic Drugs. *Angew. Chemie Int. Ed.* **56,** 13979–13984 (2017).

123. Liu, L. *et al.* Cyclic Stiffness Modulation of Cell-Laden Protein–Polymer Hydrogels in Response to User-Specified Stimuli Including Light. *Adv. Biosyst.* **2,** 1800240 (2018).

124. Zhang, W. *et al.* Optogenetic control with a photocleavable protein, PhoCl. *Nat Meth* **14,** 391–394 (2017).

125. Yin, J. *et al.* Genetically encoded short peptide tag for versatile protein labeling by Sfp phosphopantetheinyl transferase. *Proc. Natl. Acad. Sci.* **102,** 15815–15820 (2005).

126. Wong, L. S., Thirlway, J. & Micklefield, J. Direct Site-Selective Covalent Protein Immobilization Catalyzed by a Phosphopantetheinyl Transferase. *J. Am. Chem. Soc.* **130,** 12456–12464 (2008).

127. Zhao, B. *et al.* Phage Selection Assisted by Sfp Phosphopantetheinyl Transferase-Catalyzed Site-Specific Protein Labeling. in *Methods in molecular biology (Clifton, N.J.)* **1266,** 161–170 (2015).

128. Ott, W., Durner, E. & Gaub, H. E. Enzyme-Mediated, Site-Specific Protein Coupling

Strategies for Surface-Based Binding Assays. *Angew. Chemie Int. Ed.* **57,** 12666–12669 (2018).

129. Ersfeld, K. *et al.* Characterization of the tubulin-tyrosine ligase. *J. Cell Biol.* **120,** 725–732 (1993).

130. Janke, C. The tubulin code: molecular components, readout mechanisms, and functions. *J. Cell Biol.* **206,** 461–72 (2014).

131. Schumacher, D. *et al.* Versatile and Efficient Site-Specific Protein Functionalization by Tubulin Tyrosine Ligase. *Angew. Chemie Int. Ed.* **54,** 13787–13791 (2015).

132. Schumacher, D. *et al.* Broad substrate tolerance of tubulin tyrosine ligase enables one-step site-specific enzymatic protein labeling. *Chem. Sci.* **8,** 3471–3478 (2017).

133. Uttamapinant, C. *et al.* A fluorophore ligase for site-specific protein labeling inside living cells. *Proc. Natl. Acad. Sci. U. S. A.* **107,** 10914–9 (2010).

134. Plaks, J. G., Falatach, R., Kastantin, M., Berberich, J. A. & Kaar, J. L. Multisite Clickable Modification of Proteins Using Lipoic Acid Ligase. *Bioconjug. Chem.* **26,** 1104–1112 (2015).

135. Best, M., Degen, A., Baalmann, M., Schmidt, T. T. & Wombacher, R. Two-Step Protein Labeling by Using Lipoic Acid Ligase with Norbornene Substrates and Subsequent Inverse-Electron Demand Diels-Alder Reaction. *ChemBioChem* **16,** 1158–1162 (2015).

136. Baalmann, M., Best, M. & Wombacher, R. Site-Specific Protein Labeling Utilizing Lipoic Acid Ligase (LplA) and Bioorthogonal Inverse Electron Demand Diels-Alder Reaction. in *Methods in molecular biology (Clifton, N.J.)* **1728,** 365–387 (2018).

137. Miyao, H., Ikeda, Y., Shiraishi, A., Kawakami, Y. & Sueda, S. Immobilization of immunoglobulin-G-binding domain of Protein A on a gold surface modified with biotin ligase. *Anal. Biochem.* **484,** 113–121 (2015).

138. Kim, D. I. *et al.* An improved smaller biotin ligase for BioID proximity labeling. *Mol. Biol. Cell* **27,** 1188–1196 (2016).

139. Beckett, D., Kovaleva, E. & Schatz, P. J. A minimal peptide substrate in biotin holoenzyme

synthetase-catalyzed biotinylation. *Protein Sci.* **8,** 921–929 (2008).

140. Fairhead, M. & Howarth, M. Site-specific biotinylation of purified proteins using BirA. *Methods Mol. Biol.* **1266,** 171–84 (2015).

141. Wylie, R. G. *et al.* Spatially controlled simultaneous patterning of multiple growth factors in three-dimensional hydrogels. *Nat. Mater.* **10,** 799–806 (2011).

142. Li, H., Koenig, A. M., Sloan, P. & Leipzig, N. D. In vivo assessment of guided neural stem cell differentiation in growth factor immobilized chitosan-based hydrogel scaffolds. *Biomaterials* **35,** 9049–9057 (2014).

143. Dierks, T. *et al.* Multiple sulfatase deficiency is caused by mutations in the gene encoding the human C(alpha)-formylglycine generating enzyme. *Cell* **113,** 435–44 (2003).

144. Kölmel, D. K. & Kool, E. T. Oximes and Hydrazones in Bioconjugation: Mechanism and Catalysis. *Chem. Rev.* **117,** 10358–10376 (2017).

145. Carlson, B. L. *et al.* Function and structure of a prokaryotic formylglycine-generating enzyme. *J. Biol. Chem.* **283,** 20117–25 (2008).

146. Jian, H., Wang, Y., Bai, Y., Li, R. & Gao, R. Site-Specific, Covalent Immobilization of Dehalogenase ST2570 Catalyzed by Formylglycine-Generating Enzymes and Its Application in Batch and Semi-Continuous Flow Reactors. *Molecules* **21,** (2016).

147. Nguyen, G. K. T. *et al.* Butelase 1 is an Asx-specific ligase enabling peptide macrocyclization and synthesis. *Nat. Chem. Biol.* **10,** 732–738 (2014).

148. Nguyen, G. K. T. *et al.* Butelase 1: A Versatile Ligase for Peptide and Protein Macrocyclization. *J. Am. Chem. Soc.* **137,** 15398–15401 (2015).

149. Joo, S. H. Cyclic peptides as therapeutic agents and biochemical tools. *Biomol. Ther. (Seoul).* **20,** 19–26 (2012).

150. Strumillo, M. & Beltrao, P. Towards the computational design of protein post-translational regulation. *Bioorg. Med. Chem.* **23,** 2877–2882 (2015).

151. Narayanam, M. K., Liang, Y., Houk, K. N. & Murphy, J. M. Discovery of new mutually orthogonal bioorthogonal cycloaddition pairs through computational screening. *Chem. Sci.*

7, 1257–1261 (2016).

152. Kopecek, J. Hydrogel biomaterials: a smart future? *Biomaterials* **28,** 5185–92 (2007).

153. Ehrick, J. D. *et al.* Genetically engineered protein in hydrogels tailors stimuli-responsive characteristics. *Nat. Mater.* **4,** 298–302 (2005).

154. Zhou, X. X., Chung, H. K., Lam, A. J. & Lin, M. Z. Optical control of protein activity by fluorescent protein domains. *Science* **338,** 810–4 (2012).

155. Levskaya, A., Weiner, O. D., Lim, W. A. & Voigt, C. A. Spatiotemporal Control of Cell Signalling Using A Light-Switchable Protein Interaction. *Nature* **461,** 997–1001 (2009).

156. Wang, H. *et al.* LOVTRAP: an optogenetic system for photoinduced protein dissociation. *Nat. Methods* **13,** 755–8 (2016).

Chapter 2

Thesis Objectives

2.1 Overview

After a thorough review of the literature, two chemoenzymatic strategies stood out as particularly promising for biomaterial and photolithographic applications. The Sortase A enzyme acts as a molecular stapler, recognizing the LPETG sequence of amino acids and through a transpeptidation reaction installing an N-terminal GGG-tagged peptide onto the C-terminus of the protein, a process termed sortagging[1]. With advances in recombinant DNA technology, designing proteins to contain the arbitrary amino acid sequences and expressing them in hosts such as *E. coli* has become a powerful and widely employed technique[2]. In combination with powerful solid-phase peptide synthesis techniques, it becomes relatively straightforward to express LPETG-containing proteins and GGG-functionalized peptides containing a host of synthetic chemistries, including clickable handles and light-sensitive linkers[3]. By expressing each protein as a fusion with the sortase enzyme, the process is further simplified to allow purification and conjugation to occur in a single step, creating homogenous solutions of site-specifically modified proteins[4].

Using a light-sensitive protein in lieu of an exogenously installed photoreactive group enables light responsiveness to be genetically encoded. By using a newly engineered photocleavable protein (PhoCl) in combination with NMT for azide installation, photoreleasable proteins can be co-translationally produced in *E. coli*[5]. As an added benefit, PhoCl is a green fluorescent protein which allows for convenient tracking of the recombinant fusion protein. PhoCl also undergoes photoscission under visible violet light, which has deeper tissue penetrance than the UV light used in most small molecule photochemistries. Using a protein engineering and

expression strategy instead of organic synthesis could make photopatterning of bioactive proteins more accessible and scalable without specialty equipment.

For this thesis work, I propose to use the specificity of chemoenzymatic modification in combination with the power of synthetic chemistry and solid-phase peptide synthesis to achieve following aims.

2.2 Thesis Aims

Aims:

1. **Utilize chemoenzymatic methods for the site-specific modification of proteins while retaining their activity**
2. **Demonstrate reversible photochemical patterning of biomaterials using the Sortase enzyme to append synthetic peptides containing clickable handles and photosensitive moieties**
3. **Create a fully genetically encoded photocleavable linker using the PhoCl protein and the NMT enzyme to install an N-terminal azide for conjugation into hydrogels**

2.3 References

1. Parthasarathy, R., Subramanian, S. & Boder, E. T. Sortase A as a Novel Molecular 'Stapler' for Sequence-Specific Protein Conjugation. *Bioconjug. Chem.* **18,** 469–476 (2007).
2. Baneyx, F. Recombinant protein expression in Escherichia coli. *Curr. Opin. Biotechnol.* **10,** 411–421 (1999).
3. Behrendt, R., White, P. & Offer, J. Advances in Fmoc solid-phase peptide synthesis. *J. Pept. Sci.* **22,** 4–27 (2016).
4. Warden-Rothman, R., Caturegli, I., Popik, V. & Tsourkas, A. Sortase-tag expressed protein ligation: combining protein purification and site-specific bioconjugation into a single step. *Anal. Chem.* **85,** 11090–7 (2013).
5. Zhang, W. *et al.* Optogenetic control with a photocleavable protein, PhoCl. *Nat. Methods* **14,** 391–394 (2017).

Chapter 3

Bioactive Site-Specifically Modified Proteins for 4D Patterning of Gel Biomaterials

As published: Shadish, J. A.; Benuska, G. M.; DeForest, C. A. *Nat. Mater.* **2019**.
DOI: 10.1038/s41563-019-0367-7

3.1 Abstract:

Protein-modified biomaterials can be used to modulate cellular function in 3D. However, as dynamic heterogeneous control over complex cell physiology continues to be sought, strategies that permit reversible and user-defined tethering of fragile proteins to materials remain in great need. Here, we introduce a modular and robust semisynthetic approach to reversibly pattern cell-laden hydrogels with site-specifically modified proteins. Exploiting a versatile sortase-mediated transpeptidation, we generate a diverse library of homogenous, singly functionalized proteins with bioorthogonal reactive handles for biomaterial modification. We demonstrate the photoreversible immobilization of fluorescent proteins, enzymes, and growth factors to gels with excellent spatiotemporal resolution while retaining native protein bioactivity. Dynamic regulation of proliferation, intracellular mitogen-activated protein kinase signaling, and subcellularly resolved receptor endocytosis is accomplished through localized epidermal growth factor presentation. Our method broadly permits modification and patterning of a wide range of proteins, providing newfound avenues to probe and direct advanced cellular fates in 4D.

3.2 Main text:

The extracellular matrix (ECM) directs cell function through a complex choreography of biomacromolecular interactions in a tissue-dependent manner. Far from static, this hierarchical milieu of biochemical and biophysical cues presented within the native cellular niche is both spatially complex and constantly changing[1–3]. As these pericellular reconfigurations are vital for tissue morphogenesis, disease regulation, and healing, *in vitro* culture platforms that recapitulate such dynamic environmental phenomena would be invaluable for fundamental studies in cell biology, as well as in the engineering of functional human tissue[4,5]. Polymer-based hydrogels represent one such emerging and highly attractive class of *in vitro* cell culture constructs that offer promise for dynamic 3D biological studies[6–8]. Their high water content, tissue-like elasticity, and facile transport of nutrients and waste render them ideal mimics of the cell's ECM, while their optical clarity permits imaging of cell function to be performed non-destructively[9–11]. Gel materials can be produced under mild, cytocompatible conditions that enable live-cell encapsulation and are readily formulated to contain user-defined chemical functionalities, mechanical properties, and degradability[12–14]. Through appropriately developed chemistries, such niche characteristics can be modified on-demand to probe and direct cell fate in response to variable microenvironmental cues similar to those comprising the native ECM[15–17].

To emulate dynamic ECM biochemical heterogeneity *in vitro*, significant effort has been dedicated towards the creation of an expansive library of chemical strategies for hydrogel alteration. Preliminary efforts in this regard have largely focused on the exploitation of photochemical techniques to pattern bioactive small molecules and peptides spatially within synthetic hydrogel culture systems[18–23]. While such approaches have proven successful in directing relatively simple

cellular functions (such as adhesion and spreading), the ability to regulate more complex and dynamic decisions of fate using full-length proteins remains of prime interest[24–27]. Though proteins represent a powerful tool in the quest to govern cell physiology, their fragility necessitates that careful consideration be given to the chemistries employed as well as the precise site of protein modification for material tethering to ensure sustained stability and activity[28].

Installation of reactive groups onto proteins required for biomaterial decoration has been performed almost exclusively through non-specific reactions with thiols and primary amines on endogenous cysteine and lysine residues[29]. As these amino acids are not uniquely present on native proteins, bioconjugation of this type occurs randomly and stochastically to yield a heterogeneous collection of differently modified species. Such uncontrolled functionalization often leads to protein unfolding, loss of activity, and supraphysiologic doses required to elicit cellular response. Coupled with batch-to-batch variability and ambiguous extents of modification, reproducibility problems dramatically hinder laboratory and clinical translation of these ill-defined samples. Generalizable strategies to create homogeneous protein populations that can be used to modify a wide variety of biomaterial platforms while retaining native levels of bioactivity remain an open challenge.

Chemoenzymatic strategies have gained recent popularity for their unique ability to modify proteins in a manner that is site-specific, selective, and quantitative, thereby permitting the introduction of reactive chemical handles at defined locations[30–34]. We identified the sortase-mediated transpeptidation modification strategy[34,35] as one that could permit installation of

functional handles compatible with virtually any chemistry used for biomaterial formation or modification, including photoreactions that have previously proven beneficial for 4D biomaterial decoration. Sortase A, a calcium-assisted transpeptidase from *Staphylococcus aureus*, recognizes the sorting signal "LPXTG" (where X is any amino acid) and forms an acyl-enzyme intermediate with the sorting signal's threonine residue, simultaneously displacing other C-terminal amino acids[36]. The resultant thioester intermediate can be nucleophilically displaced by reaction with the N-terminal amine of a polyglycine probe, thereby conjugating a synthetic peptide onto the C-terminus while regenerating the sortase A enzyme. Critically, the polyglycine compound may contain non-natural functionality, providing a route to "sortag" a wide variety of bioorthogonal and reactive moieties site-specifically onto proteins of interest. We hypothesized that these selectively monofunctionalized species would retain significantly higher activity than those generated through random modification, and that sortagged species would likely open the door to precisely regulate advanced cellular functions throughout biomaterials that have proven inaccessible *via* conventional strategies.

To facilitate one-step protein biofunctionalization and purification, we implemented sortagging through the recently developed Sortase-Tag Enhanced Protein Ligation (STEPL) technique[37] (Figure 1a). In STEPL, the protein of interest, sorting sequence LPETG, $(GGS)_5$ flexible linker, sortase A, and a 6xHis-Tag are fused into a single protein construct which is recombinantly expressed. The flexible $(GGS)_5$ linker allows intramolecular sortagging through the encoded LPETG motif and the fused sortase domain. The sortase-LPETG intermediate is displaced by the addition of calcium and a customizable probe with an N-terminal polyglycine moiety, ligating the protein of interest to the engineered peptide while simultaneously separating it from the remaining

6xHis-functionalized sortase A. This step can be performed during immobilized metal ion affinity chromatography, where sortase A remains bound to the nickel-nitrilotriacetic acid (Ni-NTA) column, allowing for site-specific labeling and purification of proteins in a single step.

To highlight the versatility of sortase for protein modification and to generate a library of functional biomacromolecules for hydrogel modification, we constructed STEPL expression systems for six proteins spanning three distinct classes: fluorescent [Enhanced Green Fluorescent Protein (EGFP), mCherry, and mCerulean], enzymatic [β-lactamase (bla)], and growth factor [epidermal growth factor (EGF) and fibroblast growth factor (FGF)] (Supplementary Methods). As protein functional modification by sortase is defined by engineered peptide identity, we synthesized and exploited five distinct polyglycine probes (Figure 1b, Supplementary Methods) each with different reactive handles (azides, aromatic aldehydes, nitrobenzyl moieties). These bioorthogonal handles enable proteins to be conjugated to materials formed by a variety of common gel formation click chemistries (such as azide/alkyne, oxime, and hydrazine conjugation), as well as dictate how functionalized proteins interface with materials over time. With these six expression vectors and five sortaggable probes in hand, STEPL was utilized to generate all 30 possible protein-peptide conjugates (Supplementary Methods and Supplementary Table 1). Whole-protein mass spectrometry and gel shift assays revealed exceptionally high protein purity and quantitative functionalization in all cases (Figure 1c-h, Supplementary Figures 1-2 and Methods), indicating efficient generation of a diverse collection of singly modified proteins for gel patterning. This represents the largest library of sortagged proteins created for any application to date.

Having successfully generated site-specifically modified proteins bearing functionality required for gel decoration, we sought to compare the effects of sortagged modification versus stochastic N-hydroxysuccinimide (NHS) ester labeling of solvent-accessible amines, the most commonly employed method to modify proteins, using a representative protein from each class (Figure 2a-i). EGFP activity was evaluated directly through fluorescence measurements. Activity of bla was determined through a standard colorimetric assay involving hydrolysis of the chromogenic cephalosporin nitrocefin. EGF activity was quantified using HeLa cells expressing EKAREV, a Förster resonance energy transfer (FRET) reporter for mitogen-activated protein kinase (MAPK) signaling[38]. In all cases, azide-modified proteins sortagged with H-GGGGDDK(N_3)-NH_2 exhibited activity that was statistically indistinguishable from that of the native species. In contrast, protein bioactivity decreased significantly when azide tagging occurred through conventional NHS ester labeling involving 2,5-dioxopyrrolidin-1-yl 4-azidobutanoate (N_3-OSu). We note that large molar excesses of the NHS-activated compound relative to the protein of interest are required for even modest levels of protein labeling (Supplementary Figure 3), as NHS ester hydrolysis represents the dominant reaction pathway when this chemistry is performed in aqueous systems[39]. We also note that individual proteins tolerate different amounts of random modification, as higher excesses of the NHS species were required to diminish EGFP fluorescence (a protein that has been engineered for its stable structure) than bla enzyme activity. Taken together, these experiments demonstrate that sortagging allows for the creation of modified protein conjugates with higher purity and bioactivity than their NHS-modified counterparts.

Having observed that sortagging outperformed NHS chemistry when introducing reactive azides onto EGFP, bla, and EGF, we set out to quantify the bioactivity of all 30 protein-peptide library

members (Figure 2j-o). Similar to EGFP, mCherry and mCerulean activity was evaluated directly through fluorescence measurements. FGF and EGF activity were determined based on their ability to stimulate cell proliferation, quantified by increased dsDNA synthesis and content. In total, the overwhelming majority (27/30) of modified proteins exhibit bioactivity that is statistically indistinguishable from the unmodified species. Moreover, the least active conjugate (bla-*o*NB-CHO, based on the kinetically perfect bla enzyme that had previously proven highly sensitive to modification) retained 76 ± 4% of its native bioactivity. We expect the ability to singly modify functionally diverse bioactive proteins with several different reactive handles, each of which influences how they form the basis of and interact with materials, to be hugely enabling for a variety of biological applications. Additionally, the relative ease and lack of specialized equipment required for protein expression/purification, coupled with the commercial availability of sortaggable peptides containing functional handles [including azides, alkynes, (meth)acrylates, thiols, aldehydes, maleimides, allyl sulfides], render these strategies practically accessible to non-specialists.

After verifying that sortagged proteins retained native bioactivity, we sought to demonstrate their controlled incorporation within biomaterials. Exploiting strategies previously introduced by our laboratory[27], we utilized model hydrogels formed through strain-promoted azide-alkyne cycloaddition (SPAAC) between a four-arm poly(ethylene glycol) (PEG) tetrabicyclononyne and a linear PEG diazide (Supplementary Methods). Reaction kinetics and bioorthogonality permit encapsulation of live cells with high viability, while the bioinert PEG provides a non-fouling "blank slate" for subsequent decoration with proteins. The optical clarity of these materials renders

them useful for photochemical modification with biomolecules, as well as for microscopy-based assays of encapsulated cell fate.

A cytocompatible photomediated oxime ligation[27,40] was exploited to immobilize site-specifically modified proteins within hydrogels containing a 2-(2-nitrophenyl)propyloxycarbonyl (NPPOC)-photocaged alkoxyamine with spatiotemporal control (Figure 3ab). Upon mild near-UV irradiation (λ = 365 nm), NPPOC is cleaved, liberating the reactive alkoxyamine and permitting localized condensation with gel-swollen aromatic aldehyde-modified proteins to form a stable oxime linkage. Removal of unbound proteins by diffusion yields patterned gel substrates defined by user-selected light exposure locations and parameters. Gel patterning was performed with EGFP sortagged with H-GGGGDDK(CHO)-NH_2 (EGFP-CHO), whereby immobilized protein fluorescence permitted visualization and quantification. Traditional photolithographic techniques were utilized to control patterning of mask-defined shapes throughout gel thickness (Figure 3c), and to generate continuous gradients that followed a dose-dependent response predicted by NPPOC photocleavage kinetics (Figure 3df, Supplementary Figure 4). EGFP-CHO was immobilized within hydrogels with excellent 3D control through multiphoton laser-scanning lithography, whereby programmed laser rastering within the gel material dictated protein tethering location and concentration (Figure 3e, Supplementary Figure 5). These experiments represent the first demonstration of 3D patterned covalent immobilization of a site-specifically modified protein within a material.

After having demonstrated the ability to immobilize sortagged proteins within biomaterials with excellent 3D control, we shifted our efforts towards the patterned photorelease of such species from gels (Figure 3g). For this, we exploited an *ortho*-nitrobenzyl ester (*o*NB) moiety[22,41] that undergoes rapid photoscission in response to cytocompatible near-UV light (λ = 365 nm) (Figure 3h). mCherry sortagged with H-GGGGDDK(*o*NB-N_3)-NH_2 (mCherry-*o*NB-N_3) was tethered uniformly throughout materials during SPAAC-based gelation. After directed light exposure, released protein was diffused from the gel prior to sample imaging and quantification through fluorescent confocal microscopy. Mask-based photolithographic techniques were again exploited to control mCherry-*o*NB-N_3 removal and concentration throughout full gel thickness in a predictable and dose-dependent manner (Figure 3ijl, Supplementary Figures 6-7). Multiphoton lithography afforded excellent 3D patterning at user-specified regions within gels (Figure 3k, Supplementary Figure 8). These results represent the first photorelease of a site-specifically modified protein from a 2D or 3D material. In addition to controlling the location of immobilized proteins that persist within a gel, the fully defined and homogenous protein population photoreleased from the material may have significant implications for controlled drug delivery.

Sortase's versatility in introducing diverse functional handles onto different proteins of interest enables protein-patterned materials of unprecedented complexity to be photoevolved in 4D. To demonstrate dynamic material patterning, we selected three differentially modified and spectrally separated fluorescent proteins from our library of sortagged species. Gels containing the NPPOC-caged alkoxyamine were uniformly functionalized with mCherry-*o*NB-N_3. Upon directed light exposure, mCherry photorelease was performed in concert with photomediated ligation of an aldehyde-tagged protein (either mCerulean-CHO or EGFP-*o*NB-CHO) to create interconnected

biochemical patterns. The photorelease/oxime ligation sequence was repeated, enabling a third protein (either EGFP-*o*NB-CHO or mCerulean-CHO) to be immobilized (Figure 4a). This approach was utilized to create an Escher-inspired tessellation of fish/bird shapes through masked lithography (Figure 4b-e, Supplementary Figure 9), as well as a 3D schematic depiction of a cell binding to an ECM-presented ligand *via* multiphoton patterning (Figure 4f-i, Supplementary Figure 10 and Supplementary Movies 1-2). Collectively, this data demonstrates dynamic control over multiple factors in time and space with unmatched precision, as well as the first 4D regulation of site-specifically modified proteins within materials.

Encouraged that mCherry, mCerulean, and EGFP each retained their fluorescence and spatial patterning for several weeks following hydrogel tethering, we sought to determine whether we could confine enzymatic activity to user-specified gel subvolumes. Utilizing bla-*o*NB-N_3, we functionalized and patterned gels through mask-based protein photoremoval. Protein-patterned gels were treated with 7-thiophenylacetamido-3-thioacetoxymethyl-3-cephem-4-carboxylate[42] (thioacetate cefalotin, Supplementary Methods), a thiocephalosporin that eliminates a proton and a thiolate ion following β-lactam enzymatic hydrolysis that reduce a water-soluble yellow phenazine into a green water-insoluble precipitate (Figure 5a-b). Phase contrast microscopy indicated changes in localized refractive index, corresponding to enzyme-induced precipitation confined to bla-modified regions within the gel (Figure 5c, Supplementary Figure 11). We believe this to be the first successful example of confining enzymatic activity to photopatterned locations within a material, a task enabled by site-specific modification and that is likely to prove useful in enzymatically regulating critical ECM parameters.

Having demonstrated that sortagged enzymes remain active when photopatterned within gels and that gel formation and modification chemistries were cytocompatible (Supplementary Figure 12), we tested the ability of these techniques to direct local cell function using immobilized growth factors. HeLa cells transfected with the EKAREV FRET reporter for MAPK activation[38] were encapsulated in materials uniformly functionalized with EGF-*o*NB-N_3. Following complete bulk protein photorelease, cells displayed basal levels of MAPK activation; those where EGF remained immobilized elicited significantly higher (~2 fold) intracellular signaling (Figure 6d-e). The extent of FRET responses for cells in 3D was consistent with 2D studies involving soluble EGF (Supplementary Figures 13 and Supplementary Movie 3) and is not directly affected by light treatments used for gel photomodulation (Supplementary Figure 14). Taking advantage of the spatial control afforded by protein photorelease, gels with patterned EGF were created through masked exposure; MAPK activation remained high in gel subvolumes still functionalized with EGF, but returned to basal levels in those where the protein had been photoreleased. Since MAPK signaling plays a key role in regulating cell proliferation, migration, and differentiation, its patterned activation represents a powerful step towards regulating advanced cell fates with spatiotemporal control. We believe that these results are the first example of photopatterned regulation over a specific biochemical pathway in a 3D material.

An added benefit of using STEPL to modify proteins for material decoration is its compatibility with many standard protein engineering strategies. Though the photoremovable EGF proved successful in spatially regulating MAPK activation in 3D gels, its location cannot be directly visualized during experimentation. To address this limitation, we constructed an expression vector for an EGFP-EGF fusion and performed STEPL using H-GGGGDDK(*o*NB-N_3)-NH_2 to generate

EGFP-EGF-*o*NB-N_3 (Supplementary Methods, Supplementary Figure 15). Taking advantage of its fluorescence, bioactivity, tetherability, and photoreleasability, we first used this tetrafunctional protein to pattern proliferation through dynamic EGF stimulation (Figure 6ab). Epidermal A431 cells were seeded on hydrogels containing interspaced lines of EGFP-EGF-*o*NB-N_3. The immobilized EGFP-EGF remained effective in driving proliferation, where cell patterning was visibly evident by Day 3 (Supplementary Figure 16); functionalized regions yielded ~2-fold increased cell densities relative to those lacking EGF by Day 6. Control experiments with patterned EGFP exhibited homogenous surface coverage with no significant difference in cell localization based on the underlying fluorescent protein pattern at any observed time point (Supplementary Figure 17). To assay the effects of dynamically patterned EGF stimulation on cell function, a parallel experiment was constructed in which the EGFP-EGF was photoreleased from one half of each individual line-patterned gel three days into culture. Though the persistently functionalized gel portions exhibited the expected patterning in cell density, uniform surface coverage was observed across the photoreleased gel portions where presumably cells had redistributed through migration to eliminate any previously patterned heterogeneity (Supplementary Figure 18). 3D experiments involving encapsulated HeLa cells yielded similar results, where spheroid growth was enhanced with persistently patterned EGFP-EGF (Supplementary Figures 19-21). Through reversibly controlled presentation of bioactive growth factors, we demonstrate newfound ability to modulate local proliferation/migration of cells in 2D and 3D.

Armed with the EGFP-EGF-*o*NB-N_3 tetrafunctional chimera, we sought to demonstrate that photoreleased proteins could influence dynamic biological fate with resolutions previously inaccessible to biomaterials-based approaches. EGF binds cell surface protein epidermal growth

factor receptor (EGFR), yielding an activated transmembrane protein that undergoes homodimerization. Canonical EGFR activation is coupled with membrane endocytosis, receptor trafficking, and downstream signal transduction. Though matrix-bound EGF binds EGFR, growth factor tethering inhibits ligand-receptor internalization, endocytotic trafficking, and canonical signaling[43,44]. We hypothesized that photoliberation of soluble EGFR-EGF could be used to "turn on" canonical EGFR activation locally within culture, and that downstream endosome formation could be visualized with fluorescent microscopy (Figure 6c, Supplementary Figure 22). To test this hypothesis, A431 cells were encapsulated in gels modified with EGFP-EGF-*o*NB-N_3. Confocal imaging revealed a fluorescent halo initially outlining each cell, indicating ligand-receptor binding and the presence of concentrated membrane-bound EGFP-EGF (Figure 6d). Multiphoton laser scanning lithography was used to selectively release the EGFP-EGF protein on one side of an individual encapsulated cell, triggering EGFP-EGF untethering and local endocytic vesicle formation, visible in <5 minutes and concentrated in the regions of light exposure (Figure 6ef). Cells within the same imaging window (~50 μm away) were unaffected by the photoreleased EGF, further indicating that activation could be specified with single micron-scale precision. Control studies involving photoreleasable EGFP did not yield endocytosis (Supplementary Figure 23). We believe this to be the first demonstration of using dynamic materials to govern 3D fate with single cell and/or subcellular resolution.

The results described here introduce and highlight the sortase-mediated transpeptidation as a uniquely powerful strategy to create a diverse library of homogenous, singly modified proteins with non-natural functionality. Sortagged proteins can be efficiently expressed and purified by STEPL, and exhibit near-native bioactivity while affording the ability to decorate biomaterials in

a user-dictated manner. By specifying reactive handle identity, functional proteins can be reversibly immobilized within hydrogels with excellent spatiotemporal resolution through a variety of bioorthogonal photochemical reactions. Such patterned control over niche biochemical properties with homogenous proteins permits unprecedented regulation of complex biological functions and cellular pathways. We expect that this approach will prove useful in probing the effects of ECM-presented cues on single cell fate and in the design of materials for tissue engineering, potentially opening the door to multi-lineage patterning of 3D stem cell differentiation.

3.3 References:

1. Lutolf, M. P., Gilbert, P. M. & Blau, H. M. Designing materials to direct stem-cell fate. *Nature* **462,** 433–441 (2009).

2. Tibbitt, M. W. & Anseth, K. S. Dynamic Microenvironments: The Fourth Dimension. *Sci. Transl. Med.* **4,** (2012).

3. Baker, B. M. & Chen, C. S. Deconstructing the third dimension – how 3D culture microenvironments alter cellular cues. *J. Cell Sci.* **125,** (2012).

4. DeForest, C. A. & Anseth, K. S. Advances in Bioactive Hydrogels to Probe and Direct Cell Fate. *Annu. Rev. Chem. Biomol. Eng.* **3,** 421–444 (2012).

5. Burdick, J. A. & Murphy, W. L. Moving from static to dynamic complexity in hydrogel design. *Nat. Commun.* **3,** 1269 (2012).

6. Langer, R. & Vacanti, J. P. Tissue Engineering. *Science* **260,** 920–926 (1993).

7. Lee, K. Y. & Mooney, D. J. Hydrogels for tissue engineering. *Chem. Rev.* **101,** 1869–1879 (2001).

8. Langer, R. & Tirrell, D. A. Designing materials for biology and medicine. *Nature* **428,** 487–492 (2004).

9. Cushing, M. C. & Anseth, K. S. Hydrogel cell cultures. *Science* **316,** 1133–1134 (2007).

10. Tibbitt, M. W. & Anseth, K. S. Hydrogels as Extracellular Matrix Mimics for 3D Cell Culture. *Biotechnol. Bioeng.* **103,** 655–663 (2009).

11. Seliktar, D. Designing Cell-Compatible Hydrogels for Biomedical Applications. *Science*

336, 1124–1128 (2012).

12. Lutolf, M. P. & Hubbell, J. A. Synthetic biomaterials as instructive extracellular microenvironments for morphogenesis in tissue engineering. *Nat. Biotechnol.* **23,** 47–55 (2005).

13. Zhang, Y. S. & Khademhosseini, A. Advances in engineering hydrogels. *Science* **356,** (2017).

14. Tam, R. Y., Smith, L. J. & Shoichet, M. S. Engineering Cellular Microenvironments with Photo- and Enzymatically Responsive Hydrogels: Toward Biomimetic 3D Cell Culture Models. *Acc. Chem. Res.* **50,** 703–713 (2017).

15. Caliari, S. R. & Burdick, J. A. A practical guide to hydrogels for cell culture. *Nat. Methods* **13,** 405–414 (2016).

16. Rosales, A. M. & Anseth, K. S. The design of reversible hydrogels to capture extracellular matrix dynamics. *Nat. Rev. Mater.* **1,** 15012 (2016).

17. Ruskowitz, E. R. & DeForest, C. A. Photoresponsive biomaterials for targeted drug delivery and 4D cell culture. *Nat. Rev. Mater.* **3,** 17087 (2018).

18. Luo, Y. & Shoichet, M. S. A photolabile hydrogel for guided three-dimensional cell growth and migration. *Nat. Mater.* **3,** 249–253 (2004).

19. Hahn, M. S., Miller, J. S. & West, J. L. Three-dimensional biochemical and biomechanical patterning of hydrogels for guiding cell behavior. *Adv. Mater.* **18,** 2679–2684 (2006).

20. DeForest, C. A., Polizzotti, B. D. & Anseth, K. S. Sequential click reactions for synthesizing and patterning three-dimensional cell microenvironments. *Nat. Mater.* **8,** 659–664 (2009).

21. DeForest, C. A., Sims, E. A. & Anseth, K. S. Peptide-functionalized click hydrogels with independently tunable mechanics and chemical functionality for 3D cell culture. *Chem. Mater.* **22,** 4783–4790 (2010).

22. DeForest, C. A. & Anseth, K. S. Cytocompatible click-based hydrogels with dynamically tunable properties through orthogonal photoconjugation and photocleavage reactions. *Nat. Chem.* **3,** 925–931 (2011).

23. DeForest, C. A. & Anseth, K. S. Photoreversible patterning of biomolecules within click-based hydrogels. *Angew. Chemie - Int. Ed.* **51,** 1816–1819 (2012).

24. Wylie, R. G., Ahsan, S., Aizawa, Y., Maxwell, K. L., Morshead, C. M. & Shoichet, M. S. Spatially controlled simultaneous patterning of multiple growth factors in three-dimensional hydrogels. *Nat. Mater.* **10,** 799–806 (2011).

25. Mosiewicz, K. A., Kolb, L., van der Vlies, A. J., Martino, M. M., Lienemann, P. S., Hubbell, J. A., Ehrbar, M., Lutolf, M. P., Vlies, A. J. van der, Martino, M. M., Lienemann, P. S., Hubbell, J. A., Ehrbar, M. & Lutolf, M. P. In situ cell manipulation through enzymatic hydrogel photopatterning. *Nat. Mater.* **12,** 1071–1077 (2013).

26. Griffin, D. R., Borrajo, J., Soon, A., Acosta-Velez, G. F., Oshita, V., Darling, N., Mack, J., Barker, T., Iruela-Arispe, M. L. & Segura, T. Hybrid Photopatterned Enzymatic Reaction (HyPER) for in Situ Cell Manipulation. *Chembiochem* **15,** 233–242 (2014).

27. DeForest, C. A. & Tirrell, D. A. A photoreversible protein-patterning approach for guiding stem cell fate in three-dimensional gels. *Nat. Mater.* **14,** 523–531 (2015).

28. Fisher, S. A., Baker, A. E. G. & Shoichet, M. S. Designing Peptide and Protein Modified

Hydrogels: Selecting the Optimal Conjugation Strategy. *J. Am. Chem. Soc.* **139,** 7416–7427 (2017).

29. Baslé, E., Joubert, N. & Pucheault, M. Protein Chemical Modification on Endogenous Amino Acids. *Chem. Biol.* **17,** 213–227 (2010).

30. Rabuka, D. Chemoenzymatic methods for site-specific protein modification. *Curr. Opin. Chem. Biol.* **14,** 790–796 (2010).

31. Rabuka, D., Rush, J. S., deHart, G. W., Wu, P. & Bertozzi, C. R. Site-specific chemical protein conjugation using genetically encoded aldehyde tags. *Nat. Protoc.* **7,** 1052–1067 (2012).

32. Kulkarni, C., Kinzer-Ursem, T. L. & Tirrell, D. A. Selective Functionalization of the Protein N Terminus with N-Myristoyl Transferase for Bioconjugation in Cell Lysate. *ChemBioChem* **14,** 1958–1962 (2013).

33. Chen, I., Howarth, M., Lin, W. & Ting, A. Y. Site-specific labeling of cell surface proteins with biophysical probes using biotin ligase. *Nat. Methods* **2,** 99–104 (2005).

34. Guimaraes, C. P., Witte, M. D., Theile, C. S., Bozkurt, G., Kundrat, L., Blom, A. E. M. & Ploegh, H. L. Site-specific C-terminal and internal loop labeling of proteins using sortase-mediated reactions. *Nat. Protoc.* **8,** 1787–1799 (2013).

35. Mao, H., Hart, S. A., Schink, A. & Pollok, B. A. Sortase-mediated protein ligation: a new method for protein engineering. *J. Am. Chem. Soc.* **126,** 2670–1 (2004).

36. Mazmanian, S. K., Liu, G., Hung, T. T. & Schneewind, O. Staphylococcus aureus sortase, an enzyme that anchors surface proteins to the cell wall. *Science* **285,** 760–763 (1999).

37. Warden-Rothman, R., Caturegli, I., Popik, V. & Tsourkas, A. Sortase-tag expressed protein ligation: Combining protein purification and site-specific bioconjugation into a single step. *Anal. Chem.* **85,** 11090–11097 (2013).

38. Komatsu, N., Aoki, K., Yamada, M., Yukinaga, H., Fujita, Y., Kamioka, Y. & Matsuda, M. Development of an optimized backbone of FRET biosensors for kinases and GTPases. *Mol. Biol. Cell* **22,** 4647–56 (2011).

39. Hermanson, G. T. *Bioconjugate Techniques*. (Academic Press, 2013).

40. Farahani, P. E., Adelmund, S. M., Shadish, J. A. & DeForest, C. A. Photomediated Oxime Ligation as a Bioorthogonal Tool for Spatiotemporally-Controlled Hydrogel Formation and Modification. *J. Mater. Chem. B* **5,** 4435–4442 (2017).

41. Kloxin, A. M., Kasko, A. M., Salinas, C. N. & Anseth, K. S. Photodegradable Hydrogels for Dynamic Tuning of Physical and Chemical Properties. *Science* **324,** 59–63 (2009).

42. Bieniarz, C., Young, D. F. & Cornwell, M. J. Chromogenic redox assay for beta-lactamases yielding water-insoluble products. I. Kinetic behavior and redox chemistry. *Anal. Biochem.* **207,** 321–8 (1992).

43. Kuhl, P. R. & Griffith-Cima, L. G. Tethered epidermal growth factor as a paradigm for growth factor–induced stimulation from the solid phase. *Nat. Med.* **2,** 1022–1027 (1996).

44. Fan, V. H., Au, A., Tamama, K., Littrell, R., Richardson, L. B., Wright, J. W., Wells, A. & Griffith, L. G. Tethered Epidermal Growth Factor Provides a Survival Advantage to Mesenchymal Stem Cells. *Stem Cells* **25,** 1241–1251 (2007).

Acknowledgements:

The authors recognize and thank Drs. R. Seifert and D. Hailey of the University of Washington Garvey Imaging Center for their ongoing support and advice, Drs. F. Watt and T. Hiratsuka (King's College London) for helpful discussion on visualizing MAPK activation, Dr. M. Matsuda (Kyoto University) for gifting the HeLa cells transfected with EKAREV FRET sensor, as well as Drs. R. Warden-Rothman and A. Tsourkis (University of Pennsylvania) for providing the pSTEPL plasmid[37]. The authors thank B. Badeau for assistance in synthesizing N_3-*o*NB-OSu, S. Adelmund for providing BCN-OSu, and A. Im for help with protein purification optimization. We gratefully acknowledge support from S. Edgar at the UW Mass Spectrometry Center as well as that from the NIH and N. Peters at the UW W. M. Keck Microscopy Center (S10 OD016240). This work was supported by a University of Washington Faculty Startup Grant (C.A.D.), a Jaconette L. Tietze Young Scientist Research Award (C.A.D.), and a CAREER Award (DMR 1652141, C.A.D.) from the National Science Foundation.

Author contributions:

For this manuscript, J.A.S. and C.A.D. conceived and designed the experiments; J.A.S. and G.M.B. performed the experiments; J.A.S. and C.A.D. analyzed the data and prepared the figures; J.A.S. and C.A.D. wrote the paper.

Competing financial interests:

The authors declare no competing financial interests.

Material and Data Availability:

Plasmids and data that support the findings of this study are available from the corresponding author upon reasonable request.

3.4 Methods:

3.4.1 Synthesis of polyglycine probes for sortagging. Polyglycine-containing peptides [H-GGGGDDK(N_3)-NH_2, H-GGGGDDK(CHO)-NH_2, H-GGGGDDK(*o*NB-N_3)-NH_2, H-GGGG-*o*NB-DDK(CHO)-NH_2] were synthesized through standard Fmoc solid-phase methodologies involving a butyloxycarbonyl-protected N-terminal glycine and a 4-methyltrity (Mtt)-protected C-terminal lysine reside. Following selective deprotection of Mtt (1% trifluoroacetic acid, TFA, in dichloromethane, DCM) on resin, azide (N_3), aldehyde (CHO), and photoreleasable azide (*o*NB-N_3) functionality were installed respectively by condensation with 4-azidobutanoic acid, 4-formylbenzoic acid, or 4-(4-(1-(4-azidobutanoyloxy)ethyl)-2-methoxy-5-nitrophenoxy)butanoic acid and the ε-amino group of the C-terminal lysine. Resin was washed (dimethylformamide, DMF, 3x; DCM, 3x) prior to peptide cleavage/deprotection (95:5 TFA:H_2O, 20 mL, 2 hr) and precipitation (diethyl ether, 180 mL, 0 °C, 2x). The crude peptide was purified *via* semi-preparative reversed-phase high-performance liquid chromatography using a 55-minute gradient (5–100% of acetonitrile and 0.1% TFA in H_2O) and lyophilized to give final product. Peptide purity was confirmed by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Complete synthetic details and characterization are given in Supplementary Methods.

3.4.2 Generation of homogenous protein-peptide conjugates by STEPL. STEPL plasmids for EGFP, mCherry, mCerulean, bla, EGF, FGF, and EGFP-EGF were constructed from pSTEPL[37] using standard cloning techniques (Supplementary Methods) and transformed into BL21(DE3) *E. coli* (Thermo Fisher). For protein expression, transformants were grown at 37 °C in lysogeny broth containing ampicillin (100 μg mL^{-1}) until an optical density of 0.6 (λ = 600 nm). Expression was induced with isopropyl β-D-1-thiogalactopyranoside (0.5 mM) and then agitated overnight at 18

°C. Cells were harvested *via* centrifugation and lysed by sonication. Clarified lysate was loaded onto HisPur Ni-NTA resin (ThermoFisher), which was washed (20 mM Tris, 50 mM NaCl, 20 mM imidazole) to remove unbound proteins. Following treatment of the resin with polyglycine probe of interest (20x, 4 h, 37 °C), sortagged proteins were eluted and purified by dialysis (molecular weight cut-off, MWCO ~10 kDa). Protein identity and purity was confirmed by liquid chromatography-tandem mass spectrometry, sodium dodecyl sulfate polyacrylamide gel electrophoresis, and gel shift analysis. Protein concentrations were determined by ultraviolet absorption (λ = 280 nm) and bicinchoninic acid assay prior to use. Complete experimental details are given in Supplementary Methods.

3.4.3 Random modification of proteins by NHS chemistry. Statistically modified proteins were prepared by reacting N_3-OSu (0x, 0 µM; 10x, 300 µM; 100x, 3 mM; 1000x, 30 mM) with protein (30 µM) in phosphate buffered saline (PBS, 900 µL) and DMF (100 µL) overnight at room temperature. Resulting protein solutions were purified with Zeba^{TM} Spin Desalting Columns (MWCO ~ 7 kDa, Thermo Fisher) and buffer-exchanged into fresh PBS. Protein concentrations were determined by ultraviolet absorption (λ = 280 nm) and bicinchoninic acid assay prior to use.

3.4.4 Activity determination of native and modified proteins. EGFP, mCherry, and mCerulean fluorescence was quantified on a fluorescent plate reader (EGFP: $\lambda_{excitation}$ = 470 nm, $\lambda_{emission}$ = 530 nm; mCherry: $\lambda_{excitation}$ = 575 nm, $\lambda_{emission}$ = 620 nm; mCerulean: $\lambda_{excitation}$ = 430 nm, $\lambda_{emission}$ = 475 nm). Bla enzymatic activity was measured through a chromogenic assay, whereby bla (1 ng) was incubated with nitrocefin (2 mM) in PBS (100 µL). Sample absorbance (λ_{abs} = 386 nm) was

measured over time at 37 °C; the initial slope was calculated as a measure of the k_{cat} of the enzyme. To determine EGF activity, HeLa cells expressing EKAREV[38] were plated in a 96-well plate (5,000 cells well^{-1}) and allowed to attach overnight in Dulbecco's Modified Eagle Medium (DMEM) (Life Technologies) supplemented with fetal bovine serum (FBS, 10%). 16 hours prior to imaging, cells were transitioned to serum-free DMEM. Time-series images (20 sec interval) were acquired on a fluorescent confocal microscope ($\lambda_{excitation}$ = 405 nm, 1% argon laser power; $\lambda_{emission,CFP}$ = 450-484 nm, $\lambda_{emission,YFP}$ = 520-524 nm) following stimulation with EGF (12.5 nM) in serum-free DMEM. The YFP/CFP FRET response ratio was calculated for each frame (ImageJ) and plotted over time. The relative activity of each EGF species was taken as the initial rate of change of the FRET response after EGF stimulation, normalized to native EGF. Alternatively, EGF activity was determined through a commercially available DNA synthesis assay (Click-iT EdU Alexa Fluor 488, Thermo Fisher). NIH3T3 cells were placed in a 96-well plate (5,000 cells well^{-1}) and cultured overnight in DMEM containing FBS (10%). Cells were washed with PBS (2x) and transferred to serum-free DMEM. After one day, cells were transitioned to serum-free DMEM containing 5-Ethynyl-2'-deoxyuridine (10 μM) with or without EGF (100 ng mL^{-1}). After two days of culture, cells were fixed in formaldehyde and processed following kit manufacturer recommendations. The relative activity of each EGF species was taken as the percentage of cells actively synthesizing DNA after EGF stimulation, normalized to unmodified EGF. FGF activity was determined through a proliferation assay using a commercially available assay (Quant-iT PicoGreen, Thermo Fisher) for dsDNA content. NIH3T3 cells (5,000 cells well^{-1}) were cultured overnight in 96-well plate in DMEM containing FBS (10%) prior to being transferred to serum-free DMEM. After one day, cells were transferred to serum-free DMEM with or without FGF (100 ng mL^{-1}), incubated for two days, and assayed following manufacturer protocol. The relative

activity of each FGF species corresponds to the total dsDNA content, normalized to unmodified FGF.

3.4.5 Formation of SPAAC-based hydrogels. A solution of PEG-tetraBCN (M_n ~ 20,000 Da, 4 mM), N_3-PEG-N_3 (M_n ~ 3,500 Da, 8 mM, Supplementary Methods), and N_3-TEG-ONH-NPPOC (0.1 mM) was prepared in PBS. Network formation was allowed to proceed for 1 hour between Rain-X®-treated glass slides with silicone rubber spacers (McMaster-Carr, 0.5 mm thick). The slides were separated, and gels were equilibrated overnight in PBS prior to use.

3.4.6 Photopatterning conditions for protein tethering and release. For photolithographic patterning, hydrogels were exposed to collimated UV light (λ = 365 nm, 10 mW cm^{-2}, 0 – 1200 s) through a patterned chrome photomask (Photo Sciences) using a Lumen Dynamics OmniCure S1500 Spot UV Curing system equipped with an internal 365 nm band-pass filter and a second in-line 360 nm cut-on longpass filter. Protein gradients were created by moving an opaque plate attached to a programmable linear motion stage over gels during light exposure at various rates (0.3 – 2.4 mm min^{-1}). For 3D patterning experiments within gels, an Olympus FV1000 MPE BX61 Multi-photon Microscope with a 20x objective was used. ROIs were scanned 16 times with pulsed laser light (λ = 740 nm, 50% laser power) with a 2.5 μm z-interval to generate 3D patterns. After NPPOC cleavage, gels were incubated for 8 hours with aldehyde-tagged protein (100 μM in PBS) at room temperature and protected from light. To remove unbound protein, gels were gently agitated in PBS (16 h) prior to imaging. Following light exposure, gels containing photoreleasable proteins were incubated in PBS (16 h) prior to analysis. This process was repeated iteratively for

patterning of multiple proteins. Experiments involving patterning of fluorescent proteins were visualized by fluorescent confocal or multiphoton microscopy, while those involving immobilized bla were imaged by phase-contrast microscopy. For experiments involving encapsulated mammalian cells, PBS was replaced with culture media during gel formation and patterning.

3.4.7 Spatial assessment of bla enzymatic activity. SPAAC-based gels uniformly functionalized with bla-*o*NB-N_3 (30 μM) were subjected to masked light (λ = 365 nm, 10 mW cm^{-2}, 10 min). Protein-patterned gels were incubated with thioacetate cefalotin (5 mM, Supplementary Methods) and phenazine methosulfate (6.5 mM) in PBS (37 °C, 1 h). The gels were then washed with PBS (3x, 15 min) prior to visualization by phase contrast microscopy.

3.4.8 Encapsulation of MAPK FRET reporter cells. HeLa cells stably transfected with the EKAREV FRET sensor[38] were expanded in high-glucose DMEM supplemented with FBS (10%) and penicillin/streptomycin (1%) in a 5% CO_2 atmosphere at 37 °C. Culture medium was changed every two to three days, and cells were passaged at ~70% confluency. HeLa cells were encapsulated (2 x 10^6 cells mL^{-1}) in SPAAC-based gels (10 μL) modified with EGF-*o*NB-N_3 (12.5 nM) and an N_3-GRGDS-NH_2 peptide (Supplementary Methods, 100 μM) for cell attachment. Photopatterning was performed 24 hours after encapsulation. Prior to FRET imaging (16 h), culture media was replaced with serum-free DMEM containing penicillin/streptomycin (1%).

3.4.9 Visualization of patterned MAPK signaling. HeLa EKAREV-laden hydrogels were imaged using a Leica SP8X confocal microscope at a constant distance below gel surface (100

μm). EKAREV was excited using an argon laser (λ = 405 nm, 10% laser power); CFP excitation was monitored from $\lambda_{emission,CFP}$ = 450-484 nm, while YFP fluorescence was measured from $\lambda_{emission,YFP}$ = 520-524 nm. After image acquisition, background fluorescence was subtracted, and the YFP/CFP ratio for each pixel of the image was calculated using ImageJ. FRET response ratios were visualized as a blue-green-red color map.

3.4.10 Cell response to dynamically patterned EGFP-EGF. EGFP-EGF-*o*NB-N_3 (25 μM) was conjugated (25 °C, 1 hr) to PEG-tetraBCN (M_n ~ 20,000 Da, 4 mM). EGFP-EGF-modified hydrogels (5 μL) were formed (1 hr) between Rain-X®-treated glass slides with silicone rubber spacers (McMaster-Carr, 0.5 mm thick) with cell-adhesive and matrix metalloproteinase-degradable crosslinker N_3-GGRGDSPGGPQGIWGQGK(N_3)-NH_2 (Supplementary Methods, 8 mM). After removal from glass slide chamber, gels were exposed to masked collimated UV light (λ = 365 nm, 10 min, 10 mW cm^{-2}) through a slitted photomask containing 400 μm wide line features and soaked in PBS overnight. Gels were swollen (1 hr) in DMEM containing FBS (10%) prior to cell seeding. A431 cells (1 x 10^7 cells mL^{-1}) were added to the top of gels as a droplet (10 μL) and allowed to attach (1 hr, 37 °C) prior to media addition. Gels were swollen in DMEM containing FBS (10%) and penicillin/streptomycin (1%) overnight prior to imaging (condition ①). Gel media was supplemented with bovine serum albumin (BSA, 0.1%) and cells maintained in culture (37 °C, 5% CO_2). Three days after cell seeding, a portion of the hydrogels were exposed to a second round of photopatterning; one half of each treated gel was exposed to UV light (λ = 365 nm, 10 min, 10 mW cm^{-2}) while the other half was left unexposed. In all cases, gels were maintained in DMEM containing FBS (10%), penicillin/streptomycin (1%), and BSA (0.1%) for an additional three days prior to analysis (conditions ② and ③). At the end of each treatment

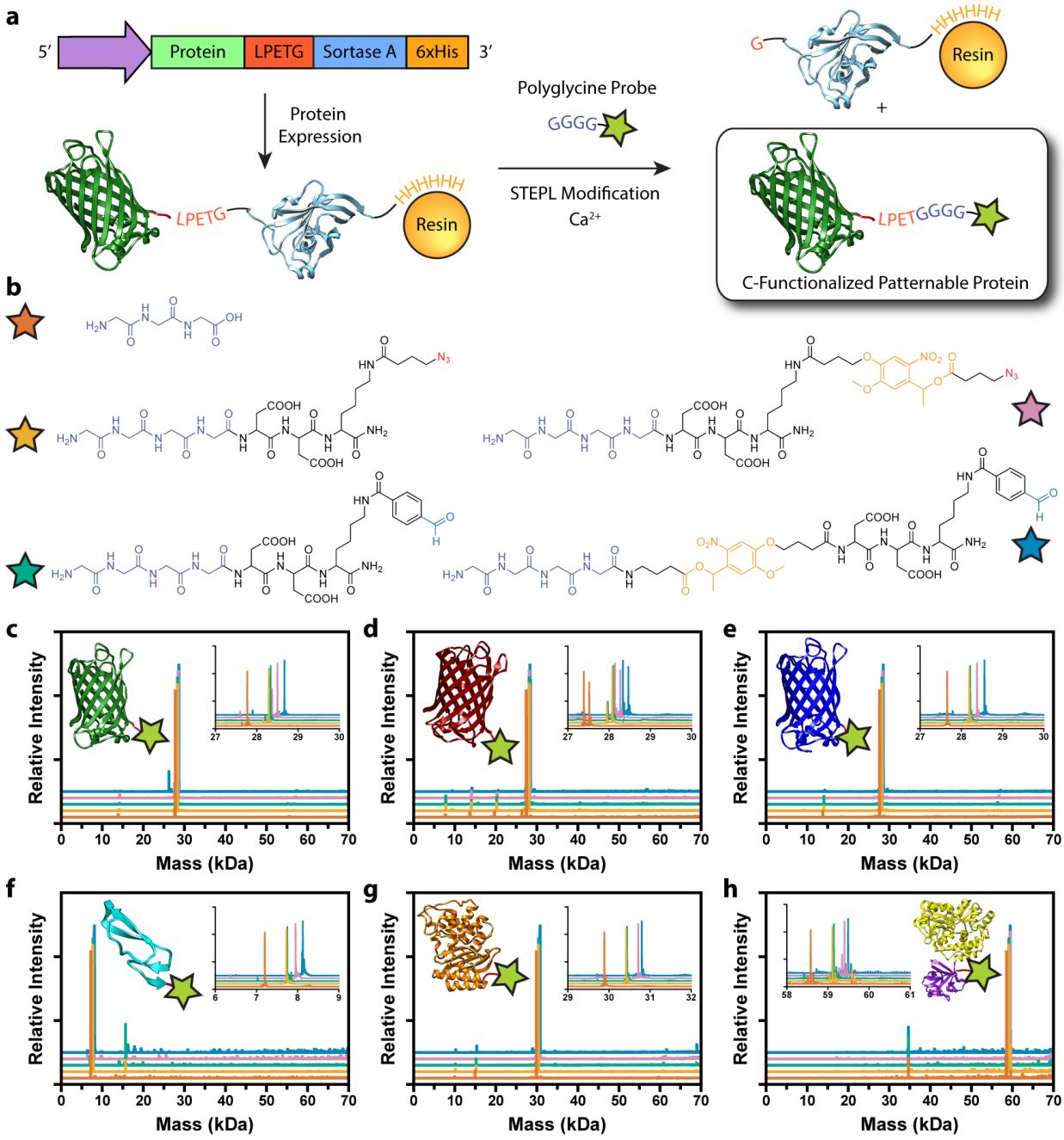
condition, cells were fixed in paraformaldehyde (1%, 30 min), stained with DAPI (0.15 μg/mL, 30 min), and visualized *via* fluorescence microscopy. Image analysis was performed for each condition using ImageJ, quantifying the normalized intensity profile for EGFP, nuclei, and optical transmission across the gel perpendicular to photopatterned lines. Whole-gel images were performed on an Epson Perfection 4490 Photo document scanner.

3.4.11 Encapsulated cell response to subcellular photoreleased EGFP-EGF. A431 cells (2 x 10^6 cells mL^{-1}) were encapsulated in SPAAC hydrogels (5 μL) containing EGFP-EGF-*o*NB-N_3 (125 nM) and N_3-GRGDS-NH_2 peptide (Supplementary Methods, 1 mM) affixed to an azide-functionalized slide (Supplementary Methods). The cells were cultured for 48 hr in DMEM supplemented with FBS (10%) and penicillin/streptomycin (1%) before transferring to phenol red-free DMEM containing FBS (10%) and penicillin/streptomycin (1%). Protein photorelease was performed *via* multiphoton laser scanning lithography (Olympus FV1000 MPE BX61 Multi-photon Microscope, 20x objective, 5x zoom, λ = 740 nm, 25% laser power, 16 scan repeats, z-interval = 2.5 μm), confined to a 40 μm x 30 μm x 5 μm subvolume bisecting an individual cell. ImageJ was used to calculate the difference in EGFP fluorescence pixel intensity accompanying photorelease.

Statistics:

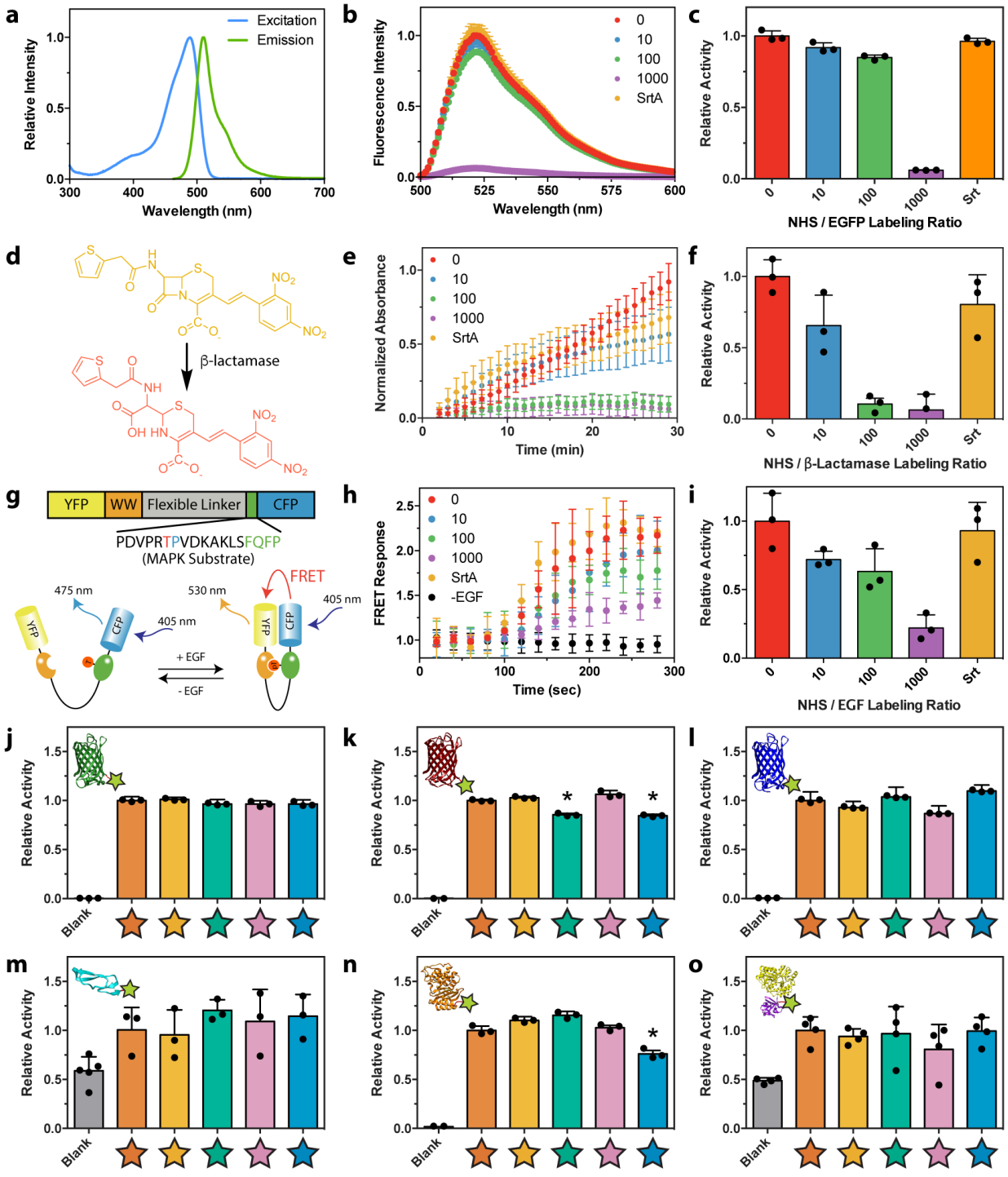
Two-tailed t-tests assuming equal variance were performed to determine p values and statistical significance.

3.5 Figures



3.5.1 Figure 1 | Generation of sortagged protein library for biomaterial modification. **a.** Sortase-Tag Expressed Protein Ligation (STEPL) enables one-step protein biofunctionalization and purification of C-modified proteins for biomaterial decoration. Proteins appended with a genetically encoded sorting signal are expressed as a fusion with the sortase enzyme and a 6xHis

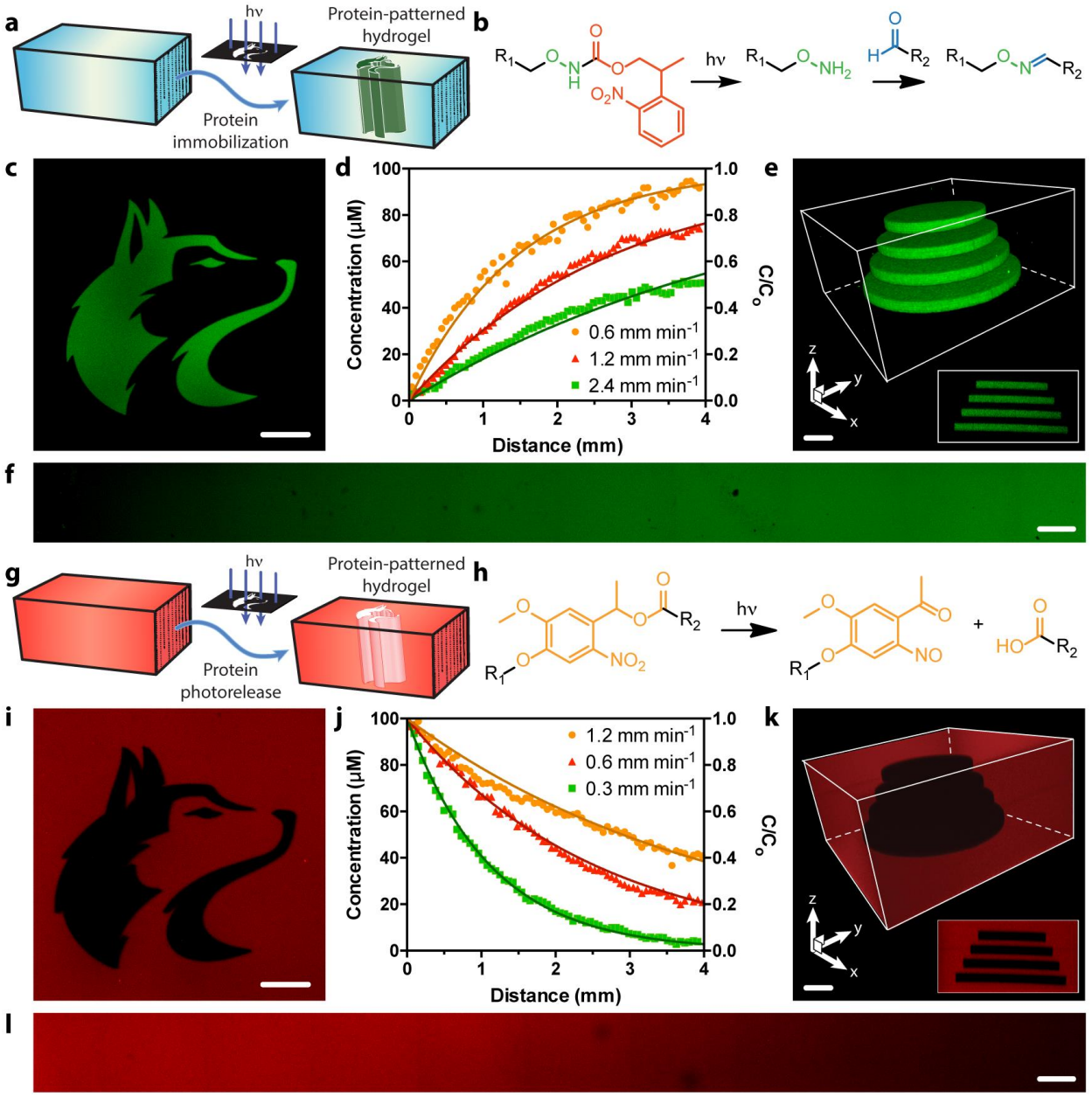
tag. Following chromatographic isolation on Ni-NTA resin, intramolecular sortagging is promoted by the addition of calcium and a polyglycine probe, catalyzing peptide ligation to the protein of interest and simultaneous displacement from the 6xHis-functionalized sortase A. Final protein functionality is defined by polyglycine compound identity. **b.** Five distinct polyglycine probes each with different reactive functional groups enabling biomaterial decoration were synthesized and exploited for STEPL: triglycine, H-GGGGDDK(N_3)-NH_2, H-GGGGDDK(CHO)-NH_2, H-GGGGDDK(*o*NB-N_3)-NH_2, H-GGGG-*o*NB-DDK(CHO)-NH_2 (denoted respectively with orange, tan, teal, pink, and blue stars). **c-h.** Whole-protein mass spectrometry of sortagged proteins (**c**, EGFP; **d**, mCherry; **e**, mCerulean; **f**, EGF; **g**, bla; **h**, FGF) indicates high sample purity and quantitative functionalization for all polyglycine probes. Curve color denotes probe identity following the same scheme as **b**. Double peaks in **d** indicate incomplete N-terminal methionine excision common to mCherry expression. Spectra in **c-h** correspond to observed masses from a single purification of each species.



3.5.2 Figure 2 | Comparing activity of differently modified proteins. a-i. Bioactivity was compared for proteins azide-tagged by STEPL with H-GGGGDDK(N_3)-NH_2 or by conventional NHS labeling with varying molar excesses of N_3-OSu (0, 10, 100, 1000x). **a.** EGFP fluorescence

was used as a surrogate readout of its activity following protein modification. **b.** Fluorescence emission spectra was determined for modified EGFP ($\lambda_{excitation}$ = 470 nm). **c.** Increasing NHS modification led to a decrease in EGFP emission ($\lambda_{excitation}$ = 470 nm, $\lambda_{emission}$ = 530 nm), while sortagged EGFP retained native-like fluorescence. **d.** Bla activity was determined by its ability to degrade a chromogenic nitrocefin substrate, which changes from yellow to red upon β-lactam cleavage. **e.** Time-course spectrophotometric analyses (λ_{abs} = 386 nm) indicate nitrocefin degradation for NHS-modified and sortagged bla. **f.** Sortagged and unmodified bla exhibited statistically indistinguishable levels of bioactivity, as indicated by similar values for k_{cat}. Bla displayed high sensitivity to NHS labeling, likely due to the presence of a critical lysine residue in its active pocket. **g.** Activity of EGF was quantitatively determined with a HeLa cell line expressing EKAREV FRET reporter for MAPK activation. Functional EGF catalyzes phosphorylation of a MAPK substrate, resulting in an intramolecular association that colocalizes a Yellow Fluorescent Protein (YFP) and Cyan Fluorescent Protein (CFP) FRET pair. **h.** YFP/CFP FRET ratios following growth factor stimulation were determined for modified EGF. **i.** Increased NHS labeling of EGF yielded decreased bioactivity, as determined by the initial rate of change in FRET response, while the sortagged growth factor exhibited native levels of activity. **j-o.** Relative bioactivity of proteins (**j**, EGFP; **k**, mCherry; **l**, mCerulean; **m**, EGF; **n**, bla; **o**, FGF) sortagged with triglycine, H-GGGGDDK(N_3)-NH_2, H-GGGGDDK(CHO)-NH_2, H-GGGGDDK(*o*NB-N_3)-NH_2, and H-GGGG-*o*NB-DDK(CHO)-NH_2 (denoted respectively with orange, tan, teal, pink, and blue stars). * denotes conjugates with a statistically significant reduction in bioactivity ($p < 0.05$), as compared to the unmodified species (unpaired two-tailed t-test, $p = 1.3 \times 10^{-4}$ for mCherry-CHO, $p = 3.5 \times 10^{-4}$ for mCherry-*o*NB-CHO, $p = 3.2 \times 10^{-3}$ for bla-*o*NB-CHO). Error bars

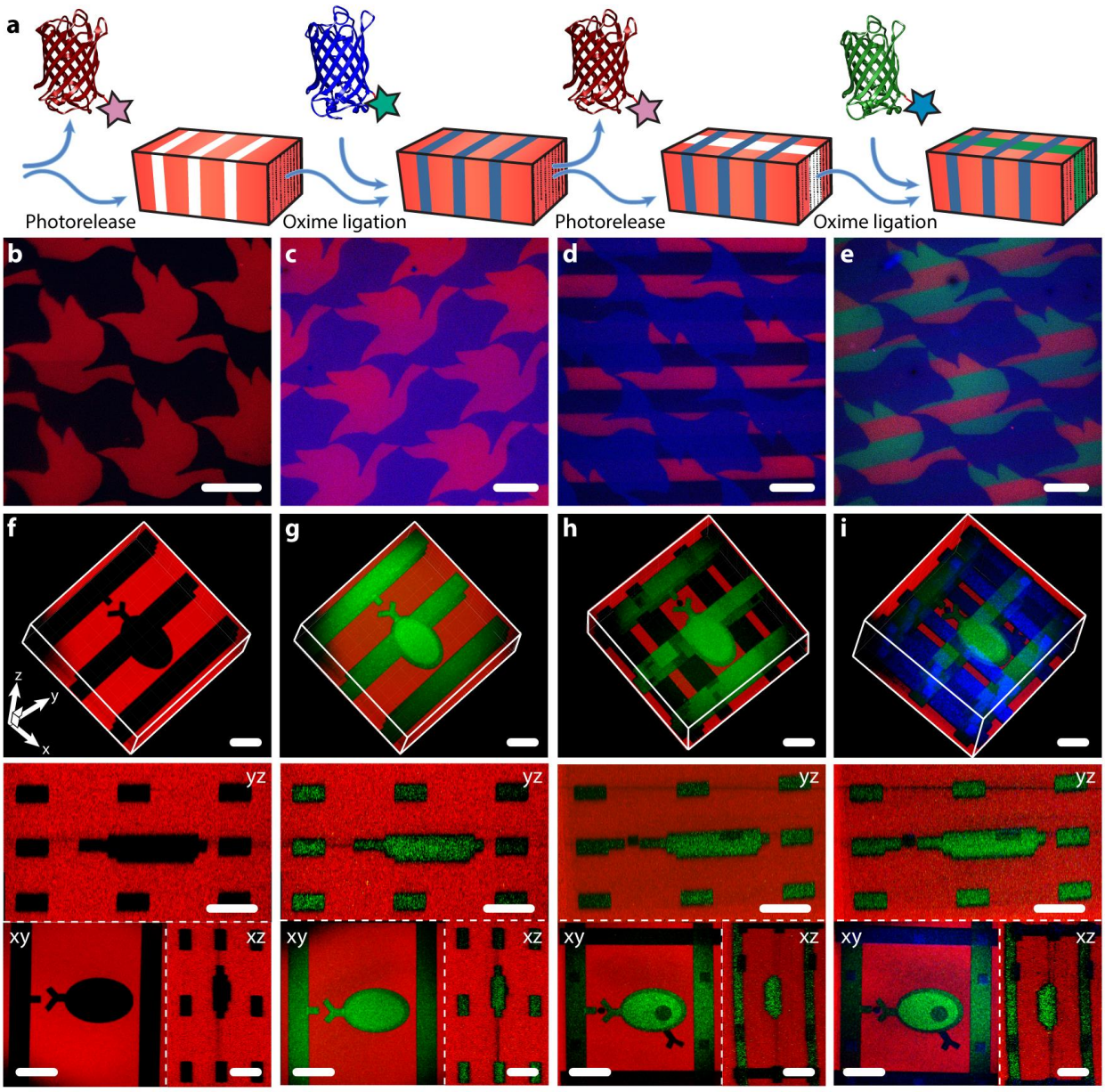
correspond to the standard deviation about the mean for biological replicate experiments (n = 3 for studies involving fluorescent proteins, β-lactamase, and EGF; n = 4 for MBP-FGF).



3.5.3 Figure 3 | Photopatterned alteration of hydrogel biomaterials with sortagged proteins. **a-f.** Photopatterned immobilization of sortagged proteins within gels. **a.** Proteins modified site-specifically with aromatic aldehydes are immobilized within SPAAC-based gels through photomediated oxime ligation. **b.** NPPOC-caged alkoxyamines distributed uniformly throughout gels (R_1) undergo irreversible β-elimination upon mild UV light exposure (λ = 365 nm or 740 nm). The deprotected alkoxyamines react with aldehyde-tagged proteins (R_2) to form stable oxime

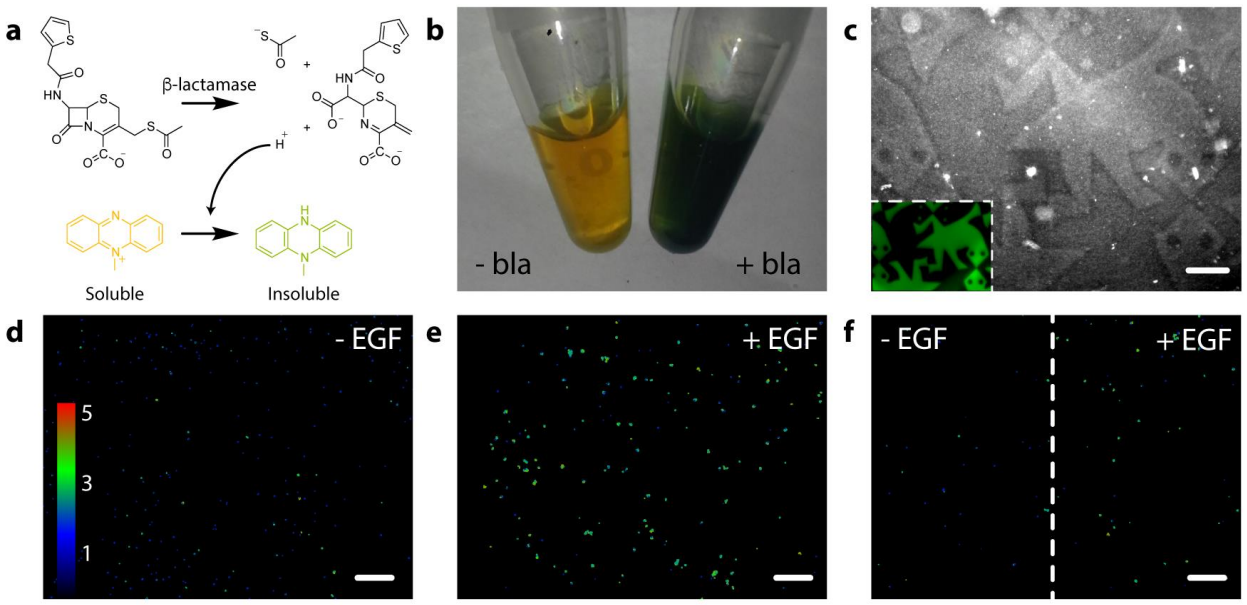
linkages. **c.** Mask-based photolithography (λ = 365 nm) was used to immobilize discrete patterns of EGFP-CHO throughout the gel thickness. **d,f.** By exposing gel surfaces to linear gradients of light exposure (created by covering samples with an opaque photomask moving at rates of 0.6, 1.2, 2.4 mm min^{-1}, 10 mW cm^{-2}, λ = 365 nm), exponential protein gradients were generated in a dose-dependent manner. C_0 represents the highest possible protein concentration that can be immobilized (determined based on the 100 μM caged alkoxyamine included during gel formulation). **e.** Full 3D control over protein tethering within gels is achieved through multiphoton laser-scanning lithographic patterning (λ = 740 nm). **g-l.** Photopatterned release of sortagged proteins from gels. **g.** Site-specifically modified proteins are released from SPAAC-based gels through *ortho*-nitrobenzyl ester (*o*NB) photocleavage. **h.** *o*NB moieties linking the sortagged proteins (R_1) and the hydrogel (R_2) undergo rapid photoscission upon mild UV light exposure (λ = 365 nm or 740 nm). **i.** Mask-based photolithography (λ = 365 nm) was used to dictate discrete patterns of protein release from gels uniformly functionalized with mCherry-*o*NB-N_3. **j,l.** By exposing gel surfaces to linear gradients of light exposure (created by covering samples with an opaque photomask moving at rates of 0.3, 0.6, 1.2 mm min^{-1}, 10 mW cm^{-2}, λ = 365 nm), exponential protein gradients were generated in a dose-dependent manner. C_0 corresponds to the initial mCherry concentration included during gel formulation (100 μM). **k.** Full 3D control over protein release within gels is achieved through multiphoton laser-scanning lithographic patterning (λ = 740 nm). Immobilized protein concentrations in **d** and **j** are determined through fluorescence correlation. Solid curves in **d** and **j** are predicted by known NPPOC/*o*NB photocleavage kinetics. Images in **c**, **e**, **f**, **i**, **k**, and **l** correspond to representative individual gels imaged by fluorescence confocal microscopy. Data in **d** and **j** was derived from experiments involving single gels for each

gradient light condition. Inset in **e** corresponds to a maximum-intensity y-projection, while that in **k** corresponds to a minimum-intensity y-projection. Scale bars = 100 μm.



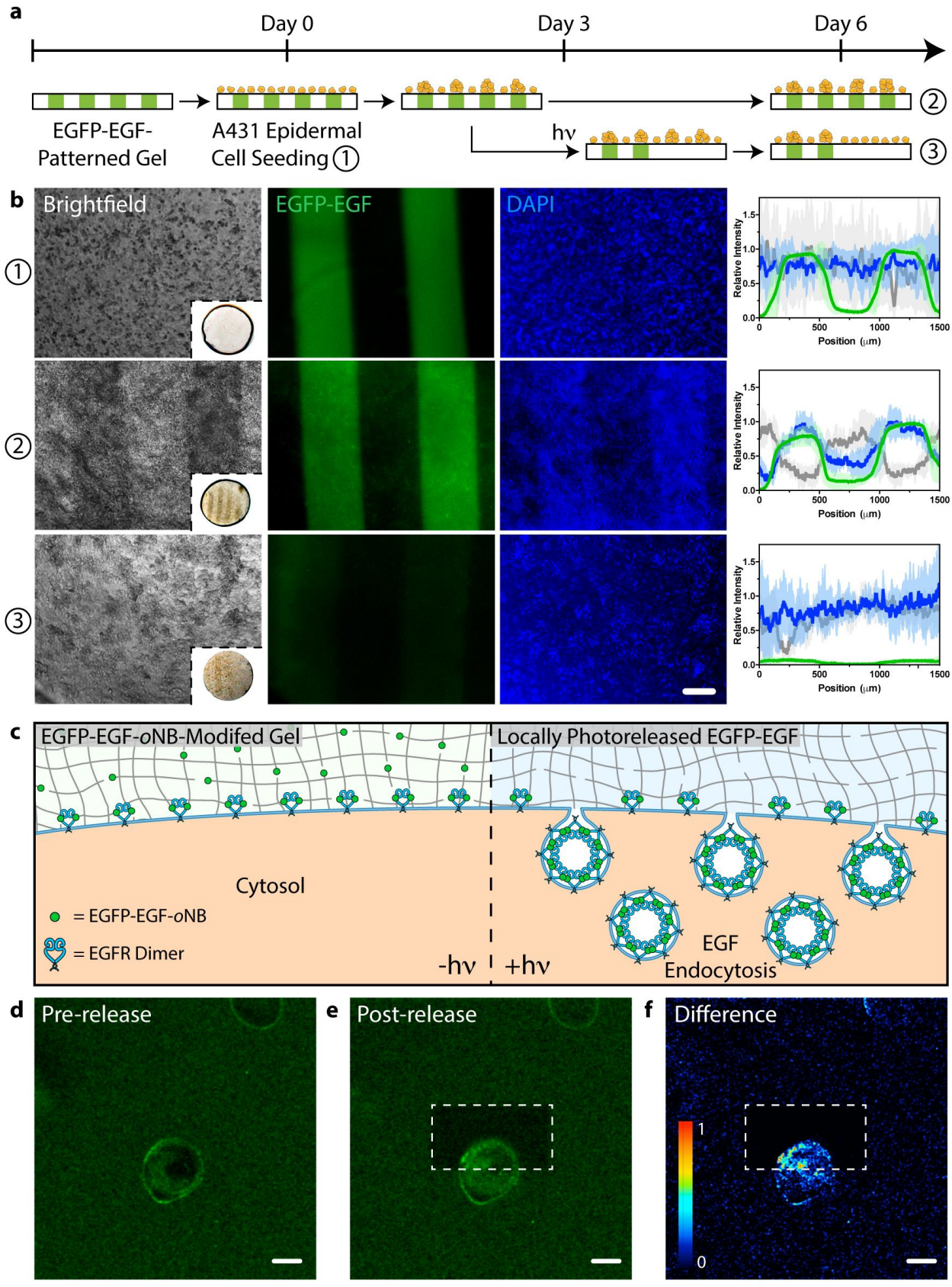
3.5.4 Figure 4 | 4D photoevolution of hydrogel biomaterials patterned with multiple sortagged proteins. a. Protein photorelease can be performed in concert with photomediated ligation of an aldehyde-tagged protein to create complex interconnected biochemical patterns. Iteration of this process enables 4D evolution of protein patterns within gels. **b-e.** Masked-based photolithographic techniques (λ = 365 nm) were utilized to control sequential protein patterning in defined shapes extending throughout the gel thickness. **b.** Following directed exposure of gels

uniformly functionalized with mCherry-*o*NB-N_3 (red), mCherry is released while uncaging sites for subsequent mCerulean-CHO (blue) immobilization by oxime ligation (**c**). **d.** A second round of directed light exposure released additional mCherry while leaving mCerulean patterns intact. **e.** EGFP-*o*NB-CHO (green) is immobilized in the uncaged alkoxyamine sites to create a trifunctional protein pattern. **f-i.** Evolution of trifunctional protein patterns was controlled in 3D space through multiphoton laser-scanning lithography (λ = 740 nm). Rows **f-i** correspond to 3D renderings of the photoevolved materials, where red channels represent minimum intensity projections and green and blue are maximum intensity projections, as well asxy, yz, and xz planar slices. **f**. Photorelease of mCherry-*o*NB-N_3 is followed by (**g**) protein backfilling with EGFP-*o*NB-CHO. **h**. Further treatment with pulsed laser light released both EGFP-*o*NB-CHO and mCherry-*o*NB-N_3 within user-specified gel sub-volumes while creating anchoring sites for (**i**) mCerulean-CHO immobilization. Images in **b**-**i** were generated using fluorescence confocal microscopy on a single gel throughout sequential patterning. Scale bars = 100 μm.



3.5.5 Figure 5 | Spatial patterning of gels with bioactive site-specifically modified enzymes and growth factors. a. Thioacetate cefalotin is hydrolyzed enzymatically by bla, eliminating proton and thiolate ions that reduce a water-soluble yellow phenazine into a green water-insoluble precipitate. **b.** In-solution treatment of thioacetate cefalotin (5 mM) and phenazine methosulfate (6.5 mM) with bla (30 µM) yields distinct color change and precipitate formation. **c.** Gels containing photopatterned regions of immobilized bla were visualized by phase contrast microscopy upon treatment with thioacetate cefalotin and phenazine methosulfate. Insoluble precipitation was confined to bla-modified gel subvolumes. Inset corresponds to an EGFP-modified gel with the same tessellated protein pattern. **d.** HeLa cells encapsulated in gels and expressing EKAREV FRET reporter for MAPK signaling exhibit basal signaling. **e.** Average intracellular MAPK activation increases ~2 fold when gels are uniformly functionalized with EGF (12.5 nM). **f.** High MAPK activation persisted in EGF-patterned gel regions, whereas no upregulation in MAPK levels is observed in unfunctionalized regions. Images in **d-f** represent color-coded FRET response normalized to basal MAPK activation, obtained from confocal z-

slices. All experiments were conducted in triplicate (n = 3) with similar results. Scale bars = 250 μm.


a
Day 0
Day 3
Day 6
EGFP-EGF-
Patterned Gel
A431 Epidermal
Cell Seeding ①
hν
②
③
b
Brightfield
EGFP-EGF
DAPI
①
②
③
Relative Intensity
Position (μm)
c
EGFP-EGF-oNB-Modifed Gel
Locally Photoreleased EGFP-EGF
Cytosol
= EGFP-EGF-oNB
= EGFR Dimer
-hν
+hν
EGF
Endocytosis
d
Pre-release
e
Post-release
f
Difference
1
0

3.5.6 Figure 6 | Modulating cell fate with a photoreleasable sortagged fluorophore-growth factor chimeric protein. a. Patterned population-level control of A431 cell proliferation is achieved through stimulation with gel-immobilized EGFP-EGF-*o*NB-N_3 in 2D culture. Protein photoremoval promotes dynamic cell redistribution. Key analytical timepoints are noted on the experimental timeline with circled numbers. **b.** Brightfield and fluorescent images highlight cell response (nuclei in blue) to patterned EGFP-EGF-*o*NB-N_3 (green) at various experimental endpoints. Relative intensity profiles are given for EGFP (green), nuclei (blue), and optical transmission (grey) across gels perpendicular to photopatterned lines, with average values (dark lines) and standard deviations (light error bars) corresponding to data from three biological replicates (n = 3). Brightfield insets depict full hydrogel (~0.5 cm in diameter). Images and analysis in condition ③ correspond to the gel half whose protein was photoreleased on Day 3. **c.** Encapsulated cells bind but cannot internalize hydrogel-tethered EGFP-EGF-*o*NB-N_3. Photoliberation of the soluble protein promotes canonical EGFR activation and associated membrane endocytosis. **d-e.** EGFP-EGF protein release is confined to gel subvolumes bisecting a single A431 cell in 3D culture. Endosome formation is visible in <5 minutes and concentrated in the regions of light exposure. Fluorescent images correspond to timepoints (**d**) immediately preceding and (**e**) 5 minutes after protein photorelease within a single gel. **f.** Difference calculations between images pre- and post-release highlight local EGFP-EGF internalization. Scale bar = 200 μm in **b** and 10 μm in **d-f**.

Chapter 4

Genetically Encoded Photocleavable Linkers for Patterned Protein Release from Biomaterials

4.1 Abstract:

Given the critical role that proteins play in almost all biological processes, there is great interest in controlling their presentation within and release from biomaterials. Despite such outstanding enthusiasm, previously developed strategies in this regard result in ill-defined and heterogeneous populations with substantially decreased activity, precluding their successful application to "fragile" species including growth factors. Here, we introduce a modular and scalable method to create monodisperse genetically encoded chimeras that enable bioactive proteins to be immobilized within and subsequently photoreleased from polymeric hydrogels. Building upon recent developments in chemoenzymatic reactions, bioorthogonal chemistry, and optogenetics, we tether fluorescent proteins, model enzymes, and growth factors site-specifically to gel biomaterials through a photocleavable protein (PhoCl) that undergoes irreversible backbone photoscission upon exposure to cytocompatible visible light ($\lambda \sim 400$ nm) in a dose-dependent manner. Mask-based and laser-scanning lithographic strategies using commonly available light sources are employed to spatiotemporally pattern protein release from hydrogels while retaining their full activity. Photopatterned epidermal growth factor presentation is exploited to promote anisotropic cellular proliferation in 3D. We expect these methods to be broadly useful for applications in diagnostics, drug delivery, and regenerative medicine.

4.2 Introduction:

Proteins represent the most functionally diverse class of biomolecules. Intimately involved in virtually all cellular signaling cascades and biological decisions of fate (e.g., proliferation, migration, differentiation), chemists, biologists, and material scientists alike have long sought to control their presentation within and release from biomaterials[1–4]. Though several chemical strategies exist to functionalize constructs indefinitely with biomacromolecules, "smart" materials that release proteins on-demand have proven uniquely versatile in regulating therapeutic deployment and to modulate cell response *in vitro* and *in vivo*[5,6]. In one particularly powerful approach, proteins are covalently tethered to materials through degradable bonds; exposure to the appropriate pre-programmed stimulus induces bond cleavage and concurrent protein release[7]. Exploiting scissile moieties sensitive to pH, reductants, or enzymes, triggered protein release from 3D biomaterials has been demonstrated using a variety of biologically relevant external stimuli and for many different functional proteins[8–11].

Although several stimuli-sensitive linkers have found utility in specifying *when* a protein payload is released, exceptionally few can also dictate *where* within a given material this occurs. Spatiotemporal control of protein release has unique utility in directing anisotropic cell behaviors and investigating the effects of dynamic biochemical presentation on cells within synthetic culture platforms[12–15]. Photochemical techniques are promising in this regard, enabling proteins to be released from materials with full 4D control based on patterned light exposure at subcellular resolutions[16–18]. Owing to their synthetic tractability and efficient photoscission under cytocompatible UV light, *ortho*-nitrobenzyl (*o*NB)- and coumarin (CM)-based esters have been the most commonly employed moiety for photomediated protein delivery[19–23]. Since *o*NB- and

CM-based linkers are non-biological and have not yet proven amenable to translational incorporation using genetic code expansion, post-synthetic strategies must be employed to install such photoactive species onto proteins[24,25]. Most commonly, activated esters of the photosensitive small molecules are coupled stochastically to primary amines present on lysine side chains and at the N-terminus, ultimately installing a variable number of photocleavable linkers through amide bonds at poorly defined locations on the protein surface. Though these approaches have been utilized to control 4D protein release within cell-laden materials, significant room for improvement exists as: 1) post-synthetic chemical modification ensures a heterogeneous protein population that suffers from batch-to-batch variability and is likely intractable for translation; 2) random modification with photoactive small molecules often results in substantially decreased protein bioactivity; and 3) synthesis of *o*NB- or CM-modified photoreleasable proteins may be difficult to scale up.

Seeking to reap the benefits of phototriggered protein release while addressing the primary drawbacks of photosensitive small molecule-based approaches, we envisioned tethering proteins of interest to biomaterials through a genetically encoded photocleavable protein known as PhoCl[26]. Recently developed by the optogenetics community as a unique tool for photoregulated gene expression, PhoCl is an engineered monomeric green fluorescent protein that undergoes irreversible β-elimination[27] and concomitant polypeptide backbone cleavage upon exposure to visible light ($\lambda \sim 400$ nm) (Scheme 1). Photocleavage of PhoCl yields a moderately sized N-terminal fragment (231 amino acids, 26 kDa) and a small C-terminal fragment (15 amino acids, 1.7 kDa), neither of which is fluorescent. Genetic fusions between PhoCl and proteins of interest

can be recombinantly produced to yield perfectly defined photocleavable fusion proteins using bacterial fermentation processes that are well-established and scalable.

To take advantage of the relatively small cleavage "scar" left on photoreleased proteins that had been fused C-terminally to PhoCl, we sought to immobilize the photoscissile chimeras to biomaterials through their N-termini. In this regard, we identified N-myristoyltransferase (NMT), an enzyme which promotes cotranslational fatty acylation on proteins bearing the "GXXXS/T" signature peptide sequence (where X is any amino acid). Since NMT tolerates many synthetic analogs of its natural myristic acid substrate, including those containing azides, alkynes, carbonyls, and halogens[28–30], we expected this to be a generalizable strategy compatible with many chemistries common to biomaterial formation and modification. Moreover, the efficiency and site-specificity of this chemoenzymatic transformation, coupled with NMT's co-translational operation and the genetic encodability of all components, provides scalable access to perfectly defined chimeric proteins that can be precisely tethered to biomaterials and subsequently photoreleased upon visible light exposure.

4.3 Results and Discussion:

Towards production of genetically encoded photoreleasable proteins, we created five expression vectors for PhoCl fused N-terminally with individual proteins from three distinct classes: mRuby, sfGFP, and mCerulean were selected for their fluorescence (red, green, and blue, respectively)[31–33]; β-lactamase (bla) was included as a model enzyme; and epidermal growth factor (EGF) was

chosen as a representative cytokine. For each fusion construct, an NMT recognition peptide sequence (MGNEASYPL)[34] and a 6xHis purification tag were installed at PhoCl's N-terminus. *E. coli* were co-transformed with a single PhoCl fusion-containing plasmid and another encoding for NMT and methionine aminopeptidase[35], an enzyme essential for myristoylation. Since bioorthogonal click reactions involving reactive azides are now widely utilized in the synthesis and decoration of biomaterials[36–39], we opted to express proteins in the presence of 12-azidododecanoic acid (12-ADA). As a synthetic analog of NMT's natural myristic acid substrate, chemoenzymatic modification involving 12-ADA yields site-specific installation of azido functionality at the N-terminus (Scheme 2). Following expression and affinity purification, the N-modified photoreleasable proteins were obtained in good yield (~15 mg L^{-1} culture, nonoptimized expression). Analysis utilizing sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and whole-protein mass spectrometry revealed high sample purity (>90%) and quantitative N-terminal azide-tagging (Supporting Information).

To determine photocleavage kinetics of the azide-modified chimeras, we investigated PhoCl response to violet light (λ = 400 nm, 10 mW cm^{-2}, 0 – 30 min) for two different classes of proteins using independent methodologies. We first quantified the disappearance of N_3-PhoCl-bla fluorescence ($\lambda_{PhoCl,excitation}$ = 380 nm, $\lambda_{PhoCl,emission}$ = 510 nm) accompanying PhoCl photoscission at 25 °C (Figure 1a,b). Under these conditions, PhoCl exhibited the expected first-order decay with a kinetic constant of 0.15 ± 0.02 min^{-1} and a half-life of 4.7 ± 0.6 min. To decouple photocleavage analysis from any potential photobleaching, SDS-PAGE analysis was subsequently performed on N_3-PhoCl-mRuby exposed to different amounts of light; the extent of photocleavage was determined by quantifying the intensities of the disappearing band from the intact fusion and the

appearance of bands corresponding to the photocleaved products (Figure 1c,d). These analyses revealed a first-order decay constant of 0.15 ± 0.03 min^{-1} and a half-life of 4.7 ± 0.8 min, values that were statistically indistinguishable from that based on PhoCl fluorescence disappearance. To ensure that PhoCl was indeed cleaving at the expected location (PhoCl's H227 residue which matures as part of its chromophore), whole protein mass spectrometry (ESI LC/MS) was performed on N_3-PhoCl-mRuby both before and after photocleavage (λ = 400 nm, 10 mW cm^{-2}, 30 min) (Figure 1e,f). Mass spectrometry revealed a singly azide-modified protein fusion of the expected mass (observed, 56,301 Da; expected, 56,262 Da) that decayed into the anticipated cleavage products (observed for PhoCl fragment, 27,973 Da; expected, 27,956 Da; observed for mRuby fragment, 28,301 Da; expected 28,326 Da) upon exposure to violet light.

Having determined the PhoCl domain's dose-dependent responsiveness to visible light, we next sought to compare the bioactivity of intact and photocleaved PhoCl fusions with the corresponding unmodified proteins of interest. As spectral overlap and potential FRET between PhoCl and mRuby/sfGFP/mCerulean render fluorescence measurements a less-than-perfect surrogate for their retained bioactivity, we focused our efforts on the more fragile and biologically relevant bla and EGF proteins. Bla activity was assessed via a colorimetric assay involving hydrolysis of the chromogenic cephalosporin nitrocefin, while EGF function was determined based on its ability to stimulate fibroblast cell proliferation as quantified by increased dsDNA synthesis and content. In both cases, the activity of the native protein was statistically indistinguishable from that of the PhoCl fusion regardless of whether or not they had been photocleaved (λ = 400 nm, 10 mW cm^{-2}, 30 min) (Figure 2).

Encouraged that the genetically encoded photoreleasable proteins retained native bioactivity, we next sought to test their utility in a biomaterials' context. Since the protein chimeras were N-terminally functionalized with reactive azides, we exploited hydrogel biomaterials formed via strain-promoted azide-alkyne cycloaddition (SPAAC) between poly(ethylene glycol) tetra-bicyclononyne (PEG-tetraBCN, M_n ~ 20 kDa, 4 mM) and a linear PEG diazide (N_3-PEG-N_3, M_n ~3.5 kDa, 8 mM) in phosphate-buffered saline (pH = 7.4). Owing to the large water content of the resultant materials, a regular mesh size that supports soluble protein diffusion, and the bioorthogonal and cytocompatible nature of SPAAC[40], these idealized step-growth polymer networks have proven beneficial for drug delivery and 3D cell culture applications[41,42]. Moreover, their optical clarity renders them useful constructs for photochemical modulation, particularly with respect to species photorelease. Finally, inclusion of azide-tagged photocleavable fusion proteins during gel formulation facilitates their site-specific conjugation uniformly throughout biomaterials at a user-defined concentration.

To test to the potential utility of PhoCl fusions for photomediated protein delivery from biomaterials, N_3-PhoCl-sfGFP and N_3-PhoCl-bla were independently immobilized (15 μM) within SPAAC gels. A subset of each gel type was exposed to light (λ = 400 nm, 10 mW cm^{-2}, 30 min) to induce photocleavage. 16 hours after light exposure, gel supernatants were analyzed for soluble protein; sfGFP was quantified by its fluorescence ($\lambda_{sfGFP,excitation}$ = 488 nm, $\lambda_{sfGFP,emission}$ = 530 nm), while bla enzyme release was assessed using the colorimetric nitrocefin assay. In both cases, supernatants of gels exposed to light indicated successful phototriggered release of active proteins,

while those from unexposed samples exhibited no non-specific release (Figure 3). These experiments highlight the potential utility of this approach for on-demand protein therapeutic delivery, as well as the strategy's absence of undesired "leaky" release.

Since PhoCl cleavage can be spatiotemporally controlled using directed light, we sought to demonstrate its versatility in creating photopatterned protein-functionalized biomaterials. As a proof of concept, N_3-PhoCl-sfGFP, N_3-PhoCl-mRuby, and N_3-PhoCl-mCerulean (15 μM) were independently incorporated into PEG hydrogels during gelation. Protein-functionalized gels were then exposed to masked violet light (λ = 400 nm, 10 mW cm^{-2}, 30 min) to trigger PhoCl cleavage in user-defined patterns (tessellated geckos for sfGFP, squares for mRuby, and an Escher-inspired fish/bird pattern for mCerulean) and released proteins were allowed to diffuse from the gels (16 hours) prior to fluorescent confocal imaging ($\lambda_{sfGFP,excitation}$ = 488 nm, $\lambda_{sfGFP,emission}$ = 500 – 590 nm; $\lambda_{mRuby,excitation}$ = 560 nm, $\lambda_{mRuby,emission}$ = 600 – 700 nm; $\lambda_{mCerulean,excitation}$ = 405 nm, $\lambda_{mCerulean,emission}$ = 425 – 475 nm) (Figure 4a,b). In each case, the fluorescence from the PhoCl domain co-localized with the fluorescent fusion partner. We anticipated this to be a convenient tool for tracking regions of protein immobilization should the protein of interest itself not be fluorescent. Results demonstrated micron-scale patterning resolution with long-term stability (>2 weeks).

Though conventional mask-based photolithography proved useful in generating binary patterned materials, we next sought to exploit PhoCl's dose-dependent photoresponsiveness to create gradients of immobilized proteins through spatially varied photorelease. Gels containing N_3-PhoCl-mRuby (15 μM) were exposed to linear gradients of light, created by moving an opaque

photomask across the gel surface at a fixed rate during light exposure (λ = 400 nm, 10 mW cm^{-2})[43]. To modulate the slope of the graded light exposure, the photomask velocity was varied (0.4 and 1.2 mm min^{-1}). After light treatment and diffusive removal of the released protein from the gel, the patterned materials were imaged using fluorescent confocal microscopy (Figure 4c). Gradient signatures for both PhoCl and mRuby fluorescence scaled as expected between the different graded light treatments and matched predictions from in-solution photocleavage kinetic analysis (Figure 1) with a first-order decay with a kinetic constant of 0.15 ± .01 min^{-1} and a half-life of 4.6 ± 0.3 min (Supporting Information). This precise control over protein release represents a powerful and facile tool for the creation of well-defined protein-patterned biomaterials.

While PhoCl cleavage can be readily controlled using mask-based photolithography, we hypothesized that the overlap in light wavelength emitted from a 405 diode laser (λ = 405 nm) with that required for photocleavage would render it amenable to laser-scanning lithographic patterning[44–46]. Here, hydrogels containing N_3-PhoCl-mCerulean (15 μM) were selectively exposed to focused laser light (λ = 405 nm) in user-defined regions of interests (ROIs) using a conventional confocal microscope. Experiments revealed that tethered proteins could be completely released under mild laser conditions (5% laser power, scan speed = 200 Hz, 64 scan repeats, 10x objective, 100 mW 405 nm diode laser) (Figure 5a**)**. Moreover, since protein release scaled with scan number repeats, stepped protein gradients with micron-scale resolution could be generated by varying the number of laser scans (0 – 64) within adjacent regions of interest (Figure 5b,c). Since diode laser lines and controlled ROI scanning are common features on most confocal microscopes, arbitrarily patterned biomaterials can be generated within minutes using equipment that is readily available in virtually all academic research environments.

Having established the ability to modulate biomaterial properties with photoreleasable bioactive proteins, we investigated the effects of patterned EGF on cell fate. HeLa cells were encapsulated within the SPAAC-based gels containing an individual photosensitive chimera (4 μM); N_3-PhoCl-EGF was included to potentially promote cell proliferation, while N_3-PhoCl-mRuby was utilized as a non-cell-stimulating control. Gels were crosslinked with an azide-flanked linear peptide (6 mM) susceptible to enzymatic cleavage by cell-secreted matrix metalloproteases and contained a cell-adhesive RGD motif. One day post encapsulation, gels were exposed to patterned violet light (λ = 400 nm, 10 mW cm^{-2}, 30 min) through a slitted photomask containing 400 μm wide line features to selectively release EGF/mRuby. Cells were maintained in culture for a total of 14 days, fixed, and their nuclei stained with 4′,6-diamidino-2-phenylindole (DAPI) prior to fluorescent and brightfield imaging (Figure 6). Areas with persistently immobilized EGF (as visualized by PhoCl fluorescence) displayed significantly larger spheroid growth and more cell density than those where EGF had been photoreleased ($p < 0.05$). As expected, control experiments involving mRuby yielded spheroids whose size did not vary throughout the patterned gel and were smaller than those in the released EGF regions. These results demonstrate that PhoCl-mediated gel patterning can be performed with high fidelity in the presence of and to govern spatiotemporal fate of encapsulated cells.

4.4 Conclusion:

In this manuscript, we have introduced a robust and versatile method to immobilize proteins site-specifically to and then subsequently photorelease them from hydrogel biomaterials. Relying on a

co-translational chemoenzymatic installation of a bioorthogonal reactive handle for N-terminal material tethering and the unique PhoCl protein that undergoes irreversible photoscission upon exposure to cytocompatible visible light, we control the immobilization and triggered delivery of fluorescent proteins, enzymes, and growth factors in a manner that preserves their activity. As the entire system is genetically encoded, perfectly monodisperse photosensitive chimeras can be generated through scalable fermentation processes. Having demonstrated the spatiotemporally dictated release of bioactive species as well as the utility of photopatterned protein presentation towards guiding anisotropic cell fates in 3D, we anticipate that these approaches will find great utility in multicellular tissue engineering and therapeutic protein delivery.

Acknowledgements:

Dr. L. Newman and Dr. R. Kahn (Emory University) for gifting the hNMT1 with Met-AP plasmid,[35] Dr. N. King for assistance with figure preparation. We gratefully acknowledge support from S. Edgar at the UW Mass Spectrometry Center as well as that from the NIH and N. Peters at the UW W. M. Keck Microscopy Center (S10 OD016240). This work was supported by a University of Washington Faculty Startup Grant (C.A.D.), a Jaconette L. Tietze Young Scientist Research Award (C.A.D.), and a CAREER Award (DMR 1652141, C.A.D.) from the National Science Foundation.

4.5 References:

1. Hubbell, J. A. Bioactive biomaterials. *Curr. Opin. Biotechnol.* **10,** 123–129 (1999).

2. Peppas, N. A., Hilt, J. Z., Khademhosseini, A. & Langer, R. Hydrogels in biology and medicine: From molecular principles to bionanotechnology. *Adv. Mater.* **18,** 1345–1360 (2006).

3. Tayalia, P. & Mooney, D. J. Controlled Growth Factor Delivery for Tissue Engineering. *Adv. Mater.* **21,** 3269–3285 (2009).

4. Seliktar, D. Designing Cell-Compatible Hydrogels for Biomedical Applications. *Science* **336,** 1124–1128 (2012).

5. Li, J. & Mooney, D. J. Designing hydrogels for controlled drug delivery. *Nat. Rev. Mater.* **1,** 16071 (2016).

6. Lu, Y., Aimetti, A. A., Langer, R. & Gu, Z. Bioresponsive materials. *Nat. Rev. Mater.* **1,** 16075 (2016).

7. Badeau, B. A. & DeForest, C. A. Programming Stimuli-Responsive Behavior into Biomaterials. *Annu. Rev. Biomed. Eng.* **21,** annurev-bioeng-060418-052324 (2019).

8. Chen, W. *et al.* In situ forming reduction-sensitive degradable nanogels for facile loading and triggered intracellular release of proteins. *Biomacromolecules* **14,** 1214–1222 (2013).

9. Purcell, B. P. *et al.* Injectable and bioresponsive hydrogels for on-demand matrix metalloproteinase inhibition. *Nat. Mater.* **13,** 653–61 (2014).

10. Ham, H. O. *et al.* In situ regeneration of bioactive coatings enabled by an evolved Staphylococcus aureus sortase A. *Nat. Commun.* **7,** 11140 (2016).

11. Guo, C. *et al.* Bio-orthogonal conjugation and enzymatically triggered release of proteins within multi-layered hydrogels. *Acta Biomater.* **56,** 80–90 (2017).

12. Tibbitt, M. W. & Anseth, K. S. Dynamic Microenvironments: The Fourth Dimension. *Sci. Transl. Med.* **4,** (2012).

13. Burdick, J. A. & Murphy, W. L. Moving from static to dynamic complexity in hydrogel

design. *Nat. Commun.* **3,** 1–8 (2012).

14. DeForest, C. A. & Anseth, K. S. Advances in bioactive hydrogels to probe and direct cell fate. *Annu. Rev. Chem. Biomol. Eng.* **3,** (2012).

15. LeValley, P. J. & Kloxin, A. M. Chemical Approaches to Dynamically Modulate the Properties of Synthetic Matrices. *ACS Macro Lett.* **8,** 7–16 (2019).

16. Ruskowitz, E. R. & DeForest, C. A. Photoresponsive biomaterials for targeted drug delivery and 4D cell culture. *Nat. Rev. Mater.* **3,** 17087 (2018).

17. Sankaran, S. & del Campo, A. Optoregulated Protein Release from an Engineered Living Material. *Adv. Biosyst.* **3,** 1800312 (2019).

18. Rapp, T. L., Highley, C. B., Manor, B. C., Burdick, J. A. & Dmochowski, I. J. Ruthenium-Crosslinked Hydrogels with Rapid, Visible-Light Degradation. *Chem. - A Eur. J.* **24,** 2328–2333 (2018).

19. Azagarsamy, M. A. & Anseth, K. S. Wavelength-Controlled Photocleavage for the Orthogonal and Sequential Release of Multiple Proteins. *Angew. Chemie Int. Ed.* **52,** 13803–13807 (2013).

20. Griffin, D. R. *et al.* Synthesis of Photodegradable Macromers for Conjugation and Release of Bioactive Molecules. *Biomacromolecules* **14,** 1199–1207 (2013).

21. Wegner, S. V., Sentürk, O. I. & Spatz, J. P. Photocleavable linker for the patterning of bioactive molecules. *Sci. Rep.* **5,** 18309 (2016).

22. DeForest, C. A. & Tirrell, D. A. A photoreversible protein-patterning approach for guiding stem cell fate in three-dimensional gels. *Nat. Mater.* **14,** 523–531 (2015).

23. Shadish, J. A., Benuska, G. M. & DeForest, C. A. Bioactive site-specifically modified proteins for 4D patterning of gel biomaterials. *Nat. Mater.* (2019). doi:10.1038/s41563-019-0367-7

24. Spicer, C. D., Pashuck, E. T. & Stevens, M. M. Achieving Controlled Biomolecule–Biomaterial Conjugation. *Chem. Rev.* **118,** 7702–7743 (2018).

25. Fisher, S. A., Baker, A. E. G. & Shoichet, M. S. Designing Peptide and Protein Modified

Hydrogels: Selecting the Optimal Conjugation Strategy. *J. Am. Chem. Soc.* **139,** 7416–7427 (2017).

26. Zhang, W. *et al.* Optogenetic control with a photocleavable protein, PhoCl. *Nat. Methods* **14,** 391–394 (2017).

27. Mizuno, H. *et al.* Photo-induced peptide cleavage in the green-to-red conversion of a fluorescent protein. *Mol. Cell* **12,** 1051–1058 (2003).

28. Devadas, B. *et al.* Substrate specificity of Saccharomyces cerevisiae myristoyl-CoA: protein N-myristoyltransferase. Analysis of fatty acid analogs containing carbonyl groups, nitrogen heteroatoms, and nitrogen heterocycles in an in vitro enzyme assay and subsequent identifi. *J. Biol. Chem.* **267,** 7224–7239 (1992).

29. Heal, W. P., Wright, M. H., Thinon, E. & Tate, E. W. Multifunctional protein labeling via enzymatic N-terminal tagging and elaboration by click chemistry. *Nat. Protoc.* **7,** 105–117 (2012).

30. Kulkarni, C., Kinzer-Ursem, T. L. & Tirrell, D. A. Selective functionalization of the protein N terminus with N-myristoyl transferase for bioconjugation in cell lysate. *ChemBioChem* **14,** 1958–1962 (2013).

31. Kredel, S. *et al.* mRuby, a bright monomeric red fluorescent protein for labeling of subcellular structures. *PLoS One* **4,** e4391 (2009).

32. Pédelacq, J. D., Cabantous, S., Tran, T., Terwilliger, T. C. & Waldo, G. S. Engineering and characterization of a superfolder green fluorescent protein. *Nat. Biotechnol.* **24,** 79–88 (2006).

33. Rizzo, M. A., Springer, G. H., Granada, B. & Piston, D. W. An improved cyan fluorescent protein variant useful for FRET. *Nat. Biotechnol.* **22,** 445–449 (2004).

34. Kulkarni, C., Lo, M., Fraseur, J. G., Tirrell, D. A. & Kinzer-Ursem, T. L. Bioorthogonal Chemoenzymatic Functionalization of Calmodulin for Bioconjugation Applications. *Bioconjug. Chem.* **26,** 2153–60 (2015).

35. Van Valkenburgh, H. A. & Kahn, R. A. Coexpression of proteins with methionine

aminopeptidase and/or N-myristoyltransferase in Escherichia coli to increase acylation and homogeneity of protein preparations. *Methods Enzymol.* **344,** 186–193 (2002).

36. Iha, R. K. *et al.* Applications of Orthogonal "Click" Chemistries in the Synthesis of Functional Soft Materials. *Chem. Rev.* **109,** 5620–5686 (2009).

37. Nimmo, C. M. & Shoichet, M. S. Regenerative biomaterials that 'click': Simple, aqueous-based protocols for hydrogel synthesis, surface immobilization, and 3D patterning. *Bioconjug. Chem.* **22,** 2199–2209 (2011).

38. Azagarsamy, M. A. & Anseth, K. S. Bioorthogonal Click Chemistry: An Indispensable Tool to Create Multifaceted Cell Culture Scaffolds. *ACS Macro Lett.* **2,** 5–9 (2013).

39. Madl, C. M. & Heilshorn, S. C. Bioorthogonal Strategies for Engineering Extracellular Matrices. *Adv. Funct. Mater.* **28,** 1706046 (2018).

40. Agard, N. J., Prescher, J. A. & Bertozzi, C. R. A strain-promoted [3 + 2] azide-alkyne cycloaddition for covalent modification of biomolecules in living systems. *J. Am. Chem. Soc.* **126,** 15046–15047 (2004).

41. DeForest, C. A., Polizzotti, B. D. & Anseth, K. S. Sequential click reactions for synthesizing and patterning three-dimensional cell microenvironments. *Nat. Mater.* **8,** 659–664 (2009).

42. DeForest, C. A. & Anseth, K. S. Cytocompatible click-based hydrogels with dynamically tunable properties through orthogonal photoconjugation and photocleavage reactions. *Nat. Chem.* **3,** 925–931 (2011).

43. Johnson, P. M., Reynolds, T. B., Stansbury, J. W. & Bowman, C. N. High throughput kinetic analysis of photopolymer conversion using composition and exposure time gradients. *Polymer (Guildf).* **46,** 3300–3306 (2005).

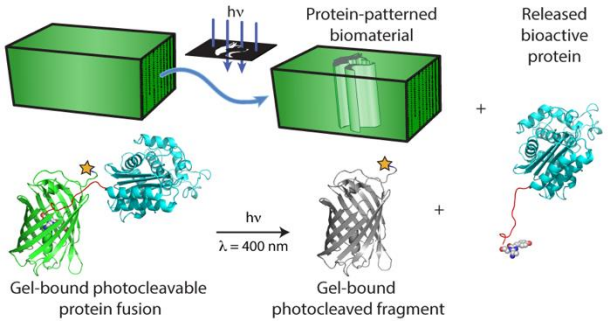
44. Hahn, M. S., Miller, J. S. & West, J. L. Laser scanning lithography for surface micropatterning on hydrogels. *Adv. Mater.* **17,** 2939–2942 (2005).

45. DeForest, C. A. & Anseth, K. S. Photoreversible patterning of biomolecules within click-based hydrogels. *Angew. Chemie Int. Ed.* **51,** 1816–1819 (2012).

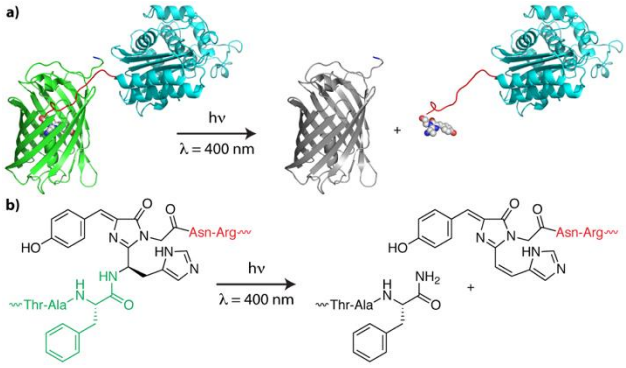
46. Arakawa, C. K., Badeau, B. A., Zheng, Y. & DeForest, C. A. Multicellular Vascularized

Engineered Tissues through User-Programmable Biomaterial Photodegradation. *Adv. Mater.* **29,** 1703156 (2017).

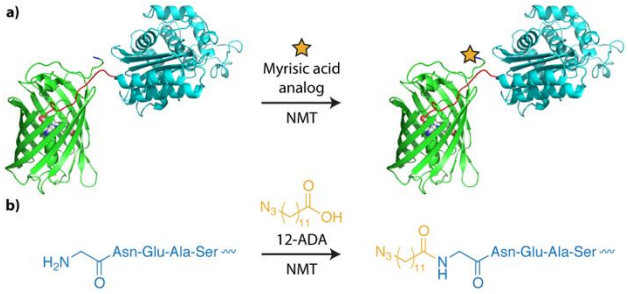
4.6 Figures



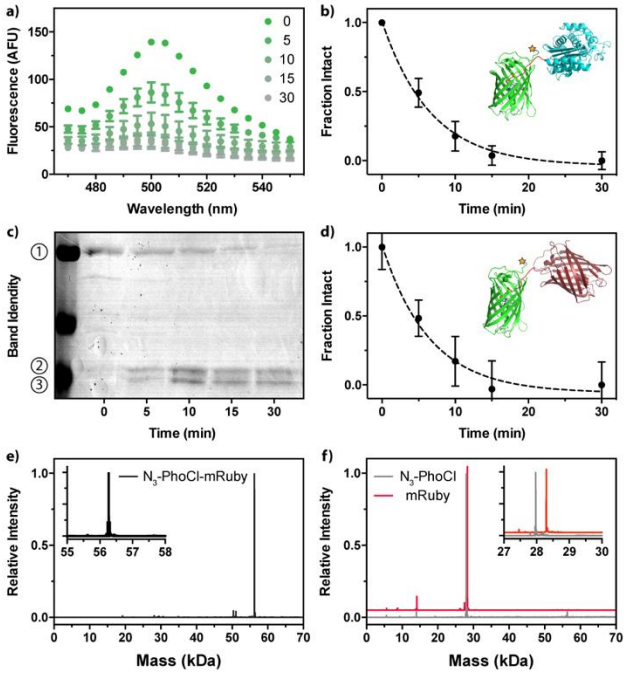
4.6.1 Graphical Abstract



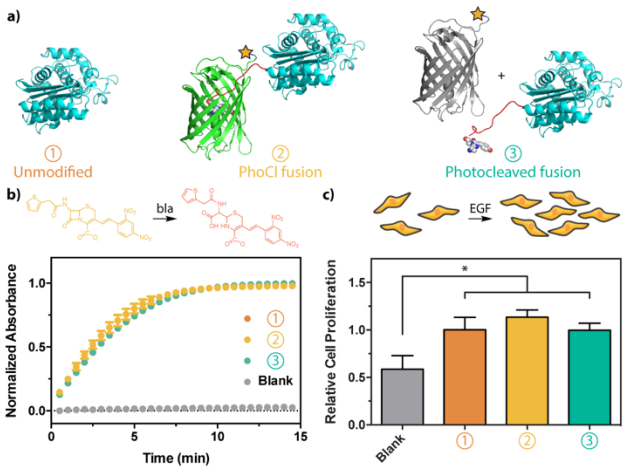
4.6.2 Scheme 1. (a) Photocleavable fusion proteins undergo irreversible photoscission near PhoCl's C-terminus in response to cytocompatible violet light (λ = 400 nm), which is accompanied by PhoCl's loss of green fluorescence. The photoreleased protein of interest is shown in blue. (b) Photoinduced β-elimination of its matured chromophore results in cleavage of PhoCl's polypeptide backbone.



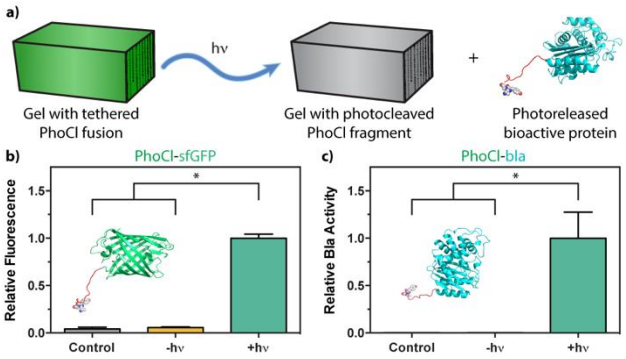
4.6.3 Scheme 2. (a) NMT-catalyzed myristoylation of PhoCl fusions bearing the N-terminal GNEASYPL sequence with (b) 12-ADA yields site-specific installation of azido functionality.



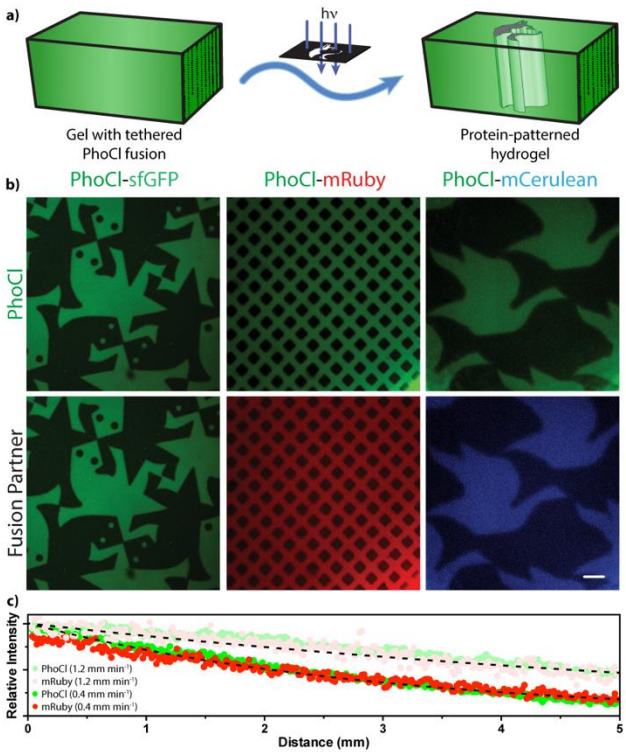
4.6.4 Figure 1. Validation and quantification of PhoCl photocleavage in response to violet light (λ = 400 nm, 10 mW cm^{-2}, 0 – 30 min). (a) N_3-PhoCl-bla fluorescence ($\lambda_{PhoCl,excitation}$ = 380 nm) was measured at various points throughout phototreatment. (b) Normalized fluorescence intensity ($\lambda_{PhoCl,emission}$ = 510 nm) quantification yields expected first-order decay behavior. (c) SDS-PAGE analysis of N_3-PhoCl-mRuby exposed to different amounts of light, with band ① denoting the intact fusion, ② mRuby, and ③ PhoCl. (d) Normalized band intensity quantification yields the first-order decay constant. (e) Whole protein mass spectrometry indicated high sample purity of N_3-PhoCl-mRuby. (f) Complete photocleavage of N_3-PhoCl-mRuby results in scission of PhoCl's chromophore and two distinct protein fragments of expected mass.



4.6.5 Figure 2. Intact and cleaved PhoCl fusions display native bioactivity. (a) Activities of the PhoCl chimera and its photocleaved products were compared with the native protein. (b) Bla activity was determined by its ability to degrade a chromogenic nitrocefin substrate, which changes from yellow to red upon β-lactam cleavage. Time-course spectrophotometric analyses (λ_{abs} = 386 nm) indicate nitrocefin degradation for all species. (c) EGF bioactivity was determined based on its ability to stimulate fibroblast cell proliferation as quantified by increased dsDNA synthesis and content.

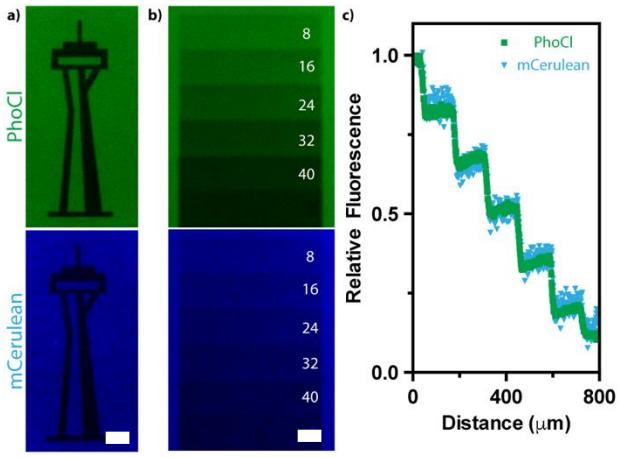


4.6.6 Figure 3. Proteins photoreleased from biomaterials remain active. (a) Photosensitive chimeras were tethered site-specifically to PEG-based hydrogels. Protein activity within the gel supernatant was assessed for PhoCl-containing samples kept in the dark or exposed to light (λ = 400 nm, 10 mW cm^{-2}, 30 min), as well as for control samples lacking protein. (b) Fluorescence measurements were utilized to assess sfGFP release. (c) A colorimetric nitrocefin assay was exploited to confirm functional bla release. Protein activities were normalized to the light-treated conditions.

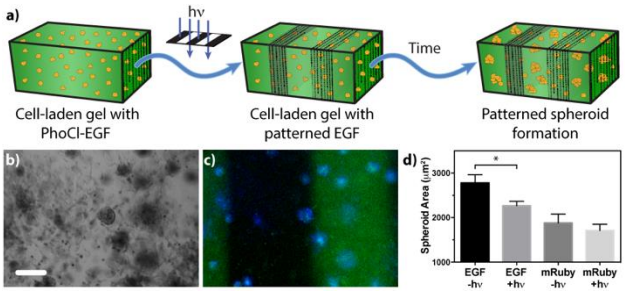


4.6.7 Figure 4. Protein-patterned gels from mask-based lithography. (a) Gels functionalized site-specifically with N_3-PhoCl-sfGFP, N_3-PhoCl-mRuby, or N_3-PhoCl-mCerulean were exposed to masked violet light (λ = 400 nm, 10 mW cm^{-2}, 30 min) to trigger PhoCl cleavage in user-defined 2D patterns through the 3D gel thickness. (b) Fluorescence from PhoCl (green) colocalized with that from sfGFP (green), mRuby (red), or mCerulean (blue). Light-treated gel regions appeared black, owing to patterned photorelease of the fluorescent protein and the PhoCl's cleavage-associated loss of fluorescence. Images correspond to a single z-slice from stacks acquired by fluorescent confocal microscopy. Scale bar = 200 μm. (c) By subjecting hydrogels modified with N_3-PhoCl-mRuby to a linear gradient of light exposure (λ = 400 nm, 10 mW cm^{-2}), well-defined gradients of protein photorelease are achieved across relatively large distances. The slope of resultant protein gradient is readily controlled by tuning that of the light, here by adjusting the

velocity by which an opaque photomask is translated over the gel (0.4 and 1.2 mm min^{-1}). Plots correspond to PhoCl and mRuby fluorescent intensity along the direction of photomask translation.



4.6.8 Figure 5. Protein-patterned gels from laser-scanning lithography. (a) Gels containing tethered N_3-PhoCl-mCerulean were selectively exposed to rastered laser scanning using a 405 nm diode laser line (λ = 405 nm, 5% laser power, scan speed = 200 Hz, 64 scan repeats) in regions of interest corresponding to the silhouetted outline of the Seattle Space Needle. Fluorescence from PhoCl (green) colocalized with that from mCerulean (blue). Light-treated gel regions appeared black, owing to patterned photorelease of the fluorescent protein and the PhoCl's cleavage-associated loss of fluorescence. Images correspond to a single z-slice from stacks acquired by fluorescent confocal microscopy. (b) By varying the number of scan repeats (0 – 64) for different regions within a single imaging window, stepped protein gradients were generated. Plots correspond to PhoCl and mCerulean fluorescent intensity along the gradient. Scale bar = 100 μm in (a) and (b).



4.6.9 Figure 6. Controlling anisotropic cell proliferation through patterned EGF. (a) HeLa cells were encapsulated in gels uniformly functionalized with N_3-PhoCl-EGF or N_3-PhoCl-mRuby. After one day, proteins were photoreleased in line patterns (400 μm wide features) using mask-based lithography. Cells were maintained in culture for a total of 14 days in the presence of the patterned protein prior to analysis. (b) Brightfield and (c) fluorescent confocal imaging (blue = nuclei, green = PhoCl-EGF) of EGF gels display increased spheroid growth and a larger cell density in protein-tethered regions. Scale bars = 200 μm. (d) Image analysis illustrates a statistically significant variation in spheroid size in functionalized vs unfunctionalized regions for EGF gels, but not for those containing patterned mRuby.

Chapter 5

5.1 Conclusions and future perspective

In this dissertation, the versatility and power of chemoenzymatic methods for protein modification were applied for the creation of functional biomaterials. In Aim 1, the power of chemoenzymatic protein modification was harnessed to create site-specifically modified proteins with high bioactivity. Synthetic peptides containing clickable handles and photoreactive groups were appended to proteins using the Sortase A enzyme to create biomaterials with reversible biochemical patterns for Aim 2. For Aim 3, the PhoCl protein was exploited as a genetically encoded photocleavable linker for hydrogel patterning. In this section, the results as they pertain to these aims will be summarized, as well as a future prospective on the technologies.

In Chapter 1, a thorough review of the literature for protein modification examined the various techniques that exist for site-selective modification of proteins. While small-molecule-based techniques have come a long way in improving their chemical and regioselectivity, enzymes have proven to be the most capable of performing modifications with the highest conversions and selectivity. Of the enzymatic methods reviewed, Sortase A and N-myristoyltransferase stood out as the most powerful for applying to biomaterials. Sortase has the ability to append nearly any functionality, as long as an N-terminal polyglycine nucleophile is maintained, making it a versatile tool in appending the chemical handles, such as azides, aldehydes, and *ortho*-nitrobenzylesters commonly used in the biomaterials community. NMT has the unique ability to quantitatively install reactive handles including azides onto the N-terminus co-translationally. This is particularly beneficial as it requires no post-synthetic steps to functionalize a protein. Having the azide installed

during protein expression allows for purification of a site-specifically modified photocleavable group that can be tethered into materials directly from *E. coli*.

In Chapter 3, the transpeptidase activity of Sortase A was used to create a library of homogenously modified proteins with a plethora of non-natural functionality. Sortase was expressed and purified the conjugated proteins using the STEPL method to give homogenous protein populations for reversible immobilization into biomaterials. To address Aim 1, the effects of protein modification on activity were investigated, and compared it to NHS treatment at different molar excesses. In every case, sortagged proteins retained more of their bioactivity than NHS modified counterparts, often completely retaining their bioactivity. To address Aim 2, photopatterning was performed to enable high spatiotemporal resolution of biomacromolecules in the presence of live cells. This control enabled unprecedented regulation of biological functions in subcellular volumes. This work has already been expanded upon in the DeForest lab by using light-sensitive proteins and peptides that respond to multiple stimuli.

In Chapter 4, the PhoCl protein was modified with the NMT enzyme for N-terminal azide incorporation, enabling its site-specific tethering to materials. For Aim 3, a library of azide-modified PhoCl-containing fusion proteins were generated that could be photoreleased from hydrogels using cytocompatible violet light. PhoCl fusion proteins could be immobilized and released from gels while retaining their activity. Cleavage was found to be predictable, with a half-life of approximately 5 minutes (λ = 400, 10 mw/cm^2). High-resolution protein patterns were formed using mask-based photolithography in conjunction with a violet LED light source as well through laser-scanning lithography utilizing the 405 nm diode laser line commonly available on a confocal microscope.

Despite the great success of chemoenzymatic protein modification in the generation of functional biomaterials, the technology was only used for the narrow goal of patterning proteins into materials with very specific chemistries. There are numerous ways that the technology could be applied to further extend the usefulness of the technique. As one example, protein immobilization and release was only performed on PEG-based hydrogels, despite the enzymatic techniques being amenable to modification of nearly any polymeric network with functional handles.

One extension would be to apply chemoenzymatic protein modification for tethering functionalized proteins into natural materials. Collagen and Matrigel are widely used by biologists due to their cytocompatability and ease of use[1]. However, they are poorly defined heterogeneous mixtures without the convenient functional handles used in synthetic materials for controlling biochemical signal presentation. However, as collagen and Matrigel are at their core just a collection of proteins, installing clickable handles could be done using one of the small molecule or enzymatic techniques outlined in Chapter 1. This would impart the synthetic functionality that is so useful in engineered materials, while maintaining the natural biocompatibility of the mixtures.

Development of further red-shifted photochemistries would impart greater cytocompatibility and tissue penetrance, as well as potentially enable orthogonal immobilization and release strategies. Red-shifted photodegradable crosslinkers made out of coumarin and ruthenium have recently been developed[2,3]. Analogous photochemistries could be implemented into the peptides used during sortagging to create proteins responsive to different wavelengths of light. This could lead to dual-release materials, where sequential exposure to different wavelengths of light would allow independent release of multiple factors. This could be beneficial in cases where sequential release of different factors from a depot would be therapeutically beneficial, such

as in wound healing[4,5]. Similarly for PhoCl, optogenetic tools are continually being engineered to shift the absorbance properties of the chromophore and to alter the cleavage kinetics[6]. Newly evolved variants of this protein could have even further redshifted absorbance which would have even deeper tissue penetrance, and could be used orthogonally to the original PhoCl protein.

While the strategies employed in Chapter 3 enable reversible patterning, proteins can only be immobilized and released once since the original end groups are not regenerated following photoscission. Careful engineering of the photoreactive pendant groups in the gel could allow for full reversibility. For example, GGG pendants with a free N-terminal amine could be stochastically incorporated into the gel. Diffusing in NPPOC-LPETG and sortase would functionalize the gel with NPPOC functionality for photopatterning. End groups could be regenerated by swelling in more sortase and NPPOC-LPETG since the LPETG motif is conserved following conjugation.

As an alternative approach, there have been numerous recent developments in optogenetics in developing light sensitive protein partners that undergo reversible dimerization in response to light, such as LovTRAP, Dropna, and Phy-Pif[7–10]. Use of these photosensitive proteins as the photoreactive linker would enable fully reversible immobilization. As an added benefit, many of these proteins are extremely sensitive to light, and the wavelengths of light used tend to be redshifted. These optogenetic tools could be modified with the chemoenzymatic methods described to create the next generation of light-responsive biomaterials. Careful engineering of the modification site will be necessary to fully harness these proteins.

While the STEPL system enables site-specific modification of full-length proteins, this modification is limited to the C-terminus. Recent investigations into natural sortase activity has revealed sortase's ability to modify interior sections of the protein through the pilin domain[11]. Incorporation of this domain into permissive internal loop should enable fully site-specific labeling

of a protein of interest. This would be beneficial in the cases of proteins whose termini are inaccessible, or where a native C-terminus are critical for function.

One limitation of this work is that in order for proteins to be conjugated and released, they must diffuse through the hydrogel matrix. While this diffusion time can be modulated by changing the mesh size of the network and the geometry of the gels, cells are still being exposed to protein cues during this time. For extremely potent signaling molecules where cells will respond to only a brief exposure to the cue, this diffusion time could be sufficient to confound the effects of the immobilized protein. Recent work has looked into including photocages on protein active domains to inhibit their function until exposed to light[12,13]. Protein activity could be blocked during diffusion, the activity of the immobilized cue could be restored spatiotemporally through photoscission of the photocage.

One advantage of the PhoCl system that was not fully utilized is the green-to-red fluorescent transition that occurs upon photocleavage. Reengineering of the fusion proteins to put the azide-tagged PhoCl domain on the C-terminus would allow for labeling of soluble vs tethered cues based upon the color of fluorescence. This could be of particular interest in tracking biological processes and the differential cellular response to immobilized and soluble cues.

The high protein purities achieved using STEPL lends itself towards the generation of new purification strategies. The transpeptidase activity of sortase could be used by encoding an N-terminal HHHHHHLPETG tag onto a protein of interest. When performing IMAC, sortase and a polyglycine could be added to cleave the protein from the resin leaving only an N-terminal glycine leftover. This would give higher purities than traditional IMAC since instead of eluting with a histidine analog, imidazole, an orthogonal reagent is added to specifically release only LPETG tagged proteins. In addition, light responsive protein tags could be added for elution strategies that

do not rely on adding an exogenous small molecule, allowing for elution directly into protein storage buffers without the need for dialysis or buffer exchange. For example, the PhoCl fusion proteins used in this work could be eluted during IMAC by violet light, giving pure protein species with a 13-amino acid tag left over.

The micron-scale 3D spatial resolution of the patterning techniques suggests the possibility to be used as an information storage scheme. Newer holographic storage media can store approximately 1 Tb/in^2, and it has been suggested that the theoretical limit for magnetic disk storage will cap out at 5 Tb/in^2 [14,15]. One major limitation of traditional information storage is that it is limited to a 2D surface. However, the optical clarity of hydrogels allow for information to be stored and retrieved in full 3D space, greatly expanding the theoretical limit for information storage in a given volume. Assuming a 1 micron resolution, the patterning technique that I employed could store ~600 Mb/in^2. Expanding this to the third dimension would give a volumetric storage of ~6 Tb/in^3 assuming a z resolution of 2.5 microns. A Blu-ray disc by comparison, has a diameter of 4.7 inches and can store up to 128 GB giving an information density of 7.4 Gb/in^2. However, multiple fluorophores and greyscale pixels could be patterned to increase the informational density by the number of fluorophores immobilized or the number of greyscale values that could be distinguished. For example, if three fluorophores were immobilized the informational density would be 18 Tb/in^3. If 10 greyscale values could be distinguished for these fluorophores the informational density would be 180 Tb/in^3, well beyond the theoretical limit for magnetic disk storage. Reversible patterning would make the system read/writeable.

Employing the logical linkers developed in the DeForest lab could impart interesting properties to the stored data[16]. For example, including an AND gate would make an "invisible ink" where the information is only revealed upon treatment with the second cue. If the bond tethering

the fluorophore to the matrix is cleaved using the same wavelength of light as for excitation, a "destroy upon reading" system could be created. This would be useful in creating watermarks that would show whether a message has been read or not, a very tantalizing prospect for cryptography.

In summary, utilizing chemoenzymatic methods for the site-specific modification of proteins has enabled the facile creation of spatiotemporally decorated proteins within biomaterials. The versatility of the Sortase A transpeptidase was used to create a library of proteins appended with synthetic peptide handles. Sortagged proteins retained their activity and were able to dictate subtle cellular responses such as MAPK signaling and vesicle internalization. Labeling a photocleavable protein, PhoCl, with an N-terminal azide via the N-myristoyltransferase enzyme allowed for co-translational functionalization of photoreleasable fusion proteins in *E. coli*. Conjugated fluorophores, enzymes, and growth factors were able to be tethered within PEG hydrogels while retaining their activity and released on demand using light. These results have important implications for the creation "smart" biomaterials that are able to fully harness the exquisite specificity and diverse functionality of full-length proteins.

5.2 References:

1. Caliari, S. R. & Burdick, J. A. A practical guide to hydrogels for cell culture. *Nat. Methods* **13,** 405–414 (2016).

2. Azagarsamy, M. A. & Anseth, K. S. Wavelength-controlled photocleavage for the orthogonal and sequential release of multiple proteins. *Angew. Chem. Int. Ed. Engl.* **52,** 13803–7 (2013).

3. Rapp, T. L., Highley, C. B., Manor, B. C., Burdick, J. A. & Dmochowski, I. J. Ruthenium-Crosslinked Hydrogels with Rapid, Visible-Light Degradation. *Chem. - A Eur. J.* **24,** 2328–2333 (2018).

4. Richardson, T. P., Peters, M. C., Ennett, A. B. & Mooney, D. J. Polymeric system for dual growth factor delivery. *Nat. Biotechnol.* **19,** 1029–1034 (2001).

5. Sohier, J. *et al.* Dual release of proteins from porous polymeric scaffolds. *J. Control. Release* **111,** 95–106 (2006).

6. Andresen, M. *et al.* Photoswitchable fluorescent proteins enable monochromatic multilabel imaging and dual color fluorescence nanoscopy. *Nat. Biotechnol.* **26,** 1035–1040 (2008).

7. Dagliyan, O., Dokholyan, N. V. & Hahn, K. M. Engineering proteins for allosteric control by light or ligands. *Nat. Protoc.* **14,** 1863 (2019).

8. Toettcher, J. E., Gong, D., Lim, W. A. & Weiner, O. D. Light Control of Plasma Membrane Recruitment Using the Phy–PIF System. in *Methods in enzymology* **497,** 409–423 (2011).

9. Zhou, X. X., Chung, H. K., Lam, A. J. & Lin, M. Z. Optical control of protein activity by fluorescent protein domains. *Science* **338,** 810–4 (2012).

10. Wang, H. *et al.* LOVTRAP: an optogenetic system for photoinduced protein dissociation. *Nat. Methods* **13,** 755–758 (2016).

11. McConnell, S. A. *et al.* Protein Labeling via a Specific Lysine-Isopeptide Bond Using the Pilin Polymerizing Sortase from *Corynebacterium diphtheriae*. *J. Am. Chem. Soc.* **140,** 8420–8423 (2018).

12. Nguyen, D. P. *et al.* Genetic Encoding of Photocaged Cysteine Allows Photoactivation of TEV Protease in Live Mammalian Cells. *J. Am. Chem. Soc.* **136,** 2240–2243 (2014).

13. Zhao, J., Lin, S., Huang, Y., Zhao, J. & Chen, P. R. Mechanism-Based Design of a Photoactivatable Firefly Luciferase. *J. Am. Chem. Soc.* **135,** 7410–7413 (2013).

14. Mallary, M., Torabi, A. & Benakli, M. One terabit per square inch perpendicular recording conceptual design. *IEEE Trans. Magn.* **38,** 1719–1724 (2002).

15. Rea, C. *et al.* Areal-Density Limits for Heat-Assisted Magnetic Recording and Perpendicular Magnetic Recording. *IEEE Trans. Magn.* **52,** 1–4 (2016).

16. Badeau, B., Comerford, M., Arakawa, C., Shadish, J. & DeForest, C. Engineered modular biomaterial logic gates for environmentally triggered therapeutic delivery. *Nat. Chem.*

Chapter 6

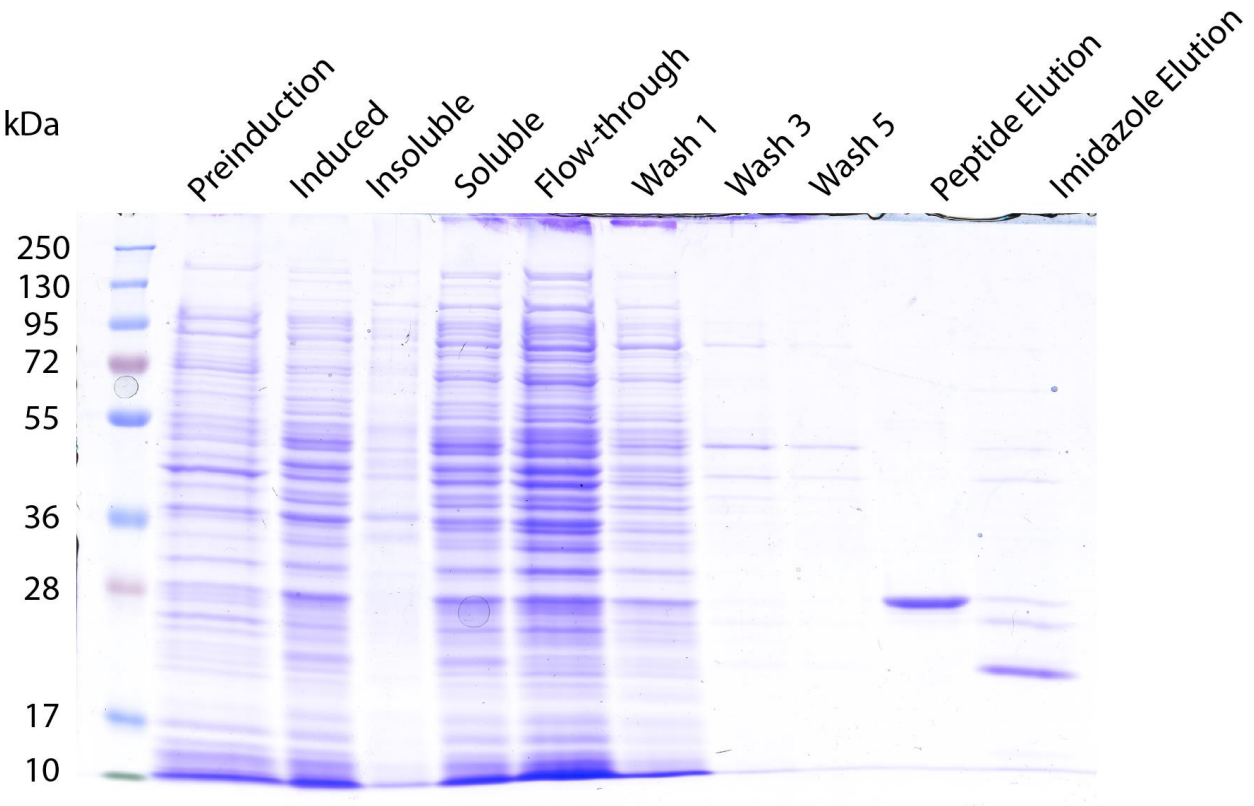
Appendix A1: Supplementary Information for

Bioactive Site-Specifically Modified Proteins for 4D Patterning of Gel Biomaterials

Jared A. Shadish, Gabrielle M. Benuska, Cole A. DeForest

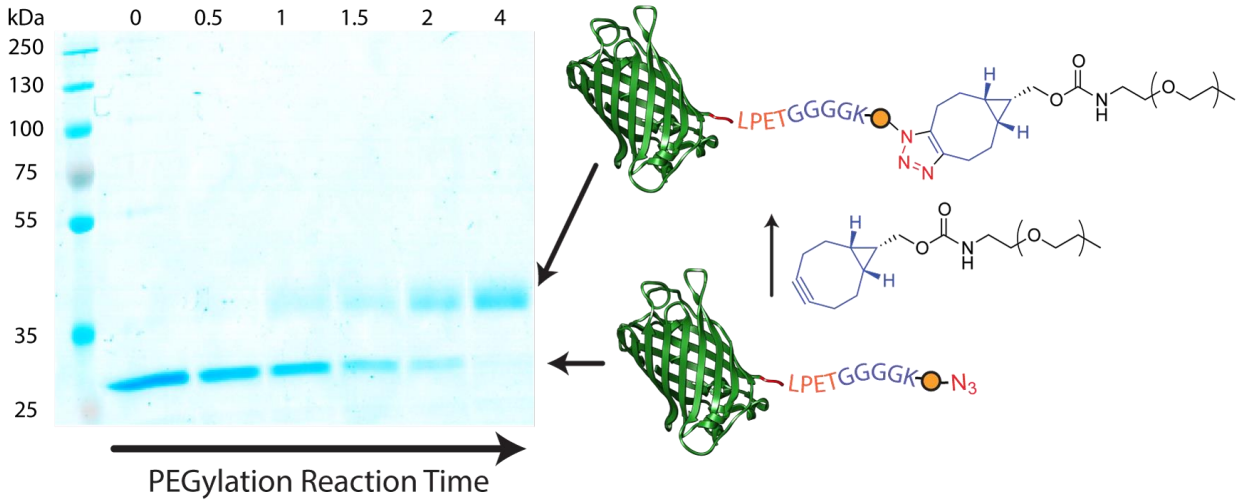
6.1 Supplementary Figures

Supplementary Figure 1:



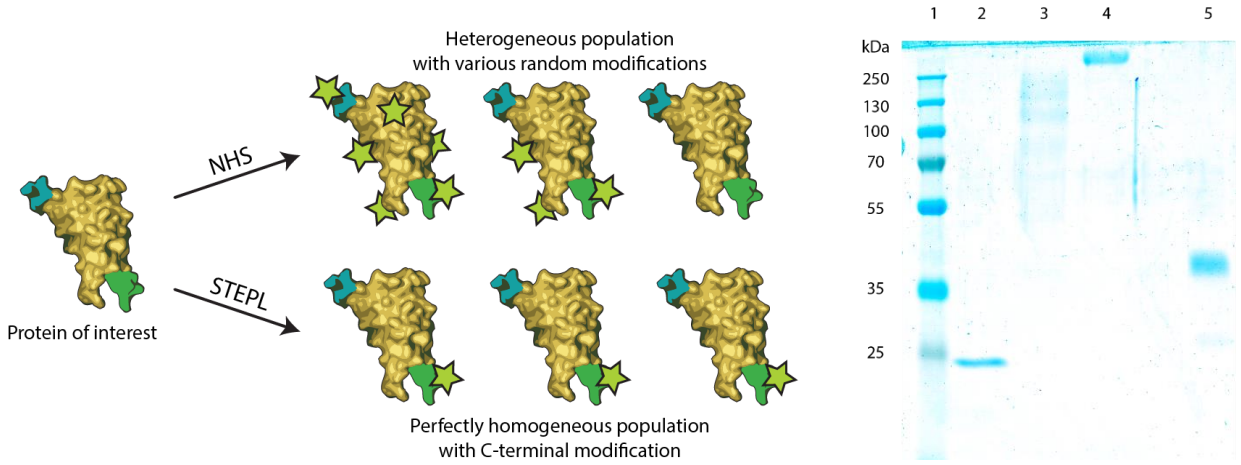
6.1.1 STEPL purification of EGFP-N3. EGFP-N_3 is generated through STEPL involving EGFP and H-GGGGDDK(N_3)-NH_2 peptide. The EGFP-SrtA-6xHis fusion appears as a ~50 kDa band post induction. After chromatographic immobilization on Ni-NTA, the resin is incubated with the polyglycine probe; the sortagged protein is collected in the eluent, while the SrtA-6xHis protein remains column-bound. Subsequent resin washes with imidazole displaces the SrtA-6xHis and regenerates the Ni-NTA. The peptide-eluted sortagged protein (EGFP-N_3) is an exceptionally pure (>95%) homogenous species appearing as a band ~29 kDa. The collected sortase enzyme appears as a band at ~17 kDa. Similar results were obtained for all protein-peptide conjugates purified throughout the reported studies.

Supplementary Figure 2:



6.1.2 Assessing purity and reactivity of sortagged proteins by SDS-PAGE gel shift analysis. To assess the ability for STEPL to generate homogenous protein samples with desired reactivity, a sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) gel shift assay[3] was performed C-terminal PEGylation by SPAAC. EGFP-*o*NB-N_3 (1 µg) was reacted with mPEG-BCN (100x, Supplementary Methods) in PBS (10 µL) for varying lengths of time (0-4 hr) at room temperature. The complete disappearance of the starting protein band after 4 hours of reaction, accompanied by the simultaneous appearance of a new band upshifted by the average molecular weight of a single PEG chain, is indicative of quantitative functionalization of purified proteins using STEPL. Similar results were obtained for $n = 3$ independent samples.

Supplementary Figure 3:



6.1.3 Comparing purity and reactivity of sortagged proteins with those randomly modified by NHS chemistry. To determine the extent of functionalization of EGFP randomly modified with azide moieties by NHS-chemistry, an SDS-PAGE gel shift assay was performed (Supplementary Figure 2**)**. Sortagged EGFP-N_3 as well as EGFP reacted with various amount of N_3-OSu (10, 100, 1000x, based on NHS:protein molar equivalents) were each treated with mPEG-BCN (100x) in PBS prior to analysis. PEGylation events resulting in protein band upshifting provide a lower bound for estimating the number of azides introduced.

Lane 1: Protein ladder of known molecular weights

Lane 2: EGFP + N_3-OSu (10x) + mPEG-BCN (100x)

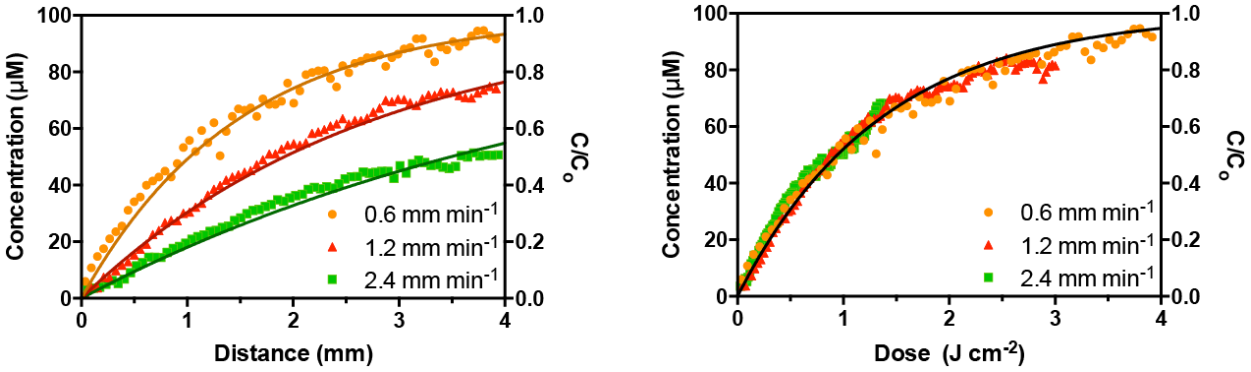
Lane 3: EGFP + N_3-OSu (100x) + mPEG-BCN (100x)

Lane 4: EGFP + N_3-OSu (1000x) + mPEG-BCN (100x)

Lane 5: EGFP-N_3 + mPEG-BCN (100x)

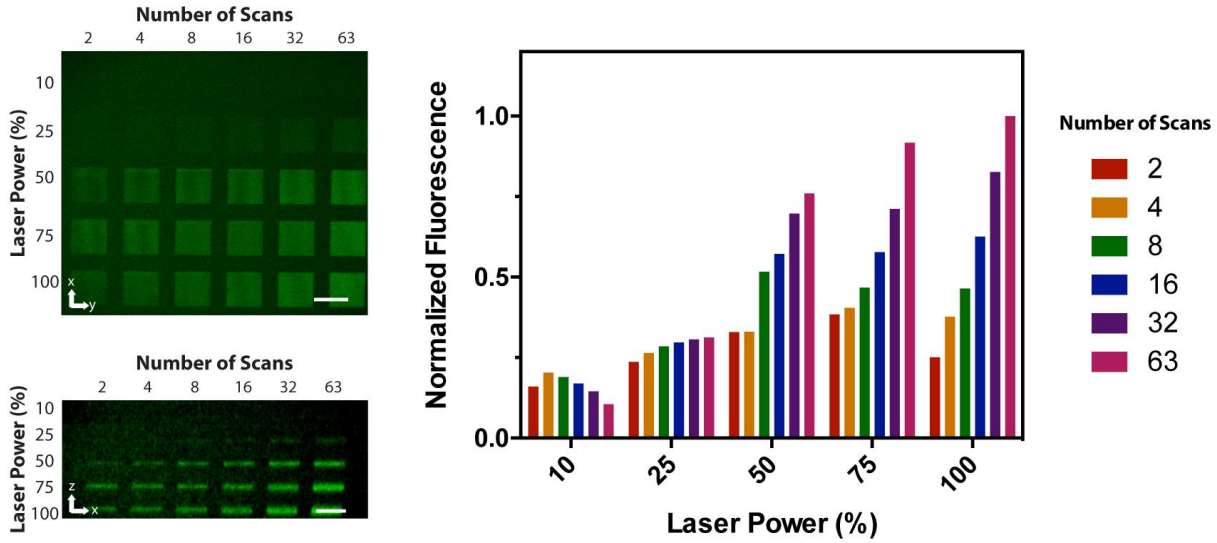
NHS treatment with N_3-OSu yielded a heterogeneous population with many species of varying azide content. 10-fold molar excess yielded no appreciable labeling; 100-fold excesses gave significant labeling but in an ill-defined manner; 1000-fold excesses resulted in uniform tagging (likely of every primary amine on the protein), whereby the extent of PEGylation prevented it from migrating far into the gel. By contrast, the sortagged variant showed only a single upshift indicating a perfectly homogenous population. Similar results were obtained for $n = 3$ independent samples.

Supplementary Figure 4:



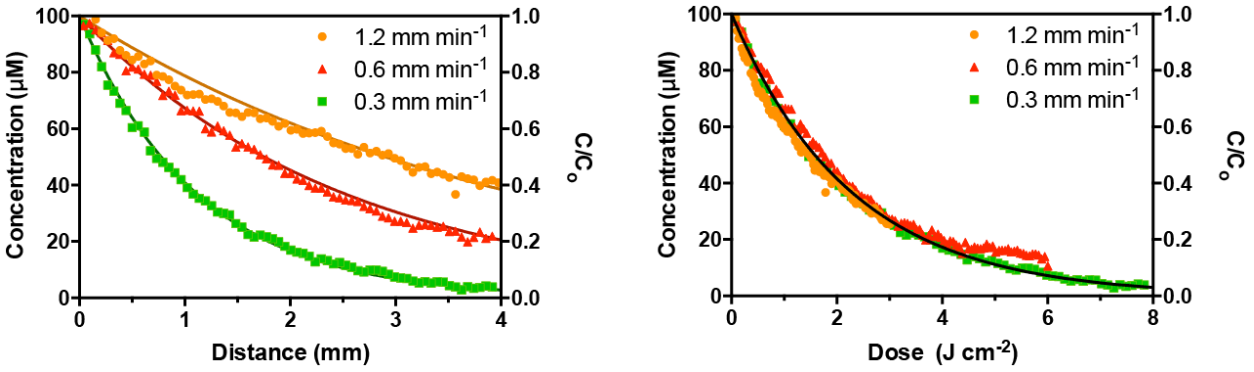
6.1.4 Dose response for photo-mediated immobilization of proteins. Dose represents the total amount of energy delivered to the system and was calculated as the product of exposure time and light intensity. Connected line represents predicted concentration based on NPPOC cleavage photokinetics[1]. Experiments were performed on a single gel for each gradient light condition.

Supplementary Figure 5:



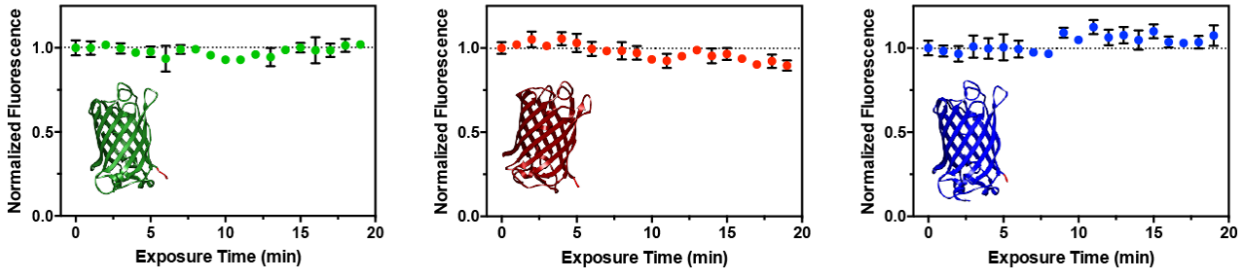
6.1.5 Quantification of 3D protein patterning via photomediated oxime ligation. Multiphoton laser-scanning lithographic techniques (λ = 740 nm) were exploited in the photomediated immobilization of EGFP-CHO within an N_3-TEG-ONH-NPPOC functionalized SPAAC-based gel over a variety of laser powers (10, 25, 50, 75, 100%) and scan repeats (2, 4, 8, 16, 32, 63). Here, 50 µm x 50 µm square region-of-interests (ROIs) were scanned at 2.5 µm z-increments for 10 µm before swelling in EGFP-CHO (100 µM) to create an array of protein-functionalized rectangular boxes, with each volume corresponding to a unique patterning condition. Samples were imaged using fluorescence microscopy (left), and quantification of immobilized protein concentration was performed for each light condition (right). Reported values correspond to the mean z-centered voxel fluorescence for each ROI across a single patterned sample. Optimal patterning conditions, maximizing total immobilized protein and z-resolution while minimizing patterning speed, were identified as 50% power with 16 scan repeats.

Supplementary Figure 6:



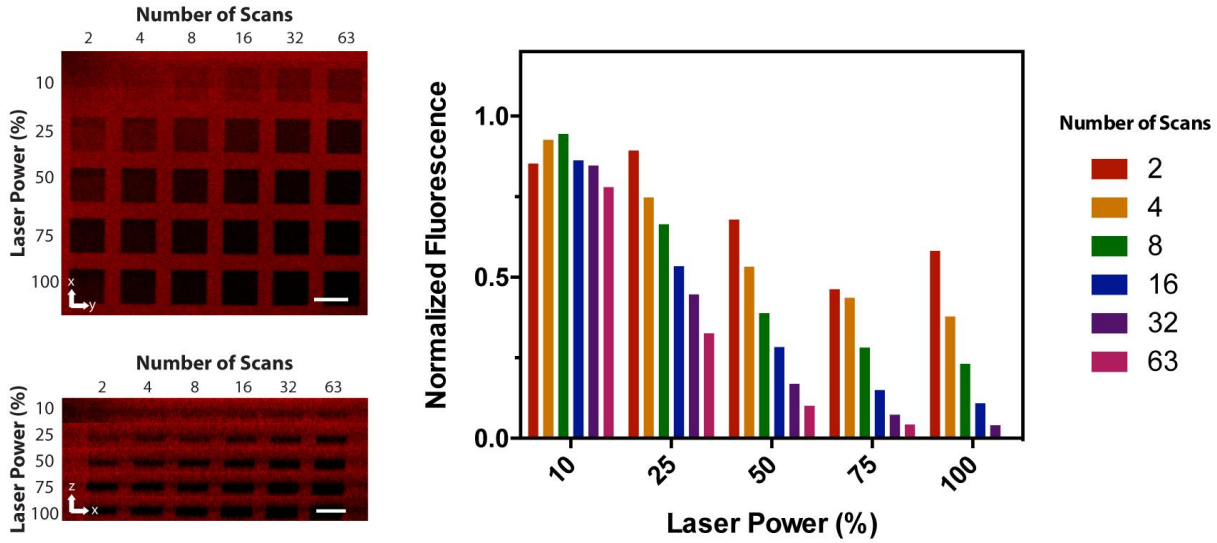
6.1.6 Dose response for photorelease of immobilized proteins. Dose represents the total amount of energy delivered to the system and was calculated as the product of exposure time and light intensity. Connected line represents predicted protein concentration based on the kinetics of *o*-nitrobenzyl ether (*o*NB) photocleavage[1]. Experiments were performed using a single gel for each gradient light condition.

Supplementary Figure 7:



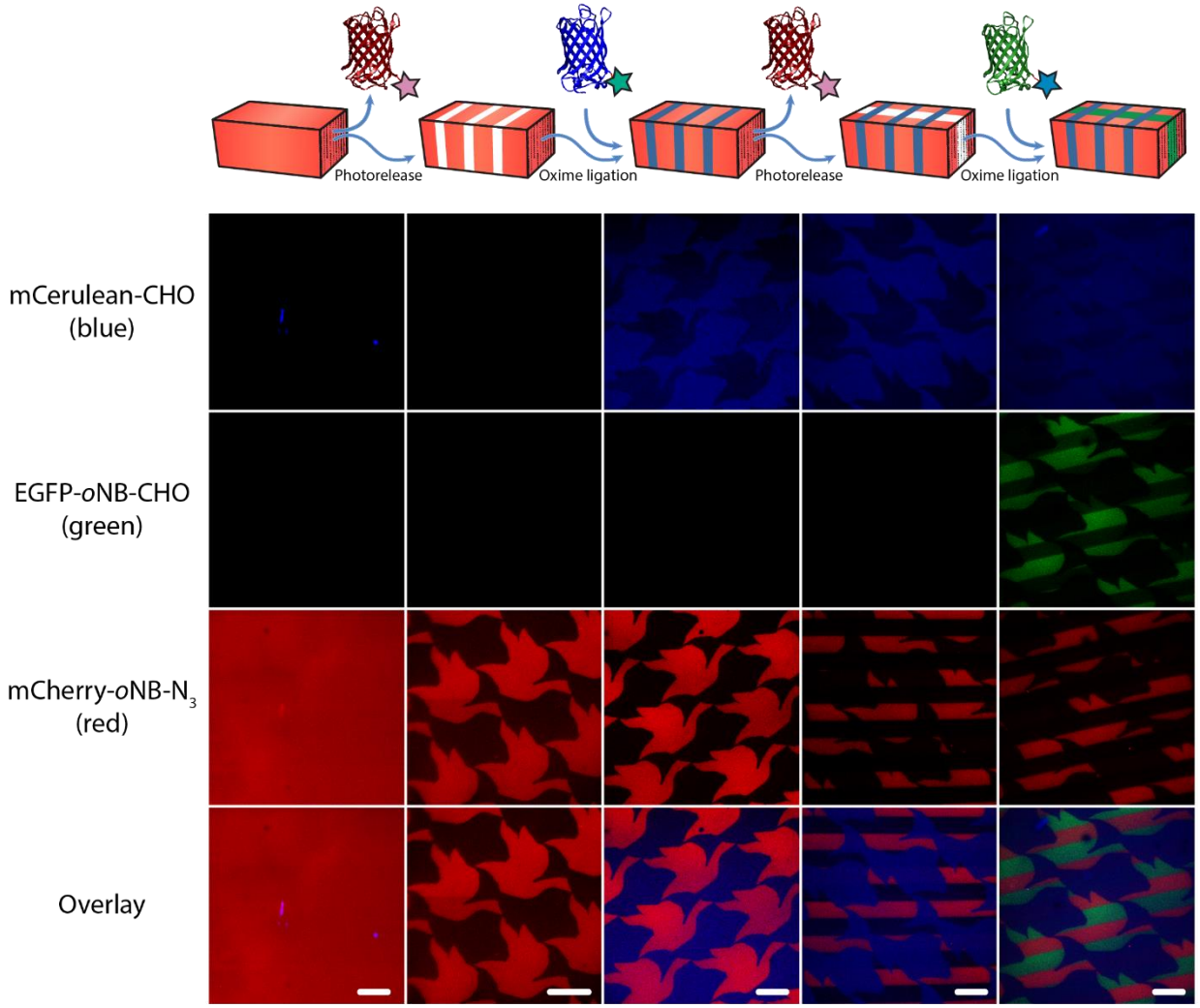
6.1.7 Assessing photobleaching of fluorescent proteins in response to UV light. To assess photobleaching kinetics of fluorescent proteins in response to patterning light conditions, samples of EGFP, mCherry, and mCerulean (each at 2 μM, in PBS) were exposed to far UV light (λ = 365 nm, 10 mW cm^{-2}) for varying amounts of time (0 – 20 min); fluorescence measurements were taken every minute and normalized to initial sample fluorescence. Analysis was performed in triplicate. Left plot corresponds to EGFP; middle to mCherry; right to mCerulean. Error bars correspond to the standard deviation about the mean for n = 3 independent samples. EGFP and mCerulean remained fully stable throughout light exposure, while mCherry showed minor photobleaching (~10%).

Supplementary Figure 8:



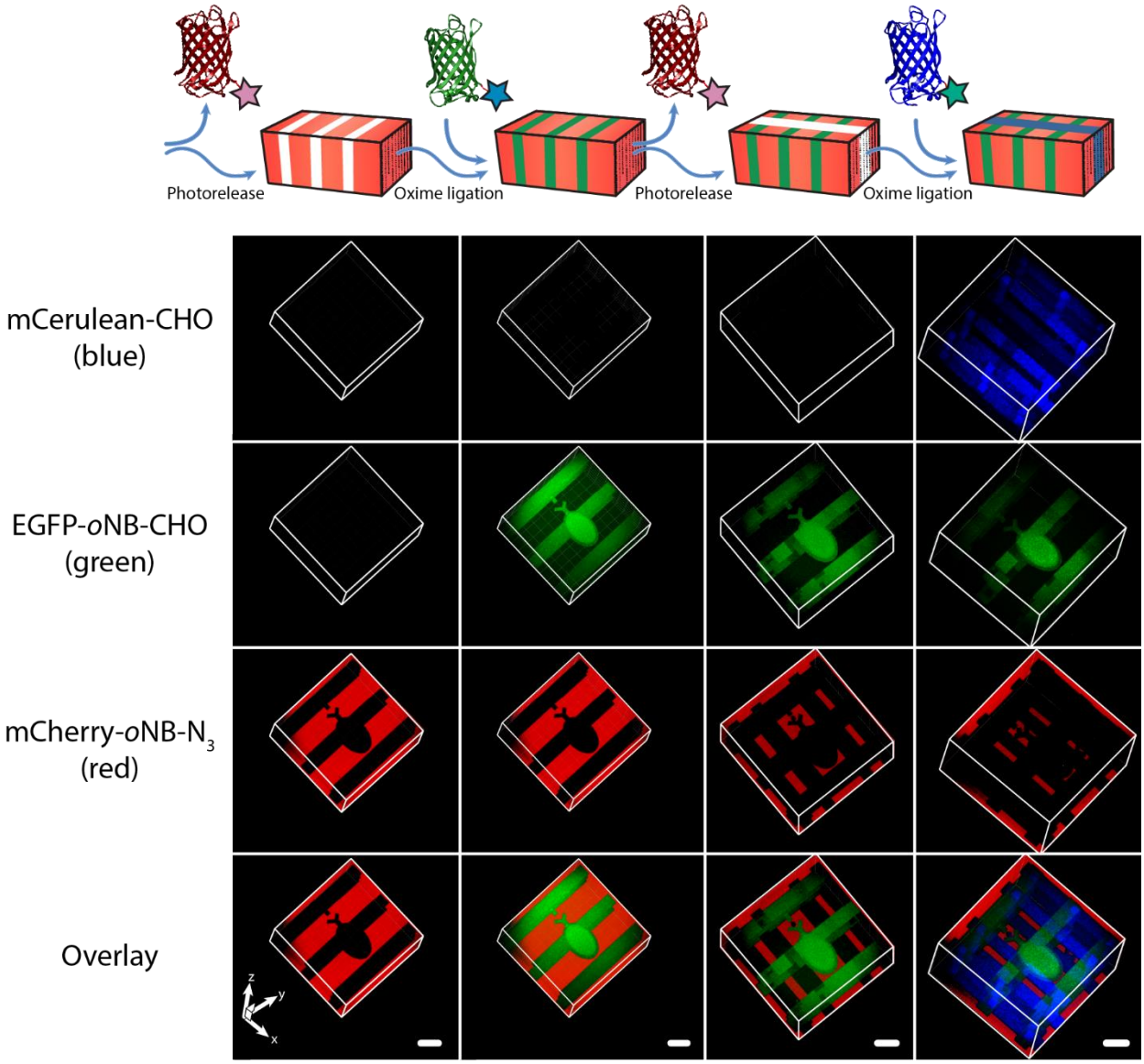
6.1.8 Quantification of 3D protein photoremoval by multiphoton laser scanning lithography. Multiphoton laser-scanning lithographic techniques (λ = 740 nm) were exploited to release immobilized mCherry-*o*NB-N_3 (initial concentration = 10 μM) from a SPAAC-based gel over a variety of laser powers (10, 25, 50, 75, 100%) and scan repeats (2, 4, 8, 16, 32, 63). Here, 50 μm x 50 μm square ROIs were scanned at 2.5 μm z-increments for 10 μm to create an array of protein-depleted rectangular boxes within gels, with each volume corresponding to a unique patterning condition. Samples were imaged using fluorescence confocal microscopy (left), and quantification of remaining immobilized protein concentration was performed for each light condition (right). Reported values correspond to the mean z-centered voxel fluorescence for each ROI across a single patterned sample. Optimal patterning conditions, maximizing total protein release and z-resolution while minimizing scan repeats, were identified as 50% power with 16 scans.

Supplementary Figure 9:



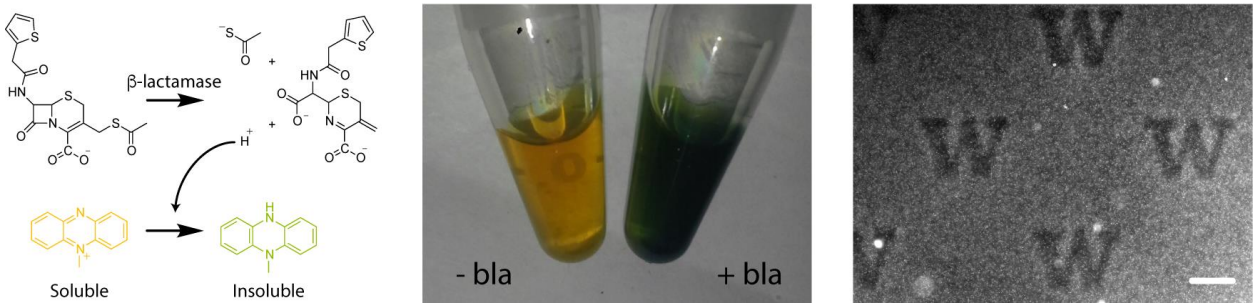
6.1.9 Split-color images of multicolor patterning generated by mask-based lithography. Images were taken throughout the photopatterning process as highlighted in Figure 4. For each stage of patterning, the red, green, and blue channels were separated to better show the localization of individual fluorescent proteins. Experiment was replicated (n = 3) with similar results. Scale bars = 100 μm.

Supplementary Figure 10:



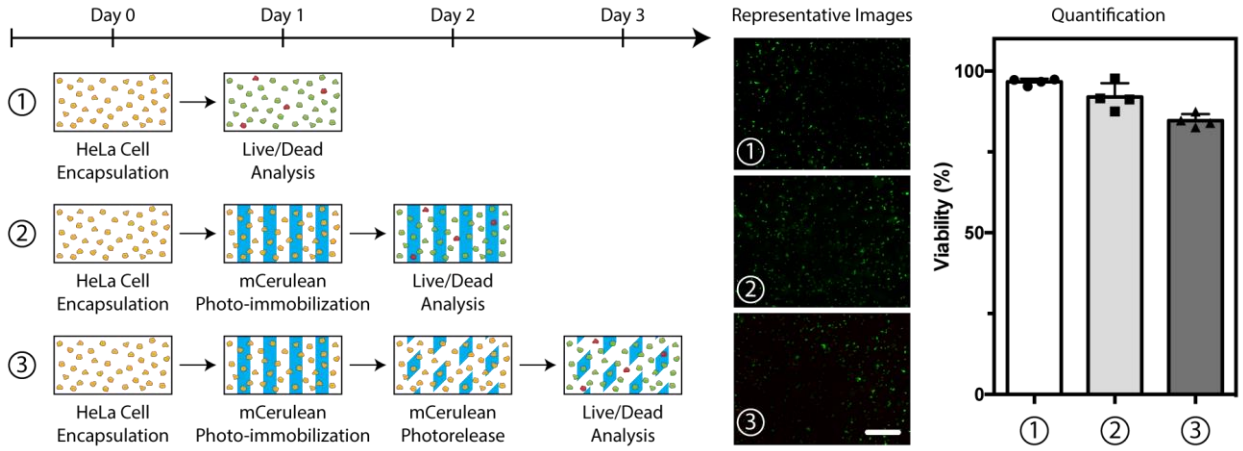
6.1.10 Split-color images of multicolor patterning by multiphoton laser scanning lithography. Images were taken throughout the photopatterning process as highlighted in Figure 4. For each stage of patterning, the red, green, and blue channels were separated to better show the localization of individual fluorescent proteins. Experiment was performed within a single gel. Scale bars = 100 μm.

Supplementary Figure 11:



6.1.11 Gel-immobilized bla remains bioactive. SPAAC-based gels uniformly functionalized with bla-*o*NB-N_3 (30 μM) were subjected to masked light (λ = 365 nm, 10 mW cm^{-2}, 10 min) through a chrome photomask containing transparent University of Washington Boundless 'W' logo features. Gels were equilibrated in PBS overnight prior to incubation with thioacetate cefalotin (5 mM, Supplementary Methods) and phenazine methosulfate (6.5 mM) in PBS (37 °C, 1 h). The gels were then washed with PBS (3x, 15 min) prior to visualization by phase contrast microscopy. Gels showed expected deposition of lightly colored precipitated product in regions where bla protein remained tethered. Experiment was replicated (n = 6) with similar results. Scale bar = 250 μm.

Supplementary Figure 12:



6.1.12 Cell viability throughout encapsulation and protein photopatterning. To determine cell viability throughout encapsulation and protein photopatterning reactions, a series of three experiments were performed:

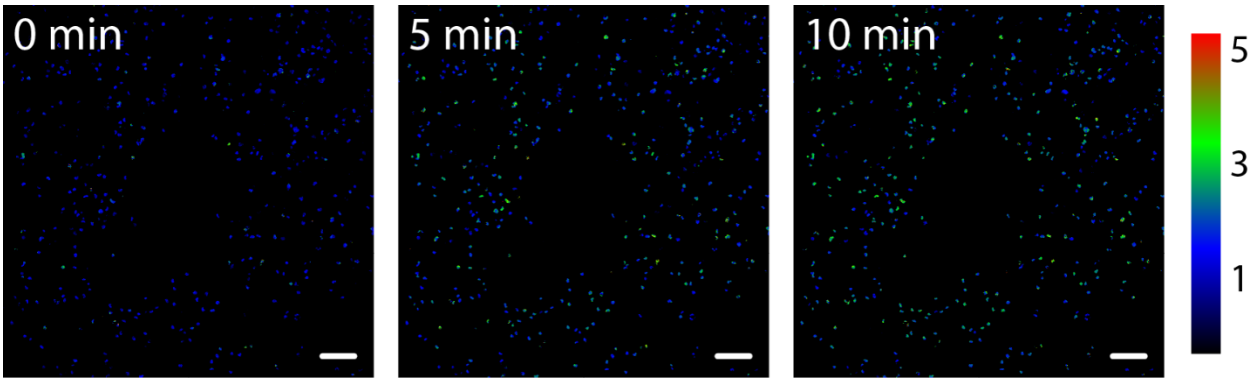
① HeLa cells were encapsulated (2 x 10^6 cells mL^{-1}) in SPAAC-based gels (10 wt%, 5 µL) crosslinked with N_3-GGRGDSPGGPQGIWGQGK(N_3)-NH_2 (Supplementary Methods, 8 mM) containing N_3-TEG-ONH-NPPOC (100 µM) attached to azide-functionalized glass slides (Supplementary Methods). Cell viability was assessed one day after encapsulation.

② HeLa cells were encapsulated in gels as in experiment ①. One day after encapsulation, gels were exposed to masked UV light (λ = 365 nm, 10 mW cm^{-2}, 10 min) through a slitted photomask containing 400 µm wide line features. Cell-laden gels were incubated in DMEM supplemented with FBS (10%), penicillin/streptomycin (1%), and mCerulean-CHO (0.3 µM). Following overnight conjugation, the gels were transferred to media lacking mCerulean-CHO. Cell viability was assessed one day after gel patterning was complete.

③ HeLa cells were encapsulated in gels as in experiment ①. One day after encapsulation, gels were exposed to masked UV light (λ = 365 nm, 10 mW cm^{-2}, 10 min) through a slitted photomask containing 400 µm wide line features. Cell-laden gels were incubated in DMEM supplemented with FBS (10%), penicillin/streptomycin (1%), and mCerulean-*o*NB-CHO (0.3 µM). Following overnight conjugation, the gels were transferred to media lacking mCerulean-*o*NB-CHO. The gels were again exposed to masked UV light (λ = 365 nm, 10 mW cm^{-2}, 10 min) through a slitted photomask containing 400 µm wide line features, this time rotated 45° from the initial protein pattern. Cells were maintained in mCerulean-free media for an additional day prior to assess viability.

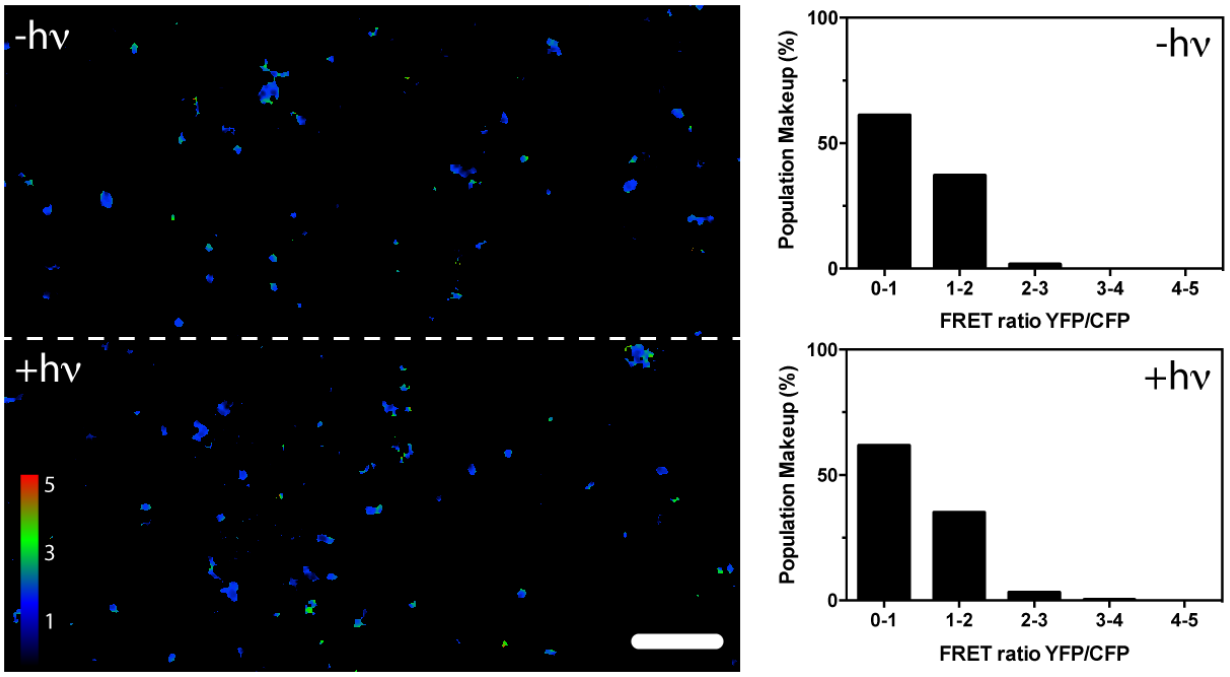
Cell viability was assessed using the LIVE/DEAD® Viability/Cytotoxicity Assay Kit (Thermo Fisher) with Calcein AM (2 µM) and Ethidium homodimer-1 (2 µM) following manufacturer protocols. Standard fluorescent microscopy and image quantification were used to determine viability for each treatment. Overall cell viabilities were determined to be 96 ± 1% for experiment ①, 92 ± 4% for experiment ②, and 84 ± 2% for experiment ③ for four biological replicates (n = 4). No significant difference using a two-tailed t-test assuming equal variance in viability was observed between UV-exposed and unexposed portions of the gel in all conditions. Bar plots correspond to the mean viability in each condition with error bars representing ±1 standard deviation. Scale bar = 250 µm.

Supplementary Figure 13:



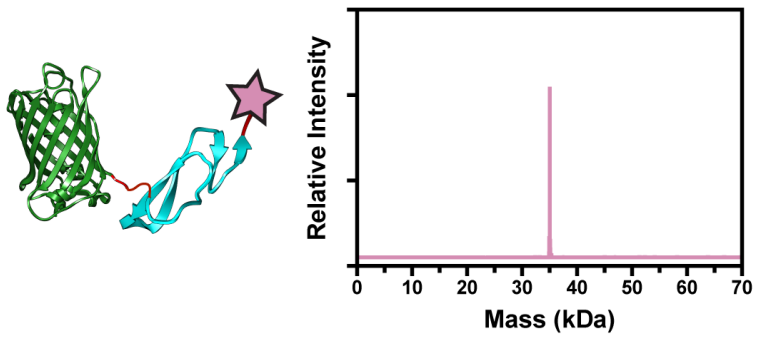
6.1.13 EGF stimulation of MAPK-reporter cells. HeLa cells with EKAREV linker treated with soluble EGF (100 ng in 600 µL total serum-free DMEM) and imaged using a Leica SP8X confocal microscope. EKAREV was excited using an argon laser (λ = 405 nm, 10% laser power); CFP excitation was monitored from $\lambda_{emission,CFP}$ = 450-484 nm, while YFP fluorescence was measured from $\lambda_{emission,YFP}$ = 520-524 nm. After image acquisition, background fluorescence was subtracted, and the YFP/CFP ratio for each pixel of the image was calculated using ImageJ. FRET response ratios were visualized as a blue-green-red color map. The left panel is prior to EGF stimulation; middle panel is 5 min post treatment; right panel is 10 min post treatment. Cells exhibit an increase in YFP/CFP ratio (represented as a color change from blue to green) upon EGF stimulation. Experiment was performed in triplicate (n = 3) with similar results. Scale bar = 100 µm.

Supplementary Figure 14:



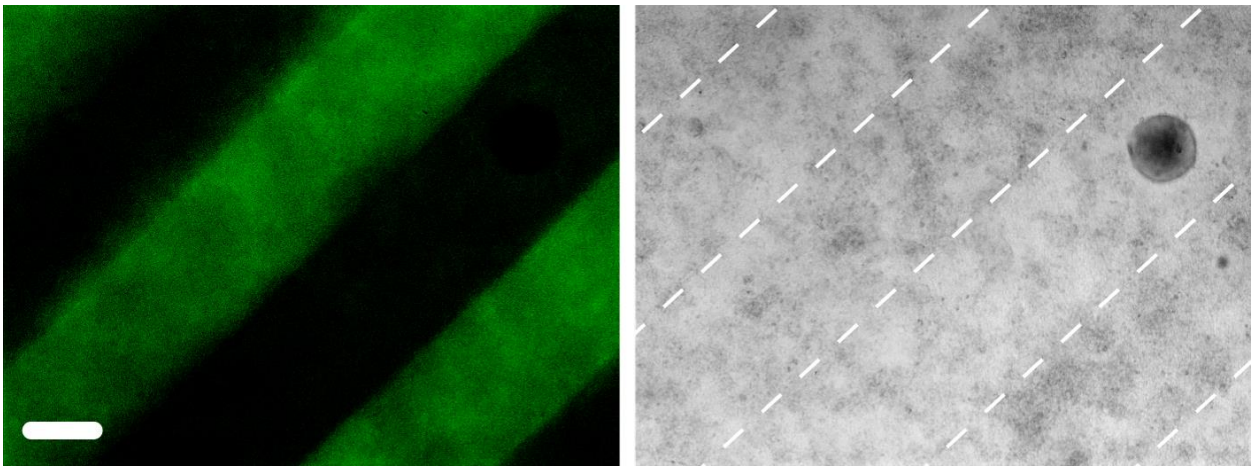
6.1.14 Effect of light treatment on EKAREV biosensor function. HeLa cells transfected with EKAREV FRET reporter for MAPK activation were encapsulated (2 x 10^6 cells mL^{-1}) in hydrogels lacking immobilized EGF. Gels were exposed to patterned light (λ = 365 nm, 10 mW cm^{-2}, 10 min), switched to serum-free media overnight, and imaged to determine FRET values. Cells exhibited the same basal level of FRET activation independent of light exposure, indicating that the EKAREV FRET reporter is suitable for use in photopatterned hydrogels. Image represents color-coded FRET response normalized to basal MAPK activation, obtained from confocal z-slices. Experiment was performed in triplicate (*n* = 3) with similar results. Scale bar = 100 µm.

Supplementary Figure 15:



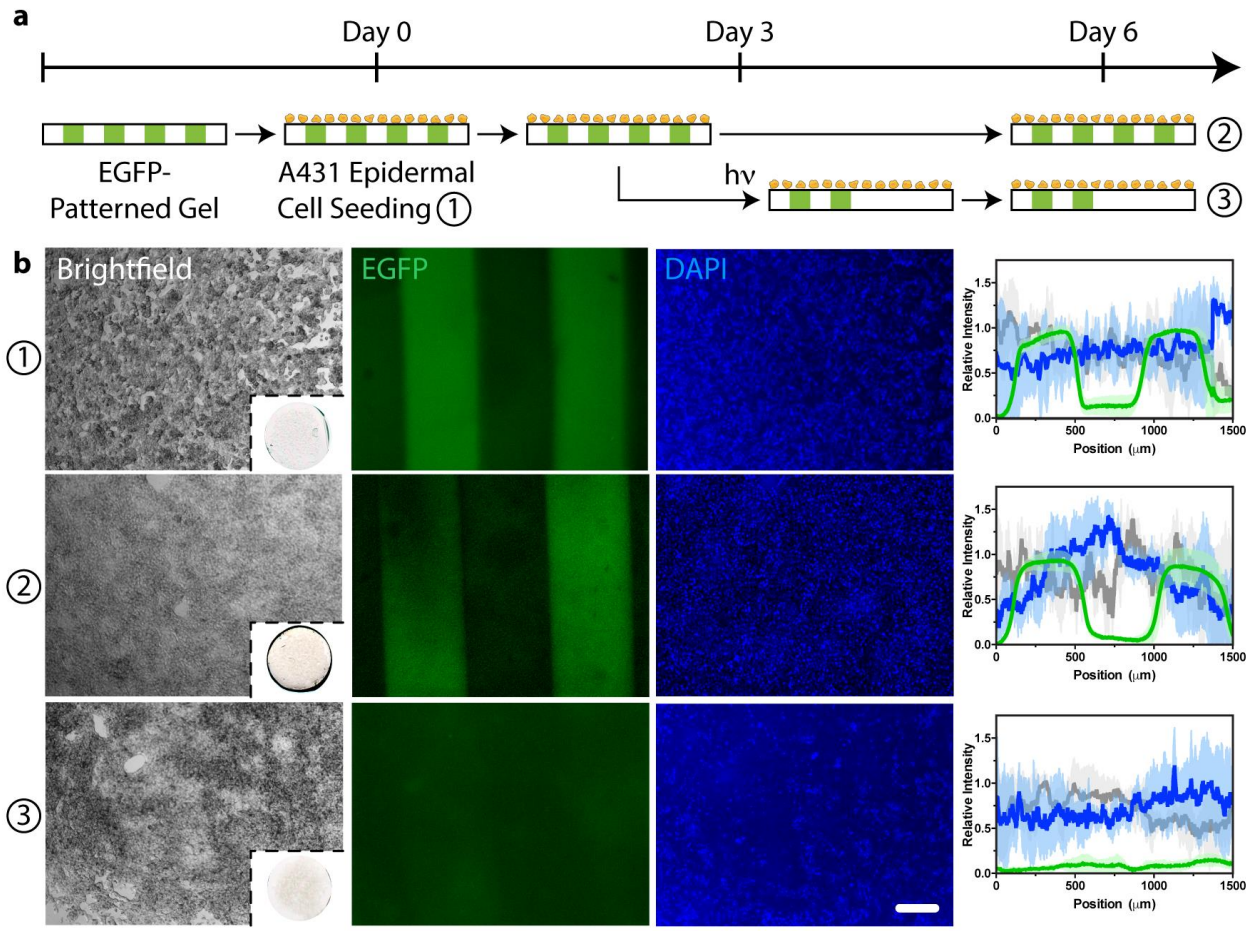
6.1.15 Mass spectrum of purified EGFP-EGF-oNB-N3. EGFP-EGF-*o*NB-N_3 was synthesized and purified following methodologies outlined in Supplementary Methods using the STEPL plasmid for EGFP-EGF sortagged with H-GGGGDDK(*o*NB-N_3)-NH_2 polyglycine probe (Supplementary Methods). Whole-protein mass spectrometry of a representative single purification indicated high sample purity and quantitative functionalization. Experiment was replicated (n = 3) with similar results.

Supplementary Figure 16:



6.1.16 Heterogenous cell density observed after three days of culture on patterned gels. For the experiment described in Figure 6a, heterogenous cell distributions is observed after Day 3 in culture on patterned gels. Fluorescent images of EGFP-EGF (green, left) correspond to regions with higher cell densities observed by brightfield microscopy (right). Experiment was replicated (n = 6) with similar results. Scale bar = 400 μm.

Supplementary Figure 17:

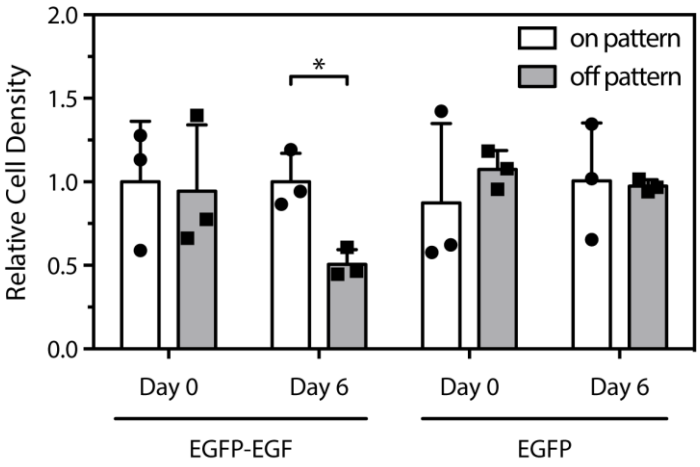


6.1.17 2D cell response to dynamically patterned EGFP on gels. EGFP-*o*NB-N_3 (25 μM) was conjugated (25 °C, 1 hr) to PEG-tetraBCN (M_n ~ 20,000 Da, 4 mM). EGFP-modified hydrogels (5 μL) were formed (1 hr) between Rain-X®-treated glass slides with silicone rubber spacers (McMaster-Carr, 0.5 mm thick) with N_3-GGRGDSPGGPQGIWGQGK(N_3)-NH_2 crosslinker (Supplementary Methods, 8 mM). After removal from the glass slide chamber, gels were exposed to masked collimated UV light (λ = 365 nm, 10 min, 10 mW cm^{-2}) through a slitted photomask containing 400 μm wide line features and soaked in PBS overnight. Gels were swollen (1 hr) in DMEM containing FBS (10%) prior to cell seeding. A431 cells (1 x 10^7 cells mL^{-1}) were added to the top of gels as a droplet (10 μL) and allowed to attach (1 hr, 37 °C) prior to media addition. Gels were swollen in DMEM containing FBS (10%) and penicillin/streptomycin (1%) overnight prior to imaging (condition ①). Gel media was supplemented with bovine serum albumin (BSA, 0.1%) and cells maintained in culture (37 °C, 5% CO_2). Three days after cell seeding, a portion of the hydrogels were exposed to a second round of photopatterning; one half of each treated gel was exposed to UV light (λ = 365 nm, 10 min, 10 mW cm^{-2}) while the other half was left unexposed.

In all cases, gels were maintained in DMEM containing FBS (10%), penicillin/streptomycin (1%), and BSA (0.1%) for an additional three days prior to analysis (conditions ② and ③).

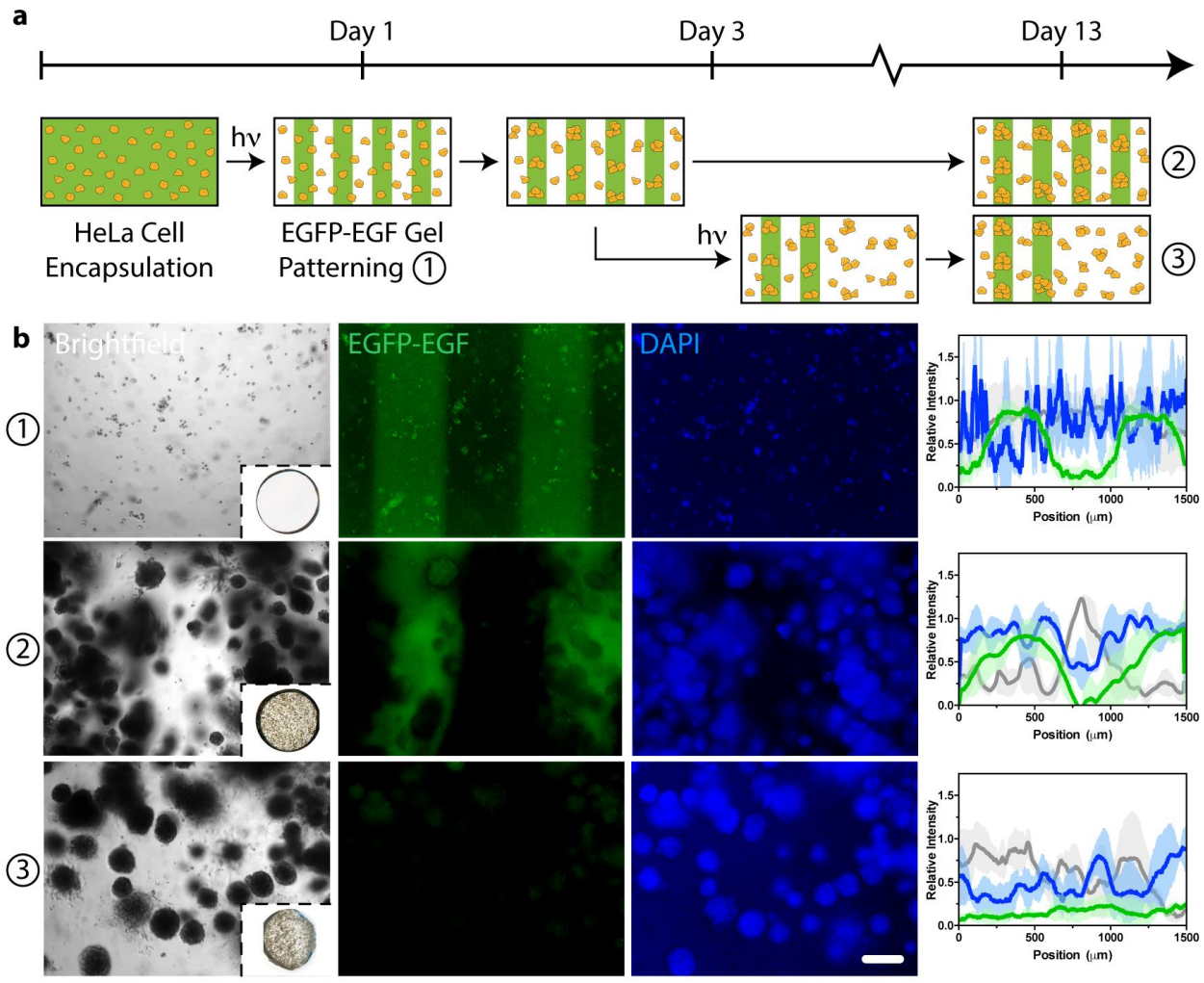
At the end of each treatment condition, cells were fixed in paraformaldehyde (1%, 30 min), stained with DAPI (0.15 µg/mL, 30 min), and visualized *via* fluorescence microscopy. Image analysis was performed for each condition using ImageJ, quantifying the normalized intensity profile for EGFP (green), nuclei (blue), and optical transmission (grey) across the gel perpendicular to photopatterned lines. Plots indicate average values (dark lines) and standard deviations (light error bars) from three biological replicates across the gel surface. Whole-gel images were performed on an Epson Perfection 4490 Photo document scanner. Unlike gels functionalized with EGFP-EGF, EGFP-only gels exhibit no variations in cell density matching the patterned region of tethered protein. Image insets correspond to gels roughly 0.5 cm in diameter. Scale bar = 200 µm.

Supplementary Figure 18:



6.1.18 Variable 2D cell response throughout patterned gels. Average 2D cell density was determined for A431 cells both on and off protein-patterned regions in conditions ① and ② of experiments comprising Figure 6a-b and Supplementary Figure 17. Cell density was calculated as the average total DAPI fluorescence in the center 200 μm portion of each region (n = 3 biological replicates). To account for basal levels in cell proliferation throughout the experiment, data has been normalized such that the highest relative cell density for experiments on a given day for a given patterned protein has a value of 1. A statistically significant difference (unpaired two-tailed t-test, $*p = 1.1 \times 10^{-2}$) was observed for cell density on versus off the EGFP-EGF pattern on day 6, but not for EGFP protein. All Day 0 data are statistically indistinguishable, indicating that the initial protein pattern did not hamper uniform cell seeding. Bar plots correspond to the mean cell density in each condition with error bars representing ±1 standard deviation.

Supplementary Figure 19:

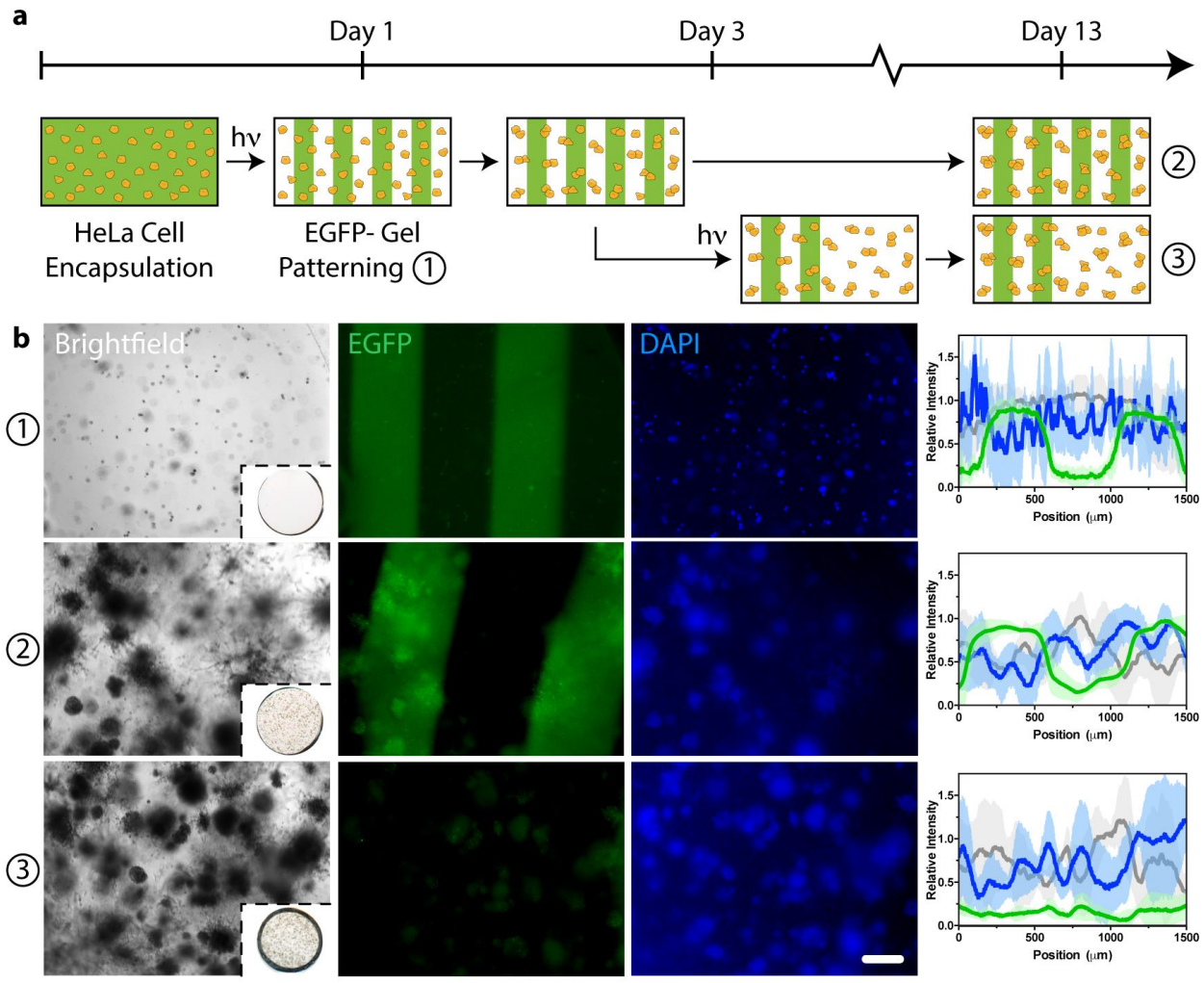


6.1.19 3D cell response to dynamically patterned EGFP-EGF within gels. EGFP-EGF-*o*NB-N_3 (10 μM) was conjugated (25 °C, 1 hr) to PEG-tetraBCN (M_n ~ 20,000 Da, 4 mM). EGFP-modified hydrogels (5 μL) were formed (1 hr) between Rain-X®-treated glass slides with silicone rubber spacers (McMaster-Carr, 0.5 mm thick) with N_3-GGRGDSPGGPQGIWGQGK(N_3)-NH_2 crosslinker (Supplementary Methods, 8 mM) and HeLa cell suspension (1 x 10^7 cells mL^{-1}). After removal from glass slide chamber, gels were maintained in DMEM containing FBS (10%) and penicillin/streptomycin (1%) overnight prior to initial photopatterning. Gels were exposed to masked collimated UV light (λ = 365 nm, 10 min, 10 mW cm^{-2}) through a slitted photomask containing 400 μm wide line features. Gels were swollen in DMEM containing FBS (10%) and penicillin/streptomycin (1%) and maintained in culture (37 °C, 5% CO_2) overnight prior to imaging (condition ①). Three days after cell encapsulation, a portion of the hydrogels were exposed to a second round of photopatterning; one half of each treated gel was exposed to UV

light (λ = 365 nm, 10 min, 10 mW cm^{-2}) while the other half was left unexposed. In all cases, gels were maintained in DMEM containing FBS (10%) and penicillin/streptomycin (1%) for an additional ten days prior to analysis (conditions ② and ③).

At the end of each treatment condition, cells were fixed in paraformaldehyde (1%, 30 min), stained with DAPI (0.15 µg/mL, 1 hr), and visualized *via* fluorescence microscopy. Image analysis was performed for each condition using ImageJ, quantifying the normalized intensity profile for EGFP-EGF (green), nuclei (blue), and optical transmission (grey) across the gel perpendicular to photopatterned lines. Plots indicate average values (dark lines) and standard deviations (light error bars) from three biological replicates across the gel volume. Cells exhibit staining in the green channel in condition ① due to EGFP-EGF binding with membrane receptors prior to being recycled. Whole-gel images were performed on an Epson Perfection 4490 Photo document scanner. Gel regions containing immobilized EGF are visually darker, attributed to enhanced spheroid growth. Image insets correspond to gels roughly 0.5 cm in diameter. Scale bar = 200 µm.

Supplementary Figure 20:

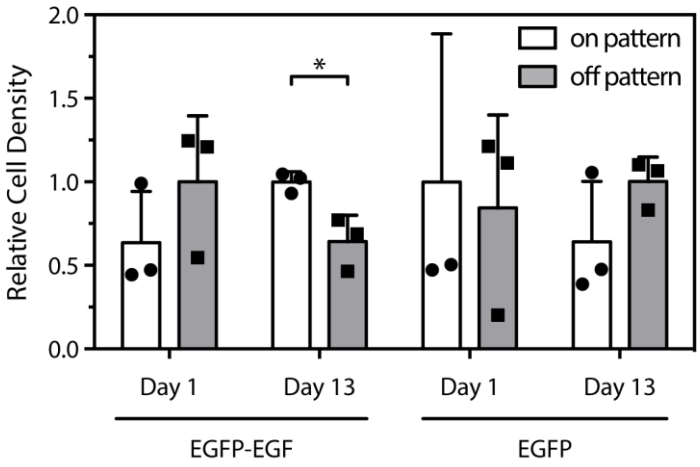


6.1.20 3D cell response to dynamically patterned EGFP within gels. EGFP-*o*NB-N_3 (10 μM) was conjugated (25 °C, 1 hr) to PEG-tetraBCN (M_n ~ 20,000 Da, 4 mM). EGFP-modified hydrogels (5 μL) were formed (1 hr) between Rain-X®-treated glass slides with silicone rubber spacers (McMaster-Carr, 0.5 mm thick) with N_3-GGRGDSPGGPQGIWGQGK(N_3)-NH_2 crosslinker (Supplementary Methods, 8 mM) and HeLa cell suspension (1 x 10^7 cells mL^{-1}). After removal from glass slide chamber, gels were maintained in DMEM containing FBS (10%) and penicillin/streptomycin (1%) overnight prior to initial photopatterning. Gels were exposed to masked collimated UV light (λ = 365 nm, 10 min, 10 mW cm^{-2}) through a slitted photomask containing 400 μm wide line features. Gels were swollen in DMEM containing FBS (10%) and penicillin/streptomycin (1%) and maintained in culture (37 °C, 5% CO_2) overnight prior to imaging (condition ①). Three days after cell encapsulation, a portion of the hydrogels were exposed to a second round of photopatterning; one half of each treated gel was exposed to UV

light (λ = 365 nm, 10 min, 10 mW cm^{-2}) while the other half was left unexposed. In all cases, gels were maintained in DMEM containing FBS (10%) and penicillin/streptomycin (1%) for an additional ten days prior to analysis (conditions ② and ③).

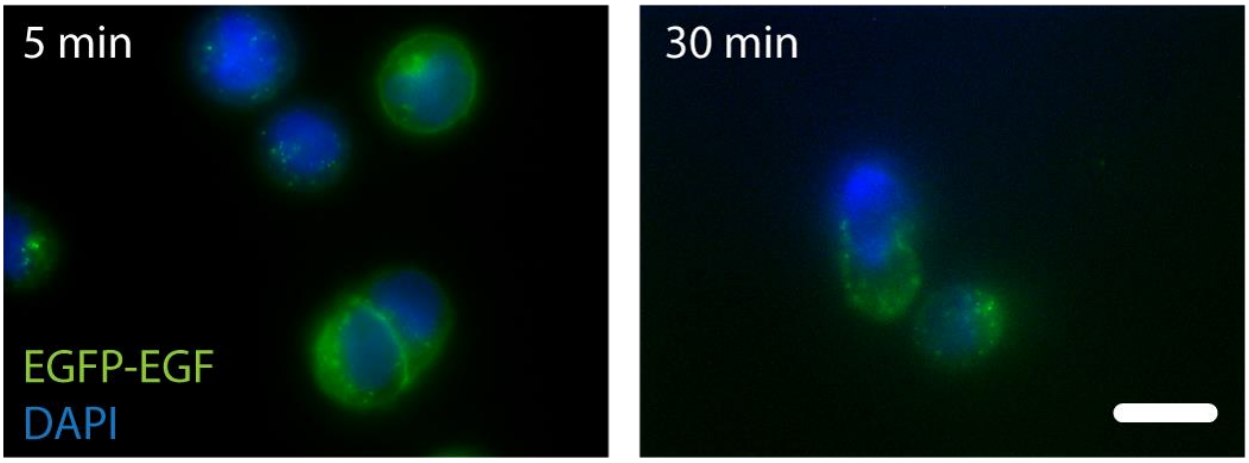
At the end of each treatment condition, cells were fixed in paraformaldehyde (1%, 30 min), stained with DAPI (0.15 µg/mL, 1 hr), and visualized *via* fluorescence microscopy. Image analysis was performed for each condition using ImageJ, quantifying the normalized intensity profile for EGFP (green), nuclei (blue), and optical transmission (grey) across the gel perpendicular to photopatterned lines. Plots indicate average values (dark lines) and standard deviations (light error bars) from three biological replicates across the gel volume. Whole-gel images were performed on an Epson Perfection 4490 Photo document scanner. EGFP-modified gels showed no increase in spheroid number or diameter between functionalized and unfunctionalized regions. Image insets correspond to gels roughly 0.5 cm in diameter. Scale bar = 200 µm.

Supplementary Figure 21:



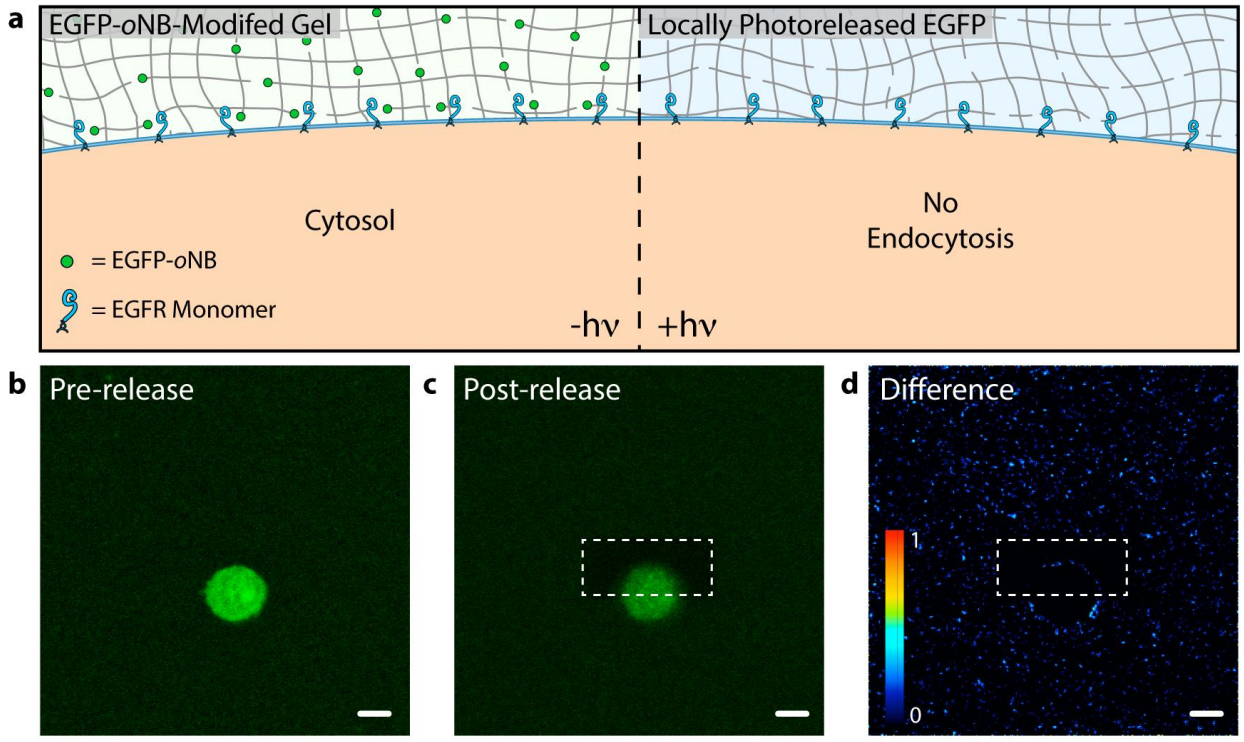
6.1.21 Variable 3D cell response throughout patterned gels. Average 3D cell density was determined for HeLa cells both on and off protein-patterned gel subvolumes in conditions ① and ② of experiments comprising Supplementary Figures 19 and 20. Cell density was calculated as the average total DAPI fluorescence in the center 200 μm portion of each region (n = 3 biological replicates). To account for basal levels in cell proliferation throughout the experiment, data has been normalized such that the highest relative cell density for experiments on a given day for a given patterned protein has a value of 1. A statistically significant difference (unpaired two-tailed t-test, $*p = 2.9 \times 10^{-2}$) was observed for cell density within versus outside the EGFP-EGF pattern on day 13, but not for EGFP protein. All Day 1 data are statistically indistinguishable, indicating that HeLa cells were encapsulated uniformly in gels. Bar plots correspond to the mean cell density in each condition with error bars representing ±1 standard deviation.

Supplementary Figure 22:



6.1.22 2D cell stimulation with soluble EGFP-EGF. To test receptor binding and internalization of the EGFP-EGF-*o*NB-N_3 fusion protein, A431 cells were plated on tissue-culture polystyrene (10^5 cells cm^{-2}) and allowed to attach overnight in DMEM supplemented with FBS (10%) and penicillin/streptomycin (1%). The media was replaced with DMEM containing EGFP-EGF-*o*NB-N_3 (125 nM) and incubated for 30 minutes at room temperature before swapping into fresh media with no added growth factor and returned to 37 °C. At either 5 or 30 minutes later, cells were fixed in paraformaldehyde (1%, 30 min), stained with DAPI (0.15 µg/mL, 30 min), and visualized *via* fluorescence microscopy. Cells display strong membrane staining at early time points, followed by more diffuse staining as the EGFR receptor is internalized. Punctate staining in the green channel corresponds to concentrated EGFP-EGF within endosomal vesicles. Experiment was replicated (n = 8) with similar results. Scale bar = 10 µm.

Supplementary Figure 23:



6.1.23 Encapsulated cell response to subcellular photoreleased EGFP. A431 cells ($1x10^7$ cells mL^{-1}) were encapsulated in SPAAC hydrogels (5 µL) containing EGFP-*o*NB-N_3 (125 nM) and N_3-GRGDS-NH_2 peptide (Supplementary Methods, 1 mM) and affixed to an azide-functionalized slide (Supplementary Methods). The cells were cultured for 48 hours in DMEM supplemented with FBS (10%) and penicillin/streptomycin (1%) before transferring to phenol red-free DMEM containing FBS (10%) and penicillin/streptomycin (1%). Protein photorelease was performed *via* multiphoton laser scanning lithography (Olympus FV1000 MPE BX61 Multi-photon Microscope, 20x objective, 5x zoom, λ = 740 nm, 25% laser power, 16 scan repeats, z-interval = 2.5 µm), confined to a 40 µm x 30 µm x 5 µm subvolume bisecting an individual cell. Images were taken (b) prior to release and (c) 30 minutes afterwards. (d) ImageJ was used to calculate the difference in pixel intensity accompanying photorelease. Gels containing EGFP-*o*NB-N_3 exhibit nonspecific staining of encapsulated cells. Photoreleased EGFP does not trigger endocytosis, as evidenced by the lack of punctate fluorescence within the cell following treatment. Experiment was performed within a single gel. Scale bar = 10 µm.

6.2 Supplementary Table 1. Identification codes for all sortagged species

Base Protein	Polyglycine Probe	Sortagged Protein Identity
EGFP	Triglycine	EGFP
	H-GGGGDDK(N_3)-NH_2	EGFP-N_3
	H-GGGGDDK(CHO)-NH_2	EGFP-CHO
	H-GGGGDDK(*o*NB-N_3)-NH_2	EGFP-*o*NB-N_3
	H-GGGG-*o*NB-DDK(CHO)-NH_2	EGFP-*o*NB-CHO
mCherry	Triglycine	mCherry
	H-GGGGDDK(N_3)-NH_2	mCherry-N_3
	H-GGGGDDK(CHO)-NH_2	mCherry-CHO
	H-GGGGDDK(*o*NB-N_3)-NH_2	mCherry-*o*NB-N_3
	H-GGGG-*o*NB-DDK(CHO)-NH_2	mCherry-*o*NB-CHO
mCerulean	Triglycine	mCerulean
	H-GGGGDDK(N_3)-NH_2	mCerulean-N_3
	H-GGGGDDK(CHO)-NH_2	mCerulean-CHO
	H-GGGGDDK(*o*NB-N_3)-NH_2	mCerulean-*o*NB-N_3
	H-GGGG-*o*NB-DDK(CHO)-NH_2	mCerulean-*o*NB-CHO
EGF	Triglycine	EGF
	H-GGGGDDK(N_3)-NH_2	EGF-N_3
	H-GGGGDDK(CHO)-NH_2	EGF-CHO
	H-GGGGDDK(*o*NB-N_3)-NH_2	EGF-*o*NB-N_3
	H-GGGG-*o*NB-DDK(CHO)-NH_2	EGF-*o*NB-CHO
bla	Triglycine	bla
	H-GGGGDDK(N_3)-NH_2	bla-N_3
	H-GGGGDDK(CHO)-NH_2	bla-CHO
	H-GGGGDDK(*o*NB-N_3)-NH_2	bla-*o*NB-N_3
	H-GGGG-*o*NB-DDK(CHO)-NH_2	bla-*o*NB-CHO

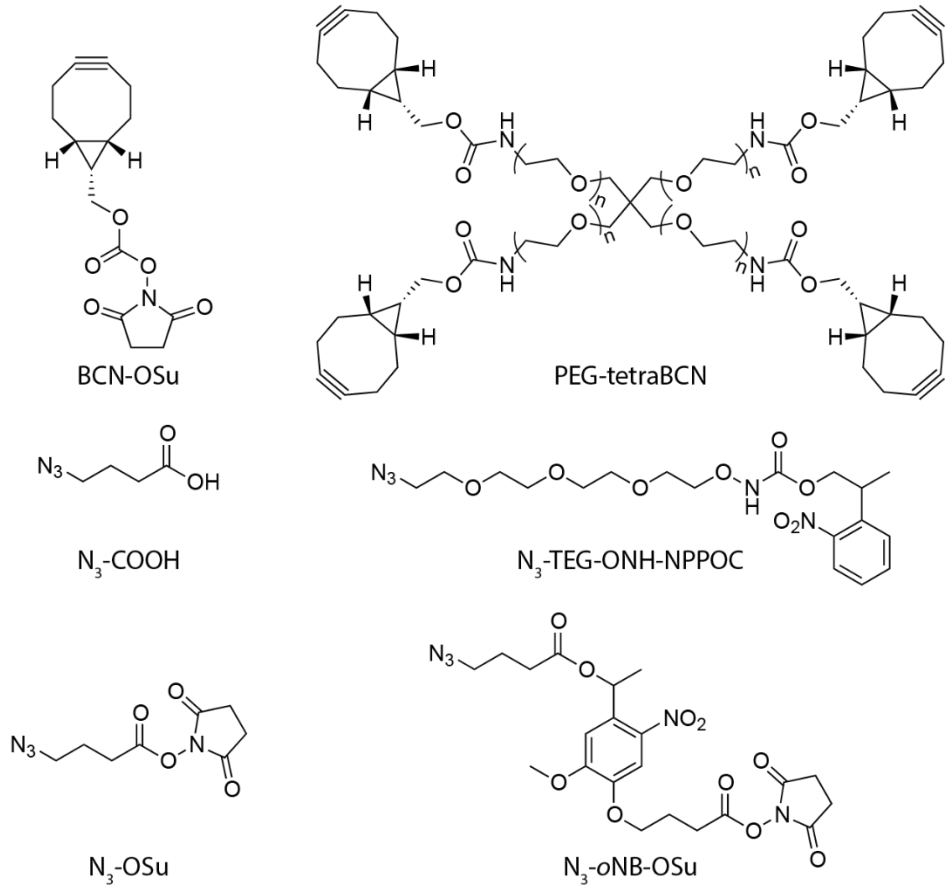
	Triglycine	MBP-FGF
	H-GGGGDDK(N_3)-NH_2	MBP-FGF-N_3
MBP-FGF	H-GGGGDDK(CHO)-NH_2	MBP-FGF-CHO
	H-GGGGDDK(*o*NB-N_3)-NH_2	MBP-FGF-N_3
	H-GGGG-*o*NB-DDK(CHO)-NH_2	MBP-FGF-CHO

6.3 Supplementary Methods:

6.3.1 General synthetic information

Chemical reagents and solvents were purchased from either Sigma-Aldrich or Fisher Scientific and used as received unless otherwise noted. Peptide synthesis reagents were purchased from either ChemPep or Chem-Impex and used as received. Deionized water (dH_2O) was generated by a U.S. Filter Corporation Reverse Osmosis System with a Desal membrane. Synthetic chemical reactions were performed under a nitrogen atmosphere in oven-dried glassware and stirred with a Teflon-coated magnetic stir bar unless otherwise noted. Solvents were removed *in vacuo* with a Büchi Rotovapor R-3 equipped with a V-700 vacuum pump and V-855 vacuum controller and a Welch 1400 DuoSeal Belt-Drive high vacuum pump. ^1H nuclear magnetic resonance (NMR) data was collected at 298 K on Bruker instruments and chemical shifts are reported relative to tetramethylsilane (TMS, $\delta = 0$). Microwave-assisted peptide synthesis was performed on a CEM Liberty 1. Semi-preparative reversed-phase high-pressure liquid chromatography (RP-HPLC) was performed on a Dionex Ultimate 3000 equipped with a variable multiple wavelength detector, automated fraction collector, and Thermo 5 µm Synchronis silica 250 x 21.2 mm C18 column. Lyophilization was performed on a LABCONCO FreeZone 2.5 Plus freeze-dryer equipped with a LABCONCO rotary vane 117 vacuum pump. Matrix-assisted laser desorption/ionization time of flight (MALDI-TOF) mass spectrometry was performed in reflectron positive ion mode or reflectron negative ion mode on a Bruker AutoFlex II using a matrix of α-cyano-4-hydroxycinnamic acid:2,5-dihydroxy benzoic acid (2:1). Whole-protein mass spectrometry was performed using a Waters Synapt – G2 QTOF. The light source for the photochemical cleavage was a Lumen Dynamics OmniCure S1500 Spot UV Curing system with an internal 365 nm filter and an external 360 nm cut-on longpass filter. Light intensity was measured using a Cole-Parmer Radiometer (Series 9811-50, $\lambda = 365$ nm). Biochemical gradients were created with the aid of a Harvard Apparatus PHD 2000 Syringe Pump. Fluorescence readings were acquired on a SpectraMax M5 spectrometer using Thermo Scientific Nunc black polypropylene 96-well plates. Confocal microscopy was performed at the University of Washington Keck Microscopy Center on a Leica SP8X confocal microscope. Multiphoton lithography was performed on an Olympus FV1000 MPE BX61 Multi-photon Microscope at the Garvey Imaging Center at the University of Washington. Polymerase chain reaction (PCR) was performed in a Bioer LifeECO thermal cycler. Protein expression was performed in a Thermo Scientific MaxQ 4000 shaker incubator. Cells were lysed using a Fisher Scientific Model 505 Sonic Dismembrator with a 1.27 cm diameter probe. Mammalian cell culture was performed in a NuAire LabGard ES NU-437 Class II Type A2 Biosafety Cabinet. Cells were maintained in a Sanyo inCu saFe® MCO-17AC incubator at 37 °C and 5% CO_2.

6.3.2 Synthesis of previously reported compounds used in this work



(1R,8S,9s)-bicyclo[6.1.0]non-4-yn-9-ylmethyl (2,5-dioxopyrrolidin-1-yl) carbonate (BCN-OSu), poly(ethylene glycol) tetrabicyclononyne (PEG-tetraBCN, M_n ~ 20,000 Da), 4-azidobutanoic acid (N_3-COOH), 2,5-dioxopyrrolidin-1-yl 4-azidobutanoate (N_3-OSu), 2-(2-nitrophenyl)propyl 2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethoxycarbamate (N_3-TEG-ONH-NPPOC), and 2,5-dioxopyrrolidin-1-yl 4-(4-(1-((4-azidobutanoyl)oxy)ethyl)-2-methoxy-5-nitrophenoxy)butanoate (N_3-*o*NB-OSu) were synthesized as previously reported[1].

6.3.3 Plasmid construction for protein expression and STEPL purification

Sources of DNA sequences

The pSTEPL plasmid[2] containing EGFP and an ampicillin resistance gene was generously donated by the Tsourkas group (University of Pennsylvania). Plasmids containing gene sequences for mCherry, Beta-lactamase, and Maltose Binding Protein sequences were donated by the Baneyx group (University of Washington). A plasmid containing the mCerulean sequence was donated by Daniel Strongin (Fred Hutchinson Cancer Research Center). EGF-bio-His was a gift from Gavin Wright (Addgene plasmid # 53340). A codon-optimized gene sequence for FGF was purchased as a gBlock (Integrated DNA Technologies).

STEPL Plasmid Construction

For each protein to be expressed in the STEPL system, polymerase chain reaction (PCR) was used to amplify gene sequences of interest and introduce relevant restriction sites (5' NdeI and 3' XhoI) for subsequent cloning.

pSTEPL plasmid and PCR products were digested (4 hr, 37 °C) with NdeI and XhoI restriction enzymes (New England BioLabs) and purified by extraction following electrophoretic separation (0.8% agarose). The digested pSTEPL and protein insert were ligated (T4 DNA ligase, 16 hr, 16 °C), and transformed into chemically competent Top10 *E. coli* (Thermo Fisher) by heat shock, and plated onto agar plates (10 g Tryptone, 5 g Yeast Extract, 10 g NaCl, 15 g agar, 1 L dH_2O) containing ampicillin (100 μg mL^{-1}). Colonies were subsequently grown overnight in Miller's Lysogeny Broth (LB, 5 mL) containing ampicillin (100 μg mL^{-1}). Plasmids were purified using a QIAprep Spin Miniprep Kit (Qiagen), and sequenced using a SimpleSeq DNA Sequencing Kit (Thermo Fisher). Plasmids corresponding to the STEPL construct of interest were purified and subsequently transformed into chemically competent BL21(DE3) *E. coli* (Promega) for expression.

Due to low solubility of the MBP-FGF-SrtA STEPL fusion protein, MBP-FGF containing a C-terminal sorting signal was also expressed separately from the sortase enzyme. The sequence for MBP was lifted out using PCR and adding a 5' NdeI site and a 3' LPETG motif followed by a HindIII site. The gBlock for FGF was ordered with a 5' HindIII site and a 3' XhoI site. Each insert was digested with their respective enzymes (New England BioLabs), and pET-21a(+) was digested with NdeI and XhoI. After gel extraction, the fragments were ligated using T4 DNA polymerase (16 °C, 16 h). The ligated product was transformed into chemically competent Top10 cells by heat shock. Purified plasmids from successful transformants were subsequently transformed into chemically competent BL21(DE3) *E. coli* for expression.

Below is a list of DNA open reading frame sequences (5' → 3') generated for and utilized in these studies. Nucleotides shown in black correspond to that for the specific protein of interest. Bases shown in orange correspond to STEPL portion common to all constructs, which features the C-terminal LPETG sortase recognition motif, a flexible $(GGS)_5$ linker, SrtA, and a 6xHis tag.

EGFP-STEPL:

ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCG
ACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCT
GACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACC
CTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCA
AGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTA
CAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGC
ATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACA
ACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAA
CATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGC
CCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACG
AGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGA
CGAGCTGTACAAGCTCGAGCTGCCGGAAACCGGTGGTGGTAGTGGTGGCTCTGGCGGTTCTGGT
GGCAGTGGCGGTAGCCAAGCTAAACCTCAAATTCCGAAAGATAAATCAAAAGTGGCAGGCTATA
TTGAAATTCCAGATGCTGATATTAAAGAACCAGTATATCCAGGACCAGCAACACCTGAACAATT
AAATAGAGGTGTAAGCTTTGCAGAAGAAAATGAATCACTAGATGATCAAAATATTTCAATTGCA
GGACACACTTTCATTGACCGTCCGAACTATCAATTTACAAATCTTAAAGCAGCCAAAAAAGGTA
GTATGGTGTACTTTAAAGTTGGTAATGAAACACGTAAGTATAAAATGACAAGTATAAGAGATGT
TAAGCCAACAGATGTAGAAGTTCTAGATGAACAAAAAGGTAAAGATAAACAATTAACATTAATT
ACTTGTGATGATTACAATGAAAAGACAGGCGTTTGGGAAAAACGTAAAATCTTTGTAGCTACAG
AAGTCAAACATCACCACCATCATCACTAA

mCherry-STEPL:

ATGGTTTCCAAGGGCGAAGAAGACAACATGGCGATCATCAAAGAATTTATGCGTTTTAAAGTTC
ACATGGAAGGTTCTGTTAACGGTCATGAGTTCGAAATTGAAGGTGAGGGTGAAGGTCGCCCGTA
CGAAGGTACCCAGACCGCGAAACTGAAAGTTACCAAAGGTGGTCCGCTGCCGTTCGCGTGGGAC
ATCCTCAGCCCGCAGTTCATGTACGGTTCTAAAGCGTACGTTAAACATCCGGCGGACATTCCAG
ACTACCTCAAACTCTCTTTCCCTGAAGGTTTCAAATGGGAACGTGTTATGAACTTCGAGGACGG
TGGTGTTGTCACCGTTACCCAGGACTCTTCTCTGCAGGACGGCGAGTTCATCTACAAGGTCAAA
CTGCGTGGCACCAACTTCCCGTCTGACGGTCCGGTTATGCAGAAAAAAACCATGGGTTGGGAAG
CGTCTTCTGAACGTATGTACCCGGAAGATGGTGCGCTGAAAGGCGAAATCAAACAGCGTCTGAA
GCTCAAAGACGGCGGTCACTACGACGCGGAGGTTAAAACCACCTACAAAGCGAAAAAGCCGGTT
CAACTGCCGGGTGCGTACAACGTTAATATCAAGCTGGACATCACCTCTCACAACGAAGACTACA
CCATCGTTGAACAGTACGAACGTGCGGAAGGCCGTCACTCTACCGGTGGTATGGACGAACTGTA
CAAGCTCGAGCTGCCGGAAACCGGTGGTGGTAGTGGTGGCTCTGGCGGTTCTGGTGGCAGTGGC
GGTAGCCAAGCTAAACCTCAAATTCCGAAAGATAAATCAAAAGTGGCAGGCTATATTGAAATTC
CAGATGCTGATATTAAAGAACCAGTATATCCAGGACCAGCAACACCTGAACAATTAAATAGAGG
TGTAAGCTTTGCAGAAGAAAATGAATCACTAGATGATCAAAATATTTCAATTGCAGGACACACT
TTCATTGACCGTCCGAACTATCAATTTACAAATCTTAAAGCAGCCAAAAAAGGTAGTATGGTGT
ACTTTAAAGTTGGTAATGAAACACGTAAGTATAAAATGACAAGTATAAGAGATGTTAAGCCAAC
AGATGTAGAAGTTCTAGATGAACAAAAAGGTAAAGATAAACAATTAACATTAATTACTTGTGAT

GATTACAATGAAAAGACAGGCGTTTGGGAAAAACGTAAAATCTTTGTAGCTACAGAAGTCAAAC
ATCACCACCATCATCACTAA

mCerulean-STEPL:

ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCG
ACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCT
GACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACC
CTGACCTGGGGCGTGCAGTGCTTCGCCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCA
AGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTA
CAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGC
ATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACGCCATCAGCGACA
ACGTCTATATCACCGCCGACAAGCAGAAGAACGGCATCAAGGCCAACTTCAAGATCCGCCACAA
CATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGC
CCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCAAGCTGAGCAAAGACCCCAACG
AGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGA
CGAGCTGTACAAGCTCGAGCTGCCGGAAACCGGTGGTGGTAGTGGTGGCTCTGGCGGTTCTGGT
GGCAGTGGCGGTAGCCAAGCTAAACCTCAAATTCCGAAAGATAAATCAAAAGTGGCAGGCTATA
TTGAAATTCCAGATGCTGATATTAAAGAACCAGTATATCCAGGACCAGCAACACCTGAACAATT
AAATAGAGGTGTAAGCTTTGCAGAAGAAAATGAATCACTAGATGATCAAAATATTTCAATTGCA
GGACACACTTTCATTGACCGTCCGAACTATCAATTTACAAATCTTAAAGCAGCCAAAAAAGGTA
GTATGGTGTACTTTAAAGTTGGTAATGAAACACGTAAGTATAAAATGACAAGTATAAGAGATGT
TAAGCCAACAGATGTAGAAGTTCTAGATGAACAAAAAGGTAAAGATAAACAATTAACATTAATT
ACTTGTGATGATTACAATGAAAAGACAGGCGTTTGGGAAAAACGTAAAATCTTTGTAGCTACAG
AAGTCAAACATCACCACCATCATCACTAA

EGF-STEPL:

ATGAACAGCGACAGCGAGTGCCCACTGAGCCACGACGGCTACTGCCTGCACGACGGCGTGTGCA
TGTACATCGAGGCCCTGGACAAGTACGCCTGCAACTGCGTCGTGGGCTACATCGGCGAGCGGTG
CCAGTACCGGGACCTGAAGTGGTGGGAGCTGAGACCTCGAGCTGCCGGAAACCGGTGGTGGTAG
TGGTGGCTCTGGCGGTTCTGGTGGCAGTGGCGGTAGCCAAGCTAAACCTCAAATTCCGAAAGAT
AAATCAAAAGTGGCAGGCTATATTGAAATTCCAGATGCTGATATTAAAGAACCAGTATATCCAG
GACCAGCAACACCTGAACAATTAAATAGAGGTGTAAGCTTTGCAGAAGAAAATGAATCACTAGA
TGATCAAAATATTTCAATTGCAGGACACACTTTCATTGACCGTCCGAACTATCAATTTACAAAT
CTTAAAGCAGCCAAAAAAGGTAGTATGGTGTACTTTAAAGTTGGTAATGAAACACGTAAGTATA
AAATGACAAGTATAAGAGATGTTAAGCCAACAGATGTAGAAGTTCTAGATGAACAAAAAGGTAA
AGATAAACAATTAACATTAATTACTTGTGATGATTACAATGAAAAGACAGGCGTTTGGGAAAAA
CGTAAAATCTTTGTAGCTACAGAAGTCAAACATCACCACCATCATCACTAA

bla-STEPL:

ATGCACCCTGAAACGCTGGTGAAAGTAAAANATGCTGAAGATCAGTTGGGTGCACGAGTGGGTT
ACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCC
AATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTGTTGACGCCGGGCAA
GAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAG
AAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGA
TAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTG
CACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATAC
CAAACGACGAGCGTGACACCACGATGCCTGCAGCAATGGCAACAACGTTGCGCAAACTATTAAC
TGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTT
GCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCG
GTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGT
AGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATA
GGTGCCTCACTGATTAAGCATTGGCTCGAGCTGCCGGAAACCGGTGGTGGTAGTGGTGGCTCTG
GCGGTTCTGGTGGCAGTGGCGGTAGCCAAGCTAAACCTCAAATTCCGAAAGATAAATCAAAAGT
GGCAGGCTATATTGAAATTCCAGATGCTGATATTAAAGAACCAGTATATCCAGGACCAGCAACA
CCTGAACAATTAAATAGAGGTGTAAGCTTTGCAGAAGAAAATGAATCACTAGATGATCAAAATA
TTTCAATTGCAGGACACACTTTCATTGACCGTCCGAACTATCAATTTACAAATCTTAAAGCAGC
CAAAAAAGGTAGTATGGTGTACTTTAAAGTTGGTAATGAAACACGTAAGTATAAAATGACAAGT
ATAAGAGATGTTAAGCCAACAGATGTAGAAGTTCTAGATGAACAAAAAGGTAAAGATAAACAAT
TAACATTAATTACTTGTGATGATTACAATGAAAAGACAGGCGTTTGGGAAAAACGTAAAATCTT
TGTAGCTACAGAAGTCAAACATCACCACCATCATCACTAA

MBP-FGF-STEPL:

ATGAAAATCGAAGAAGGTAAACTGGTAATCTGGATTAACGGCGATAAAGGCTATAACGGTCTCG
CTGAAGTCGGTAAGAAATTCGAGAAAGATACCGGAATTAAAGTCACCGTTGAGCATCCGGATAA
ACTGGAAGAGAAATTCCCACAGGTTGCGGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCA
CACGACCGCTTTGGTGGCTACGCTCAATCTGGCCTGTTGGCTGAAATCACCCCGGACAAAGCGT
TCCAGGACAAGCTGTATCCGTTTACCTGGGATGCCGTACGTTACAACGGCAAGCTGATTGCTTA
CCCGATCGCTGTTGAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAA
ACCTGGGAAGAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGT
TCAACCTGCAAGAACCGTACTTCACCTGGCCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAA
GTATGAAAACGGCAAGTACGACATTAAAGACGTGGGCGTGGATAACGCTGGCGCGAAAGCGGGT
CTGACCTTCCTGGTTGACCTGATTAAAAACAAACACATGAATGCAGACACCGATTACTCCATCG
CAGAAGCTGCCTTTAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGGGCATGGTCCAA
CATCGACACCAGCAAAGTGAATTATGGTGTAACGGTACTGCCGACCTTCAAGGGTCAACCATCC
AAACCGTTCGTTGGCGTGCTGAGCGCAGGTATTAACGCCGCCAGTCCGAACAAAGAGCTGGCAA
AAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGTCTGGAAGCGGTTAATAAAGACAAACC
GCTGGGTGCCGTAGCGCTGAAGTCTTACGAGGAAGAGTTGGCGAAAGATCCACGTATTGCCGCC
ACCATGGAAAACGCCCAGAAAGGTGAAATCATGCCGAACATCCCGCAGATGTCCGCTTTCTGGT
ATGCCGTGCGTACTGCGGTGATCAACGCCGCCAGCGGTCGTCAGACTGTCGATGAAGCCCTGAA
AGACGCGCAGACTAATTCGAGCAAGCTTGCTGCTGGTTCTATCACCACCCTGCCGGCTCTGCCG

GAAGACGGTGGTTCTGGTGCTTTCCCGCCGGGTCACTTCAAAGACCCGAAACGTCTGTACTGCA
AAAACGGTGGTTTCTTCCTGCGTATCCACCCGGACGGTCGTGTTGACGGTGTTCGTGAAAAATC
TGACCCGCACATCAAACTGCAGCTGCAGGCTGAAGAACGTGGTGTTGTTTCTATCAAAGGTGTT
TGCGCTAACCGTTACCTGGCTATGAAAGAAGACGGTCGTCTGCTGGCTTCTAAATGCGTTACCG
ACGAATGCTTCTTCTTCGAACGTCTGGAATCTAACAACTACAACACCTACCGTTCTCGTAAATA
CACCTCTTGGTACGTTGCTCTGAAACGTACCGGTCAGTACAAACTGGGTTCTAAAACCGGTCCG
GGTCAGAAAGCTATCCTGTTCCTGCCGATGTCTGCTAAATCTCTCGAGCTGCCGGAAACCGGTG
GTGGTAGTGGTGGCTCTGGCGGTTCTGGTGGCAGTGGCGGTAGCCAAGCTAAACCTCAAATTCC
GAAAGATAAATCAAAAGTGGCAGGCTATATTGAAATTCCAGATGCTGATATTAAAGAACCAGTA
TATCCAGGACCAGCAACACCTGAACAATTAAATAGAGGTGTAAGCTTTGCAGAAGAAAATGAAT
CACTAGATGATCAAAATATTTCAATTGCAGGACACACTTTCATTGACCGTCCGAACTATCAATT
TACAAATCTTAAAGCAGCCAAAAAAGGTAGTATGGTGTACTTTAAAGTTGGTAATGAAACACGT
AAGTATAAAATGACAAGTATAAGAGATGTTAAGCCAACAGATGTAGAAGTTCTAGATGAACAAA
AAGGTAAAGATAAACAATTAACATTAATTACTTGTGATGATTACAATGAAAAGACAGGCGTTTG
GGAAAAACGTAAAATCTTTGTAGCTACAGAAGTCAAACATCACCACCATCATCACTAA

MBP-*FGF*-*LPETG-His6*:

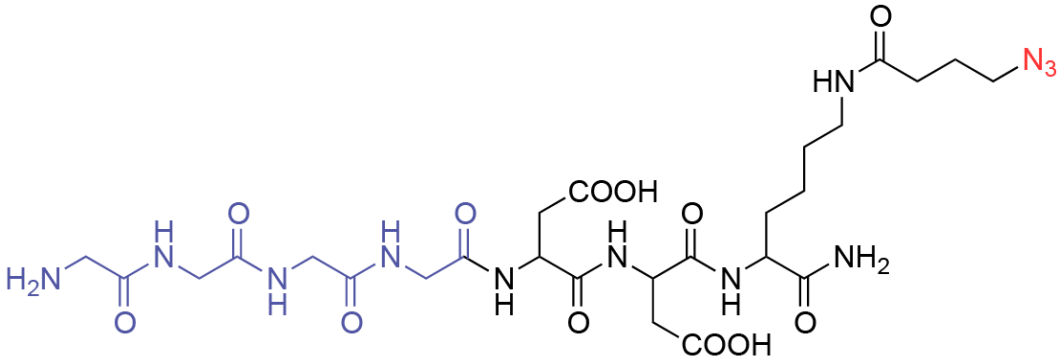
ATGAAAATCGAAGAAGGTAAACTGGTAATCTGGATTAACGGCGATAAAGGCTATAACGGTCTCG
CTGAAGTCGGTAAGAAATTCGAGAAAGATACCGGAATTAAAGTCACCGTTGAGCATCCGGATAA
ACTGGAAGAGAAATTCCCACAGGTTGCGGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCA
CACGACCGCTTTGGTGGCTACGCTCAATCTGGCCTGTTGGCTGAAATCACCCCGGACAAAGCGT
TCCAGGACAAGCTGTATCCGTTTACCTGGGATGCCGTACGTTACAACGGCAAGCTGATTGCTTA
CCCGATCGCTGTTGAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAA
ACCTGGGAAGAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGT
TCAACCTGCAAGAACCGTACTTCACCTGGCCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAA
GTATGAAAACGGCAAGTACGACATTAAAGACGTGGGCGTGGATAACGCTGGCGCGAAAGCGGGT
CTGACCTTCCTGGTTGACCTGATTAAAAACAAACACATGAATGCAGACACCGATTACTCCATCG
CAGAAGCTGCCTTTAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGGGCATGGTCCAA
CATCGACACCAGCAAAGTGAATTATGGTGTAACGGTACTGCCGACCTTCAAGGGTCAACCATCC
AAACCGTTCGTTGGCGTGCTGAGCGCAGGTATTAACGCCGCCAGTCCGAACAAAGAGCTGGCAA
AAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGTCTGGAAGCGGTTAATAAAGACAAACC
GCTGGGTGCCGTAGCGCTGAAGTCTTACGAGGAAGAGTTGGCGAAAGATCCACGTATTGCCGCC
ACCATGGAAAACGCCCAGAAAGGTGAAATCATGCCGAACATCCCGCAGATGTCCGCTTTCTGGT
ATGCCGTGCGTACTGCGGTGATCAACGCCGCCAGCGGTCGTCAGACTGTCGATGAAGCCCTGAA
AGACGCGCAGACTAATTCGAGCAAGCTTGCTGCTGGTTCTATCACCACCCTGCCGGCTCTGCCG
GAAGACGGTGGTTCTGGTGCTTTCCCGCCGGGTCACTTCAAAGACCCGAAACGTCTGTACTGCA
AAAACGGTGGTTTCTTCCTGCGTATCCACCCGGACGGTCGTGTTGACGGTGTTCGTGAAAAATC
TGACCCGCACATCAAACTGCAGCTGCAGGCTGAAGAACGTGGTGTTGTTTCTATCAAAGGTGTT
TGCGCTAACCGTTACCTGGCTATGAAAGAAGACGGTCGTCTGCTGGCTTCTAAATGCGTTACCG
ACGAATGCTTCTTCTTCGAACGTCTGGAATCTAACAACTACAACACCTACCGTTCTCGTAAATA
CACCTCTTGGTACGTTGCTCTGAAACGTACCGGTCAGTACAAACTGGGTTCTAAAACCGGTCCG

GGTCAGAAAGCTATCCTGTTCCTGCCGATGTCTGCTAAATCTCTGCCGGAAACCGGTGGTCTCG
AGCACCACCACCACCACCACTGA

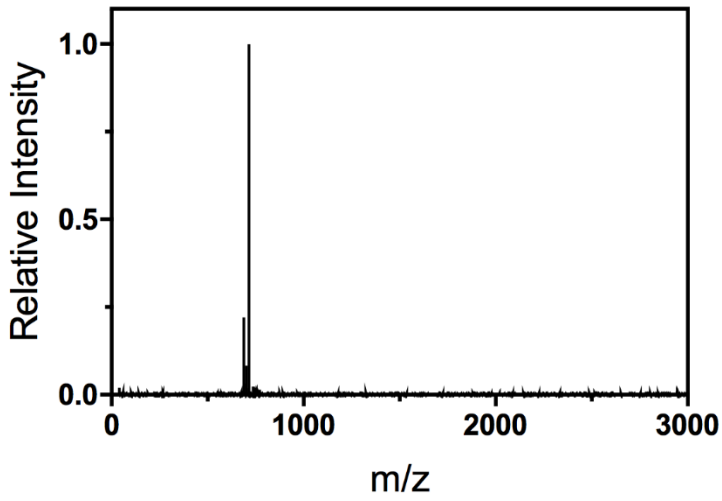
6.3.4 Fmoc solid-phase peptide synthesis

A CEM Liberty 1 was used to perform microwave-assisted Fmoc solid phase peptide synthesis (SPPS, 0.25 mmol scale). Fmoc deprotection was performed in 20% piperidine (v/v) in dimethylformamide (DMF) with 1-hydroxybenzotriazole (HOBt, 0.1 M, 90 °C, 90 sec). Amino acids were coupled to resin-bound peptides upon treatment (75 °C, 5 min) with Fmoc-protected amino acid (2 mmol, 4x), 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU, 2 mmol, 4x), and *N*,*N*-diisopropylethylamine (DIEA, 2 mmol, 4x) in a mixture of DMF (9 mL) and *N*-Methyl-2-pyrrolidone (NMP, 2 mL). Room-temperature couplings were performed without microwave assistance (1 hr).

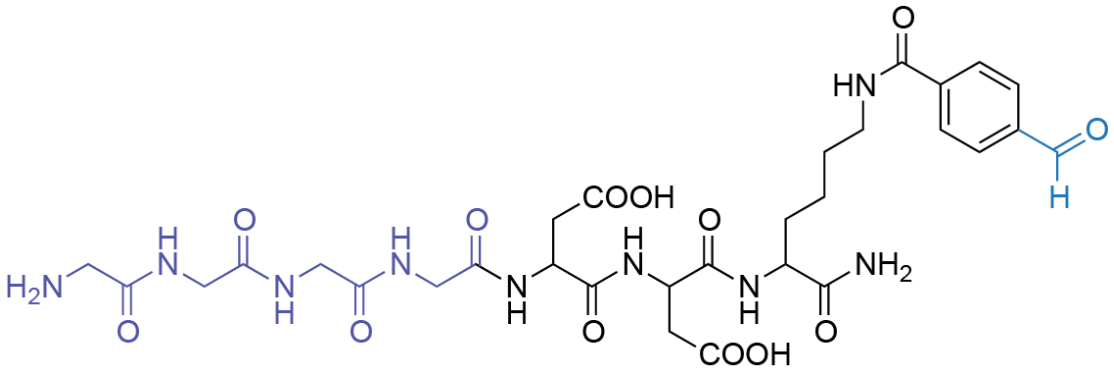
6.3.5 Synthesis of H-GGGGDDK(N_3)-NH_2



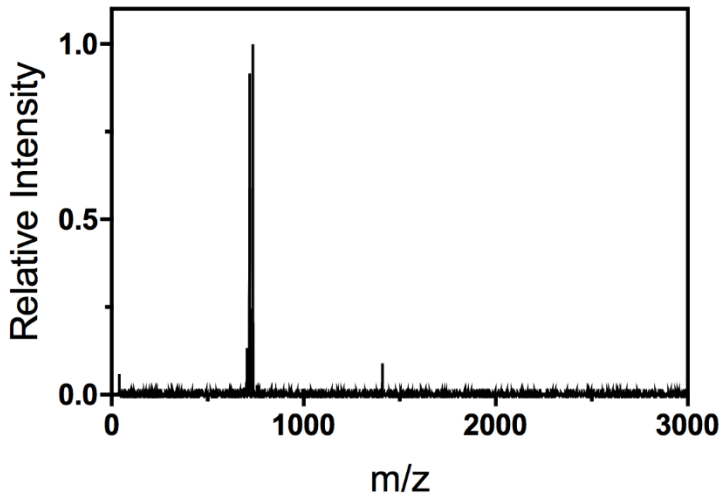
The resin-bound peptide Boc-GGGGDDK(Mtt)-NH_2 was synthesized by Fmoc SPPS (Supplementary Methods) on Rink amide resin (0.25 mmol scale). The resin was washed with DMF (3x) and dichloromethane (DCM, 3x) prior to Mtt cleavage [2 min, 15 mL, 97:2:1 DCM:Triisopropylsilane (TIS):Trifluoroacetic Acid (TFA), 9x]. Resin was washed (DCM, 3x; DMF, 3x) prior to treatment (1 hr) with N_3-COOH (4x, 1 mmol, 129 mg) by HATU coupling (3.95x, 0.988 mmol, 188 mg) and N,N-Diisopropylethylamine (DIEA, 8x, 2 mmol, 174 µL) in minimal DMF. Resin was washed (DMF, 3x; DCM, 3x) prior to peptide cleavage/deprotection (95:5 TFA:H_2O, 20 mL, 2 hr) and precipitation (diethyl ether, 180 mL, 0 °C, 2x). The crude peptide was purified *via* RP-HPLC using a 55-minute gradient from 5-100% acetonitrile:H_2O; lyophilization yielded the final product (H-GGGGDDK(N_3)-NH_2) as a white solid (62 mg, 0.087 mmol, 34% overall yield). Peptide purity was confirmed using MALDI-TOF: calculated for $C_{26}H_{43}N_{12}O_{12}^{+}$ $[M + {}^{1}H]^{+}$, 715.31; observed 715.40.



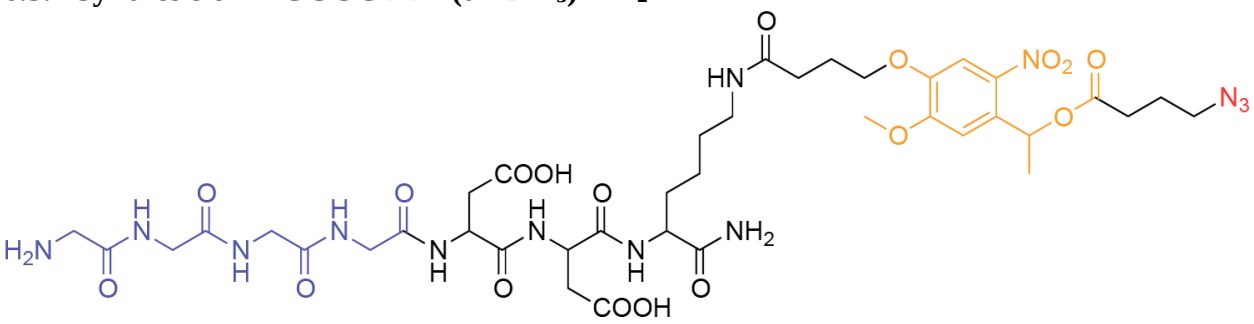
6.3.6 Synthesis of H-GGGGDDK(CHO)-NH_2



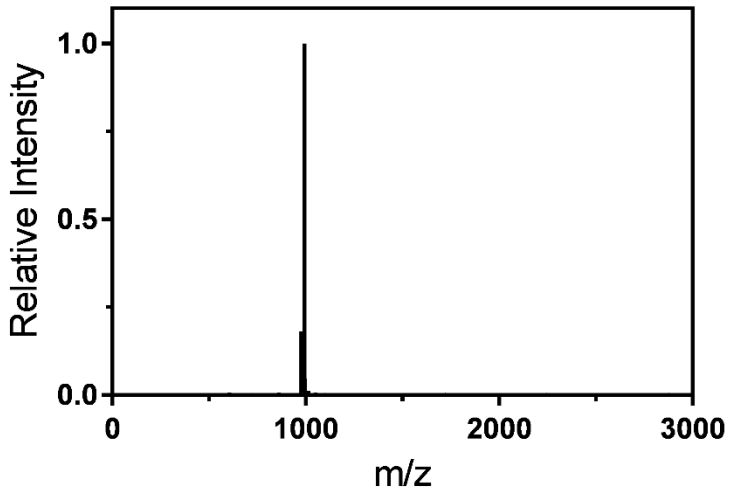
The resin-bound peptide Boc-GGGGDDK(Mtt)-NH_2 was synthesized by Fmoc SPPS (Supplementary Methods) on Rink amide resin (0.25 mmol scale). The resin was washed with DMF (3x) and dichloromethane (DCM, 3x) prior to Mtt cleavage (2 min, 15 mL, 97:2:1 DCM:TIS:TFA, 9x). Resin was washed (DCM, 3x; DMF, 3x) prior to treatment (1 hr) with 4-formylbenzoic acid (4x, 1 mmol, 150 mg) by HATU coupling (3.95x, 0.988 mmol, 188 mg) and DIEA (8x, 2 mmol, 174 µL) in minimal DMF. Resin was washed (DMF, 3x; DCM, 3x) prior to peptide cleavage/deprotection (95:5 TFA:H_2O, 20 mL, 2 hr) and precipitation (diethyl ether, 180 mL, 0 °C, 2x). The crude peptide was purified *via* RP-HPLC using a 55-minute gradient from 5-100% acetonitrile:H_2O; lyophilization yielded the final product (H-GGGGDDK(CHO)-NH_2) as a white solid (39 mg, 0.053 mmol, 21% overall yield). Peptide purity was confirmed using MALDI-TOF: calculated for $C_{30}H_{42}N_9O_{13}^+$ $[M + {}^1H]^+$, 736.29; observed 736.39.



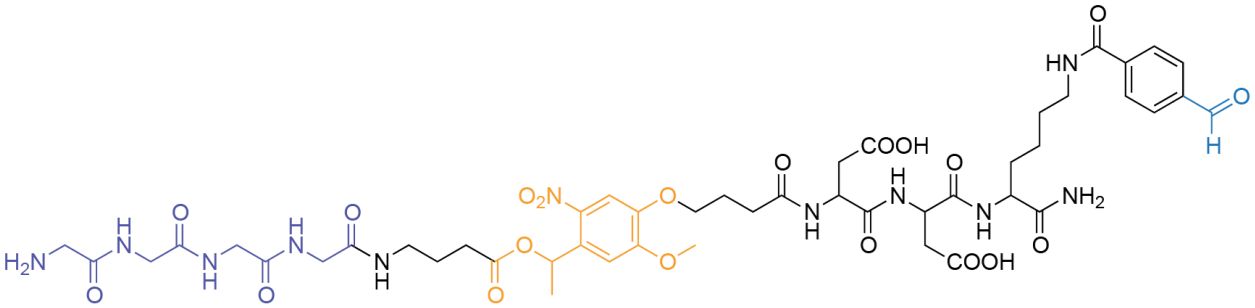
6.3.7 Synthesis of H-GGGGDDK(*o*NB-N_3)-NH_2



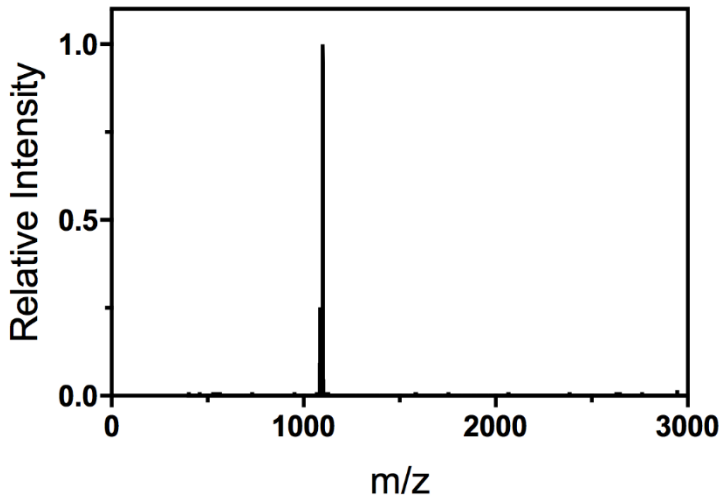
The resin-bound peptide Boc-GGGGDDK(Mtt)-NH_2 was synthesized by Fmoc SPPS (Supplementary Methods) on Rink amide resin (0.25 mmol scale). The resin was washed with DMF (3x) and dichloromethane (DCM, 3x) prior to Mtt cleavage (2x, 9 min, 15 mL, 97:2:1 DCM:TIS:TFA). Resin was washed (DCM, 3x; DMF, 3x) prior to treatment (2 hr) with N_3-*o*NB-OSu (1.5x, 0.375 mmol, 90 mg) and DIEA (4x, 1 mmol, 87 µL) in minimal DMF. Resin was washed (DMF, 3x; DCM, 3x) prior to peptide cleavage/deprotection (95:5 TFA:H_2O, 20 mL, 2 hr) and precipitation (diethyl ether, 180 mL, 0 °C, 2x). The crude peptide was purified *via* RP-HPLC using a 55-minute gradient from 5-100% acetonitrile:H_2O; lyophilization yielded the final product (H-GGGGDDK(*o*NB-N_3)-NH_2) as a yellow solid (70 mg, 0.070 mmol, 27% overall yield). Peptide purity was confirmed using MALDI-TOF: calculated for $C_{39}H_{56}N_{13}O_{18}^-$ [M - ^1H]$^-$, 994.39; observed 993.47.



6.3.8 Synthesis of H-GGGG-*o*NB-DDK(CHO)-NH_2



The resin-bound peptide H-DDK(Mtt)-NH_2 was synthesized by Fmoc SPPS (Supplementary Methods) on Rink amide resin (0.25 mmol scale). Resin was washed (DMF, 3x) prior to treatment with N_3-*o*NB-OSu (1.5x, 0.375 mmol, 90 mg) and DIEA (4x, 1 mmol, 87 µL) in minimal DMF. A Staudinger reduction involving triphenylphosphine (5 wt% in 90:10 tetrahydrofuran:H_2O, 20 mL) was utilized to reduce the N-terminal azide to a primary amine. The *o*NB-containing peptide was elaborated further by Fmoc-based methodologies to introduce Boc-GGGG-OH onto the resin-bound species at room temperature. The resin was washed with DMF (3x) and dichloromethane (DCM, 3x) prior to Mtt cleavage (2 min, 15 mL, 97:2:1 DCM:TIS:TFA, 9x). Resin was washed (DCM, 3x; DMF, 3x) prior to treatment (1 hr) with 4-formylbenzoic acid (4x, 1 mmol, 150 mg) by HATU coupling (3.95x, 0.988 mmol, 188 mg) and DIEA (8x, 2 mmol, 174 µL) in minimal DMF. Resin was washed (DMF, 3x; DCM, 3x) prior to peptide cleavage/deprotection (95:5 TFA:H_2O, 20 mL, 2 hr) and precipitation (diethyl ether, 180 mL, 0 °C, 2x). The crude peptide was purified *via* RP-HPLC using a 55-minute gradient from 5-100% acetonitrile:H_2O; lyophilization yielded the final product (H-GGGG-*o*NB-DDK(CHO)-NH_2) as a yellow solid (39 mg, 0.053 mmol, 21% overall yield). Peptide purity was confirmed using MALDI-TOF: calculated for $C_{47}H_{62}N_{11}O_{20}^-$ [M - ^1H]$^-$, 1100.42; observed 1099.95.



6.3.9 Protein expression and purification by STEPL

LB (500 mL) supplemented with ampicillin (100 μg mL^{-1}) was inoculated with an overnight cell culture (10 mL) and incubated (37 °C) with agitation (250 rev min^{-1}). After reaching an optical density at λ = 600 nm of 0.6, isopropyl β-D-1-thiogalactopyranoside was added (final concentration of 0.5 mM) and expression was continued overnight under reduced temperature (18 °C).

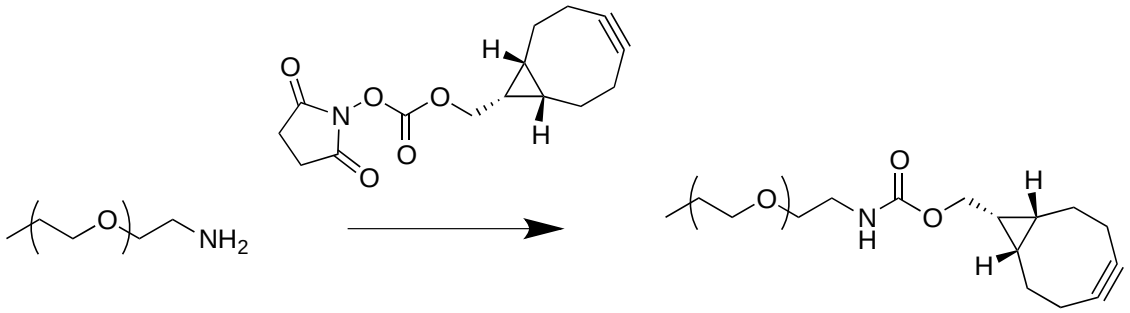
Cells were harvested *via* centrifugation (7,000 g, 10 min). The cell pellet was resuspended in lysis buffer (40 mL, 20 mM Tris, 50 mM NaCl, 10 mM imidazole, 1 mM phenylmethylsulfonyl fluoride) and sonicated on ice (6 cycles of 3 minutes at 30% amplitude 33% duty cycle and 3 min resting). Soluble and insoluble fractions were separated *via* centrifugation (5,000 g, 20 min).

Clarified lysate was applied to Ni-NTA resin (2.5 mL) and incubated under mild agitation (4 °C, 1 hr). The flow-through was discarded, and the resin was washed with wash buffer (20 mM Tris, 50 mM NaCl, 20 mM imidazole, 20 mL, 5x) and STEPL buffer (20 mM Tris, 50 mM NaCl, 20 mL, 1x). The polyglycine probe (20 molar excess) was added to resin in conjugation buffer (20 mM Tris, 50 mM NaCl, 100 μM $CaCl_2$, 2 mL) to promote intramolecular sortagging (37 °C, 4 hr).

The conjugated protein solution was collected, and the resin was washed with STEPL buffer (1 mL, 5x) to collect any remaining protein. The protein solution was dialyzed against STEPL buffer using ThermoFisher SnakeSkin Dialysis Tubing (molecular weight cut-off, MWCO ~ 10 kDa) to remove any unconjugated peptide and concentrated using an Amicon centrifugal spin column (MWCO ~ 10 kDa). Typical yields for purified proteins following STEPL were ~15 mg per liter of cell culture.

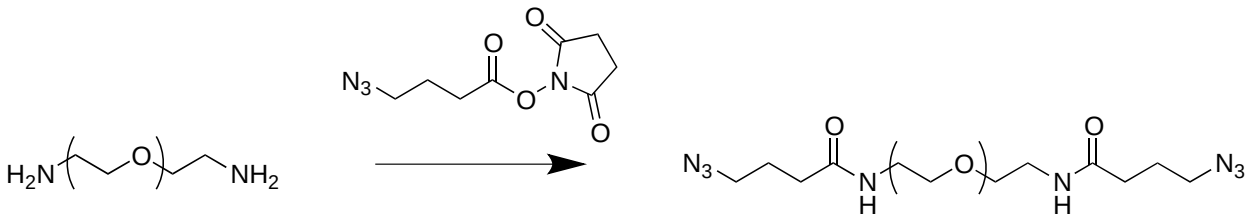
With 6 base proteins and 5 different polyglycine probes, a total of 30 sortagged proteins were generated. Identification codes denoting sortagged species identity are given in Supplementary Table 1.

6.3.10 Synthesis of mPEG-BCN



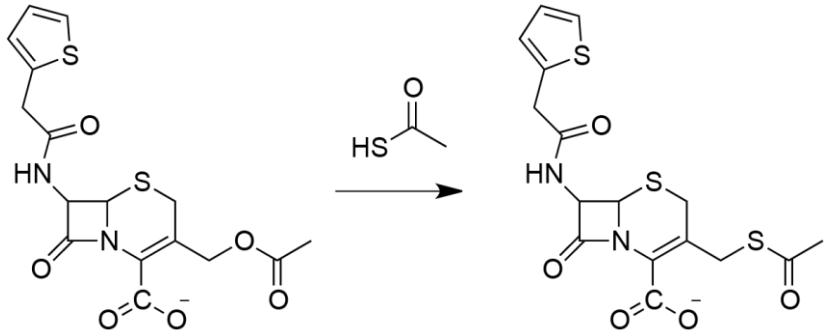
Methoxy poly(ethylene glycol) amine (M_n ~ 10,000 Da, 200 mg, 0.02 mmol NH_2, 1x, Jenkem) and BCN-OSu (8.73 mg, 0.03 mmol, 1.5x) were dissolved in DMF (2 mL). DIEA (6.9 µL, 5.1 mg, 0.08 mmol, 4x) was added to the mixture, and the reaction was stirred overnight, dissolved in water, dialyzed (MWCO ~ 2 kDa, SpectraPor), and lyophilized to yield a white powder (200 mg, quantitative yield). ^{1}H NMR (500 MHz, $CDCl_3$) δ 4.16 (d, 2H), 3.65 – 3.62 (m, 909H), 3.50 – 3.47 (m, 4H), 3.23 (s, 3H), 2.24 (m, 6H), 1.60 (m, 2H), 1.39 (m, 2H), 1.25 (m, 1H) 0.95 (m, 2H). Functionalization was found be >95% by ^{1}H-NMR by comparing integral values for characteristic BCN peaks (δ 2.24, 1.60, 1.39, 0.95) with those from the methoxy peak (δ 3.23).

6.3.11 Synthesis of PEG-diazide (N_3-PEG-N_3)



Linear poly(ethylene glycol) diamine (M_n ~ 3,500 Da, 1 g, 0.57 mmol NH_2, 1x, Jenkem) and N_3-OSu (194 mg, 0.86 mmol, 1.5x) were dissolved in dimethylformamide (5 mL). *N*,*N*-Diisopropylethylamine (398 µL, 294 mg, 2.28 mmol, 4x) was added to the mixture, and the reaction was stirred overnight, diluted in water (15 mL), dialyzed (MWCO ~ 2 kDa, SpectraPor), and lyophilized to yield a white powder (1.00 g, quantitative yield). ^{1}H NMR (500 MHz, $CDCl_3$) δ 3.75 (m, 4H), 3.65-3.61 (m, 318H), 3.28 (m, 4H), 2.35 (m, 4H), 1.86 (m, 4H). Functionalization was found to be >95% by comparing integral values for hydrogens introduced upon azide coupling (δ 3.28, 2.35, 1.86) with those from the PEG backbone (δ 3.60-3.42).

6.3.12 Synthesis of thioacetate cefalotin



Thioacetate cefalotin was synthesized based on a known synthetic route[4]. Briefly, cefalotin (1x, 0.85 mmol, 357.5 mg), thioacetic acid (2.55 mmol, 194 mg), and sodium bicarbonate (3x, 2.55 mmol, 214 mg) were dissolved in water (10 mL) and stirred (30 hr) at elevated temperatures (50 °C). The reaction mixture was frozen, lyophilized, and redissolved in methanol (5 mL). The product was recrystallized in ice-cold diethyl ether (45 mL, 0 °C, 2x), isolated by vacuum filtration, washed with ethyl acetate, and dried over vacuum to yield the final product (thioacetate cefalotin) as a light brown solid (0.554 mmol, 224 mg, 64% yield). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.64 (s, 1H), 6.99 (s, 2H), 5.55 (s, 1H), 5.03 (s, 1H), 4.02-3.99 (s, 2H), 3.90-3.86 (m, 3H), 3.84-3.81 (m, 1H), 2.38-2.34 (s, 3H).

6.3.13 Plasmid construction for EGFP-EGF expression and STEPL purification

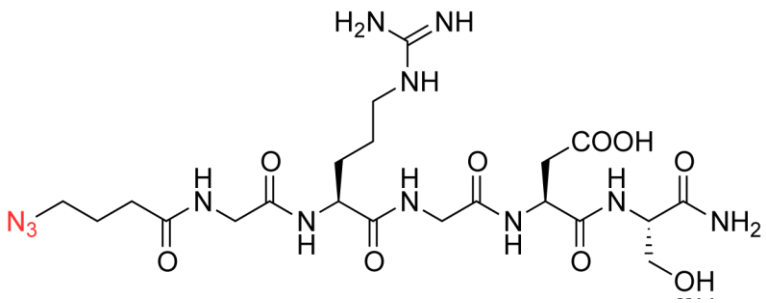
A STEPL plasmid expression system was constructed for an EGFP-EGF fusion following the strategies outlined previously. PCR was used to amplify the gene sequences of EGFP and EGF and introduce relevant restriction sites (5' NdeI and 3' BamHI for EGFP and 5' BamHI and 3' XhoI for EGF) for cloning. The EGFP-EGF STEPL plasmid was purified and subsequently transformed into chemically competent BL21(DE3) *E. coli* (Promega) for expression.

Below is a list of DNA open reading frame sequences (5' → 3') generated for and utilized in these studies. Nucleotides shown in green correspond to EGFP, while those in purple correspond to EGF. Bases shown in orange correspond to STEPL portion, which features the C-terminal LPETG sortase recognition motif, a flexible $(GGS)_5$ linker, SrtA, and a 6xHis tag.

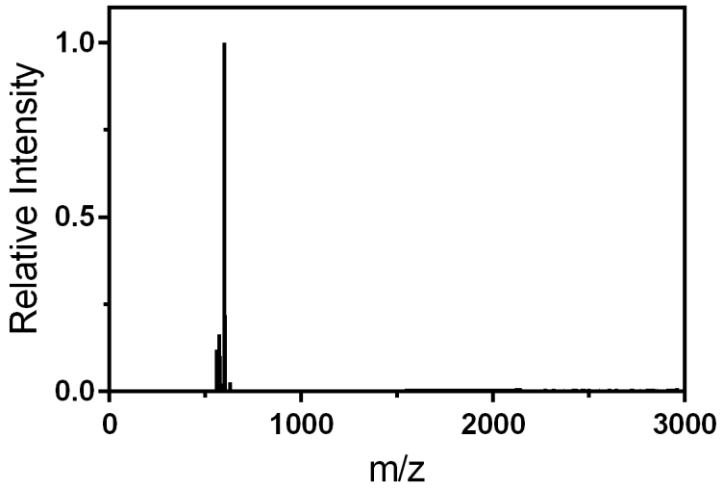
EGFP-*EGF*-*STEPL*:

ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCG
ACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCT
GACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACC
CTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCA
AGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTA
CAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGC
ATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACA
ACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAA
CATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGC
CCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACG
AGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGA
CGAGCTGTACAAGGGATCCATGAACAGCGACAGCGAGTGCCCACTGAGCCACGACGGCTACTGC
CTGCACGACGGCGTGTGCATGTACATCGAGGCCCTGNNCAAGTACGCCTGCAACTGCGTCGTGG
GCTACATCGGCGAGCGGTGCCAGTACCGGGACCTGAAGTGGTGGGAGCTGAGACTCGAGCTGCC
GGAAACCGGTGGTGGTAGTGGTGGCTCTGGCGGTTCTGGTGGCAGTGGCGGTAGCCAAGCTAAA
CCTCAAATTCCGAAAGATAAATCAAAAGTGGCAGGCTATATTGAAATTCCAGATGCTGATATTA
AAGAACCAGTATATCCAGGACCAGCAACACCTGAACAATTAAATAGAGGTGTAAGCTTTGCAGA
AGAAAATGAATCACTAGATGATCAAAATATTTCAATTGCAGGACACACTTTCATTGACCGTCCG
AACTATCAATTTACAAATCTTAAAGCAGCCAAAAAAGGTAGTATGGTGTACTTTAAAGTTGGTA
ATGAAACACGTAAGTATAAAATGACAAGTATAAGAGATGTTAAGCCAACAGATGTAGAAGTTCT
AGATGAACAAAAAGGTAAAGATAAACAATTAACATTAATTACTTGTGATGATTACAATGAAAAG
ACAGGCGTTTGGGAAAAACGTAAAATCTTTGTAGCTACAGAAGTCAAACATCACCACCATCATC
ACTAA

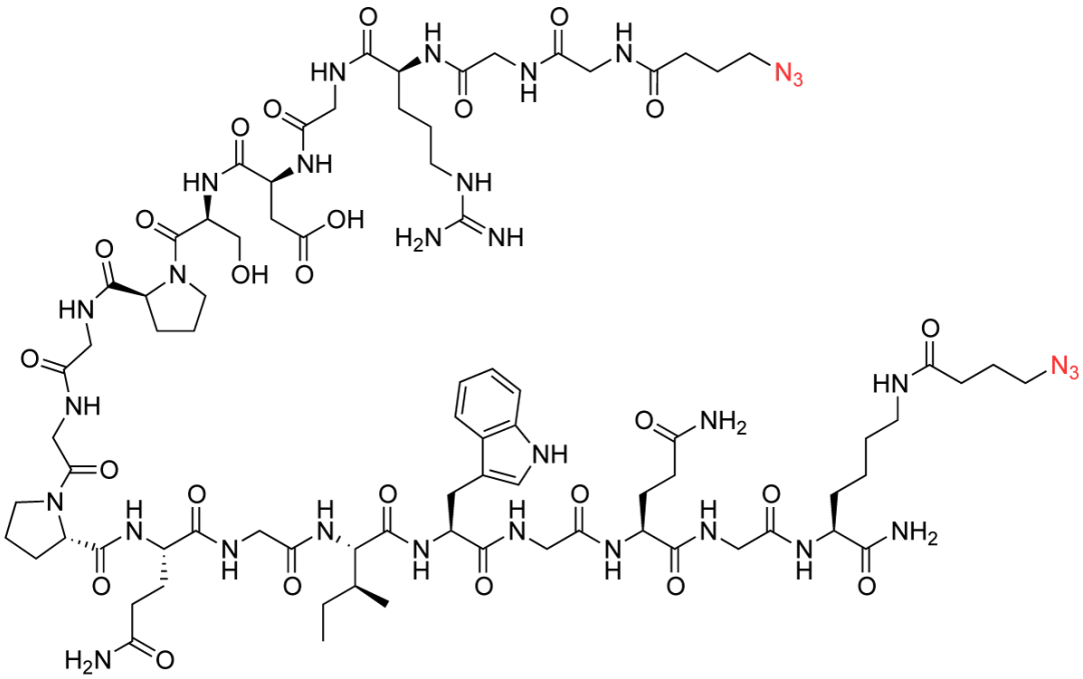
6.3.14 Synthesis of N_3-GRGDS-NH_2



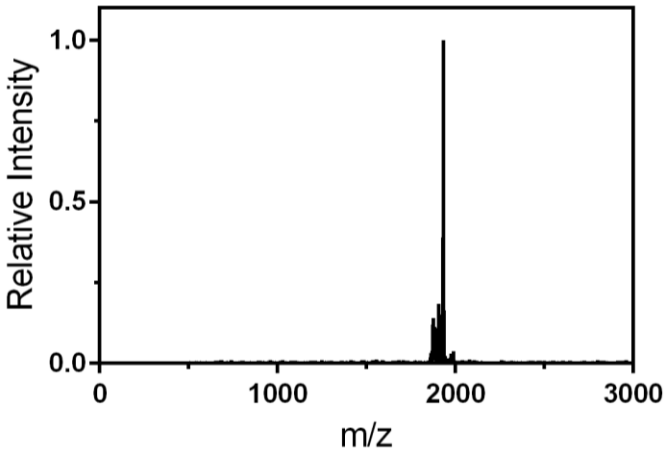
The resin-bound peptide H-GRGDS-NH_2 was synthesized by Fmoc SPPS (Supplementary Methods) on Rink amide resin (0.25 mmol scale). Resin was washed (DMF, 3x) prior to treatment (1 hr) with N_3-COOH (4x, 1 mmol, 129 mg), HATU (3.95x, 0.988 mmol, 376 mg), and DIEA (8x, 2 mmol, 174 μL) in minimal DMF. Resin was washed (DMF, 3x; DCM, 3x) prior to peptide cleavage/deprotection (95:5 TFA:H_2O, 20 mL, 2 hr) and precipitation (diethyl ether, 180 mL, 0 °C, 2x). The crude peptide was purified *via* RP-HPLC using a 55-minute gradient from 5-100% acetonitrile:H_2O; lyophilization yielded the final product (denoted N_3-GRGDS-NH_2) as a white solid (72 mg, 0.12 mmol, 48% overall yield). Peptide purity was confirmed using MALDI-TOF: calculated for $C_{21}H_{37}N_{12}O_9^+$ $[M + {}^1H]^+$, 601.59; observed 600.86.



6.3.15 Synthesis of N_3-GGRGDSPGGPQGIWGQGK(N_3)-NH_2

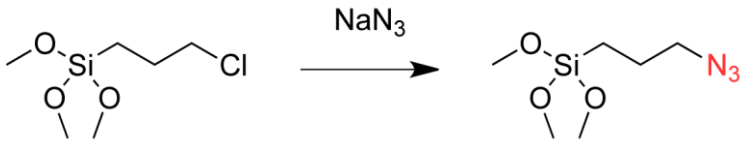


The resin-bound peptide H-GGRGDSPGGPQGIWGQGK(Dde)-NH_2 was synthesized by Fmoc SPPS (Supplementary Methods) on Rink amide resin (0.25 mmol scale). The resin was washed with DMF (3x) prior to Dde (1-(4,4-dimethyl-2,6-dioxacyclohexylidene)ethyl) deprotection (10 min, 30 mL, 2% hydrazine in DMF, 3x), yielding a peptide with free N-terminal amine and a free ε-amino group on the unprotected lysine. Resin was washed (DMF, 3x) prior to treatment (1 hr) with N_3-COOH (4x, 2 mmol, 258 mg), HATU (3.95x, 1.976 mmol, 752 mg), and DIEA (8x, 4 mmol, 348 μL) in minimal DMF. Resin was washed (DMF, 3x; DCM, 3x) prior to peptide cleavage/deprotection (95:5 TFA:H_2O, 20 mL, 2 hr) and precipitation (diethyl ether, 180 mL, 0 °C, 2x). The crude peptide was purified *via* RP-HPLC using a 55-minute gradient from 5-100% acetonitrile:H_2O; lyophilization yielded the final product [denoted N_3-GGRGDSPGGPQGIWGQGK(N_3)-NH_2] as a light yellow solid (123 mg, 0.064 mmol, 25% overall yield). Peptide purity was confirmed using MALDI-TOF: calculated for $C_{80}H_{123}N_{32}O_{25}^+$ $[M + {}^1H]^+$, 1933.02; observed 1932.53.


Relative Intensity
1.0
0.5
0.0
0
1000
2000
3000
m/z

6.3.16 Synthesis of azide-functionalized glass slides

Synthesis of (3-azidopropyl)trimethoxysilane



(3-Chloropropyl)trimethoxysilane (11.9 mL, 12.9 g, 65.3 mmol), sodium azide (6.38 g, 98 mmol, 1.5x), and dry DMF (40 mL) were added to a flame-dried round-bottom flask (250 mL) and stirred overnight at 100 °C. The reaction was cooled to room temperature and diluted with diethyl ether:water (1:1, 150 mL). The organic layer was washed with water (3x) and brine before drying over $MgSO_4$. The mixture was filtered and concentrated under vacuum to yield product (3-azidopropyl)trimethoxysilane as a pale yellow oil (12.74 g, 95% yield). ^{1}H NMR (500 MHz, $CDCl_3$) δ 3.57 (s, 9H), 3.26 (t, J = 7.0 Hz, 2H), 1.71 (m, 2H), 0.70 (m, 2H). Characterization matched that previously reported[5].

Surface cleaning and functionalization of glass slides

Glass slides were cleaned (30 min) with piranha solution (35% H_2O, 15% H_2O_2, 50% H_2SO_4) before washing with water (3x) and acetone (3x). Dry slides were treated (90 min) with (3-azidopropyl)trimethoxysilane (70 mM) and n-butylamine (70 mM) in toluene (100 mL). Slides were rinsed with toluene (3x), wiped dry, and heated (80 °C) overnight. Azide-functionalized slides were stored at room temperature and used as needed.

6.4 Supplementary References

1. DeForest, C. A. & Tirrell, D. A. A photoreversible protein-patterning approach for guiding stem cell fate in three-dimensional gels. *Nat. Mater.* **14,** 523–531 (2015).

2. Warden-Rothman, R., Caturegli, I., Popik, V. & Tsourkas, A. Sortase-tag expressed protein ligation: Combining protein purification and site-specific bioconjugation into a single step. *Anal. Chem.* **85,** 11090–11097 (2013).

3. Carrico, I. S., Carlson, B. L. & Bertozzi, C. R. Introducing genetically encoded aldehydes into proteins. *Nat. Chem. Biol.* **3,** 321–322 (2007).

4. Bieniarz, C., Young, D. F. & Cornwell, M. J. Chromogenic redox assay for beta-lactamases yielding water-insoluble products. I. Kinetic behavior and redox chemistry. *Anal. Biochem.* **207,** 321–8 (1992).

5. Godula, K., Rabuka, D., Nam, K. T. & Bertozzi, C. R. Synthesis and microcontact printing of dual end-functionalized mucin-like glycopolymers for microarray applications. *Angew. Chemie - Int. Ed.* **48,** 4973–4976 (2009).

Chapter 7

Appendix A2: Supporting Information for

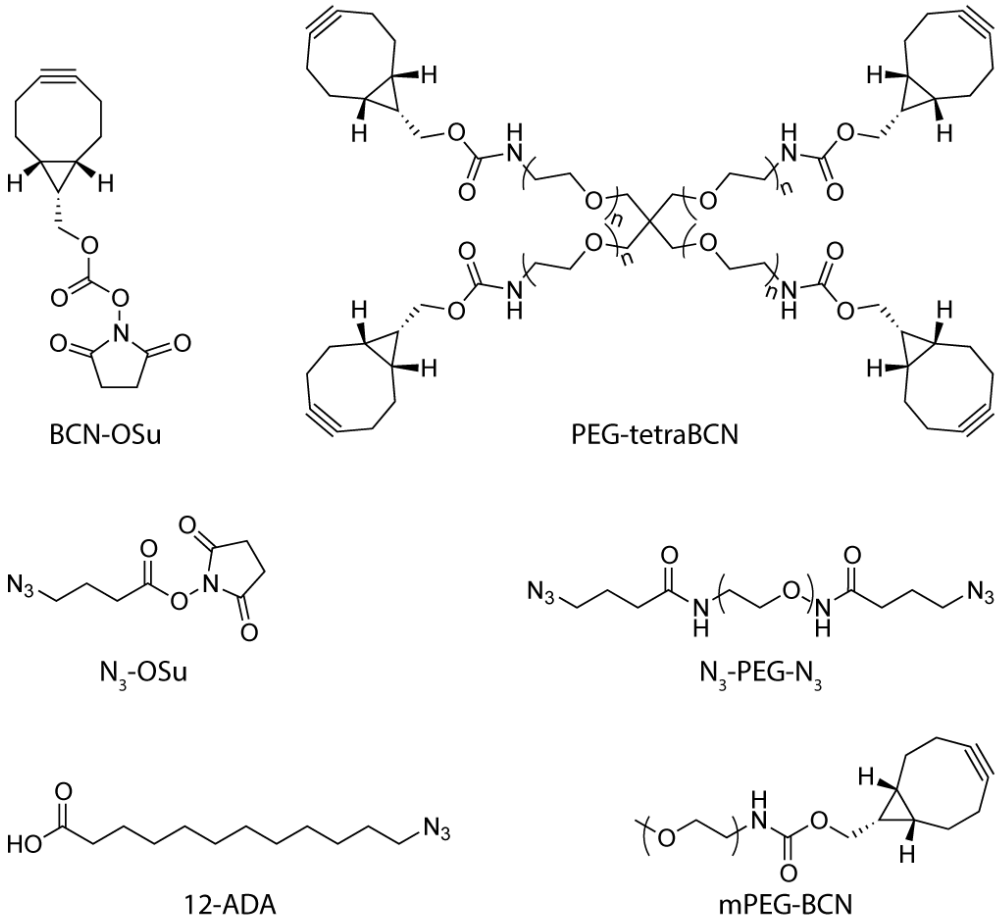
Genetically Encoded Photocleavable Linkers for Patterned Protein Release from Biomaterials

Jared A. Shadish, Alder C. Strange, Cole A. DeForest

7.1 General synthetic information

Chemical reagents and solvents were purchased from either Sigma-Aldrich or Fisher Scientific and used as received unless otherwise noted. Peptide synthesis reagents were purchased from either ChemPep or Chem-Impex and used as received. Deionized water (dH_2O) was generated by a U.S. Filter Corporation Reverse Osmosis System with a Desal membrane. Synthetic chemical reactions were performed under a nitrogen atmosphere in oven-dried glassware and stirred with a Teflon-coated magnetic stir bar unless otherwise noted. Solvents were removed *in vacuo* with a Büchi Rotovapor R-3 equipped with a V-700 vacuum pump and V-855 vacuum controller and a Welch 1400 DuoSeal Belt-Drive high vacuum pump. ^1H nuclear magnetic resonance (NMR) data was collected at 298 K on Bruker instruments and chemical shifts are reported relative to tetramethylsilane (TMS, $\delta = 0$). Lyophilization was performed on a LABCONCO FreeZone 2.5 Plus freeze-dryer equipped with a LABCONCO rotary vane 117 vacuum pump. Whole-protein mass spectrometry was performed using a Waters Synapt – G2 QTOF. The visible light source used for photomediated protein release was a Mightex Systems WheeLED Wavelength-Switchable LED Source equiped with a 5W 400 nm LED. Light intensity was measured using an OmniCure Radiometer (R2000). Biochemical gradients were created with the aid of a Harvard Apparatus PHD 2000 Syringe Pump. Fluorescence readings were acquired on a SpectraMax M5 spectrometer using Thermo Scientific Nunc black polypropylene 96-well plates. Confocal microscopy and laser scanning lithography was performed at the University of Washington Keck Microscopy Center on a Leica SP8X confocal microscope. Polymerase chain reaction (PCR) was performed in a Bioer LifeECO thermal cycler. Protein expression was performed in a Thermo Scientific MaxQ 4000 shaker incubator. Cells were lysed using a Fisher Scientific Model 505 Sonic Dismembrator with a 1.27 cm diameter probe. Mammalian cell culture was performed in a NuAire LabGard ES NU-437 Class II Type A2 Biosafety Cabinet. Cells were maintained in a Sanyo inCu saFe® MCO-17AC incubator at 37 °C and 5% CO_2.

7.2 Synthesis of previously reported compounds used in this work



(1R,8S,9s)-bicyclo[6.1.0]non-4-yn-9-ylmethyl (2,5-dioxopyrrolidin-1-yl) carbonate (BCN-OSu), poly(ethylene glycol) tetrabicyclononyne (PEG-tetraBCN, M_n ~ 20,000 Da), 2,5-dioxopyrrolidin-1-yl 4-azidobutanoate (N_3-OSu), poly(ethylene glycol) diazide (N_3-PEG-N_3, M_n ~ 3,400 Da), 12-azidododecanoic acid (12-ADA), and methoxy poly(ethylene glycol) bicycloononyne (mPEG-BCN, M_n ~ 2,000 Da or 5,000 Da) were synthesized as previously reported.[1–3]

7.3 Supplementary Methods

7.3.1 Method S1 Plasmid construction for azide-tagged PhoCl-fusion protein expression

PhoCl Plasmid Construction:

A gBlock containing the PhoCl domain with a 5' NdeI site and a 3' HindIII site and an N-terminal NMT recognition sequence and 6xHis tag (MGNEASYPLHHHHHH) was purchased from Integrated DNA Technologies. Polymerase chain reaction (PCR) was used to amplify gene sequences of interest and introduce relevant restriction sites (5' HindIII and 3' XhoI) for ligation with the pET21 expression plasmid (Novagen) and the 3' end of the PhoCl sequence.

Restriction enzymes (New England BioLabs) digested (4 hr, 37 °C) the pET21 plasmid (NdeI and XhoI), PhoCl gene sequence (NdeI and HindIII), and PCR products (HindIII and XhoI), and the digested products were purified by extraction following electrophoretic separation (0.8% agarose). A double ligation was used to insert the PhoCl and fusion protein genes into the pET21 plasmid (T4 DNA ligase, 16 hr, 16 °C). The plasmid was transformed into chemically competent Top10 *E. coli* (Thermo Fisher) by heat shock (42 °C, 45 seconds), and plated onto agar plates (10 g Tryptone, 5 g Yeast Extract, 10 g NaCl, 15 g agar, 1 L dH_2O) containing ampicillin (100 µg mL^{-1}). Colonies were subsequently grown overnight in Miller's Lysogeny Broth (LB, 5 mL) containing ampicillin (100 µg mL^{-1}). Plasmids were purified using a QIAprep Spin Miniprep Kit (Qiagen), and sequenced using a SimpleSeq DNA Sequencing Kit (Thermo Fisher). Plasmids corresponding to the construct of interest were purified and subsequently cotransformed into chemically competent BL21(DE3) *E. coli* (Promega) along with the NMT1 Met/AP plasmid (a gift from the Kahn lab at Emory University) via heat shock (42 °C, 45 seconds) and plated onto agar plates containing both ampicillin (100 µg mL^{-1}) and kanamycin (50 µg mL^{-1}). Single colonies were selected for subsequent protein expression.

Below is a list of DNA open reading frame sequences generated for and utilized in these studies. Sequences are given 5' to 3' with the NMT recognition sequence (RS1) and 6xHis tag in red, the PhoCl domain in purple, and the protein of interest in black.

RS1-6xHis-PhoCl-sfGFP:

ATGGGAAATGAGGCGTCGTATCCGTTACACCATCACCATCATCACATGGTTATCCCGGACTACTTCAAACAGTCTTT
CCCGGAAGGTTACTCTTGGGAACGTTCTATGACCTACGAAGACGGTGGTATCTGCATCGCTACCAACGACATCACCA
TGGAAGGTGACTCTTTCATCAACAAAATCCACTTCAAAGGTACCAACTTCCCGCCGAACGGTCCGGTTATGCAGAAA
CGTACCGTTGGTTGGGAAGCTAGTACCGAAAAAATGTACGAACGTGACGGTGTTCTGAAAGGTGACGTTAAAATGAA
ACTGCTGCTGAAAGGTGGTGGTCACTACCGTTGCGACTACCGTACCACCTACAAAGTTAAACAGAAACCGGTTAAAC
TGCCGGACTACCACTTCGTTGACCACCGTATCGAAATCCTGTCTCACGACAAAGACTACAACAAAGTTAAACTGTAC
GAACACGCTGTTGCTCGTAACTCTACCGACTCTATGGACGAACTGTACAAAGGTGGTTCTGGTGGTATGGTTTCTAA
AGGTGAAGAAACCATCACCTCTGTTATCAAACCGGACATGAAAAACAAACTGCGTATGGAAGGTAACGTTAACGGTC
ACGCTTTCGTTATCGAAGGTGAAGGTTCTGGTAAACCGTTCGAAGGTATCCAGACCATCGACCTGGAAGTTAAAGAA
GGTGCTCCGCTGCCGTTCGCTTACGACATCCTGACCACCGCTTTCCACTACGGTAACCGTGTTTTCACCAAATACCC
GCGTGGTGGTGGTAAGCTTATGCGTAAAGGCGAAGAGCTGTTCACTGGTGTCGTCCCTATTCTGGTGGAACTGGATG
GTGATGTCAACGGTCATAAGTTTTCCGTGCGTGGCGAGGGTGAAGGTGACGCAACTAATGGTAAACTGACGCTGAAG
TTCATCTGTACTACTGGTAAACTGCCGGTACCTTGGCCGACTCTGGTAACGACGCTGACTTATGGTGTTCAGTGCTT

TGCTCGTTATCCGGACCACATGAAGCAGCATGACTTCTTCAAGTCCGCCATGCCGGAAGGCTATGTGCAGGAACGCA
CGATTTCCTTTAAGGATGACGGCACGTACAAAACGCGTGCGGAAGTGAAATTTGAAGGCGATACCCTGGTAAACCGC
ATTGAGCTGAAAGGCATTGACTTTAAAGAAGACGGCAATATCCTGGGCCATAAGCTGGAATACAATTTTAACAGCCA
CAATGTTTACATCACCGCCGATAAACAAAAAAATGGCATTAAAGCGAATTTTAAAATTCGCCACAACGTGGAGGATG
GCAGCGTGCAGCTGGCTGATCACTACCAGCAAAACACTCCAATCGGTGATGGTCCTGTTCTGCTGCCAGACAATCAC
TATCTGAGCACGCAAAGCGTTCTGTCTAAAGATCCGAACGAGAAACGCGATCACATGGTTCTGCTGGAGTTCGTAAC
CGCAGCGGGCATCACGCATGGTATGGATGAACTGTACAAATAA

RS1-6xHis-PhoCl-mRuby

ATGGGAAATGAGGCGTCGTATCCGTTACACCATCACCATCATCACATGGTTATCCCGGACTACTTCAAACAGTCTTT
CCCGGAAGGTTACTCTTGGGAACGTTCTATGACCTACGAAGACGGTGGTATCTGCATCGCTACCAACGACATCACCA
TGGAAGGTGACTCTTTCATCAACAAAATCCACTTCAAAGGTACCAACTTCCCGCCGAACGGTCCGGTTATGCAGAAA
CGTACCGTTGGTTGGGAAGCTAGTACCGAAAAAATGTACGAACGTGACGGTGTTCTGAAAGGTGACGTTAAAATGAA
ACTGCTGCTGAAAGGTGGTGGTCACTACCGTTGCGACTACCGTACCACCTACAAAGTTAAACAGAAACCGGTTAAAC
TGCCGGACTACCACTTCGTTGACCACCGTATCGAAATCCTGTCTCACGACAAAGACTACAACAAAGTTAAACTGTAC
GAACACGCTGTTGCTCGTAACTCTACCGACTCTATGGACGAACTGTACAAAGGTGGTTCTGGTGGTATGGTTTCTAA
AGGTGAAGAAACCATCACCTCTGTTATCAAACCGGACATGAAAAACAAACTGCGTATGGAAGGTAACGTTAACGGTC
ACGCTTTCGTTATCGAAGGTGAAGGTTCTGGTAAACCGTTCGAAGGTATCCAGACCATCGACCTGGAAGTTAAAGAA
GGTGCTCCGCTGCCGTTCGCTTACGACATCCTGACCACCGCTTTCCACTACGGTAACCGTGTTTTCACCAAATACCC
GCGTGGTGGTGGTAAGCTTGTTTCTAAAGGTGAAGAACTGATCAAAGAAAACATGCGTATGAAAGTTGTTATGGAAG
GTTCTGTTAACGGTCACCAGTTCAAATGCACCGGTGAAGGTGAAGGTCGTCCGTACGAAGGTGTTCAGACCATGCGT
ATCAAAGTTATCGAAGGTGGTCCGCTGCCGTTCGCTTTCGACATCCTGGCTACCTCTTTCATGTACGGTTCTCGTAC
CTTCATCAAATACCCGGCTGACATCCCGGACTTCTTCAAACAGTCTTTCCCGGAAGGTTTCACCTGGGAACGTGTTA
CCCGTTACGAAGACGGTGGTGTTGTTACCGTTACCCAGGACACCTCTCTGGAAGACGGTGAACTGGTTTACAACGTT
AAAGTTCGTGGTGTTAACTTCCCGTCTAACGGTCCGGTTATGCAGAAAAAAACCAAAGGTTGGGAACCGAACACCGA
AATGATGTACCCGGCTGACGGTGGTCTGCGTGGTTACACCGACATCGCTCTGAAAGTTGACGGTGGTGGTCACCTGC
ACTGCAACTTCGTTACCACCTACCGTTCTAAAAAAACCGTTGGTAACATCAAAATGCCGGGTGTTCACGCTGTTGAC
CACCGTCTGGAACGTATCGAAGAATCTGACAACGAAACCTACGTTGTTCAGCGTGAAGTTGCTGTTGCTAAATACTC
TAACCTGGGTGGTGGTATGGACGAACTGTACAAATAA

RS1-6xHis-PhoCl-bla:

ATGGGAAATGAGGCGTCGTATCCGTTACACCATCACCATCATCACATGGTTATCCCGGACTACTTCAAACAGTCTTT
CCCGGAAGGTTACTCTTGGGAACGTTCTATGACCTACGAAGACGGTGGTATCTGCATCGCTACCAACGACATCACCA
TGGAAGGTGACTCTTTCATCAACAAAATCCACTTCAAAGGTACCAACTTCCCGCCGAACGGTCCGGTTATGCAGAAA
CGTACCGTTGGTTGGGAAGCTAGTACCGAAAAAATGTACGAACGTGACGGTGTTCTGAAAGGTGACGTTAAAATGAA
ACTGCTGCTGAAAGGTGGTGGTCACTACCGTTGCGACTACCGTACCACCTACAAAGTTAAACAGAAACCGGTTAAAC
TGCCGGACTACCACTTCGTTGACCACCGTATCGAAATCCTGTCTCACGACAAAGACTACAACAAAGTTAAACTGTAC
GAACACGCTGTTGCTCGTAACTCTACCGACTCTATGGACGAACTGTACAAAGGTGGTTCTGGTGGTATGGTTTCTAA
AGGTGAAGAAACCATCACCTCTGTTATCAAACCGGACATGAAAAACAAACTGCGTATGGAAGGTAACGTTAACGGTC
ACGCTTTCGTTATCGAAGGTGAAGGTTCTGGTAAACCGTTCGAAGGTATCCAGACCATCGACCTGGAAGTTAAAGAA
GGTGCTCCGCTGCCGTTCGCTTACGACATCCTGACCACCGCTTTCCACTACGGTAACCGTGTTTTCACCAAATACCC
GCGTGGTGGTGGTAAGCTTATGCACCCTGAAACGCTGGTGAAAGTAAAANATGCTGAAGATCAGTTGGGTGCACGAG
TGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATG
AGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTGTTGACGCCGGGCAAGAGCAACTCGGTCGCCGCAT
ACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAG
AATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAG
GAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGC
CATACCAAACGACGAGCGTGACACCACGATGCCTGCAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAAC
TACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCG
GCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACT
GGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATA
GACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAA

RS1-6xHis-PhoCl-EGF:

ATGGGAAATGAGGCGTCGTATCCGTTACACCATCACCATCATCACATGGTTATCCCGGACTACTTCAAACAGTCTTT
CCCGGAAGGTTACTCTTGGGAACGTTCTATGACCTACGAAGACGGTGGTATCTGCATCGCTACCAACGACATCACCA
TGGAAGGTGACTCTTTCATCAACAAAATCCACTTCAAAGGTACCAACTTCCCGCCGAACGGTCCGGTTATGCAGAAA
CGTACCGTTGGTTGGGAAGCTAGTACCGAAAAAATGTACGAACGTGACGGTGTTCTGAAAGGTGACGTTAAAATGAA
ACTGCTGCTGAAAGGTGGTGGTCACTACCGTTGCGACTACCGTACCACCTACAAAGTTAAACAGAAACCGGTTAAAC
TGCCGGACTACCACTTCGTTGACCACCGTATCGAAATCCTGTCTCACGACAAAGACTACAACAAAGTTAAACTGTAC
GAACACGCTGTTGCTCGTAACTCTACCGACTCTATGGACGAACTGTACAAAGGTGGTTCTGGTGGTATGGTTTCTAA
AGGTGAAGAAACCATCACCTCTGTTATCAAACCGGACATGAAAAACAAACTGCGTATGGAAGGTAACGTTAACGGTC
ACGCTTTCGTTATCGAAGGTGAAGGTTCTGGTAAACCGTTCGAAGGTATCCAGACCATCGACCTGGAAGTTAAAGAA
GGTGCTCCGCTGCCGTTCGCTTACGACATCCTGACCACCGCTTTCCACTACGGTAACCGTGTTTTCACCAAATACCC
GCGTGGTGGTGGTAAGCTTATGAACAGCGACAGCGAGTGCCCACTGAGCCACGACGGCTACTGCCTGCACGACGGCG
TGTGCATGTACATCGAGGCCCTGGACAAGTACGCCTGCAACTGCGTCGTGGGCTACATCGGCGAGCGGTGCCAGTAC
CGGGACCTGAAGTGGTGGGAGCTGAGACTAA

RS1-6xHis-PhoCl-mCerulean:

ATGGGAAATGAGGCGTCGTATCCGTTACACCATCACCATCATCACATGGTTATCCCGGACTACTTCAAACAGTCTTT
CCCGGAAGGTTACTCTTGGGAACGTTCTATGACCTACGAAGACGGTGGTATCTGCATCGCTACCAACGACATCACCA
TGGAAGGTGACTCTTTCATCAACAAAATCCACTTCAAAGGTACCAACTTCCCGCCGAACGGTCCGGTTATGCAGAAA
CGTACCGTTGGTTGGGAAGCTAGTACCGAAAAAATGTACGAACGTGACGGTGTTCTGAAAGGTGACGTTAAAATGAA
ACTGCTGCTGAAAGGTGGTGGTCACTACCGTTGCGACTACCGTACCACCTACAAAGTTAAACAGAAACCGGTTAAAC
TGCCGGACTACCACTTCGTTGACCACCGTATCGAAATCCTGTCTCACGACAAAGACTACAACAAAGTTAAACTGTAC
GAACACGCTGTTGCTCGTAACTCTACCGACTCTATGGACGAACTGTACAAAGGTGGTTCTGGTGGTATGGTTTCTAA
AGGTGAAGAAACCATCACCTCTGTTATCAAACCGGACATGAAAAACAAACTGCGTATGGAAGGTAACGTTAACGGTC
ACGCTTTCGTTATCGAAGGTGAAGGTTCTGGTAAACCGTTCGAAGGTATCCAGACCATCGACCTGGAAGTTAAAGAA
GGTGCTCCGCTGCCGTTCGCTTACGACATCCTGACCACCGCTTTCCACTACGGTAACCGTGTTTTCACCAAATACCC
GCGTGGTGGTGGTAAGCTTATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGG
ACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTG
AAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTGGGGCGTGCAGTG
CTTCGCCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGC
GCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAAC
CGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACGCCATCAG
CGACAACGTCTATATCACCGCCGACAAGCAGAAGAACGGCATCAAGGCCAACTTCAAGATCCGCCACAACATCGAGG
ACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAAC
CACTACCTGAGCACCCAGTCCAAGCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGT
GACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAA

7.3.2 Method S2 Protein expression and purification

LB (500 mL) supplemented with ampicillin (100 μg mL^{-1}) and kanamycin (50 μg mL^{-1}) was inoculated with an overnight cell culture (10 mL) and incubated (37 °C) with agitation (250 rev min^{-1}). After reaching an optical density at λ = 600 nm of 0.6, isopropyl β-D-1-thiogalactopyranoside (final concentration of 0.5 mM) and 12-ADA (final concentration of 1 mM) were added, and expression was continued overnight under reduced temperature (18 °C). Cells were harvested via centrifugation (4,000 g, 20 min). The cell pellet was resuspended in lysis buffer (40 mL, 20 mM Tris, 50 mM NaCl, 10 mM imidazole, 1 mM phenylmethylsulfonyl fluoride) and sonicated on ice (6 cycles, each 3 min at 30% amplitude, 33% duty cycle, and 3 min resting). Soluble and insoluble fractions were separated via centrifugation (5,000 g, 20 min). Clarified lysate was applied to Ni-NTA resin (5 mL) and incubated under mild agitation (4 °C, 1 hr). The flow-through was discarded, and the resin was washed with wash buffer (20 mM Tris, 50 mM NaCl, 20 mM imidazole, 20 mL, 5x). The protein was eluted with elution buffer (20 mM Tris, 50 mM NaCl, 250 mM imidazole 5 mL, 10x) and collected. The protein solution was dialyzed for 24 hours against Tris buffer (20 mM Tris, 50 mM NaCl, 3x 20-fold excess dialysate to sample) using ThermoFisher SnakeSkin Dialysis Tubing (molecular weight cut-off, MWCO ~ 10 kDa) to remove any residual imidazole and concentrated using an Amicon centrifugal spin column (EMD Millipore, MWCO ~ 10 kDa). Typical yields for purified proteins were ~10-20 mg L^{-1} of cell culture.

7.3.3 Method S3 Mass spectrometric analysis of azide-tagged PhoCl fusions

Protein identity and purity was confirmed by liquid chromatography-tandem mass spectrometry (LC-MS/MS). Each protein species (1 μg in buffer containing 20 mM Tris, 50 mM NaCl, pH 7.4) was injected into a LC-MS/MS (Waters Synapt – G2 HDX QTOF) using an inline POROS R1 reverse-phase column (ThermoFisher). Protein solutions were separated over a 25-minute linear gradient from 10-90% acetonitrile:H2O acetonitrile with 0.1% formic acid at 0.3 mL min^{-1}. Mass spectrum scans were taken every 1 second from 200-3,000 Da in positive mode. The chromatogram was integrated, and the full molecular weight was calculated using MaxEnt1. Expected mass of all fusion protein species account for enzymatic removal of the N-terminal Met by methionine aminopeptidase and NMT-mediated protein myristoylation with 12-ADA. Observed and expected masses of all modified proteins are given in Table S1.

7.3.4 Method S4 Cell culture conditions

HeLa cells were maintained at 37 °C and 5% CO_2 in Dulbecco's Modified Eagle Medium (DMEM) (Life Technologies) supplemented with fetal bovine serum (FBS, 10%) and Penicillin/Streptomycin (P/S, 1%). Media replacements were performed every 2-3 days. Cells were passaged using Trypsin-EDTA (0.05%, Thermo-Fisher), pelleted by centrifugation (125 g, 5 min), and resuspended (2×10^7 cells mL^{-1}) in media (DMEM containing 10% FBS and 1% P/S).

7.3.5 Method S5 Comparing activity of intact and cleaved PhoCl fusions with native protein

Bla

Bla enzymatic activity was measured through a chromogenic assay involving nitrocefin. N_3-PhoCl-bla (1 mg mL^{-1} in phosphate buffered saline, PBS, pH = 7.4) was left in the dark or exposed to light (λ = 400 nm, 10 mW cm^{-2}, 30 min). The intact N_3-PhoCl-bla, photocleaved N_3-PhoCl-bla, or unmodified bla (expressed and purified as previously reported[3]) were individually added (10 ng) to nitrocefin (2 mM in PBS, 100 μL). Sample absorbance (λ_{abs} = 386 nm) was measured over time at 37 °C; the initial slope was calculated as a measure of the enzyme activity.

EGF

EGF activity was determined by its ability to drive HeLa cell proliferation. HeLa cells were plated in 96-well plates (5,000 cells $well^{-1}$) and allowed to attach overnight. The media was then swapped for serum-free media (DMEM containing 1% P/S and 0.1% BSA). N_3-PhoCl-EGF (1 mg mL^{-1} in PBS) was left in the dark or exposed to light (λ = 400 nm, 10 mW cm^{-2}, 30 min). The intact N_3-PhoCl-EGF, photocleaved N_3-PhoCl-EGF, or unmodified EGF (expressed and purified as previously reported[3]) were individually added to each well (1 μg mL^{-1} final EGF concentration). Cells were cultured for 3 days in the presence of each EGF before assaying total DNA content with the PicoGreen assay (ThermoFisher) according to manufacturer's protocol.

7.3.6 Method S6 Formation of PhoCl-modified gels by SPAAC

PEG-tetraBCN (M_n ~ 20,000 Da, 10 mM) was pre-reacted with azide-tagged PhoCl fusion proteins (1 hour, 25 °C) in PBS. The reaction was centrifuged (10,000 rpm, 20 s) to ensure full solublity prior to use. Linear PEG diazide (N_3-PEG-N_3, M_n ~ 3,500 Da, final concentration 8 mM) was added to the PhoCl conjugated PEG-tetraBCN (final concentration 4 mM), and network formation was allowed to proceed for 1 hour between Rain-X®-treated glass slides with silicone rubber spacers (McMaster-Carr, 0.5 mm thick). Following gelation, the slides were separated and gels were equilibrated overnight in PBS prior to use.

7.3.7 Method S7 Determining protein activity following gel photorelease

sfGFP

N_3-PhoCl-sfGFP (15 µM) was pre-reacted with PEG-tetraBCN for 2 hours before gelation. Gels (5 µL, 500 µm thick) were cast between Rain-X®-treated slides and allowed to form at room temperature (1 hr). Following gelation, hydrogels were washed with PBS (1 mL, 4x over 48 hr) to remove any unconjugated protein. Hydrogels were then exposed to violet light (λ = 400 nm, 10 mW cm^{-2}, 30 min) or left in the dark. Gels were transferred to fresh PBS (500 µL). After equilibration (16 hr), the supernatant fluorescence was quantified ($\lambda_{excitation}$ = 488 nm, $\lambda_{emission}$ = 530 nm).

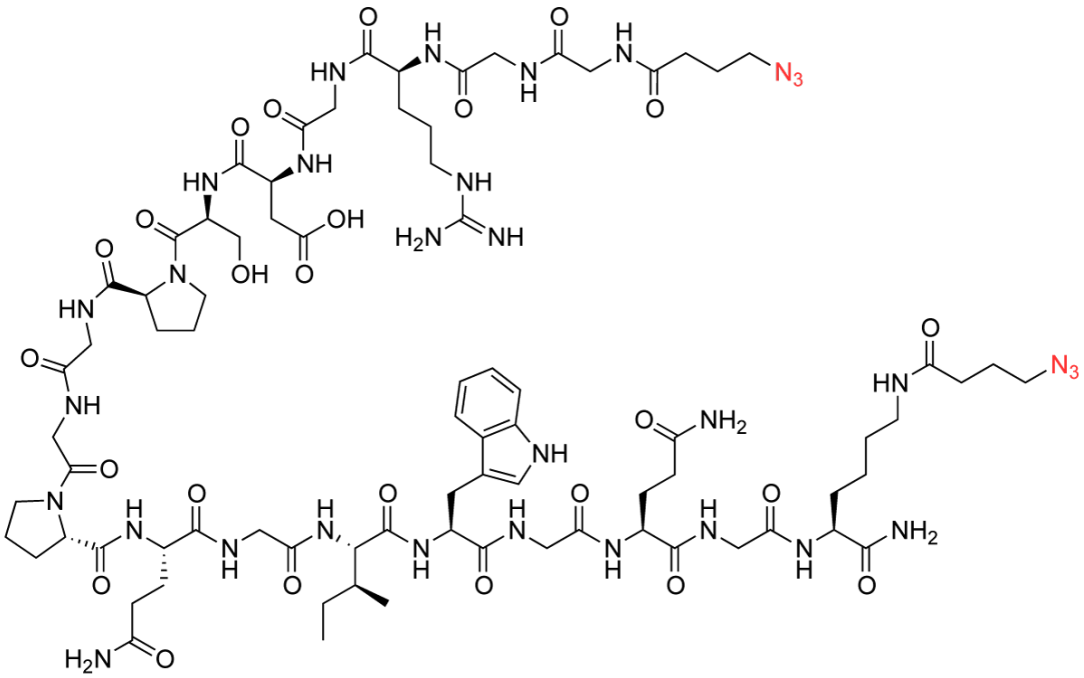
Bla

N_3-PhoCl-bla (15 µM) was pre-reacted with PEG-tetraBCN for 2 hours before gelation. Gels (5 µL, 500 µm thick) were cast between Rain-X®-treated slides and allowed to form at room temperature (1 hr). Following gelation, hydrogels were washed with PBS (1 mL, 4x over 48 hr) to remove any unconjugated protein. Hydrogels were then exposed to violet light (λ = 400 nm, 10 mW cm^{-2}, 30 min) or left in the dark. Gels were transferred to fresh PBS (500 µL). After equilibration (16 hr), supernatants (50 µL) from each gel condition were added to nitrocefin (100 µL in PBS, 0.25 mg mL^{-1}) in a clear 96-well plate. The plate was then immediately put in a plate reader maintained at 37 °C and the sample absorbance (λ_{abs} = 386 nm) was measured over time at 37 °C; the initial slope was calculated as a measure of the enzyme activity.

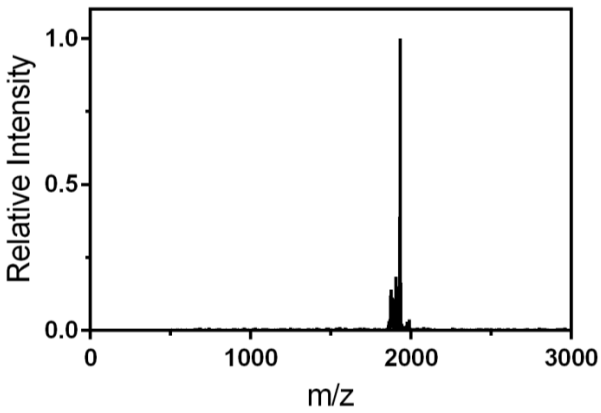
7.3.8 Method S8 Hydrogel photopatterning

N_3-PhoCl-fusion protein containing gels were exposed to collimated violet light (λ = 400 nm, 10 mW cm^{-2}, 30 min) through a patterned chrome photomask (Photo Sciences, Inc.). Released proteins were allowed to diffuse out, and the gels were incubated in PBS overnight. Gels were visualized using a Leica SP8X confocal microscope. For the generation of biochemical gradients, an opaque photomask moved across the surface of the gel during light exposure with assistance of a syringe pump[4] translated at a fixed rate (0.4 or 1.2 mm min^{-1}) under violet light (λ = 400 nm, 10 mW cm^{-2}). Confocal microscope patterning was performed on a Leica SP8x confocal microscope. Patterns were generated using the argon laser line (100 mW 405 nm diode, 5% laser power, 1.14 x 1.14 $microns^2$ $pixel^{-1}$, 200 Hz) for the given number of scans (0 – 64).

7.3.9 Method S9 Synthesis of N_3-GGRGDSPGGPQGIWGQGK(N_3)-NH_2



The resin-bound peptide H-GGRGDSPGGPQGIWGQGK(Dde)-NH_2 was synthesized by Fmoc SPPS (Supplementary Methods) on Rink amide resin (0.25 mmol scale). The resin was washed with DMF (3x) prior to Dde (1-(4,4-dimethyl-2,6-dioxacyclohexylidene)ethyl) deprotection (10 min, 30 mL, 2% hydrazine in DMF, 3x), yielding a peptide with free N-terminal amine and a free ε-amino group on the unprotected lysine. Resin was washed (DMF, 3x) prior to treatment (1 hr) with N_3-COOH (4x, 2 mmol, 258 mg), HATU (3.95x, 1.976 mmol, 752 mg), and DIEA (8x, 4 mmol, 348 μL) in minimal DMF. Resin was washed (DMF, 3x; DCM, 3x) prior to peptide cleavage/deprotection (95:5 TFA:H_2O, 20 mL, 2 hr) and precipitation (diethyl ether, 180 mL, 0 °C, 2x). The crude peptide was purified *via* RP-HPLC using a 55-minute gradient from 5-100% acetonitrile:H_2O; lyophilization yielded the final product [denoted N_3-GGRGDSPGGPQGIWGQGK(N_3)-NH_2] as a light yellow solid (123 mg, 0.064 mmol, 25% overall yield). Peptide purity was confirmed using MALDI-TOF: calculated for $C_{80}H_{123}N_{32}O_{25}^+$ $[M + {}^1H]^+$, 1933.02; observed 1932.53.

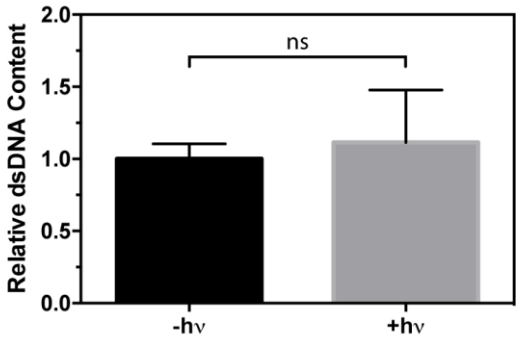

Relative Intensity
1.0
0.5
0.0
0
1000
2000
3000
m/z

7.3.10 Method S10 Controlling anisotropic 3D cell proliferation through patterned EGF

To encapsulate cells within SPAAC gels, PEG-tetraBCN (final concentration 3 mM) was pre-reacted with azide-tagged PhoCl protein (either N_3-PhoCl-EGF or N_3-PhoCl-mRuby, final concentration 4 µM) for 1 hour prior to mixing with HeLa cell suspension (final concentration 10^7 cells mL^{-1}) and N_3-GGRGDSPGGPQGIWGQGK(N_3)-NH_2 (final concentration 6 mM) crosslinker in media. Gels (5 µL) were formed between Rain-X®-treated glass slides separated with silicone rubber spacers (McMaster-Carr, 0.5 mm thick) for 1 hour at 37 °C. Formed gels were transferred to a 24-well plate and maintained in media (500 µL, DMEM containing 10% FBS and 1% P/S) at 37 °C and 5% CO_2. 24 hours post encapsulation, the media was replaced with PBS containing Ca^{2+} and Mg^{2+}, and gels were exposed to masked violet light (λ = 400 nm, 10 mW cm^{-2}, 30 min). The PBS was replaced with media (DMEM containing 10% FBS and 1% P/S) and cultures were maintained (37 °C and 5% CO_2) for 14 days (media replacement every 2 – 3 days).

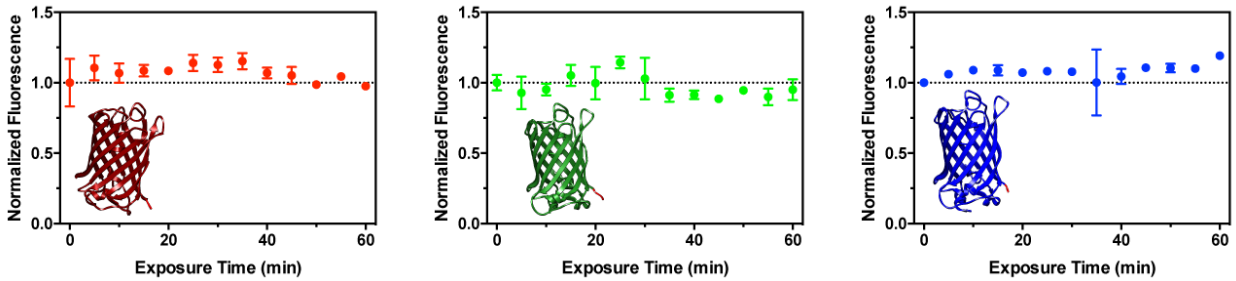
14 days after encapsulation, gels were fixed (1% paraformaldehyde, 30 min, 25 °C) and cell nuclei stained with DAPI (0.15 µg mL^{-1}, 1 hr). After imaging via fluorescence microscopy, images were analyzed using ImageJ to determine cell density and spheroid area in the exposed and unexposed regions. Areas with persistently immobilized EGF (as visualized by PhoCl fluorescence) displayed significantly larger spheroid growth and more cell density than those where EGF had been photoreleased ($p < 0.05$). Control experiments involving mRuby yielded spheroids whose size did not vary throughout the patterned gel and were smaller than those in the released EGF regions.

7.3.11 Method S11 Viability of HeLa cells after violet light exposure



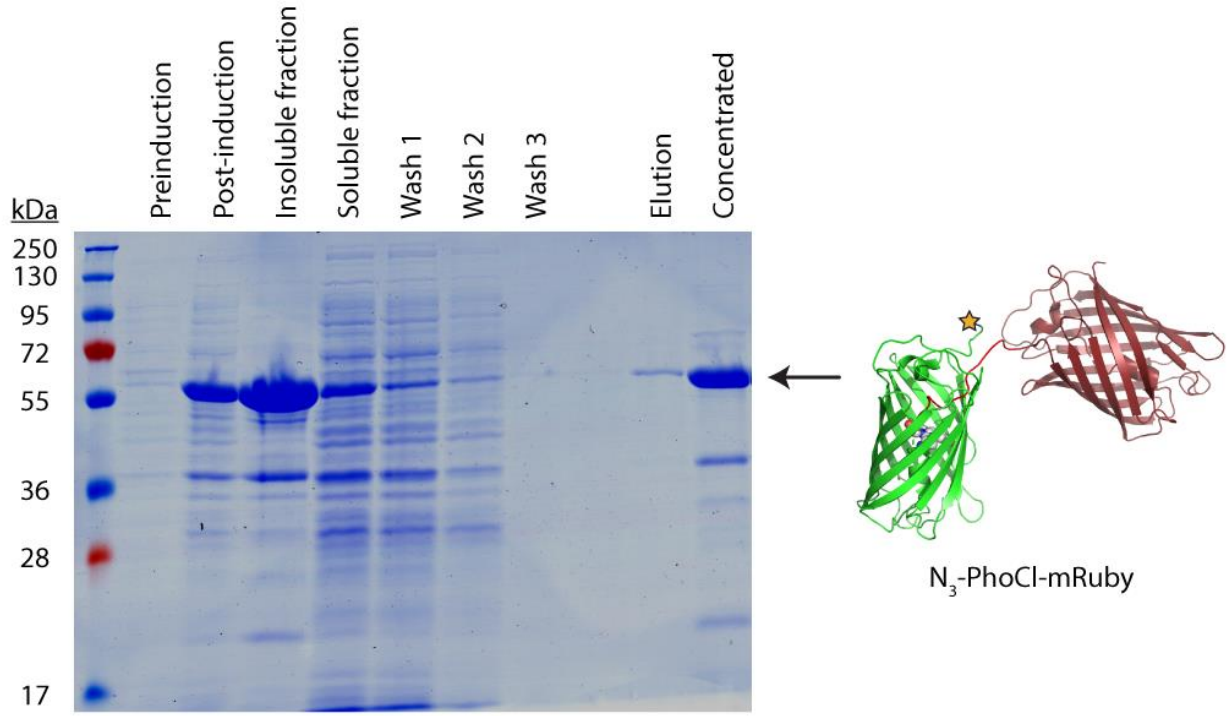
To assess the potential effect of violet light exposure on cell proliferation, HeLa cells were exposed to light (λ = 400 nm, 10 mW cm^{-2}, 30 min) in PBS containing Ca^{2+} and Mg^{2+}. Control gels were maintained in the dark for the same duration. 24 hours following light exposure, dsDNA content (a surrogate measure for cell counts) was quantified via PicoGreen analysis (ThermoFisher) following manufacture's protocol. No significant difference in dsDNA content was observed between light-exposed samples (+hν) and unexposed controls (-hν).

7.3.12 Method S12 Assessing photobleaching of fluorescent proteins in response to violet light



To assess potential loss of fluorescence from violet light exposure, photobleaching studies were performed on mRuby, sfGFP, and mCerulean (expressed and purified as previously reported[3]). Fluorophores (3 μM in PBS) were exposed to violet light (λ = 400 nm, 10 mW cm^{-2} for varying amounts of time (0 – 60 min); fluorescence measurements were taken every 5 minutes and normalized to initial sample fluorescence ($\lambda_{mRuby,excitation}$ = 558 nm, $\lambda_{mRuby,emission}$ = 605 nm; $\lambda_{sfGFP,excitation}$ = 485 nm, $\lambda_{sfGFP,emission}$ = 510 nm; $\lambda_{mCerulean,excitation}$ = 433 nm, $\lambda_{mCerulean,emission}$ = 475 nm). Analysis was performed in triplicate. Left plot corresponds to mRuby; middle to sfGFP; right to mCerulean. Error bars correspond to the standard deviation about the mean. All fluorophores retained full fluorescence after 1 hour of light treatment.

7.4 Supplementary Figures



7.4.1 Figure S1 Assessing N_3-PhoCl-mRuby expression/purification by SDS-PAGE analysis

To investigate the success of protein expression, samples were analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and visualized with Coomasie Brilliant Blue. Here, the N_3-PhoCl-mRuby fusion appears as a ~56 kDa band post induction. Following purification by immobilized metal affinity chromatography, fusion proteins were recovered in high purity (>90%). Molecular weight markers are given in kDa.

7.5 Supplementary Tables

Protein Identity	Expected mass (Da)	Observed mass (Da)
N_3-PhoCl-mRuby	56,300	56,298
N_3-PhoCl-sfGFP	56,677	56,673
N_3-PhoCl-mCerulean	56,685	56,658
N_3-PhoCl-bla	58,941	58,942
N_3-PhoCl-EGF	36,331	35,714

7.5.1 Table S1 Observed masses of azide-modified PhoCl chimera proteins

7.6 Supplementary References

(1) DeForest, C. A.; Tirrell, D. A. *Nat. Mater.* **2015**, *14* (5), 523.

(2) Liu, L.; Shadish, J. A.; Arakawa, C. K.; Shi, K.; Davis, J.; DeForest, C. A. *Adv. Biosyst.* **2018**, 1800240.

(3) Shadish, J. A.; Benuska, G. M.; DeForest, C. A. *Nat. Mater.* **2019**.

(4) Johnson, P. M.; Reynolds, T. B.; Stansbury, J. W.; Bowman, C. N. *Polymer (Guildf).* **2005**, *46* (10), 3300.

Curriculum Vitae

Jared A. Shadish

484-798-4883 | http://www.linkedin.com/in/jared-shadish | shadij@uw.edu

Education

2013-2019 **PhD Chemical Engineering** - University of Washington

2009-2013 **B.S. Chemical Engineering** - Rensselaer Polytechnic Institute

Professional Experience

Graduate Research Assistant

Advisor: Cole DeForest

- Utilized genetic engineering and microbiology techniques in order to generate, express, and purify site-specifically modifiable fusion proteins in *E. coli*.
- Performed peptide synthesis for creating protein-peptide conjugates in order to introduce functionality for photopatterning, PEGylation, tethering, and other novel uses for recombinantly expressed proteins.
- Used HPLC, mass spectrometry (LC/MS, MALDI, TOF-SIMS), microscopy (phase contrast, fluorescent, two-photon), and gel electrophoresis techniques in order to purify and characterize the various compounds and proteins generated during the course of my studies.

Undergraduate Research

Cancer Cell Biology Research Group

- Studied the chemical and mechanical cues involved in endothelial cell microvasculature formation when co-cultured with tumorigenic cell lines.
- Used standard tissue culture techniques, SDS-PAGE, Western Blotting, transwell experiments, and collagen-based gel formations.

Publications

1. **Shadish, J**., Strange, A., & DeForest, C. Genetically Encoded Photocleavable Linkers for Protein Patterning in Biomaterials. *In preparation.* (2019).
2. Gawade, P., Shadish, J., Badeau, B., & DeForest, C. Logic-Based Delivery of Site-Specifically Modified Proteins from Environmentally Responsive Hydrogel Biomaterials. *Adv. Mater.* doi:10.1002/adma.201902462 (2019).

3. **Shadish, J. A.,** Benuska, G. M. & DeForest, C. A. Bioactive site-specifically modified proteins for 4D patterning of gel biomaterials. *Nat. Mater.* doi:10.1038/s41563-019-0367-7 (2019).
4. Badeau, B.A., Comerford, M.P., Arakawa, C.K., **Shadish, J.A**. & DeForest, C.A. Engineered Modular Biomaterial Logic Gates for Environmentally Triggered Therapeutic Delivery. *Nat. Chem.*, 10, 251-258 (2018)
5. Liu, L., **Shadish, J**., Arakawa, C., Shi, K., Davis, J., & DeForest, C. Cyclic Stiffness Modulation of Cell-Laden Protein–Polymer Hydrogels in Response to User-Specified Stimuli Including Light. *Adv. Biosyst.* **2,** 1800240 (2018).
6. Swift, B. J. F., **Shadish, J. A.,** DeForest, C. A. & Baneyx, F. Streamlined Synthesis and Assembly of a Hybrid Sensing Architecture with Solid Binding Proteins and Click Chemistry. *J. Am. Chem. Soc.* 139, 3958–3961 (2017).
7. Farahani, P. E., Adelmund, S. M., **Shadish, J. A**. & DeForest, C. A. Photomediated oxime ligation as a bioorthogonal tool for spatiotemporally-controlled hydrogel formation and modification. *J. Mater. Chem. B* **5,** 4435–4442 (2017).

Patents

DeForest, C., **Shadish, J.**, Liu, L. Dynamic User-Programmable Materials Including Stimuli-Responsive Proteins 16/399,882 (2019)

Book Chapters

Tibbitt, M. W., **Shadish, J. A**. & DeForest, C. A. Photopolymers for Multiphoton Lithography in Biomaterials and Hydrogels. in *Multiphoton Lithography* 183–220 (Wiley-VCH Verlag GmbH & Co. KGaA, 2016). doi:10.1002/9783527682676.ch8

Leadership and Teaching Experience

Distinguished Young Scholars Seminar Series, Director

2016

The Distinguished Young Scholars Seminar Series is a student run and organized seminar that selects students and post docs from across the country to come to the University of Washington for meetings with faculty and students, as well as to give an hour-long seminar to the Department of Chemical Engineering. As the director, I collected applications, led a student panel for applicant selection, helped organize travel arrangements, and helped to ensure that the event ran smoothly when the speakers arrived.

Society for Biomaterials, Vice President

October 2014 – March 2015

Society for Biomaterials highlights biomaterials research done in the Pacific Northwest by bringing together graduate students, professors, and industry professionals in a day-long seminar held at the University of Washington. As Vice President I secured funding for the event,

contacted speakers, and looked for additional outreach opportunities to share the great research that goes on in the area.

Association of Chemical Engineering Graduate Students, Professional Development Chair

January 2014 - January 2015

As the chair for professional development, I worked to foster connections between industry members and graduate students at the University of Washington. I helped provide professional development opportunities in the form of public speaking competitions, sharing information about on campus career building events, and organized an annual Graduate Student Symposium where chemical engineering graduate students were given the opportunity to present their research to fellow graduate students, professors, and members of industry.

Teaching Assistant Roles

Chemical Engineering Laboratory II (2014), Transport Processes I (2015, 2016, 2017), Material and Energy Balances (2018)

Selected Presentations

2019 - Society for Biomaterials – *Genetically Encoded Photocleavable Linkers for Protein Release from Biomaterials (2019)* - **STAR award**

2018 – Gordon Research Seminar - Signal Transduction by Engineered Extracellular Matrices - *Bioactive Site-Specifically Modified Proteins for 4D Patterning of Gel Biomaterials*

2016 - Graduate Student Symposium - *Photopatterned Immobilization and Release of Site-Specifically Modified Proteins (2016)* **– 1st place oral presentation**