Abstract

Engineering protein-based biomaterials using SpyTag/SpyCatcher technology

Danielle Marie Williams

2018

The ability to create linkages using SpyTag/SpyCatcher technology is a powerful tool for protein design and engineering. SpyTag, a 13-residue peptide, and SpyCatcher, a small protein, interact spontaneously to form a covalent isopeptide bond, which has great potential for use in the design of self-assembling materials. Here we exploit the desirable properties of SpyTag/SpyCatcher in addition to other protein building blocks to develop new protein-based materials with biomedical applications.

One such project involves the design of stimulus-responsive hydrogels made entirely of protein components. We took advantage of the modular nature of TPR domains, which bind C-terminal peptides, to interact with cognate peptide cross-linkers to form an ionic hydrogel network. Through the adaptation of SpyTag/SpyCatcher technology, we have developed a concatenation scheme that allows us to create branched protein topologies with multivalent peptides that correspond to TPR domains. By mixing telechelic peptide cross-linkers with their corresponding TPR arrays, we demonstrate the formation of hydrogels that can respond to a physiological stimulus.

Other projects described in this text involve the development of geometric protein arrays. Such arrays have implications in signal amplification, biosensing, and enzyme scaffolding array. Using the modular, rod-like protein SasG and SpyTag/SpyCatcher technology, we were able to create discrete rectangular protein arrays with tunable dimensions. In a different project, SpyCatcher arrays were used to create complexes that can display various numbers of recognition elements on the same construct. As a proof of principle, we designed complexes that can bind to HER2, a relevant breast cancer biomarker.

Finally, we were able to functionalize 2D surfaces and microcapsules using SpyTag/SpyCatcher technology and a unique, amphiphilic bacterial hydrophobin, BslA. BslA self-assembles into a stable monolayer at both air-water and water-oil interfaces. We have functionalized these proteins with SpyTag and SnoopTag for use in decorating surfaces and microcapsules with proteins of interest attached to SpyCatcher.

Engineering protein-based biomaterials using

SpyTag/SpyCatcher technology

A Dissertation

Presented to the Faculty of the Graduate School

of

Yale University

In Candidacy for the Degree of

Doctor of Philosophy

By

Danielle Marie Williams

Dissertation Director: Lynne Regan

May 2018

Table of Contents

List of Figures and Tables

Glossary of Terms

AFM	Atomic Force Microscopy
BME	β-mercaptoethanol
BSA	Bovine Serum Albumin
BslA	Biofilm Surface Layer Protein A
CD	Circular Dichroism
CPEC	Circular Polymerase Extension Cloning
DI	Deionized
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic Acid
eGFP	Enhanced Green Fluorescent Protein
ELISA	Enzyme-Linked Immunosorbent Assay
ELP	Elastin-Like Polymer
EM	Electron Microscopy
FOTS	Trichloro(1H,1H,2H,2H-perfluorooctyl)silane
GST	Glutathione-S-Transferase
GFP	Green Fluorescent Protein
IPTG	Isopropyl β-D-1-Thiogalactopyranoside
LB	Lysogeny Broth
LB	Langmuir Blodgett
LS	Langmuir-Schaefer
Ni-NTA	Nickel-Nitrilotriacetic Acid
NTA	Nitrilotriacetic Acid

PBS	Phosphate Buffered Saline
PDMS	Polydimethylsiloxane
PEG	Polyethylene Glycol
PMSF	Phenylmethane Sulfonyl Fluoride
PVA	Polyvinyl Alcohol
SasG	*Staphylococcus aureus* Protein G
SDS PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
SFG	Sum Frequency Generation
TEV	Tobacco Etch Virus
TPR	Tetratricopeptide Repeat
UV	Ultraviolet
WT	Wild Type

Acknowledgements

Chapter 1: Introduction

Some text, figures, and captions in Chapter 1 are adapted from those published in:

Schloss, A.C., Williams, D.M. Regan, L. (2016) Protein-based hydrogels for tissue engineering *Adv. Exp. Med. Biol.* **940,** 167-176

Regan, L., Caballero, D., Hinrichsen, M.R., Virrueta, A., Williams, D.M., O'Hern, C.S. (2015) Protein design: Past, present, and future. *Biopolymers*. **104(4)**, 334-350

Protein engineering is a fascinating field that has interested me since I was an undergraduate in Tom Magliery's lab at Ohio State University, and I was fortunate to be able to continue these studies throughout my dissertation work under the guidance of Lynne Regan. The ability to genetically encode new functions into proteins and harness them for biomedical applications, from therapeutics to diagnostic devices, is something that is very exciting to me as I look ahead towards my career in healthcare. My time in graduate school was spent designing and creating new protein-based systems that no one else had constructed before, and while that was challenging, I am happy to have contributed to a field with such great potential for widespread impact. As science is truly a collaborative effort, I have chosen to use "we" instead of "I" throughout this document, and the work of collaborators is specifically stated at the beginning of each chapter.

1.1 Proteins as Candidates for Biomaterials

The unique features associated with protein-based materials make them attractive candidates for biomaterials. Protein-based materials confer a number of advantages over synthetic materials for biomedical applications: (1) the features required for 3-dimensional percolation and gelation are precisely encoded by the sequence, which specifies the structure; (2) genetic engineering to create virtually any desired sequence is

relatively straightforward; (3) exquisite stimuli-responsiveness can readily be controlled by appropriately engineering the interactions between protein building blocks.

1.2 SpyTag/SpyCatcher Technology

Weak protein-protein interactions are often a limitation when using protein-based materials for biotechnology. Strong linkages between protein components that can withstand a wide range of conditions would greatly increase the potential of proteins for biomedical applications (Veggiani, Zakeri, & Howarth, 2014). Through engineering bacterial adhesion proteins in *S. pyogenes*, researchers in the Howarth Lab created a protein, SpyCatcher, that rapidly and spontaneously forms an isopeptide bond with its cognate peptide, SpyTag (Figure 1.1). Components were engineered from a fibronectin-binding domain that spontaneously forms an intramolecular covalent bond in nature, and researchers were able to divide the protein into two separate parts while maintaining this activity (Zakeri et al., 2012).

Building upon this work, Arnold, Tirrell and colleagues exploited this technology to allow for easy expression of branched protein building blocks (Zhang, Sun, Tirrell, & Arnold, 2013). Typically, protein expression is typically limited to linear topologies, but the development of SpyCatcher/SpyTag technology changed that (Figure 1.2). By genetically encoding SpyTag and SpyCatcher in constructs of interest, diverse topologies can be created.
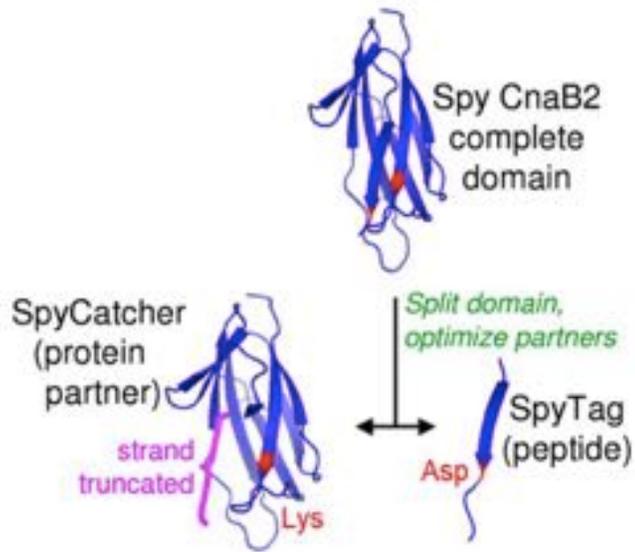
Figure 1.1: A ribbon cartoon showing design of SpyTag/SpyCatcher technology. *S. pyogenes* CnaB2 was split into a large N-terminal fragment (SpyCatcher, left) and a small C-terminal fragment (SpyTag, right). Reactive residues forming the isopeptide bond are colored red.
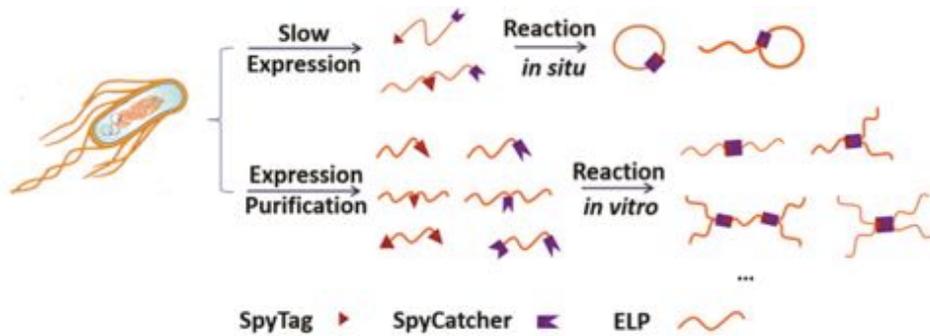
Figure 1.2: A cartoon showing the various protein topologies possible through incorporation of SpyTag/SpyCatcher genes. Depending on expression conditions and placement of Spy sequences, a number of branched and cyclic protein topologies can be created.

Since its development, many have exploited this technology for the construction of new protein-based materials (Howarth, 2017). This technology has been adapted to form hydrogel networks (Gao, Lyu, & Li, 2017; Sun, Zhang, Mahdavi, Arnold, & Tirrell, 2014), has aided in vaccine development (Brune et al., 2017; Thrane et al., 2016), and has been used to construct antibodies (Alam, 2017; Yumura et al., 2017). The connection formed by this interaction is irreversible, robust, and the genes can be encoded virtually anywhere in the construct of interest. SpyCatcher and SpyTag technology is a powerful tool to connect proteins, and it is an integral part of each project design described in this document.

1.3 Protein-Based Hydrogels

Protein-based hydrogel components are appealing for their structural designability, specific biological functionality, and stimuli-responsiveness. The tunable mechanical and structural properties of protein-based hydrogels make them excellent scaffolds for tissue engineering and repair. Moreover, using protein-based components provides the option to insert sequences associated with the promoting both cellular adhesion to the substrate and overall cell growth.

Hydrogel biocompatibility and low immunogenicity are essential for *in vivo* applications. Researchers must either choose components with known biocompatibility or engineer components to be more compatible. For protein-based hydrogels, site-specific mutagenesis and/or truncation can be used by researchers to remove immunogenic epitopes while maintaining gelation properties (Z. Liu et al., 2014). Some proteins, such as elastin-like polymers (ELPs), naturally exhibit low immunogenicity (J. Carlos

Rodriguez-Cabello, 2009) and can confer this biocompatibility onto tissue engineering scaffolds. To function properly and to be compatible with tissue growth, hydrogels for tissue engineering must meet a number of requirements, both physical and biological. The mechanical and structural properties of the gel - tensile strength, stiffness, elasticity, and so on - must be matched to the specific tissue type (Whang et al., 1999). The elastic modulus (the ratio of tensile strength to tensile stress) has been shown to directly influence to cell growth and tissue development (Georges, Miller, Meaney, Sawyer, & Janmey, 2006).

Temperature is an attractive stimulus because it is straightforward to apply. The key is to engineer stimuli-responsiveness in a regime that is compatible with physiological temperature. Woolfson and colleagues designed hydrogels using α-helical peptides with thermo-sensitive properties encoded by the types of interactions between entangled helical fibrils (Banwell et al., 2009). The propensity for the hydrogel to become stronger or weaker after heat application is dependent on whether the fibril network is formed through hydrophobic (increase in strength with increasing temperature) or hydrogen bonding (decrease in strength with increasing temperature) interactions.

Olsen, Tirrell and colleagues designed hydrogels whose formation is based on the association of α-helices into coiled-coils (Olsen, Kornfield, & Tirrell, 2010). Subsequently, Olsen and colleagues elaborated on these designs to create a thermo-sensitive hydrogel that is liquid at low temperatures (4°C) but which exhibits enhanced stiffness and durability at physiological temperature (37°) (Glassman, Chan, & Olsen, 2013). There are two key components to this design: the coiled-coil based shear thinning hydrogel midblock, and endblocks comprised of the thermo-responsive polymer poly(N-

isopropylacrylamide) (PNIPAM) (Figure 1.3). Such a shear thinning hydrogel that undergoes a transition to a more rigid, reinforced network at physiological temperatures could be used for injection, for example, in tissue repair. Hydrogels responsive to other physiological stimuli, such as ionic strength and pH, are explored in Chapter 2.
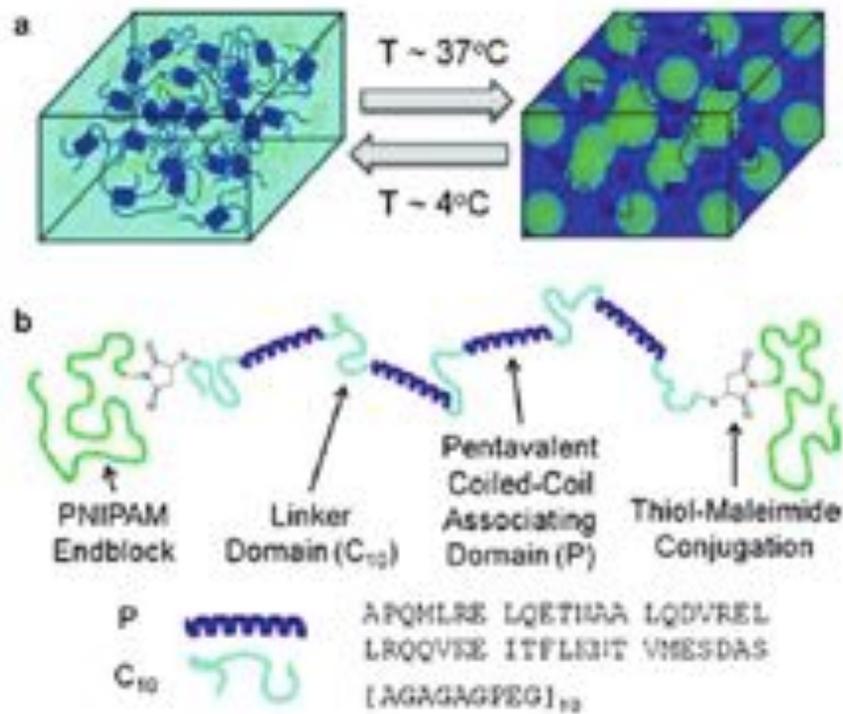
Figure 1.3 A schematic illustrating the design of a thermo-sensitive hydrogel. a) An illustration of the shear-thinning hydrogel (left) which is reinforced to become a stiffer, more rigid network at higher temperatures. b) A cartoon of the design of the dual-system hydrogel components.

1.4 Geometric Protein Arrays

DNA origami has enabled the design of an impressive diversity of 2- and 3-dimensional structures (Sacca & Niemeyer, 2012). Incorporating function, however, has proven to be more difficult. By contrast, designing structures for 'protein origami' is more involved, but functionalization is relatively straightforward. Geometric protein arrays have the potential for widespread applications in enzyme scaffolding, cargo delivery, and biosensing.

An excellent example of protein origami is demonstrated by Jerala and colleagues, who took advantage of the specificity of association between coiled-coil building blocks to form a single-chain polypeptide structure that folds into a polyhedron (Gradisar et al., 2013). The design employed six different pairs of coiled-coils (Figure 1.4). A linker sequence was chosen that included residues that would enhance flexibility and disrupt helix formation - Ser-Gly-Pro-Gly. Another crucial component of the design is the orthogonality of the pairs - unintentional crossreactivity between different coiled-coil monomers would prevent the proper assembly of the tetrahedron. The resulting 3-dimensional structure was imaged by atomic force microscopy (AFM), and the proximity of the N- and C-termini at the same vertex was confirmed by a split-fluorescent protein assay. Protein origami is attractive because such structures can be easily functionalized for use in pathway engineering, difference imaging, and novel vaccines.
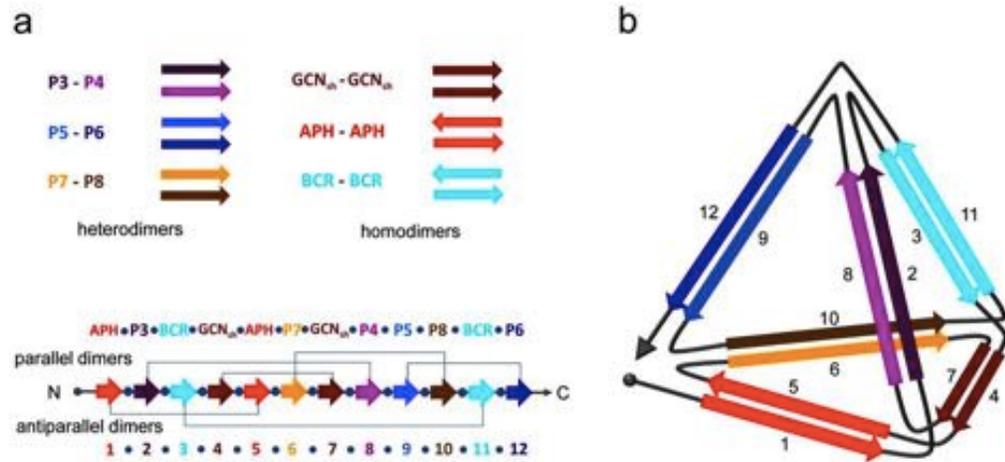
Figure 1.4 The design of a tetrahedron/trigonal pyramid using coiled-coils assembly. a) Cartoon illustrating the pyramid components – sets of heterodimeric and homodimeric parallel and anti-parallel coiled-coils. The 12 individual peptide sequences are concatenated in the indicated order, with each sequence separated by the flexible linker Ser-Gly-Pro-Gly. Gray lines indicated the interacting pairs. b) Schematic of the desired tetrahedron structure. Arrows indicate the direction of the helices in the coiled-coil pairs.

Woolfson and colleagues created self-assembled cage-like particles (SAGEs) that form spheres of roughly 100 nm in diameter (Fletcher et al., 2013). Non-covalent heterodimeric and heterotrimeric coiled-coils were employed as building blocks for the design, where different coils were connected by asymmetric disulfide bonds to form hubs that assemble into a hexagonal array upon mixing). Interestingly, instead of forming the expected flat assembly based on the hexagonal design, the structures assembled into closed spheres.

A large, 24-subunit protein cube with structural validation was designed by Yeates and colleagues (Lai et al., 2014). Their design strategy involved making fusions between natural dimeric and trimeric proteins. The particular proteins used were chosen so that the angle of the interface would satisfy the requirements for cube formation when propagated. When components were mixed, they self-assembled into a porous cube with an outer diameter of 225 Å and an inner diameter of 132 Å, as determined by x-ray crystallography. The structure was additionally validated by negative stain electron microscopy and small angle x-ray scattering (SAXS) analysis. Like the SAGE particles above, the large cavities in these protein assemblies have potential applications in delivery of molecular cargo. More geometric protein designs are discussed in Chapter 3.

1.5 Protein-Based Microcapsules

Microcapsules with the ability to compartmentalize molecular cargo are useful for a wide range of applications, including but not limited to biotechnology (Wong, Al-Salami, & Dass, 2018), food science (Sipailiene & Petraityte, 2018), and cosmetics (Casanova & Santos, 2016). The unique properties of proteins allow for the possibility of

incorporating new functions into and fine-tuning microcapsules for specific applications. In Chapter 4, we discuss the fabrication of oil-in-water microcapsules coated with a surface active, amphiphilic protein, BslA. These protein-coated capsules do not coalesce, are stable at room temperature, and have the potential to encapsulate water-insoluble cargo, which has implications in delivery of cancer drugs (Z. Liu, Robinson, Sun, & Dai, 2008). The diameter of the microcapsules can be fabricated in different sizes and functionalized with different reactive protein coatings (Schloss, 2016).

1.6 Protein Immobilization

The majority of biosensors require immobilization of the recognition element, which is often a protein (Ferrigno, 2016; Mohamad, Marzuki, Buang, Huyop, & Wahab, 2015; Sharma, Byrne, & O'Kennedy, 2016). Achieving consistent presentation of the protein recognition element on a surface while simultaneously avoiding undesirable, and potentially denaturing, interactions of that protein with the surface is currently an unsolved problem (Smith, Sapsford, Tan, & Ligler, 2011). Accomplishing consistent presentation of a native protein on a surface would increase the accessibility of that protein to the analyte and maximize the number of native proteins in a given area, both of which would increase the sensitivity of analyte detection. Several different approaches have been taken to address this important issue, but a straightforward and widely applicable strategy has not yet been established (Smith et al., 2011; Tischer & Kasche, 1999) (Boller, 2002; Bryjak, 1998; Janssen, van Langen, Pereira, van Rantwijk, & Sheldon, 2002; Kahn, 2010). The work presented in Chapter 5 aims to tackle some of these problems for facile protein immobilization to surfaces.

Chapter 2: TPR Hydrogels

Some text, figures, and captions in Chapter 2 are adapted from those published in:

This work was done in collaboration with graduate student Ashley Schloss. Specifically, Ashley cloned the SpyCatcher arrays, the MMY TPR array, and SpyTag-ELP-MEEVF, and optimized the co-expression conditions for telechelic peptide cross-linkers. She aided in the design and/or execution of all other experiments described.

2.1 Introduction

In addition to controlling the self-assembly process, the development of "smart" hydrogels that are able to respond to a physiological stimulus can be useful in many applications, such as targeted delivery of cargo. Regan and colleagues demonstrated the formation of self-assembling hydrogels by taking advantage of an ionic interaction between a protein and peptide (Grove, Osuji, Forster, Dufresne, & Regan, 2010). Tetratricopeptide repeat (TPR) domains were concatenated in arrays that formed non-covalent cross-links with their corresponding peptides attached to four-arm star polyethylene glycol (PEG) molecules (Figure 2.1). The binding interactions forming the junctions of the network were both pH- and ionic strength-dependent, allowing for reversible gelation in response to external stimuli. When placed in solutions containing 500mM NaCl, the hydrogel dissolves, but gelation can be reconstituted through the removal of salt through dialysis (Grove, Forster, Pimienta, Dufresne, & Regan, 2012). This reversibility can also be demonstrated by alteration of pH (Grove et al., 2010). Hydrogelation occurred when components were mixed in a 1:2 stoichiometric ratio (TPR

arrays: PEG-peptides). The elastic modulus was measured to be ~270 Pa, which is well above the minimum value necessary (50 Pa -100 Pa) to support mammalian cells in suspension, demonstrating potential in applications for tissue engineering and regeneration (Grove et al., 2010).
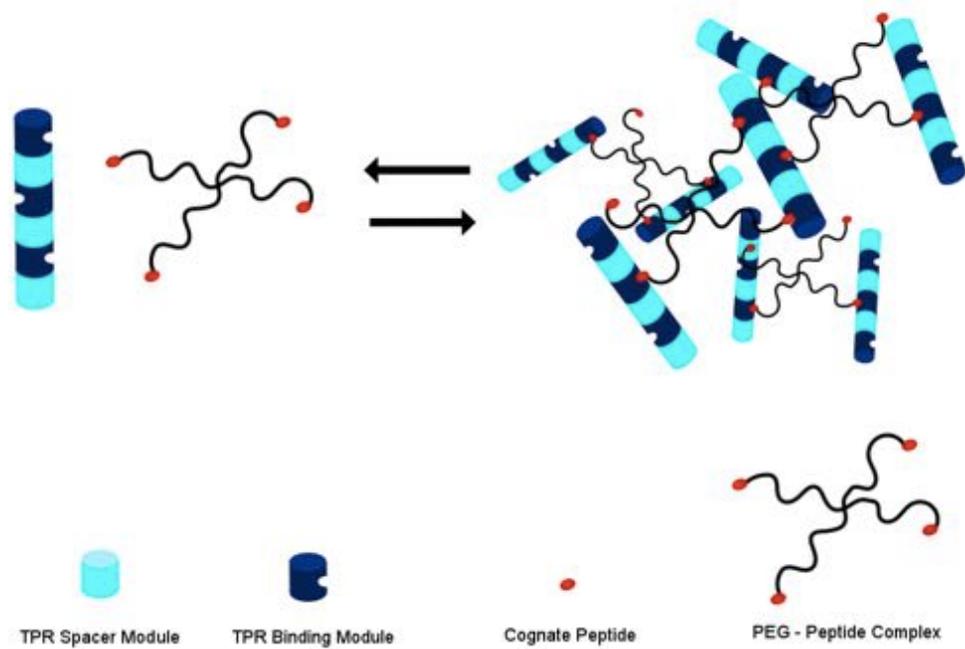
Figure 2.1 A cartoon showing the reversible formation of a TPR hydrogel network. TPR arrays are shown as blue rods with alternating binding modules (navy) and spacer modules (cyan). Corresponding peptides (red) are displayed on the ends of 4-armed PEG star molecules.

2.2 Design of New TPR Arrays

TPRs are small, 34-residue helix-turn-helix motifs that often occur in functional units of three tandem repeats (Figure 2.2a). TPR domains are highly modular and have a diverse range of binding specificities (D'Andrea & Regan, 2003). TPR domains bind peptide sequences and undergo very little conformational change upon binding, making them very modular and easy to modify the specificity and affinity of the interaction. Consensus TPRs (CTPRs) are based on idealized consensus TPR sequences and they form superhelical structures upon concatenation (Figure 2.2b) (Cortajarena, Wang, & Regan, 2010).

The modular nature of TPR domains allows for the engineering of interactions with different binding affinities/specificities that encode unique macroscopic properties. We designed new arrays using different TPR domains, as well as different concatenation schemes. TPR domains used in this study include TPR2A (a natural TPR domain of HOP), CTPR390 (Cortajarena et al., 2010), and MMY (a TPR with a hydrophobic binding pocket in the scaffold of TPR2A). MMY is a construct created previously in the lab (Jackrel, Valverde, & Regan, 2009) and binds to the hydrophobic C-terminal peptide MEEVF, while TPR2A and CTPR390 both bind to the C-terminal peptide MEEVD. Crystal structures showing the different TPR domains in complex with their cognate peptides are shown (Figure 2.2c).
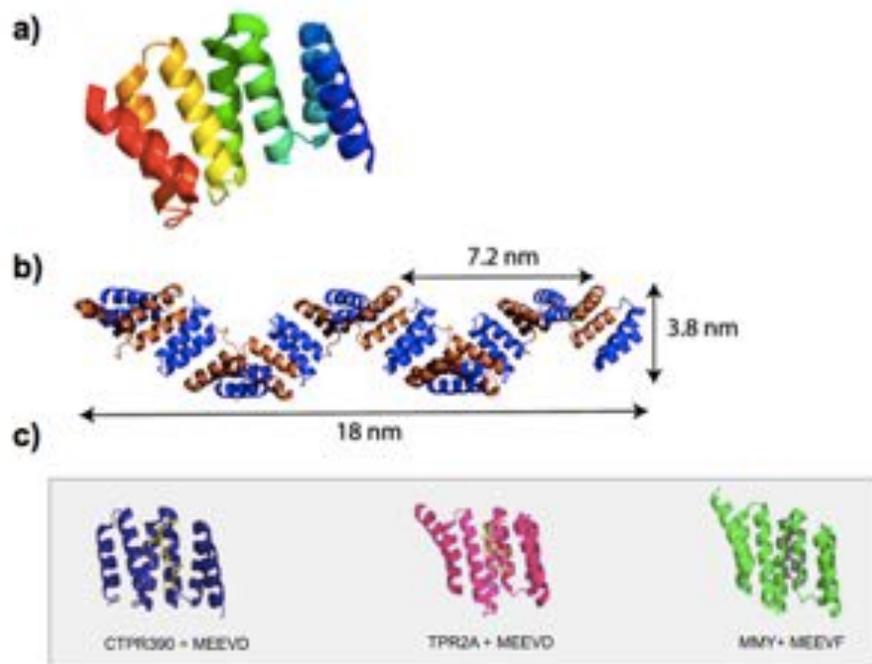
Figure 2.2 Crystal and ribbon structures of TPR domains. a) A crystal structure of the TPR1 domain from HOP (PDB: 1ELW) showing the tandem helix-turn-helix motifs characteristic of TPR domains. b) A ribbon cartoon depicting a superhelical array comprised of six CTPR domains, including the dimensions. c) Crystal structures of CTPR390 in complex with the peptide MEEVD (PDB: 3KD7, left), TPR2A in complex with the peptide MEEVD (PDB: 1ELR, middle), and MMY in complex with the peptide MEEVF (PDB: 3FWV, right).

In addition to changing the sequence identities of the TPR domains, we wanted to change the ways in which they were connected to form the arrays. Previous designs from Grove *et al.* used alternating peptide-binding and "spacer" TPR domains. In order to increase the amount of peptide-binding sites, we designed arrays that contained six consecutive binding domains with no spacers (Figure 2.3). Another design incorporated three peptide-binding domains, but instead of being connected by spacer TPRs, they were attached by a Gly_3 linker (Figure 2.3). The rationale behind this design was that the glycine linkers would increase the flexibility of the array, allowing for increased flexibility regarding the orientation of the peptide-binding domains.

TPR arrays were cloned, expressed, and purified using affinity chromatography. The MMY arrays with 6 binding domains and TPR2A arrays with six binding domains did not express well, and we therefore did not use these constructs in further experiments. Purity of the TPR arrays was verified by SDS PAGE (Figure 2.4).
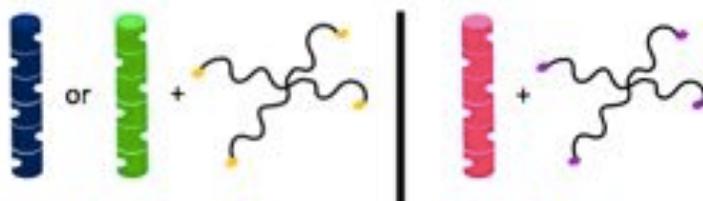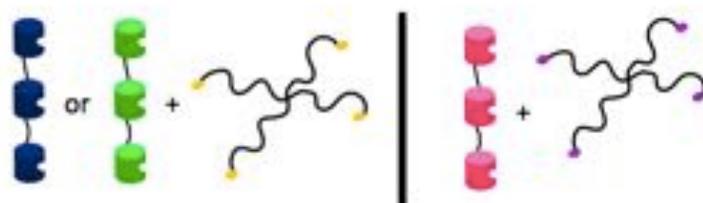
Figure 2.3 A schematic of new TPR array designs. CTPR390 domains are shown in navy, TPR2A domains are shown in green, and MMY domains are shown in pink. Multivalent cross-linkers displaying MEEVD (yellow) and MEEVF (purple) peptides at the ends of PEG stars are shown in black. The top row illustrates the first design of TPR arrays with concatenated binding domains and no spacers. The bottom rom illustrates the second design of TPR arrays consisting of three binding domains connected by glycine linkers.
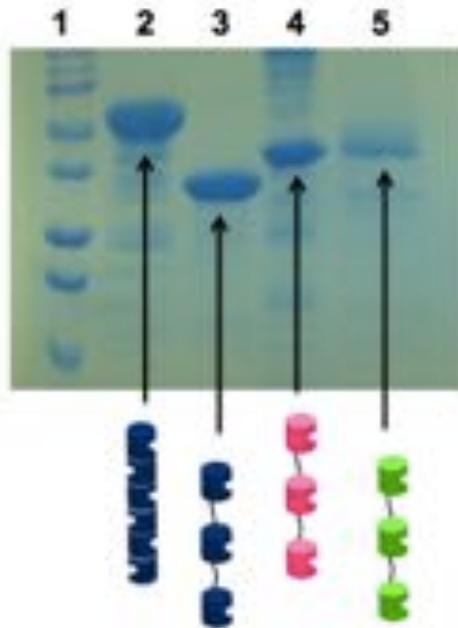
Figure 2.4 An SDS PAGE gel of purified TPR arrays. CTPR390 with six binding domains corresponds to a molecular weight of 78 kDa, CTPR390 with Gly_3 linkers corresponds to a molecular weight of 51 kDa, and MMY and TPR2A with Gly_3 linkers corresponds to a molecular weight of 41 kDa. Lane 1: Ladder; Lane 2: CTPR390 with six binding domains; Lane 3: CTPR390 with Gly_3 linkers; Lane 4: MMY with Gly_3 linkers; Lane 5: TPR2A with Gly_3 linkers.

2.3 Design of Telechelic Peptide Cross-linkers

Original designs used functionalized star PEGs as the cognate peptide "cross-linkers" for the TPR arrays. Through the adaptation of SpyTag/SpyCatcher technology, we developed a novel concatenation and branching scheme that eliminates the need for functionalized PEG. We can now create branched protein topologies with multivalent peptides that can be readily expressed in *E. coli* with the possibility of all proteins being monodisperse, and complicated chemistry previously needed to attach the peptides can be completely avoided.

To create branched proteins displaying peptides, we decided to co-express SpyTag and SpyCatcher on two separate constructs (Zhang et al., 2013). We created a construct with an N-terminal SpyTag connected to a C-terminal peptide, either MEEVD or MEEVF, by an ELP linker. These constructs did not contain a hexahistidine tag. In a different vector containing a hexahistidine tag, we cloned constructs concatenating either two, three, or four SpyCatcher domains. When co-expressed, the complexes assemble *in vivo* and can be co-purified with affinity chromatography (Figure 2.5). To optimize expression conditions, we tested different media types in addition to different temperatures and compared conditions using SDS PAGE analysis. Higher temperatures of 37 °C yielded the best results and expression was unaffected by media type (Figure 2.6).

Moving forwards, we purified SpyCatcher/SpyTag complexes displaying either two, three, or four peptides using affinity chromatography. Banding is seen on SDS PAGE gels, indicating incomplete conjugation of SpyTag-ELP-peptides to the SpyCatcher arrays (Figure 2.7).
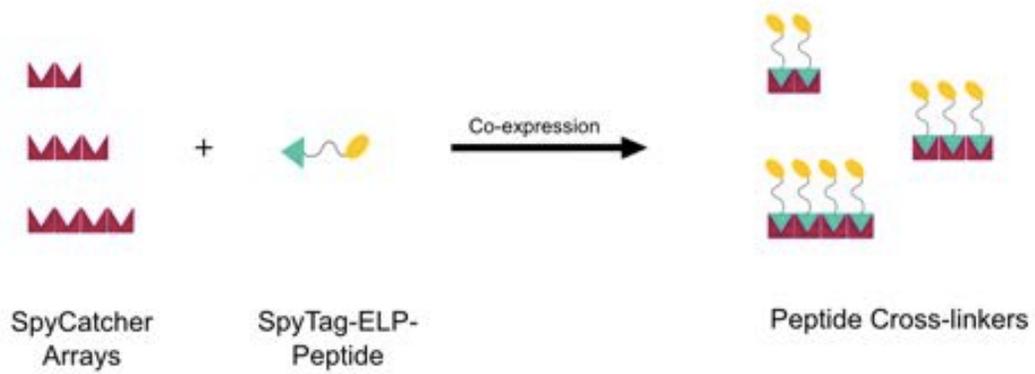
Figure 2.5 A schematic showing the co-expression of multivalent peptide cross-linkers. Arrays with two, three, or four concatenated SpyCatcher domains were co-expressed with SpyTag-ELP-peptides to yield protein complexes displaying multivalent peptides.
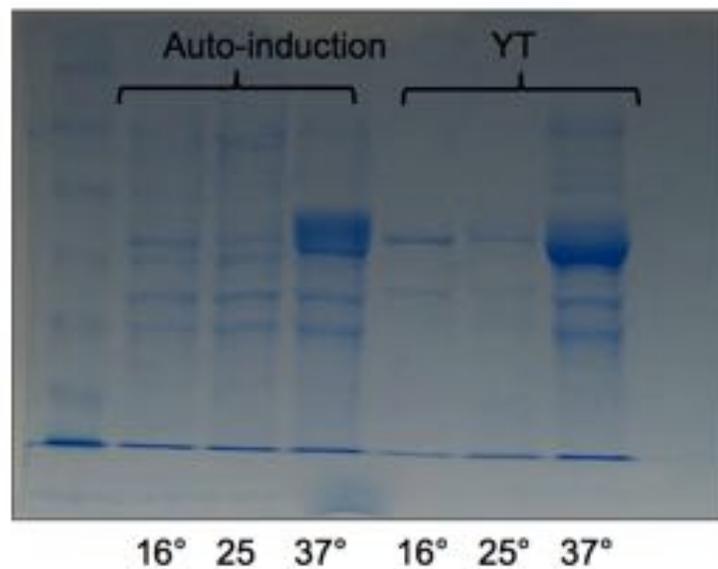
Figure 2.6 An SDS PAGE gel comparing co-expression conditions for multivalent peptide cross-linkers. Cultures expressed in auto-induction media are shown at different temperatures (left) and compared to cultures expressed in YT media at different temperatures (right). Cultures expressed at 37 °C show complete banding, indicating successful conjugation of Spy-Tag-ELP peptides to the SpyCatcher arrays.
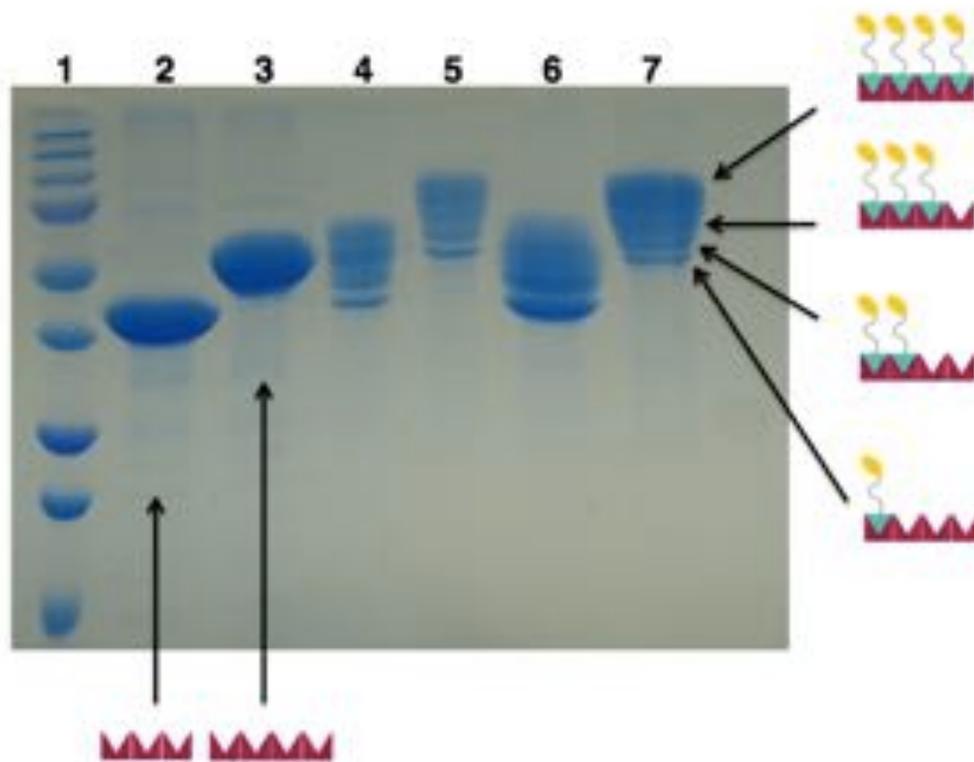
Figure 2.7 An SDS PAGE gel showing purified multivalent peptide cross-linkers. The banding observed in Lanes 4-7 is indicative of the varying efficiency of SpyTag-ELP-peptide conjugations. Cartoons on the right represent the hypothesized conjugation efficiency of each band. Lane 1: Ladder; Lane 2: SpyCatcher array with three domains expressed alone; Lane 3: SpyCatcher array with four domains expressed alone; Lane 4: SpyCatcher array with three domains co-expressed with SpyTag-ELP-MEEVD; Lane 5: SpyCatcher array with four domains co-expressed with SpyTag-ELP-MEEVD; Lane 6: SpyCatcher with three domains co-expressed with SpyTag-ELP-MEEVF; Lane 7: SpyCatcher with four domains co-expressed with SpyTag-ELP-MEEVF.

2.4 Improvements to Peptide Cross-linkers

After initial unsuccessful gelation experiments with the described constructs, we made modifications to the design. To introduce more flexibility to the SpyCatcher arrays, we incorporated a flexible (GGSGGS) linker in between each binding domain for arrays of two, three, and four. We refer to these constructs as SpyCatcher-GS. Successful co-expression of arrays with two, three, and four SpyCatcher-GS domains with SpyTag-ELP-MEEVD was observed (Figure 2.8). The improved design using peptide cross-linkers with more flexible SpyCatcher arrays is illustrated as a cartoon (Figure 2.9).
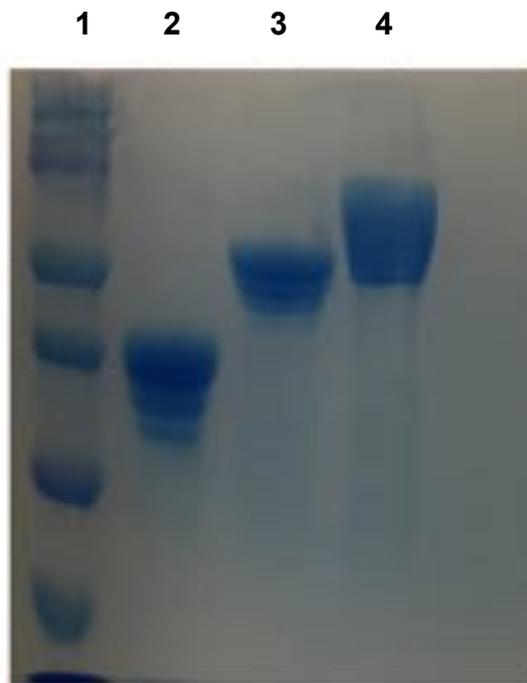
Figure 2.8 An SDS PAGE gel showing purified multivalent peptide cross-linkers made with SpyCatcher-GS arrays. Banding is observed, similar to co-expressions with previous SpyCatcher arrays. Lane 1: Ladder; Lane 2: SpyCatcher-GS array with two domains co-expressed with SpyTag-ELP-MEEVD; Lane 3: SpyCatcher-GS array with three domains co-expressed with SpyTag-ELP-MEEVD; Lane 4: SpyCatcher-GS array with four domains co-expressed with SpyTag-ELP-MEEVD.

Figure 2.9 A schematic of the formation of a TPR2A hydrogel network with improved cross-linker designs. TPR2A domains connected by Gly_3 linkers interact with multivalent, flexible peptide cross-linkers to form an ionic mesh network.

2.5 Hydrogel Formation

Hydrogel formation was successful using TPR2A arrays connected by Gly$_3$ linkers with SpyTag-ELP-MEEVD cross-linkers using the flexible SpyCatcher-GS arrays. TPR2A arrays were mixed with cross-linkers displaying either two, three, or four peptides with in a 1:1 ratio of binding sites at 10 wt/v %, and gelation was observed with all three mixtures overnight at room temperature. The gels did not occupy the entire volume of the solution (Figure 2.10a) and the hydrogels were self-supporting (Figure 2.10b).
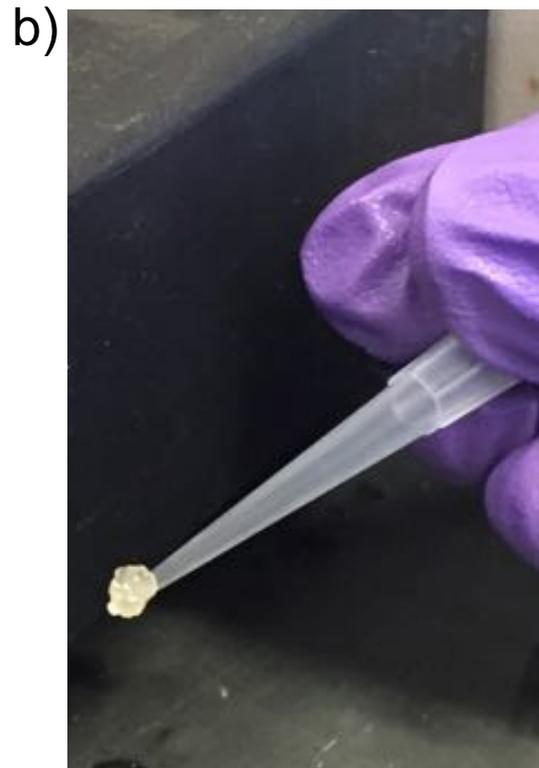
Figure 2.10 Photos of hydrogel formation. a) Hydrogels do not take up the entire volume of the mixture, but instead form at the bottom of the tube. b) Hydrogels are self-supporting.

2.6 Discussion and Conclusion

Here we have demonstrated the formation of self-supporting, hydrogels formed by ionic interactions between TPR domains and their cognate peptides. Our design allowed for gelation to occur overnight at room temperature, which is a much faster time frame than previously reported for TPR hydrogels (Grove et al., 2012). Notably, we used SpyTag/SpyCatcher technology to create multivalent peptide cross-linkers that are entirely protein-based and can be readily expressed from *E coli*. This design could be expanded to incorporate more than four peptides by increasing the amount of concatenated SpyCatcher domains. Rheological experiments are necessary to characterize the viscoelastic properties of the hydrogels, and light scattering experiments would perhaps provide insight to the dynamic interactions between the TPR domains and peptides.

Chapter 3: Creating Geometric Protein Arrays

This work was done in collaboration with graduate student Ashley Schloss. Specifically, Ashley cloned the biotinylated SpyCatcher arrays and SpyTag-GEG-SpyTag. She also optimized the co-expression conditions for HER2 sensing complexes. She aided in the design and/or execution of all other experiments described.

3.1 Introduction

Geometric protein arrays have implications in enzyme scaffolding, biosensing, and imaging. Proteins have more functionality than DNA and it is much easier to remove immunogenicity using mutagenesis. Here we present different ways of using SpyCatcher arrays to create geometric protein arrays that can be used for scaffolding and imaging, and also arrays with potential implications in biosensing.

3.2 Construction of SasG Protein Arrays

When designing a complex for use in enzyme scaffolding, the ability to know and adjust the exact distance between points of functionalization is an attractive feature. Purified from a biofilm forming bacteria, the surface protein SasG is a monomeric, modular protein that is an ideal component for this application. Potts and colleagues have biophysically characterized the beta sheet SasG proteins and determined that the protein is thin and rod-like with two major domains, G5 and E (Figure 3.1). Further investigation into the G5 and E domains of this repeat protein, they found that this protein maintains its linearity with as many as 19 repeats of G5-E domains, making it perfect for the design of linear arrays with varying lengths (Gruszka et al., 2012).
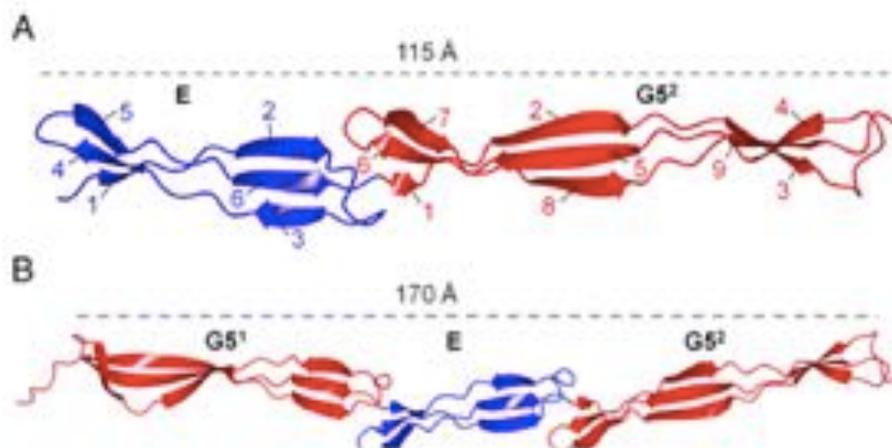
Figure 3.1 A crystal structure of SasG domains. E domains are shown in blue and G5 domains are shown in red and lengths of the segments are labeled. a) The structure of the E-G5^2 domain, which was added as a repeated unit for the long SpyTag-SasG-SpyTag construct in our study. b) The structure of G5^1-E-G5^2.

We aimed to take advantage of the stiff, rod-like properties of SasG monomers to build discrete, rectangular arrays using SpyTag/SpyCatcher technology. We first designed protein building blocks with SasG domains flanked by SpyTag peptides at the N- and C-termini. Due to the modular nature of SasG proteins, we created two constructs of different sizes. The short construct consists of only the G5-E-G5 domain of SasG, and the longer construct is the GEG domain with three extra G5-E repeats at the C-terminus. We used SpyCatcher arrays with single SpyCatcher units, two SpyCatcher units, and three SpyCatcher units as "clamps" to orient the SpyTag-SasG-SpyTag rods (Figure 2.2).
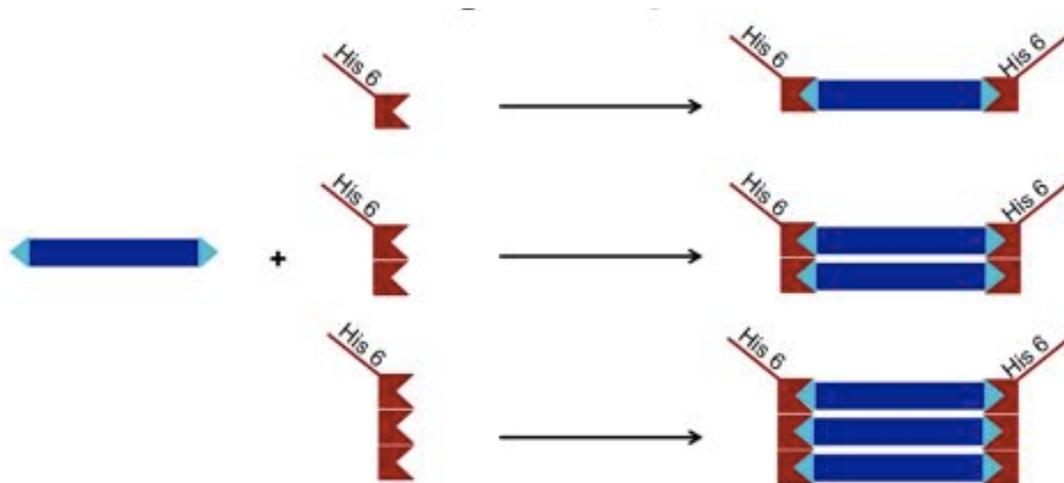
Figure 3.2 A cartoon schematic of SasG complex design. SasG rods (navy) flanked by SpyTag peptides (teal triangles) are mixed with SpyCatcher arrays (maroon crowns) with various numbers of binding domains to form rectangular complexes.

Both the SpyCatcher arrays and SpyTag-SasG-SpyTag constructs were designed with N-terminal hexahistidine tags. Each construct was expressed and purified separately, and the hexahistidine tag from SpyTag-SasG-SpyTag was removed by incubation with TEV protease. By incubating purified protein arrays *in vitro,* complexes were allowed to self-assemble into discrete geometric rectangular structures. We tested different ratios of SpyCatcher arrays and SpyTag-SasG-SpyTag rods, mixing purified components and incubating overnight at room temperature. Through SDS PAGE analysis, we determined that the best complex formation resulted from an equal ratio of SpyTag-SasG-SpyTag to SpyCatcher arrays (Figure 2.3).
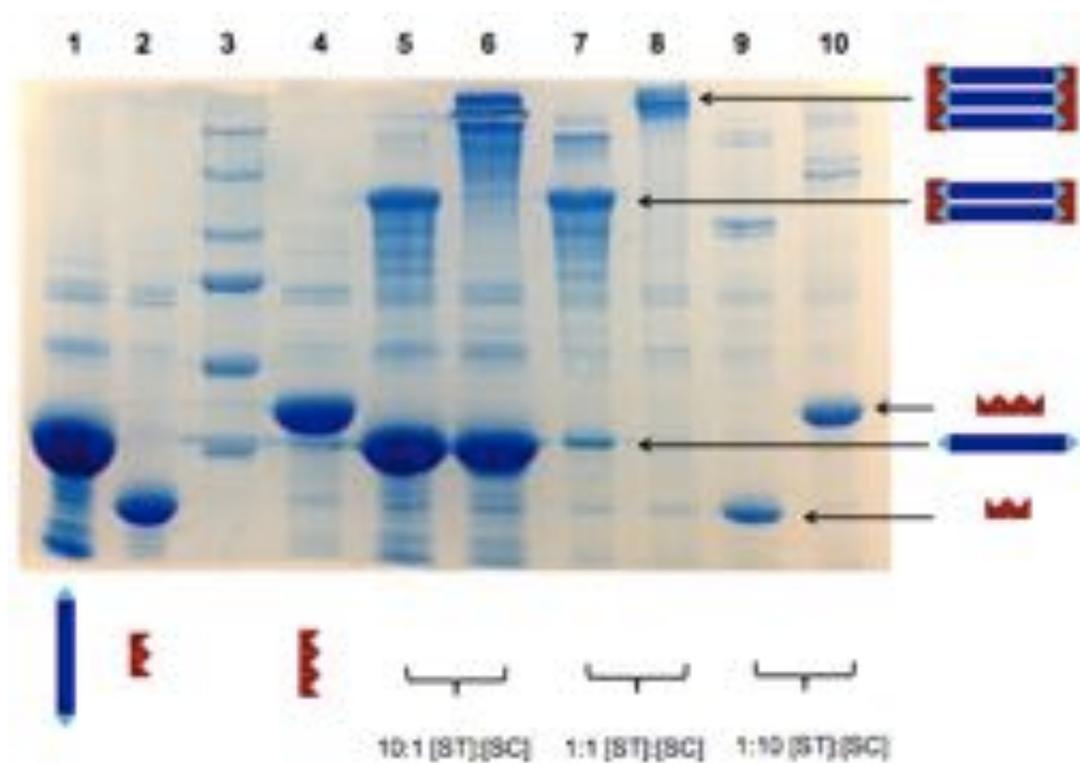
Figure 3.3 An SDS PAGE gel showing SasG complex formation at different mixing ratios. Lane 1: SpyTag-SasG-SpyTag alone; Lane 2: SpyCatcher array with two binding domains, alone; Lane 3: Molecular weight marker; Lane 4: SpyCatcher array with three binding domains; Lane 5: SasG complex formed by mixing SpyTag-SasG-SpyTag and SpyCatcher with two binding domains in a ratio of 10:1; Lane 6: SasG complex formed by mixing SpyTag-SasG-SpyTag and SpyCatcher with three binding domains in a ratio of 10:1; Lane 7: SasG complex formed by mixing SpyTag-SasG-SpyTag and SpyCatcher with two binding domains in a ratio of 1:1; Lane 8: SasG complex formed by mixing SpyTag-SasG-SpyTag and SpyCatcher with three binding domains in a ratio of 1:1; Lane 9: SasG complex formed by mixing SpyTag-SasG-SpyTag and SpyCatcher with two binding domains in a ratio of 1:10; Lane 10: SasG complex formed by mixing SpyTag-SasG-SpyTag and SpyCatcher with three binding domains in a ratio of 1:10.

Because only SpyCatcher arrays contain an affinity tag, entire complexes can be purified from an additional Ni-NTA binding step (Figure 3.4a). Complexes can be further purified using gel filtration (Figure 3.4b). Using this method, SasG complexes of many different dimensions can be assembled, depending on the number of binding domains in the SpyCatcher domain as well as the length of SpyTag-SasG-SpyTag construct (Figure 3.5).
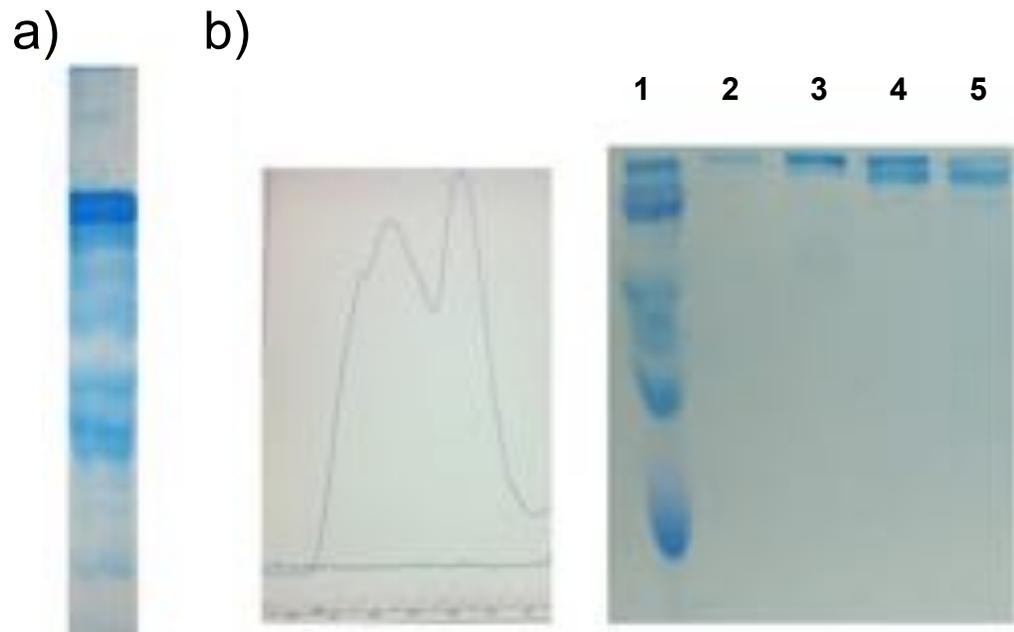
Figure 3.4 Additional purification of SasG complexes. a) An SDS PAGE gel showing a SasG complex after an additional Ni-NTA purification step. b) A gel filtration chromatograph (left) and SDS PAGE gel (right) showing the separation of two discrete bands after running over a SuperDex 10/300 column. The higher order molecular weight band in Lane 2 corresponds to the estimated molecular weight of the purified SasG complex (~200 kDa).
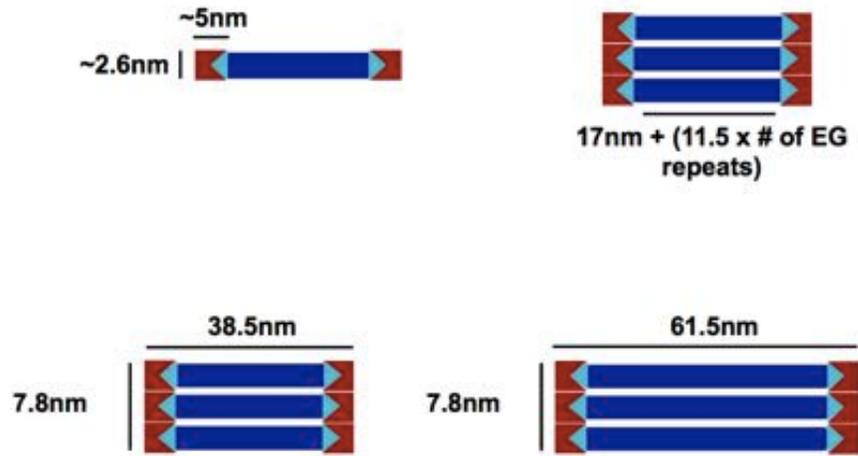
Figure 3.5 Cartoon representations of rectangular SasG complexes of various sizes. Dimensions of SasG complexes can be adjusted by using different lengths of SasG rods and SpyCatcher arrays with different numbers of binding domains (top row). The bottom row shows the estimated dimensions of our arrays using two different lengths of SasG rods.

3.3 Biotinylated SpyCatcher Arrays for Use in Biosensing

The ability to concatenate SpyCatcher domains on a single construct opens up the possibility of signal amplification for applications such as biosensing. We envisioned a protein-based system that could be used to array recognition elements attached to SpyTag peptides. As a proof of principle, we decided to use HER2 affibodies as our recognition element. Affibodies are small proteins that are engineered to mimic antibodies (Lofblom et al., 2010), and HER2 is a relevant biomarker for breast cancer (Ross, 2009). The HER2 affibody is a small, helical protein that binds to the extracellular domain of HER2 with low picomolar affinity (Eigenbrot, Ultsch, Dubnovitsky, Abrahmsen, & Hard, 2010). Its small size and ability to be expressed from *E. coli* made it a great candidate for this application. By fusing the SpyTag peptide to the HER2 affibody (referred to as HER2Af-SpyTag), we were able to array up to three affibodies at once through co-expression with SpyCatcher arrays. For this application, we added a BirA peptide sequence to the C-terminus of our SpyCatcher arrays, which allows for site-specific biotinylation by *E. coli* biotin ligase (Fairhead & Howarth, 2015). Biotinylated SpyCatcher arrays co-expressed with HER2Af-SpyTag form complexes displaying up to three affibodies (Figure 3.6). Biotinlyation allows for facile immobilization of these recognition elements to streptavidin-coated surfaces, such as silicon nanowires.
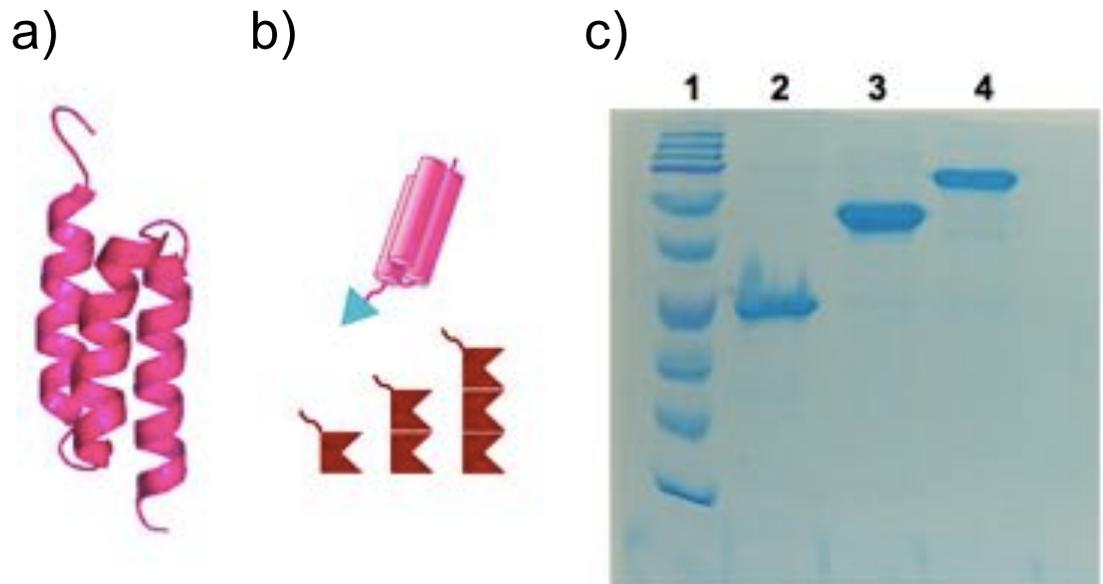
Figure 3.6 Components to form HER2 sensing complexes using SpyCatcher arrays and SpyTag fused to a HER2 affibody. a) The crystal structure of the HER2 affibody used in this design (PDB: 3MZW). b) Cartoons of protein constructs used in this design. SpyTag (teal triangle) is fused to the N-terminus of the HER2 affibody (pink helices). SpyCatcher arrays with one, two, and three binding domains (maroon crowns) contain a C-terminal BirA tag, resulting in biotinylated arrays. c) An SDS PAGE gel showing purified HER2 sensing complexes. Lane 1: Ladder; Lane 2: Purified complex displaying one affibody; Lane 3: Purified complex displaying two affibodies; Lane 4: Purified complex displaying three affibodies.

Before attaching these HER2 recognition complexes to nanowires, we first wanted to confirm that signal amplification corresponded to the number of recognition elements displayed. We also wanted to ensure that the biotinylated complexes could in fact bind to surfaces. To do this, we used eGFP-SpyTag as a recognition element instead of SpyTag-HER2. Instead of co-expressing eGFP-SpyTag with the arrays, we mixed purified constructs together in vitro to form biotinylated complexes displaying either one, two, or three eGFP molecules. As a control, we created one construct that was not biotinylated displaying three eGFP molecules (Figure 3.7). To perform the experiment, we designed a plate reader assay in which each complex was immobilized on neutravidin-coated plates. We measured eGFP fluorescence using a plate reader, and found that the signal did increase with the number of SpyCatcher binding units (Figure 3.8). We also observed that the non-biotinylated complex did not yield fluorescence signal, confirming that the fluorescence observed was due to the immobilization of biotinylated complexes.
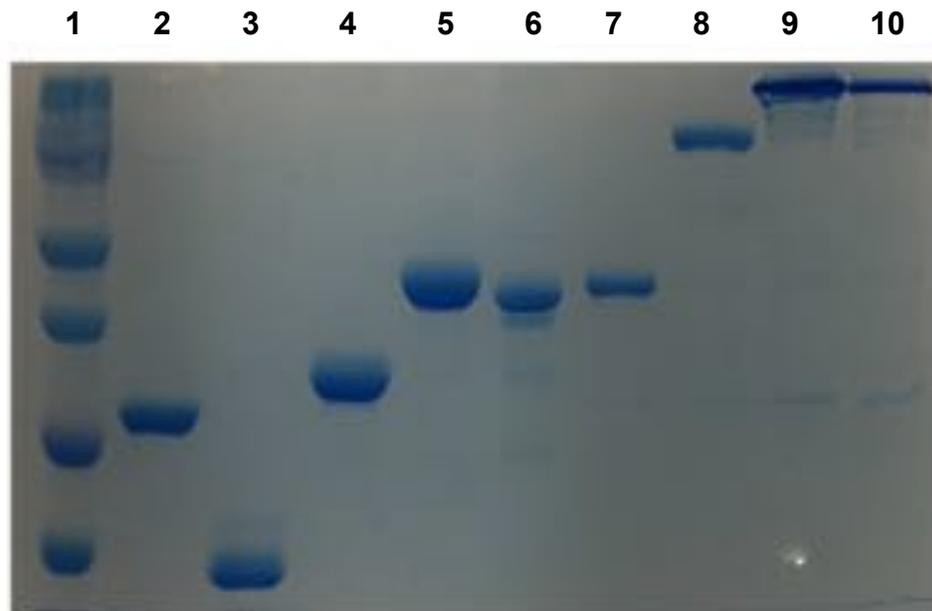
Figure 3.7 An SDS PAGE gel showing the formation of biotinylated complexes displaying eGFP. Lane 1: Ladder; Lane 2: Purified eGFP-SpyTag; Lane 3: Purified biotinylated SpyCatcher; Lane 4: Purified biotinylated SpyCatcher array with two binding domains; Lane 5: Purified biotinylated SpyCatcher array with three binding domains; Lane 6: Purified SpyCatcher array with three binding domains; Lane 7: Biotinylated complex displaying one eGFP molecule; Lane 8: Biotinylated complex displaying two eGFP molecules; Lane 9: Biotinylated complex displaying three eGFP molecules; Lane 10: Complex displaying three eGFP molecules.
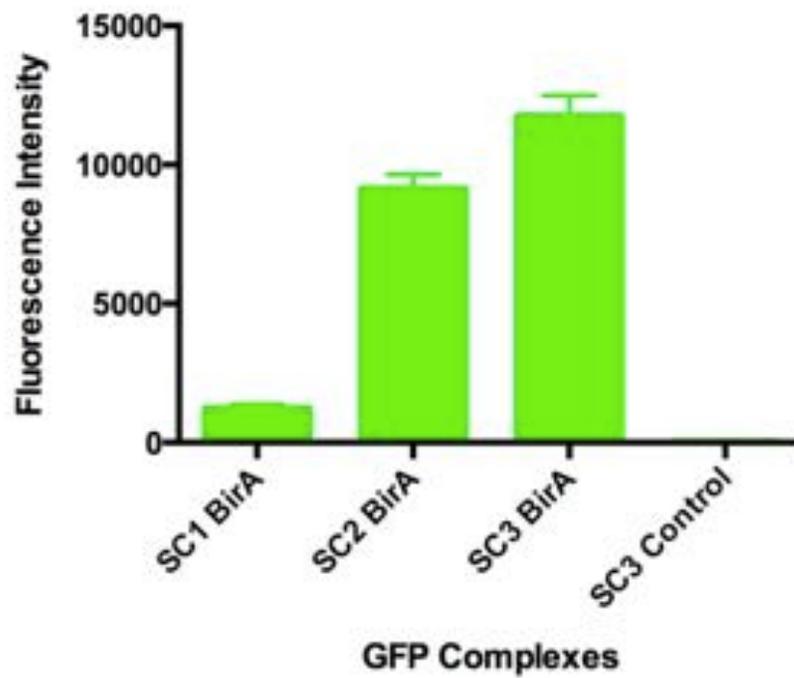
Figure 3.8 Quantification of eGFP fluorescence from plate reader assay. Fluorescence signal increased from left to right, as the complexes are displaying more eGFP molecules. No signal was observed for the non-biotinylated control.

3.4 Discussion and Conclusion

Using SpyCatcher arrays to form new protein geometries has the potential for use in enzyme scaffolding and signal amplification. For these studies, arrays were limited to three consecutive SpyCatcher domains, but the number of binding domains could be increased further. We have not yet explored the limits of how many SpyCatcher units can be concatenated, but doing so could expand the functions of these presented designs.

Rectangular SasG complexes could be made wider through the use of SpyCatcher arrays with more binding units. The length could also be increased by designing longer SasG constructs to flank with SpyTag peptides. These studies only explored two lengths of SasG units, but this length can be increased without compromising stability (Gruszka et al., 2012). In the future, we would like to be able to obtain images of these complexes either by AFM or negative-stain EM. Preliminary attempts proved to be challenging, but could be aided by some type of labeling. For example, NTA-coated gold beads can be conjugated to the hexahistidine tags on the SpyCatcher region of the complexes, which could help with identification by AFM. It is interesting to note that the SasG designs formed discrete complexes and not a gel network. This is perhaps due to the rigidity of the SasG building blocks, which may not be flexible enough to form a gel meshwork.

In the future, we'd also like to test the ability of the HER2 sensing complexes to actually detect the extracellular domain of HER2. This system could be tested using an ELISA-like assay where the sensing complexes are immobilized on a NeutrAvidin-coated plate and exposed to purified HER2 protein. The presence of HER2 could be identified by a primary antibody that binds to a different region of HER2 than the affibody. The

overall goal would involve complex immobilization to NeutrAvidin-coated silicon nanowires to increase sensitivity of HER2 detection.

Chapter 4: Formation of BslA Coated Microcapsules

Text, figures, and tables in Chapter 4 are adapted from those published in:

This work was done in collaboration with graduate students Ashley Schloss, Wei Liu, and Gilad Kaufman. Ashley cloned the BslA fusion proteins in addition to eGFP-SpyCatcher, and aided in all protein expression, purification, and microcapsule functionalization experiments. Wei Liu performed the SFG analysis and collected surface pressure-area isotherms. Gilad Kaufman built the microfluidics device and created BslA-coated microcapsules.

4.1 Introduction

Here, we describe the use of a bacterial hydrophobin, in combination with a versatile protein-protein conjugation scheme, to manufacture robust and readily functionalizable microcapsules. The bacterial hydrophobin, biofilm surface layer protein A (BslA), forms a stable, ordered monolayer at air-water or air-oil interfaces (Bromley et al., 2015; Hobley et al., 2013; Wang et al., 2016). The amphiphilic nature of BslA, which underlies such behavior, is evident from its 'bipartite' structure (Hobley et al., 2013) - a hydrophilic, classical iG fold and a hydrophobic cap (Figure 4.1). Its potential as a surface coating and emulsifier is well documented; specifically its ability to bind air, fat and water together during ice cream production produces a smoother consistency and increases the stability of the mixture, allowing ice creams to stay frozen longer in warm temperatures (Stanley-Wall & MacPhee, 2015).
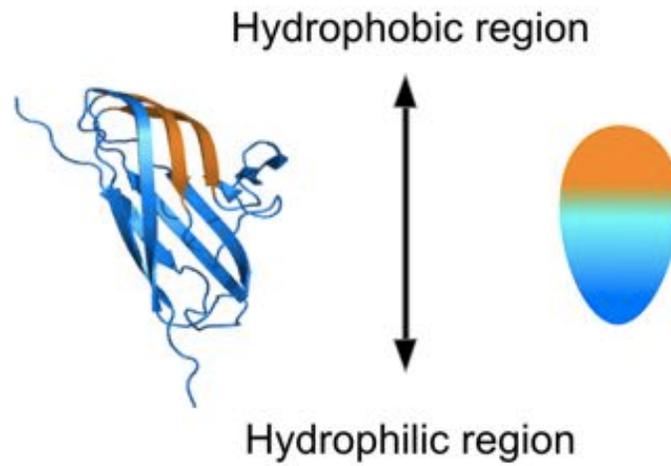
Figure 4.1 Ribbon representation and cartoon of the BslA protein. The orange coloring indicates the hydrophobic region, while the hydrophilic region is shown in blue.

4.2 Design of Protein Building Blocks

We took advantage of the SpyCatcher–SpyTag system (Zakeri et al., 2012; Zhang et al., 2013), which enables a spontaneous covalent isopeptide bond to be formed between a lysine side chain on the SpyCatcher protein and an aspartic acid side chain on the SpyTag peptide. We genetically engineered constructs that fuse the 13-residue SpyTag peptide to either the N- or C- terminus of BslA (Figure 4.2, 4.3), separated by a flexible linker (GGSGGS). BslA modified with the SpyTag peptide may then be functionalized by reaction with any protein of interest expressed as a fusion protein with SpyCatcher. For this study, we used SpyCatcher fused to the C-terminus of eGFP (Figure 4.2, 4.3) as a proof of principle.
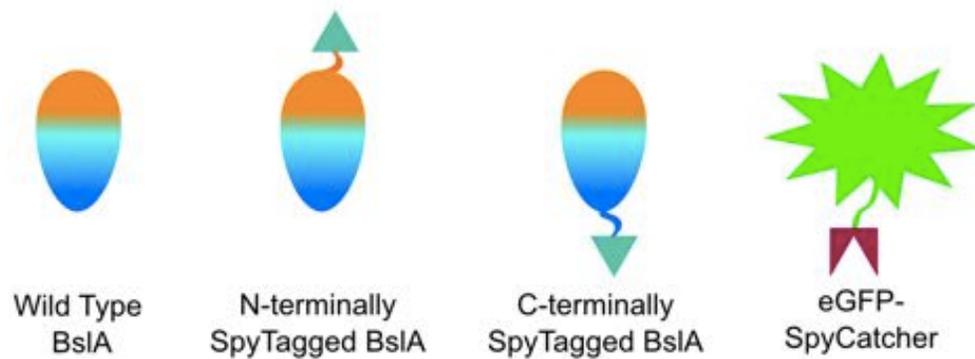
Figure 4.2 Cartoon representations of BslA fusion proteins and fluorescent fusion proteins. SpyTag (teal triangle) are attached to the C-terminus of BslA (orange and blue oval). SpyCatcher (maroon crown) and SnoopCatcher (gold crown) are attached to the C-terminus of eGFP (green starburst).
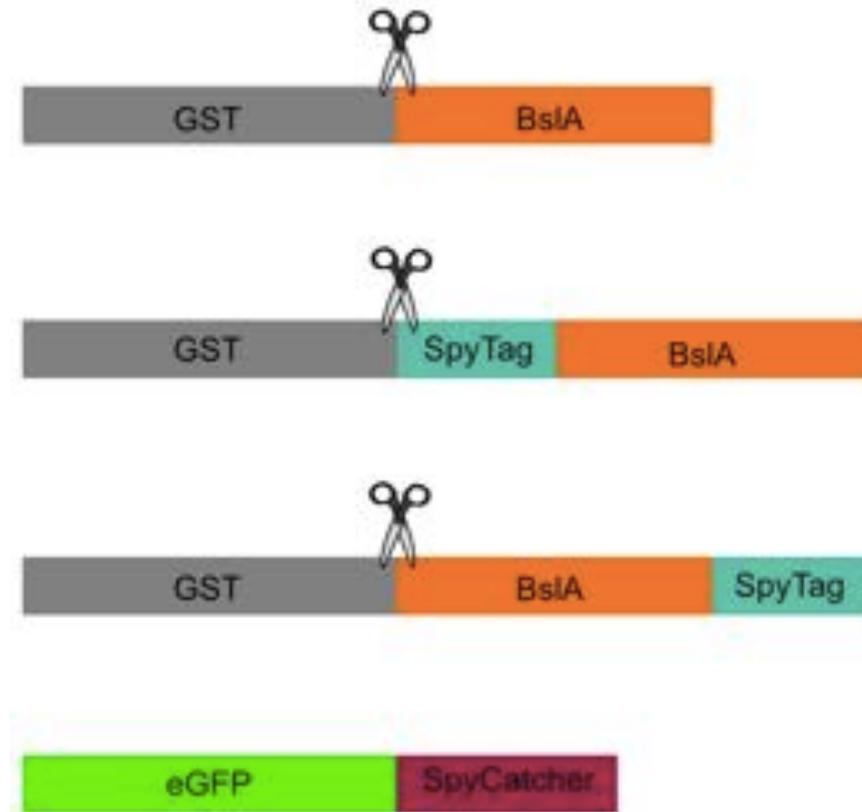
Figure 4.3 A color-coded schematic of fusion protein constructs. N-terminal GST tags (gray) are separated from BslA (orange) by a PreScission protease cleavage site, indicated by scissors cartoon. SpyTag (teal) and SnoopTag (purple) peptides are attached to the C-terminus of BslA by a GGSGGS linker. SpyCatcher (maroon) and SnoopCatcher (gold) are attached to the C-terminus of GFP (green) and mCherry (red) respectively by a GGSGGS linker.

4.3 Surface Characterization of BslA Monolayers

BslA with SpyTag attached at either end is able to form monolayers at an air-water interface. Wild-type (wt) BslA forms a well-defined monolayer at an air-water interface, the properties of which can be readily characterized using a Langmuir-Blodgett apparatus. We compared the behavior of BslA fused to SpyTag at either the N- or C-terminus with that of wt BslA, measuring surface pressure-area compression isotherms for each (Figure 4.4a). Monolayers formed by wt BslA or N- or C-terminally SpyTagged BslA all have collapsing surface pressures of ~65 mN/m, comparable to that of a typical phospholipid monolayer (55-65 mN/m). The limiting area of each monolayer was determined by extrapolation of the isotherm in the solid phase to zero surface pressure, giving values of 720 ± 12, 790 ± 4, and 900 ± 23 Å^2 for wt BslA, C-terminally SpyTagged BslA, and N-terminally SpyTagged BslA, respectively. Thus C-terminally SpyTagged BslA occupies a comparable surface area per molecule to that of wt BslA, whereas N-terminally SpyTagged BslA occupies a larger area, suggesting that the N-terminal tag causes a greater perturbation to the uniform monolayer structure than does the C-terminal tag. Moreover, the slopes of the compression isotherms of both SpyTagged BslA proteins are smaller than that of wt BslA, suggesting that the SpyTagged BslA films are more compressible than the wt BslA film (Figure 4.4a). We investigated the possible structural changes that underlie these changes in the macroscopic physical properties by molecular dynamics simulations (*vide infra*).

Characterization of the self-assembled monolayers using nonlinear surface-specific vibrational sum frequency generation (SFG) spectroscopy (Eisenthal, 1996; Richmond, 2002; Shen, 1989; Yan, Fu, Wang, & Liu, 2014) reveals the structural

similarities and differences between the different BslA constructs. Using SFG to characterize protein structure in the monolayers, we observed that spectra in the amide I region for wt BslA, C-terminally SpyTagged BslA, and N-terminally SpyTagged BslA at an air-water interface show two peaks at ~1675 and ~1690 cm^{-1}, assigned to β-turns and to the B1 mode of antiparallel β-sheets, respectively (Figure 4.4b) (Wang et al., 2016). The peak at ~1690 cm^{-1}, which is characteristically narrow in wt BslA (Wang et al., 2016), is also narrow with a full-width-at-half-maximum of less than 10cm^{-1} in the spectra of both C- and N-terminally SpyTagged BslA, indicating highly ordered structure; therefore this observation suggests that the presence of SpyTag at either end of BslA does not significantly perturb the ordered self-assembled structure at the interface. The SFG spectra in the C-H stretch region (Figure 4.4c) provide an indication of differences between the three proteins. Notably, the vibrational band of the symmetric methyl stretch for N-terminally SpyTagged BslA (~2884 cm^{-1}) was blue-shifted by ~10 cm^{-1} compared to that band in both wt BslA (~2876 cm^{-1}) and C-terminally SpyTagged BslA (~2875 cm^{-1}).
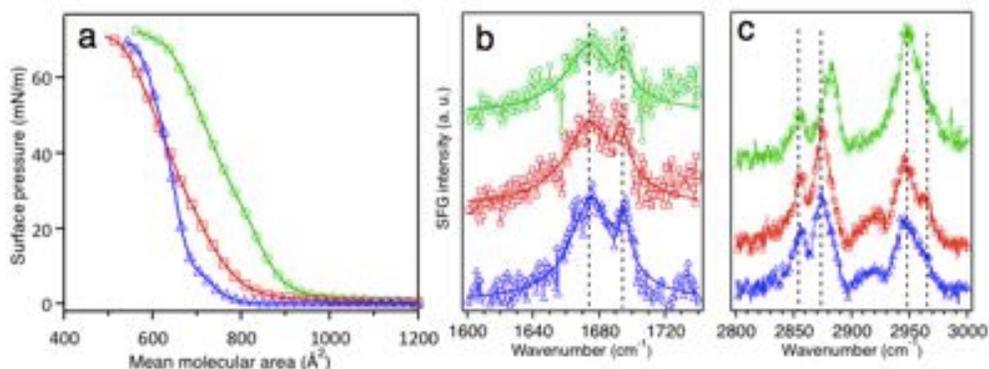
Figure 4.4 Surface characterization of wt BslA (blue triangles), C-terminally SpyTagged BslA (red squares), and N-terminally SpyTagged BslA (green circles) at the air−water interface. a) Surface pressure−area isotherms measured using a Langmuir−Blodgett apparatus. b) SFG spectra of BslA proteins, showing the Amide I region. To aid comparison, dashed lines indicate 1675 and 1690 cm^{-1}. c) SFG spectra of BslA proteins showing the C−H stretch region. Dashed lines indicate 2875, 2884, and 2942 cm^{-1}.

4.4 Fabrication of BslA Coated Microcapsules

The thin, uniform monolayer of BslA protein that is formed at an air-water or oil-water interface suggests that SpyTag peptides displayed by fusion to BslA will be readily accessible. Having established that C-terminally SpyTagged BslA forms robust monolayers, we proceeded to use this protein to fabricate microcapsules. Using a microfluidic device (Figure 4.5a). We formed oil-in-water microcapsules shown schematically in Figure 4.5b. Both wt BslA and a 3:1 mixture of wt BslA and C-terminally SpyTagged BslA form monodisperse microcapsules (Figure 4.5c) with mean diameters of 108 ± 2.4 μm and 106 ± 3.7 μm, respectively. Capsules formed from either wt BslA or a 3:1 mixture of wt BslA and C-terminally SpyTagged BslA (hereafter referred to as wt BslA capsules and SpyTagged BslA capsules, respectively) remain monodisperse and stable at room temperature for weeks.
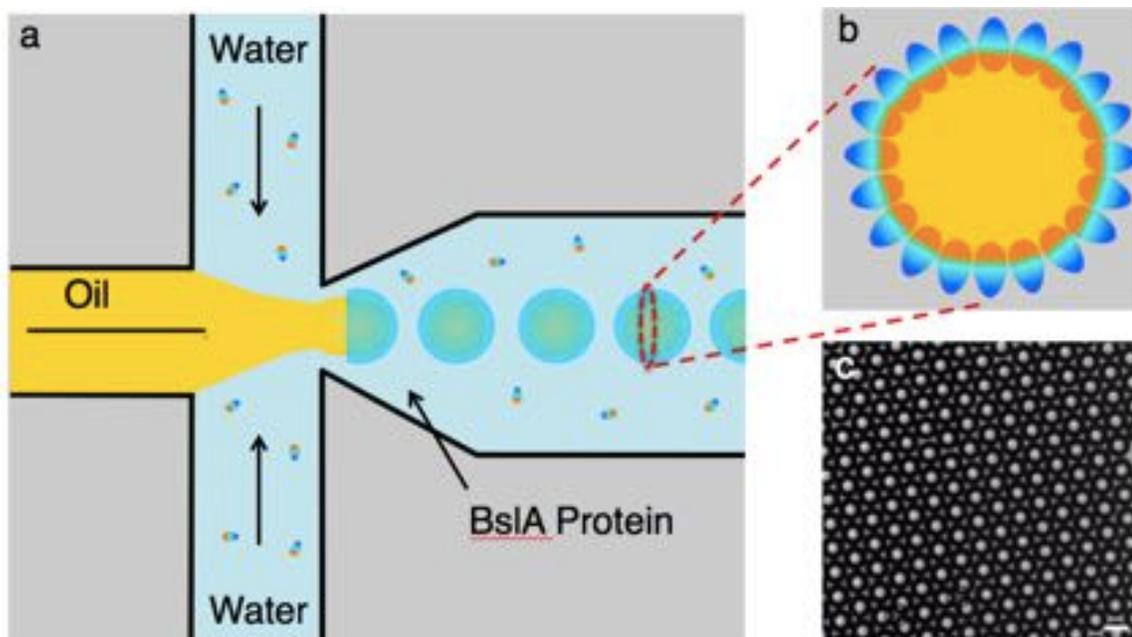
Figure 4.5 Fabrication of BslA microcapsules. a) Schematic illustration of the key region of the microfluidics device (not to scale). Aqueous phase (light blue) containing BslA protein (blue-orange ovals) and an oil phase (yellow) are brought into contact to form an oil-core microcapsule with a protein shell. b) Schematic illustration of a cross-section of an oil filled BslA capsule. The hydrophobic end of BslA (orange) faces inward, interacting with the oil and the hydrophilic end (blue) faces outward, interacting with the aqueous phase. c) Brightfield images of stable oil-in-water capsules formed with WT BslA. Scale bar is 100 μm.

Decoration of BslA Coated Microcapsules

To directly confirm that C-terminally SpyTagged BslA forms a covalent bond with SpyCatcher, we fused Glutathione S transferase (GST) to C-terminally SpyTagged BslA (GST-BslA-SpyTag) and incubated it with eGFP-SpyCatcher. The GST was fused to BslA to make the protein large enough to visualize on an SDS PAGE gel and also to prevent it from aggregating into inaccessible micelles. We analyzed the products of this reaction by boiling in SDS followed by SDS denaturing gel electrophoresis. The appearance of a new product, with a mobility consistent with a covalently linked GST-BslA-SpyTag and eGFP-SpyCatcher, is clearly evident. GST-BslA alone does not form such a complex with eGFP-SpyCatcher (Figure 4.6).

SpyTagged BslA capsules can form a covalent linkage with a SpyCatcher-enhanced green fluorescent protein (eGFP) fusion protein. To enable us to readily visualize the product, we reacted C-terminally Spy-Tagged BslA with eGFP-SpyCatcher fusion protein (Figure 4.7a). Figure 4.7b shows the results of incubating wt BslA capsules and C-terminally SpyTagged BslA capsules with eGFP-SpyCatcher. Only the capsules SpyTagged BslA capsules react to form a linkage with SpyCatcher-GFP that is resistant to washing. The attraction of labeling in this fashion is that the C-terminal SpyTag does not perturb BslA's ability to self-associate at interfaces and to form robust capsules. Such capsules may then be covalently labeled with SpyCatcher fused to any protein of interest.
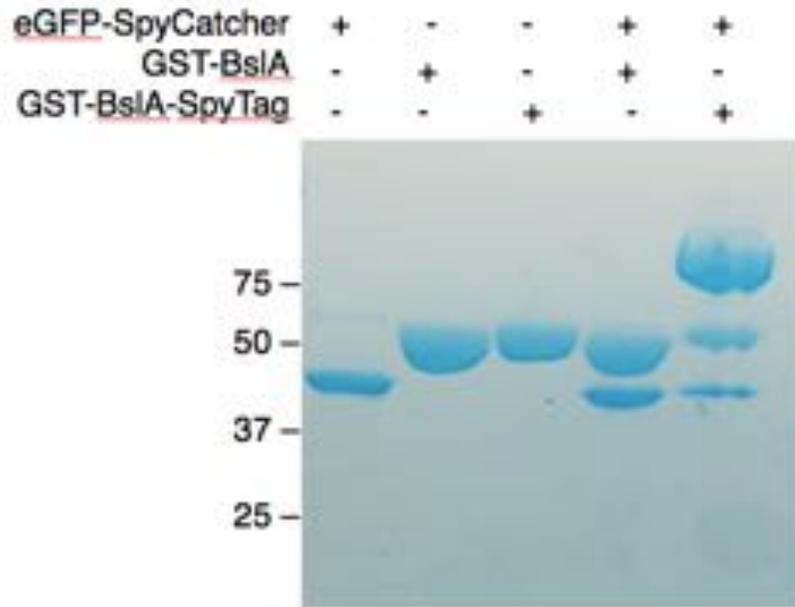
Figure 4.6 An SDS PAGE gel showing *in vitro* conjugation. C-terminally SpyTagged BslA and SpyCatcher-GFP form a covalent linkage, detected by SDS PAGE analysis. Lanes 1-3: The protein components eGFP-SpyCatcher, GST-BslA and GST-BslA-SpyTag. Lane 4: Unreacted control obtained by incubating eGFP-SpyCatcher and GST-BslA together for one hour. The presence of two distinct bands and no new higher molecular weight band indicates that the two proteins do not react to form a covalent linkage. Lane 5: Reaction product obtained by incubating eGFP-SpyCatcher and GST-BslA-SpyTag together for one hour. The appearance of a new, higher molecular weight band corresponds to the formation of the covalently linked eGFP-SpyCatcher + GST-BslA-SpyTag protein complex. This conclusion is supported by the depletion of the bands corresponding to the unreacted proteins.
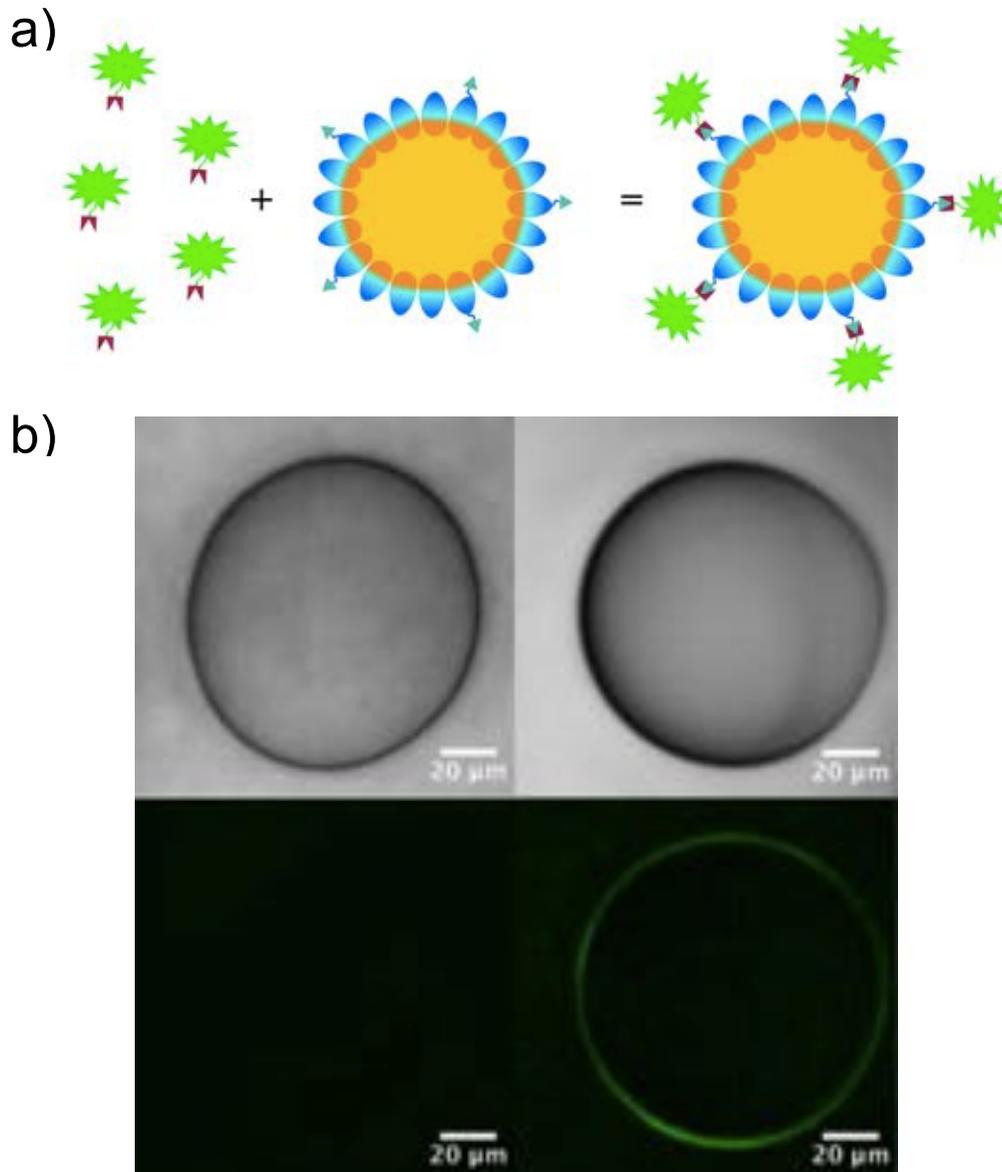
Figure 4.7 Decoration of microcapsules with eGFP. a) Cartoon illustration of the decoration of microcapsules formed from C-terminally SpyTagged BslA by reaction with eGFP-SpyCatcher. b) Brightfield and fluorescent images of wt BslA and C-terminally SpyTagged BslA after reaction with eGFP-SpyCatcher. Top: Microscope bright-field images of WT BslA (left) and C-terminally SpyTagged BslA (right) microcapsules. Bottom: Fluorescent images of the capsules after 10 min incubation with eGFP-SpyCatcher followed by washing with water, of wt BslA (left) and C-terminally SpyTagged BslA (right). eGFP labeling of the capsules only occurs via the SpyTag−SpyCatcher reaction.

Discussion and Conclusion

In summary, we have demonstrated a simple strategy to prepare monolayers and microcapsules using C-terminally SpyTagged BslA. BslA modified in this fashion retains the structural and mechanical properties of wt BslA. Moreover, it can be functionalized by the covalent attachment of any desired protein fused to SpyCatcher. We demonstrate the methodology by functionalizing the SpyTagged BslA capsules with eGFP-SpyCatcher, for ease of visualization. It is clear that this approach can be easily modified to attach a wide range of different proteins or peptides to the microspheres, such as ligands to important cellular biomarkers or receptors. One application we envision is the display of the extracellular domains of membrane proteins. This new class of functionalized surface arrays creates a novel platform for surface patterning, targeted drug delivery, and targeted imaging.

Chapter 5: Protein immobilization to surfaces using BslA

Text, figures, and tables in Chapter 5 are adapted from those submitted as:

Williams, D.M. *et al*. Facile protein immobilization using engineered surface active biofilm proteins. (2018)

This work was done in collaboration with Dr. Hadi Izadi, graduate students Gilad Kaufman and Sarah Prophet, and undergraduate student Abigail Gahm. Dr. Hadi Izadi fabricated the mold and poured PDMS stamps for microcontact printing. Gilad Kaufman provided expertise and aided in development of a microcontact printing protocol. Sarah Prophet performed surface pressure-area isotherms. Abigail Gahm contributed to experimental design and the use of the Langmuir-Schaefer attachment.

5.1 Introduction

The method presented in this chapter is straightforward, specific, and scalable, and it can be used for the surface attachment of any protein. The method successfully exploits the unique physical and chemical properties of natural proteins: BslA, which self-assembles to form a monolayer at a hydrophobic/hydrophilic interface (Hobley et al., 2013), and engineered streptococcal surface proteins (Veggiani et al., 2016; Zakeri et al., 2012; Zhang et al., 2013), which spontaneously form a covalent isopeptide bond between Lys and Asp/Asn side chains on two different polypeptides. The availability of two different protein pairs: SpyCatcher and SpyTag and SnoopCatcher and SnoopTag, which do not cross-react, provides a route for the simultaneous display of different protein recognition elements. By fusing BslA to SpyTag and SnoopTag and fluorescent proteins to SpyCatcher and SnoopCatcher, we created reactive pairs of proteins for use in our method. Key features of this method are that all the components are expressed

61

recombinantly and react spontaneously with extremely high efficiency. We anticipate that multiplexed, spatially distinct display of several recognition elements as demonstrated in this work will facilitate single-sample multi-analyte detection.

5.2 Design of Protein Building Blocks

Wild type (wt) BslA self-assembles at air-water interfaces to form robust monolayers, the properties of which have been well characterized (Bromley & MacPhee, 2017; Hobley et al., 2013; Kaufman et al., 2017; Morris, Bromley, Stanley-Wall, & MacPhee, 2016). We previously showed that the addition of the 13-residue peptide SpyTag to the C-terminus of BslA does not significantly perturb the formation of the BslA monolayer (Schloss, 2016). To allow for the possibility of displaying two recognition elements at once, we designed another construct attaching the SnoopTag peptide to the C-terminus of BslA. Thus, monolayers of BslA-SpyTag and BslA-SnoopTag can be used to simultaneously display proteins of interest attached to SpyCatcher and SnoopCatcher, respectively. For our proof of principle studies, we used eGFP attached to SpyCatcher at the C-terminus and mCherry attached to SnoopCatcher, also at the C-terminus. A total of four fusion proteins (Figure 5.1, Figure 5.2) were used in this method, in addition to wt BslA.
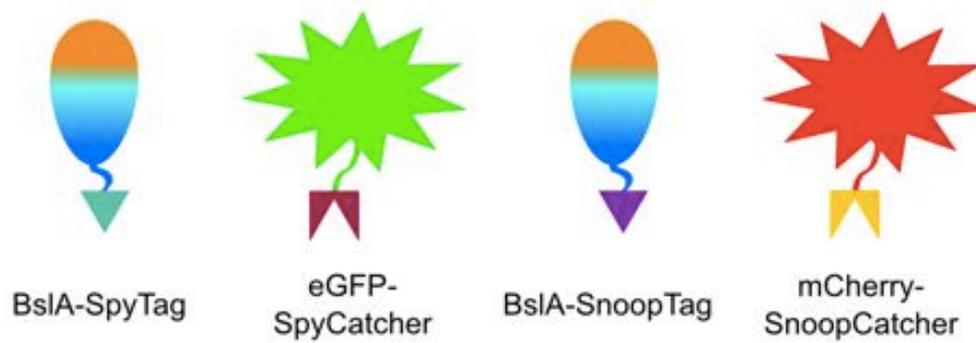
Figure 5.1 Cartoon representations of BslA fusion proteins and fluorescent fusion proteins. SpyTag (teal triangle) and SnoopTag (purple triangle) are attached to the C-terminus of BslA (orange and blue oval). SpyCatcher (maroon crown) and SnoopCatcher (gold crown) are attached to the C-terminus of eGFP (green starburst) and mCherry (red starburst), respectively.
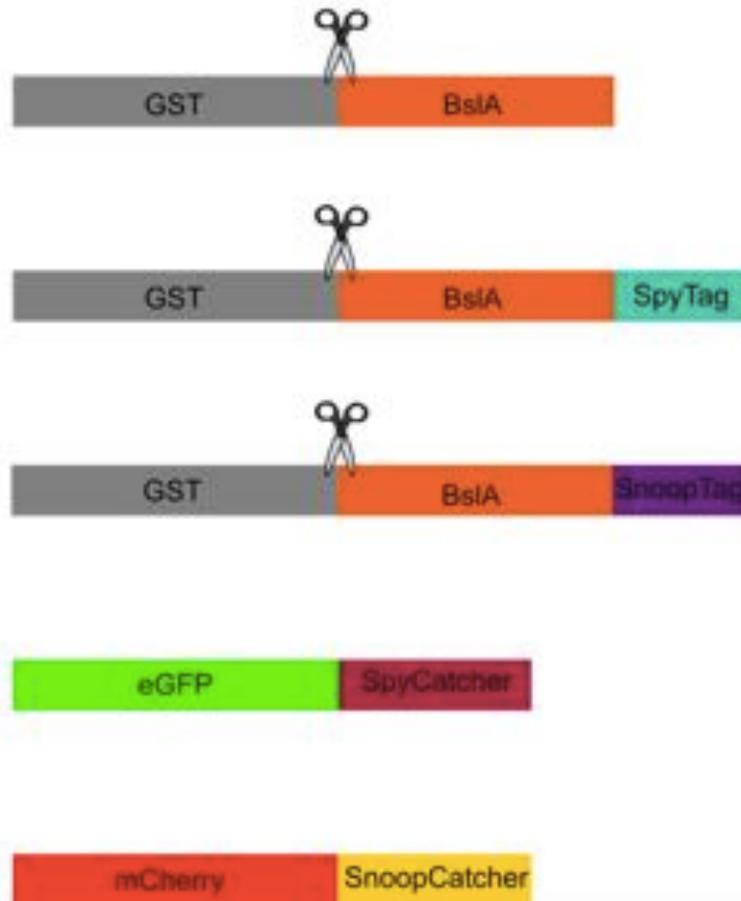
Figure 5.2 A color-coded schematic of fusion protein constructs.
N-terminal GST tags (gray) are separated from BslA (orange) by a
PreScission protease cleavage site, indicated by scissors cartoon.
SpyTag (teal) and SnoopTag (purple) peptides are attached to the
C-terminus of BslA by a GGSGGS linker. SpyCatcher (maroon)
and SnoopCatcher (gold) are attached to the C-terminus of eGFP
(green) and mCherry (red) respectively by a GGSGGS linker.

5.3 Characterization of BslA Monolayers

For the studies reported here, we used a Langmuir-Blodgett (LB) apparatus to form consistent, well-packed monolayers of BslA. We used mixtures of 25% BslA-SpyTag/75% wt BslA and 25% BslA-SnoopTag/75% wt BslA. We chose mixtures of tagged and wt BslA, as opposed to using 100% BslA-SpyTag and 100% BslA-SnoopTag, to decrease steric interference between molecules attached to the monolayer via BslA's C-terminal peptide. We characterized the behavior of such monolayers by measuring surface pressure-area isotherms using an LB apparatus (Figure 5.3).

The different protein monolayers all exhibit a similar collapse pressure of ~65 mN/m. We calculate the average area per molecule at 23 mN/m, which corresponds to the maximum surface pressure achievable before exerting mechanical compression force, to be 656 $Å^2$ for wt BslA, 753 $Å^2$ for 25% BslA-SpyTag/75% wt BslA, and 679 $Å^2$ for 25% BslA-Snooptag/75% wt BslA. The variability associated with different measurements of the same monolayer was calculated as the standard deviation of each trial, and all were minimal (around +/- 1%). The comparable collapse pressures and mean molecular areas make it clear that the behavior of these mixed monolayers is very similar to that of 100% wt BslA, indicating that the C-terminal fusions to the SpyTag and SnoopTag peptides cause little perturbation to the monolayer.
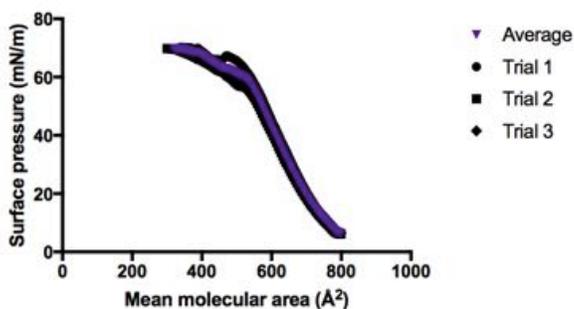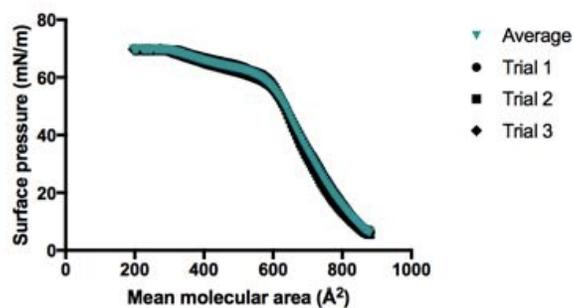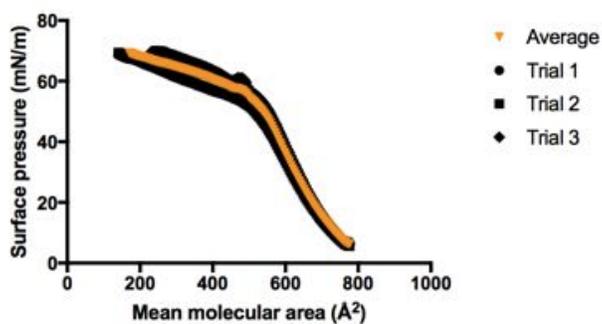
Figure 5.3 Surface pressure-area isotherms of BslA constructs. Data were obtained using a Langmuir-Blodgett apparatus. For each protein, a surface pressure versus area isotherm was measured in three independent experiments. Data from individual experiments are shown in black (circles, squares and diamonds) and the average of the three measurements is shown in colored triangles (100% wt BslA, orange; 25% BslA-SpyTag/75% wt BslA, teal; 25% BslA-SnoopTag/75% wt BslA, purple).

5.4 Microcontact Printing and Protein Deposition

We first used microcontact printing (Kane, Takayama, Ostuni, Ingber, & Whitesides, 1999) to create slides with a distinct pattern of individual hydrophobic spots on a glass surface (Figure 5.4a). PDMS micropillars 20 μm in diameter and 40 μm in height were fabricated by replica molding using micropatterned silicon molds (Izadi et al., 2016). The hydrophobic spots were created by microcontact printing of trichloro(1H,1H,2H,2H-perfluorooctyl)silane (FOTS) suspended in ethanol using a polydimethlysiloxane (PDMS) stamp. PDMS micropillar stamps were incubated with FOTS solution before being wicked away with a tissue, yielding an "inked" stamp. The stamp was placed on top of a clean glass slide with a 20 g weight on top. The weight and stamp were removed after 1 min, leaving behind a glass slide printed with a hexagonal pattern of circular, hydrophobic FOTS spots. The remainder of the surface was left untreated.

BslA protein monolayers were then transferred to these patterned glass slides using a Langmuir-Schaefer attachment (Figure 6.4b). 25% BslA-SpyTag/75% wt BslA or 25% BslA-SnoopTag/75% wt BslA were injected and allowed to equilibrate to the air-water interface of a Langmuir Blodgett trough. After equilibration, the barriers were compressed to a surface pressure of 23 mN/m, forming a protein monolayer at the air-water interface. This value was chosen based on prior studies that show that it results in the formation of monolayers that are reliably free of significant distortion (W. Liu, Li, Wang, Yan, & Leblanc, 2017). The patterned slides prepared earlier were lowered to make contact with the protein monolayer using a Langmuir-Schaefer apparatus. After making contact with the monolayer, the hydrophobic ends of 25% BslA-SpyTag/75% wt

BslA or 25% BslA-SnoopTag/75% wt BslA were transferred to the slide at the sites of the hydrophobic spots, resulting in a patterned slide displaying a protein monolayer. Slides were stored in DI water until further use.
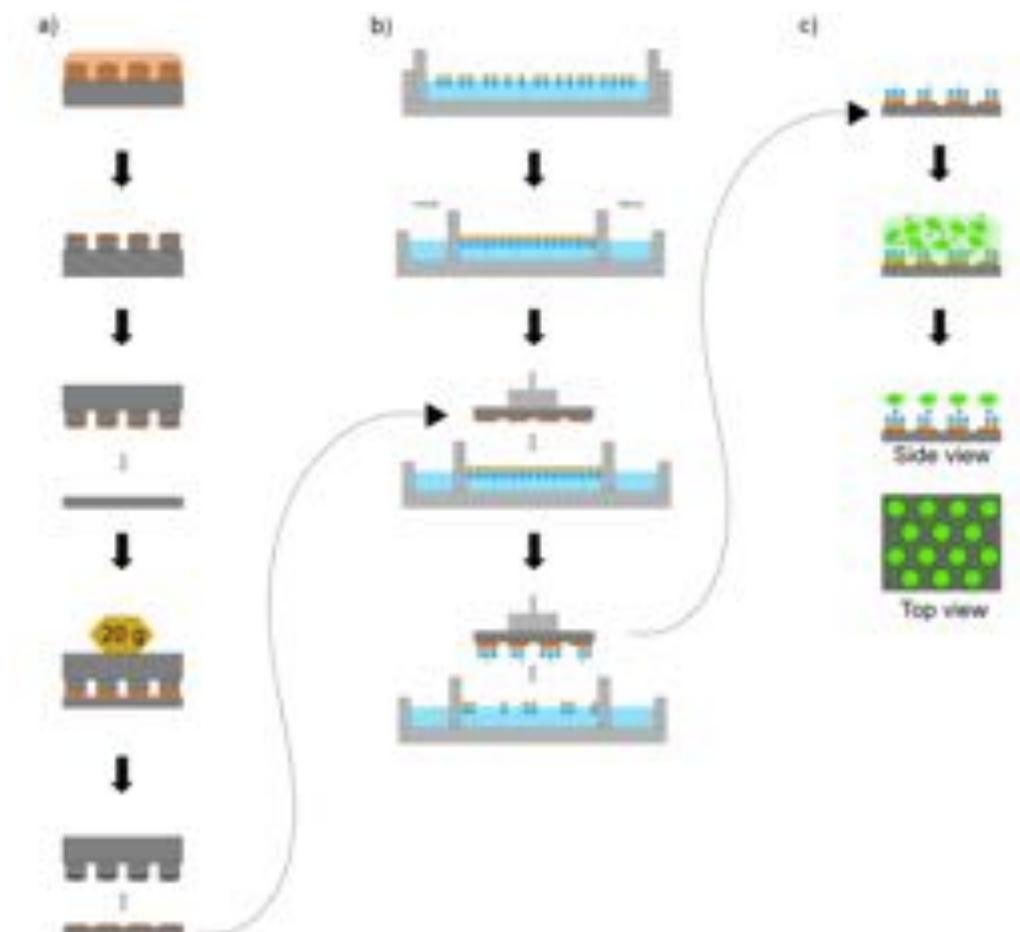
Figure 5.4 A cartoon schematic of the methods described in this chapter. (a) PDMS micropillar stamps were incubated with FOTS solution (orange) before being wicked away with a tissue, yielding an "inked" stamp. The stamp was placed on top of a glass slide with a 20g weight on top. The weight and stamp were removed, leaving behind a glass slide printed with a hexagonal pattern of circular, hydrophobic FOTS spots. (b) 25% BslA-SpyTag/75% wt BslA or 25% BslA-SnoopTag/75% wt BslA were injected and allowed to equilibrate to the air-water interface of an LB trough. After equilibration, the barriers were compressed to form a protein monolayer at the air-water interface. Patterned slides prepared from (a) were lowered to make contact with the protein monolayer using an LS apparatus. After making contact with the monolayer, the hydrophobic ends of 25% BslA-SpyTag/75% wt BslA or 25% BslA-SnoopTag/75% wt BslA were transferred to the slide at the sites of the hydrophobic spots, resulting in a patterned slide displaying a protein monolayer. (c) Patterned slides displaying a protein monolayer prepared in (b) were incubated with a solution of eGFP-SpyCatcher or mCherry SnoopCatcher (middle cartoon). Excess fluorescent proteins that did not bind to BslA-SpyTag or BslA-SnoopTag proteins were washed away with DI water to yield slides with fluorescently labeled circular spots. After rinsing, slides were wicked dry with a tissue and imaged using fluorescence microscopy.

5.5 Functionalization of Surfaces

Three different surfaces were created and tested: 25% BslA-SpyTag/75% wt BslA, 25% BslA-SnoopTag/75% wt BslA and 100% wt BslA. Surfaces were probed with either eGFP-SpyCatcher or mCherry-SnoopCatcher, rinsed with DI water (Figure 5.4c), and imaged using fluorescence microscopy (Figure 5.5). From these images it is clear that eGFP-SpyCatcher only reacts with and labels surfaces that contain BslA-SpyTag and mCherry-SnoopCatcher only reacts with and labels surfaces that contains BslA-SnoopTag. Neither eGFP-SpyCatcher nor mCherry-SnoopCatcher binds to the hydrophobic surface coated with BslA or with the non-cognate BslA-SpyTag or BslA-SnoopTag (Figure 6.5). Indeed, the BslA coatings reduce background binding to less than the background binding of the fluorescent protein to the uncoated glass slide. These observations are shown quantitatively in plots comparing the signal to background fluorescence intensity for each surface after probing (Figure 5.6). Thus, this strategy of attaching proteins to surfaces is both specific with respect to requiring a cognate SpyTag/SpyCatcher or SnoopTag/SnoopCatcher pair, and also essentially eliminates background binding to the hydrophobic surface.
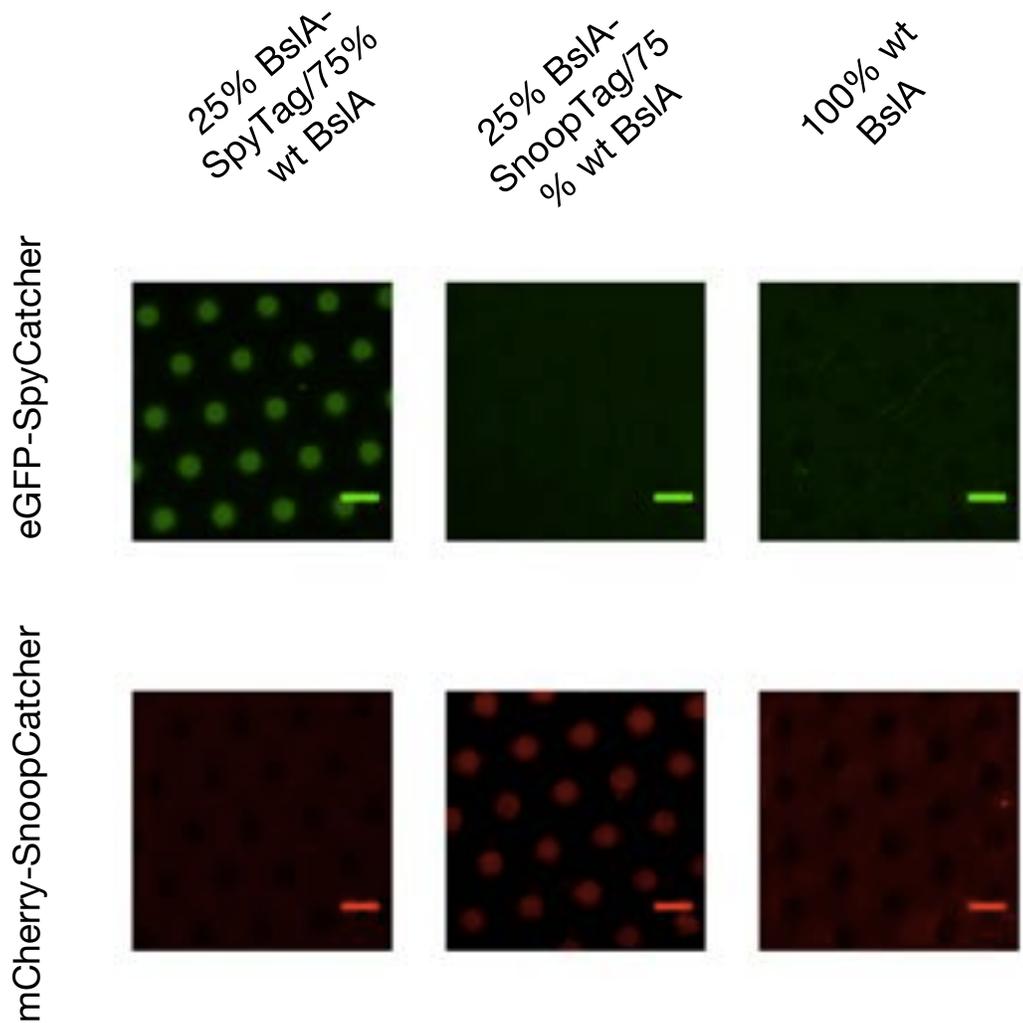
Figure 5.5 Fluorescence microscope images of functionalized patterned surfaces. Fluorescence microscope images of slides printed with a pattern of FOTS, displaying a monolayer of 25% BslA-SpyTag/75% wt BslA (left column), 25% BslA-SnoopTag/75% wt BslA (middle column), or 100% wt BslA (right column) were incubated with eGFP-SpyCatcher (top row) or mCherry-SnoopCatcher (bottom row). The images have been false colored to show eGFP fluorescence as green and mCherry fluorescence as red. All scale bars are 50 μm.
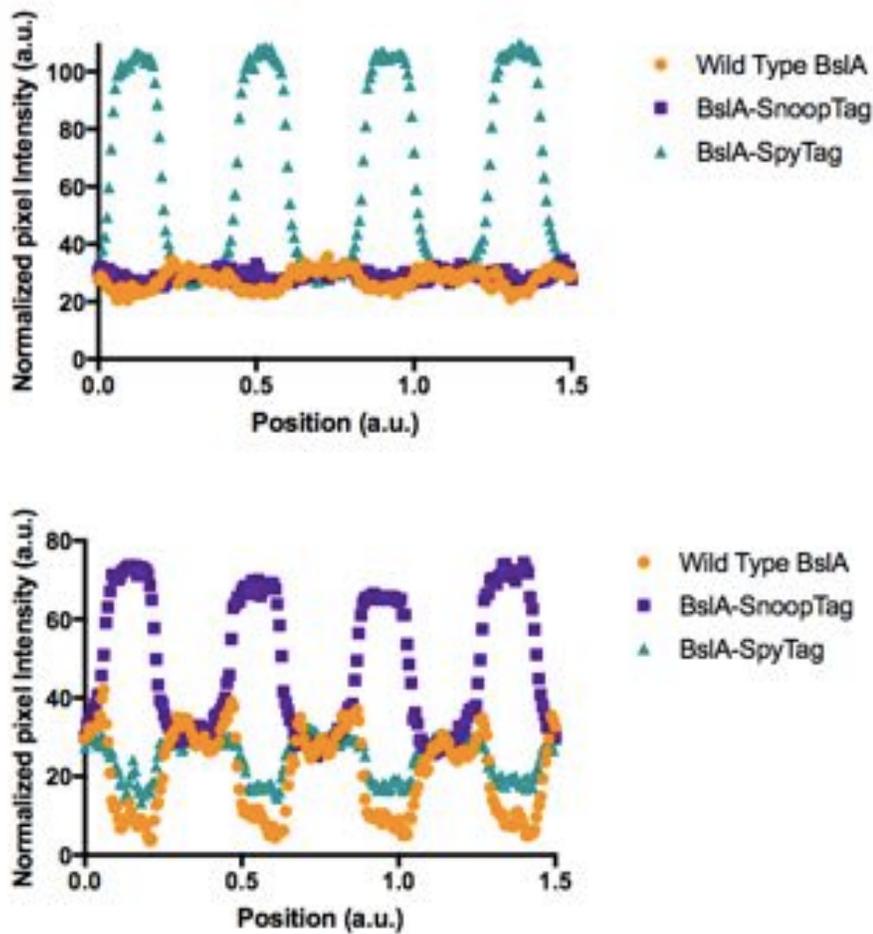
Figure 5.6 Fluorescence intensity profiles over four spots, from the images in Figure 5.5. Slides patterned with 100% wild type BslA (orange circles), 25% BslA-SpyTag/75% wt BslA (teal triangles), and 25% BslA-SnoopTag/75% wt BslA (purple squares) probed with eGFP (top) and mCherry-SnoopCatcher (bottom).

5.6 Prevention of Non-Specific Adsorption using WT BslA

In preliminary experiments, we tested the ability of a monolayer of wt BslA on the hydrophobic surface to prevent non-specific adsorption of fluorescent proteins to glass. In the absence of a BslA coating, the fluorescent protein fusions readily adsorb non-specifically to the hydrophobic surface (Figure 5.7). By contrast, such non-specific binding is effectively eliminated when the surface bears a monolayer of BslA. This is evident from the reduction of fluorescence in the area of the hydrophobic spot relative to that of non-specific binding to the glass slide (Figure 5.7). The data are noteworthy as they demonstrate that in addition to providing a novel means to attach a protein of interest to a surface, the BslA coating can also eliminate non-specific binding of proteins of interest to that surface. In the context of sensing, this is expected to reduce false negatives, i.e. to increase the confidence with which one can conclude that a species of interest is not present.
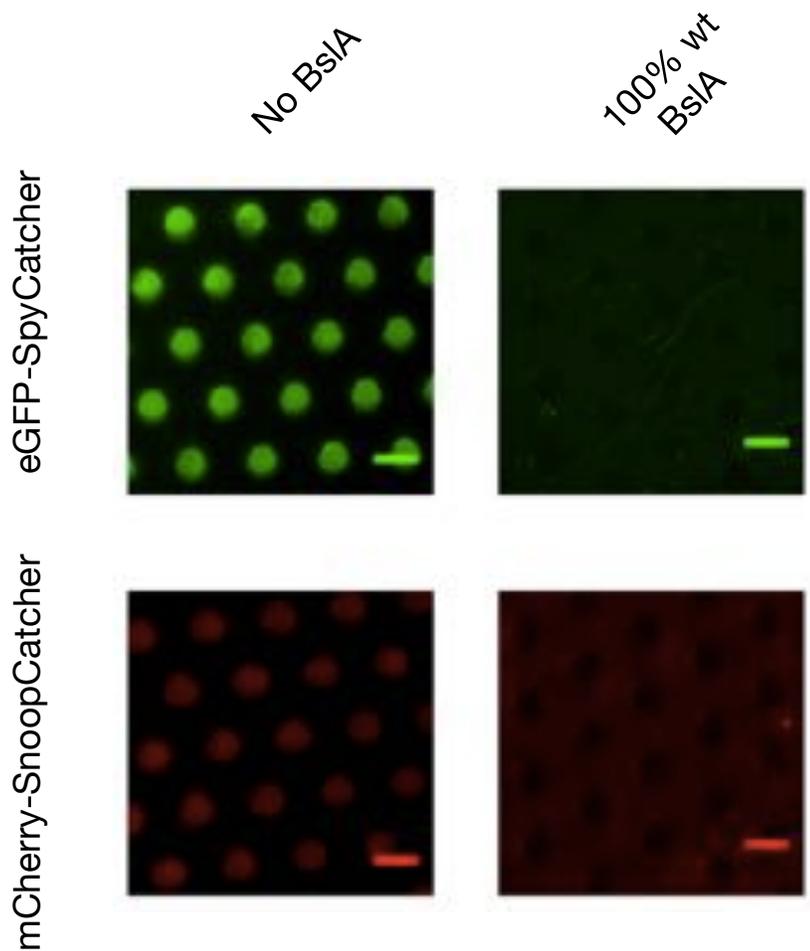
Figure 5.7 Microscope images comparing surfaces prepared with (right) and without (left) wt BslA after probing with fluorescent fusion proteins. FOTS displaying no protein monolayer (left column) and a protein monolayer of 100% wt BslA (right column) after incubation with eGFP-SpyCatcher (top row) and mCherry-SnoopCatcher (bottom row). The image has been false colored to show mCherry fluorescence as red. All scale bars are 50 μm.

5.7 Preliminary Experiments with Other BslA Fusion Proteins

Preliminary experiments suggest that the scope of this method could be broadened by changing the component fused to BslA (for example, SpyCatcher rather than SpyTag) and the component fused to the protein to be immobilized (SpyTag rather than SpyCatcher). We created BslA fusion proteins with SpyCatcher and SnoopCatcher attached to the C-terminus, in addition to fluorescent fusion proteins of eGFP with SpyTag and SnoopTag attached to the C-terminus. All fusion proteins are connected by a GGSGGS linker.

Surface pressure-area isotherms of 25% BslA-SpyCatcher/75% wt BslA show that the average area per molecule at 23 mN/m is 785 Å^2, while the average area per molecular at the same pressure is 600 Å^2 for 25% BslA-SnoopCatcher/75% wt BslA. These mixtures appear to be forming monolayers at the air-water interface and have collapse pressures of approximately 65 mN/m (Figure 5.8).
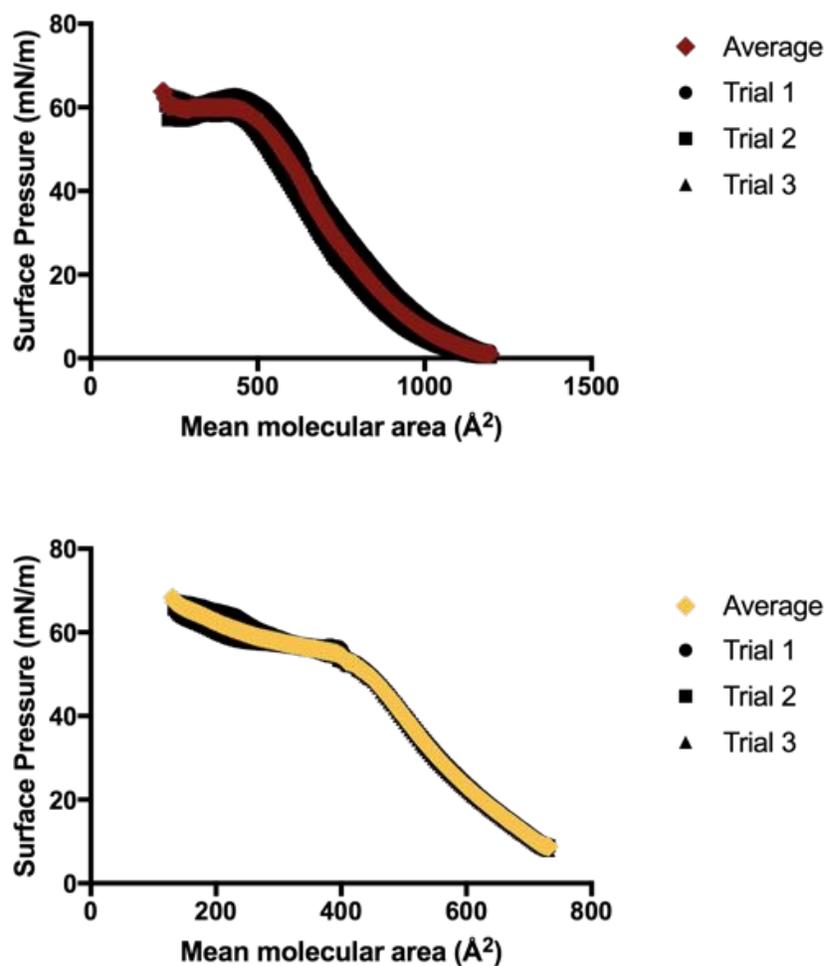
Figure 5.8 Surface pressure area isotherms of BslA-SpyCatcher and SnoopCatcher constructs. Data were obtained using a Langmuir-Blodgett apparatus. For each protein, a surface pressure versus area isotherm was measured in three independent experiments. Data from individual experiments are shown in black (circles, squares and diamonds) and the average of the three measurements is shown in colored triangles (25% BslA-SpyCatcher/75% wt BslA, maroon; 25% BslA-SnoopCatcher/75% wt BslA, gold).

5.8 Discussion and Conclusion

One can readily envision the use of the immobilization strategy presented here for practical biosensing applications, for example by immobilizing a protein against which an organism produces antibodies, and thereafter detecting the presence of those antibodies. Another example could be immobilization of recognition protein that binds to a molecule of interest, which is then detected by a second recognition protein, as in a 'sandwich ELISA' assay.

Through exploiting the self-assembling properties of natural proteins, we have created a simple but highly effective method for the immobilization of recognition elements to a surface. We demonstrated that functionalizing BslA with SpyTag and SnoopTag peptides does not significantly perturb monolayer formation. In principle, the method presented can be applied for any protein of interest attached to SpyCatcher or SnoopCatcher. When probing surfaces with fluorescent protein fusions, we did not observe any cross-reactivity of Snoop- variants with Spy- variants, and additionally saw no non-specific binding to surfaces deposited with 100% wt BslA. Moreover, the BslA coating is effective in eliminating non-specific protein surface binding, which could be useful for a broad range of other applications. Preliminary experiments indicate that it is possible for SpyCatcher/SnoopCatcher to be fused to BslA as opposed to the protein of interest, which would greatly increase the scope of this method. We anticipate the future application of the strategy we describe to immobilize therapeutically relevant proteins, towards the production of biosensors with increased sensitivity and specificity.

Chapter 6: Materials and Methods

6.1 General Methods

6.1.1 General Cloning Protocols

Restriction enzymes, Phusion polymerase, T4 DNA ligase, and appropriate reaction buffers were purchased from New England Biosciences (NEB). Restriction digests were performed for at least three hours at 37 °C. Ligations were performed overnight at 16 °C. 1 μL of the ligation reaction was heat inactivated at 65 °C and transformed into electrocompetent DH10B *E. coli* cells by electroporation. Single colonies were picked, and plasmids were purified using a Qiagen miniprep kit. Plasmids were sequenced with appropriate sequencing primers (W.M. Keck Facility, Yale University, New Haven, CT). Circular polymerase extension cloning (CPEC) was used often and is referenced throughout this chapter (Quan & Tian, 2009).

6.1.2 Protein Expression Protocols

Plasmids encoding for protein constructs were transformed into BL21(DE3) gold cells for expression using electroporation. 5 mL overnight starter cultures were grown from a single colony, added to 500 mL of Lysogeny broth (LB) or YT with 100 μg/mL of the appropriate antibiotic, and grown with shaking at 37 °C. Protein expression was induced with 1 mM isopropyl-β-thiogalactosidase (IPTG) when the cultures reached an optical density at 600 nm (OD_{600}) of ~0.6-0.8. Cells were shaken for an additional 16 hours at 30 °C, unless otherwise stated. Cells were harvested by centrifugation.

Protein expression was also achieved using auto-induction media (Studier, 2005). 5 mL overnight starter cultures grown were grown from a single colony, added to 500 mL ZYM-5052 auto-induction media (1% tryptone, 0.5% yeast extract, 25 mM Na_2HPO_4, 25 mM KH_2PO_4, 50 mM NH_4Cl, 5 mM Na_2SO_4, 0.5% glycerol, 0.05% glucose, 0.2% α-lactose, 2 mM $MgSO_4$, and 0.2 x trace elements) along with 100 μg/mL of the appropriate antibiotic, and grown with shaking at 25 °C for 16 hours, unless otherwise stated. Cells were harvested by centrifugation.

6.1.3 Protein Purification Protocols

For hexahistidine-tagged constructs, cells were resuspended and incubated in lysis buffer (50 mM Tris, pH 7.4, 150 mM NaCl, 20 mM imidazole, 1 mM phenylmethylsulfonyl fluoride (PMSF)) on ice for 30 min in the presence of one EDTA-free Complete Protease Inhibitor tablet (Roche) and sonicated at an output level of 7 in alternating pulses of 10s on and 20 s off for a total of 2 min 30 s (Misonix S-3000). Cell debris was removed by centrifugation for 50 min at 30,000 g. Clarified cell lysate was added to a gravity flow column (Bio-Rad Cat. #732-1010) with 5 mL of Ni-NTA resin (Qiagen) equilibrated with lysis buffer, capped at both ends, and allowed to bind for 1 h at 4 °C with rocking. Unbound lysate was eluted, and the resin was washed with 20 mL of Buffer A (50 mM Tris, pH 7.4, 150 mM NaCl, 1 mM phenylmethylsulfonyl fluoride (PMSF)) followed by 100 mL of Buffer B (50 mM Tris, pH 7.4, 300 mM NaCl, 1 mM phenylmethylsulfonyl fluoride (PMSF)). The desired protein was eluted with Buffer B containing 250 mM imidazole. Dialysis (into 50 mM Tris, pH 8.0, 150 mM NaCl, 1 mM

PMSF, 5 mM β-mercaptoethanol (BME) at 4 °C overnight) was used to remove imidazole, unless otherwise stated.

For GST-tagged constructs, cells were resuspended in 20 mL Buffer A in the presence of one EDTA-free Complete Protease Inhibitor tablet (Roche) and sonicated at an output level of 7 in alternating pulses of 10 s on and 20 s off for a total of 2 min 30 s (Misonix S-3000). Cell debris was removed by centrifugation for 50 min at 30,000 g. Clarified cell lysate was added to a gravity flow column (Bio-Rad Cat. #732-1010) with 2 mL of glutathione sepharose 4B resin (GE Healthcare Life Sciences), capped at both ends, and allowed to bind for 1 hour at 4 °C with rocking. Unbound lysate was eluted, and the resin was washed with 10 mL of Buffer A, followed by 10 mL Buffer B, 10 mL of Buffer C (50 mM Tris, pH 7.4, 1 M NaCl, 1 mM PMSF) and finally with 10 mL of Cleavage Buffer (50 mM Tris, pH 7.4, 150 mM NaCl, 1 mM dithiothreitol (DTT)). The desired protein was separated from the GST tag by an on-column cleavage in the same capped gravity flow column in the presence of 20 μL of PreScission protease (GE Healthcare Life Sciences) in 3 mL cleavage buffer, overnight at 4 °C with rocking. Purified proteins were collected as the eluent.

6.2 Methods for Chapter 2

6.2.1 TPR Array Cloning

The genes encoding TPR2A, MMY, and CTPR390 were cloned into pProEx-HTam using circular polymerase extension cloning CPEC. Arrays of six consecutive TPR domains were concatenated in this vector using BamHI and BglII sites. Genes encoding

TPR2A, MMY, and CTPR390 with Gly-Gly-Gly linkers at the C-terminus were cloned into pProEx-HTam. Arrays of three TPR domains connected by Gly$_3$ linkers were also concatenated in this vector using BamHI and BglII sites. The final gene for TPR arrays described here encode for N-terminally hexahistidine-tagged proteins.

6.2.2 SpyCatcher Array Cloning

The SpyCatcher gene was purchased from Addgene and cloned into the pProEx-HTam vector using CPEC. SpyCatcher arrays were constructed by concatamerization of the SpyCatcher gene in this vector using BamHI and BglII sites. We cloned arrays encoding for two, three, and four consecutive SpyCatcher genes with N-terminal hexahistidine tags.

For SpyCatcher-GS arrays, the sequence encoding for a flexible C-terminal linker (GGSGGS) was cloned onto the SpyCatcher gene using CPEC. SpyCatcher-GS arrays were constructed by concatamerization of the SpyCatcher gene in this vector using BamHI and BglII sites. We cloned arrays encoding for two, three, and four consecutive SpyCatcher-GS genes with N-terminal hexahistidine tags.

6.2.3 SpyTag-ELP-Peptide Cloning

The gene encoding SpyTag-ELP-MEEVD was ordered from GenScript and cloned into the pRSFDuet-1 vector using BamHI and HindIII restriction sites. SpyTag-ELP-MEEVF was cloned using QuikChange site-directed mutagenesis according to the manufacturer's protocol (Agilent).

6.2.4 Expression and Purification of TPR Arrays

TPR arrays were expressed in auto-induction media at 25 °C for 16 hours with shaking and purified according to the protocol for hexahistidine-tagged proteins detailed in section 6.1.3. Dialysis (into 50 mM Tris, pH 8.0, 150 mM NaCl, 1 mM PMSF, 5 mM β-mercaptoethanol (BME) at 4 °C overnight) was used to remove imidazole, and proteins were frozen and stored at -80 °C until further use. Concentrations were determined by UV absorbance at 280 nm on an HP8453 UV–vis spectrophotometer (Agilent) using extinction coefficients calculated by the ProtParam tool of ExPASy (web.expasy.org/protparam).

6.2.5 Co-expression and Purification of Telechelic Peptide Cross-linkers

SpyCatcher arrays with either two, three, or four binding domains were co-expressed with SpyTag-ELP-MEEVD or SpyTag-ELP-MEEVF in auto-induction media at 37 °C for 16 hours with shaking. Fully formed cross-linkers were purified according to the protocol for hexahistidine-tagged proteins detailed in section 6.1.3. Cross-linkers were dialyzed into deionized (DI) water to remove imidazole and salt, frozen, and lyophilized to a white powder until further use. Concentrations were determined by UV absorbance at 280 nm on an HP8453 UV–vis spectrophotometer (Agilent) using extinction coefficients calculated by the ProtParam tool of ExPASy (web.expasy.org/protparam).

6.2.6 Formation of TPR Hydrogels

Purified TPR2A arrays connected by Gly$_3$ linkers were thawed and concentrated using Amicon Ultra Centrifugal Filters (Sigma) just before use. Concentrated TPR2A arrays were pipetted directly to a solution of cross-linkers displaying two, three, or four SpyTag-ELP-MEEVD moieties in water. The final protein concentration of the solution was 10 wt/v % and 1:1 ratio by volume. Gelation occurred during an overnight incubation at room temperature. Hydrogel formation was confirmed by the inversion test.

6.3 Methods for Chapter 3

6.3.1 Cloning for SasG Constructs

The gene encoding SpyTag-eGFP-SpyTag was ordered from GenScript and cloned into pProEx-HTAm vector using BamHI and HindIII restriction sites, yielding N-terminally hexahistidine-tagged proteins. NdeI and XmaI sites are immediately 5' and 3' of the GFP sequence. The genes for the G5^1-E-G5^2 domain and the E-G5^2 domain of SasG were ordered from GenScript as well. The G5^1-E-G5^2 domain was cloned into SpyTag-eGFP-SpyTag using the NdeI and XmaI sites, replacing eGFP, resulting in SpyTag-GEG-SpyTag. To make a second, longer SasG domain flanked by SpyTag peptides, the gene encoding for the E-G5^2 domain was cloned onto the end of G5^1-E-G5^2 three consecutive times using AgeI and HindII restriction sites, resulting in a full-length SasG construct of GEGEGEGEG flanked by SpyTag peptides.

6.3.2 Expression and Purification of Proteins for SasG Complexes

SpyCatcher arrays and SpyTag-SasG-SpyTag were expressed at 25 °C in auto-induction media for 16 hours and purified using the general protocol for hexahistidine-tagged proteins outlined in section 6.1.3. Dialysis (into 50 mM Tris, pH 8.0, 150 mM NaCl, 1 mM PMSF, 5 mM β-mercaptoethanol (BME) at 4 °C overnight) was used to remove imidazole.

The hexahistidine tag was removed from SpyTag-SasG-SpyTag by incubation with TEV protease overnight at 4 °C with rocking. A second incubation with 5 mL of Ni-NTA resin for 1 hour at 4 °C with rocking was performed to remove TEV protease, allowing for elution of purified SpyTag-SasG-SpyTag without hexahistidine tags.

Concentrations were determined by UV absorbance at 280 nm on an HP8453 UV–vis spectrophotometer (Agilent) using extinction coefficients calculated by the ProtParam tool of ExPASy (web.expasy.org/protparam). SpyCatcher arrays and SpyTag-SasG-SpyTag were frozen and stored at -80 °C until further use.

6.3.3 SasG Complex Formation and SDS PAGE

SpyTag-GEG-SpyTag was mixed with SpyCatcher arrays of three or four in a molar ratio of 1:1 and left to incubate overnight at room temperature. The resulting mixtures were run on a 10% SDS PAGE gel and stained with Coomassie to evaluate SasG complex formation. Formed complexes were subjected to a second Ni-NTA resin-binding step at 4 °C for one hour and eluted, to remove impurities. If necessary, complexes were purified by gel filtration chromatography using a Superdex 200 10/300 GL column (GE Healthcare).

6.3.4 Cloning of HER2Af-SpyTag and eGFP-SpyTag

The gene encoding HER2Af-SpyTag was ordered from GenScript and cloned into the pRSFDuet-1 vector using NcoI and HindIII restriction sites. The gene encoding for eGFP-SpyTag was also ordered from GenScript and cloned into pProEx-HTam using using BamHI and HindIII restriction sites.

6.3.5 Cloning of Biotinylated SpyCatcher Arrays

For biotinylated SpyCatcher arrays, the sequence encoding for a C-terminal BirA peptide tag (GLNDIFEAQKIEWH) was fused to the end of the last SpyCatcher domain in each SpyCatcher array in pProEx-HTam described in section 6.2.2 using CPEC. We added this sequence onto SpyCatcher arrays encoding for one, two, and three consecutive SpyCatcher domains, resulting in protein constructs of SpyCatcher arrays that are biotinylated at the C-terminus.

6.3.6 Expression and Purification of HER2 Sensing Components

eGFP-SpyTag was expressed at 25 °C in auto-induction media for 16 hours and purified using the general protocol for hexahistidine-tagged proteins outlined in section 6.1.3. The hexahistidine tag was removed by incubation with TEV protease overnight at 4 °C with rocking. A second incubation with 5 mL of Ni-NTA resin for 1 hour at 4 °C with rocking was performed to remove TEV protease, allowing for elution of purified eGFP-SpyTag without hexahistidine tags. Dialysis (into 50 mM Tris, pH 8.0, 150 mM NaCl, 1 mM PMSF, 5 mM β-mercaptoethanol (BME) at 4 °C overnight) was used to remove imidazole. Concentrations were determined by UV absorbance at 280 nm on an HP8453

UV–vis spectrophotometer (Agilent) using extinction coefficients calculated by the ProtParam tool of ExPASy (web.expasy.org/protparam). Purified eGFP-SpyTag was frozen at -80 °C until further use.

HER2Af-SpyTag were co-expressed with biotinylated SpyCatcher arrays of one, two, or three binding domains in auto-induction media at 37 °C for 16 hours with shaking. Fully formed, biotinylated HER2Af-SpyTag/SpyCatcher complexes were purified according to the protocol for hexahistidine-tagged proteins detailed in section 6.1.3. Dialysis (into 50 mM Tris, pH 8.0, 150 mM NaCl, 1 mM PMSF, 5 mM β-mercaptoethanol (BME) at 4 °C overnight) was used to remove imidazole. Concentrations were determined by UV absorbance at 280 nm on an HP8453 UV–vis spectrophotometer (Agilent) using extinction coefficients calculated by the ProtParam tool of ExPASy (web.expasy.org/protparam). Complexes were frozen at -80 °C until further use.

6.3.7 Formation of Biotinylated Complexes

As described in section 6.3.6, biotinylated HER2Af-SpyTag/SpyCatcher complexes were formed in vivo and purified from *E. coli*. Formation of biotinylated eGFP-SpyTag/SpyCatcher complexes was achieved by *in vitro* mixing of eGFP-SpyTag with biotinylated SpyCatcher arrays of one, two, or three binding domains in a 1:1 molar ratio of binding sites overnight at room temperature. Formed complexes were subjected to a second Ni-NTA resin-binding step at 4 °C for one hour and eluted, to remove impurities.

6.3.8 Fluorescence Plate Reader Assay

96-well plates were coated with NeutrAvidin (Thermo Fisher Scientific) for 1 hour at 4 °C and washed three times with PBS-T. Wells were blocked with 5% BSA in TBS-T for an hour 4 °C and then washed three times with PBS-T. 50 μL of 5 μM biotinylated complexes and controls displaying eGFP were added to the wells and incubated at 4 °C for an hour. Wells were washed three times with PBS-T, and 100 μL of PBS was added to each well for plate reader analysis. Fluorescence was analyzed in a plate reader (BioTek) using a filter set appropriate for eGFP (excitation: 488/20.0, emission: 530/20.0). Experiments were performed in triplicate.

6.4 Methods for Chapter 4

6.4.1 Cloning of BslA and Fluorescent Fusion Proteins

The gene encoding $BslA_{29-176}$ was purchased from GenScript and cloned into the pGEX-6P-1 vector using the BamHI and XhoI restriction sites, to give a GST-fusion protein. There is a PreScission protease cleavage site between GST and BslA, which allows for removal of the GST tag during protein purification. The sequence encoding the SpyTag peptide (AHIVMVDAYKPTK) was cloned onto either the 5' or 3' end of the BslA gene, to create a fusion protein that includes a 6 amino acid linker (GGSGGS) between the peptide and BslA. The gene for SpyCatcher was obtained from Addgene (pDEST14-SpyCatcher, Plasmid #35044). The SpyCatcher gene was added to the 3' end of a gene encoding eGFP and both genes were cloned into pPROEX-HTam with a 6

amino acid linker (GGSGGS) between eGFP and SpyCatcher. All cloning was performed using CPEC.

6.4.2 Protein Expression and Purification

BslA proteins were expressed in auto-induction media at 25 °C for 16 hours with shaking and purified according to the GST-tagged purification protocol outlined in section 6.1.3. SpyCatcher-eGFP was expressed in auto-induction media at 25 °C for 16 hours with shaking and purified according to the hexahistidine-tagged purification protocol outlined in section 6.1.3. Dialysis (into 50 mM Tris-HCl, pH 8.0, 150 mM NaCl at 4 °C) was used to remove imidazole. Concentrations were determined by UV absorbance at 280 nm on an HP8453 UV–vis spectrophotometer (Agilent) using extinction coefficients calculated by the ProtParam tool of ExPASy (web.expasy.org/protparam). Proteins were frozen at -80 °C until further use.

6.4.3 Surface Pressure-Area Isotherms

The surface pressure-area isotherms were obtained using a Langmuir trough (KN2002, KSV Instrument Ltd, Finland). Two symmetric Teflon barriers were controlled by the KSV Nima software. The surface pressure was measured using a Langmuir-Wilhelmy balance as a function of mean molecular area, which was calculated by the software as the area between two barriers divided by the total numbers of molecules added to the system. The BslA solution was carefully spread onto the air-water interface and was allowed to equilibrate for 10 min, followed by compression using the barrier at a constant speed of 20 mm/min. For the surface pressure-area isotherms, the surface

pressure was recorded as a function of mean molecular area until the collapse point was reached.

6.4.4 SFG Spectrometer and Spectral Analyses

The SFG spectrometer used in this study was described in detail previously (Velarde & Wang, 2013; Velarde et al., 2011). Briefly, the instrument consists of two Ti:Sapphire lasers. One (Coherent, Palo Alto, CA) generated 40 fs wide pulses at 800 nm with a frequency of 1 kHz and was used as the seed beam. The seed beam was then amplified and used to pump an OPERA-Solo optical parametric amplifier (OPA) to yield IR pulses at ~30 μJ/pulse at the C-H stretch region (2800-3000 cm^{-1}). The other laser (Coherent, Palo Alto, CA) generated visible 800 nm 100 ps wide pulses at 1 kHz and was amplified to yield a final energy of ~60 μJ/pulse. The two laser systems were electronically synchronized using Synchorlock-AP (Coherent, Palo Alto, CA) to an estimated jitter less than 200 fs. Both the IR and the visible beams overlapped temporally and spatially at the air-water interface to measure sample SFG response with the IR incident angle of 55° and the visible incident angle of 45° relative to the surface normal. Reflective SFG responses were spectrally dispersed using a monochrometer (Andor Technology, Belfast, NIR, Shamrock 750 mm, 1200 lines/mm grating) and measured using a CCD camera (Andor Technology, Newton 971P, back-illuminated). All reported SFG spectra were obtained using ssp polarization configuration, that is s-polarized SFG, s-polarized visible, and p-polarized infrared. S-polarized means the lights were linearly polarized along the direction that is perpendicular to the plane of light propagation, while

p-polarized means the lights were linearly polarized along the direction that is within the plane of light propagation.

The obtained SFG spectra were fitted using the equation 1.

$$\chi^{(2)} = \chi^{(2)}_{NR} + \sum_q \chi^{(2)}_q = \chi^{(2)}_{NR} + \sum_q \frac{A_q}{\omega_{IR} - \omega_q + i\Gamma_q} \quad \text{(eq. 1)},$$

where $\chi^{(2)}$ is the experimentally measured second-order susceptibility of an interface consisting of a non-resonant term, $\chi^{(2)}_{NR}$, and a sum of vibrationally resonant terms, $\chi^{(2)}_q$; and A_q is the amplitude, Γ_q is the damping factor, ω_q is the resonant frequency of the qth vibrational mode, and ω_{IR} is the frequency of the incident IR beam.

6.4.5 Microfluidic Device Fabrication

A patterned silicon master mold was fabricated using standard photolithography methods(Xia, 1998). A polydimethylsiloxane (PDMS) pre-polymer and curing agent (Sylgard 184, Dow Corning) were mixed at a 10:1 ratio, by weight, and the mixture degassed to remove bubbles. Before the mixture was poured onto the master mold, the master mold was exposed to octadecyltrichlorosilane (Sigma-Aldrich) in a closed container for 4 hours to prevent sticking of the PDMS to the master and allow easier peeling of the PDMS from the master after curing. Next, the PDMS mixture was cured at 90°C for 2.5 hours. To form the microfluidic channels, the PDMS replica and a glass slide were exposed to oxygen plasma for 30 seconds and bonded together. The height of the microfluidic device is 90 μm and the width of injection lines for the outer and inner phases are 164 μm and 190 μm, respectively. The flow focusing junction is 51 μm wide.

6.4.6 Surface Treatment

The microfluidic devices were rendered hydrophilic by thermal immobilization of polyvinyl alcohol (PVA) onto the PDMS surface. Briefly, the microfluidic device was filled with 1 percent by weight PVA solution in water (87–90% hydrolyzed, molecular weight 30,000–70,000, Sigma-Aldrich) and incubated for 20 min at room temperature. Then, vacuum was applied to remove the PVA solution and the device is baked at 120 °C for 2 hours to thermally immobilize the PVA onto the PDMS surface.

6.4.7 Microcapsule Fabrication

Microcapsules were formed with mineral oil. BslA was flowed in the inner phase at 4 µL/min and mineral oil was flowed in the inner phase at 52 µL/min.

6.4.8 Decoration of Microcapsules with eGFP

To decorate microcapsules, wt BslA and C-terminally SpyTagged BslA microcapsules were pipetted into separate solutions of 200 μL of ~25 μM eGFP-SpyCatcher protein and incubated for 10 min at room temperature. The eGFP-SpyCatcher solution was removed by pipetting, and the capsules were washed with 600 μL of DI water. Microcapsules were imaged after incubation and each wash step using fluorescent confocal microscopy. The microscope used was an inverted Nikon Eclipse Ti-S microscope and images were taken using a GFP filter cube set (49002-ET-EGFP (FITC/Cy2) (Chroma Technology Corp)), a Nikon E Plan 40X, 0.65 NA air objective, and an Andor Zyla 5.5 sCMOS camera.

6.5 Methods for Chapter 5

6.5.1 Cloning of BslA and Fluoresecent Fusion Proteins

The gene encoding BslA$_{29-176}$ was purchased from GenScript and cloned into the pGEX-6P-1 vector (Amersham) using the BamHI and XhoI restriction sites, yielding an N-terminally tagged GST-fusion protein. A PreScission protease cleavage site between GST and BslA allows for removal of the GST tag during protein purification. The sequences encoding the SpyTag peptide and the SnoopTag peptide or were cloned onto the 3' end of the BslA gene using CPEC to create a fusion protein including a 6 amino acid linker (GGSGGS) between the peptide and BslA. The same method was employed to attach SpyCatcher and SnoopCatcher to the C-terminus of BslA, also attached by a flexible linker (GGSGGS).

To make fluorescent fusion proteins, the SpyCatcher and SnoopTag genes were added to the 3' ends of the gene encoding eGFP, SnoopCatcher was added to mCherry, in the same manner. All genes were cloned into pPROEX HTa with a 6 amino acid linker (GGSGGS) between eGFP and SpyCatcher/SnoopTag and between mCherry and SnoopCatcher. All cloning was performed using CPEC. The genes for SpyCatcher (pDEST14-SpyCatcher, Plasmid #35044) and SnoopCatcher (pET28a-SnoopCatcher, Plasmid #72322) were obtained from Addgene.

6.5.2 Protein Expression and Purification

BslA and BslA fusion proteins were expressed in auto-induction media at 25 °C for 16 hours with shaking and purified according to the GST-tagged purification protocol

outlined in section 6.1.3. Purified BslA proteins were collected as the eluant, and proteins were dialyzed into DI water. Wt BslA, BslA-SpyTag, and BslA-SnoopTag often appear opaque directly after elution from the column. If such cloudiness is observed, the protein should be left undisturbed at room temperature for 1-2 h to allow protein to dissolve back into solution before dialysis. Proteins were frozen and stored at -80 °C until further use.

eGFP-SpyCatcher and mCherry-SnoopCatcher were expressed in auto-induction media at 25 °C for 16 hours with shaking and purified according to the hexahistidine-tagged purification protocol outlined in section 6.1.3. Dialysis (into 50 mM Tris, pH 8.0, 150 mM NaCl, 1 mM PMSF, 5 mM beta-mercaptethanol (BME) at 4 °C overnight) was used to remove imidazole. Proteins were frozen and stored at -80 °C until further use.

6.5.3 Determination of Protein Concentration

Concentrations of purified fluorescent fusion proteins were determined on an HP8453 UV–vis spectrophotometer (Agilent). Extinction coefficients for eGFP-SpyCatcher and mCherry-SnoopCatcher were determined by the ProtParam tool of ExPASy (web.expasy.org/protparam).

Absorption spectra of wt BslA and BslA fusion proteins exhibited high absorbance values at 260 nm compared to absorbance values at 280 nm (data not shown), indicating nucleic acid contamination. This was not alleviated by the addition of nucleases, high salt washes during the purification process, or size exclusion chromatography post-purification. Therefore, to obtain more accurate measures of protein concentration, purified proteins in DI water were sent for quantitative amino acid analysis (SPARC BioCentre, The Hospital for Sick Children, Toronto, ON, Canada). Protein

concentration of these same samples was determined by Bradford assay (Bio-Rad, manufacturer's microtiter plate protocol). The concentrations of stock solutions were calculated from the average values of the Leu and Ile from the amino acid analysis.

6.5.4 Stamp Preparation

Cylindrical PDMS micropillars (20 μm in diameter and 40 μm in height) were fabricated by replica molding using micropatterned silicon molds. In a typical mold preparation process, first the native oxide layer of a P-type <100> silicon wafer (University Wafer) was removed by submerging the wafer in a hydrofluoric acid solution (Buffered Oxide Etch (5:1) from Avantor Performance Materials, LLC) for 30 s. After rinsing the treated silicon wafer with DI water and drying under a flow of nitrogen, AZ-9245 photoresist (Microchemicals GmbH) was spin-coated on the wafer at 4,000 rpm for 1 min using WS-400-6NPP Spin Coater (Laurell Technologies Corporation). The coated photoresist was subsequently cured at 110 °C for 2 min on a hot plate. Using a chromium photomask and with EVG 620 Mask Aligner (Electronic Visions Group), the desired pattern was then transferred on the cured photoresist. Next, the patterned photoresist was developed for 2 min in AZ 400K Developer (Microchemicals GmbH), and subsequently rinsed with DI water and dried under a flow of nitrogen. Etching the wafer was carried out by deep reactive-ion etching on Oxford Plasmalab 100 Reactive Ion Etching System (Oxford Instruments) via the conventional Bosch process. After etching, the photoresist was removed from the wafer with oxygen plasma using AutoGlow 200 plasma system (Glow Research).

In preparation for molding PDMS micropillars and to facilitate the release of the polymer from the mold, each silicon mold was coated with a self-assembled monolayer of trichloro(1H,1H,2H,2H-perfluorooctyl)silane (FOTS; Sigma). FOTS coating was carried out under vacuum at 110 °C for 1 h in presence of 200 μL of FOTS and 500-600 μL of DI water. Physically adsorbed FOTS molecules were removed from a FOTS-coated wafer by ultrasonication of the wafer in pure chloroform (ReagentPlus, ≥ 99.8%, Sigma) at 40 kHz for 10 min, using Branson B5510 Ultrasonic Cleaner (Emerson Industrial Automation) (Izadi et al., 2016).

PDMS micropillar stamps were fabricated from a two-part Sylgard 184 Silicone Elastomer Kit (Dow Corning). Degassed PDMS prepolymer with base to catalyst weight ratio of 10:1 was poured over a FOTS-coated silicon mold and within a polytetrafluoroethylene (PTFE, Teflon) spacer used to adjust the thickness of the polymer backing layer. After another degassing step for approximately 30 min, the PDMS mixture was cured at 90 °C for 2 h. Finally, the polymer and the mold were cooled down to room temperature and then the cured PDMS was gently peeled off from the mold (Izadi et al., 2016).

6.5.5 Microcontact Printing

Purchased plain pre-cleaned glass microscope slides (ThermoScientific) were cleaned by sonication in 100% acetone, followed by 100% methanol, followed by 100% isopropanol, each for 20 min at 50 W (Sonicor S-50). Slides were dried with nitrogen gas. Micropillars of PDMS stamps were inked by pipetting 50 μL of a solution of 1% w/v of FOTS (Sigma) dissolved in 100% ethanol directly onto the stamp, followed by incubation

at room temperature for 5 min. Excess solution was wicked away with a tissue and stamps were dried with nitrogen gas. Inked, dried stamps were placed in the center of cleaned glass slides with a 20 g weight on top for 1 min at room temperature, allowing transfer of the FOTS pattern from the micropillars to the slide. Stamps were removed, and the slides were not further rinsed or dried.

6.5.6 Surface Pressure-Area Isotherms

The surface pressure-area isotherms were obtained using a Langmuir trough (KN2002, KSV Instrument Ltd, Finland). Two symmetric Teflon barriers were controlled by the KSV Nima software. The surface pressure was measured using a Langmuir-Wilhelmy balance as a function of mean molecular area, which was calculated by the software as the area between two barriers divided by the total number of molecules added to the system. Stocks of wt BslA, BslA-SpyTag, and BslA SnoopTag BslA solution were prepared at equal concentrations and mixed at volumetric ratios of 1 BslA-SpyTag/SnoopTag to 3 wild type to form the 25% BslA-SpyTag/75% wt BslA and 25% BslA-SnoopTag/75% wt BslA solutions shown (Figure 2). For the 100% wt BslA experiments, protein was injected directly from the wild type stock solution (Figure 2). Protein solutions were carefully injected just below the surface, allowed to diffuse and equilibrate for 2 min, followed by compression using the barrier at a constant speed of 20 mm/min. For the surface pressure-area isotherms, the surface pressure was recorded as a function of mean molecular area until the collapse point was reached.

6.5.7 Langmuir-Schaefer deposition of protein

Wt BslA and BslA protein fusion monolayers were deposited onto patterned slides using the Langmuir-Schaefer method with the aid of a vacuum pump based horizontal dipping clamp (KN 0006, Biolin Scientific). Langmuir films were transferred to patterned glass slides at a surface pressure of 23 mN/m. Slides were lowered and raised at a speed of 1 mm/min, making contact with the surface for 1.5 min. Patterned slides with deposited BslA monolayers were stored in water until further use.

6.5.8 Fluorescence microscopy

Patterned slides with BslA monolayers were wicked dry with a tissue and incubated with 500 μL of 20-50 μM stocks of eGFP-SpyCatcher or mCherry-SnoopCatcher for 10 min at room temperature. Slides were then rinsed with 25 mL DI water, wicked dry with a tissue, and subsequently imaged using fluorescence microscopy. Negative control experiments on patterned slides with no protein deposition (Figure S2) were prepared and imaged in the same way. All images were obtained using an Axioskop epifluorescence microscope (Carl Zeiss, Thornwood, NY) with a 10X objective (1.4 NA), an AxioCam MRm camera (Carl Zeiss), and AxioVision software. Images were cropped and analyzed using ImageJ software (NIH). Pixel intensity values were computed using the ImageJ line tool drawn across four spots. Fluorescence intensity plots were normalized to the background fluorescence of the glass slide, and a consistent scalar factor was added to increase readability.

References

Alam, M. K. G., C.; Hill, W.; El-Sayed, A.; Fonge, H.; Barreto, K. (2017). Synthetic modular antibody construction using the SpyTag/SpyCatcher protein-ligase system. *Chembiochem, 18*(22), 2217-2221.

Banwell, E. F., Abelardo, E. S., Adams, D. J., Birchall, M. A., Corrigan, A., Donald, A. M., . . . Woolfson, D. N. (2009). Rational design and application of responsive alpha-helical peptide hydrogels. *Nat Mater, 8*(7), 596-600. doi: 10.1038/nmat2479

Boller, T. M., C.; Menzler, S.;. (2002). EUPERGIT oxirane acrylic beads: how to make the enzyme fit for biocatalysts. *Org Process Res Develop, 6*, 509-519.

Bromley, K. M., & MacPhee, C. E. (2017). BslA-stabilized emulsion droplets with designed microstructure. *Interface Focus, 7*(4), 20160124. doi: 10.1098/rsfs.2016.0124

Bromley, K. M., Morris, R. J., Hobley, L., Brandani, G., Gillespie, R. M., McCluskey, M., . . . MacPhee, C. E. (2015). Interfacial self-assembly of a bacterial hydrophobin. *Proc Natl Acad Sci U S A, 112*(17), 5419-5424. doi: 10.1073/pnas.1419016112

Brune, K. D., Buldun, C. M., Li, Y., Taylor, I. J., Brod, F., Biswas, S., & Howarth, M. (2017). Dual Plug-and-Display Synthetic Assembly Using Orthogonal Reactive Proteins for Twin Antigen Immunization. *Bioconjug Chem, 28*(5), 1544-1551. doi: 10.1021/acs.bioconjchem.7b00174

Bryjak, J. K., B.N. (1998). Immobilisation of trypsin on acrylic copolymers process. *Biochemistry, 33*(409-417).

Casanova, F., & Santos, L. (2016). Encapsulation of cosmetic active ingredients for topical application--a review. *J Microencapsul, 33*(1), 1-17. doi: 10.3109/02652048.2015.1115900

Cortajarena, A. L., Wang, J., & Regan, L. (2010). Crystal structure of a designed tetratricopeptide repeat module in complex with its peptide ligand. *FEBS J, 277*(4), 1058-1066. doi: 10.1111/j.1742-4658.2009.07549.x

D'Andrea, L. D., & Regan, L. (2003). TPR proteins: the versatile helix. *Trends Biochem Sci, 28*(12), 655-662. doi: 10.1016/j.tibs.2003.10.007

Eigenbrot, C., Ultsch, M., Dubnovitsky, A., Abrahmsen, L., & Hard, T. (2010). Structural basis for high-affinity HER2 receptor binding by an engineered protein. *Proc Natl Acad Sci U S A, 107*(34), 15039-15044. doi: 10.1073/pnas.1005025107

Eisenthal, K. B. (1996). Liquid interfaces probed by second-harmonic and sum-frequency spectroscopy. *Chem Rev, 96*(4), 1343-1360.

Fairhead, M., & Howarth, M. (2015). Site-specific biotinylation of purified proteins using BirA. *Methods Mol Biol, 1266*, 171-184. doi: 10.1007/978-1-4939-2272-7_12

Ferrigno, P. K. (2016). Non-antibody protein-based biosensors. *Essays Biochem, 60*(1), 19-25.

Fletcher, J. M., Harniman, R. L., Barnes, F. R., Boyle, A. L., Collins, A., Mantell, J., . . . Woolfson, D. N. (2013). Self-assembling cages from coiled-coil peptide modules. *Science, 340*(6132), 595-599. doi: 10.1126/science.1233936

Gao, X., Lyu, S., & Li, H. (2017). Decorating a Blank Slate Protein Hydrogel: A General and Robust Approach for Functionalizing Protein Hydrogels. *Biomacromolecules, 18*(11), 3726-3732. doi: 10.1021/acs.biomac.7b01369

Georges, P. C., Miller, W. J., Meaney, D. F., Sawyer, E. S., & Janmey, P. A. (2006). Matrices with compliance comparable to that of brain tissue select neuronal over glial growth in mixed cortical cultures. *Biophys J, 90*(8), 3012-3018. doi: 10.1529/biophysj.105.073114

Glassman, M. J., Chan, J., & Olsen, B. D. (2013). Reinforcement of Shear Thinning Protein Hydrogels by Responsive Block Copolymer Self-Assembly. *Adv Funct Mater, 23*(9), 1182-1193. doi: 10.1002/adfm.201202034

Gradisar, H., Bozic, S., Doles, T., Vengust, D., Hafner-Bratkovic, I., Mertelj, A., . . . Jerala, R. (2013). Design of a single-chain polypeptide tetrahedron assembled

from coiled-coil segments. *Nat Chem Biol, 9*(6), 362-366. doi: 10.1038/nchembio.1248

Grove, T. Z., Forster, J., Pimienta, G., Dufresne, E., & Regan, L. (2012). A modular approach to the design of protein-based smart gels. *Biopolymers, 97*(7), 508-517. doi: 10.1002/bip.22033

Grove, T. Z., Osuji, C. O., Forster, J. D., Dufresne, E. R., & Regan, L. (2010). Stimuli-responsive smart gels realized via modular protein design. *J Am Chem Soc, 132*(40), 14024-14026. doi: 10.1021/ja106619w

Gruszka, D. T., Wojdyla, J. A., Bingham, R. J., Turkenburg, J. P., Manfield, I. W., Steward, A., . . . Potts, J. R. (2012). Staphylococcal biofilm-forming protein has a contiguous rod-like structure. *Proc Natl Acad Sci U S A, 109*(17), E1011-1018. doi: 10.1073/pnas.1119456109

Hobley, L., Ostrowski, A., Rao, F. V., Bromley, K. M., Porter, M., Prescott, A. R., . . . Stanley-Wall, N. R. (2013). BslA is a self-assembling bacterial hydrophobin that coats the Bacillus subtilis biofilm. *Proc Natl Acad Sci U S A, 110*(33), 13600-13605. doi: 10.1073/pnas.1306390110

Howarth, M. (2017). Smart superglue in streptococci? The proof is in the pulling. *J Biol Chem, 292*(21), 8998-8999. doi: 10.1074/jbc.H117.777466

Izadi, H., Dogra, N., Perreault, F., Schwarz, C., Simon, S., & Vanderlick, T. K. (2016). Removal of Particulate Contamination from Solid Surfaces Using Polymeric Micropillars. *ACS Appl Mater Interfaces, 8*(26), 16967-16978. doi: 10.1021/acsami.5b09154

J. Carlos Rodriguez-Cabello, L. M., Matilde Alonso, F. Javier Arias, Ana M. Testera. (2009). "Recombinamers" as advanced materials for the post-oil age. *Polymer, 50*(22), 5159-5169. doi: https://doi.org/10.1016/j.polymer.2009.08.032

Jackrel, M. E., Valverde, R., & Regan, L. (2009). Redesign of a protein-peptide interaction: characterization and applications. *Protein Sci, 18*(4), 762-774. doi: 10.1002/pro.75

Janssen, M. H., van Langen, L. M., Pereira, S. R., van Rantwijk, F., & Sheldon, R. A. (2002). Evaluation of the performance of immobilized penicillin G acylase using active-site titration. *Biotechnol Bioeng, 78*(4), 425-432.

Kahn, A. A. A., M.A. (2010). Recent advances and applications of immobilized enzymes technologies: a review. *Research Journal of Biological Sciences, 5*(8), 565-575.

Kane, R. S., Takayama, S., Ostuni, E., Ingber, D. E., & Whitesides, G. M. (1999). Patterning proteins and cells using soft lithography. *Biomaterials, 20*(23-24), 2363-2376.

Kaufman, G., Liu, W., Williams, D. M., Choo, Y., Gopinadhan, M., Samudrala, N., . . . Osuji, C. O. (2017). Flat Drops, Elastic Sheets, and Microcapsules by Interfacial Assembly of a Bacterial Biofilm Protein, BslA. *Langmuir, 33*(47), 13590-13597. doi: 10.1021/acs.langmuir.7b03226

Lai, Y. T., Reading, E., Hura, G. L., Tsai, K. L., Laganowsky, A., Asturias, F. J., . . . Yeates, T. O. (2014). Structure of a designed protein cage that self-assembles into a highly porous cube. *Nat Chem, 6*(12), 1065-1071. doi: 10.1038/nchem.2107

Liu, W., Li, S., Wang, Z., Yan, E. C. Y., & Leblanc, R. M. (2017). Characterization of Surface-Active Biofilm Protein BslA in Self-Assembling Langmuir Monolayer at the Air-Water Interface. *Langmuir, 33*(30), 7548-7555. doi: 10.1021/acs.langmuir.7b01739

Liu, Z., Robinson, J. T., Sun, X., & Dai, H. (2008). PEGylated nanographene oxide for delivery of water-insoluble cancer drugs. *J Am Chem Soc, 130*(33), 10876-10877. doi: 10.1021/ja803688x

Liu, Z., Zhou, H., Wang, W., Tan, W., Fu, Y. X., & Zhu, M. (2014). A novel method for synthetic vaccine construction based on protein assembly. *Sci Rep, 4*, 7266. doi: 10.1038/srep07266

Lofblom, J., Feldwisch, J., Tolmachev, V., Carlsson, J., Stahl, S., & Frejd, F. Y. (2010). Affibody molecules: engineered proteins for therapeutic, diagnostic and biotechnological applications. *FEBS Lett, 584*(12), 2670-2680. doi: 10.1016/j.febslet.2010.04.014

Mohamad, N. R., Marzuki, N. H., Buang, N. A., Huyop, F., & Wahab, R. A. (2015). An overview of technologies for immobilization of enzymes and surface analysis techniques for immobilized enzymes. *Biotechnol Biotechnol Equip, 29*(2), 205-220. doi: 10.1080/13102818.2015.1008192

Morris, R. J., Bromley, K. M., Stanley-Wall, N., & MacPhee, C. E. (2016). A phenomenological description of BslA assemblies across multiple length scales. *Philos Trans A Math Phys Eng Sci, 374*(2072). doi: 10.1098/rsta.2015.0131

Olsen, B. D., Kornfield, J. A., & Tirrell, D. A. (2010). Yielding Behavior in Injectable Hydrogels from Telechelic Proteins. *Macromolecules, 43*(21), 9094-9099. doi: 10.1021/ma101434a

Quan, J., & Tian, J. (2009). Circular polymerase extension cloning of complex gene libraries and pathways. *PLoS One, 4*(7), e6441. doi: 10.1371/journal.pone.0006441

Richmond, G. L. (2002). Molecular bonding and interactions at aqueous surfaces as probed by vibrational sum frequency spectroscopy. *Chem Rev, 102*(8), 2693-2724.

Ross, J. S. (2009). Breast cancer biomarkers and HER2 testing after 10 years of anti-HER2 therapy. *Drug News Perspect, 22*(2), 93-106. doi: 10.1358/dnp.2009.22.2.1334452

Sacca, B., & Niemeyer, C. M. (2012). DNA origami: the art of folding DNA. *Angew Chem Int Ed Engl, 51*(1), 58-66. doi: 10.1002/anie.201105846

Schloss, A. C. Liu, W.; Williams, D.M.; Kaufman, G.; Hendrickson, H.P.; Rudshteyn, B.; Fu, L.; Wang, H.; Batista, V.S.; Osuji, C.O.; Yan, E.C.Y.; Regan, L. (2016). Fabrication of modularly functionalizable microcapsules using protein-based technologies. *ACS Biomaterials Science and Engineering, 2*(11), 1856-1851.

Sharma, S., Byrne, H., & O'Kennedy, R. J. (2016). Antibodies and antibody-derived analytical biosensors. *Essays Biochem, 60*(1), 9-18. doi: 10.1042/EBC20150002

Shen, Y. R. (1989). Surface properties probed by second-harmonic and sum-frequency generation. *Nature, 337*, 519-525.

Sipailiene, A., & Petraityte, S. (2018). Encapsulation of Probiotics: Proper Selection of the Probiotic Strain and the Influence of Encapsulation Technology and Materials on the Viability of Encapsulated Microorganisms. *Probiotics Antimicrob Proteins, 10*(1), 1-10. doi: 10.1007/s12602-017-9347-x

Smith, J. E., Sapsford, K. E., Tan, W., & Ligler, F. S. (2011). Optimization of antibody-conjugated magnetic nanoparticles for target preconcentration and immunoassays. *Anal Biochem, 410*(1), 124-132. doi: 10.1016/j.ab.2010.11.005

Stanley-Wall, N. R., & MacPhee, C. E. (2015). Connecting the dots between bacterial biofilms and ice cream. *Phys Biol, 12*(6), 063001. doi: 10.1088/1478-3975/12/6/063001

Studier, F. W. (2005). Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif, 41*(1), 207-234.

Sun, F., Zhang, W. B., Mahdavi, A., Arnold, F. H., & Tirrell, D. A. (2014). Synthesis of bioactive protein hydrogels by genetically encoded SpyTag-SpyCatcher chemistry. *Proc Natl Acad Sci U S A, 111*(31), 11269-11274. doi: 10.1073/pnas.1401291111

Thrane, S., Janitzek, C. M., Matondo, S., Resende, M., Gustavsson, T., de Jongh, W. A., . . . Sander, A. F. (2016). Bacterial superglue enables easy development of efficient virus-like particle based vaccines. *J Nanobiotechnology, 14*, 30. doi: 10.1186/s12951-016-0181-1

Tischer, W., & Kasche, V. (1999). Immobilized enzymes: crystals or carriers? *Trends Biotechnol, 17*(8), 326-335.

Veggiani, G., Nakamura, T., Brenner, M. D., Gayet, R. V., Yan, J., Robinson, C. V., & Howarth, M. (2016). Programmable polyproteams built using twin peptide superglues. *Proc Natl Acad Sci U S A, 113*(5), 1202-1207. doi: 10.1073/pnas.1519214113

Veggiani, G., Zakeri, B., & Howarth, M. (2014). Superglue from bacteria: unbreakable bridges for protein nanotechnology. *Trends Biotechnol, 32*(10), 506-512. doi: 10.1016/j.tibtech.2014.08.001

Velarde, L., & Wang, H. F. (2013). Capturing inhomogeneous broadening of the -CN stretch vibration in a Langmuir monolayer with high-resolution spectra and ultrafast vibrational dynamics in sum-frequency generation vibrational spectroscopy (SFG-VS). *J Chem Phys, 139*(8), 084204. doi: 10.1063/1.4818996

Velarde, L., Zhang, X. Y., Lu, Z., Joly, A. G., Wang, Z., & Wang, H. F. (2011). Communication: Spectroscopic phase and lineshapes in high-resolution broadband sum frequency vibrational spectroscopy: resolving interfacial inhomogeneities of "identical" molecular groups. *J Chem Phys, 135*(24), 241102. doi: 10.1063/1.3675629

Wang, Z., Morales-Acosta, M. D., Li, S., Liu, W., Kanai, T., Liu, Y., . . . Yan, E. C. (2016). A narrow amide I vibrational band observed by sum frequency generation spectroscopy reveals highly ordered structures of a biofilm protein at the air/water interface. *Chem Commun (Camb), 52*(14), 2956-2959. doi: 10.1039/c5cc05743d

Whang, K., Healy, K. E., Elenz, D. R., Nam, E. K., Tsai, D. C., Thomas, C. H., . . . Sprague, S. M. (1999). Engineering bone regeneration with bioabsorbable scaffolds with novel microarchitecture. *Tissue Eng, 5*(1), 35-51. doi: 10.1089/ten.1999.5.35

Wong, C. Y., Al-Salami, H., & Dass, C. R. (2018). Microparticles, microcapsules and microspheres: A review of recent developments and prospects for oral delivery of insulin. *Int J Pharm, 537*(1-2), 223-244. doi: 10.1016/j.ijpharm.2017.12.036

Xia, Y. W., G.M. (1998). Soft lithography. *Annu Rev Mater Sci, 28*, 153-184.

Yan, E. C., Fu, L., Wang, Z., & Liu, W. (2014). Biological macromolecules at interfaces probed by chiral vibrational sum frequency generation spectroscopy. *Chem Rev, 114*(17), 8471-8498. doi: 10.1021/cr4006044

Yumura, K., Akiba, H., Nagatoishi, S., Kusano-Arai, O., Iwanari, H., Hamakubo, T., & Tsumoto, K. (2017). Use of SpyTag/SpyCatcher to construct bispecific antibodies

that target two epitopes of a single antigen. *J Biochem, 162*(3), 203-210. doi: 10.1093/jb/mvx023

Zakeri, B., Fierer, J. O., Celik, E., Chittock, E. C., Schwarz-Linek, U., Moy, V. T., & Howarth, M. (2012). Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. *Proc Natl Acad Sci U S A, 109*(12), E690-697. doi: 10.1073/pnas.1115485109

Zhang, W. B., Sun, F., Tirrell, D. A., & Arnold, F. H. (2013). Controlling macromolecular topology with genetically encoded SpyTag-SpyCatcher chemistry. *J Am Chem Soc, 135*(37), 13988-13997. doi: 10.1021/ja4076452

Appendix

Chapter 2 Protein Sequences

Construct	Protein Sequence
TPR2A-Gly3 Array	GNHVVLHHHHHHDYDIPTTENLYFQGAMGSKQALKEKELGNDAYKKKDFDTALKHYDKAKELDPTNMTYITNQAAVYFEKGDYNKCRELCEKAIEVGRENREDYRQIAKAYARIGNSYFKEEKYKDAIHFYNKSLAEHRTPDVLKKCQQAEKILKEQGGGRSKQALKEKELGNDAYKKKDFDTALKHYDKAKELDPTNMTYITNQAAVYFEKGDYNKCRELCEKAIEVGRENREDYRQIAKAYARIGNSYFKEEKYKDAIHFYNKSLAEHRTPDVLKKCQHAEKILKEQGGGRSKQALKEKELGNDAYKKKDFDTALKHYDKAKELDPTNMTYITNQAAVYFEKGDYNKCRELCEKAIEVGRENREDYRQIAKAYARIGNSYFKEEKYKDAIHFYNKSLAEHRTPDVLKKCQHAEKILKEQRS
MMY-Gly3 Array	MSYYHHHHHHDYDIPTTENLYFQKQALKEKELGNDAYKKKDFDTALKHYDKAKELDPTNMTYIMNQAAVYFEKGDYNKCRELCEKAIEVGRENREDYRMIAYAYARIGNSYFKEEKYKDAIHFYNKSLAEHRTPKVLKKCQQAEKILKEQGGGRSKQALKEKELGNDAYKKKDFDTALKHYDKAKELDPTNMTYIMNQAAVYFEKGDYNKCRELCEKAIEVGRENREDYRMIAYAYARIGNSYFKEEKYKDAIHFYNKSLAEHRTPKVLKKCQQAEKILKEQGGGRSKQALKEKELGNDAYKKKDFDTALKHYDKAKELDPTNMTYIMNQAAVYFEKGDYNKCRELCEKAIEVGRENREDYRMIAYAYARIGNSYFKEEKYKDAIHFYNKSLAEHRTPKVLKKCQQAEKILKEQRS
CTPR390-Gly3 Array	MSYYHHHHHHDYDIPTTENLYFQGAMGSAEAWKNLGNAYYKQGDYQKAIEYYQKALELDPNNASAWYNLGNAYYKQGDYQKAIEYYQKALELDPNNAKAWYRRGNAYYKQGDYQKAIEDYQKALELDPNNRSGGGMGSAEAWKNLGNAYYKQGDYQKAIEYYQKALELDPNNASAWYNLGNAYYKQGDYQKAIEYYQKALELDPNNAKAWYRRGNAYYKQGDYQKAIEDYQKALELDPNNRSGGGMGSAEAWKNLGNAYYKQGDYQKAIEYYQKALELDPNNASAWYNLGNAYYKQGDYQKAIEYYQKALELDPNNAKAWYRRGNAYYKQGDYQKAIEDYQKALELDPNNRS
CTPR390 6 Array	MSYYHHHHHHDYDIPTTENLYFQGAMGSAEAWKNLGNAYYKQGDYQKAIEYYQKALELDPNNASAWYNLGNAYYKQGDYQKAIEYYQKALELDPNNAKAWYRRGNAYYKQGDYQKAIEDYQKALELDPNNRSMGSAEAWKNLGNAYYKQGDYQKAIEYYQKALELDPNNASAWYNLGNAYYKQGDYQKAIEYYQKALELDPNNAKAWYRRGNAYYKQGDYQKAIEDYQKALELDPNNRSMGSAEAWKNLGNAYYKQGDYQKAIEYYQKALELDPNNASAWYNLGNAYYKQGDYQKAIEYYQKALELDPNNAKAWYRRGNAYYKQGDYQKAIEDYQKALELDPNNRSMGSAEAWKNLGNAYYKQGDYQKAIEYYQKALELDPNNASAWYNLGNAYYKQGDYQKAIEYYQKALELDPNNAKAWYRRGNAYYKQGDYQKAIEDYQKALELDPNNRSMGSAEAWKNLGNAYYKQGDYQKAIEYYQKALELDPNNASAWYNLGNAYYKQGDYQKAIEYYQKALELDPNNAKAWYRRGNAYYKQGDYQKAIEDYQKALELDPNNRS
SpyCatcher	MSYYHHHHHHDYDIPTTENLYFQGAMGSAMVDTLSGLSSEQGQSGDMTIEEDSATHIKFSKRDEDGKELAGATMELRDSSGKTISTWISDGQVKDFYLYPGKYTFVETAAPDGYEVATAITFTVNEQGQVTVNGKATKGDAHIRS
SpyCatcher 2 Array	MSYYHHHHHHDYDIPTTENLYFQGAMGSAMVDTLSGLSSEQGQSGDMTIEEDSATHIKFSKRDEDGKELAGATMELRDSSGKTISTWISDGQVKDFYLYPGKYTFVETAAPDGYEVATAITFTVNEQGQVTVNGKATKGDAHIRSAMVDTLSGLSSEQGQSGDMTIEEDSATHIKFSKRDEDGKELAGATMELRDSSGKTISTWISDGQVKDFYLYPGKYTFVETAAPDGYEVATAITFTVNEQGQVTVNGKATKGDAHIRS
SpyCatcher 3 Array	MSYYHHHHHHDYDIPTTENLYFQGAMGSAMVDTLSGLSSEQGQSGDMTIEEDSATHIKFSKRDEDGKELAGATMELRDSSGKTISTWISDGQVKDFYLYPGKYTFVETAAPDGYEVATAITFTVNEQGQVTVNGKATKGDAHIRSAMVDTLSGLSSEQGQSGDMTIEEDSATHIKFSKRDEDGKELAGATMELRDSSGKTISTWISDGQVKDFYLYPGKYTFVETAAPDGYEVATAITFTVNEQGQVTVNGKATKGDAHIRSAMVDTLSGLSSEQGQSGDMTIEEDSATHIKFSKRDEDGKELAGATMELRDSSGKTISTWISDGQVKDFYLYPGKYTFVETAAPDGYEVATAITFTVNEQGQVTVNGKATKGDAHIRS

SpyCatcher 4 Array	MSYYHHHHHHDYDIPTTENLYFQGAMGSAMVDTLSGLSSEQGQSGDMTIEEDSATHIK FSKRDEDGKELAGATMELRDSSGKTISTWISDGQVKDFYLYPGKYTFVETAAPDGYEVA TAITFTVNEQGQVTVNGKATKGDAHIRSAMVDTLSGLSSEQGQSGDMTIEEDSATHIKF SKRDEDGKELAGATMELRDSSGKTISTWISDGQVKDFYLYPGKYTFVETAAPDGYEVAT AITFTVNEQGQVTVNGKATKGDAHIRSAMVDTLSGLSSEQGQSGDMTIEEDSATHIKFS KRDEDGKELAGATMELRDSSGKTISTWISDGQVKDFYLYPGKYTFVETAAPDGYEVATA ITFTVNEQGQVTVNGKATKGDAHIRSAMVDTLSGLSSEQGQSGDMTIEEDSATHIKFSK RDEDGKELAGATMELRDSSGKTISTWISDGQVKDFYLYPGKYTFVETAAPDGYEVATAI TFTVNEQGQVTVNGKATKGDAHIRS
SpyTag-ELP-MEEVD	MGSAHIVMVDAYKPTKVGVPGVGVPGVGVPGEGVPGVGVPGVGVPGVGVPG VGVPGEGVPGVGVPGVGVPGVGVPGVGVPGEGVPGVGVPGVGVMEEVD
SpyTag-ELP-MEEVF	MGSAHIVMVDAYKPTKVGVPGVGVPGVGVPGEGVPGVGVPGVGVPGVGVPG VGVPGEGVPGVGVPGVGVPGVGVPGVGVPGEGVPGVGVPGVGVMEEVF
SpyCatcher-GS	AMVDTLSGLSSEQGQSGDMTIEEDSATHIKFSKRDEDGKELAGATMELRDSSGKTIST WISDGQVKDFYLYPGKYTFVETAAPDGYEVATAITFTVNEQGQVTVNGKATKGDAHIRS GGSGGS
SpyCatcher-GS 2 Array	MSYYHHHHHHDYDIPTTENLYFQGAMGSAMVDTLSGLSSEQGQSGDMTIEEDSATHIK FSKRDEDGKELAGATMELRDSSGKTISTWISDGQVKDFYLYPGKYTFVETAAPDGYEVA TAITFTVNEQGQVTVNGKATKGDAHIRSGGSGGSAMVDTLSGLSSEQGQSGDMTIEED SATHIKFSKRDEDGKELAGATMELRDSSGKTISTWISDGQVKDFYLYPGKYTFVETAAP DGYEVATAITFTVNEQGQVTVNGKATKGDAHIRSGGSGGS
SpyCatcher-GS 3 Array	MSYYHHHHHHDYDIPTTENLYFQGAMGSAMVDTLSGLSSEQGQSGDMTIEEDSATHIK FSKRDEDGKELAGATMELRDSSGKTISTWISDGQVKDFYLYPGKYTFVETAAPDGYEVA TAITFTVNEQGQVTVNGKATKGDAHIRSGGSGGSAMVDTLSGLSSEQGQSGDMTIEED SATHIKFSKRDEDGKELAGATMELRDSSGKTISTWISDGQVKDFYLYPGKYTFVETAAP DGYEVATAITFTVNEQGQVTVNGKATKGDAHIRSGGSGGSAMVDTLSGLSSEQGQSG DMTIEEDSATHIKFSKRDEDGKELAGATMELRDSSGKTISTWISDGQVKDFYLYPGKYT FVETAAPDGYEVATAITFTVNEQGQVTVNGKATKGDAHIRSGGSGGS
SpyCatcher-GS 4 Array	MSYYHHHHHHDYDIPTTENLYFQGAMGSAMVDTLSGLSSEQGQSGDMTIEEDSATHIK FSKRDEDGKELAGATMELRDSSGKTISTWISDGQVKDFYLYPGKYTFVETAAPDGYEVA TAITFTVNEQGQVTVNGKATKGDAHIRSGGSGGSAMVDTLSGLSSEQGQSGDMTIEED SATHIKFSKRDEDGKELAGATMELRDSSGKTISTWISDGQVKDFYLYPGKYTFVETAAP DGYEVATAITFTVNEQGQVTVNGKATKGDAHIRSGGSGGSAMVDTLSGLSSEQGQSG DMTIEEDSATHIKFSKRDEDGKELAGATMELRDSSGKTISTWISDGQVKDFYLYPGKYT FVETAAPDGYEVATAITFTVNEQGQVTVNGKATKGDAHIRSGGSGGSAMVDTLSGLSS EQGQSGDMTIEEDSATHIKFSKRDEDGKELAGATMELRDSSGKTISTWISDGQVKDFYL YPGKYTFVETAAPDGYEVATAITFTVNEQGQVTVNGKATKGDAHIRSGGSGGS

Chapter 3 Protein Sequences

Construct	Protein Sequence
SpyTag-GEG-SpyTag	AHIVMVDAYKPTKMAPKTITELEKKVEEIPFKKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGVIISKGEPKEEITKDPINELTEYGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIPFEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETIAAHIVMVDAYKPTK
SpyTag-longSasG-SpyTag	AHIVMVDAYKPTKMAPKTITELEKKVEEIPFKKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGVIISKGEPKEEITKDPINELTEYGPETIAPGHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIPFEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETIAHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIPFEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETIAHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIPFEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETIAHRDEFDPKLPTGEKEEVPGKPGIKNPETGDVVRPPVDSVTKYGPVKGDSIVEKEEIPFEKERKFNPDLAPGTEKVTREGQKGEKTITTPTLKNPLTGEIISKGESKEEITKDPINELTEYGPETIAAHIVMVDAYKPTK
SpyCatcher BirA	MSYYHHHHHHDYDIPTTENLYFQGAMGSAMVDTLSGLSSEQGQSGDMTIEEDSATHIKFSKRDEDGKELAGATMELRDSSGKTISTWISDGQVKDFYLYPGKYTFVETAAPDGYEVATAITFTVNEQGQVTVNGKATKGDAHIRSGLNDIFEAQKIEWH
SpyCatcher BirA 2 Array	MSYYHHHHHHDYDIPTTENLYFQGAMGSAMVDTLSGLSSEQGQSGDMTIEEDSATHIKFSKRDEDGKELAGATMELRDSSGKTISTWISDGQVKDFYLYPGKYTFVETAAPDGYEVATAITFTVNEQGQVTVNGKATKGDAHIRSAMVDTLSGLSSEQGQSGDMTIEEDSATHIKFSKRDEDGKELAGATMELRDSSGKTISTWISDGQVKDFYLYPGKYTFVETAAPDGYEVATAITFTVNEQGQVTVNGKATKGDAHIRSGLNDIFEAQKIEWH
SpyCatcher BirA 3 Array	MSYYHHHHHHDYDIPTTENLYFQGAMGSAMVDTLSGLSSEQGQSGDMTIEEDSATHIKFSKRDEDGKELAGATMELRDSSGKTISTWISDGQVKDFYLYPGKYTFVETAAPDGYEVATAITFTVNEQGQVTVNGKATKGDAHIRSAMVDTLSGLSSEQGQSGDMTIEEDSATHIKFSKRDEDGKELAGATMELRDSSGKTISTWISDGQVKDFYLYPGKYTFVETAAPDGYEVATAITFTVNEQGQVTVNGKATKGDAHIRSAMVDTLSGLSSEQGQSGDMTIEEDSATHIKFSKRDEDGKELAGATMELRDSSGKTISTWISDGQVKDFYLYPGKYTFVETAAPDGYEVATAITFTVNEQGQVTVNGKATKGDAHIRSGLNDIFEAQKIEWH
HER2Af-SpyTag	NKEMRNAYWEIALLPNLNNQQKRAFIRSLYDDPSQSANLLAEAKKLNDAQAPKGGGAHIVMVDAYKPTK
eGFP-SpyTag	HHHHHHDYDIPTTENLYFQGAMGSSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLSPDNHYLSTQSKLSKDPNEKRDHMVLLEFVTAAGITHGMDELYKGGSGGSAHIVMVDAYKPT

Chapter 4 Protein Sequences

Construct	Protein Sequence
WT BslA	GPLGSMAESTSTKAHTESTMRTQSTASLFATITGASKTEWSFSDIELT YRPNTLLSLGVMEFTLPSGFTANTKDTLNGNALRTTQILNNGKTVRV PLALDLLGAGEFKLKLNNKTLPAAGTYTFRAENKSLSIGNKFYAEASI DVAKRSTPPTQ
BslA-C-SpyTag	GPLGSMAESTSTKAHTESTMRTQSTASLFATITGASKTEWSFSDIELT YRPNTLLSLGVMEFTLPSGFTANTKDTLNGNALRTTQILNNGKTVRV PLALDLLGAGEFKLKLNNKTLPAAGTYTFRAENKSLSIGNKFYAEASI DVAKRSTPPTQGGSGGSAHIVMVDAYKPTK
BslA-N-SpyTag	GPAHIVMVDAYKPTKGGSGGSMAESTSTKAHTESTMRTQSTASLFA TITGASKTEWSFSDIELTYRPNTLLSLGVMEFTLPSGFTANTKDTLNG NALRTTQILNNGKTVRVPLALDLLGAGEFKLKLNNKTLPAAGTYTFRA ENKSLSIGNKFYAEASIDVAKRSTPPTQ
eGFP-SpyCatcher	HHHHHHDYDIPTTENLYFQGAMGSSKGEELFTGVVPILVELDGDVN GHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLTYGVQC FSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFE GDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKV NFKIRHNIEDGSVQLADHYQQNTPIGDGPVLSPDNHYLSTQSKLSKD PNEKRDHMVLLEFVTAAGITHGMDELYKGGSGGSAMVDTLSGLSSE QGQSGDMTIEEDSATHIKFSKRDEDGKELAGATMELRDSSGKTIST WISDGQVKDFYLYPGKYTFVETAAPDGYEVATAITFTVNEQGQVTVN GKATKGDAHI

Chapter 5 Protein Sequences

Construct	Protein Sequence
WT BslA	GPLGSMAESTSTKAHTESTMRTQSTASLFATITGASKTEWSFSDIEL TYRPNTLLSLGVMEFTLPSGFTANTKDTLNGNALRTTQILNNGKTV RVPLALDLLGAGEFKLKLNNKTLPAAGTYTFRAENKSLSIGNKFYAE ASIDVAKRSTPPTQ
BslA-SpyTag	GPLGSMAESTSTKAHTESTMRTQSTASLFATITGASKTEWSFSDIEL TYRPNTLLSLGVMEFTLPSGFTANTKDTLNGNALRTTQILNNGKTV RVPLALDLLGAGEFKLKLNNKTLPAAGTYTFRAENKSLSIGNKFYAE ASIDVAKRSTPPTQGGSGGSAHIVMVDAYKPTK
BslA-SnoopTag	GPMAESTSTKAHTESTMRTQSTASLFATITGASKTEWSFSDIELTYR PNTLLSLGVMEFTLPSGFTANTKDTLNGNALRTTQILNNGKTVRVP LALDLLGAGEFKLKLNNKTLPAAGTYTFRAENKSLSIGNKFYAEASI DVAKRSTPPTQGGSGGSKLGDIEFIKVNK
BslA-SpyCatcher	HHHHHHDYDIPTTENLYFQGAMGSSKGEELFTGVVPILVELDGDVN GHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLTYGVQ CFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKF EGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGI KVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLSPDNHYLSTQSKL SKDPNEKRDHMVLLEFVTAAGITHGMDELYKGGSGGSAMVDTLSG LSSEQGQSGDMTIEEDSATHIKFSKRDEDGKELAGATMELRDSSG KTISTWISDGQVKDFYLYPGKYTFVETAAPDGYEVATAITFTVNEQG QVTVNGKATKGDAHI
BslA-SnoopCatcher	GPLGSMAESTSTKAHTESTMRTQSTASLFATITGASKTEWSFSDIEL TYRPNTLLSLGVMEFTLPSGFTANTKDTLNGNALRTTQILNNGKTV RVPLALDLLGAGEFKLKLNNKTLPAAGTYTFRAENKSLSIGNKFYAE ASIDVAKRSTPPTQGGSGGSMKPLRGAVFSLQKQHPDYPDIYGAID QNGTYQNVRTGEDGKLTFKNLSDGKYRLFENSEPAGYKPVQNKPI VAFQIVNGEVRDVTSIVPQDIPATYEFTNG KHYITNEPIPPK
eGFP-SpyTag	HHHHHHDYDIPTTENLYFQGAMGSSKGEELFTGVVPILVELDGDVN GHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLTYGVQ CFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKF EGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGI KVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLSPDNHYLSTQSKL SKDPNEKRDHMVLLEFVTAAGITHGMDELYKGGSGGSAHIVMVDA YKPT
eGFP-SnoopTag	HHHHHHDYDIPTTENLYFQGAMGSSKGEELFTGVVPILVELDGDVN GHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLTYGVQ CFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKF EGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGI KVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLSPDNHYLSTQSKL SKDPNEKRDHMVLLEFVTAAGITHGMDELYKGGSGGSKLGDIEFIK VNK
eGFP-SpyCatcher	HHHHHHDYDIPTTENLYFQGAMGSSKGEELFTGVVPILVELDGDVN GHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLTYGVQ CFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKF EGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGI KVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLSPDNHYLSTQSKL SKDPNEKRDHMVLLEFVTAAGITHGMDELYKGGSGGSAMVDTLSG LSSEQGQSGDMTIEEDSATHIKFSKRDEDGKELAGATMELRDSSG KTISTWISDGQVKDFYLYPGKYTFVETAAPDGYEVATAITFTVNEQG QVTVNGKATKGDAHI
mCherry-SnoopCatcher	MLVLHHHHHHDYDIPTTENLYFQGAMGSGIQRPTSTSSTSAAAFES RMASVSKGEEDNMAIIKEFMRFKVHMEGSVNGHEFEIEGEGEGRP YEGTQTAKLKVTKGGPLPFAWDILSPQFMYGSKAYVKHPADIPDYL KLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGEFIYKVKLRGTN FPSDGPVMQKKTMGWEASSERMYPEDGALKGEIKQRLKLKDGGH YDAEVKTTYKAKKPVQLPGAYNVNIKLDITSHNEDYTIVDQYERAE GRHSTGGMDELYKGGSGGSMKPLRGAVFSLQKQHPDYPDIYGAI DQNGTYQNVRTGEDGKLTFKNLSDGKYRLFENSEPAGYKPVQNK PIVAFQIVNGEVRDVTSIVPQDIPATYEFTNGKHYITNEPIPPK