Approaches Towards Development of Novel Fluorogenic Biosensors for Detection of Small Protein Analytes

By

Stephanie Cheung

B.A. Chemistry and Sociology
Grinnell College, 2011

SUBMITTED TO THE DEPARTMENT OF CHEMISTRY IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE IN CHEMISTRY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

FEBRUARY 2014

© 2014 Massachusetts Institute Of Technology.
All rights reserved.
Approaches Towards Development of Novel Fluorogenic Biosensors for Detection of Small Protein Analytes

by

Stephanie Cheung

Submitted to the Department of Chemistry
on February 3, 2014 in Partial Fulfillment of the
Requirements for the Degree of Master of Science
in Chemistry

Abstract

One of the most remarkable ideas in modern biology is that organization of multicellular life is orchestrated by a relatively small repertoire of signaling molecules, which mediate communication between cells. Breakdown of these communication pathways can have profound consequences, leading to a multiplicity of developmental defects and disease states. In particular, the ErbB family of receptors and ligands regulate key cellular processes such as proliferation, differentiation, and apoptosis during embryonic development, cellular homeostasis, and tumorigenesis. The generality of the ErbB-network has made it one of the most well-studied cell signaling systems. However, traditional methods used to study cellular signaling either fail to capture the dynamic nature of signaling networks or to lack the ability to quantify native signaling components. To further our understanding of the signaling processes that govern cell fate and tissue health, novel non-invasive techniques must be developed to quantitatively track native protein analytes in live cells. Herein approaches toward the development of a set novel fluorogenic biosensors capable of detecting native ErbB-ligands are reported.

Using yeast surface display, two libraries of protein-binding scaffolds were engineered to bind selectively to human EGF and human betacellulin with low nM Kd. Bio-orthogonal conjugation of these protein scaffolds to 4-DMN, a solvatochromic fluorophore, afforded a biosensor that exhibit a 3-fold fluorescence increase upon binding human EGF. Efforts to improve the fluorescent signal via alternative labeling strategies are also reported.

Thesis Supervisor: Barbara Imperiali
Title: Class of 1922 Professor of Chemistry and Professor Biology
Acknowledgements

A great number of people have helped me reach the stage of my life I am at today. Without them, none of this would have been possible. I am eternally grateful to you all.

To my parents, who risked it all to give my brother and I better opportunities in America. You raised me to be the best I could be and I cannot be more grateful.

To my brother, who was always the big brother. Thank you for always looking out for me and always being there to share in the craziness that is our family.

To Dylan, thank you for giving me courage when I needed it most. Thank you for always supporting me and believing in me.

To the Grinnell community, thank you for nurturing me into the person I am. College is a critical time in any young person’s life to discover who they are and who they want to be. You provided the best possible environment for me to explore and discover exactly what kind of person I wanted to be. To all my professors at Grinnell, I cannot tell you how much I have learned from you. You showed me what it truly means to be an educator and mentor. To all my friends, you are awesome. Even though we were going to college “in the middle of nowhere,” I had the best time of my life because of all of you.

To my friends at MIT, hang in there! You are all amazing people and I am so glad we became friends. Thank you for being there when I needed someone to talk to or someone to get dinner with – you made graduate school fun.

To all my lab mates, thank you for always being there. I will always remember the times we shared together as we puzzled over a strange result or laughed at the most recent amateur’s mistake. I’m so glad to have worked with all of you and I will miss you all!

And finally, to Barbara, thank you for everything. Thank you for giving me the opportunity to work in such a wonderful environment and with such wonderful people. I have learned so much from you.
Chapter 4: Labeling studies of hEGF-binding Fn3 .............................................. 71

Introduction ........................................................................................................... 72

Results and Discussion ......................................................................................... 73

4-1. Previous studies ............................................................................................... 73

4-2. Labeling of Fn3 5.1.07b .................................................................................. 75

4-3. Investigating the role of the N-terminal tail of Fn3 5.1.07b ......................... 80

4-4. Tail labeling Fn3 5.1.07b ................................................................................ 81

Experimental ........................................................................................................... 87

References ............................................................................................................... 97

Acknowledgements ................................................................................................ 99
List of Figures

Chapter 1

Figure 1-1. General scheme of ErbB-network signaling ....................................................... 13
Figure 1-2. Fluorogenic biosensor design .............................................................................. 17

Chapter 2

Figure 2-1. Domain structure of pro-BTC ............................................................................ 22
Figure 2-2. Secondary structure of mature BTC................................................................. 23
Figure 2-3. Structure of Fn3 domain..................................................................................... 24
Figure 2-4. Electrostatic potential maps of Fn3 5.2.18 and hBTC ......................................... 26
Figure 2-5. Erk1/2 activity of HeLa cell lysates treated with hBTC ........................................ 27
Figure 2-6. Bio-layer interferometry data of Fn3 5.2.18 binding to hBTC ......................... 29
Figure 2-7. Gel filtration chromatography traces of Fn3 5.2.18 and hBTC ......................... 30
Figure 2-8. Gel filtration chromatography traces of Fn3 5.2.13 ........................................ 31
Figure 2-9. SDS-PAGE gel images of SUMO-tagged hBTC expression in E. coli ................ 36

Chapter 3

Figure 3-1. Chemical structures of thiol-reactive 4-DMN derivatives ................................. 55
Figure 3-2. Fluorescence change of 4-DMN labeled Fn3 5.1.07b upon binding to hEGF ...... 56
Figure 3-3. Structure of Fn3 5.2.18 showing positions targeted with
  site-directed mutagenesis .................................................................................................... 58
Figure 3-4. Fluorescence change of 4-DMN labeled Fn3 5.2.18 upon binding hBTC .......... 59
Figure 3-5. SDS-PAGE of TCEP treated Fn3 5.2.18 Cys-mutants ..................................... 65

Chapter 4

Figure 4-1. Illustration of pro-EGF anchored to the cell surface ......................................... 72
Figure 4-2. Fluorescence change of 4-DMN labeled Fn3 5.1.07b Cys-mutants .................... 74
Figure 4-3. Fluorescence change of Fn3 5.1.07b labeled with
  4-DMN, BADAN, and EDANS .......................................................................................... 75
Figure 4-4. Fluorescence change of 4-DMN labeled Fn3 5.1.07b
  mutants H24C, Y26C, and Y29C .................................................................................. 76
Figure 4-5. Fluorescence change of azide-4-DMN labeled Fn3 5.1.07b ............................. 79
Figure 4-6. Fluorescence change of Fn3 5.1.07b N-terminal tail mutants .......................... 80
Figure 4-7. Fluorescence change of 4-DMNA tail-tagged Fn3 5.1.07ts mutants ................. 84
Figure 4-8. Chemical structure of 4-DMNA and 4-DMNA3 ............................................... 85
Figure 4-9. Fluorescence change of 4-DMNA2 and 4-DMNA3
  tail-tagged Fn3 5.1.07ts mutants .................................................................................. 86
Figure 4-10. SDS-PAGE of Met auxotroph expression of Fn3 5.1.07b Met-mutants ........ 91
List of Tables

Chapter 2

Table 2-1. Loops sequence alignments and $K_D$ values of hBTC-binding Fn3 mutants ................................................................. 25

Chapter 3

Table 3-1. $K_D$ values measured for Fn3 5.2.18 Cys-mutants................................................................. 58

Chapter 4

Table 4-1. Loops sequence alignments and $K_D$ values of hEGF-binding Fn3 mutants ................................................................. 73
Table 4-2. Sequences of 4-DMN sortase amenable peptides ................................................................. 82

List of Schemes

Chapter 2

Scheme 2-1. General scheme for sortase mediated ligation ................................................................. 28

Chapter 3

Scheme 3-1. Synthesis of thiol-reactive 4-DMN-V ................................................................. 56

Chapter 4

Scheme 4-1. General scheme for tail-tagging strategy ................................................................. 82
### List of abbreviations

Standard 3-letter and 1-letter codes are used for the natural amino acids. Standard 1-letter codes are used for the nucleotides.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-DMAP</td>
<td>4-N,N-dimethylaminophthalimide</td>
</tr>
<tr>
<td>4-DMN</td>
<td>4-N,N-dimethylamino- 1,8-naphthalimide</td>
</tr>
<tr>
<td>4-DMNA</td>
<td>4-N,N-dimethylamino- 1,8-naphthalimide amino acids</td>
</tr>
<tr>
<td>6-DMN</td>
<td>6-N,N-dimethylamino-2,3-naphthalimide</td>
</tr>
<tr>
<td>Å</td>
<td>angstrom</td>
</tr>
<tr>
<td>a.u.</td>
<td>arbitrary units</td>
</tr>
<tr>
<td>Boc</td>
<td>t-butoxycarbonyl</td>
</tr>
<tr>
<td>BSA</td>
<td>bovine serum albumin</td>
</tr>
<tr>
<td>BTC</td>
<td>betacellulin</td>
</tr>
<tr>
<td>CaM</td>
<td>calmodulin</td>
</tr>
<tr>
<td>CHEF</td>
<td>chelation enhanced fluorescence</td>
</tr>
<tr>
<td>DCM</td>
<td>dichloromethane</td>
</tr>
<tr>
<td>DIPEA</td>
<td>N,N-diisopropylethylamine</td>
</tr>
<tr>
<td>DMEM</td>
<td>Dulbecco's modified Eagle's medium</td>
</tr>
<tr>
<td>DMF</td>
<td>N,N-dimethylformamide</td>
</tr>
<tr>
<td>DMSO</td>
<td>dimethyl sulfoxide</td>
</tr>
<tr>
<td>DTT</td>
<td>dithiothreitol</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>EGF</td>
<td>epidermal growth factor</td>
</tr>
<tr>
<td>EGFR</td>
<td>epidermal growth factor receptor</td>
</tr>
<tr>
<td>EGTA</td>
<td>ethylene glycol tetraacetic acid</td>
</tr>
<tr>
<td>EtBr</td>
<td>ethidium bromide</td>
</tr>
<tr>
<td>FBS</td>
<td>fetal bovine serum</td>
</tr>
<tr>
<td>Fmoc</td>
<td>fluorenlymethyloxycarbonyl chloride</td>
</tr>
<tr>
<td>Fn3</td>
<td>fibronectin type III</td>
</tr>
<tr>
<td>FPLC</td>
<td>fast protein liquid chromatography</td>
</tr>
<tr>
<td>FRET</td>
<td>Förster resonance energy transfer</td>
</tr>
<tr>
<td>FT</td>
<td>flow through</td>
</tr>
<tr>
<td>GFP</td>
<td>green fluorescent protein</td>
</tr>
<tr>
<td>hBTC</td>
<td>human betacellulin</td>
</tr>
<tr>
<td>hEGF</td>
<td>human epidermal growth factor</td>
</tr>
<tr>
<td>HEPES</td>
<td>2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid</td>
</tr>
<tr>
<td>HOBT</td>
<td>hydroxybenzotriazole</td>
</tr>
<tr>
<td>HPLC</td>
<td>high-performance liquid chromatography</td>
</tr>
<tr>
<td>IPTG</td>
<td>isopropyl β-D-1-thiogalactopyranoside</td>
</tr>
<tr>
<td>K_D</td>
<td>dissociation constant</td>
</tr>
<tr>
<td>MALDI-TOF MS</td>
<td>matrix-assisted laser desorption ionization time-of-flight mass spectrometry</td>
</tr>
<tr>
<td>Mtt</td>
<td>N-Methyltrityl</td>
</tr>
<tr>
<td>NEB</td>
<td>New England Biolabs</td>
</tr>
<tr>
<td>NMP</td>
<td>N-Methyl-2-pyrrolidone</td>
</tr>
<tr>
<td>Acronym</td>
<td>Definition</td>
</tr>
<tr>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>NMR</td>
<td>nuclear magnetic resonance</td>
</tr>
<tr>
<td>NTA</td>
<td>nitrilotriacetic acid</td>
</tr>
<tr>
<td>PAGE</td>
<td>polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>PBS</td>
<td>phosphate buffered saline</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>PyBOP</td>
<td>(benzotriazol-1-yl)tripyrrolidinophosphonium hexafluorophosphate</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>sodium dodecyl sulfate polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>Sox</td>
<td>sulfonamido-oxine</td>
</tr>
<tr>
<td>SPPS</td>
<td>solid-phase peptide synthesis</td>
</tr>
<tr>
<td>SUMO</td>
<td>small ubiquitin-related modifier</td>
</tr>
<tr>
<td>tBu</td>
<td>tert-butyl</td>
</tr>
<tr>
<td>TCEP</td>
<td>tris(2-carboxyethyl)phosphine</td>
</tr>
<tr>
<td>TFA</td>
<td>trifluoroacetic acid</td>
</tr>
<tr>
<td>Tris</td>
<td>tris(hydroxymethyl)aminomethane</td>
</tr>
<tr>
<td>UV-vis</td>
<td>ultraviolet-visible</td>
</tr>
<tr>
<td>β-ME</td>
<td>beta-mercaptoethanol</td>
</tr>
<tr>
<td>ε</td>
<td>extinction coefficient</td>
</tr>
<tr>
<td>Φ</td>
<td>quantum yield</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction
I-1. The scope

Cellular signaling is the major form of information exchange and processing by cells. By employing a myriad of signaling molecules and cell surface receptors, cells can sense and respond to their environment via autocrine, paracrine, endocrine, and juxtacrine signaling. Deregulation or mutation of one or more of these signaling components can have profound consequences, leading to a multiplicity of developmental defects and disease states. Identification and characterization of these signaling components have led to numerous breakthroughs including advances in stem cell differentiation and novel drug therapies for the treatment of cancer and diabetes.1,2 Currently, measurements of signaling molecules also assist in disease diagnosis by serving as markers of human and animal health.3,4 As such, techniques to efficiently identify, track, and isolate signaling molecules have been developed and implemented both in the laboratory and the clinic.

This introductory chapter focuses one specific class of signaling molecules, the ErbB receptor-ligand family, and the present technologies used in its study. Particular attention is devoted to the ErbB ligands and methods to detect and quantify native ligand concentrations.

I-2. The ErbB network

The epidermal growth factor receptor (EGFR) family of receptor tyrosine kinases (RTKs) modulates many key aspects of cellular biology including cell proliferation, migration, differentiation and apoptosis.5,6 In mammals, this RTK family is composed of four distinct receptors (EGFR/ErbB-1, ErbB-2, ErbB-3, and ErbB-4) and 13 unique growth factor proteins including EGF and betacellulin. Each receptor is classified as a type Ia transmembrane protein consisting of an extracellular ligand-binding domain, a transmembrane domain, and a
cytoplasmic kinase domain. Ligand binding leads to receptor homo- or heterodimerization, which allows for autophosphorylation of the cytoplasmic tail and subsequent signal transduction to generate specific cellular phenotypes via transcriptional regulation (Figure 1-1). While all four receptors are capable of homo- and heterodimerization, only ErbB-1 and ErbB-4 are capable of signaling via homo-dimerization due to the fact that ErbB-2 has a nonfunctional ligand-binding domain while ErbB-3 contains an inactive kinase domain.

Figure 1-1. Metalloprotease cleavage of membrane-anchored ligand releases the free ligand into the extracellular matrix followed by ligand-activated receptor dimerization.

Ligands of the EGF family contain a consensus sequence (CX_7 CX_4,5 CX_10 CX CX_8 C) involving six Cys residues that form three disulfide bonds, termed the EGF motif. These distinctive disulfide bonds form a unique set of three loops that are involved in receptor binding and recognition. Each of the 13 ligands activates a specific subset of receptors, with each
ligand-receptor pair capable of eliciting a different cellular signal.\textsuperscript{6} Like the ErbB-receptors, each EGF-like ligand is also expressed as a type I transmembrane pre-protein consisting of an EGF motif flanked by an N-terminal extension and C-terminal transmembrane domain.\textsuperscript{6} Some ligands are capable of juxtacrine signaling while existing in their membrane-bound form, however most ErbB-ligands are only activated upon cleavage by cell-surface metalloproteases. The free ligands then take part in a combination of autocrine and paracrine signaling.

The generality of the EGFR signaling network has made it one of the most well-studied cell signaling systems. Disregulation or mutations of the ErbB-receptors or ligands have been associated with developmental defects and cancer.\textsuperscript{9,10} For example, ErbB-2 has been identified as an oncogene and overexpression in malignant breast tissues is associated with poor patient prognosis.\textsuperscript{11,12} As a result, the ErbB-network represents a major pharmacological target in cancer therapies.\textsuperscript{13} A range of therapeutics employing monoclonal antibodies (mAb) or small molecule kinase inhibitors to prevent receptor dimerization or tyrosine phosphorylation have been approved for patient use.\textsuperscript{14} However, drug efficacy differs from patient to patient and resistance often emerges after repeated administration. For example, Trastuzumab, a recombinant humanized mAb directed against ErbB-2, fails to achieve response rates > 35% in patients diagnosed with ErbB-2 overexpressing tumors when administered as a stand-alone drug.\textsuperscript{15} The mechanism of resistance development is not well understood. Several studies indicate drug susceptibility is related to the expression levels of other ErbB-receptors, ErbB-ligands, as well as other RTK receptors.\textsuperscript{15–18}
1-3. Current techniques in signaling molecule detection

Signaling molecules such as the ErbB-ligands serve as the molecular messengers of cell-cell communication. As such, the expression and secretion of these signaling components are tightly regulated and serve as important biomarkers of disease state. Identification and quantification of signaling molecules have been widely used both in the laboratory to study signaling network dynamics and in the clinic to track patient health.\(^{13,19}\)

The dominant technologies used for detecting protein analytes in biological samples include bioassays, immunoassays, and immunohistochemistry.\(^{19}\) These techniques have the advantage of direct detection of native cell signaling molecules. Bioassays are capable of measuring the activity of signaling molecules, allowing measurement of only the active forms of proteins. This is an essential feature since many signaling molecules such as the ErbB-ligands exist in multiple molecular forms. Conversely, bioassays suffer from the disadvantage of low precision and inability to distinguish between signaling molecules that elicit similar biological functions.\(^{19}\) Immunoassays such as ELISAs and flow-cytometric based immunofluorescence assays such as Quantiflow\(^{TM}\) exhibit much higher precision and specificity.\(^{20}\) Multiplexing approaches can also be used to measure up to 64 different proteins per sample. However, these methods lack the ability of bioassays to distinguish between active and inactive forms of proteins. Unlike the two previous techniques, immunohistochemistry offers a unique advantage – compatibility with fluorescence imaging to offer spatial resolution. As such, immunohistochemistry is often used to visualize the spatial distribution of signaling components within cells or tissues. Overall, these detection methods have become the gold standard for detecting and visualizing native proteins in biological samples. However, they all suffer from a similar limitation – poor temporal resolution. All three techniques require either cell fixation or
lysis, making it incompatible with live cell studies. Due to the inherently dynamic nature of cellular signaling, tracking of these signaling components with high spatial and temporal resolution is essential to understanding its biology.

The emergence of methods for fluorescently tagging proteins by genetic fusion to a fluorescent protein such as GFP has provided a general platform for the facile tracking of proteins in live cells and tissues. This method has the advantage of being genetically encodable and compatible with fluorescent microscopy to study proteins with high spatial and temporal resolution. Today, a number of GFP analogs with improved fluorescence properties spanning the visible wavelength allow its use in both FRET studies and multiplexing approaches. However, cases where fusion to GFP has altered protein function and distribution have been reported. Moreover, the use of GFP tags requires genetic manipulations of the biological sample, which cannot be applied to patient samples tested in the clinic.

1-4. Dissertation objectives

Proper exchange of signaling molecules between cells is essential to its ability to sense and respond to environmental cues. As a result, signaling events are inherently dynamic. Moreover, the relative and absolute concentrations of signaling molecules determine developmental fates and cellular homeostasis. Thus the quantification of signaling components is essential to understanding its biology. However, there currently lacks an available technique capable of detecting native signaling molecules in live cells. We propose to develop novel non-invasive tools for the quantitative detection of native protein analytes in live cells. This thesis details the efforts applied toward the development of novel fluorogenic biosensors, consisting of 1) an evolved protein-binding scaffold, Fn3, with tunable substrate specificity and binding
kinetics, and 2) a judiciously placed solvatochromic fluorophore of which the fluorescence emission signals a specific ligand-binding event (Figure 1-2). Two ErbB-ligands, human EGF and human BTC, were chose as the initial ligand targets because of their distinct biological functions and unique protein structures.

**Figure 1-2.** Schematic illustration of fluorogenic biosensor design composed of a protein-binding scaffold conjugated to a solvatochromic fluorophore.

**References:**


Chapter 2

Characterization of hBTC and Fn3-binder
Introduction

Human betacellulin (hBTC), a ligand of the ErbB family, was first isolated from the conditioned medium of mouse pancreatic insulinoma cell lines and identified as a mitogen of pancreatic β-cells.\(^1\) Like all ErbB-ligands, hBTC is expressed as a membrane anchored pre-protein. Pro-hBTC is a glycosylated 178-amino acid protein, consisting of a signal peptide (aa\(^{1-31}\)), an extracellular region (aa\(^{32-118}\)) including the EGF-motif (aa\(^{65-105}\)), a transmembrane region (aa\(^{110-139}\)), and a cytoplasmic tail (aa\(^{140-178}\)) (Figure 2-1).\(^2\)

Proteolytic cleavage of hBTC releases mature hBTC into the extracellular space where it can participate in autocrine or paracrine signaling. Uniquely, hBTC can also signal prior to cleavage through juxtacrine signaling with neighboring cells.\(^3\) Mature hBTC consists of 80-amino acids that include the signature EGF-motif (Figure 2-2). The solution structure of the EGF-motif of hBTC was solved using 2D NMR and confirms a similar β-strand fold between hBTC and other ErbB-ligands such as mouse EGF, human TGF-alpha, and human HRG-alpha\(^{e}\). Interestingly, hBTC is the only ErbB-ligand to contain a fourth disulfide bridge (\(C^{46}-C^{53}\)) in addition to the EGF motif.\(^5\) The biological significance of this extension is unknown but does not seem to affect binding to the ErbB-receptors. In fact, removal of this 30-amino acid extension has no measured effect on its mitogenic activity.\(^6\)
Cellular studies show that binding of hBTC activates both ErbB-1 and ErbB-4 homodimers and all possible receptor hetero-dimers, making it a "pan-ErbB ligand". Human BTC is predominantly expressed in the pancreas and small intestine, with lower expression levels detected in heart, lung, skeletal muscle, kidney, testis, ovary, and colon tissues. Mouse genetic knockouts of BTC show no overt detrimental effects. BTC-null mice are viable, fertile, and show no signs of pancreatic damage. However, overexpressing BTC in mice showed a range of developmental defects including disproportionate growth, reduced life span, pulmonary hemorrhage syndrome, and complex eye pathology.

The most salient biological function of BTC, however, relates to its mitogenic activity in pancreatic β-cells. Treatment with BTC and activin has been shown to convert non-insulin secreting pancreatic cells into insulin secreting cells. BTC-mediated proliferation and differentiation has also been reported in other pancreatic cells lines such as INS-1, MIN-6, and C57BL/6 mouse islets. Importantly, these effects are unique to treatment with BTC and are not seen with EGF treatment. In vivo studies have demonstrated the ability of BTC to reverse the effect of diabetes and promote β-cell regeneration in diabetic mice and rats.
Results and Discussion

2-1. Generation of Fn3 library

The tenth human type 3 fibronectin domain (Fn3) was chosen as the protein-binding scaffold for our biosensor because of its small size (~10 kDa), structural stability, high expression yields, and compatibility with thiol-labeling strategies due to lack of native cysteines. Moreover, the three exposed loops (FG, BC, and DE) on Fn3 resemble the complementarity determining regions (CDRs) of antibodies, which allow them to be engineered to recognize a specific ligand of choice (Figure 2-3). This tunable specificity of Fn3 can be easily achieved through protein engineering on yeast surface display (YSD) to yield Fn3 mutants that bind tightly to the desired ligand with high specificity.

In collaboration with Seymour de Picciotto in the Wittrup lab, a library of high affinity Fn3 mutants was engineered via YSD to bind specifically to hBTC. After five rounds of mutagenesis and selection, seven unique mutants were identified to bind to hBTC with dissociation constants (K_D) ranging from 1-7 nM (as determined using YSD) (Table 2-1). Recombinant hBTC expressed in E. coli purchased from Peprotech, which lacks the native glycosylation of the Asn residue near the N-terminus of hBTC, was used in the selection rounds.
on YSD. To ensure the Fn3 mutants do not recognize the glycosylation site as the epitope, binding of the Fn3 library to non-glycosylated hBTC was compared using YSD to a truncated hBTC mutant that lacks the N-terminal glycosylation region (hBTC<sub>30</sub>). Importantly, the reported activities of recombinant non-glycosylated hBTC and hBTC<sub>30</sub> are comparable to native glycosylated hBTC. Comparison of the Fn3 library binding to hBTC and hBTC<sub>30</sub> showed an analogous ligand affinity, thus illustrating the recognized epitope is not the glycosylation site on hBTC.

### Table 2-1. Generation 5 Fn3 mutants selected through YSD for binding to hBTC with their respective dissociation constants (K<sub>d</sub>) reported. The amino acid sequences of the three loop regions are displayed. Negatively charged residues are highlighted in blue and aromatic residues in red.

<table>
<thead>
<tr>
<th>Fn3</th>
<th>BC Loop</th>
<th>DE Loop</th>
<th>FG Loop</th>
<th>K&lt;sub&gt;d&lt;/sub&gt; (nM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.2.15</td>
<td>Y D W S F A D G S A S</td>
<td>D - - - H D D Y S D</td>
<td>6.8</td>
<td></td>
</tr>
<tr>
<td>5.2.16</td>
<td>Y D W S F A D G S F S</td>
<td>- - D W Y D Y Y S D</td>
<td>6.4</td>
<td></td>
</tr>
<tr>
<td>5.2.20</td>
<td>Y D W S F A D G S V S</td>
<td>- - D W Y D Y Y S D</td>
<td>6.3</td>
<td></td>
</tr>
<tr>
<td>5.2.12</td>
<td>Y D W S F A D G S R S</td>
<td>- - D W D D Y Y S D</td>
<td>1.4</td>
<td></td>
</tr>
<tr>
<td>5.2.18</td>
<td>Y D W S F A D G S F S</td>
<td>D - - - Y D D Y S D</td>
<td>1.2</td>
<td></td>
</tr>
<tr>
<td>5.2.10</td>
<td>Y D W S F A D G S F S</td>
<td>- - D W Y D D Y S E</td>
<td>1.2</td>
<td></td>
</tr>
<tr>
<td>5.2.13</td>
<td>Y D W S F A D G S D L</td>
<td>- - D W Y D Y E S D</td>
<td>1.1</td>
<td></td>
</tr>
</tbody>
</table>

Sequence alignment of the seven Fn3 mutants reveals a completely conserved BC loop, with highly conserved DE and FG loops. The FG loop sequences follow a similar patterning of aromatic-charged-aromatic-charged residues creating a negatively charged interaction surface. Interestingly, a complementary pattern is found on hBTC that creates a highly positively charged surface with interspaced aromatic residues (Figure 2-4). This charge complementarity between Fn3 and hBTC suggests these surfaces may form the protein-protein interaction surface between the two partners, which could serve as a preliminary structural model to aid in the optimization of fluorophore placement.
2-2. Recombinant expression and characterization of hBTC

To allow for a reliable and economical source of protein reagents to be used during the

generation of a biosensor, expression and purification strategies were established for hBTC.
While, the recombinant expression, purification, and characterization of both human BTC and rat
BTC in E. coli culture had been previously reported, the approaches either introduced extra
amino acids into the constructs or suffered from poor yields.\(^9,20,21\) Recently, the small ubiquitin-
like modifier (SUMO) protein was introduced as a new protein solubilization tag to aid with
expression of difficult proteins.\(^22\) Importantly, cleavage of the SUMO-tag affords the native
protein with no additional modifications. Following previously optimized expression protocols
established in our laboratory, hBTC and hBTC\(_{\Delta30}\) were cloned into the pE-SUMO vector and
expressed as a SUMO-fusion construct. Subsequent cleavage of the SUMO-tag afforded the free
proteins. Several E. coli cells lines were screened to optimize the expression of hBTC and
hBTC\(_{\Delta30}\). BL21-RIL (NEB) and Rosetta DE3 (Novagen) cells showed equally high expression,
while SHuffle (NEB) cells displayed low expression yields. Proteins were purified from both the soluble lysate and inclusion body fractions to yield 3-5 mg/L and 7-9 mg/L respectively.

With the recombinant proteins in hand, the activity of recombinant hBTC and hBTC\(_{\Delta 30}\) was assessed. Human BTC is known to bind the ErbB-1 receptor, which activates the MAPK/ERK signaling cascade.\(^{23}\) Thus treatment of cells expressing ErbB-1 with hBTC and hBTC\(_{\Delta 30}\) should elicit increased Erk1/2 activity. Subsequent detection of Erk1/2 activity of the treated cell lysates would indicate the relative activities of the proteins. HeLa cells were chosen for their ease of culture and high expression level of ErbB-1, while the Sox-PNT sensor was used to detect Erk1/2 activity.\(^{24,25}\) The Sox-PNT sensor binds selectively to Erk1/2 via the PNT domain from Ets-1 and senses phosphorylation using Sox-dependent CHEF (chelation enhanced fluorescence).\(^{25}\) Comparison of the activities between recombinant hBTC, hBTC\(_{\Delta 30}\), and commercial hBTC (Peprotech) demonstrated the purified proteins were equally active (Figure 2-5).

![Erk1/2 activity of HeLa cells treated with 10 ng hBTC from varying sources. Data is normalized to HeLa lysates treated with commercial hBTC (Peprotech). hBTC (soluble) = hBTC isolated in soluble lysate. hBTC (IB) = hBTC isolated from inclusion bodies and refolded.]

2.3. In vitro verification of hBTC and Fn3

The binding affinity of Fn3 5.2.18 to hBTC was validated in vitro using bio-layer interferometry (Blitz). Briefly, bio-layer interferometry is an optical analytical method that
utilizes the interference pattern of white light reflected from two surfaces, an immobilized protein layer, and an internal reference layer, to deduce ligand binding and dissociation rates. To immobilize hBTC to the protein layer, a biotin tag was appended to the C-terminus of hBTC via sortase mediated ligation (SML) (Scheme 2-1). Exposure of biotinylated hBTC to a streptavidin coated sensor tip successfully immobilized hBTC for binding measurements to Fn3 5.2.18. Unexpectedly, the association curve of Fn3 5.2.18 to hBTC did not fit to a 1:1 binding curve (Figure 2-6a). One explanation could be the biotinylation of the hBTC C-terminus impedes recognition of the epitope by Fn3. To test this hypothesis, Fn3 5.2.18 was biotinylated at the C-terminus via the same strategy. Examination of the crystal structure of native human Fn3 domain shows the C-terminus facing away from the interaction surface of the three loop regions, thus biotinylation of this site should not affect ligand binding. Importantly, X-ray structures of Fn3 mutants with highly permuted loops have been shown to maintain the general protein fold of wild-type Fn3. However, binding measurements using immobilized Fn3 yielded significantly slower kinetics that also did not correspond to a 1:1 binding model (Figure 2-6b).

Scheme 2-1. General scheme used to biotinylate hBTC via sortase mediated ligation.
The two binding experiments described suggest a more complex binding model than the expected 1:1 model. Specifically, the data imply a 2:1 binding model. Since the Blitz system can only fit to a 1:1 binding model, a different method was needed to test the 2:1 binding hypothesis. Assuming the $K_D$ values obtained on YSD are valid, it would be possible to analyze the hBTC:Fn3 complex using gel filtration chromatography (GFC). By mixing different molar ratios of hBTC:Fn3, the binding stoichiometry could be determined. When hBTC and Fn3 5.2.18 were mixed in a 1:1 molar ratio, two peaks were seen on GFC accounting for the hBTC:Fn3 complex and excess Fn3. Conversely, when hBTC and Fn3 5.2.18 were mixed in a 2:1 molar ratio, a single peak was detected (Figure 2-7a). The profiles of each peak from the two GFC experiments were confirmed on SDS-PAGE (Figure 2-7b).
Deviation from a 1:1 binding model for engineered Fn3 domains had not been previously reported. In fact, several reports have demonstrated the robust development of Fn3 mutants binding to their ligand partners in a 1:1 ratio.\textsuperscript{19,28,29} Thus the Fn3 5.2.18 mutant was regarded as an anomaly within the library. Another Fn3 mutant, 5.2.13, was then expressed and purified to test for hBTC binding. Similar GFC experiments were conducted with Fn3 5.2.13 and hBTC. However, GFC of this Fn3 alone revealed two major peaks, suggesting the presence of two oligomeric states of Fn3 in solution. Similar results were seen with SUMO-fused Fn3 5.2.13. Native-PAGE and SDS-PAGE analysis of the two peaks eluted from SUMO-Fn3 5.2.13 GFC run confirms the presence of both dimeric (45 kDa) and monomeric (22.5 kDa) states of Fn3 5.2.13 in solution (Figure 2-8). Previous reports of purified engineered Fn3 mutants had shown highly monomeric species, even in the presence of free Cys residues in the loop region.\textsuperscript{30} However, the presence of Fn3 in a dimer state can be explained by the fact that wild-type human fibronectin exists as a dimer.\textsuperscript{31}
Given the goal of the present biosensor to serve as a quantitative imaging tool, ensuring the Fn3 mutant binds to hBTC in a 1:1 binding model is essential. Based on the general observation that Fn3 mutants matured using YSD behave monomerically, with similar binding affinities to their ligand counterparts both when expressed on the yeast surface and as free proteins, it is possible the two mutants tested in the hBTC library were anomalies. In a study of the role of amino acid diversity in Fn3 library design, Gilbreth et al noted that two of the five Fn3 mutants showing greatest affinity for its target ligand displayed complex binding kinetics when expressed as a free protein. Thus it may be necessary to test the protein behavior of additional Fn3 mutants from our Fn3 library to find a mutant with the required physical properties for biosensor development.
Experimental

PBS and TE buffers:

Phosphate-buffered saline. All references to solutions prepared in PBS buffer pertain to the following recipe.

10X PBS Buffer:

80.0 g NaCl
2.0 g KCl
14.4 g Na$_2$HPO$_4$
2.4 g KH$_2$PO$_4$

Dissolved the indicated salts in approximately 800 mL Milli-Q water, then adjusted volume to 1 L. For 1X PBS, the 10X stock was diluted 10-fold with Milli-Q water and its pH adjusted to 7.4 using HCl. All 1X PBS buffers were filtered through 0.45 μm nylon membrane filter (Pall Life Sciences) before use. The final 1X PBS concentrations should be 137 mM NaCl, 10 mM phosphate, and 2.7 mM KCl.

Tris-EDTA buffer. All references to solutions in TE buffer refer to the following recipe.

1X TE buffer:

10 mL 1M Tris-HCl pH 7.5
2 mL 0.5 M EDTA pH 8.0

Dilute the indicated stock solutions in approximately 800 mL Milli-Q water. Buffer pH was adjusted to 7.5 with HCl and sterile filtered before use.

hBTC and truncated hBTC cloning:

The hBTC gene (aa$^{32-111}$) and truncated hBTC gene (aa$^{62-111}$) was synthesized by
GenScript with flanking restriction sites (Bsal and Xbal) for cloning into the pE-SUMO vector. The codon sequences were optimization for expression in E. coli. The product was delivered in pUC-57 vector as a lyophilized power and resuspended in 100 uL 1X TE buffer upon arrival. One uL of the purchased plasmid was transformed into DH5α cells (Invitrogen) and grown on LB plates supplemented with carbenicillin (50 µg/mL) according to manufacturer’s protocol. One colony was selected and its plasmid amplified and purified using a Qiagen Plasmid Miniprep kit.

The full DNA sequences of both constructs including the flanking restriction sites are given below. The hBTC and truncated hBTC (hBTC<sub>Δ30</sub>) portions are shown in red.

**hBTC**

```
GGTCTCTAGGT GACGGCAACT CTACCCGCTC CCGGAACGC AACCGGCTGC TGCTGGGCCA CCGGAAGAA AAACCTGTCTG CAACGACGAC GCAATCCAAA CGTAAAGGCC ATTTTAGCGG CTGCCCGAAA CAGTATAAAC ACTACTGCAT TAAAGGTCTT GTGCGCTTCG TGGTTGCGGA ACAAAACCCC GCCTGCGCTG TGATGAAAGG CTACATTGGT GCTCGCTGTG AAGGTGTTGA CCGTTTTTAC TCTAGAACG
```

**truncated hBTC (hBTC<sub>Δ30</sub>)**

```
GGTCTCTAGGT GTGTCGCTGG CTACCCGCTC CCGGAACGC AACCGGCTGC TGCTGGGCCA CCGGAAGAA AAACCTGTCTG CAACGACGAC GCAATCCAAA CGTAAAGGCC ATTTTAGCGG CTGCCCGAAA CAGTATAAAC ACTACTGCAT TAAAGGTCTT GTGCGCTTCG TGGTTGCGGA ACAAAACCCC GCCTGCGCTG TGATGAAAGG CTACATTGGT GCTCGCTGTG AAGGTGTTGA CCGTTTTTAC TCTAGAACG
```

A double digest of the amplified plasmid was completed with Bsal and Xbal (NEB) simultaneously following the manufacturer’s protocol. The pE-SUMO vector was digested with Bsal enzyme for 1 hour, then treated with CIP (NEB) following manufacturer’s protocol. After reaction, both the digested insert and pE-SUMO vector were loaded onto a 1% agarose gel and resolved at 110 V. The desired restriction products (the hBTC insert of the pUC-57 vector and the linearized pE-SUMO expression vector) were cut from the gel and isolated using a Wizard SV gel and PCR clean-up kit (Promega). The hBTC insert was then ligated into the pE-SUMO vector using T4 DNA ligase (NEB). A similar reaction containing only the pE-SUMO vector was used as a negative control. Both ligation reactions were then transformed into DH5α cells (Invitrogen) and plated on LB agar plates containing kanamycin (30 µg/mL). After
incubating overnight at 37°C, three colonies from the non-control plate were picked and amplified. Their respective plasmids were then sent for sequencing. All sequences were determined to be the desired product. The negative control plate showed no colonies.

The SUMO-hBTC and SUMO-hBTC\textsubscript{A30} fusion protein sequences are given below. The hBTC portions are noted in red.

**SUMO-hBTC**

\begin{verbatim}
MGHHHHHHGSVLQDSEVNQAKPEVKPEVKPEVTHINLKVSDGSEIFFKIKKTTPLRRLMEAFARKQGKEMDLSRLFLYD
GIRIQADQAPEDLDMEDNDIIIEAHREQIGGDGNSPETNGULLCGDPEENCAATTQSKRGHSRCPKQQKHYCIGKGR
CRFVVAEQTPSVCDEGYIGARCERVDFLY
\end{verbatim}

**SUMO-hBTC\textsubscript{A30}**

\begin{verbatim}
MGHHHHHHGSVLQDSEVNQAKPEVKPEVKPEVTHINLKVSDGSEIFFKIKKTTPLRRLMEAFARKQGKEMDLSRLFLYD
GIRIQADQAPEDLDMEDNDIIIEAHREQIGGRKHSRCPKQQKHYPGRRFVVAEQTPSVECIGARCERVDFLY
\end{verbatim}

**hBTC\textsubscript{A30} protein expression optimization:**

The plasmid encoding SUMO-hBTC\textsubscript{A30} was transformed into three different *E. coli* strains: BL21-RIL (NEB), Rosetta DE3 (Novagen), and Shuffle (NEB) following the manufacturer’s protocols. Transformed cells were plated onto LB-agar plates containing kanamycin (30 \(\mu\)g/mL) and grown overnight at 37°C. One colony from each cell strain was used to inoculate 5 mL LB-media starter cultures containing kanamycin (30 \(\mu\)g/mL). Each starter culture was grown overnight at 37°C with shaking at 225 rpm. One starter culture was used to inoculate a 1 L culture containing LB-media with kanamycin (30 \(\mu\)g/mL), then grown in a 37°C shaker at 225 rpm. When \(\text{OD}_{600}\) reached 0.6 – 0.8, protein expression was induced by adding 1 mL 1M IPTG. Cultures were grown for 18 – 20 hours at 16°C with shaking at 225 rpm. Cells were then harvested by centrifugation (5000 g, 4°C) for 30 min. The cell pellet was resuspended in 40 mL 1X PBS containing 0.5 M urea, transferred to 50 mL conical tubes (BD Falcon) and repelleted at 5000 x g for 30 min at 4°C. The resulting cell pellets were used immediately for
protein purification or stored at -80°C until use.

On day 1 of purification, cell pellets were thawed and resuspended in the following lysis buffer.

| 4 mL | 10X PBS buffer    |
| 4 mL | Glycerol          |
| 40 uL| Protease Inhibitor Cocktail III (calbiochem) |
| 100 uL| Triton X-100     |
| 6 mg  | DTT               |
| 40 mg | Lysozyme          |
| 40 mL | Final vol after diluting with Milli-Q water |

Once the pellets were fully resuspended, the cells were sonicated on ice at 40% amplitude, 1s on/1s off for 1 min with a Sonics Vibra Cell sonicator. Cell mixtures were allowed to cool for 1 min before a second round of sonication. A total of four rounds of sonication with 1 min cooling period in between were completed to fully lyse the cells. The lysate was clarified by centrifugation at 100,000 x g for 60 min at 4°C. The clarified lysates were then incubated for 1 hour at 4°C with 2-3 mL Ni-NTA resin (Qiagen) pre-equilibrated with 1X PBS, 0.5 M urea.

The resin was then recovered by passing the lysates through a 20 mL disposable chromatography column (Bio-Rad). The flow through was collected. The resin was washed first with 20 mL 1X PBS buffer containing 0.5 M urea, then 20 mL of wash buffer (1X PBS, 20 mM imidazole, pH 7.4) at 4°C. The protein was then eluted in 6x8 mL fractions of elution buffer (1X PBS, 250 mM imidazole, pH 7.4). All protein fractions and sample of the pelleted cell debris were loaded onto 15% SDS-PAGE and visualized by Coomassie staining (Figure 2-9).
**hBTC and truncated hBTC protein expression and purification:**

Plasmids encoding SUMO-hBTC and SUMO-hBTC\(_{Δ30}\) were transformed into BL21-RIL cells and expressed as described in "Truncated hBTC protein expression optimization." Purification of both proteins also proceeded similarly. After purification of the protein lysates, the appropriate elution fractions were combined and dialyzed against 3x4L 1X PBS, 0.5 M urea, pH 7.4. The SUMO-tag was removed by incubating 10 mL purified protein with 500 µL SUMO protease (prepared in house, 30 µM) overnight at 4°C. Ni-NTA resin (1-2 mL) pre-equilibrated with 1X PBS containing 0.5 M urea was then added to the protein solution and incubated at 4°C for 1 hour. The resin was recovered by passing through a 20 mL disposable chromatography column (Bio-Rad). The resin was then washed with 3 mL 20 mM imidazole in 1X PBS, then 3 mL 50 mM imidazole. Captured proteins were eluted in 2 mL 250 mM imidazole in 1X PBS. The flow through fraction was then concentrated using 3K MWCO ultra centrifuge filters (Millipore Amicon) prior to purification by FPLC.

Size exclusion chromatography was completed on an AKTA FPLC system equipped with UNICORN operating software. A HiLoad Superdex30 column (10x300) was used following the

---

Figure 2-9. SDS-PAGE (15% acrylamide) analysis of protein purification from (a) BL21-RIL, (b) Rosetta DE3, and (c) SHuffle cells visualized by Coomassie staining. P – cell debris pellet, FT – flow through, W – wash, 1-11 – elution fractions.
manufacturer’s protocol. Degassed 1X PBS containing 0.5 M urea was used as the running buffer. The chromatography run was monitored at 228 nm and 280 nm UV-light simultaneously and fractions of 0.5 mL were collected. Two major peaks were identified, pooled, and concentrated using 3.5k MWCO ultra centrifuge filters (Amicon). Components of the two peaks were analyzed and identified by SDS-PAGE (15% acrylamide). The second peak corresponded to hBTCΔ30 and hBTC. The purified proteins were flash frozen in liquid nitrogen and stored at -80°C.

Purification of SUMO-hBTC and SUMO-hBTCΔ30 from inclusion bodies was completed by first homogenizing the pelleted cell debris with 40 mL 1X PBS containing 0.5 M urea. The suspension was then re-pelleted by centrifugation (38,000 x g, 30 min, 4°C) before homogenizing with 40 mL resuspension buffer (1X PBS, 8 M urea, pH 7.4). Any remaining cell debris was pelleted by centrifugation at 38,000 x g for 30 min at 4°C. The supernatant was then transferred to a 50 mL conical tube (BD Falcon) and incubated with 2-3 mL Ni-NTA resin pre-equilibrated with resuspension buffer (1 hour, room temperature). Ni-NTA resin was recovered by passing the inclusion body lysate through a 20 mL disposable chromatography column (Bio-Rad) at room temperature. The flow through was collected. The resin was then washed with 20 mL resuspension wash buffer (1X PBS, 8 M urea, 20 mM imidazole, pH 7.4). The His-tagged protein was eluted in 6x10 mL fractions in resuspension buffer containing 300 mM imidazole. The collected fractions were then analyzed on SDS-PAGE (15% acrylamide). The fractions containing pure protein were then pooled and dialyzed step-wise at 4°C in the following buffers to remove the urea and imidazole.

<table>
<thead>
<tr>
<th>Dialysis buffer 1</th>
<th>1X PBS, 6 M urea, pH 7.4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dialysis buffer 2</td>
<td>1X PBS, 4 M urea, pH 7.4</td>
</tr>
<tr>
<td>Dialysis buffer 3</td>
<td>1X PBS, 2 M urea, pH 7.4</td>
</tr>
</tbody>
</table>
The SUMO-tag was then removed by incubating with SUMO-protease and purified as mentioned above. Isolation of untagged hBTC and hBTC\textsubscript{A30} was confirmed via SDS-PAGE (15% acrylamide). Protein concentration was then determined by absorbance at 280 nm and diluted to 50 uM. In a 50 mL conical tube (BD Falcon), the protein was refolded by addition of 1:10 mM GSH:GSSH to reach a final concentration of 0.1:1 mM GSH:GSSH. The protein solution was incubated overnight with gentle rotation on a rotisserie at 4°C. The redox reagents were then removed from the sample by dialysis against 3x4L 1X PBS containing 0.5 M urea. The refolded protein was then further purified on FPLC as described earlier for the soluble fraction. The protein was then flash frozen in liquid nitrogen and stored at -80°C.

**Cell culture and hBTC treatment:**

**Cell culture media and solutions:**

**Dulbecco’s Modified Eagle Media.** All references to full DMEM refer to the following solution.

- DMEM (Life Technologies, ref 11995-065)
- 10% (v/v) fetal bovine serum
- 1X Penicillin/streptomycin (Life Technologies, ref 15140-122)

**Serum-free Dulbecco’s Modified Eagle Media.** All references to serum-free DMEM refer to the following solution.

- DMEM (Life technologies, ref 11995-065)
- Penicillin/streptomycin (Life Technologies, ref 15140-122)

Dulbecco’s Phosphate Buffered Saline (DPBS) was purchased from Corning (ref 21-031-CV). Trypsin EDTA was also purchased from Corning (ref 25-051-Cl). All media and solutions were stored at 4°C. Media were warmed to 37°C prior to use while DPBS and trypsin EDTA
were warmed to 25°C prior to use. Cells were incubated at 37°C in a humidified environment with 5% CO₂.

HeLa cells were cultured in DMEM and split 1:10 every other day in a T45 tissue culture flask (BD Falcon 353135). Cells were passaged 4-6 times, then transferred to two 6-well plates (BD Falcon 353046) and grown overnight to approximately 40% confluency. Media was then removed, cells were washed with DPBS (3x2 mL), then serum-free DMEM was added. Cells were serum starved for 36-48 hours at which point cells reached 80-90% confluency. Media was then removed and cells were washed with DPBS (3x2 mL). Cells were then treated with 200 μL serum-free DMEM containing increasing amounts of commercially purchased hBTC (Peprotech): 0, 10, 25, 50, 100, and 200 ng/mL. After incubating for 10 min, the cells were washed with DPBS (3x2 mL) and collected by incubating with trypsin EDTA for 5 min. Trypsin EDTA was inactivated by addition of 5% (w/v) BSA in DPBS. All conditions were done in duplicate. Cells were then lysed by pipetting up and down with 100 μL of the following lysis buffer.

<table>
<thead>
<tr>
<th>Concentration</th>
<th>Component</th>
</tr>
</thead>
<tbody>
<tr>
<td>50 mM</td>
<td>Tris-HCl, pH 7.5</td>
</tr>
<tr>
<td>150 mM</td>
<td>NaCl</td>
</tr>
<tr>
<td>10 mM</td>
<td>Sodium pyrophosphate dibasic</td>
</tr>
<tr>
<td>50 mM</td>
<td>β-glycerophosphate disodium hydrate, pH 7.3</td>
</tr>
<tr>
<td>30 mM</td>
<td>NaF</td>
</tr>
<tr>
<td>1% (v/v)</td>
<td>Triton X-100</td>
</tr>
<tr>
<td>1 mM</td>
<td>Benzamidine</td>
</tr>
<tr>
<td>2 mM</td>
<td>EGTA</td>
</tr>
<tr>
<td>0.1 mM</td>
<td>Na₃VO₄</td>
</tr>
<tr>
<td>1 mM</td>
<td>DTT</td>
</tr>
<tr>
<td>0.1% (v/v)</td>
<td>Protease Inhibitor Cocktail III (Calbiochem)</td>
</tr>
<tr>
<td>0.1% (v/v)</td>
<td>Phosphatase Inhibitor Cocktail I (Calbiochem)</td>
</tr>
</tbody>
</table>

* Denotes these components should only be added immediately prior to use.

After incubating on ice for 1 hour, cell debris was pelleted by centrifugation (14000 x g) for 10 min at 4°C. The clarified lysate was collected, flash frozen in liquid nitrogen, and stored at -80°C. Total protein concentration of lysates was determined using the Bio-rad protein assay.
Similar procedures were used to test for activity of recombinantly expressed hBTC and hBTC\textsubscript{A30} purified from BL21-RIL cells. Cells were then treated with 200 uL serum-free DMEM containing PBS, 11 nM commercial hBTC (Peprotech), 11 nM hBTC (soluble fraction), 11 nM hBTC (insoluble fraction), 11 nM hBTC\textsubscript{A30} (soluble fraction), or 11 nM hBTC\textsubscript{A30} (insoluble fraction).

**Sox-PNT assay:**

Sox-PNT assays were done to measure Erk1/2 activity of cell lysates after treatment with hBTC following procedures reported previously.\textsuperscript{25} Briefly, lysates were diluted to a working concentration of 4 mg/mL. The appropriate volumes of 4X Sox-assay buffer (see recipe below) and ddH\textsubscript{2}O were added to a 96-well plate (Corning, ref 3992), then warmed to 30°C (approximately 10 mins in 30°C oven). The warmed plate was then loaded into a pre-warmed 30°C BioTek Synergy H1 plate reader. A background scan was completed using the single read method (temperature set at 30°C, 150 gain, \( \lambda_{ex} \) 360 nm, \( \lambda_{em} \) 485 nm). Lysates were then added to yield a final concentration of 0.13 mg/mL and mixed by pipetting up and down several times. The assay was run using kinetic method (temperature set at 30°C, 150 gain, \( \lambda_{ex} \) 360 nm, \( \lambda_{em} \) 485 nm, run time 2 hours, read every 5-10 sec). A negative control was done using PBS in place of lysate. A positive control was done using commercial Erk 1 (0.00009 mg/mL) purchased from Invitrogen.
**4X Sox-assay buffer:**

<table>
<thead>
<tr>
<th>Component</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>HEPES, pH 7.5</td>
<td>200 mM</td>
</tr>
<tr>
<td>MgCl₂</td>
<td>40 mM</td>
</tr>
<tr>
<td>EGTA</td>
<td>4 mM</td>
</tr>
<tr>
<td>DTT</td>
<td>8 mM</td>
</tr>
<tr>
<td>ATP</td>
<td>4 mM</td>
</tr>
<tr>
<td>Brjii-35P</td>
<td>0.04 %</td>
</tr>
<tr>
<td>Sox-PNT sensor</td>
<td>0.02 mM</td>
</tr>
</tbody>
</table>

*4X Sox-assay buffer was diluted to with ddH₂O and lysate to 1X for assay run.

**Sortase amenable hBTC cloning, expression, and purification:**

The sortase recognition sequence (LPRTGG) was appended to the C-terminus of hBTC in two steps by whole plasmid mutagenesis. The first whole plasmid mutagenesis added residues LPR, while the second added TGG to the vector encoding SUMO-hBTC. The following primers were obtained from Sigma-Aldrich (50 nmol scale, PAGE purified).

<table>
<thead>
<tr>
<th>Primer name</th>
<th>DNA sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>BTC_LPR_fwd</td>
<td>GAACGTGGTAGCTGAATTCTTAGGCGTGTAATCTAGAGGATCCGAA</td>
</tr>
<tr>
<td>BTC_LPR_rev</td>
<td>GAATTCCGGATCTCTAGATTAACGCGGCAAGTAAAAACATTCAACAG</td>
</tr>
<tr>
<td>BTC_TGG_fwd</td>
<td>CTGTTTTACCTGCGCGGTACCGGTGCTAAATCTAGAGGATCCGAA</td>
</tr>
<tr>
<td>BTC_TGG_rev</td>
<td>TCCTGCGATACCTGCTGGCTGCGGGCAAGTAAAAACATTCAACAG</td>
</tr>
</tbody>
</table>

Upon receiving the primers, they were resuspended in Milli-Q water to the concentration of 100 μM. The primers were further diluted to 10 μM for use. The mutagenesis reaction was assembled as follows, then the reaction was run in MJ-mini PCR machine (Bio-Rad) using the following thermocycle method.

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 μL</td>
<td>10X PFu buffer (NEB)</td>
<td></td>
</tr>
<tr>
<td>0.5 μL</td>
<td>pE-SUMO vector (SUMO-hBTC) (70 ng/μL)</td>
<td></td>
</tr>
<tr>
<td>1 μL</td>
<td>Forward primer (10 μM)</td>
<td></td>
</tr>
<tr>
<td>1 μL</td>
<td>Reverse primer (10 μM)</td>
<td></td>
</tr>
<tr>
<td>3 μL</td>
<td>DMSO</td>
<td></td>
</tr>
<tr>
<td>1 μL</td>
<td>DNTPs (10 mM)</td>
<td></td>
</tr>
<tr>
<td>1 μL</td>
<td>PfU Turbo Enzyme (NEB)</td>
<td></td>
</tr>
<tr>
<td>37.5 μL</td>
<td>Milli-Q water</td>
<td></td>
</tr>
<tr>
<td>50 μL</td>
<td>Total reaction volume</td>
<td></td>
</tr>
<tr>
<td>Step</td>
<td>Time/cycles</td>
<td>Temperature</td>
</tr>
<tr>
<td>--------------------------</td>
<td>-------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Initial activation step</td>
<td>3 min</td>
<td>95</td>
</tr>
<tr>
<td>Denaturation</td>
<td>30 sec</td>
<td>95</td>
</tr>
<tr>
<td>Annealing</td>
<td>1 min</td>
<td>55</td>
</tr>
<tr>
<td>Extension</td>
<td>1 min/kb</td>
<td>72</td>
</tr>
<tr>
<td>Number of cycles</td>
<td>Go to step 2, 25 times</td>
<td></td>
</tr>
<tr>
<td>Final denaturation</td>
<td>20 min</td>
<td>72</td>
</tr>
<tr>
<td>End of PCR</td>
<td>forever</td>
<td>4</td>
</tr>
</tbody>
</table>

Successful amplification of the hBTC gene was confirmed by running 10 uL of reaction mixture on 1% agarose gel and staining with EtBr. The template plasmid was then digested by incubating with 1 µL DpnI for 1 hour at 37°C. The reaction was then cleaned up using a Wizard SV gel and PCR clean-up kit (Promega) before transformation into DH5α cells (Invitrogen). After growing overnight on LB-agar plates containing kanamycin (30 µg/mL), three colonies were picked and its plasmid sequenced to confirm successful incorporation of the LPR residues. This new plasmid was then used as the template for the second whole plasmid mutagenesis reaction to incorporate the remaining three amino acids (TGG), following the same protocol as described. The amino acid sequence of the resulting sortase amenable hBTC is given below. The hBTC portions are noted in red, with the sortase recognition sequence in blue.

**SUMO-hBTC**

MGHHHHHHGSGLQDSEVQPEKPEVTHINLKVSDOSSIEFFKIKKTTLRRLMEAFAKRQGKEQDSL
RFYDGIQIAEQADQAPEDLDMEINDEAHEIREDGNGDSRPEIDNGLDPEIENDATTTQSKRKGHSRCP
KQYKHYCKIRCRFVVAEQTSCVCEGYIARCERVDLFYLPRTGG

Expression and purification of sortase amenable hBTC was performed as described for hBTC and hBTCΔ30 above. Except, refolded proteins were dialyzed against 3x4L 50 mM HEPES, 150 mM NaCl, 0.5 M urea, pH 7.4 instead of PBS buffer. FPLC purification was not done at this stage.
Cloning hBTC-binding Fn3 mutants 5.2.13 and 5.2.18:

Plasmids encoding the hBTC-binding Fn3 mutants 5.2.13 and 5.2.18 (pCTCON-2) were obtained from our collaborators (Seymour De Picciotto, Wittrup Lab). The genes encoding for the Fn3s were amplified by PCR to include restriction sites \((BsaI\) and \(XbaI\)) for ligation into the pE-SUMO vector. The primers were obtained from Sigma-Aldrich (25 nmol scale, purified by desalting). The primer sequences are given below.

<table>
<thead>
<tr>
<th>Primer name</th>
<th>DNA sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>BTC_13_fwd</td>
<td>GCT GGT CTC TA GGT GTT TCT GAT GTC CCG AGG</td>
</tr>
<tr>
<td>BTC_13_rev</td>
<td>GCT GGT CTC TCT AGA TCA TGT TCG GTA ATT AAT GGA</td>
</tr>
<tr>
<td>BTC_18_fwd</td>
<td>GCT GGT CTC TA GGT GTT TCT GAT GTC CCG AGG GAC</td>
</tr>
<tr>
<td>BTC_18_rev</td>
<td>GCT GGT CTC TCT AGA TCA TGT TCG GTA ATT AAT GAA GGA</td>
</tr>
</tbody>
</table>

Amplification of the two Fn3 genes were completed by the same method used for whole plasmid mutagenesis described above. The template plasmid was digested with DpnI (NEB) following manufacturer’s protocol, then cleaned up using a Wizard SV gel and PCR clean-up kit (Promega). The PCR products and pE-SUMO vector were then digested, ligated, and transformed as described above in “hBTC and truncated hBTC cloning.” Successful ligations were confirmed by sequencing of isolated plasmids. The protein sequences of the ligation products are given below. The gene coding for each Fn3 mutant is shown in red.

SUMO-Fn3 5.2.13

MGHHHHHHGLVQSEKPEVKPRKTKMTNKLKVDSQSESEIIKIKKTTPLRLLMEAFAKRQKEMDSLRFLFDGIRQADQAPEDLDMEDNDIDAEHREPQIGEYSGVPRDLEVVAATPTSLISWYDWSFADSRYITYGETGNSPVQFETVPGSALTATISGLKPGVTHINGYAVYAVIDWYDYESDPISINYRT

SUMO-Fn3 5.2.18

MGHHHHHHGLVQSEKPEVKPRKTKMTNKLKVDSQSESEIIKIKKTTPLRLLMEAFAKRQKEMDSLRFLFDGIRQADQAPEDLDMEDNDIDAEHREPQIGEYSGVPRDLEVVAATPTSLISWYDWSFADSRYITYGETGNSPVQFETVPGSALTATISGLKPGVTHINGYATGDYDDYSNDPISFINYRT
Expression and purification of hBTC-binding Fn3 mutants:

hBTC binding Fn3 5.2.13 and 5.2.18 were expressed and purified with a SUMO tag as described above for “Truncated hBTC protein expression optimization.” Removal of the SUMO tag was completed as described in “hBTC and truncated hBTC protein expression and purification.” Purified proteins were flash frozen in liquid nitrogen and stored at -80°C.

Sortase amenable Fn3 5.2.18 cloning, expression, and purification:

The sortase recognition sequence (LPRTGG) was appended to the C-terminus of hBTC-binding Fn3 5.2.18 using two consecutive PCRs. The first PCR added residues LPR, while the second added TGG to the vector encoding SUMO-Fn3 5.2.18. The following primers were obtained from Sigma-Aldrich (25 nmol scale, purified by desalting).

<table>
<thead>
<tr>
<th>Primer name</th>
<th>DNA sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>BTC_18_fwd</td>
<td>GCT GGT CTC TA GGT GTT TCT GAT GTT CCG AGG GAC</td>
</tr>
<tr>
<td>BTC_18_LPR_rev</td>
<td>ACC GGT ACG CGG CAG TG TTC GGT AAT TAA TGA</td>
</tr>
<tr>
<td>BTC_18_TGG_rev</td>
<td>CG TTC TAG A TCA GCC ACC GGT ACG CGG</td>
</tr>
</tbody>
</table>

PCR was completed following the same protocol as reported for “Sortase amenable hBTC cloning, expression, and purification.” The amplified PCR insert was then ligated into pE-SUMO vector using T4 DNA ligase (NEB) following manufacturer’s instructions. Successful ligations were confirmed by plasmid sequencing. The amino acid sequence of the resulting sortase amenable Fn3 5.2.18 is given below. The Fn3 portions are noted in red, with the sortase recognition sequence in blue.

SUMO-Fn3 5.2.18

MGHHHHHHHGSQDSEVQNEAKPEVKPEVKPETHINLKVSADGGSEIFKIKKTTLRRLMEAFARKQGKE MD SLPFLYDGIRIQADQAPELDMDNEDNIEAHREQKGGVSDVPRDLEVAAATPSTLSLLISWYSWFSADSYRITY GETGNSPVQFTVGFSATISGLKPVGDYTTITVYAVTDYDDYSDPISFINYRTLPRRTGG

Expression and purification of sortase amenable Fn3 5.2.18 was done following the same
procedures for expression and purification of Fn3 5.2.18 as described above. Except, in the final dialysis step, protein is dialyzed into 50 mM HEPES, 150 mM NaCl, 0.5 M urea, pH 7.4.

**Synthesis of GGGYKK(biotin)T peptide:**

The GGGYKK(biotin)T peptide was prepared manually using standard Fmoc-based solid-phase peptide synthesis (SPPS). The peptide was made on a 0.02 mmol scale using Fmoc-PAL PEG resin (Novabiochem, 0.18 mmol/g) in a 5 mL polypropylene syringe (Torviq, ref SF-0500). The following Fmoc-protected amino acids were used: Fmoc-Gly-OH, Fmoc-Tyr(tBu)-OH, Fmoc-Lys(Boc)-OH, Fmoc-Lys(Mtt)-OH, and Fmoc-Thr(tBu)-OH. The Fmoc-protected Lys containing an weak acid labile protecting group (Mtt) was incorporated at the n+6 position to allow for incorporation of a biotin moiety.

The resin was swelled in DCM for 10 min, then washed with DCM (2x2 mL) and DMF (3x2 mL). Fmoc deprotection was done by incubating the resin with 20% (v/v) 4-methylpiperidine in DMF (2 mL, 3x5 min). Successful deprotection was evaluated by measuring Fmoc absorbance at 300 nm. Resin was then washed with DMF (3x2 mL), DCM (3x2 mL), and DMF (3x2 mL). Coupling reactions were done using 6 equivalents of amino acid, coupling reagents and N,N-Diisopropylethylamine (DIPEA). The Fmoc-amino acid was dissolved in NMP with benzotriazole-1-yl-oxy-tris-pyrrolidino-phosphonium hexafluorophosphate (PyBOP) and hydroxybenzotriazole (HOBT) prior to addition of DIPEA. Coupling was allowed to proceed for 45 min – 1 hour at room temperature. The resin was then washed with DMF (3x2 mL), DCM (3x2 mL), and DMF (3x2 mL).

Once all amino acids had been coupled to the resin, the weak acid labile Mtt group was selectively removed by incubating with 1% TFA in DCM (2 mL) for 5 min. Removal of the Mtt
group changes the TFA solution yellow. Treatment with 1% TFA in DCM was repeated until no yellow color remained. Biotin was then coupled onto the free amine of Lys as described with the amino acids. The final Fmoc-protecting group was then removed.

Following peptide synthesis, the peptide was cleaved from the resin and the side chain protecting groups were removed by incubating the resin with 3 mL of TFA/triisopropyl saline/H$_2$O/DCM (90/2.5/2.5/5 v/v/v/v) for 3-4 hours at room temperature. The TFA was removed by evaporation and the peptide was triturated with cold diethyl ether (3x 10 mL). Peptides were purified by reverse phase HPLC with a Waters 1525 and a Waters 2489 dual-wavelength absorbance detector set at 228 and 280 nm. Separations were performed on a semi-prep C$_{18}$ YMC-Pack ODS-A column (250 x 20 mm). Analytical HPLC was done using a C$_{18}$ YMC-Pack ODS-A column (250 x 4.6 mm). HPLC conditions were 15% acetonitrile in water with 0.1% TFA for 5 min followed by a linear gradient from 15% to 70% acetonitrile in water with 0.1% TFA over 25 min. Following lyophilization, the peptide mass and purity was confirmed by LC-MS (Hewlett Packard 1100 LC, Finnigan LCQ Deca ESI-MS) running a reverse phase C$_{18}$ YMC-Pack ODS-AQ column (100 x 30 mm).

**Sortase mediated ligation:**

Biotinylation of hBTC and Fn3 5.2.18 was accomplished via sortase mediated ligation (SML). An evolved triple sortase A mutant (P94S/D160N/K196T) was used for ligation.$^{33}$ Each protein was diluted to 30-50 μM with 50 mM HEPES, 150 mM NaCl, 0.5 M urea, pH 7.4. Total reaction volume was 500 μL. The SML reaction was set up as follows.

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>hBTC-LPRTGGG/Fn3-LPRTGGG</td>
<td>1 eq</td>
</tr>
<tr>
<td>Sortase A triple mutant</td>
<td>3 eq</td>
</tr>
<tr>
<td>GGGYKK(biotin)T peptide</td>
<td>3 eq</td>
</tr>
<tr>
<td>2 M CaCl$_2$</td>
<td>10 mM</td>
</tr>
</tbody>
</table>
After incubating reaction for 2 hours at 4°C or room temperature for 30 min, the reaction was stopped by adding 10 μL 0.5M EDTA. The reaction was then desalted and concentrated using a C_{18} Sep-Pak plus short cartridge (Waters, ref 036575) to 400 μL 1:1 (v/v) acetonitrile:ddH₂O containing 0.1% TFA. The acetonitrile and TFA was removed by evaporation and 1:10 (v/v) 10X PBS was added to the protein solution. The solution was immediately loaded onto an AKTA FPLC system equipped with a HiLoad Superdex200 column (10×300) following manufacturer's protocols. Purification was completed using degassed buffer (50 mM HEPES, 150 mM NaCl, 0.5 M urea, pH 7.4) and monitored at 228 nm and 280 nm UV-light simultaneously. Fractions of 0.5 mL were collected and analyzed using SDS-PAGE (15% acrylamide). The appropriate fractions were pooled and concentrated using 3.5K MWCO ultra centrifugation filters (Amicon). Protein was then flash frozen in liquid nitrogen and stored at -80°C. Successful ligation of the biotin peptide to hBTC and Fn3 5.2.18 was confirmed by MALDI-MS (Voyager).

K_D measurements using bio-layer interferometry:

K_D measurements were completed by bio-layer interferometry on a Blitz System Package (Fortebio) following the manufacturer's protocols. Briefly, a streptavidin coated sensor tip (Fortebio ref 18-5019) was hydrated for 20 min in blitz buffer (1X PBS, 0.1% BSA, 0.00002% tween, pH 7.4). The biotinylated protein (hBTC or Fn3 5.2.18) was diluted into blitz buffer to 200 nM, then loaded onto the streptavidin sensor tip for 300 sec with shaking at 2200 rpm. Association curves were obtained by exposing the sensor to 250 μL of hBTC or Fn3 5.2.18 at varying concentrations (50 – 5000 nM) for 480 sec with shaking at 2200 rpm. A dissociation step followed where the sensor tip was exposed to 250 μL blitz buffer for 480 sec with shaking at 2200 rpm. The sensor tip was then regenerated using 250 μL 10 mM glycine, pH 1.0 (270 sec,
2200 rpm) followed by washing with 250 µL blitz buffer for 120 sec. Association and dissociation curves were globally fit to a 1:1 binding model using the Blitz software.

**Characterization of hBTC binding to Fn3 mutants using gel-filtration chromatography:**

Gel-filtration chromatography was performed on an AKTA FPLC system equipped with a Superdex200 column (10x300). The running buffer was 1X PBS, 0.5 M urea, pH 7.4 (degassed). Chromatography traces were monitored simultaneously at 228 nm and 280 nm. Fractions of 0.5 mL were collected throughout the run. First, chromatograms of each protein (hBTC, Fn3 5.2.13, Fn3 5.2.18) were collected. Both hBTC and Fn3 5.2.18 showed one major peak, but Fn3 5.2.13 showed two major peaks. The two peaks of Fn3 5.2.13 were collected and analyzed by both SDS-PAGE and native-PAGE (15% acrylamide). hBTC was mixed with Fn3 5.2.13 and Fn3 5.2.18 in an equimolar fashion, then loaded onto the FPLC. Fn3 5.2.18 was then mixed with hBTC in a 1:2 (Fn3 5.2.18:hBTC) molar ratio and loaded onto the FPLC. All major peaks were collected and analyzed by SDS-PAGE (15% acrylamide).

**References:**


Chapter 3

Labeling studies of hBTC binding Fn3
Introduction

Fluorescent probes have been extensively used in a variety of biological and biochemical studies including those aimed at investigating protein-protein interactions, visualizing subcellular environments, and quantifying enzyme activity.\textsuperscript{1,2} Applications of fluorescent dyes span the fields of academia into industry and medicine for the use of high throughput screens for drug candidates.\textsuperscript{3,4} Indeed, few other techniques offer the level of spatial and temporal resolution fluorescent techniques can provide to capture the dynamics of living systems.

Fluorescent probes capable of detecting changes in local solvent environment are especially useful in studying subcellular environmental conditions such as pH and membrane potentials.\textsuperscript{5} Solvatochromic fluorophores represent one such class of environmentally sensitive fluorophores. Solvatochromism describes the dye’s characteristic change in emission properties relative to its solvent environment. This unique feature is especially useful in studying protein-protein interactions because the binding interface of proteins often undergoes a substantial change in environment upon protein binding.\textsuperscript{6,7} In particular, the dimethylaminonaphthalimide family of fluorophores exhibits extremely low fluorescence quantum yields (\(\Phi < 0.05\)) when exposed to polar protic solvents, which presents the advantage of a low fluorescence background.\textsuperscript{8,9} This low background signal allows the dye to exhibit “switch-like” behavior to signal a binding event with potential emission intensity changes of up to 1000-fold.\textsuperscript{9} Recently, our laboratory has demonstrated the superior stability and fluorescence properties of 4-\(N,N\)-dimethylamino-1,8-naphthalimide (4-DMN), in detecting protein-peptide interactions within the calmodulin-M13 peptide model system. Binding of the M13 peptide containing a 4-DMN amino acid (4-DMNA) by calmodulin in the presence of Ca\(^{2+}\), yield a fluorescence increase of over 100-fold, a 10-fold larger fluorescence intensity change compared to many commercially
available solvatochromic fluorophores. Moreover, “activated” 4-DMN can be used in both intra- and intermolecular FRET approaches as the donor fluorophore. Importantly, 4-DMN is also compatible with two-photon imaging, allowing a versatile range of imaging strategies.

Results and Discussion

3-1. Generation of 4-DMN-V

A series of 4-DMN analogs for facile incorporation of the dye into proteins and peptides have been developed in the Imperiali lab. The amino acid variant of 4-DMN is compatible with Fmoc-based solid phase peptide synthesis enabling easy incorporation into peptide sequences. Four thiol-reactive analogs with varying linker lengths containing a maleimide or an α-bromoamide group are available for labeling free cysteine residues on proteins (Figure 3-1).

![Chemical structures of four thiol-reactive 4-DMN derivatives with their respective spacer lengths indicated.](image)

Importantly, the linker length in these derivatives has been demonstrated to be incredibly important for optimizing the fluorescent response of 4-DMN in signaling a protein-binding event. Indeed, similar effects of linker length were observed by Dr. Elke Socher in her labeling studies of an EGF-binding Fn3 mutant (unpublished data). While most of the Fn3 mutants tested by Dr. Socher showed no obvious trend relative to the length of the linker regions, one particular mutant showed different results. Specifically, the E48C mutant showed increasing levels of fluorescence change with increasing linker lengths used in the 4-DMN labeling reagent (Figure...
This trend suggested that perhaps labeling E48C with a 4-DMN derivative with an even longer linker might yield a fluorescence change of \( > 3 \)-fold. To test this hypothesis, a new thiol-reactive 4-DMN derivative (4-DMN-V) was synthesized with a linker length of \( \sim 12 \text{ Å} \) (Scheme 3-1). This new derivative has been added to the suite of thiol-reactive 4-DMN reagents used for proteins labeling studies.

**Figure 3-2.** Graph summary of fluorescent signal increases of 4-DMN labeled Fn3 mutant E48C upon binding to EGF

**Scheme 3-1.** Synthesis of 4-DMN-V (5-2)

Reagents and conditions: (a) 2-(2-(2-aminoethoxy)ethoxy)ethanol, EtOH, \( \text{N}_2 \), reflux, 84% yield; (b) \( \text{PPh}_3 \), DEAD, neopentyl alcohol, maleimide, THF, \( \text{N}_2 \), -78°C to rt, 27% yield.
3-2. Fluorescence studies of hBTC binding-Fn3

With the evolved hBTC-binding Fn3 scaffold of our biosensor design in hand, placement of a solvatochromic fluorophore was carried out. Several labeling strategies are available for incorporation of a small chemical fluorophore into the Fn3 protein scaffold. One of the most common approaches is protein labeling of Cys residues with thiol-reactive dyes because cysteine residues can be introduced site-specifically through conventional site-directed mutagenesis. Moreover, since the Fn3 scaffold is naturally devoid of cysteine residues, this labeling methodology allows specific labeling of targeted positions on the protein. The suite of five thiol-reactive 4-DMN derivatives described in the previous section was used to label all hBTC-binding Fn3 Cys-mutants. All labeling was performed on Cys-mutants of the evolved Fn3 5.2.18.

Four sites (V1C, F53, Y31, and Y23) were chosen for labeling studies because of their proximity to the binding surface of the Fn3 scaffold (Figure 3-3). First, each site was mutated to Cys using conventional site-directed mutagenesis. Next, the $K_D$ of each Cys-mutant binding to hBTC was determined by bio-layer interferometry. Since the key molecular contacts that drive the binding of Fn3 5.2.18 to hBTC is unclear, any permutation in the Fn3 mutant could potentially disrupt this binding interaction. Moreover, since native concentrations of hBTC in the body are within the nanomolar range, it is critical for our biosensor to have a $K_D$ within that range. Of the four mutants, only one showed a measured $K_D$ of greater than 1 μM (Table 3-1). As a result, this mutant (Y31C) was not pursued further in the labeling studies.
Figure 3-3. Side (left) and top (right) views of the Fn3 scaffold structure with mutations site residues V1, Y23, Y31, and F53 shown as their chemical structures in magenta. The interaction surface loops are shown in blue.

Table 3-1. Summary of \( K_D \) values determined for each Fn3 5.2.18 Cys-mutant via bio-layer interferometry.

<table>
<thead>
<tr>
<th>Fn3 5.2.18 mutant</th>
<th>( K_D )</th>
</tr>
</thead>
<tbody>
<tr>
<td>V1C</td>
<td>35 nM</td>
</tr>
<tr>
<td>F53C</td>
<td>100 nM</td>
</tr>
<tr>
<td>Y31C</td>
<td>&gt; 1 uM</td>
</tr>
<tr>
<td>Y23C</td>
<td>250 nM</td>
</tr>
</tbody>
</table>

The remaining three mutants, V1C, F23C, and Y23C, were each labeled with the set of five thiol-reactive 4-DMN derivatives. Following this, changes in fluorescence signal were measured against increasing concentrations of hBTC. None of the labeled Fn3 5.2.18 mutants showed a significant fluorescence increase upon binding to hBTC. Interestingly, one mutant (V1C) displayed a decrease in fluorescence signal upon binding to hBTC of up to 4-fold (Figure 3-4). Generation of a “turn-off” hBTC sensor utilizing fluorescence quenching as the detection
method was not anticipated, especially at the V1 position. The mutation site of V1C is present at the N-terminus of the Fn3 protein. Analysis of the crystal structure of the Fn3 scaffold shows that the N-terminal tail is extremely flexible and spatially positioned near the binding interface. Coupled with the fluorescence data, this may suggest that prior to binding hBTC, the 4-DMN fluorophore on the V1C mutant is able to associate non-specifically to a hydrophobic surface on the protein. Upon binding to hBTC, this non-specific association of the fluorophore with Fn3 is displaced and the fluorophore is exposed to a more hydrophilic environment. Importantly, since this trend was not seen with the other two mutants of Fn3 5.2.18, it seems that this is a specific characteristic of the N-terminal tail of the Fn3.

![Fluorescence Intensity Changes Chart](image)

**Figure 3-4.** Chart summarizing the fluorescence intensity changes measured for each Fn3 5.2.18 Cys-mutant labeled with each thiol-reactive 4-DMN derivative.
While only four sites on Fn3 5.2.18 were labeled with 4-DMN and tested for fluorescence change upon ligand binding, the results from these four mutants have been very insightful. First, one mutant (Y31C) showed a significant decrease in affinity toward hBTC of over 100-fold. This suggests that Y31 is a key residue governing the binding interaction between Fn3 and hBTC. Removal of this single residue reduced the binding affinity of Fn3 5.2.18 to that measured for Fn3 mutants isolated from the first 2-3 generations during YSD evolution. Interestingly, removal of another Tyr only 9 residues away (Y23) and present in the same BC loop had a much smaller effect on its affinity to hBTC. Analysis of these residues in the context of the Fn3 crystal structure shows Y23 to be well within the BC loop, while Y31 is at the edge of the beta-strand exiting the BC loop. However, the exact spatial locations of each residue in Fn3 is unknown because there lacks a crystal structure of the Fn3 5.2.18 mutant. The BC, DE, and FG loops of Fn3 have been well documented to make up the binding interaction surface even after significant permutations to its primary sequence. However, the exact residues that form these three loops change depending on the mutations introduced into the mutant. Thus, it is very likely that the residues making up the BC loop in Fn3 5.2.18 have shifted during the evolution process. Additional mutagenesis studies to residues flanking Y23 and Y31 of the BC loop may be necessary to better understand the composition of the interaction surface on Fn3 5.2.18.
Experimental

Synthesis of thiol-reactive 4-N,N-dimethyl-1,8-naphthalimide reagents:

NMR of the following derivatives: $^1$H and $^{13}$C NMR spectra were recorded on a Bruker AVANCE III 400 MHz NMR spectrophotometer operating TopSpin 3.1. Chemical shifts (δ) are referenced to residual proton in the deuterated solvent and are reported in units of parts per million (ppm). Coupling constants (J) are reported in units of Hertz (Hz). Multiplicities in $^1$H NMR are abbreviated as singlet (s), doublet (d), triplet (t), and multiplet (m).

\[
\begin{align*}
\text{H}_2\text{N} & \quad \text{O-} \\
\text{O} & \quad \text{OH} \\
\text{N} & \quad \text{O-} \\
\text{O} & \quad \text{OH}
\end{align*}
\]

6-(dimethylamino)-2-(2-(2-hydroxyethoxy)ethoxy)ethyl-1H-benzo[de]isoquinoline-1,3(2H)-dione (5-1).

Dissolved 4-N,N-dimethylamino-1,8-naphthalic anhydride, 3-1 (0.73 g, 3.05 mmol, 1 eq) was added to a 200 mL two-neck round-bottom flask equipped with a reflux condenser, magnetic stir bar, and rubber septum. Ethanol (31 mL, 200 proof) was added to the flask and heated to reflux. The anhydride remained a solid even when the suspension reached reflux. Upon addition of 2-(2-(2-aminoethoxy)ethoxy)ethanol (0.5 g, 3.35 mmol, 1.1 eq) to the reaction by syringe, the solution turned a clear, deep orange color. The reaction was refluxed for another 1.5 hours before cooling to room temperature. The reaction was concentrated to dryness by rotary evaporation and dried overnight under high vacuum to remove excess 2-(2-(2-aminoethoxy)ethoxy)ethanol. The product was isolated by flash column chromatography applying a gradient of 1:2 hexanes/ethyl
acetate to 0:1 ethyl acetate. The product was an orange oil (0.92 g, 2.5 mmol, 84% yield, $R_f = 0.1$ in 1:3 hexanes/ethyl acetate).

$^1$H NMR (400 MHz, CDCl$_3$) $\delta$: 8.42 (dd, $J = 7.3$, 1.2 Hz, 1H), 8.34 – 8.28 (m, 2H), 7.52 (dd, $J = 8.5$, 7.3 Hz, 1H), 6.98 (d, $J = 8.3$ Hz, 1H), 4.32 (t, $J = 6.2$ Hz, 2H), 3.74 (t, $J = 6.2$ Hz, 2H), 3.65 – 3.60 (m, 2H), 3.59 – 3.53 (m, 4H), 3.49 – 3.45 (m, 2H), 3.00 (s, 6H).

$^{13}$C NMR (101 MHz, CDCl$_3$) $\delta$: 164.48, 163.89, 156.84, 132.56, 131.14, 130.89, 130.10, 124.99, 124.68, 122.73, 114.51, 113.08, 72.41, 70.40, 70.01, 67.95, 61.62, 44.63, 38.77.

Prior to assembling the reaction, 5-1 (0.9 g, 2.5 mmol, 1.1 eq) was dried overnight under high vacuum to remove any trace water that may inhibit the Mitsunobu reaction. Solid triphenylphosphine (.60 g, 2.28 mmol, 1 eq) was added to a flame-dried Kjeldahl-style Schlenk flask equipped with a magnetic stir bar and rubber septum. The reaction vessel was then evacuated of air by placing under high vacuum and charged with N$_2$ gas (3x). Anhydrous THF (8 mL) was added to the Schlenk flask by syringe. The triphenylphosphine and THF mixture was stirred at room temperature until fully dissolved. The reaction was cooled to -78°C by immersing in a dry ice/isopropanol bath. Diethyl azodicarboxylate (0.36 mL, 40% in toluene, 2.28 mmol, 1 eq) was then added dropwise to the reaction by syringe over 2 min. The mixture was allowed to stir for 5 min, at which point it turned a pale yellow color. Meanwhile, 5-1 was dissolved in 12
mL anhydrous THF and slowly added to the reaction via syringe to avoid any change in temperature. Next, neopenyl alcohol (0.12 mL, 1.14 mmol, 0.5 eq) was added as a solid and allowed to dissolve. Solid maleimide (0.22 g, 2.28 mmol, 1 eq) was then added and the reaction was allowed to warm to room temperature. The solution was stirred overnight in the dark. The reaction was stopped by concentrating to dryness on the rotary evaporator. The product was isolated by flash column chromatography using 1:2 hexanes:ethyl acetate. The product was a yellow solid (0.31 g, 0.69 mmol, 27% yield, Rf = 0.4 in 1:2 hexanes/ethyl acetate).

$^1$H NMR (400 MHz, CDCl$_3$) δ: 8.53 (dd, $J = 7.3$, 1.2 Hz, 1H), 8.44 (d, $J = 8.2$ Hz, 1H), 8.41 (dd, $J = 8.5$, 1.2 Hz, 1H), 7.63 (dd, $J = 8.5$, 7.3 Hz, 1H), 7.09 (d, $J = 8.2$ Hz, 1H), 6.66 (s, 2H), 4.37 (t, $J = 6.3$ Hz, 2H), 3.74 (t, $J = 6.3$ Hz, 2H), 3.66 – 3.60 (m, 4H), 3.59 – 3.54 (m, 4H), 3.08 (s, 6H).

$^{13}$C NMR (101 MHz, CDCl$_3$) δ: 170.75, 164.70, 164.12, 157.06, 134.20, 132.77, 131.30, 131.14, 130.37, 125.35, 124.96, 123.09, 114.98, 113.38, 70.19, 70.12, 68.08, 67.83, 44.86, 38.93, 37.17, 31.67.

Cloning, expression, and purification of Fn3_5.2.18 cysteine mutants:

Site-directed mutagenesis was applied to the SUMO-Fn3 5.2.18 construct to yield the following mutants: V1C, F53C, Y31C, and Y23C. The primers used were ordered PAGE-purified from Sigma-Aldrich on the 25 mmol scale. The primer sequences are given below.

<table>
<thead>
<tr>
<th>Primer</th>
<th>DNA sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>BTC18_V1C_fwd</td>
<td>ACCGCGAACAGATTGGAGGTTGCTCTGATGTTCGGAGGGAC</td>
</tr>
<tr>
<td>BTC18_V1C_rev</td>
<td>GTCCCTCAGAATCACAGGAACACCTCCAATCTCGGTTCGGGT</td>
</tr>
<tr>
<td>BTC18_F53C_fwd</td>
<td>GTGCCCTGCTCTGCTCTGCTCTGCTCTGCTACC</td>
</tr>
<tr>
<td>BTC18_F53C_rev</td>
<td>GTTGAAGACAGGACAGACAGACACAGGCAGGCAGAC</td>
</tr>
<tr>
<td>BTC18_Y31C_fwd</td>
<td>GATTGGTCTTTTCGGACACTTTGACGATCGTACTTA</td>
</tr>
<tr>
<td>BTC18_Y31C_rev</td>
<td>TAAATGATCCCTGCAAGAGTCCGGAGAACACCTCAACTC</td>
</tr>
<tr>
<td>BTC18_Y23C_fwd</td>
<td>AGCCCTACTGATCACTGTTGTGGTGAGTTCGCC</td>
</tr>
<tr>
<td>BTC18_Y23C_rev</td>
<td>GCGAAAGAGCAATCACCACCAGCTGATCGTAGGCT</td>
</tr>
</tbody>
</table>
The PCR method used to achieve each mutant was performed as previously reported in Chapter 1, under section “Sortase amenable hBTC cloning, expression, and purification”. Expression and purification procedures used are as reported in Chapter 1, under section “Truncated hBTC protein expression optimization.”

**K_D measurements of Fn3_5.2.18 mutants:**

The binding kinetics of the four Fn3 5.2.18 Cys-mutants were determined using bio-layer interferometry as described in Chapter 1, section titled “K_D measurements using bio-layer interferometry.” Biotinylated hBTC was immobilized on a streptavidin-coated sensor tip and varying concentrations of Fn3 5.2.18 Cys-mutants (50 – 1000 nM) were exposed to the sensor.

**Labeling Fn3_5.2.18 cysteine mutants with thiol-reactive dyes:**

The total reaction volume of each labeling reaction was 400 μL. First, Fn3 5.2.18 Cys-mutants were diluted to 100 μM in 1X PBS containing 0.5 M urea, pH 7.4. The protein (400 μL) was treated with 4μL of 100 mM TCEP for 1 hour at 4°C with mixing. Next, SDS-PAGE was performed under non-reducing conditions to assess the completion of the reducing reaction. If protein containing disulfide bonds were observed, 0.4 μL of 100 mM TCEP was added to the reaction and allowed to proceed for 15 min at 4°C. Next, 4 μL of the thiol-reactive dye (100 mM in DMSO) was added to the protein solution and mixed by inverting the tube several times. The labeling reaction was allowed to proceed overnight in the dark at 4°C with gentle mixing. The next day, the reaction was stopped by addition of 4 μL β-mercaptoethanol (β-ME). The reaction was then transferred to a 0.5 mL 0.22 μm filtering device (Millipore Ultrafree-MC) and centrifuged according to manufacturer’s protocols to remove any precipitate formed during the
reaction. Excess reagents and dye was removed by loading the clarified reaction solution to a pre-equilibrated 5 mL HiTrap desalting column (GE Healthcare Life Sciences) and purified following manufacturer’s protocols. Fractions of 0.5 mL were collected and analyzed by SDS-PAGE (15% acrylamide). Prior to staining with Coomassie blue, the gel was imaged under UV illumination to detect fractions containing dye. Meanwhile, an Amicon® Ultra-4 centrifugal filter unit (3K MWCO, Millipore), was treated overnight with 2 mL 5% Tween solution. The tween solution was then removed and the filter unit was rinsed with dH₂O to remove residual tween. Fractions containing labeled protein was then concentrated using the pre-treated Amicon® Ultra-4 filter unit to 100-200 µL. Total protein concentration was determined using BioRad protein quantification assay, while the total 4-DMN dye concentration was calculated from its absorbance at 440 nm using the extinction coefficient (ε = 8800 M⁻¹cm⁻¹) that had been previously determined. Labeling efficiencies were calculated by dividing the dye concentration by the total protein concentration.

TCEP  -  +

Figure 3-5. Fn3 5.2.18 Cys-mutant before and after treatment with TCEP analyzed on SDS-PAGE (15% acrylamide) visualized with Coomassie blue. * denotes Fn3 monomer.
Fluorescence measurements:

**General procedures.** All fluorescence spectra were performed on a FluoroMax-P instrument (Horiba Jobin Yvon) at 25°C. The samples were prepared in a 100 μL quartz cuvette (Strarna Cells, Inc) and excited at the wavelengths appropriate for the fluorescent dye used. The slit widths were set at 10 nm for excitation and 5 nm for emission. Data points were collected at 5 nm increments with a 0.2 s integration period. All spectra were corrected for background fluorescence by subtracting a blank scan of the buffer system.

**4-DMN labeled Fn3 5.2.18 mutants.** Stock solutions of 4-DMN labeled Fn3 5.2.18 mutants were prepared at 5 μM in 1X PBS containing 0.5 M urea, pH 7.4. A stock solution of hBTC at 70.7 μM was also prepared in the same buffer system. Using these protein stocks, the labeled Fn3 5.2.18 mutant was diluted to 200 nM and exposed to increasing concentrations of hBTC (0, 0.5, 1.0, 2.0, 3.0, 4.0 μM) until signal saturation was reached. The samples were excited at 420 nm and emission collected from 450 – 650 nm. The fold-increase for each Fn3 5.2.18 mutant was determined by comparing the background fluorescence spectra of the labeled Fn3 5.2.18 mutant with the fluorescence spectra containing the highest concentration of hBTC measured. The largest difference in emission is reported as the fold-increase upon hBTC binding.

References:


$^1$H and $^{13}$C NMR spectra
Chapter 4

Labeling studies of hEGF-binding Fn3
**Introduction**

Epidermal growth factor (EGF) was the first component of the ErbB-family to be discovered. EGF was first noted for its role in precocious eyelid opening of newborn mice. This initial observation sparked the investigation into and discovery of the entire ErbB-signaling pathway.\(^1\) Today, EGF is known as a major signaling molecule implicated in a variety of processes including embryonic development and tumorogenesis.\(^2\) Similar to all ErbB-ligands, EGF is expressed as a 1207 amino acid pre-protein consisting of a signal peptide (aa\(^1\)\(^{-}\)22), an extracellular domain containing 9 EGF-motifs (aa\(^22\)\(^{-}\)1032), a transmembrane region (aa\(^1033\)\(^{-}\)1053), and a cytoplasmic tail (aa\(^1054\)\(^{-}\)1207) (Figure 4-1). Uniquely, pro-EGF is the only ErbB-ligand to contain 9 EGF domains in its pro-form with only the EGF domain adjacent to the transmembrane region capable of activating EGFR. Pro-EGF is anchored on the cell surface until activation and released via enzymatic cleavage at two sites to release the mature 5 kDa EGF protein.\(^3\)

![Figure 4-1. Illustration of pro-EGF anchored to the cell membrane. Figure adapted from Schneider, M.R. and E. Wolf 2009.\(^3\)](image-url)
Results and Discussion

4.1. Previous studies

Efforts to generate a fluorogenic biosensor for detection of human EGF (hEGF) began as a collaboration between Dr. Elke Socher, a postdoctoral fellow in the Imperiali lab, and Seymour De Picciotto, a graduate student in the Wittrup lab. Similar to the hBTC biosensor development, the Fn3 scaffold was engineered through YSD to bind selectively to hEGF. After five selection rounds, four unique Fn3 mutants were generated with dissociation constants measured between 5 – 25 nM (as determined by competition analysis on the yeast cell surface) (Table 4-1). Two mutants displaying a Cys residue in the BC loop were of immediate interest because of the potential compatibility with thiol-labeling strategies for fluorophore placement. The strongest binder, Fn3 5.1.07, was chosen for in vitro validation. Expression and purification of hEGF and Fn3 5.1.07 was achieved via a SUMO-fusion tag as described previously. Binding experiments via bio-layer interferometry with immobilized hEGF fit a 1:1 binding model and yielded a $K_D$ of 3 nM, in excellent agreement with that reported on YSD (5 nM).

Table 4-1. Generation 5 Fn3 mutants selected for binding to hEGF through YSD with their respective dissociation constants ($K_D$) reported. The amino acid sequences of the three loop regions are displayed. Negatively charged residues are highlighted in blue and aromatic residues in red.

<table>
<thead>
<tr>
<th>Fn3</th>
<th>BC Loop</th>
<th>DE Loop</th>
<th>FG Loop</th>
<th>$K_D$ (nM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1.01</td>
<td>H S P Y Y R F V N</td>
<td>R S I S</td>
<td>H K W R Y P F</td>
<td>25</td>
</tr>
<tr>
<td>5.1.02</td>
<td>H H P Y Y S Y A R</td>
<td>R S I S</td>
<td>H K W R Y P F</td>
<td>18</td>
</tr>
<tr>
<td>5.1.09</td>
<td>H H P Y Y S Y S C</td>
<td>R S I S</td>
<td>H K W R Y P F</td>
<td>10</td>
</tr>
<tr>
<td>5.1.07</td>
<td>H H P Y D S Y S C</td>
<td>R S V S</td>
<td>H R W R Y P F</td>
<td>5</td>
</tr>
</tbody>
</table>

With the evolved scaffold of the biosensor design in hand, labeling studies were performed to convert the protein-binding scaffold into a biosensor. The first labeling experiment involved labeling the native Cys residue of Fn3 5.1.07 with the four thiol-reactive 4-DMN...
derivatives previously developed in the Imperiali lab. No change in fluorescence signal was detected upon exposure to increasing concentrations of hEGF ligand. Since the Cys residue was located within the BC loop of the Fn3 mutant, which is believed to be part of the interaction surface, it is likely that incorporation of a fluorophore at this position might disrupt ligand binding through steric-hindrance. As such, the selection criteria of subsequent labeling sites were expanded to include residues beyond the loop regions. The Cys residue in Fn3 5.1.07 was mutated to Ala (Fn3 5.1.07b) for the subsequent Cys labeling studies. Preservation of the tight binding to hEGF of the new mutant was confirmed on YSD with a measured $K_D$ of 2.8 nM.

A total of six sites were chosen on Fn3 5.1.07b for labeling with 4-DMN: Y29, Y32, R34, E48, W80, R81 (Figure 4-2A). The six different Cys mutants were achieved through site-directed mutagenesis and subsequently labeled with four thiol-reactive 4-DMN derivatives. Three out of nine mutants showed a fluorescence increase up to 3-fold upon binding to hEGF.

Figure 4-2. (A) Amino acid sequence of Fn3 5.1.07b with Cys mutation sites in red. BC, DE, and FG loop residues are highlighted in green, purple, and orange respectively. (B) Chart summarizing the fluorescence intensity changes measured for each Cys-mutant labeled with each thiol-reactive 4-DMN derivative.
(Figure 4-2B). While these results are promising, a larger fluorescence intensity change of >10-fold was deemed to be necessary for visualizing binding events through fluorescence imaging techniques.

The performance of 4-DMN was then compared to commercially available solvatochromic fluorophores to assess if Fn3 has a hydrophobic region particularly amenable to specific fluorophore structures. The three Fn3 mutants showing the largest fluorescence increase upon ligand binding when labeled with 4-DMN (Y29C, E48C, W80C) were labeled with BADAN (6-bromoacetyl-2-dimethylaminonaphthalene) and EDANS (5-((2-aminoethyl)amino)naphthalene-1-sulfonic acid), and tested for fluorescence changes upon ligand binding. No significant fluorescence increase was observed. These fluorescence results agree with those reported by Loving and Imperiali where 4-DMN yielded the largest fluorescence intensity increase upon ligand binding relative to the other candidates (Figure 4-3).4

4-2. Labeling of Fn3 5.1.07b

Judicious placement of a solvatochromic fluorophore on the Fn3 scaffold is essential to successfully converting the protein-binding scaffold into a fluorogenic biosensor. Since Fn3
5.1.07b is a 101 amino acid protein, and only six sites had been tested during the optimization of the fluorophore placement, three additional sites were selected to further sample the sequence space of Fn3 5.1.07b. Furthermore, among the six sites tested only one was within the BC loop and none within the DE loop. Thus three sites within the BC and DE loops were chosen for mutagenesis and labeling studies: H24, Y26, and S54 (Figure 4-4A). Site-directed mutagenesis was performed to yield the respective Cys-mutants.

\[ \text{VSDVPREDLEV VAAATPTSLLI SWHHPYDYSYS AYYRITYGET} \]
\[ \text{GGNSPQEFYT VPRSVATI SGLKPGVDYT ITVYAVTHRW RYPFPISINY} \]
\[ \text{RTEIDKPSQG S} \]

**Figure 4-4.** (A) Amino acid sequence of Fn3 5.1.07b with Cys mutation sites in red. BC, DE, and FG loop residues are highlighted in green, purple, and orange respectively. (B) Chart summarizing the fluorescence intensity changes measured for each Cys-mutant labeled with each thiol-reactive 4-DMN derivative.

Given that the placement of the mutations were within the loops that make up the binding interface on the Fn3, the mutations may significantly alter the binding affinity of the Fn3 to
hEGF. Thus, the $K_D$ of the binding of each mutant to hEGF was assessed via bio-layer interferometry prior to labeling studies. The measured $K_D$ for each mutant remained in the nM range, which is the desired affinity range for our biosensor. Each mutant was then labeled with the five thiol-reactive 4-DMN derivatives in Chapter 3 for the generation of the hBTC biosensor. No significant fluorescence change in any of the labeled mutants was observed upon binding to hEGF (Figure 4-4B).

From the fluorescence results from the nine Fn3 5.1.07b Cys-mutants labeled with 4-DMN, the maximal signal increase remains at 3-fold. The potential signal increase of the fluorophore has been demonstrated to be over 1000-fold when tested in water versus dioxane. While such a large change in fluorescence cannot be expected when used in a protein system, a signal increase of 100-fold was observed in the calmodulin (CaM) and M13-peptide model system. The crystal structure of calcium-activated CaM with the M13 peptide bound shows a highly buried binding interaction surface, creating numerous hydrophobic contacts. The 4-DMN fluorophore is believed to interact with one of these hydrophobic contacts upon binding to CaM, creating a significant change in local environment to yield the 100-fold increase in fluorescence signal. The interaction surface between the Fn3 scaffold and hEGF, however, is expected to be much smaller and less buried. However, robust signal increases of over 80-fold have also been achieved in systems with small interaction surfaces such as the PDZ-domain binding with a 4-DMAP labeled target peptide. Docking of the labeled peptide onto the PDZ domain binding groove shows the fluorophore to be highly solvent exposed on one face while making hydrophobic contact with the protein on the opposing face. These results indicate that a highly buried interaction surface is not required to generate fluorescence signal changes of up to 80-fold. However, the lack of structural data available for the binding of Fn3 5.1.07b to hEGF
makes the selection and optimization of labeling sites challenging. Thus, given these constraints and the higher level of complexity in the Fn3-hEGF system, a fluorescence increase of 100-fold upon binding is not expected. A conservative 10-fold increase is expected and sufficient to generate a robust biosensor compatible with fluorescence imaging.

One factor that had not been tested thus far is the effect of linker rigidity. The effect of linker length and rigidity on optimizing the fluorescence signal has been well documented.\(^4,7,8\) In the thiol-reactive 4-DMN derivatives used thus far, fairly flexible linkers such as polyethyleneglycol have been used. While this flexibility could be beneficial for some systems, a more rigid linker may be needed in our Fn3-biosensor system. To test this hypothesis, the two best mutants, E48C and W80C, displaying a 3-fold fluorescence increase were labeled with 4-DMN using a more rigid linker. More rigidity between Fn3 and 4-DMN was achieved through formation of a triazole ring via the versatile copper-catalyzed azide-alkyne (CuAAC) reaction. A bioorthogonal alkyne methionine analogue (homopropargylglycine) was introduced site-specifically into Fn3 using a Met auxotroph strain of \textit{E. coli} developed by Tirrell and coworkers.\(^9\) The required \textit{E. coli} strain was obtained from the Tirrell Lab and the two sites of interest on Fn3 5.1.07b were mutated to give E48M and W80M. Following the published procedures from van Hest \textit{et al} (2000), Y48M and W80M containing homopropargylglycine at the Met site were expressed and purified from inclusion bodies and refolded.\(^9\) Two azide-derivatives of 4-DMN (4-DMN-I-azide and 4-DMN-II-azide) were reacted with the alkyne functionalized proteins under reaction conditions described by Finn and co-workers.\(^12\) The labeled proteins were then subjected to fluorescence measurements as described above to detect fluorescence intensity changes upon ligand binding. Similar fluorescence results were observed
between the Cys and CuAAC labeling strategies of the two Fn3 mutants, suggesting linker rigidity holds little influence in fluorophore optimization within this system (Figure 4-5).

![Diagram of 4-DMN azide derivatives](image)

**Figure 4-5.** (A) 4-DMN azide derivatives. (B) Chart summarizing the fluorescence intensity changes measured for Fn3 5.1.07b labeled with different linkers to 4-DMN at positions E48 and W80.

One important observation regarding the fluorescence data collected on the nine Fn3 5.1.07b mutants labeled with 4-DMN is high background fluorescence, suggesting 4-DMN may be nonspecifically associating with hydrophobic residues exposed on the protein surface, leading to non-ligand induced fluorescence emission. One way to circumvent this drawback is to attach the fluorophore to a more flexible region of the protein so that the fluorophore can freely associate with solvent. However, labeling at the flexible loop regions of Fn3 5.1.07b also suffered from high background fluorescence, suggesting the labeling at the loops fail to provide 4-DMN with enough rotational freedom to associate with solvent. Thus we considered labeling the tail regions of Fn3. The crystal structure of Fn3 showed the N-terminus as a promising labeling site because of its flexibility and proximity to the Fn3-EGF interaction surface. Thus
three sites on the N-terminal tail were mutated to Cys for labeling with 4-DMN: V1, S2, D3. Indeed the background fluorescence of all three mutants was well below that of previously labeled Fn3 mutants. Importantly, one of the three mutants, V1C, exhibited a 3-fold fluorescence increase upon binding to hEGF (Figure 4-6). The results from the V1C mutant support our working hypothesis that the N-terminus may be a promising site for fluorophore placement because of its proximity to the interaction surface.

![Figure 4-6](image)

**Figure 4-6.** Chart summarizing the fluorescence increase measured for Fn3 5.1.07b mutants, V1C, S2C, and D3C labeled with 4-DMN.

### 4-3. Investigating the role of the N-terminal tail of Fn3 5.1.07b

Labeling of the N-terminal of Fn3 with 4-DMN was also tested during the development of the hBTC biosensor in Chapter 3. In that case, the 4-DMN labeled Fn3 had a measured decrease in signal of up to 4-fold upon binding to hBTC. Coupled with the fluorescence results from the V1C mutant of Fn3 5.1.07b, we postulated that the N-terminal tail might be directly involved in the Fn3-ligand binding interface. To test this hypothesis, a truncated form of Fn3 5.1.07b (Fn3 5.1.07t) was generated by removing nine residues from the N-terminus of Fn3.
5.1.07b. The binding affinity of this new mutant to hEGF was then measured using bio-layer interferometry and compared to the measured $K_D$ of the parent protein. Interestingly, removal of these nine residues not only hindered the solubility and stability of the Fn3 mutant, but also significantly reduced the Fn3's binding affinity to hEGF. Fn3 5.1.07t had a measured $K_D$ of 1 $\mu$M, almost 3-orders of magnitude larger than the measured $K_D$ of the original Fn3 mutant (3 nM).

The N-terminus of the Fn3 scaffold being involved in the binding interface was not predicted by previous studies even when structural information was available for the Fn3 mutant. To our knowledge, there has only been one report of the N-terminus of an Fn3 mutant affecting its ability to bind its target ligand. However, in that case removal of 8 residues from the N-terminus improved the binding kinetics of the Fn3. Given the active role of the N-terminus in facilitating the binding kinetics of Fn3 5.1.07b to hEGF, we hypothesized that the binding event induces a large change in local environment around the N-terminal residues. Thus, by replacing the N-terminal amino acids with an unnatural amino acid such as 4-DMNA, this change in local environment would translate into a fluorescent output that signals the occurrence of the binding event. Importantly, the size of the 4-DMN chromophore is similar in size to the natural amino acid side chain of tryptophan, allowing the natural structure of the N-terminus to be minimally perturbed.

4-4. Tail labeling Fn3 5.1.07b

The ideal method to incorporate 4-DMNA into the protein structure of Fn3 5.1.07b would be through translational labeling techniques that have been developed to directly incorporate unnatural amino acids during protein translation. However, such techniques require the
generation of an orthogonal t-RNA/aaRS (aminoacyl-tRNA synthetase) pair capable of charging 4-DMNA onto a novel t-RNA. Since, such a t-RNA/aaRS pair has not been developed for 4-DMNA and development of such systems is labor intensive, an alternative method to incorporate 4-DMNA was implemented. Sortase mediated ligation (SML) has emerged as a facile method of protein-peptide ligation. The sortase enzyme recognizes a unique amino acid motif (LPXTGG) and is able to selectively ligate two fragments containing the recognition motif through a native peptide bond.\(^\text{15}\) Thus by incorporating LPXTGG at the C-terminus of a peptide containing 4-DMNA and GG into the N-terminus of Fn3 5.1.07t, 4-DMNA can be selectively incorporated into the protein structure of Fn3 5.1.07b with minimal perturbation (Scheme 4-1). Importantly, most of the native sequence of Fn3 5.1.07b is maintained in the SML product, in particular the Pro amino acid at the N5 position.

![Scheme 4-1. General scheme used to label Fn3 5.1.07t with 4-DMN containing peptides via sortase mediated ligation.](image)

<table>
<thead>
<tr>
<th>X</th>
<th>S</th>
<th>D</th>
<th>L</th>
<th>P</th>
<th>R</th>
<th>T</th>
<th>G</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>V</td>
<td>X</td>
<td>D</td>
<td>L</td>
<td>P</td>
<td>R</td>
<td>T</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>W</td>
<td>X</td>
<td>D</td>
<td>L</td>
<td>P</td>
<td>R</td>
<td>T</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>V</td>
<td>S</td>
<td>X</td>
<td>L</td>
<td>P</td>
<td>R</td>
<td>T</td>
<td>G</td>
<td>G</td>
</tr>
</tbody>
</table>

Table 4-2. Sortase amenable peptide sequences containing 4-DMNA. Positions where 4-DMNA were incorporated are denoted by a green X.
A set of four sortase amenable peptides containing 4-DMNA was synthesized through solid-phase peptide synthesis (SPPS) (Table 4-2). 4-DMNA is incorporated at the N1, N2, or N3 position while the native amino acids were maintained in the remaining positions for three peptides. The fourth peptide introduces an additional aromatic residue, Trp, at the N1 position to potentially aid with the formation of a hydrophobic contact with hEGF upon binding. Each peptide was ligated onto Fn3 5.1.07ts, a sortase amenable variant of Fn3 5.1.07t. To determine if ligation of this “synthetic tail” onto Fn3 5.1.07ts could restore its binding kinetics to hEGF, two labeled mutants were characterized using bio-layer interferometry. Both WXD and VXD labeled Fn3 5.1.07ts mutants displayed significantly improved binding kinetics ($K_D$ 50 nM and 7 nM respectively) comparable to that of native Fn3 5.1.07b.

Fluorescence measurements of the labeled Fn3 mutants were then performed to measure any changes in signal intensity upon binding hEGF. One labeled mutant, VSX-Fn3, showed a 2-fold fluorescence increase upon binding hEGF while the other three mutants showed no significant change (Figure 4-7). These results differ significantly from those measured of the Cys-labeled Fn3 5.1.07b mutants reported previously. In the Cys-mutants, only labeling of the N1 position with thiol-reactive 4-DMN-I and 4-DMN-IV showed a 2-fold and 3-fold fluorescence increase upon binding hEGF respectively. Labeling of D3C with thiol-reactive 4-DMN reagents did not yield any significant fluorescence change. These results suggest that the two labeling strategies present the 4-DMN fluorophore in considerably different environments on Fn3 5.1.07b. Moreover, the thiol-labeling strategy positioned the 4-DMN chromophore at a minimum of $\sim$5 carbon-carbon bonds away from the peptide backbone of the Fn3 protein. The tail-tagging strategy on the other hand, presents the chromophore only one methylene group away from the peptide backbone. Since the linker length may be critical to optimization of the
fluorescence signal, the effect of extending the chromophore from the peptide backbone was investigated.

![Figure 4-7. Chart summarizing the relative fluorescence signal measured for each 4-DMN labeled Fn3 5.1.07ts mutant in the presence (red) and absence (blue) of hEGF.](image)

Two new 4-DMNA derivatives, 4-DMNA2 and 4DMNA3, containing 2 and 3 methylene groups respectively were synthesized by Dr. Andrew Krueger to be compatible with Fmoc-based SPPS (Figure 4-8). Both derivatives were then incorporated into two sets of four sortase amenable peptides as described previously and ligated onto Fn3 5.1.07ts. The binding kinetics of each labeled mutant was then assessed using bio-layer interferometry. Unexpectedly, the association curves of the labeled mutants deviate significantly from that of a 1:1 binding model. However, the affinity for hEGF seems minimally perturbed because binding was easily measured within nM concentrations. The poor fit to a 1:1 binding model could be explained if the labeled Fn3 mutants contained a mixture of labeled and un-labeled Fn3 5.1.07ts. However, most mutants
showed high labeling efficiencies (> 80%), thus the presence of two protein constructs in the sample cannot fully explain the unexpected kinetics. Further investigation of the binding stoichiometry via gel filtration chromatography or crystallography will be necessary to elucidate why the association curves deviate from the expected 1:1 binding model.

Since the ability of the labeled mutants to bind to hEGF was not significantly altered, fluorescence measurements were still performed on all labeled Fn3 mutants. Two mutants labeled with 4-DMNA2 at the N2 and N3 positions (VXD2-Fn3, VSX2-Fn3) showed ~2-fold fluorescence increase upon binding hEGF, while the others showed no significant fluorescence change (Figure 4-9).

One important observation from this set of data is the comparison of the background fluorescence of labeled Fn3 mutants with that of the 4-DMN M13-peptide. The relative fluorescence of the 4-DMN M13-peptide is ~0.2, which is over 15-fold lower than the background fluorescence of any 4-DMN labeled Fn3 mutant. A key characteristic of the 4-DMN fluorophore was its exceedingly low background fluorescence in the presence of polar aqueous solvents such as water. Indeed when the M13-peptide was labeled with 4-DMNA, it had an extremely low background fluorescence signal. However, this no longer seems to be true when 4-DMNA is incorporated into Fn3 5.1.07ts. The background fluorescence of labeled Fn3 5.1.07ts

Figure 4-8. Chemical structures of 4-DMNA2 and 4-DMNA3.
is already 15 to 40-fold above that of the 4-DMN M13-peptide. Assuming the CaM/M-13 peptide system represents the maximal fluorescence output of 4-DMN in a protein system, the background fluorescence of labeled Fn3 mutants already occupies 15-40% of the maximal signal potential of the fluorophore. This means even if 4-DMN associates with the ideal hydrophobic pocket upon binding hEGF, its fluorescence signal cannot increase more than 2-3 fold above its background fluorescence. This may explain why we have not been able to measure a fluorescence increase of >3-fold in any of the 4-DMN labeled Fn3 mutants tested thus far. Further efforts focused on lowering background fluorescence either by introducing additives such as BSA, or exploring alternative Fn3 mutants or protein-binding scaffolds are needed to address this issue.

Figure 4-9. Chart summarizing the relative fluorescence signal measured for each 4-DMNA labeled Fn3 5.1.07ts mutant in the presence (red) and absence (blue) of hEGF. XSD2-Fn3, VXD2-Fn3, VSX2-Fn3, and WXD2-Fn3 were labeled with 4-DMNA2 at position X. XSD3-Fn3, VXD3-Fn3, VSX3-Fn3, and WXD3-Fn3 were labeled with 4-DMNA3 at position X.
Experimental

Cloning, expression, and purification of Fn3 5.1.07b cysteine mutants:

The gene encoding Fn3 5.1.07b was obtained from Dr. Elke Socher in the pE-SUMO vector. Site directed mutagenesis was applied to the SUMO-Fn3 5.1.07b construct to yield the following mutants: Y26C, H24C, and S54C. The primers used to generate mutant S54C were designed and ordered by Dr. Elke Socher. The primers used to generate mutants Y26C and H24C were ordered PAGE-purified from Sigma-Aldrich on the 25 mmol scale. The primer sequences are given below.

<table>
<thead>
<tr>
<th>Mutant</th>
<th>Forward Primer</th>
<th>Reverse Primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y26C</td>
<td>TGATCAGCTGGCATCATTGCGACTCTTATTTCG</td>
<td>CGAAAAGAGTCGCAAGGATGATGCAGCTGATCA</td>
</tr>
<tr>
<td>H24C</td>
<td>CAGCCTACTGATCAGCTGGCATCATTGCTTACGACTTATT</td>
<td>AATAAGAGCTGAAGGACAATGCCAGCTGATCAGTAGGCTG</td>
</tr>
</tbody>
</table>

The PCR method used to achieve each mutant was performed as previously reported in Chapter 1, section “Sortase amenable hBTC cloning, expression, and purification.” Expression and purification procedures used are as reported in Chapter 1, section “hBTCΔ30 protein expression and optimization.”

Cloning SUMO-Fn3 5.1.07b E48M and W80M:

The gene encoding Fn3 5.1.07b was obtained from Dr. Elke Socher in the pE-SUMO vector. Whole plasmid PCR was applied to the SUMO-Fn3 5.1.07b construct following procedures previously described in Chapter 1, section “Sortase amenable hBTC cloning, expression, and purification.” The primers used were ordered PAGE-purified from Sigma-Aldrich on the 25 mmol scale. The primer sequences are given below.
With the mutations in place, PCR was performed to amplify the gene encoding SUMO-Fn3 5.1.07 E48M and SUMO-Fn3 5.1.07 W80M and introduce two flanking restriction sites (Sphl, KpnI). The PCR method used was described previously Chapter 2, section “Sortase amenable hBTC cloning, expression, and purification.” The primers used were ordered from Sigma-Aldrich on the 25 mmol scale and purified by desalting. The primer sequences are given below.

\[
\begin{align*}
\text{pQ80-L fwd} & \quad \text{ATAT GCATGC ATGGGTCATCACCAT} \\
\text{pQ80-L rev} & \quad \text{TGTA GGTACC TCA GGA TCC CTG GGA TGG}
\end{align*}
\]

The amplified PCR products were then purified using the Wizard SV gel and PCR clean-up kit (Promega). A double digest of the purified PCR products and pQ80-L vector was performed using Sphl, KpnI (NEB) simultaneously following the manufacturer’s protocol. Subsequent purification and ligation of the PCR fragments into the pQ80-L vector were performed as previously described in Chapter 2, section “hBTC and hBTC\text{Δ30} cloning.”

**Test Expression of Fn3 5.1.07b E48M and W80M:**

**Buffers and media:**

**10X M9 Salts:**

- 60 g Na\textsubscript{2}HPO\textsubscript{4}
- 30 g K\textsubscript{2}HPO\textsubscript{4}
- 10 g NH\textsubscript{4}Cl
- 5 g NaCl
Dissolved the indicated salts in approximately 800 mL Milli-Q water, then adjusted volume to 1 L.

**5X amino acid (+Met).** Each of the 20 natural L-amino acids was dissolved in Milli-Q water to a final concentration of 5 M. This solution was stored at 4°C and used within 1 week.

**5X amino acid (-Met).** Except methionine, each of the remaining 19 natural L-amino acids were dissolved in Milli-Q water to a final concentration of 5 M. This solution was stored at 4°C and used within 1 week.

**1X M9AA+ media:**

- 100 mL 10X M9 salts
- 200 mL 5X amino acids (+Met)
- 1 mL 1 M MgSO₄
- 100 µL 1 M CaCl₂
- 100 µL 10 mg/mL thiamine chloride
- 4 mL 0.5 mg/mL glucose

The M9 salts and amino acid solutions were added to 696 mL Milli-Q water and sterilized by autoclaving. Once the solution had cooled, the remaining components were sterile filtered and added.

**1X M9AA- media:**

- 100 mL 10X M9 salts
- 200 mL 5X amino acids (-Met)
- 1 mL 1 M MgSO₄
- 100 µL 1 M CaCl₂
- 100 µL 10 mg/mL thiamine chloride
- 4 mL 0.5 mg/mL glucose
The M9 salts and amino acid solutions were added to 696 mL Milli-Q water and sterilized by autoclaving. Once the solution had cooled, the remaining components were sterile filtered and added.

Expression procedures. A stock of the M15MA E. coli strain was obtained from Dr. Andrew Krueger. The pQ80-L plasmids containing the gene for SUMO-Fn3 5.1.07 E48M and SUMO-Fn3 5.1.07 W80M were transformed into M15MA following procedures previously described in Chapter 2 for BL21-RIL cells. One colony of each Fn3 mutant was picked and amplified in an overnight culture (5 mL M9AA+ media). The following day, 1 mL of this culture was used to inoculate three 14 mL cultures of M9AA+ media. The cells were grown at 37°C with shaking (225 rpm) until reaching an OD$_{600}$ ~0.9 – 1.0. The cells were then pelleted by centrifugation (4000 g, 20 min, 4°C) and then washed with sterile 0.9% NaCl (2x). The cells were resuspended in 15 mL M9AA- media and allowed to grow for 15 min at 37°C with shaking (225 rpm). Then, L-homopropargylglycine (HPG) (200μL, 10 mg/mL) was added to one culture of each Fn3 mutant. To a second culture of each mutant, L-methionine (12 μL, 50 mg/mL) was added to serve as positive expression control. No amino acid was added to the third culture to serve as a negative control. The cultures were grown in a 37°C incubator (225 rpm) for another 10 min before inducing with 1 mM IPTG. Expression was allowed to proceed for 6 hours at 37°C with shaking (225 rpm). A sample from each culture (60 μL) was removed and mixed with SDS-loading dye containing β-ME. The cells were lysed by boiling the cell and dye mixture at 100°C for 10 min. All samples (8 μL) were then loaded onto a SDS-PAGE gel (15% acrylamide) and resolved at 150 V, then stained with Coomassie Blue (Figure 4-10).
Figure 4-10. Protein profile of Fn3 5.1.07b E48M and W80M overexpression analyzed by SDS-PAGE (15% acrylamide). Bands visualized using Coomassie Blue. * denotes SUMO-Fn3 overexpression.

**Fn3 5.1.07b E48M and W80M expression and purification:**

Both Fn3 methionine mutants were expressed with a SUMO-tag as described in the previous section. Expression was performed on a 0.5 L scale. Cells were harvested by centrifugation and lysed following procedures described previously (Chapter 2, section “hBTCA Δ30 protein expression and optimization”). Fn3 5.1.07b E48M and W80M were purified from inclusion bodies following procedures described in Chapter 2, section “hBTC and hBTC Δ30 protein expression and purification”. Removal of the SUMO-tag was performed in 50 mM HEPES, 150 mM NaCl, 2 M urea, pH 7.4.

**Copper-catalyzed azide-alkyne cycloaddition:**

The CuAAC reaction between Fn3 5.1.07b methionine mutants and 4-DMN azide was performed following a previously published protocol. Briefly, Fn3 5.1.07b E48M and W80M were reacted at 5 – 20 μM. Tris(3-hydroxypropyltriazolylmethyl)amine (THPTA), and two 4-DMN azide derivatives were obtained from Dr. Andrew Krueger and used without further
purification. The total reaction volume was 500 μL. THPTA was mixed with CuSO₄ immediately prior to the reaction. The reaction was set up by adding the following components to Fn3 5.1.07b Met-mutant in the given order.

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>μM</td>
<td>4-DMN azide</td>
</tr>
<tr>
<td>0.5:0.1</td>
<td>mM</td>
<td>Premixed THPTA:Cu (5:1 molar ratio)</td>
</tr>
<tr>
<td>5</td>
<td>mM</td>
<td>Aminoguanidine</td>
</tr>
<tr>
<td>5</td>
<td>mM</td>
<td>Sodium ascorbate</td>
</tr>
</tbody>
</table>

The reaction was allowed to proceed at room temperature for 45 min. The reaction was then stopped and purified using a 0.5 mL Zeba spin desalting column (7K MWCO) following manufacturer’s protocol for buffer exchange. The purified proteins were analyzed on SDS-PAGE (15% acrylamide). Prior to staining with Coomassie blue, the gel was imaged under UV illumination to detect fractions containing dye. Total protein concentration was determined using BioRad protein quantification assay, while the total 4-DMN dye concentration was calculated from its absorbance at 440 nm using the extinction coefficient (ε = 8800 M⁻¹cm⁻¹) that had been previously determined. Labeling efficiencies were calculated by dividing the dye concentration by the total protein concentration.

4-DMN amino acid derivatives:

Fmoc-protect 4-DMN amino acid derivatives were obtained from Dr. Andrew Krueger and used without further purification in solid phase peptide synthesis.

Synthesis of sortase amenable peptide series containing 4-DMNA:

Peptides were prepared using a combination of automated and manual synthesis following standard Fmoc-solid phase peptide synthesis (SPPS). The sortase recognition sequence
LPRTGGG was prepared by automated SPPS (ABI 433A Peptide Synthesizer) with Fmoc-PAL PEG resin (Novabiochem, 0.18 mmol/g) on a 0.25 mmol scale. The following Fmoc-protected amino acids were used: Fmoc-Gly-OH, Fmoc-Thr(tBu)-OH, Fmoc-Arg(Mtr)-OH, Fmoc-Pro-OH, and Fmoc-Leu(tBu)-OH. N-methyl-2-pyrrolidone (NMP) and dichloromethane (DCM) were used as solvents. PyBop and HOBT were as coupling reagents, DIPEA was used as the organic base. Each amino acid was added to an amino acid cartridge along with the two coupling reagents (4 equivalents relative to resin loading per coupling). Acyl capping was employed after each coupling step. The final peptide was left on resin with all side chain protecting groups and N-terminal Fmoc-protecting group in place. The peptide on resin was transferred to a glass vial and dried under vacuum overnight, then stored at -20°C until use. This peptide resin stock was used to synthesize all 4-DMNA containing peptides used to label the N-terminus of EGF-binding Fn3.

A total of 12 peptides were synthesized to contain the sortase recognition sequence and one variant of 4-DMN amino acid (4-DMNA-1, 4-DMNA-2, 4-DMNA-3). Their sequences are as follows:

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>4-DMNA-1</td>
<td>S</td>
<td>D</td>
<td>L</td>
<td>P</td>
<td>R</td>
<td>T</td>
</tr>
<tr>
<td>4-DMNA-1</td>
<td>V</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-DMNA-1</td>
<td>W</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-DMNA-2</td>
<td>V</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-DMNA-2</td>
<td>W</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-DMNA-3</td>
<td>V</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-DMNA-3</td>
<td>W</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

All peptides were made manually by coupling the three designated amino acids to the LPRTGGG peptide made using automated synthesis as described above. Manual synthesis was
performed as described in Chapter 2, section titled “Synthesis of GGGYKK(biotin)T peptide.” Each peptide was made on a 0.01 mmol scale in a 3 mL polypropylene syringe (Torviq, ref SF-0250). All amino acids and coupling reagents were used at 6 molar equivalences relative to the resin loading per coupling except for the 4-DMN amino acids. 4-DMN amino acids were coupled at 3 equivalences per coupling. Once all desired amino acids had been coupled, the peptide was cleaved with 1 mL TFA:thioanisole:TIPS:ddH₂O (90:5:2.5:2.5 v/v/v/v) for 4 hours.

The peptides were purified using reverse phase HPLC as described in Chapter 2, section titled “Synthesis of GGGYKK(biotin)T peptide.” HPLC conditions were 20% acetonitrile in water with 0.1% TFA for 5 min, followed by a linear gradient from 20% to 90% acetonitrile in water with 0.1% TFA for 25 min.

**Fn3 5.1.07t**

Purified Fn3 5.1.07t was obtained from Dr. Elke Socher and used without further modification in bio-layer interferometry experiments. The amino acid sequence is given below.

Fn3 5.1.07t:

```
VVAATPTSLLISWHPYDSYYYRITYGETGGNSPVQEFTVPRCVSTATISGLKPGVDYTITVYAVTHRWR
YPFPISINYRTEIDKPSQGS
```

**Fn3 5.1.07ts expression and purification:**

The gene encoding Fn3 5.1.07ts was obtained from Dr. Elke Socher in the pE-SUMO vector. The plasmid was transformed into BL21-RIL (NEB) cells and expressed. Transformation, expression, and purification procedures were performed as described in Chapter 2, section “hBTC₃₅₀ protein expression optimization.” Except the final dialysis step was done in 50 mM
HEPES, 150 mM NaCl, 0.5 M urea, pH 7.4. The SUMO-tag was not removed at this stage. SUMO-Fn3 5.1.07ts was flash frozen in 10 mL aliquots and stored at -80°C. The amino acid sequence of Fn3 5.1.07ts is given below.

Fn3 5.1.07ts:
GGVVAATPTSLISWHHPYDSYSAAYRITYGETGGNSPVQEFTVPRCVSTATISGLKPGVDYTITYAVTHWRWYPFPISINRYTEIDKPSQGS

**Sortase mediated ligation of 4-DMN peptides to Fn3 5.1.07ts:**

A one-pot reaction combining SML and SUMO-tag cleavage was performed to ligate 4-DMN peptides to the N-terminus of Fn3 5.1.07ts. Each labeling reaction used 2-3 mL SUMO-Fn3 5.1.07ts (13 – 50 μM). All protein solutions were kept at 4°C. SUMO protease (45 μM, 1% v/v) was first added to SUMO-Fn3 5.1.07ts. Then the SML components were added as follows:

<table>
<thead>
<tr>
<th>Component</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUMO-Fn3 5.1/07ts</td>
<td>1 eq</td>
</tr>
<tr>
<td>Sortase A triple mutant</td>
<td>3 eq</td>
</tr>
<tr>
<td>4-DMN LPRTGG peptide</td>
<td>3 eq</td>
</tr>
<tr>
<td>2 M CaCl₂</td>
<td>10 mM</td>
</tr>
</tbody>
</table>

The reaction was allowed to proceed for 4-6 hours at 4°C with gentle mixing in the dark. The reaction was then stopped and purified by flowing over 0.2 mL pre-equilibrated Ni-NTA resin (x3). Ni-NTA purification was done in dimmed lighting at room temperature with cold buffers. The flow through was collected. The resin was then washed with 0.8 mL 50 mM HEPES, 150 mM NaCl, 10 mM imidazole, pH 7.4. A second wash was done with increased imidazole concentration (20 mM). A third wash with 50 mM imidazole was then performed. Lastly, the SUMO-protease and sortase were eluted from the resin using 1 mL 50 mM HEPES, 150 mM NaCl, 300 mM imidazole, pH 7.4. All fractions were collected and analyzed on SDS-PAGE.
Prior to staining with Coomassie blue, the gel was imaged under UV illumination to detect fractions containing 4-DMN. Meanwhile, an Amicon® Ultra-0.5 centrifugal filter unit (3K MWCO, Millipore), was treated overnight with 2 mL 5% Tween solution. The tween solution was then removed and the filter unit was rinsed with dH2O to remove residual tween. Fractions containing labeled protein was then concentrated using the pre-treated Amicon® Ultra-4 filter unit to 50-150 μL. The total 4-DMN dye concentration was calculated from its absorbance at 440 nm using the extinction coefficient ($\varepsilon = 8800 \text{ M}^{-1} \text{cm}^{-1}$) that had been previously determined. The relative labeling efficiencies were analyzed by LCMS (Agilent Zorbax 300SB C3 column: 2.1 x 150 mm, 5 μm, linear gradient: 1-61% acetonitrile in water with 0.1% formic acid over 15 min, flow rate 0.4 mL/min). The area under the peaks corresponding to labeled and unlabeled Fn3 5.1.07t5 were compared to determined the approximate labeling efficiency of each reaction. Efficiencies ranged from 50% to quantitative labeling.

**Generating hEGF and biotinylated hEGF:**

Human EGF was obtained from Dr. Elke Socher and used without further purification in fluorescence experiments. Sortase amendable hEGF was obtained from Dr. Andrew Krueger and used without further purification in sortase mediated ligation. Human EGF was biotinylated following procedures described previously (Chapter 2, section "sortase mediated ligation").

**$K_D$ determination using bio-layer interferometry:**

The $K_D$ of Fn3 5.1.07t binding to hEGF was performed as described in Chapter 2, section “$K_D$ determination using bio-layer interferometry.” Except biotinylated hEGF was immobilized
on the streptavidin sensor tip instead of hBTC. The $K_D$ of 4-DMN labeled Fn3 5.1.07ts binding to hEGF was also determined following the same procedures.

References:


Acknowledgements

This research was supported by NIH EB010246-02 (Griffith), Perfused 3D Tissue Surrogates for Complex Cell-Cell Communication Systems, and NSF CBET-0939511 (Kamm), Science and Technology Center: Emergent Behaviors of Integrated Cellular Systems. I thank the Griffith lab and the Wittrup lab for being wonderful collaborators and Dr. Elke Socher for teaching me everything I know. The Department of Chemistry Instrumentation Facility (NSF CHE-9808061, DBI-9729592, and CHE-0234877) is also gratefully acknowledged.