

**UCSF**

**UC San Francisco Electronic Theses and Dissertations**

**Title**

DNA Programmed Assembly of Proteins: Applications in Antibody Engineering and Epidermal Growth Factor Receptor Signaling

**Permalink**

<https://escholarship.org/uc/item/3w5002f1>

**Author**

Liang, Samantha Isabel

**Publication Date**

2016

Peer reviewed|Thesis/dissertation

DNA-Programmed Assembly of Proteins: Applications in Antibody  
Engineering and Epidermal Growth Factor Receptor Signaling

by

Samantha Isabel Liang

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Biochemistry and Molecular Biology

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO



Copyright 2016

By

Samantha Isabel Liang

# ACKNOWLEDGEMENTS

Getting a PhD was like running a marathon. As I am now crossing the finish line, I am happy to have the time to look back, reflect on my how I made it to the end, and feel proud of my accomplishment. Without the proper training, stamina, and a network of support, I know that I would not have made it to the end. Many people made this possible and I would really like to thank them now.

Training: Zev Gartner taught me much about being a scientist and has been an amazing mentor during my PhD. In his lab, I was encouraged to take risks, collaborate with people all over UCSF and other institutes, and to explore and be creative on my projects. I want to thank Zev for all of his support and patience throughout the years, and also for being a great scientific role model. I also want to thank my thesis committee, Prof. Orion Weiner, Prof. Jim Wells, and Prof. Natalia Jura, for providing great guidance and encouragement. Natalia in particular has provided me with indispensable guidance for my work on EGFR signaling and I want to thank her for her time and generosity over the years. I want to thank David Rabuka at Catalent-West, formerly Redwood Biosciences, for being a collaborator from the start on the protein-DNA conjugation and antibody engineering project, and enabling those projects by helping us use the aldehyde tag technology to its full potential. Finally, I want to thank Prof. Charly Craik for all of his support on the uPAR antibody project. I am grateful that I got to work on these exciting projects and I could not have done it without the support of all of these people and labs mentioned.

Stamina: When things are going well on the lab, it is easy to push and be relentless with executing experiments – but there is no doubt that this can be just as exhausting as it is exciting, particularly when you finally run into an obstacle or feel that you have run out of creativity. And at many times I felt like I had run out of breath. So I want to thank all of my collaborators, especially Bettina Van Lengerich, Efrat Harel, Jesse McFarland, Daeha Seo, and Kelsie Eichel, for breathing life into my projects and making them more exciting with their expertise and experiments. I also want to thank my labmates in the Gartner lab who made it such a wonderful and supportive place to work. Finally, I want to thank my classmates in the Tetrad Graduate Program, who have been the most fun and talented cohort that I could have imagined. It made the grad school experience much more fun to have friends in other labs around UCSF who were also going through similar trials and tribulations at the same time. I'm grateful that I was able to interact with so many amazing people at UCSF - thank you all for keeping me going and providing me with extra boosts of energy during this long run.

Support: First and foremost, I want to thank my family. My parents, David and Jenny, encouraged me from a young age to value my education, pursue my curiosities, and foster my creativity both in artistic endeavors and in problem solving. But above all, they always stressed that nothing is more important than happiness and kindness. So while I've chugged along these few years, I always tried my best to remember to be kind, be grateful, and be happy. And it's been a great comfort to know that my parents, and my brother Kyle, would be proud of me and be my safety net no matter what happened in graduate school. I also want to thank all of my friends from childhood, high school, college, and beyond, who weren't in graduate school with me, but always listened to my problems and helped me have a lot of fun in my life

outside of lab. Thank you all for boosting my confidence, encouraging me to do things out of my comfort zone, and for all of the love and support throughout the years. And finally, a huge thank you to my husband, Robert Weber, who is my largest and most dependable source of love and laughter. He has supported me in too many ways to describe and I know that graduate school would have been much harder had I not met him.

And finally, because it is required, I also need to mention in these acknowledgements that part of this dissertation is a reproduction of material previously published and contains contributions from collaborators listed therein. Specifically, Chapter 2 is reproduced in part with permission from Liang SI, McFarland JM, Rabuka D, Gartner ZJ. A Modular Approach for Assembling Aldehyde-Tagged Proteins on DNA Scaffolds. *J Am Chem Soc*, 136(31), 10850-3, 2014. Chapter 5 is reproduced in part with permission from Weber RW, Liang SI, Selden NS, Desai TA, Gartner ZJ. Efficient Targeting of Fatty-Acid Modified Oligonucleotides to Live Cell Membranes through Stepwise Assembly. *Biomacromolecules*, 15(12), 4621-5, 2014. They are the most polished chapter of this dissertation. Please excuse the rough edges on the other chapters and look for more polished versions of these stories when they are published.

Reflecting on the last few years, I can't believe how much I have changed. I've grown so much as a scientist and as a person. Now that this marathon is done, I feel so proud that I trained, I persevered, and I made it to the end with happiness and gratefulness. I am looking forward to the next challenge.

# DNA Programmed Assembly of Proteins: Applications in Antibody Engineering and Epidermal Growth Factor Receptor Signaling

Samantha I. Liang

---

## ABSTRACT

This thesis details my work, both published and ongoing projects, on assembling proteins with modular DNA scaffolds in order to develop new strategies for engineering antibodies and probing cellular signaling complexes. The organization of cell surface receptors on a nanometer length scale can influence and regulate a cell's functional response to signals. My goal was to systematically explore the effect of receptor complex composition, valency, and geometry on cell signaling. Towards this aim, I used modular DNA-based scaffolds to combinatorially control the position of attached proteins with low nanometer resolution. This strategy required a simple and modular technique for site-specifically conjugating synthetic oligonucleotides to proteins. My approach uses the aldehyde tag, which can be genetically incorporated into proteins at either protein terminus or in an internal loop, which provides the potential for orientational control of proteins on DNA scaffolds. We designed and optimized four different reactions that generate DNA-protein conjugates to provide flexibility in linker chemistry. The protein-DNA conjugates can be efficiently assembled into structures with

greater valency and complexity. We are expanding our DNA scaffold libraries to generate collections of macromolecular assemblies varying in valency and architecture. These protein-bearing DNA-scaffolded multivalent probes may have novel activities as antibodies with unique specificities and biological activities, such as the ability to deliver protein-based effectors, increasing avidity, or modulating specificity. In one application, we are using this strategy to prepare and screen multivalent DNA-scaffolded biparatopic antibodies for inhibiting urokinase plasminogen activator receptor in triple negative breast cancer lines. Additionally, we are using these modular DNA scaffolds to investigate the effect of nanoscale organization on the signaling profiles of the Epidermal Growth Factor Receptor (EGFR) family by systematically guiding the assembly of signaling complexes on cells. By dimerizing EGFR in the absence of ligand using our DNA scaffolds, we discovered that although dimerization of EGFR is sufficient for autophosphorylation and recruitment of adaptor proteins Grb2 and SOS, it is not sufficient for Ras activation, activation of the AKT or MAPK signaling pathways, nor for the concentration of the receptor into clathrin coated pits and endocytosis.

# TABLE OF CONTENTS

<b>Chapter 1: Introduction .....</b>	<b>1</b>
Why study nanoscale organization?.....	2
Nanoscale organization of EGFR .....	4
DNA as an ideal nanoscale scaffold.....	6
Spatial mutation of EGFR.....	7
Antibody engineering on the nanoscale.....	9
Summary .....	10
References .....	14
<b>Chapter 2: A modular approach for assembling aldehyde-tagged proteins on DNA scaffolds.....</b>	<b>16</b>
Background.....	17
Results.....	18
Materials and methods .....	24
<i>Chemical reagents</i> .....	24
<i>DNA oligonucleotides</i> .....	25
<i>Expression and purification of aldehyde-tagged proteins</i> .....	27
<i>Conjugation of aldehyde-tagged proteins to DNA</i> .....	29
<i>Protein-DNA purification and assembly</i> .....	31
<i>Flow Cytometry</i> .....	32
<i>Transmission electron microscopy and data processing</i> .....	33

Acknowledgments.....	35
References .....	45
<b>Chapter 3: Modular assemblies of biparatopic antibodies on DNA scaffolds targeting</b>	
<b>urokinase plasminogen activator receptor .....</b>	<b>47</b>
Background.....	48
Approach .....	49
Results.....	52
Future directions.....	54
Materials and methods .....	55
<i>DNA oligonucleotides .....</i>	<i>55</i>
<i>Expression and purification of aldehyde-tagged proteins.....</i>	<i>56</i>
<i>Conjugation of aldehyde-tagged proteins to DNA.....</i>	<i>57</i>
<i>Protein-DNA purification and assembly.....</i>	<i>58</i>
<i>Flow Cytometry.....</i>	<i>58</i>
<i>Invasion Assay.....</i>	<i>59</i>
Acknowledgments.....	60
References .....	69
<b>Chapter 4: Nanoscale dimerization of EGFR on live cells .....</b>	<b>71</b>
Background.....	72
Results.....	74
Future directions.....	77



Materials and methods .....	78
<i>Synthesis of benzylguanine-modified oligonucleotides</i> .....	78
<i>Plasmids and lentivirus</i> .....	79
<i>Stable cell line generation</i> .....	79
<i>Stimulating cells with ligands and dimerizers</i> .....	80
<i>Western blotting</i> .....	81
<i>Confocal microscopy</i> .....	82
Acknowledgments.....	82
References .....	95
<b>Chapter 5: Efficient targeting of fatty-acid modified oligonucleotides to live cell membranes adhesion through stepwise assembly.....</b>	<b>96</b>
Background.....	97
Results.....	98
Materials and methods .....	102
<i>Synthesis of lipid-modified oligonucleotides</i> .....	102
<i>DNA labeling of cells and quantification of cell surface oligonucleotides</i> .....	104
<i>Dynamic light scattering</i> .....	106
Acknowledgments.....	107
References .....	113
<b>Appendices .....</b>	<b>114</b>
A.1: Fab genetic constructs .....	115

A.2: Synthetic DNA sequences.....	116
A.3: List of stable cell lines.....	121
A.4: List of primers for cloning genetic constructs .....	124
A.5: List of publications and patents .....	129

# TABLE OF FIGURES

## **Chapter 1: Introduction**

Figure 1.1: Techniques for determining structure-function relationships.....	12
Figure 1.2: Examples of mechanisms by which nanoscale organization of cell surface proteins can impact how a cell responds to external signals.....	13
Figure 1.3: Conceptualization of antibodies as antigen-binding domains (Fabs) and activity domains (Fc's) held together by different scaffolds.....	13

## **Chapter 2: A modular approach for assembling aldehyde-tagged proteins on DNA scaffolds**

Figure 2.1: Modular strategies for controlling antibody scaffold geometry using DNA- conjugates of aldehyde-tagged proteins.....	36
Figure 2.2: Modular and site-specific conjugation of oligonucleotides to aldehyde-tagged proteins.....	37
Figure 2.3: DNA conjugation to aldehyde-tagged immunoglobulins at the N-terminus, C- terminus, and an internal loop .....	38
Figure 2.4: Modular assembly of protein-bearing DNA multimers and their interactions with live cells.....	39
Figure 2.5: Proteins bearing the 5-residue aldehyde tag (CxPxR) are either co-expressed with Formylglycine Generating Enzyme .....	40
Figure 2.6: Three conjugations strategies .....	40
Figure 2.7: HPLC purification of protein-DNA conjugates using an anion exchange method	41

Figure 2.8: Aldehyde tagged Maltose Binding Protein (MBP) was reacted first with bifunctional aminoxy-azide linker and then with copper free click chemistry reagent DBCO-modified DNA .....	41
Figure 2.9: Purified anti-uPAR Fab-DNA conjugate was hybridized with a complementary FITC-labeled oligonucleotide .....	42
Figure 2.10: Flow cytometry analysis of non-uPar-expressing HEK cells incubated with DNA-scaffolded trimer containing Fab and AlexaFluor488-labeled MBP .....	42
Figure 2.11: Unmodified Fabs stained with uranyl acetate and imaged by transmission electron microscopy.....	43
Figure 2.12: Distribution of distances between MBP centers and the center point of trimers generated with 26 bp long oligonucleotides .....	43
Figure 2.13: Representative MALDI-MS trace for synthesized hexynyl- modified oligonucleotides .....	44

**Chapter 3: Modular assemblies of biparatopic antibodies on DNA scaffolds targeting urokinase plasminogen activator receptor**

Figure 3.1: uPAR and structural analysis for a bispecific antibody .....	61
Figure 3.2: uPAR and structural analysis for a bispecific antibody .....	62
Figure 3.3: Modular DNA-programmed assembly of Fabs to vary the flexibility and valency of biparatopic antibody scaffold .....	63
Figure 3.4: Alternative scheme for modular DNA-programmed assembly of Fabs based on using the trimer motif with a handle strand .....	64

Figure 3.5: uPAR and structural analysis for a bispecific antibody .....	65
Figure 3.6: 2G10 and 3C6 biparatopic antibodies binding to H1299 cells, which overexpress uPAR analyzed by flow cytometry .....	66
Figure 3.6: Invasion of MDA-MB-231 cells after treatment with uPAR-inhibiting Fabs .....	68

#### **Chapter 4: Nanoscale dimerization of EGFR on live cells**

Figure 4.1: Canonical mechanism for activation of EGFR and its downstream pathways .....	83
Figure 4.2: Changes in spatial organization of EGFR upon addition of ligand EGF .....	84
Figure 4.3: Mechanism of covalent conjugation of Snap-tagged receptors to benzylguanine- modified DNA .....	84
Figure 4.4: Ligand-independent dimerization of EGFR with DNA-scaffolds .....	85
Figure 4.5: Phosphorylation of EGFR is dependent of dimerization .....	86
Figure 4.6: Direct comparison of phosphorylation levels of EGFR for EGF and DNA stimulated cells .....	87
Figure 4.7: DNA-stabilized EGFR auto-phosphorylates via a similar mechanism to EGF- stimulated EGFR .....	88
Figure 4.8: DNA-dimerization of EGFR on multiple cell lines.....	89
Figure 4.9: Co-immunoprecipitation of adaptor proteins with an EGFR .....	90
Figure 4.10: Activated-RAS-GTP pull-down assays .....	90
Figure 4.11: DNA-dimerized EGFR does not stimulate endocytosis.....	91
Figure 4.12: DNA-dimerized EGFR does not get transported into clathrin-coated pits. ....	92
Figure 4.13: Super-resolution images of EGFR on cells .....	93

**Chapter 5: Efficient targeting of fatty-acid modified oligonucleotides to live cell membranes adhesion through stepwise assembly**

Figure 5.1: Step-wise assembly of fatty-acid (FA)-modified ssDNA into cell membranes.... 108

Figure 5.2: Anchor (Anch) and complementary Co-Anchor (cA) strands together enhance ssDNA targeting and retention in cell membranes..... 109

Figure 5.3: Lipid hydrophobicity affects cell labeling efficiency of Anch, cA, and pre-hybridized strands ..... 110

Figure 5.4: Modular assembly of protein-bearing DNA multimers and their interactions with live cells..... 111

Figure 5.5: Improved preparation of single cell microarrays and 3D microtissues using stepwise assembly of membrane anchored adhesive oligonucleotides..... 112

## **CHAPTER 1**

### Introduction

# WHY STUDY NANOSCALE ORGANIZATION?

Structure is inextricably linked to function in biology. Establishing relationships between structure and function requires technologies for both visualizing and systematically perturbing structures of interest. Importantly, these technologies must also interface with cells at relevant length scales in order for connections to be drawn between specific structural elements and their functional significance (Figure 1.1). For instance, a combination of crystallography and site-directed mutagenesis has revealed the relationship between structure and function of numerous macromolecules and macromolecular complexes at an atomic scale, such as the ribosome<sup>1</sup> and spliceosome<sup>2</sup>. Likewise, the combination of fluorescent microscopy and lithographically patterned surfaces with “spatial mutations” has revealed the structure and function of organized cell surface domains spanning hundreds of nanometers to tens of microns<sup>3</sup>. For example, the microscale organization of the immunological synapse<sup>4</sup>, Eph receptor complexes<sup>5</sup>, and focal adhesions<sup>6</sup> all influence the activation of downstream signaling pathways and ultimately cell behaviors. Similar techniques have also revealed the importance of organization at larger length scales. In fact, “spatial mutations” of tissue architecture have been shown to regulate many biological processes, including proliferation, apoptosis, morphogenesis and differentiation<sup>7</sup>.

Despite advances at the atomic and microscale, analyses of structure and function at the nanoscale, particularly at the cell surface, remain challenging. Frustratingly, numerous studies implicate nanoscale spatial information in the regulation of signaling pathways and cellular processes, yet the details of these systems are still obscure. These examples include



the discovery of signaling protein- rich lipid microdomains<sup>8</sup>, extracellular matrix that communicates the microenvironment's spatial information<sup>9</sup>, intracellular scaffolds that can direct complex signaling networks down specific pathways<sup>10</sup>, and the self-organization and oligomerization of cell surface proteins themselves<sup>11</sup>. In particular, we believe that the spatial organization of cellular receptors at a nanoscale level, whether its their proximity to signaling partners, stoichiometry in the signaling complex, or their subcellular compartmentalization, etc., can all impact signaling and determine how a cell responds to signals (Figure 1.2). Recent advancements in super-resolution optical microscopy techniques, such as stochastic optical reconstruction microscopy (STORM)<sup>17</sup>, now allow us to visualizing structure on the nanoscale, and appreciate the complexity of structures on this length scale. Yet there remains a need for complementary molecular tools that can introduce "spatial mutations" at this length scale so that we can perturb these structures and study their function and importance on downstream signaling pathways.

The spatial organization of cell-surface receptors can alter the outcome of an incoming signal<sup>3</sup>. For example, the microscale organization of ligands and receptors in the immunological synapse affects the levels of a specific *targeted* cytokine secreted by the activated T-cell. However, the secretion levels of *proliferative* cytokines do not vary when perturbed by the same "spatial mutations."<sup>4</sup> Likewise, in a study of FcεRI immunoreceptor clustering, a trivalent dinitrophenyl ligand with rigid DNA increased the distance between the ligands and attenuated Linker of activated T-cells (LAT) phosphorylation, but does not decrease phospholipase C $\gamma$  (PLC $\gamma$ ) phosphorylation<sup>27</sup>. These studies suggest that other

receptors integrate ligand spatial organization in ligand concentration to regulate signal intensity and specificity.

## NANOSCALE ORGANIZATION OF EGFR

Among the many molecular processes exhibiting higher order organization on the nanoscale, signaling through the Epidermal Growth Factor Receptor (EGFR) family merits further study because misregulation or mutation of components of this pathway are implicated in the pathophysiology of a number of disease<sup>12</sup>. EGFR receptors are part of a layered signaling network activated by many ligands and integrating with pathways important for development, apoptosis, and growth<sup>13</sup>. There are four receptors in the family, and EGFR, Her2, and Her3 are often misregulated in cancer and targeted by successful therapies. Yet, resistance to these chemotherapies inevitably arises and in many cases the mechanism of resistance is unclear, suggesting that we may still be missing information on how these receptors signal and are regulated despite the plethora of studies on the EGFR family.

Although genetic and biochemical studies have been invaluable for identifying and characterizing individual steps in the activation of EGFR and its downstream effectors, the detailed behavior of EGFR in a cellular context remains murky at best. For example, EGFR was thought to operate via ligand-mediated dimerization, triggering trans-activation of tyrosine kinase activity in the cytoplasm<sup>14</sup>. However, recent studies suggest dimerization alone may not be sufficient to achieve maximal activation. In fact, EGFR sometimes exists in the form of

inactive dimers prior to ligand binding in structural and diffusion based studies<sup>15</sup> and recent evidence suggests a role for EGFR multimeric cluster formation in receptor activation<sup>16</sup>. Still, it remains unclear if these higher order structures are a cause or a consequence of activation.

It has been hypothesized that these higher order structures play a role in directing EGFR signaling output. If this were the case, understanding in detail the effect of nanoscale structure on EGFR function could lead to new strategies that more precisely target or control receptor functions, and therefore cellular responses. For instance, if it is discovered that the relevant and most potent signaling of EGFR is a multimer, we could attempt to recognize or disrupt this signaling complex with a multivalent nanoscale-structure antibody-like therapeutic. However, testing this hypothesis has been a challenge due in part to a lack of tools for both for imaging and systematically perturbing the structures at the relevant length scales. The need for such tools has been felt more urgently as recent advancements in super-resolution optical microscopy techniques, such as stochastic optical reconstruction microscopy (STORM)<sup>17</sup>, have brought us closer to visualizing structure on the nanoscale. Although we can now see and appreciate nanoscale structures using new techniques, but there remains a need for complementary molecular tools that can introduce “spatial mutations” at this length scale so that we can investigate whether these nanoscale structures serve any role or are merely a consequence of other signaling steps that are more important. My goal for my PhD was to develop new strategies and technologies aimed at addressing this need in order to enable deeper study of cellular signaling complexes at a nanometer length scale.

## DNA AS AN IDEAL NANOSCALE SCAFFOLD

I proposed the synthesis of a new class of nanoscale multivalent probes for perturbing the higher order organization of cell surface receptors, including EGFR. To be widely applicable, these probes must have structurally-defined scaffolds, bioactive ligands, and be constructed via a modular assembly process. I recognized that DNA could serve as a well-precedented structural scaffold for these probes<sup>4</sup>. Unlike most polymers, DNA sequences may be designed to adopt well-defined structures at nanometer length-scales, with a resolution approaching 3.4 nm<sup>18</sup>. The field of Structural DNA Nanotechnology has already explored numerous architectures that can be adopted as scaffolds with defined valences and geometries<sup>19</sup>. Additionally, I could easily design simple 20-40 base pair sequences with minimal secondary structure that could efficiently anneal to create well defined dimers and trimers, which would already present us with many architectures without having to add too many complicated scaffold strands or introduce too much negative charge by having a large amount of DNA attached to the protein due to the negatively charged backbone of DNA.

An efficient and specific method for linking sensitive proteins to DNA was necessary to fully realize this approach. In Chapter 2, I will describe a combined genetic and chemical strategy for linking functional protein-binding domains to DNA scaffolds that I developed. Successful conjugation results in a large increase in negative charge, allowing for easy modular purification of complexes by anion-exchange chromatography. Therefore, I can prepare a large and unbiased set of reagents that vary in valency, binding affinities, and relative spacing. No current technology boasts the capabilities of simultaneously controlling all of these variables

while remaining modular in design and synthesis.

An added benefit of developing an accessible and modular protein-DNA conjugation technique is that there are many other applications for site specific protein-DNA conjugates, such as aiding with protein-structure solving by anchoring proteins of interest in defined orientations on a DNA origami grid, scaffolding enzymes that are part of a metabolic pathway together to increase pathway flux and overall final product synthesis, and also just the synthesis of antibody-siRNA or antibody-drug conjugates. We hope that our technique allows researchers to easily make protein-DNA conjugates for their applications, which is why we were careful to design a tag-based technique because capable molecular cloning skills are common for many scientists and we developed different linkers and conjugation strategies so that scientists of all levels of chemical expertise could find a method with which they feel comfortable and confident.

## SPATIAL MUTATION OF EGFR

The original goals of this project were to (1) to synthesize a wide collection of DNA-scaffolded multivalent protein structures, and (2) to apply them to directly manipulate the nanoscale organization of EGFR signaling complexes. I hoped to provide the foundation necessary to explore the functional consequences of perturbing and generating nanoscale “spatial mutations” on downstream signaling processes. This technology had the potential to explore questions about the nanoscale structure of many cell-surface signaling complexes. I

also believed that it may also provide the framework for a new class of therapeutics that can either target or create nanoscale structures to influence cell behavior.

Importantly, probes of EGFR nanoscale organization must selectively bind the receptor without unintentionally affecting its activity. I therefore originally wanted to use a well-characterized 15kDa nanobody (EGb4), which is a single domain antibody that binds to the extracellular domain of EGFR. EGb4 neither activates EGFR signaling nor competes with EGF binding<sup>20</sup>. By linking EGb4 to a number of DNA scaffolds, I hoped probe the effects of receptor organization on EGFR signaling. However, because in order to move forward with the project concurrently with developing the protein-DNA conjugation techniques, we ended up engineering cell lines to express EGFR with a Snap-tag on the extracellular N-terminus. The Snap-tag is a genetic tag that interacts covalently with benzylguanine to form a permanent linkage. So when we synthesize oligonucleotides with benzylguanine groups, these covalently conjugate to EGFR molecules on the cell surface within a few minutes.

I hypothesized that different nanoscale organizations of EGFR will lead to divergent activation levels and sensitization to EGF in subsets of downstream pathways, such as the MAPK, PLC $\gamma$ /PKC, and PI(3)K/Akt pathways. I proposed to explore this hypothesis by exposing EGFR-expressing cells to NMPs varying in size and valency, and then observing their effect on EGFR signaling. Specifically, I wanted to observe if “gain-of-dimer” mutations with DNA-scaffolded probes would (1) cause ligand-independent activation, (2) alter the sensitivity of downstream signaling pathways to EGF, and (3) influence rates of EGFR endocytosis.

Evidence suggests higher order EGFR structures play a role in modulating EGFR signaling output. With a wide collection of DNA-scaffolded probes, I was able to systematically

perturb the nanoscale organization of cell surface EGFR. We demonstrate that EGFR organization influences EGFR activation, cell sensitivity to EGF, and EGFR endocytosis. Future experiments could investigate the effect of these changes on cellular behaviors, such as proliferation and mRNA synthesis. In more advanced studies, similar experiments could be performed on cells that natively express EGFR and other ErbB receptors. Because ErbB receptors heterodimerize<sup>13</sup>, future experiments could determine if EGFR organization affects the organization and activation of other ErbB receptors. Combined with the ability to image receptor organization using super resolution microscopy, these nanoscale DNA-scaffolded probes will allow structure-function relationships for EGFR to be constructed at the nanoscale.

Significantly, the technology developed in this proposal can be applied to other signaling systems where evidence of higher-order organization exists, such as the growth hormone<sup>33</sup> or insulin receptor signaling pathways<sup>34</sup>. Understanding in detail the effect of nanoscale structure on receptor functions could lead to strategies that more precisely control receptor organization, and therefore cellular responses. As a consequence, these nanostructured molecules may provide a foundation for diagnostics or therapeutics that can either create nanoscale structures to influence cell behavior, or perhaps selectively target nanoscale structures associated with disease states.

## ANTIBODY ENGINEERING ON THE NANOSCALE

Because our DNA-protein conjugates can be assembled into nanoscale objects, we also

were interested in using them in biological and biomedical applications other than studying cell signaling. One particular application of interest is antibody engineering using DNA scaffolds. In vivo, antibodies of different classes have different biological activities and serve specialized roles during the immune response. An antibody's class is determined by its constant region, or scaffold, which encodes the valency, effector functions, and higher-order architecture of the pendant variable domains (Figure 1.3). Expanding scaffold diversity in the antibody repertoire has the potential to expand the neutralizing capacity of the immune system, i.e. by delivering new effectors, increasing avidity, or modulating specificity. In collaboration with other labs at UCSF and industry, we developed a new scheme for rapidly exploring bispecific antibody structural geometries, valencies, and flexibilities to maximize the synergistic binding of two different antibody Fabs that inhibited an extracellular signaling protein. This protein, urokinase plasminogen activator receptor (uPAR) is often upregulated in triple negative breast cancer that is also associated with increased invasiveness of tumors. Our goal was to use this as a test case to test our modular and high throughput platform for the discovery of bi- and multi-specific immunotherapies.

## SUMMARY

This thesis details my work, both published and ongoing projects, on assembling proteins with modular DNA scaffolds in order to develop new strategies for engineering antibodies and probing cellular signaling complexes. In Chapter 1, I describe a strategy



required a simple and modular technique for site-specifically conjugating synthetic oligonucleotides to proteins. In Chapter 2, I describe how we are using this technique to prepare and screen multivalent DNA-scaffolded biparatopic antibodies for inhibiting urokinase plasminogen activator receptor in triple negative breast cancer lines. Finally, in Chapter 3, I detail my work using these modular DNA scaffolds to investigate the effect of nanoscale organization on the signaling profiles of the Epidermal Growth Factor Receptor (EGFR) family by systematically guiding the assembly of signaling complexes on cells.

# FIGURES

**Figure 1.1:** Techniques for determining structure-function relationships.

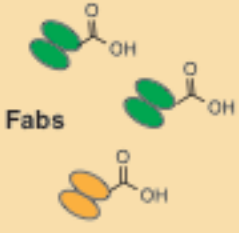
	0.1 nm	10 nm	1000 nM	
	<b>Atomic Scale</b>		<b>Nanoscale</b>	<b>Microscale</b>
<b>Example of Processes</b>	Protein structures Macromolecule complexes Enzymatic transformations		Receptor clustering Lipid domains ECM structure	Immunological synapse Eph receptor organization
<b>Visualization Techniques</b>	Nuclear magnetic resonance X-ray crystallography		Electron microscopy Atomic force microscopy Stochastic optical reconstruction microscopy	Light microscopy Fluorescence microscopy
<b>Perturbation Techniques</b>	Site-directed mutagenesis Small molecules		Antibodies ???	Lithography

Examples of visualization and perturbation tools for structure at atomic, nanometer, and micron length scales are listed. Despite advancements for visualization techniques for viewing structures at the nanoscale with super-resolution microscopy techniques, there remains a need for the development of new tools for perturbing and creating nanoscale structures in order to investigate their function in a precise and modular way.

**Figure 1.2:** Examples of mechanisms by which nanoscale organization of cell surface proteins can impact how a cell responds to external signals.



**Figure 1.3:** Conceptualization of antibodies as antigen-binding domains (Fabs) and activity domains (Fc's) held together by different scaffolds found in the human immune system, which direct their isotype and activity. All possible functions and optimized activities are not likely be captured by these 9 scaffold architectures and many opportunities could exist if we are able to engineer the scaffold of antibodies and screen for new and improved activities.

Building Blocks	Scaffold				Isotype and Activity															
 <p>Fabs</p>	<p><b>Antibody Classes (9)</b></p> <table border="1" data-bbox="639 1352 967 1686"> <tr> <td></td> <td></td> <td></td> <td></td> <td>IgG1</td> </tr> <tr> <td>IgG2</td> <td>IgG3</td> <td>IgG4</td> <td></td> <td></td> </tr> <tr> <td>IgE1</td> <td>IgA1</td> <td>IgA2</td> <td></td> <td></td> </tr> </table>								IgG1	IgG2	IgG3	IgG4			IgE1	IgA1	IgA2			<p><b>IgG:</b> Main immune response  <b>IgM:</b> Early immune response  <b>IgA:</b> mucosal areas of body (gut and respiratory tract)</p>
				IgG1																
IgG2	IgG3	IgG4																		
IgE1	IgA1	IgA2																		

## REFERENCES

- (1) Yusupov, M., Yusupova, G. & Baucom, A. Crystal structure of the ribosome at 5.5 Å resolution. *Science* (2001).
- (2) Kambach, C. & Walket, S. Structure and assembly of the spliceosomal small nuclear ribonucleoprotein particles. *Current opinion in structural biology* (1999).
- (3) Manz, B.N. & Groves, J.T. Spatial organization and signal transduction at intercellular junctions. *Nat Rev Mol Cell Biol* 11, 342-352 (2010).
- (4) Shen, K., Thomas, V.K., Dustin, M.L. & Kam, L.C. Micropatterning of costimulatory ligands enhances CD4+ T cell function. *Proc Natl Acad Sci USA* 105, 7791-7796 (2008).
- (5) Salaita, K. et al. Restriction of Receptor Movement Alters Cellular Response: Physical Force Sensing by EphA2. *Science* 327, 1380-1385 (2010).
- (6) Pampori, N. & Shattil, S. Complementary roles for receptor clustering and conformational change in the adhesive and signaling functions of integrin  $\alpha$ IIb $\beta$ 3. *The Journal of cell biology* (1998).
- (7) Nelson, C.M., Vanduijn, M.M., Inman, J.L., Fletcher, D.A. & Bissell, M.J. Tissue geometry determines sites of mammary branching morphogenesis in organotypic cultures. *Science* 314, 298-300 (2006).
- (8) Brown, D.A. & London, E. Functions of lipid rafts in biological membranes. *Annu. Rev. Cell Dev. Biol.* 14, 111-136 (1998).
- (9) Maheshwari, G., Brown, G. & Lauffenburger, D. Cell adhesion and motility depend on nanoscale RGD clustering. ... of Cell Science (2000).
- (10) Zeke, A., Lukács, M., Lim, W.A. & Reményi, A. Scaffolds: interaction platforms for cellular signalling circuits. *Trends Cell Biol* 19, 364-374 (2009).
- (11) Lemmon, M.A. & Schlessinger, J. Regulation of signal transduction and signal diversity by receptor oligomerization. *Trends Biochem. Sci.* 19, 459-463 (1994).
- (12) Wieduwilt, M.J. & Moasser, M.M. The epidermal growth factor receptor family: Biology driving targeted therapeutics. *Cell. Mol. Life Sci.* 65, 1566-1584 (2008).
- (13) Yarden, Y. & Sliwkowski, M. Untangling the ErbB signalling network. *Nat Rev Mol Cell Biol* 2, 127-137 (2001).
- (14) Zhou, M. et al. Real-time measurements of kinetics of EGF binding to soluble EGF receptor monomers and dimers support the dimerization model for receptor activation. (1993).
- (15) Chung, I. et al. Spatial control of EGF receptor activation by reversible dimerization on living cells. *Nature* 464, 783-787 (2010).
- (16) Clayton, A.H.A., Orchard, S.G., Nice, E.C., Posner, R.G. & Burgess, A.W. Predominance of activated EGFR higher-order oligomers on the cell surface. *Growth Factors* 26, 316-324 (2008).
- (17) Huang, B., Wang, W., Bates, M. & Zhuang, X. Three-Dimensional Super-Resolution Imaging by Stochastic Optical Reconstruction Microscopy. *Science* 319, 810-813 (2008).
- (18) Lin, C., Liu, Y. & Yan, H. Designer DNA Nanoarchitectures†. *Biochemistry* 48, 1663-1674 (2009).

- (19) Aldaye, F.A., Palmer, A.L. & Sleiman, H.F. Assembling Materials with DNA as the Guide. *Science* 321, 1795- 1799 (2008).
- (20) Hofman, E.G. et al. EGF induces coalescence of different lipid rafts. *J Cell Sci* 121, 2519-2528 (2008).
- (21) Niemeyer, C.M. Semisynthetic DNA-Protein Conjugates for Biosensing and Nanofabrication. *Angew. Chem. Int. Ed.* 49, 1200-1216 (2010).
- (22) Carrico, I.S., Carlson, B.L. & Bertozzi, C.R. Introducing genetically encoded aldehydes into proteins. *Nature Chemical Biology* 3, 321-322 (2007).
- (23) Wu, P. et al. Site-specific chemical modification of recombinant proteins produced in mammalian cells by using the genetically encoded aldehyde tag. *Proceedings of the National Academy of Sciences* 106, 3000 (2009).
- (24) Ning, X. et al. Protein Modification by Strain-Promoted Alkyne-Nitrone Cycloaddition. *Angew. Chem. Int. Ed.* 49, 3065-3068 (2010).
- (25) Soriano del Amo, D. et al. Biocompatible Copper (I) Catalysts for in Vivo Imaging of Glycans. *J Am Chem Soc* 2004-2021 (2010).
- (26) Seeman, N.C. Nanomaterials Based on DNA. *Annu. Rev. Biochem.* 79, 65-87 (2010).
- (27) Sil, D., Lee, J., Luo, D., Holowka, D. & Baird, B. Trivalent ligands with rigid DNA spacers reveal structural requirements for IgE receptor signaling in RBL mast cells. *ACS Chemical Biology* 2, 674-684 (2007).
- (28) Chu, C., Everiss, K. & Wikstrand, C. Receptor dimerization is not a factor in the signalling activity of a transforming variant epidermal growth factor receptor (EGFRvIII). *Biochemical ...* (1997).
- (29) Pedersen, M.W. et al. Analysis of the epidermal growth factor receptor specific transcriptome: Effect of receptor expression level and an activating mutation. *J Cell Biochem* 96, 412-427 (2005).
- (30) Verveer, P., Wouters, F. & Reynolds, A. Quantitative Imaging of Lateral ErbB1 Receptor Signal Propagation in the Plasma Membrane. *Science* (2000).
- (31) Sorkin, A. & Goh, L.K. Endocytosis and intracellular trafficking of ErbBs. *Exp Cell Res* 315, 683-696 (2009).
- (32) Wang, Q., Villeneuve, G. & Wang, Z. Control of epidermal growth factor receptor endocytosis by receptor dimerization, rather than receptor kinase activation. *EMBO Rep.* 6, 942-948 (2005).
- (33) Wells, J. Structural and functional basis for hormone binding and receptor oligomerization. *Current opinion in cell biology* 6, 163-173 (1994).
- (34) De Meyts, P. The structural basis of insulin and insulin-like growth factor-I receptor binding and negative co-operativity, and its relevance to mitogenic versus metabolic signalling. *Diabetologia* 37, 135-148 (1994).

## **CHAPTER 2**

# A modular approach for assembling aldehyde-tagged proteins on DNA scaffolds

Source: The following chapter was published in part from: Liang, S.I.; McFarland, J.M.; Rabuka, D.; Gartner, Z.J., 2014. A modular approach for assembling aldehyde-tagged proteins on DNA scaffolds. *Journal of the American Chemical Society* 136(31), 10850-10853.

Contributions: I initiated the project and performed the majority of the experiments, with help from Jesse McFarland for experiments using the full length IgG. Zev Gartner and I wrote the manuscript with editorial input from all authors. Zev Gartner supervised the project.

## BACKGROUND

DNA-protein conjugates can be assembled into nanoscale objects through the power of structural DNA nanotechnology. These motifs have the potential to revolutionize a number of biological and biomedical applications.<sup>1-3</sup> One particular application of interest is antibody engineering using DNA scaffolds.<sup>4</sup> In vivo, antibodies of different classes have different biological activities and serve specialized roles during the immune response. An antibody's class is determined by its constant region, or scaffold, which encodes the valency, effector functions, and higher-order architecture of the pendant variable domains. Expanding scaffold diversity in the antibody repertoire has the potential to expand the neutralizing capacity of the immune system, i.e. by delivering new effectors, increasing avidity, or modulating specificity.

However, systematic exploration of antibody scaffold geometry, valency, and combinatorial binding capacity is difficult with protein-based scaffolds due to the challenges associated with protein design.<sup>5-7</sup> DNA-based scaffolds, in contrast, are programmable, and can combinatorially control the position and orientation of pendant proteins with nanometer resolution (Figure 2.1A). Appropriately designed DNA scaffolds could assemble proteins that recognize specific combinations of receptors on cell surfaces,<sup>4</sup> and even deliver protein-based therapeutics specifically to these cells.<sup>8,9</sup>

To fully realize the potential of using DNA scaffolds to expand the repertoire of antibody structure and function requires more modular strategies for preparing DNA-protein conjugates. Ideal strategies would provide modularity in linkage chemistries, linkage site (e.g. termini or internal loops), and expression host. Aldehyde tagged proteins satisfy these

requirements (Figure 2.1B). Among peptide tags and self-ligating proteins,<sup>10-18</sup> the aldehyde tag uniquely combines the advantages of a short consensus sequence (5 amino acids), a bioorthogonal handle which is amenable to conjugation through a number of chemical linkers, diverse prokaryotic or eukaryotic expression hosts, and compatibility with insertion at any position in a protein's primary sequence.<sup>19-22</sup>

## RESULTS

We explored the potential of aldehyde tagged proteins to form defined DNA-protein conjugates using a model substrate (maltose binding protein, MBP) bearing a C-terminal aldehyde tag expressed in *E. coli*. Co-expression of Formylglycine Generating Enzyme (FGE) leads to the post-translational conversion of the cysteine in the aldehyde tag consensus sequence (CxPxR) to formylglycine (Figure 2.5). We also synthesized an oligonucleotide functionalized with a 5'-dimethoxytrityl (DMT)-protected aminoxy nucleophile, 1, from commercially available reagents. DMT-protection of the alkoxyamine stabilizes the product for storage but is rapidly deprotected in situ under the mildly acidic conjugation conditions. The resulting oxime product is observed as a higher molecular weight species by gel electrophoresis, and densitometry of the banding pattern indicated an 81% yield with respect to protein concentration (Figure 2.2A, lane 2). No conjugate was formed using a C→A mutation in the aldehyde tag consensus sequence (Figure S2.2). Thus, direct conjugation of aminoxy-modified DNA 1 to aldehyde-tagged proteins generates product efficiently using



only commercially available reagents.

Like other bioconjugation techniques such as thiol-maleimide coupling,<sup>23,24</sup> the oxime linkage formed between 1 and an aldehyde-tagged protein is hydrolytically unstable upon long-term incubation in serum. This observation motivated the development of alternate conjugation strategies such as the Hydrazino-iso-Pictet-Spengler (HIPS) ligation.<sup>25</sup> This recently reported reaction proceeds efficiently at near-physiological pH to form a stable covalent linkage with aldehyde tagged proteins. We therefore coupled the HIPS reagent to a 5' amino-modified oligonucleotide and incubated the product 2 with aldehyde-tagged MBP at pH 5.5 to generate a DNA-protein conjugate in 62% yield (Figure 2.2A, lane 3). While the HIPS reagent must be synthesized prior to DNA conjugation, HIPS ligation proceeds at higher pH and forms a covalent and an irreversible carbon-carbon bond between DNA and protein.<sup>26</sup>

Additionally, we explored the potential to convert the formylglycine to a more reactive functionality for cases where more rapid coupling is required. Aldehyde bearing MBP was treated with an excess of a low molecular weight bifunctional linker 3 to introduce an azide group. Excess linker drives this reaction to completion and is easily removed by gel filtration due to its low molecular weight. Subsequent coupling with alkyne-modified DNA 4 occurred upon incubation with biocompatible copper stabilizing ligands such as BTTP,<sup>27</sup> copper(II) sulfate, and sodium ascorbate with yields between 63-87% (Figure 2.2A, lane 4). Alkyne-modified DNA is inexpensive to synthesize in large quantities, allowing reaction scale-up and purification of the conjugate by anion exchange chromatography (Figure 2.2B, 2.7). The functionality and addressability of the DNA on the conjugate was verified by hybridizing it with a matching FITC-conjugated oligo (Figure 2.2C).

Conjugation of the azide-bearing protein with DNA can also proceed efficiently under copper free conditions with dibenzocyclooctyne (DBCO)-modified DNA, 5. Incubation of azide-bearing MBP with 5 generated product in 79% yield with respect to protein (Figure 2.8). Together, this combination of four conjugation strategies provides flexible means of converting aldehyde-tagged proteins into DNA-protein conjugates with diverse physicochemical properties.

A key advantage of small peptides such as the aldehyde tag is that they can be used to prepare DNA-protein conjugates at either termini or internal loops of immunoglobulins. For example, we inserted an aldehyde tag onto the C-terminus of a Fab raised against the Urokinase Plasminogen Activator Receptor (uPAR), an extracellular scaffold protein that regulates cell migration and invasion.<sup>28,29</sup> After conversion of the formylglycine to an azide using the bifunctional linker 3, the product was conjugated to 4 using BTTP-stabilized click chemistry (Figure 2.3A). The resulting DNA-protein conjugate retained its ability to specifically bind uPAR on live cells. For example, an anti-uPAR Fab-DNA conjugate hybridized with a FITC-labeled oligonucleotide was able to efficiently label uPAR-expressing H1299 cells (Figure 2.9).

We also prepared glycosylated Fc fragments of a human IgG with the aldehyde tag inserted at the N- or C-termini. Co-expression of these proteins with FGE in CHO cells yielded aldehyde-tagged protein with moderate levels of conversion. Transformation of the formylglycine to an alkyne using 3 and BTTP-stabilized click ligation to alkyne-modified oligonucleotides 4 generated a higher molecular weight species, consistent with formation of DNA-protein conjugates (Figure 2.3B). The lower yield of product observed for the C-terminally labeled site suggests that aldehyde reactivity depends on its placement within the primary

sequence of the protein. Finally, we expressed a fully glycosylated IgG containing an aldehyde tag on an internal loop in FGE-expressing CHO cells. Incubation of different concentrations of aminoxy-modified DNA 1 with the IgG resulted in the appearance of a higher molecular weight band on a reducing SDS-PAGE gel that is consistent with a DNA-conjugated light chain (Figure 2.3C). Together, these experiments indicate the necessary modularity of tag placement and expression hosts when preparing DNA-protein conjugates for assembly into nanoscale geometries on DNA-scaffolds.

As our goal of synthesizing DNA-protein conjugates is to facilitate the assembly of proteins into antibody-like geometries, we explored the efficiency with which several simple DNA motifs mimicking the geometry of antibody scaffolds could be prepared from these reagents. We conjugated aldehyde-tagged MBP to 20 and 26 base oligonucleotides designed to self-assemble into dimers and trimers, respectively (Figure 2.4A). After hybridizing in PBS for 1 h at 25 °C, SDS-PAGE analysis indicated that both dimers and trimers assembled efficiently (Figure 2.4B).

Additionally, DNA-protein conjugates assembled into small multiprotein motifs interacted efficiently with cell surfaces. For example, we used flow cytometry to analyze whether MBP/Fab heterotrimers retained their ability to interact with cells expressing uPAR. As a reporter for cell binding, we prepared an MBP-DNA conjugate modified with AlexaFluor-488 and assembled the resulting fluorescent MBP(488)-DNA conjugates with Fab-DNA to form heterotrimers. Incubation of the uPAR-expressing lung carcinoma cell line H1299 with the Fab-containing heterotrimer resulted in an increase in mean fluorescence in comparison to unlabeled cells (Figure 2.4C). To confirm that the heterotrimers were interacting specifically

with uPAR, we preincubated the H1299 cells with non-fluorescent monomeric anti-uPAR Fab prior to addition of the Fab-MBP(488)-MBP(488) trimer and saw no increase in fluorescence. Similarly, fluorescence of Human Embryonic Kidney (HEK) cells, which do not express uPAR, was unchanged after incubation with the protein heterotrimer (Figure 2.10).

We examined the spatial organization of DNA-scaffolded proteins by negative stain transmission electron microscopy. Individual molecules of MBP in protein trimers were easily identifiable as light spots with a dark halo on a salt-and-pepper background (Figure 2.4D). In contrast to MBP, antibody fragments have more distinct features that can be identified as one of two distinct shapes (Figure 2.11). Thus, in heterotrimers, a single Fab protein was identifiable alongside an MBP dimer (Figure 2.4E). We calculated the distance between the center of individual proteins and the trimer centroid as 8.12 nm and measured an average spacing of 7.03 nm  $\pm$  1.5 nm (s.d.) consistent with our estimate (Figure 2.12). The relatively large standard deviation in our measurements may indicate some conformational flexibility of the DNA scaffold. Additional spatial control and rigidity might be achieved using DNA motifs with longer persistence lengths, such as the double crossover motif.<sup>30,31</sup> Moreover, the ability to modularly insert rigid or shorter chemical linkers would provide additional spatial control in these nanostructures.

Finally, because scaffold valency plays a central role in the immune system (IgG vs. IgA vs. IgM), we explored the hierarchical assembly of simple trimer motifs into higher-order structures. For example, we used an unmodified oligonucleotide as one arm of the trimer motif to assemble MBP-Fab dimers into tetramers, where the distance between each dimer could be varied based on the length of the unmodified scaffold DNA strands (Figure 2.4F). Elaboration

of this simple strategy would allow for the assembly of scaffold protein assemblies of considerably higher valency and complexity.

In conclusion, we describe a simple and modular method for conjugating and then assembling multiple proteins onto DNA scaffolds. Our approach utilizes the aldehyde tag, which is genetically incorporated into the primary sequence of proteins expressed in both bacterial and mammalian expression systems. We tested four bioconjugation reactions that generate site-specific DNA-protein conjugates in moderate to excellent yield. The variety of strategies for conjugation of DNA to aldehyde-tagged proteins provides flexibility in linker chemistry and geometry, and is accessible to individuals with varying levels of synthetic expertise. Moreover, the ability to insert the aldehyde tag at both protein termini or in an internal loop will provide the potential for orientational control of proteins on DNA scaffolds. DNA-protein conjugates can be modularly assembled into dimeric and trimeric nanostructures resembling antibody scaffolds and interfaced with living cells. Transmission electron microscopy verified that the DNA scaffolds arranged proteins as predicted. These motifs can also be assembled hierarchically into structures of greater complexity. We anticipate expanding our DNA scaffolds libraries to generate large collections of macromolecular assemblies varying in valency and architecture that may have novel activities as nanoscale probes or antibodies with unique specificities and biological activities.

# MATERIALS AND METHODS

## CHEMICAL REAGENTS

Sodium ascorbate, copper(II) sulfate, ethylenediaminetetraacetic acid (EDTA), and bathocuproinedisulfonic acid disodium salt (BCS) were purchased from Sigma Aldrich (St. Louis, MO). HPLC grade acetonitrile and triethylamine, and acetic acid were obtained from Thermo Fisher Scientific (Waltham, MA). Controlled pore glass (CPG) support, 5-Hexyn-1-yl-(2-cyanoethyl)-(N,N-diisopropyl)-phosphoramidite (hexynyl phosphoramidite), 10-(6-oxo-6-(dibenzo[b,f]azacyclooct-4-yn-1-yl)-capramido-N-ethyl)-O-triethyleneglycol-1-[(2-cyanoethyl)-(N,N-diisopropyl)]-phosphoramidite (DBCO phosphoramidite), 10-[N-Dimethoxytrityl-aminoxyethyl]-triethyleneglycol-1-[(2-cyanoethyl)-(N,N-diisopropyl)]-phosphoramidite (aminoxy phosphoramidite), 6-(4-Monomethoxytritylamino)hexyl-(2-cyanoethyl)-(N,N-diisopropyl)-phosphoramidite (amine phosphoramidite), and columns were obtained from Glen Research (Sterling, VA). Standard phosphoramidites and DNA synthesis reagents were obtained from Azco Biotech (Oceanside, CA). Bifunctional aminoxy-azide linker O-(2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethyl)hydroxylamine, **3**, was a gift from Jason Hudak (UC Berkeley). Copper-coordinating ligand 3-[4-({bis}[(1-tert-butyl-1H-1,2,3-triazol-4-yl)methyl]amino)methyl)-1H-1,2,3-triazol-1-yl]propanol (BTTP) was a gift from Professor Peng Wu (Albert Einstein College of Medicine). All materials were used as received.

## DNA OLIGONUCLEOTIDES

DMT-protected aminoxy-modified oligonucleotides (1), hexynyl-modified oligonucleotides (4), DBCO-modified oligonucleotides (5) and amine-modified oligonucleotides were synthesized on an Applied Biosystems Expedite 8909 DNA synthesizer using their respective phosphoramidites (100 mM) and a default coupling protocol. Synthesized DMT-protected aminoxy-modified DNA (1), hexynyl-modified DNA (4), and amine-modified oligonucleotides were cleaved from solid support and deprotected using a 1:1 mixture of ammonium hydroxide/40% methylamine (AMA) for 15 minutes at 65°C. DBCO-modified DNA was cleaved and deprotected using 30% ammonium hydroxide for 2 hours at 65°C. After removing the AMA or ammonium hydroxide with a speedvac system, oligonucleotides were resuspended in 100 mM triethylamine acetate (TEAA) and filtered through 0.2- $\mu$ m spin filters. Oligonucleotides were then purified by reversed-phase high-performance liquid chromatography (HPLC) using an Agilent 1260 Infinity Series HPLC System, a 5  $\mu$ m XDB-C18 (21.2x150mm) Agilent (Santa Clara, CA) column, and 100 mM TEAA (pH 7)/acetonitrile as the mobile phase. Fractions were collected based on their absorbance at 260 and 280 nm, and fractions containing hexynyl-modified oligonucleotides were washed with Millipore MilliQ-purified water and lyophilized three times to remove residual HPLC buffer salts and resuspended in water prior to use. DNA concentrations were determined by absorbance at 260 nm. The masses of the oligonucleotides were obtained with matrix-assisted laser desorption ionization (MALDI) mass spectrometry performed on a Voyager-DE Pro, using sample spots pre-coated with nitrocellulose and a hydroxypicolinic acid matrix supplemented with ammonium citrate (representative MALDI-MS trace shown in Figure 2.13). Some DMT-

protected aminoxy-modified oligonucleotides were purchased from Integrated DNA Technologies (Coralville, Iowa).

For the synthesis of the HIPS-reagent modified oligos, 2, a 1  $\mu$ mol scale synthesis of amine-modified DNA still on solid supports was transferred to a microcentrifuge tube. 25  $\mu$ moles of the HIPS-PFP ester was resuspended in 400  $\mu$ l of dry DMF, mixed with 100  $\mu$ l of dry DIPEA, and transferred to the tube such that the liquid completely submerged the DNA-bearing beads. The reaction was incubated on a vortexer overnight. The beads were then washed 3 times with DMF, 3 times with DCM, and then dried completely on a speed vac system. The DNA was cleaved off the beads with AMA for 20 minutes at 65  $^{\circ}$ C, dried on a speed vac system, resuspended in 100 mM triethylamine acetate (TEAA), and filtered through 0.2- $\mu$ m spin filters. Oligonucleotides were then purified by reversed-phase high-performance liquid chromatography (HPLC) as described above.

DNA sequences were selected as previously described to form dimer and trimer constructs.<sup>32, 33</sup>

X = 5' modification

Dimer strands:

D1: X-5'-GTAACGATCCAGCTGTCACT-3'

D2: X-5'-AGTGACAGCTGGATCGTTAC-3'

Trimer strands:

T1: X-5'-TCGATCCGCATGACATTCGCCGTAAG-3'

T2: X-5'-CTTACGGCGAATGACCGAATCAGCCT-3'

T3: X-5'-AGGCTGATTCGGTTCATGCGGATCCA-3'



Summary of mass spectrometry results for synthesized strands of DNA:

Sequence	Expected MW [M+H]	Average Measured MW [M+H] (n=4)	StDev (n=4)
HIPS-D2	6579	6580	0.2
DBCO-D2	6729	6731	0.3
Hexynyl-D1	6238	6239	0
Hexynyl-D2	6318	6319	0
Hexynyl-T1	8092	8093	0.5
Hexynyl-T2	8101	8102	0
Hexynyl-T3	8163	8164	0

DNA strands purchased from Integrated DNA Technologies, Coralville, IA:

DMT-aminooxy-D2: DMT-aminooxy-5'-AGTGACAGCTGGATCGTTAC-3'

For Tetramer assembly, T3-16A:

5'-ACTGACTGACTGACTGAGGCTGATTCGGTTCATGCGGATCGA-3'

For Tetramer assembly, T3-16B:

5'-CAGTCAGTCAGTCAGTAGGCTGATTCGGTTCATGCGGATCGA-3'

Strands purchased from Operon, Huntsville, AL

FAM-5'-AGTGACAGCTGGATCGTTAC-3' (ordered from Operon, Huntsville, AL)

## EXPRESSION AND PURIFICATION OF ALDEHYDE-TAGGED PROTEINS

Maltose Binding Protein and Formylglycine Generating Enzyme: Aldehyde-tagged MBP and FGE were expressed as previously described.<sup>34</sup> BL21(DE3) E. coli transformed with a pET14b

plasmid containing the protein were grown in LB media supplemented with 100 µg/mL ampicillin at 37 °C. When OD<sub>600</sub> reached 0.5, the culture was cooled to 18°C. After 30 minutes, protein expression was induced with 0.1 mM IPTG and cultures were incubated at 18 °C for 14-18 h with shaking. Cells were lysed by sonication and the His6-tagged MBP or FGE was purified using Ni-NTA-agarose beads (Qiagen) under standard purification procedures. The protein was eluted in 50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, 250 mM imidazole, pH 7.4 and dialyzed into PBS with 10% glycerol for storage.

Antibody fragment for uPAR (Fab): An expression plasmid for uPAR-binding Fab 2G10 was a gift from Prof. Charles Craik (UCSF). An aldehyde tag was cloned onto the C-terminus of a His-tagged heavy chain using primers and reagents for Quikchange cloning. BL21-Gold E. coli were transformed with the Fab expression plasmid and grown in 2xYT media supplemented with 0.2% glucose and 100 µg/ml of Ampicillin at 37 °C. When OD<sub>600</sub> reached 0.6, cultures were then cooled to room temperature, induced with 1mM IPTG, and grown overnight at 30 °C. The periplasmic fraction of the cells was collected by pelleting the bacteria, resuspending the cell pellet in 20 ml of ice-cold sucrose buffer (200 mM Tris, 0.5 mM EDTA, 500 mM sucrose, pH 8), then quickly mixing 20 ml of ice cold distilled water to the resuspended cells. The cells were incubated on ice for 30 minutes, with gentle swirling of the samples every 10 minutes. The cells were then removed by centrifugation at 13000xg for 15 minutes at 4 °C and the supernatant was collected as the periplasmic fraction. The Fab was purified from the periplasmic fraction using Ni-NTA-agarose beads and standard purification procedures. Fab was eluted in 50 mM Tris, 500 mM NaCl, 500 mM imidazole, pH 8 and dialyzed into PBS for

storage. Protein concentrations were determined with a BCA protein assay (Thermo Scientific Pierce, Rockford, IL). In order to convert the cysteine of the aldehyde tag to formylglycine, the aldehyde-tagged Fabs (25-80  $\mu$ M) were incubated in 50 mM Tris, 100 mM NaCl, 2 mM DTT, pH 9 and 0.1 equivalents of *M. tuberculosis* FGE for 8-20 hours at 30 °C

Fc fragment and full-length IgG: Full-length human IgG aldehyde-tagged proteins and aldehyde-tagged Fc fragments were generated similar to previously described procedures<sup>35</sup>. Vectors bearing the heavy and light chain were constructed using standard molecular biology techniques. Protein production was performed by Redwood Bioscience using CHO-S cells and standard over-expression protocols. Full-length human IgG was purified from clarified media by affinity chromatography using Protein-A agarose resin. The eluent was dialyzed into PBS and stored at -20 °C.

### CONJUGATION OF ALDEHYDE-TAGGED PROTEINS TO DNA

Using DMT-protected aminoxy-modified DNA: Formylglycine-bearing proteins were treated with 250  $\mu$ M DMT-protected aminoxy-modified DNA, 1, in potassium acetate buffer (final concentration 100 mM, pH 4.6) and incubated overnight at 37 °C.

Using HIPS-modified DNA: Formylglycine-bearing proteins were treated with 250  $\mu$ M DMT-protected aminoxy-modified DNA, 2, in sodium acetate buffer (final concentration 100 mM, pH 5.5) and incubated overnight at room temperature.

Using hexynyl-DNA or DBCO-DNA: Formylglycine-bearing proteins were treated with 1 mM aminoxy-TEG-azide linker, 3, in potassium acetate, (final concentration 100 mM, pH 4.6) and incubated for 6-20 h at 37 °C. Reactions were then buffer exchanged into PBS with Zeba Spin Desalting columns to remove excess linker and used directly in the click reaction. If using 4, 100 µM of hexynyl-modified DNA, 300 µM of Cu(II) sulfate, 800 µM of BTTP ligand, and 2.5 mM of freshly dissolved sodium ascorbate were added to the crude azide-modified protein for 1-2 hours at room temperature. The reaction was quenched by adding 1 mM EDTA and 5 mM BCS, and then buffer exchanged into PBS + 1 mM EDTA with Zeba Spin columns for purification or storage. If using 5, 130 µM of DBCO-modified DNA was added to the azide-bearing protein overnight. For analysis, samples were denatured for SDS-PAGE and run on a 10% Tris-HCl gel, stained with either Coomassie Brilliant Blue or Sypro Red, and imaged with a Typhoon laser gel scanning system. The yield of the conjugation reactions was estimated with densitometry measurements using ImageJ software (NIH).

DNA-labeling of full-length IgG: In a 1.5 mL epi-tube was combined 2 µL of IgG (9 mg/mL in 20 mM citrate, 50 mM NaCl, pH 5.5), 1 µL of 0.5 M sodium citrate, pH 5.5, and 0, 1, 3, or 7 µL of DMT-aminoxy-DNA (2 mM in water). Water was added to bring the total reaction volume to 10 µL. The reaction was incubated for 24 h at 37 °C. Half the sample was diluted in Laemmli sample buffer containing 10 mM 2-mercaptoethanol. The samples were incubated at 95 °C for 5 min and then analyzed by SDS-PAGE. The resulting gel was stained with Coomassie Brilliant Blue and imaged with an ImageQuant LAS4000 (GE Healthcare) gel imaging station.

## PROTEIN-DNA PURIFICATION AND ASSEMBLY

Protein-DNA conjugates were purified using an HPLC system (Agilent 1200 Series) on a macroporous polystyrene anion exchange column PL-SAX 1000A 8  $\mu$ M (150 x 4.6 mm) from Agilent. The low salt buffer used was 20 mM Tris, pH 8 and high salt buffer was 20 mM Tris, pH 8 with 1 M NaCl. The conjugates were purified using a gradient method that included a ramp from 0 to 50% of high salt buffer over 15 minutes, followed by a further increase of high salt buffer from 50 to 100% in 5 minutes, then a hold at 100% high salt buffer for 5 minutes before re-equilibrating the column with the low salt buffer. Fractions were collected by monitoring absorbance at 260 and 280 nm, and protein-DNA conjugates were easily separated from both unreacted DNA and unreacted protein. A sample trace of the purification of a crude conjugation reaction is shown in Figure S6. EDTA (1mM) was added to fractions containing protein-DNA and concentrations were estimated with 280 nm absorbance values measured on a Nanodrop (Thermo Scientific, Wilmington, DE). The purity of the conjugates were verified with SDS-PAGE on a 10% Tris-HCl gel stained with Sypro Red and imaged using a Typhoon gel scanner. For assembly, each monomer was mixed together at approximately the same concentration (concentrations depended on the concentration of fractions) and incubated at room temperature for 1-2 hours. The assembly was then analyzed with SDS-PAGE for efficiency of hybridization, and if necessary, further purified with size exclusion chromatography on an HPLC using a TSKgelG3000SWxl column (TOSOH Biosciences, South San Francisco, CA) and eluted in an isocratic run of 1X PBS at 1 mL/min.

## FLOW CYTOMETRY

To generate fluorescent MBP, the amines on anion-exchange purified MBP-DNA monomers (approximately 1 mg/ml) were labeled with AlexaFluor488 succinimidyl ester (Invitrogen, resuspended at 50 µg/ul in DMSO) in 100 mM sodium bicarbonate buffer (pH 9). Reactions were then buffer exchanged twice with Zeba Spin Desalting Columns into PBS with 1 mM EDTA to remove unreacted AF488. Two AF488-labeled MBP-DNA monomers were then assembled with an anion-exchange purified Fab-DNA and purified with HPLC size exclusion chromatography as described in the previous section. Fractions containing very dilute amounts of fluorescent trimer in PBS (< 100 µg/ml) were used directly in flow cytometry experiments.

H1299 lung carcinoma cells and Human Embryonic Kidney (HEK) cells were a gift from Prof. Charles Craik and cultured in DMEM-H21 media supplemented with 10% fetal bovine serum (FBS). Cells were lifted with TrypLE (Invitrogen), washed once with fresh media and once with FACS buffer (1% BSA in PBS), and resuspended in 500 µl of FACS buffer at  $8 \times 10^5$  cells/mL. To prevent background binding of MBP to the cells, 50 µl of unlabeled purified MBP (1.64 mg/ml) was added to the cells in buffer and incubated on ice for 30 minutes. Next, a 100 µl aliquot was removed to microcentrifuge tubes for each sample (80,000 cells/sample) and all samples were centrifuged at 160xg for 5 minutes. Unlabeled cells were resuspended in either 100 µl of FACS buffer only or 100 µl of Fab-MBP(488)-MBP(488) trimer-containing FACS buffer for 2 hours at 4 °C on an orbital shaker, covered in foil to protect the samples from light. For samples that included treatment with non-fluorescent Fab, 2 µl of purified Fab (1.8 mg/ml) was added to the samples both during the MBP-blocking step and also again during the trimer-labeling step. All samples were then centrifuged at 160xg and washed 2 times in FACS buffer.

Samples were resuspended in 450  $\mu$ L of FACS buffer and analyzed on a FACSCalibur flow cytometer (BD Biosciences, San Jose, CA). FlowJo software (Tree Star, Ashland, OR) was used for data analysis.

## TRANSMISSION ELECTRON MICROSCOPY AND DATA PROCESSING

Electron Microscopy: Assembled protein-DNA dimers and trimers were visualized using a previously described protocol for negative stain electron microscopy.<sup>36</sup> Briefly, carbon-coated 400 mesh Gilder copper grids were glow-discharged using PELCO easiGlow GlowDischarge system (Ted Pella Inc., Redding, CA). 2.5  $\mu$ l of sample (< 100  $\mu$ g/ml) was deposited on a copper grid (Ted Pella Inc.) for a 30 second incubation. The sample was washed with a quick touch to 2 successive 25  $\mu$ l drops of water on Parafilm and blotted dry with filter paper after each drop. The sample was then stained in 2 successive 25  $\mu$ l drops of 0.75% uranyl acetate stain (sample was held on the second drop of stain for 30 seconds) and blotted dry after each stain. Images were collected using a Tecnai 12 TEM (FEI, Hillsboro, OR) equipped with a Gatan 4Kx4K CCD camera. The microscope was operated at 67,000x magnification (1 pixel corresponds to 0.17 nm).

Image Analysis: The coordinates of individual MBP molecules were determined by hand picking the centers of MBP molecules in 54 dimers and 51 trimers in electron microscopy images using Image J. Using Mathematica (Wolfram Research, Champaign, IL), the centroids of the dimer/trimer constructs were calculated with the coordinates of individual MBP molecules. The distance between each MBP molecule and the centroid of its corresponding dimer/trimer

was then calculated and plotted in a histogram. The Mathematica script is available upon request.

Estimation of dimer and trimer geometry: For the dimers, we estimated that the distance from the center of an MBP molecule to the center point of its corresponding dimer to be approximately 5.4 nm. This was calculated by adding the length of 10 bp of double-stranded DNA (3.4 nm) to the radius of a single MBP molecule (approximately 2 nm in our EM images). Similarly, we calculated the distance between the center of individual MBP proteins and the trimer centroid as 8.12 nm (2 nm for the MBP radius and 6.3 nm for 18 bp of DNA). The flexible tetraethylene glycol linking the protein to the oligonucleotide is approximately 1.5 nm in length if completely stretched out, but could exist in many other shorter conformations. Because the linker could place the protein either towards or away from the centroid, the linker could change the MBP to centroid distance by a range of values between -1.5 nm and +1.5 nm. However, the average effect of the linker is 0 nm, so the linker is not included in the estimates given above.

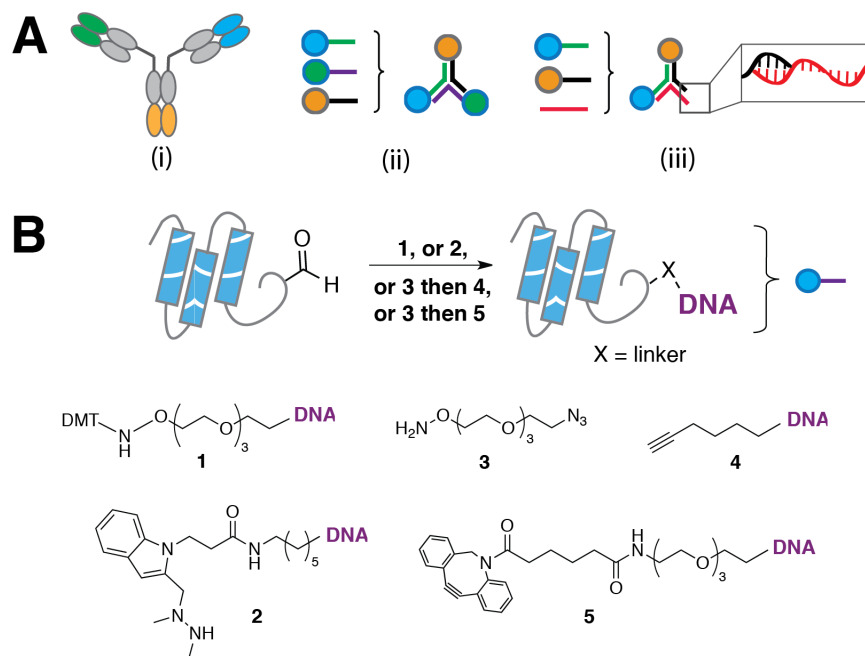


## ACKNOWLEDGMENTS

We thank Prof. Charles Craik for the uPAR Fab plasmid and Dr. Natalia Sevillano Tripero for help with Fab purification. We thank Jason Hudak for the bifunctional linker and Prof. Peng Wu for the ligand BTTP. We thank Robyn Barfield and Greg deHart for providing research materials related to FGE and MBP production. We thank Prof. Yifan Cheng and Dr. Agustin Avila-Sakar for EM training. We thank Profs. Kevan Shokat, Pam England, Danica Fujimori, and Jack Taunton for sharing instruments and facilities. We thank the members of the Gartner lab for helpful comments on the manuscript and advice on experiments. This work is partially supported by the Achievement Rewards for College Scientists Foundation Fellowship and the Genentech Foundation Fellowship to S. Liang, a UCSF CTSI-SOS pilot grant, a NIGMS Systems Biology Center Grant, and the Kimmel Family Foundation.

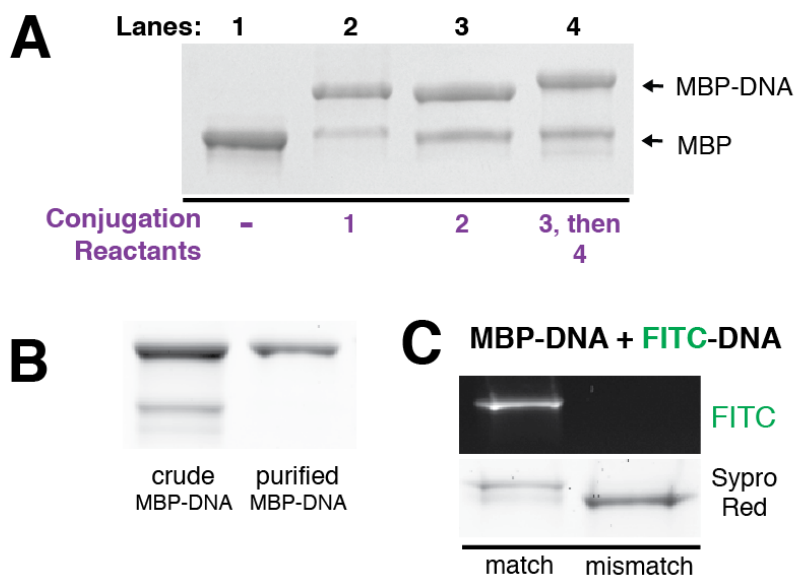
## FIGURES

**Figure 2.1:** Modular strategies for controlling antibody scaffold geometry using DNA-conjugates of aldehyde-tagged proteins.



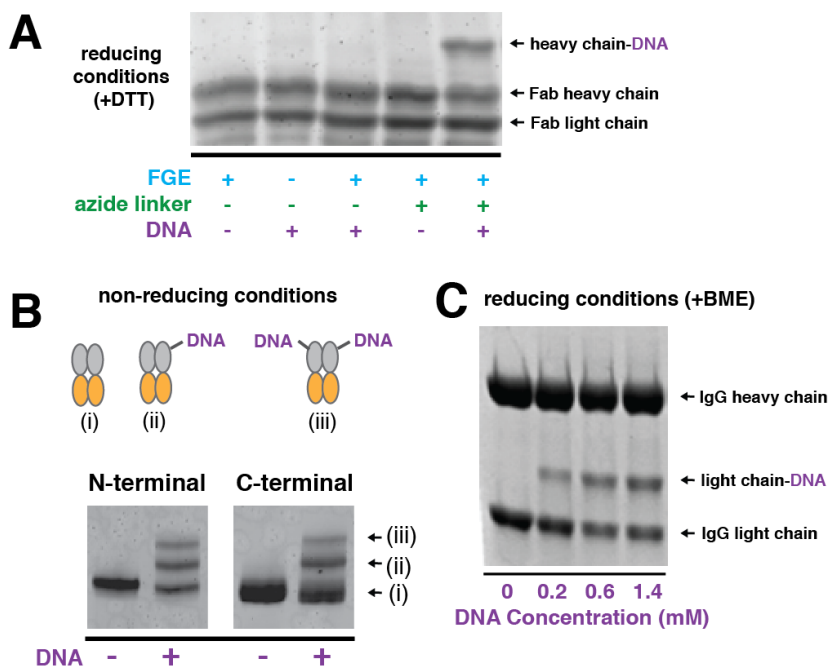
(A) (i) At least three unique functions can be incorporated on a Y-type antibody scaffold. (ii) Protein-DNA conjugates can also be used to selectively assemble antibody-bearing trimers. (iii) Additionally, the DNA may be used as a handle for assembly for of completely novel scaffold architectures. (B) The aldehyde tag can be used to site-specifically attach an oligonucleotide to protein termini or internal loops using at least four strategies: direct conjugation to DMT-protected aminoxy-modified DNA **1** or HIPS-modified DNA **2**; or indirect conjugation through bifunctional PEG **3** and subsequent copper-catalyzed triazole formation with hexynyl-modified DNA **4**, or copper-free triazole formation to cyclooctyne-modified DNA **5**.

**Figure 2.2:** Modular and site-specific conjugation of oligonucleotides to aldehyde-tagged proteins.



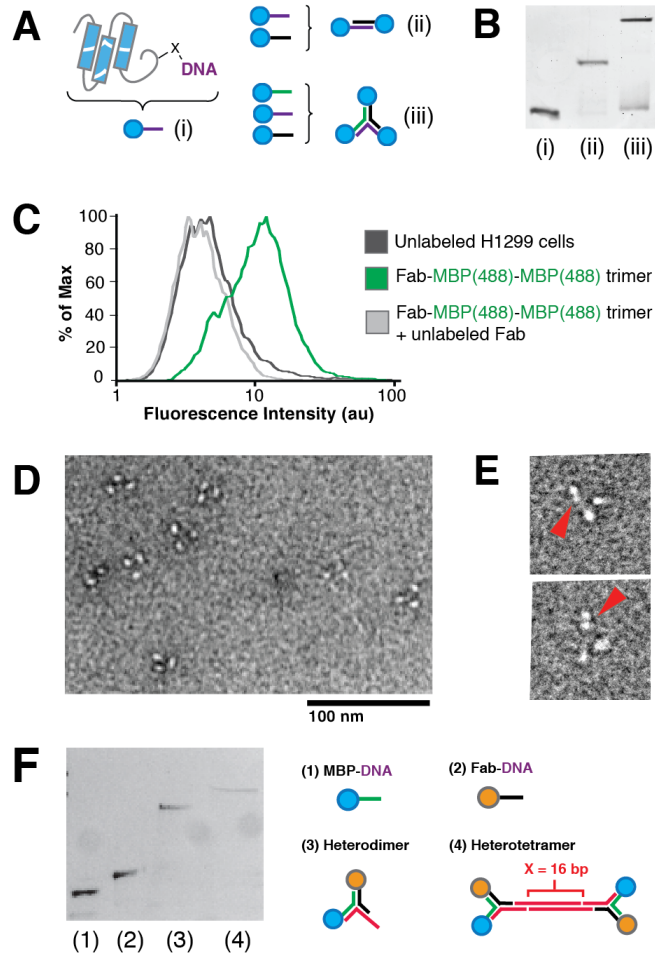
(A) Three conjugation strategies tested by incubating aldehyde-tagged Maltose Binding Protein (MBP) with the DNA conjugation reagents described and numbered in Scheme 1b. (B) MBP-DNA conjugates can be purified by anion exchange chromatography. (C) MBP-DNA conjugates incubated with complementary and non-complementary FITC-DNA only fluoresce with the complementary sequence.

**Figure 2.3:** DNA conjugation to aldehyde-tagged immunoglobulins at the N-terminus, C-terminus, and an internal loop.



(A) DNA conjugated to the N-terminus of uPAR-binding Fab expressed in *E. coli* and then treated *in vitro* with FGE. Compounds **3**, then **4** were used to label the Fab with DNA. (B) DNA-conjugated to C-terminal or N-terminal labeled Fc fragments expressed in FGE-expressing CHO cells using **3**, then **4**. (C) DNA-conjugated to an internally-labeled IgG expressed in FGE-expressing CHO cells using **1**.

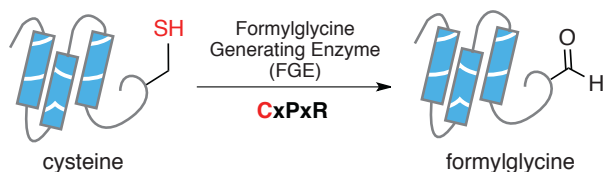
**Figure 2.4:** Modular assembly of protein-bearing DNA multimers and their interactions with live cells



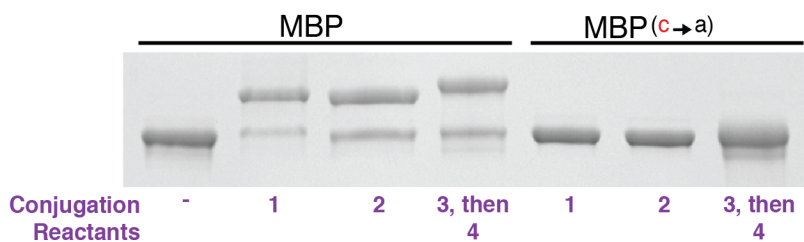
(A) Assembly of protein-DNA conjugates into dimers and trimers based on oligonucleotide sequence. (B) SDS-PAGE of DNA-linked Monomer (i), along with crude dimer (ii) and trimer (iii) assemblies. (C) Flow cytometry analysis of uPAR-expressing H1299 carcinoma cells incubated with DNA-scaffolded trimer containing Fab and Alexa488-labeled MBP. (D) A representative field of DNA-scaffolded MBP- trimers stained with uranyl acetate and imaged by transmission electron

microscopy. (E) Fab-MBP-MBP trimers. The red arrow points to the distinctly shaped Fab of the heterotrimeric construct. (F) SDS-PAGE of (1) MBP-DNA conjugate, (2) Fab-DNA conjugate, (3) a Fab and MBP bearing heterodimer, and (4) a Fab and MBP bearing heterotetramer using scaffolding strands with a variable length of X base pairs.

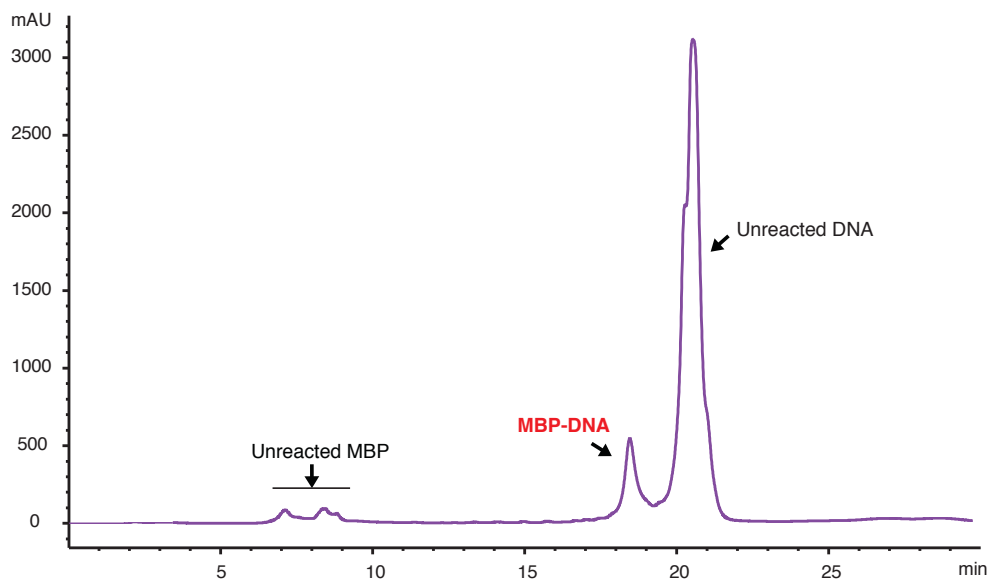
**Figure 2.5:** Proteins bearing the 5-residue aldehyde tag (CxPxR) are either co-expressed with Formylglycine Generating Enzyme (FGE) *in vivo* or treated with purified FGE *in vitro* to convert the nucleophilic cysteine to an electrophilic aldehyde.



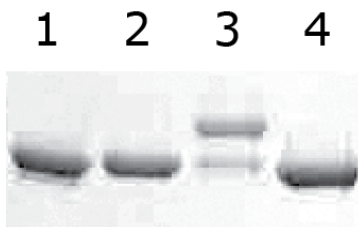
**Figure 2.6:** Three conjugations strategies tested by incubating aldehyde-tagged Maltose Binding Protein (MBP) with the DNA conjugation reagents described and numbered in Scheme 1b. When the cysteine in the aldehyde tag consensus sequence is mutated to an alanine on MBP, no conjugate is observed on an SDS-PAGE gel.



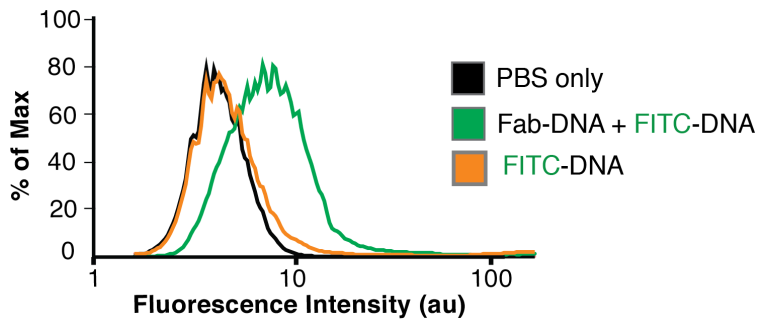
**Figure 2.7:** HPLC purification of protein-DNA conjugates using an anion exchange method. Example of the 280 nm absorbance trace using the anion exchange method on a crude MBP-DNA conjugation reaction (approximately 0.2 mg of MBP/MBP-DNA in injected sample).



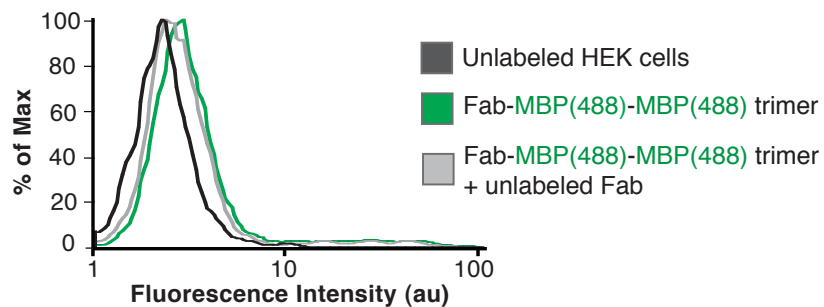
**Figure 2.8:** Aldehyde tagged Maltose Binding Protein (MBP) was reacted first with bifunctional aminoxy-azide linker **3** and then with copper free click chemistry reagent DBCO-modified DNA, **5**. Conjugate yield was determined with SDS-PAGE. Lane 1: MBP only; Lane 2: MBP and **3**; lane 3: MBP, **3**, and **5**; lane 4: MBP and **5**.



**Figure 2.9:** Purified anti-uPAR Fab-DNA conjugate (0.316 mg/ml) was hybridized with a complementary FITC-labeled oligonucleotide (0.5 mg/ml). uPAR-expressing H1299 cells were then incubated with the Fab-DNA and FITC-DNA, or FITC-DNA by itself and measured for fluorescence increase by flow cytometry.

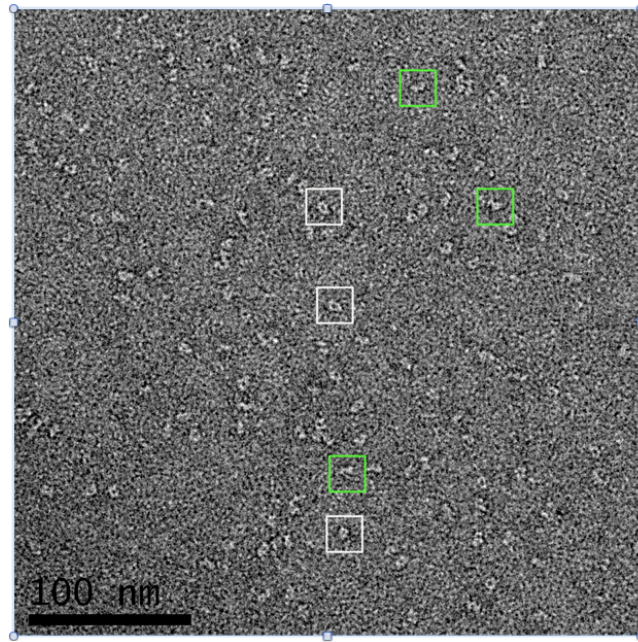


**Figure 2.10:** Flow cytometry analysis of non-uPar-expressing HEK cells incubated with DNA-scaffolded trimer containing Fab and AlexaFluor488-labeled MBP.

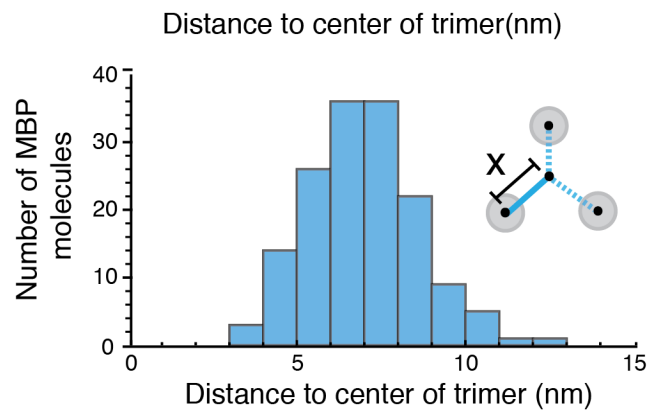




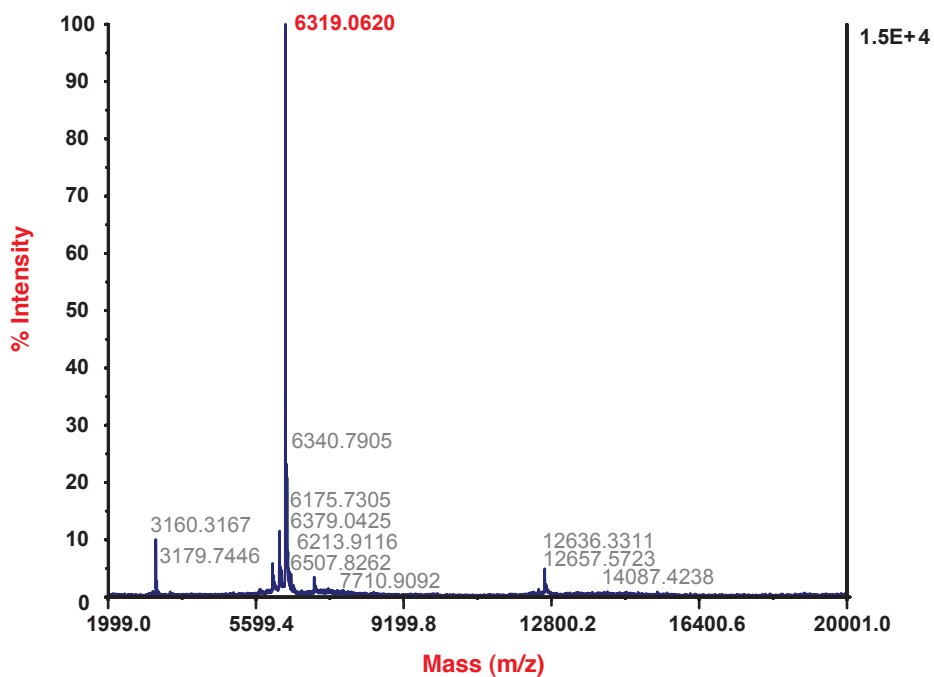
**Figure 2.11:** Unmodified Fabs stained with uranyl acetate and imaged by transmission electron microscopy. White and green squares each indicate one of two distinct shapes that can be identified as Fabs.



**Figure 2.12:** Distribution of distances between MBP centers and the center point of trimers generated with 26 bp long oligonucleotides.



**Figure 2.13:** Representative MALDI-MS trace for synthesized hexynyl-modified oligonucleotides. Sample is dimer strand "Hexynyl-D2" 20mer, expected MW [M+H]= 6238. Measurement was taken in reflector mode and positive ion mode.



## REFERENCES

- (1) Pinheiro, A. V.; Han, D.; Shih, W. M.; Yan, H. *Nature Nanotech* 2011, 6, 763–772.
- (2) Niemeyer, C. M. *Angew. Chem. Int. Ed.* 2010, 49, 1200–1216.
- (3) Simmel, F. C. *Curr. Opin. Biotechnol.* 2012, 23, 516–521.
- (4) Kazane, S. A.; Axup, J. Y.; Kim, C. H.; Ciobanu, M.; Wold, E. D.; Barluenga, S.; Hutchins, B. A.; Schultz, P. G.; Winssinger, N.; Smider, V. V. *J Am Chem Soc* 2013, 135, 340–346.
- (5) Carter, P. J. *Exp Cell Res* 2011, 317, 1261–1269.
- (6) Beck, A.; Wurch, T.; Bailly, C.; Corvaia, N. *Nat. Rev. Immunol.* 2010, 10, 345–352.
- (7) Kontermann, R. *MAbs* 2012, 4, 182–197.
- (8) Chari, R. V. J. *Acc. Chem. Res.* 2008, 41, 98–107.
- (9) Alley, S. C.; Okeley, N. M.; Senter, P. D. *Curr Opin Chem Biol* 2010, 14, 529–537.
- (10) Gautier, A.; Juillerat, A.; Heinis, C.; Corrêa, I. R.; Kindermann, M.; Beaufils, F.; Johnsson, K. *Chem. Biol.* 2008, 15, 128–136.
- (11) Keppler, A.; Gendreizig, S.; Gronemeyer, T.; Pick, H.; Vogel, H.; Johnsson, K. *Nat Biotechnol* 2003, 21, 86–89.
- (12) Los, G. V.; Encell, L. P.; McDougall, M. G.; Hartzell, D. D.; Karassina, N.; Zimprich, C.; Wood, M. G.; Learish, R.; Ohana, R. F.; Urh, M.; Simpson, D.; Mendez, J.; Zimmerman, K.; Otto, P.; Vidugiris, G.; Zhu, J.; Darzins, A.; Klauert, D. H.; Bulleit, R. F.; Wood, K. V. *ACS Chemical Biology* 2008, 3, 373–382.
- (13) George, N.; Pick, H.; Vogel, H.; Johnsson, N.; Johnsson, K. *J Am Chem Soc* 2004, 126, 8896–8897.
- (14) Chen, I.; Howarth, M.; Lin, W.; Ting, A. Y. *Nat Meth* 2005, 2, 99–104.
- (15) Lin, C.-W.; Ting, A. Y. *J Am Chem Soc* 2006, 128, 4542–4543.
- (16) Antos, J. M.; Miller, G. M.; Grotenbreg, G. M.; Ploegh, H. L. *J Am Chem Soc* 2008, 130, 16338–16343.
- (17) Fernández-Suárez, M.; Baruah, H.; Martínez-Hernández, L.; Xie, K. T.; Baskin, J. M.; Bertozzi, C. R.; Ting, A. Y. *Nat Biotechnol* 2007, 25, 1483–1487.
- (18) Lovrinovic, M.; Spengler, M.; Deutsch, C.; Niemeyer, C. M. *Mol BioSyst*, 1, 64–69.
- (19) Carrico, I. S.; Carlson, B. L.; Bertozzi, C. R. *Nature Chemical Biology* 2007, 3, 321–322.
- (20) Hudak, J. E.; Barfield, R. M.; de Hart, G. W.; Grob, P.; Nogales, E.; Bertozzi, C. R.; Rabuka, D. *Angew Chem Int Ed Engl* 2012, 51, 4161–4165.
- (21) Wu, P.; Shui, W.; Carlson, B. L.; Hu, N.; Rabuka, D.; Lee, J.; Bertozzi, C. R. *Proc. Natl. Acad. Sci. U.S.A.* 2009, 106, 3000–3005.
- (22) Rabuka, D.; Rush, J. S.; deHart, G. W.; Wu, P.; Bertozzi, C. R. *Nat Protoc* 2012, 7, 1052–1067.
- (23) Alley, S. C.; Benjamin, D. R.; Jeffrey, S. C.; Okeley, N. M.; Meyer, D. L.; Sanderson, R. J.; Senter, P. D. *Bioconjugate Chem* 2008, 19, 759–765.
- (24) Shen, B.-Q.; Xu, K.; Liu, L.; Raab, H.; Bhakta, S.; Kenrick, M.; Parsons-Reponte, K. L.; Tien, J.; Yu, S.-F.; Mai, E.; Li, D.; Tibbitts, J.; Baudys, J.; Saad, O. M.; Scales, S. J.; McDonald, P. J.; Hass, P. E.; Eigenbrot, C.; Nguyen, T.; Solis, W. A.; Fuji, R. N.; Flagella,

- K. M.; Patel, D.; Spencer, S. D.; Khawli, L. A.; Ebens, A.; Wong, W. L.; Vandlen, R.; Kaur, S.; Sliwkowski, M. X.; Scheller, R. H.; Polakis, P.; Junutula, J. R. *Nat Biotechnol* 2012, 30, 184–189.
- (25) Agarwal, P.; van der Weijden, J.; Sletten, E. M.; Rabuka, D.; Bertozzi, C. R. *Proc. Natl. Acad. Sci. U.S.A.* 2013, 110, 46–51.
- (26) Agarwal, P.; Kudirka, R.; Albers, A. E.; Barfield, R. M.; de Hart, G. W.; Drake, P. M.; Jones, L. C.; Rabuka, D. *Bioconjugate Chem* 2013, 24, 846–851.
- (27) Wang, W.; Hong, S.; Tran, A.; Jiang, H.; Triano, R.; Liu, Y.; Chen, X.; Wu, P. *Chem Asian J* 2011, 6, 2796–2802.
- (28) Duriseti, S.; Goetz, D. H.; Hostetter, D. R.; LeBeau, A. M.; Wei, Y.; Craik, C. S. *Journal of Biological Chemistry* 2010, 285, 26878–26888.
- (29) Lund, I. K.; Illemann, M.; Thurison, T.; Christensen, I. J.; Høyer-Hansen, G. *Curr Drug Targets* 2011, 12, 1744–1760.
- (30) Sa-Ardyen, P.; Vologodskii, A. V.; Seeman, N. C. *Biophys J* 2003, 84, 3829–3837.
- (31) Zheng, J.; Lukeman, P. S.; Sherman, W. B.; Micheel, C.; Alivisatos, A. P.; Constantinou, P. E.; Seeman, N. C. *Biophys Journal* 2008, 95, 3340–3348.
- (32) Hsiao, S. C.; Shum, B. J.; Onoe, H.; Douglas, E. S.; Gartner, Z. J.; Mathies, R. A.; Bertozzi, C. R.; Francis, M. B. *Langmuir* 2009, 25, 6985–6991.
- (33) Sil, D.; Lee, J.; Luo, D.; Holowka, D.; Baird, B. *ACS Chemical Biology* 2007, 2, 674–684.
- (34) Carrico, I. S.; Carlson, B. L.; Bertozzi, C. R. *Nature Chemical Biology* 2007, 3, 321–322.
- (35) Rabuka, D.; Rush, J. S.; deHart, G. W.; Wu, P.; Bertozzi, C. R. *Nat Protoc* 2012, 7, 1052–1067.
- (36) Booth, D. S.; Avila-Sakar, A.; Cheng, Y. *J Vis Exp* 2011.

## **CHAPTER 3**

Modular assemblies of biparatopic antibodies on  
DNA scaffolds targeting urokinase plasminogen  
activator receptor

## BACKGROUND

Bispecific antibodies are gaining popularity as next-generation biopharmaceuticals that can target multiple molecular cancer targets at once.<sup>1-3</sup> In addition to increasing avidity and thus therapeutic potency, combining proteins with different targets and functions can add a targeting element to an otherwise nonspecific therapeutic, improve the pharmacokinetic profile of a rapidly cleared molecule, or expand access to new therapeutically relevant targets.<sup>4,5</sup> Currently, these bispecifics are primarily generated from hybridomas or as genetic fusions with the two proteins linked at their termini. However, these limited topologies are not ideal for every protein combination as some proteins require unmodified termini for optimal bioactivity or can suffer from expression difficulties due to folding and processing issues. Additionally, controlling the geometry and valency of these proteins in a systematic fashion to identify the protein arrangement with the best therapeutic potential remains challenging, if not impossible.

Illustrating this point - inhibitory antibodies targeting two distinct subdomains of the urokinase plasminogen activator receptor (uPAR), a cell surface scaffolding protein whose overexpression contributes to the aggressive phenotype of a number of cancers due to its involvement in signaling pathways involved in migration and invasion,<sup>6-8</sup> were recently used to slow tumor growth in a mouse model of triple negative breast cancer.<sup>9</sup> These antibodies, referred to as 2G10 and 3C6, obstruct interactions between uPAR and uPA, which results in the release of metalloproteases that digest the extracellular matrix next to the cell and give the cell space to crawl, or  $\beta$ 1 integrin, which signals to the cell to activate its migration signaling

pathways, respectively<sup>10</sup> (Figure 3.1). Because these Fabs block independent functions on distinct structural domains of uPAR (Figure 3.2A), combinations of these antibodies could generate a synergistic inhibitory effect on uPAR function. A bispecific molecule is particularly attractive because 3C6 has a low binding affinity for uPAR and initial binding of uPAR by a 2G10 molecule would serve to increase the effective molarity of 3C6. However, identifying a scaffold capable of mediating this type of “biparatopic” interaction is a major challenge (Fig. 3.2b). Additionally, bispecific molecules could engage uPAR either intra- or intermolecularly (Fig. 3.2c-e), making it difficult to predict which scaffold geometric conformation is most likely to succeed in biparatopic inhibition. To overcome this challenge, we decided to use a synthetic and modular DNA-based scaffold to rapidly explore bispecific antibody structural geometries, valencies, and flexibilities that maximize synergistic binding of 2G10 and 3C6 to uPAR and inhibition of cancer cell matrix adhesion and invasion.

## APPROACH

As described in Chapter 2, I developed a simple method for assembling aldehyde-tagged Fabs using DNA scaffolds to generate multivalent molecules with precise geometries and compositions<sup>11</sup> (Figure 3.3a). The method is efficient, highly modular, and provides a means of rapidly screening tens to hundreds of scaffold geometries, providing maximum therapeutic potential in cell-based assays. In short, I developed a site-specific protein-DNA

conjugation strategy utilizing a genetically incorporated aldehyde tag, which is small (CxPxR) and not limited to placement at N or C terminus.<sup>12-15</sup> Synthesis of DNA-Fab conjugates occurs with high yield (Figure 3.3b) and combinations of conjugates efficiently assemble into predictably structured oligomers through Watson-Crick base pairing (Figure 3.3c). We used this strategy to synthesize two initial DNA-Fab conjugates in large scale and use unmodified scaffold strands to organize these conjugates into a variety of geometric conformations (Figure 3.3d and Figure 3.4). A critical and enabling feature of this strategy is that unmodified scaffold strands are fast and inexpensive to synthesize in high yield and purity. Theoretically, we could explore a large number of valencies, Fab combinations, and architectures from only a few protein-DNA conjugates (Figure 3.4). In practice, we sought to explore a few molecules first for our proof-of-concept experiments.

First, we wanted to assemble *flexible* bispecific 3C6/2G10 Fab dimers (as well as 3C6/3C6 and 2G10/2G10 dimers as controls). Two unmodified scaffold strands will template the assembly of a single DNA-3C6 and DNA-2G10 conjugate. By varying the length of the central duplex region, the overall length of the scaffold can be systematically extended (Figure 3.3e). Two nicks in the duplex permit flexibility in the construct to maximize the potential of binding. I will also assemble a library of *rigid* and *multivalent* DNA-scaffolded anti-uPAR Fabs using geometrically constrained constructs. Circularly permuting the scaffold strand around the DNA-linked Fab molecules will allow multiple inter-Fab spacings to be rapidly screened (Figure 3.3f). I will additionally explore rigid four-arm scaffolds (Figure 3.3e) and control the spacing of the arms using a central scaffold strand. Moreover, valency is easily controlled by replacing DNA-conjugated Fabs with dummy unmodified DNA strands on any of the four scaffold arms.



All constructs would be characterized by SDS-page and negative stain transmission electron microscopy. Their ability to bind to uPAR-expressing cells would be confirmed with flow cytometry experiments.

The second step would be to screen assembled anti-uPAR Fabs for synergistic inhibition of ECM binding and invasion. The activities of uPAR-binding Fabs have been characterized with assays that test for (1) cell adhesion to fibronectin mediated by uPAR interactions with  $\beta$ 1 integrin, (2) MAPK activation downstream of integrin and growth factor receptor signaling, and (3) cell motility and invasion. We wanted apply these assays to screen for the synergistic inhibition of these uPAR effector functions by DNA-scaffolded Fabs.

These assays were pretty well defined from previous projects from the Craik lab at UCSF. To screen for inhibition of cell adhesion by blocking uPAR access to  $\beta$ 1 integrin, we would add fluorescently stained MDA-MB-231 cells to fibronectin-coated 96-Well plates supplemented with RGD peptides or an RAD peptide as a control. The RGD peptide forces cells to adhere to fibronectin via uPAR-mediated activation of  $\beta$ 1 integrin. Alterations to cell adhesion caused by the addition of specific DNA-scaffolded Fabs will be quantified using a fluorescence plate reader. For the second assay, we would use an immunoblot assay to screen for DNA-scaffolded anti-uPAR Fabs that decrease phosphorylation of Erk relative to controls. We would plate cells on fibronectin-coated plates, then incubate them with uPA and DNA-scaffolded anti-uPAR Fabs. Cell lysates will be probed with anti-phospho-Erk antibodies to measure the extent of pathway activation. Finally, to screen for inhibition of matrix invasion, We would use a quantitative cell invasion assay that uses matrigel coated Fluoroblock Boyden

Chambers to measure decreases in MDA-MB-231 cell invasion overnight caused by DNA-scaffolded anti-uPAR Fabs.

We expect to identify a collection of DNA-scaffolded bispecifics that synergistically inhibit uPAR. Modification of internal bases (e.g. off the major groove of a dT residue) will allow fine-tuning of scaffold geometry for any "hits". We would then optimize the geometry of the lead scaffolds, explore homologation of scaffold chemistry to phosphorothioate backbones,<sup>16-17</sup> or peptide nucleic acids and then initiate preclinical animal studies.

## RESULTS

We expressed and purified a large batch of each of the aldehyde tagged uPAR-binding Fabs using CHO cells that were overexpressing Formylglycine Generating Enzyme. This particular batch of Fabs did not need to be treated with FGE in vitro and coupled with the DMT-aminooxy-modified oligonucleotides at 80-90% yield. We purified them using anion exchange chromatography and assembled them into dimers (Figure 3.5) by just mixing the Fabs at a 1:1 ratio at room temperature. 2G10-HetA and 3C6-HetA' dimerized efficiently enough that size exclusion chromatography was not necessary.

First we made sure that DNA-conjugation did not prevent the Fabs from binding to uPAR by performing flow cytometry assays with the Fab-DNA conjugates (Figure 3.6). We labeled 2G10 with AlexaFluor488 and 3C6 with AlexaFluor647 by buying NHS-ester conjugates to those dyes and labeling free lysines on the Fabs with the dyes. We then determined if the

Fabs had bound the cells by measuring the fluorescence in these two channels for the cells. We confirmed that the 2G10 Fab bound better with the uPAR-expressing H1299 cells compared to the 3C6 Fab. Dimerizing 3C6-DNA to 3C6-DNA to increase the avidity of 3C6 marginally improved its binding to cells. Dimerization of 3C6-HetA with 2G10-HetA' increases its binding to H1299 cells, more so than the dimerization of 3C6-HetA to 3C6-HetA or the 3C6-HetA conjugate itself. Dimerizing 3C6 to itself using DNA increases its binding to cells compared to monomer, but dimerizing 3C6 to 2G10 increases its binding even more. Dimerization of 3C6 to 2G10 significantly increases its binding to cells compared to 3C6 and 2G10 added in tandem without being dimerized by DNA. In the control, the 2G10 is not conjugated to DNA and cannot dimerized with 3C6. Overall, this confirms our hypothesis that 2G10 would facilitate the delivery of 3C6 to cells based on its increased affinity for uPAR.

Next, we wanted to do a small proof of concept experiment just to see if the a biparatopic antibody bearing both 2G10 and 3C6 would limit the invasion of a triple negative breast cancer cell line more than the two Fabs in tandem but not connected as one macromolecule (Figure 3.7). Although it was a small sample size and the variability of the assay resulted in large error bars with only three replicates, we did see some very promising trends that we hope will hold with increased replicates. When 2G10 and 3C6 were hybridized together via DNA linkers, we saw that the invasiveness of the cells was decreased even more compared to the condition of either Fab, or the Fabs in tandem. We believe that when the two Fabs are attached together the 2G10 is at least delivering the 3C6 antibody more effectively and that this is helping inhibit the invasiveness of the cells through inhibition of the integrin signaling pathway. Our hope is that increased replicates and more exploration of different

biparatopics and multivalent 2G10 and 3C6 combos will lead us to some promising combinations that will greatly inhibit uPAR signaling.

## FUTURE DIRECTIONS

This project is an ongoing collaboration between the Gartner lab, the Craik lab, and Catalent-West and will continue to be worked on by other members of these labs. More conjugates and a small biparatopic library will continue to be tested for affinity and efficacy. At the moment, we are building dimers, trimers, and tetramers of 3C6 and 2G10 with DNA, as well as directly conjugating 3C6 and 2G10 with a chemical linker as another control to bring the two domains closer together.

If successful, the biparatopic compounds generated from this project can be used as lead compounds as a first-in-class line of therapy for triple negative breast cancer. Additionally, this approach is an unexplored method to access protein geometries where the simultaneous binding of multiple components is essential. This modular and high throughput platform for the discovery of bi- and multi-specific immunotherapies represents a new approach to discover therapies with unprecedented pharmacological activities. We view this technology to be a high throughput way to screen different combinations, architectures, and valencies, but do not necessarily envision the final therapeutic to be held together with oligonucleotides due to issues with stability. However, there are many options to that could be used to replace the DNA, including an oligonucleotide with a phosphothioate backbone, a 2'-O-methyl RNA

backbone, or replacing everything with peptic nucleic acids. These bispecific compounds are an attractive "combination" therapy with potentially lower dosing, fewer side effects, and the ability to generate entirely new therapeutic effects as a consequence of their defined nanoscale geometries.

## MATERIALS AND METHODS

### DNA OLIGONUCLEOTIDES

DMT-protected aminoxy-modified oligonucleotides were synthesized on an Applied Biosystems Expedite 8909 DNA synthesizer using their respective phosphoramidites (100 mM) and a default coupling protocol. Synthesized DMT-protected aminoxy-modified DNA, were cleaved from solid support and deprotected using a 1:1 mixture of ammonium hydroxide/40% methylamine (AMA) for 15 minutes at 65°C. After removing the AMA or ammonium hydroxide with a speedvac system, oligonucleotides were resuspended in 100 mM triethylamine acetate (TEAA) and filtered through 0.2- $\mu$ m spin filters. Oligonucleotides were then purified by reversed-phase high-performance liquid chromatography (HPLC) using an Agilent 1260 Infinity Series HPLC System, a 5  $\mu$ m XDB-C18 (21.2x150mm) Agilent (Santa Clara, CA) column, and 100 mM TEAA (pH 7)/acetonitrile as the mobile phase. Fractions were collected based on their absorbance at 260 and 280 nm, and fractions containing hexynyl-modified oligonucleotides were washed with Millipore MilliQ-purified water and lyophilized three times to remove residual HPLC buffer salts and resuspended in water prior to use. DNA concentrations were

determined by absorbance at 260 nm. As a note, sometimes after lyophilization and resuspension, the DMT group will fall off of a fraction of the oligonucleotides. Because of this, the resuspended DNA will sometimes appear yellow-tinted in color, but it will still work for the conjugation reaction as expected.

Any unmodified oligonucleotides that were used as linker strands were purchased from Integrated DNA Technologies (Coralville, Iowa) and were resuspended in water at 200  $\mu$ M.

Oligonucleotide Sequences: DNA sequences were selected as previously described to form dimer, trimer, and tetramer constructs are described in the Methods section of Chapter 2 and are also included in the appendices of this thesis.

## **EXPRESSION AND PURIFICATION OF ALDEHYDE-TAGGED FABS**

Expression plasmids for uPAR-binding Fab 2G10 and 3C6 were gifts from Prof. Charles Craik (UCSF). An aldehyde tag was cloned onto the C-terminus of a His-tagged heavy chain using primers and reagents for Quikchange cloning. BL21-Gold E. coli were transformed with the Fab expression plasmid and grown in 2xYT media supplemented with 0.2% glucose and 100  $\mu$ g/ml of Ampicillin at 37 °C. When  $OD_{600}$  reached 0.6, cultures were then cooled to room temperature, induced with 1mM IPTG, and grown overnight at 30 °C. The periplasmic fraction of the cells was collected by pelleting the bacteria, resuspending the cell pellet in 20 ml of ice-cold sucrose buffer (200 mM Tris, 0.5 mM EDTA, 500 mM sucrose, pH 8), then quickly mixing 20 ml of ice cold distilled water to the resuspended cells. The cells were incubated on ice for 30 minutes, with gentle swirling of the samples every 10 minutes. The cells were then removed

by centrifugation at 13000xg for 15 minutes at 4 °C and the supernatant was collected as the periplasmic fraction. The Fab was purified from the periplasmic fraction using Ni-NTA-agarose beads and standard purification procedures. Fab was eluted in 50 mM Tris, 500 mM NaCl, 500 mM imidazole, pH 8 and dialyzed into PBS for storage. Protein concentrations were determined with a BCA protein assay (Thermo Scientific Pierce, Rockford, IL). In order to convert the cysteine of the aldehyde tag to formylglycine, the aldehyde-tagged Fabs (25-80  $\mu$ M) were incubated in 50 mM Tris, 100 mM NaCl, 2 mM DTT, pH 9 and 0.1 equivalents of M. tuberculosis FGE for 8-20 hours at 30 °C

As a part of a new collaboration with Redwood Bioscience, which has now become Catalent-West, we gave them the DNA sequences of 2G10 and 3C6 light and heavy chains and they cloned them into a mammalian CHO-expression system that has been optimized to overexpress FGE to eliminate the need for an *in vitro* FGE treatment. Much of the Fab in later experiments was produced via this protocol, which gave a better yield of protein and enabled larger scale experiments.

### CONJUGATION OF ALDEHYDE-TAGGED PROTEINS TO DNA

Formylglycine-bearing proteins were treated with 250  $\mu$ M DMT-protected aminoxy-modified DNA in potassium acetate buffer (final concentration 100 mM, pH 4.6) and incubated overnight at 37 °C. They were then buffer exchanged using Pierce Zeba Desalting columns into PBS, pH 7 for storage and purification. The yield of the conjugation reaction was analyzed via SDS-PAGE and a coomassie stain before purification.

## PROTEIN-DNA PURIFICATION AND ASSEMBLY

Protein-DNA conjugates were purified using an HPLC system (Agilent 1200 Series) on a macroporous polystyrene anion exchange column PL-SAX 1000A 8  $\mu$ M (150 x 4.6 mm) from Agilent. The low salt buffer used was 20 mM Tris, pH 8 and high salt buffer was 20 mM Tris, pH 8 with 1 M NaCl. The conjugates were purified using a gradient method that included a ramp from 0 to 50% of high salt buffer over 15 minutes, followed by a further increase of high salt buffer from 50 to 100% in 5 minutes, then a hold at 100% high salt buffer for 5 minutes before re-equilibrating the column with the low salt buffer. Fractions were collected by monitoring absorbance at 260 and 280 nm, and protein-DNA conjugates were easily separated from both unreacted DNA and unreacted protein. A sample trace of the purification of a crude conjugation reaction is shown in Figure S6. Concentrations were estimated with 280 nm absorbance values measured on a Nanodrop (Thermo Scientific, Wilmington, DE). The purity of the conjugates were verified with SDS-PAGE on a 10% Tris-HCl gel stained with Coomassie Blue and imaged using a scanner. For assembly, each monomer was mixed together at approximately the same concentration (concentrations depended on the concentration of fractions) and incubated at room temperature for 1-2 hours. The assembly was then analyzed with SDS-PAGE for efficiency of hybridization. If necessary, we could have further purified with size exclusion chromatography on an HPLC using a TSKgelG3000SWxl column (TOSOH Biosciences, South San Francisco, CA) and eluted in an isocratic run of 1X PBS at 1 mL/min, but that did not seem to be necessary in this particular application.

## FLOW CYTOMETRY



To generate fluorescent Fabs for flow cytometry, the amines on anion-exchange purified Fabs and Fab-DNA monomers (approximately 1 mg/ml) were labeled with AlexaFluor488 succinimidyl ester or AlexaFluor647 succinimidyl ester (Invitrogen, resuspended at 50 µg/ul in DMSO) in 100 mM sodium bicarbonate buffer (pH 9). Reactions were then buffer exchanged twice with Zeba Spin Desalting Columns into PBS with 1 mM EDTA to remove unreacted dyes. The Fab and Fab-DNA dye-labeled monomers were then assembled into dimers, or else used directly to label cells in flow cytometry experiments.

H1299 lung carcinoma cells and Human Embryonic Kidney (HEK) cells were cultured in DMEM-H21 media supplemented with 10% fetal bovine serum (FBS). Cells were lifted with TrypLE (Invitrogen), washed once with fresh media and once with FACS buffer (1% BSA in PBS), and resuspended in 500 µl of FACS buffer at  $1 \times 10^6$  cells/mL. Next, a 100 µl aliquot was removed to microcentrifuge tubes for each sample (100,000 cells/sample) and all samples were centrifuged at 160xg for 5 minutes. Unlabeled cells were resuspended in either 100 µl of FACS buffer only or 100 µl of Fab monomer or Fab dimer in FACS buffer for 2 hours at 4 °C on an orbital shaker, covered in foil to protect the samples from light. All samples were then centrifuged at 160xg and washed 2 times in FACS buffer. Samples were resuspended in 450 µL of FACS buffer and analyzed on a FACSCalibur flow cytometer (BD Biosciences, San Jose, CA). FlowJo software (Tree Star, Ashland, OR) was used for data analysis.

## INVASION ASSAY

The baskets of a boyden chamber assay are lined with growth-factor reduced matrigel, and cells in serum free media are placed on top of the matrigel. Then the baskets are placed

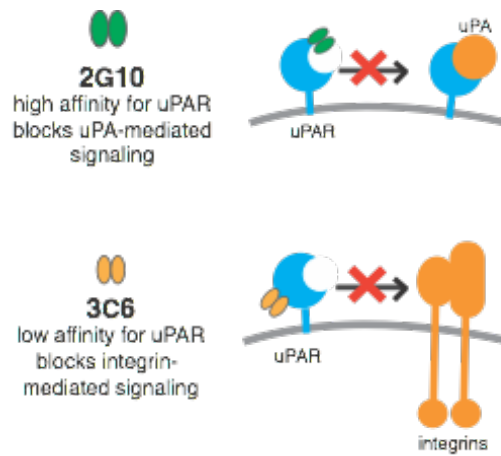
into wells with media plus serum such that there is a gradient of serum from the top of the well to the bottom. We used the cell line MDA-MB-231s (a triple negative breast cancer cell line) and the assay was ran overnight at 37C. Cells are stained with Calcein AM and a fluorescence reading is taken at the height of the basket bottom at the end of the experiment to see how many cells migrated through the matrigel and got stuck on the filter at the bottom of the basket. We did three replicates per condition, although to get significant error bars, we would need to do many more replicates in the future.

## ACKNOWLEDGMENTS

Prof. Charles Craik, Efrat Harel, Jesse McFarland, and David Rabuka have been indispensable partners on this project. The uPAR-inhibiting Fabs were originally developed in Charly Craik's lab. David Rabuka and Jesse McFarland helped with obtaining high quality aldehyde-tagged Fab. I designed the DNA scaffold strategy, synthesized and purified the Fab-DNA conjugates and did the flow cytometry experiment. Efrat Harel performed the preliminary invasion assay experiment and will be taking on a lot more of the responsibility for screening the biparatopic Fabs for uPAR-inhibiting activity as the project continues.

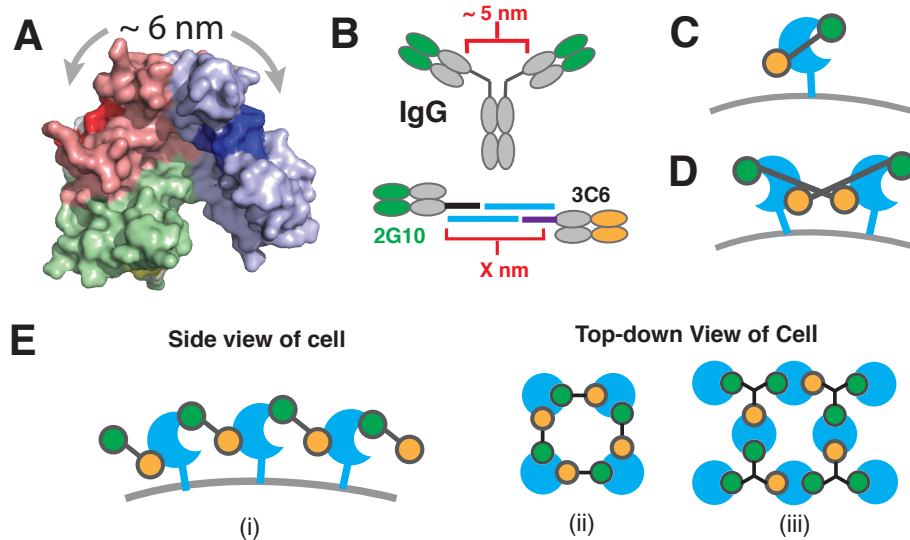
# FIGURES

**Figure 3.1:** uPAR and structural analysis for a bispecific antibody.



uPAR integrates signals for migration and proliferation and uPAR overexpression contributes to the aggressive phenotype of a number of cancers, including triple negative breast cancer. The Fab 2G10 was identified out of a phage display screen to have a high affinity for uPAR and also block uPA-mediated signaling, which results in the release of metalloproteases to digest the extracellular matrix. The Fab 3C6 was identified because it blocked the ability of uPAR to signal through integrins, which is an cellular signaling pathway that is important to adhesion and migration. 3C6 has a low affinity for uPAR, potentially because it binds a site that is close to the cell membrane and is difficult to access.

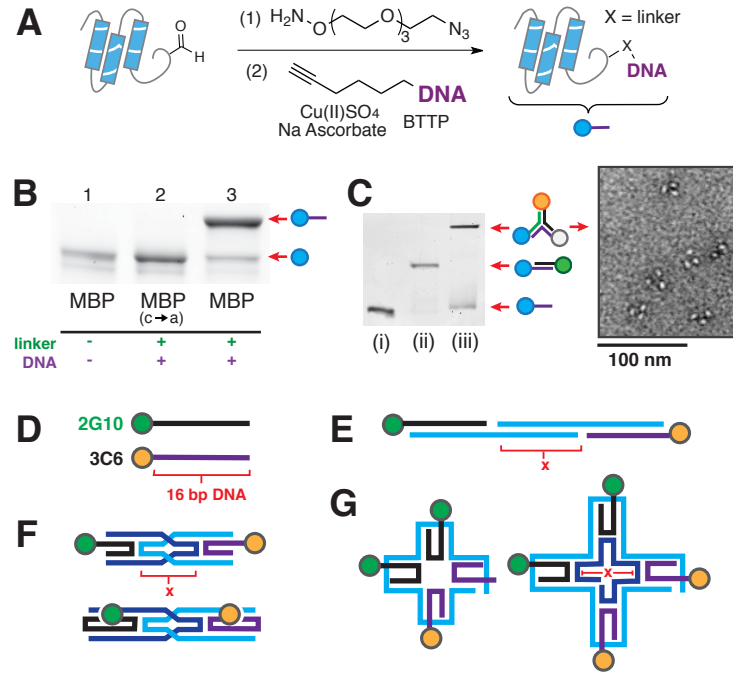
**Figure 3.2:** uPAR and structural analysis for a bispecific antibody.



(A) Crystal structure of uPAR with uPA binding region (2G10 target) shaded dark blue,  $\beta 1$  integrin interacting region (3C6 target) shaded red, and GPI membrane anchor shaded yellow.

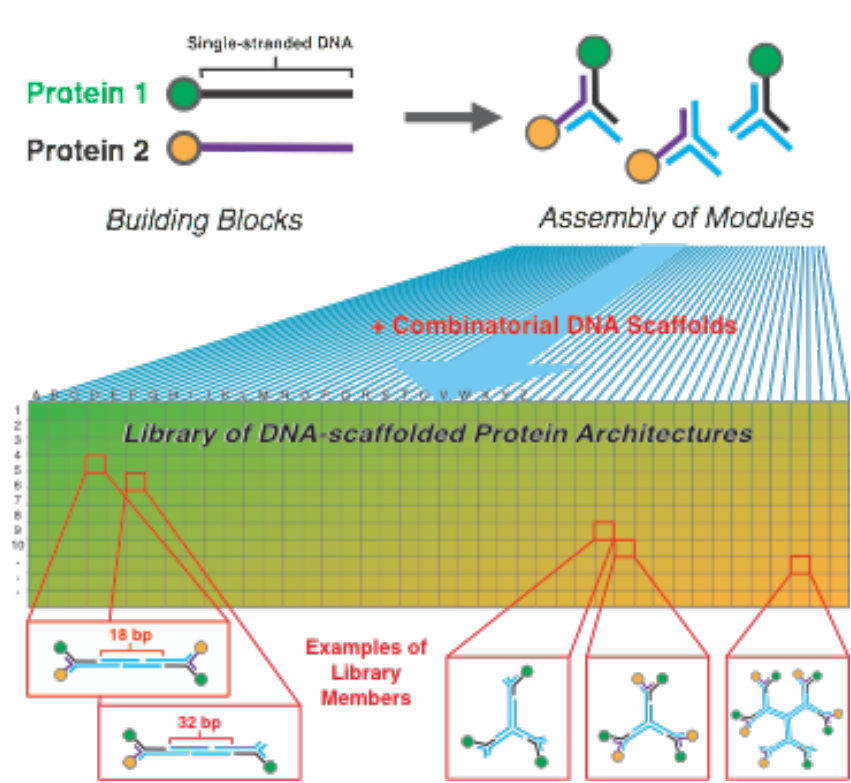
(B) Scaffold dimensions of an IgG showing fixed 5 nm scaffolding between Fabs and a depiction of a bispecific antibody with an adjustable scaffold. (C) Schematic of biparatopic binding of uPAR on a single scaffold. Note that a IgG scaffold geometry is unlikely to engage both epitopes of uPAR in this format. (D) Intermolecular engagement of uPAR via two DNA-scaffolded bispecifics in a cross motif. (E) Oligomerization of multiple uPAR receptors with multiple DNA-scaffolded bispecifics.

**Figure 3.3.** Modular DNA-programmed assembly of Fabs to vary the flexibility and valency of biparatopic antibody scaffold



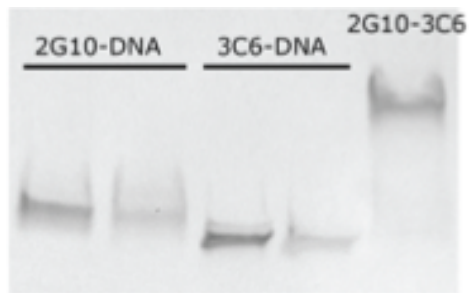
(A) Scheme for a two-step DNA-conjugation with a site-specific aldehyde tag. (B) SDS-PAGE analysis of conjugation reactions and controls reveals a greater than 80% yield. (C) SDS-PAGE analysis of monomers (i), dimers (ii), trimers (iii), and transmission electron microscopy of trimers. (D) Two unique oligonucleotides are conjugated to aldehyde-tagged 2G10 and 3C6. (E) Flexible linear heterodimer templated by 2 unmodified blue strands. The region of duplex DNA between the two template strands defines the distance "x". Three "x" lengths will be explored for each scaffold shown. (F) Locked and rigid scaffold that can be circularly permuted to adjust protein distance. (G) Trimers and tetramers assembled on 4-arm DNA scaffolds using a single long template strand.

**Figure 3.4.** Alternative scheme for modular DNA-programmed assembly of Fabs based on using the trimer motif with a handle strand.



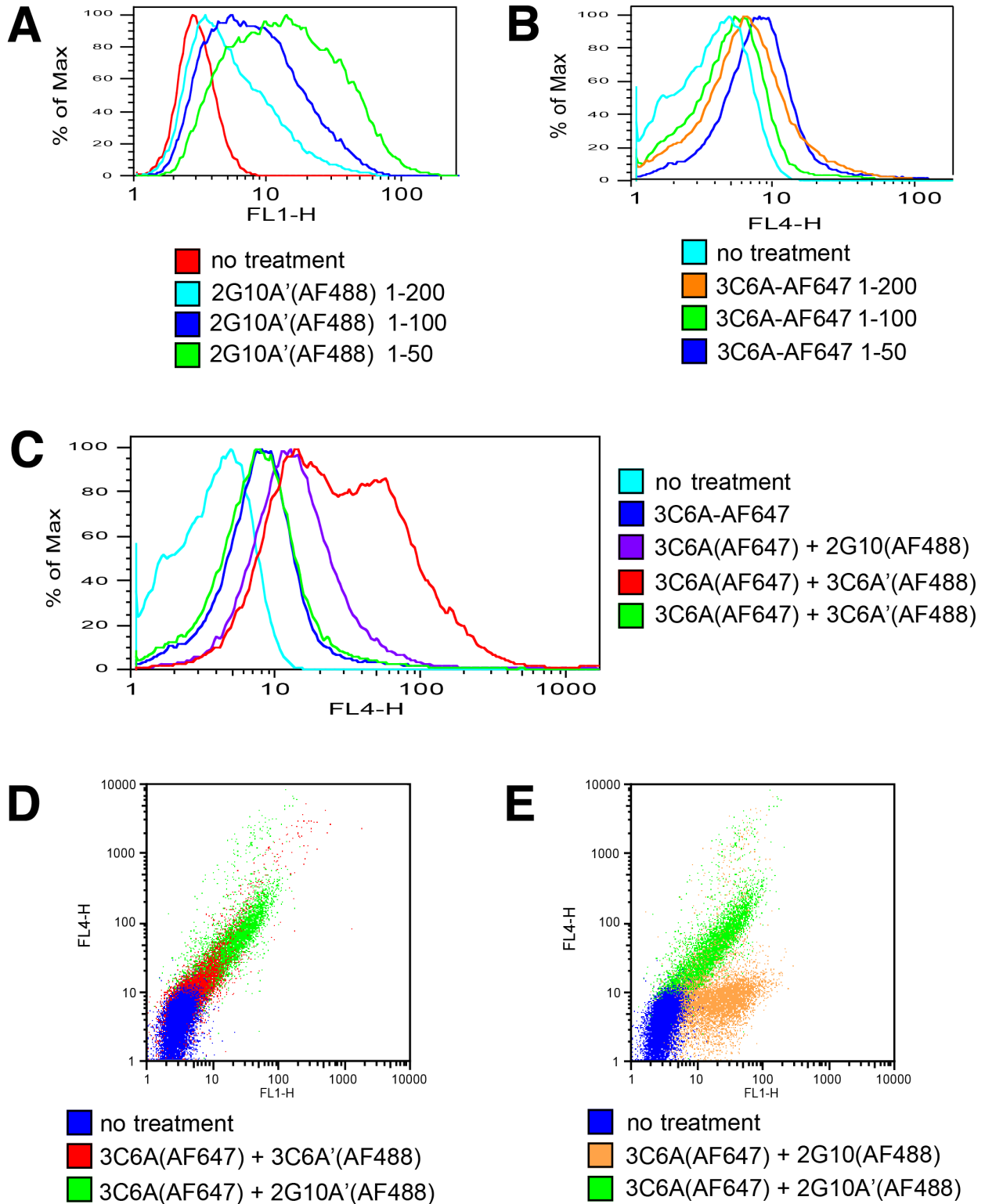
This scheme describes an alternative library of multi-specific antibodies that gives us more variety in the architecture of the scaffold as we increase the valency of our antibodies. Importantly, the strands attached to protein 1 (which would be 2G10) and protein 2 (which would be 3C6) would be the same for all of the constructs in the library. Thus only two conjugates would need to be purified and many constructs could be assembled using inexpensive and easy to synthesize unmodified “folding strands” of DNA.

**Figure 3.5:** uPAR and structural analysis for a bispecific antibody.



SDS-PAGE gel showing the purified conjugates of 2G10-DNA and 3C6-DNA. The last lane shows the efficient assembly of the 2G10-3C6 heterodimer based on the hybridization of the oligonucleotides attached to 2G10 and 3C6, which are each 20 nucleotides long.

**Figure 3.6:** 2G10 and 3C6 biparatopic antibodies binding to H1299 cells, which overexpress uPAR analyzed by flow cytometry.

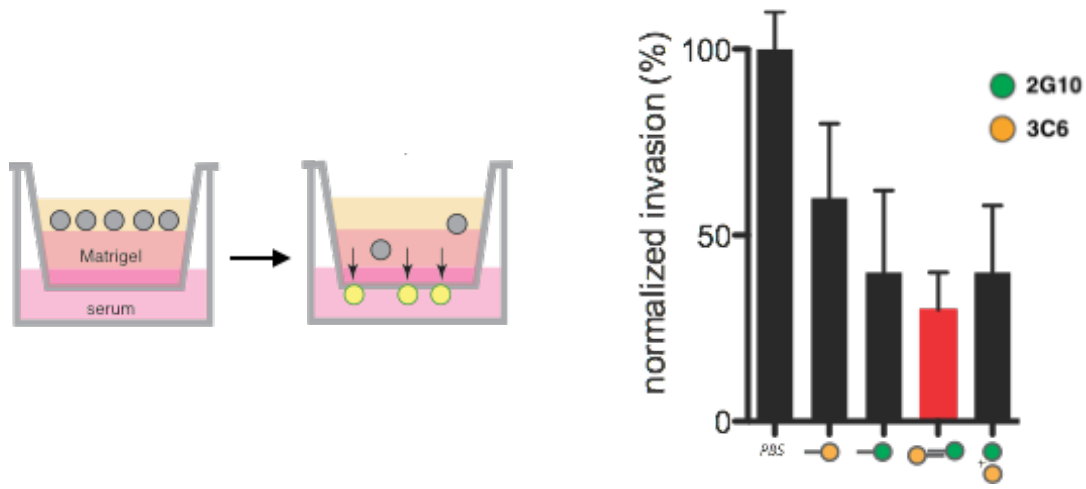




(A) Binding of 2G10-HetA' conjugate, labeled with AlexaFluor488 at increasing concentration.

(B) Binding of 3C6-HetA conjugate, labeled with AlexaFluor647 at increasing concentration. It is evident here that 3C6 has a worse affinity for uPAR than 2G10. (C) Dimerization of 3C6-HetA with 2G10-HetA' increases its binding to H1299 cells, more so than the dimerization of 3C6-HetA to 3C6-HetA or the 3C6-HetA conjugate itself. (D) Dimerizing 3C6 to itself using DNA increases its binding to cells compared to monomer, but dimerizing 3C6 to 2G10 increases its binding even more. (E) Dimerization of 3C6 to 2G10 significantly increases its binding to cells compared to 3C6 and 2G10 added in tandem without being dimerized by DNA. In the control, the 2G10 is not conjugated to DNA and cannot dimerized with 3C6.

**Figure 3.7:** Invasion of MDA-MB-231 cells after treatment with uPAR-inhibiting Fabs



Preliminary data from the invasion assay which measure the invasiveness of triple negative breast cancer cell line through extracellular matrix towards a growth-factor rich media. We showed that treatment with either 3C6-DNA or 2G10-DNA inhibited the invasion of the cells, although 2G10-DNA inhibited invasion more, most likely because of its higher affinity. We then showed the addition of non-DNA labeled 3C6 and 2G10 in tandem to cells did not reduce the invasiveness of the cells much compared to just 2G10-DNA alone. However with 2G10 and 3C6 were hybridized together via DNA linkers, we saw that the invasiveness of the cells was decreased even more compared to the condition of either Fab, or the Fabs in tandem.

## REFERENCES

- (1) Carter PJ. Introduction to current and future protein therapeutics: A protein engineering perspective. *Experimental Cell Research*. 2011 May;317(9):1261–9.
- (2) Schmidt SR. Fusion-proteins as biopharmaceuticals--applications and challenges. *Curr Opin Drug Discov Devel*. 2009 Mar;12(2):284–95.
- (3) Leader B, Baca QJ, Golan DE. Protein therapeutics: a summary and pharmacological classification. *Nat Rev Drug Discov*. 2008 Jan;7(1):21–39.
- (4) Doppalapudi VR, Huang J, Liu D, Jin P, Liu B, Li L, et al. Chemical generation of bispecific antibodies. *Proc Natl Acad Sci USA*. National Acad Sciences; 2010;107(52):22611–6.
- (5) Huang H, Lai JY, Do J, Liu D, Li L, Del Rosario J, et al. Specifically Targeting Angiopoietin-2 Inhibits Angiogenesis, Tie2-Expressing Monocyte Infiltration, and Tumor Growth. *Clinical Cancer Research*. 2011 Mar 1;17(5):1001–11
- (6) Dass K, Ahmad A, Azmi AS, Sarkar SH, Sarkar FH. Evolving role of uPA/uPAR system in human cancers. *Cancer Treatment Reviews*. 2008 Apr;34(2):122–36.
- (7) Gelder MEMM-V, Look MPM, Peters HAH, Schmitt MM, Brünner NN, Harbeck NN, et al. Urokinase-type plasminogen activator system in breast cancer: association with tamoxifen therapy in recurrent disease. *Cancer Res*. 2004 Jul 1;64(13):4563–8.
- (8) Harbeck, Kates, Schmitt, Gauger, Kiechle, Janicke, et al. Urokinase-Type Plasminogen Activator and Its Inhibitor Type 1 Predict Disease Outcome and Therapy Response in Primary Breast Cancer. *Clin Breast Cancer*. 2004 Dec 1;5(5):5–5.
- (9) LeBeau AM, Duriseti S, Murphy ST, Pepin F, Hann B, Gray JW, VanBrocklin HF, Craik CS. Targeting uPAR with antagonistic recombinant human antibodies in aggressive breast cancer. *Cancer Res*. 2013 Apr 1; 73(7):2070-81
- (10) Duriseti S, Goetz DH, Hostetter DR, LeBeau AM, Wei Y, Craik CS. Antagonistic anti-urokinase plasminogen activator receptor (uPAR) antibodies significantly inhibit uPAR-mediated cellular signaling and migration. *J Biol Chem*. 2010 Aug 27;285(35):26878–88. PMID: PMC2930687
- (11) Liang SI, McFarland JM, Rabuka D, Gartner ZJ. A Modular Approach for Assembling Aldehyde-Tagged Proteins on DNA Scaffolds. *J Am Chem Soc*, 136(31), 10850-3, 2014
- (12) Hudak JE, Barfield RM, de Hart GW, Grob P, Nogales E, Bertozzi CR, et al. Synthesis of Heterobifunctional Protein Fusions Using Copper-Free Click Chemistry and the Aldehyde Tag. *Angew Chem Int Ed Engl*. 2012;51(17):4161–5.
- (13) Carrico IS, Carlson BL, Bertozzi CR. Introducing genetically encoded aldehydes into proteins. *Nat Chem Biol*. 2007 Apr 22;3(6):321–2.

- (14) Rabuka D, Rush JS, Dehart GW, Wu P, Bertozzi CR. Site-specific chemical protein conjugation using genetically encoded aldehyde tags. *Nat Protoc.* 2012 May 10;7(6):1052–67.
- (15) Wu P, Shui W, Carlson BL, Hu N, Rabuka D, Lee J, et al. Site-specific chemical modification of recombinant proteins produced in mammalian cells by using the genetically encoded aldehyde tag. *Proc Natl Acad Sci USA.* 2009 Mar 3;106(9):3000–5. PMID: PMC2651276
- (16) Behlke MA. Chemical modification of siRNAs for in vivo use. *Oligonucleotides.* 2008 Dec;18(4):305–19
- (17) Geary RS, Henry SP, Grillone LR. Fomivirsen: clinical pharmacology and potential drug interactions. *Clin Pharmacokinet.* 2002;41(4):255–60.

## **CHAPTER 4**

Nanoscale dimerization of EGFR on live cells

## BACKGROUND

The misregulation or mutation of components of the Epidermal Growth Factor Receptor (EGFR) family are implicated in the pathophysiology of a number of disease<sup>1</sup>. EGFR receptors are part of a layered signaling network activated by many ligands and integrating with pathways important for development, apoptosis, and growth<sup>2</sup>. EGFR is targeted by many successful chemotherapies. Yet, resistance to these chemotherapies inevitably arises and in many cases the mechanism of resistance is unclear, suggesting that we may still be missing information on how these receptors signal and are regulated despite the plethora of studies on the EGFR family.

Although genetic and biochemical studies have provided much clarity on individual steps in the activation of EGFR and its downstream effectors, we still do not understand much about the spatial organization of EGFR and how that can impact signaling. Monomers of EGFR are inactive, but the binding of its ligand (EGF) causes structural changes that open up a dimerization domain on EGFR, a dimer of EGFR forms, and the intracellular kinase domains of these dimers are able to phosphorylate each other tails<sup>3</sup>. Intracellular scaffolding and signaling proteins then recognize these phosphorylated tails and lead to downstream signals – including the MapK/ERK cascade and the PI3K/AKT cascade, which are associated with cell proliferation and migration (Figure 4.1). However, recent studies suggest dimerization alone may not be sufficient to achieve maximal activation. In fact, EGFR sometimes exists in the form of inactive dimers prior to ligand binding in structural and diffusion based studies<sup>4</sup> (Figure 4.3A) and is now thought to exist in an equilibrium of monomer and inactive dimer in resting

state depending on the cell type and also even different subcellular areas on the cells.

Additionally, EGF addition causes significant rearrangement of EGFR on a cell surface in a minutes (Figure 4.2). Recent evidence suggests a role for EGFR multimeric cluster formation in receptor activation<sup>5</sup>. Still, it remains unclear if these higher order structures are a cause or a consequence of activation. This is primarily because it has been challenging to observe both the formation of higher order structures and downstream signaling at the same time in the same assay. Because EGF-stimulation causes very fast signaling and spatial changes, it is difficult to know if the nanoscale clusters are important to signaling, or simply just a transition state as EGFR moves into clathrin-coated pits and gets endocytosed to have its signals down regulated.

I hypothesized that these higher order structures play a role in directing EGFR signaling output. Although we can now see and appreciate nanoscale structures using super-resolution imaging techniques, I wanted to introduce “spatial mutations” at this length scale so that I could investigate whether these nanoscale structures serve any role or are merely a consequence of other signaling steps that are more important. My goals were to determine whether applying a “gain-of” dimerization spatial mutation to EGFR would (1) cause ligand-independent activation, (2) alter the sensitivity of downstream signaling pathways to EGF, and (3) influence rates of EGFR endocytosis for the reasons below:

Constitutive EGFR activation has been observed in cancer cells that overexpress EGFR<sup>28</sup>. Additionally, high EGFR expression levels in mouse fibroblasts leads to cellular transformation and increased cell motility independent of ligand binding<sup>9</sup>. It is possible that constitutive activation is a consequence of increased density and clustering of EGFR. I

hypothesized that dimerization would result in EGF-independent activation. To investigate this, we first stably expressed EGFR with a Snap-tag on its N-terminus on HEK cells. We also synthesized a library of benzylguanine-modified DNA oligonucleotides that would covalently link to EGFR at the Snap-tag and form dimers based on hybridization with matching DNA strands (Figure 4.3).

Studies also show that EGFR phosphorylation can propagate laterally and amplify signaling<sup>11</sup>. Nanoscale organization may contribute to this observation, affecting the cooperativity and sensitivity of receptors by a variety of mechanisms, including a change in lateral signal propagation rates. Because of this, we thought it was possible that the EGFR organizations would affect cell sensitivity to EGF and wanted to look into this.

Finally, endocytosis of EGFR plays an integral role in regulating the strength and duration of downstream signals<sup>12</sup>. Interestingly, endocytosis is dependent on receptor dimerization rather than the activation of EGFR tyrosine kinase domains<sup>13</sup>. We hypothesized that the generation of higher order EGFR structures will lead to changes in the rates of endocytosis. Using these DNA scaffolds, I will be able to systematically perturb the nanoscale organization of cell surface EGFR. The proposed experiments will demonstrate whether EGFR organization influences EGFR activation, cell sensitivity to EGF, and EGFR endocytosis.

## RESULTS



In order to organize EGFR in a ligand-independent manner, we utilized the SNAP-tag, a 20 kDa tag that we fused onto the N-terminus of EGFR and transduced in HEK293 cells, which have low amounts of endogenous EGFR. The Snap-Tag will selectively recognize its ligand benzylguanine and form a covalent bond to it. Thus, when we synthesize benzylguanine-modified DNA and add it to live cells, we can selectively and covalently label EGFR with DNA.

After characterizing this system, we first wanted to ask a simple question: **What happens when you dimerize EGFR without ligand?** We sought to answer this question by adding a 20 bp Benzylguanine-modified DNA dimer to cells expressing Snap.EGFR and synthetically dimerizing the receptors. What we observed was that dimerization of EGFR with DNA is sufficient to induce ligand independent phosphorylation of EGFR (Figure 4.4, 4.5). However, in comparison to the ligand EGF, the DNA guided dimerization does not activate the downstream pathways of pERK or pAKT significantly (Figure 4.4, 4.5).

With this bizarre result, we came up with a few ideas of how we could get this combination of phosphorylation. We thought it was possible that EGFR was not fully phosphorylated in the DNA case, that perhaps it was adopting a non-biologically relevant intracellular conformation that would prevent adaptors from binding, or that perhaps it was even just an artifact of the cell line. Our next few experiments sought to eliminate these hypotheses. Probing a little deeper, we looked at the phosphorylation of different tyrosine residues on EGFR's tail and found that the residues often associated with MAPK and AKT pathways looked pretty similar in both conditions despite there not being much downstream signaling (Figure 4.6). To address the intracellular conformation caveat, we saw that when the EGFR kinase is inhibited with gefitinib, the DNA-induced dimers are not auto-phosphorylated

(Figure 4.7A). Additionally, DNA-induced dimerization of an EGFR mutant that cannot form the asymmetric allosteric-activating dimer results in no autophosphorylation (Figure 4.7B). These results taken together suggested that the DNA-induced dimers are auto-phosphorylating via a similar mechanism to EGF-activated dimers. We also made sure that this result was not an artifact of our cell line and tested our system in 4 other cell lines, getting similar results in each of those cases (Figure 4.8).

When we performed a co-immunoprecipitation assay with EGFR with both DNA and EGF stimulation, we found the DNA-induced dimers were capable of recruiting key signaling adaptor proteins Grb2 and SOS to similar levels seen with EGF-stimulation (Figure 4.9). However, we discovered the breakdown in the signaling pathway occurred at the level of Ras activation (Figure 4.9), in that the DNA-induced dimers were not capable of activating Ras despite that being the next step in the signaling pathway after recruitment of adaptor proteins.

We thought that endocytosis might be important and found through confocal microscopy that the DNA-induced EGFR dimers were not endocytosed despite being phosphorylated (Figure 4.11) but we also found that inhibition of endocytosis did not inhibit EGF-stimulated activation of downstream pathways, leading us to believe that although DNA does not induce endocytosis, this is not the reason why it also does not induce downstream signaling. Additionally, we found that DNA-induced dimers do not get transported to clathrin-coated pits like EGF-stimulated EGFR dimers and multimers (Figure 4.12). Interestingly, we saw that EGFR still colocalizes with clathrin coated-pits when stimulated with EGF, even when the cells are treated with gefitinib, which should prevent EGFR autophosphorylation. This suggested to us that EGF causes a conformation change that signals to the receptor to form

multimers and get recruited to clathrin-coated pits. Finally, we saw that DNA also does not cause nanoscale clustering of EGFR, which we observe after EGF treatment (Figure 4.13). We believe that these nanoscale clusters are formed because of a conformational change that EGF induces and that these are an important signaling unit that serves as a platform to concentrate adaptor proteins in order to fully activate Ras and trigger the rest of the downstream signaling pathways.

## FUTURE DIRECTIONS

By being able to “stop on a dimer” and isolate this phosphorylated EGFR dimer, we can now investigate the importance of the multimers of EGFR to signaling. Our hypothesis is that you need nanoscale multimers of EGFR in order to trigger the Ras activation and downstream signaling pathways. This is supported with preliminary data where we believe that EGFR creates multimers of 6-10 EGFR molecules based on some single molecule pull-down assays that we have performed in collaboration with Tae-Young Yoon (KAIST). We are now pursuing many techniques to cluster EGFR independent of ligand addition into larger nanoscale clusters in order to rescue signaling through downstream signaling pathways, including using magnetic tweezers and watching cells that have a fluorescent phosphorylated-ERK reporter in their nuclei.

In more advanced studies, similar experiments could be performed on cells that natively express EGFR and other ErbB receptors. Because ErbB receptors heterodimerize<sup>13</sup>, future

experiments could determine if EGFR organization affects the organization and activation of other ErbB receptors. Combined with the ability to image receptor organization using super resolution microscopy, our DNA-scaffolding techniques will allow structure-function relationships for EGFR to be constructed at the nanoscale.

Significantly, the technology developed in this proposal can be applied to other signaling systems where evidence of higher-order organization exists, such as the growth hormone<sup>33</sup> or insulin receptor signaling pathways<sup>34</sup>. Understanding in detail the effect of nanoscale structure on receptor functions could lead to strategies that more precisely control receptor organization, and therefore cellular responses. As a consequence, these nanostructured molecules may provide a foundation for diagnostics or therapeutics that can either create nanoscale structures to influence cell behavior, or perhaps selectively target nanoscale structures associated with disease states.

## MATERIALS AND METHODS

### SYNTHESIS OF BENZYLGUANINE-MODIFIED OLIGONUCLEOTIDES

Benzylguanidine-conjugated N-Hydroxysuccinimide (BG-GLA-NHS) were purchased from New England Biolab (Ipswich, MA). Amine-modified oligonucleotides were synthesized on an Applied Biosystems Expedite 8909 DNA synthesizer using a amine phosphoramidite (100 mM) as the last 5' base and a default coupling protocol. Then, up to 5 columns of 1  $\mu$ mol scale synthesis of amine-modified DNA still on solid supports were transferred microcentrifuge tubes. 2 mg of the BG-GLA-NHS ester was resuspended in 400  $\mu$ l of dry DMSO, mixed with

100  $\mu$ l of dry DIPEA under argon, then quickly transferred evenly among the tubes to completely submerge the DNA-bearing beads. The reaction was parafilmmed and incubated on a vortexer overnight. The beads were then washed 3 times with DMF, 3 times with DCM, 1 time with acetonitrile, and then dried completely on a speed vac system. The DNA was cleaved off the beads with 30% ammonium hydroxide for 2 hours at 65 °C, dried on a speed vac system, resuspended in 100 mM triethylamine acetate (TEAA), and filtered through 0.2- $\mu$ m spin filters. Oligonucleotides were then purified by reversed-phase high-performance liquid chromatography (HPLC) as described in previous chapters. The BG-conjugated strand with a 3' conjugated biotin was made by synthesizing the DNA using a purchased biotin-conjugated solid support and following the same procedures as described above.

## **PLASMIDS AND LENTIVIRUS**

Lentiviral and retroviral plasmids for EGFR, Snap.EGFR, Snap.EGFR.GFP, and Snap.EGFR(VR) were a gift from Prof. Natalia Jura. The plasmid for Snap.EGFR.mEOS was made by cloning out GFP from the Snap.EGFR.GFP plasmid and replacing it with mEOS amplified using an enzymatic inverse PCR strategy.

## **STABLE CELL LINE GENERATION**

Lentivirus or retrovirus plasmids were delivered to UC San Francisco's Viracore for virus production. Cells were plated at 25% confluency on 10 cm diameter tissue culture plates and incubated in 5 ml of media with 250  $\mu$ l of virus at 37 degrees. After 6-24 hours, 5 ml of media was added to the cells. After 48 hours, safely discard the virus-contaminated media on the cells

and replace with free media with 10% serum. After 72 hours, cells are ready to be split, selected, or sorted depending on the application.

### **STIMULATING CELLS WITH LIGANDS OR DIMERIZERS**

Cells are typically plated in 12-well plates for my experiments. Cells are serum starved with 1 ml of serum-free DMEM-H21 for 6-8 hours for lowest amount of background phosphorylation of EGFR/ERK/AKT. Then, 6X stocks of the stimuli (EGF, BG-DNA, PEG) are made. For 6X BG-DNA stock (12  $\mu$ M), add 5  $\mu$ l of 240  $\mu$ M BG-20A and 5  $\mu$ l of 240  $\mu$ M BG-20A' to 90  $\mu$ l of serum free media . For 6X BG-PEG-BG stock (12  $\mu$ M), add 5  $\mu$ l of 240  $\mu$ M BG-PEG-BG to 95  $\mu$ l of serum free media. For 6X EGF, I make a 300 ng/ml stock (final concentration in experiment will be 50 ng/ml). Before starting time points, the 1 ml of serum-free media off the cells and wells are "primed" with 250  $\mu$ l of serum-free media. Make sure my 6X stocks are warm or at room temperature. Stimulate cells at various time points by adding 50  $\mu$ l of the 6X stimulus. End the experiment by moving samples onto ice, quickly washing 2 times with ice cold PBS wash, and lysing the cells with 60  $\mu$ l of lysis buffer on ice for 15 minutes. Then scrape the cells and lystates in the plate with the back of a pipette tip, transfer the lystates to microcentrifuge tubes and spin the samples at the highest speed for 10 minutes at 4  $^{\circ}$ C in order to pellet out the nuclei and other cell debris. Transfer the supernatant to a new tube and add 4X SDS-Page buffer with DTT or BME as a reducing agent. Store lysates at -20 $^{\circ}$ C and minimize freeze thaw cycles. For co-immunoprecipitation experiments, the same protocol was followed but scaled up so that the volumes worked for a 10 cm diameter tissue culture plate.

## WESTERN BLOTTING

For western blots, samples were loaded onto BioRad 4-15% gradient TGX gels in SDS-PAGE buffer. For denatured DNA-treated samples that run as a monomer, the samples were boiled immediately before loading on the gel. The gels are run for 250V for 35 minutes and the buffer is typically pretty warm after the run. For non-denatured samples where the dimer band remains intact, the samples are not boiled before loaded on the gel and the gels are run at 200V for 45-60 minutes in pre-chilled running buffer. After the run, the buffer should still be cold. All western blots are transferred using a BioRad Criterion wet transfer system onto nitrocellulose (BioRad) in a tris-glycine buffer with 20% methanol. The transfer is run at 150V in a chilled cold buffer, with an ice block and running stir bar to keep the buffer chilled during the run, for 25 or 27 minutes for a monomer or dimer transfer respectively. The blots are then washed briefly with 1X TBST and blocked in 5% milk in TBST for at least 1 hour at room temperature, rocking. The blots are then washed 3 times with 1X TBST and incubated in the primary antibody overnight on a rocker. The majority of primary antibodies used were purchased from Cell Signaling and used at 1:1000 in 5% BSA in TBST. The loading control antibodies (alpha-tubulin and beta-actin) were purchased from Millipore and used at 1:10000. The next day, the blots were washed 4 times for 5 minutes each in 1X TBST. Then, they were incubated in secondary antibodies (Invitrogen secondary antibodies conjugated to either AlexaFluor680 or AlexaFluor800 diluted 1:20000 into 5% milk in TBST) for 1-2 hours at room temperature, protected from light. After 3 more washes in 1X TBST, the blots were imaged on a Licor imaging system. Scans were then quantified and analyzed on ImageStudioLite (Licor).

## CONFOCAL MICROSCOPY

For the imaging the colocalization of EGF or DNA with an endosomal marker EEA1, we plated the cells in 8-well chamber slides and stimulated with biotinylated-EGF and BG-DNA/BG-DNA-biotin. Then we fixed the cells, permeabilized, and stained with streptavidin-conjugated Alexafluor488 and also a primary antibody for EEA1, and a secondary antibody for the EEA1 antibody that was labeled with Alexafluor647. Samples were then imaged using a Zeiss confocal microscope.

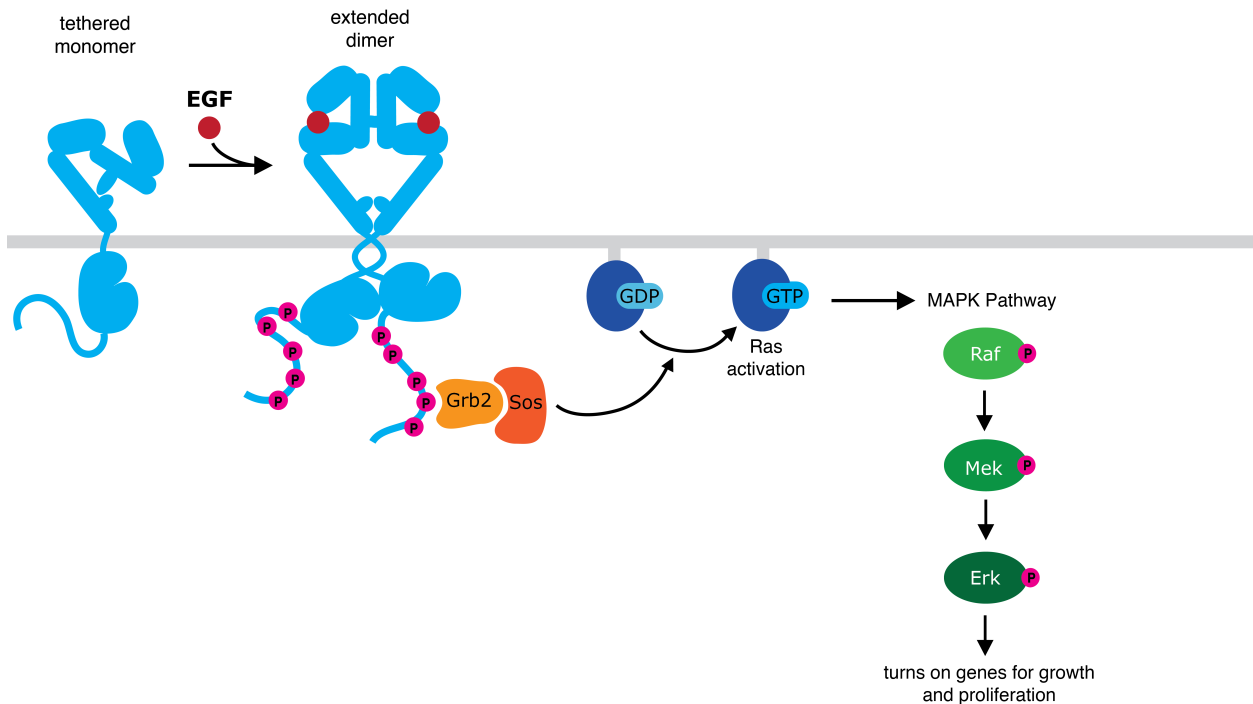
## ACKNOWLEDGMENTS

Prof. Natalia Jura and Bettina van Lengerich have been close collaborators on this product. Bettina made the majority of the plasmid constructs with the exception of one and performed the STORM experiments. They have both contributed much intellectually to the direction of this project. Prof. Mark von Zastrow and Kelsie Eichel have shared their expertise with endocytosis and TIRF microscopy with us for this project.



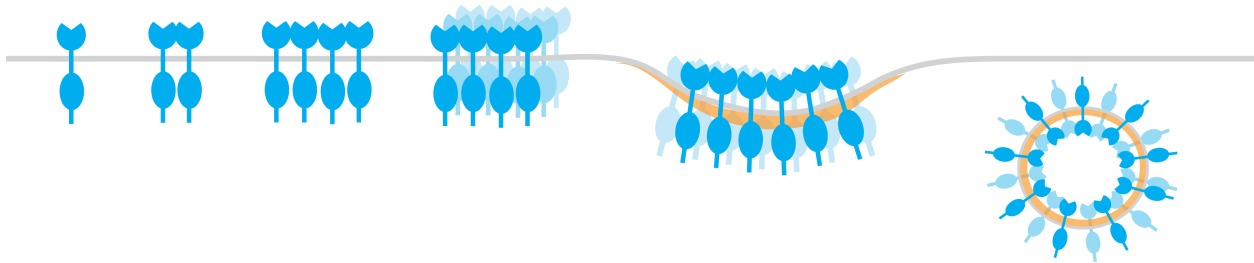
# FIGURES

**Figure 4.1:** Canonical mechanism for activation of EGFR and its downstream pathways.



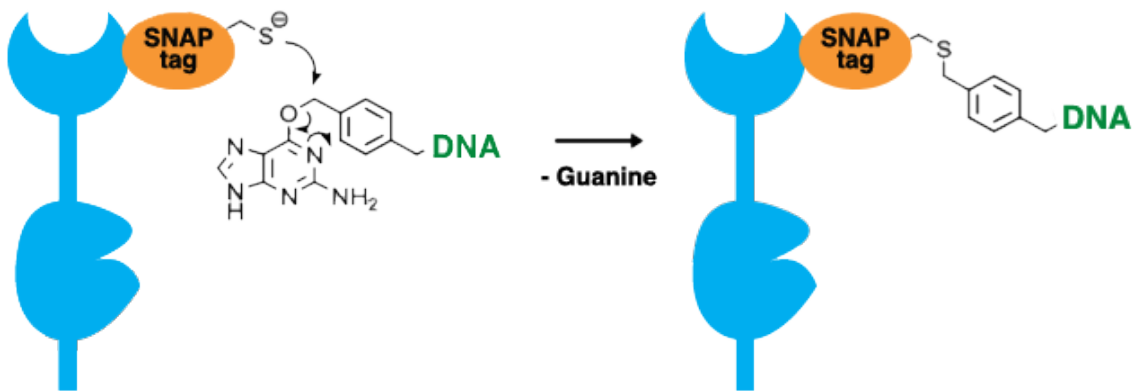
In a very simplistic explanation of how EGFR signals: monomers of EGFR are inactive, but the binding of its ligand (EGF) causes structural changes that open up a dimerization domain on EGFR, a dimer of EGFR forms, and the intracellular kinase domains of these dimers are able to phosphorylate each other tails. Intracellular scaffolding and signaling proteins then recognize these phosphorylated tails and lead to downstream signals – including the MapK/ERK cascade and the PI3K/AKT cascade, which are associated with cell proliferation and migration.

**Figure 4.2:** Changes in spatial organization of EGFR upon addition of ligand EGF.



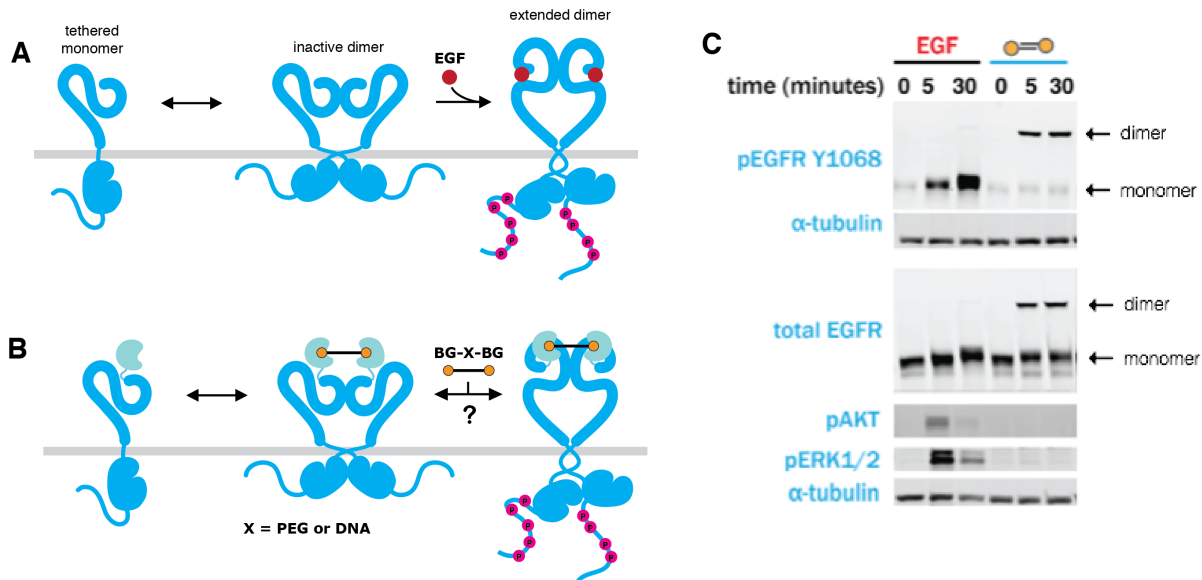
When EGFR is stimulated with EGF, it undergoes very fast changes in spatial organization. Monomers, dimers, higher order multimers, large clathrin-coated pits filled with EGFR and pinched-off endosomes with EGFR can all be observed within 15 minutes of EGF addition.

**Figure 4.3:** Mechanism of covalent conjugation of Snap-tagged receptors to benzylguanine-modified DNA.



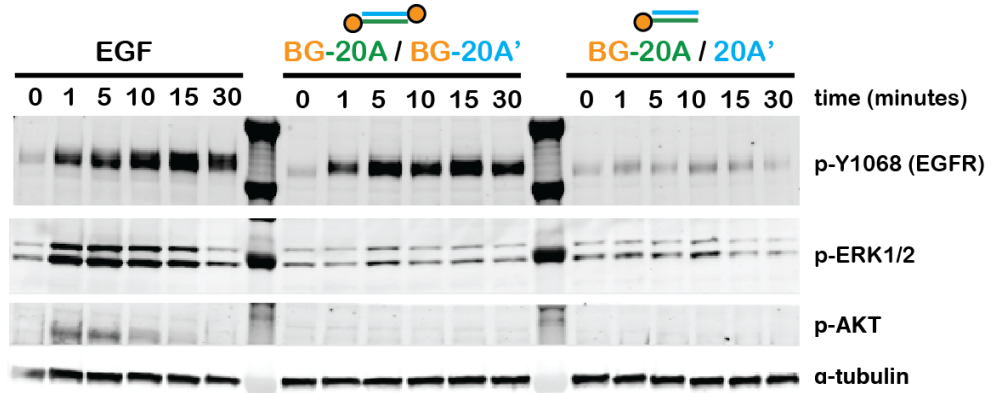
Chemical mechanism for Snap-tag conjugation is shown above. This reaction occurs at all biological temperatures and pH, and goes to completion within several minutes.

**Figure 4.4:** Ligand-independent dimerization of EGFR with DNA-scaffolds



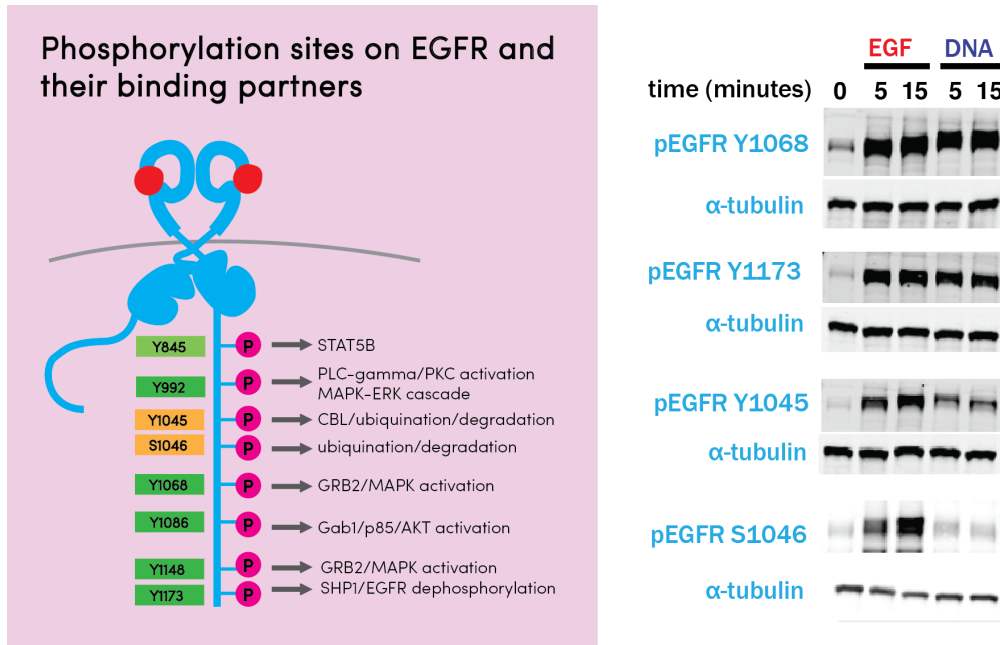
(A) EGFR exists in equilibrium between a tethered monomer and an inactive dimer that has been observed with structural and diffusion studies. Upon EGF addition, an extended active EGFR dimer is formed, and one of the kinases of the dimer serves as an allosteric activator of the other, resulting in trans-phosphorylation of EGFR's C-terminal tail tyrosines. (B) Our scheme stabilizes EGFR dimers with Benzyl-guanine modified DNA dimers in order to see if we are able to achieve ligand independent phosphorylation. (C) Western blot showing that EGF induces phosphorylation of EGFR and activation of AKT and ERK compared to background. Upon DNA stimulation, the dimer of EGFR is phosphorylated but surprisingly, no phospho-AKT or phospho-ERK is observed.

**Figure 4.5:** Phosphorylation of EGFR is dependent of dimerization



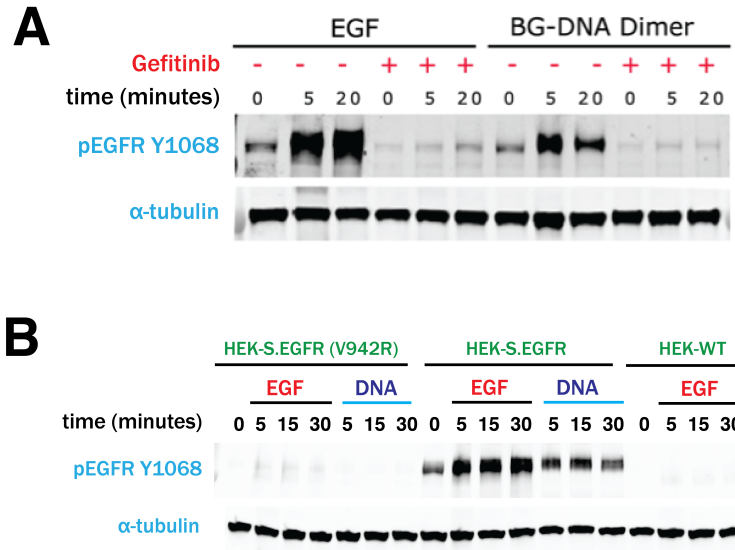
Samples were boiled prior to running to gel so that the DNA dimer is denatured and can be directly compared to the EGF-treated conditions. Levels of phospho-EGFR are comparable in both EGF-stimulated and DNA-dimerized conditions. As a control, we treated cells with DNA that only had a single benzylguanine-group such that it would not form dimers, and no phospho-EGFR was observed.

**Figure 4.6:** Direct comparison of phosphorylation levels of EGFR for EGF and DNA stimulated cells.



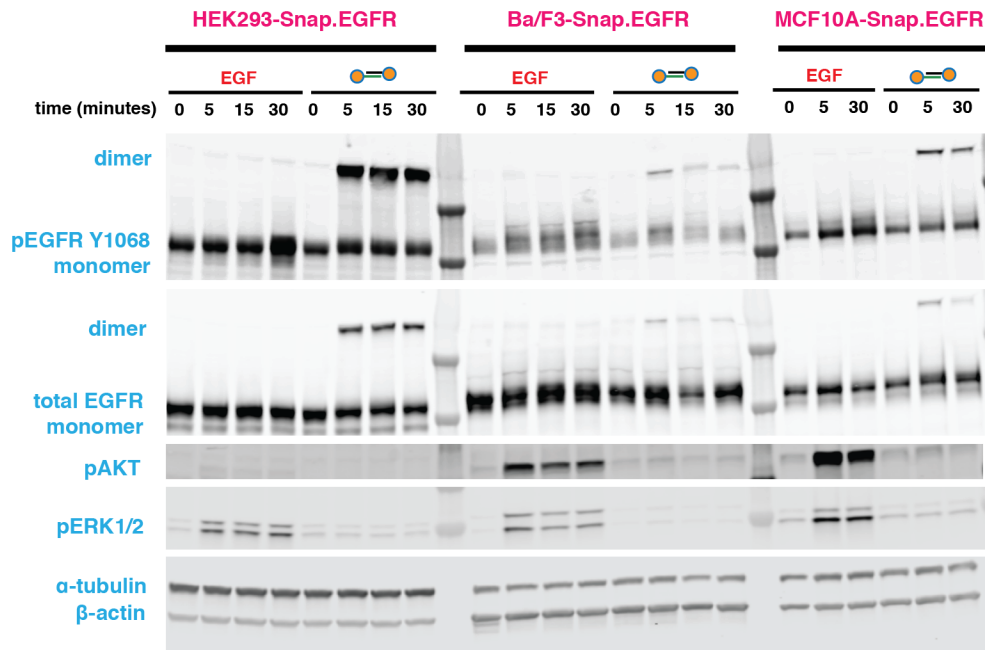
EGFR has several tyrosines on its C-terminal tail that can be phosphorylated and we wanted to find out if our DNA-induced dimer was getting fully phosphorylated. The ability to boil the DNA-dimerized samples allows us to directly compare EGFR phosphorylation levels. We did not see significant differences in phosphorylation levels over many replicates. Phospho-S1046 is a result of a different kinase and is a result of feedback from activation of downstream signaling pathways.

**Figure 4.7:** DNA-stabilized EGFR auto-phosphorylates via a similar mechanism to EGF-stimulated EGFR



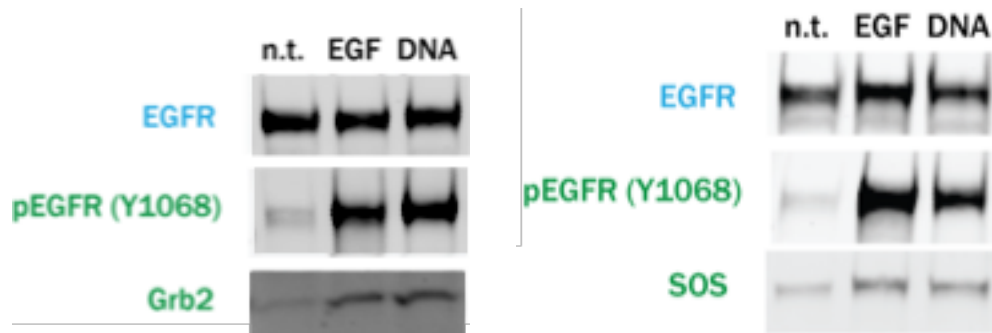
(A) When the EGFR kinase is inhibited with gefitinib, the DNA-induced dimers are not auto-phosphorylated. (B) DNA-induced dimerization of an EGFR mutant that cannot form the asymmetric allosteric-activating dimer results in no autophosphorylation. Total levels of Snap.EGFR and Snap.EGFR(V942R) are similar for the two cell lines. These results taken together suggest that the DNA-induced dimers are auto-phosphorylating via a similar mechanism to EGF-activated dimers.

**Figure 4.8:** DNA-dimerization of EGFR on multiple cell lines



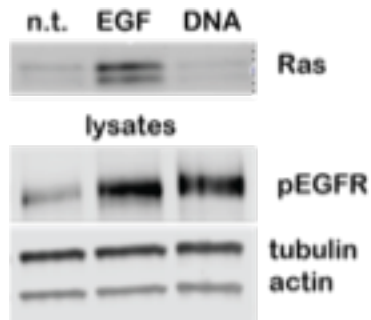
Similar results were observed for many cell lines, although HEK293 cells do show the starkest affect. Ba/F3 cells have express none of the EGFR family members and are a murine suspension cell line. MCF10A lines express low levels of wild-type EGFR and are an epithelial breast cancer cell line. Not shown are H1299 cells, a lung carcinoma cell line, which also showed a similar result.

**Figure 4.9:** Co-immunoprecipitation of adaptor proteins with an EGFR



Using a primary antibody for EGFR and doing a co-IP assay showed that the DNA-induced dimers are capable of recruiting signaling adaptor proteins, Grb2 and SOS.

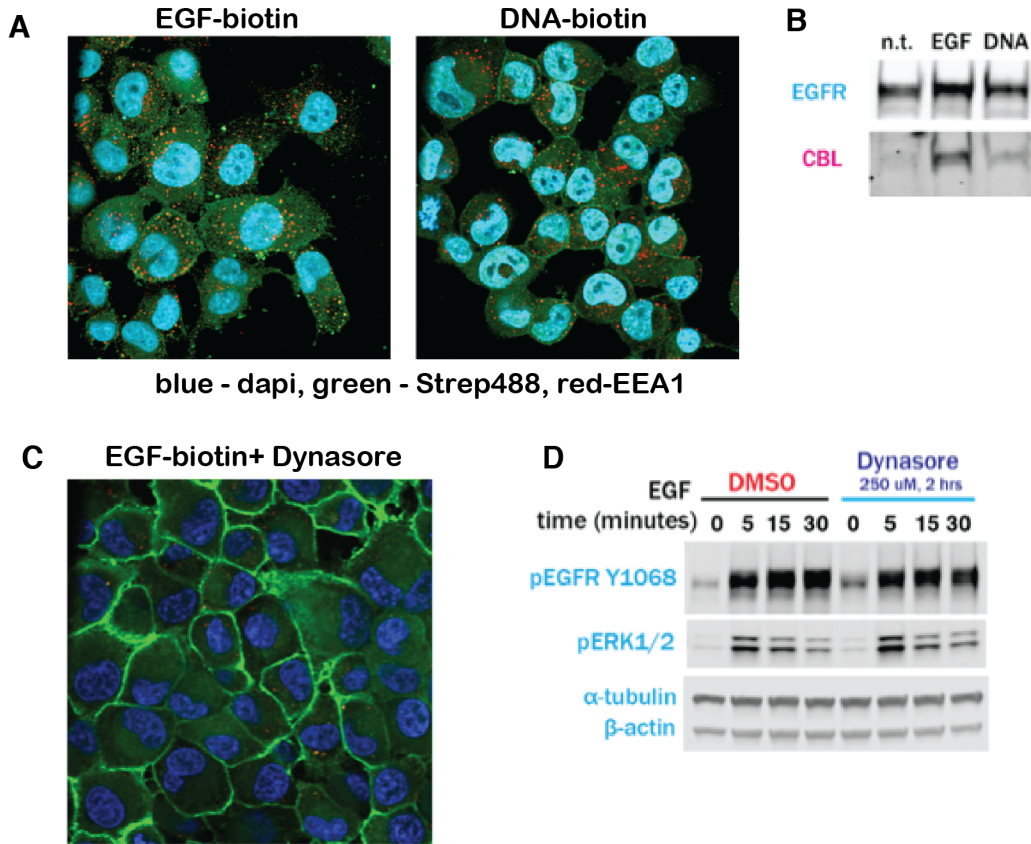
**Figure 4.10:** Activated-RAS-GTP pull-down assays



Levels of activated Ras were measured by incubating lysates with resin that binds to Ras-GTP, eluting, and then blotting with a total Ras antibody. DNA-induced dimers do not activate Ras.

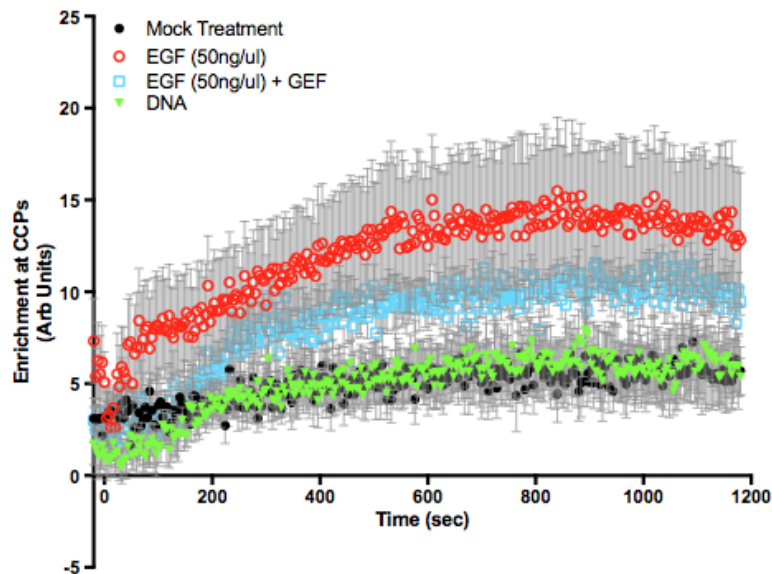


**Figure 4.11:** DNA-dimerized EGFR does not undergo endocytosis



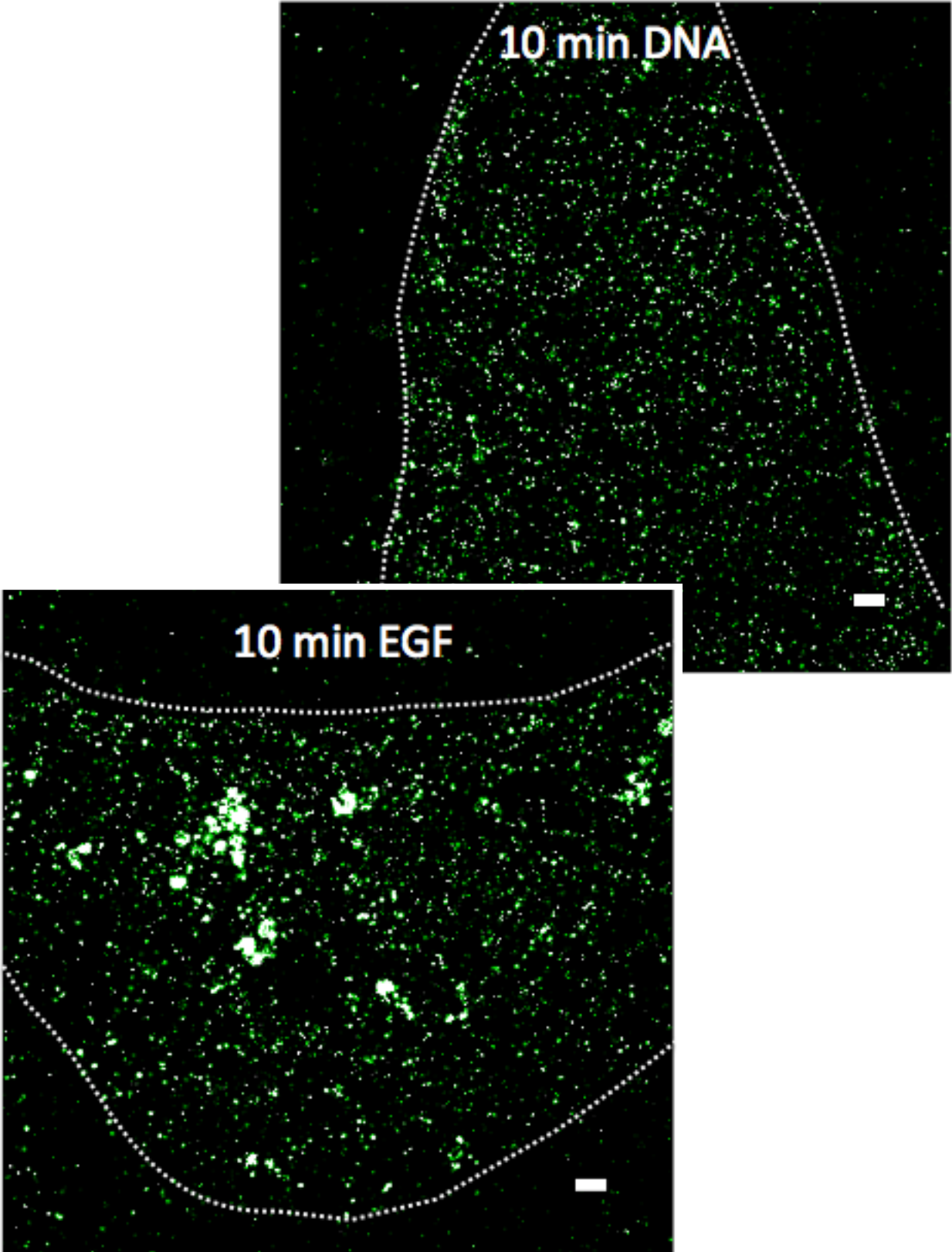
(A) Confocal images of cells stained with streptavidin-488 and an early endosomal marker after treatment with EGF-biotin and DNA-biotin show that DNA-biotin is not endocytosed like EGF-biotin is after 15 minutes. (B) Co-IP with a total-EGFR antibody shows that the DNA condition does not recruit CBL, which is involved in endocytosis (C) Treatment with dynamin inhibitor Dynasore prevents clathrin-coated pits from getting pinched off into endosomes. (D) Treatment with Dynasore does not impact EGF-induced signaling, suggesting that endocytosis is not critical to MAPK activation.

**Figure 4.12:** DNA-dimerized EGFR does not get transported into clathrin-coated pits.



This is the summary of results a set up TIRF Microscopy experiments and analyses from Kelsie Eichel (von Zastrow lab, UCSF). For each condition, 10 cells were imaged across 3 days of imaging and analyzed for the colocalization of Snap.EGFR.GFP with mCherry-clathrin heavy chain upon stimulation with EGF or DNA. The DNA does not induce EGFR transportation to clathrin-coated pits. In contrast, EGF transports EGFR to clathrin-coated pits, even upon treatment of the cells with Gefitinib, which inhibits the kinase domain of EGFR. This suggests that EGF may cause conformational changes that promote association with clathrin coated pits independent of autophosphorylation.

**Figure 4.13:** Super-resolution images of EGFR on cells



Stochastic Optical Reconstruction Microscopy (STORM) Images taken by Bettina van Lengerich (Jura lab, UCSF) of Snap.EGFR.mEOS molecules on cells after EGF or DNA treatment show that while EGF induces the nanoscale clustering of EGFR, DNA-induced dimerization does not.

## REFERENCES

- (1) Lemmon, M.A. & Schlessinger, J. Regulation of signal transduction and signal diversity by receptor oligomerization. *Trends Biochem. Sci.* 19, 459-463 (1994).
- (2) Wieduwilt, M.J. & Moasser, M.M. The epidermal growth factor receptor family: Biology driving targeted therapeutics. *Cell. Mol. Life Sci.* 65, 1566-1584 (2008).
- (3) Yarden, Y. & Sliwkowski, M. Untangling the ErbB signalling network. *Nat Rev Mol Cell Biol* 2, 127-137 (2001).
- (4) Zhou, M. et al. Real-time measurements of kinetics of EGF binding to soluble EGF receptor monomers and dimers support the dimerization model for receptor activation. (1993).
- (5) Chung, I. et al. Spatial control of EGF receptor activation by reversible dimerization on living cells. *Nature* 464, 783-787 (2010).
- (6) Clayton, A.H.A., Orchard, S.G., Nice, E.C., Posner, R.G. & Burgess, A.W. Predominance of activated EGFR higher-order oligomers on the cell surface. *Growth Factors* 26, 316-324 (2008).
- (7) Huang, B., Wang, W., Bates, M. & Zhuang, X. Three-Dimensional Super-Resolution Imaging by Stochastic Optical Reconstruction Microscopy. *Science* 319, 810-813 (2008).
- (8) Hofman, E.G. et al. EGF induces coalescence of different lipid rafts. *J Cell Sci* 121, 2519-2528 (2008).
- (9) Chu, C., Everiss, K. & Wikstrand, C. Receptor dimerization is not a factor in the signalling activity of a transforming variant epidermal growth factor receptor (EGFRvIII). *Biochemical ...* (1997).
- (10) Pedersen, M.W. et al. Analysis of the epidermal growth factor receptor specific transcriptome: Effect of receptor expression level and an activating mutation. *J Cell Biochem* 96, 412-427 (2005).
- (11) Verveer, P., Wouters, F. & Reynolds, A. Quantitative Imaging of Lateral ErbB1 Receptor Signal Propagation in the Plasma Membrane. *Science* (2000).
- (12) Sorkin, A. & Goh, L.K. Endocytosis and intracellular trafficking of ErbBs. *Exp Cell Res* 315, 683-696 (2009).
- (13) Wang, Q., Villeneuve, G. & Wang, Z. Control of epidermal growth factor receptor endocytosis by receptor dimerization, rather than receptor kinase activation. *EMBO Rep.* 6, 942-948 (2005).
- (14) Wells, J. Structural and functional basis for hormone binding and receptor oligomerization. *Current opinion in cell biology* 6, 163-173 (1994).
- (15) De Meyts, P. The structural basis of insulin and insulin-like growth factor-I receptor binding and negative co-operativity, and its relevance to mitogenic versus metabolic signalling. *Diabetologia* 37, 135-148 (1994).

## **CHAPTER 5**

# Efficient targeting of fatty-acid modified oligonucleotides to live cell membranes through stepwise assembly

Source: The following chapter was published in part from: Weber RW, Liang SI, Selden NS, Desai TA, Gartner ZJ. Efficient Targeting of Fatty-Acid Modified Oligonucleotides to Live Cell Membranes through Stepwise Assembly. *Biomacromolecules*, 15(12), 4621-5, 2014.

Contributions: I initiated the project and performed many of the proof-of-concept experiments that got the project started. As Robert Weber took over the project, I aided in making the figures for the paper, assisting in the longer and complicated experiments, and lending intellectual input. Robert Weber, Zev Gartner, and I wrote the manuscript with editorial input from all authors. Zev Gartner supervised the project.

## BACKGROUND

This was the first project that I worked on in the Gartner lab as a short 10-week project that I focused on while I was rotating through the lab. The idea for the project was developed by Zev Gartner, Nick Selden, and myself, and during this 10-week project, I synthesized many of the molecules and performed proof-of-concept experiments that showed that this project had enough promise to be pursued further. At the time, Robert Weber became interested in taking over the project, made significant intellectual contributions to improving our scheme, and did the majority of the work to finish this project and manuscript. I have included portions of the manuscript below where I feel that I contributed intellectually and those aspects mostly pertain to designing the schemes and to various aspects of characterizing the technique. Many of the applications that are included in that publication are not included in the thesis, but are very interesting and easy to find online as the paper is open-source.

Lipid-modified oligonucleotides<sup>1-3</sup> facilitate uptake of siRNA,<sup>4</sup> target DNA nanostructures to lipid bilayers,<sup>5</sup> program assembly of 3D microtissues,<sup>6,7</sup> enable preparation of live single cell microarrays,<sup>8-10</sup> and function as vaccine adjuvants and immunotherapeutics.<sup>11,12</sup> These uses are predicated on rapid, efficient, and stable partitioning of these amphiphilic molecules from solution into live cell membranes.<sup>13</sup> We recently reported an approach for incorporating dialkylglycerol modified oligonucleotides (DAG) into cell membranes.<sup>10</sup> DAG is useful for targeting DNA to the membranes of most cell lines, but suffers when targeting primary or embryonic stem cells (ESCs). Moreover, DAG and other lipid-modified oligonucleotides slowly leave the cell membrane and establish an equilibrium

with the surrounding medium.<sup>14,15</sup> This loss by re-equilibration limits the ultimate efficiency of incorporation into the bilayer over time.

To improve the concentration of lipid-anchored oligonucleotides in cell membranes, we reasoned that increasing the dialkyl anchor hydrophobicity would increase its thermodynamic stability when inserted into cell membranes.<sup>9,14-16</sup> Indeed, previous studies demonstrated that longer lipids are more stable than shorter lipids when reconstituted into synthetic lipid bilayers.<sup>3</sup> However, we found that DAG incorporation into live cell membranes (as opposed to synthetic systems) was exquisitely sensitive to alkyl chain length. The addition of even two methylenes completely inhibited partitioning into cell membranes.<sup>9</sup> We hypothesized that this was due to a competing self-aggregation reaction<sup>17,18</sup> and thus sought an alternative means of introducing greater hydrophobicity to the lipid anchors without aggregation.

## RESULTS

Previous reports show that complementary cholesterol-bearing oligonucleotides can be stably targeted to liposomes and supported lipid bilayers via hybridization.<sup>19</sup> We envisioned further increasing the hydrophobicity of the membrane anchors to further stabilize duplexes in live cells, rather than artificial lipid bilayers. To prevent aggregation of these more hydrophobic molecules, however, the two strands would need to be added sequentially to cells, rather than as a prehybridized duplex. Under conditions of stepwise addition, a first



Anchor strand (Anch) partitions into the lipid bilayer but remains in rapid equilibrium with the medium. A second, co-Anchor (cA) strand is subsequently added and also establishes rapid equilibrium between the lipid bilayer and the medium. However, upon encountering the first strand through diffusion in the phospholipid bilayer, the two strands hybridize, increase the total hydrophobicity of the now doubly anchored duplex and, thus, slowing their exchange with the medium (Figure 5.1).

To explore this strategy, we used fatty acid amides (FA) as more synthetically tractable membrane anchors than previously reported phospholipids or cholesterol.<sup>19</sup> Fatty acids are widely commercially available and do not require chemical modification before coupling. Additionally, the conjugation reaction to DNA is not highly water sensitive and requires only one reverse phase purification step after coupling. Consistent with past studies,<sup>13,20</sup> a single FA anchor does not stably label cell membranes when compared to DAG or doubly cholesterol-anchored DNA. For example, a 100 base Anch strand linked to stearic acid (C18) via a 5' amide (5'- Anch100-C18) did not yield significant DNA incorporation after incubation with cells and washing (Figure 5.2, column 3). However, addition of a second, 20 base complementary coanchor (cA-) strand linked to palmitic acid (C16) via a 3' amide (3'-cA20-C16) dramatically increased cell labeling to near that of the DAG and doubly anchored cholesterol (Figure 5.2, column 6). No increase was seen upon addition of a noncomplementary 3'-cA20-C16 strand (Figure 5.2, Column 5), indicating that at least two FA anchors, linked noncovalently through Watson-Crick base pairing in the "lock" region are necessary for stable incorporation.

We found that the number of base pairs in the lock region correlated with initial labeling and retention of oligonucleotides over time, both at 0 and 37 °C. This effect

saturated between 15 and 20 bases. Labeling was dose-dependent and occurred without altering cell viability over the examined range of 0.5 to 5  $\mu$ M. Encouragingly, even these unoptimized molecules were capable of programming cell–cell and cell–surface adhesion of model cell lines with results comparable to DAG when incorporating 60 base polythymine spacers.

These initial findings suggested we could achieve additional improvements in cell membrane incorporation by increasing the length and thus hydrophobicity of FA anchors. We therefore synthesized a series of 5'-Anch100 strands conjugated to saturated FAs between 16 and 24 carbons in length. These Anch strands behaved as predicted when added stepwise to cells in concert with 3'-cA20-C16. Increased lipophilicity of FAs enhanced the labeling efficiency and showed substantial improvement over both DAG and cholesterol linked oligonucleotides (Figure 5.3A). Anch strands with enhanced hydrophobicity also demonstrated improved retention over time at physiological temperature. Unlike the 5'-Anch100-C18 and 3'-cA20-C16 combination, stepwise addition of more hydrophobic Anch and cA strands was essential for preventing competing aggregation reactions (Figure 5.3B). Prehybridizing 5'-Anch100-C24 and 3'-cA20-C16 strands led to dramatically reduced cell membrane incorporation compared to prehybridized 5'-Anch100-C18 and 3'-cA20-C16 strands (Figure 5.3B). Together, these results support the notion that splitting the hydrophobicity of dual-anchored species across two complementary oligonucleotides added stepwise to cells prevents aggregation and improves labeling.

These data suggested we could achieve further increases in cell labeling and stability by increasing the hydrophobicity of the cA strand in addition to the Anch strand. Surprisingly,

increasing the hydrophobicity of the cA strand anchors did not yield additional gains in cell labeling. For instance, stepwise addition of 5'-Anch100-C24 and 3'-cA20-C24, which maximizes hydrophobicity for both strands, actually decreased DNA incorporation when compared to 5'-Anch100-C24/3'-cA20-C16. We investigated this effect by assaying a panel of molecules in which the fatty acid on the Anch and cA strands was varied systematically and independently. We found that increasing hydrophobicity specifically on the coanchor strand decreased labeling (Figure 5.3C). Indeed, 5'-Anch100-C24/3'-cA20-C16 inserted far more efficiently into cell membranes than 5'-Anch100-C16/3'-cA20-C24 despite containing identical number of phosphodiester bonds and methylene groups.

To explain this trend, we hypothesized that the ratio of anchor hydrophobicity to oligonucleotide length (and thus charge) determines the extent of aggregation. If this were the case, short oligonucleotides would be more prone to aggregation than equivalently modified longer oligonucleotides. To test this notion, we used dynamic light scattering (DLS) to examine the relationship between FA anchor length, oligonucleotide length, and relative aggregation. Both scattered light intensity (Figure 5.4A) and particle size correlated with the length of the FA conjugated to the cA strand. In contrast, very little light scattering was observed for any FA conjugated to the 100 base anchor strand (Figure 5.4B). These results suggested that adding additional bases to the cA strand, increasing its net size and charge, would destabilize aggregates through Coulombic or steric repulsion while simultaneously allowing for increased hydrophobicity of its FA anchor. We therefore synthesized a series of 3'-cA20-C24 strands incorporating an additional 10, 20, or 30 bases at the 5' end. Consistent with our expectations, DLS revealed an inverse relationship between the number of bases and

aggregation (Figure 5.4B). Moreover, the best of these molecules, 3'-cA50-C24, increased cell labeling in combination with 5'-Anch100-C24 to nearly 7-fold of DAG (Figure 5.4C). This combination of molecules also showed a dramatic increase in lifetime at the cell surface compared to DAG (Figure 5.4D). We calculated that the initial rate of decay of these fully optimized strands from the cell surface was nearly 100-fold lower than DAG.

Stepwise assembly of membrane-anchored oligonucleotides is a modular strategy for targeting DNA to cell membranes with improved efficiency and stability. Insertion of oligonucleotide duplexes into membranes occurs via two FA-anchors with higher net lipophilicity compared to previously reported anchors. Competing self-aggregation is prevented by separating the dual anchors between two molecules that are added sequentially to cells, as well as by balancing the ratio of hydrophobicity to oligonucleotide length. This strategy facilitates new applications such as DNA-mediated adhesion in primary cells, murine ESCs, and pancreatic  $\beta$  cells, cell types that show little to no labeling with DAG (summarized in Figure 5.5). An additional benefit of these molecules is their streamlined synthesis compared to previous methods. We anticipate that the structure/function relationships defined here will prove useful in other applications utilizing lipid-modified oligonucleotides or amphiphiles including vaccine adjuvants, siRNA delivery, and structural DNA nanotechnology.

## MATERIALS AND METHODS

### SYNTHESIS OF LIPID-MODIFIED OLIGONUCLEOTIDES

Hexadecanoic (Palmitic) acid, octadecanoic (Stearic) acid, icosanoic (Arachidic) acid, docosanoic (Behenic) acid, tetracosanoic (Lignoceric) acid, N,N-diisopropylethylamine (DIPEA), N,N-diisopropylchlorophosphoramidite (DIPC), N,N-dimethylformamide (DMF), methylamine, ammonium hydroxide, and piperidine were obtained from Sigma-Aldrich. HPLC grade acetonitrile, triethylamine, acetic acid, and dichloromethane (DCM) were obtained from Fisher Scientific. Monomethoxytritylamino)hexyl-(2-cyanoethyl)-(N,N-diisopropyl)- phosphoramidite (amine phosphoramidite), standard phosphoramidites, and DNA synthesis reagents were obtained from Azco Biotech. Controlled pore glass (CPG) support, 1-O-dimethoxytrityl-hexyl-disulfide, 1'-[(2-cyanoethyl)-(N,N-diisopropyl)]-phosphoramidite, 10-O-[1-propyl-3-N-carbamoylcholesteryl]-triethylene glycol-1-[(2-cyanoethyl)-(N,N-diisopropyl)]-phosphoramidite (5'-cholesterol-TEG phosphoramidite), (1-dimethoxytrityloxy-3-O-(N-cholesteryl-3-aminopropyl)-triethylene glycol-glycerol-2-O-succinoyl-long chain alkylamino-CPG (3'-cholesterol-TEG CPG), and 2-dimethoxytrityloxymethyl-6-fluorenylmethoxycarbonylamino-hexane-1-succinoyl-long chain alkylamino-CPG (3'-amino-modifier C7 CPG), and synthesis columns were obtained from Glen Research. All materials were used as received from manufacturer. Oligonucleotides were synthesized on an Applied Biosystems Expedite 8909 DNA synthesizer. Amino and cholesterol modified DNA strands were synthesized using amine and cholesterol phosphoramidites (100 mM), respectively, using a custom 15 min coupling protocol. For the polythymine regions of the anchor strands (Anch), the capping step was omitted in order to maximize yield. After synthesis of 5' amino-modified DNA, the MMT protecting group was deprotected manually on the synthesizer by priming alternately with deblock and dry acetonitrile three times and watching for yellow elution. To ensure complete

deprotection of the MMT group, the 5' solid supports were also resuspended in a solution of 20% acetic acid/ 80% water<sup>1</sup> shaking for 1 h at room temperature. The solid support was subsequently washed repeatedly with DMF, DCM, and acetonitrile with acetonitrile as the final wash and then dried with a speedvac system. For the 3' amino-modified CPG, a solution of 20% piperidine in dimethylformamide was prepared and used to deprotect the CPG support for 10 min at room temperature, followed by DCM and DMF washes with DCM as the final wash. This procedure was repeated twice more to ensure complete deprotection of the Fmoc protecting group prior to coupling to the fatty acid. Fatty acid conjugated oligonucleotides were synthesized by coupling the carboxylic acid moiety of the fatty acid to amino modified oligonucleotides with a 3' or 5' free amine while on the solid support. The solid support was transferred to an eppendorf tube and resuspended in a solution of dichloromethane containing 200 mM fatty acid, 400 mM DIPEA, and 200 mM DIPC. The eppendorf tubes were sealed with parafilm, crowned with a cap lock, and shaken overnight at room temperature. The next morning, they were washed with DCM and DMF repeatedly and then cleaved off the solid support. Oligonucleotides were cleaved from solid support with a 1:1 mixture of ammonium hydroxide/40% methylamine (AMA) for 1 h at 65 °C with a cap lock followed by evaporation of AMA with a speedvac system. They were then purified by HPLC as described in previous chapters. Stocks of 250 μM were prepared and from them aliquots of 50 μM were prepared for day-to-day use in order to minimize repeated freeze–thaw cycles.

## DNA LABELING OF CELLS AND QUANTIFICATION OF CELL SURFACE

### OLIGONUCLEOTIDES

For experiments, unless otherwise noted, Jurkat cells were used. Cells were pelleted at 1000 g resuspended in calcium and magnesium-free PBS (UCSF Cell Culture Facility) three times, with a final resuspension volume of 48  $\mu\text{L}$  of PBS per  $10^6$  cells. Resuspended cells were labeled with single-stranded DNA by the addition of 1  $\mu\text{L}$  of a 50  $\mu\text{M}$  solution of the anchor strand in water. Cells were gently agitated by gentle vortexing for 5 min at room temperature. Subsequently, 1  $\mu\text{L}$  of a 50  $\mu\text{M}$  solution of the coanchor strand in water was added, bringing both strands to a final concentration of 1  $\mu\text{M}$ . Cells were again gently agitated by slow vortexing for 5 min at room temperature. The cells were then pelleted and resuspended three times in ice-cold PBS to remove unbound or excess oligonucleotides. To quantify the extent of cell surface labeling, cells were incubated with 100  $\mu\text{L}$  of a 20mer complementary 6-FAM modified oligonucleotide (1  $\mu\text{g}/\text{mL}$ , Operon), which annealed to the most distal portion of the anchor strand. The strand was incubated for 30–45 min at 4  $^{\circ}\text{C}$ , protected from light. Cells were pelleted and resuspended one time in ice cold PBS before pelleting and resuspending in 100  $\mu\text{L}$  per  $10^6$  cells of LIVE/DEAD Fixable Cell Stain (Invitrogen, used per manufacturers instructions) for 15 min at 4  $^{\circ}\text{C}$  protected from light. Cells were washed one last time with ice cold PBS before flow cytometry analysis. Flow cytometry was performed on a FACSCalibur (BD Biosciences, UCSF Laboratory for Cell Analysis) and the data was analyzed using FlowJo software package (Tree Star, Inc.). For stability time course experiments, cells were incubated at 37  $^{\circ}\text{C}$  for the designated amount of time in the presence of serum-free RPMI 1640 before incubating with the fluorescent, complementary oligonucleotide. For the preannealing experiment, a 1  $\mu\text{M}$  solution of C18/C16 and a 0.3  $\mu\text{M}$  solution of C22/C16 fatty acid modified strands in room temperature PBS was prepared and gently agitated for 10 min at room

temperature. This solution was used to resuspend the cell pellet after the final wash from media and gently agitated for an additional 10 min at room temperature. This was compared to normal labeling using these same strands at these same concentrations. All reported values are the average of three or more independent measurements, with error bars indicating standard deviations.

### DYNAMIC LIGHT SCATTERING

PBS CMF (UCSF CCF) was filtered by a 0.2  $\mu\text{m}$  vacuum filter. Stock solutions of 250  $\mu\text{M}$  ss-DNA strands were diluted to 1  $\mu\text{M}$  with this filtered PBS prior to transfer to cuvette for measuring by DLS on a Wyatt Technology DynaPro Protein Solutions utilizing the DYNAMICS software package ver 6.10.1.2. Particle size was determined by cumulants analysis. All samples were prepared separately and measured in triplicate.

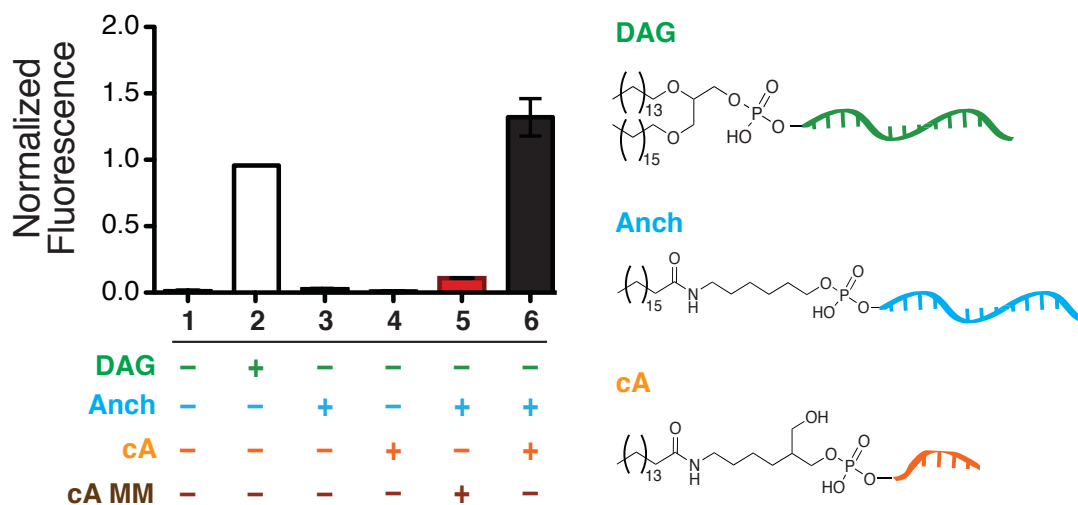


## ACKNOWLEDGMENTS

This work was supported by grants from the Department of Defense Breast Cancer Research Program (W81XWH-10-1- 1023 and W81XWH-13-1-0221 to Z.J.G.); NICHD (DP2 HD080351-01 to Z.J.G.); The Sidney Kimmel Foundation; and the UCSF Program in Breakthrough Biomedical Research. Z.J.G. is an investigator in the UCSF Center for Systems and Synthetic Biology (NIGMS Systems Biology Center Grant P50 GM081879). This work is partially supported by the Achievement Rewards for College Scientists Foundation Fellowship and the Genentech Foundation Fellowship to S.I.L. The authors thank Michael Todhunter, Noel Jee, Amanda Paulson, Matthew Thomson, Jade McPherson, Michael Broeker, and Allison Doak for technical assistance.

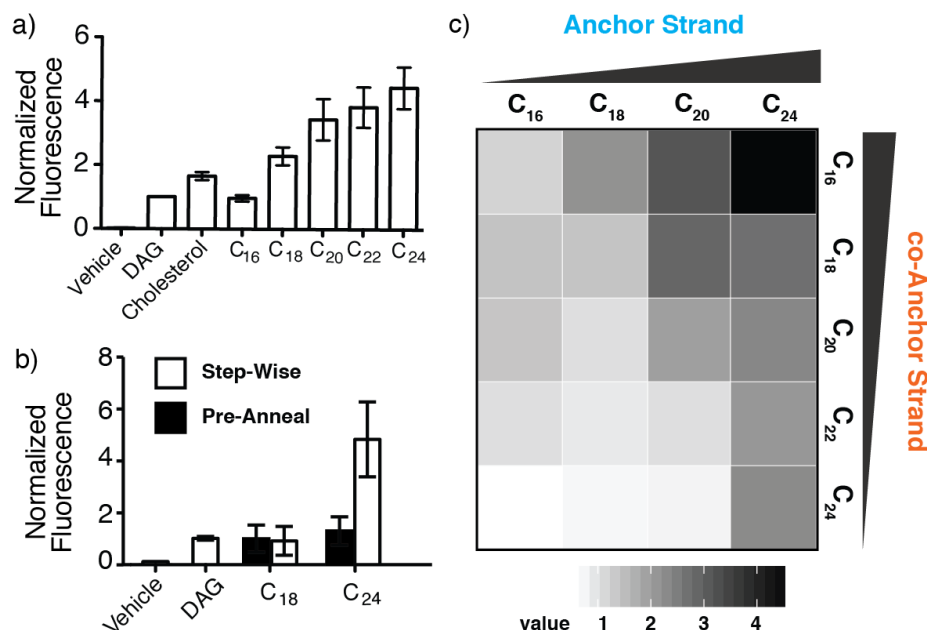


**Figure 5.2:** Anchor (Anch) and complementary Co-Anchor (cA) strands together enhance ssDNA targeting and retention in cell membranes.



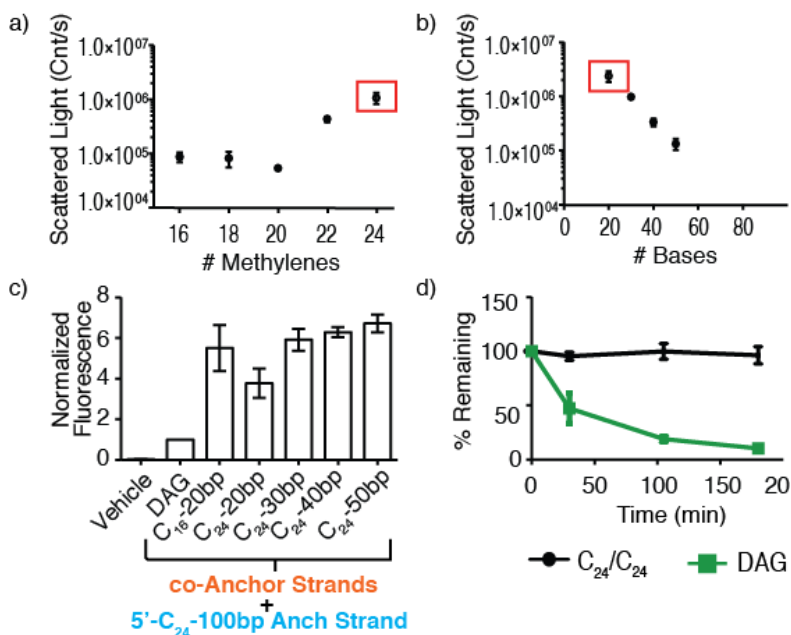
Fluorescence was measured with flow cytometry and normalized to a C<sub>16</sub> DAG-ssDNA control. cA-MM is a 20 base co-anchor strand with a DNA sequence non-complementary to the Anch strand. Error bars are standard deviation of at least 3 independent measurements.

**Figure 5.3:** Lipid hydrophobicity affects cell labeling efficiency of Anch, cA, and pre-hybridized strands.



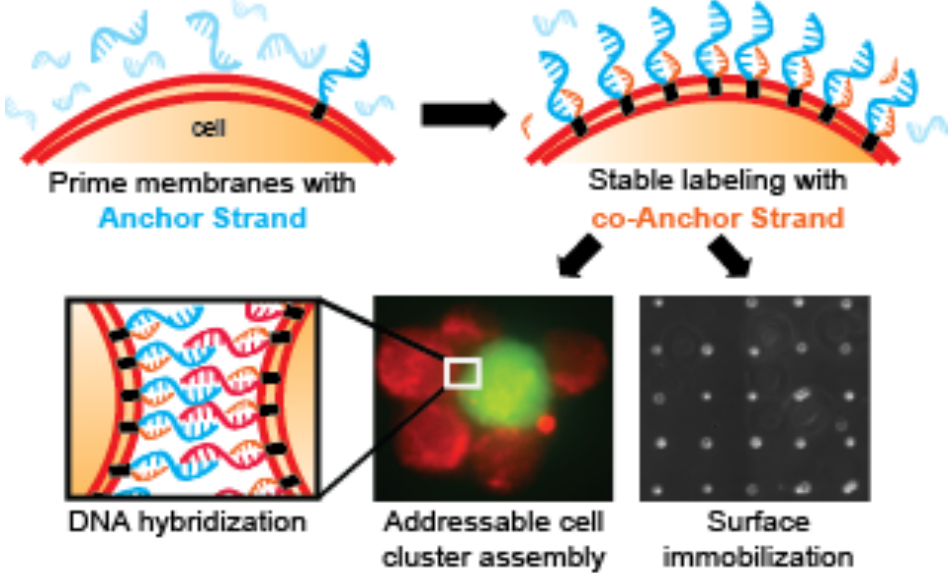
(A) Membrane incorporation compared to DAG for duplexes with 3'-cA<sub>20</sub>-C<sub>16</sub> strands and Anch strands bearing FA anchors of increasing length. (B) Membrane incorporation of 3'-cA<sub>20</sub>-C<sub>16</sub> and different Anch strands when added stepwise (white bars) or after pre-annealing (black bars). (C) Heatmap relating average membrane labeling (n=3) to combinations of Anch and cA strand FA anchor lengths. Error bars are standard deviation of at least 3 independent measurements.

**Figure 5.4:** Modular assembly of protein-bearing DNA multimers and their interactions with live cells



(A) Light scattering from solutions of 3'-cA<sub>20</sub> strands as a function of FA anchor length. b) Light scattering from solutions of 3'-cA-C<sub>24</sub> anchored oligonucleotides as a function of the number of DNA bases. The red box indicates the same strand, 3'-cA<sub>20</sub>-C<sub>24</sub>. c) Membrane incorporation compared to DAG for combinations of 3'-cA-C<sub>24</sub> with increasing numbers of DNA bases. d) Incorporation vs. time for DAG and 5'-Anch<sub>100</sub>-C<sub>24</sub>/3'-cA<sub>50</sub>-C<sub>24</sub>-DNA. Error bars are standard deviation of at least 3 independent measurements.

**Figure 5.4:** Improved preparation of single cell microarrays and 3D microtissues using stepwise assembly of membrane anchored adhesive oligonucleotides.



This strategy can be used for many programmed cell assembly applications, such as single cell microarrays prepared with murine ESCs, assemblies of MCF10A clusters, which can establish polarity after 48 hours in Matrigel, and clustering sorted ESC/ $\beta$  islet cells. These applications used to be more difficult with older DNA-cell labeling techniques which did not result in as many oligonucleotides on the cell surface.

## REFERENCES

- (1) Benkoski, J. J.; Höök, F. *J. Phys. Chem. B* **2005**, *109*, 9773.
- (2) Yoshina-Ishii, C.; Boxer, S. G. *J. Am. Chem. Soc.* **2003**, *125*, 3696.
- (3) Tumpene, J.; Ljungdahl, T.; Wilhelmsson, L. M.; Nordén, B.; Brown, T.; Albinsson, B. *J. Am. Chem. Soc.* **2009**, *131*, 2831.
- (4) Kanasty, R.; Dorkin, J. R.; Vegas, A.; Anderson, D. *Nat Mater* **2013**, *12*, 967.
- (5) Langecker, M.; Arnaut, V.; Martin, T. G.; List, J.; Renner, S.; Mayer, M.; Dietz, H.; Simmel, F. C. *Science* **2012**, *338*, 932.
- (6) Gartner, Z. J.; Bertozzi, C. R. *Proceedings of the National Academy of Sciences* **2009**, *106*, 4606.
- (7) Liu, J. S.; Gartner, Z. J. *Trends in Cell Biology* **2012**, *22*, 683.
- (8) Chandra, R. A.; Douglas, E. S.; Mathies, R. A.; Bertozzi, C. R.; Francis, M. B. *Angew. Chem. Int. Ed.* **2006**, *45*, 896.
- (9) Hsiao, S. C.; Shum, B. J.; Onoe, H.; Douglas, E. S.; Gartner, Z. J.; Mathies, R. A.; Bertozzi, C. R.; Francis, M. B. *Langmuir* **2009**, *25*, 6985.
- (10) Selden, N. S.; Todhunter, M. E.; Jee, N. Y.; Liu, J. S.; Broaders, K. E.; Gartner, Z. J. *J. Am. Chem. Soc.* **2012**, *134*, 765.
- (11) Pfaar, O.; Cazan, D.; Klimek, L.; Larenas-Linnemann, D.; Calderon, M. A. *Current Opinion in Allergy and Clinical Immunology* **2012**, *12*, 648.
- (12) Ishii, K.; Akira, S. *Trends in Immunology* **2006**, *27*, 525.
- (13) Borisenko, G. G.; Zaitseva, M. A.; Chuvilin, A. N.; Pozmogova, G. E. *Nucleic Acids Research* **2008**, *37*, e28.
- (14) Liu, H.; Kwong, B.; Irvine, D. J. *Angew. Chem. Int. Ed.* **2011**, *50*, 7052.
- (15) Palte, M. J.; Raines, R. T. *J. Am. Chem. Soc.* **2012**, *134*, 6218.
- (16) Silvius, J. R.; Leventis, R. *Biochemistry* **1993**, *32*, 13318.
- (17) Patwa, A.; Gissot, A.; Bestel, I.; Barthélémy, P. *Chem. Soc. Rev.* **2011**, *40*, 5844.
- (18) Berti, D. *Current Opinion in Colloid & Interface Science* **2006**, *11*, 74.
- (19) Pfeiffer, I.; Höök, F. *J. Am. Chem. Soc.* **2004**, *126*, 10224.
- (20) Kato, K.; Itoh, C.; Yasukouchi, T.; Nagamune, T. *Biotechnol. Prog.* **2004**, *20*, 897.
- (21) Debnath, J.; Muthuswamy, S. K.; Brugge, J. S. *Methods* **2003**, *30*, 256.
- (22) Liu, J. S.; Farlow, J. T.; Paulson, A. K.; Labarge, M. A.; Gartner, Z. J. *Cell Reports* **2012**, *2*, 1461.
- (23) Goerke, S. M.; Plaha, J.; Hager, S.; Strassburg, S.; Torio-Padron, N.; Stark, G. B.; Finkenzeller, G. *Tissue Engineering Part A* **2012**, *18*, 2395.
- (24) Saleh, F. A.; Whyte, M.; Genever, P. G. *Eur Cell Mater* **2011**, *22*, 242.
- (25) Van Hoof, D.; Mendelsohn, A. D.; Seerke, R.; Desai, T. A.; German, M. S. *Stem Cell Research* **2011**, *6*, 276.

## **Appendices**



## A.1: Fab Genetic Constructs

The uPAR-targeting Fabs 2G10 and 3C6 could be expressed from their original phage-display plasmids that the Craik lab used. I cloned the aldehyde tag on their heavy chains on the C-terminus. I did not include the full plasmid and protein sequences for these constructs as they are not yet published, but please contact the Gartner lab or the Craik lab if you would like more information on them. These Fabs also have a HIS-tag on the heavy chain. I have also cloned versions of the Fab plasmids with the aldehyde tag and the STREP-tag instead of the HIS-tag), as well as version with the Snap-tag and a HIS-tag ((sequence not included below). These versions of the plasmids did not turn out to be useful to this project, but they may be useful to future projects.

## A.2: Synthetic DNA sequences

Sequences that were synthesized in order to make controlled trimers ("Y" sequences), tetramers ("X" sequences), mismatches ("MM" sequences) are in the table below. Scaffolding sequences were typically ordered from other companies and are prefaced with an "s" in front of their numbering. For dimers or different lengths, matching strands were typically denoted as their length, plus either an "A" or an "A'," which was pronounced as "A-prime."

Name	5' to 3' sequence
(ACTG) <sub>5</sub>	ACTGACTGACTGACTGACTG
(CAGT) <sub>5</sub>	CAGTCAGTCAGTCAGTCAGT
HetA, A, or 20A	GTAACGATCCAGCTGTCACT
HetA', A', or 20A'	AGTGACAGCTGGATCGTTAC
Y26A	TCGATCCGCATGACATTGCGCGTAAG
Y26B	CTTACGGCGAATGACCGAATCAGCCT
Y26C	AGGCTGATTCGGTTCATGCGGATCGA
Y36A	ACCACTGGATCCGCATGACATTGCGCCGTAAGCACAC
Y36B	GTGTGCTTACGGCGAATGACCGAATCAGCCTGCTGA
Y36C	TCAGCAGGCTGATTCGGTTCATGCGGATCCAGTGGT
Y46A	ACCACTGGATCCGCATGAGGTAGGACGACATTGCGCCGTAAGCACAC
Y46B	GTGTGCTTACGGCGAATGTCGTACAGCACCGAATCAGCCTGCTGA
Y46C	TCAGCAGGCTGATTCGGTGCTGTCTACCTCATGCGGATCCAGTGGT
DX1A	CTACCGCACCCAGAATG

Name	5' to 3' sequence
DX1B	CGATCCGTGGCTACTG
DX1C	CAGTAGCCTGCTATCTTATGGCGTGGCAAATGAGTCGAGGACGGATCG
DX1D	CATTCTGGACGCCATAAGATAGCACCTCGACTCATTGCTGCGGTAG
15A	GTAACGATCCAGCTG
15A'	CAGCTGGATCGTTAC
40A	GTAACGATCCAGCTGTCACTTCATACGACTCACTCTAGGG
40A'	CCCTAGAGTGAGTCGTATGAGTGACAGCTGGATCGTTAC
60A	GTAACGATCCAGCTGTCACTACTGACTGACTGACTGACTGTCATACGACTCACTCTAGGG
60A'	CCCTAGAGTGAGTCGTATGACAGTCAGTCAGTCAGTCAGTAGTGACAGCTGGATCGTTAC
Y26C-D1	ACTGACTGACTGACTGACTGAGGCTGATTCGGTTCATGCGGATCGA
Y26C-D2	CAGTCAGTCAGTCAGTCAGTAGGCTGATTCGGTTCATGCGGATCGA
Y26C-D5	GTAACGATCCAGCTGTCACTAGGCTGATTCGGTTCATGCGGATCGA
Y26C-D3	ACTGATGGTAATCTGCACCTAGGCTGATTCGGTTCATGCGGATCGA
Y26C-D4	AGGTGCAGATTACCATCAGTAGGCTGATTCGGTTCATGCGGATCGA
Y26C-D6	TCATACTACGCGTAGTATGAGGCTGATTCGGTTCATGCGGATCGA
Y26C-D7	ACTGACTGACTGACTGAGGCTGATTCGGTTCATGCGGATCGA

Name	5' to 3' sequence
Y26C-D8	CAGTCAGTCAGTCAGT <b>AGGCTGATTCGGTTCATGCGGATCGA</b>
Y26C-D9	ACTGACTGACTGACTGACTGACTG <b>AGGCTGATTCGGTTCATGCGGATCGA</b>
Y26C-D10	CAGTCAGTCAGTCAGTCAGTCAGT <b>AGGCTGATTCGGTTCATGCGGATCGA</b>
Y26C-D11	ACTGACTGACTGACTGACTGACTGACTG <b>AGGCTGATTCGGTTCATGCGGATCGA</b>
Y26C-D12	CAGTCAGTCAGTCAGTCAGTCAGTCAGT <b>AGGCTGATTCGGTTCATGCGGATCGA</b>
Y26C-D13	ACTGACTGACTGACTGACTGACTGACTGACTG <b>AGGCTGATTCGGTTCATGCGGATCGA</b>
Y26C-D14	CAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGT <b>AGGCTGATTCGGTTCATGCGGATCGA</b>
0T-2xHetA'	AGTGACAGCTGGATCGTTAC <b>AGTGACAGCTGGATCGTTAC</b>
10T-2xHetA'	AGTGACAGCTGGATCGTTACTTTTTTTTTTT <b>AGTGACAGCTGGATCGTTAC</b>
20T-2xHetA'	AGTGACAGCTGGATCGTTACTTTTTTTTTTTTTTTTTTTTTTT <b>AGTGACAGCTGGATCGTTAC</b>
30T-2xHetA'	AGTGACAGCTGGATCGTTACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT <b>AGTGACAGCTGGATCGTTAC</b>
40T-2xHetA'	AGTGACAGCTGGATCGTTACTTT <b>AGTGACAGCTGGATCGTTAC</b>
30T-2xHetA'	AGTGACAGCTGGATCGTTACTTT <b>AGTGACAGCTGGATCGTTAC</b>
40T-2xHetA'	AGTGACAGCTGGATCGTTACTTT <b>AGTGACAGCTGGATCGTTAC</b>
sSL1	<b>GCTGACGACACACCATTGCGCGTAAG</b>
sSL2	ACTGACTGACTGACTGACTGACTG <b>ACCACGGTAGTGATCATGCGGATCGA</b>

Name	5' to 3' sequence
sSL3	CTTACGGCGAATGTCACTACCGTGGTCAGTCAGTCAGTCAGTCAGT AGGCTGATTCGGTTCATGCGGATCGA
sSL4	CTTACGGCGAATGTCACTACCGTGGTCAGTCAGTCAGTCAGTCAGT AGGCTGATTCGGTGTGTGTCGTCAGC
sSL5	CTTACGGCGAATGTCACTACCGTGGT
sSL6	CAGTCAGTCAGTCAGTCAGTAGGCTGATTCGGTGTGTGTCGTCAGC
sSL7	TCATACGACTCACTCTAGGGGCTGACGACACACCATTGCGCGTAAG
sSL8	CCCTAGAGTGAGTCGTATGACTTACGGCGAATGGCACAGCTGGTCA
sSL9	ACTGATGGTAATCTGCACCTTGACCAGCTGTGCTCATGCGGATCGA
sSL10	AGGTGCAGATTACCATCAGTCTTACGGCGAATGTCACTACCGTGGT
sSL11	TGACCAGCTGTGCGTGTGTCGTCAGCAGGCTGATTCGGTTCATGCG GATCGA
sSL12	TCACTACCGTGGTGACACAGCTGGTCAAGGCTGATTCGGTTCATGCG GATCGA
sSL13	GCTGACGACACACACCACGGTAGTGAAGGCTGATTCGGTTCATGCG GATCGA
X30A	TCCGTCCTAGCAAGGAGTCTGCTACCGGAA
X30B	TTCCGGTAGCAGACTAAAAGGTGGTTGAAT
X30C	ATTCAACCACCTTTTTTTAACTGCAGCAG
X30D	CTGCTGCAGTTAAAACCTTGCTAGGACGGA
X50A	CCTCGAGGGATCCGTCCTAGCAAGGGGCTGCTACCGGAAGCTTAC AGATG
X50B	CATCTGTAAGCTTCCGGTAGCAGCCTGAGCGGTGGTTGAATCACA GATG

Name	5' to 3' sequence
X50C	CATCTGTGAATTCAACCACCGCTCAACTCAACTGCAGTCTAGAACAC ATG
X50D	CATCTGTTCTAGACTGCAGTTGAGTCCTTGCTAGGACGGATCCCTCG AGG
z	ACCACGGTAGTGA
z'	TCACTACCGTGGT
y	GTGTGTCGTCAGC
y'	GCTGACGACACAC
x	TGACCAGCTGTGC
x'	GCACAGCTGGTCA
HetA'-MM5	AGGTTAGATGTATCCTTAC
HetA'-MM9	GTAACAGCTGGATCTCACT
HetA'-MM10a	GTAAGCAGCTGGATCTCACT
HetA'-MM10b	CGTATTAGAAGTATCCATTC
HetA'-MM12	GTAACGAGCTGGATGTCACT
HetA'-MM14	GTAAGTTGCTGGACGTCACT

## A.3: List of Stable Cell Lines

These cell lines were made by infecting cells with either lentivirus (human cell lines) or retrovirus (mouse cell lines). The Viracore department of UC San Francisco made most of our virus after we gave them endotoxin-free plasmid preps. Although these viral plasmids tended to be around 10-15 kilobases large, which is usually thought to be too large to package efficiently into viral particles, we found that we did not have any issue with infecting our cells. For the Snap-tagged receptor expressing cells, we sort out populations of cells expressing low, medium, and high levels of the receptor on the cell surface to use in our assays.

Cell Line	Transduced Gene(s)	Notes
HEK	Snap.EGFR	Sorted for low, med, high expressers
HEK	Snap.EGFR.GFP	Sorted for low, med, high expressers
HEK	Snap.EGFR.mEOS	Sorted for low, med, high expressers
HEK	Snap.EGFR(V942R)	Sorted for low, med, high expressers
HEK	KTR(erk)mCherry	Phospho-ERK reporter construct
3T3	Snap.EGFR	Sorted for low expressers
H1299	Snap.Transmembrane(CD20).GFP	
H1299	Snap.EGFR	Sorted for low, med, high expressers
H1299	Snap.EGFR.GFP	Sorted for low, med, high

Cell Line	Transduced Gene(s)	Notes
		expressers
H1299	KTR(erk)mCherry	Phospho-ERK reporter construct with a red fluorescent protein
H1299	KTR(erk)BFP	Phospho-ERK reporter construct with blue fluorescent protein
H1299	KTR(erk)mCherry Snap.EGFR.GFP	Bulk population is fairly heterogeneous. Daeha Seo (Jun lab) cloned out several lines that have more homogenous levels of the ERK reporter expression and Snap.EGFR.GFP expression
H1299	KTR(erk)mCherry Snap.EGFR.mEOS	Sorted for low, med, high expressers of Snap.EGFR.mEOS
Ba/F3	Snap.EGFR	Selection Marker: Puromycin. Sorted for low, med, high expressers
Ba/F3	EGFR.GFP	Puromycin resistant
MCF10A	Snap.EGFR	Sorted for low, med, high expressers
MCF10A	RTTA TetOn-Her2 Snap.Her3	Selection Markers: Blasticidin, Puromycin Tet-inducible Her2 with constant Snap.Her3 expression
MCF10A	RTTA TetOn-Her2 Her3	Selection Markers: Blasticidin, Puromycin Tet-inducible Her2 with constant Her3 expression
CHO	Snap.EGFR	
CHO	Her3	Selection Marker: Blasticidin



Cell Line	Transduced Gene(s)	Notes
CHO	Her2	Selection Marker: Puromycin
CHO	Her3(ICD)	Selection Marker: Blasticidin Truncated construct: Intracellular domain of Her3 only with Snap.tag on extracellular side
CHO	Her2(ICD)	Selection Marker: Puromycin Truncated construct: Intracellular domain of Her3 only with Snap.tag on extracellular side
CHO	Snap.Her3	Selection Marker: Blasticidin
CHO	Snap.Her2	Selection Marker: Puromycin
CHO	Snap.Her3(ICD)	Selection Marker: Blasticidin Truncated construct: Intracellular domain of Her3 only with Snap.tag on extracellular side
CHO	Snap.Her2(ICD)	Selection Marker: Puromycin Truncated construct: Intracellular domain of Her2 only with Snap.tag on extracellular side
CHO	Snap.Her3 Her2(ICD)	Selection Markers: Blasticidin, Puromycin Truncated construct: Intracellular domain of Her2 only with Snap.tag on extracellular side Full Snap.Her3 construct
CHO	Snap.Her3 Snap.Her2(ICD)	Selection Markers: Blasticidin, Puromycin Truncated construct: Intracellular domain of Her2 only with Snap.tag on extracellular side Full Snap.Her3 construct

## A.4: List of Primers for Cloning Genetic Constructs

I have opted to not include the full sequences and plasmids for many of the receptor tyrosine kinases that I worked with out of respect for my collaborators who gifted me my original plasmids and work on similar constructs, but have not yet published their results. If there is any interest in using these plasmids, the sequences are saved on the Gartner Lab server and will be happily shared upon request. My favorite methods for cloning are usually either QuikChange to put in shorter sequences, or else Enzymatic Inverse PCR (EIPCR) which involves amplifying the entire backbone of the parent plasmid and including unique restriction sites on the primers.

Primer name	Description	Sequence
sil01F	QuikChange to replace entire Myc tag with Ald tag	CATCACGGGGCCGCACTGTGCACG CCGAGCCGTGGGGCCGCATAGACT
sil01R	QuikChange to replace entire Myc tag with Ald tag	AGTCTATGCGGCCCCACGGCTCGG CGTGACAGTGCGGCCCCGTGATG
sil02F	QuikChange insert Ald tag directly after the Myc tag	GCAGAAGAGGATCTAAATCTGTGCA CGCCGAGCCGTGGGGCCGCATAGA CTGTT
sil02R	QuikChange insert Ald tag directly after the Myc tag	AACAGTCTATGCGGCCCCACGGCTC GGCGTGACAGATTTAGATCCTCTT CTGC
sil03F	IPCR Forward primer with Ald tag overhang	CATCATCATCACGGGGCCGCACTGT GCACGCCGAGCCGT
sil03R	IPCR Reverse (no myc tag)	AACAGTCTATGCGGCCCC
sil04R	IPCR Reverse (ald tag before myc tag):	CTCTTCTGAGATGAGTTTTTGTTTC

Primer name	Description	Sequence
sil07F	QuikChange to replace HisTag with Ald tag	TCTTGTGCGGCCGCACTGTGCACGC CGAGCCGTGGGGCCGCAGAACAA
sil07R	QuikChange to replace HisTag with Ald tag	TTGTTCTGCGGCCCCACGGCTCGGC GTGCACAGTGCGGCCGCACAAGA
sil08F	QuikChange to replace HisTag with Flag tag	TCTTGTGCGGCCGCAgattataaagatga tgatgataaaGGGGCCGCAGAACAA
sil08R	QuikChange to replace HisTag with Flag tag	TTGTTCTGCGGCCCCtttatcatcatcatctt tataatcTGCGGCCGCACAAGA
sil09F	QuikChange to replace HisTag with gly-ser-Ald tag	TCTTGTGCGGCCGCAAGGGTCACTGT GCACGCCGAGCCGTGGGGCCGCA GAACAA
sil09R	QuikChange to replace HisTag with gly-ser-Ald tag	TTGTTCTGCGGCCCCACGGCTCGGC GTGCACAGTGACCCTGCGGCCGCA CAAGA
sil10F	Forward NotI insert of Ald tag and Tev cleavage site	GGCCGCAGGGTCACTGTGCACGCC GAGCCGTGGGTTCAGAAAACCTGTAT TTTCAGGGAGG
sil10R	Reverse NotI insert of Ald tag and Tev cleavage site	GGCCCCTCCCTGAAAATACAGGTTT TCTGACCCACGGCTCGGCGTGCAC AGTGACCCTGC
sil11F	Forward NotI insert of StrepII tag- Ald tag-stop	GGCCGCAtggagccacccgcagttcgaaaag GGGTCAGTGTGCACGCCGAGCCGTt aaGG
sil11R	Reverse NotI insert of StrepII tag-Ald tag-stop	GGCCCCttaACGGCTCGGCGTGCAC AGTGACCCcttttcgaactgcgggtggctcca TGC
sil12F	Forward NotI insert of Ald tag- StrepII tag-stop	GGCCGCAGGGTCACTGTGCACGCC GAGCCGTGGGTTCAtggagccacccgcagt tcgaaaagtaaGG

Primer name	Description	Sequence
sil12R	Reverse NotI insert of Ald tag-StrepII tag-stop	GGCCCCttacttttcgaactgcggggtggctcca TGACCCACGGCTCGGCGTGACACAG TGACCCTCC
sil13F	Forward SfiI Fab heavy chain	GTTATTACTCGCGGCCAGCCGGCC
sil13R	Reverse Fab oligo 1 Ald tag	caTGACCCACGGCTCGGCGTGACACA GGCTCCCACAAGATTTGGGCTCAAC
sil14R	Reverse NotI oligo 2 Strep tag and stop codons	gctggcgccgcttacttttcgaactgcgggtg gctccaTGACCCACGGCTCGGC
sil15R	Reverse NotI oligo 2 Tev cleavage site	gctggcgccgcTCCCTGAAAATACAGG TTTTCTGACCCACGGCTCGGCG
sil16F	EIPCR NheI strep tag and stop codons	cgTgGCTAGCtggagccaccgcagttcgaaa agtaataaGGGGCCGCATAGACTGTT
sil16R	EIPCR NheI Ald tag addition	gcagGCTAGCTGACCCACGGCTCGG CGTGACAGGCTCCCACAAGATTTG GGCTCAAC
sil17F	EIPCR NheI anneal on his tag and NotI	cgTgGCTAGCGCGGCCGCACATCAT CATCACC
sil18F	Forward IPCR pet22B XhoI	CTCGAGCACCACCACCACC
sil18R	Reverse IPCR pet22B NcoI	GATATCCATGGCCATCGCCGGCTG
sil19F	Forward NcoI 9e10scfv	CCTTTCCATGGCGCAGGTGCAGCTG
sil19R	Reverse XhoI 9e10scfv	CCCTTCTCGAGACGGCTCGGCGTG
sil20F	Forward SpeI IPCR vector Her2ICD	tgggtACTAGTGCCAGCCCTCTGACG TCCATC
sil20R	Reverse NotI IPCR vector Her2	ccatGGCGGCCGCGGTGCTCGCGGC TCCGGGGG
sil21F	Forward SpeI IPCR vector Her3ICD	tgggtACTAGTTTtagGACAAACACTG

Primer name	Description	Sequence
		GTGCTG
sil21R	Reverse NotI IPCR vector Her3	ccatGGCGGCCCGCCTCGGAGCCCCG GGCCAGGC
sil22F	Forward NotI cloning SNAPf/CLIPf	GCACCGCGGCCCGCCatggacaaagactg cgaaatg
sil22R	Reverse SpeI cloning SNAPf/CLIPf	CCTAAACTAGTaccagcccaggcttgccca g
sil23F	Forward NotI cloning HaloTag7	CCGAGGCGGCCCGCCGAAATCGGTA CTGGCTTTCC
sil23R	Reverse SpeI cloning HaloTag7	CCTAAACTAGTACCGGAAATCTCCA GAGTAG
sil24F	Forward sequencing from end of Snap/CLIP tag	ccattctgatcccctgccac
sil25	Forward sequencing from end of Halo tag	GAGCCAGCGAACATCGTCGC
sil25F	Forward SpeI IPCR vector Her2 v2	cagtACTAGTGCCAGCCCTCTGACGT CCATCATCTCTGCGGTGG
sil25R	Reverse NotI IPCR vector Her2 v2	taagGCGGCCGCGGTGCTCGCGGCT CCGGGGGGCAAGAGGGCGAGGAG
sil26F	Forward NheI IPCR Fab vector	gtatGCTAGCcatcatcatcaccatcacgg
sil26F2	Forward NheI IPCR Fab vector v2	gtatGCTAGCcatcatcatcaccatcacggggc c
sil26R	Reverse NotII primer Fab Heavy CH1	tGCGGCCGCacaagatttgg
sil27R	Reverse NheI cloning SNAPf/CLIPf	cctaaGCTAGCaccagcccaggcttgcccag
sil28R	Reverse NheI cloning HaloTag7	cctaaGCTAGCaccggaaatctccagagtag

Primer name	Description	Sequence
sil29F	Forward SpeI IPCR vector Her2 full	tggtACTAGTcaagtgtgcaccggcag
sil30F	Forward SpeI IPCR vector Her3 full	tggtACTAGTgtgggcaactctcaggcag
sil31F	Forward NotI IPCR DT39 SEG vector	cactGCGGCCGCAtaagactctagagtcgac ctgc
sil31R	Reverse BlnI IPCR DT39 SEG vector (binds GS1 region)	Cgcctgctgagccagcagag
sil32F	Forward BlnI cloning of mEOS to replace GFP in DT39-SEG	tgctggctcagcaggcAGTGCGATTAAGC CAGACATG
sil32R	Reverse NotI cloning of mEOS	ctgaGCGGCCGCttaTCGTCTGGCATT GTCAGGC
sil33R	Reverse BlnI cloning mEOS to replace GFP in DT39-SEG	ctgaGCTCAGCtattaTCGTCTGGCATT GTCAGGC
sil34F	Forward AgeI cloning of mEOS	cgtaACCGGTagtGCGATTAAGCCAGA CATG
sil34R	Reverse XbaI cloning of mEOS	cactTCTAGAttaTCGTCTGGCATTGTC AGGC
sil35F	Forward NotI cloning of Snap.EGFR	cgtaGCGGCCGCCatgacccctccgggac gg
sil35R	Reverse XhoI cloning of Snap.EGFR	TAGCCTCGAGGTCCTGCTCCAATAA ATTCAGTGC
sil36F	Forward XhoI IPCR pENTR vector	actgCTCGAGGCTAGCGCTAGAGGG CCCTTC
sil36R	Reverse NotI IPCR pENTR vector	catGGCGGCCGCAGAGCCTGCTTTTT TGTACAAAGTTGGC

## A.5: List of Publications and Patents

**Liang SI**, van Lengerich B, Eichel K, Kwon MW, Murrow L, Patterson DM, Yoon TY, von Zastrow M, Jura N, Gartner ZJ. Epidermal Growth Factor Receptor Dimerization and Auto-phosphorylation is Not Sufficient for MAPK Pathway Activation. (*in preparation*)

**Liang SI**, McFarland JM, Rabuka D, Gartner ZJ. A Modular Approach for Assembling Aldehyde-Tagged Proteins on DNA Scaffolds. *J Am Chem Soc*, 136(31), 10850-3, 2014

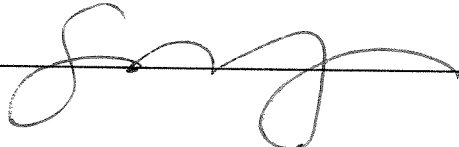
Weber RW, **Liang SI**, Selden NS, Desai TA, Gartner ZJ. Efficient Targeting of Fatty-Acid Modified Oligonucleotides to Live Cell Membranes through Stepwise Assembly. *Biomacromolecules*, 15(12), 4621-5, 2014

Patent: Publication Number WO20123067203 A1; Title: Reprogramming of Cellular Adhesion. Description: Invention of membrane-anchored oligonucleotides and their use in research and therapeutic applications, 2012

**Publishing Agreement**

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.

Author Signature  Date June 9, 2016