

ABSTRACT

TIRUTHANI, KARTHIK RAJAGOPALAN. Enabling Improved Understanding of Biological Processes through Protein Engineering. (Under the direction of Balaji M. Rao).

Protein-protein interactions are at the heart of most if not all biological processes and molecular recognition tools allow us to study these interactions. Small molecules and antibodies are currently used for characterizing and studying these protein-protein interactions. However small molecule drugs often do not have desired affinity, are not specific and many either do not have a primary target, or well studied mechanism of action. On the other hand antibodies are more specific to their target and better characterized, but many antibodies are polyclonal, have limited shelf life, are often not sequenced and can have batch to batch variability. While some issues can be alleviated through the use of recombinant antibodies currently being developed, new solutions to engineer not just specific inhibitors and biosensors and but affinity reagents in general is desirable. Specifically there is a need for reagents developed using a customized approach, tailored for the application rather than standard off the shelf reagents which may not give the desired results. In this work we demonstrate the use of protein engineering to solve these issues and present two specific applications to validate the approach.

The availability of intracellular biosensors for live-cell imaging studies is presently limited. To demonstrate the use of protein engineering we engineer biosensors for epidermal growth factor receptor (EGFR). While tandem SH2 domain from PLC γ 1 (tSH2-WT) has been used as a marker of phosphorylated EGFR we show it lacks specificity for phosphorylated EGFR and EGF-stimulated recruitment of tSH2-WT differs qualitatively

from the expected kinetics of EGFR phosphorylation. To address the limitations of the tSH2-WT biosensor, we constructed a combinatorial library through random mutagenesis of the C-terminal SH2 domain (cSH2) of PLC γ 1. The library was screened using yeast surface display to isolate a mutant protein (mSH2) with enhanced specificity for the Y992 phosphorylation site in EGFR (pY992). Accordingly, a biosensor based on mSH2 faithfully reports kinetics of EGFR phosphorylation in live-cell imaging experiments. Further we explore the use of de novo binding proteins in engineering biosensors by isolating a mutant SPY992 from the Sso7d scaffold library for the Y992 phosphorylation site in EGFR (pY992) and show that unlike mSH2 which retains some promiscuous activity it does not have any affinity for platelet derived growth factor receptor (PDGFR).

Molecular mechanisms regulating human trophoblast differentiation remain poorly understood due to difficulties in obtaining primary tissues from very early developmental stages in humans. Therefore the use of human embryonic stem cells (hESCs) as a source for generating trophoblast tissues is of significant interest. However there is controversy over whether hESC-derived cells are indeed analogous to true trophoblasts found *in vivo*. Characterization of the secretome can help identify unique trophoblast markers and the signaling pathways responsive to the microenvironment of hESCs leading to mechanistic insight into differentiation and help address these issues. The secretory pathway is an important factor in cell-cell communication and represents one way for the acquisition of the malignant phenotype of cancer cells. The secretome can provide insight into some of these mechanisms and also help in identification of potential biomarkers.

We demonstrate that the binder isolated, VB15 is capable of immunoprecipitating organelles from cells and tissues. Through electron microscopy of NIH3T3 cells we show that VB15 binds vesicles and that these vesicles often contain secreted proteins like fibronectin. So VB15 can be used to obtain secretome of tissues which we plan to eventually demonstrate through the isolation of secretory vesicles from third trimester human placenta, vCTB and STB isolated from first trimester human placenta. Our approach serves as a blueprint for engineering affinity reagents for isolation of any organelle of interest even if the proteome of the organelle is unknown.

© Copyright 2015 Karthik Rajagopalan Tiruthani

All Rights Reserved

Enabling Improved Understanding of Biological Processes through Protein Engineering

by
Karthik Rajagopalan Tiruthani

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Chemical Engineering

Raleigh, North Carolina

2015

APPROVED BY:

Balaji M. Rao
Committee Chair

Jason M. Haugh

David C. Muddiman

Gregory T. Reeves

BIOGRAPHY

Karthik Tiruthani was born in Hyderabad, India. He received his Bachelors in Technology in Mechanical Engineering (Mechatronics) in 2006 from Jawaharlal Nehru Technological University. He worked at Zetatek Industries before beginning graduate studies at North Carolina State University (NCSU) in 2007. At NCSU he earned his MS in Mechanical Engineering under the guidance of Dr. Melur Ramasubramanian in 2008 and continued on for a doctoral degree. In 2011 he joined the Protein Engineering lab of Dr. Balaji Rao to pursue his interest in biotechnology and specifically biosensor design. During the course of his study he has received Graduate Certificates in Medical Devices (JGC-MD) and in Technology, Entrepreneurship and Commercialization (TEC).

ACKNOWLEDGMENTS

I would like to thank my advisor Dr. Balaji Rao for the constant positive reinforcement and belief in me that helped a lot over the course of my graduate studies. I am grateful to Dr. Haugh for his guidance and discussion on various projects. I would like to acknowledge undergraduate researcher Stephen Ryan who over a period of two years worked on a lot of small projects that helped me progress towards completion. I would also like to thank Dr. Nimish Gera and Dr. Prasenjit Sarkar, past members of the Rao lab for their help with getting started with my research, and my colleagues Kevin, Carlos and Adam for an enjoyable lab environment. Finally I would also like to thank my parents for their support and encouragement over this journey.

TABLE OF CONTENTS

| | |
|---|-------------|
| LIST OF TABLES..... | vii |
| LIST OF FIGURES..... | viii |
| ENGINEERING AFFINITY REAGENTS – THE NEED AND THE OPPORTUNITY | 1 |
| 1.1 Introduction | 2 |
| 1.1.1 Protein Engineering..... | 3 |
| 1.1.2 Desired characteristics of engineered binding proteins | 6 |
| 1.1.3 Engineering reagents for isolation of secretory vesicles..... | 7 |
| 1.1.4 Engineering phosphospecific biosensors | 8 |
| 1.2 Thesis Overview..... | 9 |
| 1.3 References | 11 |
| DESIGN AND EVALUATION OF PROTEIN BIOSENSORS FOR LIVE CELL IMAGING OF EPIDERMAL GROWTH FACTOR RECEPTOR PHOSPHORYLATION | 21 |
| 2.1 Introduction | 22 |
| 2.2 Results..... | 24 |
| 2.2.1 Modeling EGF Receptor dynamics | 24 |
| 2.2.2 The tSH2-WT biosensor lacks specificity and does not faithfully track kinetics of EGFR phosphorylation..... | 27 |
| 2.2.3 The engineered mSH2 protein has increased specificity for pY992 of EGFR | 29 |
| 2.2.4 The mSH2 biosensor closely tracks the expected kinetics of EGFR phosphorylation and displays increased specificity for the pY992 site of EGFR in live-cell imaging experiments | 31 |
| 2.2.5 The engineered SPY992 protein has specificity for pY992 of EGFR with no affinity for pY1021 site of PDGFR..... | 33 |
| 2.3 Discussion | 34 |
| 2.4 Materials and Methods | 37 |
| 2.4.1 Library generation | 37 |
| 2.4.2 Isolation of SH2 domain mutants binding the pY992 site in EGFR | 38 |
| 2.4.3 Recombinant expression and purification of SH2 domain mutants | 39 |
| 2.4.4 Far western blotting | 40 |

| | | |
|--|--|-----------|
| 2.4.5 | Cell Culture..... | 40 |
| 2.4.6 | TIRF imaging and image analysis..... | 41 |
| 2.4.7 | Statistical analysis..... | 42 |
| 2.5 | References | 43 |
| ENGINEERING AFFINITY REAGENTS FOR EFFICIENT ISOLATION OF SECRETORY VESICLES | | 56 |
| 3.1 | Introduction | 57 |
| 3.2 | Materials and Methods | 59 |
| 3.2.1 | Construction of Sso6904 mutant library for yeast surface display | 59 |
| 3.2.2 | Cell Culture..... | 61 |
| 3.2.3 | Generation of target for negative and positive selection..... | 61 |
| 3.2.4 | Isolation of vesicle binding protein VB15 by magnetic selection and FACS..... | 62 |
| 3.2.5 | Recombinant expression and purification of VB15..... | 63 |
| 3.2.6 | Far western blotting | 64 |
| 3.2.7 | Immunofluorescence | 65 |
| 3.2.8 | Immunohistochemistry and histology analysis | 65 |
| 3.2.9 | Transmission Electron Microscopy | 65 |
| 3.2.10 | Human Placental Samples | 67 |
| 3.2.11 | Mouse tissue samples..... | 68 |
| 3.3 | Results..... | 68 |
| 3.3.1 | The engineered protein VB15 has increased specificity to fraction containing secretory vesicles..... | 68 |
| 3.3.2 | The engineered protein VB15 binds specifically to a single protein that is conserved across species..... | 69 |
| 3.3.3 | VB15 is capable of isolating organelles by immunoprecipitation | 69 |
| 3.3.4 | VB15 likely binds secretory vesicles..... | 70 |
| 3.3.5 | VB15 is capable of isolating organelles from tissues by immunoprecipitation | 70 |
| 3.4 | Discussion | 71 |
| 3.5 | References | 73 |
| TROPHOBLAST DIFFERENTIATION OF HUMAN EMBRYONIC STEM CELLS | | 90 |
| 4.1 | Introduction | 91 |
| 4.2 | HESC differentiation to trophoblast | 93 |

| | | |
|---|--|------------|
| 4.2.1 | Monolayer cultures using BMP4 and/or Activin/Nodal/TGF β inhibition..... | 93 |
| 4.2.2 | Differentiation using embryoid bodies | 95 |
| 4.3 | Markers used for characterizing hESC-derived cells | 96 |
| 4.4 | Role of differentiation media | 97 |
| 4.5 | Role of culture heterogeneity..... | 99 |
| 4.6 | Cdx2 as a trophoblast marker | 100 |
| 4.7 | Concluding remarks..... | 101 |
| 4.8 | References | 103 |
| CONCLUSIONS AND FUTURE WORK..... | | 131 |

LIST OF TABLES

| | | |
|------------------|--|-----|
| Table 2.1 | Protein Sequences of engineered biosensors | 53 |
| Table 2.2 | List of primers used | 54 |
| Table 2.3 | List of peptides used for library screening | 55 |
| Table 3.1 | DNA Sequence of Sso6904 with mutated residues | 88 |
| Table 3.2 | List of primers used | 88 |
| Table 3.3 | Isopycnic density of subcellular organelles | 89 |
| Table 4.1 | Overview of methods used for hESC differentiation to trophoblast | 112 |

LIST OF FIGURES

| | | |
|-------------------|--|----|
| Figure 1.1 | General overview of protein engineering by yeast surface display | 20 |
| Figure 2.1 | Characterization of the tSH2-WT biosensor | 47 |
| Figure 2.2 | Far western blotting analysis of cSH2 | 48 |
| Figure 2.3 | Analysis of intermediate cell populations during cSH2 library screening | 49 |
| Figure 2.4 | Characterization of specificity of the mSH2 biosensor | 50 |
| Figure 2.5 | Comparison of kinetics of EGFR phosphorylation detected using mSH2 and tSH2 –WT biosensors in c’1000 and c’1000F cell lines..... | 51 |
| Figure 2.6 | Comparison of kinetics of EGFR phosphorylation detected using mSH2 and tSH2 –WT biosensors in NR6 and NR6PAR cell lines | 52 |
| Figure 2.7 | Effect of changing binding affinity of biosensor-target interaction or biosensor concentration on readout | 52 |
| Figure 2.8 | Characterization of specificity of the SPY992 biosensor | 53 |
| Figure 3.1 | Mutated residues on Sso6904 used for library generation..... | 77 |
| Figure 3.2 | Schematic of Subcellular Fractionation using Sucrose Density Gradient Centrifugation | 78 |
| Figure 3.3 | Analysis of intermediate cell populations during isolation of secretory vesicle binders | 79 |

| | | |
|--------------------|--|----|
| Figure 3.4 | Characterization of VB15 binding to subcellular fractions | 80 |
| Figure 3.5 | Comparison of sequence of VB15 and the Sso6904 scaffold | 81 |
| Figure 3.6 | Far western blotting analysis of VB15..... | 82 |
| Figure 3.7 | Immunoprecipitation of organelles from NIH3T3 cells using VB15 | 83 |
| Figure 3.8 | Immunoprecipitation of organelles from S2 cells using VB15 | 84 |
| Figure 3.9 | TEM image of a crosssection of NIH3T3 cells labeled using VB15..... | 85 |
| Figure 3.10 | Immunoprecipitation of organelles from mouse tissues | 86 |
| Figure 3.11 | Immunohistochemistry of mouse embryos..... | 87 |

CHAPTER 1

ENGINEERING AFFINITY REAGENTS – THE NEED AND THE OPPORTUNITY

1.1 Introduction

Protein-protein interactions are at the heart of most if not all biological processes in a cell. The ability to control and study these interactions in a very specific fashion is important for a mechanistic understanding of cellular processes. Some approaches currently used to study these interactions include the use of genetic techniques (*1–9*) to increase or decrease a level of one or both the proteins. An alternate approach is the use of molecular recognition i.e. specific non-covalent interactions, either through small molecules or antibodies, to either affect protein function or in assays like microscopy, immunoprecipitation, western blotting and flow cytometry to gain insight into these interactions.

While approaches like gene insertion, deletion or mutation offer specificity, the ability to temporally affect function is lost. Deletion of genes coding for protein can have undesirable effects and result in modification of pathways other than the one of interest. Similarly, whereas the use of RNA interference using inducible expression (*10–13*) provides temporal control, it has issues with stability, delivery and off target effects (*14–16*). Small molecules represent an easy way of targeting intracellular protein-protein interactions in live cell assays with the desired temporal control, potential for spatial control and ease of delivery. However many small molecule drugs either do not have a primary target, or well-studied mechanism of action (*17*). Additionally despite significant advances in improving specificity (*18, 19*) few small molecules are usually highly specific (*20, 21*). The way small molecule drugs are obtained through high throughput phenotypic screens results in the

majority of small molecules acting on conserved sites, such as the ATP binding site in kinases which makes obtaining specificity a big challenge.

Antibodies are widely regarded as the gold standard for characterizing and studying protein-protein interactions. However their use in live cell assays as a biosensor is limited by challenges associated with delivery. Further many antibodies are polyclonal, have limited shelf life, are often not sequenced and can have batch to batch variability. Additionally tagging them with tags of interest or unnatural amino acids conferring interesting properties is also a significant challenge. While progress is being made through development of recombinant antibodies (22) to overcome these issues other alternatives are desirable.

Overall there is a need for new solutions to engineer not just specific inhibitors and biosensors and but affinity reagents in general. While projects like the Human Protein Atlas are developing affinity reagents for all human proteins (23) it is highly likely that different reagents maybe better suited for different applications and the ones generated may not be ideally suited to study protein modifications. Scientists interested in targeting proteins for e.g., only when the receptor is activated or a protein has a specific conformation or post translational modification (24) require reagents developed using a customized approach and may not get the desired results from standard off the shelf reagents.

1.1.1 Protein Engineering

Protein engineering can help overcome a lot of challenges associated with the techniques currently used in studying protein-protein interactions. Binding proteins

engineered from mutagenesis of single chain variable fragment of antibodies (scFv) (25–28), small hyperthermophilic proteins (29, 30) or domains from other proteins like 10th type III domain of human fibronectin (Fn3), SH2 domain, PDZ domain, ankyrin repeats, armadillo repeats, etc (31–37) represent another alternative to antibodies as reagents. The significant advantage of these proteins besides the cost and ease of production is their potential for use in intracellular live cell assays as a biosensor.

Specificity to the target protein can be easily engineered using protein engineering by targeting regions on a protein outside conserved regions, which are usually regions of low homology across the protein family. The targeting of specific regions on a protein through protein engineering facilitates the study of function of proteins in specific signaling pathways. Identifying binding partners to these domains thus allows a more fine-grained control of protein function. For example the protein Dishevelled interacts with over 50 different proteins (38) through three conserved domains (DIX domain, PDZ domain, DEP domain) and two conserved regions (basic region, proline rich region). Each of these interactions is involved in regulating many different functions such as β -catenin stability, secondary axis formation, axon differentiation, cell migration, etc and even in cell fate decisions. Targeting specific regions either through inhibitors or specific biosensors engineered using binding proteins could allow the study of the protein function without perturbing other pathways potentially simplifying the analysis.

The use of hyperthermophilic proteins, specifically Sso7d (a DNA binding protein from *Sulfolobus solfataricus*) in generating binding proteins to a wide range of targets has

been demonstrated in previous work (29). The basis for identification of desired proteins from combinatorial libraries is the linking of phenotype to genotype. This is achieved in cell free systems like mRNA (39–41) or ribosome (42–44) display through linking of the protein to mRNA by covalent bond or to the ribosome. In cell based systems like phage(45–47), bacteria(48–50), yeast(51–53) and mammalian(54, 55) cells this link is established by fusing the protein of interest to a cell surface protein. While each display technique is advantageous for a specific application, a combination of these techniques(56, 57) has also sometimes been used to take advantage of each. For example a combination of mRNA and yeast display used to engineer binding proteins (30) enables a higher diversity library to be initially selected using mRNA display. The enriched binding pool from mRNA display is then converted to a yeast display library to take advantage of affinity discrimination by FACS possible using yeast display. A comparison of these techniques is also available to determine the most suitable method (58). An overview of yeast display which will be used in this study is presented in Figure 1.1

Unlike small molecules proteins are not normally cell-permeable. However conjugation/fusion with cell penetrating/permeable peptides (CPP) or protein transduction domains (PTD) like antennapedia, TAT, transportan or polyarginine allows delivery into the cell (59–61) through active transport mechanisms. While the exact mechanism of transduction is not known it is thought to be through Clathrin-mediated endocytosis. Since the efficiency of endosomal escape is also unknown it leads to challenges in identifying the exact amount of active protein delivered. Other strategies frequently used are transient

transfection of plasmid DNA coding for protein of interest through the use of cationic lipids like Lipofectamine, electroporation, calcium phosphate precipitation or viral delivery (62–64). Based on an analysis of literature it appears while direct protein delivery may be ideal it is not currently reliable or repeatable enough and plasmid delivery of binding protein (with potential use of inducible expression) represents a realistic approach to achieve the desired results.

1.1.2 Desired characteristics of engineered binding proteins

Binding proteins can be custom engineered for a variety of applications depending on the affinity and specificity desired. For example an ideal biosensor should have an intermediate affinity to its target. Whereas low affinity results in very little binding and low signal, very high affinity may result in saturation of readout or perturbation of the target protein depending on the expression levels (65). On the other hand for a reagent used for immunoprecipitation the affinity can be variable depending on the target. For example while high affinity is desirable for immunoprecipitation of proteins it may not be essential for immunoprecipitation of organelles where specificity is more important since there is likely avidity between engineered binder on a bead and the organelle target which can amplify the affinity of both desired specific and undesired non-specific interactions. While affinity and specificity are key criterion depending on the application other characteristics may also become essential. For example high contrast tumor imaging requires small size reagents for rapid clearance in addition to high affinity whereas therapeutics require high serum half-life to improve their efficacy and high specificity to minimize any side effects. While there have

been many engineered proteins developed as inhibitors (31, 33, 66–68), there is a dearth of specific intracellular biosensors for live cell imaging. Similarly western blotting and immunoprecipitation reagents (32, 69–71) have been engineered but a systematic approach for engineering binding proteins for isolation of organelles is currently lacking.

1.1.3 Engineering reagents for isolation of secretory vesicles

Proteins secreted from the cell, like cytokines, growth factors and other signaling molecules play a crucial role in many cellular processes. Of particular interest here is hESC differentiation which is known to be significantly affected by the microenvironment. Characterizing the secretome can help identify the signaling pathways responsive to the microenvironment of hESCs leading to mechanistic insight into differentiation. Trophoblast formation is a major process underlying implantation of the blastocyst in the endometrium and development of the placenta. The cells of the trophoblast (TE) are precursors of all trophoblast cell types in the placenta. Identifying the secretome is especially important in hESC differentiation to trophoblast where knowledge of molecular mechanisms is lacking. We propose the engineering of binding protein that is capable of immunoprecipitating secretory vesicles. This will allow us to probe the secretome of primary placental tissue, not currently possible, allowing us to better understand trophoblast differentiation and improve on the hESC-trophoblast model. Further the binder could also help in identifying cancer biomarkers. Cancer cells acquire their malignant phenotype through the manipulation of signaling processes involved in growth, proliferation, apoptosis, angiogenesis and one of the key factors in achieving this is through secretory factors (72). So studying the

microenvironment of cancer cells can provide insight into these mechanisms and help in identification of potential biomarkers. Currently the cancer secretome is obtained through the study of cancer cell lines, tumor proximal body fluids (73). While it is understood that the biology of these cell lines is often unrepresentative of the cells they are derived from (74) they currently represent the only option. Most importantly since these cell lines have already acquired the malignant phenotype, they are unlikely to be useful in studying cancer progression. The binder developed in this study can be used to isolate secretory vesicles from tissue biopsies. This bypasses the need to culture the cells potentially altering their biology and provides more physiologically relevant information and in identification of biomarkers.

1.1.4 Engineering phosphospecific biosensors

While the eventual goal is to apply protein engineering to study human embryonic stem cell (hESC) differentiation, the tools for the development of these reagents needs to be validated and characterized. One challenge towards this end is the ability to engineer binding proteins to specific domains or epitopes on target proteins, especially when the target protein is in a specific desired state. We propose the engineering and evaluation of a biosensor to study the dynamics of phosphorylation and internalization of epidermal growth factor receptor (EGFR). This biosensor allows us to validate our approach to engineering domain and state specific binding proteins that eventually based on their affinity (65) could either be used as biosensors or inhibitors. EGFR represents a good target because its expression level is modulated by factors secreted by third trimester placenta (75) affecting trophoblast proliferation and migration. Further it is also thought to contribute to aberrant trophoblast

differentiation and failure to invade leading to preeclampsia (76). Thus the tools developed in this study can be developed further to study the pathway both in cultured trophoblast and hESC derived trophoblast to improve the hESC – trophoblast model.

1.2 Thesis Overview

The engineering of binding proteins to secretory vesicles and phosphorylation sites on EGFR is described. We explore the application of these binders derived using artificial scaffolds to study biological processes in a cell in three different ways. First as a biosensor to monitor EGFR internalization, second as an immunoprecipitation reagent to isolate secretory vesicles to characterize the secretome of cells and tissues and third as a tool to study hESC differentiation to trophoblast and potentially trophoblast biology.

Chapter 2 describes results pertaining to engineering and evaluation of a phospo and site specific biosensor to EGFR. We constructed a combinatorial library through random mutagenesis of the C-terminal SH2 domain (cSH2) of PLC γ 1. The library was screened using yeast surface display to isolate a mutant protein (mSH2) with enhanced specificity for the Y992 phosphorylation site in EGFR (pY992). Accordingly, a biosensor based on mSH2 faithfully reports kinetics of EGFR phosphorylation in live-cell imaging experiments. However mSH2 demonstrated non-specific binding to PDGFR (pY1021) since library was based on a promiscuous SH2 scaffold. To overcome this a binding protein SPY992 was engineered using the Sso7d scaffold (29) that demonstrated selective binding to pY992 without any non specific binding to PDGFR. This approach, integrating theoretical

considerations with a protein engineering strategy, can be generalized to design and evaluate suitable biosensors for other intracellular targets.

Chapter 3 describes results pertaining to engineering and evaluation of a protein capable of isolating secretory vesicles. We constructed a combinatorial library through the introduction of degenerate NNK codons into a calcium binding protein from *Sulfolobus solfataricus*, Sso6904. This library was screened using yeast surface display to isolate a protein (VB15) that we then thoroughly characterized. We demonstrate the application of this binder in isolation of secretory vesicles from NIH3T3, HEK293T and drosophila S2 cells. Further we also demonstrate the application of this binder in isolation of secretory vesicles from mouse pancreas, liver and placenta.

Chapter 4 gives an overview of previously described efforts to obtain trophoblasts from hESCs. Generating trophoblast tissues from hESCs is of significant interest because obtaining primary tissues from very early developmental stages in humans is challenging which results in the molecular mechanisms being poorly understood. However whether hESC-derived cells are analogous to true trophoblasts found *in vivo* is controversial and a discussion of some of the issues pertaining to this controversy are presented.

1.3 References

1. J. C. Miller, M. C. Holmes, J. Wang, D. Y. Guschin, Y.-L. Lee, I. Rupniewski, C. M. Beausejour, A. J. Waite, N. S. Wang, K. a Kim, P. D. Gregory, C. O. Pabo, E. J. Rebar, An improved zinc-finger nuclease architecture for highly specific genome editing., *Nat. Biotechnol.* **25**, 778–785 (2007).
2. J. a Zuris, D. B. Thompson, Y. Shu, J. P. Guilinger, J. L. Bessen, J. H. Hu, M. L. Maeder, J. K. Joung, Z.-Y. Chen, D. R. Liu, Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo., *Nat. Biotechnol.* , 1–10 (2014).
3. D. Seruggia, L. Montoliu, The new CRISPR-Cas system: RNA-guided genome engineering to efficiently produce any desired genetic alteration in animals., *Transgenic Res.* (2014), doi:10.1007/s11248-014-9823-y.
4. H. Wang, H. Yang, C. S. Shivalila, M. M. Dawlaty, A. W. Cheng, F. Zhang, R. Jaenisch, One-Step Generation of Mice Carrying Mutations in Multiple Genes by CRISPR/Cas-Mediated Genome Engineering., *Cell* **153**, 910–8 (2013).
5. C. R. Hale, S. Majumdar, J. Elmore, N. Pfister, M. Compton, S. Olson, A. M. Resch, C. V. C. Glover, B. R. Graveley, R. M. Terns, M. P. Terns, Essential features and rational design of CRISPR RNAs that function with the Cas RAMP module complex to cleave RNAs., *Mol. Cell* **45**, 292–302 (2012).
6. D. H. Kim, J. J. Rossi, Strategies for silencing human disease using RNA interference., *Nat. Rev. Genet.* **8**, 173–84 (2007).
7. D. Subramanyam, R. Blelloch, From microRNAs to targets: pathway discovery in cell fate transitions., *Curr. Opin. Genet. Dev.* **21**, 498–503 (2011).

8. D. P. Bartel, MicroRNAs: Genomics, Biogenesis, Mechanism, and Function, *Cell* **116**, 281–297 (2004).
9. F. Stegmeier, G. Hu, R. J. Rickles, G. J. Hannon, S. J. Elledge, A lentiviral microRNA-based system for single-copy polymerase II-regulated RNA interference in mammalian cells., *Proc. Natl. Acad. Sci. U. S. A.* **102**, 13212–13217 (2005).
10. S. Freundlieb, C. Schirra-Müller, H. Bujard, A tetracycline controlled activation/repression system with increased potential for gene transfer into mammalian cells., *J. Gene Med.* **1**, 4–12 (1999).
11. L. M. Fedorov, O. Y. Tyrsin, V. Krenn, E. V Chernigovskaya, U. R. Rapp, Tet-system for the regulation of gene expression during embryonic development, *Transgenic Res* **10**, 247–258 (2001).
12. A. R. Buskirk, D. R. Liu, Creating small-molecule-dependent switches to modulate biological functions, *Chem. Biol.* **12**, 151–161 (2005).
13. M. K. Pastuszka, J. A. Mackay, Biomolecular engineering of intracellular switches in eukaryotes, *J. Drug Deliv. Sci. Technol.* **20**, 163–169 (2010).
14. M. Amarzguioui, J. J. Rossi, D. Kim, Approaches for chemically synthesized siRNA and vector-mediated RNAi., *FEBS Lett.* **579**, 5974–81 (2005).
15. D. Samarsky, C. Faherty, in *RNA Interference: Principles and Applications*, J. Rossi, Ed. (The Biomedical & Life Sciences Collection, 2007).
16. D. Castanotto, J. J. Rossi, The promises and pitfalls of RNA-interference-based therapeutics., *Nature* **457**, 426–33 (2009).

17. E. Gregori-Puigjané, V. Setola, J. Hert, B. a Crews, J. J. Irwin, E. Lounkine, L. Marnett, B. L. Roth, B. K. Shoichet, Identifying mechanism-of-action targets for drugs and probes., *Proc. Natl. Acad. Sci. U. S. A.* **109**, 11178–83 (2012).
18. T. Anastassiadis, S. W. Deacon, K. Devarajan, H. Ma, J. R. Peterson, Comprehensive assay of kinase catalytic activity reveals features of kinase inhibitor selectivity., *Nat. Biotechnol.* **29**, 1039–45 (2011).
19. M. I. Davis, J. P. Hunt, S. Herrgard, P. Ciceri, L. M. Wodicka, G. Pallares, M. Hocker, D. K. Treiber, P. P. Zarrinkar, Comprehensive analysis of kinase inhibitor selectivity., *Nat. Biotechnol.* **29**, 1046–51 (2011).
20. S. Davies, H. Reddy, M. Caivano, P. Cohen, Specificity and mechanism of action of some commonly used protein kinase inhibitors., *Biochem. J.* **351**, 95–105 (2000).
21. A. C. Dar, K. M. Shokat, The evolution of protein kinase inhibitors from antagonists to agonists of cellular signaling., *Annu. Rev. Biochem.* **80**, 769–95 (2011).
22. V. Marx, Finding the right antibody for the job, *Nat. Methods* **10**, 703–707 (2013).
23. V. Marx, Calling the next generation of affinity reagents, *Nat. Methods* **10**, 829–833 (2013).
24. L. Kummer, P. Parizek, P. Rube, B. Millgramm, A. Prinz, P. R. E. Mittl, M. Kaufholz, B. Zimmermann, F. W. Herberg, A. Plückthun, Structural and functional analysis of phosphorylation-specific binders of the kinase ERK from designed ankyrin repeat protein libraries.*Proc. Natl. Acad. Sci. U. S. A.* (2012), doi:10.1073/pnas.1205399109.
25. J. S. Swers, B. a Kellogg, K. D. Wittrup, Shuffled antibody libraries created by in vivo homologous recombination and yeast surface display., *Nucleic Acids Res.* **32**, e36 (2004).

26. G. Schaefer, L. Haber, L. M. Crocker, S. Shia, L. Shao, D. Dowbenko, K. Totpal, A. Wong, C. V Lee, S. Stawicki, R. Clark, C. Fields, G. D. Lewis Phillips, R. a Prell, D. M. Danilenko, Y. Franke, J.-P. Stephan, J. Hwang, Y. Wu, J. Bostrom, M. X. Sliwkowski, G. Fuh, C. Eigenbrot, A two-in-one antibody against HER3 and EGFR has superior inhibitory activity compared with monospecific antibodies., *Cancer Cell* **20**, 472–86 (2011).
27. G. Chao, W. L. Lau, B. J. Hackel, S. L. Sazinsky, S. M. Lippow, K. D. Wittrup, Isolating and engineering human antibodies using yeast surface display., *Nat. Protoc.* **1**, 755–68 (2006).
28. B. J. Hackel, J. R. Neil, F. M. White, K. D. Wittrup, Epidermal growth factor receptor downregulation by small heterodimeric binding proteins., *Protein Eng. Des. Sel.* **25**, 47–57 (2012).
29. N. Gera, M. Hussain, R. C. Wright, B. M. Rao, Highly stable binding proteins derived from the hyperthermophilic Sso7d scaffold., *J. Mol. Biol.* **409**, 601–16 (2011).
30. M. Hussain, N. Gera, A. B. Hill, B. M. Rao, Scaffold diversification enhances effectiveness of a superlibrary of hyperthermophilic proteins, *ACS Synth. Biol.* **2**, 6–13 (2013).
31. A. Stahl, M. T. Stumpp, A. Schlegel, S. Ekawardhani, C. Lehrling, G. Martin, M. Gulotti-Georgieva, D. Villemagne, P. Forrer, H. T. Agostini, H. K. Binz, Highly potent VEGF-A-antagonistic DARPins as anti-angiogenic agents for topical and intravitreal applications., *Angiogenesis* (2012), doi:10.1007/s10456-012-9302-0.
32. Y. L. Boersma, G. Chao, D. Steiner, K. D. Wittrup, A. Plückthun, Bispecific designed ankyrin repeat proteins (DARPins) targeting epidermal growth factor receptor inhibit A431 cell proliferation and receptor recycling., *J. Biol. Chem.* **286**, 41273–85 (2011).

33. T. Kaneko, H. Huang, X. Cao, X. Li, C. Li, C. Voss, S. S. Sidhu, S. S. C. Li, Superbinder SH2 Domains Act as Antagonists of Cell Signaling, *Sci. Signal.* **5**, ra68–ra68 (2012).
34. M. Ferrer, J. Maiolo, P. Kratz, J. L. Jackowski, D. J. Murphy, S. Delagrave, J. Inglese, Directed evolution of PDZ variants to generate high-affinity detection reagents., *Protein Eng. Des. Sel.* **18**, 165–73 (2005).
35. F. Parmeggiani, R. Pellarin, A. P. Larsen, G. Varadamsetty, M. T. Stumpp, O. Zerbe, A. Caflisch, A. Plückthun, Designed armadillo repeat proteins as general peptide-binding scaffolds: consensus design and computational optimization of the hydrophobic core., *J. Mol. Biol.* **376**, 1282–304 (2008).
36. G. Varadamsetty, D. Tremmel, S. Hansen, F. Parmeggiani, A. Plückthun, Designed Armadillo Repeat Proteins: Library Generation, Characterization and Selection of Peptide Binders with High Specificity., *J. Mol. Biol.* , 1–20 (2012).
37. B. J. Hackel, A. Kapila, K. D. Wittrup, Picomolar affinity fibronectin domains engineered utilizing loop length diversity, recursive mutagenesis, and loop shuffling., *J. Mol. Biol.* **381**, 1238–52 (2008).
38. C. Gao, Y.-G. Chen, Dishevelled: The hub of Wnt signaling., *Cell. Signal.* **22**, 717–27 (2010).
39. a D. Keefe, Protein selection using mRNA display., *Curr. Protoc. Mol. Biol.* **Chapter 24**, Unit 24.5 (2001).
40. T. T. Takahashi, R. J. Austin, R. W. Roberts, mRNA display: ligand discovery, interaction analysis and beyond., *Trends Biochem. Sci.* **28**, 159–65 (2003).
41. S. W. Cotten, J. Zou, R. Wang, B.-C. Huang, R. Liu, mRNA display-based selections using synthetic peptide and natural protein libraries. *Methods Mol. Biol.* **805**, 287–97 (2012).

42. D. Lipovsek, A. Plückthun, In-vitro protein evolution by ribosome display and mRNA display., *J. Immunol. Methods* **290**, 51–67 (2004).
43. B. Ning, M. Liu, Y. Sun, Z. Sun, Y. Zhang, Construction of ribosome display library based on lipocalin scaffold and screening anticalins with specificity for estradiol, *Analyst* (2012) (available at <http://pubs.rsc.org/en/content/articlehtml/2012/an/c2an16119b>).
44. L. Lewis, C. Lloyd, Optimisation of antibody affinity by ribosome display using error-prone or site-directed mutagenesis., *Methods Mol. Biol. (Clifton, NJ)* **805**, 139–161 (2012).
45. H. R. Hoogenboom, Overview of antibody phage-display technology and its applications., *Methods Mol. Biol.* **178**, 1–37 (2002).
46. F. Tian, M.-L. Tsao, P. G. Schultz, A phage display system with unnatural amino acids., *J. Am. Chem. Soc.* **126**, 15962–3 (2004).
47. W. J. J. Finlay, L. Bloom, O. Cunningham, in *Methods in molecular biology*, Methods in Molecular Biology. D. Walls, S. T. Loughran, Eds. (Humana Press, Totowa, NJ, 2011), vol. 681, pp. 87–101.
48. S. Ståhl, M. Uhlén, Bacterial surface display: trends and progress., *Trends Biotechnol.* **15**, 185–92 (1997).
49. P. S. Daugherty, G. Chen, M. J. Olsen, B. L. Iverson, G. Georgiou, Antibody affinity maturation using bacterial surface display., *Protein Eng.* **11**, 825–32 (1998).
50. J. Rockberg, J. Löfblom, B. Hjelm, M. Uhlén, S. Ståhl, Epitope mapping of antibodies using bacterial surface display, *Nat. Methods* **5**, 1039 – 1045 (2008).
51. E. T. Boder, K. D. Wittrup, Yeast surface display for screening combinatorial polypeptide libraries., *Nat. Biotechnol.* **15**, 553–557 (1997).

52. E. T. Boder, K. D. Wittrup, Yeast surface display for directed evolution of protein expression, affinity, and stability., *Methods Enzymol.* **328**, 430–444 (2000).
53. L. Pepper, Y. Cho, A decade of yeast surface display technology: Where are we now?, *Comb. Chem. High Throughput Screen.* **11**, 127–134 (2008).
54. R. R. Beerli, M. Bauer, R. B. Buser, M. Gwerder, S. Muntwiler, P. Maurer, P. Saudan, M. F. Bachmann, Isolation of human monoclonal antibodies by mammalian cell display., *Proc. Natl. Acad. Sci. U. S. A.* **105**, 14336–41 (2008).
55. C. Zhou, F. W. Jacobsen, L. Cai, Q. Chen, W. D. Shen, Development of a novel mammalian cell surface antibody display platform., *MAbs* **2**, 508–18 (2010).
56. X. Hu, S. Kang, C. Lefort, M. Kim, M. M. Jin, Combinatorial libraries against libraries for selecting neoepitope activation-specific antibodies., *Proc. Natl. Acad. Sci. U. S. A.* **107**, 6252–7 (2010).
57. D. R. Bowley, T. M. Jones, D. R. Burton, R. a Lerner, Libraries against libraries for combinatorial selection of replicating antigen-antibody pairs., *Proc. Natl. Acad. Sci. U. S. A.* **106**, 1380–5 (2009).
58. N. Gera, M. Hussain, B. M. Rao, Protein selection using yeast surface display, *Methods* **60**, 15–26 (2013).
59. S. W. Jones, R. Christison, K. Bundell, C. J. Voyce, S. M. V Brockbank, P. Newham, M. a Lindsay, Characterisation of cell-penetrating peptide-mediated peptide delivery., *Br. J. Pharmacol.* **145**, 1093–102 (2005).
60. V. P. Torchilin, Tat peptide-mediated intracellular delivery of pharmaceutical nanocarriers., *Adv. Drug Deliv. Rev.* **60**, 548–58 (2008).

61. M. Konno, S. Masui, T. S. Hamazaki, H. Okochi, Intracellular reactivation of transcription factors fused with protein transduction domain., *J. Biotechnol.* **154**, 298–303 (2011).
62. J.-S. Kim, H. S. Chu, K. I. Park, J.-I. Won, J.-H. Jang, Elastin-like polypeptide matrices for enhancing adeno-associated virus-mediated gene delivery to human neural stem cells., *Gene Ther.* **19**, 329–37 (2012).
63. T. Helledie, V. Nurcombe, S. M. Cool, A simple and reliable electroporation method for human bone marrow mesenchymal stem cells., *Stem Cells Dev.* **17**, 837–48 (2008).
64. C. Magin-Lachmann, G. Kotzamanis, L. D’Aiuto, H. Cooke, C. Huxley, E. Wagner, In vitro and in vivo delivery of intact BAC DNA -- comparison of different methods., *J. Gene Med.* **6**, 195–209 (2004).
65. J. M. Haugh, Live-cell fluorescence microscopy with molecular biosensors: What are we really measuring?, *Biophys. J.* **102**, 2003–2011 (2012).
66. M. Kawe, P. Forrer, P. Amstutz, A. Plückthun, Isolation of intracellular proteinase inhibitors derived from designed ankyrin repeat proteins by genetic screening, *J. Biol. Chem.* **281**, 40252–40263 (2006).
67. P. Amstutz, H. Koch, H. K. Binz, S. a. Deuber, A. Plückthun, Rapid selection of specific MAP kinase-binders from designed ankyrin repeat protein libraries, *Protein Eng. Des. Sel.* **19**, 219–229 (2006).
68. E. Nordberg, M. Friedman, L. Göstring, G. P. Adams, H. Brismar, F. Y. Nilsson, S. Ståhl, B. Glimelius, J. Carlsson, Cellular studies of binding, internalization and retention of a radiolabeled EGFR-binding affibody molecule, *Nucl. Med. Biol.* **34**, 609–618 (2007).

69. L. Xu, P. Aha, K. Gu, R. G. Kuimelis, M. Kurz, T. Lam, A. C. Lim, H. Liu, P. a. Lohse, L. Sun, S. Weng, R. W. Wagner, D. Lipovsek, Directed evolution of high-affinity antibody mimics using mRNA display, *Chem. Biol.* **9**, 933–942 (2002).
70. E. Karatan, M. Merguerian, Z. Han, M. D. Scholle, S. Koide, B. K. Kay, Molecular recognition properties of FN3 monobodies that bind the Src SH3 domain, *Chem. Biol.* **11**, 835–844 (2004).
71. N. Gera, A. B. Hill, D. P. White, R. G. Carbonell, B. M. Rao, S. Karnik, Ed. Design of pH Sensitive Binding Proteins from the Hyperthermophilic Sso7d Scaffold, *PLoS One* **7**, e48928 (2012).
72. D. Hanahan, R. a Weinberg, Hallmarks of cancer: the next generation., *Cell* **144**, 646–74 (2011).
73. T. B. M. Schaaij-Visser, M. De Wit, S. W. Lam, C. R. Jiménez, The cancer secretome, current status and opportunities in the lung, breast and colorectal cancer context, *Biochim. Biophys. Acta - Proteins Proteomics* **1834**, 2242–2258 (2013).
74. J. R. Masters, Human cancer cell lines: fact and fantasy., *Nat. Rev. Mol. Cell Biol.* **1**, 233–236 (2000).
75. D. I. Sokolov, K. N. Furaeva, O. I. Stepanova, O. M. Ovchinnikova, L. P. Viazmina, G. R. Kozonov, T. U. Kuzminykh, S. a. Selkov, Changes in Functional Activity of JEG-3 Trophoblast Cell Line in the Presence of Factors Secreted by Placenta, *Arch. Med. Res.* **46**, 245–256 (2015).
76. D. R. Armant, R. Fritz, B. a. Kilburn, Y. M. Kim, J. K. Nien, N. J. Maihle, R. Romero, R. E. Leach, Reduced expression of the epidermal growth factor signaling system in preeclampsia, *Placenta* **36**, 270–278 (2015).

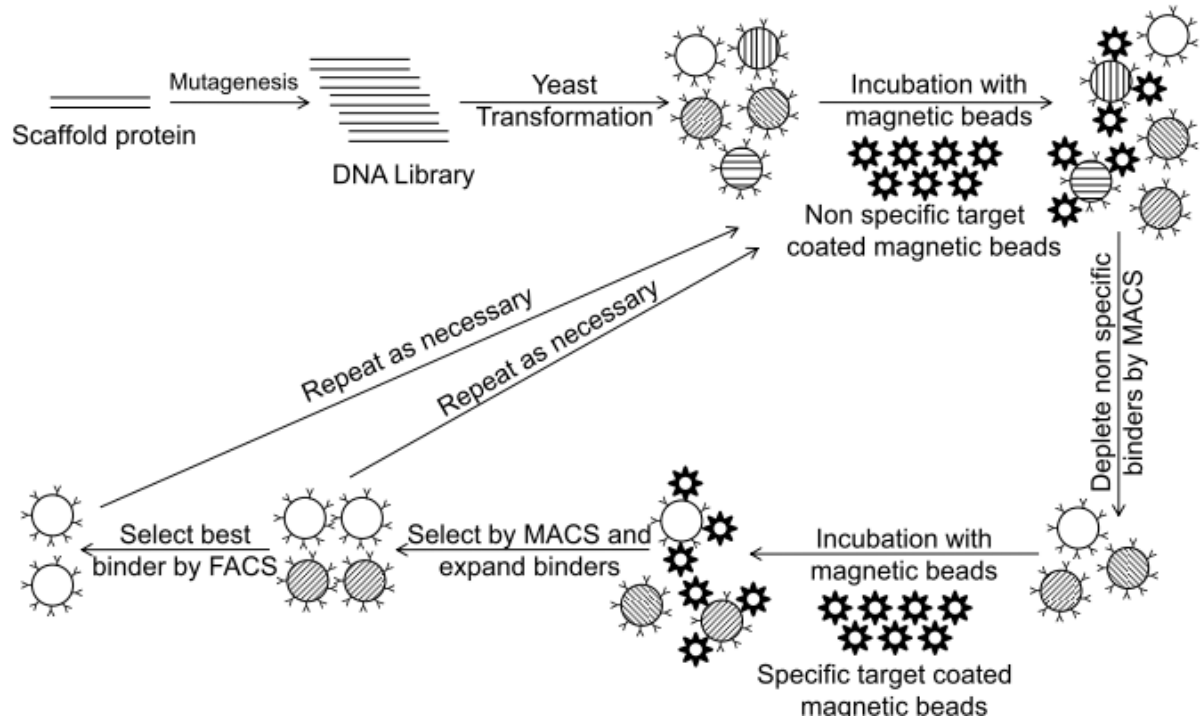


Figure 1.1: General overview of protein engineering by yeast surface display

CHAPTER 2

DESIGN AND EVALUATION OF PROTEIN BIOSENSORS FOR LIVE CELL IMAGING OF EPIDERMAL GROWTH FACTOR RECEPTOR PHOSPHORYLATION

(Adapted from manuscript in preparation by Ahmed, Tiruthani et al.)

2.1 Introduction

Receptor tyrosine kinases (RTKs) are cell surface receptors that mediate several key cellular processes including proliferation, differentiation and cell migration (1, 2). Aberrant activation of RTKs and their downstream signaling pathways is a hallmark of most human cancers. Typically, ligand binding to the extracellular domain of the receptor results in activation of tyrosine kinase function in the cytoplasmic domain, which in turn results in phosphorylation of specific residues of the receptor and receptor-associated proteins. Thus, ligand binding triggers RTK activation and initiation of downstream signaling pathways. The ability to obtain spatially resolved measurements of receptor activation using live-cell imaging can uniquely enable a mechanistic understanding of the relationship between dynamic changes in signaling and resultant cell behavior. Yet, implementation of live cell imaging using intracellular biosensors remains a significant challenge.

Intracellular biosensors contain a molecular recognition element that should specifically recognize the target(s) of interest. However, there are only a small number of highly specific protein biosensors currently available. A second and arguably greater limitation is that biosensors may significantly interfere with or otherwise modulate the signaling process they are meant to detect. In this latter context, we have previously outlined a theoretical framework for understanding the relationship between fidelity of biosensor readout and binding affinity of the biosensor for its target (3). Notably, an ideal biosensor should have an intermediate affinity to its target. Whereas low affinity results in very little binding, very high affinity may result in saturation of readout if the target species is present

at high concentration, or perturb the target if the biosensor is present in excess. Conversely, low affinities may result in decreased signal corresponding to the target. Thus, both specificity and binding affinity of molecular recognition determines the fidelity of the biosensor readout to dynamic changes in the concentration of the target species being detected. Here we apply protein engineering tools in conjunction with the aforementioned theoretical principles to design a protein biosensor targeting a specific phosphotyrosine (pY) site in the EGFR.

Ligand binding to the extracellular domain of EGFR results in receptor dimerization and subsequent phosphorylation of multiple tyrosine residues in the cytoplasmic domain of EGFR, including Y992. We sought to design a protein biosensor to measure dynamic changes in the concentration of EGFR with phosphorylated Y992 (pY992) in individual cells. Fluorescent protein fusions of the pY-binding SH2 domains have been previously used to study EGFR signaling (4, 5). Here, we evaluated the tandem SH2 domain from wild-type PLC γ 1 (denoted tSH2-WT hereafter) that has been used to study the dynamics of EGFR phosphorylation (4). Consistent with previous reports of promiscuous binding of SH2 domains (6–8), we show that tSH2-WT lacks specificity for EGFR pY992. Furthermore, the juxtaposition of two SH2 domains in tSH2-WT results in intramolecular avidity; the two SH2 domains can simultaneously bind pY-containing epitopes. This results in high effective affinity of interaction between tSH2-WT and EGFR, which in turn distorts the readout corresponding to the abundance of phosphorylated EGFR in the plasma membrane.

To overcome these limitations of tSH2-WT, we isolated a mutant SH2 domain by screening a combinatorial library generated by random mutagenesis of the C-terminal SH2 domain in tSH2-WT. The isolated mutant (denoted mSH2) exhibits improved specificity for the pY992 site in EGFR relative to tSH2-WT. Most importantly, the readout from mSH2 in live-cell total internal reflection fluorescence (TIRF) microscopy experiments is consistent with the expected dynamics of cell-surface EGFR phosphorylation and endocytosis.

However while mSH2 represented an improvement over tSH2-WT, it still demonstrated non-specific binding to PDGFR. This is potentially because it was derived from a promiscuous scaffold. While further negative selection may potentially result in an improved SH2 mutant it is likely that it may still retain some promiscuous binding to a receptor not tested. It is conceivable that using protein scaffolds with no native binding for intracellular proteins is more likely to yield a biosensor that is more specific. To this end we isolated a mutant SPY992 from a Sso7d library (9) that was shown in vitro to not have any binding to PDGFR. This approach, integrating theoretical considerations with a protein engineering strategy, can be generalized to design and evaluate suitable biosensors for specific intracellular targets.

2.2 Results

2.2.1 Modeling EGF Receptor dynamics

We adopt a highly simplified description of receptor dynamics similar to the classic model of Wiley and Cunningham (10). This model does not consider complexities such as,

e.g., receptor dimerization. In the absence of the SH2 biosensor, the dynamic equations and initial conditions are as follows.

$$\begin{aligned}\frac{dR}{dt} &= V_{syn} - k_t R - k_f [L]R + k_r C_s; & R(0) &= \frac{V_{syn}}{k_t} \\ \frac{dC_s}{dt} &= k_f [L]R - k_r C_s - k_e C_s; & C_s(0) &= 0 \\ \frac{dC_i}{dt} &= k_e C_s - k_{deg} C_i; & C_i(0) &= 0\end{aligned}$$

The model assumes that free, surface receptor (R) is produced via *de novo* synthesis at a constant rate, V_{syn} . It is consumed via basal turnover of the membrane, with first-order rate constant k_t . These two parameters determine the initial density of R . At time zero, the ligand EGF is added to the external medium with concentration $[L]$, and this concentration is held constant in the model (i.e., its depletion is considered negligible). Formation of the ligand-receptor complex on the cell surface (C_s) occurs by mass-action kinetics (forward rate constant k_f , reverse rate constant k_r). Internalization of C_s is characterized by the first-order endocytic rate constant, k_e , which is greater than k_t ; thus, ligand binding induces overall receptor downregulation. Finally, the internalized ligand-receptor complex (C_i) is degraded with first-order rate constant k_{deg} . The values of the kinetic parameters described above were assigned according to order-of-magnitude estimates: $V_{syn} = 1$ nM/min (cytosolic volume basis); $k_t = 0.01$ min⁻¹; $k_f = 0.1$ nM⁻¹min⁻¹; $k_r = k_e = 0.1$ min⁻¹; $k_{deg} = 0.05$ min⁻¹. The initial receptor expression on the surface is $V_{syn}/k_t = 100$ nM, or roughly $\sim 10^5$ molecules/cell.

The SH2 biosensor binds to the phosphorylated EGF receptor. Receptor phosphorylation is presumably rapid upon ligand binding, and therefore we assume that all

ligated receptors are phosphorylated (or, rapidly come to an equilibrium of fast phosphorylation and dephosphorylation rates). Conversely, we assume that dissociation of ligand from the receptor, or of the biosensor from an unligated receptor, is followed by rapid dephosphorylation. After internalization, the EGF-EGFR complex remains phosphorylated, with its phosphorylation sites exposed to the cytoplasm (11). Phosphorylated receptors, whether on the surface or internalized, bind the SH2 biosensor with mass-action kinetics (association rate constant k_{on} , dissociation rate constant k_{off}). With these assumptions in place, the full set of equations and initial conditions are as follows.

$$\begin{aligned}
\frac{dR}{dt} &= V_{syn} - k_t R - k_f [L]R + k_r C_s + k_{off} R^* \\
\frac{dC_s}{dt} &= k_f [L]R - k_r C_s - k_e C_s - k_{on} [B]C_s + k_{off} C_s^* \\
\frac{dC_i}{dt} &= k_e C_s - k_{deg} C_i - k_{on} [B]C_i + k_{off} C_i^* \\
\frac{dR^*}{dt} &= -k_f [L]R^* + k_r C_s^* - k_{off} R^* \\
\frac{dC_s^*}{dt} &= k_{on} [B]C_s - k_{off} C_s^* + k_f [L]R^* - k_r C_s^* - k_{eSH2} C_s^* \\
\frac{dC_i^*}{dt} &= k_{on} [B]C_i - k_{off} C_i^* + k_{eSH2} C_s^* - k_{deg} C_i^* \\
\frac{d[B]}{dt} &= -k_{on} [B](C_s + C_i) + k_{off} (R^* + C_s^* + C_i^*) + k_{deg} C_i^*
\end{aligned}$$

In this model, the new species are the receptor species with asterisk superscripts, signifying that the SH2 biosensor is bound, and $[B]$, the concentration of biosensor in the cytosol. Besides the aforementioned k_{on} and k_{off} , the other new rate constant is k_{eSH2} , the endocytic rate constant for the ligated receptor with SH2 bound. We fixed $k_{on} = 0.1 \text{ nM}^{-1} \text{ min}^{-1}$ and varied $k_{off} = K_D k_{on}$. The high-affinity variant was $k_{off} = 1 \text{ min}^{-1}$ ($K_D = 10 \text{ nM}$), whereas

the low-affinity variant was $k_{off} = 100 \text{ min}^{-1}$ ($K_D = 1 \text{ }\mu\text{M}$). Note that, in both cases, $k_{off} \gg k_r$; therefore, the abundance of R^* is negligible. As for the parameter k_{eSH2} , we investigated the hypothesis that SH2 binding interferes with endocytosis and chose $k_{eSH2} = 0.1k_e$ (i.e., $k_{eSH2} = k_i$). Finally, one must specify the initial value of $[B]$. This parameter was set in the range of $100 \text{ nM} - 1 \text{ }\mu\text{M}$ as is typical for expressed fluorescent protein fusions, as indicated in the corresponding figure caption. All calculations were performed in the freely available Virtual Cell software environment (12), using the Combined Stiff Solver (IDA/CVODE) solver option. The Biomodel ‘SH2 biosensor’ is publicly available via www.vcell.org (note that the default units of VCell are μM and s, so we have externally converted those units to nM and min).

2.2.2 The tSH2-WT biosensor lacks specificity and does not faithfully track kinetics of EGFR phosphorylation

The PLC γ 1 SH2 domain pair tSH2-WT has been shown to bind phosphorylated EGFR (pEGFR) and prevent tyrosine dephosphorylation *in vitro*, and pY992 has been identified as its major binding site(13). Further, tagged tSH2-WT has been used in live cell imaging studies(4, 14). Therefore, we chose to evaluate tSH2-WT as a biosensor to monitor dynamics of EGFR phosphorylation at pY992. Accordingly, an EGFP-tSH2-WT fusion was expressed by transient transfection in parental NR6 cells (control) or NR6 cells expressing EGFR. Subsequently, cells were treated with EGF, and membrane localization of tSH2-WT was observed using TIRF microscopy. The fluorescence signal increases substantially in response to EGF stimulation of cells expressing EGFR, but not in parental NR6 cells (**Figure**

2.1A), suggesting that tSH2-WT may be used for monitoring kinetics of EGFR phosphorylation; however, the critical attributes of specificity and fidelity have yet to be characterized for this biosensor.

To assess the specificity of tSH2-WT, we monitored the fluorescence readout upon treatment of NR6 parental cells with sodium orthovanadate, an inhibitor of protein tyrosine phosphatases. Since NR6 parental cells lack EGFR, any recruitment of tSH2-WT is attributed to recognition of other pY motifs. tSH2-WT produces a significant readout in orthovanadate-treated parental cells, indicating poor specificity (**Figure 2.1B**). To further assess specificity of tSH2-WT, we conducted far-western blotting analysis. Lysates of NR6 parental cells treated with sodium orthovanadate, and of parental and EGFR-expressing NR6 cells treated with EGF, were probed with GST-tagged tSH2-WT (**Figure 2.1C**). Although a high-intensity band corresponding to EGFR was observed in the lysate of EGF-stimulated, EGFR-expressing cells as expected; the presence of other bands confirms the promiscuity of tSH2-WT binding.

To investigate the fidelity provided by tSH2-WT as a biosensor for live-cell imaging, it is instructive to review the expected kinetics of EGFR phosphorylation, on which there is a well-established literature(15–17). In response to a saturating dose of EGF, well above its apparent binding K_D (~1 nM), receptor activation on the cell surface peaks rapidly and then decreases because of enhanced endocytosis of the active receptor complex (18, 19). These kinetics are typified by a simple mathematical model of EGFR dynamics (10–12) by which the characteristic, transient kinetics of pEGFR on the cell surface is calculated (**Fig. 2.1D**). In

contrast, the experimental readout from the tSH2-WT sensor showed little adaptation (**Fig. 2.1B**). The cooperative binding of the two SH2 domains in tSH2-WT yields high binding avidity between the tSH2-WT biosensor and pEGFR(20, 21). Kinetic model calculations further suggest that the apparent lack of adaptation seen in the experiments is consistent with high overall binding affinity (low dissociation rate constant) of tSH2-WT for pEGFR (**Figure 2.1E**). Conversely, the model also predicts that a substantial increase of the dissociation rate constant (thus reducing the affinity) of biosensor binding to pEGFR results in a readout that is consistent with the true time course of pEGFR (**Figure 2.1E** vs. **Figure 2.1D**).

Taken together, our experimental data and modeling analysis lead to two insights. First, tSH2-WT is not a suitable biosensor, as it lacks specificity and does not faithfully track the expected kinetics of pEGFR. Second, the lack of fidelity of the tSH2-WT biosensor can be attributed to its high binding affinity for pEGFR; it is predicted that use of a biosensor with lower affinity for pEGFR will alleviate the discrepancy.

2.2.3 The engineered mSH2 protein has increased specificity for pY992 of EGFR

Use of a single SH2 domain instead of the tandem SH2 domains in tSH2-WT should eliminate the high affinity for pEGFR arising from the avidity effect. However, when a GST-tagged single SH2 domain – the C-terminal SH2 domain of the pair (cSH2) – was used to probe cell lysates in far-western blotting assays, no specific EGFR band was detected (**Fig 2.2**). Further, SH2 domains are inherently promiscuous binders. To identify an improved biosensor, we generated a combinatorial library by random mutagenesis of cSH2 and screened for mutants that show specific and detectable binding to pY992 of EGFR. A library

of $\sim 10^9$ mutants was generated using yeast surface display and screened using magnetic selection and fluorescence-activated cell sorting (FACS) (22), with a synthetic peptide corresponding to the pY992 site as the target. Negative selection steps by magnetic selection were used to eliminate binders to peptides that correspond to the unphosphorylated Y992 site, other phosphorylation sites in EGFR, and putative tyrosine phosphorylation motifs in other proteins that have sequence homology to the pY motifs in EGFR. The population of cells obtained after magnetic selection and two rounds of FACS was plated, and single clones were analyzed. DNA sequencing identified a mutant (mSH2) with a single amino acid substitution (C65G) relative to cSH2. Interestingly, this residue has been implicated in determination of specificity of binding of SH2 domains(23). The complete sequence of mSH2 is shown in **Table 2.1**. Analysis of intermediate cell populations during screening of the combinatorial library is shown in **Fig 2.3**.

We used yeast surface-displayed mSH2 and synthetic peptides to assess the specificity of binding to the pY992 site in EGFR relative to other phosphorylation sites. mSH2 exhibits > 10-fold greater signal for binding to the pY992 peptide relative to the pY1148 peptide and no binding was detected for the pY1068 (**Figure 2.4A**). We further used far-western blotting analysis of NR6 parental and EGFR-expressing cells treated with EGF, and NR6 parental cells treated with sodium orthovanadate. Cell lysates were probed with GST-tagged mSH2 (**Figure 2.4B**). A comparison with the corresponding analysis with tSH2-WT (**Figure 2.1C**) indicates that mSH2 has greater specificity for EGFR. Notably, non-specific bands are absent for the case of parental NR6 cells stimulated with EGF.

However, a single high molecular weight band is seen in orthovanadate-stimulated parental NR6 cells. We hypothesized that this band could be the platelet derived growth factor β -receptor (PDGFR), since tSH2-WT is known to bind pY1021 of PDGFR(24–26). Consistent with this hypothesis, yeast cells displaying mSH2 (or tSH2-WT as a control) were labeled with a synthetic peptide corresponding to the pY1021 site in PDGFR, and both mSH2 and tSH2-WT were confirmed to bind pY1021 of PDGFR (**Figure 2.4C**). This non-specific binding of mSH2 to PDGFR is consistent with our experimental protocols for library screening, wherein a negative selection step to eliminate binders to phosphorylation motifs on PDGFR was not incorporated.

These results show that biosensors with improved specificity can be obtained through random mutagenesis of native protein domains and combinatorial screening of variants with prescribed properties. In particular, mSH2 exhibits improved specificity for EGFR relative to tSH2-WT and binds selectively to the pY992 site relative to other EGFR phosphorylation sites; however, like tSH2-WT, mSH2 retains affinity for the pY1021 site of PDGFR. It is conceivable that using protein scaffolds with no native binding for intracellular proteins (27) may be more suited for engineering intracellular biosensors.

2.2.4 The mSH2 biosensor closely tracks the expected kinetics of EGFR phosphorylation and displays increased specificity for the pY992 site of EGFR in live-cell imaging experiments

To evaluate mSH2 as a biosensor, an mSH2-tdTomato fusion was transiently transfected in NR6 parental cells, or in NR6 cells expressing either EGFR or variants thereof.

The cells were treated with EGF or sodium orthovanadate, and membrane localization of mSH2 was observed using TIRF microscopy. Compared with the EGF-stimulated translocation of tSH2-WT in EGFR-expressing cells (**Figure 2.1A**), the corresponding time course of mSH2 recruitment shows a significantly lower fold change (**Figure 2.5A**; see **Figure 2.6A** for an overlay plot). A similar comparison holds true for sodium orthovanadate treatment of NR6 parental cells that lack EGFR (**Figure 2.5B**; overlay plot shown in **Figure 2.6B**). More importantly, unlike tSH2-WT, mSH2 exhibits the transient, EGF-stimulated response that is characteristic of pEGFR on the cell surface. Finally, the intracellular concentration of the biosensor also affects readout. Nevertheless, it is important to note that higher affinity of interaction between the biosensor and pEGFR results in significantly greater distortion of the readout relative to the true pEGFR concentration, than changes in biosensor concentration (**Figure 2.7**).

To further evaluate the specificity of mSH2 for EGFR pY992, we compared the readout from mSH2 and tSH2-WT in NR6 cells expressing truncation mutants of EGFR, c'1000 or c'1000F. The cytoplasmic domain of the c'1000 mutant is truncated at residue 1000 of EGFR, leaving Y992 as the only major phosphorylation site. In the c'1000F variant, Y992 is mutated to phenylalanine and therefore cannot be phosphorylated at that site; however, it is established that c'1000F mediates intracellular signaling in response to EGF stimulation(28), through heterodimerization with ErbB2. Similar to the responses in wild-type EGFR-expressing cells, there is a substantial difference between the EGF-stimulated responses of mSH2 and tSH2-WT in cells expressing the c'1000 mutant, in terms of both

their magnitudes and shapes (**Figure 2.5C**). The most compelling contrast was observed in the c'1000F-expressing cells, however; whereas tSH2-WT was recruited in response to EGF, albeit to a lesser extent, mSH2 showed no discernable response to the activated EGFR lacking pY992 (**Figure 2.5D**).

Taken together, these results demonstrate the suitability of the engineered mutant, mSH2, for live-cell imaging experiments. Unlike tSH2-WT, the mSH2 biosensor tracks the transient kinetics of EGFR phosphorylation and exhibits greater specificity for its intended target, pY992.

2.2.5 The engineered SPY992 protein has specificity for pY992 of EGFR with no affinity for pY1021 site of PDGFR

To identify an improved biosensor with no affinity to PDGFR, we screened for mutants from Sso7d library (9) was previously derived. Mutants that showed specific and detectable binding to pY992 of EGFR and no binding to pY1021 of PDGFR or Y992 of EGFR were selected. A library of $\sim 10^8$ mutants generated using yeast surface display was screened using magnetic selection and fluorescence-activated cell sorting (FACS) (22), with a synthetic peptide corresponding to the pY992 site as the target. Negative selection steps by magnetic selection were used to eliminate binders to peptides that correspond to the unphosphorylated Y992 site, other phosphorylation sites in EGFR, and putative tyrosine phosphorylation motifs in other proteins that have sequence homology to the pY motifs in EGFR and pY1021 of PDGFR. The population of cells obtained after magnetic selection and three rounds of FACS was plated, and single clones were analyzed. The complete sequence

of SPY992 is shown in **Table 2.1**. Analysis of clone binding to phosphorylated and unphosphorylated peptide Y992 and phospho peptide pY1021 of PDGFR is shown in **Fig 2.8**. Thus biosensors with improved specificity can be obtained from non native protein domains and combinatorial screening of variants with prescribed properties. In particular, SPY992 exhibits improved specificity for EGFR and binds selectively to the pY992 site relative to other EGFR phosphorylation sites. While SPY992 has no affinity for PDGFR it is not as selective as SH2 domain for phosphotyrosine. Thus its suitability, for live-cell imaging experiments as a phospho specific biosensor needs to be validated.

2.3 Discussion

Live-cell imaging studies of intracellular dynamics are presently limited by the availability of suitable biosensors. Here we show that a previously described biosensor using the tandem SH2 domains of PLC γ 1 (tSH2-WT) suffers from two major limitations in the context of studying EGFR phosphorylation dynamics. First, tSH2-WT lacks specificity, consistent with the inherent promiscuity of SH2 domain recognition(6–8). Second, EGF-stimulated translocation of tSH2-WT, though robust in magnitude, exhibits kinetics that qualitatively deviates from the expected transient kinetics of pEGFR at the cell surface.

Analysis of a basic kinetic model shows that the inability of tSH2-WT to faithfully track cell-surface pEGFR can be explained by its high binding affinity and occupancy of more than one site on the pEGFR. Pertinently, the avidity of the two linked SH2 domains in tSH2-WT results in high overall binding affinity of pEGFR in cells. As discussed earlier, high affinity of the biosensor-target interaction can result in saturation of the biosensor

readout or perturbation of the target's dynamics(3). Indeed, mutant SH2 domains evolved for high-affinity pY binding have been shown to antagonize pY-mediated signaling(29), and binding of tSH2-WT to pY992 from EGFR prevents tyrosine dephosphorylation *in vitro*(13). Our modeling results suggest that tSH2-WT perturbs EGFR dynamics by interfering with endocytosis of the complex.

We applied protein engineering tools to overcome limitations associated with high target binding affinity and lack of specificity of the tSH2-WT biosensor. Live-cell imaging studies confirmed that an mSH2 biosensor indeed exhibits specificity for the Y992 phosphorylation site in EGFR. Most importantly, the readout from the mSH2 biosensor is qualitatively consistent with the kinetics of EGFR phosphorylation in response to EGF treatment. Phosphorylation of cell surface EGFR peaks and then shows adaptation, in large part because of ligand-induced endocytosis. The plasma membrane translocation of mSH2 shows a nearly complete recovery during the course of chronic EGF stimulation, a much greater extent of adaptation than previously inferred from population measurements for the same cell line (30). In that study, a general anti-phosphotyrosine antibody was used to quantify pEGFR, and a mild acid strip was used to distinguish cell-surface and internalized pEGFR. The quantitative differences between these measurements might reflect distinct, site-specific kinetics of EGFR phosphorylation or differential endocytosis of the various phospho-forms of EGFR. Development of other phospho-specific biosensors will allow such hypotheses to be addressed.

While mSH2 exhibits greater specificity for pY992 of EGFR, it also retains affinity for the pY1021 site of PDGFR. This promiscuous interaction can be explained by the binding of tSH2-WT to PDGFR (24–26). We used extensive negative selection steps using synthetic peptides corresponding to proteins containing phosphorylation motifs homologous to EGFR. However, negative selection against the PDGFR site, which lacks homology with the sequence surrounding Y992 of EGFR, was not carried out; SH2 domain binding is based on the consensus pYXXP motif present in both the pY992 site in EGFR and the pY1021 site in PDGFR(31). This result underscores the importance of the negative selection strategy in our approach.

To overcome this issue resulting from promiscuous interaction of SH2 domains a library derived from a non native scaffold Sso7d was used to isolate SPY992 which displays selectivity for EGFR over PDGFR. While SPY992 binds pY992 site on EGFR specifically and shows no binding to pY1021 site of PDGFR it appears to have some binding to unphosphorylated Y992. Thus it is likely to have higher background fluorescence in live cell imaging experiments and needs further validation.

In summary, we demonstrate the integration of theoretical considerations with a protein engineering strategy to design an intracellular biosensor for site-specific phosphorylation of EGFR. Our approach serves as a blueprint for the design and evaluation of intracellular biosensors, towards the development of a broad palette of highly specific live-cell imaging reagents on par with phospho-specific antibodies used for immunoblotting.

2.4 Materials and Methods

2.4.1 Library generation

The C-terminal SH2 domain from PLC γ 1 (cSH2) was amplified using primers FP1 and RP1 from a plasmid containing the Rat PLCG1 gene(4) by error prone PCR as described(32); all primers were obtained from Integrated DNA Technologies (IDT; Coralville, IA), and sequences are listed in **Table 2.2**. Nucleotide analogues 8-oxo-dGTP and dPTP (Trilink Biotechnologies, San Diego, CA) were added at a final concentration of 10 μ M to Taq 2X Master Mix (New England Biolabs (NEB), Ipswich, MA) along with 0.2 μ M of each primer and 10 ng of template DNA. The PCR product after 30 cycles was visualized on a 1% agarose gel and then used as a template for amplification by Phusion DNA polymerase (NEB) using primers PCTF1 and PCTR1 to generate at least 50 μ g of final PCR product containing the sequences necessary for homologous recombination. The PCR product was concentrated by using Pellet Paint™ (Novagen, San Diego, CA). A yeast display library was generated using the PCR product through homologous recombination as described(33). Briefly, the yeast display vector pCTCON was linearized by restriction digestion with *NheI* and *BamHI* (NEB) and concentrated using Pellet Paint. 4 μ g of linearized vector and 12 μ g of PCR product was added to 400 μ l of electrocompetent yeast cells and electroporation was performed using a 2 mm electrode gap cuvette in a Bio-Rad Gene Pulser system (Bio-Rad, Hercules, CA) at 2.5kV and 25 μ F. Four such transformations were performed, and following a one hour outgrowth in 32 ml of 1:1 mix of 1M Sorbitol and YPD (10 g/l yeast extract, 20 g/l peptone, 20 g/l dextrose), the electroporated cells were expanded in 1L of SDCAA (20 g/l

dextrose, 5 g/l Casamino acids, 6.7 g/l yeast nitrogen base, 5.40 g/l Na₂HPO₄, 7.45 g/l NaH₂PO₄). Simultaneously, an aliquot was grown on SDCAA agar plates to estimate library size.

2.4.2 Isolation of SH2 domain mutants binding the pY992 site in EGFR

Peptides listed in **Table 2.3** were obtained from Genscript (Piscataway, NJ). All peptides were at least 90% pure and contained an N-Terminal biotin moiety. The yeast display library of cSH2 mutants generated was grown in SGCAA (20 g/l galactose, 5 g/l Casamino acids, 6.7 g/l yeast nitrogen base, 5.40 g/l Na₂HPO₄, 7.45 g/l NaH₂PO₄) to induce surface protein expression. Binders to non-target species were eliminated by negative selection using magnetic sorting as described(22). Biotinylated peptides were immobilized on biotin binder beads (Life Technologies, Carlsbad, CA). An equimolar mix of all bead-immobilized peptides other than the pY992 peptide was prepared. This mix was incubated with $\sim 10^{10}$ cells of the induced cSH2 library for an hour, and the beads were removed using a magnet. The remaining cells were incubated with beads containing pY992 peptide for 30 minutes. Following three washes in 1 ml PBS-BSA (PBS + 0.1% BSA), the beads were collected and grown in SDCAA. These expanded cells were again grown in SGCAA to induce surface protein expression and FACS was performed using a MoFlo cell sorter (Beckman Coulter Inc., Brea, CA). For cytometry approximately 5×10^7 induced cells were labeled simultaneously with an anti-c-myc chicken antibody (Life technologies, Carlsbad, CA) to quantify surface protein expression and a biotinylated peptide to quantify binding to the peptide. The secondary antibodies used were Streptavidin-PE and goat-anti-

chicken conjugated to Alexa-633 (Life technologies, Carlsbad, CA). Cells obtained after the magnetic selection step were sorted twice by FACS and the final pool of cells was grown on SDCAA plates. Yeast colonies were sequenced by Genewiz (La Jolla, CA). Further analysis of the sorted population and individual clones were performed using Accuri C6 analyzer (BD Biosciences, San Jose, CA).

2.4.3 Recombinant expression and purification of SH2 domain mutants

The final mutant selected for further characterization (mSH2), cSH2 and tSH2-WT were cloned from the yeast display vector into the pET42 vector using the restriction sites *SpeI* and *XhoI* to generate GST-tagged fusions. Primers FP2 and RP2 were used to introduce the restriction sites by PCR to facilitate cloning of mSH2 and cSH2 whereas primers FP3 and RP2 were used for cloning of tSH2-WT. Sequences were verified by colony sequencing using a T7 reverse primer by Genewiz (La Jolla, CA). Subsequently, plasmids were introduced into Rosetta Cells™ (EMD Biosciences, San Diego, CA). Protein expression was induced by addition of 1mM IPTG followed by incubation at 20 °C for 12 hours with shaking at 250 rpm. After induction, cells were harvested by centrifugation (4800g, 10min), resuspended in PBS with 0.2mM PMSF (Sigma Aldrich, St. Louis, MO) and lysed by sonication (2 sec ON, 5 sec OFF) for 7 minutes. The lysate was cleared by centrifugation and filtration through a 0.22 µm filter and purified using a GSTrap FF column (GE Healthcare Bio-Sciences, Pittsburgh, PA) on a Bio-Rad Biologic LP system using a linear glutathione gradient. The eluted fractions were analyzed by SDS PAGE and pure fractions were combined and dialyzed into 4L of PBS using a 3.5kDa MWCO filter (Thermo Fisher,

Rockford, IL). Protein concentrations were determined by bicinchoninic acid (BCA) assay (Thermo scientific, Rockford, IL).

2.4.4 Far western blotting

Far western blotting was performed following the protocol of(6). Cells were lysed in lysis buffer containing protease and phosphatase inhibitors (50mM HEPES, 100mM NaCl, 10% Glycerol, 1% Triton X-100, 1mM sodium orthovanadate, 50mM β -glycerophosphate, 5mM NaF, 1mM EGTA, 10mM sodium pyrophosphate, 10 μ g/ml each of Aprotinin, Leupeptin, Pepstatin A and Chymostatin).Lysates from NR6 parental cells and EGFR-overexpressing NR6 cells, stimulated with EGF(Peprotech Rocky Hill, NJ) or sodium orthovanadate (Thermo Fisher, Rockford, IL), were electrophoresed on a 7.5% SDS-PAGE gel and transferred to PVDF membranes using standard protocols. The membranes were blocked overnight at 4°C in 5% nonfat dry milk in TBST with 1 mM sodium orthovanadate and 5 mM EDTA. Primary incubation was performed using 20 nM mSH2-GST, 20 nM tSH2-WT-GST and 100nM cSH2-GST and the membranes were then washed with TBST and blocked for 1 hour at room temperature in 5% nonfat dry milk in TBST. Anti-GST HRP (GE Healthcare Bio-Sciences, Pittsburgh, PA) was used to probe for the GST tag and SuperSignal® West Femto substrate (Thermo Fisher, Rockford, IL) was used to obtain a chemiluminescent readout.

2.4.5 Cell Culture

Parental mouse NR6 cells(34) were cultured at 37°C, 5% CO₂ in Dulbecco's MEM α medium supplemented with 7.5% fetal bovine serum, 1% sodium pyruvate, 1% non-essential

amino acids and 1% PSG (L-glutamine and antibiotics penicillin and streptomycin). Three variants of this cell line with full length EGFR (wild-type), C-terminal truncation mutants (c'1000) and Y to F mutant of c'1000 (c'1000F)(4, 35) were cultured in same media, additionally supplemented with 350 µg/ml G418 antibiotic. All tissue culture reagents were purchased from Life Technologies (Carlsbad, CA). NR6 cell lines were transiently transfected with tSH2-WT-EGFP or mSH2-tdTomato using Lipofectamine with Plus reagent (Life Technologies, Carlsbad, CA) per the manufacturer's protocol.

2.4.6 TIRF imaging and image analysis

Using PCR amplification with primers tdEF1 and tdXR1 tdTomato from tdTomato-Lifeact was cloned into *EcoRI* and *XhoI* digested pET28 vector. Then mSH2 was subcloned into this construct from the yeast display vector using *NheI* and *BamHI* to generate a tdtomato fusion. These fusions were then cloned into pBM-IRES-puro through the restriction sites *NotI* and *HindIII* using primers FP4 and tdHR2. NR6 cells were transiently transfected with fluorescent protein-tagged SH2 domains (EGFP-tSH2-WT and mSH2-tdtomato). Cells were detached with a brief trypsin-EDTA treatment and suspended in imaging buffer (20 mM HEPES pH 7.4, 125 mM NaCl, 5 mM KCl, 1.5 mM MgCl₂, 1.5 mM CaCl₂, 10 mM glucose and 2 mg/ml fatty-acid-free BSA. After centrifugation at 100 × g for 3 min, the cells were resuspended in imaging buffer and counted using a Beckman Coulter Counter. Adhesive surfaces were prepared on clean, sterile 25 × 25 mm glass cover slips (Fisher Scientific, Pittsburgh, PA), which were coated with 100 µg/ml poly-D-lysine (Sigma Aldrich, St. Louis, MO) overnight at 4°C, washed with deionized, sterile water and dried. To make a

chamber for the cells, a Teflon ring was attached to the cover slip, and 1 ml of cell suspension containing 3×10^4 cells (approximately 100 cells/mm²) was added. Cells were allowed to spread and were serum-starved in imaging buffer for 4 h prior to imaging. Mineral oil was layered on top of the buffer to prevent evaporation during the experiment. Growth factors and inhibitors were diluted in the same buffer before adding to the cells. Images were acquired using TIRF microscopy with excitation of EGFP (488nm) and tdTomato (561 nm) achieved with incident beam energy of roughly 20–50 mJ per image. A 40× water immersion objective (Zeiss Achroplan, 0.8 NA) and 0.63× camera mount were used. Digital images were acquired using a cooled charge-coupled device (ORCA ER; Hamamatsu Photonics) and MetaMorph software (Universal Imaging, West Chester, PA). In each image, the intensity of the acellular region was defined as background and subtracted from the intensity of each pixel prior to further analysis. Images were analyzed using MATLAB and ImageJ. For each cell, the background-subtracted average intensity was calculated for each frame and then normalized by the average intensity measured prior to stimulation.

2.4.7 Statistical analysis

For comparison of the time courses for different biosensors, ANOVA (two-factor with replication) analysis of normalized intensity was performed and a *p* value was obtained for the sample being the source of variation.

2.5 References

1. D. L. Cadena, G. N. Gill, Receptor tyrosine kinases., *FASEB J.* **6**, 2332–7 (1992).
2. a Ullrich, J. Schlessinger, Signal transduction by receptors with tyrosine kinase activity., *Cell* **61**, 203–212 (1990).
3. J. M. Haugh, Live-cell fluorescence microscopy with molecular biosensors: What are we really measuring?, *Biophys. J.* **102**, 2003–2011 (2012).
4. J. M. Haugh, T. Meyer, Active EGF receptors have limited access to PtdIns(4,5)P(2) in endosomes: implications for phospholipase C and PI 3-kinase signaling., *J. Cell Sci.* **115**, 303–310 (2002).
5. C. Antczak, A. Bermingham, P. Calder, D. Malkov, K. Song, J. Fetter, H. Djaballah, Domain-Based Biosensor Assay to Screen for Epidermal Growth Factor Receptor Modulators in Live Cells, *Assay Drug Dev. Technol.* **10**, 24–36 (2012).
6. P. Nollau, B. J. Mayer, Profiling the global tyrosine phosphorylation state by Src homology 2 domain binding., *Proc. Natl. Acad. Sci. U. S. A.* **98**, 13531–13536 (2001).
7. K. Machida, B. J. Mayer, P. Nollau, Profiling the global tyrosine phosphorylation state., *Mol. Cell. Proteomics* **2**, 215–233 (2003).
8. K. Machida, C. M. Thompson, K. Dierck, K. Jablonowski, S. Kärkkäinen, B. Liu, H. Zhang, P. D. Nash, D. K. Newman, P. Nollau, T. Pawson, G. H. Renkema, K. Saksela, M. R. Schiller, D. G. Shin, B. J. Mayer, High-Throughput Phosphotyrosine Profiling Using SH2 Domains, *Mol. Cell* **26**, 899–915 (2007).
9. N. Gera, M. Hussain, R. C. Wright, B. M. Rao, Highly stable binding proteins derived from the hyperthermophilic Sso7d scaffold., *J. Mol. Biol.* **409**, 601–16 (2011).
10. H. S. Wiley, D. D. Cunningham, A steady state model for analyzing the cellular binding, internalization and degradation of polypeptide ligands., *Cell* **25**, 433–440 (1981).

11. P. C. Baass, G. M. Di Guglielmo, F. Authier, B. I. Posner, J. J. M. Bergeron, Compartmentalized signal transduction by receptor tyrosine kinases, *Trends Cell Biol.* **5**, 465–470 (1995).
12. D. C. Resasco, F. Gao, F. Morgan, I. L. Novak, J. C. Schaff, B. M. Slepchenko, Virtual Cell: Computational tools for modeling in cell biology, *Wiley Interdiscip. Rev. Syst. Biol. Med.* **4**, 129–140 (2012).
13. D. Rotin, B. Margolis, M. Mohammadi, R. J. Daly, G. Daum, N. Li, E. H. Fischer, W. H. Burgess, a Ullrich, J. Schlessinger, SH2 domains prevent tyrosine dephosphorylation of the EGF receptor: identification of Tyr992 as the high-affinity binding site for SH2 domains of phospholipase C gamma., *EMBO J.* **11**, 559–567 (1992).
14. T. P. Stauffer, T. Meyer, Compartmentalized IgE receptor-mediated signal transduction in living cells, *J. Cell Biol.* **139**, 1447–1454 (1997).
15. M. G. Poirier, S. Eroglu, J. F. Marko, The bending rigidity of mitotic chromosomes., *Mol. Biol. Cell* **13**, 2170–2179 (2002).
16. H. Schmidt-Glenewinkel, E. Reinz, R. Eils, N. R. Brady, Systems biological analysis of epidermal growth factor receptor internalization dynamics for altered receptor levels, *J. Biol. Chem.* **284**, 17243–17252 (2009).
17. H. S. Wiley, J. J. Herbst, B. J. Walsh, D. a Lauffenburger, M. G. Rosenfeld, G. N. Gill, The role of tyrosine kinase activity in endocytosis, compartmentation, and down-regulation of the epidermal growth factor receptor., *J. Biol. Chem.* **266**, 11083–11094 (1991).
18. L. Goh, a D. Sorkin, Endocytosis of Receptor Tyrosine Kinases, *Cold Spring Harb. Perspect. Biol.* **5**, a017459–a017459 (2013).
19. A. Sorkin, L. K. Goh, Endocytosis and intracellular trafficking of ErbBs, *Exp. Cell Res.* **315**, 683–696 (2009).
20. E. a. Ottinger, M. C. Botfield, S. E. Shoelson, Tandem SH2 domains confer high specificity in tyrosine kinase signaling, *J. Biol. Chem.* **273**, 729–735 (1998).

21. a Chattopadhyay, M. Vecchi, Q. S. Ji, R. Mernaugh, G. Carpenter, The role of individual SH2 domains in mediating association of phospholipase C-gamma1 with the activated EGF receptor., *J. Biol. Chem.* **274**, 26091–7 (1999).
22. N. Gera, M. Hussain, B. M. Rao, Protein selection using yeast surface display, *Methods* **60**, 15–26 (2013).
23. Z. Songyang, G. Gish, G. Mbamalu, T. Pawson, L. C. Cantley, A Single Point Mutation Switches the Specificity of Group III Src Homology (SH) 2 Domains to That of Group I SH2 Domains, *J. Biol. Chem.* **270**, 26029–26032 (1995).
24. Q. S. Ji, a Chattopadhyay, M. Vecchi, G. Carpenter, Physiological requirement for both SH2 domains for phospholipase C-gamma1 function and interaction with platelet-derived growth factor receptors., *Mol. Cell. Biol.* (1999).
25. R. a Klinghoffer, B. Duckworth, M. Valius, L. Cantley, a Kazlauskas, Platelet-derived growth factor-dependent activation of phosphatidylinositol 3-kinase is regulated by receptor binding of SH2-domain-containing proteins which influence Ras activity., *Mol. Cell. Biol.* **16**, 5905–5914 (1996).
26. S. M. Pascal, a U. Singer, G. Gish, T. Yamazaki, S. E. Shoelson, T. Pawson, L. E. Kay, J. D. Forman-Kay, Nuclear magnetic resonance structure of an SH2 domain of phospholipase C-gamma 1 complexed with a high affinity binding peptide., *Cell* **77**, 461–472 (1994).
27. M. Hussain, N. Gera, A. B. Hill, B. M. Rao, Scaffold diversification enhances effectiveness of a superlibrary of hyperthermophilic proteins, *ACS Synth. Biol.* **2**, 6–13 (2013).
28. T. Sasaoka, W. J. Langlois, F. Bai, D. W. Rose, J. W. Leitner, S. J. Decker, A. R. Saltiel, G. N. Gill, M. Kobayashi, B. Draznin, J. M. Olefsky, Involvement of ErbB2 in the signaling pathway leading to cell cycle progression from a truncated epidermal growth factor receptor lacking the C-terminal autophosphorylation sites, *J. Biol. Chem.* **271**, 8338–8344 (1996).
29. T. Kaneko, H. Huang, X. Cao, X. Li, C. Li, C. Voss, S. S. Sidhu, S. S. C. Li, Superbinder SH2 Domains Act as Antagonists of Cell Signaling, *Sci. Signal.* **5**, ra68–ra68 (2012).

30. J. M. Haugh, K. Schooler, a. Wells, H. S. Wiley, D. a. Lauffenburger, Effect of Epidermal Growth Factor Receptor Internalization on Regulation of the Phospholipase C- 1 Signaling Pathway, *J. Biol. Chem.* **274**, 8958–8965 (1999).
31. Z. Songyang, S. E. Shoelson, M. Chaudhuri, G. Gish, T. Pawson, W. G. Haser, F. King, T. Roberts, S. Ratnofsky, R. J. Lechleider, SH2 domains recognize specific phosphopeptide sequences., *Cell* **72**, 767–778 (1993).
32. N. Gera, A. B. Hill, D. P. White, R. G. Carbonell, B. M. Rao, S. Karnik, Ed. Design of pH Sensitive Binding Proteins from the Hyperthermophilic Sso7d Scaffold, *PLoS One* **7**, e48928 (2012).
33. L. Benatuil, J. M. Perez, J. Belk, C. M. Hsieh, An improved yeast transformation method for the generation of very large human antibody libraries, *Protein Eng. Des. Sel.* **23**, 155–159 (2010).
34. R. M. Pruss, H. R. Herschman, Variants of 3T3 cells lacking mitogenic response to epidermal growth factor., *Proc. Natl. Acad. Sci. U. S. A.* **74**, 3918–3921 (1977).
35. P. Chen, H. Xie, M. C. Sekar, K. Gupta, a Wells, Epidermal growth factor receptor-mediated cell motility: phospholipase C activity is required, but mitogen-activated protein kinase activity is not sufficient for induced cell movement., *J. Cell Biol.* **127**, 847–857 (1994).

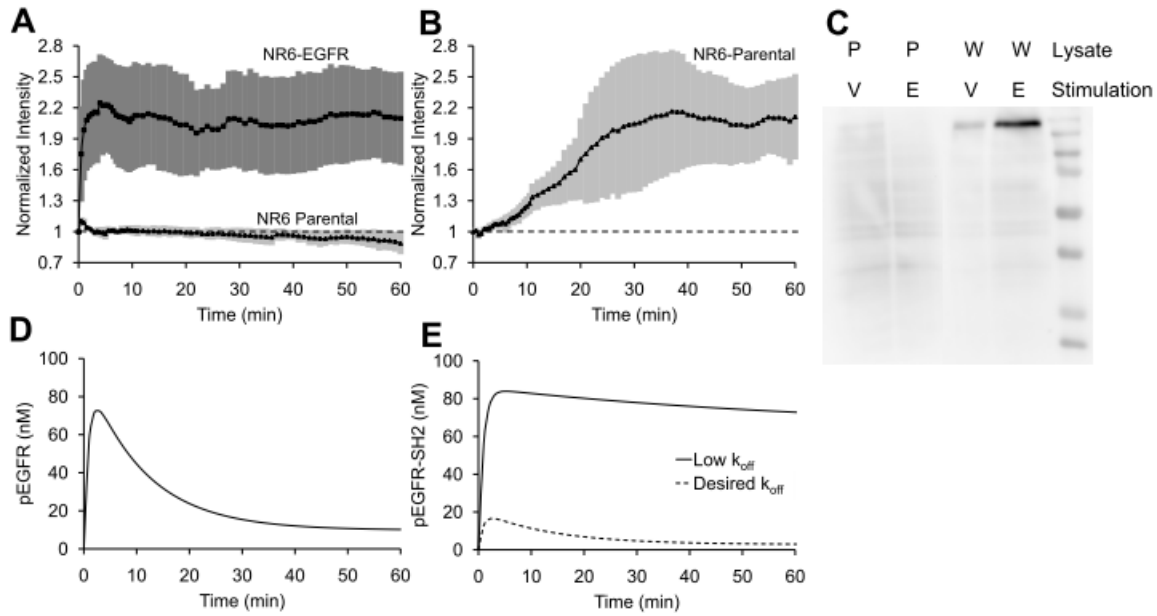


Figure 2.1: The tSH2-WT biosensor lacks specificity and does not faithfully track kinetics of EGFR phosphorylation. **(A)** Normalized recruitment of tSH2-WT in response to EGF (10 nM) stimulation of NR6 parental cells and NR6 cells expressing EGFR, as measured by TIRF microscopy ($n = 5$). **(B)** Normalized recruitment of tSH2-WT in NR6 parental cells treated with sodium orthovanadate (1 mM; $n = 7$). For **(A)** and **(B)**, error bars indicating standard deviation are shaded. **(C)** Far-western blotting analysis of NR6 cells expressing wild-type EGFR (W) or NR6 parental cells (P), stimulated with 10 nM EGF for 10 min. **(E)** or 1 mM sodium orthovanadate for 60 min. (V). Cell lysates were probed with tSH2-WT. **(D)** Kinetics of pEGFR predicted by the model. **(E)** Predicted effect of binding affinity of interaction between the biosensor and the target on biosensor readout. The desired k_{off} is higher (lower affinity), and therefore biosensor binding does not dramatically perturb receptor dynamics.

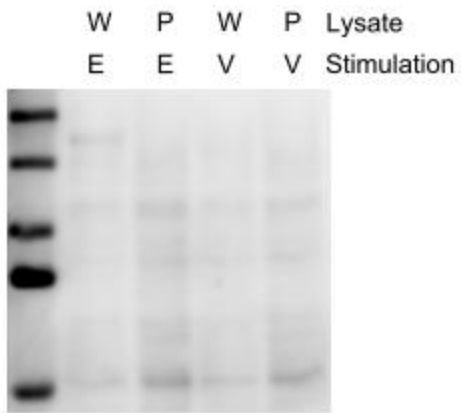


Figure 2.2: Far western blotting analysis of NR6 cells expressing wild-type EGFR (W) or NR6 parental cells (P), stimulated with 10 nM EGF for 10 min (E) or 1 mM sodium orthovanadate (V) for 60 min probed with cSH2.

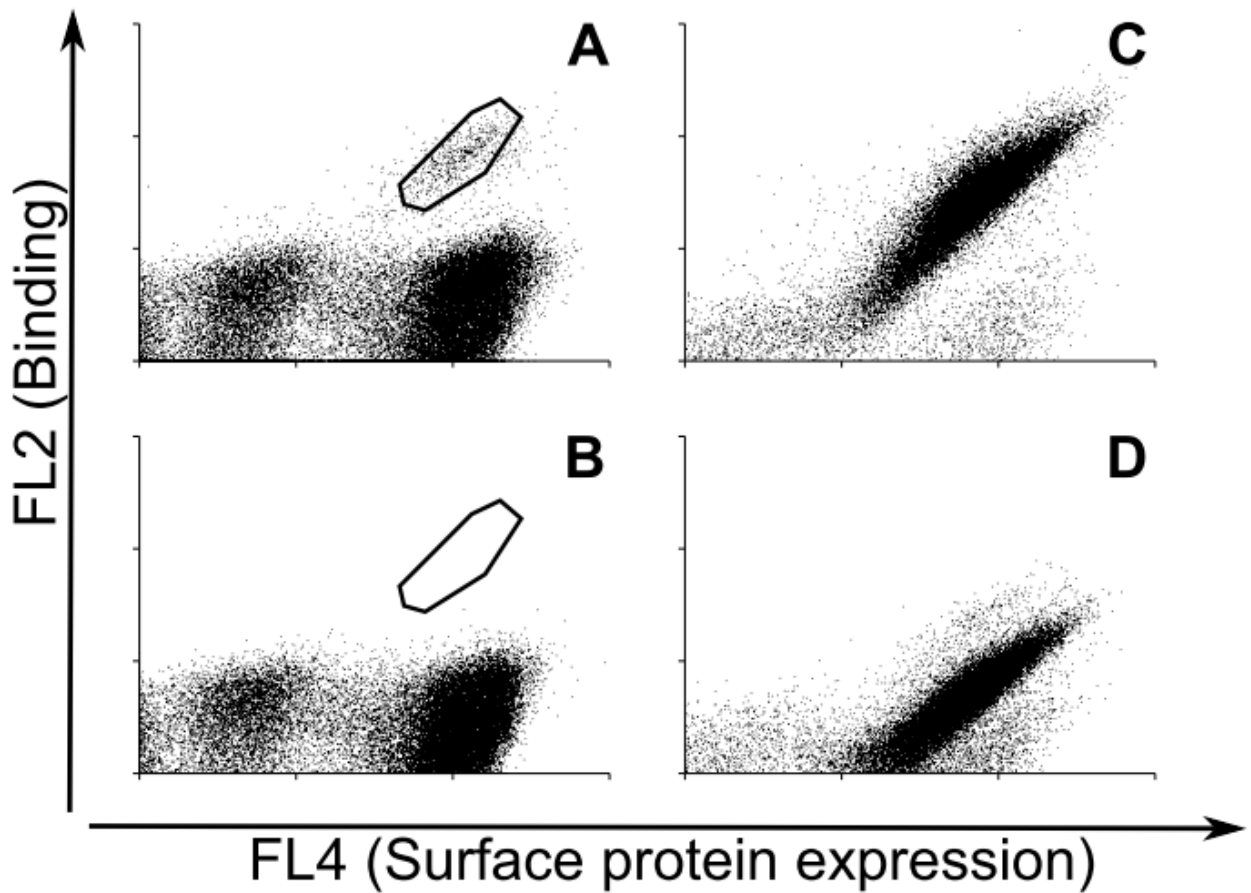


Figure 2.3: Analysis of intermediate cell populations during library screening. Yeast cells displaying cSH2 mutants were labeled with anti-c-myc antibody (chicken) and a biotinylated synthetic peptide, followed by secondary labeling with anti-chicken antibody conjugated with Alexa-633 and streptavidin-PE. Labeling of the pool of cells after magnetic sorting with (A) the pY992 peptide and (B) non-specific control (a 10 μ M equimolar mix of other peptides shown in Table 1). Sort gate is shown in A and B. Labeling of population obtained after two rounds of FACS (gated population) with (C) pY992 or (D) non-specific control

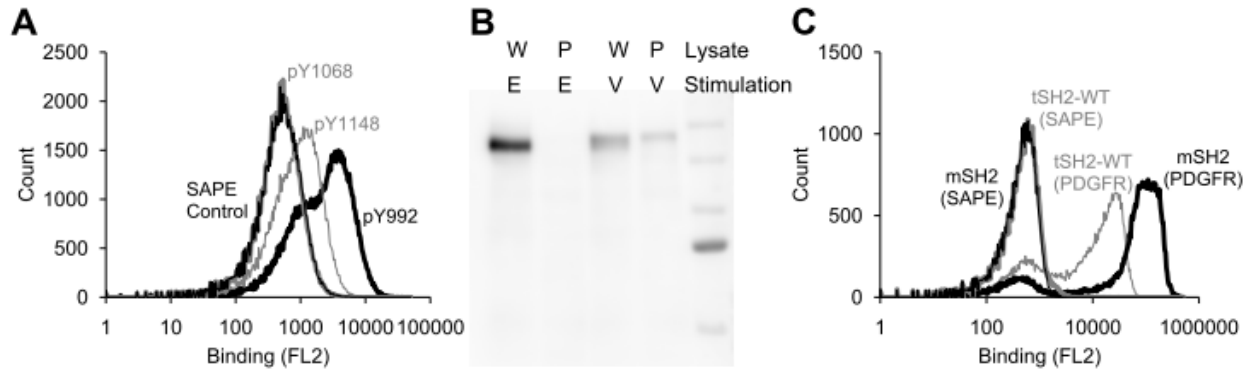


Figure 2.4: A mutant SH2 domain (mSH2) with increased specificity for pY992 of EGFR.

(A) Yeast cells displaying mSH2 were labeled with biotinylated synthetic peptides corresponding to pY992, pY1068 and pY1148 sites in EGFR, followed by secondary labeling with SA-PE. A control sample of yeast cells only labeled with SA-PE is also shown. (B) Far-western blotting analysis of NR6 cells overexpressing EGFR (W) or NR6 parental cells (P), stimulated with 10 nM EGF (E) or 1 mM sodium orthovanadate (V). Cell lysates were probed with mSH2. (C) Yeast cells displaying mSH2 or tSH2-WT were labeled with a biotinylated synthetic peptide corresponding to the pY1021 site in PDGFR, followed by secondary labeling with streptavidin-PE. A control sample of yeast cells only labeled with SA-PE is also shown.

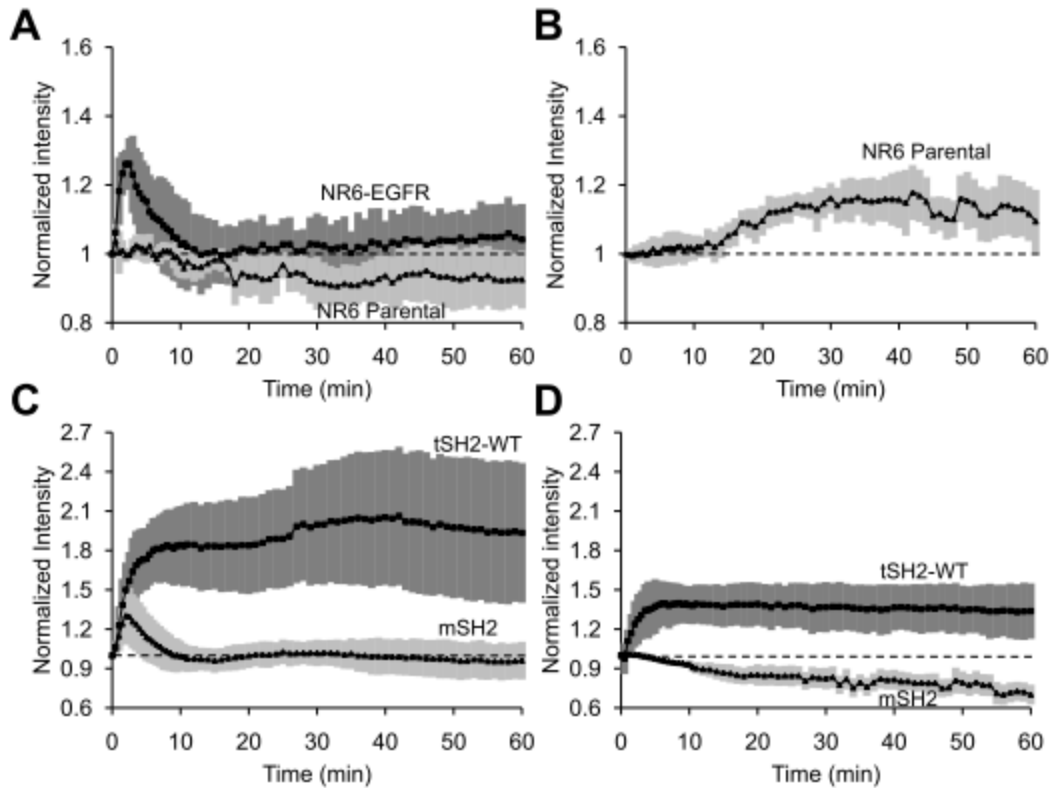


Figure 2.5: The mSH2 biosensor closely tracks expected kinetics of EGFR phosphorylation and displays increased specificity for the pY992 site of EGFR in live-cell imaging experiments. **(A)** Translocation of mSH2, measured by TIRF microscopy, in response to EGF stimulation of NR6 parental cells ($n = 7$) or NR6 cells overexpressing EGFR ($n = 6$). **(B)** mSH2 response to sodium orthovanadate treatment of NR6 ($n = 7$) parental cells. **(C)** Translocation responses of mSH2 and tSH2-WT in response to EGF stimulation of NR6 cells expressing c'1000 EGFR truncation mutant ($n = 5$). **(D)** Same as in (C) but with NR6 cells expressing c'1000F EGFR mutant with Y992F mutation ($n = 6$). Error bars indicating standard deviation are shaded in each figure.

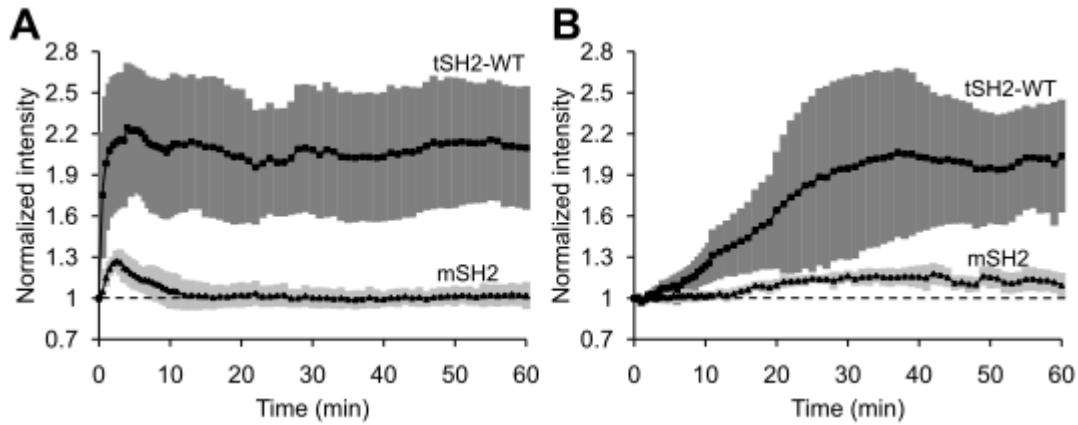


Figure 2.6: Normalized recruitment of tSH2-WT and mSH2 biosensors in response to (A) EGF treatment of EGFR-expressing NR6 cells (significantly different, $p \approx 10^{-170}$ by two-way ANOVA) and (B) sodium orthovanadate treatment of NR6 parental cells (significantly different, $p \approx 10^{-127}$ by two-way ANOVA).

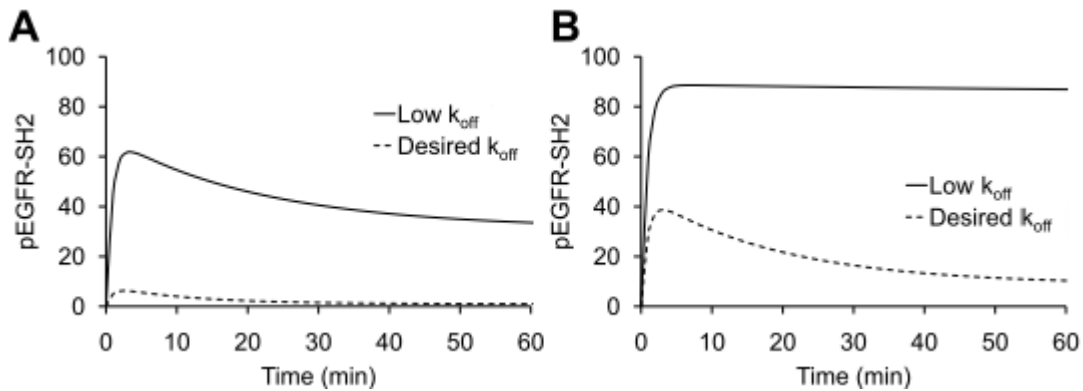


Figure 2.7: Effect of changing binding affinity of biosensor-target interaction or biosensor concentration on readout from the biosensor at (A) low and (B) high levels of biosensor expression at the k_{off} indicated

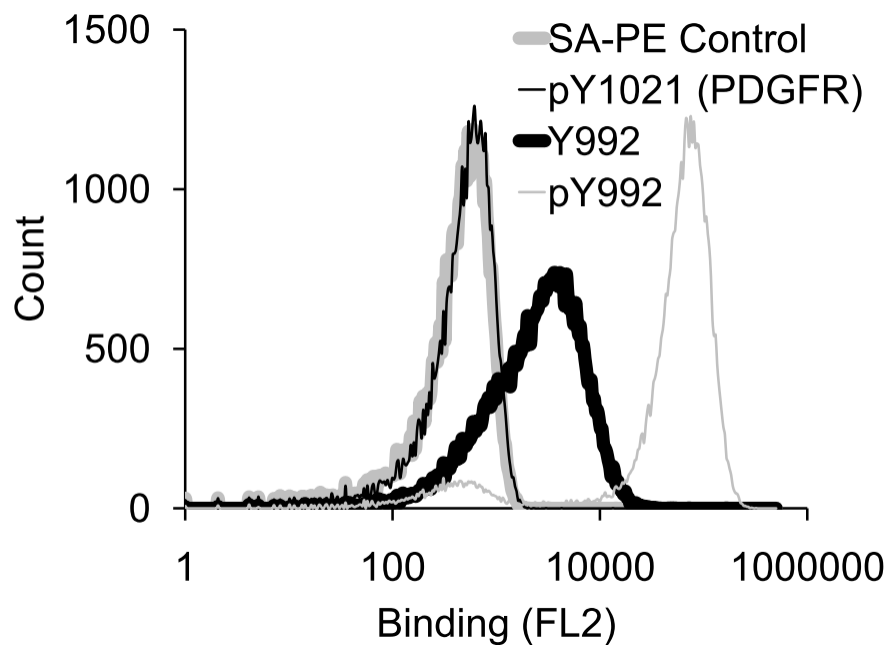


Figure 2.8: A mutant Sso7d (SPY992) with increased specificity for pY992 of EGFR. Yeast cells displaying SPY992 were labeled with biotinylated synthetic peptides corresponding to pY992, Y992 sites in EGFR and pY1021 in PDGFR, followed by secondary labeling with SA-PE. A control sample of yeast cells only labeled with SA-PE is also shown.

Table 2.1: Protein Sequences of engineered biosensors

| Protein | Sequence |
|---------|--|
| mSH2 | RLSEPVPQTNAHESKEWYHASLTRAQAEHMLMRVPRDGAFLVRKRNEP NSYAISFRAEGKIKHGRVQQEGQTVMLGNSEFDSLVDLISYYEKHPLYRK MKLRYPINEEALEKIGTAEPDYGA |
| SPY992 | MATVKFKYKGEEKQVDISKIGAVARLGKYIIFVYDLGGGKTGLGLVSEK DAPKELLQMLEKQKK |

Table 2.2: List of primers used

| Primer | Sequence (5'-3') |
|---------------|---|
| FP1 | CGCCTTTCAGAGCCTGTTCCA |
| RP1 | TGCCCCATAATCGGGTTCAGC |
| PCTF1 | AGTGGTGGTGGTGGTTCTGGTGGTGGTGGTTCTGGTGGTGGTGGTTCTG CTAGCCGCCTTTCAGAGCCTGTTCCA |
| PCTR1 | CTCGAGCTATTACAAGTCCTCTTCAGAAATAAGCTTTTGTTCCGGATCCT GCCCCATAATCGGGTTCAGC |
| FP2 | GGTCCGACTAGTCGCCTTTCAGAGCCTGTTCCA |
| RP2 | GGTCGTCTCGAGTGCCCCATAATCGGGTTCAGC |
| FP3 | GGTCCGACTAGTGGACTCAGATCTAGTGGCAGC |
| tdEF1 | TTTTGAATTCATGGTGAGCAAGGGCGAGGA |
| tdXR1 | CCCCCTCGAGTTACTTGTACAGCTCGTCCA |
| FP4 | TTTTGCGGCCGCATGCGCCTTTCAGAGCCTGTTCC |
| tdHR2 | CCCCTCTAGATTACTTGTACAGCTCGTCCA |

Table 2.3: List of peptides used for library screening. Amino acids are denoted by standard one letter code and (pY) denotes phosphotyrosine.

| Peptide | Notes | Sequence |
|---------|--|---|
| pY992 | Target sequence | ADE(pY)LIPQQGFF |
| Y992 | Phosphorylated and unphosphorylated peptides from EGFR used for negative selection | ADEYLIPQQGFF |
| pY1068 | | VPE(pY)INQSVPKR |
| Y1068 | | VPEYINQSVPKR |
| pY1086 | | NPV(pY)HNQPLNPA |
| Y1086 | | NPVYHNQPLNPA |
| pY1148 | | NPD(pY)QQDFFPKE |
| Y1148 | | NPDYQQDFFPKE |
| pY1173 | | NAE(pY)LRVAPQSS |
| Y1173 | | NAEYLRVAPQSS |
| RapGEF1 | | Phosphorylated peptides from other proteins used for negative selection |
| Dab1m | VEDPV(pY)QN | |
| Rasal2V | LSFQNPV(pY)HLNNPVP | |
| Rasal2I | LSFQNPV(pY)HLNNPIP | |

CHAPTER 3

ENGINEERING AFFINITY REAGENTS FOR EFFICIENT ISOLATION OF SECRETORY VESICLES

3.1 Introduction

Approximately 16% of the human genome is predicted to have a secreted protein product. Proteins secreted from the cell, like cytokines, growth factors and other signaling molecules play a crucial role in many physiological and pathological processes. Of particular interest here is hESC differentiation which is known to be significantly affected by the microenvironment (1, 2). In addition to the ECM used to grow hESCs the microenvironment consists of endogenously secreted cytokines and ECM. Characterizing the secretome can help identify the signaling pathways responsive to the microenvironment of hESCs leading to mechanistic insight into differentiation, especially hESC differentiation to trophoblast where such insight is lacking. Also of interest is the cancer secretome since cancer cells acquire their malignant phenotype through the manipulation of signaling processes involved in growth, proliferation, apoptosis, and angiogenesis. The secretory pathway, an important factor in cell-cell communication represents one such way for the acquisition of the malignant phenotype (3). So studying the microenvironment of cancer cells can provide insight into some of these mechanisms and also help in identification of potential biomarkers.

Current approaches to studying the secretome of hESCs focus on the use of conditioned medium. However the high protein content of serum in growth media can mask the low abundance secreted factors and cells grown in serum free media may be affected by the change in media potentially perturbing the biology of the cells. Similarly the secretome of cancer cells is obtained through the study of cancer cell lines, tumor proximal body fluids or plasma (4). While it is understood that the biology of these cell lines is often unrepresentative

of the cells they are derived from (5) they currently represent the only option. Since immortalized cell lines are likely to possess a malignant phenotype, they are not useful in studies of cancer progression and culture artifacts can likely render them unrepresentative of the primary cell physiology.

Secreted proteins usually contain a signal sequence that guides them to the endoplasmic reticulum (ER). They are transported to the cell surface from the ER through the Golgi packaged in vesicles known as secretory vesicles. Isolation of these secretory vesicles represents an alternative approach to obtain the secretome of cells. It has been demonstrated that the secretome of hESCs can be obtained by subcellular fractionation (6). While this method offers many advantages including the ability of obtaining temporal information about the secretome and obtaining the secretome of cells in co-cultures it is not suitable for isolation of secretory vesicles from tissues. Obtaining the secretome of tissues is ultimately most relevant since it bypasses the need to culture the cells potentially altering their biology and provides more physiologically relevant information. This is potentially very useful to validate the differentiated cells derived in case of controversy about their in vivo equivalent like in the case of trophoblast cells derived from hESCs (7). Additionally it represents a very interesting alternative to obtain cancer biomarkers.

Here we propose the engineering of binding protein that is capable of immunoprecipitating secretory vesicles that will allow us to probe the secretome of primary placental tissue, not otherwise possible. This will allow us to improve on the hESC-trophoblast model and also help in identifying cancer biomarkers. Since the proteome of

secretory vesicles is mostly unknown identifying a target is a challenge. Shusta et al., have previously demonstrated the use of lysates to isolate binders to unknown membrane proteins using yeast surface display (8, 9). This approach involves selectively biotinylating the target of interest. Their target of interest, membrane proteins, are biotinylated through biotinylation of the cell surface before cell lysis resulting in only the membrane proteins in the lysate being biotinylated. A modified version of this approach can be used to selectively biotinylate any organelle which can help develop affinity reagents to the desired organelle. Here we biotinylate the cell homogenate and fractionate using a sucrose gradient to obtain biotinylated organelles of interest that can be used to engineer the binder.

We demonstrate that the binder isolated, VB15 is capable of immunoprecipitating organelles from cells and tissues. Through electron microscopy of NIH3T3 cells we show that VB15 binds vesicles and that these vesicles often contain secreted proteins like fibronectin. The binder developed in this study can be used to obtain secretome of tissues which we plan to eventually demonstrate this through the isolation of secretory vesicles from third trimester human placenta, vCTB and STB isolated from first trimester human placenta and mouse placenta samples.

3.2 Materials and Methods

3.2.1 Construction of Sso6904 mutant library for yeast surface display

Eleven surface exposed residues were identified for randomization on the calcium binding protein from *Sulfolobus solfataricus* Sso6904 (10) using visual molecular dynamics

software (VMD, University of Illinois, Urbana-Champaign, IL). An ultramer was synthesized with DNA encoding the region between residues 19 and 83 with degenerate NNK codons at positions K42, I46, S50, L53, R57, K60, E61, D64, E68, Q71, and R72. An image highlighting these residues generated using UCSF Chimera (11) and the complete DNA sequence is shown in **Figure 3.1** and **Table 3.1** respectively.

Silent mutations were introduced into at positions 19, 20, 82 and 83 to introduce *HpaI* restriction sites into the protein using primers HPF1 and HPR1. This version with silent mutations was PCR amplified using primers HPYSDF1 and HPYSDR1 and cloned into yeast display vector pCTCON using the restriction sites *NheI* and *BamHI*. This construct was then linearized with *HpaI*. The ultramer was amplified using primers HPYSDF1 and HPYSDR1 to generate at least 50µg of double stranded DNA library of Sso6904 mutants. The PCR product and linearized vector were concentrated by using Pellet Paint™ (Novagen, San Diego, CA). A yeast display library was generated through homologous recombination as described in (12, 13). Briefly, 4µg of linearized vector and 12µg of PCR product were added to 400µl of electrocompetent yeast cells and electroporation was performed using a 2 mm electrode gap cuvette in a Bio-Rad Gene Pulser system (Bio-Rad, Hercules, CA) at 2.5kV and 25µF. Four such transformations were performed, and following a one hour outgrowth in 32 ml of 1:1 mix of 1M Sorbitol and YPD (10 g/l yeast extract, 20 g/l peptone, 20 g/l dextrose), the electroporated cells were expanded in 1L of SDCAA (20 g/l dextrose, 5 g/l Casamino acids, 6.7 g/l yeast nitrogen base, 5.40 g/l Na₂HPO₄, 7.45 g/l NaH₂PO₄). Simultaneously, dilutions of an aliquot were grown on SDCAA agar plates to estimate library

size. All primer were obtained from Integrated DNA Technologies (IDT; Coralville, IA), and sequences are listed in **Table 3.2**

3.2.2 Cell Culture

NIH3T3 and HEK293T cells were cultured at 37°C, 5% CO₂ in DMEM supplemented with 10% fetal bovine serum (FBS), and 1% PSG (L-glutamine and antibiotics penicillin and streptomycin). Drosophila S2 cells were cultured at 25°C in Schneider's medium with with 10% FBS, and 1% PS. All cell culture reagents were purchased from Life Technologies (Carlsbad, CA). Cells were passaged every 3 days or as needed, HEK293T and NIH3T3 cells using trypsin to dislodge them and semi adherent S2 cells by shaking/scraping.

3.2.3 Generation of target for negative and positive selection

The target sample for isolation of secretory vesicles was prepared using an isopycnic sucrose density gradient previously described (14, 15). Briefly NIH3T3, HEK293T or S2 cells were chilled on ice for 10 minutes and cells were harvested using cold 0.25% Trypsin and pelleted with a 5 minute cold centrifugation at 300g. The pellet was resuspended in a sucrose buffer (0.25M sucrose, 5mM MgCl₂, 25mM KCl, 10mM acetic acid (Thermo Fisher, Rockford, IL), 10mM triethanolamine (Sigma Aldrich, St. Louis, MO) and complete protease inhibitor tablets (Roche, Indianapolis, IN), pH 7.6) and homogenized with 5 strokes of pestle B. Nuclei were removed using 10 minute cold centrifugation at 1000g. The homogenate was then biotinylated using a Sulfo NHS LC biotinylation kit (Thermo scientific, Rockford, IL) according to standard protocols and the biotinylated homogenate

was layered on a sucrose gradient prepared in a 50ml conical tube. It was then centrifuged at 4800g for 20 hours at 4°C and 1ml fractions were collected by puncturing the bottom of the tube. A schematic of this is shown in **Figure 3.2** where the densest fractions (9ml, 1.2M – 1.6 M sucrose or 1.15 - 1.205g/cc) was designated as the Vesicle fraction and the remaining fractions of lower density designated as the Golgi fraction (11ml, 0.7M – 1.2M sucrose or 1.09 – 1.15g/cc) and Cytoplasmic fraction (<0.6M sucrose) were used for negative selections. The density of some common organelles (15–29) in these fractions is presented in **Table 3.3**.

3.2.4 Isolation of vesicle binding protein VB15 by magnetic selection and FACS

The yeast display library of Sso6904 mutants generated was grown in SGCAA (20 g/l galactose, 5 g/l Casamino acids, 6.7 g/l yeast nitrogen base, 5.40 g/l Na₂HPO₄, 7.45 g/l NaH₂PO₄) to induce surface protein expression. Binders to non-target species were eliminated by negative selection using magnetic sorting as described (19). Briefly biotin binder beads (Life Technologies, Carlsbad, CA) were loaded with biotinylated Golgi and Cytoplasmic fractions. These beads were incubated with ~10¹⁰ cells of the induced Sso6904 library for an hour, and the beads were removed using a magnet to deplete the library of non specific binders. Positive selection was performed by then incubating the unbound cells from the negative selection step with the Vesicle fraction for 30 minutes and then collecting the vesicle bound beads cells using magnetic biotin binder beads. Following three washes in 1 ml PBS-BSA (PBS + 0.1% BSA), the beads were collected and grown in SDCAA. These expanded cells were again grown in SGCAA to induce surface protein expression and FACS

was performed using a MoFlo cell sorter (Beckman Coulter Inc., Brea, CA). For flow cytometry approximately 5×10^7 induced cells were labeled individually with each of the biotinylated Golgi, Cytoplasmic and Vesicle fractions. The secondary antibody used was Streptavidin-PE (Life technologies, Carlsbad, CA). Simultaneously 5×10^7 induced cells were also labeled with an anti-c-myc chicken antibody (Life technologies, Carlsbad, CA) followed by a secondary labeling with goat-anti-chicken conjugated to Alexa-633 (Life technologies, Carlsbad, CA) to quantify surface protein expression. Six rounds of FACS were performed on the cells obtained after the magnetic selection step and the final pool of cells was grown on SDCAA plates. Yeast colonies were sequenced by Genewiz (La Jolla, CA). Further analysis of the sorted population and individual clones were performed using Accuri C6 analyzer (BD Biosciences, San Jose, CA).

3.2.5 Recombinant expression and purification of VB15

The final mutant selected for further characterization VB15 were cloned from the yeast display vector into the pET22 vector using the restriction sites *NdeI* and *XhoI* to generate His-tagged fusions. Primers NF1 and XR1 were used to introduce the restriction sites by PCR to facilitate cloning of VB15 and a Cysteine was introduced before the His tag to facilitate site specific biotinylation. Sequences were verified by colony sequencing using a T7 forward primer by Genewiz (La Jolla, CA). Subsequently, plasmids were introduced into Rosetta Cells™ (EMD Biosciences, San Diego, CA). Protein expression was induced by addition of 1mM IPTG followed by incubation at 20 °C for 12 hours with shaking at 250 rpm. After induction, cells were harvested by centrifugation (4800g, 10min), resuspended in

PBS with 0.2mM PMSF (Sigma Aldrich, St. Louis, MO) and lysed by sonication (2 sec ON, 5 sec OFF) for 7 minutes. The lysate was cleared by centrifugation and filtration through a 0.22 μ m filter and purified using a HiTrap 5ml column (GE Healthcare Bio-Sciences, Pittsburgh, PA) on a Bio-Rad (Hercules, CA) Biologic LP system using a linear imidazole gradient. The eluted fractions were analyzed by SDS PAGE and pure fractions were combined and dialyzed into 4L of PBS using a 3.5kDa MWCO filter (Thermo Fisher, Rockford, IL). Protein concentrations were determined by bicinchoninic acid (BCA) assay (Thermo scientific, Rockford, IL).

3.2.6 Far western blotting

Far western blotting was performed following standard protocols. Cells were lysed in lysis buffer containing protease and phosphatase inhibitors (50mM HEPES, 100mM NaCl, 10% Glycerol, 1% Triton X-100, 1mM sodium orthovanadate, 50mM β -glycerophosphate, 5mM NaF, 1mM EGTA, 10mM sodium pyrophosphate, 10 μ g/ml each of Aprotinin, Leupeptin, Pepstatin A and Chymostatin). Lysates from NIH3T3, HEK293T, and S2 cells were electrophoresed on a 10% SDS-PAGE gel and transferred to PVDF membranes using standard protocols. The membranes were blocked for one hour at room temperature in 5% nonfat dry milk in TBST. Primary incubation was performed overnight at 4°C using 1 μ M biotinylated VB15 in 5% nonfat dry milk in TBST. The membranes were then washed with TBST and Anti-biotin HRP (Cell Signaling, Danvers, MA) was used to probe the biotin on VB15 and SuperSignal® West Femto substrate (Thermo Fisher, Rockford, IL) was used to obtain a chemiluminescent readout.

3.2.7 Immunofluorescence

For immunofluorescence analysis, NIH3T3, HEK293T and S2 cells were passaged onto glass-bottom culture dishes (Greiner Bio-one, Monroe, NC). Cells were fixed with 4% paraformaldehyde and permeabilized with 0.1% Triton X-100. The permeabilized cells were blocked with PBS containing 5% BSA and 0.1% Triton X-100 and incubated overnight at 4°C with biotinylated VB15 alone or biotinylated VB15 and indicated primary antibody in the same blocking buffer. The cells were also labeled with mouse and rabbit isotypes corresponding to the antibodies along with unbiotinylated VB15 to ensure signal was specific. Following a wash secondary incubation was performed with Alexa 633-conjugated goat-anti-rabbit, goat-anti-mouse, or Streptavidin-PE (Life technologies, Carlsbad, CA) as appropriate, and imaged using a Zeiss LSM 710 confocal microscope.

3.2.8 Immunohistochemistry and histology analysis

Mouse embryos were fixed in and stored in 70% ethanol until paraffinization. Immunohistochemistry (IHC) and H&E staining using standard protocols were carried out at the Histology facility at NCSU. Labeling was performed using biotinylated VB15, unbiotinylated VB15 and biotinylated Sso6904. The same concentration of all proteins used for this labeling step to allow for qualitative comparison.

3.2.9 Transmission Electron Microscopy

NIH3T3 Cells grown in 24 well plates to about 80% confluence were fixed at room temperature for 20 minutes in 0.1M sodium cacodylate buffer (pH 7.4) containing 2% paraformaldehyde and 8% sucrose. Three 15 minute washes were performed in PBS

containing 0.05M glycine. The fixed cells were permeabilized with either 0.05% Saponin (Sigma Aldrich, St. Louis, MO) or 0.05% Triton X-100 in PBS2A (PBS with 0.2% BSA) at room temperature for 15 minutes. Immunogold labeling was adapted from Yi, et al. (30). Following two 15 minute washes primary incubation was performed with biotinylated VB15 and rabbit anti fibronectin antibody (Abcam, Cambridge, MA) diluted in PBS2A either individually or together overnight at 4°C. The cells were washed with in PBS2A four times for 10 minutes and blocked with 5% normal goat serum in PBS2A for 20 minutes followed by an 16 hour, 4°C incubation with goat anti-biotin IgG 1.4nm colloidal gold (Nanoprobes, Yaphank, NY) or Goat anti-rabbit IgG Fab' 1.4nm colloidal gold (Nanoprobes, Yaphank, NY). Three 10 minute washes were performed with PBS2A followed by three 5 minute washes in enhancement conditioning solution (ECS) (Aurion, Electron Microscopy Sciences, Fort Washington, PA). Enhancement was performed using R-Gent SE-EM silver enhancement reagent (Aurion, Electron Microscopy Sciences) for 90 minutes at room temperature. For double labeled samples the wash and labeling steps were repeated for the second secondary reagent and finally the samples were fixed in PBS containing 2% glutaraldehyde for 1 hour at room temperature. Finally a second enhancement silver step was performed for 60 minutes at room temperature and terminated using 3 washes in ECS.

Following enhancement the samples were processed for electron microscopy using standard protocols. Briefly cells were washed 3 times in PBS and incubated in 0.5% osmium tetroxide (Electron Microscopy Sciences, Fort Washington, PA). They were then washed in PBS and dehydrated in 30%, 50%, 75%, 100% and 100% ethanol for 10 minutes each. The

dehydrated cells were embedded in 100% PolyBed 812 resin (Polysciences, Inc., Warrington, PA) for 3 hours twice and allowed to polymerize for 24 hours at 60°C. 70nm blocks were sectioned using an Ultracut UCT ultramicrotome (Leica Microsystems, Buffalo Grove, IL) and diamond knife. The sections were mounted on 200mesh copper grids and double stained with 4% aqueous uranyl acetate for 5 minutes followed by Reynold's Lead Citrate for 8 minutes. Imaging was performed on a LEO EM-910 transmission electron microscope operating at 80kV (Carl Zeiss Microscopy, Thornwood, NY) and images were taken using a Gatan Orius SC1000 CCD camera with Digital Micrograph Software 3.11.0 (Gatan, Inc., Pleasanton, CA).

3.2.10 Human Placental Samples

First trimester human placentas (6-8 weeks) were obtained from women undergoing voluntary elective termination of pregnancy at North Carolina Memorial Hospital (UNC, Chapel Hill, NC) after obtaining approval from the Institution Review Board (IRB#12-2194) and from the patients. Placentas were handled using the protocol of Hunkapiller and Fisher, 2008 (31). Briefly, placentas were obtained and placed directly into cytowash medium. Villi were directly processed for either extraction of STBs and vCTBs or for isolation of secretory vesicles. The villi were dissociated using trypsin and the single cell suspension was homogenized as described in the target preparation section. vCTBs and STBs were isolated according to Steps 1-18 (31) and STB and vCTB cell suspensions were homogenized as described. Finally the homogenates of placenta, STB and vCTB were immunoprecipitated using biotin binder beads loaded with VB15 for isolation of secretory vesicles.

3.2.11 Mouse tissue samples

E13.5 pregnant CD1 mice were obtained from Charles River Laboratories and euthanized using CO₂. Uteri were extracted and placentas were manually isolated under the microscope (IACUC#12-010-B). Placentas were dissociated using trypsin and the single cell suspension was homogenized as described. Pancreas was isolated according to (32) and secretory vesicles were isolated from placenta and pancreas using biotin binder beads loaded with VB15 for immunoprecipitation.

3.3 Results

3.3.1 The engineered protein VB15 has increased specificity to fraction containing secretory vesicles

To identify a binder for isolation of secretory vesicles the yeast surface display library of $\sim 10^9$ Sso6904 mutants was screened using magnetic selection and fluorescence-activated cell sorting (FACS) (12). The population of cells obtained after magnetic selection and six rounds of FACS to enrich the best binding clones (enrichment along sort progression is shown in **Figure 3.3**) was plated and individual clones were analyzed. Sequencing of 30 clones revealed the only two unique clones which were further characterized for binding to each 2ml sucrose fraction (prepared as described before) and the best binder, VB15 was chosen for further analysis. We observe that VB15 exhibits remarkable selectivity for the first 5 fractions consisting the Vesicle fraction and shows no binding to any of the other fractions (shown in **Figure 3.4**). The sequence of VB15 compared to Sso6904 scaffold indicating the mutations generated using WebLogo (33) is also shown in **Figure 3.5**.

3.3.2 The engineered protein VB15 binds specifically to a single protein that is conserved across species

To further assess specificity of VB15, we conducted far-western blotting analysis. NIH3T3, HEK293T and S2 cell lysates were probed using biotinylated VB15 as primary reagent. We observe a single bright band in the lane corresponding to NIH3T3 lysate indicating that VB15 specifically binds a single protein in NIH3T3 cells (**Fig 3.6**). We also observe a band at the same molecular weight in lanes corresponding to HEK293T and S2 lysates indicating that the target protein is conserved across species. However we see other bands, albeit not as bright in the HEK293T and S2 lanes indicating either the presence of other isoforms or some non specific binding. It is likely that since NIH3T3 cell homogenates were used for majority of the sorts it is the most specific in this case. However since the binding to the target appears to be conserved across species (mouse, human and drosophila) VB15 can still be used to IP vesicles in all three cell lines. A far western blot was performed using biotinylated Sso6904 gave no specific band for any of the lysates indicating that the specificity was engineered into VB15 during the sorting process.

3.3.3 VB15 is capable of isolating organelles by immunoprecipitation

To evaluate the ability of VB15 to isolate organelles we performed an immunoprecipitation experiment using biotinylated VB15 loaded on beads against homogenates and lysates of NIH3T3 and S2 cells. The beads bound proteins and organelles were eluted in 8M urea and the eluate was electrophoresed and visualized by a silver stain using standard protocols. The results are shown in **Figure 3.7** and **Figure 3.8** respectively.

We see that in case of immunoprecipitation of lysate we obtain a single band similar to the far western blot. However in case of immunoprecipitation of homogenate we see many proteins indicating that an organelle containing proteins was isolated. This ability to immunoprecipitate not just proteins is extremely valuable for studies with isolation of vesicles. We observe the same behavior in both NIH3T3 and S2 cell lysates and homogenates further validating that VB15 binds a target that is conserved across species.

3.3.4 VB15 likely binds secretory vesicles

While flow cytometry results indicate VB15 binds selectively to fraction containing secretory vesicles it is possible since the fraction also contains other vesicles that it may bind other kinds of vesicles. To identify the organelle that VB15 binds to we performed electron microscopy experiments (shown in **Figure 3.9**) on NIH3T3 cells that were double labeled with VB15 and a secreted protein Fibronectin. We see that VB15 colocalizes to fibronectin containing vesicles and multivesicular bodies indicating that it likely binds secretory vesicles. However we also saw some non specific binding to lipid droplets which must be considered when analyzing mass spectrometry data for samples immunoprecipitated with VB15.

3.3.5 VB15 is capable of isolating organelles from tissues by immunoprecipitation

To evaluate the ability of VB15 to isolate organelles from tissues we performed an immunoprecipitation experiment using biotinylated VB15 loaded on beads against homogenates from mouse liver, pancreas and placenta obtained from E13.5 pregnant mice as described. As before the bead bound proteins and organelles were eluted in 8M urea and the eluates were visualized by silver staining (shown in **Figure 3.10**). We see that VB15 is able

to isolate organelles from all three organs (indicated by the presence of a smear of proteins) which are most likely contents secretory vesicles. Additionally we see that although the placenta had the least amount of total protein the IP from the placenta gave the highest signal indicating that the distribution of target protein is not likely uniform across all tissues. To further asses the distribution across the tissues IHC was performed using mouse embryos (shown in **Figure 3.11**) and we see that VB15 stains the entire embryo (while Sso6904 does not stain the embryo) but it is not uniformly distributed and some organs appear to have a much higher amount of target protein.

3.4 Discussion

The microenvironment of cells is critical in autocrine and paracrine communication and it is known that differentiation of hESCs is significantly affected by the cell microenvironment (1, 2). Further the secretory pathway is a key factor in cell-cell communication of cancer cells (3). So the cell secretome is desirable whether in identification of potential cancer biomarkers or to elucidate mechanistic information about hESC differentiation to trophoblast.

We constructed a combinatorial library through the introduction of degenerate NNK codons in a calcium binding protein from *Sulfolobus solfataricus*, Sso6904. This library was screened using yeast surface display to isolate a protein (VB15) that we then thoroughly characterized. We show that VB15 shows selective binding to subcellular fraction known to be enriched in secretory vesicles. We demonstrate by immunoprecipitation and western

blotting that VB15 binds a specific protein that is likely conserved across species and that it is capable of isolating organelles from cells and tissues. Through electron microscopy of NIH3T3 we show that VB15 binds vesicles. These vesicles are likely secretory vesicles since VB15 colocalized to fibronectin (a protein secreted by NIH3T3, mouse fibroblasts) containing vesicles and multivesicular bodies. Non specific binding to lipid droplets was observed however which needs to further investigation. The isopycnic density of lipid droplets and vesicles are significantly different and they are unlikely to be in the same sucrose fraction. It is possible that a surface protein may be shared between vesicles and lipid droplets.

In summary we demonstrate a protein engineering strategy for isolation of secretory vesicles. While a binder to secretory vesicles is derived in this study the approach is general and our approach serves as a blueprint for engineering affinity reagents for isolation of any organelle of interest even if the proteome of the organelle is unknown. Further work is ongoing to validate the isolation of secretory vesicles through the use of VB15 from tissue samples for e.g. human third trimester placenta samples.

Additional work is currently ongoing for the identification of the target protein using yeast two hybrid (Y2H) assays. A cDNA library of human proteins is expressed in a yeast cell using a Gal4 activation domain and this yeast library is mated with another yeast containing VB15 with a Gal4 binding domain. The zygotes containing proteins that bind VB15 are responsive to a Gal promoter and genes under Gal promoter are used in selection screens to identify the target protein.

3.5 References

1. S. C. Bendall, M. H. Stewart, P. Menendez, D. George, K. Vijayaragavan, T. Werbowetski-Ogilvie, V. Ramos-Mejia, A. Rouleau, J. Yang, M. Bossé, G. Lajoie, M. Bhatia, IGF and FGF cooperatively establish the regulatory stem cell niche of pluripotent human cells in vitro., *Nature* **448**, 1015–1021 (2007).
2. K. Onishi, P. D. Tonge, A. Nagy, P. W. Zandstra, Microenvironment-mediated reversion of epiblast stem cells by reactivation of repressed JAK-STAT signaling., *Integr. Biol. (Camb)*. **4**, 1367–76 (2012).
3. J. Zullo, K. Matsumoto, S. Xavier, B. Ratliff, M. S. Goligorsky, The cell secretome, a mediator of cell-to-cell communication, *Prostaglandins Other Lipid Mediat.* , 10–13 (2015).
4. T. B. M. Schaaij-Visser, M. De Wit, S. W. Lam, C. R. Jiménez, The cancer secretome, current status and opportunities in the lung, breast and colorectal cancer context, *Biochim. Biophys. Acta - Proteins Proteomics* **1834**, 2242–2258 (2013).
5. J. R. Masters, Human cancer cell lines: fact and fantasy., *Nat. Rev. Mol. Cell Biol.* **1**, 233–236 (2000).
6. P. Sarkar, S. M. Randall, D. C. Muddiman, B. M. Rao, Targeted proteomics of the secretory pathway reveals the secretome of mouse embryonic fibroblasts and human embryonic stem cells., *Mol. Cell. Proteomics* , 1829–1839 (2012).
7. A. S. Bernardo, T. Faial, L. Gardner, K. K. Niakan, D. Ortmann, C. E. Senner, E. M. Callery, M. W. Trotter, M. Hemberger, J. C. Smith, L. Bardwell, A. Moffett, R. a Pedersen, BRACHYURY and CDX2 mediate BMP-induced differentiation of human and mouse pluripotent stem cells into embryonic and extraembryonic lineages., *Cell Stem Cell* **9**, 144–55 (2011).

8. B. J. Tillotson, Y. K. Cho, E. V. Shusta, Cells and cell lysates: A direct approach for engineering antibodies against membrane proteins using yeast surface display, *Methods* **60**, 27–37 (2013).
9. Y. K. Cho, E. V. Shusta, Antibody library screens using detergent-solubilized mammalian cell lysates as antigen sources., *Protein Eng. Des. Sel.* **23**, 567–77 (2010).
10. Y. Feng, H. Yao, J. Wang, Solution structure and calcium binding of protein SSO6904 from the hyperthermophilic archaeon *Sulfolobus solfataricus*., *Proteins* **78**, 474–479 (2010).
11. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, UCSF Chimera - A visualization system for exploratory research and analysis, *J. Comput. Chem.* **25**, 1605–1612 (2004).
12. N. Gera, M. Hussain, B. M. Rao, Protein selection using yeast surface display, *Methods* **60**, 15–26 (2013).
13. L. Benatuil, J. M. Perez, J. Belk, C. M. Hsieh, An improved yeast transformation method for the generation of very large human antibody libraries, *Protein Eng. Des. Sel.* **23**, 155–159 (2010).
14. J. Stie, A. J. Jesaitis, Preparation of secretory vesicle-free plasma membranes by isopycnic sucrose gradient fractionation of neutrophils purified by the gelatin method., *Cytotechnology* **46**, 109–22 (2004).
15. G. Kippenberger, D. J. Palmer, M. Comer, J. Lipski, L. D. Burton, D. L. Christie, Localization of the noradrenaline transporter in rat adrenal medulla and PC12 cells: evidence for its association with secretory granules in PC12 cells., *J. Neurochem.* **73**, 1024–32 (1999).
16. P. Sarkar, T. S. Collier, S. M. Randall, D. C. Muddiman, B. M. Rao, The subcellular proteome of undifferentiated human embryonic stem cells, *Proteomics* **12**, 421–430 (2012).

17. J. R. Yates, A. Gilchrist, K. E. Howell, J. J. M. Bergeron, Proteomics of organelles and large cellular structures., *Nat. Rev. Mol. Cell Biol.* **6**, 702–14 (2005).
18. M. E. G. de Araújo, L. A. Huber, T. Stasyk, Isolation of endocytic organelles by density gradient centrifugation., *Methods Mol. Biol.* **424**, 317–31 (2008).
19. B. Brugger, W. Nickel, F. T. Wieland, Molecular mechanisms of COPI vesicle formation, *Protein Modul. Cell. Signal.* (2001).
20. T. Kobayashi, M.-H. Beuchat, J. Chevallier, A. Makino, N. Mayran, J.-M. Escola, C. Lebrand, P. Cosson, T. Kobayashi, J. Gruenberg, Separation and characterization of late endosomal membrane domains., *J. Biol. Chem.* **277**, 32157–64 (2002).
21. J. Stie, A. J. Jesaitis, Preparation of secretory vesicle-free plasma membranes by isopycnic sucrose gradient fractionation of neutrophils purified by the gelatin method, *Cytotechnology* **46**, 109–122 (2004).
22. N. C. Walworth, P. J. Novick, Purification and characterization of constitutive secretory vesicles from yeast., *J. Cell Biol.* **105**, 163–74 (1987).
23. A. Barthel, W. Nickel, C. Tonko, H. D. Söling, Sorting and budding of constitutive secretory vesicles in hepatocytes and hepatoma cells, *Adv. Enzyme Regul.* **35**, 283–288 (1995).
24. R. Beck, Z. Sun, F. Adolf, C. Rutz, J. Bassler, K. Wild, I. Sinning, E. Hurt, B. Brügger, J. Béthune, F. Wieland, Membrane curvature induced by Arf1-GTP is essential for vesicle formation., *Proc. Natl. Acad. Sci. U. S. A.* **105**, 11731–6 (2008).
25. P. Anderson, N. Kedersha, RNA granules., *J. Cell Biol.* **172**, 803–8 (2006).

26. I. Fialka, Identification of Syntenin as a Protein of the Apical Early Endocytic Compartment in Madin-Darby Canine Kidney Cells, *J. Biol. Chem.* **274**, 26233–26239 (1999).
27. K. P. Figueroa, S. M. Pulst, Identification and expression of the gene for human ataxin-2-related protein on chromosome 16, *Exp. Neurol.* **184**, 669–678 (2003).
28. K. Heimann, Specific Isoforms of Actin-binding Proteins on Distinct Populations of Golgi-derived Vesicles, *J. Biol. Chem.* **274**, 10743–10750 (1999).
29. J. Malsam, D. Gommel, F. T. Wieland, W. Nickel, A role for ADP ribosylation factor in the control of cargo uptake during COPI-coated vesicle biogenesis, *FEBS Lett.* **462**, 267–272 (1999).
30. H. Yi, J. Leunissen, G. Shi, C. Gutekunst, S. Hersch, A novel procedure for pre-embedding double immunogold-silver labeling at the ultrastructural level., *J. Histochem. Cytochem.* **49**, 279–284 (2001).
31. N. M. Hunkapiller, S. J. Fisher, Chapter 12 Placental Remodeling of the Uterine Vasculature, *Methods Enzymol.* **445**, 281–302 (2008).
32. A. C. Schloithe, C. Woods, G. T. P. Saccone, METHODS PAGE An isolated rat pancreas preparation for studying pancreatic spinal mechanosensitive and chemosensitive afferent activity Ann C Schloithe (ACMLS), CM Woods (BBioTech, PhD) and Gino T.P. Saccone (BSc (Hons), PhD), (2011), doi:10.3998/panc.2011.13.
33. G. E. Crooks, G. Hon, J. M. Chandonia, S. E. Brenner, WebLogo: A sequence logo generator, *Genome Res.* **14**, 1188–1190 (2004).

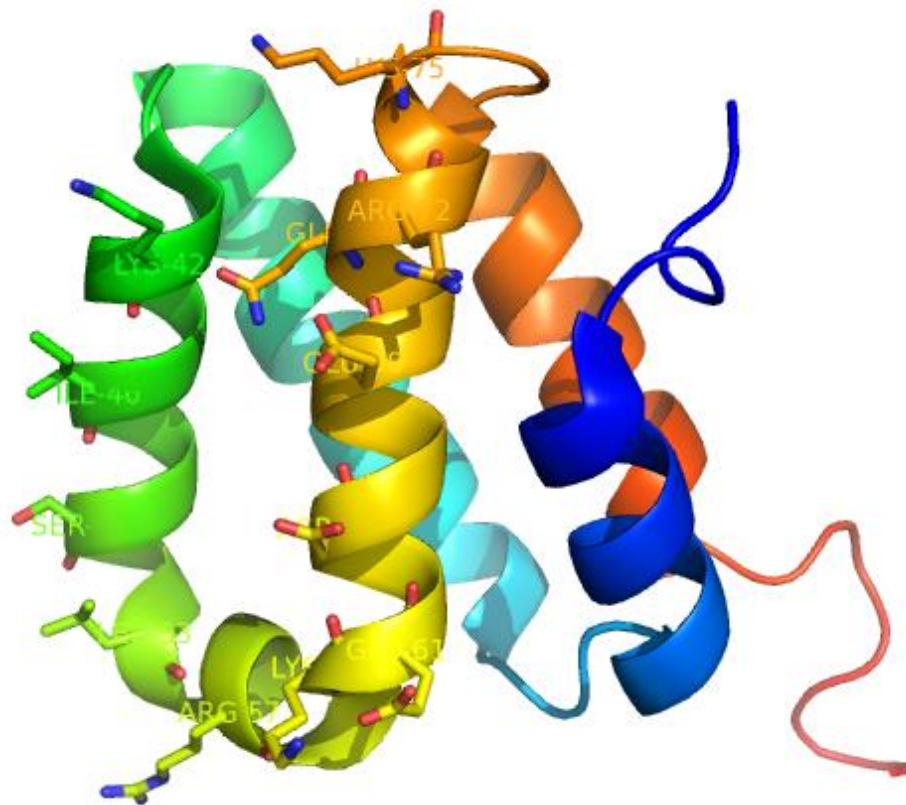


Figure 3.1: Mutated residues on Sso6904 (K42, I46, S50, L53, R57, K60, E61, D64, E68, Q71, and R72) used in the generation of the library shown using a licorice representation (PDB ID: 2KJG, Image generated using UCSF Chimera)

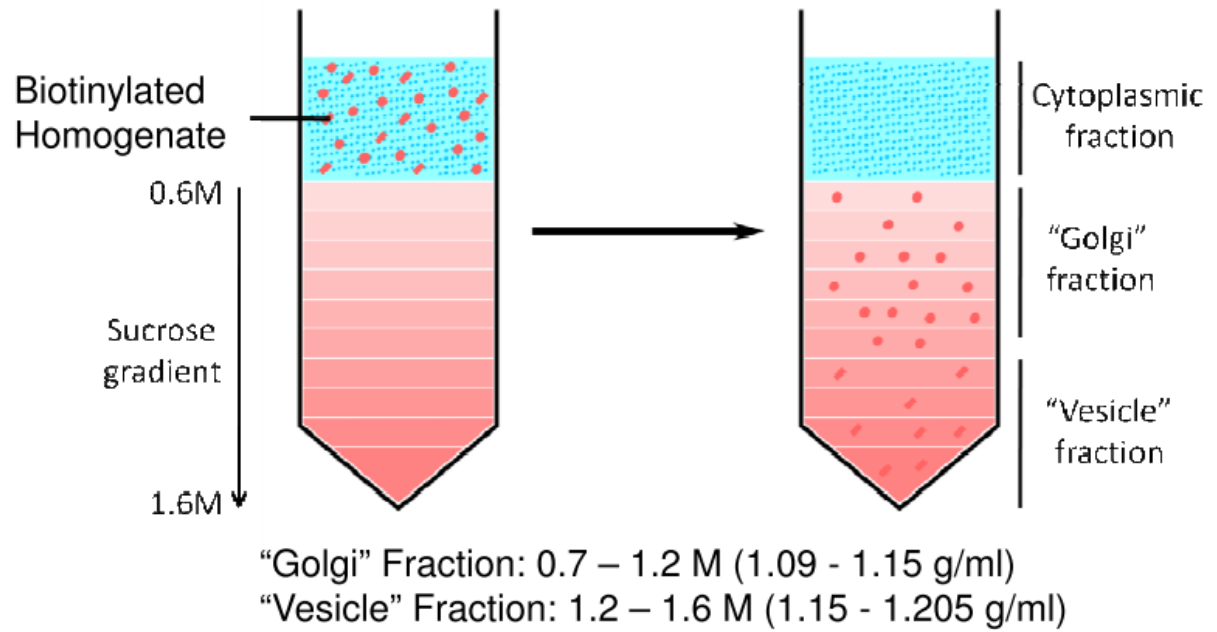


Figure 3.2: Subcellular Fractionation using Sucrose Density Gradient Centrifugation used for the isolation of fractions used for positive and negative selection in the isolation of a binding protein to secretory vesicles

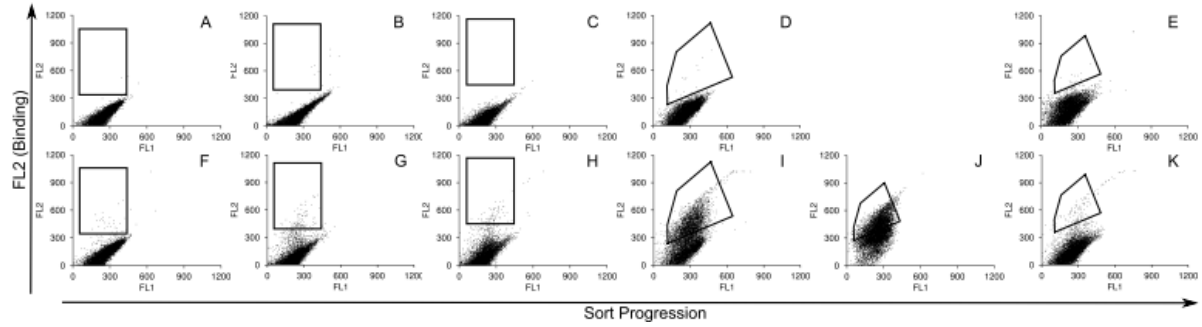


Figure 3.3: Panels A-E show the enrichment in binding to the negative control fraction (Golgi fraction). Panels F-K show enrichment in binding to the desired target (Vesicle fraction). Binding was assayed with Streptavidin-PE and the sort gate for cells sorted from the Vesicle fraction is also shown in each panel. For the last sort (Panel K) 10x (v/v) unbiotinylated homogenate was added to the biotinylated vesicle fraction to reduce non specific binding and potentially eliminate non low affinity binders.

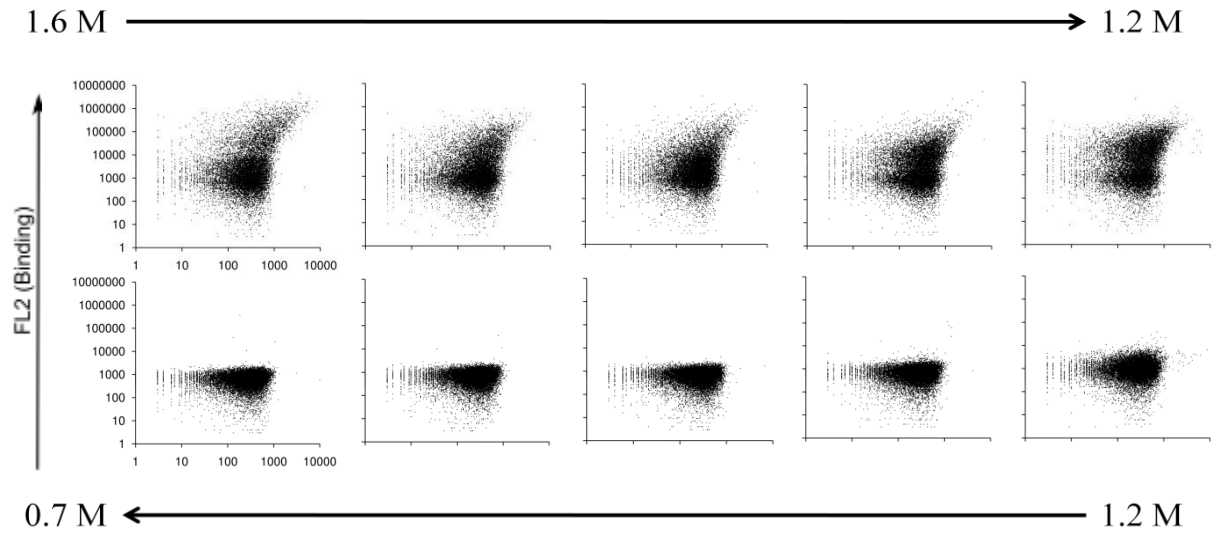


Figure 3.4: Panels in the top row show the binding of VB15 to the Vesicle fractions and panels in the bottom row show binding to the Golgi fractions. The molar concentration of sucrose gradient is as shown. Binding is assayed using Streptavidin-PE. We see that VB15 selectively binds the contents of the biotinylated Vesicle fraction but not the biotinylated Golgi fraction.



Figure 3.5: Comparison of sequence of VB15 and the Sso6904 scaffold. The sequence of the VB15 mutant isolated from the library is shown below the sequence of Sso6904 at the mutated positions (K42, I46, S50, L53, R57, K60, E61, D64, E68, Q71, and R72 on WT). The image was generated using WebLogo

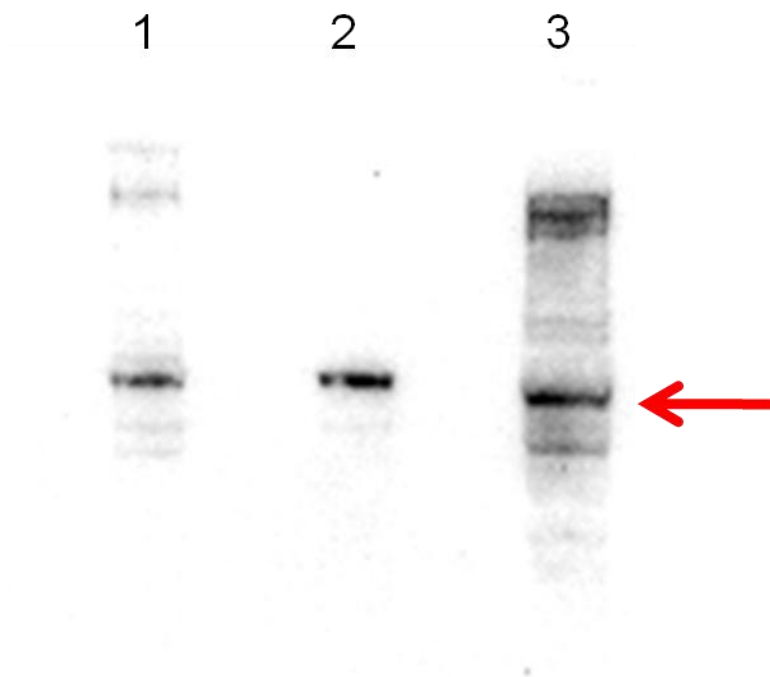


Figure 3.6: Far Western blot of cell lysates using biotinylated VB15 as primary reagent. Lanes 1, 2 and 3 were loaded with lysates of S2 (drosophila), NIH3T3 (mouse) and HEK293T (human) cell lines respectively. The band corresponding to the putative target of VB15 is indicated. The epitope of VB15 appears to be conserved on the target protein across species

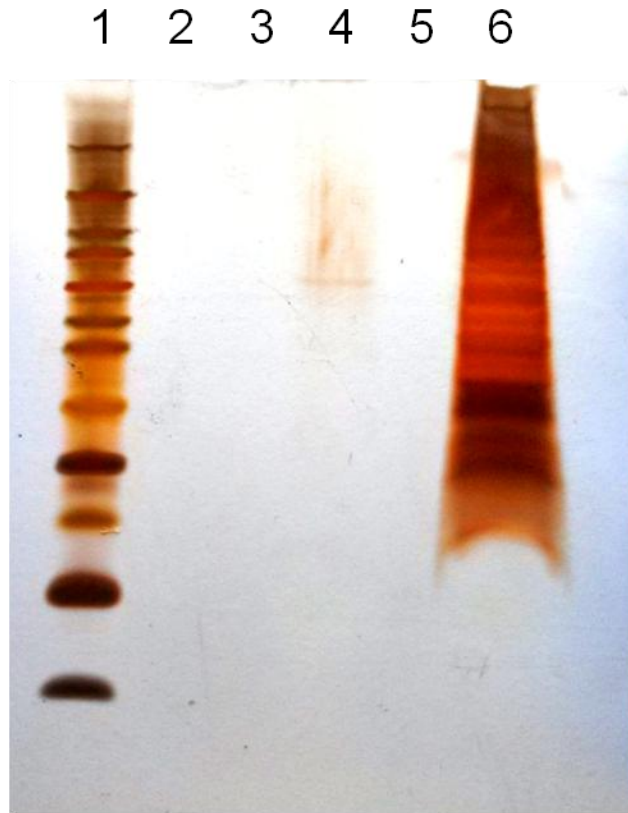


Figure 3.7: SDS PAGE gel of NIH3T3 lysate (lane 4) and homogenate (lane 6) immunoprecipitated using biotinylated VB15 loaded on beads and visualized using a silver stain. The lysate gives a single band indicating VB15 is specific and the homogenate gives lots of proteins indicating VB15 may be capable of immunoprecipitating organelles.

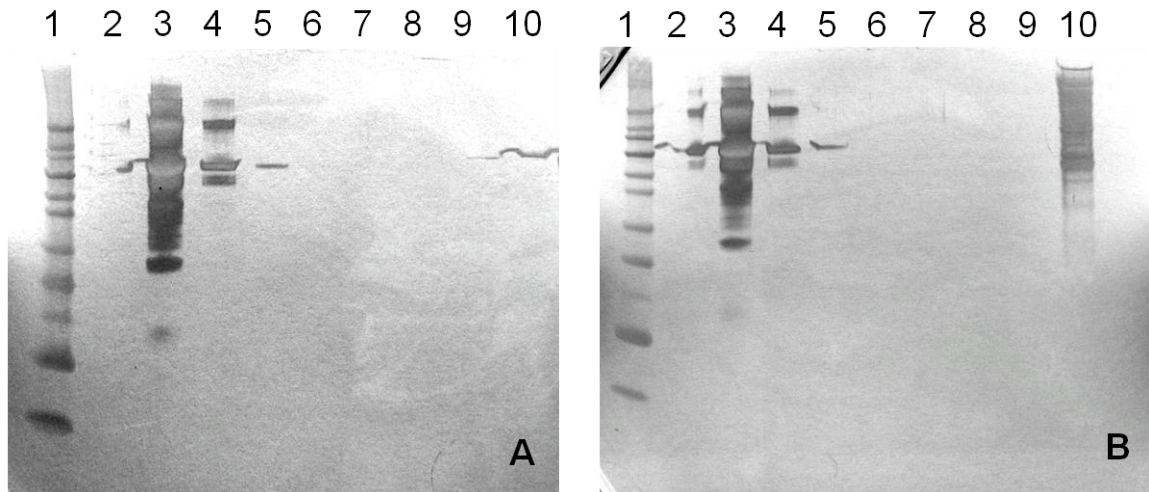


Figure 3.8: SDS PAGE gel of S2 lysate (panel A) and homogenate (panel B) immunoprecipitated using biotinylated VB15 loaded on beads and visualized using a silver stain. The lysate gives a single band (lane 10, A) indicating VB15 is specific and the homogenate gives lots of proteins (lane 10, B) indicating VB15 may be capable of immunoprecipitating organelles. Lane 1 is protein ladder and lanes 3-8 are BSA dilutions used to estimate protein content in the immunoprecipitated eluate

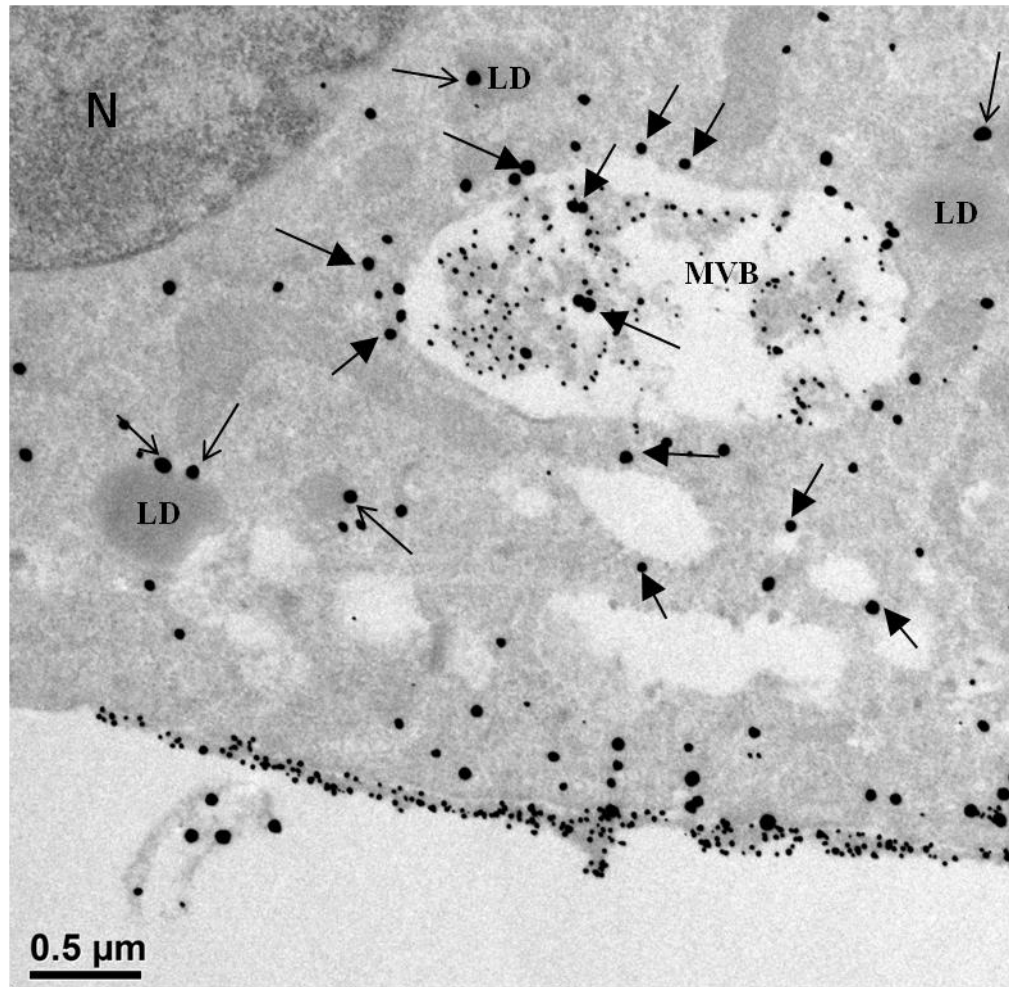


Figure 3.9: TEM image of a section showing colocalization of Fibronectin and VB15 to vesicles and (MVB) multivesicular bodies (indicated by solid arrows) and non specific binding to some (LD) lipid droplets (arrows). They are also seen colocalized at the cell periphery indicating that VB15 potentially binds secretory vesicles that dock at the plasma membrane.

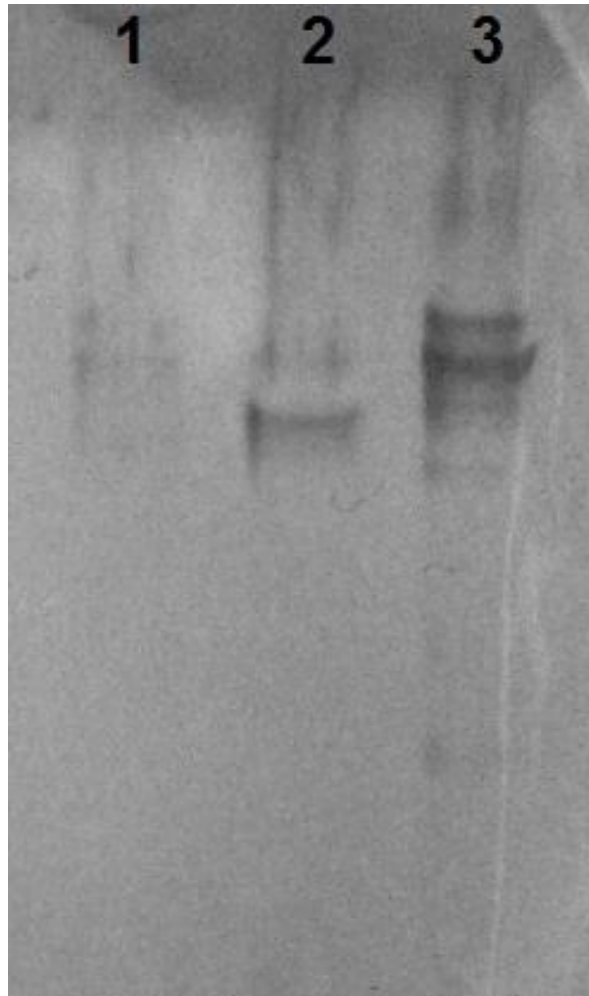


Figure 3.10: Immunoprecipitation of Liver (1), Pancreas (2) and Placenta (3) tissues from E13.5 pregnant mice with VB15 visualized using a silver stain

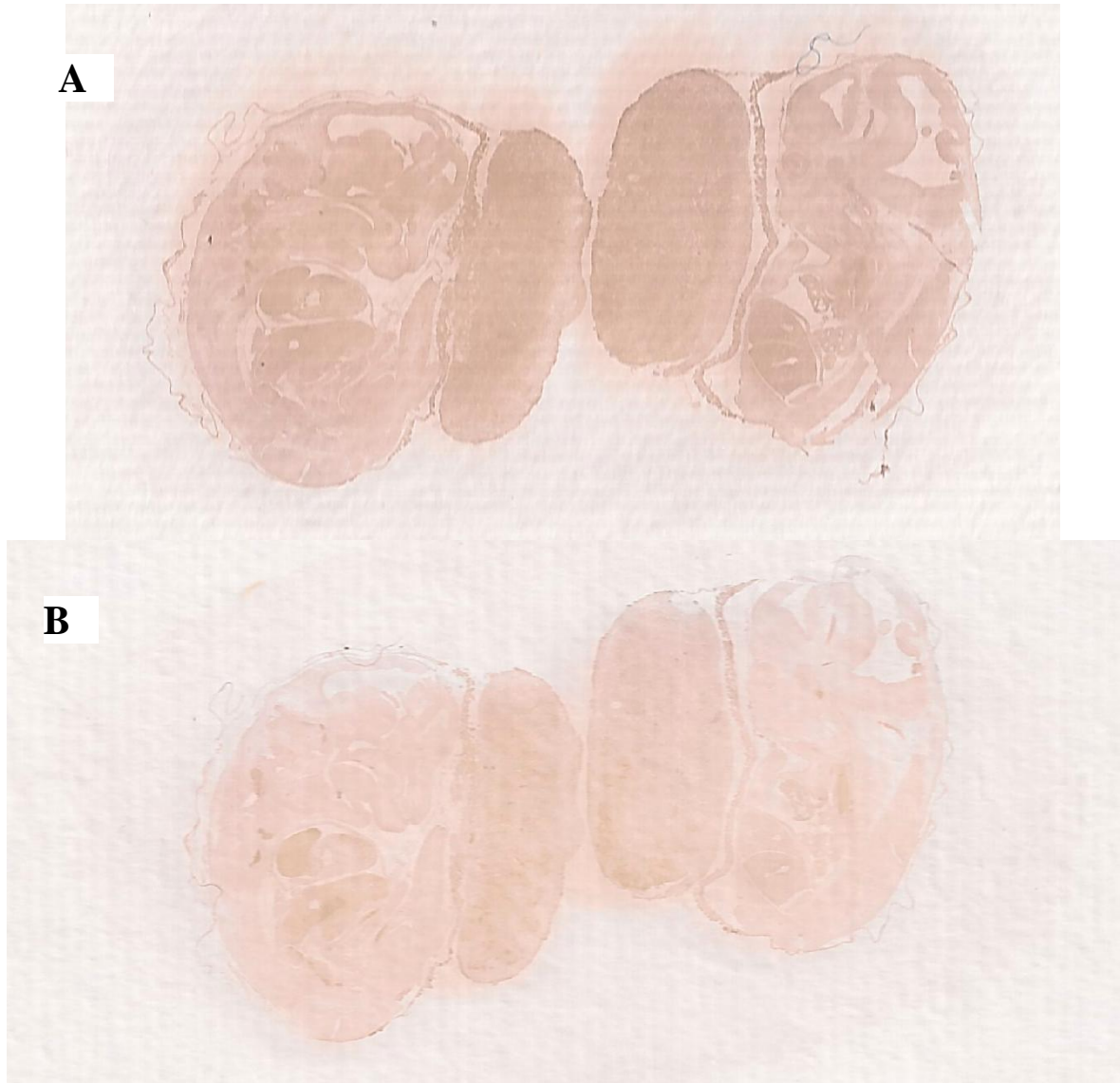


Figure 3.11: IHC of mouse embryos performed with (A) biotinylated VB15 and (B) biotinylated Sso6904 using Streptavidin-PE as secondary. We see that while VB15 stains the entire embryo (dark brown) Sso6904 does not

Table 3.1: DNA Sequence of Sso6904 with mutated (NNK) residues

| | Sequence |
|-----------------|---|
| Sso6904 Library | ATGTCAATATTAGAAGATCCAGAATTTGTAAAATTAAGACAATTTAA AGGTAAAGTAAATTTCAATTTAGTTATGCAGATACTGGATGAGATAG AACTTGATCTAAGGGGAAGTGATAATATCANNKACATCTATANNKTAT GTATATNNKAGCCATNNKGATGAGATANNKAAAAATNNKNNKTTCT ATNNKATGATTGCANNKATACTANNKNNKTATTACAAAAAATAGG CATAGAGAATGTGAATCAGTTGATACTAACTACTATAAAA |

Table 3.2: List of primers used

| Primer | Sequence (5'-3') |
|-------------|--|
| HPF1 | CGCAACGGCGTAACTTCAATTTAGTTATGCAGATACT |
| HPR1 | GCAAGCTGGGTAAACATTCTCTATGCCTATTTTTTTGTAA |
| HPYSD F1 | TCTGCTAGCATGTCAATATTAGAAGATCCAGAATTTGTAAAATTAAGACA ATTTAAAGGTAAAGTAACTTCAATTTAGTTATGC |
| HPYSD F1 | TTCAGAAATAAGCTTTTGTTCGGATCCTTTTATAGTAGTTAGTATCAACTG GTTAACATTCTCTATGCCTA |
| NF1 | GGGCATATGATGTCAATATTAGAAGATCCAGAA |
| XR1 | GGGCTCGAGACATTTTATAGTAGTTAGTATCAACTGA |

Table 3.3: Isopycnic density of some subcellular organelles

| Organelle | <u>Density (g/cm³)</u> |
|--------------------------|--|
| Nuclei | >1.30 |
| Mitochondria | 1.17 – 1.21 |
| Lysosomes | 1.20 – 1.22 |
| Peroxisomes | 1.23 |
| ER vesicles (microsomes) | 1.10 – 1.15 (smooth ER), 1.17 – 1.22 (rough ER) |
| Golgi Stacks | 1.12 – 1.15 |
| Golgi Vesicles | 1.10 – 1.13 |
| Plasma membrane sheets | 1.15 – 1.19 |
| Plasma membrane vesicles | 1.17 – 1.20 |
| COPI Vesicles | 1.16 – 1.19 |
| Early endosomes | 1.10 – 1.15 |
| Late endosomes | 1.12 – 1.15 |
| COPII Vesicles | 1.15 – 1.18 |

CHAPTER 4

TROPHOBLAST DIFFERENTIATION OF HUMAN EMBRYONIC STEM CELLS

(Adapted from Tiruthani et al. 2013, *Biotechnology Journal*, 8(4): 421–433)

4.1 Introduction

Molecular mechanisms regulating human trophoblast differentiation remain poorly understood due to difficulties in obtaining primary tissues from very early developmental stages in humans. Therefore the use of human embryonic stem cells (hESCs) as a source for generating trophoblast tissues is of significant interest. However, there is controversy over whether hESC-derived cells are indeed analogous to true trophoblasts found *in vivo*. One of the major challenges is that current studies performed on cultured trophoblast cells may not be representative of true biology and clarification or markers for identification is desirable especially to improve on the hESC-trophoblast model. Isolation of secretory vesicles through the use of VB15 (binding protein described in chapter 3) from human first trimester placenta samples, vCTB and STB from first trimester placenta and human third trimester placenta samples will greatly help address some of the issues presented herein.

Trophoblast formation is a major process underlying implantation of the blastocyst in the endometrium and development of the placenta. The cells of the trophoblast (TE) are precursors of all trophoblast cell types in the placenta. These include the villous cytotrophoblasts (vCTB), syncytiotrophoblast (STB) and extravillous cytotrophoblasts (evCTB), with vCTBs being multipotent progenitors that can differentiate to the other trophoblast cell types (1). Abnormalities in the trophoblast are associated with problems such as recurrent loss of pregnancy and preeclampsia (2–5). Yet, the study of trophoblast formation and differentiation is severely impeded by the scarce availability of primary trophoblastic tissues from very early developmental stages. Primary trophoblasts are difficult

to isolate and cannot be maintained in cell culture for long (6–15). Further, these primary cells exhibit gestational variability (16, 17), are already lineage committed (18, 19) and may also be adapted to the maternal environment (20, 21). Alternatively, studies have been carried out using carcinoma-derived cell lines and these results are combined with knowledge from other species like mice to understand trophoblast development. However, these cell lines may not be the best models for studying trophoblast development (22–25). Also, implantation and placental development in humans is different from that in other species (26) and other animal models may not be accurate (27). The isolation of multipotent trophoblast stem (TS) cells that can be maintained in cell culture and can give rise to all the differentiated trophoblast cell types – STB and evCTBs – is of immense value for studying human trophoblast development. However, unlike in mouse, the human TS cell is yet to be derived. Notably, multipotent vCTBs isolated from primary placentae cannot be sustained in long term cell culture (28). In the light of these difficulties, the use of human embryonic stem cells (hESCs) for deriving trophoblast tissues has generated significant interest. However, whether bona fide trophoblasts can indeed be derived from hESCs remains controversial. . Pertinent issues include the choice of markers used for identifying trophoblast cell types obtained from hESCs and the epigenetic state of these differentiated cells. Herein, we comprehensively review previously reported efforts to obtain trophoblasts from hESCs and discuss the merits as well as limitations of using hESCs as a source for trophoblast derivatives.

4.2 HESC differentiation to trophoblast

Differentiation of hESCs to trophoblast-like cells has been achieved using protocols that can be broadly classified into two categories: (i) treatment of hESCs in monolayer culture with Bone Morphogenetic Protein-4 (BMP4) or inhibitors of Activin/Nodal/Transforming Growth Factor- β (TGF β) signaling, and (ii) isolation of trophoblast-like cells from embryoid bodies (EBs). **Table 4.1** provides a summary of specific experimental conditions used for trophoblast differentiation of hESCs.

4.2.1 Monolayer cultures using BMP4 and/or Activin/Nodal/TGF β inhibition

Differentiation of hESCs to trophoblast-like cells was first reported by Xu et al (29) through addition of BMP4 to hESCs cultured in medium conditioned by mouse embryonic fibroblasts (MEFs; MEF-CM) (30). Differentiation by BMP4 addition is described as a wave of differentiation proceeding from the edges of colonies inwards (29, 31, 32). The primary results were the upregulation of trophoblast markers, the secretion of placental hormones and the formation of multinucleate STBs. Since then, other studies have reported similar results with the use of differing amounts of BMP4, different cell lines, other BMP ligands and differences in how often the media was refreshed (31–37). Similar studies using BMP have also been performed with different culture media as well as with inhibitors to Activin/Nodal/TGF- β and Fibroblast Growth Factor (FGF) signaling (38–40). Das et al. performed studies at two different oxygen levels and reported that oxygen accelerates this differentiation process (39). Xu et al also found that addition of TGF- β signaling inhibitors

like SB431542 increased the levels of pSMAD1/5/8 and BMP4 and that addition of FGF inhibitors like SU5402 reduced the levels of pSMAD2/3. Both increased levels of pSMAD1/5/8 and reduced levels of pSMAD2/3 aid the differentiation of hESC to trophoblast-like cells (41). Sudheer et al found that BMP4 signaling in the absence of FGF2 leads to CG α secreting syncytiotrophoblast with reduced CDX2 expression compared to the presence of FGF2 (42). Differentiation of hESC to trophoblast-like cells by BMP4 addition has also been performed in induced pluripotent stem cells (iPSCs) (35, 36).

Recently, studies have sought to further characterize trophoblast-like cells obtained through BMP treatment of hESCs and have found upregulation of mesoderm markers in these cells (43). It was found that these cells do not share certain vCTB markers such as the lack of HLA Class I antigens, the expression of ELF5 (44) and the hypomethylation of the ELF5 promoter locus. Additionally, given that these BMP-derived trophoblast-like cells are precursors for STBs, it may be possible that they may not share vCTB markers. Indeed, ELF5 expression in the trophoderm of the human embryo remains unknown, and the syncytium of the human placenta does not stain for ELF5 (44). ELF5 null mice also continue to form placental syncytia (45). Nevertheless, syncytial cells derived by BMP-treatment of hESCs may provide an important model system for studying cell fusion. Notably, all the other syncytia of the human body such as skeletal muscle, cardiac muscle and osteoclasts are mesoderm derivatives. It may be hypothesized that a conserved mechanism may exist for cell fusion and may be controlled by mesoderm-associated genes. A thorough evaluation of the

BMP-treatment-derived syncytial cells is needed for characterizing BMP-mediated trophoblast differentiation of hESCs. .

4.2.2 Differentiation using embryoid bodies

Differentiation of hESCs to trophoblast-like cells has also been carried out by Gerami-Naini et al taking advantage of spontaneous differentiation of embryoid bodies (EBs) into all lineages (46). Enrichment of trophoblast-like cells from this mixed population was achieved using different techniques such as changing media, embedding cells in matrigel drops, changing the length of culture (46), selecting for levels of CGB secretion (47–49) or cell adhesion (50). Gerami-Naini et al. observed that embedding the EBs in Matrigel drops results in higher HCG levels than growing them in suspension culture (46). Giakoumopoulos et al generated EBs in combination with either term placental fibroblasts (TPF) or dermal fibroblast from fetal foreskin (CL2F). They observed that while both the combination EBs have similar number of trophoblast cells as control EBs the TPF-EBs have much higher HCG secretion compared to control EBs (51). Additionally the size of colony seems to affect the homogeneity of differentiated cells (34). Using an EB-based approach, Harun et al. derived multipotent vCTB-like cells that can be maintained in a proliferative state for longer than 30 passages (47). Notably, neither multipotency nor proliferative capacity has been demonstrated in any of the BMP4 derived trophoblast-like cells. Udayashankar et al observed that cells at lower passages are KRT7, Ki67 positive whereas those at higher passages have reduced levels of KRT7, Ki67 but have high levels of HLA-G, CD9, MMP2, MMP9, HSD3B1 and are capable of invading matrigel or endometrial tissue (49) which suggests

potential multipotency. Along similar lines, differentiation of iPSC to trophoblast-like cells using embryoid bodies has also been reported (52). However, a significant disadvantage of using EBs is the heterogeneity/variability seen in all these cases. More importantly, these cells continue to express HLA-A, B and C antigens (47) and do not express the vCTB marker ELF5, thereby raising concerns that they may not be similar to vCTBs. Both these approaches to differentiation are further reviewed in (53–59).

4.3 Markers used for characterizing hESC-derived cells

Mouse embryonic stem cells (mESCs) contribute poorly to the TE when injected into mouse blastocysts (60, 61). MESC can be differentiated to the trophoblast lineage only through genetic manipulations such as by OCT4 knockdown (62–64), SOX2 knockdown (65) and overexpression of CDX2 (66, 67). Similar studies in hESCs by knockdown of OCT4 (68, 69) or NANOG (70) find increased expression of trophoblast associated genes such as *CGA*, *CDX2* but also markers for other lineages. It was further found that the proximal promoter locus of *Elf5* gene is differentially methylated in mESCs and mouse TS cells (44, 71). This has indicated the presence of an epigenetic barrier between ESCs and TE in mice (72, 73). These results have raised concerns that the epigenetic barrier between ESCs and TE found in mice may be conserved in humans. Thus, it raises the possibility that hESCs may not have the ability to form true trophoblasts. Indeed, Bernardo et al. found that *T* expression is upregulated by BMP4 treatment and BMP-treated hESCs lack ELF5 expression; they suggest that BMP treatment of hESCs gives extraembryonic mesoderm and not trophoblast (43). In this context, it is important to note that the proper choice of markers for identifying hESC-

derived cells as trophoblasts remains a challenging issue. HLA class I expression in trophoblast has been studied and it was observed that HLA-C and HLA-G was present in evCTBs and vCTBs were negative for HLA-A, B, C (74, 75). The HLA profile, in conjunction with the expression of other markers like CDX2, has been used in the identification and isolation of trophoblast-like cells only in a few cases (76). For example, CDX2 was assayed at the mRNA level in (34, 36, 38, 40, 42, 47, 49, 50, 69, 77–79) and at the protein level in (34, 40). However, expression of ELF5 has been demonstrated neither at the mRNA level nor at the protein level. HLA-A was assayed for at the mRNA level in (29, 47) and at the protein level in (69). HLA-G was assayed for at the mRNA level in (29, 33, 34, 42, 47, 80) and at the protein level in (39, 42, 47, 49, 79). Moreover, the *ELF5* promoter has not been assayed as being hypomethylated in any of these reports. Therefore, re-evaluation of previously reported studies through rigorous characterization of pertinent trophoblast markers such as CDX2, ELF5, HLA expression profile and methylation status of the ELF5 promoter in hESC-derived trophoblast-like cells will be immensely valuable

4.4 Role of differentiation media

It is striking to note that protocols used for trophoblast differentiation of hESCs have several differences in media composition such as the base culture medium used, and concentrations and types of cytokines or small molecule inhibitors included (**Table 4.1**). As discussed earlier, BMP has been a key component of the differentiation medium used. To understand the molecular mechanisms of trophoblast-like differentiation of hESCs, knockout studies using modified cell lines have been performed (77–79). In one study, Chen et al.

suggest that BMP signaling is necessary for trophoblast differentiation of hESCs [77]. Wu et al. further suggest that BMP signaling induced by blocking Activin/Nodal signaling alone is sufficient for trophoblast differentiation (38, 77). Lipchina et al. also found that BMP4 signaling is necessary through the regulation of BMP signaling by miR-302/367 (81).

In addition to BMP, FGF signaling seems to play an important role in trophoblast differentiation of hESCs. Wu et al. observed that addition of SB431542 represses FGF signaling (38). Greber et al. compared differentiation in MEF-CM with that in N2B27 medium and suggest that differentially expressed genes were a consequence of Activin A present in CM (82). They also observed that treatment with SB431542 or withdrawal of FGF2 reduced NANOG levels. Na et al. report that FGF2-ERK1/2 signaling is needed for mesendoderm induction and blocking of MEK1/2 using U0126 blocks induction of mesendoderm and suggest that FGF2-ERK1/2 signaling blocks BMP induced mesodermal differentiation (83). They however did not examine the effects on trophoblast related genes. Interestingly, Yu et al. report that *T* induction is reduced by removal of FGF and suggest that FGF2-ERK1/2 may be acting through NANOG to control expression of *T* (84). Bernardo et al also report similar findings (43). The effect of NANOG on *T* has been confirmed through siRNA for NANOG (70). Indeed, FGF2-associated *T* expression could possibly explain why certain studies report no differentiation or differentiation to other lineages in high FGF media but see trophoblast differentiation when they switch to media containing no FGF2 (40, 68, 84). However, Bernardo et al. observed that like *T*, CDX2 levels also depend on FGF signaling, and the increase in CDX2 expression is reduced in the absence of FGF signaling

despite upregulation of CGA, KRT7 and GCM1 (43). Sudheer et al also see reduced CDX2 upregulation in the absence of FGF2 (42). Similar results on the effects of the MAPK pathway on CDX2 were seen in mice where addition of PD98059 to 8-cell-stage embryos leads to low CDX2 expression in the blastocyst and morula (85). Ezashi et al (59) note however that the basal medium used in studies by Bernardo et al (43) was different from those used in other studies. Additionally FGF2 activity is significantly reduced or lost in the absence of heparin after culture at 37°C for a day because of protein aggregation (86). Therefore, specific culture protocols used should also be taken into account while interpreting differentiation outcomes, especially when media is not refreshed daily.

To summarize, differences in medium composition can result in variable outcomes across different protocols for trophoblast differentiation of hESCs. Therefore, the choice of medium used should be made carefully while designing experiments to probe molecular mechanisms regulating trophoblast differentiation.

4.5 Role of culture heterogeneity

Another aspect of differentiation that is often overlooked is the homogeneity, or lack thereof, of differentiated cells. HESCs are known to form a niche (87) and the extent of endogenous and exogenous signaling may vary across the colony, causing differentiation towards different lineages in the same colony (80). Cells of different lineages may secrete distinct growth factors that may influence neighboring cells in a paracrine fashion. Therefore, the result of a differentiation assay cannot be attributed solely to the exogenously added

cytokine or inhibitor. The direct microenvironment of the differentiated cells should also be characterized in the case of heterogeneous cultures (87). Moreover, different labs may observe different results due to heterogeneity of cultures. For example, varying reports have been obtained for BMP4-induced differentiation of hESCs (43, 84, 88). Therefore, it is important to quantify the extent of heterogeneity while assessing the molecular mechanisms causing differentiation. Erb et al (40) recognized the importance of autocrine and paracrine signaling in trophoblast differentiation and made an attempt to optimize the media to increase homogeneity. Yet, even in their best media conditions, the homogeneity observed was about 80% based on *CDX2* expression alone. In most reports of trophoblast differentiation in hESCs discussed in this review, heterogeneity in cell culture suggests that the results could potentially be affected by paracrine signaling from factors secreted by other differentiated cell types (87).

4.6 Cdx2 as a trophoblast marker

As detailed previously, most reports of trophoblast-like differentiation have primarily focused on using *CDX2* as a marker for trophoblast cells. However, *CDX2* expression can be seen in all the three germ layers, and is therefore not specific to trophoblast alone (89). Studies in mice suggest that while *CDX2* is required for continuation of trophoblast development, it is not necessary for initiation of trophoblast specification (90, 91). Notably, *CDX2* haploinsufficiency does not seem to affect trophoblast differentiation. TS cells could be derived from *CDX2*^{+/-} mutant mouse blastocysts (90). Ralston et al also find that a significant number of genes induced by *CDX2* could also be induced by *GATA3* (92).

In a model proposed by Wu et al the non-redundant role of CDX2 appears to be the repression of pluripotency markers *Pou5f1* and *Nanog* (91). Indeed, there is ectopic expression of *Nanog* and *Pou5f1* in the TE of *CDX2*^{-/-} mutant embryos. Since these pluripotency markers may be independently downregulated in hESCs (41), it is plausible that trophoblast differentiation can potentially proceed independent of CDX2. Further, the expression levels of syncytiotrophoblast markers such as *CGB*, *GCM1*, *KRT7* seem to be regulated independent of *CDX2* levels during hESC differentiation (42, 43). Interestingly, Hemberger et al. have found that *CDX2* transcripts are absent in second and third trimester placentas (44). Taken together, these findings raise the intriguing possibility that hESCs may give rise to differentiated trophoblast derivatives independent of CDX2 activity in cell culture.

4.7 Concluding remarks

While trophoblast-like cells obtained from hESCs share many features with human trophoblasts, certain differences have also been recently identified. A major criticism has been the upregulation of mesoderm markers upon BMP4 treatment of hESCs. It remains to be seen whether BMP4-treated hESCs are indeed mesodermal or trophoblastic. However, the fact that trophoblasts obtained by BMP4 treatment of hESCs differ from primary trophoblasts is insufficient by itself to justify the claim that that hESCs cannot form trophoblasts. Another criticism of hESC-derived trophoblasts stems from the observed hypermethylation of the *ELF5* promoter locus. If hESCs cannot overcome this epigenetic restriction on *ELF5* expression, it may be argued that hESCs cannot access the trophoblast lineage.

Nevertheless, the hypermethylation of ELF5 in BMP4-treated hESCs is insufficient evidence to suggest that trophoblast differentiation of hESCs is impossible. Indeed, studies on human blastocysts reveal that unlike in mouse, the human TE is globally hypermethylated(93); note that the placenta becomes globally hypomethylated. Whether there exists a mechanism for active demethylation of the trophoblast after implantation, and whether such a mechanism can be accessed by hESCs remains to be explored. Therefore, the ability of hESCs to form bona fide trophoblasts, analogous to those found *in vivo* during early gestation, remains an open question.

4.8 References

1. L. Vićovac, C. J. Jones, J. D. Aplin, Trophoblast differentiation during formation of anchoring villi in a model of the early human placenta in vitro., *Placenta* **16**, 41–56 (1995).
2. P. Bischof, I. Irminger-Finger, The human cytotrophoblastic cell, a mononuclear chameleon, *Int. J. Biochem. Cell Biol.* **37**, 1–16 (2005).
3. P. Banerjee, A. T. Fazleabas, Extragonadal actions of chorionic gonadotropin, *Rev. Endocr. Metab. Disord.* **12**, 323–332 (2011).
4. S. J. Fisher, The placental problem: linking abnormal cytotrophoblast differentiation to the maternal symptoms of preeclampsia., *Reprod. Biol. Endocrinol.* **2**, 53 (2004).
5. V. Minas, D. Loutradis, A. Makrigrannakis, Factors controlling blastocyst implantation, *Reprod. Biomed. Online* **10**, 205–216 (2005).
6. G. C. Douglas, B. F. King, Isolation of pure villous cytotrophoblast from term human placenta using immunomagnetic microspheres., *J. Immunol. Methods* **119**, 259–268 (1989).
7. J. J. Caulfield, I. L. Sargent, B. L. Ferry, P. M. Starkey, C. W. G. Redman, Isolation and characterisation of a subpopulation of human chorionic cytotrophoblast using a monoclonal anti-trophoblast antibody (NDOG2) in flow cytometry, *J. Reprod. Immunol.* **21**, 71–85 (1992).
8. G. Aboagye-Mathiesen, J. Laugesen, M. Zdravkovic, P. Ebbesen, Isolation and characterization of human placental trophoblast subpopulations from first-trimester chorionic villi., *Clin. Diagn. Lab. Immunol.* **3**, 14–22 (1996).
9. a Tarrade, R. Lai Kuen, a Malassiné, V. Tricottet, P. Blain, M. Vidaud, D. Evain-Brion, Characterization of human villous and extravillous trophoblasts isolated from first trimester placenta., *Lab. Invest.* **81**, 1199–211 (2001).
10. L. J. Guilbert, B. Winkler-Lowen, R. Sherburne, N. S. Rote, H. Li, D. W. Morrish, Preparation and functional characterization of villous cytotrophoblasts free of syncytial fragments, *Placenta* **23**, 175–183 (2002).
11. M. S. Manoussaka, D. J. Jackson, R. J. Lock, S. R. Sooranna, B. M. Kumpel, Flow cytometric characterisation of cells of differing densities isolated from human term placentae and enrichment of villous trophoblast cells, *Placenta* **26**, 308–318 (2005).

12. M. G. Petroff, T. a Phillips, H. Ka, J. L. Pace, J. S. Hunt, Isolation and culture of term human trophoblast cells., *Methods Mol. Med.* **121**, 203–17 (2006).
13. A. Trundley, L. Gardner, J. Northfield, C. Chang, A. Moffett, Methods for isolation of cells from the human fetal-maternal interface., *Methods Mol. Med.* **122**, 109–22 (2006).
14. G. S. J. Whitley, Production of human trophoblast cell lines., *Methods Mol. Med.* **121**, 219–28 (2006).
15. A.-C. Stenqvist, T. Chen, M. Hedlund, T. Dimova, O. Nagaeva, L. Kjellberg, E. Innala, L. Mincheva-Nilsson, An efficient optimized method for isolation of villous trophoblast cells from human early pregnancy placenta suitable for functional and molecular studies., *Am. J. Reprod. Immunol.* **60**, 33–42 (2008).
16. D. W. Morrish, J. Dakour, H. Li, Functional regulation of human trophoblast differentiation., *J. Reprod. Immunol.* **39**, 179–195 (1998).
17. J. L. James, P. R. Stone, L. W. Chamley, The effects of oxygen concentration and gestational age on extravillous trophoblast outgrowth in a human first trimester villous explant model, *Hum. Reprod.* **21**, 2699–2705 (2006).
18. J. L. James, P. R. Stone, L. W. Chamley, Cytotrophoblast differentiation in the first trimester of pregnancy: Evidence for separate progenitors of extravillous trophoblasts and syncytiotrophoblast, *Reproduction* **130**, 95–103 (2005).
19. J. L. James, P. R. Stone, L. W. Chamley, The isolation and characterization of a population of extravillous trophoblast progenitors from first trimester human placenta, *Hum. Reprod.* **22**, 2111–2119 (2007).
20. K. Red-Horse, P. M. Drake, M. D. Gunn, S. J. Fisher, Chemokine ligand and receptor expression in the pregnant uterus: reciprocal patterns in complementary cell subsets suggest functional roles., *Am. J. Pathol.* **159**, 2199–213 (2001).
21. L. A. Salamonsen, N. J. Hannan, E. Dimitriadis, Cytokines and chemokines during human embryo implantation: roles in implantation and early placentation., *Semin. Reprod. Med.* **25**, 437–44 (2007).
22. N. J. Hannan, P. Paiva, E. Dimitriadis, L. a Salamonsen, Models for study of human embryo implantation: choice of cell lines?, *Biol. Reprod.* **82**, 235–245 (2010).
23. B. Novakovic, L. Gordon, N. C. Wong, A. Moffett, U. Manuelpillai, J. M. Craig, A. Sharkey, R. Saffery, Wide-ranging DNA methylation differences of primary trophoblast cell

populations and derived cell lines: implications and opportunities for understanding trophoblast function., *Mol. Hum. Reprod.* **17**, 344–53 (2011).

24. T. Takao, K. Asanoma, K. Kato, K. Fukushima, R. Tsunematsu, T. Hirakawa, S. Matsumura, H. Seki, S. Takeda, N. Wake, Isolation and characterization of human trophoblast side-population (SP) cells in primary villous cytotrophoblasts and HTR-8/SVneo cell line., *PLoS One* **6**, e21990 (2011).

25. J. L. James, a. M. Carter, L. W. Chamley, Human placentation from nidation to 5 weeks of gestation. Part II: Tools to model the crucial first days, *Placenta* **33**, 335–342 (2012).

26. P. Bischof, In vitro models used to study implantation, trophoblast invasion and placentation, *Placenta* **18**, 67–82 (1997).

27. a. M. Carter, Animal Models of Human Placentation - A Review, *Placenta* **28**, S41–7 (2007).

28. D. W. Morrish, J. Dakour, H. Li, J. Xiao, R. Miller, R. Sherburne, R. C. Berdan, L. J. Guilbert, In vitro cultured human term cytotrophoblast: a model for normal primary epithelial cells demonstrating a spontaneous differentiation programme that requires EGF for extensive development of syncytium., *Placenta* **18**, 577–85 (1997).

29. R.-H. Xu, X. Chen, D. S. Li, R. Li, G. C. Addicks, C. Glennon, T. P. Zwaka, J. a Thomson, BMP4 initiates human embryonic stem cell differentiation to trophoblast., *Nat. Biotechnol.* **20**, 1261–4 (2002).

30. C. Xu, M. S. Inokuma, J. Denham, K. Golds, P. Kundu, J. D. Gold, M. K. Carpenter, Feeder-free growth of undifferentiated human embryonic stem cells., *Nat. Biotechnol.* **19**, 971–4 (2001).

31. S. Protocols, in *Cellular Programming and Reprogramming, Methods in Molecular Biology*, Methods in Molecular Biology. S. Ding, Ed. (Humana Press, Totowa, NJ, 2010), vol. 636, pp. 1–24.

32. R.-H. Xu, In vitro induction of trophoblast from human embryonic stem cells., *Methods Mol. Med.* **121**, 189–202 (2006).

33. L. Aghajanova, S. Shen, a. M. Rojas, S. J. Fisher, J. C. Irwin, L. C. Giudice, Comparative Transcriptome Analysis of Human Trophectoderm and Embryonic Stem Cell-Derived Trophoblasts Reveal Key Participants in Early Implantation, *Biol. Reprod.* **86**, 1–21 (2012).

34. M. Marchand, J. a Horcajadas, F. J. Esteban, S. L. McElroy, S. J. Fisher, L. C. Giudice, Transcriptomic signature of trophoblast differentiation in a human embryonic stem cell model., *Biol. Reprod.* **84**, 1258–1271 (2011).
35. R. Lister, M. Pelizzola, Y. S. Kida, R. D. Hawkins, J. R. Nery, G. Hon, J. Antosiewicz-Bourget, R. O'Malley, R. Castanon, S. Klugman, M. Downes, R. Yu, R. Stewart, B. Ren, J. a Thomson, R. M. Evans, J. R. Ecker, Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells., *Nature* **471**, 68–73 (2011).
36. K. Wolfrum, Y. Wang, A. Prigione, K. Sperling, H. Lehrach, J. Adjaye, The LARGE principle of cellular reprogramming: lost, acquired and retained gene expression in foreskin and amniotic fluid-derived human iPS cells., *PLoS One* **5**, e13703 (2010).
37. R.-H. Xu, R. M. Peck, D. S. Li, X. Feng, T. Ludwig, J. a Thomson, Basic FGF and suppression of BMP signaling sustain undifferentiated proliferation of human ES cells.*Nat. Methods* **2**, 185–90 (2005).
38. Z. Wu, W. Zhang, G. Chen, L. Cheng, J. Liao, N. Jia, Y. Gao, H. Dai, J. Yuan, L. Cheng, L. Xiao, Combinatorial signals of activin/nodal and bone morphogenic protein regulate the early lineage segregation of human embryonic stem cells., *J. Biol. Chem.* **283**, 24991–5002 (2008).
39. P. Das, T. Ezashi, L. C. Schulz, S. D. Westfall, K. a. Livingston, R. M. Roberts, Effects of FGF2 and oxygen in the BMP4-driven differentiation of trophoblast from human embryonic stem cells, *Stem Cell Res.* **1**, 61–74 (2007).
40. T. M. Erb, C. Schneider, S. E. Mucko, J. S. Sanfilippo, N. C. Lowry, M. N. Desai, R. S. Mangoubi, S. H. Leuba, P. J. Sammak, Paracrine and epigenetic control of trophectoderm differentiation from human embryonic stem cells: the role of bone morphogenic protein 4 and histone deacetylases., *Stem Cells Dev.* **20**, 1601–1614 (2011).
41. R.-H. Xu, T. L. Sampsell-Barron, F. Gu, S. Root, R. M. Peck, G. Pan, J. Yu, J. Antosiewicz-Bourget, S. Tian, R. Stewart, J. a Thomson, NANOG is a direct target of TGFbeta/activin-mediated SMAD signaling in human ESCs.*Cell Stem Cell* **3**, 196–206 (2008).
42. S. Sudheer, R. Bhushan, B. Fauler, H. Lehrach, J. Adjaye, FGF inhibition directs BMP4-mediated differentiation of Human Embryonic Stem Cells to syncytiotrophoblast., *Stem Cells Dev.* **00**, ahead of print. (2012).
43. A. S. Bernardo, T. Faial, L. Gardner, K. K. Niakan, D. Ortmann, C. E. Senner, E. M. Callery, M. W. Trotter, M. Hemberger, J. C. Smith, L. Bardwell, A. Moffett, R. a. Pedersen,

BRACHYURY and CDX2 mediate BMP-induced differentiation of human and mouse pluripotent stem cells into embryonic and extraembryonic lineages, *Cell Stem Cell* **9**, 144–155 (2011).

44. M. Hemberger, R. Udayashankar, P. Tesar, H. Moore, G. J. Burton, ELF5-enforced transcriptional networks define an epigenetically regulated trophoblast stem cell compartment in the human placenta, *Hum. Mol. Genet.* **19**, 2456–2467 (2010).

45. M. Donnison, A. Beaton, H. W. Davey, R. Broadhurst, P. L’Huillier, P. L. Pfeffer, Loss of the extraembryonic ectoderm in Elf5 mutants leads to defects in embryonic patterning., *Development* **132**, 2299–2308 (2005).

46. B. Gerami-Naini, O. V. Dovzhenko, M. Durning, F. H. Wegner, J. a. Thomson, T. G. Golos, Trophoblast differentiation in embryoid bodies derived from human embryonic stem cells, *Endocrinology* **145**, 1517–1524 (2004).

47. R. Harun, L. Ruban, M. Matin, J. Draper, N. M. Jenkins, G. C. Liew, P. W. Andrews, T. C. Li, S. M. Laird, H. D. M. Moore, Cytotrophoblast stem cell lines derived from human embryonic stem cells and their capacity to mimic invasive implantation events, *Hum. Reprod.* **21**, 1349–1358 (2006).

48. J. M. Frost, R. Udayashankar, H. D. Moore, G. E. Moore, Telomeric NAP1L4 and OSBPL5 of the KCNQ1 cluster, and the DECORIN gene are not imprinted in human trophoblast stem cells, *PLoS One* **5**, e11595 (2010).

49. R. Udayashankar, D. Baker, E. Tuckerman, S. Laird, T. C. Li, H. D. Moore, Characterization of invasive trophoblasts generated from human embryonic stem cells., *Hum. Reprod.* **26**, 398–406 (2011).

50. I. Peiffer, D. Belhomme, R. Barbet, V. Haydont, Y.-P. Zhou, N. O. Fortunel, M. Li, A. Hatzfeld, J.-N. Fabiani, J. a Hatzfeld, Simultaneous differentiation of endothelial and trophoblastic cells derived from human embryonic stem cells., *Stem Cells Dev.* **16**, 393–402 (2007).

51. M. Giakoumopoulos, L. M. Siegfried, S. V Dambaeva, M. a Garthwaite, M. C. Glennon, T. G. Golos, Placental-derived mesenchyme influences chorionic gonadotropin and progesterone secretion of human embryonic stem cell-derived trophoblasts., *Reprod. Sci.* **17**, 798–808 (2010).

52. P. Mali, Z. Ye, H. H. Hommond, X. Yu, J. Lin, G. Chen, J. Zou, L. Cheng, Improved efficiency and pace of generating induced pluripotent stem cells from human adult and fetal fibroblasts., *Stem Cells* **26**, 1998–2005 (2008).

53. T. Ezashi, B. P. V. L. Telugu, R. M. Roberts, in *Embryonic Stem Cells: The Hormonal Regulation of Pluripotency and Embryogenesis*, C. Atwood, Ed. (InTech, 2011).
54. T. G. Golos, M. Giakoumopoulos, M. a. Garthwaite, Embryonic stem cells as models of trophoblast differentiation: Progress, opportunities, and limitations, *Reproduction* **140**, 3–9 (2010).
55. G. C. Douglas, C. a. VandeVoort, P. Kumar, T. C. Chang, T. G. Golos, Trophoblast stem cells: Models for investigating trophectoderm differentiation and placental development, *Endocr. Rev.* **30**, 228–240 (2009).
56. Y. Wang, L. Cheng, Regulation of Primate Trophoblast Lineage Differentiation—Insights Learned from Human Embryonic Stem Cells, *Reprod. Biol.* **2**, 11–21 (2009).
57. R. M. Roberts, S. J. Fisher, Trophoblast stem cells., *Biol. Reprod.* **84**, 412–21 (2011).
58. M. Idelson, B. Reubinoff, M. A. Hayat, Ed. Stem Cells and Cancer Stem Cells, Volume 6, *Stem Cells* **6**, 87–99 (2012).
59. T. Ezashi, B. P. V. L. Telugu, R. M. Roberts, Model systems for studying trophoblast differentiation from human pluripotent stem cells, *Cell Tissue Res.* **349**, 809–824 (2012).
60. R. L. Gardner, Clonal analysis of early mammalian development., *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **312**, 163–178 (1985).
61. R. S. Beddington, E. J. Robertson, An assessment of the developmental potential of embryonic stem cells in the midgestation mouse embryo., *Development* **105**, 733–737 (1989).
62. H. Niwa, J. Miyazaki, a G. Smith, Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells., *Nat. Genet.* **24**, 372–6 (2000).
63. J. M. Velkey, K. S. O’Shea, Oct4 RNA interference induces trophectoderm differentiation in mouse embryonic stem cells., *Genesis* **37**, 18–24 (2003).
64. Y.-H. Loh, Q. Wu, J.-L. Chew, V. B. Vega, W. Zhang, X. Chen, G. Bourque, J. George, B. Leong, J. Liu, K.-Y. Wong, K. W. Sung, C. W. H. Lee, X.-D. Zhao, K.-P. Chiu, L. Lipovich, V. a Kuznetsov, P. Robson, L. W. Stanton, C.-L. Wei, Y. Ruan, B. Lim, H.-H. Ng, The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells., *Nat. Genet.* **38**, 431–440 (2006).

65. J. Li, G. Pan, K. Cui, Y. Liu, S. Xu, D. Pei, A dominant-negative form of mouse SOX2 induces trophoblast differentiation and progressive polyploidy in mouse embryonic stem cells, *J. Biol. Chem.* **282**, 19481–19492 (2007).
66. H. Niwa, Y. Toyooka, D. Shimosato, D. Strumpf, K. Takahashi, R. Yagi, J. Rossant, Interaction between Oct3/4 and Cdx2 determines trophoblast differentiation., *Cell* **123**, 917–29 (2005).
67. E. Tolkunova, F. Cavaleri, S. Eckardt, R. Reinbold, L. K. Christenson, H. R. Schöler, A. Tomilin, The caudal-related protein cdx2 promotes trophoblast differentiation of mouse embryonic stem cells., *Stem Cells* **24**, 139–44 (2006).
68. D. C. Hay, L. Sutherland, J. Clark, T. Burdon, Oct-4 knockdown induces similar patterns of endoderm and trophoblast differentiation markers in human and mouse embryonic stem cells., *Stem Cells* **22**, 225–235 (2004).
69. M. M. Matin, J. R. Walsh, P. J. Gokhale, J. S. Draper, A. R. Bahrami, I. Morton, H. D. Moore, P. W. Andrews, Specific knockdown of Oct4 and beta2-microglobulin expression by RNA interference in human embryonic stem cells and embryonic carcinoma cells., *Stem Cells* **22**, 659–668 (2004).
70. L. Hyslop, M. Stojkovic, L. Armstrong, T. Walter, P. Stojkovic, S. Przyborski, M. Herbert, A. Murdoch, T. Strachan, M. Lako, Downregulation of NANOG induces differentiation of human embryonic stem cells to extraembryonic lineages., *Stem Cells* **23**, 1035–1043 (2005).
71. R. K. Ng, W. Dean, C. Dawson, D. Lucifero, Z. Madeja, W. Reik, M. Hemberger, Epigenetic restriction of embryonic cell lineage fate by methylation of Elf5., *Nat. Cell Biol.* **10**, 1280–90 (2008).
72. a Ralston, J. Rossant, Genetic regulation of stem cell origins in the mouse embryo., *Clin. Genet.* **68**, 106–12 (2005).
73. H. Niwa, How is pluripotency determined and maintained?, *Development* **134**, 635–46 (2007).
74. H. Hutter, A. Hammer, A. Blaschitz, M. Hartmann, P. Ebbesen, G. Dohr, A. Ziegler, B. Uchanska-Ziegler, Expression of HLA class I molecules in human first trimester and term placenta trophoblast, *Cell Tissue Res.* **286**, 439–447 (1996).
75. R. Apps, S. P. Murphy, R. Fernando, L. Gardner, T. Ahad, A. Moffett, Human leucocyte antigen (HLA) expression of primary trophoblast cells and placental cell lines, determined

using single antigen beads to characterize allotype specificities of anti-HLA antibodies, *Immunology* **127**, 26–39 (2009).

76. a. King, L. Thomas, P. Bischof, Cell culture models of trophoblast II: Trophoblast cell lines - A workshop report, *Placenta* **21**, S113–S119 (2000).

77. G. Chen, Z. Ye, X. Yu, J. Zou, P. Mali, R. a. Brodsky, L. Cheng, Trophoblast Differentiation Defect in Human Embryonic Stem Cells Lacking PIG-A and GPI-Anchored Cell-Surface Proteins, *Cell Stem Cell* **2**, 345–355 (2008).

78. X. Yu, J. Zou, Z. Ye, H. Hammond, G. Chen, A. Tokunaga, P. Mali, Y.-M. Li, C. Civin, N. Gaiano, L. Cheng, Notch signaling activation in human embryonic stem cells is required for embryonic, but not trophoblastic, lineage commitment., *Cell Stem Cell* **2**, 461–71 (2008).

79. R. Hoya-Arias, M. Tomishima, F. Perna, F. Voza, S. D. Nimer, L3MBTL1 Deficiency Directs the Differentiation of Human Embryonic Stem Cells Toward Trophoblast, *Stem Cells Dev.* **20**, 1889–1900 (2011).

80. L. C. Schulz, T. Ezashi, P. Das, S. D. Westfall, K. a Livingston, R. M. Roberts, Human embryonic stem cells as models for trophoblast differentiation., *Placenta* **29 Suppl A**, S10–6 (2008).

81. I. Lipchina, Y. Elkabetz, M. Hafner, R. Sheridan, A. Mihailovic, T. Tuschl, C. Sander, L. Studer, D. Betel, Genome-wide identification of microRNA targets in human ES cells reveals a role for miR-302 in modulating BMP response, *Genes Dev.* **25**, 2173–2186 (2011).

82. B. Greber, H. Lehrach, J. Adjaye, Control of early fate decisions in human ES cells by distinct states of TGFbeta pathway activity. *Stem Cells Dev.* **17**, 1065–1077 (2008).

83. J. Na, M. K. Furue, P. W. Andrews, Inhibition of ERK1/2 prevents neural and mesendodermal differentiation and promotes human embryonic stem cell self-renewal, *Stem Cell Res.* **5**, 157–169 (2010).

84. P. Yu, G. Pan, J. Yu, J. a Thomson, FGF2 sustains NANOG and switches the outcome of BMP4-induced human embryonic stem cell differentiation., *Cell Stem Cell* **8**, 326–34 (2011).

85. C.-W. Lu, A. Yabuuchi, L. Chen, S. Viswanathan, K. Kim, G. Q. Daley, Ras-MAPK signaling promotes trophoblast formation from embryonic stem cells and mouse embryos., *Nat. Genet.* **40**, 921–926 (2008).

86. R. Biology, S. Barbara, EMBRYONIC STEM CELLS / INDUCED PLURIPOTENT STEM CELLS Thermal Stability of Fibroblast Growth Factor Protein Is a Determinant

Factor in Regulating Self-Renewal , Differentiation , and Reprogramming in Human Pluripotent Stem Cells, *Stem Cells* **30**, 623–630 (2012).

87. S. C. Bendall, M. H. Stewart, P. Menendez, D. George, K. Vijayaragavan, T. Werbowetski-Ogilvie, V. Ramos-Mejia, A. Rouleau, J. Yang, M. Bossé, G. Lajoie, M. Bhatia, IGF and FGF cooperatively establish the regulatory stem cell niche of pluripotent human cells in vitro., *Nature* **448**, 1015–1021 (2007).

88. A. K. K. Teo, Y. Ali, K. Y. Wong, H. Chipperfield, A. Sadasivam, Y. Poobalan, E. K. Tan, S. T. Wang, S. Abraham, N. Tsuneyoshi, L. W. Stanton, N. R. Dunn, Activin and BMP4 Synergistically Promote Formation of Definitive Endoderm in Human Embryonic Stem Cells., *Stem Cells* **30**, 631–42 (2012).

89. C. a Moskaluk, H. Zhang, S. M. Powell, L. a Cerilli, G. M. Hampton, H. F. Frierson, Cdx2 protein expression in normal and malignant human tissues: an immunohistochemical survey using tissue microarrays., *Mod. Pathol.* **16**, 913–9 (2003).

90. D. Strumpf, C.-A. Mao, Y. Yamanaka, A. Ralston, K. Chawengsaksophak, F. Beck, J. Rossant, Cdx2 is required for correct cell fate specification and differentiation of trophoderm in the mouse blastocyst, *Development* **132**, 2549–2557 (2005).

91. G. Wu, L. Gentile, T. Fuchikami, J. Sutter, K. Psathaki, T. C. Esteves, M. J. Araúzo-Bravo, C. Ortmeier, G. Verberk, K. Abe, H. R. Schöler, Initiation of trophoderm lineage specification in mouse embryos is independent of Cdx2., *Development* **137**, 4159–69 (2010).

92. A. Ralston, B. J. Cox, N. Nishioka, H. Sasaki, E. Chea, P. Rugg-Gunn, G. Guo, P. Robson, J. S. Draper, J. Rossant, Gata3 regulates trophoblast development downstream of Tead4 and in parallel to Cdx2., *Development* **137**, 395–403 (2010).

93. H. Fulka, M. Mrazek, O. Tepla, J. Fulka, DNA methylation pattern in human zygotes and developing embryos., *Reproduction* **128**, 703–8 (2004).

Table 4.1. Overview of methods used for hESC differentiation to trophoblast and characterization of the derived cells

| Cell Line(s) | Media | Treatment | Culture Protocol | Purity/Comments | Characterization - Markers | Characterization – Type of assay performed | Ref. |
|-----------------|-------|---|--|---|---|--|------|
| H1, H7, H9, H14 | CM | 1) CM+BMP4 (1, 10, 100ng/ml) 2) CM+300ng/ml BMP2 3) CM+300ng/ml BMP7 4) CM+30ng/ml GDF5 5) CM+300ng/ml noggin 6)UM | Matrigel + CM + 4ng/ml FGF2 + Factors (Passage protocol not mentioned) | 119 cells treated with BMP4 gave 322 cells with a flattened morphology (34 died) whereas 137 cells in absence of BMP4 had ES morphology (59 died) 44 syncytia (2 – 100 nuclei) among 622 cells (plated as single cells at low density) after 2 weeks of 100ng/ml BMP4 treatment | 14 genes upregulated at all time points of which 11 trophoblast or placenta development related | Microarray – H1 cells | (29) |
| | | | | | <i>CGB, GCM1, ESRRB, ASCL2, KRT7, CD9, MET, HLA-G, HLA-A, HLA-B, TERT, POU5F1</i> | RT-PCR | |
| | | | | | Change in morphology, syncytial cell fusion | microscopy | |

| Table 4.1 Continued | | | | | | | |
|---------------------|-------------|---|--------------------------------|---------------------------------------|--|--|------|
| | | | | | CGB, Estradiol, Progesterone | Immunoassay | |
| | | | | | CGB | Immunofluorescence , Flow cytometry | |
| H1 | EBs | Switched from MEF- CM/FGF2 to EB media | Spontaneous differentiation | Adherent and matrigel embedded EBs | CGB, progesterone estradiol | immunoassay | (46) |
| | | | | | VIM, CGB, KRT7, KRT8 | Immunohistochemistry | |
| H7, H14 | CM N2B27 | Oct4 siRNA, β 2- microglobulin siRNA | | | <i>CGB, CDX2, GCM1</i> | RT-PCR | (69) |
| | | | | | Morphology, multinuclear cells | Phase contrast | |
| | | | | | SSEA1, TRA- 1-60, SSEA3, HLA-A, B, C | Flow cytometry | |

| Table 4.1 Continued | | | | | | | | |
|---------------------|----|-------------|---|--|--|---|--------------------|------|
| H1, H9 | CM | Oct4 siRNA | Media switched from CM to N2B27 before transfection | | | <i>GATA4, GATA6, AFP, CDX2, GATA2, GATA3, CGA, CGB, NES, PL1, POU5F1, TERT</i> | RT-PCR | (68) |
| | | | | | | GATA6 | Immunofluorescence | |
| | | | | | | Change in morphology | microscopy | |
| H1 | CM | NANOG siRNA | | | | SSEA4 | Flow cytometry | (70) |
| | | | | | | <i>GATA4, GATA6, AFP, CDX2, GATA2, CGB, CGA, LAMB1, POU5F1, SOX2, REX1, TERT, NANOG</i> | qPCR | |
| | | | | | | GATA6, NANOG, AFP | Immunofluorescence | |
| | | | | | | GATA6, NANOG | Western blot | |

| Table 4.1 Continued | | | | | | | |
|---------------------|---|---|---|---|--|-------------------------------------|------|
| H1, H9, H14 | CM UM 1:1 CM-UM 1:1 CM-DMEM/F12, | 1) CM+BMP4 (1, 10ng/ml) | Media with BMP4 replaced daily. Passaged with 2mg/ml dispase | UM has similar level of BMP activity as CM+10ng/ml BMP4 UM + 40ng/ml FGF2 +500ng/ml Noggin maintains pluripotency and developmental potential of hESCs | <i>ID1, ID2, ID3, ID4</i> | qPCR | (37) |
| | | 2) UM+FGF2 (4, 40, 100ng/ml) | | | pSMAD1/5/8, SMAD1/5/8, BMP2/4 | Western blot | |
| | | 3) UM+Noggin (100, 500ng/ml) | | | BMP Activity | Luciferase reporter | |
| | | 4) UM + 40ng/ml FGF2 +500ng/ml Noggin | | | OCT4 | Immunofluorescence , Flow cytometry | |
| | | 5) DMEM/F12 + Serum Replacement (0, 1, 10, 20%) | | | <i>OCT4, NANOG, REX1, PAX6, CGB, FOXA1, NEUROD1, T</i> | RT-PCR | |
| | | | | | CGB | Immunoassay | |

| Table 4.1 Continued | | | | | | | |
|---------------------|--------------------------------|---|--|--|---|--|--------------|
| H1, H9, H7, H14 | CM | 1) CM+100ng/ ml BMP4 | Plate as single cells. Media with BMP4 changed every other day | | Change in morphology, syncitial cell fusion | Immunofluorescence | (32) (31) |
| | | 2) CM+300ng/ ml BMP2 | | | CGB | Immunofluorescence , Flow cytometry | |
| | | 3) CM+300ng/ ml BMP7 | | | CGB, Estradiol, Progesterone | ELISA | |
| | | 4) CM+30ng/ml GDF5 | | | | | |
| H7,H14 | HES (UM) without FGF2 | EB formation followed by selection by HCG secretion | β -HCG selection performed in TS conditioned media | Obtained a mixture of trophoblast cells. Multipotent cells propagated by regular passage. | <i>CGB, CDX2, FGF4, EOMES, HLA- G, CD9, KRT7, HLA-I, NANOG, POU5F1, SOX2, PECAM1, CDH5 FLT1</i> | RT-PCR | (47) |

Table 4.1 Continued

| | | | | | | | |
|--------|----|--|--|--|--|---|------|
| | | | | | KRT7, CGB, HLA-G, PECAM-1, MMP2, E- CADHERIN | Immunofluorescence | |
| | | | | | Cell morphology | Phase contrast | |
| | | | | | Invasive phenotype | Matrigel invasion, Endometrial co culture | |
| | | | | | CGB | ELISA | |
| | | | | | Cell fusion | time lapse microscopy | |
| H1, H9 | CM | + 10ng/ml BMP4 added starting the second day at either 4% or 20% O ₂ in presence or absence of FGF2 | BMP4 addition 24 hours after plating on matrigel. Media replaced daily. | More differentiation at atmospheric oxygen levels and without FGF2. Markers for other lineages not checked. | KRT7, CGB, CGA, OCT4, SSEA-1, HLA-G, GATA2 | Immunofluorescence | (39) |
| | | | | | Differentiation homogeneity | Phase contrast, Morphometry | |
| | | | | | multinuclear syncytia | Immunofluorescence | |
| | | | | | CGB, Progesterone | Immunoassay | |

| Table 4.1 Continued | | | | | | | |
|---------------------|----|--|--|------|--|--|------|
| hES3 | UM | EB formation followed by plating on untreated TC plate | Adhesion based selection EB medium (0.6% methylcellulose in KO-DMEM, 10%FCS, 0.5mM L-glutamine, 0.05mM β-mercaptoethanol, 0.5% NEAA, 0.5% insulin, transferrin, selenium, 10ng/ml FGF2) | ~70% | <i>CGA, CDX1, CDX2, HAND1, INSL4, TERT, POU5F1, ZFP42, GDF3, LEFTY1, NODAL, PECAM1, GATA2, KDR, PAX3, MAP2, PAX6, NEFL, NEUROD1, AFP, ONECUT1, PROX1</i> | qPCR | (50) |
| | | | | | GB25, INSL4, KRT7, CGB, GB17, AFP, PECAM-1, VE-CADHERIN | Immunofluorescence | |
| | | | | | CGB | microparticle enzyme immunoassay(MEIA) | |

| Table 4.1 Continued | | | | | | | |
|--|--------------------------|--|---|---|--|--------------------|------|
| MP2 (iPS clone from IMR90) | CM on matrigel | EB culture 50ng/ml BMP4 for 7 days | | EB derived cells very heterogeneous whereas BMP4 treated cells were less heterogeneous | Morphology | Phase contrast | (52) |
| | | | | | Differentiation | AP staining | |
| | | | | | TROMA-I, OCT4, TRA- 1-60, | Immunofluorescence | |
| H1, H9 | 1) hESC (UM) 2) CM | DNMAML1- GFP transfected to inhibit Notch signaling. BMP4 addition to CM for 10 days | Media changed daily and cells passaged every 3-5 days. | 5-6% of cells in DNMAML+ EBs (<1% in control) BMP4 still heterogeneous but better than control EBs | <i>CDX2, CGA, CGB, GATA4, PAX6, T, AFP, EOMES, POU5F1, LFNG, HEY1, HEY2, HES5, NANOG, TALI</i> | qPCR | (78) |
| | | | | | CD34, β -HCG, TROMA-I | cytometry | |
| | | | | | CGB | ELISA | |
| | | | | | TROMA-I, OCT4 | Immunofluorescence | |

| Table 4.1 Continued | | | | | | | |
|-----------------------------------|----|--|--|--|---|--------------------|------|
| H1, H9, AR1/2-C1 (derived), G-GFP | CM | 1) Spontaneous differentiation of EBs 2) 50ng/ml BMP4 | | EB generates heterogeneous population with all lineages except trophoblast | <i>CGA, CGB, CDX2, AFP, CD34, MSII, PAX6, GATA4, TALI, SERPINA1</i> | RT-PCR | (77) |
| | | | | | TROMA – I, P – CADHERIN | immunofluorescence | |
| | | | | | pSMAD1/5/8 | Western blot | |
| | | | | | BMP4 induced signaling | Luciferase assay | |
| | | | | | CGB | ELISA | |
| H1, H9 | CM | +10ng/ml BMP4 (No FGF2) | | Mixture of HCG+ and HLA-G+ cells obtained | Cell morphology | Phase contrast | (80) |
| | | | | | <i>CGA, CGB, OCT4, KRT7</i> | Immunofluorescence | |
| | | | | | <i>OCT4, CGA, NANOG, CGB, SOX2, KRT7, LEFTY2, KRT8</i> | Microarray data | |

Table 4.1 Continued

| | | | | | | | |
|----|-------|---|--|--|---|------------------------------------|------|
| H9 | TeSR1 | 10 μ M SB431542 10 μ M SU5402 \pm 100ng/ml FGF2 \pm 10ng/ml ActivinA 100ng/ml BMP4 100ng/ml Noggin | | Observed markers for other lineages, especially mesoderm in all conditions | BMP signaling | BMP responsive luciferase reporter | (41) |
| | | | | | CGB | immunoassay | |
| | | | | | SMAD1/5/8, BMP4, pSMAD1/5/8, SMAD2/3, pMEK1/2, pSMAD2/3 | Western blot | |
| | | | | | <i>POU5F1</i> , <i>NANOG</i> | qPCR | |
| | | | | | OCT4 | Flow cytometry | |
| | | | | | Markers for all lineages compared | Microarray | |
| | | | | | | | |

Table 4.1 Continued

| | | | | | | | |
|--------------------|----|--|---|--|---|------------------------------------|------|
| H1, HUES- 17 | CM | <p>1) CM+ SB431542 (0, 1, 10μM) 2) CM+ BMP4(0, 1, 10, 50ng/ml) 3) CM + Follistatin (0, 3, 30, 300ng/ml) 4) CM+10ng/ml BMP4 + ActivinA (1, 10, 100ng/ml)</p> | <p>hESCs cells passaged approximatel y once a week by incubation in 1 mg/ml collagenase IV for 30 min at 37 °C.</p> | <p>Also saw some upregulation of neurectoderm markers indicating possible heterogeneity in culture</p> | <p><i>CGA, CGB</i> <i>CGB, CGA,</i> <i>CDX2, GCM1,</i> <i>GATA2,</i> <i>MSX2, BMP2,</i> <i>BMP4, BMP7,</i> <i>FGF2, FGF4,</i> <i>FGF8, WNT3</i></p> | <p>Immunofluorescence qPCR</p> | (38) |
|--------------------|----|--|---|--|---|------------------------------------|------|

Table 4.1 Continued

| | | | | | | | |
|--|--|--|--|--|---|--------------|--|
| | | | | | <i>POU5F1,</i> <i>NANOG,</i> <i>LEFTYA,</i> <i>LEFTYB,</i> <i>NODAL,</i> <i>FGF2, FGF4,</i> <i>FGF8, WNT3,</i> <i>EOMES, NFH,</i> <i>SERPINA1,</i> <i>SLC25A20,</i> <i>SOX1, SOX3,</i> <i>NEUROG2,</i> <i>NES</i> | qPCR | |
| | | | | | OCT4, SMAD2, pSMAD2, SMAD1, pSMAD1 | Western blot | |
| | | | | | CGB, Estradiol, Progesterone | Immunoassay | |

| Table 4.1 Continued | | | | | | | |
|------------------------------|----------------------------|---|--|----------------------------------|---|--|------|
| H1 | UM, CM before EBs are made | Combination EBs (H1 with either term placental fibroblast TPF or dermal fibroblasts from fetal foreskin CI2F) | Switched to EB media | ~40% | VIM, KRT7, CGB, MKI67 | Immunohistochemistry | (51) |
| | | | | | CGB, progesterone | radioimmunoassay and enzyme immune assay | |
| | | | | | VIM, KRT7 | Flow Cytometry | |
| | | | | | OCT3/4 | Immunofluorescence | |
| H1, H9, AFiPS C4, 5,6,10, 41 | CM | 1) 100ng/ml BMP2/4(No FGF2) for 5 days 2) 10ng/ml BMP4 + 10µM SB for 7 days | Media switched to N2B27 for experiments and replaced daily | No specific test for homogeneity | <i>CDX2, KRT7, HAND1, FOXF1, GATA3, ID2, POU5F1, NANOG, SOX2, TERT, CER1, GDF3, FGF4, DPPA4, DNMT3B, LEFTY1</i> | qPCR, microarray | (36) |

| Table 4.1 Continued | | | | | | | |
|---------------------|------------|---|--|---|---|--|------|
| | | | | | CGB, OCT4, NANOG, SSEA4, TRA-1-60 | Immunofluorescence | |
| | | | | | Change in morphology | Phase contrast | |
| H7S14, Shef4 | hESC media | EB formation followed by selection by HCG secretion (47) | | | Not characterized assumed to be TS cell (EB-TS) based on Harun et al. | | (48) |
| H7, Shef4 | hESC | TSCM+25ng/ml FGF4 + 1µg/ml Heparin TSCM(30% TS media + 70% CM) | Grown on MEFs and then EBs. Media switched from hESC to TSCM+25ng/ml FGF4 + 1µg/ml Heparin | Cytotrophoblasts exhibit cell fusion and invasive phenotypes (multipotent) Heterogenous mixture of both obtained at high passage numbers | KRT7, MKI67, HLA-G, VWF, HSD3B1 | Immunofluorescence | (49) |
| | | | | | MMP2, MMP9 | Immunofluorescence , Western blot, gelatine zymography | |
| | | | | | Endometrial cell invasion | Coculture of cells and trophoblast vesicles with primary endometrial cells | |

| Table 4.1 Continued | | | | | | | |
|---------------------|--|--|---|---|---|----------------------------------|------|
| | | | | | <i>CDX2, CD9, KRT7, CGB, TIMP1, TIMP2, POU5F1, SOX17, GATA4</i> | RT-PCR | |
| | | | | | Formation of multinucleated cells | Time lapse micrographs | |
| | | | | | CGB secretion | ELISA | |
| H7 | EMIM MIM DSR mTESR StemPro RPMI | 1) 10µM SB431542 + 100ng/ml BMP4 2) 500ng/ml BMP4 | Media changed daily, Accutase or Collagenase IV passaging. Media contains Pen/Strep | About 80% homogeneity even in the best case | <i>CDX2, T, GATA6, POU5F1</i> | qPCR | (40) |
| | | | | | Multinucleated cells, general morphology changes | Phase contrast, Texture analysis | |
| | | | | | OCT4, CDX2, NES, T, GATA6, HDACs1,2,3 | Immunofluorescence | |

Table 4.1 Continued

| | | | | | | | |
|----|---------------------------|---|--|------------------------------------|---|--------------------|------|
| H9 | HES media (Similar to UM) | Spontaneous differentiation in HES media with FGF withdrawn CM + 100ng/ml BMP4 | L3MBTL1 Knockdown (Cells cultured for 2 weeks) Cells fed daily, Dispase Passaging | Mesoderm markers also upregulated. | <i>CGB, CDX2, HAND1, KRT7, KRT8, GCM1, POU5F1, NANOG, AFP, SOX1, PAX6, NEUROD1, MAP2, SOX2, HNF4A, PECAM1, FOXA2, MIXL1, RUNX1, RUNX2</i> | qPCR | (79) |
| | | | | | CGB, Progesterone | Immunoassay | |
| | | | | | HLA-G | Flow cytometry | |
| | | | | | HAND1, OCT4 | immunofluorescence | |
| | | | | | pSMAD1/5/8 | Western blot | |

Table 4.1 Continued

| | | | | | | | |
|--------|----|--|---|--|--|--------------------|------|
| H7, H9 | CM | CM + 100ng/ml BMP4 (FGF2 withdrawn) (However GEO dataset GSE30915 describes them as grown on MEF feeders which is not consistent with the paper) | Media changed daily, passaged with accutase (day 0, 2, 4, 6, 8, 10, 12 RNA) | Single cells more homogenous than colonies | <i>CDX2, KRT7, ID2, HLA-G, CGA, CGB, HCG, CER1, CTSL2, LEFTY2, HAND1, SMAD9, POU5F1, NANOG, SOX2</i> | qPCR | (34) |
| | | | | | CDX2, KRT7 | Immunofluorescence | |
| | | | | | CGB | Immunoassay | |
| | | | | | 670 trophoblast enriched genes, 449 hESC enriched genes | Microarray | |
| | | | | | Change in morphology | Phase contrast | |

| Table 4.1 Continued | | | | | | | |
|---------------------|----|---|---|--|--|------|------|
| H1, H9 | CM | 1) 10ng/ml BMP4 | Media switched from CM to N2B27 for experiments | Not multipotent. Differentiation to syncytiotrophoblast in the absence of FGF2 | <i>CDX2, CGB, KRT7, HLA-G, T, EOMES, CDH1, VIM, SNAI2, GATA6, SOX7, SOX17, CER1, NCAM1, PAX6, POU5F1, NANOG, ERVW-1, FST, GATA3, GCM1, GREM2, HAND1, ERVFRD-1, HOPX, ID2, IPL, SOX1, P53</i> | qPCR | (42) |
| | | 2) 10ng/ml BMP4 + 20μM SB | | | | | |
| | | 3) 10ng/ml BMP4 + 20μM SB + 20μM SU5402 | | | | | |
| | | 4) 10ng/ml BMP4 + 20μM SB + 4-20ng/mlFGF2 | | | | | |
| | | 5) 20μM SU5402 | | | | | |
| | | 6) 10ng/ml BMP4 + 20μM SU5402 | | | | | |
| | | 7) 20μM SB431542 + 20μM SU5402 | | | | | |

Table 4.1 Continued

| | | | | | | | |
|--------|----|--|--|--|--|-----------------|------|
| | | | | | SMAD2, pSMAD2 | Western blot | |
| | | | | | CGB | ELISA | |
| | | | | | Comparison to undifferentiated cells and Placenta, Gene expression | Microarray | |
| | | | | | Changes in morphology | Phase contrast | |
| | | | | | Multinuclear cells | TEM | |
| H7, H9 | CM | | | Same as (34) but also includes comparison with trophoctoderm from 13 human blastocysts | Principal Component analysis and hierarchical clustering | Microarray data | (33) |
| | | | | | <i>HLA-G, CDX2, ID2</i> | qPCR | |

CHAPTER 5

CONCLUSIONS AND FUTURE WORK

There is a need for new solutions to engineer not just specific inhibitors and biosensors and but affinity reagents in general. Further it is likely as we dig deeper different kinds of reagents are needed. These reagents have to be developed using a customized approach tailored to the application and desired results may not be obtained from standard off the shelf reagents. The use of hyperthermophilic protein scaffolds like Sso7d and Sso6904 presented here represents a promising alternative to antibodies in generation of these reagents. The yeast surface display system used here can be used to engineer high diversity ($\sim 10^9$) scaffold libraries that can be used to isolate and characterize these reagents.

Chapter 1 provides a general introduction to protein engineering and information about different approaches and scaffolds used to isolate binding proteins. Additionally information about desired characteristics of binding proteins is presented. Further an outline of our approach to solving some of the current issues with reagent unavailability through the use of protein engineering is also presented.

In chapter 2 we outline an approach integrating theoretical considerations with a protein engineering strategy to design and evaluate suitable biosensors for phosphorylated epidermal growth factor receptor (EGFR) which can be generalized to other intracellular targets. The biosensor developed in this study can be used to study EGFR internalization and phosphorylation dynamics. However we observe that the biosensor derived also has binding to phosphorylated platelet derived growth factor β -receptor (PDGFR) highlighting the importance of negative selection.

We hypothesized that using protein scaffolds with no native binding for intracellular proteins may be more suited for engineering intracellular biosensors and proceeded to develop biosensors for site-specific phosphorylation of EGFR. These binding proteins, SPY992 and SPY1148 targeting the tyrosine residues at 992 and 1148 on EGFR are based on the scaffold Sso7d. Future work will be focused on characterizing these biosensors to study both the PLC-DAG and the RAS-RAF-MAPK pathways. Further work can be focused on engineering biosensors to targets to which biosensors currently do not exist of which there are many. This can help improve our understanding of the mechanistic details of signaling pathways and cellular processes overall.

In chapter 3 we describe our approach for isolation of secretory vesicles to study the secretome of primary tissues. In this work we characterized the binder to demonstrate that it specifically binds vesicles and that these vesicles contain secretory proteins. Our approach serves as a blueprint for engineering affinity reagents for isolation of any organelle of interest even if the proteome of the organelle is unknown. Additional work is currently ongoing for identifying the target protein on the secretory vesicle using yeast two hybrid (Y2H) assays. Further the immunoprecipitation of vesicles from tissues is being optimized to ensure compatibility with mass spectrometry which can result in a wealth of invaluable data.

In chapter 4 an overview of methods previously used for differentiation of hESCs to trophoblast is presented. Further information about the markers and assays used to characterize these trophoblast-like cells is presented highlighting a controversy over whether these cells represent true in vivo trophoblast analogues. The secretome of trophoblast tissue

obtained in chapter 3 can potentially help identify better markers that may help characterize these cells and elucidate mechanisms and signaling pathways involved in differentiation to trophoblast.

Despite the availability of plethora of antibodies majority of the proteome remains uncharacterized. Some of the approaches presented here can be improved on to engineer affinity reagents to other organelles of interest like the Golgi, early endosomes, microvesicles, to which good affinity reagents currently do not exist. This will help validate hypotheses associated with processes involving them and gain a better understanding of these processes. Cancer is a leading cause of death and while early detection can significantly improve outcomes one of the current challenges is the identification of biomarkers. The reagent developed in Chapter 3, VB15 can be improved to increase its specificity and the cancer secretome data can help in understanding of mechanisms involved in cancer metastasis and identification of cancer biomarkers.