UNIVERSITY OF CALIFORNIA,
IRVINE

Design Recombinant Protein Adhesion Molecules Target for Therapeutic Nanoparticles and Tumor Biomarkers

THESIS

submitted in partial satisfaction of the requirements
for the degree of

MASTER OF SCIENCE

in Chemical and Biochemical Engineering

by

Yuanqing Xue

Thesis Committee:
Associate Professor Jered B. Haun, Chair
Associate Professor James P. Brody
Assistant Professor Han Li

2017

DEDICATION

To

my parents and friends

in recognition of their love and support

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

This thesis became into reality with the support of many individuals. I would like to deeply thank Prof. Jered Haun for his guidance, support and patience for every difficulty I had during the projects.

I also would like to thank my committee members Prof. James Brody and Prof. Han Li for their insight and valuable opinions.

Much of my work could not have been done without the help of my lab mates Maha Rahim, Phinix Dominh, Esther Yu-Tin Chen, Yuting Ji and Xiaolong Qiu for giving me trainings of using experiment instruments.

Thank you all for all Haun lab mates for creating such a happy family and support each other all the time.

ABSTRACT OF THE THESIS

Design Recombinant Protein Adhesion Molecules Target for Therapeutic Nanoparticles and Tumor Biomarkers

By

Yuanqing Xue

Master of Chemical and Biochemical Engineering

University of California, Irvine, 2017

Professor Jered B. Haun, Chair

A wide spectrum of therapeutic nanoparticles have been investigated as delivery systems to improve the pharmacological properties of drugs or imaging agents. By utilizing the advantageous properties of nanoparticles such as high surface to volume ratio and unique optical properties, improved delivery has been shown in several different disease applications. Furthermore, nanoparticles can be associated with proteins that can provide the ability to specifically target selected areas or diseased tissues without exposing the rest of the body. In previous work, our lab engineered recombinant protein constructs containing single-chain antibodies to study nanoparticle adhesion to inflammatory molecules. The goal of this work is to advance the recombinant protein expression construct towards two specific goals: 1) site-specific modification of the recombinant proteins with small molecule chemistries using different enzyme/peptide tag systems and 2) create a new panel of adhesion receptors based on I-domains of the integrin LFA-1 for targeting the inflammatory molecule ICAM-1. Specifically, three different enzyme/tag systems are tested and compared using yeast surface display and soluble protein, including

Sfp synthase/S6 tag, and Sortase A/LPETG tag, and Lipoic Acid Ligase (LplA)/LAP2 tag. We also insert wild type and mutant I-domains obtained from the integrin Lymphocyte function-associated antigen-1 (LFA-1) into the plasmid vector and transformed into yeast. The establishment of I-domain targeting panel and enzymatic conjugation method will provide a solid foundation for further optimization of adhesion therapeutic nanoparticles.

CHAPTER 1 INTRODUCTION

Traditional drug admission routes are normally oral intake or injection which limited the therapeutic efficiency of drugs. Drugs interact with healthy tissue tend to cause strong side effects and shorten drug lifetime. To overcome these downsides issues, targeted drug delivery concept was brought up introduced. Targeted drug delivery is a method of delivering medication to a patient in a manner that increases the concentration of the medication in some parts of the body relative to others. Incorporation of the drug into nanoparticle can protect it against degradation in vitro and in vivo, the release can be controlled and it also offers a platform for drug targeting. Nanoparticle loaded with the drug would solely target for diseased tissue thereby avoiding interaction with healthy tissue.[1]

Ideally, for anticancer drugs to be effective in cancer treatment, they should first be able to reach the desired tumor tissues through the penetration of barriers in the body with minimal loss of their volume or activity in the blood circulation. Second, after reaching the tumor tissue, drugs should have the ability to selectively kill tumor cells without affecting normal cells with a controlled release mechanism of the active form. These two basic strategies are also associated with improvements in patient survival and quality of life by increasing the intracellular concentration of drugs and reducing dose-limiting toxicities simultaneously.[2]

There are main two drug targeting mechanisms: 1) passive targeting; 2) active targeting. Passive targeting consists of the transport of nanocarriers through leaky tumor capillary fenestrations into the tumor interstitium and cells by convection or passive diffusion. Nevertheless, to passively reach the tumor, some limitations exist. The passive

targeting depends on the degree of tumor vascularization and angiogenesis. Thus, extravasation of nanocarriers need to vary with tumor types and anatomical sites which give rise to more challenges in drug design. In active targeting, targeting ligands are attached at the surface of the nanocarriers for binding to appropriate receptors expressed at the target site. This avoids the need of variation of the size of nanocarriers. Targeting ligands are either monoclonal antibodies (mAbs) and antibody fragments or other cancer associated antigens.[3]

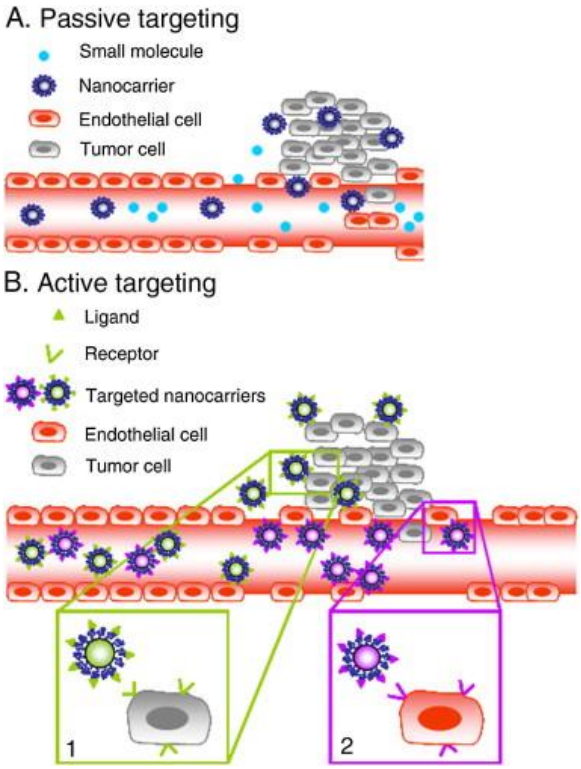


Figure 1[3], Passive targeting of nanocarriers.
(1) Nanocarriers reach tumors selectively through the leaky vasculature surrounding the tumors. (2) Schematic representation of the influence of the size for retention in the tumor tissue. Drugs alone diffuse freely in and out the tumor blood vessels because of their small size and thus their effective concentrations in the tumor decrease rapidly. By contrast, drug-loaded nanocarriers cannot diffuse back into the blood stream because of their large size, resulting in progressive accumulation: the EPR effect. B. Active targeting strategies. Ligands grafted at the surface of nanocarriers bind to receptors (over)expressed by (1) cancer cells or (2) angiogenic endothelial cells.

Careful consideration should be given in the choice of targeting ligand that is used to attach drugs targeted for cancer cells. The targeted antigen or receptor should have a high

density on the surface of the target cells. For example, a receptor density of 105 ERBB2 receptors per cell was required for an improved therapeutic effect of anti-ERBB2-targeted liposomal doxorubicin over non-targeted liposomal doxorubicin in a metastatic breast cancer model.[4] The antigen or receptor should also be able to tolerate cell metabolism and circulation. Circulating shed antigen will compete with the target cells for binding of the targeted therapeutics, and any complexes that form would be rapidly cleared from the circulation.[4]

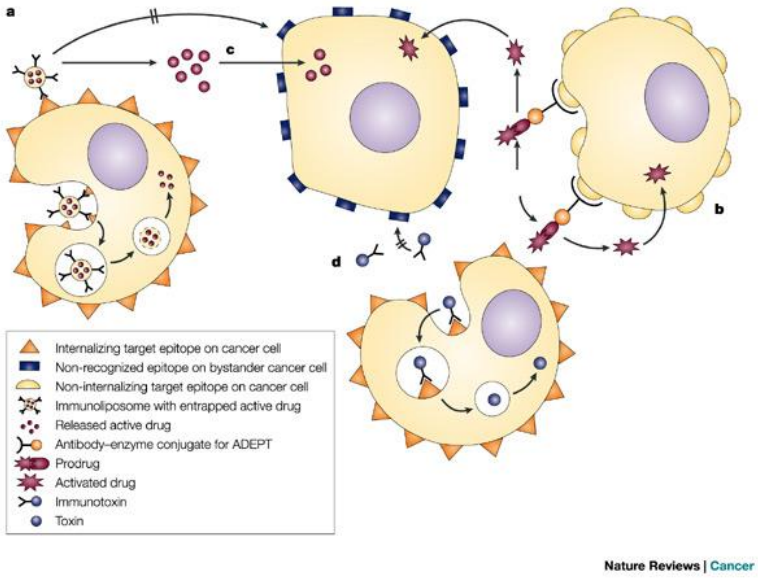
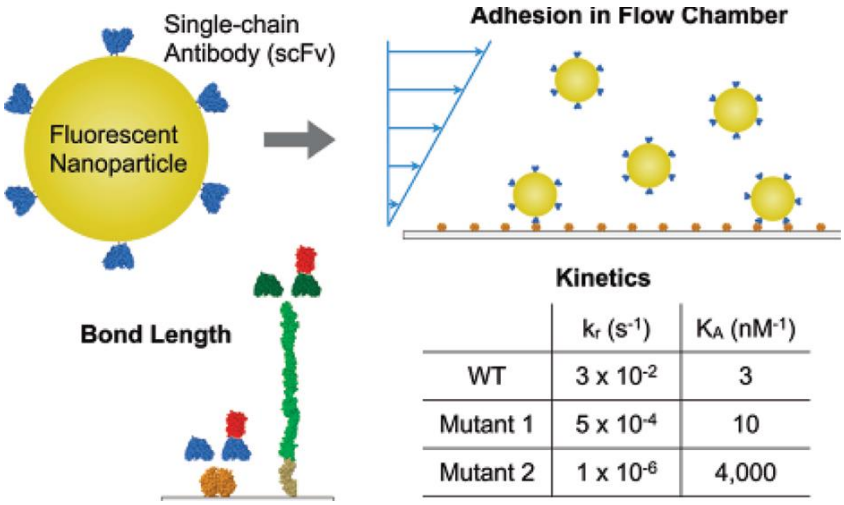


Figure 2[4], Binding of the ligand-targeted therapeutics (LTTs) to their target epitopes
In the case of some antibodies, this will promote receptor-mediated internalization of the LTT and, following release of the therapeutic intracellularly, lead to cytotoxicity (for example, immunoliposomes and immunotoxins). b | Binding of LTTs linked to non-internalizing antibodies will result in the LTT remaining attached at the target-cell surface (for example, ADEPT (antibody-directed enzyme–prodrug therapy)). c | All the cancer cells will preferably express the target epitope; however, some of the cancer cells might not. Drug that is released into the tumor interstitial space might be taken up non-selectively by cancer cells that do not express the target epitope; this results in cytotoxicity by the 'bystander effect' (for example, immunoliposomes and ADEPT). d | Immunotoxins must be internalized to show cytotoxicity, so no opportunity for a bystander effects exists.

We have established a platform of 4420 scFv fused fluorescent protein {mcherry or eGFP}, biotin acceptor protein (AviTag), six-histidine tag (6His), c-myc epitope in expression

vector. Avitag peptide (GLNDIFEAQKIEWHE) can be efficiently biotinylated by the E.coli enzyme, biotin ligase. The recombinant single chain antibody is biotinylated and then functionalized onto avidin coated nanoparticles. Molecular binding interactions are measured in flow chamber under fluid flow. [5]



	k_r (s^{-1})	K_A (nM^{-1})
WT	3×10^{-2}	3
Mutant 1	5×10^{-4}	10
Mutant 2	1×10^{-6}	4,000

Figure 3[5], Model of targeted delivery system

In this study, we developed target receptor by insert S6 and LPETG peptide tag into already constructed PCT302 and PRS314 expression vector based on previous study mentioned above. The recombinant protein then produced in yeast and attached on cell surface after yeast surface display. Enzymatic reactions are performed in vitro with corresponding ligands {Coenzyme A, GGGK peptide} and enzymes {Sfp, SortaseA} on yeast strain EBY100. The construction of these recombinant proteins thus optimize the binding efficiency between nanoparticles and targeted receptors.

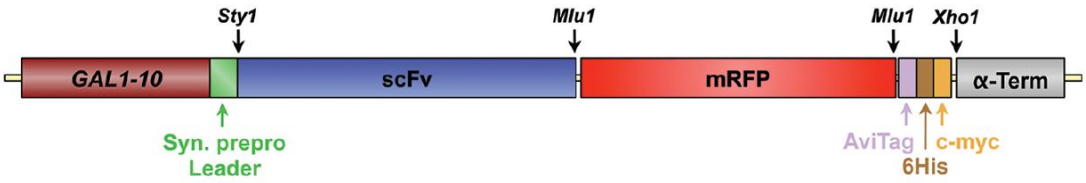


Figure 4, Construct of scFv with peptide linkers

Construct including GAL1-10 as promoter, synthetic prepro leader as sequence to direct the protein, mRFP as fluorophore, AviTag as biotin acceptor protein, 6His tag used for protein purification, cmyc as epitope tag, and untranslated region. Avitag was replaced by S6, LPETG tags using digestion cloning.

CHAPTER 2 Enzyme-Mediated cell surface labeling by Sfp

1.1 Background

Site-specific chemical labeling in living cells has allowed a wide range of study in live cells by spatial and temporal modification of fluorescent proteins. Fluorescent proteins can genetic encode with targeting protein to track the localization and movement of fusion proteins in cells and thus gives a tool for researchers to understand the function of proteins and their native environment.[6]

The discovery of green fluorescent protein (GFP) is a milestone for the revolution of cell fluorescent labeling and imaging. The fusion of protein of interests (POIs) with engineered, self-modifying enzymes such as SNAP-,2 CLIP-,3 TMP-,4,5 or Halo-tags6 also significantly extended the scope to site-specific labeling. Enzyme-based tags provide covalent labelling reactions with higher specificity and shorter reaction time. This gives advantage for cell analysis that is usually accompanied by fast biological process.[7] However, the fusion of enzymes can disturb protein trafficking because of the massive size of the enzymes. To avoid this shortcoming, the focus is to discover smaller tags.

Recently, short peptide tags S6 (GDSLSWLLRLLN, 12 amino acids in length) were discovered from phage display library and demonstrated as an efficient substrate for site specific living cell surface labeling. The reaction is catalyzed by Sfp phosphopantetheinyl transferases (PPTases). S6 peptide tag can be fused to N- or C-termini of proteins to post-translationally modify the peptidyl carrier protein (PCP) or acyl carrier protein (ACP) domains. These domains usually fused with other small molecule probes with diverse structures. PCP and ACP proteins are 80-100 residues in size which stand as a key part in various kinds of synthases. In order to make this enzymatic labeling reaction possible, PCP

and ACP need to be post translationally modified by PPTase to add on a phosphopantetheinyl (Ppant) prosthetic group derived from Coenzyme A (CoASH). After the addition of CoASH substrate, Sfp has been found capable of covalently transfering small molecule entities such as biotin, peptide or fluorophore peptide onto PCP or ACP tags with high efficiency.[8] This property has been widely explored in application for cell surface site specific labeling with target protein.

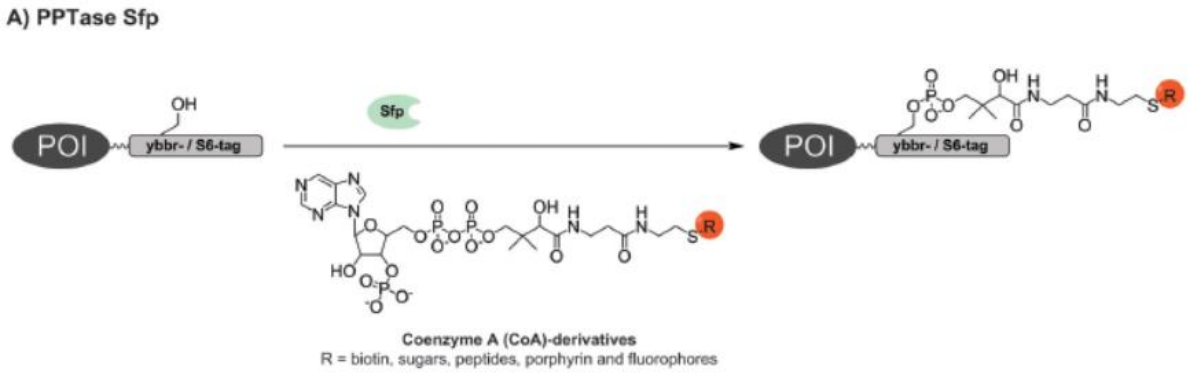


Figure 5[7] Protein labelling via phosphopantetheinyl transferase (PPTase) Sfp

In a previous study, researchers constructed the phase-display peptide library for screening ACP tags S1-S7 and kinetic characterized of Sfp- and AcpS-catalyzed peptide labeling reaction by biotin-CoA. They found that the S6 tag has the highest reaction specificity with Sfp. The S6 tag then fused with epidermal growth factor receptor (EGFR) in expression vector and transfected into mammalian cells for N-terminal S6-EGFR surface protein expression, which would be exposed to surface labeling with CoA-fluorophore. The result was positive which indicates that S6 tag fusing with receptor by Sfp is suitable for cell surface site-specific labeling and imaging.[8]

PPTase Sfp catalyzed CoA-S6 tag labeling system has the following combined advantages: 1) the small size of S6 peptide tag avoids the disturbance of target protein and cell function. 2) One step labeling protein of interest with diverse structure. 3) Direct and highly specific labeling on cell surface without intracellular background fluorescence.[8]

In this paper, we replaced avitag with S6 tag and fused with mcherry fluorescent protein which acts as first labeling to assure protein expression. The recombinant protein was constructed in PCT320 vector with an Aga2p mating protein gene which enables protein fused on C-terminal of Aga2p tethered on yeast cells to surface. This allows the detection of recombinant protein and peptide epitope specific labeling detection by flow cytometry. We used photostable fluorescent substrate CoA 488 as a secondary labeling to test the conjugation efficiency of CoA substrate binding with S6 fusion protein catalyzed by Sfp.

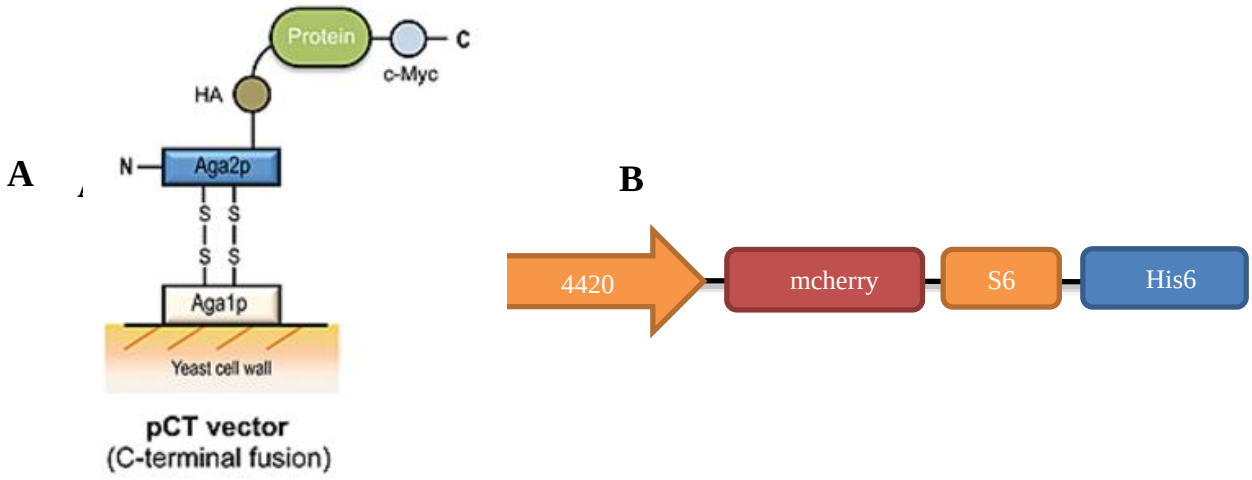


Figure 6[9], Schematic of yeast surface display strategies
A. pCT (C-terminal Aga2p fusion) B. gene sequence of recombinant protein between HA and cymc, which is 4420 scFv, mcherry fluorescent protein, S6 tag, His6 tag.

1.2 Yeast surface display of S6-mcherry recombinant protein

In this paper, S6-mcherry recombinant protein was expressed in EBY100 yeast strain transformed with PCT plasmid showed above.

The transformation and protein induction protocols are showed in appendix.

1.3 Test cymc staining by flow cytometry

Flow cytometry is a wildly used method for analyzing the protein expression on cell surface and intracellular molecule characterizing and defining different cell types in a heterogeneous cell populations, assessing the purity of isolated subpopulations and analyzing cell size and volume. It allows simultaneous multi-parameter analysis of single cells.

It is predominantly used to measure fluorescence intensity produced by fluorescent-labeled antibodies detecting proteins, or ligands that bind to specific cell-associated molecules such as propidium iodide binding to DNA, cancer cell surface specific epitope conjugation with florescent receptor.

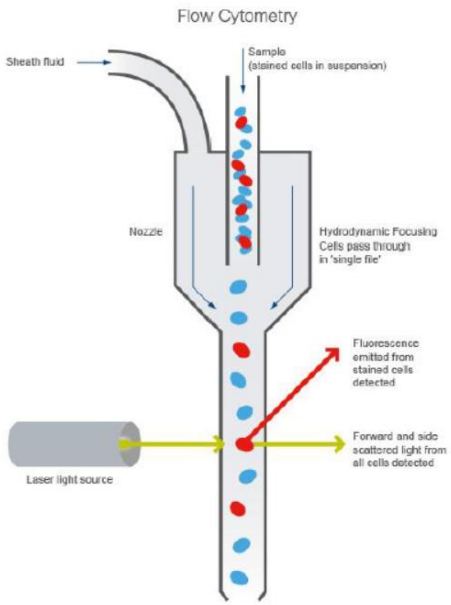


Figure 7, Overview of the flow cytometer. Sheath fluid focuses the cell suspension, causing cells to pass through a laser beam one cell at a time. Forward and side scattered light is detected, as well as fluorescence emitted from stained cells.

Since cymc tag is on the end of protein sequence, by testing cymc expression using two steps labeling along with mcherry signal, we can confirm S6 tag peptide expressed properly for following enzymatic reaction test. For each colony, set one sample for primary antibody negative control, and one for positive reaction.

To avoid spillover between mcherry and FITC signal due to two emission spectrums overlap, we set up compensation controls in order to analyze the population of multicolor stained cells more precisely. Following here are compensation controls:

Unstained control: use cell line without mcherry gene and did not perform cymc staining.

Mcherry single stain control: use cell line with mcherry gene and did not perform cymc staining.

FITC single stain control: use cell line without mcherry gene and performed cymc staining.

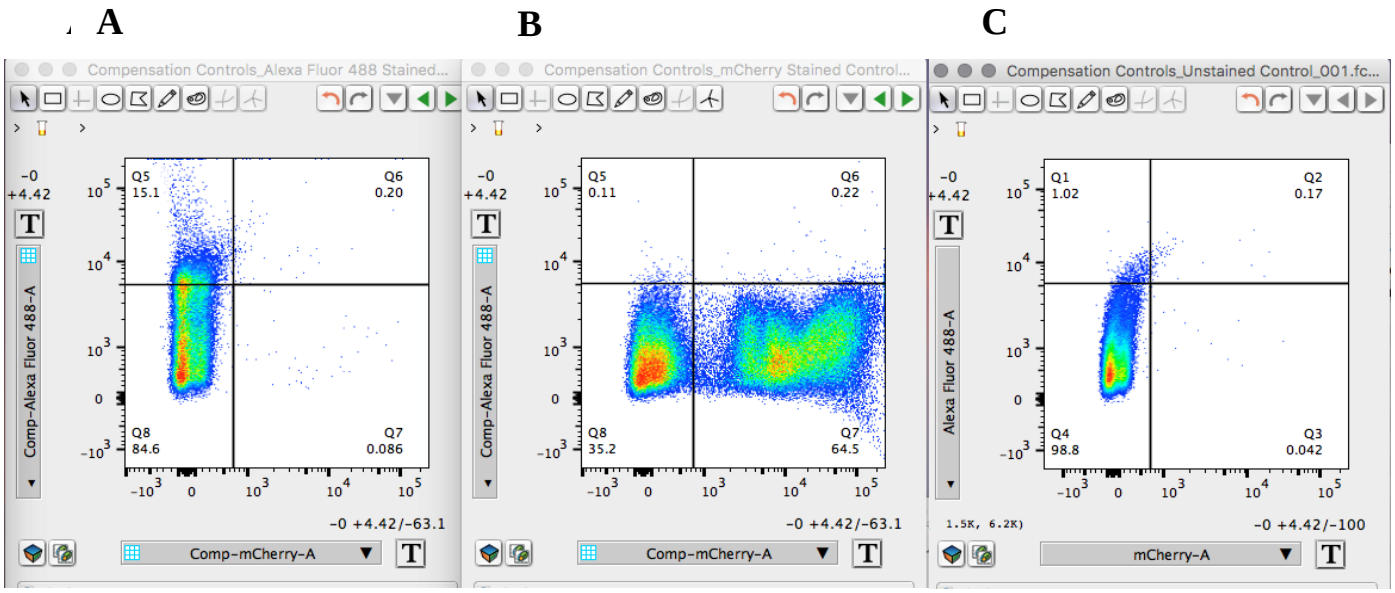


Figure 8, compensation controls for cymc staining (PCT-4420-mcherry-S6-His6-cymc transformed yeast stain) A) FITC single stain control; B) mcherry single stain control; C) Unstained control

The base lines are drew based on FITC single stain control and mcherry single stain control. For the convenience of further analysis.

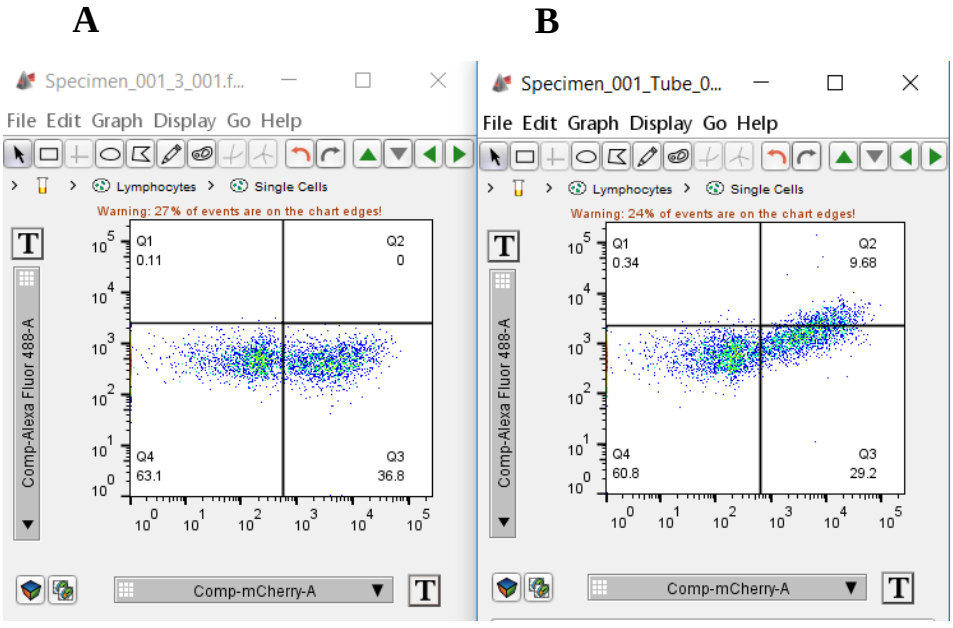


Figure 9, Flow cytometry result for cymc staining yeast expressed PCT-4420-mcherry-S6-His6-cymc recombinant protein, A) without anti-cymc Igg1 antibody, with anti-Igg1 antibody; B) with anti-cymc Igg1 antibody and anti-Igg1 antibody

As shown in figure 9A, without adding anti cymc Igg1 antibody which is the primary antibody, there was little population above AF488=5×10^3 base line, means anti-Igg1 antibody as the secondary antibody did not has binding ligand in the sample and was washed away during wash step. Compare to figure 9B, after adding primary antibody, cell population in mcherry and AF488 both positive region (top right) increased. In both samples, mcherry expressed significantly noticing the dense cell population on the right of mcherry=10^3 base line. This indicates the transformed yeast cell indeed expressed recombinant protein. Since S6 tag is between mcherry and cymc tag, we can confirm S6 tag expression.

1.4 Measurement of Sfp catalyzed S6 tag-CoA conjugation

To measure the Sfp catalyzed enzymatic reaction, the reaction conditions are as following: 1uM Sfp synthase, 10uM CoA-Alexa Fluor 488 substrate, 10mM $MgCl_2$, 50mM HEPES, 2×10^6 yeast cells with recombinant protein expression, pH 7.5. The mixture was incubated in 37℃ at dark for 1 hour.

Table 1 S6 reaction system

	Enzyme Sfp	S6 Peptide	CoA substrate	Mg	HEPES	Reaction temp.	Reaction time
A	1uM	2 mil cells	10uM	10mM	50mM	37℃	1h
B	0	2 mil cells	10uM	10mM	50mM	37℃	1h
C	1uM	2 mil cells	10uM	10mM	50mM	37℃	1h
D	0	2 mil cells	10uM	10mM	50mM	37℃	1h

The reaction result was characterized by flow cytomery.

A B

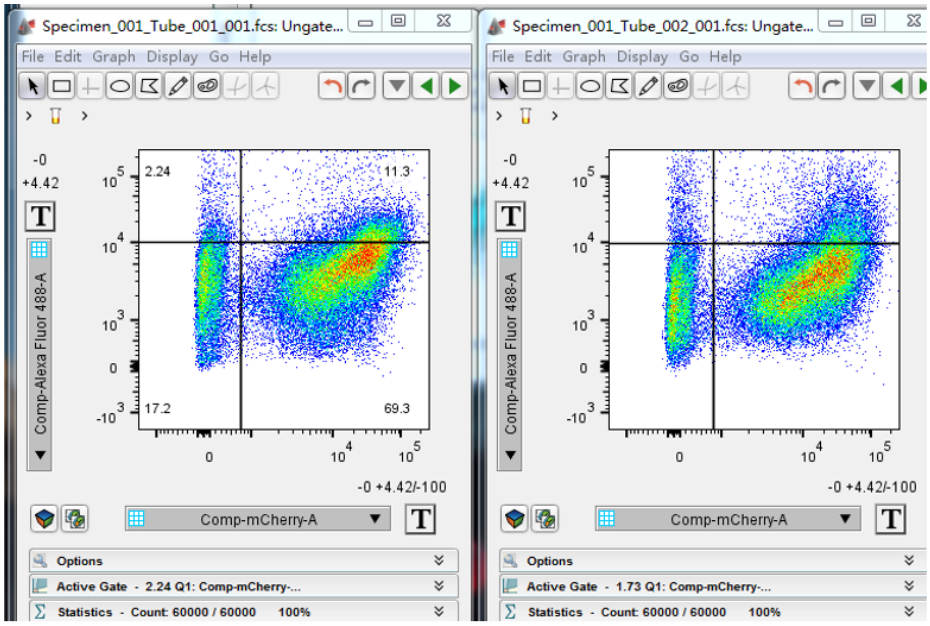


Figure 10, Flow cytometry result of yeast expressed PCT-4420-mcherry-S6-His6- cymc recombinant protein A) Reaction without Sfp. B) Reaction with Sfp.

C D

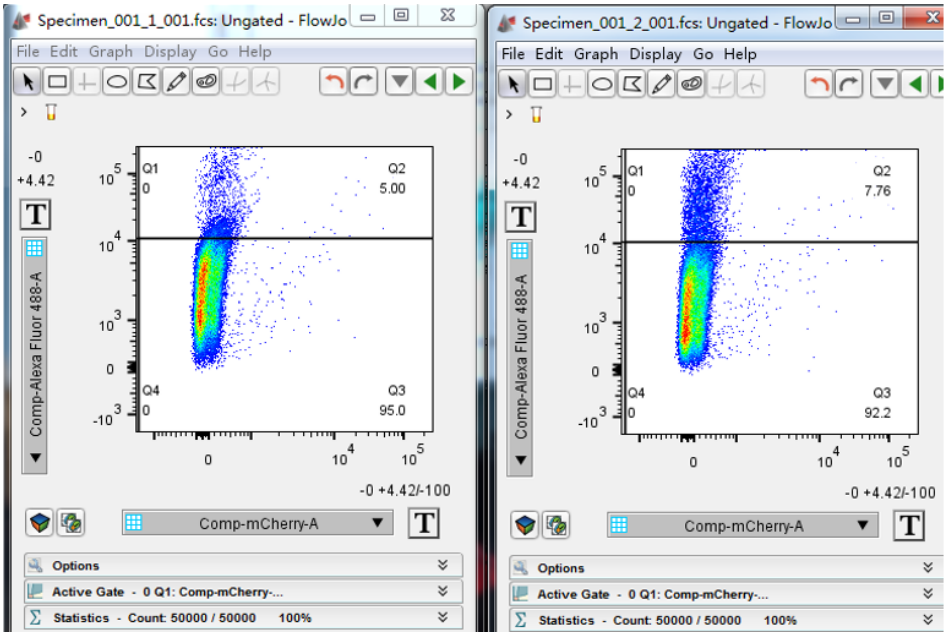


Figure 11, Flow cytometry result of yeast expressed PCT-4420 -S6-His6-cymc C) Reaction without Sfp. D) Reaction with Sfp.

From Figure 10 and Figure 11, the flow cytometry results are showed separately for yeast expressed two kinds of S6 recombinant protein and with the present/ absent of Sfp. As showed in Figure 10, yeast cells expressed S6 recombinant protein fused with mcherry. Compare without and with Sfp synthase graph respectively, there is an obvious shift of cell population from Alexa Fluor 488 (AF488) negative section to AF488 positive section. Mean AF488 intensity for both AF488 and mcherry positive section (top right) raised up 30% (from 21938 to 30462). In Figure 11, cell line expressed S6 recombinant protein without mcherry. Mean fluorescent intensity of AF488 for both fluorophore positive section increased 85% (from 3038 to 5629) from without Sfp sample to with Sfp synthase sample. The population also rise significantly. This proved both 4420 scFv-S6 and 4420 scFv-mcherry-S6 recombinant protein can both conjugate with CoA substrate by Sfp synthase. The difference of conjugation efficiency between two kinds of protein may cause by spatial interference of mcherry in the reaction or unspecific binding on cell surface. For further analysis we will purify fusion protein secreted by BJ5464 yeast strain and run conjugation to eliminate cell binding background.

CHAPTER 3 Cell surface site-specific binding catalyzed by Sortase A

2.1 Background

Molecular biology development in recent years has placed an increasing focus on exploring methods for chemo-selective ligation and site-specific protein labeling. This discovery has a vital role to enhance the understanding of cellular pathways and molecular mechanisms. The introduction of small molecules like fluorescent tags, cross-linking agents into proteins at well-defined positions offers a useful tool in developing novel target therapeutic agents. A great deal of research has proved that post-translational protein modification is a simple and powerful new method in the protein engineering field to introduce unnatural amino acids and protein biophysical probes. Twenty years ago, a protein semisynthesis method was used to ligate a phosphotyrosine peptide to the C-terminus of tyrosine kinase by intercepting a thioester within the protein of interest and an N-terminal cysteine containing synthetic peptide. Then the chemical ligation of peptides is achieved.[13] However, specific labeling at the N terminus of a protein is often the only option available, either because of the constraints imposed by the protein's topology or because the N terminus acts as an essential function for cellular membrane anchoring.[14]

Sortase A (Staphylococcus aureus) is a kind of transpeptidase expressed by Gram-positive bacteria. They are essential in cell wall biosynthesis and covalent attachment of proteins to the peptidoglycan cell wall.[15] In this paper, we use Ca^{2+}-dependent Sortase A for peptide cleavage and ligation for site-specific labeling on living cells. The N-terminal of Sortase recognition motif LPXTG sequence (X being any amino acid labeling described here) is engineered with an exposed glycine. Its C terminus is decorated with a functional group of interest (fluorophores, biotin, lipids, nucleic acids, carbohydrates and so on). After the LPETG

motif is recognized by Sortase A, the catalytic cysteine residue in the enzyme's active site serves as a nucleophile and cleaves between the threonine and glycine residues, forming a thioester intermediate with the peptide probe.[16] The intermediate then reacts with N terminus of a glycine nucleophile to create a new chemical bond between substrate and peptide probe which serves as the new nucleophile.

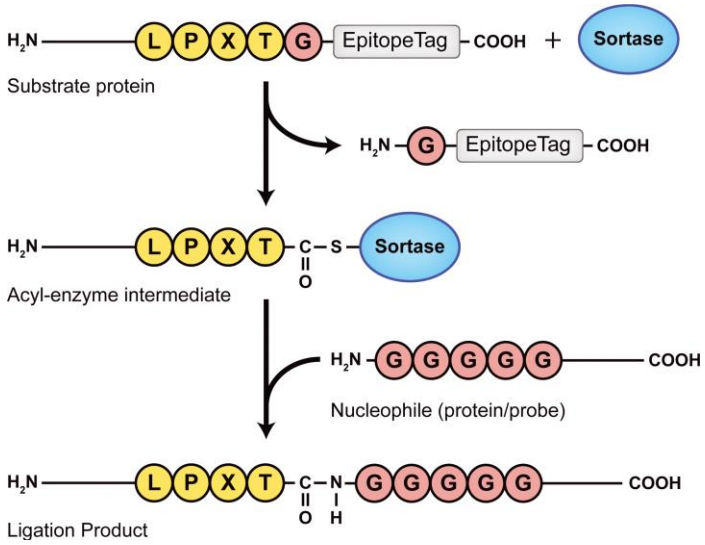


Figure 12[16], Schematic representation of sortase-mediated protein ligation. A substrate protein containing an LPXTG motif near the C-terminus followed by an epitope tag is cleaved between the threonine and the glycine by sortase, thereby releasing the epitope tag (top). The formed acylenzyme intermediate can be resolved by an (oligo) glycine-based nucleophile (middle), resulting in the release of sortase and the formation of a substrate-nucleophile ligation product (bottom).

Sortase A ligation method has been applied to several areas giving advantage of level of selectivity for N-terminal labeling and can provide a labeled protein product in high yield (>90%) and good purity,[14]including preparation for protein-drug conjugation, precise quantitative imaging and therapeutic nanoparticle design, and others. The peptide can be replaced by any molecules accessible by chemical synthesis (e.g., fluorophores, biotin, cross-linkers, lipids, carbohydrates, nucleic acids) decorated on C terminus, provided oligoglycine on its N terminus is accessible as incoming nucleophile. By relying on a common mechanistic

principle, sortagging affords ready access to a wealth of site-specific modifications: C-terminal, internal loop regions, N-terminal and formation of cyclized (poly) peptides.[17]

Here we used GGG-FL (GK (FITC)-4) {sequence GGGK (FITC)} as the nucleophile for LPETG recombinant peptide. Fluorophore FITC is attached to the C terminus of the GGGK leaving N terminus glycine free to conjugate with C terminus of the acyl-enzyme intermediate. The reaction is operated on yeast cell surface with protein anchored on cell surface by AgA2 anchor, followed by mcherry protein, LPETG tag, His6 and cymc tag fused in PCT expression vector with 4420 scFv. Flow cytometry analysis was used to qualify cell surface protein-GGGK conjugation efficiency.

2.2 Yeast surface display of S6-mcherry recombinant protein

4420-mcherry-LPETG-His6- cymc recombinant protein was expressed in EBY100 yeast strain transformed with PCT plasmid. The experiment protocol can refer to appendix B

2.3 Test cymc staining by flow cytometry

Cymc staining assay followed the same protocol as 4420-mcherry-S6-His6-cymc staining in chapter 2. For each colony, set one sample for primary antibody negative control, and one for positive reaction.

Three samples were set up for compensation:

Unstained control: use cell line without mcherry gene and did not perform cymc staining.

Mcherry single stain control: use cell line with mcherry gene and did not perform cymc staining.

FITC single stain control: use cell line without mcherry gene and performed cymc staining.

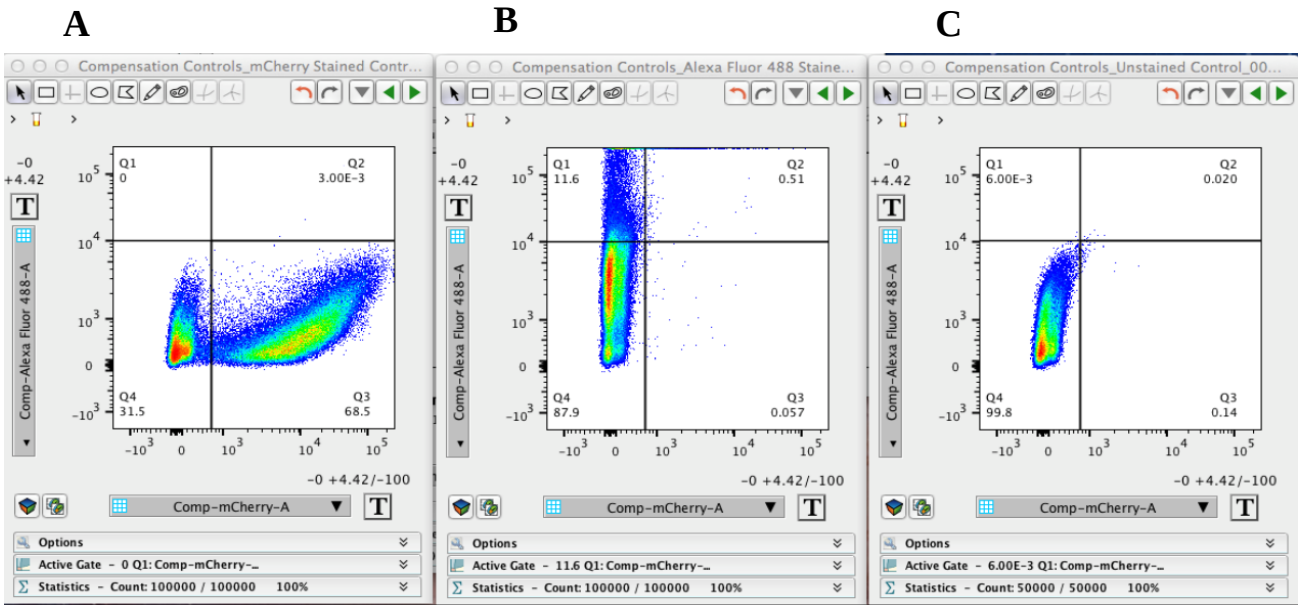


Figure 13, Compensation controls for cymc staining for PCT-4420-mcherry-LPETG-His6-cymc transformed yeast stain. A) mcherry single stain control; B) FITC single stain control; C) Unstained control

The base lines are drew based on FITC single stain control and mcherry single stain control to separate positive and negative population. For the convenience of further analysis.

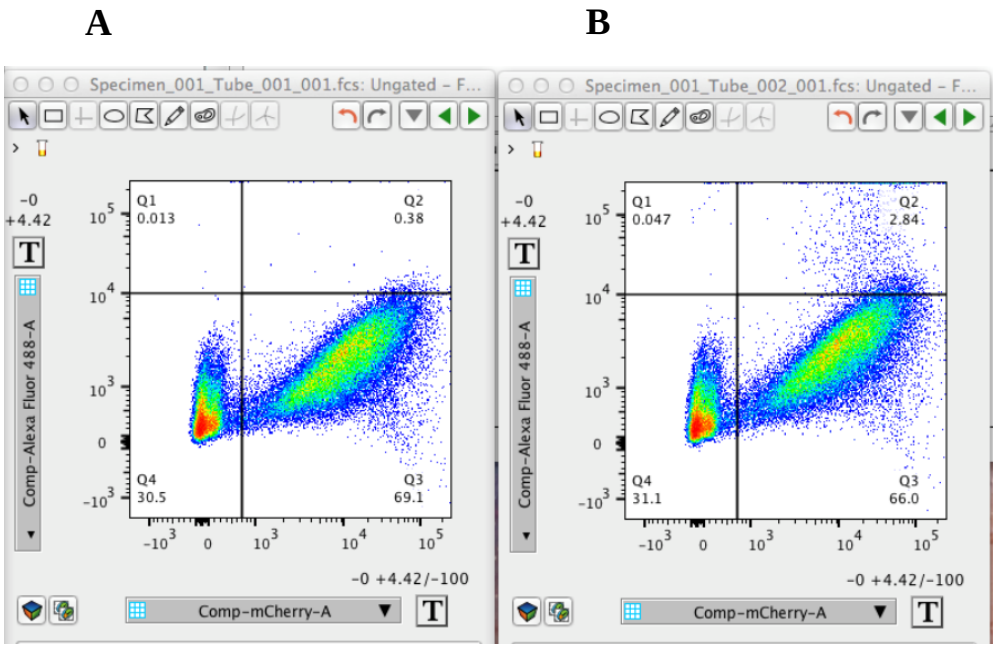


Figure 14, Flow cytometry result for cymc staining for PCT-4420-mcherry-LPETG-His6-cymc transformed yeast stain. A) Without anti-cymc Igg1 antibody, with anti-Igg1 antibody; B) with anti-cymc Igg1 antibody and anti-Igg1 antibody

As shown in Figure 14 A, without adding anti cymc Igg1 antibody which is the primary antibody, there was little population above AF488=10^4 base line, means anti-Igg1 antibody as the secondary antibody did not has binding ligand in the sample and was washed away during wash step. Compare to Figure 14 B, after adding primary antibody, cell population in mcherry and AF488 both positive region (top right) increased from 0.38% to 2.84%. In both samples, mcherry expressed significantly noticing the dense cell population on the right of mcherry=10^3 base line. This indicates the transformed yeast cell indeed expressed recombinant protein. Since LPETG tag is between mcherry and cymc tag, we can confirm LPETG tag expression.

2.4 Measurement of Sortase A catalyzed LPETG tag-GGGK conjugation

For Sortase A mediated ligation, substrate protein LPET- with exposed C-terminus and a target protein GGGK with glycine residue on N-terminus must be solvent-exposed and must be sterically accessible to Sortase A. The reaction was set in sortase buffer (20 mM Tris-base, 50 mM NaCl, pH 7.5), and since the Sortase A is a calcium-assisted transpeptidase, 100 μM Ca^{2+} was also added in the reaction solution. The reaction conditions are as following: 0.266mM GGGK, 100uM Sortase A, $CaCl_2$ 100uM, 2× 10^6 yeast cells with recombinant protein expression, pH 7.5. The mixture was incubated in 37℃ at dark for 3 hours.

Table 2 Sortase A reaction system

	Enzyme Sortase A	LPETG Peptide	GGGK substrate	Ca^{2+}	Sortase buffer	Reactio n temp.
A	100uM	2 mil cells	0.266mM	10mM	50mM	37℃
B	0	2 mil cells	0.266mM	10mM	50mM	37℃

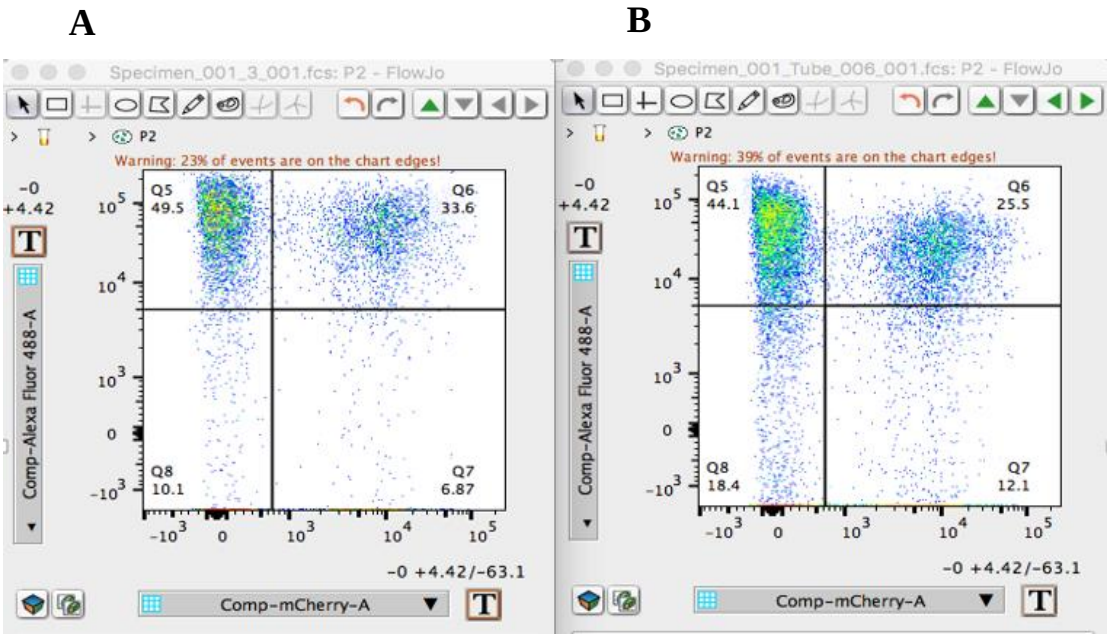


Figure 15, Flow cytometry result for sortase A reaction system A) reaction with Sortase A. B) reaction without Sortase A.

The reaction result was characterized by Flow cytomery.

By comparing figure A and B, cell population in figure A has an obvious up-shift compared to figure B, which means yeast cells were stained with higher density of GGGK-FITG protein substrate after adding Sortase A, leading to higher overall fluorescent intensity of AF488 in figure A. In both mcherry and AF488 positive section (top right), the percentage of cell population raised from 25.5% to 33.6% along with 101% increase of fluorescent intensity of AF488 (from 28258 to 56943). The AF488 background in figure B might cause by unspecific binding of GGGK-FITC to protein on cell surface. This demonstrated that Sortase A can be used for catalyze cell surface site-specific protein-protein conjugation, which has a promising future application for the development of small bio-probe and protein-drug binding. In further research, we plan to reduce fluorescent background by incubating purified recombinant protein with GGGK-FITG substrate and Sortase A to obtain solid results. This reaction system will be used in research for targeting protein tags conjugate on therapeutic nanoparticles

CHAPTER 4 Recombinant plasmid construct for inserted domains binding for ICAM-1

3.1 background

Integrin is a family of heterodimers which function as transmembrane receptors to facilitate cell-extracellular matrix (ECM) adhesion. They have two subunits: α (alpha) and β (beta). Upon ligand binding, integrin activate signal transduction pathways that mediate cellular signals such as regulation of the cell cycle, organization of the intracellular cytoskeleton, and movement of new receptors to the cell membrane.[21] Lymphocyte function-associated antigen-1 (LFA-1, αLβ2 integrin) found on leukocytes has been confirmed in earlier research to bind endothelial cells by conjugation with intercellular adhesion molecule-1 (ICAM-1), which is up-regulated on endothelial cells at sites of inflammation. Integrin ICAM-1 adhesiveness plays a vital role in leucocyte trafficking, migration and cell signaling, which dynamically regulates immune response.[22,23] ICAM-1 molecule belongs to the immunoglobulin superfamily and is induced by inflammatory cytokines which often are found on inflammatory tissue.[24]

Inserted (I) domain is the major ligand binding site which has a structure about 200 amino acids in size and resides in the integrin α subunit. α I domain can be expressed independently of other integrin domains and adopts the dinucleotide-binding or Rossmann fold, with α-helices surrounding a central β-sheet. The structure of I domain binds to a portion of ICAM-1, combined with a complementary structure containing the remaining portion of ICAM-1 providing a topological view of the αLβ2-ICAM-1 interaction as it might take place during cell-cell interactions.[22]

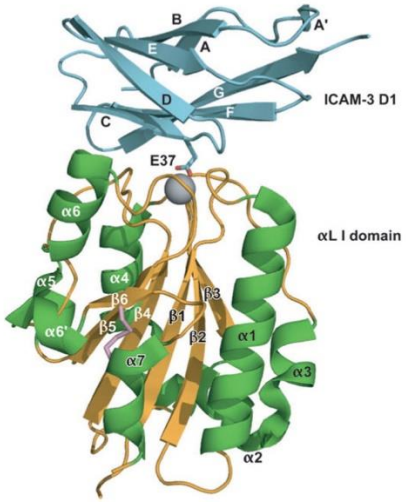


Figure 16[22], A mutant, high-affinity αL I domain (gold β-sheet and coil and green α-helices) in complex with domain 1 of ICAM-3 (cyan). The Mg^{2+} is shown as a gray sphere. The side chain of the key integrin-binding residue, Glu37 of ICAM-3, is shown. The mutationally introduced K287C/K294C disulfide bond that stabilizes the open conformation is shown in pink. ICAM-3 domain 2 is omitted for clarity. [From Protein Data Bank (PDB) ID code 1T0P (7).]

In a previous study, researchers have displayed wild-type I domain and mutants on the surface of yeast and validated expression using ICAM-1 as a binding ligand to investigate integrin antagonists, conformational change of I domain subunit structure during interaction and the effect of metal ions like Mg^{2+}, Ca^{2+}, Mn^{2+}.[26] Affinity and kinetic of wild type and mutants binding with ICAM-1 have also been measured using yeast surface display system with surface plasmon resonance (SPR) as reported before. The data is shown in Figure 17 below. The affinities of the mutants cluster into three groups of high, intermediate, and low affinity. For some of I domain mutants have increased affinity for ICAM-1 by 9,000-fold compared to wild-type.[25] The combination of two activating mutations (F265S/F292G) leads to an increase of 200,000-fold in affinity to intercellular adhesion molecule-1 and slower disassociate rate. The slower disassociation rate of this mutant may indicate some subtle improvements in the ligand bound structures, and as a consequence, a lower free energy of

the complex. Mutant combination K287C/K294C showed the highest binding affinity and lowest disassociation affinity among all mutants.

αL I Domain Design and Binding to ICAM-1

I Domain	Open Cβ - Cβ	Closed Distance (Å)	k_{on} ($M^{-1}s^{-1} \times 10^{-3}$)	k_{off} (s^{-1})	K_D (μM)	Class
K287C/K294C	3.8	9.1	115 ± 7	0.014 ± 0.001	0.15 ± 0.016	High
E284C/E301C	7.0	12.5	105 ± 3	0.045 ± 0.006	0.36 ± 0.04	High
L161C/F299C	8.1	11.4	133 ± 10	0.43 ± 0.07	3.0 ± 0.44	Inter.
K160C/F299C	7.8	8.0	103 ± 15	0.77 ± 0.07	8.4 ± 2.4	Inter.
L161C/T300C	13.0	14.9	89 ± 12	0.76 ± 0.07	9.4 ± 2.4	Inter.
K160C/T300C	12.8	10.9	3.4 ± 0.9	1.2 ± 0.08	450 ± 210	Low
L289C/K294C	8.0	3.9	2.3 ± 0.3	3.6 ± 0.34	1600 ± 170	Low
Wild-type	N/A	N/A	3.1 ± 0.1	4.6 ± 0.36	1500 ± 200	Low

Figure 17[27], All I Domain Design and Binding to ICAM-1

Here we cloned wild-type I domain, F265S mutant I domain and K287C/K294C mutant I domain into PCT302 expression vector with mcherry fluorescent protein. The peptide sequences are shown with the secondary-structure assignment.

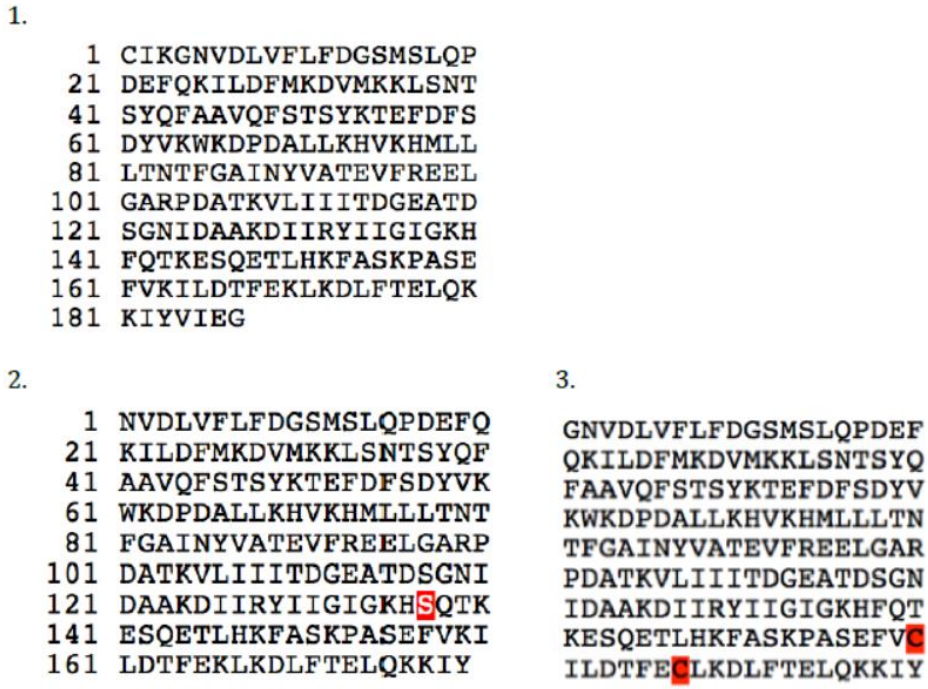


Figure 18, DNA sequence of I domains.(1) Wild-type I domain sequence is shown from residue 125-311; (2) Single mutant F265S sequence is shown from residue 129-307, in which the 265th amino acid mutated from Phenylalanine (F) to Serine (S) (highlight in red); (3) Combination mutants K287C/K294C sequence is shown from residue 128-307, in which the 287th and 294th amino acids mutated from Lysine (K) to Cysteine (C)(highlight in red).

3.2 Materials and Methods

3.2.1 Reagents, Materials and Kits

All materials and reagents used in this paper were listed in table 6 in appendix A

3.2.2 Strains and Media

Escherichia coli strain DH5α (New England Biolabs) was used for recombinant plasmids gene cloning.

3.2.3 Primer and Plasmid Design

Since all three I domain sequences (wild-type, F265S mutant, K287C/K294C) have the same first 25 base pairs of starting and ending sequence, they were amplified by PCR using the same set of sense (GAGGAGGCTAGCGGCAACGTAGACCTGGTATTTC) and antisense

(CTCCTCACGCGTATAGATCTTCTTCTGCAGCTCAGT) primers. The sense primer introduces Nhe1 restriction site on 5' prime end while antisense primer introduces Mlu1 3' prime end (cut sites are underlined) to assist cloning. PCR templates for each sequences are plasmids made by former graduate student. All three I domain sequences were cloned into PCT-4420-mcherry expression vector for yeast surface display. Below are expected sequence constructs.

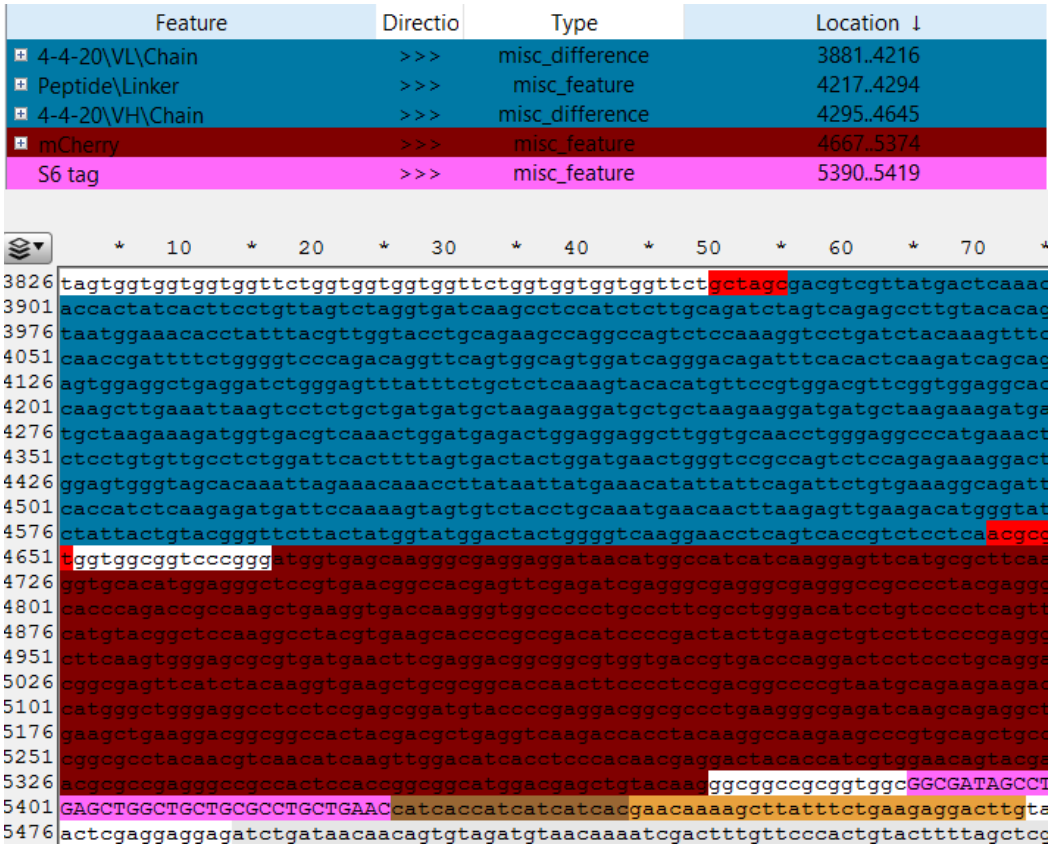


Figure 19, The original sequence of PCT vector with 4420-mcherry-S6-His6-cymc insert. Forward restriction site Nhe1 and reverse restriction site Mlu1 used for cloning are highlighted red.

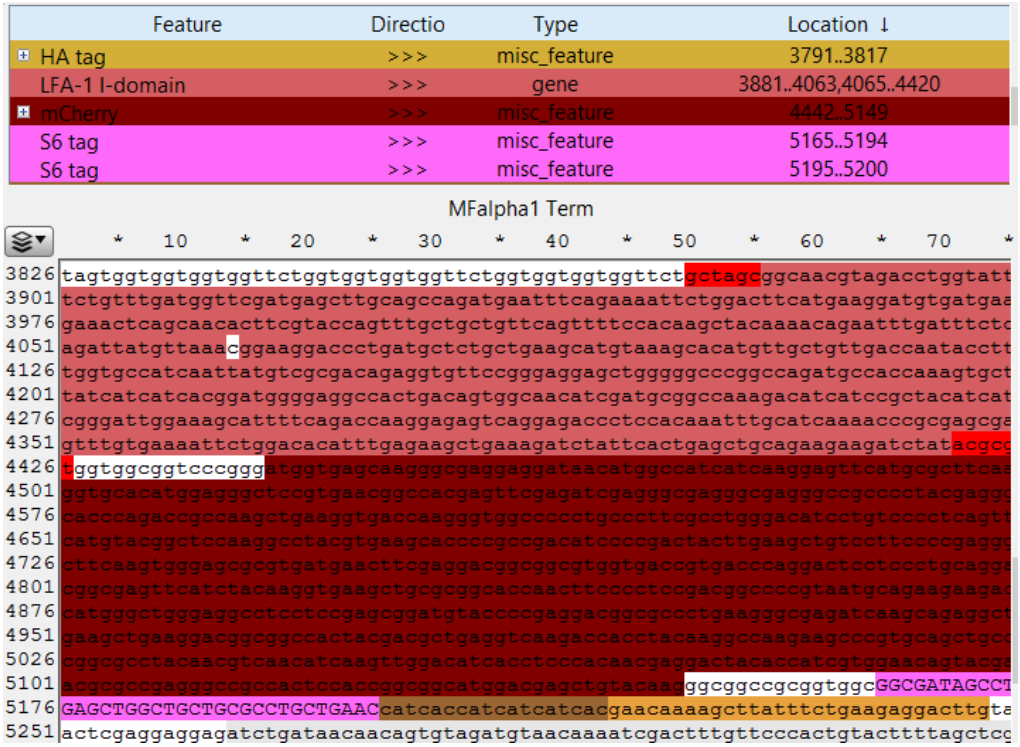


Figure 20, Predicted sequence of PCT-I domain wild-type-S6-His6-cymc

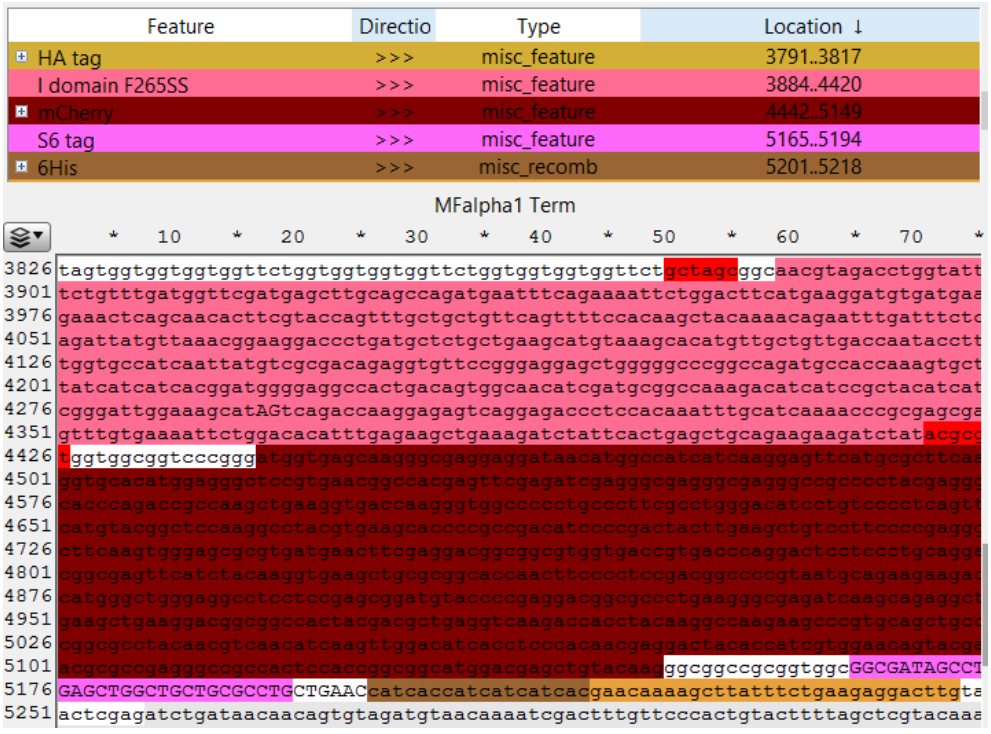


Figure 21, Predicted sequence of PCT-I domain F265S-S6-His6-cymc

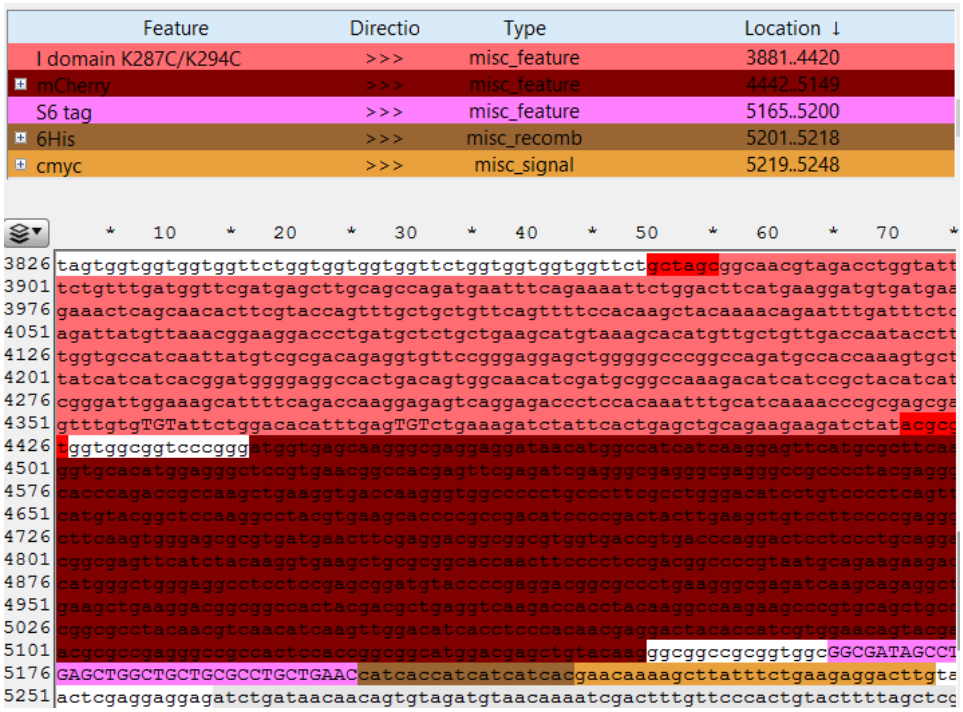


Figure 22, Predicted sequence of PCT-I domain K287C/K294C-S6-His6-cymc

3.2.4 Digestion, ligation and transformation methods

PCR products of three I domain are around 560bp in size and were extracted by electrophoresis. PCR product and PCT-4420-mcherry backbone were double digested by Mlu1 and Nhe1 digestion enzyme.

T4 ligase was used to ligate three types of insert into PCT backbone. T4 ligase catalyzes the formation of a phosphodiester bond between juxtaposed 5' phosphate and 3' hydroxyl termini in duplex DNA facilitated by ATP which provide in 10X T4 ligase buffer. The restriction enzyme Nhe1 and Mlu1 created sticky ends on both insert and backbone during digestion. Then we used gel electrophoresis and PCR purification kit to collect digested backbone and insert separately. T4 ligase joints insert and backbone into intact duplex plasmid by ligate two sets of cut sites.

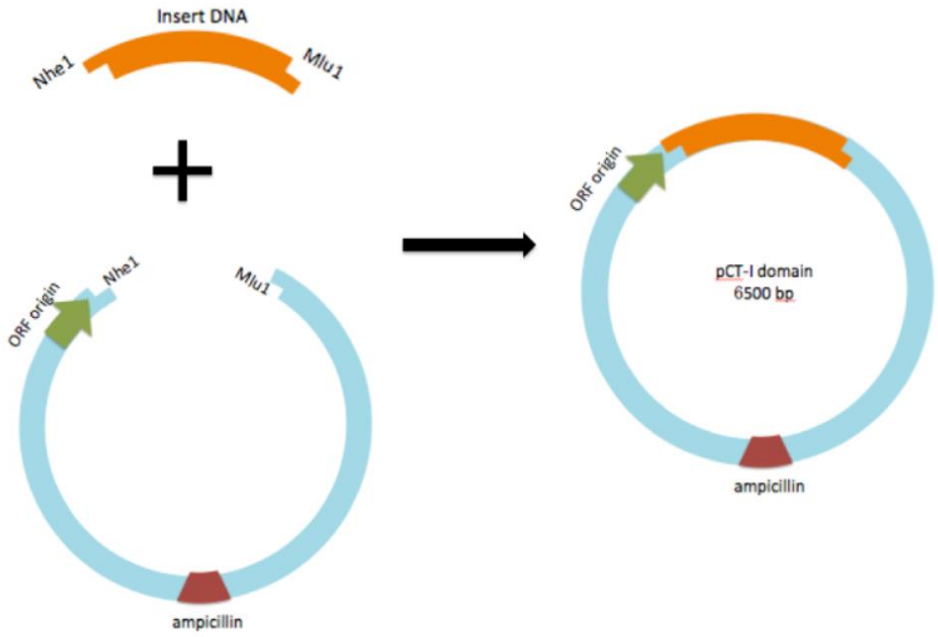


Figure 23, Ligation model

Thereafter, the intact plasmids were transformed into E. coli strain DH5α using heat shock method. The plasmids sequence were then confirmed by DNA sequencing (GENEWIZ)

3.3 results

According to DNA sequence in Figure 20-22, the size of I domains are about 540 base pairs, and we applied the same set of primers to PCR all three types of I domains. The PCR product was collect using gel electrophoresis.

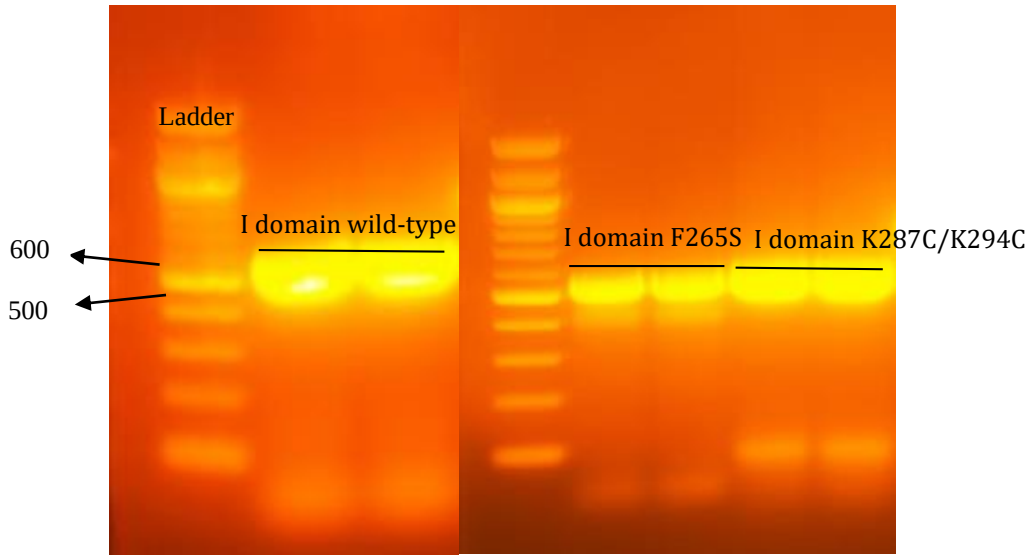


Figure 24, PCR product in gel electrophoresis. From left to right are I domain wild-type, I domain F265S and I domain K287C/K294C. 100bp DNA ladder.

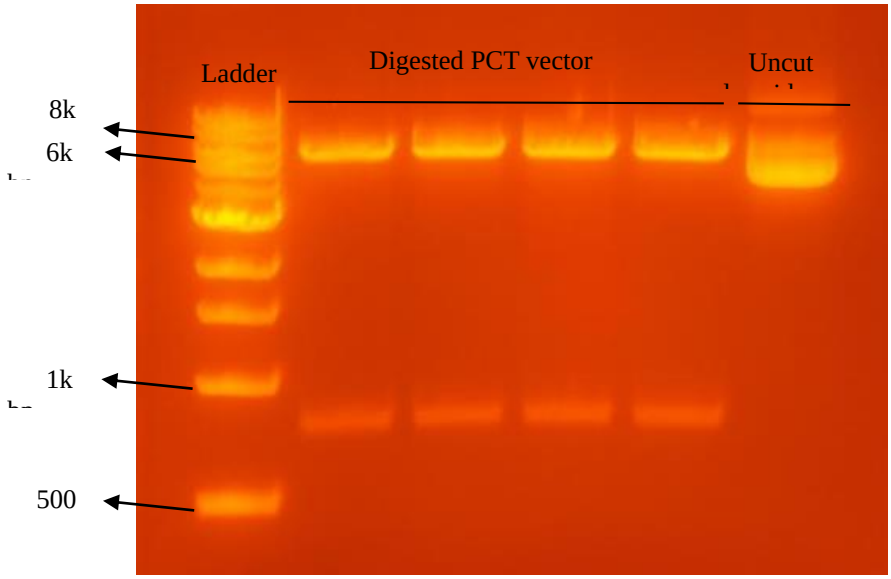


Figure 25， Gel result for PCT-4420-mcherry-S6-His6-c-yMc double digestion.

According to Figure 19, the original insert scFv 4420 in PCT vector is about 750bp. In figure3.10, a band around 750bp appeared on the bottom of each lane, which is 4420 insert

cut off from the plasmid. In the last lane is the uncut plasmid as negative control. Intact plasmid that supercoiling run faster because they sustain less friction than linear DNA.

The digested I domains and pCT backbones were collected by gel extraction kit (QIAGEN) and concentrations were tested by Nanodrop 2000 (Thermo Scientific).

Table 3 Concentrations of digested inserts and vectors

I domains	Concentration ng/ul	PCT	Concentration ng/ul
Wild-type	45.6	**No.1 vector**	12.1
F265S	53.3	**No.2 vector**	17.4
K287C/K294C	47.8		

The collected fragments were then ligated using the molar ratio 3:1 for inserts and vectors, the ligation systems are shown in Table 4

Table 4 Ligation reaction system

I domain wild type		**I domain F265S**		**I domain K287C/K294C**	
Insert	1ul	Insert	1.2ul	Insert	1ul
Vector	16ul	Vector	15.8ul	Vector	16ul
T4 ligase	1ul	T4 ligase	1ul	T4 ligase	1ul
10X ligase buffer	5ul	10X ligase buffer	5ul	10X ligase buffer	5ul

The fragments were ligated follow the scheme in Figure 23 for 4 hours under room temperature and followed by heat shock transformation. After overnight incubation, most

transformants on LB-Ampicillin plates are transformation positive since an anti-ampicillin gene was inserted in PCT vector. We picked 10 colonies on each plates for colony PCR (table 6) to screening for positive colonies using PCT backbone forward and reverse primers. A control was set as a comparison. The control PCR sample was added with PCT-4420-S6- His6-cymc purified plasmid and the same primers.

The colony suspensions with positive results were replated onto new LB-Amp plates and minipreped for storage and sequencing. The correction of inserts wild-type I domain, F265S I domain, and K287C/K294C I domain were confirmed by DNA sequencing (Figure 3.11-13).

Figure 26， Sequencing result for PCT-wild-type I domain-mcherry

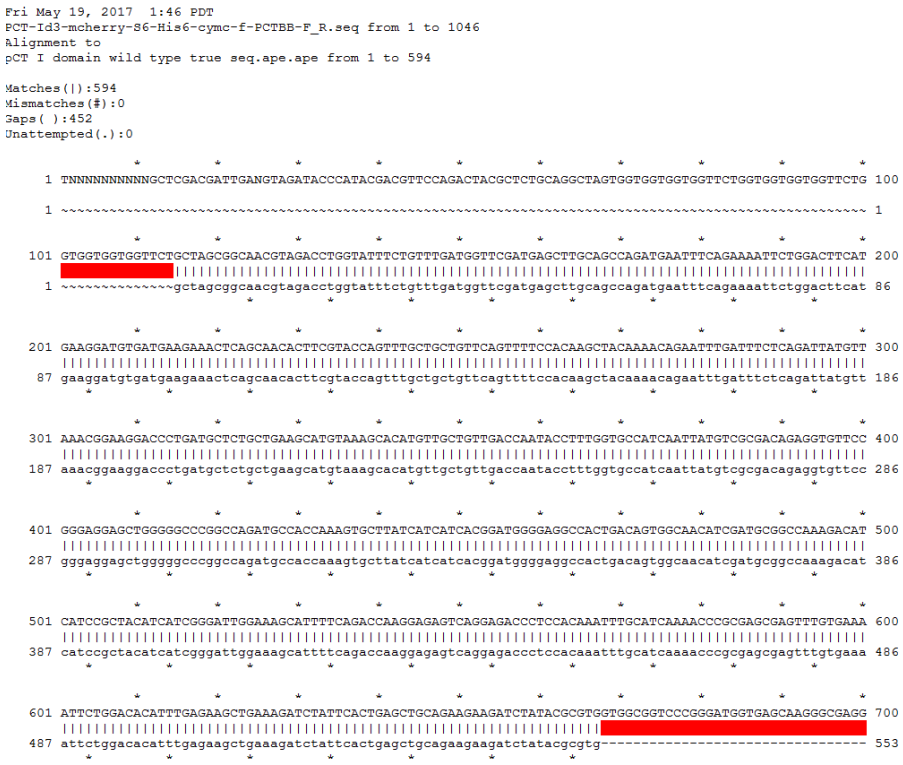


Figure 27， Sequencing result for PCT-I domain F265S–mcherry

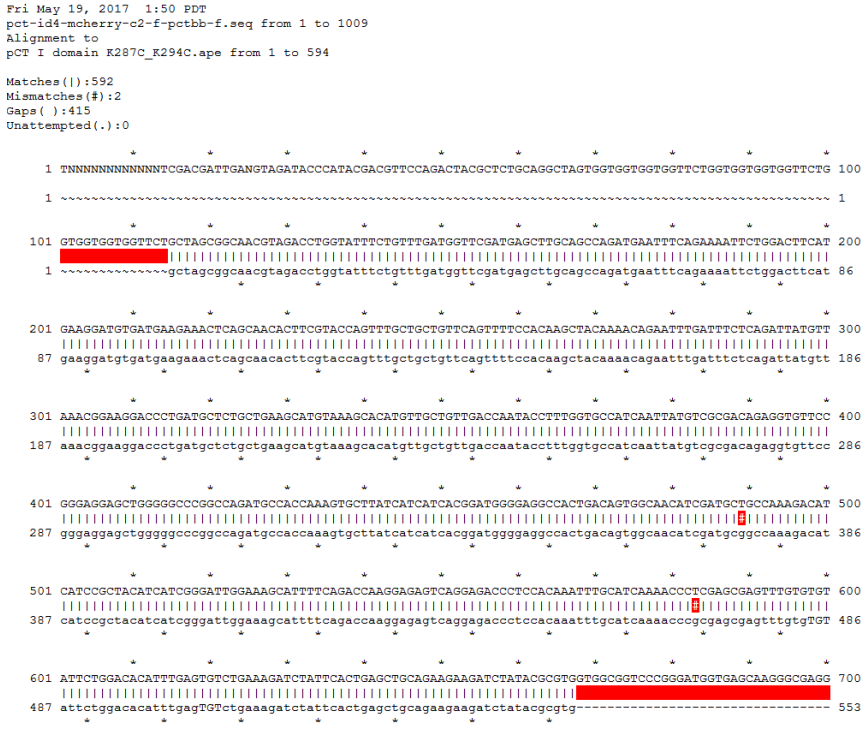


Figure 28， Sequencing result for PCT-I domain K287C/K294C–mcherry.

As results showed, all the base groups paired successfully except for PCT-I domain K287C/K294C–mcherry. The two mutations in plasmid were confirmed still code for the same amino acids as expected. Therefore do not affect further research. Thus wild-type I domain, F265S I domain and K287C/K294C I domain have been inserted in pCT, the recombinant plasmids have been confirmed to be structured. The confirmed colonies were picked into LB-Amp broth to culture overnight, and stored with 50X glycerol under -80℃. Both stored cell culture and clonal population will be further used for making protein production and testing binding assay.

Chapter 5 Flagelliform and GGS flexible linker construction

4.1 background

Spider silk has been noted for years for its extraordinary mechanical property. Flagelliform spider silk is from the N. clavipes flagelliform gland encoding a protein spiral silk[30] which has outstanding mechanical properties compared to other spider silk like dragline silk and minor ampullate silk. Earlier research has showed that the strength of flagelliform spider silk is 1×109 N/m2, with >200 % elongation and requires 4×10^5 J/kg energy to break.[31] Flagelliform silk forms the core fiber of the catching spiral orb-web which weaving spiders rely on in their aerial nets to entrap flying prey. A key mechanical feature of orb-web is high elasticity.[32] This discovery gives us an opportunity to apply flagelliform silk peptide as a short linker between two tags discussed in chapters above. The tag-spider silk linker-tag structure is the final design of protein domain which connects therapeutic nanoparticles and cancer cells. We also constructed one of the most common flexible peptide linker (GGS)n consisting primarily of stretches of Gly and Ser residues. By adjusting the copy number "n", the length of this GS linker can be optimized to achieve appropriate separation of the functional domains, or to maintain necessary inter-domain interactions.[33] These two types of linkers provide flexibility and allow for mobility of the connecting nanoparticle and cells, as well as reducing the undesired interactions. This improves the mechanical and kinetic properties of the targeted drug delivery platform. The kinetic modeling was done using computer simulation tools. On experimental side, we fused flexible linkers within expression vector between S6 and LPETG tags along with mcherry fluorescent protein and His6 tag. The dynamic data will be measured using flow chamber after the recombinant linker conjugated with surface modified nanoparticles.

4.2 Materials and Methods

The flagelliform protein is composed largely of iterated sequences. In the dominant we designed repeat of this protein is Gly-Pro-Gly-Gly-Ala with 25 copy number. We also used 25 copy number for GGS linker with Gly-Gly-Ser repeat. The fusion flexible linker with flagelliform25 and GGS25 insert sequence construct are shown below.

Because of the repeats in both flagelliform25 and GGS25, it is impossible to PCR the product since the primers do not have a specific binding region. We ordered gblocks (genewiz) with sequence (TCTGCCATAGAGTACTGATCACGCGTGGTCCTGGAGGAGCTGGACCAGGTGGTGCAGGACCGGGTG GAGCCGGGCCGGGAGGGGCAGGCCCCGGAGGCGCACGTACGTTCAGATAGACCATGCAGTC) and (AGCAAGGGCGAGGAGGATAAACGCGTGGCGGTAGTGGAGGAAGTGGAGGGTCAGGTGGATCTGGT GGCAGTGGTGGAAGTGGAGGTAGCGGCGGGAGTGGCCGTACGCACCCCGCCGACATCCCCGA) for flagelliform25 and GGS25 respectively. Two restriction sites Bsiw1 and Mlu1 was inserted on both ends of linkers for cloning.

The cloning process followed the same protocol in Chapter 3 except PCR. The gblock DNA was dissolved in 10X TAE buffer and added straight to digestion mixture after calculation.

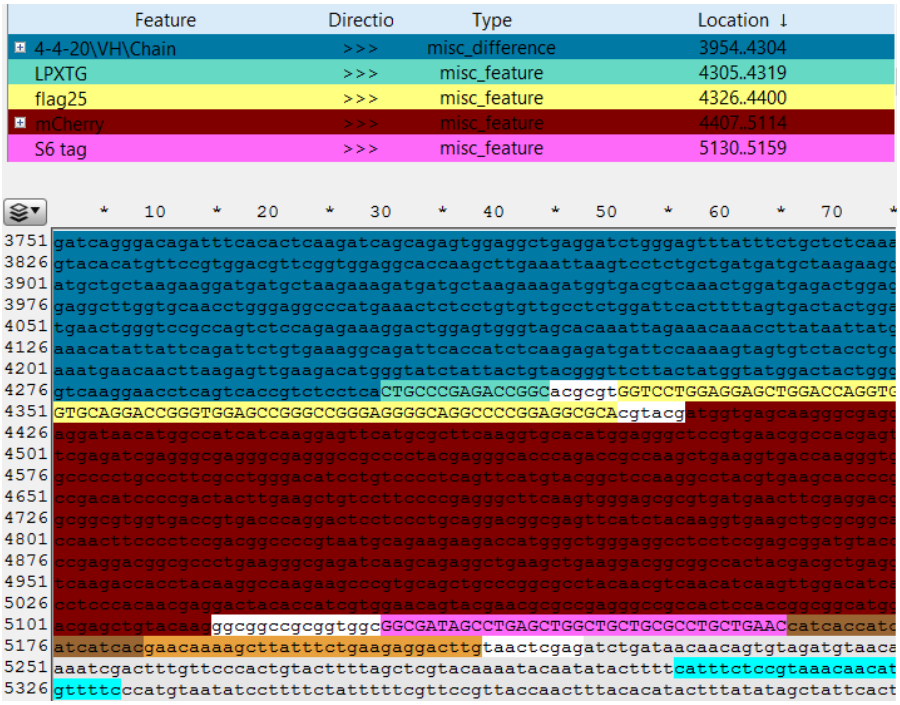


Figure 29， Flagelliform25 in PRS expression vector

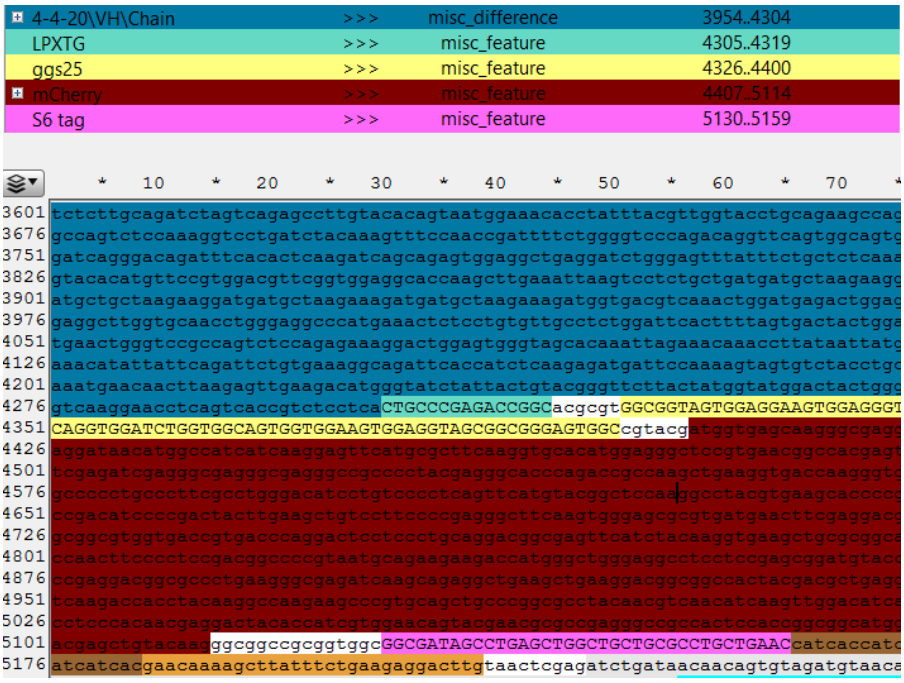


Figure 30， GGS25 linker in PRS expression vector

4.3 Results and discussion

Flagelliform25 and GGS25 were successfully cloned into PRS vector for future protein production. We also designed flagelliform50, GGS50 and stiff linkers to compare the dynamic properties of different linkers.

CHAPTER 6 Summary and Conclusions

To construct the protein and therapeutic nanoparticle adhesion system, the Sortase A and Sfp reaction systems were primarily confirmed by conjugation reaction by flow cytometry. First we successfully constructed recombinant DNA expression vectors with peptide tags LPETG and S6 along with mcherry fluorescent protein, then the plasmids were transformed into yeast strain EBY100 for yeast surface display. Recombinant protein expression was confirmed by both cymc staining and mcherry expression and the expression was confirmed by the high fluorescent intensity of two fluorophores. We demonstrated the conjugation of S6 peptide tag with CoA substrate via Sfp catalyzed system, and ligation between LPETG peptide tag and GGGK by Sortaese A cleavage and mediation. The C-terminus of GGGK protein can be further modified with any biomarkers like biotin, azide or fluorophores for further research.

In the meantime, insert domains for integrin including a wild type and two mutants were successfully cloned into yeast recombinant protein vectors for yeast surface display to prepare for the test of their conjugation efficiency with ICAM-1-biotin and avidin-oregon green system. This reaction system will be further developed to act as a binding ligand for nanoparticles with corresponding receptor.

We also developed the idea of inserting flagelliform linker and flexible peptide linker GGS between two peptide tags to improve the kinetic parameters for nanoparticle binding. The two linkers were cloned into PRS protein secretion vector using gblock gene fragments and sequenced correctly.

Therefore, new targeting protein platforms I domain- linker-mcherry-S6/LPETG or S6 – linker-mcherry-LPETG were designed for high binding efficacy with both nanoparticles and receptor on tumor or tissue.

Moreover, S6 and LPETG recombinant protein were purified after yeast secretion using yeast strain BJ5464 by Histrap column affinity chromatography. The purified protein will be used to test the conjugation again in order to achieve solid confirmation of the reaction system. Furthermore the I domains will be tested by cell-cell binding between CHO-ICAM-1 and EBY100 yeast cells with I domain protein attached to cell surface. After we confirm the most efficient platform, we will scale up protein production for nanoparticle conjugation and flow chamber kinetic measurement.

Reference

1. Muller, Rainer H., and Cornelia M. Keck. "Challenges and solutions for the delivery of biotech drugs–a review of drug nanocrystal technology and lipid nanoparticles." Journal of biotechnology 113.1 (2004): 151-170.

2. Cho, Kwangjae, et al. "Therapeutic nanoparticles for drug delivery in cancer."Clinical cancer research 14.5 (2008): 1310-1316.

3. Danhier, Fabienne, Olivier Feron, and Veronique Preat. "To exploit the tumor micro-environment: passive and active tumor targeting of nanocarriers for anti-cancer drug delivery." Journal of Controlled Release 148.2 (2010): 135-146.

4. Allen, Theresa M. "Ligand-targeted therapeutics in anticancer therapy."Nature Reviews Cancer 2.10 (2002): 750-763.

5. Haun, Jered B., et al. "Using engineered single-chain antibodies to correlate molecular binding properties and nanoparticle adhesion dynamics." Langmuir27.22 (2011): 13701-13712.

6. Marks, Kevin M., and Garry P. Nolan. "Chemical labeling strategies for cell biology." Nature methods 3.8 (2006): 591-596.

7. Lotze, Jonathan, et al. "Peptide-tags for site-specific protein labelling in vitro and in vivo." Molecular BioSystems 12.6 (2016): 1731-1745.

8. Zhou, Zhe, et al. "Genetically encoded short peptide tags for orthogonal protein labeling by Sfp and AcpS phosphopantetheinyl transferases." ACS chemical biology 2.5 (2007): 337-346.

9. Lim, Sungwon, et al. "Dual display of proteins on the yeast cell surface simplifies quantification of binding interactions and enzymatic bioconjugation reactions." Biotechnology Journal (2017).

10. Vivero-Pol, L., George, N., Krumm, H., Johnsson, K., and Johnsson, N. (2005) Multicolor imaging of cell surface proteins, J. Am. Chem.Soc. 127, 12770–12771.

11. Bidlingmaier, Scott, and Bin Liu. "Construction of yeast surface-displayed cDNA libraries." cDNA Libraries: Methods and Applications (2011): 199-210.

12. BD Biosciences, San Jose, CA. "An Introduction to Compensation for Multicolor Assays on Digital Flow Cytometers." BD Biosciences. Technical Bulletin. August 2009

13. Muir, Tom W., Dolan Sondhi, and Philip A. Cole. "Expressed protein ligation: a general method for protein engineering." Proceedings of the National Academy of Sciences 95.12 (1998): 6705-67

14. Theile, Christopher S., et al. "Site-specific N-terminal labeling of proteins using sortase-mediated reactions." Nature protocols 8.9 (2013): 1800-1807.

15. Paterson, G.K. & Mitchell, T.J. The biology of Gram-positive sortase enzymes. Trends Microbiol. 12, 89–95 (2004).

16. Strijbis, Karin, Eric Spooner, and Hidde L. Ploegh. "Protein ligation in living cells using sortase." Traffic 13.6 (2012): 780-789.

17. Guimaraes, Carla P., et al. "Site-specific C-terminal and internal loop labeling of proteins using sortase-mediated reactions." Nature protocols 8.9 (2013): 1787-1799.

18. Popp, Maximilian W., et al. "Sortagging: a versatile method for protein labeling." Nature chemical biology 3.11 (2007): 707-708.

19. Wagner, Koen, et al. "Bispecific antibody generated with sortase and click chemistry has broad antiinfluenza virus activity." Proceedings of the National Academy of Sciences 111.47 (2014): 16820-16825.

20. Mao, Hongyuan, et al. "Sortase-mediated protein ligation: a new method for protein engineering." Journal of the American Chemical Society 126.9 (2004): 2670-2671.

21. Giancotti, Filippo G., and Erkki Ruoslahti. "Integrin signaling." Science 285.5430 (1999): 1028-1033.

22. Luo, Bing-Hao, Christopher V. Carman, and Timothy A. Springer. "Structural basis of integrin regulation and signaling." Annu. Rev. Immunol. 25 (2007): 619-647.

23. Pepper, Lauren R., Daniel A. Hammer, and Eric T. Boder. "Rolling Adhesion of α L I Domain Mutants Decorrelated from Binding Affinity." Journal of molecular biology 360.1 (2006): 37-44.

24. Crucian, Brian, Mayra Nelman-Gonzalez, and Clarence Sams. "Rapid flow cytometry method for quantitation of LFA-1-adhesive T cells." Clinical and vaccine immunology 13.3 (2006): 403-408.

25. Shimaoka, Motomu, et al. "Structures of the αL I domain and its complex with ICAM-1 reveal a shape-shifting pathway for integrin regulation." Cell 112.1 (2003): 99-111.

26. Li, Rui, et al. "Two functional states of the CD11b A-domain: correlations with key features of two Mn2+-complexed crystal structures." The Journal of cell biology 143.6 (1998): 1523-1534.

27. Shimaoka, M., Lu, C., Palframan, R. T., von Andrian, U. H., McCormack, A., Takagi, J. & Springer, T. A. (2001). Reversibly locking a protein fold in an active conformation with a

disulfide bond: integrin alphaL I domains with high affinity and antagonist activity in vivo. Proc. Natl Acad. Sci. USA, 98, 6009–6014.

28. Jin, Moonsoo, et al. "Directed evolution to probe protein allostery and integrin I domains of 200,000-fold higher affinity." Proceedings of the National Academy of Sciences 103.15 (2006): 5758-5763.

29. Huth, Jeffrey R., et al. "NMR and mutagenesis evidence for an I domain allosteric site that regulates lymphocyte function-associated antigen 1 ligand binding." Proceedings of the National Academy of Sciences 97.10 (2000): 5231-5236.

30. Lewis, Randolph V. "Spider silk: ancient ideas for new biomaterials." Chemical Reviews 106.9 (2006): 3762-3774.

31. Gosline, John M., Mark W. Denny, and M. Edwin DeMont. "Spider silk as rubber." Nature 309.5968 (1984): 551-552.

32. Hayashi, Cheryl Y., and Randolph V. Lewis. "Evidence from flagelliform silk cDNA for the structural basis of elasticity and modular nature of spider silks." Journal of molecular biology 275.5 (1998): 773-784.

33. Chen, Xiaoying, Jennica L. Zaro, and Wei-Chiang Shen. "Fusion protein linkers: property, design and functionality." Advanced drug delivery reviews 65.10 (2013): 1357-1369.

Appendix A

Gene Cloning and Propagation

A.1 PCR (typical protocol using Vent Polymerase)

Prepare master mix (50 ml volume): 5 ml Thermopol buffer, 1 ml Vent polymerase, 1 ml dNTPs, 1 ml forward and reverse primers at 100 mM concentration, 1 ml plasmid DNA and 40 ml ultrapure water.

Thermocycling: 94ºC for 2 min to denature, 25 cycles of 94ºC for 30 sec, primer anneal (typically 55-62ºC) for 30 sec and polymerization at 72ºC for 45 sec, and 72ºC for 5 min to ensure all copies are full length.

A.2 Creation of an Artificial Gene by Extension PCR (limited to approx 150-200 bp)

Round 1: generation of small amount product. Mix 2 ml of each oligo (at 0.1 to 10 mM) and add 1 ml to 5 ml Thermopol buffer, 1 ml Vent, 1 ml dNTPs and 42 ml DI. Thermocycling: 30 cycles of 94ºC for 45 sec, anneal for 45 sec and 72ºC for 30 sec, followed by 5 min at 72ºC.

Round 2: amplify product from round 1. Combine 1 ml product taken directly from round 1, 5 ml Thermopol buffer, 1 ml Vent, 1 ml dNTPs, 1 ml each primer at 1-100 mM and 40 ml DI. Thermocycling: same as above.

A.3 Electrophoresis and Gel Extraction
A.4

Prepare 0.5 g agarose (or 1 g low-melting temperature agarose) in 50 ml TAE. Melt agarose in microwave, allow to cool and add 2.5 ml ethidium bromide. Pour into gel box and add comb.

Dilute sample with loading buffer and add to gel (up to 30 ml for 8 well, 10 ml for 15 well). Separate samples using an empty lane unless the product sizes are well seperated.

Run at 105 V for 45 min (1%) or 1.5 hrs (2%).

Cut out bands with a clean razor blade, place in microcentrifuge tube and weigh.

Extract the bands using the Qiagen kit per instructions.

A.5 Plasmid Digestion

Prepare master mix (40 ml): 4 ml digestion buffer, 2 ml enzyme(s), 0.4 ml 100x BSA (if necessary) and 3.6 ml ultrapure water (or 5.6 ml if using only 1 enzyme). Add 28 ml insert from GE or 20 ml vector (+ 8 ml DI) from miniprep.

Update: digest 1000ng DNA, for each enzyme, add 10 Units.

Compare the enzymes' condiontions, and choose the optimal buffer

Incubate at 37ºC for 1 hour. (compare the enzymes' conditions, and pay attention to the temperature. For example, digest DNA with Xbal and BsiWl. BsiW1 requires 55C. Set the temperature at 37C for 1hour and then increase it to 55C for another 1hr.)

Purify vector by electrophoresis as outlined above, except run the gel for at least 1 hour. Inserts can be purified by gel extraction or simply using a Qiagen PCR clean-up kit.

A.6 Ligation

Prepare master mix (20 ml): 4 ml Ligase buffer, 1 ml T4 DNA ligase, 9 ml ultrapure water, 1 ml digested vector and 5 ml digested insert. Ideally should have a 3:1 molar ratio of insert to vector.

Incubate at room temperature for 1-2 hrs (usually 1.5). Note recommended to use at lower temperature.(16C, overnight) Can be transformed directly into E. coli or frozen and used at a later time.

A.7 Transformation into E. coli (based on DH5a from Invitrogen)

Add 2 ml ligation mixture to 50 ml DH5a cells, let sit 30 min on ice.

Heat shock for 45 sec at 42ºC, place back on ice for 2 min and then add 0.95 ml SOC media.

Incubate 1 hr at 37ºC and plate cells onto LB-Amp. If using low transformation efficiency clones, spin sample and plate all cells. Grow overnight at 37ºC.

A.8 Colocy PCR

Scrape a single colony using a pipet tip or small ring and dip into 50 ml ultrapure water.

Perform typical PCR except use ½ the volumes listed above and add 5 ml of the cells suspension. Run PCR as usual with annealing at 55ºC. Check for product on a 1% agarose gel.

Positive samples can be plated (5 ml) and grown up (remaining 40 ml) in 3 ml LB-Amp for later use and sequencing, respectively.

Media Recipes

LB Broth:

Base media (for 1 L): 25 g LB powder, filter sterilize or autoclave.

Supplement: Ampicillin or other antimicrobial agent, typically used at 50 mg/ml

LB Plates:

40 g LB agar powder/L, autoclave.

Supplements: Add ampicillin or other antimicrobial agent (typically at 50 mg/ml) once the media has cooled to 50°C (touchable by bare hand).

Reagent	Company
Restriction enzyme Nhe1	New England Biolabs
Restriction enzyme Xho1	New England Biolabs
100 bp DNA ladder	New England Biolabs
1 kb DNA ladder	New England Biolabs
T4 DNA ligase	New England Biolabs
Vent DNA polymerase	New England Biolabs
SOC outgrowth medium	New England Biolabs
Agarose	Invitrogen
Ethidium bromide	Invitrogen
LB broth base	Invitrogen
LB agar powder	Invitrogen
Ampicillin	Invitrogen
PCR tubes	Fisher Scientific
Razor blades	Fisher Scientific
QIAquick gel extraction kit	Qiagen
QIAprep spin miniprep kit	Qiagen
QIAquick PCR purification kit	Qiagen

Appendix B

Protocol: Yeast Culture and Recombinant Protein Production

B.1 Growth

Streak a single colony on YPD or SD-CAA plates and incubate at 30ºC for 2 days.

Inoculate 3 ml liquid culture with a single colony and grow overnight (YPD) or for 1.5 days (SD-CAA) at 30ºC while shaking at 200-225 RPM.

B.2 Transformation (Lithium Acetate method from Gietz RD and RA Woods, Methods in Enzymology, 2002)

Inoculate 3 ml YPD with BJ5464 (transformation with pRS plasmid) or EBY100 (pCT plasmid) and grow overnight.

Heat single-stranded carrier DNA (2 mg/ml) in a boiling water bath for 5 min and then place on ice.

Spin yeast at 3000 RPM for 1 min, resuspend in 1 ml ultrapure water and transfer to a microcentrifuge tube. Spin again and discard the supernatant.

Layer the following on top of the pellet: 240 ml PEG 3500 (50% w/v), 36 ml LiAc (1.0 M), 50 ml carrier DNA (after vortexing) and 34 ml plasmid DNA (0.1 to 1 mg mixed in ultrapure water, typically use 2 ml). Vortex to mix and resuspend cells (pipet it necessary).

Incubate at 42ºC for 1 hr before centrifugation and aspiration of the transformation mix. Add 1 ml ultrapure water, resuspend and plate 100 ml on selective media (SD-CAA).

B.3 Surface Display of Recombinant Protein

Inoculate 3 ml of SD-CAA with EBY100 yeast transformed with a pCT-based plasmid. Grow overnight at 30ºC while shaking at 225 RPM.

Measure concentration based on absorbance at 600 nm, which is typically 4-6 absorbance units for an overnight culture. 1 absorbance unit is approximately 107 cells/ml.

Transfer 3x107 cells to a microcentrifuge tube, centrifuge, resuspend in 1 ml SG-CAA media, spin, and resuspend in 3 ml SG-CAA.

Grow for 16 hours at 20ºC while shaking at 225 RPM.

B.4 Soluble Expression of Recombinant Protein (small volume assessment).

Inoculate 3 ml of SD-CAA+Ura with BJ5464 yeast transformed with a pRS-based plasmid. Grow 24 hrs at 30ºC while shaking at 225 RPM.

Centrifuge cells, resuspend in 1 ml SG-CAA media, spin, and resuspend in 3 ml SG-CAA+Ura+0.1% BSA.

Grow for 3 days at 20ºC (alternatively grow 2 days at 30ºC) while shaking at 225 RPM.

Recover supernatant by centrifuging at 3000xg for 10 min.

B.5 Soluble Expression of Recombinant Protein (scale-up).

Inoculate 50 ml of SD-CAA+Ura with BJ5464 yeast transformed with a pRS-based plasmid. Grow at least 24 hrs at 30ºC while shaking at 225 RPM.

Centrifuge cells and resuspend in 1 L SD-CAA+Ura, continue to grow overnight (at least 16 hours).

Centrifuge cells and resuspend in 1 L SG-CAA+Ura+0.1% BSA. Grow for 3 days at 20ºC (alternatively grow 2 days at 30ºC) while shaking at 225 RPM.

Place on ice for at least 1 hour to aggregate ammonium sulfate. Recover supernatant by centrifuging at 3000xg for 10 min.

B.6 Cymc staining for flow cytometry

Transfer 2× 106 yeast cells into sterilized eppendorf tube. Add 500ul ice cold PBS+ to wash the cells. Spin down the cells at 3000rpm for 1 min, remove supernatant.

Add 1ul primary antibody (purified anti- cymc Igg1 monoclonal mouse antibody 9E10) in samples for positive reaction.

Add 150ul PBS+ in all the samples, vortex to suspend cells. Vortex samples at room temperature for 30 min.

Centrifuge down the cells at 3000rpm, 1 min. remove supernatant. Wash away excess primary antibody by adding 500ul PBS+ to resuspend the cells and centrifuge. Repeat 3 times.

Add secondary antibody (purified anti-Igg1 mouse antibody- FITC) to all the samples with 150ul PBS+. Wrap the tube with foil to avoid light. Vortex samples at room temperature for 30 min.

Wash away excess secondary antibody by adding 500ul PBS+ to suspend cells and centrifuge. Repeat 3 times. Add 500ul PBS+ to samples after the final wash. Resuspend the cells and transfer to FACS tubes.

Media Recipes

Complex Media: YPD

Base media (for 1 L): 10 g yeast extract, 20 g peptone, 20 g dextrose

Filter sterilize

For plates: Autoclave bacto-agar (15 g/L) in 90% of the final volume. Filter i carbon source, yeast nitrogen base and amino acids in 10% of final volume and add to autoclaved mixture once it has cooled.

Minimal Media: SD-CAA and SG-CAA

Base media (for 1 L): 20 g dextrose or galactose, 6.7 g Yeast N2 Base w/o Amino Acids, 5 g Casamino Acids (-ADE, -URA, -TRP), 5.4 g Na_2HPO_4, 7.46 g NaH_2PO_4

Supplements: Uracil (20 mg/L), Tryptophan (20 mg/L), BSA (1 g/L)

Filter sterilize. Note that the buffer salts do not dissolve well, can autoclave in half of the total volume (subtract ~20 ml for the solids volume).

For plates: Autoclave buffer salts, agar (15-18 g/L) and sorbitol (182 g/L) in 80% of the final volume. Filter sterilize carbon source, yeast nitrogen base and amino acids in 10% of final volume and add to autoclaved mixture once it has cooled below 50°C (touchable by bare hand).

Strains

BJ5464 or BJ5465: supplement with uracil, as well as tryptophan unless transformed with a pRS-based plasmid.

EBY100: supplement with tryptopan unless transformed with a pCT-based plasmid.