

**HISTONE MIMETIC PEPTIDE MEDIATED  
NUCLEAR DELIVERY OF THERAPEUTIC PROTEIN  
TO INFLAMMATORY BREAST CANCER CELLS**

by

Qirun Li

A thesis submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Honors Degree in Chemical Engineering with Distinction

Spring 2021

© 2021 Qirun Li  
All Rights Reserved

**HISTONE MIMETIC PEPTIDE MEDIATED  
NUCLEAR DELIVERY OF THERAPEUTIC PROTEIN  
TO INFLAMMATORY BREAST CANCER CELLS**

by

Qirun Li



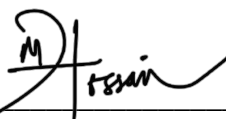
Approved: \_\_\_\_\_

Millicent Sullivan, Ph.D.  
Professor in charge of thesis on behalf of the Advisory Committee



Approved: \_\_\_\_\_

Wilfred Chen, Ph.D.  
Committee member from the Department of Chemical Engineering



Approved: \_\_\_\_\_

Zubaer Hossain, Ph.D.  
Committee member from the Board of Senior Thesis Readers

Approved: \_\_\_\_\_

Michael Chajes, Ph.D.  
Dean, University Honors Program

## **ACKNOWLEDGMENTS**

I would like to thank my advisors, Dr. Sullivan and Dr. Chen, for guidance over these past four years, as well as the opportunity to conduct research in both lab groups; my mentor, Rachel Lieser, for training me to be a researcher and for answering my endless list of questions; my colleagues in the lab, all of you have been a great help throughout my time here; and of course, my BL21s, thank you for expressing all my proteins.

## TABLE OF CONTENTS

LIST OF TABLES .....	vii
LIST OF FIGURES .....	viii
ABSTRACT .....	x
1 THERAPEUTIC PROTEIN .....	1
1.1 Therapeutic Protein Delivery .....	1
1.2 Genomic editing <i>in vivo</i> .....	2
1.3 Histone mimetic peptide.....	2
1.4 New strategies .....	4
2 DELIVERY SYSTEM CONSTRUCTION .....	7
2.1 Delivery System .....	7
2.1.1 Unnatural amino acids modified mCherry protein .....	7
2.1.2 SpyTag-SpyCatcher ligation system on mCherry .....	8
2.2 Material and methods .....	9
2.2.1 Protein construct.....	9
2.2.1.1 Protein expression .....	9
2.2.1.2 Protein purification.....	10
2.2.1.3 Protein characterization .....	11
2.2.1.3.1 Bradford assay .....	11
2.2.1.3.2 Measurements of mCherry concentration ..	11
2.2.1.3.3 SDS-PAGE gel .....	12
2.2.1.4 Protein modification .....	13
2.2.1.4.1 Unnatural amino acids incorporation .....	13
2.2.1.4.2 Ester-amine chemistry .....	14
2.2.2 Peptide .....	16
2.2.2.1 Synthesis strategy .....	16
2.2.2.2 Peptide cleavage .....	17
2.2.2.3 Peptide characterization.....	18
2.2.3 Peptide attachment.....	18

2.2.3.1	Copper-catalyzed click chemistry .....	18
2.2.3.2	Conjugation evaluation.....	20
2.3	Results .....	21
2.3.1	Protein synthesis.....	21
2.3.2	Peptide synthesis .....	22
2.3.3	Click chemistry conjugation.....	25
2.3.3.1	Fluorescent dye conjugation.....	25
2.3.3.2	Peptide conjugation .....	26
2.3.3.3	yCD ligation .....	29
2.3.4	Conclusion.....	30
3	DELIVERY STUDY .....	32
3.1	Introduction .....	32
3.1.1	Inflammatory breast cancer .....	32
3.1.2	yCD.....	33
3.2	Materials and methods.....	34
3.2.1	Cell Maintenance.....	34
3.2.2	Uptake study .....	34
3.2.2.1	Cell Preparation .....	34
3.2.2.2	Protein Delivery.....	34
3.2.2.2.1	Cell internalization study.....	34
3.2.2.2.2	Cell fixing.....	35
3.2.2.2.3	Cell staining.....	35
3.2.2.2.4	Live imaging.....	35
3.2.2.3	Flow cytometry.....	36
3.2.2.4	Fluorescent microscopy .....	36
3.3	Results .....	37
3.3.1	Fluorescent microscopy.....	37
3.3.2	Flow Cytometry.....	42
3.3.3	Confocal .....	45
3.4	Conclusion.....	46

4	ENDOSOMAL ESCAPE .....	47
4.1	Gaining access to the nucleus.....	47
4.1.1	Peptide strategy .....	48
4.1.2	Gal8 visualization.....	48
4.2	Methods.....	49
4.2.1	Construction of GALA construct .....	49
4.2.1.1	Plasmid editing .....	49
4.2.1.2	Agarose gel.....	49
4.2.2	Endosomal escape assay.....	50
4.3	Results .....	50
4.3.1	Diagnostic digestion .....	50
4.3.2	Protein construction.....	51
4.3.3	Delivery study .....	52
4.4	Conclusion.....	53
	REFERENCES .....	55
A	MATLAB SCRIPTS FOR GALECTIN-8 QUANTIFICATION .....	58
B	MATLAB SCRIPTS FOR CELL OUTLINING.....	60
C	Molecular Cloning Primer .....	62

## LIST OF TABLES

Table C1: Primer sequence used in this study.....	62
---	----

## LIST OF FIGURES

<b>Figure 1.1:</b> H3 peptide functionalized mCherry protein, POI denotes protein of interest. ....	5
<b>Figure 1.2:</b> Functionalized GALA-mCherry-SpyCatcher and SpyTag-E2.....	6
<b>Figure 2.1:</b> Chemical structure of p-azido-L-phenylalanine.....	14
<b>Figure 2.2:</b> Ester-amine chemistry attachment of peptides.....	15
<b>Figure 2.3:</b> Mechanism of azidobutyric acid NHS ester reacting with amine group.....	15
<b>Figure 2.4:</b> Chemical structures of side chain protecting groups.....	17
<b>Figure 2.5:</b> Attachment peptide via copper catalyzed azide-alkyne cycloaddition ...	19
<b>Figure 2.6:</b> Mechanism of copper catalyzed azide-alkyne cycloaddition.....	19
<b>Figure 2.7:</b> SDS-PAGE gel showing mCherry purification.....	21
<b>Figure 2.8:</b> Incomplete H3 peptide .....	23
<b>Figure 2.9:</b> Mass spectroscopy verification of the first half of H3 peptides.....	24
<b>Figure 2.10:</b> Mass spectroscopy verification of H3 peptides.....	24
<b>Figure 2.11:</b> Dye conjugation results, upper half showing stained SDS-PAGE gel against white background, lower half showing the same SDS-PAGE gel under UV. ....	25
<b>Figure 2.12:</b> SDS-PAGE gel of H3 peptide conjugation results, where the number refers to the amount of azide functional groups present .....	26
<b>Figure 2.13:</b> Mass spectrometry verification of mCherry conjugation.....	27
<b>Figure 2.14:</b> Conjugation comparison under different temperature condition .....	27
<b>Figure 2.15:</b> Conjugation efficiency with and without sodium sulfate addition.....	28
<b>Figure 2.16:</b> Conjugation efficiency of UAA incorporation compared to ester-amine chemistry .....	29
<b>Figure 2.17:</b> Ligation efficiency results .....	30

<b>Figure 3.1:</b> 5-Fluorecytosine conversion into 5-fluouracil via yCD.....	33
<b>Figure 3.2:</b> Fluorescent microscopy images showing xH3-mCherry-SpyCatcher protein internalization immediately following four hours of internalization.....	38
<b>Figure 3.3:</b> Fluorescent microscopy images showing 4H3-mCherry-SpyCatcher protein internalization following incubation for 24 hours after protein internalization.....	39
<b>Figure 3.4:</b> Fluorescent microscopy images showing protein internalization of xH3-mCherry-SpyCatcher, incubated for additional 20 hours after protein internalization.....	40
<b>Figure 3.5:</b> Fluorescent microscopy images showing protein internalization of xH3-mCherry-SpyCatcher-SpyTag-yCD, immediately following four hours of protein internalization. ....	41
<b>Figure 3.6:</b> Fluorescent microscopy images showing protein internalization of xH3-mCherry-SpyCatcher, immediately following four hours of protein internalization.....	41
<b>Figure 3.7:</b> Flow cytometry results for mCherry-SpyCatcher in IBC cells .....	42
<b>Figure 3.8:</b> Flow cytometry results for mCherry-yCD in IBC cells .....	43
<b>Figure 3.9:</b> Flow cytometry results comparison between IBC cells and normal breast epithelial cells. ....	44
<b>Figure 3.10:</b> Confocal microscopy images showing protein internalization of xH3-mCherry-SpyCatcher, incubated for additional 20 hours after protein internalization.....	45
<b>Figure 4.1:</b> Diagnostic digestion of GALA-GS-mCherry-h6 construct, left most lane is ladder and all others are plasmids from different colonies. ....	51
<b>Figure 4.2:</b> Fluorescent microscopy images showing protein internalization of GALA-xH3-mCherry-SpyCatcher, immediately following four hours of protein internalization. ....	52
<b>Figure 4.3:</b> Fluorescent microscopy images showing gal8 detection of endosomal disruption, immediately following four hours of protein internalization.....	53

## ABSTRACT

Therapeutic proteins are a novel class of biopharmaceuticals, ideal for disease treatment due to their versatility and sophisticated functionality. A major challenge in utilization of therapeutic proteins is delivery of the protein, especially when the protein needs to reach a specific sub-cellular compartment such as the nucleus. This study seeks to improve nuclear delivery of proteins into inflammatory breast cancer (IBC) cells. A histone mimetic peptide (histone H3 tail) was used to target caveolar uptake pathways to gain entry to the nucleus. Click chemistry was used to attach synthetic alkyne-functionalized histone mimetic peptides to azide functionalized fluorescent protein. The protein construct was functionalized either through unnatural amino acids (UAA) incorporation or ester-amine chemistry. The protein also contained a ligation system for easy attachment of therapeutic protein. Additional fluorescent proteins were constructed to add GALA peptide functionalization to promote endosomal escape. Delivery of the H3-conjugated protein construct to IBC cells demonstrated the peptide's ability to enhance protein uptake. Without GALA functionalization, most proteins were trapped in lysosome. Constructs with GALA functionalization exhibited endosomal escape activity based on reporter assays but may require multiple peptides in order to function well. Additional experiments are necessary to determine the relative uptake in IBC cells as compared to normal cells.

## **Chapter 1**

### **THERAPEUTIC PROTEIN**

#### **1.1 Therapeutic Protein Delivery**

Proteins are macromolecules made of amino acid residues, despite its simple basic biochemical makeup, they can catalyze a wide range of reactions. Proteins with bio-catalytical activities are called enzymes and they are the key to most biochemical processes in human body. Hence, dysfunctional enzymes are the origin of many hard-to-treat diseases. Therapeutic protein is a critical area of interest for development of novel therapeutics, especially for cancer, genetic disorders, autoimmune and infectious diseases <sup>[1]</sup>. Some key advantages of utilizing therapeutic proteins are their potency and specificity. With exciting development in directed evolution techniques, the library of therapeutically relevant enzymes is expanding exponentially.

Despite their immense potential, there are multiple challenges associated with protein delivery. Unlike traditional small molecule drugs, proteins cannot freely diffuse across the cell membrane, and it will be readily digested if administered via the GI tract. If administered intravenously, the therapeutic protein can have immunogenicity, with repeat administration often leading to formation of anti-drug antibodies that could interfere with the protein, neutralize its effect, and cause undesired responses to the patient <sup>[1]</sup>.

For these therapeutic proteins to maintain full bioactivity, it is critical to preserve all levels of protein structure and keep correct folding while delivering the

protein to the desired site. Some therapeutic proteins may even require access to specific sub-cellular compartments to be functional <sup>[2]</sup>.

## **1.2 Genomic editing *in vivo***

The clustered regularly interspaced short palindromic repeats (CRISPR)-Cas9 endonuclease system is a powerful tool for genomic editing derived from bacteria immune system <sup>[3]</sup>. Cas9 can be used to knockdown oncogenes or genes that prevent cancer cell death, subsequently induce apoptosis. It has been demonstrated using Cas9 to knockdown expression of poly(ADP-ribose) polymerase-1 (PARP-1) and induce apoptosis in ovarian cancer. PARP-1 is a DNA repair enzyme required for a highly error-prone DNA repair pathway and is often overexpressed in breast and ovarian cancer. In ovarian cancer xenografted mice, significant tumor volume reductions were observed following delivery of Cas9 loaded cancer exosome with guide RNA targeting PARP-1 gene <sup>[2]</sup>.

Current strategies for delivery of Cas9 protein are largely *in vitro*, via delivery of plasmid DNA encoding the constructs. In a clinical trial using Cas9 to treat sickle cell disease, patients would receive autologous hematopoietic stem and progenitor cells edited with Cas9 *in vitro*. This procedure requires patients to undergo myeloablation to suppress bone marrow activity prior to receiving edited cells <sup>[3]</sup>. It may be desired to directly deliver Cas9 enzyme into the nucleus of the cell, which will be degraded by the cells after use.

## **1.3 Histone mimetic peptide**

Cells contain natural mechanisms that shuttle proteins into correct location needed to function. These mechanisms can be hijacked and exploited for transport

of therapeutic protein, thus delivering protein into the nucleus of the cells. One of the main mechanisms for delivering macromolecules into the cell is endocytic uptake.

Two of the most characterized mechanisms for receptor-mediated endocytic uptake are clathrin-dependent endocytosis and caveolae-mediated endocytosis. Clathrin-dependent endocytosis require assembly of clathrin coated pits, of which can bud off from the cell membrane forming a vesicle. The resulting vesicles will fuse with early endosome and go through acidification process that eventually leads to fusing with lysosomes and degradation of the cargo. Caveolae-mediated uptake utilizes a similar mechanism where a vesicle is formed with caveolin proteins. These caveolar vesicles could either fuse with early endosome or caveosomes. The neutral caveosomes allow entry into the retrograde transport pathway. <sup>[4]</sup> While both mechanisms allow entry into the cell, the uptake routes are crucial to the final fate of the cargos delivered. Clathrin vesicles traffic through early endosome to late endosome and eventually lysosome, while the caveolar vesicles may traffic through early endosomes but can directly enter Golgi and endoplasmic reticulum (ER) <sup>[5]</sup>.

Caveolin-1 and -2 are important scaffolding protein in caveolae-mediated endocytosis. In IBC cells, caveolin-1 and -2 promoters are often hypomethylated and will lead to an overexpression of caveolin-1 and -2 <sup>[6]</sup>. While caveolin expression is commonly seen for breast epithelial cells, the levels of caveolin expression are significantly elevated for IBC cells <sup>[5]</sup>. This makes caveolin-1 an attractive therapeutic target for treatment of IBC.

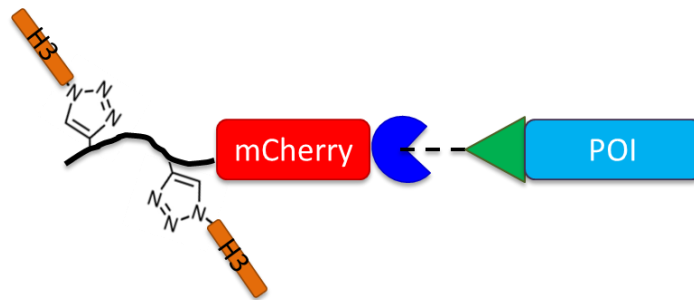
Histone mimetic peptide H3 tail contains mammalian N-terminal residues 1-25 and can utilize caveolar uptake pathways to enter the cell. Previously, our lab group used H3 peptide with poly(ethylenimine) (PEI) to construct polyplexes for DNA

delivery. A mixture of PEI was mixed with H3 peptide solution and desired DNA to form the polyplexes via self-assembly based on charge-interactions. The resulting H3-targeted polyplexes were delivered into IBC cells and normal breast epithelial cells and the polyplexes were able to selectively transfect IBC cells at a 4-fold higher level compared to normal breast epithelial cells [5].

Pulse and chase experiments were done to examine the trafficking routes of the H3 targeted polyplexes in Chinese hamster ovary (CHO) cells. It was found that the H3-targeted polyplexes utilizes both clathrin-dependent endocytosis and caveolae-mediated endocytosis. Some polyplexes were able to avoid endosomal escape and enter the nucleus during postmitotic redistribution of the endoplasmic reticulum membranes [7].

#### **1.4 New strategies**

This project aims to utilize H3 peptide in a protein-based construct via introduction of novel functional groups onto the peptide and protein to enable click chemistry conjugation. The protein construct also contained a ligation system for easy attachment of therapeutic proteins. Azide functional groups can be introduced into the protein construct either via unnatural amino acids incorporation, or ester-amine chemistry. And the terminal alkyne functionalization is introduced via solid state peptide synthesis.

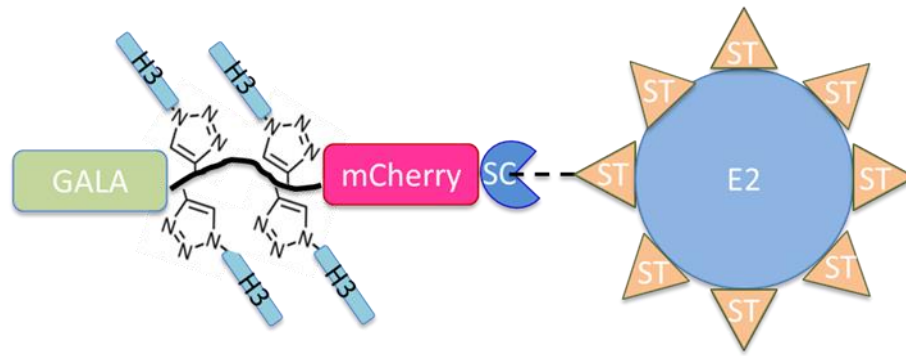


**Figure 1.1:** H3 peptide functionalized mCherry protein, POI denotes protein of interest.

Compared to polyplexes, native proteins are expected to have different trafficking behaviors. H3 peptide contains a nuclear localization signal (NLS), of which should allow active transport of H3 functionalized protein into the nucleus. The NLS can interact with GTP dependent importin protein and thus enter the nucleus in a cell-cycle-dependent manner <sup>[7]</sup>.

In early study, it was found that majority of the proteins delivered are colocalized with lysosome, suggesting entrapment of the protein in endosomal acidification pathway and eventually shuttled into the lysosome. To combat endosomal entrapment and gain access to the cytosol, additional protein constructs were made to add GALA functionalization. GALA peptide is a pH-responsive endosomal escape peptide that changes its conformation under acidic condition, going from a random coil to an  $\alpha$ -helix structure. This conformation change enables GALA peptide to insert itself into the endosomal membrane and thus opening up a pore and allow endosomal escape <sup>[8]</sup>. Multiple GALA peptides are required to achieve endosomal escape, our lab has previously engineered E2 nanocage with ligation tags. This E2 nanocage is a hollow structure with 60 repeat units, which will help increase the

amount of GALA peptides available per construct to increase endosomal escape efficacy <sup>[9]</sup>.



**Figure 1.2:** Functionalized GALA-mCherry-SpyCatcher and SpyTag-E2

## Chapter 2

### DELIVERY SYSTEM CONSTRUCTION

#### 2.1 Delivery System

##### 2.1.1 Unnatural amino acids modified mCherry protein

Our lab previously engineered mCherry, a stable and photobleaching resistant model fluorescent protein, allowing site-specific conjugation of peptides <sup>[10]</sup>. This is achieved via unnatural amino acid incorporation, which introduced novel functional handles into the protein with spatial control via amber stop codon (TAG) suppression. An orthogonal tRNA synthetase able to charge *p*-azido-phenylalanine and the corresponding tRNA sequence able to recognize and suppress amber stop codon are introduced. This pair of tRNA synthetase and tRNA allow incorporation of *p*-azido-phenylalanine in place of amber stop codon <sup>[11]</sup>.

The peptides are attached to the protein using click chemistry, which allows clustering of the desired ligand in a site-specific manner. Ligand clustering has been shown to be highly effective in targeted delivery of protein. Using epithelial growth factor receptor (EGFR) targeting peptide against EGFR-overexpressing inflammatory breast cancer (IBC) cells, an enhanced uptake in IBC cells was observed over normal breast epithelial cells <sup>[10]</sup>.

In addition to unnatural amino acid incorporation, azidobutyric acid EHS ester can react with exposed surface amine groups in mCherry allowing random incorporation of functional ligands through click chemistry. The fluorescent protein enables easy tracking of the cargo location using fluorescent microscopy, while the color of the protein also aids the purification process.

### 2.1.2 SpyTag-SpyCatcher ligation system on mCherry

The mCherry protein was engineered additionally to allow easy attachment of therapeutic protein via fusing mCherry protein with SpyCatcher, a protein partner to a peptide sequence SpyTag. SpyCatcher can react with SpyTag and form an amide bond in near quantitative yield under a diverse pH, temperature, and buffer conditions. Since the ligation requires SpyTag peptide sequence, this reaction tolerates other functional groups well <sup>[12]</sup>. This SpyCatcher-SpyTag ligation system has been previously used to incorporate yeast cytosine deaminase (yCD), a prodrug conversion enzyme, into the mCherry construct. Delivery of EGFR targeted mCherry-yCD construct into IBC cells, combined with the prodrug, were able to elicit a 3-fold difference in cell death between normal breast epithelial cells and IBC cells. <sup>[10]</sup>. In addition, attachment of yCD to mCherry construct did not compromise delivery significantly.

Another benefit of SpyCatcher system is attachment to SpyTag functionalized E2 nanocage, a 60-mer self-assembling nanocage, derived from pyruvate dehydrogenase complex, which assembles into a hollow cage structure with 12 openings <sup>[13]</sup>. Our lab has previously engineered the E2 nanocage to display SpyTag, allowing ligation of SpyCatcher containing construct onto the E2 nanocage, drastically increase the numbers of available functional ligands per construct <sup>[9]</sup>.

Some challenges are present in construction of the H3-functionlized protein construct. Most significant impact on cell uptake level of the H3 functionalized protein construct is the numbers of H3 peptide successfully conjugated via click chemistry. H3 peptides has eight positive surface charge at neutral pH and has an isoelectric point around pH over 11. Thus, the conjugation of H3 peptides clustered together on protein construct could be impacted by charge repulsion. Failure to incorporate several H3

peptides would lead to a sharp drop in uptake level in both IBC cells and normal breast epithelial cells. In addition, it is worth noting that yCD is a relatively small protein, while nuclear delivery may require a much larger therapeutic enzyme such as Cas9, which could impact ligation process and the subsequent delivery.

## **2.2 Material and methods**

### **2.2.1 Protein construct**

#### **2.2.1.1 Protein expression**

A protease deficient strain of *E. coli*, BL21(DE3) was used to carry out protein expression. This strain also contained the necessary machinery for T7 expression of recombinant proteins. Competent BL21 cells were made using calcium chloride methods and transformed with the desired plasmid DNA using heat shock methods. The resulting cells were cultured on agar plate containing antibiotic selection agents and incubated at 37 °C overnight.

Following selection, a random colony was picked off the plate and added to Terrific Broth (TB) media solution containing antibiotics in a test tube to make a pre-culture. The pre-culture was shaken and incubated at 37 °C for several hours. The resulting culture was diluted with additional TB media and antibiotics in a flask, and incubated at 37 °C for one to two hours, until optical density (OD) reach between 0.6-0.8. Protein expression was then induced via addition of Isopropyl  $\beta$ -d-1-thiogalactopyranoside (IPTG). If incorporating UAA, also add p-azido-phenylalanine into the solution, also include a culture without addition of UAA as control. The product cultures were incubated overnight, the ODs were measured, and the cells were collected via a centrifuge. The resulting cells were resuspended in 1x phosphate

buffered saline (PBS), and disrupted in a probe sonicator. Cell lysate were obtained via a centrifuge, and the dead cells were discarded.

#### **2.2.1.2 Protein purification**

The mCherry-SpyCatcher construct contained a 6xHis tag, of which can coordinate to metal ions strongly, allowing purification through nickel charged nitrilotriacetic acid (Ni-NTA) resin. Following resuspending the *E. coli* biomass in 1x PBS, prior to sonication, 1 M imidazole solution in 1x PBS was added to make the overall concentration of the imidazole in the cell solution to be approximately 10 mM. The column is packed with Ni-NTA resin and equilibrated with 10 mM imidazole solution before loading in the supernatants from homogenization. Flow through was collected for further analysis. The column was then washed with 20 mM imidazole solution to elute any non-specific binding. Finally, the desired protein was eluted using 250 mM imidazole, since mCherry protein had its signature pink color, the movement of the desired protein construct was visible in the column. Following eluting the product, the column was regenerated using 2-(N-Morpholino)ethanesulfonic acid(MES) buffer and deionized water. Regenerated column was stored in a 50% slurry consisted of the resin and 20% ethanol in water at 4 °C until next use. Columns were only used to purify one specific type of the construct, respectively, to avoid cross contamination. The resulting protein elutes were dialyzed in 1x PBS overnight, and stored in the fridge for further analysis.

### **2.2.1.3 Protein characterization**

#### **2.2.1.3.1 Bradford assay**

To measure protein concentration in the supernatant produced from the homogenized cell solution, four dilutions of protein standards are constructed via diluting bovine serum albumin (BSA) solution using serial dilution technique, covering a range from 0.5-0.0625 mg/mL of BSA. 10  $\mu$ L of the 10x diluted supernatant samples and the BSA standards were added to three wells per sample, respectively, followed by 200  $\mu$ L of Bradford dye in each well. The resulting solutions were incubated for 5 minutes before measuring absorbance at 595 nm, the readings were averaged between the three wells containing the same sample. The BSA standard is used to construct a calibration curve, which was then used to calculate the protein concentration in the supernatant.

#### **2.2.1.3.2 Measurements of mCherry concentration**

The fluorescence measurements of supernatants were obtained at 610 nm to verify the expression of mCherry. The purified protein elutes from the Ni-NTA column was dialyzed and the absorbance is tested using a Nanodrop spectrophotometer at 280 nm, and the concentration is calculated from beer-lambert law.

$$A = \epsilon bc \quad (2.1)$$

Where A is absorbance,  $\epsilon$  is molar extinction coefficient, in  $M^{-1}cm^{-1}$ , b is pathlength, in cm, and c is concentration in M. The molar extinction coefficient is obtained by expressing a large amount of the protein, followed by lyophilization and resuspension in 1x PBS to construct a standard curve, where the molar extinction coefficient can be calculated. For subsequent modification, the molar extinction coefficient for the non-

modified protein is used to estimate the concentration of the modified protein, which may be less accurate and likely an overestimation since the modification should increase molar extinction coefficient due to addition functional groups that strongly absorb in the aromatic region.

#### **2.2.1.3.3 SDS-PAGE gel**

Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) is used to analyze the protein solutions, by separation of protein in the gel according to size difference of the protein. The gel is constructed from acrylamide solution in tris buffer, SDS and water, polymerized using ammonium persulfate (APS) oxidizing tetramethylethylenediamine (TEMED), which is a radical stabilizer, the resulting free radical will in turn promotes acrylamide polymerization. The gel was separated into two regions, one at pH 6.8 with lower concentration of acrylamide, being the stacking gel on top, allowing sample loading with easy movement of protein through the gel and prepare entry into the resolving gel on the bottom. The resolving gel was at pH 8.8, and contained a higher concentration of acrylamide, allowing protein to move through at different rate based their size difference.

To prepare protein solution, the samples were diluted to no more than 1  $\mu\text{g}/\mu\text{L}$  in protein as determined from the Bradford assay, for analysis of the protein purified from the supernatant, and all subsequent modification, the concentration is scaled based on the volume ratio to the original supernatant and diluted accordingly to keep the concentration consistent across the gel. A dye solution containing SDS, dithiothreitol (DTT), and bromophenol blue is added to the diluted protein solution, SDS will coat the protein once denatured, DTT will reduce disulfide bonds in protein, and bromophenol blue will indicate the progress of protein movement in the gel. The

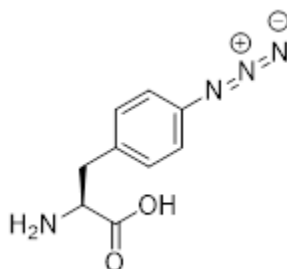
resulting solutions were mixed and heated at 95 °C for at least 10 minutes to denature the protein. The resulting protein should be a linear molecule coated with SDS, and thus will only be differentiated by their relative size in SDS-PAGE. The denatured protein solutions were loaded into the gel and electrophoresis were performed in SDS buffer. The voltage was kept at 60 V until the bromophenol blue front reaches the end of the stacking gel and raised to 120 V until the bromophenol blue front is near the edge of the gel or coming out of the gel. The finished polyacrylamide gel was stained in Coomassie dye solution, which can stain both the proteins and the gel itself, the proteins are visualized by de-staining the gel in a mixture of methanol, acetic acid and water. Destained gel is stored in the de-staining solution and imaged using a gel documentation system.

#### **2.2.1.4 Protein modification**

##### **2.2.1.4.1 Unnatural amino acids incorporation**

Unnatural amino acids are incorporated into the protein via amber stop codon (TAG) repression, which is the least used stop codon and rarely terminates essential genes<sup>[11]</sup>. A plasmid is introduced into BL21 encoding nonsense suppressor tRNA and aminoacyl-tRNA synthetase that can charge p-Azido-L-phenylalanine onto the nonsense tRNA, from there, azido functional group can be incorporated with spatial control. Existing mCherry constructs obtained from Lieser et al<sup>[10]</sup> containing either one, two or four amber stop codons separated by flexible linker sequence. In place of amber stop codon, p-azido-L-phenylalanine (pAzF) could be incorporated, and since the amber stop codons were placed at the N-terminus, while the 6xHis tag was placed at the C-terminus, unsuccessful incorporation of desired amount of pAzF will lead to

premature termination of the polypeptide synthesized, which will not have the purification tags attached.

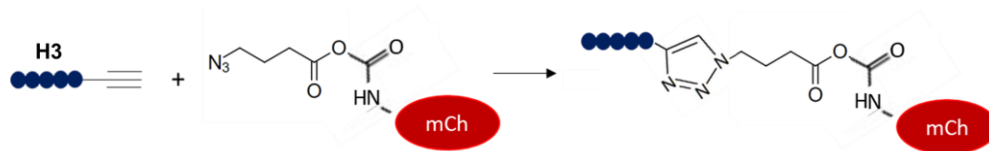


**Figure 2.1:** Chemical structure of p-azido-L-phenylalanine

Through UAA incorporation, azide function groups were introduced into mCherry protein clustered at the N terminus of the protein, separated by short flexible linkers. During expression of the protein, along with addition of IPTG, pAzF stock solution is added to make the final concentration reach 1mM per amber stop codon in the construct.

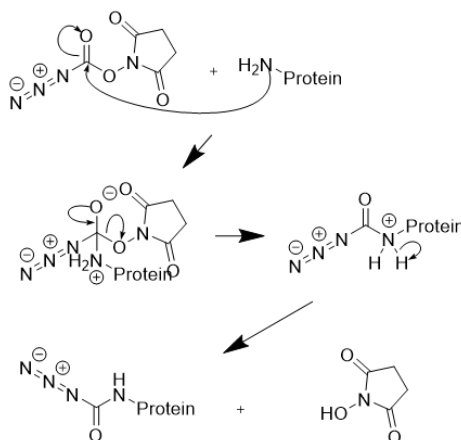
#### **2.2.1.4.2 Ester-amine chemistry**

In addition to amber stop codon containing construct, another amber stop codon free mCherry construct is made, leaving only the flexible linker sequence. This construct is denoted GS-mCherry-SpyCatcher, and subsequently used for ester-amine chemistry.



**Figure 2.2:** Ester-amine chemistry attachment of peptides

In this reaction, the surface amine, sourced from exposed lysine residues on mCherry protein, is reacted with azidobutyric acid NHS ester. Azido groups introduced by NHS ester were randomly spread out on the protein surface, with little spatial control.



**Figure 2.3:** Mechanism of azidobutyric acid NHS ester reacting with amine group.

In addition, since the surface amine are critical to protein solubility, subsequent incorporation of peptides through click chemistry could greatly impact the solubility of the protein construct. It was reported an average of two GE11 peptides can be incorporated into the mCherry construct without compromising protein solubility <sup>[10]</sup>.

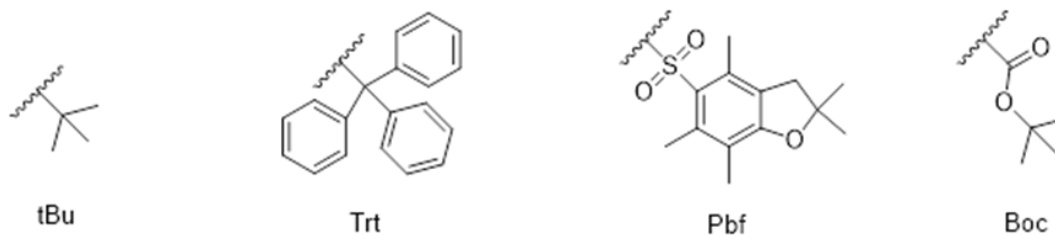
In experimentation, the numbers of H3 peptides tolerated are greater since H3 peptide is more water soluble than GE11 peptide.

## **2.2.2 Peptide**

Peptide synthesis is a highly repetitive and labor-intensive process, which led to systems being developed to automate the process. Using a peptide synthesizer, the H3 peptide was synthesized via solid state peptide synthesis (SSPS). The resulting peptides were cleaved from solid support, purified using high performance liquid chromatography (HPLC), the resulting fractions were lyophilized. The mass of purified H3 peptide was verified via mass spectrometry.

### **2.2.2.1 Synthesis strategy**

Starting with solid resin support, one amino acid residue would be added per cycle, the peptide was synthesized from C-terminus to the N-terminus. The N terminus of amino acids were protected with the base labile Fmoc (fluorenylmethyloxycarbonyl) group, and the sidechain functional groups were protected with acid labile groups. For most amino acids used, the side chain protecting groups were commonly tBu (tert-butyl) group, with some being Boc (tert-butyloxycarbonyl) group, Trt (trityl) group or Pbf (2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-sulfonyl) group, respectively. The structures of side chain protecting groups are shown in Figure 2.1, with Trt group and Pbf group being bulkier than the other groups, suggesting a harder removal later.



**Figure 2.4:** Chemical structures of side chain protecting groups.

For each cycle, an Fmoc protected amino acid was introduced into the system with the C terminus exposed. The exposed C terminus would then be coupled to the exposed N terminus of the existing peptide (or linker, for the starting amino acid), with the help of coupling agent HBTU. Following attachment of the new amino acid residue, the N terminus of the new peptide would be Fmoc protected. From there, a mild base, piperidine, was introduced to remove the Fmoc group, getting the peptide ready to move onto the next cycle.

Following this cycle, a modified version of the H3 peptide is synthesized with the complete sequence being ARTKQTARKSTGGKAPRKQLATKAAG', where G' denotes propargylglycine, which contained a terminal alkyne for click chemistry conjugation.

### 2.2.2.2 Peptide cleavage

After synthesis cycles have finished, the resulting resins were lyophilized and stored in the fridge, with peptide still attached to the linker. Peptides need to be cleaved and removed from solid resin support for subsequent purification. Since linker between the resin and the peptide and the side chain protecting groups all acid labile, thus the cleavage is done in acidic condition, which would also remove side chain

protecting groups. A cleavage cocktail is made from 95% trifluoroacetic acid for cleavage, 2.5% Triisopropyl silane for scavenging peptide groups removed, with the rest 2.5% being HPLC grade water. The cleavage is done for 3 hours at room temperature, and the resulting solution with resins were filtered through a syringe with glass filter. In a fume hood, the solution is evaporated by blowing nitrogen gas. Cooled ethyl ether was used to precipitate the peptides, and the solution was centrifuged, and ethyl ether is poured off. The resulting peptides are dried by blowing nitrogen gas, redissolved in water, and lyophilized to obtain dried peptide.

### **2.2.2.3 Peptide characterization**

Lyophilized peptides are purified via high performance liquid chromatography (HPLC), using acetonitrile in water, through reverse phase chromatography, high concentration of water is used initially, slowly increase the percentage of acetonitrile used to elute the peptide. The output solution from HPLC was subjected to a spectrophotometer, and the major peaks were collected manually based on real time spectrophotometer reading. The resulting fractions were lyophilized and verified via mass spectroscopy.

### **2.2.3 Peptide attachment**

#### **2.2.3.1 Copper-catalyzed click chemistry**

Azide and terminal alkyne can undergo copper-catalyzed azide-alkyne cycloaddition (CuAAC) to form a highly stable triazole ring. This reaction is fast, efficient, specific, and able to tolerate many different functional groups, it can be done bioorthogonally and is ideal for attachment of biomolecules.



containing ligands were used since the purification process would also retain unreacted mCherry protein. For the peptide attachment, the concentration of all catalysts was kept consistent, where the mCherry concentration was lowered to 20  $\mu\text{M}$ , and the peptide concentration was lowered to 100  $\mu\text{M}$ .

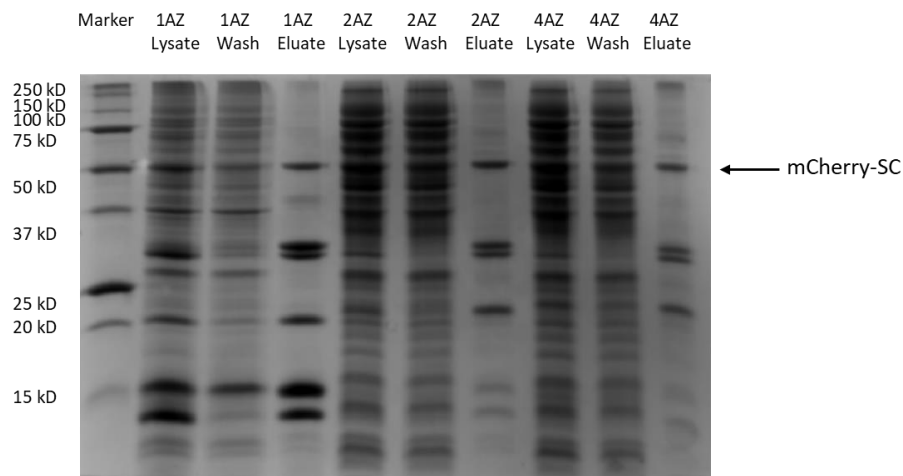
### **2.2.3.2 Conjugation evaluation.**

Conjugation efficiency was evaluated via SDS-PAGE, by comparing the protein before and after reaction. A much lower conjugation efficiency was observed for H3 peptide compared to GE11 peptide when reacted to the same batch of protein. It was then proposed that since H3 peptide contains eight positive surface charge under neutral condition, compared to no charge on GE11 peptide, the repulsion force of like charge could have impacted the conjugation efficiency. In ionic solution, charged species can form an electrical double layer around charged biomolecule, effectively shield the charges in a certain radius, of which the effective radius where electrostatic effects can persist is referred to as Debye length refers electrostatic effects can persist. The reaction proceeds in 1x PBS, of which has a Debye length of approximately 0.7 nm<sup>[13]</sup>, and the length of (GGGGS)<sub>3</sub> linker was calculated to be about 5.7 nm<sup>[14]</sup>. Thus, the GGGGS linker used in our construct should have a length of approximately 1.9 nm, of which is outside of the Debye length of the solution, but since the linker is flexible and H3 is highly charged, the charged species in 1x PBS may not be enough to completely shield the charge. An increase in ionic species was achieved via addition of sodium sulfate to the reaction mixture, of which improved conjugation efficiency. A final concentration of approximately 0.7 mM of sodium sulfate was used and is well tolerated by the protein, a high concentration of sodium sulfate could lead to protein denaturation.

## 2.3 Results

### 2.3.1 Protein synthesis

The proteins were synthesized using BL21 bearing desired plasmids, encoding for UAA incorporation machineries and xAz-mCherry-SpyCatcher construct, where the numbers of azide groups incorporated into the protein varies. Following expression of desired protein, the cells were homogenized and centrifuged to obtain the cell lysate. The lysate was purified through Ni-NTA resin column and the flowthrough from the lysate as well as the washing step were collected and analyzed. The eluates were collected and dialyzed in 1x PBS solution before analysis via SDS-PAGE.



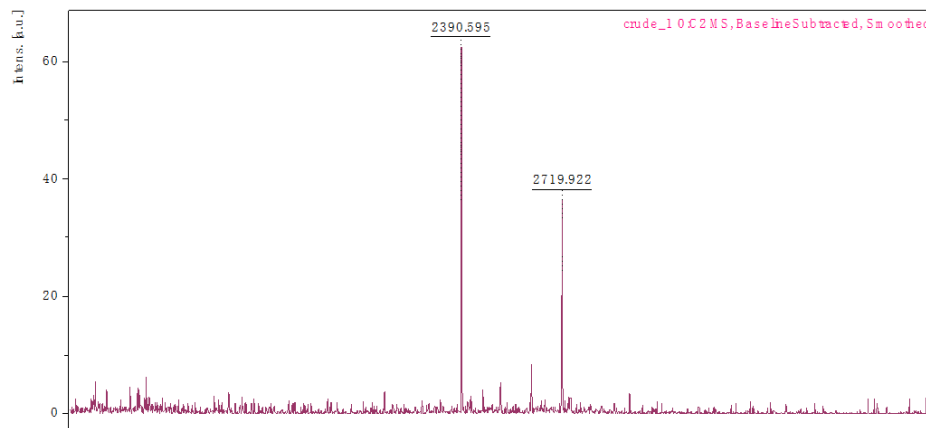
**Figure 2.7:** SDS-PAGE gel showing mCherry purification.

Compared to the lysate, the elution from the Ni-NTA column was successfully purified, eliminating the majority of the genomic protein. In the eluate solution, besides the major mCherry-SpyCatcher construct band, which has a molecular weight

of approximately 50 kD, other smaller bands can be observed as well, including truncated mCherry protein that still containing the 6xHis tag.

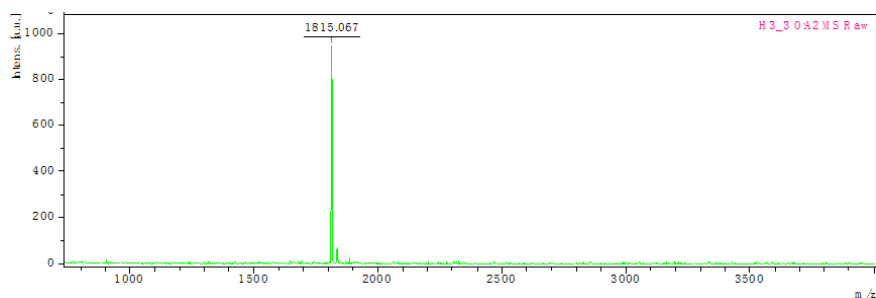
### **2.3.2 Peptide synthesis**

Initial attempts of peptide synthesis resulted in incomplete peptide sequences. Mass spectrometry was used to identify the major product isolated and a brief calculation was done using MATLAB to narrow down the identity of the major production peak. Assuming several amino acids residues were not incorporated, list of possible product masses were generated by removing amino acids from the complete sequence. Generated possible product mass list were searched through to find masses that are close to major product peak, and several possibilities were present, most possibilities suggested threonine and serine being the lost residues. Fresh reagents were used for the next attempts, and the coupling time for threonine and serine were increased from the standard coupling procedure. It is assumed that the incomplete peptide does not contain any remaining Fmoc or other protective groups. This assumption was made based on the cleavage cocktail used should be able to completely remove all protecting groups. However, it is worth noting the major product peak identity could also be incomplete peptides with protecting group still attached, if the deprotection was not complete during peptide cleavage.



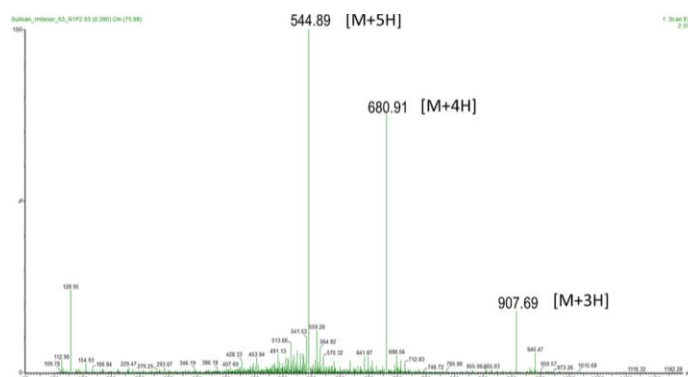
**Figure 2.8:** Incomplete H3 peptide

Threonine used in peptide synthesis was protected by a bulky Trt group and was also identified as one of the problematic amino acid residues to incorporate in previous synthesis. In addition, the second threonine residue was approximately in the middle of the complete peptide sequence. Therefore, the next attempt at H3 peptide synthesis was paused following incorporation of the second threonine in the sequence, a minimal amount of incomplete peptide was obtained, cleaved from the resins, and purified. The incomplete sequence should contain Fmoc-TGGKAPRKQLATKAAG', with the expected mass being 1814.871 g/mol, matching the high intensity experimental peak at 1815.067. The Fmoc protecting group is still present since the synthesis was paused prior to deprotection of the last incorporated residue.



**Figure 2.9:** Mass spectroscopy verification of the first half of H3 peptides

Following verification of the first half, the synthesis was allowed to proceed and the resulting peptides were cleaved off the resin and purified via HPLC. Major product peaks were collected, lyophilized, and the masses were verified via mass spectroscopy. The final targeted sequence is ARTKQTARKSTGGKAPRKQLATKAAG', which should have a mass of 2718.577, this corresponds to the major product peaks shown.



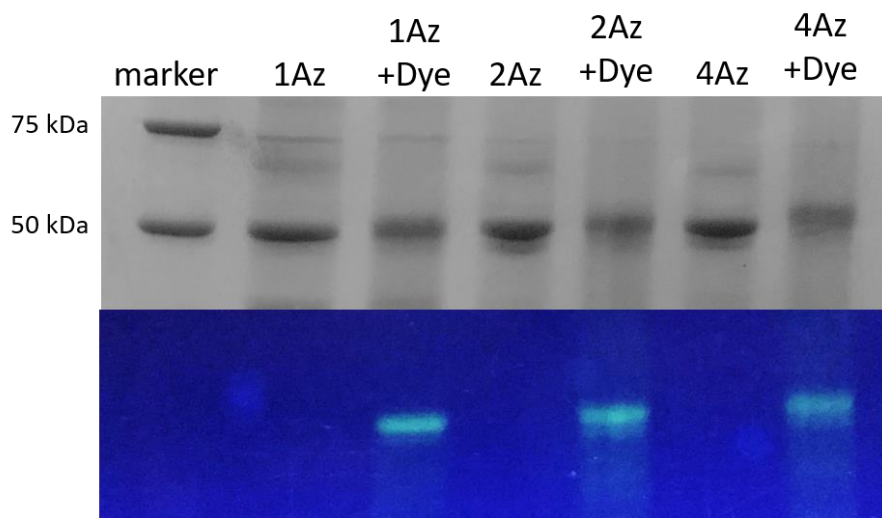
**Figure 2.10:** Mass spectroscopy verification of H3 peptides.

Lyophilized peptides from HPLC purification was redissolved in a mixture of water and acetonitrile, with minimal amount of DMSO, reaching a final concentration of 5 mM. The resulting solution was aliquoted into small vials and stored at -20 °C.

### 2.3.3 Click chemistry conjugation

#### 2.3.3.1 Fluorescent dye conjugation

Alexa 488 dye with a terminal alkyne functionalization is reacted with the purified xAz-mCherry-SpyCatcher construct bearing different numbers of UAA incorporation. This step is done to verify the reactivity of the azide functional group in the pAzF incorporated.

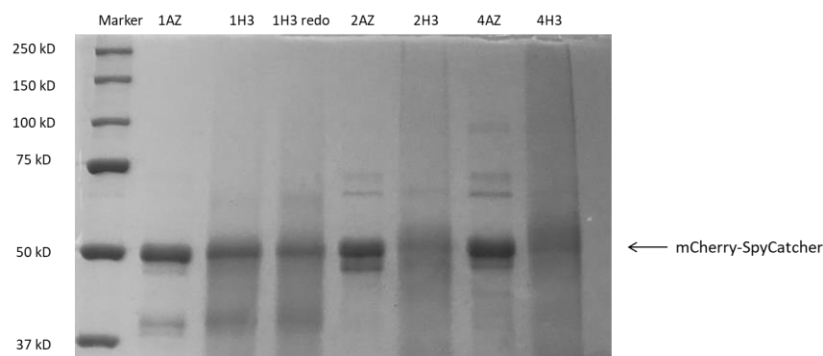


**Figure 2.11:** Dye conjugation results, upper half showing stained SDS-PAGE gel against white background, lower half showing the same SDS-PAGE gel under UV.

Following click chemistry conjugation, the Alexa 488 fluorescent dye can be detected in the major protein band, suggesting successful conjugation, and thus confirming reactivity of the expressed protein.

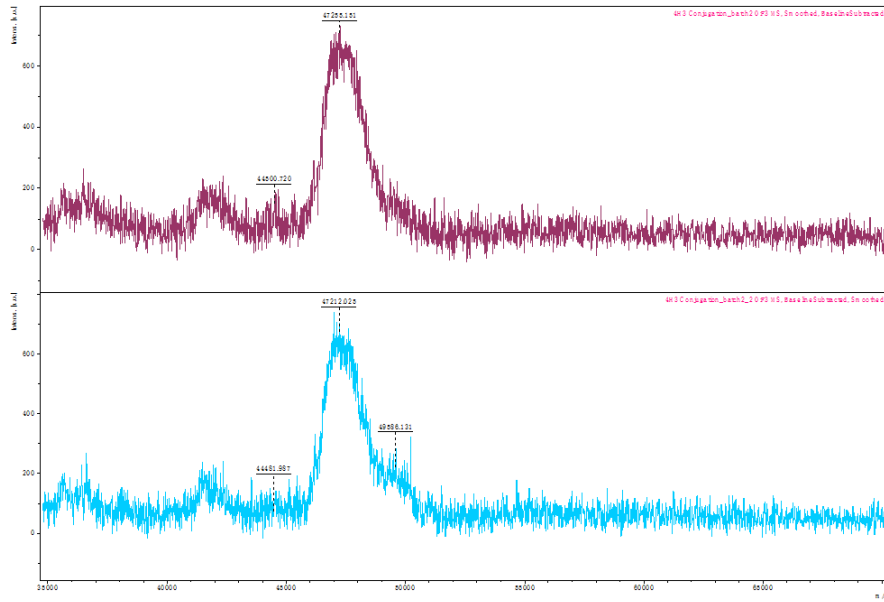
### 2.3.3.2 Peptide conjugation

Synthesized H3 peptide with terminal alkyne functionalized was reacted with xAz-mCherry-SpyCatcher, the resulting solution were analyzed along with the original stock solution of mCherry to evaluate conjugation efficiency.



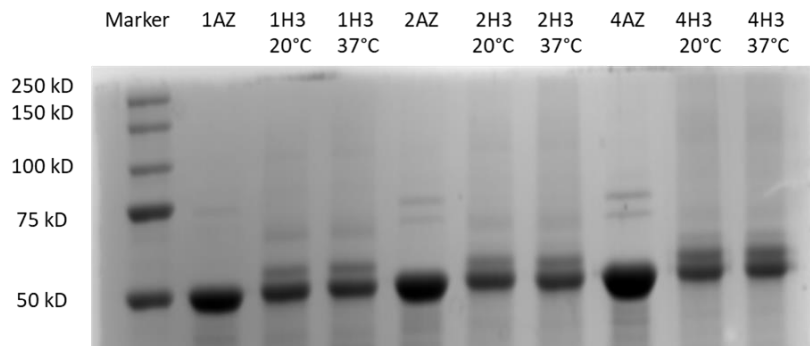
**Figure 2.12:** SDS-PAGE gel of H3 peptide conjugation results, where the number refers to the amount of azide functional groups present

A significant increase in mass was observed for the 4Az group, suggesting successful conjugation of the peptides. However, only limited success was achieved in conjugation in following attempts. Poor conjugation efficiency was observed across the board despite different numbers of azide functional groups present in the protein, with mostly one peptide being attached as verified by mass spectrometry.



**Figure 2.13:** Mass spectrometry verification of mCherry conjugation

From here, multiple attempts were done to troubleshoot the reaction, including changing the reaction pH, temperature, protein/peptide concentration, and adding zinc sulfate to supply sacrificial metal.



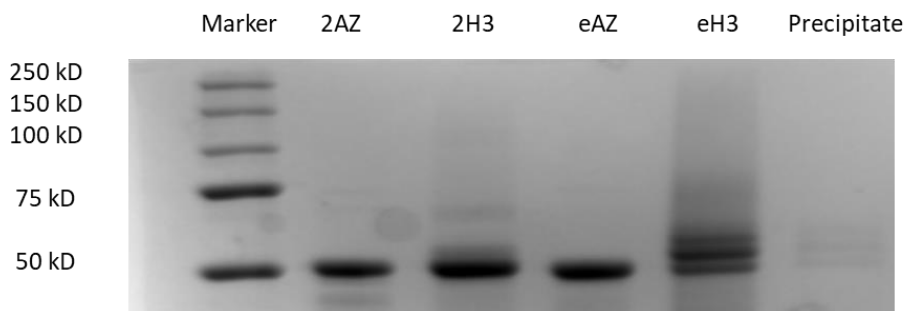
**Figure 2.14:** Conjugation comparison under different temperature condition

It was observed that conjugation efficiency remained low while one peptide can be attached, combined with H3 having 8 positive surface charge, it was proposed charge repulsion is the reason for the poor conjugation. Addition of sodium sulfate was able to improve the conjugation efficiency.



**Figure 2.15:** Conjugation efficiency with and without sodium sulfate addition

The reaction protocol was modified for all subsequent conjugation to improve conjugation efficiency. In addition, ester-amine chemistry is used to introduce azide functional group, where a high number of peptides can be incorporated but also compromised protein solubility. Upon trial and error, a molar ratio of 5 is used for subsequent reaction to reduce precipitation. The amount of peptide tolerated by mCherry protein is much higher for H3 peptide due to its high solubility in water, when compared to the relatively more insoluble GE11 peptide.

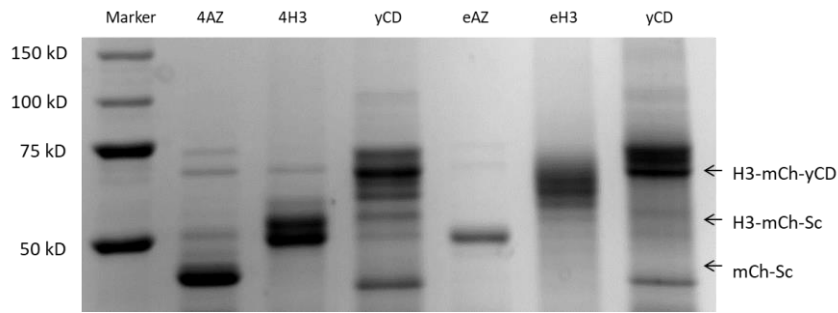


**Figure 2.16:** Conjugation efficiency of UAA incorporation compared to ester-amine chemistry

H3 conjugated protein is purified via Ni-NTA resin column, dialyzed in 1x PBS and stored in 4 °C.

### 2.3.3.3 yCD ligation

Our lab previously constructed yCD-SpyTag, of which can be ligated onto mCherry-SpyCatcher protein in quantitative yields under mild conditions. The yCD-SpyTag protein was synthesized in BL21, extracted via sonication, and stored in cell lysate. Since no purification tag is present in the yCD protein, the cell lysate will be directly used for reaction with mCherry-SpyCatcher construct, the resulting solution was incubated in 4 °C overnight, and purified via Ni-NTA column, dialyzed and analyzed.



**Figure 2.17:** Ligation efficiency results

The ligation can be done rapidly at room temperature in several hours based on reported kinetics, but in operation, it is easier to react in the fridge overnight. This is done to achieve near complete ligation to ensure most SpyCatcher is reacted, since unreacted mCherry protein will still have the 6xHis tag and will be retained by the Ni-NTA column. Ligation of yCD-SpyTag did not compromise purification.

### 2.3.4 Conclusion

Protein constructs were expressed in BL21, with azide functional group incorporated either through UAA incorporation or ester-amine chemistry. The reactivity of the UAAs incorporated was verified by click chemistry conjugation with fluorescent dye. H3 peptides were synthesized with terminal alkyne functionalization, purified via HPLC, and stored in frozen stock solution.

For click chemistry conjugation, the construct functionalized by ester-amine chemistry was able to incorporate more peptides than construct using UAA incorporation. For the UAA incorporation construct, poor H3 conjugation efficiency was seen with an average number of one H3 peptide incorporated. A range of different experiments were done under different pH, temperature, and buffer conditions to seek

increase in conjugation efficiency. It was proposed charge repulsion was the major barrier for peptide conjugation, which led to an attempt of increasing salt concentration via direct addition of sodium phosphate solution. The resulting functionalized protein were purified and dialyzed in 1x PBS solution. Concentration of mCherry-SpyCatcher in the dialyzed solution was determined based on absorbance measurement and kept at 4 °C until use.

## **Chapter 3**

### **DELIVERY STUDY**

#### **3.1 Introduction**

##### **3.1.1 Inflammatory breast cancer**

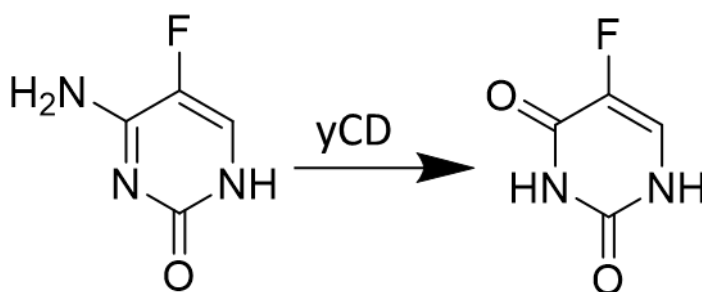
IBC is a very aggressive type of breast cancer, representing around 1% to 6% of all breast cancer and often has poor prognosis <sup>[15]</sup>. IBC is marked by cancer cells blocking lymph vessels, which is the driver of the apparent inflammation. IBC progresses in a matter of weeks or months; at diagnosis, IBC is either stage III or IV disease <sup>[16]</sup>. Current treatment of IBC includes chemotherapy, surgery and radiation therapy, 5-year survival rate has been improved from below 5% to around 30 % with novel therapeutics <sup>[15]</sup>.

Newer strategies for IBC treatment focus on targeted therapy. For breast cancer, the most characterized molecular target is estrogen receptor, leading to development of hormonal therapeutics. However, more than half of IBC patients diagnosed are estrogen receptor negative <sup>[15]</sup>, eliminating the possibility for hormonal treatment. Epidermal growth factor receptor family proteins are also expressed commonly on IBC, notably EGFR and HER2. The expression of these proteins has been reported to inversely correlate with estrogen receptor status <sup>[16]</sup>.

The targeting strategy in this study is focused on H3 interaction with caveolar uptake pathways. They are commonly overexpressed in tumor associated endothelial cells. Caveolin-1 and -2, structural proteins necessary for caveolae formation, are overexpressed in IBC cells <sup>[5]</sup>, making caveolin-1 and -2 suitable therapeutic target. It is also worth noting that. Normal breast epithelial cells also express caveolin-1, but the level of expression is significantly lower compared to IBC cells <sup>[5]</sup>.

### 3.1.2 $\gamma$ CD

Yeast cytosine deaminase is a prodrug conversion enzyme. It can catalyze the conversion of 5-fluorocytosine (5-FC), a nontoxic antifungal medication; into 5-fluouracil (5-FU), an FDA approved chemotherapeutic agent. 5-FU is extensively used for treatment of breast cancer among others. 5-FU is structurally similar to uracil and can be processed in place of uracil to produce a range of active metabolites. These active metabolites can participate in RNA synthesis and incorporate 5-FU in place of uracil and disrupt subsequent function of the RNA in cellular process <sup>[17]</sup>. In addition, thymidylate synthase, an essential enzyme catalyzing the conversion of deoxyuridine monophosphate to deoxythymidine monophosphate, can be inhibited by one of the active metabolites produced from 5-FU <sup>[17]</sup>. This inhibition of thymidylate synthase will lead to depletion of thymine the cell and cause DNA damage in rapidly dividing cells.



**Figure 3.1:** 5-Fluorecytosine conversion into 5-fluouracil via  $\gamma$ CD

Following delivery of EGFR targeted mCherry- $\gamma$ CD construct into IBC cells, 5-FC was supplied, and resulted in a 3-fold difference in cell death between the normal breast epithelial cells and IBC cells <sup>[10]</sup>.

## **3.2 Materials and methods**

### **3.2.1 Cell Maintenance**

SUM149 cell line was used as model for IBC cells, and MCF10A cell line was used as model for normal breast epithelial cells. Both cell lines were passaged every week. Cells were washed with Dulbecco's phosphate-buffered saline (DPBS), trypsinized with 0.25% trypsin, neutralized with media solution. 10% of the cells were seeded back into another flask while the rest were seeded in plates for delivery study.

The media used is Ham's F-12 for SUM149, and a mixture of Ham's F-12 and Dulbecco's Modified Eagle Medium (DMEM) for MCF10A. Both media were supplemented with fetal bovine serum for cell growth.

### **3.2.2 Uptake study**

#### **3.2.2.1 Cell Preparation**

Prior to seeding, the plates were prepped with collagen solution for two hours to allow attachment of the cells.

Following obtaining trypsinized IBC cells and normal breast epithelial cells in their respective media, the cells were stained with trypan blue and counted using a hemocytometer under a microscope. Numbers of cells were calculated and seeded into 8-well glass plates at 50,000 cells/cm<sup>3</sup>.

#### **3.2.2.2 Protein Delivery**

##### **3.2.2.2.1 Cell internalization study**

The seeded plates of IBC cells and breast epithelial cells were incubated overnight in their respective media. The desired protein were prepared by diluting

protein down to 1-3  $\mu\text{M}$  in warm media solution. The spent media was replaced with diluted protein solutions and incubated for four hours. Protein solutions were aspirated off and the cells were washed three times with 1x DPBS. Depending on the difference in total incubation time, the cells were either resuspended in 1x DPBS for immediate fixing, or resupplied fresh media solution for longer incubation periods.

#### **3.2.2.2.2 Cell fixing**

All operation with cell fixing steps is in a fume hood except the shaking step, of which the cells are covered by glass caps and stored in a secondary container covered with tin foil. All wastes were disposed in designated formaldehyde chemical waste box. The cells were fixed by incubating at room temperature with 4% paraformaldehyde buffer solution and shaken for 15 minutes. Spent solution was aspirated off and collected in a designated waste container. The cells were washed with 1x DPBS three times and stained with DAPI dye.

#### **3.2.2.2.3 Cell staining**

All steps with fluorescent stains were done while protected from light to avoiding photobleaching. DAPI stain was diluted to 300  $\mu\text{M}$  in 1x DPBS and incubated with fixed cells for 5 minutes. The resulting cells were washed with 1x DPBS three times and resuspended in 1x DPBS for imaging.

#### **3.2.2.2.4 Live imaging**

For live imaging, lysotracker and nuclear stain were diluted in cell media, and incubated with live cells on the shaker for 5 minutes. Resulting cells was resuspended in live cell imaging solution and imaged under the fluorescent microscope immediately.

### **3.2.2.3 Flow cytometry**

Cells were plated for flow cytometry by seeding 250,000 cells per well and incubating overnight. Spent media were removed, and proteins were diluted in cell media to 1  $\mu$ M and added to the cells, incubated for four hours. Flow cytometry tubes were prepared on ice. Incubation solutions were removed, and cells were washed with 1x DPBS twice. Following addition of trypsin, the cells were incubated for 5 to 10 minutes. The resulting solutions were neutralized by adding cell media, and the cell solutions were collected in centrifuge tubes. Spent media was removed via centrifuge and aspirated off. The cells were resuspended in 300  $\mu$ L of cold 1x DPBS, mixed with pipette, and pushed through filter cap on the flow cytometry tubes. Cells were kept on ice before analysis through the flow cytometer.

In the flow cytometer, the cells were flown through the detector individually, the detector can measure a range of data, including width in two directions to determine cells size, of which is used to determine cell viability.

### **3.2.2.4 Fluorescent microscopy**

For normal cell internalization study, the cells were fixed, either immediately following taking off protein, or incubated for additional time in fresh media solution before fixing. The resulting fixed cells were washed and stained with DAPI prior to imaging.

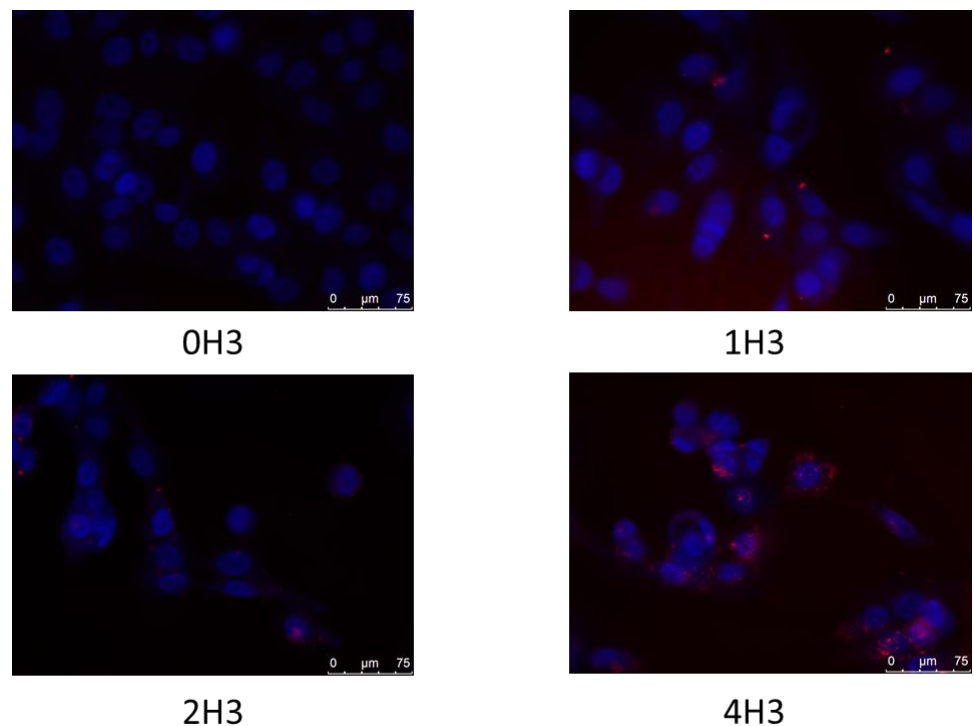
For colocalization study with lysosome, the cells were stained in lysotracker and nuclear stain solutions and imaged live. Confocal microscopy was done in a similar fashion.

Fluorescent microscopy was done to obtain a normal phase image, a DAPI stain in blue, mCherry in red, lysotracker in green. The images were processed in MATLAB using automated scripts and applications.

### **3.3 Results**

#### **3.3.1 Fluorescent microscopy**

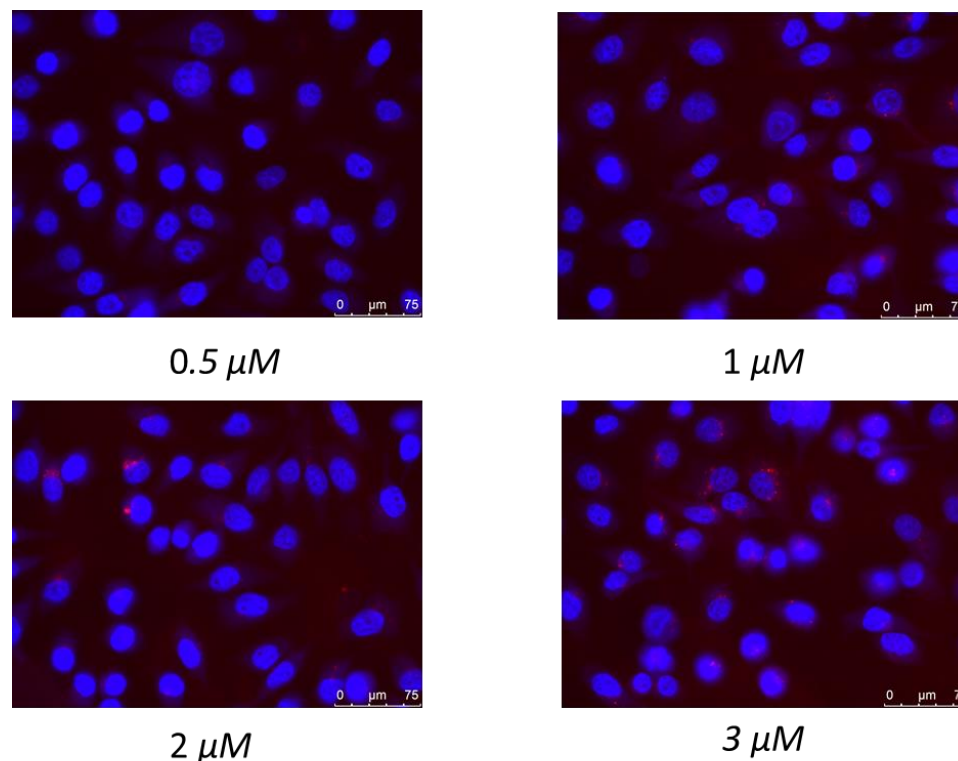
UAA functionalized protein constructs containing a maximum of 1 2 or 4 H3 peptides were delivered into IBC cells at 3  $\mu$ M concentration. The cells were fixed and imaged immediately after four hours of internalization study. In the images, blue color is DAPI nuclear stain showing the nucleus, and red color is mCherry-SpyCatcher protein. Due to insufficient click chemistry conjugation, the numbers represent a maximum possible number of peptides incorporated per construct and may not necessarily represent actual numbers of peptides present.



**Figure 3.2:** Fluorescent microscopy images showing xH3-mCherry-SpyCatcher protein internalization immediately following four hours of internalization.

Increasing numbers of UAA incorporated led to an increase in uptake level of the protein. 1H3 group was not able to elicit a significant amount of uptake and was not used for subsequent study. A good level of uptake was seen for 4H3 group, but large batch to batch variation was also observed due to inconsistent click chemistry conjugation.

To determine an ideal concentration for protein delivery, a range of different concentrations were explored for 4H3-mCherry-SpyCatcher construct in IBC cells. Following taking off the protein solution, IBC cells were incubated for another 24 hours in fresh media solution before fixing and imaging.

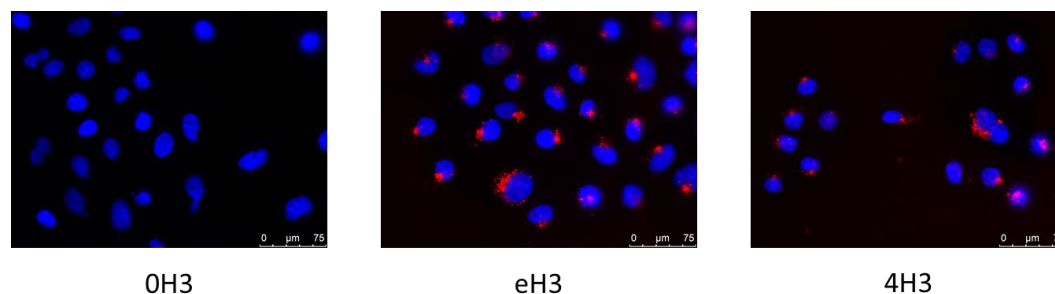


**Figure 3.3:** Fluorescent microscopy images showing 4H3-mCherry-SpyCatcher protein internalization following incubation for 24 hours after protein internalization.

Proteins were able to stay in the cell after 24 hours of incubation in fresh media. 2 μM and 3 μM concentration group produced good results, while 3 μM impacted cell viability, likely due to high protein uptake. 2 μM concentration was used for subsequent study.

In improvement of click chemistry conjugation, ester-amine functionalized constructs were used to allow random incorporation of H3 peptide over the surface of mCherry. Ester-amine construct led to more H3 peptide being incorporated, since the numbers of available azide groups introduced were greater, up to 7 to 8 H3 peptides could potentially be incorporated. In actual construction, it was observed with

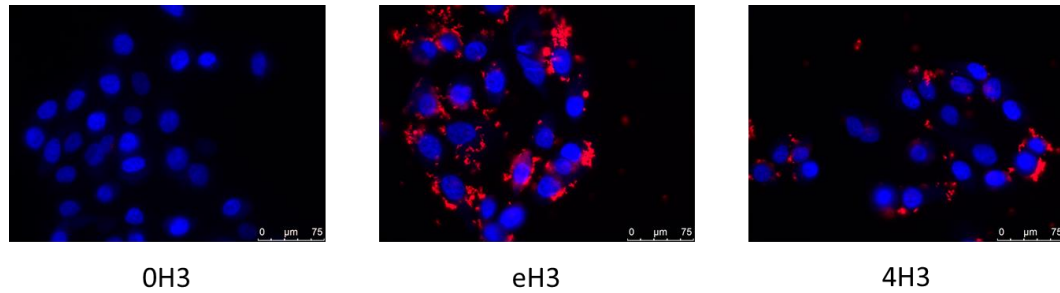
optimized reaction parameters that the 4H3 group (UAA incorporation) would have around 2 to 3 H3 incorporated, where the eH3 group (ester amine chemistry) would have 3 to 5 H3 peptides incorporated. The proteins were delivered into IBC cells, and the cells were incubated for another 20 hours before imaging.



**Figure 3.4:** Fluorescent microscopy images showing protein internalization of xH3-mCherry-SpyCatcher, incubated for additional 20 hours after protein internalization.

A higher level of uptake was achieved using ester-amine functionalized constructs, compared to UAA functionalized constructs, due to increased numbers of H3 peptides incorporated. Proteins stayed in the cell 20 hours following removal of the protein.

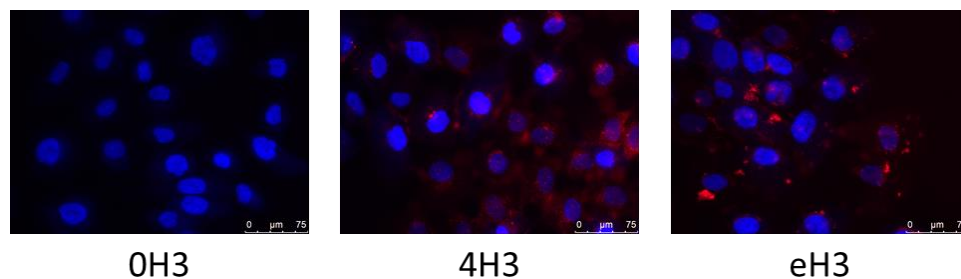
Following optimizing click chemistry parameters, yCD-SpyTag protein was ligated onto the mCherry-SpyCatcher construct via the SpyCatcher/SpyTag ligation system and delivered into IBC cells. The cells were fixed and imaged immediately following four hours of protein internalization.



**Figure 3.5:** Fluorescent microscopy images showing protein internalization of xH3-mCherry-SpyCatcher-SpyTag-yCD, immediately following four hours of protein internalization.

Similar uptake pattern was seen with yCD construct compared to without. Ester-amine groups again showed higher level of uptake due to higher numbers of H3 peptides. Ligation with yCD and subsequent purification process did not impact the protein constructs trafficking significantly.

H3 functionized mCherry-SpyCatcher constructs were also delivered into normal breast epithelial cells, fixed and imaged immediately following four hours of protein internalization study.

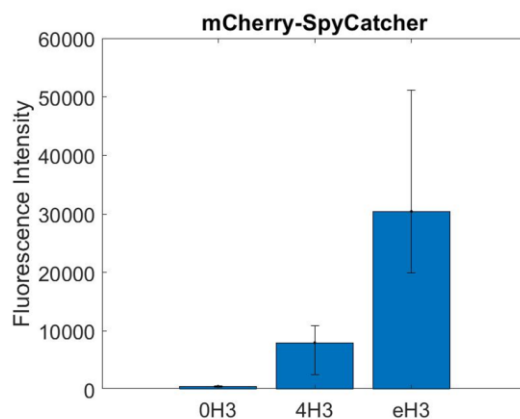


**Figure 3.6:** Fluorescent microscopy images showing protein internalization of xH3-mCherry-SpyCatcher, immediately following four hours of protein internalization.

Protein uptake is significant for both H3 functionalized constructs, likely due to positively charged H3 peptides facilitating entry through negatively charged cell membrane. The UAA functionalized groups were spread out around the cell edge instead of concentrated near the nucleus, this is different from the pattern seen for the ester-amine functionalized group as well as in IBC cells.

### 3.3.2 Flow Cytometry

Following delivery of mCherry construct, fluorescent measurements were done using a flow cytometer, passing one cell through the detector at a time. A blank sample was prepared, and all subsequent measurements were normalized against the blank fluorescence intensity.

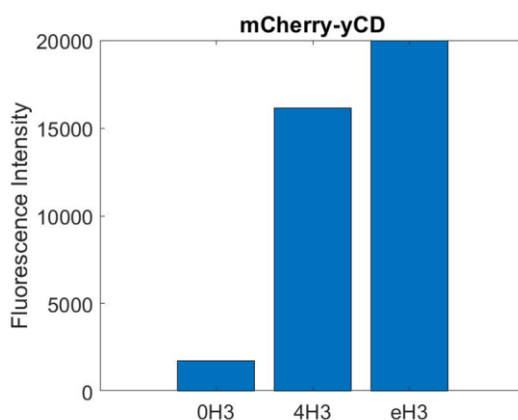


**Figure 3.7:** Flow cytometry results for mCherry-SpyCatcher in IBC cells

Elevated uptake was seen for both the ester-amine functionalized constructs and UAA functionalized constructs. Reaction with UAA constructs were not fully optimized and led to much lower uptake, full optimization of conjugation procedure

would allow further increase in uptake. Ester-amine chemistry led to a much higher uptake due to higher numbers of H3 peptides incorporated and is consistent with fluorescent microscopy results. Large variation is seen due to different batches of protein used in the experiments, further optimization of the expression protocol and conjugation parameters are needed.

Following ligation of yCD to mCherry construct, the protein was delivered into IBC cells and the fluorescent intensities of the cells were evaluated using flow cytometry.

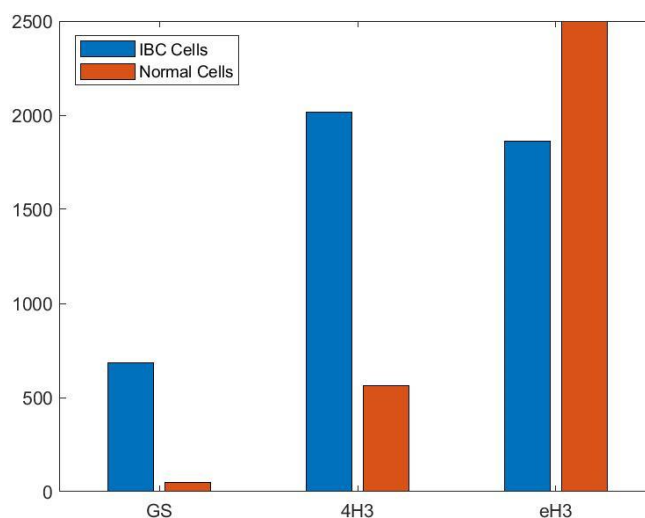


**Figure 3.8:** Flow cytometry results for mCherry-yCD in IBC cells

Uptake level did not change significantly for yCD containing constructs compared to SpyCatcher construct. Additional experiments are necessary to produce statistically significant data. The slightly higher value for UAA functionalized constructs is likely due to better conjugation efficiency for the batch of protein instead of actual improvement of uptake. The slightly lower value for ester-amine

functionalized constructs could be due to solubility issues with highly functionalized constructs.

Ester-amine and UAA functionalized constructs were delivered into both IBC cells and normal breast epithelial cells to compare the relative differences in uptake. The cells were fixed in formalin solution, washed and resuspended in 1x DPBS before imaging days later.



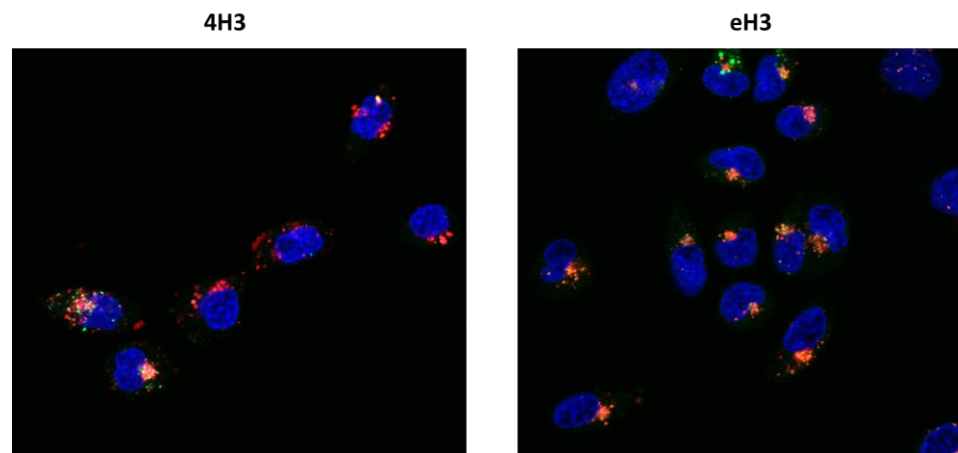
**Figure 3.9:** Flow cytometry results comparison between IBC cells and normal breast epithelial cells.

No overall trend can be concluded due to low numbers of data points. IBC cells seem to have higher uptake overall, but the ester-amine group performed unexpectedly. Based SDS-PAGE results, it is likely that ester-amine functionalized protein construct contains more H3 peptides per construct, and logically should lead to elevated uptake. However, the elevated uptake was only observed in normal breast

epithelial cells but not for IBC cells. This could be due to clustering of the H3 peptides helping to reduce non-specific uptake. Since the cells were fixed and stored in the fridge prior to flow cytometry, the fluorescent intensity suffered significantly compared to cells imaged live. Additional data is required to produce statistically significant data.

### 3.3.3 Confocal

Confocal microscopy was done following 20 hours incubation time post protein delivery to allow time for entry into the nucleus. The cells were stained in nuclear stain and lysotracker solution, and imaged live. The red color is mCherry construct, the blue color is nuclear stain, and the green color is lysotracker. Confocal microscopy takes images at the same height level, blocking out lights that are out of focus. Thus, red mCherry construct colocalized with blue nuclear stain suggest nuclear delivery of the protein was achieved.



**Figure 3.10:** Confocal microscopy images showing protein internalization of xH3-mCherry-SpyCatcher, incubated for additional 20 hours after protein internalization.

Large amount of mCherry protein is colocalized with lysotracker, suggesting majority of the protein was trapped in the lysosome following 20 hours after delivery. Some mCherry protein can be seen in the nucleus, suggesting minimal amount of nuclear delivery.

### **3.4 Conclusion**

Once successfully conjugated onto the mCherry-SpyCatcher construct, the H3 peptides were able to enhance protein uptake in IBC cells. Increasing numbers of peptides attached will significantly elevate uptake level. Some amount of the protein construct was able to stay in the cell at least until 24 hours post taking off protein solution. In addition, the ligation with yCD through the SpyCatcher/SpyTag ligation system did not significantly affect uptake. However, it is worth noting yCD is a relatively small therapeutic protein around 17 kD in mass, compared to around 50 kD for mCherry-SpyCatcher itself. Ligation of a much larger protein could still impact delivery.

From confocal microscopy, minimal amount of nuclear delivery can be seen with some bright puncta colocalized with nuclear stain. High colocalization for mCherry with lysotracker was seen, suggestion large amount of the protein construct being trapped in the lysosome. It was hypothesized that in order to gain entry to the nucleus, an endosomal escape method is needed.

## Chapter 4

### ENDOSOMAL ESCAPE

#### 4.1 Gaining access to the nucleus

Therapeutic proteins may require access to specific cellular compartments in order to fully function as desired. It is often that only a small fraction of nanocarriers would be able shuttle biologics to desired intracellular destination <sup>[18]</sup>. Particularly, access to cytosol and the nucleus are the two of the key interests in therapeutic protein delivery. Protein cargos can be internalized by the cell through endocytosis, where part of the cell membrane engulfs the cargos and trap the cargos in a vesicle named endosome. The endosome will go through an acidifying process and eventually fuse with lysosome and digest all cargos included. Alternatively, caveolae mediated uptake could avoid fusing with early endosome and fuse with neutral caveosomes instead, of which allow the cargos to enter retrograde transport to enable access to the nucleus <sup>[5]</sup>.

H3 peptide were used to in PEI polyplexes for gene delivery, it was observed that H3 peptide can utilize caveolae mediated uptake, enter retrograde transport, and enter the nucleus in post mitosis redistribution of the endoplasmic reticulum. The mechanisms for H3 peptides could potentially be a function of histone methyltransferases that regulates retrograde transport of cargo from endosomes <sup>[19]</sup>. It is expected that for native protein, the trafficking pattern would be different from polyplexes. Since H3 peptide contains nuclear localization signal (NLS), once exposed in the cytosol, H3 tagged polyplexes could be shuttled into the nucleus. It is proposed that once endosomal escape is achieved, the protein constructs would be delivered into the nucleus via importin protein.

#### **4.1.1 Peptide strategy**

Endosomal escape peptides were either obtained from natural sources or engineered based on rational protein engineering. Since these peptides are able to disrupt endosomal membrane, some of them that can also disrupt cell membrane. These peptides are also known as cell-penetrating peptides, which would enhance uptake in all cells significantly, and is undesired for therapeutic protein delivery.

One of the peptide that is not able to disrupt cell membrane is GALA peptide. It is an engineered peptide that is proposed to go through a conformational change under acidic condition. In the acidic endosome environment, GALA peptide will change from a random coil structure to an  $\alpha$ -helix, of which can be inserted into endosomal membrane. Multiple helical GALA can form an endosomal pore, allowing the protein construct to escape, an average of 10 GALA peptides are required for pore formation [8].

#### **4.1.2 Gal8 visualization**

Galectin-8 (Gal8) is a  $\beta$ -galactosidase binding protein containing carbohydrate recognition domains. Gal8 has high affinity to glycans exclusively found on the inside of endosome, thus when the endosome is disrupted, Gal8 will be recruited and signal endosomal escape. Gal8-YFP construct was incorporated into a triple negative breast cancer cell line MDA-MB-231, allowing detection of endosomal disruption. Via delivering siRNA using polymeric material, Gal8 endosomal recruitment was correlated with intracellular siRNA bioactivity [20]. In this study, Gal8 intensity will be used to evaluate endosomal escape efficiency.

## **4.2 Methods**

### **4.2.1 Construction of GALA construct**

#### **4.2.1.1 Plasmid editing**

The plasmids were edited using restriction enzymes, and desired edits were inserted through designed primer sequences. The primer sequences were oligonucleotides synthesized artificially with complimentary sequence containing desired edits, with matching overhangs to the restriction enzyme cut site, and also contain additional restriction enzyme cut site for diagnostic digestion.

The primers were annealed at 95 °C in buffer solution for two minutes and allowed to cool. The primers were treated with T4 kinase to phosphorylate the 5' end for subsequent ligation. Existing plasmid containing mCherry-SpyCatcher construction with four amber stop codon were used as the base constructs. The plasmids were replicated in NEB 5- $\alpha$  strain and isolated using plasmid extraction kit.

The plasmids were digested using restriction enzyme, isolated using agarose gel, to obtain the vector. Desired edits were then introduced through adding the treated primers and ligating using T4 ligase. The resulting new plasmids were transformed into NEB 5- $\alpha$  cells and plated on antibiotic selection LB plate. The plates were incubated overnight at 37 °C, surviving colonies were randomly picked and grown in LB media, followed by plasmid extraction, diagnostic digestion, and sequencing.

#### **4.2.1.2 Agarose gel**

Dissolve 0.1 weight% agarose powders in 1x TAE (Tris-Acetate-EDTA) buffer solution in microwave, allow gels to cool and add ethidium bromide solution, pour the gel into mold and wait for at least 15 minutes to set. Mixing loading dye with

plasmid solution, load into the gel and run the gel in 1x TAE solution at 90 volts for 30 minutes. Visualize the gel in gel documentation system under UV. If recovering plasmids, cut out the desired bands under UV light, and recover plasmid via agarose gel extraction kit.

#### **4.2.2 Endosomal escape assay**

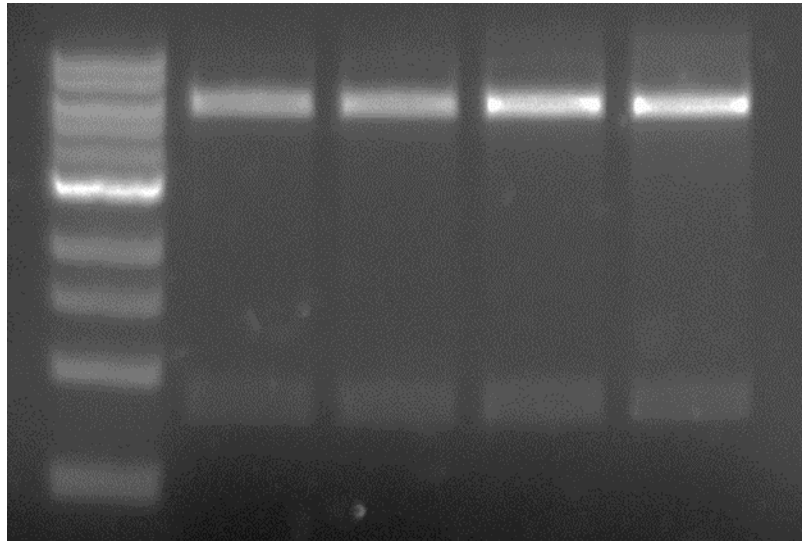
When the endosomal membrane is disrupted, Galectin8-YFP concentrates into bright, punctate fluorescent spots, allowing visualization under fluorescent microscope in green channel. A triple negative breast cancer cell line, MDA-MB-231 strain harboring gal8 edits were used to perform the assay, following normal protein internalization protocol.

The images were processed in MATLAB based on scripts from Duvall group <sup>[20]</sup>, reprogrammed into batch processing scripts and individual image processing application. Gal8 intensity was integrated and normalized with cell counts.

### **4.3 Results**

#### **4.3.1 Diagnostic digestion**

Edited plasmids were transformed into NEB 5- $\alpha$  strain and isolated. Digestions were done with an existing restriction enzyme in the original construct, and another restriction enzyme targeting a cutting site that was previously missing and introduced by the primer sequence. If the edits were successful, the plasmids will be cut into two fragments, showing two bands.



**Figure 4.1:** Diagnostic digestion of GALA-GS-mCherry-h6 construct, left most lane is ladder and all others are plasmids from different colonies.

Plasmid samples that can be cut were subsequently sent to sequencing to verify successful edits were made in the desired location. Following verification, frozen cells stocks were made from transforming finished plasmids into NEB 5- $\alpha$  strain for subsequent plasmid production.

#### **4.3.2 Protein construction**

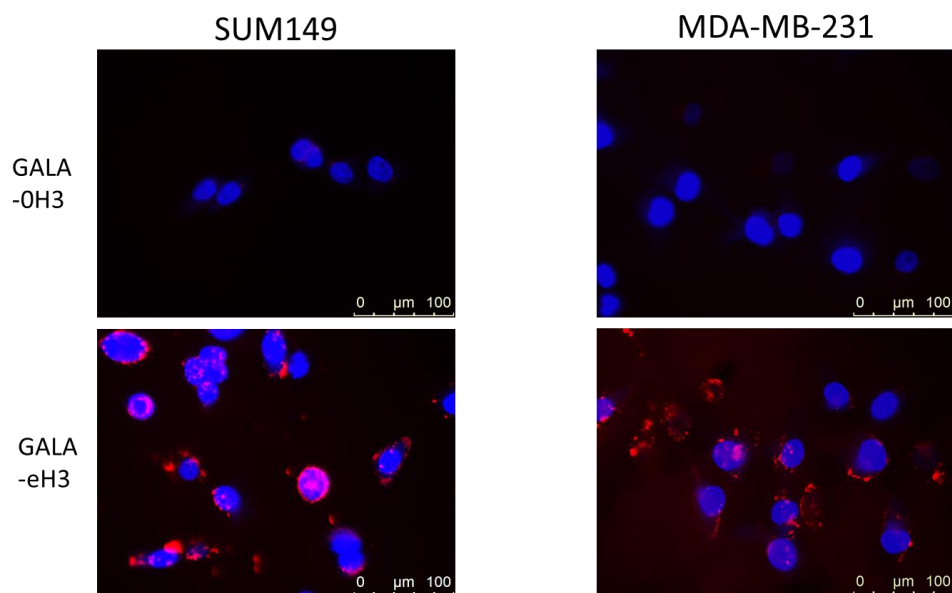
GALA-xAz-mCherry-SpyCatcher protein was expressed in BL21. Numbers of azide groups incorporated can be either 0 or 4, based on numbers of amber stop codons included in the plasmid. Fusion of GALA peptide did not compromise expression level significantly. On the other hand, removal of amber stop codon resulted in sharp elevation of expression level.

UAA functionalized constructs containing 4 azide groups were used in subsequent click chemistry. Constructs without UAA were functionalized using ester-

amine chemistry and used in click chemistry. H3 peptides were conjugated onto GALA containing construct using optimized protocol. The resulting proteins were purified and dialyzed to prepare for delivery. The UAA functionalized constructs were not successfully conjugated to H3 peptides, additional optimization is required.

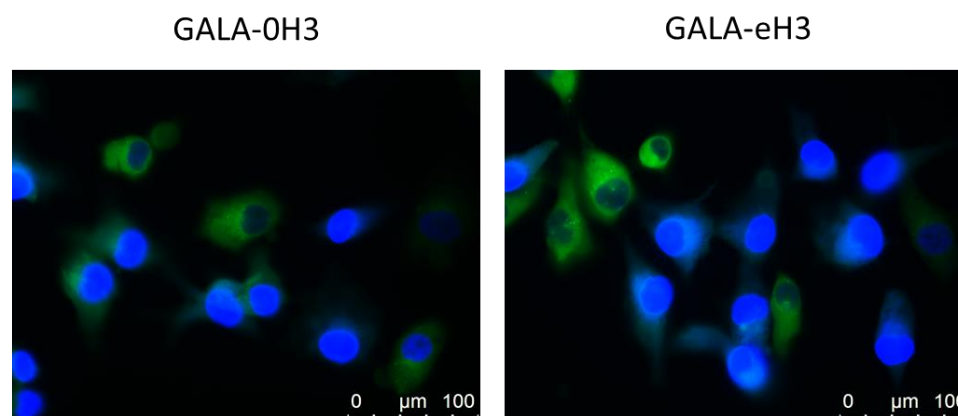
### 4.3.3 Delivery study

H3 functionalized GALA containing constructs were delivered into IBC cells and Gal8 expressing triple negative breast cancer cells. The cells were fixed and imaged immediately following four hours of protein internalization. The IBC cells were used to verify uptake, while the triple negative breast cancer cells were used to evaluate endosomal escape activity.



**Figure 4.2:** Fluorescent microscopy images showing protein internalization of GALA-xH3-mCherry-SpyCatcher, immediately following four hours of protein internalization.

Similar level of uptake was produced in both IBC cells and triple negative breast cancer cells. Compared to non-GALA-containing constructs, a larger number of overlapping mCherry spots and nuclear stain were observed. Suggesting improvement in nuclear entry, but to verify nuclear delivery, confocal microscopy is needed,



**Figure 4.3:** Fluorescent microscopy images showing gal8 detection of endosomal disruption, immediately following four hours of protein internalization.

Following delivery of GALA containing constructs, bright Gal8 puncta were observed, suggesting endosomal disruption occurred, most likely due to GALA activity.

#### 4.4 Conclusion

GALA peptide was successfully introduced into the protein construct using restriction enzymes. The proteins were expressed in BL21, where the fusion of GALA peptides did not impact protein expression significantly. Azide functionalized proteins were modified with click chemistry to introduce H3 functionalization.

Following delivery of GALA containing constructs into the cells, similar uptake level was seen for IBC cells and triple negative breast cancer cells. In Gal8 expressing triple negative breast cancer cells, Gal8 recruitment was observed, suggesting endosomal disruption.

Additional experiments can be done to increase numbers of GALA peptides available per construct, including fusing additional GALA repeats onto the existing construct, and ligation of mCherry-SpyCatcher construct onto SpyTag functionalized E2 nanocage.

Successful conjugation of H3 peptides onto mCherry constructs allows entry into IBC cells and normal breast epithelial cells. The relative uptake level in both types of cells requires additional verification, while preliminary data suggest a difference in uptake level for clustered H3 peptides. Only a very small fraction of H3 conjugated mCherry constructs were able to make it into the nucleus, while a large portion were trapped in the lysosome. Addition of GALA peptide into the mCherry construct led to endosomal escape activity, but additional numbers of GALA peptides may be needed to gain access to the cytosol. Attachment of a small therapeutic protein yCD did not compromise uptake level significantly, while additional verification is needed for larger sized therapeutic protein, such as Cas9.

## REFERENCES

1. De Groot, A.; Scott, D. Immunogenicity of Protein Therapeutics. *Trends in Immunology* 2007, 28 (11), 482-490.
2. Kim, S.; Yang, Y.; Oh, S.; Hong, Y.; Seo, M.; Jang, M. Cancer-Derived Exosomes As A Delivery Platform Of CRISPR/Cas9 Confer Cancer Cell Tropism-Dependent Targeting. *Journal of Controlled Release* 2017, 266, 8-16.
3. Frangoul, H., Altshuler, D., Cappellini, M., Chen, Y., Domm, J., Eustace, B., Foell, J., de la Fuente, J., Grupp, S., Handgretinger, R., Ho, T., Kattamis, A., Kernytsky, A., Lekstrom-Himes, J., Li, A., Locatelli, F., Mapara, M., de Montalembert, M., Rondelli, D., Sharma, A., Sheth, S., Soni, S., Steinberg, M., Wall, D., Yen, A. and Corbacioglu, S., 2021. CRISPR-Cas9 Gene Editing for Sickle Cell Disease and  $\beta$ -Thalassemia. *New England Journal of Medicine*, 384(3), pp.252-260.
4. Ross, N.; Munsell, E.; Sabanayagam, C.; Sullivan, M. Histone-Targeted Polyplexes Avoid Endosomal Escape And Enter The Nucleus During Postmitotic Redistribution Of ER Membranes. *Molecular Therapy - Nucleic Acids* 2015, 4, e226.
5. Ross, N.; Sullivan, M. Overexpression Of Caveolin-1 In Inflammatory Breast Cancer Cells Enables IBC-Specific Gene Delivery And Prodrug Conversion Using Histone-Targeted Polyplexes. *Biotechnology and Bioengineering* 2016, 113 (12), 2686-2697.
6. V. Munsell, E.; L. Ross, N.; O. Sullivan, M. Journey To The Center Of The Cell: Current Nanocarrier Design Strategies Targeting Biopharmaceuticals To The Cytoplasm And Nucleus. *Current Pharmaceutical Design* 2016, 22 (9), 1227-1244.
7. Leader, B., Baca, Q. and Golan, D., 2008. Protein therapeutics: a summary and pharmacological classification. *Nature Reviews Drug Discovery*, 7(1), pp.21-39.
8. Li, W. GALA: A Designed Synthetic Ph-Responsive Amphipathic Peptide With Applications In Drug And Gene Delivery. *Advanced Drug Delivery Reviews* 2004, 56 (7), 967-985.

9. Swartz, A.; Chen, W. Spytag/Spycatcher Functionalization of E2 Nanocages With Stimuli-Responsive Z-ELP Affinity Domains For Tunable Monoclonal Antibody Binding And Precipitation Properties. *Bioconjugate Chemistry* 2018, 29 (9), 3113-3120.
10. Lieser, R.; Chen, W.; Sullivan, M. Controlled Epidermal Growth Factor Receptor Ligand Display on Cancer Suicide Enzymes Via Unnatural Amino Acid Engineering For Enhanced Intracellular Delivery In Breast Cancer Cells. *Bioconjugate Chemistry* 2019, 30 (2), 432-442.
11. Schultz, K.; Supekova, L.; Ryu, Y.; Xie, J.; Perera, R.; Schultz, P. A Genetically Encoded Infrared Probe. *Journal of the American Chemical Society* 2006, 128 (43), 13984-13985.
12. Zakeri, B.; Fierer, J.; Celik, E.; Chittock, E.; Schwarz-Linek, U.; Moy, V.; Howarth, M. Peptide Tag Forming A Rapid Covalent Bond to A Protein, Through Engineering A Bacterial Adhesin. *Proceedings of the National Academy of Sciences* 2012, 109 (12), E690-E697.
13. Vacic, A.; Criscione, J.; Rajan, N.; Stern, E.; Fahmy, T.; Reed, M. Determination of Molecular Configuration By Debye Length Modulation. *Journal of the American Chemical Society* 2011, 133 (35), 13886-13889.
14. Chen, X.; Zaro, J.; Shen, W. Fusion Protein Linkers: Property, Design And Functionality. *Advanced Drug Delivery Reviews* 2013, 65 (10), 1357-1369.
15. Robertson, F.; Bondy, M.; Yang, W.; Yamauchi, H.; Wiggins, S.; Kamrudin, S.; Krishnamurthy, S.; Le-Petross, H.; Bidaut, L.; Player, A.; Barsky, S.; Woodward, W.; Buchholz, T.; Lucci, A.; Ueno, N.; Cristofanilli, M. Inflammatory Breast Cancer: The Disease, The Biology, The Treatment. *CA: A Cancer Journal for Clinicians* 2010, 60 (6), 351-375.
16. Masuda, H.; Zhang, D.; Bartholomeusz, C.; Doihara, H.; Hortobagyi, G.; Ueno, N. Role Of Epidermal Growth Factor Receptor In Breast Cancer. *Breast Cancer Research and Treatment* 2012, 136 (2), 331-345.
17. Longley, D.; Harkin, D.; Johnston, P. 5-Fluorouracil: Mechanisms Of Action And Clinical Strategies. *Nature Reviews Cancer* 2003, 3 (5), 330-338.
18. Varkouhi, A.; Scholte, M.; Storm, G.; Haisma, H. Endosomal Escape Pathways for Delivery Of Biologicals *Journal of Controlled Release* 2011, 151 (3), 220-228.

19. Reilly, M.; Larsen, J.; Sullivan, M. Histone H3 Tail Peptides And Poly(Ethylenimine) Have Synergistic Effects For Gene Delivery. *Molecular Pharmaceutics* 2012, 9 (5), 1031-1040.
20. Kilchrist, K.; Dimobi, S.; Jackson, M.; Evans, B.; Werfel, T.; Dailing, E.; Bedingfield, S.; Kelly, I.; Duvall, C. Gal8 Visualization Of Endosome Disruption Predicts Carrier-Mediated Biologic Drug Intracellular Bioavailability. *ACS Nano* 2019.

## Appendix A

### MATLAB SCRIPTS FOR GALECTIN-8 QUANTIFICATION

Two image files in tiff format were read into MATLAB as RGB matrix, first one being blue DAPI stain image, the second one being green galectin-8 image. The number of cells was counted from DAPI image, using an adjustable threshold. Galectin-8 spots were also identified based on another adjustable threshold, and then the pixel intensity around the bright spot were integrated. The results were normalized against numbers of cells present in the image.

```
clc, clear, clear all;
%input of this program is assumed to be tiff images
%two tiff images is a group
%assume scale bar is not present or removed via codes
%tiff image 1 - nuclear stain
%tiff image 2 - gal8
%input images is assumed to be in the same folder as the code
%images must be ordered correctly for batch processing
list = dir('*.tif');
picnum = length(list);
picname = strings(picnum,1);
for i = 1:picnum
    picname(i) = list(i).name;
end
DataCells={'Run Number','Image Label Figure 1',...
    'Image Label Figure 2','Figure Group','Cell Count','Gal8
Intensity',...
    'Normalized Gal8 Intensity'};
for i = 1:picnum/2
    nucnum = i;
    galnum = i+picnum/2;
    nucname = picname(nucnum);
    nucname = cell2mat(nucname);
    nucname = nucname(6:end-4);
    nucname = string(nucname);
    info1 = imfinfo(picname(nucnum));
    info2 = imfinfo(picname(galnum));
    nuc = imread(picname(nucnum),'Info',info1);
    gal8 = imread(picname(galnum),'Info',info2);
```

```

nuc = nuc(:,:,3);
gal8 = gal8(:,:,2);
se1=strel('disk',1);
se2=strel('disk',2);
se3=strel('disk',3);
se4=strel('disk',4);
se5=strel('disk',5);
se6=strel('disk',6);
se7=strel('disk',7);
se8=strel('disk',8);
se9=strel('disk',9);
se10=strel('disk',10);
se25=strel('disk',25);
galectin8_threshold = 30; % threshold should be manually adjusted
nuclear_threshold = 25; % threshold should be manually adjusted
gal8th = imtophat(gal8,se25);
gal8pos1 = gal8th > galectin8_threshold;
gal8pos2 = imopen(gal8pos1,se1);
circlelayer = xor(imdilate(gal8pos2,se8),imdilate(gal8pos2,se5));
comp = cat(3,circlelayer.*0,gal8,nuc);
circled = cat(3,circlelayer.*2^8,gal8,nuc);
nthr = nuc>nuclear_threshold;
nuc = (imopen(nthr,se25));
basinmap = watershed(~nuc);
cellarea = imdilate(nuc, se3) ;
cellmap = basinmap;
cellmap(~cellarea) = 0;
nucmap = (label2rgb(cellmap));
numcell = max(cellmap(:));
galsum = sum(gal8(gal8pos2));
runnum = strcat('Run ',num2str(i));
DataCells =
[DataCells;{runnum},{picname(nucnum)},{picname(galnum)},{nucname},...
{double(numcell)},{double(galsum)},{double(galsum)./double(numcell)}];
imwrite(nucmap, strcat(picname(nucnum), 'nucmap', '.png'), 'png');

imwrite(circled, strcat(picname(nucnum), 'composite_circled', '.png'), 'png'
);
imwrite(comp, strcat(picname(nucnum), 'composite', '.png'), 'png');
end
xlswrite('Gal8 Images.xlsx',DataCells,1)

```

## Appendix B

### MATLAB SCRIPTS FOR CELL OUTLINING

Two image files in tiff format were read into MATLAB as RGB matrix, first one being phase image, the second one being fluorescence image. Cell edges were detected and outlined on both the original phase image and the fluorescence image.

```
clc, clear, clear all;
%input of this program is assumed to be tiff images
%tiff image 1 - phase image
%tiff image 2 - fluorescent image
%input images is assumed to be in the same folder as the code
%images must be ordered correctly for batch processing
fnamephase = 'phaseimage.tif';
fnamefluor = 'fluorescenceimage.tif';
im1 = imread(fnamephase);
im2 = rgb2gray(im1);
im3 = imread(fnamefluor);
im2eq = adapthisteq(im2);
[~,threshold] = edge(im2eq,'Sobel');
FudgeFactor = 0.7; % can be adjusted to make edge more clear
EdgeFig = edge(im2eq,'Sobel',threshold*FudgeFactor);
Edgeclean = bwareaopen(EdgeFig,5);
DilationLength = 3; % can be adjusted for a cleaner image
se90 = strel('line',DilationLength,90);
se0 = strel('line',DilationLength,0);
DilaFig = imdilate(Edgeclean,[se90 se0]);
MorphSize = 7; % can be adjusted to remove small noises
se5 = strel('disk',MorphSize);
HoleFig = imclose(DilaFig,se5);
HoleSize = 500; % can be adjusted to adjust hole size recognized
FilledFig = bwareaopen(HoleFig,HoleSize);
Figoutline = bwperim(FilledFig);
Figoutline3(:, :, 1) = Figoutline;
Figoutline3(:, :, 2) = Figoutline;
Figoutline3(:, :, 3) = Figoutline;
PoutFig = im1;
PoutFig(Figoutline3) = 255;
imshow(PoutFig) % output outlined phase image
FoutFig = im3;
FoutFig(Figoutline3) = 255;
```

```
imshow(FoutFig) % output outlined fluorescence image
```

## Appendix C

### Molecular Cloning Primer

Table C1: Primer sequence used in this study

Sequence Name	Main sequence identity	Overhang Site 1	Overhang Site 2	Introduced Site	Bases	Sequence
NotI-GALA-BamHI_F1	GALA	NotI			58	GGCCCGGTGGGAAGCAGCTTTGGCAGAGGCTCTTGCGGAGGCCCTGGCCGAGCACTTG
NotI-GALA-BamHI_F2	GALA	BamHI		KpnI	52	AGTGCTCGGCCAGGGCTCCGCAAGAGCCTTGCCAAAGTGCTTCCCACGC
NotI-GALA-BamHI_R1	GALA	NotI		KpnI	52	GATCCGTTACCTGCAGTAATGCTTCCAGGGCTCGGCCAGAGCTTCTGCCA
NotI-GALA-BamHI_R2	GALA	BamHI			51	AGTGCTCGGCCAGGGCTCCGCAAGAGCCTTGCCAAAGTGCTTCCCAGC
SacII-G4S1_2-NheI_F	GGGGS	SacII	NheI	MfeI	39	GGGGTGGCGGAGGGAGCGGTGGCGGTGGCAGCCAATTGG
SacII-G4S1_2-NheI_R	GGGGS	SacII	NheI	MfeI	45	CTAGCCAATTGGCTGCCACCGCCACCGTCCCTCCGCCACCCCGC

Primers were constructed with overhangs to match the cut sites on the plasmid, F denotes forward strand, and R denotes reverse strand. The primers also contains new cut sites that was originally not present in the base plasmid, for subsequent diagnostic digestion.