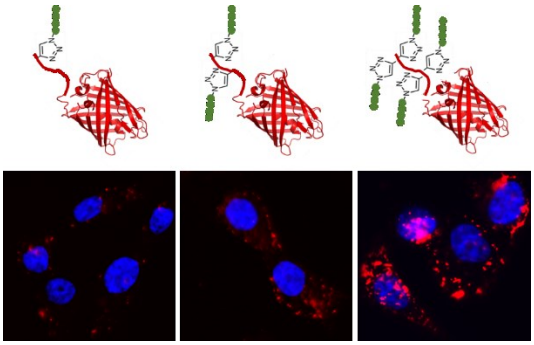
Controlled EGFR ligand display on cancer suicide enzymes via UAA engineering for enhanced intracellular delivery in breast cancer cells

Rachel M. Lieser, Wilfred Chen, and Millicent O. Sullivan**

Department of Chemical and Biomolecular Engineering, University of Delaware, 150 Academy Street, Newark, DE 19716 United States of America
Rachel M. Lieser email: rmlieser@udel.edu
Corresponding authors: Wilfred Chen email: wilfred@udel.edu and Millicent O. Sullivan email: msullivan@udel.edu

Abstract

Proteins are ideal candidates for disease treatment because of their high specificity and potency. Despite this potential, delivery of proteins remains a significant challenge due to the intrinsic size, charge, and stability of proteins. Attempts to overcome these challenges have most commonly relied on direct conjugation of polymers and peptides to proteins via reactive groups on naturally occurring residues. While such approaches have shown some success, they allow limited control of the spacing and number of moieties coupled to proteins, which can hinder bioactivity and delivery capabilities of the therapeutic. Here, we describe a strategy to site-specifically conjugate delivery moieties to therapeutic proteins through unnatural amino acid (UAA) incorporation, in order to explore the effect of epidermal growth factor receptor (EGFR)-targeted ligand valency and spacing on internalization of proteins in EGFR-overexpressing inflammatory breast cancer (IBC) cells. Our results demonstrate the ability to enhance targeted protein delivery by tuning a small number of EGFR ligands per protein and clustering these ligands to promote multivalent ligand-receptor interactions. Furthermore, the tailorability of this simple approach was demonstrated through IBC-targeted cell death via the delivery of yeast cytosine deaminase (yCD), a prodrug converting enzyme.



Introduction

Therapeutic proteins are one of the fastest growing sectors on the pharmaceutical market because of their sophisticated functional properties, capacity for highly specific recognition of biological binding partners, and relevance to multiple diseases.[1] In fact, biologics represented 35% of all new FDA approvals between 2010 and 2016, and the market is poised for further evolution and growth.[2] However, despite enormous interest and significant investments in developing advanced protein therapeutics, protein delivery remains a major limitation. Stability issues have hindered or halted clinical advancement of many antibodies.[3] Most therapeutically-relevant enzymes are incompatible with the extracellular environment, leading to aggregation, activity loss, and/or renal clearance.[4] Furthermore, most proteins are membrane impermeable, and therefore active strategies are necessary to deliver proteins transcellularly or intracellularly, while also targeting the correct cells and subcellular compartments. As a result, the pipeline has a dearth of intracellular protein drug candidates, even though intracellular proteins comprise over 60% of the human proteome and have predicted therapeutic applications ranging from neurological disorders[5] to lysosomal storage disease[6] to cancer[7, 8, 9, 10, 11].

Engineering efforts to address these issues and improve protein delivery often rely on modifying proteins through direct conjugation of biocompatible polymers or targeting ligands that can increase protein stability, alter protein biodistribution, and/or improve cellular uptake.[12] For example, polyethylene glycol (PEG) has been conjugated to a number of FDA-approved protein drugs to improve protease resistance and reduce renal clearance.[13] Bioconjugation is typically accomplished using naturally-occurring reactive residues (*e.g.*, lysines) within the protein sequence, or by using genetic fusion to the protein termini.[14, 15] While both methods have the capacity to enhance various aspects of delivery as compared to proteins in their native form, the

inability to chemically modify proteins with site-specificity often affects the activity of the protein and thereby significantly hinders pharmacological action.[16] For example, one study conjugated a single PEG molecule to distinct locations in human growth factor using a site-specific modification technique and observed a threefold difference in activity *in vivo,* depending on the location of the polymer.[17]

Non-site-specific modification approaches also do not offer control over design variables that can be important determinants of targeting efficacy. Particularly, the importance of ligand spacing and density has been demonstrated for multivalent ligand-receptor interactions via ligand clustering[18] and synergistic receptor binding via dual ligand co-functionalization.[19, 20] The importance of multivalent effects in delivery has been clearly demonstrated using nanocarriers. For example, folate ligands were incorporated in micelles within clusters of varying valencies for delivery to folate-overexpressing cancer cells. The dissociation constant decreased two orders of magnitude when micelles had an average of three ligands per cluster as compared to micelles with an average of 1.5 ligands per cluster.[21] In other examples involving nanocarriers, clustering cell-binding ligands was shown to increase cell-binding affinity by up to 1000-fold.[22, 23] Furthermore, optimal ligand cluster sizes and densities have been proposed for maximal receptor specificity and affinity. [24] While the advantages of ligand clustering have been demonstrated in nanoparticle systems, the effect of ligand clustering for protein internalization has been sorely lacking due to conjugation limitations. These examples point out the clear need for new strategies to modify proteins with well-defined ligand arrays.

A potential strategy to accomplish this goal is the use of a site-specific and multivalent conjugation method for insertion of multiple chemical modification sites into proteins. Previous work has demonstrated the ability to insert biorthogonal reactive residues into proteins through

unnatural amino acid (UAA) incorporation with nonsense codon replacement, enabling direct protein conjugation with simple 'click' chemistries.[25] To date, over 70 UAAs with an array of structures have been successfully incorporated into the genetic code of multiple organisms.[26] UAA incorporation has been used in a number of applications including protein labeling, biosensing, vaccine development, and antibody-drug conjugation.[27, 28, 29]

Herein, we explored the potential of this approach for improved intracellular delivery of protein therapeutics. We specifically sought to determine whether UAA-modification could be used to control the spacing and number of receptor-binding ligands that were inserted into therapeutically relevant enzymes. UAA incorporation was used to attach clusters of the high-affinity epidermal growth factor receptor (EGFR) targeting peptide GE11[30] into fluorescent proteins and suicide enzymes for delivery to inflammatory breast cancer (IBC) cells. IBC is an aggressive subtype of breast cancer with a less than 50% survival rate beyond 5 years,[31] and IBC associated with EGFR overexpression is particularly aggressive. Current drug delivery approaches in IBC are limited due to severe adverse side effects, and hence mechanisms to improve targeted drug uptake would have significant benefits.

We first quantified protein uptake using the fluorescent model protein mCherry. By varying the spacing and number of GE11 peptides linked to mCherry via UAA conjugation, we demonstrated an 18-fold increase in protein uptake when four GE11 peptides were clustered in the protein as compared to proteins with a single GE11 peptide. In addition, uptake in healthy breast epithelial cells was found to be 4-fold lower than uptake in IBC cells, demonstrating the ability to not only deliver large amounts of protein, but also selectively target IBC cells. Furthermore, we modularized the GE11-mCherry for attachment to therapeutic enzyme cargoes using SpyCatcher-SpyTag bioconjugation.[32] SpyCatcher-SpyTag was used to link GE11-mCherry to yeast cytosine

deaminase (yCD), a suicide enzyme that converts the non-toxic prodrug 5-fluorocytosine (5-FC) into the widely used chemotherapeutic 5-fluorouracil (5-FU). The linkage strategy fully preserved yCD enzyme activity, and moreover, co-delivery of yCD and 5-FC resulted in a threefold difference in cell death between the normal and cancerous breast epithelial cells. These results demonstrate the benefits of UAA incorporation for controlling targeting peptide presentation and maximizing cargo protein activity, with key benefits relevant to prodrug therapeutics and a wide range of other intracellular protein therapies.

Results and Discussion

Unnatural amino acid (UAA) incorporation allows site-specific addition of pAzF

Previous studies have demonstrated successful incorporation of p-Azido-l-phenylalanine (pAzF) in *E. coli* though amber stop codon suppression with the pULTRA-CNF suppressor plasmid system.[25] Here, one, two, or four pAzF UAAs were incorporated onto the N-terminus of the fluorescent protein mCherry, and the resulting proteins were termed 1Az-, 2Az-, and 4Az-mCherry. Flexible linkers made up of glycine and serine residues (G4S1) separated the UAAs. The fluorescence intensities of the cell lysates indicated that mCherry was preferentially expressed for all three constructs only when *E. coli* cells were grown with pAzF present in the culture medium (Figure 1A). Compared to 1Az-mCherry, 1.8-fold and 2.9-fold decreases in expression were observed for 2Az-mCherry and 4Az-mCherry, respectively. Decreased expression levels with an increased number of UAAs is a common outcome of UAA-linked protein expression.[33] The addition of multiple amber stop codons in the encoding gene increases the possibility of early termination of translation.[34] Importantly, expression of mCherry with up to four UAAs still resulted in significant yields. Growing cells in the absence of pAzF in the culture medium resulted

in low levels of mCherry expression in any samples, indicating that a majority of the full-length protein in the pAzF-grown samples likely contained the UAA (Figure S1A). The three proteins were purified using His-tag Ni-NTA chromatography (Figure S1B). Incorporation of pAzF was confirmed through copper-catalyzed alkyne-azide cycloaddition (CuAAC) with Alexa Fluor® 488 alkyne dye. Successful conjugation of the Alexa Fluor 488 dye to 1Az-, 2Az-, and 4Az-mCherry was confirmed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), using fluorescence analysis and Coomassie staining (Figure 1B).

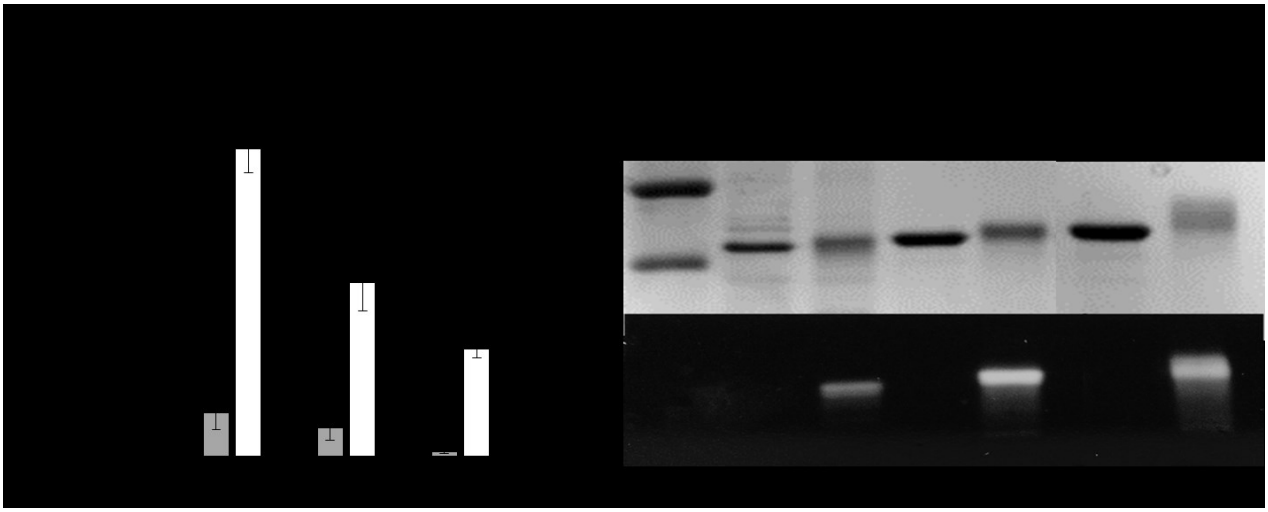


Figure 1. Unnatural amino acid incorporation of pAzF. (A) Expression of full-length mCherry with 1,2, or 4 pAzF residues on the N-terminus in the presence (white) or absence (grey) of pAzF. Fluorescence is normalized to lysate concentration. (B) SDS-PAGE analysis of Alexa Fluor 488-alkyne conjugation to azide groups of mCherry by Coomassie staining to detect total protein (top) and UV transillumination to detect Alexa Fluor 488 fluorescence (bottom).

CuAAC 'click' chemistry conjugates EGFR targeting ligand to protein

The GE11 targeting peptide was chosen to target EGFR, which is overexpressed in a number of different cancer cells including IBC cells.[35], [36] The GE11 peptide exhibits high affinity toward EGFR (K_D = 22 nM) and has been utilized for delivery of an array of nanoparticles in multiple types of EGFR-overexpressing cancer cells.[30, 37, 38, 39] For instance, liposomes loaded with

doxorubicin and functionalized with PEG-GE11 demonstrated a 2.2-fold increase in A549 tumor cell accumulation in *in vivo* models as compared to liposomes functionalized only with PEG.[40]

Here, GE11 was synthesized with an N-terminal propargyl-glycine through Fmoc solid phase peptide synthesis, purified through HPLC, and the final product was confirmed through MALDI-TOF mass spectrometry (Figure S2). Subsequently, CuAAC was used to conjugate the propargyl-GE11 peptide on to the azido groups of mCherry. The reaction yield of propargyl-GE11 with 1pAzF-, 2pAzF-, and 4pAzF-mCherry was estimated via SDS-PAGE analysis (Figure S3A) and MALDI-TOF mass spectrometry (Figure S3B). The change in product size following propargyl-GE11 conjugation to 1Az- and 2Az-mCherry confirmed that a majority of the final product in each protein had all reactive groups modified with the GE11 peptide. Reaction with 4Az-mCherry indicated a majority of the final product had either three or four GE11 peptides. We suspect that the lower modification efficiency of 4GE11-mCherry is due to both limitations in UAA incorporation, *i.e.*, that only three out of the four encoded pAzF groups were actually incorporated in the protein product, as the incorporation of multiple UAAs can increase the likelihood of amber codon read-through; as well as steric hindrance due to the close proximity of the GE11 peptides. The CuAAC reaction itself is unlikely to be the issue as CuAAC linkage typically results in high yields.[41] Usage of an *E. coli* cell line developed specifically for UAA incorporation, rather than BL21(DE3) [42, 43], or increasing the spacing of the UAAs, would likely improve the reaction efficiency.

EGFR overexpression in IBC cells provides targeting opportunity

SUM149 cells are a patient-tumor derived IBC cell line that expresses high levels of EGFR; MCF10A cells are a normal breast epithelial cell line that expresses lower levels of EGFR.[44] The

EGFR expression levels in each cell line were confirmed through immunostaining, using an Alexa Fluor 488 conjugated anti-EGFR antibody (Figure 2A). Flow cytometry analyses indicated that IBC SUM149 expressed approximately 5-fold higher levels of EGFR than MCF10A cells (Figure 2B). Histograms of cellular Alexa Fluor 488 fluorescence based on flow cytometry indicated that a majority of the IBC SUM149 cells expressed EGFR at high levels (Figure 2C), whereas a majority of MCF10A cells expressed only basal levels of EGFR (Figure 2D). The small population of MCF10A cells expressing high levels of EGFR was likely caused by variations in EGFR expression that occur during the cell cycle[45] and cellular heterogeneity.

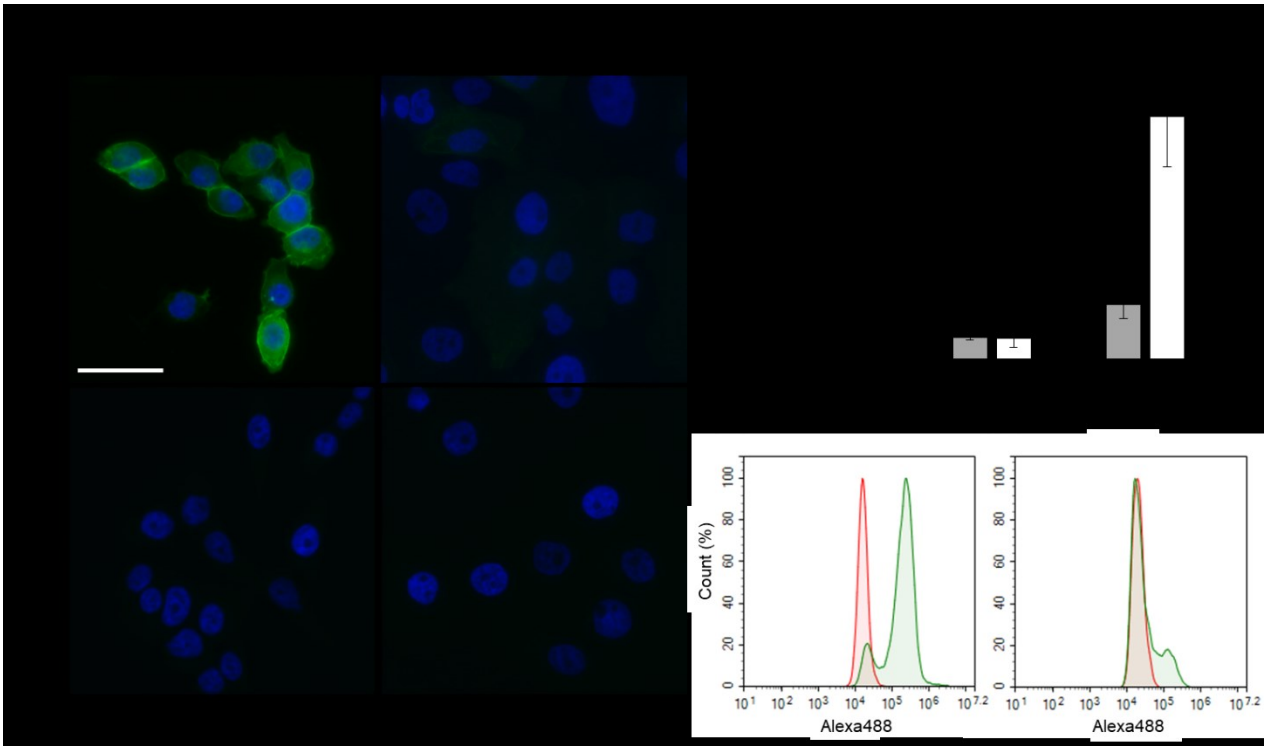


Figure 2. EGFR immunostaining. (A) Fluorescence microscopy images of IBC SUM149 and MCF10A cells incubated with Alexa Fluor 488-conjugated anti-EGFR antibody or IgG antibody control. The scale bar represents 50 µm. (B) Median fluorescence intensity of EGFR and IgG immunostained IBC SUM149 (white) and MCF10A (grey) cells from flow cytometry analysis. Results are shown as the median ± standard deviation of data obtained from three independent experiments. *Indicates a statistically significant difference in EGFR expression in IBC SUM149 and MCF10A cells ($p < 0.05$). (C) Flow cytometry histograms of IBC SUM149 cells stained with

rabbit IgG (red) and anti-EGFR (green). (D) Flow cytometry histograms of MCF10A cells stained with rabbit IgG (red) and anti-EGFR (green).

Control of ligand number provides tunable targeted uptake in IBC cells

To determine if ligand valency played a role in targeted cellular uptake, the GE11-mCherry proteins were delivered to IBC SUM149 cells and MCF10A cells. Fluorescence microscopy was used to visual cell internalization (Figure 3A and Figure S5). Corresponding phase images are provided in the supplemental information (Figure S4). From flow cytometry, 2.9-fold, 13.6-fold, and 40.7-fold increases in cellular association were observed in IBC SUM149 cells for 1GE11-, 2GE11-, and 4GE11-mCherry as compared to uptake levels of mCherry lacking any GE11, termed 0GE11-mCherry (Figure 3B). In addition, GE11-mCherry at all valencies exhibited higher uptake levels in IBC SUM149 cells as compared to MCF10A cells, which express significantly less EGFR. Cellular association of 1GE11-mCherry was 1.8-fold higher in IBC SUM149 cells as compared to MCF10A cells, although this difference was not statistically significant. The association levels of 2GE11-mCherry and 4GE11-mCherry were 4.7-fold and 4.1-fold higher in IBC SUM149 cells *vs.* MCF10A cells, respectively. Although internalization increased in MCF10A cells with increased GE11 number, the difference in internalization levels between IBC SUM149 and MCF10A cells remained significant at higher valencies.

We expect that the non-linearity observed between ligand number and cellular uptake was the result of ligand clustering. Previous studies have demonstrated improved uptake *via* clustering when targeting EGF receptors. For example, affibody molecules against EGFR demonstrated significantly higher uptake in EGFR-positive A431 cells when the affibodies were clustered through a heptamerization domain, as compared to monomeric affibodies.[46] An additional study demonstrated a similar phenomenon by clustering HER2 ligands on liposomes; uptake was

enhanced in both HER2-overexpressing cells and cells with lower HER2 expression levels as compared to liposomes with uniformly distributed ligands.[47]

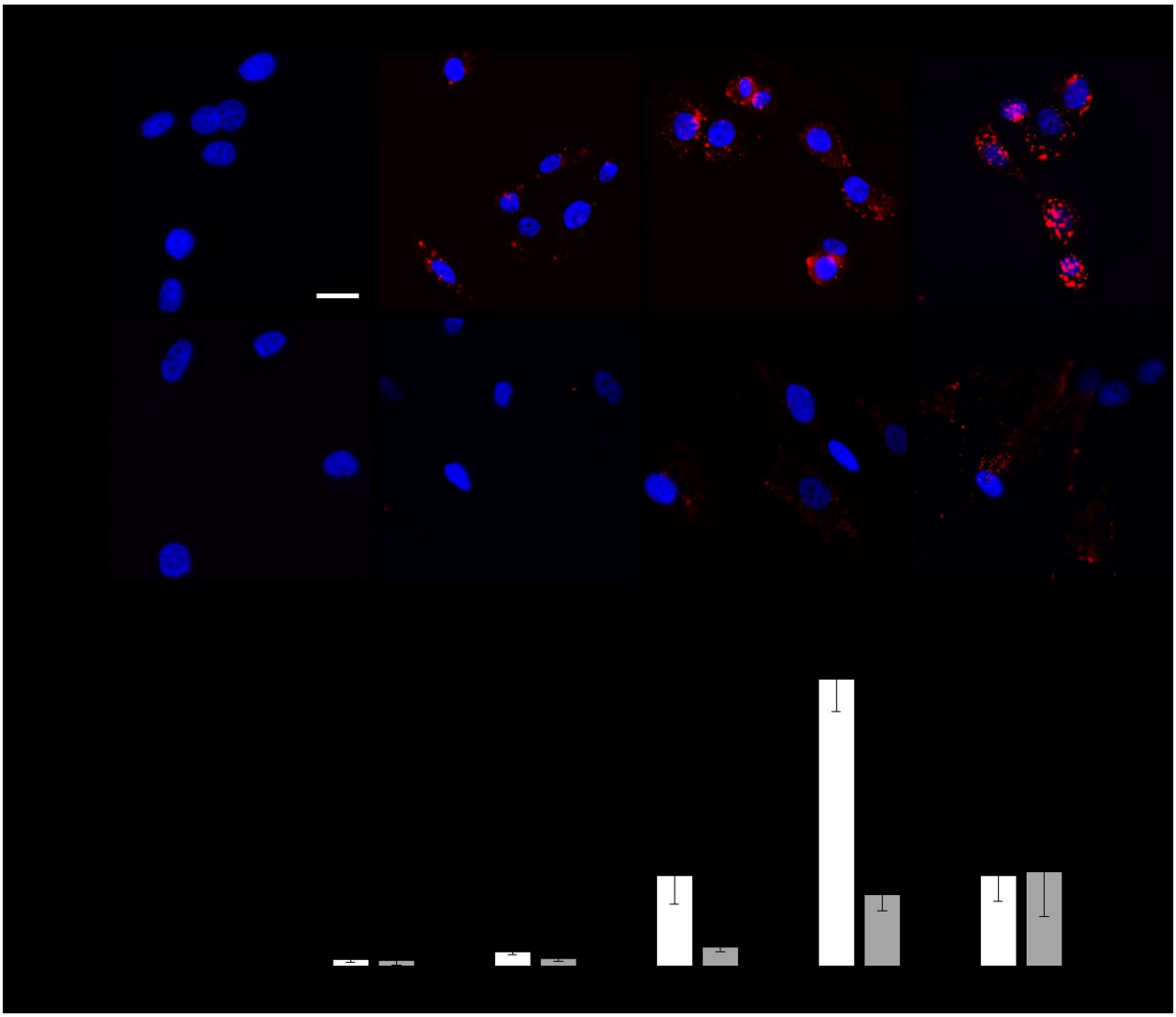


Figure 3. GE11-mCherry cellular internalization with varying GE11 decoration densities. (A) Fluorescence micrographs of IBC SUM149 (white) and MCF10A (grey) cells incubated with GE11-mCherry constructs. The scale bar represents 25 μm. (B) Fluorescence intensity of GE11-mCherry uptake from flow cytometry analysis. Results are shown as the mean ±standard deviation of data obtained from three independent experiments. *Indicates a statistically significant difference in uptake between IBC SUM149 and MCF10A cells ($p < 0.05$). **Indicates a statistically significant difference in uptake between mCherry constructs in IBC SUM149 cells ($p < 0.05$).

To determine whether UAA-mediated GE11 targeting offered benefits in cell uptake and cell specificity as compared with common cell penetrating peptide (CPP) strategies in proteins, protein internalization levels also were compared to uptake levels of a Tat-mCherry fusion protein. Tat is an HIV-derived CPP that is able to elicit high levels of non-specific cellular uptake due to its positive charge.[48] Tat fusions have been widely used to stimulate intracellular delivery of proteins.[49] Tat-mCherry showed comparable levels of uptake in IBC SUM149 cells and MCF10A cells demonstrating its non-specificity. In addition, Tat-mCherry showed 3.2-fold lower cell association levels as compared to 4GE11-mCherry in IBC SUM149 cells, demonstrating that the UAA/GE11 approach offered significant improvements in not only targeting specificity, but also cellular uptake efficiency.

EGF inhibition demonstrates EGFR mediated uptake

EGF (K_D = 2 nM) was used in competitive binding assays to confirm whether or not GE11-mCherry uptake was EGFR specific. EGF was incubated with IBC SUM149 cells for 1 h prior to delivery of GE11-mCherry. Fluorescence microscopy and flow cytometry showed significantly reduced uptake levels of all GE11-mCherry constructs when EGF was pre-incubated with IBC SUM149 cells, suggesting that GE11-mCherry uptake was EGFR mediated (Figure 4A). Specifically, 11.3-fold, 5.4-fold, and 2.6-fold drops in cell association were observed when IBC SUM149 cells were pre-incubated with excess concentrations of EGF followed by delivery of 1GE11-, 2GE11-, or 4GE11-mCherry (Figure 4B), respectively. These results suggest that the number of GE11 peptides has an impact on the ability of EGF to inhibit mCherry internalization. One possible explanation for this is ligand clustering, which has been shown to enhance the apparent ligand-receptor binding affinity.[50] An alternative possibility is based on the charge of the

GE11 peptide. GE11 has a positive charge of one, meaning that the overall charge of GE11-mCherry increases as the number is increased. Cationic charge has been shown to elicit cellular uptake, and so the increased charge could lead to non-specific mCherry uptake.[51] Nonetheless, EGFR appeared to play an important role in the uptake mechanism of the protein.

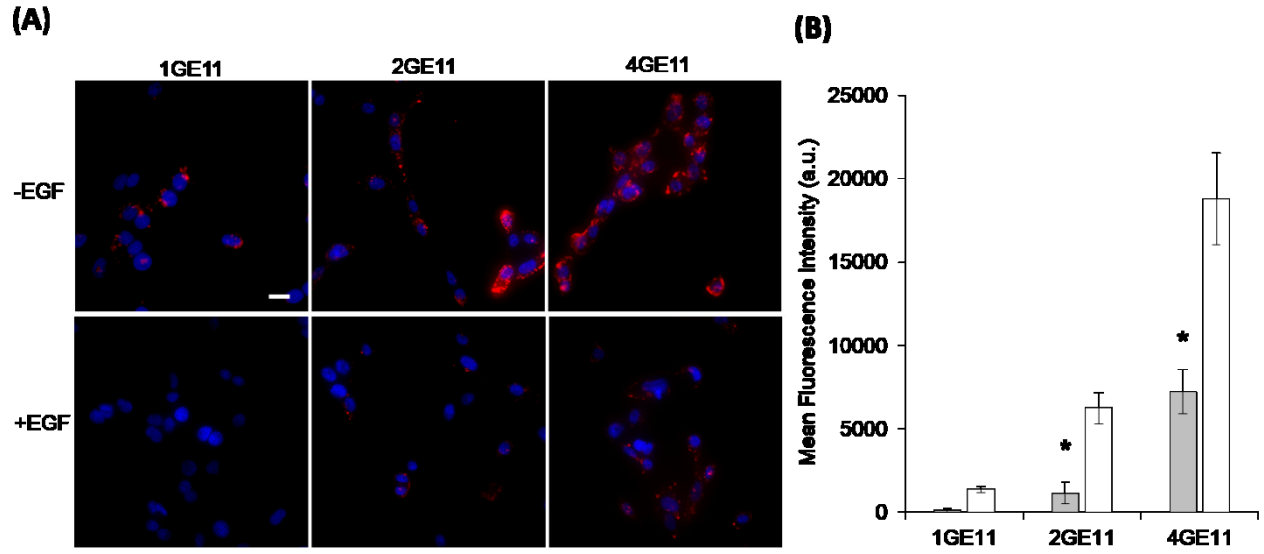


Figure 4. GE11-mCherry EGFR Inhibition. (A) Fluorescent microscopy images of GE11-mCherry uptake (-EGF) in IBC SUM149 compared to uptake following EGF pre-incubation (+EGF). Scale bar represents 25 μM. (B) Mean fluorescent intensity of GE11-mCherry association with and without EGF inhibition from flow cytometry analysis. Results are shown as the mean ±standard deviation of data obtained from three independent experiments. *Indicates a statistically significant difference in uptake between IBC SUM149 cells pre-incubated or in the absence of EGF ($p < 0.05$).

Ligand clustering impacts cellular uptake

To determine whether ligand clustering influenced EGFR binding and mCherry uptake, we generated a GE11-mCherry construct in which the GE11 peptide was spaced throughout the protein rather than clustered on the N-terminus. To do this, the GE11 peptide was conjugated to mCherry using reactive amines on surface lysine residues (Figure S7A). First, azidobutyric acid NHS ester was reacted in tenfold excess with amine groups in mCherry so that azido groups were available on the surface. Secondly, azido groups were reacted with alkyne-GE11 using CuAAC to

incorporate approximately four GE11 peptides per protein (Figure S7B). This reaction resulted in mCherry protein that was highly insoluble, likely due to the loss of charge on the mCherry protein resulting from reacting the lysine residues. Applying the randomly modified protein to IBC SUM149 cells resulted in the appearance of protein aggregates on the surface of the cells, but the protein was not internalized at high levels (Figure S7C). Extensive optimization was needed in order to procure protein that was soluble, further highlighting the usefulness of UAA incorporation for ligand conjugation.

We hypothesized that this decrease in solubility was due to changing the overall charge of the protein by reacting surface lysine residues. To improve the solubility of GE11-mCherry conjugated through NHS ester-amine chemistry, azidobutyric acid NHS ester was limited so that only an average of two azido groups were reacted per protein. Following CuAAC with alkyne-GE11, the protein appeared soluble (Figure S7E). From SDS-PAGE, this protein appeared to have similar numbers of GE11 per mCherry as compared with 2GE11-mCherry reacted via the UAA approach (Figure S7D). While there is little control over GE11 addition using this method, it was expected that a majority of the protein would have non-clustered GE11 peptides due to the spacing of available lysine residues.

Both the clustered and non-clustered 2GE11-mCherry proteins were administered to IBC SUM149 cells. Cellular internalization and association was detected with fluorescence microscopy (Figure 5A) and flow cytometry (Figure 5B), respectively. The results of these experiments demonstrated that the clustered 2GE11-mCherry was uptaken at significantly higher levels as compared to the non-clustered version, which demonstrated comparable uptake levels to 1GE11-mCherry; these results suggest that GE11 ligand clustering plays a clear role in enhancing EGFR-mediated uptake. These results are comparable to a study that demonstrated improved DNA

delivery with polyplexes functionalized with a branched divalent GE11 ligand as compared with individual GE11 peptides.[52] It should be noted that while these finding appear to be a result of ligand clustering, because different conjugation techniques were used to generate the clustered and non-clustered proteins, changes in protein properties, such solubility, could also be responsible for some of the differences observed.

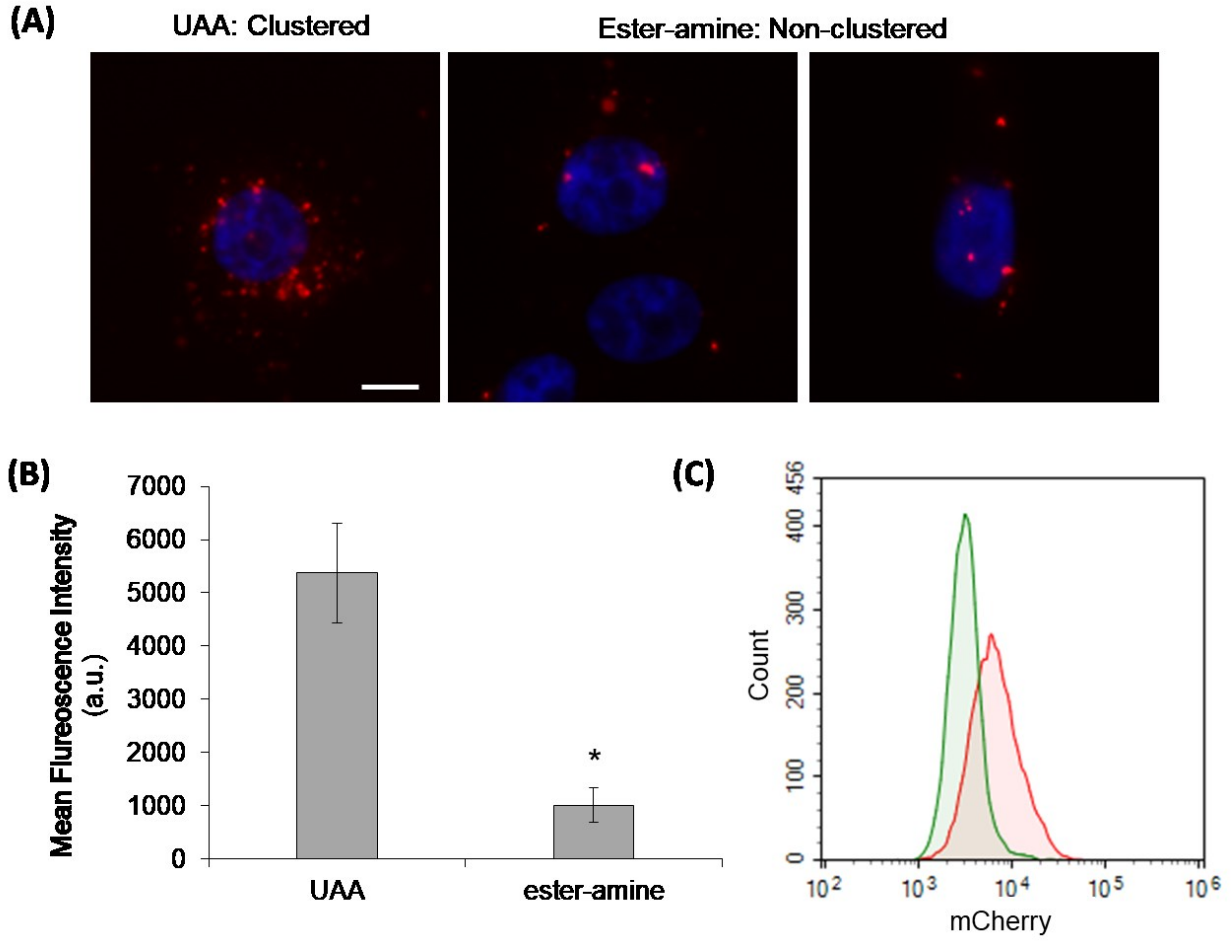


Figure 5. Internalization with non-clustered 2GE11-mCherry. (A) Fluorescence microscopy images of clustered and non-clustered 2GE11-mCherry cellular internalization. Scale bar represents 15 µm. (B) Flow cytometry analyses of non-clustered and clustered 2GE11-mCherry cellular association. Results are shown as the mean ± standard deviation of data collected from three independent experiments. *Indicates a statistically significant difference in uptake between clustered and non-clustered 2GE11-mCherry ($p < 0.05$). (C) Flow cytometry histograms of clustered (red) and non-clustered (green) 2GE11-mCherry uptake in IBC SUM149 cells.

Plug-and-play incorporation of yCD suicide enzyme with SpyCatcher-SpyTag

The SpyCatcher-SpyTag protein coupling system was used to facilitate plug-and-play conjugation of therapeutically-relevant enzymes to GE11-mCherry. The lysine on SpyCatcher and aspartic acid on SpyTag are able to form an isopeptide bond within minutes and the reaction demonstrates high yields in diverse solution conditions.[32] A variety of protein cargoes could be delivered by simply fusing SpyTag, a small 13 amino acid peptide tag, to a choice therapeutic protein and coupling it to the GE11-mCherry-SpyCatcher constructs. To enable SpyCatcher-SpyTag-mediated enzyme attachment to the GE11-modified mCherry delivery constructs, SpyCatcher was fused to the C-terminus of 1Az-, 2Az-, and 4Az-mCherry. Expression levels were not significantly altered with the addition of the SpyCatcher fusion protein (Figure S8A). Subsequently, GE11 was conjugated to the Az-mCherry-SpyCatcher constructs using CuAAC chemistry. This reaction showed yields similar to those seen with the Az-mCherry constructs (Figure S8B).

To demonstrate the plug-and-play enzyme attachment strategy, SpyCatcher-SpyTag was used to attach the prodrug converting enzyme, yCD, to 1GE11-, 2GE11, or 4GE11-mCherry-SpyCatcher. yCD catalyzes the deamination of cytosine to uracil, enabling rapid conversion of 5-FC, a non-toxic prodrug, into 5-FU, a toxic chemotherapeutic.[53, 54] 5-FU is an FDA-approved chemotherapeutic that has been used widely in the treatment of colorectal and breast cancers due to its capacity to inhibit DNA replication;[55] however, off-target side effects of 5-FU, such as fatigue, nausea, and cognitive impairment[56] have motivated interest in prodrug conversion strategies. SpyTag was fused to the N-terminus of yCD, and the SpyTag-yCD protein was coupled to GE11-mCherry-SpyCatcher to form GE11-mCherry-yCD. SDS-PAGE analysis confirmed that a majority of the full length mCherry protein was coupled to yCD (Figure S8B).

Delivery of suicide enzyme elicits IBC-targeted cell death

To assess the internalization of 1GE11-, 2GE11-, and 4GE11-mCherry-yCD, uptake experiments were performed in IBC SUM149 cells and MCF10A cells. Flow cytometry experiments demonstrated that the GE11-mCherry-yCD constructs showed on average roughly twofold lower uptake compared to GE11-mCherry proteins in both cells lines (Figure 6A). This reduction of uptake is likely due to changes in the molecular weight and surface properties from the addition of the yCD. Despite this difference, the effect of GE11 valency remained. Uptake in IBC SUM149 demonstrated 2.4-fold, 7.4-fold, and 19.3-fold increases for 1GE11-, 2GE11-, and 4GE11-mCherry-yCD as compared to 0GE11-mCherry-yCD. In addition, uptake was 2.3-fold, 2.8-fold, and 3.1-fold higher in IBC SUM149 than MCF10A cells, again demonstrating the targeting capacity of this approach.

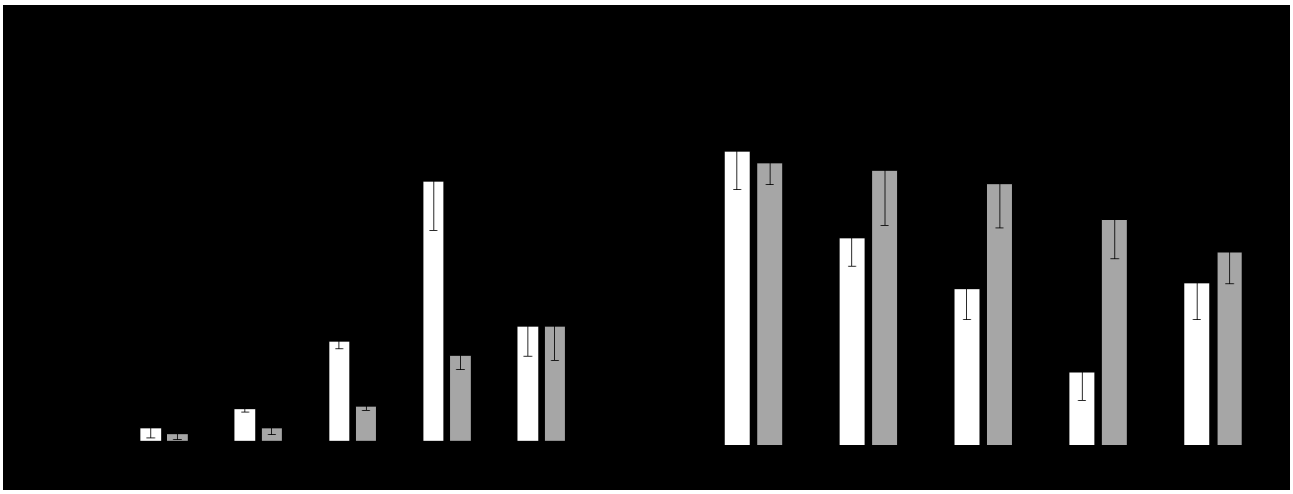


Figure 6. yCD delivery and cell viability. (A) Flow cytometry analyses of mCherry-yCD cellular association in IBC SUM149 (white) and MCF10A (grey) cells. Results are shown as the mean ± standard deviation of data collected from three independent experiments. *Indicates a statistically significant difference in uptake between IBC SUM149 and MCF10A cells ($p < 0.05$). **Indicates a statistically significant difference in uptake between mCherry-yCD constructs in IBC SUM149 cells ($p < 0.05$). (B) MTT assay to assess the viability of SUM149 (white) and MCF10A (grey) cells following delivery of yCD and treatment with 5-FC. 5-FU viability was subtracted as background and the results were normalized to 5-FC viability. Results are shown as the mean ± standard deviation of data obtained from four independent experiments. *Indicates a statistically significant difference ($p < 0.05$) in viability between IBC SUM149 and MCF10A cells.

The activity of yCD was tested in both cell types to determine whether GE11-mCherry-yCD could selectively decrease viability of IBC SUM149 cells when treated with the prodrug 5-FC. Both types of cells were incubated with the GE11-mCherry-yCD proteins. Extracellular protein was removed through multiple wash steps, and 5-FC was subsequently administered. Following treatment, cell viability was measured using MTT assays. In the absence of yCD, cell death did not occur despite the presence of 5-FC, with 91% cell viability in IBC SUM149 cells, and 90% cell viability in MCF10A cells as compared to the same cell types untreated. Direct treatment with 5-FU induced significant decreases in viability for both IBC SUM149 cells and MCF10A cells, with cell viabilities of 45% and 47%, respectively, following 5-FU treatment. Delivery of yCD proteins without 5-FC did not impact viability in either cell type (Figure S9).

To understand the effect of yCD/5-FC treatment, the cell viability of 5-FU was set to zero and sample viability was normalized to the viability of 5-FC (Figure 6B). In IBC SUM149 cells, a clear reduction in viability was evident as the GE11 valency was increased, with 4GE11-mCherry-yCD/5-FC treatment inducing toxicity levels that were nearly the same as the toxicity levels following direct treatment with 5-FU. Treatment of MCF10A cells resulted in a similar effect, in that higher GE11 valences produced increased levels of off-target cell death; however, viability at all GE11 densities was higher for MCF10A cells compared to IBC SUM149 cells. Specifically, 1GE11-, 2GE11-, and 4GE11-mCherry-yCD/5-FC treatment produced 1.4-, 1.7-, and 2.9-fold differences in cell viability, respectively, between the two cell types. Treatment with Tat-mCherry-yCD and 5-FC also resulted in elevated levels of cell death in both cell types, but there was not a significant difference in viability between IBC SUM149 cells and MCF10A cells. These data demonstrate the targeting capabilities of GE11-mCherry-yCD proteins, and they further

demonstrate the ability to control viability selectively in IBC SUM149 cells and MCF10A cells using alterations in GE11 valency.

Conclusions

Herein, we demonstrated the ability to control intracellular protein delivery to IBC cells by controlling the valency of EGFR targeting peptides through UAA incorporation. Furthermore, the SpyCatcher-SpyTag conjugation system was employed for plug-and-play therapeutic cargo protein attachment. Though this method was applied to prodrug cancer therapy, this approach could be tailored to a multitude of applications by selecting alternative targeting peptides and therapeutic proteins. In addition, while the scope of this study focused on targeting peptides, site-directed conjugation of hydrophilic polymers and endosomolytic peptides could be further explored to address additional challenges associated with *in vivo* cytosolic protein delivery.

Using this approach, delivery of the prodrug converting enzyme, yCD, resulted in significant IBC-targeted cell death when treated with 5-FC, with the levels of cell death controllable through alterations in ligand number. Our results demonstrate the importance of controlling the location of delivery molecules conjugated to proteins for improved cell specificity, delivery efficiency, and pharmacological activity of protein drugs, a phenomenon that has not been studied extensively in proteins due to conjugation limitations. While UAA incorporation is currently limited to laboratory scale production, as the ability to site-specifically modify proteins becomes more feasible, further understanding of ligand and polymer display will likely play an important role in improving delivery and efficacy of advance protein therapeutics.

Experimental Procedures

Materials

All DNA primers used to perform polymerase chain reaction (PCR) were purchased from IDT (Coralville, IA). Restriction enzymes, T4 DNA ligase, and Q5 DNA polymerase for DNA cloning were purchased from NEB (Ipswich, MA). Bacterial culture medium ingredients were purchased from Fisher Scientific (Pittsburgh, PA). Zyppy plasmid kit was purchased from Zymo Research (Irvine, CA) for DNA purification following digestion or gel electrophoresis. Antibiotics and isopropyl-β-D-1-thiogalactopyranoside (IPTG) were purchased from Sigma-Aldrich (St. Louis, MO). p-Azido-l-phenylalanine (4-azido-l-phenylalanine, $\geq$ 98 % (HPLC)) was purchased from Chem-Impex International Inc. (Wood Dale, IL). Reagents for SDS-PAGE were purchased from BIO-RAD (Hercules, CA). Amino acids and resin for peptide synthesis were purchased from MilliporeSigma (Burlington, MA) and CEM Corporation (Matthews, NC). All solvents for peptide synthesis were purchased from Fisher Chemical (Fair Lawn, NJ). Dulbecco's' Phosphate Buffered Saline (DPBS, 1X), Ham's F-12, and Dulbecco's Modification of Eagle's Medium/Ham's F-12 50/50 Mix were purchased from Thermo Fisher Scientific (Grand Island, NY).

Construction of expression plasmids

Constructs were prepared using standard molecular cloning techniques. A gene fragment encoding mCherry was amplified with PCR using primers amber-mCherry-F and amber-mCherry-R (Table S1), in which an amber codon and NheI cut site were genetically fused to the N-terminus of the mCherry gene fragment. The PCR product was inserted into pET22b-w3-CFP-his6[57] using SacII and XhoI sites to yield pET22b-amber-mCherry-his6. Gene fragments encoding amber-G_4S_1-amber and (amber-G_4S_1-amber)$_2$, were annealed using primers 2amber-F, 2amber-R,

4amber-F, 4amber-R (Table S1) followed by T4 Polynucleotide Kinase treatment. The DNA fragments were inserted into pET22b-amber-mCherry-his6 using SacII and NheI sites to yield pET22b-2amber-mCherry-his6, and pET22b-4amber-mCherry-his6, respectively. All plasmids were transformed into *Escherichia coli* NEB5α (NEB, Ipswich, MA) [*fhuA2* Δ(*argF-lacZ*)*U169 phoA 23 glnV44* Φ*80* Δ(*lacZ*)*M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17*]. Bacteria were grown on 25 g/L Luria-Bertani Broth (LB, 10 g/L tryptone, 5 g/L yeast extract, 5 g/L sodium chloride) and 15 g/L agar plates supplemented with 100 μg/mL ampicillin. Positive clones were co-transformed with pULTRA-CNF (a gift from Prof. Peter G. Schultz,[25]) into *E. coli* strain BL21(DE3) (EMD Millipore, Madison, WI) [*F- ompT hsdSB(rB- mB-) gal dcm (DE3) Δ(srlrecA)306::Tn10 (TetR)*]. Plasmid sequences and primers can be found in the Supporting Information (Figure S10 and Table S1).

Expression and purification of proteins

Proteins were expressed in Terrific Broth (TB) media (12 g/L tryptone, 24 g/L yeast extract, 0.4% (v/v) glycerol, 9.4 g/L potassium phosphate monobasic, 2.2 g/L potassium phosphate dibasic) supplemented with 100 μg/mL ampicillin and 100 μg/mL spectinomycin. Cultures were inoculated with an overnight culture from a single colony to an OD600 of 0.05 and allowed to grow at 37ºC in a shake flask to an OD600 of 0.06 – 0.08. Expression was then induced with 1 mM IPTG and supplemented with 1 mM pAzF. Cultures were grown overnight at either 20ºC (1Az-mCherry-his6 and Tat-mCherry-his6) or 37ºC (2Az-mCherry-his6 and 4Az-mCherry-his6).

Cells were pelleted with centrifugation at 4000g for 10 minutes at 4ºC. Spent media was removed and cells were resuspended in 1 x phosphate buffered saline (PBS, pH 7.4) with 10 mM imidazole to an OD600 of 20. Cells were lysed via sonication and centrifuged at 10,000g for 15

min at 4°C to collect soluble protein. Proteins were purified using His-Bind Ni-NTA resin gravity column from Thermo Fisher (Pittsburgh, PA) according to the manufacturer's protocol. After purification, proteins were dialyzed overnight in 1x PBS.

Site-specific conjugation of Alexa Fluor 488

Alexa Fluor 488 Alkyne was purchased from Molecular Probes (Eugene, Oregon). 100 μM Alexa Fluor 488 Alkyne was reacted to 25 μM azido-mCherry in PBS (pH 6.5) with 250 μM $CuSO_4$, 1.25 mM THPTA ligand, and 5 mM sodium ascorbate for 1 hour at room temperature.[58] Dye conjugation was confirmed via SDS-PAGE. Alexa Fluor 488 was detected with a Typhoon laser scanner (Marlborough, Massachusetts) and the gel was subsequently stained with Coomassie Brilliant Blue.

Synthesis of GE11 peptide

GE11 with an N-terminal linker (HAIYPRHYHWYGYTPQNVI) was synthesized using solid phase peptide synthesis[59] A propargyl glycine was conjugated to the N-terminus for CuAAC. The peptide was purified by reverse-phase high-performance liquid chromatography (not shown) with >90% purity and the final product was confirmed with MALDI-TOF mass spectrometry. From MALDI the molecular weight of the final product (2510.47 g/mol) matched the expected molecular weight (2510.77 g/mol).

Site-specific conjugation of GE11 peptide to pAzF-mCherry proteins

GE11 was conjugated to mCherry constructs with CuAAC. Briefly, 150 μM alkyne-GE11 was reacted to 25 μM azido-mCherry in 1xPBS (pH 6.5) with 250 μM CuSO4, 1.25 mM THPTA

ligand, and 5 mM sodium ascorbate for 1 hour at room temperature.[58] The protein-peptide conjugate was then purified with His-Bind Ni-NTA resin and dialyzed overnight in 1x PBS. The products were analyzed with SDS-PAGE and MALDI mass spectrometry. Samples were filtered with a 0.22 μM syringe filter before being used in cell studies.

Cell culture

IBC SUM149 cells (a gift from Kenneth van Golen[60]) were grown in Ham's F12 medium supplemented with 5% FBS, 1% (v/v) penicillin/streptomycin, 1% (v/v) mycoplasma antibiotic supplement, 1% (v/v) glutamine, 5 μg/mL insulin, 2.5 μg/mL transferrin, 200 ng/mL selenium, and 1 μg/mL hydrocortisone according to previously established methods.[61, 62] MCF10A cells, purchased from ATCC (Manassa, Virginia), were grown in 50/50 DMEM/Ham's F12 medium supplemented with 5% FBS, 1% (v/v) penicillin/streptomycin, 50 μg/mL bovine pituitary extract, 10 μg/mL insulin, 0.5 μg/mL hydrocortisone, 100 ng/mL cholera toxin, and 20 ng/mL epidermal growth factor.

Cellular internalization of GE11-mCherry protein

First, 5 x 10^4 IBC SUM149 and MCF10A cells were seeded in 8-well plates with a collagen film (1.5 mg/mL Collagen I Bovine Protein in 0.02 M acetic acid) and incubated for 24 hours at 37ºC. Cells were incubated with 1 μM of protein for 3 hours. Media was removed and cells were washed three times in 1x DPBS (pH 7.4). Cells were fixed with 4% paraformaldehyde for 15 minutes, treated with DAPI (300 nM) for 10 minutes, and rinsed three times with 1x DPBS. Internalization was observed at 40x magnification on a Leica DM6000 fluorescence microscope

(Wetzlar, Germany) with 350/50 nm excitation and 460/50 nm emission for DAPI and 545/25 nm excitation and 605/70 nm emission for mCherry.

Flow cytometry was used as a quantitative analysis of GE11-mCherry uptake in IBC SUM149 and MCF10A cells. Cells were seeded in six-well plates at a density of 3 x 10^5 cells per well and incubated overnight at 37ºC. Medium was replaced and cells were incubated with 1 μM of protein for 3 hours. Cells were washed three times in 1x DPBS and trypsinized. Following tyrpsinization, cells were neutralized with the appropriate cell media and centrifuged at 800 rpm for 4 minutes. Cells were resuspended in cold 1x DPBS and analyzed by flow cytometry (NovoCyte, ACEA Biosciences, Inc., San Diego, CA, USA). Fluorescence intensity of 1 x 10^4 cells was measured with 488 nm laser and 660 nm emission for mCherry. The median fluorescence intensity of three replicates was reported along with a histogram of one replicate for each sample (Figure S6).

EGFR immunostaining

Cells were seeded in 8-well plates with a collagen film and incubated for 24 hours before fixation and DAPI staining as previously described. Following fixation, cells were blocked with 3% BSA in 1x DPBS for 20 minutes at room temperature then incubated with 2.5 μg/mL Alexa Fluor 488 anti-EGFR Antibody, from Thermo Fisher, or Alexa Fluor 488 Rabbit IgG Isotype control from Cell Signaling Technology (Danvers, MA) in 1x DPBS for one hour. Unreacted antibody was removed by washing three times with 1x DPBS and labelled cells were imaged at 40x magnification on the fluorescence microscope with 470/40 nm excitation and 525/50 nm emission.

Flow cytometry was used as a quantitative analysis of EGFR level in IBC SUM149 and MCF10A cells. Cells were seeded in six-well plates at a density of 3 x 10^5 cells per well and incubated overnight at 37ºC. Cells were trypsinized and spun down at 800 rpm for 4 minutes. Media was removed and cells were incubated with 10% formalin for 10 minutes before being spun down and undergoing one wash step. 2.5 μg/mL of Alexa Fluor 488 conjugated anti-EGFR or IgG control in DPBS were incubated with the cells at room temperature for 1 hour followed by a wash step and resuspension in cold 1x DPBS. Fluorescence of 1 x 10^4 cells was measured with 488 nm laser and 532 nm emission for Alexa Fluor 488. The median fluorescence intensity of three replicates was reported along with a histogram of one replicate for each sample.

EGFR inhibition

IBC SUM149 cells were seeded at 5 x 10^4 cells per well in an 8-well collagen coated plate. Following overnight incubate at 37ºC, media was replaced and cells were incubated with 100 μM of EGF for one hour. Cells were washed three times with DPBS and incubated with 1 μM of GE11-mCherry protein for three hours. Media was removed and cells were washed three times in 1x DPBS before fixation and DAPI staining. Fluorescence imaging and flow cytometry analyses were conducted as described above.

SpyTag/SpyCatcher conjugation of yCD to GE11-mCherry

A gene fragment encoding SpyCatcher was amplified with PCR using primers SpyCatcher-F and SpyCatcher-R (Table S1). The PCR product was inserted into pET22b-w3-amber-mCherry-his6 constructs using XhoI and BlpI sites to yield pET22b-amber-mCherry-SpyCatcher-his6. Identical methods of plasmid preparation, protein expression, and purification were used as

described above for pET22b-amber-mCherry- his6. CuAAC was again used to conjugate GE11 to Az-mCherry-SpyCatcher using the same reactant concentrations as above. Following CuAAC, the protein was purified and concentrated with his-tag chromatography and dialyzed in 1x PBS. To couple the yCD, 15 μM GE11-mCherry-SpyCatcher was reacted with 15 μM of yCD in 1 x PBS (pH 7.4) for two hours followed again by purification with his-tag chromatography.

Prodrug Treatment and Cell Viability

In a 96 well plate, cells were seeded at 10,000 cells/well and incubated overnight. 1 μM of mCherry-yCD protein was added to cells and incubated for three hours. Cells were washed three times with 1x DPBS to remove protein that was not internalized, treated with 500 μg/mL of 5-FC and incubated for 48 hours at 37ºC. Following incubation, MTT Cell Proliferation assays from Thermo Fisher were performed according to the manufacturer's protocol.

Statistical Analyses

Results were reported as mean ± standard deviation except where noted. All experiments were replicated at least three times with unique protein batches, except for Tat constructs which were repeated with two unique batches of protein. Statistical significance was determined with an unequal variance T-test. Significance was accepted at $p < 0.05$.

Conflicts of Interest

The authors declare no competing conflicts of interest.

Acknowledgements

Thank you to Qirun Li for assisting with gel preparation and protein expression. RML was supported by grants from the National Science Foundation (1510817 and 1144726). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the view of the National Science Foundation.

Supporting Information

SDS-PAGE of Az-mCherry expression and purification; SDS-PAGE and MALDI-TOF mass spectrometry of GE11 conjugation by CuAAC; Flow cytometry histograms of mCherry and mCherry-yCD internalization; SDS-PAGE and reaction scheme of non-clustered GE11 conjugation via amine/NHS ester chemistry; Az-mCherry-SpyCatcher expression and GE11-mCherry-SpyCatcher conjugation to SpyTag-yCD; MTT assay of cell viability following yCD/5-FC treatment; DNA oligos and nucleotide sequences used in this study.

Abbreviations

UAA, unnatural amino acid; pAzF, 4-azido-l-phenylalanine; CuAAC, copper-catalyzed alkyne-azide cycloaddition; IBC, inflammatory breast cancer; EGFR, epidermal growth factor receptor; yCD, yeast cytosine deaminase; EGF, epidermal growth factor; 5-FC, 5-fluorocytosine; 5-FU, 5-fluorouracil; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; MALDI-TOF, Matrix-assisted laser desorption/ionization- time of flight.

1. Lagasse, H.; Alexaki, A.; Simhadri, V.; Katagiri, N.; Jankowski, W.; Sauna, Z.; Kimichi-Sarfaty, C., Recent advances in (therapeutic protein) drug development. *F1000Research* **2017,** *6* (113).
2. Analysis and Insight into Critical Drug Development Issues. *Tufts Center for the Study of Drug Development Impact Report* **2017**.
3. AL, D.; RJ, M., Formulation and delivery issues for monoclonal antibody therapeutics. *Advanced Drug Delivery Reviews* **2006,** *58* (5), 686-706.
4. LH., E.; S, C.; JA., C., Nanoparticles for Intracellular Delivery of Therapeutic Enzymes. *Journal of Pharmaceutical Sciences* **2014,** *103* (6), 1863-1871.
5. Nagahara, A.; Tuszynski, M., Nature Reviews Drug Discovery. *Nature Reviews Drug Discovery* **2011,** *10*, 209-219.
6. Grubb, J.; Vogler, C.; Sly, W., New Strategies for Enzyme Replacement Therapy for Lysosomal Storage Diseases. *Rejuvenation Research* **2010,** *13* (2), 229-236.
7. Yang, Y.; Kitagaki, J.; Dai, R.; Tsai, Y.; Lorick, K.; Ludwig, R.; Pierre, S.; Jensen, J.; Davydov, I.; Oberoi, P.; Li, C.; Kenten, J.; Beutler, J.; Vousden, K.; Weissman, A., Inhibitors of ubiquitin-activating enzyme (E1), a new class of potential cancer therapeutics. *Cancer Research* **2007,** *67* (19), 9472-9481.
8. Cramer, S.; Saha, A.; Liu, J.; Tadi, S.; Tiziani, S.; Yan, W.; Triplett, K.; Lamb, C.; Alters, S.; Rowlinson, S.; Zhang, Y.; Keating, M.; Huang, P.; DiGiovanni, J.; Georgiou, G.; Stone, E., Systemic depletion of L-cyst(e)ine with cyst(e)inase increases reactive oxygen species and suppresses tumor growth. *Nature Medicine* **2017,** *23* (1), 120-127.
9. Li, C.; Pazgier, M.; Li, C.; Yuan, W.; Liu, M.; Wei, G.; Lu, W.-Y.; Lu, W., Systemic Mutational Analysis of Peptide Inhibition of the p53-MDM2/MDMX Interations. *Journal of Molecular Biology* **2010,** *398* (2), 200-213.
10. Los, M.; Panigrahi, S.; Rashedi, I.; Mandal, S.; Stetefeld, J.; Essmann, F.; Schulze-Osthoff, K., Apoptin, a tumor-selective killer. *Biochimica et Biophysica Acta* **2009,** *1793* (8), 1335-1342.
11. Gaynor, A.; Chen, W., Induced prodrug activation by conditional protein degradation. *Journal of Biotechnology* **2017,** *260*, 62-66.
12. Mitragotir, S.; Burke, P.; Langer, R., Overcoming the challenges in administering biopharmaceuticals: formulation and delivery strategies. *Nature Reviews Drug Discovery* **2014,** *13* (9), 655-672.
13. Banerjee, S.; Aher, N.; Patil, R.; Khandare, J., Poly(ethylene glycol)-Prodrug Conjugates: Concept, Design, and Applications. *Journal of Drug Delivery* **2012**.
14. Kalkhof, S.; Sinz, A., Chances and pitfalls of chemical cross-linking with amine-reactive N-hydroxysuccinimide esters. *Analytical and Bioanalytical Chemistry* **2008,** *392* (1), 305-312.
15. Dinca, A.; Chien, W.-M.; Chin, M., Intracellular Delivery of Proteins with Cell-Penetrating Peptides for Therapeutic Uses in Human Disease. *International Journal of Molecular Sciences* **2016,** *17* (2), 263.
16. Veronese, F.; Mero, A., The Impact of PEGylation on Biological Therapies. *Biodrugs* **2008,** *22* (5), 315-329.
17. Cho, H.; Daniel, T.; Buechler, Y.; Litzinger, D.; Maio, Z.; Hays Putnam, A.-M.; Kraynov, V.; Sim, B.-C.; Bussell, S.; Javahishvili, T.; Kaphle, S.; Viramontes, G.; Ong, M.; Chu, S.; GC, B.; Lieu, R.; Knudsen, N.; Castiglioni, P.; Norman, T.; Axelrod, D.; Hoffman, A.; Schultz, P.; DiMarchi, R.; Kimmel, B., Optimized clinical performance of growth hormone with an expanded genetic code. *PNAS* **2011,** *108* (22), 9060-9065.

18. Liu, J.; Liu, M.; Zheng, B.; Yao, Z.; Xia, J., Affinity Enhancement by Ligand Clustering Effect Inspired by Peptide Dendrimers-Shank PDZ Proteins Interactions. *PLoS ONE* **2016,** *11* (2).
19. Mardilovich, A.; Kokkoli, E., Biomimetic Peptide-Amphiphiles for Functional Biomaterials: The Role of GRGDSP and PHSRN. *Biomacromolecules* **2004,** *5*, 950-957.
20. Almeda, D.; Wang, B.; Auguste, D., Minimizing antibody surface density on liposomes while sustaining cytokine-activated EC targeting. *Biomaterials* **2015,** *41*, 37-44.
21. Poon, Z.; Chen, S.; Engler, A.; Lee, H. i.; Atas, E.; von Maltzahn, G.; Bhatia, S.; Hammond, P., Ligand-clustered "patchy" nanoparticles for modulated cellular uptake and in vivo tumor targeting. *Angewandte Chemie* **2010,** *49* (40), 7266-7270.
22. J, L.; GE, W.; B, Z.; PS, A.; DM, E.; VR, M.; R, R., Computational model for nanocarrier binding to endothelium validated using in vivo, in vitro, and atomic force microscopy experiments. *Proceedings of the National Academy of Sciences of the United States of America* **2010,** *107* (38), 16530-16535.
23. G, T.; RO, D.; Magnani; JL; JT, P., Nanomolar E-Selectin Inhibitors: 700-Fold Potentiation of Affinity by Multivalent Ligand Presentation. *Journal of the American Chemical Society* **2001,** *123* (41), 10113-10114.
24. DR, E.; A, P.; Popik; V; A, T., Effect of ligand density, receptor density, and nanoparticle size on cell targeting. *Nanomedicine : nanotechnology, biology, and medicine* **2013,** *9* (2), 194-201.
25. Chatterjee, A.; Sun, S.; Furman, J.; Xiao, H.; Schultz, P., A versatile platform for single- and multiple-unnatural amino acid mutagenesis in Escherichia coli. *Biochemistry* **2013,** *52* (10), 1828-1837.
26. Liu, C.; Schultz, P., Adding New Chemistries to the Genetic Code. *Annual Review of Biochemistry* **2010,** *79*, 413-444.
27. Niu, W.; Guo, J., Expanding the chemistry of fluorescent protein biosensors through genetic incorporation of unnatural amino acids. *Molecular BioSystems* **2013,** *9* (12), 2961-2970.
28. Wang, N.; Li, Y.; Niu, W.; Sun, M.; Cerny, R.; Li, Q.; Guo, J., Construction of a live-attenuated HIV-1 vaccine through genetic code expansion. *Angewandte Chemie International Edition* **2014,** *5* (53), 4867-4871.
29. Axup, J.; Bajjuri, K.; Ritland, M.; Hutchins, B.; Kim, C.; Kazane, S.; Halder, R.; Forsyth, J.; Santidrian, A.; Stafin, K.; Lu, Y.; Tran, H.; Seller, A.; Biroc, S.; Szydlik, A.; Pinkstaff, J.; Tian, F.; Sinha, S.; Felding-Habermann, B.; Smider, V.; Schultz, P., Synthesis of site-specific antibody-drug conjugates using unnatural amino acids. *PNAS* **2012,** *109* (40), 16101-16106.
30. Li, Z.; Zhao, R.; Wu, X.; Sun, Y.; Yao, M.; Li, J.; Xu, Y.; Gu, J., Identification and characterization of a novel peptide ligand of epidermal growth factor receptor for targeted delivery of therapeutics. *The FASEB Journal* **2005,** *19* (14), 1978-1985.
31. Kleer, C.; van Golen, K.; Merajver, S., Molecular biology of breast cancer metastasis: Inflammatory breast cancer: clinical syndrome and molecular determinants. *Breast Cancer Research* **2000,** *2* (6), 423-429.
32. 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. *PNAS* **2012,** *109* (12), E690-E697.
33. Wals, K.; Ovaa, H., Unnatural amino acid incorporation in *E. coli*: current and future applications in the design of therapeutic proteins. *Frontiers in Chemistry* **2014,** *2* (15).

34. Nakamura, Y.; Ito, K., How protein reads the stop codon and terminates translation. *Genes to Cells* **1998,** *3*, 265-278.
35. Lev, D. C.; Kim, L. S.; Melnikova, V.; Ruiz, M.; Ananthaswamy, H. N.; Price, J. E., Dual blockade of EGFR and ERK1/2 phosphorylation potentiates growth inhibition of breast cancer cells. **2004,** *91* (4), 795-802.
36. Masuda, H.; Zhang, D.; Bartholomeusz, C.; Doihara, H.; Hortobagyi, G. N.; Ueno, N. T., Role of epidermal growth factor receptor in breast cancer. *Breast cancer research and treatment* **2016,** *136* (2), 331-345.
37. Fan, M.; Liang, X.; Yang, D.; Pan, X.; Li, Z.; Wang, H.; Bizhi, S., Epidermal growth factor receptor-targeted peptide conjugated phospholipid micelles for doxorubicin delivery. *Journal of Drug Targeting* **2015,** *24* (2), 111-119.
38. Mondal, G.; Kumar, V.; Shukla, S.; Singh, P.; Mahato, R., EGFR-Targeted Polymeric Mixed Micelles Carrying Gemcitabine for Treating Pancreatic Cancer. *Biomacromolecules* **2016,** *17* (1), 301-313.
39. Chen, J.; Ouyang, J.; Chen, Q.; Deng, C.; Meng, F.; Zhang, J.; Cheng, R.; Lan, Q.; Zhong, Z., EGFR and CD44 Dual-Targeted Multifunctional Hyaluronic Acid Nanogels Boost Protein Delivery to Ovarian and Breast Cancers In Vitro and In Vivo. *ACS Applied Materials & Interfaces* **2017,** *9*, 24140-24147.
40. Cheng, L.; Huang, F.-Z.; Cheng, L.-F.; Zhu, Y.-Q.; Hu, Q.; Li, L.; Wei, L.; Chen, D.-W., GE11-modified liposomes for non-small cell lung cancer targeting: preparation ex vitro and in vivo evaluation. *International Journal of Nanomedicine* **2014,** *9*, 921-935.
41. Liang, L.; Astruc, D., The copper(I)-catalyzed alkyne-azide cycloaddition (CuAAC) "click" reaction and its applications. An overview. *Coordination Chemistry Reviews* **2011,** *255* (23), 2933-2945.
42. Wu, I.-L.; Patterson, M. A.; Carpenter Desai, H. E.; Mehl, R. A.; Giorgi, G.; Conticello, V. P., Multiple site-selective insertions of noncanonical amino acids into sequence-repetitive polypeptides. *Chembiochem* **2013,** *14* (8), 968-978.
43. Johnson, D. B. F.; Xu, J.; Shen, Z.; Takimoto, J. K.; Schultz, M. D.; Schmitz, R. J.; Xiang, Z.; Ecker, J. R.; Briggs, S. P.; Wang, L., RF1 knockout allows ribosomal incorporation of unnatural amino acids at multiple sites. *Nature Chemical Biology* **2011,** *7* (11), 779-786.
44. Willmarth, N.; Baillo, A.; Dziubinski, M.; Wilson, K.; Riese, D.; Ethier, S., Altered EGFR localization and degradation in human breast cancer cells with an amphiregulin/EGFR autocrine loop. *Cell Signalling* **2009,** *21* (2), 212-219.
45. Wee, P.; Wang, Z., Epidermal Growth Factor Receptor Cell Proliferation Signaling Pathways. *Cancers* **2017,** *9* (5), 52-97.
46. Kim, D.; Yan, Y.; Valencia, A.; Liu, R., Heptameric Targeting Ligands against EGFR and HER2 with High Stability and Avidity. *PLoS ONE* **2012,** *7* (8), e43077.
47. Sempkowski, M.; Charles, Z.; Zofia Menzenski, M.; Kevrekidis, I.; Bruchertseifer, F.; Morgenstern, A.; Sofou, S., Sticky Patches on Lipid Nanoparticles Enable the Selective Targeting and Killing of Untargetable Cancer Cells. *Langmuir* **2016,** *32* (33), 8329-8338.
48. Brooks, H.; Lebleu, B.; Vives, E., Tat peptide-mediated cellular delivery: back to basics. *Advanced Drug Delivery Reviews* **2005,** *57* (4), 559-577.
49. Becker-Hapak, M.; McAllister, S.; Dowdy, S., TAT-mediated protein transduction into mammalian cells. *Methods* **2001,** *24* (3), 247-256.
50. Care, B.; Soula, H., Impact of receptor clustering on ligand binding. *BMC Systems Biology* **2011,** *5* (48).

51. Frohlich, E., The role of surface charge in cellular uptake and cytotoxicity of medical nanoparticles. *International Journal of Nanomedicine* **2012,** *7*, 5577-5591.
52. D, L.; YM, L.; J, K.; MK, L.; WJ, K., Enhanced tumor-targeted gene delivery by bioreducible polyethylenimine tethering EGFR divalent ligands. *Biomaterials Science* **2015,** *3* (7), 1096-1104.
53. Polak, A.; Eschenhof, E.; Fernex, M.; Scholer, H., Metabolic Studies with 5-Fluorocytosine-6-^{14}C in Mouse, Rat, Rabbit, Dog and Man. *Chemotherapy* **1976,** *22*, 137-153.
54. Noordhuis, P.; Holwerda, U.; Van der Wilt, C.; Van Groeningen, C.; Smid, K.; Meijer , S.; Pinedo, H.; Peters, G., 5-Fluorouracil incorporation into RNA and DNA in relation to thymidylate synthase inhibition of human colorectal cancers. *Annals of Oncology* **2004,** *15* (7), 1025-1032.
55. Longley, D.; Harkin, P.; Johnston, P., 5-Fluorouracil: mechanisms of action and clinical strategies. *Nature Reviews Cancer* **2003,** *3*, 330-338.
56. Wigmore, P.; Mustafa, S.; El-Beltagy, M.; L, L.; J, U.; G, B., Effects of 5-FU. *Advances in Experimental Medicine and Biology* **2010,** *678*, 157-64.
57. Kim, H.; Chen, W., A non-chromatographic protein purification strategy using Src 3 homology domains as generalized capture domains. *Journal of Biotechnology* **2016,** *234*, 27-34.
58. Presolski, S.; Phong Hong, V.; Finn, M. G., Copper-Catalyzed Azide-Alkyne Click Chemistry for Bioconjugation. *Current Protocols in Chemical Biology* **2011,** *3* (4), 153-162.
59. Amblard, M.; Fehrentz, J.; Martinez, J.; Subra, G., Methods and protocols of modern solid phase Peptide synthesis. *Molecular Biotechnology* **2006,** *33* (3), 239-254.
60. Lehman, H.; Van Laere, S.; van Golen, C.; Vermeulen, P.; Dirix, L.; van Golen, K., Regulation of inflammatory breast cancer cell invasion through Akt1/PKBalpha phosphorylation of RhoC GTPase. *Molecular Cancer Research* **2012,** *10* (10), 1306-1318.
61. Van den Eynden, G.; Van Laere, S.; Van der Auwera, I.; Merajver, S.; Van Marck, E.; van Dam, P.; Vermeulen, P.; Dirix, L.; van Golen, K., Overexpression of caveolin-1 and -2 in cell lines and in human samples of inflammatory breast cancer. *Breast Cancer Research and Treatment* **2006,** *95*, 219–228.
62. 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.