Protective Antigen-mediated Delivery of Biomolecules

By

Zeyu (Mike) Lu

B.A. Chemistry
Wabash College, 2012

SUBMITTED TO THE DEPARTMENT OF CHEMISTRY IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY IN CHEMISTRY
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2018

Signature of Author: Signature redacted

Department of Chemistry
August 31, 2018

Certified by: Signature redacted

Bradley L. Pentelute
Pfizer-Laubach Career Development Professor of Chemistry
Thesis Supervisor

Accepted by: Signature redacted

Robert W. Field
Haslam and Dewey Professor of Chemistry
Chair, Departmental Committee on Graduate Students

This doctoral thesis has been examined by a committee of the Department of Chemistry as follows:

Signature redacted

Matthew D. Shoulders
Whitehead CD Associate Professor of Chemistry
Thesis Committee Chair

Signature redacted

Bradley L. Pentelute
Pfizer-Laubach Career Development Professor of Chemistry
Thesis Supervisor

Signature redacted

John M. Essigmann
William R. and Betsy P. Leitch Professor of Chemistry and Biological Engineering

Protective Antigen-mediated Delivery of Biomolecule

By

Zeyu (Mike) Lu

Submitted to the Department of Chemistry
On 31 August 2018 in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

Abstract

The intracellular delivery of therapeutic biomolecules such as oligonucleotides and proteins remains a key challenge today. Protective antigen, a naturally evolved protein translocase derived from Bacillus anthracis, has shown promise as a platform of protein delivery due to its ability to form a transmembrane pore that allows the cargo to have cytosolic access. We and others have used the LF_N/PA system to deliver a wide variety of natural and non-natural peptides and proteins.

Despite the significant progress made with the LF_N/PA delivery platform, some aspects including cargo selection and targeting still remain limited. In the first part of the thesis, we greatly expand the application of the platform by demonstration of efficient delivery of peptide nucleic acids (PNAs), an oligonucleotide analog. Using this technology, we successfully exploited a cancer-specific gene dependency by the intracellular delivery of an anti-sense PNA in a receptor-dependent manner.

In addition to exploiting new types of cargo for delivery, we developed a new strategy to target the LF_N/PA system to specific cell types. In the second part of the thesis, we chemically conjugated a full-length immunoglobulin G (IgG) to a mutant PA (mPA). Significantly, we took advantage of the fact that PA activation is protease-dependent and created highly specific delivery vehicles that can only be activated by the concurrent presence of two entities on the cell surface. We showed a protein toxin delivered by these IgG-mPA variants effectively inhibited cell growth in different cancer cell lines and exhibited a significantly increased therapeutic window over previously reported PA variants both in vitro and in vivo.

In the last part of the thesis, we explored the possibility of simplifying the LF_N/PA system by directly ligating protein cargos to PA. In the absence of LF_N, the chemically created single-component system significantly increased the amount of delivered cargo. Moreover, the single-component system combined with a short N-terminal polylysine tag further improved the delivery efficiency by more than 100-fold. Our findings raise the prospect of a simpler PA-mediated delivery platform.

Thesis Supervisor: Bradley L. Pentelute
Pfizer-Laubach Career Development Professor of Chemistry

Acknowledgements

First and foremost, I would like to thank my Ph.D. advisor Prof. Brad Pentelute. I have come a long way since I first joined his lab. From the beginning, Brad, with his contagious enthusiasm about science, has encouraged me to think big. As a result of this, I have been constantly trying out new ideas and experiments outside of my comfort zone. During the process, Brad has provided me with the financial and intellectual support that I needed to succeed. Most importantly, he taught me how to be a better scientist and it has been a great pleasure working with him.

I would like to thank my thesis committee chair Prof. Matt Shoulders for the guidance and support during our annual meetings. I am also grateful that Prof. John Essigmann agreed to join my thesis committee the last minute after Prof. Alice Ting left.

My Ph.D. would not have been possible without the help from my labmates and collaborators. I want to thank Dr. Xiaoli Liao, Dr. Amy Rabideau, and Dr. Jingjing Ling for teaching me about the anthrax system as well as new experimental techniques. I also want to thank my other collaborators Dr. Brent Paolella, Prof. Rameen Beroukhim, Ms. Yiran Cheng, and Mr. Alex Loftis for directly working with me on various projects. I've also had many thought-provoking discussions with Dr. Peng Dai, Dr. Chi Zhang, and Dr. Nick Truex. These discussions had a profound impact on me during my graduate school. I've also enjoyed my interaction with many former and current lab members, including but not limited to Alex V., Anthony Q., Azin, Justin, Mark, Mette, Michael S., Suan.

Last but not least, I want to thank my family for having my back throughout my Ph.D. Although I have not been able to spend much time with them, they have been extremely understanding and supportive.

Table of Contents

List of Figures

Scope of Thesis

This thesis aims to develop novel strategies to improve different aspects of the protective antigen-mediated delivery, including cargo selection, targeting, and overall complexity.

Chapter 1 reports the delivery of a class of anti-sense oligonucleotide using the LF_N/PA system to target cancer cells. Delivery of therapeutic oligonucleotides into cells remains a major challenge. We studied protective antigen (PA), a component of anthrax toxin, as a non-viral, lipid-free delivery platform for peptide nucleic acids (PNAs). PNAs are versatile agents used to perturb different processes in the cytosol of cells. PNAs of different lengths and sequences were covalently linked to the 263-residue N-terminal domain of lethal factor (LF_N) via sortase-mediated ligation. We found the LF_N/PA system was able to transport all PNAs under study efficiently into the cytosol of mammalian cells. Furthermore, using this technology we successfully exploited a cancer-specific dependency on SF3B1. Over 10% of cancers have lost one or more copies of SF3B1, which renders these cancers susceptible to further SF3B1 suppression. This dependency on SF3B1 is one example of a class of over 120 CYCLOPS (Copy number alterations Yielding Cancer Liabilities Owing to Partial losS) dependencies in cancer. Anti-sense PNAs delivered by the LF_N/PA system suppressed SF3B1 expression and reduced cell viability of cancer cells with partial copy-loss of SF3B1, but not cells without SF3B1 loss. Moreover, PNA was delivered selectively to a HER2-overexpressing cancer cell line using a HER2-binding PA variant, demonstrating cell type-specific targeting. We conclude that our PA-mediated delivery system is a powerful tool for oligonucleotide-based antisense technologies and provides a potentially generalizable approach to target SF3B1 and dozens of other dependencies in cancer.

Chapter 2 describes a novel and widely applicable strategy to target the LF_N/PA delivery system to specific cell types of interest. Immunoglobulin G (IgG) is currently one of the most

important classes of therapeutics. Due to its wide application and favorable in vivo properties, there has been significant interest in using IgG as a drug carrier. Although some success has been achieved with small molecules and self-translocating protein toxins, the antibody-mediated delivery of the majority of proteins has been elusive. Here, we report a targeted protein delivery system by chemically conjugating a mutant protein translocase mPA to an IgG molecule. Using this system, we demonstrated the delivery of different protein cargos by different antibodies and overcame antibody drug resistance in refractory cancer cell lines. Significantly, we took advantage of the fact that PA activation is protease-dependent and created highly specific delivery vehicles that can only be activated by the concurrent presence of two entities on cell surface. We showed that a protein toxin delivered by these IgG-mPA variants had a significantly increased therapeutic window and effectively inhibited tumor growth both in vitro and in vivo.

Chapter 3 reports a novel single-component PA-mediated delivery system. Anthrax toxin is a tripartite toxin derived from *Bacillus anthracis*. The effector proteins lethal factor (LF) and edema factor (EF) are delivered into the cytoplasm of cells by protective antigen (PA). It has been shown that the N-terminal region of LF (LF_N) is capable of PA-dependent translocation. Here we report a single-component PA-mediated delivery system by chemically conjugating the cargo protein to PA. We showed that, in the absence of LF_N, the single-component approach can significantly increase the amount of delivered cargo. Moreover, we discovered that, when combined with a short N-terminal polylysine tag, the single-component system improved the delivery efficiency by more than 100-fold. Our findings raise the possibility of a simpler PA-mediated delivery platform.

Chapter 1: Targeting cancer gene dependencies via anthrax-mediated delivery of peptide nucleic acids

Author Contributions

Zeyu Lu, Brenton R. Paollela, Rameen Beroukhim, and Bradley L. Pentelute conceived the work; Brenton R. Paollela performed analyses on gene dependencies; Zeyu Lu prepared the PNAs, peptides, and proteins and performed the PNA translocation assays and some of the cell viability assays; Brenton R. Paollela, Meredith S. Brown, and John Busanovich performed western blot and some of the cell viability assays; Zeyu Lu and Brenton R. Paollela wrote the manuscript with input from Rameen Beroukhim and Bradley L. Pentelute.

1.1. Introduction

Due to their affinity and specificity for RNA and DNA targets, therapeutic oligonucleotides are promising drug candidates for a wide range of diseases including cancer, Alzheimer's and cardiovascular disorders[1]. Peptide nucleic acids (PNAs) are nucleic acid analogs in which the sugar-phosphate backbone is replaced by N-(2-aminoethyl)glycine units[2]. PNAs form strong complexes with complementary strands of DNA or RNA and, due to their pseudopeptide backbone, are not degraded by nucleases or proteases, thus making PNAs attractive candidates for antisense and anti-gene therapeutics[3,4].

One set of attractive therapeutic targets for PNAs are CYCLOPS (Copy number alterations Yielding Cancer Liabilities Owing to Partial losS) genes. CYCLOPS genes are cell-essential genes that often undergo partial copy-number loss in cancer cells. This partial copy-number loss leaves the cancer cells with lower expression of the gene relative to normal cells, generating a therapeutic window whereby further partial suppression of the gene can kill the cancer cells while leaving normal cells intact[5]. We have identified 124 candidate CYCLOPS genes; a typical cancer will exhibit partial loss of up to 47 such genes, suggesting these may be a rich source of potential therapeutic targets[5,6]. These features of CYCLOPS genes—their abundance and the sensitivity of cancers to their partial suppression—render CYCLOPS genes ideal targets for PNA therapeutics.

However, the uncharged PNA molecules are difficult to deliver across biological barriers such as the endosomal and plasma membranes. Various protocols such as co-transfection with DNA, microinjection, conjugation to peptides and electroporation have been devised, but none have been met with major success[7].

Fortunately, nature has evolved mechanisms to transport biomolecules into cells across these biological barriers[8]. One delivery system is anthrax toxin from *Bacillus anthracis*, a three-

component system containing the receptor-binding, pore-forming protein protective antigen (PA) as well as the enzymatic moieties lethal factor (LF) and edema factor (EF). PA_{83} binds to receptors on host cells and is cleaved by furin-family protease[9]. The resulting fragment (PA_{63}) self-assembles into the ring-shaped heptameric or octameric pre-pore,[10,11] forming complexes with LF and EF with high affinity. The complexes are then endocytosed and acidification triggers conformational rearrangement of the PA pre-pore to form an ion-conductive ß-barrel transmembrane pore[12]. LF and EF then translocate through the PA pore into the cytosol from N- to C-terminus.

This anthrax toxin system has been disarmed by removal of the C-terminal catalytic lobes of LF and subsequently used for delivery of proteins and other variants into cells[13,14]. Often cargo is covalently fused to the C-terminus of LF_N and in the presence of PA hitches a ride through the transporter formed in the endosome.

While the LF_N/PA system has been shown to transport various polypeptide cargo attached to the C-terminus of LF_N into mammalian cells[15,16], here we show for the first time the successful intracellular delivery of PNAs using LF_N/PA. Furthermore, we demonstrate the utility of the LF_N/PA system to effectively target a CYCLOPS dependency on SF3B1, an essential spliceosome component that undergoes partial loss in 11% of all cancers (https://portals.broadinstitute.org/tcga/home).

1.2. Results

1.2.1. Many cancer gene dependencies cannot be targeted using existing small molecules

To understand the general scale and features of candidate therapeutic targets in cancer, we integrated data describing the level of essentiality of 5711 genes from an shRNA screen across 216 cancer cell lines[17], with genome-wide genetic profiles of these cell lines (Figure 1.2.1a). We call genes that are essential in a particular cell "gene dependencies" for that cell. We hypothesized that gene dependencies that are only present in cells with specific somatic genetic alterations represent candidate cancer-specific dependencies, because normal cells would not share these genetic alterations. As cancer-specific dependencies, we therefore consider these genes to be potential therapeutic targets in cancer.

We identified gene dependencies associated with either mutations in six known driver genes (*KRAS*, *PIK3CA*, *BRAF*, *NRAS*, *CTNNB1* and *EGFR*) or with any copy-number alteration genome-wide. Mutations in five of the six driver genes (with the exception of EGFR) were associated with increased dependence on that driver; we also identified 342 additional gene dependencies associated with mutations in these driver genes. We previously identified 174 dependencies associated with copy-number alterations, that were previously published[6]. These dependencies include 124 CYCLOPS dependencies on genes that have undergone partial loss and 50 additional dependencies on genes that were in most cases not themselves altered by the copy-number change. Integration of the mutation and copy number associated dependencies from our previous work yielded 486 unique context-specific gene dependencies, comprising 8% of all genes analyzed.

We found that the vast majority (87%) of these candidate cancer gene dependencies do not belong to a straightforward "druggable" protein target classes. We classified each gene dependency

by its protein target class using the IUPHAR drug targets and family database (http://www.guidetopharmacology.org/targets.jsp). Of the 486 context-specific gene dependencies, 423 were not classified as belonging to a protein family that could be targeted by a small molecule (Figure 1.2.1b). On average, each cell line had 12.7 dependencies that were in an existing IUPHAR druggable target class and 82.1 dependencies that were not (Figure 1.2.1c, $p<10^{-15}$). These data suggest that a substantial fraction of all gene dependencies would benefit from alternative and generalizable approaches to target cancer vulnerabilities.

Antisense gene suppression represents one tractable approach to target cancer gene dependencies, but methods for delivery of such molecules remain limited. PNAs are stable molecules with the capacity to suppress gene translation or transcription. However, PNAs cannot cross cell membranes on their own. We investigated whether we could overcome obstacles to anti-sense PNA delivery by the LF_N/PA platform to enable their uptake into cells.

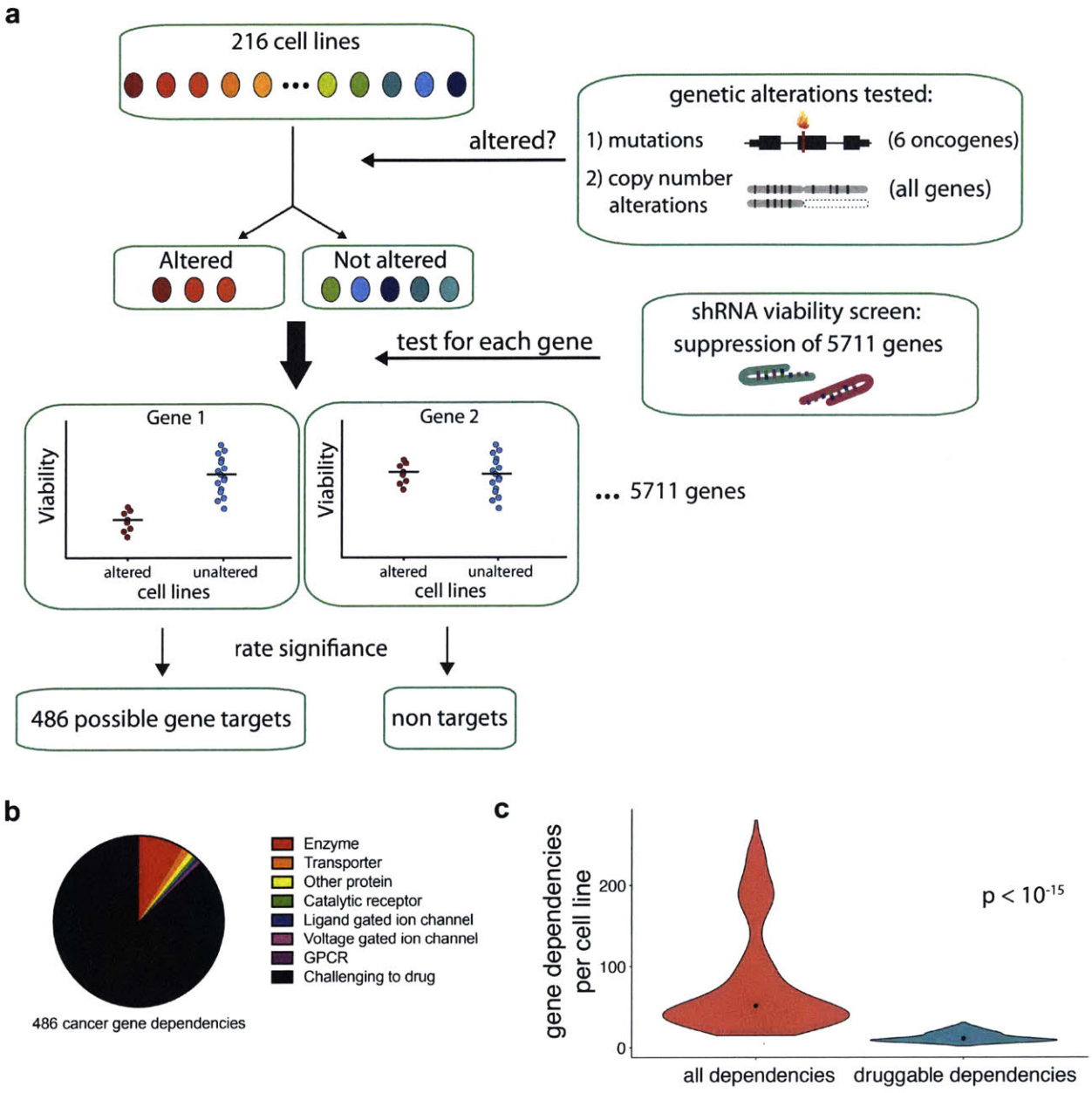

Figure 1.2.1. Many cancer gene dependencies are challenging to target with existing therapeutic approaches. a) Schematic describing our approach to detect context-specific cancer gene dependencies. b) Pie chart indicating the IUPAR target classification for the 486 cancer gene dependencies identified. c) Violin plot quantitating the number of potent gene dependencies per cell line. Candidate gene dependencies that belonged to an existing IUPHAR target class were considered "druggable". Width of plots represents relative sample density, dots represent sample median.

1.2.2. Protective antigen efficiently delivers PNAs of different lengths

We investigated the delivery of various PNA constructs into the cytosol of mammalian cells using the LF_N/PA system (Figure 1.2.2a). First, we synthesized a panel of PNAs of different lengths (Figure 1.2.2b). Each PNA was conjugated to the C-terminus of LF_N-DTA or LF_N using sortase-mediated ligation[18] to give LF_N-DTA-PNA or LF_N-PNA conjugates (LDP or LP respectively; Figures 1.7.1-2). DTA is the A chain of diphtheria toxin that catalyzes the ADP-ribosylation of elongation factor 2 and inhibits protein synthesis.

We investigated the translocation efficiency of LF_N-DTA conjugates using a protein synthesis inhibition assay[19,20]. CHO-K1 cells were incubated with serial dilutions of purified LDP1-9 conjugates in the presence of 20 nM of PA for 2 hours and then incubated with ^{3}H-Leu medium. ^{3}H-Leu incorporation within proteins indicated the translocation efficiency of cytosolic LF_N-DTA-PNA conjugates. We found that all PNA cargo translocated in the presence of PA, regardless of PNA length and sequence (Figure 1.2.2c). We further verified that the LDPs translocated via the LF_N/PA mechanism using controls that disrupt cellular entry. Cells treated with LDP9 in the presence of PA[F427H], a PA mutant that arrests translocation in the endosome[21,22] or Bafilomycin A1, an inhibitor of endosomal acidification, did not show any protein synthesis inhibition, indicating a functional PA and an acidic endosome are essential for translocation. After confirming translocation, we then proceeded to work with conjugates containing only LF_N, as the toxic protein DTA was removed to eliminate any interference with normal cellular functions.

We further investigated the delivery of the PNA conjugates by Western blot. HEK-293T cells were incubated with 250 nM of LP9 or a scrambled version of LP9 in the presence of PA for 12 hours. We collected the cytosolic lysates using digitonin as described before[23] and immunostained for the early endosome marker, Rab5, and cytosolic marker, Erk1/2 to confirm

cytosolic extraction. Western blots indicated cytosolic delivery of LP9 and its variant (Figures 1.2.2d and 1.7.4). The band intensity of LP9 was comparable to 5 ng of purified wild-type LF_N, indicating efficient cytosolic translocation. We observed a linear relationship between protein concentration and anti-LF staining intensity (Figure 1.7.5), and estimated a total of 3.0 ng of LP9 was delivered to 1 million cells, resulting in 51,000 molecules per cell or 84 nM cytosolic concentration. Additionally, the molecular mass shift observed for LP9 compared to LF_N indicates the PNA cargo was intact. In order to confirm most of the delivered protein was in the cytosol, we lysed the cells by total lysis buffer containing 1% NP-40. As expected, a similar amount of LP9 was observed in the cytosolic fraction and total cell lysate, indicating efficient translocation of LF_N conjugates from the endosome into the cytosol. As predicted, delivery with PA[F427H] did not allow for LP9 delivery, indicating the translocation process requires a functional PA. In addition to LP9, all LDP constructs studied were also subjected to western blot and showed similar results (Figure 1.7.6). As an additional control, Bafilomycin A1, an inhibitor of endosomal acidification, was used to validate the translocation pathway. These findings show that the LF_N/PA translocation system can be used as a non-viral lipid-free platform to deliver various PNAs.

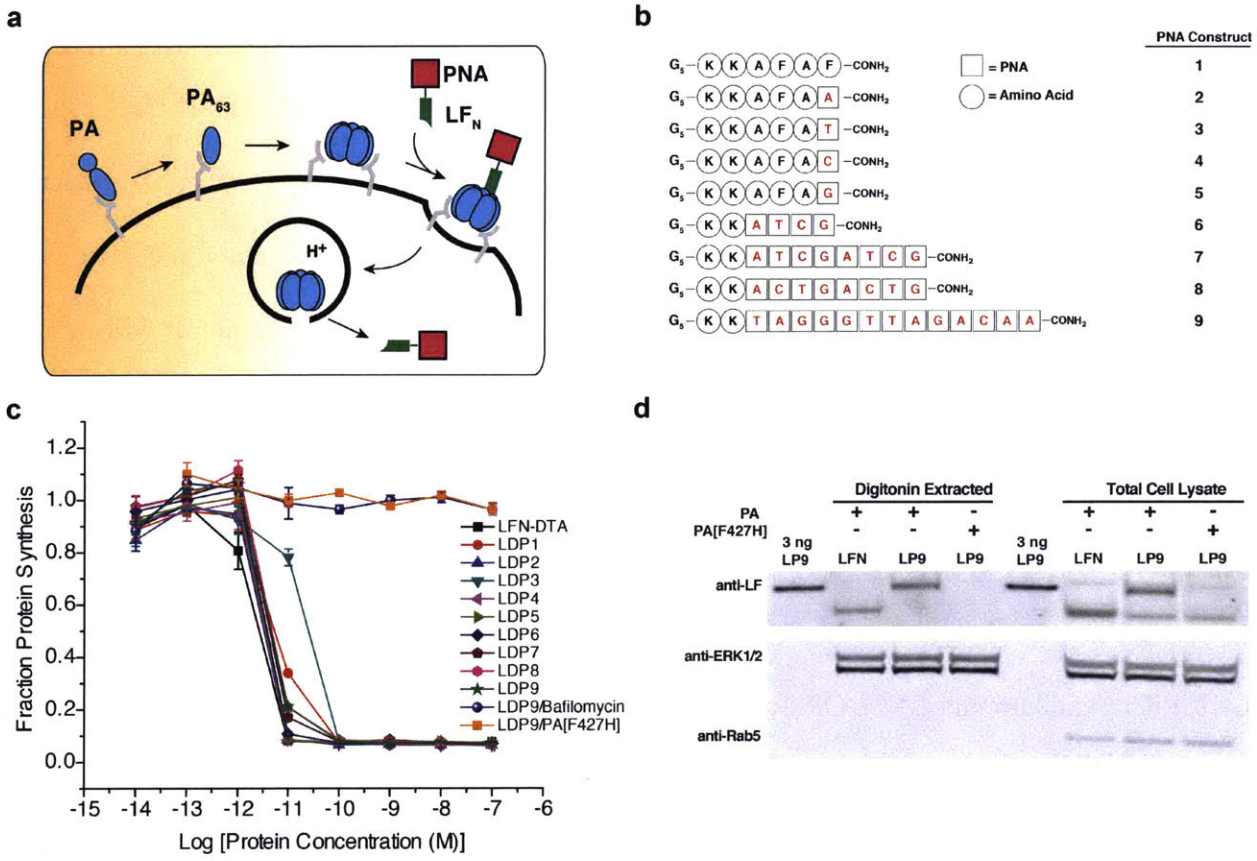

Figure 1.2.2. The non-toxic LF_N/PA platform can efficiently deliver PNAs of various lengths into cells. a) Mechanism of translocation of LF_N/PA system into cells. The square represents the PNA to be delivered. b) Peptide nucleic acids 1-9 from N-terminus to C-terminus. All PNAs have N-terminal Gly_5 tags. PNA9 is a 13-mer molecule used as a model to evaluate endosomal translocation. c) Protein synthesis inhibition assay using CHO-K1 cells treated with varying concentrations of LDP1-9 in the presence of 20 nM PA. After the treatment, the uptake of ^{3}H-Leucine is measured by scintillation counter to determine the level of protein synthesis inhibition. The radioactive counts were normalized to cells treated with only PA. PA[F427H] is a PA mutant that is incapable of translocation. Bafilomycin A1 is a compound that inhibits endosomal acidification. d) Western blot showing the translocation of LP9. HEK-293T cells were treated with 250 nM of LP9 and 40 nM of PA or PA[F427H] for 12 hours. The cells were subsequently subjected to digitonin extraction or total lysis (~1 million cells per lane). Purified LP was also run on the gel for the purpose of quantification.

1.2.3. Anthrax-delivered PNAs effectively exploit cancer cell dependencies

We next evaluated the ability of the LF_N/PA PNA system to target CYCLOPS cancer vulnerabilities. We have previously identified 124 CYCLOPS genes across the genome. Cancers that lose one or more copies of these genes are more sensitive to further suppression of that gene compared to normal cells with two copies of the gene[5] (Figure 1.2.3a). A typical cancer will exhibit partial loss of up to 47 CYCLOPS genes, suggesting they are a rich source of potential therapeutic targets. We have determined the mechanisms of selective vulnerability for two CYCLOPS genes, the proteasome component *PSMC2*, and the splicing factor *SF3B1*.

We previously generated isogenic cells differing only in their SF3B1 copy-number using CRISPR to validate this CYCLOPS gene dependency[6]. We designate cells lacking alterations in SF3B1 copy-number as $CRISPR^{neutral}$, while cells that exhibit partial SF3B1 copy-loss are designated $CRISPR^{copy\text{-}loss}$. In both cases in which CYCLOPS vulnerabilities have been validated, small molecules were unsuccessful in targeting activity of the gene product or the complex in which it resides. Consistent with our previous observations[6], we found that $CRISPR^{copy\text{-}loss}$ cells do not exhibit enhanced sensitivity towards splicing inhibition by Splicesostatin A[24], a small molecule that binds SF3B1, when compared to $CRISPR^{neutral}$ (Figure 1.7.7). These data suggest an unmet need for novel methods to target CYCLOPS gene dependencies that could also be expanded to suppress the hundreds of candidate gene dependencies we identified.

We next established the SF3B1 CYCLOPS vulnerability prior to testing LF_N/PA PNA targeting of SF3B1. We confirmed by immunoblot that SF3B1 copy loss reduces SF3B1 protein levels (Fig. 1.7.8). We then characterized the sensitivity of $CRISPR^{neutral}$ and $CRISPR^{copy\text{-}loss}$ cells to partial SF3B1 suppression by siRNA to validate the SF3B1 CYCLOPS phenotype. Transfections of increasing concentrations of siRNAs targeting SF3B1, ranging from 12.5-200

nM, resulted in dose-dependent SF3B1 mRNA suppression (Figure 1.7.9a). SF3B1 suppression was greater in cells with SF3B1 copy-loss because they express reduced levels of SF3B1 at baseline. Consistent with the CYCLOPS phenotype, partial *SF3B1* suppression in $CRISPR^{copy\text{-}loss}$ cells reduced their viability at 25-100 nM siRNA concentrations, which were tolerated by $CRISPR^{neutral}$ cells (Figure 1.7.9b).

We therefore investigated whether the SF3B1 CYCLOPS vulnerability could be exploited by the non-viral lipid-free LF_N/PA system. We synthesized LF_N-PNAs complementary to SF3B1 and characterized their abilities to bind and suppress SF3B1 expression. We synthesized three anti-sense SF3B1 PNAs and one sense SF3B1 PNA as a negative control (Figure 1.2.3b). One of the antisense SF3B1 PNA locations was designed to hybridize close to the start codon (Figure 1.2.3c), which has been reported to enhance translation inhibition[25]. The PNA that optimally suppressed SF3B1 expression (named anti-sense SF3B1-PNA) hybridizes to the 144 position on SF3B1 (Figure 1.7.10).

We confirmed pharmacodynamic activity of anti-sense SF3B1-PNA. Anti-sense SF3B1-PNA suppressed SF3B1 expression in both $CRISPR^{neutral}$ and $CRISPR^{copy\text{-}loss}$ cell lines (Figure 1.2.3d), and did so in a dose-dependent manner (Figure 1.2.3e). We also detected LF_N in cellular lysates by immunoblot, validating the mechanism of cytosolic delivery for the sense and anti-sense SF3B1-PNAs (Figures 1.2.3e and 1.7.11).

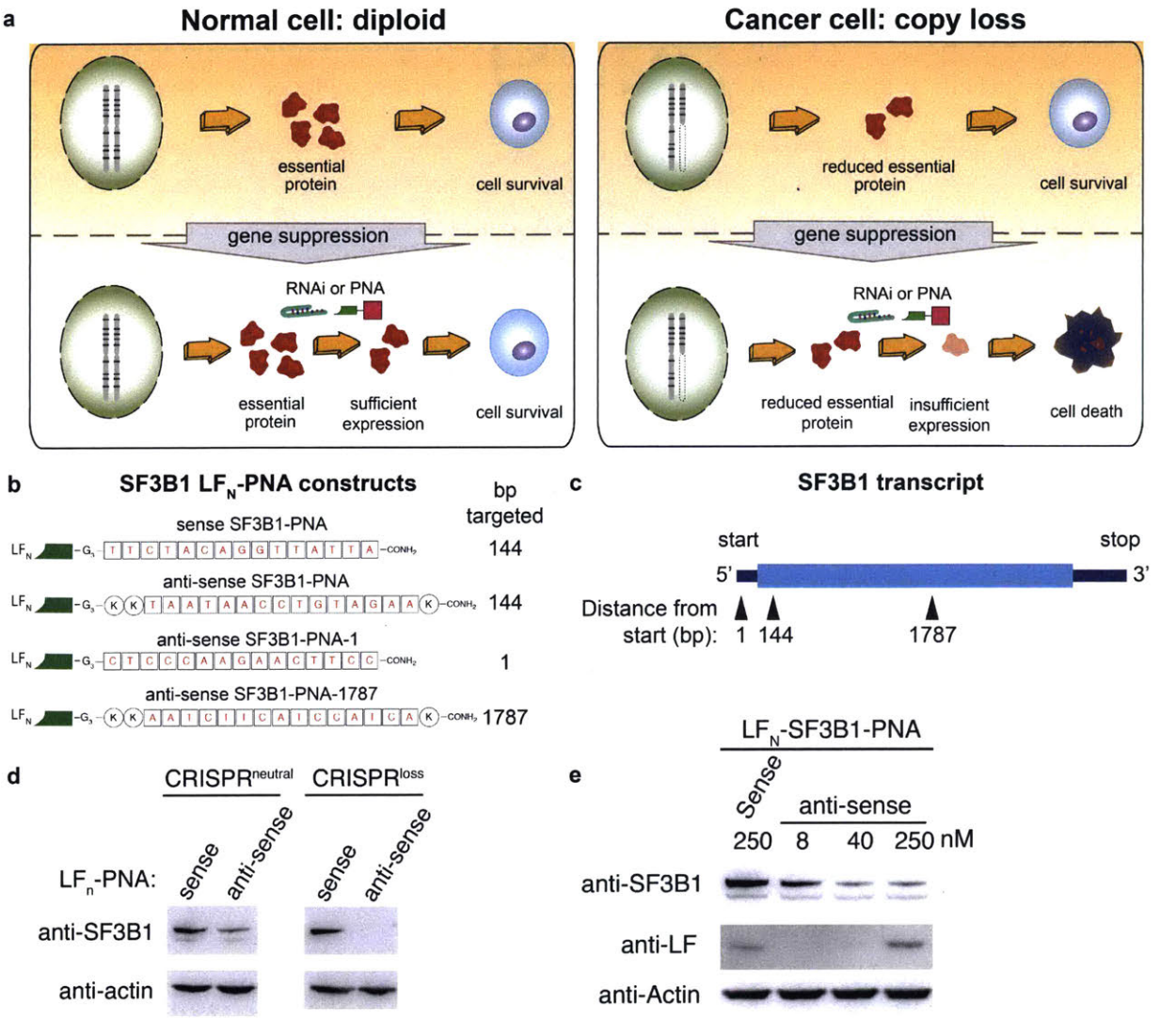

Figure 1.2.3. Targeting SF3B1 by anti-sense SF3B1-PNAs leads to suppression of SF3B1. a) Schematic illustrating the CYCLOPS gene dependency concept. Normal cells express excess amounts of an essential protein that allows them to tolerate partial gene suppression by RNAi or PNAs. Whereas cancer cells with partial copy loss express less of the essential protein and are sensitized to further gene suppression resulting in cell death. b) Schematic of SF3B1 PNA sequences synthesized and conjugated to LF_N. c) Schematic illustrating the location of anti-sense SF3B1 PNAs on the SF3B1 mRNA transcript. Arrowheads mark the locations of the PNAs relative to their distance from the translational start site. d) SF3B1 immunoblot blot demonstrating suppression of SF3B1 in both $CRISPR^{neutral}$ and $CRISPR^{copy\text{-}loss}$ cells by anti-SF3B1 LF_N-PNAs. e) SF3B1 and anti-LF immunoblot in $CRISPR^{neutral}$ cells with 8, 40, or 250 nM LF_N-conjugated antisense SF3B1 PNA in the presence of 50 nM PA. 250 nM LF_N-conjugated SF3B1 sense PNA was used as a control. Lysates were collected four days after LF_N/PA PNA delivery.

Therefore, we tested whether LF_N/PA delivered anti-SF3B1 PNAs could target the SF3B1 CYCLOPS vulnerability and selectively kill $SF3B1^{loss}$ cells. $CRISPR^{neutral}$ and $CRISPR^{copy\text{-}loss}$ cells were treated with 7-1000 nM anti-sense SF3B1-PNA in the presence of 50 nM PA, while 1000 nM of sense SF3B1-PNA treatments served as controls. $CRISPR^{copy\text{-}loss}$ cells had significantly reduced viability upon anti-sense SF3B1-PNA/PA exposure from 125-1,000 nM that was not observed in $CRISPR^{neutral}$ cells or with sense SF3B1-PNA controls (Figures 1.2.4a-b). These data suggest that the LF_N/PA delivery platform can be used to effectively target CYCLOPS vulnerabilities, including those that cannot be targeted with existing small molecules.

Native anthrax receptors (ANTXR1 and ANTXR2) are ubiquitously expressed across human tissues, suggesting that LF_N/PA mediated delivery of PNAs could target genetic dependencies in many cell lines and diverse tumor types. To evaluate the general applicability of the LF_N/PA PNA delivery system, we tested anti-sense SF3B1 suppression in nine additional cell lines from breast and blood lineages. These cell lines represent varied cellular contexts including adherent and non-adherent growth conditions, and molecular subtypes of cancer from unique cells of origin.

Delivery of anti-sense SF3B1 PNAs by LF_N/PA resulted in significantly reduced cell viability in $SF3B1^{loss}$ cell lines with naturally occurring SF3B1 copy loss from breast and blood cancers. SF3B1 suppression by PNA delivery resulted in ~50% reduced viability in $SF3B1^{loss}$ cell lines, BT549 and HCC1954, but not $SF3B1^{neutral}$ cell lines HS578T and MCF7 (Figure 1.2.4c). We also observed significantly reduced cell viability upon anti-sense SF3B1 PNA delivery in blood cancer cell lines with SF3B1 copy-loss (Figure 1.2.4d).

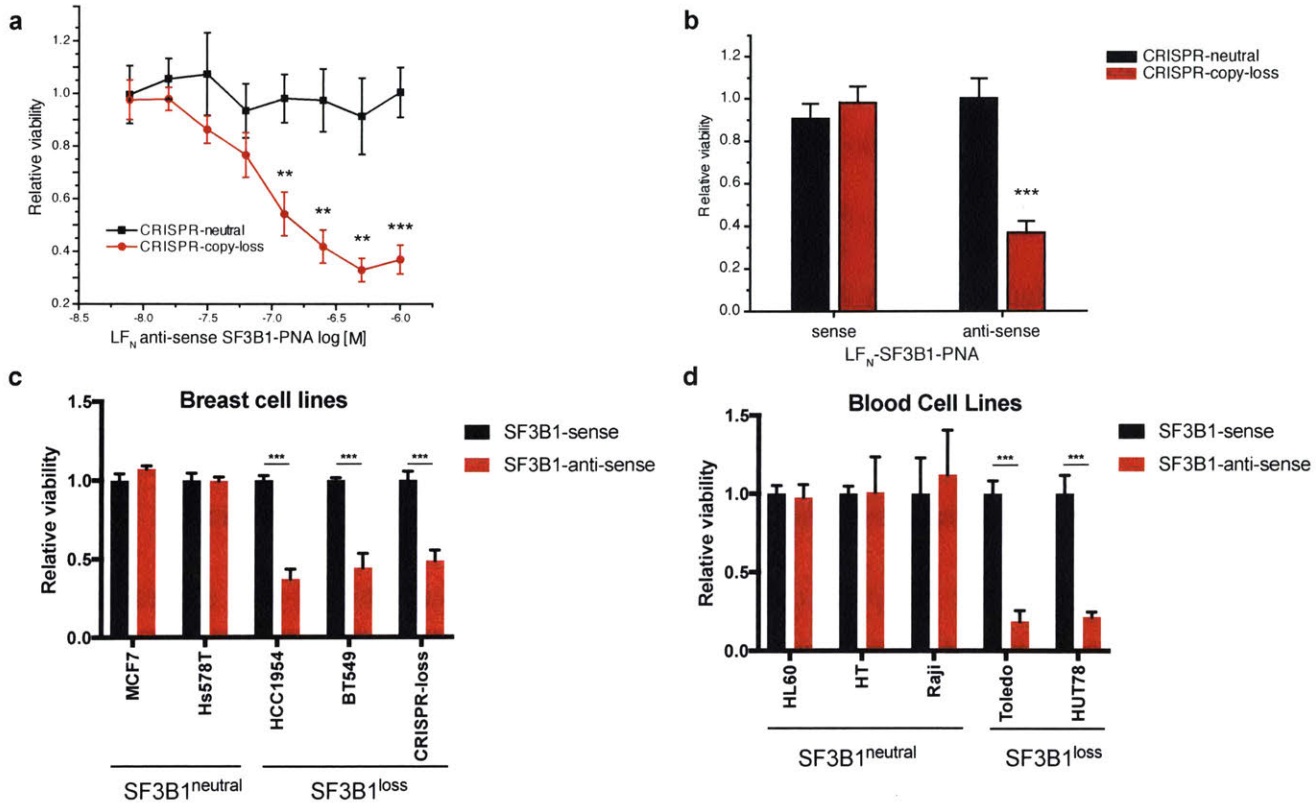

Figure 1.2.4. Anthrax-mediated delivery of anti-sense SF3B1-PNAs induces cell death in SF3B1^{loss} cells across multiple cell types. a) Dose response curve for relative cell viability of CRISPR^{neutral} (black) and CRISPR$^{copy\text{-}loss}$ (red) cells to LF_N/PA mediated delivery of anti-sense SF3B1 PNAs. Cell viability calculated as the fold change in Cell Titer-Glo luminescence relative to PNA only (no delivery system) control. Representative experiment performed in duplicate. b) Viability of cells treated with LF_N-conjugated SF3B1 sense in parallel experiments. c,d) Breast and blood cell lines with SF3B1^{neutral} or SF3B1^{loss} were treated with 100 nM of sense or antisense LF_N-SF3B1 in the presence of 50 nM wild-type PA. Viability of cells was measured by Cell Titer-Glo and normalized to untreated cells. All experiments were performed in triplicate. For all panels *p<0.05, **p<0.01, and ***p<0.001.

One significant advantage of the LF_N/PA system compared to other delivery systems is its ability to target specific cell types as demonstrated by others previously. To this end, we treated $ERBB2^{neutral}$ BT549 and $ERBB2^{amp}$ HCC1954 cells with PNAs and a known chimeric PA variant that is fused with an affibody (PA-HER2) that targets ERBB2. Although both cell lines were sensitive to anti-sense PNA delivery by wild type PA (Figure 1.2.4c), only $ERBB2^{amp}$ HCC1954 cells were susceptible to anti-sense SF3B1-PNAs delivered by PA-HER2 (Figure 1.2.5a). As controls, PA-HER2 treatment without PNA or with sense SF3B1-PNA did not yield any significant change in cell viability. Consistent with the specificity of the HER2-targeting affibody, treatment of $ERBB2^{neutral}$ Cal51-$SF3B1^{loss}$ cells, known to be sensitive to SF3B1 suppression, did not exhibit reduced cell viability with PA-HER2 delivered anti-sense SF3B1 PNAs but did have reduced viability when treated with wild-type PA targeted PNAs (Figure 1.2.5b). Taken together, these data suggest that the PA system is a potent PNA delivery system that can also be readily engineered to retarget PNA delivery to endogenous cancer cell specific receptors, such as ERBB2.

Since cell penetrating peptide (CPP)-mediated delivery of PNAs have been previously described[26], we directly evaluated the potency of the LF_N/PA system compared to a commonly used CPP TAT. Anti-sense SF3B1 PNA was conjugated to TAT via "click chemistry" and the resulted conjugate TAT-PNA was purified by HPLC. $SF3B1^{loss}$ Cal51 or HCC1954 cells were incubated with TAT-PNA or LF_N-PNA in the presence of PA or PA-HER2. Unmodified PNA was included as a negative control. Both cell lines exhibited dose-dependent viability reduction as the result of LF_N-PNA/PA treatment and only $ERBB2^{amp}$ HCC1954 showed susceptibility to LF_N-PNA/PA-HER2 treatment (Figures 1.2.5c-d), consistent with our previous observations (Figures 1.2.5a-5b). However, neither TAT-PNA nor unmodified PNA resulted in any appreciable effect on viability up to a maximum of 5 uM. These results suggest LF_N/PA system is 100- to 1000-fold

more potent for PNA delivery than TAT-conjugated CPPs, possibly due to the interaction between the CPP and serum and the superior ability of the LF_N/PA system to release PNA from the endosome.

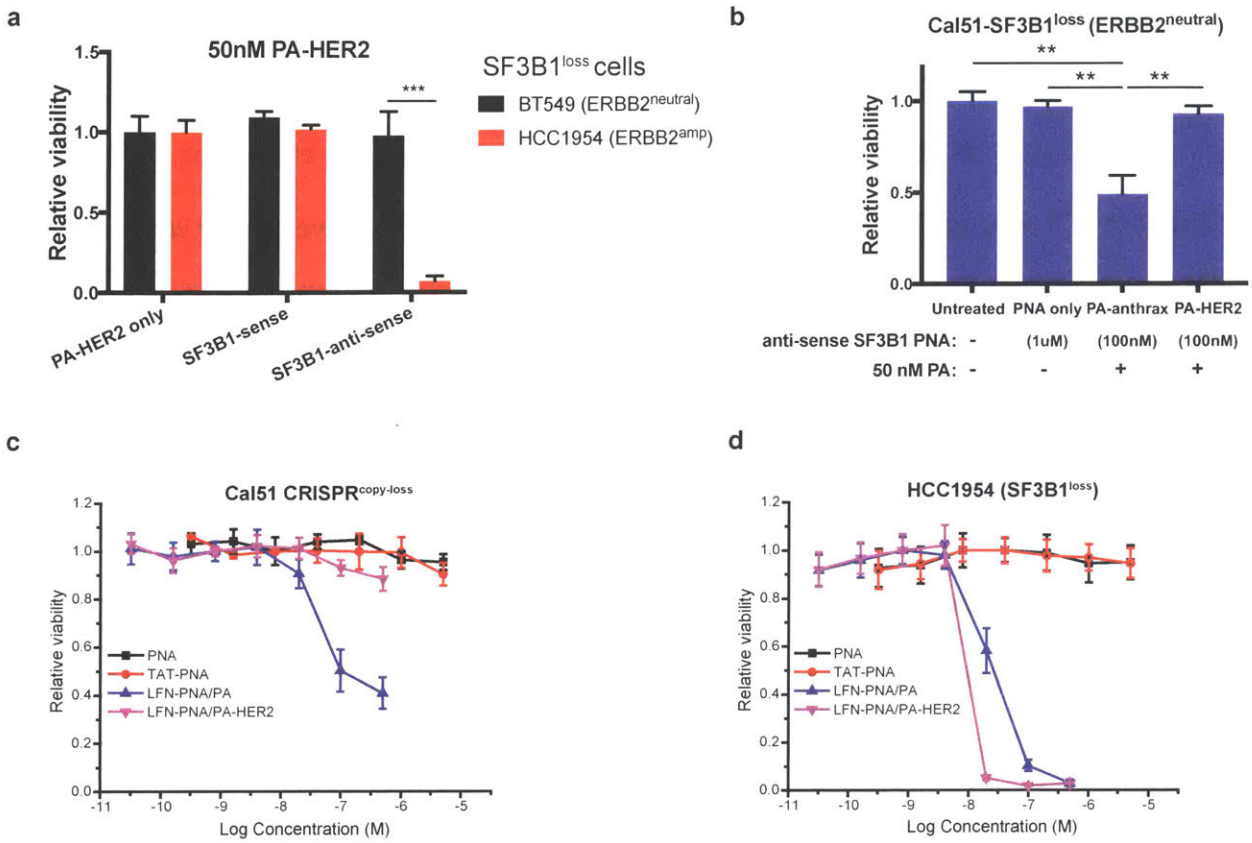

Figure 1.2.5. Cancer specific retargeting of Anthrax-mediated anti-sense SF3B1 PNAs induces cell death in ERBB2 amplified breast cancer cells. a) Viability of ERBB2^{neutral} BT549 and ERBB2^{loss} HCC1954 cells after treatment with sense or antisense LF_N-SF3B1 in the presence of PA-HER2, a ERBB2-specific PA variant. b) Viability of ERBB2^{neutral} Cal51 with SF3B1^{loss} after delivery of sense LF_N-SF3B1 using wild-type PA or PA-HER2, non LF_N-conjugated antisense SF3B1 PNA ("PNA only") was used as a control. c,d) Dose response curves of SF3B1^{loss} Cal51(ERBB2^{neutral}) and HCC1954(ERBB2^{amp}) cells treated with antisense TAT-PNA or antisense LF_N-PNA in the presence of PA or PA-HER2. For all panels, error bars represent +-SD, *$p<0.05$, **$p<0.01$ and ***$p<0.001$.

We then determined the specificity of anti-SF3B1 PNA delivery through a rescue experiment. Specifically, we asked whether enhanced SF3B1 expression could prevent the loss of cell viability upon SF3B1 suppression. We first confirmed increased expression of SF3B1 in CRISPR$^{copy\text{-}loss}$ cells transfected with an SF3B1 expression vector relative to GFP controls, in the context of treatment with anti-sense SF3B1-PNA (Figure 1.2.6a). We then found that this exogenous SF3B1 expression increased cell viability upon LF_N/PA delivery of anti-sense SF3B1 PNAs, relative to GFP-expressing controls (Figure 1.2.6b).

We further confirmed that anti-sense SF3B1-PNAs reduced viability of SF3B1^{loss} cells by interfering with splicing. To this end, we quantified splicing activity using dual luciferase reporters. One vector contained a luciferase ORF interrupted by an intron (Luc-I), which required splicing for its function, while the second vector was a control luciferase ORF that does not require splicing for its activity (Luc-ORF) (Figure 1.2.6c). Upon anti-sense SF3B1-PNA delivery, *SF3B1^{loss}* cells had significantly decreased splicing activity compared to *SF3B1^{neutral}* cells with SF3B1 suppression (Figure 1.2.6d). We also observed a modest, but nonetheless significant, decrease in splicing activity in SF3B1^{neutral} cells treated with antisense SF3B1 PNAs compared to control SF3B1 sense PNAs. These data indicate that anthrax-mediated delivery of PNAs can be used to target the SF3B1 CYCLOPS gene dependency.

In addition to targeting SF3B1, I also investigated the delivery of anti-sense PNA to inhibit telomerase as another application for this technology. Since it is not a part of the CYCLOPS work, its information can be found in section 1.7.2.

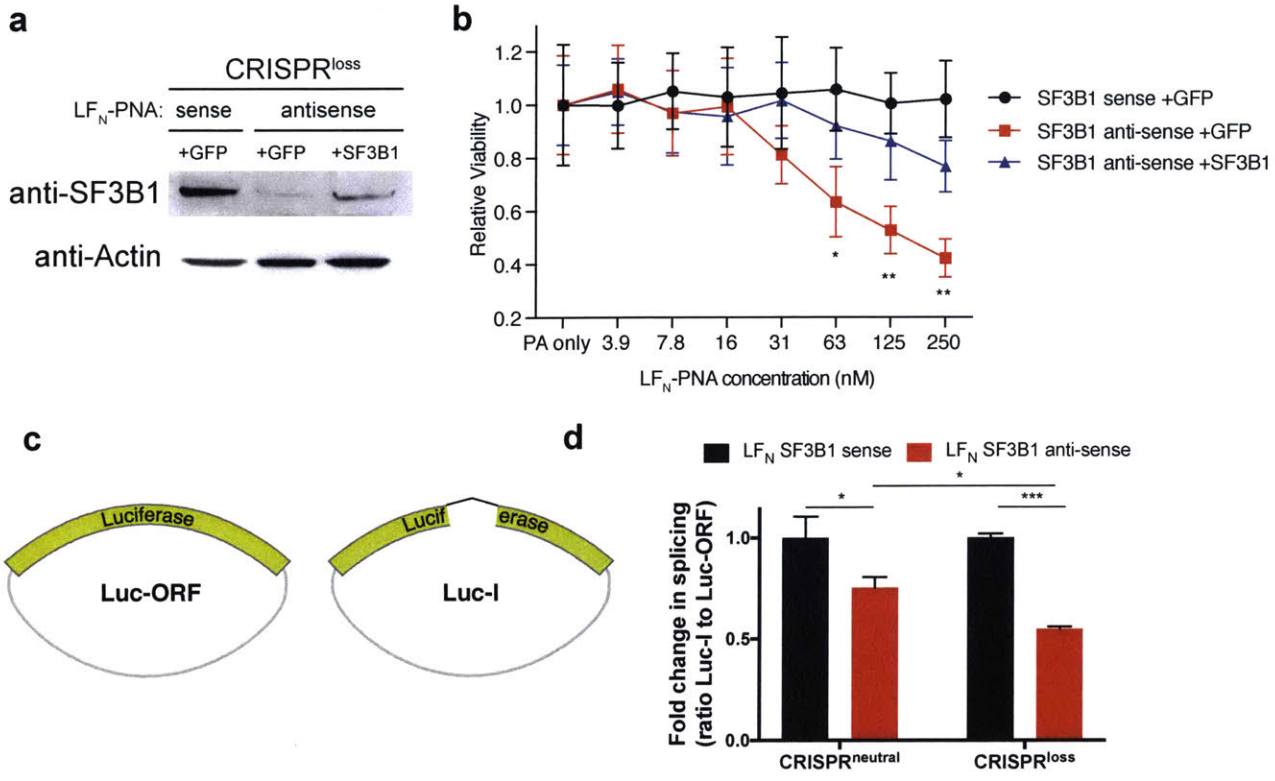

Figure 1.2.6. Specificity of LF_N/PA mediated delivery of anti-sense SF3B1 PNAs. a) SF3B1 immunoblot in $CRISPR^{copy\text{-}loss}$ cells expressing either GFP or exogenous SF3B1. Cells were subsequently treated with sense or anti-sense LF_N-SF3B1 PNAs in the presence of PA. b) Dose response curve of relative cell viability in $CRISPR^{copy\text{-}loss}$ cells upon LF_N/PA delivery of SF3B1 PNAs expressing either GFP or exogenous SF3B1 as in (a). $CRISPR^{copy\text{-}loss}$ cells with anti-sense SF3B1 delivery expressing GFP (red) had decreased viability compared to $CRISPR^{copy\text{-}loss}$ cells with the same anti-sense PNA with enhanced SF3B1 expression (blue). $CRISPR^{copy\text{-}loss}$ cells expressing GFP with sense targeting SF3B1 PNAs served as a control (black). Cell viability calculated as the fold change in Cell Titer-Glo luminescence relative to PA only (no LF_N-PNA) control. Representative experiment performed in triplicate. c) Schematic of luciferase reporters used to evaluate changes in pre-mRNA splicing. Luc-ORF contains a luciferase open reading frame that does not require splicing for its activity. Luc-I contains a luciferase open reading frame with an intronic sequence that requires splicing for its activity. d) Quantitation of splicing activity after LF_N/PA SF3B1-sense (black) or SF3B1-anti-sense (red) PNA delivery. Splicing activity was quantified as the ratio of Luc-I to Luc-ORF signal after normalizing to SF3B1 sense PNA as a control. For all panels, error bars represent +-SD, *p<0.05, **p<0.01 and ***p<0.001.

1.3. Discussion

Recent profiling of cancer gene dependencies by loss of function screens have enabled systematic identification of candidate therapeutic targets in cancer[17,27]. However, our analyses suggest the number of cancer gene dependencies that can be directly targeted with traditional small molecule approaches remains limited. Therefore, substantial opportunity exists for alternative therapeutic approaches to target the spectrum of gene dependencies identified. While many modalities will be needed, potential therapeutic approaches that can phenocopy gene suppression would have broad appeal.

Here we developed an anthrax-derived LF_N/PA system that could efficiently deliver PNAs into the cytosol. Our data indicate that PNAs could 'piggy back' on the C-terminus of LF_N and translocate through PA into the cytosol. The translocation efficiency was found to be independent of the PNA length for the constructs tested, demonstrating a unique advantage of the LF_N/PA system. The PA nanomachinery is quite promiscuous and once translocation is initiated by LF_N, the C-terminal extension with PNAs can translocate through the 12 Å pore[28]. In addition, this assay allowed us to quantify the amount of material translocated. Contrary to previous delivery techniques where most of delivered material is trapped in the endosome[7], we found most of the internalized LF_N-PNA is in the cytosol. These studies suggest that the LF_N/PA system is highly efficient with approximate cytosolic delivery as high as around 100 nM.

The efficiency and versatility of the LF_N/PA platform allowed us to deliver an anti-sense PNA that suppresses the CYCLOPS gene, SF3B1, to specifically target $SF3B1^{loss}$ cancer cells in a receptor-specific manner. PNAs and the LF_N/PA system are ideally suited as a generalizable non-viral lipid-free strategy to target CYCLOPS genes for three reasons: (i) The LF_N/PA system enables superior endosomal release of PNA while showing no obvious toxicity up to the highest

concentration we tested; (ii) While the native LF_N/PA delivery system can be used to deliver PNA cargos to most cell types due to the ubiquitous expression of the ANTXR1/2 receptors, PA can be readily engineered to target other receptors such as ERBB2 as we demonstrated; (iii) The ability of LF_N/PA to deliver PNAs of varying length and sequence indicates they can be used to target and credential the over 100 candidate CYCLOPS genes previously identified[5,6].

SF3B1 is a promising therapeutic target in cancer beyond its role as CYCLOPS gene dependency. SF3B1 mutations have been found in many cancer types including CLL[29], myelodysplastic syndrome[30,31], breast[32], uveal melanoma[33], lung[34], and others[35]. Recent work suggests that at least some of these cancers have increased sensitivity to spliceosome modulators structurally related to spliceostatin A[36], which are in clinical development[37]. Furthermore MYC-driven cancers display sensitivity to spliceosome modulators and partial SF3B1 suppression[38]. These observations raise the possibility that PNA-mediated suppression of *SF3B1* may have therapeutic benefits in *SF3B1* mutant and *MYC*-driven cancers, suggesting further research into the cancer types susceptible to SF3B1 PNA targeting.

In summary, we have demonstrated successful cytosolic delivery of the PNA constructs studied. We verified the targeting of a CYCLOPS gene dependency by anthrax-mediated PNA delivery in a panel of cancer cell lines across two different tissue types. Furthermore, we showed a re-engineered PA is able to target PNA specifically to HER2-overexpressing cancer cells. This is the first time the LF_N/PA system has been used to deliver PNAs and our results show that LF_N/PA system has the potential to become a potent and generalizable method with the potential for engineered specificity to bring the promise of PNAs to fruition and bolster oligonucleotide-based therapy for cancer therapeutics and beyond.

1.4. Experimental

1.4.1. Identification of cancer gene dependencies

Cancer gene dependencies were identified from analyses using a genome-scale shRNA viability screen from project Achilles (version 2.4.3, https://portals.broadinstitute.org/achilles/datasets/5/download). 216 cell lines were profiled for their sensitivity to suppression of 5711 genes and had mutation calls from hybrid capture sequencing of commonly mutated cancer driver genes (*KRAS*, *PIK3CA*, *BRAF*, *NRAS*, *CTNNB1* and *EGFR)*. Briefly, each cell line was classified as sensitive to suppression of each gene using an ATARiS[39] gene dependency threshold of >-0.75. For each gene, we performed a two-class comparison between cell lines with and without mutation of each of the driver genes above. Significance was calculated using fisher's exact tests and false discovery rate adjusted p values were calculated using the Benjamini and Hochberg method[40]. Gene dependencies identified from the mutation of those driver genes we then combined with list of copy number associated gene dependencies identified in a previous analysis from our group[6].

1.4.2. Estimation of the druggability of cancer gene dependencies

Classification of candidate cancer gene dependencies in traditionally druggable protein target families was facilitated by integrating information from The British Pharmacological Society (BPS) and the International Union of Basic and Clinical Pharmacology (IUPHAR) database on targets and families accessed on March 23rd 2017 (http://www.guidetopharmacology.org/DATA/targets_and_families.csv). If a cancer gene dependency was found as a member of the protein target it was considered a candidate druggable target, otherwise it was categorized at "challenging to drug"

1.4.3. Protein expression and purification

His_6-SUMO-LF_N-DTA(C186S)-LPSTGG-His_6, His_6-SUMO-LF_N-LPSTGG-His_6, SrtA*-His_6, wild-type protective antigen (PA), and PA[F427H] were expressed in *E. coli* BL21 (DE3) cells at New England Regional Center of Excellence/Biodefense and Emerging Infectious Diseases (NERCE). The cells were resuspended in 50 mM Tris-HCl, 150 mM NaCl, pH 7.5 buffer containing Roche protease inhibitor cocktail and lysed by sonication. The supernatant was subsequently loaded on the HisTrap FF crude Ni-NTA column (GE Healthcare, UK) and eluted with 500 mM imidazole. The protein was then changed into 20 mM Tris-HCl, 150 nM NaCl, pH 8.5 buffer using a HiPrep 26/10 Desalting column (GE Healthcare). WT PA and PA[F427H] were purified from the periplasm of *E. coli* BL21 (DE3) cells.

1.4.4. Synthesis of peptide nucleic acids (PNAs)

Peptide nucleic acids (PNAs) were synthesized on 0.05 mmol scale aminomethyl resin with a Rink amide linker using Fmoc synthesis. All PNAs have 5 glycines at their N-terminus for the purpose of sortase-mediated conjugation and 2 lysines after the glycines to increase their solubility. After synthesis, PNAs were cleaved from resin with TFA/m-cresol (4:1) for 2 hours and dried under N_2(g). The dried PNAs were then triturated three times with cold ether, dissolved in 0.1 % TFA in Acetonitrile (B)/0.1 % TFA in H_2O (A), and lyophilized.

1.4.5. Peptide and PNA purification

The crude peptides and PNAs were dissolved in 99:1 A:B and purified by semi-preparative RP-HPLC with Agilent Zorbax 300SB C_{18} column (9.4 x 250 mm, 5 um) at a flow rate of 4 mL/min using the gradient of 1-31% B over 100 min. Multiple rounds of purifications might apply if deemed necessary. HPLC fractions were subsequently analyzed by either MALDI or LC-MS. The final yields of the PNAs ranged from 5% to 10%.

1.4.6. Sortase-mediated ligation

The PNAs were conjugated to LF_N-DTA-LPSTGG-H_6 and LF_N-LPSTGG-H_6 by 10 uM SrtA* in sortagging buffer (10 mM CaCl2, 50 mM Tris-HCl, 150 mM NaCl, pH 7.5) for 30 minutes at room temperature. For each reaction, 100 uM of LF_N-DTA-LPSTGG-H_6 or LF_N-LPSTGG-H_6 and 500 μM of G_5-PNA were used. Simultaneously, the reaction mixture was incubated with Ni-NTA beads to get rid of all the starting materials as well as H_6-SrtA*. The supernatant was then subjected to 4 rounds of 20 mM Tris-HCl, 150 mM NaCl, pH7.5 buffer exchange to remove the excess G_5-PNA.

1.4.7. Analytical liquid chromatography-mass spectrometry (LC-MS)

All LC-MS' were performed on an Agilent 6520 Accurate-Mass quadrupole time-of-flight (Q-TOF) liquid chromatography-mass spectrometry system. We used Agilent Zorbax 300SB C_3 column (2.1 x 150 mm, 5 um) with 0.1% formic acid in H_2O at solvent A' and 0.1% formic acid in Acetonitrile as solvent B'. A gradient of 5-65% B' over 15 min or 1-61% over 12 min was used depending on the solubility of the construct.

1.4.8. Protein synthesis inhibition assay

The CHO-K1 cells were maintained in F-12K media supplemented with 10% (v/v) fetal bovine serum at 37 °C and 5% CO_2. The cells were plated at 30K cells per well in a 96-well plate one day before the assay. On the day of experiment, the cells were treated with indicated concentrations of LDPs and 20 nM of PA for 30 minutes at 37 °C. They were then washed with PBS and incubated with medium supplemented with 1 μCi/mL ^{3}H-leucine (Perkin Elmer, MA) for 1 hour at 37 °C. The incorporation of ^{3}H-Leu into cellular proteins was measured to determine the inhibition of protein synthesis by scintillation counter. All counts were normalized to cells treated with only PA and all experiments were done in triplicate.

1.4.9. Cytosolic protein extraction and western blot

CHO-K1 or HEK-293T cells were treated with 250 nM of LDPs or LPs and 40 nM of PA or PA[F427H] for 12 hours at 37 °C. The cells were then trypsinized and washed with PBS. For cytosolic extraction, ~1 million cells were resuspended in 50 - 100 ug/mL digitonin in 75 mM NaCl, 1 mM NaH_2PO_4, 8 mM Na_2HPO_4, 250 mM sucrose supplemented with Roche protease inhibitor cocktail on ice for 10 minutes and spun down at 13,000 rpm for 5 minutes at 4 °C. For total cell lysis, lysis buffer (pH 7.5) with 25 mM Tris, 150 mM NaCl, 1% v/v NP-40 supplemented with Roche protease inhibitor cocktail was used. Cells were lysed on ice for 30 minutes and spun down for 10 minutes at 4 °C. Subsequently, the supernatants from both methods were run on an SDS-PAGE gel and then transferred onto a nitrocellulose membrane soaked in transfer buffer (48 mM Tris, 39 mM glycine, 0.0375% SDS, 20% methanol) using a TE 70 Semi-Dry Transfer Unit (GE). After the transfer is finished, the membrane was blocked with LI-COR blocking buffer at RT for 2 hours before it was blotted against anti-LF or anti-DTA or anti-ERK1/2 and anti-Rab5 in TBST overnight at 4 °C. The membrane was then washed and incubated with appropriate secondary antibodies for 1 h at RT and imaged by LI-COR Odyssey infrared imaging system. The signal intensity of each band was quantified by drawing a rectangle around it and measuring the average signal within the rectangle using the analysis tool in Image Studio (LI-COR Biosciences).

1.4.10. Generation of isogenic SF3B1^{loss} cells by CRISPR Cas9

The diploid breast cancer cell line, Cal51, was used to create daughter cell lines that differed only in *SF3B1* copy number. We designed gRNAs that target intergenic loci upstream and downstream of the *SF3B1* locus to delete the entire gene. The 5' gRNA targeted one allele of a heterozygous SNP (rs3849362) to bias towards heterozygous *SF3B1* deletion because homozygous SF3B1 deletion is lethal. Oligonucleotides with BbsI overhangs were annealed and cloned into

pX458 that also encodes Cas9 and GFP (Addgene plasmid #48138). For the 5' guide targeting rs3849362, 5' CACCGCGCATTATAGATTATGGCCC (forward) and 5' AAACGGGCCATAATCTATAATGCGC (reverse). For the 3' targeting guide: 5'CACCGCGGAGTTTCATCCGTTACAC (forward), AAACGTGTAACGGATGAAACTCCGC (reverse). pX458 plasmids with either the 5' or 3' targeting guides were co-transfected into cells and GFP+ single cells were plated at 1 cell per well by FACS.

1.4.11. PNA and siRNA delivery for cell viability assays using Cell Titer Glo

Estimation of cell viability was performed using ATP luminescence. Briefly, 1000 cells per well were plated in 96 well plates, allowed to recover for 4-16 hours, and treated with the indicated LF_N/PA system components. Cell viability was assessed 7 days after PNA treatment. For siRNA transfection, the same plating conditions were used, except a pool of 4 SF3B1 siRNAs were transfected in 6 well plated 24 hours prior to plating using Oligofectamine. The sequences for the siRNAs were first reported in (Massiello et al., 2006)[41] and are as follows:

siSF3B1 #1 (sense) 5'-GGA AUU AGA UGC UAU GUU CUU, (antisense) 5'-GAA CAU AGC AUC UAA UUC CUU.

siSF3B1 #2 (sense) 5'-GCA AAC GAG UCA AAC CAU AUU, (antisense) 5'-UAU GGU UUG ACU CGU UUG CUU

siSF3B1 #3 (sense) 5'-GAA CCG CUA UUG AUU GAU GUU, (antisense) 5'-CAU CAA UCA AUA GCG GUU CUU

siSF3B1 #4 (sense) 5'-GUA GAA UGU UGC AAU AUU GUU, (antisense) 5'-CAA UAU UGC AAC AUU CUA CUU

siLacZ control (sense) 5'-UGU UCG CAU UAU CCG AAC CUU, (antisense) 5'-GGU UCG GAU AAU GCG AAC AUU

1.4.12. Determination of cellular sensitivity to chemical modulation of the spliceosome

Cells were plated at 10,000 cells per well in 96 well plates. 24 hours after seeding, cells were treated with the indicated concentrations of splicing modulator. Relative cell viability was expressed as the percent change in Cell Titer-Glo luminescence normalized to vehicle (DMSO) treated cells. Spliceostatin A was a gift of Dr. Minoru Yoshida[24].

1.4.13. Splicing luciferase reporter assays

Polyclonal, stable CRISPR^{neutral} and CRISPR$^{copy\text{-}loss}$ cells were generated by transfection and hygromycin selection. Cells expressed either CMV-LUC2CP/intron/ARE (Luc-I in this manuscript, Addgene plasmid # 62858) or CMV-LUC2CP/ARE (Luc-ORF in this manuscript, Addgene plasmid # 62857) which were gifts from Gideon Dreyfuss[42]. To calculate the fold change in splicing activity, the ratio of Luc-I:Luc-ORF was calculated after normalizing to sense SF3B1 PNA treated cells.

1.5. Significance

Targeted therapeutics have provided notable successes for the treatment of cancer, however there remains a scarcity of effective therapies. In order to augment the number of potential therapeutic strategies, a comprehensive catalog of candidate cancer therapeutic targets is needed. Furthermore, versatile and effective delivery methods for these therapeutics presents a major barrier to success. To address these challenges, we analyzed genome-wide shRNA viability screens and identified 486 genes that are required for cancer cell survival that are associated with a cancer-specific somatic genetic alteration. Strikingly, we estimate that only 13% of these cancer gene dependencies are members of "druggable" protein classes. However, antisense gene suppression, such as through peptide-nucleic acid conjugates (PNAs), represents a generalizable approach to target cancer gene dependencies. Here we developed a non-viral, lipid-free delivery system using engineered components of anthrax toxin. We demonstrate efficient delivery of anti-sense PNAs across cell membranes and into the cytosol and target gene suppression. In proof of concept studies, we demonstrate that anthrax delivered PNAs can target cancer specific gene dependencies on the essential spliceosome component *SF3B1*. This dependency on SF3B1 is one of over 120 examples of CYCLOPS (Copy number alterations Yielding Cancer Liabilities Owing to Partial LosS) dependencies, indicating a vast set of dependencies that may be exploited by our PNA delivery system.

1.6. Acknowledgements

This work was funded by MIT start-up funds, MIT Reed Fund, Damon Runyon Cancer Research Foundation Innovation Award, and National Science Foundation (NSF) CAREER Award (CHE-1351807) for B.L.P, NIH grants R01 CA188228 (R.B.), F32 CA180653 (B.R.P.), the Sontag Foundation (R.B. & B.L.P.), the Melanoma Research Alliance (R.B.), The Gray Matters Brain Cancer Foundation (R.B.), and the Friends of DFCI (B.R.P.). We thank R. J. Collier (Harvard) for his continued contribution of laboratory equipment, the NERCE facility (grant: U54 AI057159) for expression of toxin proteins. We also thank Prof. Douglas Lauffenburger, and Jingjing Ling for their comments.

1.7. Appendix

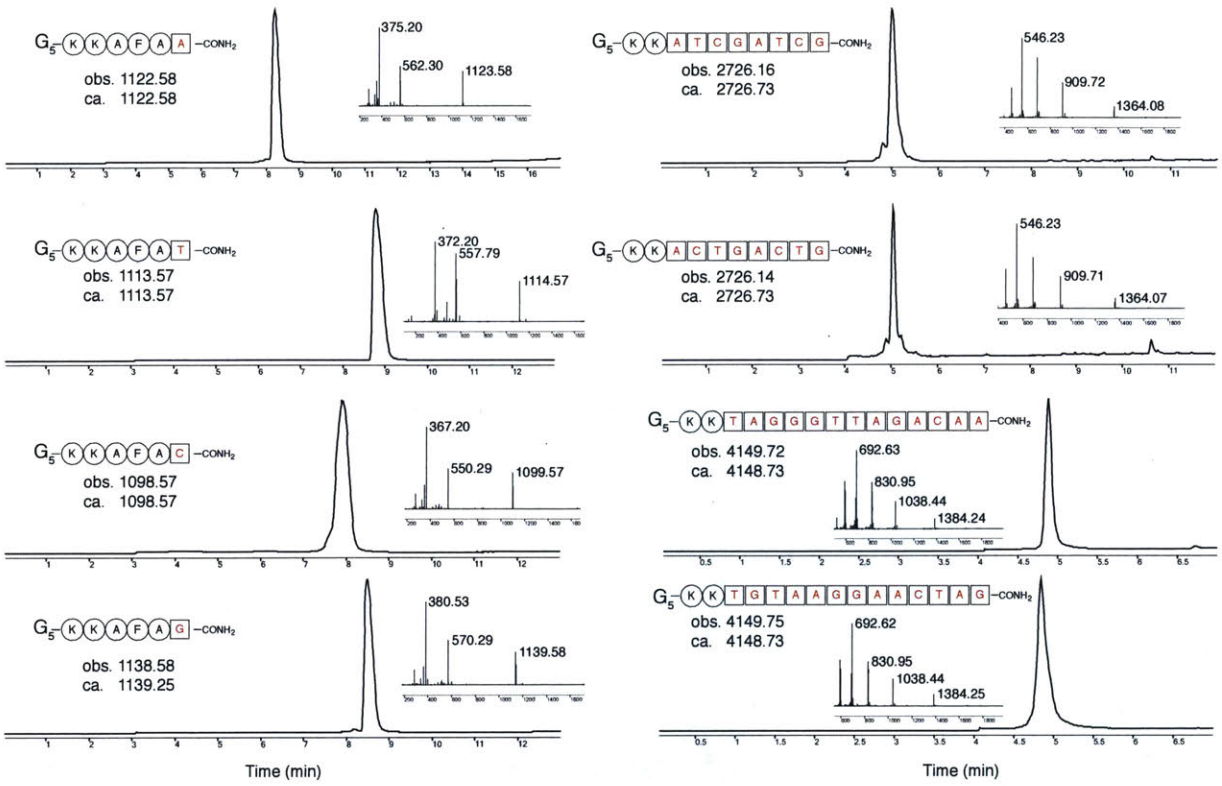

Figure 1.7.1. LC-MS traces of the PNA constructs used in this study. All the PNAs have N-terminal G5 tags. All LC-MS' were performed on an Agilent 6520 Accurate-Mass quadrupole time-of-flight (Q-TOF) liquid chromatography-mass spectrometry system. We used Agilent Zorbax 300SB C_3 column (2.1 x 150 mm, 5 um) with 0.1% formic acid in H_2O at solvent A' and 0.1% formic acid in Acetonitrile as solvent B'. A gradient of 5-65% B' over 15 min or 1-61% over 12 min was used depending on the solubility of the construct.

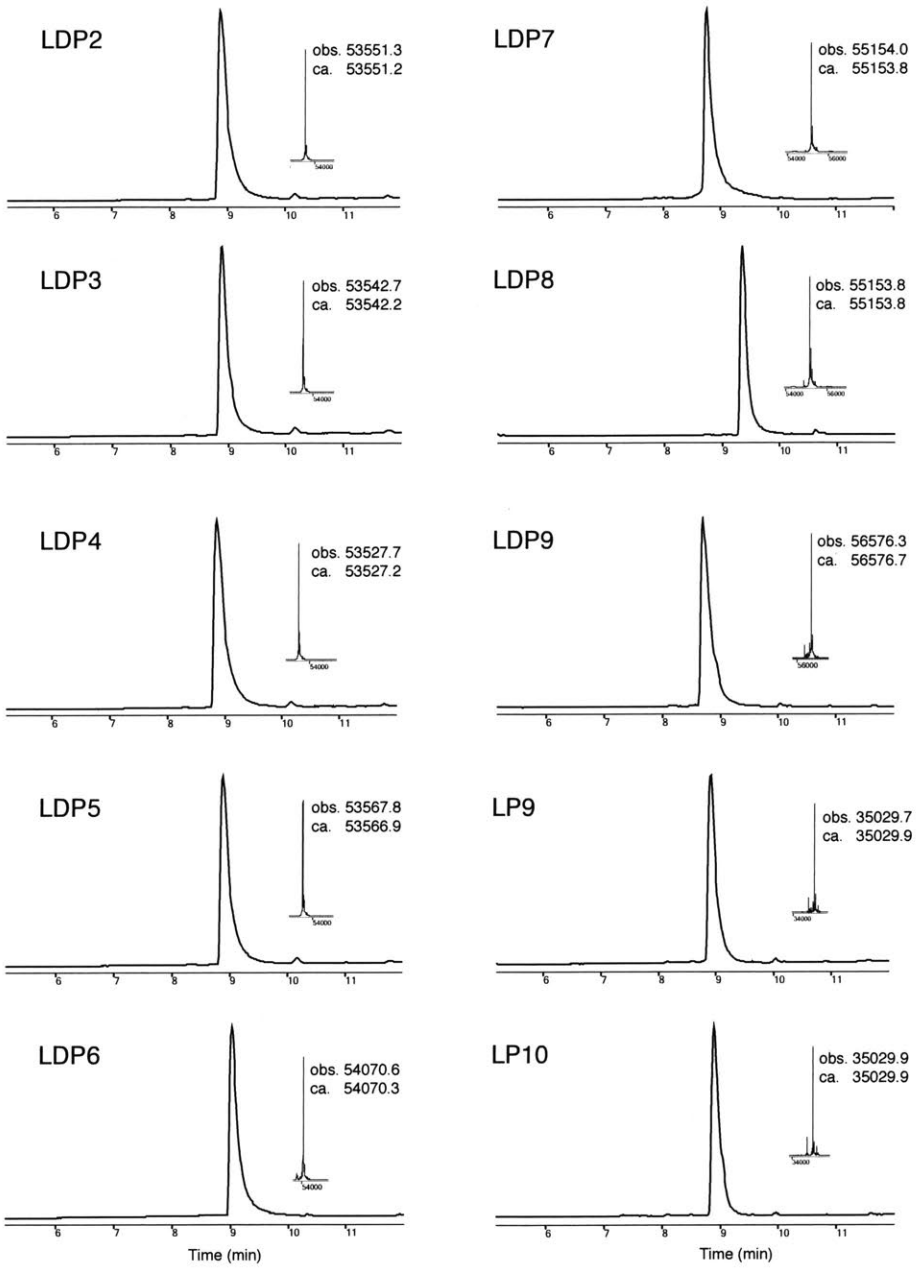

Figure 1.7.2. LC-MS traces of LDPs and LPs used in this study. The LC-MS' were performed as described in Supplementary Figure 1.

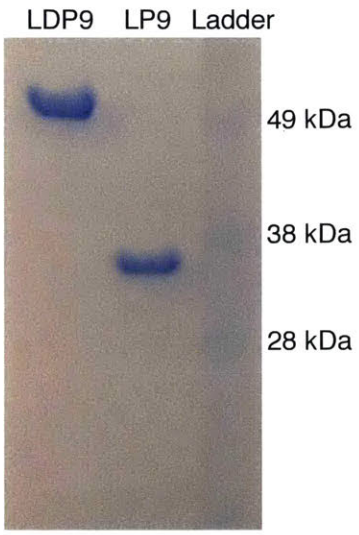

Figure 1.7.3. Coomassie stained SDS-PAGE gel of LF_N-DTA-PNA9 (LDP9) and LF_N-PNA9 (LP9).

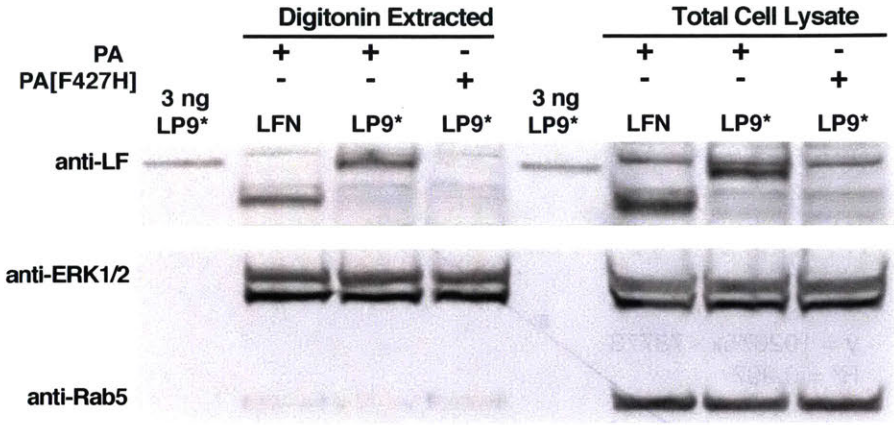

Figure 1.7.4. Western blot showing the translocation of LP9*, a scrambled version of LP9. HEK-293T cells were treated with 250 nM of LP9 and 40 nM of PA or PA[F427H] for 12 hours. The cells were subsequently subjected to digitonin extraction or total lysis (~1 million cells per lane). Purified LP was also run on the gel for the purpose of quantification.

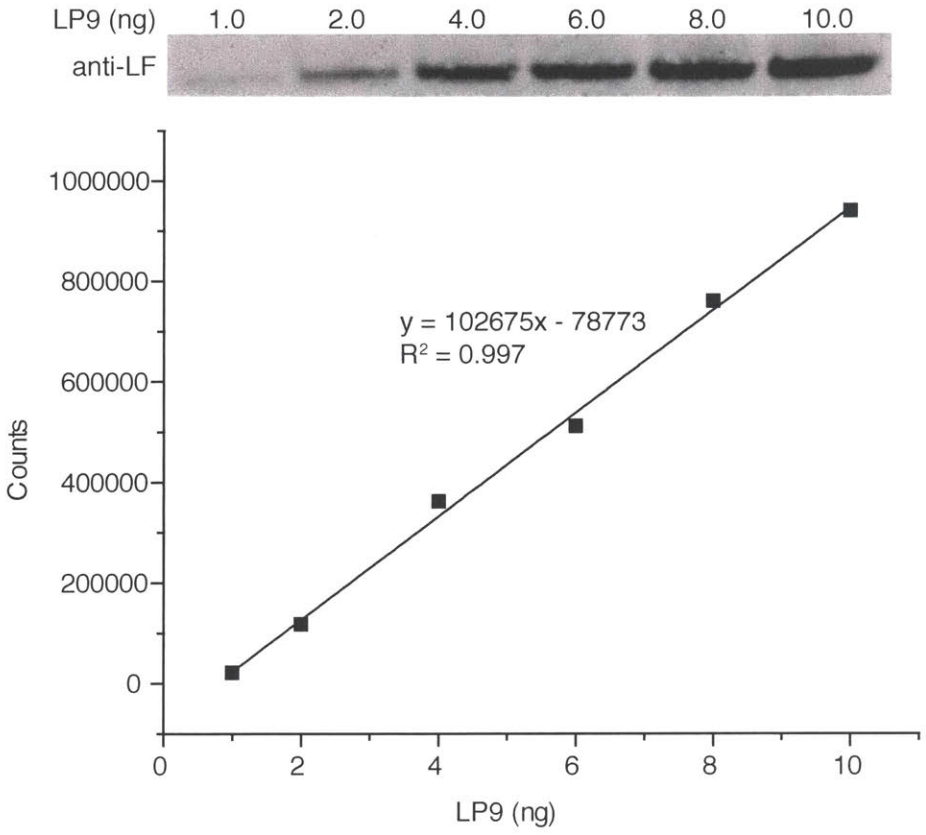

Figure 1.7.5. Linear relationship between band intensity and the amount of protein loaded. 1.0-10.0 ng of pure LP9 was loaded on the gel and the signal intensity of each band was analyzed by Image Studio (LI-COR Biosciences).

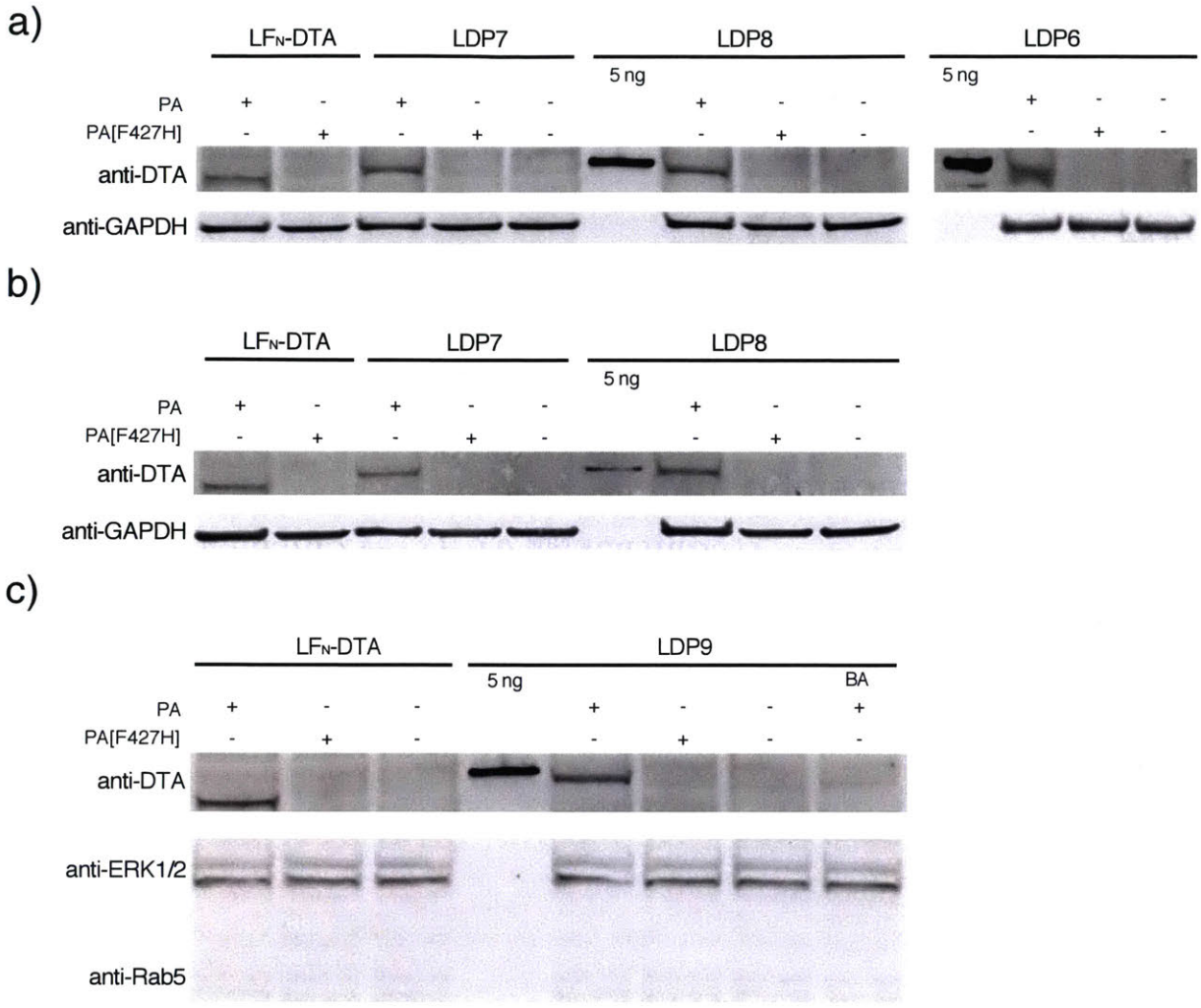

Figure 1.7.6. Western blot showing the translocation of LDP6-9. Cells were treated with 100 nM of LDP and 20 nM of PA or PA[F427H] for 12 hours. The cells were subsequently subjected to digitonin extraction. Purified LDP was also run on the gel as a loading control. a) Translocation of LDP6-8 in HEK-293T cells. b) Translocation of LDP7 and LDP8 in CHO-K1 cells. c) Translocation of LDP9 in HEK-293T cells. BA stands for Bafilomycin A1, which inhibits the acidification of the endosome. It was used as a negative control to validate the translocation pathway.

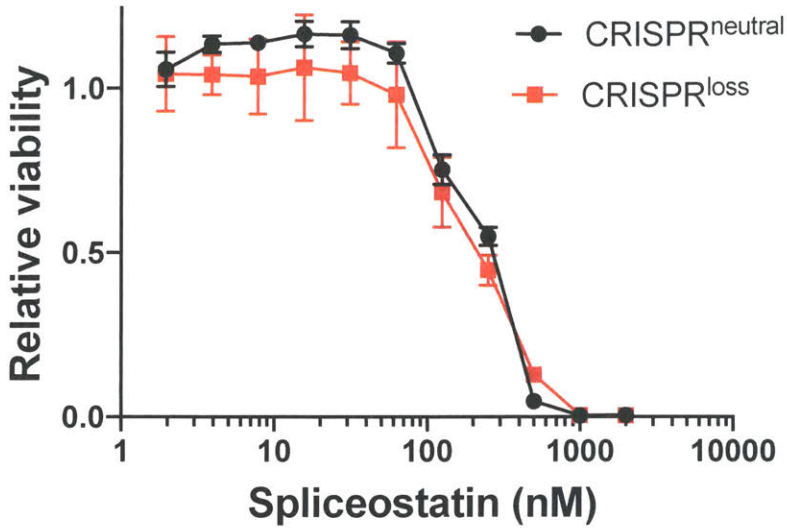

Figure 1.7.7. Dose response curve for CRISPR^{neutral} (black) and CRISPR$^{copy\text{-}loss}$ cells (red) upon treatment with the splicing modulator, Spliceostatin A. Relative viability was quantified by Cell titer glo and normalized to DMSO vehicle control. Representative experiment performed in triplicate.

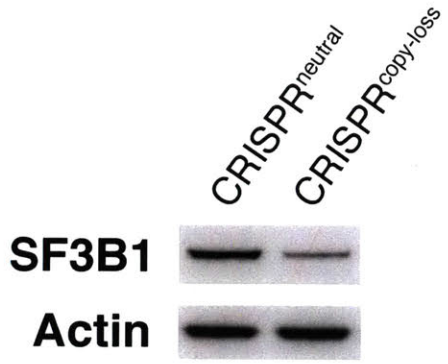

Figure 1.7.8. SF3B1 immunoblot in CRISPR^{neutral} and CRISPR$^{copy\text{-}loss}$ cells. Isogeneic Cal51 cells treated with transient transfection of Cas9 sgRNAs upstream and downstream of the SF3B1 locus and screened for single copy SF3B1 deletion. Cells were then lysed and evaluated for the effects of partial SF3B1 copy loss on protein levels.

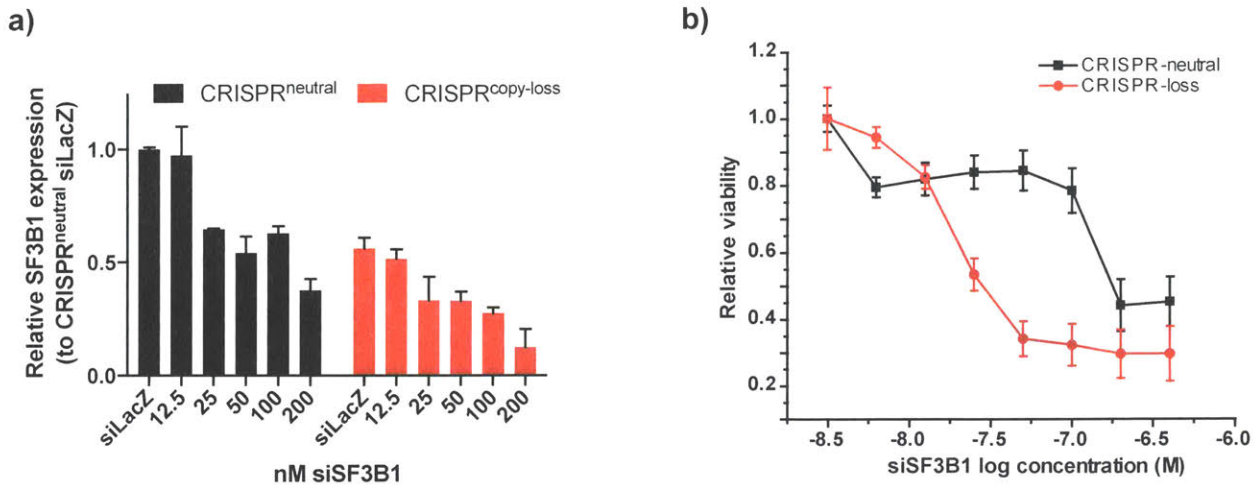

Figure 1.7.9. Sensitivity of CRISPR^{neutral} and CRISPR^{loss} Cal51 cells in the presence of different concentrations of siRNA that suppresses SF3B1. a) SF3B1 mRNA expression level upon siRNA transfection by quantitative PCR with the indicated concentrations of siRNA. Relative SF3B1 expression was normalized to siLacZ transfected cells in CRISPR^{neutral} cells (black) and CRISPR^{loss} cells (red). b) Dose response curve for relative cell viability of CRISPR^{neutral} (black) and CRISPR$^{copy\text{-}loss}$ (red) cells upon siRNA-mediated SF3B1 suppression. Cell viability calculated as the fold change in Cell Titer-Glo luminescence relative to siLacZ control. Representative experiment performed in triplicate. For all panels, Error bars represent +- SD.

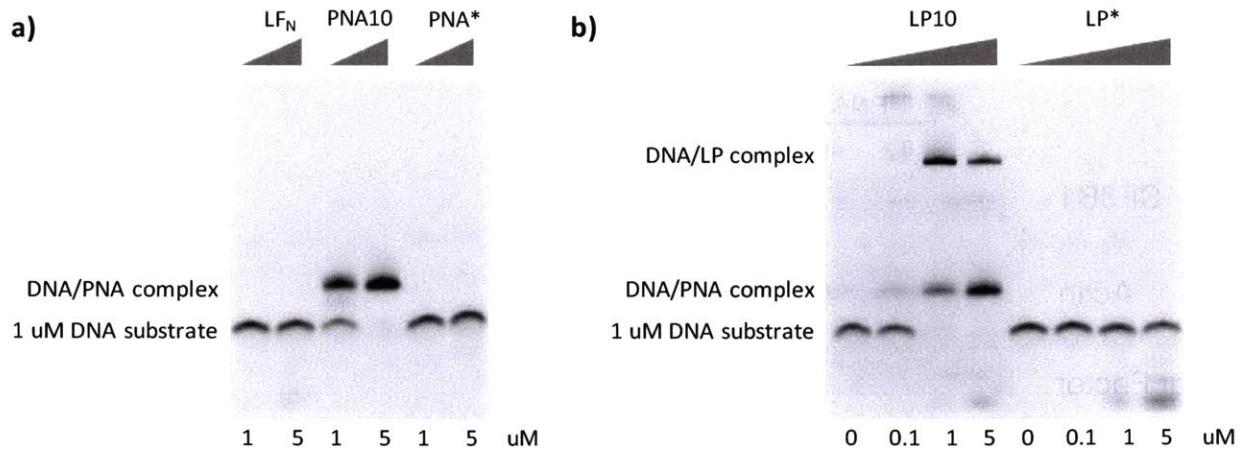

Figure 1.7.10. The *in vitro* binding of PNA10 to its DNA substrate measured by electrophoretic mobility shift assay. 1 uM of a DNA-Cy3 probe that hybridizes with the anti-sense SF3B1 PNA was incubated with indicated concentration of PNA10 (anti-sense SF3B1 PNA) or LP10 (anti-sense SF3B1 PNA conjugated to LF_N) for 2 hours at room temperature. The mixture was then subjected to gel electrophoresis and imaged with a Typhoon FLA 7000 scanner (GE). a) The binding to the DNA substrate by LF_N only, PNA10, or a negative control PNA* measured by the band shift. b) The binding to the DNA substrate by LP10 or the negative control LF_N-PNA* (LP*). The DNA/PNA complex band observed resulted from some residual PNA10 after the conjugation of LF_N and PNA10.

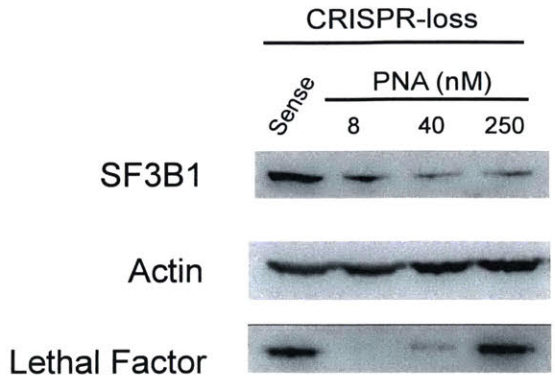

Figure 1.7.11. Immunoblot after treatment with LF_N sense or anti-sense SF3B1 PNAs in $CRISPR^{loss}$ cells. Western probed with SF3B1 and anti-LF antibodies with 8, 40, or 250 nM LFN-conjugated antisense SF3B1 PNA in the presence of 50 nM PA. 250 nM LFN-conjugated SF3B1 sense PNA was used as a control. Lysates were collected four days after LFN/PA PNA delivery. Actin was used as a loading control.

1.7.1. Supplementary notes:

Definition of terms

Gene dependency – A gene, that when suppressed by RNAi, results in decreased cellular fitness.

Somatic genetic alteration – A genetic change that deviates from the normal germline DNA sequence, including mutations or copy number changes. These alterations accumatuate in cells and contribute to cellular transformation.

Cancer gene dependency – A gene dependency that is associated with a cancer specific somatic genetic alteration. The association with a genetic alteration suggests the gene is only required for cancer cell, but not normal cell survivial because the normal cells lack the associated genetic event

Driver gene – A gene, that when altered by a somatic genetic alteration (i.e. mutation or copy number change), contributes to cellular transformation and cancer.

1.7.2. Inhibition of telomerase by LF_N/PA delivered anti-sense PNA

In addition to targeting SF3B1, we investigated the delivery of anti-sense PNA to inhibit telomerase. Telomerase catalyzes the addition of telomeric repeats onto the 3' end of a specific sequence of DNA using its RNA component as the template[43,44] and it has been shown to play a key role in cancer proliferation[45]. Telomeric Repeat Amplification Protocol (TRAP) is a PCR-based assay that detects the activity of telomerase by amplifying and measuring telomeric repeats, which is the product of telomerase[45] (Figure 1.7.12). The TRAP assays were performed using the TRAPeze kit (Millipore).

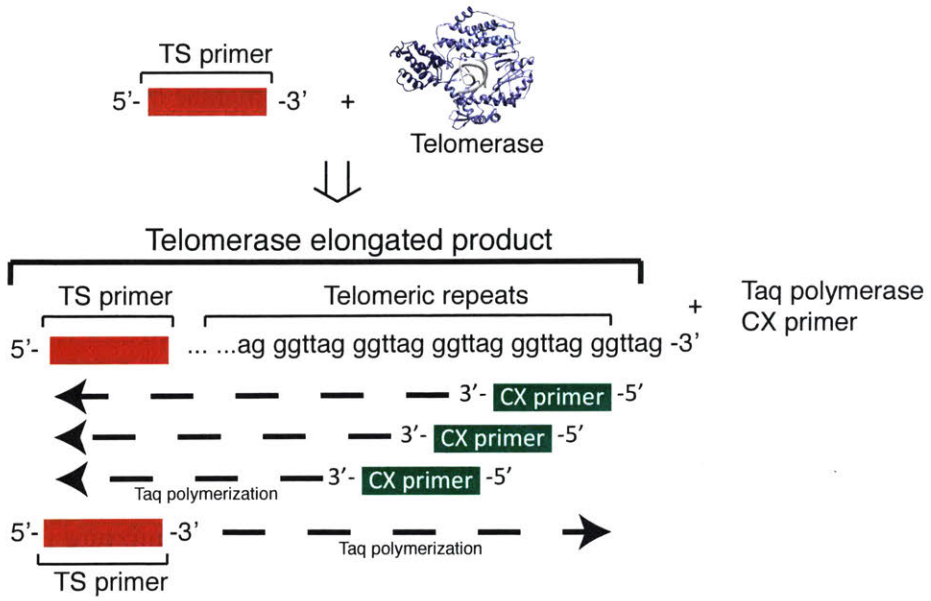

Figure 1.7.12. A schematic diagram showing how Telomeric Repeat Amplification Protocol (TRAP) assays work.

LP9 is one of the model PNAs we used to characterize PNA delivery by LF_N/PA and it is also a reported inhibitor of telomerase[46]. We first tested the inhibition of telomerase by LP9 using cell lysate. Both HEK-293T (human embryonic kidney) and A-431 epidermoid carcinoma cells were used for the assays and both have been shown to be dependent on telomerase function. One million cells were lysed using CHAPS buffer and a subset (600 cells) were analyzed for (4 μl of 150 cell equivalents/μl) telomerase activity. The cellular extracts were preincubated at 25 °C for 30 minutes with LP9 or LP10 at the indicated concentrations. Following preincubation, 2 μl of the cellular extract containing LP9 or LP10 was added to a PCR tube with a reaction mixture (Tris buffer, TS primer, CX primer, IS primer, dNTP, and *Taq* polymerase) and incubated for 30 min at 30 °C to allow the elongation of the TS primer. Then the amplification step was performed through 30 cycles of PCR (30 seconds at 95 °C, 30 seconds at 50 °C, 45 seconds at 72 °C) and the resulted products were loaded onto a 10% nondenaturing polyacrylamide gel. Gels were stained with ethidium bromide, imaged by UV and processed with ImageStudio. For each condition, the total fluorescence intensity was summed and normalized to the internal standard for the quantification. LP9 was shown to effectively inhibit telomerase activity from both HEK-293T and A-431 lysate at concentrations above 10 nM (Figure 1.7.13). Unlike LP9, LP10, a scrambled version of LP9, showed no inhibition of telomerase activity even at 1 μM, indicating that complementary binding between the PNA and the template region of human telomerase RNA (hTR) is necessary for this action. The absence of any product bands in the no-lysate control confirmed the bands were due to the telomerase activity instead of any nonspecific *Taq* activity.

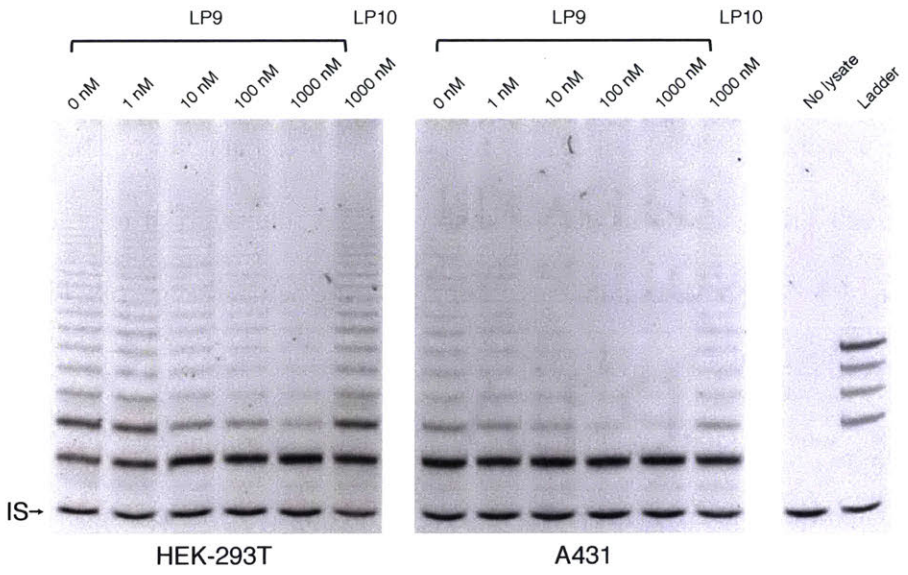

Figure 1.7.13. TRAP assay showing inhibition of telomerase activity by LP9 in both HEK-293T and A431 cells. LP10 was used to confirm the complementation is required for the inhibition. The 36-bp internal standard (IS) strand is noted. The lane denoted "No lysate" represents a negative control to which no cell lysate was added.

After confirming the inhibition of telomerase by LP9 with cell extracts, we proceeded to inhibit telomerase inside cancer cells by translocating LP9 with PA. A-431 cells were incubated with 250 nM of LP9 in the presence of 40 nM of PA for 60 hours prior to harvesting. The cells were subsequently subjected to TRAP assays as described above. 35% inhibition of telomerase activity was observed for the LP9 + PA treated sample, as compared to the negative controls (PNA9 only, PNA9 +PA, LP10 +PA, LF_N + PA, LP9 + PA[F427H], and LP9 only) (Figures 1.7.14-15). A previous report found that 8 lysine residues covalently attached to the C-terminus of a PNA designed to cause exon skipping was delivered into cells by PA.[23] However, we failed to observe inhibition of telomerase activity with PNA9-K_8 event in the presence of PA. Moreover, a cell-penetrating TAT peptide was also conjugated to PNA9 and it too failed to inhibit telomerase activity in the cell. The fact that no inhibition was observed when either PA or LF_N was absent indicates their indispensable roles during translocation. Finally, we used LP10 with a mismatched PNA sequence as an additional negative control to confirm that inhibition occurred through the hybridization between LP9 and the template region of hTR.

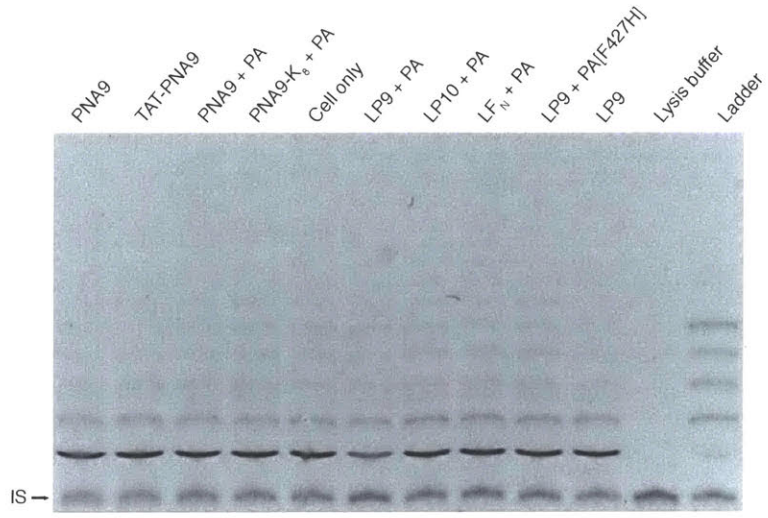

Figure 1.7.14. Telomeric repeat amplification protocol (TRAP) assays. A431 cells were treated with 250 nM of indicated compounds for all experiments for 60 hours. After treatment, cells were trypsinized, washed, and assayed for telomerase activity. All experiments were carried out three times and for each trial TRAP products were run on a 10% non-denaturing acrylamide gel (a representative gel from one of the trials is shown).

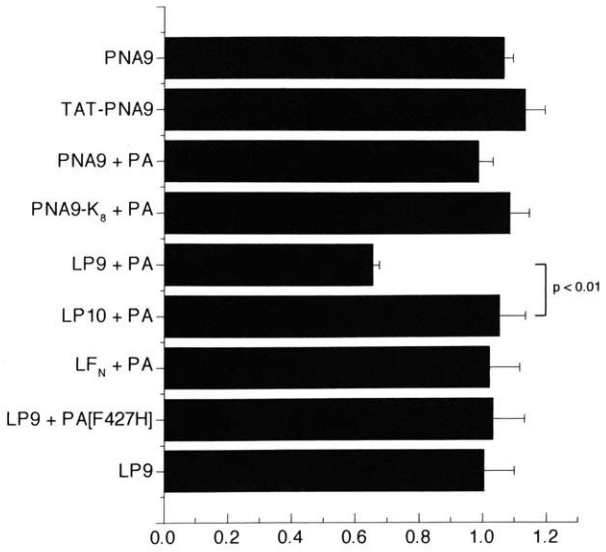

Figure 1.7.15. Telomeric repeat amplification protocol (TRAP) assays. A431 cells were treated with 250 nM of indicated compounds for all experiments in the presence or absence of 40 nM PA for 60 hours. After treatment, cells were trypsinized, washed, and assayed for telomerase activity. All experiments were carried out three times and for each trial TRAP products were run on a 10% non-denaturing acrylamide gel (Figure S8). Telomerase activity was quantified by UV fluorescence from ethidium bromide and processed with ImageStudio. For each condition, the total fluorescence intensity was summed and normalized to the internal standard with in the same well for the quantification. All the data points from each trial were normalized to cell only. P values were calculated using the Student's t-Test.

1.8. References

1. Gambari, R. *Expert Opin Ther Pat* **2014**, *24*, 267.
2. Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991**, *254*, 1497.
3. Demidov, V. V.; Potaman, V. N.; Frankkamenetskii, M. D.; Egholm, M.; Buchard, O.; Sonnichsen, S. H.; Nielsen, P. E. *Biochem Pharmacol* **1994**, *48*, 1310.
4. Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Freier, S. M.; Driver, D. A.; Berg, R. H.; Kim, S. K.; Norden, B.; Nielsen, P. E. *Nature* **1993**, *365*, 566.
5. Nijhawan, D.; Zack, T. I.; Ren, Y.; Strickland, M. R.; Lamothe, R.; Schumacher, S. E.; Tsherniak, A.; Besche, H. C.; Rosenbluh, J.; Shehata, S.; Cowley, G. S.; Weir, B. A.; Goldberg, A. L.; Mesirov, J. P.; Root, D. E.; Bhatia, S. N.; Beroukhim, R.; Hahn, W. C. *Cell* **2012**, *150*, 842.
6. Paolella, B. R.; Gibson, W. J.; Urbanski, L. M.; Alberta, J. A.; Zack, T. I.; Bandopadhayay, P.; Nichols, C. A.; Agarwalla, P. K.; Brown, M. S.; Lamothe, R.; Yu, Y.; Choi, P. S.; Obeng, E. A.; Heckl, D.; Wei, G.; Wang, B.; Tsherniak, A.; Vazquez, F.; Weir, B. A.; Root, D. E.; Cowley, G. S.; Buhrlage, S. J.; Stiles, C. D.; Ebert, B. L.; Hahn, W. C.; Reed, R.; Beroukhim, R. *Elife* **2017**, *6*.
7. Koppelhus, U.; Nielsen, P. E. *Advanced Drug Delivery Reviews* **2003**, *55*, 267.
8. *The Comprehensive Sourcebook of Bacterial Protein Toxins*; 3rd ed.; Academic Press, 2006.
9. Klimpel, K. R.; Molloy, S. S.; Thomas, G.; Leppla, S. H. *P Natl Acad Sci USA* **1992**, *89*, 10277.
10. Milne, J. C.; Furlong, D.; Hanna, P. C.; Wall, J. S.; Collier, R. J. *J Biol Chem* **1994**, *269*, 20607.
11. Kintzer, A. F.; Thoren, K. L.; Sterling, H. J.; Dong, K. C.; Feld, G. K.; Tang, I. I.; Zhang, T. T.; Williams, E. R.; Berger, J. M.; Krantz, B. A. *J Mol Biol* **2009**, *392*, 614.
12. Krantz, B. A.; Trivedi, A. D.; Cunningham, K.; Christensen, K. A.; Collier, R. J. *J Mol Biol* **2004**, *344*, 739.
13. Rabideau, A. E.; Liao, X.; Akcay, G.; Pentelute, B. L. *Scientific reports* **2015**, *5*, 11944.
14. Liao, X.; Rabideau, A. E.; Pentelute, B. L. *Chembiochem : a European journal of chemical biology* **2014**, *15*, 2458.
15. Ballard, J. D.; Collier, R. J.; Starnbach, M. N. *P Natl Acad Sci USA* **1996**, *93*, 12531.
16. Hobson, J. P.; Liu, S. H.; Rono, B.; Leppla, S. H.; Bugge, T. H. *Nat Methods* **2006**, *3*, 259.
17. Cowley, G. S.; Weir, B. A.; Vazquez, F.; Tamayo, P.; Scott, J. A.; Rusin, S.; East-Seletsky, A.; Ali, L. D.; Gerath, W. F.; Pantel, S. E.; Lizotte, P. H.; Jiang, G.; Hsiao, J.; Tsherniak, A.; Dwinell, E.; Aoyama, S.; Okamoto, M.; Harrington, W.; Gelfand, E.; Green, T. M.; Tomko, M. J.; Gopal, S.; Wong, T. C.; Li, H.; Howell, S.; Stransky, N.; Liefeld, T.; Jang, D.; Bistline, J.; Hill Meyers, B.; Armstrong, S. A.; Anderson, K. C.; Stegmaier, K.; Reich, M.; Pellman, D.; Boehm, J. S.; Mesirov, J. P.; Golub, T. R.; Root, D. E.; Hahn, W. C. *Scientific data* **2014**, *1*, 140035.
18. Chen, I.; Dorr, B. M.; Liu, D. R. *P Natl Acad Sci USA* **2011**, *108*, 11399.
19. Collier, R. J.; Kandel, J. *J Biol Chem* **1971**, *246*, 1496.
20. Wilson, B. A.; Collier, R. J. *Curr Top Microbiol* **1992**, *175*, 27.
21. Krantz, B. A.; Melnyk, R. A.; Zhang, S.; Juris, S. J.; Lacy, D. B.; Wu, Z. Y.; Finkelstein, A.; Collier, R. J. *Science* **2005**, *309*, 777.
22. Sun, J.; Lang, A. E.; Aktories, K.; Collier, R. J. *P Natl Acad Sci USA* **2008**, *105*, 4346.
23. Adam, S. A.; Marr, R. S.; Gerace, L. *J Cell Biol* **1990**, *111*, 807.
24. Kaida, D.; Motoyoshi, H.; Tashiro, E.; Nojima, T.; Hagiwara, M.; Ishigami, K.; Watanabe, H.; Kitahara, T.; Yoshida, T.; Nakajima, H.; Tani, T.; Horinouchi, S.; Yoshida, M. *Nature chemical biology* **2007**, *3*, 576.

25. Doyle, D. F.; Braasch, D. A.; Simmons, C. G.; Janowski, B. A.; Corey, D. R. *Biochemistry* **2001**, *40*, 53.
26. Bendifallah, N.; Rasmussen, F. W.; Zachar, V.; Ebbesen, P.; Nielsen, P. E.; Koppelhus, U. *Bioconjug Chem* **2006**, *17*, 750.
27. Hart, T.; Chandrashekhar, M.; Aregger, M.; Steinhart, Z.; Brown, K. R.; MacLeod, G.; Mis, M.; Zimmermann, M.; Fradet-Turcotte, A.; Sun, S.; Mero, P.; Dirks, P.; Sidhu, S.; Roth, F. P.; Rissland, O. S.; Durocher, D.; Angers, S.; Moffat, J. *Cell* **2015**, *163*, 1515.
28. Blaustein, R. O.; Finkelstein, A. *J Gen Physiol* **1990**, *96*, 943.
29. Wang, L.; Lawrence, M. S.; Wan, Y.; Stojanov, P.; Sougnez, C.; Stevenson, K.; Werner, L.; Sivachenko, A.; DeLuca, D. S.; Zhang, L.; Zhang, W.; Vartanov, A. R.; Fernandes, S. M.; Goldstein, N. R.; Folco, E. G.; Cibulskis, K.; Tesar, B.; Sievers, Q. L.; Shefler, E.; Gabriel, S.; Hacohen, N.; Reed, R.; Meyerson, M.; Golub, T. R.; Lander, E. S.; Neuberg, D.; Brown, J. R.; Getz, G.; Wu, C. J. *The New England journal of medicine* **2011**, *365*, 2497.
30. Papaemmanuil, E.; Cazzola, M.; Boultwood, J.; Malcovati, L.; Vyas, P.; Bowen, D.; Pellagatti, A.; Wainscoat, J. S.; Hellstrom-Lindberg, E.; Gambacorti-Passerini, C.; Godfrey, A. L.; Rapado, I.; Cvejic, A.; Rance, R.; McGee, C.; Ellis, P.; Mudie, L. J.; Stephens, P. J.; McLaren, S.; Massie, C. E.; Tarpey, P. S.; Varela, I.; Nik-Zainal, S.; Davies, H. R.; Shlien, A.; Jones, D.; Raine, K.; Hinton, J.; Butler, A. P.; Teague, J. W.; Baxter, E. J.; Score, J.; Galli, A.; Della Porta, M. G.; Travaglino, E.; Groves, M.; Tauro, S.; Munshi, N. C.; Anderson, K. C.; El-Naggar, A.; Fischer, A.; Mustonen, V.; Warren, A. J.; Cross, N. C.; Green, A. R.; Futreal, P. A.; Stratton, M. R.; Campbell, P. J. *The New England journal of medicine* **2011**, *365*, 1384.
31. Yoshida, K.; Sanada, M.; Shiraishi, Y.; Nowak, D.; Nagata, Y.; Yamamoto, R.; Sato, Y.; Sato-Otsubo, A.; Kon, A.; Nagasaki, M.; Chalkidis, G.; Suzuki, Y.; Shiosaka, M.; Kawahata, R.; Yamaguchi, T.; Otsu, M.; Obara, N.; Sakata-Yanagimoto, M.; Ishiyama, K.; Mori, H.; Nolte, F.; Hofmann, W. K.; Miyawaki, S.; Sugano, S.; Haferlach, C.; Koeffler, H. P.; Shih, L. Y.; Haferlach, T.; Chiba, S.; Nakauchi, H.; Miyano, S.; Ogawa, S. *Nature* **2011**, *478*, 64.
32. Ellis, M. J.; Ding, L.; Shen, D.; Luo, J.; Suman, V. J.; Wallis, J. W.; Van Tine, B. A.; Hoog, J.; Goiffon, R. J.; Goldstein, T. C.; Ng, S.; Lin, L.; Crowder, R.; Snider, J.; Ballman, K.; Weber, J.; Chen, K.; Koboldt, D. C.; Kandoth, C.; Schierding, W. S.; McMichael, J. F.; Miller, C. A.; Lu, C.; Harris, C. C.; McLellan, M. D.; Wendl, M. C.; DeSchryver, K.; Allred, D. C.; Esserman, L.; Unzeitig, G.; Margenthaler, J.; Babiera, G. V.; Marcom, P. K.; Guenther, J. M.; Leitch, M.; Hunt, K.; Olson, J.; Tao, Y.; Maher, C. A.; Fulton, L. L.; Fulton, R. S.; Harrison, M.; Oberkfell, B.; Du, F.; Demeter, R.; Vickery, T. L.; Elhammali, A.; Piwnica-Worms, H.; McDonald, S.; Watson, M.; Dooling, D. J.; Ota, D.; Chang, L. W.; Bose, R.; Ley, T. J.; Piwnica-Worms, D.; Stuart, J. M.; Wilson, R. K.; Mardis, E. R. *Nature* **2012**, *486*, 353.
33. Harbour, J. W.; Roberson, E. D.; Anbunathan, H.; Onken, M. D.; Worley, L. A.; Bowcock, A. M. *Nat Genet* **2013**, *45*, 133.
34. Imielinski, M.; Berger, A. H.; Hammerman, P. S.; Hernandez, B.; Pugh, T. J.; Hodis, E.; Cho, J.; Suh, J.; Capelletti, M.; Sivachenko, A.; Sougnez, C.; Auclair, D.; Lawrence, M. S.; Stojanov, P.; Cibulskis, K.; Choi, K.; de Waal, L.; Sharifnia, T.; Brooks, A.; Greulich, H.; Banerji, S.; Zander, T.; Seidel, D.; Leenders, F.; Ansen, S.; Ludwig, C.; Engel-Riedel, W.; Stoelben, E.; Wolf, J.; Goparju, C.; Thompson, K.; Winckler, W.; Kwiatkowski, D.; Johnson, B. E.; Janne, P. A.; Miller, V. A.; Pao, W.; Travis, W. D.; Pass, H. I.; Gabriel, S. B.; Lander, E. S.; Thomas, R. K.; Garraway, L. A.; Getz, G.; Meyerson, M. *Cell* **2012**, *150*, 1107.
35. Cancer Genome Atlas Research Network. Electronic address, w. b. e.; Cancer Genome Atlas Research, N. *Cell* **2017**, *169*, 1327.

36. Obeng, E. A.; Chappell, R. J.; Seiler, M.; Chen, M. C.; Campagna, D. R.; Schmidt, P. J.; Schneider, R. K.; Lord, A. M.; Wang, L.; Gambe, R. G.; McConkey, M. E.; Ali, A. M.; Raza, A.; Yu, L.; Buonamici, S.; Smith, P. G.; Mullally, A.; Wu, C. J.; Fleming, M. D.; Ebert, B. L. *Cancer Cell* **2016**, *30*, 404.
37. Lee, S. C.; Abdel-Wahab, O. *Nat Med* **2016**, *22*, 976.
38. Hsu, T. Y.; Simon, L. M.; Neill, N. J.; Marcotte, R.; Sayad, A.; Bland, C. S.; Echeverria, G. V.; Sun, T.; Kurley, S. J.; Tyagi, S.; Karlin, K. L.; Dominguez-Vidana, R.; Hartman, J. D.; Renwick, A.; Scorsone, K.; Bernardi, R. J.; Skinner, S. O.; Jain, A.; Orellana, M.; Lagisetti, C.; Golding, I.; Jung, S. Y.; Neilson, J. R.; Zhang, X. H.; Cooper, T. A.; Webb, T. R.; Neel, B. G.; Shaw, C. A.; Westbrook, T. F. *Nature* **2015**, *525*, 384.
39. Shao, D. D.; Tsherniak, A.; Gopal, S.; Weir, B. A.; Tamayo, P.; Stransky, N.; Schumacher, S. E.; Zack, T. I.; Beroukhim, R.; Garraway, L. A.; Margolin, A. A.; Root, D. E.; Hahn, W. C.; Mesirov, J. P. *Genome Res* **2013**, *23*, 665.
40. Benjamini, Y.; Hochberg, Y. *Journal of the Royal Statistical Society. Series B (Methodological)* **1995**, 289.
41. Massiello, A.; Roesser, J. R.; Chalfant, C. E. *FASEB J* **2006**, *20*, 1680.
42. Younis, I.; Berg, M.; Kaida, D.; Dittmar, K.; Wang, C.; Dreyfuss, G. Mol Cell Biol 2010, 30, 1718.
43. Greider, C. W.; Blackburn, E. H. Nature 1989, 337, 331.
44. Greider, C. W.; Blackburn, E. H. Cell 1985, 43, 405.
45. Kim, N. W.; Piatyszek, M. A.; Prowse, K. R.; Harley, C. B.; West, M. D.; Ho, P. L. C.; Coviello, G. M.; Wright, W. E.; Weinrich, S. L.; Shay, J. W. *Science* **1994**, *266*, 2011.
46. Norton, J. C.; Piatyszek, M. A.; Wright, W. E.; Shay, J. W.; Corey, D. R. *Nat Biotechnol* **1996**, *14*, 615.

Chapter 2: Antibody-directed Cytosolic Delivery of Proteins enabled by Orthogonal Targeting

2.1. Introduction

Proteins such as hormones, cytokines and antibodies have emerged as one of the most important classes of drugs[1,2]. However, most proteins cannot be used against intracellular targets due to their inability to reach the cytosol[3,4]. Although gene delivery has been extensively used to introduce exogenous proteins into cells, it risks permanent disruption of the endogenous genome[5]. Several approaches including peptides[6], polymers[7], and nanoparticles[8,9] have been developed for direct protein delivery but most of them either lack specificity or have limited endosomal release, two key features of an ideal protein delivery strategy.

Monoclonal antibodies (mAbs) are ideal molecules for targeting. They have been exploited in almost every branch of biomedical research and are widely used in humans for the treatments of immune-related diseases and cancer as one of the most successful and important therapeutic strategies[10–12]. Today more than 70 mAb drugs have been approved and more than 550 antibodies are in clinical development[13]. Among them, the predominant form is immunoglobulin G (IgG)[14]. IgG is composed of two distinct functional units: the antigen binding fragment (Fab), which confers antigen specificity, and the constant fragment (Fc), which interacts with the immune cells and extends the antibody half-life in vivo[13]. Due to their great versatility and clinical relevance, there has been a considerable amount of interest in using IgG molecules as drug carriers to target specific type of cells[15]. Antibody drug conjugates (ADCs) and immunotoxins are two examples of using mAbs or mAb fragments to deliver cytotoxic moieties to cancer cells[16,17]. During its general mechanism of action, the ADC is endocytosed via the antibody-receptor interaction into the endosome/lysosome, where the small molecule drug is released from the antibody and diffuses out of the endosome into the cytosol[18,19]. In the case of immunotoxin, after its release from the antibody, the protein toxin self-transports into the cytosol due to its unique natural toxin-derived property[20].

However, for the majority of proteins, the cytosolic delivery using mAbs has remained elusive due to their inability to escape the endosome.

Fortunately, nature has evolved several types of translocation systems for delivery of proteins into host cells[21]. Among them, anthrax lethal toxin is a good example of using a translocase to deliver protein toxins from the endosome to the cytosol. It is composed of three proteins, the translocase protective antigen (PA), and the effector proteins lethal factor (LF)[22] and edema factors (EF)[23]. During its mechanism of action, PA_{83} its receptors on host cell surface and then is proteolytically cleaved into PA_{63} by furin-family protease[24]. Upon cleavage, PA_{63} oligomerizes into a heptameric[25] or octameric[26] pre-pore and three or four LF molecules bind to the pre-pore through their N-terminal domain LF_N with 1-2 nM affinity[27,28]. The entire complex is subsequently endocytosed and the acidic pH in the endosome causes conformational change of the pre-pore to form a trans-membrane pore, which enables the translocation of LF from endosome into cytosol[29–31].

By hijacking this cytosolic delivery machinery, researchers have demonstrated the cellular delivery of a wide variety of molecules including proteins, peptides, and small molecules. These studies have shown LF_N/PA system can efficiently deliver enzymes or enzymatic domains such as beta-lactamase[32], Diphtheria toxin A (DTA)[33], Pseudomonas exotoxin A[34], Ras protease RRSP[35], etc. Moreover, antibody mimics[36] as well as non-natural entities such as mirror image polypeptides and proteins[37] have also been translocated into cells using the LF_N/PA system. In addition to different cargos, retargeted PAs have also been investigated[38,39]. Notably, the Leppla group has reported PA mutants with altered protease sites that can only be activated by matrix metalloproteinase (MMP)[40] or urokinase-type plasminogen activator (uPA)[41], both of which are proteases overexpressed on tumor cells. In this study, we report a novel generalizable antibody-directed protein delivery system by chemically conjugating the protein translocase PA to IgG molecules using the

strategy diagrammed in Fig. 1a. Taking advantage of the modularly designed system, we investigated the delivery of different cargo proteins by different antibodies. Based on the unique property of PA, we created IgG PA conjugates that are able to orthogonally target specific antigen and protease on cell surface simultaneously and demonstrated the efficient delivery of DTA in different cancer models both in vitro and in vivo.

2.2. Results

2.2.1. Modular design of antibody protective antigen conjugate

When we first sought out to design the antibody PA conjugate, we rationalized that an ideal such system would meet four criteria: 1. Both the antibody and PA should be able to carry out their functions as antigen binder and protein translocase in the conjugate (Figure 2.2.1a). 2. The conjugate should be relatively stable in a biological environment. 3. The conjugate should be a homogeneous pure substance. 4. The conjugate should be easy to prepare.

Recombinant expression is a common strategy used to create fusion proteins. Two conceivable ways could be explored for this purpose, with PA expressed either on the N-terminus or C-terminus of the heavy chain/light chain (Figure 2.2.1b). However, significant challenges exist with both approaches. If PA is on the N-terminus of the heavy chain or light chain, there is a high probability that the binding of the antibody will be affected due to the proximity of PA to the antigen binding region. If PA is on the C-terminus, PA_{63} will be cleaved off antibody upon binding to the cell surface and unable to form the pre-pore. This is because PA is only activated after its N-terminal PA_{20} is proteolytically cleaved by furin-family protease to give PA_{63}, leading to formation of pre-pore. Due to these pitfalls, we concluded the recombinant strategy is not suitable in this particular case because it cannot satisfy the first and the most important criterion listed above, to retain the functionalities of both proteins.

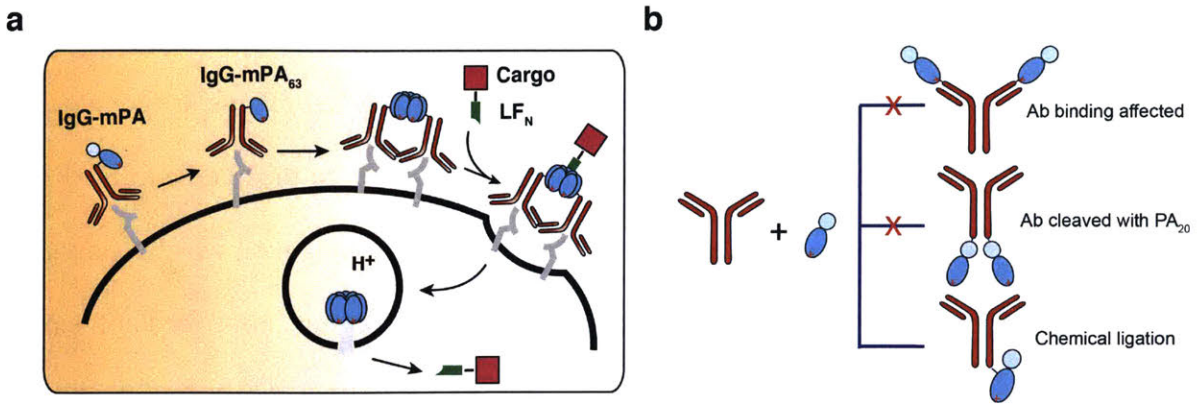

Figure 2.2.1. Modular design of immunoglobulin (IgG) protective antigen (PA) conjugate. a) The proposed mechanism of antibody-mediated delivery by the IgG-mPA system. b) Different possible strategies of conjugating mPA to IgG. Recombinant expression was ruled out due the potential impact on the functionality of IgG or PA.

Chemical conjugation is an alternative to recombinant expression and it gives us more freedom to design the conjugate. In order to preserve the functions of both proteins, we conjugated the C-terminus of the heavy chain, away from the antigen binding site, to the middle of the PA_{63} domain of PA using sortase-mediated ligation. We first created a PA variant on which a polyglycine linker can be installed for conjugation. A single cysteine mutation was introduced at the previously reported lysine 563 position[42] of PA(N682A, D683A), a known PA mutant with ablated native receptor-binding ability while retaining its translocase activity[43]. We were able to express and purify the resulted mutant protein mPAC from Escherichia coli Bl21(DE3) in high yield (Figure 2.6.1). A short D-peptide linker with an N-terminal tri-glycine was crosslinked to the cysteine of mPAC through bromoacetamide (Figure 2.2.2a). The resulted G_3-mPA was subsequently ligated by a triple mutant sortase (SrtA*)[44] onto an IgG molecule with inserted C-terminal LPSTGG tag at its heavy chains. This reaction gave a product mixture of both mono- and dual-mPA tagged IgGs, which was then purified to homogeneous IgG-mPA by size exclusion and anion exchange chromatography with a final yield of greater than 50% (Figure 2.2.2b).

This conjugation strategy was designed to fulfill all four criteria we originally sought for the system. The attachment sites on both mPA and IgG were carefully chosen to avoid interference with their functions as antigen binder and translocase. The C-termini of the heavy chains were used for conjugation to minimize the potential impact of mPA on the IgG molecule and the cysteine was introduced in the PA_{63} domain so that IgG would remain attached after protease-mediated activation of mPA. We performed an in-vitro SDS-resistant pore formation assay to confirm the activity of mPA after attachment to IgG (Figure 2.6.7). Additionally, unlike ADCs or immunotoxins, of which the payloads need to be released from the antibody at the target site, IgG-mPA needs to remain intact during its mechanism of action and the cargo protein comes as a separate entity.

Toward this end, the short bromoacetamide D-peptide linker between IgG and mPA was designed to minimize potential cleavage in a biological environment. Indeed, IgG-mPA prepared with this linker showed no appreciable degradation in 10% serum up to 7 days whereas the maleimide linked conjugate started showing cleavage after 8 hours under the same condition, consistent with previous reports[45,46] (Figure 2.6.8). A longer poly GGS maleimide linker was also tested and exhibited similar activity to the short maleimide linker, indicating the length of the linker does not play a major role (Figure 2.6.9). All IgG-mPA conjugates used in this chapter were thus prepared with the bromoacetamide D linker unless stated otherwise. Moreover, in order to obtain homogeneous IgG-mPA, we took advantage of the pI difference between IgG-mPA and IgG-$(mPA)_2$ and successfully separated them by anion exchange. IgG-mPA was chosen to limit the possible effect on the properties of the IgG, although IgG-$(mPA)_2$ is also active (Figure 2.6.10). We also found that IgG-$(mPA)_n$ prepared by non-specific conjugation using a NHS-Maleimide linker was capable of delivery despite its reduced potency (Figures 2.6.11-12). Last but not least, the entire process of conjugation and purification takes less than half a day and can be scaled up easily, fulfilling the fourth requirement.

Since the cleavage site of mPA can be engineered to target different proteases, mPA variants with different protease specificities were also generated for IgG conjugation. Unlike the recombinant method in which each protein needs to be prepared individually, our conjugation strategy enables rapid combinations of different antibodies and mPA variants in a modular fashion (Figure 2.2.2c). The resulted IgG-mPA variant would require the simultaneous presence of the IgG-specific antigen and the PA-specific protease for activation as demonstrated later in the results (Figure 2.2.2d).

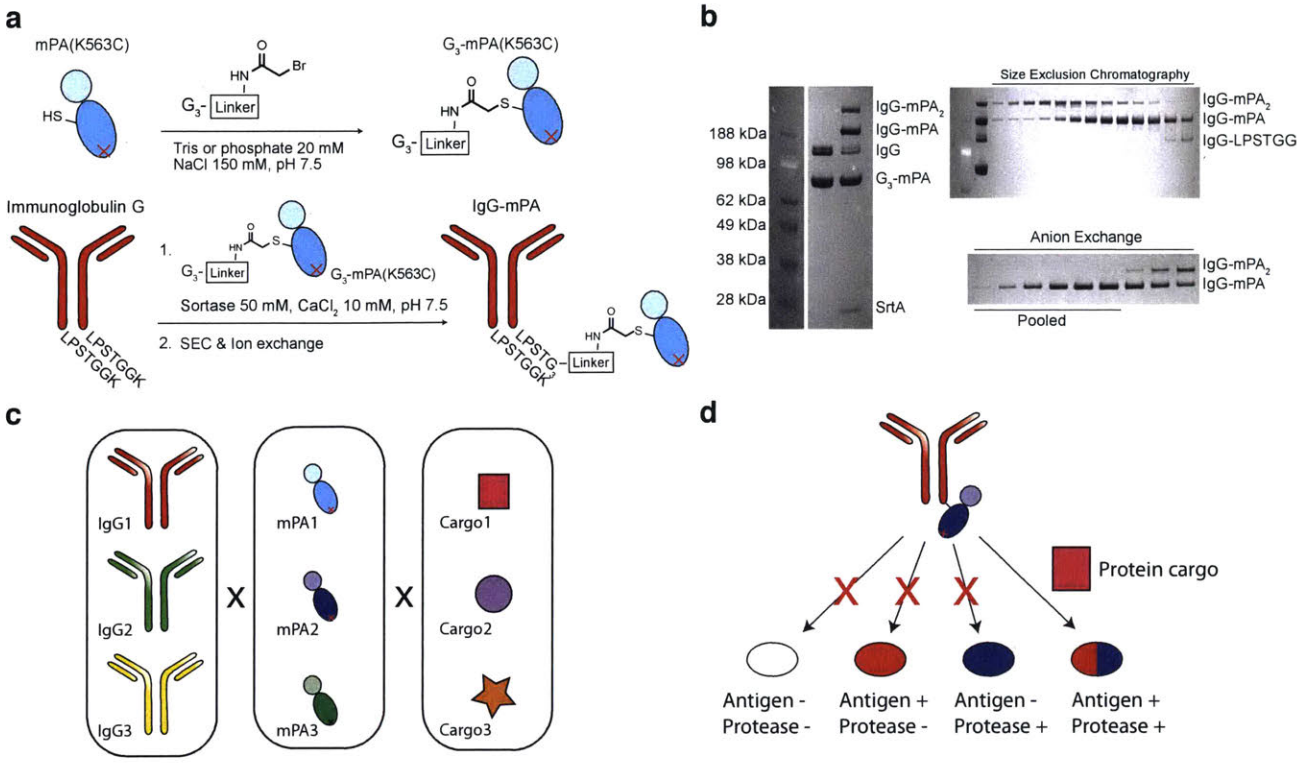

Figure 2.2.2. Preparation of IgG-mPA conjugate using sortase-mediated ligation. a) mPA(K563C) was put on a polyglycine linker by either maleimide or bromoacetamide before sortagging onto IgG with LPSTGG tags. b) The resulted product mixture of IgG-mPA and IgG-$(mPA)_2$ was purified to homogeneous IgG-mPA by standard size exclusion and anion exchange chromatography. The IgG in this experiment is Trastuzumab. The preparation of other IgG-mPA variants are described in the supplementary information. c) Free combinations of different antibodies and different mPA variants enable orthogonally targeted delivery of different protein cargos. d) Schematic showing IgG-mPA variant delivers a protein cargo to specific cell type by dual targeting of particular antigen and protease.

2.2.2. Antibody-mediated delivery of different protein cargos in specific cell types

Epidermal growth factor receptor (EGFR) and HER2 are among the most studied cell surface receptor tyrosine kinases. They are implicated in tumor development and are overexpressed in a wide variety of human cancers including breast, ovarian, lung, colon, etc[47–51]. The high interest in these receptors as anticancer targets has led to the development of various therapeutics from small molecule inhibitors to large antibodies. Trastuzumab (Tmab) and cetuximab (Cmab) are two humanized recombinant monoclonal antibodies that bind to HER2 and EGFR respectively[52,53]. They both have been approved by FDA years ago and are therefore suitable for our study.

We first investigated the ability of Tmab-mPA, prepared as described above, to deliver LF_N-DTA into HER2 positive cells. LF_N-DTA (LD) is a fusion protein between LF_N, the PA pre-pore binding domain and DTA, the catalytic domain of diphtheria toxin. The DTA moiety inhibits protein synthesis and causes cell death by ribosylating the eukaryotic elongation factor 2 within the cytosol[54]. We found that 10 nM LD in the presence of Tmab-mPA effectively reduced the viability of HER2-overexpressing BT474 cells[55], indicating potent cytosolic delivery of DTA, whereas LD in the presence of unconjugated Tmab/mPA only exhibited similar antiproliferative effect to Tmab-mPA or Tmab alone (Figure 2.2.3a). We then examined whether the translocase function of PA is required for the DTA delivery by using mPA(F427A), a mutant PA that is incapable of translocation but still binds LF_N[56]. As expected, LD with Tmab-mPA(F427A) failed to show any enhanced activity. To be thorough, we conjugated LD directly to Tmab (Figure 2.6.13). The conjugate Tmab-LD also did not show any increased killing, confirming mere endocytosis of LD is not enough to trigger DTA activity.

We next wanted to see if such delivery strategy can be expanded to another antibody. EGFR overexpressing A431 cells were treated with Cmab-mPA in the presence of LD under the same

conditions. Just like in the case of Tmab-mPA, only Cmab-mPA combined with LD showed significant antiproliferative effect in A431 cells, whereas no other treatment showed any difference from Cmab alone, verifying both Cmab and functional PA are both required for the cytosolic delivery of DTA (Figure 2.2.3b). These results provided strong evidence that this IgG-mPA cytosolic delivery strategy can be applied to different antibodies targeting different antigens.

We then investigated the specificity of this delivery method in different cell lines. When we treated BT549, a triple negative breast cancer (TNBC) cell line lacking HER2[57], with Tmab-mPA and LD under the same conditions as BT474, no significant reduced viability was observed up to the highest concentration tested (Figure 2.2.3c). Likewise, when we challenged EGFR negative Jurkat cells with Cmab-mPA and LD, the cells showed very little change of viability (Figure 2.2.3d). In contrast, significant toxicities were observed when we treated both cell lines with wildtype PA and LD (Figure 2.6.14). In addition, to further verify the specificities of IgG-mPA conjugates, same cell lines were treated with Tmab-mPA or Cmab-mPA in the presence of LD. We observed major toxicity difference between the two conjugates in good accordance with the relative HER2 and EGFR expression level (Figure 2.6.15). Taken together, these data suggest IgG and mPA can retain their separate functions in the conjugate and the delivery of the protein cargo depends on both of them.

After verifying the delivery of DTA by IgG-mPA, we further explored whether such a system can be used to deliver other protein cargos. Edema factor (EF), one of the native effector proteins of anthrax toxin, is a calmodulin and Ca^{2+} dependent adenylate cyclase, which increases intracellular concentration of cyclic AMP (cAMP)[23]. cAMP has been found to be involved in a variety of biological processes including the immune system[58]. As a result, several drugs have been or are being developed to modulate cAMP level to attenuate autoimmune and inflammatory

diseases[59]. Here we demonstrate the ability to increase intracellular cAMP concentration through antibody-directed delivery of EF. EGFR-expressing TNBC MDA-MB-231 cells[55] were incubated with 20 nM of EF in the presence of Cmab-mPA or other controls for 2 hours and the cAMP levels were subsequently measured by a commercial ELISA-based competition assay. We observed significant elevation of cAMP level only when Cmab-mPA was used as the delivery vehicle. Translocation-disabled Cmab-mPA(F427A) failed to enhance cAMP concentration, suggesting entrapment of EF in the endosome. Unconjugated Cmab/mPA also showed similar level of cAMP to the untreated background, indicating conjugation is required for the activity of IgG-mPA (Figure 2.3.3e). Additionally, Tmab-mPA with EF failed to increase the cAMP level, confirming again the delivery is antibody-dependent. When we replaced the cargo protein with LD, little change of cAMP level was detected, demonstrating the effect is cargo specific.

We further expanded the protein cargo library to a third protein. Ras is an oncoprotein that plays a critical role in cancer development and Ras/Rap1-specific endopeptidase (RRSP) is a recently characterized Ras protease that rapidly cleaves Ras once delivered into cells by wildtype PA[35]. MDA-MB-231 cells were treated with LF_N-RRSP (LR) in the presence of Cmab-mPA and other controls for 24 hours. The cells were subsequently lysed and analyzed by Western blots. As shown in Figure 2.2.3f, Ras exhibited dose-dependent degradation as a result of RRSP induced cleavage. ERK is a downstream target protein activated by Ras and a key player in cell growth regulation. While the total amount of ERK remained unchanged across different conditions, phospho-ERK (pERK) showed dose-dependent suppression similar to Ras in the presence of LR and Cmab-mPA, indicating ERK1/2 dephosphorylation caused by RRSP. On the contrary, just like in the previous experiments, Cmab-mPA(F427A), unconjugated Cmab/mPA and Tmab-mPA did not show any obvious activity compared to the background.

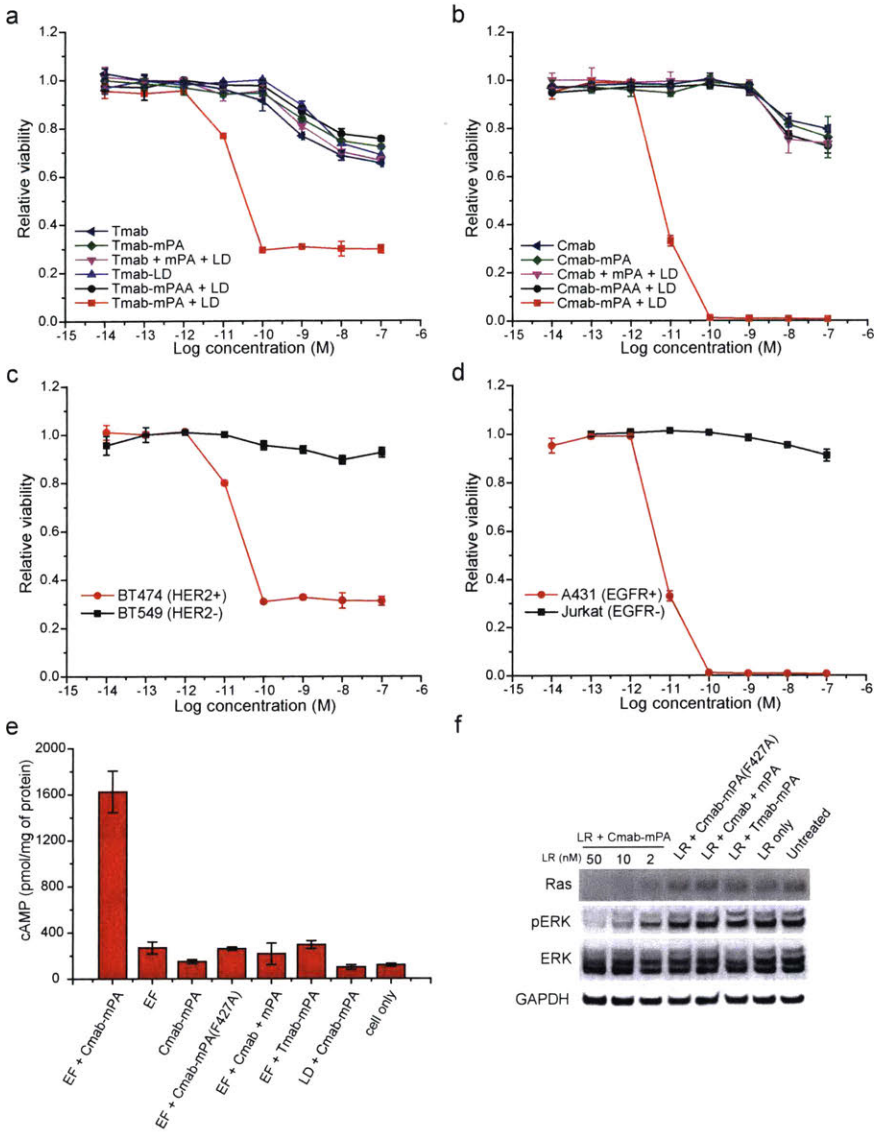

Figure 2.2.3. IgG-mPA delivers different protein cargos to different cell types in a receptor-specific manner. a) HER2 expressing BT474 cells were treated with ten-fold serial dilutions of Tmab or Tmab conjugates with or without 10 nM LD in medium containing 10% FBS for 72 hours. mPAA stands for mPA(F427A), a translocation-deficient mutant of mPA. The cell viabilities were measured by CellTitor-Glo assay and normalized to untreated cells. b) EGFR expressing A431 cells were treated with Cmab or Cmab conjugates the same way as described in a. c) Viability comparison between HER2 positive BT474 and HER2 negative BT549 cells after treatment with Tmab-mPA and 10 nM LD. d) Viability comparison between EGFR positive A431 and EGFR negative Jurkat cells after treatment with Cmab-mPA and 10 nM LD. e) EGFR expressing TNBC MDA-MB-231 cells were treated with 20 nM EF or LD and different Cmab or Tmab conjugates for 2 hours. The intracellular cAMP concentrations were measured by an ELISA-based competition assay. f) MDA-MB-231 were incubated with different concentrations of LR in the presence of 50 nM Cmab-mPA for 24 hours. 50 nM of LR with different Cmab or Tmab conjugate controls were also included.

2.2.3. Overcome drug resistance in different HER2 and EGFR expressing cancer cell lines with Trastuzumab and Cetuximab mediated delivery of diphtheria toxin A

Since the approval of Tmab (Herceptin) for the treatment of HER2-overexpressing breast cancer, a significant number of patients with this type of breast cancer have developed resistance to Tmab treatment[60,61]. Similarly, EGFR inhibitors such as Gefitinib and Cmab (Erbitux) have also been found to be ineffective in EGFR-positive cancer patients[62–64]. Such resistance is often attributed to RAS mutations that lead to constitutive activation of EGFR signaling[65]. In addition, previous studies have shown cancer cells often still retain HER2 or EGFR expression when they become refractory to HER2- or EGFR-based therapies[64,66,67]. Based on these results, we hypothesized the drug resistance could be overcome by Tmab- and Cmab-directed delivery of DTA, which induces cellular toxicity through a different mechanism.

Upon the intoxication of LD+Tmab-mPA, HER2-expressing ovarian cancer cell line SKOV-3 displayed dose-dependent cell growth inhibition but showed little response to Tmab treatment (Figure 2.2.4a). Likewise, EGFR-positive human non-small-cell lung cancer (NSCLC) cell lines A549 and H441, colorectal cancer cell line HCT-116, TNBC cell lines BT549 and MDA-MB-231 all showed high susceptibilities to LD+Cmab-mPA and no significant response to Cmab or Gefitinib treatment, consistent with previous studies[68–70] (Figures 2.2.4b-f).

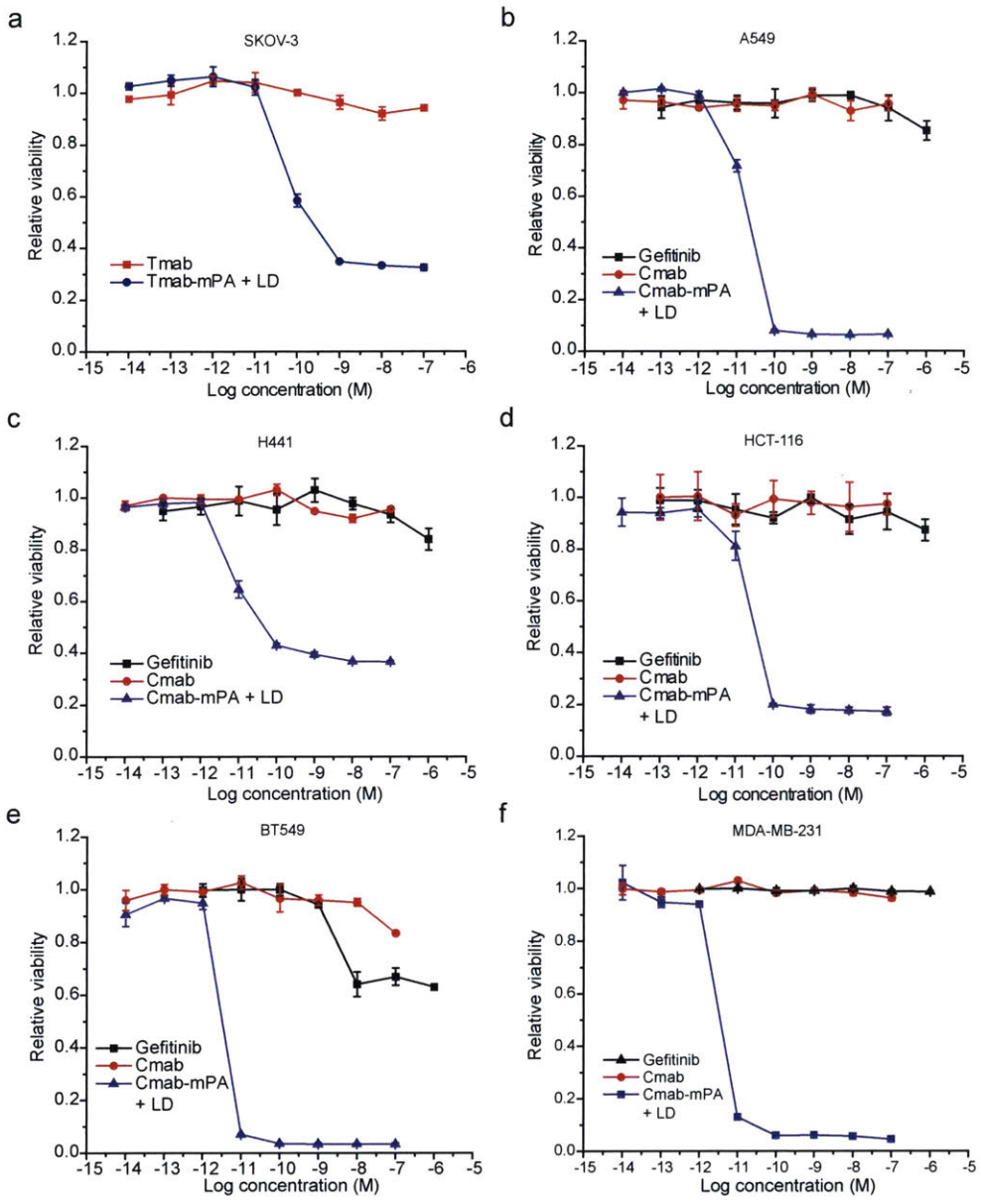

Figure 2.2.4. Resistance to HER2 or EGFR based therapies can be overcome by Tmab or Cmab delivered DTA. Different refractory cancer cell lines were treated with either IgG or IgG-mPA/LD in medium containing 10% FBS for 72 hours. The Cmab resistant cell lines were also treated with Gefitinib, an FDA approved EGFR inhibitor. The relative cell viabilities were measured by CellTitor Glo assay and normalized to untreated cells. a) Human ovarian cancer cell line SKOV-3. b, c) Human non-small-cell lung cancer (NSCLC) cell lines A549 and H441. d) Human colon cancer cell line HCT-116. e, f) Triple negative breast cancer (TNBC) cell lines BT549 and MDA-MB-231.

2.2.4. Orthogonal targeting of cancer specific receptors and proteases by IgG-mPA variants

Although HER2 and EGFR are often overexpressed in cancer cells, they also play essential roles in regular cellular functions and therefore are present in normal tissues[71,72]. In fact, it is rare to find a single antigen exclusively on the cells of interest. This could be particularly problematic when the delivered cargo is a highly toxic protein due to the possibility of off-target effect. However, the specificity of a delivery system can be significantly enhanced if it requires two entities simultaneously at the target site for activation. During PA-mediated translocation, after binding to its cellular receptors, PA can only oligomerize upon cleavage by furin-family proteases after the sequence RKKR[24]. Leveraging this requirement of a protease during PA activation, the Leppla group has developed PA variants with re-engineered cleavage site susceptible to MMP or uPA[40,41], both of which are over produced by tumor tissues and implicated in cancer cell growth and metastasis. Here we combined these PA variants with an IgG molecule using the modular conjugation strategy and created Cmab-mPA-MMP and Cmab-mPA-uPA, which orthogonally target EGFR and MMP or uPA. These dual targeting IgG-mPA variants require the concurrent presence both the antigen and a specific cell surface protease for activation (Figure 2.2.2d), providing another degree of selectivity for the delivery system. Using these variants, we demonstrated highly specific delivery of DTA into particular cancer cell lines.

We first verified the susceptibilities of Cmab-mPA, Cmab-mPA-MMP and Cmab-mPA-uPA to proteases furin, MMP and uPA respectively. The conjugates were incubated individually with furin, uPA, or MMP-9 for 3 hours at 37°C and analyzed by reducing SDS-PAGE gel, which reduced the conjugates to mPA tagged heavy chain (HC-mPA), HC, and light chain (LC). The cleavage of mPA was indicated by the downward shift of HC-mPA to HC-mPA_{63} and the appearance of PA_{20} band (Figure 2.2.5a). As expected, each conjugate variant was only vulnerable to its

specific protease. The band between HC-mPA$_{63}$ and HC in MMP-9 treated Cmab-MMP was possibly caused by subsequent cleavage of HC-mPA$_{63}$ by MMP-9.

We then tested the hypothesis that Cmab-mPA-MMP and Cmab-mPA-uPA can target cells with both EGFR and their corresponding proteases. LD was again used as a model protein cargo due to its robust viability readout. TNBC cell lines BT549 and MDA-MB-231, which express EGFR, MMP-9, and uPA[73,74], were treated with Cmab-mPA-MMP or Cmab-mPA-uPA in the presence or absence of LD for 72 hours. Both cell lines showed high sensitivities to the IgG-mPA variants that is comparable to Cmab-mPA, indicating the modified cleavage site did not significantly affect IgG-mPA activity when the targeted protease is present (Figures 2.2.5b-c). In contrast, H2030 cells, which are EGFR positive but lack MMP and uPA[75], showed complete resistance to LD with Cmab-mPA-MMP and Cmab-mPA-uPA but not with Cmab-mPA (Figure 2.2.5d). In addition, more attenuated LD toxicity was observed when normal human endothelial HMEC-1 cells were treated with Cmab-mPA-MMP and Cmab-mPA-uPA when compared to Cmab-mPA, consistent with our assumption that the dual-targeting further improves the specificity of the delivery system. Meanwhile, HMEC-1 still remained highly susceptible to wild-type PA + LD, confirming the cell line is still sensitive to LD delivery (Figure 2.2.5e).

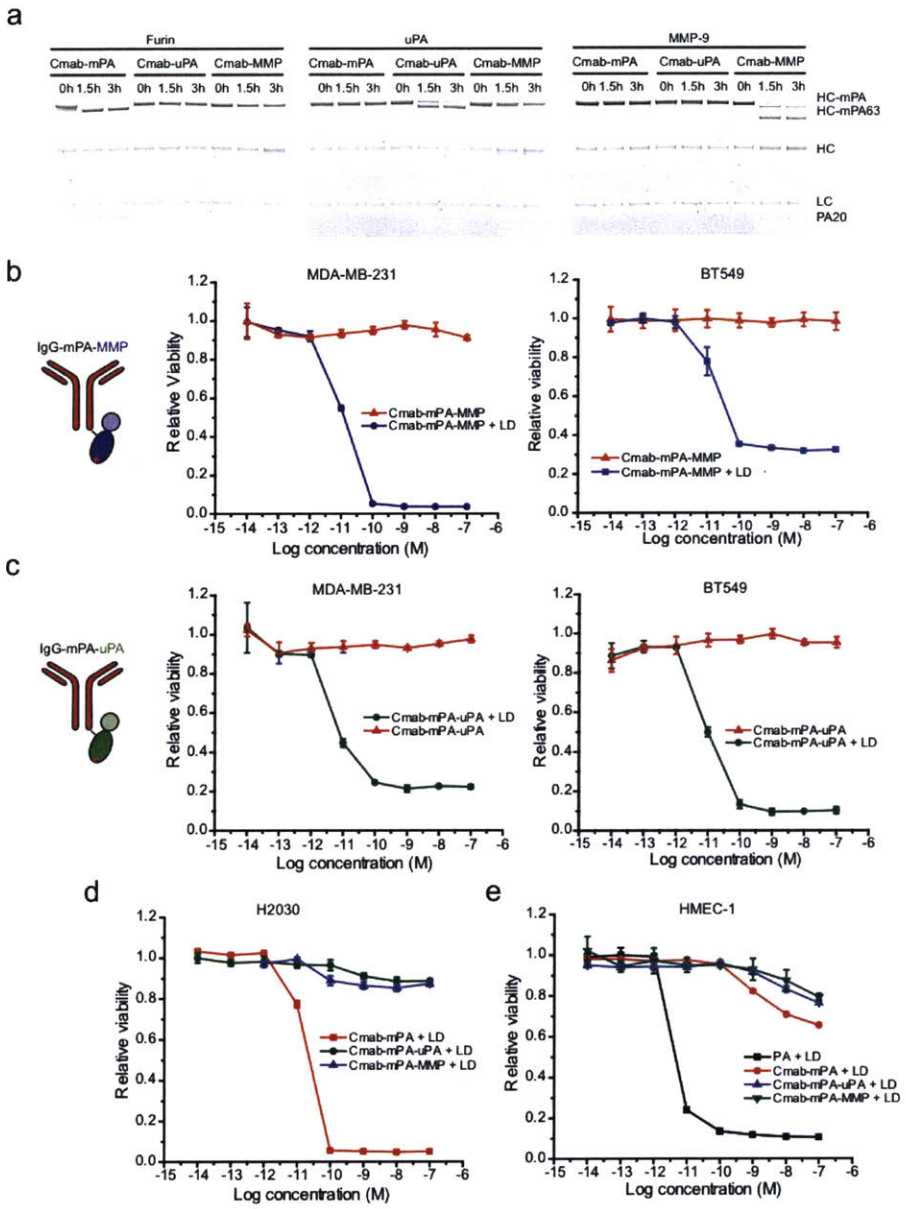

Figure 2.2.5. Orthogonal targeting IgG-mPA variants require the presence of both the antigen and the protease for cargo delivery and hence have less off-target effect. a) In vitro cleavage of Cmab-mPA, Cmab-mPA-uPA (Cmab-uPA), and Cmab-mPA-MMP (Cmab-MMP) by furin, uPA, or MMP-9. The Cmab conjugate was incubated with the corresponding protease at 37°C for the times indicated. The samples were analyzed by SDS-PAGE and stained with coomassie dye. b) EGFR/uPA positive MDA-MB-231 and BT549 cells were treated with or without 10 nM LD in the presence of ten-fold serial dilutions of Cmab-mPA-uPA for 72 hours before subjecting to CellTitor Glo assay. c) MDA-MB-231 and BT549 cells, which also express MMP, were treated with Cmab-mPA-MMP and LD as well. d) EGFR positive and uPA/MMP negative H2030 cells were incubated with serial dilutions of all three Cmab conjugates and 10 nM LD. e) Cytotoxicities of LD delivered by wild-type PA or different Cmab-mPA variants were tested in normal human microvascular endothelial cell line HMEC-1.

2.2.5. Protein toxin delivered by orthogonal targeting Cmab-mPA variants showed significantly improved safety profile and pharmacokinetic properties

We next investigated whether the attenuated toxicity observed in cell culture was able to translate into an animal model. Female nude mice were challenged intravenously with 1 mg/kg of LD alone or in combination with increasing amounts of Cmab-mPA-uPA or Cmab-mPA-MMP. The injected mice were closely monitored for any outward signs of toxicity for 48 h. After the mice were sacrificed, the tissue samples from kidney and liver were also subjected to microscopic examinations. Cmab-mPA-uPA or Cmab-mPA-MMP combined with LD showed no outward or histological signs of toxicity up to the highest doses (15 mg/kg) tested (Figure 2.2.6a), significantly higher than any of the previously reported PA variants.

In order to study the pharmacokinetic properties of the Cmab-mPA variants, we labeled wild-type PA, Cmab-mPA-uPA, or Cmab with radioactive ^{89}Zr and injected mice (n = 4) intravenously with 1 mg/kg of each protein. PET images acquired 6 h post injection showed similar biodistribution pattern between Cmab and Cmab-mPA-uPA, with some accumulation in liver and significant portion still remained in circulation (Figure 2.2.6b). In contrast, most of the WT PA was found in liver, indicating a fast clearance. In parallel, blood samples in different treatment groups were collected at different time points for radioactivity measurements by gamma counter. Consistent with PET imaging results, Cmab-mPA-uPA showed significantly prolonged clearance time compared to the WT PA (Figure 2.2.6c).

In addition to Tmab and Cmab, we also conjugated mPA to Daclizumab as a mean to target regulatory T cells. Due to time constraint, only proof-of-concept experiments were carried out. Please see section 2.6.5 for detailed information.

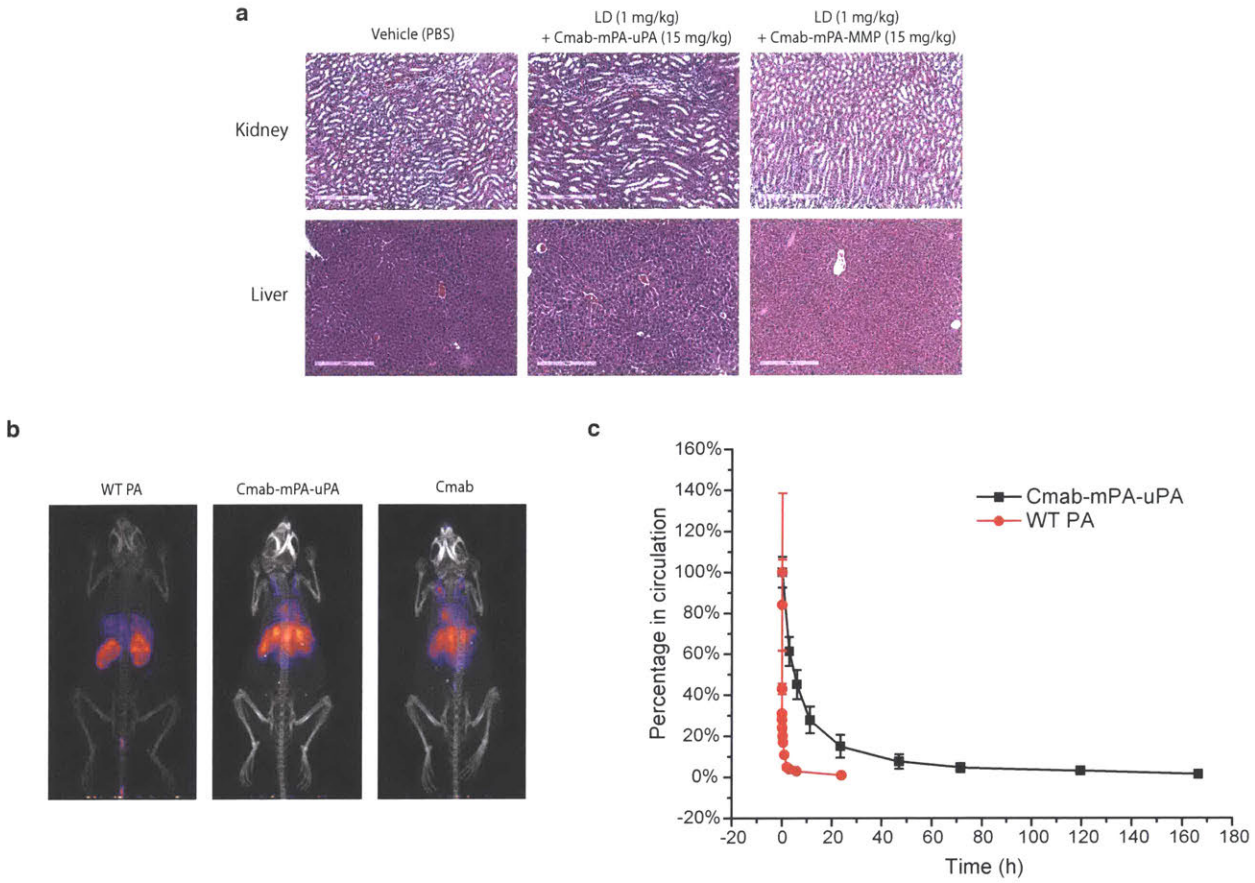

Figure 2.2.6. In vivo studies of Cmab-mPA variants. a) Microscopic appearance of kidney and liver of mice 48 h after i.v. injection with PBS or 1 mg/kg of LD plus 15 mg/kg of Cmab-mPA variants. b) Whole-body PET imaging of mice 6 h after i.v. injection of ^{89}Zr-labeled wild-type PA, Cmab, or Cmab-mPA-uPA (1 mg/kg). c) Comparison of relative blood concentration between Cmab-mPA-uPA and wild-type PA. Blood samples (n=4) were taken at different time points and measured for their radioactivity as described in the methods section. Relative %ID/g was normalized to time 0.

2.3. Discussion

Targeted intracellular protein delivery poses a major challenge since the vehicle needs to not only reach the target site but also release the cargo into the cytosol. Monoclonal antibodies possess the necessary versatility and clinical properties to be ideal targeting ligands. However, their application in protein delivery has been limited due to its lack of endosomal release. In this work, we report a novel targeted protein delivery platform by modularly combining IgG molecules with the protein translocase PA through a chemical ligation strategy. We demonstrated the delivery vehicle is entirely IgG-guided and requires the function of the translocase for activity. We used this method to deliver proteins with a variety of functions, including ribosyltransferase DTA, adenylate cyclase EF, and Ras protease RRSP.

Pathogenic bacteria often express toxins that are capable of delivering effector proteins into the cytosol of cells. PA is a protein translocase used by Bacillus anthracis to deliver its cytotoxic payload lethal factor (LF). After nearly 40 years of extensive mechanistic and structural studies, a model for PA-mediated protein translocation has been established. Biochemical and biophysical analyses have revealed how the N-terminal, PA-binding domain of LF (LF_N) interacts with the PA prepore to initiate protein translocation[27]. Using PA/LF_N as a delivery platform, we and others have previously demonstrated its efficient cytosolic delivery of a wide variety of natural and unnatural entities, including Pseudomonas exotoxin A (PE), β lactamase, dihydrofolate reductase, GFP, DARPin, L and D affibody, L and D peptides, etc[76]. Based on our in vitro pore formation assay of IgG-mPA and the data on the delivery of three protein cargos with diverse functions, we believe the translocation mechanism of the IgG-mPA system is similar to that of the wild type PA and therefore should be compatible with the types of cargos that have previously been studied. The

free combinations of different IgGs with protein cargos can potentially greatly expand the scope of protein delivery.

When cancer cells are targeted for delivery, the payload is typically cytotoxic. Immunotoxins for example are protein toxins engineered to target cancer cells through ligands or antibodies. However, they are often associated with dose-limiting toxicities as the result of off-target damage to other tissues and cells[20,77]. This is because most of the tumor cell surface antigens or receptors are also present on normal cells and even a small amount of such toxins can cause serious harm due to their catalytic nature[78]. Here we took advantage of the protease-gated activation of PA and the fact that most aggressive tumors require particular proteases such as uPA and MMPs for the dissolution of extracellular matrix, which is a prerequisite for tumor invasion and metastasis[79,80]. By replacing furin-dependent mPA with uPA- or MMP-dependent mPA variants in the IgG-mPA conjugate, we created a highly specific delivery system that requires the simultaneous presence of IgG-specific antigen and mPA-specific protease for activation. When we used a highly potent toxin as the cargo of delivery, the dual-targeted system displayed significantly improved selectivity compared to previous studies both in vitro and in vivo. Moreover, combinations of IgGs and mPA variants needed little effort thanks to the modular design of the conjugation strategy, which made the process easy and highly reproducible among different IgGs and mPAs. With an enormous selection of monoclonal antibodies and many cell surface proteases expressed by human cells and pathogens, we envision this delivery system can be easily tailored for a variety of diseases.

One potential limitation with this delivery platform is that PA and LF_N are both bacteria-derived foreign proteins and therefore can induce antibody generation that limits their repeated

dosages in a clinical setting. However, recent studies have found that an immunosuppressive regimen that consists of pentostatin and cyclophosphamide (PC) can prevent the immunogenic response against PA/LF_N by depleting lymphocytes, particularly B cells, in immunocompetent C57BL/6J mice[81,82]. In addition, another group has reported induced immune tolerance to recombinant immunotoxins by co-administering them with nanoparticles containing rapamycin (SVP-R)[83,84]. It is conceivable that the combination of IgG-mPA with these strategies will further improve its utility as a protein drug delivery platform in a broad spectrum of diseases.

2.4. Experimental

2.4.1. Materials

Fmoc-protected L- and D-amino acids used for peptide synthesis were purchased from ChemImpex. Plasmids of Trastuzumab and Cetuximab were generously provided by Dr. Dane Wittrup (Koch Institute for Integrative Cancer Research, Cambridge, MA). The plasmid encoding LF_N-RRSP was a generous gift from Dr. Karla Satchell (Northwestern University). All antibodies used were purchased from Cell Signaling unless stated otherwise. All media used in cell culture were purchased from Thermo Fisher.

2.4.2. Synthesis and purification of peptide linkers

Peptide linkers were synthesized on ChemMatrix resin with a Rink amide linker using either manual flow peptide synthesis or automated flow peptide synthesizer as previously described[85,86]. The bromoacetamide or maleimide was coupled on resin as described in the supplementary information. The crude peptides were cleaved by TFA/H_2O (97.5:2.5) and purified by semi-preparative RP-HPLC with Agilent Zorbax 300SB C_{18} column (9.4 x 250 mm, 5 um) at a flow rate of 4 mL/min using the gradient of 1-31% acetonitrile over 80 min. Pure HPLC fractions were subsequently pooled and analyzed by LC-MS (supplementary figure).

2.4.3. Construction of PA and IgG mutants

The plasmids of mPA (N682A, D683A), mPAC (N682A, D683A, K563C), and mPACA (N682A, D683A, K563C, F427A) were prepared from wild type PA using QuickChange multi site-directed mutagenesis kit (Agilent) according to the manufacturer's protocol. The furin cleavage site ^{164}RKKR^{167} in mPAC was replaced by the uPA substrate sequence PGSGRSA or gelatinase substrate sequence GPLGMLSQ to give mPAC-uPA or mPAC-MMP using QuickChange

single site-directed mutagenesis kit (Agilent). Using the same kit, the sortase-specific conjugation tag LPSTGG was inserted at the C-terminus of the heavy chain of Trastuzumab and Cetuximab.

2.4.4. Expression and purification of PA mutants, EF, and LF_N-RRSP

All non-IgG proteins were expressed in E. coli BL21 (DE3) cells from Thermo Fisher. The expression of PA variants was done at New England Regional Center of Excellence/Biodefense and Emerging Infectious Diseases (NERCE) and purified as previously described[31]. In short, cell pellet was first resuspended in sucrose buffer (20 mM Tris pH 8.5, 1 mM EDTA, 20% sucrose) and then incubated with 5 mM $MgSO_4$. The supernatant was purified by an anion exchange column, after which the pure fractions were pooled and concentrated based on the analysis of SDS-PAGE. EF and LF_N-RRSP were expressed in Champion PET-SUMO vector with a His tag in E. coli BL21(DE3) and were induced at OD 0.7~0.9 with 0.4 mM IPTG at 20°C for 16 h. The proteins were purified by a HisTrap FF Ni-NTA column. The SUMO was cleaved by SUMO protease and removed by size exclusion chromatography.

2.4.5. Expression and purification of IgGs

Trastuzumab and cetuximab in gWiz vector were transiently transfected using PEI and expressed in FreeStyle 293-F cells (Thermo Fisher) according to the manufacturer's protocol. The IgGs were subsequently purified from the medium by Protein A affinity chromatography (Genscript) and stored in PBS at -80°C.

2.4.6. Conjugation of mPA and IgGs

mPAC or its variants (400 uM) was first conjugated with the peptide linker (2 mM) in 20 mM Tris and 150 mM NaCl (pH 8.5). After the reaction went to completion in 1 h, the excess of peptide was removed by three rounds of buffer exchange with 30K Amicon ultra-15 centrifugal filter (Millipore). The resulted G_3-mPA (100 uM) was incubated with IgG-LPSTGG (40 uM) in

the presence of 5 uM of triple mutant sortase (SrtA*) in sortase buffer (50 mM Tris, 150 mM NaCl, 10 mM $CaCl_2$, pH 7.5) for 1 h at room temperature. The reaction mixture was then loaded onto HiLoad 16/600 Superdex 200 pg size exclusion column (GE). The fractions containing IgG-mPA and IgG-$(mPA)_2$ were pooled and purified again by 5-mL HiTrap Q anion exchange column (GE) to give pure IgG-mPA.

2.4.7. IgG-mPA stability

Cmab-mPA and Cmab-M-mPA prepared with bromoacetamide linker and maleimide linker respectively were incubated in MEM supplemented with 10% FBS at 37°C. Samples from different time points were analyzed by western blot using anti-PA antibody (Santa Cruz Biotechnology).

2.4.8. In vitro cleavage of Cetuximab-mPA variants by furin, uPA, and MMP-9

10 μg of different Cetuxiamb-mPA variants were incubated with 1 μl of furin (>2000 unit/ml, Sigma), 1 μg of uPA (Millipore), or 0.2 μg of MMP-9 (Millipore) in a total reaction volume of 40 μl. The cleavage reactions were performed in similar conditions as previously described[40,41]. Furin cleavage was done in 25 mM HEPES (pH 7.5), 150 mM NaCl, 0.2 mM EDTA, 0.2 mM EGTA, 1.0 mM $CaCl_2$, and 1.0 mM $MgCl_2$. uPA cleavage was done in 20 mM Tris·HCl (pH 7.5) and 150 mM NaCl. MMP-9 cleavage was carried out in 50 mM HEPES (pH 7.5), 10 mM $CaCl_2$, 200 mM NaCl, 0.05% (v/v) Brij-35, and 50 μM $ZnSO_4$. Aliquots at different time points were analyzed by reducing SDS-PAGE using Bolt 4-12% Bis-Tris plus gel (Thermo Fisher).

2.4.9. Cell culture

All cell lines were purchased from American Type Culture Collection (Manassas, VA). The cells were grown at 37°C in a 5% CO_2 environment. All cells were maintained according to the instructions on ATCC except that MDA-MB-231 cells were grown in DMEM with 10% FBS.

2.4.10. Cell viability assays

Cells were seeded in a 96-well plate at a density of 5×10^3 per well and allowed to attach overnight. The next day, the cells were treated with 10-fold serial dilutions of different protein cargos with or without various IgG-PA variants for 72 h. The cell viability was measured by CellTiter-Glo luminescent assay (Promega) following the manufacturer's protocol. The relative viability was normalized to cells without treatment.

2.4.11. cAMP assay

MDA-MB-231 cells (1.5×10^4) were plated in a 96-well plate on the previous day of treatment. Cells were incubated with 20 nM EF or LD in the presence or absence of 100 nM different IgG-mPA variants for 2 h. Cells were lysed in 0.1 M HCl with 0.5% Triton X-100. A colorimetric competitive ELISA kit (Enzo Life Science) was used to measure the intracellular cAMP levels according to the manufacturer's instructions.

2.4.12. Western blotting

MDA-MB-231 cells were plated at 2×10^5 cells/well in a 12-well plate. After 24 h, the cells were treated with 50 nM LF_N-RRSP (LR) in the presence or absence of different IgG-mPA variants for another 24 h. Subsequently, the medium was removed and the cells were lysed by RIPA buffer (10 mM pH 8.0 Tris-Cl, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS, 140 mM NaCl, 1 mM PMSF) supplemented with Roche protein inhibitor for 20 minutes on ice. The lysate was separated by SDS-PAGE and transferred onto a PVDF membrane (Bio-Rad) using a Bio-Rad Trans-Blot turbo transfer system. The membrane was blocked with LI-COR blocking buffer on a shaker at room temperature for 1 h before the membrane was incubated with different primary antibodies in TBST buffer and 0.1% Tween overnight at 4°C. On the next day, the mem-

brane was washed by TBST three times and then incubated with the appropriate fluorescent secondary antibody in TBST on a shaker for 1 h at RT. The imaging was done by ChemiDoc MP imaging system (Bio-Rad). Serial detection of proteins was done by stripping the membrane with stripping buffer (Thermo Fisher) and repeating the above staining procedure.

2.4.13. Histopathological analysis

All animal work was conducted in accordance with a protocol approved by the MIT Committee on Animal Care. We used a dose-escalation protocol to characterize the potential toxicity of LD in combination with dual-targeting Cetuximab-mPA variants. Female nude mice between 8 to 12 weeks (n = 3) were anesthetized by isoflurane inhalation and injected intravenously with either a single dose of vehicle PBS or LD (1 mg/kg). For combination dose, a second injection was performed to administer the 3 mg/kg or 15 mg/kg of Cetuximab-mPA-uPA or Cetuximab-mPA-MMP. The mice were closely monitored for any signs of toxicity. After 48 h, the mice were euthanized, tissues were fixed in 4% formalin, embedded in paraffin, and sections were stained with hematoxylin and eosin (H&E) for microscopic analysis by a pathologist.

2.4.14. PK study

Female nude mice between 8 to 12 weeks were injected i.v. through the tail vain with wild type PA or Cetuximab-mPA-uPA labeled with ^{89}Zr (Washington University School of Medicine in St. Louis) at 1 mg/kg (n = 4). 5-15 mg of blood samples at indicated time points were collected by tail snipping and the radioactivity was measured by PerkinElmer Wizard 2 gamma counter. The PET scans were acquired on a G8 PET/CT preclinical imaging system (PerkinElmer). After 166 h, the mice were sacrificed by CO_2 followed by cervical dislocation. The organs and tissues were collected and measured by the gamma counter.

2.5. Acknowledgements

This work was funded by MIT start-up funds, MIT Reed Fund, Damon Runyon Cancer Research Foundation Innovation Award, National Science Foundation (NSF) CAREER Award (CHE-1351807) for B.L.P, F32 CA180653 (B.R.P.), Novartis Early Career Award (Novartis Institutes for BioMedical Research Incorporated), Unrestricted Grant in Synthetic Organic Chemistry (Bristol Myers SQUIBB), the Friends of DFCI (B.R.P.). We thank R. J. Collier (Harvard) for his continued contribution of laboratory equipment and the NERCE facility (grant: U54 AI057159) for expression of toxin proteins. The animal work was partially supported by the Cancer Center Support Grant P30CA14051 from the National Cancer Institute.

2.6. Appendix

2.6.1. Amino acid sequences of mPAC and mPACA

PA[N682A, D683A, K563C] (mPAC)

MEVKQENRLLNESESSSQGLLGYYFSDLNFQAPMVVTSSTTGDLSIPSSELENIPSENQYF
QSAIWSGFIKVKKSDEYTFATSADNHVTMWVDDQEVINKASNSNKIRLEKGRLYQIKIQ
YQRENPTEKGLDFKLYWTDSQNKKEVISSDNLQLPELKQKSSNS**RKKR**STSAGPTVPDR
DNDGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSTASDPYSDFEK
VTGRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRTH
TSEVHGNAEVHASFFDIGGSVSAGFSNSNSSTVAIDHSLSLAGERTWAETMGLNTADTA
RLNANIRYVNTGTAPIYNVLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYYPSKNLAPI
ALNAQDDFSSTPITMNYNQFLELEKTKQLRLDTDQVYGNIATYNFENGRVRVDTGSNW
SEVLPQIQETTARIIFNGKDLNLVERRIAAVNPSDPLETTKPDMTLKEALKIAFGFNEPNG
NLQYQGKDITEFDFNFDQQTSQNI**C**NQLAELNATNIYTVLDKIKLNAKMNILIRDKRFHY
DRNNIAVGADESVVKEAHREVINSSTEGLLLNIDKDIRKILSGYIVEIEDTEGLKEVINDR
YDMLNISSLRQDGKTFIDFKKY**AA**KLPLYISNPNYKVNVYAVTKENTIINPSENGDTSTN
GIKKILIFSKKGYEIG

PA[N682A, D683A, K563C, F427A] (mPACA)

MEVKQENRLLNESESSSQGLLGYYFSDLNFQAPMVVTSSTTGDLSIPSSELENIPSENQYF
QSAIWSGFIKVKKSDEYTFATSADNHVTMWVDDQEVINKASNSNKIRLEKGRLYQIKIQ
YQRENPTEKGLDFKLYWTDSQNKKEVISSDNLQLPELKQKSSNS**RKKR**STSAGPTVPDR
DNDGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSTASDPYSDFEK
VTGRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRTH
TSEVHGNAEVHASFFDIGGSVSAGFSNSNSSTVAIDHSLSLAGERTWAETMGLNTADTA
RLNANIRYVNTGTAPIYNVLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYYPSKNLAPI
ALNAQDD**A**SSTPITMNYNQFLELEKTKQLRLDTDQVYGNIATYNFENGRVRVDTGSNW
SEVLPQIQETTARIIFNGKDLNLVERRIAAVNPSDPLETTKPDMTLKEALKIAFGFNEPNG
NLQYQGKDITEFDFNFDQQTSQNI**C**NQLAELNATNIYTVLDKIKLNAKMNILIRDKRFHY
DRNNIAVGADESVVKEAHREVINSSTEGLLLNIDKDIRKILSGYIVEIEDTEGLKEVINDR
YDMLNISSLRQDGKTFIDFKKY**AA**KLPLYISNPNYKVNVYAVTKENTIINPSENGDTSTN
GIKKILIFSKKGYEIG

*The Furin cleavage sites are highlighted in blue and the mutated residues are highlighted in red.

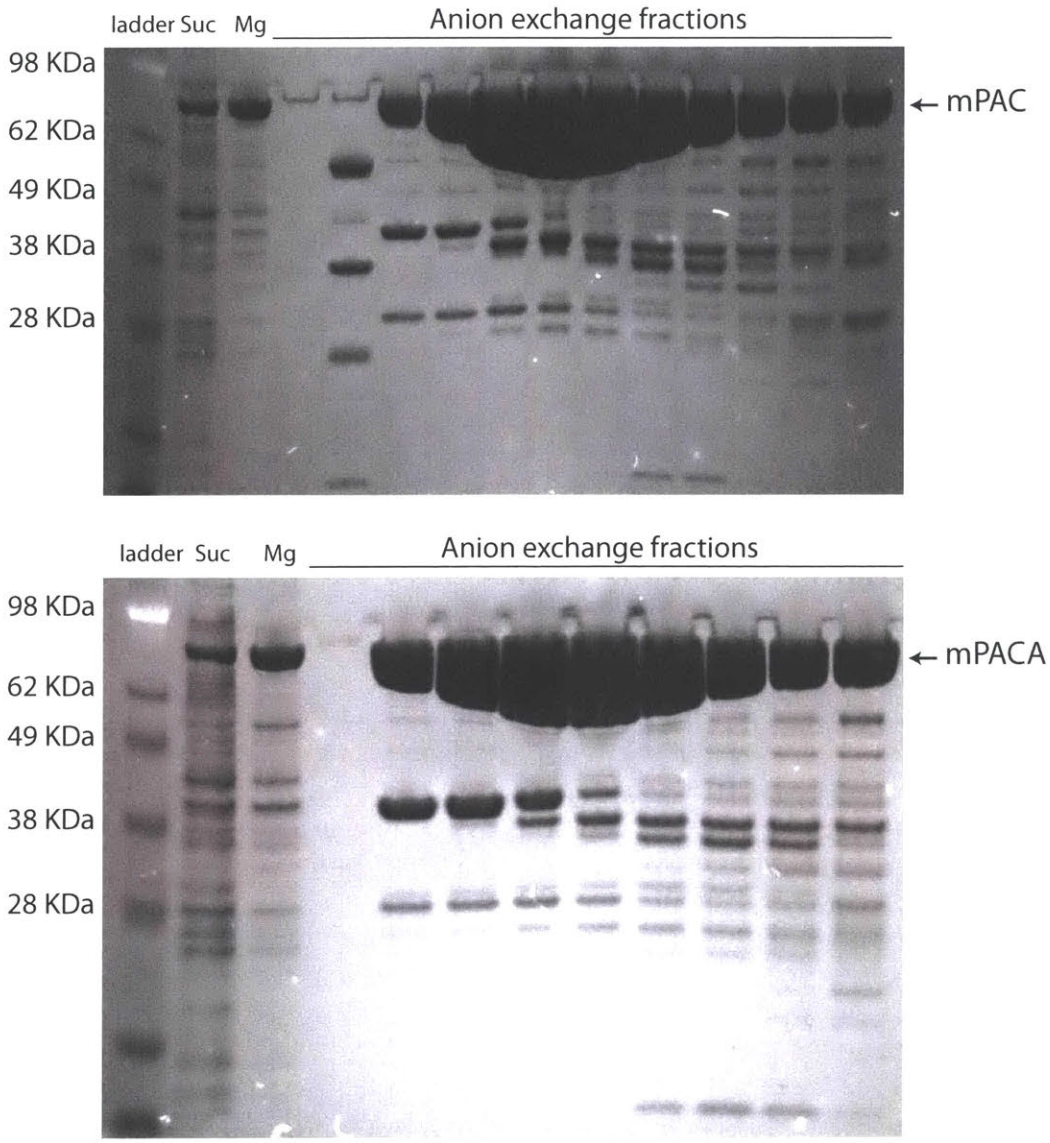

Figure 2.6.1. The SDS-PAGE analyses of anion exchange chromatography of mPAC (N682A, D683A, K563C), and mPACA (N682A, D683A, K563C, F427A). The gels were overloaded due to the high protein concentration. Suc is the supernatant from sucrose buffer resuspension. Mg is the supernatant from Mg buffer extraction.

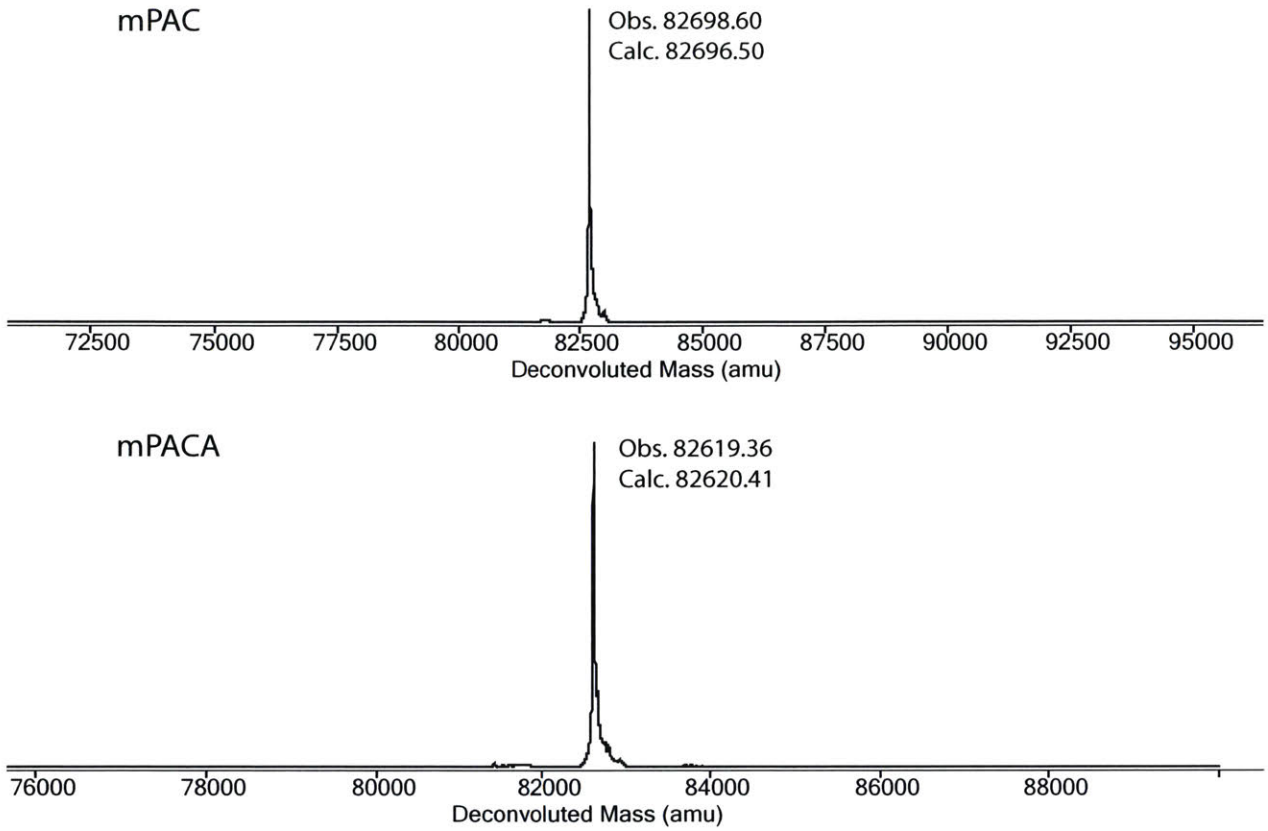

Figure 2.6.2. Deconvoluted mass traces of mPAC and mPACA from LC-MS. 200 ng of each protein was injected into Zorbax 300SB C_3 column with a method of 5-65% acetonitrile over 15 minutes. The TIC peak was integrated and the mass was deconvoluted using maximum entropy algorithm.

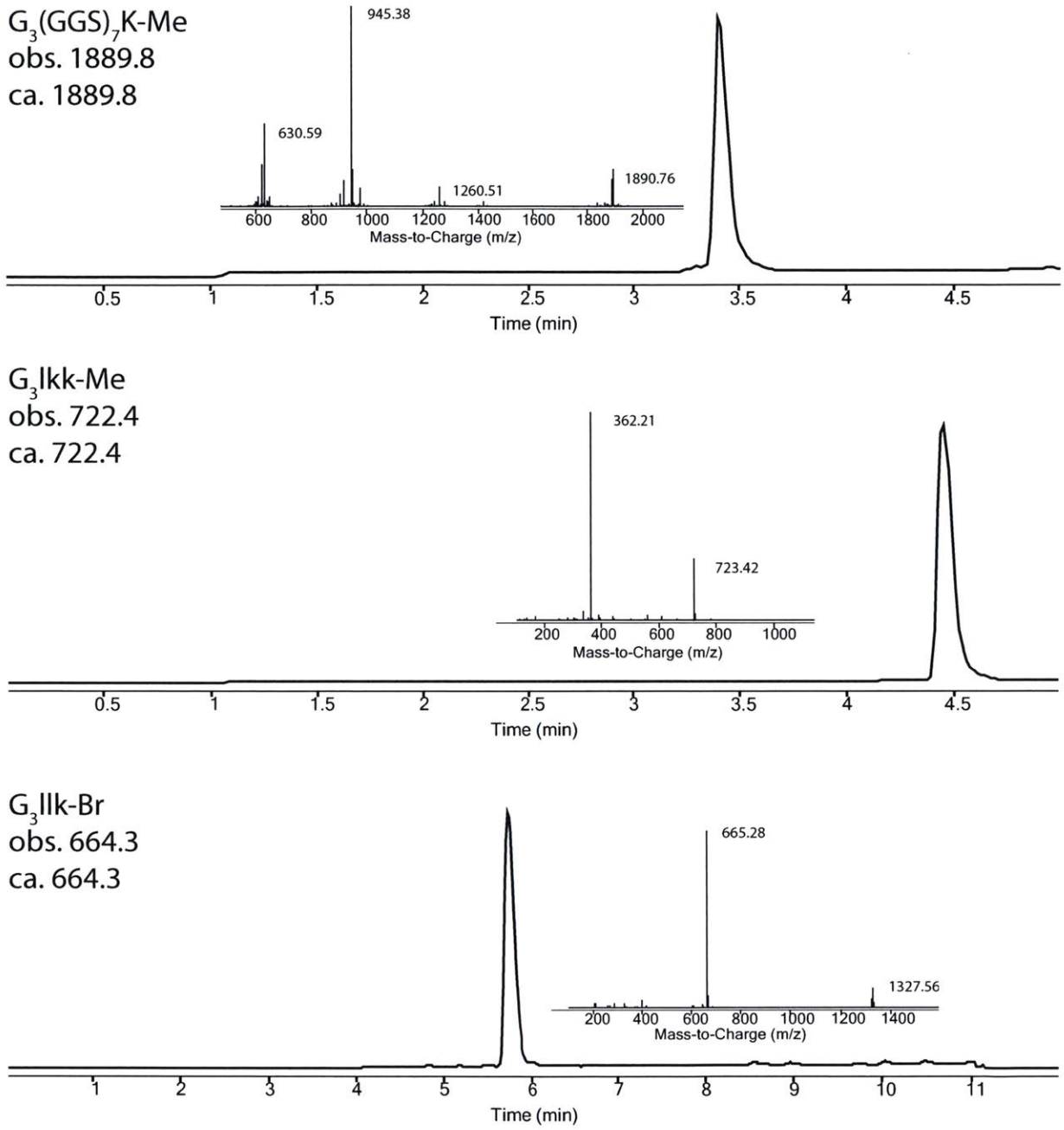

Figure 2.6.3. LC-MS traces of the peptide linkers used in this study. D-amino acids are indicated by the lowercase letters. Maleimide (Me) was coupled on the side chain of L- or D-lysine by treating the resin with 1.5 equivalents of N-γ-maleimidobutyryl-oxysuccinimide ester (GMBS, Thermo Fisher) dissolved in DMF for 3 h after lysine deprotection. Similarly, Br was coupled by incubating the resin with 0.4 M bromoacetic acid, 0.38 M HATU, and 0.38 M DIEA in DMF for 20 minutes.

2.6.2. Amino acid sequence of Trastuzumab

Trastuzumab-LC
ADIQMTQSPSSLSASVGDRVTITCRASQDVNTAVAWYQQKPGKAPKLLIYSASFLYSGV
PSRFSGSRSGTDFTLTISSLQPEDFATYYCQQHYTTPPTFGQGTKVEIKRTVAAPSVFIFPP
SDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSST
LTLSKADYEKHKV YACEVTHQGLSSPVTKSFNRGEC

Trastuzumab-HC
EVQLVESGGGLVQPGGSLRLSCAASGFNIKDTYIHWVRQAPGKGLEWVARIYPTNGYT
RYADSVKGRFTISADTSKNTAYLQMNSLRAEDTAVYYCSRWGGDGFYAMDYWGQGT
LVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFP
AVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTCPPCPA
PELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTK
PREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQV
YTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS
KLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK

Trastuzumab-HC-LPSTGG
EVQLVESGGGLVQPGGSLRLSCAASGFNIKDTYIHWVRQAPGKGLEWVARIYPTNGYT
RYADSVKGRFTISADTSKNTAYLQMNSLRAEDTAVYYCSRWGGDGFYAMDYWGQGT
LVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFP
AVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTCPPCPA
PELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTK
PREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQV
YTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS
KLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSP**LPSTG**GK

*The LPSTG insertion is highlighted in red.

2.6.3. Amino acid sequence of Cetuximab

Cetuximab-LC
DILLTQSPVILSVSPGERVSFSCRASQSIGTNIHWYQQRTNGSPRLLIKYASESISGIPSRFS
GSGSGTDFTLSINSVESEDIADYYCQQNNNWPTTFGAGTKLELKRTVAAPSVFIFPPSDE
QLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLS
KADYEKHKVYACEVTHQGLSSPVTKSFNRGEC

Cetuximab-HC
QVQLKQSGPGLVQPSQSLSITCTVSGFSLTNYGVHWVRQSPGKGLEWLGVIWSGGNTD
YNTPFTSRLSINKDNSKSQVFFKMNSLQSNDTAIYYCARALTYYDYEFAYWGQGTLVT
VSAASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAV
LQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTCPPCPAPE
LLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPR
EEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTL
PPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLT
VDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK

Cetuximab-HC-LPSTGG
QVQLKQSGPGLVQPSQSLSITCTVSGFSLTNYGVHWVRQSPGKGLEWLGVIWSGGNTD
YNTPFTSRLSINKDNSKSQVFFKMNSLQSNDTAIYYCARALTYYDYEFAYWGQGTLVT
VSAASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAV
LQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTCPPCPAPE
LLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPR
EEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTL
PPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLT
VDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSP**LPSTG**GK

*The LPSTG insertion is highlighted in red.

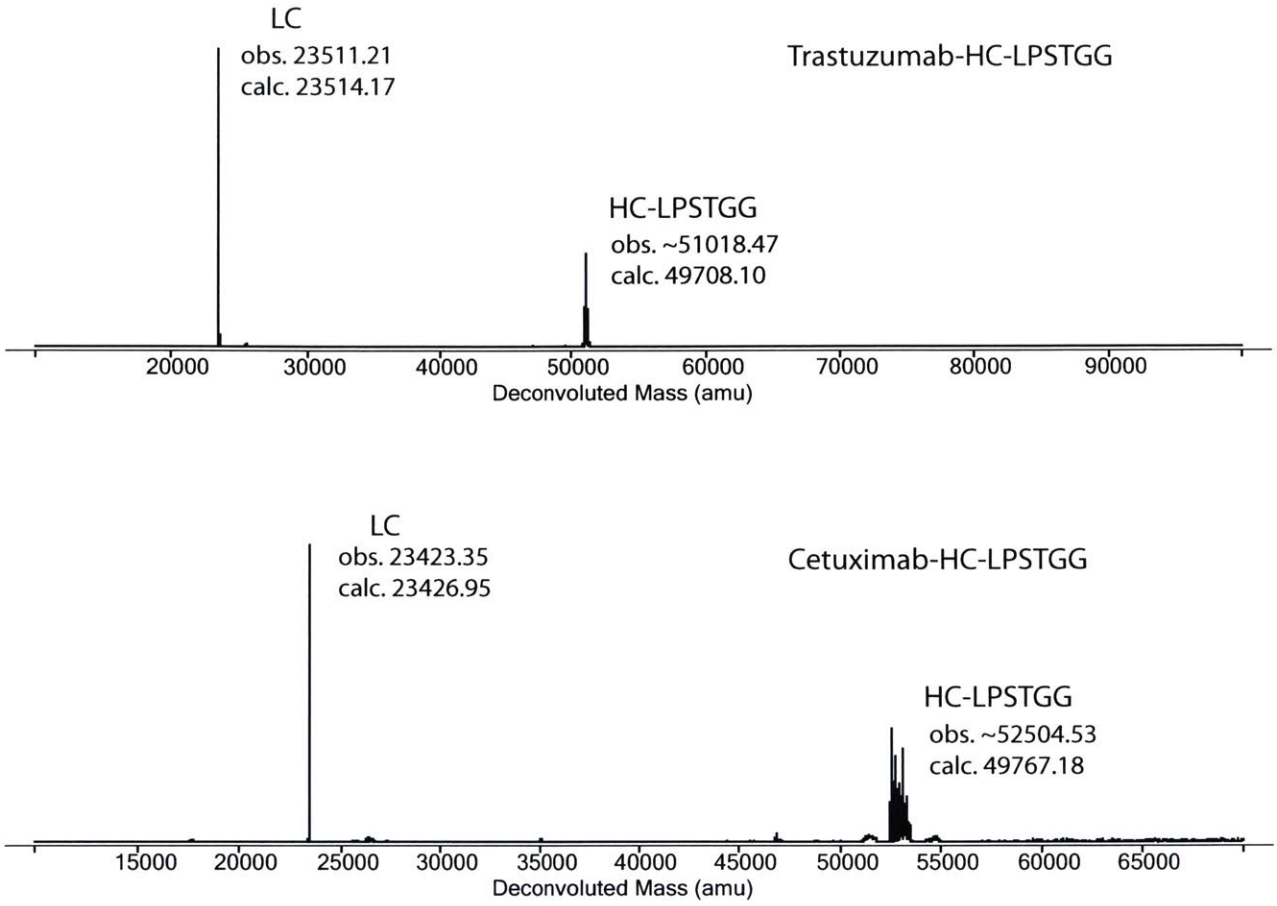

Figure 2.6.4. The LC-MS deconvoluted masses of the light chains (LCs) and heavy chains (HCs) of Trastuzumab-HC-LPSTGG and Cetuximab-HC-LPSTGG. The IgGs were reduced by 20 mM TCEP in PBS at 37°C for 20 minutes before LC-MS analysis. The mass difference and heterogeneity of the HCs was due to post translational modification.

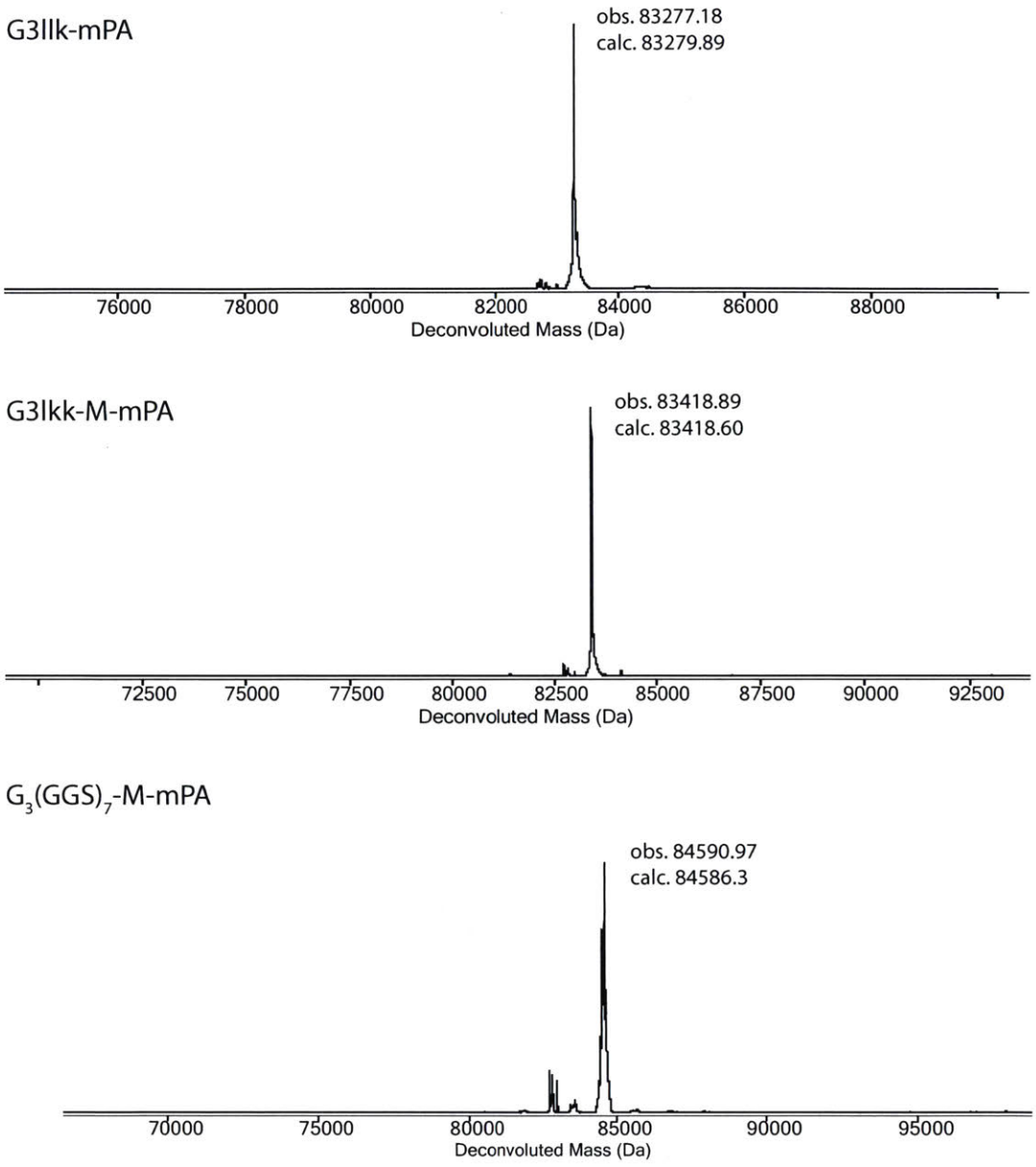

Figure 2.6.5. LC-MS deconvoluted masses of mPAs conjugated with different peptide linkers. The conjugations were performed as described in the methods section. 200 ng of each protein was injected into Zorbax 300SB C_3 column with a method of 5-65% acetonitrile over 15 minutes. The TIC peak was integrated and the mass was deconvoluted using maximum entropy algorithm.

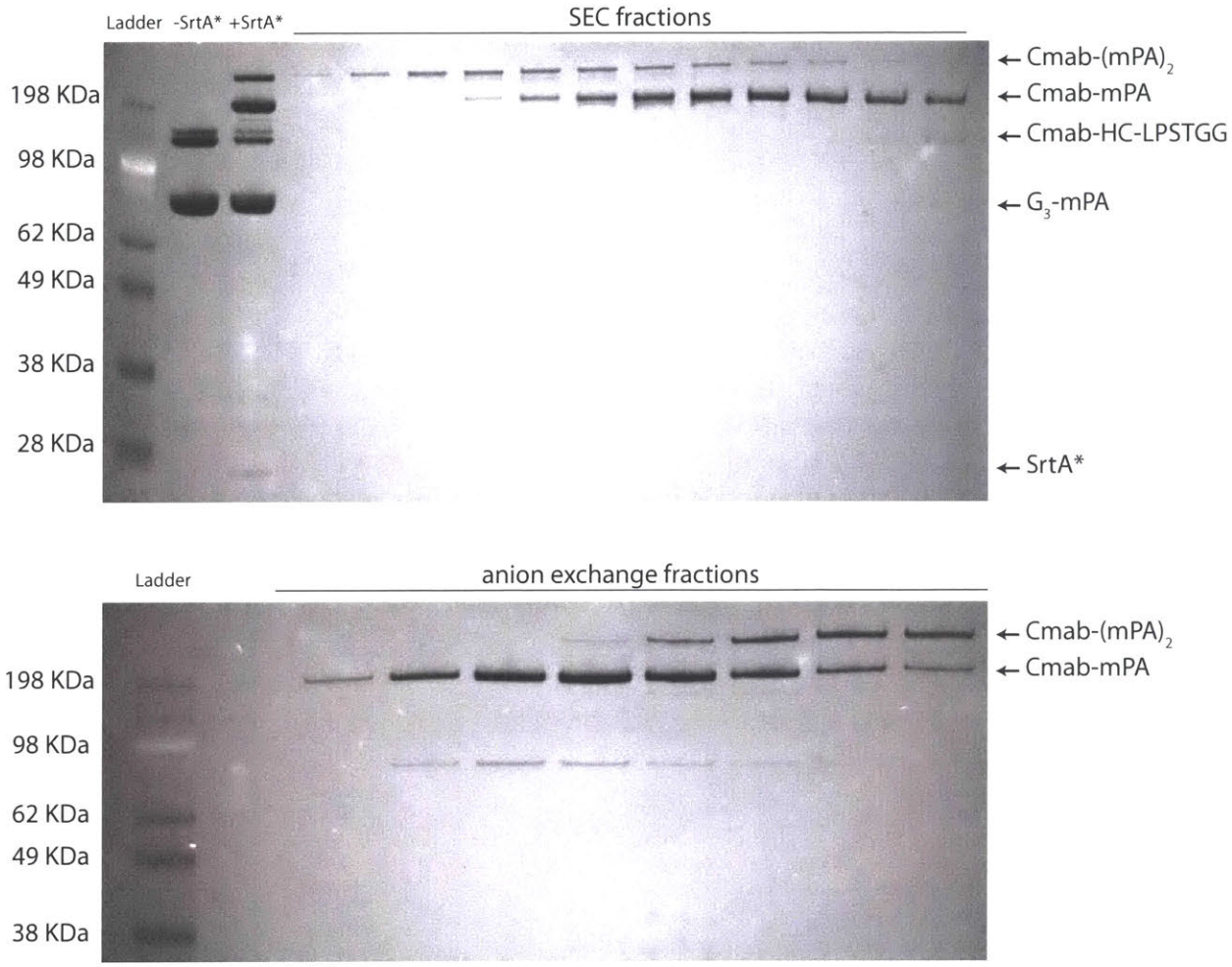

Figure 2.6.6. SDS-PAGE analysis of Cmab-mPA prepared as described in the methods section. In short, Cmab-HC-LPSTGG was reacted with three equivalents of G_3-mPA in the presence of SrtA*. After 1 h at RT, the reaction mixture was sequentially purified by size exclusion and anion exchange chromatography to give pure Cmab-mPA.

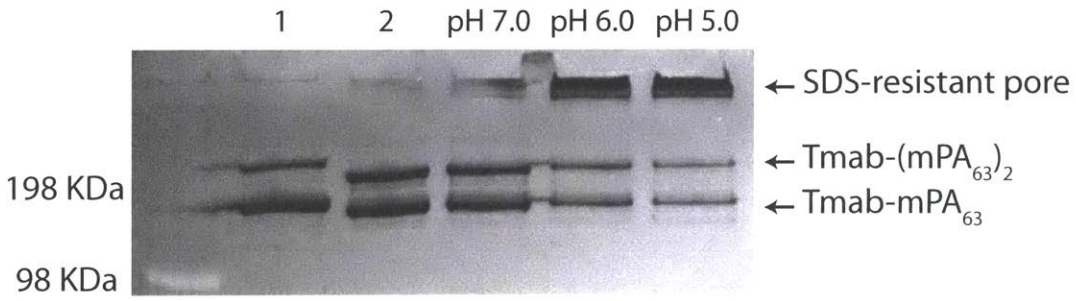

Figure 2.6.7. In vitro SDS-resistant pore formation assay analysed by SDS-PAGE. The SDS-resistant pore formation assay was performed as described previously. Mixture of Tmab-$(mPA)_2$ and Tmab-mPA (lane 1) was first nicked by Trypsin and purified by anion exchange chromatography (lane 2). The resulted mixture of Tmab-$(mPA_{63})_2$ and Tmab-mPA_{63} were incubated with Tris or MES buffers at different pHs for 30 minutes at RT and then resolved by SDS-PAGE. Protein bands were visualized by Coomassie blue staining.

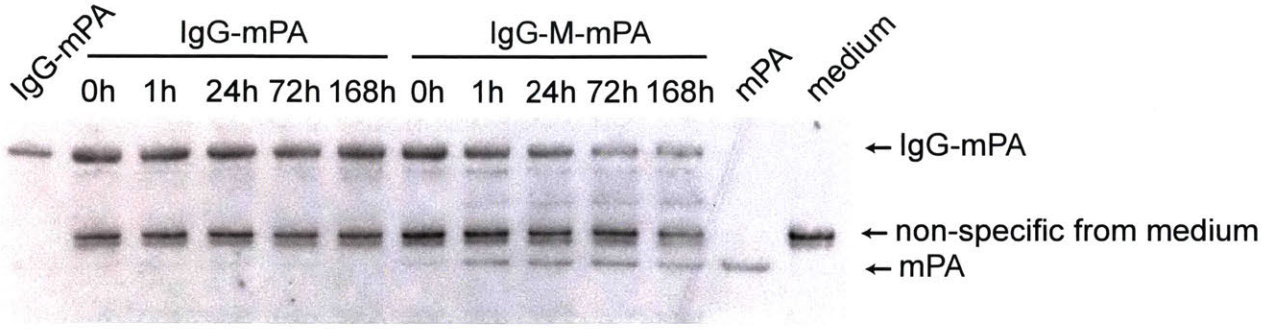

Figure 2.6.8. Stability comparison between IgG mPA conjugates with different linkers. Antibody mPA conjugates prepared with bromoacetamide linker (IgG-mPA) or maleimide linker (IgG-M-mPA) was incubated in MEM with 10% fetal bovine serum at 37°C for a week. Samples were taken from indicated time points and analyzed by anti-PA western blot. IgG-M-mPA showed significant degradation whereas IgG-mPA mostly remained intact. Trastuzumab is used in this experiment.

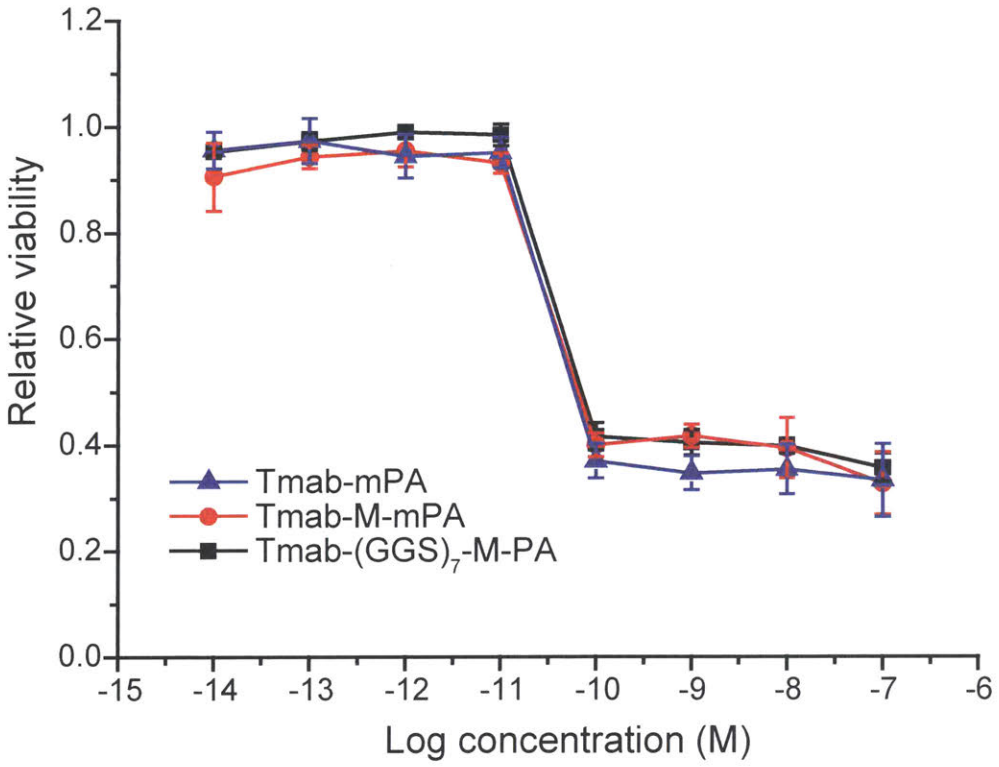

Figure 2.6.9. Activity comparison between Tmab-mPA, Tmab-M-mPA, Tmab-$(GGS)_7$-M-mPA. BT474 cells were treated with 10 nM LD and different Tmab-mPA variants for 72 h before being subject to CellTiter-Glo. The relative viability is normalized to no treatment. Each data point is a triplicate.

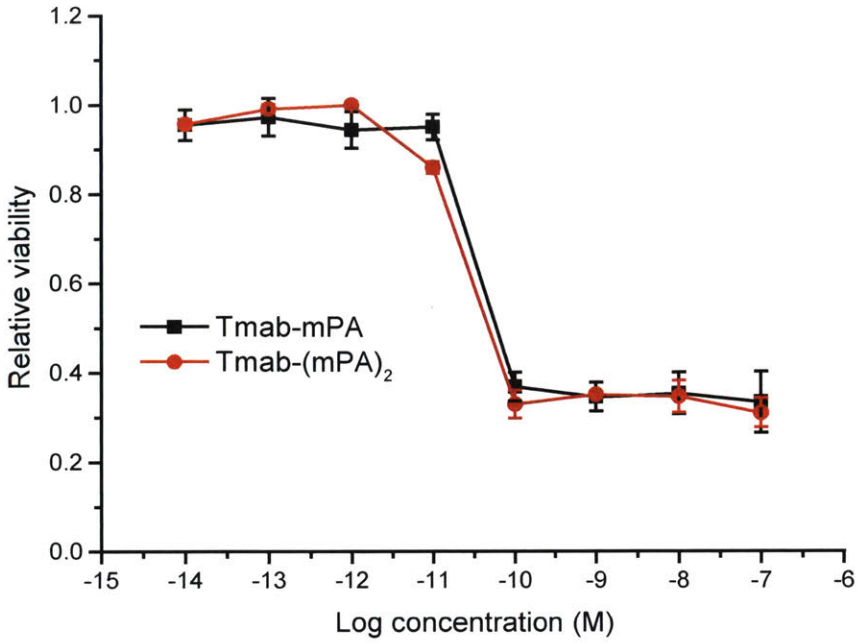

Figure 2.6.10. Activity comparison between Tmab-mPA and Tmab-$(mPA)_2$ in the presence of 10 nM LD.

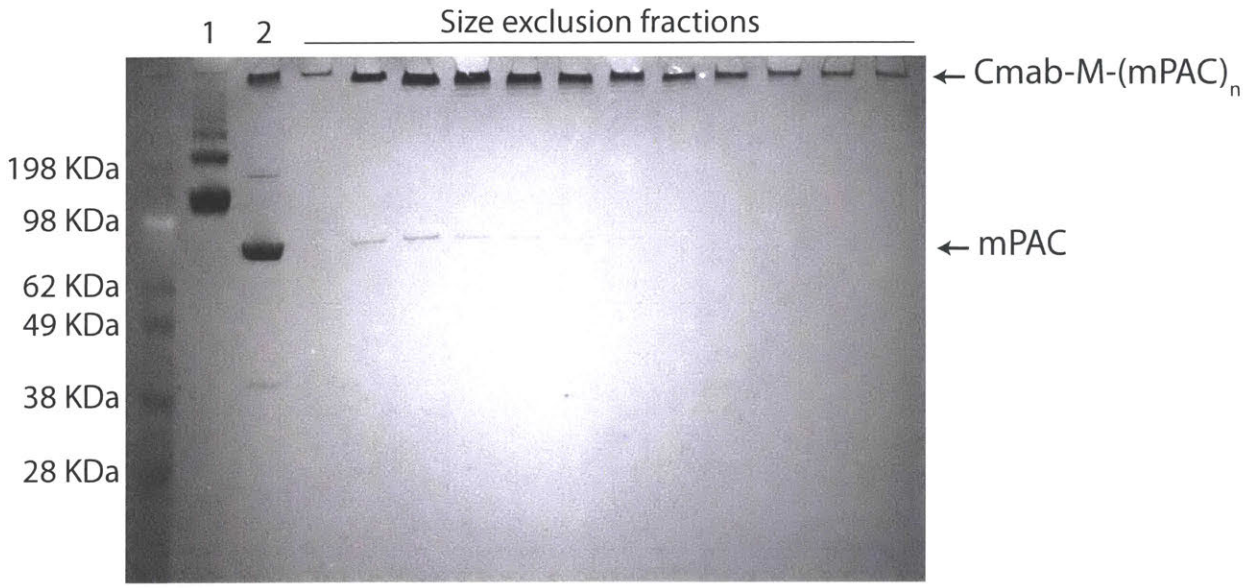

Figure 2.6.11. SEC purification of Cmab-M-(mPAC)$_n$ analyzed by SDS-PAGE. Cmab was first reacted with 50 equivalents of SM(PEG)4 (Thermo Fisher) dissolved in DMSO for 1 h at RT (lane 1). The extra linker was then removed by buffer exchange. The resulted Cmab was incubated with 40 equivalents of mPAC for 1 h at RT (lane 2) before separation by size exclusion. The average number of mPA per antibody was calculated to be about 7 based on final yield of the reaction assuming the antibody is 100% recovered after purification.

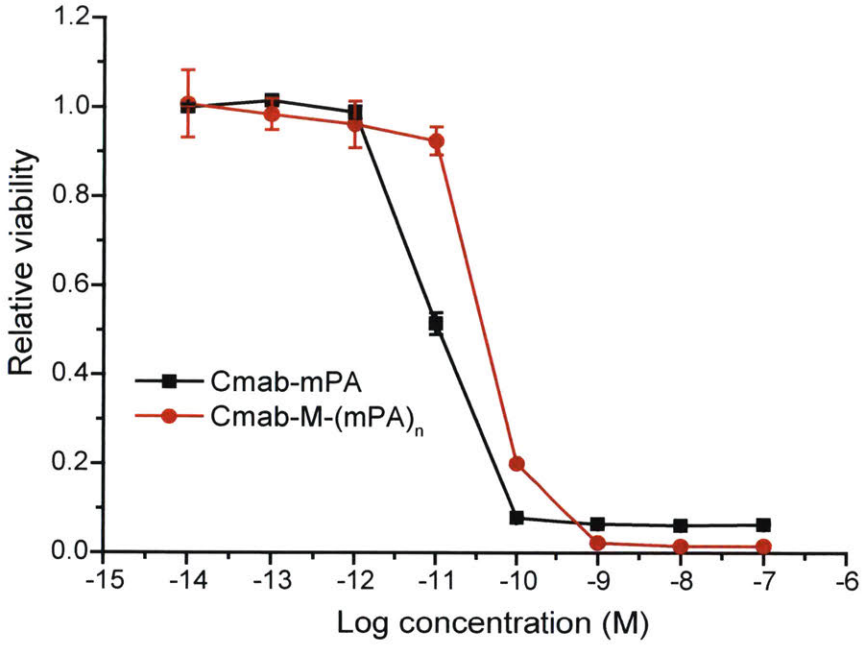

Figure 2.6.12. Activity comparison between Cmab-mPA and Cmab-M-(mPA)$_n$. A549 cells were treated with both proteins at the indicated concentrations in the presence of 10 nM LD for 72 h. The cell vability was measured by CellTiter-Glo and normalized to no treatment group. Each data point is a triplicate.

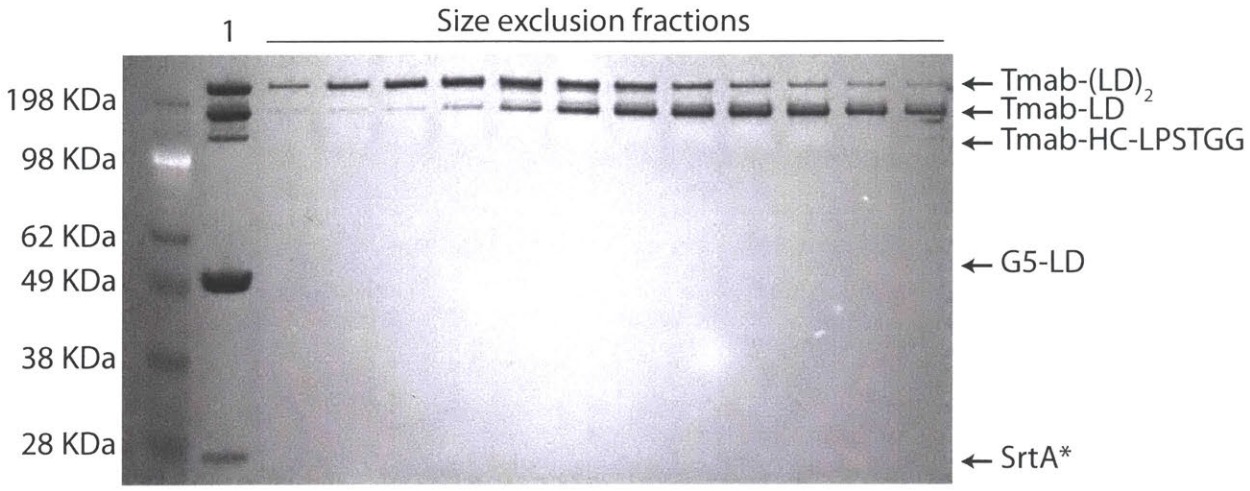

Figure 2.6.13. SDS-PAGE analysis of Tmab-LD purified by SEC. Tmab-LD was prepared similarly to IgG-mPA. Tmab-HC-LPSTGG was reacted with 5 equivalents of G_5-LD in the presence of SrtA* (lane 1) for 30 minutes at RT before purification by size exclusion chromatography.

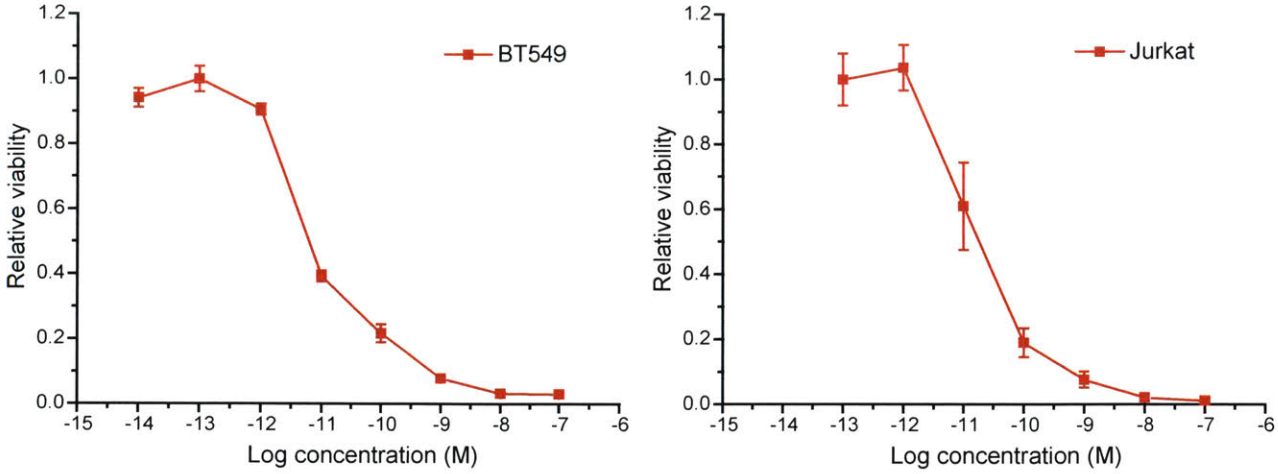

Figure 2.6.14. Viability of BT549 and Jurkat cells treated with WT PA in the presence of LD. Cells were treated with 10-fold serial dilutions of wild-type PA in the presence of 10 nM of LD. Viability was measured after 72 h by CellTiter-Glo as described in the methods section.

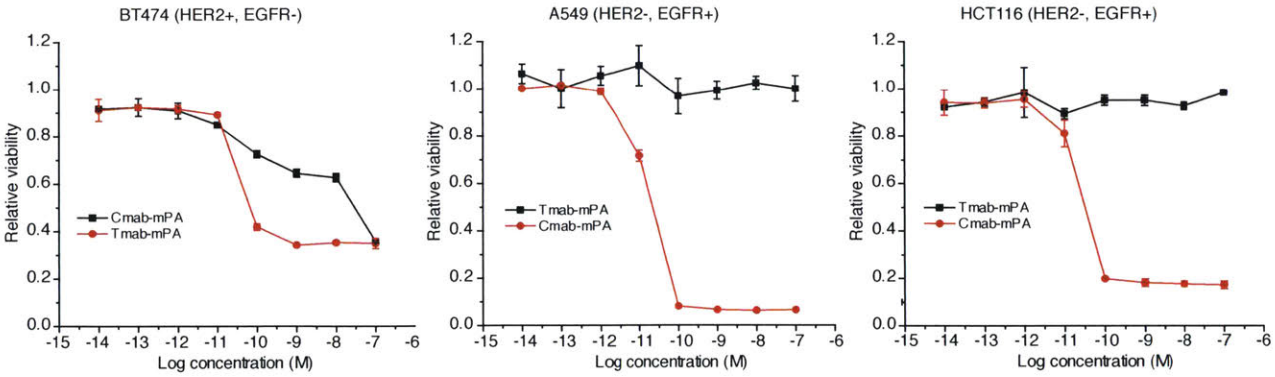

Figure 2.6.15. Different responses to Cmab-mPA and Tmab-mPA in the presence of 10 nM of LD in different cell lines. The cells were treated with 10-fold serial dilutions of Cmab-mPA or Tmab-mPA with 10 nM of LD. Viability was measured by CellTiter-Glo as described in the methods section.

2.6.4. Amino acid sequence of mPAC-uPA and mPAC-MMP

mPAC-uPA

MEVKQENRLLNESESSSQGLLGYYFSDLNFQAPMVVTSSTTGDLSIPSSELENIPSENQYF
QSAIWSGFIKVKKSDEYTFATSADNHVTMWVDDQEVINKASNSNKIRLEKGRLYQIKIQ
YQRENPTEKGLDFKLYWTDSQNKKEVISSDNLQLPELKQKSSNS**PGSGRSA**STSAGPTV
PDRDNDGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSTASDPYSD
FEKVTGRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTS
RTHTSEVHGNAEVHASFFDIGGSVSAGFSNSNSSTVAIDHSLSLAGERTWAETMGLNTA
DTARLNANIRYVNTGTAPIYNVLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYYPSKN
LAPIALNAQDDFSSTPITMNYNQFLELEKTKQLRLDTDQVYGNIATYNFENGRVRVDTG
SNWSEVLPQIQETTARIIFNGKDLNLVERRIAAVNPSDPLETTKPDMTLKEALKIAFGFNE
PNGNLQYQGKDITEFDFNFDQQTSQNI**C**NQLAELNATNIYTVLDKIKLNAKMNILIRDKR
FHYDRNNIAVGADESVVKEAHREVINSSTEGLLLNIDKDIRKILSGYIVEIEDTEGLKEVIN
DRYDMLNISSLRQDGKTFIDFKKY**AA**KLPLYISNPNYKVNVYAVTKENTIINPSENGDTS
TNGIKKILIFSKKGYEIG

mPAC-MMP

MEVKQENRLLNESESSSQGLLGYYFSDLNFQAPMVVTSSTTGDLSIPSSELENIPSENQYF
QSAIWSGFIKVKKSDEYTFATSADNHVTMWVDDQEVINKASNSNKIRLEKGRLYQIKIQ
YQRENPTEKGLDFKLYWTDSQNKKEVISSDNLQLPELKQKSSNS**GPLGMLSQ**STSAGPT
VPDRDNDGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSTASDPYS
DFEKVTGRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTST
SRTHTSEVHGNAEVHASFFDIGGSVSAGFSNSNSSTVAIDHSLSLAGERTWAETMGLNT
ADTARLNANIRYVNTGTAPIYNVLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYYPSK
NLAPIALNAQDDFSSTPITMNYNQFLELEKTKQLRLDTDQVYGNIATYNFENGRVRVDT
GSNWSEVLPQIQETTARIIFNGKDLNLVERRIAAVNPSDPLETTKPDMTLKEALKIAFGFN
EPNGNLQYQGKDITEFDFNFDQQTSQNI**C**NQLAELNATNIYTVLDKIKLNAKMNILIRDK
RFHYDRNNIAVGADESVVKEAHREVINSSTEGLLLNIDKDIRKILSGYIVEIEDTEGLKEVI
NDRYDMLNISSLRQDGKTFIDFKKY**AA**KLPLYISNPNYKVNVYAVTKENTIINPSENGDT
STNGIKKILIFSKKGYEIG

*The Furin cleavage sites are replaced by uPA and MMP substrate sequences as highlighted in blue and the mutated residues are highlighted in red.

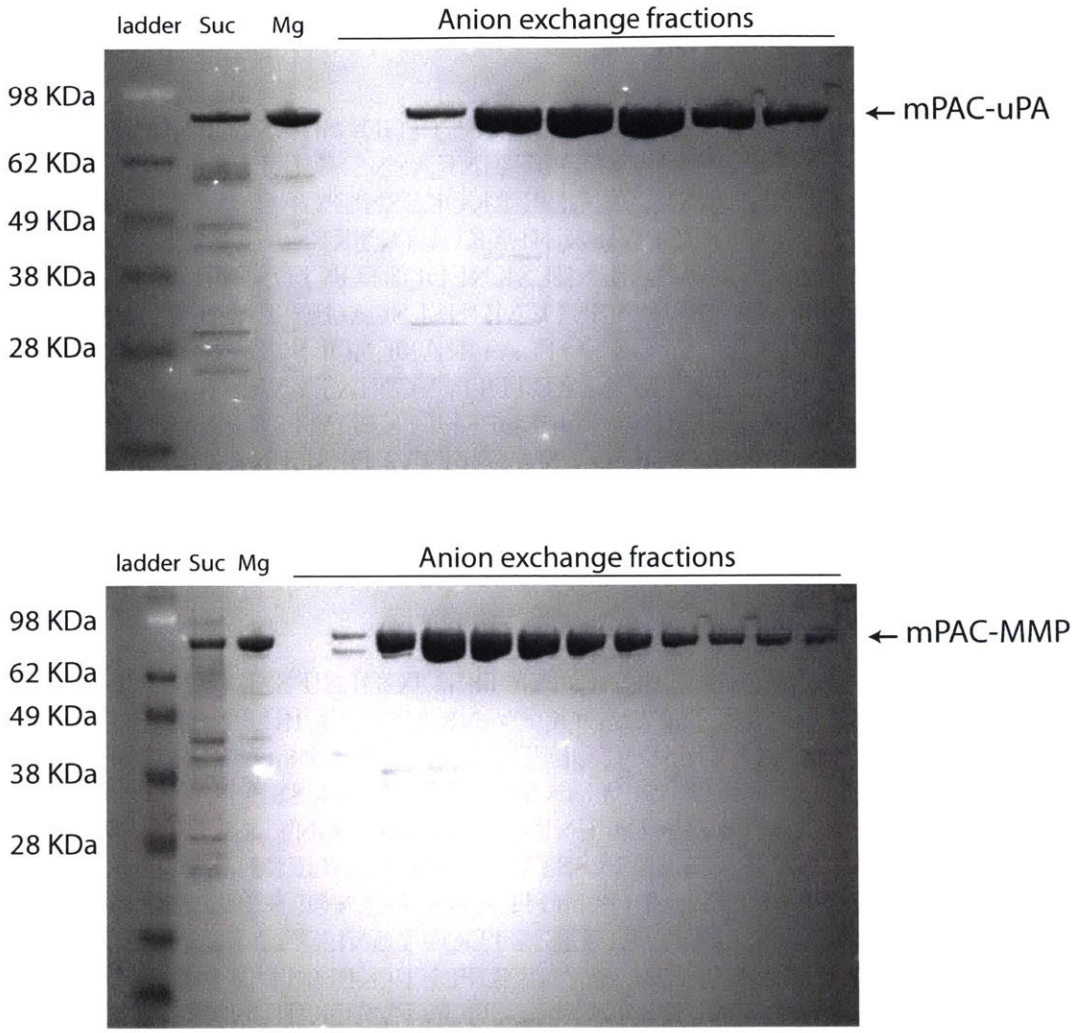

Figure 2.6.16. SDS-PAGE analysis of mPAC-uPA and mPAC-MMP purified by anion exchange chromatography.

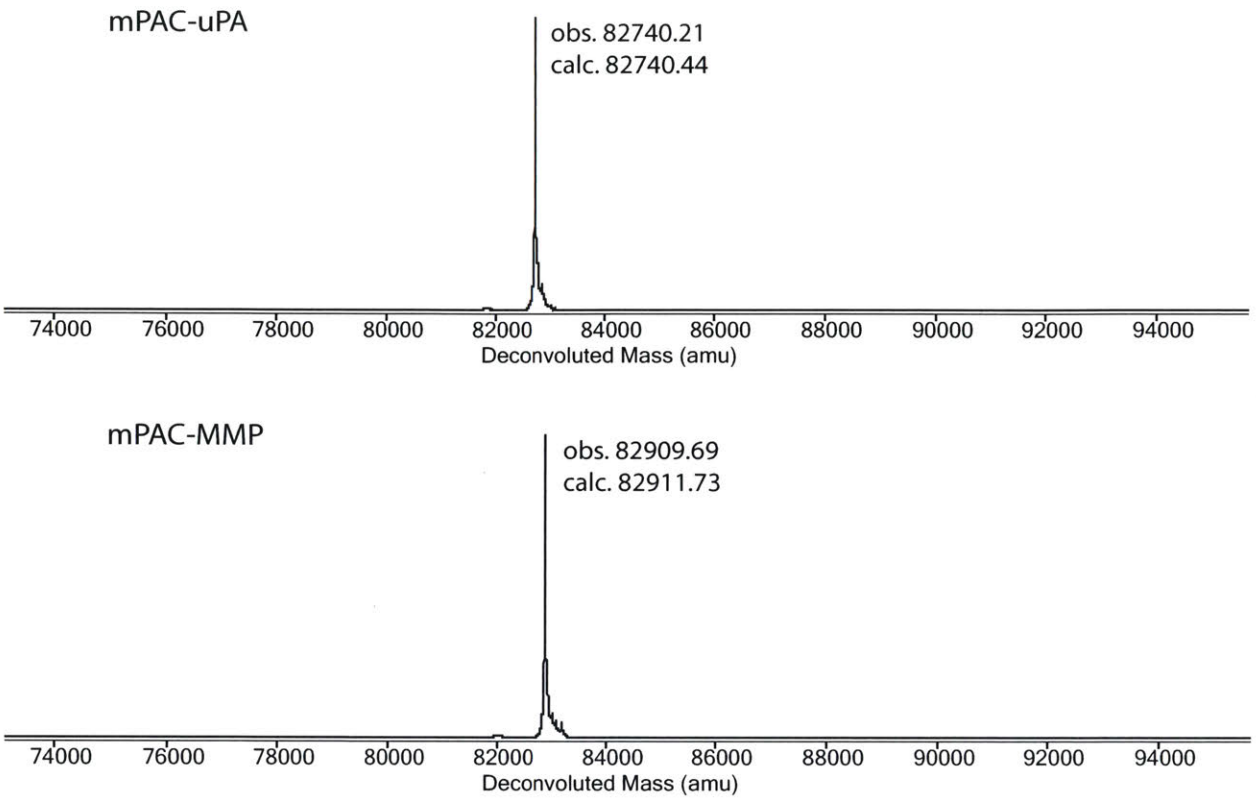

Figure 2.6.17. Deconvoluted mass traces of mPAC-uPA and mPAC-MMP from LC-MS. 200 ng of each protein was injected into Zorbax 300SB C_3 column with a method of 5-65% acetonitrile over 15 minutes. The TIC peak was integrated and the mass was deconvoluted using maximum entropy algorithm.

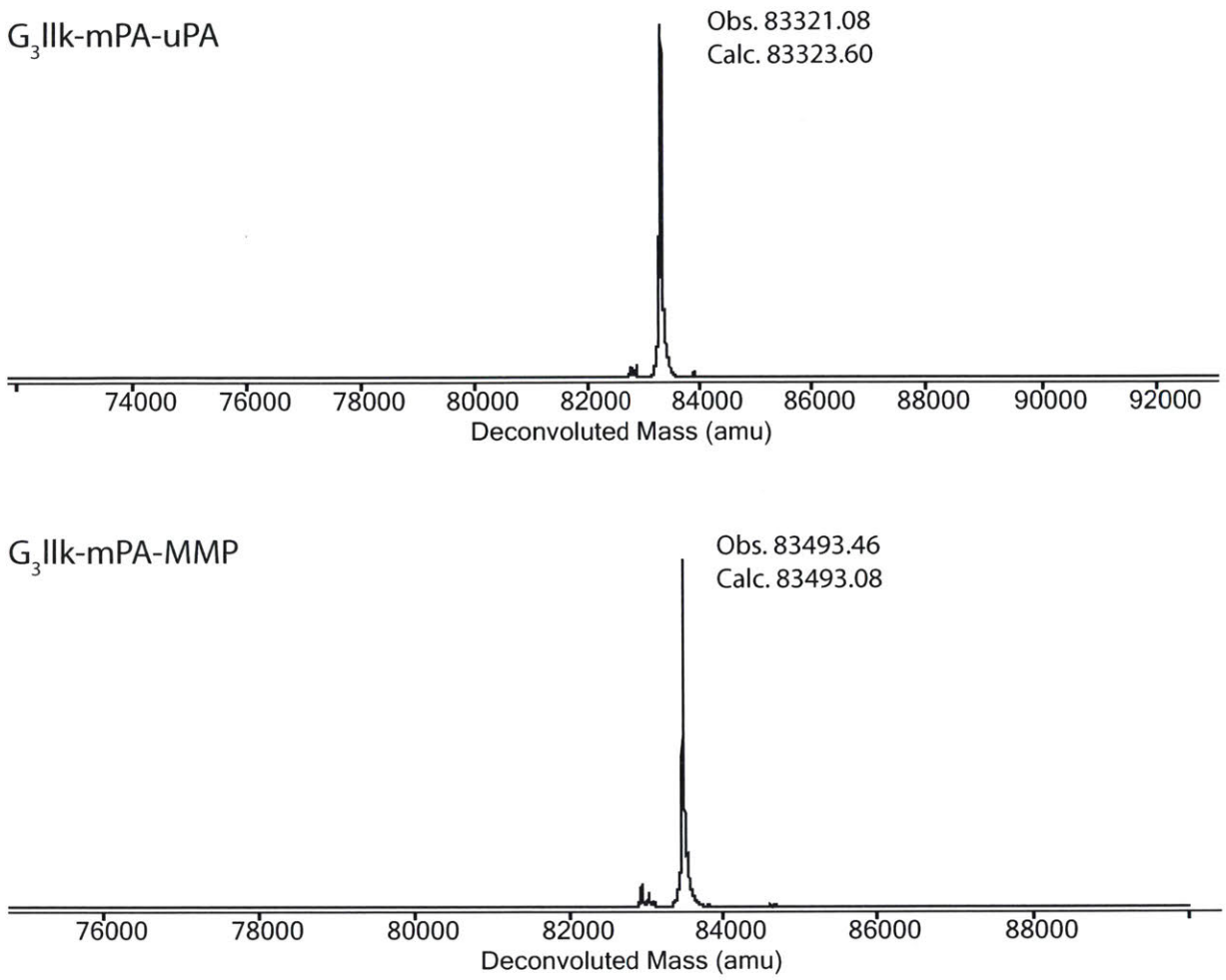

Figure 2.6.18. Deconvoluted masses of G$_3$llk-mPAC-uPA and G$_3$llk-mPAC-MMP. G$_3$llk-Br conjugation to mPAC variants was performed as described in the methods section.

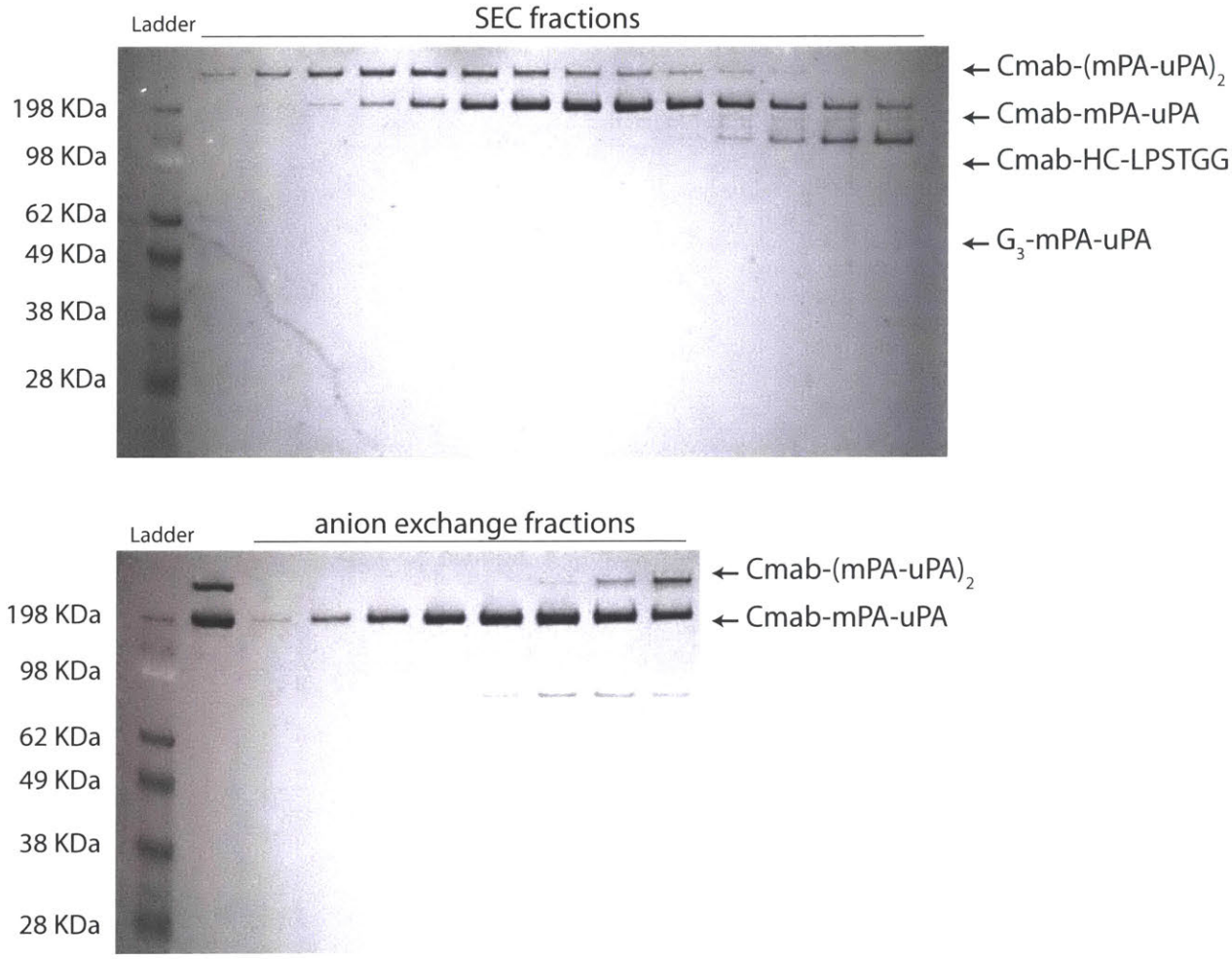

Figure 2.6.19. SDS-PAGE analysis of Cmab-mPA-uPA purification. Cmab-mPA-uPA was prepared as described in the methods section. In short, Cmab-HC-LPSTGG was reacted with three equivalents of G_3-mPA-uPA in the presence of SrtA*. After 1 h at RT, the reaction mixture was sequentially purified by size exclusion and anion exchange chromatography to give pure Cmab-mPA-uPA.

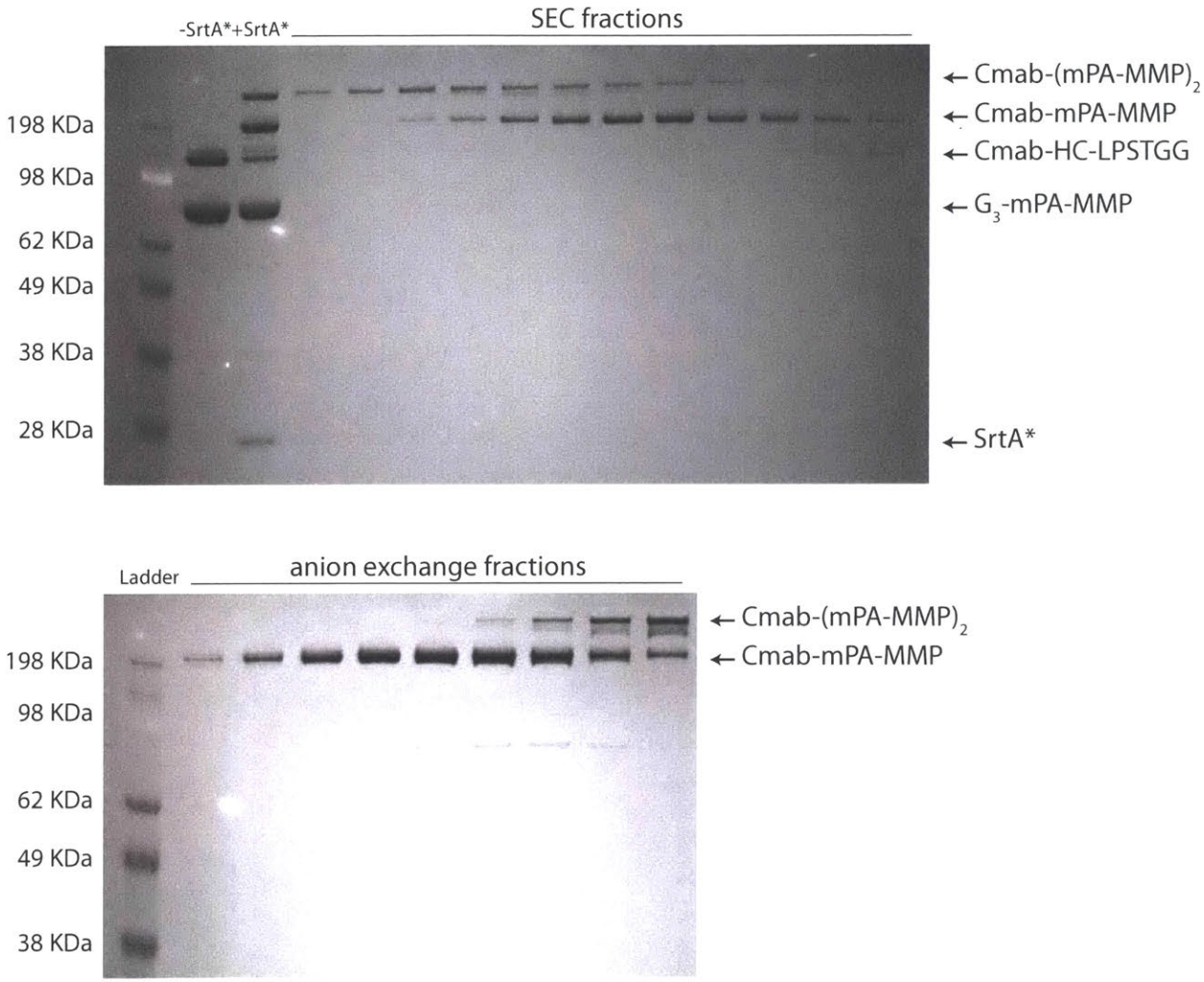

Figure 2.6.20. SDS-PAGE analysis of Cmab-mPA-MMP purification. Cmab-mPA-MMP was prepared as described in the methods section. In short, Cmab-HC-LPSTGG was reacted with three equivalents of G_3-mPA-MMP in the presence of SrtA*. After 1 h at RT, the reaction mixture was sequentially purified by size exclusion and anion exchange chromatography to give pure Cmab-mPA-MMP.

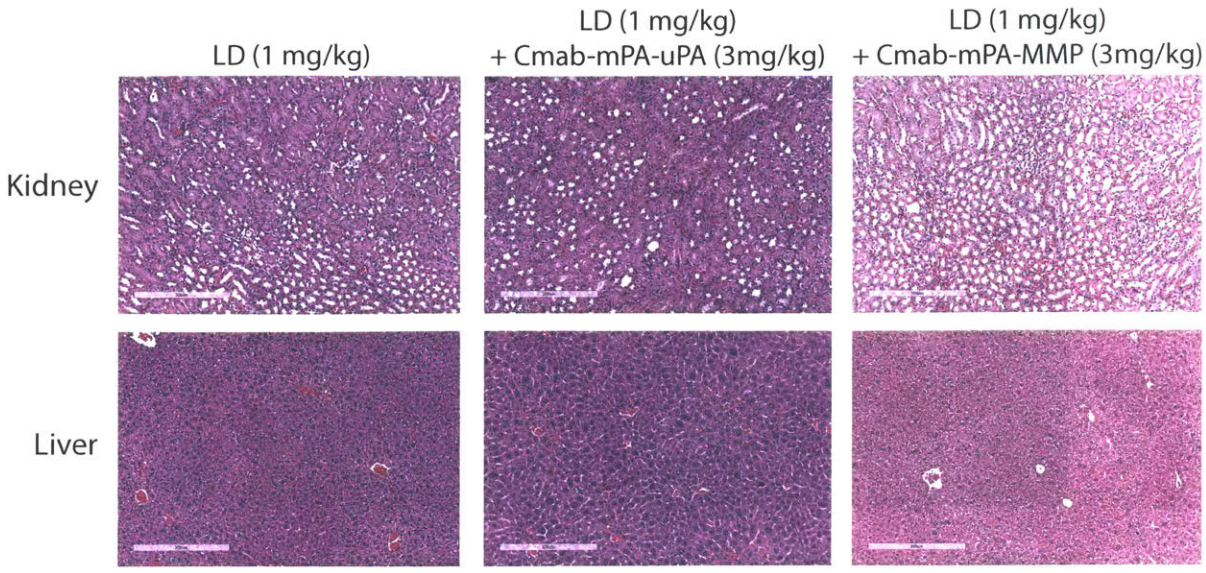

Figure 2.6.21. Histological analysis of kidney and liver tissues from mice injected with 1 mg/kg of LD alone or in combination with Cmab-mPA-uPA or Cmab-mPA-MMP (1 mg/kg). The tissue samples were prepared as described in the methods section.

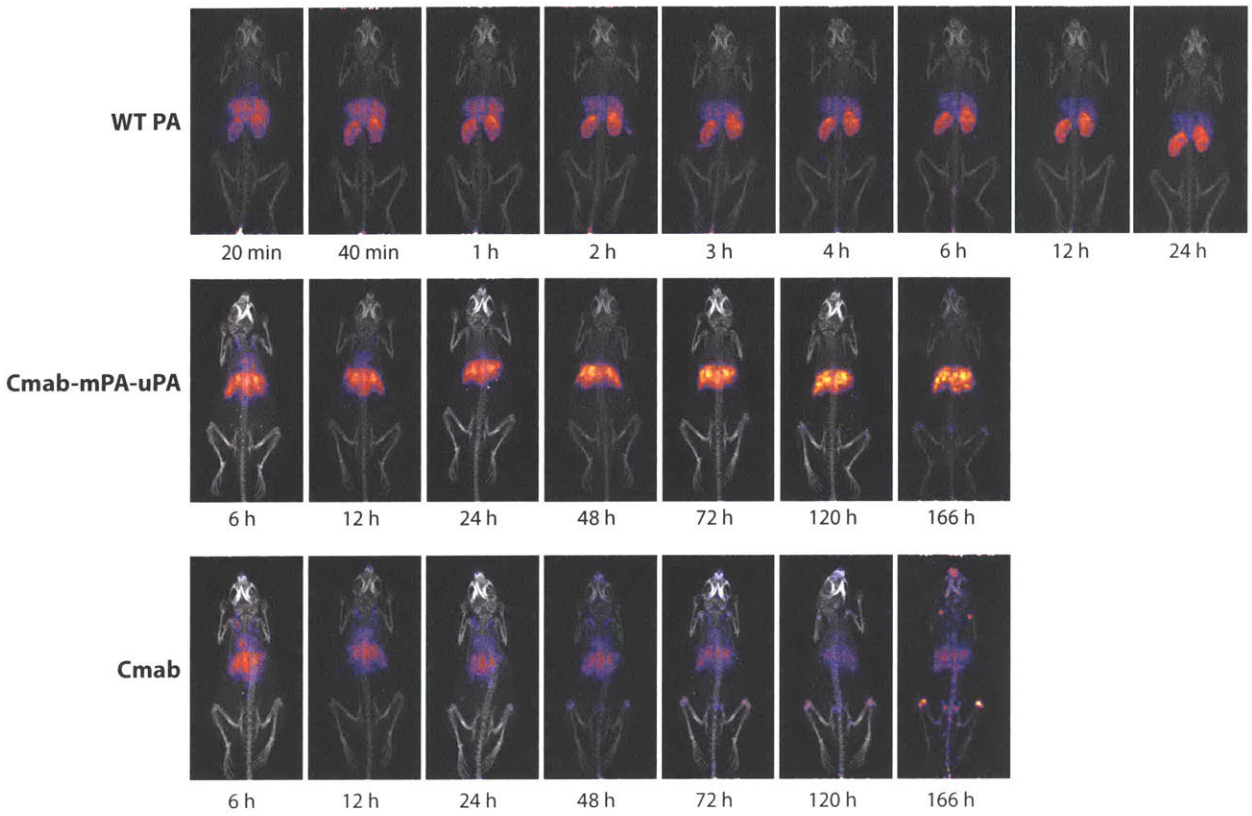

Figure 2.6.22. Change of biodistribution of different proteins over time. Whole-body PET images were acquired as described in the methods section. Shorter time scale was used for wild-type PA due to its rapid clearance time.

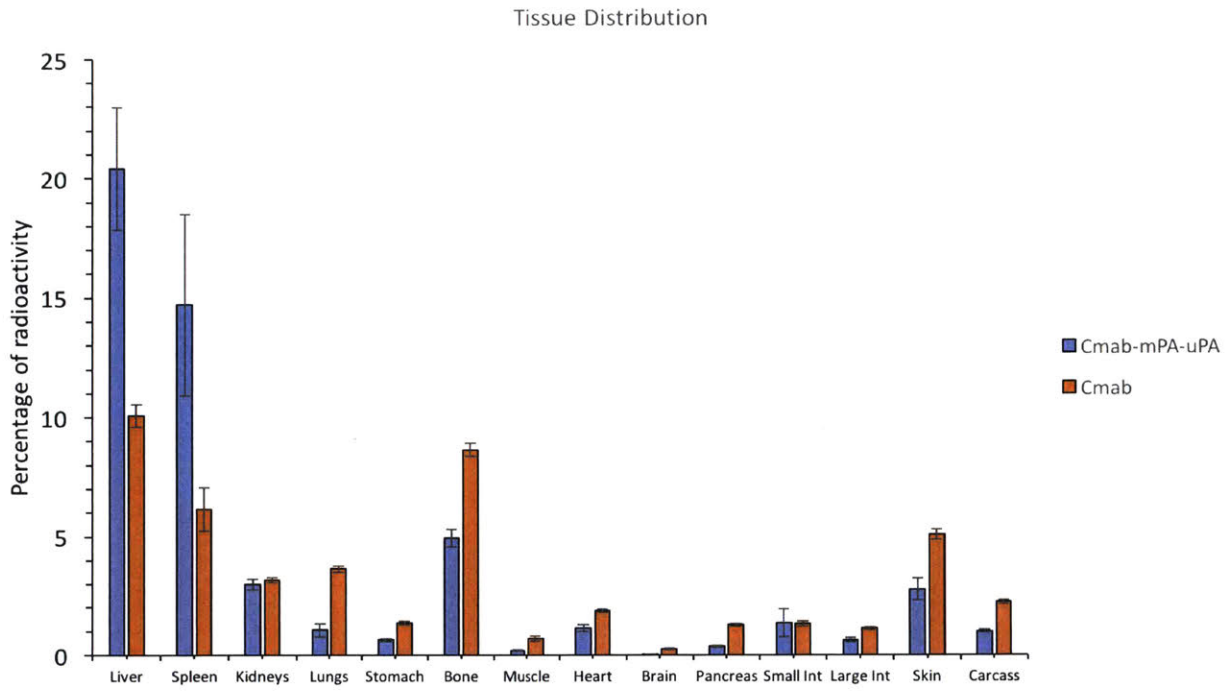

Figure 2.6.23. Tissue distribution of Cmab and Cmab-mPA-uPA 166 h post injection. Tissues from different organs were measured for radioactivity as described in the methods section.

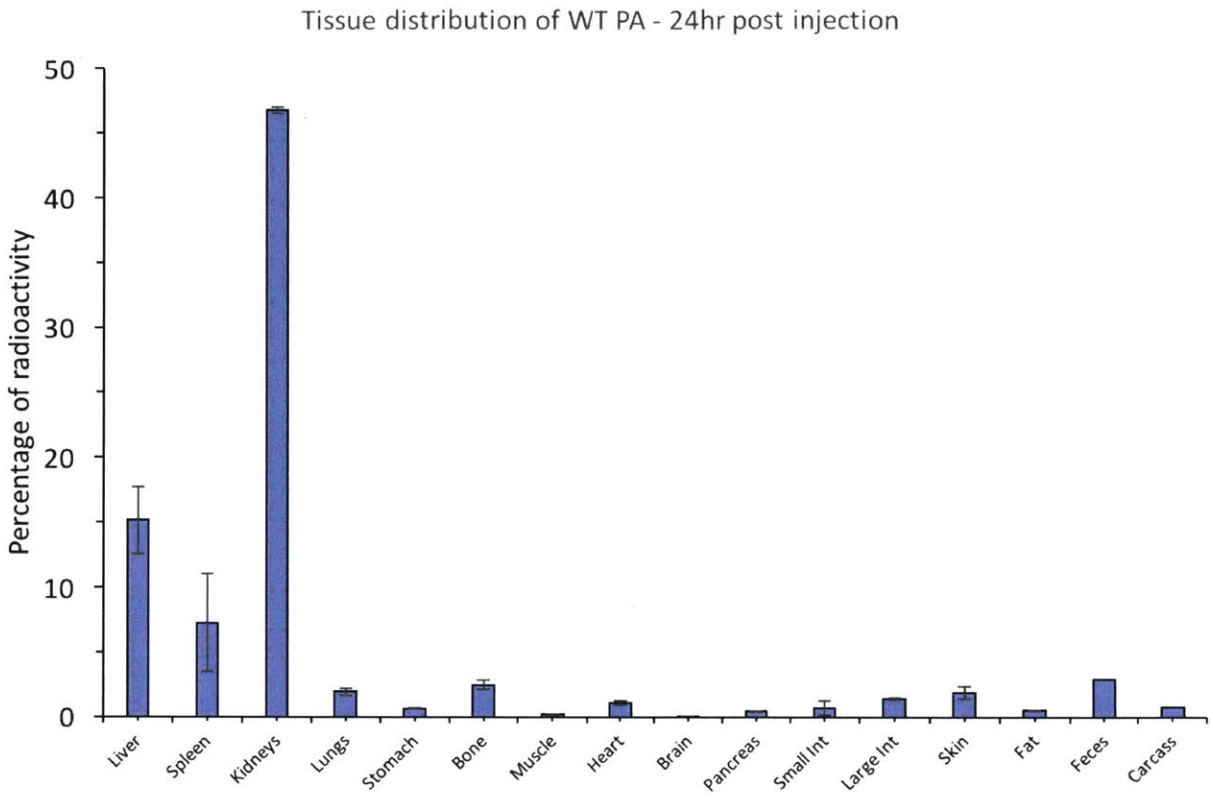

Figure 2.6.24. Tissue distribution of WT PA 24 h post injection. Tissues from different organs were measured for radioactivity as described in the methods section.

2.6.5. Targeting of activated T cells by LD delivered by Daclizumab-mPA

Great progress has been made in taking advantage of immune system to suppress tumor in the past decade. But at the same time, new challenge has emerged that it is difficult to effectively evoke autoimmunity. Accumulating evidence indicates depleting Treg cells can elicit strong anti-tumor immune response. Here we report a novel approach to target and remove Treg cells by using our IgG-mPA system directed by Daclizumab (Dmab), an IgG molecule that binds to CD25, the alpha subunit of the IL-2 receptor of Treg cells. We investigated the ability of Dmab-mPA to deliver LD into activated Jurkat cells. Compared with Dmab only, Dmab-mPA with LD was able to significantly reduce viability of Jurkat cells. Surprisingly, Dmab-mPA itself was also toxic to activated Jurkat cells. One explanation is that the oligomerization of CD25 mediated by Dmab-mPA on cell surface might have an effect on the viability of the Jurkat cells. Another possibility is that the pore-forming Dmab-mPA significantly permeablized the endosome and caused toxicity.

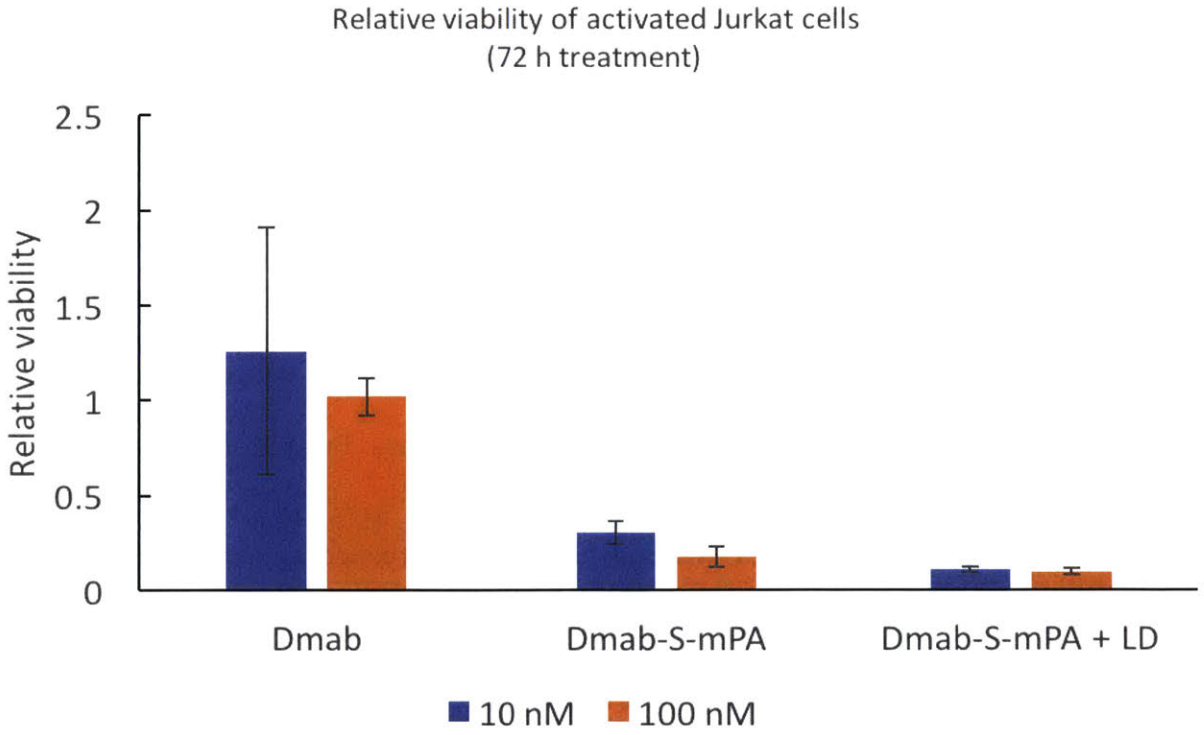

Figure 2.6.25. Relative viability of activated Jurkat cells treated with Dmab-mPA + LD and other controls. Jurkat cells were seeded in a 96-well v-bottom plate at a density of 2×10^4 per well and activated by Wash Dynabeads from Life Technologies according to the manufacturer's instructions. 2×10^4 Jurkat cells with 50 μL medium were incubated with 2 μL of pre-washed and resuspended Dynabeads in a humidified CO_2 incubator at 37°C for 24 h. After activation, Jurkat cells were treated 10-fold serial dilutions of different Dmab, Dmab-mPA, or Dmab-mPA + LD for 72 h. The cell viability was measured as described in the methods section.

2.7. References

1. Leader, B., Baca, Q. J. & Golan, D. E. Protein therapeutics: A summary and pharmacological classification. *Nature Reviews Drug Discovery* **7,** 21–39 (2008).
2. Lagassé, H. A. D. *et al.* Recent advances in (therapeutic protein) drug development. *F1000Research* **6,** 113 (2017).
3. Akishiba, M. *et al.* Cytosolic antibody delivery by lipid-sensitive endosomolytic peptide. *Nat. Chem.* **9,** 751–761 (2017).
4. Zhang, Z. *et al.* The fluorination effect of fluoroamphiphiles in cytosolic protein delivery. *Nat. Commun.* **9,** (2018).
5. Zuris, J. A. *et al.* Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. *Nat. Biotechnol.* **33,** 73–80 (2015).
6. Guidotti, G., Brambilla, L. & Rossi, D. Cell-Penetrating Peptides: From Basic Research to Clinics. *Trends in Pharmacological Sciences* **38,** 406–424 (2017).
7. Ng, D. Y. W. *et al.* Constructing hybrid protein zymogens through protective dendritic assembly. *Angew. Chemie - Int. Ed.* **53,** 324–328 (2014).
8. Scaletti, F. *et al.* Protein delivery into cells using inorganic nanoparticle–protein supramolecular assemblies. *Chem. Soc. Rev.* (2018). doi:10.1039/C8CS00008E
9. Mout, R. *et al.* General Strategy for Direct Cytosolic Protein Delivery via Protein-Nanoparticle Co-engineering. *ACS Nano* **11,** 6416–6421 (2017).
10. Weiner, L. M., Surana, R. & Wang, S. Monoclonal antibodies: Versatile platforms for cancer immunotherapy. *Nature Reviews Immunology* **10,** 317–327 (2010).
11. Weiner, G. J. Building better monoclonal antibody-based therapeutics. *Nature Reviews Cancer* **15,** 361–370 (2015).
12. Scott, A. M., Wolchok, J. D. & Old, L. J. Antibody therapy of cancer. *Nature Reviews Cancer* **12,** 278–287 (2012).
13. Carter, P. J. & Lazar, G. A. Next generation antibody drugs: Pursuit of the 'high-hanging fruit'. *Nature Reviews Drug Discovery* **17,** 197–223 (2018).
14. Kaplon, H. & Reichert, J. M. Antibodies to watch in 2018. *MAbs* **10,** 183–203 (2018).
15. Sliwkowski, M. X. & Mellman, I. Antibody therapeutics in cancer. *Science* **341,** 1192–1198 (2013).
16. Akbari, B. *et al.* Immunotoxins in cancer therapy: Review and update. *International Reviews of Immunology* **36,** 207–219 (2017).
17. Beck, A., Goetsch, L., Dumontet, C. & Corvaïa, N. Strategies and challenges for the next generation of antibody-drug conjugates. *Nature Reviews Drug Discovery* **16,** 315–337 (2017).
18. Chalouni, C. & Doll, S. Fate of Antibody-Drug Conjugates in Cancer Cells. *Journal of Experimental and Clinical Cancer Research* **37,** (2018).
19. Kalim, M. *et al.* Intracellular trafficking of new anticancer therapeutics: Antibody–drug conjugates. *Drug Design, Development and Therapy* **11,** 2265–2276 (2017).
20. Alewine, C., Hassan, R. & Pastan, I. Advances in Anticancer Immunotoxin Therapy. *Oncologist* **20,** 176–185 (2015).
21. Alouf, J. & Popoff, M. *The Comprehensive Sourcebook of Bacterial Protein Toxins. The Comprehensive Sourcebook of Bacterial Protein Toxins* (2006). doi:10.1016/B978-0-12-088445-2.X5000-8
22. Pannifer, A. D. *et al.* Crystal structure of the anthrax lethal factor. *Nature* **414,** 229–233

(2001).
23. Leppla, S. H. *et al.* Anthrax toxin edema factor: a bacterial adenylate cyclase that increases cyclic AMP concentrations of eukaryotic cells. *Proc. Natl. Acad. Sci.* **79,** 3162–3166 (1982).
24. Klimpel, K. R., Molloy, S. S., Thomas, G. & Leppla, S. H. Anthrax toxin protective antigen is activated by a cell surface protease with the sequence specificity and catalytic properties of furin. *Proc. Natl. Acad. Sci.* **89,** 10277–10281 (1992).
25. Milne, J. C., Furlong, D., Hanna, P. C., Wall, J. S. & Collier, R. J. Anthrax protective antigen forms oligomers during intoxication of mammalian cells. *J. Biol. Chem.* **269,** 20607–20612 (1994).
26. Kintzer, A. F. *et al.* The Protective Antigen Component of Anthrax Toxin Forms Functional Octameric Complexes. *J. Mol. Biol.* **392,** 614–629 (2009).
27. Feld, G. K. *et al.* Structural basis for the unfolding of anthrax lethal factor by protective antigen oligomers. *Nat. Struct. Mol. Biol.* **17,** 1383–1390 (2010).
28. Mogridge, J., Cunningham, K. & Collier, R. J. Stoichiometry of anthrax toxin complexes. *Biochemistry* **41,** 1079–1082 (2002).
29. Nassi, S., Collier, R. J. & Finkelstein, A. PA63 channel of anthrax toxin: An extended ??-barrel. *Biochemistry* **41,** 1445–1450 (2002).
30. Lacy, D. B., Wigelsworth, D. J., Melnyk, R. A., Harrison, S. C. & Collier, R. J. Structure of heptameric protective antigen bound to an anthrax toxin receptor: A role for receptor in pH-dependent pore formation. *Proc. Natl. Acad. Sci.* **101,** 13147–13151 (2004).
31. Miller, C. J., Elliott, J. L. & Collier, R. J. Anthrax protective antigen: Prepore-to-pore conversion. *Biochemistry* **38,** 10432–10441 (1999).
32. Hu, H. & Leppla, S. H. Anthrax toxin uptake by primary immune cells as determined with a lethal factor-beta-lactamase fusion protein. *PLoS One* **4,** e7946 (2009).
33. Milne, J. C., Blanket, S. R., Hanna, P. C. & Collier, R. J. Protective antigen-binding domain of anthrax lethal factor mediates translocation of a heterologous protein fused to its amino- or carboxy-terminus. *Mol. Microbiol.* **15,** 661–666 (1995).
34. Arora, N., Klimpel, K. R., Singh, Y. & Leppla, S. H. Fusions of anthrax toxin lethal factor to the ADP-ribosylation domain of Pseudomonas exotoxin A are potent cytotoxins which are translocated to the cytosol of mammalian cells. *J. Biol. Chem.* **267,** 15542–15548 (1992).
35. Antic, I., Biancucci, M., Zhu, Y., Gius, D. R. & Satchell, K. J. F. Site-specific processing of Ras and Rap1 Switch i by a MARTX toxin effector domain. *Nat. Commun.* **6,** (2015).
36. Liao, X., Rabideau, A. E. & Pentelute, B. L. Delivery of antibody mimics into mammalian cells via anthrax toxin protective antigen. *ChemBioChem* **15,** 2458–2466 (2014).
37. Rabideau, A. E., Liao, X. & Pentelute, B. L. Delivery of mirror image polypeptides into cells. *Chem. Sci.* **6,** 648–653 (2015).
38. Mechaly, A., McCluskey, A. J. & John Collier, R. Changing the receptor specificity of anthrax toxin. *MBio* **3,** (2012).
39. McCluskey, A. J., Olive, A. J., Starnbach, M. N. & Collier, R. J. Targeting HER2-positive cancer cells with receptor-redirected anthrax protective antigen. *Mol. Oncol.* **7,** 440–451 (2013).
40. Liu, S., Netzel-Arnett, S., Birkedal-Hansen, H. & Leppla, S. H. Tumor cell-selective cytotoxicity of matrix metalloproteinase-activated anthrax toxin. *Cancer Res.* **60,** 6061–6067 (2000).

41. Liu, S., Bugge, T. H. & Leppla, S. H. Targeting of Tumor Cells by Cell Surface Urokinase Plasminogen Activator-dependent Anthrax Toxin. *J. Biol. Chem.* **276,** 17976–17984 (2001).
42. Mourez, M. *et al.* Mapping dominant-negative mutations of anthrax protective antigen by scanning mutagenesis. *Proc. Natl. Acad. Sci.* **100,** 13803–13808 (2003).
43. Rosovitz, M. J. *et al.* Alanine-scanning mutations in domain 4 of anthrax toxin protective antigen reveal residues important for binding to the cellular receptor and to a neutralizing monoclonal antibody. *J. Biol. Chem.* **278,** 30936–30944 (2003).
44. Chen, I., Dorr, B. M. & Liu, D. R. A general strategy for the evolution of bond-forming enzymes using yeast display. *Proc. Natl. Acad. Sci.* **108,** 11399–11404 (2011).
45. Alley, S. C. *et al.* Contribution of linker stability to the activities of anticancer immunoconjugates. *Bioconjug. Chem.* **19,** 759–765 (2008).
46. Lyon, R. P. *et al.* Self-hydrolyzing maleimides improve the stability and pharmacological properties of antibody-drug conjugates. *Nat. Biotechnol.* **32,** 1059–1062 (2014).
47. Ciardiello, F. & Tortora, G. EGFR Antagonists in Cancer Treatment. *N. Engl. J. Med.* **358,** 1160–1174 (2008).
48. Berchuck, A. *et al.* Overexpression of HER-2/neu is associated with poor survival in advanced epithelial ovarian cancer. *Cancer Res.* **50,** 4087–91 (1990).
49. Gravalos, C. & Jimeno, A. HER2 in gastric cancer: A new prognostic factor and a novel therapeutic target. *Annals of Oncology* **19,** 1523–1529 (2008).
50. Arteaga, C. L. *et al.* Treatment of HER2-positive breast cancer: current status and future perspectives. *Nat. Rev. Clin. Oncol.* **9,** 16–32 (2011).
51. Yewale, C., Baradia, D., Vhora, I., Patil, S. & Misra, A. Epidermal growth factor receptor targeting in cancer: A review of trends and strategies. *Biomaterials* **34,** 8690–8707 (2013).
52. Kirkpatrick, P., Graham, J. & Muhsin, M. Cetuximab. *Nat. Rev. Drug Discov.* **3,** 549–550 (2004).
53. Yamashiro, H., Toyama, K., Bando, H., Saji, S. & Toi, M. Trastuzumab treatment for breast cancer. *The New England Journal of Medicine* **354,** 2186; author reply 2186 (2006).
54. Wilson, B. a & Collier, R. J. Diphtheria toxin and Pseudomonas aeruginosa exotoxin A: active-site structure and enzymic mechanism. *Curr. Top. Microbiol. Immunol.* **175,** 27–41 (1992).
55. Subik, K. *et al.* The expression patterns of ER, PR, HER2, CK5/6, EGFR, KI-67 and AR by immunohistochemical analysis in breast cancer cell lines. *Breast Cancer Basic Clin. Res.* **4,** 35–41 (2010).
56. Sellman, B. R., Nassi, S. & Collier, R. J. Point Mutations in Anthrax Protective Antigen that Block Translocation. *J. Biol. Chem.* **276,** 8371–8376 (2001).
57. Weigelt, B., Lo, A. T., Park, C. C., Gray, J. W. & Bissell, M. J. HER2 signaling pathway activation and response of breast cancer cells to HER2-targeting agents is dependent strongly on the 3D microenvironment. *Breast Cancer Res. Treat.* **122,** 35–43 (2010).
58. Serezani, C. H., Ballinger, M. N., Aronoff, D. M. & Peters-Golden, M. Cyclic AMP: Master regulator of innate immune cell function. *American Journal of Respiratory Cell and Molecular Biology* **39,** 127–132 (2008).
59. Raker, V. K., Becker, C. & Steinbrink, K. The cAMP pathway as therapeutic target in autoimmune and inflammatory diseases. *Frontiers in Immunology* **7,** (2016).
60. Vu, T. & Claret, F. X. Trastuzumab: Updated Mechanisms of Action and Resistance in Breast Cancer. *Front. Oncol.* **2,** (2012).

61. Pohlmann, P. R., Mayer, I. A. & Mernaugh, R. Resistance to trastuzumab in breast cancer. *Clinical Cancer Research* **15,** 7479–7491 (2009).
62. Ware, K. E. *et al.* A mechanism of resistance to gefitinib mediated by cellular reprogramming and the acquisition of an FGF2-FGFR1 autocrine growth loop. *Oncogenesis* **2,** (2013).
63. Van Emburgh, B. O., Sartore-Bianchi, A., Di Nicolantonio, F., Siena, S. & Bardelli, A. Acquired resistance to EGFR-targeted therapies in colorectal cancer. *Mol. Oncol.* **8,** 1084–94 (2014).
64. Brand, T. M., Iida, M. & Wheeler, D. L. Molecular mechanisms of resistance to the EGFR monoclonal antibody cetuximab. *Cancer Biology and Therapy* **11,** 777–792 (2011).
65. Misale, S. *et al.* Emergence of KRAS mutations and acquired resistance to anti-EGFR therapy in colorectal cancer. *Nature* **486,** 532–536 (2012).
66. Ritter, C. A. *et al.* Human breast cancer cells selected for resistance to trastuzumab in vivo overexpress epidermal growth factor receptor and ErbB ligands and remain dependent on the ErbB receptor network. *Clin. Cancer Res.* **13,** 4909–4919 (2007).
67. Nahta, R., Yuan, L. X. H., Zhang, B., Kobayashi, R. & Esteva, F. J. Insulin-like growth factor-I receptor/human epidermal growth factor receptor 2 heterodimerization contributes to trastuzumab resistance of breast cancer cells. *Cancer Res.* **65,** 11118–11128 (2005).
68. Corkery, B., Crown, J., Clynes, M. & O'Donovan, N. Epidermal growth factor receptor as a potential therapeutic target in triple-negative breast cancer. *Ann. Oncol.* **20,** 862–867 (2009).
69. Mukohara, T. *et al.* Differential effects of gefitinib and cetuximab on non-small-cell lung cancers bearing epidermal growth factor receptor mutations. *J. Natl. Cancer Inst.* **97,** 1185–1194 (2005).
70. Napolitano, S. *et al.* Primary and acquired resistance of colorectal cancer to anti-EGFR monoclonal antibody can be overcome by combined treatment of regorafenib with cetuximab. *Clin. Cancer Res.* **21,** 2975–2983 (2015).
71. Press, M. F., Cordon-Cardo, C. & Slamon, D. J. Expression of the HER-2/neu proto-oncogene in normal human adult and fetal tissues. *Oncogene* **5,** 953–62 (1990).
72. Yano, S. *et al.* Distribution and Function of EGFR in Human Tissue and the Effect of EGFR Tyrosine Kinase Inhibition. *Anticancer Research* **23,** 3639–3650 (2003).
73. Mehner, C. *et al.* Tumor cell-produced matrix metalloproteinase 9 (MMP-9) drives malignant progression and metastasis of basal-like triple negative breast cancer. *Oncotarget* **5,** 2736–2749 (2014).
74. Huang, S., New, L., Pan, Z., Han, J. & Nemerow, G. R. Urokinase plasminogen activator/urokinase-specific surface receptor expression and matrix invasion by breast cancer cells requires constitutive p38alpha mitogen-activated protein kinase activity. *J Biol Chem* **275,** 12266–12272 (2000).
75. Abi-Habib, R. J. *et al.* A urokinase-activated recombinant anthrax toxin is selectively cytotoxic to many human tumor cell types. *Mol. Cancer Ther.* **5,** 2556–2562 (2006).
76. Rabideau, A. E. & Pentelute, B. L. Delivery of Non-Native Cargo into Mammalian Cells Using Anthrax Lethal Toxin. *ACS Chemical Biology* **11,** 1490–1501 (2016).
77. Pastan, I., Hassan, R., FitzGerald, D. J. & Kreitman, R. J. Immunotoxin therapy of cancer. *Nature Reviews Cancer* **6,** 559–565 (2006).
78. Yamaizumi, M., Mekada, E., Uchida, T. & Okada, Y. One molecule of diphtheria toxin fragment a introduced into a cell can kill the cell. *Cell* **15,** 245–250 (1978).

79. Nozaki, S. *et al.* Targeting urokinase-type plasminogen activator and its receptor for cancer therapy. *Anti-Cancer Drugs* **17,** 1109–1117 (2006).
80. Stetler-Stevenson, W. G., Aznavoorian, S. & Liotta, L. A. Tumor Cell Interactions with the Extracellular Matrix During Invasion and Metastasis. *Annu. Rev. Cell Biol.* **9,** 541–573 (1993).
81. Liu, S. *et al.* Anti-tumor activity of anthrax toxin variants that form a functional translocation pore by intermolecular complementation. *Oncotarget* **8,** 65123–65131 (2017).
82. Liu, S. *et al.* Solid tumor therapy by selectively targeting stromal endothelial cells. *Proc. Natl. Acad. Sci.* **113,** E4079–E4087 (2016).
83. Mazor, R. *et al.* Tolerogenic nanoparticles restore the antitumor activity of recombinant immunotoxins by mitigating immunogenicity. *Proc. Natl. Acad. Sci.* 201717063 (2018). doi:10.1073/pnas.1717063115
84. Kishimoto, T. K. *et al.* Improving the efficacy and safety of biologic drugs with tolerogenic nanoparticles. *Nat. Nanotechnol.* **11,** 890–899 (2016).
85. Mijalis, A. J. *et al.* A fully automated flow-based approach for accelerated peptide synthesis. *Nat. Chem. Biol.* **13,** 464–466 (2017).
86. Simon, M. D. *et al.* Rapid flow-based peptide synthesis. *ChemBioChem* **15,** 713–720 (2014).

Chapter 3: Single-component anthrax delivery system enabled by chemical conjugation

3.1. Introduction

Anthrax toxin is a bacterial toxin from Bacillus anthracis. It consists of two effector proteins, lethal factor (LF)[1] and edema factor (EF)[2], and a translocase, protective antigen (PA)[3]. Over the past 40 years, extensive biochemical and biophysical research has shed light on how these individual components act together to fulfill the biological function of anthrax toxin. Once the toxin is released by the bacteria, PA binds to two receptors on cell surface with nanomolar or picomolar affinity[4–6]. The bound PA is subsequently cleaved by cell-surfce furin to a 63 kDa form (PA_{63})[7] and oligomerizes into the PA heptameric or octameric prepore[8,9], which three or four LF or EF molecules dock to with high affinity[10,11]. The entire complex is then endocytosed and the acidic pH in the endosome converts the prepore into transmembrane pore by triggering its conformational rearrangement[12–14]. The pore translocates the LF or EF molecules to the cytoplasm, where the effector proteins catalyze their respective reactions[2,15].

Previous findings have revealed that the N-terminal domain of LF (LF_N, residues 1-254) interacts with two adjacent subunits of the PA prepore and is responsible for initiating PA-mediated translocation. Mechanistic studies show that key electrostatic interactions between the first α-helix and β-sheet of LF_N and PA help partially unfold the N-terminal region of the protein and poise it for N- to C-terminal translocation[11]. Based on these findings, a variety of cargos have been fused to the C-terminus of LFN and translocated via PA. These cargos include beta-lactamase[16], Diphtheria toxin A (DTA)[17], Pseudomonas exotoxin A[18], Ras protease RRSP[19], etc. In addition, antibody mimics[20] as well as non-natural entities such as mirror image polypeptides and proteins[21] have also been delivered into cells using the LF_N/PA system.

Since the function of LF_N can be dissected into co-localizing the cargo to the PA pore and facilitating the subsequent translocation, it is conceivable that this function may be partially replaced by covalent conjugation of the cargo protein to PA. In this study, we ligated the cargo protein with or without LF_N directly to PA through a specific sortase-mediated reaction. We also investigated the effect of fusing the cargo protein with an N-terminal polylysine tag, which has been previously shown to facilitate PA-dependent translocation[22,23]. Our results demonstrate that in the absence of LF_N, the chemically generated single-component cargo-PA conjugate with appropriate linker significantly enhanced delivery efficiency compared to the traditional dual-component system.

3.2. Results

3.2.1. Design of single-component anthrax system

A common strategy to combine two proteins together is recombinant expression. However, this strategy only allows one protein to be fused at the N- or C-terminus of the other protein. During the mechanism of PA, the C-terminal domain needs to bind to cell surface receptors and the N-terminal PA_{20} domain needs to be proteolytically cleaved by furin protease for activation. Therefore, recombinant expression of the cargo protein with PA risks disruption of the function of PA or the cargo protein.

We sought to use chemical conjugation as an alternative strategy to attach a cargo protein to PA. A previously known functional PA mutant PA[K563C] (PAC) was first prepared by recombinant expression in Escherichia coli Bl21(DE3) and purified with anion exchange chromatography. Two different peptides, one short D-peptide and one long L-peptide, were synthesized using Fmoc chemistry, each with an N-terminal triglycine and a C-terminal maleimide (Figures 3.2.1b and 3.6.1). The linkers were designed with different stabilities in the endosome with the long one susceptible to endosomal protease cathepsin B[24] and this was to test the cleavage requirement of the cargo protein. The peptides were then conjugated to a cargo protein with a C-terminal LPSTGG recognition motif via sortase-mediated ligation using an evolved sortase (SrtA*)[28]. The resulting maleimide-containing protein was purified and covalently ligated again to PAC through thiol-maleimide coupling. The product mixture was separated by size exclusion chromatography to give homogeneous protein-PA conjugate (Figures 3.2.1c-e).

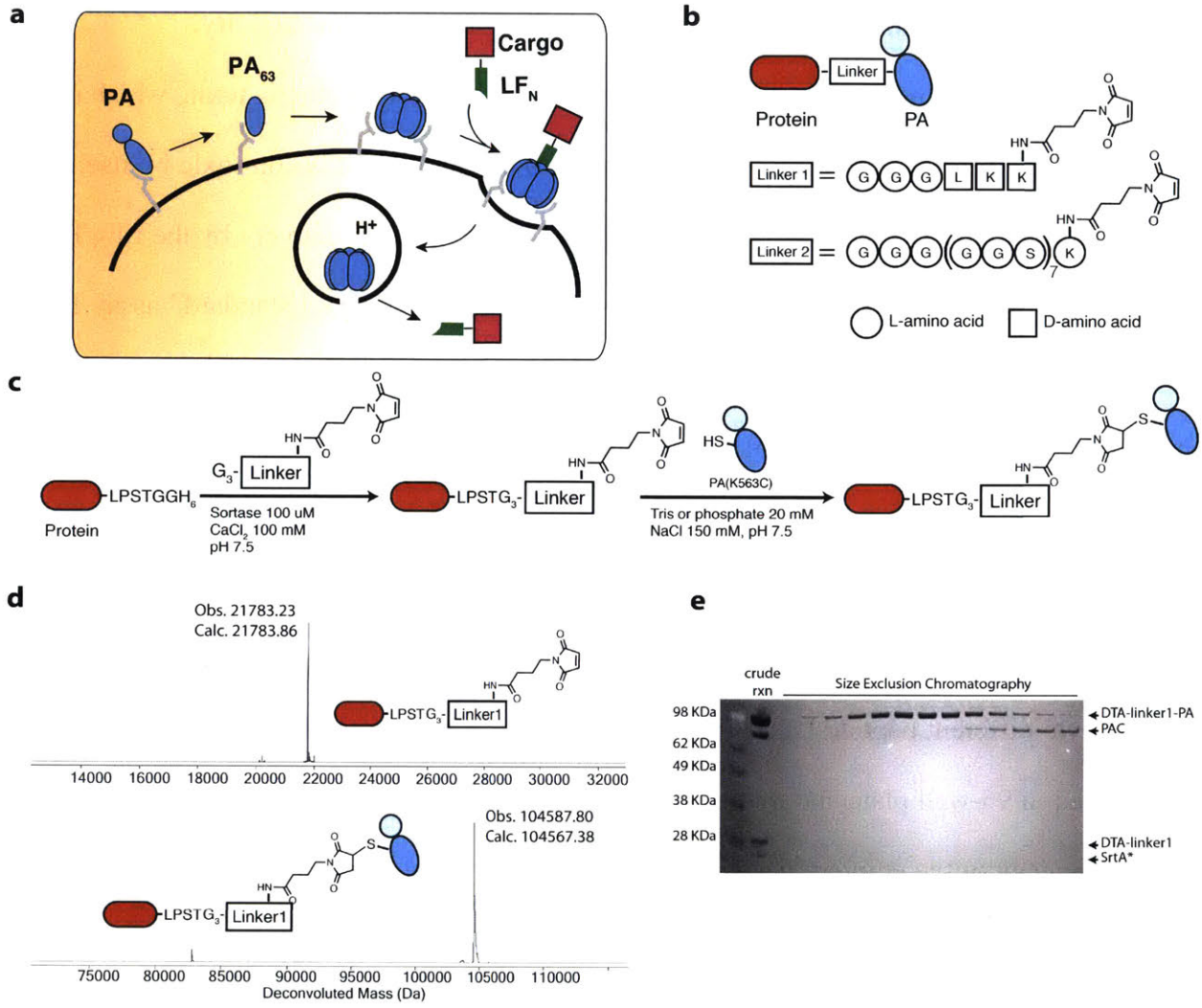

Figure 3.2.1. Design and preparation of the single-component anthrax delivery system. a) Mechanism of LF_N/PA-mediated translocation. b) Schematic showing the design of cargo protein-PA conjugate with two different linkers. Linker1 and linker2 contain either L- or D-amino acids respectively. Both linkers have N-terminal triglycine tag and C-terminal maleimide for conjugation. c) Cargo protein with a C-terminal LPSTGG tag was labeled with one of the peptide linkers via sortase-mediated ligation and then attached to PAC through maleimide-thiol coupling. The final product was purified by size exclusion chromatography. d) LC-MS deconvoluted mass traces of a representative protein cargo DTA conjugated to PA via peptide linker1. e) SDS-PAGE analysis of DTA-linker1-PA purification by size exclusion chromatography.

3.2.2. Single-component anthrax delivery system showed significant activity.

Diphtheria toxin A chain (DTA) is the catalytic domain of diphtheria toxin, which ribosylates the eukaryotic elongation factor 2 (EF-2) with ADP. Although it is non-toxic by itself, it becomes an extremely potent protein toxin upon cytosolic delivery. Its delivery by the LF_N/PA system has been extensively characterized and has been used as the "gold standard" assay for PA-mediated cytosolic translocation in previous studies. We used DTA as the reporter protein to study the activity of the single-component anthrax-mediated delivery system due to its robust effect on cell viability.

We first investigated whether the PA delivery system can still be functional when the LF_N cargo protein conjugate is directly attached to PA. LF_N-DTA-LPSTGG (LD) was ligated to PAC with the two different peptide linkers and purified as described above (Figure 3.2.2a). HeLa cells were plated in 96-well plate and treated with 5-fold serial dilutions of different proteins. After 72 h, the efficiency of the delivery was measured by the change of cell viability, which correlates to the amount of translocated DTA. We hypothesized that the cargo LD can be released from PA and translocated in case of the longer L-linker2 due to its susceptibility to cleavage in endosome, whereas the less vulnerable shorter D-linker1 should have a lower delivery efficiency. However, to our surprise, both LD-linker1-PA and LD-linker2-PA displayed highly efficient delivery of DTA, with LD-linker1-PA more active than LD-linker2-PA and almost as active as LD + PA (Figure 3.2.2b). LD protein by itself did not cause any cell death as expected, confirming this process is PA-dependent. While this observation may indicate both peptide linkers can be cleaved in endosome prior to translocation, it is also possible that LD is getting released in the medium especially considering the known instability of maleimide in serum.

In order to study whether the translocase function of PA is required for the delivery of DTA, we also conjugated LD to PA[F427A] (PAA), a mutant PA that's incapable of translocation but maintains the other functions. As expected, we found that PAA arrested translocation in both single-component LD-linker1-PAA and two-component LD + PAA (Figure 3.2.2c), indicating the translocation is mediated by the translocase activity of PA.

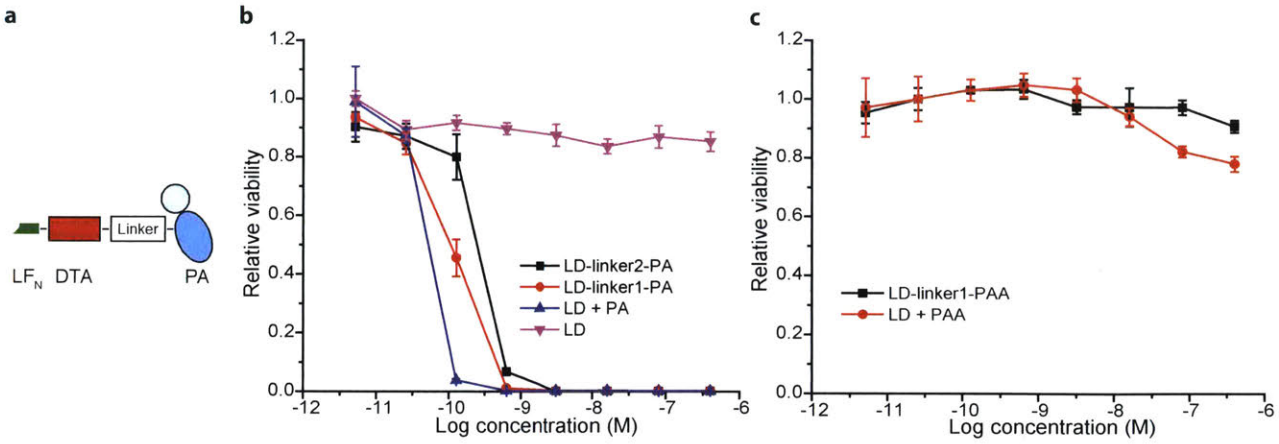

Figure 3.2.2. Delivery of LF_N-DTA (LD) mediated by the single-component anthrax system. a) Design of single-component LD-PA conjugate. b) Relative viability of HeLa cells after treatment with different proteins. The cells were incubated with 5-fold serial dilutions of LD-linker1-PA ($IC_{50} = 4.71 \times 10^{-11}$ M), LD-linker2-PA ($IC_{50} = 1.22 \times 10^{-10}$ M), and positive control LD + PA ($IC_{50} = 5.27 \times 10^{-10}$ M), or negative control LD for 72 h. The cell viability was measured by CellTiter-Glo assay and normalized to the group with no treatment. In the case of LD + PA, equimolar of LD and PA were used and diluted simultaneously to keep a 1:1 ratio. c) Cells were treated with LD conjugated to PA[F427A] (PAA), a translocation-deficient mutant of PA, under the same conditions. Each data point represents a triplicate.

3.2.3. Single-component anthrax delivery system has significantly improved activity in the absence of LF_N.

Past biophysical and biochemical studies have shed light on the two-fold function of the first 254 residues of lethal factor LF_N: 1. It co-localizes the cargo protein with PA by binding to the PA prepore through its interaction with two adjacent subunits of the oligomer; 2. It also initiates translocation by partially unfolding its N-terminal region and poising the protein for translocation from N- to C-terminus.

Based on the activity data of LD-PA, we wanted to study whether the function of LF_N can be partially replaced by co-localizing the cargo protein with PA through chemical ligation. DTA-LPSTGG was conjugated to PAC using the same method as described above (Figure 3.2.3a). HeLa cells were incubated with 5-fold serial dilutions of DTA-linker1-PA or DTA-linker2-PA for 72 h and the cell viability was measured by CellTiter-Glo. Both conjugates exhibited much more enhanced activity when compared to the control of separate DTA and PA (Figure 3.2.3b). DTA by itself did not show any activity even at the highest concentration used. In comparison, DTA-PAA conjugates of different linkers as well as DTA + PAA failed to inhibit cell growth (Figure 3.2.3c). Taken together, these data suggest covalent attachment of cargo protein to PA partially restores the function of LF_N and significantly improve the delivery efficiency of the cargo. These results also indicate the cleavage of the maleimide was not significant enough to mask the effect of co-localization, since the free DTA is much less potent than the single-component in the presence of PA. One possible explanation for the superior activity of DTA-linker1-PA to DTA-linker2-PA is the longer and more flexible linker2 led to intramolecular DTA-PA interaction, which inhibited one or a few steps during the mechanism of PA translocation.

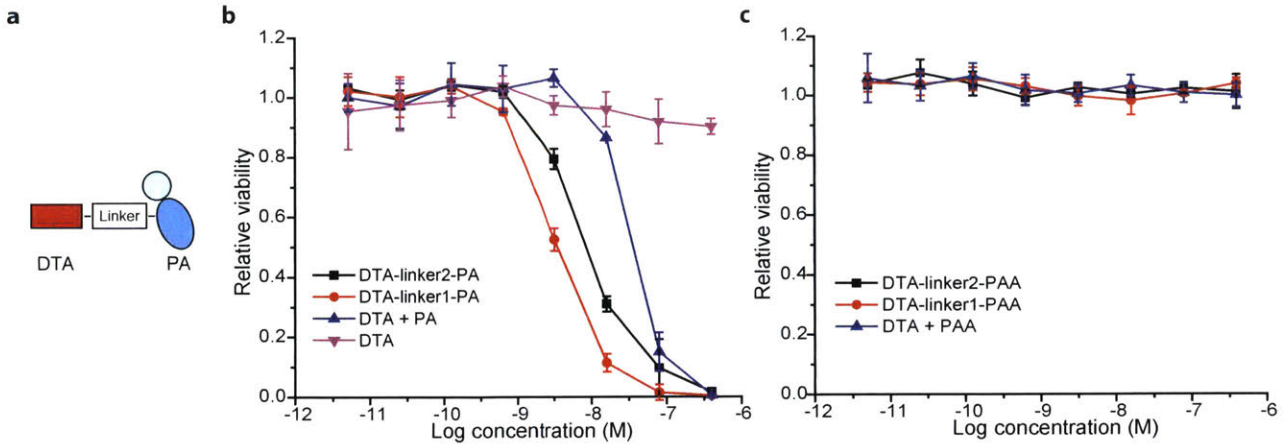

Figure 3.2.3. Delivery of DTA in the absence of LF_N. a) Design of DTA-PA conjugate. DTA without the N-terminal LFN was conjugated directly to PA. b) Relative viability of HeLa cells after incubation with DTA-PA. HeLa cell were treated with 5-fold serial dilutions of DTA-linker1-PA ($IC_{50} = 3.32 \times 10^{-9}$ M), DTA-linker2-PA ($IC_{50} = 8.14 \times 10^{-9}$ M), and DTA +PA ($IC_{50} = 3.38 \times 10^{-8}$ M) for 72 h before subject to CellTiter-Glo assay. c) Relative viability of HeLa cells after treatment with DTA-PAA. Cells were treated with DTA-PAA conjugates under the same conditions and measured for their relative viability. Each data point is a triplicate.

3.2.4. N-terminal polylysine further improves the efficiency of the single-component system.

It has been reported previously that N-terminal polylysine tag can mediate translocation through the PA pore. Here we investigated whether an N-terminal Lys tag can further potentiate the activity of the single-component PA delivery system. K_5DTA-LPSTGG was conjugated to PAC using the same strategy as described above (Figure 3.2.4a). We treated HeLa cells with serial dilutions of K_5DTA-PA conjugates with K_5DTA + PA as the control. We found that separate K_5DTA and PA increased the delivery efficiency by nearly 10-fold compared to without the K_5 tag, consistent with the previous finding (Figure 3.2.4b). More importantly, the single-component K_5DTA-linker1-PA further increased the activity by nearly another order of magnitude, in good agreement with our hypothesis. On the other hand, K_5DTA-linker2-PA failed to show any significantly enhanced delivery compared to the two-component system or DTA-linker2-PA. This might be due to stronger nonspecific interaction between K_5DTA and PA caused by the polycationic tag. To confirm the delivery is still PA-dependent, we again conjugated K_5DTA to PAA and treated HeLa cells under the same condition. Neither K_5DTA-linker1-PAA or K_5DTA-linker2-PAA showed any antiproliferative effect even at the highest concentration tested (Figure 3.2.4c). Similarly, co-treatment of K_5DTA and PAA also had little impact on the cell growth. Taken together, these results indicate the covalent conjugation strategy with proper linker significantly improves the delivery efficiency of polylysine tagged cargo protein. They also suggest that the K_5 tag contributed more to the initiation of translocation than to the binding to PA since it increased the activity of the unconjugated system by about the same extent as the conjugated system.

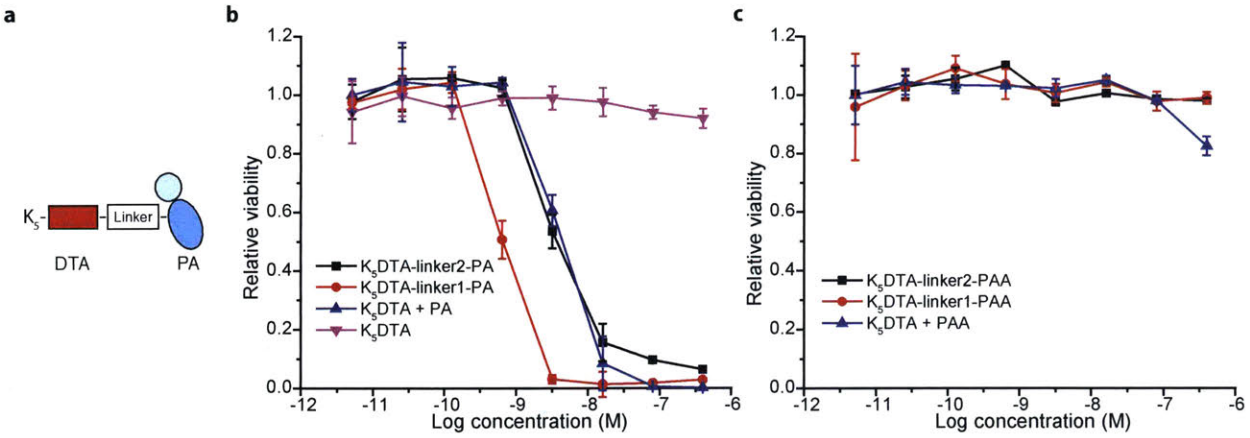

Figure 3.2.4. Delivery of polylysine tagged DTA with the single-component system. a) Schematic representation of K_5DTA-PA. The N-terminal K_5 tag was fused to DTA by recombinant expression. b) Relative viability of HeLa cells treated with K_5DTA-linker1-PA (IC_{50} = 6.40 × 10^{-10} M), K_5DTA-linker2-PA (IC_{50} = 3.74 × 10^{-9} M), and K_5DTA + PA (IC_{50} = 3.89 × 10^{-9} M) for 72 h. c) Relative viability of HeLa cells after 72-h K_5DTA-PAA or K_5DTA + PAA treatment. The cell viability was measured by CellTiter-Glo and normalized to untreated cells. Each data point is a triplicate.

3.3. Discussion

Here we report a novel single-component anthrax delivery system that can significantly enhance the delivery efficiency of the protein cargo in the absence of LF_N. The single-component system was generated by conjugating the cargo protein, which is labeled with C-terminal maleimide by sortase-mediated ligation, site-specifically to the sole cysteine in PAC via maleimide-thiol crosslinking to give a homogeneous product with an exact cargo-to-PA ratio of one. Another alternative would be to use a non-specific conjugation strategy such as a bifunctional linker with NHS ester and maleimide. Although this approach could potentially increase the cargo load per PA molecule, it only works when the cargo protein has no cysteine, which severely limits its application. Moreover, some conjugated cargos could be near places that are important for PA function, such as the receptor binding region or the interface between subunits during oligomerization.

When we sought out to design the single-component PA delivery system, our goal was to engineer it so that the cargo protein is only released from PA in the endosome. For this purpose, two linkers were synthesized, a long poly-GGS linker (linker2) that is susceptible to cleavage by the endosomal protease cathepsin B[24] and a short D-peptide linker (linker1) as a non-cleavable control. In order to test the cleavabilities of the poly-GGS linker, we conjugated LD to PA using both linkers and hypothesized that only LD-linker2-PA would be active. However, to our surprise, both conjugates displayed efficient delivery of LD within one order of magnitude to LD + PA. With this observation, we came up with three possible explanations: 1. Both linkers can be cleaved efficiently in the endosome due to its harsh environment. 2. Both linkers were cleaved in the medium due to the reported instability of maleimide[25] and the observed activity completely came from the LD released outside of cell. 3. LD-PA conjugate can translocate and the linker cleavage is not required.

In order to investigate whether the single-component system is truly effective, we removed LF_N and conjugated DTA directly to PA. Since DTA does not translocate through PA very well, if it is cleaved outside of cell, DTA-PA should have comparable activity to DTA + PA. Moreover, if there is no cleavage, DTA-PA should not perform better than DTA + PA due to the extra cargo load. To our delight, when we treated the cells with these constructs, we observed superior activity of DTA-PA to DTA + PA, eliminating the possibilities of premature cleavage and DTA-PA translocation. Fusion of the N-terminal K_5 tag further demonstrated that the single-component PA system significantly improved the translocation of the LF_N-null cargo. Another observation from these experiments was that linker1 was consistently better than linker2 and the difference widened in the case of K_5 tagged DTA. Since it is unlikely that the short D-linker is more susceptible to cleavage in endosome, a more plausible explanation would be the longer and more flexible linker causes intramolecular interaction between the cargo protein and PA, which inhibited the translocation process. This effect would become more prominent with K_5 tagged cargo due to the stronger interaction between the tag and the negatively-charged cation-selective PA pore. This observation also suggests that the linker cleavage is the rate-limiting step.

It was discovered years ago by the Collier group that fusing a stretch of basic amino acids to the N-terminus of DTA can potentiate its PA-dependent translocation[23]. It is believed that the poly-lys tag is drawn by the negatively charged pore by electrostatic attraction and inserts itself into the lumen of the pore. The cargo protein is then translocated through the pore by charge state-dependent Brownian ratcheting process. Here we show that the covalent linkage of the cargo to PA significantly enhanced the translocation with or without the K_5 tag. Our results suggest that the poly-lys is not enough to fully co-localize the tagged cargo with the PA prepore and it contributes to the improved activity primarily by facilitating the initiation of translocation. Another point

worth discussion is that since K5DTA-PA has some binding affinity to the prepore/pore and the linker cleavage is likely the rate-limiting step, the intact K5DTA-PA will compete with free K5DTA and as a result the potential of the single-component system was probably not fully realized. In the future, we envision the system can be further improved by optimizing the linker for endosomal cleavage and other cargos in addition to DTA will delivered using this strategy.

3.4. Experimental

3.4.1 Synthesis and purification of peptide linkers

Fmoc-protected amino acids were purchased from ChemImpex. Peptide linker1 and linker2 were synthesized using either manual flow peptide synthesis or automated flow peptide synthesizer on ChemMatrix resin with a Rink amide linker as previously described[26,27]. The maleimide was coupled on the side chain of L- or D-lysine by treating the Lys-deprotected resin with 1.5 equivalents of N-γ-maleimidobutyryl-oxysuccinimide ester (GMBS, Thermo Fisher) dissolved in DMF for 3 h. The crude peptides were cleaved by TFA/H_2O (97.5:2.5) at RT for 2 h and purified by semi-preparative RP-HPLC with Agilent Zorbax 300SB C_{18} column (9.4 x 250 mm, 5 um) at a flow rate of 4 mL/min using the gradient of 1-31% acetonitrile over 80 min. HPLC fractions were subsequently analyzed by LC-MS and pure ones were pooled and lyophilized (supplementary figure).

3.4.2. Protein expression and purification

All proteins were expressed in E. coli BL21 (DE3) cells from Thermo Fisher. PA[K563C] (PAC) and K_5DTA-LPSTGG were prepared from WT PA and DTA-LPSTGG respectively using QuickChange single site-directed mutagenesis kit (Agilent). The expression of PAC was done at New England Regional Center of Excellence/Biodefense and Emerging Infectious Diseases (NERCE) and purified as previously described[14]. 15g to 25g of cell pellet was resuspended in sucrose buffer (20 mM Tris pH 8.5, 1 mM EDTA, 20% sucrose) and incubated with 5 mM $MgSO_4$ buffer. The supernatant was purified by an anion exchange chromatography and pure fractions were pooled. LF_N-DTA-LPSTGG (LD), DTA-LPSTGG, and K_5DTA-LPSTGG in Champion PET-His_6-SUMO vector were also expressed in E. coli BL21(DE3). The bacteria were grown at 37°C until OD reached 0.7~0.9 and induced with 0.4 mM IPTG at 30°C for 16 h. The proteins

were purified by a HisTrap FF crude Ni-NTA column (GE) and desalted by buffer exchange. The SUMO was cleaved by SUMO protease and removed by size exclusion chromatography.

3.4.3. Generation of single-component anthrax delivery system

Cargo proteins with C-terminal LPSTGG tag (50 uM) were first conjugated with poly-glycine maleimide peptide linkers (1.5 mM) by a triple mutant sortase (SrtA*)[28] and purified by Ni beads pull-down as previously described. 1.5 equivalents of maleimide labeled cargo protein (30 uM) was then reacted with PAC (20 uM) in 20 mM Tris, 150 mM NaCl buffer (pH 8.5) at RT for 1 h and the resulted mixture was loaded onto HiLoad 16/600 Superdex 200 pg size exclusion column (GE). The fractions were analyzed by SDS-PAGE stained with Coomassie dye and the fractions containg cargo-PA were pooled and concentrated.

3.4.4. Cell viability assays

HeLa cells were purchased from American Type Culture Collection (Manassas, VA) and maintained in MEM supplemented with 10% at 37°C in a 5% CO_2 environment. On the day prior to treatment, the cells were seeded in a 96-well plate at a density of 5×10^3 per well and allowed to attach overnight. The next day, the cells were treated with 5-fold serial dilutions of different proteins for 72 h. The cell viability was measured by CellTiter-Glo luminescent assay (Promega) following the manufacturer's protocol. The relative viability was normalized to cells without treatment (n=3).

3.5. Acknowledgements

This work was generously funded by MIT start-up funds, MIT Reed Fund, Damon Runyon Cancer Research Foundation Innovation Award, National Science Foundation (NSF) CAREER Award (CHE-1351807) for B.L.P, F32 CA180653 (B.R.P.), Novartis Early Career Award (Novartis Institutes for BioMedical Research Incorporated), Unrestricted Grant in Synthetic Organic Chemistry (Bristol Myers SQUIBB), the Friends of DFCI (B.R.P.). We also thank R. J. Collier (Harvard) for his continued contribution of laboratory equipment and the NERCE facility (grant: U54 AI057159) for expression of toxin proteins.

3.6. Appendix

3.6.1. The amino acid sequences of the PA mutants used in this study

PA[K563C] (PAC)

MEVKQENRLLNESESSSQGLLGYYFSDLNFQAPMVVTSSTTGDLSIPSSELENIPSENQYF
QSAIWSGFIKVKKSDEYTFATSADNHVTMWVDDQEVINKASNSNKIRLEKGRLYQIKIQ
YQRENPTEKGLDFKLYWTDSQNKKEVISSDNLQLPELKQKSSNSRKKRSTSAGPTVPDR
DNDGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSTASDPYSDFEK
VTGRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRTH
TSEVHGNAEVHASFFDIGGSVSAGFSNSNSSTVAIDHSLSLAGERTWAETMGLNTADTA
RLNANIRYVNTGTAPIYNVLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYYPSKNLAPI
ALNAQDDFSSTPITMNYNQFLELEKTKQLRLDTDQVYGNIATYNFENGRVRVDTGSNW
SEVLPQIQETTARIIFNGKDLNLVERRIAAVNPSDPLETTKPDMTLKEALKIAFGFNEPNG
NLQYQGKDITEFDFNFDQQTSQNICNQLAELNATNIYTVLDKIKLNAKMNILIRDKRFHY
DRNNIAVGADESVVKEAHREVINSSTEGLLLNIDKDIRKILSGYIVEIEDTEGLKEVINDR
YDMLNISSLRQDGKTFIDFKKYNDKLPLYISNPNYKVNVYAVTKENTIINPSENGDTSTN
GIKKILIFSKKGYEIG

PA[F427A, K563C] (PACA)

MEVKQENRLLNESESSSQGLLGYYFSDLNFQAPMVVTSSTTGDLSIPSSELENIPSENQYF
QSAIWSGFIKVKKSDEYTFATSADNHVTMWVDDQEVINKASNSNKIRLEKGRLYQIKIQ
YQRENPTEKGLDFKLYWTDSQNKKEVISSDNLQLPELKQKSSNSRKKRSTSAGPTVPDR
DNDGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSTASDPYSDFEK
VTGRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRTH
TSEVHGNAEVHASFFDIGGSVSAGFSNSNSSTVAIDHSLSLAGERTWAETMGLNTADTA
RLNANIRYVNTGTAPIYNVLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYYPSKNLAPI
ALNAQDDASSTPITMNYNQFLELEKTKQLRLDTDQVYGNIATYNFENGRVRVDTGSNW
SEVLPQIQETTARIIFNGKDLNLVERRIAAVNPSDPLETTKPDMTLKEALKIAFGFNEPNG
NLQYQGKDITEFDFNFDQQTSQNICNQLAELNATNIYTVLDKIKLNAKMNILIRDKRFHY
DRNNIAVGADESVVKEAHREVINSSTEGLLLNIDKDIRKILSGYIVEIEDTEGLKEVINDR
YDMLNISSLRQDGKTFIDFKKYNDKLPLYISNPNYKVNVYAVTKENTIINPSENGDTSTN
GIKKILIFSKKGYEIG

*The mutated residues are highlighted in red.

3.6.2. The amino acid sequences of the LF_N-DTA-LPSTGG, DTA-LPSTGG, and K_5DTA-LPSTGG

LF_N-DTA-LPSTGG (LD)

AGGHGDVGMHVKEKEKNKDENKRKDEERNKTQEEHLKEIMKHIVKIEVKGEEAVKKE
AAEKLLEKVPSDVLEMYKAIGGKIYIVDGDITKHISLEALSEDKKKIKDIYGKDALLHEH
YVYAKEGYEPVLVIQSSEDYVENTEKALNVYYEIGKILSRDILSKINQPYQKFLDVLNTIK
NASDSDGQDLLFTNQLKEHPTDFSVEFLEQNSNEVQEVFAKAFAYYIEPQHRDVLQLYA
PEAFNYMDKFNEQEINLSLEELKDQRSGRELERGADDVVDSSKSFVMENFSSYHGTKPG
YVDSIQKGIQKPKSGTQGNYDDDWKGFYSTDNKYDAAGYSVDNENPLSGKAGGVVKV
TYPGLTKVLALKVDNAETIKKELGLSLTEPLMEQVGTEEFIKRFGDGASRVVLSLPFAEG
SSSVEYINNWEQAKALSVELEINFETRGKRGQDAMYEYMAQASAGNRLPSTGGHHHHH

DTA-LPSTGG

ADDVVDSSKSFVMENFSSYHGTKPGYVDSIQKGIQKPKSGTQGNYDDDWKGFYSTDNK
YDAAGYSVDNENPLSGKAGGVVKVTYPGLTKVLALKVDNAETIKKELGLSLTEPLMEQ
VGTEEFIKRFGDGASRVVLSLPFAEGSSSVEYINNWEQAKALSVELEINFETRGKRGQDA
MYEYMAQASAGNRLPSTGGHHHHHH

K_5DTA-LPSTGG

A**KKKKK**ADDVVDSSKSFVMENFSSYHGTKPGYVDSIQKGIQKPKSGTQGNYDDDWKG
FYSTDNKYDAAGYSVDNENPLSGKAGGVVKVTYPGLTKVLALKVDNAETIKKELGLSL
TEPLMEQVGTEEFIKRFGDGASRVVLSLPFAEGSSSVEYINNWEQAKALSVELEINFETR
GKRGQDAMYEYMAQASAGNRLPSTGGHHHHHH

*The polylysine tag is highlighted in red. The N-terminal alanine was introduced to facilitate SUMO cleavage based on the manufacturer's protocol.

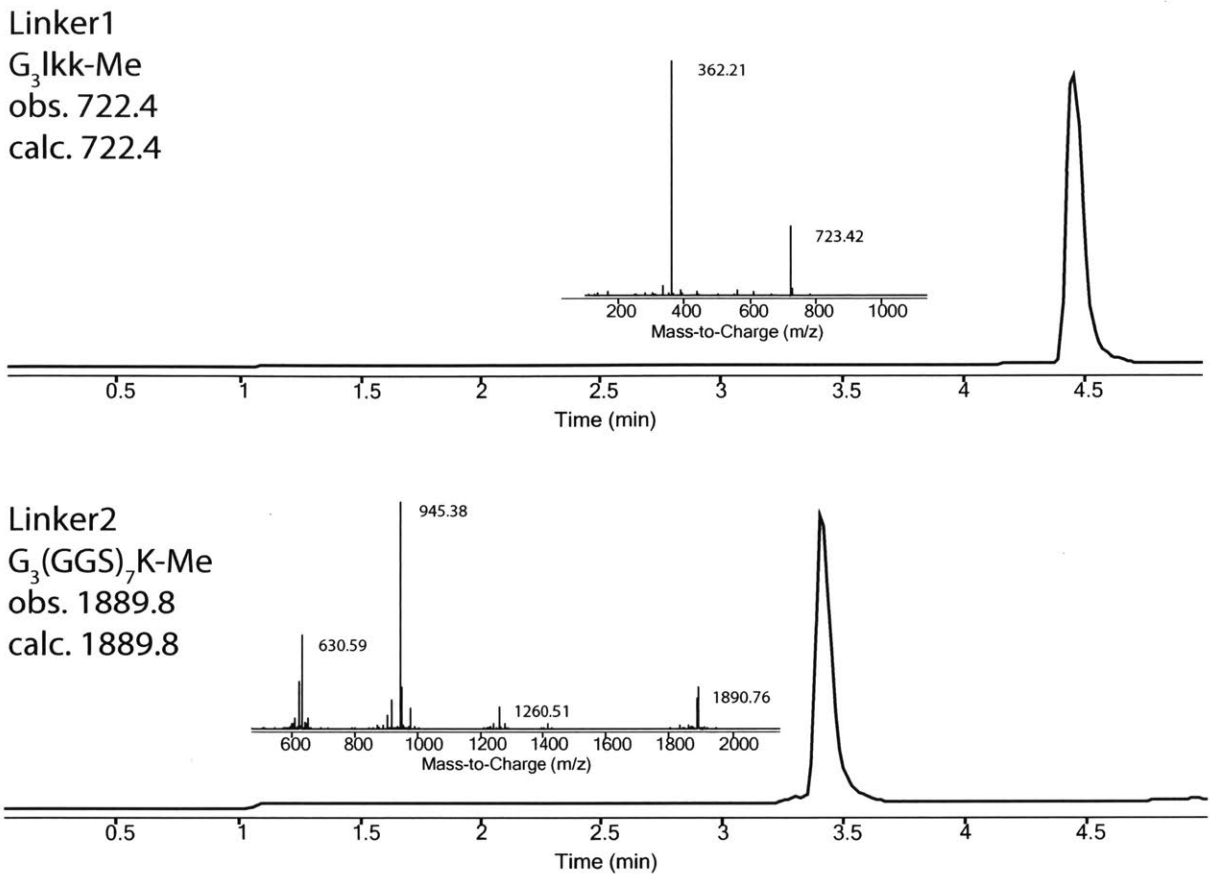

Figure 3.6.1. LC-MS traces and mass analyses of the peptide linkers used in this study. D-amino acids are indicated by the lowercase letters. About 200 ng of each peptide was injected into Zorbax 300SB C_3 column with a method of 5-65% acetonitrile over 5 minutes.

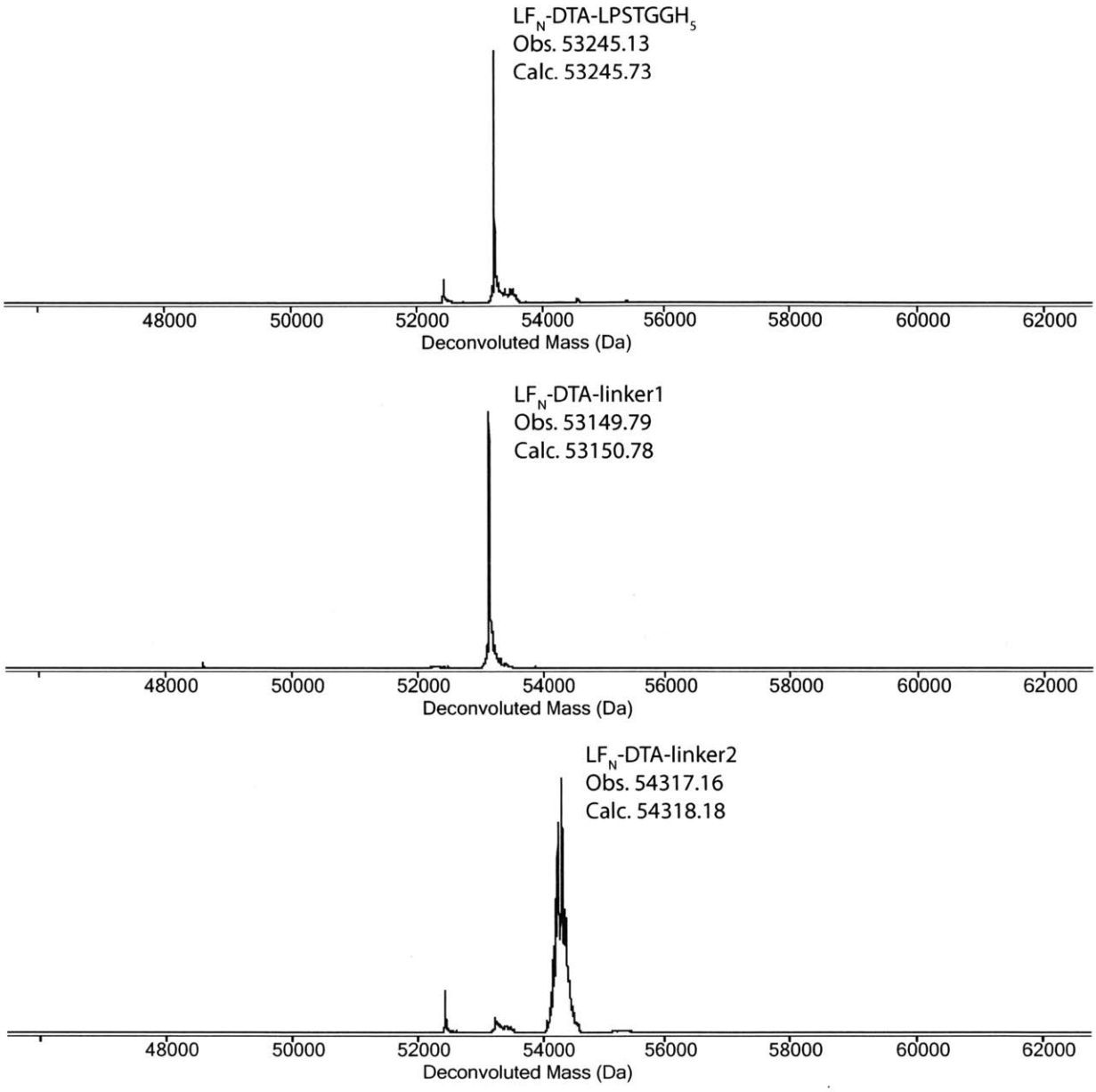

Figure 3.6.2. LC-MS deconvoluted masses of LD, LD-linker1, and LD-linker2. The peptide linkers were conjugated to LD as described in the manuscript. 200 ng of each protein was injected into Zorbax 300SB C_3 column with a method of 5-65% acetonitrile over 15 minutes. The TIC peak was integrated and the mass was deconvoluted using maximum entropy algorithm.

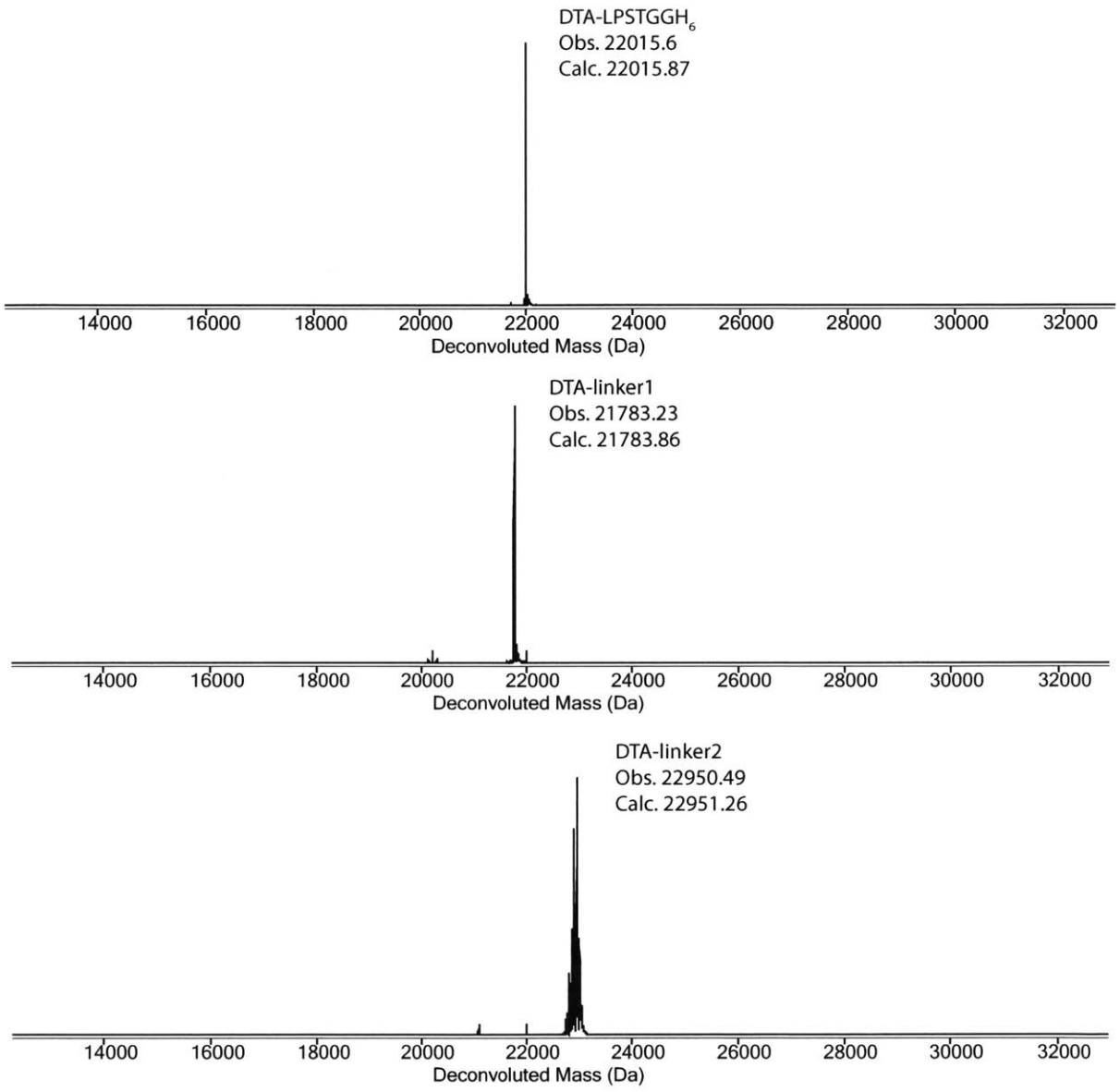

Figure 3.6.3. LC-MS deconvoluted masses of DTA-LPSTGG, DTA-linker1, and DTA-linker2. The peptide linkers were conjugated to DTA-LPSTGG as described in the manuscript. 200 ng of each protein was injected into Zorbax 300SB C_3 column with a method of 5-65% acetonitrile over 15 minutes. The TIC peak was integrated and the mass was deconvoluted using maximum entropy algorithm.

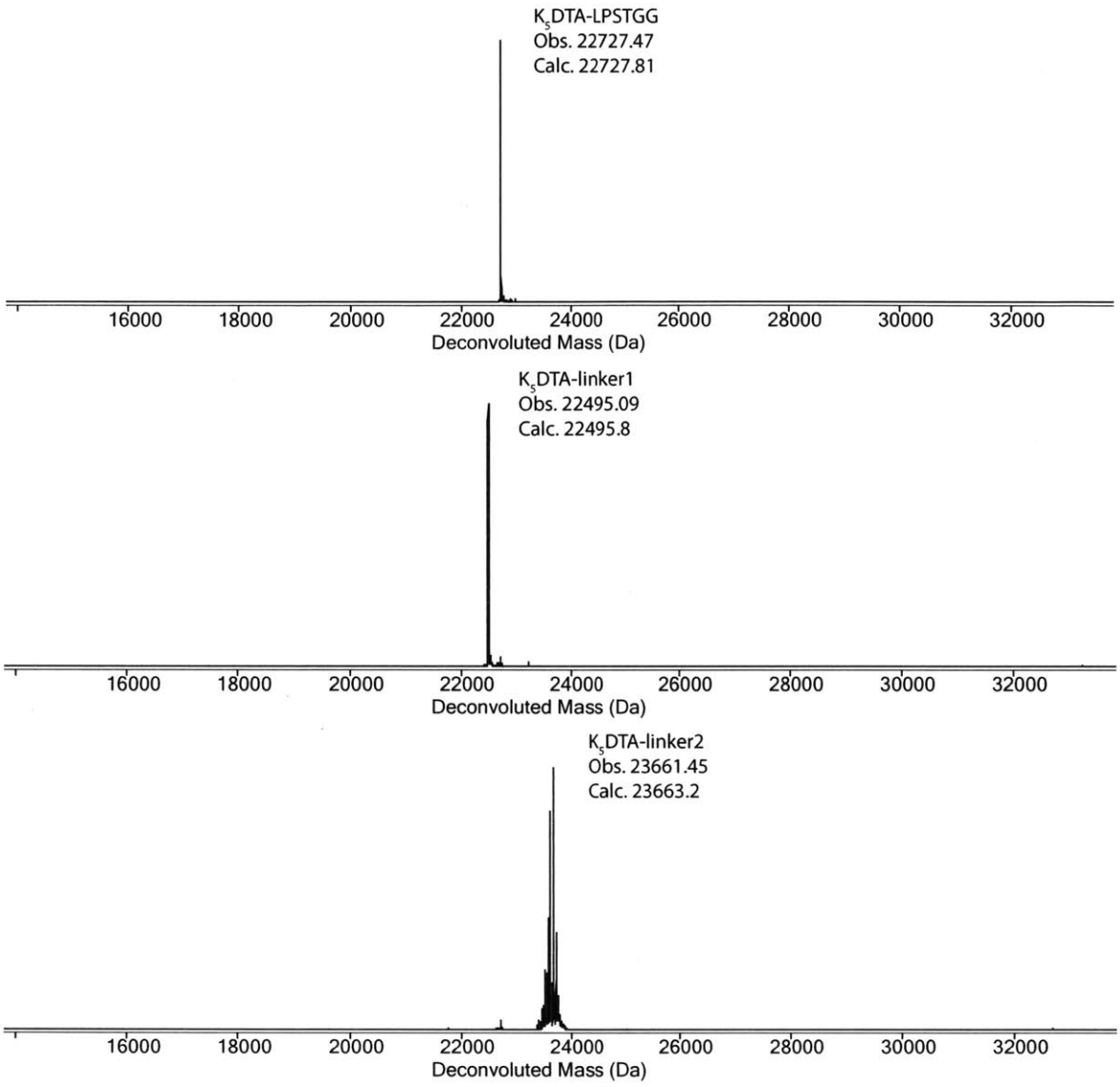

Figure 3.6.4. LC-MS deconvoluted masses of K_5DTA-LPSTGG, K_5DTA-linker1, and K_5DTA-linker2. The peptide linkers were conjugated to K_5DTA-LPSTGG as described in the manuscript. 200 ng of each protein was injected into Zorbax 300SB C_3 column with a method of 5-65% acetonitrile over 15 minutes. The TIC peak was integrated and the mass was deconvoluted using maximum entropy algorithm.

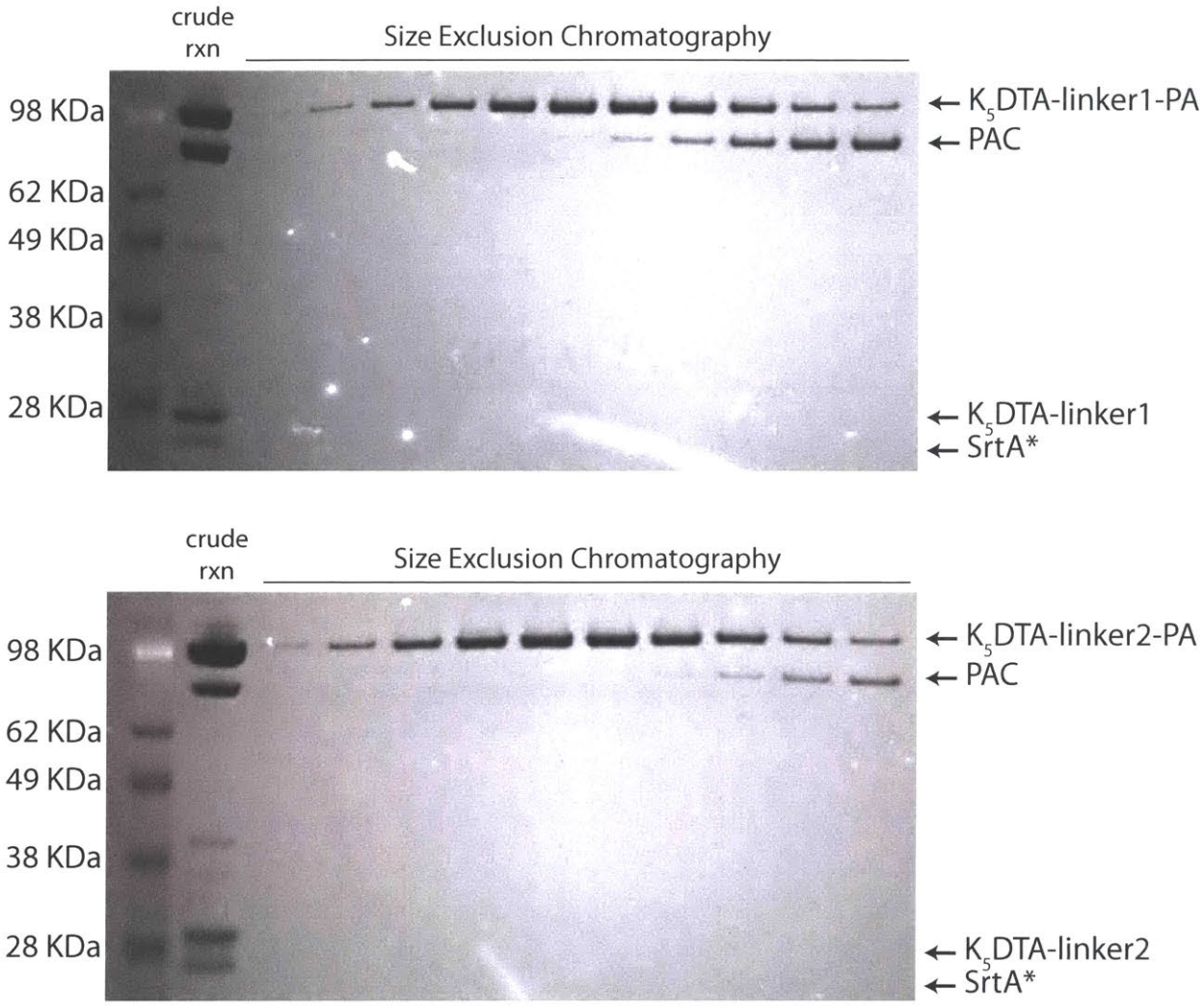

Figure 3.6.5. Preparation of K_5DTA-PA conjugates analyzed by SDS-PAGE. K_5DTA with either linker1 or linker2 was conjugated to PAC as described in the methods section. The resulted crude reaction mixture was separated by size exclusion column. The fractions were run on SDS-PAGE and stained by Coomassie blue. The pure fractions were pooled, concentrated, QCed by LC-MS, and stored in -80°C.

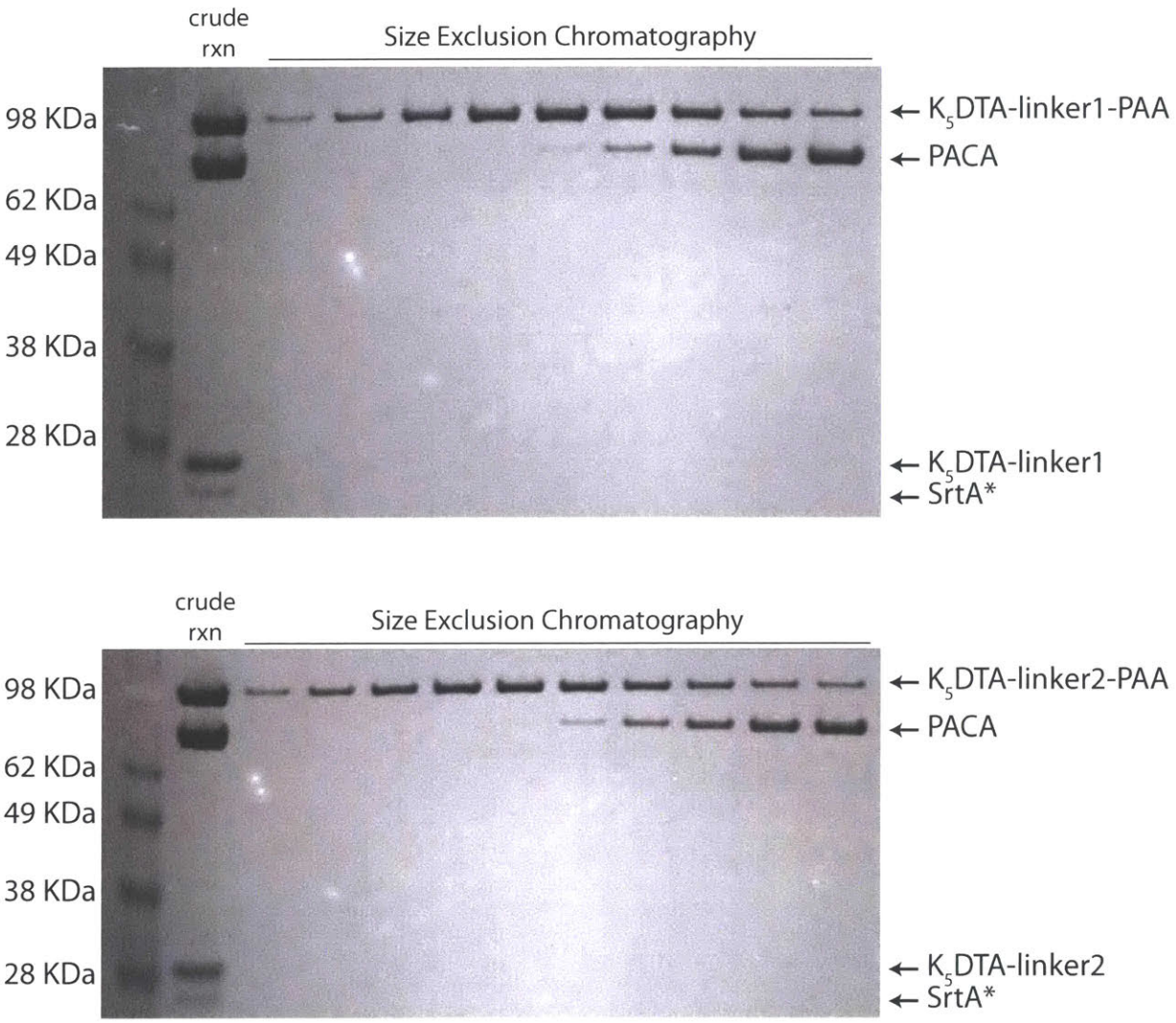

Figure 3.6.6. Preparation of K_5DTA-PAA conjugates analyzed by SDS-PAGE. K_5DTA with either linker1 or linker2 was conjugated to PACA as described in the methods section. The resulted crude reaction mixture was separated by size exclusion column. The fractions were run on SDS-PAGE and stained by Coomassie blue. The pure fractions were pooled, concentrated, QCed by LC-MS, and stored in -80°C.

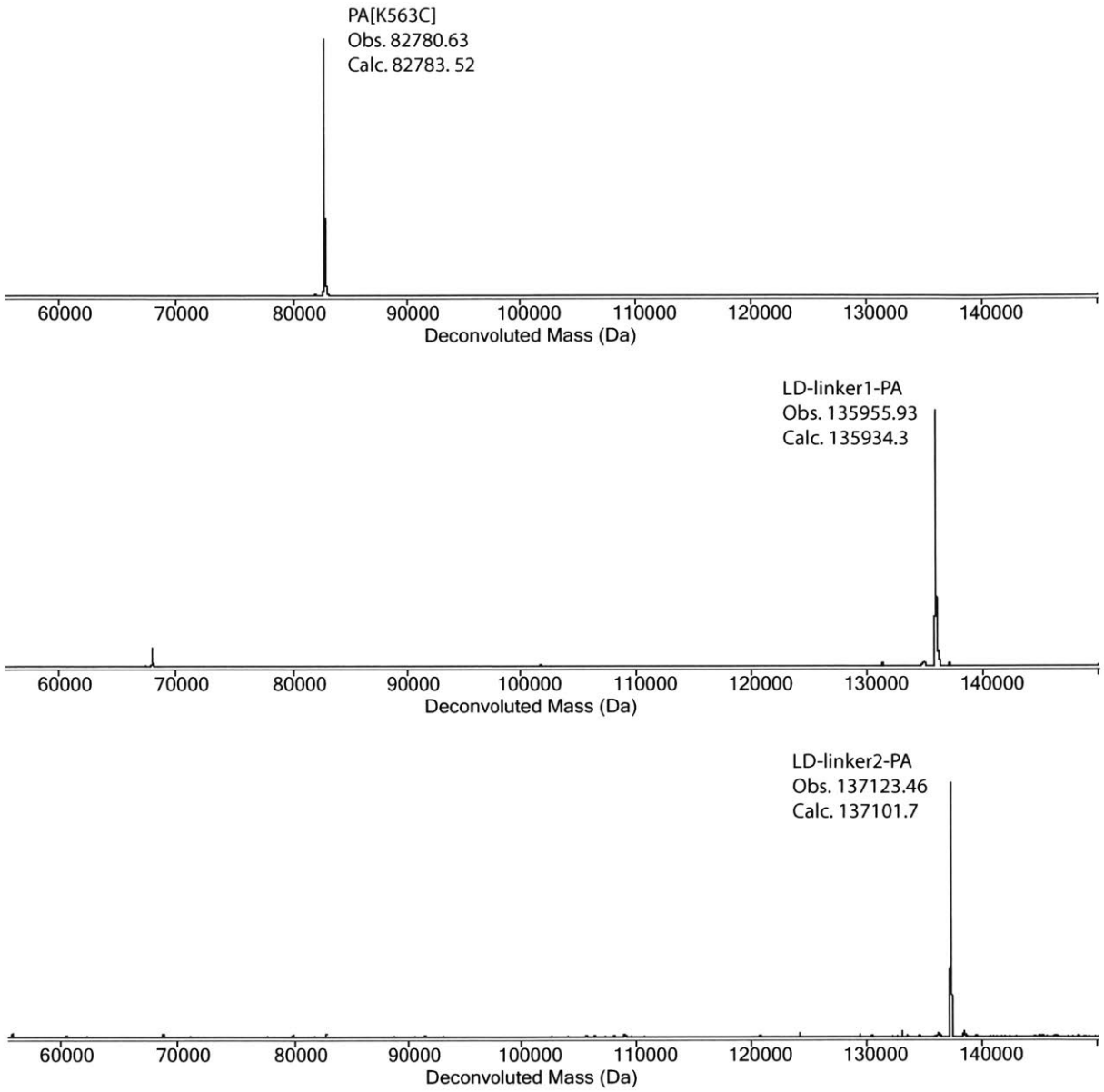

Figure 3.6.7. LC-MS deconvoluted masses of PAC, LD-linker1-PA, and LD-linker2-PA. LDs with different linkers were conjugated to PAC as described in the manuscript. 200 ng of each protein was injected into Zorbax 300SB C_3 column with a method of 5-65% acetonitrile over 15 minutes. The TIC peak was integrated and the mass was deconvoluted using maximum entropy algorithm.

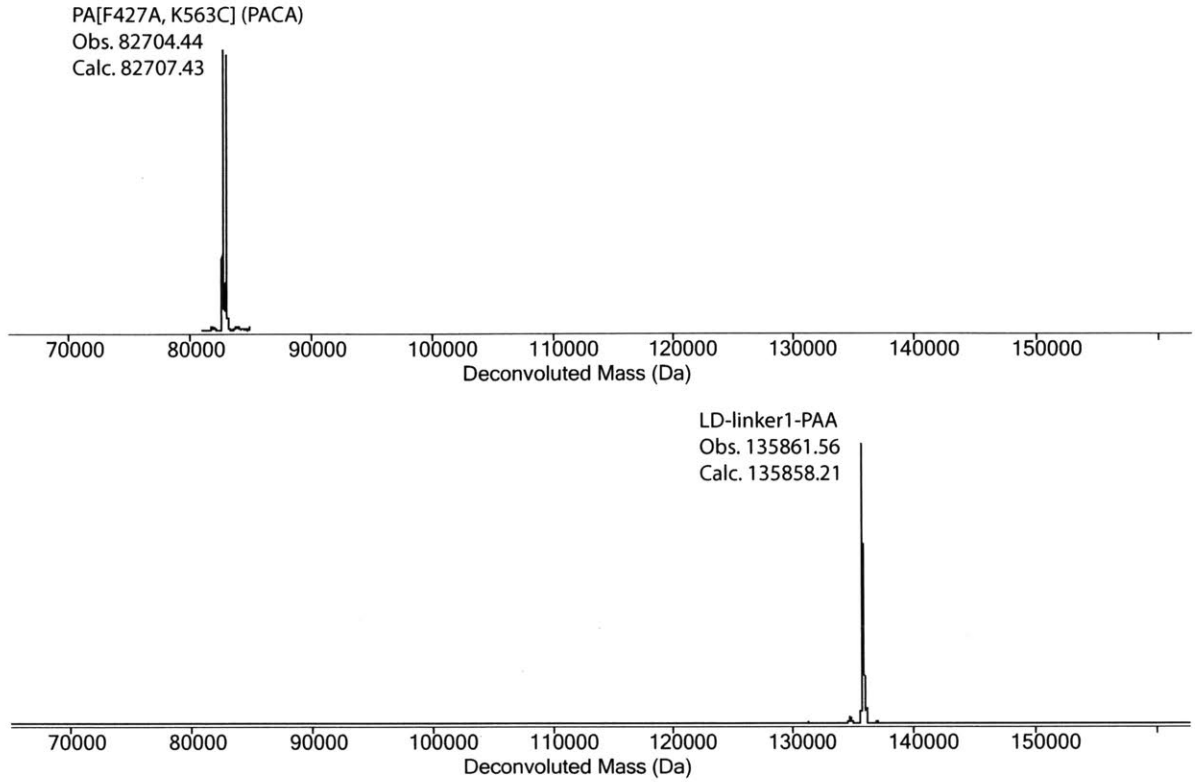

Figure 3.6.8. LC-MS deconvoluted masses of PACA and LD-linker1-PAA. PACA glutathione adduct (+305.43 Da) was also found in this batch of PACA. However, the adduct did not react with LD-linker1 due to the capped cysteine as expected. LD-linker1 was conjugated to PACA as described in the manuscript. 200 ng of each protein was injected into Zorbax 300SB C_3 column with a method of 5-65% acetonitrile over 15 minutes. The TIC peak was integrated and the mass was deconvoluted using maximum entropy algorithm.

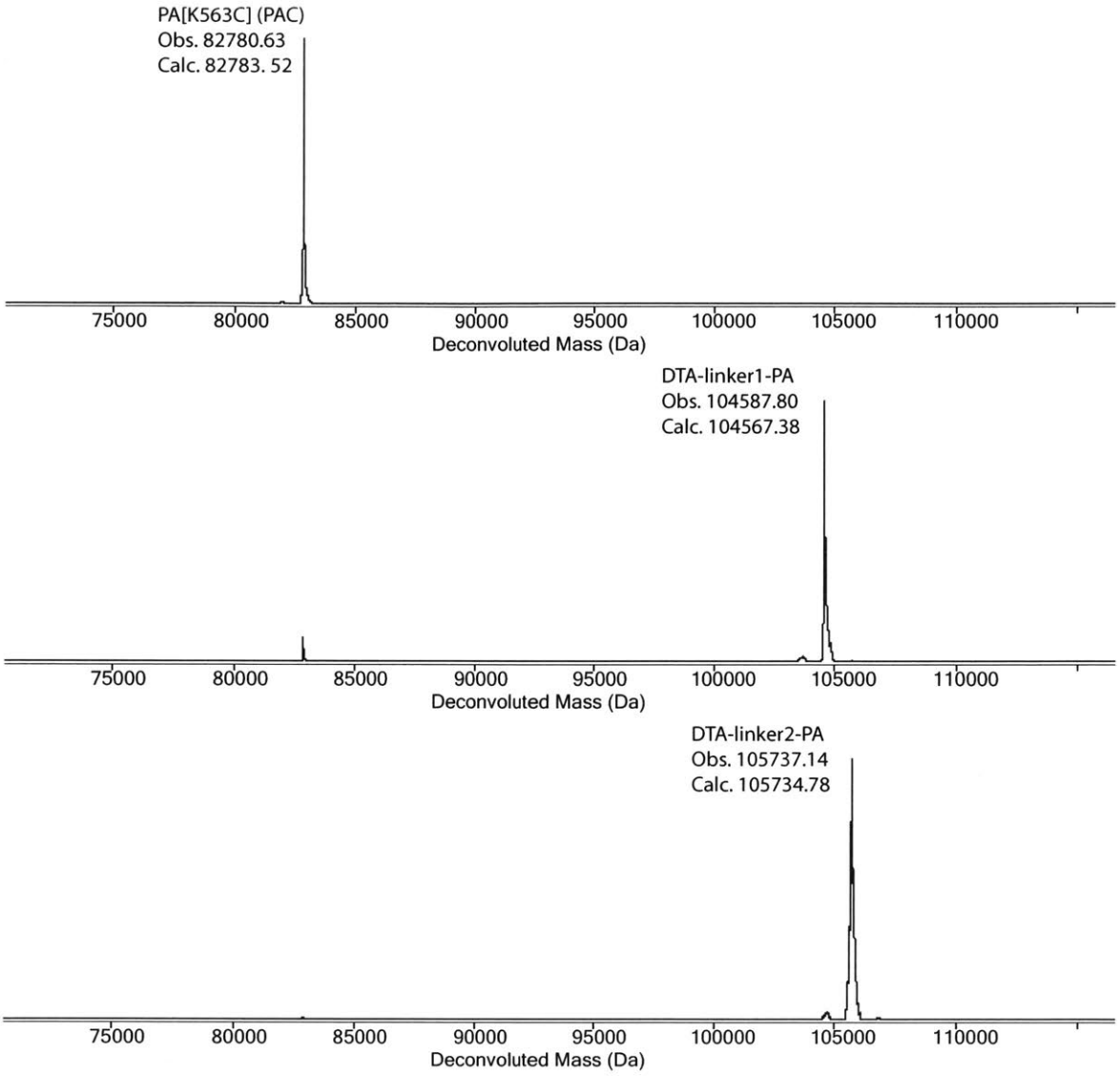

Figure 3.6.9. LC-MS deconvoluted masses of PAC, DTA-linker1-PA, and DTA-linker2-PA. DTAs with different linkers were conjugated to PAC as described in the manuscript. 200 ng of each protein was injected into Zorbax 300SB C_3 column with a method of 5-65% acetonitrile over 15 minutes. The TIC peak was integrated and the mass was deconvoluted using maximum entropy algorithm.

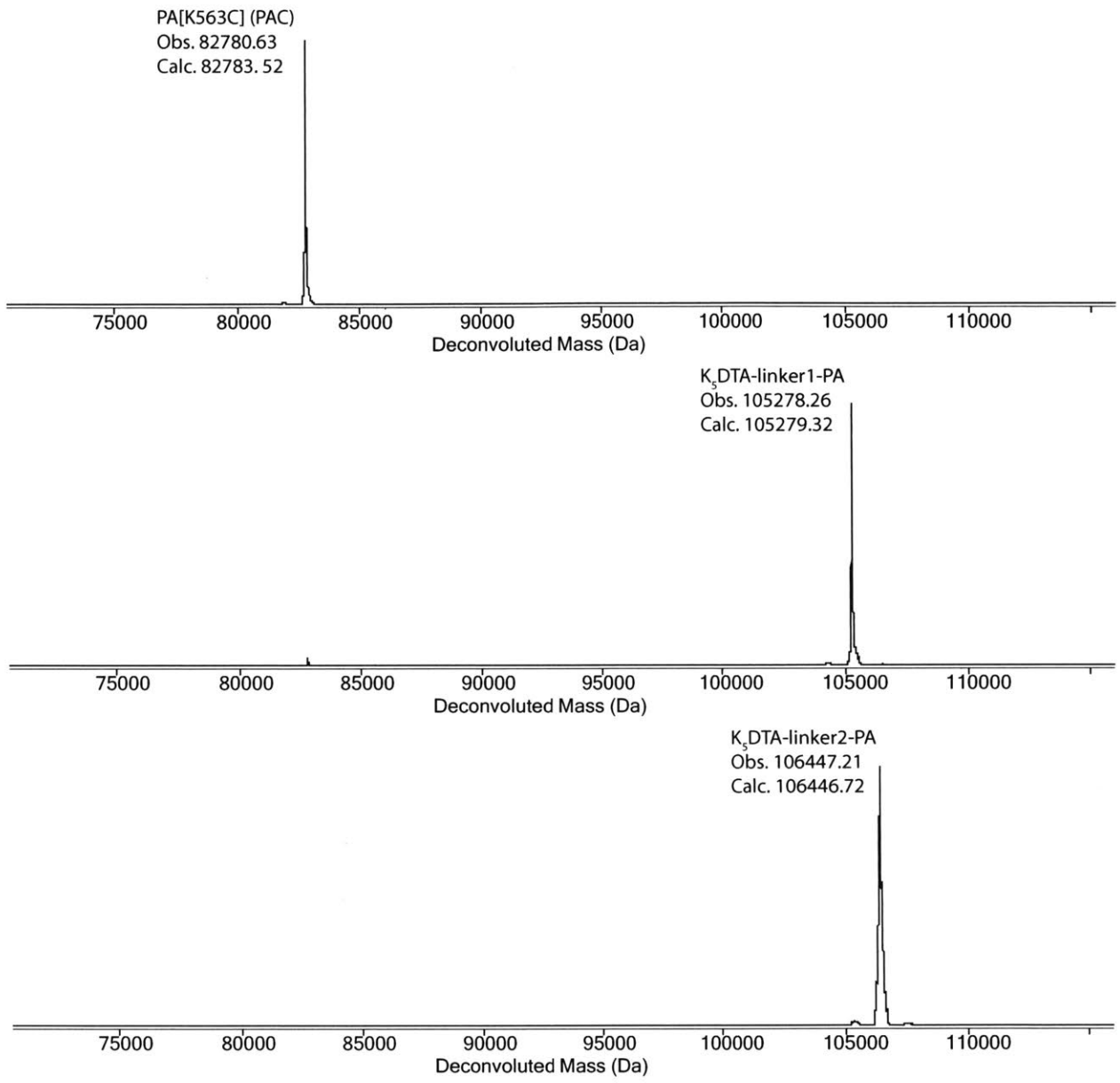

Figure 3.6.10. LC-MS deconvoluted masses of PAC, K_5DTA-linker1-PA, and K_5DTA-linker2-PA. K_5DTAs with different linkers were conjugated to PAC as described in the manuscript. 200 ng of each protein was injected into Zorbax 300SB C_3 column with a method of 5-65% acetonitrile over 15 minutes. The TIC peak was integrated and the mass was deconvoluted using maximum entropy algorithm.

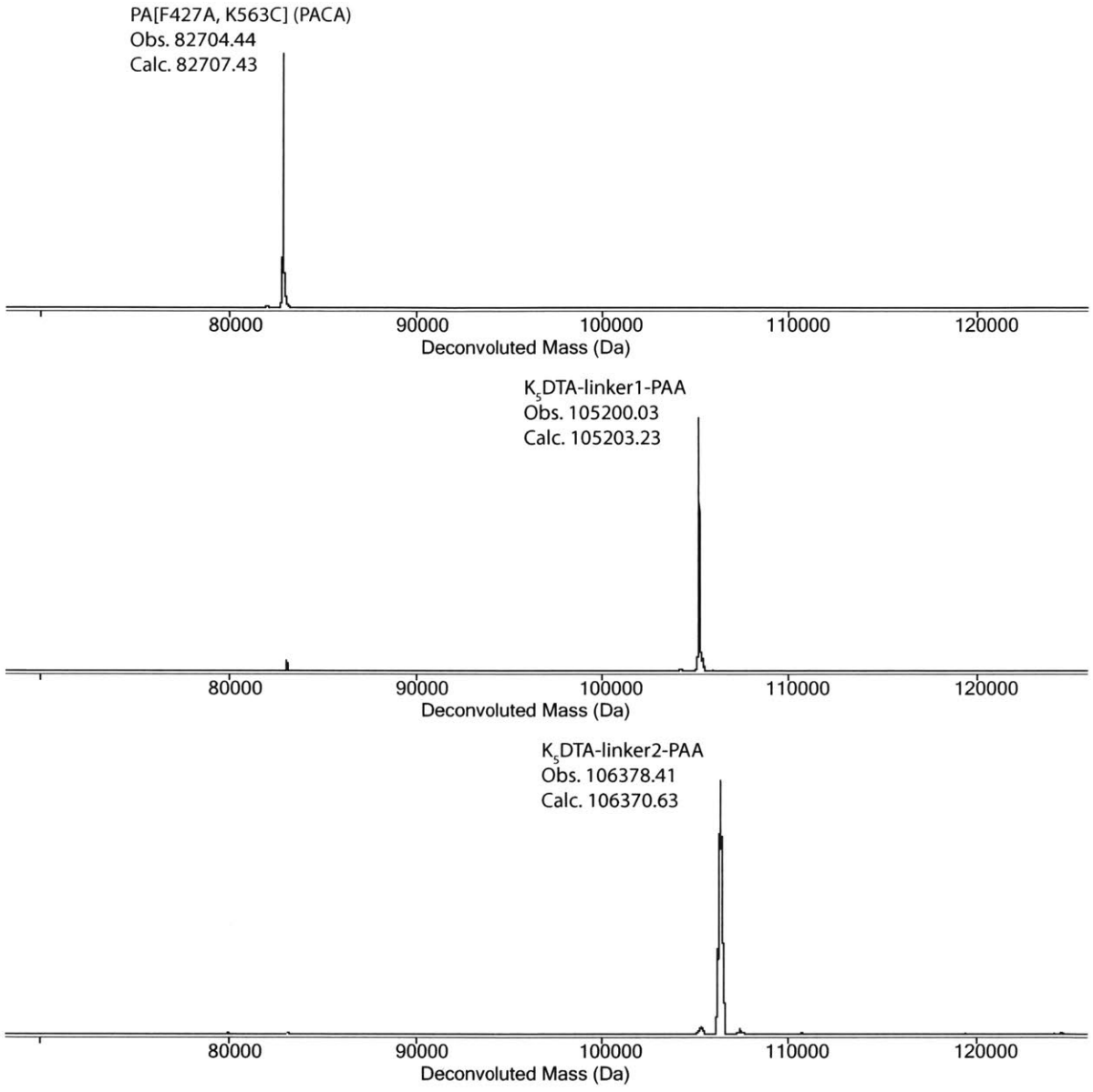

Figure 3.6.11. LC-MS deconvoluted masses of PACA, K_5DTA-linker1-PAA, and K_5DTA-linker2-PAA. K_5DTAs with different linkers were conjugated to PACA as described in the manuscript. 200 ng of each protein was injected into Zorbax 300SB C_3 column with a method of 5-65% acetonitrile over 15 minutes. The TIC peak was integrated and the mass was deconvoluted using maximum entropy algorithm.

3.7. References

1. Pannifer, A. D. *et al.* Crystal structure of the anthrax lethal factor. *Nature* **414,** 229–233 (2001).
2. Leppla, S. H. Anthrax toxin edema factor: a bacterial adenylate cyclase that increases cyclic AMP concentrations of eukaryotic cells. *Proc. Natl. Acad. Sci. U. S. A.* **79,** 3162–3166 (1982).
3. Milne, J. C. & Collier, R. J. pH-dependent permeabilization of the plasma membrane of mammalian cells by anthrax protective antigen. *Mol. Microbiol.* **10,** 647–653 (1993).
4. Scobie, H. M., Rainey, G. J. A., Bradley, K. A. & Young, J. A. T. Human capillary morphogenesis protein 2 functions as an anthrax toxin receptor. *Proc Natl Acad Sci U S A* **100,** 5170–5174 (2003).
5. Lacy, D. B., Wigelsworth, D. J., Scobie, H. M., Young, J. A. T. & Collier, R. J. Crystal structure of the von Willebrand factor A domain of human capillary morphogenesis protein 2: an anthrax toxin receptor. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 6367–72 (2004).
6. Bradley, K. A., Mogridge, J., Mourez, M., Collier, R. J. & Young, J. A. Identification of the cellular receptor for anthrax toxin. *Nature* **414,** 225–229 (2001).
7. Klimpel, K. R., Molloy, S. S., Thomas, G. & Leppla, S. H. Anthrax toxin protective antigen is activated by a cell surface protease with the sequence specificity and catalytic properties of furin. *Proc. Natl. Acad. Sci.* **89,** 10277–10281 (1992).
8. Milne, J. C., Furlong, D., Hanna, P. C., Wall, J. S. & Collier, R. J. Anthrax protective antigen forms oligomers during intoxication of mammalian cells. *J. Biol. Chem.* **269,** 20607–20612 (1994).
9. Kintzer, A. F. *et al.* The Protective Antigen Component of Anthrax Toxin Forms Functional Octameric Complexes. *J. Mol. Biol.* **392,** 614–629 (2009).
10. Mogridge, J., Cunningham, K. & Collier, R. J. Stoichiometry of anthrax toxin complexes. *Biochemistry* **41,** 1079–1082 (2002).
11. Feld, G. K. *et al.* Structural basis for the unfolding of anthrax lethal factor by protective antigen oligomers. *Nat. Struct. Mol. Biol.* **17,** 1383–1390 (2010).
12. Nassi, S., Collier, R. J. & Finkelstein, A. PA63 channel of anthrax toxin: An extended ??-barrel. *Biochemistry* **41,** 1445–1450 (2002).
13. Krantz, B. A., Trivedi, A. D., Cunningham, K., Christensen, K. A. & Collier, R. J. Acid-induced unfolding of the amino-terminal domains of the lethal and edema factors of anthrax toxin. *J. Mol. Biol.* **344,** 739–756 (2004).
14. Miller, C. J., Elliott, J. L. & Collier, R. J. Anthrax protective antigen: Prepore-to-pore conversion. *Biochemistry* **38,** 10432–10441 (1999).
15. Duesbery, N. S. *et al.* Proteolytic inactivation of MAP-kinase-kinase by anthrax lethal factor. *Science (80-.).* **280,** 734–737 (1998).
16. Hu, H. & Leppla, S. H. Anthrax toxin uptake by primary immune cells as determined with a lethal factor-beta-lactamase fusion protein. *PLoS One* **4,** e7946 (2009).
17. Milne, J. C., Blanket, S. R., Hanna, P. C. & Collier, R. J. Protective antigen-binding domain of anthrax lethal factor mediates translocation of a heterologous protein fused to its amino- or carboxy-terminus. *Mol. Microbiol.* **15,** 661–666 (1995).
18. Arora, N., Klimpel, K. R., Singh, Y. & Leppla, S. H. Fusions of anthrax toxin lethal factor to the ADP-ribosylation domain of Pseudomonas exotoxin A are potent cytotoxins which are translocated to the cytosol of mammalian cells. *J. Biol. Chem.* **267,** 15542–15548

(1992).
19. Antic, I., Biancucci, M., Zhu, Y., Gius, D. R. & Satchell, K. J. F. Site-specific processing of Ras and Rap1 Switch i by a MARTX toxin effector domain. *Nat. Commun.* **6,** (2015).
20. Liao, X., Rabideau, A. E. & Pentelute, B. L. Delivery of antibody mimics into mammalian cells via anthrax toxin protective antigen. *ChemBioChem* **15,** 2458–2466 (2014).
21. Rabideau, A. E., Liao, X. & Pentelute, B. L. Delivery of mirror image polypeptides into cells. *Chem. Sci.* **6,** 648–653 (2015).
22. Sharma, O. & Collier, R. J. Polylysine-mediated translocation of the diphtheria toxin catalytic domain through the anthrax protective antigen pore. *Biochemistry* **53,** 6934–6940 (2014).
23. Blanke, S. R., Milne, J. C., Benson, E. L. & Collier, R. J. Fused polycationic peptide mediates delivery of diphtheria toxin A chain to the cytosol in the presence of anthrax protective antigen. *Proc Natl Acad Sci U S A* **93,** 8437–8442 (1996).
24. Thompson, D. B., Villaseñor, R., Dorr, B. M., Zerial, M. & Liu, D. R. Cellular uptake mechanisms and endosomal trafficking of supercharged proteins. *Chem. Biol.* **19,** 831–843 (2012).
25. Lyon, R. P. *et al.* Self-hydrolyzing maleimides improve the stability and pharmacological properties of antibody-drug conjugates. *Nat. Biotechnol.* **32,** 1059–1062 (2014).
26. Mijalis, A. J. *et al.* A fully automated flow-based approach for accelerated peptide synthesis. *Nat. Chem. Biol.* **13,** 464–466 (2017).
27. Simon, M. D. *et al.* Rapid flow-based peptide synthesis. *ChemBioChem* **15,** 713–720 (2014).
28. Chen, I., Dorr, B. M. & Liu, D. R. A general strategy for the evolution of bond-forming enzymes using yeast display. *Proc. Natl. Acad. Sci.* **108,** 11399–11404 (2011).