The Discovery and Chemical Synthesis of Peptides and Proteins that Cross Biological Barriers

By
Charlotte E. Farquhar

Bachelor of Science in Chemistry
Duke University, Durham, NC, 2016

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 2023

Signature of Author……
Charlotte E. Farquhar
Department of Chemistry
08/03/2023

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

Accepted by……...……
Adam Willard
Associate Professor
Graduate Officer

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

Ronald T. Raines
Roger and Georges Firmenich Professor of Natural Products Chemistry
Thesis Committee Chair

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

Laura L. Kiessling
Novartis Professor of Chemistry

The Discovery and Chemical Synthesis of Peptides and Proteins that Cross Biological Barriers

By
Charlotte E. Farquhar

Submitted to the Department of Chemistry on July 20, 2023
In Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy in Chemistry

Abstract:

The precise delivery of therapeutic agents to specific tissues or cells remains a complex task. Cell-penetrating peptides (CPPs) and proteins are used as carrier modules, tissue targeting agents, and probes to better understand disease states. In this work, we first investigate CPPs for peptide-mediated delivery of macromolecules. Despite the existence of thousands of CPP sequences, none have been approved by regulatory authorities for macromolecule delivery. Here, we demonstrate a method for in-cell penetration selection-mass spectrometry (in-cell PS-MS) to discover novel peptides from a synthetic library. These all-D, non-canonical peptides can deliver macromolecule cargo to the cytosol with high peptide stability and low toxicity. In-cell PS-MS introduces a method to discover unnatural synthetic peptides to delivery therapeutically relevant cargo into subcellular compartments.

The chromatographic enrichment of a peptide library through cation-exchange chromatography also generates novel, non-canonical peptide sequences for delivery of macromolecular cargo. We developed a method based on testing fractionated or pooled peptide libraries for the discovery of novel peptides with low toxicity, which can be modified to produce novel CPPs for oligonucleotide delivery with efficacy at nanomolar concentrations.

Cell penetrating peptides have also previously shown efficacy in crossing the blood-brain barrier (BBB), allowing a small, soluble peptide (BTP-7) to cross the BBB and target human glioblastoma (GBM). We demonstrate that conjugation of BTP-7 to camptothecin improves drug solubility in aqueous solutions, retains drug efficacy against patient-derived GBM stem cells, enhances BBB permeability, and enables therapeutic targeting to intracranial GBM, leading to higher toxicity in GBM cells compared to normal brain tissues, and ultimately prolongs survival in mice bearing intracranial xenografts of patient-derived GBM.

Dipeptide repeat proteins (DPRs), are characteristic in amyotrophic lateral sclerosis and frontotemporal dementia. Synthesis of these proteins by biological or chemical means has previously proved difficult, due to their toxicity and propensity towards aggregation. Herein we report the chemical synthesis of four DPRs using automated fast-flow peptide synthesis. Structural assays demonstrate in-vitro aggregation, especially in proline-rich DPRs. Human neuroblastoma cells cultured with arginine-rich DPRs with longer repeat lengths show reduced cell viability, thereby reproducing the cytotoxic property of endogenous DPRs. This research demonstrates the potential of automated flow synthesis for synthesizing difficult-to-access proteins and gives us new insight into the behavior of arginine-rich DPRs at the cell membrane. Collectively, this work demonstrates successful approaches to address the cellular delivery of therapeutic modalities for the future development of novel therapeutics.

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

Acknowledgements

All of the research reported here was made possible by Professor Brad Pentelute. Brad, thank you for creating a science playground, for giving us to tools and resources to create novel ideas, and for fostering a lab full of collaboration and mutual support. I have grown so much under your mentorship, both as a scientist and as a person, and I am so glad to have gotten this opportunity to work in your lab. Andrei, nothing is possible without you. Thank you for keeping everything moving, helping me fix every instrument in lab, and turning my writing into something readable.

Thank you to my thesis committee chair, Professor Raines, for five years of support and mentorship. You always have helpful recommendations on the next steps for these projects whenever I feel stuck. (I also still make our first-year graduate students read through the notes from your ChemBio tutorial class). Thank you also to Professor Kiessling, both for being a member of this committee and for becoming a mentor and role model for me. I admire you for your work as both an outstanding professor and a champion of DEI.

This work would have been impossible without Carly Scissel, who was a mentor to me when I first joined the lab and remains a good friend even now. Carly and Beck Holden taught me about CPPs and how to make them, and they helped shape my proposed project (Chapter 5). More than that, Carly taught me everything from proposing a project to writing a manuscript, and I could have done none of this without her. I cannot think of a better person to become a chemistry professor in the future.

On a similar note, Professors Choi-Fong Cho and Kohei Sato have been wonderful mentors in the project areas of Glioblastoma targeting and di-peptide repeat proteins, respectively. The research I have done with them (Chapters 3 and 4) was both exciting and impactful, and I am grateful to them both for welcoming me onto their teams. I am especially grateful to Fong for teaching me how to write grants and collaboratively develop a project into a new space.

While this thesis is submitted under my name, this research would have been impossible without the contributions of the many co-authors of these works and others. I want to especially mention Carly Schissel, Choi-Fong Cho, Kohei Sato, Jacob Rodriguez, Alex Callahan, Chengxi Li, Coralee Backlund, Beck Holden, Jorge Jimenez-Macias, Ed Miller, Yenchun Lee, Niklas von Spreckelsen, Benjamin Scott, Colin Fadzen, Justin Wolfe, Martin Gazvoda, Heemal Dhanjee, Nick Truex, Joseph Brown, Michael Lee, Sebastian Pomplun, Chris Shugrue, Max Distler, Nick Vecchiarello, Xuyu Tan, and the rest. Thank you all for five years of collaboration and learning.

I could not have survived 5 years of graduate school without the constant cheery, silly, and (occasionally) deeply loving support of my friends and colleagues in the Pentelute group. When I came to the group, I was immediately welcomed with open arms and supported by the friendship and mentorship of Carly, Beck, Azin, and Nina. These four were both amazing and inspiring, and without them to catch me when I fell, I might not have stood back up.

I looked forward to coming into lab, even when the science wasn't going well, because of the large circus of clowns parading by my desk every day. To all of the clowns in the Pentelute lab, and especially my beloved 7-11 crew Dio, Joeseph-moseph, Michael-wave, Jakey-snakey, Heemal, Azine Sabeti, A-aron, Sarah, Amanda, Corgytha, Ghanathan, Wayniel, Nedward, Maxabillion, Nate-the-great, Stevo, Tombo, Xiyun, Yehlin, Maysa, David, Ching-pei, and Simon the frog-man. And of course to all of those who keep us serious, especially Alex, Nina, Aurelie, and Nathalie. Thank you for being there for me when I need it most. Thank you especially to Nathalie and Ed, who read through this monstrosity and kindly edited it into something useful.

I want to acknowledge all of the people I've had the pleasure of on-boarding, training, and/or mentoring. I've called you my lab family, and you might have thought I was kidding, but I promise I really value you all to that extent. To my beloved mentees Yehlin, Gha, and Maysa, and to my adopted post-doc sons Ed, Max, Shuai, and baby Nate: I am so excited to see the amazing science you all will generate over your careers, and I consider myself quite lucky that I have gotten to be a small part of it. Thank you all, as well, for all the help you've contributed to this thesis, through both emotional and research support.

I would like to recognize my good friends Sloane and Ellen (and, I suppose, their husbands Jacob and Alex). The stupid lists, the bizarre charts, the temporary pet spiders George I - George VII, and many Saturday rantings created a "Capital F" friendship. You represent four big reasons that I am willing to endure more Boston winters and stay up here in the frigid north after graduation.

My first year at MIT wasn't nearly as bad as I thought it would be due to the presence of an amazing first-year cohort. To Natalie, Mikaila, Katherine, Gino, Cliff, Nile, Levi, Keith, Victoria, JoLynn, Griff and the rest, spending the last 5 years with you all has, alone, made the MIT PhD worth the effort. Especially Katherine, Brant, Amy, and Alex, who had the terrible misfortune to live with me. And Gino, who for some reason still does. Thank you for giving me something wonderful to come home to every day.

Throughout my time at MIT I have also been heavily involved in a number of student groups. Thank you Jesús, Tony, and Francesca for welcoming me into CADI with open arms, and thank you to all CADI members, especially Valerie, Gisele, and Amanda, for keeping us going. To everyone I've worked with in WIC+, CGSC, and the DEIC, thank you all so much for letting me become a part of something so much bigger than myself. I hope that, together, we've created a supportive and inclusive community here. I really believe that the work we've done together will help future students find a great place here at MIT. Levi, Katherine, Azin, Francesca, Nile, Spencer, Angie, Kayla, and Stephanie, especially, thank you for everything you've done for our department.

I have to thank all of my Chi Omega sisters from Duke, and especially those of you who still persist in texting me despite my demonstrated reputation as a terrible texter. I'm not exactly a "sister" anymore, but I still feel just as close to you all. Katya, our friendship has made me so much stronger (both literally and figuratively). Caroline, our Marvel movie weekends have been essential in forcing me to actually take a day off. Kyla, our Sunday pilates and knitting keeps me sane (and in great shape!) even from across the country. Chase, I am still eagerly awaiting our dinosaur-themed wedding in the future. Emily, we'll definitely make it back to the beach someday (once I am medically cleared to fly, probably). And Gracie, even though you chose CalTech over MIT, I still can't wait to see the amazing professor you will become.

I literally would not be here without my parents Will and Michele Farquhar. Mom and Dad, thank you for being there for me, for supporting me when I'm down, and for encouraging me to always push forward. Jamie-Tom and Doug'n'Ali, you are the best siblings anyone could ask for. To baby Jack Farquhar, welcome to a wonderful and wacky family. We all love you so much!

Finally, I want to give a strange sort of thanks to Corey Z., who hit me with his Toyota Corolla one dark night in October. After months of rehabilitation I am still not fully recovered, but I have certainly learned how lucky I am to be at a supportive place like MIT chemistry, and especially the Pentelute Group. The kindness and patience everyone has shown me in over the past year has been an incredible reminder of how wonderful our MIT community can be, and I will be forever grateful to Professor Brad Pentelute for his constant and enduring support during my most difficult time.

Table of Contents

List of Figures

Chapter 1. Background and Overview

1.1 The development of peptides as therapeutics

Since initial studies on peptide hormones such as oxytocin, insulin and vasopressin, peptides have been of great interest for their therapeutic value.[1] The first chemical synthesis of oxytocin by solution-phase synthesis represented a milestone in our ability to synthetically access these peptides.[2] Due to the inherent difficulties in solution-phase peptide synthesis, it took another twelve years before a biologically active form of the longer peptide hormone insulin was successfully synthesized.[3] Merrifield's development of solid-phase peptide synthesis, recognized with the Nobel Prize in Chemistry in 1984, reduced the time required for the assembly of insulin from months to weeks, marking a turning point in our ability to access peptides for therapeutic purposes.

Subsequent innovations in peptide synthesis, including the development of Fmoc chemistry, acid-labile linkers such as Rink, and solid supports with improved handling properties, have since facilitated the synthetic preparation of peptides.[4,5] Of particular relevance to this body of work, the development of flow-based peptide synthesis, both in semi-automated and automated setups, has opened the door to peptide and protein synthesis at a higher purity and on a shorter time scale than previous methods.[6–8]

Taken together, these advances in peptide synthesis increase the convenience of generating peptides for therapeutic exploration. With the increasing synthetic accessibility of amide bonds, the therapeutic scope of peptides has also expanded. Apart from several analogs of insulin, more than 60 peptide drugs have been approved around the world. These drugs range from synthetic peptide hormones to sequences bearing non-natural modifications such as liraglutide.[9,10] The therapeutic areas of these approved drugs include modulating the endocrine system, interfering with protein-protein interactions, and antimicrobial activity.[1,9–11] Among the peptide therapeutics still in clinical development there is an even wider range of applications, which include delivery of small molecules and macromolecular drugs both through tissue targeting and by crossing biological barriers.[12–14]

A major challenge for the successful translation of preclinical peptide candidates is associated with their limited bioavailability, often due to insufficient stability in vivo or the inability to reach the target tissue.[15–17] The proteolytic degradation of peptide drugs through intra- and extra-cellular proteases limits the half-life of peptides and thus quickly reduces the overall drug concentration. Strategies to make peptides more resistant against proteolytic degradation include protection of the N- and C-termini, substitution of unnatural or non-canonical amino acids, backbone modifications, or reversal of chirality.[15,16,18–21] Among these strategies, the preparation of peptides incorporating D-amino acids is one of the most effective methods to increase proteolytic resistance.[22] Throughout this body of work, we have applied these strategies, especially the synthesis of peptides with non-canonical and/or D-amino acids, to develop peptides with a higher resistance to proteases.

There are three main strategies to address the limited availability of peptide drugs in their target tissues: increasing peptide circulation time, modifying the peptide to become cell-permeable, or targeting the peptide to its destination tissue. Peptide circulation time can be extended by increasing the overall molecule size or via binding to serum proteins to diminish renal clearance.[18] Liraglutide represents a clinical success of this model: The fatty acid chain appended to the GLP1-like peptide extends the peptide half-life in vivo through reversible binding to albumin.[10]

Alternatively, peptides that can cross the cell membrane, also known as cell-penetrating peptides (CPPs), can more rapidly enter biological tissues before proteolytic cleavage or renal clearance.[23] Peptides ranging from cationic helices to non-polar cyclic peptides have the ability to enter cells, and increasing the cell permeability of peptide drugs is a highly active area of research.[23–28] Finally, targeting a peptide to a specific tissue can increase the likelihood of that peptide exiting the bloodstream and entering a biological tissue of interest.[12,29] Peptide cell permeability will be the focus of the remaining sections. proteins that can cross biological barriers represents the bulk of the work presented herein.

1.2 Cell-penetrating peptides

As the interest in therapeutic peptides increases, there is an increasing need for peptides that can cross the plasma membrane, either to increase bioavailability in specific tissues or to reach a target within the cytosol. In the late 1980s, research into proteins that can cross the plasma membrane revealed protein transduction domains in the Tat protein of HIV-1 and the Antennapedia homeodomain, which were truncated into the two peptides TAT and penetratin, respectively.[30–32] These became the first cell-penetrating peptides (CPPs), which are herein defined as peptides of 3-60 residues that can cross the plasma membrane with or without an attached cargo molecule.

Since the original discovery of penetratin and TAT from protein domains, more than one thousand different CPP sequences have been discovered.[33] Many of these peptides were similarly derived from natural peptides or proteins, such as CPPs isolated from heparin-binding proteins,[34] DNA-binding proteins,[35,36]homeoproteins,[37,38] and viral proteins responsible for cell entry.[39,40] In addition to those sequences found within larger proteins, some CPPs are directly derived from natural peptides. Antimicrobial peptides (AMPs), which are found in venoms, toxins, and saliva represent natural CPPs, although many can damage the plasma membrane during translocation.[41]

In the search for effective CPPs, many candidates emerged from rational design of natural CPPs such as the "penetramax" peptide derived from penetratin and a number of AMPs modified for reduced toxicity.[42,43] Modifications can further include creating chimeric peptides that combine CPPs, signaling peptides, and/or localizing peptides for synergistic action.[44,45] Further CPP design efforts include using synthetic molecular evolution to combine existing peptides[46,47] or prediction of novel cell-penetrating peptides based on the properties of the existing peptides.[48–52]

Finally, and most relevant to this work, CPPs can be discovered de novo from large libraries. Phage display and related technologies have been used to successfully discover novel CPPs.[53,54] While these are high-throughput methods for CPP discovery, these display methods often lack the ability to conjugate the cargo of interest directly to the CPPs during screening. The methods for CPP discovery described in Chapters 1 and 2 focus on screening synthetic peptide

libraries conjugated to the cargo of interest, presenting the next step in discovering CPPs from combinatorial libraries.

Despite the large number and variety of existing CPPs, discovery efforts are still ongoing as the cell penetration of a CPP is highly dependent on the attached cargo. The CPPs discovered for small-molecule delivery may not effectively deliver larger cargoes,[51,55,56] and CPP efficacy may differ even between various larger cargoes such as anti-sense oligonucleotides.[52] The properties of the cargo can affect the propensity of a CPP towards different cell entry mechanisms, which in turn alters the overall cell penetration of the CPP-cargo conjugate. [57,58]

CPPs can enter cells through energy-dependent or energy-independent mechanisms. A given CPP can often enter through either method, depending on the concentration and the cargo attached. [57,59] The energy-independent mechanisms are particularly common for uncharged, cyclic peptides like cyclosporin. In these cases, the peptides are able to directly permeate the cell membrane much like smaller, more lipophilic drug-like molecules.[25] For highly charged or amphipathic CPPs, energy-independent uptake more often results from interactions between the cationic residues and negatively charged phospholipid head groups and membrane sugars.[60,61] Cationic, energy-independent entry can also cause membrane disruption, involving the formation of particle-like compounds of CPP and membrane components,[62] or the formation of pores in the plasma membrane.[63] Energy-dependent cell penetration relies on various forms of endocytosis into the cell. Endocytosis can be receptor mediated, as in the case of clathrin or caveolin pits or peptides that rely on receptor internalization and recycling. Clathrin-mediated endocytosis is a common entry mechanism for CPP-oligonucleotide conjugates, and thus highly relevant for the work described in the following chapters.[44,64,65] Receptor-independent mechanisms, such as micropinocytosis and lipid-raft-mediated endocytosis, are additional entry mechanisms for CPP conjugates.[27] Interactions with membrane sugars, such as heparin sulfate, are likely to facilitate CPP clustering on the cell membrane to aide in triggering endocytosis.[66]

Once endocytosed, peptides in the endosomes can face challenges such as rapid degradation by proteases and sequestration from the cytosol. [67] This leads to the next important factor for cell entry: endosomal escape. Especially for those peptides or peptide-drug conjugates

that must reach the cytosol or nucleus for therapeutic effect, endosomal escape can be just as crucial as endocytosis. Endosomal escape can involve direct disruption of the endosomal membrane, such as the pore-formation mechanism discussed above. There is additional evidence to suggest that the negatively charged phospholipid bis(monoacylglycero)phosphate, which is enriched in the late-stage endosome, may interact with CPPs to induce lipid mixing and membrane leakage.[68] The low pH of the endosome, especially in the late stage, may also explain why these basic CPPs are more likely to disrupt the endosomal membrane than the plasma membrane.[69] Depending on the cargo and entry mechanism, CPPs can cause toxicity, often due to disruption of the plasma membrane.[41,70,71] This is especially present in AMP-derived CPPs, many of which are known to lyse eukaryotic cells.[41] Cationic, and especially arginine-rich CPPs are further associated with high toxicity as frequently observed in kidney samples from in vivo studies.[70] Therefore, an important consideration in the discovery of novel cell-penetrating peptides for therapeutic cargo delivery is ensuring low toxicity.

Herein we also discuss the C9orfDPRs, a series of dipeptide repeat proteins (DPRs) with the ability to enter neighboring cells when secreted.[72] These proteins, however, are not in the category of peptides and proteins being studied for cargo delivery. Instead, they are being studied for their relevance to neurodegenerative disorders.[73] In the case of these arginine-rich DPRs, their cell entry is paired with high toxicity.[74] These peptides cause toxicity both after their endocytic cell entry and through direct disruption of the plasma membrane, which we demonstrate in Chapter 5. While the synthesis and study of these DPRs is useful for better understanding the mechanisms behind the two disorders ALS and FTD, they also represent a type of peptide the CPP field is eager to avoid: peptides with high toxicity and membrane disruption during cell penetration.

1.3 Cell penetrating peptides for crossing the blood-brain barrier

The limited bioavailability of peptide drugs within the target tissue is particularly crucial within the brain due to the added factor of the blood-brain barrier (BBB). Unlike other tissues, endothelial cells in the brain vasculature form tight junctions and have an extensive ensemble of transporters and efflux pumps to prevent molecules from crossing from the bloodstream into the brain.[75,76] In order to reach tissues within the brain, such as neurons and glial cells, molecules must cross not only the endothelial cells at the capillaries, but also pericytes and astrocytic endfeet.[75,77] Taken together, these factors prevent most macromolecular therapeutics and more than 98% of existing small molecule drugs from reaching the brain.[78]

Due to the importance of macromolecular therapeutics such as antibody-drug conjugates and antisense oligonucleotides for treating disorders in the central nervous system, there has been much research into crossing the blood-brain barrier.[79,80] Invasive technologies include direct brain injection and intrathecal injections, as well as deep brain stimulation and focused ultrasound.[76,77] Less invasive methods are under investigation, which include carrier-mediated processes that can transport a compound across the BBB after intravenous or oral dosing.[81] These methods primarily rely on transcytosis across endothelial cells.

Receptor-mediated transcytosis relies on attaching a carrier molecule to the therapeutic cargo that can be recognized by endothelial cells and transported across the BBB. This can involve covalent conjugation between carrier and cargo or encapsulation of the cargo within a liposome or nanoparticle. [76,77,82] Peptides often play a role as the receptor-recognition module, especially virus-derived peptides or peptides that recognize the transferrin receptor.[83,84] Transcytosis can also occur independent of a receptor through the process of adsorptive transcytosis. This method was proposed based on the increased ability of cationized albumin to cross the BBB over the native protein.[85] The mechanisms of cationic adsorptive transcytosis are similar to those discussed above for CPPs,[86] except that BBB-penetrating peptides must both enter and exit the endothelial cells, as well as the neighboring pericytes and astrocytes. This has opened a role for CPPs as BBB shuttles.[87] Although not all CPPs can shuttle cargo across the BBB through transcytosis, many known CPPs have demonstrated CNS penetration.[87,88]

In an effort to increase the number of therapeutics able to cross the BBB, a number of in vitro assays for BBB penetration have been developed. The transwell model is one of the simplest and most common models for BBB penetration. This assay assesses transport across a flat layer of endothelial cells or a co-culture of endothelial and other BBB cells. [89] Microfluidic models like the BBB-on-a-chip model incorporate the fluid flow that more closely resembles the capillary tissues. Cells are grown on a fabricated chip, and permeability across the BBB is measured by built-in sensors.[90] Finally, co-culture models and organoids can more accurately model the BBB. [91,92] The BBB spheroid model used to measure BBB penetration relies on the co-culture of endothelial cells, pericytes, and astrocytes, and penetration is assessed by the quantity of compound reaching the center of the spheroid.[88] These in-vitro assays, however, continuously fail to correlate with results from in vivo studies.[88,91,92] Therefore, our approach to determine BBB penetration begins with the BBB-spheroid as an in vitro method, but is grounded mainly in a series of in vivo studies using a mouse model. The results from in vivo studies presented in Chapter 4 demonstrate more concretely that our glioma-targeting peptide crosses the BBB to enter the target tissue.

1.4 Cell-penetrating peptides for oligonucleotide delivery

Cell-penetrating peptides offer a promising solution to the challenging delivery of antisense oligonucleotides (ASOs). Antisense oligonucleotides first demonstrated therapeutic efficacy as potential antiviral compounds, inhibiting viral replication in vitro.[93] Moreover, these synthetic nucleotide strands exhibit base-pairing with RNA that can be leveraged for gene-specific knock-down, alternate splicing, and even gene activation.[94] Subsequent chemical modifications of the ASO backbone have conferred additional features including stability, increased RNA hybridization, and reduced immunogenicity, resulting in a range of molecular structures within the ASO category. [94]

Among their many uses, ASOs have shown great therapeutic benefit over the past decade in modifying the splicing of pre-mRNA.[95] Nusinersen, a modified oligonucleotide featuring 2'-O-Methyl bases and phosphorothioate linkages (2'-MOE PS), oligonucleotide, is a clinically approved treatment for spinal muscular atrophy (SMA). This drug hybridizes to the pre-mRNA to block a splicing silencer element, leading to the increased inclusion of exon 7 in the mRNA transcript.[79] Nusinersen requires intrathecal injection to cross the BBB and exhibits low accumulation in the nucleus due to poor cell permeability.

Similarly, there are approved ASOs for splice correction in Duchenne Muscular Dystrophy (DMD). Eteplirsen, Golodirsen, Viltolarsen, and Casimersen are all FDA-approved therapeutics with a phosphorodiamidate morpholino oligomer (PMO) backbone, which is shown in **Fig. 1.1a**.[96,97] DMD stems from frameshift or nonsense mutations in the dystrophin gene that generate a non-functional copy of the dystrophin protein.[98] Dystrophin is an important protein in muscle damage recovery, and patients with DMD face muscle weakening and early death.[98] The PMO drugs approved for DMD treatments engage in exon skipping to restore the reading frame resulting in a truncated but functional copy of the dystrophin protein. Currently there are approved drugs for skipping exons 51, 53, and 45, with more ASO sequences currently in development.[99] Like the drugs for SMA, however, these PMO drugs for DMD are not cell-permeable and require IV injections at high doses.[100]

There is an unmet and pressing need for more effective delivery of PMO drugs into the nucleus of muscle cells. Cell-penetrating peptides have demonstrated the ability to enhance PMO delivery both in murine models and in human clinical trials.[101,102] Sarepta Therapeutics, which produced three previous PMO drugs, is moving forward with clinical trials of SRP-5051, a CPP-PMO conjugate.[102] While these peptide-PMO conjugates can increase the efficacy of PMO drugs, they can also generate potentially harmful or toxic effects. [101] There are no approved peptide-ASO conjugates for splice correction to date. There has, however, been extensive research into discovering novel CPPs for PMO conjugation. This led to the development of in vitro assays for PMO delivery, paired with assays to evaluate toxicity.[103] While many assays for CPP discovery rely on fluorophore delivery or otherwise look for evidence of peptides crossing the plasma membrane, the discovery of effective CPP-PMO conjugates for gene therapy requires a more functional readout. Instead of looking for CPP penetration, these functional assays assess splice correction of a target gene product. [103] One commonly used in vitro assay relies on the HeLa 654 cell line, which has been stably transfected with a coding sequence for enhanced green fluorescent protein (eGFP) that is interrupted by the IVS2-654 β-globin intron containing an aberrant splice site (**Fig. 1.1b**). Delivery of the 18-mer IV654 PMO antisense oligonucleotide to the nucleus restores eGFP fluorescence by covering the aberrant splice site, allowing correct splicing of the β-globin intron out of the eGFP transcript. Delivery of PMO to the nucleus can be assessed by measuring eGFP fluorescence.[51] This functional in vitro assay is used to evaluate the delivery efficiency of the Peptide-PMO conjugates discovered in Chapters 2 and 3.

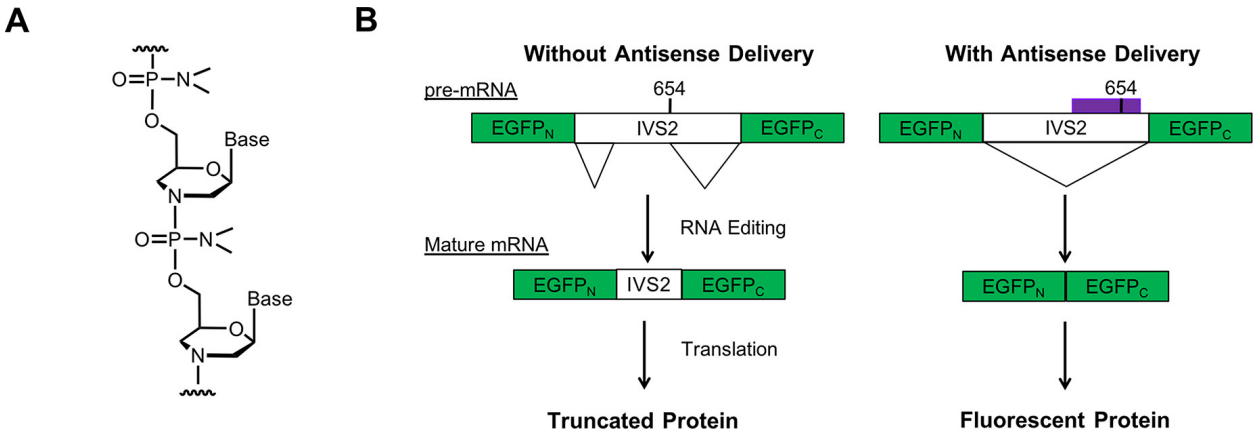


Figure 1.1 A model in-vitro assay for nuclear PMO delivery.
(a) The backbone structure of phosphorodiamidate morpholino oligonucleotides (PMOs), the delivery of which is the focus of Chapters 1 and 2. (b) An in vitro assay for PMO-mediated exon skipping. HeLa-654 cells are stably transfected with an eGFP construct that contains a mutant intron with an aberrant splice site. In the absence of PMO, a nonfluorescent protein is expressed. If PMO IVS2-654 is present, it hybridizes to the mutant intron, blocks the aberrant splice site, alters pre-mRNA splicing, and produces functional mRNA that is translated into fluorescent eGFP. Fluorescence read by flow cytometry can then quantify PMO delivery to the nucleus. Figure 1.1 adapted from Wolfe and Fadzen et al.[51]

The potential toxicity of CPP-cargo constructs must also be assessed in vitro, since both cargo and peptide can affect conjugate toxicity.[55,56] Overall cell viability can be measured by assessing cellular metabolic activity. The MTT assay, cell-titer Glo assay, and others measure the number of active cells remaining in culture after CPP-PMO treatment (**Fig. 1.2a**).[104] CPPs may also induce more acute membrane disruption, which can be measured through assessing the integrity of the plasma membrane. The lactate dehydrogenase leakage assay (**Fig 1.2b**) detects the presence of cytosolic proteins in the external cell media, indicating plasma membrane disruption.[105] These cell assays are used throughout the following chapters to ensure that PMO-CPP conjugates have low toxicity and membrane disruption. These in-vitro assays are also used in Chapter 5 to measure toxicity of arginine-rich DPR proteins, which demonstrate both acute toxicity at the cell membrane and long-term lowered cell viability.

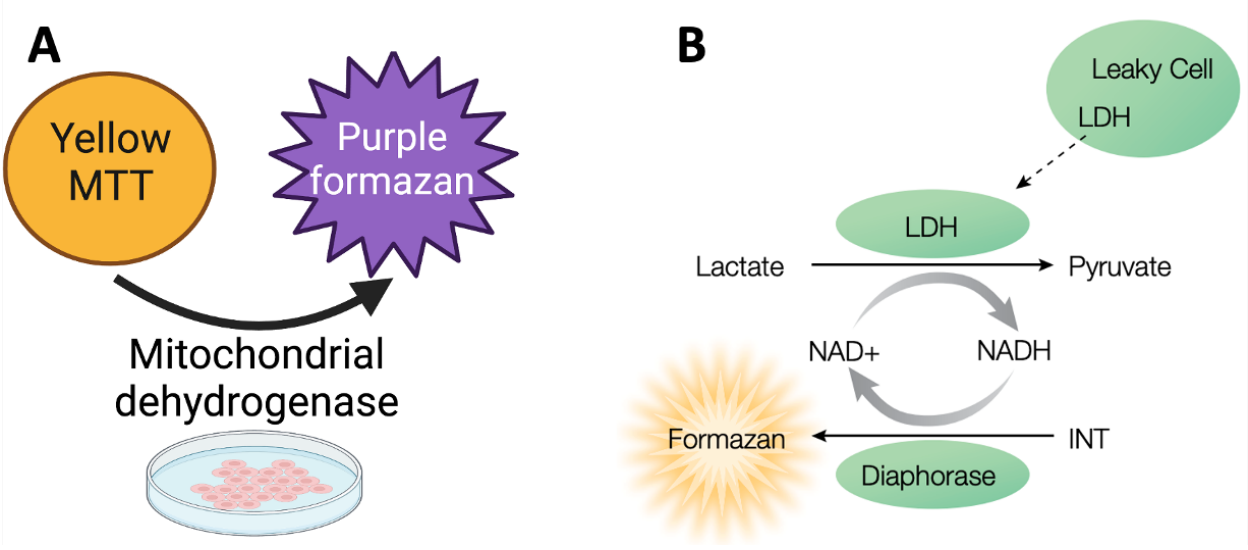


Figure 1.2 In-vitro assays for cell viability and membrane permeability.
(a) The Promega LDH assay assesses membrane permeability by measuring the quantity of the cytosolic protein lactate dehydrogenase (LDH) in the external cell media. Through an enzyme-coupled assay, the

presence of LDH results in the production of formazan, which can be measured spectroscopically. (b) The MTT assay measures cell viability by quantifying metabolically active cells. Live cells have mitochondrial dehydrogenase, which can convert MTT to purple formazan crystals. Subsequent cell lysis and dissolution of the formazan crystals allows quantification of formazan spectroscopically. Figure 1.2B adapted from Promega CytoTox 96™ Nonradioactive Cytotoxicity Assay literature.

Throughout these studies, we have discovered and assessed numerous PMO-CPP compounds for both efficacy and toxicity in vitro. Our work was aimed at the discovery of peptides with high delivery efficacy and low toxicity, and we have used sequence-activity relationships to increase the gap between efficacy and toxicity (also referred to as the therapeutic window) where necessary. We hope to see the results from in vitro assays translate to in vivo activity to improve the delivery of PMO drugs in patients affected by DMD.

1.5 Summary and outlook

The contents of this thesis are aimed at the discovery of novel cell-penetrating peptides, and the novel synthesis of existing compounds that can cross the cell membrane. **Chapter 1** introduces therapeutic peptides with a special focus on peptides able to cross cell membranes and other biological barriers and the assays often used to measure peptide penetration and toxicity. **Chapter 2** focuses on the in-cell penetration-selection mass-spectrometry platform, which enables the discovery of novel CPPs from within cells after treatment with a peptide library. **Chapter 3** centers on the cation-exchange pooling platform, which uses chromatographic separation of a peptide library to discover new CPPs. **Chapter 4** converges on using the BTP-7 peptide, which penetrates the BBB and targets glioma cells, to deliver camptothecin into cancerous glioma tissues within the brain. **Chapter 5** discusses the first chemical synthesis of 200-residue dipeptide repeat proteins (DPRs) to study their biophysical properties and toxicity in vitro. Finally, the **appendix** describes a mass-spectrometry-based platform to quantify endosomal entrapment and cytosolic entry of cell-penetrating peptides conjugated to an oligonucleotide cargo.

In addition to these works within this thesis, I have also been involved in a number of additional projects as a collaborator. The work presented in Chapter 2 relies on protocols developed prior, especially the isoseramox cleavable linker and the gentle cell lysis to differentiate endosomally trapped peptides from those present in the cytosol (as discussed in the appendix).[64, 106,107] While Chapters 2 and 3 focus on delivery of PMO, CPPs are used to deliver a much wider variety of therapeutic cargoes. The delivery of peptide nucleic acids (PNAs) can be performed in a single-shot synthesis of a PNA-CPP construct, without a requirement for an additional linker between peptide and cargo.[108] CPPs can also deliver antigens to dendritic cells and carry platinum drugs across the BBB into tumor cells.[109,110] This ability of peptides to carry therapeutics across the BBB is discussed thoroughly in Chapter 4. The BTP-7 peptide is an effective Camptothecin delivery agent,[111] and we have we have also previously studied BTP-7 for use as a tumor imaging agent.[112] Beyond peptides, we have also studied proteins and larger compounds with targeting or cell-penetrating properties. Chapter 5 covers the synthesis of dipeptide repeat proteins,[113] and we have also used synthetic protein or antibody conjugates as targeting modalities for small-molecule drugs.[114] Taken together, the. chapters within this thesis represent those works to which I have made substantial contribution, but I have also been lucky to participate in a number of additional

projects relating to the discovery and chemical synthesis of peptides and proteins that cross biological barriers.

1.6 References

(1) Wang, L.; Wang, N.; Zhang, W.; Cheng, X.; Yan, Z.; Shao, G.; Wang, X.; Wang, R.; Fu, C. Therapeutic Peptides: Current Applications and Future Directions. *Signal Transduction and Targeted Therapy 2022 7:1* **2022**, *7* (1), 1–27. https://doi.org/10.1038/s41392-022-00904-4.

(2) Du Vigneaud, V.; Ressler, C.; Swan, J. M.; Roberts, C. W.; Katsoyannis, P. G.; Gordon, S. The Synthesis of an Octapeptide Amide With The Hormonal Activity of Oxytocin. *J Am Chem Soc* **1953**, *75* (19), 4879–4880.

(3) Kung, Y. T.; Du, Y. C.; Huang, W. T.; Chen, C. C.; Ke, L. T. Total Synthesis of Crystalline Bovine Insulin. *Scientia Sinica (English Edition)* **1965**, *14*, 1710–1716.

(4) Behrendt, R.; White, P.; Offer, J. Advances in Fmoc Solid-Phase Peptide Synthesis. *Journal of Peptide Science* **2016**, *22* (1), 4–27. https://doi.org/10.1002/PSC.2836.

(5) M Jaradat, san M. Thirteen Decades of Peptide Synthesis: Key Developments in Solid Phase Peptide Synthesis and Amide Bond Formation Utilized in Peptide Ligation. *Amino Acids* **2018**, *50*, 39–68. https://doi.org/10.1007/s00726-017-2516-0.

(6) Simon, M. D.; Heider, P. L.; Adamo, A.; Vinogradov, A. A.; Mong, S. K.; Li, X.; Berger, T.; Policarpo, R. L.; Zhang, C.; Zou, Y.; Liao, X.; Spokoyny, A. M.; Jensen, K. F.; Pentelute, B. L. Rapid Flow-Based Peptide Synthesis. *ChemBioChem* **2014**, *15* (5), 713–720. https://doi.org/10.1002/CBIC.201300796.

(7) Mijalis, A. J.; Thomas, D. A.; Simon, M. D.; Adamo, A.; Beaumont, R.; Jensen, K. F.; Pentelute, B. L. A Fully Automated Flow-Based Approach for Accelerated Peptide Synthesis. *Nat Chem Biol* **2017**, *13* (5), 464–466. https://doi.org/10.1038/nchembio.2318.

(8) Hartrampf, N.; Saebi, A.; Poskus, M.; Gates, Z. P.; Callahan, A. J.; Cowfer, A. E.; Hanna, S.; Antilla, S.; Schissel, C. K.; Quartararo, A. J.; Ye, X.; Mijalis, A. J.; Simon, M. D.; Loas, A.; Liu, S.; Jessen, C.; Nielsen, T. E.; Pentelute, B. L. Synthesis of Proteins by Automated Flow Chemistry. *Science (1979)* **2020**, *368* (6494), 980–987.

(9) Henninot, A.; Collins, J. C.; Nuss, J. M. The Current State of Peptide Drug Discovery: Back to the Future? *J Med Chem* **2017**, *61*, 1382–1414. https://doi.org/10.1021/acs.jmedchem.7b00318.

(10) Knudsen, L. B.; Lau, J. The Discovery and Development of Liraglutide and Semaglutide. *Front Endocrinol (Lausanne)* **2019**, *10* (APR), 155. https://doi.org/10.3389/FENDO.2019.00155.

(11) Craik, D. J.; Fairlie, D. P.; Liras, S.; Price, D. The Future of Peptide-Based Drugs. *Chem Biol Drug Des* **2013**, *81* (1), 136–147. https://doi.org/10.1111/CBDD.12055.

(12) Todaro, B.; Ottalagana, E.; Luin, S.; Santi, M. Targeting Peptides: The New Generation of Targeted Drug Delivery Systems. *Pharmaceutics* **2023**, *15* (6), 1648. https://doi.org/10.3390/PHARMACEUTICS15061648.

(13) Samec, T.; Boulos, J.; Gilmore, S.; Hazelton, A.; Alexander-Bryant, A. Peptide-Based Delivery of Therapeutics in Cancer Treatment. *Mater Today Bio* **2022**, *14*, 100248. https://doi.org/10.1016/J.MTBIO.2022.100248.

(14) Xie, J.; Bi, Y.; Zhang, H.; Dong, S.; Teng, L.; Lee, R. J.; Yang, Z. Cell-Penetrating Peptides in Diagnosis and Treatment of Human Diseases: From Preclinical Research to Clinical Application. *Front Pharmacol* **2020**, *11*, 541971. https://doi.org/10.3389/FPHAR.2020.00697/BIBTEX.

(15) Di, L. Strategic Approaches to Optimizing Peptide ADME Properties. *AAPS Journal* **2015**, *17* (1), 134–143. https://doi.org/10.1208/s12248-014-9687-3.

(16) Soto, C. A. and C. Converting a Peptide into a Drug: Strategies to Improve Stability and Bioavailability. *Current Medicinal Chemistry*. **2002**, 963–978. https://doi.org/http://dx.doi.org/10.2174/0929867024606731.
(17) Lee, A. C. L.; Harris, J. L.; Khanna, K. K.; Hong, J. H. A Comprehensive Review on Current Advances in Peptide Drug Development and Design. *Int J Mol Sci* **2019**, *20* (10). https://doi.org/10.3390/ijms20102383.
(18) Pollaro, L.; Heinis, C. Strategies to Prolong the Plasma Residence Time of Peptide Drugs. *MedChemComm*. **2010**, 319–324. https://doi.org/10.1039/c0md00111b.
(19) Kelly, P. A.; Wang, H.; Napoli, K. L.; Kahan, B. D.; Strobel, H. W. Metabolism of Cyclosporine by Cytochromes P450 3A9 and 3A4. *Eur J Drug Metab Pharmacokinet* **1999**, *24* (4), 321–328. https://doi.org/10.1007/BF03190040.
(20) Weiss, H. M.; Wirz, B.; Schweitzer, A.; Amstutz, R.; Perez, M. I. R.; Andres, H.; Metz, Y.; Gardiner, J.; Seebach, D. ADME Investigations of Unnatural Peptides: Distribution of A14C-Labeledβ3-Octaarginine in Rats. *Chem Biodivers* **2007**, *4* (7), 1413–1437. https://doi.org/10.1002/cbdv.200790121.
(21) John, H.; Maronde, E.; Forssmann, W.-G.; Meyer, M.; Adermann, K. N-Terminal Acetylation Protects Glucagon-like Peptide GLP-1-(7-34)-Amide from DPP-IV-Mediated Degradation Retaining CAMP- and Insulin-Releasing Capacity. *Eur J Med Res* **2008**, *13* (2), 73–78.
(22) Kremsmayr, T.; Aljnabi, A.; Blanco-Canosa, J. B.; Tran, H. N. T.; Emidio, B.; Muttenthaler, M. On the Utility of Chemical Strategies to Improve Peptide Gut Stability. *J. Med. Chem* **2022**, 6191–6206. https://doi.org/10.1021/acs.jmedchem.2c00094.
(23) Tripathi, P. P.; Arami, H.; Banga, I.; Gupta, J.; Gandhi, S. Cell Penetrating Peptides in Preclinical and Clinical Cancer Diagnosis and Therapy. *Oncotarget* **2018**, *9* (98), 37252–37267. https://doi.org/10.18632/oncotarget.26442.
(24) Lian, W.; Jiang, B.; Qian, Z.; Pei, D. Cell-Permeable Bicyclic Peptide Inhibitors against Intracellular Proteins. *J Am Chem Soc* **2014**, *136* (28), 9830–9833. https://doi.org/10.1021/ja503710n.
(25) Ahlbach, C. L.; Lexa, K. W.; Bockus, A. T.; Chen, V.; Crews, P.; Jacobson, M. P.; Lokey, R. S. Beyond Cyclosporine A: Conformation-Dependent Passive Membrane Permeabilities of Cyclic Peptide Natural Products. *Future Med Chem* **2015**, *7* (16), 2121–2130. https://doi.org/10.4155/fmc.15.78.
(26) Alex, A.; Millan, D. S.; Perez, M.; Wakenhut, F.; Whitlock, G. A. Intramolecular Hydrogen Bonding to Improve Membrane Permeability and Absorption in beyond Rule of Five Chemical Space. *Medchemcomm* **2011**, *2* (7), 669–674. https://doi.org/10.1039/c1md00093d.
(27) LeCher, J. C.; Nowak, S. J.; McMurry, J. L. Breaking in and Busting out: Cell-Penetrating Peptides and the Endosomal Escape Problem. *Biomol Concepts* **2017**, *8* (3–4), 131. https://doi.org/10.1515/BMC-2017-0023.
(28) Bechara, C.; Sagan, S. Cell-Penetrating Peptides: 20 Years Later, Where Do We Stand? *FEBS Lett* **2013**, *587* (12), 1693–1702. https://doi.org/10.1016/J.FEBSLET.2013.04.031.
(29) Ruoslahti, E.; Bhatia, S. N.; Sailor, M. J. Targeting of Drugs and Nanoparticles to Tumors. *J Cell Biol* **2010**, *188* (6), 759–768. https://doi.org/10.1083/jcb.200910104.
(30) Fischer, P. M.; Zhelev, N. Z.; Wang, S.; Melville, J. E.; Fåhraeus, R.; Lane, D. P. Structure-Activity Relationship of Truncated and Substituted Analogues of the Intracellular Delivery Vector Penetratin. *J Pept Res* **2000**, *55* (2), 163–172.

(31) Fawell, S.; Seery, J.; Daikh, Y.; Moore, C.; Chen, L. L.; Pepinsky, B.; Barsoum, J. Tat-Mediated Delivery of Heterologous Proteins into Cells. *Proc Natl Acad Sci U S A* **1994**, *91* (2), 664. https://doi.org/10.1073/PNAS.91.2.664.

(32) Nagahara, H.; Vocero-Akbani, A. M.; Snyder, E. L.; Ho, A.; Latham, D. G.; Lissy, N. A.; Becker-Hapak, M.; Ezhevsky, S. A.; Dowdy, S. F. Transduction of Full-Length TAT Fusion Proteins into Mammalian Cells: TAT-P27Kip1 Induces Cell Migration. *Nat Med* **1998**, *4* (12), 1449–1452. https://doi.org/10.1038/4042.

(33) Agrawal, P.; Bhalla, S.; Usmani, S. S.; Singh, S.; Chaudhary, K.; Raghava, G. P. S.; Gautam, A. CPPsite 2.0: A Repository of Experimentally Validated Cell-Penetrating Peptides. *Nucleic Acids Res* **2016**, *44*, D1098–D1103. https://doi.org/10.1093/nar/gkv1266.

(34) De Coupade, C.; Fittipaldi, A.; Chagnas, V.; Michel, M.; Carlier, S.; Tasciotti, E.; Darmon, A.; Ravel, D.; Kearsey, J.; Giacca, M.; Cailler, F. Novel Human-Derived Cell-Penetrating Peptides for Specific Subcellular Delivery of Therapeutic Biomolecules. *Biochem J* **2005**, *390*, 407–418. https://doi.org/10.1042/BJ20050401.

(35) Reynolds, F.; Weissleder, R.; Josephson, L. Protamine as an Efficient Membrane-Translocating Peptide. *Bioconjug Chem* **2005**, *16* (5), 1240–1245.

(36) Ziegler, A.; Seelig, J. High Affinity of the Cell-Penetrating Peptide HIV-1 Tat-PTD for DNA †. *Biochemistry* **2007**, *46*, 8138–8145. https://doi.org/10.1021/bi700416h.

(37) Balayssac, S.; Burlina, F.; Convert, O.; Bolbach, G.; Chassaing, G.; Lequin, O. Comparison of Penetratin and Other Homeodomain-Derived Cell-Penetrating Peptides: Interaction in a Membrane-Mimicking Environment and Cellular Uptake Efficiency. *Biochemistry* **2006**, *45* (5), 1408–1420. https://doi.org/10.1021/bi0518390.

(38) Derossi, D.; Joliot, A. H.; Chassaing, G.; Prochiantz, A. The Third Helix of the Antennapedia Homeodomain Translocates through Biological Membranes. *J Biol Chem* **1994**, *269* (14), 10444–10450.

(39) Durzyńska, J.; Przysiecka, Ł.; Nawrot, R.; Barylski, J.; Nowicki, G.; Warowicka, A.; Musidlak, O.; Goździcka-Józefiak, A. Viral and Other Cell-Penetrating Peptides as Vectors of Therapeutic Agents in Medicine. *Journal of Pharmacology and Experimental Therapeutics* **2015**, *354* (1), 32–42. https://doi.org/10.1124/JPET.115.223305.

(40) Huey, R.; Hawthorne, S.; McCarron, P. The Potential Use of Rabies Virus Glycoprotein-Derived Peptides to Facilitate Drug Delivery into the Central Nervous System: A Mini Review. *J Drug Target* **2017**, *25* (5), 379–385. https://doi.org/10.1080/1061186X.2016.1223676.

(41) Splith, K.; Neundorf, I. Antimicrobial Peptides with Cell-Penetrating Peptide Properties and Vice Versa. *European Biophysics Journal* **2011**, *40* (4), 387–397. https://doi.org/10.1007/s00249-011-0682-7.

(42) Thaker, H. D.; Sgolastra, F.; Clements, D.; Scott, R. W.; Tew, G. N. Synthetic Mimics of Antimicrobial Peptides from Triaryl Scaffolds. *J Med Chem* **2011**, *54* (7), 2241–2254. https://doi.org/10.1021/jm101410t.

(43) Diedrichsen, R. G.; Harloff-Helleberg, S.; Werner, U.; Besenius, M.; Leberer, E.; Kristensen, M.; Nielsen, H. M. Revealing the Importance of Carrier-Cargo Association in Delivery of Insulin and Lipidated Insulin. *Journal of Controlled Release* **2021**, *338*, 8–21. https://doi.org/10.1016/J.JCONREL.2021.07.030.

(44) Fadzen, C. M.; Holden, R. L.; Wolfe, J. M.; Choo, Z. N.; Schissel, C. K.; Yao, M.; Hanson, G. J.; Pentelute, B. L. Chimeras of Cell-Penetrating Peptides Demonstrate Synergistic

Improvement in Antisense Efficacy. *Biochemistry* **2019**, *58* (38), 3980–3989. https://doi.org/10.1021/acs.biochem.9b00413.
(45) Majidi, A.; Nikkhah, M.; Sadeghian, F.; Hosseinkhani, S. Development of Novel Recombinant Biomimetic Chimeric MPG-Based Peptide as Nanocarriers for Gene Delivery: Imitation of a Real Cargo. *European Journal of Pharmaceutics and Biopharmaceutics* **2016**, *107*, 191–204. https://doi.org/10.1016/J.EJPB.2016.06.017.
(46) Kauffman, W. B.; Guha, S.; Wimley, W. C. Synthetic Molecular Evolution of Hybrid Cell Penetrating Peptides. *Nat Commun* **2018**, *9* (1). https://doi.org/10.1038/s41467-018-04874-6.
(47) Wimley, W. C. Application of Synthetic Molecular Evolution to the Discovery of Antimicrobial Peptides. *Adv Exp Med Biol*, **2019**, *1117,* 241–255. https://doi.org/10.1007/978-981-13-3588-4_13.
(48) Hällbrink, M.; Kilk, K.; Elmquist, A.; Lundberg, P.; Lindgren, M.; Jiang, Y.; Pooga, M.; Soomets, U.; Langel, Ü. Prediction of Cell-Penetrating Peptides. *Int J Pept Res Ther* **2005**, *11* (4), 249–259. https://doi.org/10.1007/s10989-005-9393-1.
(49) Sanders, W. S.; Johnston, C. I.; Bridges, S. M.; Burgess, S. C.; Willeford, K. O. Prediction of Cell Penetrating Peptides by Support Vector Machines. *PLoS Comput Biol* **2011**, *7* (7), 1002101. https://doi.org/10.1371/journal.pcbi.1002101.
(50) Mooney, C.; Haslam, N. J.; Pollastri, G.; Shields, D. C. Towards the Improved Discovery and Design of Functional Peptides: Common Features of Diverse Classes Permit Generalized Prediction of Bioactivity. *PLoS One* **2012**, *7* (10). https://doi.org/10.1371/journal.pone.0045012.
(51) Wolfe, J. M.; Fadzen, C. M.; Choo, Z. N.; Holden, R. L.; Yao, M.; Hanson, G. J.; Pentelute, B. L. Machine Learning to Predict Cell-Penetrating Peptides for Antisense Delivery. *ACS Cent Sci* **2018**, *4* (4), 512–520. https://doi.org/10.1021/acscentsci.8b00098.
(52) Schissel, C. K.; Mohapatra, S.; Wolfe, J. M.; Fadzen, C. M.; Bellovoda, K.; Wu, C. L.; Wood, J. A.; Malmberg, A. B.; Loas, A.; Gómez-Bombarelli, R.; Pentelute, B. L. Deep Learning to Design Nuclear-Targeting Abiotic Miniproteins. *Nat Chem* **2021**, *13* (10), 992–1000. https://doi.org/10.1038/S41557-021-00766-3.
(53) Wu, C. H.; Liu, I. J.; Lu, R. M.; Wu, H. C. Advancement and Applications of Peptide Phage Display Technology in Biomedical Science. *Journal of Biomedical Science* **2016**, *23* (1), 1–14. https://doi.org/10.1186/S12929-016-0223-X.
(54) Gao, S.; Simon, M. J.; Hue, C. D.; Morrison, B.; Banta, S. An Unusual Cell Penetrating Peptide Identified Using a Plasmid Display-Based Functional Selection Platform. *ACS Chem. Biol* **2011**, *6*, 484–491. https://doi.org/10.1021/cb100423u.
(55) El-Andaloussi, S.; Järver, P.; Johansson, H. J.; Langel, Ü. Cargo-Dependent Cytotoxicity and Delivery Efficacy of Cell-Penetrating Peptides: A Comparative Study. *Biochemical Journal* **2007**, *407* (2), 285–292. https://doi.org/10.1042/BJ20070507.
(56) Hymel, H. C.; Rahnama, A.; Sanchez, O. M.; Liu, D.; Gauthier, T. J.; Melvin, A. T. How Cargo Identity Alters the Uptake of Cell-Penetrating Peptide (CPP)/Cargo Complexes: A Study on the Effect of Net Cargo Charge and Length. *Cells* **2022**, *11* (7), 1195. https://doi.org/10.3390/CELLS11071195.
(57) Lundin, P.; Johansson, H.; Guterstam, P.; Holm, T.; Hansen, M.; Langel, Ü.; EL Andaloussi, S. Distinct Uptake Routes of Cell-Penetrating Peptide Conjugates. *Bioconjug Chem* **2008**, *19* (12), 2535–2542. https://doi.org/10.1021/bc800212j.

(58) Tünnemann, G.; Martin, R. M.; Haupt, S.; Patsch, C.; Edenhofer, F.; Cardoso, M. C. Cargo-Dependent Mode of Uptake and Bioavailability of TAT-Containing Proteins and Peptides in Living Cells. *FASEB journal* **2006**, *20* (11), 1775–1784. https://doi.org/10.1096/FJ.05-5523COM.

(59) Guterstam, P.; Madani, F.; Hirose, H.; Takeuchi, T.; Futaki, S.; EL Andaloussi, S.; Gräslund, A.; Langel, Ü. Elucidating Cell-Penetrating Peptide Mechanisms of Action for Membrane Interaction, Cellular Uptake, and Translocation Utilizing the Hydrophobic Counter-Anion Pyrenebutyrate. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **2009**, *1788* (12), 2509–2517. https://doi.org/10.1016/J.BBAMEM.2009.09.014.

(60) Su, Y.; Waring, A. J.; Ruchala, P.; Hong, M. Membrane-Bound Dynamic Structure of an Arginine-Rich Cell-Penetrating Peptide, the Protein Transduction Domain of HIV TAT, from Solid-State NMR. *Biochemistry* **2010**, *49* (29), 6009. https://doi.org/10.1021/BI100642N.

(61) Herce, H. D.; Garcia, A. E.; Cardoso, M. C. Fundamental Molecular Mechanism for the Cellular Uptake of Guanidinium-Rich Molecules. *J Am Chem Soc* **2014**, *136* (50), 17459–17467. https://doi.org/10.1021/JA507790Z.

(62) Hirose, H.; Takeuchi, T.; Osakada, H.; Pujals, S.; Katayama, S.; Nakase, I.; Kobayashi, S.; Haraguchi, T.; Futaki, S. Transient Focal Membrane Deformation Induced by Arginine-Rich Peptides Leads to Their Direct Penetration into Cells. *Molecular Therapy* **2012**, *20* (5), 984. https://doi.org/10.1038/MT.2011.313.

(63) Herce, H. D.; Garcia, A. E.; Litt, J.; Kane, R. S.; Martin, P.; Enrique, N.; Rebolledo, A.; Milesi, V. Arginine-Rich Peptides Destabilize the Plasma Membrane, Consistent with a Pore Formation Translocation Mechanism of Cell-Penetrating Peptides. *Biophys J* **2009**, *97* (7), 1917. https://doi.org/10.1016/J.BPJ.2009.05.066.

(64) Schissel, C. K.; Farquhar, C. E.; Loas, A.; Malmberg, A. B.; Pentelute, B. L. In-Cell Penetration Selection-Mass Spectrometry Produces Noncanonical Peptides for Antisense Delivery. *ACS Chem Biol.* **2023**, *18,* 615-628. https://doi.org/10.1101/2022.04.13.488231.

(65) McClorey, G.; Banerjee, S. Cell-Penetrating Peptides to Enhance Delivery of Oligonucleotide-Based Therapeutics. *Biomedicines* **2018**, *6* (2). https://doi.org/10.3390/BIOMEDICINES6020051.

(66) Fuchs, S. M.; Raines, R. T. Pathway for Polyarginine Entry into Mammalian Cells. *Biochemistry* **2004**, *43* (9), 2438. https://doi.org/10.1021/BI035933X.

(67) Authier, F.; Posner, B. I.; Bergeron, J. J. M. Minireview Endosomal Proteolysis of Internalized Proteins. *FEBS 16979 FEBS Letters* **1996**, *389*, 55–60. https://doi.org/10.1016/0014-5793(96)00368-7.

(68) Yang, S. T.; Zaitseva, E.; Chernomordik, L. V.; Melikov, K. Cell-Penetrating Peptide Induces Leaky Fusion of Liposomes Containing Late Endosome-Specific Anionic Lipid. *Biophys J* **2010**, *99* (8), 2525–2533. https://doi.org/10.1016/J.BPJ.2010.08.029.

(69) Erazo-Oliveras, A.; Muthukrishnan, N.; Baker, R.; Wang, T. Y.; Pellois, J. P. Improving the Endosomal Escape of Cell-Penetrating Peptides and Their Cargos: Strategies and Challenges. *Pharmaceuticals*. **2012**, *5*, 1177–1209. https://doi.org/10.3390/ph5111177.

(70) Li, Q.; Xu, M.; Cui, Y.; Huang, C.; Sun, M. Arginine-Rich Membrane-Permeable Peptides Are Seriously Toxic. *Pharmacol Res Perspect* **2017**, *5*, e00334. https://doi.org/10.1002/prp2.334.

(71) Birch, D.; Christensen, M. V.; Staerk, D.; Franzyk, H.; Nielsen, H. M. Stereochemistry as a Determining Factor for the Effect of a Cell-Penetrating Peptide on Cellular Viability and

Epithelial Integrity. *Biochemical Journal* **2018**, *475* (10), 1773–1788. https://doi.org/10.1042/BCJ20180155.

(72) Flores, B. N.; Dulchavsky, M. E.; Krans, A.; Sawaya, M. R.; Paulson, H. L.; Todd, P. K.; Barmada, S. J.; Ivanova, M. I. Distinct C9orf72-Associated Dipeptide Repeat Structures Correlate with Neuronal Toxicity. *PLoS One* **2016**. https://doi.org/10.1371/journal.pone.0165084.

(73) Schmitz, A.; Pinheiro Marques, J.; Oertig, I.; Maharjan, N.; Saxena, S. Emerging Perspectives on Dipeptide Repeat Proteins in C9ORF72 ALS/FTD. *Front Cell Neurosci* **2021**, *15*. https://doi.org/10.3389/FNCEL.2021.637548.

(74) Kanekura, K.; Harada, Y.; Fujimoto, M.; Yagi, T.; Hayamizu, Y.; Nagaoka, K.; Kuroda, M. Characterization of Membrane Penetration and Cytotoxicity of C9orf72-Encoding Arginine-Rich Dipeptides. *Scientific Reports 2018 8:1* **2018**, *8* (1), 1–11. https://doi.org/10.1038/s41598-018-31096-z.

(75) Serlin, Y.; Shelef, I.; Knyazer, B.; Friedman, A. Anatomy and Physiology of the Blood-Brain Barrier. *Semin Cell Dev Biol* **2015**, *38*, 2–6. https://doi.org/10.1016/j.semcdb.2015.01.002.

(76) Tajes, M.; Ramos-Fernández, E.; Weng-Jiang, X.; Bosch-Morató, M.; Guivernau, B.; Eraso-Pichot, A.; Salvador, B.; Fernàndez-Busquets, X.; Roquer, J.; Muñoz, F. J. The Blood-Brain Barrier: Structure, Function and Therapeutic Approaches to Cross It. *Mol Membr Biol* **2014**, *31* (5), 152–167. https://doi.org/10.3109/09687688.2014.937468.

(77) Pardridge, W. M. Drug Transport across the Blood-Brain Barrier. *J Cereb Blood Flow Metab* **2012**, *32* (11), 1959–1972. https://doi.org/10.1038/jcbfm.2012.126.

(78) Wu, D.; Chen, Q.; Chen, X.; Han, F.; Chen, Z.; Wang, Y. The Blood–Brain Barrier: Structure, Regulation, and Drug Delivery. *Signal Transduction and Targeted Therapy 2023 8:1* **2023**, *8* (1), 1–27. https://doi.org/10.1038/s41392-023-01481-w.

(79) Li, Q. Nusinersen as a Therapeutic Agent for Spinal Muscular Atrophy. *Yonsei Med J* **2020**, *61* (4), 273. https://doi.org/10.3349/YMJ.2020.61.4.273.

(80) Abounader, R.; Schiff, D. The Blood-Brain Barrier Limits the Therapeutic Efficacy of Antibody-Drug Conjugates in Glioblastoma. *Neuro Oncol* **2021**, *23* (12), 1993–1994. https://doi.org/10.1093/NEUONC/NOAB223.

(81) Pardridge, W. M. Blood-Brain Barrier Drug Targeting: The Future of Brain Drug Development. *Mol Interv* **2003**, *3* (2), 90–105. https://doi.org/10.1124/mi.3.2.90.

(82) Anselmo, A. C.; Mitragotri, S. Nanoparticles in the Clinic: An Update. *Bioeng Transl Med* **2019**, *4* (3). https://doi.org/10.1002/btm2.10143.

(83) Sharma, G.; Lakkadwala, S.; Modgil, A.; Singh, J.; Sharma, G.; Lakkadwala, S.; Modgil, A.; Singh, J. The Role of Cell-Penetrating Peptide and Transferrin on Enhanced Delivery of Drug to Brain. *Int J Mol Sci* **2016**, *17* (6), 806. https://doi.org/10.3390/ijms17060806.

(84) Oswald, M.; Geissler, S.; Goepferich, A. Targeting the Central Nervous System (CNS): A Review of Rabies Virus-Targeting Strategies. *Mol Pharm* **2017**, *14* (7), 2177–2196. https://doi.org/10.1021/acs.molpharmaceut.7b00158.

(85) Kumagai, A. K.; Eisenberg, J. B.; Pardridge, W. M. Absorptive-Mediated Endocytosis of Cationized Albumin and a Beta-Endorphin-Cationized Albumin Chimeric Peptide by Isolated Brain Capillaries. Model System of Blood-Brain Barrier Transport. *Journal of Biological Chemistry* **1987**, *262* (31), 15214–15219. https://doi.org/10.1016/S0021-9258(18)48160-4.

(86) Hervé, F.; Ghinea, N.; Scherrmann, J. M. CNS Delivery via Adsorptive Transcytosis. *AAPS Journal* **2008**, *10* (3), 455–472. https://doi.org/10.1208/s12248-008-9055-2.

(87) Oller-Salvia, B.; Sánchez-Navarro, M.; Giralt, E.; Teixidó, M. Blood–Brain Barrier Shuttle Peptides: An Emerging Paradigm for Brain Delivery. *Chem Soc Rev* **2016**, *45* (17), 4690–4707. https://doi.org/10.1039/C6CS00076B.
(88) Cho, C.-F.; Wolfe, J. M.; Fadzen, C. M.; Calligaris, D.; Hornburg, K.; Chiocca, E. A.; Agar, N. Y. R.; Pentelute, B. L.; Lawler, S. E. Blood-Brain-Barrier Spheroids as an in Vitro Screening Platform for Brain-Penetrating Agents. *Nat Commun* **2017**, *8*, 15623. https://doi.org/10.1038/ncomms15623.
(89) Stone, N. L.; England, T. J.; O'Sullivan, S. E. A Novel Transwell Blood Brain Barrier Model Using Primary Human Cells. *Front Cell Neurosci* **2019**, *13*, 455689. https://doi.org/10.3389/FNCEL.2019.00230/BIBTEX.
(90) van der Helm, M. W.; van der Meer, A. D.; Eijkel, J. C. T.; van den Berg, A.; Segerink, L. I. Microfluidic Organ-on-Chip Technology for Blood-Brain Barrier Research. *Tissue Barriers* **2016**, *4* (1). https://doi.org/10.1080/21688370.2016.1142493.
(91) Helms, H. C.; Abbott, N. J.; Burek, M.; Cecchelli, R.; Couraud, P.-O.; Deli, M. A.; Förster, C.; Galla, H. J.; Romero, I. A.; Shusta, E. V; Stebbins, M. J.; Vandenhaute, E.; Weksler, B.; Brodin, B. In Vitro Models of the Blood-Brain Barrier: An Overview of Commonly Used Brain Endothelial Cell Culture Models and Guidelines for Their Use. *J Cereb Blood Flow Metab* **2016**, *36* (5), 862–890. https://doi.org/10.1177/0271678X16630991.
(92) Kaisar, M. A.; Sajja, R. K.; Prasad, S.; Abhyankar, V. V.; Liles, T.; Cucullo, L. New Experimental Models of the Blood-Brain Barrier for CNS Drug Discovery. *Expert Opin Drug Discov* **2017**, *12* (1), 89–103. https://doi.org/10.1080/17460441.2017.1253676.
(93) Stephenson, M. L.; Zamecnik, P. C. Inhibition of Rous Sarcoma Viral RNA Translation by a Specific Oligodeoxyribonucleotide. *Proc Natl Acad Sci U S A* **1978**, *75* (1), 285–288. https://doi.org/10.1073/PNAS.75.1.285.
(94) Chan, J. H. P.; Lim, S.; Wong, W. S. F. Antisense Oligonucleotides: From Design to Therapeutic Application. *Clin Exp Pharmacol Physiol* **2006**, *33* (5–6), 533–540. https://doi.org/10.1111/J.1440-1681.2006.04403.X.
(95) Havens, M. A.; Hastings, M. L. Splice-Switching Antisense Oligonucleotides as Therapeutic Drugs. *Nucleic Acids Res* **2016**, *44* (14), 6549. https://doi.org/10.1093/NAR/GKW533.
(96) Roshmi, R. R.; Yokota, T. Pharmacological Profile of Viltolarsen for the Treatment of Duchenne Muscular Dystrophy: A Japanese Experience. *Clin Pharmacol* **2021**, *13*, 235. https://doi.org/10.2147/CPAA.S288842.
(97) Aartsma-Rus, A.; Krieg, A. M. FDA Approves Eteplirsen for Duchenne Muscular Dystrophy: The Next Chapter in the Eteplirsen Saga. *Nucleic Acid Ther* **2017**, *27* (1), 1–3. https://doi.org/10.1089/nat.2016.0657.
(98) Yiu, E. M.; Kornberg, A. J. Duchenne Muscular Dystrophy. *J Paediatr Child Health* **2015**, *51* (8), 759–764. https://doi.org/10.1111/jpc.12868.
(99) Aartsma-Rus, A. The Future of Exon Skipping for Duchenne Muscular Dystrophy. *Hum Gene Ther* **2023**, *34*, 372–378.
(100) Charleston, J. S.; Schnell, F. J.; Dworzak, J.; Donoghue, C.; Lewis, S.; Chen, L.; Young, G. D.; Milici, A. J.; Voss, J.; DeAlwis, U.; Wentworth, B.; Rodino-Klapac, L. R.; Sahenk, Z.; Frank, D.; Mendell, J. R. Eteplirsen Treatment for Duchenne Muscular Dystrophy: Exon Skipping and Dystrophin Production. *Neurology* **2018**, *90* (24), e2146–e2154. https://doi.org/10.1212/WNL.0000000000005680.

(101) Tsoumpra, M. K.; Fukumoto, S.; Matsumoto, T.; Takeda, S.; Wood, M. J. A.; Aoki, Y. Peptide-Conjugate Antisense Based Splice-Correction for Duchenne Muscular Dystrophy and Other Neuromuscular Diseases. *EBioMedicine* **2019**, *45*, 630. https://doi.org/10.1016/J.EBIOM.2019.06.036.
(102) Sheikh, O.; Yokota, · Toshifumi. Pharmacology and Toxicology of Eteplirsen and SRP-5051 for DMD Exon 51 Skipping: An Update. *Arch Toxicol* **2022**, *96*, 1–9. https://doi.org/10.1007/s00204-021-03184-z.
(103) Boisguerin, P.; O'Donovan, L.; Gait, M. J.; Lebleu, B. In Vitro Assays to Assess Exon Skipping in Duchenne Muscular Dystrophy. *Methods in Molecular Biology* **2015**, *1324*, 317–329. https://doi.org/10.1007/978-1-4939-2806-4_20.
(104) Ghasemi, M.; Turnbull, T.; Sebastian, S.; Kempson, I. The Mtt Assay: Utility, Limitations, Pitfalls, and Interpretation in Bulk and Single-Cell Analysis. *Int J Mol Sci* **2021**, *22* (23). https://doi.org/10.3390/IJMS222312827/S1.
(105) Saar, K.; Lindgren, M.; Hansen, M.; Eiríksdóttir, E.; Jiang, Y.; Rosenthal-Aizman, K.; Sassian, M.; Langel, Ü. Cell-Penetrating Peptides: A Comparative Membrane Toxicity Study. *Anal Biochem* **2005**, *345* (1), 55–65. https://doi.org/10.1016/J.AB.2005.07.033.
(106) Schissel, C. K.; Farquhar, C. E.; Malmberg, A. B.; Loas, A.; Pentelute, B. L. Cell-Penetrating D-Peptides Retain Antisense Morpholino Oligomer Delivery Activity. *ACS Bio Med Chem Au.* **2022**, *2* (2), 150–160.
(107) Pomplun, S.; Shugrue, R.; Schmitt, M.; Schissel, C. K.; Farquhar, C. E.; Pentelute, B. L. Secondary Amino Alcohols: Traceless Cleavable Linkers for Use in Affinity Capture and Release. *Angew. Chemie Int. Ed.* **2020**, *59* (28), 11566–11572.
(108) Li, C.; Callahan, A. J.; Phadke, K. S.; Bellaire, B.; Farquhar, C. E.; Zhang, G. et al. Automated Flow Synthesis of Peptide-PNA Conjugates. *ACS Cent. Sci.* **2022**, *8* (2), 205–213.
(109) Backlund, C. M*.; Holden, R. L.*; Moynihan, K. D.; Garafola, D.; Farquhar, C. E.; Mehta, N. K.; Maiorino, L.; Pham, S.; Iorgulescu, J. B.; Reardon, D. A.; et al. Cell-Penetrating Peptides Enhance Peptide Vaccine Accumulation and Persistence in Lymph Nodes to Drive Immunogenicity. *Proc. Natl. Acad. Sci. U. S. A.* 2022, *119* (32). **These authors contributed equally*
(110) Jimenez-Macias, J. L.; Lee, Y. C.; Miller, E.; Finkelberg, T.; Zdioruk, M.; Berger, G.; Farquhar, C. E.; Nowicki, M. O.; Cho, C. F.; Fedeles, B. I.; et al. A Pt(IV)-Conjugated Brain Penetrant Macrocyclic Peptide Shows Pre-Clinical Efficacy in Glioblastoma. *J. Control. Release.* **2022**, *352*, 623–636.
(111) Cho, C.-F.; Farquhar, C. E.; Fadzen, C. M.; Scott, B.; Zhuang, P.; von Spreckelsen, N.; Loas, A.; Hartrampf, N.; Pentelute, B. L.; Lawler, S. E. A Tumor-Homing Peptide Platform Enhances Drug Solubility, Improves Blood-Brain Barrier Permeability and Targets Glioblastoma. *Cancers (Basel).* **2022**, *14* (9), 2207.
(112) Spreckelsen, N.; Fadzen, C. M.; Hartrampf, N.; Ghotmi, Y.; Wolfe, J. M.; Dubey, S.; Yang, B. Y.; Kijewski, M. F.; Wang, S.; Farquhar, C.; et al. Targeting Glioblastoma Using a Novel Peptide Specific to a Deglycosylated Isoform of Brevican. *Adv. Ther.* **2021**, 2000244.
(113) Sato, K.; Farquhar, C.E.; Rodriguez, J.J.L.; Pentelute, B.L. Automated fast-flow synthesis of Chromosome 9 open reading frame 72 dipeptide repeat proteins. *J. Am. Chem. Soc.* **2023**, *145* (24), 12992–12997
(114) *Delivery* Gazvoda, M.; Dhanjee, H. H.; Rodriguez, J.; Brown, J. S.; Farquhar, C. E.; Truex, N. L.; Loas, A.; Buchwald, S. L.; Pentelute, B. L. Palladium-Mediated Incorporation of

Carboranes into Small Molecules, Peptides, and Proteins. *J. Am. Chem. Soc.* **2022,** *144(17)*, 7852-7860.

Chapter 2. In-Cell Penetration Selection-Mass Spectrometry Produces Noncanonical Peptides for Antisense Delivery

The work presented in this chapter has been reproduced from the following manuscript:

Schissel, C. K.*; Farquhar, C. E*; Malmberg, A. B.; Loas, A.; Pentelute, B. L. In-cell penetration selection—mass spectrometry produces noncanonical peptides for antisense delivery. *ACS Chem Bio.* **2023**, *18* (3), 615-628. **These authors contributed equally*

2.1 Introduction

After 30 years of investigation, therapies involving cell-penetrating peptides (CPPs) are beginning to advance to late-stage clinical trials.[1,2] These sequences, composed typically of fewer than 20 amino acids and endowed with diverse physicochemical properties, are able to penetrate the cellular membrane and at times deliver otherwise non-penetrant cargo.[3] Because of these properties, CPPs have potential applications for the treatment of disease, including cancer, genetic disorders, inflammation, and diabetes. An example is SRP-5051, in which an antisense phosphorodiamidate morpholino oligomer (PMO) is conjugated to a CPP, and is currently under investigation in a Phase 2 clinical trial for splice-correction efficacy in Duchenne muscular dystrophy patients.[4] In addition, there have been recent FDA approvals for polycationic peptides, including Daxxify which includes a 35-mer cationic peptide sequence to stabilize botulinum toxin while it interacts with extracellular receptors.[5] Despite these recent advances, to our knowledge no CPP-based therapy has reached the commercial market yet.[6]

While there are several limitations that have slowed the clinical advancement of CPPs, one pressing challenge is the empirical design of novel, more efficient sequences. Historically, CPPs, also known as protein transduction domains, were derived from transmembrane portions of viral and transcriptional proteins. For example, the polyarginine peptide TAT was derived from the HIV-trans-activator of transcription protein and was found to penetrate into the nucleus and target gene expression.[7,8] From this and similar sequences, synthetic peptides could be designed, including some tailored for delivery of PMO cargo such as Bpep, which relies on a polyarginine sequence to improve uptake and the unnatural residues β-alanine and 6-amino-hexanoic acid to reduce endosomal trapping.[9,10] Beyond empirical design using derivatives of polyarginine sequences, the rational design of new sequences remains challenging. Methods involving some rational design include synthetic molecular evolution[11,12] and in silico methods.[13–18] The latter include our own recent work that leverages machine learning to design new sequences using a model trained with a combinatorial library tested for the desired activity: nuclear localization.[19–21] Finally, another common strategy involves screening platforms employing libraries from phage or mRNA display.[22–25] For example, a screening platform identified several "phylomer" CPPs from bacterial and viral genomes that were then shown to deliver antisense cargo *in vivo*.[26] Still, a persistent

limitation with these approaches is the difficulty of incorporating D-chiral or unnatural amino acids, the result of which is that current well-studied CPPs typically rely on canonical L-amino acids. While amino acid stereochemistry may or may not affect membrane penetration,[27–30] incorporation of unnatural residues would provide access to an augmented chemical space and additional protease stability.[29,31] Unnatural amino acids are more easily incorporated into synthetic one-bead one-compound (OBOC) libraries, although discovery of CPPs by these methods often relies on lyposomes.[32] Therefore, improved screening platforms are needed to discover enhanced, unnatural peptides to access greater chemical diversity and proteolytic stability, and require biologically relevant screening conditions, such as in-cell selection and inclusion of the specific cargo to be delivered.

Classic affinity selection involves screening peptide ligands from synthetic libraries (OBOC), phage or mRNA display against immobilized protein targets, and decoding hits.[33,34] These methods advanced to biologically relevant conditions in on-cell selection platforms for the discovery of new ligands with affinity for the external surface of cells and tissues.[35–37] Again, biological display techniques are restricted to the use of mostly natural amino acids, limiting the resulting library diversity and proteolytic stability[38,39], and even those mirror image techniques that allow D-peptide discovery still have difficulty incorporating non-canonical residues.[40,41] Screening of a synthetic one-bead one-compound (OBOC) library eases the incorporation of non-canonical and D-residues. Our group has recently demonstrated that in vivo affinity selection-mass spectrometry (AS-MS) could identify an erythrocyte-targeting D-peptide.[42] Such label-free techniques applied to the cell surface can be used to discover novel, non-canonical, D-peptide binders without the addition of display scaffolds or encoding tags.

While most works have focused on affinity screening at the cell surface, there has been some success pushing these techniques to discover peptides that cross the cell membrane. As mentioned, phage display and encoded peptide libraries have been used to discover novel cell penetrating peptides, but these methods have limited advancements for discovery of peptides that deliver cargo to subcellular compartments.[23,26] Recently, the first example of a DNA-encoded small molecule library screen inside living cells resulted in several chemical motifs that bind the

over-expressed protein targets inside oocytes.[43] This strategy of in-cell selection (rather than on-cell) would be beneficial for discovery of peptides that can deliver macromolecules to the cytosol.

Here we have combined CPP library design and AS-MS selection approaches into a new method: in-cell penetration selection-mass spectrometry (in-cell PS-MS, **Fig. 2.1**). Bringing together our expertise in both MS-based selection methodologies and in-cell localization, this technique enables direct recovery of "hit" peptides that deliver a specific type of antisense cargo into the cytosol of cells. Our PS-MS methodology allows the detection of non-canonical peptide-cargo conjugates in-cell, with only the addition of a small biotin handle for extraction. Our library was composed of D amino acids as well as 10 noncanonical residues which enhance serum stability by avoiding endogenous proteases. In addition, the PS-MS platform allows spatial resolution, as we can extract peptides from the cytosolic fraction as well as from whole cell lysates, which includes endosomes. Resulting peptides are tested in a validation assay that quantifies delivery of the cargo to the nucleus, providing an additional layer of spatial resolution. While the PMO-CPP library used in the screening demonstrated antisense delivery activity as an average of all PMO constructs within it, our in cell PS-MS method selects for individual sequences which enter the cell cytosol at higher concentrations. This method led to the discovery of an antisense delivery peptide, Cyto1a, isolated from the cytosol of cells. This peptide was more active and more efficiently localized to the nucleus compared to peptides that were isolated from the whole cell extracts, which include endosomes.

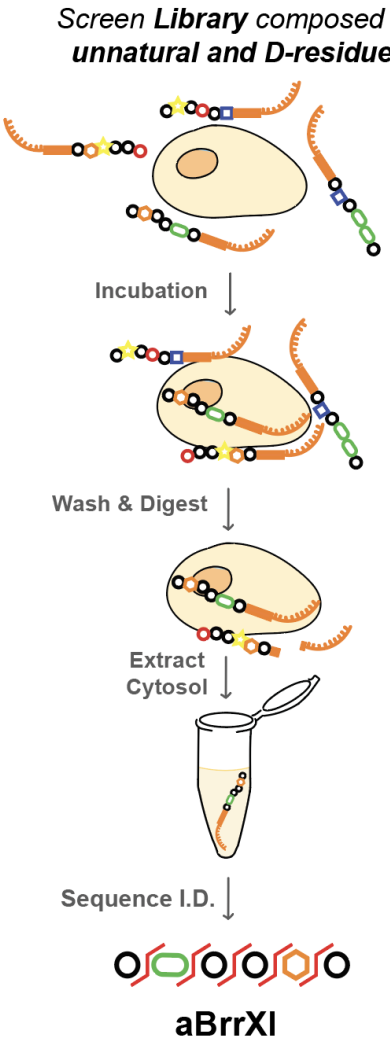


Figure 2.1 In-cell penetration selection—mass spectrometry identifies noncanonical peptides that access the cytosol.

In-cell PS-MS identifies noncanonical peptides that carry macromolecular cargo into the cytosol of cells. A library of unnatural and D-chirality peptide-antisense conjugates are screened in live cells. Extracellular constructs are removed, and the cytosol and whole cell lysates are extracted. Internalized peptides are affinity captured, released, and identified by sequencing with tandem mass spectrometry.

2.2 Results and Discussion

2.2.1 Library preparation

The library was prepared with a "CPP-like" C-terminal sequence and six variable positions containing D- and unnatural amino acids (**Fig. 2.2a**). Split-and-pool synthesis afforded 0.016 μg of peptide per bead for a low-redundancy, 95,000-member library with a theoretical diversity greater than 10^8. A D-kwkk motif, derived from the established cell-penetrating peptide penetratin,[44] was installed at the C-terminus to give the library a boost in activity. We have previously shown that these fixed constraints and C-terminal charge also increase peptide recovery in AS-MS.[33,45] Unnatural amino acids were chosen to expand the chemical diversity and potentially enhance cell penetration and proteolytic stability of the library peptides. The unnatural residues included those with non-α backbones to promote endosomal escape (γ-aminobutyric acid and β-alanine),[9] residues with hydrophobic and aromatic functionality to increase membrane penetration (homoleucine, norleucine, naphthylalanine, and diphenylalanine),[46,47] and additional charged residues and arginine analogues to enhance membrane penetration (diaminobutyric acid, aminopiperidine-carboxylic acid, aminomethylphenylalanine, and 2-amino-4-guanidinobutanoic acid)[48,49] (**Fig. 2.2b**). The oxidative cleavable linker isoseramox was installed by reductive amination immediately following the variable region as previously reported,[50] followed by a trypsin cleavage site to prevent the recovery of non-internalized peptides. Finally, azidolysine and biotin capped the N-terminus of the sequences to allow for PMO conjugation and affinity capture, respectively. Following cleavage from the resin, a portion of the library was conjugated by azide-alkyne cycloaddition to a model PMO derivatized with dibenzocyclooctyne (DBCO), monitored by LC-MS. Quality control analysis of the library by Orbitrap nano-liquid chromatography-tandem mass spectrometry (nLC-MS/MS) confirmed successful synthesis and exhibited a range of incorporated residues (**Fig. 2.2c**). The canonical and noncanonical residues are evenly incorporated and distributed throughout the variable region, with the exception of guanidyl residues, which showed reduced incorporation consistent with previous observations in solid-phase peptide synthesis.[51] This library design ensured the isolation of 10-mer peptides with a native N-terminus, suitable for sequencing via tandem mass spectrometry.[50]

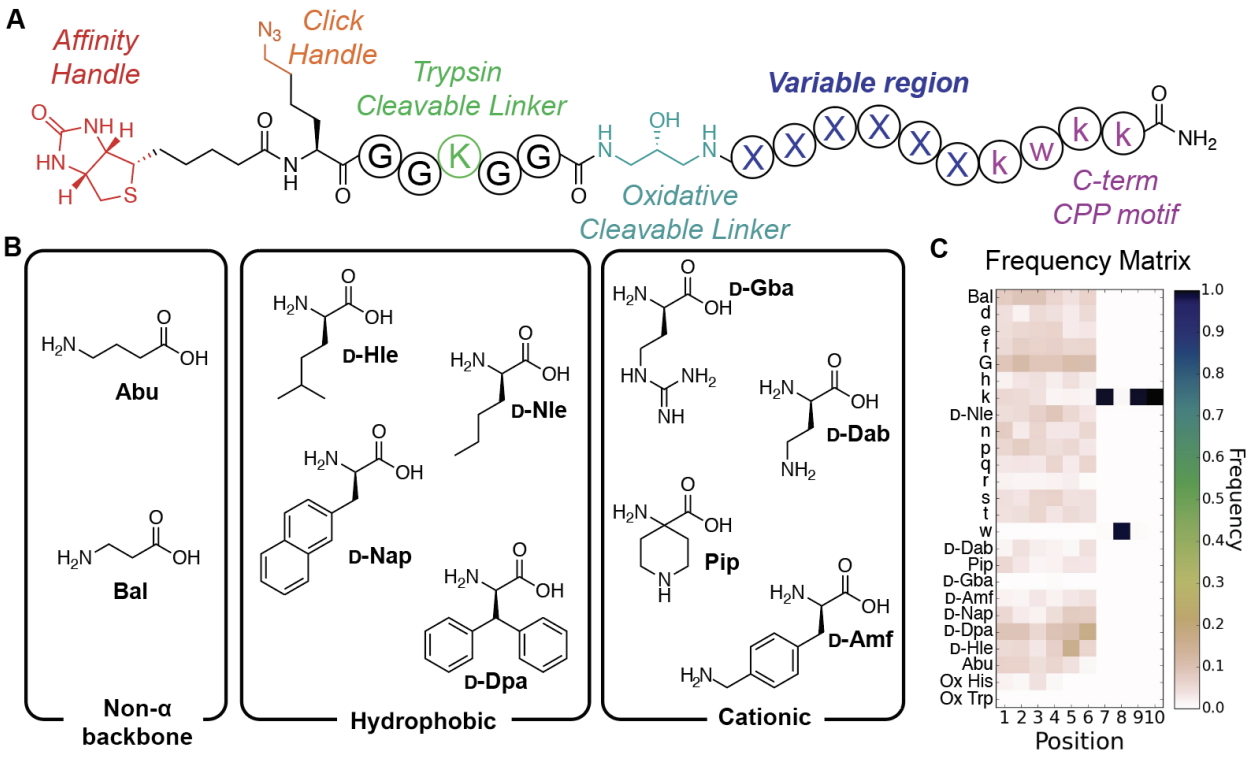


Figure 2.2 A combinatorial library was prepared with unnatural and d-amino acids.
(a) Design of the library. An N-terminal biotin and azidolysine provide an affinity handle and conjugation handle, respectively. A trypsin-cleavable linker prevents isolation of extracellular conjugates, and isoseramox cleavable linker permits oxidative cleavage of native peptides for MS/MS analysis. Finally, there are six variable positions within the library peptides, with a "CPP-like" motif capping the C-terminus. (b) Structures of the unnatural monomers used. All natural-backbone monomers were in D-form. (c) Heat map of the quality control showing relative abundance of the various amino acids of the sequence up to the isoseramox linker. Positions 7-10 show the D-kwkk motif, with positions 1-6 showing the varied composition of the variable region. Abu (γ-aminobutyric acid), Bal (β-alanine), D-Hle (homoleucine), D-Nle (norleucine), D-Nap (naphthylalanine), D-Dpa (diphenylalanine), D-Dab (diaminobutyric acid), Pip (aminopiperidine-carboxylic acid), D-Amf (aminomethylphenylalanine), and D-Gba (2-amino-4-guanidinobutanoic acid).

We then performed a series of *in vitro* experiments to confirm that the peptides within the library had nuclear-localizing activity. The phenotypic assay used detects the amount of active PMO delivered to the nucleus by resulting in corrective splicing to produce enhanced green fluorescent protein (EGFP), quantified by flow cytometry (**Fig 2.3a**). First, PMO-library aliquots demonstrated a concentration-dependent increase in activity (**Fig. 2.3b**). At the same time, the library at these concentrations did not exhibit any membrane disruption or toxicity as determined by a lactate dehydrogenase (LDH) release assay (**Fig. 2.3c**). We also found that testing the same concentrations of library aliquots containing an increased number of sequences, thereby increasing

diversity, showed no difference in activity as well as no demonstrated membrane toxicity as determined by the LDH release assay (**Fig. 2.3d**).

The PMO delivery activity of the library is likely energy-dependent, similar to PMO-CPP conjugates previously investigated.[19,20,52] A 1,000-member portion of the PMO-library was incubated with cells at 4 ˚C, conditions that arrest energy-dependent uptake. After incubation with the PMO-CPP conjugates, each well was washed extensively with PBS and heparin in order to disrupt and remove membrane-bound conjugates.[28] The cells were warmed back up to 37 ˚C and the assay continued in standard format and analyzed by flow cytometry. The significant decrease in library PMO delivery relative to PMO alone at 4 ˚C for both 5 μM and 20 μM library incubation conditions suggests energy-dependent uptake for the PMO-CPPs (**Fig. 2.3e**). Previous PMO-CPPs discovered in our laboratory are hypothesized to enter cells via clathrin-mediated endocytosis, as demonstrated through incubation with a panel of chemical endocytosis inhibitors, suggesting this could be a likely mechanism of energy-dependent uptake for a PMO-CPP library. [19,20,52] To confirm this, a 1,000 member library was tested with a series of endocytosis inhibitors in a pulse-chase format EGFP assay, in which HeLa 654 cells were pre-incubated with the chemical inhibitors to arrest various endocytosis pathways before PMO-CPPs were added. Following 3 h co-incubation, cells were washed extensively with heparin to dissociate membrane-bound constructs.[19,20,52] Results from experimental conditions were compared with a no inhibitor control where the cells were incubated with the PMO-peptide also for 3h. Activity of the library sample was reduced by 10 μM chlorpromazine (**Fig. 2.3f**), a known inhibitor of clathrin-mediated endocytosis.

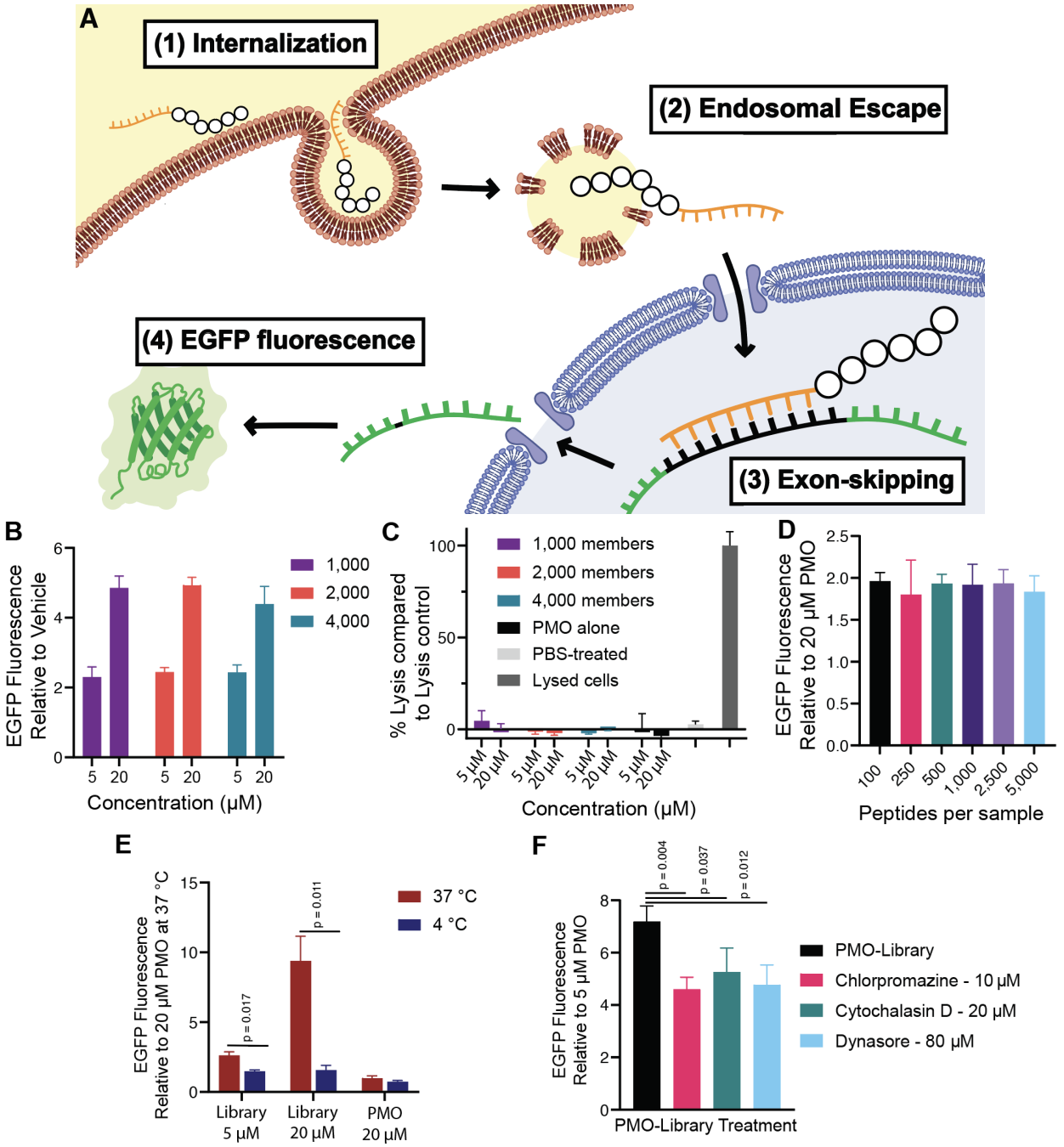


Figure 2.3 The CPP library can deliver PMO regardless of member size and enters via active transport.

(a) Graphic demonstrating the cell-based PMO activity assay. EGFP fluorescence correlates with the amount of PMO successfully delivered to the nucleus. (b) Bar graph showing EGFP fluorescence following HeLa 654 cell treatment with 5 or 20 μM of PMO-Library containing ~1000, ~2000, or ~4000 members for 22 h prior to flow cytometry. Results are given as the mean EGFP fluorescence of cells treated with PMO-peptide relative to the fluorescence of cells alone. PMO-Library samples show concentration-dependent PMO delivery at all library sizes tested. (c) Bar graph showing LDH release following HeLa 654 cell treatment with 5 or 20 μM PMO-Library for 22 h. Results are given as LDH release above vehicle relative to fully lysed cells. No compounds showed LDH release significantly above vehicle-treated cells. (d) Bar graph showing EGFP fluorescence following HeLa 654 cell treatment with 20 μM PMO-Library of

varying member sizes or 20 μM PMO alone for 22 h prior to flow cytometry. Results are given relative to the fluorescence of PMO-treated cells. There is no significant difference in EGFP fluorescence between the libraries of different sizes. (e) Bar graph showing EGFP fluorescence following HeLa 654 cell incubation at 4 ˚C or 37 ˚C for 30 min prior to treatment with 20 μM PMO-Library or 20 μM PMO alone for 2 h at the indicated temperature. After treatment, cells were washed with 0.1 mg/mL heparin and incubated in media for 22 h prior to flow cytometry. Results are given relative to the fluorescence of PMO-treated cells. There was a significant difference between the 4 ˚C and 37 ˚C treatment conditions at 5 μM ($p = 0.017$) and 20 μM ($p = 0.011$) of PMO-library. (f) Bar graph showing EGFP fluorescence following HeLa 654 cell treatment with different endocytosis inhibitors. The cells were pre-incubated for 30 min with the inhibitors, then treated with 10 μM PMO-library (1,000 members) for 3 h. The cells were washed with 0.1 mg/mL heparin and the media was exchanged for fresh, untreated media for 22 h prior to flow cytometry. At 10 μM chlorpromazine, EGFP fluorescence significantly decreased ($p = 0.004$). Bars represent mean ± SD, N = 3. P-values determined by two-sided unpaired student's t-test.

2.2.2 In-cell penetration selection-mass spectrometry

We subjected the library to the in-cell penetration selection-mass spectrometry platform (PS-MS) to discover sequences that are present in the cytosol, making them more likely to access the nucleus and the RNA target of the PMO cargo (**Fig. 2.4**). Our hypothesis is that while the peptides are screened in a mixture, meaning there is a possibility that some different sequences would act cooperatively with each other to enter the cell, this method would still reveal certain sequences that enter the cytosol at higher concentrations than others and thus have greater PMO delivery activity. The protocol for extracting biotinylated sequences was adapted from our recent method of profiling mixtures of PMO-D-CPPs from the cytosol and whole cell using MALDI-ToF.[30,53] Confluent HeLa cells in a 12-well plate were treated with 20 μM of biotin-library or PMO-biotin-library (~10^3 members, 3.5 nmol individual peptide per bead) for 1 h, and washed with PBS and heparin to dissociate membrane-bound conjugates. Cells were then lifted and extracellular conjugates digested with Trypsin, pelleted, and washed with PBS. Cells were gently lysed using either RIPA buffer (for whole cell extraction) or digitonin buffer (for cytosolic extraction).[30,54] Digitonin extraction is selective to cholesterol-rich membranes, leaving endosomes and some organelles intact while permeabilizing the cell membrane and semi-permeabilizing the nucleus.[55] Exclusion of endosomes in the cytosolic fraction was confirmed by Western blot, in which the endosomal marker, Rab5, is absent (**Appendix I**). Half of the no-treatment lysates were spiked with library as positive controls.

Biotinylated species in the lysates were affinity captured with magnetic streptavidin beads, and ultimately released by oxidative cleavage using brief incubation with sodium periodate. We

had previously used this cleavable linker to recover a single PMO-CPP conjugate from inside HeLa cells, and it was found to reliably cleave library peptides from streptavidin beads to isolate native peptides for sequencing by mass spectrometry.[50] The isolated peptides were desalted by solid-phase extraction and analyzed via Orbitrap tandem mass spectrometry using a mixed fragmentation method optimized for cationic peptides, consisting of electron-transfer dissociation (ETD), higher-energy ETD, and higher-energy collisional dissociation (HCD). Sequences matching the library design were then identified using a Python script.[33,45]

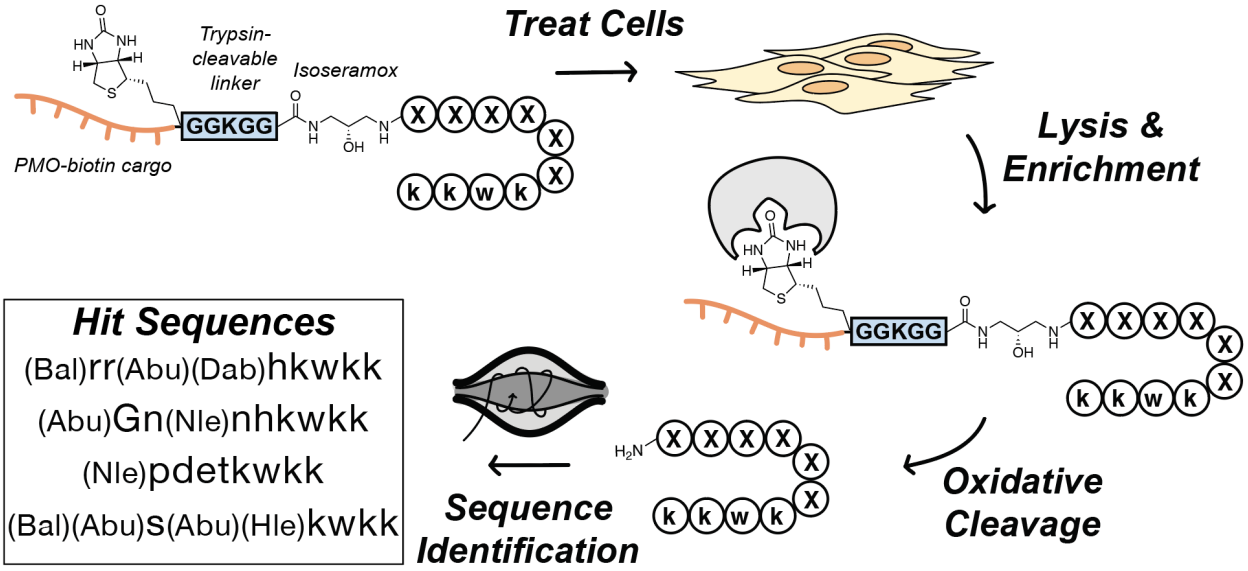


Figure 2.4 Workflow of in-cell penetration selection-mass spectrometry.
HeLa cells were treated with 20 μM PMO-biotin-library or biotin-library (1,000 members) for 1 h at 37 ˚C. Cells were then extensively washed with PBS and 0.1 mg/mL heparin before lysis with RIPA (whole cell extract) or digitonin (cytosolic extract). Lysates were incubated with magnetic streptavidin beads, and the C-terminal native peptides were cleaved from the beads under oxidative conditions. The peptides were desalted through solid-phase extraction and sequenced by LC-MS/MS. Hit PMO-delivering sequences were then identified as those peptides found only in the PMO-library fractions that do not overlap with peptides found in the cell only control or the samples treated with the biotin-library.

2.2.3 Hit peptides identified from in cell PS-MS show nuclear PMO delivery, but are not solely responsible for the PMO delivery activity of the library

Several hit peptides were selected for experimental validation and showed differential activities depending on the fraction in which they were found. We selected two sequences found exclusively in the cytosolic extract (Cyto1a, Cyto1b) and two from only the whole cell extract (WC1c, WC1d), with sequences shown in **Fig. 2.5a-b**. These peptides were synthesized via semi-automated solid-phase fast-flow peptide synthesis[56] with identical sequences to the library design

with the exception of a d-Ser residue to replace the isoseramox linker. These sequences were tested first in a concentration-response EGFP assay. The sequences extracted from the cytosol showed significantly increased activity compared to the sequences from the whole cell lysate, with Cyto1a showing an approximate EC50 of 25 μM compared to WC1c with EC50 of 400 μM (**Fig. 2.5c).** It was also confirmed that these sequences did not exhibit membrane toxicity at the concentrations tested (**Fig. 2.5d**). The peptides showed a positive correlation between charge and activity, with the highest performing peptide (Cyto1a) having a charge of +7, compared to WC1c with a charge of +2. This trend of positive charge correlating with PMO activity has been observed consistently in our lab.[19–21,52] Although the hit peptides do not show higher activity than the parent peptide penetratin in its native l-form, they do show less toxicity than penetratin at 25 μM in an LDH release assay (**Fig. 2.5d**), demonstrating their potential utility as PMO delivery vehicles. We also compared Cyto1a and WC1c to the known endosomal escape peptide Bpep,[30] composed of eight Arg residues interspaced with non-α-backbone residues β-alanine and 6-aminohexanoic acid, in its d-form which shows an EC50 closer to 3 μM (**Fig. 2.5e**). Interestingly, Cyto1a shares some similar motifs to Bpep, namely two Arg residues flanked by two non-α-backbone residues. If these motifs are responsible for endosomal escape, it is not surprising that Cyto1a was found in the cytosol and confirmed to have significant PMO delivery activity.

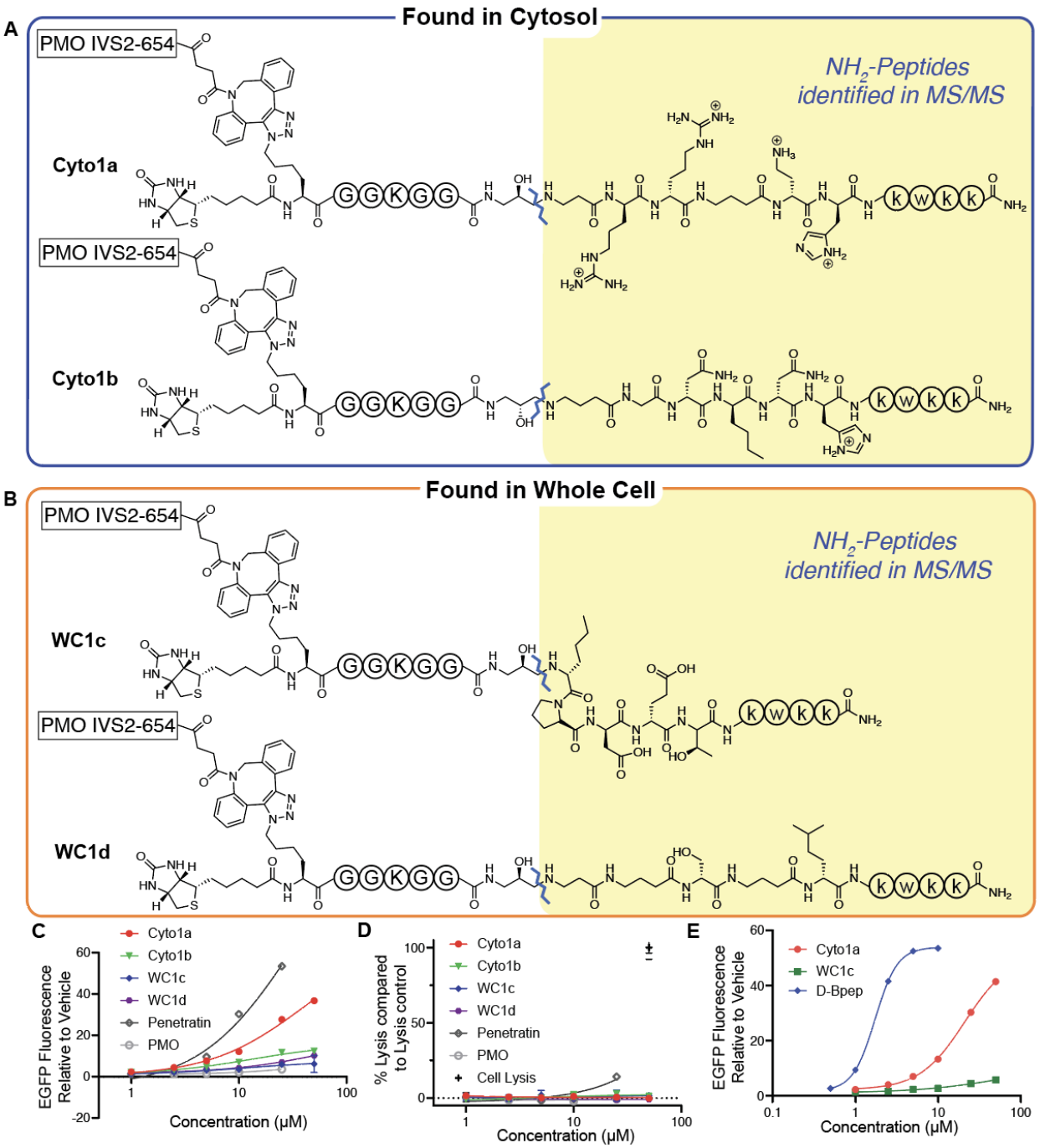


Figure 2.5 Candidate peptides from PS-MS demonstrate PMO delivery.
Shown are the sequences of the four candidate peptides, grouped by their extraction from either (s) cytosol or (b) whole-cell lysate. On the N-terminus are biotin and PMO IVS2-654 linked through a dibenzocyclooctyne-azidolysine. The fixed residues of the sequence are shown within circles while the regions unique to each sequence are fully drawn structures. The isoseramox linker is cleaved during library screening, and is substituted for d-serine in the PMO-peptides synthesized for hit validation. (c) Concentration-response graphs showing increase in EGFP fluorescence following HeLa 654 cells treatment with 1, 2.5, 5, 10, 25, or 50 μM PMO-CPP for 22 h prior to flow cytometry. Results are given as the mean EGFP fluorescence of cells treated with PMO-peptide relative to the fluorescence of cells treated with

vehicle only. (d) Cell supernatant from (e) was tested for LDH release. Results are given as percent LDH release above vehicle relative to fully lysed cells. Only l-penetratin showed significant ($p = 0.008$) LDH release above vehicle-treated cells. (e) Concentration-response graphs showing increase in EGFP fluorescence following HeLa 654 cells treatment with 0.5, 1, 2.5, 5, 10, 25, or 50 μM PMO-CPP for 22 h. Based on the average of three biological replicates, Cyto1a has a significantly lower EC50 (25 μM) compared to WC1c with EC50 of approximately 400 μM ($p = 0.0013$) estimated by constraining Emax to the Emax of D-bpep and Cyto1a. Bars represent mean ± SD, N = 3..

The discovered sequences rely on their D-chirality and noncanonical nature for serum stability, but not for PMO delivery. Although the candidate peptides contain mostly D-peptides, the trypsin-cleavable GGKGG sequence contains an L-lysine which may be cleaved in cells. We compared activity, toxicity, and serum stability between analogs of Cyto1a and WC1c which were all D-chirality (including the trypsin cleavage site), all L-chirality, and all L- and canonical residues. The canonical analogs contained natural substitutions for their unnatural residues (i.e., Lys for Dab and Leu for Nle). These changes in chirality and small changes in structure did not significantly impact the PMO delivery or membrane toxicity of the peptide constructs (**Fig. 2.6a-b**). In contrast, the serum stability of these sequences is greatly impacted. After a 9-hour incubation in 25% mouse serum, the L- and fully canonical sequences are significantly more degraded than the original an all D-sequences (**Fig. 2.6c-d**). This protease stability gives these newly discovered candidate peptides an advantage over the parent peptide penetratin, which is fully degraded in these conditions, and likely higher stability than other commonly used CPPs consisting of canonical L-amino acids. We would expect this additional stability to improve both the peptides' lifetime inside the cell, as well as in the serum in future in vivo studies.

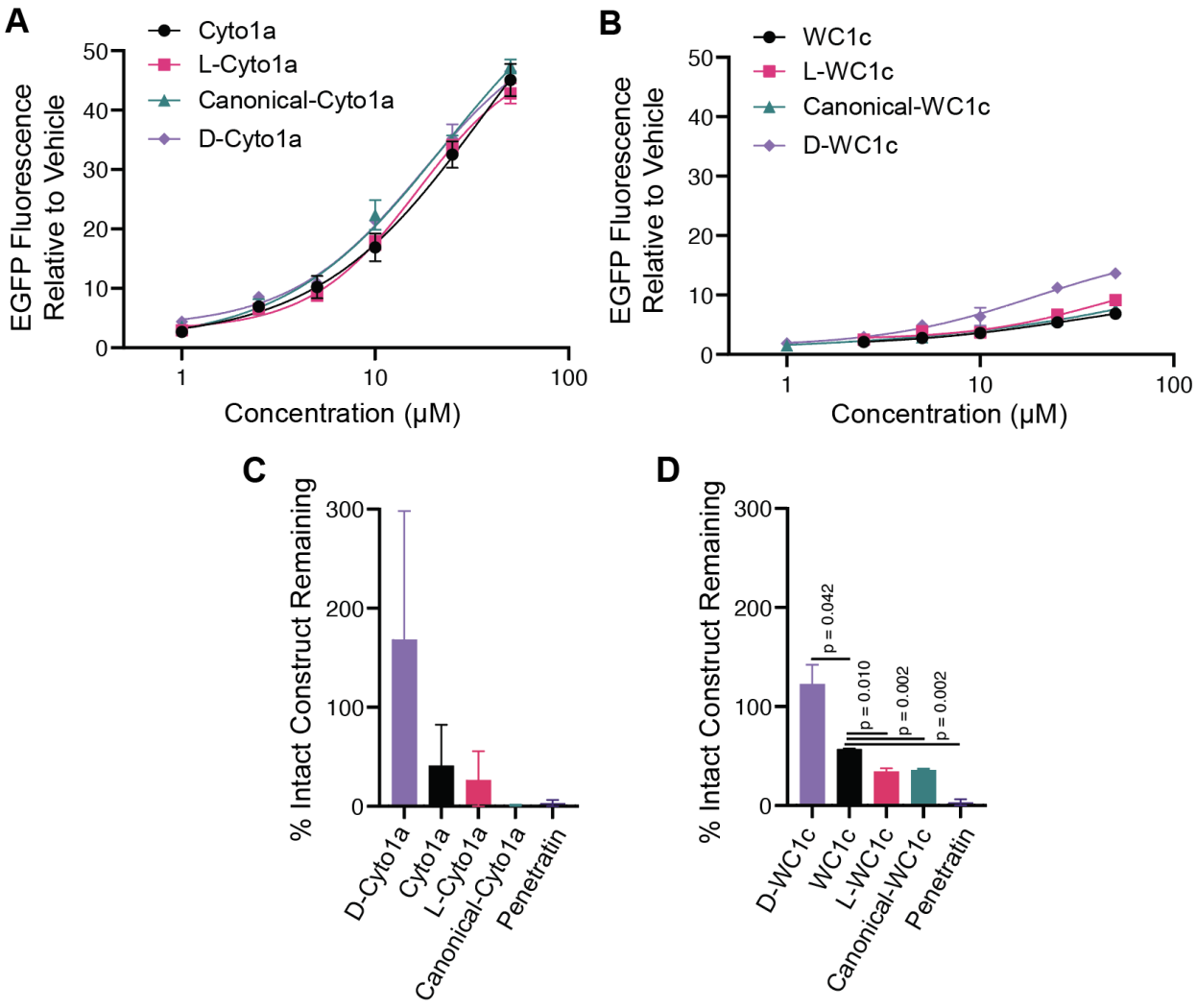


Figure 2.6 Candidate peptides Cyto1a and WC1c show higher protease stability than parent peptide penetratin or L/canonical versions of the candidate peptides.

(a-b) HeLa 654 cells were treated with 1, 2.5, 5, 10, 25, or 50 μM PMO-CPP for 22 h prior to flow cytometry. Results are given as the mean EGFP fluorescence of cells treated with PMO-peptide relative to the fluorescence of cells treated with vehicle only. Bars represent mean ± SD, N = 3. Variants of Cyto1a (a) or WC1c (b) were tested. L- indicates a peptide synthesized with all L-amino acids, D indicates a peptide synthesized with all-D amino acids (including the lysine in the trypsin cleavable GGKGG linker), and Canonical indicates a peptide synthesized with L, canonical amino acids. PMO-peptide variants were incubated in 25% mouse serum for 9 hours before quenching in 4 M guanidine hydrochloride. Samples were run on LCMS and peptide remaining was quantified through the Extracted Ion chromatogram and % peptide remaining was calculated relative to sample present at the 0 h timepoint. Variants of Cyto1a (c) or WC1c (d) were tested. At 9 hours, the concentration of PMO-WC1c was significantly less than PMO-D-WC1c ($p = 0.0421$) and significantly greater than PMO-L-WC1c, PMO-canonical WC1c, or PMO-penetratin ($p = 0.0101$, $p = 0.0019$, and $p = 0.0018$ respectively).

2.2.4 Hit peptide demonstrates high endosomal escape activity and PMO delivery

These PMO-CPPs likely enter the cell via clathrin-mediated endocytosis as we have found for previous constructs and the parent library. Cyto1a was tested with a series of chemical endocytosis inhibitors in the same pulse-chase format EGFP assay used to assess the endocytosis mechanism of the 1,000-member library sample. Activity of Cyto1a was impacted by 10 µM chlorpromazine (**Fig. 2.7a**). Chlorpromazine is a known inhibitor of clathrin-mediated endocytosis, indicating that this conjugate may participate in this pathway. It appeared that Cytochalasin D, Dynasore, and Wortmannin had no effect on activity. In addition to chlorpromazine, the 4 ˚C condition also significantly impacted the activities of each conjugate without causing toxicity, indicating that, like the entire library sample, uptake of the four hit peptides is energy-dependent, and the peptides are most likely entering the cells through endocytosis (**Fig. 2.7b**).

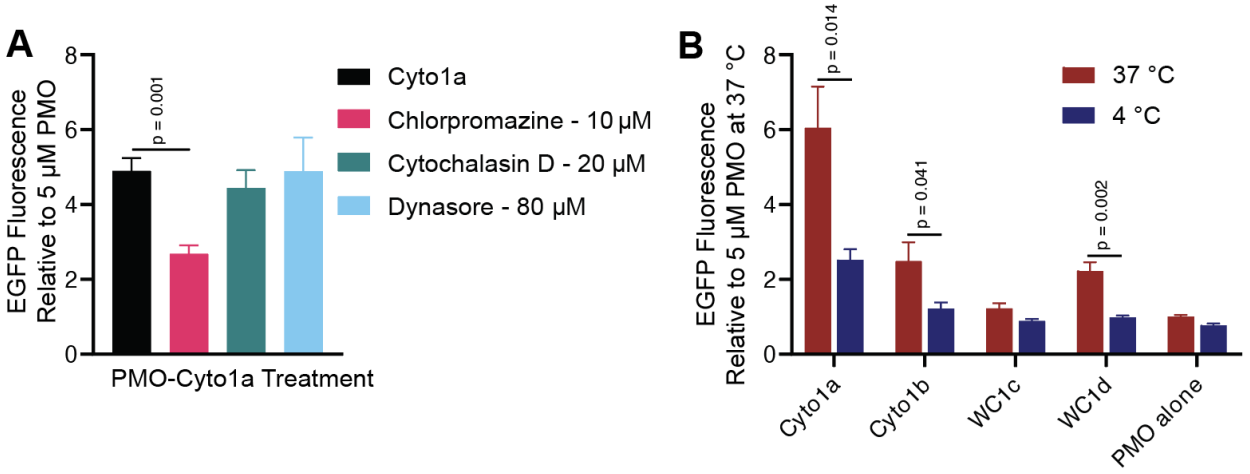


Figure 2.7 Hit peptides likely deliver PMO via clathrin-mediated endocytosis.
(a) Plot of EGFP mean fluorescence intensity relative to PMO for cells treated with different endocytosis inhibitors. The cells were pre-incubated for 30 min with the indicated compound and then 5 µM PMO-Cyto1a was added. After treatment with the construct for 3 h, the cells were washed with 0.1 mg/mL heparin and the media was exchanged for fresh, untreated media for 22 h prior to flow cytometry. Results are given as the mean EGFP fluorescence of cells treated with PMO-peptide relative to the fluorescence of cells treated with 5 µM PMO. At 10 µM chlorpromazine, EGFP fluorescence significantly decreased ($p = 0.001$). (b) Plot of EGFP mean fluorescence intensity relative to PMO for cells incubated with PMO-CPPs at 4 ˚C or 37 ˚C. The cells were pre-incubated for 30 min at 4 ˚C or 37 ˚C, followed by the addition of PMO-peptide conjugate to each well at a concentration of 5 µM. After incubation at 4˚C or 37 ˚C for 2 h, the cells were washed with 0.1 mg/mL heparin and the media was exchanged for fresh, untreated media for 22 h prior to flow cytometry. Results are given as the mean EGFP fluorescence of cells treated with PMO-peptide relative to the fluorescence of cells treated with 5 µm PMO at 37˚C. Bars represent mean ± SD, N = 3.

We further investigated the differences in PMO delivery activity and total uptake between the sequences found in the cytosol versus the whole cell lysate and compared them to a benchmark compound, PMO-D-Bpep, using flow cytometry.[30] For this purpose, several SulfoCy5-labeled PMO-CPPs were generated and tested to ensure the fluorophore did not impact PMO delivery activity (**Fig. 2.8a**). Comparing the results of the EGFP assay of conjugates with and without the fluorophore, no significant differences were found between the constructs' EC50 values (**Fig. 2.8b-d**).

After confirming the Cy5 label did not interfere with efficacy, we then looked at the uptake and nuclear delivery of the PMO-SulfoCy5-CPPs. This study was performed by monitoring both EGFP fluorescence elicited by nuclear-localized compound and Cy5 fluorescence to measure total cellular uptake of each analog. Each conjugate demonstrated similar concentration-dependent increases in both EGFP and Cy5 fluorescence without showing cytotoxicity (**Fig. 2.8e-g**). WC1c showed low fluorescence signal in both the EGFP and Cy5 channels. On the other hand, Cyto1a showed higher fluorescence signals in each channel, indicating greater uptake and nuclear localization compared to WC1c. Moreover, when comparing relative fluorescence between these two constructs, it appears that WC1c has poorer PMO activity compared to total uptake, while Cyto1a has greater PMO activity compared to total uptake. This property is seen even more dramatically with our positive control peptide Bpep, which has well-studied endosomal escape properties. This readout underscores the utility of the PS-MS methodology in identifying peptides that can efficiently escape endosomes.

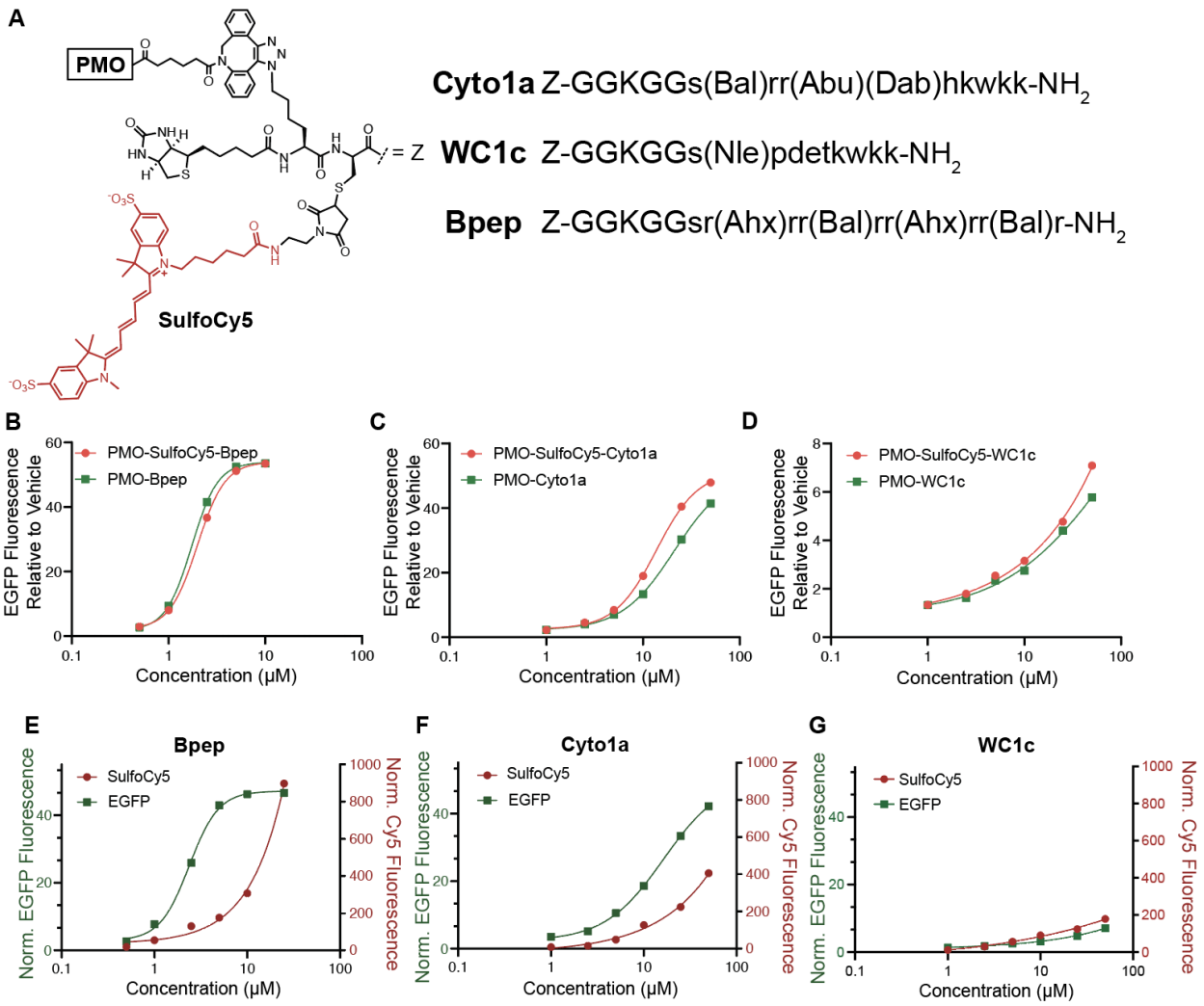


Figure 2.8 SulfoCy5 label does not impact PMO delivery and reveals differential PMO delivery efficiency.

(a) Sequences of PMO-SulfoCy5-CPP constructs, with the N-terminal cargo fully drawn out (Z). Lowercase letters denote d-amino acids. (b-d) HeLa 654 cells were treated with 1, 2.5, 5, 10, 25, or 50 μM PMO-CPP or PMO-SulfoCy5-CPP for 22 h prior to flow-cytometry. Results are given as the mean EGFP fluorescence of cells treated with PMO-peptide relative to the fluorescence of cells treated with vehicle only. (b) Treatment with d-Bpep constructs. (c) Treatment with Cyto1a constructs. (d) Treatment with WC1c constructs. Cyto1a showed no significant difference with the addition of SulfoCy5 at 5 μM and below ($p < 0.05$), while WC1c and d-Bpep showed no significant difference between the SulfoCy5 and standard constructs at 25 μM and below ($p < 0.05$). No peptides had significant differences in EC50 between SulfoCy5 and unlabeled constructs. Bal (beta-Alanine), Abu (γ-aminobutyric acid), Dab (d-diaminobutyric acid), Nle (d-Norleucine), Ahx (6-aminohexanoic acid). (e-g) HeLa 654 cells were treated with 1, 2.5, 5, 10, 25, or 50 μM PMO-SulfoCy5-CPP for 22 h prior to flow cytometry. Results are given as the mean fluorescence of cells treated with PMO-SulfoCy5-peptide relative to the fluorescence of cells treated with vehicle only for each channel. (e) Treatment with d-Bpep constructs. (f) Treatment with Cyto1a constructs. (g) Treatment with WC1c constructs. Bars represent mean ± SD, N = 3.

The uptake of the fluorescent conjugates into HeLa cells was also evaluated via confocal microscopy. HeLa cells were treated with 5 μM or 25 μM of SulfoCy5-labeled conjugates for 30 min, followed by a wash in complete media. This shorter incubation, compared to the 22 h PMO activity assay, is required here to view the early stages of uptake and endosomal escape, which occurs within the first hour of incubation. Hoechst and LysoTracker Green in complete media were added to the cells immediately prior to imaging. D-Bpep was again used as a control as it is a sequence known to efficiently escape the endosome and localize to the nucleus.[30] Indeed, at both concentrations diffuse fluorescence was observed in the cytosol and nucleus of PMO-Bpep-treated cells, in addition to punctate fluorescence that co-localized with LysoTracker, suggesting accumulation in endosomes (**Fig. 2.9**). Cyto1a also demonstrated perinuclear localization, especially at the higher concentration (**Fig. 2.9b**). In contrast, WC1c showed significantly reduced overall fluorescence inside the cell, and exclusively as punctate fluorescence within endosomes.

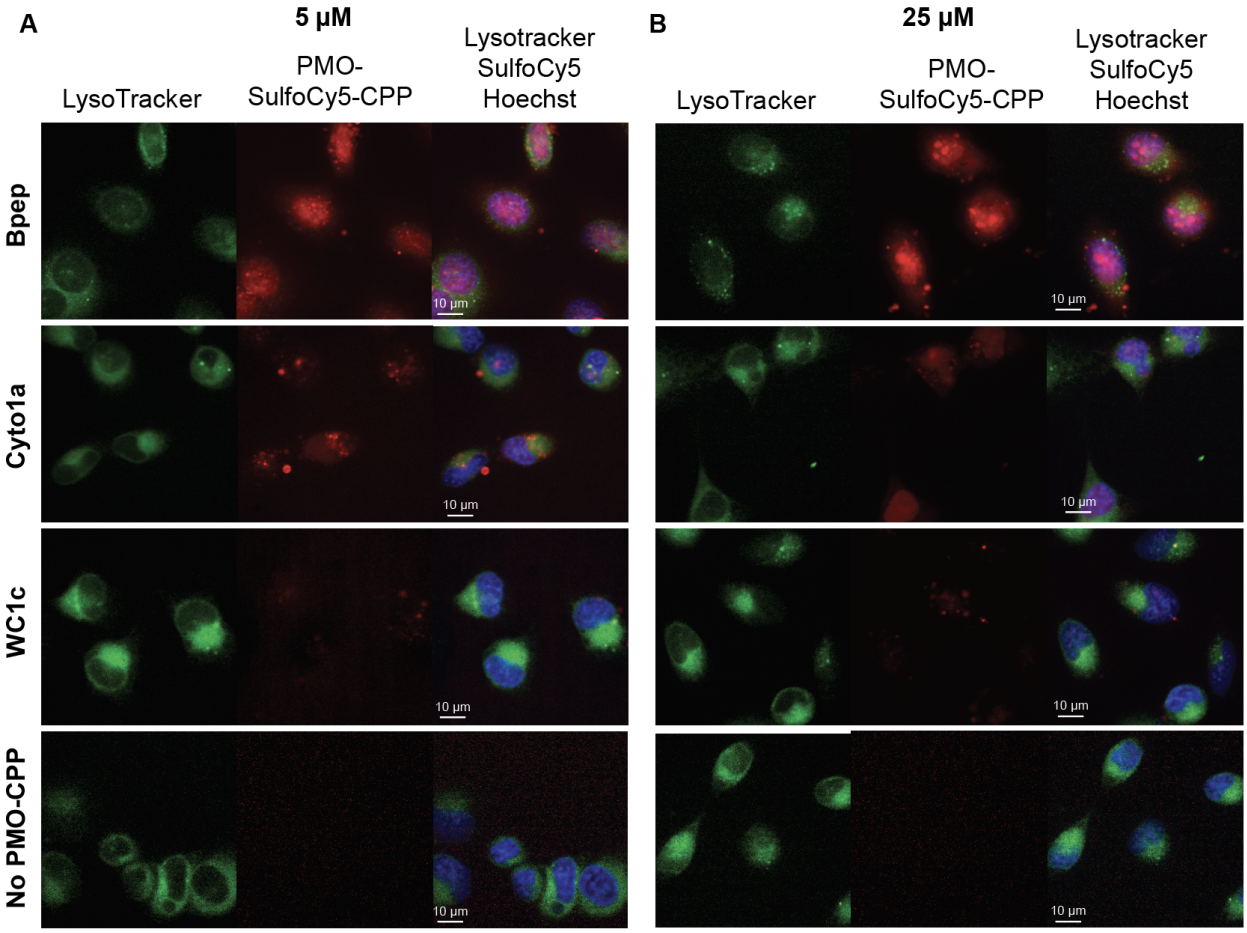


Figure 2.9 Cyto1a localizes to cytosol and nucleus.
Confocal micrographs of HeLa cells treated with (a) 5 µM or (b) 25 µM PMO-SulfoCy5-Bpep, PMO-SulfoCy5-Cyto1a, or PMO-SulfoCy5-WC1c. Hoechst labels the nuclei and Lyostracker Green labels the endosomes. SulfoCy5-labeled PMO-CPPs can be observed in the endosomes, cytosol, and nuclei of the cells.

2.3 Conclusions

Affinity selection-mass spectrometry (AS-MS) techniques have traditionally been used to probe protein-protein interactions in vitro.[34] Our group has recently shown that chemical libraries may reach the diversity of other display techniques for identification of peptide binders to proteins[33], and that this strategy can be applied for cell-surface selection in vivo.[42] In this work, we have added an additional spatial element to this strategy, by extracting the cytosol for in-cell selection of fully synthetic peptide libraries conjugated to a model antisense cargo. By comparing these sequences to those found in a whole cell extract, we can exclude sequences that accumulate in the endosomes.

In-cell PS-MS combined with subcellular fractionation resulted in the identification of a novel, noncanonical peptide capable of accessing the cytosol and delivering PMO to the nucleus. Cyto1a, like the positive control peptide Bpep, was able to deliver PMO to the nucleus by escaping endosomes. Furthermore, Cyto1a does not appear to permeabilize the endosome to allow the escape of other endosomal cargo, nor does it demonstrate cell membrane toxicity at the measured concentrations in an LDH release assay. All peptides discovered through this novel platform demonstrated lower toxicity and higher protease stability than the CPP penetratin, which contributed the fixed "CPP-like" C-terminal region in the library. Endowed with lower toxicity, higher stability, and superior chemical diversity provided by the noncanonical residues, the peptides discovered with the in-cell PS-MS platform show advantages over the library's parent peptide. The few active PMO-CPPs individually sequenced and validated are not solely responsible for the overall cell penetration of the library, however. In fact, it is more likely that the library's overall cell penetration activity arises from the combined activity of a number of peptides at low concentration, including the hits discovered with our platform, despite the variable activity of the individual CPPs within the library. Cooperative activities of peptides in a mixture is an avenue that warrants further investigation. We hypothesized that screening a library of peptides for cytosolic penetration would identify certain sequences that accessed the cytosol with higher efficacy. Indeed, in cell PS-MS screening of the library resulted in the isolation of ~2% of the peptides from the cytosol, indicating that these specific sequences entered the cytosol at higher concentrations than other library members.

Current investigations in our laboratory aim to further combine this method with orthogonal approaches in order to better focus the selection on successful peptides. As such, we envisage using chromatographic fractionation of the peptide libraries to enrich for penetrant peptides within the library before conducting the PS-MS uptake assay. This chromatographic pre-enrichment would allow for screening of an overall less active, but more diverse library, perhaps obviating the need for a pre-installed fixed C-terminal penetrating motif.

2.4 Experimental Section

2.4.1 Reagents and Solvents

H-Rink Amide-ChemMatrix resin was obtained from PCAS BioMatrix Inc. (St-Jean-sur-Richelieu, Quebec, Canada) and 300 μm TentaGel resin was obtained from Rapp Polymere (Tuebingen, Germany). 1-[Bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-b]pyridinium-3-oxid-hexafluorophosphate (HATU), Fmoc-L-Lys(N_3), Fmoc-β-Ala-OH, Fmoc-norleucine, *N*α-Fmoc-*N*γ-Boc-D-2,4-diaminobutyric acid, Fmoc-D-homoleucine, Fmoc-3,3-diphenyl-D-alanine, Fmoc-3-(1-naphthyl)-D-alanine, Fmoc-4-(Boc-aminomethyl)-D-phenylalanine, 1-Boc-piperidine-4-Fmoc-amino-4-carboxylic acid, Fmoc-γ-aminobutyric acid, and 2-(Fmoc-amino)-4-(bis-Boc-guanidino)-D-butyric acid were purchased from Chem-Impex International (Wood Dale, IL). Fmoc-protected D-amino acids (Fmoc-Arg(Pbf)-OH; Fmoc-Asn(Trt)-OH; Fmoc-Asp-(O*t*-Bu)-OH; Fmoc-Gln(Trt)-OH; Fmoc-Glu(O*t*-Bu)-OH; Fmoc-Gly-OH; Fmoc-His(Trt)-OH; Fmoc-Lys(Boc)-OH; Fmoc-Phe-OH; Fmoc-Pro-OH; Fmoc-Ser(But)-OH; Fmoc-Thr(*t*-Bu)-OH; Fmoc-Trp(Boc)-OH), were purchased from the Novabiochem-line from MilliporeSigma. Sulfo-Cyanine5 maleimide was purchased from Lumiprobe (Cockeysville, MD), and 7-diethylaminocoumarin-3-carboxylic acid was purchased from AAT Bioquest (Sunnyvale, CA). Dibenzocyclooctyne acid was purchased from Click Chemistry Tools (Scottsdale, AZ). Cytochalasin D was obtained from Santa Cruz Biotech. Hoechst 33342 and LysoTracker™ Green were purchased from ThermoFisher Scientific (Walthan, MA). The LDH Assay kit was purchased from Promega (Madison, WI). PMO IVS2-654 was provided by Sarepta Therapeutics (Cambridge, MA). Peptide synthesis-grade *N,N*-dimethylformamide (DMF), CH_2Cl_2 (DCM), diethyl ether, and HPLC-grade acetonitrile were obtained from VWR International (Radnor, PA). All other reagents were purchased from Sigma-Aldrich (St. Louis, MO). Milli-Q water was used exclusively.

2.4.2 Liquid chromatography-mass spectrometry

LC-MS analyses were performed on an Agilent 6550 iFunnel Q-TOF LC-MS system (abbreviated as 6550) coupled to an Agilent 1290 Infinity HPLC system. Mobile phases were: 0.1% formic acid in water (solvent A) and 0.1% formic acid in acetonitrile (solvent B). The following LC-MS method was used for characterization:

Method A: 1-61% B over 6 min, Zorbax C3 column (6550)

LC: Agilent Zorbax C3 RRHD column: 2.1 × 50 mm, 1.8 μm, column temperature: 40 °C, gradient: 0-1 min 1% B, 1-6 min, 1-61% B, 6-7 min, 91% B, 7-8 min, 1% B; flow rate: 0.5 mL/min.

MS: Positive electrospray ionization (ESI) extended dynamic range mode in mass range 300–3000 m/z. MS is on from 1 to 6 min.

Data were processed using the Agilent MassHunter software package. Y-axis in all chromatograms shown represents total ion chromatogram (TIC).

2.4.3 General peptide preparation

Fast-flow Peptide Synthesis. Peptides were synthesized using an automated fast-flow peptide synthesizer for L-peptides [59] and a semi-automated fast-flow peptide synthesizer for D-peptides and non-canonical L-peptides as previously reported.[56] Briefly, all reagents were flowed at 40 mL/min over 150 mg of ChemMatrix Rink Amide Resin at 90ºC. Amino acids were activated with HATU or PYAOP and excess diisopropylethylamine. Fmoc was removed with 20% (v/v) piperidine, and extensive DMF washes were performed between each step.

Peptide Cleavage, Deprotection, and Purification. Each peptide was deprotected and cleaved from resin by treatment with 5 mL of 94% trifluoroacetic acid (TFA), 2.5% thioanisole, 2.5% water, and 1% triisopropylsilane (TIPS) (v/v) at room temperature for 2 to 4 h. The cleaved peptide was triturated as previously reported.[30]

Peptide Purification. The peptides were purified by mass-directed semi-preparative reversed-phase HPLC as previously described, and peptide purity was confirmed by LC-MS.[30]

Preparation of PMO-Peptides. PMO-peptides were prepared as previously described through an aqueous strain-promoted azide-alkyne cycloaddition between PMO-DBCO and azido-peptides.30 The purity of the final construct was confirmed by LC-MS to be >95%.

2.4.4 Preparation of peptide libraries

The peptide library was prepared as previously described through split-and-pool solid-phase synthesis on 300 μm TentaGel resin (0.23 mmol/g) for a 95,000 member library.[45]

2.4.5 EGFP Assay

HeLa 654 cells were obtained from the University of North Carolina Tissue Culture Core facility and assayed as previously described.[30] Briefly, PMO-peptides were dissolved in PBS before being diluted in MEM to the target concentration and incubated on cells plated 18 hours prior at 5,000 cells/well in a 96-well plate for 22 hours. Flow cytometry analysis was carried out on a BD LSRII flow cytometer. Cells that were positive for propidium iodide or fell outside of the main population in their forward/side scatter readings were excluded from analysis through gating.

Flow cytometry data was analyzed using FlowJo and the mean fluorescence intensities of the samples were graphed and analyzed with Graphpad Prism7.

2.4.6 Endocytosis Inhibition Assay

Chemical endocytosis inhibitors were used to probe the mechanism of delivery of PMO by these peptides in a pulse-chase format. We have conducted such analyses on similar PMO-peptide constructs previously with comparable outcomes.52 18 h prior to treatment, HeLa 654 cells were plated at a density of 5,000 cells per well in a 96-well plate in MEM supplemented with 10% FBS and 1% penicillin-streptomycin. For the PMO constructs, cells were preincubated with various chemical endocytosis inhibitors for 45 minutes before treatment with PMO-CPP constructs for three hours. Treatment media was then replaced with fresh media and the cells were incubated for 22 h at 37 °C and 5% CO2. Cells were then lifted as previously described and EGFP synthesis was measured by flow cytometry. The panel of endocytosis inhibitors included: 10, 20, and 80 μM chlorpromazine (CPZ), which is demonstrated to interfere with clathrin-mediated endocytosis; 20 μM cytochalasin D (CyD), which inhibits phagocytosis and micropinocytosis; 200 nM wortmannin (Wrt), which alters various endocytosis pathways by inhibiting phosphatidylinositol kinases; 50 μM EIPA (5-(N-ethyl-N-isopropyl)amiloride), which inhibits micropinocytosis; and 80 μM Dynasore (Dyn), which also inhibits clathrin-mediated endocytosis.[60,61]

2.4.7 LDH Assay

Cytotoxicity assays were performed on the same HeLa 654 cells used for flow cytometry as previously described.30 Cell supernatant following treatment for the PMO delivery activity assay was transferred to a new 96-well plate for analysis of LDH release with the CytoTox 96 assay kit (Promega). The measurement of vehicle-treated cells was subtracted from each

measurement, and % LDH release was calculated as % cytotoxicity = 100 × Experimental LDH Release (OD490) / Maximum LDH Release (OD490).

2.4.8 Serum Stability Assay

Serum stability assays were performed by incubating the PMO-peptide in 25% mouse serum at 50 μM. Timepoints were taken in duplicate at 0, 1, 3, 9, and 27 h. Each timepoint was quenched in 4 M guanidine hydrochloride to dissociate peptide from serum proteins, and then diluted to 1 M guanidine hydrochloride for LC-MS analysis. 0.5 μL of each sample was injected onto the LCMS and analyzed by method B. For each compound, remaining peptide in serum was quantified by the Extracted Ion Chromatogram (EIC) using MassHunter software. The ion count was then normalized to the 0 h EIC to determine the % peptide remaining in serum.

2.4.9 Microscopy

HeLa cells were plated at a density of 8,000 cells/well in a 96-well cover glass-bottomed plate and incubated overnight in MEM supplemented with 10% (v/v) fetal bovine serum (FBS) and 1% (v/v) penicillin-streptomycin at 37 °C and 5% CO_2. For standard localization imaging, cells were treated with PMO-Sulfo-Cy5-CPP conjugates at 5 μM or 25 μM for 30 min in supplemented MEM as above. Each well was washed with media and incubated in fresh media for 1 h before 1.5 μM Hoechst (nuclear) and 50 nM Lysotracker Green (endosomal) fluorescent tracking dyes were added, and imaged immediately. The endosomal release experiment was adapted from a previously reported protocol.[58] Cells were treated with 50 μM 7-Diethylaminocoumarin-3-carboxylic acid (DEAC)-k5 for 1 h at 37 °C and 5% CO_2 before being washed with media. Then, PMO-Sulfo-Cy5-CPP conjugates were added as before. Sytox Green was added immediately before imaging in order to exclude observation of nonviable cells. Imaging was performed at the Whitehead Institute's Keck Imaging Facility on an RPI Spinning Disk Confocal Microscope at 40x objective.

2.4.10 Uptake Assay

<u>Cell treatment</u>

HeLa cells in a 6-well or 12-well plate were co-incubated with library at 20 μM at 37ºC following a previously described procedure.[30]

Lysis

To acquire whole cell lysate, 50 µL RIPA buffer, protease inhibitor cocktail, water was added to the cell pellet, mixed gently, and placed on ice for 1 h. To extract the cytosol, 50 µL digitonin buffer (0.05 mg/mL digitonin, 250 mM sucrose, PBS) was added to a cell pellet, pipetted up and down a single time to resuspend the cells, and placed on ice for 10 min. Cytosol or whole-cell extracts were then prepared as previously described.30 10µg protein from the untreated control samples was analyzed by SDS-PAGE gel electrophoresis and then transferred to a nitrocellulose membrane for immunostaining with anti-Erk1/2 and anti-Rab5.[30]

Penetration Selection

10 µL Dynabeads™ MyOne™ Streptavidin T1 (Thermo Fisher) were transferred to tubes in a magnet stand and washed with PBS. Cell extracts were added to the corresponding bead-containing tube and rotated at 4 °C for 2 h. To two of the cell-only sample lysates was added each 0.5 µL of PMO-biotin-library and 0.5 µL of biotin-library and was combined with 50 µL of Streptavidin beads. Following pulldown, the beads were washed with 6 M guanidinium chloride (GuHCl, pH 6.8, 2 x 200 µL) and suspended in 100 µL PBS. Then, NaIO4 (1 mM in H2O, 2 µL) was added and the beads incubated for 5 min in absence of light, followed by quench solutions: Na2SO3 (100 mM, 5 µL) and NH2OH (100 mM, 5 µL). The supernatants were transferred to new tubes and the beads were washed with 6 M GuHCl (2 x 100 µL). Pooled supernatant fractions were then desalted by solid-phase extraction (SPE) using C18 ZipTips, lyophilized, and rehydrated in 10 µL 1 M GuHCl in water containing 0.1% formic acid.

Orbitrap LC-MS/MS

Analysis was performed on an EASY-nLC 1200 (Thermo Fisher Scientific) nano-liquid chromatography handling system connected to an Orbitrap Fusion Lumos Tribrid Mass Spectrometer (Thermo Fisher Scientific). The standard nano-LC method was run at 40 °C and a flow rate of 300 nL/min with the following gradient: 1% solvent B in solvent A ramping linearly to 41% B in A over 55 min, where solvent A = water (0.1% FA), and solvent B = 80% acetonitrile, 20% water (0.1% FA). Fragmentation for secondary MS was carried out with higher-energy collisional dissociation and electron transfer/higher-energy collisional dissociation energy. Instrument parameters were set as previously described.[33]

De novo peptide sequencing and filtering

De novo peptide sequencing of the acquired data was performed in PEAKS 8 (BioInformatics Solutions Inc.). Using PEAKS, spectra were prefiltered to remove noise and sequenced. All non-canonical amino acids were sequenced as post-translational modifications based on the canonical amino acid most closely matching their molecular mass. Norleucine and β-Alanine were sequenced as leucine and alanine, respectively. In addition, His-oxide and Trp-oxide were allowed as variable post-translational modifications. Twenty candidate sequence assignments were created for each secondary scan.

An automated Python-based routine was used for postprocessing data analysis to eliminate noise, synthetic impurities, duplicates, resolve certain sequencing ambiguities, and to select the best candidate sequence assignment for each MS/MS scan.45 The script eliminates all sequence candidates of length other than 9 or 10 and all candidates not bearing the C-terminal KWKK motif.

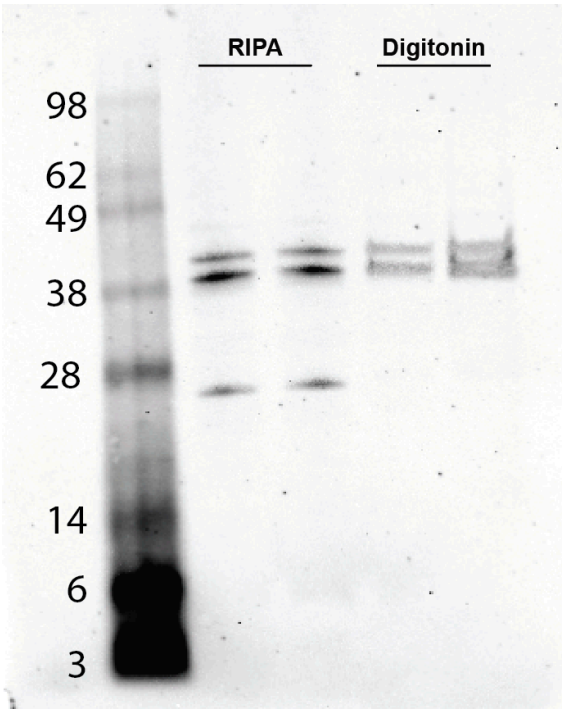
2.4.11 Statistics

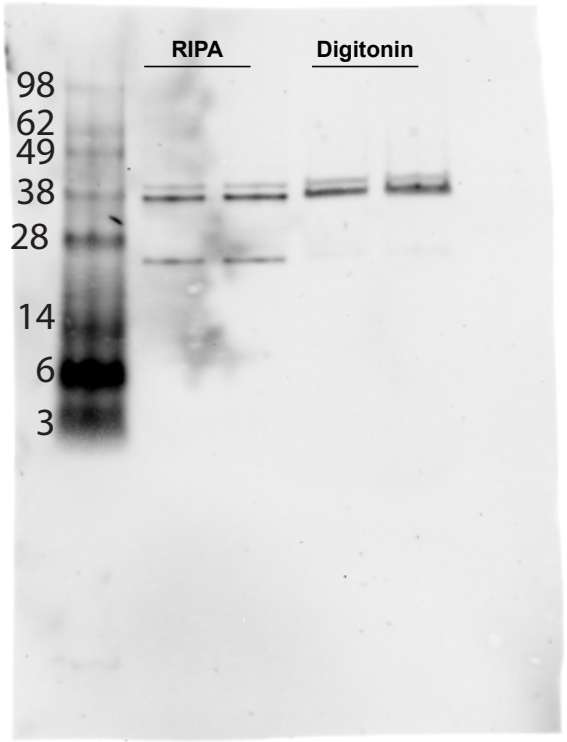
Statistical analysis and graphing were performed using Prism (Graphpad) or Excel (Microsoft). Concentration-response curves were fitted using Prism nonlinear regression. All experiments were performed in triplicate unless otherwise indicated, with at least two biological replicates. Significance for activities between constructs was determined using a student's two-sided, unpaired t-test.

2.5 Acknowledgements

This research was funded by Sarepta Therapeutics. C.K. Schissel (4000057398) and C.E. Farquhar (4000057441) acknowledge the National Science Foundation Graduate Research Fellowship (NSF Grant No. 1122374) for research support. We acknowledge support from the Swanson Biotechnology Center Flow Cytometry Facility at the Koch Institute for Integrative Cancer Research at MIT through the use of their flow cytometers (NCI Cancer Center Support Grant P30-CA14051). We thank C. Rogers and B. Braswell at the Whitehead Institute's Keck Imaging facility for use of the confocal microscope.

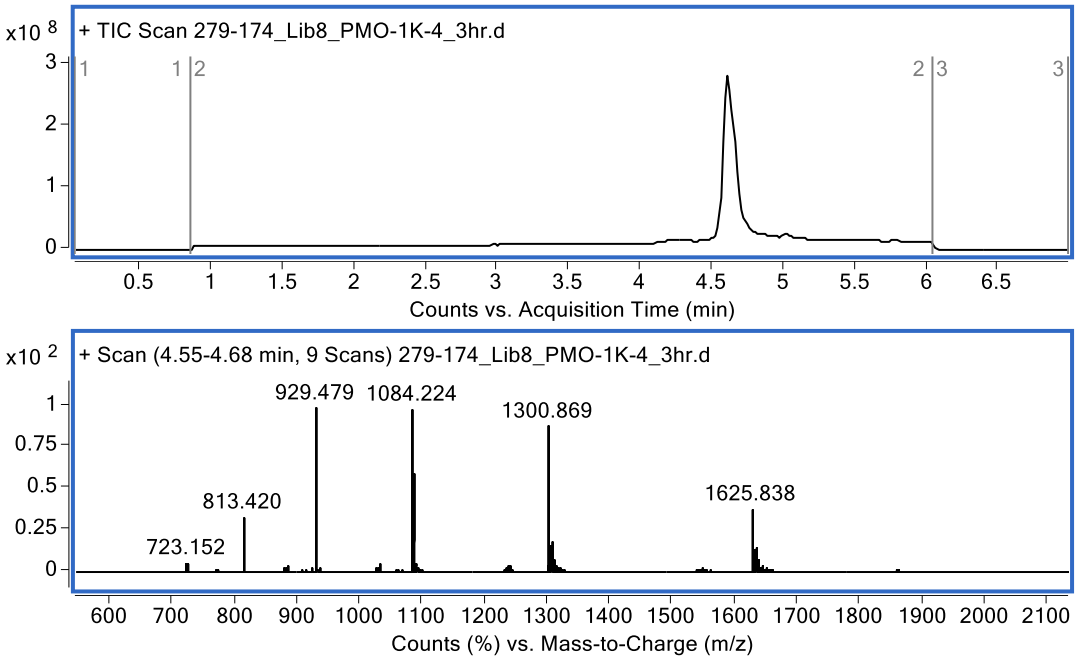
2.6 Appendix 1: Full Western Blot images



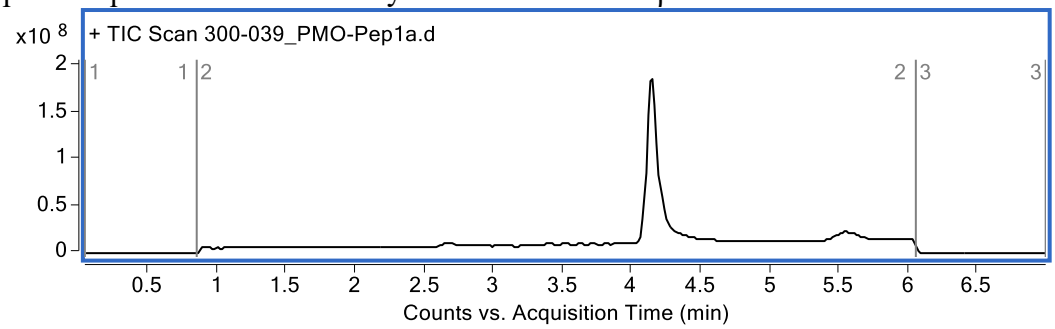


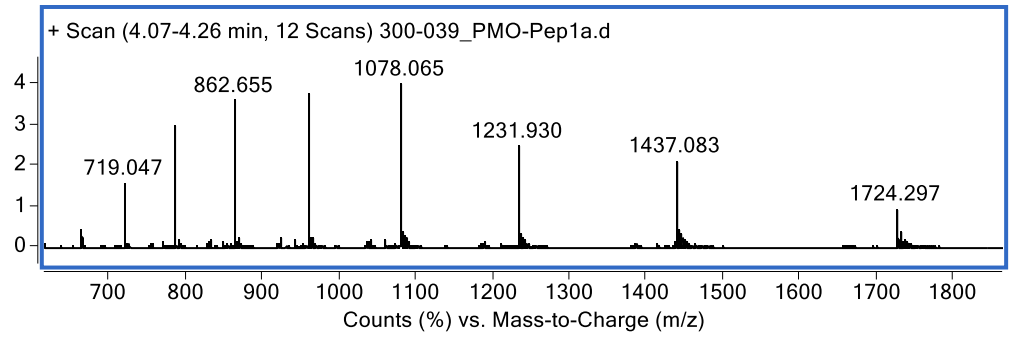
2.7 Appendix II: Characterization of compounds

PMO-DBCO
Mass Expected: 6500.0
Mass Observed: 6499.9

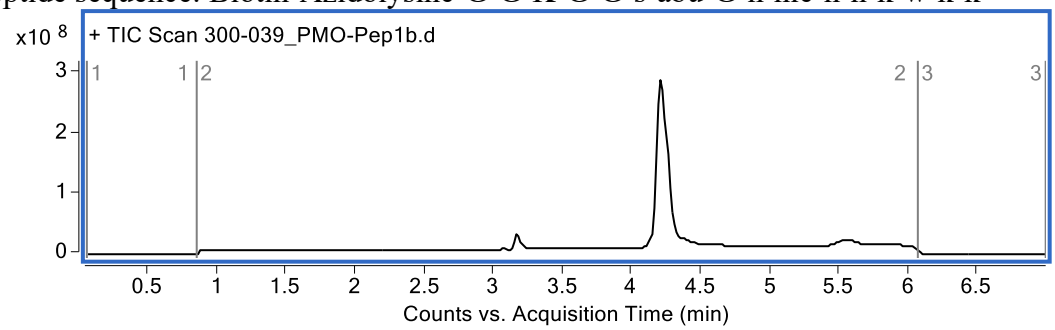


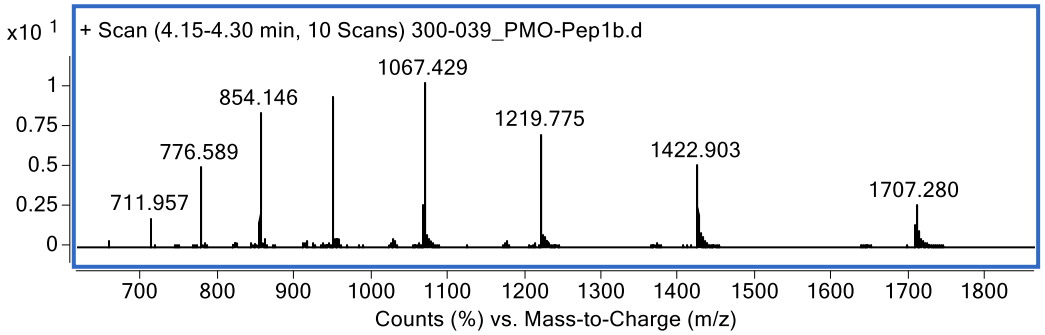
PMO-Cyto1a
Mass Expected: 8617.4
Mass Observed: 8617.5
Peptide sequence: Biotin-Azidolysine-G-G-K-G-G-s-βal-r-r-abu-dab-h-k-w-k-k



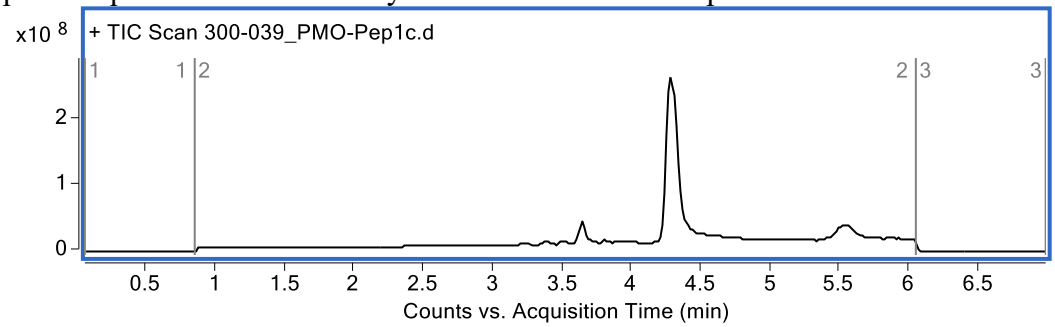


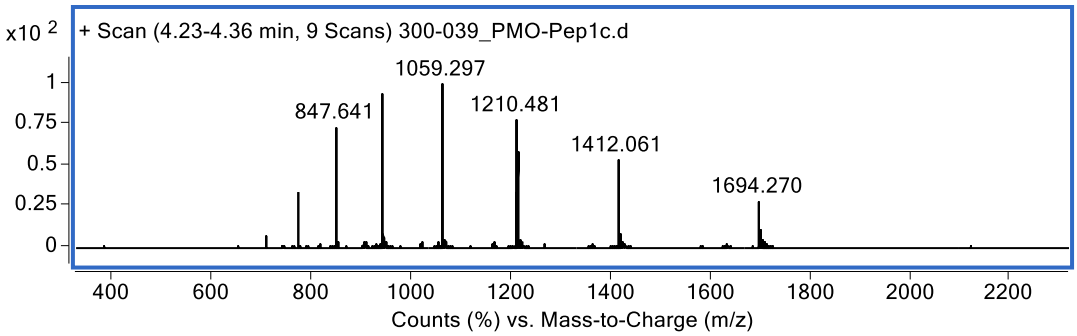
PMO-Cyto1b
Mass Expected: 8532.4
Mass Observed: 8532.4
Peptide sequence: Biotin-Azidolysine-G-G-K-G-G-s-abu-G-n-nle-n-h-k-w-k-k





PMO-WC1c
Mass Expected: 8467.3
Mass Observed: 8467.3
Peptide sequence: Biotin-Azidolysine-G-G-K-G-G-s-nle-p-d-e-t-k-w-k-k



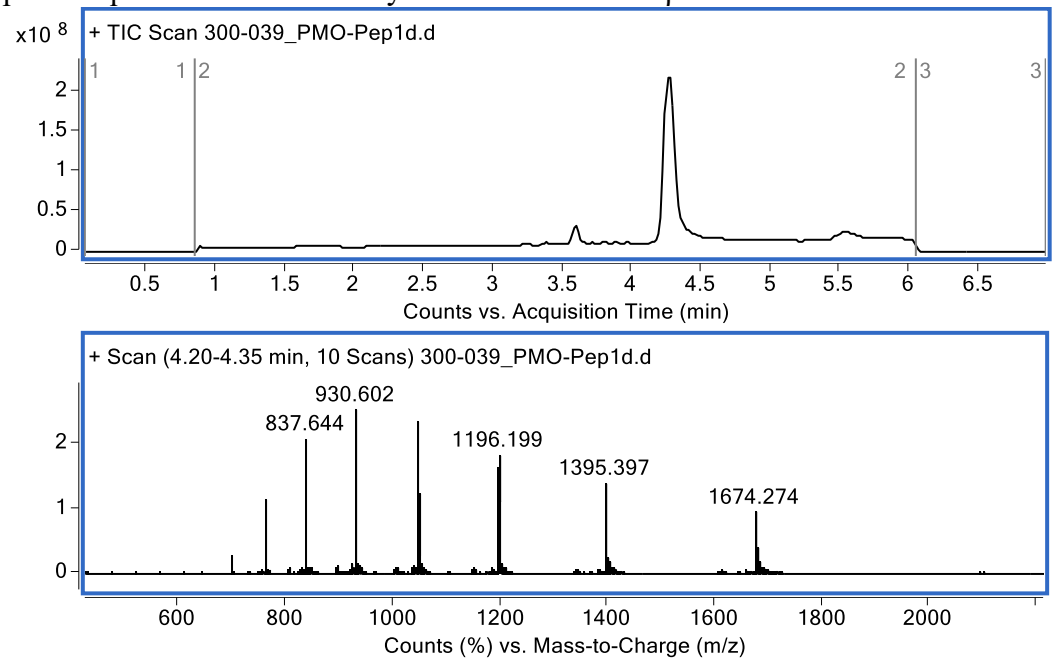


PMO-WC1d

Mass Expected: 8367.2

Mass Observed: 8367.2

Peptide sequence: Biotin-Azidolysine- G-G-K-G-G-s-βal-abu-s-abu-hle-k-w-k-k

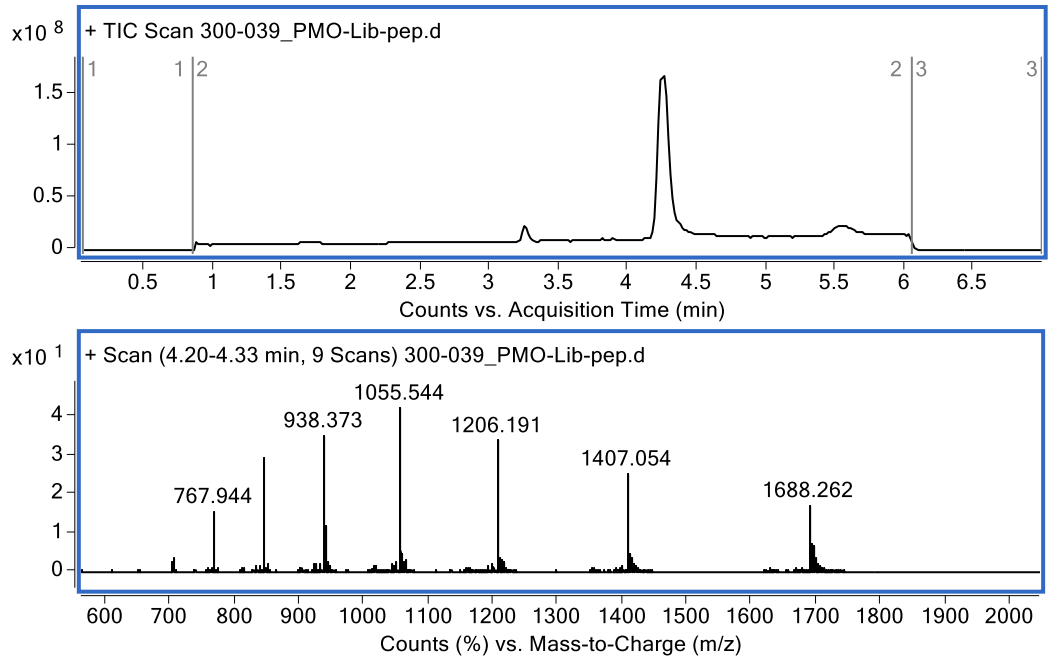


PMO-library-peptide

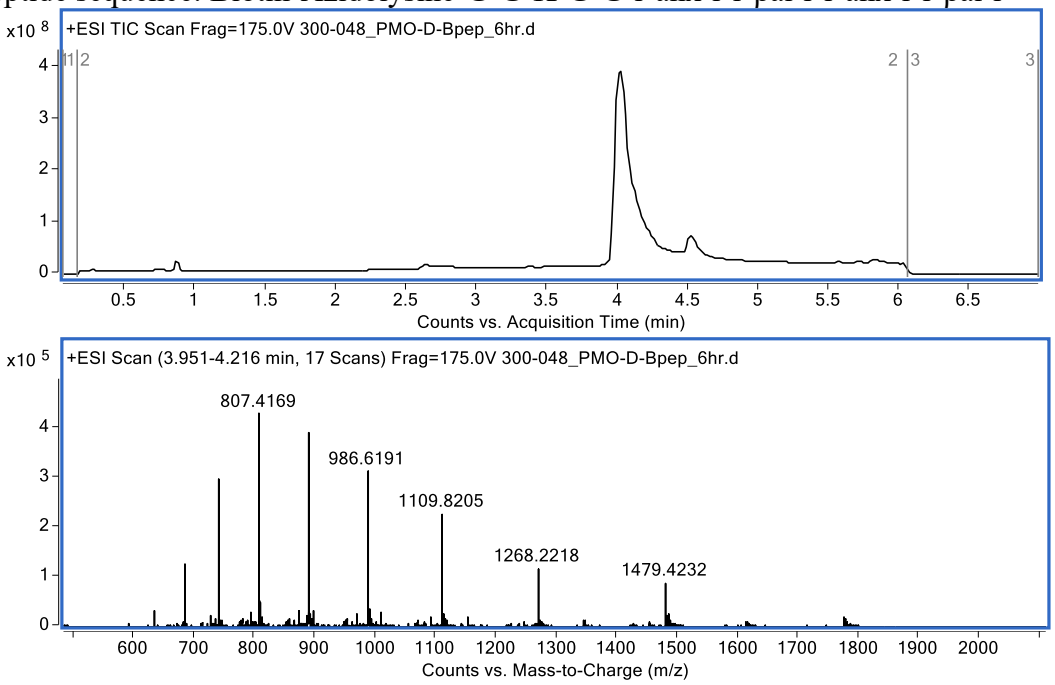
Mass Expected: 8437.1

Mass Observed: 8437.1

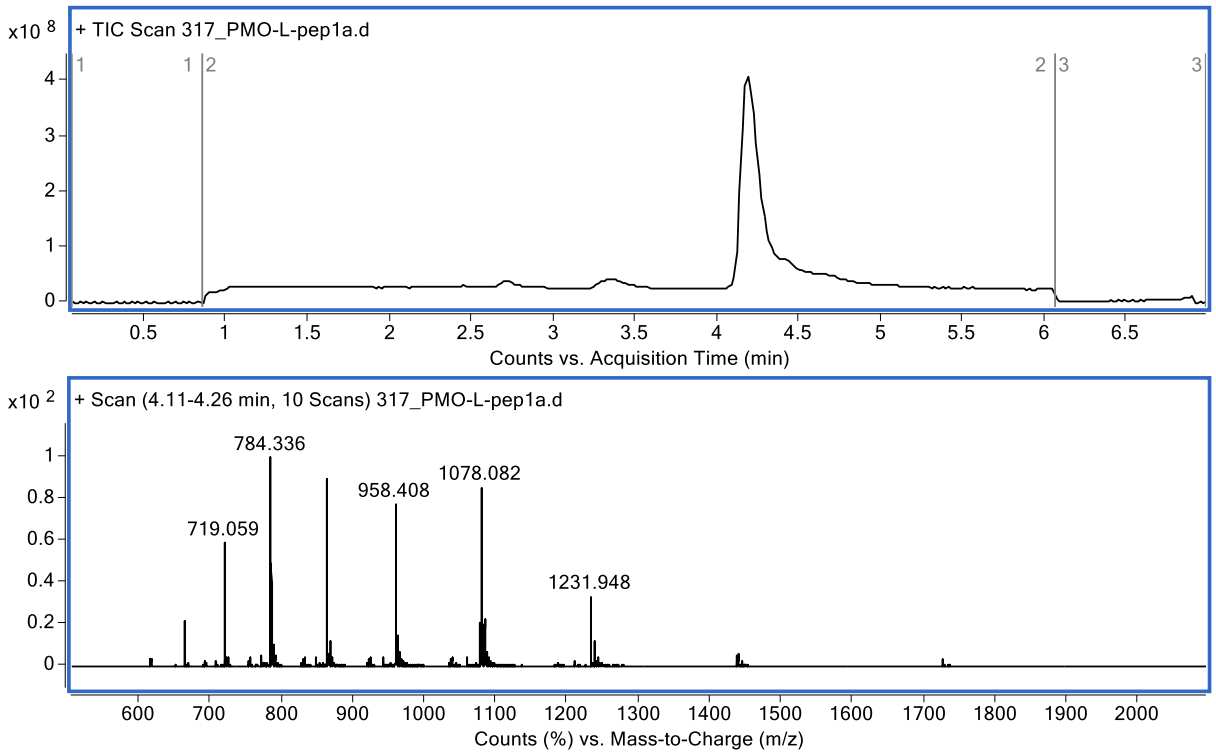
Peptide sequence: Biotin-Azidolysine-G-G-K-G-G-s-G-βal-n-d-p-βal-k-w-k-k



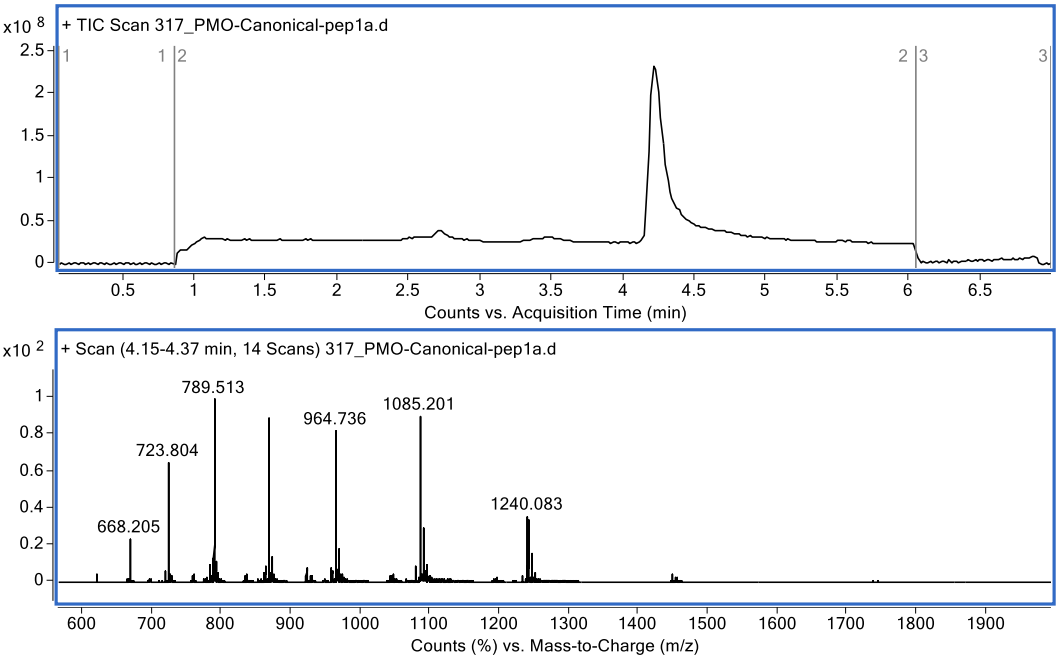
PMO-D-Bpep
Mass Expected: 8871.9
Mass Observed: 8871.8
Peptide sequence: Biotin-Azidolysine-G-G-K-G-G-r-ahx-r-r-βal-r-r-ahx-r-r-βal-r



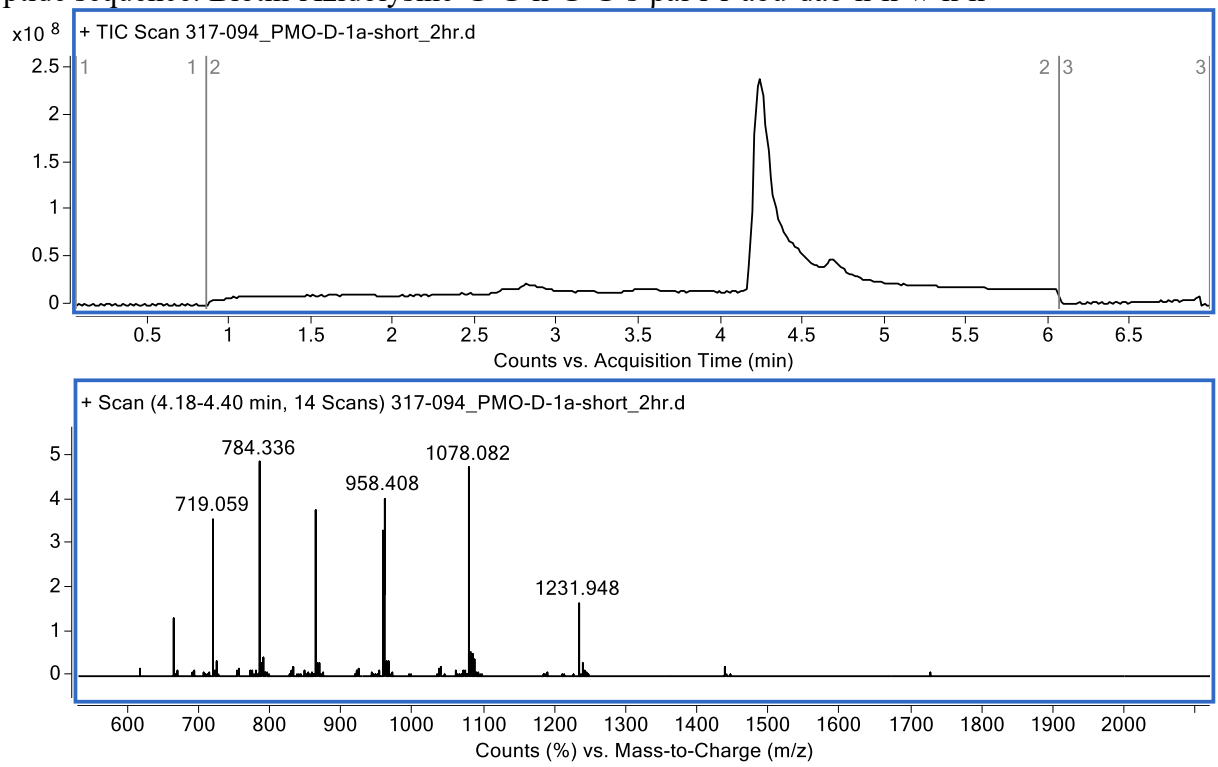
PMO-L-Cyto1a
Mass Expected: 8617.4
Mass Observed: 8618.0
Peptide sequence: Biotin-Azidolysine-G-G-K-G-G-S-βal-R-R-abu-Dab-H-K-W-K-K



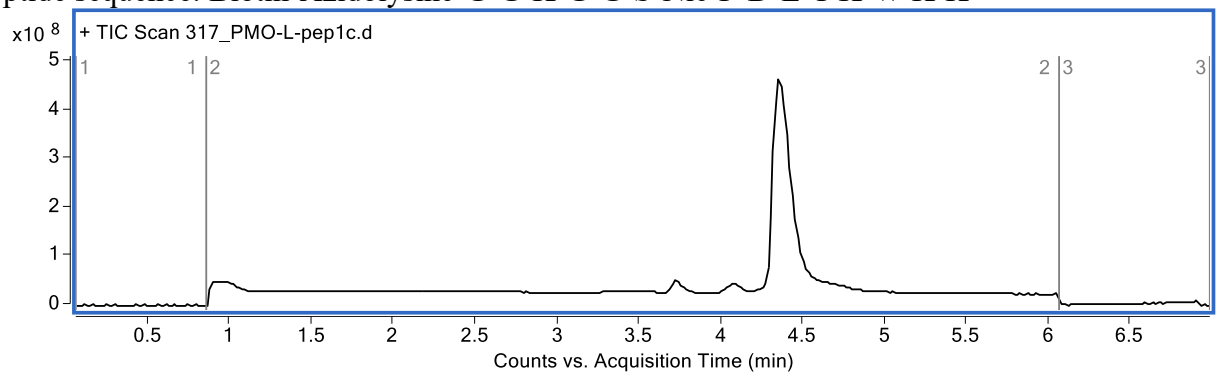
PMO-Cannonical-Cyto1a
Mass Expected: 8674.5
Mass Observed: 8674.8
Peptide sequence: Biotin-Azidolysine-G-G-K-G-G-S-A-R-R-G-G-K-H-K-W-K-K

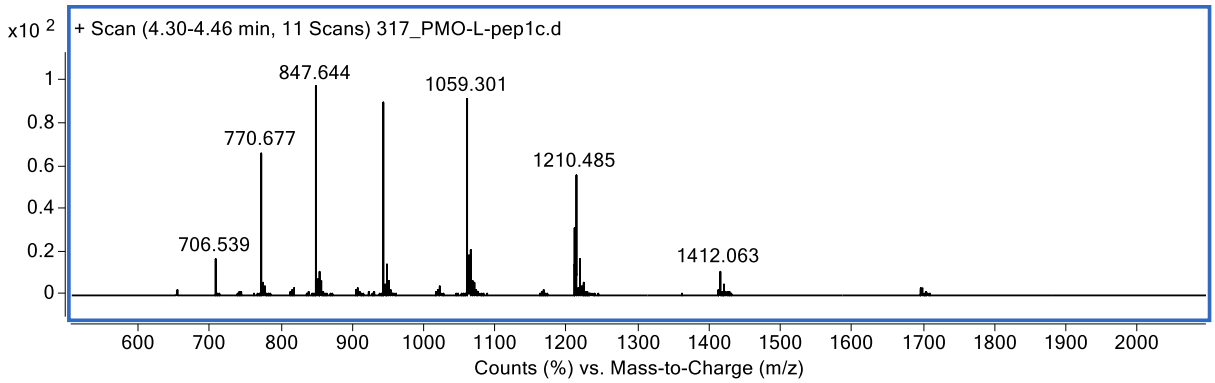


PMO-D-Cyto1a
Mass Expected: 8617.4
Mass Observed: 8618.0
Peptide sequence: Biotin-Azidolysine-G-G-k-G-G-s-βal-r-r-abu-dab-h-k-w-k-k

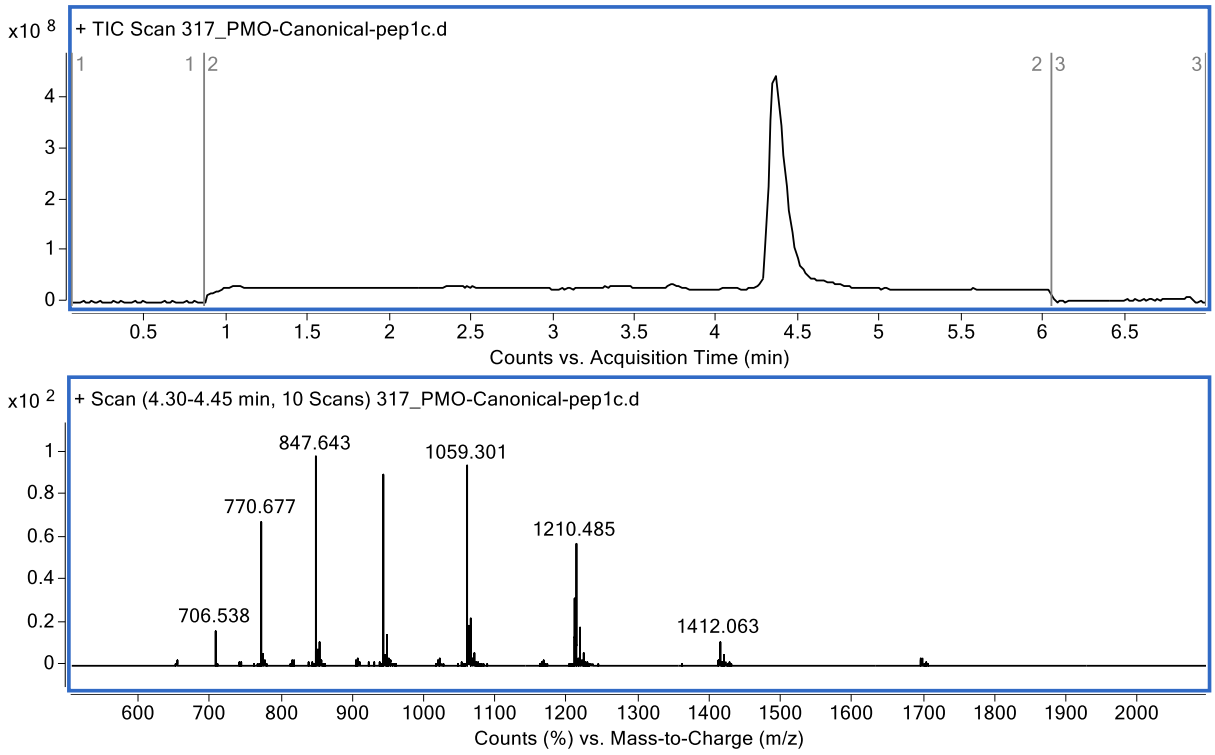


PMO-L-WC1c
Mass Expected: 8467.3
Mass Observed: 8467.5
Peptide sequence: Biotin-Azidolysine-G-G-K-G-G-S-Nle-P-D-E-T-K-W-K-K

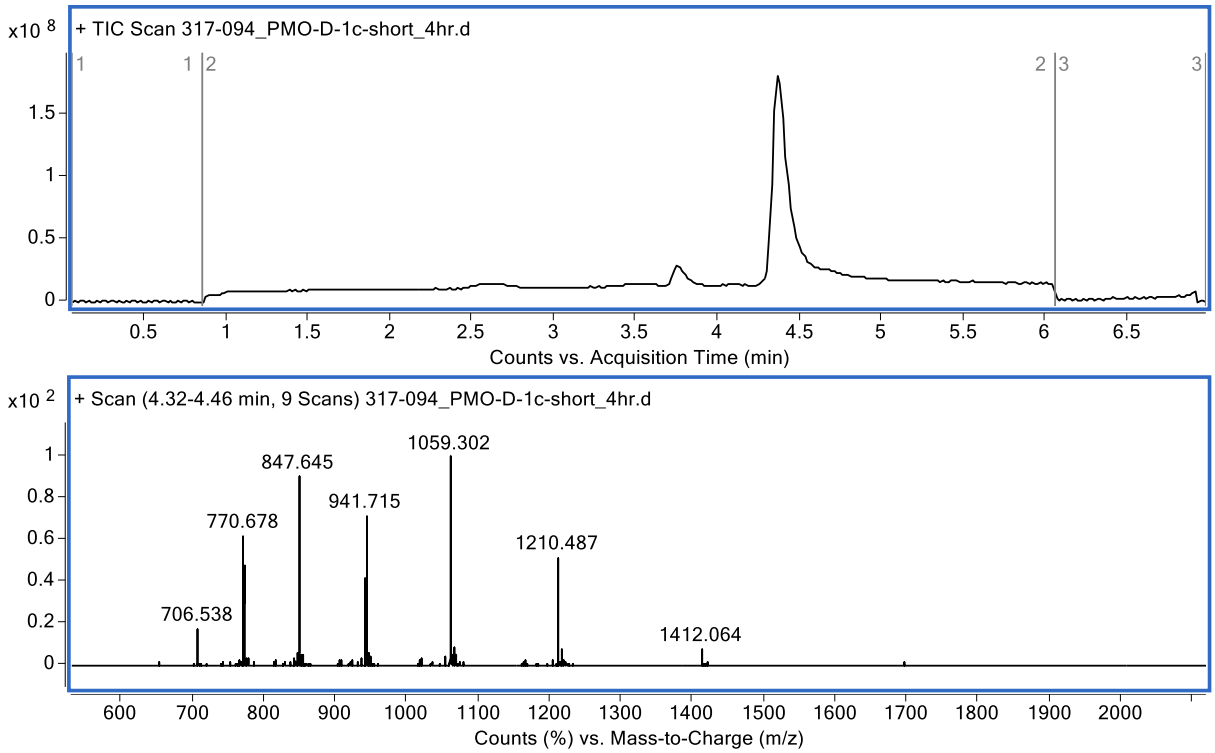




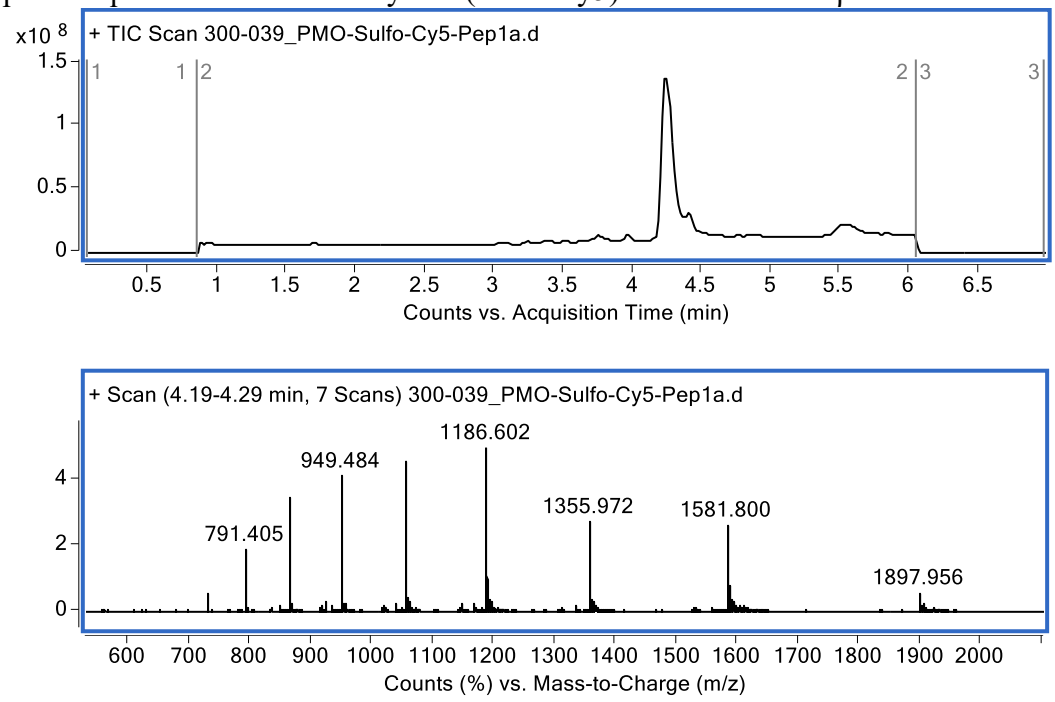
PMO-Cannonical-WC1c
Mass Expected: 8467.3
Mass Observed: 8467.4
Peptide sequence: Biotin-Azidolysine-G-G-K-G-G-S-L-P-D-E-T-K-W-K-K



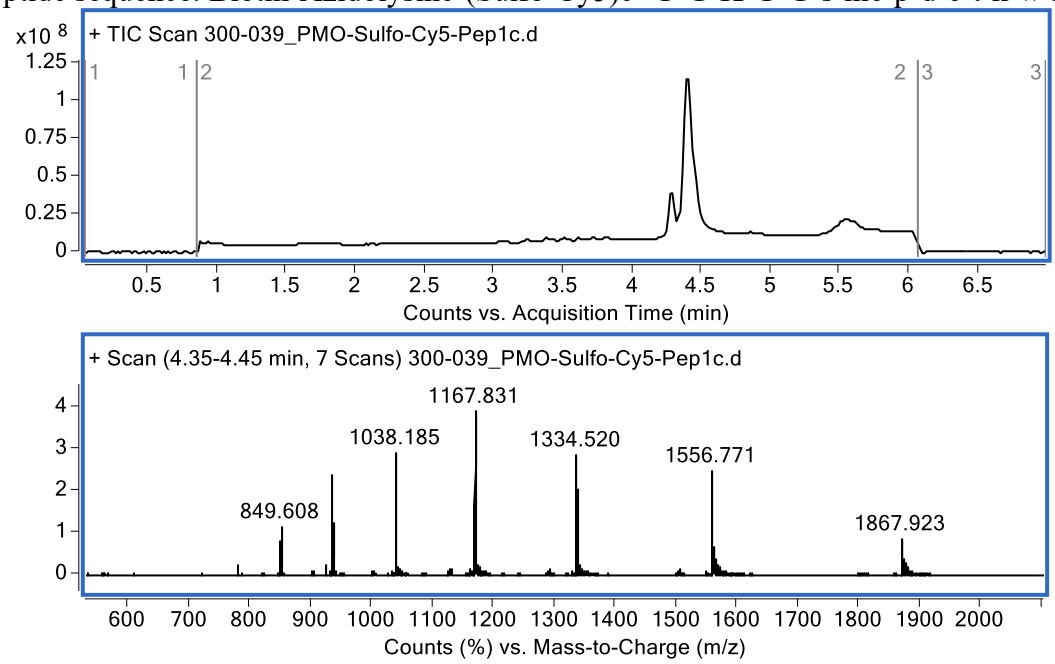
PMO-D-WC1c
Mass Expected: 8467.3
Mass Observed: 8467.4
Peptide sequence: Biotin-Azidolysine-G-G-k-G-G-s-nle-p-d-e-t-k-w-k-k



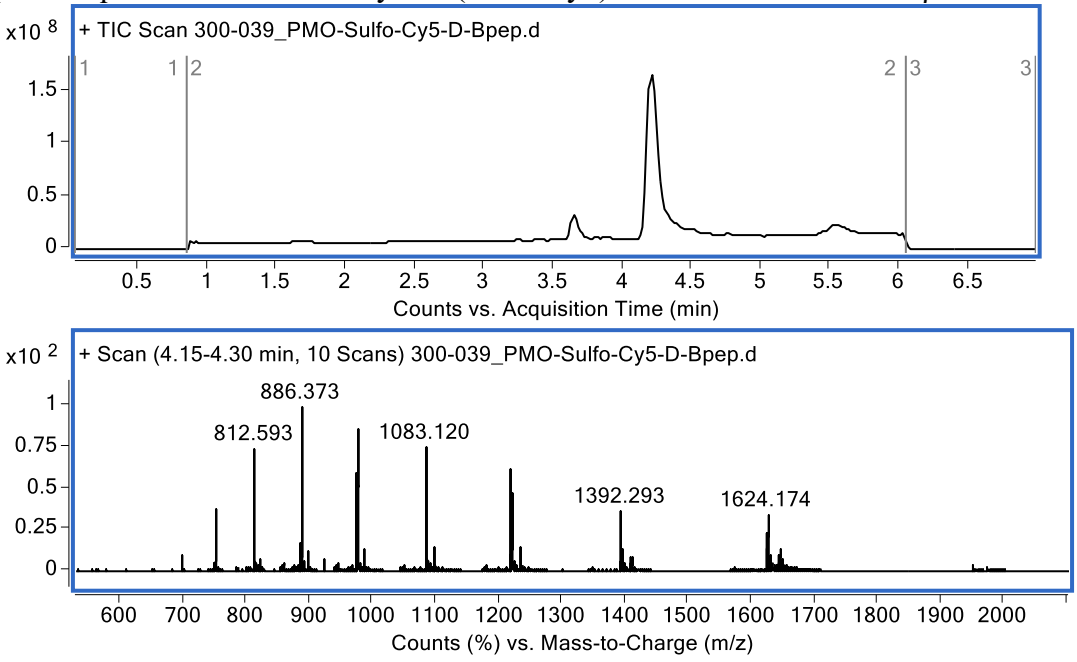
PMO-Sulfo-Cy5-Cyto1a
Mass Expected: 9485.6
Mass Observed: 9485.7
Peptide sequence: Biotin-Azidolysine-(Sulfo-Cy5)c-G-G-K-G-G-s-βal-r-r-abu-dab-h-k-w-k-k



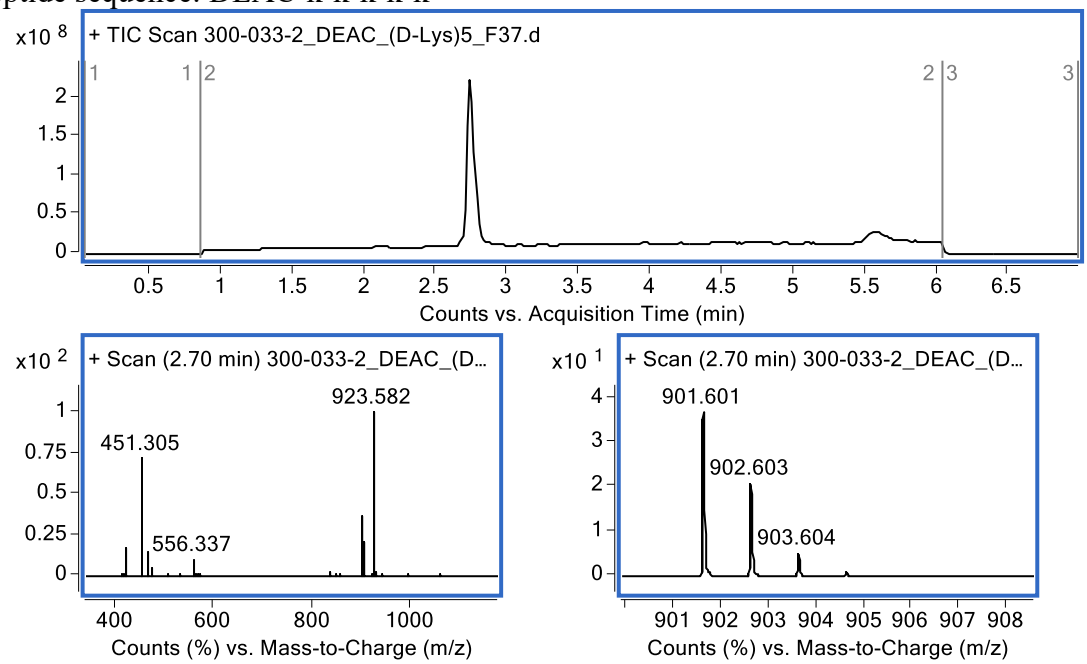
PMO-Sulfo-Cy5-WC1c
Mass Expected:9335.3
Mass Observed: 9335.5
Peptide sequence: Biotin-Azidolysine-(Sulfo-Cy5)c- G-G-K-G-G-s-nle-p-d-e-t-k-w-k-k



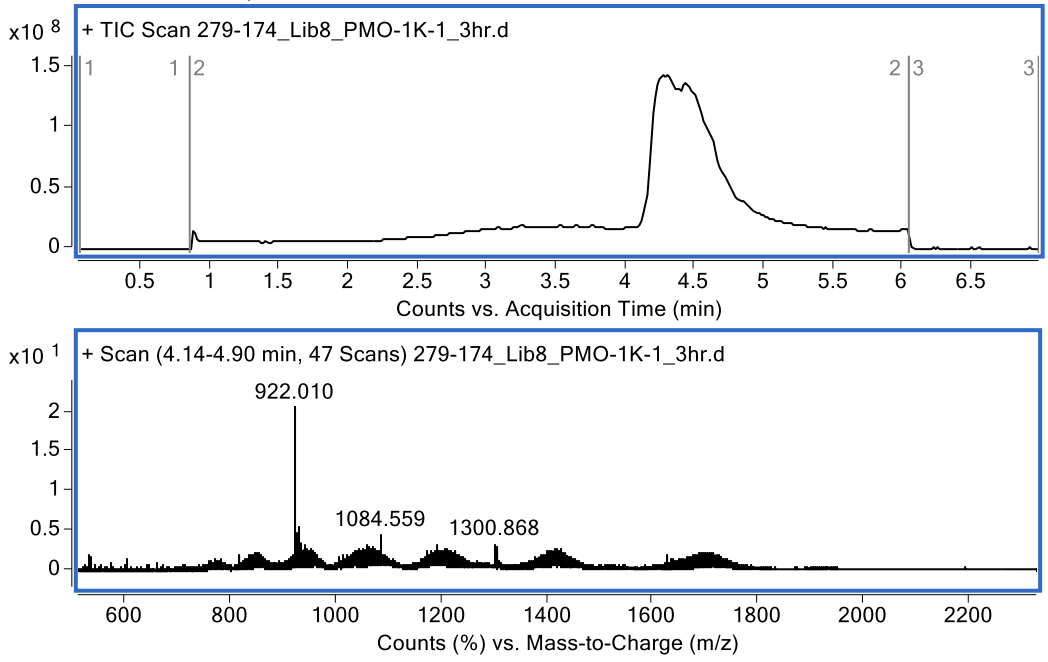
PMO-Sulfo-Cy5-D-Bpep
Mass Expected: 9739.9
Mass Observed: 9739.8
Peptide sequence: Biotin-Azidolysine-(Sulfo-Cy5)c-G-G-K-G-G-r-ahx-r-r-βal-r-r-ahx-r-r-βal-r

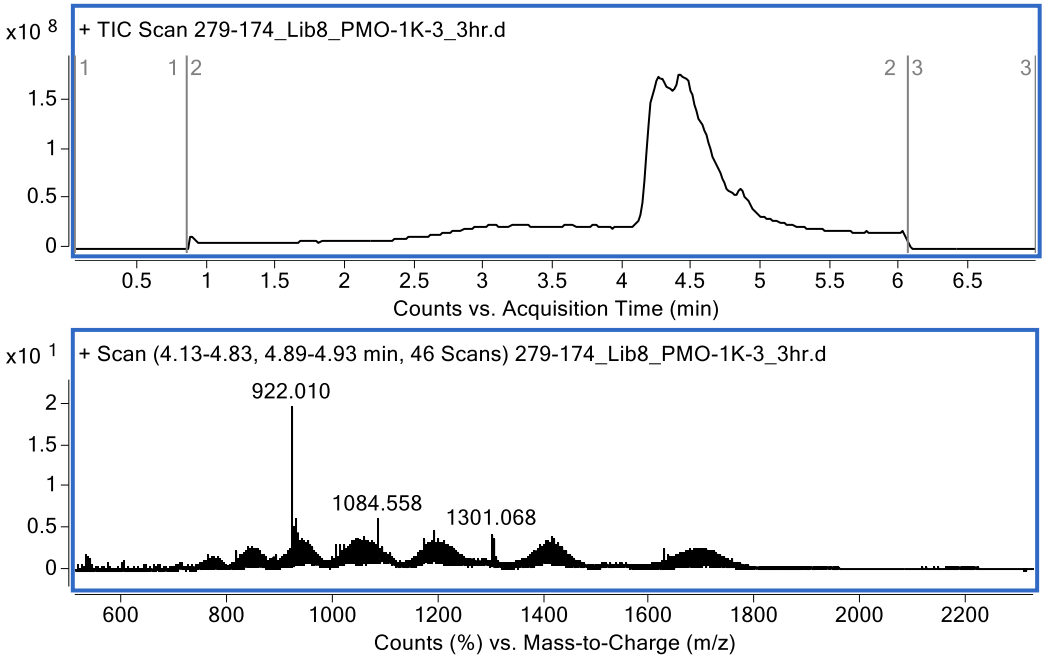


DEAC-k$_5$
Mass Expected: 900.591
Mass Observed: 900.593
Peptide sequence: DEAC-k-k-k-k-k



PMO-Library (1,000 member)
The libraries shown with greater than 1,000 members were produced by combining 1,000 member libraries. Two of the 1,000-member PMO-libraries are shown here.




x10 8
+ TIC Scan 279-174_Lib8_PMO-1K-3_3hr.d
1
1
2
2
3
3
1.5
1
0.5
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
5.5
6
6.5
Counts vs. Acquisition Time (min)

x10 1
+ Scan (4.13-4.83, 4.89-4.93 min, 46 Scans) 279-174_Lib8_PMO-1K-3_3hr.d
922.010
1084.558
1301.068
2
1.5
1
0.5
0
600
800
1000
1200
1400
1600
1800
2000
2200
Counts (%) vs. Mass-to-Charge (m/z)

2.8 References

(1) Tripathi, P. P.; Arami, H.; Banga, I.; Gupta, J.; Gandhi, S. Cell Penetrating Peptides in Preclinical and Clinical Cancer Diagnosis and Therapy. *Oncotarget* **2018**, *9* (98), 37252–37267. https://doi.org/10.18632/oncotarget.26442.

(2) Xie, J.; Bi, Y.; Zhang, H.; Dong, S.; Teng, L.; Lee, R. J.; Yang, Z. Cell-Penetrating Peptides in Diagnosis and Treatment of Human Diseases: From Preclinical Research to Clinical Application. *Frontiers in Pharmacology* **2020**, *11*, 697. https://doi.org/10.3389/fphar.2020.00697.

(3) Bechara, C.; Sagan, S. Cell-Penetrating Peptides: 20years Later, Where Do We Stand? *FEBS Letters* **2013**, *587* (12), 1693–1702. https://doi.org/10.1016/j.febslet.2013.04.031.

(4) Paulk, L. R. MSG 2021 Meeting Information and Abstracts. *RRNMF Neuromuscular Journal* **2021**, *2* (4). https://doi.org/10.17161/rrnmf.v2i4.15849.

(5) Research, C. for D. E. and. Novel Drug Approvals for 2022. *FDA* **2023**.

(6) Kim, G. C.; Cheon, D. H.; Lee, Y. Challenge to Overcome Current Limitations of Cell-Penetrating Peptides. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **2021**, *1869* (4), 140604. https://doi.org/10.1016/j.bbapap.2021.140604.

(7) Frankel, A. D.; Pabo, C. O. Cellular Uptake of the Tat Protein from Human Immunodeficiency Virus. *Cell* **1988**, *55* (6), 1189–1193. https://doi.org/10.1016/0092-8674(88)90263-2.

(8) Green, M.; Loewenstein, P. M. Autonomous Functional Domains of Chemically Synthesized Human Immunodeficiency Virus Tat Trans-Activator Protein. *Cell* **1988**, *55* (6), 1179–1188. https://doi.org/10.1016/0092-8674(88)90262-0.

(9) Jearawiriyapaisarn, N.; Moulton, H. M.; Buckley, B.; Roberts, J.; Sazani, P.; Fucharoen, S.; Iversen, P. L.; Kole, R. Sustained Dystrophin Expression Induced by Peptide-Conjugated Morpholino Oligomers in the Muscles of Mdx Mice. *Molecular Therapy* **2008**, *16* (9), 1624–1629. https://doi.org/10.1038/mt.2008.120.

(10) Wu, B.; Moulton, H. M.; Iversen, P. L.; Jiang, J.; Li, J.; Li, J.; Spurney, C. F.; Sali, A.; Guerron, A. D.; Nagaraju, K.; Doran, T.; Lu, P.; Xiao, X.; Lu, Q. L. Effective Rescue of Dystrophin Improves Cardiac Function in Dystrophin-Deficient Mice by a Modified Morpholino Oligomer. *PNAS* **2008**, *105* (39), 14814–14819. https://doi.org/10.1073/pnas.0805676105.

(11) Wimley, W. C. Synthetic Molecular Evolution of Cell Penetrating Peptides. In *Cell Penetrating Peptides: Methods and Protocols*; Langel, Ü., Ed.; Methods in Molecular Biology; Springer US: New York, NY, 2022; pp 73–89. https://doi.org/10.1007/978-1-0716-1752-6_5.

(12) Kauffman, W. B.; Guha, S.; Wimley, W. C. Synthetic Molecular Evolution of Hybrid Cell Penetrating Peptides. *Nature Communications* **2018**, *9* (1), 2568. https://doi.org/10.1038/s41467-018-04874-6.

(13) Porosk, L.; Gaidutšik, I.; Langel, Ü. Approaches for the Discovery of New Cell-Penetrating Peptides. *Expert Opin Drug Discov* **2021**, *16* (5), 553–565. https://doi.org/10.1080/17460441.2021.1851187.

(14) Lee, E. Y.; Wong, G. C. L.; Ferguson, A. Machine Learning-Enabled Discovery and Design of Membrane-Active Peptides. *Bioorganic & medicinal chemistry* **2018**, *26* (10), 2708–2718. https://doi.org/10.1016/j.bmc.2017.07.012.

(15) Manavalan, B.; Subramaniyam, S.; Shin, T. H.; Kim, M. O.; Lee, G. Machine-Learning-Based Prediction of Cell-Penetrating Peptides and Their Uptake Efficiency with Improved Accuracy. *J. Proteome Res.* **2018**, *17* (8), 2715–2726. https://doi.org/10.1021/acs.jproteome.8b00148.

(16) Pandey, P.; Patel, V.; George, N. V.; Mallajosyula, S. S. KELM-CPPpred: Kernel Extreme Learning Machine Based Prediction Model for Cell-Penetrating Peptides. *J. Proteome Res.* **2018**, *17* (9), 3214–3222. https://doi.org/10.1021/acs.jproteome.8b00322.

(17) Su, R.; Hu, J.; Zou, Q.; Manavalan, B.; Wei, L. Empirical Comparison and Analysis of Web-Based Cell-Penetrating Peptide Prediction Tools. *Brief. Bioinformatics* **2020**, *21* (2), 408–420. https://doi.org/10.1093/bib/bby124.

(18) Kumar, V.; Raghava, G. P. S. In Silico Design of Chemically Modified Cell-Penetrating Peptides. *Methods Mol Biol* **2022**, *2383*, 63–71. https://doi.org/10.1007/978-1-0716-1752-6_4.

(19) Schissel, C. K.; Mohapatra, S.; Wolfe, J. M.; Fadzen, C. M.; Bellovoda, K.; Wu, C.-L.; Wood, J. A.; Malmberg, A. B.; Loas, A.; Gómez-Bombarelli, R.; Pentelute, B. L. Deep Learning to Design Nuclear-Targeting Abiotic Miniproteins. *Nat. Chem.* **2021**, *13* (10), 992–1000. https://doi.org/10.1038/s41557-021-00766-3.

(20) López-Vidal, E. M.; Schissel, C. K.; Mohapatra, S.; Bellovoda, K.; Wu, C.-L.; Wood, J. A.; Malmberg, A. B.; Loas, A.; Gómez-Bombarelli, R.; Pentelute, B. L. Deep Learning Enables Discovery of a Short Nuclear Targeting Peptide for Efficient Delivery of Antisense Oligomers. *JACS Au* **2021**, *1* (11), 2009-2020. https://doi.org/10.1021/jacsau.1c00327.

(21) Wolfe, J. M.; Fadzen, C. M.; Choo, Z.-N.; Holden, R. L.; Yao, M.; Hanson, G. J.; Pentelute, B. L. Machine Learning To Predict Cell-Penetrating Peptides for Antisense Delivery. *ACS Cent Sci* **2018**, *4* (4), 512–520. https://doi.org/10.1021/acscentsci.8b00098.

(22) Gao, S.; Simon, M. J.; Hue, C. D.; Morrison, B.; Banta, S. An Unusual Cell Penetrating Peptide Identified Using a Plasmid Display-Based Functional Selection Platform. *ACS Chem. Biol.* **2011**, *6* (5), 484–491. https://doi.org/10.1021/cb100423u.

(23) Kamide, K.; Nakakubo, H.; Uno, S.; Fukamizu, A. Isolation of Novel Cell-Penetrating Peptides from a Random Peptide Library Using in Vitro Virus and Their Modifications. *International Journal of Molecular Medicine* **2010**, *25* (1), 41–51. https://doi.org/10.3892/ijmm_00000311.

(24) Bowen, J.; Schloop, A. E.; Reeves, G. T.; Menegatti, S.; Rao, B. M. Discovery of Membrane-Permeating Cyclic Peptides via MRNA Display. *Bioconjugate Chem.* **2020**, *31* (10), 2325–2338. https://doi.org/10.1021/acs.bioconjchem.0c00413.

(25) Jirka, S. M. G.; 't Hoen, P. A. C.; Diaz Parillas, V.; Tanganyika-de Winter, C. L.; Verheul, R. C.; Aguilera, B.; de Visser, P. C.; Aartsma-Rus, A. M. Cyclic Peptides to Improve Delivery and Exon Skipping of Antisense Oligonucleotides in a Mouse Model for Duchenne Muscular Dystrophy. *Mol Ther* **2018**, *26* (1), 132–147. https://doi.org/10.1016/j.ymthe.2017.10.004.

(26) Hoffmann, K.; Milech, N.; Juraja, S. M.; Cunningham, P. T.; Stone, S. R.; Francis, R. W.; Anastasas, M.; Hall, C. M.; Heinrich, T.; Bogdawa, H. M.; Winslow, S.; Scobie, M. N.; Dewhurst, R. E.; Florez, L.; Ong, F.; Kerfoot, M.; Champain, D.; Adams, A. M.; Fletcher, S.; Viola, H. M.; Hool, L. C.; Connor, T.; Longville, B. A. C.; Tan, Y.-F.; Kroeger, K.; Morath, V.; Weiss, G. A.; Skerra, A.; Hopkins, R. M.; Watt, P. M. A Platform for Discovery of Functional Cell-Penetrating Peptides for Efficient Multi-Cargo Intracellular Delivery. *Scientific Reports* **2018**, *8* (1), 12538. https://doi.org/10.1038/s41598-018-30790-2.

(27) Birch, D.; Christensen, M. V.; Staerk, D.; Franzyk, H.; Nielsen, H. M. Stereochemistry as a Determining Factor for the Effect of a Cell-Penetrating Peptide on Cellular Viability and Epithelial Integrity. *Biochem J* **2018**, *475* (10), 1773–1788. https://doi.org/10.1042/BCJ20180155.

(28) Verdurmen, W. P. R.; Bovee-Geurts, P. H.; Wadhwani, P.; Ulrich, A. S.; Hällbrink, M.; van Kuppevelt, T. H.; Brock, R. Preferential Uptake of L- versus D-Amino Acid Cell-Penetrating Peptides in a Cell Type-Dependent Manner. *Chemistry & Biology* **2011**, *18* (8), 1000–1010. https://doi.org/10.1016/j.chembiol.2011.06.006.

(29) Najjar, K.; Erazo-Oliveras, A.; Brock, D. J.; Wang, T.-Y.; Pellois, J.-P. An L- to d-Amino Acid Conversion in an Endosomolytic Analog of the Cell-Penetrating Peptide TAT Influences Proteolytic Stability, Endocytic Uptake, and Endosomal Escape. *J Biol Chem* **2017**, *292* (3), 847–861. https://doi.org/10.1074/jbc.M116.759837.

(30) Schissel, C. K.; Farquhar, C. E.; Malmberg, A. B.; Loas, A.; Pentelute, B. L. *Cell-Penetrating D-Peptides Retain Antisense Morpholino Oligomer Delivery Activity*; 2021; p 2021.09.30.462617. https://doi.org/10.1101/2021.09.30.462617.

(31) Kremsmayr, T.; Aljnabi, A.; Blanco-Canosa, J. B.; Tran, H. N. T.; Emidio, N. B.; Muttenthaler, M. On the Utility of Chemical Strategies to Improve Peptide Gut Stability. *J Med Chem* **2022**, *65* (8), 6191–6206. https://doi.org/10.1021/acs.jmedchem.2c00094.

(32) Carney, R. P.; Thillier, Y.; Kiss, Z.; Sahabi, A.; Heleno Campos, J. C.; Knudson, A.; Liu, R.; Olivos, D.; Saunders, M.; Tian, L.; Lam, K. S. Combinatorial Library Screening with Liposomes for Discovery of Membrane Active Peptides. *ACS Comb. Sci.* **2017**, *19* (5), 299–307. https://doi.org/10.1021/acscombsci.6b00182.

(33) Quartararo, A. J.; Gates, Z. P.; Somsen, B. A.; Hartrampf, N.; Ye, X.; Shimada, A.; Kajihara, Y.; Ottmann, C.; Pentelute, B. L. Ultra-Large Chemical Libraries for the Discovery of High-Affinity Peptide Binders. *Nat Commun* **2020**, *11* (1), 3183. https://doi.org/10.1038/s41467-020-16920-3.

(34) Zuckermann, R. N.; Kerr, J. M.; Siani, M. A.; Banville, S. C.; Santi, D. V. Identification of Highest-Affinity Ligands by Affinity Selection from Equimolar Peptide Mixtures Generated by Robotic Synthesis. *Proc Natl Acad Sci U S A* **1992**, *89* (10), 4505–4509. https://doi.org/10.1073/pnas.89.10.4505.

(35) Beck, S.; Jin, X.; Yin, J.; Kim, S.-H.; Lee, N.-K.; Oh, S.-Y.; Jin, X.; Kim, M.-K.; Kim, E.-B.; Son, J.-S.; Kim, S.-C.; Nam, D.-H.; Kim, S.-H.; Kang, S.-K.; Kim, H.; Choi, Y.-J. Identification of a Peptide That Interacts with Nestin Protein Expressed in Brain Cancer Stem Cells. *Biomaterials* **2011**, *32* (33), 8518–8528. https://doi.org/10.1016/j.biomaterials.2011.07.048.

(36) Wu, C.-H.; Kuo, Y.-H.; Hong, R.-L.; Wu, H.-C. α-Enolase–Binding Peptide Enhances Drug Delivery Efficiency and Therapeutic Efficacy against Colorectal Cancer. *Science Translational Medicine* **2015**, *7* (290), 290ra91-290ra91. https://doi.org/10.1126/scitranslmed.aaa9391.

(37) Wu, C.-H.; Liu, I.-J.; Lu, R.-M.; Wu, H.-C. Advancement and Applications of Peptide Phage Display Technology in Biomedical Science. *Journal of Biomedical Science* **2016**, *23* (1), 8. https://doi.org/10.1186/s12929-016-0223-x.

(38) Ren, Y.; Zhan, C.; Gao, J.; Zhang, M.; Wei, X.; Ying, M.; Liu, Z.; Lu, W. A D-Peptide Ligand of Integrins for Simultaneously Targeting Angiogenic Blood Vasculature and Glioma Cells. *Mol Pharm* **2018**, *15* (2), 592–601. https://doi.org/10.1021/acs.molpharmaceut.7b00944.

(39) Wei, X.; Zhan, C.; Shen, Q.; Fu, W.; Xie, C.; Gao, J.; Peng, C.; Zheng, P.; Lu, W. A D-Peptide Ligand of Nicotine Acetylcholine Receptors for Brain-Targeted Drug Delivery. *Angew Chem Int Ed Engl* **2015**, *54* (10), 3023–3027. https://doi.org/10.1002/anie.201411226.

(40) Huang, L.; Xie, J.; Bi, Q.; Li, Z.; Liu, S.; Shen, Q.; Li, C. Highly Selective Targeting of Hepatic Stellate Cells for Liver Fibrosis Treatment Using a D-Enantiomeric Peptide Ligand of Fn14 Identified by Mirror-Image MRNA Display. *Mol Pharm* **2017**, *14* (5), 1742–1753. https://doi.org/10.1021/acs.molpharmaceut.6b01174.

(41) Eckert, D. M.; Malashkevich, V. N.; Hong, L. H.; Carr, P. A.; Kim, P. S. Inhibiting HIV-1 Entry: Discovery of D-Peptide Inhibitors That Target the Gp41 Coiled-Coil Pocket. *Cell* **1999**, *99* (1), 103–115. https://doi.org/10.1016/s0092-8674(00)80066-5.

(42) Loftis, A. R.; Zhang, G.; Backlund, C.; Quartararo, A. J.; Pishesha, N.; Hanna, C. C.; Schissel, C. K.; Garafola, D.; Loas, A.; Collier, R. J.; Ploegh, H.; Irvine, D. J.; Pentelute, B. L. An in Vivo Selection-Derived d-Peptide for Engineering Erythrocyte-Binding Antigens That Promote Immune Tolerance. *PNAS* **2021**, *118* (34), e2101596118. https://doi.org/10.1073/pnas.2101596118.

(43) Petersen, L. K.; Christensen, A. B.; Andersen, J.; Folkesson, C. G.; Kristensen, O.; Andersen, C.; Alzu, A.; Sløk, F. A.; Blakskjær, P.; Madsen, D.; Azevedo, C.; Micco, I.; Hansen, N. J. V. Screening of DNA-Encoded Small Molecule Libraries inside a Living Cell. *J. Am. Chem. Soc.* **2021**, *143* (7), 2751–2756. https://doi.org/10.1021/jacs.0c09213.

(44) Fischer, P. m.; Zhelev, N. z.; Wang, S.; Melville, J. e.; Fåhraeus, R.; Lane, D. p. Structure–Activity Relationship of Truncated and Substituted Analogues of the Intracellular Delivery Vector Penetratin. *The Journal of Peptide Research* **2000**, *55* (2), 163–172. https://doi.org/10.1034/j.1399-3011.2000.00163.x.

(45) Vinogradov, A. A.; Gates, Z. P.; Zhang, C.; Quartararo, A. J.; Halloran, K. H.; Pentelute, B. L. Library Design-Facilitated High-Throughput Sequencing of Synthetic Peptide Libraries. *ACS Comb. Sci.* **2017**, *19* (11), 694–701. https://doi.org/10.1021/acscombsci.7b00109.

(46) Schmidt, S.; Adjobo-Hermans, M. J. W.; Kohze, R.; Enderle, T.; Brock, R.; Milletti, F. Identification of Short Hydrophobic Cell-Penetrating Peptides for Cytosolic Peptide Delivery by Rational Design. *Bioconjugate Chem.* **2017**, *28* (2), 382–389. https://doi.org/10.1021/acs.bioconjchem.6b00535.

(47) Som, A.; Reuter, A.; Tew, G. N. Protein Transduction Domain Mimics: The Role of Aromatic Functionality. *Angewandte Chemie International Edition* **2012**, *51* (4), 980–983. https://doi.org/10.1002/anie.201104624.

(48) Lönn, P.; Dowdy, S. F. Cationic PTD/CPP-Mediated Macromolecular Delivery: Charging into the Cell. *Expert Opin Drug Deliv* **2015**, *12* (10), 1627–1636. https://doi.org/10.1517/17425247.2015.1046431.

(49) Najjar, K.; Erazo-Oliveras, A.; Mosior, J. W.; Whitlock, M. J.; Rostane, I.; Cinclair, J. M.; Pellois, J.-P. Unlocking Endosomal Entrapment with Supercharged Arginine-Rich Peptides. *Bioconjugate Chem.* **2017**, *28* (12), 2932–2941. https://doi.org/10.1021/acs.bioconjchem.7b00560.

(50) Pomplun, S.; Shugrue, C. R.; Schmitt, A. M.; Schissel, C. K.; Farquhar, C. E.; Pentelute, B. L. Secondary Amino Alcohols: Traceless Cleavable Linkers for Use in Affinity Capture and Release. *Angewandte Chemie International Edition* **2020**, *59* (28), 11566–11572. https://doi.org/10.1002/anie.202003478.

(51) Behrendt, R.; White, P.; Offer, J. Advances in Fmoc Solid-Phase Peptide Synthesis. *Journal of Peptide Science* **2016**, *22* (1), 4–27. https://doi.org/10.1002/psc.2836.
(52) Fadzen, C. M.; Holden, R. L.; Wolfe, J. M.; Choo, Z.-N.; Schissel, C. K.; Yao, M.; Hanson, G. J.; Pentelute, B. L. Chimeras of Cell-Penetrating Peptides Demonstrate Synergistic Improvement in Antisense Efficacy. *Biochemistry* **2019**, *58* (38), 3980-3989. https://doi.org/10.1021/acs.biochem.9b00413.
(53) Burlina, F.; Sagan, S.; Bolbach, G.; Chassaing, G. A Direct Approach to Quantification of the Cellular Uptake of Cell-Penetrating Peptides Using MALDI-TOF Mass Spectrometry. *Nature protocols* **2006**, *1* (1), 200.
(54) Rabideau, A. E.; Pentelute, B. L. Delivery of Non-Native Cargo into Mammalian Cells Using Anthrax Lethal Toxin. *ACS Chem. Biol.* **2016**, *11* (6), 1490–1501. https://doi.org/10.1021/acschembio.6b00169.
(55) Ogawa, Y.; Imamoto, N. Methods to Separate Nuclear Soluble Fractions Reflecting Localizations in Living Cells. *iScience* **2021**, *24* (12), 103503. https://doi.org/10.1016/j.isci.2021.103503.
(56) Simon, M. D.; Heider, P. L.; Adamo, A.; Vinogradov, A. A.; Mong, S. K.; Li, X.; Berger, T.; Policarpo, R. L.; Zhang, C.; Zou, Y.; Liao, X.; Spokoyny, A. M.; Jensen, K. F.; Pentelute, B. L. Rapid Flow-Based Peptide Synthesis. *ChemBioChem* **2014**, *15* (5), 713–720. https://doi.org/10.1002/cbic.201300796.
(57) Verdurmen, W. P. R.; Thanos, M.; Ruttekolk, I. R.; Gulbins, E.; Brock, R. Cationic Cell-Penetrating Peptides Induce Ceramide Formation via Acid Sphingomyelinase: Implications for Uptake. *Journal of Controlled Release* **2010**, *147* (2), 171–179. https://doi.org/10.1016/j.jconrel.2010.06.030.
(58) Brock, D. J.; Kustigian, L.; Jiang, M.; Graham, K.; Wang, T.-Y.; Erazo-Oliveras, A.; Najjar, K.; Zhang, J.; Rye, H.; Pellois, J.-P. Efficient Cell Delivery Mediated by Lipid-Specific Endosomal Escape of Supercharged Branched Peptides. *Traffic* **2018**, *19* (6), 421–435. https://doi.org/10.1111/tra.12566.
(59) Hartrampf, N.; Saebi, A.; Poskus, M.; Gates, Z. P.; Callahan, A. J.; Cowfer, A. E.; Hanna, S.; Antilla, S.; Schissel, C. K.; Quartararo, A. J.; Ye, X.; Mijalis, A. J.; Simon, M. D.; Loas, A.; Liu, S.; Jessen, C.; Nielsen, T. E.; Pentelute, B. L. Synthesis of Proteins by Automated Flow Chemistry. *Science* **2020**, *368* (6494), 980–987. https://doi.org/10.1126/science.abb2491.
(60) Kjeken, R.; Mousavi, S. A.; Brech, A.; Griffiths, G.; Berg, T. Wortmannin-Sensitive Trafficking Steps in the Endocytic Pathway in Rat Liver Endothelial Cells. *Biochem J* **2001**, *357* (Pt 2), 497–503.
(61) Dutta, D.; Donaldson, J. G. Search for Inhibitors of Endocytosis: Intended Specificity and Unintended Consequences. *Cellular Logistics* **2012**, *2* (4), 203–208. https://doi.org/10.4161/cl.23967.

Chapter 3. Chromatographic Enrichment of a Peptide Library Produces Novel, Non-Canonical Cell-Penetrating Peptides for Antisense Oligonucleotide Delivery

The work presented in this chapter has been reproduced from the following manuscript:

Farquhar, C. E.; Schissel, C.K.; Loas, A.; Malmberg, A.B.; Pentelute, B.L. Chromatographic Enrichment of a Peptide Library Produces Novel, Non-Canonical Cell-Penetrating Peptides for Antisense Oligonucleotide Delivery. *(Manuscript in preparation)*

3.1 Introduction

Duchene muscular dystrophy (DMD) is the most prevalent genetic neuromuscular disorder, affecting more than 1 in 5,000 newborn boys.[1] Mutations in the X-linked dystrophin gene lead to a loss in dystrophin expression, causing severe muscle wasting and eventual respiratory and cardiac failure.[2] Exon-skipping antisense oligonucleotide drugs can restore dystrophin function by binding and covering a splice site on the pre-mRNA transcript, causing the spliceosome to remove the mutation-containing exon and produce an in-frame truncated but functional copy of dystrophin.[3] The clinically approved exon-skipping drugs for DMD are nonnatural phosphorodiamidate morpholino oligonucleotides (PMO), which hybridize to pre-mRNA for splice-correction. Despite PMO efficacy in the cell, animal and human models consistently demonstrate that PMO drugs suffer from poor cell permeability and therefore must be given at frequent, large doses of at least 50mg/kg to reach full distribution to muscle tissues.[4] This high dose requirement driven by poor cell penetration prevents more widespread use of PMO drugs.

Cell-penetrating peptide (CPP) conjugates can greatly increase the delivery of antisense oligonucleotides,[5,6] and conjugation of a PMO drug to an optimized CPP may improve cell-penetration. More than one thousand CPP sequences have been discovered and applied to the effective delivery of cargoes including fluorophores, proteins, and anti-sense oligonucleotides.[7] Discovery efforts are still underway, however, because CPP efficacy is cargo-dependent[8] and known CPPs are not universally effective at delivering a PMO cargo. As a result, there remains an unmet need for novel CPPs that can effectively deliver a PMO cargo into cells.

To meet this need, we have developed an assay platform for the discovery of novel CPPs from a large library of diverse peptide sequences. It has been previously demonstrated that a library of different peptide sequences conjugated to PMO can deliver PMO to the cell nucleus.[9] In this previous case, individual penetrant sequences were isolated from the cell and sequenced, generating the novel CPP Cyto1a. This previous study met difficulty, however, in the size of libraries able to be tested. Because peptides had to be extracted from the cell prior to sequencing, the assay required treatment with a high concentration of each individual peptide to ensure enough material after cell lysis. However, the upper limit of total peptide concentration, above which cells may exhibit non-specific endocytosis,[10] meant that the overall size of the library had to remain

small (<1,000 peptides). The small library size of this previous experimental protocol may have limited the discovery of novel, diverse sequences.

Instead of directly treating cells with a large library, we envisioned separating the library into smaller pools for testing. Studies on existing CPPs have demonstrated that positive charge, especially from arginine residues, improve CPP function by enabling interactions with negatively charged phosphate heads and sugars in the cell membrane.[10–12] Because charge has an influence on penetration, we used this as the basis of separation to partition a large library into smaller library pools for testing. This fractionation of the library provides two major benefits: 1) resulting peptide pools can be sequenced prior to cell treatment, resulting in a much lower amount of each peptide needed for treatment and 2) partitioning a library into smaller segments by charge allows us to investigate whether the importance of charge in cell penetration applies to PMO-delivery, since CPP efficacy is cargo-dependent.[8]

In order to most effectively separate a peptide library based on charge, a chromatographic resolution strategy was envisioned (**Fig. 3.1**). Using this platform, a diverse library, containing both canonical and non-canonical amino acids, could be chromatographically resolved by charge via high performance liquid chromatography with a strong cation exchange column. Each smaller pool of peptides would then be conjugated to PMO via strain-promoted azide-alkyne cycloaddition (SPAAC), and then subsequently assayed for PMO cargo delivery. Library fractions that show high PMO delivery into cells contain novel and effective cell-penetrating peptides, which are analyzed by de novo MS/MS sequencing to identify potential CPPs. This smaller number of sequences from the library can then be individually synthesized and tested for PMO delivery to validate novel CPPs. The application of this platform in HeLa cells with a PMO cargo resulted in four novel, non-canonical and non-toxic **CXP peptides** for PMO delivery. Linearly combining these peptides and performing subsequent sequence-activity relationship studies produced **CXD1B**, a peptide with nanomolar in-vitro efficacy for nuclear PMO delivery.

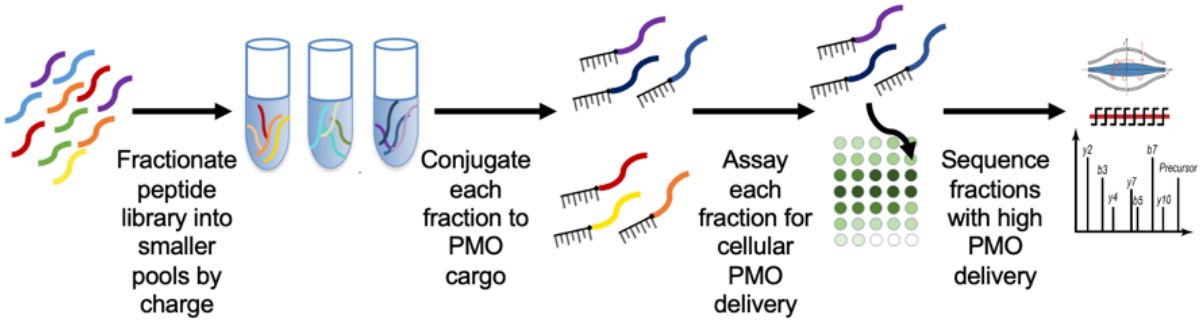


Figure 3.1 Outline of the chromatographic resolution platform for discovering novel peptides for PMO delivery

3.2 Results and Discussion

3.2.1 Library preparation

We first designed and synthesized a diverse peptide library, following the construct design shown in **Fig. 3.2**. A conserved C-terminal KWKK motif, derived from the CPP penetratin, was incorporated into the library design to improve bulk properties of the peptides. The three positive charges improve binding to a cation exchange column, and the tryptophan serves to improve membrane penetration of cationic peptides and allows the library to be quantified through UV-VIS spectroscopy at 280 nm.[13-15] The N-terminal 5-azidopentanoic acid allows the library to be coupled with PMO-DBCO through a strain-promoted azide-alkyne cycloaddition. The interior of the library would be composed of ten variable positions consisting of 16 natural amino acids, as well as 6 additional noncanonical amino acids (**Fig. 3.2**), which serve to expand the chemical space of the constructs.

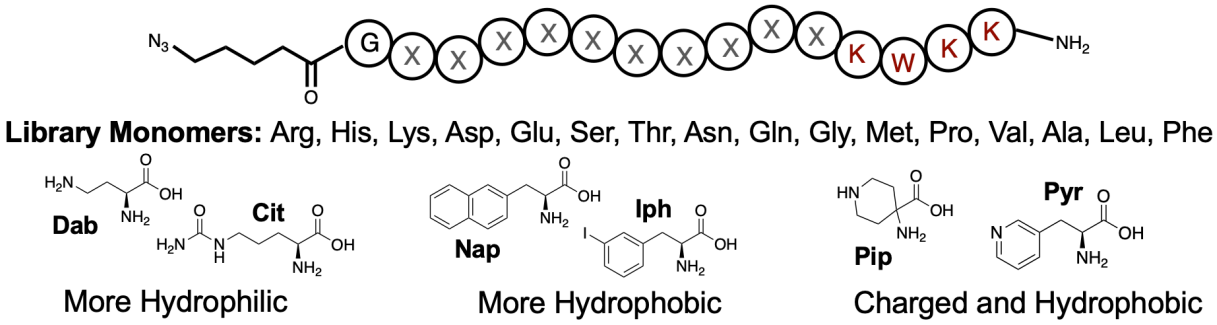


Figure 3.2 Library design and monomer set of the CPP library.
The sequence in red is derived from the existing CPP penetratin, and the remaining library (represented by **X**) consists of a variable region containing any of the 22 monomers. The Glycine at the N-terminus separates the N-terminal azide from the variable region. The library is amidated at the C-terminus. Unnatural monomers include **Dab** (L-2,4-diaminobutyric acid), **Cit** (L-citrulline), **Nap** (3-(2-naphthyl)-L-alanine), **Iph** (4-iodo-L-phenylalaine), **Pip** (4-Aminopiperidine-4-carboxylic acid), and **Pyr** (3-(4'-pyridyl)-L-alanine).

3.2.2 Chromatographic resolution and initial uptake assay

A library consisting of 3,000 individual peptide sequences was accessed through a split-pool technique via solid-phase peptide synthesis.[16] This 3,000-member library was then separated by cation-exchange chromatography with an ammonium acetate salt gradient, and peptides were collected in small fractions as they eluted off the column. The chromatogram of the separation is

shown in **Fig. 3.3** in green. These smaller fractions were combined into 10 pools, each containing about 300 individual peptide sequences based on UV quantification of the tryptophan residue. These larger pools, numbered 01-10 based on their cation exchange elution time, were then conjugated to the PMO cargo via SPAAC, and the solvent and ammonium acetate salts were removed by lyophilization.

The ten library pools, each containing approximately 300 different PMO-peptide sequences, were tested for PMO delivery in a cellular assay with HeLa 654 cells. The HeLa 654s cell line has been stably transfected with an eGFP coding sequence that is interrupted by the IVS2-654 β-globin intron containing an aberrant splice site. The 18-mer IV654 antisense oligonucleotide restores eGFP fluorescence by sterically blocking the aberrant splice site, allowing correct splicing of the eGFP transcript.[17] Measuring eGFP fluorescence through flow cytometry provides a functional readout of PMO delivery to the nucleus through splice-correction of an altered eGFP transcript. The PMO delivery of the ten library pools is shown in the blue bar graph in **Fig. 3.3**, relative to the nuclear PMO delivery of un-conjugated PMO. The five most cationic peptide pools (based on cation exchange retention time) showed significant delivery of the PMO cargo, with fraction 10 showing the highest delivery.

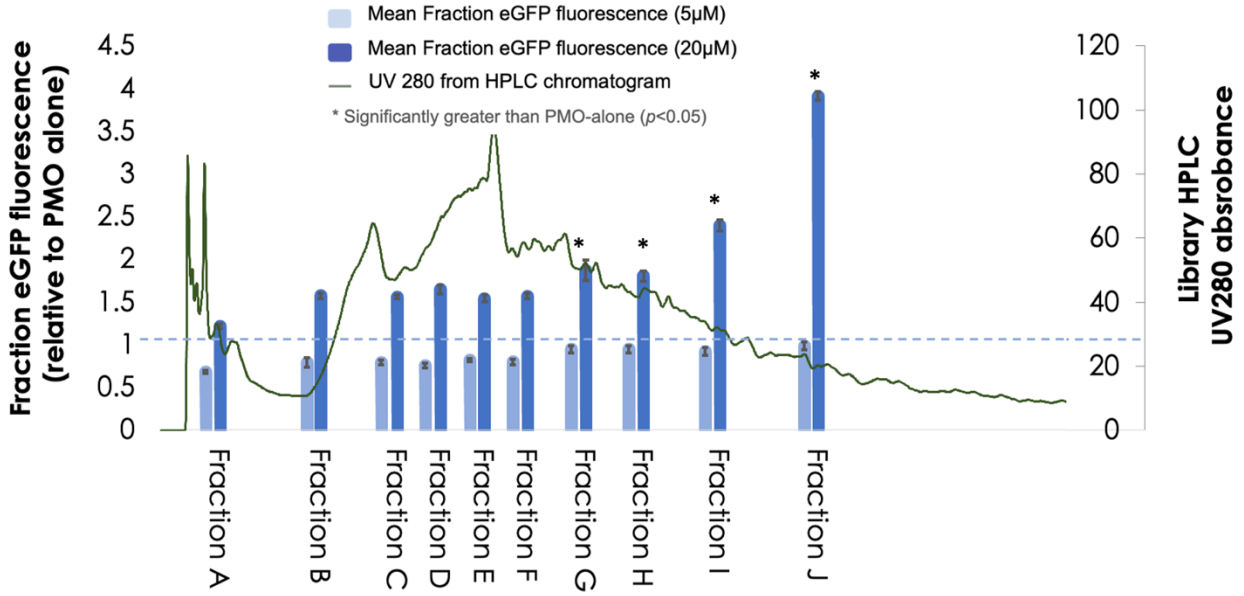


Figure 3.3 eGFP fluorescence of HeLa 654 cells treated with 5 or 20 μM total PMO-peptide from the library fractions.
The peptides were resolved through cation exchange chromatography (green chromatogram) in ammonium acetate buffer (pH 5) with a 20-500 mM gradient over 75 minutes. The x-axis represents time (minutes) for both the green chromatogram and the blur bar graphs. The eGFP fluorescence of each peptide-PMO fraction

relative to the fluorescence of 5 or 20 μM of PMO alone determined via flow cytometry is superimposed on the chromatogram from the CX separation of the library peptides. PMO-peptide fractions G through J improved PMO delivery significantly over the PMO-654 alone at 20 μM ($p<0.05$).

The most penetrant pool of peptides, fraction J, was sequenced by MS/MS. Almost 100 peptide sequences of high sequencing confidence were recovered. The peptide sequences with the highest sequencing fidelity and lowest arginine content, to avoid arginine-derived peptide toxicity, were selected for further testing.[18] These nine peptides were named CXP, to designate that they came from cation-exchange pooling of a library. The pure peptides were generated via solid-phase peptide synthesis, conjugated to the PMO cargo, and tested for PMO delivery in HeLa 654 cells. All nine peptides showed significant delivery relative to unconjugated PMO at 20 μM (**Fig. 3.4**), which corroborates the high PMO delivery seen by treatment with fraction J.

We have previously shown that the presence of a single highly active CPP does not change whether a 100-member peptide library delivers a PMO cargo into HeLa 654 cells.[9] Based on these previous results, we would expect fraction J to contain numerous CPPs that contribute to the significant PMO delivery of the entire library pool.[9] Indeed, all nine peptides tested showed significant PMO delivery. This correlates with the results of the previous study, suggesting that for a library pool to show significant PMO delivery, it must contain multiple cell-penetrating peptides.

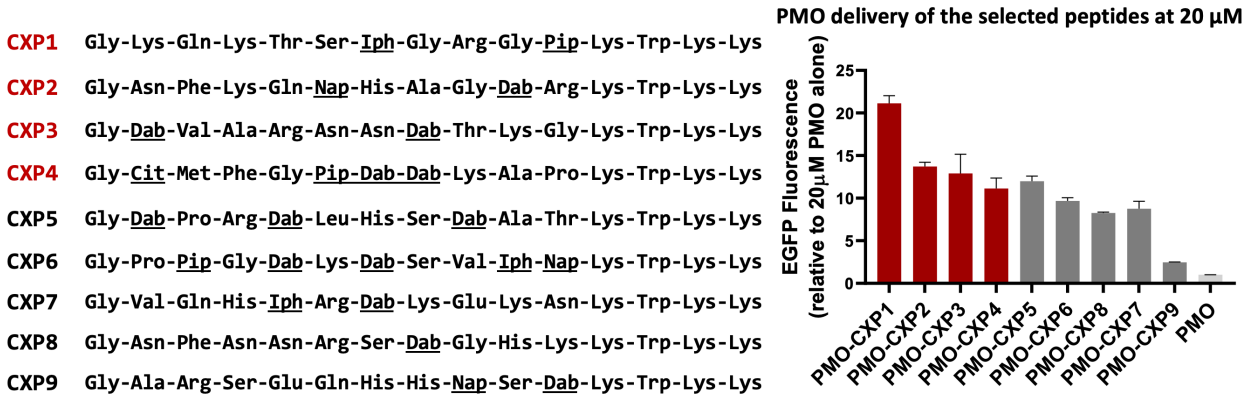


Figure 3.4 The PMO delivery of the eleven peptide sequences chosen from fraction J.
The nine peptides were conjugated to PMO-DBCO via N-term azidopentanoic acid. The eGFP fluorescence of each peptide-PMO relative to the fluorescence of 20 μM of PMO 654 alone determined via flow cytometry. All PMO-peptides improved PMO delivery significantly over the PMO-654 alone at 20 μM ($p<0.05$). Penetratin, a known CPP, was included as a positive control.

3.2.3 Hit library peptides show high PMO delivery

Four peptides in particular, **CXP1**, **CXP2**, **CXP3**, and **CXP4**, showed high levels of PMO delivery, and were brought forward for further testing. The peptides showed dose-dependent delivery of PMO (**Fig. 3.5a**) with no significant toxicity within the tested concentration range (**Fig. 3.5b**). These four peptides demonstrate that the CXP assay platform allows for discovery of novel and efficacious sequences for PMO delivery. Although the peptides do not show PMO delivery as efficacious as the CPP penetratin, on which this library was based, **CXP1** does approach this high PMO delivery. In addition, The CXP peptides all show lower toxicity at the membrane when compared to the parent penetratin peptide (Fig **3.5b**). For therapeutic use, low toxicity can be as important as high PMO delivery.

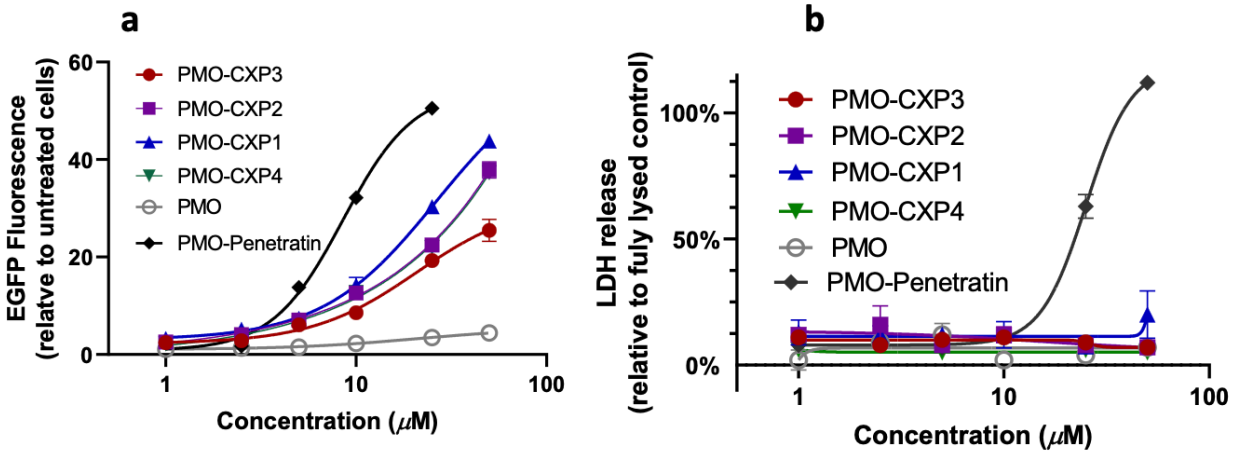


Figure 3.5 PMO delivery and membrane toxicity of the four most effective CXP peptides. (a) HeLa 654 cells were treated with 1, 2.5, 5, 10, 25, or 50 μM PMO-CPP for 22 h prior to flow cytometry. Results are given as the mean EGFP fluorescence of cells treated with PMO-peptide relative to the fluorescence of cells treated with vehicle only. Bars represent mean ± SD, N = 3. (b) Cell supernatant from (a) was tested for LDH release as a measure of membrane permeability. Results are given as percent LDH release above vehicle relative to fully lysed cells.

3.2.4 Chimeric peptides demonstrate improved PMO delivery over parent CXP peptides

It has been previously demonstrated that CPPs can more effectively deliver PMO as chimeras, a linear combination of multiple peptides.[17] In the previous case, these chimeric peptides were highly efficacious, but also showed potent toxicity in cellular assays. Because the peptide **CXP1** showed high PMO delivery and low toxicity, it became a strong candidate for a linear chimera. The chimeric versions of **CXP1** in combination with **CXP3** and **CXP4,** as well as **a**

CXP1-CXP1 dimer, were generated through SPPS (**Fig. 3.6a). CXP2** was not included due to its increased toxicity at high doses (**Appendix I**). The dimeric peptide **CXP1-CXP1** showed a nearly 100-fold increase in PMO delivery efficacy compared to the CXP1 parent peptide (**Fig 3.6b,** blue), and the other CXP-based chimeras also showed a strong improvement in PMO delivery relative to their parent peptides (**Fig. 3.6b**, green and orange). While these chimeric peptides reach maximum efficacy in delivery before they begin to show significant membrane toxicity (**Fig 3.6c**), they do demonstrate toxicity at lower concentrations than the parent CXP peptides. Nevertheless, these chimeras show higher activity and less toxicity than those previously generated and tested for PMO delivery within the Pentelute lab.[17]

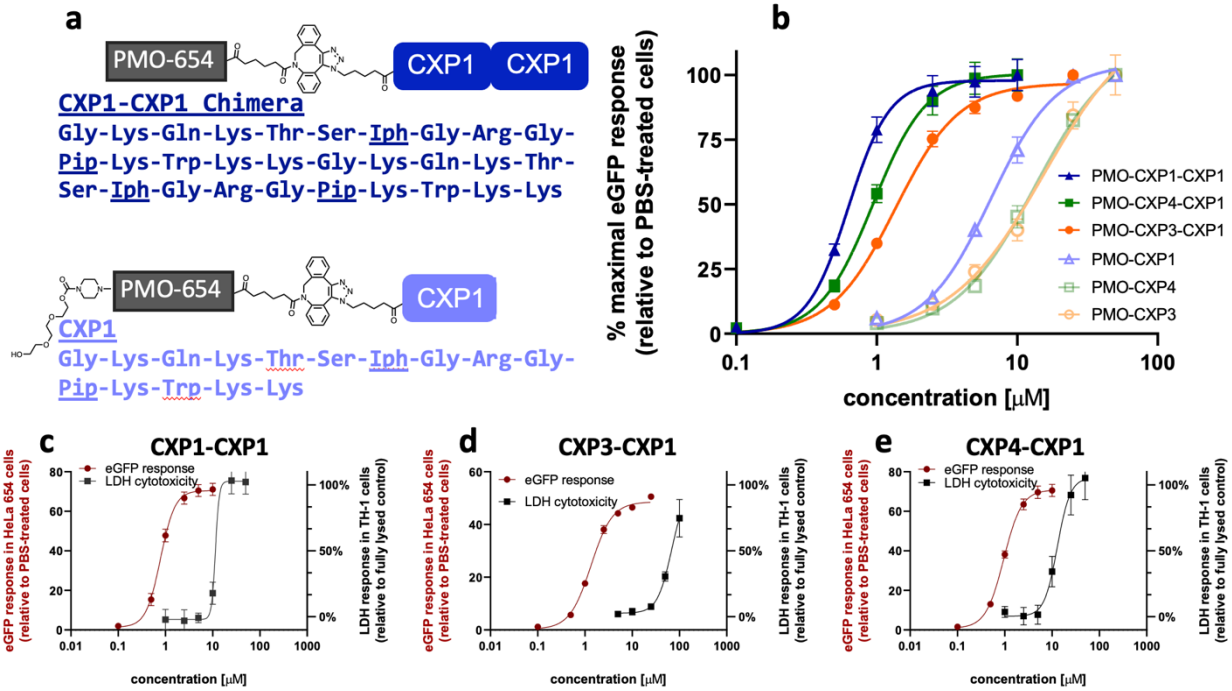
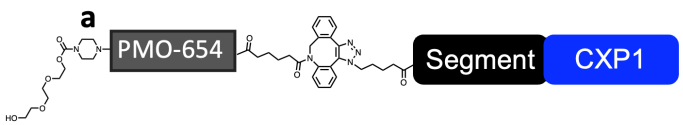


Figure 3.6 Chimeric peptides show high efficacy for PMO delivery.
(a) The sequence and structure of the chimeric PMO-CXP1-CXP1 (top) and singular PMO-CXP1 (bottom). (b) HeLa 654 cells were treated with 1, 2.5, 5, 10, 25, or 50 μM PMO-CPP for 22 h prior to flow cytometry. Results are given as the mean EGFP fluorescence of cells treated with PMO-peptide relative to the fluorescence of cells treated with vehicle only. Bars represent mean ± SD, N = 3. (c-e) Therapeutic windows of the three chimeric peptides. PMO-delivery was measured through the HeLa 654 EGFP assay, and toxicity was measured through LDH release at 18 hours in TH-1 renal cells.

3.2.5 Sequence-activity relationship studies allow the rational design of two CXD1-derived peptides with high PMO delivery and lower toxicity

The relationship between the CXP1 sequence and PMO delivery was probed to further increase the efficacy of the constructs. Because small peptide segments have previously shown very low PMO delivery,[19] we instead used a chimeric strategy to determine the functions of different segments of the CXP1 sequence. Short, overlapping segments (named SA, SB, SC, and SD) of the CXP1 sequence with a net 2+ charge were appended to PMO on the N-terminus via SPACC and the CXP1 sequence on the C-terminus, as shown in **Fig. 3.7a**. We tested each segment for PMO delivery and toxicity to determine which portions of the sequence contribute to the penetration and toxicity of the CXP1-CXP1 chimera. Segments A and B showed the highest PMO delivery (**Fig. 3.7b**), while peptide **SA-CXP1** had the toxicity closest to the parent CXP1-CXP1 peptide (**Fig. 3.7c**). This suggests that the "Iph-Gly-Arg" motif across segments A and B is likely important for penetration, while the "Gly-Pip" motif unique to segment A may be contributing to peptide toxicity.



SA-CX1
Gly-Iph-Gly-Arg-Gly-Pip-Lys-Gln-Lys-Thr-Ser-Iph-Gly-Arg-Gly-Pip-Lys-Trp-Lys-Lys

SB-CX1
Gly-Lys-Thr-Ser-Iph-Gly-Arg-Lys-Gln-Lys-Thr-Ser-Iph-Gly-Arg-Gly-Pip-Lys-Trp-Lys-Lys

SC-CX1
Gly-Lys-Gln-Lys-Thr-Ser-Lys-Gln-Lys-Thr-Ser-Iph-Gly-Arg-Gly-Pip-Lys-Trp-Lys-Lys

SD-CX1
Gly-Lys-Trp-Lys-Lys-Gln-Lys-Thr-Ser-Iph-Gly-Arg-Gly-Pip-Lys-Trp-Lys-Lys

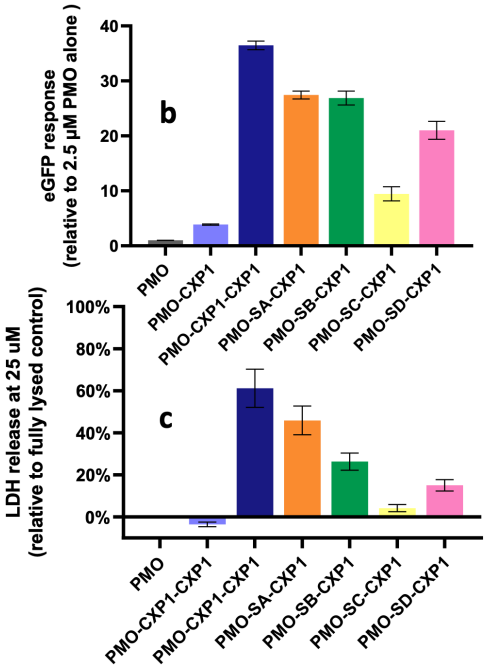


Figure 3.7 The segmented chimeras of CXP1 show key regions for efficacy and toxicity.
(a) The sequence and structure of the chimeric PMO-SX-CXP1 peptides. (b) HeLa 654 cells were treated with 2.5 µM PMO-CPP for 22 h prior to flow cytometry. Results are given as the mean EGFP fluorescence of cells treated with PMO-peptide relative to the fluorescence of cells treated with vehicle only. Bars represent mean ± SD, N = 3. (c) Supernatant from cells treated in (b) was tested for LDH release at 25 µM as a measure of membrane permeability. Results are given as percent LDH release above vehicle relative to fully lysed cells.

The results of the sequence-activity relationship assays allowed the rational design of two new peptide sequences based on the original CXP1. These "CXD" cation-exchange designed peptide sequences are shorter than the CXP1-CXP1 chimera and lack the potentially toxic "Gly-Pip" motif. These two sequences (**Fig. 3.8**) were synthesized and demonstrated high PMO delivery, especially the **CXD1B** sequence, which is longer and has a higher net charge. This peptide showed high PMO delivery and lower toxicity than the CXP1-CXP1 sequence, resulting in a wider therapeutic window. In designing **CXD1B**, we were able to apply the results from the sequence-activity relationship study to rationally design a novel, cationic peptide for PMO delivery.

CXD1A
Gly-Lys-Thr-Ser-Iph-Gly-Arg-Lys-Gln-Lys-Thr-Ser-Iph-Gly-Arg-Lys-Trp-Lys-Lys

CXD1B
Gly-Lys-Gln-Lys-Thr-Ser-Iph-Gly-Arg-Lys-Trp-Lys-Gln-Lys-Thr-Ser-Iph-Gly-Arg-Lys-Trp-Lys-Lys

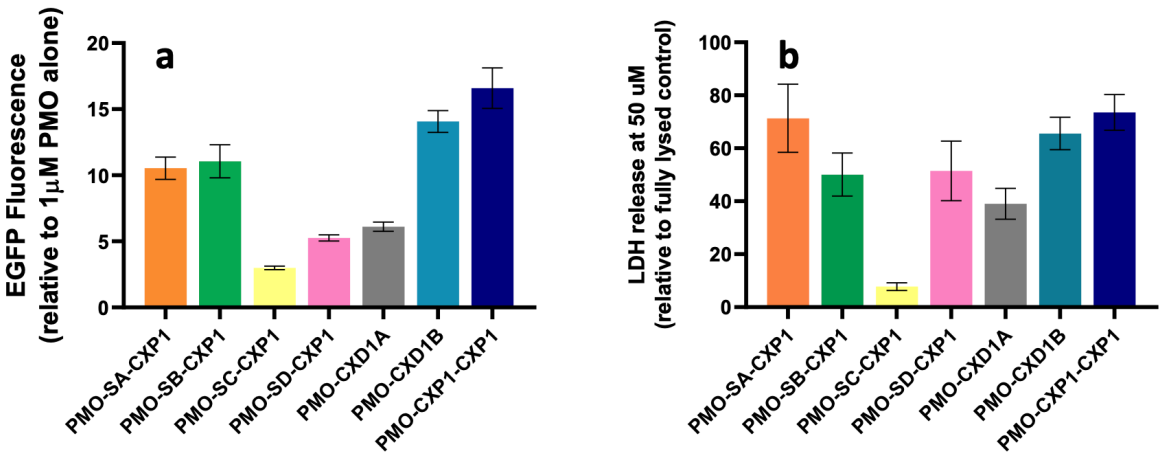


Figure 3.8 The designed peptides CXD1A and CXD1B show high PMO delivery and lower toxicity than longer chimeras.
(a) The sequence and structure of the chimeric PMO-SX-CXP1 peptides. (b) HeLa 654 cells were treated with 1 μM PMO-CPP for 22 h prior to flow cytometry. Results are given as the mean EGFP fluorescence of cells treated with PMO-peptide relative to the fluorescence of cells treated with vehicle only. Bars represent mean ± SD, N = 3. (c) Supernatant from cells treated in (b) was tested for LDH release at 50 μM as a measure of membrane permeability. Results are given as percent LDH release above vehicle relative to fully lysed cells.

3.3 Conclusions

The chromatographic resolution of a peptide library into smaller pools enables novel peptide discovery. The library pools resolved by charge demonstrate the immense importance of positive charge to CPPs for PMO delivery, which we hope to integrate with machine learning models in the future to enable further discovery.20,21 The peptides sequenced from the most penetrant pool all demonstrated effective PMO delivery, and four CXP peptides showed efficacious delivery of the PMO cargo with low toxicity.

The higher efficacy of the chimeric peptides underscores the effectiveness of the linear chimera strategy in generating CPPs. The CXD1B sequence demonstrated high PMO delivery with nanomolar efficacy and lower toxicity than previous chimeric PMO-peptide constructs. We hope that the wide therapeutic windows of the CXP1 and CXD1B peptides will enable in-vivo delivery of PMO cargoes.

3.4 Experimental section

3.4.1 Reagents and Solvents

H-Rink Amide-ChemMatrix resin was obtained from PCAS BioMatrix Inc. (St-Jean-sur-Richelieu, Quebec, Canada) and TentaGel was obtained from Rapp Polymere (Tuebingen, Germany). 1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium-3- oxid-hexafluorophosphate (HATU), Nα-Fmoc-Nγ-Boc-L-2,4-diaminobutyric acid, Fmoc-L-Iodophenylalanine, Fmoc-3-(1-naphthyl)-D-alanine, 1-Boc-piperidine-4-Fmoc-amino-4-carboxylic acid, Fmoc-L-citrulline, Fmoc-3-(4'-pyridyl)-L-alanine, and 5-azidopentanoic acid were purchased from Chem-Impex International (Wood Dale, IL). Fmoc-protected L-amino acids (Fmoc-Arg(Pbf)-OH; Fmoc-Asn(Trt)-OH; Fmoc-Asp-(Ot-Bu)-OH; Fmoc-Gln(Trt)-OH; Fmoc-Glu(Ot-Bu)-OH; Fmoc-Gly-OH; Fmoc-His(Trt)-OH; Fmoc-Lys(Boc)-OH; Fmoc-Phe-OH; Fmoc-Pro-OH; Fmoc-Ser(But)-OH; Fmoc-Thr(t-Bu)-OH; Fmoc-Trp(Boc)-OH), were purchased from the Novabiochem-line from MilliporeSigma. Dibenzocyclooctyne acid was purchased from Click Chemistry Tools (Scottsdale, AZ). Peptide synthesis-grade N,N-dimethylformamide (DMF), CH2Cl2 (DCM), diethyl ether, and HPLC-grade acetonitrile were obtained from VWR International (Radnor, PA). The LDH Assay kit was purchased from Promega (Madison, WI). All other reagents were purchased from Sigma-Aldrich (St. Louis, MO). Milli-Q water was used exclusively.

3.4.2 General peptide preparation

Fast-flow Peptide Synthesis:

Peptides were synthesized on a 0.1 mmol scale using a semi-automated fast-flow peptide synthesizer as previously reported.[45] 1 mmol of amino acid was combined with 2.5 mL of 0.4 M HATU and 500 μL of DIEA and mixed before being delivered to the reactor containing resin via syringe pump at 6 mL/min. The reactor was submerged in a water bath heated to 70 ºC. An HPLC pump delivered either DMF (20 mL) for washing or 20% piperidine/DMF (6.7 mL) for Fmoc deprotection, at 20 mL/min.

Peptide cleavage and deprotection:

Each peptide was subjected to simultaneous global side-chain deprotection and cleavage from resin by treatment with 5 mL of 94% trifluoroacetic acid (TFA), 2.5% thioanisole, 2.5% water, and 1% triisopropylsilane (TIPS) (v/v) at room temperature for 2 to 4 h. The cleavage cocktail was first concentrated by bubbling N_2 through the mixture, and cleaved peptide was precipitated and triturated with 40 mL of cold ether (chilled in dry ice). The crude product was pelleted by centrifugation for three minutes at 4,000 rpm and the ether was decanted. This wash step was repeated two more times. After the third wash, the pellet was dissolved in 50% water and 50% acetonitrile containing 0.1% TFA, filtered through a fritted syringe to remove the resin and lyophilized.

Peptide Purification:

The peptides were dissolved in water and acetonitrile containing 0.1% TFA, filtered through a 0.22 μm nylon filter and purified by mass-directed semi-preparative reversed-phase HPLC. Solvent A was water with 0.1% TFA additive and Solvent B was acetonitrile with 0.1% TFA additive. A linear gradient from 5 to 45% B that changed at a rate of 0.5% B/min was used. Most of the peptides were purified on an Agilent Zorbax SB C18 column: 9.4 x 250 mm, 5 μm. Based on target ion mass data recorded for each fraction, only pure fractions were pooled and lyophilized. The purity of the fraction pool was confirmed by LC-MS.

Preparation of peptide library:

Split-and-pool synthesis was carried out on 180 μm TentaGel resin (0.28 mmol/g) for a 50,000 member library. Splits were performed by suspending the resin in DCM and dividing it evenly (via pipetting) among 22 plastic fritted syringes on a vacuum manifold. Couplings were carried out as follows: solutions of Fmoc-protected amino acids (10 equivalents relative to the resin loading), PyAOP (0.38 M in DMF; 0.95 eq. relative to amino acid), and DIEA (1.1 eq. for histidine; 3 eq. for all other amino acids) were each added to individual portions of resin. Couplings were allowed to proceed for 60 min. Resin portions were recombined and washed with DCM and DMF. Fmoc removal was carried out by treatment of the resin with 20% piperidine in DMF (1x flow wash; 2x 10 min batch treatments). Resin was washed again with DMF and DCM before the next split.

After synthesis, the library was separated into 3,000-member portions by mass. Each portion was conjugated to 5-azidopentanoic acid through a coupling with PYAOP as described above. Library

peptides were then cleaved as described above and lyophilized to generate libraries of 3,000 peptide sequences.

Preparation of PMO-DBCO:

PMO IVS2-654 (50 mg, 8 μmol) obtained from Sarepta Therapeutics was dissolved in 150 μL DMSO. To the solution was added a solution containing 2 equivalents of dibenzocyclooctyne acid (5.3 mg, 16 μmol) activated with HBTU (37.5 μL of 0.4 M HBTU in DMF, 15 μmol) and DIEA (2.8 μL, 16 μmol) in 40 μL DMF (Final reaction volume = 0.23 mL). The reaction proceeded for 25 min before being quenched with 1 mL of water and 2 mL of ammonium hydroxide. The ammonium hydroxide hydrolyzed any ester formed during the course of the reaction. After 1 hour, the solution was diluted to 40 mL in water/acetonitrile and purified using reverse-phase HPLC (Agilent Zorbax SB C3 column: 21.2 x 100 mm, 5 μm) and a linear gradient from 2 to 60% B (solvent A: water; solvent B: acetonitrile) over 58 min (1% B / min). Using mass data about each fraction from the instrument, only pure fractions were pooled and lyophilized. The purity of the fraction pool was confirmed by LC-MS.

Conjugation of PMO to peptides:

PMO-DBCO (1 eq, 5 mM, water) was conjugated to azido-peptides (1 eq, 5 mM, water) or azido-peptide library (1eq, 1mM, water) at room temperature for 2 h, or 12 hours for peptide library. Reaction progress was monitored by LC-MS and additional stock of azido-peptide was added until all PMO-DBCO was consumed. The purity of the final construct was confirmed by LC-MS to be >95%.

3.4.3 Liquid chromatography—mass spectrometry

LC-MS analyses:

Analysis was performed on an Agilent 6550 iFunnel Q-TOF LC-MS system (abbreviated as 6550) coupled to an Agilent 1290 Infinity HPLC system. Mobile phases were: 0.1% formic acid in water (solvent A) and 0.1% formic acid in acetonitrile (solvent B). The following LC-MS method was used for characterization:

1-61% B over 6 min, Zorbax C3 column (6550)

LC: Agilent EclipsePlus C18 RRHD column: 2.1 × 50 mm, 1.8 µm, column temperature: 40 °C, gradient: 0-1 min 1% B, 1-6 min, 1-61% B, 6-7 min, 91% B, 7-8 min, 1% B; flow rate: 0.5 mL/min. MS: Positive electrospray ionization (ESI) extended dynamic range mode in mass range 300–3000 m/z. MS is on from 1 to 6 min.

All data were processed using Agilent MassHunter software package. Y-axis in all chromatograms shown represents total ion current (TIC) unless noted.

Orbitrap LC-MS/MS:

Analysis was performed on an EASY-nLC 1200 (Thermo Fisher Scientific) nano-liquid chromatography handling system connected to an Orbitrap Fusion Lumos Tribrid Mass Spectrometer (Thermo Fisher Scientific). Samples were run on a PepMap RSLC C18 column (2 µm particle size, 15 cm × 50 µm ID; Thermo Fisher Scientific, P/N ES901). A nanoViper Trap Column (C18, 3 µm particle size, 100 Å pore size, 20 mm × 75 µm ID; Thermo Fisher Scientific, P/N 164946) was used for desalting. The standard nano-LC method was run at 40 °C and a flow rate of 300 nL/min with the following gradient: 1% solvent B in solvent A ramping linearly to 41% B in A over 55 min, where solvent A = water (0.1% FA), and solvent B = 80% acetonitrile, 20% water (0.1% FA). Positive ion spray voltage was set to 2200 V. Orbitrap detection was used for primary MS, with the following parameters: resolution = 120,000; quadrupole isolation; scan range = 150–1200 m/z; RF lens = 30%; AGC target = 250%; maximum injection time = 100 ms; 1 microscan. Acquisition of secondary MS spectra was done in a data-dependent manner: dynamic exclusion was employed such that a precursor was excluded for 30 s if it was detected four or more times within 30 s (mass tolerance: 10.00 ppm); monoisotopic precursor selection used to select for peptides; intensity threshold was set to 2×10^4; charge states 2–10 were selected; and precursor selection range was set to 200–1400 m/z. The top 15 most intense precursors that met the preceding criteria were subjected to subsequent fragmentation. Two fragmentation modes— higher-energy collisional dissociation (HCD), and electron-transfer/higher-energy collisional dissociation (EThcD)—were used for acquisition of secondary MS spectra. Detection was performed in the Orbitrap (resolution = 30,000; quadrupole isolation; isolation window = 1.3 m/z; AGC target = 2×10^4; maximum injection time = 100 ms; 1 microscan). For HCD, a stepped collision energy of 3, 5, or 7% was used. For EThcD, a supplemental activation collision energy of 25% was used.

De novo peptide sequencing and filtering:

De novo peptide sequencing of the acquired data was performed in PEAKS 8 (BioInformatics Solutions Inc.). Using PEAKS, spectra were prefiltered to remove noise, and sequenced. All non-canonical amino acids were sequenced as post-translational modifications based on the canonical amino acid most closely matching their molecular mass. Twenty candidate sequence assignments were created for each secondary scan.

3.4.4 Library fractionation and pooling

Library separation:

The 3,000 member peptide library was dissolved in loading buffer (10mM ammonium acetate, pH 5, 20% Acetonitrile). The peptides were then resolved on a Propac SCX-10 column (Thermo Fisher).

Solvent A was 10mM ammonium acetate with 10% acetonitrile, pH 5, and Solvent B was 1M ammonium acetate with 10% acetonitrile, pH 5. A linear gradient from 1 to 75% B over 75 minutes was used, with a 10 minute hold at 1% B to allow compound to load onto the column. Fractions eluting off the column were collected every minute, with a total volume of 1mL.

Library desalting:

All fractions were lyophilized overnight, and then re-dissolved in 1 mL of water. Fractions were then frozen and re-lyophilized to remove remaining ammonium acetate buffer. This process was repeated for at least 3 lyophilization cycles to fully remove the volatile buffer.

Library pooling:

After desalting, the peptide concentration on each fraction was measure spectroscopically at 280nM, based on the absorbance of the single tryptophan residue in each peptide library member. Neighboring peptide fractions were combined based on peptide concentration to generate 10 pools with approximately equal peptide concentration. This resulted in pools containing about 300 peptide sequences, based on the original library of 3,000 peptide sequences.

3.4.5 In-vitro evaluation of PMO-peptides

EGFP Assay:

HeLa 654 cells obtained from the University of North Carolina Tissue Culture Core facility were maintained in DMEM supplemented with 10% (v/v) fetal bovine serum (FBS) and 1% (v/v) penicillin-streptomycin at 37 °C and 5% CO_2. 18 h prior to treatment, the cells were plated at a density of 5,000 cells per well in a 96-well plate in DMEM supplemented with 10% FBS and 1% penicillin-streptomycin.

For individual peptide testing, PMO-peptides were dissolved in PBS without Ca^{2+} or Mg^{2+} at a concentration of 1 mM (determined by UV absorbance of the PMO) before being diluted in DMEM. Cells were incubated at the designated concentrations in triplicate for 22 h at 37 °C and 5% CO_2. Next, the treatment media was removed, and the cells were washed once before being incubated with 0.25 % Trypsin-EDTA for 15 min at 37 °C and 5% CO_2. Lifted cells were transferred to a V-bottom 96-well plate and washed once with PBS, before being resuspended in PBS containing 2% FBS and 2 μg/mL propidium iodide (PI). Flow cytometry analysis was carried out on a BD LSRII flow cytometer. Gates were applied to the data to ensure that cells that were positive for propidium iodide or had forward/side scatter readings that were sufficiently different from the main cell population were excluded. Each sample was capped at 5,000 gated events.

Analysis was conducted using Graphpad Prism 7 and FlowJo. For each sample, the mean fluorescence intensity (MFI) and the number of gated cells was measured. To report activity, triplicate MFI values were averaged and normalized to the PMO alone condition.

LDH assay:

Cytotoxicity assays were performed in HeLa 654 cells or TH-1 renal cells. TH-1 cells were maintained in DMEM supplemented with 10% (v/v) fetal bovine serum (FBS) and 1% (v/v) penicillin-streptomycin at 37 °C and 5% CO_2H-1. Cells were plated at 5,000 cells/well in a 96-well plate 18-hours prior to treatment, and treated with PMO-peptide compounds as described above. HeLa cell LDH assays were performed concurrently on the cells treated for Flow cytometry, so all plating and treatment protocols match those described above.

17 hours after cell treatment, the Promega CytoTox lysis solution was added to 3 wells of each plate as a fully lysed control (100% LDH-release). 18 hours after treatment, cell supernatant was transferred to a new 96-well plate for analysis of LDH release. To each well of the 96-well plate containing supernatant was added CytoTox 96 Reagent (Promega). The plate was shielded from

light and incubated at room temperature for 30 min. Equal volume of Stop Solution was added to each well, mixed, and the absorbance of each well was measured at 490 nm. The measurement of vehicle-treated cells was subtracted from each measurement, and % LDH release was calculated as % cytotoxicity = 100 × Experimental LDH Release (OD490) / Maximum LDH Release (OD490).

3.5 Acknowledgements

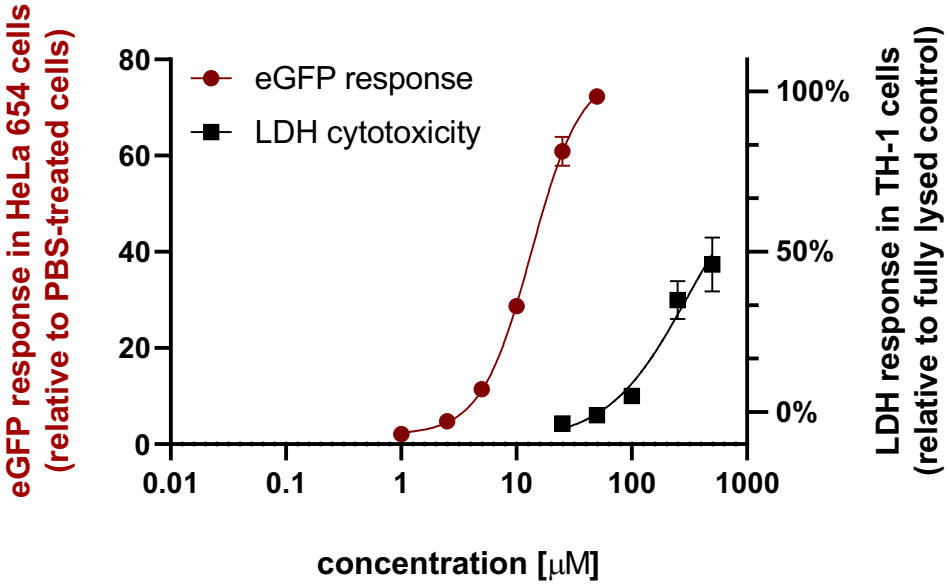
This research was funded by Sarepta Therapeutics. C.K. Schissel (4000057398) and C.E. Farquhar (4000057441) acknowledge the National Science Foundation Graduate Research Fellowship (NSF Grant No. 1122374) for research support. We acknowledge support from the Swanson Biotechnology Center Flow Cytometry Facility at the Koch Institute for Integrative Cancer Research at MIT through the use of their flow cytometers (NCI Cancer Center Support Grant P30-CA14051).

3.6 Appendix I: Therapeutic windows of selected PMO-peptides

For all therapeutic window studies, PMO-delivery was measured through the HeLa 654 EGFP assay, and toxicity was measured through LDH release at 18 hours in TH-1 renal cells.

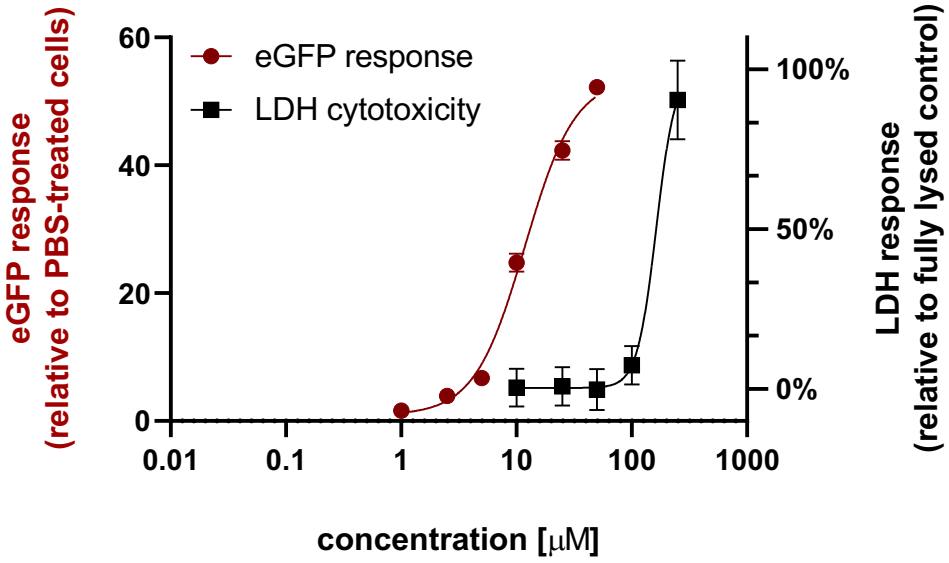
PMO-CXP1

Peptide sequence: 5azido-Gly-Lys-Gln-Lys-Thr-Ser-Iph-Gly-Arg-Gly-Pip-Lys-Trp-Lys-Lys



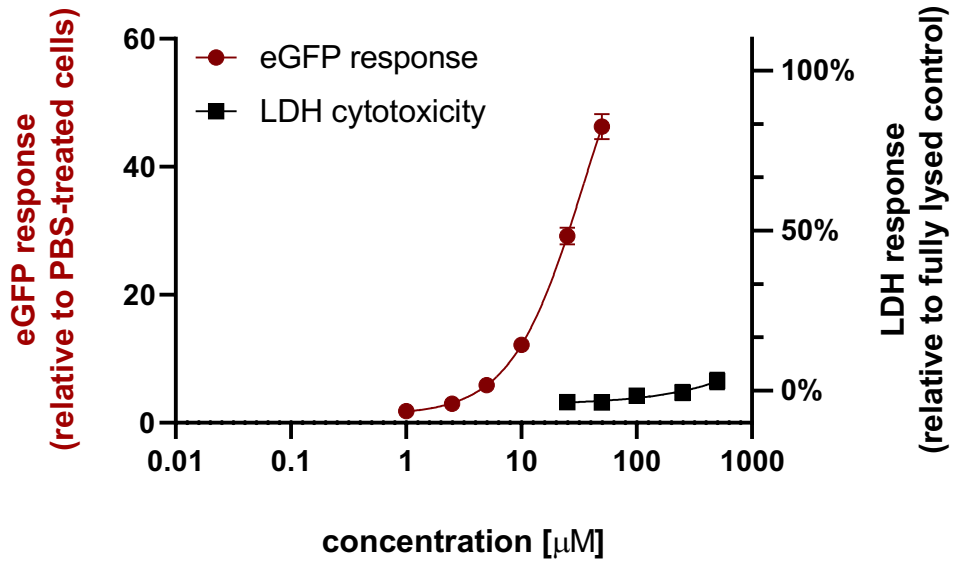
PMO-CXP2

Peptide sequence: 5azido-Gly-Asn-Phe-Lys-Gln-Nap-His-Ala-Gly-Dab-Arg-Lys-Trp-Lys-Lys



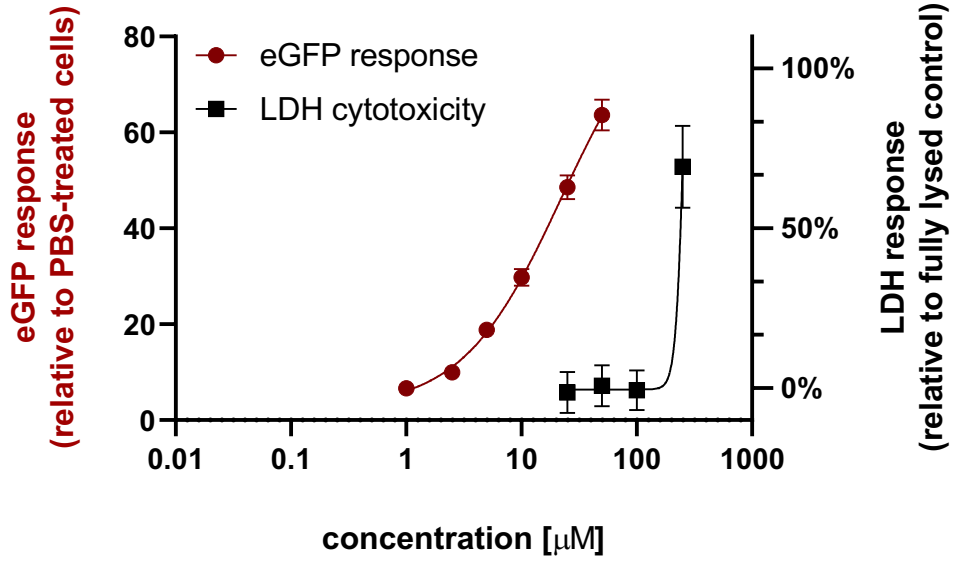
PMO-CXP3

Peptide sequence: 5azido-Gly-Dab-Val-Ala-Arg-Asn-Asn-Dab-Thr-Lys-Gly-Lys-Trp-Lys-Lys



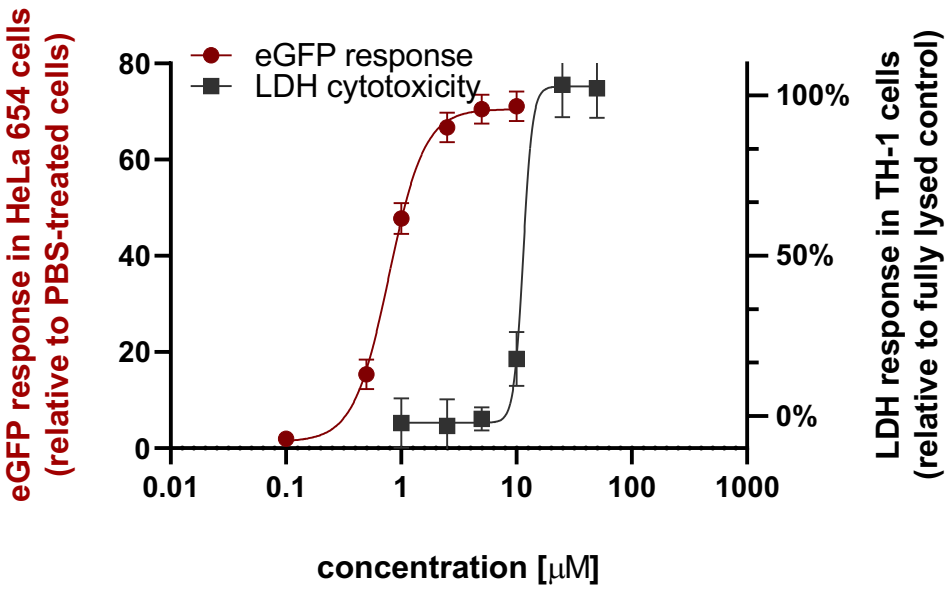
PMO-CXP4

Peptide sequence: 5azido-Gly-Cit-Met-Phe-Gly-Pip-Dab-Dab-Lys-Ala-Pro-Lys-Trp-Lys-Lys



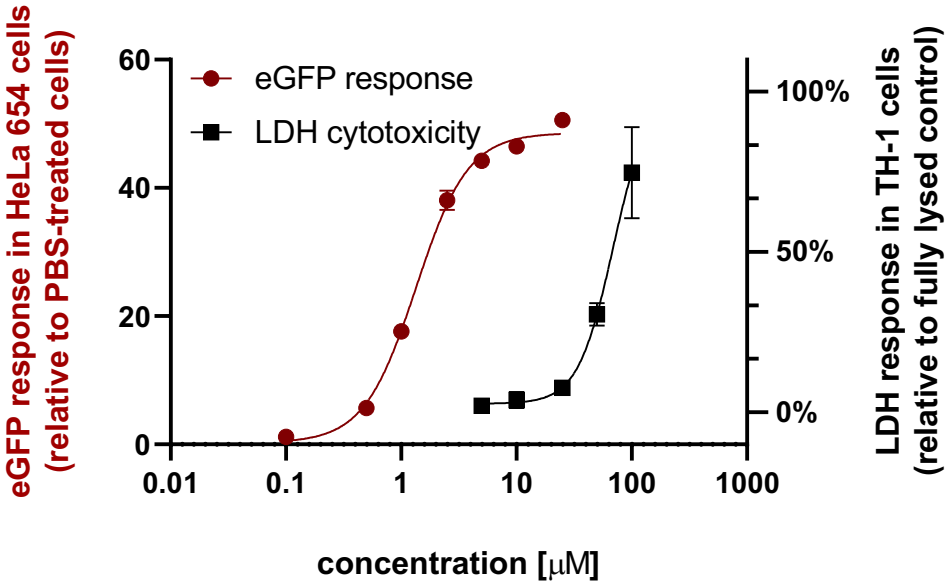
PMO-CXP1-CXP1

Peptide sequence: 5azido-Gly-Lys-Gln-Lys-Thr-Ser-Iph-Gly-Arg-Gly-Pip-Lys-Trp-Lys-Lys- Gly-Lys-Gln-Lys-Thr-Ser-Iph-Gly-Arg-Gly-Pip-Lys-Trp-Lys-Lys



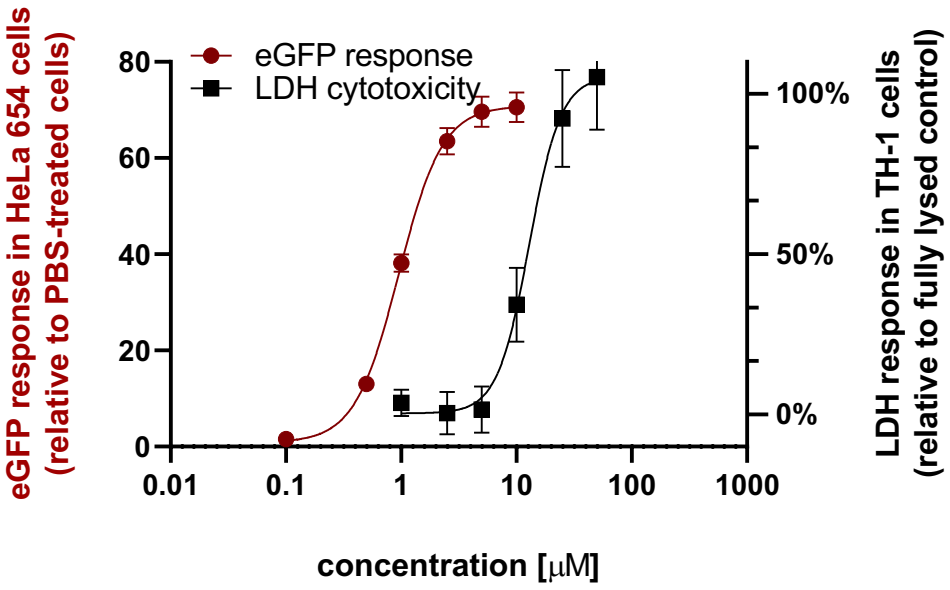
PMO-CXP3-CXP1

Peptide sequence: 5azido-Gly-Dab-Val-Ala-Arg-Asn-Asn-Dab-Thr-Lys-Gly-Lys-Trp-Lys-Lys-Gly-Lys-Gln-Lys-Thr-Ser-Iph-Gly-Arg-Gly-Pip-Lys-Trp-Lys-Lys



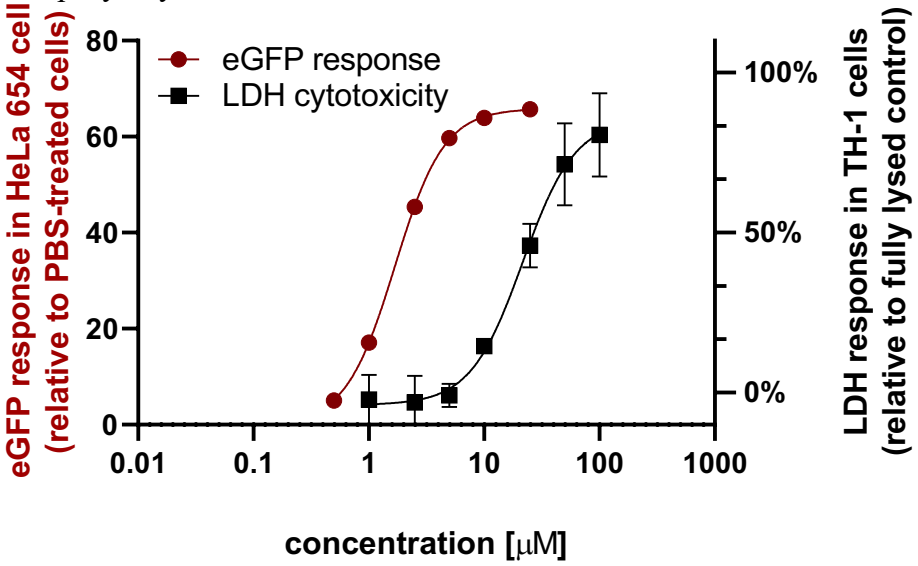
PMO-CXP4-CXP1

Peptide sequence: 5azido-Gly-Cit-Met-Phe-Gly-Pip-Dab-Dab-Lys-Ala-Pro-Lys-Trp-Lys-Lys-Gly-Lys-Gln-Lys-Thr-Ser-Iph-Gly-Arg-Gly-Pip-Lys-Trp-Lys-Lys



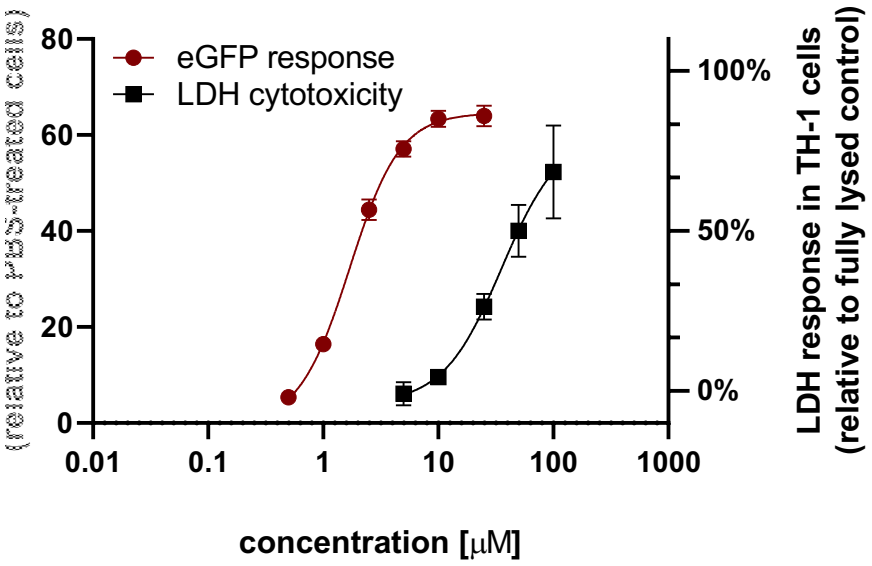
PMO-SA-CXP1

Peptide sequence: 5azido-Gly-Iph-Gly-Arg-Gly-Pip-Lys-Gln-Lys-Thr-Ser-Iph-Gly-Arg-Gly-Pip-Lys-Trp-Lys-Lys



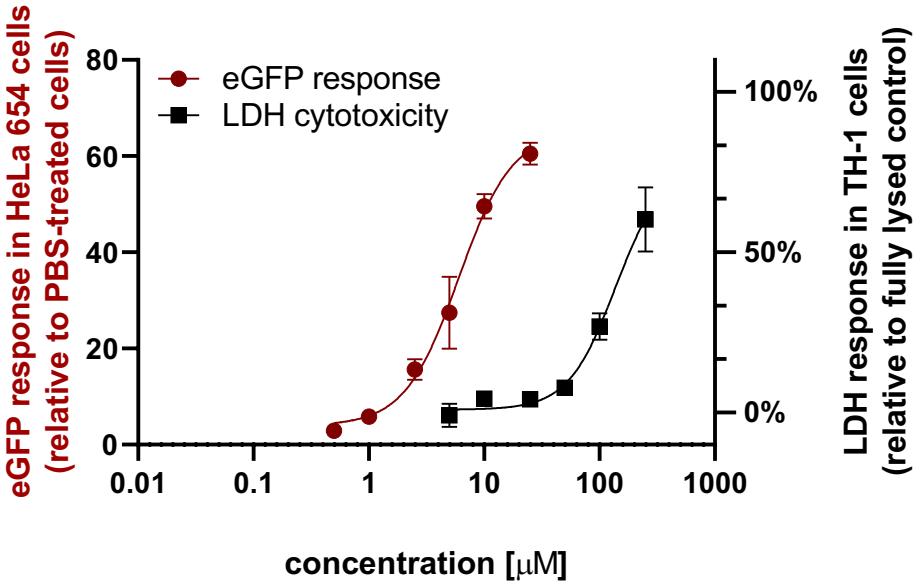
PMO-SB-CXP1

Peptide sequence: 5azido-Gly-Lys-Thr-Ser-Iph-Gly-Arg-Lys-Gln-Lys-Thr-Ser-Iph-Gly-Arg-Gly-Pip-Lys-Trp-Lys-Lys



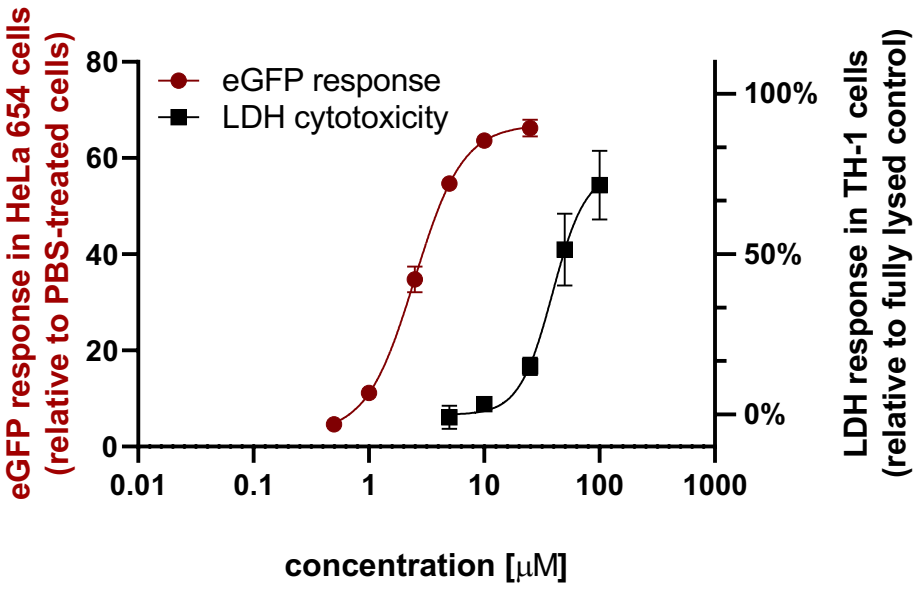
PMO-SC-CXP1

Peptide sequence: 5azido-Gly-Lys-Gln-Lys-Thr-Ser-Lys-Gln-Lys-Thr-Ser-Iph-Gly-Arg-Gly-Pip-Lys-Trp-Lys-Lys



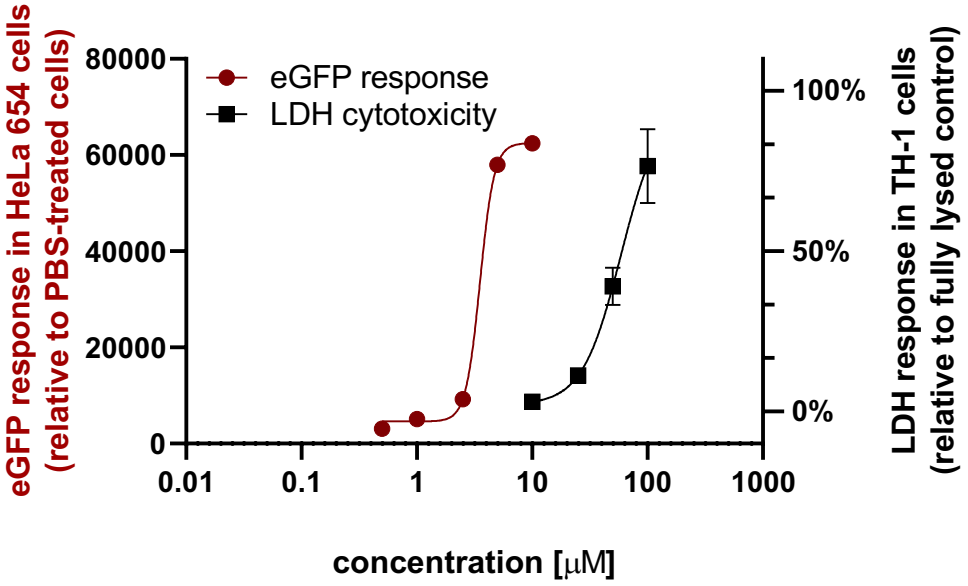
PMO-SD-CXP1

Peptide sequence: 5azido-Gly-Lys-Trp-Lys-Lys-Gln-Lys-Thr-Ser-Iph-Gly-Arg-Gly-Pip-Lys-Trp-Lys-Lys



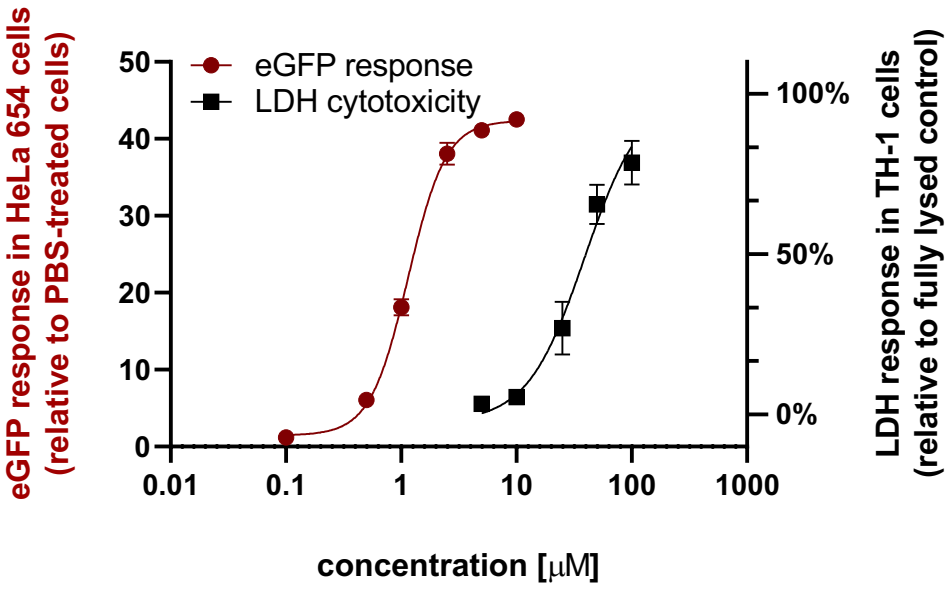
PMO-CXD1A

Peptide sequence: 5azido-Gly-Lys-Thr-Ser-Iph-Gly-Arg-Lys-Gln-Lys-Thr-Ser-Iph-Gly-Arg-Lys-Trp-Lys-Lys



PMO-CXD1B

Peptide sequence: 5azido-Gly-5azido-Gly-Lys-Gln-Lys-Thr-Ser-Iph-Gly-Arg-Lys-Trp-Lys-Gln-Lys-Thr-Ser-Iph-Gly-Arg-Lys-Trp-Lys-Lys

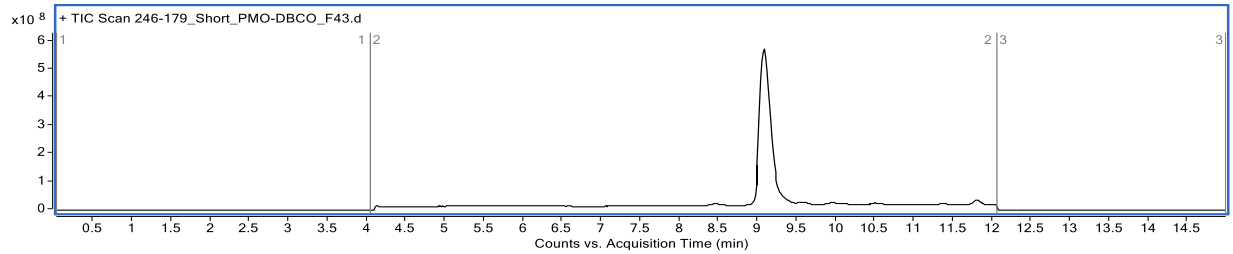

eGFP response in HeLa 654 cells
(relative to PBS-treated cells)
eGFP response
LDH cytotoxicity
50
40
30
20
10
0
100%
50%
0%
0.01
0.1
1
10
100
1000
concentration [μM]
LDH response in TH-1 cells
(relative to fully lysed control)

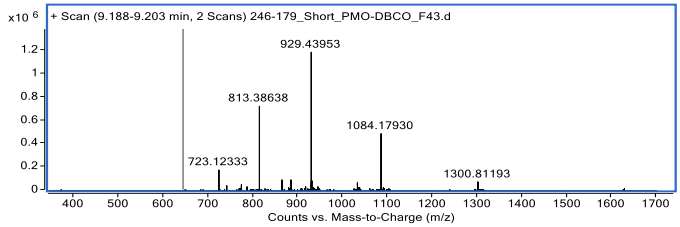
3.7 Appendix II: Characterization of compounds

PMO-DBCO

Mass Expected: 6500.0

Mass Observed: 6499.9



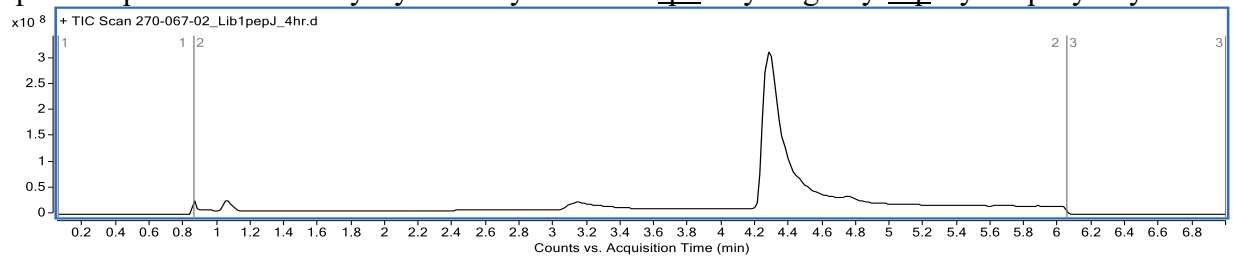


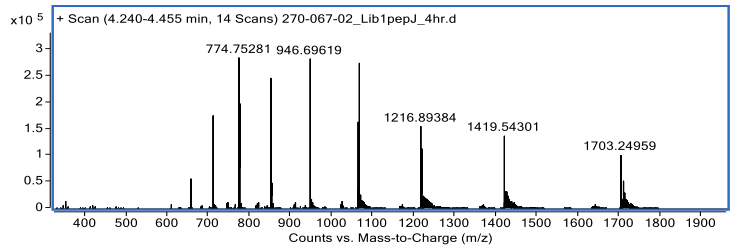
PMO-CXP1

Mass Expected: 8512.10

Mass Observed: 8512.06

Peptide sequence: 5azido-Gly-Lys-Gln-Lys-Thr-Ser-Iph-Gly-Arg-Gly-Pip-Lys-Trp-Lys-Lys



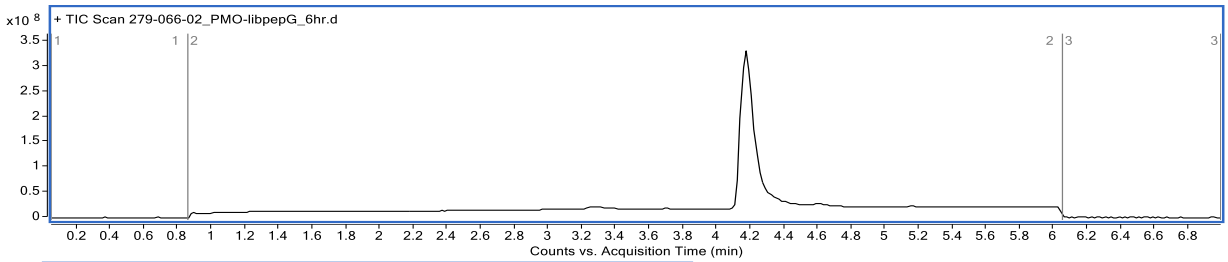


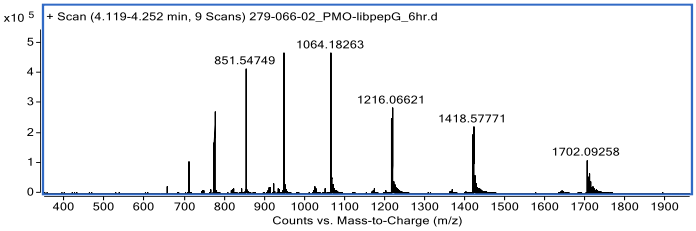
PMO-CXP2

Mass Expected: 8506.32

Mass Observed: 8506.35

Peptide sequence: 5azido-Gly-Asn-Phe-Lys-Gln-Nap-His-Ala-Gly-Dab-Arg-Lys-Trp-Lys-Lys



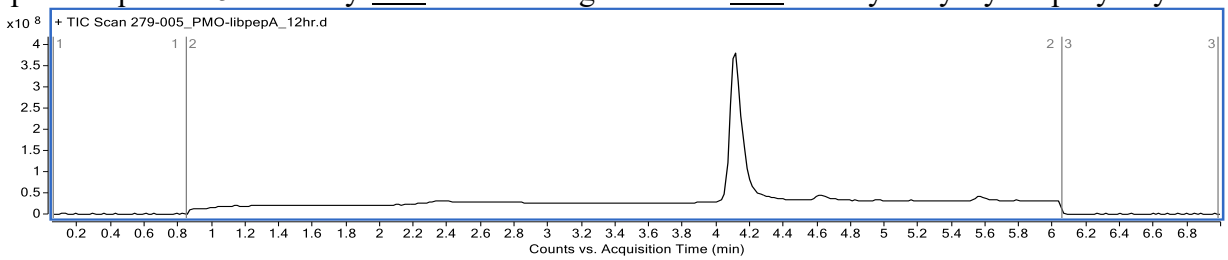


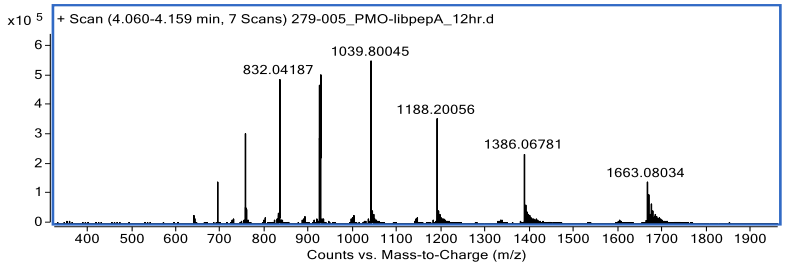
PMO-CXP3

Mass Expected: 8311.09

Mass Observed: 8311.25

Peptide sequence: 5azido-Gly-Dab-Val-Ala-Arg-Asn-Asn-Dab-Thr-Lys-Gly-Lys-Trp-Lys-Lys



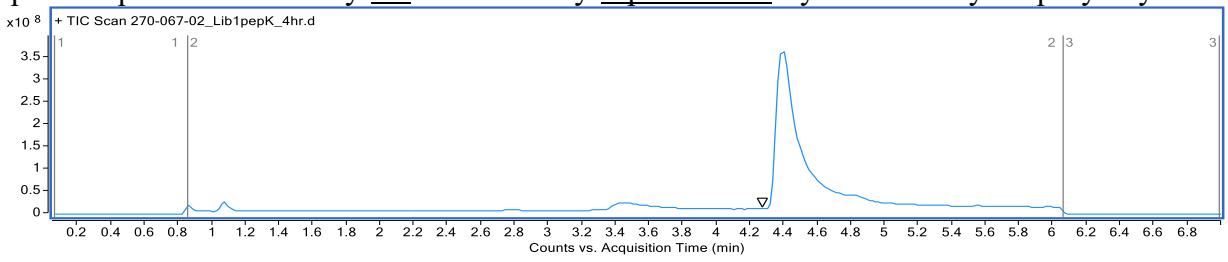


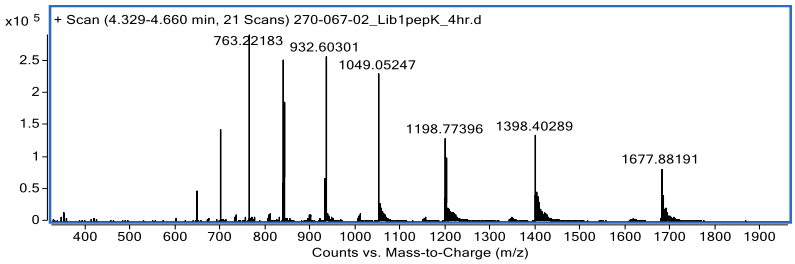
PMO-CXP4

Mass Expected: 8385.29

Mass Observed: 8385.33

Peptide sequence: 5azido-Gly-Cit-Met-Phe-Gly-Pip-Dab-Dab-Lys-Ala-Pro-Lys-Trp-Lys-Lys



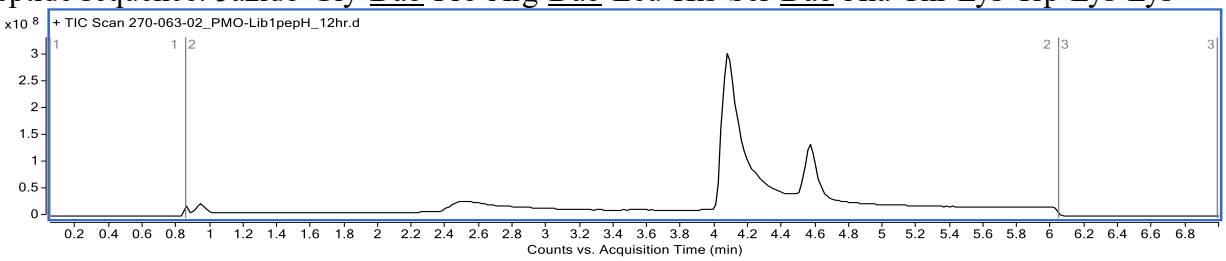


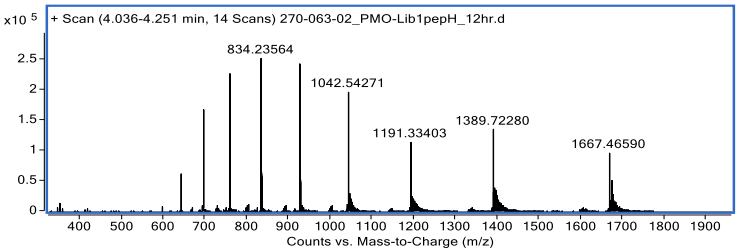
PMO-CXP5

Mass Expected: 8333.13

Mass Observed: 8333.10

Peptide sequence: 5azido-Gly-Dab-Pro-Arg-Dab-Leu-His-Ser-Dab-Ala-Thr-Lys-Trp-Lys-Lys



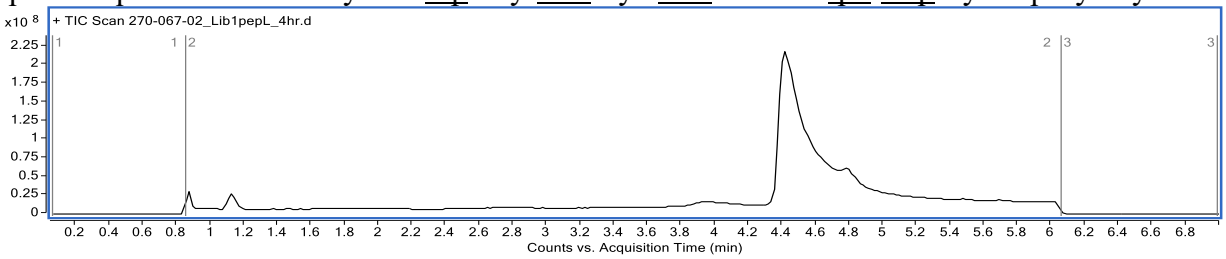


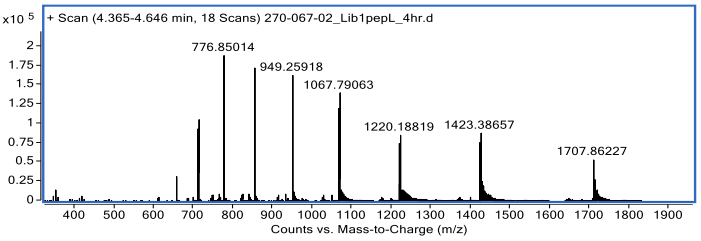
PMO-CXP6

Mass Expected: 8535.16

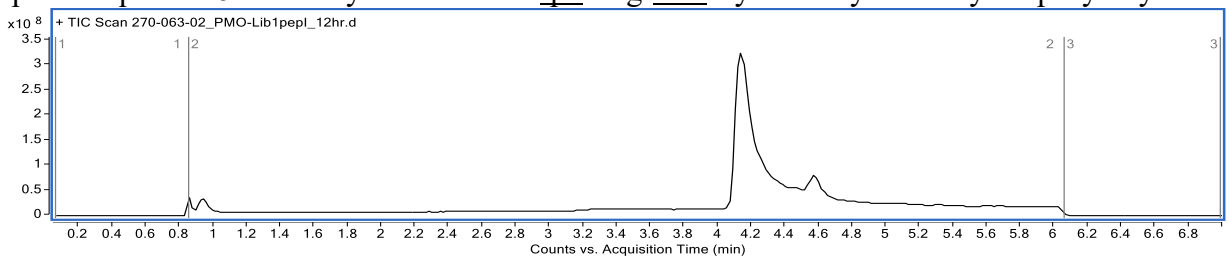
Mass Observed: 8535.16

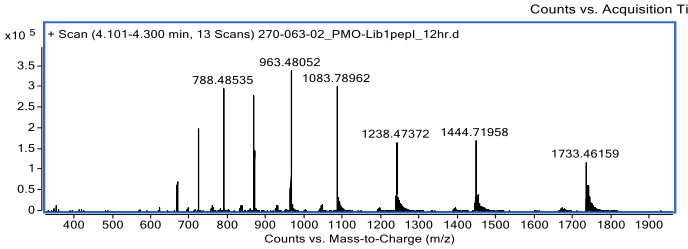
Peptide sequence: 5azido-Gly-Pro-Pip-Gly-Dab-Lys-Dab-Ser-Val-Iph-Nap-Lys-Trp-Lys-Lys



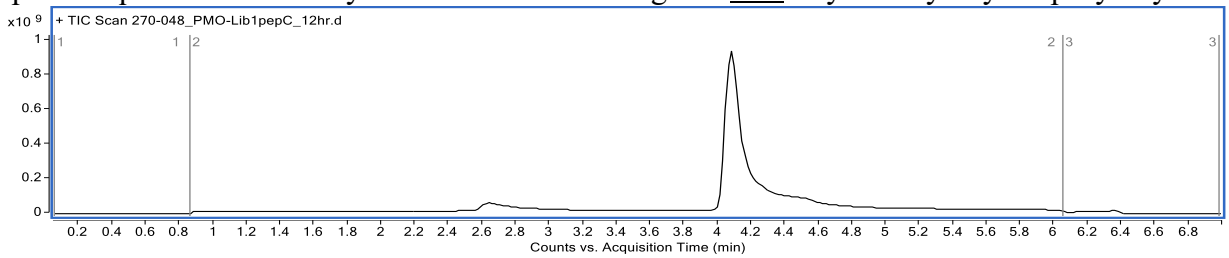


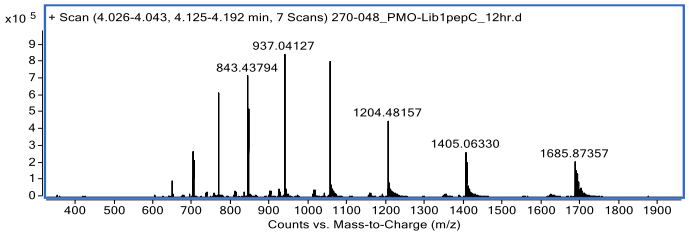
PMO-CXP7
Mass Expected: 8663.27
Mass Observed: 8663.24
Peptide sequence: 5azido-Gly-Val-Gln-His-Iph-Arg-Dab-Lys-Glu-Lys-Asn-Lys-Trp-Lys-Lys



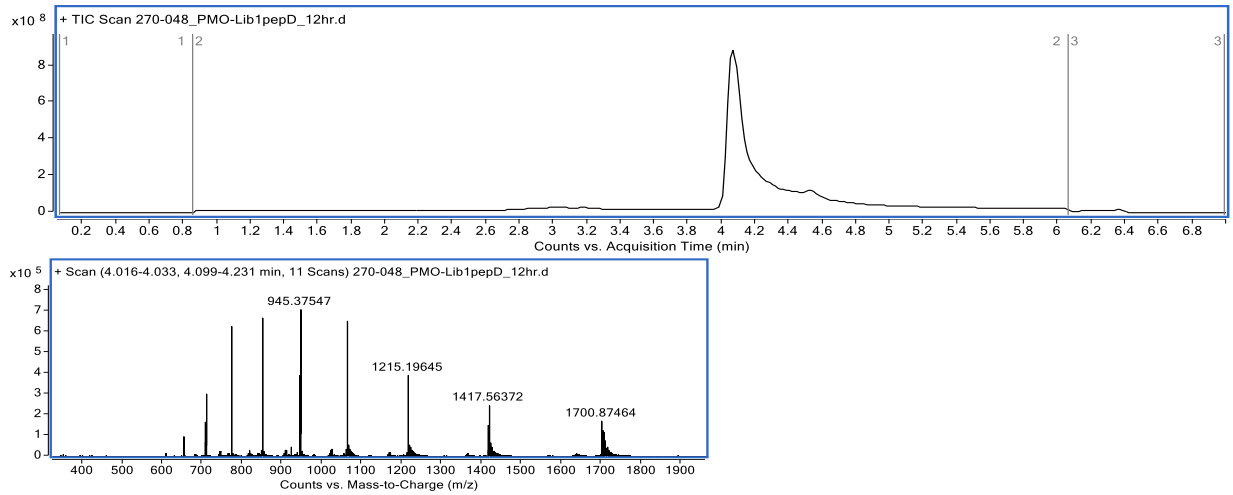


PMO-CXP8
Mass Expected: 8425.17
Mass Observed: 8425.31
Peptide sequence: 5azido-Gly-Asn-Phe-Asn-Asn-Arg-Ser-Dab-Gly-His-Lys-Lys-Trp-Lys-Lys





PMO-CXP9
Mass Expected: 8500.23
Mass Observed: 8500.38
Peptide sequence: 5azido-Gly-Ala-Arg-Ser-Glu-Gln-His-His-Nap-Ser-Dab-Lys-Trp-Lys-Lys

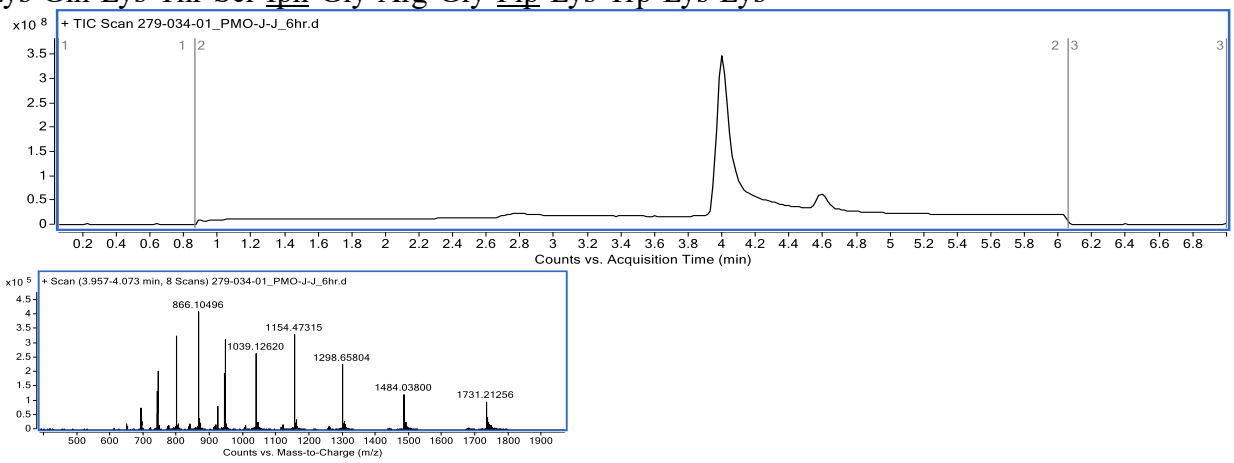


PMO-CXP1-CXP1

Mass Expected: 10382.07

Mass Observed: 10382.09

Peptide sequence: 5azido-Gly-Lys-Gln-Lys-Thr-Ser-Iph-Gly-Arg-Gly-Pip-Lys-Trp-Lys-Lys- Gly-Lys-Gln-Lys-Thr-Ser-Iph-Gly-Arg-Gly-Pip-Lys-Trp-Lys-Lys

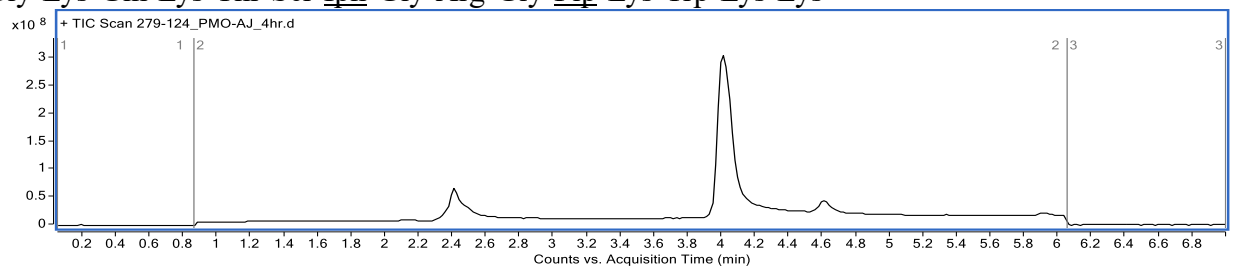


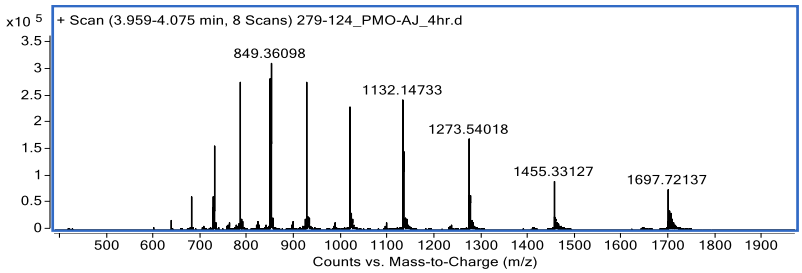
PMO-CXP3-CXP1

Mass Expected: 10181.05

Mass Observed: 10181.02

Peptide sequence: 5azido-Gly-Dab-Val-Ala-Arg-Asn-Asn-Dab-Thr-Lys-Gly-Lys-Trp-Lys-Lys-Gly-Lys-Gln-Lys-Thr-Ser-Iph-Gly-Arg-Gly-Pip-Lys-Trp-Lys-Lys



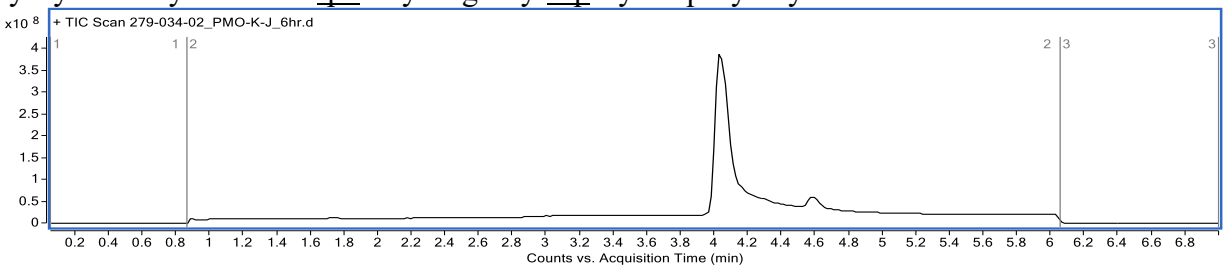


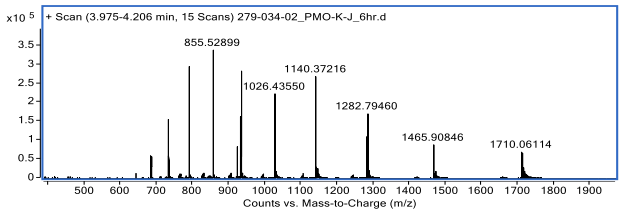
PMO-CXP4-CXP1

Mass Expected: 10255.25

Mass Observed: 10255.28

Peptide sequence: 5azido-Gly-Cit-Met-Phe-Gly-Pip-Dab-Dab-Lys-Ala-Pro-Lys-Trp-Lys-Lys-Gly-Lys-Gln-Lys-Thr-Ser-Iph-Gly-Arg-Gly-Pip-Lys-Trp-Lys-Lys



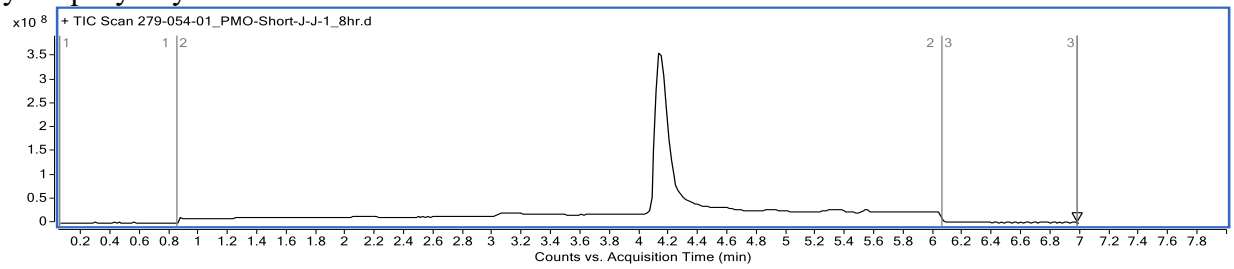


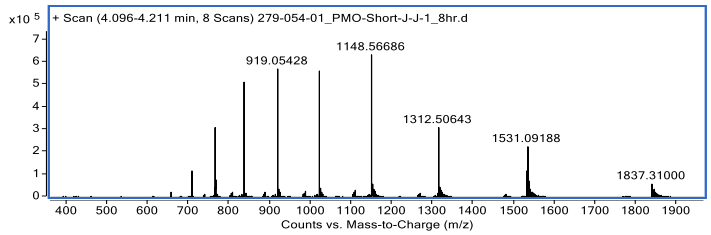
PMO-SA-CXP1

Mass Expected: 9181.63

Mass Observed: 9181.66

Peptide sequence: 5azido-Gly-Iph-Gly-Arg-Gly-Pip-Lys-Gln-Lys-Thr-Ser-Iph-Gly-Arg-Gly-Pip-Lys-Trp-Lys-Lys



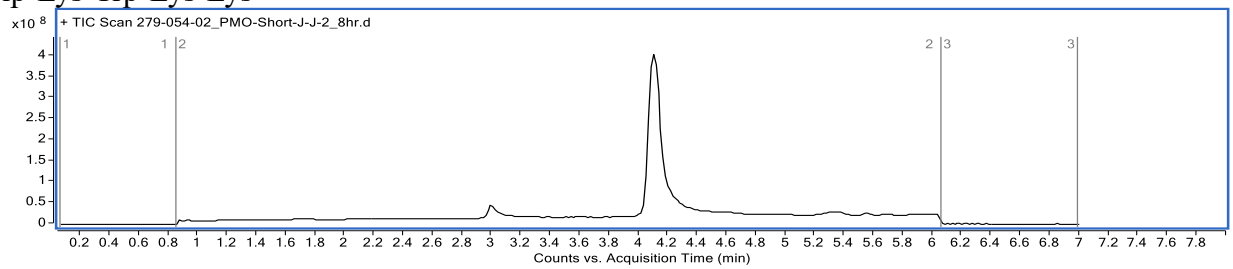


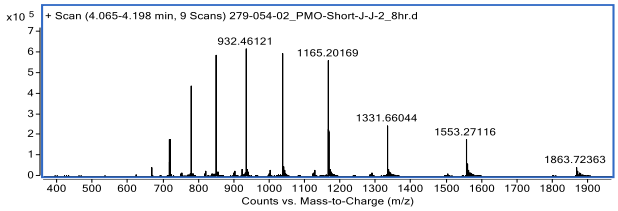
PMO-SB-CXP1

Mass Expected: 9314.78

Mass Observed: 9314.82

Peptide sequence: 5azido-Gly-Lys-Thr-Ser-Iph-Gly-Arg-Lys-Gln-Lys-Thr-Ser-Iph-Gly-Arg-Gly-Pip-Lys-Trp-Lys-Lys



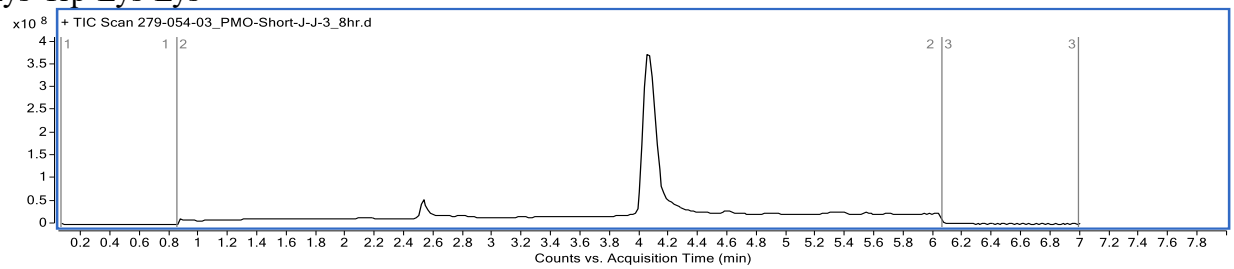


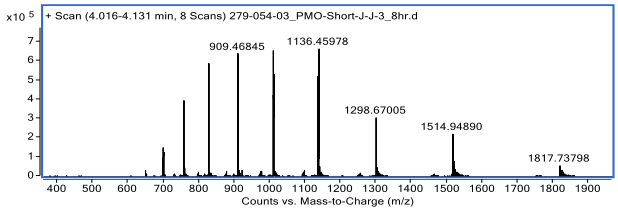
PMO-SC-CXP1

Mass Expected: 9084.78

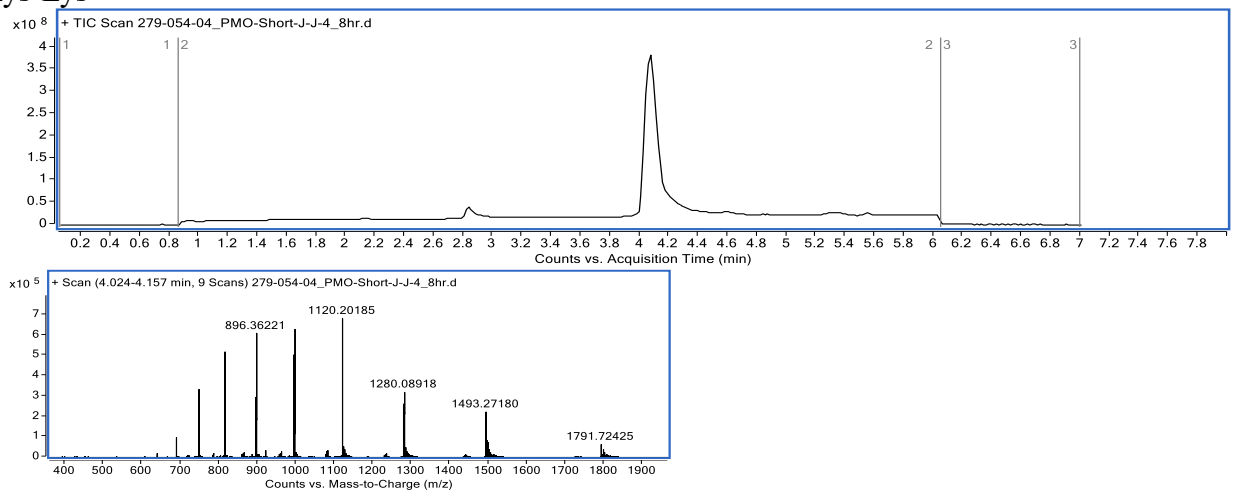
Mass Observed: 9084.80

Peptide sequence: 5azido-Gly-Lys-Gln-Lys-Thr-Ser-Lys-Gln-Lys-Thr-Ser-Iph-Gly-Arg-Gly-Pip-Lys-Trp-Lys-Lys

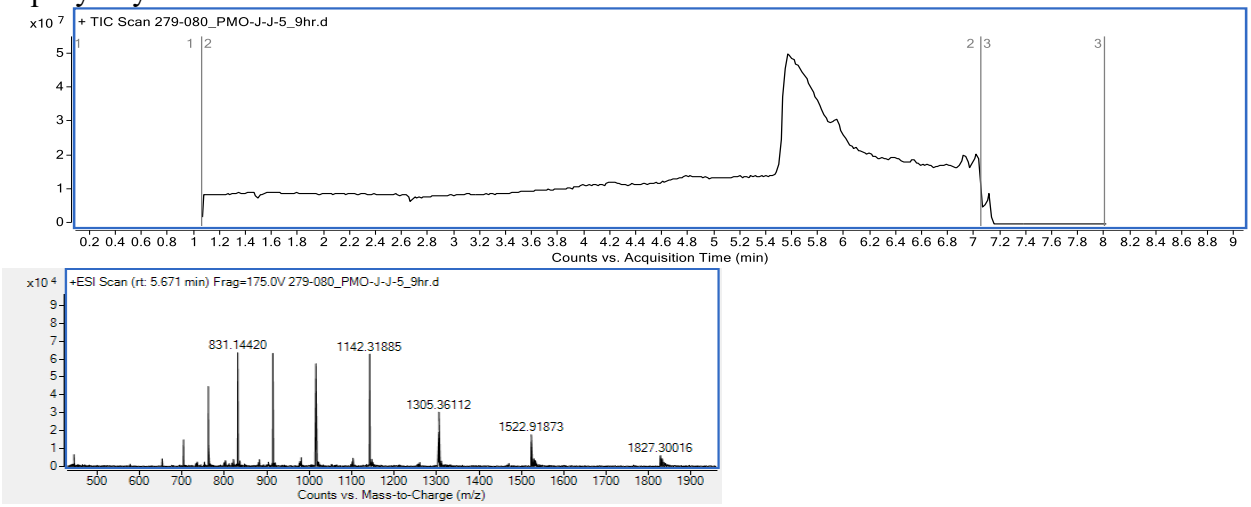




PMO-SD-CXP1
Mass Expected: 8954.67
Mass Observed: 8954.69
Peptide sequence: 5azido-Gly-Lys-Trp-Lys-Lys-Gln-Lys-Thr-Ser-Iph-Gly-Arg-Gly-Pip-Lys-Trp-Lys-Lys

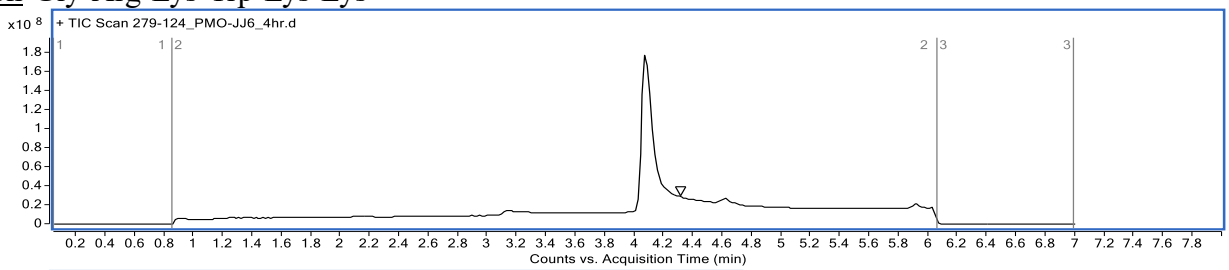


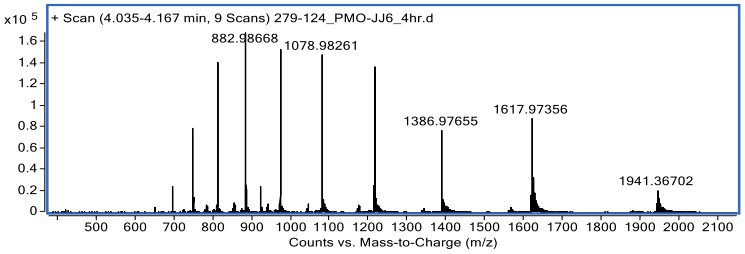
PMO-CXD1A
Mass Expected: 9131.58
Mass Observed: 9131.67
Peptide sequence: 5azido-Gly-Lys-Thr-Ser-Iph-Gly-Arg-Lys-Gln-Lys-Thr-Ser-Iph-Gly-Arg-Lys-Trp-Lys-Lys



PMO-CXD1B
Mass Expected: 9702.26
Mass Observed: 9702.12

Peptide sequence: 5azido-Gly-Lys-Gln-Lys-Thr-Ser-Iph-Gly-Arg-Lys-Trp-Lys-Gln-Lys-Thr-Ser-Iph-Gly-Arg-Lys-Trp-Lys-Lys





3.8 References

(1) Falzarano, M. S.; Scotton, C.; Passarelli, C.; Ferlini, A. Duchenne Muscular Dystrophy: From Diagnosis to Therapy. *Molecules* **2015**, *20* (10), 18168–18184. https://doi.org/10.3390/molecules201018168.

(2) Findlay, A. R.; Wein, N.; Kaminoh, Y.; Taylor, L. E.; Dunn, D. M.; Mendell, J. R.; King, W. M.; Pestronk, A.; Florence, J. M.; Mathews, K. D.; et al. Clinical Phenotypes as Predictors of the Outcome of Skipping around DMD Exon 45. *Ann. Neurol.* **2015**, *77* (4), 668–674. https://doi.org/10.1002/ana.24365.

(3) Charleston, J. S.; Schnell, F. J.; Dworzak, J.; Donoghue, C.; Lewis, S.; Chen, L.; Young, G. D.; Milici, A. J.; Voss, J.; DeAlwis, U.; et al. Eteplirsen Treatment for Duchenne Muscular Dystrophy: Exon Skipping and Dystrophin Production. *Neurology* **2018**, *90* (24), e2146–e2154. https://doi.org/10.1212/WNL.0000000000005680.

(4) Moulton, H. M.; Moulton, J. D. Morpholinos and Their Peptide Conjugates: Therapeutic Promise and Challenge for Duchenne Muscular Dystrophy. *Biochim Biophys Acta*. **2010**, *1798* (12), 2296-2303. https://doi.org/10.1016/j.bbamem.2010.02.012.

(5) McClorey, G.; Banerjee, S. Cell-Penetrating Peptides to Enhance Delivery of Oligonucleotide-Based Therapeutics. *Biomedicines* **2018**, *6* (2), 51. https://doi.org/10.3390/biomedicines6020051.

(6) Blain, A. M.; Greally, E.; Mcclorey, G.; Manzano, R.; Betts, C. A.; Godfrey, C.; O'donovan, L.; Coursindel, T.; Gait, M. J.; Wood, M. J.; et al. Peptide-Conjugated Phosphodiamidate Oligomer-Mediated Exon Skipping Has Benefits for Cardiac Function in Mdx and Cmah-/-Mdx Mouse Models of Duchenne Muscular Dystrophy. *PLoS One*. **2018**, *13*(6), e0198897. https://doi.org/10.1371/journal.pone.0198897.

(7) Agrawal, P.; Bhalla, S.; Usmani, S. S.; Singh, S.; Chaudhary, K.; Raghava, G. P. S.; Gautam, A. CPPsite 2.0: A Repository of Experimentally Validated Cell-Penetrating Peptides. *Nucleic Acids Res*. **2016,** *44* (D1), D1098–D1103. https://doi.org/10.1093/nar/gkv1266.

(8) Hymel, H. C.; Rahnama, A.; Sanchez, O. M.; Liu, D.; Gauthier, T. J.; Melvin, A. T. How Cargo Identity Alters the Uptake of Cell-Penetrating Peptide (CPP)/Cargo Complexes: A Study on the Effect of Net Cargo Charge and Length. *Cells.* **2022**, *11* (7), 1195

(9) Schissel, C. K.*; Farquhar, C. E*; Malmberg, A. B.; Loas, A.; Pentelute, B. L. In-cell penetration selection—mass spectrometry produces noncanonical peptides for antisense delivery. *ACS Chem Bio.* **2023**, *18* (3), 615-628.

(10) Verdurmen, W.; Thanos, M.; Ruttekolk, I.; Gulbins, E.; Brock, R. Cationic cell-penetrating peptides induce ceramide formation via acid sphingomyelinase: implications for uptake. *J. Control Release.* **2010,** *147* (2), 171-179.

(11) Ramaker, K.; Henkel, M.; Krause, T.; Röckendorf, N.; Frey, A. Cell Penetrating Peptides: A Comparative Transport Analysis for 474 Sequence Motifs. *Drug Deliv*. **2018**, *25* (1), 928–937. https://doi.org/10.1080/10717544.2018.1458921.

(12) Lönn, P.; Dowdy, S. F. Cationic PTD/CPP-Mediated Macromolecular Delivery: Charging into the Cell. *Expert Opin. Drug Deliv*. **2015**, *12* (10), 1627–1636. https://doi.org/10.1517/17425247.2015.1046431.

(13) Jobin, M.-L.; Blanchet, M.; Henry, S.; Chaignepain, S.; Manigand, C.; Castano, S.; Lecomte, S.; Burlina, F.; Sagan, S.; Alves, I. D. The Role of Tryptophans on the Cellular Uptake and Membrane Interaction of Arginine-Rich Cell Penetrating Peptides. *Biochim. Biophys. Acta - Biomembr*. **2015,** *1848* (2), 593–602.

https://doi.org/10.1016/J.BBAMEM.2014.11.013.
(14) Bechara, C.; Pallerla, M.; Zaltsman, Y.; Burlina, F.; Alves, I. D.; Lequin, O.; Sagan, S. Tryptophan within Basic Peptide Sequences Triggers Glycosaminoglycan-Dependent Endocytosis. *FASEB J.* **2013**, *27* (2), 738–749. https://doi.org/10.1096/fj.12-216176.
(15) Derossi, D.; Joliot, A. H.; Chassaing, G.; Prochiantz, A. The Third Helix of the Antennapedia Homeodomain Translocates through Biological Membranes. *J. Biol. Chem.* **1994**, *269* (14), 10444–10450.
(16) Quartararo, A.J., Gates, Z.P., Somsen, B.A. *et al.* Ultra-large chemical libraries for the discovery of high-affinity peptide binders. *Nat. Commun.* **2020**, *11*, 3183. https://doi.org/10.1038/s41467-020-16920-3
(17) Fadzen, C. M.; Holden, R. L.; Wolfe, J. M.; Choo, Z. N.; Schissel, C. K.; Yao, M.; Hanson, G. J.; Pentelute, B. L. Chimeras of Cell-Penetrating Peptides Demonstrate Synergistic Improvement in Antisense Efficacy. *Biochemistry* **2019,** *58* (38), 3980–3989. https://doi.org/10.1021/acs.biochem.9b00413.
(18) Li, Q.;Xu, M.;Cui, Y.;Huang, C.; Sun, M. Arginine-rich membrane-permeable peptides are seriously toxic. *Pharmacol. Res. Perspect.* **2017**, *5* (5), e00334.
(19) Wolfe, J. M.; Fadzen, C. M.; Choo, Z. N.; Holden, R. L.; Yao, M.; Hanson, G. J.; Pentelute, B. L. Machine Learning to Predict Cell-Penetrating Peptides for Antisense Delivery. *ACS Cent. Sci*. **2018**, *4* (4), 512–520. https://doi.org/10.1021/acscentsci.8b00098.
(20) Schissel, C. K.*; Mohapatra, S.*; Wolfe, J. M.; Fadzen, C. M.; Bellovoda, K.; Wu, C.-L.; Wood, J. A.; Malmberg, A. B.; Loas, A.; Gómez-Bombarelli, R.; Pentelute, B. L. Deep Learning to Design Nuclear-Targeting Abiotic Miniproteins. Nat. Chem. 2021, 13, 992-1000.
(21) Lopez-Vidal, E. M.*; Schissel, C. K.*; Mohapatra, S.; Bellovoda, K.; Wu, C.-L.; Wood, J. A.; Malmberg, A. B.; Loas, A.; Gómez-Bombarelli, R.; Pentelute, B. L. Discovery of a Short Nuclear Targeting peptide for PMO delivery using Machine Learning. JACS Au. 2021, 1, 2009-2020.

Chapter 4. A Tumor-Homing Peptide Platform Enhances Drug Solubility, Improves Blood–Brain Barrier Permeability and Targets Glioblastoma

The work presented in this chapter has been reproduced from the following manuscript:

Cho, C.-F.; Farquhar, C. E.; Fadzen, C. M.; Scott, B.; Zhuang, P.; von Spreckelsen, N.; Loas, A.; Hartrampf, N.; Pentelute, B. L.; Lawler, S. E. A Tumor-Homing Peptide Platform Enhances Drug Solubility, Improves Blood-Brain Barrier Permeability and Targets Glioblastoma. *Cancers (Basel).* **2022**, *14* (9), 2207.

4.1 Introduction

High-grade gliomas, including grade III and grade IV glioblastoma (GBM), remain among the most difficult cancers to treat. GBM is the most common and deadliest primary malignant brain tumor, with a 5-year survival rate of only 5%.[1] Current stand-ard-of-care which consists of surgery followed by chemotherapy and/or radiotherapy is non-curative, with almost all patients having recurrence with more aggressive tumors.[2] There is no standard-of-care for recurrent GBM. Malignant GBM cells invade into sur-rounding brain tissues, often escaping surgical resection and radiotherapy. Additionally, although a defining feature of GBM is abnormal angiogenesis leading to disorganized and leaky blood vessels, a significant subgroup of GBM cells are protected from chemotherapeutics by the blood–brain barrier (BBB).[3,4]

Most experimental GBM drugs from clinical trials over the past few decades have not made a major impact in limiting disease progression, and patients continue to suffer significant morbidity and death. The majority of therapies often fail clinically due to poor drug solubility,[5] lack of tumor selectivity leading to undesired side effects [6], poor permeability across the blood–brain barrier (BBB), as well as extensive intra- and inter-tumor heterogeneity. The development of precision medicines with improved solubility and bioavailability, BBB penetrance, and the ability to target cancerous GBM cells selectively with minimal effect on healthy tissues would be a valuable strategy towards enhancing therapeutic efficacy.

While GBM is well-known for cellular and genetic heterogeneity,[7] the tumor extracellular matrix (ECM) has less spatial variability, making it an attractive target for therapeutic intervention strategies compared to specific membrane receptors that are only expressed on select tumor cell populations.[8] An ECM glycoprotein expressed exclu-sively in the central nervous system (CNS) called brevican (Bcan) is upregulated in GBM,[9-11] and is linked to increased tumor invasion and aggressiveness.[12] The degly-cosylated isoform, called dg-Bcan, is only found in human high-grade glioma tissue samples,[10,13] making it an attractive GBM-specific marker for the development of novel therapeutic targeting strategies. Previously, we have screened a combinatorial D-amino acid peptide library to identify an octameric dg-Bcan-Targeting Peptide, called BTP-7, that is stable in serum, binds dg-Bcan specifically, crosses the BBB, and prefer-entially homes to intracranial

GBM xenografts in mice established using patient-derived GBM stem cells (GSC).[13] Positron emission tomography (PET) imaging using [18F]-radiolabeled BTP-7 reveals higher tumor-to-brain uptake ratio (sustained over several hours) in comparison to values reported from similar studies evaluating clinical radiotracers such as [18F]FDG and [18F]FET.[13] These studies highlight BTP-7 as a promising tumor-specific agent for delivery of functional cargoes to GBM, warranting further investigation and development.

In this manuscript, we explore the potential of BTP-7 for the development of a targeting platform for GBM treatment. Camptothecin (CPT) is a potent topoisomerase I inhibitor that has shown promising results in preclinical models but failed to translate well in the clinic, mostly due to its insoluble nature.[14,15] Here, we show that chemical functionalization of CPT with BTP-7 (a positively charged peptide under physiological conditions) yields a peptide–drug conjugate (BTP-7-CPT) that is water-soluble and has BBB-penetrating properties. We demonstrate that BTP-7-CPT exhibits potency against patient-derived GBM stem cells in vitro, increases drug delivery to tumors in an intra-cranial patient-derived xenograft (PDX) mouse model of GBM resulting in enhanced tumor toxicity compared to healthy brain tissues, and ultimately prolongs survival in animals.

4.2 Results

4.2.1 Conjugation of CPT with BTP-7 enhances drug solubility

BTP-7 has a high isoelectric point of 10 and is soluble in aqueous solution.[13] Previously, we have observed that conjugation of Cy5.5, a hydrophobic fluorescent dye to BTP-7, enhances its water solubility.[13] As a proof-of-concept that BTP-7 functionalization improves the solubility of hydrophobic drugs, we examine here the physiochemical effect of conjugating BTP-7 to CPT, an insoluble chemotherapeutic.[14,15]

BTP-7 containing an aminohexanoic acid linker (X) and a cysteine at the C-terminus was attached to a thiol on CPT via a previously reported disulfide linker[17] (workflow illustrated in **Fig. 4.1a**). A detailed description of the synthetic route is provided as Supplementary Information. Within the reducing environment of the cellular cytosol, the disulfide linker is reduced to release the active CPT drug metabolite that is toxic to cells (**Fig. 4.1b**).[21] Using the same synthetic route, a scrambled version of the peptide was conjugated to CPT (Scr-7-CPT) as a control conjugate to assess for specific binding of the BTP-7 conjugate.

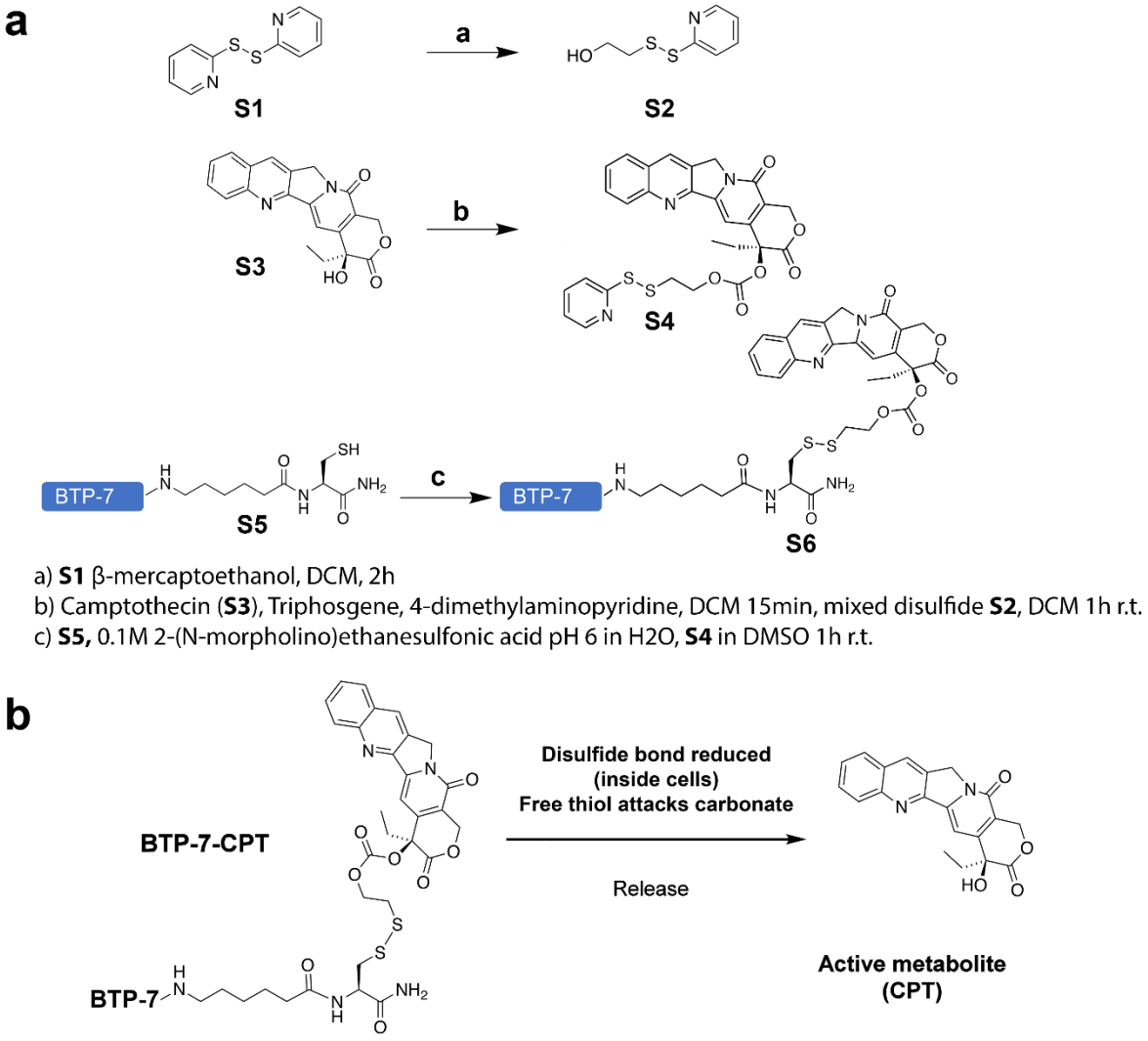


Figure 4.1 Synthesis and cleavage of the BTP7-CPT drug construct.
(a) Chemical synthesis workflow showing the conjugation of BTP-7 to camptothecin (CPT) and (b) cleavage mechanism to release active CPT drug upon cell internalization

To test the solubility of each compound in aqueous solution, a 10 mM stock of each compound in pure DMSO was diluted in saline solution (0.9% sodium chloride in water) to a final concentration of either 100 μM or 1.3 mM. The samples were incubated at room temperature for 5 min and centrifuged to remove any precipitate. The supernatant was analyzed by liquid chromatography–mass spectrometry (LC-MS). To quantify solubility, the extracted ion count from samples diluted in saline was normalized to samples diluted in pure DMSO.

As expected, the native unmodified CPT drug displayed low solubility, with 30% of soluble compound present in saline at 100 µM relative to in pure DMSO (**Fig. 4.2a**). CPT was practically undetectable in saline solution at a higher concentration of 1.3 mM (**Fig. 4.2b**). Conjugation of CPT to BTP-7 resulted in an increase in drug solubility at both concentrations (**Fig. 4.2**). Indeed, we demonstrated that BTP-7-CPT was more soluble than irinotecan (CPT-11) (**Fig. 4.2a**), a known water-soluble derivative of CPT that has been approved by the FDA for the treatment of advanced colorectal and pancreatic cancer.[22] The scramble Scr-7-CPT conjugate was also soluble in aqueous solution, similar to that of BTP-7-CPT (**Fig. 4.2**). This result indicates that the enhanced solubility is attributed to the intrinsic amino acid composition of the peptide, and not its specific sequence.

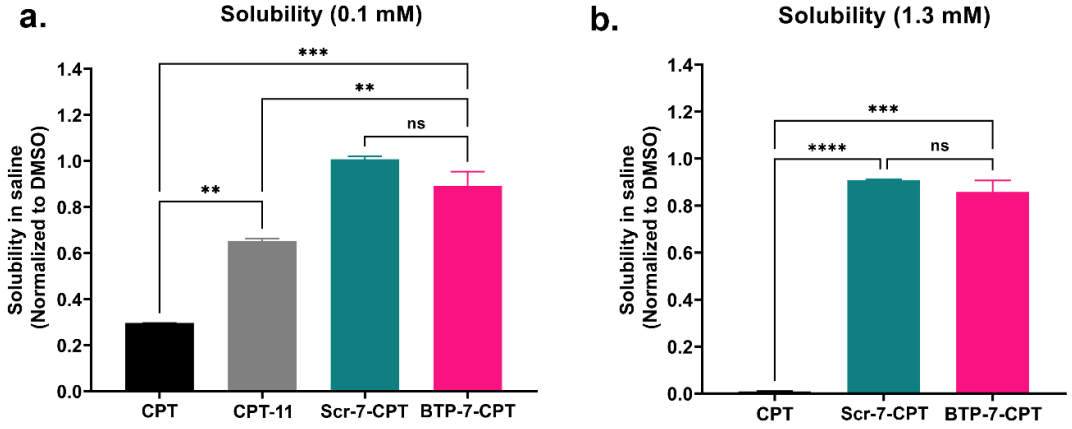


Figure 4.2 BTP-7 conjugation enhances drug solubility.
(a) BTP-7 (and Scr-7) conjugation to CPT enhances drug solubility at 0.1 mM (low conc.) and (b) 1.3 mM (high conc.). CPT, CPT-11, BTP-7-CPT, and Scr-7-CPT stocks (in DMSO) were diluted in either DMSO or saline. Samples were centrifuged to pellet insoluble compound, and the resulting supernatant was analyzed by LC-MS. Solubility is measured by comparing the extracted ion chromatogram of the sample di-luted in saline vs. in DMSO. Statistical significance was determined using a two-way ANOVA and Tukey's multiple comparisons test (** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$, ns = not statisti-cally significant).

4.2.2 BTP-7-CPT binds dg-Bcan Protein

Next, we investigated BTP-7-CPT drug specificity and potency. Binding kinetics analyses using the Octet RED platform showed that BTP-7-CPT bound to purified re-combinant dg-Bcan protein (dissociation constant, KD = 6.1 µM) (**Fig. 3.3b**). This result is consistent with our previous observations which showed that BTP-7 was able to retain binding to dg-Bcan protein after being chemically linked to other molecules (i.e., a fluorophore, or radioisotope).[13] As expected, unmodified CPT did not exhibit any af-finity to recombinant dg-Bcan (**Fig. 3.3b**). Altogether, these

findings further highlight BTP-7 as a promising agent for therapeutic functionalization to enable dg-Bcan targeting.

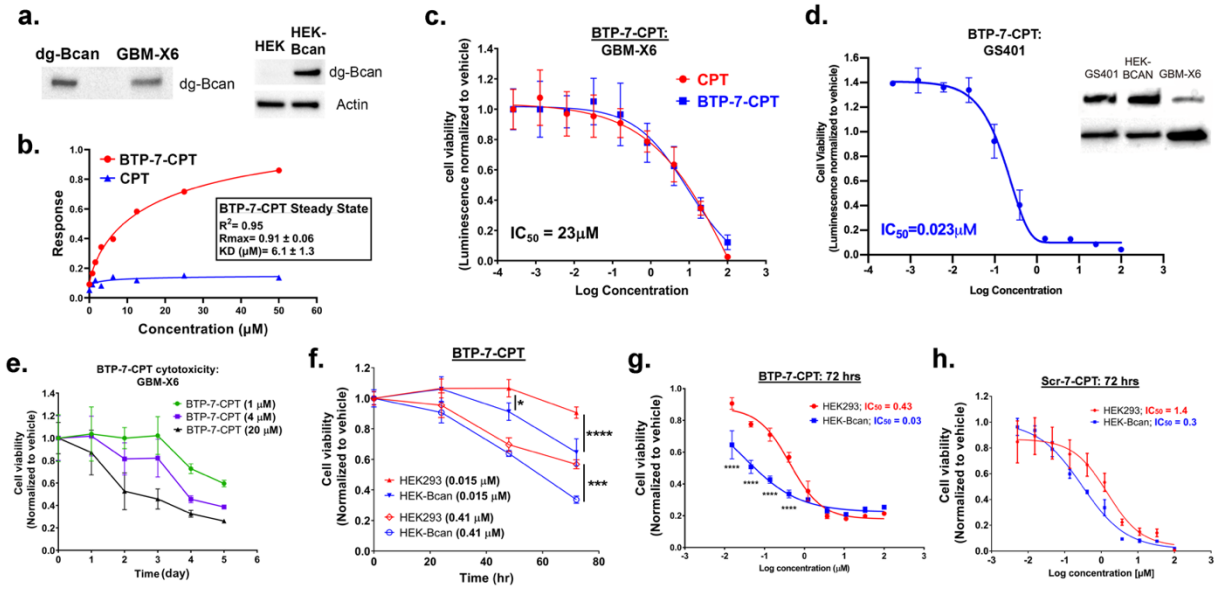


Figure 4.3 Brevican binding and cytotoxicity of BTP-7-CPT.
(a) Western blot showing expression of dg-Bcan in patient-derived GBM-X6 cells, as well as in Bcan-overexpressing HEK cells (HEK-Bcan). (b) Steady-state binding kinetics analysis of BTP-7-CPT (red) and unmodified CPT (blue) to recombinant dg-Bcan protein using the Octet (ForteBio) platform. KD (dissociation constant) and Rmax (maximal response in response unit) were calculated using a non-linear regression (one site specific binding) fit in the Graphpad Prism software. (c) Luminescent cell viability (CellTiter-Glo) of GBM-X6 cells treated with BTP-7-CPT (nwells = 3) for 72 h. (d) Luminescent cell viability (CellTiter-Glo) of GS401 recurrent GBM cells treated with BTP-7-CPT (nwells = 3) for 72 h. All IC50 values were measured through the non-linear 'log(inhibitor) vs. response—variable slope (four parameters)' fit. Western blot (right) shows higher expression of dg-Bcan in GS401 cells than in GBM-X6. (e) CellTiter-Glo assay of GBM-X6 cells treated with BTP-7-CPT at 1, 4, or 20 μm (nwells = 3) over 5 days. (f) CellTiter-Glo assay of HEK293 (red) or Bcan-overexpressing HEK (HEK-Bcan) cells (blue) treated with BTP-7-CPT at 0.015, or 41 μm (nwells = 3) over 3 days. (g,h) CellTiter-Glo assay of HEK-Bcan cells (blue) and control HEK293 cells (red) in the presence of (g) BTP-7-CPT or (h) Scr-7-CPT after 72 h. Statistical significance was determined using a two-way ANOVA, Sidak's multiple comparisons test (* $p < 0.05$, *** $p < 0.001$, **** $p < 0.0001$).

4.2.3 Cytotoxicity of BTP-7-CPT

To evaluate BTP-7-CPT cytotoxicity in vitro, we used a well-characterized pa-tient-derived xenoline (GBM-X6 cells)[23,24] and a recurrent GSC model (GS401 cells).[13] Both GBM-X6 and GS401 cells express endogenous dg-Bcan (**Fig. 4.3a,d**). Using the CellTiter-Glo luminescent cell viability assay, we showed that BTP-7-CPT was toxic to GBM-X6 cells with potency comparable with unmodified CPT (IC50 = 23 μM) (**Fig 4.3c**). The recurrent GS401 cells, which have higher

dg-Bcan expression, were more sensitive to BTP-7-CPT cytotoxicity (IC50 = 0.023 μM) compared to GBM-X6 cells (**Fig. 4.3c,d**). These suggest that chemically linking CPT to BTP-7 did not alter the drug's potency, and that increased drug sensitivity is correlated with higher dg-Bcan expression, likely due to greater cellular uptake as previously demonstrated.[13] We further demonstrated that BTP-7-CPT was toxic to GBM-X6 cells in a dose- and time-dependent manner (**Fig. 4.3e**). To investigate BTP-7-CPT specificity, we measured drug response in human embryonic kidney (HEK293) cells that do not express endogenous brevican, as well as in engineered HEK293 cells with stable brevican overexpression (HEK-Bcan) using the same CellTiter-Glo assay. We found that both cell types responded to BTP-7-CPT in a dose- and time-dependent manner, although HEK-Bcan cells displayed higher sensitivity to BTP-7-CPT treatment compared to HEK293 cells (IC50, HEK = 0.43 μM vs. IC50, HEK-Bcan = 0.03 μM) (**Fig. 4.3f,g**). No significant difference was observed with the scramble drug con-jugate Scr-7-CPT (**Fig. 4.3h**). These data further support the correlation of dg-Bcan ex-pression in cells and drug sensitivity.

4.2.4 Analysis of BBB permeability

We previously showed that BTP-7 conjugated to a fluorescent dye (Cy5.5) was able to cross the BBB in vitro using human BBB organoids and in vivo in mice.[13] To examine the potential of BTP-7 for enhancing CPT delivery across the BBB, we employed the in vitro human BBB organoid model.[19,20] CPT and its derivatives share a heterocyclic functional group that can be excited at 360 nm with fluorescence emission in the 440–700 nm range.[25] Using confocal fluorescence microscopy, we measured the level of influx of each drug conjugate into the BBB organoids. BTP-7-CPT displayed the highest influx into the organoids, comparable to that of CPT-11 which is known to cross the BBB (**Fig. 4.4a,b**).[26] Consistent with previous findings,[13] the influx of the scramble Scr-7-CPT conjugate into the organoids was lower than that of BTP-7-CPT, but higher than unmodified CPT (**Fig. 4.4a,b**). These data suggest that the enhanced BBB permeability observed is likely attributable to the intrinsic amino acid composition and/or structural properties of the peptide, although further studies to examine their specific mechanism of penetration are warranted. Finally, we demonstrated that the BBB organoids were able to restrict the entry of dextran (4.4 kDa) in the presence of these cytotoxic compounds (at 10 μM), in-dicating that the integrity of the organoids' surface barrier remained intact for the du-ration of the experiment (**Fig. 4.4c**).

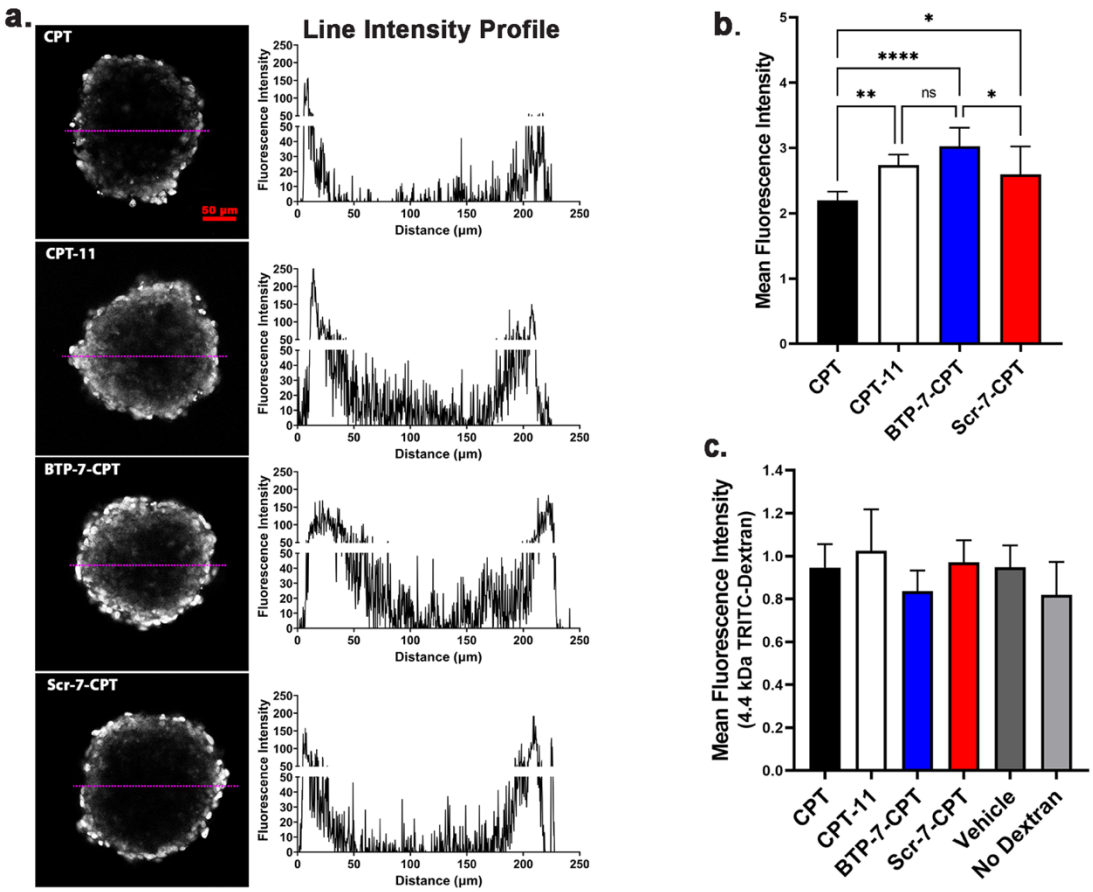


Figure 4.4 Conjugation of BTP-7 to CPT enhances permeability in human BBB organoids. (a) Fluorescence images showing the level of influx (penetration across the organoid surface) of BTP-7-CPT, Scr-7-CPT, CPT (camptothecin), and CPT-11 (irinotecan) (at 10 µM) in BBB organoids. Line profile (depicted by magenta line) showing the fluorescence (drug) level through the organoid. (b) Bar graph quantifying the mean fluorescence intensity of each drug conjugate, showing BTP-7-CPT had higher level of organoid influx compared to CPT and Scr-7-CPT (n = 8 organoids, t = 3 h). (c) Co-incubation of BBB organoids with each drug conjugate (from (b)) and 4.4 kDa TRITC-dextran (at 10 µg/mL) did not affect the level of dextran influx (all data were not significantly different compared to the vehicle control), indicating that the presence of the drug conjugate did not alter the organoid's surface integrity (n = 6–8 organoids, t = 3 h). Statistical significance was determined using a one-way ANOVA and Tukey's multiple comparisons test (* $p < 0.05$, ** $p < 0.01$, **** $p < 0.0001$).

4.2.5 Stability of peptide–drug conjugate in serum

Prior to animal testing, we performed ex vivo serum stability assays to measure the stability of BTP-7-CPT, Scr-7-CPT, and CPT in mouse serum. We have previously demonstrated that the native BTP-7 peptide (composed entirely of D-amino acids) linked to an aminohexanoic acid linker at the C-terminus (BTP-7-X) and the scrambled variant, Scr-7-X, were stable in mouse

serum.[13] We observed that the concentration of both BTP-7-CPT and Scr-7-CPT compound declined over the same time frame in 25% mouse serum compared to in PBS, with the scramble Scr-7-CPT drug conjugate showing a slightly longer half-life (half-life$_{mouse}$ = 2.7 h) compared to BTP-7-CPT (half-life$_{mouse}$ = 1.7 h) (Fig. 3.5a,b). We postulated that the source of BTP-7-CPT degradation in serum originated from the di-sulfide linkage connecting BTP-7 to CPT, which was primarily designed to be cleaved in the reducing environment of the cell to release the toxic CPT drug into the tumor. Indeed, we showed that while BTP-7-X was stable in 25% human serum, with approximately 40% of the peptide still intact after 24 h of incubation (half-life$_{human}$ = 4.4 h) (**Fig. 4.5c,d**),[13] the addition of a cysteine (Cys) residue to the C-terminus of BTP-7-X (BTP-7-X-Cys) resulted in a rapid decline of detectable compound (half-life$_{human}$ = 0.16 h) (**Fig. 4.5c,d**), likely due to the ability of Cys to form disulfides with serum proteins, which would result in the loss of BTP-7-X-Cys in the supernatant after the purification step to remove serum proteins. Similarly, we observed a relatively rapid decrease in BTP-7-CPT and Scr-7-CPT level in human serum (half-life$_{human}$ = 0.25 and 0.45 h, respectively) compared to BTP-7-X (**Fig. 4.5c,d**), suggesting compound instability that could be attributed to the likelihood of di-sulfide exchange occurring between the drug conjugates and serum protein.

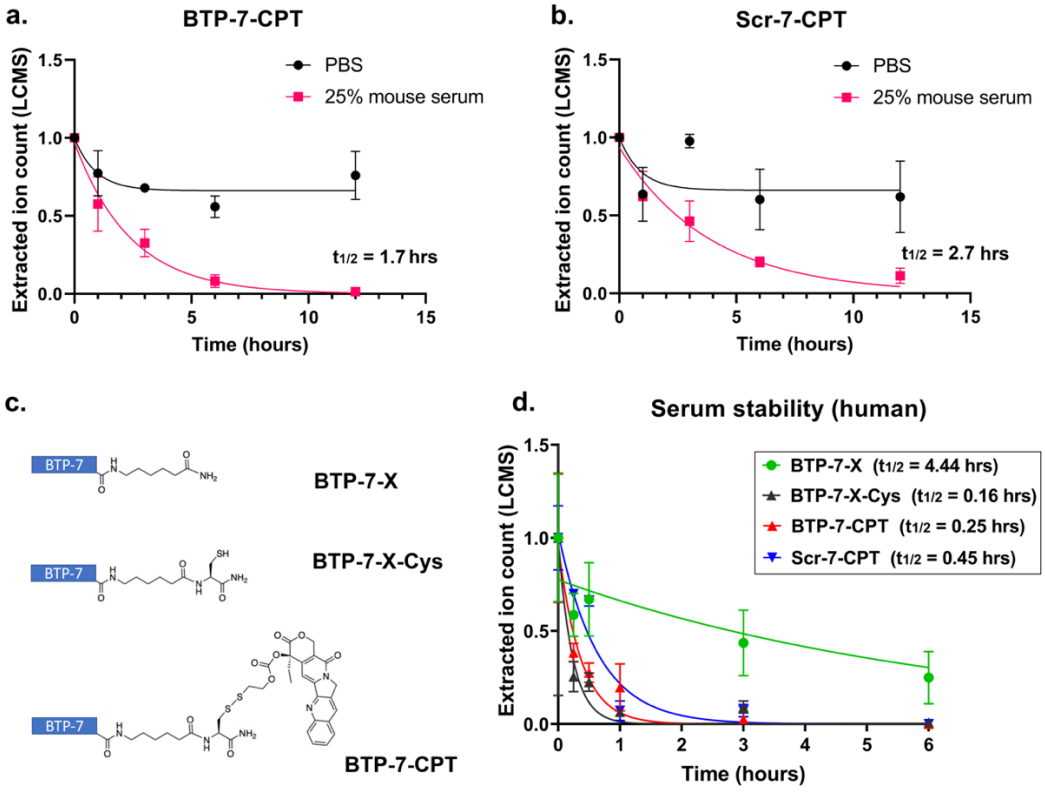


Figure 4.5 Stability of peptide–drug conjugates in serum.
(a,b) LC-MS analysis showing the level of (a) BTP-7-CPT and (b) Scr-7-CPT detected (extracted ion count) in PBS and in 25% human serum over 12 h (n = 3). (c) Schematics of BTP-7 conjugated to an aminohexanoic acid linker (X) at the C-terminus (BTP-7-X), BTP-7-X functionalized with a cysteine (Cys) residue at the C-terminus (BTP-7-X-Cys), and BTP-7-CPT drug conjugate. (d) Stability of compounds from (c) in 25% human serum. All datasets were normalized to the EIC at t = 0 (n = 2).

4.2.6 Delivery of CPT to intracranial GBM tumor using BTP-7

To investigate the efficacy of BTP-7-CPT in an orthotopic GBM mouse model, a patient-derived GBM-X6 tumor xenograft was established in the right hemisphere of the frontal lobe of each mouse. Tumor formation was confirmed by MRI at day 24 post xenotransplantation (**Fig. 4.6a**). At day 25, mice were randomized into three treatment groups to receive either vehicle (control), Scr-7-CPT, or BTP-7-CPT therapy (intraperitoneal (i.p.) 10 mg/kg, every 2 days). At day 47, we observed a decrease in tumor size in mice treated with BTP-7-CPT or Scr-7-CPT in comparison to vehicle (**Fig. 4.6a**). Ex vivo analysis of brain cryo-sections showed that tumor tissues in both treatment groups displayed a higher level of phospho-H2AX (p-H2AX), indicating

greater DNA damage in the treatment groups compared to the control (**Fig. 4.6b–d**). As expected, tumor tissues from the BTP-7-CPT group had a higher level of p-H2AX than the Scr-7-CPT group, underscoring the ability of BTP-7 to improve drug targeting to GBM (**Fig. 4.6b–d**). Importantly, both the treated groups showed negligible DNA damage in non-cancerous brain tissues from the left hemisphere of the brain (**Fig. 4.6b–d**).

Next, we evaluated their therapeutic efficacy using either a late-stage (day 25–49) or early-stage/long-duration (day 18–68) therapy intervention (i.p. 10 mg/kg, every 2 days). In both therapy regimens, treatment with BTP-7-CPT or Scr-7-CPT extended survival compared to the control animals that received only vehicle (**Fig. 4.6e,f**). Mice treated with BTP-7-CPT showed survival benefit over mice treated with Scr-7-CPT (late-stage (73 vs. 67 days, $p < 0.01$); early-stage/long-duration (88 vs. 82 days, $p < 0.05$)) (**Fig. 4.6e,f**). We further verified the safety of BTP-7-CPT by administering healthy nude mice with higher drug doses (25, 50, or 100 mg/kg). All animals showed no signs of morbidity and maintained their weight over 14 days, demonstrating that BTP-7-CPT is safe in mice even at higher therapeutic doses. Altogether, these results highlight that BTP-7-CPT is well-tolerated, preferentially targets patient-derived GBM xenograft in vivo without harming healthy brain tissues, and prolongs survival in mice.

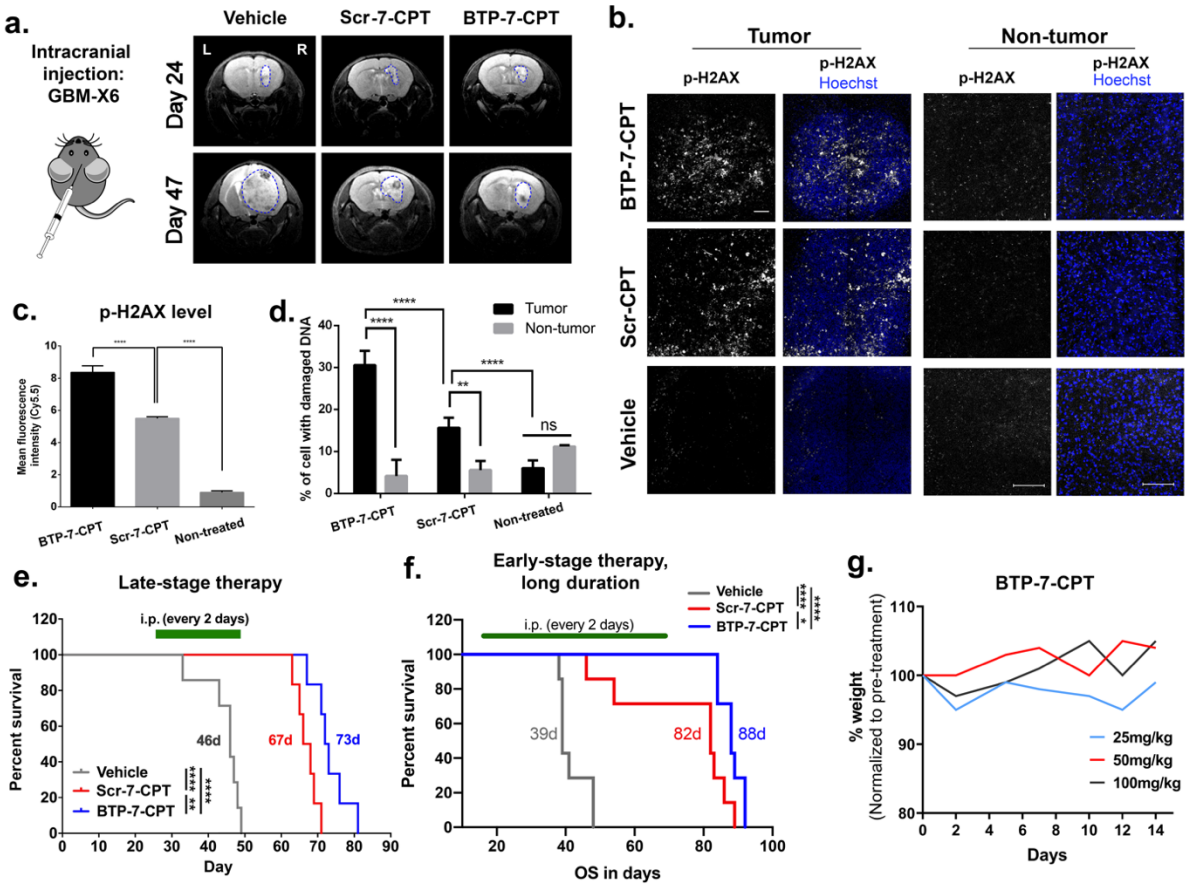


Figure 4.6 BTP-7 enhances drug delivery to GBM and prolongs survival in orthotopic PDX mouse model of GBM.

(a) Representative MRI scans (coronal view) of mouse brain showing GBM-X6 (delineated with blue lines) after orthotopic xenotransplantation of patient-derived GBM-X6 cells in the right frontal lobe of the mouse brain at day 24 (before treatment) and day 47 (after treatment). Both the BTP-7-CPT and Scr-7-CPT groups showed reduced tumor burden compared to the control vehicle group. (b,c,d) Ex vivo analysis of brain cryo-sections from mice bearing GBM-X6 tumors after treatment. (b) Immunofluorescence staining for phospho-H2AX in tumor and non-tumor areas of mice brains harvested at day 49 after 13 treatments (nuclei counterstained with Hoechst dye (blue)). Scale bar: 100 microns. (c) Quantification of phospho-H2AX signal within the tumor areas shown in (b) (n = 3, one-way ANOVA and Tukey's multiple comparisons test). (d) Quantification of the number of nuclei positive for phospho-H2AX signal in the tumor and non-tumor tissues of (b) (n = 3, two-way ANOVA and Sidak's multiple comparison test (**** $p < 0.0001$, ** $p < 0.01$). (e,f) Kaplan–Meier survival plot of mice bearing GBM-X6 tumors that received (e) late-stage treatment (i.p.; 10 mg/kg dose starting at Day 25 to 49 post tumor implantation) or (f) early-stage and longer treatment duration (i.p.; 10 mg/kg dose starting from day 18 to day 68 post tumor implantation) with BTP-7-CPT (blue), Scr-7-CPT (red), or vehicle (gray). Treatment was performed every 2 days. In both studies, a significant difference (late-stage $p < 0.0092$; early-stage $p < 0.022$) is observed between the BTP-7-CPT and Scr-7-CPT treatment group, as well as between the drug conjugate and vehicle group ($p < 0.0001$) as determined by the Log-rank (Mantel–Cox) test. (g) Graph showing the weights of the mice treated with BTP-7-CPT at 25, 50, and 100 mg/kg, indicating no decline in weight over more than 2 weeks.

4.3 Discussion

Despite significant progress in understanding the molecular biology underlying GBM, few advances have been achieved over the past two decades in the development of efficacious therapies for treating this morbid disease.[27] The poor prognosis and severe treatment side effects associated with GBM continue to drive efforts towards the development of new therapies. The precision medicine field has grown rapidly in recent years. However, clinical efficacy against GBM has been disappointing, mostly attributed to the significant inter- and intra-tumor genomic heterogeneity of GBM cells, limited delivery across the BBB, and the rapid development of resistant GBM phenotypes.[28,29] While many relevant genomic variants continue to be attractive targets, none of them have been strongly linked with clear prognostic and predictive value.[6]

Our efforts are focused on developing a peptide platform targeting the deglycosylated isoform of Bcan (dg-Bcan) that is found exclusively in the ECM of all high-grade glioma (including GBM) tissues analyzed to date. This isoform is also distributed throughout the tumor.[10,13] The specificity, ubiquity, and accessibility of dg-Bcan in the GBM tumor microenvironment make it a very promising marker for the development of new targeted therapies. The biological roles of dg-Bcan remain to be fully investigated. Several studies have been reported on the major glycosylated isoform of the full-length Bcan protein that is expressed in normal CNS and is associated with synaptic stabilization and plasticity.[30] Bcan is upregulated and secreted by malignant gliomas, and is implicated in tumor invasion, progression, and poorer prognosis.[31-34] It has multiple isoforms produced by glycosylation, cleavage, and alternative splicing. However, only the dg-Bcan isoform is uniquely expressed in human high-grade gliomas and absent in non-cancerous adult brain tissues.[35]

The mechanism of action of each Bcan isoform is poorly understood, although cleavage of the full-length Bcan is known to be necessary to promote glioma progression.[36] At the molecular level, Bcan and its cleaved products are secreted by the tumor cells, and activate EGFR/mitogen-activated protein kinase (MAPK) signaling and fibronectin production in glioma cells to enhance cell adhesion, migration, and invasive characteristics.[37] It is possible that the lack of glycosylation in dg-Bcan may regulate cleavage and signaling mechanisms, although these have not yet been

fully elucidated. It is also not clear why dg-Bcan is retained on the cell surface without being released as a soluble protein.[10,13] However, this unique spatial expression provides an excellent and accessible target that is specific to GBM cells. Bcan is found to accumulate at the tumor invasion front.[38] It has been reported that while normal brain tissues exhibit radiation-induced downregulation of Bcan, GBM cells retain or upregulate Bcan expression after irradiation.[39,40] In another study, GBM tissues obtained from patients after receiving standard-of-care (maximal safe resection followed by radiotherapy/concomitant and maintenance temozolomide chemotherapy) and bevacizumab monotherapy upon tumor recurrence, exhibit high Bcan expression with no significant difference between the 'poor' or 'better' prognostic group.[41] Indeed, we have previously observed a high level of dg-Bcan in both the primary and corresponding recurrent patient-derived GSC cultures,[13] further highlighting this protein as an important marker for targeting strategies against both the primary GBM tumor and the disease recurrence. The expression of Bcan following other forms of therapy (i.e., immunotherapy or oncolytic viral therapy) and its role in therapeutic resistance and GBM recurrence warrant further investigations.

Peptides are attractive tools for rationally designed targeting agents as they are small, cost-effective, scalable, and can be easily modified to select for desirable properties. With respect to GBM, this would include enhanced tumor specificity and BBB penetration.[42-43] They have significant potential for delivering therapeutics or diagnostics, while exhibiting faster diffusion and clearance rates.[44] BTP-7 is a small hydrophilic BBB-permeable peptide, and is made up entirely of d-amino acid residues which confer resistance to proteolytic degradation, altogether ensuing high peptide bioavailability in vivo. Clinical trials with CPT were discontinued in the 1970s due to its poor water solubility, high toxicity, and low response rate, leading to the development of CPT analogues with higher aqueous solubility (i.e., CPT-11) in the 1980s that were eventually approved by the FDA.[45] We found that BTP-7-CPT conjugate was even more soluble than CPT-11 (**Fig. 4.2a**). Collectively, our data show that BTP-7 binds dg-Bcan specifically, is internalized by GBM cells, is able to penetrate the BBB, and enhances therapeutic targeting to GBM tumors leading to prolonged survival in a PDX mouse model of GBM.[13]

In our efficacy studies in mice, we chose a treatment dose of 10 mg/kg, which is well below the highest dose tested (100 mg/kg) in our acute toxicity study, to ensure minimal risk of toxicity

over continuous administrations. Overall, BTP-7-CPT treatment appears safe and well-tolerated in mice. Both BTP-7-CPT and its scramble counterpart Scr-7-CPT showed efficacy in our PDX GBM model compared to the control mice that received the vehicle, although BTP-7 enabled a higher level of drug targeting to the tumor compared to Scr-7, leading to increased tumor toxicity relative to healthy brain tissues. The Scr-7 peptide was initially devised to investigate the amino acid sequence specificity of BTP-7. However, we found that Scr-7-CPT also reached the tumor, which is not surprising given the enhanced permeability and retention (EPR) effect of GBM, the ability of Scr-7-CPT to cross the BBB (albeit less permeable than BTP-7) (**Fig. 4.4**),[13] as well as higher Scr-7-CPT stability in mouse serum compared to BTP-7-CPT (**Fig. 4.5a,b**). Nevertheless, the enhanced GBM-targeting properties of BTP-7-CPT compared to Scr-7-CPT offered some survival benefit in vivo.

While native BTP-7 appears stable over 12 h in human serum,[13] the BTP-7-CPT conjugate is completely degraded within 1 h (**Fig. 4.5d**), suggesting a high likelihood of premature degradation of the disulfide linker in the bloodstream prior to reaching the tumor, despite this strategy having been successfully utilized for antibody–drug conjugates (ADC) in other studies.[46] Even though our previous pharmacokinetics analysis of [18F]BTP-7 radiotracer using PET imaging indicated that BTP-7 was almost immediately taken up by the GBM tumor post intravenous administration, future studies will aim to improve the linker stability in human serum as we continue to develop potential BTP-7-derived therapeutics for GBM therapy. Enzyme-labile peptide linkers cleaved through lysosomal hydrolysis upon cell internalization represent one potential option. Specifically, the valine–citrulline linker is cleaved by the lysosomal protease Cathepsin B that is upregulated in GBM.[47] This linker has been shown to have high serum stability and is used in several ADCs that are currently either in clinical trials or have been approved by the FDA.[46] Alternatively, ester linkers that are cleavable by esterases inside cells could be investigated, as they are considered stable cleavable linkers and have been employed for targeted delivery of CPT derivatives that are currently on the market.[48] We will also continue optimizing future therapeutic candidates based on the BTP-7 peptide platform for improved binding affinity to dg-Bcan, BBB permeability, and potency.

We have described the design of the first BTP-7 therapeutic conjugate and shown that BTP-7 could provide a robust platform for targeting various anti-cancer therapeutics to GBM. In addition to potent small molecules, other payloads that may be delivered with our platform include drug-loaded nanoparticles, radionuclides for targeted radiotherapy, or even potent lead candidate drugs that may have previously failed in clinical trials due to poor solubility. Taken together, BTP-7 is a new modality that opens the door to possibilities for targeted therapeutic approaches for GBM.

4.4 Experimental Section

4.4.1 Reagents and solvents

For peptide synthesis, N^{α}-Fmoc protected D-amino acids (a.a., Fmoc-D-Ala-OH, Fmoc-D-Arg(Pbf)-OH; Fmoc-D-Asn(Trt)-OH; Fmoc-D-Asp-(O*t*-Bu)-OH; Fmoc-D-Cys(Trt)-OH; Fmoc-D-Gln(Trt)-OH; Fmoc-D-Glu(O*t*-Bu)-OH; Fmoc-D-His(Boc)-OH; Fmoc-D-Ile-OH; Fmoc-D-Leu-OH; Fmoc-D-Lys(Boc)-OH; Fmoc-D-Lys(alloc)-OH, Fmoc-D-Met-OH; Fmoc-D-Phe-OH; Fmoc-D-Pro-OH; Fmoc-D-Ser(*t*-But)-OH; Fmoc-D-Thr(*t*-Bu)-OH; Fmoc-D-Trp(Boc)-OH; Fmoc-D-Tyr(*t*-Bu)-OH; Fmoc-D-Val-OH) and Fmox-L-Lys(biotin)-OH were purchased through Advanced ChemTech (Louisville, KY). Chem-Impex (Wood Dale, IL) and Peptides International (Louisville, KY). N^{α}-Fmoc protected L-amino acids (Fmoc-L-Ala-OH, Fmoc-L-Arg(Pbf)-OH; Fmoc-L-Asn(Trt)-OH; Fmoc-L-Asp-(O*t*-Bu)-OH; Fmoc-L-Cys(Trt)-OH; Fmoc-L-Gln(Trt)-OH; Fmoc-L-Glu(O*t*-Bu)-OH; Fmoc-Gly-OH, Fmoc-L-Ile-OH; Fmoc-L-Leu-OH; Fmoc-L-Lys(Boc)-OH; Fmoc-L-Met-OH; Fmoc-L-Phe-OH; Fmoc-L-Pro-OH; Fmoc-L-Ser(*t*-But)-OH; Fmoc-L-Thr(*t*-Bu)-OH; Fmoc-L-Trp(Boc)-OH; Fmoc-L-Tyr(*t*-Bu)-OH; Fmoc-L-Val-OH;) were purchased from the Novabiochem-line through Millipore Sigma (Darmstadt, Germany). Fmoc-L-His(Boc)-OH was bought from CEM. Succinimidyl 4,4'-azipentanoate (NHS-Diazirine) was purchased from Thermo Fisher Scientific (Waltham, MA) (#26167). Amino acids in peptide sequences are abbreviated with one letter code; capitalized letters refer to L-amino acids, lowercase letters refer to D-amino acids. 6-aminohexanoic acid, Irinotecan (CPT-11), and (S)-(+)-Camptothecin were purchased from ChemImpex (Wood Dale, IL).

H-Rink Amide-ChemMatrix resin was obtained from PCAS BioMatrix Inc. (St-Jean-sur-Richelieu, Quebec, Canada). 4-pentynoic acid, Fmoc-L-propargylglycine, 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU), and 2-(7-aza-1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU) were purchased from Chem Impex (Wood Dale, IL) and P3 Biosystems (Louisville, KY). (7-Azabenzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate (PyAOP) was purchased from P3 Biosystems (Louisville, KY). AldraAmine trapping agents (for 1000–4000 mL DMF), N-methyl pyrrolidinone (NMP), triisopropylsilane (TIPS), *t*-butylmethyl ether (TBME), Diisopropylethylamine (DIEA; 99.5%, biotech grade), piperidine (ACS reagent, ≥99.0%), trifluoroacetic acid (HPLC grade, ≥99.0%), triisopropylsilane (≥98.0%), formic acid (FA, ≥95.0%), dimethyl sulfoxide (DMSO, HPLC grade, ≥99.7%) were purchased from Sigma-Aldrich.

N,N-Dimethylformamide (DMF), dichloromethane (DCM), and HPLC-grade acetonitrile were from EMD Millipore (Billerica, MA). All solvents used for HPLC-MS were purchased from EMD and Fluka (Darmstadt, Germany). Cy5.5-azide was obtained from Lumiprobe (Hallandale Beach, FL). All other chemicals and reagents were purchased from Sigma-Aldrich (St. Louis, MO). Water was deionized using a Milli-Q Reference water purification system (EMD Millipore, Billerica, MA). Nylon 0.22 μm syringe filters were TISCH brand SPEC17984.

The following molecular biology supplies and reagents were used in our experimental studies: Recombinant human brevican (R&D Systems Minneapolis, MN; Cat. # 4009-BC-050), Streptavidin-coated pink-fluorescent magnetic particles (2.0–2.9 μm, Spherotech, Lake Forest, IL; Cat. # FSVM-2058-2), *O*-glycosidase (Roche; 11347101001), Neuraminidase (Sialidase) (Sigma-Aldrich, St. Louis, MO); Cat. # 10269611001), PNGase F (New England Biolabs, Ipswich, MA; Cat. # P0704S), Octet Ni-NTA Biosensors (Forte Bio, Fremont, CA; Cat. # 18-5101), CellTiter-Glo 3D cell viability assay (Promega, Madison, WI; Cat. # G983); Heparin Solution (STEMCELL Technologies; Cat. # 07980); Phospho-H2AX (Cell Signaling Technology; Cat. # 9718), Hoechst dye (Life Technologies; Cat. # H3570), Horseradish peroxidase (HRP)-conjugated anti-rabbit IgG (GE Healthcare; Cat. # NA934V), SuperSignal West Femto Maximum Sensitivity chemiluminescent substrate (Thermo Fisher Scientific; Cat. # 34096), clear-bottom black-well 96-well plates (Greiner BIO-ONE; Cat. # 655090).

4.4.2 Cell lines and culture conditions

GBM stem cells (GSC) GBM-X6 were cultured as neurospheres in Neurobasal medium (Invitrogen, Carlsbad, CA) supplemented with 2% B27 (Invitrogen), 1% glutamine (Invitrogen), epidermal growth factor (EGF) (20 ng/ml; PeproTech, Rocky Hill, NJ), and fibroblast growth factor-2 (FGF) (20 ng/ml; PeproTech, Rocky Hill, NJ). The GBM-X6 cells used in our studies had been passaged once in a mouse brain, where the cells were implanted intracranially and allowed to form a solid tumor over 30 days. Then, the GBM-X6 tumor was excised, dissociated in a tissue culture flask and cultured in Neurobasal medium as above. Human embryonic kidney (HEK) cells was cultured in DMEM (Invitrogen, Carlsbad, CA) supplemented with 10% FBS. Cells were grown in T25 or T75 vented-cap tissue culture flasks (Sarstedt AG and Co). HEK-Bcan cells were generated by cloning hBCAN cDNA into the pcDNA3.1 vector and stably transfected in HEK293 cells. Primary human astrocytes were purchased from Lonza Bioscience, Basel, Switzerland and

cultured in Astrocyte Growth Medium (AGM; Lonza Bioscience, Basel, Switzerland) (consisting of astrocyte basal medium supplemented with hEGF (human epidermal growth factor), insulin, ascorbic acid, GA-1000 (Gentamicin, Amphotericin-B), L-glutamine and 1% fetal bovine serum (FBS)). Astrocytes were grown in T75 Cell+ vented-cap tissue culture flasks (Sarstedt AG and Co, Nümbrecht Germany). Human brain microvascular pericytes (HBVP) (ScienCell Research Laboratories, Carlsbad, CA) were cultured in Pericyte Medium (ScienCell Research Laboratories, Carlsbad, CA) containing 2% FBS, pericyte growth supplement and penicillin-streptomycin. Immortalized human cerebral microvascular endothelial cells (hCMEC/D3) (Cedarlane Laboratories, Burlington, Canada) were maintained in culture in Endothelial Growth Medium (EGM-2) containing hEGF, hydrocortisone, GA-1000, FBS, VEGF, hFGF-B, R^3-IGF-1, ascorbic acid and heparin (Lonza Bioscience, Basel, Switzerland). For BBB organoid formation in low-attachment co-culture condition and functional assays, the organoids were maintained in EGM-2 (Lonza Bioscience, Basel, Switzerland) supplemented with 2% human serum (Valley Biomedical, Winchester, VA; Cat. # HS1021) and with the elimination of VEGF supplementation (this media formulation will henceforth be known as 'BBB working media'). All cells were cultured in a humidified incubator at 37 ^oC with 5% CO_2, and 95% natural air. For cell dissociation, StemPro Accutase Cell Dissociation Reagent (Thermo Fisher Scientific, Waltham, MA; Cat. # A1110501) was used on GSC cultured as neurospheres, and Trypsin-EDTA (0.05% v/v, with phenol red) (Thermo Fisher Scientific, Waltham, MA; Cat. # 25300054) was used for detaching adherent cells. All cell lines were regularly tested for mycoplasma contamination.

4.4.3 Liquid chromatography–mass spectrometry (LC-MS)

For mass spectrometry analysis, the filtered peptide solution (10 μL of a 1mg/mL solution) was diluted in 50% acetonitrile in water with 0.1% TFA (90 μL) to a final concentration of approximately 0.1 mg/mL. LCMS chromatograms and associated high resolution mass spectra were acquired using an Agilent 6520 Accurate-Mass Q-TOF LCMS (abbreviated as 6520) or an Agilent 6550 iFunnel Q-TOF LCMS system (abbreviated as 6550). Solvent compositions used in the LCMS are water with 0.1% formic acid additive (solvent A) and acetonitrile with 0.1% formic acid additive (solvent B).

4.4.4 Mass-directed reversed-phase high performance liquid chromatography (RP-HPLC)

For RP-HPLC purification, the crude lyophilized peptides were dissolved in water with 0.1% TFA additive containing a minimal amount of acetonitrile for solubility (e.g. 5% acetonitrile). All samples were filtrated through a Nylon 0.22 μm syringe filter prior to purification. For all HPLC purifications, a gradient of acetonitrile with 0.1 % TFA additive (solvent B) and water with a 0.1% TFA additive (solvent A) was used unless otherwise noted. Specific purification conditions such as column and gradient are specified for each case.

4.4.5 NMR Spectroscopy

Proton nuclear magnetic resonance (^{1}H NMR) spectra were recorded in 5 mm tubes on a Bruker Avance Neo spectrometer in deuterated solvents at room temperature. Chemical shifts (δ scale) are expressed in parts per million (ppm) and are calibrated using residual protic solvent as an internal reference (DMSO: δ = 2.50 ppm). Data for ^{1}H NMR spectra are reported as follows: chemical shift (δ ppm) (multiplicity, coupling constants (Hz), integration). Couplings are expressed as: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet or combinations thereof. Carbon chemical shifts (δ scale) are also expressed in parts per million (ppm) and are referenced to the central carbon resonance of the solvent (DMSO: δ = 39.52 ppm). In order to assign the ^{1}H and ^{13}C NMR spectra, a range of 2D NMR experiments (COSY, HSQC, HMBC, NOESY) were used as appropriate.

4.4.6 Octet binding kinetic analysis

The FortéBio OctetRed384 was used to study the binding kinetics of each peptide to recombinant human brevican (in PBS and 0.1 mM EDTA, pH 6.8). The brevican protein was deglycosylated prior to experimental use as detailed above. All binding kinetics assays were performed within the OctetRed instrument under agitation at 1000 rpm in 0.9% NaCl irrigation with 0.05% Tween (working buffer). Assays were performed at 30°C in solid black 384-well plates (Geiger Bio-One). The final volume for all the solutions was 80 μl/well. Firstly, Ni-NTA biosensors were soaked for 10 min in working buffer. Before loading the protein onto each biosensor, a baseline was established in working buffer for 60 s. Deglycosylated brevican, dg-Bcan (50 μg/ml; His-tagged) was loaded on the surface of each biosensor for 180 s. Typical capture levels varied slightly between 0.5 and 2 nm, and variability within run did not exceed 0.1 nm. Reference

biosensors were exposed to the 'de-glycosylation buffer' lacking brevican protein during the loading step as internal controls. A 60 s biosensor washing step was applied. Then, biosensors were exposed to the analyte (peptide) in working buffer (ranging between 0–10 μM BTP) for 300 s during the 'association' step. Finally, the biosensors were exposed to working buffer (without peptide) during the 'dissociation' step for 600 s. Binding affinity of each peptide was assessed through steady state analysis, where the response unit was plotted over peptide concentration. All data was plotted and curves were fitted using the non-linear 'one-site specific binding' fit, and the dissociation constant (K_D) value was calculated using the GraphPad Prism software.

4.4.7 Cell Culture

GBM-X6 is a xenoline that was established initially by direct implantation of freshly resected human GBM tissues into the flanks of immunocompromised athymic nude mice. GSCs are normally cultured and maintained as neurospheres in the absence of serum to preserve their stem-like state. GS401 is a GSC derived from a recurrent patient from Erasmus Medical Center [13]. We maintained both GBM-X6 and GS401 cultures in neurobasal (NB) supplemented with 2% B27, 1% glutamine, 20 ng/mL of fibroblast growth factor-2 (FGF), and 20 ng/mL of epidermal growth factor (EGF) [16]. Human embryonic kidney (HEK) and dg-Bcan-overexpressing HEK cells (HEK-Bcan) were cultured in DMEM supplemented with 10% FBS. All cells were cultured in a 37 °C incubator with 5% CO_2 and 95% natural air.

4.4.8 Synthesis of BTP-7-CPT

Disulfide-cleavable CPT prodrug was synthesized as described [17]. The pyridyldithiol arm of the prodrug allows for conjugation to free thiols via disulfide exchange, enabling CPT to be attached to a cysteine residue on BTP-7.

4.4.9 Analysis of Compound Solubility

Dried powder of each compound was first resuspended in pure dimethyl sulfoxide (DMSO) to form a 10 mM stock. Each stock was then diluted in saline (0.9% NaCl) to form an aqueous solution of 100 μM (final DMSO 1% (*v*/*v*)) or 1.3 mM (final DMSO 10% (*v*/*v*)) of each compound. The samples were incubated at room temperature for 5 min, then centrifuged at 10,000 rcf for 10 min to pellet any precipitates. 5 μL of supernatant was diluted into LC-MS vials containing 100 μL of 50:50 water:acetonitrile with 0.1% TFA additive and analyzed by LC-MS. The presence of

soluble compound in each supernatant was determined by quantifying the ion count in the extracted ion chromatogram (EIC) using the MassHunter software. Solubility was measured by comparing the total EIC of the sample diluted in saline vs. in pure DMSO. Each sample was measured in duplicates.

4.4.10 Western Blot Analysis

Western blotting was performed as previously described [13]. GBM-X6 and GS401 cells cultured as neurosphere suspensions were collected from the culture flask using a pipette. HEK293 and Bcan-overexpressing HEK293 (HEK-Bcan) cells cultured as adherent cells were harvested from their flask using a cell scraper. Cell lysates were extracted using NP40 cell lysis buffer, the samples were separated on an SDS-PAGE gel, and then transferred onto a polyvinylidene difluoride (PVDF) membrane by Western blotting. The pan-brevican (Bcan) primary antibody[13] was used to detect the deglycosylated isoform of the Bcan protein. Horseradish peroxidase (HRP)-conjugated anti-rabbit IgG was used as a secondary antibody, and SuperSignal West Femto Maximum Sensitivity chemiluminescent substrate was used for enhanced chemiluminescence (ECL) detection.

4.4.11 Cell Viability Assay

The CellTiter-Glo® luminescent cell viability assay was used. GBM-X6 and GS401 glioma stem cells [18] (cultured in Neurobasal growth media), or HEK293 cells (cultured in supplemented DMEM media) were washed once with PBS, dissociated to form a single cell suspension using either StemPro Accutase for GBM-X6 and GS401 neurospheres, or 0.05% (v/v) Trypsin-EDTA for HEK293 cells, resuspended in the appropriate growth media, and then counted using a hemocytometer. Cells were seeded at the desired number (typically 30,000 GBM-X6 cells per well; 5000–20,000 HEK cells per well; 8000 GS401 cells per well) into clear-bottom black-well 96-well plates in triplicates. Stock solutions (10 mM) of CPT, BTP-7-CPT, or scramble Scr-7-CPT were prepared in DMSO. A 1:3 serial dilution (in DMSO) of each stock was performed, and then 1 µL of each compound was added onto cells in each well containing 99 µL of media (in triplicate), so that a final concentration range of 0–100 µM (1% *v/v* DMSO) was achieved. Vehicle (DMSO with no drug) was added as a control. The plates were returned into a 37 °C tissue culture incubator. At every 24 h after incubation, a plate was removed from the incubator and 100 µL of CellTiter-Glo reagent was added to each well using a multi-channel pipette. The plate was incubated (in the

dark) for 10 min on a nutating mixer, and then analyzed in a luminescence plate reader (POLARstar Omega, BMG Labtech, Offenburg, Germany). Each dataset was normalized to the vehicle (control), and plotted using GraphPad Prism. IC_{50} values were obtained using a non-linear fit (four parameters). Each experiment was repeated 3 times to ensure reproducibility.

4.4.12 Permeability Analysis In Vitro Using Human BBB Organoids

Multicellular BBB organoids were established for 48 h as previously described [19,20]. Briefly, primary human astrocytes, primary human brain vascular pericytes (HBVP), and hCMEC/D3 brain endothelial cells were co-cultured in endothelial cell media (ECM) supplemented with 2% human serum in a sterile 96-well plate coated with 1% agarose. The cells were allowed to self-assemble into multicellular BBB organoids in a 37 °C incubator with 5% CO_2 and 95% natural air for 48 h. The organoids were then pooled together in a microfuge tube and incubated with each drug conjugate at a final concentration of 10 μM in EGM containing 2% human serum for 3 h in a 37 °C incubator under constant rotation (n = 6–8 organoids per group). To examine if the drug conjugate affects the overall integrity of the BBB organoid surface, the organoids were also co-incubated with fluorescent dextran (TRITC-dextran (4.4 kDa)) at a final concentration of 10 μg/mL. The organoids were then washed three times with 1 mL of BBB working medium (see Supplementary Materials and Methods) and fixed in 3.7% formaldehyde in PBS for 10 min at room temperature. The organoids were then washed three times with 1 mL of media, transferred onto a thin-well chambered cover glass and imaged by confocal fluorescence microscopy. Quantification of peptide and dextran permeability was performed using the Fiji software, where the mean fluorescence intensity of the core of each organoid (at a depth between 50–90 μm) was measured and plotted using GraphPad Prism.

4.4.13 Ex Vivo Serum Stability Assay

Mouse serum was obtained from 8–10-week-old athymic female mice. Mice were euthanized by CO2 asphyxiation with secondary decapitation to ensure death. An incision was made from the neck to the abdomen and the thorax was cut to expose the heart. An amount of 100 μL of heparin was then injected directly into the right atrium to prevent the blood from clotting. A cardiac puncture was performed by piercing the left ventricle with a 21G needle to remove 400–1000 μL of blood. Blood from five mice was collected and combined in a 10-mL centrifuge tube

and centrifuged for 5 min at 500 rcf at 4 °C. The top layer containing the serum (opaque) was removed and stored at −20 °C.

To test for serum stability, a 10 mM stock of BTP-7-CPT and Scr-7-CPT was prepared in DMSO. Each compound (1.4 µL) was transferred into a microcentrifuge tube containing 200 µL of PBS with 25% mouse serum, or in PBS alone, and incubated at 37 °C (final concentration 70 µM). At timepoints of 0, 1, 3, 6, and 12 h, 10 µL of each sample was removed from the microcentrifuge tube and transferred to a different microcentrifuge tube containing 20 µL of a solution of 2 M guanidine hydrochloride (GuHCl). Then, 10 µL of the sample from each time point was purified via solid-phase extraction with Millipore C18 10 µL ziptips to separate remaining peptide from serum proteins. Samples containing the peptide were eluted with 10 µL of 70% acetonitrile in water containing 0.1% TFA into an LC-MS vial containing 20 µL of water with 0.1% TFA additive and analyzed via LC-MS. The amount of intact compound was determined through quantifying the ion count in the extracted ion chromatogram (EIC) using the MassHunter software. Each timepoint was performed in duplicate.

4.4.14 Intracranial GBM Implantation and Efficacy Studies

All animal protocols were reviewed and approved by the in-house Institutional Animal Care and Use Committee (IACUC). GBM-X6 GSC (100,000 cells) resuspended in 2µL of PBS were inoculated into the right striatum of 6–8-week-old athymic female mice using a stereotactic frame as described. T2-weighted MRI of the brain was performed the day before treatment to ensure uniform tumor formation. The animals were then randomly assigned into groups and injected intraperitoneally (i.p.) with either BTP-7-CPT, Scr-7-CPT, or vehicle control (DMSO in 0.9% NaCl) at 10 mg/kg (n = 6–7 mice). We employed either a late-stage treatment regimen (i.p. injection every 2 days from Day 25 to 49 post tumor implantation), or early-stage and longer treatment regimen (i.p. injection every 2 days from day 18 to day 68 post tumor implantation). At day 47 in the ‘late-stage’ therapy study, brains were imaged again by T2-weighted MRI. The Kaplan–Meier survival graph was plotted using GraphPad Prism.

4.4.15 Ex Vivo Immunofluorescence Staining of GBM Tissue Sections

At day 49 in the ‘late-stage’ therapy study, one animal from each treatment group was sacrificed by CO2 asphyxiation and transcardial perfusion. Brains were frozen and cryosectioned into 16 µm sections and immune-stained for phospho-H2AX. The rabbit anti-phospho-H2AX was

used. Anti-rabbit Alexa-Fluor 647 secondary antibody (Invitrogen) and Hoechst dye were used for detection. Brain tissue sections were fixed with methanol for 1 min. The slides were then washed twice with PBS + 0.025% Triton X-100. Tissue sections were blocked with 10% normal goat serum diluted in PBS + 0.025% Triton X-100. Primary antibody against phospho-H2AX was added to the blocking solution (at 1:100 dilution) and the tissues were incubated overnight at 4 °C in a dark humidified box. The slides were then washed three times (2 min each) with PBS + 0.025% Triton X-100. Tissues were then incubated with Alexa Fluor secondary antibody and Hoechst 33,342 (both at 1:1000 dilution) in blocking solution. Then, the slides were washed three times (2 min each) with PBS + 0.025% Triton X-100. Vectashield mounting medium for fluorescence was applied onto the tissues, and a coverslip was mounted onto the slide. The edges of the coverslip were sealed with clear nail polish, and the tissues were imaged using an epifluorescence microscope under a 20x objective (n = 10).

4.4.16 Statistical Analysis

All datasets are presented as means ± SD. Significance is indicated on each graph and the statistical tests performed are indicated in the figure legends. The significance is represented as follows: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

4.4.17 Use of Human Specimens

All patient-derived cells were obtained from patients who were undergoing surgical treatment at the Mayo Clinic in Rochester or Erasmus Medical Center. All subjects had provided informed consent and signed consent forms that were approved by the Institutional Review Board (IRB). Our lab received the cells in a de-identified manner.

4.5 Acknowledgements

Contributions: Conceptualization, C.-F.C., S.E.L. and B.L.P.; methodology, C.-F.C., C.E.F., C.M.F., B.S., N.v.S. and N.H.; validation, C.-F.C., C.E.F., C.M.F., B.S., P.Z., N.v.S. and N.H.; formal analysis, C.-F.C., C.E.F., C.M.F., B.S., P.Z., N.v.S. and N.H.; investigation, C.-F.C., C.E.F., C.M.F., B.S., P.Z., N.v.S. and N.H.; resources, C.-F.C., S.E.L. and B.L.P.; data curation, C.-F.C., C.E.F., C.M.F., B.S., P.Z. and N.v.S.; writing—original draft preparation, C.-F.C., C.E.F., C.M.F. and B.S.; writing—review and editing, C.-F.C., C.E.F., C.M.F., B.S., N.v.S., A.L., N.H., B.L.P. and S.E.L.; supervision, C.-F.C., A.L., S.E.L. and B.L.P.; funding acquisition, C.-F.C., S.E.L. and B.L.P. All authors have read and agreed to the published version of the manuscript.

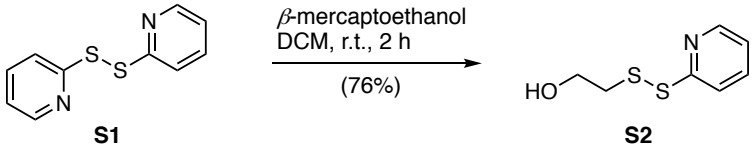
Funding: C.-F.C. is supported by grants from the Brigham Research Institute, Sperling Family Foundation, Harvard Stem Cell Institute and Department of Defense (DOD grant number W81XWH1910791). N.v.S. is supported by the Deutsche Forschungsgemeinschaft (DFG) (research fellowship grant number 400975596). S.E.L. is supported by a grant from the National Cancer Institute (R01CA237063). B.L.P. is supported by a grant from the National Cancer Institute (R01CA237063). C.E.F. is supported by the National Science Foundation (NSF) Graduate Research Fellowship program (fellowship number 4000057441).

The study was conducted in accordance with institutional IRB approval from Dana Farber Cancer Institute (Title: Integrated Tissue and Clinical Data Bank for Patients with Neurological Disorders, Protocol No: 10-417, Protocol Version No./Date: v37:12/04/2016). Animal experiments described in this proposal are included in the BWH Center of Comparative Medicine Animal Experimentation Protocol 2019N000074 (Choi-Fong Cho (PI), BWH IACUC approved 2019).

We acknowledge Zaimara Canas-Sniffin for administrative help in grant applications and Sylwia Wojcik for helping to manage daily lab activities.

4.6 Appendix I: Synthesis of Camptothecin-linker and conjugation to BTP-7

4.6.1 2-(pyridin-2-yldisulfaneyl)ethan-1-ol (S2)



S1 **S2**

2-Mercaptoethanol (500 µL, 7.10 mmol, 1 eq.) and 2,2'-dipyridyl disulfide (S1, 4.70 g, 21.4 mmol, 3 eq.) were dissolved in DCM (20 mL) and stirred at room temperature for 3 hrs. Afterwards, the reaction mixture was concentrated under reduced pressure and flash column chromatography [hexanes/EtOAc, 10:1 to 2:1] afforded mixed disulfide S2 (1.01 g, 5.39 mmol, 76%) as a yellow solid.

Rf = 0.45 [hexanes/ethyl acetate, 1:1].

1H NMR(500 MHz, DMSO-d6) δ 8.50 – 8.42 (m, 1H), 7.88 – 7.75 (m, 2H), 7.24 (ddd, J = 6.6, 4.8, 2.1 Hz, 1H), 4.99 (t, J = 5.5 Hz, 1H), 3.62 (q, J = 6.1 Hz, 2H), 2.92 (t, J = 6.3 Hz, 2H) ppm.

13C NMR (500 MHz, DMSO-d6) δ 159.5, 149.5, 137.8, 121.1, 119.3, 59.1, 41.2 ppm.

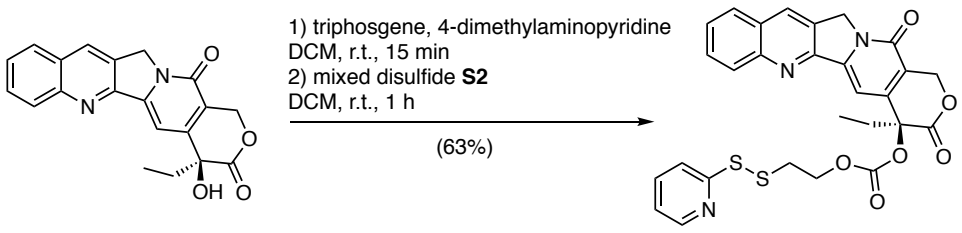
FTIR (thin film): ṽ = 3288 (br), 2920 (w), 2863 (w), 1574 (m), 1415 (s), 1043 (m), 754 (s) cm–1.

HRMS (ESI):

calcd. for C7H10NOS2+: 188.0198 [M+H]+

found: 188.0269 [M+H]+.

4.6.2 2-pyridinyldithioethyl carbonate Camptothecin (S4)



S3 **S4**

Camptothecin (S3, 250 mg, 0.718 mmol, 1 eq.), triphosgene (82.6 mg, 0.278 mmol, 0.387 eq.) and 4-dimethylaminopyridine (459 mg, 3.75 mmol, 5.21 eq.) were combined in dry DCM (10 mL), and after 15 min, mixed disulfide S2 was added (148 mg, 0.790 mmol, 1.10 eq.) and the reaction

was stirred at room temperature for 4 hrs. Flash column chromatography [DCM:Acetone, 20:1 to 4:1] afforded S4 (254 mg, 0.452 mmol, 63%).

Rf = 0.27 [hexanes/ethyl acetate, 1:3].

1H NMR (500 MHz, DMSO-d6) δ 8.68 (s, 1H), 8.39 (ddd, J = 4.8, 1.9, 0.9 Hz, 1H), 8.18 – 8.09 (m, 2H), 7.85 (ddd, J = 8.5, 6.8, 1.5 Hz, 1H), 7.81 – 7.71 (m, 1H), 7.71 – 7.63 (m, 2H), 7.15 (ddd, J = 7.4, 4.8, 1.1 Hz, 1H), 7.09 (s, 1H), 5.52 (d, J = 2.0 Hz, 2H), 5.29 (s, 2H), 4.33 (t, J = 6.0 Hz, 2H), 3.20 – 3.08 (m, 2H), 2.25 – 2.13 (m, 2H), 0.92 (t, J = 7.4 Hz, 3H) ppm.

13C NMR (500 MHz, DMSO-d6) δ 167.1, 158.6, 156.5, 152.7, 152.2, 149.6, 147.9, 146.3, 144.7, 137.7, 131.6, 130.4, 129.8, 129.0, 128.5, 128.0, 127.8, 121.3, 119.4, 119.2, 94.4, 77.9, 66.5 66.8, 50.3, 36.8, 30.3, 7.6 ppm.

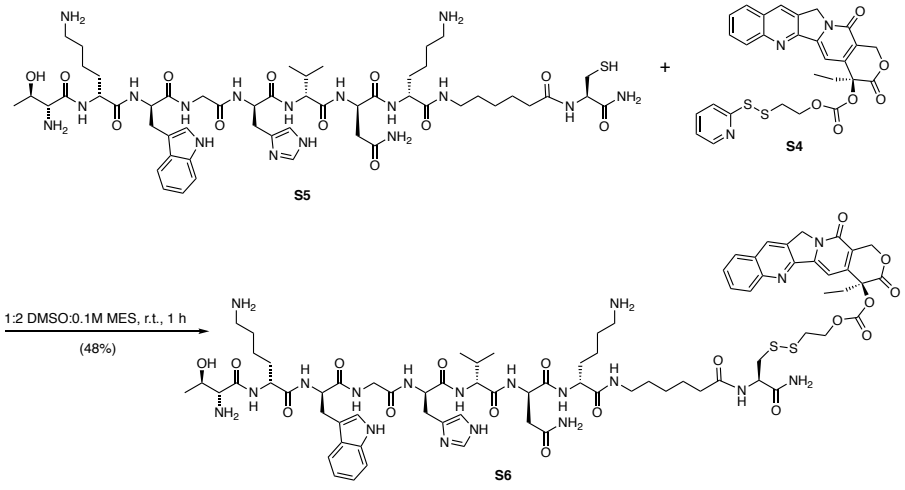
FTIR (thin film): ṽ = 2992.83 (w), 1748.22 (m), 1667.56 (m), 1748.22 (m), 1667.56 (m),1253.34 (m), 666.08 (s) cm–1.

HRMS (ESI):

calcd. for C28H23N3O6S2+: 562.1101 [M+H]+;

found: 562.1104 [M+H]+.

4.6.3 BTP-7-Camptothecin (S6)



The pyridyldithiol arm of S4 allows for conjugation to free thiols via disulfide exchange, enabling S4 to be attached to BTP-7 with a C-terminal cysteine S5. To perform this conjugation, the peptide (S5, 50 mg, 32 μmol, 1.9 eq.) in 2-(N-morpholino)ethanesulfonic acid (MES, 0.1 M, pH 6, 14 mL)

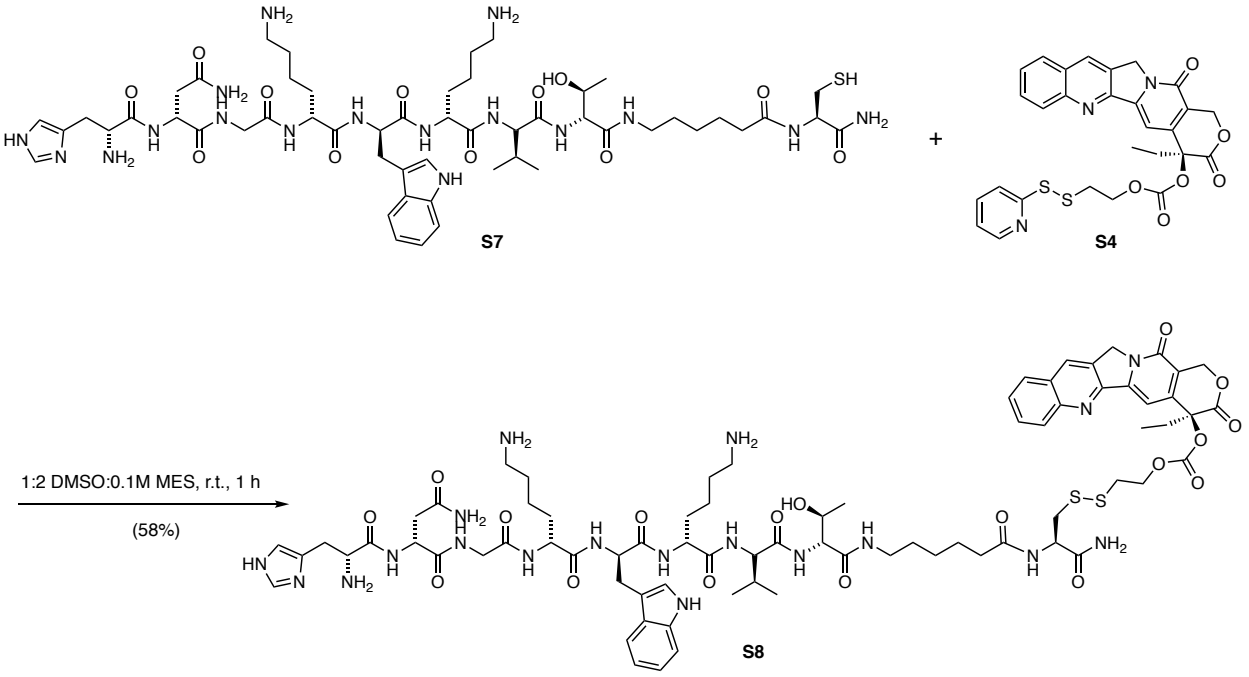
was combined with S4 (9.5 mg, 17 µmol, 1 eq.) in DMSO (7 mL) and stirred at room temperature for 1 hr. The reaction was quenched with trifluoroacetic acid (0.20 mL, 2.6 mmol, 150 eq.). The camptothecin-BTP-7 (S6) conjugate was purified by RP-high-performance liquid chromatography (RP-HPLC) in 100 mM ammonium acetate buffer (pH = 5) and an acetonitrile gradient (15–30% over 75 min) (15 mg, 8.1 µmol, 48%). Purity of each RP-HPLC fraction was confirmed by LC-MS.

HRMS (LC-MS):

calcd. for C76H103N19O18S2+: 817.8661 [M+2H]2+

found: 817.8666 [M+2H]2+.

4.6.4 Scrambled BTP-7-Camptothecin (S8)



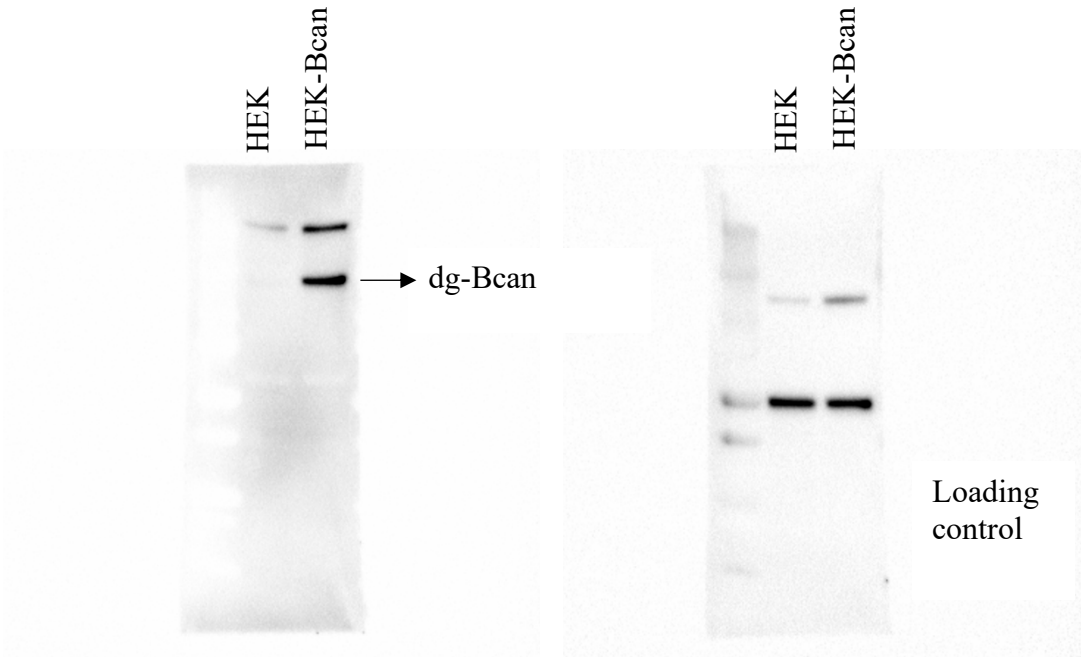
The conjugation and purification protocols given for S6 were used to conjugate the scrambled BTP-7 peptide (S7) to camptothecin to synthesize the camptothecin-scramble (S8) (18 mg, 9.7 µmol, 58%).

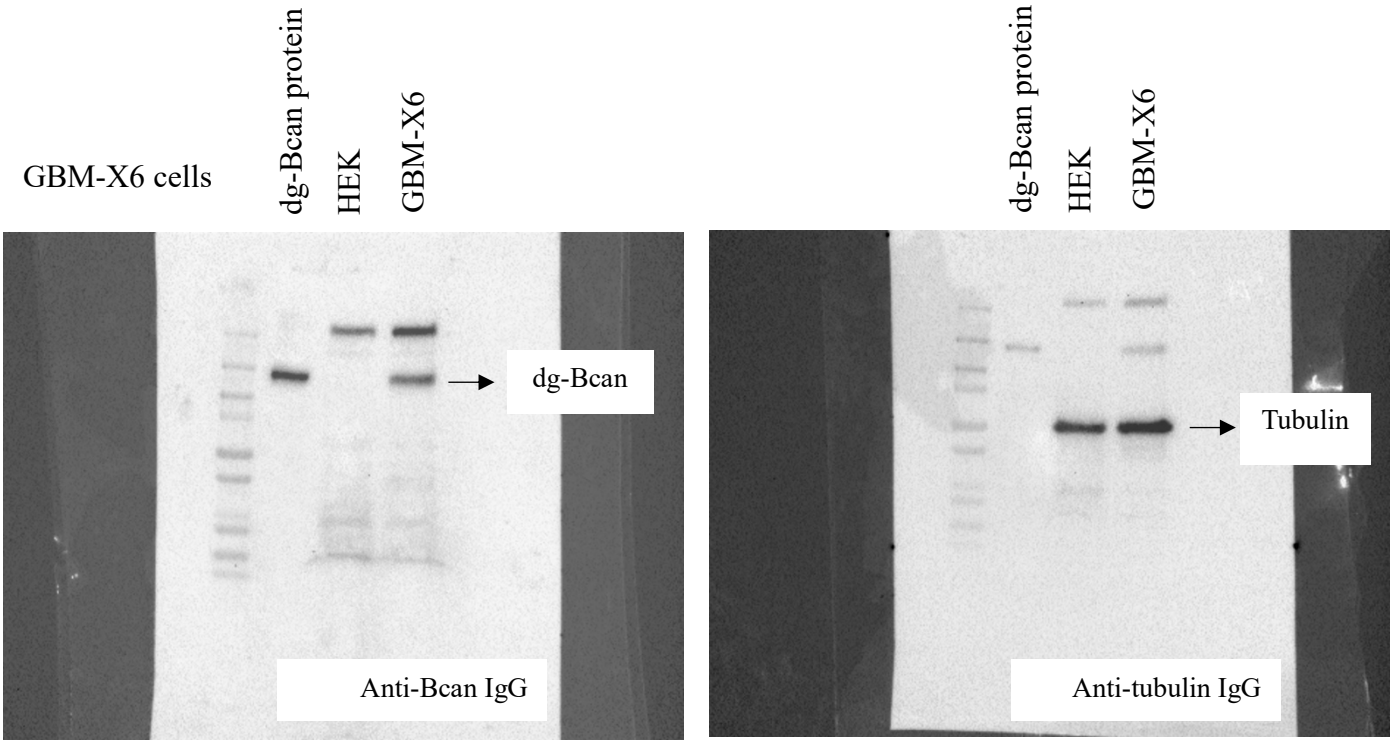
HRMS (LCMS):

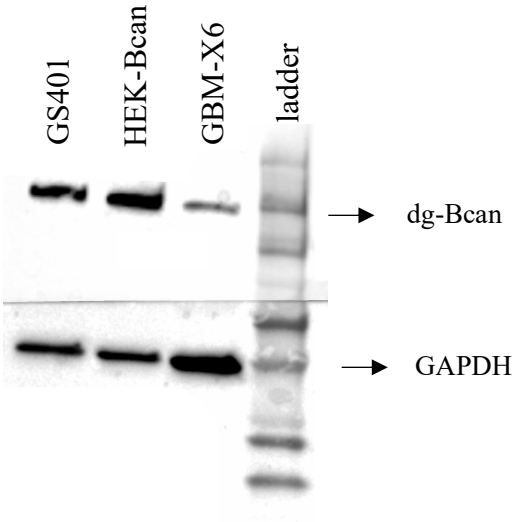
calcd. for C76H103N19O18S2+: 817.8661 [M+2H]2+

Found: 817.8669 [M+2H]2+.

4.7 Appendix II: Full Western Blot images

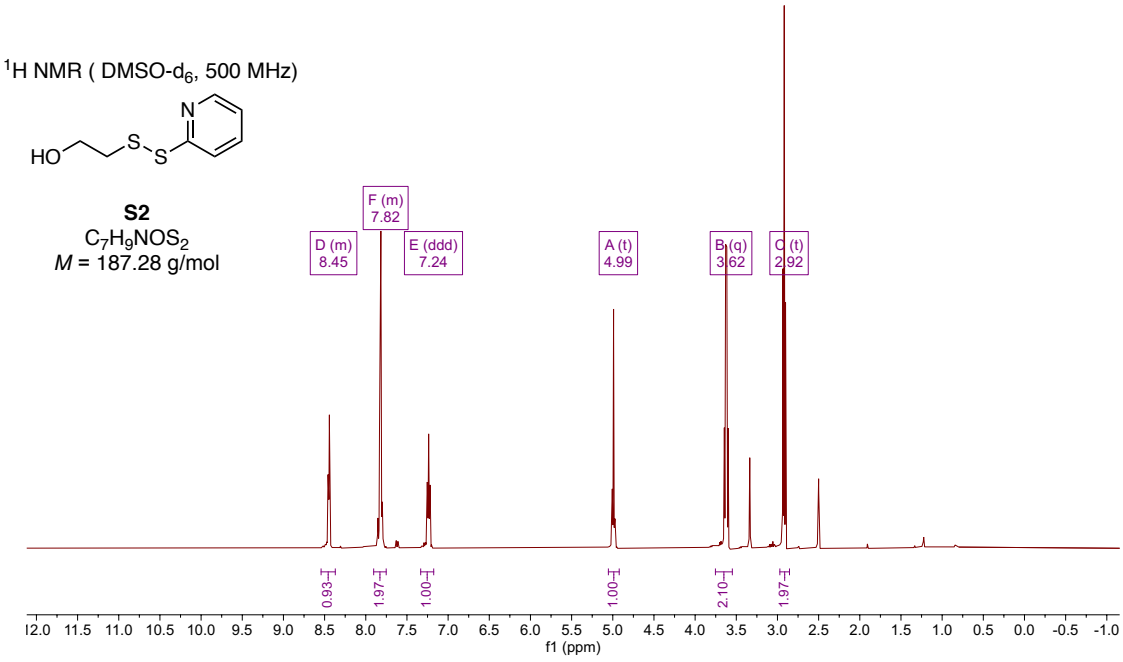


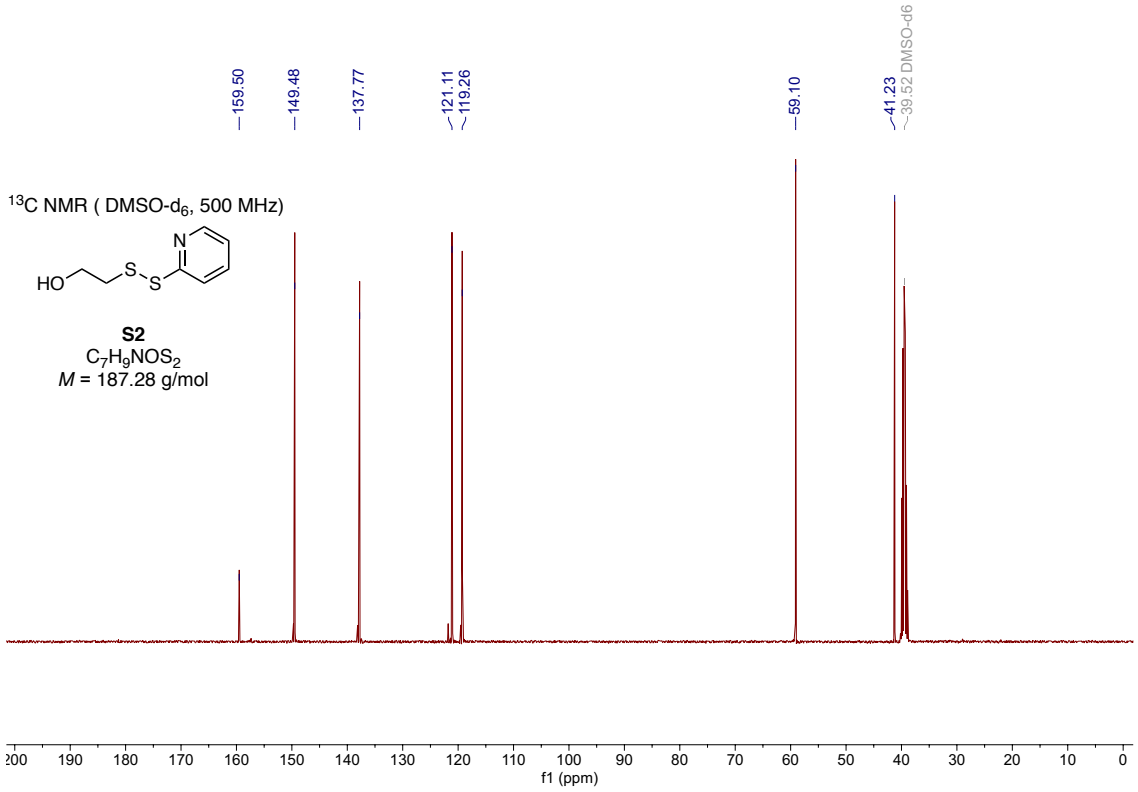


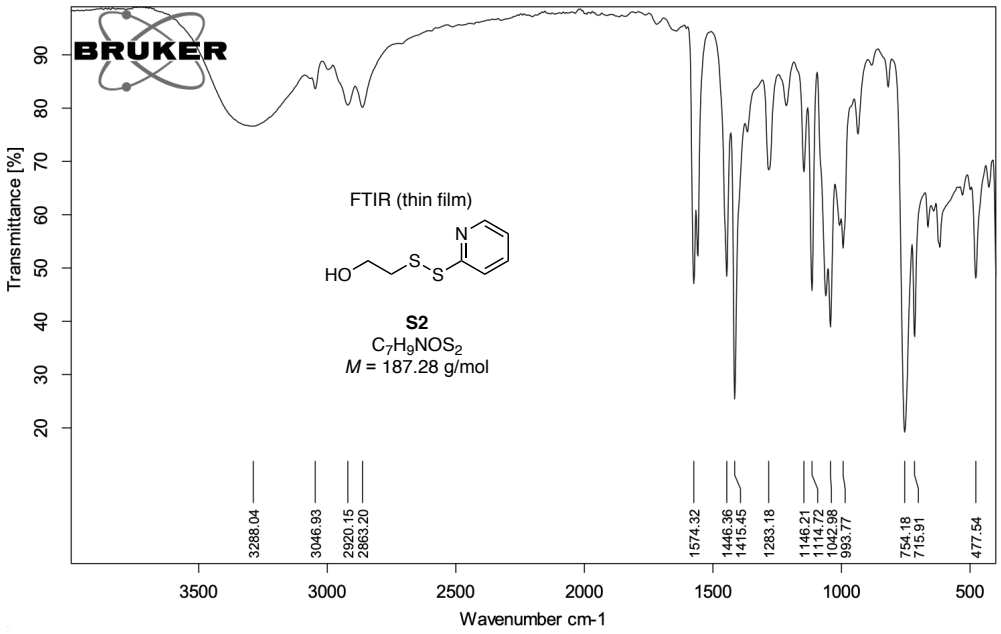

GS401
HEK-Bcan
GBM-X6
ladder
dg-Bcan
GAPDH

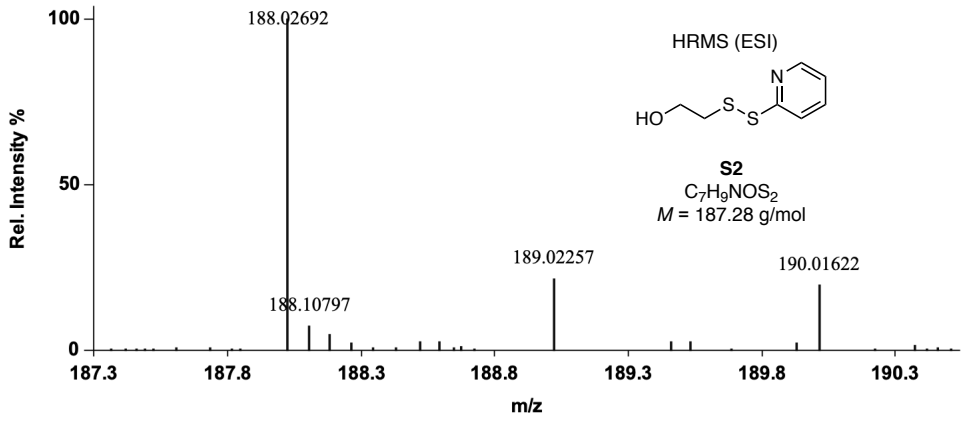
4.8 Appendix III: Characterization of compounds

2-(pyridin-2-yldisulfaneyl)ethan-1-ol (S2)

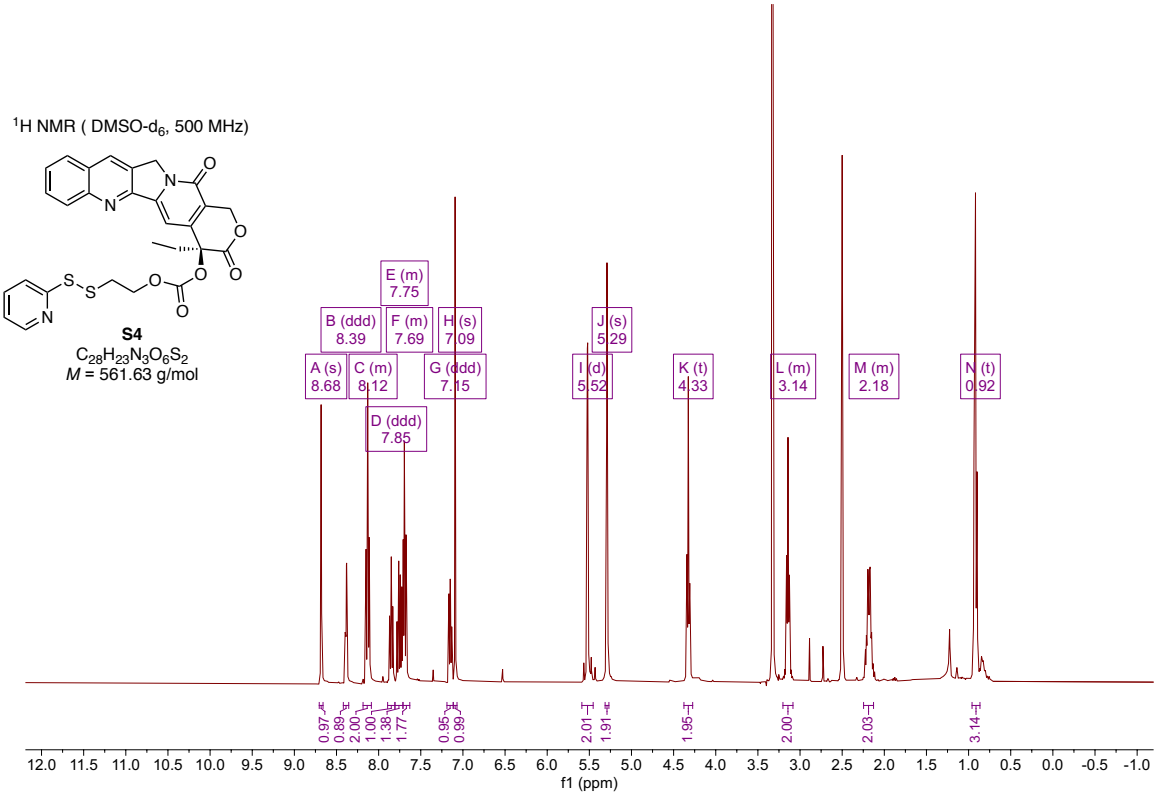


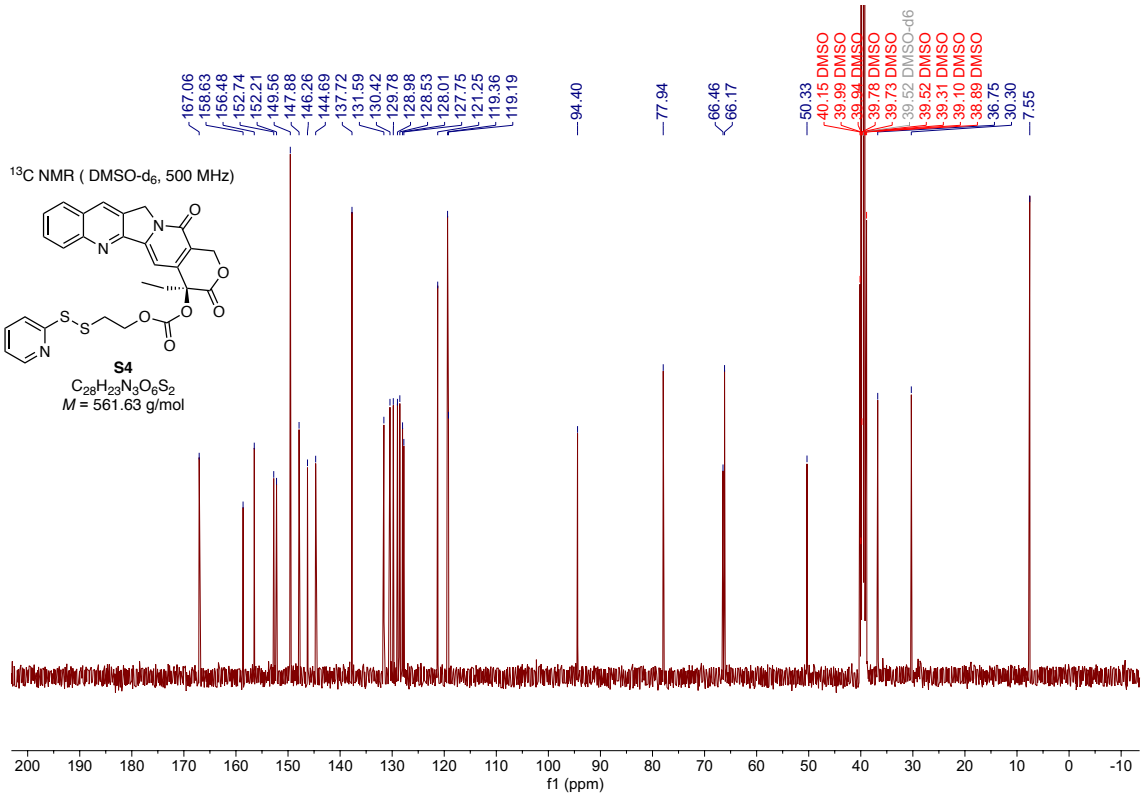

159.50
149.48
137.77
121.11
119.26
59.10
41.23
39.52 DMSO-d6
^{13}C NMR (DMSO-d$_6$, 500 MHz)
HO
S
S
N
S2
C$_7$H$_9$NOS$_2$
M = 187.28 g/mol
200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)

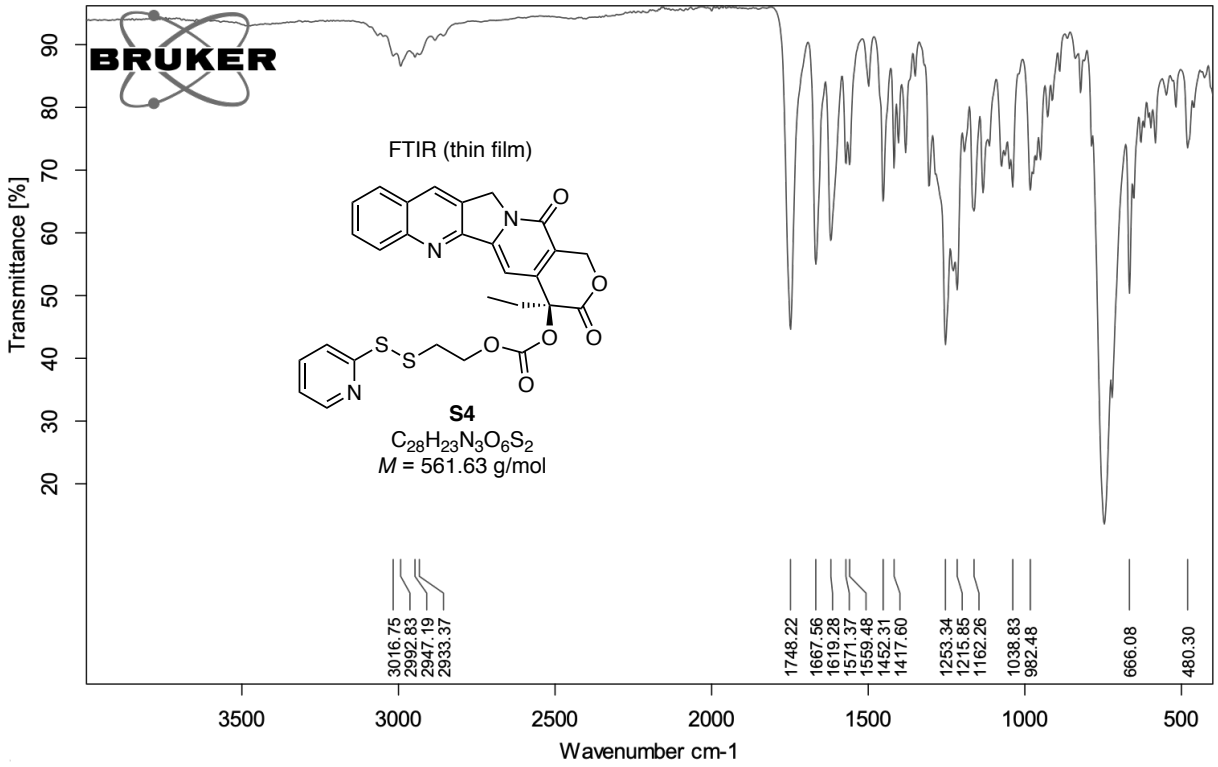

BRUKER
FTIR (thin film)
HO
S
S
N
S2
C$_7$H$_9$NOS$_2$
M = 187.28 g/mol
Transmittance [%]
20 30 40 50 60 70 80 90
3288.04
3046.93
2920.15
2863.20
1574.32
1446.36
1415.45
1283.18
1146.21
1114.72
1042.98
993.77
754.18
715.91
477.54
3500 3000 2500 2000 1500 1000 500
Wavenumber cm-1

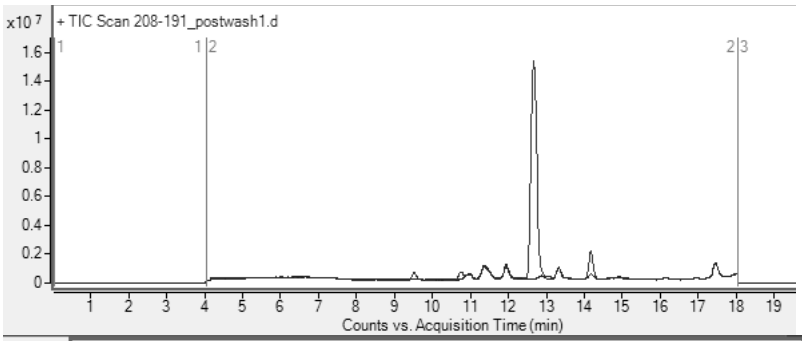


2-pyridinyldithioethyl carbonate Camptothecin (S4)

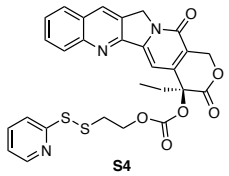



167.06
158.63
156.48
152.74
152.21
149.56
147.88
146.26
144.69
137.72
131.59
130.42
129.78
128.98
128.53
128.01
127.75
121.25
119.36
119.19
94.40
77.94
66.46
66.17
50.33
40.15 DMSO
39.99 DMSO
39.94 DMSO
39.78 DMSO
39.73 DMSO
39.52 DMSO
39.52 DMSO
39.31 DMSO
39.10 DMSO
38.89 DMSO
36.75
30.30
7.55
13C NMR (DMSO-d6, 500 MHz)
S4
C28H23N3O6S2
M = 561.63 g/mol
200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10
f1 (ppm)

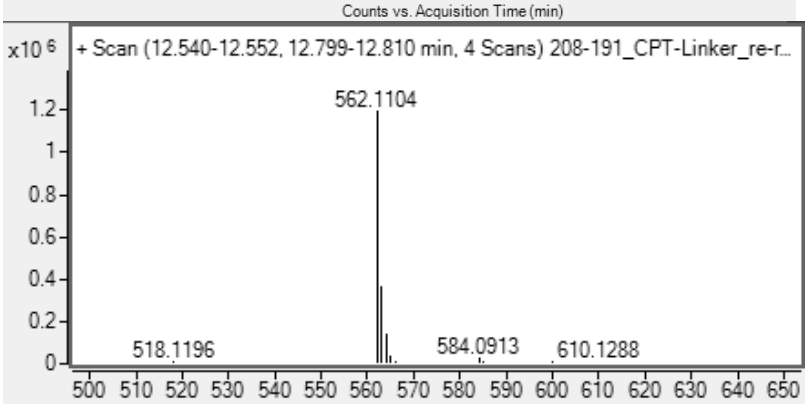

BRUKER
FTIR (thin film)
S4
C28H23N3O6S2
M = 561.63 g/mol
Transmittance [%]
20 30 40 50 60 70 80 90
3016.75
2992.83
2947.19
2933.37
1748.22
1667.56
1619.28
1571.37
1559.48
1452.31
1417.60
1253.34
1215.85
1162.26
1038.83
982.48
666.08
480.30
3500 3000 2500 2000 1500 1000 500
Wavenumber cm-1



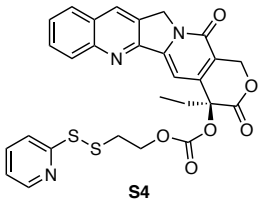
LCMS (C_4, 1-91% MeCN, overlay with blank chromatogram)



S4
$C_{28}H_{23}N_3O_6S_2$
M = 561.63 g/mol

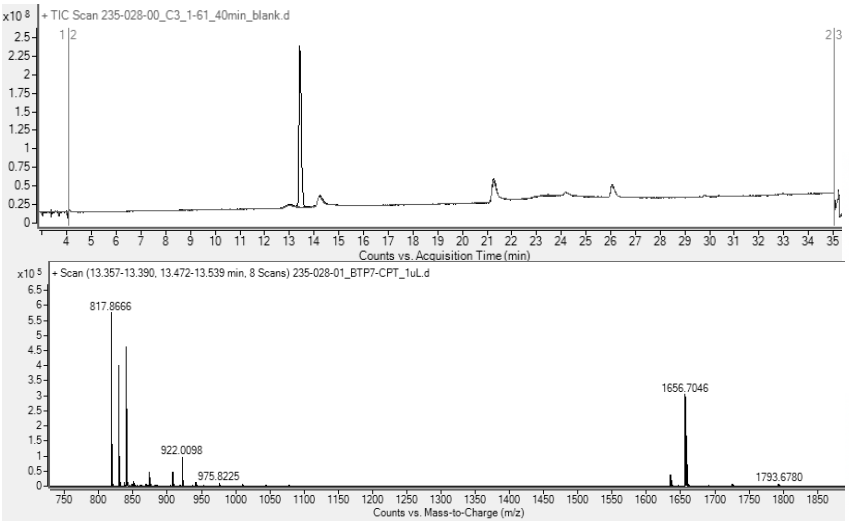


HRMS (ESI)



S4
$C_{28}H_{23}N_3O_6S_2$
M = 561.63 g/mol

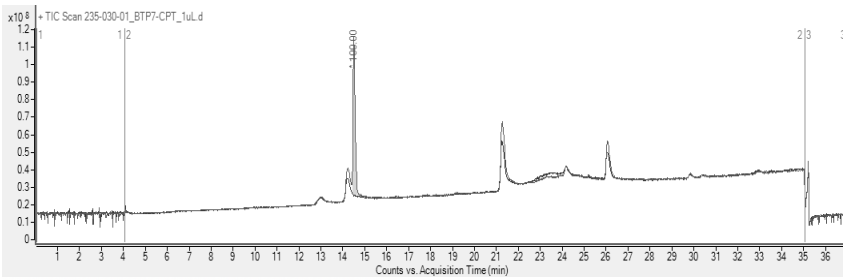
BTP-7-CPT (S6)



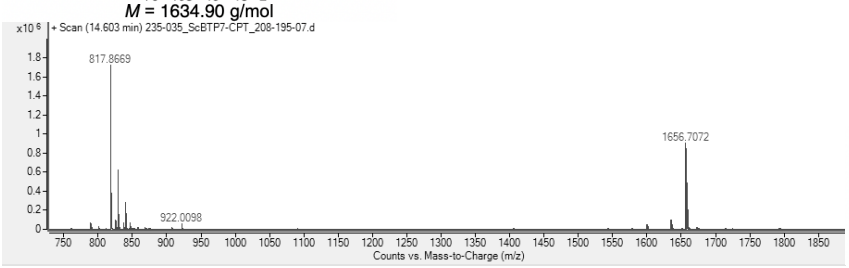
LCMS (C4, 1-61% MeCN, overlayed with blank)
BTP7-Campthotecin
S6
$C_{76}H_{103}N_{19}O_{18}S_2$
M = 1634.90 g/mol

HRMS (ESI)
BTP7-Campthotecin
S6
$C_{76}H_{103}N_{19}O_{18}S_2$
M = 1634.90 g/mol

Scramble Scr-7-CPT (S8)



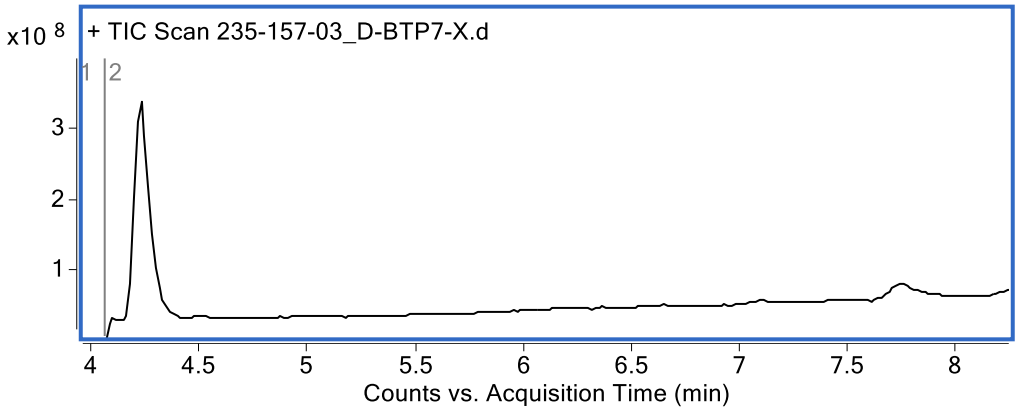
LCMS (C4, 1-61% MeCN, overlayed with blank)
Scrambled BTP7-Campthotecin
S8
$C_{76}H_{103}N_{19}O_{18}S_2$
M = 1634.90 g/mol

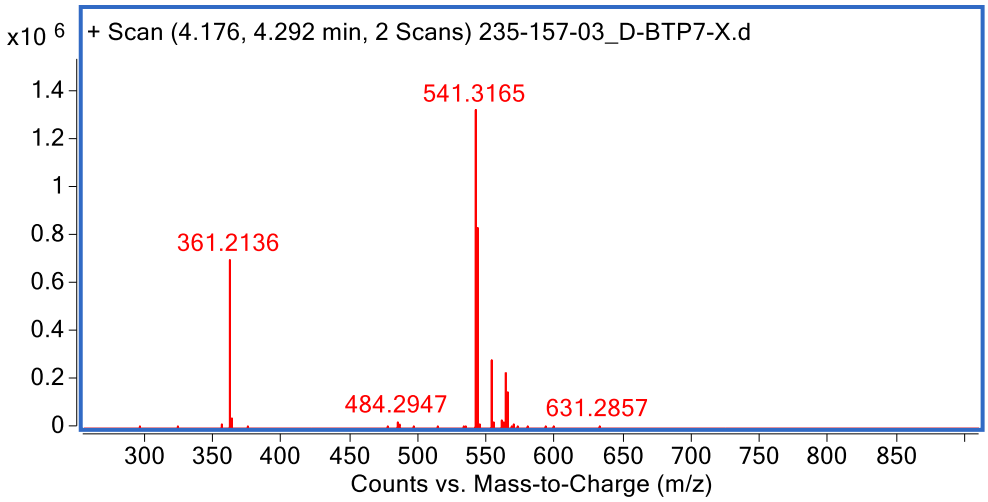


HRMS (ESI)
Scrambled BTP7-Campthotecin
S8
$C_{76}H_{103}N_{19}O_{18}S_2$
M = 1634.90 g/mol

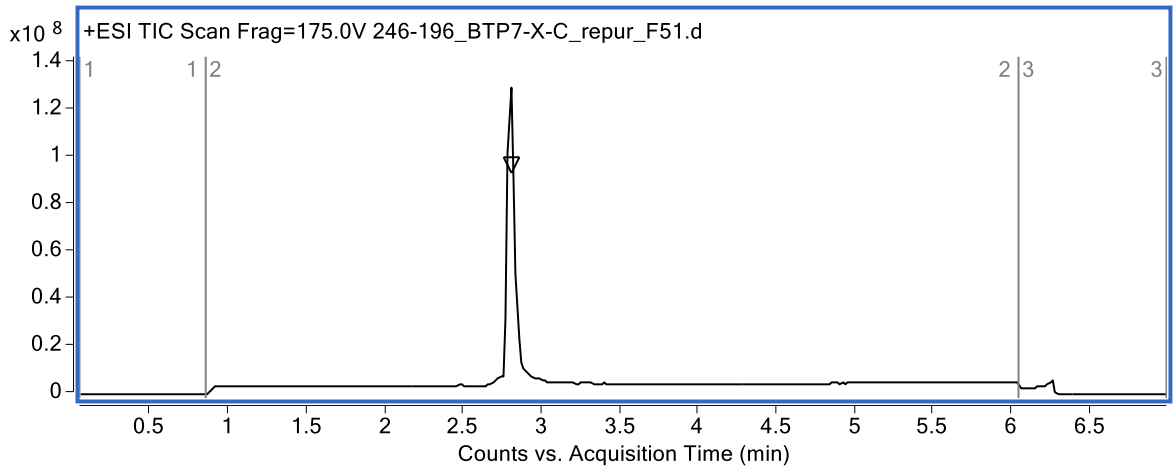
Peptide: BTP-7-X (X = aminohexanoic acid)

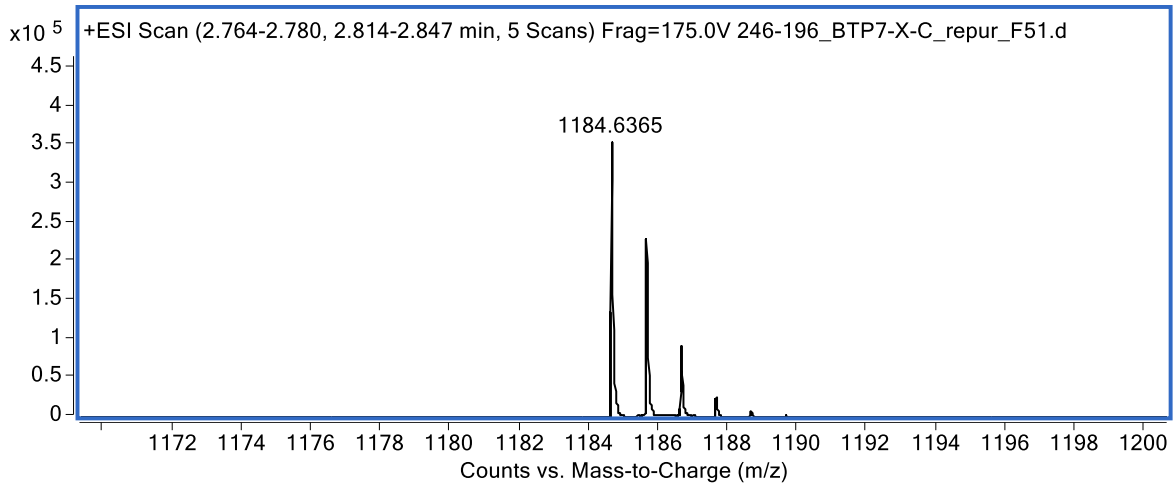
Sequence: tkwGhvnk-X
Mass expected: 1080.6 Da
Mass observed: 1080.6 Da





Peptide: BTP-7-X-C (X = aminohexanoic acid)
Sequence: tkwGhvnk-XC
Mass expected: 1183.6 Da
Mass observed: 1183.6 Da




+ESI Scan (2.764-2.780, 2.814-2.847 min, 5 Scans) Frag=175.0V 246-196_BTP7-X-C_repur_F51.d
x10 5
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
1184.6365
1172
1174
1176
1178
1180
1182
1184
1186
1188
1190
1192
1194
1196
1198
1200
Counts vs. Mass-to-Charge (m/z)

4.9 References

(1) Ostrom, Q.T.; Gittleman, H.; Truitt, G.; Boscia, A.; Kruchko, C.; Barnholtz-Sloan, J.S. CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2011–2015. *Neuro-Oncol.* **2018**, *20*, iv1–iv86. https://doi.org/10.1093/neuonc/noy131.

(2) Jain, K.K. A Critical Overview of Targeted Therapies for Glioblastoma. *Front. Oncol.* **2018**, *8*, 419. https://doi.org/10.3389/fonc.2018.00419.

(3) Louis, D.N. Molecular Pathology of Malignant Gliomas. *Annu. Rev. Pathol.* **2006**, *1*, 97–117. https://doi.org/10.1146/annurev.pathol.1.110304.100043.

(4) Löscher, W.; Potschka, H. Drug Resistance in Brain Diseases and the Role of Drug Efflux Transporters. *Nat. Rev. Neurosci.* **2005**, *6*, 591–602. https://doi.org/10.1038/nrn1728.

(5) Gala, U.H.; Miller, D.A.; Williams, R.O. Harnessing the Therapeutic Potential of Anticancer Drugs through Amorphous Solid Dispersions. *Biochim. Biophys. Acta (BBA)-Rev. Cancer* **2020**, *1873*, 188319. https://doi.org/10.1016/j.bbcan.2019.188319.

(6) Touat, M.; Idbaih, A.; Sanson, M.; Ligon, K.L. Glioblastoma Targeted Therapy: Updated Approaches from Recent Biological Insights. *Ann. Oncol.* **2017**, *28*, 1457–1472. https://doi.org/10.1093/annonc/mdx106.

(7) Patel, A.P.; Tirosh, I.; Trombetta, J.J.; Shalek, A.K.; Gillespie, S.M.; Wakimoto, H.; Cahill, D.P.; Nahed, B.V.; Curry, W.T.; Martuza, R.L.; et al. Single-Cell RNA-Seq Highlights Intratumoral Heterogeneity in Primary Glioblastoma. *Science* **2014**, *344*, 1396–1401. https://doi.org/10.1126/science.1254257.

(8) Raavé, R.; van Kuppevelt, T.H.; Daamen, W.F. Chemotherapeutic Drug Delivery by Tumoral Extracellular Matrix Targeting. *J Control Release* **2018**, *274*, 1–8. https://doi.org/10.1016/j.jconrel.2018.01.029.

(9) Jaworski, D.M.; Kelly, G.M.; Hockfield, S. BEHAB, a New Member of the Proteoglycan Tandem Repeat Family of Hyaluronan-Binding Proteins That Is Restricted to the Brain. *J. Cell Biol.* **1994**, *125*, 495–509.

(10) Viapiano, M.S.; Bi, W.L.; Piepmeier, J.; Hockfield, S.; Matthews, R.T. Novel Tumor-Specific Isoforms of BEHAB/Brevican Identified in Human Malignant Gliomas. *Cancer Res.* **2005**, *65*, 6726–6733. https://doi.org/10.1158/0008-5472.CAN-05-0585.

(11) Jaworski, D.M.; Kelly, G.M.; Piepmeier, J.M.; Hockfield, S. BEHAB (Brain Enriched Hyaluronan Binding) Is Expressed in Surgical Samples of Glioma and in Intracranial Grafts of Invasive Glioma Cell Lines. *Cancer Res.* **1996**, *56*, 2293–2298.

(12) Lu, R.; Wu, C.; Guo, L.; Liu, Y.; Mo, W.; Wang, H.; Ding, J.; Wong, E.T.; Yu, M. The Role of Brevican in Glioma: Promoting Tumor Cell Motility in Vitro and in Vivo. *BMC Cancer* **2012**, *12*, 607. https://doi.org/10.1186/1471-2407-12-607.

(13) Spreckelsen, N.; Fadzen, C.M.; Hartrampf, N.; Ghotmi, Y.; Wolfe, J.M.; Dubey, S.; Yang, B.Y.; Kijewski, M.F.; Wang, S.; Farquhar, C.; et al. Targeting Glioblastoma Using a Novel Peptide Specific to a Deglycosylated Isoform of Brevican. *Adv. Ther.* **2021**, *4*, 2000244. https://doi.org/10.1002/adtp.202000244.

(14) Morgan, M.T.; Nakanishi, Y.; Kroll, D.J.; Griset, A.P.; Carnahan, M.A.; Wathier, M.; Oberlies, N.H.; Manikumar, G.; Wani, M.C.; Grinstaff, M.W. Dendrimer-Encapsulated Camptothecins: Increased Solubility, Cellular Uptake, and Cellular Retention Affords

Enhanced Anticancer Activity in Vitro. *Cancer Res.* **2006**, *66*, 11913–11921. https://doi.org/10.1158/0008-5472.CAN-06-2066.
(15) Oberlies, N.H.; Kroll, D.J. Camptothecin and Taxol: Historic Achievements in Natural Products Research. *J. Nat. Prod.* **2004**, *67*, 129–135. https://doi.org/10.1021/np030498t.
(16) Lee, J.; Kotliarova, S.; Kotliarov, Y.; Li, A.; Su, Q.; Donin, N.M.; Pastorino, S.; Purow, B.W.; Christopher, N.; Zhang, W.; et al. Tumor Stem Cells Derived from Glioblastomas Cultured in BFGF and EGF More Closely Mirror the Phenotype and Genotype of Primary Tumors than Do Serum-Cultured Cell Lines. *Cancer Cell* **2006**, *9*, 391–403. https://doi.org/10.1016/j.ccr.2006.03.030.
(17) Henne, W.A.; Doorneweerd, D.D.; Hilgenbrink, A.R.; Kularatne, S.A.; Low, P.S. Synthesis and Activity of a Folate Peptide Camptothecin Prodrug. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 5350–5355. https://doi.org/10.1016/j.bmcl.2006.07.076.
(18) Giannini, C.; Sarkaria, J.N.; Saito, A.; Uhm, J.H.; Galanis, E.; Carlson, B.L.; Schroeder, M.A.; James, C.D. Patient Tumor EGFR and PDGFRA Gene Amplifications Retained in an Invasive Intracranial Xenograft Model of Glioblastoma Multiforme. *Neuro-Oncol.* **2005**, *7*, 164–176. https://doi.org/10.1215/S1152851704000821.
(19) Cho, C.-F.; Wolfe, J.M.; Fadzen, C.M.; Calligaris, D.; Hornburg, K.; Chiocca, E.A.; Agar, N.Y.R.; Pentelute, B.L.; Lawler, S.E. Blood-Brain-Barrier Spheroids as an in Vitro Screening Platform for Brain-Penetrating Agents. *Nat. Commun.* **2017**, *8*, 15623. https://doi.org/10.1038/ncomms15623.
(20) Bergmann, S.; Lawler, S.E.; Qu, Y.; Fadzen, C.M.; Wolfe, J.M.; Regan, M.S.; Pentelute, B.L.; Agar, N.Y.R.; Cho, C.-F. Blood-Brain-Barrier Organoids for Investigating the Permeability of CNS Therapeutics. *Nat. Protoc.* **2018**, *13*, 2827–2843. https://doi.org/10.1038/s41596-018-0066-x.
(21) Saito, G.; Swanson, J.A.; Lee, K.-D. Drug Delivery Strategy Utilizing Conjugation via Reversible Disulfide Linkages: Role and Site of Cellular Reducing Activities. *Adv. Drug Deliv. Rev.* **2003**, *55*, 199–215.
(22) Bailly, C. Irinotecan: 25 Years of Cancer Treatment. *Pharmacol. Res.* **2019**, *148*, 104398. https://doi.org/10.1016/j.phrs.2019.104398.
(23) Giannini, C.; Sarkaria, J.N.; Saito, A.; Uhm, J.H.; Galanis, E.; Carlson, B.L.; Schroeder, M.A.; James, C.D. Patient Tumor EGFR and PDGFRA Gene Amplifications Retained in an Invasive Intracranial Xenograft Model of Glioblastoma Multiforme. *Neuro-Oncol.* **2005**, *7*, 164–176. https://doi.org/10.1215/S1152851704000821.
(24) Friedman, G.K.; Haas, M.C.; Kelly, V.M.; Markert, J.M.; Gillespie, G.Y.; Cassady, K.A. Hypoxia Moderates γ(1)34.5-Deleted Herpes Simplex Virus Oncolytic Activity in Human Glioma Xenoline Primary Cultures. *Transl. Oncol.* **2012**, *5*, 200–207. https://doi.org/10.1593/tlo.12115.
(25) Lu, J.; Liu, C.; Wang, P.; Ghazwani, M.; Xu, J.; Huang, Y.; Ma, X.; Zhang, P.; Li, S. The Self-Assembling Camptothecin-Tocopherol Prodrug: An Effective Approach for Formulating Camptothecin. *Biomaterials* **2015**, *62*, 176–187. https://doi.org/10.1016/j.biomaterials.2015.05.046.
(26) Goldwirt, L.; Beccaria, K.; Carpentier, A.; Farinotti, R.; Fernandez, C. Irinotecan and Temozolomide Brain Distribution: A Focus on ABCB1. *Cancer Chemother. Pharm.* **2014**, *74*, 185–193. https://doi.org/10.1007/s00280-014-2490-0.
(27) Polivka, J.; Polivka, J.; Holubec, L.; Kubikova, T.; Priban, V.; Hes, O.; Pivovarcikova, K.; Treskova, I. Advances in Experimental Targeted Therapy and Immunotherapy for Patients

with Glioblastoma Multiforme. *Anticancer Res.* **2017**, *37*, 21–33. https://doi.org/10.21873/anticanres.11285.
(28) Zhou, Y.; Wu, W.; Bi, H.; Yang, D.; Zhang, C. Glioblastoma Precision Therapy: From the Bench to the Clinic. *Cancer Lett.* **2020**, *475*, 79–91. https://doi.org/10.1016/j.canlet.2020.01.027.
(29) Johnson, B.E.; Mazor, T.; Hong, C.; Barnes, M.; Aihara, K.; McLean, C.Y.; Fouse, S.D.; Yamamoto, S.; Ueda, H.; Tatsuno, K.; et al. Mutational Analysis Reveals the Origin and Therapy-Driven Evolution of Recurrent Glioma. *Science* **2014**, *343*, 189–193. https://doi.org/10.1126/science.1239947.
(30) Frischknecht, R.; Seidenbecher, C.I. Brevican: A Key Proteoglycan in the Perisynaptic Extracellular Matrix of the Brain. *Int. J. Biochem. Cell Biol.* **2012**, *44*, 1051–1054.
(31) Hu, B.; Kong, L.L.; Matthews, R.T.; Viapiano, M.S. The Proteoglycan Brevican Binds to Fibronectin after Proteolytic Cleavage and Promotes Glioma Cell Motility. *J. Biol. Chem.* **2008**, *283*, 24848–24859. https://doi.org/10.1074/jbc.M801433200.
(32) Dwyer, C.A.; Bi, W.L.; Viapiano, M.S.; Matthews, R.T. Brevican Knockdown Reduces Late-Stage Glioma Tumor Aggressiveness. *J. Neurooncol.* **2014**, *120*, 63–72. https://doi.org/10.1007/s11060-014-1541-z.
(33) Tsidulko, A.Y.; Kazanskaya, G.M.; Volkov, A.M.; Suhovskih, A.V.; Kiselev, R.S.; Kobozev, V.V.; Gaytan, A.S.; Krivoshapkin, A.L.; Aidagulova, S.V.; Grigorieva, E.V. Chondroitin Sulfate Content and Decorin Expression in Glioblastoma Are Associated with Proliferative Activity of Glioma Cells and Disease Prognosis. *Cell Tissue Res* **2020**, *379*, 147–155. https://doi.org/10.1007/s00441-019-03127-2.
(34) Virga, J.; Bognár, L.; Hortobágyi, T.; Zahuczky, G.; Csősz, É.; Kalló, G.; Tóth, J.; Hutóczki, G.; Reményi-Puskár, J.; Steiner, L.; et al. Prognostic Role of the Expression of Invasion-Related Molecules in Glioblastoma. *J. Neurol. Surg. A Cent. Eur. Neurosurg.* **2016**, *78*, 12–19. https://doi.org/10.1055/s-0036-1584920.
(35) Viapiano, M.S.; Matthews, R.T. From Barriers to Bridges: Chondroitin Sulfate Proteoglycans in Neuropathology. *Trends Mol. Med.* **2006**, *12*, 488–496.
(36) Viapiano, M.S.; Hockfield, S.; Matthews, R.T. BEHAB/Brevican Requires ADAMTS-Mediated Proteolytic Cleavage to Promote Glioma Invasion. *J. Neurooncol.* **2008**, *88*, 261–272. https://doi.org/10.1007/s11060-008-9575-8.
(37) Hu, B.; Kong, L.L.; Matthews, R.T.; Viapiano, M.S. The Proteoglycan Brevican Binds to Fibronectin after Proteolytic Cleavage and Promotes Glioma Cell Motility. *J. Biol. Chem.* **2008**, *283*, 24848–24859. https://doi.org/10.1074/jbc.M801433200.
(38) Glass, R.; Synowitz, M.; Kronenberg, G.; Walzlein, J.-H.; Markovic, D.S.; Wang, L.-P.; Gast, D.; Kiwit, J.; Kempermann, G.; Kettenmann, H. Glioblastoma-Induced Attraction of Endogenous Neural Precursor Cells Is Associated with Improved Survival. *J. Neurosci.* **2005**, *25*, 2637–2646.
(39) Grigorieva, E.V. Radiation Effects on Brain Extracellular Matrix. *Front. Oncol.* **2020**, *10*, 576701. https://doi.org/10.3389/fonc.2020.576701.
(40) Gupta, K.; Burns, T.C. Radiation-Induced Alterations in the Recurrent Glioblastoma Microenvironment: Therapeutic Implications. *Front. Oncol.* **2018**, *8*, 503. https://doi.org/10.3389/fonc.2018.00503.
(41) Virga, J.; Szivos, L.; Hortobágyi, T.; Chalsaraei, M.K.; Zahuczky, G.; Steiner, L.; Tóth, J.; Reményi-Puskár, J.; Bognár, L.; Klekner, A. Extracellular Matrix Differences in

Glioblastoma Patients with Different Prognoses. *Oncol. Lett.* **2019**, *17*, 797–806. https://doi.org/10.3892/ol.2018.9649.
(42) Ladner, R.C.; Sato, A.K.; Gorzelany, J.; Souza, M. Phage Display-Derived Peptides as Therapeutic Alternatives to Antibodies. *Drug Discov. Today* **2004**, *9*, 525–529. https://doi.org/10.1016/S1359-6446(04)03104-6.
(43) Zhang, X.-X.; Eden, H.S.; Chen, X. Peptides in Cancer Nanomedicine: Drug Carriers, Targeting Ligands and Protease Substrates. *J. Control. Release* **2012**, *159*, 2–13. https://doi.org/10.1016/j.jconrel.2011.10.023.
(44) Kruger, R.P. The Coming Peptide Tidal Wave. *Cell* **2017**, *171*, 497. https://doi.org/10.1016/j.cell.2017.10.010.
(45) Venditto, V.J.; Simanek, E.E. Cancer Therapies Utilizing the Camptothecins: A Review of the in Vivo Literature. *Mol. Pharm.* **2010**, *7*, 307–349. https://doi.org/10.1021/mp900243b.
(46) Kalim, M.; Chen, J.; Wang, S.; Lin, C.; Ullah, S.; Liang, K.; Ding, Q.; Chen, S.; Zhan, J. Intracellular Trafficking of New Anticancer Therapeutics: Antibody–Drug Conjugates. *Drug Des. Dev. Ther.* **2017**, *11*, 2265.
(47) Doronina, S.O.; Toki, B.E.; Torgov, M.Y.; Mendelsohn, B.A.; Cerveny, C.G.; Chace, D.F.; DeBlanc, R.L.; Gearing, R.P.; Bovee, T.D.; Siegall, C.B. Development of Potent Monoclonal Antibody Auristatin Conjugates for Cancer Therapy. *Nat. Biotechnol.* **2003**, *21*, 778–784.
(48) Bargh, J.D.; Isidro-Llobet, A.; Parker, J.S.; Spring, D.R. Cleavable Linkers in Antibody-Drug Conjugates. *Chem. Soc. Rev.* **2019**, *48*, 4361–4374. https://doi.org/10.1039/c8cs00676h.

Chapter 5. Automated Fast-Flow Synthesis of Chromosome 9 Open Reading Frame 72 Dipeptide Repeat Proteins

The work presented in this chapter has been reproduced from the following manuscript:

Sato, K.; Farquhar, C.E.; Rodriguez, J.J.L.; Pentelute, B.L. Automated fast-flow synthesis of Chromosome 9 open reading frame 72 dipeptide repeat proteins. *J. Am. Chem. Soc.* **2023**, *145* (24), 12992–12997.

5.1 Introduction

Amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) are refractory neurodegenerative diseases in which ALS causes progressive muscle paralysis, and FTD causes personality and behavioral changes.[1] These two diseases share many pathologic, genetic, and clinical features, where half of ALS patients are estimated to develop the aspects of FTD.[2] The most common genetic cause of ALS and FTD is a mutation in the chromosome 9 open frame 72 (*c9orf72*) gene with the expansion of the hexanucleotide (GGGGCC) repeat sequence.[3,4] The mutated sequence is further translated by the non-canonical mechanism to produce five dipeptide repeat proteins (DPRs): poly-glycine-arginine (poly-GR), poly-proline-arginine (poly-PR), poly-glycine-proline (poly-GP), poly-proline-alanine (poly-PA), and poly-glycine-alanine (poly-GA).[5,6]

Since the discovery of DPRs, their neurogenerative properties have been extensively studied using DPR-expressing disease models.[7–11] However, the fundamental physicochemical properties of DPRs remain largely unknown due to their limited availability. Although recombinant expression has been the most powerful approach to obtaining proteins, these techniques are limited by scale. In addition, the expression of toxic peptides and proteins, such as highly cationic ones, is often challenging.[12,13] Therefore, a synthetic approach to obtain such DPRs is of great importance to further understand the pathogenic mechanisms of ALS/FTD and to develop new therapies targeting DPRs. In particular, the chemical synthesis of DPRs longer than 36 repeats (72 amino acids) is crucial, as this has been described as the minimum length to exhibit toxicity *in vivo*.[8,11] Although several groups have reported the chemical synthesis of DPRs, their maximum length is limited to up to 30 repeats (60 amino acids),[14–16] which is reasonable considering the general difficulty of synthesizing peptides longer than 50 amino acids by standard solid-phase peptide synthesis (SPPS).[17] Additionally, chemoselective ligation reactions would not be readily applicable to the synthesis of long DPRs due to the lack of reactive residues.[18] We thus set the goal of this work to overcome these problems and to establish the chemical synthesis of long DPRs to investigate their properties.

5.2 Results and Discussion

5.2.1 Synthesis of the DPRs

Our research group has recently developed an automated fast-flow peptide synthesis (AFPS) system that can complete a single amide coupling reaction in 40 seconds and allow the direct production of peptide chains up to 164 amino acids in hours (**Fig. 5.1a**).[19–22] Here, we performed AFPS of DPRs on a Rink amide resin following the previously optimized synthetic conditions.[20] As activation reagents, we used hexafluorophosphate azabenzotriazole tetramethyl uronium (HATU) for coupling glycine and proline, and (7-azabenzotriazol-1-yloxy)trispyrrolidinophosphonium hexafluorophosphate (PyAOP) for alanine and arginine with extended coupling time. Removal of fluorenylmethylmethoxy-carbonyl (Fmoc) groups was performed using piperidine, and the deprotection reaction was monitored by in-line ultraviolet-visible detection of the eluent (**Fig. 5.1b**).

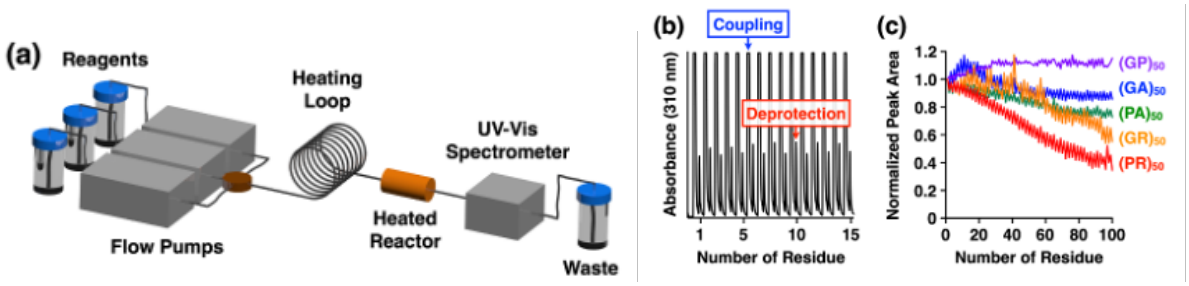


Figure 5.1 Rapid flow synthesis of dipeptide repeat proteins
(a) Schematic illustration of the automated fast-flow peptide synthesizer (AFPS) used in this study. (b) Absorbance recorded at 310 nm during the synthesis of poly-GP. (c) Normalized peak area of Fmoc deprotections obtained during the synthesis of five DPRs with 50 repeats (100 amino acids).

Fig. 5.1c shows the normalized area of the deprotection peaks during the synthesis of five DPRs with 50 repeats (100 amino acids). We found that the arginine-containing DPRs **(GR)$_{50}$** and **(PR)$_{50}$** showed relatively smaller peak areas compared to the DPRs variants without arginine **(GP)$_{50}$**, **(GA)$_{50}$**, and **(PA)$_{50}$**, indicating the low coupling efficiency of arginine residues. This observation is consistent with our previous machine learning study showing that arginine residues can have reduced coupling efficiency, probably due to their bulky aromatic 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl (Pbf) protective groups.[23] Nevertheless, AFPS followed by purification using reverse-phase high-performance liquid chromatography (RP-HPLC) allowed the isolation of **(GR)$_{50}$** and **(PR)$_{50}$** in 2% (7.2 mg) and 1% (4.4 mg) yields,

respectively, which were still sufficient for the subsequent physicochemical and biological studies. The purity of **(GR)$_{50}$**, **(PR)$_{50}$**, and their shorter repeats was confirmed by analytical RP-HPLC and electrospray-ionization time-of-flight (ESI-TOF) mass spectrometry (**Fig. 5.2a,b**). For poly-GP and poly-PA, we were even able to synthesize and isolate **(GP)$_{100}$** and **(PA)$_{100}$** in 3% (11.3 mg) and 5% (23.7 mg) yield, respectively. Since poly-GP and poly-PA were poorly ionizable by ESI, we used matrix-assisted laser desorption ionization (MALDI)-TOF mass spectrometry and analytical RP-HPLC to confirm their purity (**Fig. 5.2c,d**). We also attempted to synthesize poly-GA and isolated **(GA)$_{10}$** and **(GA)$_{20}$**; However, we were unable to purify **(GA)$_{50}$** and **(GA)$_{100}$** due to their poor solubility. However, MALDI-TOF mass spectra of crude samples clearly showed the target mass peaks, indicating that AFPS of poly-GA itself was also successful.

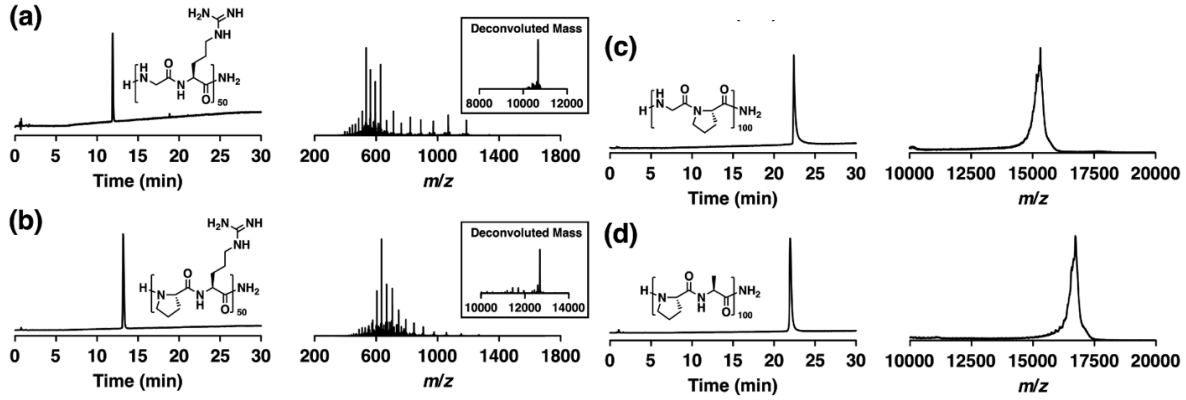


Figure 5.2 Analytical RP-HPLC traces and mass spectra of DPRs.
(a) **(GR)$_{50}$**, (b) **(PR)$_{50}$**, (c) **(GP)$_{100}$**, and (d) **(PA)$_{100}$** (**(GR)$_{50}$** and **(PR)$_{50}$**: ESI-TOF mass spectrometry, **(GP)$_{100}$** and **(PA)$_{100}$**: MALDI-TOF mass spectrometry).

5.2.2 In-vitro structure and aggregation of the synthesized DRPs

With the chemically synthesized DPRs in hand, we then performed circular dichroism (CD) spectroscopy of soluble DPRs (poly-GR, poly-PR, poly-GP, and poly-PA) in phosphate-buffered saline (PBS) at different temperatures to investigate their secondary structures. Previously, the secondary structure of poly-GA has been extensively studied and its propensity to form fibrous aggregates through the formation of β-sheet structures has been reported.[14,24–26] However, the structures of other DPRs have not been studied in detail. As shown in **Fig. 4.3a**, **(GR)$_{50}$** exhibited featureless CD spectra regardless of the measurement temperature, indicating its random coil-rich structure, consistent with previous reports.[10,24]

We then measured CD spectra of **(PR)$_{50}$**, **(GP)$_{100}$**, and **(PA)$_{100}$** at 25 °C (**Fig. 5.3b–3d**, blue curves) and observed a negative peak at 200 nm, which has been previously characterized as flexible or random coil configurations.[10,24,26] However, as we increased the temperature, the peak top showed a red shift associated with the decrease in intensity around 220–230 nm, with an apparent isodichroic point at 215 nm. These results demonstrated that the proline-containing DPRs **(PR)$_{50}$**, **(GP)$_{100}$**, and **(PA)$_{100}$** do not simply adopt disordered structures but show a transition from one state to another upon heating.[27,28] We then found that the observed thermal transition of **(PR)$_{50}$**, **(GP)$_{100}$**, and **(PA)$_{100}$** closely resembles the denaturation profile of collagens.[29] It is noteworthy that approximately 1/3 of the constituent amino acids of collagens are prolines or hydroxyprolines, which contribute to their formation of polyproline II-like helical secondary structures.[29,30] Thus, it is reasonable to assume that **(PR)$_{50}$**, **(GP)$_{100}$**, and **(PA)$_{100}$**, whose half of the constituent amino acids are prolines, also adopt similar secondary structures. In addition, we measured the CD spectra of proline-containing DPRs with shorter repeat lengths and found that they exhibited substantially the same spectral profiles. These results suggest that the secondary structures formed within poly-PR, poly-GP, and poly-PA might be present partially rather than entirely throughout the chains.

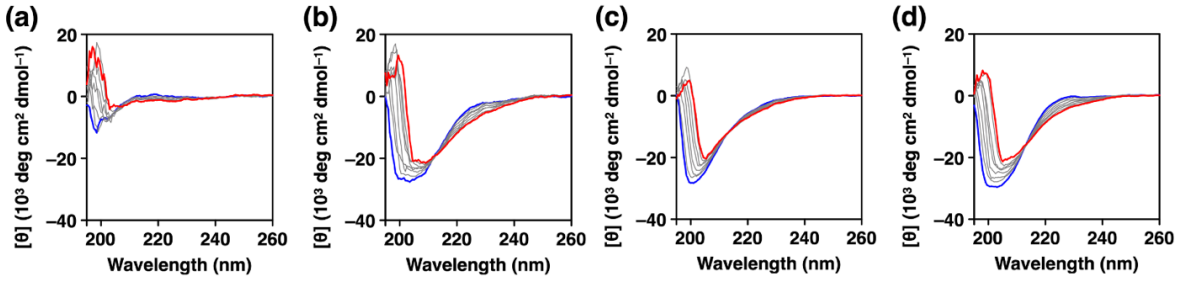


Figure 5.3 CD spectra of the DPRs.
(a) **(GR)$_{50}$**, (b) **(PR)$_{50}$**, (c) **(GP)$_{100}$**, and (d) **(PA)$_{100}$** in PBS ([DPR] = 0.1 mg mL^{-1}) upon heating from 25 °C (blue) to 95 °C (red). Spectra were recorded at every 10 °C.

In addition to the secondary structures, previous studies have also shown that DPRs can self-assemble into aggregates under physiological conditions.[24,26] However, a comprehensive understanding has been lacking due to the unavailability of DPRs with variable repeat lengths. Here, we took advantage of AFPS, which allowed us to obtain the libraries of DPRs, and systematically investigated their aggregation property by size exclusion chromatography (SEC) using PBS as an eluent.

As expected, poly-GP, poly-PA, and poly-PR showed shorter retention times as their molecular weight increased (**Fig. 5.4a–4c**). However, the SEC profiles of poly-GR did not show sharp peaks under the same analytical conditions, likely due to their strong interaction with the SEC matrix (agarose and dextran). We then evaluated the apparent molecular weight of poly-GP, poly-PA, and poly-PR based on the retention times of globular protein standards. As shown in **Fig. 5.4d** (red), poly-PR showed a linear correlation between the derived apparent molecular weight and the actual molecular weight (R^2 = 0.97), indicating that poly-PR exists as a monomeric protein and adopts a globular shape. Interestingly, the apparent molecular weight of poly-GP and poly-PA showed an exponential increase (R^2 = 0.99, **Fig. 5.4d** purple and green). Considering that CD spectral studies of poly-GP and poly-PA did not show any obvious length-dependent changes, the observed non-linear increase in the apparent molecular weight should not be due to the differences in protein folding but most likely to the enhanced aggregation property of longer DPRs through multivalent interactions.[31,32]

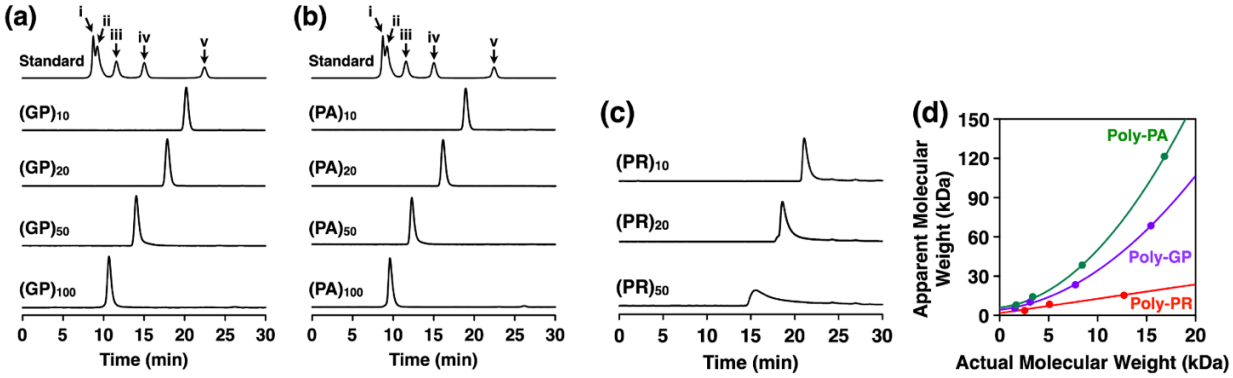


Figure 5.4 SEC traces of the DPRs.
(a) poly-GPs, (b) poly-PAs, and (c) poly-PRs at 25 °C ([DPR] = 1.0 mg mL–1 in PBS, eluent: PBS). The molecular weights of the globular protein standards are (i) 670 kDa, (ii) 158 kDa, (iii) 44 kDa, (iv) 17 kDa, and (v) 1.35 kDa, respectively. (d) Plots and their corresponding fitted curves of the actual molecular weight of DPRs as a function of their derived apparent molecular weight using protein standards.

5.2.3 Cellular toxicity of exogenous DPRs

Finally, we investigated the cytotoxicity that synthetic DPRs might possess. Following the previously reported protocol,[14] we performed the MTT cell viability assay using human neuroblastoma cells in the presence of synthesized poly-DPRs. We pre-dissolved synthetic DPRs with different repeat length and concentration in cell media, added the solution of DPRs to BE(2)-

C cells and incubated for 48 h, and then quantified the cell viability relative to untreated cells. As shown in **Fig. 5.5a** and **5.5b**, poly-GR and poly-PR exhibited lower cell viability with increasing repeat length and concentration. In contrast, poly-GP and poly-PA did not show significant changes in cell viability regardless of their repeat length and concentration (**Fig. 5.5c,d**). This tendency is consistent with previous reports on endogenous DPRs expressed *in vivo*,[8,9] indicating that the synthetic DPRs successfully reproduced their intrinsic cytotoxicity.

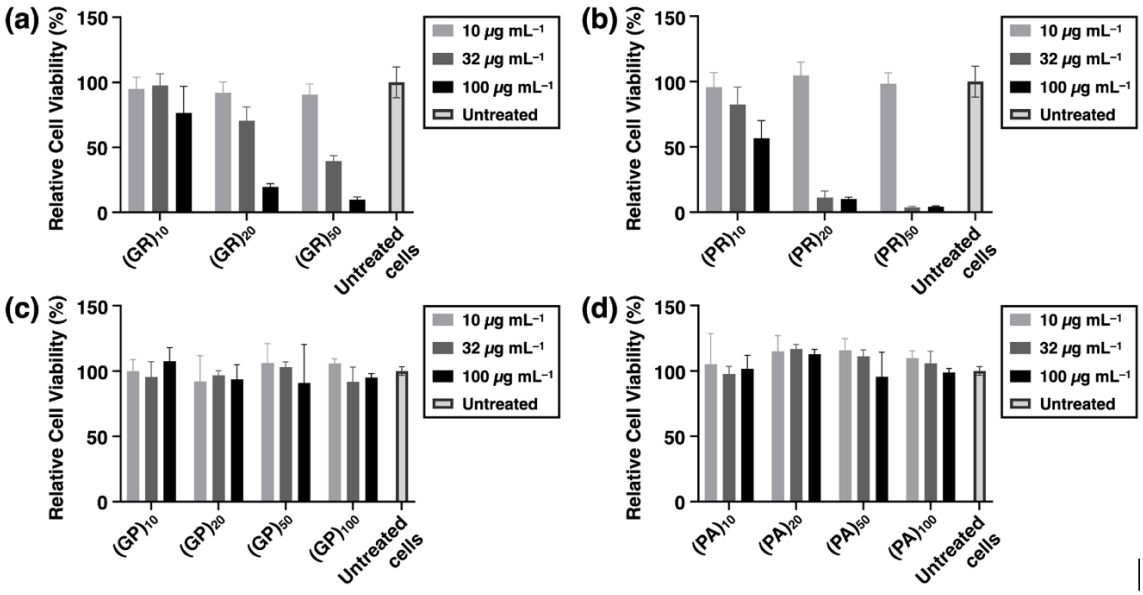


Figure 5.5 Relative viability of human neuroblastoma BE(2)-C cells 48 h after treatment with DPRs.

Cells were treated with media containing (a) poly-GR, (b) poly-PR, (c) poly-GP, and (d) poly-PA with different repeat length and concentrations. The cell viability was quantified by MTT assay as percentage viable relative to untreated cells. Bars represent the standard deviation of the mean based on n=2 replicates

This tendency of the arginine-rich DPRs towards toxicity was further supported by a lactate dehydrogenase (LDH) release assay, which tests membrane integrity by measuring the amount of the cytosolic protein LDH in the external cell medium. As shown in **Fig. 5.6a and 5.6b**, poly-GR and poly-PR once again exhibited dose- and length- dependent toxicity. In contrast, the poly GP and poly-PA compounds showed no significant damage to the cell membrane.

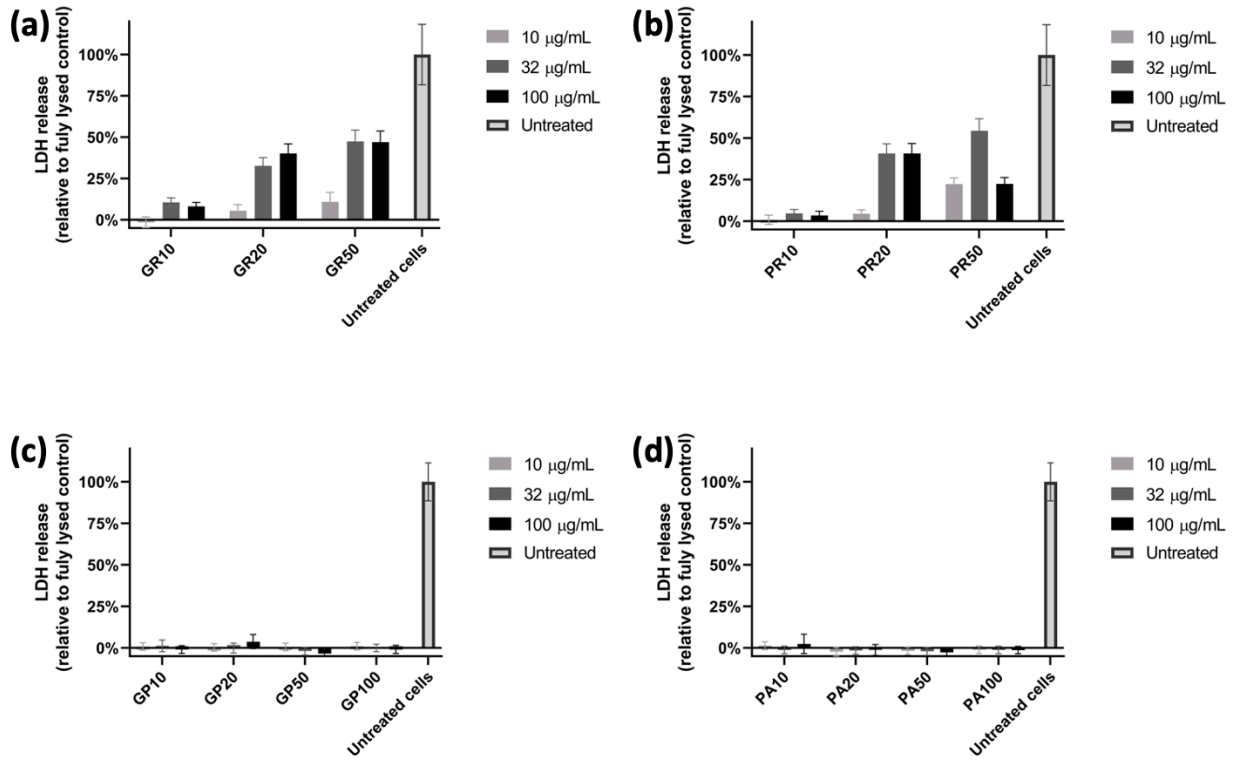


Figure 5.6 Relative membrane integrity of human neuroblastoma BE(2)-C cells 8 h after treatment with DPRs.

Cells were treated with media containing (a) poly-GR, (b) poly-PR, (c) poly-GP, and (d) poly-PA with different repeat length and concentrations. 8 hours after treatment, cell media was removed and LDH protein levels were measured. Results are given as LDH release above vehicle relative to fully lysed cells.

While the membrane leakage demonstrated in the LDH assay can correlate with overall reduced cell viability, it can also represent acute toxicity at the cell membrane.[33] To further investigate this possibility, we repeated the LDH assay at two hours after treatment. The poly-GR peptides showed significant membrane disruption at 2-hours (**Fig 5.7a**), indicating an acute toxicity effect, and likely a disruption of the cell membrane. This effect was also present for the longer poly-PR compounds (**Fig 5.7b**). Intriguingly, by the long-term cell viability measurement in **Fig. 5.5b** the poly-PR series appears more detrimental to cell viability, while the poly-GR peptides seem to have a more acute toxicity at the membrane (**Fig 5.7a**). Arginine-rich DPRs have previously demonstrated the ability to damage neighboring, healthy cells.[8-9, 34] While previous studies have focused on the DPRs entering cells via endocytosis, these results suggest that DPRs may have an additional toxic effect directly at the plasma membrane.

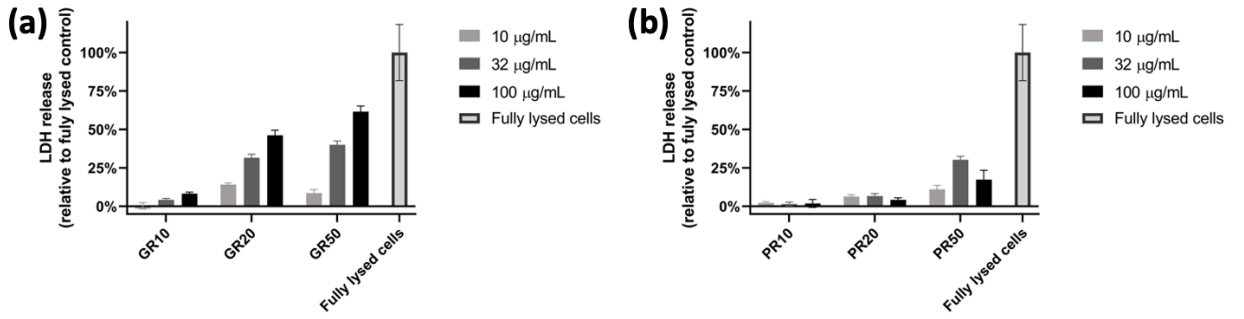


Figure 5.7 Relative membrane integrity of human neuroblastoma BE(2)-C cells 2 h after treatment with DPRs.

Cells were treated with media containing (a) poly-GR and (b) poly-PR with different repeat length and concentrations. 2 hours after treatment, cell media was removed and LDH protein levels were measured. Results are given as LDH release above vehicle relative to fully lysed cells.

5.3 Conclusions

In conclusion, we have successfully synthesized *c9orf72* DPRs with up to 100 repeats (200 amino acids) using AFPS. CD spectroscopy of synthetic DPRs revealed that poly-PR, poly-GP, and poly-PA might adopt polyproline II-like helical secondary structures. Furthermore, SEC analysis of synthetic DPRs showed that poly-GP and poly-PA form larger aggregates as their repeat length increases. Finally, a cell viability assay using human neuroblastoma cells demonstrated the increased cytotoxicity of longer poly-GR and poly-PR. The LDH membrane permeability assay further demonstrated toxicity at the cell membrane on a much shorter timescale. We believe that this research has demonstrated the potential of AFPS to synthesize peptides and proteins with limited availability, which is necessary to study their pathogenic mechanisms in detail and to construct disease models for drug discovery.

5.4 Experimental Section

5.4.1 Automated fast-flow synthesis of DPRs

Automated fast-flow synthesis of DPRs was performed using Fmoc-protected amino acids (Fmoc-Gly-OH, Fmoc-Ala-OH, Fmoc-Pro-OH, and Fmoc-Arg(Pbf)-OH) on H-Rink Amide resin (ChemMatrix, 0.17 mmol g–1 loading, typical scale: 150 mg, 0.0255 mmol), following the previously established methods:[20] flow-rate = 40 mL min–1, temperature = 90 °C, and 85−90 °C (reactor). The 50 mL min–1 pump head pumps 400 μL of liquid per pump stroke; the 5 mL min–1 pump head pumps 40 μL of liquid per pump stroke. The standard synthetic cycle involves a first step of prewashing the resin at elevated temperatures for 60 s at 40 mL min–1. During the coupling step, three HPLC pumps are used: a 50 mL min–1 pump head pumps the activating agent, a second 50 mL min–1 pump head pumps the amino acid, and a 5 mL min–1 pump head pumps DIEA. The first two pumps are activated for 8 pumping strokes in order to prime the coupling agent and amino acid before the DIEA pump is activated. The three pumps are then actuated together for a period of 7 pumping strokes, after which the activating agent pump and amino acid pump are switched using a rotary valve to select DMF. The three pumps are actuated together for a final 8 pumping strokes, after which the DIEA pump is shut off, and the other two pumps continue to wash the resin for another 40 pump strokes. During the deprotection step, two HPLC pumps are used. Using a rotary valve, one HPLC pump selects deprotection stock solution and DMF. The pumps are activated for 13 pump strokes. Both solutions are mixed in a 1:1 ratio. Next, the rotary valves select DMF for both HPLC pumps, and the resin is washed for an additional 40 pump strokes. The coupling−deprotection cycle is repeated for all additional monomers.

5.4.2 Cleavage of DPRs from resin

After synthesis, the peptidyl resin was washed with dichloromethane (3 × 5 mL) and dried in a vacuum chamber overnight, and weighed. Approximately 10 mL of cleavage solution (95% TFA, 2.5% water, 2.5% TIPS) was added to the peptidyl resin inside the fritted syringe. The cleavage was kept at room temperature for 2–4 h, with occasional shaking. After this time, the cleavage mixture was transferred to a tube (through the syringe frit, keeping the resin in the syringe), and the resin was washed with an additional 5 mL of TFA. Ice cold diethyl ether (90 mL) was added to the cleavage mixture, and the precipitate was collected by centrifugation. After the supernatant was discarded, residual ether was allowed to evaporate, and the peptide was dissolved in 50% acetonitrile in water with 0.1% TFA, frozen, and then lyophilized until dry.

5.4.3 Preparative RP-HPLC of DPRs

Preparative reverse-phase high-performance liquid chromatography (RP-HPLC) was performed at 25 °C on an Agilent model 1260 Infinity system equipped with 1260 Prep Pumps, 1260 MWD VL, 1260 FC-PS, and 6130 Quadrupole LC/MS using H_2O/CH_3CN gradient containing 0.1% TFA (v/v) as an eluent. Purification of **$(GR)_{10}$**, **$(GR)_{20}$**, **$(GR)_{50}$**, **$(PR)_{10}$**, **$(PR)_{20}$**, and **$(PR)_{50}$** was performed using an ACE 5 C18-AR column (10.0 x 250) at a flow rate of 4.0 mL min^{-1}. Up to 40 mg of crude dipeptide repeat proteins were dissolved in deionized H_2O containing 0.1% TFA (v/v, 5.0 mL) and filtered through 0.22 μm nylon filter prior to the injection.

Purification of **$(GP)_{10}$**, **$(GP)_{20}$**, **$(GP)_{50}$**, **$(PA)_{10}$**, **$(PA)_{20}$**, and **$(PA)_{50}$** was performed using a ZORBAX PrepHT 300SB-C18 column (C_{18} stationary phase, 7 μm, 300 Å pore size, 21.2 x 250 mm) at a flow rate of 20.0 mL min^{-1}. Up to 100 mg of crude dipeptide repeat proteins were dissolved in deionized H_2O containing 0.1% TFA (v/v, 5.0 mL) and filtered through 0.22 μm nylon filter prior to the injection.

Purification of **$(GP)_{100}$** and **$(PA)_{100}$** was performed using a ZORBAX 300SB-C3 column (C_3 stationary phase, 5 μm, 300 Å pore size, 9.4 x 250 mm) at a flow rate of 4.0 mL min^{-1}. Up to 100 mg of crude dipeptide repeat proteins were dissolved in deionized H_2O containing 0.1% TFA (v/v, 5.0 mL) and filtered through 0.22 μm nylon filter prior to the injection.

Purification of **$(GA)_{10}$** and **$(GA)_{20}$** was performed using a ZORBAX PrepHT 300SB-C18 column (C_{18} stationary phase, 7 μm, 300 Å pore size, 21.2 x 250 mm) at a flow rate of 20.0 mL min^{-1}. Up to 100 mg of crude dipeptide repeat proteins were dissolved in 200 μL of hexafluoroisopropanol (HFIP) at 60 °C, diluted with 5.0 mL of aqueous guanidinium chloride (6 M) and filtered through 0.22 μm nylon filter prior to the injection.

5.4.4 Analytical RP-HPLC of DPRs

Analytical RP-HPLC was performed at 25°C on an Agilent model 1200 Series system equipped with G1322A degasser, G1311A quat pump, G1329A ALS, and G1365B MWD using H_2O/CH_3CN gradient containing 0.1% TFA (v/v) as an eluent. Purified **$(GR)_{10}$**, **$(GR)_{20}$**, **$(GR)_{50}$**, **$(PR)_{10}$**, **$(PR)_{20}$**, **$(PR)_{50}$** **$(GP)_{10}$**, **$(GP)_{20}$**, **$(GP)_{50}$**, **$(PA)_{10}$**, **$(PA)_{20}$**, and **$(PA)_{50}$** were dissolved in deionized H_2O containing 0.1% TFA (v/v) (10 μL, 1.0 mg mL^{-1}) and analyzed by Kinetex® LC column (C_{18} stationary phase, 2.6 μm, 100 Å pore size, 2.1 x 100 mm) at a flow rate of 0.375 mL min^{-1}.

For **(GP)$_{100}$** and **(PA)$_{100}$**, purified samples were dissolved in deionized H_2O containing 0.1% TFA (v/v) (10 μL, 1.0 mg mL^{-1}) and analyzed by ZORBAX 300SB-C3 column (C_3 stationary phase, 5 μm, 300 Å pore size, 2.1 x 150 mm) at a flow rate of 0.5 mL min^{-1}.

5.4.5 Mass spectrometry of DPRs

Electrospray ionization (ESI) mass spectrometry was performed on an Agilent model 6550 iFunnel Q-TOF LC/MS spectrometer. Matrix-assisted laser deposition ionization time-of-flight (MALDI-TOF) mass spectrometry was performed on a Brucker autoflex III smartbeam spectrometer.

5.4.6 SEC of DPRs

Size exclusion chromatography (SEC) was performed at 25 °C on an Agilent model 1290 Infinity II system equipped with 1290 MCT, 1260 Bio-inert Pump, 1260 DAD WR, and 1260 Bio-FC-AS using phosphate-buffered saline (PBS) as an eluent. Purified **(GR)$_{10}$**, **(GR)$_{20}$**, **(GR)$_{50}$**, **(PR)$_{10}$**, **(PR)$_{20}$**, **(PR)$_{50}$** **(GP)$_{10}$**, **(GP)$_{20}$**, **(GP)$_{50}$**, **(GP)$_{100}$**, **(PA)$_{10}$**, **(PA)$_{20}$**, **(PA)$_{50}$**, and **(GP)$_{100}$** were dissolved in PBS (100 μL, 1.0 mg mL^{-1}) and analyzed by Superdex^{TM} 75 Increase 12/300 GL column using PBS as an eluent at a flow rate of 0.8 mL min^{-1}.

5.4.7 CD spectroscopy of DPRs

Circular dichroism (CD) spectra were recorded on a JASCO model J-1500 CD spectrometer. Purified **(GR)$_{10}$**, **(GR)$_{20}$**, **(GR)$_{50}$**, **(PR)$_{10}$**, **(PR)$_{20}$**, **(PR)$_{50}$** **(GP)$_{10}$**, **(GP)$_{20}$**, **(GP)$_{50}$**, **(GP)$_{100}$**, **(PA)$_{10}$**, **(PA)$_{20}$**, **(PA)$_{50}$**, and **(GP)$_{100}$** were dissolved in PBS (350 μL, 0.1 mg mL^{-1}) and their CD spectra were recorded using a quartz cell of 0.1 mm optical path length. CD spectra were recorded from 300 to 190 nm with the scanning speed of 50 nm min^{-1} and data integration time of 4 s. Each spectrum was recorded for 3 times and averaged.

5.4.8 MTT assay using human neuroblastoma BE(2)-C cells

Human neuroblastoma BE(2)-C cells (ATCC, catalog no. CRL-2268) were incubated at 37 °C under 5% CO_2 and cultured in a 1:1 mixture of Eagles Minimum Essential Medium (EMEM) and F12 medium with 10% FBS. Twenty-four hours prior to treatment, cells were plated at a density of 40,000 cells/well in a 96-well plate in 1:1 EMEM:F12 with 10% FBS.

The media was discarded before treatment, and cells were washed twice in unsupplemented 1:1 EMEM:F12. The dipeptide repeats were diluted from stock solutions (10mg/mL in PBS) to the target treatment concentrations in unsupplemented medium. Cells were treated in duplicate with 100µL of dipeptide repeat in medium at the indicated concentration (10, 32, or 100 µg/mL), and then incubated for an additional 48 hours.

For the MTT assay, 10uL of 5.5mg/mL MTT solution was added to the cells and incubated for an additional 3 hours. The medium was then removed, and DMSO (20µL) was added to lyse the formazan crystals. The plate was then read on a biotek Epoch Spectrophotometer at 590 nm. The measured background from the plate and DMSO alone was subtracted from each measurement, and % formazan production was calculated as % viability = 100 × Experimental formazan production (OD590) / formazan production in vehicle-treated cells (OD590).

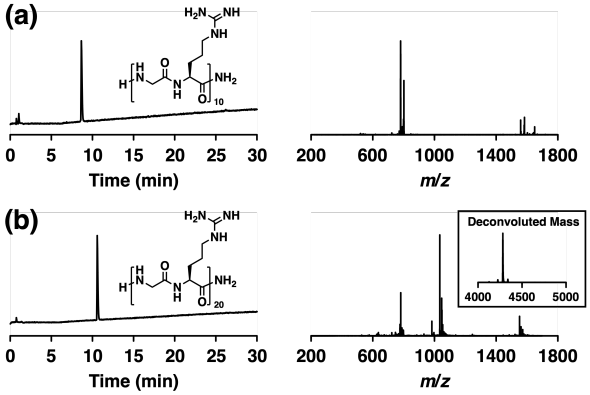
5.4.9 LDH assay using human neuroblastoma BE(2)-C cells

Cytotoxicity assays were performed using the BE(2)-C cell line cultured and treated as above. 2 or 8 hours after treatment, supernatant was transferred to a new 96-well plate for analysis of LDH release with the CytoTox 96 assay kit (Promega). The measurement of vehicle-treated cells was subtracted from each measurement, and % LDH release was calculated as % cytotoxicity = 100 × Experimental LDH Release (OD490) / Maximum LDH Release (OD490).

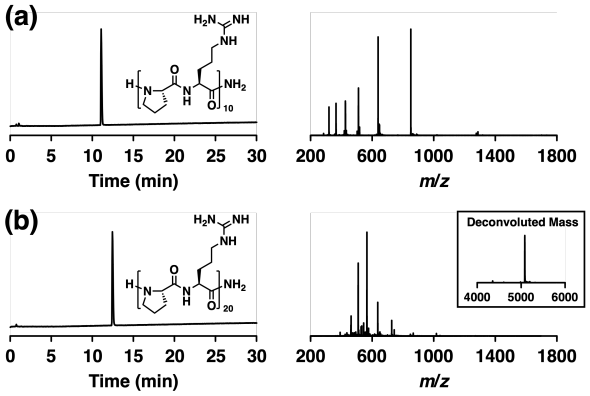
5.5 Acknowledgements

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript. The authors thank Department of Chemistry Instrumentation Facility (DCIF) and Biophysical Instrumentation Facility (BIF), Department of Biology, Massachusetts Institute of Technology for MALDI-TOF mass spectrometry and circular dichroism spectroscopy. This work was supported by Grant-in-Aid for Early-Career Scientists (21K14670 to K.S.) and Grant-in-Aid for Transformative Research Areas "Molecular Cybernetics" (21H05872 to K.S.).

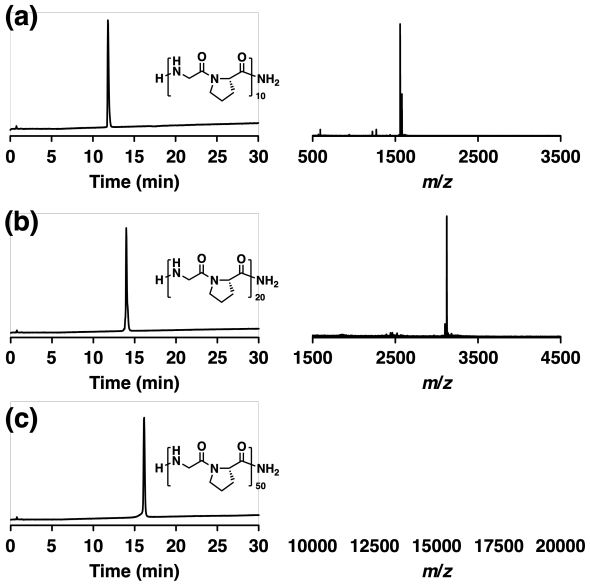
5.6 Appendix I: Characterization of compounds



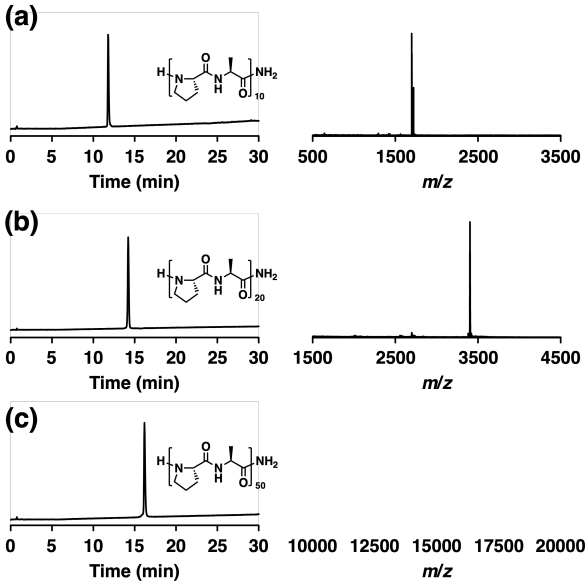
Analytical RP-HPLC traces of and ESI-TOF mass spectra of (a) (GR)10 and (b) (GR)20.



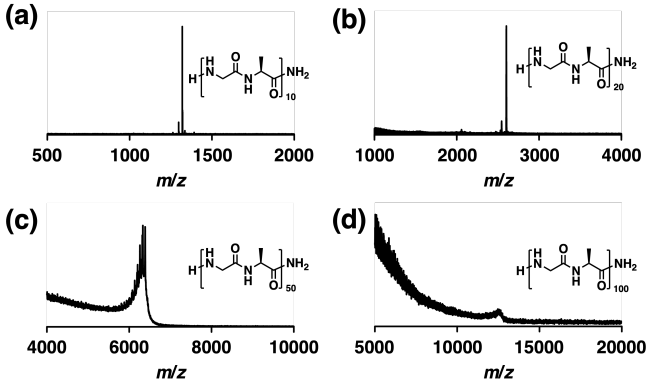
Analytical RP-HPLC traces of and ESI-TOF mass spectra of (a) $(PR)_{10}$ and (b) $(PR)_{20}$.



Analytical RP-HPLC traces of and MALDI-TOF mass spectra of (a) $(GP)_{10}$, (b) $(GP)_{20}$, and (c) $(GP)_{50}$.



Analytical RP-HPLC traces of and MALDI-TOF mass spectra of (a) (PA)10, (b) (PA)20, and (c) (PA)50.



MALDI-TOF mass spectra of (a) purified (GA)10, (b) purified (GA)20, (c) crude (GA)50, and (d) crude (GA)100.

Analytical RP-HPLC traces of poly-GAs could not be obtained due to their poor solubility.

5.7 References

(1) Lattante, S.; Ciura, S.; Rouleau, G. A.; Kabashi, E. Defining the genetic connection linking amyotrophic lateral sclerosis (ALS) with frontotemporal dementia (FTD). *Trends Genet.* **2015**, *31*, 263–273.

(2) Ringholz, G. M.; Appel, S. H.; Bradshaw, M.; Cooke, N. A.; Mosnik, D. M.; Schulz, P. E. Prevalence and patterns of cognitive impairment in sporadic ALS. *Neurology* **2005**, *65*, 586–590.

(3) DeJesus-Hernandez, M.; Mackenzie, I. R.; Boeve, B. F.; Boxer, A. L.; Baker, M.; Rutherford, N. J.; Nicholson, A. M.; Finch, N. A.; Flynn, H.; Adamson, J.; Kouri, N.; Wojtas, A.; Sengdy, P.; Hsiung, G.-Y. R.; Karydas, A.; Seeley, W. W.; Josephs, K. A.; Coppola, G.; Geschwind, D. H.; Wszolek, Z. K.; Feldman, H.; Knopman, D. S.; Petersen, R. C.; Miller, B. L.; Dickson, D. W.; Boylan, K. B.; Graff-Radford, N. R.; Rademakers, R. Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of *C9ORF72* Causes Chromosome 9p-Linked FTD and ALS. *Neuron* **2011**, *72*, 245–256.

(4) Renton, A. E.; Majounie, E.; Waite, A.; Simón-Sánchez, J.; Rollinson, S.; Gibbs, J. R.; Schymick, J. C.; Laaksovirta, H.; van Swieten, J. C.; Myllykangas, L.; Kalimo, H.; Paetau, A.; Abramzon, Y.; Remes, A. M.; Kaganovich, A.; Scholz, S. W.; Duckworth, J.; Ding, J.; Harmer, D. W.; Hernandez, D. G.; Johnson, J. O.; Mok, K.; Ryten, M.; Trabzuni, D.; Guerreiro, R. J.; Orrell, R. W.; Neal, J.; Murray, A.; Pearson, J.; Jansen, I. E.; Sondervan, D.; Seelaar, H.; Blake, D.; Young, K.; Halliwell, N.; Callister, J. B.; Toulson, G.; Richardson, A.; Gerhard, A.; Snowden, J.; Mann, D.; Neary, D.; Nalls, M. A.; Peuralinna, T.; Jansson, L.; Isoviita, V.-M.; Kaivorinne, A.-L.; Hölttä-Vuori, M.; Ikonen, E.; Sulkava, R.; Benatar, M.; Wuu, J.; Chiò, A.; Restagno, G.; Borghero, G.; Sabatelli, M.; ITALSGEN Consortium; Heckerman, D.; Rogaeva, E.; Zinman, L.; Rothstein, J. D.; Sendtner, M.; Drepper, C.; Eichler, E. E.; Alkan, C.; Abdullaev, Z.; Pack, S. D.; Dutra, A.; Pak, E.; Hardy, J.; Singleton, A.; Williams, N. M.; Heutink, P.; Pickering-Brown, S.; Morris, H. R.; Tienari, P. J.; Traynor, B. J. A Hexanucleotide Repeat Expansion in *C9ORF72* Is the Cause of Chromosome 9p21-Linked ALS-FTD. *Neuron* **2011**, *72*, 257–268.

(5) Mori, K.; Arzberger, T.; Grässer, F. A.; Gijselinck, I.; May, S.; Rentzsch, K.; Weng, S.-M.; Schludi, M. H.; van der Zee, J.; Cruts, M.; Van Broeckhoven, C.; Kremmer, E.; Kretzschmar, H. A.; Haass, C.; Edbauer, D. Bidirectional transcripts of the expanded *C9orf72* hexanucleotide repeat are translated into aggregating dipeptide repeat proteins. *Acta Neuropathol*. **2013**, *126*, 881–893.

(6) Zu, T.; Liu, Y.; Bañez-Coronel, M.; Reid, T.; Pletnikova, O.; Lewis, J.; Miller, T. M.; Harms, M. B.; Falchook, A. E.; Subramony, S. H.; Ostrow, L. W.; Rothstein, J. D.; Troncoso, J. C.; Ranum, L. P. W. RAN proteins and RNA foci from antisense transcripts in *C9ORF72* ALS and frontotemporal dementia. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, E4968–E4977.

(7) Wen, X.; Tan, W.; Westergard, T.; Krishnamurthy, K.; Markandaiah, S. S.; Shi, Y.; Lin, S.; Shneider, N. A.; Monaghan, J.; Pandey, U. B.; Pasinelli, P.; Ichida, J. K.; Trotti, D. Antisense Proline-Arginine RAN Dipeptides Linked to *C9ORF72*-ALS/FTD Form Toxic Nuclear Aggregates that Initiate In Vitro and In Vivo Neuronal Death. *Neuron* **2014**, *84*, 1213–1225.

(8) Mizielinska, S.; Grönke, S.; Niccoli, T.; Ridler, C. E.; Clayton, E. L.; Devoy, A.; Moens, T.; Norona, F. E.; Woollacott, I. O. C.; Pietrzyk, J.; Cleverley, K.; Nicoll, A. J.; Pickering-Brown, S.; Dols, J.; Cabecinha, M.; Hendrich, O.; Fratta, P.; Fisher, E. M. C.; Partridge, L.; Isaacs, A. M. *C9orf72* repeat expansions cause neurodegeneration in *Drosophila* through arginine-rich proteins. *Science* **2014**, *345*, 1192–1194.

(9) Lee, K.-H.; Zhang, P.; Kim, H. J.; Mitrea, D. M.; Sarkar, M.; Freibaum, B. D.; Cika, J.; Coughlin, M.; Messing, J.; Molliex, A.; Maxwell, B. A.; Kim, N. C.; Temirov, J.; Moore, J.; Kolaitis, R.-M.; Shaw, T. I.; Bai, B.; Peng, J.; Kriwacki, R. W.; Taylor, J. P. C9orf72 Dipeptide Repeats Impair the Assembly, Dynamics, and Function of Membrane-Less Organelles. *Cell* **2016**, *167*, 774–788.
(10) Freibaum, B. D.; Taylor, J. P. The Role of Dipeptide Repeats in C9ORF72-Related ALS-FTD. *Front. Mol. Neurosci.* **2017**, *10*, 35.
(11) Sharpe, J. L.; Harper, N. S.; Garner, D. R.; West, R. J. H. Modeling *C9orf72*-Related Frontotemporal Dementia and Amyotrophic Lateral Sclerosis in *Drosophila*. *Front. Cell. Neurosci.* **2021**, *15*, 770937.
(12) Chen, Y. Q.; Zhang, S. Q.; Li, B. C.; Qiu, W.; Jiao, B.; Zhang, J.; Diao, Z. Y. Expression of a cytotoxic cationic antibacterial peptide in *Escherichia coli* using two fusion partners. *Protein Expr. Purif.* **2008**, *57*, 303–311.
(13) Rosano, G. L.; Ceccarelli, E. A. Recombinant protein expression in *Escherichia coli*: advances and challenges. *Front. Microbiol.* **2014**, *5*, 172.
(14) Chang, Y.-J.; Jeng, U.-S.; Chiang, Y.-L.; Hwang, I.-S.; Chen, Y.-R. The Glycine-Alanine Dipeptide Repeat from C9orf72 Hexanucleotide Expansions Forms Toxic Amyloids Possessing Cell-to-Cell Transmission Properties. *J. Biol. Chem.* **2016**, *291*, 4903–4911.
(15) Boeynaems, S.; Bogaert, E.; Kovacs, D.; Konijnenberg, A.; Timmerman, E.; Volkov, A.; Guharoy, M.; De Decker, M.; Jaspers, T.; Ryan, V. H.; Janke, A. M.; Baatsen, P.; Vercruysse, T.; Kolaitis, R.-M.; Daelemans, D.; Taylor, J. P.; Kedersha, N.; Anderson, P.; Impens, F.; Sobott, F.; Schymkowitz, J.; Rousseau, F.; Fawzi, N. L.; Robberecht, W.; Van Damme, P.; Tompa, P.; Van Den Bosch, L. Phase Separation of *C9orf72* Dipeptide Repeats Perturbs Stress Granule Dynamics. *Mol. Cell* **2017**, *65*, 1044–1055.
(16) Babu, M.; Favretto, F.; de Opakua, A. I.; Rankovic, M.; Becker, S.; Zweckstetter, M. Proline/arginine dipeptide repeat polymers derail protein folding in amyotrophic lateral sclerosis. *Nat. Commun.* **2021**, *12*, 3396.
(17) Kent, S. B. H. Total chemical synthesis of proteins. *Chem. Soc. Rev.* **2009**, *38*, 338–351.
(18) Agouridas, V.; El Mahdi, O.; Diemer, V.; Cargoët, M.; Monbaliu, J.-C. M.; Melnyk, O. Native Chemical Ligation and Extended Methods: Mechanisms, Catalysis, Scope, and Limitations. *Chem. Rev*. **2019**, *119*, 7328–7443.
(19) Mijalis, A. J.; Thomas III, D. A.; Simon, M. D.; Adamo, A.; Beaumont, R.; Jensen, K. F.; Pentelute, B. L. A fully automated flow-based approach for accelerated peptide synthesis. *Nat. Chem. Biol.* **2017**, *13*, 464–466.
(20) Hartrampf, N.; Saebi, A.; Poskus, M.; Gates, Z. P.; Callahan, A. J.; Cowfer, A. E.; Hanna, S.; Antilla, S.; Schissel, C. K.; Quartarano, A. J.; Ye, X.; Mijalis, A. J.; Simon, M. D.; Loas, A.; Liu, S.; Jessen, C.; Nielsen, T. E.; Pentelute, B. L. Synthesis of proteins by automated flow chemistry. *Science* **2020**, *368*, 980–987.
(21) Pomplun, S.; Jbara, M.; Schissel, C. K.; Hawken, S. W.; Boija, A.; Li, C.; Klein, I.; Pentelute, B. L. Parallel Automated Flow Synthesis of Covalent Protein Complexes That Can Inhibit MYC-Driven Transcription. *ACS Cent. Sci.* **2021**, *7*, 1408–1418.
(22) Jbara, M.; Pomplun, S.; Schissel, C. K.; Hawken, S. W.; Boija, A.; Klein, I.; Rodriguez, J.; Buchwald, S. L.; Pentelute, B. L. Engineering Bioactive Dimeric Transcription Factor Analogs via Palladium Rebound Reagents. *J. Am. Chem. Soc.* **2021**, *143*, 11788–11798.

(23) Mohapatra, S.; Hartrampf, N.; Poskus, M.; Loas, A.; Gómez-Bombarelli, R.; Pentelute, B. L. Deep Learning for Prediction and Optimization of Fast-Flow Peptide Synthesis. *ACS Cent. Sci.* **2020**, *6*, 2277–2286.
(24) Flores, B. N.; Dulchavsky, M. E.; Krans, A.; Sawaya, M. R.; Paulson, H. L.; Todd, P. K.; Barmada, S. J.; Ivanova, M. I. Distinct *C9orf72*-Associated Dipeptide Repeat Structures Correlate with Neuronal Toxicity. *PLoS ONE* **2016**, *11*, e0165084.
(25) Guo, Q.; Lehmer, C.; Martínez-Sánchez, A.; Rudack, T.; Beck, F.; Hartmann, H.; Pérez-Berlanga, M.; Frottin, F.; Hipp, M. S.; Hartl, F. U.; Edbauer, D.; Baumeister, W.; Fernández-Busnadiego, R. *In Situ* Structure of Neuronal *C9orf72* Poly-GA Aggregates Reveals Proteasome Recruitment. *Cell* **2018**, *172*, 696–705.
(26) Brasseur, L.; Coens, A.; Waeytens, J.; Melki, R.; Bousset, L. Dipeptide repeat derived from C9orf72 hexanucleotide expansions forms amyloids or natively unfolded structures *in vitro*. *Biochem. Biophys. Res. Commun.* **2020**, *526*, 410–416.
(27) Shimada, H.; Caughey, W. S. Dynamic Protein Structures. *J. Biol. Chem.* **1982**, *257*, 11893–11900.
(28) Ryle, M. J.; Lanzilotta, W. N.; Seefeldt, L. C.; Scarrow, R. C.; Jensen, G. M. *J. Biol. Chem.* **1996**, *271*, 1551–1557.
(29) Lopes, J. L. S.; Miles, A. J.; Whitmore, L.; Wallace, B. A. Distinct circular dichroism spectroscopic signatures of polyproline II and unordered secondary structures: Applications in secondary structure analyses. *Protein Sci.* **2014**, *23*, 1765–1772.
(30) Adzhubei, A. A.; Sternberg, M. J. E.; Makarov, A. A. Polyproline-II Helix in Proteins: Structure and Function. *J. Mol. Biol.* **2013**, *425*, 2100–2132.
(31) Mammen, M.; Choi, S.-K.; Whitesides, G. M. Polyvalent Interactions in Biological Systems: Implications for Design and Use of Multivalent Ligands and Inhibitors. *Angew. Chem. Int. Ed.* **1998**, *37*, 2754–2794.
(32) Fasting, C.; Schalley, C. A.; Weber, M.; Seitz, O.; Hecht, S.; Koksch, B.; Dernedde, J.; Graf, C.; Knapp, E.-W.; Haag, R. Multivalency as a Chemical Organization and Action Principle. *Angew. Chem. Int. Ed.* **2012**, *51*, 10472–10498
(33) Lobner, D. Comparison of the LDH and MTT assays for quantifying cell death: validity for neuronal apoptosis? *J. Nerurosci. Methods*. **2000**, *96* (2), 147-152.
(34) Marchi, p.; Marrone, L.; Brasseur, L.; Coens, A. et al. C9ORF72 DPRs spread from iAstrocytes to motor neurons. *Life Sci. Alliance.* **2022**, *5* (9) e202101276. DOI: 10.26508/lsa.202101276

Appendix. Cell-Penetrating D-Peptides Retain Antisense Morpholino Oligomer Delivery Activity

The work presented in this chapter has been reproduced from the following manuscript:

Schissel, C. K.; Farquhar, C. E.; Malmberg, A. B.; Loas, A.; Pentelute, B. L. Cell-Penetrating D-Peptides Retain Antisense Morpholino Oligomer Delivery Activity. *ACS Bio Med Chem Au* **2022**, *2* (2), 150–160.

A.1 Introduction

Cell-penetrating peptides (CPPs) can help treat disease by enhancing the delivery of cell-impermeable cargo. CPPs are a class of peptides that are capable of directly entering the cell cytosol.[1–3] These sequences can deliver covalently bound cargo, offering therapeutic potential to macromolecules otherwise restricted to extracellular targets. Although CPPs have been widely studied since their discovery, the field lacks robust methodology to quantify cell entry and penetration efficacy. This dearth of knowledge is due to the complicated mechanisms of CPP cell entry and the many variables that affect CPP efficacy in any given assay—such as peptide concentration, cell type, temperature, treatment time, and cargo.[4] For example, for the well-studied CPP penetratin (RQIKIWFQNRRMKWKK), the reported ratio between intracellular and extracellular concentration ranges from 0.6:1.0 to 95.0:1.0.[5,6] In addition, it is challenging to determine subcellular localization once a peptide is internalized, despite advances in fluorescence, immunoblot, and mass spectrometry detection.[7] The choice of CPP cargo adds an additional confounding factor, as studies in our laboratory have demonstrated that the cell-penetrating ability of more than ten common CPPs differs when bound to a cyanine dye versus a macromolecular drug, with no discernable trend.[8] Therefore, effective development of CPPs requires new methodology for understanding CPP cell entry and subcellular localization that can be carried out on the CPP-cargo conjugate.

A therapeutic macromolecule that would benefit from enhanced delivery is phosphorodiamidate morpholino oligomer (PMO), which has recently reached the market as an antisense "exon skipping" therapy for Duchenne muscular dystrophy (DMD). One of the drugs, eteplirsen, is a 10 kDa synthetic antisense oligomer that must reach the nucleus and bind pre-mRNA for its therapeutic effect. However, studies have shown that two-thirds of eteplirsen is cleared renally within 24 h of administration.[9,10] Several CPPs have been shown to increase PMO uptake, and recent clinical trial results have shown that once-monthly dosing of SRP-5051 resulted in higher muscle concentration, increased exon-skipping and dystrophin production at 12 weeks as compared to once weekly dosing of eteplirsen after 24 weeks in a different study.[11] Although a wide variety of CPPs have been tested for PMO delivery, they have been limited to the native L-

form and studied predominantly with an activity-based assay, forgoing quantitative information on the amount of material inside the cell.[12]

Because PMO is proteolytically stable and invisible to the immune system,[10] attachment to an L-peptide may disrupt the desirable characteristics of PMO; however, conjugates using mirror image peptides may retain these characteristics.[13–15] D-peptides have been explored as CPPs and have been found to display at times greater activities to their L-counterparts, despite suggested chiral binding interactions between CPPs and the cell membrane.[16,17] A study on TAT endosomalytic peptide analogs found that the full D-form decreased uptake, but enhanced endosomal escape and proteolytic stability compared to the native form.[15,18] Another reported that a cyclic D-peptide, when co-administered with insulin, enhanced its oral bioavailability and therapeutic effect in the gut.[19] Some reports are contentious as to whether mutations to D-amino acids are detrimental to CPP activity, and certainly CPPs that depend on a chiral interaction or highly ordered structure to enter the membrane would lose efficacy from D-amino acid substitutions.[20,21] Previous studies have suggested that efficient CPPs for PMO delivery lack secondary structure and can enter the cell through clathrin-mediated endocytosis.[22,23] While it has been found that PMO-D-CPPs have enhanced proteolytic stability over their L-counterparts,[24] their PMO delivery activity has not been investigated. In addition to their underexplored potential, we are interested in studying D-CPPs as potential therapeutic carrier moieties because their fully noncanonical sequence is resistant to proteolysis and may go unrecognized by the host immune system.[15,20] Despite this potential, mirror image peptides have not yet been fully explored for the delivery of PMO.

The proteolytic stability of mirror image peptides would simplify their characterization after uptake into cells, providing orthogonal information to activity-based assays. Currently, the main method used to characterize PMO-CPP internalization is an in vitro assay in which successful delivery of the active oligomer to the nucleus results in green fluorescence.[25] While this is an excellent assay to measure PMO-CPP activity, this assay does not give information on the quantity of material inside the cell. Especially for conjugates with a known endocytic mechanism, understanding endosomal escape is crucial. High concentrations of peptides trapped in the endosome would not be apparent by the activity assay alone, but this loss of active peptide could

be of great therapeutic detriment.[26] Therefore, an additional assay that reveals relative quantities of material in different parts of the cell would provide a valuable metric of CPP delivery, in parallel with our current activity assay. This could be achieved by extracting the cytosolic fraction of treated cells using a mild detergent, such as digitonin, and comparison to the whole cell fraction.[27] Comparing activity and relative quantity would be a valuable metric for efficiency, where a high efficiency peptide is one with a high ratio of antisense activity to internal concentration.

Existing methods to quantify uptake into cells include fluorescence, immunoblot, and mass spectrometry, but it is still challenging to distinguish between endosomal and cytosolic localization. Several assays can differentiate between endosomal and cytosolic localization using indirect quantification via a readout generated by a delivered cargo, including the chloroalkane penetration assay (CAPA),[28] GFP complementation assays,[29] and more recently the NanoClick[30] assay and SLEEQ[31] assay. Direct quantification of the cytosolic concentration of a fluorescently-labeled protein construct is possible using fluorescence correlation spectroscopy.[32]

Mass spectrometry is a direct quantification tool that would give information about the concentration of peptides recovered from biological mixtures with limited labeling required. Past studies have illustrated how MALDI-TOF mass spectrometry is a practical tool for absolute and relative quantification of peptides and proteins. For example, using an internal standard of a similar molecular weight is sufficient for generation of a calibration curve.[33] Quantitation of total uptake of L-CPPs was achieved using heavy atom-labeled internal standards.[34,35] While this assay provided information regarding whole cell uptake of CPPs and CPP-peptide conjugates, it is limited by the need for heavy-atom labeling and the rapid degradation of L-peptides.[36] A method for circumnavigating the need for spike-in of heavy atom-labeled standards was developed for the relative quantification of phosphopeptides.[37] However, the proteolytic stability of D-peptides would facilitate their recovery and analysis as a mixture from inside cells and animals, allowing for the use of a new metric of antisense delivery efficiency.

Here we report that compared to the native L-forms, the mirror image forms of several sequences were equally able to deliver antisense molecules to the nucleus, but their increased proteolytic stability simplified mass spectrometry-based characterization following cytosolic

delivery. Cytosolic delivery can be quantified based on the recovery of intact constructs from inside the cell. We profiled the uptake of biotinylated CPPs and PMO-CPPs to determine their relative concentrations in the whole cell and cytosol using careful extraction with digitonin and direct detection via MALDI-TOF. By comparing PMO delivery activity to relative internal concentration, we can derive a new metric for cargo delivery efficiency that may be useful for future development of CPPs for PMO delivery.

A.2 Results and discussion

A.2.1 Mirror image peptides have same PMO delivery activity as native forms

We first established that several mirror image peptides could deliver a model PMO molecule to the nucleus of cells with similar efficacy to their native L-forms. We selected commonly studied, cationic peptides with precedented PMO delivery activity, but not known for their dependence on secondary structure or receptor interaction, and synthesized them in their L and D-form. The first iteration of these constructs contained a biotin linked through a 6-aminohexanoic acid residue and a trypsin-cleavable motif in between the peptide and the cargo. The cargo portion contained an azide for conjugation to PMO, and a biotin for use in affinity capture (**Fig. A.1a**). The second generation of these constructs did not include an aminohexanoic acid linker, and did include a single tryptophan residue for quantitation by UV-Vis (**Fig. A.1c**). The final constructs included cationic (all peptides), oligoarginine (R8, BPEP, Bac7, TAT, TATp), and amphipathic (Penetratin, Hell11-7, DPV6, DPV7, MPG) sequences.

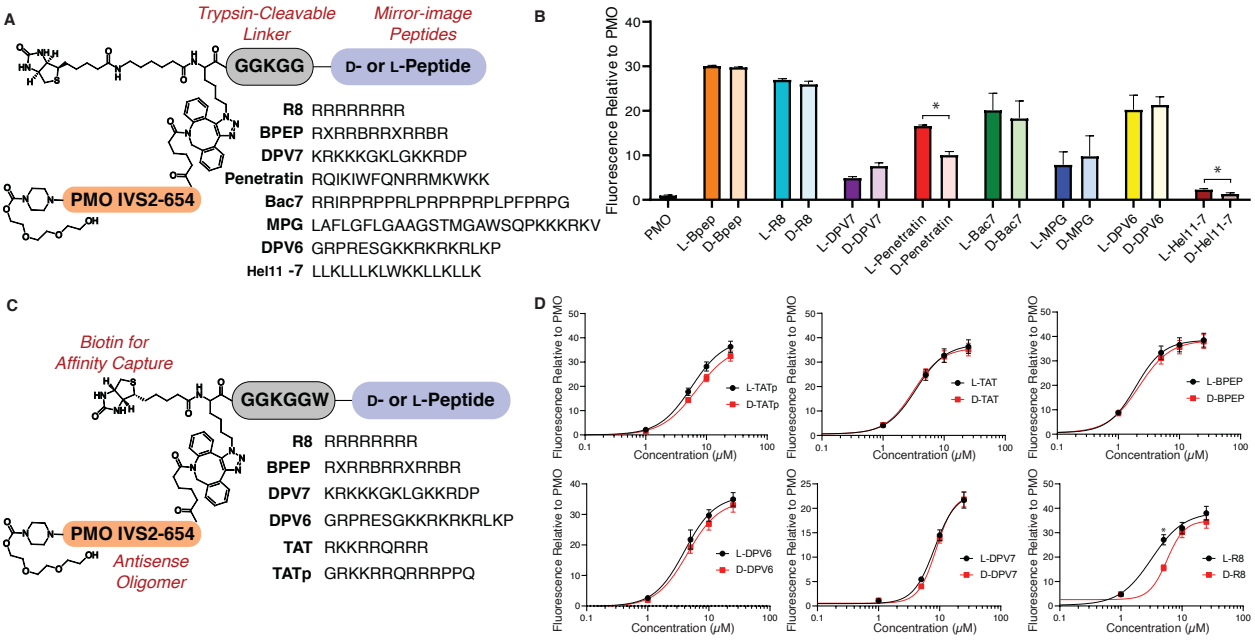


Figure A.1 Mirror image cell-penetrating peptides have similar PMO delivery activity as their native counterparts.
(a) Construction of the first generation conjugates studied, including the eight studied CPP sequences. Macromolecular cargo PMO IVS2-654 is attached to the N-terminus of the peptides, along with a biotin handle for subsequent affinity capture. A trypsin-cleavable linker connects the cargo to the peptides, and a 6-aminohexanoic acid moiety is between the biotin and the peptide. (b) Shown is activity data from the

EGFP 654 assay conducted with D- and L-forms of the first generation sequences at 5 μM. (c) Construction of the second generation conjugates used in subsequent experiments. The aminohexanoic acid linker was removed, and each peptide contains a single tryptophan residue for quantitation by UV-Vis. (b) Dose-response curves of D- and L-forms of several sequences at varying concentrations. Bars and points represent mean ± SD, N = 3 distinct samples from a representative biological replicate. Replicate experiments showed similar results and are shown in the SI. Statistical significance between the D- and L-samples at each concentration was determined using a two-tailed unpaired t-test and indicated by *p<0.01, with blank indicating not statistically significant.

In total, fourteen PMO-peptide sequences were synthesized in their D- and L-form and tested in an activity-based in vitro assay.[22,23] To synthesize the constructs, L-peptides were synthesized via automated fast-flow peptide synthesis, and D-peptides were synthesized using semi-automated fast flow peptide synthesis. Azido-lysine and biotin moieties were added to the N-terminus of the peptides manually, and the peptides were simultaneously cleaved and deprotected before purification via RP-HPLC. PMO was modified with a dibenzocyclooctyne (DBCO) moiety and purified before attachment to the azido-peptides via strain-promoted azide-alkyne cycloaddition in water. Purified constructs were then tested using an activity-based readout in which nuclear delivery results in fluorescence. Briefly, HeLa cells stably transfected with an EGFP gene interrupted by a mutated intron of ß-globin (IVS2-654) produce a non-fluorescent EGFP protein. Successful delivery of PMO IVS2-654 to the nucleus results in corrective splicing and EGFP synthesis. The amount of PMO delivered to the nucleus is therefore correlated with EGFP fluorescence, quantified by flow cytometry. Activity is reported as mean fluorescence intensity (MFI) relative to PMO alone. This activity assay provides indirect information on how much active PMO is delivered to the nucleus.

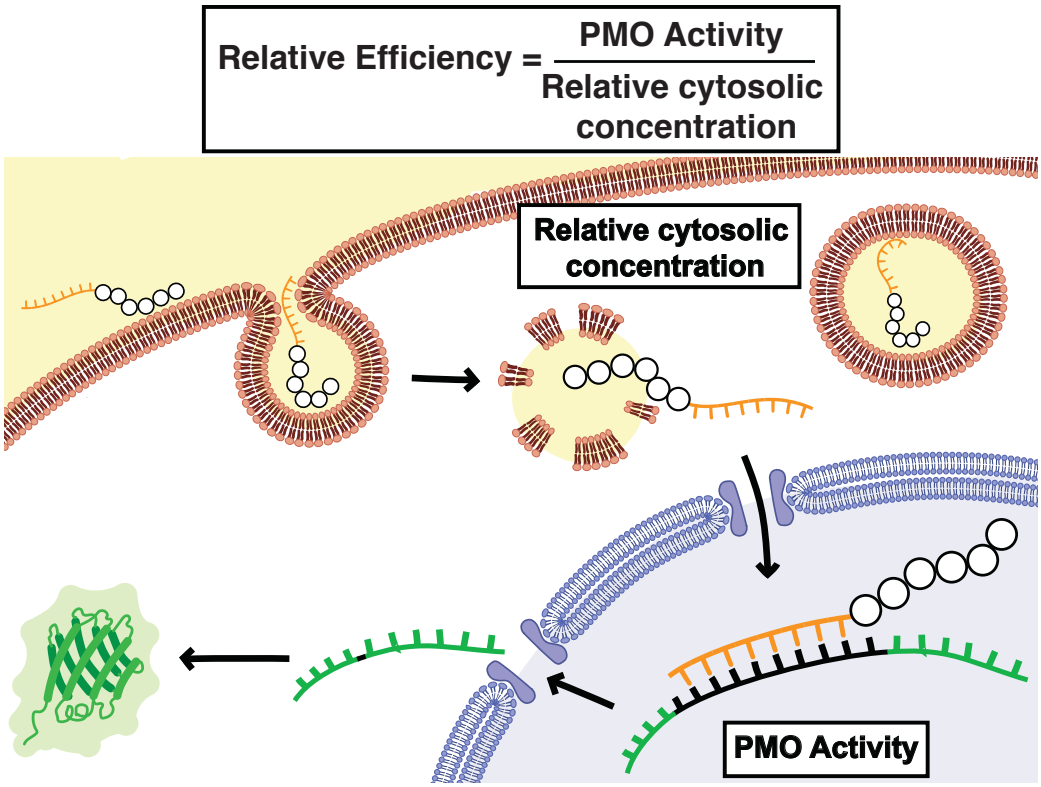


Figure A.2 Schematic of how PMO-CPPs may enter the cell to perform exon-skipping activity.
While activity assays give information of how much active PMO reached its target, a mass spectrometry-based assay may give information on the amount of PMO-CPP located in the cell and cytosol. Comparing these metrics provides a new estimate of CPP efficiency.

Dose-response studies with sequences in the d- and l-form confirmed similar activities between mirror image CPPs. From our initial proof-of-concept experiment, involving eight sequences in l- and d-form tested at a single concentration (5 μM), there was not a significant difference between the activities of the mirror image peptides with the exception of Penetratin and Hel11-7 (**Fig A.1b**). The second generation constructs (**Fig. A.1c**) were tested at varying concentrations in the EGFP 654 assay, and the results further suggested that mirror image peptides shared nearly identical PMO delivery activities (**Fig. A.1d**). Interestingly, while PMO-d-R8 showed similar activity to PMO-l-R8, it did not demonstrate the closeness of activity observed with the other sequences. The trend has been observed before, albeit without attached cargo, that l-R8 entered cells more efficiently than d.[20] Though striking similarities were observed for the selected peptides, this similarity is certainly not expected for all CPPs and cargoes, including those that rely on secondary structure or receptor-mediated uptake. These activities correspond with how much PMO is delivered to the nucleus, but do not provide information regarding the total amount

of material inside the cell. Relative efficiency of a PMO-CPP could be characterized by comparing activity to internal concentration, as discussed later (**Fig. A.2**).

A.2.2 Mirror image peptides are proteolytically stable

Our primary motivation for investigating mirror image peptides for transporting PMO was that the D-form would be stable against proteolysis and thus would match this property of the PMO cargo. D-peptides are indeed stable against degradation, illustrated by a time-course study in which both forms of PMO-CPPs were incubated in 25% human serum. While the studied PMO-D-CPPs remained intact 24 h later, the L-forms rapidly degraded into multiple fragments, leaving the parent construct as a minor product after only one hour of incubation (**Fig. A.3**). The major degradation products correspond to loss of C-terminal RRRPPQ and KKRRQRRRPPQ motifs of TATp (sequence: GGKGGWGRKKRRQRRRPPQ). This observation furthers the notion that L-peptides are not suitable for investigation using mass spectrometry after recovery from a biological setting. However, D-peptide conjugates can be recovered from a biological environment such as serum without suffering degradation, simplifying their characterization via mass-spectrometry.

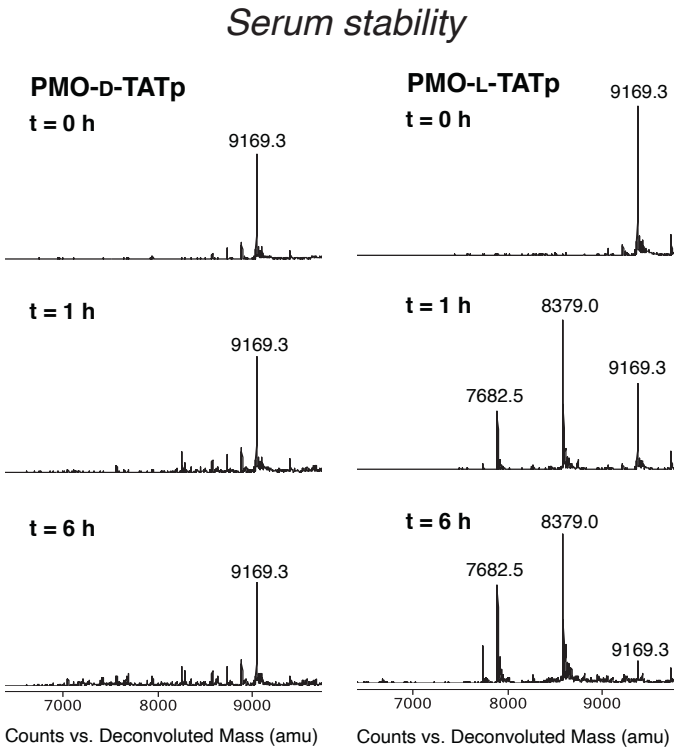


Figure A.3 Mirror image cell-penetrating peptides remain proteolytically stable.

Mass spectra of PMO-D- and L- conjugates following incubation with human serum. The L-variant is degraded within 1 h, while the D-variant is stable after 6 h.

A.2.3 Mirror image PMO-peptides can be recovered from inside cells

The proteolytic stability of d-peptides simplifies the recovery of mixtures of intact constructs after being internalized into cells. While MALDI-ToF has been used previously to analyze the quantity of l-peptides and protein-peptide conjugates recovered from inside cells, it has not yet been used to profile PMO-d-peptides, or mixtures of more than three conjugates at a time.[34,35] MALDI-ToF has also been used to quantify many different delivered cargos, rather than the CPP.[38–40] The use of d-peptides would facilitate the analysis of a mixture of conjugates, because without degradation, only the parent peak would be observed. Moreover, this platform has not yet been used to study peptides recovered from sub-cellular fractions. Through a series of experiments, we show that mixtures of intact PMO-d-CPPs can be recovered from the cytosol of cells and analyzed by MALDI-ToF to estimate their relative abundances without the need for isotope labeling or standard curves (**Fig. A.4a**).

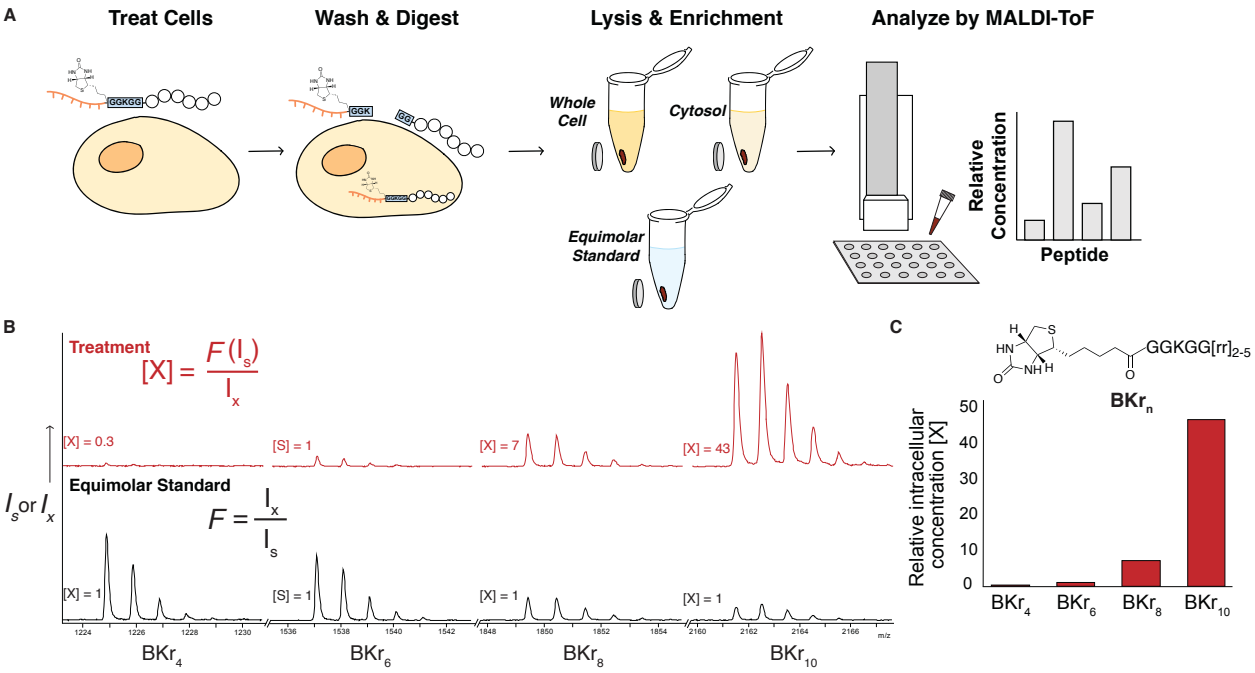


Figure A.4 Uptake assay reveals relative concentrations of intact construct inside the cell.
(a) Workflow of the uptake assay; cells are treated with PMO-D-CPPs, washed, and lysed to extract the whole cell lysate or the cytosol. A trypsin-cleavable L-linker ensures extracellular constructs are not recovered. Constructs are immobilized on magnetic streptavidin beads, washed, and plated directly for MALDI analysis. Intensities of analytes compared to an equimolar standard provide relative concentrations.

(b) MALDI-TOF mass spectra displaying ions of intact biotinylated D-polyarginine peptides, isolated after internalization into HeLa cells. Spectra show ions corresponding to intact conjugates in the equimolar spike-in (black) and the whole cell lysate of cells treated with equimolar mixture (red). (c) By comparing relative intensities in the equimolar standard, the response factor (*F*) was determined and used to calculate the fold change in concentration in the experimental samples, shown as bar graph, normalized to BKr_4. Also shown is the equation used to determine relative concentration: *I* (intensity), *[X]* (sample concentration), *F* (response factor), *X* (sample), *S* (standard).

Our first step was to recapitulate a known empirical trend that more Arg residues leads to greater uptake. We began our assay with a simple model system of four polyarginine peptides with a trypsin cleavable linker and a single biotin label. HeLa cells were incubated with biotin-K-D-Arg_4, D-Arg_6, D-Arg_8, and D-Arg_{10} for 1 h. The cells were then washed extensively with PBS and heparin and trypsinized to lift the cells as well as to cleave the L-linker on the extracellular constructs, preventing their recovery. The treatment with heparin is known to release peptides bound to the membrane but that are not internalized[20], although it is possible that membrane-bound peptides are recovered in the whole cell lysate. The whole cell lysate was then prepared using RIPA buffer. Fully intact biotinylated peptides were captured with magnetic streptavidin Dynabeads, washed, and plated directly for MALDI analysis. Also plated were Dynabeads incubated in an equimolar mixture of the same constructs as determined by UV-Vis (Fig. A.4b).

The relative concentration of peptides on the beads can be estimated by determining the analyte's response factor (F) from the equimolar standard (**Fig. A.4b**). In the standard, each analyte's concentration is 1 mM, and each analyte's response factor (F) is determined by normalizing their intensities to an internal standard (S). Here, BKr6 was selected as the arbitrary standard, where F = 1. The response factor of each analyte should remain consistent across samples that contain the same analytes,[33,37] and was used to calculate the fold change in concentration in the experimental samples. The relative concentrations [X] of the analytes, normalized to the 'internal standard' BKr6 are shown as a bar graph (**Fig. A.4c**). There is a clear increase in concentration of the constructs with more Arg residues, with BKr10 having 40-fold greater concentration than BKr6. This trend of greater number of Arg residues leading to greater uptake is already well documented in the literature.[41,42]

A.2.4 PMO-d-CPPs can be extracted from cytosol and analyzed by MADI

Next, we found that these PMO-CPP constructs likely enter via energy-dependent endocytosis. When evaluating CPP delivery efficiency, considering the mechanism of uptake and potential for endosomal entrapment is necessary. In contrast with CPPs alone which can enter via passive diffusion,[41,43] we hypothesized that the uptake mechanism of these constructs was likely endocytosis, considering the size of the PMO cargo as well as previous studies that we have conducted on similar conjugates.[22,23] Using a panel of chemical endocytosis inhibitors, we performed a pulse-chase EGFP 654 assay format in which cells were pre-incubated with inhibitors before treatment with the L- and D-forms of PMO-DPV7 and PMO-Bac7. Analysis by flow cytometry revealed that chlorpromazine reduced activity in a dose-dependent manner (**Fig. A.5a**). Chlorpromazine is an inhibitor of clathrin-mediated endocytosis and has previously been observed to inhibit activity of similar PMO constructs. While it is possible that multiple uptake mechanisms are occurring, these PMO-CPP conjugates are likely taken up by active transport.

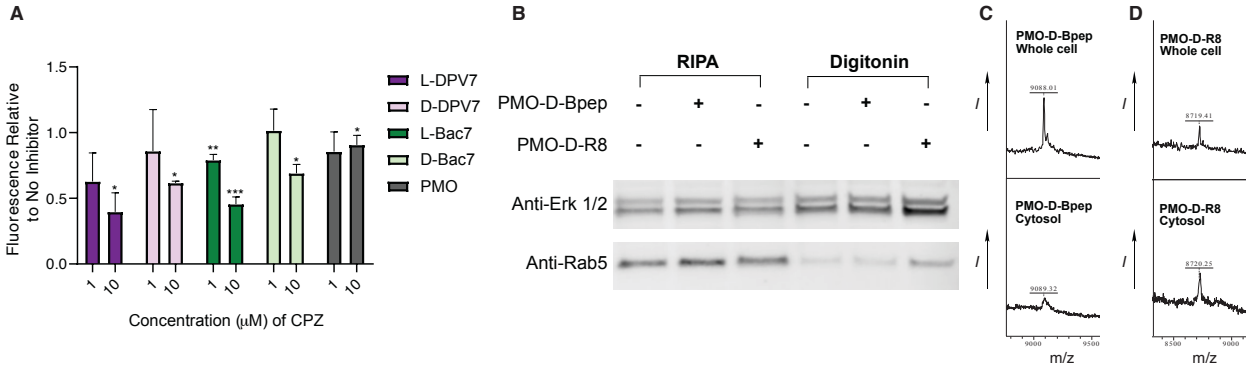


Figure A.5 PMO-D-CPPs enter via endocytosis, and can be detected in whole cell and cytosolic lysate by MALDI-TOF.
(a) Shown are PMO activities from an EGFP 654 assay in which cells are treated in a pulse-chase format with chemical endocytosis inhibitors followed by PMO-conjugates at 5 μM. Chlorpromazine (CPZ) produces a dose-dependent inhibition of PMO activity. Bars represent group mean ± SD, N = 3 distinct samples from a single biological replicate. *p<0.05, **p<0.005 compared to cells treated without inhibitor. (b) Western blot demonstrating extraction of whole cell lysate and cytosolic fraction with RIPA and digitonin buffer, respectively. Erk 1/2 is a cytosolic marker, whereas Rab5 is a late endosomal marker. Lysates shown are from untreated cells, as well as cells treated with PMO-D-Bpep and PMO-D-R8. (c) and (d) show example MALDI spectra following uptake analysis of lysates from (B) containing PMO-d-Bpep and PMO-d-R8, respectively. Intact construct was detected in the whole cell (top) as well as cytosolic (bottom) fractions.

Knowing that PMO-CPPs enter via endocytosis, we assert that extracting the cytosol and comparing to whole cell lysate is critical when evaluating relative concentrations of constructs

internalized into cells. Not all endocytosed compounds are able to escape the endosome, and endosomal entrapment would lead to less active PMO delivered into the cytosol, measured by a lower cytosolic concentration relative to whole-cell lysate. Therefore, we then extracted biotinylated PMO-CPPs from the cytosol as well as the whole cell lysate following internalization and detected them by Western blot and MALDI. Individually, we incubated PMO-D-R8 or PMO-D-Bpep at 5 μM with HeLa cells in a 12-well plate for 1 h before washing with heparin and digesting with trypsin. The cytosol was extracted using Digitonin buffer, which selectively permeabilizes the outer membrane. RIPA buffer was used to prepare whole cell lysates. To confirm cytosolic extraction, a portion of each sample was analyzed via Western blot using a cytosolic marker (Erk 1/2) and a late-endosomal marker (Rab5). Samples of cytosolic extract from both treated and untreated cells have markedly reduced Rab5 while all samples contain Erk 1/2 (**Fig. A.5a**). Trace contamination of Rab5 is observed in the cytosolic extract although the amount is significantly reduced compared to the whole cell lysate. Finally, as with the biotinylated peptides, the PMO-CPPs were then extracted from the samples with Streptavidin coated magnetic Dynabeads, washed extensively, and analyzed via MALDI-TOF. We indeed detected PMO-D-R8 and PMO-D-Bpep in their respective samples, presenting the first instance of an intact peptide-oligonucleotide conjugate being extracted from cells and analyzed by mass spectrometry (**Fig. A.5c-d).** Moreover, we found that incubation at lower temperature appeared to inhibit cytosolic localization of PMO-D-CPPs, but resulted in equal relative concentrations between whole cell and cytosolic fractions for biotin-D-CPPs, which may be able to passively diffuse through the membrane.

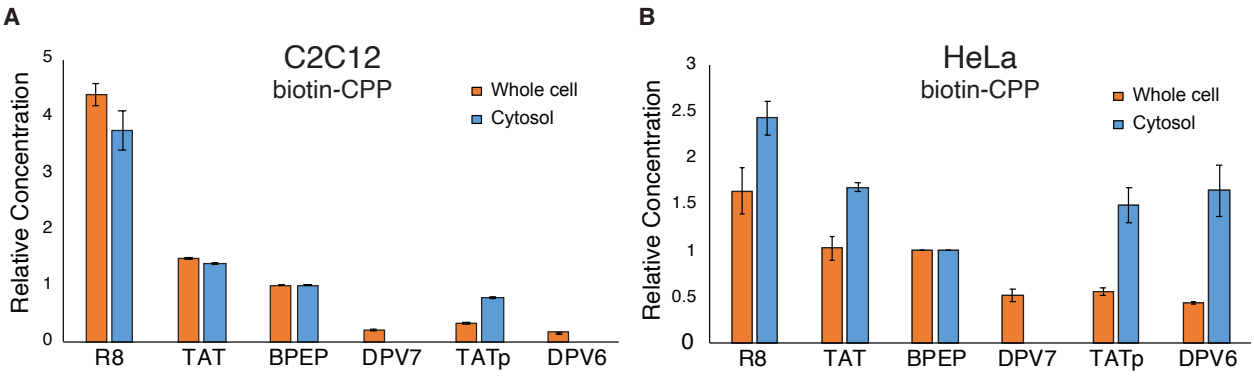


Figure A.6 Uptake of biotin-CPPs can be profiled in different cell lines.

(a) Bar graph showing concentrations of biotin-CPPs relative to BPEP in the whole cell and cytosolic extracts of C1C12 mouse myoblast cells following 1 h treatment at 37 °C. (b) Bar graph showing concentrations of biotin-CPPs in the whole cell and cytosolic extracts of HeLa cells following 1 h treatment at 37 °C. Bars represent group mean ± SD N = 2 replicate samples. Relative concentration is normalized to Bpep.

Next, we expanded the analytes tested and profiled the uptake of six biotin-D-CPPs in the cytosol of distinct cell lines. Biotin-D-R8, TAT, Bpep, DPV7, TATp, and DPV6 were profiled by MALDI in both HeLa (**Fig. A.6a**) and C2C12 mouse myoblast (**Fig. A.6b**) cell lines. We observed different uptake patterns between the two cell lines; polyarginine was significantly more abundant in the C2C12 cells compared to the other peptides, and DPV6 and DPV7 were not detected in the cytosol. In HeLa, polyarginine had the highest relative concentration in both cytosol and whole cell, and DPV7 was again not detected in the cytosol. However, TAT, TATp, and DPV6 all had similar relative cytosolic concentrations. With these experiments we demonstrated that this profiling platform could determine the relative concentration of six intact peptides extracted from whole cell and cytosol of two distinct cell lines.

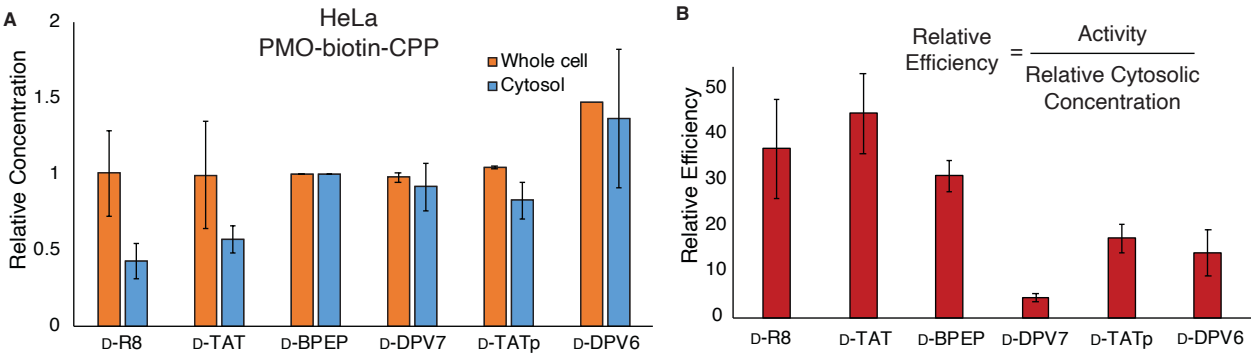


Figure A.7 Mass spectrometry-based profiling combined with activity gives new efficiency metric for PMO-CPPs.

(a) Bar graph showing concentrations of PMO-biotin-CPPs relative to BPEP in the whole cell and cytosolic extracts of Hela cells. Relative concentration is normalized to BPEP. (b) Bar graph showing relative efficiency (PMO activity / relative cytosolic concentration) of PMO-CPPs. Bars show group mean ± SD, N = 2 distinct samples from a single biological replicate, except for the whole cell condition of (A) in which N = 2 distinct samples from two independent biological replicates.

Finally, we profiled uptake of six PMO-D-CPPs. PMO-D-R8, TAT, BPEP, DPV7, TATp, and DPV6 were profiled in HeLa cells as usual. Extraction of cytosol was confirmed by Western blot (SI) and the samples were analyzed by MALDI. Relative concentrations were normalized to BPEP. In the whole cell extract, the relative concentrations were generally consistent across

peptides, with DPV6 having 1.5-fold higher relative concentration. On the other hand, relative concentrations in the cytosol were lower for the polyarginine peptides R8 and TAT, and highest for BPEP and DPV6 (**Fig. A.7a**). By comparing the relative cytosolic concentration to overall peptide activity, the relative efficiencies demonstrate an effective metric for determining which of the six CPPs can effectively deliver PMO for nuclear splice-correction activity (**Fig. A.7b**).

A.3 Conclusions

Here we show that the all-D version of several known cell-penetrating peptides can deliver an antisense oligonucleotide to the nucleus of cells at the same efficiency as their native L-form. While the respective activities were nearly identical in cells, the D-forms were resistant to serum proteolysis. This stability enabled recovery and mass spectrometric analysis of the D-peptides from cytosolic and whole cell lysates. By comparing PMO activity to relative internalized concentrations, we obtained a metric for PMO delivery efficiency.

While it would not be expected that the mirror image of all CPPs would retain delivery activity, the striking similarities observed for the mirror image peptides selected in this study suggests that D-peptides should be studied further. Despite similar activity in cells, the difference in proteolytic stability indicates the potential for differential activities in animals. Proteolytic degradation of peptide therapeutics has been considered a major weakness limiting therapeutic investigation.[44] However, mutations to D-amino acids, or the study of entirely D-peptides, has the potential to improve pharmacokinetic properties of proteinogenic therapies. Avoidance of cleavage sites can enhance proteolytic stability and half-life in vivo, and can be achieved by integration of unnatural amino acids, including D-amino acids.[45] An all-D polyproline CPP was shown to be efficacious in mice[18], and PMO-D-CPPs were previously found to be the most stable among tested B-peptide analogs.[15] Moreover, it is likely that intact PMO-CPP conjugates are able to enter the nucleus as opposed to a cleaved PMO only.[22] Despite these promising findings, the activities of PMO-D-CPPs had not previously been explored. We found that in HeLa cells, several CPP sequences had nearly identical delivery activities in their D- and L-forms, while the D-form remained completely proteolytically stable.

Besides potential in vivo applications, the proteostability of D-peptides also simplified their direct analysis by mass spectrometry following recovery from biological milieu. Previous reports demonstrated that biotinylated L-peptides could be recovered from inside whole cell lysate and quantified by MALDI-ToF, and cleavage products were identified and accounted for.[7,34] Here, by focusing only on fully intact constructs, we analyzed a mixture of six individual biotinylated D-peptides from inside cells by observing only the intact parent ions. Moreover, we demonstrated relative quantification of a mixture of biotinylated peptides recovered from cytosolic extract.

Analysis of the cytosolic portion is critical when studying and developing cell-penetrating peptides, considering that whole cell uptake does not correlate with cytosolic delivery due to the possibility of endosomal entrapment.

Finally, we showed relative quantification of intact biotinylated PMO-D-CPPs in whole cell lysate. By combining relative internal concentrations with PMO delivery activity, we obtained a metric for relative delivery efficiency. Delivery efficiency would be a useful metric for comparing CPPs delivering active therapeutic cargo as it takes into account both the activity of the cargo as well as the internal concentration. A highly efficient peptide would have high activity with low internal concentration, and thus a high relative efficiency. This metric decouples activity measurements from endosomal escape, allowing the direct comparison of cytosolic concentration and activity. Using these metrics when investigating CPPs would be useful in narrowing the scope of sequences early in development to exclude sequences which accumulate in the endosomes or otherwise do not efficiently deliver active cargo.

Exploration of mirror image CPPs would enter into a largely untapped chemical space. We recently described a machine learning-based platform for the discovery of nuclear-targeting peptides containing unnatural amino acids.[23] While the unnatural residues included in this work were non-alpha-amino acids, future applications of this method could include D-amino acid substitutions, or fully mirror image CPPs in order to discover highly active, novel sequences that are completely stable.

A.4 Experimental Section

A.4.1 Reagents and solvents

H-Rink Amide-ChemMatrix resin was obtained from PCAS BioMatrix Inc. (St-Jean-sur-Richelieu, Quebec, Canada). 1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium-3- oxid-hexafluorophosphate (HATU), 4-pentynoic acid, 5-azidopentanoic acid, Fmoc-β-Ala-OH, Fmoc-6-aminohexanoic acid, and Fmoc-L-Lys(N3) were purchased from Chem-Impex International (Wood Dale, IL). PyAOP was purchased from P3 BioSystems (Louisville, KY). Fmoc-protected L-amino acids (Fmoc-Ala-OHxH2O, Fmoc-Arg(Pbf)-OH; Fmoc-Asn(Trt)-OH; Fmoc-Asp-(Ot-Bu)-OH; Fmoc-Gln(Trt)-OH; Fmoc-Glu(Ot-Bu)-OH; Fmoc-Gly-OH; Fmoc-His(Trt)-OH; Fmoc-Ile-OH; Fmoc-Leu-OH; Fmoc-Lys(Boc)-OH; Fmoc-Met-OH; Fmoc-Phe-OH; Fmoc-Pro-OH; Fmoc-Ser(But)-OH; Fmoc-Thr(t-Bu)-OH; Fmoc-Trp(Boc)-OH; Fmoc-Tyr(t-Bu)-OH; Fmoc-Val-OH), were purchased from the Novabiochem-line from Sigma Millipore. Fmoc-protected D-amino acids were purchased from Chem Impex (Wood Dale, IL). Peptide synthesis-grade N,N-dimethylformamide (DMF), CH2Cl2, diethyl ether, and HPLC-grade acetonitrile were obtained from VWR International (Radnor, PA). All other reagents were purchased from Sigma-Aldrich (St. Louis, MO). Milli-Q water was used exclusively.

A.4.2 General method for peptide preparation

Fast-flow Peptide Synthesis: Peptides were synthesized on a 0.1 mmol scale using an automated fast-flow peptide synthesizer for L-peptides and a semi-automated fast-flow peptide synthesizer for D-peptides. Automated synthesis conditions were used as previously reported.[1] Briefly, a 100 mg portion of ChemMatrix Rink Amide HYR resin was loaded into a reactor maintained at 90 ºC. All reagents were flowed at 40 mL/min with HPLC pumps through a stainless-steel loop maintained at 90 ºC before introduction into the reactor. For each coupling, 10 mL of a solution containing 0.4 M amino acid and 0.38 M HATU in DMF were mixed with 600 µL diisopropylethylamine and delivered to the reactor. Fmoc removal was accomplished using 10.4 mL of 20% (v/v) piperidine. Between each step, DMF (15 mL) was used to wash out the reactor. To couple unnatural amino acids or to cap the peptide (e.g. with 4-pentynoic acid), the resin was incubated for 30 min at room temperature with amino acid (1 mmol) dissolved in 2.5 mL 0.4 M HATU in DMF with 500 µL diisopropylethylamine. After completion of the synthesis, the resin was washed 3 times with dichloromethane and dried under vacuum.

Semi-automated synthesis was carried out as previously described.[2] 1 mmol of amino acid was combined with 2.5 mL 0.4 M HATU and 500 µL DIEA and mixed before being delivered to the reactor containing resin via syringe pump at 6 mL/min. The reactor was submerged in a water bath heated to 70 ºC. An HPLC pump delivered either DMF (20 mL) for washing or 20 % piperidine/DMF (6.7 mL) for Fmoc deprotection, at 20 mL/min.

Peptide Cleavage and Deprotection: Each peptide was subjected to simultaneous global side-chain deprotection and cleavage from resin by treatment with 5 mL of 94% trifluoroacetic acid (TFA), 2.5% thioanisole, 2.5% water, and 1% triisopropylsilane (TIPS) (v/v) at room temperature for 2 to 4 hours. The cleavage cocktail was first concentrated by bubbling N_2 through the mixture, and cleaved peptide was precipitated and triturated with 40 mL of cold ether (chilled in dry ice). The crude product was pelleted by centrifugation for three minutes at 4,000 rpm and the ether was decanted. This wash step was repeated two more times. After the third wash, the pellet was dissolved in 50% water and 50% acetonitrile containing 0.1% TFA, filtered through a fritted syringe to remove the resin and lyophilized.

Peptide Purification: The peptides were dissolved in water and acetonitrile containing 0.1% TFA, filtered through a 0.22 µm nylon filter and purified by mass-directed semi-preparative reversed-phase HPLC. Solvent A was water with 0.1% TFA additive and Solvent B was acetonitrile with 0.1% TFA additive. A linear gradient that changed at a rate of 0.5% B/min was used. Most of the peptides were purified on an Agilent Zorbax SB C18 column: 9.4 x 250 mm, 5 µm. Using mass data about each fraction from the instrument, only pure fractions were pooled and lyophilized. The purity of the fraction pool was confirmed by LC-MS.

A.4.3 PMO-DBCO synthesis

PMO IVS-654 (50 mg, 8 µmol) was dissolved in 150 µL DMSO. To the solution was added a solution containing 2 equivalents of dibenzocyclooctyne acid (5.3 mg, 16 µmol) activated with HBTU (37.5µL of 0.4 M HBTU in DMF, 15 µmol) and DIEA (2.8 µL, 16 µmol) in 40 µL DMF (Final reaction volume = 0.23 mL). The reaction proceeded for 25 min before being quenched with 1 mL of water and 2 mL of ammonium hydroxide. The ammonium hydroxide hydrolyzed any ester formed during the course of the reaction. After 1 hour, the solution was diluted to 40 mL in water/acetonitrile and purified using reverse-phase HPLC (Agilent Zorbax SB C3 column: 21.2 x

100 mm, 5 μm) and a linear gradient from 2 to 60% B (solvent A: water; solvent B: acetonitrile) over 58 min (1% B / min). Using mass data about each fraction from the instrument, only pure fractions were pooled and lyophilized. The purity of the fraction pool was confirmed by LC-MS.

Conjugation to peptides

PMO-DBCO (1 eq, 5 mM, water) was conjugated to azido-peptides (1.5 eq, 5 mM, water) at room temperature for 2 h. Reaction progress was monitored by LCMS and purified when PMO-DBCO was consumed. Purification was conducted using mass-directed HPLC (Solvent A: 100 mM ammonium acetate in water, Solvent B: acetonitrile) with a linear gradient that changed at a rate of 0.5% B/min, on an Agilent Zorbax SB C13 column: 9.4 x 250 mm, 5 μm. Using mass data about each fraction from the instrument, only pure fractions were pooled and lyophilized. The purity of the fraction pool was confirmed by LC-MS.

A.4.4 EGFP assay

HeLa 654 cells were maintained in MEM supplemented with 10% (v/v) fetal bovine serum (FBS) and 1% (v/v) penicillin-streptomycin at 37 °C and 5% CO2. 18 h prior to treatment, the cells were plated at a density of 5,000 cells per well in a 96-well plate in MEM supplemented with 10% FBS and 1% penicillin-streptomycin.

PMO-peptides were dissolved in cation-free PBS at a concentration of 1 mM (determined by UV) before being diluted in MEM. Cells were incubated at the designated concentrations for 22 h at 37 °C and 5% CO2. The treatment media was removed, and the cells were washed once before being incubated with 0.25 % Trypsin-EDTA for 15 min at 37 °C and 5% CO2. Lifted cells were transferred to a V-bottom 96-well plate and washed once with PBS, before being resuspended in PBS containing 2% FBS and 2 μg/mL propidium iodide (PI). Flow cytometry analysis was carried out on a BD LSRII flow cytometer at the Koch Institute. Gates were applied to the data to ensure that cells that were positive for propidium iodide or had forward/side scatter readings that were sufficiently different from the main cell population were excluded. Each sample was capped at 5,000 gated events.

Analysis was conducted using Graphpad Prism 7 and FlowJo. For each sample, the mean fluorescence intensity (MFI) and the number of gated cells was measured. To report activity, triplicate MFI values were averaged and normalized to the PMO alone condition.

A.4.5 Endocytosis inhibition assay

Chemical endocytosis inhibitors were used to probe the mechanism of delivery of PMO by these peptides in a pulse-chase format. We have conducted such analysis on similar PMO-peptide constructs previously with comparable outcomes.[3] For the PMO constructs, HeLa 654 cells were preincubated with various chemical inhibitors for 30 minutes before treatment with PMO-CPP constructs for three hours. The panel included: a panel of endocytosis inhibitors including: chlorpromazine (CPZ), which is demonstrated to interfere with clathrin-mediated endocytosis; cytochalasin D (CyD), which inhibits phagocytosis and micropinocytosis; wortmannin (Wrt), which alters various endocytosis pathways by inhibiting phosphatidylinositol kinases; EIPA (5-(N-ethyl-Nisopropyl) amiloride), which inhibits micropinocytosis; and Dynasore (Dyn), which also inhibits clathrin-mediated endocytosis.[4,5] Treatment media was then replaced with fresh media and the cells were incubated for 22 hours at 37 °C and 5% CO_2. Cells were then lifted as previously described and EGFP synthesis was measured by flow cytometry.

A.4.6 LDH assay

Cytotoxicity assays were performed in HeLa 654 cells. Cell supernatant following treatment for flow cytometry was transferred to a new 96-well plate for analysis of LDH release. To each well of the 96-well plate containing supernatant was added CytoTox 96 Reagent (Promega). The plate was shielded from light and incubated at room temperature for 30 minutes. Equal volume of Stop Solution was added to each well, mixed, and the absorbance of each well was measured at 490 nm. The measurement of untreated cells was subtracted from each treated cell measurement, and % LDH release was calculated as % cytotoxicity = 100 × Experimental LDH Release (OD490) / Maximum LDH Release (OD490).

A.4.7 Uptake assay

The uptake assay was adapted from a previously reported protocol[35] Briefly, cells were plated in a 12-well plate the evening before the experiment. On the day of, cells were treated at

varying concentrations of PMO-biotin-peptide or biotin-peptide at varying temperatures and durations. Treatment media was removed, and cells were washed with fresh media before incubated with porcine heparin for 5 min to dissociate membrane-bound peptide. Cells were lifted and membrane-bound peptide was cleaved by incubation in trypsin for 5 min. Collected cells were pelleted and washed before being digested with RIPA buffer (for whole cell extract) or digitonin buffer (for cytosolic extract) on ice. Supernatant was collected and protein concentration was quantified using a BCA protein assay kit. A portion of some samples were retained for analysis by Western blot. The remaining supernatant was incubated with magnetic streptavidin Dynabeads overnight at 4 °C.

Dynabeads were washed with a series of buffers: 2 x 100 µL Buffer A (50 mM Tris-HCl (pH 7.4) and 0.1 mg/mL BSA), 2 x 100 µL Buffer B (50 mM Tris-HCl (pH 7.4), 0.1 mg/mL BSA, and 0.1% SDS), 2 x 100 µL Buffer C (50 mM Tris-HCl (pH 7.4), 0.1 mg/mL BSA, and 1 M NaCl), and 2 x 100 µL water. Washed beads were then resuspended in CHCA MALDI matrix and plated directly on a MALDI plate. The plate was analyzed using MALDI-ToF on a high-resolution Bruker Autoflex LRF Speed mass spectrometer in linear positive mode.

Relative concentrations of peptides in the mixture were determined as follows. Analytes in a mixture ionize according to their response factor (F). F was determined by normalizing the intensities of each analyte to one analyte in the control sample, where the concentration of each analyte is arbitrarily set to 1. The values of F is then used in the experimental spectra containing the same mixture of analytes to determine their relative concentrations.

$$\frac{I_x}{[X]} = F\left(\frac{I_s}{[S]}\right)$$

A.4.8 Serum stability assay

Each PMO-peptide was dissolved in PBS to a concentration of 1 mM, as confirmed by UV-Vis. PMO-peptide was then added to a solution of either PBS or PBS containing 25% human serum to a final concentration of 50 µM and incubated at 37 °C. 10 µL aliquots were removed at varying timepoints (t = 0, 1h, 6 h, 24 h), and quenched with 20 µL 1M guanidinium hydrochloride and 50 mM EDTA. 50 µL ice-cold acetonitrile was then added and the aliquots were flash frozen until LCMS analysis. Samples were thawed, and a portion of the aqueous layer was diluted before

analysis by LC-qTOF. The mass spectrum for the PMO-peptides was analyzed by deconvolution, to best demonstrated whether the analyte had stayed intact or degraded.

A.4.9 Statistics

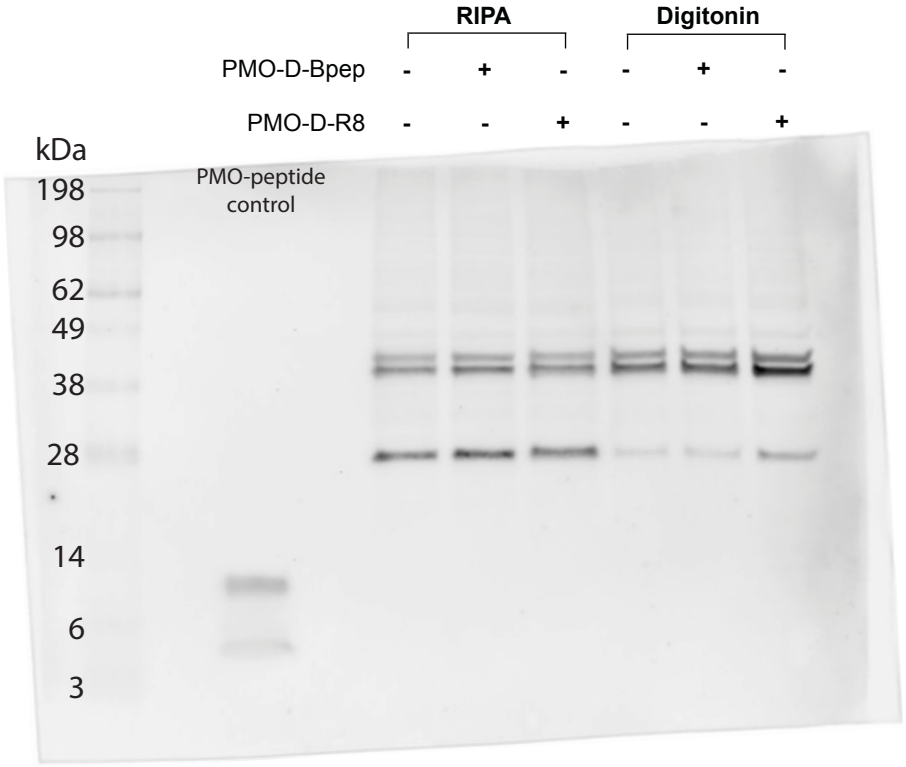
Statistical analysis and graphing was performed using Prism (Graphpad) or Excel (Microsoft). Concentration-response curves were fitted using Prism using nonlinear regression. The listed replicates for each experiment indicates the number of distinct samples measured for a given assay. Significance for activities between constructs was determined using a student's two-sided, unpaired t-test.

A.5 Acknowledgements

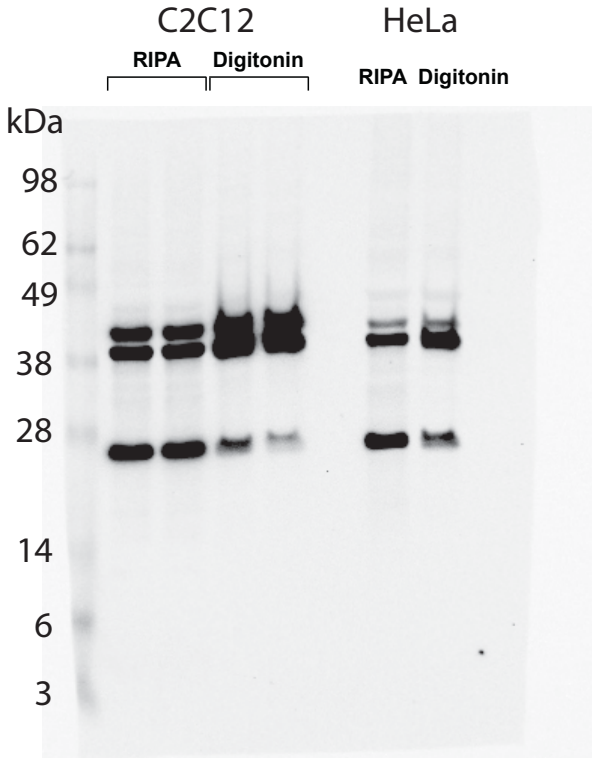
This research was funded by Sarepta Therapeutics. C.K.S. (4000057398) and C.E.F. (4000057441) acknowledge the National Science Foundation Graduate Research Fellowship (NSF Grant No. 1122374) for research support. We thank the Swanson Biotechnology Center Flow Cytometry Facility at the Koch Institute for the use of their flow cytometers, and MIT's Department of Chemistry Instrumentation Facility for the use of the MALDI-ToF.

A.6 Appendix I: Gel Images

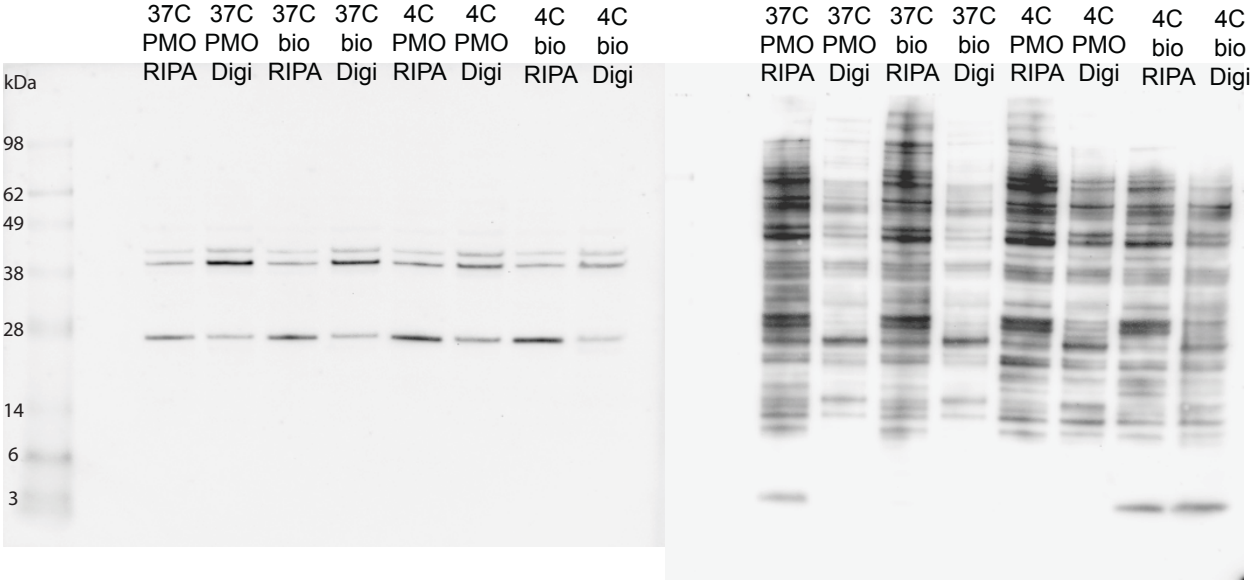
Full gel image corresponding to that shown in Figure A.4.



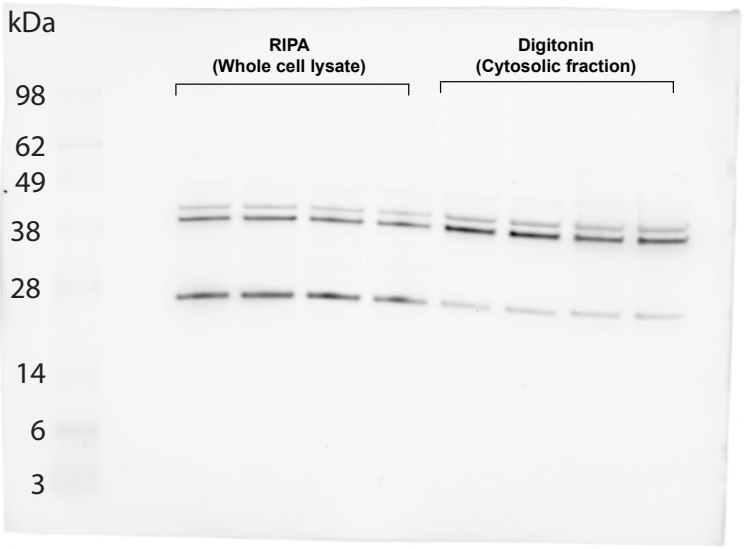
Full gel image corresponding to the experiment described in Figure A.5.



Full gel image. IR800 (left) and Chemiluminescence (right)

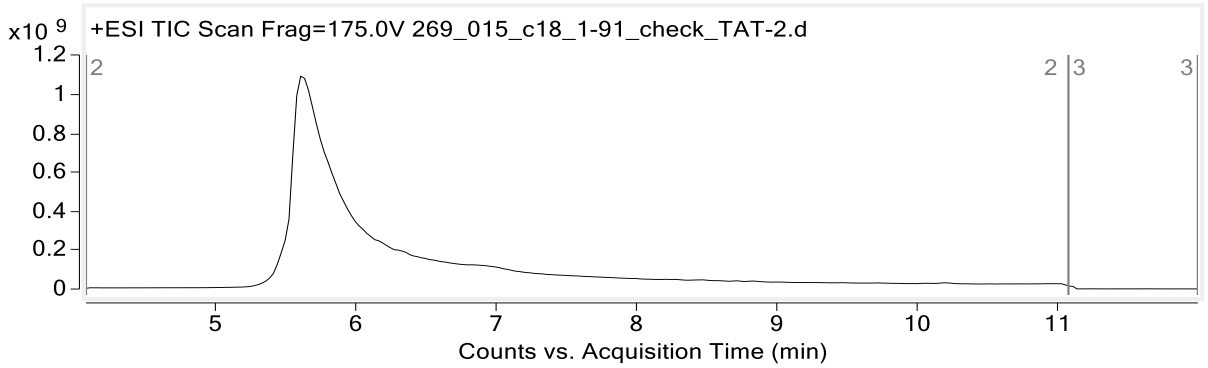


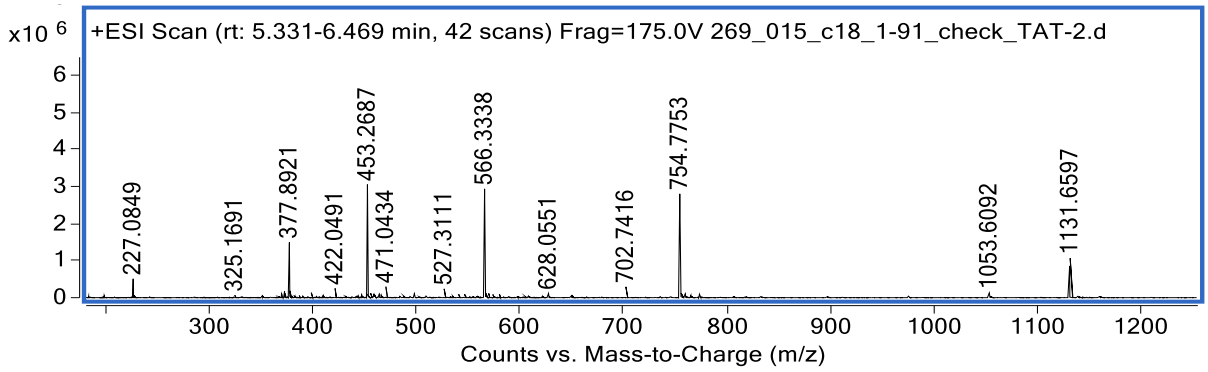
Full gel image



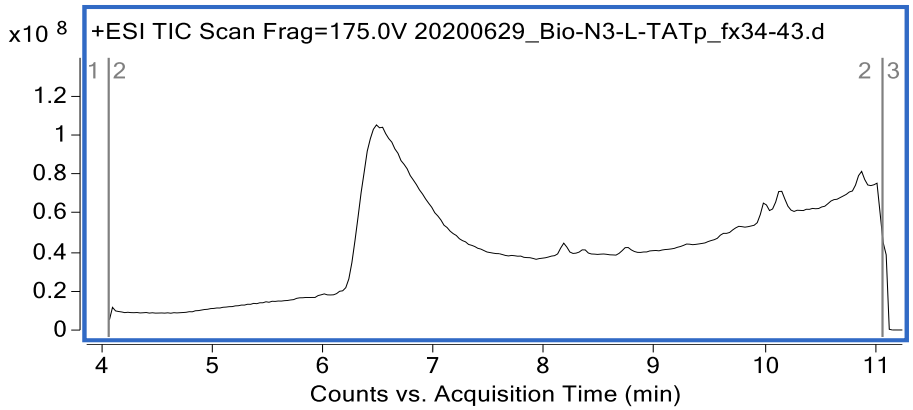
A.7 Appendix II: Compound characterization

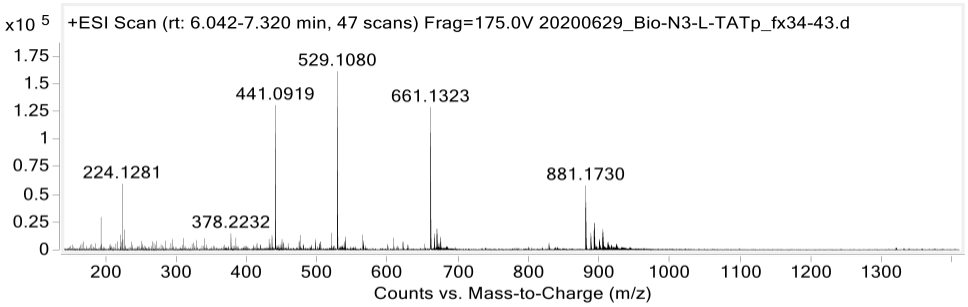
Peptide: Bio-D-TAT
Sequence: Biotin-Lys(N_3)-GGKGGWRKKRRQRRR
Calculated monoisotopic mass: 2260.3 Da
Observed monoisotopic mass: 2260.4 Da



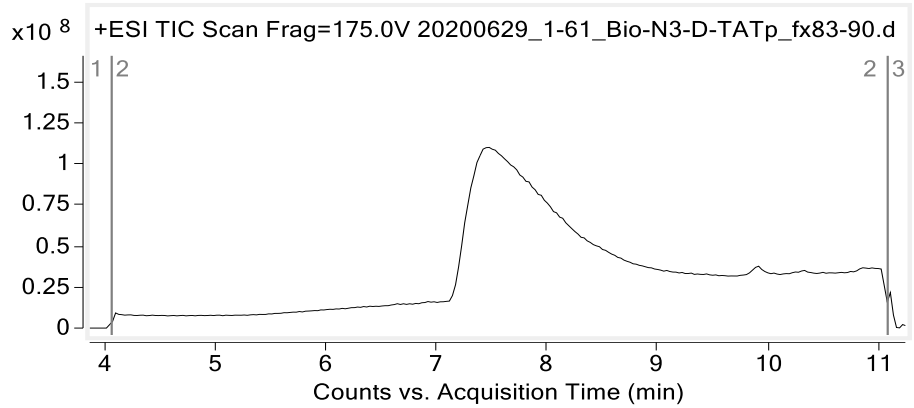


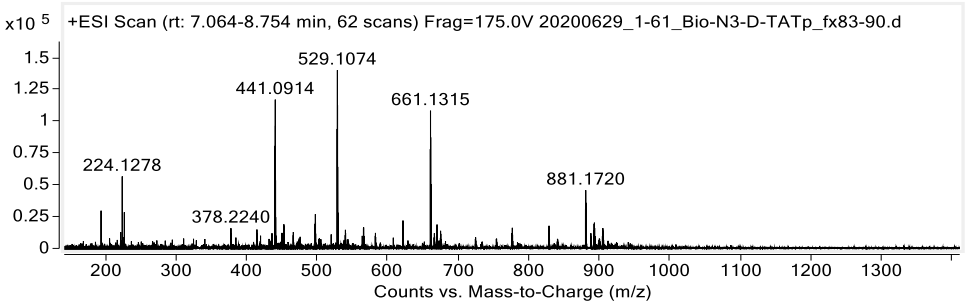
Peptide: Bio-L-TAT
Sequence: Biotin-Lys(N_3)-GGKGGWRKKRRQRRR
Calculated monoisotopic mass: 2260.3 Da
Observed monoisotopic mass: 2260.4 Da



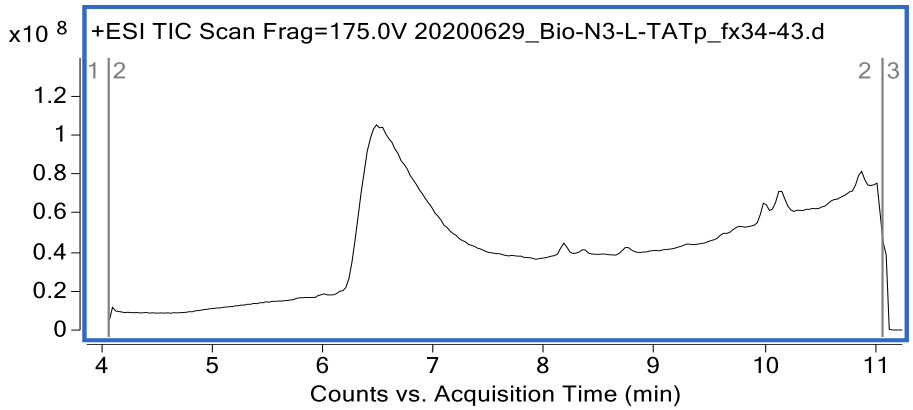


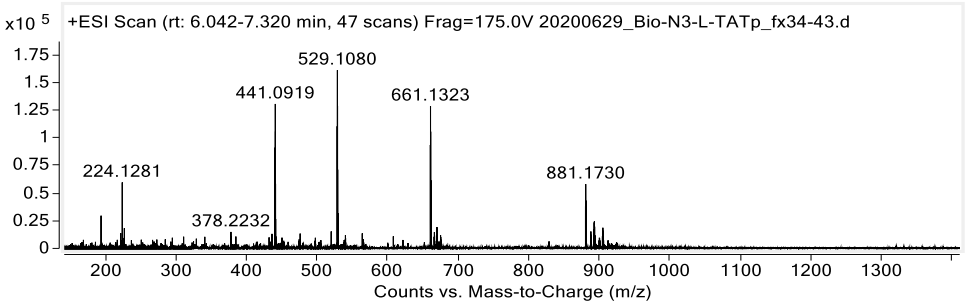
Peptide: Bio-D-TATp
Sequence: Biotin-Lys(N_3)-GGKGGWGRKKRRQRRRPPQ
Calculated monoisotopic mass: 2639.5 Da
Observed monoisotopic mass: 2639.5 Da



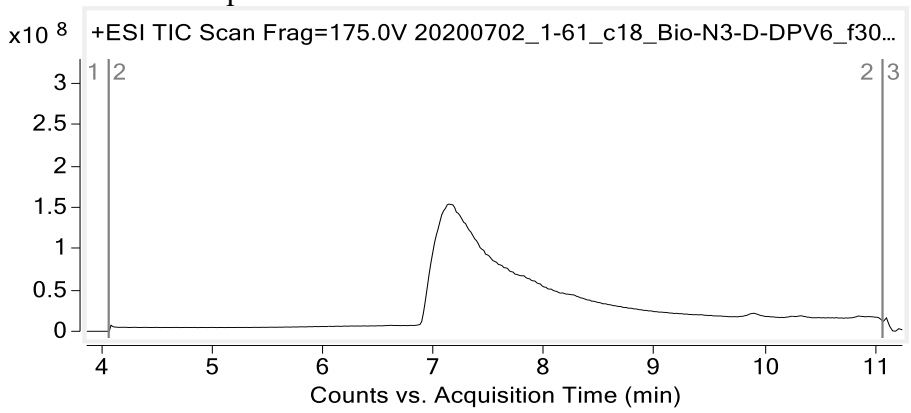


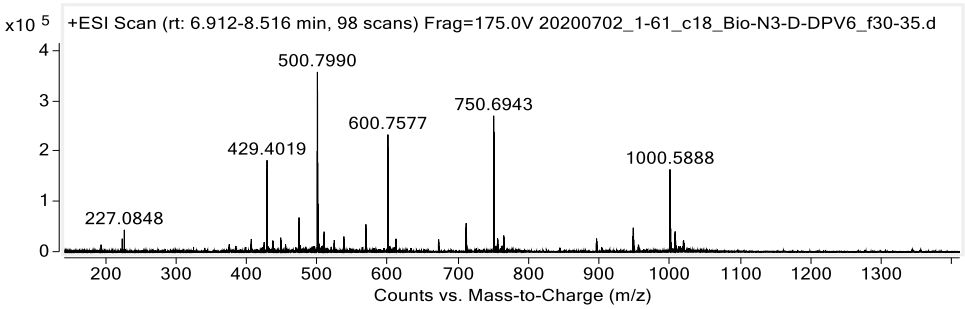
Peptide: Bio-L-TATp
Sequence: Biotin-Lys(N_3)-GGKGGWGRKKRRQRRRPPQ
Calculated monoisotopic mass: 2639.5 Da
Observed monoisotopic mass: 2639.5 Da



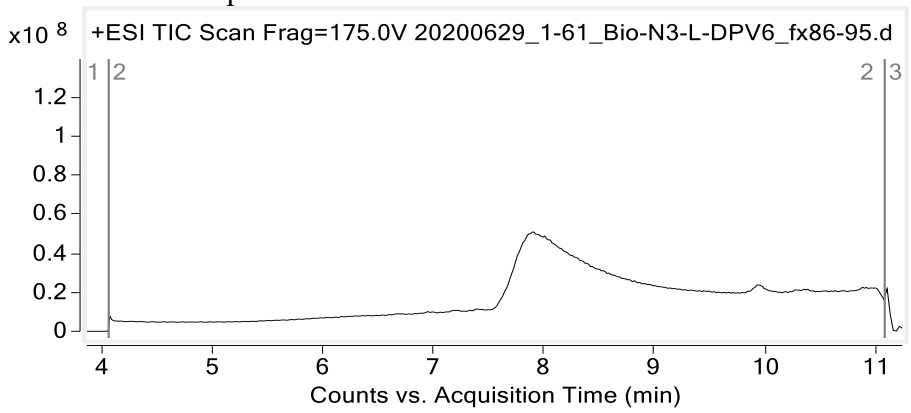


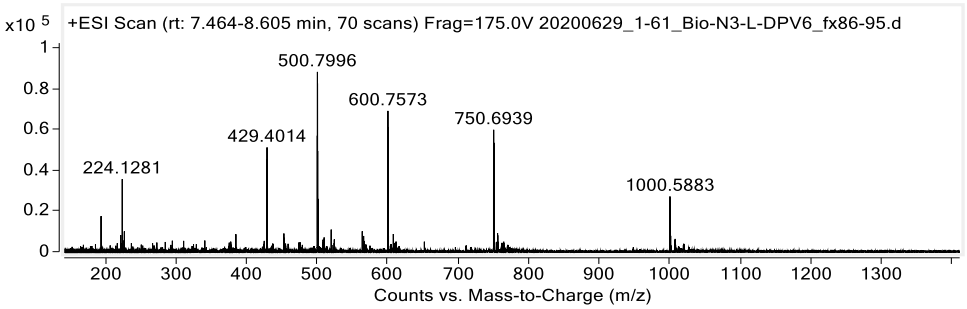
Peptide: Bio-D-DPV6
Sequence: Biotin-Lys(N_3)-GGKGGWGRPRESGKKRKRKRLKP
Calculated monoisotopic mass: 2997.7 Da
Observed monoisotopic mass: 2997.8 Da



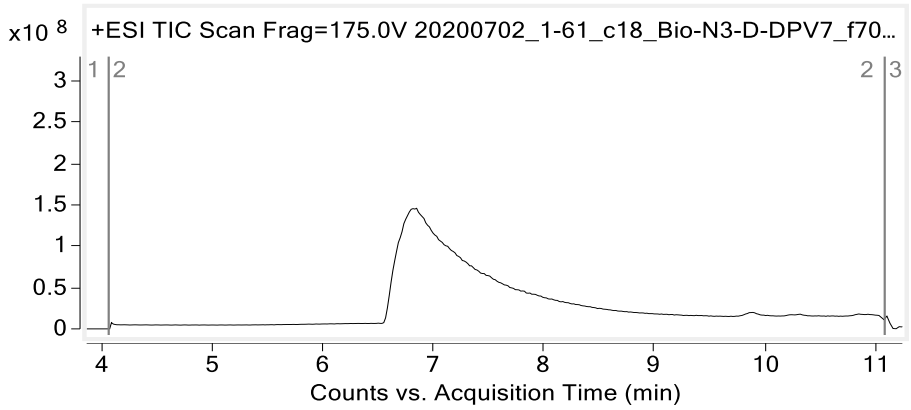


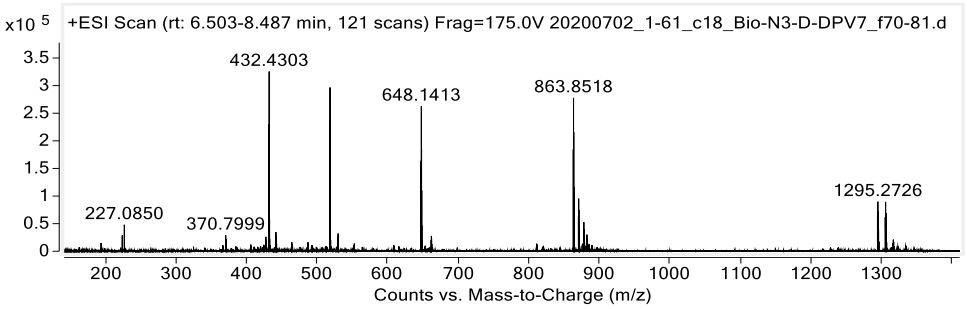
Peptide: Bio-L-DPV6
Sequence: Biotin-Lys(N_3)-GGKGGWGRPRESGKKRKRKRLKP
Calculated monoisotopic mass: 2997.7 Da
Observed monoisotopic mass: 2997.8 Da



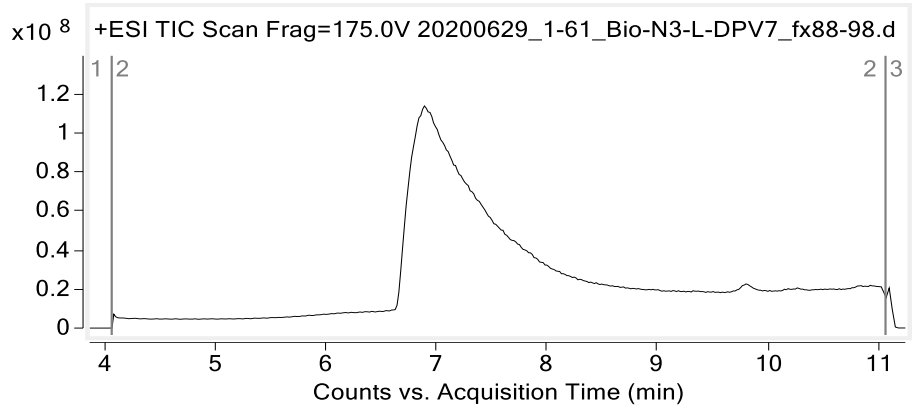


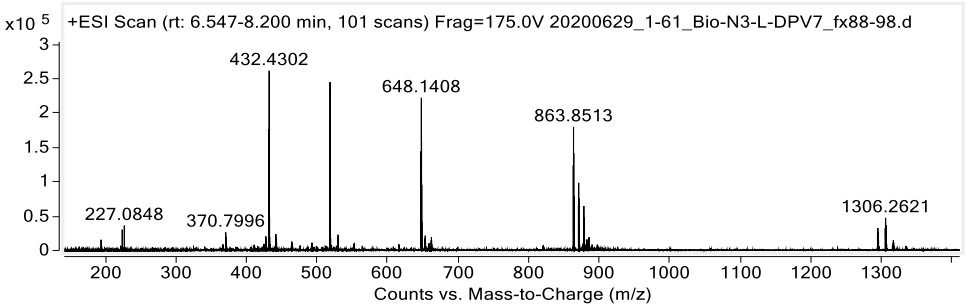
Peptide: Bio-D-DPV7
Sequence: Biotin-Lys(N_3)-GGKGGWKRKKKGKLGKKRDP
Calculated monoisotopic mass: 2587.5 Da
Observed monoisotopic mass: 2587.5 Da



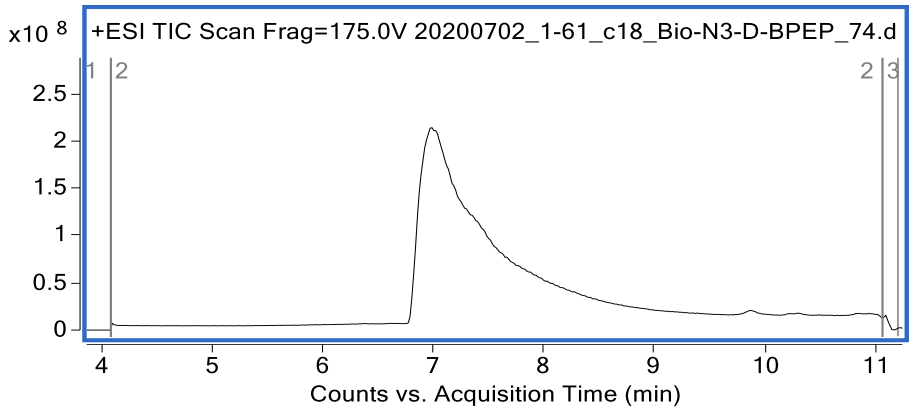


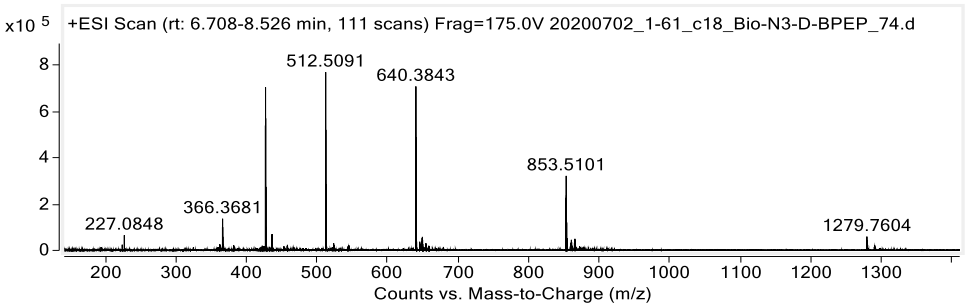
Peptide: Bio-L-DPV7
Sequence: Biotin-Lys(N_3)-GGKGGWKRKKKGKLGKKRDP
Calculated monoisotopic mass: 2587.5 Da
Observed monoisotopic mass: 2587.5 Da



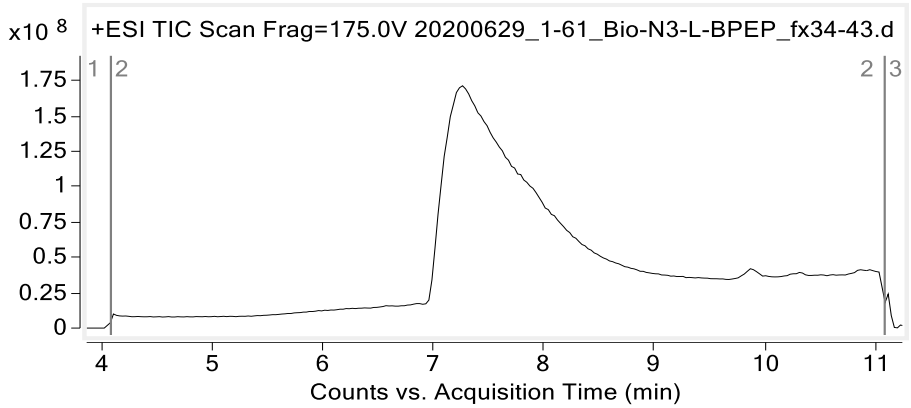


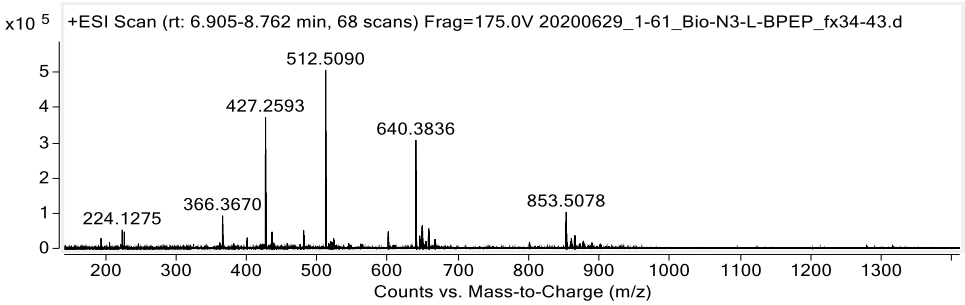
Peptide: Bio-D-BPEP
Sequence: Biotin-Lys(N_3)-GGKGGWRXRRBRRXRRBR
Calculated monoisotopic mass: 2556.5 Da
Observed monoisotopic mass: 2556.5 Da



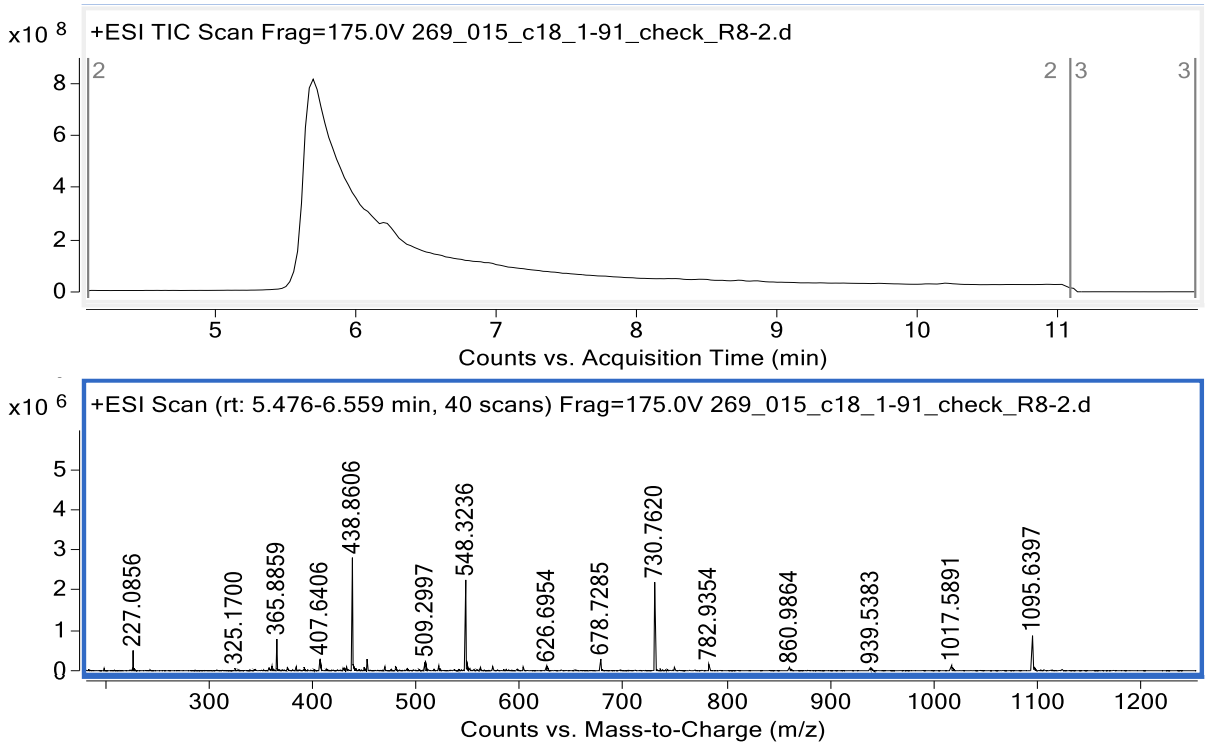


Peptide: Bio-L-BPEP
Sequence: Biotin-Lys(N_3)-GGKGGWRXRRBRRXRRBR
Calculated monoisotopic mass: 2556.5 Da
Observed monoisotopic mass: 2556.5 Da

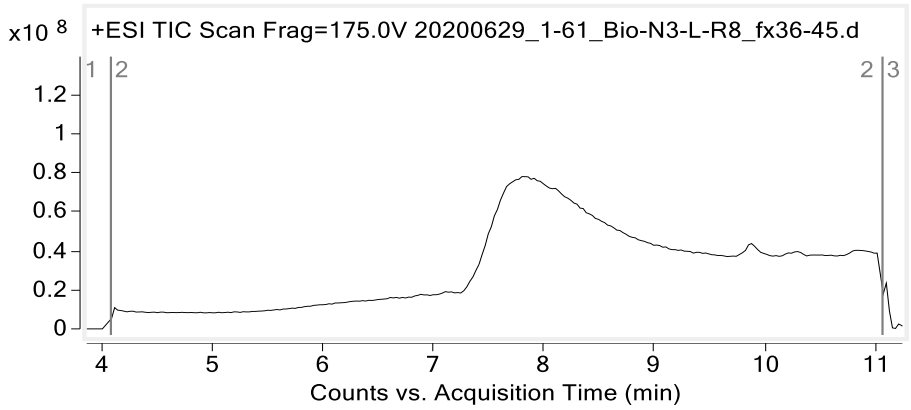


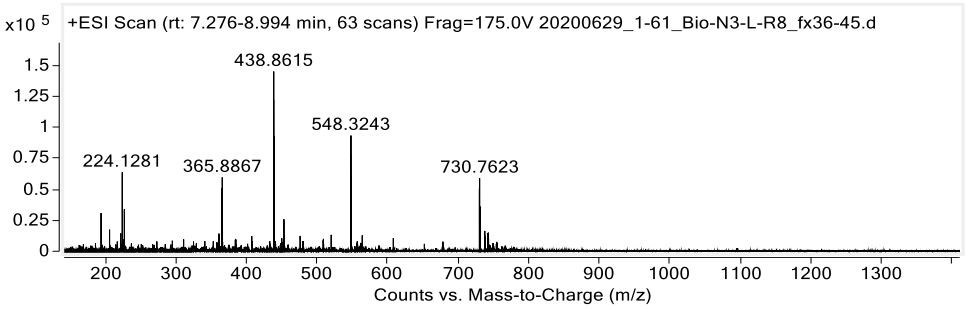


Peptide: Bio-D-R8
Sequence: Biotin-Lys(N_3)-GGKGGWRRRRRRRR
Calculated monoisotopic mass: 2188.3 Da
Observed monoisotopic mass: 2188.3 Da

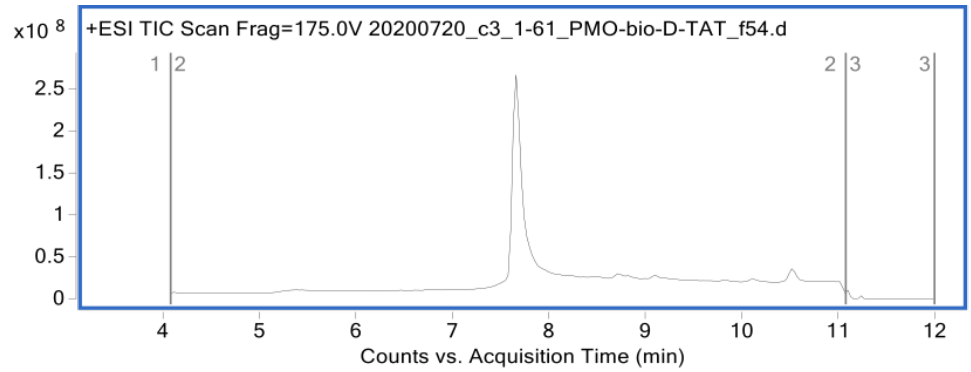


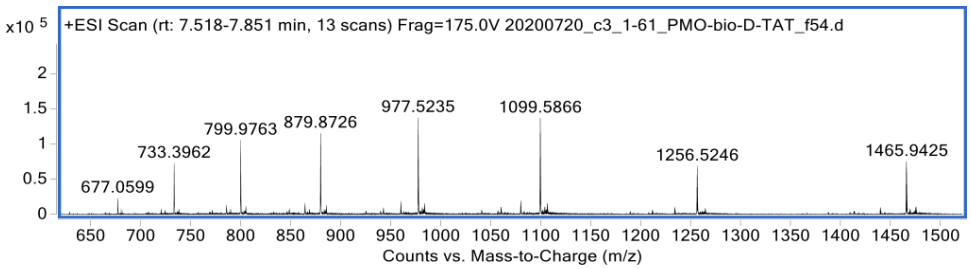
Peptide: Bio-L-R8
Sequence: Biotin-Lys(N_3)-GGKGGWRRRRRRRR
Calculated monoisotopic mass: 2188.3 Da
Observed monoisotopic mass: 2188.3 Da



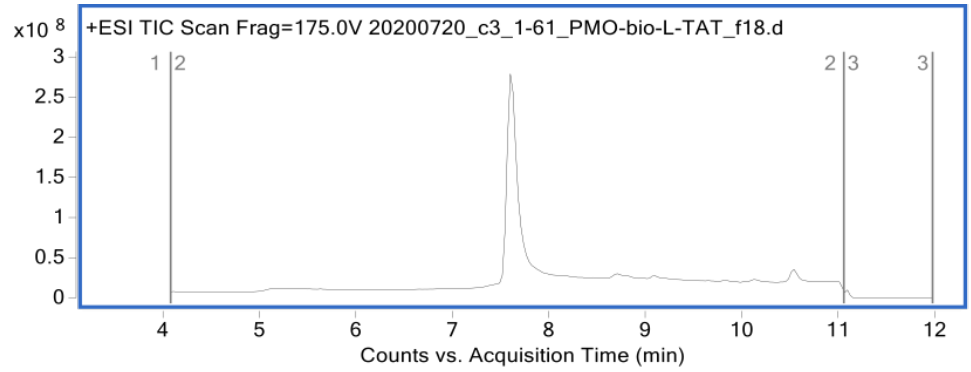


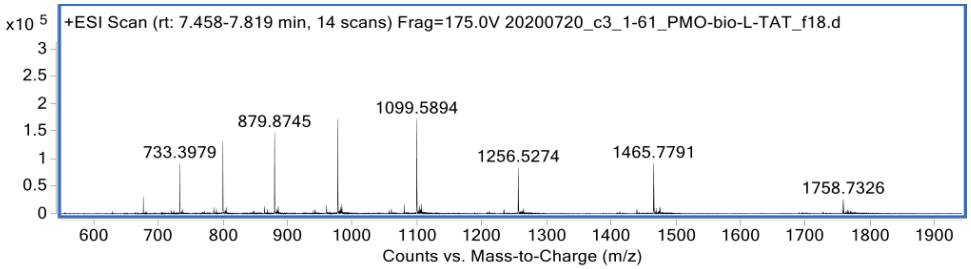
Peptide: PMO-Bio-D-TAT
Sequence: Biotin-Lys(PMO)-GGKGGWRKKRRQRRR
Calculated mass:8789.6 Da
Observed mass: 8789.8 Da



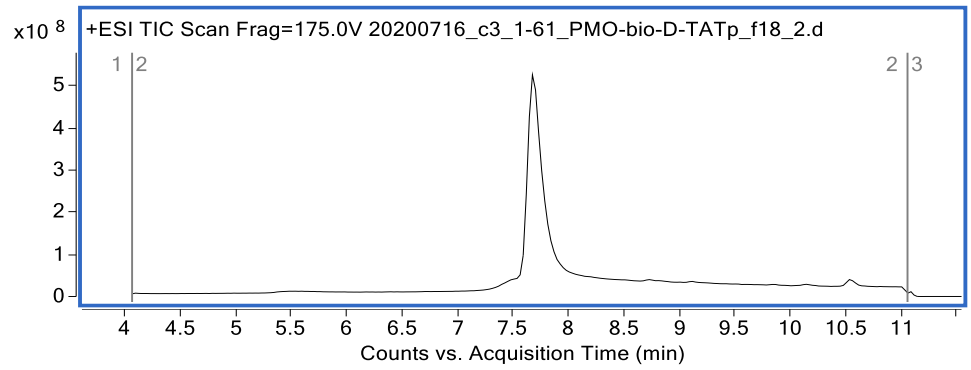


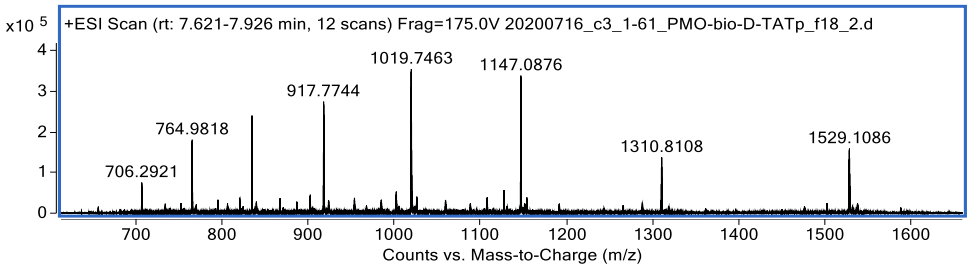
Peptide: PMO-Bio-L-TAT
Sequence: Biotin-Lys(PMO)-GGKGGWRKKRRQRRR
Calculated mass: 8789.6 Da
Observed mass: 8789.7 Da



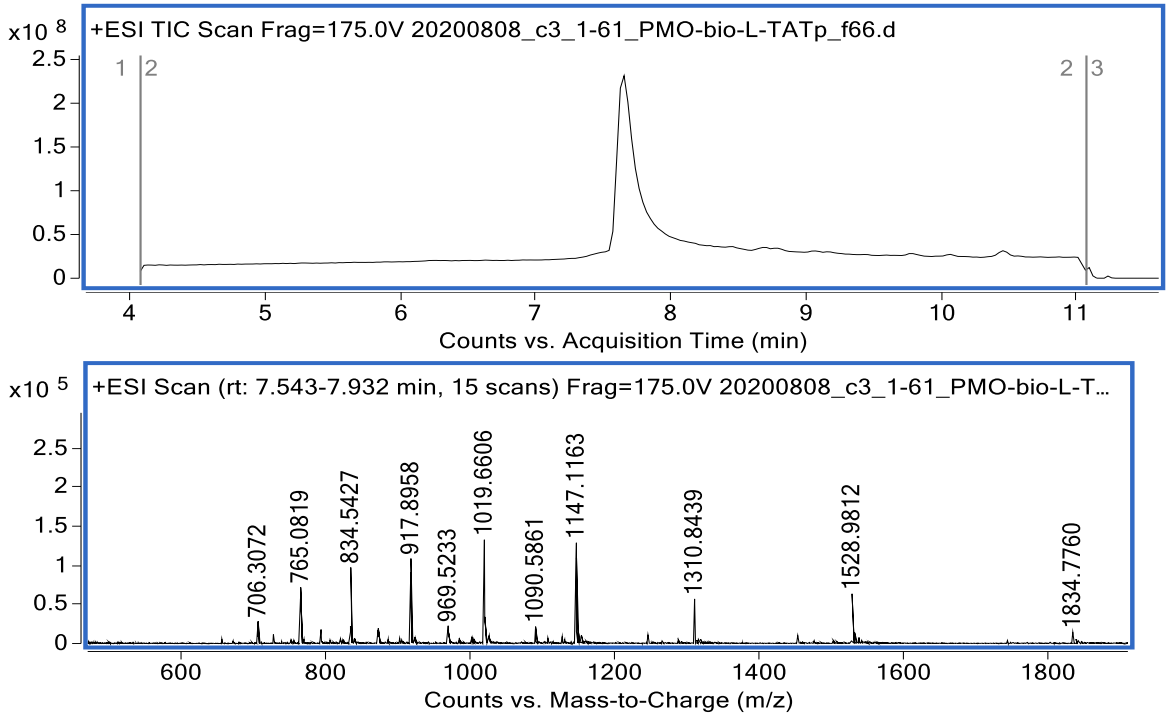


Peptide: PMO-Bio-D-TATp
Sequence: Biotin-Lys(PMO)-GGKGGWGRKKRRQRRRPPQ
Calculated mass: 9169.0 Da
Observed mass: 9169.1 Da

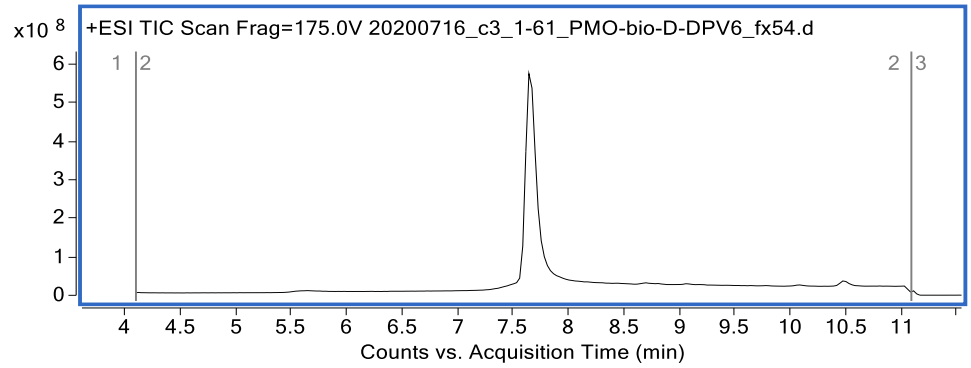


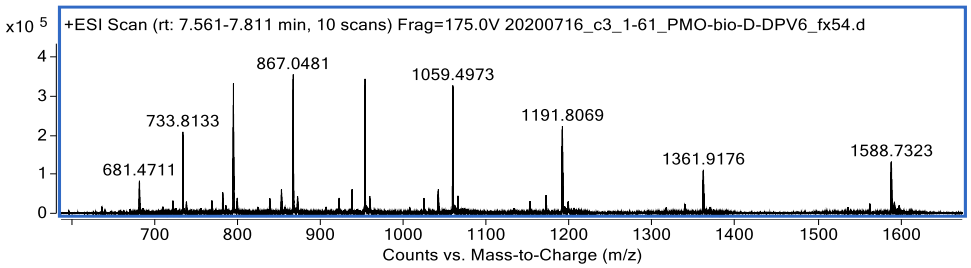


Peptide: PMO-Bio-L-TATp
Sequence: Biotin-Lys(PMO)-GGKGGWGRKKRRQRRRPPQ
Calculated mass: 9169.0 Da
Observed mass: 9169.1 Da

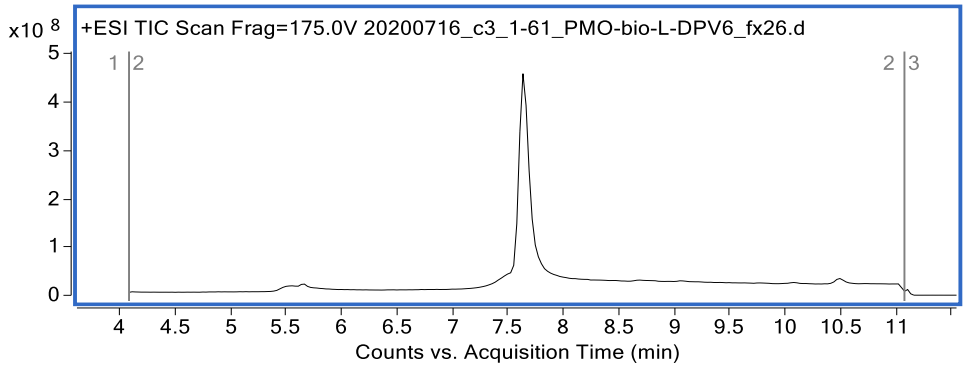


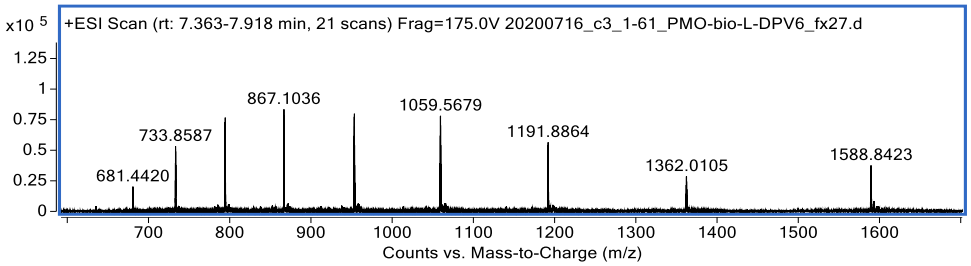
Peptide: PMO-Bio-D-DPV6
Sequence: Biotin-Lys(PMO)-GGKGGWGRPRESGKKRKRKRLKP
Calculated mass: 9527.5 Da
Observed mass: 9527.2 Da



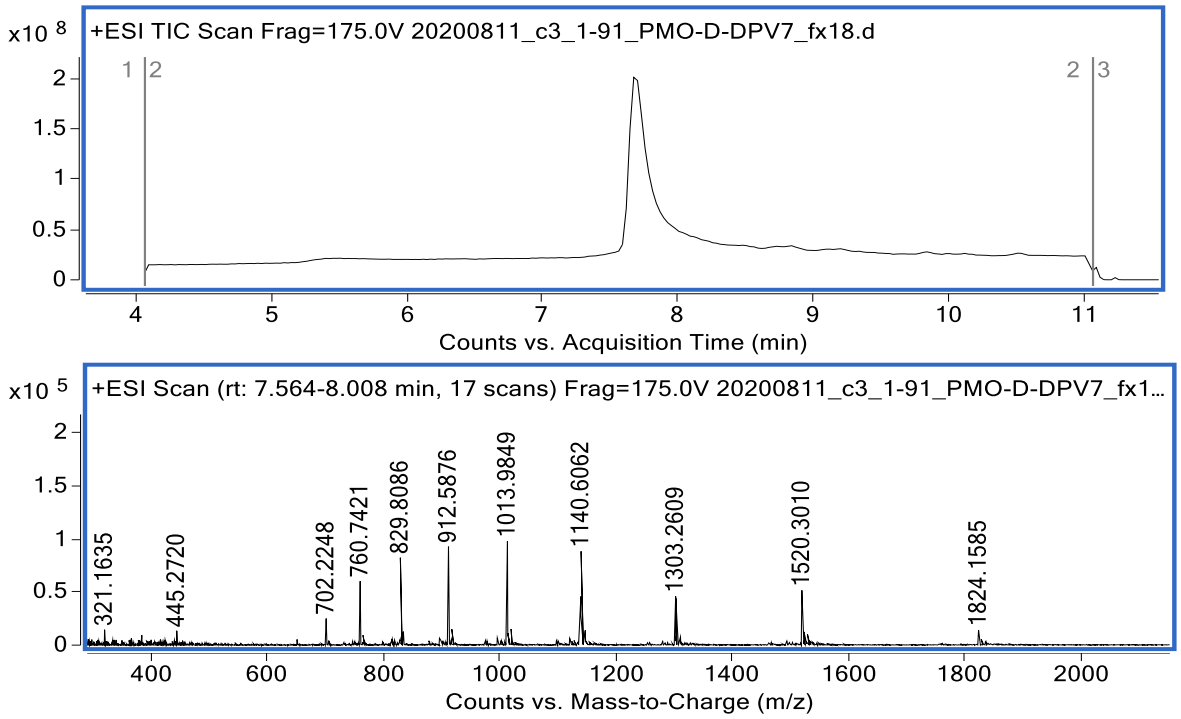


Peptide: PMO-Bio-L-DPV6
Sequence: Biotin-Lys(PMO)-GGKGGWGRPRESGKKRKRKRLKP
Calculated mass: 9527.5 Da
Observed mass: 9527.4 Da

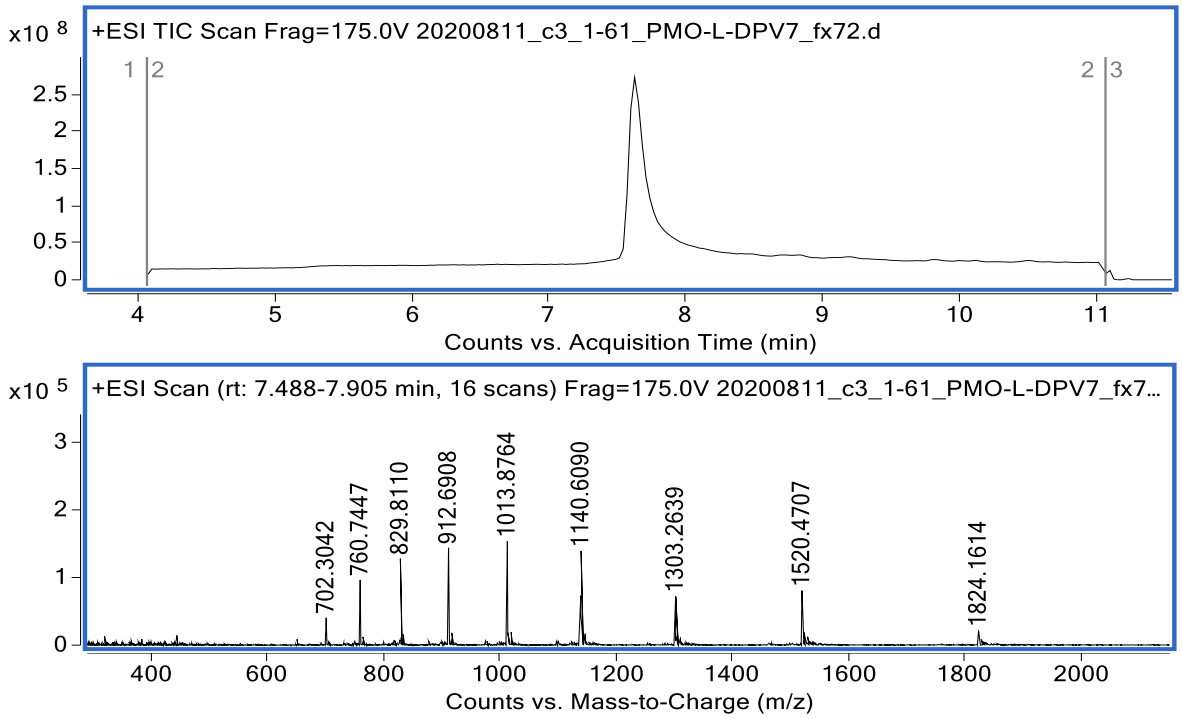




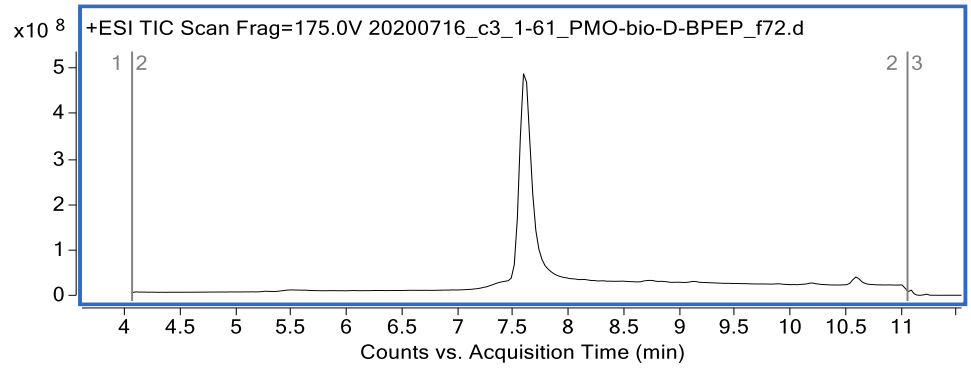
Peptide: PMO-Bio-D-DPV7
Sequence: Biotin-Lys(PMO)-GGKGGWKRKKKGKLGKKRDP
Calculated mass: 9117.1 Da
Observed mass: 9117.1 Da

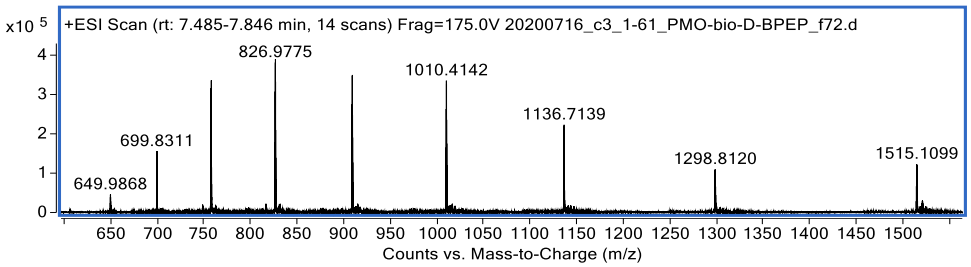


Peptide: PMO-Bio-L-DPV7
Sequence: Biotin-Lys(PMO)-GGKGGWKRKKKGKLGKKRDP
Calculated mass: 9117.1 Da
Observed mass: 9117.1 Da

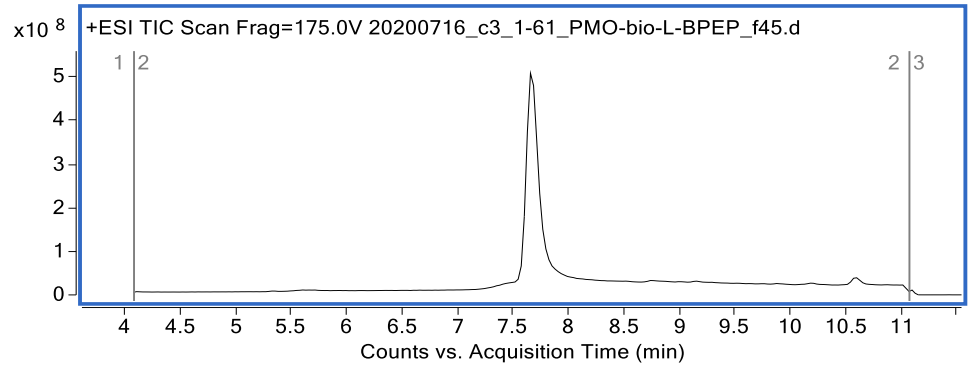


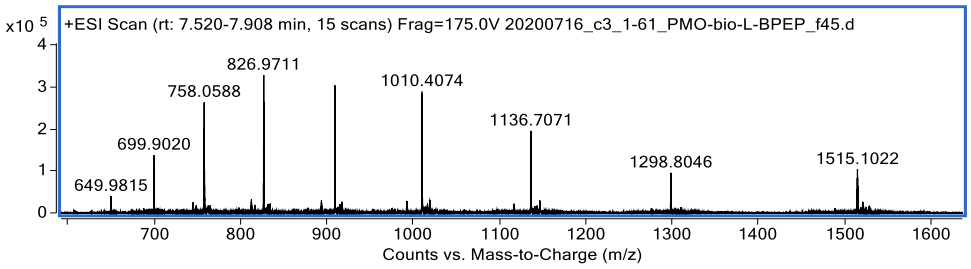
Peptide: PMO-Bio-D-BPEP
Sequence: Biotin-Lys(PMO)-GGKGGWRXRRBRRXRRBR
Calculated mass: 9086.0 Da
Observed mass: 9086.1 Da



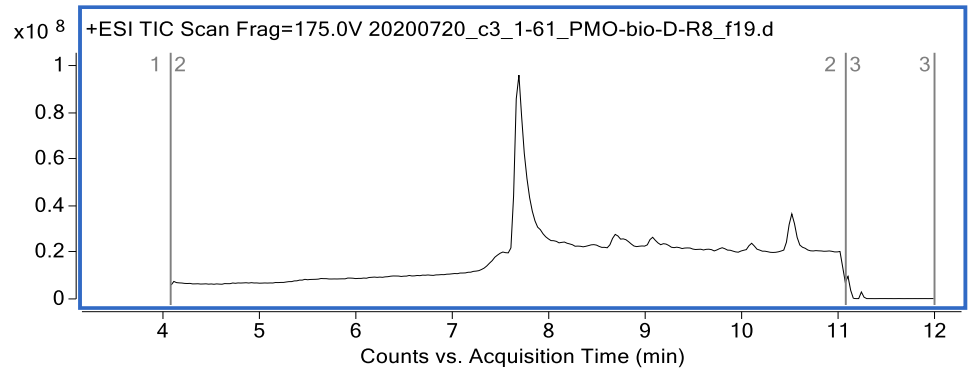


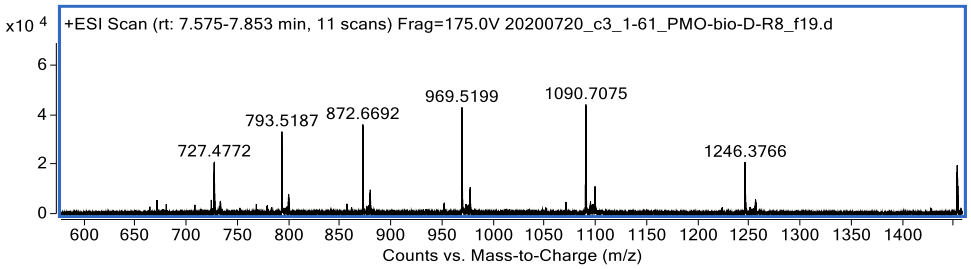
Peptide: PMO-Bio-L-BPEP
Sequence: Biotin-Lys(PMO)-GGKGGWRXRRBRRXRRBR
Calculated mass: 9086.0 Da
Observed mass: 9085.8 Da



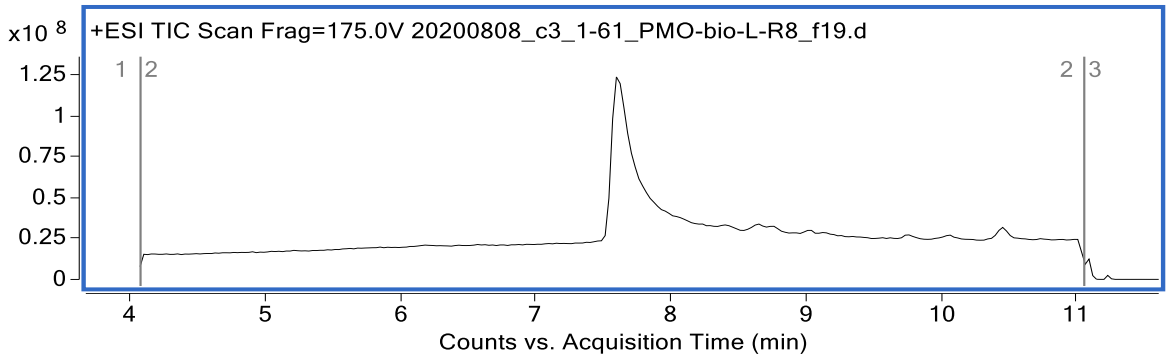


Peptide: PMO-Bio-D-R8
Sequence: Biotin-Lys(PMO)-GGKGGWRRRRRRRRR
Calculated mass: 8717.5 Da
Observed mass: 8717.7 Da





Peptide: PMO-Bio-L-R8
Sequence: Biotin-Lys(PMO)-GGKGGWRRRRRRRR
Calculated mass: 8717.5 Da
Observed mass: 8717.6 Da



A.8 References

(1) Wolfe Justin M.; Fadzen Colin M.; Holden Rebecca L.; Yao Monica; Hanson Gunnar J.; Pentelute Bradley L. Perfluoroaryl Bicyclic Cell-Penetrating Peptides for Delivery of Antisense Oligonucleotides. *Angew. Chem. Int. Ed.* **2018,**. https://doi.org/10.1002/anie.201801167.

(2) Oehlke, J.; Scheller, A.; Wiesner, B.; Krause, E.; Beyermann, M.; Klauschenz, E.; Melzig, M.; Bienert, M. Cellular Uptake of an α-Helical Amphipathic Model Peptide with the Potential to Deliver Polar Compounds into the Cell Interior Non-Endocytically. *Biochim. Biophys. Acta BBA - Biomembr*. **1998,** *1414* (1), 127–139. https://doi.org/10.1016/S0005-2736(98)00161-8.

(3) Margus, H.; Padari, K.; Pooga, M. Cell-Penetrating Peptides as Versatile Vehicles for Oligonucleotide Delivery. *Mol. Ther. J. Am. Soc. Gene Ther*. **2012,** *20* (3), 525–533. https://doi.org/10.1038/mt.2011.284.

(4) Reissmann Siegmund. Cell Penetration: Scope and Limitations by the Application of Cell-penetrating Peptides. *J. Pept. Sci.* **2014,** *20 (*10), 760–784. https://doi.org/10.1002/psc.2672.

(5) Fischer, R.; Waizenegger, T.; Köhler, K.; Brock, R. A Quantitative Validation of Fluorophore-Labelled Cell-Permeable Peptide Conjugates: Fluorophore and Cargo Dependence of Import. *Biochim. Biophys. Acta BBA - Biomembr*. **2002**, *1564* (2), 365–374. https://doi.org/10.1016/S0005-2736(02)00471-6.

(6) Lindgren, M. E.; Hällbrink, M. M.; Elmquist, A. M.; Langel, U. Passage of Cell-Penetrating Peptides across a Human Epithelial Cell Layer in Vitro. *Biochem. J.* **2004**, *377* (Pt 1), 69–76. https://doi.org/10.1042/BJ20030760.

(7) Illien, F.; Rodriguez, N.; Amoura, M.; Joliot, A.; Pallerla, M.; Cribier, S.; Burlina, F.; Sagan, S. Quantitative Fluorescence Spectroscopy and Flow Cytometry Analyses of Cell-Penetrating Peptides Internalization Pathways: Optimization, Pitfalls, Comparison with Mass Spectrometry Quantification. *Sci. Rep.* **2016,** *6,* 36938. https://doi.org/10.1038/srep36938.

(8) Wolfe, J. M.; Fadzen, C. M.; Choo, Z.-N.; Holden, R. L.; Yao, M.; Hanson, G. J.; Pentelute, B. L. Machine Learning To Predict Cell-Penetrating Peptides for Antisense Delivery. ACS Cent. Sci. **2018**, 4 (4), 512–520. https://doi.org/10.1021/acscentsci.8b00098.

(9) Baker, D. E. Eteplirsen. *Hosp. Pharm.* **2017,** *52* (4), 302–305. https://doi.org/10.1310/hpj5204-302.

(10) Lim, K. R. Q.; Maruyama, R.; Yokota, T. Eteplirsen in the Treatment of Duchenne Muscular Dystrophy. *Drug Des. Devel. Ther.* **2017**, *11,* 533–545. https://doi.org/10.2147/DDDT.S97635.

(11) Inc, S. T. Sarepta Therapeutics Announces Positive Clinical Results from MOMENTUM, a Phase 2 Clinical Trial of SRP-5051 in Patients with Duchenne Muscular Dystrophy Amenable to Skipping Exon 51 http://www.globenewswire.com/news-release/2020/12/07/2140613/0/en/Sarepta-Therapeutics-Announces-Positive-Clinical-Results-from-MOMENTUM-a-Phase-2-Clinical-Trial-of-SRP-5051-in-Patients-with-Duchenne-Muscular-Dystrophy-Amenable-to-Skipping-Exon-5.html

(12) Kurrikoff, K.; Vunk, B.; Langel, Ü. Status Update in the Use of Cell-Penetrating Peptides for the Delivery of Macromolecular Therapeutics. *Expert Opin. Biol. Ther.* **2021**, *21* (3), 361–370. https://doi.org/10.1080/14712598.2021.1823368.

(13) Dintzis, H. M.; Symer, D. E.; Dintzis, R. Z.; Zawadzke, L. E.; Berg, J. M. A Comparison of the Immunogenicity of a Pair of Enantiomeric Proteins. *Proteins Struct. Funct. Bioinforma.* **1993**, *16* (3), 306–308. https://doi.org/10.1002/prot.340160309.

(14) Kroenke, M. A.; Weeraratne, D. K.; Deng, H.; Sloey, B.; Subramanian, R.; Wu, B.; Serenko, M.; Hock, M. B. Clinical Immunogenicity of the D-Amino Acid Peptide Therapeutic Etelcalcetide: Method Development Challenges and Anti-Drug Antibody Clinical Impact Assessments. *J. Immunol. Methods* **2017,** *445*, 37–44. https://doi.org/10.1016/j.jim.2017.03.005.

(15) Najjar, K.; Erazo-Oliveras, A.; Brock, D. J.; Wang, T.-Y.; Pellois, J.-P. An L- to d-Amino Acid Conversion in an Endosomolytic Analog of the Cell-Penetrating Peptide TAT Influences Proteolytic Stability, Endocytic Uptake, and Endosomal Escape. *J. Biol. Chem.* **2017**, *292* (3), 847–861. https://doi.org/10.1074/jbc.M116.759837.

(16) Ma, Y.; Gong, C.; Ma, Y.; Fan, F.; Luo, M.; Yang, F.; Zhang, Y.-H. Direct Cytosolic Delivery of Cargoes in Vivo by a Chimera Consisting of D- and L-Arginine Residues. *J. Controlled Release* **2012**, *162* (2), 286–294. https://doi.org/10.1016/j.jconrel.2012.07.022.

(17) Henriques, S. T.; Peacock, H.; Benfield, A. H.; Wang, C. K.; Craik, D. J. Is the Mirror Image a True Reflection? Intrinsic Membrane Chirality Modulates Peptide Binding. *J. Am. Chem. Soc.* **2019**, *141* (51), 20460–20469. https://doi.org/10.1021/jacs.9b11194.

(18) Pujals, S.; Sabidó, E.; Tarragó, T.; Giralt, E. All-D Proline-Rich Cell-Penetrating Peptides: A Preliminary in Vivo Internalization Study. *Biochem. Soc. Trans*. **2007,** *35* (4), 794–796. https://doi.org/10.1042/BST0350794.

(19) Ito, S.; Torii, Y.; Chikamatsu, S.; Harada, T.; Yamaguchi, S.; Ogata, S.; Sonoda, K.; Wakayama, T.; Masuda, T.; Ohtsuki, S. Oral Coadministration of Zn-Insulin with d-Form Small Intestine-Permeable Cyclic Peptide Enhances Its Blood Glucose-Lowering Effect in Mice. *Mol. Pharm.* **2021,** *18* (4), 1593–1603.

(20) Verdurmen, W. P. R.; Bovee-Geurts, P. H.; Wadhwani, P.; Ulrich, A. S.; Hällbrink, M.; van Kuppevelt, T. H.; Brock, R. Preferential Uptake of L- versus D-Amino Acid Cell-Penetrating Peptides in a Cell Type-Dependent Manner. *Chem. Biol.* **2011**, *18* (8), 1000–1010. https://doi.org/10.1016/j.chembiol.2011.06.006.

(21) Knox, S. L.; Wissner, R.; Piszkiewicz, S.; Schepartz, A. Cytosolic Delivery of Argininosuccinate Synthetase Using a Cell-Permeant Miniature Protein. ACS Cent. Sci. **2021**, 7 (4), 641–649. https://doi.org/10.1021/acscentsci.0c01603.

(22) Fadzen, C. M.; Holden, R. L.; Wolfe, J. M.; Choo, Z.-N.; Schissel, C. K.; Yao, M.; Hanson, G. J.; Pentelute, B. L. Chimeras of Cell-Penetrating Peptides Demonstrate Synergistic Improvement in Antisense Efficacy. *Biochemistry* **2019.**

(23) Schissel, C. K.; Mohapatra, S.; Wolfe, J. M.; Fadzen, C. M.; Bellovoda, K.; Wu, C.-L.; Wood, J. A.; Malmberg, A. B.; Loas, A.; Gómez-Bombarelli, R.; Pentelute, B. L. Deep Learning to Design Nuclear-Targeting Abiotic Miniproteins. *Nat. Chem.* **2021**, 1–9. https://doi.org/10.1038/s41557-021-00766-3.

(24) Youngblood, D. S.; Hatlevig, S. A.; Hassinger, J. N.; Iversen, P. L.; Moulton, H. M. Stability of Cell-Penetrating Peptide−Morpholino Oligomer Conjugates in Human Serum and in Cells. *Bioconjug. Chem.* **2007,** *18* (1), 50–60. https://doi.org/10.1021/bc060138s.

(25) Sazani, P.; Gemignani, F.; Kang, S.-H.; Maier, M. A.; Manoharan, M.; Persmark, M.; Bortner, D.; Kole, R. Systemically Delivered Antisense Oligomers Upregulate Gene Expression in Mouse Tissues. *Nat. Biotechnol.* **2002**, *20* (12), 1228–1233. https://doi.org/10.1038/nbt759.

(26) Erazo-Oliveras, A.; Muthukrishnan, N.; Baker, R.; Wang, T.-Y.; Pellois, J.-P. Improving the Endosomal Escape of Cell-Penetrating Peptides and Their Cargos: Strategies and Challenges. *Pharmaceuticals* **2012**, *5* (11), 1177–1209. https://doi.org/10.3390/ph5111177.
(27) Holden, P.; Horton, W. A. Crude Subcellular Fractionation of Cultured Mammalian Cell Lines. *BMC Res. Notes* **2009**, *2* (1), 243. https://doi.org/10.1186/1756-0500-2-243.
(28) Peraro, L.; Deprey, K. L.; Moser, M. K.; Zou, Z.; Ball, H. L.; Levine, B.; Kritzer, J. A. Cell Penetration Profiling Using the Chloroalkane Penetration Assay. *J. Am. Chem. Soc.* **2018,** *140* (36), 11360–11369. https://doi.org/10.1021/jacs.8b06144.
(29) Schmidt, S.; Adjobo-Hermans, M. J. W.; Wallbrecher, R.; Verdurmen, W. P. R.; Bovée-Geurts, P. H. M.; van Oostrum, J.; Milletti, F.; Enderle, T.; Brock, R. Detecting Cytosolic Peptide Delivery with the GFP Complementation Assay in the Low Micromolar Range. Angew. *Chem. Int. Ed*. **2015,** *54* (50), 15105–15108.
(30) Peier, A.; Ge, L.; Boyer, N.; Frost, J.; Duggal, R.; Biswas, K.; Edmondson, S.; Hermes, J. D.; Yan, L.; Zimprich, C.; Sadruddin, A.; Kristal Kaan, H. Y.; Chandramohan, A.; Brown, C. J.; Thean, D.; Lee, X. E.; Yuen, T. Y.; Ferrer-Gago, F. J.; Johannes, C. W.; Lane, D. P.; Sherborne, B.; Corona, C.; Robers, M. B.; Sawyer, T. K.; Partridge, A. W. NanoClick: A High Throughput, Target-Agnostic Peptide Cell Permeability Assay. *ACS Chem. Bio*l. **2021**, *16* (2), 293–309. https://doi.org/10.1021/acschembio.0c00804.
(31) Teo, S. L. Y.; Rennick, J. J.; Yuen, D.; Al-Wassiti, H.; Johnston, A. P. R.; Pouton, C. W. Unravelling Cytosolic Delivery of Cell Penetrating Peptides with a Quantitative Endosomal Escape Assay. *Nat. Commun.* **2021**, *12* (1), 3721. https://doi.org/10.1038/s41467-021-23997-x.
(32) Wissner, R. F.; Steinauer, A.; Knox, S. L.; Thompson, A. D.; Schepartz, A. Fluorescence Correlation Spectroscopy Reveals Efficient Cytosolic Delivery of Protein Cargo by Cell-Permeant Miniature Proteins. *ACS Cent. Sci.* **2018**, *4* (10), 1379–1393. https://doi.org/10.1021/acscentsci.8b00446.
(33) Duncan, M. W. Practical Quantitative Biomedical Applications of MALDI-TOF Mass Spectrometry.
(34) Burlina, F.; Sagan, S.; Bolbach, G.; Chassaing, G. Quantification of the Cellular Uptake of Cell-Penetrating Peptides by MALDI-TOF Mass Spectrometry. *Angew. Chem. Int. Ed.* **2005**, *44* (27), 4244–4247. https://doi.org/10.1002/anie.200500477.
(35) Burlina, F.; Sagan, S.; Bolbach, G.; Chassaing, G. A Direct Approach to Quantification of the Cellular Uptake of Cell-Penetrating Peptides Using MALDI-TOF Mass Spectrometry. *Nat. Protoc.* **2006,** *1* (1), 200.
(36) Aubry, S.; Aussedat, B.; Delaroche, D.; Jiao, C.-Y.; Bolbach, G.; Lavielle, S.; Chassaing, G.; Sagan, S.; Burlina, F. MALDI-TOF Mass Spectrometry: A Powerful Tool to Study the Internalization of Cell-Penetrating Peptides. *Biochim. Biophys. Acta BBA - Biomembr.* **2010,** *1798* (12), 2182–2189. https://doi.org/10.1016/j.bbamem.2009.11.011.
(37) Ho, H.-P.; Rathod, P.; Louis, M.; Tada, C. K.; Rahaman, S.; Mark, K. J.; Leng, J.; Dana, D.; Kumar, S.; Lichterfeld, M.; Chang, E. J. Studies on Quantitative Phosphopeptide Analysis by MALDI Mass Spectrometry Without Label, Chromatography or Calibration Curves. *Rapid Commun. Mass Spectrom. RCM* **2014,** *28* (24), 2681–2689. https://doi.org/10.1002/rcm.7063.
(38) Aussedat, B.; Sagan, S.; Chassaing, G.; Bolbach, G.; Burlina, F. Quantification of the Efficiency of Cargo Delivery by Peptidic and Pseudo-Peptidic Trojan Carriers Using

MALDI-TOF Mass Spectrometry. *Biochim. Biophys. Acta BBA - Biomembr.* **2006**, *1758* (3), 375–383. https://doi.org/10.1016/j.bbamem.2006.01.012.

(39) Bode, S. A.; Thévenin, M.; Bechara, C.; Sagan, S.; Bregant, S.; Lavielle, S.; Chassaing, G.; Burlina, F. Self-Assembling Mini Cell-Penetrating Peptides Enter by Both Direct Translocation and Glycosaminoglycan-Dependent Endocytosis. C*hem. Commun.* **2012**, *48* (57), 7179–7181. https://doi.org/10.1039/C2CC33240J.

(40) Aussedat, B.; Dupont, E.; Sagan, S.; Joliot, A.; Lavielle, S.; Chassaing, G.; Burlina, F. Modifications in the Chemical Structure of Trojan Carriers: Impact on Cargo Delivery. *Chem. Commun.* **2008,** *N12*, 1398–1400. https://doi.org/10.1039/B800433A.

(41) Brock, R. The Uptake of Arginine-Rich Cell-Penetrating Peptides: Putting the Puzzle Together. *Bioconjug. Chem.* 2014, *25* (5), 863–868. https://doi.org/10.1021/bc500017t.

(42) Fuchs, S. M.; Raines, R. T. Polyarginine as a Multifunctional Fusion Tag. *Protein Sci.* **2005**, *14* (6), 1538–1544. https://doi.org/10.1110/ps.051393805.

(43) Richard, J. P.; Melikov, K.; Vives, E.; Ramos, C.; Verbeure, B.; Gait, M. J.; Chernomordik, L. V.; Lebleu, B. Cell-Penetrating Peptides. A Reevaluation of the Mechanism of Cellular Uptake. J. *Biol. Chem.* 2003, *278* (1), 585–590. https://doi.org/10.1074/jbc.M209548200.

(44) Böttger, R.; Hoffmann, R.; Knappe, D. Differential Stability of Therapeutic Peptides with Different Proteolytic Cleavage Sites in Blood, Plasma and Serum. *PLOS ONE* **2017**, *12* (6), e0178943. https://doi.org/10.1371/journal.pone.0178943.

(45) Werner, H. M.; Cabalteja, C. C.; Horne, W. S. Peptide Backbone Composition and Protease Susceptibility: Impact of Modification Type, Position, and Tandem Substitution. *Chembiochem Eur. J. Chem. Biol.* **2016**, *17* (8), 712–718.

(46) Hartrampf, N.; Saebi, A.; Poskus, M.; Gates, Z. P.; Callahan, A. J.; Cowfer, A. E.; Hanna, S.; Antilla, S.; Schissel, C. K.; Quartararo, A. J.; Ye, X.; Mijalis, A. J.; Simon, M. D.; Loas, A.; Liu, S.; Jessen, C.; Nielsen, T. E.; Pentelute, B. L. Synthesis of Proteins by Automated Flow Chemistry. *Science* **2020**, *368* (6494), 980–987.

(47) Simon, M. D.; Heider, P. L.; Adamo, A.; Vinogradov, A. A.; Mong, S. K.; Li, X.; Berger, T.; Policarpo, R. L.; Zhang, C.; Zou, Y.; Liao, X.; Spokoyny, A. M.; Jensen, K. F.; Pentelute, B. L. Rapid Flow-Based Peptide Synthesis. *ChemBioChem* **2014**, *15* (5), 713–720. https://doi.org/10.1002/cbic.201300796.